A Valószínűségszámítás I. előadássorozat ötödik előadása.
2003. március 4.

Összefoglaló:

A várható érték tulajdonságainak bizonyításánál szükségünk van az alábbi segédtételre, amely megegyezik a Fazekas könyvben szereplő 2.6 Állítással a 61. oldalon.

Segédtétel. *Legyen ξ egy diszkrét eloszlású valószínűségi változó egy $(\Omega,\mathcal{A},P)$ valószínűségi mezőn, amely valamilyen $x_1,x_2,\dots,$ értékeket vesz fel, és legyen $B_k=\{\omega:\xi(\omega)=x_k\}$, $k=1,2,\dots,$ a ξ valószínűségi változó értékei által meghatározottt teljes eseményrendszer. Legyen C_j, $j=1,2,\dots,$ egy ennél finomabb teljes eseményrendszer, azaz tegyük fel, hogy a $C_j\in\mathcal{A}$ halmazok diszjunktak, $\bigcup\limits_{j=1}^\infty C_j=\Omega$, és minden C_j halmazhoz tartozik olyan B_k halmaz, amelyre $C_j\subset B_k$. Ha $C_j\subset B_k$, akkor rendeljük hozzá a C_j halmazhoz az $y_j=x_k$ számot. A $\sum\limits_{k=1}^\infty x_kP(B_k)$ és $\sum\limits_{j=1}^\infty y_jP(C_j)$ összegek egyszerre abszolut konvergensek, és ha ezek az összegek abszolut konvergensek, akkor*

$$E\xi=\sum_{k=1}^\infty x_kP(B_k)=\sum_{j=1}^\infty y_jP(C_j).$$

Bizonyítás: Ha a $\sum\limits_{k=1}^\infty x_kP(B_k)$ összeg abszolut konvergens, azaz létezik olyan $L<\infty$ szám, amelyre $\sum\limits_{k=1}^\infty |x_k|P(B_k)=L$ akkor minden N indexre

$$\sum_{j=1}^N |y_j|P(C_j)\le\sum_{k=1}^\infty |x_k|P(B_k)=L.$$

Ezért $\sum\limits_{j=1}^\infty |y_j|P(C_j)\le L$, azaz a $\sum\limits_{j=1}^\infty y_jP(C_j)$ összeg is abszolut konvergens.

Ha a $\sum\limits_{j=1}^\infty y_jP(C_j)$ összeg abszolut konvergens, akkor $\sum\limits_{j=1}^\infty |y_j|P(C_j)\le L<\infty$ alkalmas L számmal, ahonnan tetszőleges N indexre

$$\sum_{k=1}^N |x_k|P(B_k)=\sum_{k=1}^N\sum_{j\colon\, C_j\subset B_k}^\infty |y_j|P(C_j)\le\sum_{j=1}^\infty |y_j|P(C_j)\le L<\infty,$$

ahonnan a $\sum\limits_{k=1}^\infty x_kP(B_k)$ összeg is abszolut konvergens. Ezzel beláttuk, hogy a két tekintett összeg egyszerre abszolut konvergens.

Ha a fenti sorok abszolut konvergensek akkor minden $\varepsilon > 0$ számhoz létezik olyan $N = N(\varepsilon)$ index, hogy $\sum\limits_{k=N}^{\infty} |x_k| P(B_k) < \varepsilon$. Ekkor

$$\sum_{k=N}^{\infty} \sum_{j\colon\, C_j \subset B_k} |y_j| P(C_j) < \varepsilon.$$

Legyen $U_k = \{j\colon C_j \subset B_k\}$ minden $k = 1, 2, \dots$ számra. Ekkor minden U_k halmaznak van olyan véges $V_k = V_k(\varepsilon) \subset U_k$ részhalmaza, amelyre

$$|x_k| \sum_{j \in U_k \setminus V_k} P(C_j) \le \varepsilon 2^{-k}$$

Ezért

$$\left| \sum_{j\colon\, j \in V_k}^{\infty} y_j P(C_j) - \sum_{j\colon\, C_j \subset B_k}^{\infty} y_j P(C_j) \right| \le \sum_{j \in U_k \setminus V_k} |y_j| P(C_j), \quad k = 1, 2, \dots,$$

és

$$\left| \sum_{j\colon\, j \in V_k}^{\infty} y_j P(C_j) - x_k P(B_k) \right| \le \sum_{j \in U_k \setminus V_k} |y_j| P(C_j), \quad k = 1, 2, \dots,$$

ahonnan

$$\left| \sum_{j=1}^{\infty} y_j P(C_j) - \sum_{k=1}^{N-1} \sum_{j \in V_k} y_j P(C_j) \right| \le \sum_{k=N}^{\infty} \sum_{j \in V_k} |y_j| P(C_j) + \sum_{k=1}^{N-1} \sum_{j \in U_k \setminus V_k} |y_j| P(C_j) \le 2\varepsilon,$$

és

$$\left| \sum_{k=1}^{\infty} x_k P(B_k) - \sum_{k=1}^{N-1} \sum_{j \in V_k} y_j P(C_j) \right| \le \sum_{k=N}^{\infty} |x_k| P(B_k) + \sum_{k=1}^{N-1} \sum_{j \in U_k \setminus V_k} |y_j| P(C_j) \le 2\varepsilon.$$

Innen

$$\left| \sum_{j=1}^{\infty} y_j P(C_j) - \sum_{k=1}^{\infty} x_k P(B_k) \right| \le 4\varepsilon.$$

Mivel ez az állítás minden $\varepsilon > 0$ számra érvényes, innen következik a Segédtétel állítása.

