

Valószínűségszámítás B előadásjegyzet

Tóth Dávid (BME SZIT)

2024 (utolsó frissítés: 2024.04.15.)

Tartalomjegyzék

Bevezetés	5
1. A véletlen matematikai modellje	6
1.1. Valószínűségi mezők	6
1.1.1. Eseményterek és események	7
1.1.2. Klasszikus valószínűség	11
1.1.3. Valószínűségi mérték	23
1.2. Feltételes valószínűség	30
1.2.1. Feltételes valószínűség és függetlenség	30
1.2.2. A feltételes valószínűség alkalmazásai	36
2. Diszkrét valószínűségi változók	43
2.1. Diszkrét valószínűségi változók eloszlása	44
2.2. Együttes eloszlás és függetlenség	57
2.3. A várható érték	65
2.3.1. A várható érték definíciója	66
2.3.2. Feltételes várható érték	72
2.3.3. Transzformált várható értéke	81
2.4. Variancia és szórás	87
2.5. Generátorfüggvények (kiegészítő anyag)	89
2.6. Nevezetes diszkrét eloszlások	96
2.6.1. A Poisson-eloszlás	96
2.6.2. Negatív binomiális eloszlás (kiegészítő anyag)	102
3. Abszolút folytonos valószínűségi változók	105
3.1. Geometriai valószínűségi mezők	105
3.2. Valószínűségi változók eloszlásfüggvénye	111
3.3. Abszolút folytonos változók	118
3.3.1. A sűrűségfüggvény fogalma	118
3.3.2. Várható érték és szórás	122
3.4. Nevezetes folytonos eloszlások	125
3.4.1. Egyenletes eloszlás	126
3.4.2. Folytonos örökifjú eloszlások	127
3.4.3. Normális eloszlás	129
3.5. Határeloszlás-tételek	133
3.5.1. A nagy számok törvénye	134
3.5.2. A de Moivre–Laplace-tétel	135
3.5.3. A centrális határeloszlás tétele	139

4. Statisztikai módszerek	142
4.1. Statisztikai alapfogalmak	142
4.2. Pontbecslések	147
4.3. Intervallumbecslések	149
4.4. Hipotézisvizsgálat	153
Táblázatok	157
A gyakorlatok, feladatok végeredményei	159
A gyakorlatok, feladatok megoldásai	165
Irodalomjegyzék	180
Tárgymutató	181

Bevezetés

Ez a jegyzet a BME Villamosmérnöki és Informatikai Karának üzemmérnök-informatikus képzésén tartott Valószínűségi számítás B tárgy előadásainak anyagát öleli fel. Célja, hogy a valószínűségi számítás elméletébe, ill. annak legegyszerűbb gyakorlati (első sorban statisztikai) alkalmazásaiba rövid betekintést nyújtson. A tárgyalásmódot alapvetően a kurzus keretei alakították, és bár a jegyzet szándék szerint bárki számára ízelítőt kínálhat a matematika ezen kiemelkedően fontos területéből, a tantárgy szabta korlátok miatt azonban kompromisszumokkal kellett élnünk. Így például számos alapvető fontosságú téma tárgyalását elhagyjuk, a szükséges technikai eszköztárat pedig a lehetőségekhez mérten minimalizáljuk. A jegyzet a középiskolai matematika tantárgy anyagán felül lényegében egy bevezető kalkulus tárgy anyagának ismeretét tételezi fel. Használni fogjuk tehát a határérték fogalmát és annak tulajdonságait, a végtelen sorokra vonatkozó legalapvetőbb ismereteket, és szükségünk lesz a derivált és a Riemann-integrál fogalmára is.

Az anyag feldolgozásához és a jó szemléletmód kialakításához természetesen elengedhetetlenek az absztrakt fogalmak tartalmát megvilágító példák, melyekből a jegyzet is tartalmaz jónéhányat. Ezen felül minden anyagrészben ill. azok után található néhány gyakorlat, esetenként egy-két nehezebb feladat is. A nehézségi különbségeket az elnevezésbeli különbségek is jelzik (vagyis a gyakorlatok szándékaink szerint könnyebbek, míg a feladatok nagyobb kihívást jelenthetnek). Azt javasoljuk, hogy az olvasó próbálja meg ezeket az adott rész feldolgozása közben vagy után, az elméleti anyag tárgyalása során kidolgozott példák mintájának alapján *önállóan* megcsinálni. Minden gyakorlat és feladat végeredménye (amennyiben az egy számszerű végeredmény) megtalálható a jegyzet végén, némelyikhez pedig részletes kidolgozott megoldás is található ott (ez utóbbit mindig egy "(M)" jelöli a gyakorlat vagy feladat szövege előtt). Ezekon felül számos további feladat és segédanyag található a tantárgy honlapján (<https://cs.bme.hu/bvalszam>).

Miért érdemes a valószínűségi számítással foglalkozni?

A valószínűségi számítás a matematika legfontosabb ágainak egyike. A véletlen jelenségek átszövik az élet szinte minden területét, bár a hétköznapi életben ez talán csak ritkán tudatosodik bennünk. Pedig a véletlen fontos szerepet játszik az olyan szituációk nagy részében, ahol esélyeket latolgatunk, ha tehát valamely esemény végső kimenetelét nem tudjuk pontosan meghatározni vagy megbecsülni, például mert azt áttekinthetetlenül sok tényező befolyásolja, vagy pedig ezeket a tényezőket nem ismerjük pontosan. A legkézenfekvőbb példák ezekre természetesen a szerencsejátékok, nem csoda, hogy a valószínűségi számítás korai problémái is erről a területről származnak. Az általunk tárgyalt egyszerű példák jelentős része is ezekhez kapcsolódik, hiszen a kockadobás vagy a lottószámok húzása jól ismert (de nem feltétlenül mindenki által jól értett) jelenségek.

E tudományterület alkalmazásai azonban jóval túlmutatnak a szerencsejátékok körén. A valószínűségszámítás talán legfontosabb alkalmazási területe a statisztika, amelynek eszköztárát lépten-nyomon használja szinte minden tudomány, így az informatika is. Ráadásul számos esetben előfordul, hogy bizonyos adatok, valamint az azokból levont statisztikai következtetések a nagyközönség számára is hozzáférhetőek, ezért különösen fontos, hogy ezekből képesek legyünk a megfelelő információkat kiszűrni, vagy pedig egy rossz, megtévesztő érvelést felismerni. Alkalmazásokat persze magán a matematikán belül is bőségesen találunk, elég csak a véletlent használó algoritmusok elméletét említeni. Ezen kurzus fókuszában persze a valószínűségszámítás elmélete áll, és a fontosabb alkalmazások közül itt csupán egy-két statisztikai módszer bemutatására szorítkozunk.

Az alábbiakban tárgyalt elmélet ma már ugyan klasszikusnak számít, azonban annak precíz megalapozásához egészen a XX. század elejéig kellett várni. Ennek az egyik oka, hogy az ehhez szükséges eszköztár (nevezetesen a mértékelmélet) is ekkor kristályosodott ki, azonban ez a matematika sok más klasszikus ágához mérten késői érés talán annak is betudható, hogy a véletlen jelenségek megértéséhez és helyes leírásához gyakran az intuíciónkkal szöges ellentétben álló szemléletmód kialakítása szükséges. Hogy erről az olvasót is meggyőzzük, bemutatunk egy jól ismert példát (a példa mögött rejlő matematikai modellt az első fejezetben részletesen is elemezni fogjuk).

A Monty Hall-paradoxon

Képzeljük el, hogy egy vetélkedőben három csukott ajtó között kell választanunk, melyekből pontosan az egyik mögött rejtőzik egy nyeremény, míg a másik kettőből választva nem nyerünk semmit (elterjedt az a megfogalmazás, hogy az egyik mögött egy autó, míg a másik kettő mögött egy-egy kecske rejtőzik, ettől itt a precízség kedvéért eltérünk). Azt is tudjuk, hogy a nyereményt véletlenszerűen helyezték el, így tehát logikusnak tűnik az a feltételezés, hogy bármelyik ajtót választva egyforma esélyünk van a nyeremény megszerzésére.

Csavarjunk most egyet a játékszabályokon, és tegyük fel, hogy az ajtó kiválasztása után kinyitnak egyet a másik kettő ajtó közül, amely mögött nincsen nyeremény. Ezt követően dönthetünk, hogy maradunk az eredeti választásunknál, vagy pedig inkább a másik, még csukott ajtót mögé szeretnénk benézni. Mi a jó stratégia? Megéri-e vajon váltani, vagy előnyösebb inkább az eredeti választásunknál maradni? Van-e egyáltalán bármi különbség az esélyeink közt a két esetben? Az utóbbi kérdésen elgondolkodva úgy érvelhetünk, hogy két csukott ajtó maradt, így hát a kettő közül valamelyik mögé rejtették a nyereményt véletlenszerűen, mindegy tehát, hogy melyiket választjuk, az esélyünk a nyeresésre 50%.

A fent bemutatott érvelés azonban hibás. A hiba ott rejtőzik, hogy bár eredetileg valóban nem tudtunk semmit a nyeremény helyéről azon felül, hogy véletlenszerűen választották, viszont az egyik ajtó kinyitásával plusz információt kaptunk, és mivel a nyitott ajtó függ az első választásunktól, így ez a szituáción lényegesen változtat.

Ez talán első hallásra hihetetlennek tűnik, ezért (a pontos matematikai elemzést későbbre halasztva) most a szemléletesség kedvéért kicsit megváltoztatjuk a szituációt, hogy a kapott plusz információ jól érzékelhető legyen. Tegyük fel, hogy most *ezer* ajtó közül választhatunk, melyek közül pontosan az egyik mögött van nyeremény. Miután kiválasztottunk egyet, a maradék 999 ajtó közül kinyitnak 998-at, amelyek mögött nincs nyeremény. Ezután dönthetünk, hogy maradunk az eredeti választásunknál, vagy a maradék 999 ajtóból csukva maradt egy szem ajtóra váltunk. Mit tenne az olvasó ebben a helyzetben?

A jegyzet felépítése

Az alábbiakban a Valószínűségszámítás B tárgy előadásainak felépítését követjük. Az első fejezet a véletlen jelenségek matematikai modelljével, azaz a valószínűségi mezők leírásával foglalkozik. Ezután a második fejezetben a véletlentől függő mennyiségekkel és az ezeket leíró valószínűségi változókkal foglalkozunk, az ezekkel kapcsolatos legfontosabb fogalmakat először csak az ún. diszkrét esetben tárgyaljuk. A harmadik fejezet a folytonos valószínűségi változókkal foglalkozik, ennek a résznek a fő célja a normális eloszlás és ezzel kapcsolatban a centrális határeloszlás tételének bemutatása. Végül a negyedik fejezetben a tanultak néhány egyszerű statisztikai alkalmazását mutatjuk be.

A jegyzet anyagát természetesen nagyban befolyásolta a kar által a BSc képzésben a mérnökinformatikusoknak oktatott Valószínűségszámítás c. tárgy felépítése, az ott szerzett tapasztalatok alapján abból sok mindent átemeltünk ide, de ugyanúgy sok mindent meg is változtattunk. Számos gondolat, példa vagy feladat bekerült Ketskeméty László, Pintér Márta vagy Mészáros Szabolcs egyetemi jegyzeteiből és feladatgyűjteményeiből (és persze a terület hatalmas irodalmából sok egyéb mű is elkerülhetetlenül hatást gyakorolt a szövegre). Mindazonáltal reményeink szerint az alábbiakban kifejezetten a Valószínűségszámítás B tárgy hallgatóira szabott témákat és feladatokat tárgyalunk.

A jegyzetben (minden igyekezet ellenére) természetesen maradhattak hibák, amelyek a `toth.david.akos@vik.bme.hu` címen jelezhetők.

1. fejezet

A véletlen matematikai modellje

Tekintsük a következő kísérletet: dobjunk fel egy érmét ötször egymás után, és jegyezzük fel a dobások eredményét. A kísérlet kimeneteleként így egy öt tagból álló *fej-írás* sorozat adódik. A pontos eredményt persze nem jósolhatjuk meg előre, hiszen azt számos tényező befolyásolja, amelyeket nem látunk át. A kimenetelt tehát *véletlenszerűnek* tekintjük.

Tegyük fel, hogy elvégezve a kísérletet öt fejet dobtunk. Milyen következtetést vonhatunk le az eredményből? Az intuíciónk talán azt súgja, hogy ez az eredmény túl szabályos ahhoz, hogy egy véletlen kísérlet eredménye legyen. Erre persze azt az ellenérvet hozhatjuk fel, hogy bár minden egyes dobásnál csak 50% esélyünk volt éppen fejet dobni, de egy tetszőleges fej-írás sorozatnál is pontosan 50% esélyünk van épp a soron következő eredményt dobni, tehát ez a kimenetel semmivel sem valószínűtlenebb, mint bármely másik. Mégis, talán a nyugtalanító érzés nem múlik el, és azt gondoljuk, hogy a következő dobásnál ezek után most már valószínűleg egy írás kell következzen, hiszen hat fej egymás után már tényleg elég valószínűtlen. De képzeljük el, hogy valaki most csatlakozik hozzánk, hogy megfigyelje a következő dobást, és semmit sem tud az előző eredményekről. Az ő szemszögéből a fej valószínűsége megint csak 50%.

Kinek van igaza? Lehet-e itt egyáltalán objektíven valószínűségekről beszélni? Ha igen, hogyan definiálhatók azok, és mik a helyes következtetések? A fenti kérdések megválaszolásához szükségünk van egy olyan eszköztárra, amely az összes lehetséges véletlen kimenetel áttekintésére alkalmas. Ennek alapját a valószínűségi mező fogalma adja, melyet a fejezet első szakaszában definiálunk. Szeretnénk persze olyan modellt, amit aztán adaptálhatunk az esetlegesen rendelkezésre álló plusz információkhoz. Éppen ezt teszi lehetővé a feltételes valószínűség fogalma, melyet a fejezet második pontjában tárgyalva eljutunk a sztochasztikus függetlenség definíciójához. Már ez az eszköztár is elegendő lesz ahhoz, hogy a fent leírt situációt elemezzük. Ezen felül tárgyaljuk még az ún. Bayes-tételt, melynek a fejezet végén bemutatjuk egy gyakorlati alkalmazását is.

1.1. Valószínűségi mezők

Ahhoz, hogy a matematika eszköztárát alkalmazni tudjuk, modellekre van szükség. Ahogy a térbeli alakzatok leírásához a három dimenziós euklideszi tér vagy az időben változó mennyiségek leírásához a függvény fogalmát használjuk, úgy lesznek segítségünkre a véletlen jelenségek kezeléséhez az ún. *valószínűségi mezők*. Ezek a fenti példáknál összetettebb és talán kevésbé szemléletes struktúrák, melyeket ezért az alábbiakban több lépésben, számos példán illusztrálva vezetünk be.

1.1.1. Eseményterek és események

A fejezet első példájában egy véletlen kísérletet írtunk le, melynek a kimenetele a véletlentől függött. Most ezeket a lehetséges kimeneteket együtt szeretnénk kezelni. Tulajdonképpen az, hogy ezek pontosan hogyan adódtak (tehát maga a kísérlet) nem is feltétlenül lényeges, sőt, véletlen események nem feltétlenül csak klasszikus értelemben vett kísérletek eredményeként figyelhetők meg. Például egy városban az egy napon történt közlekedési balesetek összessége tekinthető egy véletlen eseményeknek, mégsem mondható, hogy azért közlekedünk, hogy megfigyeljük a baleseteket, ráadásul a körülmények ebben az esetben nem is reprodukálhatók.

Éppen ezért a modellünk magukból a lehetséges eseményekből épül fel. Persze a jól használható modell választását befolyásolhatja, hogy az eseményeket milyen folyamat eredményezi, de egyetlen modell akár több szituációban is jól használható lehet. Például egy kockadobás lehetséges eredményeit célszerű lehet az 1, 2, 3, 4, 5 és 6 számokkal leírni, de ugyanez a modell megfelel akkor is, ha egy program kimenetelét szeretnénk kezelni, amely véletlenszerűen generál egy egész számot 1 és 6 között. Egy érme ötszöri feldobása esetén pedig 5 hosszúságú $F - I$ sorozatokkal dolgozhatunk, egy kimenetel lehet például $FIFFI$.

A modellünk tehát egyszerűen úgy áll elő, hogy a lehetséges kimenetek mindegyikéhez páronként különböző matematikai objektumokat (számokat, sorozatokat, halmazokat, stb.) rendelünk, és ezek összessége alkotja az ún. *eseményteret*. Az eseménytér jelölésére az alábbiakban szinte kivétel nélkül az Ω jelölést használjuk, ez tehát egyszerűen egy (az adott szituációtól függő, általunk választott) nem üres halmaz, melynek elemeit *kimeneteknek* vagy *elemi eseményeknek* fogjuk nevezni, és tipikusan ω -val jelöljük őket.

A fenti definíció látszólag óriási szabadságot ad nekünk a tekintetben, hogy hogyan is válasszuk a modellt. Ez részben igaz, azonban a konkrét példákban gyakran természetesen adódik az eseménytér. Persze az is előfordul, hogy több kézenfekvő lehetőség közül választhatunk, de esetenként nem mindegyik egyformán praktikus. A valószínűségszámításban sokszor éppen az jelenti az egyik nagy kezdeti nehézséget, hogy megtaláljuk a megfelelő modellt egy adott szituációhoz. Minden esetre a klasszikus példákban tipikusan mindig ugyanazzal a választással fogunk élni:

Példák.

- Dobjunk egy érmével, ekkor a lehetséges kimenetek a *fej* ill. az *írás*, ennek megfelelően az eseménytér

$$\Omega = \{F, I\}.$$

- Tekintsünk most egy hat oldalú dobókockát, ezzel dobva az

$$\Omega = \{1, 2, 3, 4, 5, 6\}$$

eseménytér egy elemét kapjuk.

- Az ötöslottón véletlenszerűen húznak 5 különböző számot az 1, 2, ..., 90 számok közül. Ekkor a lehetséges kimenetek olyan számötösök, ahol minden szám 1 és 90 közt van. Azaz ha $A = \{1, 2, \dots, 90\}$, akkor Ω az A halmaz 5 elemű részhalmazaiából áll, vagyis

$$\Omega = \{B \subset A : |B| = 5\}. \quad (\text{Itt } |B| \text{ jelöli a } B \text{ halmaz elemszámát.})$$

Figyeljük meg, hogy ebben a példában nem lényeges, hogy a számokat milyen sorrendben húzták ki, pusztán az számít, hogy melyik öt számról van szó. Ezért a számok sorrendjében nem is teszünk különbséget, hanem csak az öt szám halmazát tekintjük.

A véletlen jelenségekkel kapcsolatban persze számos kérdést feltehetünk. Például egy kockadobás esetén nem csak a konkrét eredményről beszélhetünk, hanem esetleg csak az érdekelhet bennünket, hogy 3-nál nagyobbat dobunk-e. A lottóhúzásnál kíváncsiak lehetünk például arra, hogy mekkora eséllyel húzzák ki a kedvenc számunkat.

A fentiek persze ismét véletlen események, azonban nyilván nem volna célszerű minden egyes esetben egy külön eseményteret definiálni, ha egyszer a kérdés egy már modellezett jelenségre vonatkozik. Ehelyett ezeket az eseményeket a már definiált eseményterek segítségével fogjuk leírni. Azt például, hogy 3-nál nagyobbat dobunk, kifejezhetjük úgy is, hogy 4-et, 5-öt vagy 6-ot dobunk, azaz azon elemi események felsorolásával, amikre a fenti állítás igaz, tehát az Ω egy *részhalmazának* megadásával.

1.1.1. Definíció. Legyen Ω egy eseménytér, ekkor az Ω részhalmazait *eseményeknek* nevezzük. Azt mondjuk, hogy az $A \subset \Omega$ esemény egy adott konkrét $\omega \in \Omega$ kimenetel esetén *bekövetkezik*, ha $\omega \in A$.

1.1.1. Példa. Tekintsük ismét a kockadobás példáját, ekkor tehát $\Omega = \{1, 2, 3, 4, 5, 6\}$. Legyen E az az esemény, hogy párosat dobunk, jelölje továbbá P azt, hogy a dobás értéke prímszám. Ekkor az E és P események Ω következő részhalmazaival írhatók le:

$$E = \{2, 4, 6\}, \quad P = \{2, 3, 5\}.$$

Azt az eseményt, hogy egész számot dobunk, maga az Ω halmaz írja le (amely persze részhalmaza önmagának). Talán egy kicsit kevésbé természetes, de tekinthetjük azt az eseményt is, hogy irracionális számot dobunk. Nyilván egyik lehetséges kimenetelre sem teljesül ez az állítás, így tehát ezt az eseményt az üres halmaz adja meg, melynek jele \emptyset .

Az utolsó két esemény, tehát maga az Ω és az üres halmaz minden eseménytér esetén megjelenik, hiszen ezek mindig részhalmazok, tehát a fenti definíció értelmében események lesznek. Mivel Ω minden egyes kimenetel esetén bekövetkezik, ezért ezt *biztos eseménynek* nevezzük. Továbbá, mivel tetszőleges $\omega \in \Omega$ esetén $\omega \notin \emptyset$, ezért az üres halmaz neve *lehetetlen esemény*.

1.1.2. Példa. Tekintsük a lottóhúzás példáját, és legyen H az az esemény, hogy kihúzzák a hármas számot. Ekkor H az $A = \{1, 2, \dots, 90\}$ halmaz azon 5 elemű részhalmazaiából fog állni (tehát azon 5 elemű részhalmazok halmaza), amelyek tartalmazzák a hármas számot is.

Műveletek eseményekkel

Az eseményeket gyakran úgy adjuk meg, hogy leszűkítjük a kimenetek halmazát valamilyen tulajdonság alapján. Ezen tulajdonságokat természetesen logikai műveletek segítségével összekapcsolhatjuk, ezzel újabb eseményeket definiálva, az ezek által eredményezett események pedig valójában halmazelméleti műveletek végeredményeként kaphatók meg.

A legegyszerűbb ezt először egy példán illusztrálni. Az 1.1.1. példában az E eseményt az a tulajdonság definiálta, hogy a kockával páros számot dobunk, míg a P eseményt az, hogy a dobott szám prím. Ekkor tekinthetjük a következő eseményt:

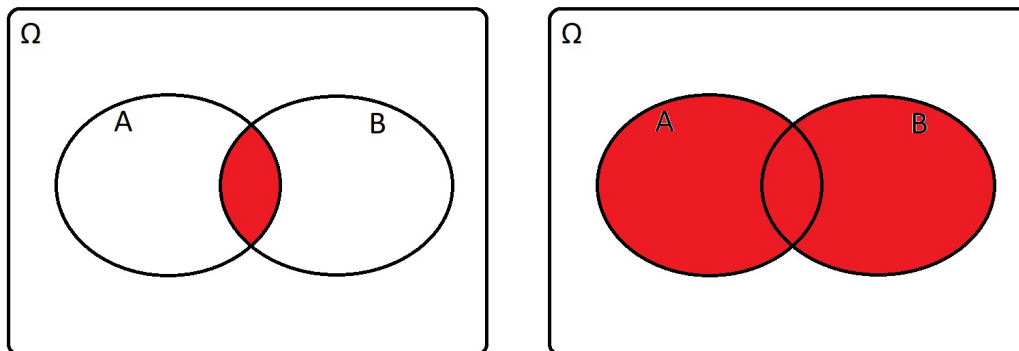
$$\text{a dobott szám páros és prím: } E \cap P = \{2\}.$$

A logikai és műveletnek tehát a halmazelméleti *metszet* felel meg. Általánosan, ha $A, B \subset \Omega$ tetszőleges események, akkor $A \cap B$ azon kimenetelekből áll, melyek mind A -ban, mind B -ben benne vannak.

Visszatérve a fenti példához, kössük össze a két eseményt most a *vagy* logikai művelettel:

a dobott szám páros *vagy* prím: $E \cup P = \{2, 3, 4, 5, 6\}$.

Azaz a logikai *vagy* az események *uniójával* fejezhető ki: $A \cup B$ éppen azon kimenetek halmaza, amelyek az A és B események legalább egyikében benne vannak.



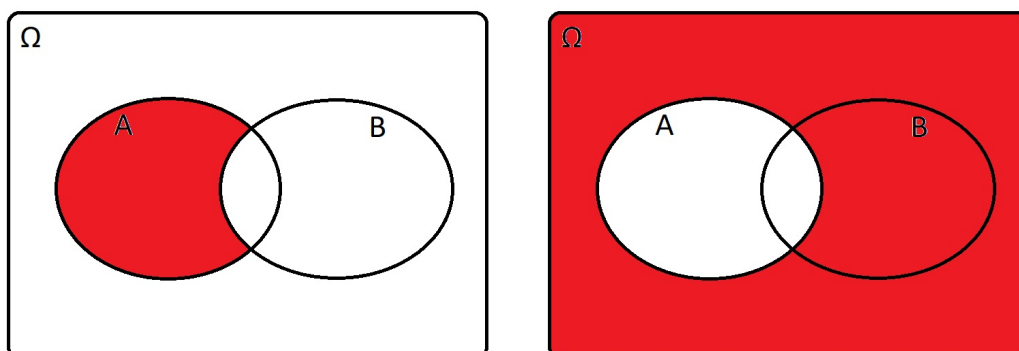
1.1. ábra. Az $A \cap B$ ill. az $A \cup B$ események

Definiáljuk továbbá két esemény *különbségét*: $A \setminus B$ azon kimenetelekből áll, amelyek benne vannak A -ban, de nincsenek benne B -ben. A fenti példában:

a dobott szám páros, *de nem* prím: $E \setminus P = \{4, 6\}$.

Végül a *negáció* halmazelméleti megfelelője a *komplementer*: $\bar{A} := \Omega \setminus A$ azon Ω -beli kimenetek halmaza, amik nincsenek A -ban. Ez utóbbi definíció kapcsán érdemes megjegyezni, hogy a komplementerképzés mindig valamilyen alaphalmazra vonatkozóan történik. Itt tehát az A eseményről feltesszük, hogy az az Ω eseménytér részhalmaza, és az Ω -ra vonatkozóan képezzük a komplementerét. A példánkban:

a dobott szám *nem* páros: $\bar{E} = \{1, 3, 5\}$.



1.2. ábra. Az $A \setminus B$ ill. az \bar{A} események

Az eseményeket persze nem muszáj logikai állítások segítségével definiálni. Ezek mindig egy Ω eseménytér részhalmazai, a fenti műveletek pedig ezáltal definiálva vannak rajtuk. A

gyakorlatban azonban többnyire nem az elemeik felsorolásával adjuk meg őket, hanem a fenti példában látottakhoz hasonló leírásokkal. Éppen ezért is hasznos a fenti logikai műveletek és a halmazműveletek összekapcsolása.

Az imént definiált műveleteknek számos fontos tulajdonsága van, ezek közül most (a teljesség igénye nélkül) felsorolunk néhányat. Azzal a feltételezéssel élve, hogy az olvasó ezek egy részével már korábban is találkozott, az alábbi néhány egyszerű állítás igazolását elhagyjuk (illetve az olvasóra bízunk).

Az unió- illetve a metszetképzés is asszociatív és kommutatív:

$$\begin{aligned} (A \cup B) \cup C &= A \cup (B \cup C) & (A \cap B) \cap C &= A \cap (B \cap C), \\ A \cup B &= B \cup A & A \cap B &= B \cap A. \end{aligned}$$

Érvényes továbbá a következő disztributív szabály is:

$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C).$$

Megjegyezzük azt is, hogy a különbségképzés felírható a metszet és a komplementer segítségével a következőképp: $A \setminus B = A \cap \bar{B}$. A komplementerképzésre vonatkozóan is megadunk még néhány jól használható azonosságot. Először is nyilvánvalóan teljesül $\overline{\bar{A}} = A$. Továbbá, az alábbi két szabályt *de Morgan-azonosságoknak* nevezzük:

$$\overline{A \cup B} = \bar{A} \cap \bar{B}, \quad \overline{A \cap B} = \bar{A} \cup \bar{B}.$$

A fenti állítások több eseményre vonatkozó analogonja is érvényes:

$$\overline{A_1 \cup \dots \cup A_n} = \bigcap_{i=1}^n \bar{A}_i = \bar{A}_1 \cap \dots \cap \bar{A}_n, \quad \overline{\bigcap_{i=1}^n A_i} = \bigcup_{i=1}^n \bar{A}_i,$$

sőt, valójában itt n esemény helyett akár (megszámlálhatóan) végtelen sok esemény uniójának ill. metszetének komplementerét is vehetnénk.

A későbbiekben kiemelten fontos szerephez jut az a szituáció, amikor két eseménynek nincs közös eleme:

1.1.2. Definíció. Legyenek $A, B \subset \Omega$ események. Azt mondjuk, hogy A és B egymást kizáróak (vagy *diszjunktak*), ha metszetük a lehetetlen esemény, azaz ha $A \cap B = \emptyset$ teljesül.

Például az A és \bar{A} események egymást kizáróak minden A eseményre, továbbá $\Omega = A \cup \bar{A}$ is teljesül.

A szakaszt egy technikai megjegyzéssel zárjuk. Az 1.1.1 definícióban valójában nem voltunk egészen precízek. Ha az olvasó más forrásokat is kézbe vesz, a fentieknél komplikáltabb definíciókkal találkozhat. Ennek az az oka, hogy ha teljes általánosságban szeretnénk felépíteni a valószínűségszámítást, akkor az események definíciójában nem engedhetünk meg tetszőleges részhalmazokat. Azonban egyrészt számos esetben nem feltétlenül szükséges korlátozásokkal élnünk (ami lényegesen leegyszerűsíti az anyag tárgyalását). Amennyiben például az eseménytér véges vagy megszámlálhatóan végtelen elemi eseményből áll, akkor ez lényegében nem okoz problémát, azaz egy tetszőleges részhalmaz eseménynek tekinthető. Ennek és a következő fejezetnek a példái kivétel nélkül ilyenek lesznek. Másrészt a fent definiált műveletek általában sem vezetnek ki az események köréből, tehát két esemény uniója, metszete, különbsége, ill. egy esemény komplementere minden esetben esemény lesz, ebből kifolyólag

pedig az összes alábbi állításunk általában is érvényes. Az olvasónak tehát nem kell ezen a ponton a technikai részletekkel foglalkoznia, első megközelítésben nyugodtan képzelheti azt, hogy minden egyes részhalmaz esemény. A precíz definícióhoz a harmadik fejezetben vissza fogunk térni, annak fényében pedig esetleg érdemes újra végiggondolni, hogy az alábbiak miért maradnak érvényben általában is.

Gyakorlatok, feladatok

1.1.1. Gyakorlat. Egyszer dobunk egy szabályos kockával. Jelölje E azt az eseményt, hogy páros számot dobunk, P azt, hogy prímszámot dobunk, N pedig azt, hogy legfeljebb négyet (azaz négyet vagy kevesebbet) dobunk. Legyen továbbá A_i az az esemény, hogy a dobás eredménye i ($i = 1, \dots, 5$), és jelölje még B azt, hogy háromnál nagyobbat dobunk. Fejezzük ki az A_i és a B eseményeket az E , P és N események ill. a halmazműveletek segítségével.

1.1.2. Gyakorlat. Két szabályos dobókockával dobunk. Hogyan definiálnánk az Ω eseményteret? Jelölje S_i azt az eseményt, hogy a dobott számok összege *legalább* i ($i = 2, 3, \dots, 12$), és legyen MAX_j az az esemény, hogy a dobott számok maximumának értéke j ($j = 1, 2, \dots, 6$). Fejezzük ki a fenti események és a halmazműveletek segítségével az alábbi eseményeket:

$$A = \{\text{a dobott számok összege } 7\}, \quad B = \{\text{két darab } 2\text{-est dobunk}\},$$

$$C = \{\text{a dobott számok mindegyike } 1\text{-es vagy } 6\text{-os}\}.$$

1.1.2. Klasszikus valószínűség

Az előző pontban már megalkottuk a véletlen események matematikai modelljét, nincs más hátra, mint hogy definiáljuk ezen események valószínűségét. Ez a valószínűség egyszerűen egy szám lesz: skálázzuk az esélyeket a lehetetlentől a biztosig. Megadhatnánk a skálát például a $[0; 100]$ intervallumon is (mint ahogy a gyakorlatban sokszor ezt is tesszük, amikor százalékban fejezünk ki esélyeket), de a matematika szemszögéből nézve mégis célszerűbb a $[0; 1]$ intervallumot választani. Tehát a valószínűség egy 0 és 1 közötti szám lesz, de hogyan határozhatjuk meg ezt?

A definícióhoz különféle megfontolások vezethetnek. Amennyiben például egy megismételhető véletlen kísérletet sokszor elvégzünk, akkor kézenfekvő a különböző kimenetek és a kísérletek számának arányát, azaz az egyes kimenetek *relatív gyakoriságát* tekinteni. Intuitívan gondolhatunk úgy egy valószínűségre, mint egy olyan számra, amely körül ez a relatív gyakoriság ingadozik. Például ha sokszor dobunk egy szabályos kockával, akkor azt tapasztalhatjuk, hogy mind a hat lehetséges eredmény nagyjából az esetek 1/6 részében adódik. Ez az intuíció olyannyira közel áll az igazsághoz, hogy valójában ez egy matematikai tételnek, az ún. *nagy számok törvényének* egy (nem túl pontos) megfogalmazása. Ezt a tételt (pontosabban annak egyik változatát) a későbbiekben tárgyalni fogjuk. A különböző kimenetek számának összege természetesen az összes kísérlet számát adja, ennek megfelelően a relatív gyakoriságok összege mindig 1 lesz. Hasonlóképp az egyes kimenetek valószínűségének összegétől is elvárjuk a fenti tulajdonságot.

A kockadobás esetében nem csak tapasztalati megfontolások vezethetnek el a valószínűség definíciójához. Ha egy kocka szabályos, az azt jelenti, hogy homogén anyageloszlású és szimmetrikus, így semmilyen fizikai ok nem látszik arra, hogy valamely oldalára többször essen, mint egy másikra. Vagyis ideális esetben minden kimenetel egyformán valószínű kell legyen. Mivel pedig 6 lehetséges kimenetel van, így ez a valószínűség csak 1/6 lehet (feltéve

persze, hogy a 6 kimenetel egyike mindig bekövetkezik). Itt felvetődhet, hogy az ideális eset a valóságban sosem fordul elő. Nincs teljesen szimmetrikus dobókocka, és az anyageloszlás sem lehet tökéletesen homogén. Ha azonban elég közel vagyunk az ideálhoz, akkor az apró eltérések elhanyagolásával vétett hiba valójában jelentéktelenül kicsi, és persze matematikailag lényegesen könnyebb az ideális esettel dolgozni.

Ebben a szakaszban feltesszük, hogy - akárcsak a kockadobásnál - minden itt tekintett eseménytérben egyforma valószínűséggel adódnak az egyes kimenetelek. Az eddigi példánk közül ilyennek tekinthető a pénzérme feldobása, ahol tehát (hacsak mást nem mondunk) a továbbiakban mindig feltesszük, hogy $1/2$ valószínűséggel kaphatunk fejet ill. írást. Egy másik ilyen példa még a lottóhúzás is. Itt a lehetséges számötösök száma $\binom{90}{5}$ (ennek az ún. binomiális együtthatónak a definícióját később, az (1.1) formulában ill. az azt megelőző bekezdésekben adjuk meg), így tehát egy adott számötössel a nyerési esélyünk $\frac{1}{\binom{90}{5}}$. Az alábbiakban további hasonló klasszikus példákat is fogunk látni.

Általában azt mondhatjuk, hogy ha n lehetséges kimenetel van, azaz $|\Omega| = n$, és minden egyes kimenetel egyformán valószínű, akkor ezek valószínűsége $1/n$. Természetesen nem csak az egyes kimenetelek valószínűségéről szeretnénk beszélni, hanem ezt egy tetszőleges $A \subset \Omega$ esemény esetén is szeretnénk meghatározni. Ezen esemény nem más, mint néhány kimenetel halmaza, tehát A egyszerűen ezen kimenetelek valamelyikének bekövetkezését jelenti, továbbá benne minden egyes kimenetel $1/n$ -nel növeli az esélyét annak, hogy A bekövetkezik, így A valószínűségére $|A| \cdot \frac{1}{n} = |A| / |\Omega|$ adódik. Ez persze ezen a ponton még nem egy matematikai érvelés volt, csupán egy olyan gondolatmenet, ami elvezet a valószínűség *definíciójához*:

1.1.3. Definíció. Legyen Ω egy véges eseménytér, és legyen $A \subset \Omega$. Defináljuk az A esemény $\mathbb{P}(A)$ valószínűségét a

$$\mathbb{P}(A) := \frac{|A|}{|\Omega|}$$

formulával. Ekkor azt mondjuk, hogy az Ω eseménytér, a rajta megadott események (tehát Ω részhalmazai) és a fenti formulával definiált valószínűségeik együttesen egy *klasszikus valószínűségi mezőt* alkotnak.

Mielőtt további konkrét példákat tekintenénk, megemlítjük a fent definiált valószínűség néhány fontos tulajdonságát. Először is, mivel $A \subset \Omega$ esetén $0 \leq |A| \leq |\Omega|$ mindig teljesül, így $\mathbb{P}(A)$ értéke valóban a $[0; 1]$ intervallumban lesz.

Továbbá a biztos esemény, azaz Ω mindig bekövetkezik, és ezzel összhangban

$$\mathbb{P}(\Omega) = \frac{|\Omega|}{|\Omega|} = 1$$

is mindig teljesül. Ez természetesen megfelel a valószínűséggel szemben támasztott azon elvárásunknak, hogy az értékészlet, vagyis $[0; 1]$ intervallum jobb széle a teljes bizonyosságot tükrözze. Ugyanez a helyzet a másik végtelennel, a lehetetlen eseményre

$$\mathbb{P}(\emptyset) = \frac{|\emptyset|}{|\Omega|} = \frac{0}{|\Omega|} = 0.$$

A következő kérdés, amivel foglalkozunk, hogy az eseményeken végzett műveletek milyen viszonyban vannak a valószínűséggel. Itt az egyik fontos észrevétel, hogy ha $A, B \subset \Omega$ egymást kizáró események, tehát $A \cap B = \emptyset$, akkor $|A \cup B| = |A| + |B|$, hiszen az utóbbi

összegben minden egyes elemet, ami A -ban vagy B -ben van, pontosan egyszer számolunk. Tehát ha A és B egymást kizáróak, akkor

$$\mathbb{P}(A \cup B) = \frac{|A \cup B|}{|\Omega|} = \frac{|A| + |B|}{|\Omega|} = \mathbb{P}(A) + \mathbb{P}(B).$$

A korábbiakban már láttuk, hogy bármely $A \subset \Omega$ esemény esetén A és \bar{A} egymást kizáróak, továbbá $A \cup \bar{A} = \Omega$, így tehát a fentiek szerint

$$1 = \mathbb{P}(\Omega) = \mathbb{P}(A \cup \bar{A}) = \mathbb{P}(A) + \mathbb{P}(\bar{A})$$

érvényes, átrendezéssel pedig ebből

$$\mathbb{P}(\bar{A}) = 1 - \mathbb{P}(A)$$

adódik.

A fent igazolt tulajdonságokat a későbbi számolások és a feladatmegoldások során is sokszor alkalmazni fogjuk, ráadásul a későbbiekben fontos elvi jelentőségük is lesz, ezért most még egyszer összefoglaljuk őket:

1.1.1. Állítás. *Tekintsünk egy klasszikus valószínűségi mezőt egy Ω eseménytérrel, legyen továbbá $A \subset \Omega$ egy tetszőleges esemény. Ekkor*

(i) $0 \leq \mathbb{P}(A) \leq 1$,

(ii) $\mathbb{P}(\Omega) = 1$ és $\mathbb{P}(\emptyset) = 0$,

(iii) ha továbbá $B \subset \Omega$ olyan esemény, melyre A és B egymást kizáróak, akkor

$$\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B)$$

teljesül. Speciálisan

$$\mathbb{P}(A) + \mathbb{P}(\bar{A}) = \mathbb{P}(\Omega) = 1.$$

A (iii) tulajdonsággal kapcsolatban felmerül a kérdés, hogy mit mondhatunk akkor, ha az A és B események nem (vagy nem feltétlenül) egymást kizáróak. Az érvelésen ekkor sem kell sokat változtatni. Az $A \cup B$ elemszámát ekkor nem az $|A| + |B|$ összeg adja, hiszen ebben kétszer is megszámoltuk azokat az elemeket, amelyek A -ban és B -ben is benne vannak, tehát az $A \cap B$ esemény elemeit. Ezen elemek számát tehát még le kell vonnunk az összegből, vagyis a helyes összefüggés a következő: $|A \cup B| = |A| + |B| - |A \cap B|$. Ez az ún. (kombinatorikai) *szita-formula*, pontosabban annak két halmazra felírt változata. Ebből a valószínűségekre a következő összefüggést kapjuk:

$$\mathbb{P}(A \cup B) = \frac{|A \cup B|}{|\Omega|} = \frac{|A| + |B| - |A \cap B|}{|\Omega|} = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B).$$

A fenti összefüggést szintén *szita-formulának*, ill. *Poincaré-formulának* is nevezzük. A következő szakaszban, illetve később, a második fejezetben még egyszer visszatérünk majd ennek egy általánosabb formájára.

1.1.3. Példa. Tekintsük az 1.1.1. példában definiált eseményeket. Legyen tehát E az az esemény, hogy egy dobókockával párosat dobunk, P pedig az, hogy a dobott szám prímszám. Itt persze az eseménytér ismét, $\Omega = \{1, 2, 3, 4, 5, 6\}$, a korábban látottak alapján pedig

$$\mathbb{P}(E) = \frac{|E|}{|\Omega|} = \frac{3}{6} = \frac{1}{2}, \quad \mathbb{P}(P) = \frac{|P|}{|\Omega|} = \frac{3}{6} = \frac{1}{2}.$$

Számoljuk ki most a $\mathbb{P}(E \cup P)$ valószínűséget. Ezt persze úgy is megtehetjük, hogy meghatározzuk az $E \cup P$ esemény elemeit. Ezek azon kimenetek lesznek, amelyek párosak vagy prímszámok. Tehát $E \cup P = \{2, 3, 4, 5, 6\}$, így

$$\mathbb{P}(E \cup P) = \frac{|E \cup P|}{|\Omega|} = \frac{5}{6}.$$

Használhatjuk ehelyett a szita-formulát is:

$$\mathbb{P}(E \cup P) = \mathbb{P}(E) + \mathbb{P}(P) - \mathbb{P}(E \cap P).$$

Itt az $E \cap P$ esemény azon kimenetekből áll, amik párosak és prímek is egyszerre, ez pedig csak a 2-re teljesül, tehát a keresett valószínűség $\frac{1}{2} + \frac{1}{2} - \frac{1}{6} = \frac{5}{6}$. Bár a második módszerünk ebben a példában komplikáltabbnak tűnik, de sokszor azért használható jól a szita-formula, mert két esemény metszetét gyakran sokkal kényelmesebb leírni, mint az uniójukat.

Számoljuk ki végül a $\mathbb{P}(E \cup \bar{P})$ valószínűséget. Ismét a szita-formulát használva

$$\mathbb{P}(E \cup \bar{P}) = \mathbb{P}(E) + \mathbb{P}(\bar{P}) - \mathbb{P}(E \cap \bar{P}) = \mathbb{P}(E) + 1 - \mathbb{P}(P) - \mathbb{P}(E \cap \bar{P}) = \frac{1}{2} + 1 - \frac{1}{2} - \frac{2}{6} = \frac{2}{3},$$

hiszen $E \cap \bar{P}$ azon kimeneteket tartalmazza, amelyek párosak és nem prímek, tehát a 4-et és a 6-ot.

Eseményterek Descartes-szorzata

Tegyük fel, hogy két szabályos dobókockával dobunk. Mi lehet ez esetben az eseménytér? Egy kimenetelt persze megadhatunk úgy, ha megadjuk, hogy melyik számból mennyit dobtunk. Ez a modell azonban nem feltétlenül praktikus. Egyrészt így a két kockát nem tudjuk megkülönböztetni. Ha például egy hatost és egy egyest dobtunk, nem tudjuk, hogy melyik kockával melyik számot dobtuk. Persze könnyen előfordulhat, hogy a valóságban is két egyforma (pontosabban annak látszó) kockánk van, így még indokoltnak is tűnhet, hogy ne különböztessük meg őket. Ezzel a modellel azonban más gond is lehet. Ha sokszor elvégezzük ezt a kísérletet, megfigyelhetjük, hogy az egyes-hatos kombináció nagyjából kétszer olyan gyakran jön ki, mint mondjuk a két darab hatos. Ennek egyszerűen az az oka, hogy az előbbi - a két kockát tudatosan megkülönböztetve (pl. különbözőképp megjelölve) - akkor is kijön, hogy ha az első kockával dobtunk egyest és a másodikkal hatost, de fordítva is. Ezzel szemben két hatos csak akkor fordulhat elő, ha mindkét kockával hatost dobtunk.

A fő probléma tehát a fent leírt eseménytérrel az, hogy nem egyformán valószínűek az egyes kimenetek, így tehát nem egy klasszikus valószínűségi mezőt kapunk. Ahogy ezt a következő pontban majd részletesen is tárgyaljuk, ez tulajdonképpen technikailag nem okoz gondot. Mégis, sokszor átláthatóbbá válik a helyzet, ha klasszikus valószínűségi mezővel dolgozunk. Ezt a fenti példában egyszerűen elérhetjük azzal, ha a két kockát megkülönböztetjük, és a kimeneteleink olyan $(i; j)$ rendezett számpárok lesznek, ahol az első szám az

első kocka, a második szám pedig a második kocka eredményét adja meg (a rendezettség itt arra vonatkozik, hogy a számpárokban szereplő számok sorrendjét is figyelembe vesszük). Életszerűnek tűnik azt feltételezni, hogy a két dobás eredménye nem befolyásolja egymást, tehát az első kockával az esetek kb. $1/6$ részében adódik bármelyik eredmény az lehetséges számok közül, és ettől függetlenül a második kockára is ugyanez igaz, azaz egy adott $(i; j)$ párt az esetek nagyjából $\frac{1}{6} \cdot \frac{1}{6} = \frac{1}{36}$ részében kapunk. Vagyis tekinthetjük úgy, hogy minden lehetséges kimenetel egyformán valószínű.

Vegyük észre, hogy az utóbbi a megközelítésnek egy olyan előnye is van, hogy az eseményterünk "gazdagabb" lesz abban az értelemben, hogy többféle eseményt tudunk vele leírni (és persze minden olyan eseményt le tudunk írni, amit az elsővel). Valóban, a második modellben meg tudjuk különböztetni az első és második kocka által dobott számokat, míg az elsőben (ha azok különbözők, akkor) nem. Ugyanakkor természetesen a második modellel is leírhatók az első modell kimenetelei, így az azok által leírható események is. Ugyanis azt, hogy a dobott számok i és j , a második modellben az $\{(i; j), (j; i)\}$ esemény fejezi ki.

A második modellünk elemei, azaz a rendezett párok valójában egy ún. Descartes-szorzat elemei. Ha $\Omega = \{1, 2, 3, 4, 5, 6\}$, akkor az Ω -nak az önmagával vett $\Omega \times \Omega$ Descartes-szorzata éppen azon rendezett párokból áll, amelynek első és második eleme is az Ω elemeinek valamelyike. Ez a konstrukció könnyedén általánosítható több halmazra is:

1.1.4. Definíció. Legyenek $\Omega_1, \dots, \Omega_n$ tetszőleges nem üres halmazok. Ekkor ezen halmazok *Descartes-szorzata* az

$$\Omega_1 \times \Omega_2 \times \dots \times \Omega_n := \{(\omega_1, \dots, \omega_n) : \omega_i \in \Omega_i\}$$

halmaz, azaz azon rendezett n -esek halmaza, melyeknek az i -edik eleme az Ω_i halmaz eleme minden $1 \leq i \leq n$ esetén.

Megjegyezzük, hogy a fenti definícióban nem követeljük meg, hogy a halmazaink végesek legyenek, azonban ebben és a következő fejezetben is többnyire ilyen példákkal fogunk találkozni. Amennyiben ez teljesül, akkor a Descartes-szorzat elemszáma

$$|\Omega_1 \times \Omega_2 \times \dots \times \Omega_n| = |\Omega_1| \cdot |\Omega_2| \cdot \dots \cdot |\Omega_n|,$$

hiszen egy rendezett n -es i -edik elemére $|\Omega_i|$ lehetséges választásunk van minden $1 \leq i \leq n$ esetén egymástól függetlenül, így az egyes lehetőségek száma összeszorozódik. Speciálisan, ha $|\Omega_1| = |\Omega_2| = \dots = |\Omega_n| = k$, akkor a Descartes-szorzatuk elemszáma k^n .

A fenti dobókockás példában $|\Omega| = 6$ teljesül, így $|\Omega \times \Omega| = 36$. Itt tehát egy eseménytér önmagával vett Descartes-szorzata szolgált eseménytérként. Számos hasonló példa van, ahol ez a konstrukció jól használható, tipikusan olyan esetekben alkalmazzuk, amikor egymástól független véletlen eseményeket szeretnénk együtt kezelni, azaz egyetlen valószínűségi mezővel leírni. A fő példánkat ebben a fejezetben az érme- vagy kockadobás adja, később további fontos példákat is fogunk látni.

1.1.4. Példa. Feldobunk egy érmét ötször egymás után. Jelölje A azt az eseményt, hogy fej és írás is szerepel a dobások között. Hogyan lehetne modellezni ezt az eseményt?

Nincs okunk feltételezni, hogy egy dobás eredménye befolyásolná bármelyik azt követő dobást, így tehát (hosszú távon) egymástól függetlenül mindegyik (nagyjából) az esetek felében ad fejet vagy írást, egy adott fej-írás sorozat tehát az esetek (nagyjából) $(1/2)^5 = 1/32$ részében adódik.

Legyen $\Omega = \{F, I\}$, ekkor egy ilyen dobássorozatot egy öt hosszú $F - I$ sorozat ír le, azaz éppen az $\Omega' := \Omega \times \Omega \times \Omega \times \Omega \times \Omega$ szorzathalmaz. Ez tehát egy megfelelő jelölt az eseménytérre, a fentiek szerint pedig ennek minden eleme éppen $(1/2)^5 = 1/32$ valószínűséggel kell adódjon kimenetelként. Mivel $32 = 2^5 = |\Omega|^5$ éppen e szorzathalmaz elemszáma, így tehát egy klasszikus valószínűségi mezőről beszélünk.

Az A esemény most azon sorozatokból áll, amelyekben az F és I elemek mindegyike előfordul. Ezeket felsorolni persze fáradságos munka, de észrevehetjük, hogy könnyű azokat a sorozatokat felsorolni, amik nincsenek benne A -ban. Valóban ebből összesen kettő van, a csupa fej ill. a csupa írás sorozat. Azaz $\bar{A} = \{FFFFF, IIIII\}$, és így

$$\mathbb{P}(\bar{A}) = \frac{|\bar{A}|}{|\Omega'|} = \frac{2}{32} = \frac{1}{16}, \quad \mathbb{P}(A) = 1 - \mathbb{P}(\bar{A}) = 1 - \frac{1}{16} = \frac{15}{16} = 0,9375.$$

1.1.5. Példa. Dobjunk egymás után egy 6 oldalú, ill. egy ikozaéder alakú 20 oldalú dobókockával (melyeknek lapjai 1-től 6-ig, ill. 1-től 20-ig vannak számozva). Mi a valószínűsége, hogy a dobott számok összege páros?

A dobások itt sem befolyásolják egymást, így hát az $\Omega = \Omega_1 \times \Omega_2$ eseménytérrel dolgozhatunk, ahol $\Omega_1 = \{1, 2, 3, 4, 5, 6\}$ és $\Omega_2 = \{1, 2, \dots, 20\}$, és egy tetszőleges (i, j) kimenetel, ahol $1 \leq i \leq 6$ és $1 \leq j \leq 20$, éppen $\frac{1}{6 \cdot 20} = \frac{1}{120}$ eséllyel adódik.

Legyen A az az esemény, hogy a dobott számok összege páros. Ez kétféleképp fordulhat elő, mégpedig vagy úgy, hogy mindkét dobásunk páratlan, vagy pedig úgy, hogy mindkét dobás páros. Megszámoljuk azokat az (i, j) párokat, amik ezeket a feltételeket teljesítik. Ha i és j is páratlan, akkor i -t háromféleképp választhatjuk (az értéke lehet 1, 3 vagy 5), és bármelyiket is választjuk, a j értéke ettől függetlenül még tízféleképp választható. Vagyis $3 \cdot 10 = 30$ olyan pár van, melynek mindkét tagja páratlan. Hasonlóképp adódik, hogy 30 olyan pár van, amelynek mindkét tagja páros, és így $|A| = 2 \cdot 30 = 60$, vagyis

$$\mathbb{P}(A) = \frac{|A|}{|\Omega|} = \frac{60}{120} = \frac{1}{2}.$$

1.1.6. Példa. Tegyük fel, hogy hatszor dobunk egy (hat oldalú) dobókockával. A kimeneteleink ekkor 6 hosszú sorozatok, melyek minden tagja egy 1 és 6 közti szám. (Az eseménytér az előző példában definiált Ω_1 halmaz önmagával vett 6-szoros szorzata.) Legyen K az az esemény, hogy a mind a hat dobásunk különböző.

Meghatározzuk K valószínűségét, ehhez megszámloljuk, hogy hány eleme van. Egy K -ban lévő sorozat az 1, 2, 3, 4, 5, 6 számok mindegyikét pontosan egyszer tartalmazza, ez tehát ezen számok egy sorba rendezése, más szóval *permutációja*. Hányféle sorrendje létezik ezen elemeknek? Az első helyre 6-féle számot választhatunk, ha ezt fixáljuk, akkor a második helyre már csak 5-féle szám kerülhet, ezután a harmadik helyre 4-féle, és így tovább. Ez összesen

$$6 \cdot 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1 = 720 = 6! \quad (6 \text{ faktoriális})$$

lehetőség, vagyis

$$\mathbb{P}(K) = \frac{720}{6^6} = \frac{5}{324} \approx 0,0154,$$

tehát annak az esélye, hogy hat különböző számot dobunk, csupán 1,54%.

Az utóbbi példában szereplő okoskodás könnyedén általánosítható n különböző elem esérére:

1.1.5. Definíció. Egy n elemű halmaz elemeinek egy sorba rendezését az elemek egy *permutációjának* nevezzük.

A fenti érvelést 6 helyett n elemre elmondva könnyen adódik, hogy n különböző elem permutációinak száma az első n pozitív egész szorzata:

$$n \cdot (n - 1) \cdot \dots \cdot 2 \cdot 1 = n! \quad (n \text{ faktoriális}).$$

Urnamodellek

Számos olyan szituációval találkozhatunk, amikor egy véletlen esemény néhány elemnek egy adott halmazból való véletlenszerű kiválasztásaként írható le. Ilyen például az, amikor kihúzzunk néhány lapot egy megkevert kártyapakliból, vagy amikor kihúzzák a lottózszámokat. Ezek az események tipikusan leírhatók egy ún. urnamoddellel: adott egy urna, benne n golyó 1-től n -ig számozva, melyeket összekeverünk, majd véletlenszerűen húzzunk belőlük (vakon) k darabot. Feltesszük, hogy egyforma méretű, alakú és azonos tömegeloszlású golyókkal dolgozunk, vagyis a golyók közt a húzás során nem tudunk különbséget tenni.

Ezt az általános problémát esetekre fogjuk bontani, mert az egyes speciális eseteket más-hogyan kell kezelni. Ezeket az eseteket aszerint különböztetjük meg, hogy az egyes húzások után visszatesszük-e a kihúzott golyót az urnába (a következő húzás előtt ismét összekeverve a golyókat), illetve, hogy a kimenetelnél figyelembe vesszük-e, hogy milyen sorrendben húztuk ki az egyes golyókat.

Húzás visszatevéssel és a sorrend figyelembevételével

Ebben a szituációban a kimeneteleink k hosszú sorozatok, melyeknek minden tagja egy 1 és n közti szám, ezek száma pedig n^k . (Az eseménytér tehát az $\{1, 2, \dots, n\}$ halmaz önmagával k -szor vett Descartes-szorzata.) Azt feltételezzük, hogy minden egyes húzásnál egyforma valószínűséggel húzzuk ki mindegyik golyót (hiszen a golyókat újrakeverjük minden húzás után), így pedig minden lehetséges kimenetel $1/n^k$ valószínűséggel adódik.

Vegyük észre, hogy ez éppen a korábbi példáinkkal analóg szituáció. Egy érme feldobása megfeleltethető egy olyan kísérletnek, amikor két golyó közül választunk vakon. Azaz egy érme k -szor való feldobása megfeleltethető k darab húzásnak 2 egyforma golyó közül visszatevéssel. Ugyanígy, egy kockával való dobás megfeleltethető egy húzásnak, ahol 6 egyforma golyó közül választunk vakon, a többszöri dobás pedig egy többszöri húzásnak. Ugyan az eseménytér az iménti példákban nem feltétlenül pontosan ugyanaz, de annak elemszáma és a párba állított események valószínűségei megegyeznek, így a valószínűségszámítás szempontjából az egymásnak megfeleltetett példák lényegében azonosak.

Húzás visszatevés nélkül és a sorrend figyelembevételével

Ha n golyóból úgy húzzunk ki k darabot, hogy a húzott golyókat nem tesszük vissza, de számontartjuk a húzások sorrendjét, akkor a kimeneteleinket olyan k hosszú sorozatokkal írhatjuk le, melyeknek minden tagja egy 1 és n közötti egész szám, és ezek a tagok különbözők (itt persze feltesszük, hogy $k \leq n$). Természetesen az egyes sorozatok egyformán valószínűek (ha valamelyik sorozat valószínűbb volna a többinél, akkor átszámozva a golyókat ugyanezt kapnánk minden sorozatra). Tehát ebben az esetben is egy klasszikus valószínűségi mezőt kapunk.

Számoljuk meg a lehetséges kimenetek számát. Az első húzásnál n különböző golyóból választhatunk, ezután a másodiknál már az első húzástól függetlenül csak $(n - 1)$ -ből (persze

a fennmaradó lehetőségek függnek az első húzástól, de ezek száma nem), és így folytatva, a k -adik húzásra már csak $n - k + 1$ lehetőségünk marad, így összesen

$$n \cdot (n - 1) \cdots (n - k + 1) = \frac{n!}{(n - k)!}$$

lehetséges kimenetel van, és így persze mindegyikük valószínűsége $\frac{(n-k)!}{n!}$.

Természetesen a valószínűségi modelltől elvonatkoztatva általában is beszélhetünk egy n elemű halmaz elemeiből képzett k elemű sorozatokról:

1.1.6. Definíció. Egy n elemű halmazból k különböző elem ($k \leq n$) egy adott sorrendben való kiválasztását az n elem egy k -adosztályú ismétlés nélküli variációjának nevezzük. Ezek száma $n!/(n - k)!$.

Vegyük észre, hogy a $k = n$ esetben az előző definíció éppen az n elem egy permutációját adja. Ekkor a fenti képlet szerint a permutációk száma $n!/(n - n)! = n!/0!$ volna. Hogy ezen (valamint néhány később említésre kerülő) formula ebben a speciális esetben is érvényben maradjon, ezért a $0!$ értékét 1-nek *definiáljuk*.

Az ismétlés nélküli jelző a fenti definícióban természetesen arra utal, hogy a sorba rendezett k elem különböző. Ezzel szemben az első urnamodellünkben olyan sorozatokról beszélünk, ahol megengedtük az ismétlődést, a fentiek mintájára tehát ezeket ismétléses variációknak fogjuk nevezni.

1.1.7. Definíció. Egy n elemű halmazból k (nem feltétlenül különböző) elem egy adott sorrendben való kiválasztását az n elem egy k -adosztályú ismétléses variációjának nevezzük. Ezek száma n^k .

Húzás visszatevés és a sorrend figyelembevétele nélkül

Tekintsük most azt a szituációt, amikor kihúzunk n golyóból k darabot, a golyókat nem tesszük vissza a húzás után, de a húzás eredményénél nem vesszük figyelembe azt, hogy milyen sorrendben húztuk ki a golyókat, csak azt, hogy melyik k golyót húztuk ki. Egy gyakran előkerülő példa erre a lottóhúzás. A korábbiakhoz hasonlóan láthatjuk, hogy minden kimenelt egyformán valószínűnek tekinthetünk, hiszen, ha bármelyik k golyó húzása valószínűbb lenne a többinél, akkor átszámozva a golyókat és így lényegében ugyanazt a kezdeti szituációt előidézve (mégpedig, hogy a golyók közt a húzásnál nem tudunk különbséget tenni) bármelyik k -asra ugyanez adódna.

Adjuk meg most a lehetséges kimenetek számát. Tulajdonképpen az $\{1, \dots, n\}$ halmaz k elemű részhalmazainak számát keressük. Ezekre a kombinatorikában külön elnevezést használnak:

1.1.8. Definíció. Egy n elemű halmaz k elemű részhalmazait az n elem k -adosztályú ismétlés nélküli kombinációinak nevezzük.

Egy adott részhalmazból persze többféle sorrendben kihúzhatjuk a golyókat, éppen ezért célszerű először számításba venni a sorrendet. A fentiek alapján tudjuk, hogy $n!/(n - k)!$ darab különböző k hosszú sorozatot képezhetünk az n különböző elemből. Számoljuk meg, hogy hány esetben szerepel ugyanaz a k elem különböző sorozatokban. Mivel k különböző elemről beszélünk, ezekből nyilván $k!$ különböző sorozatot képezhetünk, tehát minden k -ast

ennyiszor számoltunk meg az összes k hosszú sorozat összeszámolásánál, így azok számát még $k!$ -sal kell osztanunk. Vagyis a különböző k -asok száma

$$(1.1) \quad \binom{n}{k} := \frac{n!}{k!(n-k)!},$$

azaz minden k -as $1/\binom{n}{k}$ valószínűséggel adódik. Megjegyezzük, hogy a fenti formula az $n = k = 0$ esetben a $\binom{0}{0} = 1$ definíciót adja, továbbá $n < k$ és $k < 0$ esetén a binomiális együttható értékét 0-nak *definiáljuk*.

A fenti számot *binomiális együtthatónak* is nevezik, és a bal oldalon álló jelölést a következőképp olvassuk ki: " n alatt a k ". Az fenti számok elnevezését az ún. *binomiális tételben* játszott szerepük indokolja. Ugyanis a tétel a következő formulát adja egy kéttagú összeg n -edig hatványára:

$$(1.2) \quad (x+y)^n = \binom{n}{0}x^n + \binom{n}{1}x^{n-1}y + \binom{n}{2}x^{n-2}y^2 + \dots + \binom{n}{n-1}xy^{n-1} + \binom{n}{n}y^n = \sum_{k=0}^n \binom{n}{k}x^{n-k}y^k,$$

ahol x és y tetszőleges valós számok és n egy pozitív egész. A binomiális együtthatók alábbi egyszerű tulajdonságai azonnal következnek az (1.1) képletből:

$$\binom{n}{k} = \binom{n}{n-k} \text{ minden } 0 \leq k \leq n \text{ esetén,} \quad \binom{n}{0} = \binom{n}{n} = 1, \quad \binom{n}{1} = \binom{n}{n-1} = n.$$

Magának a binomiális tételnek is számos következménye van a binomiális együtthatókra nézve. Hogy a fő gondolatmenettől ne térjünk el túlságosan, itt most csak a legegyszerűbb megadására szorítkozunk (és néhány továbbit tárgyalunk az 1.1.7. feladat megoldásában). Ha tehát az (1.2) formulában az $x = y = 1$ választással élünk, akkor

$$2^n = (1+1)^n = \sum_{k=0}^n \binom{n}{k} \cdot 1^{n-k} \cdot 1^k = \sum_{k=0}^n \binom{n}{k}$$

adódik, tehát a binomiális együtthatók összege (fix n esetén) 2^n .

1.1.7. Példa. Mi a valószínűsége, hogy az ötöslottón minden kihúzott szám páros? Az ötöslottón 90 számból húznak ki 5-öt, tehát az összes lehetőség száma, azaz az Ω eseménytér elemszáma $\binom{90}{5}$. Legyen A az az esemény, hogy az összes kihúzott szám páros. Számoljuk meg, hogy hány ilyen kimenetel van. Ha minden kihúzott szám páros, akkor az összes számot az 1 és 90 közti páros számok halmazából húzták. Ezekből 45 darab van, tehát $\binom{45}{5}$ -féleképp tudunk 5 páros számot húzni. Azaz

$$\mathbb{P}(A) = \frac{|A|}{|\Omega|} = \frac{\binom{45}{5}}{\binom{90}{5}} = \frac{45 \cdot 44 \cdot 43 \cdot 42 \cdot 41}{90 \cdot 89 \cdot 88 \cdot 87 \cdot 86} = \frac{7 \cdot 41}{4 \cdot 29 \cdot 89} = \frac{287}{10324} \approx 0,0278.$$

Tehát nagyjából 2,78% valószínűséggel húznak csupa páros számot. Nem meglepő tehát, hogy a magyar ötöslottón 2022-ben ez mindössze egyszer történt meg.

Binomiális együtthatók közötti összefüggések nem pusztán a binomiális tételből származtathatók, más egyszerű kombinatorikus megfontolások is gyakran vezetnek olyan eredményekre, melyek magából a definiáló (1.1) formulából esetleg csak nehezen látszanak. Ezek az érvelések sokszor nagyon kényelmesen interpretálhatók valószínűségi modelleken keresztül. Erre mutatunk most egy egyszerű példát:

1.1.8. Példa. Véletlenszerűen kihúzzunk $1 \leq k \leq n$ darab golyót visszatevés és a sorrend figyelembevétele nélkül egy urnából, melyben n piros és 1 fehér golyó található. Legyen F az az esemény, hogy kihúzzuk a fehér golyót. Összesen $\binom{n+1}{k}$ lehetőségünk van a k golyó húzására, az pedig, hogy a fehér golyót kihúzzuk $\binom{n}{k-1}$ esetben következik be, hiszen a fehér mellett a maradék $k-1$ golyót a fennmaradó n golyóból ennyiféleképp választhatjuk, tehát

$$\mathbb{P}(F) = \frac{\binom{n}{k-1}}{\binom{n+1}{k}}.$$

Hasonlóképpen a komplementer eseménynek, vagyis annak a valószínűsége, hogy nem húzzuk ki a fehér golyót,

$$\mathbb{P}(\bar{F}) = \frac{\binom{n}{k}}{\binom{n+1}{k}},$$

hiszen $\binom{n}{k}$ -féleképp húzhatunk k golyót az n darab piros közül. Mivel továbbá $\mathbb{P}(F) + \mathbb{P}(\bar{F}) = 1$ is érvényes, így a fenti formulákat behelyettesítve, majd a nevezőkkel felszorozva az

$$(1.3) \quad \binom{n}{k-1} + \binom{n}{k} = \binom{n+1}{k}$$

összefüggés adódik.

Hogy még jobban érzékeltessük a fenti érvelés eleganciáját, levezetjük az (1.3) formulát az (1.1) definícióból is:

$$\begin{aligned} \binom{n}{k-1} + \binom{n}{k} &= \frac{n!}{(k-1)!(n-k+1)!} + \frac{n!}{k!(n-k)!} \\ &= \frac{k \cdot n!}{k!(n-k+1)!} + \frac{(n-k+1) \cdot n!}{k!(n-k+1)!} \\ &= \frac{(n+1) \cdot n!}{k!(n-k+1)!} = \frac{(n+1)!}{k!(n+1-k)!} = \binom{n+1}{k}. \end{aligned}$$

Az utóbbi számolás ugyan nem különösebben nehéz, az azonban nem feltétlenül látszik belőle, hogy tulajdonképpen mit is jelent az (1.3) egyenlet. A két érvelés összevetése jól mutatja, hogy ha egy jelenségnek jó modellt választunk és ezáltal a kapcsolódó matematikai fogalomnak szemléletes jelentést kölcsönzünk, akkor az összefüggések megsejtése (és gyakran az igazolásuk is) természetessé válhat. A fenti példában egy esetszétválasztás (vagy kihúzzuk a fehér golyót, vagy nem) matematikai képletekre való "lefordítása" eredményezte az (1.3) egyenletet. Egy hasonló példa található még az 1.1.8. feladatban.

Melyik modellt válasszuk?

Sok esetben a fenti kérdésre logikusan következik a válasz az adott szituációból. Például a lottóhúzás esetén világos, hogy nem szükséges figyelembe venni a húzás sorrendjét. Vannak azonban olyan esetek, amikor a modell nem ilyen egyértelmű. Tekintsük a következő problémát: egy urnában van két sárga és két piros golyó. Visszatevés nélkül húzva két golyót milyen valószínűséggel húzunk különböző színű golyókat?

A fenti szituációban nem tudjuk, hogy a húzás eredményénél figyelembe vesszük-e a sorrendet vagy sem. Melyik modellt válasszuk hát, hogy a jó valószínűséget kapjuk? A (kevésbé pontos) válasz az, hogy mindegy. Ha precízebbek akarunk lenni, akkor azt mondhatjuk, hogy

amennyiben a kérdéses esemény a lehetséges kimenetek segítségével leírható, úgy a fenti két modell bármelyike ugyanazt a valószínűséget fogja adni.

Tekintsük először azt a megoldást, ahol nem vesszük figyelembe a húzások sorrendjét. Ekkor két golyót a négyből $\binom{4}{2} = 6$ -féleképp húzhatunk (ez tehát az eseményterünk elemszáma), két különböző színűt pedig úgy kaphatunk, ha egyet választunk a sárgákból, egyet pedig ettől függetlenül a pirosakból. Ez tehát $\binom{2}{1} \cdot \binom{2}{1} = 2 \cdot 2 = 4$ lehetőség, vagyis a keresett valószínűség $4/6 = 2/3$.

Oldjuk meg most a feladatot a másik modellel, vagyis tegyük fel, hogy számít a húzás sorrendje. Ekkor $4 \cdot 3 = 12$ -féle kimenetel lehetséges. Továbbá különböző színű golyókat úgy húzhatunk, ha vagy először egy sárgát és másodszor egy pirosat húzunk, vagy pedig először egy pirosat és másodszor egy sárgát. Mindkét eset $2 \cdot 2 = 4$ -féleképp következhet be, tehát összesen 8 esetben húzunk különböző színű golyókat. Vagyis a keresett valószínűség $8/12 = 2/3$.

Arra biztatjuk az olvasót, gondolja át, hogy mi az oka a fenti jelenségnek, tehát hogy miért választhatunk szabadon a két modell között. Végezetül megemlítjük, hogy természetesen vannak olyan esetek, amikor nem használható mindkét modell, mivel nem tudjuk mindkettőben magát a vizsgált eseményt leírni. Például azt az eseményt, hogy elsőre sárgát, másodikra pedig pirosat húzunk, csak abban a modellben kezelhetjük, ahol a húzás sorrendjét is figyelembe vesszük.

Húzás visszatevéssel a sorrend figyelembevétele nélkül (kiegészítő anyag)

Ez a szituáció egy fokkal komplikáltabb az előzőeknél, és voltaképp a későbbiek megértéséhez ennek tárgyalása nem szükséges, pusztán a teljesség kedvéért részletezzük ezt az esetet is. Az itt fellépő probléma megértéséhez tekintsük a következő példát: egy kockával ötször dobunk, és most csak arra vagyunk kíváncsiak, hogy milyen számokat dobtunk, arra nem, hogy milyen sorrendben. A kimeneteket megadhatjuk például úgy, hogy növekvő sorrendben felsoroljuk a dobott számokat:

pl. 1, 1, 2, 2, 3; 1, 1, 1, 1, 1; 1, 2, 3, 4, 5, 6.

Rögtön szembetűnő lehet, hogy például az utolsó két kimenetel nem lehet egyformán valószínű, hiszen a csupa egyes dobás csak akkor valósulhat meg, ha minden alkalommal egyest dobunk, míg az utolsó lehetőség többféleképp adódhat, ha figyelembe vesszük, hogy milyen sorrendben következtek a számok.

Nem célszerű tehát a fenti kimeneteket elemi eseményeknek tekinteni, hiszen így a szituációt nem tudjuk klasszikus valószínűségi mezővel modellezni. Egy kézenfekvő megoldás azonban már a fenti megfontolások alapján is könnyen adódik. Amennyiben a dobások sorrendjét is figyelembe vesszük (tehát az első urnamodellel leírt valószínűségi mezőt használjuk), úgy a fenti eredmények már kifejezhetők eseményekként. Például az 1, 1, 2, 2, 3 eredményhez tartozó A esemény azon dobássorozatokból áll, amelyekben két darab 1-es, két darab 2-es és egy 3-as szerepel.

Ahhoz, hogy az A esemény valószínűségét kiszámoljuk, meg kell határoznunk ezen dobássorozatok számát. Mindezt most teljes általánosságban csináljuk meg. Azaz meghatározzuk az olyan sorozatok számát, amelyekben az n_1, n_2, \dots, n_m különböző elemek szerepelnek, és az n_i elem pontosan k_i darab helyen fordul elő.

1.1.9. Definíció. *Ismétléses permutáció* alatt néhány, egymástól nem feltétlenül különböző elemnek egy sorba rendezését értjük.

Visszatérve a fenti szituációhoz, összesen $n = k_1 + k_2 + \dots + k_m$ elemet kell sorba rendeznünk. Habár az elemek közt lehetnek azonosak (az n_i elemből pontosan k_i fordul elő), ideiglenesen különböztessük meg az összes elemet egymástól. Ekkor persze $n!$ különböző sorrendet kapunk. Ezekből kettő akkor adja pontosan ugyanazt a rendezést az eredeti helyzetben, ha (az ideiglenesen megkülönböztetett) n_i elemek helye megegyezik mindkét rendezésben minden $1 \leq i \leq m$ esetén. Tehát a megkülönböztetett elemek $n!$ -féle sorba rendezése közt minden egyes megkülönböztetés nélküli elrendezést annyiszor számoltunk meg, ahányféleképp az egyes elemeket - fixálva azok helyét - egymás között átrendezhetjük. Ezt az n_i elem esetén $k_i!$ -féleképp tehetjük meg, az egyes elemek pedig egymástól függetlenül rendezhetők át, összesen tehát $k_1! \cdot k_2! \cdot \dots \cdot k_m!$ alkalommal számoltunk meg minden elrendezést.

Összegezve az eddigieket, a különböző ismétléses permutációk száma

$$\frac{n!}{k_1! \cdot \dots \cdot k_m!}.$$

Ez a fenti példánk esetén azt adja, hogy az A eseményt $\frac{5!}{2! \cdot 2! \cdot 1!} = 30$ sorozat alkotja, és így

$$\mathbb{P}(A) = \frac{30}{6^5} = \frac{5}{1296}.$$

Gyakorlatok, feladatok

1.1.3. Gyakorlat. Két kockával dobunk. Legyenek

$$A = \{\text{az összeg } 7\} \quad B = \{\text{mindegyik páros}\} \quad C = \{\text{van közöttük hármas}\}$$

események. Számoljuk ki a $\mathbb{P}(A \cap (B \cup \overline{C}))$ és $\mathbb{P}((A \cup C) \cap \overline{B})$ valószínűségeket.

1.1.4. Gyakorlat. Egy szabályos érmével hatszor dobunk. Mennyi a valószínűsége, hogy

- először az ötödik dobásra kapunk fejet?
- pontosan két fejet dobunk?
- legalább két fejet dobunk?
- a fejek száma páros/páratlan?
- legalább 2 fejet vagy legalább 3 írást dobunk?

1.1.5. Gyakorlat. Mekkora a valószínűsége, hogy az ötöslottón

- pontosan 2 találatunk lesz?
- pontosan k találatunk lesz?
- 13 a legkisebb kihúzott szám?
- a legnagyobb kihúzott szám 80 és 90 közé esik (a 80-at és a 90-et is beleszámítva)?

1.1.6. Gyakorlat. Egy urnában 3 piros, 3 sárga és 3 kék golyó van. Véletlenszerűen húzva 3 golyót az urnából visszatevés nélkül, mi a valószínűsége annak, hogy

- 3 különböző színű golyót húzunk?
- 3 egyforma színű golyót húzunk?

1.1.7. Feladat. (M) A nyúl és a teknősbéka versenyeznek. Miután felálltak a startvonalra, pénzfeldobással döntenek el, hogy ki haladhat előre. Ha fejet dobnak, akkor a nyúl lép előre egy métert, ha írást, akkor a teknős halad előre ugyanennyit. Ezek után mindig megismétlik

az érmedobást a fenti szabályokat alkalmazva. Mi a valószínűsége, hogy 101 dobás után

- a) a nyúl a teknős előtt áll?
- b) a nyúl távolsága a startvonalától (méterben mérve) páros szám?

1.1.8. Feladat. (M) Egy kalapba beteszünk n cédulát, melyekre az $1, \dots, n$ számokat írjuk (minden cédulára pontosan egy számot, és minden számot pontosan egyszer). Visszatevés nélkül kihúzzunk ezek közül véletlenszerűen k darabot ($1 \leq k \leq n$).

- a) Mi a valószínűsége, hogy a legnagyobb kihúzott szám éppen n ?
- b) Mi a valószínűsége, hogy növekvő sorrendbe rendezve a kihúzott számokat az i -edik szám éppen j ($1 \leq i \leq k$, $i \leq j \leq n - k - i$)?
- c) Mutassuk meg, hogy $0 \leq l \leq k \leq n$ esetén $\binom{n+1}{k+1} = \sum_{m=0}^n \binom{m}{l} \binom{n-m}{k-l}$ teljesül.
- d) A c) részben a $k = l = 1$ ill. a $k = 2$, $l = 1$ választásokat használva adjunk képletet az első n pozitív egész összegére ill. az első n négyzetszám összegére.

1.1.9. Feladat. (M) Egy 10 emeletes házban 7 ember egyszerre érkezik ugyanahhoz a lift-hez, és indul felfelé a földszintről. Tegyük fel, hogy mindegyikük (egymástól függetlenül) ugyanolyan valószínűséggel tarthat bármelyik emeletre. Mi a valószínűsége, hogy lesz olyan emelet, ahol a liftbe beszálló 7 emberből többen is kiszállnak?

1.1.3. Valószínűségi mérték

A klasszikus valószínűségi mezők - bár számos példa leírására alkalmasak - messze nem elegendők céljainkhoz. Ez talán már az előző szakaszban is világossá vált, például amikor két kockadobás lehetséges kimeneteleit kíséreltük meg leírni. Ha ugyanis ezt a kockák megkülönböztetése nélkül, pusztán a két eredmény megadásával tennénk, akkor - mint azt korábban megjegyeztük - a kimenetek nem azonos valószínűséggel adódnának, ilyen módon pedig nem kapnánk klasszikus valószínűségi mezőt.

A fenti problémát ugyan sikerült orvosolnunk, de más esetekben erre nincs lehetőség a fogalmak általánosítása nélkül. Egyrészt egy klasszikus valószínűségi mező esetén az Ω eseménytér egy véges halmaz kell legyen, a későbbiekben (főként a 3. fejezetben, de valójában már korábban is) számos olyan példát fogunk látni, ahol ez a feltétel nem teljesül. De véges eseményterek esetén is adódnak problémák. Például egy cinkelt kocka viselkedését nem tudjuk a klasszikus mezővel leírni, hiszen az a különböző kimenetek tekintetében kifejezetten különbözőképp viselkedik.

Ahhoz, hogy a megfelelő általánosságban használható fogalmakat kapjunk, az eddigieknél absztraktabb definíciókra van szükség. Elvonatkoztatunk tehát attól, hogy az események valószínűségét konkrétan milyen formula definiálta, és arra koncentrálnunk, hogy milyen matematikai objektumot kaptunk ezzel a definícióval, és az milyen tulajdonságokkal bír. Először is, a valószínűségek az eseményekhez rendelt számok, vagy másképp fogalmazva, egy valószínűség valójában egy az események halmazán értelmezett, valós értékű *függvény*.

Céljainknak persze nem felel meg akármilyen függvény. Először is, mivel valószínűségekről szeretnénk beszélni, a függvényértékek a $[0; 1]$ intervallumban kell legyenek. Legyen Ω egy eseménytér, és jelölje \mathcal{A} az események halmazát. Amit keresünk, az tehát egy $\mathbb{P} : \mathcal{A} \rightarrow [0; 1]$ függvény. Azt szeretnénk persze, hogy az 1 függvényérték a teljes bizonyosságot reprezentálja, míg 0 legyen a valószínűség akkor, ha valami nem következhet be. Ennek megfelelően teljesülnie kell a $\mathbb{P}(\Omega) = 1$ és $\mathbb{P}(\emptyset) = 0$ egyenlőségeknek. Az is elvárható, hogy ha két esemény egyszerre nem következhet be (tehát vagy az egyik, vagy a másik következik be), akkor

annak a valószínűsége, hogy egyik a kettőből bekövetkezik, az egyes valószínűségek összege legyen. Ezt a matematika nyelvén úgy fogalmazhatjuk meg, hogy ha A és B egymást kizáró események, tehát $A \cap B = \emptyset$ teljesül, akkor a $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B)$ összefüggésnek fenn kell állnia.

Vegyük észre, hogy ezek éppen az 1.1.1. állításban felsorolt, a klasszikus valószínűségekre érvényes tulajdonságok. Az általános esetben lényegében ezeket fogjuk megkövetelni a valószínűségtől. Persze egy fogalom megalkotásánál érdemes minél kevesebb feltételt szabni, így például csak az fogjuk feltenni, hogy $\mathbb{P}(\Omega) = 1$ teljesül, a $\mathbb{P}(\emptyset) = 0$ pedig majd a feltételeinkből következni fog. Továbbá, mivel végtelen eseménytereket is kezelni szeretnénk, ezért az 1.1.1. állítás (iii) tulajdonságának megszámlálhatóan végtelen sok halmazra kiterjesztett változatát fogjuk előírni. Minden készen áll tehát az alábbi definícióhoz:

1.1.10. Definíció. Legyen Ω eseménytér, \mathcal{A} pedig ezen eseménytér eseményeinek halmaza. Tegyük fel, hogy a $\mathbb{P} : \mathcal{A} \rightarrow [0; 1]$ függvényre a következők teljesülnek:

(i) $\mathbb{P}(\Omega) = 1$,

(ii) (σ -additivitás) ha $A_1, \dots, A_n, \dots \in \mathcal{A}$ páronként egymást kizáró események, vagyis $A_i \cap A_j = \emptyset$ teljesül minden $i, j \in \mathbb{N}^+$, $i \neq j$ esetén, akkor

$$\mathbb{P}\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mathbb{P}(A_i).$$

Ekkor a \mathbb{P} függvényt az \mathcal{A} halmazon értelmezett *valószínűségi mértéknek* nevezzük. Továbbá, ha \mathbb{P} valószínűségi mérték az \mathcal{A} halmazon, akkor az $(\Omega, \mathcal{A}, \mathbb{P})$ hármast *valószínűségi mezőnek* nevezzük.

A fenti definícióban nem voltunk teljesen precízek az \mathcal{A} halmaz tekintetében. Valójában az események halmazának egy speciális szerkezettel kell bírnia, általában fel kell tenni, hogy \mathcal{A} egy ún. σ -algebra. Ezekkel a technikai részletekkel ebben és a következő fejezetben nem foglalkozunk, az olvasó itt nyugodtan feltételezheti, hogy \mathcal{A} az Ω összes részhalmazának halmaza. A 3. fejezetben röviden visszatérünk még erre a kérdésre.

Megjegyezzük, hogy az 1.1.1. állításban felsorolt tulajdonságokból levezethető, hogy a klasszikus valószínűség teljesíti a fenti definícióban felsorolt (i) és (ii) tulajdonságokat, vagyis a klasszikus valószínűség is egy valószínűségi mérték. Másképp szólva a fenti definíció a klasszikus valószínűség egy általánosítása. A definícióban szereplő (i) tulajdonság teljesülése a klasszikus valószínűség esetén az 1.1.1. állítás része, a σ -additivitás pedig viszonylag könnyen levezethető ebből az állításból és abból a feltételből, hogy klasszikus valószínűségi mezők esetén véges eseménytérrel dolgozunk. A részletek kidolgozását itt elhagyjuk.

A alábbiakban igazoljuk a valószínűségi mérték néhány alapvető tulajdonságát.

1.1.2. Állítás. Legyen $\mathbb{P} : \mathcal{A} \rightarrow [0; 1]$ egy valószínűségi mérték egy eseménytér eseményeinek \mathcal{A} halmazán. Ekkor

(i) $\mathbb{P}(\emptyset) = 0$,

(ii) (additivitás) ha $A_1, \dots, A_n \in \mathcal{A}$ páronként egymást kizáró események, akkor

$$\mathbb{P}(A_1 \cup \dots \cup A_n) = \mathbb{P}(A_1) + \dots + \mathbb{P}(A_n),$$

speciálisan, ha $A, B \in \mathcal{A}$ és $A \cap B = \emptyset$, akkor

$$\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B),$$

(iii) minden $A \in \mathcal{A}$ -ra $\mathbb{P}(\bar{A}) = 1 - \mathbb{P}(A)$,

(iv) ha $A, B \in \mathcal{A}$ és $B \subset A$, akkor $\mathbb{P}(B) \leq \mathbb{P}(A)$.

A (iv) tulajdonságnál fennálló szituációban, tehát ha az A és B eseményekre $B \subset A$ teljesül, azt mondjuk, hogy a B esemény maga után vonja A -t. A szóhasználatot az indokolja, hogy minden kimenetel, ami B -ben van, egyben A -ban is benne van, tehát ha egy kimenetelre B bekövetkezik, akkor ez maga után vonja az A bekövetkezését is.

Bizonyítás. Az (i) tulajdonság igazolásához használjuk a valószínűségi mérték σ -additivitását az $A_i = \emptyset$ választással minden $i \geq 1$ index esetén. Ekkor az A_i halmazok páronként diszjunktak, így

$$\mathbb{P}(\emptyset) = \mathbb{P}\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mathbb{P}(A_i) = \sum_{i=1}^{\infty} \mathbb{P}(\emptyset).$$

Itt a bal oldalon egy 0 és 1 közötti szám, a jobb oldalon pedig egy végtelen összeg áll, melyben minden tag a bal oldalon álló szám. Ennek a végtelen összegnek tehát egy véges értéket kell adnia, ez pedig csak akkor lehetséges, ha $\mathbb{P}(\emptyset) = 0$.

