6. Valószínűségi változók viszonya

Amikor valószínűségi változókról volt szó, mindig egymagukban vizsgáltuk az egyes példákat. Ilyenkor elégséges volt az eloszlásukkal foglalkozni, azaz diszkrét esetben a $\mathbb{P}(X = k)$ alakú valószínűségek sorozatát, folytonos esetben pedig az F_X eloszlásfüggvényt vagy az f_X sűrűségfüggvényt nézni. Az eloszlás minden lényeges tulajdonságot elmondott a valószínűségi változóról.

De ne keverjük össze az eloszlást magával a valószínűségi változóval: attól, hogy 100 dobásból mind a dobott fejek száma, mind a dobott írások száma $B(100; \frac{1}{2})$ binomiális eloszlású, még nyilván nem mondhatjuk, hogy mindig ugyanannyi fejet dobnánk, mint írást. Ez a megkülönböztetés különösen lényeges, ha több valószínűségi változóról beszélünk egyszerre.

Ebben a fejezetben két valószínűségi változó függetlenségét, illetve lineáris összefüggőségük mértékét vizsgáljuk.

6.1. Függetlenség

Események függetlenségének a fogalmát már bevezettük a 2. előadáson: A és B események függetlenek, ha $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$. Definiáljuk most ezt felhasználva valószínűségi változók függetlenségét.

⚠ **6.1.1. Definíció.** Legyenek X és Y valószínűségi változók az $(\Omega, \mathcal{F}, \mathbb{P})$ valószínűségi mezőn. Azt mondjuk, hogy X és Y **függetlenek**, ha minden $x, y \in \mathbb{R}$ esetén az $\{X < x\}$ és $\{Y < y\}$ események függetlenek.

A valószínűségi mezőt azért kell emlegessük, mert előfordulhatna, hogy $X : \Omega_1 \to \mathbb{R}$, míg $Y : \Omega_2 \to \mathbb{R}$ alakú függvény, azaz más valószínűségi mezőn vannak értelmezve. Ilyen esetben nem tudunk X és Y függetlenségéről beszélni, mert „más világban élnek”.

6.1.2. Példa. Egy kockadobás eredménye és a ma leeső csapadék mennyisége, amik intuitívan függetlenek, a fenti definíció értelmében is függetlenek, ahogy ezt események függetlenségénél már megjegyeztük.

De a függetlenség nem minden esetben ilyen nyilvánvaló. Például legyen Z egy egyenletes eloszlású valószínűségi változó az $\{1, 2, 3, \dots, 11, 12\}$ halmazon, jelölje X a Z hármas maradékát, és Y a Z négyes maradékát. Belátható, hogy ekkor X és Y függetlenek.

Hogyan tudjuk ellenőrizni két valószínűségi változó függetlenségét? Ebben az előadásban a diszkrét esetre koncentrálunk. Ekkor a következő állítás szolgáltat módszert a függetlenség ellenőrzésére.

6.1.3. Állítás. *Két diszkrét valószínűségi változó pontosan akkor független, ha minden $x, y \in \mathbb{R}$ esetén az $\{X = x\}$ és $\{Y = y\}$ események függetlenek, azaz*

$$\mathbb{P}(\{X = x\} \cap \{Y = y\}) = \mathbb{P}(X = x)\mathbb{P}(Y = y).$$

Megjegyzés. A definícióból az is következik, hogy minden X-szel és Y-nal kifejezhető halmazpár független, például $\{X = x\}$ és $\{1 \leq Y \leq 5\}$ független események. Általánosan: nézhetjük az X által **generált σ-algebrát**, azaz a legkisebb olyan – $\sigma(X)$-el jelölt – halmazt, aminek elemei az $\{X < x\}$ események ($x \in \mathbb{R}$), és teljesíti a σ-algebra definícióját. Ekkor X és Y függetlensége ekvivalens azzal, hogy bármilyen $A \in \sigma(X)$ és $B \in \sigma(Y)$ események függetlenek.

