

Az október 12.-i gyakorlat témája

Az előző gyakorlaton tárgyaltuk azt, hogy több megfigyelt véletlen érték összegének a várható értékét hogyan számolhatjuk ki. Fontos annak a vizsgálata is, hogy a kiszámolt várható érték milyen jó közelítése a valódi (véletlen) értéknek. Ennek mérésére vezették be a szórásnégyzet fogalmát. A szórásnégyzetet először csak független valószínűségi változók összegére számoljuk ki. Az általános, nem feltétlenül független valószínűségi változók összegének az esete is érdekes és fontos, de ennek vizsgálatát későbbre halasztjuk. Ahhoz, hogy programunkat végrehajtsuk, először meg kell értenünk a független valószínűségi változók fogalmát. Egyelőre csak diszkrét eloszlású valószínűségi változókkal foglalkozunk, és ezek függetlenségét definiáljuk.

Diszkrét eloszlású valószínűségi változók függetlensége. *Legyenek ξ_1, \dots, ξ_k diszkrét eloszlású valószínűségi változók ugyanazon az (Ω, \mathcal{A}, P) valószínűségi mezőn, amelyek valamilyen x_1, x_2, \dots értékeket vesznek fel. Azt mondjuk, hogy ezek az $\xi_1, \xi_2, \dots, \xi_k$ valószínűségi változók függetlenek, ha*

$$P(\xi_1 = x_{j_1}, \xi_2 = x_{j_2}, \dots, \xi_k = x_{j_k}) = P(\xi_1 = x_{j_1}) P(\xi_2 = x_{j_2}) \cdots P(\xi_k = x_{j_k})$$

minden $j_s = 1, 2, \dots, 1 \leq s \leq k$, indexre.

- 1.) Legyenek ξ_1, \dots, ξ_k független, diszkrét valószínűségi változók egy (Ω, \mathcal{A}, P) valószínűségi mezőn, amelyek valamilyen x_1, x_2, \dots értékeket vesznek fel. Legyenek $A_1 \subset \{x_1, x_2, \dots\}, A_2 \subset \{x_1, x_2, \dots\}, \dots, A_k \subset \{x_1, x_2, \dots\}$ tetszőleges halmazok. Ekkor

$$P(\xi_1 \in A_1, \xi_2 \in A_2, \dots, \xi_k \in A_k) = P(\xi_1 \in A_1) P(\xi_2 \in A_2) \cdots P(\xi_k \in A_k).$$

Továbbá mutassuk meg, hogy tetszőleges $\{j_1, \dots, j_s\} \subset \{1, \dots, k\}$ indexhalmazra a $\xi_{j_1}, \dots, \xi_{j_s}$ valószínűségi változók függetlenek.

Megoldás: A függetlenség miatt

$$\begin{aligned} & P(\xi_1 \in A_1, \xi_2 \in A_2, \dots, \xi_k \in A_k) \\ &= \sum_{x_1 \in A_1} \sum_{x_2 \in A_2} \cdots \sum_{x_k \in A_k} P(\xi_1 = x_1, \xi_2 = x_2, \dots, \xi_k = x_k) \\ &= \sum_{x_1 \in A_1} \sum_{x_2 \in A_2} \cdots \sum_{x_k \in A_k} P(\xi_1 = x_1) P(\xi_2 = x_2) \cdots P(\xi_k = x_k). \end{aligned}$$

Másrészt

$$\begin{aligned} & P(\xi_1 \in A_1) P(\xi_2 \in A_2) \cdots P(\xi_k \in A_k) \\ &= \sum_{x_1 \in A_1} P(\xi_1 = x_1) \sum_{x_2 \in A_2} P(\xi_2 = x_2) \cdots \sum_{x_k \in A_k} P(\xi_k = x_k). \end{aligned}$$

Elvégezve a beszorzásokat kapjuk a feladat első állítását.

A második állítást megkapjuk az első speciális eseteként $A_{j_u} = \{x_{j_u}\}$, $1 \leq u \leq s$ és $A_j = \{x_1, x_2, \dots\}$, $u \in \{1, \dots, k\} \setminus \{j_1, \dots, j_s\}$ választással.

A teljesség kedvéért vezessük be események függetlenségének a definícióját is.

Két esemény függetlenségének a definíciója. Azt mondjuk, hogy egy A és B esemény független, ha $P(A \cap B) = P(A)P(B)$.

Ez a definíció azonban önmagában nem kielégítő számunkra. Beszélni akarunk több esemény függetlenségéről is. Ezért a következő definíciót is bevezetjük.