A (ii) tulajdonság a következőképp adódik az imént igazolt (i) állításból és a mérték σ -additivitásából. Legyenek $A_1, \dots, A_n \in \mathcal{A}$ páronként diszjunkt halmazok, és legyen minden $i > n$ esetén $A_i = \emptyset$. Ekkor az $A_1, \dots, A_n, A_{n+1}, \dots \in \mathcal{A}$ halmazok továbbra páronként diszjunktak, és így

$$\mathbb{P}(A_1 \cup \dots \cup A_n) = \mathbb{P}\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mathbb{P}(A_i) = \sum_{i=1}^n \mathbb{P}(A_i) + \sum_{i=n+1}^{\infty} \mathbb{P}(A_i).$$

Itt az első egyenlőség azért áll fenn, mert az első n halmaz uniójához már csak üres halmazokat veszünk hozzá, míg a második egyenlőség a \mathbb{P} σ -additivitásából következik. Viszont a jobb oldalon álló második összegben minden tag $\mathbb{P}(\emptyset)$, ez pedig a fent látottak szerint 0, így a (ii) állítást beláttuk.

A (iii) állításhoz legyen $A \in \mathcal{A}$. Ekkor A és \bar{A} egymást kizáróak, azaz $A \cap \bar{A} = \emptyset$, továbbá $\Omega = A \cup \bar{A}$, így a valószínűségi mérték definíciójában szereplő (i) tulajdonság és a mérték additivitása miatt

$$1 = \mathbb{P}(\Omega) = \mathbb{P}(A \cup \bar{A}) = \mathbb{P}(A) + \mathbb{P}(\bar{A}),$$

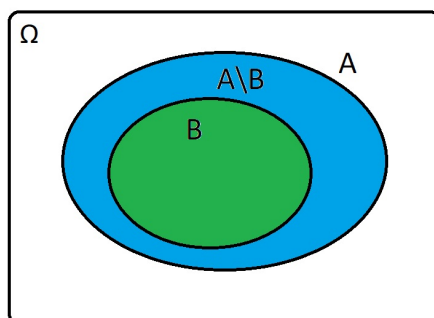
ezt az egyenletet átrendezve pedig éppen a bizonyítandó állítást kapjuk.

Végül a (iv) állítás igazolásához legyen $A, B \in \mathcal{A}$, melyekre $B \subset A$ teljesül. Az A elemeit két csoportra oszthatjuk aszerint, hogy a B halmaz egy eleméről van szó vagy sem, és mivel A tartalmazza az egész B -t, így $A = B \cup (A \setminus B)$ teljesül (lásd az 1.3. ábrát). A jobb oldalon egymást kizáró események uniója áll, így a \mathbb{P} additivitása miatt

$$\mathbb{P}(A) = \mathbb{P}(B \cup (A \setminus B)) = \mathbb{P}(B) + \mathbb{P}(A \setminus B) \geq \mathbb{P}(B),$$

mert $\mathbb{P}(A \setminus B) \geq 0$. □

Ahogy a klasszikus valószínűségi mező esetén, úgy itt is felmerül a kérdés, hogy mit mondhatunk két esemény uniójának valószínűségéről, amennyiben azok nem (feltétlenül) egymást kizáróak. A válasz természetesen ugyanaz, mint a korábbi speciális esetben, és az általános formula igazolása általában sem sokkal komplikáltabb.



1.3. ábra. $B \subset A \implies A = B \cup (A \setminus B)$

1.1.3. Állítás (Szita-formula 2 eseményre). *Legyenek A, B egy valószínűségi mező tetszőleges eseményei, ekkor*

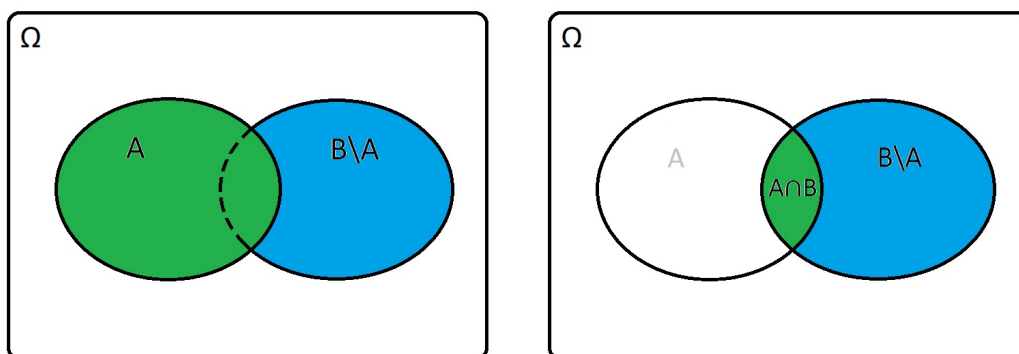
$$\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B).$$

A klasszikus esetben ez abból következett, hogy az A és B halmazok elemszámát külön-külön megszámlálva a közös elemeket kétszer számoltuk. Az általános esetben is analóg módon igazolható az állítás, itt elemszám helyett az egyes halmazok mértékét adjuk össze, a közös részt kétszer "megmérve" így. Mindez precízen a következőképp néz ki:

Bizonyítás. Az $A \cup B$ esemény 2 egymást kizáró részre bontható úgy, hogy csoportosítjuk az elemeit aszerint, hogy benne vannak A -ban vagy sem. Utóbbi kimenetek a $B \setminus A$ halmazt alkotják, így tehát $A \cup B = A \cup (B \setminus A)$ (lásd 1.4. ábrát alább), és a valószínűségi mérték additivitása miatt $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B \setminus A)$. Csempésszük be a jobb oldalra a $\mathbb{P}(A \cap B)$ értéket egyszer pozitív, egyszer pedig negatív előjellel, a kifejezés értékét így nem változtatva:

$$(1.4) \quad \mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B \setminus A) + \mathbb{P}(A \cap B) - \mathbb{P}(A \cap B).$$

Végül, mivel $B = (B \setminus A) \cup (A \cap B)$ a fentiekhez hasonlóan a B esemény egy egymást kizáró eseményekre való felbontását adja (lásd az 1.4. ábrát), így $\mathbb{P}(B) = \mathbb{P}(B \setminus A) + \mathbb{P}(A \cap B)$, ezt pedig (1.4) jobb oldalán behelyettesítve a bizonyítandó állítást kapjuk. \square



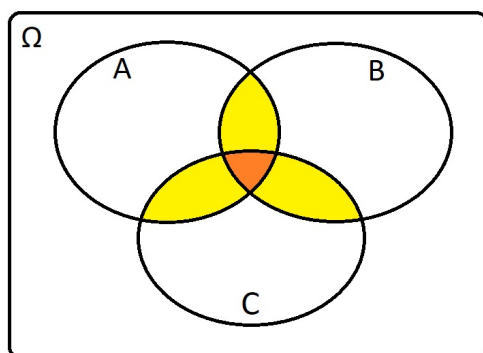
1.4. ábra. $A \cup B = A \cup (B \setminus A)$, $B = (B \setminus A) \cup (A \cap B)$

A szita-formulát szokás Poincaré-formulának is nevezni, és a fenti állítás kettőnél több halmaz uniójára is általánosítható. Az alábbiakban megadjuk a három halmazra vonatkozó

változatot, az általánosított, n halmaz uniójára vonatkozó verzió megtalálható pl. az [5] jegyzet 1.4. szakaszában. Az alapötlet három esemény esetén is az, hogy három halmazt külön-külön megmérve a halmazpárok közös részeit duplán számoljuk, így azok mértékét le kell vonni, ekkor viszont valójában azt a részt, ami mindhárom eseményben benne van, végül egyszer sem számoljuk, ezért annak mértékét még hozzá kell adni az egészhez. A precíz bizonyítást itt elhagyjuk (illetve az olvasóra bízunk), viszont megjegyezzük, hogy a később a 2.3.4. példában a várható érték egy alkalmazásaként egy alternatív bizonyítást nyerünk a következő állításra.

1.1.4. Állítás (Szita-formula 3 eseményre). *Legyenek A, B, C egy valószínűségi mező tetszőleges eseményei, ekkor*

$$\mathbb{P}(A \cup B \cup C) = \mathbb{P}(A) + \mathbb{P}(B) + \mathbb{P}(C) - \mathbb{P}(A \cap B) - \mathbb{P}(B \cap C) - \mathbb{P}(C \cap A) + \mathbb{P}(A \cap B \cap C).$$



1.5. ábra. A $\mathbb{P}(A) + \mathbb{P}(B) + \mathbb{P}(C)$ összegben egyes részeket többször megmérünk

Példák valószínűségi mértékre

A valószínűségi mértékre eddig az egyetlen konkrét példánk a klasszikus valószínűség volt. Hogyan adhatunk további példákat? Emlékeztetünk, hogy az általánosítás szükségességének illusztrálására többek között azt a szituációt használtuk, amikor egy szabálytalan dobókocka viselkedését szeretnénk leírni. Tegyük fel, hogy egy ilyen kockával nagyon sokszor dobva az esetek $\frac{16}{100}$ részében dobunk egyest vagy kettést, továbbá az esetek $\frac{17}{100}$ részében dobunk hármaszt, négyest, ötöst és hatost. Ekkor jogosnak érezhetjük azt mondani, hogy a különböző dobások valószínűségeit a fenti törtek adják. Ha tehát ezt a szituációt modellezni szeretnénk az $\Omega = \{1, 2, 3, 4, 5, 6\}$ eseményteret használva, akkor az egyetlen kimenetelből álló eseményekre

$$\mathbb{P}(\{1\}) = \mathbb{P}(\{2\}) = 0,16 \quad \text{és} \quad \mathbb{P}(\{3\}) = \mathbb{P}(\{4\}) = \mathbb{P}(\{5\}) = \mathbb{P}(\{6\}) = 0,17$$

kell teljesülnön.

Vegyük észre, hogy ha egy valószínűségi mérték a fenti egyenlőségeket teljesíti, akkor ezek már meghatározzák annak értékét az Ω egy tetszőleges részhalmaza, azaz minden esemény esetén. Valóban, bármely esemény ilyen kimenetelekből álló (egy elemű) események páronként diszjunkt uniója, és így a mérték additivitása miatt annak valószínűsége az öt alkotó kimenetelek valószínűségének összege kell legyen. Például az $E = \{2, 4, 6\}$ eseményre

$$E = \{2\} \cup \{4\} \cup \{6\},$$

és így a $\mathbb{P}(E) = 0,16 + 2 \cdot 0,17 = 0,5$ egyenlőségnek teljesülnie kell.

Figyeljük meg, hogy a kimenetek valószínűségének megadása ugyan meghatározza az egyetlen lehetséges \mathbb{P} függvényt, de azt nem mutattuk meg, hogy ez a valóban valószínűségi mérték, tehát teljesíti az 1.1.10. definícióban szereplő (i) és (ii) feltételeket. Az (i) tulajdonság, vagyis $\mathbb{P}(\Omega) = 1$ teljesülése nyilvánvaló a fenti példában, mert a kimenetek valószínűségeinek összege 1. A (ii) tulajdonság igazolása sem nehéz, de ezt itt nem részletezzük. A két tulajdonság teljesülése viszont azt jelenti, hogy valóban valószínűségi mértéket adtunk meg.

A fenti módszer szerencsére gond nélkül általánosítható, amennyiben az Ω eseménytér véges vagy megszámlálhatóan végtelen (azaz az elemei a pozitív egészekkel indexelve felsorolhatók). Ebben és a következő fejezetben csak ilyen példákkal fogunk találkozni, az ilyen eseménytereket *diszkrétnek* fogjuk hívni.

A diszkrét esetben tehát elegendő megadni a kimenetek valószínűségét, azaz hozzárendelni minden kimenetelhez egy súlyt a $[0; 1]$ intervallumban. Természetesen arra figyelni kell, hogy ezeknek a valószínűségeknek 1 legyen az összege, hiszen a $\mathbb{P}(\Omega) = 1$ egyenletnek érvényesnek kell lennie. Mindez azonban már elegendő ahhoz, hogy egy valószínűségi mértéket egyértelműen meghatározzon, ahol egy tetszőleges esemény valószínűsége az őt alkotó kimenetek valószínűségének összege. Ezt a következő állításban formálisan is leírjuk.

1.1.5. Állítás. *Legyen Ω egy véges vagy megszámlálhatóan végtelen halmaz. Legyen továbbá $p : \Omega \rightarrow [0; 1]$ egy tetszőleges súlyfüggvény, amely egy $\omega \in \Omega$ kimenetelhez a $p(\omega) \in [0; 1]$ súlyt rendel. Tegyük fel, hogy*

$$\sum_{\omega \in \Omega} p(\omega) = 1$$

teljesül. Definiáljuk a \mathbb{P} függvény értékét egy $A \subset \Omega$ eseményre a következőképp:

$$(1.5) \quad \mathbb{P}(A) := \sum_{\omega \in A} p(\omega).$$

Ekkor a \mathbb{P} függvény egy valószínűségi mérték az Ω eseményeinek halmazán. Továbbá a \mathbb{P} egyértelműen meghatározott abban az értelemben, hogy ez az egyetlen valószínűségi mérték Ω -n, amelyre $\mathbb{P}(\{\omega\}) = p(\omega)$ teljesül minden $\omega \in \Omega$ esetén.

1.1.9. Példa. Feldobunk két pénzérmét. Az egyes dobásokat ezúttal (kivételesen) nem különböztetjük meg, a kimenetek például megadhatók úgy, ha megadjuk a fejek számát. Dolgozzunk tovább az $\Omega = \{0, 1, 2\}$ eseménytérrel. Mivel 1 fejet kétféleképp is kaphatunk, így értelemszerűen az egyes kimenetekhez a következőképp rendelünk valószínűségeket (súlyokat):

$$p(0) = p(2) = \frac{1}{4}, \quad p(1) = \frac{1}{2}.$$

A fenti állítás értelmében a p valószínűségi súlyfüggvény egyértelműen meghatároz egy valószínűségi mértéket az Ω eseményeinek halmazán. Például annak a valószínűsége, hogy dobunk fejet, az (1.5) formula szerint

$$\mathbb{P}(\{1, 2\}) = p(1) + p(2) = \frac{3}{4}.$$

Ebben a példában a p súlyfüggvényt azért választottuk így, hogy jól használható modellt kapjunk a tényleges pénzdobásra. Elvi szempontból azonban semmi akadályja annak, hogy teljesen más valószínűségeket rendeljünk az egyes kimenetekhez (például a konstans $\frac{1}{3}$ súly

éppen a klasszikus valószínűséget adná), azonban más választások nem volnának alkalmasak a valóság jó leírására. Minden esetre azt érdemes kiemelni, hogy az eseménytérből nem következik maga valószínűségi mérték választása, ezt az Ω halmaztól függetlenül határozhatjuk meg, éppen ezért kezeljük ezeket az objektumokat külön egymástól. Amikor a valószínűségi mérték is rögzített, akkor már nem csak egy eseménytérrel, hanem egy valószínűségi mezőről beszélhetünk.

Az 1.1.5. állítás igazolása ugyan nem különösebben nehéz, de a technikai részletek igényelnek némi körültekintést, és ezt a bizonyítást itt nem tárgyaljuk. Maga az állítás azt a csalóka érzést keltheti, hogy tulajdonképpen megoldottuk a véletlen jelenségek modellezésének problémáját, azonban az Ω számosságára vonatkozó feltétel az állításban óriási korlátozást jelent. Valójában a leggyakoribb alkalmazások többségéhez nem elegendőek a megszámlálható (tehát véges vagy megszámlálhatóan végtelen) eseményterek. A leggyakrabban előkerülő ilyen példák a véges halmazokon túl a természetes számok \mathbb{N} halmaza, az egész számok \mathbb{Z} halmaza, sőt belátható, hogy a racionális számok \mathbb{Q} halmaza is megszámlálható. Ha például azonban egy valós intervallumban felvett véletlen értékről beszélünk (amely tehát egy tetszőleges valós, akár irracionális szám lehet), akkor a lehetséges értékek halmaza, vagyis az intervallum elemeinek száma már nem megszámlálható. Ez az oka annak, hogy a 3. fejezetben már komolyabb eszköztárat kell majd bevetnünk, de a véletlen egyes törvényei jól megérthetők a diszkrét eseményterek példáin keresztül is. Ráadásul, mivel a törvényszerűségek levezetése során sokszor csak a mérték általános tulajdonságait használjuk fel, számos eredményünk (mint például az 1.1.2. állítás) a bizonyításával együtt érvényes az általános esetben is.

Gyakorlatok, feladatok

1.1.10. Gyakorlat. Jelölje A azt az eseményt, hogy az ötöslottón minden kihúzott szám értéke legalább 10, legyenek továbbá B_9 ill. B_{10} azon események, hogy kihúzzák a 9-es ill. 10-es számot. Határozzuk meg a $\mathbb{P}(A \cup B_{10})$ és $\mathbb{P}(A \cup B_9 \cup B_{10})$ valószínűségeket.

1.1.11. Gyakorlat. Egy termék próbagyártása során két szempontból vizsgálják a késztermékeket. Az A esemény azt jelenti, hogy egy véletlenszerűen kiválasztott mintadarab anyaghibás, a B pedig az az esemény, hogy a kiválasztott gyártmány mérethibás. Tudjuk, hogy $\mathbb{P}(A) = 0,15$, $\mathbb{P}(B) = 0,3$ és $\mathbb{P}(A \cap B) = 0,08$. Mennyi annak a valószínűsége, hogy

- egy termék anyaghibás, de nem mérethibás?
- egy termék hibátlan?

1.1.12. Gyakorlat. Az A , B és C események mindegyike $\frac{1}{3}$ valószínűséggel következik be, továbbá az A és B , ill. a B és C események egymást kizáróak. Adjuk meg $A \cap C$ valószínűségét, ha tudjuk, hogy $\frac{1}{9}$ annak a valószínűsége, hogy az A , B és C események egyike sem következik be.

1.1.13. Feladat. Mutassuk meg, hogy egy tetszőleges A eseményre $\mathbb{P}(A)\mathbb{P}(\bar{A}) \leq \frac{1}{4}$ teljesül.

1.1.14. Feladat. Igazoljuk, hogy bármilyen $n \geq 1$ egészre és $A_1, \dots, A_n \subset \Omega$ eseményekre teljesülnek a következő egyenlőtlenségek:

$$\mathbb{P}(A_1 \cup \dots \cup A_n) \leq \mathbb{P}(A_1) + \dots + \mathbb{P}(A_n), \quad \mathbb{P}(A_1 \cap \dots \cap A_n) \geq \mathbb{P}(A_1) + \dots + \mathbb{P}(A_n) - n + 1.$$

1.2. Feltételes valószínűség

Ebben a pontban azt vizsgáljuk meg, hogy a véletlen eseményekhez rendelt valószínűségeket hogyan változtatják meg bizonyos, az eredeti helyzethez képest új feltételezések vagy a szituációval kapcsolatban nyert plusz információk. Ez utóbbiakra már a jegyzet bevezetőjében leírt Monty Hall-paradoxon esetén is láttunk példát. Továbbá ezen fejezet elején már volt szó a következő szituációról. Tegyük fel, hogy többször feldobunk egy érmét, de az első néhány esetben nem kapunk egyetlen fejet sem. Ezek után egy új megfigyelő csatlakozik a kísérlethez. Az ő szemszögéből a következő dobás az első megfigyelt véletlen esemény. Változtat-e valamit a helyzeten, ha eláruljuk neki, hogy ezidáig egyetlen fejet sem dobtunk?

Mindkét példa elemzéséhez szükségünk lesz a feltételes valószínűség fogalmára, a fenti második példa pedig egy speciális esetet ír le, mégpedig azt, amikor az új információk nincsenek hatással a valószínűségekre, az a feltételtől függetlenül változatlan marad. A szakasz második felében pedig a feltételes valószínűség néhány alkalmazását, nevezetesen a teljes valószínűség tételét, a Bayes-tételt és a szorzási szabályt tárgyaljuk.

1.2.1. Feltételes valószínűség és függetlenség

Tekintsük a kockadobás kísérletét. Tegyük fel, hogy egy dobás után valaki elárulja nekünk, hogy a dobott szám páros, azaz bekövetkezett az $E = \{2, 4, 6\}$ esemény. Milyen valószínűségi modellt használjuk ebben az esetben? Természetesen megváltoztathatjuk az eseményteret, hiszen itt már csak három kimenetel lehetséges, tehát élhetünk az $\Omega = E$ választással. Az is értelemszerű, hogy ez a három kimenetel ekkor ugyanolyan valószínű, tehát mindhárom eset $1/3$ eséllyel következik be.

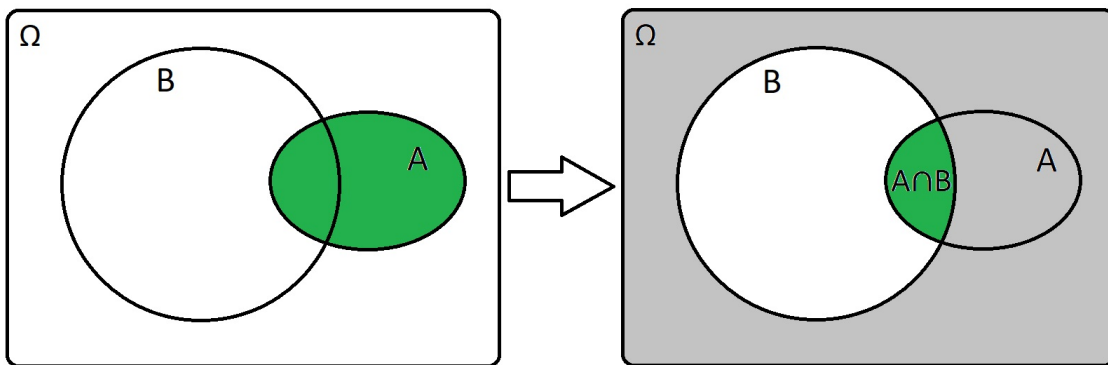
A fenti választás azonban nem feltétlenül praktikus, például lehetséges, hogy további dobások esetén ez az információ nem áll majd rendelkezésünkre, és akkor ismét más modellt kell alkalmaznunk. Az is lehet, hogy megtudjuk, hogy páros-e az eredmény, de ekkor a két eshetőség két különböző eseményteret eredményezhet, ez pedig feleslegesen komplikálttá és nehézkesé teszi a kísérletek precíz matematikai leírását.

Az eseménytér megváltoztatása helyett a *feltételes valószínűség* fogalmát fogjuk használni, amely oly módon szűkíti le az eseményteret a feltételezetten bekövetkező eseményre, hogy az azon kívül eső kimenetek valószínűségét nullává teszi. Ezzel ekvivalens, hogy a feltételben szereplő esemény 1 valószínűséggel bekövetkezik. Ezt pedig úgy érjük el, hogy egyszerűen leosztunk annak valószínűségével:

1.2.1. Definíció. Legyenek A, B egy valószínűségi mező eseményei, ahol $\mathbb{P}(B) > 0$ teljesül. Ekkor az A eseménynek a B eseményre vett *feltételes valószínűsége*

$$(1.6) \quad \mathbb{P}(A | B) := \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} \quad ("A \text{ valószínűsége, feltéve } B").$$

Tehát az A esemény B -re vett feltételes valószínűségét úgy kapjuk, hogy az A és B közös részének valószínűségét a B esemény valószínűségéhez arányítjuk. A B esemény valószínűségének természetesen pozitívnak kell lennie, hiszen osztani szeretnénk vele. Az, hogy a fenti tört számlálójában az A -nak csak a B -be eső részét vesszük figyelembe, azon elvárásunkat tükrözi, hogy a B -n kívül eső kimeneteknek és így bármely B -től diszjunkt eseménynek a valószínűsége 0 legyen.



1.6. ábra. A feltételes valószínűség képzése

1.2.1. Példa. A fenti példa folytatásaként tekintsünk egy kockadobást, és tegyük fel, hogy a dobás páros. Határozzuk meg ebben annak valószínűségét, hogy kettést dobunk. Most is a szokásos $\Omega = \{1, 2, 3, 4, 5, 6\}$ eseményteret használjuk, \mathbb{P} pedig a klasszikus valószínűséget jelöli. A fenti definíció értelmében a kettés dobás valószínűsége az E feltétel mellett

$$\mathbb{P}(\{2\} | E) = \frac{\mathbb{P}(\{2\} \cap E)}{\mathbb{P}(E)} = \frac{\mathbb{P}(\{2\})}{\mathbb{P}(E)} = \frac{1/6}{1/2} = \frac{1}{3},$$

ami megfelel az előzetes várakozásainknak.

Legyen most A az az esemény, hogy 3-nál nagyobbat dobunk, ekkor az A valószínűsége az E feltétel mellett

$$\mathbb{P}(A | E) = \frac{\mathbb{P}(A \cap E)}{\mathbb{P}(E)} = \frac{\mathbb{P}(\{4, 5, 6\} \cap \{2, 4, 6\})}{\mathbb{P}(E)} = \frac{\mathbb{P}(\{4, 6\})}{\mathbb{P}(E)} = \frac{2/6}{1/2} = \frac{2}{3}.$$

Továbbá (az elvárásainknak megfelelően)

$$\mathbb{P}(E | E) = \frac{\mathbb{P}(E \cap E)}{\mathbb{P}(E)} = \frac{\mathbb{P}(E)}{\mathbb{P}(E)} = 1, \quad \mathbb{P}(\{1\} | E) = \frac{\mathbb{P}(\{1\} \cap E)}{\mathbb{P}(E)} = \frac{\mathbb{P}(\emptyset)}{\mathbb{P}(E)} = 0.$$

Figyeljük meg, hogy feltételes valószínűség minden esetben egy 0 és 1 közötti szám, hiszen $A \cap B \subset B$ miatt az 1.1.2. állításban szereplő (iv) tulajdonság szerint $\mathbb{P}(A \cap B) \leq \mathbb{P}(B)$ teljesül, és így az (1.6) formulában lévő tört értéke legfeljebb 1 lehet (és persze nemnegatív számok hányadosa nemnegatív). A fogalom elnevezése azonban messze nem csak emiatt jogos, ahogy az az alábbi állításból kiderül:

1.2.1. Állítás. Legyen $(\Omega, \mathcal{A}, \mathbb{P})$ egy valószínűségi mező, legyen továbbá $B \in \mathcal{A}$ egy esemény, melyre $\mathbb{P}(B) > 0$ teljesül. Definiáljuk a $\mathbb{P}_B : \mathcal{A} \rightarrow [0; 1]$ függvényt a $\mathbb{P}_B(A) = \mathbb{P}(A | B)$ formulával minden $A \in \mathcal{A}$ esetén. Ekkor \mathbb{P}_B egy valószínűségi mérték az Ω eseménytér eseményeinek \mathcal{A} halmazán (vagyis $(\Omega, \mathcal{A}, \mathbb{P}_B)$ egy valószínűségi mező).

A fenti állítás tehát azt mutatja, hogy a feltételes valószínűség segítségével egy valószínűségi mértékből újabbakat konstruálhatunk. Következik belőle, hogy mindaz, amit a valószínűségi mértékekre általában belátunk, igaz a feltételes valószínűségekre is (hiszen az utóbbi az előbbinek egy speciális esete). Például

$$\mathbb{P}(\bar{A} | B) = 1 - \mathbb{P}(A | B)$$

teljesül minden A és B eseményre, feltéve persze, hogy $\mathbb{P}(B) > 0$ (és így a feltételes valószínűség definiált).

Független események

A $\mathbb{P}(A | B)$ feltételes valószínűség és a $\mathbb{P}(A)$ valószínűség összehasonlítása képet ad arról, hogy hogyan befolyásolja a B bekövetkezése az esélyeket az A esemény bekövetkezésére. Egy fontos speciális eset az, amikor ez a valószínűség a B bekövetkezésétől függetlenül ugyanaz marad, tehát

$$\mathbb{P}(A) = \mathbb{P}(A | B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}.$$

A fenti egyenlőséget $\mathbb{P}(B)$ -vel beszorozva ekkor a

$$(1.7) \quad \mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$$

összefüggés adódik. Ez az, amit a függetlenség *definíciójához* felhasználunk:

1.2.2. Definíció. Legyenek A és B egy valószínűségi mező eseményei, ekkor A és B egymástól *függetlenek*, ha az (1.7) egyenlőség teljesül rájuk.

Megjegyezzük, hogy $\mathbb{P}(B) > 0$ esetén az (1.7) egyenletből is következik a $\mathbb{P}(A) = \mathbb{P}(A | B)$ összefüggés, ezek tehát ekkor ekvivalensek egymással. Sőt, ha $\mathbb{P}(A) > 0$ is teljesül, akkor hasonlóan látható, hogy mindkét egyenlet ekvivalens azzal, hogy $\mathbb{P}(B | A) = \mathbb{P}(B)$. Összefoglalva tehát, $\mathbb{P}(A), \mathbb{P}(B) > 0$ esetén

$$\mathbb{P}(A | B) = \mathbb{P}(A) \iff \mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B) \iff \mathbb{P}(B | A) = \mathbb{P}(B),$$

és mindhárom egyenlet éppen azt jelenti, hogy A és B függetlenek. Jól látszik a fenti ekvivalens összefüggésekből, hogy a fogalom szimmetrikus (vagyis A és B függetlensége ugyanazt jelenti, mint B és A függetlensége), persze ez már az (1.7) egyenletből is nyilvánvaló. Vegyük azonban észre, hogy míg a feltételes valószínűségekhez a feltételben szereplő eseménynek pozitív valószínűségűnek kellett lennie, addig a függetlenség definíciójában ezt nem követeljük meg, az (1.7) egyenlet mindkét oldala értelmes akkor is, ha az A és B események bármelyike 0 valószínűségű.

1.2.1. Gyakorlat. Mutassuk meg, hogy ha egy valószínűségi mező A és B eseményei közül legalább az egyik valószínűsége 0, akkor A és B függetlenek.

Az 1.2.2. definíció a matematikai modellünkben szereplő eseményekre vonatkozik. Persze a való életben is használjuk a függetlenség fogalmát, azt mondtuk például, hogy két kocka dobásánál az egyik dobás eredménye nem befolyásolja a másikat. Ugyanezt tapasztalhatjuk például érmedobások esetén is. Az eseményekre definiált függetlenség fogalma akkor használható, ha jól illeszkedik az imént felsorolt valós esetekben használt függetlenségfogalomhoz, vagyis ha a tapasztalati úton egymástól függetlennek ítélt jelenségek a modellben is azok a fenti definíció értelmében. Szerencsére ez a helyzet, és ezt most egy példán keresztül is illusztráljuk.

1.2.2. Példa. Tekintsük a következő kísérletet: kétszer egymás után feldobunk egy szabályos pénzérmét. Ha ezt sokszor egymás után végrehajtjuk, azt tapasztaljuk, hogy a második dobás nagyjából az esetek felében lesz fej vagy írás akkor is, ha az első dobásra fejet kaptunk, és akkor is, ha elsőre írás adódott. Azaz az első dobás eredménye semmilyen módon nem befolyásolja az esélyeket a második dobásnál.

Vizsgáljuk most meg ugyanezen eseményeket a matematikai modellünkben. A szokásos eseményterünk az $\Omega = \{F, I\} \times \{F, I\}$ szorzathalmaz, azaz a kimenteleink rendezett fej-írás

párok, és mindegyikük $1/4$ valószínűséggel következik be. Legyen A az az esemény, hogy elsőre fejet dobunk, B pedig az, hogy a második dobás fej, azaz

$$A = \{(F, F), (F, I)\}, \quad B = \{(F, F), (I, F)\}.$$

Tehát

$$\mathbb{P}(A \cap B) = \mathbb{P}(\{(F, F)\}) = \frac{1}{4} = \frac{1}{2} \cdot \frac{1}{2} = \mathbb{P}(A)\mathbb{P}(B),$$

így A és B függetlenek az 1.2.2. definíció értelmében (is). Mivel pozitív valószínűségű eseményekről van szó, ugyanezt a következő számolással is igazolhattuk volna:

$$\mathbb{P}(B | A) = \frac{\mathbb{P}(B \cap A)}{\mathbb{P}(A)} = \frac{\mathbb{P}(\{(F, F)\})}{\mathbb{P}(\{(F, F), (F, I)\})} = \frac{1/4}{1/2} = \frac{1}{2} = \mathbb{P}(B).$$

Továbbá annak a valószínűsége, hogy másodszorra írást dobunk, ha tudjuk, hogy elsőre fejet kaptunk

$$\mathbb{P}(\bar{B} | A) = 1 - \mathbb{P}(B | A) = \frac{1}{2} = \mathbb{P}(\bar{B}),$$

így A és \bar{B} is függetlenek. Hasonlóképp adódik ugyanez az \bar{A} és B , ill. az \bar{A} és \bar{B} eseményekre is. Vagyis az, hogy az A esemény bekövetkezik vagy sem, semmilyen hatással nincs a B bekövetkezésének esélyeire, ami persze teljes összhangban van a várakozásainkkal.

1.2.2. Gyakorlat. Mutassuk meg a fenti példában látottakhoz hasonlóan, hogy ha egy pénzérmével egymás után hatszor dobunk, akkor feltéve, hogy az első öt dobás eredménye fej lett, annak a valószínűsége, hogy hatodikra szintén fejet kapunk, változatlanul $1/2$.

Az 1.2.2. példában azt tapasztaltuk, hogy két független esemény esetén bármelyiknek (esetleg mindkettőnek) a komplementerét véve szintén független eseményeket kaptunk. A következő állítás azt mutatja, hogy ez nem egy véletlen egybeesés eredménye, hanem egy általánosan érvényes tény.

1.2.2. Állítás. *Tegyük fel, hogy egy valószínűségi mező A és B eseményei egymástól függetlenek. Ekkor \bar{A} és B és függetlenek. Következésképp az A és \bar{B} ill. \bar{A} és \bar{B} eseménypárok szintén függetlenek.*

Bizonyítás. Az A és B események függetlensége a $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$ összefüggés teljesülését jelenti. Ebből kiindulva szeretnénk belátni, hogy \bar{A} és B is függetlenek, azaz, hogy az előző egyenlet rájuk is teljesül. Induljunk ki az egyenlet jobb oldalából:

$$(1.8) \quad \mathbb{P}(\bar{A})\mathbb{P}(B) = (1 - \mathbb{P}(A))\mathbb{P}(B) = \mathbb{P}(B) - \mathbb{P}(A)\mathbb{P}(B) = \mathbb{P}(B) - \mathbb{P}(A \cap B),$$

ahol az A és B függetlenségét az utolsó egyenlőségnél használtuk ki. Korábban (például a szita-formula bizonyításában) már láttuk, hogy $B = (B \setminus A) \cup (A \cap B)$ a B halmaz egy egymást kizáró események uniójára való felbontása. Azt pedig már az eseményeken definiált műveletek tulajdonságainak vizsgálatánál említettük, hogy $B \setminus A = \bar{A} \cap B$, így a valószínűségi mérték tulajdonságai szerint $\mathbb{P}(B) = \mathbb{P}(\bar{A} \cap B) + \mathbb{P}(A \cap B)$ teljesül. Ezt átrendezve kapjuk, hogy az (1.8) egyenlőség jobb oldalának értéke $\mathbb{P}(\bar{A} \cap B)$, tehát az \bar{A} és B események függetlenek.

Végezetül, mivel a függetlenség szimmetrikus reláció, azaz ha A és B függetlenek, akkor B és A is azok, így a fent látottak szerint \bar{B} és A is függetlenek. Ebből pedig - ismét megcserélve az események szerepét, majd alkalmazva a bizonyított állítást - azt kapjuk, hogy \bar{A} és \bar{B} is független események. \square

Az eddigiek alapján az eseményének függetlenségének fogalma jól modellezi azt, ha bizonyos valós jelenségek nem befolyásolják egymást. Azonban, mivel a valószínűségszámításban a függetlenséget a valószínűségek szintjén definiáljuk, azaz ez pusztán az (1.7) egyenlőség teljesülését jelenti, előfordulhat, hogy az szemléletesen összefüggőnek tűnő eseményeike is érvényes.

1.2.3. Példa. Dobjunk egy szabályos dobókockával, és legyen A az az esemény, hogy a dobott szám kisebb, mint 5, P pedig az, hogy a dobott szám prím. Azaz

$$A = \{1, 2, 3, 4\}, \quad \mathbb{P}(A) = \frac{4}{6} = \frac{2}{3}, \quad P = \{2, 3, 5\}, \quad \mathbb{P}(P) = \frac{3}{6} = \frac{1}{2},$$

továbbá

$$\mathbb{P}(A \cap P) = \mathbb{P}(\{2, 3\}) = \frac{2}{6} = \frac{1}{3} = \frac{2}{3} \cdot \frac{1}{2} = \mathbb{P}(A)\mathbb{P}(P),$$

tehát A és P függetlenek.

Mindezek után felmerülhet a kérdés, hogy az általunk alkotott fogalom használható-e, ha kettőnél több esemény függetlenségéről szeretnénk beszélni. Ha esetleg ehhez nem elégséges az eszköztárunk, akkor hogyan lehetne azt kiterjeszteni, hogy több eseményre is működjön?

Legyenek A , B és C egy valószínűségi mező eseményei. A három esemény függetlenségét még nem definiáltuk, azt minden esetre elvárhatjuk, hogy ha a három esemény független egymástól, akkor közülük bármely kettő az legyen. Elégséges-e vajon ez ahhoz, hogy a három esemény ne befolyásolja egymást?

Az utóbbi mondat tartalmát ismét a szokásos példán keresztül világítjuk meg. Tegyük fel, hogy háromszor feldobunk egy pénzérmét. Legyen A az az esemény, hogy az első dobás fej, B az, hogy a második fej, C pedig az, hogy harmadszorra fejet dobunk. Azt várjuk, hogy mindegyik dobás a többi eredményétől függetlenül kb. az esetek $1/2$ részében lesz fej, így összességében az esetek $\frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{8}$ részében kapunk 3 darab fejet. Ezt valószínűségek segítségével a

$$(1.9) \quad \mathbb{P}(A \cap B \cap C) = \mathbb{P}(A)\mathbb{P}(B)\mathbb{P}(C)$$

egyenlet írja le. Vegyük észre, hogy ez a két esemény függetlenségét definiáló (1.7) egyenlet kézenfekvő általánosítása.

Az első kérdés tehát: elvárható-e általában az (1.9) egyenlet teljesülése, ha feltesszük, hogy az eseményeink páronként függetlenek (tehát hogy közülük bármely kettő független)? A válasz sajnos negatív, ahogy azt a következő példa mutatja.

1.2.4. Példa. Dobjunk fel két érmét, és legyen A az az esemény, hogy az első dobás fej, B az, hogy a második dobás fej, C pedig az, hogy a dobott fejek száma páros. Ekkor

$$A = \{(F, F), (F, I)\}, \quad B = \{(F, F), (I, F)\}, \quad C = \{(F, F), (I, I)\},$$

tehát $\mathbb{P}(A) = \mathbb{P}(B) = \mathbb{P}(C) = \frac{1}{2}$. Továbbá

$$\mathbb{P}(A \cap B) = \mathbb{P}(B \cap C) = \mathbb{P}(A \cap C) = \mathbb{P}(\{(F, F)\}) = \frac{1}{4},$$

és így a $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$, $\mathbb{P}(B \cap C) = \mathbb{P}(B)\mathbb{P}(C)$ és $\mathbb{P}(A \cap C) = \mathbb{P}(A)\mathbb{P}(C)$ egyenletek mindegyike teljesül, azaz az A , B és C események páronként függetlenek. Azonban

$$\mathbb{P}(A \cap B \cap C) = \mathbb{P}(\{(F, F)\}) = \frac{1}{4} \neq \mathbb{P}(A)\mathbb{P}(B)\mathbb{P}(C) = \frac{1}{8},$$

tehát (1.9) ebben az esetben nem teljesül.

1.2.3. Gyakorlat. (M) Mutassunk példát olyan A, B, C eseményekre, melyekre teljesül az (1.9) összefüggés, de nem igaz rájuk, hogy páronként függetlenek.

A fenti példa és gyakorlat tanulsága szerint a függetlenség fogalmának három eseményre való kiterjesztéséhez nem elegendő csupán a páronkénti függetlenséget vagy éppen az (1.9) egyenlet teljesülését elvárni. Valójában az a jó feltétel, ha ezek mindegyikét elvárjuk, és háromnál több esemény esetén is hasonlóképp kell eljárni:

1.2.3. Definíció. Legyenek A_1, \dots, A_n egy valószínűségi mező eseményei. Azt mondjuk, hogy A_1, \dots, A_n (együttesen) függetlenek, ha minden $\emptyset \neq I \subset \{1, \dots, n\}$ indexhalmaz esetén

$$(1.10) \quad \mathbb{P} \left(\bigcap_{i \in I} A_i \right) = \prod_{i \in I} \mathbb{P}(A_i)$$

teljesül, vagyis ha tetszőlegesen kiválasztunk belőlük néhány (de legalább egy) eseményt (azaz ezek indexeinek I halmazát), a kiválasztott események metszetének a valószínűsége ezen események valószínűségének szorzata.

Megjegyezzük, hogy a fenti definícióban (a tömörség kedvéért) megengedünk egy elemű I indexhalmazokat is, ekkor az egy elemű metszet definíció szerint magát a halmazt jelenti, míg az egy elemű szorzat értéke egyszerűen az a szám, amit a produktum jel után írunk (ez esetben a kiválasztott esemény valószínűsége). Az egy elemű indexhalmazokra a fenti egyenlőség nyilvánvalóan igaz bármely n eseményre, és így nem ad hozzá semmit a definíció tartalmához, ezen esetek megengedése pusztán a megfogalmazást egyszerűsíti. Kizárjuk viszont az üres indexhalmaz esetét (habár voltaképp némi további diszkusszió árán akár ezt is megengedhetnénk).

Az 1.2.2 állítás könnyedén általánosítható az együttes függetlenség esetére a következőképp. Ha az A_1, \dots, A_n események együttesen függetlenek, akkor ezek közül néhánynak a komplementerét véve (a többit változatlanul hagyva) az így kapott n esemény szintén együttesen független. Ennek igazolása nem tartalmaz lényeges új gondolatot az 1.2.2 állítás bizonyításához képest, a fellépő nehézségek pusztán technikaiak, így ezt a gondolatmenetet nem részletezzük.

A szakaszt egy példával zárjuk. Általában fáradságos munka lehet belátni események együttes függetlenségét, hiszen ehhez számos egyenlet teljesülését ellenőrizni kell. A problémák azonban sokszor itt is csak technikai jellegűek, az együttes függetlenség pedig számos esetben szemléletesen nyilvánvaló. Éppen ezért a következő példát leszámítva ennek részletezésével általában nem foglalkozunk.

1.2.5. Példa. Tegyük fel, hogy 10-szer dobunk egy szabályos érmével, legyen F_i az az esemény, hogy az i -edik dobás fej. Megmutatjuk, hogy az F_1, F_2, \dots, F_{10} események együttesen függetlenek. Először is, a 10 hosszú $F - I$ sorozatok száma, azaz az Ω eseménytér elemszáma 2^{10} . Továbbá, az F_i események elemszáma 2^9 , hiszen ezek olyan kimenetelekből állnak, amelyeknél az i -edik dobás eredménye rögzített (fej), a többi 9 dobás eredményére viszont egymástól függetlenül mindig 2 lehetséges választásunk van, ez tehát összesen 2^9 lehetőség. Vagyis

$$\mathbb{P}(F_i) = \frac{|F_i|}{|\Omega|} = \frac{2^9}{2^{10}} = \frac{1}{2}$$

minden $1 \leq i \leq 10$ esetén.

Rögzítsünk most az F_1, \dots, F_{10} események közül néhányat, jelölje ezek indexeinek halmazát I . Be kell látnunk az (1.10) egyenlet teljesülését. A jobb oldalon álló szorzat

$$\prod_{i \in I} \mathbb{P}(F_i) = \frac{1}{2^{|I|}}$$

azaz a kiválasztott események számának megfelelő hatványra kell emelni az események (azonos) valószínűségét. Határozzuk meg most az (1.10) bal oldalán álló metszet elemszámát. Ezt olyan kimenetek (10 hosszú fej-írás sorozatok) alkotják, melyekben $|I|$ darab rögzített helyen fej áll. A maradék $10 - |I|$ helyen egymástól függetlenül egyenként kétféle elem állhat, összesen tehát $2^{10-|I|}$ ilyen kimenetel van. Azaz

$$\mathbb{P}\left(\bigcap_{i \in I} F_i\right) = \frac{|\bigcap_{i \in I} F_i|}{|\Omega|} = \frac{2^{10-|I|}}{2^{10}} = \frac{1}{2^{|I|}} = \prod_{i \in I} \mathbb{P}(F_i),$$

tehát (1.10) teljesül egy tetszőleges $\emptyset \neq I \subset \{1, \dots, 10\}$ indexhalmaz esetén, és így az F_1, \dots, F_{10} események együttesen függetlenek.

Gyakorlatok, feladatok

1.2.4. Gyakorlat. Számoljuk ki annak a valószínűségét, hogy két kockával dobva mindkét érték páros, feltéve, hogy összegük legalább tíz.

1.2.5. Gyakorlat. Két szabályos kockával dobunk. Tekintsük a következő eseményeket:

$$A = \{\text{a dobott számok összege páros}\},$$

$$B = \{\text{a dobott számok különbségének abszolút értéke legalább három}\},$$

$$C = \{\text{a dobott számok szorzata legfeljebb 4}\}.$$

Állapítsuk meg, hogy függetlenek-e az A és B illetve az A és C eseménypárok. Mennyi $\mathbb{P}(A | B)$ és $\mathbb{P}(A | C)$ értéke?

1.2.6. Gyakorlat. Tegyük fel, hogy az A , B és C együttesen független események valószínűségei $\mathbb{P}(A) = 0,3$, $\mathbb{P}(B) = 0,4$, és $\mathbb{P}(C) = 0,8$. Számoljuk ki a következő események valószínűségeit:

- mindhárom fenti esemény bekövetkezik,
- legalább az egyik esemény bekövetkezik,
- egyik esemény sem következik be.

1.2.7. Feladat. Tegyük fel, hogy az A és a B események közül legalább az egyik mindig bekövetkezik. Ha $\mathbb{P}(A | B) = 0,2$ és $\mathbb{P}(B | A) = 0,5$, akkor mennyi $\mathbb{P}(A)$, $\mathbb{P}(B)$ illetve $\mathbb{P}(A | \bar{B})$? Független-e A és B ?

1.2.2. A feltételes valószínűség alkalmazásai

Számos esetben előfordul, hogy a rendelkezésre álló információk egy valószínűségi modellben feltételes valószínűségként interpretálhatók, ezért érdemes megvizsgálni, hogy ezek segítségével hogyan számolhatók ki más események valószínűségei. Ugyancsak gyakori helyzet, hogy egyes események valószínűségének kiszámolása bizonyos feltételek mellett lényegesen leegyszerűsödik, ilyenkor érdemes a számolást esetekre bontani. Az alábbiakban ezeket a szituációkat vizsgáljuk meg közelebbről, és bemutatunk három példát a feltételes valószínűség alkalmazására.

A szorzási szabály

Egy urnában 5 piros és 5 fehér golyó van. Két golyót húzva az urnából visszatevés nélkül, mi a valószínűsége, hogy elsőre piros, másodikra pedig fehér golyót húzunk? Természetesen egyszerű megszámlálni a jó húzásokat, ezekből éppen $5 \cdot 5 = 25$ van, hiszen ennyiféleképp alkothatunk olyan rendezett párokat, ahol az első golyó a pirosak közül kerül ki, a második pedig a fehérek közül. Összesen $10 \cdot 9 = 90$ különböző lehetőségünk van a húzásra, tehát a keresett valószínűség $\frac{25}{90} = \frac{5}{18}$.

Ugyanerre az eredményre más gondolatmenet is elvezet minket. Piros-fehér húzáshoz először természetesen pirosat kell húznunk, ennek a valószínűsége $\frac{5}{10}$. Ha viszont elsőre pirosat húzunk, akkor az urnában 9 golyó marad, amiből 5 darab fehér. Tehát ezen feltétel mellett $\frac{5}{9}$ valószínűséggel húzunk fehéret másodszorra, azaz a valószínűség $\frac{5}{10} \cdot \frac{5}{9} = \frac{5}{18}$.

Írjuk le a fenti példát formálisan. Legyen P_1 az az esemény, hogy elsőre pirosat húzunk, F_2 pedig az, hogy másodszorra fehéret. A kérdéses valószínűség ekkor a $P_1 \cap F_2$ esemény valószínűsége, és a második számolásunkat a $\mathbb{P}(P_1 \cap F_2) = \mathbb{P}(P_1)\mathbb{P}(F_2 | P_1)$ egyenlet írja le.

Vegyük észre, hogy a fenti egyenlet valójában a feltételes valószínűség

$$\mathbb{P}(B | A) = \frac{\mathbb{P}(B \cap A)}{\mathbb{P}(A)}$$

definíciójának átrendezése az $A = P_1$ és $B = F_2$ eseményekre, vagyis a

$$\boxed{\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B | A)}$$

formula általában is érvényes, amennyiben a feltételes valószínűség definiált (azaz $\mathbb{P}(A) > 0$). Ezt az összefüggést *szorzási szabálynak* nevezzük.

A fenti példánkban tulajdonképpen ennek a szabálynak az alkalmazása nem volt létfontosságú, hiszen más érveléssel is egyszerűen kijött az eredmény. Vannak viszont olyan szituációk, amikor egyes események bekövetkezése lényegesen megváltoztatja a körülményeket, és így a feltételes valószínűséggel való számolás jóval kényelmesebbé válik.

Változtassunk most egy kicsit a fenti példán. Tegyük fel, hogy az urnában 5 piros és 5 fehér golyó van, és kétszer húzunk. Azonban az első húzás után a húzott golyót visszatesszük az urnába, és beteszünk még egy, a húzott golyó színével megegyező színű golyót. Mi a valószínűsége most, hogy elsőre pirosat, másodszorra pedig fehéret húzunk?

Mivel ebben az esetben az első húzás színe szerint máshogy alakulnak a körülmények a második húzásnál, ezért ezúttal a szorzási szabály alkalmazása a célravezető út, hiszen azt könnyen ki tudjuk számolni, hogy milyen valószínűséggel húzunk fehéret, ha elsőre pirosat húztunk. Ekkor ugyanis 11 golyó lesz az urnában, melyből 5 fehér, így (a fenti jelöléseket megtartva) $\mathbb{P}(F_2 | P_1) = \frac{5}{11}$, azaz

$$\mathbb{P}(P_1 \cap F_2) = \mathbb{P}(P_1)\mathbb{P}(F_2 | P_1) = \frac{5}{10} \cdot \frac{5}{11} = \frac{5}{22}.$$

A szituáció tovább bonyolódik, ha több eseményünk is van, melyek közül egyesek bekövetkezése befolyásolja néhány másik bekövetkezési esélyeit. Ha az események együttes bekövetkezésének esélyét (azaz a metszetük valószínűségét) szeretnénk meghatározni, akkor a következő formulát használhatjuk:

1.2.3. Állítás (Szorzási szabály). *Legyenek A_1, \dots, A_n egy valószínűségi mező eseményei, ekkor*

$$(1.11) \quad \mathbb{P}(A_1 \cap \dots \cap A_n) = \mathbb{P}(A_1)\mathbb{P}(A_2 | A_1)\mathbb{P}(A_3 | A_1 \cap A_2) \cdots \mathbb{P}(A_n | A_1 \cap \dots \cap A_{n-1})$$

érvényes, amennyiben a fenti feltételes valószínűségek definiáltak.

A fenti állítás $n = 2$ esetén éppen a már korábban látott formulát adja. Az állítás bizonyítása könnyen adódik a feltételes valószínűségek definíciójának behelyettesítésével. Ezt most az $n = 3$ esetben részletezzük (de általában sem nehezebb):

$$\mathbb{P}(A_1)\mathbb{P}(A_2 | A_1)\mathbb{P}(A_3 | A_1 \cap A_2) = \mathbb{P}(A_1) \cdot \frac{\mathbb{P}(A_2 \cap A_1)}{\mathbb{P}(A_1)} \cdot \frac{\mathbb{P}(A_3 \cap (A_1 \cap A_2))}{\mathbb{P}(A_1 \cap A_2)} = \mathbb{P}(A_1 \cap A_2 \cap A_3)$$

a metszet kommutativitása és asszociativitása miatt.

Vegyük észre, hogy ha az A_1, \dots, A_n események (együttesen) függetlenek volnának, akkor az (1.11) egyenlőség jobb oldalán a feltételeket elhagyhatnánk, és így éppen a függetlenség egyik definiáló formuláját kapnánk. Az általános eset kezelése azonban éppen az (1.11) formula segítségével történik.

1.2.6. Példa. Módosítsuk most ismét egy kicsit a fenti példánkat. Tegyük fel, hogy egy urnában 5 piros és 5 fehér golyó van, és 3 golyót húzunk egymás után. Az első két húzás után azonban visszatesszük a húzott golyót, továbbá még egy, a húzott golyó színével megegyező színű golyót is. Mi a valószínűsége, hogy három piros golyót húzunk?

Legyenek P_1, P_2 ill. P_3 azok az események, hogy elsőre, másodikra ill. harmadikra piros golyót húztunk. A kérdés tehát a $P_1 \cap P_2 \cap P_3$ esemény valószínűsége. A szorzási szabály szerint

$$\mathbb{P}(P_1 \cap P_2 \cap P_3) = \mathbb{P}(P_1)\mathbb{P}(P_2 | P_1)\mathbb{P}(P_3 | P_1 \cap P_2).$$

Először 10 golyóból 5 darab piros van az urnában, tehát $\mathbb{P}(P_1) = \frac{5}{10}$. Ha elsőre pirosat húztunk, akkor a második húzásnál 11 golyó lesz az urnában, amelyekből 6 lesz piros. Tehát ebben az esetben $\frac{6}{11}$ eséllyel húzunk pirosat másodszorra, ez tehát a $\mathbb{P}(P_2 | P_1)$ valószínűség. Végül, ha az első két húzásunk piros (azaz a $P_1 \cap P_2$ esemény bekövetkezik), akkor az urnában immár 12 golyó lesz a harmadik húzás előtt, közülük pedig 7 lesz piros. Tehát $\mathbb{P}(P_3 | P_1 \cap P_2) = \frac{7}{12}$, vagyis a keresett valószínűség

$$\mathbb{P}(P_1 \cap P_2 \cap P_3) = \frac{5}{10} \cdot \frac{6}{11} \cdot \frac{7}{12} = \frac{7}{44}.$$

1.2.8. Gyakorlat. Egy matematikusnak 10 pár különböző színű zoknija van. Három egymás utáni napon is késve indul el otthonról, ami azt jelenti, hogy véletlenszerűen választ a (párosítatlan) zoknik közül. Mi a valószínűsége, hogy mindhárom napon sikerül azonos színű zoknikat vennie, feltéve, hogy mindig tisztát vesz fel, és a három nap alatt nem volt zoknimosás?

A teljes valószínűség tétele

Tekintsük ismét az előző szakasz példáiban látott szituációt. Kétszer húzunk egy urnából, amelyben 5 piros és 5 fehér golyó van, továbbá az első húzás után a kihúzott golyóval együtt visszarakunk még egy ugyanolyan színű golyót. Mi a valószínűsége, hogy másodszorra piros golyót húzunk?

Itt abba a nehézségbe ütközünk, hogy második húzásnál az esélyek függenek az első húzás eredményétől, ami viszont az előző szakasszal ellentétben nem ismert. Az első húzás eredményét is fixálva persze már könnyebb dolgunk van. Korábban kiszámoltuk ezt a valószínűséget akkor, ha az első húzás eredménye is piros, és hasonlóképp számolható a valószínűség a másik esetben. Összesen tehát két különböző esetet különböztethetünk meg.

Hogyan kaphatjuk meg a kérdéses valószínűséget e két esetből? Vegyük észre, hogy a két esetünket egymást kizáró események határozzák meg: az első húzás vagy piros, vagy

fehér, de egyszerre csak az egyik teljesülhet. Az is igaz viszont, hogy valamelyik mindenképp teljesül. Ha tehát P_i jelöli azt, hogy az i -edik húzás piros, F_i pedig azt, hogy fehér, akkor F_1 vagy P_1 legalább egyike mindig bekövetkezik, de egyszerre csak az egyik következhet be. Következésképp P_2 bekövetkezése a $P_1 \cap P_2$ vagy az $F_1 \cap P_2$ események bekövetkezését jelenti:

$$P_2 = (P_1 \cap P_2) \cup (F_1 \cap P_2).$$

Itt persze ismét két egymást kizáró esemény uniója szerepel, így valószínűségi mérték tulajdonságai ill. az előző szakasz eredményei alapján

$$\begin{aligned} \mathbb{P}(P_2) &= \mathbb{P}(P_1 \cap P_2) + \mathbb{P}(F_1 \cap P_2) \\ &= \mathbb{P}(P_2 | P_1)\mathbb{P}(P_1) + \mathbb{P}(P_2 | F_1)\mathbb{P}(F_1) = \frac{6}{11} \cdot \frac{1}{2} + \frac{5}{11} \cdot \frac{1}{2} = \frac{1}{2}. \end{aligned}$$

A példában látott gondolatmenetet most általánosítani fogjuk. A fenti megoldás kulcsa az volt, hogy *esetszétválasztással* visszavezettük a problémát könnyebben kezelhető szituációkra. Az esetszétválasztásnál olyan eseteket tekintettünk, amelyek egyike biztosan bekövetkezik, de közülük egyszerre csak egy következhet be, ez pedig lehetővé tette a kérdéses valószínűség összegekre való bontását. Az ilyen esetszétválasztást eredményező eseményeknek külön nevet adunk:

1.2.4. Definíció. Legyenek $A_1, \dots, A_n \subset \Omega$ egy eseménytér páronként egymást kizáró eseményei, melyekre $\Omega = A_1 \cup \dots \cup A_n$ teljesül. Ekkor az A_1, \dots, A_n eseményekből álló rendszert *teljes eseményrendszernek* nevezzük.

A teljes eseményrendszer tehát nem más, mint az Ω eseménytér egy páronként diszjunkt eseményekre való felbontása. Egy ilyen teljes eseményrendszer szerinti esetszétválasztás lehetővé teszi a fenti gondolatmenet végigvitelét, amely a következő eredményt adja:

1.2.4. Tétel. (*A teljes valószínűség tétele*) Legyenek A_1, \dots, A_n és B egy valószínűségi mező eseményei. Tegyük fel, hogy A_1, \dots, A_n teljes eseményrendszert alkot, és $\mathbb{P}(A_i) > 0$ teljesül minden $1 \leq i \leq n$ esetén. Ekkor

$$\mathbb{P}(B) = \mathbb{P}(B | A_1)\mathbb{P}(A_1) + \dots + \mathbb{P}(B | A_n)\mathbb{P}(A_n).$$

Bizonyítás. Mivel egy teljes eseményrendszer eseményei páronként egymást kizáróak, így a $B \cap A_1, \dots, B \cap A_n$ események is szükségképp azok, hiszen közülük bármely kettő különböző esemény egy közös eleme egyben két különböző A_i és A_j eseménynek is eleme volna, ami lehetetlen.

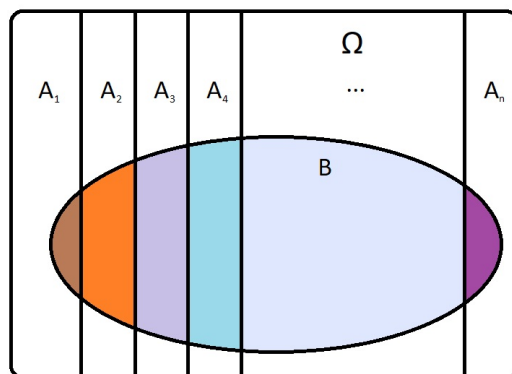
Továbbá, mivel Ω minden egyes eleme benne van valamelyik A_i eseményben, így B összes eleme egyben valamelyik $B \cap A_i$ eseménynek is eleme, azaz

$$B = (B \cap A_1) \cup \dots \cup (B \cap A_n)$$

a B esemény egy páronként diszjunkt eseményekre való felbontását adja (1.7. ábra). Ekkor a valószínűségi mérték additivitása miatt

$$\mathbb{P}(B) = \mathbb{P}(B \cap A_1) + \dots + \mathbb{P}(B \cap A_n).$$

Végül a szorzási szabály miatt $\mathbb{P}(B \cap A_i) = \mathbb{P}(B | A_i)\mathbb{P}(A_i)$ teljesül minden $1 \leq i \leq n$ esetén, ezt a fenti egyenlőség jobb oldalán minden tagnál behelyettesítve adódik az állítás. \square



1.7. ábra. $B = (B \cap A_1) \cup \dots \cup (B \cap A_n)$

1.2.7. Példa. Tegyük fel, hogy egy vírusfertőzés kimutatására készített teszt 95% valószínűséggel mutatja ki a fertőzöttséget egy beteg tesztalany esetén. Azonban a teszt időnként fals pozitív eredményt is produkál, azaz nem fertőzött személyeket tesztelve 0,1% valószínűséggel szintén pozitív eredményt kapunk. Ha tudjuk, hogy a lakosságnak 0,01%-a fertőzött, akkor véletlenszerűen kiválasztva és letesztelve valakit mi a valószínűsége, hogy pozitív lesz az eredmény?

Az a két esemény, hogy a véletlenszerűen választott tesztalany fertőzött vagy sem, egy teljes eseményrendszert alkot, hiszen egyszerre csak az egyik állapot állhat fent, de valamelyik biztosan fennáll. Jelölje F azt, hogy a tesztalany fertőzött, ekkor \bar{F} az az esemény, hogy az alany egészséges. Alkalmazható tehát a teljes valószínűség tétele az F és \bar{F} teljes eseményrendszerre és arra eseményre, hogy a teszt eredménye pozitív. Jelölje ez utóbbi eseményt POZ , ekkor a tétel szerint

$$\mathbb{P}(POZ) = \mathbb{P}(POZ | F)\mathbb{P}(F) + \mathbb{P}(POZ | \bar{F})\mathbb{P}(\bar{F}).$$

A fenti egyenlőség most azért előnyös számunkra, mert éppen a jobb oldalon álló valószínűségek állnak rendelkezésünkre: egy beteg páciens esetén 95% eséllyel mutatja ki a fertőzést a teszt, azaz $\mathbb{P}(POZ | F) = 0,95$. Hasonlóképpen tudjuk, hogy $\mathbb{P}(POZ | \bar{F}) = 0,001$. Mivel továbbá a lakosság 0,01%-a fertőzött, így egy véletlenszerűen választott személy fertőzöttségének valószínűsége $\mathbb{P}(F) = 0,0001$, és így következképp $\mathbb{P}(\bar{F}) = 1 - \mathbb{P}(F) = 0,9999$. Tehát behelyettesítve

$$\mathbb{P}(POZ) = 0,95 \cdot 0,0001 + 0,001 \cdot 0,9999 = 0,0010949.$$

1.2.9. Feladat. Az eddig rendelkezésünkre álló eszköztár segítségével elemezzük a jegyzet bevezetőjében bemutatott Monty Hall-paradoxont. Melyik stratégiával mennyi esélyünk van a nyeresésre?

A Bayes-tétel

Az 1.2.7. példa gondolatmenetét folytatva próbáljuk meg megítélni, hogy a példában látott teszt mennyire jó. A megadott adatok azt sugallják, hogy a teszt hatékony, hiszen nagy eséllyel kimutatja a betegséget egy fertőzött személynél, és csak nagyon kis valószínűséggel ítélt betegnek egy egészséges embert.

Közelítsük meg egy kicsit máshogy ezt a kérdést. Tegyük fel, hogy egy véletlenszerűen választott személyen elvégzett teszt pozitív eredményt adott. Mennyire lehetünk bizonyosak

abban, hogy az illető valóban beteg? Az 1.2.7. példa jelöléseit használva tehát a $\mathbb{P}(F \mid POZ)$ feltételes valószínűsége vagyunk kíváncsiak. Vegyük észre, hogy a rendelkezésünkre álló $\mathbb{P}(POZ \mid F)$ feltételes valószínűségben éppen fel van cserélve a fenti két esemény. Szerencsére egy egyszerű összefüggés alapján az egyikből megkapható a másik, éppen ezt írja le a következő állítás:

1.2.5. Tétel ((Egyszerű) Bayes-tétel). *Legyenek A és B egy valószínűségi mező pozitív valószínűségű eseményei. Ekkor*

$$\mathbb{P}(A \mid B) = \frac{\mathbb{P}(B \mid A)\mathbb{P}(A)}{\mathbb{P}(B)}.$$

Bizonyítás. Az állítás egyszerűen adódik a feltételes valószínűség definíciója alapján:

$$\frac{\mathbb{P}(B \mid A)\mathbb{P}(A)}{\mathbb{P}(B)} = \frac{\frac{\mathbb{P}(B \cap A)}{\mathbb{P}(A)} \cdot \mathbb{P}(A)}{\mathbb{P}(B)} = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} = \mathbb{P}(A \mid B).$$

□

A tétel alapján ill. az előző szakasz eredményeit felhasználva most már megválaszolhatjuk, hogy a fenti példánkban egy pozitív teszt esetén mennyire lehetünk biztosak abban, hogy a tesztelt személy fertőzött:

$$\mathbb{P}(F \mid POZ) = \frac{\mathbb{P}(POZ \mid F)\mathbb{P}(F)}{\mathbb{P}(POZ)} = \frac{0,95 \cdot 0,0001}{0,0010949} \approx 0,0868.$$

Vagy úgy is fogalmazhatunk, hogy egy pozitív teszt ellenére még mindig 91,32% valószínűséggel egészséges a tesztelt alany. Mindezek alapján talán már nem is tűnik annyira hatékonynak a tesztünk.

Fontos leszögezni, hogy a fenti példában kapott eredmény akkor érvényes, ha valóban véletlenszerűen választottuk a személyt, akin a tesztet elvégezzük. Nyilván egyéb információ birtokában (például, ha a tesztelt személynek a vírusfertőzésre jellemző tünetei vannak) annak fgyelembevételével ez a valószínűség jelentősen módosulhat.

A fenti eredmény egyszerűen adódott a Bayes-tételből, de a munka egy részét voltaképpen már az előző szakaszban elvégeztük, amikor a $\mathbb{P}(POZ)$ valószínűség kiszámításához a teljes valószínűség tételét használtuk. Ezt a két lépést gyakran nem választjuk szét, éppen ezért megfogalmazzuk a Bayes-tételnek a teljes valószínűség tételével kombinált általános alakját is:

1.2.6. Tétel (Bayes-tétel). *Legyenek A_1, \dots, A_n és B egy valószínűségi mező pozitív valószínűségű eseményei. Tegyük fel, hogy az A_1, \dots, A_n események teljes eseményrendszert alkotnak, ekkor*

$$\mathbb{P}(A_1 \mid B) = \frac{\mathbb{P}(B \mid A_1)\mathbb{P}(A_1)}{\sum_{i=1}^n \mathbb{P}(B \mid A_i)\mathbb{P}(A_i)}.$$

Bizonyítás. Alkalmazzuk a $\mathbb{P}(A_1 \mid B)$ feltételes valószínűsége az egyszerű Bayes-tételt, majd a nevezőben a $\mathbb{P}(B)$ valószínűsége a teljes valószínűség tételét. □

A Bayes-tétel utóbbi alakja jobban megvilágítja az okokat a fenti példánk némileg meglepő eredménye mögött. Alkalmazzuk tehát a fenti formulát a $\mathbb{P}(F | POZ)$ valószínűsége, ahol a korábban a nevezőben lévő $\mathbb{P}(POZ)$ valószínűséget összeg alakban írjuk fel:

$$\begin{aligned}\mathbb{P}(F | POZ) &= \frac{\mathbb{P}(POZ | F)\mathbb{P}(F)}{\mathbb{P}(POZ)} = \frac{\mathbb{P}(POZ | F)\mathbb{P}(F)}{\mathbb{P}(POZ | F)\mathbb{P}(F) + \mathbb{P}(POZ | \bar{F})\mathbb{P}(\bar{F})} \\ &= \frac{0,95 \cdot 0,0001}{0,95 \cdot 0,0001 + 0,001 \cdot 0,9999} = \frac{0,000095}{0,000095 + 0,0009999}.\end{aligned}$$

Az utóbbi tört számlálójában a nevezőben lévő összeg egyik tagja áll, mégpedig a kisebbik, ami valójában jóval kisebb a másiknál, ezért lesz a tört értéke kicsi. Ezt a jelenséget pedig három tényező együttesen okozza. Egyrészt nagyon kevés a populációban a fertőzött, következésképp nagyon magas az nem fertőzött emberek aránya. Ilyen arányok mellett a fals pozitív 0,1%-os valószínűsége igazándiból nem elég jó, mert még így is viszonylag nagy eséllyel adódik ilyen hamis eredmény. Továbbá, hiába mutatja ki a teszt a betegséget a fertőzötteknél 95% valószínűségével, ez nem elegendő ahhoz, hogy a fenti tényezőket ellensúlyozza.

Gyakorlatok, feladatok

1.2.10. Gyakorlat. Feldobunk egy szabályos kockát, majd egy szabályos érmét annyiszor, amennyit a kocka mutat.

- Mennyi a valószínűsége, hogy csak írást dobunk?
- Feltéve, hogy csak írást dobunk, mi a valószínűsége, hogy a kockával hatost dobtunk?

1.2.11. Gyakorlat. (M) Feldobunk két szabályos dobókockát és ha k darab hatos van a dobások közt, akkor k piros és $2 - k$ sárga golyót teszünk egy (kezdetben üres) dobozba. Ezután kétszer húzunk visszatevéssel: mindkét húzásra piros golyót húzunk. Mit tippelnénk k értékére? Mekkora esélyünk van eltalálni?

1.2.12. Gyakorlat. Egy gépjárműveket biztosító társaság az ügyfeleit három osztályba sorolja: jó sofőr, átlagos sofőr, rossz sofőr. A társaság tapasztalata alapján a jó, átlagos és rossz sofőrök 0,05, 0,15, illetve 0,3 eséllyel lesznek baleset részesei egy év alatt. Hogyha az ügyfelek 20%-a jó sofőr, 50%-a átlagos sofőr és 30%-a rossz sofőr, akkor egy véletlenszerűen választott ügyfél milyen eséllyel lesz baleset részese a jövő év folyamán? Ha tudjuk, hogy egy adott ügyfélnek nem volt tavaly balesete, milyen valószínűséggel jó, átlagos illetve rossz sofőr?

1.2.13. Gyakorlat. Egy vizsgakérdésben három lehetséges válaszból kell kiválasztani az egyetlen helyeset. Egy hipotetikus hallgató p valószínűséggel tudja a helyes választ, míg ha nem tudja tippel (egyenlő eséllyel választva a három válasz közül). Feltéve, hogy a hallgató helyesen válaszolt, mi a valószínűsége, hogy tudta is a választ? Mennyi ez a valószínűség $p = \frac{1}{4}$ esetén?

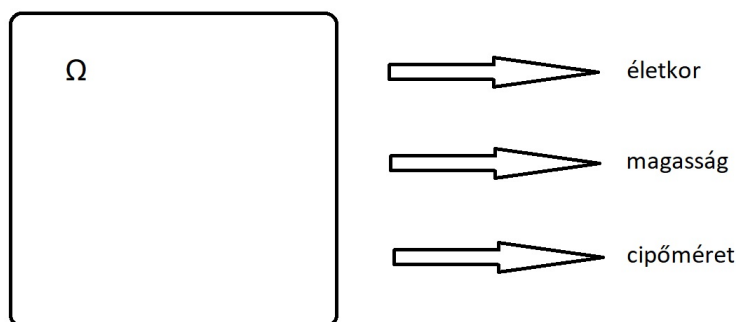
2. fejezet

Diszkrét valószínűségi változók

Az előző fejezetben felépítettük a valószínűségi mezők modelljét, ezek azonban céljainkhoz még önmagukban nem elegendők. A valószínűségszámítás alkalmazásaiban tipikusan véletlen mennyiségekkel dolgozunk. Voltaképp egy kockadobás eredményére is számként gondolunk, de a gyakorlati alkalmazásokban számtalan további példára lelhetünk. Véletlen mennyiség lehet egy csoportból véletlenszerűen választott ember életkora, egy termék tesztelésénél a kiválasztott mintában lévő hibás termékek darabszáma, egy olyan program futásideje, amelynek futása a véletlentől függ, egy csatornán való információküldésnél a hibásan továbbított bitek száma, egy mérés hibája, egy részvény árfolyama, stb.

Az eddigiek alapján tekinthetnénk véletlen kimeneteleknek e számértékeket, az eseményterünket definiálhatnánk ezek összességéeként. Könnyen meggyőzhetjük magunkat arról, hogy ez nem a legcélszerűbb eljárás. Tekintsük a fenti példák közül egy véletlenszerűen választott ember életkorát. Bár ez egy véletlen érték, itt voltaképpen nem egy számot választunk véletlenszerűen, hanem egy embercsoport egy tagját. Az életkor a választott ember egy *jellemzője*. Természetesen emellett más jellemzők is érdekelhetnek minket. Beszélhetünk a választott ember cipőméretéről, magasságáról, vizsgálhatjuk, hogy a kettő között van-e összefüggés.

Mivel a felsorolt attribútumok egyazon személyhez tartoznak, nem célszerű ezeket egymástól teljesen függetlenül, külön-külön eseményterekben kezelni. Ehelyett a véletlenszerűen választott személyhez kapcsoljuk őket, azaz voltaképpen egyetlen eseménytér elemeihez rendelünk hozzá különböző értékeket, tulajdonságokat, melyeket aztán együtt kezelhetünk. Ezt a célt ún. *valószínűségi változók* segítségével valósíthatjuk meg, melyek a fentiek értelmében hozzárendelések, vagy másképp szólva az eseménytéren értelmezett *függvények*.



2.1. ábra. A valószínűségi változók az eseménytér kimeneteleihez különböző jellemző értékeket hozzárendelő függvények

Ez és a következő fejezet a valószínűségi változók és a hozzájuk kapcsolódó legfontosabb fogalmak alapvető tulajdonságait ill. ezek néhány egyszerű alkalmazását tárgyalja. Ebben a fejezetben kizárólag olyan esetekre szorítkozunk, amikor az eseménytér és így a valószínűségi változók értékkészlete is véges vagy megszámlálhatóan végtelen. Ahogy már korábban is, ez a feltétel a technikai problémák elkerülésének és az alkalmazott eszköztár minimalizálásának lehetőségét teremti meg számunkra, hogy a megfelelő szemlélet kialakítására koncentrálhassunk. Az alapfogalmak bemutatása után azt tárgyaljuk, hogy hogyan kezelhetünk egyszerre több változót, és hogyan fogalmazhatjuk meg azt a valószínűségszámítás nyelvén, hogy a változók értéke nem befolyásolja egymást. Ezután a várható érték és a szórás fogalmát mutatjuk be, majd eszköztárunkat a generátorfüggvényekkel bővítjük. Végül részletesen tárgyalunk néhány fontos nevezetes valószínűségeloszlást, egyes esetekben pedig az eloszlások közötti kapcsolatokról is szót ejtünk.

2.1. Diszkrét valószínűségi változók eloszlása

Az egész fejezetben minden eseménytérrel feltesszük, hogy megszámlálható (véges vagy megszámlálhatóan végtelen). Továbbá, az első fejezetben látottakhoz hasonlóan feltesszük, hogy egy eseménytér összes részhalmeza eseményt alkot.

2.1.1. Definíció. Legyen Ω egy megszámlálható eseménytér, ekkor egy $X : \Omega \rightarrow \mathbb{R}$ függvényt *diszkrét valószínűségi változónak* nevezünk. Az X valószínűségi változó *egyszerű*, ha $\text{ran } X$ (azaz az X értékkészlete) véges. A valószínűségi változókat (ebben a jegyzetben) tipikusan nagy latin betűkkel jelöljük.

Megjegyzés. Az Ω eseménytér megszámlálható volta valójában nem szükséges a diszkrét valószínűségi változó definíciójában, elegendő volna pusztán azt feltenni, hogy az értékkészlet megszámlálható. A plusz feltételezéssel itt az egyszerűség kedvéért élünk, az általánosabb definícióhoz szükséges technikai részletekkel a következő fejezetben foglalkozunk majd.

2.1.1. Példa. Az alábbiak diszkrét valószínűségi változók:

- Egy szabályos kockával dobunk, legyen X maga az eredmény. Itt $\Omega = \{1, 2, 3, 4, 5, 6\}$, és minden $\omega \in \Omega$ esetén $X(\omega) = \omega$.
- Legyen $Y = X^2$, ahol X a fent definiált valószínűségi változó. Ekkor tehát Y egy kockadobás eredményének a négyzete.
- Feldobunk egy szabályos pénzérmét kétszer egymás után, legyen Z a fejek száma. Tehát $\Omega = \{FF, FI, IF, II\}$, továbbá $Z(FF) = 2$, $Z(FI) = Z(IF) = 1$, valamint $Z(II) = 0$.
- Legyen $A \subset \Omega$ egy esemény, legyen továbbá

$$\mathbb{1}_A(\omega) = \begin{cases} 1, & \text{ha } \omega \in A, \\ 0, & \text{ha } \omega \notin A. \end{cases}$$

Az $\mathbb{1}_A$ valószínűségi változót az A esemény *indikátorváltozójának* nevezzük.

- Egymástól függetlenül elvégzünk egy kísérletet n -szer egymás után. Legyen U egy adott (fix) esemény bekövetkezéseinek száma. Például: n -szer dobunk egy érmével, és megszámláljuk, hogy hányszor következett be az az esemény, hogy fejet kapunk. Vagy n -szer dobunk egy kockával, és megszámláljuk, hogy hányszor dobtunk páros számot. Ez tehát a fenti Z változó általánosítása.

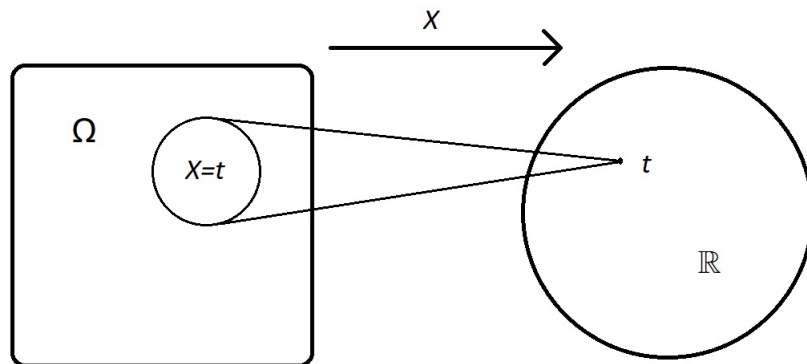
- Addig végzünk egymás után többször függetlenül egy kísérletet, amíg egy adott esemény be nem következik. Legyen V a szükséges kísérletek száma.

Vegyük észre, hogy a fent definiált X, Y, Z, U és $\mathbb{1}_A$ valószínűségi változók egyszerűek (azaz az értékészletük véges), míg $\text{ran } V = \mathbb{N}^+$ a pozitív egészek halmaza.

A valószínűségi változók viselkedését az általuk felvett értékek segítségével jellemezhetjük. Beszélhetünk például arról, hogy milyen valószínűséggel lesz egy dobás értéke 3-nál nagyobb, vagy hogy egy üzenet küldésénél mi a valószínűsége, hogy legfeljebb 10 bit hibásodik meg.

Ezt a problémát egységesen fogjuk kezelni minden diszkrét valószínűségi változóra. Ebben az esetben a fentiekhez hasonló kérdések megválaszolásához elegendő tudni, hogy a lehetséges értékeket külön-külön milyen valószínűséggel veszi fel egy változó. Ehhez nagyon hasonló jelenséggel már találkoztunk az 1.1.2. állításban, ahol azt láttuk, hogy egy megszámlálható eseménytérben egy valószínűségi mértéket egyértelműen meghatároz az, hogy az egyes kimenetek milyen valószínűek.

2.1.2. Definíció. Legyen $X : \Omega \rightarrow \mathbb{R}$ egy diszkrét valószínűségi változó, $t \in \mathbb{R}$, ekkor az $\{X = t\}$ halmaz az Ω eseménytér mindazon ω elemeiből álló részhalmaza (vagyis azon ω -k alkotta esemény), melyekre $X(\omega) = t$ teljesül.



2.2. ábra. Az $\{X = t\}$ esemény

2.1.2. Példa. Kétszer dobunk egy szabályos érmével, legyen Z a fejek száma. Ekkor

$$\{Z = 1\} = \{FI, IF\}.$$

2.1.3. Definíció. A fenti definícióhoz hasonlóan, egy $X : \Omega \rightarrow \mathbb{R}$ diszkrét valószínűségi változó és egy $t \in \mathbb{R}$ valós szám esetén az $\{X \leq t\}$ esemény azon $\omega \in \Omega$ kimenetelekből áll, melyekre $X(\omega) \leq t$ teljesül. Analóg módon definiálhatjuk az $\{X < t\}$, $\{X \geq t\}$, $\{X > t\}$, $\{s < X \leq t\}$, stb. eseményeket is. Általában, ha $H \subset \mathbb{R}$ egy tetszőleges halmaz, akkor a H halmaz X általi *ősképe* az $X^{-1}(H) = \{X \in H\}$ halmaz, amely azon $\omega \in \Omega$ kimenetelekből áll, melyekre $X(\omega) \in H$ teljesül.

2.1.3. Példa. Kétszer dobunk egy szabályos érmével, legyen ismét Z a fejek száma. Ekkor

$$\{Z < 2\} = \{Z = 0\} \cup \{Z = 1\} = \{II, FI, IF\}.$$

Amennyiben Ω -n egy valószínűségi mérték is adott, akkor egy $X : \Omega \rightarrow \mathbb{R}$ diszkrét valószínűségi változó egy valószínűségi mezőn van értelmezve. Ekkor, mivel a fenti definíciókban

definiált halmazok események, beszélhetünk például a $\mathbb{P}(X = 1) := \mathbb{P}(\{X = 1\})$ vagy a $\mathbb{P}(X < 2) := \mathbb{P}(\{X < 2\})$ valószínűségekről. Ezekben az esetekben a jelölés egyszerűsítése érdekében a halmazokat jelölő kapcsos zárójeleket általában elhagyjuk (tehát például a fenti két valószínűségnek tipikusan az első írásmódját használjuk).

Mivel X értéke egy adott $\omega \in \Omega$ -ra egyértelműen definiált, így $s, t \in \mathbb{R}$, $s \neq t$ esetén az $\{X = s\}$ és $\{X = t\}$ események egymást kizáróak, ezért tehát

$$\mathbb{P}(X = s \text{ vagy } X = t) = \mathbb{P}(X \in \{s, t\}) = \mathbb{P}(X = s) + \mathbb{P}(X = t)$$

teljesül. Ebből már látszik, hogy az X értékeinek segítségével jellemezhető események valószínűségeinek meghatározásához elegendő a $\mathbb{P}(X = t)$ valószínűségeket ismerni, ahol t végigfut az X értékkészletén. Általában egy tetszőleges $H \subset \mathbb{R}$ halmazra

$$(2.1) \quad \mathbb{P}(X \in H) = \sum_{t \in H \cap \text{ran } X} \mathbb{P}(X = t)$$

teljesül.

2.1.4. Példa. Legyen X egy kockadobás eredménye. Ekkor

$$\mathbb{P}(X = 1) = \mathbb{P}(X = 2) = \mathbb{P}(X = 3) = \mathbb{P}(X = 4) = \mathbb{P}(X = 5) = \mathbb{P}(X = 6) = \frac{1}{6},$$

továbbá pl.

$$\begin{aligned} \mathbb{P}(\text{párosat dobunk}) &= \mathbb{P}(X = 2 \text{ vagy } X = 4 \text{ vagy } X = 6) \\ &= \mathbb{P}(X = 2) + \mathbb{P}(X = 4) + \mathbb{P}(X = 6) = \frac{3}{6} = \frac{1}{2}. \end{aligned}$$

2.1.4. Definíció. Ha $X : \Omega \rightarrow \mathbb{R}$ egy diszkrét valószínűségi változó, akkor minden $t \in \text{ran } X$ értékre legyen $f_X(t) := \mathbb{P}(X = t)$, minden más $t \in \mathbb{R} \setminus \text{ran } X$ valós számra pedig legyen $f_X(t) := 0$. Az így kapott $f_X : \mathbb{R} \rightarrow [0; 1]$ függvényt az X változó *súlyfüggvényének* nevezzük.

A fentiek értelmében tehát egy X diszkrét valószínűségi változó f_X súlyfüggvénye meghatározza az X értékeivel kifejezhető események valószínűségét, és ezzel a jelölésmóddal egy tetszőleges $H \subset \mathbb{R}$ halmaz esetén a (2.1) formula a

$$\mathbb{P}(X \in H) = \sum_{t \in H \cap \text{ran } X} f_X(t)$$

alakba írható. Azt mondjuk, hogy az X súlyfüggvénye meghatározza az X *eloszlását*. Fontos megjegyezni, az f_X függvény megadása természetesen magában foglalja az X értékkészletének megadását is, tehát ehhez mind a felvehető értékeket, mind azok felvételének valószínűségeit meg kell adnunk.

Később (a következő fejezetben) majd látni fogjuk, hogy az X eloszlását, vagyis az utóbbi formula bal oldalán szereplő valószínűségeket nem csak a súlyfüggvény értékeiből lehet egyértelműen meghatározni, tehát a súlyfüggvény megadása az eloszlás megadásának *egyik*, de nem az *egyetlen* módja.

Megjegyzés. A magyar nyelvű irodalomban rendkívüli sokszínűség uralkodik mind a súlyfüggvény elnevezésének, mind a jelölések tekintetében. Az f_X súlyfüggvényt gyakran nevezik valószínűségi tömegfüggvénynek, valószínűségi függvénynek, sőt, valószínűségeloszlásnak is.

Bár a valószínűségi mértékek esetén is beszéltünk súlyfüggvényekről, a valószínűségi változók esetén továbbra is következetesen a súlyfüggvény megnevezést fogjuk használni. Ahol esetleg a szövegkörnyezetből nem egyértelmű, ott mindig kiemeljük, hogy egy változó súlyfüggvényéről beszélünk-e, vagy pedig egy eseménytér kimenetelein értelmezett súlyfüggvényről. Az előbbieket következetesen f , míg utóbbiakat p betűvel fogjuk jelölni.

2.1.5. Példa. Legyen Y egy kockadobás eredményének négyzete. Ekkor Y az 1, 4, 9, 16, 25 és 36 értékeket veheti fel, mindegyiket $\frac{1}{6}$ valószínűséggel, tehát a súlyfüggvénye:

$$f_Y(1) (= \mathbb{P}(Y = 1) =) f_Y(4) = f_Y(9) = f_Y(16) = f_Y(25) = f_Y(36) = \frac{1}{6}.$$

2.1.6. Példa. Legyen A egy esemény, melyre $\mathbb{P}(A) = p$. Ekkor az A -hoz tartozó $\mathbb{1}_A$ indikátor valószínűségi változó eloszlása:

$$f_{\mathbb{1}_A}(1) = \mathbb{P}(\mathbb{1}_A = 1) = \mathbb{P}(A) = p, \quad f_{\mathbb{1}_A}(0) = \mathbb{P}(\mathbb{1}_A = 0) = \mathbb{P}(\bar{A}) = 1 - \mathbb{P}(A) = 1 - p.$$

A binomiális eloszlás

2.1.7. Példa. Dobjunk fel egy kockát ötször egymás után, és legyen X a dobott hatosok száma. Mennyi az X változó súlyfüggvényének értéke például a 2 helyen? Azaz: mennyi az $f_X(2) = \mathbb{P}(X = 2)$ valószínűség? Minden dobásnál $\frac{1}{6}$ a hatos valószínűsége, míg $\frac{5}{6}$ eséllyel más lesz az eredmény. Egy olyan dobássorozat tehát, ahol pontosan kettő, előre fixált dobásnál kapunk hatost, $(\frac{1}{6})^2 \cdot (\frac{5}{6})^3$ valószínűséggel adódik, hiszen az egyes dobások függetlenek egymástól. A két hatosdobás "helyét" (azaz a két dobás sorszámát) $\binom{5}{2} = 10$ -féleképp választhatjuk ki, és így a keresett valószínűség $10 \cdot (\frac{1}{6})^2 \cdot (\frac{5}{6})^3 = \frac{625}{3888} \approx 0,1608$.

