Valószínűségszámítás jegyzet 2. rész

1. Valószínűségi változó, eloszlásfüggvény, diszkrét és folytonos eset. Az eloszlásfüggvény négy tulajdonsága. Intervallumok valószínűségei. Diszkrét eloszlás. Sűrűségfüggvény.

A kísérletek pontosabb megértése, vizsgálata során a valószínűségek számítása után felmerülnek ennél bonyolultabb kérdések, amelyek a kísérlet kimeneteleinek egyfajta kiértékelésére vonatkoznak. Ezeknek a vizsgálatára új fogalmat kell bevezetnünk:

1. Definíció. *Egy függvényt, ami a kísérlet kimeneteleihez valós számokat rendel,* **valószínűségi változónak** *nevezünk, amennyiben bármilyen valós x értékre az az esemény, hogy ennek a függvénynek az értéke kisebb lesz, mint x, egy megfigyelhető esemény. Formálisan leírva $X : \Omega \to \mathbf{R}$ akkor valószínűségi változó, ha minden x-re $\{X < x\}$ esemény megfigyelhető. (Az $\{X < x\}$ eseményt sokszor nívóhalmaznak is nevezik, és $\{\omega \in \Omega | X(\omega) < x\}$ módon formalizálják.)*

Tehát definiálható olyan függvény a kimeneteken, ami nem lesz valószínűségi változó. Ennek ellenére az általunk vizsgált függvények mind rendelkeznek majd ezzel a tulajdonsággal (valós életből származó egyszerű példán csak ilyen merül fel). A valószínűségi változó tehát számszerűsíti az eseményteret. A valószínűségi változó értékeinek a segítségével fejezünk ki eseményeket. (Az így kifejezhető események mentén egy új eseményteret is definiálhatunk, ami matematikai modellje lesz a kísérletnek.)

A kockadobás esetén legegyszerűbb módon úgy definiálhatunk egy X valószínűségi változót, hogy az értékének a dobás kimenetelét adjuk. Ezen felül persze rengeteg különböző valószínűségi változó értelmezhető ugyanezen a kísérleten. Például: Y = páros kimenet esetén 1, egyébként 0; Z = prím dobás esetén 25 egyébként a dobott érték négyzete.

Kevésbé egyértelműek az olyan esetek, amikor a kísérletnek nincs köze számokhoz. Például ha egy napon való esőzéshez rendelek értékeket: eső esetén 0, felhős égbolt esetén 100, napsütés esetén 500. Ha azonban úgy fogalmazom meg, hogy ez a valószínűségi változó a napelemem által termelt energiamennyiséget adja meg, máris látszik, hogy ennek a valószínűségi változónak is van értelme.

2. Definíció. *Egy X valószínűségi változónak az* **eloszlásfüggvénye** $F_X(x) = P(X < x)$

A szabályos kockadobás esetén az előző példában definiált X valószínűségi változónak adjuk meg az eloszlásfüggvényét! Könnyen látható, hogy az $1, 2, 3, 4, 5$ és 6 értékeknél lesz csak változás, ezeken a pontokon kívül nem. Az egyes intervallumokra a megfelelő valószínűségeket kiszámolva megkapjuk, hogy $F_X(x) = 0$, ha $x \leq 1$, $F_X(x) = i/6$, ha $i < x \leq i+1$ és $1 < x \leq 6$, végül pedig $F_X(x) = 1$, ha $x > 6$.

3. Tulajdonság(ok). *Minden F_X eloszlásfüggvény rendelkezik az alábbi tulajdonságokkal. Ami ennél is érdekesebb, hogy tetszőleges függvény, ami rendelkezik ezekkel a tulajdonságokkal, az eloszlásfüggvény (azaz lehet hozzá megfelelő kísérletet és valószínűségi változót definiálni).*

1. *monoton növekvő:* $\forall x_1 < x_2 : F_X(x_1) \leq F_X(x_2)$
2. *balról folytonos:* $\forall x_0 \in \mathbf{R} : lim_{x \to x_0 -} F_X(x) = F(x_0)$
3. $lim_{x \to -\infty} F_X(x) = 0$
4. $lim_{x \to \infty} F_X(x) = 1$

Könnyen látható, hogy az előző példában kapott F_X teljesíti a tulajdonságokat. Azt is vegyük észre, hogy F_X jobbról nem folytonos.

4. Tétel. *Balról zárt, jobbról nyílt intervallum valószínűsége számítható az eloszlásfüggvényből az alábbi képlettel (feltéve, hogy $x < y$ valós számok):*

$$P(X \in [x, y)) = F_X(y) - F_X(x)$$

Bizonyítás: Ha a jobb oldalba behelyettesítjük a definíciót, akkor a $P(X < y) - P(X < x)$ kifejezést kapjuk. Itt azonban a második esemény része az elsőnek, így a valószínűségek különbsége éppen az események különbségének a valószínűsége, azaz $P(X < y$ és $X \geq x)$. Ez pedig pontosan a keresett valószínűség.

A szabályos kockadobásra ismert eloszlásfüggvényünk alapján számoljuk ki annak a valószínűségét, hogy a dobott érték nem kisebb mint 3, de kisebb mint 6: $P(X \in [3, 6)) = F_X(6) - F_X(3) = \frac{5}{6} - \frac{2}{6} = \frac{3}{6}$.

1.1. Diszkrét eloszlás

Azt mondjuk, hogy egy X valószínűségi változó **diszkrét**, ha legfeljebb megszámlálhatóan sok különböző értéket vehet fel. Diszkrét valószínűségi változó eloszlásfüggvénye egyértelműen számolható a $p_i = P(X = i)$ értékekből (az összes lehetséges $i \in R_X$ értékre megadjuk p_i értéket). Ezt az információt egyszerűbb is kezelni, ezt hívjuk a diszkrét valószínűségi változó eloszlásának.

Gondoljuk végig, hogy miként néz ki egy diszkrét valószínűségi változó eloszlásfüggvénye: A kockadobásos példákon is láttuk, hogy a lehetséges értékeknél szakadása van a függvénynek, ezeken a helyeken kívül viszont konstans (ha koordinátarendszerben ábrázolnánk, akkor vízszintes vonal lenne). A szakadások nagysága minden esetben az értékhez tartozó valószínűség (azaz i-ben a szakadás nagysága $F_X(i+0) - F_X(i) = p_i$). A szakadási helyeken felvett érték a balról folytonosságnak megfelelően a korábbi érték (nem pedig a valószínűséggel megnövelt). Az ezekkel a tulajdonságokkal bíró függvényeket szokásosan **lépcsős függvény**eknek nevezik.

Természetesen $0 < p_i \leq 1$ (hiszen valószínűség, és a lehetséges kimenetelekről beszélünk). Emellett $\sum_{i \in R_X} p_i = 1$, hiszen az összes lehetséges esemény valószínűségét összegezzük (kizáró módon).

5. Következmény. *Ha egy kísérletnek megszámlálható a kimeneteleinek a száma, akkor egyértelműen csak diszkrét valószínűségi változót lehet rajta definiálni.*

Ez a következtetés fordítva nem igaz, nézzük erre a következő példát: *Tegyük fel, hogy a $(0,1)$ intervallumból véletlen választott szám kísérletét figyeljük. Definiáljuk az X valószínűségi változónkat úgy, hogy az értéke legyen 0, ha a kísérlet eredménye kisebb, mint $\frac{1}{2}$ és legyen 1 egyébként. Az így kapott X valószínűségi változó diszkrét.*

1.2. Folytonos eloszlás

Azt mondjuk, hogy egy X valószínűségi változó **folytonos**, ha létezik olyan $f_X(x) \geq 0$ integrálható függvény, amire igaz, hogy $F_X(x) = \int_{-\infty}^{x} f_X(y)dy$. Az ilyen tulajdonságú $f_X(x)$ függvény az X valószínűségi változó **sűrűségfüggvény**e.

Legyen az X valószínűségi változónk értéke a $(0,1)$ intervallumból véletlen választott szám. Az így kapott X valószínűségi változó folytonos. A sűrűség függvénye $f_X(t) = 1$ ha $0 < t < 1$, egyébként 0. Könnyen látható, hogy az eloszlásfüggvény 0 és 1 közötti számok esetén pont a szám értékét adja ($F_X(t) = \int_0^t 1dx = t$), ami megfelel az elképzelésünknek, miszerint $P(X < t) = t$.

6. Következmény. *A definícióból következnek*

1. *Folytonos valószínűségi változó eloszlásfüggvénye folytonos.*

2. *Folytonos valószínűségi váltózóra igaz, hogy minden érték valószínűsége* 0. *Azaz:* $\forall t : P(X = t) = 0$.

