

# 1.gyakorlat

2015.09.11.

- I. János, Jakab, József, Joli és Jenő egy együttest alkotnak, mely öt hangszeren játszik. Ha mindegyikük tud mind az 5 hangszeren játszani, hányféle elrendezés lehetséges? És, ha János, Jakab és Joli mind az 5 hangszeren játszhat, de József és Jenő mindketten dobolni és zongorázni tudnak?
- II. Hányféle (esetleg értelmetlen, de különböző) szót lehet kirakni a *MISSISSIPPI* betűiből (mindet pontosan egyszer használva)?
- III. Hányféleképpen lehet sorba rendezni 8 kék, 5 fekete és 2 piros tollat, ha az azonos színű tollak között nem teszünk különbséget?
- IV. Hányféle rendszámot lehet kiadni a mai magyar rendszerben (pl. EWA-590)? És ha kiadjunk 5 ronda szót?
- V.
  - a) Hányféleképpen ülhet le egy sorban 4 lány és 3 fiú?
  - b) Hányféleképpen ülhet le egy sorban 4 lány és 3 fiú, ha a lányok egymás mellett ülnek, és a fiúk is egymás mellett ülnek?
  - c) És ha csak a fiúk kell, hogy egymás mellett üljenek?
  - d) Hányféleképpen ülhetnek le, ha azonos neműek nem ülhetnek egymás mellé?
- VI. Egy társaság 8 nőből és 7 férfiből áll. Belőlük kell egy 4 nőből és 3 férfiből álló bizottságot alakítanunk. Hányféle különböző bizottság lehetséges, ha
  - a) a férfiak közül 2 nem hajlandó egy bizottságban dolgozni?
  - b) egy nő és egy férfi nem hajlandó egy bizottságban dolgozni?
- VII. Egy árverésen 4 műgyűjtő vásárolt összesen 5 Dalit, 6 van Goghot, 7 Picassót. Ha egy tudósító csak annyit jegyez fel, hogy melyik gyűjtő hány Dalit, van Goghot, és Picassót vásárolt, akkor hányféle különböző feljegyzés születhetett?
- VIII. Másik oldalon.

| Kísérlet                                                 | $\Omega$ (absztrakt esemény-tér) | $ \Omega $ | E (esemény) | $\mathbb{P}(E) =$ |
|----------------------------------------------------------|----------------------------------|------------|-------------|-------------------|
| Esik-e ma az eső?                                        |                                  |            |             |                   |
| Lóverseny 7 lóval. Mi a befutók sorrendje?               |                                  |            |             |                   |
| Két érmével dobunk.                                      |                                  |            |             |                   |
| Két kockával dobunk.                                     |                                  |            |             |                   |
| $k$ db különböző golyót $n$ urnába szeretnénk beletenni. |                                  |            |             |                   |
| $k$ db egyforma golyót $n$ urnába szeretnénk beletenni.  |                                  |            |             |                   |
| A kockával addig dobunk, amíg 6-os nem lesz.             |                                  |            |             |                   |
| Egy eszköz élettartama (pl. hónapokban)                  |                                  |            |             |                   |

## 2-3.gyakorlat

2015.09.18. és 2015.09.25.

- I. Mi a valószínűsége, hogy egy véletlenszerűen kiválasztott 6 jegyű szám jegyei mind különbözők? Mi az eseménytér?
- II. Egy urnában 6 piros és 5 kék golyó van. Véletlenül hármat húzva (visszatevés nélkül) mi a valószínűsége, hogy pontosan 1 piros és 2 kék golyót húzunk? Számoljuk ki a valószínűséget úgy, ha a húzások sorrendjére tekintettel vagyunk, és úgy is, ha nem! Írjuk fel mindkét esetben az eseményteret!
- III. 10 házaspárt véletlenszerűen leültetünk egy kerek asztalhoz. Mi a valószínűsége, hogy egy férj és felesége sem ül egymás mellé?
- IV. Egy ládában van
- 5 jó égő, mely 1 hónapnál tovább bírja az üzemet
  - 10 erősen használt égő, ezek 2 nap múlva kiégnek
  - 10 már kiégett égő.
- Egy égőt találmra választva és kipróbálva, felgyullad. Mi a valószínűsége, hogy 1 hét után is működik majd? A feladatot oldd meg a feltételes valószínűség definíciójával és redukált eseménytérrel is!
- V. Három kockát feldobunk. Feltéve, hogy a dobott számok között nincs két egyforma, mennyi a valószínűsége, hogy (legalább) az egyiken 6-os van?
- VI. Két kockával dobunk. Tekintsük a következő eseményeket:
- $E = \{\text{A dobások összege } 6\}$   
 $F = \{\text{A dobások összege } 7\}$   
 $G = \{\text{Az első kocka } 3\text{-at mutat}\}$   
 $H = \{\text{A második kocka } 4\text{-et mutat}\}$
- a) Független-e  $E$  és  $G$ ?
- b) Független-e  $F$  és  $G$ ?
- c) Független-e  $F$ ,  $G$  és  $H$ ?
- d) Kettőnél több esemény esetében a páronkénti függetlenségből már következik-e az események (teljes) függetlensége?
- VII. Egy biztosító két csoportba osztja az autóvezető ügyfeleit:
- a jó vezetők alkotják az ügyfelek 70%-át, ők minden évben 0,2 valószínűséggel okoznak balesetet
  - a rossz vezetők alkotják az ügyfelek 30%-át, ők minden évben 0,4 valószínűséggel okoznak balesetet.
- a) Mi a valószínűsége, hogy a biztosító egy ügyfele balesetet okoz idén?
- b) Mi a valószínűsége, hogy a biztosító egy ügyfele jó vezető, ha tudjuk, hogy balesetet okozott idén?
- VIII. Egy urnában 8 kék és 4 piros golyó van. Húzunk belőle kettőt.
- a) Mi a valószínűsége, hogy mindkettő kék?
- b) Mi a valószínűsége, hogy a második kihúzott golyó piros?