Az előző előadásban megfogalmazott $E(\xi_1 + \xi_2) = E\xi_1 + E\xi_2$ reláció bizonyítása.
Ha a ξ_1 és ξ_2 valószínűségi változók $x_1, x_2, \dots$ értékeket vesznek fel, akkor tekintsük az $\{\omega\colon \xi_1 = x_j, \xi_2 = x_k\}$, $j, k = 1, 2, \dots,$ eseményeket. A Segédtétel alapján

$$E\xi_1 = \sum_{j=1}^{\infty} \sum_{k=1}^{\infty} x_j P(\xi_1 = x_j, \xi_2 = x_k),$$

és

$$E\xi_2 = \sum_{j=1}^{\infty}\sum_{k=1}^{\infty} x_k P(\xi_1 = x_j, \xi_2 = x_k).$$

E két azonosságot összeadva, és alkalmazva megint a Segédtételt kapjuk, hogy

$$E\xi_1 + E\xi_2 = \sum_{j=1}^{\infty}\sum_{k=1}^{\infty} (x_j + x_k) P(\xi_1 = x_j, \xi_2 = x_k) = E(\xi_1 + \xi_2),$$

mert az $\{\omega\colon \xi_1(\omega) = x_j, \xi_2(\omega) = x_k\}$, $j, k = 1, 2, \dots$, halmazok egy finomítását adják az $\{\omega\colon \xi_1(\omega) + \xi_2(\omega) = y\}$, ($y = x_j + x_k$ alkalmas x_j és x_k számmal) alakú halmazokból álló particiónak.

Tétel. *Legyen ξ diszkrét eloszlású valószínűségi változó, amely bizonyos $x_1, x_2, \dots$ értéket vesz fel és $g(x)$ valós függvény. Ekkor*

$$Eg(\xi) = \sum_{j=1}^{\infty} g(x_j) P(\xi = x_j),$$

feltéve, hogy a $\sum\limits_{j=1}^{\infty} g(x_j) P(\xi = x_j)$ összeg abszolut konvergens.

Bizonyítás: Vezessük be az $x_u \sim x_v$ relációt, ha $g(x_u) = g(x_v)$. Ekkor $\sim$ ekvivalenciareláció, és az $x_1, x_2, \dots,$ értékeket megszámlálható sok $D_1, D_2, \dots,$ osztályba tudjuk sorolni úgy, hogy x_u és x_v akkor kerül ugyanabba az osztályba, ha $x_u \sim x_v$. Definiáljuk a z_k számot úgy, hogy legyen $z_k = g(x_u)$, ha $x_u \in D_k$. Vezessük be továbbá a $B_k = \{\omega\colon \xi(\omega) \in D_k\}$ és $C_j = \{\omega\colon \xi(\omega) = x_j\}$ halmazokat az $(\Omega, \mathcal{A}, P)$ valószínűségi mezőn. Ekkor a várható érték definiciója, illetve a Segédtétel alapján

$$Eg(\xi) = \sum_k z_k P(C_k) = \sum_{j=1}^{\infty} g(x_j) P(B_j) = \sum_{j=1}^{\infty} g(x_j) P(\xi = x_j),$$

feltéve, hogy a $\sum\limits_{j=1}^{\infty} g(x_j) P(\xi = x_j)$ összeg abszolut konvergens.

Tétel. *Legyenek $\xi_1, \xi_2, \dots, \xi_n$ független, diszkrét eloszlású valószínűségi változók, amelyek mindegyikére létezik az $E\xi_j$, $1 \le j \le n$ várható érték. Ekkor az $E\xi_1\xi_2 \cdots \xi_n$ várható érték is létezik, és*

$$E\xi_1\xi_2 \cdots \xi_n = E\xi_1 E\xi_2 \cdots E\xi_n.$$

Megjegyés: A fent megfogalmazott eredmény rendkívül fontos tulajdonsága a *független* valószínűségi változóknak. A Fazekas könyvben ez az eredmény csak az $n = 2$ esetben van megfogalmazva a 2.21 Tételben a 66. oldalon. Ha jól tudunk számolni független valószínűségi változókkal, akkor az általános eredmény visszavezethető erre a speciális esetre, de egyszerűbbnek láttam közvetlenül az általános esettel foglalkozni.