Ez utóbbi gondolatmenet általánosabban is elmondhatjuk. Tegyük fel, hogy egymástól függetlenül elvégzünk egy kísérletet n -szer egymás után. Legyen X egy fix p valószínűségű esemény bekövetkezéseinek a száma eme kísérletek során. Ekkor X értékkészlete a $\{0, 1, \dots, n\}$ halmaz. Az $\{X = k\}$ esemény azon lehetséges kísérletsorozatokról áll, melyeknél pontosan k -szor következik be a vizsgált esemény. Ez azt jelenti, hogy az a többi $n - k$ esetben nem következik be (vagyis ezeknél az esemény $1 - p$ valószínűségű komplementere következik be). Mivel a kísérleteket egymástól függetlenül végezzük, így egyetlen ilyen kísérletsorozat valószínűsége az egyes kísérletekben adódó eredmények valószínűségeinek szorzata, azaz $p^k(1 - p)^{n-k}$.

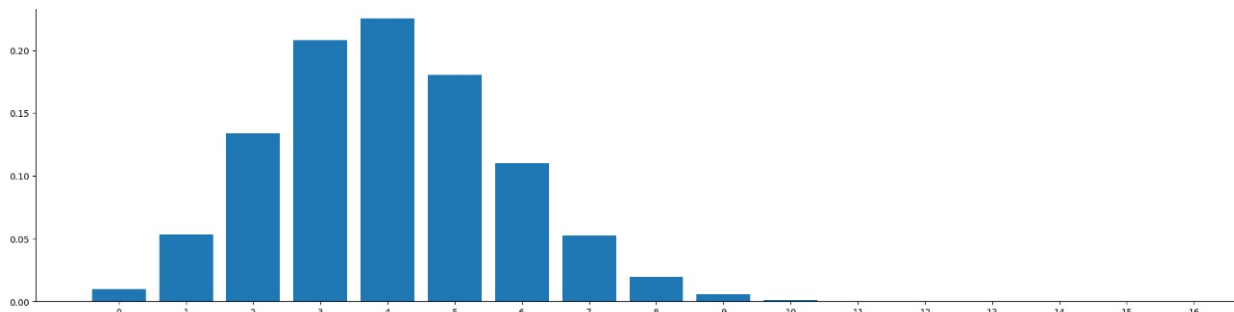
Hány olyan kísérletsorozat van, ahol éppen k -szor következik be a vizsgált esemény (éppen k darab "sikeres kísérlet van")? A sikeres kísérletek sorszámainak halmaza $\binom{n}{k}$ -féleképp választható (ez a választás pedig már egyértelműen meghatározza a (maradék) "sikertelen" kísérleteket). Tehát az $\{X = k\}$ eseményt $\binom{n}{k}$ darab azonos $p^k(1 - p)^{n-k}$ valószínűségű, páronként egymást kizáró esemény alkotja, ezért

$$(2.2) \quad f_X(k) = \mathbb{P}(X = k) = \binom{n}{k} p^k (1 - p)^{n-k}.$$

A fenti súlyfüggvény egy nevezetes eloszlást ad meg:

2.1.5. Definíció. Az X diszkrét valószínűségi változó *binomiális eloszlású* az $n \in \mathbb{N}^+$ és $p \in [0, 1]$ paraméterekkel, ha értékkészlete a $\{0, 1, \dots, n\}$ halmaz, és az $f_X(k)$ súlyfüggvény értékét a (2.2) formula jobb oldalán szereplő kifejezés adja meg minden $0 \leq k \leq n$ egész esetén. Jelölés: $X \sim \text{Bin}(n; p)$.

A fenti utolsó példában tehát X egy binomiális eloszlású változó volt. Bár sok esetben ilyen változók kísérletsorozatok modelljeiből adódnak, ez nem szükségszerűen van így. Az eloszlás csak a különböző értékek felvételének *valószínűségét* írja le, hogy magát a valószínűségi változót hogyan konstruáljuk, arról semmit nem mond.



2.3. ábra. A $Bin(16, \frac{1}{4})$ eloszlás súlyfüggvényének értékei

A fenti ábra egy $Bin(16; \frac{1}{4})$ eloszlású változó súlyfüggvényének értékét mutatja. Jól látható, hogy az egyes k értékekhez tartozó valószínűségek egy darabig nőnek, majd egy ponttól kezdve csökkenő sorozatot alkotnak. A következő állítás mutatja, hogy ez egy tetszőleges binomiális eloszlású változó esetén így van, sőt, azt is könnyen meghatározhatjuk, hogy melyik értéket veszi fel a változó a legnagyobb valószínűséggel. Az állítás bizonyítása megtalálható pl. a [2] könyvben.

2.1.1. Állítás. Legyen $X \sim Bin(n; p)$ egy binomiális eloszlású valószínűségi változó. Ekkor az $f_X(k)$ súlyfüggvény monoton növekvő a $[0; \lfloor (n+1)p \rfloor]$ intervallum által tartalmazott egészek halmazán, míg monoton csökkenő az $[\lfloor (n+1)p \rfloor; n]$ intervallumon. Következésképp a változó súlyfüggvénye az $m = \lfloor (n+1)p \rfloor$ értékre maximális. Amennyiben $(n+1)p$ egy pozitív egész (és így $m = \lfloor (n+1)p \rfloor = (n+1)p$ teljesül), akkor $f_X(m) = f_X(m-1)$, és a maximum pontosan ezen két értékre vétetik fel. Egyéb esetben a maximumhely egyértelmű.

A geometriai eloszlás

2.1.8. Példa. Addig ismétlünk egymás után többször függetlenül egy kísérletet, amíg egy p valószínűségű esemény be nem következik. Ilyen példák a következők: addig dobálunk egy érmét, amíg fejet nem kapunk (ebben az esetben $p = \frac{1}{2}$), vagy addig dobálunk egy kockát, amíg hatost nem dobunk (itt $p = \frac{1}{6}$). Legyen X a szükséges kísérletek száma. Ha k egy pozitív egész, akkor $f_X(k) = \mathbb{P}(X = k)$ tehát annak a valószínűsége, hogy az első $k-1$ kísérlet során a vizsgált esemény nem következik be (tehát az $1-p$ valószínűségű komplementere következik be), majd a k -edik kísérlet során a vizsgált esemény bekövetkezik. Mivel a kísérletek egymástól függetlenek, így

$$(2.3) \quad f_X(k) = (1-p)^{k-1}p.$$

2.1.6. Definíció. Az X diszkrét valószínűségi változó *geometriai eloszlású* $p \in (0; 1)$ paraméterrel, ha értékészlete a pozitív egészek halmaza, és (2.3) teljesül minden $k \in \mathbb{N}^+$ pozitív egész esetén. Jelölés: $X \sim Geo(p)$.

Az 1.2.2. példában és az 1.2.2. gyakorlatban láttuk, hogy egy pénzérmével egymás után néhány fejet dobva ez nem befolyásolja annak az esélyeit, hogy a következő dobásnál is fejet

(vagy írást) dobjunk. Bár mindezt a konkrét esetekben számolással is alátámasztottuk, a jelenség mögöttes oka az, hogy a dobások egymástól függetlenek, és így a már megtörtént események nem lesznek hatással a soron következőkre.

Ugyanez az érvelés általánosan is elmondható a 2.1.8. példában szereplő szituációban, tehát ha egymástól függetlenül ismételnünk egy kísérletet egy adott esemény bekövetkezéséig. Ha a kísérlet k -szor sikertelen volt, azaz ha tudjuk, hogy a példában szereplő X változó értéke k -nál nagyobb, az nincs hatással arra, hogy ezután még hányszor kell azt elvégezni a következő sikerig. Ezt írja le formálisan a következő állítás:

2.1.2. Állítás. Legyen $X \sim \text{Geo}(p)$ valószínűségi változó. Legyen továbbá $k, n \in \mathbb{N}^+$, ekkor

$$(2.4) \quad \mathbb{P}(X > k + n \mid X > k) = \mathbb{P}(X > n).$$

A fenti állításban szereplő tulajdonságnak nevet is adunk:

2.1.7. Definíció. Az X valószínűségi változót *örökifjúnak* nevezzük az \mathbb{N}^+ halmazon, amennyiben $\text{ran } X = \mathbb{N}^+$, és a (2.4) tulajdonság teljesül X -re minden $k, n \in \mathbb{N}^+$ esetén.

Úgy fogalmazhatunk tehát, hogy egy X geometriai eloszlású valószínűségű változó örökifjú az \mathbb{N}^+ halmazon. A fenti érvelés (még ha remélhetőleg meggyőző is) nem egy precíz indoklás, így az előző állítás még bizonyításra szorul.

Bizonyítás. Írjuk fel a feltételes valószínűség definíciója alapján az (2.4) egyenlőség bal oldalát:

$$\mathbb{P}(X > k + n \mid X > k) = \frac{\mathbb{P}(\{X > k + n\} \cap \{X > k\})}{\mathbb{P}(X > k)}.$$

Vizsgáljuk meg a jobb oldal számlálójában szereplő eseményt. Mivel $n > 0$, így minden egyes kimenetel esetén, amelyre $X > k + n$ teljesül, $k + n > k$ miatt egyben $X > k$ is igaz. Azaz $\{X > k + n\} \subset \{X > k\}$, ezért

$$\{X > k + n\} \cap \{X > k\} = \{X > k + n\}.$$

Így tehát

$$(2.5) \quad \mathbb{P}(X > k + n \mid X > k) = \frac{\mathbb{P}(X > k + n)}{\mathbb{P}(X > k)} = \frac{1 - \mathbb{P}(X \leq k + n)}{1 - \mathbb{P}(X \leq k)} = \frac{1 - \sum_{i=1}^{k+n} f_X(i)}{1 - \sum_{j=1}^k f_X(j)}.$$

Most egy olyan trükköt mutatunk be, ami nagyon gyakran sokat egyszerűsít a számolásokon. A fenti egyenlőség jobb oldala mutatja, hogy a bal oldalon álló feltételes valószínűség kizárólag az X eloszlásától, vagyis az $f_X(k)$ súlyoktól függ. Tehát elegendő egy konkrét (de tetszőleges), p paraméterű geometriai eloszlású valószínűségi változó esetén kiszámolni ezt az értéket, az eredmény pedig az összes azonos eloszlású változó esetén érvényes lesz.

Tekintsük a 2.1.8. példában szereplő X valószínűségi változót. Az X tehát a szükséges kísérletek száma egy olyan kísérletsorozatban, melyben egymástól függetlenül ismételjük a kísérleteket, és egy p valószínűségű esemény első bekövetkezésére várunk. A példában láttuk, hogy $X \sim \text{Geo}(p)$ teljesül.

Amint azt már az állítás előtt is említettük, pontosan akkor végzünk több, mint k kísérletet, ha az első k kísérlet sikertelen volt. Azaz $\mathbb{P}(X > k) = (1 - p)^k$, és hasonlóképp $\mathbb{P}(X > k + n) = (1 - p)^{k+n}$, így tehát

$$\mathbb{P}(X > k + n \mid X > k) = \frac{(1 - p)^{k+n}}{(1 - p)^k} = (1 - p)^n = \mathbb{P}(X > n),$$

és ezzel az állítást beláttuk. □

2.1.1. Feladat. (M) Lássuk be a fenti állítást a konkrét valószínűségi változó használata nélkül, azaz fejezzük be az érvelést a (2.5) egyenlőségből pusztán az $f_X(k)$ értékeket használva.

A fenti állításnak valójában a megfordítása is igaz, tehát az örökifjú eloszlások az \mathbb{N}^+ halmazon pontosan a geometriai eloszlások. Az alábbi tétel bizonyítása megtalálható például az [5] jegyzetben.

2.1.3. Tétel. *Ha egy X valószínűségi változó örökifjú az \mathbb{N}^+ halmazon, akkor $X \sim Geo(p)$ teljesül valamilyen $p \in (0; 1)$ számra.*

A súlyfüggvény jellemzése

Tekintsünk egy tetszőleges $X : \Omega \rightarrow \mathbb{R}$ diszkrét valószínűségi változót. Vegyük észre, hogy ha t végigfut az X értékészletén, akkor az $\{X = t\}$ páronként diszjunkt események uniója éppen a teljes Ω eseménytér (hiszen minden egyes $\omega \in \Omega$ -ra felveszi X valamelyik $t \in \text{ran } X$ értéket). Azaz ezek az események teljes eseményrendszert alkotnak, speciálisan az $f_X(t) = \mathbb{P}(X = t)$ valószínűségek összege, amint t végigfut a $\text{ran } X$ halmazon, szükségképpen az Ω biztos esemény valószínűsége, azaz 1. Beláttuk tehát a következőt:

2.1.4. Állítás. *Ha X egy diszkrét valószínűségi változó, f_X pedig a hozzá tartozó súlyfüggvény, akkor*

$$(2.6) \quad \sum_{t \in \text{ran } X} f_X(t) = 1.$$

A fenti egyenlőség egyszerűen látható például a 2.1.5. példában vagy egy indikátorváltozó esetén, de persze igaz binomiális vagy geometriai eloszlású változókra is, és persze ez utóbbi esetekben is leellenőrizhető akár egy egyszerű számolással is. Legyen $X \sim Bin(n; p)$, ekkor tehát

$$\sum_{k=0}^n f_X(k) = \sum_{k=0}^n \binom{n}{k} p^k (1-p)^{n-k}.$$

A jobb oldal a binomiális tétel (lásd az (1.2) formulát a 19. oldalon) miatt nem más, mint $(p + 1 - p)^n = 1^n = 1$, tehát (2.6) valóban teljesül. Egy $Y \sim Geo(p)$ geometriai eloszlású valószínűségi változó esetén a fenti egyenlőség egy geometriai sor összegzéséből adódik (innen az eloszlás neve):

$$\sum_{k=1}^{\infty} f_Y(k) = \sum_{k=1}^{\infty} (1-p)^{k-1} p = p \sum_{i=0}^{\infty} (1-p)^i = \frac{p}{1 - (1-p)} = 1.$$

A (2.6) egyenlőség valójában pontosan karakterizálja a súlyfüggvényt. Azaz, amennyiben p_1, \dots, p_n, \dots olyan $[0; 1]$ -beli számok egy véges vagy megszámlálhatóan végtelen sorozata, melyek összege 1, akkor megadható olyan diszkrét valószínűségi változó, amelynek súlyfüggvénye éppen ezeket az értékeket veszi fel. Ennek konstrukciója valójában nem különösebben komplikált, azt javasoljuk az olvasónak, hogy próbálkozzon meg egy megfelelő változót találni.

Műveletek valószínűségi változók között

Azonos eseménytereken definiált különböző valószínűségi változókat együtt is kezelhetünk. Például a valós számokon értelmezett műveletek, azaz az összeadás, kivonás, szorzás és osztás segítségével két változóból újabbakat készíthetünk:

2.1.8. Definíció. Legyenek $X, Y : \Omega \rightarrow \mathbb{R}$ ugyanazon eseménytéren értelmezett diszkrét valószínűségi változók. Ekkor

- $X + Y : \Omega \rightarrow \mathbb{R}$ az a diszkrét valószínűségi változó, amely minden $\omega \in \Omega$ esetén az $X(\omega) + Y(\omega)$ értéket veszi fel,
- $X \cdot Y : \Omega \rightarrow \mathbb{R}$ az a diszkrét valószínűségi változó, amely minden $\omega \in \Omega$ esetén az $X(\omega) \cdot Y(\omega)$ értéket veszi fel,
- ha $Y(\omega) \neq 0$ teljesül minden $\omega \in \Omega$ -ra, akkor $X/Y : \Omega \rightarrow \mathbb{R}$ az a diszkrét valószínűségi változó, amely minden $\omega \in \Omega$ esetén az $X(\omega)/Y(\omega)$ értéket veszi fel.

2.1.9. Példa. Dobjunk kétszer egy szabályos kockával. Legyen X az első, míg Y a második dobás eredménye, ekkor az $X + Y$ valószínűségi változó értéke a dobások értékének összege. Mi lesz $X + Y$ eloszlása? Az összeg lehetséges értékei az 2 és 12 közötti egész értékek. A 2 és 12 kizárólag egyféleképp adódhat összegként, mégpedig, ha két darab egyest ill. két darab hatost dobunk, ezen események valószínűsége pedig $\frac{1}{36}$:

$$f_{X+Y}(2) = \mathbb{P}(X + Y = 2) = \frac{1}{36} = f_{X+Y}(12) = \mathbb{P}(X + Y = 12).$$

A 3 érték már kétféleképp is adódhat összegként: $3 = 1 + 2 = 2 + 1$. Ezek különböző előállítások, mert a dobások sorrendjét is figyelembe vesszük. Ugyanez érvényes a 11-re is: $11 = 5 + 6 = 6 + 5$, tehát

$$f_{X+Y}(3) = f_{X+Y}(11) = \frac{2}{36} = \frac{1}{18}.$$

Hasonló módon megvizsgálva a lehetséges előállítások számát, a következő valószínűségeket kapjuk a további értékekre:

$$f_{X+Y}(4) = f_{X+Y}(10) = \frac{3}{36} = \frac{1}{12}, \quad f_{X+Y}(5) = f_{X+Y}(9) = \frac{4}{36} = \frac{1}{9},$$

$$f_{X+Y}(6) = f_{X+Y}(8) = \frac{5}{36}, \quad f_{X+Y}(7) = \frac{6}{36} = \frac{1}{6}.$$

2.1.10. Példa. Tekintsük a 2.1.7. példa általánosított verziójában szereplő $X \sim \text{Bin}(n; p)$ valószínűségi változót, ami tehát egy n független kísérletből álló kísérletsorozatban egy adott p valószínűségű esemény bekövetkezéseinek számát adja meg. Legyen A_i az az esemény, hogy az i -edik kísérletnél bekövetkezik a vizsgált esemény. Ekkor X előáll, mint az A_i -k indikátorváltozóinak összege: $X = \mathbb{1}_{A_1} + \dots + \mathbb{1}_{A_n}$, hiszen $X = k$ éppen akkor teljesül, ha pontosan k következik be az A_i események közül, azaz, ha az azokhoz tartozó k darab indikátor az 1, a többi pedig a 0 értéket veszi fel.

Általában, ha A_1, \dots, A_n egy valószínűségi mező tetszőleges együttesen független eseményei, melyekre $\mathbb{P}(A_i) = p$ teljesül minden $1 \leq i \leq n$ -re, akkor könnyen látható, hogy az $\mathbb{1}_{A_1} + \dots + \mathbb{1}_{A_n}$ összeg azt adja meg, hogy a fenti eseményekből hány darab következik be, és ennek eloszlása binomiális, méghozzá az n és p paraméterekkel.

A következő példában valószínűségi változók hányadosának az eloszlását határozzuk meg. Látni fogjuk, hogy ez nem feltétlenül nagyon egyszerű feladat, de a következő szakaszban részletesen elemzünk majd olyan szituációkat, ahol ez a probléma szisztematikusan kezelhető. Ugyanakkor a mostani példában a számolás melléktermékeként egyben egy számelméleti alkalmazást is nyerünk majd.

2.1.11. Példa. Egy kalapba beteszünk n cédulát, melyekre az $1, \dots, n$ számokat írjuk (minden cédulára pontosan egy számot, és minden számot pontosan egyszer). Visszatevés nélkül kihúzzunk ezek közül véletlenszerűen két darabot, legyen X értéke a kihúzott cédulákra írt számok közül a kisebbik, míg Y értéke a nagyobbik szám.

Meghatározzuk az X, Y ill. X/Y változók eloszlását. A húzás eredményét olyan $(i; j)$ rendezett számpárokkal írhatjuk le, ahol $1 \leq i < j \leq n$. Ezek száma $\binom{n}{2}$, és minden ilyen számpár $1/\binom{n}{2}$ valószínűséggel adódik.

Az X értékészlete az $\{1, 2, \dots, n-1\}$ halmaz. Az $\{X = i\}$ esemény éppen azt jelenti, hogy a nagyobb húzott szám értéke i -nél nagyobb, tehát a nagyobb értékre $n-i$ lehetőség van. Azaz

$$\mathbb{P}(X = i) = \frac{n-i}{\binom{n}{2}}$$

minden $1 \leq i < n$ esetén.

Az X értékészlete a $\{2, \dots, n\}$ halmaz, hiszen a nagyobbik szám érték szükségképp legalább 2. A fentiekhez hasonlóan adódik, hogy ha a nagyobb j érték adott (azaz $Y = j$), akkor a kisebb értékre $j-1$ lehetőség van, azaz

$$\mathbb{P}(Y = j) = \frac{j-1}{\binom{n}{2}}$$

minden $2 \leq j \leq n$ esetén.

Vizsgáljuk most meg az X/Y változó lehetséges értékeit. Világos, hogy minden ilyen érték egy racionális szám, amelyet egyszerűsítés után a/b alakba írhatunk, ahol $1 \leq a < b \leq n$, továbbá a és b relatív prímek (azaz $(a, b) = 1$). Természetesen minden ilyen tört előáll az X és Y hányadosaként, ezekből áll tehát az X/Y értékészlete.

Egy a/b törtet azonban több $(i; j)$ pár esetén megkaphatunk, hiszen az $(ma)/(mb)$ törtek értéke minden m pozitív egész esetén a/b . Itt m értékére az $mb \leq n$ egyenlőtlenségnek teljesülnie kell, hiszen n -nél nagyobb számot nem húzhatunk. Tehát az m lehetséges értékei a pozitív egészek 1 és n/b között, ezek száma pedig $\lfloor \frac{n}{b} \rfloor$, vagyis

$$(2.7) \quad \mathbb{P}\left(\frac{X}{Y} = \frac{a}{b}\right) = \frac{\lfloor \frac{n}{b} \rfloor}{\binom{n}{2}}$$

minden $1 \leq a < b \leq n$, $(a, b) = 1$ esetén.

Ha a fenti valószínűségeket most összeadjuk minden lehetséges a, b értékre, akkor természetesen 1-et kapunk, hiszen a változó súlyfüggvényének értékeit összegezzük. A fenti valószínűségekben azonban a tört nevezője mindig azonos, ezzel átszorozva pedig

$$\sum_{\substack{1 \leq a < b \leq n \\ (a, b) = 1}} \left\lfloor \frac{n}{b} \right\rfloor = \binom{n}{2}$$

adódik. Figyeljük meg, hogy a (2.7) valószínűség független a értékétől, tehát a fenti összegben egy fix b esetén minden a -ra ugyanazt a tagot összegezzük:

$$(2.8) \quad \binom{n}{2} = \sum_{1 < b \leq n} \sum_{1 \leq a < b, (a, b) = 1} \left\lfloor \frac{n}{b} \right\rfloor = \sum_{b=2}^n \varphi(b) \left\lfloor \frac{n}{b} \right\rfloor,$$

ahol φ az ún. *Euler-féle fí függvény*, amelynek értéke minden b pozitív egész esetén a b -hez relatív prím pozitív egészek száma b -ig.

A (2.8) egyenlőség ebben a formában még hasznunkra válhat majd további számolásoknál, azonban ezt a formulát úgy szokás megadni, hogy a $b = 1$ értékhez tartozó tagot is hozzávesszük mindkét oldalhoz, tehát az egyenlethez hozzáadjuk a $\varphi(1)\left[\frac{n}{1}\right] = n$ értéket:

$$(2.9) \quad \sum_{b=1}^n \varphi(b) \left[\frac{n}{b}\right] = \binom{n}{2} + n = \frac{n(n-1)}{2} + n = \frac{n(n+1)}{2} = \binom{n+1}{2}.$$

Ennek a formulának van egy egyszerű következménye. Írjuk fel a fenti egyenletet n -re és $(n-1)$ -re is, majd az utóbbit vonjuk ki az előbbiből:

$$\begin{aligned} n &= \binom{n+1}{2} - \binom{n}{2} = \sum_{b=1}^n \varphi(b) \left[\frac{n}{b}\right] - \sum_{b=1}^{n-1} \varphi(b) \left[\frac{n-1}{b}\right] \\ &= \varphi(n) + \sum_{b=1}^{n-1} \varphi(b) \left\{ \left[\frac{n}{b}\right] - \left[\frac{n-1}{b}\right] \right\}. \end{aligned}$$

Az utolsó különbség pontosan akkor 1, ha n -et b -vel maradékosan osztva lesz a hányados, mint ha $(n-1)$ -et osztanánk b -vel maradékosan, azaz ha b osztója n -nek. Egyéb esetekben pedig a különbség 0. Azaz összefoglalva

$$(2.10) \quad \sum_{d|n} \varphi(d) = n,$$

ahol a bal oldalon az n pozitív osztóira összegzünk. Ezt az összeget a φ *összegzési függvényének* nevezzük.

A φ függvény a számelmélet egy alapvető függvénye, aminek jelentősége messze nem csak elméleti: fontos szerep jut neki például az RSA nyílt kulcsú titkosító algoritmus matematikai hátterének felépítésében, de számos más számelméleti algoritmusban (pl. prímtesztben) is. Minderről az érdeklődő olvasó részletesen olvashat pl. a [10] jegyzet első vagy a [3] könyv ötödik fejezében.

Magának a (2.10) formulának fontos szerep jut például a φ függvény átlagos értékének vizsgálatában, lásd pl. a [3] könyv 6.7. szakaszát. Természetesen a φ összegzési függvényének értéke más módon is levezethető, sőt, esetenként a (2.9) formulát vezetik le az összegzési függvényből teljes indukcióval, habár a fenti kombinatorikus érvelés jóval természetesebb. Egy másik lehetőség a (2.10) egyenlőség levezetésére a φ függvény multiplikatív használatára: ha n, m relatív prím pozitív egészek, akkor $\varphi(nm) = \varphi(n)\varphi(m)$ (lásd pl. az 1.5.4. Lemmát a [10] jegyzetben). Belátható, hogy multiplikatív függvény összegzési függvénye is multiplikatív (lásd a [3] könyv 6. fejezetét), így azt meghatározzák annak prímhatvány helyen felvett értékei. Ezek egyszerűen kiszámolhatók, innen pedig már könnyen adódik az érték tetszőleges n -re. Másrészt viszont magának a φ függvénynek a multiplikatívítása belátható az összegzési függvény (tehát az $F(n) = n$ függvény) multiplikatívításából az ún. megfordítási függvény segítségével, felhasználva, hogy egy multiplikatív függvény megfordítási függvénye szintén multiplikatív. A részleteket illetően ismét csak a [3] könyv hatodik fejezetére hivatkozunk.

*Szorzatmértékek (kiegészítő anyag)

A szakasz zárásaként néhány kiegészítő megjegyzést teszünk a felépített valószínűségi modellel kapcsolatban. Az olvasónak talán feltűnt, hogy míg a valószínűségi változókra adott

első példákban részletesen tárgyaltuk, hogy milyen eseménytérre vannak értelmezve, ill. hogy az egyes kimenetek esetén hogyan kapjuk a változó értékét, ezek a részletek már elmaradtak, vagy legalábbis csak szemléletesen voltak körülírva (többek között) a binomiális és a geometriai eloszlás példáinál.

Ez a jelenség általában is megfigyelhető, nem csupán e jegyzet hátralévő részében, hanem a témával kapcsolatos irodalomban általában is. Ennek oka esetenként egyszerűen az, hogy a háttérben megbúvó modell lényegét tekintve nagyon egyszerű, máskor viszont éppen az, hogy túl bonyolult. Azaz, sokszor az eseménytér és a rajta értelmezett valószínűségi mérték precíz definíciójának technikai részletei feleslegesen komplikálttá és hosszadalmassá teszik a tárgyalást, és inkább csak elfednek, mintsem megvilágítanak fontos és egyszerű gondolatokat. Ez a helyzet pl. a binomiális eloszlás esetén. Vannak azonban olyan esetek is, amikor a jó konstrukció megalkotása egyáltalán nem triviális feladat, tehát nem pusztán technikai jellegű nehézségekbe ütközünk, és a részletek tárgyalása egyszerűen túl messzire vezetne. Ilyen pl. a geometriai eloszlás esete.

Mivel ez a jegyzet bevezető jellegű, és (többek közt) egy jó szemlélet kialakítását tűztük ki célul, így kezdetben inkább csak elvéve adunk meg technikai érveléseket. Azonban fontos hangsúlyozni, hogy ha valaki a valószínűségszámítás mélyebb eredményeit alkalmazni szeretné, akkor elkerülhetetlen, hogy egy ponton elmélyüljön a részletekben. Esetünkben például, ha valószínűségi változókkal dolgozunk, akkor képesnek kell lennünk megadni az eseményteret és megmutatni, hogy a valószínűség, amivel számolunk, valóban egy valószínűségi mértékből származik, különben semmi sem jogosítana fel minket a kiépített eszköztár használatára, tehát a levezetéseink helytelenek volnának, és rossz eredményekre vezethetnének. A 3.1. szakaszban látni fogunk egy konkrét példát is arra, hogy pusztán a szemléletünkre támaszkodva hogyan vezethetnek tévútra érveléseink.

A modell precíz konstrukcióját most megadjuk a binomiális eloszlás (egyik) példájánál, a geometriai eloszlás esetében viszont ezt elhagyjuk (az érdeklődő olvasó megtalálja azt a [4] könyv első fejezetében). Az alábbiakban n független kísérleteket szeretnénk modellezni, így azt a módszert alkalmazhatjuk, amit már az 1.1. szakaszban is használtunk ilyen esetben, azaz eseményterek Descartes-szorzatával dolgozunk. Fontos megjegyezni, hogy általában nem klasszikus valószínűségi mezőből indulunk ki, ezért most az 1.1. szakasz példáját általánosítanunk kell.

Legyenek tehát $(\Omega_i, \mathcal{A}_i, \mathbb{P}_i)$ diszkrét valószínűségi mezők $(1 \leq i \leq n)$. Nincs szükség rá, hogy ezek megegyezzenek, csupán azt fogjuk modellezni, hogy a kísérleteink függetlenek egymástól, nem kell, hogy ezek a kísérletek egyformák legyenek. Azt feltételezzük tehát, hogy az i -edik kísérlet kimenetele az Ω_i eseménytér eleme, és így a kísérletsorozat az

$$\Omega = \Omega_1 \times \cdots \times \Omega_n$$

szorzathalmaz írja le. Belátható, hogy amennyiben mindegyik Ω_i megszámlálható halmaz, úgy a fenti szorzat is az, ennek megfelelően a korábbiak szerint minden részhalmazát eseménynek tekinthetjük. (Megjegyezzük, hogy ez a körülmény leegyszerűsíti a dolgunkat, az általános esetben az események megadása ennél komplikáltabb.)

Tegyük fel, hogy az i -edik kísérletben az $A_i \subset \Omega_i$ esemény bekövetkezését vizsgáljuk. Ezt az eseményt a modellünkben Ω szorzathalmaz következő részhalmaza írja le:

$$A'_i := \Omega_1 \times \cdots \times \Omega_{i-1} \times A_i \times \Omega_{i+1} \times \cdots \times \Omega_n,$$

tehát olyan n -eseket veszünk, melyeknek az i -edik eleme az A_i halmazban van, minden $j \neq i$ esetén a j -edik elem pedig Ω_j tetszőleges eleme lehet (ez jelenti azt, hogy csak az i -edik kísérlet eredményét tekintjük).

Mivel a kísérleteket függetlennek gondoljuk, ezért az A'_i eseményeknek a modellünkben is függetlennek kell lenniük (az 1.2.2. definíció értelmében), azaz ha \mathbb{P} egy megfelelő valószínűségi mérték volna az Ω szorzathalmaz eseményein, akkor teljesülnie kellene rá (többek között) a

$$(2.11) \quad \mathbb{P}(A'_1 \cap \dots \cap A'_n) = \mathbb{P}(A'_1) \dots \mathbb{P}(A'_n)$$

egyenlőségnek. Vegyük észre, hogy a bal oldalon lévő metszet éppen az $A_1 \times \dots \times A_n$ szorzathalmaz, továbbá mivel A'_i és A_i ugyanazon eseményt hivatottak modellezni, így az egyetlen értelmes választás a $\mathbb{P}(A'_i) = \mathbb{P}_i(A_i)$ definíció volna.

A fenti érvelés azonban már az összes lehetséges kimenetel valószínűségét meghatározza. Valóban, ha $\Omega_i = \{\omega_{i,1}, \omega_{i,2}, \dots\}$ az i -edik eseménytér kimenetelinek egy felsorolása, akkor a szorzathalmaz kimenetelei az $(\omega_{1,k_1}, \omega_{2,k_2}, \dots, \omega_{n,k_n})$ n -esek, ahol $k_i \in \mathbb{N}^+$ minden $1 \leq i \leq n$ esetén (k_i persze nem feltétlenül lehet akármilyen egész akkor, ha az Ω_i halmaz véges). Ha most az $A_i = \{\omega_{i,k_i}\}$ választással élünk, akkor a fentiek szerint teljesülnie kell a

$$\mathbb{P}(\{(\omega_{1,k_1}, \omega_{2,k_2}, \dots, \omega_{n,k_n})\}) = \mathbb{P}_1(\{\omega_{1,k_1}\}) \cdot \mathbb{P}_2(\{\omega_{2,k_2}\}) \dots \mathbb{P}_n(\{\omega_{n,k_n}\})$$

egyenletnek. A jelölések egyszerűsítése érdekében legyen $p_{i,k_i} = \mathbb{P}_i(\{\omega_{i,k_i}\})$. Azt kaptuk tehát, hogy a

$$(2.12) \quad \mathbb{P}(\{(\omega_{1,k_1}, \omega_{2,k_2}, \dots, \omega_{n,k_n})\}) = \prod_{i=1}^n p_{i,k_i}$$

az egyetlen értelmes lehetséges definíció a kimenetek valószínűségeire. Az 1.1.5. állítás szerint a fenti értékek pontosan akkor határoznak meg egy valószínűségi mértéket az Ω szorzathalmazon, ha az összegük 1, ebben az esetben pedig ez a mérték egyértelmű. Ezeknek szorzatoknak az összege

$$\sum_{k_1} \sum_{k_2} \dots \sum_{k_n} \prod_{i=1}^n p_{i,k_i} = \prod_{i=1}^n \left(\sum_{k_i} p_{i,k_i} \right)$$

ahol a fenti szummákban az Ω_i halmazok fenti felsorolásában előforduló összes (véges vagy végtelen sok) k_i indexre szummázunk. A fenti egyenlet voltaképp a jobb oldali zárójelek felbontásának eredményét fejezi ki, ami még végtelen összeg esetén is megtehető (mivel csak nemnegatív számokat összegzünk). A jobb oldalon lévő szummák azonban 1-et adnak (és így persze a szorzatuk is), hiszen itt minden esetben egy eseménytér összes kimenetelének valószínűségeit adjuk össze. Ezzel tehát beláttuk, hogy a (2.12) formula segítségével egyértelműen megadunk egy valószínűségi mértéket az Ω szorzathalmazon, ezt a \mathbb{P}_i mértékek *szorzatának* nevezzük, és a $\mathbb{P}_1 \otimes \dots \otimes \mathbb{P}_n$ szimbólummal jelöljük.

Az 1.1.5. állítás tetszőleges esemény valószínűségét is megadja, az az általa tartalmazott kimenetek valószínűségeinek összege. Speciálisan, ha $A_i \subset \Omega_i$ egy esemény Ω_i -ben, akkor

$$\begin{aligned} \mathbb{P}(A_1 \times \dots \times A_n) &= \sum_{\omega_1 \in A_1} \dots \sum_{\omega_n \in A_n} \mathbb{P}(\{\omega_1, \dots, \omega_n\}) = \sum_{\omega_1 \in A_1} \dots \sum_{\omega_n \in A_n} \prod_{i=1}^n \mathbb{P}_i(\{\omega_i\}) \\ &= \prod_{i=1}^n \left(\sum_{\omega_i \in A_i} \mathbb{P}_i(\{\omega_i\}) \right) = \prod_{i=1}^n \mathbb{P}_i(A_i) = \prod_{i=1}^n \mathbb{P}(A'_i). \end{aligned}$$

Megjegyezzük, hogy Ω -nak nem minden részhalmaza írható a fenti szorzat alakba.

Tehát az A'_i eseményekre előzetes szándékainknak megfelelően teljesül a (2.11) egyenlet. Hasonlóképp látható, hogy az együttes függetlenséget definiáló többi egyenlet is érvényes, tehát az A_i eseményeknek a szorzattérben megfelelő A'_i események valóban együttesen függetlenek.

Végezetül, ha X egy kísérlet során azon kísérletek számát hivatott modellezni, ahol az i -edik kísérletben bekövetkezett az A_i esemény, akkor egy $(\omega_1, \dots, \omega_n)$ kimenetel esetén X azon indexek számát jelenti, ahol $\omega_i \in A_i$ teljesül. A $\mathbb{P}(X = k)$ valószínűség tehát azon események valószínűségének összege, ahol $\omega_i \in A_i$ pontosan k darab indexre teljesül. Tegyük fel az egyszerűség kedvéért, hogy ez az első k index, ekkor ez utóbbi esemény éppen

$$A_1 \times \dots \times A_k \times \overline{A_{k+1}} \times \dots \times \overline{A_n},$$

ennek valószínűsége pedig

$$\mathbb{P}_1(A_1) \dots \mathbb{P}_k(A_k)(1 - \mathbb{P}_{k+1}(A_{k+1})) \dots (1 - \mathbb{P}_n(A_n)).$$

Az ezekkel analóg szorzatokat kell tehát az indexek minden fixált k -asára összegezni. A szakasz korábbi példáiban a \mathbb{P}_i mértékek és a $\mathbb{P}_i(A_i)$ valószínűségek is megegyeztek, így adódott a formula a binomiális eloszlás súlyfüggvényére.

Látható, hogy időnként a precíz definíciók és levezetések megadása kissé hosszadalmas és talán fáradságos munka, ráadásul ebben az esetben nem sok fontos adalékkal szolgál a szemléletes (de nem precíz) érveléshez képest. Éppen ezért az ilyen részleteket a jegyzet további részében legtöbbször elhagyjuk. Megadjuk viszont a technikai részleteket, amikor azok olyan eszköztárat vezetnek be vagy használnak fel, aminek a jó ismerete az olvasó számára is nagyon hasznos, esetleg nélkülözhetetlen lesz a későbbiek megértéséhez.

Gyakorlatok, feladatok

2.1.2. Gyakorlat. Egy urnában 3 piros, 5 fehér golyó van. Addig húzunk az urnából *visszatétel nélkül*, amíg piros golyót nem húzunk. Jelölje X a húzások számát. Adjuk meg a $\mathbb{P}(X < 4)$ és a $\mathbb{P}(X \geq 3)$ valószínűségeket.

2.1.3. Gyakorlat. (M) Két szabályos dobókockával dobunk, jelölje X a dobott számok összegét, Y pedig a szorzatukat. Függetlenek-e az $\{X \text{ páros}\}$ és az $\{Y \leq 4\}$ események?

2.1.4. Gyakorlat. Egy vizsgán a kiosztott tesztlapon 10 feleletválasztós kérdés szerepel. Mindegyik kérdésre pontosan egy válasz jó a felkínált 4 lehetőség közül, és csak egyet szabad választani. Találomra kitöltünk egy ilyen tesztlapon (mindenféle előzetes tudás nélkül).

- Mi a valószínűsége, hogy egy találatunk sem lesz?
- Mekkora valószínűséggel érhetünk el legalább 4 találatot?
- Hány helyes választ kapunk a legnagyobb valószínűséggel? És ha 11 kérdés van a tesztben?

2.1.5. Gyakorlat. Tegyük fel, hogy az $\text{Indicator}(p)$ függvény p valószínűséggel 1 értéket ad vissza, és $1 - p$ valószínűséggel 0-t. Mi a következő függvény visszatérési értékének eloszlása?

```
def randomexperiments():
    k = 1
    while Indicator(0.5) == 0:
        k += 1
    return k
```

2.1.6. Gyakorlat. Addig dobunk egy szabályos kockával, amíg 3-nál kisebb számot nem kapunk. Jelölje X az ehhez szükséges dobások számát.

a) Melyik valószínűség nagyobb: $\mathbb{P}(2 \leq X \leq 3)$ vagy $\mathbb{P}(X > 3)$?

b) Mi a valószínűsége, hogy legalább 10 dobásra lesz szükségünk, feltéve, hogy az első 6 dobás esetén 2-nél nagyobb számot kapunk?

2.1.7. Gyakorlat. (M) Béla úgy dönt, hogy próbára teszi a szerencséjét, és mostantól minden héten vesz három darab, ugyanolyan típusú sorsjegyet (minden héten ugyanolyan) egészen addig, amíg legalább az egyikkel nem nyer valamit (tehát minden héten megvesz hármat, és ezután azonnal lekaparja őket). Tudható, hogy a sorsjegyekből minden tizedik rejt valamilyen nyereményt. Mi a valószínűsége, hogy Béla legfeljebb a harmadik hét után abbahagyja a sorsjegyvásárlást?

2.1.8. Feladat. A gyönyörű királylány kezéért n királyfi verseng, akiktől minden nap kap egy-egy randévumeghívást (minden királyfitól minden nap pontosan egyet). A királylányt nagyon lekötik udvari teendői, így meghagyja a szolgájának, hogy függetlenül attól, hogy épp melyik királyfi hívja meg és hogy a többieknek mit mondott, p valószínűséggel fogadja el a meghívást a nevében, $1 - p$ valószínűséggel pedig utasítsa el. Mivel a királylány rendkívül elfoglalt, így ha nem muszáj, nem akar egy nap több randévúra is elmenni, de egy királyfival azért minden nap szívesen találkozik. Hogyan válassza a p számot, hogy a terve megvalósulásának valószínűségét maximalizálja? Mihez tart ez a (maximális) valószínűség, ha $n \rightarrow \infty$?

2.1.9. Feladat. (M) Egy pénzérmét addig dobálunk, amíg egymás után 2 egyforma eredményt nem kapunk. Adjuk meg a szükséges dobások számának eloszlását. Mi lesz az eloszlás, ha egymás után 3 egyforma eredményig várunk?

2.2. Együttes eloszlás és függetlenség

A valószínűségi változók fogalmának egyik előnye, hogy különböző véletlen mennyiségek egyazon modellben (tehát ugyanazon eseménytér segítségével) kezelhetők. Ebben a szakaszban megvizsgáljuk, hogyan írható le több változó együttes viselkedése. Az különböző értékek párhuzamos nyilvántartásának természetes módja, ha azokat egy sorozatba rendezzük: n változó értékeit tehát egy rendezett n -es, azaz egy n hosszú vektor írja le.

2.2.1. Definíció. Legyen Ω egy megszámlálható eseménytér, ekkor egy $\underline{X} : \Omega \rightarrow \mathbb{R}^n$ függvényt *diszkrét valószínűségi vektorváltozónak* nevezünk.

Ebben a jegyzetben \mathbb{R}^n elemeit sorvektorokkal adjuk meg. Egy diszkrét $\underline{X} = (X_1, \dots, X_n)$ vektorváltozó koordinátái maguk is $\Omega \rightarrow \mathbb{R}$ függvények, tehát diszkrét valószínűségi változók. Megfordítva, ha adott n darab, ugyanazon eseménytéren értelmezett diszkrét valószínűségi változó, azokat egy n hosszú vektorba rendezve egy vektorváltozót kapunk.

Ahhoz hasonlóan, ahogy ez egyetlen változónál történt, egy diszkrét vektorváltozó viselkedése leírható azáltal, ha megadjuk, hogy egyes értékeket (vagyis egyes konkrét szám n -eseket) milyen valószínűséggel vesz fel. Az értékészletet, amely tehát ebben az esetben \mathbb{R}^n egy részhalmaza, itt is $\text{ran } \underline{X}$ jelöli.

2.2.2. Definíció. Ha $\underline{X} = (X_1, \dots, X_n) : \Omega \rightarrow \mathbb{R}^n$ egy valószínűségi mezőn értelmezett diszkrét valószínűségi vektorváltozó, akkor minden $\underline{t} = (t_1, \dots, t_n) \in \text{ran } \underline{X}$ értékre legyen

$$f_{\underline{X}}(\underline{t}) := \mathbb{P}(\underline{X} = \underline{t}) = \mathbb{P}(X_1 = t_1, \dots, X_n = t_n).$$

Az így kapott $f_{\underline{X}} : \text{ran } \underline{X} \rightarrow [0; 1]$ függvényt az \underline{X} vektorváltozó *súlyfüggvényének* nevezzük.

Ha adott egy $\underline{X} = (X_1, \dots, X_n)$ vektorváltozó súlyfüggvénye, akkor ezzel a korábbiakhoz hasonlóan kifejezhető bármely az adott változó értékeinek segítségével leírható esemény valószínűsége. Mivel itt voltaképp a koordináták, azaz több változó viselkedését írjuk le, ezért azt mondjuk, hogy a súlyfüggvény megadja az X_1, \dots, X_n változók *együttes eloszlását*.

Mindezt most két egyszerű valószínűségi változó esetén fogjuk illusztrálni, jelölje ezeket X és Y . Ekkor az együttes eloszlás kényelmesen megadható táblázatos formában. A táblázatot úgy konstruáljuk, hogy egyes oszlopai az egyik, míg a sorai a másik változó által felvett (véges sok) lehetséges értékekhez tartoznak, és az $s \in \text{ran } X$ oszlopának ill. a $t \in \text{ran } Y$ sorának metszetébe a $f_{X,Y}(s, t) = \mathbb{P}(X = s, Y = t)$ valószínűség kerül.

2.2.1. Példa. Legyenek X és Y olyan valószínűségi változók, melyekre $\text{ran } X = \{2, 3, 5\}$ és $\text{ran } Y = \{0, 1, 2\}$, továbbá az együttes eloszlásukat a következő táblázat adja meg:

$Y \backslash X$	2	3	5
0	0,05	0,15	0,1
1	0,1	0,2	0,1
2	0,05	0,2	0,05

Itt példának okáért a 2 értékhez tartozó oszlop és az 1 értékhez tartozó sor metszetében a $f_{X,Y}(2, 1) = \mathbb{P}(X = 2, Y = 1) = 0,1$ valószínűség szerepel, és a súlyfüggvény többi lehetséges értéke is hasonlóan olvasható le.

Számoljuk ki a $\mathbb{P}(X \leq 4, Y > 0)$ valószínűséget. Az előző feltétel pontosan akkor teljesül, ha X a 2 vagy 3, míg Y az 1 vagy 2 értékeket veszi fel, tehát éppen az $(2; 1)$, $(2; 2)$, $(3; 1)$ és $(3; 2)$ értékpárookra. Természetesen az $\{(X; Y) = (s; t)\}$ események páronként egymást kizáróak, hiszen a vektorváltozó egyszerre egyetlen értéket vesz fel, így tehát a keresett valószínűség négy ilyen esemény valószínűségének, azaz a súlyfüggvény négy értékének összege:

$$\begin{aligned} \mathbb{P}(X \leq 4, Y > 0) &= f_{X,Y}(2, 1) + f_{X,Y}(2, 2) + f_{X,Y}(3, 1) + f_{X,Y}(3, 2) \\ &= 0,1 + 0,05 + 0,2 + 0,2 = 0,55. \end{aligned}$$

Egy vektorváltozó súlyfüggvényét pontosan úgy jellemezhetjük, ahogy az az előző szakaszban egy valós értékű változó esetén történt. Ahogy a fenti példában is láttuk, az $\{X = s, Y = t\}$ események, ahol s végigfut az X értékkészletén, t pedig az Y értékkészletén, általában is páronként egymást kizáróak, és persze lefedik a teljes eseményteret, így teljes eseményrendszert alkotnak. Ugyanez igaz n változó esetén is az $\{X_1 = t_1, \dots, X_n = t_n\}$ eseményekre, ahol t_i végigfut az X_i értékkészletén. Ezért a súlyfüggvény lehetséges értékeinek (és így egy együttes eloszlást megadó táblázatban szereplő valószínűségeknek) az összege szükségszerűen a teljes eseménytér valószínűsége, azaz 1. (Ellenőrizzük ezt a fenti táblázatra!) Másrészt, nemnegatív számok egy táblázata, ahol az egyes elemek összege 1, mindig egy együttes eloszlást ad meg.

Marginális eloszlások

Ha több változó együttes eloszlása adott, akkor ebből könnyedén számolható az egyes változók eloszlása külön-külön. Ezeket az eloszlásokat az (X_1, \dots, X_n) vektorváltozó *marginális eloszlásainak* (vagy *peremeloszlásainak*) nevezzük.

A marginális eloszlás meghatározását először két változó esetén tárgyaljuk, bár az általános eset pusztán a jelölések tekintetében komplikáltabb. Legyen tehát adott az X és Y

változók együttes eloszlása. Az X marginális eloszlásához az $f_X(s) = \mathbb{P}(X = s)$ valószínűségeket kell meghatározni, ahol s végigfut az X értékkészletén. Mivel az $\{Y = t\}$ események, ahol t végigfut az Y értékkészletén, teljes eseményrendszert alkotnak, így hasonlóan ahhoz, ahogy ez az 1.2.4. tétel bizonyításában történt, az $\{X = s\}$ eseményt az Y értékei szerint páronként egymást kizáró részeseményekre oszthatjuk:

$$\{X = s\} = \bigcup_{t \in \text{ran } Y} \{X = s, Y = t\},$$

és így

$$f_X(s) = \mathbb{P}(X = s) = \sum_{t \in \text{ran } Y} \mathbb{P}(X = s, Y = t) = \sum_{t \in \text{ran } Y} f_{X,Y}(s, t).$$

Vegyük észre, hogy (egyszerű valószínűségi változók esetén) a jobb oldali összegben az együttes eloszlás táblázatának s -hez tartozó oszlopelemei szerepelnek. Hasonlóan, a $\mathbb{P}(Y = t)$ valószínűség meghatározásához X értékei szerint bonthatjuk páronként kizáró részekre az $\{Y = t\}$ eseményt, így a táblázat t -hez tartozó sorolemeit összeadva adódik a keresett valószínűség:

$$f_Y(t) = \mathbb{P}(Y = t) = \sum_{s \in \text{ran } X} f_{X,Y}(s, t).$$

A 2.2.1. példában szereplő változók marginális eloszlásai:

$$f_X(2) = \mathbb{P}(X = 2) = f_{X,Y}(2, 0) + f_{X,Y}(2, 1) + f_{X,Y}(2, 2) = 0,05 + 0,1 + 0,05 = 0,2$$

$$f_X(3) = \mathbb{P}(X = 3) = 0,15 + 0,2 + 0,2 = 0,55$$

$$f_X(5) = \mathbb{P}(X = 5) = 0,1 + 0,1 + 0,05 = 0,25$$

$$f_Y(0) = \mathbb{P}(Y = 0) = f_{X,Y}(2, 0) + f_{X,Y}(3, 0) + f_{X,Y}(5, 0) = 0,05 + 0,15 + 0,1 = 0,3$$

$$f_Y(1) = \mathbb{P}(Y = 1) = 0,1 + 0,2 + 0,1 = 0,4$$

$$f_Y(2) = \mathbb{P}(Y = 2) = 0,05 + 0,2 + 0,05 = 0,3.$$

Az általános esetben, tehát n darab X_1, \dots, X_n változó esetén a fenti eljárásn pusztán annyit kell változtatni, hogy a vizsgált eseményeket több másik változó értéke szerint bontjuk szét. Az X_i marginális eloszlásához most is az $f_{X_i}(t_i) = \mathbb{P}(X_i = t_i)$ valószínűségeket kell meghatározni, ahol t_i végigfut az X_i értékkészletén, és ezúttal az $\{X_i = t_i\}$ eseményt a többi $n - 1$ változó különböző értékei szerint bomlik páronként egymást kizáró részeseményekre:

$$\{X_i = t_i\} = \bigcup_{\substack{s_j \in \text{ran } X_j \\ j \neq i}} \{X_1 = s_1, \dots, X_{i-1} = s_{i-1}, X_i = t_i, X_{i+1} = s_{i+1}, \dots, X_n = s_n\},$$

és így

$$f_{X_i}(t_i) = \mathbb{P}(X_i = t_i) = \sum_{\substack{s_j \in \text{ran } X_j \\ j \neq i}} f_{\underline{X}}(s_1, \dots, s_{i-1}, t_i, s_{i+1}, \dots, s_n).$$

Diszkrét valószínűségi változók függetlensége

Több változó együttes eloszlásának ismerete elvileg lehetőséget ad arra, hogy a véletlen mennyiségek közötti különféle összefüggéseket felismerjük. A regresszióanalízis a statisztika azon ága, ami ilyen összefüggések felderítésével foglalkozik. Persze a statisztikában az eloszlásokra vonatkozóan csak becslések állnak rendelkezésünkre, de ezek helyességének megítélését éppen a mögöttes valószínűségi számítási eszköztár teszi lehetővé. Ennek a felépítésével majd csak jóval később foglalkozunk, de az első lépést már most megtesszük annak az esetnek a vizsgálatával, amikor a változók között egyáltalán nincs összefüggés.

Emlékeztetünk, hogy véletlen *események* függetlensége intuitívan azt jelenti, hogy egy vagy több esemény bekövetkezése egy másik esemény bekövetkezésének esélyeire semmilyen hatással nincs. Ugyanezen jelenség megfogalmazható *valószínűségi változók esetén* is: ezek függetlensége azt jelenti, hogy egy vagy több változó által felvett értékek nem befolyásolják egy másik változó értékeit. Ilyen lehet például két független kockadobás eredménye.

Azonban az talán nem meglepő, hogy itt a precíz (absztrakt) fogalom megalkotásánál nem elegendő változónként egy-egy eseményről beszélni. A 2.1.3. gyakorlatban láttuk, hogy ha két kockával dobva X jelöli a dobott számok összegét, Y pedig a szorzatukat, akkor az $\{X \text{ páros}\}$ és az $\{Y \leq 4\}$ események függetlenek, holott a két mennyiség nyilván nem független egymástól, hiszen például ha a szorzat és így szükségszerűen mindkét dobás értéke páratlan, akkor az összeg mindenképp páros kell legyen.

Több valószínűségi változót tehát akkor fogunk függetlennek nevezni, ha a segítségükkel kifejezhető események mind függetlenek egymástól. Ez a definíció azonban abból a szempontból kényelmetlen volna, hogy általában nagyon nehézkes az összes szóba jövő eseményre leellenőrizni a függetlenséget. Szerencsére belátható, hogy elegendő az események egy lényegesen korlátozott halmazára szorítkozni. Először az egyszerűség kedvéért két változó esetét tekintjük át:

2.2.3. Definíció. Az $X, Y : \Omega \rightarrow \mathbb{R}$ ugyanazon valószínűségi mezőn értelmezett diszkrét valószínűségi változókat *függetlennek* nevezünk, ha minden $s \in \text{ran } X$ és $t \in \text{ran } Y$ esetén az $\{X = s\}$ és $\{Y = t\}$ események függetlenek, azaz ha

$$f_{X,Y}(s, t) = \mathbb{P}(X = s, Y = t) = \mathbb{P}(X = s)\mathbb{P}(Y = t) = f_X(s)f_Y(t)$$

teljesül.

Igazolható, hogy amennyiben két változó független a fenti definíció értelmében, akkor bármely két, külön-külön az egyes változók segítségével kifejezhető esemény független egymástól. Tehát a fenti definíció éppen az általunk leírni kívánt fogalmat adja.

Vegyük észre, hogy a definíció feltétele a súlyfüggvények fogalma segítségével egészen tömören kifejezhető, hiszen az éppen az $f_{X,Y} = f_X \cdot f_Y$ egyenlet teljesülését jelenti. Ennek az egyenletnek minden lehetséges $s \in \text{ran } X$ ill. $t \in \text{ran } Y$ értéket behelyettesítve fenn kell állnia, ami egyben azt is jelenti, hogy ha találunk akár egyetlen értékpárt, amelyre az egyenlet nem teljesül, akkor ezzel belátjuk, hogy a változók összefüggők. Függetlenség esetén viszont ennek igazolása jóval hosszadalmasabb lehet, hiszen minden értékpár esetén igazolni kell az egyenlőséget.

2.2.2. Példa. Dobunk egy szabályos kockával, jelölje X a dobott szám kettővel vett osztási maradékát (tehát a paritását), Y pedig a hármas maradékát. Ekkor $\text{ran } X = \{0, 1\}$ és $\text{ran } Y = \{0, 1, 2\}$, továbbá X minden lehetséges értékét $1/2$, míg Y minden lehetséges értékét

1/3 eséllyel veszi fel. Viszont könnyen látható, hogy ha fixáljuk a dobott szám paritását és hármast maradékát is, az egyértelműen meghatározza azt. Így tehát

$$f_{X,Y}(s,t) = \mathbb{P}(X = s, Y = t) = \frac{1}{6} = \frac{1}{2} \cdot \frac{1}{3} = \mathbb{P}(X = s)\mathbb{P}(Y = t) = f_X(s)f_Y(t)$$

teljesül minden $s \in \text{ran } X$ és $t \in \text{ran } Y$ esetén, azaz X és Y függetlenek.

Két egyszerű valószínűségi változó együttes eloszlásának táblázatos megadása ezt a vizsgálatot viszonylag kényelmessé teszi. Ugyanis az egyik változó egy s értékéhez tartozó oszlopának, valamint a másik változó t értékéhez tartozó sorának metszetében a marginális eloszlások súlyfüggvényeinek szorzata kell álljon. Ezek a marginális eloszlások kényelmesen adminisztrálhatók az oszlopok alján vagy a sorok végén, így a fenti tulajdonság ellenőrzése egyszerűvé válik. Ezt a következő példán illusztráljuk.

2.2.3. Példa. Legyenek X és Y olyan valószínűségi változók, melyekre $\text{ran } X = \{0, 1, 2\}$ és $\text{ran } Y = \{0, 1\}$, továbbá az együttes eloszlásukat a következő táblázat adja meg:

$X \backslash Y$	0	1	2
0	1/12	1/8	1/8
1	1/6	1/4	1/4

Amint azt már korábban láttuk, az egyes oszlopokban illetve sorokban álló valószínűségeket összeadva a marginális eloszlások megfelelő súlyfüggvényértékei adódnak, amiket az adott oszlop aljára vagy az adott sor végére írhatunk. Például az első sorban lévő valószínűségek összege $\frac{1}{12} + \frac{1}{8} + \frac{1}{8} = \frac{1}{3}$, ami épp a $\mathbb{P}(Y = 0)$ valószínűség:

$X \backslash Y$	0	1	2	
0	1/12	1/8	1/8	1/3 = $\mathbb{P}(Y = 0) = f_Y(0)$
1	1/6	1/4	1/4	2/3 = $f_Y(1)$
	1/4 = $f_X(0)$	3/8 = $f_X(1)$	3/8 = $f_X(2)$	

Most már könnyen ellenőrizhető, hogy az egyes oszlopokhoz ill. sorokhoz tartozó f_X és f_Y értékek szorzata minden esetben a metszetben lévő szám, vagyis a két változó független.

2.2.1. Gyakorlat. Döntsük el, hogy a 2.2.1. példában megadott változók függetlenek-e.

Megjegyzés. Ha az olvasó jártas a lineáris algebrában, akkor az együttes eloszlás táblázatára egy mátrixként tekinthet. Legyen az X értékészlete $\text{ran } X = \{s_1, \dots, s_n\}$, míg az Y értékészlete $\text{ran } Y = \{t_1, \dots, t_m\}$. Legyen továbbá $\underline{u} = (f_X(s_1), \dots, f_X(s_n))$, és $\underline{v} = (f_Y(t_1), \dots, f_Y(t_m))$. A két (egyszerű) változó pontosan akkor független egymástól, ha az együttes eloszlás táblázatát az $\underline{v}^T \cdot \underline{u}$ szorzat adja, ahol \underline{v}^T az \underline{v} vektor transzponáltja. Ennek a mátrixnak a sorai ill. az oszlopai egymás skalárszorosai, így tehát egy 1 rangú mátrixot kapunk. Úgy fogalmazhatunk tehát, hogy két egyszerű változó pontosan akkor független, ha az együttes eloszlásuk táblázata egy 1 rangú mátrix. Jól látható az utolsó példában, hogy a táblázat második sora az első kétszerese, ebből pedig rögtön adódik, hogy a mátrix rangja 1, azaz a változók függetlenek.

Kettőnél több változó függetlenségének fogalma a két változó esetének egyszerű általánosításaként adódik. Ekkor azonban, ahogy kettőnél több esemény esetén is, páronkénti ill. együttes függetlenséget is vizsgálhatunk:

2.2.4. Definíció. Az $X_1, \dots, X_n : \Omega \rightarrow \mathbb{R}$ ugyanazon valószínűségi mezőn értelmezett diszkrét valószínűségi változókat *páronként* ill. *együttesen függetlennek* nevezzük, ha minden $t_1 \in \text{ran } X_1, \dots, t_n \in \text{ran } X_n$ esetén az $\{X_1 = t_1\}, \dots, \{X_n = t_n\}$ események páronként ill. együttesen függetlenek.

A következő állítás egy egyszerű példát mutat együttesen független változókra, amely egy későbbi alkalmazásban jó szolgálatot tesz majd számunkra:

2.2.1. Állítás. Legyenek $A_1, \dots, A_n \subset \Omega$ egy valószínűségi mező együttesen független eseményei. Ekkor a hozzájuk tartozó $\mathbb{1}_{A_1}, \dots, \mathbb{1}_{A_n}$ változók együttesen függetlenek.

Bizonyítás. Az állításban szereplő változók értékkészlete a $\{0, 1\}$ halmaz, tehát az összes olyan

$$\{\mathbb{1}_{A_1} = t_1\}, \dots, \{\mathbb{1}_{A_n} = t_n\}$$

eseményhalmazról, ahol minden $1 \leq i \leq n$ esetén $t_i = 0$ vagy 1 , be kell látnunk, hogy együttesen függetlenek. Vezessük be egy A esemény és $t \in \{0, 1\}$ esetén az

$$A^t = \begin{cases} A, & \text{ha } t = 1 \\ \bar{A}, & \text{ha } t = 0 \end{cases}$$

jelölést. Ekkor egy tetszőleges $\emptyset \neq I \subset \{1, \dots, n\}$ indexhalmazra

$$\mathbb{P} \left(\bigcap_{i \in I} \{\mathbb{1}_{A_i} = t_i\} \right) = \mathbb{P} \left(\bigcap_{i \in I} A_i^{t_i} \right).$$

Az 1.2.3. definíció után tett megjegyzések szerint ha az A_1, \dots, A_n események közül néhánynak a komplementerét vesszük, az így kapott események továbbra is együttesen függetlenek maradnak. Ez éppen azt jelenti, hogy az $A_1^{t_1}, \dots, A_n^{t_n}$ események együttesen függetlenek, és így az utóbbi valószínűség éppen

$$\prod_{i \in I} \mathbb{P}(A_i^{t_i}) = \prod_{i \in I} \mathbb{P}(\mathbb{1}_{A_i} = t_i),$$

és mivel ez minden $\emptyset \neq I$ indexhalmazra és minden $t_1, \dots, t_n \in \{0, 1\}$ szám n -esre igaz, így az $\mathbb{1}_{A_1}, \dots, \mathbb{1}_{A_n}$ változók együttesen függetlenek. \square

Konvolúció (kiegészítő anyag)

A szakasz hátralévő részében megvizsgáljuk, hogy hogyan számolható ki valószínűségi változók összegének, szorzatának ill. hányadosának az eloszlása abban a speciális esetben, amikor a szóban forgó változók függetlenek. Látni fogjuk, hogy ebben az esetben ha nem is egyszerűen, de szisztematikusan elvégezhetőek ezek a számolások az egyes változók eloszlásának ismeretében.

Legyen először $Z = X + Y$, ahol $X, Y : \Omega \rightarrow \mathbb{R}$ független valószínűségi változók. A Z értékkészlete azon $n \in \mathbb{R}$ számokból áll, amelyek előállnak két olyan szám összegeként, amelyek egyike X , másika pedig Y értékkészletében van. Egy ilyen n számra

$$\mathbb{P}(Z = n) = \mathbb{P} \left(\bigcup_{\substack{k \in \text{ran } X, l \in \text{ran } Y \\ k+l=n}} \{X = k, Y = l\} \right) = \sum_{\substack{k \in \text{ran } X, l \in \text{ran } Y \\ k+l=n}} \mathbb{P}(X = k, Y = l),$$

mivel az unióban szereplő események páronként egymást kizáróak. Most a változók függetlenségét felhasználva

$$\begin{aligned}\mathbb{P}(Z = n) &= \sum_{\substack{k \in \text{ran } X, l \in \text{ran } Y \\ k+l=n}} \mathbb{P}(X = k, Y = l) = \sum_{\substack{k \in \text{ran } X, l \in \text{ran } Y \\ k+l=n}} \mathbb{P}(X = k)\mathbb{P}(Y = l) \\ &= \sum_{k \in \text{ran } X} \mathbb{P}(X = k)\mathbb{P}(Y = n - k) = \sum_{k \in \text{ran } X} f_X(k)f_Y(n - k) = \sum_{l \in \text{ran } Y} f_X(n - l)f_Y(l).\end{aligned}$$

Vegyük észre, hogy az utóbbi összefüggés akkor is érvényes, amennyiben az n szám nincs az $X + Y$ értékkészletében, hiszen ekkor minden $k \in \text{ran } X$ esetén szükségképp $n - k \notin \text{ran } Y$ teljesül (ill. $l \in \text{ran } Y$ esetén $n - l \notin \text{ran } X$), így az utóbbi összeg értéke egyszerűen 0, ami ebben az esetben megegyezik a $\mathbb{P}(Z = n)$ valószínűséggel. A következőt kaptuk tehát:

2.2.2. Állítás. *Legyenek $X, Y : \Omega \rightarrow \mathbb{R}$ független valószínűségi változók f_X ill. f_Y súlyfüggvénnyel, ekkor az $X + Y$ változó súlyfüggvénye*

$$f_{X+Y}(t) = (f_X * f_Y)(t),$$

ahol az $f_X * f_Y$ függvényt a

$$(f_X * f_Y)(t) = \sum_{s \in \text{ran } X} f_X(s)f_Y(t - s) = \sum_{s \in \text{ran } Y} f_X(t - s)f_Y(s)$$

formulák definiálják minden $t \in \mathbb{R}$ esetén. Az $f_X * f_Y$ függvényt az f_X és f_Y függvények konvolúciójának nevezzük.

2.2.4. Példa. Legyenek $X, Y \sim \text{Geo}(p)$ független valószínűségi változók. Lehet például X egy p valószínűségű esemény bekövetkezéséig végzett kísérletek száma, amely kísérletet az első bekövetkezés után újra ismételni kezdünk az adott esemény második bekövetkezéséig, és az első siker után végzett kísérletek számát jelöli Y . (Bizonyítsuk be, hogy így valóban független, azonos paraméterű geometriai eloszlású változókat kapunk!) Ekkor $X + Y$ értékkészlete az 1-nél nagyobb egészek halmaza (ez a konkrét példában az összes elvégzett kísérlet számát jelenti). A fenti állítás szerint pedig tetszőleges $n > 1$ egész esetén

$$\begin{aligned}f_{X+Y}(n) &= \mathbb{P}(X + Y = n) = \sum_{k=1}^{n-1} \mathbb{P}(X = k)\mathbb{P}(Y = n - k) \\ &= \sum_{k=1}^{n-1} (1 - p)^{k-1}p(1 - p)^{n-k-1}p = \sum_{k=1}^{n-1} (1 - p)^{n-2}p^2 = (n - 1)(1 - p)^{n-2}p^2.\end{aligned}$$

Ez az eloszlás szoros kapcsolatban áll egy ún. negatív binomiális eloszlással, amelyről részletesen a 2.6.2. szakaszban lesz szó.

2.2.5. Példa. Legyenek $X \sim \text{Bin}(n; p)$ és $Y \sim \text{Bin}(m; p)$ független, binomiális eloszlású valószínűségi változók, amelyek eloszlásának második paramétere megegyezik. Meghatározzuk $X + Y$ eloszlását, először valójában a 2.2.2. állítás használatával nélkül. Ehhez ugyanazt a trükköt fogjuk használni, amivel már a 2.1.2. állítás bizonyításában is találkoztunk: mivel az összeg eloszlása csak az egyes változók eloszlásától és nem maguktól a változóktól függ, így elég az eloszlást egyetlen konkrét példa esetén kiszámolni.

Tekintsünk tehát egy $n + m$ kísérletből álló független kísérletsorozatot, ahol egy p valószínűségű esemény bekövetkezéseinek számát figyeljük. Jelöljük X ill. Y a bekövetkezések számát az első n ill. az utolsó m kísérlet során. Világos, hogy ekkor $X \sim \text{Bin}(n; p)$ és $Y \sim \text{Bin}(m; p)$, és a két változó független egymástól. Továbbá az $X + Y$ változó értéke az esemény bekövetkezéseinek száma az összes kísérlet során, tehát $X + Y \sim \text{Bin}(n + m; p)$.

Bár az eredményünk levezetéséhez nem volt szükségünk a 2.2.2. állításra, érdemes megfontolni, hogy mi adódik az állítás használatával. Legyenek tehát ismét $X \sim \text{Bin}(n; p)$ és $Y \sim \text{Bin}(m; p)$ függetlenek, ekkor a 2.2.2. állítás szerint minden $0 \leq k \leq n + m$ esetén

$$\begin{aligned} f_{X+Y}(k) &= \sum_{i=0}^n f_X(i) f_Y(k-i) = \sum_{i=0}^n \binom{n}{i} p^i (1-p)^{n-i} \cdot \binom{m}{k-i} p^{k-i} (1-p)^{m-k+i} \\ &= \sum_{i=0}^n \binom{n}{i} \binom{m}{k-i} p^k (1-p)^{n+m-k} = p^k (1-p)^{n+m-k} \cdot \left(\sum_{i=0}^n \binom{n}{i} \binom{m}{k-i} \right). \end{aligned}$$

Másrészről a fentiek szerint ez az érték a $\text{Bin}(n + m; p)$ eloszlás súlyfüggvényének értéke a k helyen, azaz

$$\binom{n+m}{k} p^k (1-p)^{n+m-k} = p^k (1-p)^{n+m-k} \cdot \left(\sum_{i=0}^n \binom{n}{i} \binom{m}{k-i} \right)$$

teljesül minden $0 \leq k \leq n + m$ esetén. Ha most p -t egy 0-tól és 1-től különböző paraméternek választjuk, akkor a $p^k (1-p)^{n+m-k}$ (0-tól különböző) számmal leosztva az

$$(2.13) \quad \binom{n+m}{k} = \sum_{i=0}^n \binom{n}{i} \binom{m}{k-i}$$

egyenlőséget nyerjük. Ez a formula tehát érvényes tetszőleges pozitív n, m és $0 \leq k \leq n + m$ esetén. Persze az nincs kizárva, hogy bizonyos tagok értéke a jobb oldalon szereplő összegben 0 legyen, nevezetesen, ha $k - i < 0$ vagy $k - i > m$ teljesül.

Figyeljük meg, hogy bár a (2.13) egyenlőség egy kombinatorikai állítás, de - ellentétben pl. az (1.3) ill. a (2.8) formulák bizonyításaival - az ide vezető valószínűségszámítási levezetésből nem (feltétlenül) látszik rögtön, hogy hogy lehetne azt egy kombinatorikai (leszámlálási) érveléssé átalakítani. Tulajdonképpen a fenti bizonyítás lényeges elemeit kiszűrve (a valószínűségszámítás fogalmaitól megszabadulva) egyszerűen beláthatjuk az állítást a binomiális tétel felhasználásával. Van azonban egy egyszerű kombinatorikai érvelés is, amely abból indul ki, hogy $\binom{n}{k}$ nem más, mint egy n elemű halmaz k elemű részhalmazainak száma. Arra biztatjuk az olvasót, hogy próbálja meg ezeket az alternatív levezetéseket maga megtalálni.

Végül a teljesség kedvéért megadjuk a 2.2.2. állításnak a független változók szorzatára illetve hányadosára vonatkozó analogonját. Az alábbi formulák a 2.2.2. állítás bizonyításához hasonlóan igazolhatók, ezért a részleteket elhagyjuk.

2.2.3. Állítás. *Legyenek $X, Y : \Omega \rightarrow \mathbb{R}$ független valószínűségi változók, melyeknek súlyfüggvénye f_X ill. f_Y . Ekkor $0 \neq t \in \mathbb{R}$ esetén*

$$f_{XY}(t) = \sum_{s \in \text{ran } X \setminus \{0\}} f_X(s) f_Y(t/s) = \sum_{s \in \text{ran } Y \setminus \{0\}} f_X(t/s) f_Y(s),$$

ahol f_{XY} az XY szorzat súlyfüggvénye. Továbbá, ha Y nem vesz fel 0 értéket, akkor minden $t \in \mathbb{R}$ esetén

$$f_{\frac{X}{Y}}(t) = \sum_{s \in \text{ran } Y} f_X(ts) f_Y(s),$$

ahol $f_{\frac{X}{Y}}$ az X/Y hányados súlyfüggvénye.

Gyakorlatok, feladatok

2.2.2. Gyakorlat. Kétszer dobunk egy szabályos dobókockával. Jelölje X a hatosok, Y pedig a páros eredmények számát. Adjuk meg X és Y együttes eloszlását. Függetlenek-e az X és Y változók?

2.2.3. Gyakorlat. Az X és Y valószínűségi változók együttes eloszlását tartalmazza az alábbi táblázat. Határozzuk meg a p értékét. Mennyi a $\mathbb{P}(X \leq 0, Y = 1)$ valószínűség? Független-e X és Y ?

$Y \backslash X$	-1	0	1
-1	p	3p	6p
1	5p	15p	30p

2.2.4. Gyakorlat. Az X és Y valószínűségi változók együttes eloszlását tartalmazza az alábbi táblázat.

- Függetlenek-e az $\{X = 2\}$ és $\{Y = 2\}$ események?
- Függetlenek-e az $\{X < 2\}$ és $\{Y < 2\}$ események?
- Függetlenek-e az X és Y változók?

$Y \backslash X$	0	1	2
0	1/10	1/10	1/10
1	1/10	1/10	3/10
2	1/20	1/20	1/10

2.2.5. Feladat. Legyenek $X \sim \text{Geo}(p)$ és $Y \sim \text{Geo}(q)$ független valószínűségi változók, ahol $p \neq q$. Határozzuk meg az $X + Y$ változó eloszlását.

2.3. A várható érték

Egy valószínűségi változó értéke egy véletlen kimenetel függvénye, de ezt a véletlenszerű viselkedést jellemezhetjük különböző módokon. Az eloszlás megadása persze egy teljes jellemzést ad, de gyakran elegendő annak néhány egyszerűbb paraméterét vizsgálni. Szakaszunk egyik fő motiváló példájában is ilyen szituációval találkozunk: tekintsünk egy olyan programot, amelynek a futása a véletlentől függ úgy, hogy a véletlen a program kimenetelének helyességét nem, de a futásidejét befolyásolja. Mikor tekinthetünk egy ilyen programot a gyakorlat szempontjából megfelelőnek? Nyilván elvárjuk, hogy még ha elvben tud is rossz futásidőt produkálni, azt csak nagyon ritkán, elhanyagolhatóan kis valószínűséggel tegye, vagyis többnyire, azaz *átlagosan* gyorsan lefusson. A szakasz végén részletesen elemezzük a Randomized Quick Sort rendezőalgoritmust, ami éppen a fent leírt tulajdonsággal bír. Mivel

az algoritmus futásidejét leíró lépésszám egy véletlen mennyiség, azt egy valószínűségi változóval fogjuk modellezni. Az "átlagos lépésszámot" pedig ennek a változónak a *várható értéke* fogja leírni, amit az alábbiakban definiálunk.

2.3.1. A várható érték definíciója

Egy változó átlagos értékére egy olyan számként gondolhatunk, ami körül sok ugyanolyan eloszlású változót "kiértékelve" kapott véletlen számsorozat átlaga "ingadozik". Az előző mondatnak később precíz értelmet fogunk majd adni. Minden esetre talán nem meglepő, hogy ez az "átlagos érték" valójában a változó lehetséges értékeinek a hozzájuk tartozó valószínűségekkel, azaz a súlyfüggvény értékeivel súlyozott átlaga lesz. Mindezt formálisan a következő definícióban fogalmazzuk meg:

2.3.1. Definíció. Legyen X egy egyszerű valószínűségi változó. Ekkor az

$$\mathbb{E}(X) = \sum_{t \in \text{ran } X} t \cdot f_X(t) = \sum_{t \in \text{ran } X} t \cdot \mathbb{P}(X = t)$$

kifejezést az X *várható értékének* nevezzük.

Vegyük észre, hogy mivel a fenti mennyiséget egyszerű valószínűségi változóra definiáltuk, így annak értékészlete és ezzel együtt a fenti összeg is véges. Az is világos továbbá, hogy a várható érték pusztán az X valószínűségi változó eloszlásától, azaz az $f_X(t)$ értékektől függ, nem konkrétan az X változótól.

2.3.1. Példa. Számoljuk ki néhány egyszerű példára a várható értéket:

- Legyen X egy kockadobás eredménye. Ekkor

$$\begin{aligned} \mathbb{E}(X) &= 1 \cdot f_X(1) + 2 \cdot f_X(2) + 3 \cdot f_X(3) + 4 \cdot f_X(4) + 5 \cdot f_X(5) + 6 \cdot f_X(6) \\ &= \frac{1}{6} \cdot (1 + 2 + 3 + 4 + 5 + 6) = \frac{21}{6} = \frac{7}{2} = 3,5. \end{aligned}$$

- Legyen $Y = X^2$, azaz egy kockadobás eredményének négyzete. A 2.1.5. példában meghatároztuk az Y változó eloszlását. Ennek alapján az Y várható értéke:

$$\begin{aligned} \mathbb{E}(Y) &= 1 \cdot f_Y(1) + 4 \cdot f_Y(4) + 9 \cdot f_Y(9) + 16 \cdot f_Y(16) + 25 \cdot f_Y(25) + 36 \cdot f_Y(36) \\ &= \frac{1}{6} \cdot (1 + 4 + 9 + 16 + 25 + 36) = \frac{91}{6}. \end{aligned}$$

- Legyen A egy esemény, melyre $\mathbb{P}(A) = p$. Ekkor az A -hoz tartozó $\mathbb{1}_A$ indikátor valószínűségi változó várható értéke:

$$\mathbb{E}(\mathbb{1}_A) = 0 \cdot \mathbb{P}(\mathbb{1}_A = 0) + 1 \cdot \mathbb{P}(\mathbb{1}_A = 1) = \mathbb{P}(A) = p.$$

A várható érték kiszámolásánál gyakran sokat egyszerűsít a számolásokon, ha az adott változót összeg alakban tudjuk felírni, ugyanis a várható érték ezekben az esetekben a különösen szépen viselkedik:

2.3.1. Tétel (A várható érték linearitása). *Legyenek $X, Y : \Omega \rightarrow \mathbb{R}$ ugyanazon valószínűségi mezőn értelmezett egyszerű valószínűségi változók, $c \in \mathbb{R}$ pedig egy tetszőleges valós szám. Ekkor*

$$\mathbb{E}(X + Y) = \mathbb{E}(X) + \mathbb{E}(Y), \quad \mathbb{E}(cX) = c \cdot \mathbb{E}(X).$$

Ha az olvasó ismeri a vektorterek elméletét, akkor a fenti tulajdonság elnevezése minden bizonnyal ismerősen cseng. (Amennyiben viszont az olvasó nem találkozott még vektorterekkel, akkor ezt a bekezdést nyugodtan ugorja át.) Vegyük észre, hogy egy adott Ω eseménytéren értelmezett egyszerű valószínűségi változók a pontonkénti összeadásra és a valós számokkal mint skalárokkal való (pontonkénti) szorzásra nézve vektorteret alkotnak. A várható érték pedig a fenti tétel értelmében egy ezen vektortéren értelmezett és a valós számok halmazába képező lineáris leképezés, más néven egy lineáris funkcionál.

Mint említettük, a fenti tétel alkalmazása sok esetben megkönnyíti a dolgunkat. Ez egyben azt is jelenti, hogy némi technikai nehézség magának a tételnek a bizonyításában van elrejtve:

Bizonyítás. A második állítás bizonyításával kezdünk. Tegyük fel először, hogy $c \in \mathbb{R} \setminus \{0\}$ egy nullától különböző valós szám. Legyen az X véges értékészlete $\text{ran } X = \{t_1, \dots, t_n\}$, ekkor a cX változó értékészlete $\text{ran } cX = \{ct_1, \dots, ct_n\}$. Tehát

$$\mathbb{E}(cX) = \sum_{k=1}^n ct_k \cdot \mathbb{P}(cX = ct_k) = c \cdot \sum_{k=1}^n t_k \cdot \mathbb{P}(X = t_k) = c \cdot \mathbb{E}(X).$$

Ha viszont $c = 0$, akkor cX a konstans 0 változó, tehát $\text{ran } cX = \{0\}$, és ezt az értéket cX 1 valószínűséggel veszi fel, így a várható érték definíciója szerint

$$\mathbb{E}(cX) = 0 = 0 \cdot \mathbb{E}(X) = c \cdot \mathbb{E}(X).$$

Térjünk rá most első állítás bizonyítására. A definícióból indulunk ki:

$$\mathbb{E}(X + Y) = \sum_{m \in \text{ran } X+Y} m \cdot \mathbb{P}(X + Y = m) = \sum_{m \in \text{ran } X+Y} m \cdot \mathbb{P}(\cup_{s+t=m} \{X = s, Y = t\}).$$

Itt a jobb oldalon egymást kizáró események uniója áll, így az unió valószínűsége a valószínűségek összege lesz:

$$\sum_{m \in \text{ran } X+Y} \sum_{s+t=m} (s + t) \cdot \mathbb{P}(X = s, Y = t).$$

Vegyük észre, hogy itt minden (s, t) pár, ahol $s \in \text{ran } X$ és $t \in \text{ran } Y$, pontosan egyszer szerepel, hiszen az m összeg és az összeg egyik tagja már meghatározza a másik tagot. Ez az összeg tehát nem más, mint

$$\begin{aligned} & \sum_{s \in \text{ran } X} \sum_{t \in \text{ran } Y} (s + t) \cdot \mathbb{P}(X = s, Y = t) = \\ & = \sum_{s \in \text{ran } X} \sum_{t \in \text{ran } Y} s \cdot \mathbb{P}(X = s, Y = t) + \sum_{s \in \text{ran } X} \sum_{t \in \text{ran } Y} t \cdot \mathbb{P}(X = s, Y = t) \\ & = \sum_{s \in \text{ran } X} s \sum_{t \in \text{ran } Y} \mathbb{P}(X = s, Y = t) + \sum_{t \in \text{ran } Y} t \sum_{s \in \text{ran } X} \mathbb{P}(X = s, Y = t). \end{aligned}$$

Ahogy azt már a marginális eloszlások kiszámításánál is láttuk, a $\mathbb{P}(X = s, Y = t)$ valószínűségek összege az fenti első kettős összeg belső összegében éppen $\mathbb{P}(X = s)$, és hasonlóképp, a második kettős összeg belső összege $\mathbb{P}(Y = t)$. Így tehát végeredményben

$$\mathbb{E}(X + Y) = \sum_{s \in \text{ran } X} s \cdot \mathbb{P}(X = s) + \sum_{t \in \text{ran } Y} t \cdot \mathbb{P}(Y = t) = \mathbb{E}(X) + \mathbb{E}(Y).$$

□

2.3.2. Példa. Dobjunk kétszer egy kockával, legyen X az első, Y pedig a második dobás eredménye. Ekkor a dobott számok összege $X + Y$, ennek várható értéke fenti tétel szerint

$$\mathbb{E}(X + Y) = \mathbb{E}(X) + \mathbb{E}(Y) = 2 \cdot 3,5 = 7.$$

Ezt persze kiszámolhatjuk a várható érték definíciója alapján is. A 2.1.9. példában meghatároztuk az $X + Y$ változó eloszlását, ez alapján

$$\begin{aligned} \mathbb{E}(X + Y) &= 2 \cdot \frac{1}{36} + 3 \cdot \frac{2}{36} + 4 \cdot \frac{3}{36} + 5 \cdot \frac{4}{36} + 6 \cdot \frac{5}{36} + 7 \cdot \frac{6}{36} \\ &\quad + 8 \cdot \frac{5}{36} + 9 \cdot \frac{4}{36} + 10 \cdot \frac{3}{36} + 11 \cdot \frac{2}{36} + 12 \cdot \frac{1}{36} \\ &= \frac{2 + 6 + 12 + 20 + 30 + 42 + 40 + 36 + 30 + 22 + 12}{36} = \frac{252}{36} = 7. \end{aligned}$$

2.3.3. Példa. Kiszámoljuk a binomiális eloszlású valószínűségi változók várható értékét. A várható érték definíciója után már megjegyeztük, hogy az pusztán a súlyfüggvény értékeitől függ, tehát csak az eloszlástól, nem pedig konkrétan a változótól. Már korábban is láttunk példát arra, hogy ilyen esetben egy konkrét változóra elegendő a számolást elvégezni, és az eredmény általában is érvényes lesz.

Az $n \in \mathbb{N}^+$ és $p \in [0; 1]$ paraméterű binomiális eloszlás várható értékéhez tekintsük a 2.1.10. példában látott $X \sim \text{Bin}(n; p)$ változót, aminek értéke egy n független kísérletből álló kísérletsorozatban egy adott p valószínűségű esemény bekövetkezéseinek száma. A példában megmutattuk, hogy

$$X = \mathbf{1}_{A_1} + \cdots + \mathbf{1}_{A_n},$$

ahol A_i az az esemény, hogy az i -edik kísérletnél bekövetkezik a vizsgált esemény. Mivel $\mathbb{P}(A_i) = p$ teljesül minden $1 \leq i \leq n$ esetén, így a várható érték linearitása, valamint az indikátor várható értéke szerint az n , p paraméterű binomiális eloszlású változók várható értéke

$$\mathbb{E}(X) = \mathbb{E}(\mathbf{1}_{A_1} + \cdots + \mathbf{1}_{A_n}) = \mathbb{E}(\mathbf{1}_{A_1}) + \cdots + \mathbb{E}(\mathbf{1}_{A_n}) = np.$$

Hogy még jobban érzékelhető legyen a fenti módszer egyszerűsége, levezetjük ezt az eredményt pusztán az $f_X(k)$ értékek és a definíció felhasználásával. A várható érték definíciója szerint

$$\mathbb{E}(X) = \sum_{k=0}^n k \binom{n}{k} p^k (1-p)^{n-k} = \sum_{k=1}^n k \binom{n}{k} p^k (1-p)^{n-k}.$$

Itt $1 \leq k \leq n$ esetén

$$k \binom{n}{k} = k \cdot \frac{n!}{k!(n-k)!} = n \cdot \frac{(n-1)!}{(k-1)!(n-1-(k-1))!} = n \binom{n-1}{k-1},$$

tehát

$$\mathbb{E}(X) = \sum_{k=1}^n n \binom{n-1}{k-1} p^k (1-p)^{n-1-(k-1)} = np \sum_{j=0}^{n-1} \binom{n-1}{j} p^j (1-p)^{n-1-j}.$$

Az utolsó szumma a binomiális tétel szerint $(p + (1-p))^{n-1} = 1^{n-1} = 1$, így tehát a várható érték np .