3. *Folytonos* X *valószínűségi változó esetén egy intervallumba esés valószínűsége a sűrűségfüggvény integrálja az adott intervallumon:* $P(X \in (a, b)) = \int_a^b f_X(t)dt$.

4. *Folytonos valószínűségi váltózó* f_X *sűrűségfüggvényének folytonossági pontjaiban az értéke pont az eloszlásfüggvény deriváltja az adott helyen (azaz* $F'_X(t) = f_X(t)$*).*

Fontos megjegyezni, hogy nem csak diszkrét és folytonos valószínűségi változók léteznek, de mi a tárgy keretein belül csak ilyenekkel fogunk foglalkozni. Ezen kívül érdemes elgondolkodni a sűrűségfüggvény egy pontban való megváltoztatásának hatásán: az integrál nem változik, így a kapott függvény ugyanannak az eloszlásnak sűrűségfüggvénye. Tehát a sűrűségfüggvény nem egyértelmű, de mi a lehetséges sűrűségfüggvények közül mindig a legkevesebb szakadással rendelkezőt fogjuk tekinteni.

7. Következmény. *A sűrűségfüggvényre teljesül, hogy* $\int_{-\infty}^{\infty} f_X(y)dy = 1$ *(hiszen ez pont a biztos esemény valószínűsége).*

8. Megjegyzés. *Folytonos eloszlás esetén a 4. Tétel eredménye könnyen kiterjeszthető (a 6. Következmény második pontjából következik):*

$$P(X \in [a, b)) = P(X \in [a, b]) = P(X \in (a, b)) = P(X \in (a, b]) = F_X(b)-F_X(a).$$

2. Várható érték diszkrét esetben, nevezetes diszkrét v.v.: binomiális, Poisson, geometriai. A binomiális eloszlás közelítése a Poisson-eloszlással. A geometriai eloszlás örökifjú tulajdonsága.

9. Definíció. *Akkor mondjuk, hogy egy sorozat* **abszolút konvergens***, ha a sorozat elemeinek abszolút értékeinek összege véges. Azaz* $\sum_{i\in R_X} |x_i| < \infty$.

10. Definíció. *Egy* X *diszkrét valószínűségi változó (aminek az értékkészlete* R_X *és* $i \in R_X$ *esetén* $P(X = i) = p_i$*)* $E(X)$ *várható értéke a lehetséges értékeinek a valószínűségekkel súlyozott átlaga, ha ez a sor abszolút konvergens. Azaz:*

$$E(X) = \sum_{i\in R_X} i \cdot p_i.$$

Ha a sor nem abszolút konvergens, akkor azt mondjuk, hogy nem létezik a várható érték.

Legyen az X valószínűségi változónk a szabályos kockadobás értéke. Ekkor a várható érték 3,5, hiszen $E(X) = \sum_{i \in R_X} i \cdot p_i = \sum_{i=1}^{6} i \cdot \frac{1}{6} = \frac{21}{6} = \frac{7}{2}$.

Tehát lehetséges olyan valószínűségi változót definiálni, aminek a képlet szerint kaphatnánk várható értéket, de mégis azt mondjuk, hogy nem létezik a várható érték. Nézzünk egy ilyen példát is.

Legyen az X valószínűségi változónk $j = 1, 2, \ldots$ esetén $\frac{2^j}{j}$ értékű $\frac{1}{2^{j+1}}$ valószínűséggel, és $-\frac{2^j}{j}$ értékű $\frac{1}{2^{j+1}}$ valószínűséggel. Ekkor a sor összege

$$\sum_{i \in R_X} i \cdot p_i = \sum_{j=1}^{\infty} \left(\frac{2^j}{j} \cdot \frac{1}{2^{j+1}} - \frac{2^j}{j} \cdot \frac{1}{2^{j+1}} \right) = 0.$$

Azonban a sor elemeinek abszolút értékeinek összege

$$\sum_{i \in R_X} |i| \cdot p_i = \sum_{j=1}^{\infty} \left(\frac{2^j}{j} \cdot \frac{1}{2^{j+1}} + \frac{2^j}{j} \cdot \frac{1}{2^{j+1}} \right) = \sum_{j=1}^{\infty} \frac{1}{j} = \infty.$$

Tehát ennek a valószínűségi változónak definíció szerint nem létezik várható értéke, annak ellenére, hogy a képlet szerint 0-nak vehetnénk.

11. Megjegyzés. *A nem abszolút konvergens sorok esetén azért nem lehet várható értéket definiálni, mert olyan esetben a sorrend megváltoztatása az összeget is megváltoztathatja.*

Tekintsük a

$$\sum_{i=1}^{\infty} (-1)^{i+1} \frac{1}{i}$$

sort. Ennek a határértéke $ln(2)$. Azonban ha átrendezzük, $1 + \frac{1}{3} - \frac{1}{2} + \frac{1}{5} + \frac{1}{7} - \frac{1}{4} + \ldots$, *akkor a kapott sor határértéke $\frac{3}{2} ln(2)$.*

12. Következmény. *Ha egy X diszkrét valószínűségi változó értékkészlete R_X, és az $i \in R_X$ értéket p_i valószínűséggel veszi fel, akkor*

$$min_{a \in \mathbf{R}} \sum_{i \in R_X} (a - i)^2 \cdot p_i = \sum_{i \in R_X} (E(X) - i)^2 \cdot p_i.$$

Tegyük fel, hogy egy kísérlet elvégzése előtt mondanunk kell egy a számot, és a valószínűségi változónk kapott értékétől való eltérést (azaz $|X - a|$ összeget) ki kéne fizetnünk. Ekkor a mediánt érdemes a-nak értékként adni. A **medián** olyan m érték, amire $P(X \leq m) \geq \frac{1}{2}$ és $P(X \geq m) \geq \frac{1}{2}$.

Ha azonban $(X-a)^2$ összeget kéne fizetnünk, akkor az előző következmény szerint $a = E(X)$-el járunk a legjobban. Azaz a várható értékre úgy is lehet tekinteni, mint az optimális fogadásra a négyzeteskülönbség-fizetős játékban. Másik fontos tulajdonsága a várható értéknek, hogy ha sokszor elvégeznénk a kísérletet, és kiátlagolnánk a kapott eredményeket, akkor a várható értékhez közeli számot kapnánk. Ezt a két tulajdonságot később még precízebb formában is látni fogjuk.

Ezután ebben a fejezetben olyan diszkrét valószínűségi változókat vizsgálunk meg, amelyek a valóságban gyakran fordulnak elő.

13. Definíció. *Az* **indikátor valószínűségi változó** *csak* 0 *vagy* 1 *értéket vehet fel.*

Könnyen definiálható bármely kísérlet bármely A eseményének segítségével egy X indikátor valószínűségi változó: legyen az értéke 1 ha bekövetkezett az esemény, és legyen 0 egyébként. Ebben az esetben $p_0 = P(\overline{A})$ és $p_1 = P(A)$. Jelölése $X \in I(A)$ vagy $X \in I(p)$.

14. Tétel. *Az $X \in I(p)$ indikátor valószínűségi változó várható értéke $E(X) = p$.*

Ez az állítás nagyon egyszerűen bizonyítható a definíciókat felhasználva: $E(X) = 1 \cdot p + 0 \cdot (1-p) = p$.

A szabályos kockadobás kísérleten definiáljuk X-et úgy, mint a hatos dobás indikátora, azaz 1 az értéke, ha a dobás hatos, és 0 egyébként. Ekkor $E(X) = \frac{1}{6}$.

15. Definíció. *Legyen A egy K kísérleten definiált p valószínűségű esemény. Végezzük el a K kísérletet n-szer (egymástól függetlenül), és definiáljuk az X valószínűségi változót úgy, hogy az értéke legyen az A bekövetkezéseinek a száma. Ekkor X* **binomiális valószínűségi változó** *n és p paraméterekkel. X értékkészlete: $R_X = \{0, 1, \ldots, n\}$. Jelölése: $X \in Bin(n,p)$.*

Tehát hasonlóan az indikátor valószínűségi változóhoz binomiálisat is tudunk bármely kísérlet segítségével generálni. Az igazság az, hogy a binomiális valószínűségi változó n darab indikátor valószínűségi változó összegéből is megkapható (ha a kísérleteket egymástól függetlenül végeztük el). Vegyük észre, hogy a kísérlet sorozat is tekinthető egy kísérletnek. Ez esetben az eseménytér a lehetséges kimenetsorozatokból áll, azaz Ω^n (ahol Ω az eredeti kísérlet eseménytere).