Bizonyítás: Tegyük fel, hogy a $\xi_1, \dots, \xi_n$ valószínűségi változók $x_1, x_2, \dots,$ értékeket vesznek fel. Ekkor

$$\begin{aligned} E\xi_1 \cdots E\xi_n &= \sum_{j_1=1}^{\infty} x_{j_1} P(\xi_1 = x_{j_1}) \cdots \sum_{j_n=1}^{\infty} x_{j_n} P(\xi_n = x_{j_n}) \\ &= \sum_{j_1=1}^{\infty} \cdots \sum_{j_n=1}^{\infty} x_{j_1} \cdots x_{j_n} P(\xi_1 = x_{j_1}) \cdots P(\xi_n = x_{j_n}) \\ &= \sum_{j_1=1}^{\infty} \cdots \sum_{j_n=1}^{\infty} x_{j_1} \cdots x_{j_n} P(\xi_1 = x_{j_1}, \cdots, \xi_n = x_{j_n}), \end{aligned}$$

a ξ_j, $1 \le j \le n$ valószínűségi változók függetlensége miatt, valamint azért mert a fenti sorok mindegyike abszolut konvergens. (Ugyanis abszolut konvergens sorok szorzata is abszolut konvergens.) Viszont a Segédtétel alapján a fenti azonosság jobboldalán szereplő kifejezés $E\xi_1\xi_2 \cdots \xi_n$, ezért

$$E\xi_1\xi_2 \cdots \xi_n = E\xi_1 E\xi_2 \cdots E\xi_n.$$

A szórásnégyzet fogalma.

Azt hogy valószínűségi változók körülbelül mekkora értéket vesznek fel az (egyelőre csak diszkrét eloszlású valószínűségi változók esetén tárgyalt) várható érték méri megfelelő módon. Annak mérésére, hogy mekkora a valószínűségi változó ingadozása a várható értéke körül vezették be a szórásnégyzet (angolul variance) fogalmát. Ennek definiciója a következő:

A szórásnégyzet definiciója: *Legyen ξ olyan valószínűségi változó, amelyre $E\xi^2 < \infty$. Ekkor a ξ valószínűségi változó szórásnégyzetét a*

$$\operatorname{Var} \xi = E \left(\xi - E\xi\right)^2$$

képlet definiálja. Ha $E\xi^2 = \infty$ akkor a $\operatorname{Var} \xi$ szórásnégyzetet nem definiáljuk vagy azt mondjuk, hogy $\operatorname{Var} \xi = \infty$.

Lemma.

$$\operatorname{Var} \xi = E\xi^2 - (E\xi)^2.$$

Továbbá, $(E\xi)^2 \le E\xi^2$.

Bizonyítás:

$$\begin{aligned} 0 \le \operatorname{Var} \xi = E\left(\xi - E\xi\right)^2 &= E\left(\xi^2 - 2\xi E\xi + (E\xi)^2\right) \\ &= E\xi^2 - 2E\xi E\xi + (E\xi)^2 = E\xi^2 - (E\xi)^2 \end{aligned}$$

a várható érték tulajdonságai miatt. Összehasonlítva a bal és jobboldalt kapjuk a megfogalmazott egyenlőtlenséget.

Megjegyzés: Ha precízek vagyunk felmerülhet a kérdés, teljesen kifogástalan-e a fenti bizonyítás. Az okozhat esetleg problémát, hogy csak akkor definiáltuk a várható értéket, ha a definiáló összeg abszolut konvergens. Ezt viszont csak az $E\xi^2$ kifejezést definiáló összegről tettük fel, ugyanakkor az $E\xi$-vel is szabadon számoltunk. Nem okoz-e ez gondot?

Az, hogy a fenti számolás jogos, például a következő módon látható. Vegyen fel a ξ valószínűségi változó $x_1, x_2, \dots$ értékeket, és definiáljuk a ξ valószínűségi változó ξ_n közelítéseit, $n = 1, 2, \dots$, a következő módon: $\xi_n(\omega) = \xi(\omega)$, ha $\xi(\omega) = x_k$, $1 \le k \le n$, $\xi_n(\omega) = 0$, ha $\xi(\omega) = x_k$, és $k > n$. Ekkor a $\xi_n(\omega)$ illetve ennek abszolut értékével a $|\xi_n(\omega)|$ valószínűségi változókkal szabadon végrehajthatjuk a fenti számolásokat. Innen következik, hogy $(E|\xi_n(\omega)|)^2 \le E\xi_n^2(\omega) \le E\xi^2(\omega)$. Innen viszont $n \to \infty$ határátmenettel kapjuk, hogy $(E|\xi|)^2 = \lim\limits_{n\to\infty} (E|\xi_n(\omega)|)^2 \le E\xi^2(\omega)$. Ez azt jelenti, hogy az $E\xi$ várható értéket meghatározó összeg is abszolut konvergens, és a fenti számolásokat elvégezhetjük.