2.3.4. Példa. Az indikátorváltozók egy további alkalmazásaként bebizonyítjuk a 3 halmazra kimondott szita-formulát. Legyenek A, B és C egy valószínűségi mező tetszőleges eseményei. A három esemény uniójának valószínűsége a de Morgan-azonosságok segítségével felírható a következőképp:

$$(2.14) \quad \mathbb{P}(A \cup B \cup C) = 1 - \mathbb{P}(\overline{A \cup B \cup C}) = 1 - \mathbb{P}(\overline{A} \cap \overline{B} \cap \overline{C}).$$

Vegyük észre, hogy az $\overline{A} \cap \overline{B} \cap \overline{C}$ indikátorváltozója éppen az $\mathbb{1}_{\overline{A}} \cdot \mathbb{1}_{\overline{B}} \cdot \mathbb{1}_{\overline{C}}$, hiszen ennek értéke pontosan akkor 1, ha mindhárom tényező 1, azaz, ha mindhárom esemény bekövetkezik, különben pedig 0 lesz az érték. Mivel egy esemény valószínűsége az hozzá tartozó indikátorváltozó várható értéke, így

$$\mathbb{P}(\overline{A} \cap \overline{B} \cap \overline{C}) = \mathbb{E}(\mathbb{1}_{\overline{A}} \cdot \mathbb{1}_{\overline{B}} \cdot \mathbb{1}_{\overline{C}}).$$

Az A, B, C események komplementereinek indikátorai könnyen kifejezhetők maguknak az eseményeknek az indikátoraival:

$$\mathbb{1}_{\overline{A}} \cdot \mathbb{1}_{\overline{B}} \cdot \mathbb{1}_{\overline{C}} = (1 - \mathbb{1}_A)(1 - \mathbb{1}_B)(1 - \mathbb{1}_C).$$

A zárójeleket felbontva, továbbá ismét felhasználva, hogy indikátorok szorzata a megfelelő események metszetének az indikátora, a következőt kapjuk:

$$\begin{aligned} \mathbb{1}_{\overline{A}} \cdot \mathbb{1}_{\overline{B}} \cdot \mathbb{1}_{\overline{C}} &= 1 - \mathbb{1}_A - \mathbb{1}_B - \mathbb{1}_C + \mathbb{1}_A \cdot \mathbb{1}_B + \mathbb{1}_B \cdot \mathbb{1}_C + \mathbb{1}_A \cdot \mathbb{1}_C - \mathbb{1}_A \cdot \mathbb{1}_B \cdot \mathbb{1}_C \\ &= 1 - \mathbb{1}_A - \mathbb{1}_B - \mathbb{1}_C + \mathbb{1}_{A \cap B} + \mathbb{1}_{B \cap C} + \mathbb{1}_{A \cap C} - \mathbb{1}_{A \cap B \cap C}. \end{aligned}$$

Mindkét oldal várható értéket véve, felhasználva az indikátor várható értékét és a linearitást

$$\mathbb{P}(\overline{A} \cap \overline{B} \cap \overline{C}) = 1 - \mathbb{P}(A) - \mathbb{P}(B) - \mathbb{P}(C) + \mathbb{P}(A \cap B) + \mathbb{P}(B \cap C) + \mathbb{P}(A \cap C) - \mathbb{P}(A \cap B \cap C)$$

adódik. Végül visszahelyettesítve a (2.14) egyenletbe éppen a szita-formula állítását kapjuk.

Egy geometriai eloszlású valószínűségi változó várható értékét még nem értelmeztük, hiszen annak értékészlete megszámlálhatóan végtelen. Ennek az esetnek a kezelése technikailag valamivel nehezebb az eddigieknél. Az alábbiakban röviden részletezzük a fellépő problémák természetét, amikkel - mint látni fogjuk - a helyes definíciót használva valójában nem kell foglalkoznunk.

Egy megszámlálhatóan végtelen értékészlettel rendelkező diszkrét valószínűségi változó várható értéke formálisan ugyanúgy definiálható a

$$(2.15) \quad \sum_{t \in \text{ran } X} t \cdot f_X(t)$$

összeg segítségével. A probléma az, hogy ez az összeg nem feltétlenül konvergens. Egyáltalán ahhoz, hogy részletösszegeket és határértéket definiálhassunk, szükségünk van az értékészlet egy felsorolására: $\text{ran } X = \{t_1, \dots, t_n, \dots\}$. Ezután tekinthetjük a

$$\sum_{k=1}^{\infty} t_k \cdot f_X(t_k) = \lim_{N \rightarrow \infty} \sum_{k=1}^N t_k \cdot f_X(t_k)$$

határértéket, amennyiben az létezik. Azonban előfordulhat, hogy a jobb oldalon lévő határérték végtelen, és persze egy valószínűségi változó negatív értékeket is felvehet, így még az is megeshet, hogy a fenti részösszegek sem egy számhoz, sem a $\pm\infty$ -hez nem konvergálnak. Mint ahogy az is megtörténhet, hogy ugyan konvergens sort kapunk, de a határérték függ attól, hogy az elemeket konkrétan hogyan soroltuk fel.

Ahhoz, hogy ilyen problémákkal ne szembesüljünk, azt fogjuk megkövetelni, hogy a (2.15) sor abszolút konvergens legyen, azaz tagok abszolút értékeinek összege is konvergáljon (és ebben az esetben mindegy, hogy ezeket a nemnegatív tagokat milyen sorrendben összegezzük). Ebből szerencsére már következik, hogy az abszolút érték nélkül vett határérték is létezik és nem függ az elemek felsorolásától, így a (2.15) formula is értelmet nyer. Egy-egy ellenpéldán kívül ebben a jegyzetben kizárólag olyan esetekkel fogunk foglalkozni, ahol a (2.15) sor abszolút konvergens, és így létezik a várható érték a következő definíció szerinti értelemben:

2.3.2. Definíció. Legyen X egy diszkrét valószínűségi változó, amelynek értékészlete megszámlálhatóan végtelen, és amelyre

$$\sum_{t \in \text{ran } X} |t| \cdot f_X(t) < \infty$$

teljesül. Ekkor azt mondjuk, hogy *véges az X várható értéke*, amelyet az

$$(2.16) \quad \mathbb{E}(X) = \sum_{t \in \text{ran } X} t \cdot f_X(t)$$

összeg definiál.

A várható érték linearitása ebben az esetben is érvényes annyi megkötéssel, hogy fel kell tenni, hogy az egyes változók várható értéke véges (azaz a fenti sor abszolút konvergens). Természetesen egy egyszerű valószínűségi változó várható értéke mindig egy véges szám, így tehát egy diszkrét valószínűségi változó várható értékének végelessége alatt a továbbiakban azt értjük, hogy a változó vagy egyszerű, vagy pedig a (2.16)-ban szereplő végtelen sor abszolút konvergens. A két esetet tehát együtt kezeljük a következő tételben, amelynek bizonyítása lényegében megegyezik az egyszerű változókra kimondott tétel bizonyításával, az abban szereplő átalakítások érvényessége most az abszolút konvergencia következménye. A pontos részleteket elhagyjuk.

2.3.2. Tétel. *Legyenek $X, Y : \Omega \rightarrow \mathbb{R}$ ugyanazon valószínűségi mezőn értelmezett diszkrét valószínűségi változók, melyeknek véges a várható értékük. Legyen továbbá $c \in \mathbb{R}$ egy tetszőleges valós szám. Ekkor*

$$\mathbb{E}(X + Y) = \mathbb{E}(X) + \mathbb{E}(Y), \quad \mathbb{E}(cX) = c \cdot \mathbb{E}(X).$$

Megjegyezzük még, hogy amennyiben a (2.16) sor nem abszolút konvergens, akkor általában nem tudjuk definiálni a várható értéket. Abban az esetben azonban, ha a változó csak nemnegatív értékeket vehet fel, akkor a sor részletösszegei monoton nőnek, így azok egy divergens sor esetén a végtelenhez divergálnak. Ekkor azt mondjuk, hogy a várható érték végtelen ($\mathbb{E}(X) = \infty$).

2.3.5. Példa. Meghatározzuk a geometriai eloszlás várható értékét. Legyen $X \sim Geo(p)$, ahol $p \in (0; 1)$, ekkor

$$(2.17) \quad \mathbb{E}(X) = \sum_{k=1}^{\infty} k \cdot f_X(k) = \sum_{k=1}^{\infty} k \cdot (1-p)^{k-1}p,$$

feltéve persze, hogy ez a sor konvergens. A konvergencia sokféleképp belátható, következik például a hányadoskritériumból. A sor egymást követő tagjainak hányadosa

$$\frac{(k+1)(1-p)^k p}{k(1-p)^{k-1}p} = \frac{k+1}{k} \cdot (1-p).$$

Ha k -val tartunk a végtelenbe, akkor a fenti kifejezés határértéke $1-p < 1$, és így a sor valóban konvergens.

Az érték meghatározásához írjuk a (2.17) jobb oldalán lévő összeget a következőképp:

$$p + \sum_{k=2}^{\infty} k \cdot (1-p)^{k-1}p = p + \sum_{i=1}^{\infty} (i+1) \cdot (1-p)^i p = p + \sum_{i=1}^{\infty} i(1-p)^i p + \sum_{i=1}^{\infty} (1-p)^i p.$$

A jobb oldalon álló utolsó szumma értékét könnyedén meghatározhatjuk, hiszen $(1-p)$ -t kiemelve, éppen az X súlyfüggvényének értékei kerülnek az összegbe, ezek összege pedig 1:

$$\sum_{i=1}^{\infty} (1-p)^i p = (1-p) \sum_{i=1}^{\infty} (1-p)^{i-1} p = (1-p) \sum_{i=1}^{\infty} f_X(i) = (1-p) \cdot 1 = (1-p).$$

A másik szummából szintén kiemelhetünk $(1-p)$ -t, és így ismét egy ismerős kifejezést kapunk:

$$\sum_{i=1}^{\infty} i(1-p)^i p = (1-p) \sum_{i=1}^{\infty} i(1-p)^{i-1} p = (1-p) \cdot \mathbb{E}(X).$$

Mindezt felhasználva $\mathbb{E}(X) = p + (1-p)\mathbb{E}(X) + (1-p) = 1 + (1-p)\mathbb{E}(X)$, átrendezés után pedig

$$\mathbb{E}(X) = \frac{1}{p}$$

adódik.

A szakasz zárásaként egy olyan eredményt mutatunk be, ami akkor is hasznos információt szolgáltat, ha egy változó eloszlását nem ismerjük pontosan, de a várható érték ismert. Ennek segítségével ugyanis becsléseket kaphatunk bizonyos valószínűségekre:

2.3.3. Tétel (Markov-egyenlőtlenség). *Legyen $X \geq 0$ egy nemnegatív (diszkrét) valószínűségi változó és $a > 0$. Ekkor*

$$(2.18) \quad \mathbb{P}(X \geq a) \leq \frac{\mathbb{E}(X)}{a}.$$

Megjegyezzük, hogy (2.18) csak akkor ad nemtriviális becslést, ha $a > \mathbb{E}(X)$.

Bizonyítás. Ha $\mathbb{E}(X) = \infty$, akkor (2.18) nyilván teljesül, ezért a továbbiakban feltesszük,

hogy $\mathbb{E}(X) < \infty$. Ha X az Ω eseménytéren van értelmezve, akkor definiáljuk az $Y : \Omega \rightarrow \mathbb{R}$ változót a következőképp: minden $\omega \in \Omega$ -ra legyen

$$Y(\omega) = \begin{cases} a, & \text{ha } X(\omega) \geq a, \\ 0 & \text{különben.} \end{cases}$$

Másképp írva $Y = a \cdot \mathbf{1}_{X \geq a}$, ahol $\mathbf{1}_{X \geq a}$ az $\{X \geq a\}$ eseményhez tartozó indikátorváltozó.

Vegyük észre, hogy $X \geq Y$ érvényes, ebből pedig $X - Y \geq 0$. Egy nemnegatív változó várható értéke mindig nemnegatív (hiszen a definiáló összegben minden tag nemnegatív), és feltevésünk szerint $\mathbb{E}(X) < \infty$, továbbá $\mathbb{E}(-Y) = -a \cdot \mathbb{P}(X \geq a) < \infty$, ezért a várható érték linearitása miatt $\mathbb{E}(X - Y) = \mathbb{E}(X) - \mathbb{E}(Y) \geq 0$, vagyis $\mathbb{E}(X) \geq \mathbb{E}(Y) = a \cdot \mathbb{P}(X \geq a)$. Az $a > 0$ számmal elosztva az egyenlőtlenséget éppen a (2.18) egyenlőtlenséget kapjuk. \square

Gyakorlatok, feladatok

2.3.1. Gyakorlat. Tegyük fel, hogy az 5-ös lottó nyereményei rögzítettek: az 5-ös találat 1 milliárd, a 4-es 6 millió, a 3-as 35 ezer, míg a 2-es kétezer forintot nyer. Egy szelvényvel játszva mennyi a nyereményünk várható értéke?

2.3.2. Gyakorlat. Jelölje X egy kockadobás eredményét. Mennyi $\mathbb{E}((X - 3)^2)$?

2.3.3. Feladat. Egy urnában egy piros és egy fehér golyó van. Kihúzzunk egy golyót, és ha az piros, akkor megállunk. Ellenkező esetben a kihúzott fehér golyót egy újabb fehér golyóval együtt visszarakjuk az urnába, és ugyanilyen szabály szerint folytatjuk a húzást. Legyen Y az első piros golyóig húzott golyók száma. Határozzuk meg Y eloszlását és várható értékét.

2.3.4. Feladat. (M) Tekintsük a 2.1.9. feladat példáit, azaz dobáljunk egy pénzérmét addig, amíg egymás után 2 ill. 3 egyforma eredményt nem kapunk. Mi lesz a dobások számának várható értéke az egyes esetekben?

2.3.2. Feltételes várható érték

Ahogy azt már korábban láttuk, a kimenetekkel kapcsolatos információk módosíthatják az események valószínűségeit, ezzel együtt pedig az eseménytéren értelmezett valószínűségi változók eloszlásait és így a várható értékeket is. Vagyis ha $A \subset \Omega$ egy pozitív valószínűségű esemény és $X : \Omega \rightarrow \mathbb{R}$ egy (diszkrét) valószínűségi változó, akkor az A esemény bekövetkezése mellett az X eloszlását a $\mathbb{P}(X = k \mid A)$ feltételes valószínűségek írják le, ez tehát az X eloszlása a $\mathbb{P}_A(\cdot) = \mathbb{P}(\cdot \mid A)$ valószínűségi mértékre nézve. Ha pedig adott egy feltételes eloszlás, akkor arra nézve a várható érték is definiált:

2.3.3. Definíció. Legyen $X : \Omega \rightarrow \mathbb{R}$ egy diszkrét valószínűségi változó az $(\Omega, \mathcal{A}, \mathbb{P})$ valószínűségi mezőn, és legyen $A \subset \Omega$ egy esemény, melyre $\mathbb{P}(A) > 0$ teljesül. Ekkor az X változónak az A -ra vett $\mathbb{E}(X \mid A)$ *feltételes várható értéke* az X -nek a $\mathbb{P}(\cdot \mid A)$ valószínűségi mérték szerinti várható értéke, amennyiben az véges, azaz

$$\mathbb{E}(X \mid A) = \sum_{k \in \text{ran } X} k \cdot \mathbb{P}(X = k \mid A).$$

Rögtön megjegyezzük, hogy a feltételes várható érték egyszerűen egy speciálisan választott mértékre vonatkozó várható érték, tehát érvényes rá minden, amit a várható értékre belátunk, azaz például a *feltételes várható érték lineáris*: ha $X, Y : \Omega \rightarrow \mathbb{R}$ diszkrét valószínűségi

változók, $A \subset \Omega$ pedig egy pozitív valószínűségű esemény, akkor feltéve, hogy az $\mathbb{E}(X | A)$ és $\mathbb{E}(Y | A)$ feltételes várható értékek végesek,

$$\mathbb{E}(X + Y | A) = \mathbb{E}(X | A) + \mathbb{E}(Y | A) \quad \text{ill.} \quad \mathbb{E}(cX | A) = c\mathbb{E}(X | A)$$

érvényes, ahol $c \in \mathbb{R}$ egy tetszőleges valós szám.

Ahogy a feltételes valószínűség esetén is, ez a fogalom hasznos lehet olyan szituációkban, amikor az $\{X = k\}$ események valószínűségének kiszámítása valamilyen esemény teljesülése esetén lényegesen leegyszerűsödik. Mivel a feltételben szereplő esemény szabadon választható, így - még ha ez talán ezen a ponton nem is látszik világosan - a feltételes várható érték rendkívül széles körűen alkalmazható. Az alkalmazásokhoz azonban szükségünk lesz arra, hogy szét tudjuk bontani különböző esetekre a számolást. Ehhez a következő állítás lesz segítségünkre:

2.3.4. Állítás. *Legyen $Y : \Omega \rightarrow \mathbb{R}$ egy diszkrét valószínűségi változó, melynek véges a várható értéke, legyen továbbá $A \subset \Omega$ egy pozitív valószínűségű esemény. Ekkor $\mathbb{E}(Y | A)$ is véges, és*

$$\mathbb{E}(Y | A) = \frac{1}{\mathbb{P}(A)} \mathbb{E}(Y \mathbf{1}_A),$$

ahol $\mathbf{1}_A$ az A eseményhez tartozó indikátorváltozó.

Bizonyítás. A feltételes várható érték definíciója szerint

$$\mathbb{E}(Y | A) = \sum_{k \in \text{ran } Y} k \cdot \mathbb{P}(Y = k | A) = \sum_{k \in \text{ran } Y} k \cdot \frac{\mathbb{P}(\{Y = k\} \cap A)}{\mathbb{P}(A)}.$$

Az A esemény felírható a hozzá tartozó $\mathbf{1}_A$ indikátorváltozó segítségével $\{\mathbf{1}_A = 1\}$ alakban, így $k \neq 0$ esetén

$$\mathbb{P}(\{Y = k\} \cap A) = \mathbb{P}(Y = k, \mathbf{1}_A = 1) = \mathbb{P}(Y \cdot \mathbf{1}_A = k)$$

teljesül. Amennyiben a 0 is az Y értékkészletében van, akkor viszont $0 \cdot \mathbb{P}(\{Y = k\} \cap A) = 0 \cdot \mathbb{P}(Y \cdot \mathbf{1}_A = 1) = 0$ teljesül, így a fenti utolsó összeg

$$\mathbb{E}(Y | A) = \frac{1}{\mathbb{P}(A)} \sum_{k \in \text{ran } Y} k \cdot \mathbb{P}(Y \cdot \mathbf{1}_A = k)$$

alakba írható. Itt ketté bonthatjuk az Y értékkészletét aszerint, hogy benne vannak-e az elemei az $Y \cdot \mathbf{1}_A$ szorzat értékkészletében:

$$\begin{aligned} \frac{1}{\mathbb{P}(A)} \sum_{k \in \text{ran } Y} k \cdot \mathbb{P}(Y \cdot \mathbf{1}_A = k) &= \frac{1}{\mathbb{P}(A)} \sum_{k \in \text{ran } Y \cap \text{ran } Y \mathbf{1}_A} k \cdot \mathbb{P}(Y \cdot \mathbf{1}_A = k) \\ &\quad + \frac{1}{\mathbb{P}(A)} \sum_{k \in \text{ran } Y \setminus \text{ran } Y \mathbf{1}_A} k \cdot \mathbb{P}(Y \cdot \mathbf{1}_A = k). \end{aligned}$$

Vegyük észre, hogy a jobb oldalon lévő második összegben az $Y \cdot \mathbf{1}_A = k$ egyenlőség nem teljesülhet (hiszen $k \notin \text{ran } Y \mathbf{1}_A$), így az ott szereplő valószínűségek mind nullák. Tehát a várható érték definíciója szerint

$$(2.19) \quad \mathbb{E}(Y | A) = \frac{1}{\mathbb{P}(A)} \sum_{k \in \text{ran } Y \mathbf{1}_A} k \cdot \mathbb{P}(Y \cdot \mathbf{1}_A = k) = \frac{1}{\mathbb{P}(A)} \mathbb{E}(Y \mathbf{1}_A).$$

Hátra van még a fenti számítási lépések jogosságának indoklása (hiszen itt az összegek végtelenek is lehetnek). Ehhez megjegyezzük, hogy $k \neq 0$ esetén az $Y \cdot \mathbf{1}_A = k$ egyenlőségből következik az $Y = k$ egyenlőség is, hiszen ekkor az indikátorváltozó csak 1 értéket vehet fel. Azaz, ha $k \neq 0$, akkor

$$\{Y \cdot \mathbf{1}_A = k\} \subset \{Y = k\} \implies \mathbb{P}(Y \cdot \mathbf{1}_A = k) \leq \mathbb{P}(Y = k).$$

Mivel $k = 0$ esetén a (2.19) egyenlőségben szereplő összeg megfelelő tagja egyszerűen 0, így

$$\begin{aligned} 0 \leq \sum_{k \in \text{ran } Y \mathbf{1}_A} |k| \cdot \mathbb{P}(Y \mathbf{1}_A = k) &\leq \sum_{k \in \text{ran } Y \mathbf{1}_A} |k| \cdot \mathbb{P}(Y = k) \\ &\leq \sum_{k \in \text{ran } Y} |k| \cdot \mathbb{P}(Y = k) = \mathbb{E}(|Y|) < \infty, \end{aligned}$$

mert az állításban tett feltétel szerint az Y várható értéke véges. A fenti első sor abszolút konvergenciája biztosítja, hogy $\mathbb{E}(Y \mathbf{1}_A)$ véges, ill. hogy a bizonyításban szereplő számolási lépések végtelen összeg esetén is helyesek, és ebből egyben következik (a számolásokat fordított sorrendben elvégezve) az $\mathbb{E}(Y | A)$ várható érték végeessége, valamint a bizonyítandó egyenlőség. \square

A következő példán bemutatjuk, hogy hogyan alkalmazható a fenti állítás az esetek szétválasztására.

2.3.6. Példa. Béla minden héten lottózik, egyes heteken az ötöslottón, máskor a hatoslottón játszik, de mindig 1 darab szelvényt vesz. Hogy melyik játékban vesz részt, azt pedig minden esetben egy szabályos érme feldobásával dönti el, ha fejet dob, akkor az ötöslottót, írás esetén a hatoslottót választja. Az ötöslottón 90 számból húznak öt darabot, míg a hatoslottón 45 számból húznak hatot.

Tegyük fel, hogy egy ötöslottószelvény ára 400 forint, továbbá, hogy egy adott héten az ötöslottón ötös találattal 300 millió forintot, négyes találattal egymillió forintot, hármas találattal ötvenezer forintot, kettes találattal pedig ötezer forintot lehet nyerni. (Az egyszerűség kedvéért tegyük fel most azt is, hogy a megnyert összegben semelyik játékban nem kell más, ugyanannyi találatot elérő nyertesekkel osztózni.) A hatoslottószelvények ára szintén 400 forint, ott a hatos találat 1,3 milliárd forintot, az ötös találat négyszázezer forintot, a négyes találat tízezer forintot, míg a hármas találat háromezer forintot fizet.

Meghatározzuk a Béla által a játékon elért heti profit várható értékét (a profit alatt játék során nyert nyereménynek és a kiadásnak a különbségét értjük), jelölje ezt a profitot X . Az X várható értékét persze könnyen kiszámolhatjuk azon feltétel mellett, hogy Béla az ötös- vagy a hatoslottón játszik, tehát hogy fejet dob (F) vagy írást (I):

$$\begin{aligned} \mathbb{E}(X | F) &= 300\,000\,000 \cdot \frac{1}{\binom{90}{5}} + 1\,000\,000 \cdot \frac{\binom{5}{4} \cdot \binom{85}{1}}{\binom{90}{5}} \\ &\quad + 50\,000 \cdot \frac{\binom{5}{3} \cdot \binom{85}{2}}{\binom{90}{5}} + 5\,000 \cdot \frac{\binom{5}{2} \cdot \binom{85}{3}}{\binom{90}{5}} - 400 \approx -230,52 \text{ Ft}, \end{aligned}$$

illetve

$$\begin{aligned}\mathbb{E}(X | I) &= 1\,300\,000\,000 \cdot \frac{1}{\binom{45}{6}} + 400\,000 \cdot \frac{\binom{6}{5} \cdot \binom{39}{1}}{\binom{45}{6}} \\ &\quad + 10\,000 \cdot \frac{\binom{6}{4} \cdot \binom{39}{2}}{\binom{45}{6}} + 3\,000 \cdot \frac{\binom{6}{3} \cdot \binom{39}{3}}{\binom{45}{6}} - 400 \approx -147,93 \text{ Ft.}\end{aligned}$$

Érezhető, hogy mivel a fej és az írás egyformán valószínű, így a profit várható értéke e két várható érték átlaga kell legyen. Ezt most le is fogjuk vezetni, ehhez először írjuk az átlagot a következő alakba:

$$-189,225 \approx \frac{\mathbb{E}(X | F) + \mathbb{E}(X | I)}{2} = \mathbb{E}(X | F)\mathbb{P}(F) + \mathbb{E}(X | I)\mathbb{P}(I).$$

Az utóbbi összeg tagjai a 2.3.4. állítás szerint

$$\mathbb{E}(X\mathbb{1}_F) + \mathbb{E}(X\mathbb{1}_I) = \mathbb{E}(X\mathbb{1}_F + X\mathbb{1}_I) = \mathbb{E}(X(\mathbb{1}_F + \mathbb{1}_I)) = \mathbb{E}(X),$$

ahol használtuk a várható érték linearitását, illetve hogy az F és az I események közül mindig pont az egyik következik be, így az indikátorok összege mindig $1 + 0 = 1$. Tehát a profit várható értéke kb. $-189,23$ Ft.

A fenti gondolatmenet kulcsa a 2.3.4. állításon túl az volt, hogy az indikátorváltozók összege azonosan 1 volt, ez pedig azért volt érvényes, mert a megfelelő események közül mindig pontosan az egyik következett be, azaz az események teljes eseményrendszert alkottak. Ezekre a szituációkra a fentivel analóg állítás érvényes:

2.3.5. Tétel (A teljes várható érték tétele). *Legyen $Y : \Omega \rightarrow \mathbb{R}$ egy diszkrét valószínűségi változó, melynek véges a várható értéke, legyen továbbá $A_1, \dots, A_n \subset \Omega$ egy pozitív valószínűségű eseményekből álló teljes eseményrendszer. Ekkor $\mathbb{E}(Y | A_i) < \infty$ minden $1 \leq i \leq n$ esetén, és*

$$\mathbb{E}(Y) = \sum_{i=1}^n \mathbb{E}(Y | A_i)\mathbb{P}(A_i).$$

Bizonyítás. A bizonyítás a fenti példában látott gondolatmenet alapján történik. Mivel A_1, \dots, A_n egy teljes eseményrendszer, így az ezen eseményekhez tartozó $\mathbb{1}_{A_i}$ indikátorváltozók összege azonosan 1. Vagyis $\mathbb{E}(Y) = \mathbb{E}(Y(\mathbb{1}_{A_1} + \dots + \mathbb{1}_{A_n})) = \mathbb{E}(Y\mathbb{1}_{A_1} + \dots + Y\mathbb{1}_{A_n})$. A 2.3.4. állítás szerint az $\mathbb{E}(Y | A_i)$ és így az $\mathbb{E}(Y\mathbb{1}_{A_i})$ várható értékek végesek, így a várható értékek linearitása ill. a 2.3.4. állításban bizonyított egyenlőség miatt

$$\mathbb{E}(Y) = \mathbb{E}(Y\mathbb{1}_{A_1}) + \dots + \mathbb{E}(Y\mathbb{1}_{A_n}) = \mathbb{E}(Y | A_1)\mathbb{P}(A_1) + \dots + \mathbb{E}(Y | A_n)\mathbb{P}(A_n),$$

ezzel a tételt beláttuk. □

Megjegyezzük, hogy a fenti állításnak, valamint például a teljes valószínűség tételének megfogalmazható egy végtelen sok eseményből álló teljes eseményrendszerre vonatkozó változata is. Azaz ha $A_1, \dots, A_n, \dots \subset \Omega$ egy megszámlálhatóan végtelen sok pozitív valószínűségű eseményből álló teljes eseményrendszer, akkor a teljes valószínűség tétele

$$\mathbb{P}(B) = \sum_{i=1}^{\infty} \mathbb{P}(B | A_i)\mathbb{P}(A_i)$$

alakba írható minden $B \subset \Omega$ eseményre, ha továbbá $Y : \Omega \rightarrow \mathbb{R}$ egy véges várható értékű diszkrét valószínűségi változó, akkor

$$\mathbb{E}(Y) = \sum_{i=1}^{\infty} \mathbb{E}(Y | A_i) \mathbb{P}(A_i).$$

Míg az első egyenlőség a mérték σ -additivitásának alkalmazásával lényegében ugyanúgy bizonyítható, mint a véges sok eseményből álló eseményrendszerre kimondott 1.2.4. tétel, az utóbbi egyenlőség bizonyítása már egy kicsit kényelmetlenebb, de az abszolút konvergencia sorokra vonatkozó alapvető eredmények segítségével a 2.3.5. tétel és a 2.3.4. állítás bizonyításai könnyen módosíthatók úgy, hogy a fenti állítást nyerjük. Az igazán hatékony és elegáns módszer azonban nem ez, hanem egy mértékelméleti eszköztár kiépítése és alkalmazása volna, amely azonban túlmutat a jegyzet keretein.

2.3.7. Példa. A teljes várható érték tételének alkalmazásaként ismét kiszámoljuk a geometriai eloszlás várható értékét. A most következő számolás pontosan ugyanarra az egyenletre vezet majd, amit az előző szakaszban kaptunk. Ez nem véletlen, ugyanis ez a levezetés lényegét tekintve nem különbözik a korábbtól, viszont megmutatja a számolás mögött rejlő okokat, míg a korábbi levezetés ezen okok feltüntetése nélkül valószínűleg trükkös átalakítások egymásutánjának tűnik.

A teljes várható érték tételét fogjuk alkalmazni egy konkrét $X \sim Geo(p)$ változóra és egy megfelelően választott eseményre (a várható érték végessége ugyanúgy látható, mint az előző szakaszban, ezzel itt külön nem foglalkozunk). Legyen tehát X a dobások száma egy független kísérletsorozatban, amelyet egy p valószínűségű esemény bekövetkezéséig végzünk. Tekintsük továbbá azt az eseményt, hogy már az első kísérletünk is sikeres, azaz az $\{X = 1\}$ eseményt. Világos, hogy $\mathbb{E}(X | X = 1) = 1$ érvényes, továbbá ha az első kísérletünk sikertelen, akkor a kísérletek függetlensége, vagy még pontosabban az eloszlás örökifjú tulajdonsága miatt a második kísérlet előtt éppen abban a szituációban vagyunk, mint a kísérletsorozat elején, vagyis ezen feltétel mellett a kísérletek számának várható értéke $\mathbb{E}(X | X > 1) = 1 + \mathbb{E}(X)$.

Ez az érvelés formálisan a következőképp írható le a feltételes várható érték ill. az örökifjú tulajdonság definíciójának felhasználásával:

$$\mathbb{E}(X | X > 1) = \sum_{k=1}^{\infty} k \cdot \mathbb{P}(X = k | X > 1).$$

Először vegyük észre, hogy $k = 1$ esetén a jobb oldalon a $\mathbb{P}(X = 1 | X > 1) = 0$ valószínűséget kapjuk. Ha $k > 1$, akkor az eloszlás örökifjúságának felhasználásához átírjuk a fenti összegben szereplő valószínűségeket a következő alakba:

$$\begin{aligned} \mathbb{P}(X = k | X > 1) &= \mathbb{P}(X > k - 1 | X > 1) - \mathbb{P}(X > k | X > 1) \\ &= \mathbb{P}(X > k - 2) - \mathbb{P}(X > k - 1) = \mathbb{P}(X = k - 1), \end{aligned}$$

azaz

$$\begin{aligned} \mathbb{E}(X | X > 1) &= \sum_{k=2}^{\infty} k \cdot \mathbb{P}(X = k - 1) = \sum_{k=1}^{\infty} (k + 1) \mathbb{P}(X = k) \\ &= \sum_{k=1}^{\infty} k \cdot \mathbb{P}(X = k) + \sum_{k=1}^{\infty} \mathbb{P}(X = k) = \mathbb{E}(X) + 1, \end{aligned}$$

ahol a várható érték definíciója mellett felhasználtuk, hogy az eloszlás súlyfüggvényének értékeit összegezve 1-et kapunk.

Végezetül a teljes várható érték tételét alkalmazva

$$\begin{aligned}\mathbb{E}(X) &= \mathbb{E}(X \mid X = 1)\mathbb{P}(X = 1) + \mathbb{E}(X \mid X > 1)\mathbb{P}(X > 1) \\ &= 1 \cdot p + (\mathbb{E}(X) + 1)(1 - p) = 1 + (1 - p) \cdot \mathbb{E}(X),\end{aligned}$$

ezt átrendezve pedig $\mathbb{E}(X) = \frac{1}{p}$ adódik.

A Randomized Quick Sort algoritmus (kiegészítő anyag)

Térjünk most rá a szakasz elején említett Randomized Quick Sort rendezőalgoritmus tárgyalására. Ez - mint a neve is mutatja - egy véletlent használó algoritmus, azaz futása véletlenszerű értékektől függ. Az már önmagában egy nem triviális probléma, hogy a véletlen viselkedés számítógép segítségével hogyan szimulálható, de ezzel ebben a jegyzetben nem foglalkozunk, hanem feltesszük, hogy egy program képes egyenletesen véletlenszerűen generálni egy egész számot egy megadott intervallumban.

Alapvetően két különböző típusú véletlent használó algoritmust különböztetünk meg. Egy algoritmust *Monte Carlo típusú algoritmusnak* nevezünk, ha futása a véletlentől függ, és pozitív valószínűséggel helytelen eredményt produkál. Első ránézésre talán nem világos, hogy miért lehet hasznos egy ilyen algoritmus, de ezek alkalmazhatósága egyszerűen azon múlik, hogy a gyakorlatban sokszori független futtatás esetén rendkívül kicsi, voltaképp elhanyagolható annak a valószínűsége, hogy mindig rossz eredményt kapunk. Az érdeklődő olvasó az [5] jegyzet 2.3. szakaszában talál egy példát egy Monte Carlo típusú gráfelméleti algoritmusra, a mélyebb részletek iránt érdeklődőknek pedig a [6] könyvet ajánljuk.

Az itt tárgyalt példa nem a fenti kategóriába tartozik, azaz minden esetben helyes eredményt produkál. Ami viszont a véletlentől függ, az a futásidő. Az ilyen algoritmusokat *Las Vegas típusú algoritmusoknak* nevezzük. Egy program tényleges futásideje természetesen számos tényezőtől függ magán az algoritmuson kívül is, de egy algoritmus tervezésénél értelemszerűen az eljárásban leírt (elemi) lépések számát vizsgáljuk. Mivel itt csak egy rövid kitekintést teszünk, így ebben a jegyzetben nem foglalkozunk azzal a kérdéssel sem, hogy mit tekinthetünk egy algoritmus elemi lépésének, minden esetre a következőkben egész számok összehasonlítását, tömbelemek indexének növelését, számok tömbbe való írását, sőt, egy véletlen szám generálását is egyetlen lépésként fogjuk kezelni.

Amit viszont fontos kiemelni, hogy a lépések száma nyilvánvalóan függ az input méretétől, hiszen nagyobb bemenet esetén már annak végigolvasása is több lépést igényel. Az input méretén tipikusan azt értjük, hogy a bemenet leírásához hány bájtira van szükségünk. Mivel az alábbiakban egy egész számokból álló tömb lesz a bemenet, és ezen számokat korlátosnak tesszük fel, így mindegyikük fix (konstans) számú bájt segítségével leírható. Következésképp a bemenet mérete alatt itt a tömbben szereplő számok számát fogjuk érteni, és az eljárásunk lépésszámát is ennek függvényében vizsgáljuk. A fenti kérdések részletesebb vizsgálatával kapcsolatban a [8] könyvet ajánljuk az olvasónak.

A Randomized Quick Sort egy rekurzív rendező algoritmus, amelynek bemenete egy n egész számból álló tömb, kimenete pedig egy olyan tömb, amely ugyanezen számokat tartalmazza nagyság szerint nem csökkenő sorrendben. Az alábbiakban feltesszük, hogy a bemenetben minden szám különböző, és ennek megfelelően nem csökkenő helyett növekvő sorrendet

követelünk meg a kimenetben. A bemenet méretének a fenti megjegyzések értelmében az n számot fogjuk tekinteni.

Mivel az algoritmus lépésszáma egy véletlen pozitív egész szám, így ezt a lépésszámot egy X diszkrét valószínűségi változóval modellezhetjük. Valójában az algoritmus hatékonysága szempontjából azok a lépések lesznek a relevánsak, amikor két tömbelemet hasonlítunk össze, ugyanis a többi lépés figyelembe vételével is az összehasonlítások számának egy konstanssorosát kapjuk a lépésszám felső becslésére. A lépésszám helyett tehát a továbbiakban az összehasonlítások számát fogjuk vizsgálni.

Egy ilyen eljárás nyilván akkor használható jól a gyakorlatban, ha az esetek "többségben" a futásidő "elég jó", még ha kis valószínűséggel esetleg "a lehető legrosszabb" eset is bekövetkezhet. Ugyan az X változó eloszlásának a pontos elemzése bonyolult, azonban ehelyett megelégszünk az átlagos lépésszám, azaz az $\mathbb{E}(X)$ várható érték becslésével. Figyelemre méltó, hogy ez tehát az eloszlás pontos ismerete nélkül is megtehető, sőt annak valószínűségét is becsülni tudjuk majd, hogy a futásidő rosszabb az átlagosnál. Megjegyezzük, hogy ez a várható érték egészen pontosan kiszámolható, de valójában ennek a lényeges következménye éppen az a becslés, amit mi alább a feltételes várható érték alkalmazásaként megkapunk. A pontos érték meghatározása megtalálható pl. az [5] jegyzetben.

Az algoritmus lépéseit a következőképp írhatjuk le:

RANDOMIZED QUICK SORT

Input: páronként különböző pozitív egészek egy $S = (x_1, \dots, x_n)$ sorozata.

1. Válasszuk ki egyenletesen véletlenszerűen az S egy Y elemét.
2. Hasonlítsuk össze az S -nek minden, az Y -tól különböző elemét Y -nal, és adjuk meg az ezekből álló $S_{<} = (y_1, \dots, y_k)$ ill. $S_{>} = (z_1, \dots, z_l)$ sorozatokat, ahol
 - $y_i < Y$ ill. $z_j > Y$ érvényes minden $1 \leq i \leq k$ és $1 \leq j \leq l$ esetén,
 - az elemek sorrendje egymáshoz képest ugyanaz $S_{<}$ -ben és $S_{>}$ -ban, mint S -ben volt,
 - az S -nek minden Y -tól különböző eleme szerepel vagy $S_{<}$ -ben, vagy $S_{>}$ -ban.
3. Hajtsuk végre rekurzívan az algoritmust az $S_{<}$ és $S_{>}$ sorozatok közül azokra, amelyek nem üresek.
4. Adjuk vissza az elemeket a következő sorrendben: először a már rendezett $S_{<}$ sorozat elemeit, majd az Y elemet, végül a már rendezett $S_{>}$ sorozat elemeit.

Output: Az S elemei növekvő sorrendben.

Figyeljük meg, hogy a 2. lépésben kapott $S_{<}$ ill. $S_{>}$ sorozatok minden esetben kevesebb elemből állnak az S sorozatnál, továbbá amennyiben $S_{<}$ vagy $S_{>}$ az üres sorozat, akkor az már biztosan rendezett. Így tehát az eljárás biztosan véget ér véges sok lépés után, és világos, hogy mindig helyes eredményt ad.

Nézzük meg az algoritmus működését egy példán: ha a bemenet

$$S = (17, 6, 56, 23, 4, 19, 58, 28, 14, 43, 62, 3),$$

és az 1. lépésben az $Y = 23$ elemet választjuk, akkor a 2. lépésben az

$$S_{<} = (17, 6, 4, 19, 14, 3), \quad \text{ill.} \quad S_{>} = (56, 58, 28, 43, 62)$$

sorozatokat kapjuk, majd ezeket rekurzívan ugyanígy rendezzük.

Térjünk rá az összehasonlítások számára. Először is figyeljük meg, hogy szerencsétlen esetben a lehető legegyszerűbb forgatókönyv is előfordulhat, nevezetesen hogy minden elemet minden másikkal összehasonlítunk, ez tehát $\binom{n}{2} = \frac{n^2-n}{2}$ összehasonlítást jelent. Ráadásul mindez úgy is megtörténhet, hogy már eleve egy rendezett sorozat az input: ha a kiindulási sorozat $S = (1, 2, \dots, n)$, továbbá mindig éppen az utolsó elemet választjuk az 1. lépésben, akkor $S_{>}$ mindig üres lesz, míg a k -adik (rekurzív) hívásnál az $S_{<} = (1, \dots, n-k)$ sorozatot kapjuk. Azt is vegyük azonban észre, hogy egy ilyen futás valószínűsége rendkívül kicsi, mindössze $\frac{1}{n!}$.

Az összehasonlítások számának vizsgálatát nagyban leegyszerűsíti két egyszerű észrevétel. Az egyik, hogy ha ugyanazok az elemek más sorrendben szerepelnek a kezdeti sorozatban, az a lépésszám eloszlásán nem változtat, mivel az összehasonlítás alapjául szolgáló Y elemet egyenletesen véletlenszerűen választjuk. Tehát egy konkrét futásnál kapott Y elemek sorozata egy más kezdeti sorrend esetén is ugyanolyan valószínűséggel adódik, bár maguk az $S_{<}$ ill. $S_{>}$ sorozatok a két futás egyes lépéseinél nem lesznek azonosak, csak egymás permutációi, de így is ugyanolyan összehasonlításszámot kapunk a két futásnál.

Továbbá azt is megfigyelhetjük, hogy a sorozatokban szereplő számok értéke az algoritmus futásának szempontjából irreleváns, hiszen csak a számok közt fennálló relációk játszanak szerepet az egyes lépéseknél. Mindebből tehát következik, hogy **az összehasonlítások számának várható értéke kizárólag az n hossz függvénye**, nem függ sem a konkrét elemektől, sem azok sorrendjétől. Ezt figyelembe véve kimondhatjuk a szakasz fő eredményét:

2.3.6. Tétel. *Jelölje X_n az összehasonlítások számát a Randomized Quick Sort algoritmus futása során egy n hosszúságú bemenetnél, ekkor $\mathbb{E}(X_n) \leq 2nH_n$, ahol*

$$(2.20) \quad H_n = \sum_{k=1}^n \frac{1}{k} = \ln n + \gamma + o(1),$$

itt $\gamma \approx 0,55721 \dots$ az úgynevezett Euler-konstans, az $o(1)$ (ejtsd: "kis ordó 1") jelölés pedig azt jelenti, hogy $H_n - \ln n - \gamma \rightarrow 0$, ha $n \rightarrow \infty$.

Bizonyítás. Mivel a fenti megjegyzések szerint $\mathbb{E}(X_n)$ csak az n számtól függ, így vezessük be erre a $C(n)$ jelölést a továbbiakban. Világos, hogy $C(0) = C(1) = 0$, $n = 0, 1$ esetén tehát teljesül az állítás, és a továbbiakban feltehető, hogy $n > 1$.

Mivel az első választott elemmel minden más elemet összehasonlítunk, így ez $n-1$ összehasonlítást jelent. Ezek után futtatjuk az algoritmust az $S_{<}$ ill. $S_{>}$ sorozatokra, jelölje az itt történő összehasonlítások számát $X_{S_{<}}$ ill. $X_{S_{>}}$. Ezeket a jelöléseket használva az $X_n = n-1 + X_{S_{<}} + X_{S_{>}}$ egyenlőség adódik.

A tétel kimondása előtti megjegyzések szerint feltehető, hogy az input az $1, 2, \dots, n$ számokból áll. Legyen Y az első lépésben kiválasztott elem. Ha $Y = k$, akkor $S_{<}$ hossza $k-1$, $S_{>}$ hossza pedig $n-k$, így a feltételes várható érték linearitása miatt

$$\begin{aligned} \mathbb{E}(X_n \mid Y = k) &= n-1 + \mathbb{E}(X_{S_{<}} \mid Y = k) + \mathbb{E}(X_{S_{>}} \mid Y = k) \\ &= n-1 + C(k-1) + C(n-k). \end{aligned}$$

Ebből a teljes várható érték tétele szerint

$$\begin{aligned} C(n) &= \mathbb{E}(X_n) = \sum_{k=1}^n \mathbb{E}(X_n | Y = k) \mathbb{P}(Y = k) \\ &= \frac{1}{n} \left(n(n-1) + \sum_{k=1}^n [C(k-1) + C(n-k)] \right) = n-1 + \frac{2}{n} \sum_{k=1}^{n-1} C(k), \end{aligned}$$

ahol a $C(0) = 0$ egyenlőséget is kihasználtuk az utolsó lépésben. A jobb oldalon lévő tagok nagy része kiejthető, ha a formulát n helyett $(n-1)$ -re is felírjuk:

$$C(n-1) = n-2 + \frac{2}{n-1} \sum_{k=1}^{n-2} C(k).$$

Megjegyezzük, hogy ez $n = 2$ esetén is érvényes, hiszen ekkor mindkét oldal 0-t ad. A következő különbségben tehát a jobb oldalon számos tag kiesik:

$$nC(n) - (n-1)C(n-1) = n(n-1) - (n-1)(n-2) + 2C(n-1) = 2(n-1) + 2C(n-1),$$

amit átrendezve

$$\frac{C(n)}{n+1} = \frac{2(n-1)}{n(n+1)} + \frac{C(n-1)}{n} < \frac{2}{n} + \frac{C(n-1)}{n}$$

adódik. Az utóbbi egyenlőtlenséget egymás után $(n-1)$ -szer alkalmazva

$$\frac{C(n)}{n+1} < 2 \cdot \left(\frac{1}{n} + \dots + \frac{1}{2} \right) = 2(H_n - 1),$$

vagyis

$$C(n) < 2(n+1)(H_n - 1) = 2(nH_n + H_n - n - 1) < 2nH_n.$$

A (2.20) egyenlőség jól ismert az analízisből, ezt itt nem bizonyítjuk. □

A domináns tag a fenti felső becslésben $2n \ln n$, és tulajdonképpen belátható, hogy ez önmagában egy felső becslés az összehasonlítások számának várható értékére (lásd az [5] jegyzetet), valamint megmutatható az is, hogy nagyságrendjét tekintve ez valójában optimális a kizárólag összehasonlításra alapuló algoritmusok között (lásd a [8] könyvet). Igaz, ezt a lépésszámot csak átlagos értelemben kaptuk meg, de könnyen becslést kaphatunk a rossz esetek valószínűségére például a Markov-egyenlőtlenség (2.3.3. tétel) segítségével. Ha ugyanis $f(n)$ bármilyen függvény, ami gyorsabban nő a logaritmusnál, azaz $\ln n / f(n) \rightarrow 0$ teljesül, ha $n \rightarrow \infty$ (például ilyen az $f(n) = n^\varepsilon$ függvény bármilyen kicsi pozitív $\varepsilon > 0$ esetén), akkor a (2.18) egyenlőtlenség szerint

$$\mathbb{P}(X_n \geq n \cdot f(n)) \leq \frac{\mathbb{E}(X_n)}{n \cdot f(n)} \leq \frac{2n \ln n}{n \cdot f(n)} = \frac{2 \ln n}{f(n)} \rightarrow 0,$$

ha $n \rightarrow \infty$. Vagyis a "rossz esetek" valószínűsége tetszőlegesen kicsi, ha n elég nagy. Mindez összhangban van azzal a tapasztalattal, hogy a Quick Sort a gyakorlatban nagyon hatékony algoritmusnak bizonyul.

Gyakorlatok, feladatok

2.3.5. Gyakorlat. Négyyszer feldobunk egy szabályos érmét, legyen X a dobások során kapott fejek száma, és legyen A az az esemény, hogy először a második dobásnál kapunk fejet. Határozzuk meg az $\mathbb{E}(Y | A)$ feltételes várható értéket.

2.3.6. Gyakorlat. Háromszor dobunk egy szabályos kockával. Jelölje X a legkisebb, Y pedig a legnagyobb értéket. Adjuk meg az $\mathbb{E}(X | Y = 3)$ feltételes várható értéket.

2.3.7. Gyakorlat. Egy fájl összes előforduló példányát keressük, amik a két meghajtónk egyikén (és csak egyikén) vannak, $\frac{2}{3}$ eséllyel az elsőn, egyébként a másodikon. A keresés lefuttatása az első meghajtó esetén 7 percre tart, a második esetén 3 percre. Tegyük fel, hogy a fájl összes példányát csak az öt tartalmazó meghajtó teljes átnézése után találjuk meg. Várhatóan mennyi időbe telik megtalálni a fájlokat, ha az átnézést

a) az első meghajtóval, b) a második meghajtóval kezdjük?

2.3.8. Gyakorlat. Egy bányában eltévedt egy ember, aki egy csomópontban áll, ahonnan három irányba mehet tovább. Az első irányt választva kijut a bányából egy óra alatt, a második irányt választva visszaérkezik ugyanide 2 óra alatt, a harmadik irányt választva pedig 3 óra alatt ér vissza ugyanide. Feltéve, hogy minden esetben, amikor választania kell, akkor egyenlő valószínűséggel választja bármelyik opciót, várhatóan mennyi idő múlva jut ki a bányából?

2.3.9. Gyakorlat. Egy társasjátékban egy szabályos dobókockával dobva döntjük el, hogy hányat lépünk a bábuval. A kockával egészen addig dobunk, míg hatostól különböző számot nem kapunk, a lépések száma pedig az összes dobás értékének összege. Határozzuk meg a lépések számának várható értékét egy ilyen dobássorozatnál.

2.3.10. Gyakorlat. A gyönyörű királylány kezéért sokan versengenek, őt azonban eléggé lekötik udvari teendői, így meghagyja szolgálójának, hogy minden randevúmeghívást (minden egyéb körülménytől függetlenül) csak 56,25% eséllyel fogadjon el a nevében. Az állhatatos királyfi elhatározza, hogy minden egyes nap elhívja randevúra a királylányt, amíg az igent nem mond. Nem tudván a királylány által felállított szabályokról, úgy próbálja meg pozitív irányba befolyásolni a döntést, hogy minden egyes nap rózsát visz a királylánynak, az első nap egy szálal, ezután minden másik nap kétszer annyit, mint a megelőző napon. Mennyi azon rózsaszálak számának a várható értéke, amit a királylány a királyfitól összesen kap?

2.3.11. Feladat. Egy 52 lapos franciakártya-paklihoz hozzákeverünk 2 jokert. Az így kapott pakli összekeverése után a pakli tetejéről húzunk egészen addig, amíg jokert nem kapunk. Mennyi a húzott lapok számának várható értéke?

2.3.3. Transzformált várható értéke

A későbbiek során számos alkalommal dolgozunk majd valószínűségi változók transzformáltjával. Ha $X : \Omega \rightarrow \mathbb{R}$ egy diszkrét valószínűségi változó, akkor a transzformáltja alatt egy $g(X)$ valószínűségi változót értünk, ahol $g : \mathbb{R} \rightarrow \mathbb{R}$ valamilyen függvény. Bár a következő fejezetekben majd más fontos függvény is előfordul, itt most az első lényeges példánk a $g(t) = t^2$ függvény lesz, vagyis egy X változó esetén az X^2 változót tekintjük.

Sok esetben érdekes és fontos kérdés a transzformált változó eloszlásának a meghatározása, de ebben a pontban pusztán a transzformált várható értékére koncentrálnunk, amelynek majd

a szórás meghatározásában lesz fontos szerepe (lásd a következő szakaszt). Az $g(X)$ változó várható értéke persze kiszámolható a $g(X)$ eloszlásából, de az alábbi tétel fontosságát éppen az adja, hogy az $\mathbb{E}(g(X))$ értéket pusztán az X eloszlása segítségével fejezi ki:

2.3.7. Tétel (Transzformált várható értéke). *Legyen X egy diszkrét valószínűségi változó, $g : \mathbb{R} \rightarrow \mathbb{R}$ pedig egy függvény, melyre az $\mathbb{E}(g(X))$ várható érték véges. Ekkor*

$$\mathbb{E}(g(X)) = \sum_{t \in \text{ran } X} g(t) f_X(t) = \sum_{t \in \text{ran } X} g(t) \mathbb{P}(X = t).$$

Bizonyítás. A várható érték definíciója szerint

$$\mathbb{E}(g(X)) = \sum_{s \in \text{ran } g(X)} s \cdot \mathbb{P}(g(X) = s) = \sum_{s \in \text{ran } g(X)} s \cdot \mathbb{P}(X \in g^{-1}(s)),$$

ahol $g^{-1}(s)$ az $\{s\}$ halmaz g általi ősképe, azaz $\{t \in \mathbb{R} : g(t) = s\}$ (hasonlóan ahhoz, mint ahogy ezt a fogalmat már a valószínűségi változókra is definiáltuk).

Egy fix $s \in \text{ran } g(X)$ esetén a fenti összeg egy tagja

$$s \cdot \mathbb{P}(X \in g^{-1}(s)) = \sum_{t \in \text{ran } X \cap g^{-1}(s)} g(t) \cdot \mathbb{P}(X = t) = \sum_{t \in \text{ran } X \cap g^{-1}(s)} g(t) \cdot f_X(t),$$

hiszen minden, az utóbbi szummában szereplő t értékre $s = g(t)$. Behelyettesítve tehát

$$\mathbb{E}(g(X)) = \sum_{s \in \text{ran } g(X)} \sum_{t \in \text{ran } X \cap g^{-1}(s)} g(t) \cdot f_X(t).$$

Végül még azt kell észrevenni, hogy itt $\text{ran } X$ minden tagját pontosan egyszer soroltunk fel, mégpedig egy adott $t \in \text{ran } X$ értéket a külső szumma $s = g(t)$ értékhez tartozó tagjában. Megszámlálhatóan végtelen értékészlet esetén az átrendezések jogosságát a fenti sorok abszolút konvergenciája biztosítja, így az állítást beláttuk. \square

Különös jelentősége miatt a fenti formulát külön felírjuk a $g(t) = t^2$ függvény esetén. Ha $\mathbb{E}(X^2)$ véges, akkor ennek értékét az

$$(2.21) \quad \mathbb{E}(X^2) = \sum_{t \in \text{ran } X} t^2 \cdot f_X(t) = \sum_{t \in \text{ran } X} t^2 \cdot \mathbb{P}(X = t)$$

formula adja meg.

2.3.8. Példa. Egy urnában 2 piros és 3 fehér golyó van. Visszatevés nélkül húzunk 2 golyót, jelölje X a kihúzott piros golyók számát. Megadjuk az $\mathbb{E}(X^2)$ várható értéket. Ehhez írjuk fel az X eloszlását:

$$\mathbb{P}(X = 0) = \frac{\binom{3}{2}}{\binom{5}{2}} = \frac{3}{10}, \quad \mathbb{P}(X = 1) = \frac{\binom{2}{1} \cdot \binom{3}{1}}{\binom{5}{2}} = \frac{6}{10} = \frac{3}{5}, \quad \mathbb{P}(X = 2) = \frac{\binom{2}{2}}{\binom{5}{2}} = \frac{1}{10}.$$

A fenti tétel szerint tehát

$$\mathbb{E}(X^2) = 0^2 \cdot \mathbb{P}(X = 0) + 1^2 \cdot \mathbb{P}(X = 1) + 2^2 \cdot \mathbb{P}(X = 2) = 1 \cdot \frac{3}{5} + 4 \cdot \frac{1}{10} = \frac{3}{5} + \frac{2}{5} = 1.$$

A fenti példában voltaképp könnyedén meghatározhattuk volna az X^2 eloszlását is, és az alapján is számolhattunk volna várható értéket. A 2.3.7. tétel által adott formula azonban azért kényelmes, mert sokszor előfordul, hogy az X eloszlása eleve adott, és ebből így $\mathbb{E}(g(X))$ mechanikusan számolható. Továbbá, bár ebben a példában a $t \mapsto t^2$ függvény az X és az X^2 értékészlete között egy kölcsönösen egyértelmű megfeleltetést ad, ez általában nincs így, ami a transzformált eloszlásának számolását akár kényelmetlenné is teheti.

2.3.12. Gyakorlat. Legyen X egy véletlenszerűen választott hónap sorszáma, áprilistól decemberig. Tegyük fel, hogy annak az esélye, hogy az i -edik hónapot választjuk ki, éppen $\frac{i}{72}$ ($i = 4, \dots, 12$). Legyen $Y = (-1)^X$. Határozzuk meg $\mathbb{E}(Y)$ -t az Y ill. az X eloszlásával számolva is.

Térjünk most rá a nevezetes eloszlásokra. A binomiális eloszlás esetében az $\mathbb{E}(X^2)$ értékének felhasználását voltaképp ki fogjuk küszöbölni a következő szakaszban, ezért annak meghatározását feladatként tűzzük ki. Kiszámoljuk viszont ezt a várható értéket a geometriai eloszlás esetén.

2.3.13. Feladat. (M) Legyen $X \sim \text{Bin}(n; p)$, határozzuk meg az $\mathbb{E}(X^2)$ várható értéket. Az előző szakaszban látottakhoz hasonlóan számoljuk ki a fenti várható értéket indikátorváltozók segítségével és a transzformált várható értékére vonatkozó formula segítségével is.

2.3.9. Példa. Legyen $X \sim \text{Geo}(p)$, meghatározzuk az $\mathbb{E}(X^2)$ értéket. A (2.21) formula által adott sor konvergenciája belátható pl. úgy, ahogy az X várható értékénél eljártunk. Az érték meghatározásához viszont nem közvetlenül fogjuk ezt a formulát használni, hanem ahogy az előző szakaszban a várható érték kiszámolásánál is, a teljes várható érték tételét fogjuk alkalmazni az $\{X = 1\}$ és $\{X > 1\}$ eseményekből álló teljes eseményrendszerrel. A szemléletes érveléshez választunk egy konkrét X -et: legyen ez a kísérletek száma egy kísérletsorozatban, ahol a kísérleteket egymástól függetlenül végezzük egy p valószínűségű esemény bekövetkezéséig.

Most $\mathbb{E}(X^2 \mid X = 1) = 1$ nyilván teljesül, továbbá ha $X > 1$, az azt jelenti, hogy az első kísérlet sikertelen volt, utána pedig az eloszlás örökifjúsága miatt lényegében a kezdeti szituációból indulunk, tehát

$$(2.22) \quad \mathbb{E}(X^2 \mid X > 1) = \mathbb{E}((X + 1)^2) = \mathbb{E}(X^2) + 2\mathbb{E}(X) + 1.$$

A teljes várható érték tétele alapján

$$\begin{aligned} \mathbb{E}(X^2) &= \mathbb{E}(X^2 \mid X = 1)\mathbb{P}(X = 1) + \mathbb{E}(X^2 \mid X > 1)\mathbb{P}(X > 1) \\ &= 1 \cdot p + (\mathbb{E}(X^2) + 2\mathbb{E}(X) + 1)(1 - p) = (1 - p)\mathbb{E}(X^2) + \frac{2(1 - p)}{p} + 1 \\ &= (1 - p)\mathbb{E}(X^2) + \frac{2 - p}{p}, \end{aligned}$$

ezt az egyenletet átrendezve pedig $\mathbb{E}(X^2) = \frac{2 - p}{p^2}$ adódik.

A fenti szemléletes érvelést (ahogy azt az előző szakaszban az X várható értékének kiszámolásánál is tettük) könnyű precíz formális levezetéssé alakítani a transzformált várható

értékére vonatkozó tétel segítségével, azt a $\mathbb{P}(\cdot | X > 1)$ feltételes mértékre felírva:

$$\begin{aligned}\mathbb{E}(X^2 | X > 1) &= \sum_{k=1}^{\infty} k^2 \cdot \mathbb{P}(X = k | X > 1) \\ &= 1^2 \cdot \mathbb{P}(X^2 = 1 | X > 1) + \sum_{k=2}^{\infty} k^2 \cdot \mathbb{P}(X = k | X > 1) \\ &= 1 \cdot 0 + \sum_{k=2}^{\infty} k^2 \cdot \mathbb{P}(X = k - 1) = \sum_{k=1}^{\infty} (k + 1)^2 \cdot \mathbb{P}(X = k).\end{aligned}$$

Most az utolsó összegben a zárójelet felbontva, az összeget három külön összegre bontva, ezután a transzformált várható értékére vonatkozó tételt, a várható érték definícióját ill. a súlyfüggvény értékeinek összegéről tanultakat figyelembe véve adódik a (2.22) egyenlet.

Számos esetben előfordul, hogy több változó segítségével definiálunk újabbakat, ilyenkor tehát egy vektorváltozó transzformáltját tekintjük. Az egyik legegyszerűbb, a későbbiekben fontos szerepet játszó érték két változó szorzatának a várható értéke, így először ennek kiszámolását részletezzük.

Legyenek tehát $X, Y : \Omega \rightarrow \mathbb{R}$ ugyanazon valószínűségi mezőn értelmezett valószínűségi változók. Ekkor $XY : \Omega \rightarrow \mathbb{R}$ is egy valószínűségi változó, és amennyiben ennek várható értéke véges, akkor definíció szerint ez (az összeg várható értékének levezetésében látottakhoz hasonlóan)

$$\begin{aligned}\mathbb{E}(XY) &= \sum_{m \in \text{ran } XY} m \cdot \mathbb{P}(XY = m) = \sum_{m \in \text{ran } XY} m \cdot \mathbb{P}\left(\bigcup_{st=m} \{X = s, Y = t\}\right) \\ &= \sum_{m \in \text{ran } XY} m \cdot \sum_{st=m} \mathbb{P}(X = s, Y = t) = \sum_{s \in \text{ran } X} \sum_{t \in \text{ran } Y} st \cdot \mathbb{P}(X = s, Y = t) \\ (2.23) \quad &= \sum_{s \in \text{ran } X} \sum_{t \in \text{ran } Y} st \cdot f_{X,Y}(s, t).\end{aligned}$$

2.3.10. Példa. Tekintsük ismét a 2.2.1. példában látott valószínűségi változókat, amelyekre tehát $\text{ran } X = \{2, 3, 5\}$ és $\text{ran } Y = \{0, 1, 2\}$, továbbá az együttes eloszlásuk

	X			
Y		2	3	5
0		0,05	0,15	0,1
1		0,1	0,2	0,1
2		0,05	0,2	0,05

A fenti formula szerint

$$\begin{aligned}\mathbb{E}(XY) &= 2 \cdot 0 \cdot 0,05 + 3 \cdot 0 \cdot 0,15 + 5 \cdot 0 \cdot 0,1 \\ &\quad + 2 \cdot 1 \cdot 0,1 + 3 \cdot 1 \cdot 0,2 + 5 \cdot 1 \cdot 0,1 \\ &\quad + 2 \cdot 2 \cdot 0,05 + 3 \cdot 2 \cdot 0,2 + 5 \cdot 2 \cdot 0,05 \\ &= 0,2 + 0,6 + 0,5 + 0,2 + 1,2 + 0,5 = 3,2.\end{aligned}$$

Egy fontos speciális esetben, nevezetesen független változók esetén a szorzat várható értéke könnyen kiszámolható az egyes tényezők várható értékéből. Ugyanis ekkor $f_{X,Y} = f_X \cdot f_Y$, azaz a fentiek szerint

$$\begin{aligned}\mathbb{E}(XY) &= \sum_{s \in \text{ran } X} \sum_{t \in \text{ran } Y} st \cdot f_{X,Y}(s,t) = \sum_{s \in \text{ran } X} \sum_{t \in \text{ran } Y} st \cdot f_X(s)f_Y(t) \\ &= \left(\sum_{s \in \text{ran } X} s \cdot f_X(s) \right) \left(\sum_{t \in \text{ran } Y} t \cdot f_Y(t) \right) = \mathbb{E}(X)\mathbb{E}(Y).\end{aligned}$$

Fontos megjegyezni, hogy a fenti átalakítások csak akkor helyesek, ha a jobb oldalon szereplő mindkét várható érték véges (azoknak a végességéből viszont már következik a szorzat várható értékének végessége, legalábbis független változók esetén). Az utóbbi technikai megjegyzésben foglaltakat is figyelembe véve beláttuk tehát a következőt:

2.3.8. Állítás. *Ha X és Y független diszkrét valószínűségi változók, továbbá az $\mathbb{E}(X)$ és $\mathbb{E}(Y)$ várható értékek végesek, akkor $\mathbb{E}(XY)$ is létezik, és $\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y)$ teljesül.*

A fenti tétel állítása természetesen általában (vagyis a függetlenségi feltétel nélkül) nem érvényes. Az olvasó a 2.3.10. (és 2.2.1.) példában szereplő marginális eloszlásokat kiszámolva könnyen meggyőződhet arról, hogy ebben az esetben $\mathbb{E}(XY) \neq \mathbb{E}(X)\mathbb{E}(Y)$ (összhangban a 2.2.1. gyakorlat eredményével, tehát azzal, hogy a változók nem függetlenek).

Megjegyezzük, hogy egy X változó várható értékének végességéből általában nem következik például az X^2 várható értékének végessége. Ugyanígy, az X és Y valószínűségi változók várható értékének végességéből általában nem következik az XY várható értékének a végessége, vagy még nagyobb általánosságban fogalmazva, nem következik bármilyen $g(X)$ vagy $h(X, Y)$ transzformált várható értékének a végessége (itt $g : \mathbb{R} \rightarrow \mathbb{R}$ és $h : \mathbb{R}^2 \rightarrow \mathbb{R}$ tetszőleges függvények). A jegyzet hátralévő részében viszont túlnyomórészt olyan változókat vizsgálunk, ahol ilyen problémák nem merülnek fel.

Végezetül megadjuk a 2.3.7. tételnek a vektorváltozókra vonatkozó analogonját:

2.3.9. Tétel. *Legyen $\underline{X} : \Omega \rightarrow \mathbb{R}^n$ egy diszkrét valószínűségi vektorváltozó, $h : \mathbb{R}^n \rightarrow \mathbb{R}$ pedig egy függvény, melyre az $\mathbb{E}(h(\underline{X}))$ várható érték véges. Ekkor*

$$\mathbb{E}(h(\underline{X})) = \sum_{\underline{t} \in \text{ran } \underline{X}} h(\underline{t})f_{\underline{X}}(\underline{t}) = \sum_{\underline{t} \in \text{ran } \underline{X}} h(\underline{t})\mathbb{P}(\underline{X} = \underline{t}).$$

A fenti tétel bizonyítása nem tartalmaz lényeges új gondolatot a 2.3.7. tétel bizonyításához képest, csak a technikai részletek kopmplikáltabbak, ezért a bizonyítás helyett inkább néhány példára koncentrálnak. Ha $n = 2$ és X, Y valószínűségi változók, akkor ezeknek egy $h(X, Y)$ transzformáltjára példa a $h(X, Y) = XY$ transzformált is, és vegyük észre, hogy az $\mathbb{E}(XY)$ várható értékre adott (2.23) formula pontosan a 2.3.9. tétel állítása ebben a speciális esetben.

2.3.11. Példa. A szakasz zárásaként még visszatérünk a 2.1.11. példában vázolt szituációhoz: egy kalapba beteszünk n cédulát, melyekre rendre az $1, \dots, n$ számokat írjuk, majd visszatevés nélkül kihúzzunk belőlük két darabot. Legyen X értéke a kihúzott cédulákra írt számok közül a kisebbik, míg Y értéke a nagyobbik szám. A 2.1.11. példában kiszámoltuk az X/Y hányados eloszlását, most pedig meghatározzuk a várható értékét. Ehhez immár több lehetséges út közül választhatunk, használhatjuk az X/Y változó (2.7) formulával megadott eloszlását, de akár a 2.3.9. tételt is.

Elsőként az utóbbit választjuk. Vegyük észre, hogy az $f_{X,Y}$ súlyfüggvény nagyon egyszerű, hiszen minden lehetséges $(a; b)$ számpár, ahol $a < b$, ugyanakkora, nevezetesen $1/\binom{n}{2}$ valószínűséggel adódik, így a tétel által adott formula viszonylag kényelmes számolásra vezet:

$$\begin{aligned}\mathbb{E}(X/Y) &= \sum_{1 \leq a < b \leq n} \frac{a}{b} \cdot \frac{1}{\binom{n}{2}} = \sum_{b=2}^n \sum_{a=1}^{b-1} \frac{a}{b} \cdot \frac{1}{\binom{n}{2}} = \frac{1}{\binom{n}{2}} \sum_{b=2}^n \frac{1}{b} \sum_{a=1}^{b-1} a \\ &= \frac{1}{\binom{n}{2}} \sum_{b=2}^n \frac{1}{b} \cdot \frac{(b-1)b}{2} = \frac{1}{2\binom{n}{2}} \sum_{b=2}^n (b-1) = \frac{1}{2\binom{n}{2}} \sum_{b=1}^{n-1} b = \frac{1}{2\binom{n}{2}} \cdot \frac{n(n-1)}{2} = \frac{1}{2}.\end{aligned}$$

Annak demonstrálására, hogy a 2.3.9. tétel mennyire egyszerűsítette le a számolást, kiszámoljuk ugyanezt a várható értéket annak definíciója alapján az X/Y eloszlásából. A változó lehetséges értékei az a/b alakú számok, ahol $1 \leq a < b \leq n$, továbbá a és b relatív prímek. Most a várható érték definíciója és a (2.7) formula szerint

$$\mathbb{E}(X/Y) = \sum_{1 \leq a < b \leq n, (a,b)=1} \frac{a}{b} \cdot \frac{\lfloor \frac{n}{b} \rfloor}{\binom{n}{2}} = \frac{1}{\binom{n}{2}} \sum_{b=2}^n \frac{1}{b} \cdot \lfloor \frac{n}{b} \rfloor \sum_{1 \leq a < b, (a,b)=1} a.$$

Mivel $(a, b) = 1$ esetén $(b-a, b) = 1$ is teljesül, így a jobb oldalon szereplő belső összeg valójában $\frac{b \cdot \varphi(b)}{2}$, ahol φ az Euler-féle φ -függvény, tehát a (2.8) egyenlőséget is felhasználva

$$\mathbb{E}(X/Y) = \frac{1}{2\binom{n}{2}} \sum_{b=2}^n \lfloor \frac{n}{b} \rfloor \varphi(b) = \frac{1}{2\binom{n}{2}} \cdot \binom{n}{2} = \frac{1}{2}.$$

A kapott eredmény talán némileg meglepő: a fenti számolások egyike sem mutatja igazán, hogy átlagosan miért épp a $[0; 1]$ intervallum közepét adják ezek a törtek. Hogy ezt megvilágítsuk, térjünk vissza az

$$\mathbb{E}(X/Y) = \sum_{1 \leq a < b \leq n, (a,b)=1} \frac{a}{b} \cdot \mathbb{P}\left(\frac{X}{Y} = \frac{a}{b}\right) = \sum_{1 \leq a < b \leq n, (a,b)=1} \frac{a}{b} \cdot \frac{\lfloor \frac{n}{b} \rfloor}{\binom{n}{2}}$$

formulához. Vegyük észre, hogy a fenti összegben lévő valószínűség pusztán a b számtól függ, a -tól nem. Figyelembe véve, hogy $(a, b) = 1$ esetén $(b-a, b) = 1$ is teljesül, azt kapjuk, hogy az $\frac{a}{b}$ és az $1 - \frac{a}{b}$ értékeket azonos valószínűséggel veszi fel az X/Y változó. Ezek az értékek az $\frac{1}{2}$ számra szimmetrikusan helyezkednek el a számegyenesen, és mivel ez tetszőleges $1 \leq a < b \leq n$, $(a, b) = 1$ esetén teljesül, azt kapjuk, hogy az eloszlás *szimmetrikus az $\frac{1}{2}$ -re*. Ebből adódóan

$$\mathbb{E}(X/Y) = \sum_{1 \leq a < b \leq n, (a,b)=1} \frac{a}{b} \cdot \mathbb{P}\left(\frac{X}{Y} = \frac{a}{b}\right) = \sum_{1 \leq a < b \leq n, (a,b)=1} \left(1 - \frac{a}{b}\right) \cdot \mathbb{P}\left(\frac{X}{Y} = \frac{a}{b}\right),$$

és a fenti két összeget összeadva

$$2\mathbb{E}(X/Y) = \sum_{1 \leq a < b \leq n, (a,b)=1} \mathbb{P}\left(\frac{X}{Y} = \frac{a}{b}\right) = 1, \quad \mathbb{E}(X/Y) = \frac{1}{2}.$$

Megjegyezzük, hogy teljes várható érték tételének segítségével egy ettől különböző szemléletes érveléssel is alátámaszthatjuk ezt az eredményt:

2.3.14. Feladat. Számoljuk ki az utóbbi példában szereplő $\mathbb{E}(X/Y)$ várható értéket a teljes várható érték tételének segítségével is. (Segítség: mi lesz ez a várható érték, ha fixáljuk a nagyobbik számot?)

2.4. Variancia és szórás

A várható érték leírja ugyan egy valószínűségi változó átlagos viselkedését, az átlag azonban arról semmit nem mond, hogy az értékek ettől milyen messze esnek. Ha például egy változó a ± 10 értékeket veszi fel $\frac{1}{2} - \frac{1}{2}$ valószínűséggel, akkor bár a várható értéke 0 lesz, de a változó értékei ettől lényegesen eltérnek.

A *szórás* ill. a *szórásnégyzet* (vagy más néven *variancia*) olyan, az eloszlásra jellemző fontos számszerű adatok, amik a várható értéktől való eltérést mérik. A két érték közti kapcsolat már az elnevezésből is látszik: a szórás a szórásnégyzet nemnegatív gyöke. E két különböző fogalom használatának az az oka, hogy míg az abszolút eltérésnél jobban kezelhető az átlagos négyzetes eltérés, az eltérés mértékegysége viszont a szórásnál egyezik meg az eredeti mértékegységgel. Ez később a statisztikai példákban is látható lesz.

2.4.1. Definíció. Legyen X egy diszkrét valószínűségi változó, melynek véges a várható értéke. Ha az $\mathbb{E}((X - \mathbb{E}(X))^2)$ várható érték is véges, akkor ezt az X *szórásnégyzetének* (vagy *varianciájának*) nevezzük, és $\mathbb{D}^2(X)$ -el jelöljük. A szórásnégyzet nemnegatív gyökét az X *szórásának* nevezzük és $\mathbb{D}(X)$ -el jelöljük.

Egy valószínűségi változó szórásnégyzete tehát (ha véges) a várható értékétől vett négyzetes eltéréseinek a várható értéke. Így tehát egy nemnegatív valószínűségi változó várható értékeként maga is nemnegatív, ezért valóban van értelme a nemnegatív négyzetgyökekről beszélni (azaz a fenti definíció értelmes).

Megjegyezzük még, hogy egy egyszerű valószínűségi változó esetén a várható érték és a szórásnégyzet is mindig véges, ekkor tehát ezzel a technikai kérdéssel nem kell foglalkozni. Ha azonban az értékészlet végtelen, akkor már óvatosabbnak kell lenni, de az ezzel kapcsolatos apróbb technikai részletekbe az alábbiakban nem fogunk belemenni (azon túl, hogy az állításokat pontosan fogalmazzuk meg).

2.4.1. Állítás. Legyen X egy diszkrét valószínűségi változó, amelynek véges a szórásnégyzete. Ekkor tetszőleges $c \in \mathbb{R}$ esetén

$$\mathbb{D}(X + c) = \mathbb{D}(X), \quad \mathbb{D}(cX) = |c| \mathbb{D}(X).$$

Bizonyítás. Az első állítást elegendő az X szórásnégyzetére bizonyítani, míg a második állítás esetében a vele ekvivalens $\mathbb{D}^2(cX) = c^2 \mathbb{D}^2(X)$ állítást igazoljuk. Tekintsünk a c számra egy konstans valószínűségi változóként (ami minden $\omega \in \Omega$ -hoz a c értéket rendeli), ekkor persze $c = \mathbb{E}(c)$ érvényes. Ezért

$$\begin{aligned} \mathbb{D}^2(X) &= \mathbb{E}((X - \mathbb{E}(X))^2) = \mathbb{E}((X + c - [\mathbb{E}(X) + \mathbb{E}(c)])^2) \\ &= \mathbb{E}((X + c - \mathbb{E}(X + c))^2) = \mathbb{D}^2(X + c), \end{aligned}$$

ahol a várható érték linearitását és a szórásnégyzet definícióját használtuk. Hasonlóan,

$$\begin{aligned} \mathbb{D}^2(cX) &= \mathbb{E}((cX - \mathbb{E}(cX))^2) = \mathbb{E}((cX - c\mathbb{E}(X))^2) \\ &= \mathbb{E}(c^2(X - \mathbb{E}(X))^2) = c^2 \mathbb{E}((X - \mathbb{E}(X))^2) = c^2 \mathbb{D}^2(X), \end{aligned}$$

ahol a második és a negyedik egyenlőségnél használtuk a várható érték linearitását. \square

A fenti definíció ugyan szemléletes, de a szórásnégyzet kiszámolásához leggyakrabban a következő formulát használjuk:

2.4.2. Állítás. Legyen X egy diszkrét valószínűségi változó. Az X -nek pontosan akkor véges a szórásnégyzete, ha $\mathbb{E}(X^2)$ véges, és ekkor

$$\mathbb{D}^2(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2.$$

A végesség kérdését (a korábbi megjegyzéseinkkel összhangban) nem részletezzük, egyetlen fontos megjegyzésen kívül: az iménti állításból adódóan $\mathbb{E}(X^2)$ vagy $\mathbb{D}^2(X)$ végessége maga után vonja $\mathbb{E}(X)$ végességét! A fenti formula egyszerűen adódik a következő számolásból (használva a várható értékek végességét és a várható érték linearitását):

$$\begin{aligned} \mathbb{D}^2(X) &= \mathbb{E}((X - \mathbb{E}(X))^2) = \mathbb{E}(X^2 - 2\mathbb{E}(X) \cdot X + \mathbb{E}(X)^2) \\ &= \mathbb{E}(X^2) - 2\mathbb{E}(X) \cdot \mathbb{E}(X) + \mathbb{E}(X)^2 = \mathbb{E}(X^2) - \mathbb{E}(X)^2. \end{aligned}$$

2.4.1. Példa. Legyen X egy kockadobás eredménye. A 2.3.1. példában kiszámoltuk, hogy $\mathbb{E}(X) = \frac{7}{2}$ és $\mathbb{E}(X^2) = \frac{91}{6}$, így

$$\mathbb{D}^2(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2 = \frac{91}{6} - \frac{49}{4} = \frac{35}{12}, \quad \mathbb{D}(X) \approx 1,71.$$

2.4.2. Példa. Legyen A egy esemény, melyre $\mathbb{P}(A) = p$, és legyen $\mathbb{1}_A$ a hozzá tartozó indikátorváltozó. Mivel $\mathbb{1}_A$ csak a 0 és a 1 értékeket veszi fel, így valójában $\mathbb{1}_A^2 = \mathbb{1}_A$, tehát

$$\mathbb{D}^2(\mathbb{1}_A) = \mathbb{E}(\mathbb{1}_A^2) - \mathbb{E}(\mathbb{1}_A)^2 = \mathbb{E}(\mathbb{1}_A) - \mathbb{E}(\mathbb{1}_A)^2 = p - p^2 = p(1 - p).$$

Ahogy azt már korábban megtapasztaltuk, egy binomiális eloszlású változót indikátorváltozók összegeként felírva az eloszlás várható értékének meghatározása lényegesen leegyszerűsödik a várható érték linearitásának köszönhetően. Ezt az utat szeretnénk tehát követni ezen eloszlás szórásnégyzetének esetén is, azonban a gond az, hogy egy összeg szórásnégyzete általában *nem* a szórásnégyzetek összege. Szerencsére bizonyos speciális esetekben ez a probléma egyszerűen elhárul, ugyanis (például) független változókra a kívánt tulajdonság teljesül:

2.4.3. Állítás. Ha X és Y független diszkrét valószínűségi változók, melyekre $\mathbb{D}^2(X)$ és $\mathbb{D}^2(Y)$ véges, akkor $\mathbb{D}^2(X + Y) = \mathbb{D}^2(X) + \mathbb{D}^2(Y)$.

Bizonyítás. Először is megjegyezzük, hogy a szórásnégyzet definíciója ill. a 2.4.2. állítás szerint a szórásnégyzetek végességéből $\mathbb{E}(X)$, $\mathbb{E}(Y)$, $\mathbb{E}(X^2)$, $\mathbb{E}(Y^2)$ végessége is következik. Mivel továbbá az X és Y változók függetlenek, ezért a 2.3.8. állítás szerint $\mathbb{E}(XY)$ is véges, és $\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y)$ teljesül.

A 2.4.2. állításban szereplő formula alapján

$$\begin{aligned} \mathbb{D}^2(X + Y) &= \mathbb{E}((X + Y)^2) - \mathbb{E}(X + Y)^2 \\ &= \mathbb{E}(X^2 + 2XY + Y^2) - (\mathbb{E}(X) + \mathbb{E}(Y))^2 \\ &= \mathbb{E}(X^2) + 2\mathbb{E}(XY) + \mathbb{E}(Y^2) - (\mathbb{E}(X)^2 + 2\mathbb{E}(X)\mathbb{E}(Y) + \mathbb{E}(Y)^2) \\ &= \mathbb{E}(X^2) - \mathbb{E}(X)^2 + \mathbb{E}(Y^2) - \mathbb{E}(Y)^2 + 2(\mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y)) \\ &= \mathbb{E}(X^2) - \mathbb{E}(X)^2 + \mathbb{E}(Y^2) - \mathbb{E}(Y)^2 + 0 = \mathbb{D}^2(X) + \mathbb{D}^2(Y), \end{aligned}$$

ahol a várható érték linearitásának alkalmazásánál felhasználtuk az egyes várható értékek végességét is. \square

A fentihez nagyon hasonló számolások megadják az analóg állítást n változó esetén is:

2.4.4. Állítás. Ha X_1, \dots, X_n páronként független diszkrét valószínűségi változók, melyek szórásnégyzete véges, akkor

$$\mathbb{D}^2(X_1 + \dots + X_n) = \mathbb{D}^2(X_1) + \dots + \mathbb{D}^2(X_n).$$

Figyeljük meg, hogy a fenti állításban elegendő a változóknak a páronkénti függetlenségét feltenni, tehát nem kell, hogy együttesen függetlenek legyenek. Ennek oka könnyen látszik, ha végiggondoljuk a bizonyítás menetét. Az előző bizonyításban látott számolást az n tagú összegre végrehajtva az első lépésében ugyanis az $(X_1 + \dots + X_n)^2$ várható értéke, illetve az $(\mathbb{E}(X_1) + \dots + \mathbb{E}(X_n))^2$ kifejezés jelenik meg, ami miatt az $\mathbb{E}(X_i X_j)$ ill. $\mathbb{E}(X_i)\mathbb{E}(X_j)$ tagok keletkeznek, ezek pedig már páronkénti függetlenség esetén is kiejtik egymást.

2.4.3. Példa. A fentiek alkalmazásaként kiszámoljuk a binomiális eloszlás szórásnégyzetét. Legyen $X \sim B(n; p)$, ekkor X eloszlása megegyezik egy $\mathbb{1}_{A_1} + \dots + \mathbb{1}_{A_n}$ összeg eloszlásával, ahol A_1, \dots, A_n együttesen független, p valószínűségű események, $\mathbb{1}_{A_1}, \dots, \mathbb{1}_{A_n}$ pedig az ezekhez tartozó indikátor valószínűségi változók. A 2.2.1. állításban láttuk, hogy az A_i események együttes függetlensége esetén ezek az indikátorok is együttesen (és így persze páronként is) függetlenek, ezért (felhasználva a 2.4.2. példa eredményét is)

$$\mathbb{D}^2(X) = \mathbb{D}^2(\mathbb{1}_{A_1} + \dots + \mathbb{1}_{A_n}) = \mathbb{D}^2(\mathbb{1}_{A_1}) + \dots + \mathbb{D}^2(\mathbb{1}_{A_n}) = np(1 - p).$$

2.4.4. Példa. Végezetül kiszámoljuk egy $X \sim Geo(p)$ geometriai eloszlású változó szórásnégyzetét. A munka orozslánrészét már elvégeztük a 2.3.5. és a 2.3.9. példákban. Ezek eredményei és a 2.4.2. állítás szerint

$$\mathbb{D}^2(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2 = \frac{2-p}{p^2} - \frac{1}{p^2} = \frac{1-p}{p^2}.$$

Gyakorlatok, feladatok

2.4.1. Gyakorlat. Az X és Y diszkrét valószínűségi változók együttes eloszlását tartalmazza az alábbi táblázat. Számoljuk ki a $\mathbb{D}^2(X)$ és a $\mathbb{D}^2(Y)$ szórásnégyzeteket. Határozzuk meg az $X + Y$ változó eloszlását és szórásnégyzetét is.

	X		
Y		-1	1
	-1	1/10	1/5
	1	3/10	2/5

2.4.2. Gyakorlat. Egy oktaéder alakú szabályos dobókocka lapjai az $1, 2, \dots, 8$ számokkal vannak számozva. Dobjunk ezzel a dobókockával kétszer egymástól függetlenül. Határozzuk meg az a dobott számok összegének várható értékét és szórást.

2.4.3. Gyakorlat. Legyen $X \sim Geo(\frac{1}{3})$. Adjuk meg az $\mathbb{E}((3 - X)^2)$, a $\mathbb{D}(5 - 2X)$ és az $\mathbb{E}((X + 1)(X - 2))$ mennyiségeket.

