A sztochasztika alapjai fizikusoknak
5. feladatsor: kovariancia, nevezetes eloszlások
Megoldások

2. A megtakarított pénzünket értékpapírba fektetjük, 20 darabot vásárolunk az A vállalat és 10 darabot a B vállalat részvényeiből. Egy év múlva a két vállalat részvényei várható értékben 700 illetve 1500 dollárt fognak majd érni, az árfolyamok szórása pedig 20 illetve 80 dollár.

(a) Tegyük fel, hogy a részvények árfolyama független egymástól. Várhatóan menynyit fog majd érni a portfóliónk egy év múlva? Mennyi a portfólió értékének a szórása?

(b) Tegyük fel, hogy a részvények árfolyama nem független egymástól. Az árfolyamok közötti korrelációs együttható függvényében írjuk fel formulával és ábrázoljuk a portfólió értékének várható értékét és varianciáját!

(c) Milyen kapcsolat van a korrelációs együttható és a befektetés kockázata között? Ha én egy kockázatkerülő befektető vagyok, akkor pozitív vagy negatív korrelációjú értékpapírokból állítsak össze portfóliót?

Megoldás. Jelölje X az A vállalat, Y pedig a B vállalat részvényének értékét egy év múlva. A feladat szerint $\mathbf{E}(X) = 700$, $\mathbf{D}(X) = 20$, $\mathbf{E}(Y) = 1500$, és $\mathbf{D}(Y) = 80$. Egy év múlva a portfóliónk értéke $20X + 10Y$.

(a) Az összeg várható értéke *mindig* a várható értékek összege, az összeg szórásnégyzete pedig akkor egyenlő a szórásnégyzetek összegével, ha a változók függetlenek (na jó, korrelálatlanság is elég). Tehát, *mindig*

$$\mathbf{E}(20X + 10Y) = 20\mathbf{E}(X) + 10\mathbf{E}(Y) = 29000,$$

és ha X és Y függetlenek, akkor

$$\begin{aligned}\mathbf{D}^2(20X + 10Y) &= \mathbf{D}^2(20X) + \mathbf{D}^2(10Y) = (20)^2\mathbf{D}^2(X) + (10)^2\mathbf{D}^2(Y)\\ &= 160000 + 640000 = 800000,\end{aligned}$$

azaz $\mathbf{D}(20X + 10Y) = 400\sqrt{5}$.

(b) Láttuk,

$$\mathbf{E}(20X + 10Y) = 20\mathbf{E}(X) + 10\mathbf{E}(Y) = 29000,$$

azaz a várható érték nem függ a korrelációtól.

A szórásnégyzet (variancia) az általános esetben, használva a korreláció $\rho(X,Y) = \mathbf{Cov}(X,Y)/(\mathbf{D}(X)\mathbf{D}(Y))$ definícióját, és hogy a kovariancia olyan mint a szorzás, bilineáris

$$\begin{aligned}\mathbf{D}^2(20X+10Y) &= \mathbf{Cov}(20X+10Y, 20X+10Y)\\ &= \mathbf{Cov}(20X,20X) + 2\mathbf{Cov}(20X,10Y) + \mathbf{Cov}(10Y,10Y)\\ &= \mathbf{D}^2(20X) + 2\mathbf{D}(X)\mathbf{D}(Y)\rho + \mathbf{D}^2(10Y)\\ &= 400^2 + 2\cdot 400\cdot 800\cdot \rho + 800^2\\ &= 400^2(5+4\rho).\end{aligned}$$

Ez egyszerű lineáris függvény. Figyeljünk, hogy $\rho \in [-1,1]$.

(c) Nyilván a kockázatos befektetés azért kockázatos, mert nagyon ingadozik a várható értéke körül, azaz a szórása nagyobb. Tehát, ha én kockázatkerülő vagyok, akkor minél kisebb szórásnégyzetű portfóliót szeretnék összerakni. Ezt úgy érhetem el, ha ρ a korreláció negatív.

4. Egy könyvben az egyes oldalakon a sajtóhubák száma egymástól független, Poisson(2) eloszlást követ. Határozzuk meg annak a valószínűségét, hogy a 30. és 31. oldalon sincs hiba! Adjuk meg az ezeken az oldalakon található sajtóhubák várható értékét és szórását!