Feladat:

Minden m valós számra

$$E(\xi - m)^2 = E(\xi - E\xi)^2 + ((E\xi) - m)^2 .$$

Következésképpen

$$\operatorname{Var}\xi = \inf_{-\infty<m<\infty} E\left[(\xi - m)^2\right].$$

Megoldás:

$$\begin{aligned} E(\xi - m)^2 &= E\left[(\xi - E\xi) + (E\xi - m)\right]^2 \\ &= E(\xi - E\xi)^2 - 2E(\xi - m)E(\xi - E\xi) + E(\xi - m)^2 \\ &= E(\xi - E\xi)^2 + ((E\xi) - m)^2 , \end{aligned}$$

mert $2E(\xi - m)E(\xi - E\xi) = 2E(\xi - m)(E\xi - E\xi) = 0$. Ebből az azonosságból adódik, hogy $E(\xi - m)^2 \ge E(\xi - E\xi)^2$ minden m valós számra. Viszont $m = E\xi$ esetén az egyenlőtlenség két oldala megegyezik.

Lemma. Minden a és b valós számra

$$\operatorname{Var}(a\xi + b) = a^2 \operatorname{Var}\xi.$$

Bizonyítás:

$$\begin{aligned} \operatorname{Var}(a\xi + b) &= E\left[(a\xi + b) - E(a\xi + b)\right]^2 = E\left[(a\xi + b) - aE(\xi - b\right]^2 \\ &= E\left[a^2(\xi - E\xi)^2\right] = a^2 E(\xi - E\xi)^2 = a^2 \operatorname{Var}\xi. \end{aligned}$$

A következő eredmény rendkívül hasznos összefüggést ad, ha független valószínűségi változók összegének szórásnégyzetét akarjuk kiszámítani. Hangsúlyozzuk, hogy ebben az eredményben fontos a tekintett valószínűségi változók függetlensége. Később megfogalmazzuk ezen eredmény egy olyan általánosítását, amely akkor is érvényes, ha az összeadandók nem feltétlenül függetlenek. Látni fogunk olyan példákat is, amelyek megoldásában ezt az általánosabb eredményt érdemes alkalmazni.

Tétel. Ha $\xi_1, \xi_2, \dots, \xi_n$ *független valószínűségi változók, akkor*

$$\mathrm{Var}\,(\xi_1 + \xi_2 + \dots + \xi_n) = \mathrm{Var}\,\xi_1 + \mathrm{Var}\,\xi_2 + \dots + \mathrm{Var}\,\xi_n.$$

Következmény. Ha $\xi_1, \xi_2, \dots, \xi_n$ *független valószínűségi változók,* $c_1, c_2, \dots, c_n$ *valós számok, akkor*

$$\mathrm{Var}\,(c_1\xi_1 + c_2\xi_2 + \dots + c_n\xi_n) = c_1^2\mathrm{Var}\,\xi_1 + c_2^2\mathrm{Var}\,\xi_2 + \dots + c_n^2\mathrm{Var}\,\xi_n.$$

A Tétel bizonyítása. Vezessük be a $\bar{\xi}_j = \xi_j - E\xi_j$ valószínűségi változókat. Ekkor

$$\begin{aligned}
\mathrm{Var}\,(\xi_1 + \xi_2 + \dots + \xi_n) &= E(\bar{\xi}_1 + \bar{\xi}_2 + \dots + \bar{\xi}_n)^2 \\
&= E\left(\bar{\xi}_1^2 + \bar{\xi}_2^2 + \dots + \bar{\xi}_n^2 + 2\sum_{j=1}^{n-1}\sum_{k=j+1}^{n} \bar{\xi}_j\bar{\xi}_k\right) \\
&= E\bar{\xi}_1^2 + E\bar{\xi}_2^2 + \dots + E\bar{\xi}_n^2 + 2\sum_{j=1}^{n-1}\sum_{k=j+1}^{n} E\bar{\xi}_j\bar{\xi}_k \\
&= E\bar{\xi}_1^2 + E\bar{\xi}_2^2 + \dots + E\bar{\xi}_n^2 = \mathrm{Var}\,\xi_1 + \mathrm{Var}\,\xi_2 + \dots + \mathrm{Var}\,\xi_n,
\end{aligned}$$

mert $E\bar{\xi}_j\bar{\xi}_k = E\bar{\xi}_j E\bar{\xi}_k = 0$ minden $1 \le j < k \le n$ számra a ξ_j illetve a $\bar{\xi}_j$, $1 \le j \le n$, valószínűségi változók függetlensége miatt.

Feladat:

Egy szabályos pénzdarabot feldobunk 100-szor egymás után. Számoljuk ki a fejdobások számának a várható értékét és szórásnégyzetét. Számítsuk ki ezt egyrészt direkt módon felírva és kiszámolva a várható értéket és szórásnégyzetet kifejező összegeket, másrészt egyszerűbben az előző tétel segítségével.

