Valószínűségszámítás jegyzet 3. rész

1. Markov- és Csebisev-egyenlőtlenség. Együttes- és vetületi eloszlásfüggvény, függetlenség. Diszkrét és folytonos eset. Együttes- és vetületi eloszlás. Együttes- és vetületi sűrűségfüggvény.

A várható érték és szórás segítségével képesek vagyunk becsléseket adni bizonyos valószínűségekre. A két legnevezetesebb ilyen összefüggést mondja ki az alábbi két tétel.

1. Tétel. *(Markov-egyenlőtlenség) Legyen X egy valószínűségi változó, aminek értékkészlete $R_X \subseteq \mathbf{R}_\mathbf{0}^+$ (azaz $X \geq 0$), és létezik várható értéke. Ekkor bármely $a > 0$ valós számra:*

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

Ekvivalensen $P(X \geq aE(X)) \leq \frac{1}{a}$ és $P(X < aE(X)) \geq 1 - \frac{1}{a}$ is teljesül.

Bizonyítás diszkrét esetre:

$$E(X) = \sum_{i \in R_X} i \cdot p_i = \sum_{i>0, i \in R_X} i \cdot p_i \geq \sum_{i \geq a, i \in R_X} i \cdot p_i \geq \sum_{i \geq a, i \in R_X} a \cdot p_i = aP(X \geq a)$$

Bizonyítás folytonos esetre:

$$E(X) = \int_{-\infty}^{\infty} y f_X(y) dy = \int_0^{\infty} y f_X(y) dy \geq \int_a^{\infty} y f_X(y) dy$$

$$\int_a^{\infty} y f_X(y) dy \geq \int_a^{\infty} a f_X(y) dy = a \cdot P(X \geq a)$$

Tekintsük a szabályos kockadobás X értékét. Becsüljük felülről Markov-egyenlőtlenséggel, hogy mi a valószínűsége a 3-nál nem kisebb értékű dobásnak.

$$P(X \geq 3) \leq \frac{\frac{7}{2}}{3} = \frac{7}{6}$$

Ez a becslés nem mond sokat, hiszen egynél nagyobb értéket kaptunk egy valószínűségre felső becslésként. Vizsgáljuk meg a 6-nál nem kisebb dobás esetét:

$$P(X \geq 6) \leq \frac{\frac{7}{2}}{6} = \frac{7}{12} \approx 0{,}583.$$

Ez már érdekesebb becslés, de mivel tudjuk, hogy a valószínűség pontosan $\frac{1}{6}$, így nem igazán hasznos.

Válasszunk 100 000 ledet egy gyártósorról minőségellenőrzésre. Ha egy led 1% eséllyel hibás, becsüljük meg annak a valószínűségét, hogy legalább 5000 hibását választunk. Ezt a valószínűséget már nehezen számolnánk ki, hiszen $Y \in Bin(100000, 0{,}01)$ esetén nagy kitevőkkel és csúnya binomiális együtthatókkal kéne dolgozni (és sok ilyet összegezni). Próbálkozhatnánk Poisson-eloszlással vagy normális eloszlással közelíteni, de most a Markov-egyenlőtlenséget fogjuk használni, ahol $E(Y) = 1000$:

$$P(Y \geq 5 \cdot 1000) \leq \frac{1}{5} = 0{,}2.$$

Ez a becslés nagyon egyszerűen adott számunkra egy felső korlátot egy nagyon nehezen számolható valószínűségre.

Nem csak nemnegatív valószínűségi változókra tudjuk a Markov-egyenlőtlenséget alkalmazni:

2. Következmény. *Legyen X egy valószínűségi változó. Ekkor bármely $a > 0$ valós számra fennállnak:*

$$P(|X| \geq a) \leq \frac{E(|X|)}{a}.$$

$$P(X^2 \geq a) \leq \frac{E(X^2)}{a}.$$

Bizonyítás: A Markov-egyenlőtlenséget felírhatjuk az $|X|$ és X^2 valószínűségi változókra, hiszen ezek nemnegatív értékűek. Behelyettesítve pont az állítást kapjuk.

3. Tétel. *A Markov-egyenlőtlenségnek létezik egy kicsivel erősebb formája is: Legyen X egy valószínűségi változó és $g(x)$ egy nemnegatív értékű monoton növő függvény. Ekkor bármely $a > 0$ valós számra:*

$$g(a)P(X \geq a) \leq E(g(X)).$$

Bizonyítás folytonos esetre:

$$g(a)P(X \geq a) = \int_a^\infty g(a) f_X(y) dy \leq \int_a^\infty g(y) f_X(y) dy \leq E(g(X))$$

Bizonyítás diszkrét esetre:

$$g(a)P(X \geq a) = \sum_{i \geq a} g(a) p_i \leq \sum_{i \geq a} g(i) p_i \leq E(g(X))$$

4. Tétel. *(Csebisev-egyenlőtlenség) Legyen X egy valószínűségi változó, aminek létezik $\sigma(X)$ szórása. Ekkor bármely $a > 0$ valós számra:*

$$P(|X - E(X)| \geq a) \leq \frac{\sigma^2(X)}{a^2}.$$

Ekvivalensen $P(|X - E(X)| \geq a\sigma(X)) \leq \frac{1}{a^2}$ és $P(|X - E(X)| < a\sigma(X)) \geq 1 - \frac{1}{a^2}$ is teljesül.

Bizonyítás: Az $Y = |X - E(X)|$ valószínűségi változóra írjuk fel a 3. Tételt:

$$g(a)P(|X - E(X)| \geq a) \leq E(g(|X - E(X)|)).$$

Ekkor ha a $g(x) = x^2$, ha $x > 0$, egyébként 0 függvényt írjuk be (ami monoton növő):

$$a^2 P(|X - E(X)| \geq a) \leq E((|X - E(X)|)^2) = E((X - E(X))^2) = \sigma^2(X).$$

András és Bíbor játszanak egy játékot: dobnak egy szabályos kockával, és ha a kapott X szám nagyobb, mint 3, akkor András fizet Bíbornak $X - \frac{7}{2}$ eurót, ellenkező esetben Bíbor fizet Andrásnak $\frac{7}{2} - X$ eurót. Adjunk felső korlátot annak a valószínűségére, hogy legalább másfél euró cserél gazdát egy dobásnál! Korábban már kiszámoltuk, hogy a kockadobás szórása $\frac{35}{12}$. Ekkor a kérdéses valószínűség

$$P(|X - E(X)| \geq 1{,}5) \leq \frac{\left(\frac{35}{12}\right)^2}{1{,}5^2} = \left(\frac{35}{18}\right)^2.$$

Ez megint nem érdekes becslés, hiszen nagyobb mint 1 (miközben a kérdéses valószínűség $\frac{2}{3}$).

Nézzük az előző példában látott Y valószínűségi változót a hibás ledek számára, és becsüljük meg ugyanazt a valószínűséget. Ez esetben (kihasználva, hogy az $\{Y - 1000 \leq -4000\}$ lehetetlen esemény):

$$P(Y \geq 5000) = P(Y - 1000 \geq 4000) = P(|Y - 1000| \geq 4000).$$

Ekkor felírva a Csebisev-egyenlőtlenséget ($\sigma^2(X) = 100000 \cdot 0{,}01 \cdot 0{,}99$):

$$P(|Y - 1000| \geq 4000) \leq \frac{990}{4000^2} = 0{,}000061875.$$

Látható, hogy most a Markov-egyenlőtlenséggel kapott 0,2-es becslésnél sokkal erősebb felső korlátot kaptunk.