Dobjunk egymás után 10-szer szabályos kockával, és legyen X a dobott hatosok száma. Ekkor $X \in Bin(10, \frac{1}{6})$.

Adjuk meg $X \in Bin(n, p)$ eloszlását: a lehetséges értékek $0, 1, ..., n$. Mennyi lesz az egyes értékekhez tartozó valószínűség? Egy rögzített kimenet sorrend ugyanakkora valószínűségű, mint egy másik, amiben a sikerek száma ugyanannyi. Így rögzített i esetén a lehetséges sorrendek számát kell megszorozni annak a valószínűségével, hogy az első i kísérlet esetén bekövetkezett A, de a következő $n-i$ esetén nem. Tehát a kérdéses valószínűség

$$p_i = \binom{n}{i} p^i (1-p)^{n-i}.$$

Legyen X az előző példában definiált valószínűségi változó: $X \in Bin(10, \frac{1}{6})$. Ekkor $i = 0, 1, ..., 10$: $p_i = \binom{10}{i} (\frac{1}{6})^i (\frac{5}{6})^{10-i}$.

16. Tétel. *Az $X \in Bin(n, p)$ valószínűségi változó várható értéke $E(X) = n \cdot p$.*

Az előző példákban használt X kockadobás sorozatunk várható értéke $E(X) = \frac{10}{6} = \frac{5}{3}$, hiszen $X \in Bin(10, \frac{1}{6})$.

17. Tétel. *Ha $X \in Bin(n, p)$, akkor X legvalószínűbb értéke az $\lfloor (n+1)p \rfloor$ lesz (ahol $\lfloor x \rfloor$ alsó egészrészt jelöl).*

Legyen X az előző példákból ismert valószínűségi változó: $X \in Bin(10, \frac{1}{6})$. Ekkor X legvalószínűbb értéke $\lfloor (10+1)\frac{1}{6} \rfloor = 1$ lesz. Az is könnyen látható, hogy 11 dobás esetén már a 2 hatos kimenet lenne a legvalószínűbb (az igazság az, hogy ebben az esetben az 1 hatos kimenet pont olyan valószínű mint a 2, de ezt a tételből nem tudjuk leolvasni).

18. Definíció. *Ha az X valószínűségi változó értékkészlete a természetes számok halmaza ($\{0, 1, 2, ...\}$), és létezik egy $\lambda > 0$ paraméter, hogy a p_i valószínűségeket az alábbi képlet adja, akkor λ paraméterű* **Poisson-eloszlás***ának nevezzük, és $X \in Poi(\lambda)$-ként jelöljük:*

$$p_i = e^{-\lambda} \frac{\lambda^i}{i!}.$$

Ahhoz, hogy jobban megértsük, hogy milyen esetekben merül fel Poisson-eloszlás, először a következő tételre lesz szükségünk:

19. Tétel. *Ha* $\lambda = n \cdot p$ *konstans, akkor* $lim_{n\to\infty}\binom{n}{i}p^i(1-p)^{n-i} = e^{-\lambda}\frac{\lambda^i}{i!}$. *(vegyük észre, hogy mivel* λ *konstans, így* n *növekedésével* p *tartani fog* 0*-hoz).*

Tehát a binomiális eloszlás valószínűségeinek határértékei adják a Poisson-eloszlás valószínűségeit, ahol a várható érték ugyanaz marad (hiszen $n \cdot p$ pontosan a várható értéke a binomiális eloszlásnak). Könnyen látható, hogy ha n úgy tartana végtelenhez, hogy közben p konstans, akkor minden korlátos i-re p_i tartana 0-hoz. Ezért csakis úgy van értelme a binomiális eloszlásból végtelenig terjedőt definiálni, hogy p-nek tartania kell nullához. Ez azt is jelenti, hogy elég nagy n és elég kicsi p esetén a binomiális eloszlású valószínűségi változó helyett érdemes Poisson-eloszlásúval modellezni a kísérletet (ezzel könnyebben számolhatóvá tesszük az egyes valószínűségeket).

Ismételjük meg 10^7*-szer (tízmillió) azt a kísérletet, hogy egyszerre dobunk 9 kockával. Legyen* X *azoknak az eseteknek a száma, amikor mind a 9 dobás hatos. Ekkor* $X \in Bin(10^7, (\frac{1}{6})^9)$*, azonban jól közelíthetjük a viselkedését, ha azt mondjuk, hogy* $X \in Poi(1)$ *(mivel* $10^7(\frac{1}{6})^9 \approx 0{,}99229$*).*

20. Tétel. *Az* $X \in Poi(\lambda)$ *valószínűségi változó várható értéke* $E(X) = \lambda$*.*

Ez a várható érték nem lep meg minket, hiszen a binomiális valószínűségi változó sorozat, amivel tartunk a Poisson-eloszlásunkhoz, éppen $n \cdot p$ várható értékű, ami pedig a fentiek alapján megegyezik λ-val.

Az előzőnél életszerűbb példa, ha X*-et úgy definiáljuk, hogy legyen András szúnyogcsípéseinek a száma egy tavaszi tóparti túra után. Feltehető, hogy minden szúnyog, aki közelében elsétált, egyforma* p *valószínűséggel szívta a vérét. Mivel a szúnyogok számát azonban becsülni is nehéz, így nem tudjuk binomiális eloszlással kezelni a feladatot. Ha azonban a tapasztalatok szerint András átlagosan* 1 *szúnyogcsípéssel szokott hazatérni, az azt jelentené, hogy a várható érték* 1 *közeli, így* X *ismét egy* $Poi(1)$ *valószínűségi változónak tekinthető. Ekkor tehát annak a valószínűsége, hogy csípés nélkül tért haza*

$$p_0 = e^{-1}\frac{1^0}{0!} = \frac{1}{e} \approx 0{,}36788.$$

Persze azt is kiszámolhatjuk, hogy mi a valószínűsége, hogy pontosan 2 *csípéssel tér haza:* $p_2 = e^{-1}\frac{1^2}{2!} = \frac{1}{2e} \approx 0{,}18394$*. Annak a valószínűségét, hogy háromnál kevesebb csípése lesz, egy összegzéssel kaphatjuk meg:* $P(X < 3) = P(X = 0) + P(X = 1) + P(X = 2) = \frac{1}{e} + \frac{1}{e} + \frac{1}{2e} = \frac{5}{2e} \approx 0{,}9197$*.*

21. Definíció. *Ha egy kísérletet addig ismételgetünk, amíg egy $p > 0$ valószínűségű A esemény be nem következik, és az X valószínűségi változó értéke az elvégzett kísérletek száma, akkor X-et p paraméterű* **geometriai eloszlás***ú valószínűségi változónak nevezzük. Jelölése $X \in G(p)$.*

Emlékeztethet minket a geometriai eloszlás a binomiálisra, hiszen ugyanúgy egy kísérletet ismételgetünk, és ugyanúgy egy esemény sikerességét vizsgáljuk. A különbség, hogy míg a binomiális esetén előre ismertük a próbálkozások számát (n), és a sikerekét számoltuk, a geometriai esetén a sikerek számát ismerjük (1), és a próbálkozásokét számoljuk.

Dobjunk addig szabályos kockával, amíg hatost nem kapunk. Legyen X a dobások száma. Ekkor $X \in G(\frac{1}{6})$.

A geometriai eloszlásnak nincs köze a geometriai valószínűségi mezőhöz, a kettő nem keverendő! Nézzünk egy példát, amiben mindkettő szerepel: *Robin Hood 100 méter távolságról egy r sugarú kör alakú céltáblára lő, amin a nyila egyenletesen csapódik be valahova (a táblát mindig eltalálja). Addig próbálkozik, amíg el nem találja a középső $\frac{r}{4}$ sugarú koncentrikus kört. Írjunk fel egy X valószínűségi változót a lövések számára!*
Egy lövés sikeressége geometriai valószínűségi mezővel számolható: $P(siker) = \frac{T_{r/4}}{T_r} = \frac{\frac{r^2}{16}\pi}{r^2\pi} = \frac{1}{16}$. Tehát $X \in G(\frac{1}{16})$.

Adjuk meg az $X \in G(p)$ eloszlását: a lehetséges értékek $1, 2, \ldots$. Mennyi lesz az egyes értékekhez tartozó valószínűség? Egy i kimenet azt jelenti, hogy kezdetben $i-1$ sikertelen, végül egy sikeres kísérletet látunk. Tehát a kérdéses valószínűség

$$p_i = p(1-p)^{i-1}.$$

Vizsgáljuk a kockadobásos példában definiált $X \in G(\frac{1}{6})$ valószínűségi változót. Annak a valószínűsége, hogy a harmadik dobás lesz az első hatos $p_3 = \frac{1}{6}(\frac{5}{6})^2 = \frac{25}{216} \approx 0{,}1574$. Annak a valószínűsége, hogy legalább három dobás szükséges $P(X > 2) = 1 - P(X = 1) - P(X = 2) = 1 - \frac{1}{6} - \frac{1}{6}\frac{5}{6} = \frac{25}{36}$ (ezt a valószínűséget persze számolhattunk volna úgy is, hogy a vele ekvivalens B esemény valószínűségét nézzük: $B =\{$az első két dobás nem hatos$\}$).