Több esemény függetlenségének definíciója: Az A_1, \dots, A_n események akkor (teljesen) függetlenek, ha az $\{1, \dots, n\}$ indexhalmaz minden $\{j_1, \dots, j_k\} \subset \{1, \dots, n\}$ részhalmazára

$$P(A_{j_1} \cap \dots \cap A_{j_k}) = P(A_{j_1}) \dots P(A_{j_k}).$$

Események végtelen A_1, A_2, \dots sorozata akkor és csak akkor független, ha tetszőleges pozitív n egész számra az A_1, \dots, A_n események függetlenek.

Speciálisan $n = 3$ esetben ez a definíció a következőt jelenti: Az A , B és C események akkor függetlenek, ha a következő azonosságok mindegyike teljesül:

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

Megjegyzés: Alább bevezetjük események páronkénti függetlenségének az irodalomban szintén használt fogalmát. Fontos, hogy események páronkénti függetlenségének és (teljes) függetlenségének a fogalmát meg tudjuk különböztetni egymástól.

Események páronkénti függetlenségének a definíciója. Legyen A_1, A_2, \dots , események (véges vagy végtelen) sorozata egy valószínűségi mezőn. Azt mondjuk, hogy ezek az események páronként függetlenek, ha tetszőleges különböző számokból álló (j, k) , $j \neq k$, indexekre $P(A_j \cap A_k) = P(A_j)P(A_k)$.

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 definíciója a következő:

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

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

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

Lemma.

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

2.) Minden m valós számra

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

Következésképpen

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

Megoldás:

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

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

Lemma. Minden a és b valós számra

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

Bizonyítás:

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

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

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

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

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

Idézzük fel még a következő eredményt valószínűségi változók várható értékének a kiszámításáról:

Tétel. Legyen ξ diszkrét eloszlású valószínűségi változó, amely valamely x_1, x_2, \dots értékeket vesz fel, és legyen $f(x)$ valamilyen függvény. Ekkor

$$Ef(\xi) = \sum f(x_k)P(\xi = x_k).$$

Speciálisan,

$$E\xi = \sum x_k P(\xi = x_k)$$

$$E\xi^2 = \sum x_k^2 P(\xi = x_k)$$

- 3.) 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, $0 \leq k \leq 100$, $p_k = \binom{100}{k}2^{-100}$. Ezért a dobások számának várható értéke $E\xi = \sum_{k=0}^{100} k \binom{100}{k}2^{-100}$, ahol ξ jelöli a fejdobások számát megadó

valószínűségi változót. Továbbá $E\xi^2 = \sum_{k=0}^{100} k^2 \binom{100}{k}2^{-100}$, a szórásnégyzet pedig

$\text{Var}\xi = E\xi^2 - (E\xi)^2$. Ezeket az összegeket közvetlenül kiszámíthatjuk. Valóban, $k \binom{100}{k} = 100 \binom{99}{k-1}$, $E\xi = \sum_{k=0}^{100} k \binom{100}{k}2^{-100} = 50 \sum_{k=0}^{99} k \binom{99}{k}2^{-99} = 50 \left(\frac{1}{2} + \frac{1}{2}\right)^{99} = 50$. Továbbá, $k^2 \binom{100}{k} = [k(k-1) + k] \binom{100}{k} = 100 \cdot 99 \cdot \binom{98}{k-2} + 100 \cdot \binom{99}{k-1}$, $E\xi^2 = 2^{-100} \cdot 100 \cdot 99 \cdot \sum_{k=0}^{98} \binom{98}{k} + 2^{-100} \cdot 100 \cdot \sum_{k=0}^{99} \binom{99}{k} = 2^{-100} (100 \cdot 99 \cdot (1+1)^{98} + 100 \cdot (1+1)^{99}) = 25 \cdot 99 + 50$, $\text{Var}\xi = E\xi^2 - (E\xi)^2 = 25$.

Valójában a vizsgált várható értéket és szórásnégyzetet egyszerűbben is kiszámíthatjuk. Vezessük be a következő ξ_j valószínűségi változókat: $\xi_j = 1$, ha a j -ik dobás eredménye fej $\xi_j = 0$, ha a j -ik dobás eredménye írás, $1 \leq j \leq 100$. Ekkor

a fejdobások száma $\xi = \sum_{j=1}^{100} \xi_j$, $E\xi = \sum_{j=1}^{100} E\xi_j$, $\text{Var } \xi = \sum_{j=1}^{100} \text{Var } \xi_j$. Mivel $E\xi_j = \frac{1}{2}$, $E\xi_j^2 = \frac{1}{2}$, $\text{Var } \xi_j = \frac{1}{4}$, $1 \leq j \leq 100$, ezért $E\xi = \frac{1}{2} \cdot 100 = 50$, $\text{Var } \xi = \frac{1}{4} \cdot 100 = 25$.

- 4.) Egy szabályos dobókockát feldobunk 100-szor egymás után. Számoljuk ki a dobás-eredmények összegének várható értékét és szórásnégyzetét.