Megoldás. Először is felidézzük a Poisson-eloszlás definícióját (rosszabb esetben megnézzük a képletgyűjteményben). Eszerint X Poisson-eloszlású $\lambda = 2$ paraméterrel, ha

$$\mathbf{P}(X=k) = \frac{\lambda^k}{k!}e^{-\lambda} = \frac{2^k}{k!}e^{-2}, \quad k = 0,1,2,\ldots.$$

Jelölje X a 30. oldalon, Y a 31. oldalon levő hibák számát. A feladat szerint ezek független, Poisson-eloszlásúak 2 paraméterrel. Az, hogy egyik oldalon sincsen hiba, azt jelenti, hogy $X=0$ és $Y=0$. Ennek a valószínűsége, a függetlenség miatt

$$\mathbf{P}(X=0, Y=0) = \mathbf{P}(X=0)\cdot\mathbf{P}(Y=0) = e^{-2}\cdot e^{-2} = e^{-4}.$$

Tudjuk, hogy λ-paraméterű Poisson várható értéke és szórásnégyzete is λ, ezért

$$\mathbf{E}(X+Y) = \mathbf{E}(X) + \mathbf{E}(Y) = 4,$$

és mivel független változók összegének szórásnégyzete egyenlő a szórásnégyzetek összegével, ezért

$$\mathbf{D}^2(X+Y) = \mathbf{D}^2(X) + \mathbf{D}^2(Y) = 4,$$

azaz $\mathbf{D}(X+Y) = 2$.

5. Egy augusztusi éjszakán megfigyelhető csillaghullások száma Poisson–eloszlást követ. Annak a valószínűsége, hogy egy éjszaka egyetlen hullócsillagot sem látunk 0,1. Várhatóan hány hullócsillag figyelhető meg egy éjszaka?

Megoldás. Itt is Poisson-eloszlás van. Most ki kell találnunk az adatokból a paramétert, mert nincs expliciten megadva, mint az előbb. Legyen X a megfigyelt hullócsillagok száma egy este. Ekkor X ~Poisson(λ). Az, hogy nem látunk hullócsillagot azt jelenti, hogy $X = 0$. Azaz, azt tudjuk, hogy

$$\mathbf{P}(X = 0) = 0,1.$$

Na de a formula szerint ez éppen $\lambda^0/0! \cdot e^{-\lambda} = e^{-\lambda}$. Tehát $e^{-\lambda} = 0,1$ ahonnan $\lambda = -\log 0,1 = \log 10$ (itt a log = ln a természetes alapú logaritmus). A Poisson paramétere tetszőleges pozitív szám lehet, nem kell, hogy egész legyen! Poisson várható értéke a paramétere, azaz

$$\mathbf{E}(X) = \lambda = \log 10.$$

6. Egy biztosítótársaság felmérte, hogy egy év során egy családi ház 0,0002 valószínűséggel gyullad ki. Mennyi a valószínűsége, hogy 2008-ban egy faluban, ahol 15000 ház van, négynél kevesebb tűz üt ki? (Közelítsünk Poisson–eloszlással!)

Megoldás. Itt is Poisson-eloszlással kell számolni. A pontos eloszlás binomiális $n = 15000$ és $p = 0{,}0002$ paraméterekkel, de a binomiális jól közelíthető Poisson-eloszlással ha p kicsi és n nagy.

Először meg kell határoznunk a λ paramétert. Mivel a tűz valószínűsége egy háznál 0,0002, és a faluban 15000 ház van, ezért az összes tűz várható értéke

$$15000 \cdot 0{,}0002 = 3.$$

Tehát egy olyan Poisson-eloszlás kell nekünk, aminek a várható értéke 3. A várható érték éppen a paraméter, tehát $\lambda = 3$. Ezek szerint annak a valószínűsége, hogy négynél kevesebb tűz üt ki az

$$\begin{aligned}\mathbf{P}(X < 4) &= \mathbf{P}(X = 0) + \mathbf{P}(X = 1) + \mathbf{P}(X = 2) + \mathbf{P}(X = 3)\\ &= \sum_{k=0}^{3} \frac{3^k}{k!} e^{-3} = e^{-3}\left(1 + 3 + \frac{9}{2} + \frac{27}{6}\right)\\ &= e^{-3}13.\end{aligned}$$

8. Egy villanykörte élettartama exponenciális eloszlású, átlagosan 2 évig működik. Mennyi a valószínűsége, hogy legalább egy évig fog működni egy új villanykörte? Mennyi a valószínűsége, hogy legalább még egy évig fog működni egy már fél éve működő? Mennyi időt él meg a villanykörték 90%-a?