22. Tétel. *Az $X \in G(p)$ valószínűségi változó várható értéke $E(X) = \frac{1}{p}$.*

Tekintsük az előző példákban vizsgált $X \in G(\frac{1}{6})$ valószínűségi változót. A képlet szerint a várható értéke $\frac{1}{\frac{1}{6}} = 6$, azaz várhatóan a hatodik kockadobás lesz az első hatos.

A geometriai valószínűségi változónak van egy olyan érdekes tulajdonsága, hogy abból az információból, hogy egy bizonyos értéknél nagyobb a kapott érték, az eloszlás nem változik meg, csak eltolódik. Ezt formalizálja az alábbi tétel:

23. Tétel. *A geometriai eloszlás* **örökifjú***:* $\forall b \in \{0, 1, 2, ...\}, \forall a :$
$P(X = a + b|X > b) = P(X = a)$*.*

Bizonyítás: $P(X = a+b|X > b) = \frac{P(X=a+b, X>b)}{P(X>b)} = \frac{(1-p)^{a+b-1}p}{(1-p)^b} = (1-p)^{a-1}p = P(X = a)$

Ha a geometriai eloszlás mögötti kísérletsorozatra gondolunk, akkor könnyen megérthetjük ennek az okát. Attól még, hogy egymás után k darab kísérlet esetén nem következett be a várt kimenet, a következő kísérlet éppen olyan lesz, mintha most indulna egy kísérletsorozat, aminek ő az első eleme (hiszen a kísérleteket egymástól függetlenül végezzük). Vegyük észre, hogy a képletből következik a $P(X > a + b|X > b) = P(X > a)$ és a $P(X < a + b|X > b) = P(X < a)$ alakú formula is.

Tekintsük az előző példákban vizsgált $X \in G(\frac{1}{6})$ *valószínűségi változót. Ha tudjuk, hogy az első* 15 *dobás során nem láttunk hatost, akkor mi annak a valószínűsége, hogy szükség lesz 18-adik dobásra is? A képlet szerint* $P(X > 17|X > 15) = P(X > 2)$ *a keresett valószínűség, ezt pedig egy korábbi példában már kiszámoltuk:* $\frac{25}{36}$*.*

Érdemes itt megemlíteni azt a téves elképzelést, miszerint ha egy kaszinóban a rulett játéknál színre akarunk fogadni, akkor nagyobb az esélyünk a pirosra, ha már régóta csak fekete jött ki (ezt hívják a "Szerencsejátékosok tévedése"-nek). Ez a fajta tévedés gyakori probléma (sok ezzel ekvivalens módon felmerül a mindennapokban). Az örökifjúságból könnyen látható, hogy a korábbiak kimenetelének sikertelensége nem növeli a következő kísérlet sikerességének esélyét.

A diszkrét eloszlások témakörének lezárásaként emlékezzünk vissza a klasszikus valószínűségi mezőre. Természetesen ilyen eseménytéren csakis diszkrét eloszlású valószínűségi változót definiálhatunk (hiszen az eseménytér megszámlálható elemszámú). Nézzünk egy ilyen valószínűségi változó kategóriát:

24. Definíció. *Egy* X *valószínűségi változó* **diszkrét egyenletes** *eloszlású, ha az értékkészletének* (R_X) *elemeit azonos valószínűséggel veszi fel. Ezt* $X \in DU(R_X)$*-rel fogjuk jelölni.*

Azaz ha egy klasszikus valószínűségi mezőn olyan valószínűségi változót definiálunk, ami minden elemi eseményhez különböző értéket rendel, akkor diszkrét egyenletes eloszlásról beszélhetünk (ha két kimenethez rendelhetné ugyanazt az értéket, akkor a kapott valószínűségi változó már nem lenne diszkrét egyenletes eloszlású). Könnyen látható, hogy ekkor minden érték felvételének a valószínűsége $\frac{1}{|R_X|}$.

25. Tétel. *Az $X \in DU(R_X)$ valószínűségi változó várható értéke*

$$E(X) = \frac{1}{|R_X|} \sum_{e \in R_X} e.$$

Korábban sok példát láttunk már klasszikus valószínűségi mezőn definiált diszkrét egyenletes eloszlásra, például a szabályos kockadobás példáját. A várható érték definíciójánál pontosan ennek a valószínűségi változónak számoltuk ki a várható értékét.

3. Várható érték folytonos esetben, nevezetes folytonos v.v.: egyenletes, exponenciális, normális. Szimuláció egyenletes eloszlással. Az exponenciális eloszlás örökifjú tulajdonsága. A standard normális eloszlás, lineáris transzformáció.

3.1. Várható érték folytonos esetben

Az előző fejezet elején definiáltuk a várható értéket, de csak a diszkrét valószínűségi változók esetén. Nézzük meg, hogyan lehetne a folytonos esetre hasonló tulajdonságokkal rendelkező függvényt megadni.

26. Definíció. *Egy $f_X(t)$ sűrűségfüggvényű X folytonos valószínűségi változó $E(X)$* **várható érték***e a lehetséges értékeinek a sűrűségfüggvénnyel súlyozott integrálja, azaz*

$$E(X) = \int_{-\infty}^{\infty} t \cdot f_X(t) dt,$$

amennyiben ez az integrál abszolút konvergens (tehát létezik az $\int_{-\infty}^{\infty} |t| \cdot f_X(t) dt$ integrálnak határértéke). Ha az integrál nem abszolút konvergens, akkor azt mondjuk, hogy nem létezik a várható érték.

Belátható, hogy ez a definíció megfelelő kiterjesztése a diszkrét esetnek, így egyforma tulajdonságokkal fognak rendelkezni. Ezzel egy későbbi fejezetben még részletesen foglalkozni fogunk.

Legyen X az a valószínűségi változó, aminek sűrűségfüggvénye a $(0,1)$ intervallumon 1 azon kívül 0. Ekkor $E(X) = \int_{-\infty}^{\infty} t \cdot f_X(t) dt = \int_0^1 t \cdot 1 dt = [\frac{t^2}{2}]_0^1 = \frac{1}{2}$.

3.2. Egyenletes eloszlás (folytonos)

Emlékezzünk vissza a geometriai valószínűségi mezőre. A klasszikus valószínűségi mezőhöz hasonlóan volt benne egyfajta egyenletesség, de nem lehetett fix p valószínűsége a pontoknak, hiszen végtelen sok van belőlük (és az "összegük" 1-et kéne adjon). Könnyen kikövetkeztethető, hogy ezeknek a kezeléséhez már folytonos valószínűségi változót kell definiálnunk.

27. Definíció. *Ha az X valószínűségi változó értékkészlete egy $I = (a, b)$ intervallum, és ennek bármely részintervallumába pontosan akkora valószínűséggel esik, amekkora a részintervallum hosszának és I hosszának aránya, akkor azt mondjuk, hogy az X valószínűségi változó egyenletes eloszlású az I intervallumon. Jelölése $X \in U(a, b)$.*

28. Következmény. *A definícióból levezethetők az alábbiak:*

1. *Az $X \in U(a, b)$ valószínűségi változónak a sűrűségfüggvénye:*

$$f_X(t) = \frac{1}{b-a} \text{ , ha } t \in (a,b) \text{ ; egyébként } 0.$$

2. *Az $X \in U(a, b)$ valószínűségi változónak a eloszlásfüggvénye:*

$$F_X(t) = 0 \text{ , ha } t < a; \quad \frac{t-a}{b-a} \text{ , ha } t \in (a,b) \text{ ; egyébként } 1.$$

Bizonyítsuk be a két képletet:

1. *Emlékeztetjük az olvasót, hogy a sűrűségfüggvény alapvetően nem egyértelmű (megszámlálható sok pontban megváltoztatva nem változik az általa definiált eloszlás), de mi a jegyzet folyamán minden esetben a legkevesebb szakadási ponttal rendelkezők közül választunk. Ha az (a, b) intervallumon kívül is lenne nemnulla értéke (és nem szakadási pontban), akkor ennek megfelelően kicsi környezetén integrálva a sűrűségfüggvényt azt kapnánk, hogy a valószínűségi változónk pozitív valószínűséggel esik ide (azaz (a, b)-n kívülre), ami lehetetlen. Az (a, b) intervallumon belül az egyenletes eloszlás definíciója miatt egy c konstansnak kell lennie a sűrűségfüggvénynek. A teljes eseménytér valószínűsége pedig 1, tehát $1 = \int_{-\infty}^{\infty} f_X(t)dt$. Esetünkben ez azt jelenti, hogy $\int_a^b c dt = 1$. Ebből azt kapjuk, hogy $b \cdot c - a \cdot c = (b-a)c = 1$, azaz $c = \frac{1}{b-a}$. (Vegyük észre, hogy a $t = a$ vagy b helyeken f_X értékét vehetjük $\frac{1}{b-a}$-nek is, hiszen így még a szakadások száma sem változik.)*

2. *Idézzük fel az eloszlásfüggvény definícióját:* $F_X(t) = P(X < t)$. *Ebből azonnal adódik, hogy az a érték előtt* 0, *a b érték után pedig* 1 *kell legyen az értéke. A kettő között a valószínűség az intervallumok hosszának arányából számolható. A kedvező intervallum egy t helyen* $(t-a)$ *hosszú, a teljes intervallum pedig* $(b-a)$ *hosszú. Ezekből azonnal adódik a képlet.*

Válasszunk egy számot egyenletesen a $(0,1)$ *intervallumról. Ekkor a számot egy* $X \in U(0,1)$*-el modellezhetjük. Ez a valószínűségi változó már korábban is felmerült a példáinkban, így sok tulajdonsága ismert lehet, de most a sűrűségfüggvény és eloszlásfüggvény segítségével fogunk dolgozni. Mi a valószínűsége, hogy* $\frac{1}{3}$*-nál kisebb számot választunk?* $F_X(\frac{1}{3}) = \frac{\frac{1}{3}-0}{1-0} = \frac{1}{3}$. *Mi a valószínűsége, hogy* $X \in (\frac{1}{3}, \frac{2}{3})$*? A 6. Következmény alapján integráljuk a sűrűségfüggvényt a megfelelő intervallumon:* $\int_{\frac{1}{3}}^{\frac{2}{3}} \frac{1}{1-0} dt = \frac{2}{3} - \frac{1}{3} = \frac{1}{3}$. *Ez utóbbit a 8. Megjegyzés-ben szereplő, intervallumok valószínűségeire vonatkozó formulával is számolhatjuk:* $P(X \in (\frac{1}{3}, \frac{2}{3})) = F_X(\frac{2}{3}) - F_X(\frac{1}{3}) = \frac{2}{3} - \frac{1}{3}$.

Amikor számítógéppel akarunk valószínűségi változót előállítani, akkor legtöbbször csak közelíteni tudjuk a viselkedést. A valószínűségi változók között azonban bizonyos transzformációkkal létezik egyfajta "átjárás", ezt alapozza meg a következő tétel:

29. Tétel. *Legyen* $X \in U(0,1)$*, és legyen az* $F_Y(t)$ *eloszlásfüggvény szigorúan monoton (ahol nem* 0 *vagy* 1*). Ekkor az* $Y = F_Y^{-1}(X)$ *módon definiált valószínűségi változó eloszlásfüggvénye* $F_Y(t)$.

Bizonyítás: A szigorú monotonitásból következik, hogy létezik az $F_Y^{-1}(t)$ *inverz. Ekkor* $P(Y < t) = P(F_Y^{-1}(X) < t) = P(X < F_Y(t)) = F_X(F_Y(t)) = F_Y(t)$.

Eszerint, ha a számítógépünk képes a $(0,1)$ intervallumon egyenletes eloszlású valószínűségi változót generálni, akkor ennek segítségével bármelyik folytonos, szigorúan monoton eloszlásfüggvényű valószínűségi változót elő tudjuk vele állítani. Ez egy gyakori felhasználása az eloszlásfüggvény inverzének, amit sokszor **kvantilis**nek (angolul quantile) is szoktak hívni. Figyeljük meg, hogy a kvantilis értéke az $\frac{1}{2}$ helyen pontosan a medián (persze most nem törődünk olyan esetekkel, amikor a kvantilis nem létezik).

A tételben valószínűségi változót adtunk egy függvénynek bemenetként. Ez nem jelent egyebet, mint egy új valószínűségi változót. Ennek egy kiértékelését úgy kaphatjuk meg, hogy elvégezzük az eredeti kísérletet, kiértékeljük az eredeti valószínűségi változót, majd a kapott éréknek vesszük a függvény szerinti képét. Ekkor ha egy $[a, b)$ intervallum valószínűsége p volt, az intervallum K képének a valószínűsége legalább p lesz (akkor lesz nagyobb, ha van olyan $[a, b)$-től diszjunkt pozítív valószínűségű intervallum, aminek a képe része K-nak).

Valójában bármely halmaz valószínűsége a halmaz ősképének a valószínűségével egyezik meg.

3.3. Exponenciális eloszlás

A geometriai eloszláshoz nagyon hasonló dolgot definiálhatunk, ha egy kísérlet első sikerességének idejét vizsgáljuk. A kísérlet azonban most nem ismételt egyszeri dolog, így az időt nem daraboljuk megszámlálható sok részre, hanem mint intervallum tekintünk rá.

30. Definíció. *Az X folytonos valószínűségi változó $\lambda > 0$ paraméterű* **exponenciális eloszlás***ú, ha értékkészlete a pozitív valós számok halmaza, sűrűségfüggvénye pedig:*

$$f_X(t) = \lambda e^{-\lambda t} \text{ , ha } t \geq 0 \text{ ; egyébként } 0.$$

Jelölése $X \in Exp(\lambda)$.

A fentiek alapján tehát akkor beszélhetünk exponenciális eloszlásról, ha valamire való várakozásunk idejét vizsgáljuk.

Például a legközelebbi telefonhívásig; egy gép első meghibásodásáig; a következő ügyfél érkezéséig vagy a következő villámlásig eltelt idő exponenciális eloszlásúnak tekinthető.

31. Következmény. *Az $X \in Exp(\lambda)$ valószínűségi változó eloszlásfüggvénye:*

$$F_X(t) = 1 - e^{-\lambda t} \text{ , ha } t \geq 0 \text{ ; egyébként } 0.$$

Bizonyítás: A folytonos eloszlás definíciója szerint a sűrűségfüggvényt integrálva kapjuk az eloszlásfüggvény értékét. A $t < 0$ esetekre egyszerűen adódik 0, a többi:

$$F_X(t) = \int_{-\infty}^{t} f_X(y)dy = \int_0^t \lambda e^{-\lambda y} dy = [-e^{-\lambda y}]_0^t = -(e^{-\lambda t} - 1) = 1 - e^{-\lambda t}$$

32. Tétel. *Az $X \in Exp(\lambda)$ valószínűségi változó várható értéke:*

$$E(X) = \frac{1}{\lambda}$$

Bizonyítás: parciális integrálást alkalmazunk

$$E(X) = \int_0^\infty \lambda t e^{-\lambda t} dt = [\lambda t(-\frac{e^{-\lambda t}}{\lambda})]_0^\infty - \int_0^\infty \lambda(-\frac{e^{-\lambda t}}{\lambda})dt =$$

$$= 0 + \int_0^\infty e^{-\lambda t} dt = [-\frac{e^{-\lambda t}}{\lambda}]_0^\infty = -(0 - \frac{1}{\lambda}) = \frac{1}{\lambda}$$

Könnyű összekeverni az exponenciális és a Poisson-eloszlás várható értékét, hiszen mindegyik esetén λ-val jelöljük a paramétert, de míg az exponenciálisnál $\frac{1}{\lambda}$, a Poissonnál egyszerűen λ a várható érték.

A valóságban felmerülő példákban a várható értékből következtethetünk a λ paraméter értékére (hasonlóan a Poisson-eloszláshoz). Legyen most X egy gép első meghibásodásáig eltelt idő. A gyártó adatai szerint átlagosan 5 év leteltével jelentkezik az első hiba. Ekkor X exponenciális eloszlású $\lambda = \frac{1}{5}$ paraméterrel, azaz $X \in Exp(\frac{1}{5})$.
Mi annak a valószínűsége, hogy egy gép egy éven belül meghibásodik? $P(X < 1) = F_X(1) = 1 - e^{-\frac{1}{5}1} \approx 1 - 0{,}8187 = 0{,}1813$.
Mi annak a valószínűsége, hogy pontosan a harmadik év során lesz az első hiba? $P(2 < X < 3) = \int_2^3 \lambda e^{-\lambda t} dt = [-e^{-\lambda t}]_2^3 = -(e^{-\frac{1}{5}3} - e^{-\frac{1}{5}2}) \approx 0{,}6703 - 0{,}5488 = 0{,}1215$.