Megoldás: A fejdobások száma 0 és 100 között van. Annak valószínűsége, hogy k fejdobás következik be, $0 \le k \le 100$, $p_k = \binom{100}{k} 2^{-100}$. Ezért a dobások számának várható értéke $E\xi = \sum\limits_{k=0}^{100} k\binom{100}{k}2^{-100}$, ahol ξ jelöli a fejdobások számát megadó valószínűségi változót. Továbbá $E\xi^2 = \sum\limits_{k=0}^{100} k^2\binom{100}{k}2^{-100}$, a szórásnégyzet pedig $\mathrm{Var}\,\xi = E\xi^2 - (E\xi)^2$. Ezeket az összegeket közvetlenül kiszámíthatjuk. Valóban, $k\binom{100}{k} = 100\binom{99}{k-1}$, $E\xi = \sum\limits_{k=0}^{100} k\binom{100}{k}2^{-100} = 50\sum\limits_{k=0}^{99} k\binom{99}{k}2^{-99} = 50\left(\frac{1}{2}+\frac{1}{2}\right)^{99} = 50$. Továbbá, $k^2\binom{100}{k} = [k(k-1)+k]\binom{100}{k} = 100\cdot 99\cdot\binom{98}{k-2} + 100\cdot\binom{100}{99}$, $E\xi^2 = 2^{-100}\cdot 100\cdot 99\cdot\sum\limits_{k=0}^{98}\binom{98}{k} + 2^{-100}\cdot 100\cdot\sum\limits_{k=0}^{99}\binom{99}{k} = 2^{-100}\left(100\cdot 99\cdot(1+1)^{98} + 100\cdot(1+1)^{99}\right) = 25\cdot 99+50$, $\mathrm{Var}\,\xi = E\xi^2 - (E\xi^2) = 25$. Valójában a vizsgált várható értéket és szórásnégyzetet egyszerűbben is kiszámíthatjuk. Vezessük be a ξ_j valósz1nűségi változókat, $\xi_j = 1$, ha a j-ik dobás eredménye fej $\xi_j = 0$, ha a j-ik dobás eredménye írás, $1 \le j \le 100$. Ekkor a fejdobások száma $\xi = \sum\limits_{j=1}^{100}\xi_j$, $E\xi = \sum\limits_{j=1}^{100} E\xi_j$, $\mathrm{Var}\,\xi = \sum\limits_{j=1}^{100}\mathrm{Var}\,\xi_j$. Mivel $E\xi_j = \frac{1}{2}$, $E\xi_j^2 = \frac{1}{2}$, $\mathrm{Var}\,\xi_j = \frac{1}{4}$, $1 \le j \le 100$, ezért $E\xi = \frac{1}{2}\cdot 100 = 50$, $\mathrm{Var}\,\xi = \frac{1}{4}\cdot 100 = 25$.

Egy szabályos dobókockát feldobunk 100-szor egymás után. Számoljuk ki a dobáseredmények összegének várható értékét és szórásnégyzetét.

Megoldás: Legyenek $\eta_1,\dots,\eta_{100}$ *független* valószínűségi változók, amelyekre $P(\eta_j = k) = \frac{1}{6}$, $1 \le j \le 100$, $1 \le k \le 6$. Ekkor a kiszámítandó várható érték és szórásnégyzet $E\sum\limits_{j=1}^{100}\eta_j = \sum\limits_{j=1}^{100}E\eta_j$, $\mathrm{Var}\sum\limits_{j=1}^{100}\eta_j = \sum\limits_{j=1}^{100}\mathrm{Var}\,\eta_j$. (A második reláció felhasználja a tekintett valószínűségi változók függetlenségét.) $E\eta_j = \frac{1}{6}(1+2+3+4+5+6) = 3.5$, $E\eta_j^2 = \frac{1}{6}(1+4+9+16+25+36) = \frac{91}{6}$, $\mathrm{Var}\,\eta_j = \frac{35}{12}$, $E\xi = 350$, $\mathrm{Var}\,\xi = \frac{3500}{12}$.

Megfogalmazzuk az előző tétel egy olyan általánosítását, amely lehetővé teszi, hogy kiszámoljuk nem feltétlenül független valószínűségi változók összegének a szórásnégyzetét. Ennek érdekében először be kell vezetnünk egy új fogalmat, két valószínűségi változó kovarianciafüggvényét. Ez azt méri hogy milyen erős a kapcsolat két valószínűségi változó között.

A kovarianciafüggvény definiciója. *Legyen ξ és η két valószínűségi változó, amelyek mindegyikének létezik szórásnégyzete, azaz $E\xi^2 < \infty$ és $E\eta^2 < \infty$. Ekkor a ξ és η kovarianciafüggvényét* $\mathrm{Cov}\,(\xi,\eta)$*-t a*

$$\mathrm{Cov}\,(\xi,\eta) = E\left[(\xi - E\xi)(\eta - E\eta)\right]$$

kifejezés definiálja. Ha a $E\xi^2 < \infty$ és $E\eta^2 < \infty$ feltételek valamelyike nem teljesül, akkor nem definiáljuk a $\mathrm{Cov}\,(\xi,\eta)$ *kovariancifüggvényt.*

Megjegyzés: $|(\xi - E\xi)||(\eta - E\eta)| \le \dfrac{|\xi - E\xi|^2}{2} + \dfrac{|\eta - E\eta|^2}{2}$, ahonnan $E|(\xi - E\xi)||(\eta -$

$E\eta)| \leq \dfrac{E|\xi - E\xi|^2}{2} + \dfrac{E|\eta - E\eta|^2}{2} < \infty$. Innen következik, hogy az $E\xi^2 < \infty$ és $E\eta^2 < \infty$ feltételek teljesülése esetén a $\operatorname{Cov}(\xi, \eta)$-t definiáló várható érték valóban létezik.