Megoldás: Legyenek $\eta_1, \dots, \eta_{100}$ független valószínűségi változók, amelyekre $P(\eta_j = k) = \frac{1}{6}$, $1 \leq j \leq 100$, $1 \leq k \leq 6$. Ekkor a kiszámítandó várható érték és szórásnégyzet $E \sum_{j=1}^{100} \eta_j = \sum_{j=1}^{100} E\eta_j$, $\text{Var } \sum_{j=1}^{100} \eta_j = \sum_{j=1}^{100} \text{Var } \eta_j$. (A második reláció felhasználja a tekintett valószínűségi változók függetlenségét.) $E\eta_j = \frac{1}{6}(1+2+3+4+5+6) = \frac{7}{2}$, $E\eta_j^2 = \frac{1}{6}(1+4+9+16+25+36) = \frac{91}{6}$, $\text{Var } \eta_j = E\eta_j^2 - (E\eta_j)^2 = \frac{91}{6} - \frac{49}{4} = \frac{182-147}{12} = \frac{35}{12}$, $E\xi = 350$, $\text{Var } \xi = \frac{3500}{12}$.

- 5.) Egy szabályos dobókockát feldobunk egymás után addig, amíg meg nem jelenik a 100. hatos dobás. Számoljuk ki a dobások számának a várható értékét és szórásnégyzetét.

Megoldás: Jelölje T_1 azt a valószínűségi változót, hogy hányszor kellett a kockát feldobni addig, amíg az első hatos dobás megjelent (az első hatos dobást is beleszámítva ebbe a dobássorozatba), T_2 azt, hányszor kellett a kockát az első hatos dobás után feldobni, amíg a második hatos dobás megjelent, és általában legyen T_j az a valószínűségi változó, amelyik azt számolja meg, hogy hány dobás következett be a $j-1$ -k és j -ik hatos dobás között. (A j -ik hatos dobást beszámítjuk a $j-1$ -ik hatos dobást viszont nem számítjuk be ebbe a sorozatba.) Ekkor a minket érdeklő valószínűségi változó, a dobások száma a 100. hatos dobásig $T = \sum_{j=1}^{100} T_j$, továbbá, a

T_j valószínűségi változók függetlenek, és $P(T_j = k) = \left(\frac{5}{6}\right)^{k-1} \frac{1}{6}$, $k = 1, 2, \dots$. Ezért $ET = \sum_{j=1}^{100} ET_j$, $\text{Var } T = \sum_{j=1}^{100} \text{Var } T_j$, $ET_j = \sum_{k=1}^{\infty} k \frac{1}{6} \left(\frac{5}{6}\right)^{k-1}$, $ET_j^2 = \sum_{k=1}^{\infty} k^2 \frac{1}{6} \left(\frac{5}{6}\right)^{k-1}$, $\text{Var } T_j = ET_j^2 - (ET_j)^2$.

Az ET_j várható érték egy lehetséges kiszámolási módja: Deriváljuk a

$$\sum_{k=0}^{\infty} x^k = \frac{1}{1-x}, \quad |x| < 1$$

azonosságot. Azt kapjuk, hogy

$$(*) \quad \sum_{k=0}^{\infty} kx^{k-1} = \frac{1}{(1-x)^2}, \quad |x| < 1,$$

ahonnan $x = \frac{5}{6}$ helyettesítéssel

$$\sum_{k=0}^{\infty} k \left(\frac{5}{6}\right)^{k-1} = 6^2,$$

és $ET_j = 6$, $ET = 100ET_1 = 600$. Hasonlóan, a (*) azonosságot x -szel beszorozva, deviválva, majd az $x = \frac{5}{6}$ behelyettesítést alkalmazva azt kapjuk, hogy

$$\sum_{k=1}^{\infty} k^2 x^{k-1} = \frac{1}{(1-x)^2} + \frac{2x}{(1-x)^3}, \quad |x| < 1,$$

$$ET_j^2 = \frac{1}{6}(36 + 360) = 66, \quad \text{Var } T_j = ET_j^2 - (ET_j)^2 = 30, \quad \text{Var } T = 100\text{Var } T_1 = 3000.$$

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

Megoldás: Vezessük be a következő ξ_j , $1 \leq j \leq 20$, 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 $\xi = \sum_{j=1}^{20} \xi_j$ összeg várható értékét kell kiszámolnunk. Továbbá

$E\xi_j = E\xi_1 = \frac{2}{5}$, mert $E\xi_j = 1 \cdot P(\xi_j = 1) = p(\xi_1 = 1)$. Innen a 20 dobásban kihúzott piros golyók számának várható értéke $20 \cdot \frac{2}{5} = 8$.

Házi feladat:

Egy szabályos dobókockát feldobunk egymás után százszor. Tekintsük a páros értékű dobások összegét, és számítsuk ki ennek várható értékét és szórásnégyzetét.