Megoldás. Exponenciális eloszlásunk van, először meg kell határozni a paramétert.

Legyen egy villanykörte évben mért élettartama X. Ekkor a feladat szerint X exponenciális eloszlású és várható értéke 2. Tudjuk (vagy megnézzük a képletgyűjteményben), hogy az exponenciális várható értéke a paraméter reciproka. Azaz, ha $X \sim$Exp(λ), akkor $\mathbf{E}(X) = 1/\lambda$. Az, hogy egy villanykörte átlagosan 2 évig működik, éppen azt jelenti, hogy az élettartam várható értéke 2. Azaz $2 = 1/\lambda$, ahonnan kapjuk, hogy $\lambda = 1/2$.

Az, hogy egy új villanykörte legalább egy évig jó, azt jelenti, hogy az élettartam nagyobb, mint 1. Tehát a kérdés, hogy mennyi $\mathbf{P}(X > 1)$ (mivel folytonos eloszlásunk van, mindegy, hogy ≥ 1 vagy > 1). Az exponenciális eloszlásfüggvénye $\mathbf{P}(X \leq x) = 1 - e^{-\lambda x}$, ha $x \geq 0$, és 0 különben, ezért

$$\mathbf{P}(X > 1) = 1 - \mathbf{P}(X \leq 1) = e^{-\frac{1}{2}} \approx 0{,}6.$$

A második kérdésnél tudjuk, hogy már fél éve működik az izzónk. Ez egy feltételes valószínűség, hiszen van egy részinformációm, nevezetesen $\{X > 1/2\}$. A kérdés, hogy mennyi a valószínűsége, hogy még legalább 1 évig él, azaz $\{X > 3/2\}$. A feltételes valószínűség definíciója szerint

$$\mathbf{P}(X > 3/2 | X > 1/2) = \frac{\mathbf{P}(X > 3/2, X > 1/2)}{\mathbf{P}(X > 1/2)} = \frac{\mathbf{P}(X > 3/2)}{\mathbf{P}(X > 1/2)} = \frac{e^{-\frac{1}{2}\cdot\frac{3}{2}}}{e^{-\frac{1}{2}\cdot\frac{1}{2}}} = e^{-\frac{1}{2}}.$$

Most lényegében beláttuk, hogy az exponenciális eloszlás örökifjú, ami azt jelenti, hogy tetszőleges $s, t > 0$ esetén $\mathbf{P}(X > s + t | X > s) = \mathbf{P}(X > t)$, azaz ha a változóm már s időt megélt (feltétel!), akkor annak a valószínűsége, hogy még t-t megél az ugyanannyi, mint annak a valószínűsége, hogy t-t megél (nincs feltétel!).

A harmadik kérdésen kicsit el kell mélázni. Az, hogy a időt megél az izzók 90%-a azt jelenti, hogy ha egyszerre becsavarok 100 izzót, akkor a idő után még 90 db világít. Másképpen, 10 égett ki a-ig, azaz annak a valószínűsége, hogy egy izzó élettartam a-nál kisebb, az 0,1. Tehát a kérdés az az a szám, melyre

$$\mathbf{P}(X \leq a) = 0{,}1.$$

(Mivel folytonos eloszlásunk van, ezért mindegy, hogy hol van szigorú egyenlőtlenség, és hol nincs.) Beírva az eloszlásfüggvény képletét ($\lambda = 1/2$, ezt már kiszámoltuk!)

$$0{,}1 = 1 - e^{-\frac{a}{2}},$$

vagyis $a = 2 \log \frac{10}{9} \approx 0{,}21$. Azaz 0,21 évet él meg a villanykörték 90%-a.

11. A skót bakák mellkasának körmérete $N(88, 10)$ eloszlást követ. Mekkora hányaduk fér bele 84-es zubbonyba?

Megoldás. Végre egy normális eloszlás. Most mindkét paraméter meg van adva. A várható érték $\mu = 88$, a szórásnégyzet $\sigma^2 = 10$. Arra figyeljünk, hogy a szórásnégyzet a 10, nem a szórás! Azt fogjuk használni, hogy ha X normális eloszlású μ és σ^2 paraméterekkel, akkor $Z = (X - \mu)/\sigma$ standard normális eloszlású, ami meg táblázatba van szedve. Figyeljünk oda, hogy mi σ és mi σ^2!