Lemma.

$$\operatorname{Cov}(\xi, \eta) = E\xi\eta - E\xi E\eta.$$

Következmény. *Ha ξ és η két független valószínűségi változó, akkor* $\operatorname{Cov}(\xi, \eta) = 0$.

Lemma bizonyítása.

$$\operatorname{Cov}(\xi, \eta) = E[\xi\eta - \eta E\xi - \xi E\eta + E\xi E\eta] = E\xi\eta - E\xi E\eta - E\xi E\eta + E\xi E\eta = E\xi\eta - E\xi E\eta.$$

Most megfogalmazzuk az alábbi eredményt valószínűségi változók összegének a szórásnégyzetéről:

Tétel. *Ha $\xi_1, \ldots, \xi_n$ valószínűségi változók ugyanazon a valószínűségi mezőn, amelyek mindegyikének létezik szórásnégyzete, akkor*

$$\begin{aligned}
\operatorname{Var}\left(\sum_{j=1}^{n} \xi_j\right) &= \sum_{j=1}^{n} \operatorname{Var} \xi_j + 2 \sum_{1 \leq j < k \leq n} (E\xi_j\xi_k - E\xi_j E\xi_k) \\
&= \sum_{j=1}^{n} \operatorname{Var} \xi_j + 2 \sum_{1 \leq j < k \leq n} (\operatorname{Cov}(\xi_j, \xi_k) \\
&= \sum_{j=1}^{n} \operatorname{Var} \xi_j + \sum_{1 \leq j, k \leq n,\ j \neq k} (\operatorname{Cov}(\xi_j, \xi_k).
\end{aligned}$$

Bizonyítás.

$$\begin{aligned}
\operatorname{Var}\left(\sum_{j=1}^{n} \xi_j\right) &= E\left(\sum_{j=1}^{n} \xi_j\right)^2 - \left(E\sum_{j=1}^{n} \xi_j\right)^2 = E\left(\sum_{j=1}^{m}\sum_{k=1}^{n} \xi_j\xi_k\right) - \left(\sum_{j=1}^{m}\sum_{k=1}^{n} E\xi_j E\xi_k\right) \\
&= \left(\sum_{j=1}^{m}\sum_{k=1}^{n} E\xi_j\xi_k\right) - \left(\sum_{j=1}^{m}\sum_{k=1}^{n} E\xi_j E\xi_k\right) \\
&= \sum_{j=1}^{n} \left(E\xi_j^2 - (E\xi_j)^2\right) + 2 \sum_{1 \leq j < k \leq n} (E\xi_j\xi_k - E\xi_j E\xi_k).
\end{aligned}$$

és $E\xi_j^2 - (E\xi_j)^2 = \operatorname{Var} \xi_j$, $E\xi_j\xi_k - E\xi_j E\xi_k = \operatorname{Cov}(\xi_j, \xi_k)$.

Megmutatjuk a (részben a gyakorlaton), hogy a fenti eredmény segítségével viszonylag egyszerűen meg tudjuk oldani a következő két feladatot.

Feladat:

Legyen egy urnában 20 piros és 30 fehér golyó. Húzzunk ki 20 golyót visszatevés nélkül. Számoljuk ki a kihúzott piros golyók számának várható értékét és szórásnégyzetét.

Megoldás: Vezessük be a következő ξ_j, $1 \le j \le 20$, valószínűségi változókat: $\xi_j(\omega) = 1$, ha a j-ik húzás eredménye piros, $\xi_j(\omega) = 0$, ha a j-ik húzás eredménye fehér. Ekkor a $\xi = \sum\limits_{j=1}^{20} \xi_j$ összeg várható értékét és szórásnégyzetét kell kiszámolnunk. Továbbá $E\xi_j = E\xi_1 = \frac{2}{5}$, $\operatorname{Var} \xi_j = \operatorname{Var} \xi_1 = \frac{2}{5} - \frac{4}{25} = \frac{6}{25}$, $E\xi_j\xi_k - E\xi_j E\xi_k = E\xi_1\xi_2 - E\xi_1 E\xi_2 = \frac{2}{5} \cdot \frac{19}{49} - \frac{4}{25} = -\frac{6}{1245}$, ha $j \ne k$. Innen a 20 dobásban kihúzott piros golyók számának várható értéke $20 \cdot \frac{2}{5} = 8$ és szórásnégyzete $20 \cdot \frac{6}{25} - 380 \cdot \frac{6}{1245} = \frac{144}{49}$.