Sokszor nem egyetlen véletlenen alapuló értékre vagyunk kíváncsiak, inkább értékek egy halmazára. Ezt tipikusan vektorban tároljuk. Így jutunk el a valószínűségi vektorváltozókhoz.

5. Definíció. *Egy* $\underline{X} \in \Omega \to \mathbf{R}^n$ *függvény* **valószínűségi vektorváltozó** *(*$\underline{X} = (X_1, X_2, ..., X_n)$*), ha az* $A = \{\underline{X} < \underline{x}\}$ *esemény megfigyelhető bármely* $\underline{x} \in \mathbf{R}^n$ *esetén (*$\underline{x} < \underline{y}$ *azt jelenti, hogy minden* i *koordinátában* $x_i < y_i$*).*
Az $\underline{X}$ *valószínűségi vektorváltozó* **eloszlásfüggvénye** *az az* $\mathbf{R}^n \to \mathbf{R}$ *függvény, amire:*

$$F_{\underline{X}}(\underline{x}) = P(\underline{X} < \underline{x}) = P(X_1 < x_1, X_2 < x_2, ..., X_n < x_n).$$

6. Tulajdonság(ok). *Az alábbi tulajdonságok minden valószínűségi vektorváltozó eloszlásfüggvényére igazak (*$\underline{x} \to \underline{y} - 0$ *azt jelenti, hogy minden egyes koordináta balról tart a másikhoz):*

1. *monoton növekvő: ha* $\underline{x} \leq \underline{y}$*, akkor* $F_{\underline{X}}(\underline{x}) \leq F_{\underline{X}}(\underline{y})$*,*
2. *balról folytonos:* $\lim_{\underline{x} \to \underline{y}-0} F_{\underline{X}}(\underline{x}) = F_{\underline{X}}(\underline{y})$*,*
3. *ha az egyik koordináta tart mínusz végtelenhez, akkor az értéke 0-hoz tart:* $\forall i: \ \lim_{x_i \to -\infty} F_{\underline{X}}(x_1, x_2, ..., x_n) = 0$*,*
4. *ha minden koordináta tart a végtelenhez, akkor az értéke 1-hez tart:* $\lim_{\forall i: x_i \to \infty} F_{\underline{X}}(x_1, x_2, ..., x_n) = 1$*.*

Dobjunk két kockával, és legyen az X_1 *valószínűségi változó a nagyobbik,* X_2 *pedig a kisebbik dobás értéke. Az* $\underline{X} = (X_1, X_2)$ *valószínűségi vektorváltozó eloszlása az alábbi táblázatból kiolvasható:*

X_2 \ X_1	1	2	3	4	5	6
1	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{2}{36}$
2	0	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{2}{36}$
3	0	0	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{2}{36}$
4	0	0	0	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{2}{36}$
5	0	0	0	0	$\frac{1}{36}$	$\frac{2}{36}$
6	0	0	0	0	0	$\frac{1}{36}$

Ahogy korábban is láttuk, a diszkrét esetben nem kényelmes az eloszlásfüggvény megadása, ezért itt is csak a valószínűségekkel adtuk meg az eloszlást.

Nézzünk egy folytonos esetet is. Legyen $X, Y \in U(0,1)$ *két egymástól függetlenül választott érték. Mi lesz az együttes eloszlásfüggvényük?*

$F_{(X,Y)}(x, y) = P(X < x, Y < y) = P(X < x) \cdot P(Y < y) = x \cdot y$ *, ha* $0 \leq x, y \leq 1$

ha $x < 0$ vagy $y < 0$ akkor az értéke 0, ha $x > 1, y \in [0, 1]$ akkor az értéke y, ha $y > 1, x \in [0, 1]$ akkor az értéke x, ha pedig $1 < x, y$ akkor az értéke 1. Persze mindezt kicsit egyszerűbben így is írhatnánk:

$$F_{(X,Y)}(x, y) = \max\{0, \min\{1, x\}\} \cdot \max\{0, \min\{1, y\}\}.$$

7. Következmény. *Az előző tulajdonságok szerint egy valószínűségi vektorváltozó minden koordinátája valószínűségi változó. Amikor a koordinátákban szereplő valószínűségi változókról beszélünk, akkor a vektorváltozó eloszlásfüggvényét az* **együttes eloszlásfüggvény***üknek nevezzük. Ha a vektor koordinátáinak egy részhalmazának az eloszlásáról beszélünk, akkor ezt* **vetületi eloszlás***nak (vagy peremeloszlásnak) hívjuk.*

Nézzük meg az előző példában szereplő két kockás valószínűségi vektorváltozót. Diszkrét kétdimenziós esetben a fentihez hasonló táblázatból könnyen számolható a két vetületi eloszlás, méghozzá a sorösszegek és oszlopösszegek segítségével:

X_2 \ X_1	1	2	3	4	5	6	Σ
1	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{11}{36}$
2	0	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{9}{36}$
3	0	0	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{7}{36}$
4	0	0	0	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{5}{36}$
5	0	0	0	0	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$
6	0	0	0	0	0	$\frac{1}{36}$	$\frac{1}{36}$
Σ	$\frac{1}{36}$	$\frac{3}{36}$	$\frac{5}{36}$	$\frac{7}{36}$	$\frac{9}{36}$	$\frac{11}{36}$	1

A táblázat minden sorának utolsó eleme annak a valószínűsége, hogy X_2 megegyezik a sor első elemével, míg a táblázat minden oszlopának utolsó eleme annak a valószínűsége, hogy X_1 megegyezik az oszlop első elemével.

8. Tétel. *Az együttes eloszlásból egyértelműen meghatározhatók a vetületi eloszlások (de ez fordítva nem igaz). Pontosabban, ha $I = \{i_1, i_2, ...i_k\}$ részhalmaza $\{1, 2, ..., n\}$-nek, akkor*

$$\lim_{\forall i \notin I: x_i \to \infty} F_{\underline{X}}(x_1, x_2, ..., x_n) = F_{(X_{i_1}, X_{i_2}, ..., X_{i_k})}(x_{i_1}, x_{i_2}, ..., x_{i_k})$$

Nézzünk egy példát arra, hogy visszafelé nem lehet következtetni a vetületi eloszlásokból. Tegyük fel, hogy ismerjük két 0,5 valószínűségű esemény indikátor változóját: $I_A, I_B \in I(\frac{1}{2})$. Mi lehet az együttes eloszlásuk? A két esemény akár meg is egyezhet, de lehet egymás komplementere is. Ezt a két lehetőséget mutatja az alábbi két táblázat.

$I_B \backslash I_A$	0	1	Σ
0	$\frac{1}{2}$	0	$\frac{1}{2}$
1	0	$\frac{1}{2}$	$\frac{1}{2}$
Σ	$\frac{1}{2}$	$\frac{1}{2}$	1

$I_B \backslash I_A$	0	1	Σ
0	0	$\frac{1}{2}$	$\frac{1}{2}$
1	$\frac{1}{2}$	0	$\frac{1}{2}$
Σ	$\frac{1}{2}$	$\frac{1}{2}$	1

Látszik, hogy a vetületi eloszlások megegyeznek (utolsó sor és utolsó oszlop), de az együttes eloszlás teljesen más.