2.5. Generátorfüggvények (kiegészítő anyag)

Ebben a szakaszban egy hasznos eszközt mutatunk be, amely számos probléma kezelésénél jó szolgálatot tesz. Ilyenek pl. a várható érték és a szórásnégyzet kiszámítása, független

valószínűségi változók összegének vizsgálata, az ún. elágazó folyamatok vizsgálata, eloszlások közelítése stb. Az alapgondolat az, hogy egy valószínűségi eloszlást leíró információt egy könnyen kezelhető matematikai objektumban "kódolunk el". Erre többféle lehetőség is adódik, mi most az eloszlás generátorfüggvényével ismerkedünk meg:

2.5.1. Definíció. Legyen $\{a_k\}_{k \in \mathbb{N}}$ egy valós számsorozat, ekkor a hozzá tartozó *generátorfüggvényt* az

$$A(z) = \sum_{k=0}^{\infty} a_k z^k$$

hatványsor definiálja, amennyiben ez konvergens valamilyen $(-r; r)$ intervallumon, ahol $r > 0$ valós szám. Speciálisan, ha $a_k = f_X(k)$ valamilyen X valószínűségi változóra, amely nemnegatív egész értékeket vesz fel, akkor az ehhez a sorozathoz tartozó

$$P_X(z) = \sum_{k=0}^{\infty} f_X(k) z^k = \sum_{k=0}^{\infty} \mathbb{P}(X = k) z^k$$

függvényt az X változó *generátorfüggvényének* nevezzük.

Ahhoz, hogy egy változó generátorfüggvényével dolgozni tudjunk, először is meg kell mutatnunk, hogy az őt definiáló hatványsor valamilyen $(-r; r)$ intervallumon konvergál. Valóban, $|z| \leq 1$ esetén

$$|P_X(z)| \leq \sum_{k=0}^{\infty} f_X(k) |z|^k \leq \sum_{k=0}^{\infty} f_X(k) = 1,$$

így a $P_X(z)$ hatványsor konvergenciasugara legalább 1. Megjegyezzük, hogy ebből a számolásból, valamint a 2.3.7. tétel bizonyításából adódóan minden fix $|z| \leq 1$ esetén $P_X(z) = \mathbb{E}(z^X)$ teljesül. Valóban, a $P(z)$ -t definiáló sor abszolút konvergenciája garantálja a tétel bizonyításában lévő számolás helyességét, amely éppen a fenti egyenlőséget adja a $g(X) = z^X$ választással. Gyakran éppen az $\mathbb{E}(z^X)$ várható értékkel definiálják az X változó generátorfüggvényét.

A hatványsorok általános elméletéből ismertek a következők:

2.5.1. Tétel. Legyen X egy valószínűségi változó, amely csak nemnegatív értéket vesz fel. Ekkor a $P_X(z)$ generátorfüggvény folytonos a $[-1; 1]$ zárt intervallumon, végtelenszer differenciálható a $(-1; 1)$ intervallumon, továbbá minden $z \in (-1; 1)$ esetén

$$P_X^{(n)}(z) = \sum_{k=n}^{\infty} k(k-1) \dots (k-n+1) f_X(k) z^{k-n},$$

ahol $P_X^{(n)}(z)$ jelöli a generátorfüggvény n -edik deriváltjának értékét a z helyen ($n = 0$ esetén magát a $P_X(z)$ függvényt). Következésképp

$$f_X(k) = \frac{P_X^{(k)}(0)}{k!}$$

érvényes minden $k \in \mathbb{N}$ esetén, vagyis a generátorfüggvény egyértelműen meghatározza az X eloszlását. Minden $n \in \mathbb{N}$ esetén érvényes továbbá a

$$(2.24) \quad \lim_{z \rightarrow 1^-} P_X^{(n)}(z) = \sum_{k=n}^{\infty} k(k-1) \dots (k-n+1) f_X(k)$$

egyenlőség, ahol a bal oldalon az 1-ben vett bal oldali határérték szerepel, továbbá az egyenlőséget úgy értjük, hogy amennyiben egyik oldal véges, akkor a másik is, és ekkor egyenlőség teljesül, ha pedig egyik oldal ∞ , akkor a másik is.

A fenti tétel egyes állításainak bizonyítása megtalálható pl. a [9] könyv 8. fejezetében (vagy bármely más bevezető analízis könyvben), a maradék állítások pedig az ott bizonyított tételek egyszerű, közvetlen következményei. Mindezekből egyszerűen adódik az alábbi

2.5.2. Következmény. *Legyen X egy valószínűségi változó, amely csak nemnegatív értéket vesz fel, és legyen a generátorfüggvénye $P_X(z)$. Ekkor $\mathbb{E}(X) < \infty$ pontosan akkor teljesül, ha $\lim_{z \rightarrow 1^-} P'_X(z) < \infty$, ahol $P'_X(z)$ a generátorfüggvény deriváltja, és ebben az esetben*

$$\lim_{z \rightarrow 1^-} P'_X(z) = P'_X(1) = \mathbb{E}(X).$$

Továbbá $\mathbb{D}(X^2) < \infty$ pontosan akkor teljesül, ha $\lim_{z \rightarrow 1^-} P''_X(z) = P''_X(1) < \infty$, ahol $P''_X(z)$ a generátorfüggvény második deriváltja. Ebben az esetben $P'_X(1) = \mathbb{E}(X) < \infty$, és

$$\mathbb{D}^2(X) = P''_X(1) + P'_X(1) - P'^2(1).$$

Bizonyítás. Az első állítás közvetlenül adódik a (2.24) formulából az $n = 1$ választással. Ha $\mathbb{D}^2(X) < \infty$, akkor a 2.4.2. állítás szerint $\mathbb{E}(X) < \infty$ és így a tétel első része szerint $P'_X(1) = \mathbb{E}(X)$, valamint $\mathbb{E}(X^2) < \infty$ is teljesül. Továbbá

$$\mathbb{E}(X^2) - \mathbb{E}(X) = \sum_{k=1}^{\infty} k^2 f_X(k) - \sum_{k=1}^{\infty} k f_X(k) = \sum_{k=1}^{\infty} (k^2 - k) f_X(k) = \lim_{z \rightarrow 1^-} P''_X(z) = P''_X(1)$$

a (2.24) formulát $n = 2$ választással alkalmazva (ahol persze a számolásban azt is kihasználjuk, hogy a fenti értékek mind végesek). Tehát a 2.4.2. állítás szerint

$$\mathbb{D}^2(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2 = \mathbb{E}(X^2) - \mathbb{E}(X) + \mathbb{E}(X) - \mathbb{E}(X)^2 = P''_X(1) + P'_X(1) - P'^2(1),$$

ezzel az állítást beláttuk. \square

A fentiek első alkalmazásaként most meghatározzuk néhány nevezetes eloszlás generátorfüggvényét, és ezek segítségével (újra) kiszámoljuk a az eloszlások várható értékét és szórását.

2.5.1. Példa. Legyen $X \sim \text{Bin}(n; p)$, ahol $n > 0$ egész és $p \in [0; 1]$, ekkor (mivel $k > n$ esetén $f_X(k) = 0$)

$$P_X(z) = \sum_{k=0}^n f_X(k) z^k = \sum_{k=0}^n \binom{n}{k} p^k (1-p)^{n-k} z^k = \sum_{k=0}^n \binom{n}{k} (pz)^k (1-p)^{n-k} = (pz + 1 - p)^n,$$

ahol az utolsó egyenlőségénél a binomiális tételt használtuk. Továbbá

$$P'_X(z) = np(pz + 1 - p)^{n-1}, \quad P''_X(z) = n(n-1)p^2(pz + 1 - p)^{n-2}$$

(a második formula $n = 1$ esetén is érvényes, kivéve a $p = 1, z = 0$ esetet). Most a 2.5.2. következmény szerint

$$\mathbb{E}(X) = P'_X(1) = np(p \cdot 1 + 1 - p)^{n-1} = np,$$

továbbá

$$\mathbb{D}^2(X) = P''_X(1) + P'_X(1) - P'^2(1) = n(n-1)p^2 + np - n^2p^2 = np - np^2 = np(1-p).$$

2.5.2. Példa. Legyen $X \sim Geo(p)$, ahol $p \in (0; 1)$, ekkor $|z| \leq 1$ esetén

$$P_X(z) = \sum_{k=1}^{\infty} f_X(k)z^k = \sum_{k=1}^{\infty} (1-p)^{k-1}pz^k = pz \sum_{j=0}^{\infty} (1-p)^j z^j = \frac{pz}{1-(1-p)z}.$$

Tehát

$$P'_X(z) = \frac{p(1-(1-p)z) + pz(1-p)}{(1-(1-p)z)^2} = \frac{p}{(1-(1-p)z)^2},$$

így

$$\mathbb{E}(X) = P'_X(1) = \frac{p}{(1-(1-p) \cdot 1)^2} = \frac{p}{p^2} = \frac{1}{p}.$$

2.5.1. Gyakorlat. (M) Határozzuk meg a geometriai eloszlás szórásnégyzetét a generátorfüggvénye segítségével.

A következő tétel különösen kényelmes eszközzé teszi a generátorfüggvényeket, amikor független változók összegének eloszlását vizsgáljuk.

2.5.3. Tétel. Legyenek X_1, \dots, X_n független, nemnegatív egész értékű valószínűségi változók, melyek generátorfüggvényei $P_{X_1}(z), \dots, P_{X_n}(z)$. Legyen még $Y = X_1 + \dots + X_n$ a változók összege $P_Y(z)$ generátorfüggvénnyel. Ekkor

$$P_Y(z) = P_{X_1}(z) \dots P_{X_n}(z) = \prod_{k=1}^n P_{X_k}(z)$$

érvényes minden $z \in [0; 1]$ esetén.

Bizonyítás. Teljes indukcióval bizonyítunk, és a tétel állítását először két változóra igazoljuk. Legyenek tehát X_1 és X_2 nemnegatív egész értékű független változók $P_{X_1}(z)$ ill. $P_{X_2}(z)$ generátorfüggvényekkel. Ekkor

$$\begin{aligned} P_{X_1}(z)P_{X_2}(z) &= \left(\sum_{k=0}^{\infty} f_{X_1}(k)z^k \right) \left(\sum_{l=0}^{\infty} f_{X_2}(l)z^l \right) = \sum_{m=0}^{\infty} z^m \left(\sum_{k=0}^m f_{X_1}(k)f_{X_2}(m-k) \right) \\ &= \sum_{m=0}^{\infty} (f_{X_1} * f_{X_2})(m) \cdot z^m = \sum_{m=0}^{\infty} f_{X_1+X_2}(m)z^m = P_{X_1+X_2}(z) \end{aligned}$$

a 2.2.2. állítás szerint.

Legyenek most X_1, \dots, X_n függetlenek, ahol $n > 2$, és tegyük fel, hogy az állítás teljesül n -nél kevesebb valószínűségi változóra. Ekkor az $X_1 + \dots + X_{n-1}$ és X_n változók függetlenek, így a két változóra bizonyított egyenlőség ill. az indukciós feltevés szerint

$$P_{X_1+\dots+X_n}(z) = P_{X_1+\dots+X_{n-1}}(z)P_{X_n}(z) = P_{X_1}(z) \dots P_{X_{n-1}}(z)P_{X_n}(z),$$

ezzel az állítást beláttuk. □

2.5.3. Példa. A fenti tétel használatának illusztrálásaként (ismét) kiszámoljuk független binomiális eloszlású változók összegének eloszlását. Legyenek tehát $X \sim Bin(n; p)$ valamint $Y \sim Bin(m; p)$ független változók, ekkor a 2.5.1. példa szerint ezen változók generátorfüggvényei

$$P_X(z) = (pz + 1 - p)^n, \quad P_Y(z) = (pz + 1 - p)^m,$$

tehát a 2.5.3. tétel szerint az $X + Y$ változó generátorfüggvénye

$$P_{X+Y}(z) = P_X(z)P_Y(z) = (pz + 1 - p)^n(pz + 1 - p)^m = (pz + 1 - p)^{n+m}.$$

Ez utóbbi függvény viszont éppen egy $Bin(n + m; p)$ eloszlású változó generátorfüggvényével egyezik meg, és mivel a generátorfüggvény egyértelműen meghatározza az eloszlást, így következésképp $X + Y \sim Bin(n + m; p)$.

Eloszlások közelítése

A következőkben több példát is fogunk látni olyan szituációra, amikor egy bonyolult formulákkal megadott eloszlást más, a gyakorlati számítások szempontjából praktikusabb eloszlással közelíthetünk. Talán a legfontosabb példa erre a normális eloszlással való közelítés (lásd a 3.5.3. szakaszt), ehhez azonban előbb a folytonos eloszlások definíciójára van szükség. Van azonban egy diszkrét eloszlás, amely sok szituációban jól használható más eloszlások közelítésére. Ez a ún. Poisson-eloszlás, amely többek között az imént említett tulajdonsága miatt is számos gyakorlati alkalmazásban kap fontos szerepet, és amellyel a következő szakaszban foglalkozunk részletesen.

Általában azt mondhatjuk, hogy egy X diszkrét változó akkor közelít (más szóval approximál) jól egy Y diszkrét változót, ha a $\mathbb{P}(X = k)$ valószínűségek minden szóba jövő k értékre közel vannak a $\mathbb{P}(Y = k)$ valószínűségekhez. Tipikusan valamilyen eloszláscsaládból származó X_n sorozatot fogunk tekinteni, ahol n maga is az eloszláscsalád egy paramétere, amit viszonylag nagyoknak fogunk választani. Amennyiben a $\mathbb{P}(X_n = k)$ valószínűségek tartanak a $\mathbb{P}(Y = k)$ valószínűséghez, ha n tart a végtelenhez, akkor nagy n értékekre ezek a valószínűségek már közel lesznek egymáshoz.

2.5.2. Definíció. Legyenek $Y, X_1, \dots, X_n, \dots$ nemnegatív értékű diszkrét valószínűségi változók. Ekkor azt mondjuk az X_n változók sorozata *eloszlásban konvergál* az Y változóhoz, ha $\mathbb{P}(X_n = k) \rightarrow \mathbb{P}(Y = k)$ teljesül minden k -ra, ha $n \rightarrow \infty$. Ennek jele $X_n \xrightarrow{d} Y$.

A változók értékészletének megszorítására a fenti definícióban valójában nem feltétlenül volna szükség, de mivel a diszkrét esetben csak ilyen példákkal találkozunk, továbbá az alábbi eredmény is ilyen esetekre vonatkozik, így átmenetileg megelégszünk ezzel a speciális definícióval. Később ezt a fogalmat ennél sokkal általánosabban is be fogjuk vezetni.

A generátorfüggvények fontos szerepét az approximációknál a következő tétel biztosítja:

2.5.4. Tétel. Legyenek $Y, X_1, \dots, X_n, \dots$ nemnegatív értékű diszkrét valószínűségi változók, melyeknek generátorfüggvényei $P_Y(z), P_{X_1}(z), \dots, P_{X_n}(z), \dots$. Ekkor a következő állítások ekvivalensek:

- (i) $\mathbb{P}(X_n = k) \rightarrow \mathbb{P}(Y = k)$ minden $k \in \mathbb{N}$ -re, ha $n \rightarrow \infty$;
- (ii) $\mathbb{P}(X_n \in A) \rightarrow \mathbb{P}(Y \in A)$ minden $A \subset \mathbb{N}$ -re, ha $n \rightarrow \infty$;
- (iii) $P_{X_n}(z) \rightarrow P_Y(z)$ minden $z \in [0; 1]$ -re, ha $n \rightarrow \infty$;
- (iv) $P_{X_n}(z) \rightarrow P_Y(z)$ minden $z \in [0; \eta)$ esetén valamilyen $\eta \in (0; 1)$ -re ha $n \rightarrow \infty$.

Bizonyítás. (i) \Rightarrow (ii) Legyen $A \subset \mathbb{N}$. Azt fogjuk kihasználni, hogy elég nagy N -re az X_n -ek értéke csak kis valószínűséggel esik N fölé, így az $A \cap [N; \infty)$ halmazba is csak kis valószínűséggel eshet. Ha pedig N -et fixáljuk, akkor az $A \cap [0; N - 1]$ halmaz végessége miatt a $\mathbb{P}(X_n = k) \rightarrow \mathbb{P}(Y = k)$ állítást már csak véges sok k -ra kell használnunk.

Fixáljunk tehát egy $\varepsilon > 0$ számot, és válasszuk az $N \in \mathbb{N}$ számot olyan nagyra, hogy $\mathbb{P}(Y \in [N; \infty)) < \varepsilon/4$ teljesüljön. Ilyen N létezik, hiszen

$$(2.25) \quad \mathbb{P}(Y \in [N; \infty)) = \sum_{k \in \text{ran } Y \cap [N; \infty)} f_Y(k) = \sum_{k=N}^{\infty} \mathbb{P}(Y = k),$$

és mivel a $\sum_{k=0}^{\infty} \mathbb{P}(Y = k)$ sor konvergens (az értéke 1), így (2.25)-ben a jobb oldalon álló összeg 0-hoz tart.

Most minden $0 \leq k < N$ esetén legyen n_k olyan, hogy $|\mathbb{P}(Y = k) - \mathbb{P}(X_n = k)| < \varepsilon/(4N)$ teljesüljön minden $n > n_k$ esetén. Ha most $n > m := \max\{n_0, \dots, n_{N-1}\}$, akkor

$$\sum_{k=0}^{N-1} |\mathbb{P}(Y = k) - \mathbb{P}(X_n = k)| < N \cdot \frac{\varepsilon}{4N} = \frac{\varepsilon}{4}.$$

Következésképp $n > m$ esetén

$$\begin{aligned} \sum_{k=N}^{\infty} \mathbb{P}(X_n = k) &= \left| 1 - \sum_{k=0}^{N-1} \mathbb{P}(X_n = k) + \sum_{k=0}^{N-1} \mathbb{P}(Y = k) - \sum_{k=0}^{N-1} \mathbb{P}(Y = k) \right| \\ &\leq \left| 1 - \sum_{k=0}^{N-1} \mathbb{P}(Y = k) \right| + \sum_{k=0}^{N-1} |\mathbb{P}(Y = k) - \mathbb{P}(X_n = k)| \\ &= \sum_{k=N}^{\infty} \mathbb{P}(Y = k) + \sum_{k=0}^{N-1} |\mathbb{P}(Y = k) - \mathbb{P}(X_n = k)| < \frac{\varepsilon}{4} + \frac{\varepsilon}{4} = \frac{\varepsilon}{2}, \end{aligned}$$

és így

$$\begin{aligned} |\mathbb{P}(Y \in A) - \mathbb{P}(X_n \in A)| &= |\mathbb{P}(Y \in A \cap [0; N-1]) + \mathbb{P}(Y \in A \cap [N; \infty)) - \\ &\quad - \mathbb{P}(X_n \in A \cap [0; N-1]) - \mathbb{P}(X_n \in A \cap [N; \infty))| \\ &\leq \sum_{k=0}^{N-1} |\mathbb{P}(Y = k) - \mathbb{P}(X_n = k)| + \sum_{k=N}^{\infty} \mathbb{P}(X_n = k) + \sum_{k=N}^{\infty} \mathbb{P}(Y = k) \\ &< \frac{\varepsilon}{4} + \frac{\varepsilon}{2} + \frac{\varepsilon}{4} = \varepsilon. \end{aligned}$$

Ez pedig éppen azt jelenti, hogy $\mathbb{P}(X_n \in A) \rightarrow \mathbb{P}(Y \in A)$, ha $n \rightarrow \infty$.

(ii) \Rightarrow (i) Ha $k \in \mathbb{N}$, akkor alkalmazzuk a (ii) állítást az $A = \{k\}$ eseményre.

(i) \Rightarrow (iii) Lényegében lemásoljuk az (i) \Rightarrow (ii) rész bizonyítását annyi különbséggel, hogy az A eseményre itt nincs is szükségünk (vagy úgy is tekinthetjük, hogy az $A = \mathbb{N}$ választással élünk).

Legyen $z \in [0; 1]$, ekkor választhatunk egy olyan N egészt, amelyre

$$\sum_{k=N}^{\infty} f_Y(k) z^k \leq \sum_{k=N}^{\infty} \mathbb{P}(Y = k) < \frac{\varepsilon}{4}.$$

Hasonlóképp, elég nagy n -re

$$\sum_{k=0}^{N-1} |f_Y(k) z^k - f_{X_n}(k) z^k| \leq \sum_{k=0}^{N-1} |\mathbb{P}(Y = k) - \mathbb{P}(X_n = k)| < \frac{\varepsilon}{4}.$$

Ilyen n -ekre

$$\sum_{k=N}^{\infty} f_{X_n}(k)z^k \leq \sum_{k=N}^{\infty} \mathbb{P}(X_n = k) < \frac{\varepsilon}{2},$$

tehát

$$|P_Y(z) - P_{X_n}(z)| \leq \sum_{k=0}^{N-1} |f_Y(k)z^k - f_{X_n}(k)z^k| + \sum_{k=N}^{\infty} f_{X_n}(k)z^k + \sum_{k=N}^{\infty} f_Y(k)z^k < \varepsilon.$$

(iii) \Rightarrow (iv) Triviális.

(iv) \Rightarrow (i) Azt kell belátnunk, hogy $f_{X_n}(k) \rightarrow f_Y(k)$ teljesül minden k -ra, ha $n \rightarrow \infty$. A k szerinti teljes indukcióval bizonyítunk. Mivel $P_{X_n}(0) = f_{X_n}(0)$, $P_Y(0) = f_Y(0)$, és (iv) szerint $P_{X_n}(0) \rightarrow P_Y(0)$, ha $n \rightarrow \infty$, $k = 0$ -ra az állítás teljesül.

Legyen most $k > 0$, és tegyük fel, hogy k -nál kisebb értékekre az állítás teljesül. Defináljuk a

$$Q_{X_n,k}(z) = \frac{P_{X_n}(z) - \sum_{j=0}^{k-1} f_{X_n}(j)z^j}{z^k} = \sum_{j=0}^{\infty} f_{X_n}(j+k)z^j$$

ill. a

$$Q_{Y,k}(z) = \frac{P_Y(z) - \sum_{j=0}^{k-1} f_Y(j)z^j}{z^k} = \sum_{j=0}^{\infty} f_Y(j+k)z^j$$

függvényeket. Ekkor (iv), valamint az indukciós feltevés miatt minden $0 < z < \eta$ számra $Q_{X_n,k}(z) \rightarrow Q_{Y,k}(z)$ teljesül, ha $n \rightarrow \infty$. Továbbá, ha $0 < \delta < \eta$, akkor

$$|Q_{X_n,k}(0) - Q_{X_n,k}(\delta)| = \sum_{j=1}^{\infty} f_{X_n}(j+k)\delta^j \leq \sum_{j=1}^{\infty} \delta^j = \frac{\delta}{1-\delta},$$

továbbá hasonlóképp $|Q_{Y,k}(0) - Q_{Y,k}(\delta)| \leq \frac{\delta}{1-\delta}$, így ha $\varepsilon > 0$ egy rögzített kicsi szám, és a δ értékét elég kicsire választjuk, akkor a fenti különbségek $\varepsilon/3$ -nál kisebbek (n -től függetlenül). Fixáljunk egy ilyen $\delta > 0$ -t, és legyen most n_0 olyan, hogy $|Q_{X_n,k}(\delta) - Q_{Y,k}(\delta)| < \varepsilon/3$ teljesüljön minden $n > n_0$ esetén. Ekkor

$$\begin{aligned} |f_{X_n}(k) - f_Y(k)| &= |Q_{X_n,k}(0) - Q_{Y,k}(0)| \\ &\leq |Q_{X_n,k}(0) - Q_{X_n,k}(\delta)| + |Q_{X_n,k}(\delta) - Q_{Y,k}(\delta)| + |Q_{Y,k}(\delta) - Q_{Y,k}(0)| < 3 \cdot \frac{\varepsilon}{3} = \varepsilon. \end{aligned}$$

Mivel ε tetszőleges volt, $f_{X_n}(k) \rightarrow f_Y(k)$ teljesül, ha $n \rightarrow \infty$, és ezzel az állítást beláttuk. \square

Gyakorlatok, feladatok

2.5.2. Feladat. Legyen X egy nemnegatív egész értékű valószínűségi változó, és legyen $q_k = \mathbb{P}(X > k)$ minden $k \in \mathbb{N}$ esetén. Mutassuk meg, hogy a $\{q_k\}_{k \in \mathbb{N}}$ sorozathoz tartozó $Q(z)$ generátorfüggvény konvergál a $(-1; 1)$ intervallumon. Mutassuk meg továbbá, hogy minden $z \in (-1; 1)$ esetén $Q(z) = \frac{1-P_X(z)}{1-z}$, ahol $P_X(z)$ az X változó generátorfüggvénye.

2.5.3. Feladat. Egy urnában N darab páronként különböző színű golyó van. Visszatevéssel húzunk véletlenszerűen az urnából egészen addig, amíg a kihúzott golyók közt már r különböző színű szerepel ($0 < r \leq N$). Legyen X_r a húzások száma.

a) Írjuk fel X_r -t független geometriai eloszlású változók összegeként, és határozzuk meg az X_r várható értékét, szórásnégyzetét és generátorfüggvényét.

b) Mutassuk meg, hogy

$$\mathbb{P}(X_r = k + 1) = \frac{r - 1}{N} \mathbb{P}(X_r = k) + \frac{N - r + 1}{N} \mathbb{P}(X_{r-1} = k)$$

érvényes minden $1 \leq k$ és $1 < r \leq N$ esetén. Határozzuk meg ismét az X_r generátorfüggvényét ennek az összefüggésnek a segítségével.

2.5.4. Feladat. (M) Tekintsük a 2.1.9. feladat második példáját, azaz dobáljunk egy pénzérmét addig, amíg egymás után 3 egyforma eredményt nem kapunk. Adjuk meg a dobások számának generátorfüggvényét, és számoljuk ki a segítségével a dobások számának várható értékét és szórásnégyzetét.

2.6. Nevezetes diszkrét eloszlások

Az alábbiakban további, eddig nem tárgyalt nevezetes eloszlásokkal ismerkedünk meg. Fontossága miatt a sort a Poisson-eloszlással kezdjük, amely a gyakorlat szempontjából is kiemelt szereppel bír, hiszen számos szituációban jól használható modellt ad. Ennek egyik oka az, hogy több más eloszlás ún. határeloszlásaként éppen a Poisson-eloszlást kapjuk, amire példákat is fogunk látni a szakasz során. Ezután a fejezet zárásaként megismerkedünk a negatív binomiális eloszlással.

2.6.1. A Poisson-eloszlás

A Poisson-eloszlás gyakran jó modellt ad, ha nagyon sok egymástól független, kis valószínűségű esemény esetén azt vizsgáljuk, hogy ezekből hány darab következett be. Ilyen példák lehetnek a következők:

- egy biztosítóhoz bejelentett káresek száma egy év alatt: ebben az esetben rengeteg ügyfele lehet a biztosítónak, melyek mindegyike lényegében a többitől függetlenül, kis valószínűséggel jelent be káresetet;
- nyomtatási hibák száma egy könyvben, ahol nagyon sok karakter található, és ezeket egymástól függetlenül kis eséllyel nyomtatják rosszul;
- hozzáférések száma egy szerver esetén egy adott időszak alatt: rengeteg felhasználó van, akik általában (lényegében) függetlenül, kis valószínűséggel szeretnének hozzáférni adatokhoz.

A fenti megfogalmazás persze több ponton sem precíz, nem világos, hogy mit értünk "nagyon sok" esemény ill. "kis" valószínűség alatt, és a következő példa mutatja, hogy ezzel óvatosnak kell lennünk. Találkoztunk már korábban is olyan szituációval, amikor n független eseményből a bekövetkezők számát vizsgáltuk, nevezetesen a binomiális eloszlásnál, ahol egy kísérletsorozatról beszéltünk, és egy konkrét p valószínűségű esemény bekövetkezéseinek számát tekintettük. Bármilyen kicsi is a p valószínűség, amennyiben pozitív, úgy az eloszlás np várható értéke végtelenhez tart, ha $n \rightarrow \infty$, tehát az átlagosan bekövetkező sikerek száma az n -nel együtt minden határon túl nő.

Nem várhatjuk tehát, hogy ezt a szituációt egyetlen konkrét eloszlás leírja nagy n esetén. A fenti helyzet elkerülése érdekében megválaszthatjuk a p értékét az n függvényében, azonban itt is vigyázni kell, hiszen ha túl kis p -ket választunk, akkor előfordulhat, hogy az np várható érték 0-hoz tart, ami határeloszlásként azt a triviális eloszlást adná, amelynél 1 valószínűséggel 0 a bekövetkezett események száma.

Van viszont egy köztes eset, amikor egymáshoz viszonyítva n nagy, p kicsi, de az np szorzat se nem kicsi, se nem nagy. Egészen pontosan az n számhoz úgy választjuk a p_n valószínűséget, hogy a $\lambda_n = np_n$ szorzat egy pozitív λ számhoz konvergáljon, ha $n \rightarrow \infty$. Ilyenkor az n és p_n paraméterekkel adott binomiális eloszlás jól közelíthető egy meghatározott eloszlással, amelyhez valójában eloszlásban konvergál (lásd a 2.5.2. definíciót). Egy gyakori választás a $p_n = \lambda/n$ valószínűség valamely pozitív λ esetén, hiszen ilyenkor az $np_n = \lambda$ szorzat konstans (és így konvergens).

2.6.1. Tétel. Legyen $p_n \in [0; 1]$ egy olyan sorozat, amelyre $np_n \rightarrow \lambda \in (0; \infty)$, ha $n \rightarrow \infty$. Legyen továbbá $S_n \sim \text{Bin}(n; p_n)$ egy binomiális eloszlású valószínűségi változó n és p_n paraméterekkel, ekkor

$$(2.26) \quad \lim_{n \rightarrow \infty} \mathbb{P}(S_n = k) = e^{-\lambda} \frac{\lambda^k}{k!}$$

teljesül minden $k \in \mathbb{N}$ esetén. Speciálisan, a fenti egyenlőség teljesül $p_n = \lambda/n$ esetén, ahol $\lambda > 0$ egy (fix) pozitív valós szám.

Először is vegyük észre, hogy a (2.26) egyenlet jobb oldalán szereplő értékek egy valószínűségeloszlást adnak meg. Valóban, ha $f(k) = e^{-\lambda} \frac{\lambda^k}{k!}$ minden $k \in \mathbb{N}$ -re, akkor

$$\sum_{k=0}^{\infty} f(k) = \sum_{k=0}^{\infty} e^{-\lambda} \frac{\lambda^k}{k!} = e^{-\lambda} \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} = e^{-\lambda} e^{\lambda} = 1,$$

ahol felhasználtuk az e^{λ} sorfejtését, tehát a 2.1.4. állítás utáni második bekezdésben tett megjegyzés szerint $f(k)$ egy valószínűségi változó súlyfüggvénye, amelynek értékei a nemnegatív egész számok.

2.6.1. Definíció. Legyen X egy olyan valószínűségi változó, melyre $\text{ran } X = \mathbb{N}$, továbbá minden $k \in \mathbb{N}$ esetén $f_X(k) = \mathbb{P}(X = k) = e^{-\lambda} \frac{\lambda^k}{k!}$ valamely $\lambda > 0$ számra. Ekkor azt mondjuk, hogy X Poisson-eloszlású λ paraméterrel, ennek jelölése $X \sim \text{Pois}(\lambda)$.

A fenti és a 2.5.2. definíció alapján tehát úgy fogalmazhatunk, hogy a 2.6.1. tételben szereplő S_n változók eloszlásban konvergálnak egy $X \sim \text{Pois}(\lambda)$ Poisson-eloszlású változóhoz (emlékeztetünk ennek jelölésére: $S_n \xrightarrow{d} X$).

A 2.6.1. tétel bizonyítása. Legyen $k \in \mathbb{N}$, a binomiális eloszlás súlyfüggvényének értéke

$$\begin{aligned} \mathbb{P}(S_n = k) &= \binom{n}{k} \cdot p_n^k \cdot (1 - p_n)^{n-k} = \frac{n(n-1) \dots (n-k+1)}{k!} \cdot p_n^k \cdot (1 - p_n)^{n-k} \\ &= \frac{n(n-1) \dots (n-k+1)}{n^k} \cdot \frac{(np_n)^k}{k!} \cdot \left(1 - \frac{np_n}{n}\right)^n \cdot (1 - p_n)^{-k} \end{aligned}$$

A fenti szorzatban az első tört két 1 főgyütthetős k -adfokú polinom hányadosa, amely így 1-hez tart, ha $n \rightarrow \infty$ (bármely fix k -ra). Mivel továbbá $np_n \rightarrow \lambda$, így $(np_n)^k/k! \rightarrow \lambda^k/k!$, ha $n \rightarrow \infty$.

Az utolsó tényező határértékéhez vegyük észre, hogy $p_n = np_n/n$, és az utóbbi tört számlálója korlátos (mert konvergens), így tehát n -nel a végtelenbe tartva $p_n \rightarrow 0$, és így $(1 - p_n)^{-k} \rightarrow 1^{-k} = 1$.

Hátra van még a harmadik tényező határértékének meghatározása. Ha ott np_n helyére ennek határértékét, azaz λ -t íránk, akkor ez egy nevezetes sorozat volna $e^{-\lambda}$ határértékkel. Mivel np_n nagy n esetén tetszőlegesen közel van λ -hoz, így az itt szereplő sorozatnak is lesz a határértéke, ami a fentiekkel együtt éppen az állítást adja.

Mindez precízen a következőképp látható: legyen $\varepsilon > 0$ tetszőleges. Legyen továbbá n olyan nagy, hogy $p_n < 1$, $\frac{\lambda}{n} < 1$, $|\lambda - np_n| < \varepsilon/2$ és végül

$$(2.27) \quad \left| \left(1 - \frac{\lambda}{n}\right)^n - e^{-\lambda} \right| < \frac{\varepsilon}{2}$$

teljesüljenek. Ekkor a háromszög-egyenlőtlenség miatt

$$\left| \left(1 - \frac{np_n}{n}\right)^n - e^{-\lambda} \right| \leq \left| \left(1 - \frac{np_n}{n}\right)^n - \left(1 - \frac{\lambda}{n}\right)^n \right| + \left| \left(1 - \frac{\lambda}{n}\right)^n - e^{-\lambda} \right|.$$

A jobb oldalon a második tag az n választása miatt legfeljebb $\varepsilon/2$, továbbá

$$\begin{aligned} & \left| \left(1 - \frac{np_n}{n}\right)^n - \left(1 - \frac{\lambda}{n}\right)^n \right| = \\ & = \left| 1 - \frac{np_n}{n} - \left(1 - \frac{\lambda}{n}\right) \right| \cdot \left| \left(1 - \frac{np_n}{n}\right)^{n-1} + \left(1 - \frac{np_n}{n}\right)^{n-2} \left(1 - \frac{\lambda}{n}\right)^1 + \dots + \left(1 - \frac{\lambda}{n}\right)^{n-1} \right| \\ & = \frac{|\lambda - np_n|}{n} \left\{ (1 - p_n)^{n-1} + (1 - p_n)^{n-2} \left(1 - \frac{\lambda}{n}\right) + \dots + \left(1 - \frac{\lambda}{n}\right)^{n-1} \right\} \leq \frac{\varepsilon/2}{n} \cdot n \cdot 1 = \frac{\varepsilon}{2}, \end{aligned}$$

tehát a (2.27) egyenlőtlenség bal oldala legfeljebb ε , és így az állítást beláttuk. \square

A fenti tételt pl. a $p_n = \lambda/n$ valószínűségekkel alkalmazva legalább szemléletesen világos, hogy mivel a binomiális eloszlású S_n változók átlagos értéke az $np_n = \lambda$ konstans, és az eloszlásuk a λ paraméterű Poisson eloszláshoz tart, így ez utóbbi eloszlás várható értéke is λ kell legyen. Mindezt persze precíz számolással is alátámaszthatjuk: ha $X \sim Pois(\lambda)$, akkor

$$\begin{aligned} \mathbb{E}(X) &= \sum_{k=0}^{\infty} k \cdot e^{-\lambda} \cdot \frac{\lambda^k}{k!} = e^{-\lambda} \sum_{k=0}^{\infty} k \cdot \frac{\lambda^k}{k!} = e^{-\lambda} \sum_{k=1}^{\infty} k \cdot \frac{\lambda^k}{k!} \\ &= \lambda \cdot e^{-\lambda} \sum_{k=1}^{\infty} \frac{\lambda^{k-1}}{(k-1)!} = \lambda \cdot e^{-\lambda} \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} = \lambda \cdot e^{-\lambda} \cdot e^{\lambda} = \lambda. \end{aligned}$$

A szórásnégyzetre is elmondható az előbbi szemléletes érvelés: az S_n valószínűségi változók szórásnégyzete $np_n(1 - p_n) = \lambda(1 - \lambda/n) \rightarrow \lambda$, ha $n \rightarrow \infty$, és nagy n -re ez a szórásnégyzet közel kell legyen a Poisson-eloszlás szórásnégyzetéhez, tehát ha $X \sim Pois(\lambda)$, akkor

$$\mathbb{D}^2(X) = \lambda$$

kell teljesüljön. Ez az állítás megint csak igazolható a várható érték esetében imént látott precíz levezetéshez hasonlóan, ezt a számolást azonban az olvasóra bízunk.

2.6.1. Feladat. Legyen $X \sim Pois(\lambda)$, számoljuk ki az $\mathbb{E}(X^2)$ várható értéket, és ennek segítségével mutassuk meg, hogy $\mathbb{D}^2(X) = \lambda$.

Megjegyezzük, hogy a fenti szemléletes érveléseinkről megmutatható, hogy valójában azok teljesen korrekt matematikai levezetések. Ennek igazolását (hangsúlyozottan nehéz) feladatként tűzzük ki, melynek a(z inkább csak a teljesség kedvéért) megadott mintamegoldásában a korábbiaknál nagyobb analitikus eszköztárat használunk.

2.6.2. Feladat.* (M) Mutassuk meg, hogy ha $X_n \xrightarrow{d} Y$, az $\mathbb{E}(X_n)$ várható értékek és $\mathbb{D}^2(X_n)$ szórásnégyzetek végesek és e két számsorozat konvergens, akkor az Y várható értéke és szórásnégyzete is véges, és $\mathbb{E}(X_n) \rightarrow \mathbb{E}(Y)$ ill. $\mathbb{D}^2(X_n) \rightarrow \mathbb{D}^2(Y)$, ha $n \rightarrow \infty$.

2.6.1. Példa. Tegyük fel, hogy egy előadáson 100 ember vesz részt. Mi a valószínűsége, hogy közülük aznap legalább egy személynek születésnapja van?

Úgy tekintjük, hogy az előadásra egymástól függetlenül érkeztek a résztvevők (például senki nem hozta magával az ikertestvérét), illetve most az egyszerűség kedvéért azt is feltesszük, hogy egy véletlenszerűen választott embernek azonos valószínűséggel esik a születésnapja az év 365 napjának bármelyikére.

Ekkor tehát a szituáció egy 100 független kísérletből álló kísérletsorozatként fogható fel: 100-szor kiválasztunk véletlenszerűen és egymástól függetlenül egy-egy embert, és megvizsgáljuk, hogy bekövetkezik-e az a $p = \frac{1}{365}$ valószínűségű esemény, hogy az illetőnek az adott napon van a születésnapja. A kísérletsorozat során a "sikerek" S száma egy $Bin(100; \frac{1}{365})$ eloszlású változó, amelynek értéke tehát azon emberek száma, akiknek az adott napon születésnapja van.

A kérdés ekkor a $\mathbb{P}(S > 0)$ valószínűség, amit persze könnyedén kiszámolhatunk áttérve az esemény komplementerére:

$$\mathbb{P}(S > 0) = 1 - \mathbb{P}(S = 0) = 1 - \left(\frac{364}{365}\right)^{100} \approx 0,2399.$$

Nézzük meg, mit ad erre a Poisson-eloszlással való közelítés. A 2.6.1. tételben az np_n szorzatok sorozat közelített a határeloszlás paraméteréhez, és bár most nem eloszlások egy sorozatáról beszélünk, logikus választás magát a $100 \cdot \frac{1}{365} = \frac{20}{73}$ szorzatot választani a közelítő eloszlás paraméterének. Ezekben az esetekben általában is ezzel a választással fogunk élni, vagyis egy $Bin(n; p)$ eloszlást egy $\lambda = np$ paraméterű Poisson-eloszlással fogunk közelíteni.

Legyen tehát $X \sim Pois(\frac{20}{73})$, ekkor a fenti valószínűsége a

$$\mathbb{P}(X > 0) = 1 - \mathbb{P}(X = 0) = 1 - e^{-20/73} \cdot \frac{\left(\frac{20}{73}\right)^0}{0!} = 1 - e^{-20/73} \approx 0,2396$$

közelítést adjuk. Látható, hogy a becslésünk ebben az esetben kifejezetten jó.

Általában a 2.6.1. tétel feltételei nem szükségszerűen adottak egy olyan szituációban, ahol a Poisson-eloszlással approximálunk, tehát az alkalmazásokban nem feltétlenül egy eloszlássorozatáról beszélünk, hanem egyetlen (pl. binomiális) eloszlást közelítünk. Felmerül ezért a kérdés, hogy a közelítés során elkövetett hibára adható-e valamilyen becslés. Szerencsére van ilyen eredmény, sőt, a súlyfüggvényértékek abszolút eltéréseinek *összegét* is becsülni tudjuk:

2.6.2. Tétel. Legyen $S \sim Bin(n; p)$ binomiális eloszlású valószínűségi változó, $\lambda = np$ és $X \sim Pois(\lambda)$, ekkor

$$\sum_{k=0}^n |\mathbb{P}(S = k) - \mathbb{P}(X = k)| \leq \frac{2\lambda^2}{n}.$$

Ez az állítás tehát a fenti példában a

$$2 \cdot \left(\frac{100}{365}\right)^2 \cdot \frac{1}{100} = \frac{200}{365^2} \approx 0,0015$$

felső becslést adja a hibára tetszőleges, az S változó segítségével kifejezhető esemény esetén!

A Poisson-eloszlást persze messze nem csak a binomiális eloszlás közelítésére alkalmas, ahogy azt már a szakasz bevezető példáiban is igyekeztünk érzékeltetni. Egy tipikus alkalmazásban például egy adott időintervallumban bekövetkező független események számát vizsgáljuk, ilyen volt a bevezető első és harmadik példája is, most pedig egy újabb hasonló példán szemléltetjük az eloszlás használatát.

2.6.2. Példa. Egy városban egy hónap alatt átlagosan 10 tüzeset történik. Mi a valószínűsége, hogy a következő héten nem történik tüzeset?

Egy adott héten a bekövetkező tüzeseteket egymástól függetlennek és kis valószínűségűnek tehetjük fel, így tehát Poisson-eloszlást használhatunk. A fenti információ a tüzesetek átlagos számát adja meg, amely tehát 4 hetente 10, azaz hetente 2,5. Ha Poisson-eloszlással szeretnénk a heti tüzesetek X számát modellezni, akkor az átlag itt a várható értéknek felel meg, amely pedig megegyezik az eloszlás paraméterével, tehát egy $X \sim Pois(2,5)$ eloszlású valószínűségi változót használhatunk.

Az, hogy egy adott héten nem történik tüzeset, éppen az $\{X = 0\}$ eseménynek felel meg, ennek valószínűsége pedig

$$\mathbb{P}(X = 0) = e^{-2,5} \approx 0,0821.$$

Előfordul, hogy nem az esetek átlagos száma az az információ, ami rendelkezésünkre áll. Tegyük fel most, hogy egy városban 4 hétből átlagosan egy olyan van, amikor nem történik tüzeset, és számoljuk ki annak a valószínűségét, hogy a következő héten kettőnél többször riasztják tűzhoz a tűzoltókat.

Milyen $Y \sim Pois(\lambda)$ eloszlással modellezhetnénk ezt a szituációt? Vegyük észre, hogy itt valószínűsége annak a valószínűségét írtuk le, hogy egy héten nem történik baleset, ez ugyanis a fentiek szerint $1/4$, mindez pedig a $\mathbb{P}(Y = 0) = e^{-\lambda} = 0,25$ egyenletben foglalható össze. Ebből már kifejezhető az eloszlás paramétere:

$$\lambda = -\ln 0,25 \approx 1,3863,$$

a keresett valószínűség pedig

$$\begin{aligned} \mathbb{P}(Y > 2) &= 1 - (\mathbb{P}(Y = 0) + \mathbb{P}(Y = 1) + \mathbb{P}(Y = 2)) \\ &= 1 - e^{-\lambda} - e^{-\lambda}\lambda - e^{-\lambda}\frac{\lambda^2}{2} \approx 1 - 0,25 \cdot \left(1 + 1,3863 + \frac{1,3863^2}{2}\right) \approx 0,1632. \end{aligned}$$

Sok esetben előfordul, hogy több különböző eloszlást kell kezelnünk egy példában. Maradva a legutolsó szituációnál, határozzuk meg, hogy a következő három hónapban várhatóan hány olyan hét lesz, amikor kettőnél több tüzeset történik.

Vegyük észre, hogy itt már nem (csak) a tüzesetek számoljuk, a kérdés azon *hetek számára* vonatkozik az elkövetkező 12 hétből, ahol egy adott esemény (nevezetesen $\{Y > 2\}$) bekövetkezik. Ez tehát egy binomiális eloszlású Z változóval modellezhető, melynek paraméterei $n = 12$ és $p = \mathbb{P}(Y > 2) \approx 0,1632$. Ennek várható értéke $\mathbb{E}(Z) = np \approx 12 \cdot 0,1632 \approx 1,95$, azaz várhatóan nagyjából két héten fog előfordulni kettőnél több tüzeset.

A Poisson-eloszlás alkalmazásának jogosságát az utóbbi (és más hasonló) példában a következő érvelés mutatja. Tegyük fel, hogy független események bekövetkezésének számát vizsgáljuk egy egységnyi hosszúságú időintervallumban. Képzeljük el, hogy a vizsgált intervallumot felosztjuk n darab diszjunkt $1/n$ hosszúságú részre. A véletlen események bekövetkezésének ideje véges sok pontot ad ezen intervallumban, és olyan szituációkat tekintünk, ahol feltehető, hogy minden részintervallumban ugyanakkora p_n valószínűséggel következik be legalább egy esemény, továbbá az, hogy egy adott részintervallumban bekövetkezik-e legalább egy esemény, független minden más intervallumtól.

Ha azt vizsgáljuk, hogy az egyes intervallumok foglaltak vagy üresek-e (vagyis hogy bekövetkezik-e ott legalább egy esemény vagy sem), akkor egy n kísérletből álló független kísérletsorozatot kapunk, és a foglalt szakaszok száma egy $X_n \sim Bin(n; p_n)$ eloszlású változó. Az $\{X_n = 0\}$ esemény azt jelenti minden n esetén, hogy egyetlen vizsgált esemény sem következett be. Ennek valószínűsége $p_0 = (1 - p_n)^n$, és ez persze minden n esetén egy pozitív (nak feltételezett) konstans (ebből valójában $p_n = 1 - \sqrt[n]{p_0}$).

Ha finomítjuk a modellt, tehát n -nel tartunk a végtelenbe, akkor $(1 - p_n)^n$ konvergens (mert konstans), de az $\ln(1 - x)$ hatványsorát használva

$$\ln p_0 = \ln(1 - p_n)^n = n \ln(1 - p_n) = -n \sum_{k=1}^{\infty} \frac{p_n^k}{k} = -np_n \sum_{k=1}^{\infty} \frac{p_n^{k-1}}{k}.$$

Mivel $p_n \rightarrow 0$, ha $n \rightarrow \infty$, így az utolsó összeg határértéke 1, azaz $np_n \rightarrow \lambda$ valamely $\lambda > 0$ számra.

Feltéve azt is, hogy különálló pontokkal dolgozunk, ezek elég nagy n esetén különböző intervallumba kerülnek, és ekkor X_n értéke már a pontok számával egyezik meg, tehát határeloszlásban a pontok számának eloszlását kapjuk, ez pedig a 2.6.1. tétel szerint a $Pois(\lambda)$ eloszlás.

Megjegyezzük, hogy olyan esetek is előfordulnak, amikor nem tehető fel, hogy a pontjaink különbözők, ekkor az ún. összetett Poisson-eloszlással dolgozhatunk, melyről bővebben pl. a [2] könyvben olvashatunk.

A Poisson-eloszlás generátorfüggvénye (kiegészítő anyag)

Kiszámoljuk a Poisson-eloszlás várható értékét és szórásnégyzetét az eloszlás generátorfüggvényének segítségével is, amely definíció szerint

$$P(z) = \sum_{k=0}^{\infty} e^{-\lambda} \frac{\lambda^k}{k!} z^k = e^{-\lambda} \sum_{k=0}^{\infty} \frac{(\lambda z)^k}{k!} = e^{-\lambda} e^{\lambda z} = e^{\lambda(z-1)},$$

ismét csak az exponenciális függvény sorfejtését használva. Ebből a 2.5.2. következmény szerint

$$\mathbb{E}(X) = P'(1) = \lambda \cdot e^{\lambda(1-1)} = \lambda,$$

illetve

$$\mathbb{D}^2(X) = P''(1) + P'(1) - P'^2(1) = \lambda^2 \cdot e^{\lambda(1-1)} + \lambda - \lambda^2 = \lambda.$$

A generátorfüggvény egyszerű alakjából fakadóan a 2.5.3. tétel alapján könnyen adódik a következő

2.6.3. Tétel. *Legyenek $X_1 \sim Pois(\lambda_1), \dots, X_n \sim Pois(\lambda_n)$ független valószínűségi változók, ekkor $X_1 + \dots + X_n \sim Pois(\lambda_1 + \dots + \lambda_n)$.*

Bizonyítás. A 2.5.3. tétel szerint az összeg generátorfüggvénye

$$P_{X_1+\dots+X_n}(z) = \prod_{i=1}^n P_{X_i}(z) = \prod_{i=1}^n e^{\lambda_i(z-1)} = e^{(\lambda_1+\dots+\lambda_n)(z-1)},$$

ez pedig éppen a $\lambda_1 + \dots + \lambda_n$ paraméterű Poisson-eloszlás generátorfüggvénye, és mivel a generátorfüggvény egyértelműen meghatározza az eloszlást, így a tételt beláttuk. \square

Gyakorlatok, feladatok

2.6.3. Gyakorlat. Egy számítógépes szervizben egy hónap húsz munkanapjából átlagosan kettőn nincsen reklamáció. Poisson-eloszlást feltételezve mennyi annak a valószínűsége, hogy egy adott napon legalább három reklamáció érkezik?

2.6.4. Gyakorlat. Egy egyetemen nagyon sok telefonkészülék van, amelyek egymástól függetlenül romlanak el azonos, egyenként kis valószínűséggel. Az év 365 napjából átlagosan 10 olyan nap van, hogy egyetlen készülék sem romlik el. Várhatóan hány telefon romlik el egy nap? Mennyi azon napok számának várható értéke, amelyeken 2 vagy 2-nél több telefon romlik el?

2.6.5. Gyakorlat. Egy városban az utakon 25% az olyan napok aránya, amikor egyetlen baleset sem történik. Rengeteg autó közlekedik, nagyságrendileg minden nap ugyanannyi, és minden autó egymástól független, egyforma valószínűséggel okoz balesetet. Mennyi annak a valószínűsége, hogy a jövő héten pontosan 2 napon lesz 1-nél több baleset?

2.6.6. Gyakorlat. Egy könyvtárban minden könyvet minden hónapban legfeljebb egy ember kölcsönöz ki, legfeljebb egyszer és csak arra a hónapra. A könyvtárba egyre több könyvet hoznak, de a látogatók száma nem növekszik, ezért ahogy telik az idő, egy-egy könyvet egyre kisebb valószínűséggel vesz ki valaki egy adott hónapban. A megfigyelés kezdetétől számított k -adik hónapban (ahol $k \geq 1$ egy egész szám) $3k$ könyv van a könyvtárban, és mindegyiket egymástól függetlenül $1/k$ valószínűséggel veszi ki valaki.

a) Határozzuk meg (pontosan) annak a valószínűségét, hogy a 68. hónapban legalább 3 könyvet vesznek ki a könyvtárból, és adjuk meg a 68. hónap során kivett könyvek számának szórását.

b) Közelítsük nagy k esetén annak a valószínűségét alkalmas approximációval, hogy a k -adik hónapban legalább 3 könyvet vesznek ki.

2.6.7. Feladat. Bizonyítsuk be az utóbbi tételt a 2.2.2. állítás segítségével is.

2.6.2. Negatív binomiális eloszlás (kiegészítő anyag)

Tekintsünk egy kísérletsorozatot, ahol a kísérleteket egymástól függetlenül végezzük, és minden esetben egy $0 < p < 1$ valószínűségű esemény bekövetkezésére várunk. Láttuk, hogy ha X_1 jelöli az első sikerig elvégzett kísérletek számát, akkor $X_1 \sim Geo(p)$. Legyen most X_r az r -edik sikerig elvégzett kísérletek száma. Mi lesz X_r eloszlása? Nyilván $r > 1$ esetén

$$\mathbb{P}(X_r = 1) = \dots = \mathbb{P}(X_r = r - 1) = 0.$$

Ha $k \geq r$ egész, akkor $X_r = k$ esetén a k -adik kísérlet sikeres kell legyen. A maradék $k - 1$ kísérletből $r - 1$ sikeres és így $k - 1 - (r - 1) = k - r$ sikertelen. A sikeres kísérletek sorszámát

$\binom{k-1}{r-1}$ -féleképp választhatjuk, és persze minden egyes választás meghatározza a sikertelen kísérletek sorszámát. Ezek egymástól függetlenül p ill. $1-p$ valószínűséggel következnek be, vagyis

$$(2.28) \quad \mathbb{P}(X_r = k) = \binom{k-1}{r-1} p^r (1-p)^{k-r}.$$

Vegyük észre, hogy az $r = 1$ esetben a fenti formula (a várakozásoknak megfelelően) a geometriai eloszlást adja.

Ha a formulákat egységesíteni szeretnénk, akkor definiálhatjuk az általánosított binomiális együtthatót a következőképp: legyen $\alpha \in \mathbb{R}$ tetszőleges valós szám, és $k \in \mathbb{N}$, ekkor

$$\binom{\alpha}{k} := \frac{\alpha(\alpha-1)\dots(\alpha-k+1)}{k!}.$$

Ennek az általánosításnak az egyik előnye az, hogy az (2.28) formula $1 \leq k < r$ esetén is érvényben marad.

Az irodalomban azonban bevett szokás, hogy az r -edik sikerig elvégzett kísérletek X_r száma helyett az r -edik sikerig bekövetkezett kudarcok Y_r számát tekintik. Mivel érvényes az $Y_r = X_r - r$ összefüggés, így megszűnik az a kényelmetlen helyzet, amit az jelentett, hogy az X_r legkisebb felvett értéke r lehet, és ez némileg egyszerűsít bizonyos számolásokat is. Az Y_r változók értékkészlete minden $r \in \mathbb{N}^+$ esetén a természetes számok \mathbb{N} halmaza, és

$$\mathbb{P}(Y_r = k) = \mathbb{P}(X_r = k+r) = \binom{k+r-1}{r-1} p^r (1-p)^k.$$

2.6.2. Definíció. Az X valószínűségi változó *negatív binomiális* eloszlású $r \in \mathbb{N}^+$ és $p \in (0; 1)$ paraméterekkel, ha

$$(2.29) \quad f_X(k) = \mathbb{P}(X = k) = \binom{k+r-1}{r-1} p^r (1-p)^k$$

teljesül minden $k \in \mathbb{N}$ esetén. Jelölése: $X \sim NB(r; p)$.

Bár a fenti gondolatmenet nagyon szemléletes, de a formulák alapján első ránézésre nem feltétlenül nyilvánvaló, hogy itt valóban egy valószínűségeloszlást definiáltunk. Ezt az alábbiakban ellenőrizni fogjuk. Itt ismét az általánosított binomiális együtthatók kerülnek képbe, ugyanis a binomiális tétel, annak legalábbis bizonyos speciális esetei nagyobb általánosságban is érvényesek. Idézzük fel (ismét) a tétel állítását: tetszőleges $a, b \in \mathbb{R}$ valós és $n \in \mathbb{N}^+$ pozitív egész számokra

$$(a+b)^n = \sum_{k=0}^n \binom{n}{k} a^k b^{n-k}.$$

(Emlékeztetünk, hogy definíció szerint $0^0 = 1$, így a formula valóban tetszőleges a, b számokra érvényes.) Ha most $k > n$ egész számokat is megengedünk a szummában, akkor az általánosított binomiális együttható definícióját használva valójában 0 értékű tagokat veszünk hozzá az eredeti összegünkhöz, így az - formálisan legalábbis - egy végtelen összeggé válik. Ahhoz persze, hogy értelmes kifejezéseket kapjunk az összegben, az is szükséges, hogy a b szám negatív hatványai is definiálva legyenek, hiszen $n-k$ értéke $k > n$ esetén negatív lesz. Ehhez elég, ha $b \neq 0$ teljesül, ekkor azonban a b^n számmal le is oszthatjuk az egyenletet, és a binomiális tételt végül az

$$(1+x)^n = \sum_{k=0}^{\infty} \binom{n}{k} x^k$$

formába írhatjuk, ahol $x = a/b$. Ez az az alak, ami (bizonyos feltételek mellett) általánosítható tetszőleges valós kitevőre:

2.6.4. Tétel (Általánosított binomiális tétel). *Legyen $\alpha \in \mathbb{R}$, ekkor minden $x \in \mathbb{C}$, $|x| < 1$ esetén*

$$(1 + x)^\alpha = \sum_{k=0}^{\infty} \binom{\alpha}{k} x^k.$$

Ezt a tételt itt nem bizonyítjuk, viszont a segítségével ellenőrizzük, hogy a (2.29) formula valóban egy valószínűségeloszlást definiál. E formula jobb oldalán álló kifejezéseket összegezve minden $k \in \mathbb{N}$ -re a következő adódik:

$$(2.30) \quad \sum_{k=0}^{\infty} \binom{k+r-1}{r-1} p^r (1-p)^k = p^r \sum_{k=0}^{\infty} \binom{k+r-1}{k} (1-p)^k,$$

továbbá a $q = 1 - p$ jelöléssel a fenti tétel szerint

$$\begin{aligned} p^{-r} = (1-q)^{-r} &= \sum_{k=0}^{\infty} \binom{-r}{k} (-q)^k = \sum_{k=0}^{\infty} \frac{-r(-r-1)\dots(-r-k+1)}{k!} (-q)^k \\ &= \sum_{k=0}^{\infty} \frac{(k-1+r)\dots(r+1)r}{k!} q^k = \sum_{k=0}^{\infty} \binom{k-1+r}{k} (1-p)^k, \end{aligned}$$

vagyis a (2.30) jobb oldala $p^r p^{-r} = 1$.

A fenti számolás egy melléktermékeként kapjuk, hogy a (2.29) formula a következő (gyakran használatos) alakba írható:

$$f_X(k) = \binom{-r}{k} (-1)^k p^r (1-p)^k.$$

Az előző formulában lévő binomiális együtthatók felső paramétere negatív, innen a "negatív binomiális" elnevezés. Vegyük észre, hogy ez utóbbi képlet valójában tetszőleges pozitív valós r esetén egy valószínűségeloszlást definiál, hiszen a fenti számolás lényeges része érvényben marad (azon részek kivételével, ahol az r a binomiális együtthatók alsó paraméterében szerepel). Az r pozitivitása ahhoz szükséges, hogy valóban nemnegatív súlyokat kapjunk. Ez jól látszik például a

$$f_X(k) = \binom{k+r-1}{k} p^r (1-p)^k$$

ekvivalens definícióból ($k \in \mathbb{N}$). Eszerint tehát a fenti képlet(ek) segítségével tetszőleges $r \in \mathbb{R}^+$ és $p \in (0; 1)$ paraméterekre definiáltuk a negatív binomiális eloszlást (bár annak az egész értékekhez tartozó szemléletes interpretációja tetszőleges valós r esetén nem adható). Az áttekinthetőség kedvéért külön kiemeljük itt ezt az általános definíciót:

2.6.3. Definíció. Az X valószínűségi változó *negatív binomiális* eloszlású $r \in \mathbb{R}^+$, $p \in (0; 1)$ paraméterekkel, ha

$$f_X(k) = \binom{-r}{k} (-1)^k p^r (1-p)^k = \binom{k+r-1}{k} p^r (1-p)^k$$

teljesül minden $k \in \mathbb{N}$ esetén. Jelölése: $X \sim NB(r; p)$.

3. fejezet

Abszolút folytonos valószínűségi változók

A diszkrét valószínűségi változók után ebben a fejezetben egy másik fontos osztállyal ismerkedünk meg: az abszolút folytonos valószínűségi változókkal. Ezekre egészen természetes példák adódnak egyszerű geometriai megfontolások alapján, amikhez azonban olyan (ún. geometriai) valószínűségi mezőkre lesz szükségünk, amelyek már nem tárgyalhatók a második fejezetben rögzített keretek között. Így a fejezet első szakaszában az ún. Bertrand-paradoxonon keresztül ismerkedünk meg a valószínűségi mezők egy újabb csoportjával, amelyekben a kimeneteleket geometriai objektumok, azaz az egyenes, a sík ill. a tér pontjai alkotják. Ezeknek segítségével olyan változókat definiálhatunk, amelyek nem diszkrét, így a valószínűségi változók fogalmát általánosabban kell majd megfogalmaznunk. Ezek eloszlásának leírására is új eszközöket vezetünk be, mégpedig a valószínűségi változók eloszlásfüggvényét, melyekkel az eloszlás mellett a változók függetlenségének fogalma is leírható.

A legfontosabb példáink azonban olyan változók lesznek, amelyeknél a diszkrét változók eloszlásának leírására használt súlyfüggvény fogalmának létezik egy analogonja, amelyet sűrűségfüggvénynek neveznek. Ezeknek a kezeléséhez a differenciál- és integrálszámítás eszköztárát kell majd mozgósítanunk. Mindennek segítségével bevezetjük majd a diszkrét esetben definiált fontos fogalmak és eloszlások, azaz a várható érték, a szórás, ill. az egyenletes és geometriai eloszlások folytonos megfelelőjét.

A fejezet második részében az egyik legjelentősebb abszolút folytonos eloszlás, az ún. normális eloszlás leírásával foglalkozunk részletesen. Ezzel kapcsolatban tárgyaljuk majd a valószínűségi számítás legfontosabb tételeinek egyikét is: a centrális határeloszlás tételét, melynek a következő fejezetben bemutatjuk néhány statisztikai alkalmazását. A szakasz végén egy másik jelentős tétel tárgyalása sem maradhat el: az ún. nagy számok törvényének segítségével precízen megfogalmazhatjuk azt, amire a valószínűség szemléletes definíciójában már építettünk, nevezetesen, hogy a valószínűség az a szám, amely körül a relatív gyakoriságok "ingadoznak". Ezzel fejezzük be a valószínűségi számítás alapjainak felépítését.

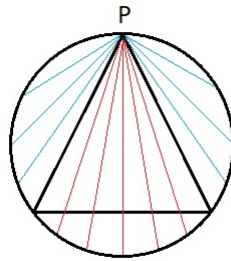
3.1. Geometriai valószínűségi mezők

Ebben a szakaszban olyan példákat vizsgálunk, melyek összekötik a geometria és valószínűségi számítás fogalmait. Amint azt alább látni fogjuk, ez a törekvésünk rövid úton ellentmondásokat szülhet, legalábbis ha nem vagyunk elég óvatosak. A következő klasszikus példán demonstráljuk, hogy a precizitást elhanyagolva hogyan vezethet tévútra az intuíció.

A Bertrand-paradoxon

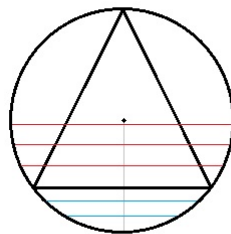
Tekintsük a következő problémát: véletlenszerűen kiválasztjuk egy egység sugarú kör egy húrját, mi a valószínűsége, hogy a húr hosszabb, mint a körbe írható szabályos háromszög egy oldala? A paradoxon elnevezést az indokolja, hogy az alábbi három érvelés különböző végeredményeket ad.

1. megoldás. Válasszuk ki először véletlenszerűen a körvonal egy P pontját, majd válasszunk ettől függetlenül egy másik Q pontot is a körvonalon. A két pont meghatároz egy véletlenszerűen választott húr. Ha a körbe berajzoljuk azt a szabályos háromszöget, melynek egyik csúcsa P (és a másik két csúcs is a körvonalon van), akkor könnyen láthatóan pontosan akkor lesz a PQ húr hosszabb a háromszög oldalánál, ha Q a háromszög P -től különböző két csúcsa által meghatározott rövidebb köríven van. Ez P -től függetlenül a körvonal $1/3$ része, így tehát a keresett valószínűség $1/3$.



3.1. ábra. A jó (piros) és rossz (kék) esetek a paradoxon első megoldásában

2. megoldás. Válasszuk ki véletlenszerűen a kör egy sugarát, majd válasszunk azon véletlenszerűen egy pontot. E kettő egyértelműen meghatároz egy húr, mégpedig azt, ami merőleges a sugárra, és keresztülmegy a választott ponton. Tekintsük azt a körbe írt szabályos háromszöget, amelynek egyik oldala merőleges a sugárra. Ez az oldal a sugarat könnyen láthatóan annak felezőpontjában metszi. A választott húr tehát akkor hosszabb a körbe írt szabályos háromszög oldalánál, ha a sugáron véletlenszerűen választott pont a sugár felezőpontja és a kör középpontja közt van. Mindez a sugár választásától függetlenül teljesül, és eszerint tehát a húr $1/2$ valószínűséggel lesz nagyobb a háromszög oldalánál.



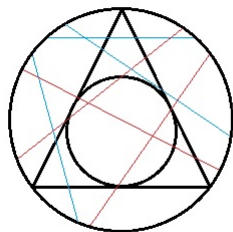
3.2. ábra. A jó (piros) és rossz (kék) esetek a paradoxon második megoldásában

3. megoldás. Válasszuk ki a körlapon egy P pontot véletlenszerűen. Ez egyértelműen meghatároz egy húr következőképp: húzzunk egy sugarat a kör középpontjából a P ponton

keresztül, és válasszuk azt a húrt, ami merőleges erre a sugárra és átmegy a P ponton. Figyeljük meg, hogy ez a gondolatmenet csak akkor érvényes, ha P és a kör középpontja különböző. Azokat a húrokat tehát ilyen módon nem kapjuk meg, amik átmennek a kör középpontján. Egy kézenfekvő megoldás erre a problémára, hogy a fent említett húrokat egyszerűen kizárjuk a vizsgálat alól, azokat tehát sosem választjuk ki. Nem nehéz meggondolni, hogy az előző két okoskodás ezzel a megszorítással ugyanolyan végeredményre vezet (ahogy később fogalmazni fogunk: pusztán egy 0 valószínűségű eseménytől tekintünk el).

A 2. megoldásban látott okoskodás azt adja, hogy ha a kör középpontja körül egy $1/2$ sugarú kört rajzolunk (amely egyben a nagyobb körbe írt szabályos háromszögek beírt köre), akkor pontosan az ezen körlapra eső P pontok adnak olyan húrt, amely hosszabb a szabályos háromszög oldalánál. A keresett valószínűség tehát a kis és a nagy kör területének aránya,

$$\text{azaz } \frac{0,5^2 \cdot \pi}{1^2 \cdot \pi} = \frac{1}{4}.$$



3.3. ábra. A jó (piros) és rossz (kék) esetek a paradoxon harmadik megoldásában

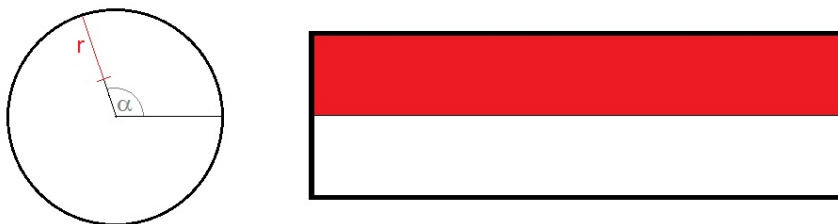
Melyik megoldás a jó a fentiek közül? A válasz valójában az, hogy mindegyik. A különböző végeredmények abból adódnak, hogy a három esetben különbözőképp generáljuk a húrt, azaz valójában három teljesen különböző problémát kezeltünk.

Visszafele haladunk a különböző megoldások elemzésében. A harmadiknál van a legegyszerűbb dolgunk. A húr választása valójában megegyezik egy pont választásával a körlapon (pontosabban a kilyukasztott körlapon, hiszen a középpontot nem választhatjuk). A jó eseteket az azonos középpontú $1/2$ sugarú körlap pontjai reprezentálják, valószínűségként pedig e két körlap területének arányát tekintjük annak megfelelően, mint mikor egy véges eseménytérben a "jó eset/összes eset" képletet alkalmaztuk a klasszikus valószínűségi mérték definiálásánál.

A második megoldásnál két pontot választunk, egyet a körvonalon, egyet pedig az általa meghatározott sugáron. A második pont választása függ az elsőtől, de valójában a választásunk két függetlenül választott számmal is leírható. Az egyszerűség kedvéért tegyük fel, hogy a kör középpontja az origó. Ekkor a körvonalon választott pont megadható a szöggel, amit a választott pontba húzott sugár és az x tengely egymással bezár, azaz egy $\alpha \in [0; 2\pi)$ intervallumból választott számmal. Ha ezen felül még választunk egy számot a $r \in [0; 1]$ intervallumból, és ezt mint távolságot felmérjük a választott pontból az origóba húzott sugárra, akkor így megkapjuk a második pontunkat (és egyben a húrt is).

Ahelyett, hogy két számról beszéljünk, azt is mondhatjuk, hogy a $[0; 2\pi) \times [0; 1]$ téglalapon választunk véletlenszerűen egy pontot (a két szám pedig ennek két koordinátája lesz). A jó esetek azok a számpárok lesznek, ahol a második koordináta a $(0,5; 1]$ intervallumban van, ami épp azt jelenti, hogy a pont a sugáron a felezőpont és a középpont közt van. Azaz

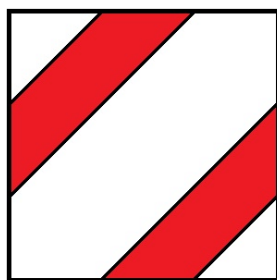
a jó eseteket éppen a $[0; 2\pi) \times (0,5; 1]$ téglalap írja le. Ennek területét és a nagy téglalap területével osztva pedig éppen az $1/2$ valószínűség adódik.



3.4. ábra. A második megoldásnál egy téglalap egy pontját választjuk ki véletlenszerűen

Az első megoldásnál két pontot választunk a körvonalon egymástól függetlenül, egyenletesen véletlenszerűen. Ismét feltéve, hogy a kör középpontja az origó, mindkét pont egyértelműen jellemezhető egy-egy szöggel, amit a választott pontokba húzott sugarak és az x tengely egymással bezárnak, azaz egy-egy $[0, 2\pi)$ intervallumba eső számmal. Ismét mondhatjuk tehát azt, hogy a $[0, 2\pi) \times [0, 2\pi)$ négyzeten választunk véletlenszerűen egy pontot (a két szög pedig ennek két koordinátája lesz).

Ha az első koordináta egy fix α szám, akkor a beírt háromszög oldalánál hosszabb húrokhoz α -nál több mint 120° -kal, de kevesebb mint 240° -kal kell nagyobboknak lennie a második koordinátának. Azaz egy fix α -ra az $\{\alpha\} \times (\alpha + 2\pi/3; \alpha + 4\pi/3)$ szakasz írja le a jó eseteket. Vigyázni kell azonban, hogy ez a szakasz kinyúlhat a négyzetből, ezért ezt " 2π -vel maradékosan osztva" kell tekintenünk, azaz - szemléletesen fogalmazva - a kinyúló szakaszt a "négyzet alján" kell folytatni, ahogy az az alábbi ábrán látható. Vagy precízen fogalmazva, amennyiben a fenti szakasznak van olyan része, aminél a második koordináta legalább 2π , akkor ezt a részt el kell tolni lefelé 2π -vel.



3.5. ábra. A harmadik megoldásnál egy négyzet egy pontját választjuk ki véletlenszerűen

Az ábrán lévő négyzet területe $(2\pi)^2 = 4\pi^2$. A két piros rész nyilván azonos területű, elegendő az egyikét kiszámolni, és azt kétszerezni. Ezt megkaphatjuk, ha a felső sarok, azaz $4\pi/3$ befogójú derékszögű háromszög területéből kivonjuk a fehér csücsök területét (amely egy $2\pi/3$ befogójú derékszögű háromszög). Tehát piros terület:

$$2 \cdot \left[\left(\frac{4\pi}{3} \right)^2 \cdot \frac{1}{2} - \left(\frac{2\pi}{3} \right)^2 \cdot \frac{1}{2} \right] = \frac{4\pi^2}{3},$$

ennek és a négyzet területének aránya pedig a várakozásainknak megfelelően $1/3$.

A geometriai valószínűségi mezők definíciója

A fenti három példában a kimenetek három különböző halmazból kerültek ki, tehát három különböző eseménytérrel beszéltünk (holott látszólag ugyanazt a problémát írtuk le). Ezek az eseményterek, és így az események is a sík részhalmazai voltak, a valószínűséget pedig az adott esemény és a teljes eseménytér területeinek arányaként definiáltuk.

Ezek speciális esetei az ún. *geometriai valószínűségi mezők*nek, melyek általában a következőképp írhatók le:

- Ω (eseménytér): az egyenes/sík/tér véges hosszúságú/területű/térfogatú részhalmazai;
- az események az Ω részhalmazai (valójában nem mindegyik, lásd a megjegyzést alább);
- egy $A \subset \Omega$ esemény $\mathbb{P}(A)$ valószínűsége az A és Ω hosszának/területének/térfogatának aránya.

Megjegyzés. Itt nem térünk ki rá részletesen, de valójában nem lehet az egyenes/sík/tér egy tetszőleges részhalmazának hosszát/területét/térfogatát definiálni, azaz az eddigi példáinktól eltérően nem lehet az Ω egy tetszőleges részhalmaza esemény. Ennek a problémának a kezelése a mértékelmélet témakörébe tartozik, amely a valószínűségszámítás precíz matematikai felépítéséhez elengedhetetlen eszköztár is egyben. Itt azonban a technikai részletektől eltekintünk, megelégszünk azzal, hogy a későbbiekben általunk tekintett részhalmazokra ezek a mennyiségek gond nélkül definiálhatók, azaz ezek mindig események lesznek. Annyit azonban még megemlítünk, hogy az események halmaza minden esetben zárt a szokásos halmazműveletekre (azaz az a megszámlálható unióra és metszetre, ill. a különbség- és komplementerképzésre) nézve. Ezt a későbbiekben hallgatólagosan gyakran kihasználjuk majd.

A geometriai valószínűségi mezők valójában nem csak geometriai problémák kezelésénél segítenek. Rengeteg olyan véletlen mennyiség van, amelynek értékét célszerű egy intervallum egy elemeként kezelni. Ezek közül a legegyszerűbb eset az, amikor egyenesen véletlenszerűen szeretnénk választani egy számot pl. a $[0; 1]$ intervallumból (függetlenül bármilyen geometriai problémától). Bár a valóságban ezt az intervallumot nem osztjuk korlátlanul kicsi részekre, a matematika szemszögéből nézve célszerű az ilyen mennyiségeket a valós számok segítségével modellezni, tehát az intervallum egy tetszőleges elemét lehetséges kimenetelnek fogjuk tekinteni, a valószínűségi mérték pedig - legalábbis, ha a kimenetelt egyenesen véletlenszerűen választjuk - meg fog egyezni a fent definiált geometriai valószínűséggel.

Magasabb dimenziós eseményterek is egyszerűen adódnak, ha például függetlenül választunk két számot egy-egy intervallumból. Ebben az esetben a kimenetelt (ahogy a Bertrand-paradoxon első és második megoldásában is történt) az intervallumok Descartes-szorzata által leírt téglalaphoz való választással modellezhetjük, ahol a két függetlenül választott számot a választott pont két koordinátája adja. Bár szemléletesen nyilvánvaló, hogy a koordináták értékei egy ilyen választásnál nem befolyásolják egymást, de később a valószínűségi változók függetlenségének (általános) definíciója után ez a tény majd újból megerősítést nyer.

Hasonlóképp járhatunk el 3 (vagy több) függetlenül (de egyenesen véletlenszerűen) választott szám esetén, ekkor az eseménytereink téglatestek (vagy általában magasabb dimenziós téglatestek) lesznek. Ebben a jegyzetben nem tekintünk 3-nál magasabb dimenziós eseménytereket.

A választott számok viszonyának jellemzése ilyen esetekben a megfelelő téglalap vagy téglatest különböző részhalmazainak vizsgálatára vezethető vissza. Ezt lényegében már a Bertrand-paradoxon lehetséges megoldásainál is láthattuk, az alábbiakban pedig két újabb példán illusztráljuk a fent leírtakat.

3.1.1. Példa. Válasszunk két számot egymásól függetlenül és egyenletesen a $[0; 1]$ intervallumból. Mi a valószínűsége, hogy az második szám nagyobb az elsőnél?

A fentieknek megfelelően ebben az esetben a két szám választása megfelel egy $P = (x, y)$ pont választásának az $\Omega = [0; 1] \times [0; 1]$ egységnégyzetből, a választott számok pedig a pont x és y koordinátái lesznek. Jelölje x az első, y pedig a második választott számot, ekkor a kérdés a $\mathbb{P}(x < y)$ valószínűség. Az $x < y$ egyenlőtlenség által leírt esemény azon $P \in \Omega$ pontokból áll, amelyek koordinátáira az teljesül. Világos, hogy ezek éppen az $x = y$ egyenes fölé eső síkrésznek azon pontjai, amik az Ω egységnégyzetben is benne vannak, vagyis az $A = (0; 0)$, $B = (0; 1)$ ill. $C = (1; 1)$ pontok által meghatározott háromszög pontjai. A geometriai valószínűségi mezőben ezen esemény valószínűsége az ABC háromszög és az Ω egységnégyzet területeinek az aránya, vagyis $1/2$.

3.1.2. Példa. Válasszunk (ismét) két számot egymásól függetlenül és egyenletesen a $[0; 1]$ intervallumból. Mi a valószínűsége, hogy a választott számok négyzetösszege 1-nél kisebb?

Ismét az $\Omega = [0; 1] \times [0; 1]$ egységnégyzetből választunk egy $P = (x, y)$ pontot. Vegyük észre, hogy az $x^2 + y^2$ négyzetösszeg éppen a P pont origótól mért távolságának a négyzete. Ennek értéke pontosan akkor lesz 1-nél kisebb, ha a gyöke, vagyis a P és az origó távolsága kisebb, mint 1. Ez pedig pontosan azt jelenti, hogy a pont rajta van az origó középpontú, 1 sugarú körlapon. Ennek az Ω egységnégyzetbe eső része egy negyedkör, aminek a területe az egységkör területének negyede, és mivel az egységnégyzet területe 1, így

$$\mathbb{P}(x^2 + y^2 < 1) = \frac{T(\text{negyedkör})}{T(\Omega)} = T(\text{negyedkör}) = \frac{\pi}{4},$$

ahol $T(\cdot)$ a területet jelöli.

Végezetül még kiemeljük egy (a korábbi példákhoz mérten) különös tulajdonságát a geometriai valószínűségi mezőknek. Az egy pontból álló halmazok hossza/területe/térfogata minden esetben 0, így ezekben az esetekben $\mathbb{P}(\{x\}) = 0$ teljesül minden $x \in \Omega$ elemi esemény esetén. Hasonlóan, két és három dimenzióban a(z egy dimenziós) szakaszok területe/térfogata és így ezek valószínűsége is 0, továbbá az analóg állítás érvényes három dimenzióban a síklapokra.

Gyakorlatok, feladatok

3.1.1. Gyakorlat. Egy 10 cm oldalhosszúságú négyzetre leejtünk egy 3 cm átmérőjű kör alakú pénzdarabot úgy, hogy a pénzdarab középpontja benne legyen a négyzetben. Tegyük fel, hogy a pénzdarab középpontja egyenletes valószínűséggel eshet akárhova (azaz egy bármilyen $x \text{ cm}^2$ területű részbe esés valószínűsége $x/100$). Mennyi a valószínűsége, hogy a pénzdarab lefedi a négyzet egy csúcsát?

3.1.2. Gyakorlat. Választunk egy pontot véletlenszerűen az $(1; 1)$, $(1; -1)$, $(-1; -1)$ és $(-1; 1)$ pontok által meghatározott négyzeten. Legyen A az az esemény, hogy a választott pont az origó középpontú, 1 sugarú körre esik, továbbá legyen B az az esemény, hogy a választott pont mindkét koordinátája pozitív. Döntsük el, hogy függetlenek-e az A és B események.

3.1.3. Gyakorlat. Véletlenszerűen választunk egy pontot a $(\pm 10; \pm 10)$ csúcspontok által meghatározott négyzeten. Mekkora az esélye, hogy a $(-1; -1)$, $(-1; 7)$, $(5; -1)$ pontokat összekötő háromszög, ennek origóra vett középpontos tükörképe, vagy a $(\pm 2; \pm 2)$ pontokat összekötő négyzet közül legalább az egyik tartalmazza a pontunkat?

3.1.4. Gyakorlat. A $[0; 1]$ intervallumon találomra kiválasztunk két számot egymástól függetlenül. Mennyi a valószínűsége, hogy az egyik szám több, mint kétszerese a másiknak?

3.1.5. Gyakorlat. Anita és Bálint megbeszélik, hogy találkozni. Egyikük sem túl határozott vagy precíz ember, ezért csak annyiban állapodnak meg, hogy délelőtt 10 és 11 óra között találkoznak egy meghatározott helyen. Azonban sajnos a türelem sem az erősségük, így az érkezéstől számított 20 perc elteltével mindig elunják a várakozást, és továbbállnak. Mennyi a találkozás valószínűsége, ha mindketten egy véletlenszerű időpontban érkeznek?

3.2. Valószínűségi változók eloszlásfüggvénye

Az alábbiakban az előző szakaszban definiált geometriai valószínűségi mezők segítségével konstruálunk véletlen mennyiségeket modellező valószínűségi változókat. Óvatosnak kell azonban lennünk: ezek a változók nem kezelhetők a második fejezetben felépített eszköztár segítségével. Ennek okai könnyen világossá válnak, amint a szakasz legfontosabb példáinak definícióját megadjuk:

- Válasszunk egy számot (pontot) egyenletesen véletlenszerűen a $[0; 1]$ intervallumban. Jelölje a szám értékét Y . Ekkor tehát $\Omega = [0; 1]$, Y pedig az az $\Omega \rightarrow \mathbb{R}$ függvény, ami az eseménytér minden eleméhez önmagát rendeli.
- Válasszunk egy pontot egyenletesen véletlenszerűen a síkon az origó középpontú egység sugarú K körlapon. Legyen Z a választott pont és az origó távolsága. Ekkor tehát $\Omega = K$, Z pedig az az $\Omega \rightarrow \mathbb{R}$ függvény, amire $(x, y) \in \Omega$ esetén $Z((x, y)) = \sqrt{x^2 + y^2}$.

Az 1.1.3. szakasz végén tett megjegyzéseink értelmében a fenti első eseménytér, vagyis a $[0; 1]$ intervallum nem megszámlálhatóan végtelen, és mivel az origó középpontú egységkör tartalmazza (lényegében) ezt az intervallumot az x tengely egyenesén, így e körlap pontjainak száma sem lehet megszámlálhatóan végtelen. Az ilyen halmazok kezelése ugyan már eleve kívül esik a második fejezet eszköztárának határain, de a 2.1.1. definíció után tett megjegyzésben azt is említettük, hogy a fogalmat teljes általánosságban definiálva valójában nem az eseménytér, hanem a változó értékészletének számossága határozza meg, hogy az adott változó diszkrét-e. Azonban világos, hogy ebben az esetben $\text{ran } Y = \text{ran } Z = [0; 1]$, és így az Y és Z függvények *nem diszkrét valószínűségi változók*, mert az értékészletük nem megszámlálható.

A fenti példák és a diszkrét változók közti lényeges különbséget jól illusztrálja a következő probléma. Vegyük észre, hogy tetszőleges $t \in [0; 1]$ esetén az $\{Y = t\}$ esemény az egy pontból álló $\{t\}$ esemény lesz, ha $t \notin [0; 1]$, akkor pedig a lehetetlen esemény. Ezért tehát (az előző szakasz utolsó bekezdése szerint) $\mathbb{P}(Y = t) = 0$ minden $t \in \mathbb{R}$ esetén. A Z változó esetében minden $t \in (0; 1]$ esetén a $\{Z = t\}$ esemény egy t sugarú *körvonal*, $\{Z = 0\}$ az origóból álló esemény, míg $t \notin [0; 1]$ esetén $\{Z = t\} = \emptyset$. Ezen halmazok területe azonban 0, így $\mathbb{P}(Z = t) = 0$ ebben az esetben is igaz minden $t \in \mathbb{R}$ -re.

Mivel a fenti események mind 0 valószínűségűek, így a változók viselkedéséről nem hozhatunk elegendő információt. Látjuk tehát, hogy a diszkrét valószínűségi változókkal ellentétben egy általános X változó esetén az $\{X = t\}$ események valószínűségei nem feltétlenül írják le az eloszlást. Ezt a problémát úgy kezeljük, hogy általános esetben a $\mathbb{P}(X < t)$ valószínűségeket fogjuk használni az eloszlás megadására. Később látni fogjuk, hogy ezek ismerete elegendő az események valószínűségeinek meghatározásához.

Ezen a ponton felmerül még egy fontos technikai probléma, amire itt röviden kitérünk. Az előző szakaszban a geometriai valószínűségi mezők definiálása után tett megjegyzés szerint egy geometriai valószínűségi mező esetén nem tekinthetjük az Ω összes részhalmazát eseménynek. Fel kell tennünk tehát a kérdést, hogy az $\{X < t\}$ halmazok, tehát Ω azon ω elemei, melyekre $X(\omega) < t$ teljesül, egyáltalán eseményt alkotnak-e? Vagyis: értelmes-e a $\mathbb{P}(X < t)$ valószínűségről beszélni? A fenti példákban ez így van, hiszen azokra

$$\{Y < t\} = \begin{cases} \emptyset, & \text{ha } t \leq 0 \\ [0; t), & \text{ha } 0 < t \leq 1 \\ [0; 1], & \text{ha } t > 1, \end{cases}$$

illetve

$$\{Z < t\} = \begin{cases} \emptyset, & \text{ha } t \leq 0 \\ K_t, & \text{ha } 0 < t \leq 1 \\ K_1, & \text{ha } t > 1, \end{cases}$$

ahol K_t jelöli az origó középpontú, t sugarú körlapot. Mivel ezen halmazoknak létezik a hossza/területe, így a fenti események valószínűsége definiált:

$$(3.1) \quad \mathbb{P}(Y < t) = \begin{cases} 0, & \text{ha } t \leq 0, \\ \frac{l([0; t))}{l([0; 1])} = t, & \text{ha } 0 < t \leq 1, \\ 1, & \text{ha } t > 1, \end{cases}$$

ahol $l(\cdot)$ jelöli az intervallum hosszát, illetve

$$(3.2) \quad \mathbb{P}(Z < t) = \begin{cases} 0, & \text{ha } t \leq 0, \\ \frac{T(K_t)}{T(K_1)} = \frac{t^2\pi}{\pi} = t^2, & \text{ha } 0 < t \leq 1, \\ 1, & \text{ha } t > 1, \end{cases}$$

ahol $T(\cdot)$ a területet jelöli.