Feldobunk egy dobókockát, majd ezután egy szabályos pénzdarabot annyi alkalommal, amennyi a kockadobás eredménye volt. (Az egyes dobások eredményei függetlenek egymástól.) Számoljuk ki a fejdobások számának várható értékét és szórásnégyzetét.

Végül tárgyaltuk azt, hogyan lehet kiszámolni független, diszkrét eloszlású valószínűségi változók összegének az eloszlását, majd definiáltuk diszkrét eloszlású valószínűségi változók eloszlásainak a konvolucióját, és néhány megjegyzést tettünk ezzel kapcsolatban.

Legyen ξ és η két független, diszkrét eloszlású valószínűségi változó, amelyek valamilyen $x_1, x_2, \dots$ értékeket vesznek fel. Ekkor ezek összege a $\xi + \eta$ valószínűségi változó szintén diszkrét és eloszlását ki tudjuk számolni a következő módon:

$$P(\xi + \eta = x) = \sum_{(x_j, x_k):\ x_j + x_k = x} P(\xi_j = x_j, \eta = x_k) = \sum_{(x_j, x_k):\ x_j + x_k = x} P(\xi_j = x_j) P(\eta = x_k)$$

Ez a képlet egyszerűsödik abban a fontos speciális esetben, ha a független ξ és η valószínűségi változók csak egész értékeket vesznek fel, és még egyszerűbb, ha csak nem negatív egész értékeket vesznek fel. Azt az eredményt, amelyet ebben a speciális esetben kapunk fogalmazzuk meg tétel formájában.

Tétel. *Legyen ξ és η két független, egész értékeket felvevő valószínűségi változó. Ekkor a $\xi + \eta$ valószínűségi változó is csak egész értékeket vesz fel, és*

$$P(\xi + \eta = n) = \sum_{k=-\infty}^{\infty} P(\xi = k) P(\eta = n - k), \quad n = 0, \pm 1, \pm 2, \dots.$$

Ez a formula tovább egyszerűsödik, ha a független ξ és η valószínűségi változók csak nem negatív egész értékeket vesznek fel. Ekkor

$$P(\xi+\eta=n)=\sum_{k=0}^{n} P(\xi=k)P(\eta=n-k), \quad \text{ha } n\ge 0,$$

és

$$P(\xi+\eta=n)=0, \quad \text{ha } n<0.$$

Bizonyítás: A tétel első képlete azonnal adódik a tétel előtt tárgyalt általános esetből abban az esetben, ha a valószínűségi véltozók $x_1, x_2, \dots$ értékei egész számok. A második képlet az első képlet következménye, ha észrevesszük, hogy elég csak azon $(k, n-k)$ párokra összegezni, amelyekre $P(\xi=k)P(\eta=n-k)\neq 0$. Ez azt jelenti, hogy abban az esetben, ha ξ és η nem negatív egészeket vesznek fel, akkor az összegezésből kihagyhatjuk azokat a tagokat, amelyekre $k<0$ vagy $n-k<0$, azaz $k>n$.

A fenti eredmény miatt érdemes bevezetni a következő definiciót.

Diszkrét, egész értékű eloszlásások konvoluciójának a definiciója. *Legyen*

$$\mathcal{P}=\{p_k\colon k=0,\pm1,\pm2,\dots\} \quad \text{és} \quad \mathcal{Q}=\{q_k\colon k=0,\pm1,\pm2,\dots\}$$

két az egész számokon értelmezett diszkrét eloszás, azaz tegyük fel, hogy $p_k\ge 0$, $q_k\ge 0$, $k=0,\pm1,\pm2,\dots$, $\sum\limits_{k=-\infty}^{\infty} p_k=1$ és $\sum\limits_{k=-\infty}^{\infty} q_k=1$. Ekkor a $\mathcal{P}$ és $\mathcal{Q}$ eloszlások konvoluciója, az a $\mathcal{R}=\mathcal{P}*\mathcal{Q}=\{r_n\colon n=0,\pm1,\pm2,\dots\}$ eloszlás, amelyre

$$r_n=\sum_{k=-\infty}^{\infty} p_k q_{n-k}, \quad n=0,\pm1,\pm2,\dots.$$

Abban a speciáis esetben, ha $p_k=0$ és $q_k=0$ $k<0$ esetén, akkor $r_n=0$ $n<0$ esetén, és

$$r_n=\sum_{k=0}^{n} p_k q_{n-k}, \quad n=0,1,2,\dots.$$

Kiegészítés:

Érdemes a következő észrevételt tenni. Legyen $f(x)$ és $g(x)$ két függvény, amelyek hatványsorba fejthetőek, azaz,

$$f(x)=\sum_{k=0}^{\infty} a_k x^k, \quad g(x)=\sum_{k=0}^{\infty} b_k x^k,$$

ha $-A < x < A$ valamilyen $A > 0$ számmal. Ekkor az $f(x)g(x)$ szorzat is sorba fejthető, azaz

$$f(x)g(x) = \sum_{n=0}^{\infty} c_n x^n$$

(ugyanabban a $-A < x < A$ intervallumban), és

$$c_n = \sum_{k=0}^{n} a_k b_{n-k}.$$

Vegyük észre, hogy a fenti hatványsorokban szereplő a_k, b_k és c_n együtthatók között hasonló reláció érvényes mint ami a konvolució definiciójában szerepelt. Ezenkívül érdemes tudni a következő analízisben bizonyított eredményeket.

a.) Ha az $f(x)$ függvény hatványsorba fejthető, azaz $f(x) = \sum\limits_{k=0}^{\infty} a_k x^k$ valamilyen $-A < x < A$ intervallumon, $A > 0$, akkor az $f(x)$ függvény értékei a $-A < x < A$ intervallumon meghatározzák az a_n együtthatókat. (Sőt azok explicit módon kiszámíthatóak az $n!a_n = \left.\dfrac{d^n f(x)}{dx^n}\right|_{x=0}$ képlet segítségével.)

b.) Ha $f_n(x)$, $n = 1, 2, \dots$, függvényeknek egy olyan sorozata, amelyek mindegyike előállítható $f_n(x) = \sum\limits_{k=0}^{\infty} a_{k,n} x^k$ hatványsor alakban valamilyen $-A < x < A$, $A > 0$ intervallumon, ahol az $A > 0$ szám nem függ az n indextől, továbbá létezik az $f(x) = \lim\limits_{n\to\infty} f_n(x)$ határérték minden $-A < x < A$ számra, akkor az $f(x)$ határfüggvény is hatványsorba fejthető $f(x) = \sum\limits_{k=0}^{\infty} a_k x^k$ alakban a $-A < x < A$ intervallumon. Továbbá $a_k = \lim\limits_{n\to\infty} a_{k,n}$ minden $k = 0, 1, 2, \dots$ együtthatóra.

Később látni fogjuk, hogy a hatványsorok fent említett tulajdonságai jól használhatóak bizonyos valószínűségi vizsgálatokban. Ugyanis egy nem-negatív egész értékékeket felvevő ξ valószínűségi változó $p_k = P(\xi = k)$, $k = 0, 1, 2, \dots$, eloszlásához hozzárendelhetjük az $f(x) = \sum\limits_{k=0}^{\infty} p_k x_k$ függvényt, amelyet ezen eloszlás generátorfüggvényének neveznek. Fogalmazzuk meg pontosabban e fogalom definicióját:

Diszkrét, nem negatív egész értékű eloszlás generátorfüggvényének a definiciója: *Legyen adva egy* $\mathcal{P} = \{p_0, p_1, p_2 \dots\}$ *eloszlás a nem-negatív egész számokon, azaz olyan számsorozat, amelyenk tagjaira* $p_k \ge 0$ *minden* $0 \le k \le \infty$ *számra, és* $\sum\limits_{k=1}^{\infty} p_k = 1$. *(Az elnevezés oka az, hogy létezik olyan* ξ *nem-negatív egész értékű valószínűségi változó, amelyre* $P(\xi = k) = p_k$ *minden* $k = 0, 1, 2, \dots$ *számra.) Ennek az eloszlásnak a generátorfüggvénye az*

$$F(x) = \sum_{k=0}^{\infty} p_k x^k$$

függvény.

A fent mondottak alapján nemcsak az igaz, hogy egy $\mathcal{P}$ eloszlás meghatározza $F(x)$ generátorfüggvényét, hanem megfordítva az $F(x)$ generátorfüggvény segítségével ki lehet számolni azt az eloszlást, amelynek ez a generátorfüggvénye, mivel

$$k!p_k = \left.\frac{dF^k(x)}{dx^k}\right|_{x=0}, \qquad k = 0, 1, 2, \dots.$$

Egy eloszlás generátorfüggvénye mindig konvergens a $-1 \le x \le 1$ intervallumban. Ha egy $\mathcal{P}$ eloszlás generátorfüggvénye $F(x)$ egy $\mathcal{Q}$ eloszlás generátorfüggvénye $G(x)$, akkor $\mathcal{P} * \mathcal{Q}$ konvolúciójuk generátorfüggvénye az $F(x)G(x)$ függvény. Továbbá, ha $\mathcal{P}_n$, $n = 1, 2, \dots$ a nem-negatív egész számokon felvett eloszlások $F_n(x)$ generátorfüggvényei pontonként konvergálnak egy $F(x)$ függvényhez valamely $-A \le x \le A$ intervallumban, akkor az $F(x)$ függvény valamely a nem negatív egész számokon vett $\mathcal{P}$ eloszlás generátorfüggvénye, és a $\mathcal{P}_n$ eloszlások konvergálnak a $\mathcal{P}$ eloszláshoz.

Ezek a tények lehetővé teszik, hogy bizonyos valószínűségi vizsgálatokat viszonylag egyszerűen végrehajthassunk a generátorfüggvények segítségével.