Most végre minden eszközünk megvan ahhoz, hogy precízen is definiáljuk a függetlenséget, amit korábban már oly sok esetben feltételeztünk valamilyen módon. Emlékezzünk az események függetlenségére: A és B pontosan akkor független, ha $P(AB) = P(A)P(B)$.

9. Definíció. *Az X és Y valószínűségi változók pontosan akkor* **függetlenek***, ha bármely x, y valós számpárra $F_{X,Y}(x, y) = F_X(x)F_Y(y)$.*

Vegyük észre, hogy ez az eloszlásfüggvények definíciója szerint azzal ekvivalens, hogy $P(X < x, Y < y) = P(X < x)P(Y < y)$. *Egy nem sokkal ezelőtti példában ezt már ki is használtuk, amikor $X, Y \in U(0, 1)$ egymástól függetlenül választott számokról beszéltünk.*

Ez pedig azt jelenti, hogy két valószínűségi változó pontosan akkor független, ha az egyik segítségével definiált bármely megfigyelhető esemény független a másik segítségével definiált bármely megfigyelhető eseménytől. Ez a definíció az alábbi két módon terjeszthető ki 2-nél magasabb dimenzióra (azaz 2-nél több valószínűségi változóra).

10. Definíció. *Az $X_1, X_2, ..., X_n$ valószínűségi változók pontosan akkor* **páronként függetlenek***, ha bármely X_i, X_j pár független ($i \neq j$).*

11. Definíció. *Az $X_1, X_2, ..., X_n$ valószínűségi változók pontosan akkor* **teljesen függetlenek***, ha bármely $I = \{i_1, i_2, ...i_k\}$ részhalmazára $\{1, 2, ..., n\}$-nek, és bármely $\{x_{i_1}, x_{i_2}, ..., x_{i_k}\}$ valós számsorozatra*

$$F_{(X_{i_1}, X_{i_2}, ..., X_{i_k})}(x_{i_1}, x_{i_2}, ..., x_{i_k}) = \Pi_{i \in I} F_{X_i}(x_i).$$

12. Következmény. *Ha az $X_1, X_2, ..., X_n$ valószínűségi változók teljesen függetlenek, akkor az együttes eloszlás számítható a vetületi eloszlásokból:*

$$F_{(X_1,X_2,...,X_n)}(x_1,x_2,...,x_n) = \Pi_{i=1}^{n} F_{X_i}(x_i).$$

Az általunk korábban használt függetlenség a teljes függetlenség volt, ahol a szó szoros értelmében egymástól függetlenül elvégzett kísérleteken definiált valószínűségi változókat vizsgáltunk. Nézzünk olyan példát, amikor egyazon kísérleten definiált valószínűségi változók függetlenek:

Amikor bevezettük az események függetlenséget (a jegyzet 3. fejezetében), beláttuk, hogy a szabályos kockadobás esetén az $A = \{$5-öt vagy 6-ot dobunk$\}$ és a $B = \{$páratlant dobunk$\}$ események függetlenek. Ekkor viszont könnyű belátni, hogy az ő indikátorváltozóik is függetlenek lesznek. Vizsgáljuk ezt meg az alábbi táblázattal:

$I_B \backslash I_A$	0	1	Σ
0	$\frac{1}{3}$	$\frac{1}{6}$	$\frac{1}{2}$
1	$\frac{1}{3}$	$\frac{1}{6}$	$\frac{1}{2}$
Σ	$\frac{2}{3}$	$\frac{1}{3}$	1

A függetlenség a diszkrét esetben azzal ekvivalens, hogy az egyes cellák tartalma pontosan a vetületi eloszlások megfelelő értékeinek szorzata (azaz a sorösszegének és oszlopösszegének szorzata). Mivel ez minden egyes cellára teljesül, így a fentiek függetlenek. Például

$$P(I_A = 0, I_B = 0) = \frac{1}{3} = \frac{2}{3} \cdot \frac{1}{2} = P(I_A = 0)P(I_B = 0).$$

Ahogy az előző néhány példában már láttuk, a diszkrét együttes eloszlás táblázatosan megadható 2 valószínűségi változó esetén. Általánosan az alábbit tekintjük az együttes eloszlásnak:

13. Definíció. *Az $X_1, X_2, ..., X_n$ diszkrét valószínűségi változók* **együttes eloszlás***a a $P(X_1 = x_1, X_2 = x_2, ...X_n = x_n)$ típusú valószínűségek összessége. Ha $I = \{i_1, i_2, ...i_k\}$ részhalmaza $\{1, 2, ..., n\}$-nek, akkor a $P(X_{i_1} = x_{i_1}, X_{i_2} = x_{i_2}, ...X_{i_k} = x_{i_k})$ valószínűségek összessége egy vetületi eloszlást határoz meg.*

14. Következmény. *A diszkrét eloszlás összes valószínűségét összegezve 1-et kapunk (hiszen az a biztos esemény valószínűsége).*

Ahogy már az előző példában is kihasználtuk, diszkrét esetben a függetlenséget egyszerűbb a valószínűségeken ellenőrizni, nem pedig az eloszlásfüggvényeken. Ezt az alábbi tétel miatt tehetjük meg:

15. Tétel. *Az X_1 és X_2 diszkrét valószínűségi változók pontosan akkor függetlenek, ha bármely x_1, x_2 valós számpárra*

$$P(X_1 = x_1, X_2 = x_2) = P(X_1 = x_1)P(X_2 = x_2).$$

Ebből a páronként függetlenség egyértelműen ellenőrizhető, de a teljesen függetlenséghez az alábbira van szükség:

16. Tétel. *Az $X_1, X_2, ..., X_n$ diszkrét valószínűségi változók pontosan akkor teljesen függetlenek, ha bármely $I = \{i_1, i_2, ...i_k\}$ részhalmazára $\{1, 2, ..., n\}$-nek, és bármely $\{x_{i_1}, x_{i_2}, ..., x_{i_k}\}$ valós számsorozatra*

$$P(\Pi_{j\in I} X_j = x_j) = \Pi_{j\in I} P(X_j = x_j).$$

Emlékezzünk, hogy egy folytonos valószínűségi változó eloszlását legjobban a sűrűségfüggvényének segítségével tudtuk jellemezni. A többdimenziós esetre is tudunk sűrűségfüggvényt bevezetni, az alábbi módon:

17. Definíció. *Az $X_1, X_2, ..., X_n$ folytonos valószínűségi változók* **együttes sűrűségfüggvény***e az $f_{(X_1,X_2,...,X_n)}(x_1, x_2, ..., x_n)$ többváltozós integrálható függvény, ha*

$$F_{(X_1,X_2,...,X_n)}(x_1, x_2, ..., x_n) =$$
$$= \int_{-\infty}^{x_1} \int_{-\infty}^{x_2} \cdots \int_{-\infty}^{x_n} f_{(X_1,X_2,...,X_n)}(y_1, y_2, ..., y_n)\ dy_n dy_{n-1} \dots dy_1.$$

Legyenek $X, Y, Z \in U(0, 1)$ független valószínűségi változók. Ekkor az együttes eloszlásuk:

$$F_{(X,Y,Z)}(x, y, z) = P(X < x, Y < y, Z < z) = P(X < x)P(Y < y)P(Z < z).$$

Mindez a lényeges részen, azaz ahol $0 < x, y, z < 1$

$$F_{(X,Y,Z)}(x, y, z) = x \cdot y \cdot z.$$

Könnyen látható, hogy ekkor az $f_{(X,Y,Z)}(x, y, z) = 1$ ahol $0 < x, y, z < 1$ (egyébként 0) megfelelő együttes sűrűségfüggvény.