Fontos különbség az események függetlenségével szemben, hogy eseményekre ugyanannyi fáradtság volt leellenőrizni a függetlenséget és a nem-függetlenséget, hiszen mindkét esetben csak ki kellett számoljuk a metszet, illetve a két esemény külön-külön vett valószínűségét. Ezzel szemben valószínűségi változókra a függetlenséget cáfolni általában jóval egyszerűbb, mint igazolni: ha találunk egy $\{X = x\}$ és $\{Y = y\}$ eseményt, amelyek nem függetlenek, akkor a valószínűségi változók sem azok.

Miért tud hasznos lenni a függetlenség? Például, mert segíthet kiszámolni a várható értéket.

⚠ **6.1.4. Állítás.** *Ha X és Y független valószínűségi változók, és $\mathbb{E}(XY)$, $\mathbb{E}(X)$ és $\mathbb{E}(Y)$ létezik, akkor*

$$\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y).$$

Bizonyítás. Csak arra az esetre bizonyítunk, amikor X és Y egyszerű valószínűségi változók. Az általános eset határátmenet segítségével igazolható, ettől itt eltekintünk.

Először legyen X és Y indikátor valószínűségi változó, azaz $X = \mathbf{1}_A$ és $Y = \mathbf{1}_B$ valamilyen A és B eseményekre. Ekkor

$$\mathbb{E}(\mathbf{1}_A \mathbf{1}_B) = \mathbb{E}(\mathbf{1}_{A\cap B}) = \mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B) = \mathbb{E}(\mathbf{1}_A)\mathbb{E}(\mathbf{1}_B),$$

vagyis az állítás ebben a speciális esetben teljesül.

Nézzük az általánosabb esetet: tegyük fel, hogy X és Y egyszerű valószínűségi változó. Ekkor X és Y felírható indikátor valószínűségi változók lineáris kombinációjaként:

$$X = \sum_{k\in \mathrm{Ran}(X)} k \cdot \mathbf{1}_{\{X=k\}} \qquad \text{és} \qquad Y = \sum_{l\in \mathrm{Ran}(Y)} l \cdot \mathbf{1}_{\{Y=l\}}.$$

Az előző bekezdést és a várható érték additivitását felhasználva kapjuk, hogy

$$\mathbb{E}(XY) = \mathbb{E}\Bigg(\sum_{k\in \mathrm{Ran}(X)} k \cdot \mathbf{1}_{\{X=k\}} \sum_{l\in \mathrm{Ran}(Y)} l \cdot \mathbf{1}_{\{Y=l\}}\Bigg) = \sum_{k\in \mathrm{Ran}(X)} \sum_{l\in \mathrm{Ran}(Y)} k \cdot l \cdot \mathbb{E}\big(\mathbf{1}_{\{X=k\}}\mathbf{1}_{\{Y=l\}}\big) =$$

$$= \sum_{k\in \mathrm{Ran}(X)} \sum_{l\in \mathrm{Ran}(Y)} k \cdot l \cdot \mathbb{E}\big(\mathbf{1}_{\{X=k\}}\big)\mathbb{E}\big(\mathbf{1}_{\{Y=l\}}\big) = \mathbb{E}\Bigg(\sum_{k\in \mathrm{Ran}(X)} k \cdot \mathbf{1}_{\{X=k\}}\Bigg)\mathbb{E}\Bigg(\sum_{l\in \mathrm{Ran}(Y)} l \cdot \mathbf{1}_{\{Y=l\}}\Bigg),$$

ahol a jobb oldal éppen $\mathbb{E}(X)\mathbb{E}(Y)$, ahogy állítottuk. □

Megjegyzés. Felmerülhetne, hogy miért nem az állításban szereplő, kellemesebb egyenlettel definiáltuk valószínűségi változók függetlenségét. Azért, mert az $\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y)$ teljesülése gyengébb tulajdonság, nem következik belőle a valószínűségi változók függetlensége. Amit ehelyett felhasználhatnánk, az a következő állítás: ha minden nemnegatív valós f és g függvények esetén $\mathbb{E}(f(X)g(Y)) = \mathbb{E}(f(X))\mathbb{E}(g(Y))$, akkor X és Y függetlenek.