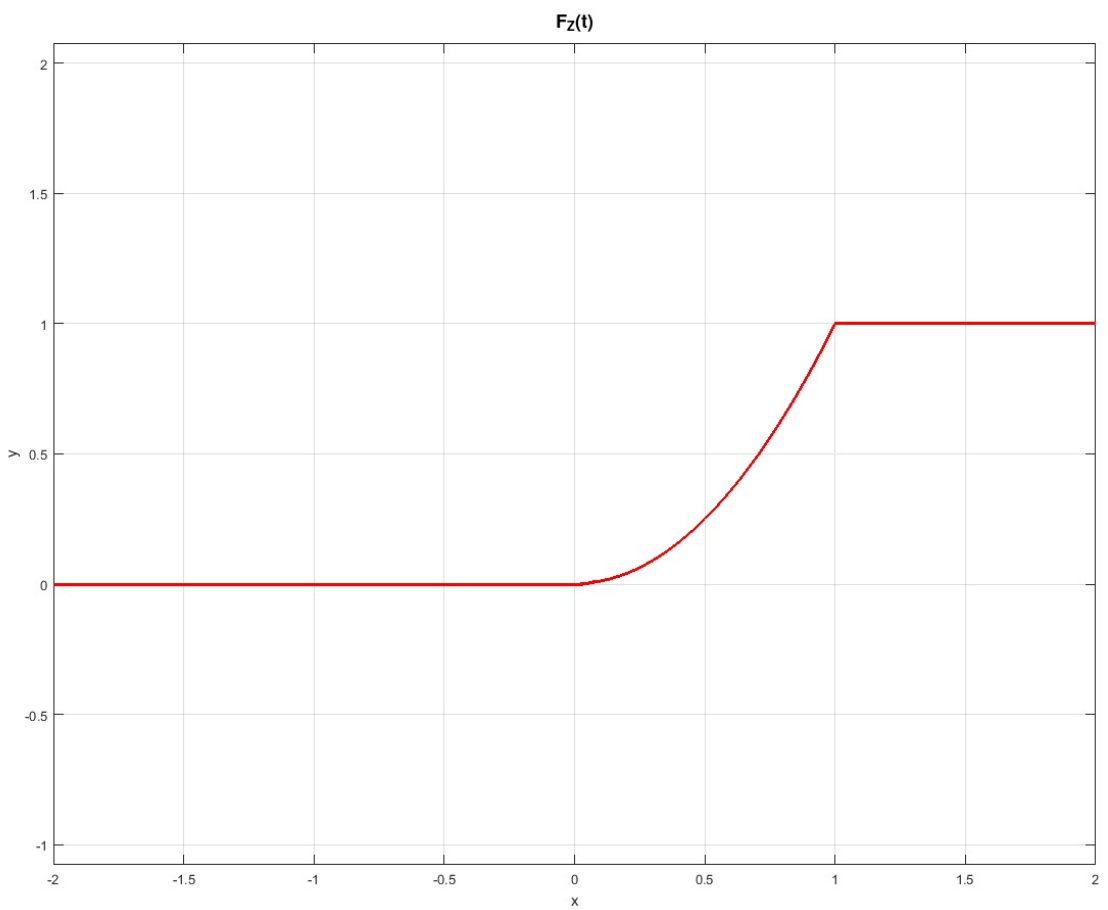
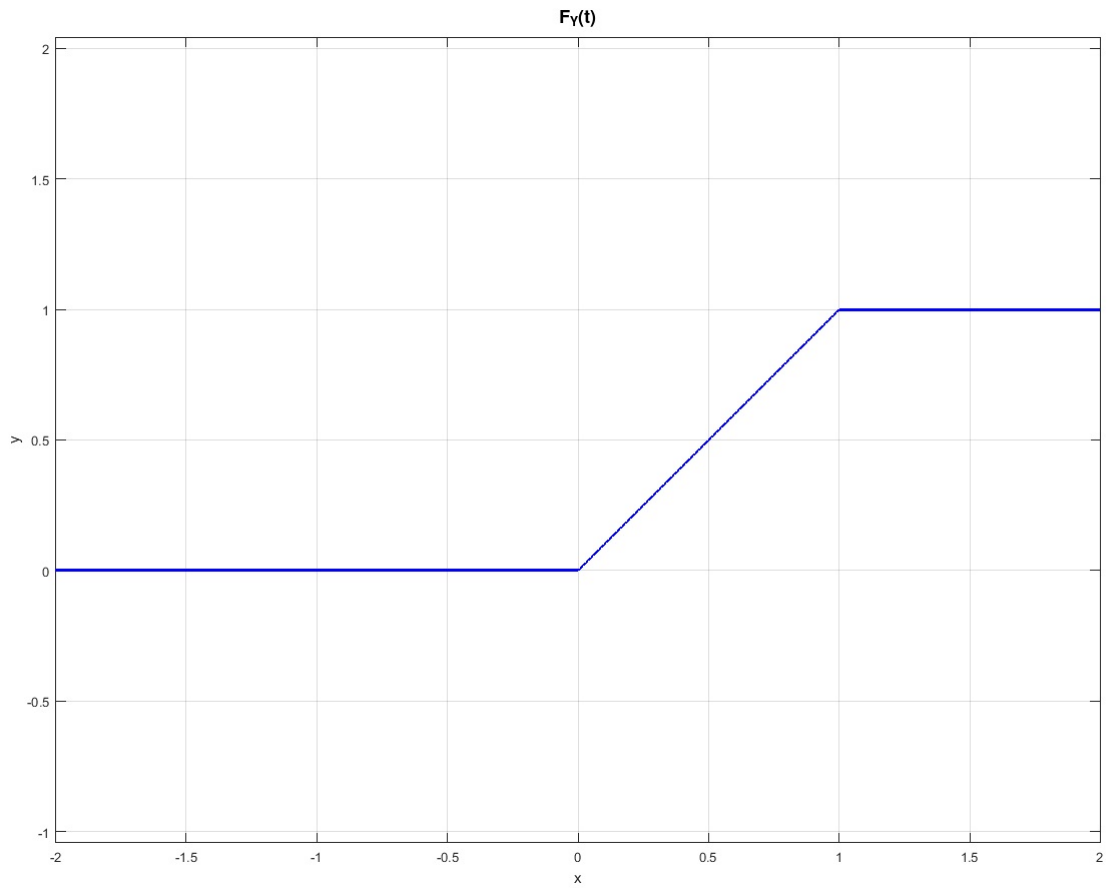
Ahhoz, hogy az általános esetben is beszélhessünk a $\mathbb{P}(X < t)$ valószínűségekről, *meg fogjuk követelni*, hogy ezek értelmesek legyenek. Azaz mostantól kezdve **csak olyan** $X : \Omega \rightarrow \mathbb{R}$ **függvényeket nevezünk valószínűségi változónak, amikre az $\{X < t\} \subset \Omega$ halmaz egyben esemény is minden egyes $t \in \mathbb{R}$ esetén**, és így értelmesek a $\mathbb{P}(X < t)$ valószínűségek. Ezekkel a technikai részletekkel azonban a későbbiekben nem fogunk külön foglalkozni, hanem előrebozsátjuk, hogy minden általunk (eddig vagy a későbbiekben) tekintett $\Omega \rightarrow \mathbb{R}$ függvény teljesíti ezt a feltételt, és így valószínűségi változó. Minden készen áll most az eloszlásfüggvény definíciójának kimondásához:

3.2.1. Definíció. Legyen $X : \Omega \rightarrow \mathbb{R}$ egy valószínűségi változó, ekkor azt az $F_X : \mathbb{R} \rightarrow [0; 1]$ függvényt, melyet az

$$F_X(t) = \mathbb{P}(X < t)$$

formula definiál minden $t \in \mathbb{R}$ -re, az X valószínűségi változó *eloszlásfüggvényének* nevezzük.

A (3.1) és (3.2) formulák tehát az Y és Z valószínűségi változók eloszlásfüggvényét adják meg, ezek grafikonja látható az alábbi ábrákon.



3.6. ábra. Az Y és Z változók eloszlásfüggvényei

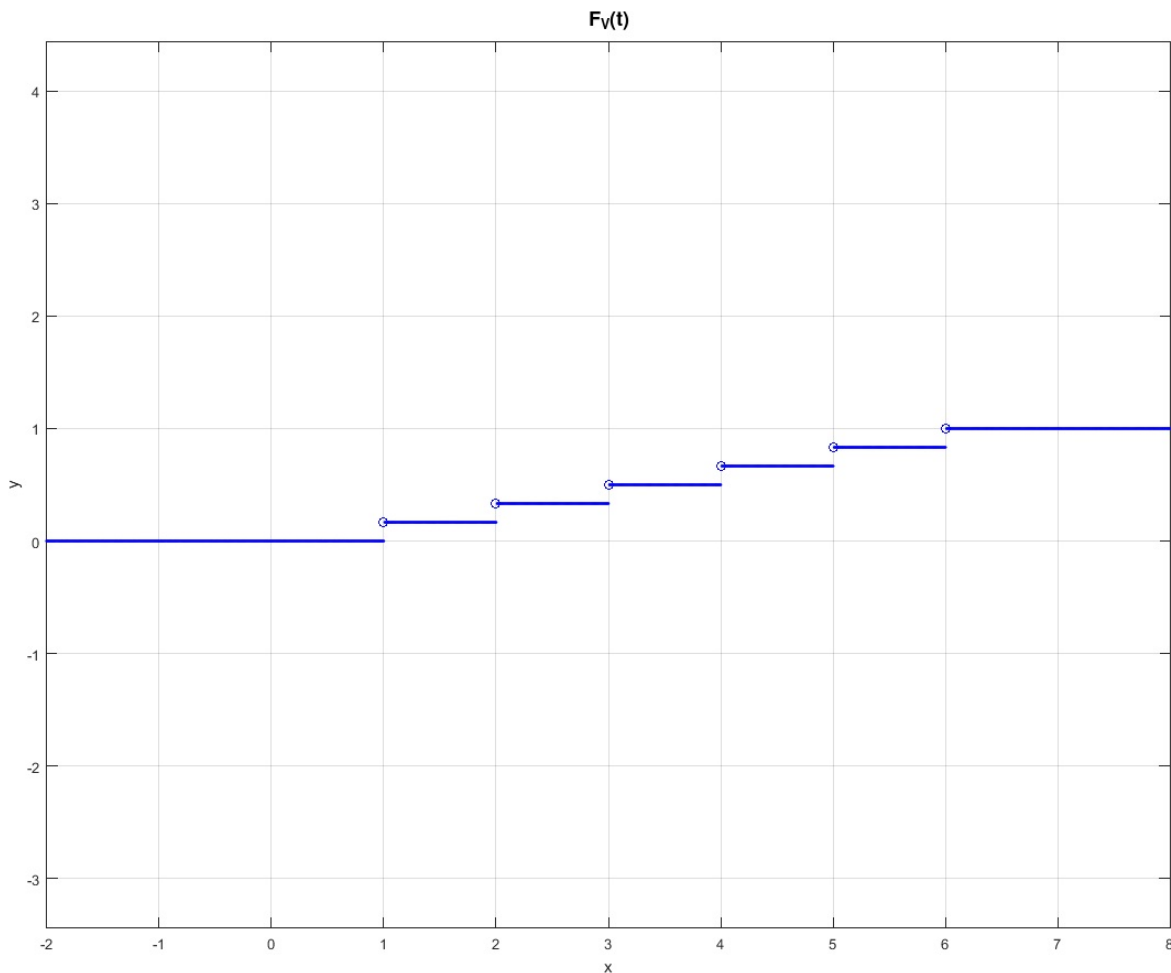
Legyen most V egy kockadobás eredménye, ekkor

$$\mathbb{P}(V < t) = \sum_{1 \leq k < \min\{t, 7\} \text{ egész}} \mathbb{P}(V = k).$$

A fenti összeg minden $t \leq 1$ -re egy üres összeg, azaz definíció szerint 0, majd $t \in (1; 2]$ esetén egy 1 tagú összeget kapunk, melynek értéke $\mathbb{P}(V = 1) = \frac{1}{6}$. Ha $2 < t \leq 3$, akkor már a $\mathbb{P}(V = 1) + \mathbb{P}(V = 2) = \frac{2}{6}$ összeget kapjuk, és könnyen láthatóan az eredmény mindig $\frac{1}{6}$ -dal nő, amint t átugorja a következő egész számot. Ez így meg egészen 6-ig, $t > 6$ esetén már mindig 1-et kapunk. Mindezt egy formulával is megadhatjuk:

$$F_V(t) = \begin{cases} 0, & \text{ha } t \leq 1, \\ \frac{[t] - 1}{6}, & \text{ha } 1 < t \leq 6, \\ 1, & \text{ha } t > 6, \end{cases}$$

ahol $[t]$ jelöli a t felső egészrészét, azaz a legkisebb egészt, melynek értéke legalább t . A formulánál azonban sokkal beszédesebb a V eloszlásfüggvényének grafikonja:



3.7. ábra. A V változó eloszlásfüggvénye

Jól látható, hogy míg Y és Z eloszlásfüggvénye folytonos, addig V eloszlásfüggvényére ez nem igaz. Megjegyezzük, hogy diszkrét valószínűségi változók eloszlásfüggvénye sosem folytonos.

Vannak azonban olyan tulajdonságok, amik mindhárom függvény esetén teljesülnek. Talán a legszembeötlőbbek egyike, hogy ezek a függvények (nem szigorúan) monoton növekvők. Ez könnyen láthatóan bármely X valószínűségi változó eloszlásfüggvényére igaz. Ha ugyanis $s, t \in \mathbb{R}$ és $s < t$, akkor minden olyan $\omega \in \Omega$ -ra, melyre $X(\omega) < s$ teljesül, arra egyben $X(\omega) < t$ is igaz, azaz $\{X < s\} \subset \{X < t\}$. Ezért ekkor

$$F_X(s) = \mathbb{P}(X < s) \leq \mathbb{P}(X < t) = F_X(t)$$

is teljesül az 1.1.2. állítás (iv) pontja szerint.

Nem ez az egyetlen közös jellemzője az eloszlásfüggvényeknek, a következő tétel pontosan karakterizálja, hogy mely függvények állnak elő ilyen módon.

3.2.1. Tétel. *Egy $F : \mathbb{R} \rightarrow [0; 1]$ függvény pontosan akkor áll elő egy valószínűségi változó eloszlásfüggvényeként, ha a következő három tulajdonság teljesül rá:*

1. F (nem feltétlenül szigorúan) monoton növekvő,
2. F balról folytonos, azaz minden $t \in \mathbb{R}$ -re az F baloldali határértéke t -ben $F(t)$,
3. $\lim_{t \rightarrow \infty} F(t) = 1$ és $\lim_{t \rightarrow -\infty} F(t) = 0$.

A tétel kimondása előtt éppen azt láttuk be, hogy ha F egy eloszlásfüggvény, akkor az első tulajdonság teljesül rá. A tétel többi állítását nem bizonyítjuk, de könnyen ellenőrizhető, hogy a fenti példákra valóban teljesül mindhárom tulajdonság.

Az eloszlásfüggvény értéke egy esemény valószínűségéeként van definiálva, de segítségével más események valószínűségei is kifejezhetők:

3.2.2. Állítás. *Legyen X egy valószínűségi változó, $s, t \in \mathbb{R}$, $s < t$, ekkor*

$$F_X(t) - F_X(s) = \mathbb{P}(s \leq X < t).$$

Bizonyítás. Az állítás igazolásához csupán azt kell megmondolni, hogy

$$\{X < t\} = \{X < s\} \cup \{s \leq X < t\}$$

egy diszjunkt felbontás, azaz a jobb oldalon egymást kizáró események uniója áll. Ezért a valószínűségi mérték additivitása miatt

$$\mathbb{P}(X < t) = \mathbb{P}(X < s) + \mathbb{P}(s \leq X < t)$$

teljesül. Ebből az egyenletet átrendezve és az eloszlásfüggvény definícióját használva adódik az állítás. \square

Az eloszlásfüggvényt azért vezettük be, mert egyes valószínűségi változóknál a $\mathbb{P}(X = t)$ valószínűségek nem voltak elégségesek az eloszlás jellemzésére. A diszkrét valószínűségi változóknál azonban ez nem így volt, ott a fenti valószínűségek bármely, a változó értékeinek segítségével kifejezhető esemény valószínűségének kiszámolásához elegendők voltak. Vajon tartalmaz-e az eloszlásfüggvény a diszkrét esetben elegendő információt ehhez? A válasz

természetesen az, hogy igen, hiszen maguk a $\mathbb{P}(X = t)$ valószínűségek is kifejezhetők a segítségével:

$$\mathbb{P}(X = t) = \mathbb{P}(X \leq t) - \mathbb{P}(X < t) = \lim_{s \rightarrow t+0} F_X(s) - F_X(t),$$

ahol a jobb oldalon az első tagban jobb oldali határértéket veszünk. Ezt a formulát nem bizonyítjuk, és a diszkrét esetben a gyakorlatban továbbra is elsősorban a $\mathbb{P}(X = t)$ valószínűségeket használjuk az eloszlás megadásához.

Együttes eloszlás és függetlenség

Az alábbiakban nagyon röviden és vázlatosan tárgyaljuk, hogy hogyan vizsgálható egyszerre több valószínűségi változó viszonya. Ahogy a diszkrét esetben a súlyfüggvény többváltozós általánosítása kézenfekvő módon írta le több valószínűségi változó együttes eloszlását, úgy az általános esetben analóg módon kaphatjuk meg azt az eloszlásfüggvény általánosításával. Azaz, ha $\underline{X} = (X_1, \dots, X_n)$ egy valószínűségi vektorváltozó, vagyis $X_1, \dots, X_n : \Omega \rightarrow \mathbb{R}$ azonos valószínűségi mezőn értelmezett tetszőleges valószínűségi változók, akkor az *együttes eloszlásfüggvényüket* az

$$F_{\underline{X}}(t_1, \dots, t_n) = \mathbb{P}(X_1 < t_1, \dots, X_n < t_n)$$

formula definiálja. Hasonlóan az egy változós esethez, az együttes eloszlásfüggvény megadása elegendő ahhoz, hogy minden, az X_1, \dots, X_n változók segítségével leírható esemény valószínűségét meghatározzuk, tehát azt mondjuk, hogy az eloszlásfüggvény leírja a változók együttes eloszlását.

Fontos speciális eset az, amikor a változók viselkedése nincs hatással egymásra, azaz a változók függetlenek egymástól. A függetlenséget azonban általában nem tudjuk megadni a súlyfüggvény segítségével, de szerencsére az eloszlásfüggvény erre a célra is jól használható:

3.2.2. Definíció. Az $X, Y : \Omega \rightarrow \mathbb{R}$ ugyanazon valószínűségi mezőn értelmezett valószínűségi változókat *függetlennek* nevezzük, ha minden $s, t \in \mathbb{R}$ esetén az $\{X < s\}$ és $\{Y < t\}$ események függetlenek, azaz ha

$$F_{X,Y}(s, t) = \mathbb{P}(X < s \text{ és } Y < t) = \mathbb{P}(X < s)\mathbb{P}(Y < t) = F_X(s)F_Y(t)$$

teljesül.

Természetesen felmerül a kérdés, hogy a fenti definíció ugyanazt a fogalmat adja-e a diszkrét esetben, mint amit a 2.2.3. definícióban kaptunk. Szerencsére belátható, hogy ez így van, tehát két diszkrét változó pontosan akkor független a 2.2.3. definíció értelmében, ha az utóbbi definíció értelmében azok. Vagyis az új függetlenségfogalom a korábbiak az általánosítása.

3.2.1. Példa. Az előző szakaszban megelőlegeztük azt a tényt, hogy egy téglalapon egyenletesen véletlenszerűen választott pont koordinátái egymástól függetlenek. Ezt most be is látjuk, azaz megmutatjuk, hogy a fenti definícióban megadott tulajdonság teljesül.

Legyen $\Omega = [a; b] \times [c; d]$, és válasszunk ezen a téglalapon véletlenszerűen egy pontot. Jelölje továbbá X a pont első, míg Y a pont második koordinátáját, ekkor $X : \Omega \rightarrow [a; b]$ és $Y : \Omega \rightarrow [c; d]$ valószínűségi változók, és minden $s \in [a; b]$ és $t \in [c; d]$ esetén

$$\mathbb{P}(X < s) = \frac{T([a; s] \times [c; d])}{T(\Omega)} = \frac{(s-a)(d-c)}{(b-a)(d-c)} = \frac{s-a}{b-a},$$

$$\mathbb{P}(Y < t) = \frac{T([a; b] \times [c; t])}{T(\Omega)} = \frac{(b-a)(t-c)}{(b-a)(d-c)} = \frac{t-c}{d-c},$$

továbbá

$$\mathbb{P}(X < s, Y < t) = \frac{T([a; s] \times [c; t])}{T(\Omega)} = \frac{(s-a)(t-c)}{(b-a)(d-c)},$$

ahol $T(\cdot)$ jelöli a területet. Eszerint minden $s \in [a; b]$ és $t \in [c; d]$ esetén $\mathbb{P}(X < s, Y < t) = \mathbb{P}(X < s)\mathbb{P}(Y < t)$ teljesül.

Ha $s < a$ vagy $t < c$, akkor könnyen láthatóan az előbbi egyenlet mindkét oldala 0 lesz, ha pedig $s > b$ és $t > d$ akkor a fenti valószínűségek értéke 1, így az egyenlet ekkor is teljesül. Ha $s \in [a; b]$ és $t > d$, akkor

$$\mathbb{P}(X < s, Y < t) = \frac{T([a; s] \times [c; d])}{T(\Omega)} = \mathbb{P}(X < s) = \mathbb{P}(X < s) \cdot 1 = \mathbb{P}(X < s)\mathbb{P}(Y < t),$$

a maradék egy eset pedig, tehát amikor $s > b$ és $t \in [c; d]$, ugyanígy kezelhető. Ezzel tehát az állítást beláttuk.

Végezetül definiáljuk több változó függetlenségét az általános esetben is:

3.2.3. Definíció. Az $X_1, \dots, X_n : \Omega \rightarrow \mathbb{R}$ ugyanazon valószínűségi mezőn értelmezett valószínűségi változókat *páronként/együttesen függetlennek* nevezzük, ha minden $t_1, \dots, t_n \in \mathbb{R}$ esetén az $\{X_1 < t_1\}, \dots, \{X_n < t_n\}$ események páronként/együttesen függetlenek.

Gyakorlatok, feladatok

3.2.1. Gyakorlat. Válasszunk egy pontot véletlenszerűen az $(1; 0)$, $(0; 1)$ és $(-1; 0)$ csúcsok által meghatározott egyenlő szárú háromszög belsejében, és jelölje X a választott pont és az x tengely távolságát.

a) Adjuk meg X eloszlásfüggvényét.

b) Számoljuk ki a $\{0,25 \leq X < 0,5\}$ esemény valószínűségét.

3.2.2. Feladat. A $[0; 1]$ intervallumon véletlenszerűen kiválasztunk két számot. Legyen X a két szám távolsága. Adjuk meg az X eloszlásfüggvényét.

3.2.3. Feladat. (M) Véletlenszerűen választunk két számot a $[-1; 1]$ intervallumból. Legyen X a két szám összege. Mi a valószínűsége annak, hogy X pozitív? Határozzuk meg X eloszlásfüggvényét.

3.2.4. Feladat. Az egységnégyzeten taláломra kiválasztunk egy P pontot. Jelölje X a P -hez legközelebbi oldal és a P pont távolságát. Határozzuk meg X eloszlásfüggvényét.

3.2.5. Gyakorlat. Jelölje X az ötös lottón kihúzott öt szám közül a legkisebbet. Adjuk meg X eloszlásfüggvényének értékét a 4 ill. 25 helyeken. Folytonos-e ez az eloszlásfüggvény?

3.2.6. Gyakorlat. Legyen X egy kockadobás eredménye. Határozzuk meg az $Y = (X - 3)^2$ eloszlásfüggvényét.

3.2.7. Gyakorlat. Eloszlásfüggvények-e az alábbi hozzárendelési szabályú $\mathbb{R} \rightarrow \mathbb{R}$ függvények?

a) $F(t) = \begin{cases} 1 & \text{ha } t > 0, \\ 0 & \text{egyébként.} \end{cases}$ b) $F(t) = \begin{cases} 1 - e^{-at} & \text{ha } t > 0, \\ 0 & \text{egyébként.} \end{cases} \quad (a \in \mathbb{R})$

c) $F(t) = 1 - e^{-t^2}$ d) $F(t) = \frac{1}{\pi} \arctg(t) + \frac{1}{2}$

3.3. Abszolút folytonos változók

Ebben a szakaszban a valószínűségi változók egy speciális osztályával ismerkedünk meg, mégpedig olyan változókkal, amelyek esetében létezik a diszkrét változóknál definiált súlyfüggvény egy "folytonos" analogonja. Az előző fejezet példáinak nagy része is ilyen, ezek körül egy konkrét változót elemezve fogunk eljutni az ún. sűrűségfüggvény ill. az abszolút folytonos valószínűségi változók fogalmához. Ezek után a diszkrét esetben már látott várható érték és szórás fogalmait definiáljuk az abszolút folytonos változók esetén is, majd néhány nevezetes eloszlás elemzése következik.

3.3.1. A sűrűségfüggvény fogalma

Tekintsük a következő, az előző szakaszból már ismert valószínűségi változót: válasszunk egy pontot véletlenszerűen a síkon az origó középpontú egység sugarú körlapon, és legyen Z a választott pont távolsága az origótól. Először is emlékeztetünk, hogy

$$(3.3) \quad F_Z(t) = \mathbb{P}(Z < t) = \begin{cases} 0, & \text{ha } t \leq 0, \\ t^2, & \text{ha } 0 < t \leq 1, \\ 1, & \text{ha } t > 1. \end{cases}$$

Az előző szakaszban láttuk, hogy ennek segítségével felírhatók más események valószínűségei is, például ha $s, t \in \mathbb{R}$, $s < t$, akkor $\mathbb{P}(s \leq Z < t) = F_Z(t) - F_Z(s)$. Továbbá, mivel $\mathbb{P}(Z = s) = 0$ minden $s \in \mathbb{R}$ esetén, valamint

$$\{s \leq Z < t\} = \{Z = s\} \cup \{s < Z < t\}$$

egy diszjunkt felbontás (azaz a jobb oldalon egymást kizáró események uniója áll), így

$$\mathbb{P}(s < Z < t) = \mathbb{P}(s \leq Z < t) - \mathbb{P}(Z = s) = \mathbb{P}(s \leq Z < t) = F_Z(t) - F_Z(s)$$

is érvényes, tehát ebben az esetben mindegy, hogy az intervallum végpontját a fenti eseményben számításba vesszük vagy sem. Hasonló érveléssel adódik ugyanez a másik végpontra, tehát az $[s; t]$, $[s; t)$, $(s; t]$ és $(s; t)$ intervallumokba ugyanolyan eséllyel esik a Z változó. Ez érvényes minden olyan valószínűségi változóra, ami az adott intervallum végpontjainak értékét 0 valószínűséggel veszi fel. Látni fogjuk, hogy az alább definiált abszolút folytonos valószínűségi változóknál ez így lesz.

Habár $\mathbb{P}(Z = t) = 0$ érvényes minden $t \in \mathbb{R}$ esetén, mégis van arról információnk, hogy Z milyen eséllyel lesz *közel* t -hez, például annak egy kis $\varepsilon > 0$ sugarú környezetében. A fentiek szerint ugyanis

$$\mathbb{P}(t - \varepsilon < Z < t + \varepsilon) = F_Z(t + \varepsilon) - F_Z(t - \varepsilon).$$

Ez az érték persze kis ε -ra kicsi lesz, hiszen F_Z folytonossága miatt a fenti különbség mindkét tagja $F_Z(t)$ -hez, különbségük pedig így 0-hoz tart, ahogy ε tart 0-hoz. Mindez összhangban van azzal az általunk nem bizonyított, de szemléletes ténnyel, hogy a $\mathbb{P}(t - \varepsilon < Z < t + \varepsilon)$ valószínűség $\mathbb{P}(Z = t) = 0$ -hoz tart, ahogy a t körüli intervallum hossza 0-hoz tart.

A fenti valószínűség azonban már sokkal informatívabb, ha az intervallum hosszát is figyelembe vesszük, pontosabban, ha a valószínűség és az intervallum hosszának arányát tekintjük, azaz a

$$\frac{\mathbb{P}(t - \varepsilon < Z < t + \varepsilon)}{2\varepsilon} = \frac{F_Z(t + \varepsilon) - F_Z(t - \varepsilon)}{2\varepsilon}$$

kifejezést. Ha $\varepsilon \rightarrow 0$, akkor ennek határértékét - amennyiben létezik - tekinthetjük egyfajta súlynak, ami meghatározza, hogy milyen valószínűséggel esik Z a t közelébe. Hamarosan látni fogjuk, hogy ez valóban egy jó kép.

A határérték meghatározásához figyeljük meg, hogy

$$\begin{aligned} \frac{F_Z(t + \varepsilon) - F_Z(t - \varepsilon)}{2\varepsilon} &= \frac{F_Z(t + \varepsilon) - F_Z(t) + F_Z(t) - F_Z(t - \varepsilon)}{2\varepsilon} \\ &= \frac{1}{2} \cdot \frac{F_Z(t + \varepsilon) - F_Z(t)}{(t + \varepsilon) - t} + \frac{1}{2} \cdot \frac{F_Z(t) - F_Z(t - \varepsilon)}{(t - \varepsilon) - t}, \end{aligned}$$

ez pedig éppen az F_Z eloszlásfüggvény t -beli deriváltjához tart, ha ε tart 0-hoz - legalábbis amennyiben a derivált létezik. Az $F_Z(t)$ -re adott (3.3) formula alapján könnyen látható, hogy F_Z deriválható minden $t \neq 0, 1$ esetén, hiszen itt vagy egy konstanssal, vagy egy polinomfüggvénnyel egyezik meg. Az is könnyen látható, hogy a 0 pontban a jobb- és baloldali deriváltak léteznek, és mindkettő 0-val egyenlő, így a derivált létezik a $t = 0$ pontban is. Nem ez a helyzet viszont a $t = 1$ pontban, hiszen, habár a bal és jobb oldali deriváltak léteznek, azok nem egyeznek meg:

$$\lim_{t \rightarrow 1-0} \frac{F_Z(t) - F_Z(1)}{t - 1} = \lim_{t \rightarrow 1-0} \frac{t^2 - 1}{t - 1} = \lim_{t \rightarrow 1-0} t + 1 = 2,$$

de

$$\lim_{t \rightarrow 1+0} \frac{F_Z(t) - F_Z(1)}{t - 1} = \lim_{t \rightarrow 1+0} \frac{1 - 1}{t - 1} = 0.$$

Definiáljuk az $f_Z(t)$ függvényt a következőképp: legyen $f_Z(t) = F'_Z(t)$, ha $t \neq 1$, és legyen $f_Z(1) = 0$. Tehát

$$(3.4) \quad f_Z(t) = \begin{cases} 2t, & \text{ha } 0 < t < 1 \\ 0 & \text{különben.} \end{cases}$$

Ezt a függvényt a későbbi definíciók alapján a Z sűrűségfüggvényének fogjuk hívni. A Newton–Leibniz-formula szerint ekkor

$$(3.5) \quad \mathbb{P}(s < Z < t) = F_Z(t) - F_Z(s) = \int_s^t f_Z(y) dy$$

érvényes minden olyan $s, t \in \mathbb{R}$, $s < t$ számpárra, ahol az F_Z függvény deriválható az $(s; t)$ intervallumon és folytonos az $[s; t]$ zárt intervallumon, vagyis ez igaz minden $s < t \leq 1$ és $1 \leq s < t$ esetén. Azonban ha $s < 1 < t$, akkor kettébonthatjuk az $(s; t)$ intervallumot, és tagonként alkalmazva a Newton–Leibniz-formulát láthatjuk, hogy (3.5) ekkor is teljesül:

$$\begin{aligned} \int_s^t f_Z(y) dy &= \int_s^1 f_Z(y) dy + \int_1^t f_Z(y) dy \\ &= F_Z(1) - F_Z(s) + F_Z(t) - F_Z(1) = F_Z(t) - F_Z(s). \end{aligned}$$

Ebből valóban látszik, amit már korábban említettünk: ha f_Z folytonos t -ben, és $\varepsilon > 0$ kicsi, akkor

$$\mathbb{P}(t - \varepsilon < Z < t + \varepsilon) = F_Z(t + \varepsilon) - F_Z(t - \varepsilon) = \int_{t-\varepsilon}^{t+\varepsilon} f_Z(y) dy \approx 2\varepsilon \cdot f_Z(t),$$

vagyis $f_Z(t)$ -re valóban tekinthetünk egyfajta súlyként, ami leírja, hogy milyen valószínűséggel esik a Z értéke a t egy kis környezetébe. Ebből a szempontból f_Z a diszkrét változónál látott súlyfüggvény analogonjának tekinthető, amely konkrétan a $\mathbb{P}(X = t)$ valószínűséget írta le.

Mielőtt rátérnénk az abszolút folytonosság általános definíciójára, még egy formulát igazolunk a Z változóra. Az f_Z sűrűségfüggvényt az F_Z eloszlásfüggvényből deriválással kaptuk, a Newton–Leibniz-formula pedig ennek a folyamatnak a megfordítását adja. Ha most a (3.5) formulában s -sel tartunk a mínusz végtelenhez, akkor $F_Z(s)$ 0-hoz tart (pontosabban már minden $s \leq 0$ esetén 0), így az integrálás segítségével ténylegesen visszakapjuk az eloszlásfüggvény értékét a sűrűségfüggvényből:

$$F_Z(t) = \int_{-\infty}^t f_Z(y) dy.$$

Ez az a formula, aminek segítségével az abszolút folytonos valószínűségi változókat *definiáljuk*:

3.3.1. Definíció. Az X valószínűségi változót *abszolút folytonosnak* nevezzük, ha létezik olyan $f_X : \mathbb{R} \rightarrow [0; \infty)$ nemnegatív függvény, amelyre az $\int_{-\infty}^{\infty} f_X(y) dy$ Riemann-integrál létezik, és minden $t \in \mathbb{R}$ esetén

$$F_X(t) = \int_{-\infty}^t f_X(y) dy,$$

ahol F_X az X eloszlásfüggvénye. Ekkor az f_X függvényt az X változó *sűrűségfüggvényének* nevezzük.

Megjegyzés. Ezen a kurzuson a későbbiekben az abszolút folytonos változókra gyakran csak folytonos valószínűségi változókként hivatkozunk, és az abszolút jelzőt elhagyjuk. Az irodalomban azonban folytonos valószínűségi változó alatt olyan változót értenek, amelyeknek az eloszlásfüggvénye folytonos. Mivel az abszolút folytonos valószínűségi változók eloszlásfüggvénye egy sűrűségfüggvény integrálfüggvényeként áll elő, így folytonos. Tehát az abszolút folytonos valószínűségi változók a fenti értelemben folytonosak, azonban általában ez fordítva nem igaz. Mivel mi a továbbiakban kizárólag abszolút folytonos változókkal foglalkozunk, a két elnevezést szinonimaként használjuk.

Vegyük észre, hogy a fenti definícióban lévő f_X sűrűségfüggvény nem egyértelmű. Ha például véges sok ponton megváltoztatjuk az értékét, akkor ez a változtatás az f_X integrálját nem változtatja meg, így a fenti definiáló tulajdonság továbbra is érvényben marad, tehát az így kapott függvény szintén sűrűségfüggvény lesz.

A fentiek szerint az Z változó (abszolút) folytonos (ehhez persze még ellenőrizni kell, hogy f_Z integrálható a teljes valós egyenesen, de mivel $t \notin (0; 1)$ esetén $f_Z(t) = 0$, ez nyilvánvalóan teljesül). A Z változó példáján illusztrált módszer általában is alkalmazható a folytonosság igazolására:

3.3.1. Tétel. Legyen X olyan valószínűségi változó, amelynek F_X eloszlásfüggvénye folytonos és véges sok pont kivételével mindenhol deriválható. Ekkor az X valószínűségi változó abszolút folytonos, és az

$$f_X(t) = \begin{cases} F_X'(t), & \text{ahol a derivált létezik,} \\ 0 & \text{különben} \end{cases}$$

függvény sűrűségfüggvénye X -nek.

A Z változóra alkalmazott gondolatmenet kis módosításokkal tulajdonképpen a fenti tétel bizonyítását az általános esetben is megadja, de ezt itt nem részletezzük. Viszont néhány további, a konkrét példával kapcsolatban látott állítást kimondunk az általános esetben is.

3.3.2. Állítás. *Legyen X egy folytonos valószínűségi változó, amelynek sűrűségfüggvénye f_X . Ekkor tetszőleges $s, t \in \mathbb{R}$, $s < t$ esetén*

$$\mathbb{P}(s \leq X < t) = \int_s^t f_X(y) dy.$$

Bizonyítás. Az előző szakaszban láttuk, hogy

$$\mathbb{P}(s \leq X < t) = F_X(t) - F_X(s).$$

Mivel X folytonos valószínűségi változó, így az eloszlásfüggvényének értékei felírhatók a sűrűségfüggvény segítségével:

$$F_X(t) - F_X(s) = \int_{-\infty}^t f_X(y) dy - \int_{-\infty}^s f_X(y) dy = \int_s^t f_X(y) dy,$$

ahol felhasználtuk az integrál határokra vonatkozó additivitását. \square

Megjegyezzük, hogy általában is igaz minden X folytonos valószínűségi változóra, hogy tetszőleges $t \in \mathbb{R}$ esetén $\mathbb{P}(X = t) = 0$, és így a fenti állításban szereplő eseményben mindkét egyenlőtlenségnél elhagyhatjuk vagy megengedhetjük az egyenlőséget, tehát

$$\mathbb{P}(s < X < t) = \mathbb{P}(s \leq X < t) = \mathbb{P}(s < X \leq t) = \mathbb{P}(s \leq X \leq t),$$

és ezeket a fenti integrál segítségével írhatjuk fel. Továbbá a fenti állításban megengedhetjük az $s = -\infty$ és $t = \infty$ értéket is, ekkor a fenti egyenlőtlenségek úgy értendők, hogy X -re nem szabunk alsó ill. felső korlátot, az integrálásnál pedig a megfelelő improprius integrált tekintjük.

Ahhoz hasonlóan, ahogy az eloszlásfüggvények esetén történt, a sűrűségfüggvényeket is karakterizálhatjuk:

3.3.3. Tétel. *Egy $f : \mathbb{R} \rightarrow [0; \infty)$ nemnegatív függvény pontosan akkor sűrűségfüggvénye egy X folytonos valószínűségi változónak, ha Riemann-integrálható a valós egyenesen, és*

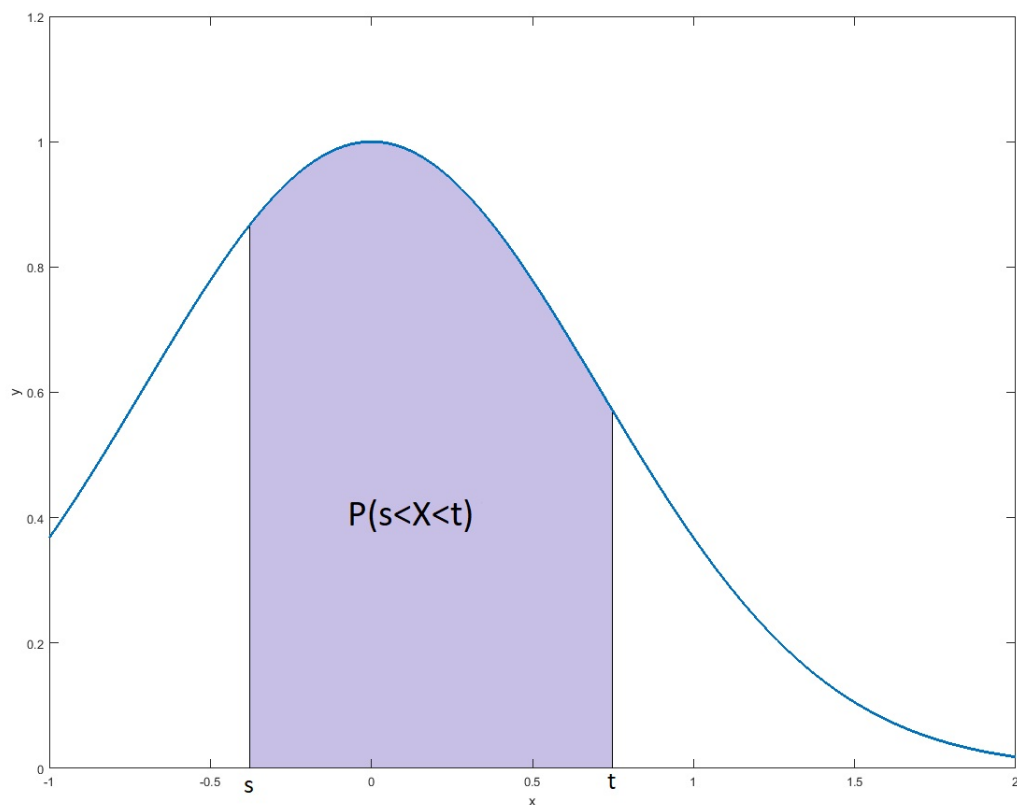
$$\int_{-\infty}^{\infty} f(y) dy = 1.$$

Az világos, hogy ha f egy X változó sűrűségfüggvénye, akkor a fenti egyenletnek teljesülnie kell, hiszen

$$\int_{-\infty}^{\infty} f(y) dy = \lim_{t \rightarrow \infty} \int_{-\infty}^t f(y) dy = \lim_{t \rightarrow \infty} F_X(t) = 1$$

az eloszlásfüggvény tulajdonságai alapján. A tétel másik irányát nem bizonyítjuk.

A sűrűségfüggvény és a diszkrét változók súlyfüggvénye közötti analógia a fenti tétel állításában is tetten érhető, ha összevetjük azt a súlyfüggvény (2.6) karakterizációjával. Ebben az esetben és a későbbiekben is gyakran előfordul, hogy a diszkrét ill. abszolút folytonos változókra vonatkozó egymásnak megfelelő állítások vagy definíciók formálisan csak abban különböznek, hogy egy (esetleg végtelen) összeget egy valós intervallumon vett integrálra cserélünk (persze az előbbi esetben a súlyfüggvény, míg utóbbi esetben a sűrűségfüggvény szerepel a formulákban). Ez is azt tükrözi, hogy az integrálás valójában az összeadásnak egyfajta megfelelője. A valós számok "túl sokan vannak" ahhoz, hogy össze tudjuk adni mindegyiküket, de az integrál végeredményben ezt teszi lehetővé, összeg helyett egy görbe alatti területet adva.



3.8. ábra. A sűrűségfüggvény egy részének görbe alatti területe egy valószínűséget fejez ki

3.3.2. Várható érték és szórás

A diszkrét valószínűségi változók esetén a várható érték a változó által felvett értékek valószínűségekkel, azaz a súlyfüggvény értékeivel súlyozott összege volt. A folytonos esetben ilyen összeg nem írható fel, azonban a súlyfüggvény szerepét most is a sűrűségfüggvény veszi át. Továbbá, ahogy az előző szakasz végén azt már előrebocsátottuk, összeg helyett itt is integrálást alkalmazunk:

3.3.2. Definíció. Legyen X abszolút folytonos valószínűségi változó, aminek sűrűségfüggvénye f_X . Ekkor X várható értéke

$$\mathbb{E}(X) = \int_{-\infty}^{\infty} t f_X(t) dt,$$

amennyiben a fenti integrál abszolút konvergens, azaz az

$$\int_{-\infty}^{\infty} |t| f_X(t) dt$$

integrál véges.

Megjegyezzük, hogy ha egy improprius integrál abszolút konvergens, akkor konvergens is, így a definícióban szereplő feltétel teljesülésekor a várható érték is véges.

3.3.1. Példa. Számoljuk ki az előző szakaszban szereplő Z valószínűségi változó várható értékét. Az f_Z sűrűségfüggvény értékét a (3.4) formula adja meg. Ebből látszik, hogy a $(0; 1)$ intervallumon kívül ez a függvény 0, így a várható érték kiszámításánál valójában csak ezen a véges intervallumon kell integrálnunk (és így persze abszolút konvergencia lesz az integrál). A fenti definícióba behelyettesítve, majd a Newton–Leibniz-formulát alkalmazva:

$$\mathbb{E}(Z) = \int_{-\infty}^{\infty} t f_Z(t) dt = \int_0^1 2t^2 dt = \left[\frac{2t^3}{3} \right]_0^1 = \frac{2}{3}.$$

Ahogy a diszkrét esetben, a várható érték linearitása a folytonos esetben is érvényes:

3.3.4. Tétel. *Legyenek $X, Y : \Omega \rightarrow \mathbb{R}$ ugyanazon valószínűségi mezőn értelmezett abszolút folytonos valószínűségi változók, melyekre $\mathbb{E}(X)$ és $\mathbb{E}(Y)$ véges. Legyen továbbá $c \in \mathbb{R}$ egy tetszőleges valós szám, ekkor*

$$\mathbb{E}(X + Y) = \mathbb{E}(X) + \mathbb{E}(Y), \quad \mathbb{E}(cX) = c \cdot \mathbb{E}(X).$$

Az tételben szereplő egyenlőségek úgy értendők, hogy amennyiben az X ill. Y várható értéke véges, akkor az egyenlőségek mindkét oldala véges, és azok megegyeznek. Megjegyezzük, hogy a várható érték tetszőleges - azaz nem csak diszkrét vagy folytonos - esetben is definiálható (lásd pl. a 4.3. szakaszt az [5] jegyzetben), és a linearitás általában is érvényes, ha az X és Y várható értéke véges. Ezt tehát akár olyan esetekben is alkalmazhatjuk, mikor az egyik változó folytonos, a másik pedig diszkrét.

A folytonos esetben a szórás ugyanúgy fogjuk definiálni a várható érték segítségével, ahogy az a diszkrét esetben is történt. Ehhez most is szükségünk lesz az X^2 változó várható értékére, ezt pedig a transzformált várható értékére vonatkozó, a diszkrét esettel analóg állítás segítségével számolhatjuk ki:

3.3.5. Tétel (Transzformált várható értéke a folytonos esetben). *Legyen X egy folytonos valószínűségi változó f_X sűrűségfüggvénnyel, $g : \mathbb{R} \rightarrow \mathbb{R}$ pedig egy olyan függvény, melyre $g(X)$ szintén valószínűségi változó. Ha az $\mathbb{E}(g(X))$ várható érték véges, akkor*

$$\mathbb{E}(g(X)) = \int_{-\infty}^{\infty} g(t) f_X(t) dt.$$

A tételt nem bizonyítjuk, viszont azt megjegyezzük, hogy itt az $g(X)$ változóról nem tesszük fel, hogy folytonos, a tétel teljesen általánosan érvényes, amennyiben ez egy valószínűségi változó, azaz $\{g(X) < t\}$ halmaz egyben esemény is. Ez az összes általunk tekintett esetben teljesülni fog, így ezen technikai részletekkel a jövőben nem foglalkozunk.

Ahogy már említettük, a legfontosabb példánk a tétel alkalmazására az X^2 változó várható értéke lesz, ahol tehát a $g(t) = t^2$ választással élünk. Éppen ezért ebben a speciális esetben külön felírjuk a tétel állítását: ha X egy folytonos valószínűségi változó, amelyre $\mathbb{E}(X^2)$ véges, akkor

$$(3.6) \quad \mathbb{E}(X^2) = \int_{-\infty}^{\infty} t^2 f_X(t) dt.$$

3.3.2. Példa. Az előző példa folytatásaként kiszámoljuk az előző szakaszban tekintett Z valószínűségi változó négyzetének a várható értékét. Emlékeztetünk, hogy az f_Z sűrűségfüggvény értékét a (3.4) formula adja meg, amely a $(0; 1)$ intervallumon kívül 0, így az előző formulába behelyettesítve csak ezen a véges intervallumon kell integrálnunk:

$$\mathbb{E}(Z^2) = \int_{-\infty}^{\infty} t^2 f_Z(t) dt = \int_0^1 2t^3 dt = \left[\frac{t^4}{2} \right]_0^1 = \frac{1}{2}.$$

Ahogy már említettük, a szórásnégyzetet és a szórást a diszkrét esethez hasonlóan definiálhatjuk a folytonos változókra, sőt, valójában teljesen általános változókra is a várható érték segítségével. Mivel a szórásra vonatkozó állítások igazolásánál csak a várható érték tulajdonságait használtuk, amelyek pedig a folytonos esetben ill. teljesen általánosan is érvényesek, így a szórás tulajdonságai itt is érvényben maradnak. Ezek igazolása lényegében szóról szóra megegyezik a diszkrét esetben látottakkal, ezért a bizonyítások részletezését mellőzzük, és megelégszünk a definíció és a főbb állítások összefoglalásával.

3.3.3. Definíció. Legyen X egy valószínűségi változó, melynek véges a várható értéke. Ha az $\mathbb{E}((X - \mathbb{E}(X))^2)$ várható érték is véges, akkor ezt az X szórásnégyzetének (vagy varianciájának) nevezzük, és $\mathbb{D}^2(X)$ -el jelöljük. A szórásnégyzet nemnegatív gyökét az X szórásának nevezzük és $\mathbb{D}(X)$ -el jelöljük.

A szórásnégyzet általában is egy nemnegatív valószínűségi változó várható értéke, és így nemnegatív, valamint minden $c \in \mathbb{R}$ esetén érvényesek a

$$\mathbb{D}(X + c) = \mathbb{D}(X), \quad \mathbb{D}(cX) = |c| \mathbb{D}(X)$$

transzformációs szabályok. Ezek úgy értendők, hogy ha a szórás véges, akkor a formulák mindkét oldala is, és azok egyenlők. A szórásnégyzet esetén az ezekkel ekvivalens formulák $\mathbb{D}^2(X + c) = \mathbb{D}^2(X)$ ill. $\mathbb{D}^2(cX) = c^2 \mathbb{D}^2(X)$. Továbbá egy X egy valószínűségi változónak pontosan akkor véges a szórásnégyzete, ha $\mathbb{E}(X^2)$ véges, és ekkor

$$\mathbb{D}^2(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2$$

is érvényes. A továbbiakban ezt a formulát használjuk a szórás(négyzet) meghatározására.

3.3.3. Példa. Tekintsük ismét az előző szakasz Z változóját. Az előző két példában meghatároztuk az $\mathbb{E}(Z)$ ill. az $\mathbb{E}(Z^2)$ várható értékeket, így tehát az utóbbi formulát alkalmazva

$$\mathbb{D}^2(Z) = \mathbb{E}(Z^2) - \mathbb{E}(Z)^2 = \frac{1}{2} - \left(\frac{2}{3}\right)^2 = \frac{1}{2} - \frac{4}{9} = \frac{1}{18}, \quad \mathbb{D}(Z) = \frac{1}{3\sqrt{2}} = \frac{\sqrt{2}}{6}.$$

Végezetül rátérünk a függetlenség és a szórásnégyzet kapcsolatára. Általában is igaz, hogy ha X és Y független valószínűségi változók, továbbá $\mathbb{E}(X)$, $\mathbb{E}(Y)$ véges, akkor $\mathbb{E}(XY)$ is véges, és fennáll az $\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y)$ egyenlőség. Ebből a diszkrét esetben látottakhoz hasonlóan következik, hogy az X és Y véges szórású és független valószínűségi változók esetén

$$\mathbb{D}^2(X + Y) = \mathbb{D}^2(X) + \mathbb{D}^2(Y)$$

érvényes. Természetesen most is igaz az analóg állítás n változó esetén, azaz ha X_1, \dots, X_n páronként független valószínűségi változók, melyek szórása véges, akkor

$$\mathbb{D}^2(X_1 + \dots + X_n) = \mathbb{D}^2(X_1) + \dots + \mathbb{D}^2(X_n).$$

Gyakorlatok, feladatok

3.3.1. Gyakorlat. Válasszunk egy pontot véletlenszerűen az $(1; 0)$, $(0; 1)$ és $(-1; 0)$ csúcsok által meghatározott egyenlő szárú háromszög belsejében, és jelölje X a választott pont és az x tengely távolságát. Adjuk meg az X változó sűrűségfüggvényét, várható értékét és szórását. (Lásd a 3.2.1. gyakorlatot.)

3.3.2. Gyakorlat. (M) Véletlenszerűen választunk két számot egymástól függetlenül a $[-1; 1]$ intervallumból. Legyen X a két szám összege. Határozzuk meg X sűrűségfüggvényét, várható értékét és szórását. (Lásd a 3.2.3. feladatot.)

3.3.3. Gyakorlat. Az X folytonos valószínűségi változó sűrűségfüggvénye:

$$f_X(t) = \begin{cases} \frac{\alpha}{\sqrt{t}}, & \text{ha } t \in (0; 1), \\ 0, & \text{különben,} \end{cases}$$

ahol $\alpha \in \mathbb{R}$ egy paraméter. Határozzuk meg α értékét és a $\mathbb{P}\left(\frac{1}{9} < X < \frac{1}{4}\right)$ valószínűséget.

3.3.4. Gyakorlat. Egy benzinkút üzemanyagtartályát hetente teletöltik. Jelölje X a heti fogyasztást (százezer literekben), melynek sűrűségfüggvénye:

$$f_X : t \mapsto \begin{cases} 5(1-t)^4 & \text{ha } 0 < t < 1, \\ 0 & \text{egyébként.} \end{cases}$$

Mekkora legyen a tartály kapacitása, hogy annak a valószínűsége, hogy a héten kifogy az üzemanyag, kisebb legyen 0,05-nél? Mekkora az átlagos heti fogyasztás?

3.4. Nevezetes folytonos eloszlások

Ebben a szakaszban megismerkedünk három fontos folytonos eloszláscsaláddal, és leírjuk a legalapvetőbb tulajdonságait. A legegyszerűbb példánk az egyenletes eloszlás lesz, amely a diszkrét egyenletes eloszlásnak (amely esetén minden kimenetel egyformán valószínű) a folytonos megfelelője. Valójában ilyen eloszlással már a geometriai valószínűségi mezők tárgyalása során is találkoztunk. Ezután megismerkedünk a geometriai eloszlás folytonos analogójával, az ún. exponenciális eloszlással, amelyre szintén érvényes az örökifjú tulajdonság, amelyet az \mathbb{N}^+ halmaz helyett most az egész $[0, \infty)$ intervallumra definiálni fogunk. Végezetül a valószínűségszámítás egyik legfontosabb eloszláscsaládjával, az ún. normális eloszlásokkal foglalkozunk. Ezek alapvető szerepet játszanak mind a következő szakaszban, mind a valószínűségszámítás egyszerű statisztikai alkalmazásait tárgyaló következő fejezetben.

Általában egy X valószínűségi változó eloszlását az eloszlásfüggvényével adhatjuk meg. A 3.2.2. állításban láttuk, hogy azon esemény valószínűsége, hogy X értéke egy adott intervallumba esik, ennek segítségével könnyen felírható, valójában pedig igaz az is, hogy bármely "hasonlóképp felírt" esemény valószínűsége kiszámolható az eloszlásfüggvény értékeiből. Azaz: ha ismerjük az eloszlásfüggvényt, akkor ismerjük az eloszlást. Emlékeztetünk arra is, hogy míg a diszkrét valószínűségi változók esetén a $\mathbb{P}(X = t)$ valószínűségek megadása az eloszlásfüggvény megadásával ekvivalens, általában, speciálisan az abszolút folytonos esetben ez nincs így. Az abszolút folytonos valószínűségi változók esetén azonban a sűrűségfüggvény kiváltja a súlyfüggvény szerepét, azaz az eloszlás megadása (az eloszlásfüggvény felírásán kívül) történhet a sűrűségfüggvény megadásával is, hiszen egyrészt abból az eloszlásfüggvény integrálás segítségével adódik, másrészt a különböző események valószínűségei - szintén integrálás segítségével - a sűrűségfüggvényből közvetlenül megkaphatók. Ez utóbbira láttunk is példát a 3.3.2. állításban, amikor a $\mathbb{P}(s < X < t)$ valószínűséget adtuk meg ilyen módon, tehát az f_X sűrűségfüggvényt az $(s; t)$ intervallumon integrálva. Valójában ez a módszer szintén működik más, "hasonlóképp felírt" eseményekre is, de ezzel itt részletesen nem foglalkozunk. Az alábbiakban tehát megadjuk a fent említett három nevezetes eloszlás eloszlás- és sűrűségfüggvényeit, és megadjuk az eloszlások legfontosabb paramétereit és fontosabb speciális jellemzőit.

3.4.1. Egyenletes eloszlás

3.4.1. Definíció. Az X valószínűségi változót *egyenletes eloszlásúnak* nevezzük az $(a; b)$ intervallumon, ha sűrűségfüggvénye

$$f_X(t) = \begin{cases} \frac{1}{b-a}, & \text{ha } t \in (a; b), \\ 0 & \text{különben.} \end{cases}$$

Ennek jelölése: $X \sim U(a; b)$.

Erre az eloszlásra valójában már láttunk példát, hiszen a 3.2. szakaszban látott legegyszerűbb példánk, nevezetesen az Y valószínűségi változó egyenletes eloszlású a $(0; 1)$ intervallumon. Valóban, a (3.1) formula által adott eloszlásfüggvényt deriválva (leszámítva a 0 és 1 pontokat, ahol a függvény nem deriválható) megkapjuk az Y sűrűségfüggvényét:

$$f_Y(t) = \begin{cases} 1, & \text{ha } t \in (0; 1), \\ 0 & \text{különben.} \end{cases}$$

Vegyük észre, hogy annak a valószínűsége, hogy egy $(a; b)$ intervallumon egyenletes eloszlású X valószínűségi változó az intervallum egy $(s; t)$ részintervallumába esik, éppen a két intervallum arányával egyezik meg. Ha ugyanis $a < s < t < b$, akkor a 3.3.2. állítás (ill. az azt követő megjegyzés) szerint

$$\mathbb{P}(s < X < t) = \int_s^t f_X(y) dy = \int_s^t \frac{1}{b-a} dy = \frac{t-s}{b-a}.$$

Számoljuk ki az egyenletes eloszláshoz tartozó eloszlásfüggvényt. Ha $X \sim U(a; b)$, akkor

$$F_X(t) = \int_{-\infty}^t f_X(y) dy.$$

Ez nyilván 0, ha $t \leq a$, hiszen az a -tól balra az f_X függvény értéke 0. Ha viszont $t > b$, akkor valójában a és b között integrálunk, hiszen b -tól jobbra f_X ismét 0. Az $(a; b)$ intervallumon integrálva a konstans $1/(b-a)$ függvényt az eredmény 1 lesz. Végezetül, ha $t \in (a; b]$, akkor

$$F_X(t) = \int_a^t \frac{1}{b-a} dy = \frac{t-a}{b-a}.$$

Összefoglalva:

$$F_X(t) = \begin{cases} 0, & \text{ha } t \leq a, \\ \frac{t-a}{b-a}, & \text{ha } a \leq t < b, \\ 1, & \text{ha } t > b. \end{cases}$$

A következő lépésben megadjuk a fenti X változó várható értékét is. A definíció alapján

$$\mathbb{E}(X) = \int_{-\infty}^{\infty} t f_X(t) dt = \int_a^b \frac{t}{b-a} dt = \left[\frac{t^2}{2(b-a)} \right]_a^b = \frac{b^2 - a^2}{2(b-a)} = \frac{(b-a)(b+a)}{2(b-a)} = \frac{a+b}{2}.$$

Továbbá a (3.6) formula szerint

$$\begin{aligned}\mathbb{E}(X^2) &= \int_{-\infty}^{\infty} t^2 f_X(t) dt = \int_a^b \frac{t^2}{b-a} dt = \left[\frac{t^3}{3(b-a)} \right]_a^b \\ &= \frac{b^3 - a^3}{3(b-a)} = \frac{(b-a)(b^2 + ab + a^2)}{3(b-a)} = \frac{a^2 + ab + b^2}{3},\end{aligned}$$

és így az X szórásnégyzete

$$\begin{aligned}\mathbb{D}^2(X) &= \mathbb{E}(X^2) - \mathbb{E}(X)^2 = \frac{a^2 + ab + b^2}{3} - \frac{(a+b)^2}{4} \\ &= \frac{4a^2 + 4ab + 4b^2 - 3a^2 - 6ab - 3b^2}{12} = \frac{a^2 - 2ab + b^2}{12} = \frac{(b-a)^2}{12}.\end{aligned}$$

3.4.2. Folytonos örökifjú eloszlások

A nevezetes folytonos eloszlások tárgyalását olyan eloszlásokkal folytatjuk, amik tipikusan egy várakozási időt írnak le. Ilyenre már láttunk példát a 2.1. szakaszban tárgyalt geometriai eloszlások esetén. Ezek interpretálhatók úgy, mint független kísérletek sorozata esetén egy adott esemény bekövetkezéséig szükséges próbálkozások száma, ami valójában felfogható diszkrét várakozási időként is. A diszkrét jelző alatt itt pusztán azt értjük, hogy a felvett értékek egész számok (a diszkrét halmazok fogalmát pontosabban is meg lehetne adni, de ezzel itt nem foglalkozunk). Természetes kérdés, hogy ha a (véletlentől függő) várakozási időt folytonosan szeretnénk mérni, tehát értékeként egy tetszőleges nemnegatív valós számot megengednénk, akkor milyen eloszlást kapnánk. Látni fogjuk, hogy az ún. exponenciális eloszlású valószínűségi változók megfelelők lesznek erre a célra.

3.4.2. Definíció. Egy X folytonos valószínűségi változó *exponenciális eloszlású* $\lambda > 0$ paraméterrel, ha sűrűségfüggvénye

$$f_X(t) = \begin{cases} \lambda e^{-\lambda t}, & \text{ha } t > 0, \\ 0 & \text{különben.} \end{cases}$$

Ennek jelölése $X \sim \text{Exp}(\lambda)$

A fenti f_X függvény valóban sűrűségfüggvény, hiszen nemnegatív, és

$$\int_{-\infty}^{\infty} f_X(t) dt = \int_0^{\infty} \lambda e^{-\lambda t} dt = [-e^{-\lambda t}]_0^{\infty} = 1$$

teljesül. Számoljuk ki a hozzá tartozó $F_X(t)$ eloszlásfüggvényt. Ez persze 0, ha $t \leq 0$, pozitív t esetén pedig (a sűrűségfüggvény definíciója alapján)

$$\int_{-\infty}^t f_X(y) dy = \int_0^t \lambda e^{-\lambda y} dy = [-e^{-\lambda y}]_0^t = 1 - e^{-\lambda t},$$

azaz

$$F_X(t) = \begin{cases} 1 - e^{-\lambda t}, & \text{ha } t > 0, \\ 0 & \text{különben.} \end{cases}$$

Az exponenciális eloszlás bizonyos értelemben a geometriai eloszlás folytonos megfelelője. Hogy ezt lássuk, először definiáljuk az örökifjú tulajdonságot a folytonos esetben is:

3.4.3. Definíció. Az X valószínűségi változót *örökifjúnak* nevezzük a $[0; \infty)$ halmazon, amennyiben $\mathbb{P}(X \in [0; \infty)) = 1$ (azaz X értékei 1 valószínűséggel nemnegatívak), és

$$\mathbb{P}(X > s + t \mid X > s) = \mathbb{P}(X > t)$$

teljesül X -re minden $s, t \in [0; \infty)$ esetén.

3.4.1. Állítás. Ha $X \sim \text{Exp}(\lambda)$, akkor X örökifjú eloszlású a $[0; \infty)$ halmazon.

Bizonyítás. Az állítás egyszerű számolással adódik. Először is

$$\mathbb{P}(X \in [0; \infty)) = \mathbb{P}(X \geq 0) = 1 - \mathbb{P}(X < 0) = 1 - F_X(0) = 1,$$

továbbá ha $s, t \in [0; \infty)$, akkor (a geometriai eloszlásnál látottakhoz hasonlóan) felhasználva t nemnegativitását

$$\mathbb{P}(X > s + t \mid X > s) = \frac{\mathbb{P}(\{X > s + t\} \cap \{X > s\})}{\mathbb{P}(X > s)} = \frac{\mathbb{P}(X > s + t)}{\mathbb{P}(X > s)}$$

adódik. Ha most $t = 0$, akkor az utolsó tört éppen $1 = \mathbb{P}(X \geq 0) = \mathbb{P}(X > 0) = \mathbb{P}(X > t)$. Ha pedig $s = 0$, akkor az utolsó tört értéke szintén $\mathbb{P}(X > t)$. Ha viszont mindkét érték pozitív, akkor a valószínűségeket felírhatjuk az X eloszlásfüggvényének segítségével:

$$\begin{aligned} \frac{\mathbb{P}(X > s + t)}{\mathbb{P}(X > s)} &= \frac{1 - \mathbb{P}(X < s + t)}{1 - \mathbb{P}(X < s)} = \frac{1 - F_X(s + t)}{1 - F_X(s)} = \frac{e^{-\lambda(s+t)}}{e^{-\lambda s}} \\ &= e^{-\lambda t} = 1 - F_X(t) = 1 - \mathbb{P}(X < t) = \mathbb{P}(X > t), \end{aligned}$$

ahol az első és az utolsó lépésekben a komplementer eseményre való áttérésnél az egyenlőségjelet az X folytonossága miatt elhagytuk. \square

A diszkrét esethez hasonlóan most is igaz a fenti állítás megfordítása:

3.4.2. Tétel. Tegyük fel, hogy egy X valószínűségi változó örökifjú a $[0; \infty)$ halmazon. Ekkor $X \sim \text{Exp}(\lambda)$ teljesül valamilyen $\lambda > 0$ számra.

A geometriai és exponenciális eloszlások közötti hasonlóság a várható értékük esetén is megfigyelhető. Emlékeztetünk, hogy egy p paraméterű geometriai eloszlású valószínűségi változó várható értéke $1/p$. Ha pedig $X \sim \text{Exp}(\lambda)$, akkor parciális integrálást alkalmazva

$$\begin{aligned} \mathbb{E}(X) &= \int_{-\infty}^{\infty} t \cdot f_X(t) dt = \int_0^{\infty} \lambda t e^{-\lambda t} dt = [-t e^{-\lambda t}]_0^{\infty} + \int_0^{\infty} e^{-\lambda t} dt \\ &= 0 + \frac{1}{\lambda} \int_0^{\infty} \lambda e^{-\lambda t} dt = \frac{1}{\lambda} \int_{-\infty}^{\infty} f_X(t) dt = \frac{1}{\lambda}. \end{aligned}$$

Továbbá a (3.6) formula szerint

$$\begin{aligned} \mathbb{E}(X^2) &= \int_{-\infty}^{\infty} t^2 f_X(t) dt = \int_0^{\infty} t^2 \lambda e^{-\lambda t} dt = [-t^2 e^{-\lambda t}]_0^{\infty} + 2 \int_0^{\infty} t e^{-\lambda t} dt \\ &= 0 + \frac{2}{\lambda} \int_0^{\infty} t \lambda e^{-\lambda t} dt = \frac{2}{\lambda} \int_{-\infty}^{\infty} t f_X(t) dt = \frac{2}{\lambda} \cdot \mathbb{E}(X) = \frac{2}{\lambda^2}, \end{aligned}$$

és így az X szórásnégyzete

$$\mathbb{D}^2(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2 = \frac{2}{\lambda^2} - \frac{1}{\lambda^2} = \frac{1}{\lambda^2}.$$

3.4.1. Példa. Az örökifjú tulajdonság alapján az exponenciális eloszlás bizonyos esetekben valóban jól modellezhet várakozási időt. Tekintsük a következő példát: hullócsillagot várva kémleljük az eget éjszaka. Tegyük fel, hogy a hullócsillag észleléséig hátralévő idő nem függ az eddig eltelt várakozási idő hosszától. Ez tehát azt jelenti, hogy a várakozási időt (pl. órákban mérve) egy örökifjú, tehát a fenti tétel értelmében egy $X \sim Exp(\lambda)$ exponenciális eloszlású valószínűségi változóval írhatjuk le.

Tegyük fel továbbá, hogy $1 - e^{-2/3} \approx 0,4869$ eséllyel látunk hullócsillagot az első 20 percben. Meg tudjuk-e mondani, hogy mekkora az esélye, hogy az első egy órában látunk hullócsillagot? Az eddigi információk alapján (az eloszlást és a fent megadott valószínűséget felhasználva)

$$1 - e^{-2/3} = \mathbb{P}(X < 1/3) = F_X(1/3) = 1 - e^{-\lambda/3}$$

adódik, amiből $\lambda = 2$ következik. Ebből

$$\mathbb{P}(X < 1) = F_X(1) = 1 - e^{-2} \approx 0,8647.$$

Azt is láthatjuk, hogy átlagosan $\mathbb{E}(X) = 1/\lambda = 1/2$ órát kell várnunk az első hullócsillagig.

Gyakorlatok, feladatok

3.4.1. Gyakorlat. Legyen X exponenciális eloszlású valószínűségi változó, amiről tudjuk, hogy $\mathbb{P}(X > 3) = e^{-6}$.

a) Mi X eloszlásának paramétere? b) $\mathbb{P}(X < 2) = ?$ c) $\mathbb{E}(X) = ?$ d) $\mathbb{D}(X) = ?$

3.4.2. Gyakorlat. Tegyük fel, hogy egy adott mosógéptípus átlagosan 2 évig bírja az első meghibásodásig, és az első meghibásodás időpontja folytonos, örökifjú eloszlást követ. Mi a valószínűsége, hogy az első 3 év során nem hibásodik meg, ha tudjuk, hogy az első 2 évben hibátlanul működött?

3.4.3. Gyakorlat. Egy radioaktív atom élettartama években mérve exponenciális eloszlású valószínűségi változó. Az atom 32 év leforgása alatt 0,5 valószínűséggel bomlik el. Mennyi az esélye, hogy az atom nem bomlik el 24 év alatt? Mennyi időn belül bomlik el az atom 0,95 valószínűséggel?

3.4.4. Feladat. Mutassuk meg, hogy ha $X \sim Exp(\lambda)$, akkor a $\lceil X \rceil$ valószínűségi változó (ahol $\lceil \cdot \rceil$ a felső egészrész jelöli) geometriai eloszlású $1 - e^{-\lambda}$ paraméterrel.

3.4.3. Normális eloszlás

Az alábbiakban megismerkedünk a jegyzetben szereplő legfontosabb eloszláscsaláddal, az ún. normális eloszlásokkal. Előzetes szemléletes motiváció helyett itt csak megemlítjük, hogy a hátralévő szakaszokban ezek központi szerephez jutnak majd. Különösképpen igaz ez az ún. sztenderd normális eloszlásra, így a szakaszt ennek ismertetésével kezdjük, majd segítségével előállítjuk az eloszláscsalád többi tagját is. Ez a tárgyalásmód lényegesen leegyszerűsíti a számolásokat, és rámutat a különböző eloszlások kapcsolatára is.

3.4.4. Definíció. Az X folytonos valószínűségi változót *sztenderd normális eloszlásúnak* nevezzük, ha sűrűségfüggvénye

$$\varphi(t) := \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} \quad (t \in \mathbb{R}).$$

Ennek jelölése: $X \sim N(0; 1)$.

A jegyzet hátralévő részére a φ jelölést rögzítjük. Ezen a ponton egyelőre az sem világos, hogy ez a függvény valóban egy eloszlást ad meg. Ehhez az

$$\int_{-\infty}^{\infty} \varphi(t) dt = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt = 1$$

egyenlőséget kellene ellenőrizni, amely valóban teljesül (tehát, mivel $\varphi(t) > 0$ is teljesül minden t -re, valóban egy valószínűségeloszlást definiáltunk), a részletes számolást azonban itt nem fogjuk elvégezni. Az érdeklődő olvasó megtalálja azt mind a [2] könyvben, mind az [5] jegyzetben.

3.4.3. Állítás. *Legyen $X \sim N(0; 1)$, ekkor $\mathbb{E}(X) = 0$ és $\mathbb{D}^2(X) = 1$.*

Bizonyítás. A várható érték definíciója szerint

$$\mathbb{E}(X) = \int_{-\infty}^{\infty} t\varphi(t) dt = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} t e^{-\frac{t^2}{2}} dt = \frac{1}{\sqrt{2\pi}} \left[-e^{-\frac{t^2}{2}} \right]_{-\infty}^{\infty} = 0.$$

A szórásnégyzet meghatározásához először (szokás szerint) kiszámoljuk X^2 várható értékét a transzformált várható értékére vonatkozó állítás segítségével:

$$\begin{aligned} \mathbb{E}(X^2) &= \int_{-\infty}^{\infty} t^2 \varphi(t) dt = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} t^2 e^{-\frac{t^2}{2}} dt = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} t \cdot \frac{\partial}{\partial t} \left(-e^{-\frac{t^2}{2}} \right) dt \\ &= \left[-\frac{1}{\sqrt{2\pi}} t e^{-\frac{t^2}{2}} \right]_{-\infty}^{\infty} + \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt = 0 + \int_{-\infty}^{\infty} \varphi(t) dt = 1, \end{aligned}$$

hiszen φ az X sűrűségfüggvénye. Így tehát $\mathbb{D}^2(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2 = 1 - 0 = 1$. □

A következőkben a sztenderd normális eloszlású valószínűségi változók lineáris transzformáltjának eloszlását fogjuk vizsgálni. Ehhez a következő, valójában jóval általánosabb állítást fogjuk használni.

3.4.4. Lemma. *Legyen X egy folytonos valószínűségi változó, melynek sűrűségfüggvénye $f_X(t)$. Legyenek továbbá $\sigma, \mu \in \mathbb{R}$, $\sigma > 0$, ekkor az $Y = \sigma X + \mu$ transzformált is folytonos valószínűségi változó, melynek sűrűségfüggvénye*

$$f_Y(t) = \frac{1}{\sigma} f_X \left(\frac{t - \mu}{\sigma} \right).$$

Bizonyítás. Az eloszlásfüggvény definíciója szerint

$$F_Y(t) = \mathbb{P}(Y < t) = \mathbb{P}(\sigma X + \mu < t) = \mathbb{P} \left(X < \frac{t - \mu}{\sigma} \right)$$

minden $t \in \mathbb{R}$ esetén, ahol felhasználtuk, hogy $\sigma > 0$, és így az osztásnál az egyenlőtlenség iránya nem változik. Ezt a valószínűséget kifejezhetjük az X változó f_X sűrűségfüggvényének segítségével a következőképp:

$$\mathbb{P} \left(X < \frac{t - \mu}{\sigma} \right) = \int_{-\infty}^{\frac{t - \mu}{\sigma}} f_X(s) ds.$$

Most az integrálban az $s = \frac{u-\mu}{\sigma}$ helyettesítést végezzük, és így ds helyett $\frac{1}{\sigma}du$ -t írunk, továbbá ha $s = \frac{t-\mu}{\sigma}$, akkor $u = \sigma s + \mu = t$, így tehát az integrálás felső határa t lesz. Azaz

$$F_Y(t) = \int_{-\infty}^t \frac{1}{\sigma} f_X\left(\frac{u-\mu}{\sigma}\right) du,$$

ezért a folytonos valószínűségi változó ill. a sűrűségfüggvény definíciója alapján adódik az állítás. \square

A lemma alapján azonnal adódik, hogy ha $X \sim N(0; 1)$, és $\sigma, \mu \in \mathbb{R}$, $\sigma > 0$, akkor az $Y = \sigma X + \mu$ valószínűségi változó folytonos, és a sűrűségfüggvénye

$$(3.7) \quad f_Y(t) = \frac{1}{\sigma} \varphi\left(\frac{t-\mu}{\sigma}\right) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(t-\mu)^2}{2\sigma^2}}.$$

3.4.5. Definíció. Egy Y folytonos valószínűségi változó *normális eloszlású* μ és $\sigma^2 > 0$ paraméterekkel, ha sűrűségfüggvénye a (3.7) jobb oldalán álló függvény. Ennek jele $Y \sim N(\mu; \sigma^2)$.

A sztenderd normális eloszlás jelölésénél a μ és σ^2 helyére 0 és 1 került, vagyis a várható érték és a szórásnégyzet. Ennek alapján már sejthető, hogy ez általában is így van, és ez jelen helyzetben már könnyedén igazolható. Egy $N(\mu; \sigma^2)$ eloszlású változó ezen jellemzőinek meghatározásához ugyanis elegendő azokat egy konkrét változó esetén kiszámolni, ilyet pedig a fentiek szerint tudunk konstruálni. Legyen ugyanis X egy sztenderd normális eloszlású valószínűségi változó, ekkor $\sigma X + \mu$ a látottak alapján normális eloszlású μ és σ^2 paraméterekkel (itt σ -t a σ^2 pozitív gyökének választjuk), így ezen eloszlás várható értéke

$$(3.8) \quad \mathbb{E}(\sigma X + \mu) = \sigma \mathbb{E}(X) + \mu = 0 + \mu = \mu$$

a várható érték linearitása ill. a sztenderd normális eloszlás várható értéke alapján, továbbá

$$(3.9) \quad \mathbb{D}^2(\sigma X + \mu) = \mathbb{D}^2(\sigma X) = \sigma^2 \mathbb{D}^2(X) = \sigma^2$$

a szórásnégyzet tulajdonságai ill. a sztenderd normális eloszlás szórásnégyzete miatt. Azaz: **a normális eloszlás paraméterei a várható érték és a szórásnégyzet.**

Láttuk, hogy a sztenderd normális eloszlás transzformáltjaként előáll bármely normális eloszlás. A konstrukció azonban fordítva is működik, bármely normális eloszlásból előállítható a sztenderd normális. Legyen ugyanis $Y \sim N(\mu; \sigma^2)$, és tekintsük az $X = \frac{Y-\mu}{\sigma}$ változót. Ekkor

$$\mathbb{E}(X) = \mathbb{E}\left(\frac{Y-\mu}{\sigma}\right) = \frac{\mathbb{E}(Y) - \mu}{\sigma} = 0$$

és

$$\mathbb{D}^2(X) = \mathbb{D}^2\left(\frac{Y-\mu}{\sigma}\right) = \mathbb{D}^2\left(\frac{Y}{\sigma}\right) = \frac{1}{\sigma^2} \mathbb{D}^2(Y) = 1.$$

Amit be kell még látni, hogy X is normális eloszlású, ami azonban a fenti lemma alapján adódik, amely szerint X sűrűségfüggvénye

$$\sigma f_Y(\sigma t + \mu) = \sigma \left(\frac{1}{\sigma} \varphi\left(\frac{(\sigma t + \mu) - \mu}{\sigma}\right) \right) = \varphi(t).$$

A fent definiált X változót az Y *sztenderdizáltjának* nevezzük.

Innen már egyszerűen látszik, hogy ha $Y \sim N(\mu; \sigma^2)$, továbbá $\rho, \nu \in \mathbb{R}$, $\rho > 0$, akkor $\rho Y + \nu$ is normális eloszlású. Legyen ugyanis X az Y sztenderdizáltja, ekkor $Y = \sigma X + \mu$ teljesül, így $\rho Y + \nu = \rho(\sigma X + \mu) + \nu = \rho\sigma X + \rho\mu + \nu$. Mivel $X \sim N(0; 1)$, így az eddigiek alapján adódik a következő

3.4.5. Állítás. Legyen $Y \sim N(\mu; \sigma^2)$, továbbá legyenek $\rho, \nu \in \mathbb{R}$ valós számok, ahol $\rho > 0$. Ekkor $\rho Y + \nu \sim N(\rho\mu + \nu; \rho^2\sigma^2)$.

Az állítás lényegesebb része, hogy a $\rho Y + \nu$ transzformált is normális eloszlású, amennyiben Y az. A paraméterek (ha valaki nem szeretné fejben tartani) egyszerűen adódnak a várható érték linearitásából és a szórásnégyzet tulajdonságaiból, ahogy azt a (3.8)-ban és (3.9)-ben is láttuk.

A normális eloszlás eloszlásfüggvénye

Ahogy a normális eloszlás definíciója esetén is, az eloszlásfüggvények vizsgálatát is a sztenderd normális eloszlás eloszlásfüggvényével kezdjük, amely a folytonos változók definíciója szerint

$$\Phi(t) := \int_{-\infty}^t \varphi(s) ds.$$

A Φ jelölést a továbbiakban rögzítjük.

Nem véletlen, hogy a fenti integrál kiszámolását nem részleteztük tovább. Ennek oka, hogy a Φ függvény (bár a fenti integrál segítségével definiált, azonban) nem fejezhető ki zárt alakban elemi függvények segítségével, és így a Newton–Leibniz-formulát nem tudjuk közvetlenül alkalmazni. A Φ értékei így számoló- ill. számítógép vagy táblázat segítségével határozhatók meg. Mivel ez a függvény a páros φ sűrűségfüggvény integrálja, így kielégíti a következő egyenletet. (Egy $f : \mathbb{R} \rightarrow \mathbb{R}$ függvényt párosnak nevezünk, ha $f(-t) = f(t)$ teljesül minden $t \in \mathbb{R}$ -re.)

3.4.6. Állítás. Minden $t \in \mathbb{R}$ esetén $\Phi(-t) = 1 - \Phi(t)$.

Bizonyítás. Ha $t \in \mathbb{R}$, akkor

$$\begin{aligned} \Phi(-t) &= \int_{-\infty}^{-t} \varphi(s) ds = \int_{-\infty}^{\infty} \varphi(s) ds - \int_{-t}^{\infty} \varphi(s) ds \\ &= 1 - \int_t^{-\infty} -\varphi(-u) du = 1 - \int_{-\infty}^t \varphi(u) du = 1 - \Phi(t), \end{aligned}$$

ami éppen az állítás. □

Megjegyzés. A fenti számolás mutatja, hogy az állításban szereplő formula érvényes minden páros sűrűségfüggvénnyel rendelkező eloszlás eloszlásfüggvényére, hiszen csak azt használtuk ki a fenti számolásban, hogy φ -nek a valós egyenesen vett integrálja 1, ill. hogy $\varphi(-t) = \varphi(t)$.

A fenti állítást gyakran alkalmaznunk kell, ha táblázatból keressük ki a Φ értékeit, hiszen éppen ennek az egyenletnek köszönhetően egy táblázatban elegendő megadni a függvény pozitív helyen felvett értékeit. Egy ilyen táblázat található a jegyzet végén szereplő "Táblázatok" c. szakaszban. A Φ függvényre nem csak a sztenderd normális eloszlás esetén van szükség, ugyanis segítségével kifejezhető tetszőleges normális eloszlás eloszlásfüggvénye:

3.4.7. Állítás. Legyen $Y \sim N(\mu; \sigma^2)$, ekkor az Y eloszlásfüggvénye $F_Y(t) = \Phi\left(\frac{t-\mu}{\sigma}\right)$.

Bizonyítás. Jelölje X az Y sztenderdizáltját, ekkor $Y = \sigma X + \mu$ (ahol σ a σ^2 pozitív gyöke). Az eloszlásfüggvény definíciója alapján

$$F_Y(t) = \mathbb{P}(Y < t) = \mathbb{P}(\sigma X + \mu < t) = \mathbb{P}\left(X < \frac{t - \mu}{\sigma}\right) = \Phi\left(\frac{t - \mu}{\sigma}\right),$$

mert Φ az X eloszlásfüggvénye. □

A normális eloszlás alkalmazásai

A normális eloszlást a gyakorlatban számos esetben alkalmazhatjuk. Ennek okai a következő szakaszban, a centrális határeloszlás tételének tárgyalásakor valamivel tisztábban látszanak majd, minden esetre az alkalmazásokban tipikusan akkor találkozunk a normális eloszlással, amikor nagyszámú, független, de egymagában elhanyagolható méretű hatás eredményeként létrejövő valószínűségi változó eloszlását modellezzük. Ilyen lehet egy fizikai mérés eredménye, pl. egy terület átlaghőmérséklete egy adott hónapban.

3.4.2. Példa. Egy tartályból a gyártási folyamat végeztével mintát veszünk. Tegyük fel, hogy a minta hőmérséklete Celsius-fokban mérve $N(-2; 1,69)$ eloszlású. Mi a valószínűsége, hogy a minta hőmérséklete nagyobb, mint 0°C ?

Jelölje a minta hőmérsékletét Y , ekkor a $\mathbb{P}(Y > 0)$ mennyiséget keressük. A komplementer eseményre áttérve, az Y folytonosságát és az előző állítást is felhasználva

$$\mathbb{P}(Y > 0) = 1 - \mathbb{P}(Y < 0) = 1 - \Phi\left(\frac{0 - (-2)}{1,3}\right) = 1 - \Phi\left(\frac{2}{1,3}\right) \approx 0,0618.$$

Gyakorlatok, feladatok

3.4.5. Gyakorlat. Egy gép a beállítása szerint 2 kg lisztet adagol a zacskókba, de a technológia következtében a zacskóba került liszt mennyisége $N(\mu; 0,002^2)$ eloszlást követ. Előzetes megfigyelésekből lehet tudni, hogy 0,01 annak a valószínűsége, hogy a zacskóban a liszt mennyisége kevesebb 2 kg-nál. Mennyi μ értéke?

3.4.6. Gyakorlat. Egy normális eloszlású valószínűségi változó 0,2 valószínűséggel vesz fel 10-nél kisebb értéket és 0,3 valószínűséggel 14-nél nagyobb értéket. Mik az eloszlás paraméterei?

3.4.7. Gyakorlat. Texasban a hőmérsékletet Fahrenheit fokokban mérik. Megállapították, hogy az ottani hőmérséklet eloszlása nyaranta $N(86; 16)$. Hogyan változik meg az eloszlás, ha áttérünk Celsius-skálára?

$$\left(\frac{5}{9} (^{\circ}F - 32)\right) = ^{\circ}C$$

3.4.8. Gyakorlat. Legyen $X \sim N(10; 9)$ és $Z = \left(\frac{X-10}{3}\right)^2$. Határozzuk meg a Z változó eloszlásfüggvényét.

3.4.9. Feladat. (M) Mutassuk meg, hogy $t > 0$ esetén $\frac{t\varphi(t)}{t^2 + 1} < 1 - \Phi(t) < \frac{\varphi(t)}{t}$.