Persze a sűrűségfüggvény nem egyértelmű, így itt is érdemes kikötni, hogy szeretnénk minél kevesebb szakadású sűrűségfüggvényt választani (azaz csak ott szakadhat ahol szükséges). Ez azt jelenti, hogy az eloszlásfüggvény folytonossági pontjaiban a sűrűségfüggvény értéke az eloszlásfüggvény deriváltjaként kapható meg (az összes változó szerint deriválunk).

18. Definíció. *Az $X_1, X_2, ..., X_n$ folytonos valószínűségi változók együttes sűrűségfüggvényének egy k-dimenziós* **vetületi sűrűségfüggvény***e az $X_{i_1}, X_{i_2}, ...X_{i_k}$ változók együttes sűrűségfüggvénye (ha $\{i_1, i_2, ...i_k\}$ részhalmaza $\{1, 2, ..., n\}$-nek).*

Az előző példában látott $X, Y, Z \in U(0, 1)$ független valószínűségi változók együttes eloszlásfüggvényét láttuk, hogy $F_{(X,Y,Z)}(x, y, z) = x \cdot y \cdot z$ (a lényeges részen). Ugyan azzal a számolási menettel belátható, hogy az (X, Y) 2-dimenziós vetületi eloszlásfüggvény $F_{(X,Y)}(x, y) = x \cdot y$ (a lényeges részen). Ekkor az $f_{(X,Y,Z)}(x, y, z) = 1$ együttes sűrűségfüggvényhez nagyon hasonló lesz a vetületi sűrűségfüggvény: $f_{(X,Y)}(x, y) = 1$ (ahol $0 < x, y < 1$,egyébként 0).

19. Következmény. *A sűrűségfüggvényt a teljes tartományon integrálva 1-et kapunk (hiszen ez az integrál a biztos esemény valószínűsége). A sűrűségfüggvény nemnegatív (csak akkor lehetne negatív, ha nem kötnénk meg a minimális szakadási számot).*

A korábbi példákban látott $X, Y, Z \in U(0, 1)$ független valószínűségi változók együttes sűrűségfüggvényéről láttuk, hogy $f_{(X,Y,Z)}(x, y, z) = 1$ (ahol $0 < x, y, z < 1$,egyébként 0). Számoljuk ki a teljes tartomány valószínűségét:

$$P(\Omega) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{(X,Y,Z)}(x, y, z)dxdydz = \int_0^1\int_0^1\int_0^1 1dxdydz =$$

$$= \int_0^1\int_0^1 [x]_0^1 dydz = \int_0^1\int_0^1 1dydz = \int_0^1 [y]_0^1 dz = \int_0^1 1dz = [z]_0^1 = 1.$$

A vetületi sűrűségfüggvény könnyen számolható az együttesből, csupán a felesleges változók szerint ki kell integrálnunk a teljes tartományon:

20. Tétel. *Ha $\{i_1, i_2, ...i_k\}$ részhalmaza $\{1, 2, ..., n\}$-nek, és a komplementer részhalmaz $\{j_1, j_2, ..., j_{n-k}\}$. Ekkor:*

$$f_{(X_{i_1}, X_{i_2}, ..., X_{i_k})}(y_{i_1}, y_{i_2}, ..., y_{i_n}) =$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} f_{(X_1,X_2,...,X_n)}(y_1, y_2, ..., y_n)\ dy_{j_1}dy_{j_2}\ldots dy_{j_{n-k}}.$$

Az előző példákban látott $X, Y, Z \in U(0, 1)$ független valószínűségi változók együttes sűrűségfüggvénye $f_{(X,Y,Z)}(x, y, z) = 1$ (ahol $0 < x, y, z < 1$,egyébként 0), és az (X, Y) vetületi sűrűségfüggvény: $f_{(X,Y)}(x, y) = 1$ (ahol $0 < x, y < 1$,egyébként 0). Ellenőrizzük ezt le a tétel szerint is:

$$f_{(X,Y)}(x, y) = \int_{-\infty}^{\infty} f_{(X,Y,Z)}(x, y, z)dz = \int_0^1 1dz = [z]_0^1 = 1.$$

A folytonos valószínűségi változók függetlenségét sem mindig érdemes a definíció alapján (azaz az eloszlásfüggvényekből) megállapítani, itt azzal ekvivalensen a sűrűségfüggvények tudnak segíteni:

21. Tétel. *Az X_1 és X_2 folytonos valószínűségi változók pontosan akkor függetlenek, ha bármely x_1, x_2 valós számpárra $f_{(X_1,X_2)}(x_1,x_2) = f_{X_1}(x_1)f_{X_2}(x_2)$.*

Ebből a tételből már egyértelművé válhat mindenki számára, hogy az eddigi példáink, ahol a független $X, Y, Z \in U(0,1)$ valószínűségi változókat vizsgáltuk, miért egyszerűsödött le ennyire. Az együttes sűrűségfüggvény egyszerűen az ismert sűrűségfüggvények szorzata:

$$f_{(X,Y,Z)}(x,y,z) = f_X(x)f_Y(y)f_Z(z) = 1 \text{ (ahol } 0 < x,y,z < 1, \text{egyébként } 0\text{)}.$$

Valójában az előző tételnél többet használtunk ki a példában, hiszen nem kettő, hanem három változó együttes sűrűségfüggvényét bontottuk szorzatra. Ugyan a tétel jól használható a páronként függetlenség ellenőrzésére és felhasználására, de a teljesen függetlenséghez az alábbira van szükség (és ez az amit a fenti példa is használt):

22. Tétel. *Az $X_1, X_2, ..., X_n$ folytonos valószínűségi változók pontosan akkor teljesen függetlenek, ha bármely $I = \{i_1, i_2, ...i_k\}$ részhalmazára $\{1, 2, ..., n\}$-nek, és bármely $\{x_{i_1}, x_{i_2}, ..., x_{i_k}\}$ valós számsorozatra*

$$f_{(X_{i_1},X_{i_2},...X_{i_k})}(x_{i_1}, x_{i_2}, ...x_{i_k}) = \Pi_{j\in I} f_{X_j}(x_j).$$

Az eddigi szóhasználatunk, amikor valószínűségi változókról függetlenséget feltételeztünk, az valójában mindig a teljes függetlenség volt. Innentől kezdve precízen, minden esetben kiírjuk a függetlenség típusát.

Amennyiben az előző példákban páronként független (de nem feltétlen teljesen független) $X, Y, Z \in U(0,1)$ valószínűségi változókat vizsgáltunk volna, akkor lehetne közöttük valamilyen függés. Például Z binárisan felírva lehetne az X és Y bináris alakjainak bitenkénti XOR-ja (kizáró vagy). Belátható, hogy ekkor tényleg Z is $U(0,1)$ eloszlású, és hárman páronként függetlenek. Azonban az együttes eloszlásuk és sűrűségfüggvényük ez esetben nem egyezne meg a fent kiszámítottal: például

$$F_{(X,Y,Z)}\left(\frac{1}{2}, \frac{1}{2}, \frac{3}{4}\right) = \frac{1}{4} \neq \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{3}{4},$$

hiszen ha X és Y is kisebb mint $\frac{1}{2}$, akkor Z nem lehet nagyobb, mint $\frac{1}{2}$ (az első biten alkalmazott kizáró vagy miatt).