Legyen X egy skót baka mellkasának körmérete. Erre úgy gondolunk, hogy a skót katonák összességéből véletlenszerűen kiválasztunk egyet. Ezzel a jelöléssel a feladat éppen a $\{X \leq 84\}$ esemény valószínűségét kérdezi. Ez pedig

$$\mathbf{P}(X \leq 84) = \mathbf{P}\left(\frac{X-88}{\sqrt{10}} \leq \frac{84-88}{\sqrt{10}}\right) = \mathbf{P}(Z \leq -1{,}29) = \Phi(-1{,}26).$$

Az standard normális eloszlásfüggvény táblázatában negatív x-ek nincsenek. De nem baj, tudjuk, hogy a sűrűség páros függvény, amiből következik, hogy

$$\Phi(-x) = 1 - \Phi(x), \quad x > 0.$$

Tehát

$$\Phi(-1{,}26) = 1 - \Phi(1{,}26) = 1 - 0{,}8962 = 0{,}1038.$$

Azaz a skót bakák kb. 10%-a fér bele egy 84-es zubbonyba.

12. Egy munkadarabokat készítő gép 40 cm-re van beállítva. A hiba normális eloszlást követ 0 várható értékkel. Annak a valószínűsége, hogy egy munkadarab nagyobb, mint 40,5 cm, 0,05. Mennyi a szórás?

Megoldás. Jelölje X egy munkadarab hibáját (előjellel, tehát lehet negatív is), azaz a munkadarab hossza $40 + X$ cm. Tudjuk, hogy X normális eloszlású. Most nincs megadva mindkét paramétere a normális eloszlásnak. Azt tudjuk, hogy 0 a várható érték, azaz $\mu = 0$. Legyen σ^2 a szórásnégyzet. Azt tudjuk még, hogy

$$\mathbf{P}(X > 0{,}5) = 0{,}05,$$

hiszen a munkadarab pontosan akkor nagyobb, mint 40,5, ha a hiba nagyobb, mint 0,5. Megint a normális eloszlás skálázási tulajdonságát használjuk. Eszerint $X/\sigma = Z$ standard normális. Tehát

$$\mathbf{P}(X > 0{,}5) = \mathbf{P}(X/\sigma > 0{,}5/\sigma) = \mathbf{P}(Z > 0{,}5/\sigma) = 1 - \Phi(0{,}5/\sigma) = 0{,}05.$$

Innen adódik, hogy

$$\Phi(0{,}5/\sigma) = 0{,}95$$

Most azt kell megkeresni a normális eloszlás táblázatában, hogy hol veszi fel a 0,95 értéket. Ez valahol 1,64 és 1,65 között történik, legyen 1,65. Tehát azt kaptuk, hogy

$$\frac{0{,}5}{\sigma} = 1{,}65,$$

ahonnan
$$\sigma = 0{,}3.$$
Azaz a szórás 0,3.

16. Az FC Barcelona passzolási hatékonysága 2019. áprilisában $p = 0{,}85$ (azaz egy passz ekkora valószínűséggel sikeres). Adjuk meg annak a valószínűségét, hogy a Liverpool elleni 525 passzból legalább 460 sikeres.