3.5. Határeloszlás-tételek

Ebben a szakaszban a valószínűségi számítás két alapvető és kiemelten fontos tételével foglalkozunk. Pontosabb lenne tételek helyett tételkörökről beszélni, hiszen az alábbi állításoknak többféle változata létezik, melyek lényegében ugyanazt a jelenséget írják le különböző módokon vagy épp más-más feltételek mellett. Bár a határeloszlás-tételek témaköréből itt csupán ízelítőt adunk, és a tárgyalt állítások bizonyításai is túlmutatnak a kurzus keretein, de a terület fontossága és alkalmazásai miatt röviden mégis foglalkoznunk kell vele. Először az ún. nagy számok törvényével ismerkedünk meg, majd a de Moivre-Laplace tételt tárgyaljuk, amely speciális esete a szakasz végén ismertetett centrális határeloszlás tételének. Utóbbi két tétel rávilágít arra is, hogy miért játszik a normális eloszlás központi szerepet a valószínűségi számításban és alkalmazásaiban.

3.5.1. A nagy számok törvénye

Tekintsünk egy dobókockát. Hogyan tudjuk megállapítani, hogy mennyi a 6-os dobás valószínűsége? Egy ideális (homogén és szimmetrikus) kocka esetén természetesen szimmetriai megfontolások egyszerűen vezethetnek olyan következtetésre, hogy hosszú távon bármilyen eredmény, így a 6-os is ugyanolyan gyakorisággal fordul elő, azaz a 6-os dobás valószínűsége $1/6$ -nak tekinthető.

De mi mondható akkor, ha a kocka nem teljesen szimmetrikus vagy nem teljesen homogén? Egy kellően aszimmetrikus kocka esetén már néhány dobás után feltűnhet, hogy bizonyos eredmények gyakrabban adódnak, mint mások. Ha aztán a dobássorozatot folytatjuk, és ez nem változik, akkor az az érzés alakul ki bennünk, hogy bizonyos kimenetek valószínűbbek, mint mások. Amennyiben a dobásokat egészen hosszan ismétljük, és az előfordulások számát feljegyezzük, akkor azt tapasztalhatjuk, hogy ezek aránya az összes dobás számához mérten lényegében állandósulni látszik, vagy pontosabban valamilyen érték körül ingadozik, még hozzá a dobások számának növelésével egyre kisebb mértékben. Ezt az értéket joggal tekinthetjük az adott kimenetel valószínűségének, hiszen intuitíven éppen így definiáljuk azt: az a szám, ami megadja, hogy hosszú távon milyen arányban adódik egy bizonyos kimenetel. Ebből persze az is következik, hogy így nem tudjuk pontosan meghatározni a valószínűséget, csak megközelíteni tudjuk azt.

A fenti gondolatmenet helyességéről azonban elméleti érvelésekkel is meg kell győződnie, hiszen matematikai igazságokat tapasztalati úton megsejteni ugyan lehet, de bizonyítani nem. Adjunk matematikai modellt a fenti kísérletsorozatra. Ez tehát egy dobókockánál adott kimenetek egy eloszlása, és például a 6-os dobás valószínűségét próbáljuk meg kísérletekkel megállapítani. Ezen a kísérletek egymás utáni dobások, melyek eredménye nem befolyásol más dobásokat. Legyen A_i az az esemény, hogy az i -edik dobásnál 6-ost dobunk. Feltételezünk szerint ezek az események együttesen függetlenek, és mindegyikének valószínűsége egy p szám, amit meg szeretnénk határozni. A kísérletsorozatunk egy adott kimenetele most egy B_1, \dots, B_n eseménysorozattal modellezhető, ahol $B_i = A_i$ vagy $B_i = \bar{A}_i$ minden i -re. Ezek az események az 1.2.3. definíció után tett megjegyzéseink értelmében szintén együttesen függetlenek, és mindegyikük a valószínűsége p vagy $1 - p$ attól függően, hogy épp A_i vagy annak komplementere következik be az i -edik dobásnál.

A 6-osok relatív gyakorisága kifejezhető a hozzájuk tartozó $\mathbb{1}_{A_1}, \dots, \mathbb{1}_{A_n}$ indikátorváltozókkal, amelyek a fentiekből adódóan azonos eloszlásúak és a 2.2.1. állítás szerint együttesen függetlenek. Mivel a 6-osok száma egyszerűen ezen változók összege, így a relatív gyakoriság éppen

$$\frac{\sum_{i=1}^n \mathbb{1}_{A_i}}{n}.$$

Erről a kifejezésről szeretnénk belátni, hogy az A_i események (közös) valószínűségéhez tart.

Itt óvatosságnak kell lennünk, hiszen a fenti kifejezés valójában valószínűségi változók átlaga, ami maga is egy valószínűségi változó, tehát egy függvény, míg a valószínűség egyetlen szám. Ez azonban valójában nem fog gondot okozni, de jobban megvilágítja a lényegét, ha elvonatkoztatunk az adott szituációtól, és általánosan tekintjük az X_1, \dots, X_n, \dots együttesen független, azonos eloszlású valószínűségi változók egy sorozatát, amelyek persze egyazon valószínűségi mezőn vannak értelmezve. Jegyezzük meg, hogy az együttes függetlenséget csak véges sok valószínűségi változóra definiáltuk, ebben az esetben pedig végtelen sorozatról van szó. De ez a probléma könnyen orvosolható: itt függetlenség alatt egyszerűen azt értjük,

hogy a változók bármely véges részhalmaza független. Tekintsük tehát a változók

$$\frac{\sum_{i=1}^n X_i}{n}$$

átlagát. A változók "átlagos viselkedéséről" - legalábbis intuitívan - a várható érték ad információt, tehát ennek az átlagnak a változók (azonos) várható értékéhez kéne tartania.

Hogyan tarthat tehát egy függvény egy számhoz? A valószínűségszámításban többféle konvergenciafogalmat is szokás értelmezni, mi most a függvények pontonkénti konvergenciáját használjuk, aminek segítségével az ún. 1 valószínűségű konvergenciát kapjuk. Ahogy az X_i függvények, úgy az átlagaik sorozatának tagjai is egy adott Ω eseménytéren értelmezett függvények, melyeket kiértékelhetünk egy adott $\omega \in \Omega$ pontban, és az így kapott számsorozat határértéke (amennyiben létezik és véges) egy ω -hoz rendelt szám, tehát ez az ún. pontonkénti limesz szintén egy függvényt definiál Ω -n. A nagy számok erős törvénye éppen azt mondja ki, hogy bizonyos feltételek mellett ez a pontonkénti határérték egy konstans, mégpedig az X_i -k (minden i -re azonos) várható értéke, legalábbis "majdnem mindig". A "majdnem mindig" kifejezés a valószínűség nyelvén úgy fogalmazható meg, hogy az adott eseménynek 1 a valószínűsége, innen tehát az 1 valószínűségű konvergencia név. Minden adott tehát a szakasz fő eredményének kimondásához:

3.5.1. Tétel (A nagy számok erős törvénye). *Legyen X_1, \dots, X_n, \dots egy $(\Omega, \mathcal{A}, \mathbb{P})$ eseménytéren értelmezett független, azonos eloszlású valószínűségi változók egy sorozata, melyekre $\mathbb{D}(X_i) = \sigma < \infty$. Ekkor $\mathbb{E}(X_i) = \mu$ is véges, és*

$$\mathbb{P}\left(\lim_{n \rightarrow \infty} \frac{\sum_{i=1}^n X_i}{n} = \mu\right) = 1,$$

ahol a határérték úgy értendő, mint az

$$\omega \mapsto \frac{\sum_{i=1}^n X_i(\omega)}{n} \quad (\omega \in \Omega)$$

függvények pontonkénti határértéke.

A tételt tehát a korábbi példánkra alkalmazva, azaz (a fenti jelöléseket használva) az $X_i = \mathbb{1}_{A_i}$ választással azt kapjuk, hogy a 6-os dobások relatív gyakorisága 1 valószínűséggel tart az $\mathbb{1}_{A_i}$ indikátorváltozók várható értékéhez, ami pedig az A_i eseményeknek, tehát a 6-os dobásnak a valószínűsége. Ez tehát, noha kezdeti intuíciónkkal összhangban van, nem pusztán tapasztalati tény, hanem immár egy matematikai törvény.

Összességében tehát úgy fogalmazhatunk, hogy a valószínűségi változókat kiátlagolva egy "jó" becslést kapunk a várható értékre. Hogy mit jelent a jó becslés, és hogy hány változó átlagát kell tekintenünk, az tulajdonképpen már a matematikai statisztika témakörébe tartozik, amellyel a következő fejezetben fogunk foglalkozni.

3.5.2. A de Moivre–Laplace-tétel

Míg a nagy számok törvénye információt ad a változók átlagáról, ebben és a következő szakaszban a változók összegének *eloszlásával* foglalkozunk. Vizsgáljuk meg ezt a kérdést először a fenti első példánkban (pontosabban annak általános alakjában). Legyenek tehát az A_i események együttesen függetlenek, ahol $1 \leq i \leq n$, továbbá tegyük fel, hogy $\mathbb{P}(A_i) = p$

teljesül minden i -re. Tekintsük a hozzájuk tartozó indikátorváltozók $S_n = \sum_{i=1}^n \mathbb{1}_{A_i}$ összegét. Ennek eloszlását jól ismerjük a 2.1.10. példából: $S_n \sim B(n; p)$.

Az egyszerűség kedvéért először tovább specializáljuk a problémát, azaz a $p = \frac{1}{2}$ esettel fogunk foglalkozni. Tegyük fel tehát, hogy egy szabályos pénzérmét n -szer feldobunk, és jelölje X a fejek számát. Ekkor $X \sim B(n, \frac{1}{2})$, és tetszőleges $0 \leq k \leq n$ egész esetén

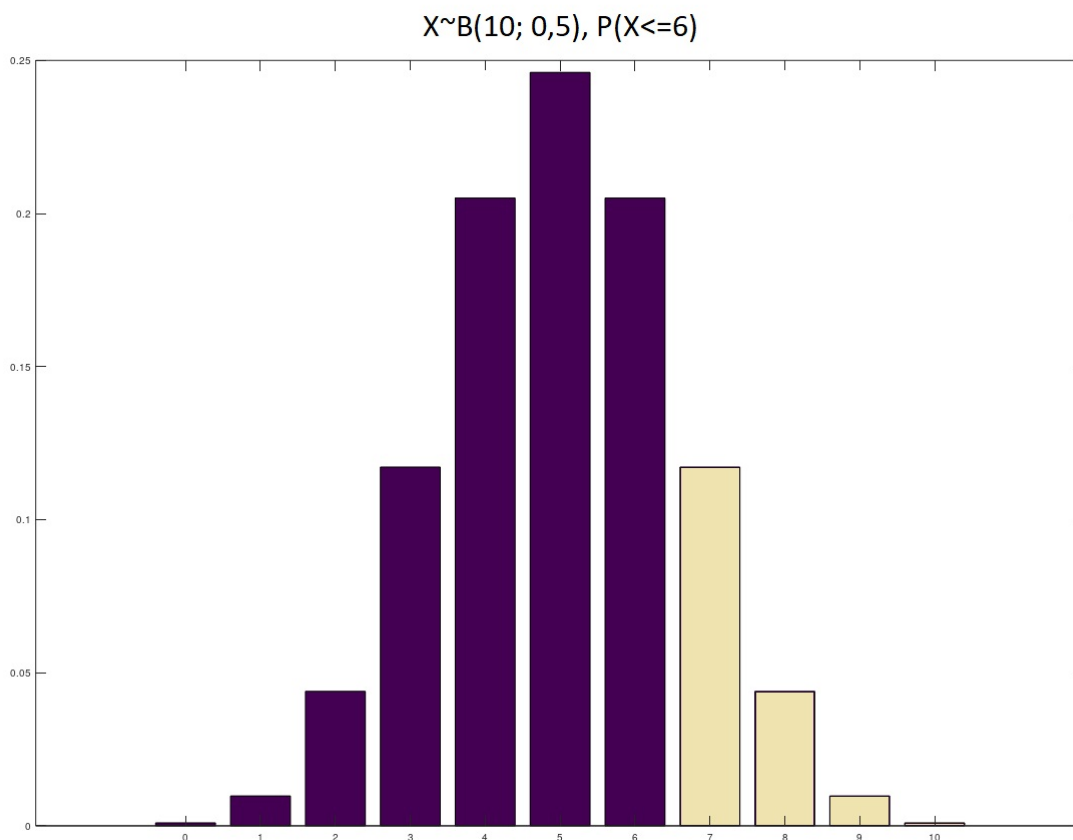
$$\mathbb{P}(X = k) = \binom{n}{k} \cdot \frac{1}{2^n}.$$

Ez a valószínűség persze könnyen számolható, amíg n viszonylag kicsi. De már ebben az esetben is jóval fáradtságosabb lehet a

$$(3.10) \quad \mathbb{P}(X \leq k) = \sum_{i=0}^k f_X(i) = \sum_{i=0}^k \mathbb{P}(X = i) = \frac{1}{2^n} \sum_{i=0}^k \binom{n}{i} = \sum_{i=0}^k \frac{n!}{2^n i! (n-i)!}$$

valószínűség kiszámolása, hiszen itt egy $k+1$ tagú összeget számolunk. Ráadásul ha n nagy, akkor az összegben szereplő binomiális együtthatók számolása sem egyszerű.

Ahelyett, hogy a fenti valószínűséget pontosan kiszámolnánk, egy jó közelítést fogunk adni rá. A számolásokat itt nem részletezzük, de vizuálisan jó képet kaphatunk arról, hogy mi is történik pontosan. Az (3.10) összegben k db valószínűséget adunk össze. Ezeket a valószínűségeket egy oszlopdiagramon ábrázolva látható, hogy ez tulajdonképpen k darab téglalap területének az összegeként is felfogható, ahol a téglalapok oldalainak hossza 1 ill. $\mathbb{P}(X = i)$. A következő ábra szemlélteti az $n = 10$, $k = 6$ esetet.



3.9. ábra. A binomiális eloszlás súlyfüggvényének értékei megadhatók téglalapok területeként

Hogy analitikusan jobban kezelhető legyen ez az összeg, először is a X értékészletét "szimmetrizáljuk", azaz eltoljuk balra $\frac{n}{2}$ -vel. Az $X - \frac{n}{2}$ változó már a 0-ra szimmetrikus $[-\frac{n}{2}; \frac{n}{2}]$ intervallumban veszi fel az értékeit. Vegyük észre, hogy valójában a várható értékkel toltunk el, hiszen $\mathbb{E}(X) = \frac{n}{2}$, tehát voltaképp úgy transzformáltuk a változónkat, hogy a transzformált várható értéke 0 legyen, hiszen

$$(3.11) \quad \mathbb{E}(X - \mathbb{E}(X)) = \mathbb{E}(X) - \mathbb{E}(\mathbb{E}(X)) = 0.$$

Emellett szükség lesz még egy normalizálásra is: a fenti különbséget leosztjuk az X szórásával, azaz $\mathbb{D}(X) = \frac{1}{2}\sqrt{n}$ -nel. A szórás tulajdonságai alapján

$$(3.12) \quad \mathbb{D}\left(\frac{X - \mathbb{E}(X)}{\mathbb{D}(X)}\right) = \mathbb{D}\left(\frac{X}{\mathbb{D}(X)}\right) = \frac{1}{\mathbb{D}(X)}\mathbb{D}(X) = 1,$$

azaz az $\frac{X - \mathbb{E}(X)}{\mathbb{D}(X)}$ egy olyan transzformáltja X -nek, melynek várható értéke 0 és szórása 1. Persze itt X eloszlását nem használtuk ki, azaz (3.11) és (3.12) igaz egy tetszőleges valószínűségi változóra, amelynek véges a várható értéke, ill. véges és pozitív a szórása (a pozitivitás is szükséges, hiszen osztunk vele). Vegyük észre, hogy ugyanezt csináltuk, amikor egy normális eloszlású változót sztenderd normális eloszlásúvá transzformáltunk. Ebben az esetben ezt sztenderdizálásnak neveztük, és ezt az elnevezést általában is használjuk:

3.5.1. Definíció. Legyen X egy olyan valószínűségi változó, melyre $\mathbb{E}(X)$ és $\mathbb{D}(X)$ véges, továbbá $\mathbb{D}(X) > 0$. Ekkor az

$$\frac{X - \mathbb{E}(X)}{\mathbb{D}(X)}$$

változót az X sztenderdizáltjának nevezzük.

Térjünk most vissza a $n, \frac{1}{2}$ paraméterű binomiális eloszláshoz, és fejezzük ki a $\mathbb{P}(X \leq k)$ valószínűséget a sztenderdizált segítségével:

$$(3.13) \quad \mathbb{P}(X \leq k) = \mathbb{P}\left(\frac{X - \frac{n}{2}}{\frac{1}{2}\sqrt{n}} \leq \frac{k - \frac{n}{2}}{\frac{1}{2}\sqrt{n}}\right).$$

A sztenderdizált az értékeit már a $[-\sqrt{n}; \sqrt{n}]$ intervallumon veszi fel, és a szomszédos értékek között $2/\sqrt{n}$ a különbség. Éppen ezért, ha fenti valószínűséget egy oszlopdiagramon szereplő téglalapok területének összegével szeretnénk kifejezni, akkor - mivel a téglalapok szélessége $2/\sqrt{n}$ -szerese az X esetében adódó 1 szélességnek - a terület (azaz a valószínűség) akkor marad változatlan, ha a téglalapok magasságát $\sqrt{n}/2$ -vel szorozzuk.

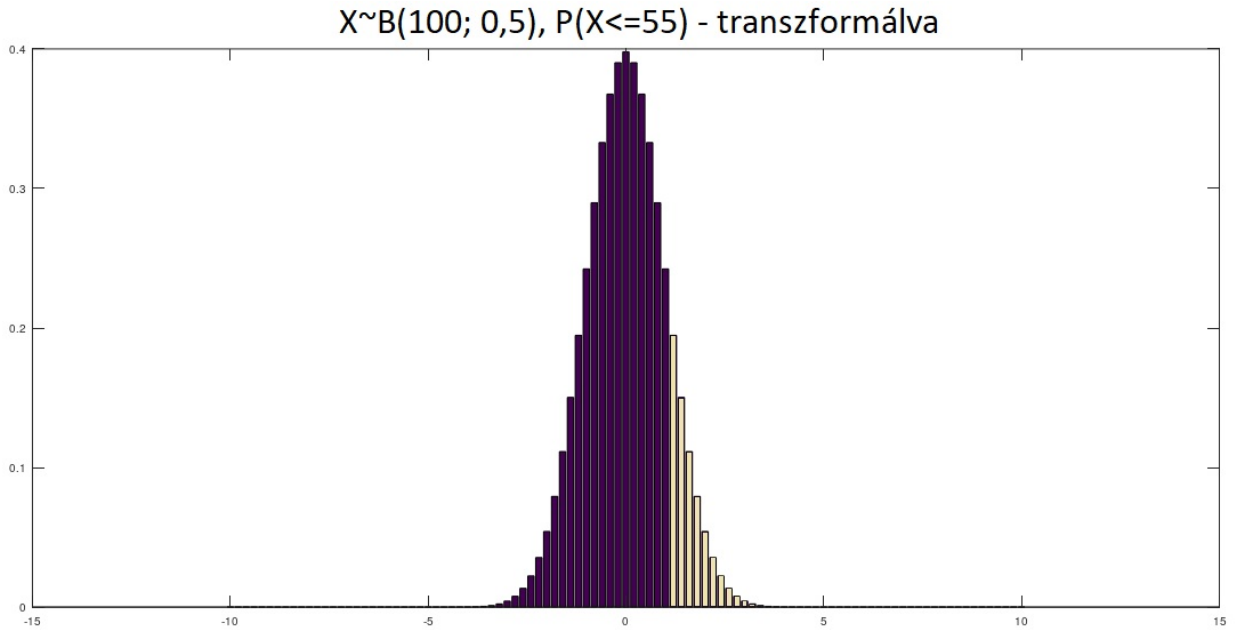
Ennek a transzformációnak az eredménye látható az alábbi ábrán az $n = 100$ értékre. Ha most a téglalapok helyett csak azok magasságát ábrázoljuk a grafikonon, akkor megfigyelhető, hogy az így kapott pontok jól illeszkednek egy speciális görbére, nevezetesen a sztenderd normális eloszlás

$$\varphi(t) := \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}}$$

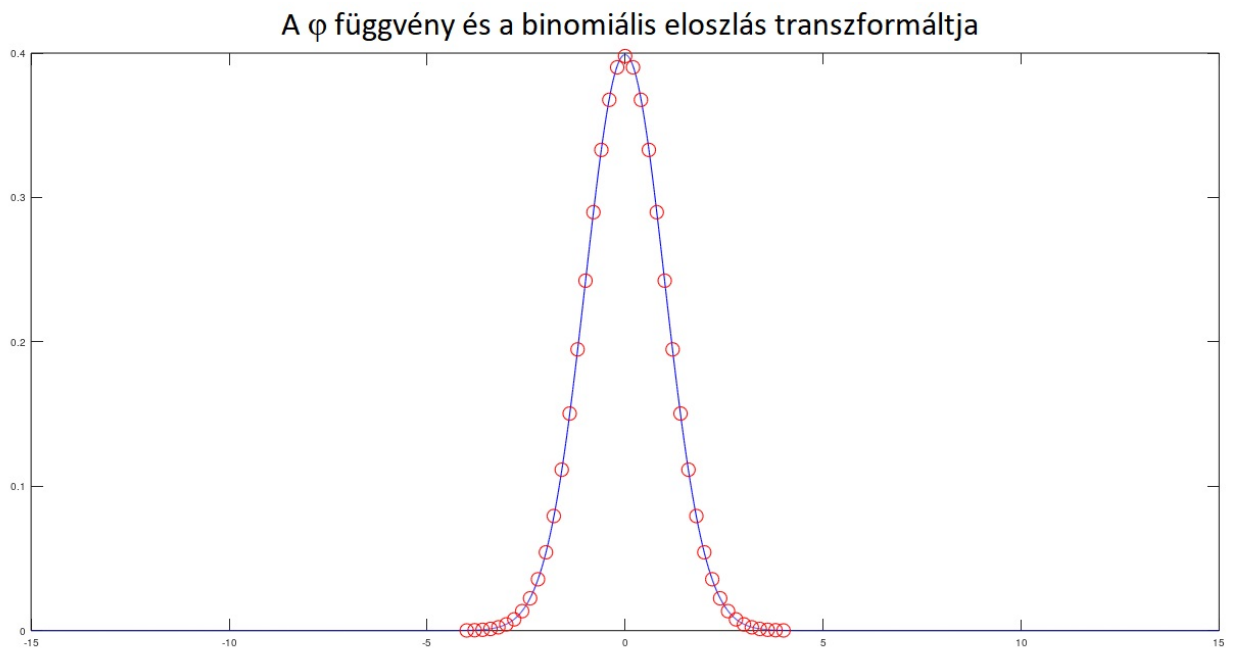
sűrűségfüggvényének grájára, és így a téglalapok területe valójában a φ integrálját közelíti. Egészen pontosan az igaz, hogy ha n tart a végtelenbe, akkor egy fix $t \in \mathbb{R}$ -re a $\mathbb{P}\left(\frac{X - \mathbb{E}(X)}{\mathbb{D}(X)} \leq t\right)$ valószínűség tart a

$$\Phi(t) := \int_{-\infty}^t \varphi(s) ds$$

integrálhoz, vagyis a sztenderd normális eloszlás eloszlásfüggvényének t helyen felvett értékéhez. Azaz a (3.13) valószínűséget nagy n esetén jól tudjuk közelíteni a sztenderd normális eloszlással (a $t = 2(k - \frac{n}{2})n^{-\frac{1}{2}}$ értéket véve).



3.10. ábra. A binomiális eloszlás súlyfüggvényértékeinek transzformált diagramja



3.11. ábra. A binomiális eloszlás súlyfüggvényértékei transzformálva jól illeszkednek a sztenderd normális eloszlás φ sűrűségfüggvényének grájfjára

Habár a fentiekben mi egy speciális binomiális eloszlású változóval dolgoztunk, valójában ugyanilyen eredmények adódnak egy tetszőleges $X \sim B(n; p)$ változó esetén, ahol $p \in (0; 1)$. Természetesen ekkor $\mathbb{E}(X) = np$, tehát ezzel kell eltolni X -et, továbbá a $\mathbb{D}(X) = \sqrt{np(1-p)}$ értékkel kell osztanunk a sztenderdizálásnál. Továbbá, bár a (3.13) valószínűségben csak egy felső korlátot szabtuk a változónak, amely az integrálás felső határát adja, de ugyanígy vizsgálhatunk egy alsó korlátot is, ez természetesen a közelítésnél az integrálás alsó határa lesz. Minden adott ahhoz, hogy kimondjuk a következő tételt.

3.5.2. Tétel (de Moivre–Laplace). *Legyen $X_n \sim B(n; p)$ egy binomiális eloszlású valószínűségi változó, ahol $p \in (0; 1)$ rögzített szám. Ekkor*

$$\begin{aligned} \lim_{n \rightarrow \infty} \mathbb{P} \left(a < \frac{X_n - \mathbb{E}(X_n)}{\mathbb{D}(X_n)} < b \right) &= \lim_{n \rightarrow \infty} \mathbb{P} \left(a < \frac{X_n - np}{\sqrt{np(1-p)}} < b \right) \\ &= \int_a^b \varphi(s) ds = \Phi(b) - \Phi(a), \end{aligned}$$

ahol $a, b \in \mathbb{R}$. Továbbá a $b = \infty$ és $a = -\infty$ értékek is megengedhetők, ekkor a fenti egyenlőség jobb oldalán a $\lim_{b \rightarrow \infty} \Phi(b) = 1$ ill. a $\lim_{a \rightarrow -\infty} \Phi(a) = 0$ értékek szerepelnek a megfelelő tag helyett.

A tétel precíz bizonyítása természetesen jóval túlmutat e kurzus keretein, viszont a fenti állítás csupán speciális esete egy lényegesen általánosabb tételnek, a centrális határeloszlás tételének, amely a következő szakasz témája.

Gyakorlatok, feladatok

3.5.1. Gyakorlat. Egy bizonyos csavar esetében a selejtes darabok aránya 5%. Egy üzlet 1000 darabot vásárolt a kérdéses csavarból. Mennyi a valószínűsége annak, hogy több mint 60 selejtes csavar lesz köztük?

3.5.2. Gyakorlat. Egy termékbemutató szervezésekor 1000 meghívót küldenek szét. A tapasztalat szerint a meghívottak egymástól függetlenül 0,1 valószínűséggel fogadják el a meghívást és jelennek meg a rendezvényen. Mekkora teremben kell a rendezvényt megtartani, ha azt akarják, hogy a megjelentek mind le tudjanak ülni legalább 90%-os valószínűséggel?

3.5.3. A centrális határeloszlás tétele

A következőkben szintén változók összegének az eloszlásával foglalkozunk, azonban ezúttal már nem szorítkozunk egy speciális eloszlásra. A fenti de Moivre–Laplace-tétel általánosítását fogjuk megadni, ehhez pedig emlékeztetünk rá, hogyan lehet a binomiális eloszlás összeg formában előállítani (lásd az előző szakasz elejét is). Legyenek tehát az A_i események együttesen függetlenek, ahol $1 \leq i \leq n$, továbbá tegyük fel, hogy $\mathbb{P}(A_i) = p$ teljesül minden i -re. Tekintsük a hozzájuk tartozó indikátorváltozók $S_n = \sum_{i=1}^n \mathbb{1}_{A_i}$ összegét, ennek eloszlása $S_n \sim B(n; p)$, és a de Moivre–Laplace-tétel értelmében egy fix p paraméter esetén

$$\mathbb{P} \left(\frac{S_n - \mathbb{E}(S_n)}{\mathbb{D}(S_n)} < t \right) \rightarrow \Phi(t) \quad \text{ha} \quad n \rightarrow \infty.$$

Ahogy a nagy számok törvényének tárgyalásánál, az általános(abb) esetben itt is egy X_1, \dots, X_n, \dots sorozatot tekintünk, ahol az X_i -k együttesen független, azonos eloszlású valószínűségi változók. Az előző szakaszhoz hasonlóan most is a $\sum_{i=1}^n X_i$ összeg sztenderdizáltját szeretnénk kezelni, így azt biztosan fel kell tennünk, hogy a $\mathbb{D}(X_i) = \sigma$ szórás véges és pozitív, mert sztenderdizálásnál az összeg szórásával osztanunk kell. Ekkor persze $\mathbb{E}(X_i) = \mu < \infty$ is teljesül, és az összeg várható értéke

$$\mathbb{E}\left(\sum_{i=1}^n X_i\right) = \sum_{i=1}^n \mathbb{E}(X_i) = n\mu.$$

Továbbá a változók függetlensége miatt

$$\mathbb{D}^2\left(\sum_{i=1}^n X_i\right) = \sum_{i=1}^n \mathbb{D}^2(X_i) = n\sigma^2,$$

tehát az összeg sztenderdizáltja

$$\frac{\sum_{i=1}^n X_i - \mathbb{E}(\sum_{i=1}^n X_i)}{\mathbb{D}(\sum_{i=1}^n X_i)} = \frac{\sum_{i=1}^n X_i - n\mu}{\sqrt{n}\sigma}.$$

3.5.3. Tétel (A centrális határeloszlás tétele). *Legyen X_1, \dots, X_n, \dots együttesen független, azonos eloszlású valószínűségi változók egy sorozata, melyekre $0 < \mathbb{D}(X_i) = \sigma < \infty$ minden i -re. Ekkor $\mathbb{E}(X_i) = \mu < \infty$, és minden $t \in \mathbb{R}$ esetén*

$$\mathbb{P}\left(\frac{\sum_{i=1}^n X_i - n\mu}{\sqrt{n}\sigma} < t\right) \rightarrow \Phi(t) \quad \text{ha} \quad n \rightarrow \infty,$$

ahol Φ a sztenderd normális eloszlás eloszlásfüggvénye.

A de Moivre–Laplace-tétel jól láthatóan adódik ebből az $X_i = \mathbb{1}_{A_i}$ választással, de a fenti eredmény lényegesen erősebb, hiszen az X_i változók eloszlása tetszőleges lehet (véges szórással). Éppen ebből adódik a tétel széleskörű alkalmazhatósága, bár a változók azonos eloszlásának megkövetelése valójában lényeges megszorítás. Azonban szerencsére a fenti tételnek léteznek jóval általánosabb változatai, ahol ettől a megszorítástól megszabadulhatunk (a részleteket lásd pl. a [7] könyvben). Ez (ha nem is teljesen precízen, de szemléletesen) megerősíti a normális eloszlás alkalmazásairól elmondottakat, miszerint nagyszámú, független, de egymagában elhanyagolható méretű hatás összességéként létrejövő valószínűségi változó eloszlását normális eloszlással modellezhetjük.

3.5.1. Példa. Egy projektorhoz van összesen 100 darab égőnk, melyek élettartama egymástól független exponenciális eloszlású 50 óra várható értékkel. Tegyük fel, hogy az égőket egymás után használjuk, azonnal kicserélve azt, amelyik kiégett. Mi annak valószínűsége, hogy legalább 5250 órán keresztül használható az égőkkel a projektor?

Jelölje X_i az i -edik égő élettartamát, ekkor az X_1, \dots, X_{100} változók egymástól független, azonos eloszlású valószínűségi változók, ahol az eloszlás paramétere $\lambda = 1/\mathbb{E}(X_i) = \frac{1}{50}$. Az égők teljes élettartama a változók összege, így tehát a

$$\mathbb{P}\left(\sum_{i=1}^{100} X_i > 5250\right)$$

valószínűséget keressük. A centrális határeloszlás tételének alkalmazásához átalakítjuk a fenti eseményt úgy, hogy egyrészt abban az összeg sztenderdizáltja szerepeljen, másrészt (a tétel fenti alakjához hasonlóan) az egyenlőtlenség a másik irányba álljon. Ehhez szükség van még a változók szórására: $\mathbb{D}(X_i) = 1/\lambda = 50$. Tehát

$$\begin{aligned} \mathbb{P}\left(\sum_{i=1}^{100} X_i > 5250\right) &= 1 - \mathbb{P}\left(\sum_{i=1}^{100} X_i < 5250\right) \\ &= 1 - \mathbb{P}\left(\frac{\sum_{i=1}^{100} X_i - 100 \cdot 50}{\sqrt{100 \cdot 50}} < \frac{5250 - 100 \cdot 50}{\sqrt{100 \cdot 50}}\right) \\ &= 1 - \mathbb{P}\left(\frac{\sum_{i=1}^{100} X_i - 100 \cdot 50}{\sqrt{100 \cdot 50}} < \frac{1}{2}\right) \approx 1 - \Phi\left(\frac{1}{2}\right) \approx 0,3085, \end{aligned}$$

mivel a sztenderdizált közelítőleg sztenderd normális eloszlású. A komplementerre való áttérésnél hallgatólagosan felhasználtuk azt a(z eddig nem említett) tényt, hogy folytonos valószínűségi változók összege is folytonos, így a komplementer eseménynél is szigorú egyenlőtlenség írható.

Gyakorlatok, feladatok

3.5.3. Gyakorlat. Az L méretű tojások átlagos tömege 68 g, 4 g szórással. Ha egy tálcán 25 tojás van, akkor mennyi a valószínűsége, hogy az összsúlyuk legalább 1,65 kg?

3.5.4. Gyakorlat. Adottak az $X_1, X_2, \dots, X_{12} \sim U(0; 1)$ együttesen független valószínűségi változók. Ezek segítségével állítsunk elő közelítőleg $N(5; 4)$ normális eloszlású véletlen számot.

3.5.5. Feladat. Béla követőket gyűjt egy online platformon. Kezdetben 1000 követője van. Tegyük fel, hogy ez minden héten vagy 1,1-szeresére nő, vagy nem változik, vagy 0,9-szeresére csökken. Mindhárom kimenetel valószínűsége egyenként $\frac{1}{3}$. Az egyes hetek változásai egymástól függetlenek.

- Határozzuk meg a követői számának várható értékét 104 hét (kb. két év) elteltével.
- Közelítőleg mennyi az esélye, hogy Bélának kevesebb követője lesz 104 hét múlva, mint kezdetben?

4. fejezet

Statisztikai módszerek

A statisztika (mint tudomány) célja, hogy kísérleti megfigyelések eredményei alapján, a valószínűségi számítás eszközeit felhasználva minél többet feltárjon a háttérben zajló véletlen jelenség természetéről, és ezek alapján segítse például a kockázatelemzést vagy jövőbeni döntések meghozatalát. Ezen kurzus keretei pusztán egy nagyon rövid betekintést engednek meg a tudomány alapvető módszereibe, hiszen mind az idő rövide, mind az erősen korlátozott matematikai eszköztár határt szab lehetőségeinknek. Mindazonáltal ez is elegendő arra, hogy olyan alapvető statisztikákat, mint az átlag vagy a tapasztalati szórásnégyzet, bizonyos részletességgel tárgyaljunk, segítségükkel pedig becsléseket adhassunk eloszlások paramétereire. Végül még egy nagyon rövid ízelítőt adunk a hipotézisvizsgálatok elméletéből. Mindezekről a témákról (és még sok másokról) részletesen olvashatunk például az [1] könyvben.

4.1. Statisztikai alapfogalmak

Egy statisztikus kiindulópontja tipikusan valamilyen megfigyelés. A megfigyelés tárgyát képező egyedek összességét *statisztikai sokaságnak* nevezzük. Statisztikai sokaság lehet pl. egy egyetem hallgatóinak halmaza vagy egy terület időjárása egyes napokon. Mivel a sokaságot alkotó egyedeket tipikusan vagy véletlenszerűen választjuk, vagy pedig előfordulásuk számos tényező által befolyásolt és így véletlen jelenségnek tekinthető, a magát az egész sokaságot a statisztikában egy valószínűségi mező, tehát az egyedek által alkotott Ω eseménytér, az azon kijelölt események, illetve egy valószínűségi mérték modellezi.

A statisztikai sokaság egyedeire vonatkozó tulajdonságokat, jellemzőket *statisztikai ismérveknek* nevezzük. Ezek az ismérvek különféle kategóriákba sorolhatók attól függően, hogy milyen típusú jellemzőit írják le az egyedeknek. Szokás például időbeli, területi, minőségi ill. mennyiségi ismérvekről beszélni. A különféle kategóriák ellenére ezen tulajdonságok mindegyikét a sokaságot modellező valószínűségi mezőn értelmezett valószínűségi változóként kezelhetjük. Mennyiségi ismérveket leíró változók értelmezése általában kézenfekvő, beszélhetünk például az egyedek magasságáról egy populációban vagy egy terület napi középhőmérsékletéről. De ez a modell a minőségi jellemzők leírásánál sem jelent korlátot. Ilyen esetekben az egyedeket különféle kategóriákba soroljuk, melyek persze jelölhetők számokkal is, úgynevezett kategorikus valószínűségi változókat adva ilyen módon. Például egy populáció egyedeinek szemszínét a kék, zöld, barna, stb. kategóriákba sorolhatjuk. Ha ezeket az $1, 2, 3, \dots$ számokkal azonosítjuk, akkor ilyen módon egy $X : \Omega \rightarrow \mathbb{R}$ valószínűségi változót kapunk. A fejezet további részében szinte kizárólag mennyiségi ismérveket modellező változókkal foglalkozunk.

A matematikai alapprobléma tehát a következő: adott egy valószínűségi mező és egy azon értelmezett valószínűségi változó, amelynek eloszlása (az ún. *háttereeloszlás*) nem ismert, a feladatunk pedig ennek (vagy az eloszlás jellemzőinek) meghatározása. Sokszor előfordul, hogy különféle megfontolások alapján a háttereeloszlás egy adott eloszláscsalád tagjának tekinthető. Például egy várakozási idő gyakran tekinthető exponenciális eloszlásúnak, amennyiben arra az örökifjú tulajdonság vonatkoztatható. Ezen feltétel mellett a feladat az eloszlás paraméterének meghatározására redukálódik. Hasonlóképp, számos véletlen mennyiség tekinthető normális eloszlású valószínűségi változónak, ekkor az eloszlás meghatározásához elegendő annak paramétereit, azaz a várható értéket és a szórásnégyzetet megadni. Első lépésként tehát - az eseményteret és az eseményeket fixálva - kijelöljük azoknak a lehetséges valószínűségi mértékeknek az összességét, amelyek bármelyikével együtt az eseménytér valószínűségi mezőt alkot, és amelyek közt ott van az is, ami a háttereeloszlást adja.

4.1.1. Definíció. Egy eseménytér, az azon kijelölt események, továbbá valószínűségi mértékek egy családja együttesen *statisztikai mezőt* alkot, ha a mértékcsalád bármely tagja az eseménytérrel és az eseményekkel együtt egy valószínűségi mezőt ad meg. Ha a mértékcsalád tagjai valós paraméterekkel paraméterezhetők, akkor *paraméteres statisztikai mezőről* beszélünk. A paramétereket tipikusan egy valós vektorként adjuk meg, k különböző paraméter esetében tehát ezek \mathbb{R}^k elemei. A mértékcsalád tagjait leíró lehetséges paraméterek $\Theta \subset \mathbb{R}^k$ halmazát *paramétertérnek* nevezzük.

Mivel egy teljes sokaságot tipikusan nem tudunk áttekinteni, ezért a fenti feladat megoldásához a gyakorlatban egy ún. *statisztikai mintát* használunk, amely a sokaság egyedeinek egy részhalmazából és az ezen egyedekhez tartozó jellemzőkből áll. Ezt a mintát egymástól független mérések vagy az egyedek egymástól független választásai adják. Éppen ezért, ha X jelöli a *hátterváltozót*, aminek eloszlása a háttereeloszlás, akkor egy n elemű minta elemeit az X_1, \dots, X_n együttesen független valószínűségi változókkal modellezzük, amelyek eloszlása azonos az X eloszlásával. Fontos kiemelni, hogy egy konkrét mintánál számadatokkal dolgozunk, nem pedig valószínűségi változókkal. Egy ilyen konkrét mérés eredményéből származó adatsor a minta *realizációja*. Az elméleti modellben azonban a mintavételt mégis változókkal írjuk le, és így egy realizáció ezen változók egy kiértékelését jelenti. Ez garantálja, hogy az elméleti modelltől kapott eredmények minden egyes realizációra érvényesek legyenek. Megjegyezzük, hogy a minta egy realizációját néha (az egyszerűség kedvéért pongyolán) szintén mintának nevezzük.

Természetesen különböző realizációk különböző eredményeket szolgáltatnak. Ha mondjuk különböző dobássorozatok eredményéből akarunk egy dobókockával való dobás eredményének eloszlására következtetni, akkor más-más relatív gyakoriságok jönnek ki az egyes kimenetelekre az egyes dobássorozatok esetén. Éppen ezért nem mondhatjuk, hogy a háttereeloszlást egyik vagy másik kísérletsorozat eredménye adja. Véletlen eloszlásokra vagy véletlen mennyiségekre tehát általában nem tudunk a minták alapján pontosan következtetni. Azaz a statisztikai következtetések (bár logikai következtetések, de) természetükből adódóan becslések. Persze a becslések hibáját is kezelni kell, ez a hiba pedig kontrollálható. Azonban a véletlen tulajdonságaiból adódóan a teljes bizonyosságot itt is fel kell áldoznunk. Amit viszont előírhatunk, az az, hogy a kapott eredményünk nagy valószínűséggel egy bizonyos hibahatáron belül közelítse a keresett értéket. Mindez a valószínűségszámítás eszközeivel teljesen precízen megfogalmazható, és azt is látni fogjuk, hogy hogyan érhetünk el pontosabb becslést, vagy hogy milyen mértékben kell feláldoznunk a pontosságot a nagyobb bizonyosságért cserébe.

Habár a mintavételezés technikájának részleteivel itt nem foglalkozunk, de röviden megemlítenénk néhány alapvető követelményt ezzel kapcsolatban. A minta természetesen általában

tipikusan kis elemszámú a sokaság elemszámához mérten. Éppen azért dolgozunk egy mintával, mert nem feltétlenül van lehetőségünk vagy kapacitásunk a teljes sokaság áttekintésére. Azonban arra is ügyelni kell, hogy ez az elemszám elegendően nagy legyen ahhoz, hogy a kapott eredményt kellően megalapozottnak tekinthessük. Egy 10 elemű, bizonyos esetekben kicsinek tekinthető minta például gyakran viszonylag nagy valószínűséggel viselkedhet az átlagostól eltérő módon. Könnyen előfordulhat, hogy egy dobókockával dobva az első tíz dobás úgy alakul, hogy nem fordul elő 4-es, 5-ös vagy 6-os a dobott számok közt. Ha viszont ez száz egymást követő dobásnál történik, az már azt a gyanút kelti, hogy az adott eredmények nem egy megfelelően vett véletlen mintából származnak. Persze az is előfordulhat, hogy egyszerűen nincs lehetőségünk elég nagy mintával dolgozni. Ezért is fontos, hogy az eredményeink olyan információt is szolgáltatassanak, ami a minta elemszámának függvényében ad becslést a pontosságra.

Megemlítjük még, hogy a mintának *reprezentatív*nak kell lennie abban az értelemben, hogy a vizsgálat tárgyát képező sokasághoz szerkezetében, főbb jellemzőiben hasonlatosnak kell lennie. Itt nem foglalkozunk azzal, hogy ez hogyan érhető el.

A mintából kapott adatsort sokszor érdemes úgy átalakítani, hogy az praktikus legyen a feldolgozásnál. Az alább definiált tapasztalati eloszlásfüggvény kiszámolásához például érdemes a minta elemeit nagyság szerint növekvő sorrendbe rendezni. Így kapjuk az ún. *rendezett mintát*, melynek elemeit egy kiindulásul szolgáló X_1, \dots, X_n minta esetén X_1^*, \dots, X_n^* jelöli, ahol tehát az X_i^* változók értékeinek halmaza megegyezik az X_i -k értékeinek halmazával (egy fix realizációnál), továbbá $X_1^* \leq X_2^* \leq \dots \leq X_n^*$ teljesül.

4.1.2. Definíció. Legyen X_1, \dots, X_n egy független, azonos eloszlású minta, ekkor a mintához tartozó F_n^* *tapasztalati (vagy empirikus) eloszlásfüggvény* értéke egy $t \in \mathbb{R}$ helyen

$$F_n^*(t) = \frac{\sum_{i=1}^n \mathbb{1}_{\{X_i < t\}}}{n}.$$

A tapasztalati eloszlásfüggvény t helyen vett értékének meghatározásához tehát azt kell megszámlálni, hogy a minta hány elemének értéke esik t alá, ennek a számnak és a minta elemszámának az aránya adja a függvény értékét. Vegyük észre, hogy ha pontosan i darab érték esik t alá, az azt jelenti, hogy $X_i^* < t$, de $X_{i+1}^* \geq t$. Természetesen külön kezelendő az az eset, amikor már a legkisebb érték is legalább t , illetve amikor a legnagyobb érték is kisebb t -nél:

$$F_n^*(t) = \begin{cases} 0, & \text{ha } t \leq X_1^*, \\ \frac{i}{n}, & \text{ha } X_i^* < t \leq X_{i+1}^* \quad (1 \leq i \leq n-1), \\ 1, & \text{ha } t > X_n^*. \end{cases}$$

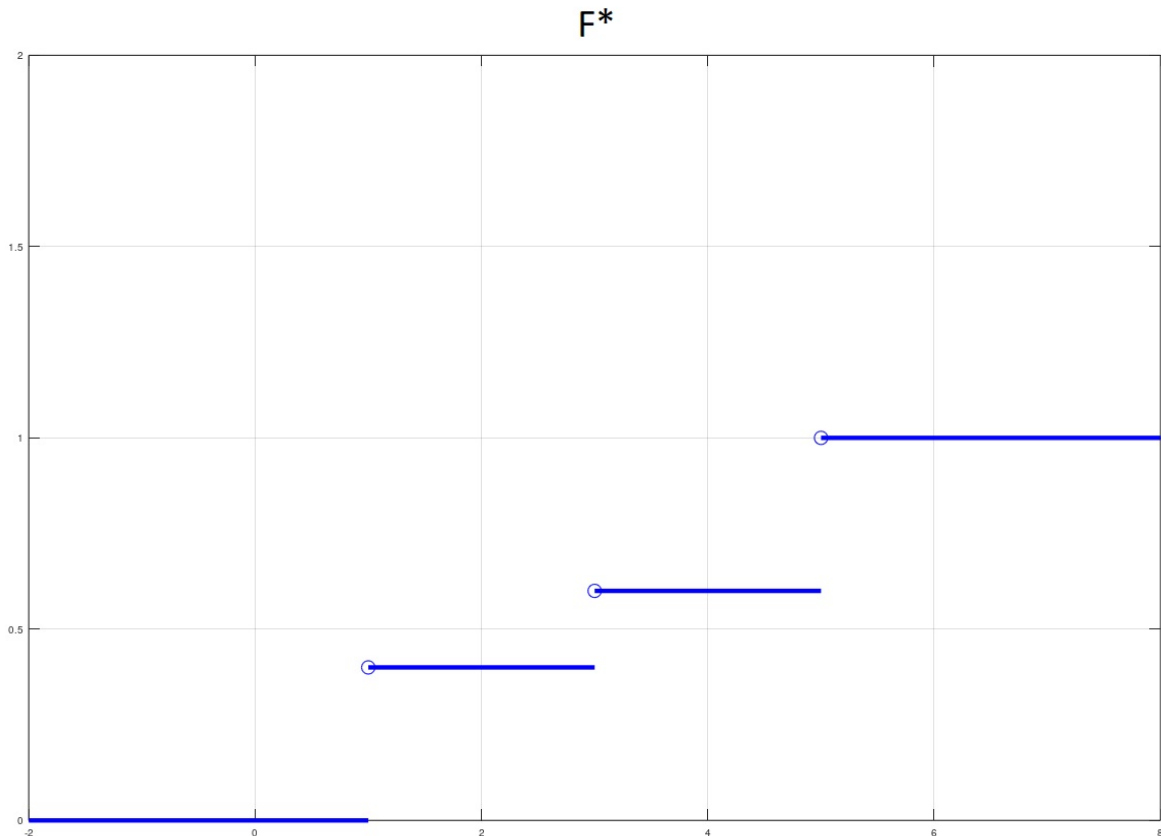
Fontos megjegyezni, hogy bár a tapasztalati eloszlásfüggvény egy valószínűségi változó, de a minta egy konkrét realizációját behelyettesítve már egy $\mathbb{R} \rightarrow \mathbb{R}$ függvényt kapunk.

4.1.1. Példa. Ötször dobunk egy dobókockával, jelölje a dobások eredményét X_1, \dots, X_5 . Ezek tehát független, azonos eloszlású valószínűségi változók, amelyek egy 5 elemű mintát alkotnak. Tegyük fel, hogy egy konkrét kísérlet során a 3, 1, 5, 5, 1 eredmények adódnak. Ez tehát a minta egy realizációja. Erre a realizációra a rendezett minta értékei:

$$X_1^* = 1, \quad X_2^* = 1, \quad X_3^* = 3, \quad X_4^* = 5, \quad X_5^* = 5.$$

Az eloszlásfüggvény tehát erre a realizációra a következő:

$$F_5^*(t) = \begin{cases} 0, & \text{ha } t \leq 1, \\ \frac{2}{5}, & \text{ha } 1 < t \leq 3, \\ \frac{3}{5}, & \text{ha } 3 < t \leq 5, \\ 1, & \text{ha } t > 5. \end{cases}$$



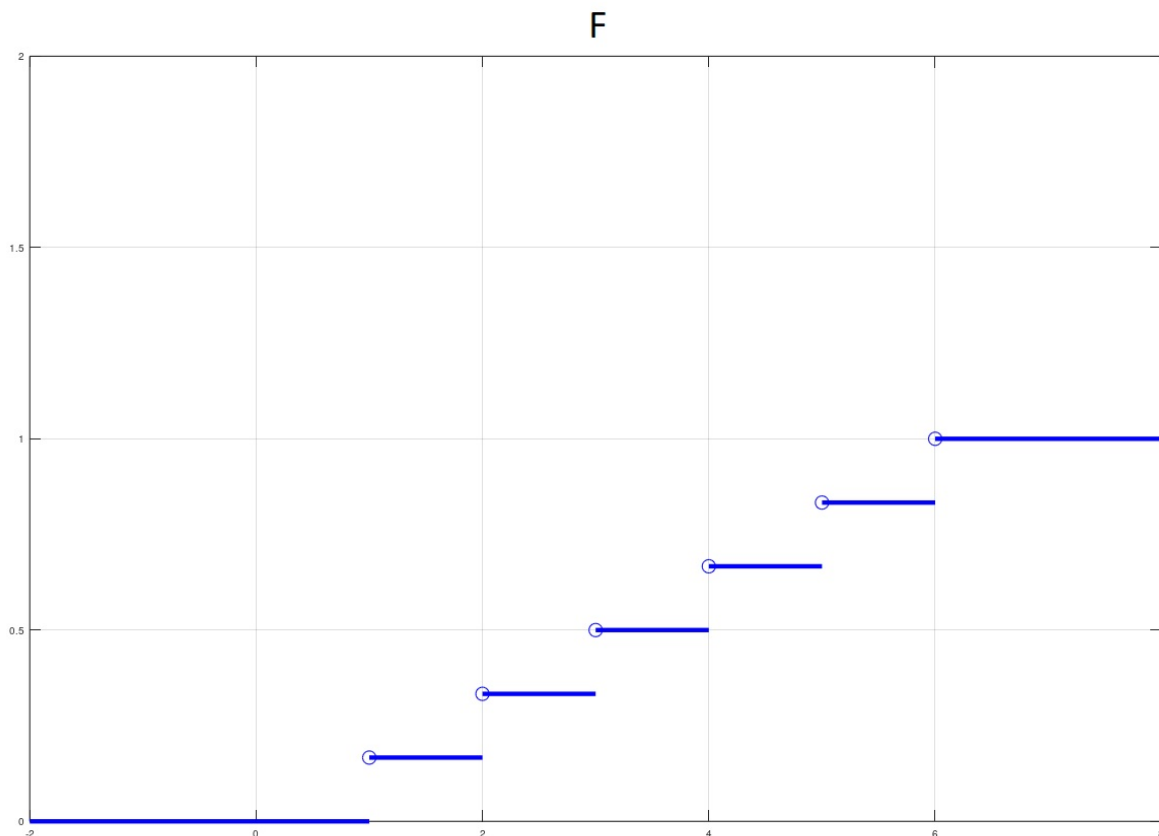
4.1. ábra. A kockadobás-kísérlethez tartozó tapasztalati eloszlásfüggvény

Ha a kocka szabályos, akkor magának a háttéreloszlásnak az eloszlásfüggvénye

$$F(t) = \begin{cases} 0, & \text{ha } t \leq 1, \\ \frac{i}{6}, & \text{ha } i < t \leq i + 1 \quad (1 \leq i \leq 5), \\ 1, & \text{ha } t > 6. \end{cases}$$

Habár ilyen kis mintaelemszám esetén a tapasztalati és a tényleges eloszlásfüggvény közt nagy különbség lehet, ez a különbség a mintaelemszám növelésével fokozatosan csökken, sőt valójában a tapasztalati eloszlásfüggvény értéke az eloszlásfüggvény értékéhez tart.

4.1.1. Állítás. Minden $t \in \mathbb{R}$ esetén $\mathbb{P}(\lim_{n \rightarrow \infty} F_n^*(t) = F(t)) = 1$, ahol F_n^* a tapasztalati eloszlásfüggvény, F pedig a háttéreloszlás eloszlásfüggvénye.



4.2. ábra. A kockadobás valószínűségi modelljéhez tartozó eloszlásfüggvény

Bizonyítás. Az $F_n^*(t)$ valószínűségi változó nem más, mint az $\mathbb{1}_{\{X_i < t\}}$ indikátorok átlaga. Mivel az X_i -k azonos eloszlásúak, és eloszlásuk megegyezik az X háttérváltozó eloszlásával, így

$$\begin{aligned} \mathbb{P}(\mathbb{1}_{\{X_i < t\}} = 1) &= \mathbb{P}(X_i < t) = \mathbb{P}(X < t), \\ \mathbb{P}(\mathbb{1}_{\{X_i < t\}} = 0) &= \mathbb{P}(\overline{\{X_i < t\}}) = 1 - \mathbb{P}(X_i < t) = 1 - \mathbb{P}(X < t), \end{aligned}$$

azaz a fenti indikátorváltozók is azonos eloszlásúak. Továbbá az együttes függetlenség definíciója (lásd a 3.2.3. definíciót) és a 2.2.1. állítás alapján az X_i -k együttes függetlensége miatt ezek az indikátorok is együttesen függetlenek, és mivel a szórásuk is véges, így alkalmazható rájuk a nagy számok erős törvénye (azaz a 3.5.1. tétel): az indikátorok átlaga egy valószínűséggel tart a (közös) várható értékükhöz. Mivel $\mathbb{E}(\mathbb{1}_{\{X_i < t\}}) = \mathbb{P}(X_i < t) = \mathbb{P}(X < t) = F(t)$, így az állítást beláttuk. \square

A fenti állításnál azonban több is igaz. Azt már látjuk, hogy minden egyes t -re a tapasztalati eloszlásfüggvény és a háttéreloszlás eloszlásfüggvénye közel lesz egymáshoz, amennyiben a minta elemszáma nagy. De az, hogy milyen nagynak kell választani az elemszámot, elvben függhetne a t konkrét értékétől. Ez azonban nem így van, a fenti konvergencia egyenletes, ez a *statisztika alaptételének* állítása.

4.1.2. Tétel (Glivenko-Cantelli, a statisztika alaptétele). *Az F_n^* tapasztalati eloszlásfüggvény 1 valószínűséggel egyenletesen tart a háttéreloszlás F eloszlásfüggvényéhez. Azaz: ha a minta elemszáma elég nagy, akkor F_n^* értéke 1 valószínűséggel tetszőlegesen közel kerül egyenletesen*

(azaz egyszerre minden $t \in \mathbb{R}$ -re) az F -hez. Formalizálva:

$$\sup_{t \in \mathbb{R}} |F_n^*(t) - F(t)| \rightarrow 0, \quad (n \rightarrow \infty)$$

teljesül 1 valószínűséggel.

Az eddigiek tehát úgy összegezhetők, hogy ha elég nagy a mintánk, akkor a háttéreloszlás tetszőleges pontossággal megközelíthető.

4.2. Pontbecslések

A továbbiakban paraméteres statisztikai mezőket tekintünk. Ha adott egy X_1, \dots, X_n minta, akkor ennek segítségével szeretnénk a háttéreloszlás paramétereit (vagy esetleg annak függvényeit) becsülni. Ehhez n elemű minta esetén egy $T : \mathbb{R}^n \rightarrow \mathbb{R}$ függvényt fogunk használni (amely persze mindig olyan lesz, hogy $T(X_1, \dots, X_n)$ is egy valószínűségi változó). Ekkor a mintaelemek $T(X_1, \dots, X_n)$ függvényét (is) *statisztikának* nevezzük. Megjegyezzük, hogy az alábbiakban tipikusan az elemszámot bármilyen nagynak választhatjuk, így a T függvényt minden egyes n -re definiálni kell. Vegyük észre továbbá, hogy egy konkrét x_1, \dots, x_n realizációra $T(x_1, \dots, x_n)$ is egy számértéket ad.

Ebben a szakaszban megismerkedünk néhányval a leggyakrabban használt statisztikák közül, nevezetesen a mintaátlaggal, a tapasztalati szórásnégyzettel és a korrigált tapasztalati szórásnégyzettel. Ahogy az ezen a ponton már bizonyára sejthető, ezeket a háttéreloszlás várható értékének ill. a szórásnégyzetének becslésére használhatjuk.

4.2.1. Definíció. Legyen X_1, \dots, X_n független, azonos eloszlású n elemű minta, ekkor az

$$\bar{X} = \frac{X_1 + \dots + X_n}{n}$$

statisztikát *mintaátlagnak* nevezzük.

Egy konkrét realizáció esetén a mintaátlagot \bar{x} fogja jelölni. Ha hangsúlyozni szeretnénk az elemszámot, akkor \bar{X}_n -t ill. \bar{x}_n -t írunk. Az előző szakasz 4.1.1. példájában a kockadobások átlaga

$$\frac{3 + 1 + 5 + 5 + 1}{5} = 3.$$

Mivel egy valószínűségi változó várható értéke a változó átlagos értékét hivatott jellemezni, kézenfekvőnek tűnik a háttéreloszlás várható értékét a mintaátlaggal becsülni. Felmerül a kérdés, hogy mennyire lesz jó ez a becslés. Egy becslés jóságát természetesen különböző kritériumokkal mérhetjük, amelyből itt most egyet tárgyalunk részletesen.

4.2.2. Definíció. A $T(X_1, \dots, X_n)$ statisztika *torzítatlan becslés* a θ paraméterre, ha

$$\mathbb{E}(T(X_1, \dots, X_n)) = \theta$$

teljesül.

4.2.1. Állítás. *A mintaátlag torzítatlan becslés a háttéreloszlás várható értékére (amennyiben az véges).*

Bizonyítás. A várható érték linearitása miatt

$$\mathbb{E}(\bar{X}_n) = \frac{1}{n} \sum_{i=1}^n \mathbb{E}(X_i) = \frac{1}{n} \cdot n \cdot \mathbb{E}(X) = \mathbb{E}(X),$$

ahol X a háttérváltozó. □

Tehát a mintaátlag átlagosan jól viselkedik, de felmerül a kérdés, hogy mennyire lesznek közel a várható értékhez a tényleges értékek. Ha az X háttérváltozó szórásnégyzete véges, akkor a minta függetlensége miatt

$$(4.1) \quad \mathbb{D}^2(\bar{X}) = \mathbb{D}^2\left(\frac{X_1 + \dots + X_n}{n}\right) = \frac{1}{n^2} \sum_{i=1}^n \mathbb{D}^2(X_i) = \frac{1}{n} \mathbb{D}^2(X),$$

vagyis ekkor a mintaátlag szórása 0-hoz tart, ha n tart végtelenhez, ami épp azt jelenti, hogy a tényleges értékek a várható érték körül koncentrálódnak, ha az elemszám nagy.

Térjünk most rá a szórásnégyzet becslésére. A szórásnégyzet lényegében a várható értéktől való átlagos négyzetes eltérés, így most logikusnak tűnhet ezt az mintaátlagtól való átlagos négyzetes eltéréssel becsülni.

4.2.3. Definíció. Legyen X_1, \dots, X_n független, azonos eloszlású n elemű minta, ekkor az

$$S^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2$$

statisztikát a minta *tapasztalati szórásnégyzetének* nevezzük. Ennek S gyöke a *tapasztalati szórás*.

Egy konkrét realizáció esetén a tapasztalati szórásnégyzetet s^2 fogja jelölni. Ha hangsúlyozni szeretnénk az elemszámot, akkor S_n^2 -et ill. s_n^2 -et írunk. A gyakorlatban sokszor hasznosabb a fenti formulát átalakítani:

$$\begin{aligned} S^2 &= \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2 = \frac{1}{n} \sum_{i=1}^n (X_i^2 - 2X_i \cdot \bar{X} + \bar{X}^2) \\ &= \frac{1}{n} \sum_{i=1}^n X_i^2 - 2\bar{X} \cdot \frac{1}{n} \sum_{i=1}^n X_i + \frac{1}{n} \cdot n \cdot \bar{X}^2 \\ &= \bar{X}^2 - 2\bar{X}^2 + \bar{X}^2 = \bar{X}^2 - \bar{X}^2. \end{aligned}$$

A tapasztalati szórásnégyzetet tehát úgy kapjuk, ha a mintaelemek négyzetének átlagából kivonjuk a mintaátlag négyzetét. A 4.1.1. példában a mintaátlag 3, a mintaelemek négyzete pedig rendre 9, 1, 25, 25, 1, tehát

$$s^2 = \frac{9 + 1 + 25 + 25 + 1}{5} - 3^2 = \frac{61}{5} - 9 = \frac{16}{5}.$$

A tapasztalati szórásnégyzet azonban nem lesz torzítatlan becslése a szórásnégyzetnek, mert (a fenti állítást és (4.1)-et is felhasználva)

$$\begin{aligned} \mathbb{E}(S_n^2) &= \frac{1}{n} \sum_{i=1}^n \mathbb{E}(X_i^2) - \mathbb{E}(\bar{X}^2) = \frac{1}{n} \sum_{i=1}^n (\mathbb{D}^2(X_i) + \mathbb{E}(X_i)^2) - (\mathbb{D}^2(\bar{X}) + \mathbb{E}(\bar{X})^2) \\ &= \frac{1}{n} \cdot n \cdot (\mathbb{D}^2(X) + \mathbb{E}(X)^2) - \left(\frac{1}{n} \mathbb{D}^2(X) + \mathbb{E}(X)^2\right) = \frac{n-1}{n} \mathbb{D}^2(X). \end{aligned}$$

4.2.4. Definíció. A $T(X_1, \dots, X_n)$ statisztika *aszimptotikusan torzítatlan* becslés θ -ra, ha

$$\lim_{n \rightarrow \infty} \mathbb{E}(T(X_1, \dots, X_n)) = \theta.$$

A fenti definíció értelmében tehát a tapasztalati szórásnégyzet aszimptotikusan torzítatlan becslés a szórásnégyzetre, hiszen $\lim_{n \rightarrow \infty} (n-1)/n = 1$. Vegyük azonban észre, hogy némi korrekcióval itt is torzítatlan becslést kaphatunk:

4.2.5. Definíció. Az $S_n^{*2} := \frac{n}{n-1} S_n^2$ statisztikát *korrigált tapasztalati szórásnégyzetnek* nevezzük. Ennek S_n^* gyöke a *korrigált tapasztalati szórás*.

Szokásos módon egy realizációra az s_n^{*2} ill. s_n^* jelöléseket használjuk, esetlegesen az elem-számot a jelölésből elhagyjuk. A fentiek értelmében a korrigált tapasztalati szórásnégyzet torzítatlan becslése a szórásnégyzetnek, hiszen

$$\mathbb{E}(S_n^{*2}) = \frac{n}{n-1} \mathbb{E}(S_n^2) = \frac{n}{n-1} \cdot \frac{n-1}{n} \cdot \mathbb{D}^2(X) = \mathbb{D}^2(X).$$

Végül számoljuk ki a korrigált tapasztalati szórásnégyzetet a 4.1.1 példára. Mivel láttuk, hogy $s^2 = \frac{16}{5}$, a mintaelemszám pedig 5, így $s^{*2} = \frac{5}{4} \cdot \frac{16}{5} = 4$.

4.3. Intervallumbecslések

Az előző szakaszban bemutatott pontbecsléseinktől természetesen nem várhatjuk, hogy a háttéreloszlás paramétereit pontosan megadják. Most azt fogjuk megvizsgálni, hogy mit mondhatunk a tényleges értéktől való eltérésről. Bár ez látszólag csak a szemszögön való árnyalatnyi változtatás, mégis kényelmesebb a becslt pontot tekinteni kiindulópontnak, és azt kérdezni, hogy ettől milyen messze esik a tényleges érték. Vagyis: a becslt érték körül mekkora intervallumot kell venni, hogy abba a tényleges érték nagy valószínűséggel beleessen? Az intervallum két végpontját egy-egy statisztika segítségével fogjuk kijelölni:

4.3.1. Definíció. A $(T_1(X_1, \dots, X_n); T_2(X_1, \dots, X_n))$ statisztikapárral definiált intervallum *legalább $1 - \varepsilon$ szintű konfidenciaintervallum* a θ paraméterre, ha

$$\mathbb{P}(T_1(X_1, \dots, X_n) < \theta < T_2(X_1, \dots, X_n)) \geq 1 - \varepsilon$$

teljesül. Ha itt a fenti egyenlőtlenség helyett szigorú egyenlőség teljesül, akkor pontosan $1 - \varepsilon$ szintű konfidenciaintervallumról beszélünk.

Az $\varepsilon > 0$ értéket az eljárás során előre rögzítjük, és úgy konstruáljuk az intervallum két végpontját, hogy $1 - \varepsilon$ valószínűséggel abba essen a tényleges paraméter. Ez a gyakorlatban akkor ad jól használható eljárást, ha az ε -t elég kicsinek választjuk (pl. gyakori az $\varepsilon = 0,05$ vagy az $\varepsilon = 0,01$ választás, ekkor azt mondjuk, hogy 95%-os ill. 99%-os szintű konfidenciaintervallumot keresünk). De azt is figyelembe kell vennünk, hogy a bizonyosság növeléséért cserébe a pontosságot kell feláldoznunk, nagyon kicsi ε esetén nagyon nagy lehet az intervallum hossza, ami így esetleg túl kevés információt szolgáltat számunkra. Viszont, mint azt hamarosan látni fogjuk, a minta elemszámának növelése is egy eszköz lehet a pontosság növelésére.

Konfidenciaintervallum szerkesztése normális eloszlás várható értékére ismert szórás esetén

Legyen most $X \sim N(\mu; \sigma^2)$ a háttérváltozó, ahol σ^2 ismert. Ebben az esetben a konfidenciaintervallum középpontját a mintaátlagnak fogjuk választani, az intervallumot pedig $(\bar{X} - r_\varepsilon; \bar{X} + r_\varepsilon)$ alakban keressük. Itt az $r_\varepsilon > 0$ sugár természetesen az előzetesen rögzített ε -tól is függ. Az r_ε értékét a következő érvelésből kaphatjuk meg: azt szeretnénk, hogy

$$1 - \varepsilon = \mathbb{P}(\bar{X} - r_\varepsilon < \mu < \bar{X} + r_\varepsilon)$$

teljesüljön. Alakítsuk át az egyenlet jobb oldalát:

$$\begin{aligned} \mathbb{P}(\bar{X} - r_\varepsilon < \mu < \bar{X} + r_\varepsilon) &= \mathbb{P}(-r_\varepsilon < \mu - \bar{X} < r_\varepsilon) = \mathbb{P}(-r_\varepsilon < \bar{X} - \mu < r_\varepsilon) \\ &= \mathbb{P}\left(-r_\varepsilon < \frac{\sum_{i=1}^n X_i - n\mu}{n} < r_\varepsilon\right) \\ &= \mathbb{P}\left(-\frac{r_\varepsilon\sqrt{n}}{\sigma} < \frac{\sum_{i=1}^n X_i - n\mu}{\sqrt{n}\sigma} < \frac{r_\varepsilon\sqrt{n}}{\sigma}\right). \end{aligned}$$

Ezen a ponton felhasználjuk a következő, bizonyítás nélkül közölt tételt:

4.3.1. Tétel. *Legyenek $X_1 \sim N(\mu_1; \sigma_1^2)$ és $X_2 \sim N(\mu_2; \sigma_2^2)$ független, normális eloszlású valószínűségi változók, ekkor $X_1 + X_2 \sim N(\mu_1 + \mu_2; \sigma_1^2 + \sigma_2^2)$.*

A fenti állítás értelmében a mintaelemek $\sum_{i=1}^n X_i$ összege normális eloszlású $n\mu$ várható értékkel és $n\sigma^2$ szórással, tehát az utolsó valószínűségről ennek éppen a sztenderdizáltja szerepel, ami így sztenderd normális eloszlású. Ezért ez a valószínűség felírható a Φ eloszlásfüggvény segítségével, és a következőt kapjuk:

$$1 - \varepsilon = \Phi\left(\frac{r_\varepsilon\sqrt{n}}{\sigma}\right) - \Phi\left(-\frac{r_\varepsilon\sqrt{n}}{\sigma}\right) = 2\Phi\left(\frac{r_\varepsilon\sqrt{n}}{\sigma}\right) - 1,$$

azaz

$$1 - \frac{\varepsilon}{2} = \Phi\left(\frac{r_\varepsilon\sqrt{n}}{\sigma}\right).$$

Mivel a Φ függvény deriváltja a sztenderd normális eloszlás φ sűrűségfüggvénye, ez utóbbi pedig minden valós t -re pozitív, így a Φ függvény szigorúan monoton növe az \mathbb{R} -en, ebből kifolyólag pedig kölcsönösen egyértelmű leképezés \mathbb{R} és a $(0; 1)$ intervallum között. Tehát létezik a $\Phi^{-1} : (0; 1) \rightarrow \mathbb{R}$ inverzfüggvény, amit a fenti egyenletre alkalmazva

$$\frac{r_\varepsilon\sqrt{n}}{\sigma} = \Phi^{-1}\left(1 - \frac{\varepsilon}{2}\right), \quad \text{azaz} \quad \boxed{r_\varepsilon = \Phi^{-1}\left(1 - \frac{\varepsilon}{2}\right) \cdot \frac{\sigma}{\sqrt{n}}}$$

adódik. Ez utóbbi formula adja tehát a konfidenciaintervallum sugarát.

A fentiekből jól látszik, hogy rögzített ε mellett a mintaelemszám növelésével az intervallum sugara csökken. Míg ha a mintaelemszám fix, akkor az ε csökkentésével $1 - \varepsilon/2$ nő. Mivel Φ szigorúan monoton növe, így könnyen meggondolható, hogy Φ^{-1} is az, vagyis ε csökkentése az intervallum sugarának növekedését eredményezi.

Megjegyezzük, hogy a centrális határeloszlás tétele szerint a

$$\frac{\sum_{i=1}^n X_i - n\mu}{\sqrt{n}\sigma}$$

sztenderdizált (véges és pozitív szórás mellett) közelítőleg sztenderd normális eloszlású lesz tetszőleges háttéreloszlás esetén, ha a mintaelemszám elegendően nagy. Így tehát a fenti eredmény ekkor is jól használható.

Konfidenciaintervallum szerkesztése normális eloszlás várható értékére ismeretlen szórás esetén

Az alábbiakban azt az esetet vizsgáljuk, amikor a normális eloszlású háttérváltozó σ szórása nem ismert. Ez a szakasz vázlatosabb és technikai szempontból (ill. a tömörség miatt is) valamivel nehezebb a korábbi anyagrészeknél.

Ha a háttéreloszlás szórása nem ismert, akkor kézenfekvőnek tűnhet a korábbi számolásban a szórást a korrigált tapasztalati szórással becsülni, azaz a

$$\frac{\sum_{i=1}^n X_i - n\mu}{\sqrt{n}\sigma}$$

kifejezés helyett a

$$\frac{\sum_{i=1}^n X_i - n\mu}{\sqrt{n}S_n^*} = \sqrt{n} \cdot \frac{\bar{X} - \mu}{S_n^*}$$

változót használni. Vegyük azonban észre, hogy míg az előbbi törtről tudjuk, hogy sztenderd normális eloszlású változó, az utóbbi változó eloszlását (egyelőre) nem ismerjük, ezért most egy rövid kitérőt teszünk.

Először röviden megismerkedünk két újabb eloszlással. Tekintsük az $X_1, \dots, X_n \sim N(0; 1)$ együttesen független, sztenderd normális eloszlású valószínűségi változókat, ekkor az

$$Y_n := X_1^2 + \dots + X_n^2$$

valószínűségi változó eloszlását n szabadságfokú centrált χ^2 (ejtsd: "khi négyzet") eloszlásnak nevezzük. Jelölése $Y_n \sim \chi^2(n)$. Továbbá, ha az $X \sim N(0; 1)$ és az $Y_n \sim \chi^2(n)$ független valószínűségi változók, akkor a

$$Z_n := \sqrt{n} \cdot \frac{X}{\sqrt{Y_n}}$$

változó eloszlását n szabadságfokú Student-eloszlásnak (vagy t -eloszlásnak) nevezzük. Ennek jele $Z_n \sim t(n)$.

Mivel ezen eloszlásokat illetően mindössze egyetlen egy alkalmazásra szorítkozunk, így a tulajdonságaik részletes leírását itt mellőzzük. Mindazonáltal megemlítjük, hogy a gyakorlatok közt már érintőlegesen találkozhattunk a Student-eloszlással. Ha $Z_n \sim t(n)$, akkor Z_n sűrűségfüggvénye

$$g_n(z) = \frac{1}{\sqrt{\pi n}} \frac{\Gamma(\frac{n+1}{2})}{\Gamma(\frac{n}{2})} \left(1 + \frac{z^2}{n}\right)^{-\frac{n+1}{2}},$$

ahol

$$\Gamma(a) = \int_0^\infty y^{a-1} e^{-y} dy,$$

ha $a > 0$. Bár ez utóbbi Γ függvény önmagában is rendkívül fontos, itt csak annyit jegyzünk meg róla, hogy egyrészt a fenti előállításból láthatóan $a > 0$ esetén pozitív értéket vesz fel (hiszen egy pozitív függvény integrálja), valójában pedig pozitív egész n esetén $\Gamma(n) = (n-1)!$ érvényes, továbbá viszonylag könnyen kiszámolható a $\Gamma(\frac{1}{2}) = \sqrt{\pi}$ érték is.

Tekintsük most az $n = 1$ esetet, ekkor a fenti sűrűségfüggvény az előző megjegyzések alapján a

$$g_1(z) = \frac{1}{\pi} \cdot \frac{1}{1+z^2}$$

alakot ölti. Az $\frac{1}{1+z^2}$ függvénynek a primitív függvénye az $\arctg(z)$ függvény, így tehát a $t(1)$ eloszlás eloszlásfüggvénye

$$G_1(y) = \frac{1}{\pi} \int_{-\infty}^y \frac{1}{1+z^2} dz = \frac{1}{\pi} \arctg(y) + \frac{1}{2},$$

amelyről a 3.2.7. gyakorlat d) részében már beláttuk, hogy eloszlásfüggvény.

Térjünk vissza a kezdeti problémánkra. Ha adott egy $X_1, \dots, X_n \sim N(\mu; \sigma^2)$ független, normális eloszlású minta, akkor egyrészt a 4.3.1. tétel szerint

$$\sum_{i=1}^n X_i \sim N(n\mu; n\sigma^2),$$

így a normális eloszlás transzformáltjáról tanultak, egészen pontosan a 3.4.5. állítás alapján $\bar{X} - \mu \sim N(0; \sigma^2/n)$, vagyis

$$\frac{\sqrt{n}}{\sigma} (\bar{X} - \mu) \sim N(0; 1).$$

Belátható továbbá, hogy egy normális eloszlásból származó független minta esetén \bar{X} és S_n^{*2} függetlenek, valamint

$$\frac{(n-1)S_n^{*2}}{\sigma^2} \sim \chi^2(n-1),$$

így definíció szerint

$$(4.2) \quad \sqrt{n-1} \cdot \frac{\sqrt{n}(\bar{X} - \mu)}{\sigma} \cdot \sqrt{\frac{\sigma^2}{(n-1)S_n^{*2}}} = \sqrt{n} \cdot \frac{\bar{X} - \mu}{S_n^*} \sim t(n-1).$$

Ezt felhasználva a következőképp szerkeszthetünk $1 - \varepsilon$ szintű konfidenciaintervallumot a háttéreloszlás várható értékére: ismét $(\bar{X} - r_\varepsilon; \bar{X} + r_\varepsilon)$ alakban keressük az intervallumot, így az

$$1 - \varepsilon = \mathbb{P}(\bar{X} - r_\varepsilon < \mu < \bar{X} + r_\varepsilon) = \mathbb{P}(-r_\varepsilon < \bar{X} - \mu < r_\varepsilon)$$

egyenlőségből indulunk ki. Transzformáljuk most úgy a középső változót, hogy az (4.2) változó kerüljön a helyére:

$$1 - \varepsilon = \mathbb{P}\left(-\frac{r_\varepsilon \sqrt{n}}{S_n^*} < \sqrt{n} \cdot \frac{\bar{X} - \mu}{S_n^*} < \frac{r_\varepsilon \sqrt{n}}{S_n^*}\right).$$

Keressük tehát egy 0 körüli szimmetrikus $(-c; c)$ intervallumot, amelybe a $Z_{n-1} := \sqrt{n} \cdot \frac{\bar{X} - \mu}{S_n^*}$ változó $1 - \varepsilon$ valószínűséggel beleesik. Azonban

$$\mathbb{P}(-c < Z_{n-1} < c) = G_{n-1}(c) - G_{n-1}(-c),$$

ahol G_{n-1} a Z_{n-1} $n-1$ szabadságfokú Student-eloszlású változó eloszlásfüggvénye. Mivel a változó g_{n-1} sűrűségfüggvénye páros, így a normális eloszlás eloszlásfüggvényének tárgyalásánál bizonyított 3.4.6. állítás után tett megjegyzés szerint

$$G_{n-1}(-c) = 1 - G_{n-1}(c)$$

most is érvényes. Tehát keressük az $1 - \varepsilon = 2G_{n-1}(c) - 1$ egyenlet megoldását. Mivel a G_{n-1} függvény g_{n-1} deriváltja mindenütt pozitív (ez látszik a fenti képletből), így - ahogy azt a

Φ függvény esetén is láttuk - a G_{n-1} függvény szigorúan monoton növény, ezért kölcsönösen egyértelmű, és létezik az inverzfűggvénye. Mindezt összerakva tehát

$$\frac{r_\varepsilon \sqrt{n}}{S_n^*} = G_{n-1}^{-1} \left(1 - \frac{\varepsilon}{2} \right)$$

adódik. Vezessük be a $t_{\varepsilon/2}(n-1) = G_{n-1}^{-1} \left(1 - \frac{\varepsilon}{2} \right)$ jelölést, ezt a $t(n-1)$ eloszlás $1 - \varepsilon/2$ kvantilisének nevezzük, értékét pedig a jegyzet végén szereplő táblázatból olvashatjuk ki. Vegyük észre, hogy itt két paraméterünk is van, n ill. ε , ezért a Student-eloszlás táblázatai a kvantiliseket csak a tipikusan használt értékekre (pl. kis n paraméterekre) tartalmazzák. A jegyzetben szereplő táblázatban az $1 - \varepsilon$ szignifikanciaszint oszlopához és az n szabadságfok sorához tartozó cellában a $t_{\varepsilon/2}(n)$ érték található.

Összefoglalva az eddigieket, az $1 - \varepsilon$ szignifikanciaszintű intervallum sugarát az

$$r_\varepsilon = \frac{t_{\varepsilon/2}(n-1) S_n^*}{\sqrt{n}}$$

formula határozza meg. Megjegyezzük (a technikai részleteket ezúttal teljesen mellőzve), hogy nagy mintaelemszám esetén a korrigált tapasztalati szórás jól közelíti a háttéreloszlás szórását (azt korábban láttuk, hogy S_n^{*2} torzítatlan becslés σ^2 -re), így ebben az esetben a fenti képlet helyett az előző szakaszban kapott képlet alkalmazható, pontosabban abban a szórás helyére a realizációból kapott s^* értéket kell beírni. Azaz: "nagy" minta esetén az

$$r_\varepsilon = \frac{\Phi^{-1} \left(1 - \frac{\varepsilon}{2} \right) s_n^*}{\sqrt{n}},$$

"kis" mintaelemszám esetén pedig az

$$r_\varepsilon = \frac{t_{\varepsilon/2}(n-1) s_n^*}{\sqrt{n}}$$

képlet javasolt. A kis ill. nagy jelzők pontosítása persze szükséges: gyakorlati statisztika-könyvek $n \geq 30$ -tól beszélnek nagy mintáról.

4.4. Hipotézisvizsgálat

Tegyük fel, hogy egy üzembn cukrot csomagolnak, és 1 kg-os csomagokat szeretnének gyártani. A csomagba kerülő cukor mennyiségét számos tényező befolyásolja, így előfordulhat, hogy valamelyest eltér 1 kg-tól. A gyártási folyamatot minden esetre úgy szeretnék beállítani, hogy a csomagban lévő cukor *várható értéke* 1 kg legyen.

A beállítások elvégzése után tesztelni szeretnék, hogy azok megfelelőek-e. Vesznek tehát egy 25 elemű mintát, és megállapítják, hogy a csomagban lévő cukor mennyisége átlagosan 0,98 kg. Mire következtethetnek ezek alapján? Az eltérés lehet a véletlen játéka, hiszen az egyes esetekben különbözhet a tényleges mennyiség a várható értéktől. Hogyan dönthetik el, hogy a kapott eredmény a véletlen műve, vagy pedig a beállítás rossz?

A fenti szituációt most általánosan fogjuk kezelni. Az ilyen típusú problémát *hipotézisvizsgálatnak* hívjuk, hiszen van egy előzetes feltevésünk (hipotézisünk) az eloszlásra vonatkozóan, és erről egy konkrét X_1, \dots, X_n minta alapján szeretnénk eldönteni, hogy helyes-e. Az eloszlásra vonatkozó feltevést *nullhipotézisnek* nevezzük, ezt H_0 fogja jelölni. Az H_1 -gyel jelölt

ellenhipotézis a nullhipotézis ellentéte, ez tehát egyszerűen a "nem igaz, hogy H_0 " állítás. Azaz a hipotézisvizsgálat során el kell döntenünk, hogy a H_0 nullhipotézist elfogadjuk (és ekkor persze H_1 -et elvetjük), vagy pedig az ellenhipotézist fogadjuk el, és H_0 -t elvetjük.

Természetesen, mivel a háttéreloszlásról csak korlátozott információnk van (a mintán keresztül), így teljesen bizonyosan sosem dönthetünk, döntésünk helyességének csupán a valószínűségéről beszélhetünk. Ekkor persze pozitív valószínűséggel előfordulhat, hogy helytelenül döntünk. Alapvetően kétféle hibát különböztetünk meg. Lehetséges, hogy a H_0 nullhipotézis fennáll, azt mégis elvetjük. Ezt *elsőfajú hibának* nevezzük. Ha a nullhipotézis nem áll fenn, de azt mégis elfogadjuk, akkor *másodfajú hibáról* beszélünk.

Látni fogjuk, hogy nem feltétlenül tudjuk egyszerre mindkét típusú hibát kontrollálni. Általában olyan szituációkat kell kezelni, ahol az elsőfajú hiba súlyosabbnak bizonyul, mint a másodfajú, éppen ezért a tárgyalt módszereink is olyanok, hogy az elsőfajú hiba valószínűségét vagyunk képesek uralni. Ha ennek valószínűsége ε , azaz

$$\mathbb{P}(H_0 \text{ fennáll, de elvetjük}) = \varepsilon,$$

akkor azt mondjuk, hogy a döntés *szignifikanciaszintje* (vagy másképp mondva *megbízhatósági szintje*) $1 - \varepsilon$. Ezt a későbbiekben mindig előre rögzíteni fogjuk. A fenti példában a nullhipotézisünk az, hogy az üzem dolgozói jól végezték el a beállításokat. Szeretnénk lehetőség szerint nagyon alacsonyra (ε kicsire) csökkenteni annak valószínűségét, hogy ártatlanul megvádoljuk őket az ellenkezőjével. A másodfajú hiba itt azt jelenti, hogy a beállításnál hibáztak, mégis felmentjük a "bűnösöket". Ezt tekintjük tehát a kevésbé súlyos hibának.

Az eljárást, amivel H_0 helyességéről döntünk, *statisztikai próbának* nevezzük. Itt ún. *paraméteres próbákkal* foglalkozunk, ahol a feltevésünk az eloszlás valamilyen paraméterével kapcsolatos. Bár a paraméteres próbák elmélete lényegesen általánosabban is kidolgozható, mi most egyetlen speciális esetre szorítunk, mégpedig arra, amikor a H_0 nullhipotézis a háttéreloszlás várható értékére vonatkozik. Egészen pontosan - ahogy a fenti példában is - az $\mathbb{E}(X) = \mu_0$ hipotézis érvényességét szeretnénk eldönteni. Ekkor az ellenhipotézis egyszerűen az, hogy $\mathbb{E}(X) \neq \mu_0$.

A próba elvégzéséhez mindig kiszámolunk egy $T(X_1, \dots, X_n)$ statisztikát, melynek értékészletét a diszjunkt \mathcal{X}_e *elfogadási tartományra* és \mathcal{X}_k *kritikus tartományra* osztjuk. Azaz $\mathcal{X}_e \cap \mathcal{X}_k = \emptyset$, és

$$\text{ran } T(X_1, \dots, X_n) = \mathcal{X}_e \cup \mathcal{X}_k.$$

Ha a konkrét realizációra a statisztika értéke az elfogadási tartományba esik, akkor elfogadjuk, ellenkező esetben elvetjük H_0 -t. Természetesen a tartományokat az előzetesen rögzített $1 - \varepsilon$ szignifikanciaszint alapján határozzuk meg, mégpedig úgy, hogy

$$\mathbb{P}(H_0 \text{ fennáll, de } T(X_1, \dots, X_n) \notin \mathcal{X}_e) = \varepsilon$$

teljesüljön.

A próba menetét tehát a következőképp összegezzük:

- az alapprobléma ismeretében kiválasztunk egy megfelelő T próbastatisztikát,
- rögzítjük az $1 - \varepsilon$ szignifikanciaszintet,
- meghatározzuk az ehhez tartozó elfogadási és kritikus tartományokat,
- kiszámoljuk a próbastatisztikát a minta konkrét realizációjára,
- ha ez az érték az elfogadási tartományba esik, akkor elfogadjuk H_0 -t, különben pedig elvetjük.