33. Tétel. *Az exponenciális eloszlás örökifjú. Azaz $X \in Exp(\lambda)$ esetén $P(X < a + b|X > b) = P(X < a)$.*

Bizonyítás:

$$P(X < a + b|X > b) = \frac{P(b < X < a + b)}{P(X > b)} = \frac{F_x(a + b) - F_X(b)}{1 - F_X(b)} =$$

$$\frac{1 - e^{-\lambda(a+b)} - (1 - e^{-\lambda b})}{1 - (1 - e^{-\lambda b})} = \frac{e^{-\lambda b} - e^{-\lambda(a+b)}}{e^{-\lambda b}} = 1 - e^{-\lambda a} = P(X < a)$$

34. Megjegyzés. *Ha egy X eloszlás folytonos és örökifjú, akkor $X \in Exp(\lambda)$ egy megfelelő $\lambda > 0$ konstanssal.*

Tehát nincs másik folytonos eloszlás ami rendelkezhet az örökifjúsággal.

35. Tétel. *Az $X \in Exp(\lambda)$ valószínűségi változó kvantilise $Q(p) = -\frac{ln(1-p)}{\lambda}$*

Bizonyítás: Csak $t > 0$ esetekre kell megnézni, a többi érték számunkra lényegtelen (a kvantilis nekünk a $(0,1)$ intervallumon egyenletes eloszlásból való generáláshoz kell). $F_X(Q(p)) = 1 - e^{-\lambda Q(p)} = 1 - e^{-\lambda(-\frac{ln(1-p)}{\lambda})} = 1 - e^{ln(1-p)} = 1 - (1-p) = p$. Azaz tényleg inverze az eloszlásfüggvénynek.

36. Következmény. *Egy $X \in U(0,1)$ valószínűségi változó segítségével előállított $Y = -\frac{ln(1-X)}{\lambda}$ valószínűségi változó éppen λ paraméterű exponenciális eloszlású (azaz $Y \in Exp(\lambda)$).*

Mindez egyszerű következménye a 29. Tételnek.

Korábban már beszéltünk róla, hogy a geometriai eloszlásnak a kiterjesztéseként lehet tekinteni az exponenciálisra, ezt specifikálja a következő tétel:

37. Tétel. *Ha $X \in Exp(\lambda)$ és $Y = [X]$ (X egészrésze), akkor $Y \in G(1 - e^{-\lambda})$.*

3.4. Normális eloszlás

Az utolsó nevezetes eloszlás, amit vizsgálni fogunk, egy kicsit bonyolultabbnak tűnhet, mint az eddigiek. Hogy jobban megértsük, miért is fontos nekünk, először gondoljunk vissza a Poisson-eloszlásra. A binomiális eloszlásból úgy definiáltuk, hogy a kísérletek száma tartott végtelenhez, de a várható érték konstans maradt. Egy másik lehetőség, ha most nem kötjük meg, hogy a várható érték konstans maradjon, hanem inkább kivonjuk azt. Azaz tegyük fel, hogy $X_n \in Bin(n,p)$, és vizsgáljuk az $Y_n = X_n - n \cdot p$ valószínűségi változó sorozat határértékét ($n \to \infty$). Sajnos az így definiált valószínűségi változó még nem lesz számunkra megfelelő, mert túlságosan hektikusan tud viselkedni, azaz túl nagy valószínűséggel lenne nagyon nagy vagy nagyon kicsi értéke. Ez legfőképp ott zavarna minket, hogy nem létezne várható értéke (mert nem abszolút konvergens a megfelelő integrál [lásd a folytonos valószínűségi változók várható értékének definícióját]). Ha azonban a $Z_n = \frac{X_n - n \cdot p}{\sqrt{n}}$ valószínűségi változó sorozatot vizsgáljuk, akkor már egy értelmezhető várható értékkel rendelkező eloszlást kapunk (később ezt még részletezzük a de Moivre–Laplace-tételnél). Erről a fajta eloszlásról fogunk beszélni ebben a fejezetben, azaz a normális eloszlásról.

Tegyük fel, hogy egy fizikai kémiai kísérlet során 10^{24} alkalommal történhet meg egy reakció, mind egymástól függetlenül, p valószínűséggel. Mekkora a valószínűsége annak, hogy a reakciók száma legfeljebb 1% mértékben tér el a várható reakciószámtól (azaz $p \cdot 10^{24}$-től). Ez a fajta kérdés normális eloszlással kezelhető.

38. Definíció. *Egy X valószínűségi változó* **normális eloszlás***ú, ha léteznek $\sigma > 0$ és μ valós paraméterek, amire X sűrűségfüggvénye*

$$f_X(t) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(t-\mu)^2}{2\sigma^2}}.$$

Jelölése: $X \in N(\mu, \sigma)$.

Normális eloszlásúnak tekinthető nagyjából minden olyan folytonos valószínűségi változó, aminek az értéke nagyon sok tényező által befolyásolt. Például az emberek testmagassága, gyufaszál hossza, vonat beérkezési ideje az állomásra, fűszálak össztömege egy négyzetméteren, emberek testsúlya...

A normális eloszlás eloszlásfüggvénye (azaz a fenti sűrűségfüggvény integrálja) nem fejezhető ki elemi függvények segítségével. Ebből kifolyólag szükségünk lesz a következő definíciókra és eredményekre.

39. Definíció. *Az $X \in N(0, 1)$ eloszlású valószínűségi változót* **standard normális** *eloszlásúnak nevezzük. A standard normális valószínűségi változó sűrűségfüggvénye $\varphi(t) = \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dy$, eloszlásfüggvénye $\Phi(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{t} e^{-\frac{y^2}{2}} dy$.*

A standard normális eloszlás eloszlásfüggvénye sem fejezhető ki elemi függvények segítségével, de ezt táblázatokból kiolvashatjuk. A következő képlet segít abban, hogy a standard normális eloszlás eloszlásfüggvényének segítségével hogyan tudjuk a nem standard eseteket kezelni:

40. Tulajdonság(ok). *Ha $X \in N(\mu, \sigma)$, akkor az eloszlásfüggvényére igaz, hogy*

$$F_X(t) = \Phi\left(\frac{t-\mu}{\sigma}\right)$$

Bizonyítás:

$$F_X(t) = \int_{-\infty}^{t} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(y-\mu)^2}{2\sigma^2}} dy$$

Ezt helyettesítéses integrállal tudjuk átalakítani. Hogy a helyettesítéses integrálás (határozott esetben) ne okozzon problémát senkinek, ismételjük át a legfontosabb részeit:

41. Megjegyzés. *Az alábbiak szerint lehetséges egy integrálban $y = h(x)$-et helyettesíteni:*

$$\int_a^b f(h(x))h'(x)dx = \int_{h(a)}^{h(b)} f(y)dy.$$

Amire érdemes odafigyelni, hogy az integrálás alsó és felső határa össze tud „csúszni” , amennyiben a helyettesítési függvény csökkenő. Ilyenkor érdemes az $\int_{h(b)}^{h(a)} -f(y)dy$ *formát használni (a kettő ekvivalens).*

A fenti $F_X(t)$*-be helyettesítsünk* $y = x\sigma + \mu$ *szerint, azaz* $x = \frac{y-\mu}{\sigma}$*. Ekkor* $y < t \iff x < \frac{t-\mu}{\sigma}$*. Ezen felül* $\frac{dy}{dx} = \sigma$*, azaz* $dy = \sigma dx$*.*

$$F_X(t) = \int_{-\infty}^{\frac{t-\mu}{\sigma}} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x\sigma+\mu-\mu)^2}{2\sigma^2}} \sigma dx = \int_{-\infty}^{\frac{t-\mu}{\sigma}} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx = \Phi\left(\frac{t-\mu}{\sigma}\right)$$

Ezt az eredményt felhasználva bármely normális eloszlású valószínűségi változónak egy intervallumba való esésének valószínűségét meg tudjuk határozni csupán a Φ függvény táblázatnak segítségével.

z	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

1. táblázat. A $\Phi(z)$ függvény értékei, ahol $0 \leq z \leq 3,49$

Tegyük fel, hogy az emberek testmagassága centiméterben $X \in N(170, 10)$ *eloszlású valószínűségi változóval modellezhető. Mi a valószínűsége, hogy egy véletlenszerűen választott ember magassága* 170 *cm és* 180 *cm közé esik?*

$$P(X > 170, X < 180) = F_X(180) - F_X(170) = \Phi\left(\frac{180-170}{10}\right) - \Phi\left(\frac{170-170}{10}\right)$$

$$= \Phi(1) - \Phi(0) \approx 0{,}8413 - 0{,}5 = 0{,}3413$$

Érdemes megjegyezni, hogy a φ sűrűségfüggvény, így természetesen teljesíti a sűrűségfüggvényekről ismert tulajdonságokat (pl. nemnegatív, integrálja 1), ezen felül folytonos is. Φ pedig eloszlásfüggvény, így teljesíti a folytonos valószínűségi változók eloszlásfüggvényeiről ismert tulajdonságokat (monoton növekvő, 1-hez tart, folytonos).