2. Konvolúció. Változók lineáris kombinációjának várható értéke, független változók összegének, különbségének szórása.

Az együttes eloszlások ismerete lehetőséget ad arra, hogy ne csak független valószínűségi változók esetén mondhassunk valamit a belőlük képzett valószínűségi változókról. A legegyszerűbb ilyen eset két valószínűségi változó összege, más néven **konvolúció**ja.

23. Tétel. *Ha X és Y diszkrét valószínűségi változók, és $Z = X + Y$, akkor*

$$P(Z = z) = \sum_{x \in R_X} P(X = x, Y = z - x).$$

24. Tétel. *Ha X és Y folytonos valószínűségi változók, együttes sűrűségfüggvényük $f_{(X,Y)}(x, y)$ és $Z = X + Y$, akkor*

$$f_Z(z) = \int_{-\infty}^{\infty} f_{(X,Y)}(t, z - t)dt.$$

25. Következmény. *Ha X és Y függetlenek, akkor diszkrét esetben*

$$P(X + Y = z) = \sum_{x \in R_X} P(X = x)P(Y = z - x),$$

míg folytonos esetben

$$f_{X+Y}(z) = \int_{-\infty}^{\infty} f_{(X)}(t)f_Y(z - t)dt.$$

Korábban már vizsgáltuk a példát, amiben két szabályos kockával dobunk, és X_1 a nagyobbik, X_2 pedig a kisebbik értéke. Az $\underline{X} = (X_1, X_2)$ valószínűségi vektorváltozó eloszlása az alábbi táblázatból kiolvasható:

X_2 \ X_1	1	2	3	4	5	6
1	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{2}{36}$
2	0	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{2}{36}$
3	0	0	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{2}{36}$
4	0	0	0	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{2}{36}$
5	0	0	0	0	$\frac{1}{36}$	$\frac{2}{36}$
6	0	0	0	0	0	$\frac{1}{36}$

Ekkor az $X_1 + X_2$ eloszlása a 23. Tétel képletének segítségével pont a balról jobbra emelkedő átlók összegzésével számolható. Például

$$P(X_1+X_2=6) = P(X_1=1, X_2=5)+P(X_1=2, X_2=4)+P(X_1=3, X_2=3)+$$

$$+P(X_1=4, X_2=2)+P(X_1=5, X_2=1) = 0+0+\frac{1}{36}+\frac{2}{36}+\frac{2}{36} = \frac{5}{36}.$$

$X_1 + X_2$	2	3	4	5	6	7	8	9	10	11	12
$P()$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

Nézzünk egy folytonos példát is. Legyen az X és Y folytonos valószínűségi változók együttes sűrűségfüggvénye $f_{(X,Y)}(x, y) = 4x \cdot y$ ha $0 < x, y < 1$ (egyébként 0). Mi lesz az $X + Y$ sűrűségfüggvénye?
A sűrűségfüggvényt a 24. Tétel képletének segítségével határozhatjuk meg:

$$f_Z(z) = \int_{-\infty}^{\infty} f_{(X,Y)}(t, z-t)dt = \int_{max(0,z-1)}^{min(1,z)} 4t(z-t)dt.$$

Figyeljük meg, hogy az integrálási határok elbonyolodtak. Mivel a sűrűségfüggvény mindkét változójában 0 és 1 között vett csak fel nem 0 értéket, így a konvolúciós képlet behelyettesítésénél figyelni kellett, hogy az alábbi négy kritérium teljesüljön: $0 < t$; $1 > t$; $0 < z - t$; $1 > z - t$. Ez a négy feltétel határozta meg az integrálási határokat. Így két esetre bomlik az integrál:

$$0 < z \leq 1 : f_Z(z) = \int_0^z 4t(z-t)dt = \left[2t^2 z - \frac{4}{3}t^3\right]_0^z = 2z^3 - \frac{4}{3}z^3 = \frac{2}{3}z^3$$

$$1 < z < 2 : f_Z(z) = \left[2t^2 z - \frac{4}{3}t^3\right]_{z-1}^1 = 2z - \frac{4}{3} - \left(2z(z-1)^2 - \frac{4}{3}(z-1)^3\right) =$$

$$= 2z - \frac{4}{3} - \left(2z^3 - 4z^2 + 2z - \frac{4}{3}z^3 + 4z^2 - 4z + \frac{4}{3}\right) =$$

$$= -2z^3 + \frac{4}{3}z^3 + 4z - \frac{8}{3} = 4z - \frac{2}{3}z^3 - \frac{8}{3}$$

Vegyük észre, hogy X és Y függetlenek, hiszen a 20. Tétel szerint $f_X(x) = \int_0^1 4xydy = [2y^2x]_0^1 = 2x$ (ahol $0 < x < 1$, egyébként 0) és szimmetrikusan $f_Y(y) = \int_0^1 4xydx = [2x^2y]_0^1 = 2y$ (ahol $0 < y < 1$, egyébként 0), azaz az együttes sűrűségfüggvény tényleg a vetületi sűrűségfüggvények szorzata ($4xy = 2x \cdot 2y$).

Ennek megfelelően alkalmazható a 25. Következmény. Láthatjuk, hogy a fenti integrálás esetén ez csak annyit jelent, hogy $4t(z-t) = 2t \cdot 2(z-t)$.

Ha az összegnél bonyolultabb dolgot szeretnénk művelni a valószínűségi változóinkkal, arra is vannak megfelelő képletek, de ezekre nem érdemes időt szánni, mivel a bizonyításuk rövidebb, mint a képletük. Amire még szükségünk lehet, az valószínűségi változókból generált valószínűségi változó várható értéke. Ez teljesen analóg módon kezelhető az 1 dimenziós esettel:

26. Tétel. *Legyenek $X_1, X_2, ..., X_n$ valószínűségi változók, $\underline{X} = (X_1, X_2, ..., X_n)$ és $Y = g(\underline{X})$ valószínűségi változó (ahol $g : \mathbf{R}^n \to \mathbf{R}$ függvény), aminek létezik várható értéke.*

- *Ha $X_1, X_2, ..., X_n$ folytonosak, akkor $E(Y) =$*
$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} g(x_1, x_2, ..., x_n) \cdot f_{\underline{X}}(x_1, x_2, ..., x_n) dx_1 dx_2 \dots dx_n.$$
- *Ha $X_1, X_2, ..., X_n$ diszkrétek, akkor $E(Y) =$*
$$\sum_{(x_1, x_2, ..., x_n) \in R_{\underline{X}}} g(x_1, x_2, ..., x_n) \cdot P(X_1 = x_1, X_2 = x_2, ..., X_n = x_n).$$

Emlékezzünk a teljes valószínűség tételére, ami azt mondta ki, hogy teljes eseményrendszer mentén feldarabolva egy eseményt, a valószínűségét megkaphatjuk a darabok valószínűségeinek összegéből. Az előző tételt felhasználva bebizonyíthatjuk ennek a tételnek egy folytonos alakját. Ehhez szükségünk lesz még a feltételes valószínűség kiterjesztésére:

27. Definíció. $P(A|\underline{X} = \underline{x}) = \lim_{\underline{\epsilon} \to 0} P(A|\underline{x} \le \underline{X} < \underline{x} + \underline{\epsilon})$

Emlékezzünk, hogy a feltételes valószínűség eredeti definíciójával nem lehetne folytonos $\underline{X}$ eloszlás esetén az $\underline{X} = \underline{x}$ eseményre feltételezni, hiszen az 0 valószínűségű. Ezt felhasználva már felírhatjuk a teljes valószínűség tétel folytonos alakját:

28. Tétel. *(Folytonos Teljes Valószínűség Tétel) Legyen $\underline{X}$ folytonos valószínűségi vektorváltozó $f_{\underline{X}}(\underline{x})$ sűrűségfüggvénnyel. Ekkor bármely A eseményre igaz az alábbi:*
$$P(A) = \int_{-\infty}^{\infty} P(A|\underline{X} = \underline{x}) f_{\underline{X}}(\underline{x}) d\underline{x}.$$

Bizonyítás: Használjuk fel a 26. Tételt $g(\underline{x}) = P(A|\underline{X} = \underline{x})$ esetén:

$$E(P(A|\underline{X} = \underline{X})) = \int_{-\infty}^{\infty} P(A|\underline{X} = \underline{x}) \cdot f_{\underline{X}}(\underline{x}) d\underline{x}.$$

A bal oldalt jobban megvizsgálva pontosan a tétel képletét kapjuk, mivel $\underline{X} = \underline{X}$ a biztos esemény, másrészt pedig konstansnak a várható értéke önmaga.

A Mosó márkájú mosógépek átlagosan 2 évig bírják az első meghibásodásig. András és Bíbor egyszerre vettek mosógépet. Mi annak a valószínűsége, hogy Bíbor gépe legalább 2 évvel később lesz először hibás, mint ahogy Andrásé? Legyen A valószínűségi változó András, B pedig Bíbor gépének első meghibásodásának ideje, ekkor $A, B \in Exp(\frac{1}{2})$. Legyen a C esemény, hogy Bíbor gépe legalább két évvel tovább bírja ($C = \{A + 2 < B\}$). Használjuk fel a folytonos teljes valószínűség tételt:

$$P(C) = \int_{-\infty}^{\infty} P(C|B = b) f_B(b) db = \int_{0}^{\infty} P(A < b - 2) \frac{1}{2} e^{-\frac{1}{2}b} db =$$

Írjuk be a valószínűséget az exponenciális eloszlás eloszlásfüggvényének segítségével, de figyeljünk, hogy $b < 2$ esetén 0-lesz (azaz az integrálási határon is módosítanunk kell):

$$= \int_{2}^{\infty} (1 - e^{-\frac{1}{2}(b-2)}) \frac{1}{2} e^{-\frac{1}{2}b} db = \frac{1}{2} \int_{2}^{\infty} e^{-\frac{1}{2}b} - e^{-b+1} db =$$

$$= \frac{1}{2} \left[-2 \cdot e^{-\frac{1}{2}b} + e^{-b+1} \right]_{2}^{\infty} = 0 - (-e^{-1} + \frac{1}{2} e^{-1}) = \frac{1}{2e}.$$

Mindezt az örökifjúság felhasználásával egyszerűbben is ki tudjuk számolni, ha a folytonos teljes valószínűség tétel előtt még alkalmazunk egy diszkrét teljes valószínűség tételt is:

$$P(C) = P(C|A > B)P(A > B) + P(C|A = B)P(A = B) + P(C|A < B)P(A < B)$$

Most használjuk ki, hogy $P(A = B) = 0$ hiszen független, folytonos valószínűségi változók. Így $P(A > B) = P(A < B) = \frac{1}{2}$, hiszen azonos eloszlásúak. Eközben $P(C|A > B) = 0$ hiszen ha Bíbor gépe előbb hibásodik meg, mint Andrásé, akkor nem tud két évvel később meghibásodni (először), mint Andrásé. Most használjuk fel a folytonos teljes valószínűség tételt, de most A-ra feltételezzünk:

$$P(C) = P(C|A < B)P(A < B) = \frac{1}{2} \int_{-\infty}^{\infty} P(A + 2 < B | A < B, A = a) f_A(a) da =$$

Most használjuk fel az örökifjúságot is:

$$= \frac{1}{2} \int_{-\infty}^{\infty} P(a + 2 < B | a < B) f_A(a) da = \frac{1}{2} \int_{-\infty}^{\infty} P(2 < B) f_A(a) da =$$

Mivel a valószínűség nem függ a-tól, kihozhatjuk az integrál elé. Így azonban egy sűrűségfüggvényt integrálunk a teljes tartományon, ami 1*:*

$$= \frac{1}{2}P(2 < B)\int_{-\infty}^{\infty} f_A(a)da = \frac{1}{2}P(2 < B) = \frac{1}{2e}.$$

29. Tétel. *Legyenek* $X_1, X_2, ..., X_n$ *véges várható értékű valószínűségi változók, és* $a_1, a_2, ..., a_n$ *valós számok. Ekkor*

$$E(a_1X_1 + a_2X_2 + ... + a_nX_n) = a_1E(X_1) + a_2E(X_2) + ... + a_nE(X_n).$$

Egy szerencsejátékban egymás után 10-szer dobunk szabályos kockával, és minden dobás után megkapjuk a dobás értékének annyiszorosát forintban, ahányadik volt a dobás. Mi a várható X *nyereményünk?*
Legyen K *egy szabályos kockadobás értéke. A fenti tétel alapján a kérdéses várható érték az alábbi:*

$$E(X) = E(K) + 2E(K) + ... + 10(E(K)) = 55E(K) = 55\frac{7}{2} = 192{,}5$$

Azaz ha mondjuk 190 *forint lenne a beszálló a játékba, akkor megérné játszani, míg* 195 *forintért már nem.*

Hasonlót már korábban is említettünk, hogy a várható érték összegre és konstansszorosra jól viselkedik. Most nézzük meg, hogy mi a helyzet, ha nem konstanssal, hanem másik valószínűségi változóval akarnánk szorozni:

30. Tétel. *Legyenek* $X_1, X_2, ..., X_n$ *teljesen független, véges várható értékű valószínűségi változók. Ekkor*

$$E(\Pi_{i=1}^n X_i) = \Pi_{i=1}^n E(X_i).$$

Egy szerencsejátékban egymás után 10-szer dobunk szabályos kockával (egymástól teljesen függetlenül), és az X *nyereményünk a dobások szorzata. Mi a várható nyereményünk?*
Legyen K *egy szabályos kockadobás értéke. A fenti tétel alapján a kérdéses várható érték az alábbi:*

$$E(X) = E(K)^{10} = \left(\frac{7}{2}\right)^{10} = 275854{,}74$$

Azaz ha mondjuk 276 *ezer forint lenne a beszálló a játékba, akkor nem érné meg játszani, míg* 275 *ezer forintért már igen (feltéve, hogy merünk kockáztatni ekkora összeget).*

Transzformált valószínűségi változó szórását vizsgálva csak azt láttuk korábban, hogy a konstans szorzó abszolút értékben kiszervezhető (és hogy additív konstans eltűnik). Még valószínűségi változók összegére sem tudtunk mit mondani, azonban az előző tételt felhasználva már adódik az alábbi:

31. Tétel. *Legyenek $X_1, X_2, ..., X_n$ páronként független, véges szórású valószínűségi változók és $a_1, a_2, ..., a_n$ valós számok. Ekkor*

$$\sigma^2(a_1X_1 + a_2X_2 + ... + a_nX_n) = a_1^2\sigma^2(X_1) + a_2^2\sigma^2(X_2) + ... + a_n^2\sigma^2(X_n).$$

Ennek a tételnek a segítségével az előző példa játékának szórását nem fogjuk tudni meghatározni, de az azelőttiét igen. Azaz egymás után 10-szer dobunk szabályos kockával, és minden dobás után megkapjuk a dobás értékének annyiszorosát forintban, ahányadik volt a dobás. Ekkor a fenti tétel alapján a nyeremény szórásnégyzete az alábbi:

$$\sigma^2(X) = \sigma^2(K) + 4\sigma^2(K) + ... + 100\sigma^2(K) = 385\sigma^2(K) = 385\frac{35}{12} \approx 1122{,}92$$

A fejezetet nevezetes konvolúciókkal zárjuk:

32. Tétel. *Legyenek $X \in Bin(n_1, p)$ és $Y \in Bin(n_2, p)$ független valószínűségi változók. Ekkor $X + Y \in Bin(n_1 + n_2, p)$.*

33. Tétel. *Legyenek $X \in Poi(\lambda_1)$ és $Y \in Poi(\lambda_2)$ független valószínűségi változók. Ekkor $X + Y \in Poi(\lambda_1 + \lambda_2)$.*

34. Tétel. *Legyenek $X \in N(\mu_1, \sigma_1)$ és $Y \in N(\mu_2, \sigma_2)$ független valószínűségi változók. Ekkor $X + Y \in N(\mu_1 + \mu_2, \sqrt{\sigma_1^2 + \sigma_2^2})$.*

3. Kovariancia, korrelációs együttható. A korrelációs együttható és a kovariancia tulajdonságai. Kapcsolat a függetlenség és a korrelálatlanság között.

Az előző fejezetben láttuk, hogy több valószínűségi változó viszonya lehet független és nem független. A függetlenség esetén a peremeloszlások definiálják az együttest, de ha nem függetlenek, akkor csak az együttes eloszlás segítségével tudunk összetett eseményeket vizsgálni (azaz olyan eseményeket, amik mindkettő változó értékétől függnek). Arról eddig nem esett szó, hogy milyen mértékű a függés.

Legyen X és Y két $(0,1)$-en egyenletes eloszlású független valószínűségi változó. Ha $Z_1 = X$, akkor egyértelmű, hogy X és Z_1 nem független, sőt naivan azt mondanánk, hogy nagyon erős a függés kettejük között. Ha $Z_2 = \frac{1}{10}X + \frac{9}{10}Y$, akkor X és Z_2 között kisebb mértékű a függés, mint Y és Z_2 között, és mindkettő kisebb, mint az X és Z_1 közötti. (Könnyű látni, hogy $Z_2 \in (0,1)$, de ez sajnos nem lesz egyenletes.)

A következő definíciók segítségével egyfajta mértéket kapunk az összefüggőségre.

35. Definíció. *Legyen X és Y két valószínűségi változó, aminek létezik szórása. Ekkor az X és Y* **kovarianciája***:*

$$cov(X,Y) = E\left((X - EX)\cdot(Y - EY)\right).$$

Nem véletlen, ha ez a formula emlékeztet minket valamire, hiszen a szórásnégyzet definíciója nagyon hasonló:

36. Következmény. $cov(X,X) = \sigma^2(X)$

Bizonyítás:

$$cov(X,X) = E\left((X - EX)\cdot(X - EX)\right) = E\left((X - EX)^2\right) = \sigma^2(X)$$

A szórásnégyzet képletét egyszerűen átalakítottuk a Steiner-formulával, amivel analóg módon most is ekvivalens formulát kaphatunk:

37. Következmény. *(Steiner)* $cov(X,Y) = E(XY) - EXEY$

Bizonyítás: A várható érték linearitását felhasználva

$$cov(X,Y) = E\left((X - EX)\cdot(Y - EY)\right) = E\left(XY - YEX - XEY + EXEY\right) =$$

$$= E(XY) - E(YEX) - E(XEY) + E(EXEY) = E(XY) - EXEY.$$

Az alábbi tulajdonságok mind egyszerű következményei a várható érték tulajdonságainak.

38. Tulajdonság(ok). *Legyenek a, b, c, d valós számok, és X, Y, Z három valószínűségi változó, aminek létezik szórása. Ekkor:*

1. $cov(X, Y) = cov(Y, X)$,
2. $cov(aX + b, cY + d) = a \cdot c \cdot cov(X, Y)$,
3. $cov(X, a) = 0$,
4. $cov(X + Y, Z) = cov(X, Z) + cov(Y, Z)$.

Ahogy már említettük, a kovariancia egyfajta függőségi mérték, így nem meglepő az alábbi következmény:

39. Következmény. *Ha X és Y függetlenek, akkor $cov(X, Y) = 0$*

Bizonyítás: A függetlenség miatt $E(XY) = EXEY$, így:

$$cov(X, Y) = E(XY) - EXEY = EXEY - EXEY = 0.$$

Legyenek X, Y, Z_1, Z_2 az előző példában definiált valószínűségi változók. Tudjuk, hogy ezek várható értéke $\frac{1}{2}$, X, Y, Z_1 szórásnégyzete $\frac{1}{12}$ és második momentuma $\frac{1}{3}$ (hiszen Steiner-formula segítségével $\frac{1}{12} + \left(\frac{1}{2}\right)^2$). Ekkor

$$cov(X, Y) = 0,$$

$$cov(X, Z_1) = cov(X, X) = \sigma^2(X) = \frac{1}{12},$$

$$cov(X, Z_2) = E(XZ_2) - EXEZ_2 = E(X(\frac{1}{10}X + \frac{9}{10}Y)) - \frac{1}{4} =$$

$$\frac{1}{10}E(X^2) + \frac{9}{10}EXEY - \frac{1}{4} = \frac{1}{30} - \frac{1}{40} = \frac{1}{120},$$

$$cov(Y, Z_2) = E(YZ_2) - EYEZ_2 = E(Y(\frac{1}{10}X + \frac{9}{10}Y)) - \frac{1}{4} =$$

$$= \frac{1}{10}EXEY + \frac{9}{10}E(Y^2) - \frac{1}{4} = \frac{1}{40} + \frac{9}{30} - \frac{1}{4} = \frac{9}{120}.$$

Ez egész közel áll a naiv elképzelésünkhöz, mert tényleg pontosan akkor nagyobb az érték, amikor nagyobb összefüggést gondoltunk. Azonban jobban örülnénk, ha X-nek az önmagával vett összefüggése nem $\frac{1}{12}$ volna, hanem valami X-től független konstans.

Az alábbi definíció egyfajta normálása a kovarianciának, így általánosabban tekinthető összefüggőségi mértéknek:

40. Definíció. *Legyen X és Y két valószínűségi változó, aminek létezik szórása, és $\sigma_X, \sigma_Y > 0$. Ekkor az X és Y* **korrelációs együttható***ja:*

$$R(X,Y) = \frac{cov(X,Y)}{\sigma(X)\sigma(Y)}.$$

41. Következmény. $R(X,X) = 1$

Bizonyítás:
$$R(X,X) = \frac{cov(X,X)}{\sigma(X)\sigma(X)} = \frac{\sigma^2(X)}{\sigma(X)\sigma(X)} = 1$$

Amikor a korrelációs együttható értéke 0, akkor azt mondjuk, hogy a valószínűségi változók **korrelálatlanok**.