6.2. Diszkrét együttes eloszlás

Diszkrét valószínűségi változók esetén a függetlenségük vizsgálatához a $\mathbb{P}(X = k,\ Y = l)$ (azaz a $\mathbb{P}(\{X = k\} \cap \{Y = l\})$) valószínűségekre, vagyis a változók úgynevezett együttes eloszlására van szükségünk. (A folytonos esettel a 8. előadáson fogunk foglalkozni.)

6.2.1. Példa. Legyen X és Y olyan valószínűségi változók, ahol $\mathrm{Ran}(X) = \{2, 3, 5\}$, $\mathrm{Ran}(Y) = \{0, 1, 2\}$, és a $\mathbb{P}(X = k,\ Y = l)$ valószínűségeket a következő táblázat foglalja össze. Független-e X és Y, illetve mennyi $\mathbb{E}(XY)$?

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

Egy fentihez hasonló táblázattal megadott együttes eloszlás pontosan akkor lehet két valószínűségi változó együttes eloszlása, ha a benne szereplő számok nemnegatívak, és **összegük** 1. Leellenőrizhetjük, hogy ez a példában teljesül.

A függetlenség kiszámolásához szükségünk van a $\mathbb{P}(X = k)$ illetve a $\mathbb{P}(Y = l)$ mennyiségekre, vagyis az X és Y eloszlására.

6.2.2. Definíció. Legyenek X és Y egyszerű valószínűségi változók. Ha adott X és Y **együttes eloszlás**a, vagyis a $\mathbb{P}(X = k,\ Y = l)$ valószínűség minden $k \in \mathrm{Ran}(X)$ és $l \in \mathrm{Ran}(Y)$ esetén, akkor X és Y eloszlásait az együttes eloszlás **marginális eloszlás**ainak nevezzük.

A marginális eloszlásokat a valószínűség additivitása miatt a következőképp számolhatjuk ki:

$$\mathbb{P}(X=k)=\sum_{l\in \mathrm{Ran}(Y)}\mathbb{P}(X=k,\ Y=l)\qquad \mathbb{P}(Y=l)=\sum_{k\in \mathrm{Ran}(X)}\mathbb{P}(X=k,\ Y=l),$$

vagyis a táblázat sor- és oszlopösszegei adják az X és Y valószínűségi változók eloszlásait. A példa esetében így

$$\mathbb{P}(X=2)=0{,}2\ ,\qquad \mathbb{P}(X=3)=0{,}55\ ,\qquad \mathbb{P}(X=5)=0{,}25\ ,$$
$$\mathbb{P}(Y=0)=0{,}3\ ,\qquad \mathbb{P}(Y=1)=0{,}4\ ,\qquad \mathbb{P}(Y=2)=0{,}3.$$

Tehát a függetlenség definíciójából adódóan X és Y nem független, hiszen például $\mathbb{P}(X=5,\ Y=0)=0{,}1$, de $\mathbb{P}(X=5)\cdot\mathbb{P}(Y=0)=0{,}075$.

Számoljuk ki a fenti példában szereplő X és Y esetén az $\mathbb{E}(XY)$ mennyiséget is. Ehhez új definícióra nincs szükség, hiszen XY valószínűségi változó, értékkészlete $\{k\cdot l \mid k\in \mathrm{Ran}(X),\ l\in \mathrm{Ran}(Y)\}$. Így

$$\mathbb{E}(XY)=\sum_{m\in \mathrm{Ran}(XY)} m\cdot\mathbb{P}\Big(\bigcup_{\substack{k\in \mathrm{Ran}(X)\\ l\in \mathrm{Ran}(Y)\\ k\cdot l=m}}\{X=k,\ Y=l\}\Big)=\sum_{k\in \mathrm{Ran}(X)}\sum_{l\in \mathrm{Ran}(Y)} k\cdot l\cdot\mathbb{P}(X=k,\ Y=l)=$$
$$=0\cdot 0{,}05+0\cdot 0{,}15+0\cdot 0{,}1+2\cdot 0{,}1+3\cdot 0{,}2+5\cdot 0{,}1+4\cdot 0{,}05+6\cdot 0{,}2+10\cdot 0{,}05=3{,}2.$$

Vegyük észre, hogy ugyan a változók nem függetlenek, az $\mathbb{E}(XY)$ mennyiség így is kiszámolható.