Vegyük észre, hogy a táblázat csak pozitív z értékekre adja meg Φ értékét. Ez a következő tulajdonság miatt nem okoz problémát:

42. Tulajdonság(ok). *φ szimmetrikus az y tengelyre, Φ pedig középpontosan szimmetrikus a $(0, \frac{1}{2})$ pontra. Azaz:*

1. $\varphi(t) = \varphi(-t) \qquad \forall t \in \mathbf{R}$.
2. $\Phi(0) = \frac{1}{2}$.
3. $\Phi(t) = 1 - \Phi(-t) \qquad \forall t \in \mathbf{R}$.

Bizonyítás:

1. $\varphi(t) = \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dy = \frac{1}{\sqrt{2\pi}} e^{-\frac{(-t)^2}{2}} dy = \varphi(-t)$
2. $\Phi(0) = \int_{-\infty}^{0} \varphi(y)dy$, *ez pedig egy $y = -x$ helyettesítés után (lásd a 41. Megjegyzést) használjuk az 1. tulajdonságot:*
$$\Phi(0) = \int_{\infty}^{-0} \varphi(-y)dy = \int_{0}^{\infty} \varphi(-y)dy = \int_{0}^{\infty} \varphi(y)dy.$$
Azaz $\Phi(0) + \Phi(0) = \int_{-\infty}^{\infty} \varphi(y)dy$. A sűrűségfüggvénynek a teljes tartományon vett integrálja 1, így $\Phi(0) = \frac{1}{2}$.
3. $\Phi(t) = \int_{-\infty}^{t} \varphi(y)dy = 1 - \int_{t}^{\infty} \varphi(y)dy$ *Ismét helyettesítsünk $y = -x$-et:* $\Phi(t) = 1 - \int_{-\infty}^{-t} \varphi(-x)dx = 1 - \int_{-\infty}^{-t} \varphi(x)dx = 1 - \Phi(-t)$

Ezek segítségével és az előző példánkban szereplő $X \in N(170, 10)$ valószínűségi változóval kiszámolhatjuk, hogy mi a valószínűsége annak, hogy egy véletlenszerűen választott ember magassága 130 cm és 150 cm közé esik.

$$P(X > 130, X < 150) = F_X(150) - F_X(130) = \Phi\left(\frac{150-170}{10}\right) - \Phi\left(\frac{130-170}{10}\right)$$

$$= \Phi(-2) - \Phi(-4) = 1 - \Phi(2) - (1 - \Phi(4)) = \Phi(4) - \Phi(2) \approx 1 - 0{,}9772 = 0{,}0228.$$

Az is könnyen látható, hogy a szimmetria miatt ez megegyezik annak a valószínűségével, hogy egy véletlenszerűen választott ember magassága 190 cm és 210

cm közé esik.
Melyik eseménynek nagyobb a valószínűsége egy véletlenül választott ember magasságára: $A =${160 cm és 200 cm közé esik} vagy $B =${140 cm és 180 cm közé esik}? A szimmetria miatt a két valószínűség pontosan megegyezik.

Normális valószínűségi változó a 40. Tulajdonság szerint a standard normálisból kifejezhető. Ezt a tulajdonságot máshogy is meg tudjuk fogalmazni:

43. Következmény. *Bármely két normális eloszlás között lineáris a kapcsolat:*

1. *Ha* $X \in N(0,1)$, *akkor* $Y = \sigma X + \mu$ *esetén* $Y \in N(\mu, \sigma)$.
2. *Ha* $Y \in N(\mu, \sigma)$, *akkor* $X = \frac{Y-\mu}{\sigma}$ *esetén* $X \in N(0,1)$.
3. *Ha* $N_1 \in N(\mu_1, \sigma_1)$, *akkor* $N_2 = \frac{\sigma_2}{\sigma_1} N_1 + \mu_2 - \frac{\sigma_2}{\sigma_1}\mu_1$ *esetén* $N_2 \in N(\mu_2, \sigma_2)$.

Bizonyítás:

1. $F_Y(t) = P(\sigma X + \mu < t) = P(X < \frac{t-\mu}{\sigma}) = \Phi(\frac{t-\mu}{\sigma})$, *ami a 40. Tulajdonság szerint* $N(\mu, \sigma)$ *eloszlású valószínűségi változó eloszlásfüggvénye.*
2. $F_X(t) = P(\frac{Y-\mu}{\sigma} < t) = P(Y < t\sigma + \mu) = F_Y(t\sigma + \mu)$, *ez pedig a 40. Tulajdonság szerint* $\Phi(\frac{t\sigma+\mu-\mu}{\sigma}) = \Phi(t)$. *Tehát* X *standard normális eloszlású.*
3. *A második tulajdonság szerint* $S = \frac{N_1-\mu_1}{\sigma_1}$ *standard normális eloszlású. Vegyük észre, hogy* $N_2 = \sigma_2 S + \mu_2$, *ami az első tulajdonság szerint ekkor* $N(\mu_2, \sigma_2)$ *eloszlású.*

44. Tétel. *Az* $X \in N(\mu, \sigma)$ *valószínűségi változó várható értéke* $E(X) = \mu$.

4. Transzformált változó várható értékének kiszámolása. Szórás, momentumok. Várható értékre, szórásra vonatkozó tételek. Steiner-formula. Nevezetes eloszlások várható értékei, szórásai.

Korábban láttuk a várható érték definícióját diszkrét és folytonos esetre is. Most vizsgáljuk meg, hogy mit mondhatunk a transzformált valószínűségi változók várható értékéről!

45. Tulajdonság(ok). *Legyen X tetszőleges valószínűségi változó, ekkor:*

1. $E(aX + b) = aE(X) + b$.
2. *Legyen X diszkrét, R_X értékkészlettel, és $g(x)$ egy függvény. Ekkor $g(X)$ várható értéke az alábbi, amennyiben a sor abszolút konvergens:*
$$E(g(X)) = \sum_{i \in R_X} g(i) \cdot p_i.$$
3. *Legyen X folytonos, f_X sűrűségfüggvénnyel, és $g(x)$ egy függvény. Ekkor $g(X)$ várható értéke az alábbi, amennyiben az integrál abszolút konvergens:*
$$E(g(X)) = \int_{-\infty}^{\infty} g(y) \cdot f_X(y) dy.$$
4. *Legyen Y is egy valószínűségi változó, ekkor:*
$$E(X + Y) = E(X) + E(Y).$$
5. *Legyen $c \in \mathbf{R}$ konstans. Ha $C = c$ vagy C egy olyan valószínűségi változó, ami 1 valószínűséggel a c értéket veszi fel, akkor $E(C) = c$. Így persze*
$$E(E(X)) = E(X).$$

Ha kíváncsiak lennénk a szabályos kockadobás értékének 10-szeresénél eggyel nagyobb szám várható értékére, akkor az $Y = 10X + 1$ valószínűségi változót kell vizsgálnunk (ahol X a kockadobás értéke). Az első tulajdonság szerint $E(Y) = 10\frac{7}{2} + 1 = 36$. A negyedik tulajdonságból pedig azt is tudjuk, hogy például két kockadobás értékének összegének várható értéke $\frac{7}{2} + \frac{7}{2} = 7$.

Talán egy kicsit még izgalmasabb, ha az X kockadobás és egy $Z \in U(0, 1)$ egyenletes eloszlású valószínűségi változót adnánk össze. A negyedik tulajdonságból rögtön adódik hogy $E(X + Z) = 4$.

A várható érték abban segített nekünk, hogy tudjunk valamit az eloszlásunk átlagos értékéről. Sokszor ezen felül az is érdekelne minket, hogy ettől az értéktől várhatóan mennyire tér el a valószínűségi változónk értéke. Ezt mérhetnénk akár $E(|X - EX|)$ meghatározásával, azonban az abszolút érték kezelése nem egyszerű, ezért érdemesebb az ehhez közel álló, következő értéket figyelnünk (ez persze nem a legfontosabb érv, valójában egy dimenzió felett az Euklideszi távolsághoz hasonlóan a különbség négyzete játszik lényeges szerepet az eltérés mértékének meghatározásában):

46. Definíció. *Egy X valószínűségi változó* **szórás***án a*
$$\sigma(X) = \sqrt{E((X - E(X))^2)}$$

értéket értjük, amennyiben a gyök alatti várható érték létezik. A **szórásnégyzet***et sokszor* **varianciá***nak hívjuk:* $\sigma^2(X) = E((X - E(X))^2)$.