42. Következmény. *Ha X és Y pozitív szórású függetlenek, akkor korrelálatlanok is (azaz $R(X,Y) = 0$).*

Bizonyítás: A függetlenség miatt $cov(X,Y) = 0$, amiből azonnal látható, hogy a korrelációs együttható is 0.

Legyenek az X, Y, Z_1, Z_2 az előző példákban látott valószínűségi változók. A kovarianciákat már kiszámoltuk, így a korrelációs együtthatók a szórások segítségével könnyen számolhatók. Z_2 szórásnégyzete:

$$\sigma^2 Z_2 = \frac{1}{100}\sigma^2 X + \frac{81}{100}\sigma^2 Y = \frac{82}{1200} = \frac{41}{600}.$$

$$R(X,Y) = \frac{cov(X,Y)}{\sigma(X)\sigma(Y)} = 0,$$

$$R(X,Z_1) = \frac{cov(X,X)}{\sigma(X)\sigma(X)} = \frac{\frac{1}{12}}{\frac{1}{\sqrt{12}}\frac{1}{\sqrt{12}}} = 1,$$

$$R(X,Z_2) = \frac{cov(X,Z_2)}{\sigma(X)\sigma(Z_2)} = \frac{\frac{1}{120}}{\frac{1}{\sqrt{12}}\sqrt{\frac{41}{600}}} = \frac{1}{\sqrt{82}},$$

$$R(Y,Z_2) = \frac{cov(Y,Z_2)}{\sigma(Y)\sigma(Z_2)} = \frac{\frac{9}{120}}{\frac{1}{\sqrt{12}}\sqrt{\frac{41}{600}}} = \frac{9}{\sqrt{82}}.$$

Ezek a számok már pontosabban írják le, amit szerettünk volna, ahol 1 jelenti a teljes függést. (Figyeljük meg, hogy Z_2-nek két egymástól független valószínűségi változóval való korrelációjának az összege 1-nél nagyobb, ami kicsit ellentmond az elképzeléseinknek.)

Sajnos a kovariancia és korreláció mindezek ellenére nem tökéletes a függetlenség mérésére, mert a 42. Következmény megfordítása nem igaz.

43. Tétel. *Annak ellenére, hogy $R(X, Y) = 0$, lehetséges, hogy X és Y nem független.*

Bizonyítás: Nézzünk egy példát, ami igazolja a tétel állítását:
Legyen $X \in U(-1, 1)$ és $Y = X^2$. Ekkor természetesen X és Y nem függetlenek, valamint $E(X) = 0$. Lássuk mi a helyzet a kovarianciával:

$$cov(X, Y) = E(XY) - EXEY = E(X^3) - EXE(X^2) =$$

$$= \int_{-1}^{1} t^3 \cdot \frac{1}{2} dt - 0 \cdot E(X^2) = \left[\frac{t^4}{8}\right]_{-1}^{1} = \frac{1}{8} - \frac{1}{8} = 0.$$

Ekkor persze a korreláció is 0.

44. Tétel. *Ha X és Y valószínűségi változók standardizáltjai X' és Y', akkor $R(X, Y) = cov(X', Y') = R(X', Y')$.*

Bizonyítás:

$$R(X', Y') = \frac{cov(X', Y')}{\sigma(X')\sigma(Y')} = \frac{E(X'Y') - EX'EY'}{1 \cdot 1} = E(X'Y') - 0 =$$

$$= E\left(\frac{(X - EX)(Y - EY)}{\sigma(X)\sigma(Y)}\right) = \frac{E\left((X - EX)(Y - EY)\right)}{\sigma(X)\sigma(Y)} =$$

$$\frac{cov(X, Y)}{\sigma(X)\sigma(Y)} = R(X, Y).$$

45. Tétel. *Tegyük fel, hogy az X és Y valószínűségi változóknak létezik szórása. Ekkor*

$$\sigma^2(X \pm Y) = \sigma^2(X) + \sigma^2(Y) \pm 2cov(X, Y).$$

Bizonyítás: $\sigma^2(X \pm Y) =$

$$= E((X \pm Y)^2) - E^2(X \pm Y) = E(X^2 \pm 2XY + Y^2) - E^2(X) \mp 2EXEY - E^2(Y) =$$

$$= E(X^2) - E^2(X) + E(Y^2) - E^2(Y) \pm 2E(XY) \mp 2EXEY =$$

$$= \sigma^2(X) + \sigma^2(Y) \pm 2cov(X, Y).$$

46. Következmény. *Tegyük fel, hogy az $X_1, X_2, ..., X_n$ valószínűségi változóknak létezik szórása. Ekkor*

$$\sigma^2\left(\sum_{i=1}^{n} X_i\right) = \sum_{i,j=1}^{n} cov(X_i, X_j) = \sum_{i=1}^{n} \sigma^2(X_i) + \sum_{i,j=1;i\neq j}^{n} cov(X_i, X_j).$$

47. Tétel. *Tegyük fel, hogy az X és Y valószínűségi változóknak pozitív a szórása. Ekkor $-1 \leq R(X,Y) \leq 1$.*

Bizonyítás: Legyen X' és Y' a két valószínűségi változónk standardizáltja. Ekkor a korábbiak szerint:

$$0 \leq \sigma^2(X' \pm Y') = \sigma^2(X') + \sigma^2(Y') \pm 2cov(X', Y') = 2 \pm 2R(X,Y).$$

Külön vizsgálva a plusz és mínusz eseteket azonnal következik, hogy $-1 \leq R(X,Y)$, és $R(X,Y) \leq 1$.

Ahogyan korábban láttuk, a korreláció 0 értéke szükséges, de nem elégséges feltétele a függetlenségnek. Ennél szerencsésebb a helyzet a másik véglettel, azaz amikor a korreláció abszolút értéke 1:

48. Tétel. *Tegyük fel, hogy az X és Y valószínűségi változóknak pozitív a szórása. Ekkor $R(X,Y) = \pm 1$ akkor és csak akkor teljesül, ha létezik a, b valós számpár, hogy $P(Y = aX + b) = 1$ (ahol $R(X,Y)$ előjele az a előjelével egyezik meg).*

Valószínűségi vektorváltozók esetén a várható érték és a szórás szerepét az alábbi struktúrák veszik át:

49. Definíció. *Az $\underline{X} = (X_1, X_2, ..., X_n)$ valószínűségi vektorváltozó* **várható érték vektora** *$E\underline{X} = (EX_1, EX_2, ..., EX_n)$,* **kovarianciamátrixa** *pedig az a mátrix, aminek az i-edik sorának j-edik eleme $cov(X_i, X_j)$.*

50. Következmény. *A 46. Következmény szerint az $X_1, X_2, ..., X_n$ valószínűségi változók összegének szórásnégyzete éppen a kovarianciamátrix elemeinek az összege, hiszen:*

$$\sigma^2\left(\sum_{i=1}^{n} X_i\right) = \sum_{i,j=1}^{n} cov(X_i, X_j).$$