6.3. Kovariancia

Ahogy az a példa esetében is látható, nem független valószínűségi változók esetében is lehet az összefüggésük mértéke alacsony (azaz intuitívan a $\mathbb{P}(X=k)\mathbb{P}(Y=l)$ szorzatok elég közel vannak az egyes $\mathbb{P}(X=k,\ Y=l)$ valószínűségekhez). Hogyan tudnánk mérni valószínűségi változók összefüggésének fokát? Erre több lehetőség is van,[27] kezdjük a kovariancia fogalmával.

6.3.1. Definíció. Az X és Y valószínűségi változók **kovarianciá**ját a következőképp definiáljuk:

$$\mathrm{cov}(X,Y)\overset{\mathrm{def}}{=}\mathbb{E}\big((X-\mathbb{E}X)(Y-\mathbb{E}Y)\big),$$

feltéve, hogy a várható érték létezik és véges.

6.3.2. Állítás. *Ha* $\mathrm{cov}(X,Y)$ *értelmes, akkor* $\mathrm{cov}(X,Y)=\mathbb{E}(XY)-\mathbb{E}(X)\mathbb{E}(Y)$.

Bizonyítás. A definíciót kibontva kapjuk, hogy

$$\mathbb{E}\big((X-\mathbb{E}X)(Y-\mathbb{E}Y)\big)=\mathbb{E}(XY)-\mathbb{E}\Big(\mathbb{E}(X)Y\Big)-\mathbb{E}\Big(X\mathbb{E}(Y)\Big)+\mathbb{E}\Big(\mathbb{E}(X)\mathbb{E}(Y)\Big)=$$
$$=\mathbb{E}(XY)+(-1-1+1)\mathbb{E}(X)\mathbb{E}(Y)=\mathbb{E}(XY)-\mathbb{E}(X)\mathbb{E}(Y),$$

ami épp a belátandó állítás. □

6.3.3. Következmény. *Legyen* X *és* Y *valószínűségi változó, amire* $\mathrm{cov}(X,Y)$ *értelmes.*

(1) Ha Y *konstans, akkor* $\mathrm{cov}(X,Y)=0$.
(2) Ha X *és* Y *függetlenek, akkor* $\mathrm{cov}(X,Y)=0$.
(3) Attól, hogy $\mathrm{cov}(X,Y)=0$, *még nem feltétlenül teljesül, hogy* X *és* Y *független.*

Bizonyítás. Jelölje Y konstans értékét $c\in\mathbb{R}$. Az előző állítás szerint

$$\mathrm{cov}(X,Y)=\mathbb{E}(XY)-\mathbb{E}(X)\mathbb{E}(Y)=\mathbb{E}(Xc)-\mathbb{E}(X)c=0.$$

A második ponthoz felhasználhatjuk az előző alfejezet második állítását, így $\mathrm{cov}(X,Y)=\mathbb{E}(XY)-\mathbb{E}(X)\mathbb{E}(Y)=0$.

---

[27] lásd még például: mediántól vett átlagos abszolút eltérés (mean absolute error); távolság-kovariancia.

A harmadik állításhoz legyen $\mathrm{Ran}(X) = \{-1, 0, 1\}$, amely értékeket $\frac{1}{4}$, $\frac{1}{2}$ és $\frac{1}{4}$ valószínűségekkel veszi fel X. Legyen $Y = |X|$. Kiszámolható, hogy $\mathrm{cov}(X, Y) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y) = 0 - 0 \cdot \frac{1}{2} = 0$, pedig a változók nem függetlenek, hiszen $\mathbb{P}(X = 0)\mathbb{P}(Y = 1) = \frac{1}{2} \cdot \frac{1}{2}$, míg $\mathbb{P}(X = 0, Y = 1) = 0$. □

6.3.4. Példa.

(1) Már láttuk, hogy $\mathbb{E}(XY) = 3{,}2$. Kiszámolható, hogy $\mathbb{E}(X) = 3{,}8$ és $\mathbb{E}(Y) = 1$, így $\mathrm{cov}(X, Y) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y) = 3{,}2 - 3{,}8 \cdot 1 = -0{,}6$.