Megoldás. Ez a legklasszikusabb példa a de Moivre–Laplace-tétel alkalmazására. Jelölje S a sikeres passzok számát. Mivel összesen 525 passz van, és mindegyik passz egymástól függetlenül $0,85$ valószínűséggel sikeres, ezért S eloszlása binomiális $n = 525$ és $p = 0{,}85$ paraméterekkel. A de Moivre–Laplace-tétel szerint
$$\frac{S - np}{\sqrt{np(1-p)}}$$
közelítőleg standard normális eloszlású, azaz
$$\mathbf{P}\left(a \le \frac{S - np}{\sqrt{np(1-p)}} \le b\right) \approx \Phi(b) - \Phi(a).$$
A pontos állítás az, hogy ha $n \to \infty$, akkor az $\approx$ helyett $=$ van. Ha csak felső korlát van, akkkor $a = -\infty$ és $\Phi(-\infty) = 0$, ha csak alsó, akkor $b = \infty$ és $\Phi(\infty) = 1$. A feladat kérdése a $\mathbf{P}(S \ge 460)$ valószínűség. Egyszerű átalakítással elérjük, hogy $(S - np)/\sqrt{np(1-p)}$ jelenjen meg. Az $np = 525 \cdot 0{,}85 \approx 446$ és $\sqrt{np(1-p)} \approx 8{,}2$ értékeket beírva
$$\begin{aligned}\mathbf{P}(S \ge 460) = \mathbf{P}\left(\frac{S - np}{\sqrt{np(1-p)}} \ge \frac{460 - np}{\sqrt{np(1-p)}}\right) = \mathbf{P}\left(\frac{S - np}{\sqrt{np(1-p)}} \ge 1{,}7\right)\\ \approx 1 - \Phi(1{,}7) = 1 - 0{,}955 = 0{,}045.\end{aligned}$$
Tehát a keresett valószínűsége 0,045.

18. Chicago és Los Angeles között két vasútvonal van, melyek mindegyikén egy-egy vonat közlekedik. Mindkét vonat egyidőben indul, lényegében egyformán kényelmes és k személyes. Tegyük fel, hogy 1000 utas egymástól függetlenül $1/2 - 1/2$ valószínűséggel választ vonatot. Legalább mekkora legyen az ülőhelyek k száma, hogy 0,01-nél kisebb legyen annak a valószínűsége, hogy lesz olyan utas, akinek nem jut ülőhely?

Megoldás. Ez is de Moivre–Laplace. Jelölje S az A társasággal utazók számát. Ekkor a B-vel utazók száma $1000 - S$. Mivel $n = 1000$ ember egymástól függetlenül $1/2 - 1/2$ valószínűséggel dönt A ill. B mellett, ezért S binomiális eloszlású $n = 1000$ és $p = 1/2$ paraméterekkel. Legyen k az ülőhelyek száma (mindkét vonaton). Az, hogy lesz olyan, akinek nem jut hely, azt jelenti, hogy vagy az A társaságnál túl sokan vannak, vagy a B-nél. Azaz, vagy $S > k$ (A vonat betelik) vagy $1000 - S > k$ (B vonat betelik). A kérdés a legkisebb olyan k érték, melyre

$$\mathbf{P}(S > k \text{ vagy } S < 1000 - k) \leq 0{,}01.$$

Vegyük észre, hogy $k \geq 500$ kell legyen, és ekkor csak az egyik vonat telhet be, azaz a fenti valószínűségben szereplő két esemény egymást kizáró, ahonnan

$$\mathbf{P}(S > k \text{ vagy } S < 1000 - k) = \mathbf{P}(S > k) + \mathbf{P}(S < 1000 - k).$$

Használjuk a de Moivre–Laplace-tételt. Beírjuk az $np = 500$, $\sqrt{np(1-p)} = 5\sqrt{10} \approx 15{,}8$ értékeket. Ekkor

$$\mathbf{P}(S > k) = \mathbf{P}\left(\frac{S - np}{\sqrt{np(1-p)}} > \frac{k - np}{\sqrt{np(1-p}}\right) \approx 1 - \Phi((k - 500)/15{,}8),$$

és ugyanígy (csak felhasználjuk, hogy $\Phi(-x) = 1 - \Phi(x)$ ha $x > 0$, és hogy $500 - k \leq 0$)

$$\mathbf{P}(S < 1000 - k) = \mathbf{P}\left(\frac{S - np}{\sqrt{np(1-p)}} > \frac{1000 - k - np}{\sqrt{np(1-p}}\right) \approx 1 - \Phi((k - 500)/15{,}8).$$

Azaz, keressük azt a legkisebb k értéket melyre

$$2\left[1 - \Phi((k - 500)/15{,}8)\right] \leq 0{,}01.$$

Átrendezve,

$$\Phi((k - 500)/15{,}8) \geq 0{,}995.$$

Kikeressük a normális eloszlás táblázatából, hogy hol veszi fel Φ a 0,995 értéket. Ez 2,57, azaz

$$\frac{k - 500}{15{,}8} \geq 2{,}57,$$

amit átrendezve

$$k \geq 540{,}6.$$

Mivel k egész, ezért 541 a legkisebb olyan k, amire a feltétel teljesül.