A képletbe be tudjuk helyettesíteni a diszkrét vagy folytonos várható érték definícióját, ekkor a következőket kapjuk (a 45. Tulajdonságot felhasználva):

47. Következmény. *Ha X diszkrét valószínűségi változó, aminek létezik szórásnégyzete, akkor* $\sigma^2(X) = \sum_{i \in R_X} (i - E(X))^2 \cdot p_i$.
Ha X folytonos valószínűségi változó, aminek létezik szórásnégyzete, akkor $\sigma^2(X) = \int_{-\infty}^{\infty} (y - E(X))^2 \cdot f_X(y) dy$.

Vizsgáljuk a szabályos kockadobás X értékét. A következményben szereplő képlet szerint $\sigma^2(X) = \sum_{i \in R_X} (i - E(X))^2 \cdot p_i$ Helyettesítsük be az értékeket, és a korábban kiszámolt $E(X) = \frac{7}{2}$ várható értéket:

$$\sigma^2(X) = \sum_{i=1}^{6} \left(i - \frac{7}{2}\right)^2 \cdot \frac{1}{6} = \frac{1}{6}\left(\frac{25}{4} + \frac{9}{4} + \frac{1}{4} + \frac{1}{4} + \frac{9}{4} + \frac{25}{4}\right) = \frac{35}{12}.$$

Vizsgáljuk az $Y \in U(0,1)$ valószínűségi változót. A következményben szereplő képlet szerint $\sigma^2(Y) = \int_{-\infty}^{\infty} (y - E(Y))^2 \cdot f_Y(y) dy$. Helyettesítsük be a sűrűségfüggvényt, és a korábban kiszámolt $E(Y) = \frac{1}{2}$ várható értéket:

$$\sigma^2(Y) = \int_0^1 (y - \frac{1}{2})^2 \cdot 1 dy = \int_0^1 y^2 - y + \frac{1}{4} dy = \left[\frac{y^3}{3} - \frac{y^2}{2} + \frac{y}{4}\right]_0^1 = \frac{1}{12}$$

Emlékezzünk, hogy a 46. Definícióban szereplő képlethez hasonló szerepelt a várható érték definíciójánál a 12. Következményben. A kettőt összevonva úgy fogalmazhatjuk meg, hogy az $E((X - a)^2)$ kifejezés minimuma σ^2, amit $a = E(X)$-ben vesz fel.

48. Tulajdonság(ok). *Az X valószínűségi változó lineáris transzformáltjának szórása:* $\sigma(aX + b) = |a|\sigma(X)$ *(azaz* $\sigma^2(aX + b) = a^2\sigma^2(X)$*).*

A szórás kiszámítása legtöbbször egyszerűbb a következő formulával, mint a definícióval:

49. Tétel. *(Steiner-formula)* $\forall a \in \mathbf{R}$ *esetén igaz, hogy* $\sigma^2(X) = E((X - a)^2) - (E(X - a))^2$. *Így természetesen:*

$$\sigma^2(X) = E(X^2) - (E(X))^2.$$

Bizonyítás: Csak az $a = 0$ esetet fogjuk bizonyítani. A definíció szerint

$$\sigma^2(X) = E((X - E(X))^2) = E(X^2 - 2X \cdot E(X) + (E(X))^2).$$

Ha felhasználjuk a 45. Tulajdonságban látottakat, akkor azt kapjuk, hogy

$$\sigma^2(X) = E(X^2) - E(2X \cdot E(X)) + E((E(X))^2) = E(X^2) - 2E(X) \cdot E(X) + (E(X))^2.$$

A rendezés során felhasználtuk, hogy $E(X)$ egy konstans. A jobb oldal egyenlő a Steiner-formula jobb oldalával, így a két érték valóban megegyezik.

A formulából nagyon egyszerűen számolható az indikátor valószínűségi változó szórásnégyzete. Ha $X \in I(p)$ indikátor valószínűségi változó, akkor $X \simeq X^2$, hiszen a 0 és az 1 számok önmaguk négyzetei. Így a Steiner-formula szerint $\sigma^2(X) = E(X) - (E(X))^2 = p - p^2 = p(1-p)$.

A Steiner-formulában szerepel a valószínűségi változó négyzetének várható értéke. A következő definíció nevet ad többek közt ennek is:

50. Definíció. *Egy X valószínűségi változó n-**edik momentum**án az n-edik hatványának a várható értékét értjük, azaz $E(X^n)$-t. (Azt mondjuk, hogy létezik a momentum, ha létezik a hozzá tartozó várható érték.)*

51. Tétel. *Pontosan akkor létezik egy valószínűségi változónak szórása, ha létezik második momentuma.*

A várható érték lineáris transzformált esetén a 45. Tulajdonságban látottak szerint viselkedik, a szórás pedig a 48. Tulajdonság szerint. A kettőt felhasználva a következőt kapjuk:

52. Következmény. *Az $X' = \frac{X - E(X)}{\sigma(X)}$ valószínűségi változó várható értéke 0 szórása pedig 1.*

Ezt az X' valószínűségi változót szokás X standardizáltjának hívni. Vegyük észre, hogy amikor normális eloszlásból standard normálist akartunk előállítani, ott is ugyanezt tettük.

4.1. Nevezetes eloszlások szórásai

53. Tétel. *Ha $X \in I(p)$ indikátor valószínűségi változó, akkor $\sigma^2(X) = p(1-p)$.*

54. Tétel. *Ha $X \in Bin(n,p)$ binomiális eloszlású valószínűségi változó, akkor*

$$\sigma^2(X) = np(1-p).$$

55. Tétel. *Ha* $X \in Poi(\lambda)$ *Poisson-eloszlású valószínűségi változó, akkor*

$$\sigma^2(X) = \lambda.$$

56. Tétel. *Ha* $X \in G(p)$ *geometriai eloszlású valószínűségi változó, akkor*

$$\sigma^2(X) = \frac{1-p}{p^2}.$$

57. Tétel. *Ha* $X \in U(a,b)$ *egyenletes eloszlású valószínűségi változó, akkor*

$$\sigma^2(X) = \frac{(b-a)^2}{12}.$$

Bizonyítás:

$$E(X^2) = \int_{-\infty}^{\infty} y^2 f_X(y) dy = \int_a^b y^2 \frac{1}{b-a} dy = \left[\frac{x^3}{3(b-a)}\right]_a^b =$$

$$= \frac{b^3 - a^3}{3(b-a)} = \frac{a^2 + ab + b^2}{3}$$

Így Steiner-formula segítségével:

$$\sigma^2(X) = \frac{a^2+ab+b^2}{3} - \left(\frac{a+b}{2}\right)^2 = \frac{4a^2 + 4ab + 4b^2 - 3a^2 - 6ab - 3b^2}{12} =$$

$$\frac{a^2 - 2ab + b^2}{12} = \frac{(a-b)^2}{12}.$$

58. Tétel. *Ha* $X \in Exp(\lambda)$ *exponenciális eloszlású valószínűségi változó, akkor*

$$\sigma^2(X) = \frac{1}{\lambda^2}.$$

Bizonyítás: A második momentumot most parciális integrálással számolhatjuk:

$$E(X^2) = \int_{-\infty}^{\infty} y^2 f_X(y) dy = \int_0^{\infty} y^2 \lambda e^{-\lambda y} dy =$$

$$= \left[y^2 \left(-\frac{1}{\lambda}\lambda e^{-\lambda y}\right)\right]_0^{\infty} - \int_0^{\infty} 2y \left(-\frac{1}{\lambda}\lambda e^{-\lambda y}\right) dy = 0 + \frac{2}{\lambda}\int_0^{\infty} y\lambda e^{-\lambda y} dy =$$

$$= \frac{2}{\lambda}E(X) = \frac{2}{\lambda^2}$$

Így Steiner-formula segítségével:

$$\sigma^2(X) = \frac{2}{\lambda^2} - \left(\frac{1}{\lambda}\right)^2 = \frac{1}{\lambda^2}.$$

59. Tétel. *Ha* $X \in N(\mu, \sigma)$ *normális eloszlású valószínűségi változó, akkor*

$$\sigma^2(X) = \sigma^2.$$