(2) Legyen X egyenletes eloszlású az $\{1, 2, \ldots, 10\}$ halmazon, illetve Y egyenletes eloszlású az $\{1, -1\}$ halmazon. Tegyük fel, hogy X és Y függetlenek. Ekkor

$$\mathrm{cov}(X,\ 0{,}9 \cdot X + 0{,}1 \cdot Y) = \mathbb{E}(0{,}9 \cdot X^2 + 0{,}1 \cdot XY) - \mathbb{E}(X)\mathbb{E}(0{,}9 \cdot X + 0{,}1 \cdot Y) =$$
$$= 0{,}9 \cdot \mathbb{E}(X^2) + 0{,}1 \cdot \mathbb{E}(XY) - 0{,}9 \cdot \mathbb{E}(X)^2 - 0{,}1 \cdot \mathbb{E}(XY)$$
$$= 0{,}9 \sum_{k=1}^{10} k^2 \frac{1}{10} - 0{,}9 \cdot \left(\frac{11}{2}\right)^2 \approx 7{,}425.$$

A példából is látható, hogy a várható érték additivitása könnyíthet a kovariancia kiszámolásán.

Megjegyzés. Adódik a kérdés, hogy ha a kovariancia nulla volta nem is karakterizálja a függetlenséget, akkor miért ezt a definíciót nézzük? Ennek a fő oka, hogy a kovariancia szimmetrikus és bilineáris, azaz

$$\mathrm{cov}(X, Y) = \mathrm{cov}(Y, X) \qquad \text{és} \qquad \mathrm{cov}(X, aY + bZ) = a \cdot \mathrm{cov}(X, Y) + b \cdot \mathrm{cov}(X, Z) \qquad (a, b \in \mathbb{R}),$$

ha a fenti kovarianciák léteznek. Így a kovariancia a vektorok skaláris szorzatának rokonfogalma.

6.4. Variancia és szórás

Speciális eset a kovariancia számolásában, amikor $Y = X$.

⚠ **6.4.1. Definíció.** Egy X valószínűségi változó **szórásnégyzet**e, vagy más néven **varianciá**ja:

$$\mathrm{cov}(X, X) = \mathbb{E}\big((X - \mathbb{E}X)^2\big) = \mathbb{E}(X^2) - \mathbb{E}(X)^2.$$

Jelölés: $\mathbb{D}^2(X)$ (alternatív jelölése: $\mathrm{Var}(X)$). Egy valószínűségi változónak nem mindig létezik szórásnégyzete (hiszen lehet olyan eset, hogy már $\mathbb{E}(X)$ is értelmetlen), de ha létezik, akkor nemnegatív. A szórásnégyzet négyzetgyökét **szórás**nak hívjuk, jelölése: $\mathbb{D}(X)$.

Megjegyzés. Más szavakkal, X szórásnégyzete az X-nek az átlagos értékétől való négyzetes eltérése. A vektoros analógiát felhasználva, ha a kovarianciát a vektorok skaláris szorzatával állítjuk párhuzamba, akkor a szórásnégyzet a vektor hossznégyzetének, míg a szórás a vektor hosszának feleltethető meg.

Nyilván ez a mennyiség nem X-nek a saját magával való összefüggőségéről szolgáltat információt, hanem arról, hogy X értékei mennyire terülnek szét az átlaga körül. Ilyen „szétterülést” mérő számot többféleképp definiálhatnánk, például nézhetnénk az $\mathbb{E}(|X - \mathbb{E}X|)$-et is. Hogy mégis a szórásnégyzet a népszerű mérőszám erre, annak az egyik oka az alábbi állítás.

⚠ **6.4.2. Állítás.** *Ha X és Y független, akkor* $\mathbb{D}^2(X + Y) = \mathbb{D}^2(X) + \mathbb{D}^2(Y)$.