- c) Mi a valószínűsége, hogy az első golyó kék, ha tudjuk, hogy másodiknak pirosat fogunk húzni?
- IX. Pistike elindul a Zrínyi Ilona matematika versenyen, ahol minden egyes feladatnál 5 lehetséges megoldás közül kell kiválasztania a helyeset. Pistike minden kérdésre  $p$  valószínűséggel tudja a helyes választ,  $1 - p$  valószínűséggel pedig tippel, azaz az 5 lehetőség közül egyenletesen választ.
- a) Mi a valószínűsége, hogy ha helyesen válaszol, akkor tényleg tudja a helyes megoldást?
- b) Mi a valószínűsége, hogy Pistike eltalálja a helyes megoldást?
- c) Mi a valószínűsége, hogy az első öt feladat megoldását tudja, ha tudjuk, hogy mind az öt válasza helyes lett, és a kérdésekre egymástól függetlenül ad választ?
- X. Egy felmérés szerint a lakosság 0,5%-a HIV+. Egy "megbízhatónak számító" diagnosztikai eljárás egy HIV+ beteg esetén 95%-os valószínűséggel kimutatja a fertőzést, míg 5% valószínűséggel nem. A teszt egészséges ember esetén 99%-os valószínűséggel negatív, míg 1% eséllyel fals pozitív. Mi a valószínűsége, hogy ha a teszt eredménye pozitív, akkor a páciens valóban beteg?  
Hf: Mi a valószínűsége, hogy ha a teszt eredménye negatív, akkor a páciens valóban nem fertőzött?
- XI. Móricka feldob egy (szabálytalan) érmét 11-szer egymás után, a fej valószínűsége  $p$ . Nézzük a következő eseményeket:  
 $A = \{\text{a dobott érmék között pontosan 8 fej van}\}$   
 $B = \{\text{a negyedik dobás eredménye írás}\}$
- a) Add meg az eseményteret! Hány elemű?
- b)  $\mathbb{P}(A), \mathbb{P}(B) = ?$
- c) Van-e olyan  $p$ , amire  $A$  és  $B$  események függetlenek egymástól?
- XII. Egy játékos annyiszor lőhet egy léggömbre, ahány 6-ost dobott egymás után egy kockával (pl. ha dob egy 6-ost, majd egy 2-est, akkor egyszer lőhet). Mennyi annak a valószínűsége, hogy szétlővi a léggömböt, ha egy lövésnél  $\frac{1}{1000}$  valószínűséggel talál?

## XII. feladat megoldása:

Legyen  $A := \{\text{szétlővi a léggömböt}\}$  esemény, kérdés a  $\mathbb{P}(A)$  valószínűség.  
Alkalmazzuk a teljes valószínűség tételét (TVT), feltételezzünk be arra, hogy (pontosan!) hány-szor lőhet,  $A_i := \{\text{pontosan } i\text{-szer lőhet}\}$ , ahol  $i = 0, 1, 2, \dots$

$$\mathbb{P}(A) = \sum_{i=1}^{\infty} \mathbb{P}(A|A_i) \cdot \mathbb{P}(A_i)$$

(A szummát igazából nulláról kéne indítani, de az  $i = 0$  esetben a feltételes valószínűség annak az eseménynek a valószínűsége, hogy eltalálja a léggömböt, feltéve, hogy egyszer sem lőhet, ez nyilván egy nulla szorzó.)