Egymintás u -próba

Tegyük fel, hogy az $X \sim N(\mu; \sigma^2)$ háttérváltozó normális, a σ szórás pedig ismert. Meg szeretnénk állapítani, hogy az X várható értéke megegyezik-e egy adott μ_0 értékkel. Vagyis

$$H_0 : \mu = \mu_0, \quad H_1 : \mu \neq \mu_0.$$

A próbát azért nevezzük egymintásnak, mert egyetlen háttérváltozó eloszlásáról döntünk egyetlen minta alapján.

Korábban a háttéreloszlás várható értékét a mintaátlaggal becsültük. Tulajdonképpen most is pontosan ugyanazt tesszük, az elfogadási tartomány meghatározásához pedig az \bar{X} körüli, $1 - \varepsilon$ szintű szignifikanciaszintű r_ε sugarú konfidenciaintervallumból indulunk ki. Ha $\mu = \mu_0$ teljesül, akkor

$$1 - \varepsilon = \mathbb{P}(\bar{X} - r_\varepsilon < \mu_0 < \bar{X} + \varepsilon) = \mathbb{P}(\bar{X} \in (\mu_0 - r_\varepsilon; \mu_0 + r_\varepsilon)),$$

tehát a $(\mu_0 - r_\varepsilon; \mu_0 + \varepsilon)$ megfelelő elfogadási tartomány lesz.

A korábban látottak alapján az r_ε sugár a szignifikanciaszint mellett a minta elemszámától és az háttéreloszlás szórásától is függ. Hogy ezt elkerüljük, módosítjuk a statisztikát, aminek köszönhetően az elfogadási tartományt már a szignifikanciaszint önmagában meghatározza, a mintától és az eloszlástól való függést pedig teljes egészében a statisztikában kódoljuk el:

$$\bar{X} \in (\mu_0 - r_\varepsilon; \mu_0 + r_\varepsilon) \iff |\bar{X} - \mu_0| < r_\varepsilon = \Phi^{-1}\left(1 - \frac{\varepsilon}{2}\right) \cdot \frac{\sigma}{\sqrt{n}},$$

ez pedig pontosan akkor teljesül, ha

$$\left| \sqrt{n} \cdot \frac{\bar{X} - \mu_0}{\sigma} \right| < \Phi^{-1}\left(1 - \frac{\varepsilon}{2}\right) =: u_{\varepsilon/2},$$

ahol az $u_{\varepsilon/2}$ a sztenderd normális eloszlás $(1 - \frac{\varepsilon}{2})$ -kvantilise. Ezt *kritikus értékek* nevezzük, mert ez választja el az elfogadási és a kritikus tartományt, melyek definícióját hamarosan megadjuk. Előbb azonban rögzítsük a próbastatisztikát:

$$u(X_1, \dots, X_n) = \sqrt{n} \cdot \frac{\bar{X} - \mu_0}{\sigma}.$$

Vegyük észre, hogy ez a statisztika a szignifikanciaszinttől már független, így az csak az elfogadási tartományt befolyásolja, amit pedig a $(-u_{\varepsilon/2}; u_{\varepsilon/2})$ intervallum definiál (a kritikus tartomány pedig nyilván ennek komplementere). Ilyen módon az ε értékét és a többi paramétert szétválasztottuk. Ez hasznos, ha egy konkrét számolásban az ε értékét vagy akár az n mintaelemszámot módosítani szeretnénk a többi paraméter fixálása mellett, hiszen ekkor vagy csak a statisztika értékét, vagy csak az elfogadási tartományt kell újraszámolni.

Összefoglalva: ha $X \sim N(\mu; \sigma^2)$, ahol σ ismert, akkor a $H_0 : \mu = \mu_0$ nullhipotézist az $1 - \varepsilon$ szignifikanciaszint mellett pontosan akkor fogadjuk el, ha

$$|u(X_1, \dots, X_n)| < u_{\varepsilon/2}.$$

4.4.1. Példa. Tekintsük a bevezető példánkat. Tegyük fel, hogy a csomagban lévő cukor várható értéke normális eloszlást követ 0,05 szórással. Döntsünk 95%-os szignifikanciaszint mellett. A 25 elem mintára a mintaátlag 0,98 kg, így a próbastatisztikánk értéke

$$u(X_1, \dots, X_n) = 5 \cdot \frac{0,98 - 1}{0,05} = -2.$$

Mivel $u_{0,025} = 1,96$, így a nullhipotézist elvetjük.

Vizsgáljuk meg, hogy mi történik 99%-os szignifikanciaszint mellett. Ekkor a próbat statisztika nem változik, míg $u_{0,005} = 2,58$, így ha szigorúbbak vagyunk, akkor már H_0 -t el kell fogadjuk.

Egymintás t -próba

Legyen most a háttérváltozó $X \sim N(\mu; \sigma^2)$, ahol a σ szórást ezúttal nem ismerjük. Ugyanúgy, ahogy az előző esetben, $H_0 : \mu = \mu_0$ és így $H_1 : \mu \neq \mu_0$ valamilyen $\mu_0 \in \mathbb{R}$ számra.

Rögzítsünk először egy ε -t. Az eljárásunk itt is ugyanaz, mint az imént: az \bar{X} középpontú, r_ε sugarú konfidenciaintervallumból indulunk ki. Ha $\mu = \mu_0$ teljesül, akkor

$$1 - \varepsilon = \mathbb{P}(\bar{X} - r_\varepsilon < \mu_0 < \bar{X} + \varepsilon).$$

A fenti egyenlőtlenségekben az r_ε értékét behelyettesítve

$$\bar{X} - \frac{t_{\varepsilon/2}(n-1)S_n^*}{\sqrt{n}} < \mu_0 < \bar{X} + \frac{t_{\varepsilon/2}(n-1)S_n^*}{\sqrt{n}}$$

adódik, amit átrendezve a következő ekvivalens alakot kapjuk:

$$\left| \sqrt{n} \cdot \frac{\bar{X} - \mu_0}{S_n^*} \right| < t_{\varepsilon/2}(n-1),$$

ahol $t_{\varepsilon/2}(n-1)$ az $n-1$ szabadságfokú Student-eloszlás $(1 - \varepsilon/2)$ -kvantilise.

Ebben az esetben tehát a

$$t(X_1, \dots, X_n) = \sqrt{n} \cdot \frac{\bar{X} - \mu_0}{S_n^*}$$

próbat statisztikát választjuk, az elfogadási tartomány pedig a $(-t_{\varepsilon/2}(n-1); t_{\varepsilon/2}(n-1))$ intervallum.

A másodfajú hiba

Végül nagyon röviden kitérünk a másodfajú hiba esélyére. Ez akkor fordul elő, ha a nullhipotézis nem áll fent, mi mégis elfogadjuk. Ekkor persze a hiba valószínűsége nem csak a korábbi paraméterektől, de magától a μ várható értéktől is függ. Kiszámolható, hogy a másodfajú hiba valószínűsége akkor nagy, ha μ a μ_0 közelében van. Ez persze nem meglepő, ha a tippünk csak egy kicsit rossz, akkor elég nagy az esélye, hogy a hiba nem tűnik fel a minta alapján.

Ha a szignifikanciaszintet növeljük (vagyis ε -t csökkentjük), akkor kiderül, hogy a másodfajú hiba valószínűsége nő. Az elsőfajú hiba esélyének csökkentésével párhuzamosan tehát azt az árat fizetjük, hogy a kevésbé súlyos másodfajú hiba viszont valószínűbb.

Az is belátható, hogy a mintaelemszám növelésével a másodfajú hiba valószínűsége 0-hoz tart, ilyen módon tehát mégis lehetséges valamiféle kontroll a másodfajú hiba felett is (még ha a gyakorlatban arról nincs is szó, hogy előírhatnánk a valószínűségét). A részletes elemzés megtalálható az [1] könyvben.

Táblázatok

A sztenderd normális eloszlás eloszlásfüggvényének táblázata

x	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0,0	0,5000	0,5040	0,5080	0,5120	0,5160	0,5199	0,5239	0,5279	0,5319	0,5359
0,1	0,5398	0,5438	0,5478	0,5517	0,5557	0,5596	0,5636	0,5675	0,5714	0,5753
0,2	0,5793	0,5832	0,5871	0,5910	0,5948	0,5987	0,6026	0,6064	0,6103	0,6141
0,3	0,6179	0,6217	0,6255	0,6293	0,6331	0,6368	0,6406	0,6443	0,6480	0,6517
0,4	0,6554	0,6591	0,6628	0,6664	0,6700	0,6736	0,6772	0,6808	0,6844	0,6879
0,5	0,6915	0,6950	0,6985	0,7019	0,7054	0,7088	0,7123	0,7157	0,7190	0,7224
0,6	0,7257	0,7291	0,7324	0,7357	0,7389	0,7422	0,7454	0,7486	0,7517	0,7549
0,7	0,7580	0,7611	0,7642	0,7673	0,7704	0,7734	0,7764	0,7794	0,7823	0,7852
0,8	0,7881	0,7910	0,7939	0,7967	0,7995	0,8023	0,8051	0,8078	0,8106	0,8133
0,9	0,8159	0,8186	0,8212	0,8238	0,8264	0,8289	0,8315	0,8340	0,8365	0,8389
1,0	0,8413	0,8438	0,8461	0,8485	0,8508	0,8531	0,8554	0,8577	0,8599	0,8621
1,1	0,8643	0,8665	0,8686	0,8708	0,8729	0,8749	0,8770	0,8790	0,8810	0,8830
1,2	0,8849	0,8869	0,8888	0,8907	0,8925	0,8944	0,8962	0,8980	0,8997	0,9015
1,3	0,9032	0,9049	0,9066	0,9082	0,9099	0,9115	0,9131	0,9147	0,9162	0,9177
1,4	0,9192	0,9207	0,9222	0,9236	0,9251	0,9265	0,9279	0,9292	0,9306	0,9319
1,5	0,9332	0,9345	0,9357	0,9370	0,9382	0,9394	0,9406	0,9418	0,9429	0,9441
1,6	0,9452	0,9463	0,9474	0,9484	0,9495	0,9505	0,9515	0,9525	0,9535	0,9545
1,7	0,9554	0,9564	0,9573	0,9582	0,9591	0,9599	0,9608	0,9616	0,9625	0,9633
1,8	0,9641	0,9649	0,9656	0,9664	0,9671	0,9678	0,9686	0,9693	0,9699	0,9706
1,9	0,9713	0,9719	0,9726	0,9732	0,9738	0,9744	0,9750	0,9756	0,9761	0,9767
2,0	0,9772	0,9778	0,9783	0,9788	0,9793	0,9798	0,9803	0,9808	0,9812	0,9817
2,1	0,9821	0,9826	0,9830	0,9834	0,9838	0,9842	0,9846	0,9850	0,9854	0,9857
2,2	0,9861	0,9864	0,9868	0,9871	0,9875	0,9878	0,9881	0,9884	0,9887	0,9890
2,3	0,9893	0,9896	0,9898	0,9901	0,9904	0,9906	0,9909	0,9911	0,9913	0,9916
2,4	0,9918	0,9920	0,9922	0,9925	0,9927	0,9929	0,9931	0,9932	0,9934	0,9936
2,5	0,9938	0,9940	0,9941	0,9943	0,9945	0,9946	0,9948	0,9949	0,9951	0,9952
2,6	0,9953	0,9955	0,9956	0,9957	0,9959	0,9960	0,9961	0,9962	0,9963	0,9964
2,7	0,9965	0,9966	0,9967	0,9968	0,9969	0,9970	0,9971	0,9972	0,9973	0,9974
2,8	0,9974	0,9975	0,9976	0,9977	0,9977	0,9978	0,9979	0,9979	0,9980	0,9981
2,9	0,9981	0,9982	0,9982	0,9983	0,9984	0,9984	0,9985	0,9985	0,9986	0,9986
3,0	0,9987	0,9987	0,9987	0,9988	0,9988	0,9989	0,9989	0,9989	0,9990	0,9990
3,1	0,9990	0,9991	0,9991	0,9991	0,9992	0,9992	0,9992	0,9992	0,9993	0,9993
3,2	0,9993	0,9993	0,9994	0,9994	0,9994	0,9994	0,9994	0,9995	0,9995	0,9995
3,3	0,9995	0,9995	0,9995	0,9996	0,9996	0,9996	0,9996	0,9996	0,9996	0,9997
3,4	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9998

A Student-eloszlás kvantiliseinek táblázata

Szabadsági fok	Szignifikanciaszint ($1 - \varepsilon$)				
	0,9	0,95	0,98	0,99	0,999
1	6,314	12,706	31,821	63,657	636,619
2	2,920	4,303	6,965	9,925	31,599
3	2,353	3,182	4,541	5,841	12,924
4	2,132	2,776	3,747	4,604	8,610
5	2,015	2,571	3,365	4,032	6,869
6	1,943	2,447	3,143	3,707	5,959
7	1,895	2,365	2,998	3,499	5,408
8	1,860	2,306	2,896	3,355	5,041
9	1,833	2,262	2,821	3,250	4,781
10	1,812	2,228	2,764	3,169	4,587
11	1,796	2,201	2,718	3,106	4,437
12	1,782	2,179	2,681	3,055	4,318
13	1,771	2,160	2,650	3,012	4,221
14	1,761	2,145	2,624	2,977	4,140
15	1,753	2,131	2,602	2,947	4,073
16	1,746	2,120	2,583	2,921	4,015
17	1,740	2,110	2,567	2,898	3,965
18	1,734	2,101	2,552	2,878	3,922
19	1,729	2,093	2,539	2,861	3,883
20	1,725	2,086	2,528	2,845	3,850
21	1,721	2,080	2,518	2,831	3,819
22	1,717	2,074	2,508	2,819	3,792
23	1,714	2,069	2,500	2,807	3,768
24	1,711	2,064	2,492	2,797	3,745
25	1,708	2,060	2,485	2,787	3,725
26	1,706	2,056	2,479	2,779	3,707
27	1,703	2,052	2,473	2,771	3,690
28	1,701	2,048	2,467	2,763	3,674
29	1,699	2,045	2,462	2,756	3,659
30	1,697	2,042	2,457	2,750	3,646
40	1,684	2,021	2,423	2,704	3,551
50	1,676	2,009	2,403	2,678	3,496
60	1,671	2,000	2,390	2,660	3,460
80	1,664	1,990	2,374	2,639	3,416
100	1,660	1,984	2,364	2,626	3,390
500	1,648	1,965	2,334	2,586	3,310

A gyakorlatok, feladatok végeredményei

1. fejezet

1.1.1. $A_1 = \bar{E} \cap \bar{P}$, $A_2 = E \cap P$, $A_3 = N \cap \bar{E} \cap P$, $A_4 = N \cap E \cap \bar{P}$, $A_5 = \bar{N} \cap P$,
 $B = (E \setminus P) \cup (\bar{N} \cap P)$

1.1.2. pl. $\Omega = \{(i, j) : 1 \leq i, j \leq 6\}$ jó,
de Ω elemei lehetnek a dobott számokból álló halmazok is
 $A = S_7 \cap \bar{S}_8$, $B = MAX_2 \cap S_4$, $C = S_{12} \cup MAX_1 \cup (S_7 \cap \bar{S}_8 \cap MAX_6)$

1.1.3. $\frac{1}{9}$, $\frac{15}{36}$

1.1.4. a) $\frac{1}{32}$ b) $\frac{15}{64}$ c) $\frac{57}{64}$ d) $\frac{1}{2}, \frac{1}{2}$ e) 1

1.1.5. a) $\frac{\binom{5}{2} \cdot \binom{85}{3}}{\binom{90}{5}} \approx 0,0225$

b) $\frac{\binom{5}{k} \cdot \binom{85}{5-k}}{\binom{90}{5}}$ ha $k = 0, 1, 2, 3, 4, 5$

c) $\frac{\binom{77}{4}}{\binom{90}{5}}$ d) $1 - \frac{\binom{79}{5}}{\binom{90}{5}}$

1.1.6. a) $\frac{3^3}{\binom{9}{3}} = \frac{27}{84}$ b) $\frac{3}{84}$

1.1.7. a) $\frac{1}{2}$, b) $\frac{1}{2}$

1.1.8. a) $\frac{\binom{n-1}{k-1}}{\binom{n}{k}}$, b) $\frac{\binom{j-1}{i-1} \binom{n-i}{k-i}}{\binom{n}{k}}$, d) $\sum_{m=1}^n m = \frac{n(n+1)}{2}$, $\sum_{m=1}^n m^2 = \frac{n(n+1)(2n+1)}{6}$

1.1.9. 0,93952

1.1.10. $\mathbb{P}(A \cup B_{10}) \approx 0,6026$, $\mathbb{P}(A \cup B_9 \cup B_{10}) \approx 0,6556$

1.1.11. a) 0,07 b) 0,63

1.1.12. $\mathbb{P}(A \cap C) = \frac{1}{9}$

1.2.4. $\frac{1}{2}$

1.2.5.

A és B nem függetlenek: $\mathbb{P}(A \cap B) = \frac{1}{9}$, $\mathbb{P}(A)\mathbb{P}(B) = \frac{1}{2} \cdot \frac{1}{3} = \frac{1}{6}$, $\mathbb{P}(A | B) = \frac{1}{3}$

A és C függetlenek $\mathbb{P}(A | C) = \mathbb{P}(A) = \frac{1}{2}$

1.2.6. a) 0,096, b) 0,916, c) 0,084.

1.2.7. $\mathbb{P}(A) = \frac{1}{3}$, $\mathbb{P}(B) = \frac{5}{6}$, $\mathbb{P}(A | \bar{B}) = 1$, nem független A és B

1.2.8. $\frac{1}{4845} \approx 0,0002$

1.2.9. 3 ajtó esetén: ha nem váltunk, a nyerségi esélyünk $\frac{1}{3}$, ha váltunk, akkor $\frac{2}{3}$

1.2.10. a) $\frac{21}{128}$, b) $\frac{1}{63}$

1.2.11. Feltéve, hogy mindkét húzott golyó piros, annak a legnagyobb a valószínűsége, hogy 1 db hatost dobtunk, és ez a valószínűség $\frac{5}{7}$ (tehát erre érdemes tippelni).

Továbbá $\frac{2}{7}$ valószínűséggel dobtunk a fenti feltétel mellett 2 db hatost, és 0 valószínűséggel 0 db hatost.

1.2.12. 0,175; $\approx 0,2303$; $\approx 0,5152$; $\approx 0,2545$

1.2.13. $\mathbb{P}(\text{tudja a választ} | \text{helyesen válaszolt}) = \frac{3p}{1+2p}$, $p = \frac{1}{4}$ esetén ennek értéke $\frac{1}{2}$

2. fejezet

2.1.2. $\mathbb{P}(X < 4) = \frac{23}{28}$, $\mathbb{P}(X \geq 3) = \frac{5}{14}$

2.1.3. függetlenek

2.1.4. a) $\left(\frac{3}{4}\right)^{10} \approx 0,0563$ b) $\approx 0,2241$

c) a legnagyobb valószínűséggel 2 helyes válaszunk lesz

11 kérdés esetén 2 és 3 választ is ugyanolyan és maximális valószínűséggel kapunk

2.1.5. $Geo(0,5)$

2.1.6. a) $\mathbb{P}(2 \leq X \leq 3) = \frac{10}{27}$, $\mathbb{P}(X > 3) = \frac{8}{27}$ b) $\frac{8}{27}$

2.1.7. $1 - \left(\frac{9}{10}\right)^9 = \frac{612\,579\,511}{1\,000\,000\,000} \approx 0,6126$

2.1.8. $p = \frac{1}{N}$; a határérték: $\frac{1}{e}$

2.1.9. az első esetben az $f(k) = \frac{1}{2^{k-1}}$, ha $k \geq 2$ egész,

a második esetben $f(k) = \frac{F_{k-2}}{2^{k-1}}$, ha $k \geq 3$ egész, ahol F_k a k -edik Fibonacci-szám

2.2.1. nem függetlenek

2.2.2. X és Y nem függetlenek, X és Y együttes eloszlása:

$X \backslash Y$	0	1	2
0	1/4	0	0
1	1/3	1/6	0
2	1/9	1/9	1/36

2.2.3. a) $p = \frac{1}{60}$ b) $\mathbb{P}(X \leq 0, Y = 1) = \frac{1}{3}$ c) függetlenek

2.2.4. a) függetlenek b) függetlenek c) nem függetlenek

2.2.5. $\text{ran}(X + Y) = \mathbb{N} \setminus \{0, 1\}$, $k \geq 2$ esetén $f_{X+Y}(k) = pq \frac{(1-p)^{k-1} - (1-q)^{k-1}}{q-p}$

2.3.1. $\approx 154,1527$

2.3.2. $\frac{19}{6}$

2.3.3. $\text{ran} Y = \mathbb{N}^+$, $\mathbb{P}(Y = k) = \frac{1}{k(k+1)}$, $\mathbb{E}(Y) = \infty$

2.3.4. 3 ill. 7

2.3.5. 2

2.3.6. $\frac{27}{19}$

2.3.7. a) 8 b) $7\frac{2}{3}$

2.3.8. 6

2.3.9. 3,7

2.3.10. 8

2.3.11. $\frac{55}{3}$

2.3.12. $\frac{1}{9}$

2.3.13. $\mathbb{E}(X^2) = np + n^2p^2 - np^2$

2.4.1. $\mathbb{D}^2(X) = \frac{24}{25}$, $\mathbb{D}^2(Y) = \frac{21}{25}$

$f_{X+Y}(-2) = \frac{1}{10}$, $f_{X+Y}(0) = \frac{1}{2}$, $f_{X+Y}(2) = \frac{2}{5}$, $\mathbb{D}^2(X + Y) = \frac{41}{25}$

2.4.2. X =dobott számok összege, $\mathbb{E}(X) = 9$, $\mathbb{D}(X) = \sqrt{\frac{21}{2}} \approx 3,2404$

2.4.2. $\mathbb{E}((3 - X)^2) = 6$, $\mathbb{D}(5 - 2X) = 2\sqrt{6} \approx 4,8990$, $\mathbb{E}((X + 1)(X - 2)) = 10$

2.5.3. a) $X_r = Y_1 + \dots + Y_r$, ahol $Y_1 = 1$, és $Y_i = X_i - X_{i-1}$ minden $1 < i \leq r$ esetén,

$$Y_i \sim Geo\left(\frac{N - i + 1}{N}\right)$$

$$\mathbb{E}(X_r) = N \cdot \sum_{i=1}^r \frac{1}{N - i + 1}$$

$$\mathbb{D}^2(X_r) = N \cdot \sum_{i=1}^r \frac{i - 1}{(N - i + 1)^2}$$

$$P_{X_r}(z) = \frac{N!}{N^r(N - r)!} \cdot \frac{z^r}{\prod_{i=1}^r \left(1 - \frac{i-1}{N}z\right)}$$

b) A rekurziós formulából

$$P_{X_r}(z) = \frac{N - r + 1}{N} \cdot z \cdot \frac{1}{\left(1 - \frac{r-1}{N}z\right)} \cdot P_{X_{r-1}}(z),$$

ahol $P_{X_1}(z) = z$. Innen teljes indukcióval adódik a generátorfüggvény.

2.5.4. X a dobások száma, $\mathbb{E}(X) = 7$, $\mathbb{D}^2(X) = 22$, $P'_X(z) = \frac{z^3}{4 - 2z - z^2}$.

2.6.3. 0,4046

2.6.4. 3,5973; 319,027

2.6.5. 0,2583

2.6.6. a) a keresett valószínűség $\approx 0,5785$, a keresett szórás $\approx 1,7193$

b) egy $Pois(3)$ eloszlású változóval közelített valószínűség $\approx 0,5768$

3. fejezet

3.1.1. 0,0707

3.1.2. függetlenek

3.1.3. $\frac{23}{200}$

3.1.4. $\frac{1}{2}$

3.1.5. $\frac{5}{9}$

3.2.1.

$$a) F_X(t) = \begin{cases} 0, & \text{ha } t \leq 0, \\ 2t - t^2, & \text{ha } 0 < t \leq 1, \\ 1, & \text{ha } 1 < t. \end{cases}$$

$$\text{b) } \mathbb{P}(0,25 \leq X < 0,5) = \frac{5}{16}$$

3.2.2. az előző feladat a) részével azonos a megoldás

$$3.2.3. \mathbb{P}(X > 0) = \frac{1}{2}$$

$$F_X(t) = \begin{cases} 0, & \text{ha } t \leq -2, \\ \frac{(2+t)^2}{8}, & \text{ha } -2 < t \leq 0, \\ 1 - \frac{(2-t)^2}{8}, & \text{ha } 0 < t \leq 2, \\ 1, & \text{ha } t > 2. \end{cases}$$

3.2.4.

$$F_X(t) = \begin{cases} 0, & \text{ha } t \leq 0, \\ 4t - 4t^2, & \text{ha } 0 < t \leq \frac{1}{2}, \\ 1, & \text{ha } \frac{1}{2} < t. \end{cases}$$

3.2.5. 0,1593; 0,7967; nem folytonos

3.2.6.

$$F_Y(t) = \begin{cases} 0, & \text{ha } t \leq 0, \\ \frac{1}{6}, & \text{ha } 0 < t \leq 1, \\ \frac{1}{2}, & \text{ha } 1 < t \leq 4, \\ \frac{5}{6}, & \text{ha } 4 < t \leq 9, \\ 1, & \text{ha } t > 9, \end{cases}$$

3.2.7. a) igen, b) pontosan akkor eloszlásfüggvény, ha $a > 0$, c) nem, d) igen.

3.3.1.

$$f_X(t) = \begin{cases} 2 - 2t, & \text{ha } t \in (0; 1), \\ 0, & \text{különben,} \end{cases} \quad \mathbb{E}(X) = \frac{1}{3}, \quad \mathbb{D}(X) = \frac{1}{3\sqrt{2}} \approx 0,2357$$

3.3.2.

$$f_X(t) = \begin{cases} \frac{2+t}{4}, & \text{ha } t \in (-2; 0), \\ \frac{2-t}{4}, & \text{ha } t \in (0; 2), \\ 0, & \text{különben,} \end{cases} \quad \mathbb{E}(X) = 0, \quad \mathbb{D}(X) = \sqrt{\frac{2}{3}} \approx 0,8165$$

$$3.3.3. \alpha = \frac{1}{2} \quad \mathbb{P}\left(\frac{1}{9} < X < \frac{1}{4}\right) = \frac{1}{6}$$

3.3.4. 45,072 l, 16,667 l

3.4.1. a) $\lambda = 2$ b) $\mathbb{P}(X < 2) = 0,9817$ c) $\mathbb{E}(X) = \frac{1}{2}$ d) $\mathbb{D}(X) = \frac{1}{2}$

3.4.2. 0,6065

3.4.3. 0,5946; 138,3017 év

3.4.5. 2,0047

3.4.6. $\mu = 12,47, \sigma^2 = 8,657$

3.4.7. $N\left(30; \left(\frac{20}{9}\right)^2\right)$

3.4.8.

$$F_Z(t) = \begin{cases} 2\Phi(\sqrt{t}) - 1, & \text{ha } t > 0, \\ 0 & \text{különben.} \end{cases}$$

3.5.1. $\approx 0,0735$

3.5.2. egy 113 fős terem elég

3.5.3. $\approx 0,9938$

3.5.4. $2 \sum_{i=1}^{12} X_i - 7$

3.5.5. a) 1000 b) 0,6628

A gyakorlatok, feladatok megoldásai

1. fejezet

1.1.7. a) Az, hogy 101 dobás után a nyúl áll előrébb, éppen azt jelenti, hogy legalább 51 fejet, azaz lefeljebb 50 írást dobtak. Bármilyen 101 hosszú fej-írás sorozat $1/2^{101}$ valószínűséggel adódik. Azon sorozatok száma, ahol pontosan k darab írás van, éppen $\binom{101}{k}$, hiszen az írások dobásainak sorszámát kiválasztva már meghatározzuk a sorozatot. Így tehát annak a keresett valószínűség

$$\frac{1}{2^{101}} \cdot \left[\binom{101}{0} + \binom{101}{1} + \binom{101}{2} + \cdots + \binom{101}{50} \right]$$

A fenti összeg kiszámolásához felhasználjuk a binomiális együtthatók szimmetriatulajdonságát: $\binom{101}{k} = \binom{101}{101-k}$. Ezt az összeget minden egyes tagjára alkalmazva

$$\binom{101}{0} + \binom{101}{1} + \binom{101}{2} + \cdots + \binom{101}{50} = \binom{101}{101} + \binom{101}{100} + \binom{101}{99} + \cdots + \binom{101}{51}$$

adódik. Azonban láttuk, hogy az összes binomiális együttható összege 2^{101} , így

$$2^{101} = \sum_{k=0}^{101} \binom{101}{k} = \sum_{k=0}^{50} \binom{101}{k} + \sum_{k=51}^{101} \binom{101}{k} = 2 \cdot \sum_{k=0}^{50} \binom{101}{k},$$

vagyis

$$\sum_{k=0}^{50} \binom{101}{k} = 2^{100},$$

és így a keresett valószínűség $\frac{2^{100}}{2^{101}} = \frac{1}{2}$.

Vegyük észre, hogy a fenti érvelés 101 helyett tetszőleges $n = 2m + 1$ páratlan szám esetén elmondható, ebben az esetben a binomiális együtthatókra a következő összefüggést kapjuk:

$$\sum_{k=0}^m \binom{n}{k} = \sum_{k=m+1}^n \binom{n}{k} = 2^{n-1}.$$

b) Hasonlóképp a fentiekhez, a páros sok fejdobás valószínűsége

$$\frac{1}{2^{101}} \cdot \left[\binom{101}{0} + \binom{101}{2} + \binom{101}{4} + \cdots + \binom{101}{100} \right].$$

Most is alkalmazhatjuk a binomiális együtthatók szimmetriáját, így azt kapjuk, hogy

$$\binom{101}{0} + \binom{101}{2} + \binom{101}{4} + \cdots + \binom{101}{100} = \binom{101}{101} + \binom{101}{99} + \binom{101}{97} + \cdots + \binom{101}{1}.$$

A két oldal összege ismételten az összes együttható összegét adja, vagyis ismét azt kapjuk, hogy mindkét oldal értéke 2^{100} , tehát a keresett valószínűség ismét $1/2$.

A fentivel analóg egyenlőség általában is megadható tetszőleges (nem csak páratlan) n esetén. Ehhez a binomiális tételt (lásd az (1.2) formulát) alkalmazzuk: az $x = -1$, $y = 1$ választás esetén

$$0 = (-1 + 1)^n = \sum_{k=0}^n \binom{n}{k} \cdot (-1)^k \cdot 1^{n-k} = \sum_{k=0}^n (-1)^k \cdot \binom{n}{k}$$

adódik. Tehát váltakozó előjellel összeadva a fenti binomiális együtthatókat 0-t kapunk. Ez pedig (a fenti egyenlőséget átrendezve) azt adja, hogy páros ill. páratlan k -ra összegezve ugyanazt a számot kapjuk, ami így szükségképp 2^{n-1} (mivel a teljes összeg 2^n).

1.1.8 a) Az, hogy a legnagyobb kihúzott szám n , éppen azt jelenti, hogy rajta kívül még az első $n - 1$ számból húztunk $k - 1$ darabot, ezt pedig $\binom{n-1}{k-1}$ -féleképp tehetjük meg. Mivel összesen $\binom{n}{k}$ -féleképp választhatunk, így a keresett valószínűség

$$\mathbb{P}(\text{a legnagyobb szám } n) = \frac{\binom{n-1}{k-1}}{\binom{n}{k}}.$$

b) Ha nagyság szerint az i -edik szám j , az éppen azt jelenti, hogy az első $j - 1$ számból választottunk $i - 1$ darabot, ezektől függetlenül pedig az utolsó $n - j$ számból pedig $k - i$ darabot, tehát a keresett valószínűség

$$\frac{\binom{j-1}{i-1} \binom{n-i}{k-i}}{\binom{n}{k}},$$

ahol a $\binom{0}{0} = 1$ definíciót is használjuk.

c) Mivel a húzásnál a nagyság szerinti i -edik szám mindenképp i és $n - k - i$ között van, és persze azok az események, hogy az i -edik húzás értéke egy ebbe az intervallumba eső j egész szám, páronként egymást kizáróak, így a valószínűségük összege 1 kell legyen. Ez könnyen levezethető indukcióval a az 1.1.1. állítás (iii) részéből illetve a $\mathbb{P}(\Omega) = 1$ egyenlőségből (lásd még az 1.1.2. állítás (ii) részét is). Tehát a fenti valószínűségeket az $i \leq j \leq n - k - i$ értékekre összegezve 1-et kapunk, $\binom{n}{k}$ -val átszorozva pedig

$$\binom{n}{k} = \sum_{j=i}^{n-k-i} \binom{j-1}{i-1} \binom{n-j}{k-i}$$

adódik. Felírva ezt $(n+1)$ -re és $(k+1)$ -re, továbbá az $m = j - 1$ és az $l = i - 1$ helyettesítéseket használva, valamint az összeghez további esetleges 0 tagokat hozzávéve (amikor $m < l$ ill. $m > n - k - l - 1$ teljesül)

$$\binom{n+1}{k+1} = \sum_{m=l}^{n-k-l-1} \binom{m}{l} \binom{n-m}{k-l} = \sum_{m=0}^n \binom{m}{l} \binom{n-m}{k-l},$$

ahol tehát kihasználtuk, hogy $\binom{a}{b}$ definíció szerint 0, ha $a < b$.

d) Ha a fenti formulában a $k = l = 1$ választással élünk, akkor a

$$\sum_{m=0}^n \binom{m}{1} \binom{n-m}{0} = \sum_{m=1}^n m = \binom{n+1}{2} = \frac{(n+1)n}{2}$$

formula adódik. Hasonlóképp a $k = 2, l = 1$ esetben

$$\begin{aligned} \binom{n+2}{3} &= \frac{(n+2)(n+1)n}{6} = \sum_{m=0}^{n+1} \binom{m}{1} \binom{n+1-m}{1} = \sum_{m=1}^n m(n+1-m) \\ &= (n+1) \sum_{m=1}^n m - \sum_{m=1}^n m^2 = \frac{(n+1)^2 n}{2} - \sum_{m=1}^n m^2 \end{aligned}$$

adódik, ezt átrendezve pedig

$$\sum_{m=1}^n m^2 = \frac{(n+1)^2 n}{2} - \frac{(n+2)(n+1)n}{6} = \frac{n(n+1)}{2} \left(n+1 - \frac{n+2}{3} \right) = \frac{n(n+1)(2n+1)}{6}.$$

1.1.9. Mind a hét ember választ egy emeletet a 10-ből, vagyis a kimeneteleink olyan 7 hosszú sorozatok lesznek, ahol minden elem egy szám 1 és 10 között. Ilyenből összesen 10^7 darab van, és mindegyikük egyformán valószínű. A kérdés tehát azon kimenetek száma, ahol ez a sorozat tartalmaz legalább két egyforma számot. Egyszerűbb megszámolni azokat a sorozatokat, ahol minden elem különböző. Magára a 7 különböző számra $\binom{10}{7}$ választási lehetőség van, és persze minden ilyen 7 elemű számhalmaznál figyelembe kell venni, hogy melyik számot melyik ember választotta. Ez voltaképp a 7 szám egy sorba rendezését (permutációját) jelenti, ezekből pedig $7!$ darab van. Vagyis $\binom{10}{7} \cdot 7!$ féleképp történhet meg, hogy mindenki máshol száll ki, így a keresett valószínűség

$$1 - \frac{\binom{10}{7} \cdot 7!}{10^7} = 1 - \frac{9!}{6 \cdot 10^6} = 0,93952.$$

1.2.3 Példát mutatunk olyan A, B, C eseményekre, amelyekre teljesül a $\mathbb{P}(A \cap B \cap C) = \mathbb{P}(A)\mathbb{P}(B)\mathbb{P}(C)$ egyenlet, de a

$$\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B), \quad \mathbb{P}(B \cap C) = \mathbb{P}(B)\mathbb{P}(C), \quad \mathbb{P}(C \cap A) = \mathbb{P}(C)\mathbb{P}(A)$$

egyenletek közül pontosan kettő teljesül. Ehhez tekintsünk egy tetszőleges olyan A eseményt, melyre $\mathbb{P}(A) = 0$ teljesül (pl. az $A = \emptyset$ jó választás, de bármi más nulla valószínűségű esemény megfelel). Ekkor $A \cap B \subset A, A \cap C \subset A$ ill. $A \cap B \cap C \subset A$ miatt az **1.1.2.** állítás (iv) pontja szerint

$$\mathbb{P}(A \cap B) \leq \mathbb{P}(A) = 0, \quad \mathbb{P}(A \cap C) \leq \mathbb{P}(A) = 0, \quad \mathbb{P}(A \cap B \cap C) \leq \mathbb{P}(A) = 0$$

is teljesül, és mivel a fenti valószínűségek nemnegatív számok, így szükségképp

$$\mathbb{P}(A \cap B) = 0, \quad \mathbb{P}(A \cap C) = 0, \quad \mathbb{P}(A \cap B \cap C) = 0,$$

ebből pedig azonnal következik, hogy

$$\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B) = 0, \quad \mathbb{P}(A \cap C) = \mathbb{P}(A)\mathbb{P}(C) = 0, \quad \mathbb{P}(A \cap B \cap C) = \mathbb{P}(A)\mathbb{P}(B)\mathbb{P}(C) = 0.$$

Vegyük észre, hogy a B és C eseményekről nem használtunk ki semmit, ha tehát ezeket olyan eseményeknek választjuk, melyek nem függetlenek egymástól, akkor a $\mathbb{P}(B \cap C) = \mathbb{P}(B)\mathbb{P}(C)$ egyenlet *nem* teljesül.

Mutatunk egy olyan konstrukciót is, ahol az A, B, C események mind pozitív valószínűségűek, a $\mathbb{P}(A \cap B \cap C) = \mathbb{P}(A)\mathbb{P}(B)\mathbb{P}(C)$ egyenlet teljesül, de az események nem páronként

függetlenek. Ehhez tekintsünk egy olyan cinkelt hat oldalú dobókockát, ahol az egyes valószínűsége $1/27$, a kettese és a hármase $7/27$, az összes többi kimenetel pedig $4/27$. (Nyilván ez egy valószínűségi eloszlás példája, de elvi jelentősége ennek nincs, ilyen matematikai modell minden korlát nélkül konstruálható.)

Tekintsük továbbá az

$$A = \{1, 4, 5\}, \quad B = \{1, 5, 6\}, \quad C = \{1, 4, 6\}$$

eseményeket. Ekkor

$$\mathbb{P}(A) = \mathbb{P}(B) = \mathbb{P}(C) = \frac{1}{27} + \frac{4}{27} + \frac{4}{27} = \frac{1}{3},$$

továbbá $A \cap B \cap C = \{1\}$, így $\mathbb{P}(A \cap B \cap C) = \frac{1}{27} = \left(\frac{1}{3}\right)^3 = \mathbb{P}(A)\mathbb{P}(B)\mathbb{P}(C)$. Ugyanakkor $A \cap B = \{1, 5\}$, így $\mathbb{P}(A \cap B) = \frac{5}{27} \neq \frac{1}{9} = \mathbb{P}(A)\mathbb{P}(B)$. Hasonlóképp látható, hogy a B és C , valamint az A és C eseménypárok nem függetlenek.

1.2.11 Kiszámoljuk a

$$\begin{aligned} &\mathbb{P}(0 \text{ db hatost dobunk} \mid \text{két piros golyót húzunk}), \\ &\mathbb{P}(1 \text{ db hatost dobunk} \mid \text{két piros golyót húzunk}), \\ &\mathbb{P}(2 \text{ db hatost dobunk} \mid \text{két piros golyót húzunk}) \end{aligned}$$

feltételes valószínűségeket. Arra logikus tippelni, ami a feltétel mellett legnagyobb valószínűséggel bekövetkezik.

A fenti valószínűségek meghatározásához a Bayes-tételt hívjuk segítségül:

$$\begin{aligned} &\mathbb{P}(0 \text{ db hatost dobunk} \mid \text{két piros golyót húzunk}) = \\ &= \frac{\mathbb{P}(\text{két piros golyót húzunk} \mid 0 \text{ db hatost dobunk}) \cdot \mathbb{P}(0 \text{ db hatost dobunk})}{\mathbb{P}(\text{két piros golyót húzunk})}. \end{aligned}$$

Ha feltesszük, hogy 0 darab hatost dobtunk, az éppen azt jelenti, hogy egyetlen piros golyó sincs a dobozban, tehát ezen feltétel mellett 0 valószínűséggel húzunk két piros golyót:

$$\mathbb{P}(\text{két piros golyót húzunk} \mid 0 \text{ db hatost dobunk}) = 0,$$

és így a jobb oldal számlálója, tehát az egész tört eredménye is 0, vagyis

$$\mathbb{P}(0 \text{ db hatost dobunk} \mid \text{két piros golyót húzunk}) = 0.$$

A második valószínűséget is hasonlóképp számíthatjuk ki:

$$\begin{aligned} &\mathbb{P}(1 \text{ db hatost dobunk} \mid \text{két piros golyót húzunk}) = \\ &= \frac{\mathbb{P}(\text{két piros golyót húzunk} \mid 1 \text{ db hatost dobunk}) \cdot \mathbb{P}(1 \text{ db hatost dobunk})}{\mathbb{P}(\text{két piros golyót húzunk})}. \end{aligned}$$

Itt azonban a $\mathbb{P}(\text{két piros golyót húzunk})$ valószínűség kiszámolásához szükségünk lesz a teljes valószínűség tételére. Ehhez vezessük be a következő jelölést: legyen H_i ($i = 0, 1, 2$) az az esemény, hogy i darab hatost dobunk. Ekkor persze H_0 , H_1 és H_2 páronként egymást

kizáróak, továbbá egyikük mindig bekövetkezik, azaz ezek teljes eseményrendszert alkotnak. Legyen még P_2 az, hogy 2 piros golyót húzunk, ekkor tehát a teljes valószínűség tétele szerint

$$\mathbb{P}(P_2) = \mathbb{P}(P_2 | H_0)\mathbb{P}(H_0) + \mathbb{P}(P_2 | H_1)\mathbb{P}(H_1) + \mathbb{P}(P_2 | H_2)\mathbb{P}(H_2)$$

A fentiek szerint az első tag a jobb oldalon 0. A $\mathbb{P}(P_2 | H_1)$ valószínűségben a H_1 feltétel épp azt jelenti, hogy 1 piros és 1 sárga golyó van az urnában, tehát minden húzásnál $1/2$ valószínűséggel húzunk pirosat, és mivel a húzásaink függetlenek, így $\mathbb{P}(P_2 | H_1) = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$.

Továbbá pontosan 1 darab hatost $\frac{10}{36} = \frac{5}{18}$ valószínűséggel dobunk, tehát

$$\mathbb{P}(P_2 | H_1)\mathbb{P}(H_1) = \frac{1}{4} \cdot \frac{5}{18} = \frac{5}{72}.$$

A H_2 feltétel mellett kizárólag piros golyók vannak a dobozban, így $\mathbb{P}(P_2 | H_2) = 1$, továbbá $\mathbb{P}(H_2) = \frac{1}{36}$, vagyis

$$\mathbb{P}(P_2) = 0 + \frac{5}{72} + \frac{1}{36} = \frac{7}{72}.$$

Mindezek alapján

$$\mathbb{P}(1 \text{ db hatost dobunk} | \text{két piros golyót húzunk}) = \mathbb{P}(H_1 | P_2) = \frac{\mathbb{P}(P_2 | H_1)\mathbb{P}(H_1)}{\mathbb{P}(P_2)} = \frac{5/72}{7/72} = \frac{5}{7},$$

és

$$\mathbb{P}(2 \text{ db hatost dobunk} | \text{két piros golyót húzunk}) = \mathbb{P}(H_2 | P_2) = \frac{\mathbb{P}(P_2 | H_2)\mathbb{P}(H_2)}{\mathbb{P}(P_2)} = \frac{1/36}{7/72} = \frac{2}{7},$$

vagyis arra érdemes tippelni, hogy 1 darab hatost dobtunk, és ekkor $5/7$ a valószínűsége, hogy helyesen tippelünk.

2. fejezet

2.1.1. Induljunk ki a (2.5) formulából:

$$\mathbb{P}(X > k + n | X > k) = \frac{\mathbb{P}(X > k + n)}{\mathbb{P}(X > k)} = \frac{1 - \sum_{i=1}^{k+n} f_X(i)}{1 - \sum_{j=1}^k f_X(j)} = \frac{1 - \sum_{i=1}^{k+n} (1-p)^{i-1}p}{1 - \sum_{j=1}^k (1-p)^{j-1}p}.$$

Az utolsó tört számlálójában és nevezőjében lévő összegekben a p tényező kiemelhető, és ekkor 1 kezdőtagú, $(1-p)$ kvóciensű mértani sorozat első $k+n$ ill. első k tagját kell összegeznünk. Írjuk fel ezt először a nevezőre, használjuk a mértani sorozat összegképletét:

$$\begin{aligned} 1 - \sum_{j=1}^k (1-p)^{j-1}p &= 1 - p \cdot (1 + (1-p) + (1-p)^2 + \dots + (1-p)^{k-1}) \\ &= 1 - p \cdot \frac{1 - (1-p)^k}{1 - (1-p)} = 1 - p \cdot \frac{1 - (1-p)^k}{p} = (1-p)^k. \end{aligned}$$

A nevezőben k helyett $(k+n)$ -nel számolva az $(1-p)^{k+n}$ kifejezést kapjuk, és így

$$\mathbb{P}(X > k+n \mid X > k) = \frac{(1-p)^{k+n}}{(1-p)^k} = (1-p)^n.$$

Ugyanez az érvelés adja, hogy a jobb oldal itt éppen $\mathbb{P}(X > n)$.

2.1.3. Jelöljük a dobások kimenetelét rendezett párokkal, ekkor az eseménytér:

$$\Omega = \{(i; j) : 1 \leq i, j \leq 6\}$$

Ha X a dobott számok összege, akkor az $\{X \text{ páros}\}$ esemény azon kimenetelekből áll, melyekre a két dobás eredményének paritása megegyezik (mindkettő páros vagy mindkettő páratlan). Ilyen párokból $2 \cdot 3^2 = 18$ darab van, így

$$\mathbb{P}(X \text{ páros}) = \frac{|\{X \text{ páros}\}|}{|\Omega|} = \frac{18}{36} = \frac{1}{2}.$$

Ugyanezt megkaphatjuk a **2.1.9.** példa alapján is, ahol kiszámoltuk X eloszlását. A dobott számok összege pontosan akkor páros, ha a 2, 4, 6, 8, 10 vagy 12 értékek valamelyikét veszi fel, azaz

$$\begin{aligned} \mathbb{P}(X \text{ páros}) &= f_X(2) + f_X(4) + f_X(6) + f_X(8) + f_X(10) + f_X(12) \\ &= \frac{1}{36} + \frac{3}{36} + \frac{5}{36} + \frac{5}{36} + \frac{3}{36} + \frac{1}{36} = \frac{18}{36} = \frac{1}{2}. \end{aligned}$$

Ha Y a dobott számok szorzata, akkor

$$\{Y \leq 4\} = \{(1; 1), (1; 2), (2; 1), (1; 3), (3; 1), (1; 4), (4; 1), (2; 2)\}.$$

Tehát

$$\mathbb{P}(Y \leq 4) = \frac{|\{Y \leq 4\}|}{|\Omega|} = \frac{8}{36} = \frac{2}{9}.$$

Továbbá a két esemény metszete $\{(1; 1), (1; 3), (3; 1), (2; 2)\}$, ezért

$$\mathbb{P}(X \text{ páros és } Y \leq 4) = \frac{|\{X \text{ páros}\} \cap \{Y \leq 4\}|}{|\Omega|} = \frac{4}{36} = \frac{1}{9}.$$

Mivel

$$\mathbb{P}(X \text{ páros}) \cdot \mathbb{P}(Y \leq 4) = \frac{1}{2} \cdot \frac{2}{9} = \frac{1}{9} = \mathbb{P}(X \text{ páros és } Y \leq 4),$$

így a függetlenség definíciója szerint a két esemény független.

2.1.7. A sorsjegyekkel egymástól függetlenül $\frac{1}{10}$ valószínűséggel lehet nyerni, azaz egy sorsjegy $\frac{9}{10}$ valószínűséggel nem nyer. Így $(\frac{9}{10})^3$ annak a valószínűsége, hogy Béla egy adott héten nem nyer, mert a sorsjegyeket egymástól függetlenül választottuk ki. Tehát $p = 1 - (\frac{9}{10})^3$ annak a valószínűsége, hogy Béla egy adott héten nyer.

Legyen X azon hetek száma, amikor Béla játszik. Ekkor X egy valószínűségi változó. Mivel az első olyan hétig játszik, amikor nyer, és az egyes heteken a játékok egymástól függetlenek, így $X \sim Geo(p)$. A keresett valószínűség

$$\mathbb{P}(X \leq 3) = \mathbb{P}(X = 1) + \mathbb{P}(X = 2) + \mathbb{P}(X = 3) = f_X(1) + f_X(2) + f_X(3).$$

A súlyfüggvény képletét behelyettesítve ez a valószínűség

$$p + (1-p)p + (1-p)^2p = \left(1 - \left(\frac{9}{10}\right)^3\right) \left(1 + \left(\frac{9}{10}\right)^3 + \left(\frac{9}{10}\right)^6\right) = 1 - \left(\frac{9}{10}\right)^9 \approx 0,6126.$$

Egy második lehetséges megoldás a következő. A sorsjegyekkel egymástól függetlenül $\frac{1}{10}$ valószínűséggel lehet nyerni, azaz egy sorsjegy $\frac{9}{10}$ valószínűséggel nem nyer. Írjuk fel a keresett valószínűséget a komplementere segítségével:

$$\mathbb{P}(\text{Béla az első három hét valamelyikén nyer}) = 1 - \mathbb{P}(\text{Béla az első három héten nem nyer}).$$

Az utóbbi valószínűség éppen annak a valószínűsége, hogy 9 egymás után egymástól függetlenül választott sorsjegy nem nyer, azaz $\left(\frac{9}{10}\right)^9$, tehát a keresett valószínűség $1 - \left(\frac{9}{10}\right)^9$.

2.1.9. Legyen X a szükséges dobások száma abban az esetben, amikor 2 egyforma eredményre várunk, ekkor persze $X \geq 2$. Ha $X = k$, akkor az első $k - 1$ dobásból álló sorozatban fejek és irások váltakoznak, ilyen sorozatból pedig 2 darab van. Azaz

$$f_X(k) = \mathbb{P}(X = k) = \frac{2}{2^k} = \frac{1}{2^{k-1}}.$$

teljesül minden $k \geq 2$ -re.

Legyen Y a szükséges dobások száma abban az esetben, amikor 3 egyforma eredményre várunk. Ekkor $Y \geq 3$ mindenképp teljesül. Ha $Y = k$, akkor az első $k - 3$ elembe nincs sem FFF , sem III , továbbá $k > 3$ esetén az első $k - 3$ dobás egyértelműen meghatározza, hogyan végződik a sorozat. Tehát $Y = k$ és $k > 3$ esetén a lehetséges $F - I$ sorozatok száma megegyezik azon $k - 3$ hosszú sorozatok C_{k-3} számával, amikben nincs FFF vagy III , $Y = 3$ esetén pedig két lehetséges sorozat van.

Legyen A_n azon n hosszúságú sorozatok száma, amikben nincs FFF vagy III , és két egyforma elemmel végződnek, és legyen B_n azon n hosszúságú sorozatok száma, amikben nincs FFF vagy III , és nem igaz rájuk, hogy két egyforma elemmel végződnek. Ekkor $C_n = A_n + B_n$, $A_1 = 0$, $A_2 = 2$, $B_1 = 2$, $B_2 = 2$, $C_1 = 2$, $C_2 = 4$ továbbá $n > 2$ esetén $A_n = B_{n-1}$, $B_n = A_{n-1} + B_{n-1} = C_{n-1}$, vagyis $n > 2$ esetén

$$C_n = C_{n-1} + C_{n-2}.$$

Legyen most $C_0 = 2$, ekkor a fenti rekurzió $n = 2$ esetén is érvényes. Vegyük észre, hogy itt valójában $C_n = 2F_{n+1}$, ahol F_n jelöli a Fibonacci-sorozat n -edik tagját, és így

$$f_Y(k) = \mathbb{P}(Y = k) = \frac{C_{k-3}}{2^k} = \frac{F_{k-2}}{2^{k-1}}.$$

teljesül minden $k \geq 3$ -ra. Ezt az eloszlást *Fibonacci-eloszlásnak* is szokás nevezni. Vegyük észre, hogy a fentiekből következik a

$$\sum_{i=1}^{\infty} \frac{F_i}{2^{i+1}} = 1$$

összefüggés. Ezt persze egyszerűen belátható a rekurzív definíció alapján. Jelölje ugyanis a fenti összeget a . Ekkor könnyen látható, hogy $a < \infty$ (miért?), és így

$$\begin{aligned} a &= \frac{1}{2} + \frac{1}{4} + \sum_{i=3}^{\infty} \frac{F_i}{2^{i+1}} = \frac{1}{2} + \frac{1}{4} + \sum_{i=3}^{\infty} \frac{F_{i-1} + F_{i-2}}{2^{i+1}} \\ &= \frac{1}{4} + \frac{1}{8} + \frac{1}{2} \left(\sum_{i=1}^{\infty} \frac{F_i}{2^{i+1}} - \frac{1}{4} \right) + \frac{1}{4} \sum_{i=1}^{\infty} \frac{F_i}{2^{i+1}} \\ &= \frac{1}{4} + \frac{1}{8} + \frac{a}{2} - \frac{1}{8} + \frac{a}{4} = \frac{3a}{4} + \frac{1}{4}, \end{aligned}$$

amit átrendezve $a = 1$ adódik.

2.3.4. Legyen X a szükséges dobások száma abban az esetben, amikor 2 egyforma eredményre várunk. A 2.1.9. feladatban meghatároztuk az X eloszlását, amiből látható, hogy $X - 1$ geometriai eloszlású $\frac{1}{2}$ paraméterrel, így $\mathbb{E}(X) = \mathbb{E}(X - 1) + 1 = 2 + 1 = 3$.

Legyen Y a szükséges dobások száma abban az esetben, amikor 3 egyforma eredményre várunk. Ekkor $Y \geq 3$, és $k \geq 3$ esetén $f_Y(k) = \frac{F_{k-2}}{2^{k-1}}$, ahol F_n az n -edik Fibonacci-szám. Definíció szerint

$$\begin{aligned} \mathbb{E}(Y) &= \sum_{k=3}^{\infty} k \cdot \frac{F_{k-2}}{2^{k-1}} = \sum_{k=1}^{\infty} (k+2) \cdot \frac{F_k}{2^{k+1}} = \frac{3}{4} + \frac{1}{2} + \sum_{k=3}^{\infty} (k+2) \cdot \frac{F_{k-2} + F_{k-1}}{2^{k+1}} \\ &= \frac{5}{4} + \frac{1}{4} \sum_{k=3}^{\infty} k \cdot \frac{F_{k-2}}{2^{k-1}} + \frac{1}{2} \sum_{k=3}^{\infty} \frac{F_{k-2}}{2^{k-1}} + \sum_{k=3}^{\infty} ((k+1) + 1) \cdot \frac{F_{k-1}}{2^{k+1}} \\ &= \frac{5}{4} + \frac{1}{4} \mathbb{E}(Y) + \frac{1}{2} \sum_{k=3}^{\infty} f_Y(k) + \frac{1}{2} \sum_{k=4}^{\infty} (k+1) \cdot \frac{F_{k-2}}{2^{k-1}} \\ &= \frac{5}{4} + \frac{1}{4} \mathbb{E}(Y) + \frac{1}{2} + \frac{1}{2} \left(\sum_{k=3}^{\infty} (k+1) \cdot \frac{F_{k-2}}{2^{k-1}} - 1 \right) \\ &= \frac{5}{4} + \frac{1}{4} \mathbb{E}(Y) + \frac{1}{2} + \frac{1}{2} \mathbb{E}(Y) = \frac{7}{4} + \frac{3}{4} \mathbb{E}(Y), \end{aligned}$$

ezt átrendezve $\mathbb{E}(Y) = 7$ adódik.

Némi plusz számolás árán egyszerűsíthető (vagy akár ki is kerülhető) a rekurzív érvelés, ha az

$$F_n = \frac{1}{\sqrt{5}} \left(\left(\frac{1+\sqrt{5}}{2} \right)^n - \left(\frac{1-\sqrt{5}}{2} \right)^n \right)$$

explicit képletet használjuk, és így

$$\begin{aligned}\mathbb{E}(Y) &= \sum_{k=3}^{\infty} k \cdot \frac{F_{k-2}}{2^{k-1}} = \frac{1}{2\sqrt{5}} \sum_{k=1}^{\infty} (k+2) \left(\left(\frac{1+\sqrt{5}}{4} \right)^k - \left(\frac{1-\sqrt{5}}{4} \right)^k \right) \\ &= \frac{1+\sqrt{5}}{4\sqrt{5}} \sum_{k=0}^{\infty} \left(\frac{1+\sqrt{5}}{4} \right)^k - \frac{1-\sqrt{5}}{4\sqrt{5}} \sum_{k=0}^{\infty} \left(\frac{1-\sqrt{5}}{4} \right)^k + \\ &\quad + \frac{1}{2\sqrt{5}} \sum_{k=1}^{\infty} k \left(\left(\frac{1+\sqrt{5}}{4} \right)^k - \left(\frac{1-\sqrt{5}}{4} \right)^k \right).\end{aligned}$$

Az első két végtelen összeg egy mértani sorösszeg, így a fenti kifejezés első két tagja

$$\begin{aligned}\frac{1+\sqrt{5}}{4\sqrt{5}} \cdot \frac{1}{1-\frac{1+\sqrt{5}}{4}} - \frac{1-\sqrt{5}}{4\sqrt{5}} \cdot \frac{1}{1-\frac{1-\sqrt{5}}{4}} &= \frac{1+\sqrt{5}}{\sqrt{5}(3-\sqrt{5})} - \frac{1-\sqrt{5}}{\sqrt{5}(3+\sqrt{5})} \\ &= \frac{(1+\sqrt{5})(3+\sqrt{5}) - (1-\sqrt{5})(3-\sqrt{5})}{\sqrt{5}(3+\sqrt{5})(3-\sqrt{5})} = \frac{8\sqrt{5}}{4\sqrt{5}} = 2.\end{aligned}$$

A harmadik összeget is kettébontjuk, ehhez vezessük be a $p = \frac{1+\sqrt{5}}{4}$ és $q = \frac{1-\sqrt{5}}{4}$ jelöléseket, ekkor az említett összeg

$$\frac{1}{2\sqrt{5}} \left(\sum_{k=1}^{\infty} kp^k - \sum_{k=1}^{\infty} kq^k \right)$$

alakba írható. A két tagot külön kezeljük:

$$\sum_{k=1}^{\infty} kp^k = p \sum_{k=0}^{\infty} (k+1)p^k = p \sum_{k=0}^{\infty} p^k + p \sum_{k=0}^{\infty} kp^k = \frac{p}{1-p} + p \sum_{k=1}^{\infty} kp^k,$$

vagyis

$$\sum_{k=1}^{\infty} kp^k = \frac{p}{(1-p)^2} = \frac{1+\sqrt{5}}{4} \cdot \frac{16}{(3-\sqrt{5})^2} = \frac{4(1+\sqrt{5})}{(3-\sqrt{5})^2}$$

Ugyanígy

$$\sum_{k=1}^{\infty} kq^k = \frac{q}{(1-q)^2} = \frac{4(1-\sqrt{5})}{(3+\sqrt{5})^2}.$$

Ezek különbsége

$$4 \cdot \left[\frac{(1+\sqrt{5})(14+6\sqrt{5}) - (1-\sqrt{5})(14-6\sqrt{5})}{(3-\sqrt{5})^2(3+\sqrt{5})^2} \right] = 4 \cdot \frac{40\sqrt{5}}{4^2} = 10\sqrt{5}.$$

Vagyis $\mathbb{E}(X) = 2 + \frac{1}{2\sqrt{5}} \cdot 10\sqrt{5} = 7$.

A $\sum_{k=1}^{\infty} kp^k$ és $\sum_{k=1}^{\infty} kq^k$ összegeket ki lehet számolni hatványsorok segítségével is. Ugyanis $|x| < 1$ esetén

$$\sum_{k=0}^{\infty} x^k = \frac{1}{1-x},$$

ezt pedig x szerint deriválva

$$\sum_{k=1}^{\infty} k \cdot x^{k-1} = \frac{1}{(1-x)^2}$$

adódik. Az egyenletet x -szel szorozva az $x = p$ ill. $x = q$ helyettesítéssel adódik az eredmény.

2.3.13. Ha $X \sim \text{Bin}(n; p)$, akkor X eloszlása megegyezik egy $\mathbb{1}_{A_1} + \dots + \mathbb{1}_{A_n}$ összeg eloszlásával, ahol A_1, \dots, A_n együttesen független, p valószínűségű események, $\mathbb{1}_{A_1}, \dots, \mathbb{1}_{A_n}$ pedig a hozzájuk tartozó indikátorváltozók. Így $\mathbb{E}(X^2)$ azonos az utóbbi összeg négyzetének várható értékével. Ennek meghatározásához jegyezzük meg, hogy

$$\begin{aligned} (\mathbb{1}_{A_1} + \dots + \mathbb{1}_{A_n})^2 &= \mathbb{1}_{A_1}^2 + \dots + \mathbb{1}_{A_n}^2 + 2 \cdot (\mathbb{1}_{A_1} \cdot \mathbb{1}_{A_2} + \mathbb{1}_{A_1} \cdot \mathbb{1}_{A_3} + \dots + \mathbb{1}_{A_{n-1}} \cdot \mathbb{1}_{A_n}) \\ &= \mathbb{1}_{A_1} + \dots + \mathbb{1}_{A_n} + 2 \cdot (\mathbb{1}_{A_1 \cap A_2} + \mathbb{1}_{A_1 \cap A_3} + \dots + \mathbb{1}_{A_{n-1} \cap A_n}), \end{aligned}$$

hiszen egy $\mathbb{1}_{A_i} \cdot \mathbb{1}_{A_j}$ szorzat $i \neq j$ esetén pontosan akkor 1, ha mind az A_i , mind az A_j esemény bekövetkezik (egyébként pedig a szorzat 0). Tehát a várható érték linearitása és az A_i események függetlensége miatt

$$\begin{aligned} \mathbb{E}((\mathbb{1}_{A_1} + \dots + \mathbb{1}_{A_n})^2) &= \mathbb{E}(\mathbb{1}_{A_1}) + \dots + \mathbb{E}(\mathbb{1}_{A_n}) + 2 \cdot (\mathbb{E}(\mathbb{1}_{A_1 \cap A_2}) + \dots + \mathbb{E}(\mathbb{1}_{A_{n-1} \cap A_n})) \\ &= np + 2 \cdot (\mathbb{P}(A_1 \cap A_2) + \mathbb{P}(A_1 \cap A_3) + \dots + \mathbb{P}(A_{n-1} \cap A_n)) \\ &= np + 2 \cdot (\mathbb{P}(A_1)\mathbb{P}(A_2) + \dots + \mathbb{P}(A_{n-1})\mathbb{P}(A_n)) \\ &= np + 2 \binom{n}{2} p^2 = np + n(n-1)p^2 = np + n^2 p^2 - np^2, \end{aligned}$$

Határozzuk meg ezt a várható értéket a transzformált várható értékére vonatkozó formula segítségével is:

$$\mathbb{E}(X^2) = \sum_{k=0}^n k^2 \binom{n}{k} p^k (1-p)^{n-k} = \sum_{k=1}^n k^2 \binom{n}{k} p^k (1-p)^{n-k}.$$

Itt $1 \leq k \leq n$ esetén

$$k^2 \binom{n}{k} = k^2 \cdot \frac{n!}{k!(n-k)!} = kn \cdot \frac{(n-1)!}{(k-1)!(n-1-(k-1))!} = kn \binom{n-1}{k-1},$$

tehát

$$\begin{aligned} \mathbb{E}(X^2) &= \sum_{k=1}^n kn \binom{n-1}{k-1} p^k (1-p)^{n-1-(k-1)} = np \sum_{j=0}^{n-1} (j+1) \binom{n-1}{j} p^j (1-p)^{n-1-j} \\ &= np \left(\sum_{j=0}^{n-1} j \binom{n-1}{j} p^j (1-p)^{n-1-j} + \sum_{j=0}^{n-1} \binom{n-1}{j} p^j (1-p)^{n-1-j} \right) \\ &= np((n-1)p + 1) = n^2 p^2 - np^2 + np \end{aligned}$$

hiszen az utolsó előtti szumma egy $\text{Bin}(n-1; p)$ eloszlású változó várható értékét adja.

2.5.1. A **2.5.2** példában meghatároztuk az X generátorfüggvényét és annak deriváltját, ez utóbbi

$$P'_X(z) = \frac{p(1 - (1-p)z) + pz(1-p)}{(1 - (1-p)z)^2} = \frac{p}{(1 - (1-p)z)^2},$$

így

$$P_X''(z) = \frac{2p(1-p)}{(1-(1-p)z)^3},$$

vagyis a 2.5.2. következmény szerint

$$\begin{aligned} \mathbb{D}^2(X) &= P_X''(1) + P_X'(1) - P_X'^2(1) = \frac{2p-2p^2}{p^3} + \frac{1}{p} - \frac{1}{p^2} \\ &= \frac{2p-2p^2+p^2-p}{p^3} = \frac{p-p^2}{p^3} = \frac{1-p}{p^2}. \end{aligned}$$

2.5.4. Ebben a megoldásban a 2.1.9. feladat megoldásának jelöléseit használjuk (de a súlyfüggvény értékeit nem fogjuk felhasználni). Legyen X a szükséges dobások száma. Az eloszlás súlyfüggvényére az ott látottaknak megfelelően

$$f_X(k) = \mathbb{P}(X = k) = \frac{C_{k-3}}{2^k}$$

teljesül minden $3 \leq k$ -ra, ahol $C_0 = C_1 = 2$, és minden $2 \leq n$ -re

$$C_n = C_{n-1} + C_{n-2}.$$

Ebből fakadóan pedig $5 \leq k$ esetén $f_X(k) = \frac{1}{2}f_X(k-1) + \frac{1}{4}f_X(k-2)$ érvényes.

Ezért az X

$$P_X(z) = \sum_{k=3}^{\infty} f_X(k)z^k,$$

generátorfüggvényére

$$\begin{aligned} P_X(z) &= \frac{2z^3 + z^4}{2^3} + \sum_{k=5}^{\infty} f_X(k)z^k = \frac{2z^3 + z^4}{2^3} + \sum_{k=5}^{\infty} \left(\frac{1}{2}f_X(k-1) + \frac{1}{4}f_X(k-2) \right) z^k \\ &= \frac{2z^3 + z^4}{2^3} - \frac{z^4}{2^3} + \frac{z}{2} \sum_{k=3}^{\infty} f_X(k)z^k + \frac{z^2}{4} \sum_{k=3}^{\infty} f_X(k)z^k = \frac{z^3}{2^2} + \frac{z}{2}P_X(z) + \frac{z^2}{4}P_X(z). \end{aligned}$$

Innen

$$P_X(z) \left(1 - \frac{z}{2} - \frac{z^2}{4} \right) = \frac{z^3}{4}, \quad P_X(z) = \frac{z^3}{4-2z-z^2}, \quad P_X'(z) = \frac{3z^2}{(4-2z-z^2)^2} + \frac{2z^3(z+1)}{(4-2z-z^2)^2},$$

a várható érték tehát $\mathbb{E}(X) = P'(1) = 7$. Továbbá

$$\begin{aligned} P_X''(z) &= \frac{6z(4-2z-z^2) + 6z^2(z+1)}{(4-2z-z^2)^2} + \\ &\quad + \frac{(8z^3 + 6z^2)(4-2z-z^2)^2 + 4(z+1)(4-2z-z^2)(2z^4 + 2z^3)}{(4-2z-z^2)^4}, \end{aligned}$$

azaz $\mathbb{D}^2(X) = P_X''(1) + P_X'(1) - P_X'^2(1) = 64 + 7 - 49 = 22$.

2.6.2. A 2.5.4. tétel (i) \implies (iii) irányának bizonyítása valójában azt adja, hogy ha $X_n \xrightarrow{d} Y$,

akkor a P_{X_n} generátorfüggvények az egész komplex egységkörtön egyenletesen konvergálnak a P_Y generátorfüggvényhez. Mivel a nyílt egységkörtön holomorf függvényekről van szó, így következésképp a P'_{X_n} ill. P''_{X_n} függvények az egész nyílt egységkörtön pontonként (és annak minden kompakt részhalmazán egyenletesen) konvergálnak a P'_Y ill. P''_Y függvényekhez. Ha most a $P'_{X_n}(1)$ ill. $P''_{X_n}(1)$ számsorozatok végesek és konvergensek (vagy ami ezzel ekvivalens: az X_n változók várható értékei és szórásnégyzetei végesek és konvergensek), akkor korlátosak is, és így a $P'_{X_n}(z)$ ill. $P''_{X_n}(z)$ függvényekre n -től független korlát adható az egész egységkörtön. Ekkor viszont $P'_Y(1)$ és $P''_Y(1)$ is véges kell legyen, máskülönben ezek az 1 közelében tetszőlegesen nagy értéket felvennének, és így ez igaz lenne P'_{X_n} -re ill. P''_{X_n} -re is nagy n esetén.

A 2.5.4. tétel (i) \implies (iii) irányának bizonyításában látottakhoz nagyon hasonlóan bizonyítható az is, hogy ha $X_n \xrightarrow{d} Y$, továbbá a P_{X_n} és P_Y generátorfüggvényekre a $P'_{X_n}(1)$, $P'_Y(1)$, $P''_{X_n}(1)$ és $P''_Y(1)$ értékek végesek minden n -re, akkor $P'_{X_n}(1) \rightarrow P'_Y(1)$ és $P''_{X_n}(1) \rightarrow P''_Y(1)$, tehát a 2.5.2. következmény alapján a határeloszlás várható értékét ill. szórásnégyzetét a X_n változók várható értékének ill. szórásnégyzetének határértékeként kapjuk.

3. fejezet

3.2.3. Két szám független választása a $[-1; 1]$ intervallumon valójában ugyanaz, mint egy $P = (x; y)$ pont választása a $\Omega = [-1; 1] \times [-1; 1]$ négyzeten. Az első kérdés megválaszolásához az $x + y > 0$ feltétel által megadott síkrész Ω -ba eső részének területét és a négyzet területét kell elosztanunk egymással, mert (geometriai valószínűségi mezőről lévén szó) az esemény valószínűsége $T_{\text{kedvező}}/T_{\text{összes}}$.

A megadott feltétel azzal ekvivalens, hogy $y > -x$, ez pedig Ω azon pontjaira teljesül, amik az $y = -x$ egyenes felett helyezkednek el, azaz az $A = (-1; 1)$, $B = (1; 1)$ és $C = (1; -1)$ pontok által meghatározott ABC derékszögű háromszög pontjaira (leszámítva az átfogót a fenti feltételben szereplő szigorú egyenlőség miatt). Tehát annak a valószínűsége, hogy a választott pontra $x + y > 0$ teljesül, megegyezik a háromszög területének és a $[-1; 1] \times [-1; 1]$ négyzet területének arányával. A háromszög területe 2, a négyzet területe 4, azaz a keresett valószínűség $1/2$.

Meghatározzuk az $X(P) = x + y$ változó eloszlásfüggvényét. Mivel mindkét választott szám legalább -1 , így az összegük nem lehet (-2) -nél kisebb, tehát $F_X(t) = P(X < t) = 0$, ha $t \leq -2$. Hasonlóképp, mindkét szám legfeljebb 1, így az összegük legfeljebb 2 lehet, tehát $F_X(t) = P(X < t) = 1$, ha $t > 2$.

Az $F_X(t) = P(X < t)$ érték annak a valószínűsége, hogy a választott $(x; y)$ pontra $x + y < t$, azaz $y < -x + t$ teljesül, tehát hogy a pont az $y = -x + t$ egyenes alatt (és az Ω négyzetben) helyezkedik el. Ha $t \in (-2; 0]$, akkor ez pontosan az A , C és $D = (-1; -1)$ pontok által meghatározott ACD derékszögű háromszögnek az $y = -x + t$ egyenes alá eső része, amely szintén egy derékszögű háromszög, melynek mindkét befogója $2 + t$ hosszúságú (itt $t \leq 0$). Ennek területe $\frac{(2+t)^2}{2}$, a $\mathbb{P}(X < t)$ valószínűség pedig ezen terület és a négyzet területének aránya. Vagyis ha $t \in (-2; 0]$, akkor $F_X(t) = \frac{(2+t)^2}{8}$.

Ha viszont $t \in (0; 2]$, akkor $X < t$ pontosan azokra a pontokra teljesül, amelyek az ACD háromszögben, vagy pedig az ABC háromszögben, de az $y = -x + t$ egyenes alatt helyezkednek el. Ez utóbbi síkrész területe megkapható, ha az ABC háromszög területéből levonjuk az $y = -x + t$ egyenes fölé eső rész területét, amely szintén egy derékszögű háromszög

$2 - t$ befogókkal. Tehát az $X < t$ feltétel által meghatározott síkrész területe

$$T(ACD\Delta) + T(ABC\Delta) - \frac{(2-t)^2}{2} = 4 - \frac{(2-t)^2}{2}.$$

Leosztva a négyzet területével $F_X(t) = 1 - \frac{(2-t)^2}{8}$. Összefoglalva tehát

$$F_X(t) = \begin{cases} 0, & \text{ha } t \leq -2, \\ \frac{(2+t)^2}{8}, & \text{ha } -2 < t \leq 0, \\ 1 - \frac{(2-t)^2}{8}, & \text{ha } 0 < t \leq 2, \\ 1, & \text{ha } t > 2. \end{cases}$$

3.3.2. Az X változó azonos a **3.2.3.** feladatban definiálttal, ahol meghatároztuk az F_X eloszlásfüggvényt (lásd a **3.2.3.** feladat megoldását). Mivel $F_X(t)$ véges sok pont kivételével mindenhol deriválható, így ebben az esetben a sűrűségfüggvényt megkaphatjuk az eloszlásfüggvényből, ha azt deriváljuk azokban a pontokban, ahol lehet. A maradék véges sok pontban a sűrűségfüggvény értéke tetszőlegesen megválasztható (pl. nullának). Ezért

$$f_X(t) = \begin{cases} \frac{2+t}{4}, & \text{ha } t \in (-2; 0), \\ \frac{2-t}{4}, & \text{ha } t \in (0; 2), \\ 0 & \text{különben} \end{cases}$$

sűrűségfüggvénye X -nek. (Valójában a derivált a $t = 0$ pontban is létezik, és a sűrűségfüggvény definíciójában az első két sor valamelyikében a 0 pontot is hozzávehetjük az intervallumhoz, de természetesen az is megfelelő, ha ott pl. 0 -nak (vagy bármi másnak) definiáljuk $f_X(t)$ -t.)

Az X várható értéke definíció szerint

$$\mathbb{E}(X) = \int_{-\infty}^{\infty} t f_X(t) dt = \frac{1}{4} \int_{-2}^0 (2t + t^2) dt + \frac{1}{4} \int_0^2 (2t - t^2) dt.$$

A számolást innen a Newton–Leibniz-formula segítségével nem nehéz elvégezni, azonban egy egyszerű észrevétel lényegesen leegyszerűsíti azt. Vegyük észre ugyanis, hogy az $f_X(t)$ függvény páros, azaz $f_X(t) = f_X(-t)$ teljesül minden t -re. Következésképp a $g(t) = t f_X(t)$ függvény páratlan, hiszen $g(-t) = -t f_X(-t) = -t f_X(t) = -g(t)$. Ha pedig egy páratlan függvényt egy 0 -ra szimmetrikus intervallumon integrálunk, akkor az eredmény szükségképp 0 , vagyis $\mathbb{E}(X) = 0$.

Mivel a várható érték 0 , a szórásnégyzet meg fog egyezni az X^2 várható értékével, ami a transzformált várható értékére vonatkozó formula szerint

$$\begin{aligned} \mathbb{E}(X^2) &= \int_{-\infty}^{\infty} t^2 f_X(t) dt = \frac{1}{4} \int_{-2}^0 (2t^2 + t^3) dt + \frac{1}{4} \int_0^2 (2t^2 - t^3) dt \\ &= \frac{1}{4} \left[\frac{2t^3}{3} + \frac{t^4}{4} \right]_{-2}^0 + \frac{1}{4} \left[\frac{2t^3}{3} - \frac{t^4}{4} \right]_0^2 = \frac{4}{3} - 1 + \frac{4}{3} - 1 = \frac{8}{3} - 2 = \frac{2}{3}, \end{aligned}$$

vagyis $\mathbb{D}(X) = \sqrt{\frac{2}{3}}$.

Megjegyezzük, hogy ezt a számolást kikerülhetjük, hogy ha felírjuk az X változót $Y + Z$ alakban, ahol Y az első, míg Z a második választott szám. Mivel ezek függetlenek és azonos eloszlásúak, így

$$\mathbb{E}(X) = \mathbb{E}(Y + Z) = \mathbb{E}(Y) + \mathbb{E}(Z) = 2\mathbb{E}(Y),$$

továbbá

$$\mathbb{D}^2(X) = \mathbb{D}^2(Y + Z) = \mathbb{D}^2(Y) + \mathbb{D}^2(Z) = 2\mathbb{D}^2(Y).$$

A 3.4.1. szakaszban belátjuk, hogy Y és Z ún. egyenletes eloszlású valószínűségi változók a $[-1; 1]$ intervallumon, és kiszámoljuk a várható értéküket és szórásnégyzetüket, eszerint pedig $\mathbb{E}(Y) = \mathbb{E}(Z) = 0$ és $\mathbb{D}^2(Y) = \mathbb{D}^2(Z) = \frac{1}{3}$, ahonnan a feladat eredménye már következik.

3.4.9 Legyen $X \sim N(0; 1)$, ekkor az $1 - \Phi(t)$ kifejezés éppen a $\mathbb{P}(X \geq t)$ valószínűséget adja, ez pedig a sztenderd normális eloszlás sűrűségfüggvényével kifejezve

$$1 - \Phi(t) = \int_t^\infty \varphi(s) ds.$$

Mivel $t > 0$, így $s \geq t$ esetén $\frac{s}{t} \geq 1$, ezért (a $\varphi(s) > 0$ egyenlőtlenséget is kihasználva) az integrált függvényt pontonként felülről becslve egy felső becslést kapunk az előbbi integrálra:

$$\int_t^\infty \varphi(s) ds \leq \int_t^\infty \frac{s}{t} \varphi(s) ds = \frac{1}{t} \int_t^\infty \frac{1}{\sqrt{2\pi}} s e^{-s^2/2} ds.$$

Itt a jobb oldalon integrált függvénynek egyszerűen meghatározható egy primitív függvénye, így a Newton–Leibniz-formula alkalmazható:

$$\frac{1}{t} \int_t^\infty \frac{1}{\sqrt{2\pi}} s e^{-s^2/2} ds = \frac{1}{t} \left[-\frac{1}{\sqrt{2\pi}} e^{-s^2/2} \right]_t^\infty = \frac{1}{\sqrt{2\pi}t} e^{-t^2/2} = \frac{\varphi(t)}{t},$$

ezzel a felső becslést beláttuk.

Az alsó becsléshez tekintsük a két oldal különbségét:

$$g(t) = 1 - \Phi(t) - \frac{t\varphi(t)}{t^2 + 1},$$

azt kell belátnunk, hogy $g(t) > 0$ teljesül minden $t > 0$ esetén. Először is vegyük észre, hogy $g(0) = \frac{1}{2} > 0$ teljesül. Belátjuk, hogy $g'(t) < 0$ teljesül minden $t \in \mathbb{R}$ esetén, tehát a $g(t)$ függvény szigorúan monoton csökkenő. Valóban,

$$g'(t) = -\Phi'(t) - \frac{(\varphi(t) + t\varphi'(t))(t^2 + 1) - 2t^2\varphi(t)}{(t^2 + 1)^2} = -\varphi(t) - \frac{\varphi(t)(1 - t^2)}{(t^2 + 1)^2} - \frac{t\varphi'(t)(t^2 + 1)}{(t^2 + 1)^2},$$

és $\varphi'(t) = -\frac{1}{\sqrt{2\pi}} t e^{-t^2/2} = -t\varphi(t)$ miatt a fenti kifejezés

$$-\frac{\varphi(t)((t^2 + 1)^2 + 1 - t^2)}{(t^2 + 1)^2} + \frac{\varphi(t)t^2(t^2 + 1)}{(t^2 + 1)^2} = \frac{\varphi(t)}{(t^2 + 1)} = \frac{-2\varphi(t)}{(t^2 + 1)^2},$$

ami negatív minden valós t -re. Továbbá

$$\lim_{t \rightarrow \infty} g(t) = \lim_{t \rightarrow \infty} (1 - \Phi(t)) - \lim_{t \rightarrow \infty} \frac{t}{t^2 + 1} \cdot \varphi(t) = 0,$$

hiszen $\Phi(t)$ egy eloszlásfüggvény, a második limeszben pedig a szorzat mindkét tagja 0-hoz tart. Ekkor viszont $g(t)$ csak pozitív értéket vehet fel, ugyanis ha valamely $t_0 \in \mathbb{R}$ esetén $g(t_0) \leq 0$ teljesülne, akkor a szigorúan monoton csökkenés miatt $g(t_1) < 0$ teljesül bármely $t_1 > t_0$ esetén, és ekkor minden $t \geq t_1$ -re $g(t) < g(t_1) < 0$, ez viszont ellentmond annak, hogy a végtelenben a g határértéke 0. Ezzel az alsó becslést is beláttuk.

Irodalomjegyzék

- [1] Bolla Marianna, Krámlí András, *Statisztikai következtetések elmélete*, második kiadás, Typotex (2012)
- [2] William Feller, *Bevezetés a valószínűségszámításba és alkalmazásaiba*, Műszaki Könyvkiadó (1978)
- [3] Freud Róbert, Gyarmati Edit, *Számelmélet*, Nemzeti Tankönyvkiadó, Budapest (2000, 2006)
- [4] Achim Klenke, *Probability theory: a comprehensive course*, Springer Science & Business Media (2013)
- [5] Mészáros Szabolcs, *Valószínűségszámítás*, BME VIK elektronikus előadásjegyzet (2021), letölthető a https://cs.bme.hu/bvalszam/jegyzet_msz.pdf linkről
- [6] M. Mitzenmacher, E. Upfal, *Probability and Computing*, Cambridge Univ. Press (2005)
- [7] Prékopa András, *Valószínűségelmélet*, Műszaki könyvkiadó (1972)
- [8] Rónyai Lajos, Ivanyos Gábor, Szabó Réka, *Algoritmusok*, Typotex kiadó (1999)
- [9] Walter Rudin *Principles of mathematical analysis*, 3rd edition, McGraw-Hill, New York (1976)
- [10] Szeszlér Dávid, *Bevezetés a számításelméletbe 1*, BME VIK elektronikus előadásjegyzet (2014), letölthető a https://cs.bme.hu/bsz1/jegyzet/bsz1_jegyzet.pdf linkről

Tárgymutató

- örökifjú eloszlás
 - diszkrét, 49
 - folytonos, 128
- összegzési függvény, 53
- általánosított binomiális tétel, 104
- őskép, 45

- aszimptotikusan torzítatlan becslés, 149

- Bayes-tétel, 41
- Bertrand-paradoxon, 106
- binomiális együttható, 19
- binomiális eloszlás, 47
- binomiális tétel, 19
- biztos esemény, 8

- centrális határeloszlás tétele, 140

- de Moivre–Laplace-tétel, 139
- de Morgan-azonosságok, 10
- Descartes-szorzat, 15

- egyenletes eloszlás, 126
- egymintás t -próba, 156
- egymintás u -próba, 155
- egymást kizáró események, 10
- egyszerű valószínűségi változó, 44
- együttes eloszlás, 58
- együttes eloszlásfüggvény, 116
- együttesen független események, 35
- elemi esemény, 7
- elfogadási tartomány, 154
- ellenhipotézis, 154
- eloszlásfüggvény, 112
- elsőfajú hiba, 154
- esemény, 8
- eseménytér, 7
- Euler-féle χ^2 függvény, 53
- exponenciális eloszlás, 127

- feltételes valószínűség, 30
- feltételes várható érték, 72
- Fibonacci-eloszlás, 171
- független események, 32
- független valószínűségi változók
 - diszkrét eset, 60
 - általános eset, 116

- generátorfüggvény, 90
- geometriai eloszlás, 48
- geometriai valószínűségi mezők, 109
- Glivenko–Cantelli-tétel, 146

- háttéreloszlás, 143

- indikátorváltozó, 44

- kimenetel, 7
- klasszikus valószínűségi mező, 12
- kombináció, 18
- konfidenciaintervallum, 149
- konvolúció, 63
- korrigált tapasztalati szórás, 149
- korrigált tapasztalati szórásnégyzet, 149
- kritikus tartomány, 154

- Las Vegas típusú algoritmus, 77
- lehetetlen esemény, 8

- marginális eloszlás, 58
- Markov-egyenlőtlenség, 71
- minta realizációja, 143
- mintaátlag, 147
- Monte Carlo típusú algoritmus, 77
- másodfajú hiba, 154

- nagy számok erős törvénye, 135
- nagy számok törvénye, 11
- negatív binomiális eloszlás, 103
- normális eloszlás, 131
- nullhipotézis, 153

paraméterter, 143
 permutáció
 ismétlés nélküli, 17
 ismétléses, 21
 Poincaré-formula, 13
 Poisson-eloszlás, 97

 rendezett minta, 144
 reprezentatív minta, 144

 statisztika, 147
 statisztika alaptétele, 146
 statisztikai mező, 143
 paraméteres, 143
 statisztikai minta, 143
 statisztikai próba, 154
 statisztikai sokaság, 142
 szignifikanciaszint, 154
 szita-formula, 13
 szorzatmérték, 55
 szorzási szabály, 37
 sztenderd normális eloszlás, 129
 sztenderdizált, 137
 szórás, 124
 diszkrét esetben, 87
 szórásnégyzet, 124
 diszkrét esetben, 87
 súlyfüggvény, 46
 sűrűségfüggvény, 120

 tapasztalati eloszlásfüggvény, 144
 tapasztalati szórás, 148
 tapasztalati szórásnégyzet, 148
 teljes eseményrendszer, 39
 teljes valószínűség tétele, 39
 teljes várható érték tétele, 75
 torzítatlan becslés, 147
 transzformált várható értéke
 diszkrét esetben, 82
 folytonos esetben, 123

 valószínűség, 12
 valószínűségi mező, 24
 valószínűségi mérték, 24
 valószínűségi vektorváltozó, 57
 valószínűségi változó, 112
 abszolút folytonos, 120
 diszkrét, 44
 valószínűségi változó eloszlása, 46
 variancia, 124
 diszkrét esetben, 87
 variáció, 18
 várható érték
 diszkrét változóé, 66
 folytonos változóé, 122
 várható érték linearitása
 diszkrét esetben, 67
 folytonos esetben, 123