Bizonyítás. A szórásnégyzet definícióját kibontva:

$$\mathbb{D}^2(X + Y) = \mathbb{E}\big((X + Y)^2\big) - \big(\mathbb{E}(X + Y)\big)^2 =$$
$$= \mathbb{E}(X^2) + \mathbb{E}(Y^2) + 2\mathbb{E}(XY) - \mathbb{E}(X)^2 - \mathbb{E}(Y)^2 - 2\mathbb{E}(X)\mathbb{E}(Y)$$
$$= \mathbb{D}^2(X) + \mathbb{D}^2(Y) + 2\mathrm{cov}(X, Y).$$

Mivel X és Y függetlenek, így $\mathrm{cov}(X, Y) = 0$, amiből az állítás már következik. □

Megjegyzés. A bizonyításból látható, hogy a fenti állítást függetlenség nélkül is kimondhattuk volna, csak úgy valamivel bonyolultabb eredményt kapunk:

$$\mathbb{D}^2(X+Y) = \mathbb{D}^2(X) + \mathbb{D}^2(Y) + 2\mathrm{cov}(X,Y).$$

További elemi tulajdonságai a szórásnégyzetnek:

6.4.3. Állítás. *Tegyük fel, hogy $\mathbb{D}(X)$ létezik és véges. Ekkor tetszőleges $c \in \mathbb{R}$ esetén*

$$\mathbb{D}(X+c) = \mathbb{D}(X) \qquad \textit{és} \qquad \mathbb{D}(cX) = |c|\mathbb{D}(X),$$

azaz a szórás eltolás-invariáns és abszolút homogén.

Bizonyítás. A szórásnégyzet definícióját kibontva

$$\mathbb{D}^2(X+c) = \mathbb{E}\Big(\big(X+c-\mathbb{E}(X+c)\big)^2\Big) = \mathbb{E}\big((X-\mathbb{E}X)^2\big) = \mathbb{D}^2(X), \qquad \text{illetve}$$

$$\mathbb{D}^2(cX) = \mathbb{E}\Big(\big(cX-\mathbb{E}(cX)\big)^2\Big) = \mathbb{E}\big(c^2(X-\mathbb{E}X)^2\big) = c^2\mathbb{D}^2(X),$$

amely egyenletekből gyökvonással adódik az állítás. □

6.4.4. Példa.

(1) Legyen K egyenletes eloszlású valószínűségi változó az $\{1,2,3,4,5,6\}$ halmazon. (Vajon miért jelöljük K-val?) Ekkor a 3. előadás példája szerint $\mathbb{E}(K^2) = \frac{91}{6}$, míg $\mathbb{E}(K) = \frac{7}{2}$, ezért

$$\mathbb{D}^2(K) = \frac{91}{6} - \left(\frac{7}{2}\right)^2 = \frac{182-147}{12} = \frac{35}{12} \approx 2{,}9167, \quad \text{és} \quad \mathbb{D}(K) = \sqrt{\mathbb{D}^2(K)} \approx 1{,}7078.$$

(2) Vizsgáljuk az $\mathbf{1}_A$ indikátor valószínűségi változót, és jelölje az A esemény valószínűségét p. Ekkor

$$\mathbb{D}^2(\mathbf{1}_A) = \mathbb{E}(\mathbf{1}_A^2) - \mathbb{E}(\mathbf{1}_A)^2 = \mathbb{E}(\mathbf{1}_A) - \mathbb{E}(\mathbf{1}_A)^2 = \mathbb{P}(A) - \mathbb{P}(A)^2 = p(1-p).$$

(3) Legyen $X \sim B(n;p)$. Bár $\mathbb{D}^2(X)$ kibontható a definíció alapján is, célszerűbb felhasználni, hogy felírható $X = \mathbf{1}_{A_1} + \cdots + \mathbf{1}_{A_n}$ alakban, ahol $A_1, \ldots, A_n$ együttesen független, p valószínűségű események. Így a fenti állítás miatt

$$\mathbb{D}^2(X) = \mathbb{D}^2(\mathbf{1}_{A_1} + \cdots + \mathbf{1}_{A_n}) = \mathbb{D}^2(\mathbf{1}_{A_1}) + \cdots + \mathbb{D}^2(\mathbf{1}_{A_n}) = np(1-p).$$

(4) Legyen $T \sim \mathrm{Geo}(p)$. Ekkor

$$\mathbb{D}^2(T) = \mathbb{E}(T^2) - \mathbb{E}(T)^2 = \sum_{k=1}^{\infty} k^2(1-p)^{k-1}p - \left(\frac{1}{p}\right)^2 = \frac{1-p}{p^2},$$

ahol az első szumma kiszámolható például hatványsorok deriválásával, ettől itt eltekintünk.

(5) Legyen $Y \sim \mathrm{Pois}(\lambda)$. Ekkor

$$\mathbb{D}^2(Y) = \mathbb{E}(Y^2) - \mathbb{E}(Y)^2 = \mathbb{E}(Y^2 - Y) + \mathbb{E}(Y) - \mathbb{E}(Y)^2 = \sum_{k=0}^{\infty}(k^2-k)\frac{\lambda^k}{k!}e^{-\lambda} + \lambda - \lambda^2 =$$

$$= \lambda^2 \sum_{k=2}^{\infty} \frac{\lambda^{k-2}}{(k-2)!}e^{-\lambda} + \lambda - \lambda^2 = \lambda^2 \cdot 1 + \lambda - \lambda^2 = \lambda.$$

(6) A szórás definíciója akkor is értelmes, ha a valószínűségi változó például folytonos. Legyen $Z \sim \mathrm{Exp}(\lambda)$ valamilyen λ pozitív valósra. Kiszámolható (és a 10. előadáson ki is számoljuk), hogy ekkor $\mathbb{D}(Z) = \frac{1}{\lambda}$.

6.5. Korreláció

Fentebb utaltunk rá, hogy a kovariancia segíthet mérni valószínűségi változók összefüggését. De hogyan kell ezt érteni, ha a kovariancia épp nem 0? Például az első példából adódó $\operatorname{cov}(X,Y) = -0{,}6$ értéknek mi a jelentése, mennyire függnek ettől össze az X és Y változók?

Ezt pusztán a kovariancia alapján nem tudjuk megválaszolni, ahhoz egy ebből származtatott mennyiség lesz a segítségünkre.

⚠ **6.5.1. Definíció.** Legyenek X és Y valószínűségi változók. Ha $\operatorname{cov}(X,Y)$, $\mathbb{D}(X)$ és $\mathbb{D}(Y)$ értelmes, akkor X és Y **korreláció**ja:

$$\operatorname{corr}(X,Y) \stackrel{\text{def}}{=} \frac{\operatorname{cov}(X,Y)}{\mathbb{D}(X)\mathbb{D}(Y)}.$$

Belátható, hogy $-1 \le \operatorname{corr}(X,Y) \le 1$ mindig teljesül. A szélsőséges esetekben X és Y közt tökéletes lineáris összefüggés áll fent, azaz teljesül a következő állítás.

6.5.2. Állítás. *Legyenek X és Y valószínűségi változók. Ha $\operatorname{corr}(X,Y) \in \{1,-1\}$, akkor az $Y = aX + b$ állítás 1 valószínűséggel teljesül valamilyen a és b valós számokra, ahol az a előjele megegyezik $\operatorname{corr}(X,Y)$ előjelével.*

6.5.3. Példa. A diszkrét együttes eloszlás részben vizsgált példa esetében

$$\operatorname{corr}(X,Y) = \frac{\operatorname{cov}(X,Y)}{\mathbb{D}(X)\mathbb{D}(Y)} = \frac{-0{,}6}{\sqrt{9{,}2}\cdot\sqrt{0{,}6}} \approx -0{,}2554.$$

Ennek a szemléletes jelentése az, hogy X és Y szívesebben tér el ellentétes irányba az átlagától, mint azonos irányba, de a köztük lévő lineáris összefüggés relatíve alacsony (legalábbis amennyire a 0,25 az 1-hez képest alacsony).

Ahogy kovariancia esetében is, a korreláció nulla mivolta nem jelenti, hogy a két valószínűségi változó független volna. Valójában a korreláció a két változó közti lineáris összefüggés fokát méri. Más szavakkal, hiába függ össze két valószínűségi változó, ha az összefüggésük nemlineáris, azt a korreláció nem fogja észrevenni. Például megadható olyan X valószínűségi változó, amire $\operatorname{corr}(X,X^2) = 0$.