Az  $A_i$  esemény pontosan akkor következik be, ha az első  $i$  darab kockadobás mind 6-os, majd dobunk egy nem 6-ost, így:

$$\mathbb{P}(A_i) = \left(\frac{1}{6}\right)^i \cdot \frac{5}{6}.$$

(A gyakorlaton csak annak a valószínűségét számoltuk ki, hogy dobunk legalább  $i$  darab 6-ost az elején, majd mindegy mit dobunk (akár 6-ost is dobhatunk). Ezek az események azonban nem alkotnak teljes eseményrendszert, nem kizáróak, így nem alkalmazható rájuk a TVT.)

A  $\mathbb{P}(A|A_i)$  feltételes valószínűséget a legkönnyebben úgy lehet kiszámolni, ha észrevesszük, hogy a komplementer esemény könnyen számolható:

$$\mathbb{P}(A|A_i) = 1 - \mathbb{P}(\bar{A}|A_i) = 1 - \left(\frac{999}{1000}\right)^i,$$

hiszen a komplementer esemény az, hogy nem lövi szét a léggömböt, feltéve, hogy  $i$ -szer lő, vagyis egyszer sem találja el, feltéve, hogy  $i$ -szer lő, a lövések függetlenek, és egy lövésnél a "nem eltalálás" valószínűsége  $1 - \frac{1}{1000} = \frac{999}{1000}$ .

A megoldás:

$$\mathbb{P}(A) = \sum_{i=1}^{\infty} \left(1 - \left(\frac{999}{1000}\right)^i\right) \cdot \left(\frac{1}{6}\right)^i \cdot \frac{5}{6}$$

Igazából ennél a pontnál kész a feladat, meg szeretném magyarázni, hogy mit rontottam el, ehhez kicsit nézzük a fentiek mögé formálisan:

Legyen  $E_k := \{\text{(pontosan a) } k\text{-adikra lövi szét a léggömböt}\}$  esemény. Ekkor  $A = \cup_{k=1}^{\infty} E_k$ , és

$$\begin{aligned} \mathbb{P}(A|A_i) &= \mathbb{P}(\cup_{k=1}^i E_k | A_i) = \\ &= \begin{cases} 1 - \mathbb{P}(\cap_{k=1}^i \bar{E}_k | A_i) = 1 - \left(\frac{999}{1000}\right)^i & \text{ezt csináltuk előbb} \\ \sum_{k=1}^i \mathbb{P}(E_k | A_i) & \text{, hiszen kizárók és } \mathbb{P} \text{ } \sigma\text{-additív} \end{cases} \end{aligned}$$

Én a

$$\mathbb{P}(A|A_i) = \sum_{k=1}^i \mathbb{P}(E_k | A_i)$$

módon próbáltam számolni, csak valahogy nagyon rosszul...

Könnyen látható (gondoljunk bele mit jelent az  $E_k | A_i$  esemény), hogy

$$\mathbb{P}(E_k | A_i) = \left(\frac{999}{1000}\right)^{k-1} \cdot \frac{1}{1000}$$

Innen:

$$\mathbb{P}(A|A_i) = \sum_{k=1}^i \left(\frac{999}{1000}\right)^{k-1} \cdot \frac{1}{1000}$$

Használva a mértani sorozatok összegképletét megkapjuk az  $\mathbb{P}(A|A_i) = 1 - \left(\frac{999}{1000}\right)^i$  eredményt.

## 4-5.gyakorlat

2015.10.02. és 2015.10.09.

- I. Három szabályos érmével dobva jelölje  $X$  a kapott fejek számát. Milyen értékeket vehet fel a valószínűségi változó? Milyen valószínűségekkel jöhetnek ki ezek az értékek? Egyszerűbben egyben a két kérdés: mi  $X$  eloszlása?
- II. Legyen  $X$  nemnegatív egész értékű valószínűségi változó (vv.), azaz  $X = 0, 1, 2, 3, \dots$ , továbbá legyen  $\lambda > 0$  valós paraméter, a súlyfüggvény

$$p(i) = \mathbb{P}(X = i) = c \cdot \frac{\lambda^i}{i!}.$$

- a) Milyen  $c$ -re lesz ez egy súlyfüggvény (avagy eloszlás)?
- b)  $\mathbb{P}(X = 0) = ?$
- c)  $\mathbb{P}(X > 2) = ?$

- III. **Negatív binomiális eloszlás.** Független kísérleteket végzünk, mind  $p$  valószínűséggel sikeres,  $1 - p$  valószínűséggel sikertelen. Jelölje  $X$  azt, hogy hanyadikra végezzük az  $r$ -edik sikeres kísérletet ( $X \sim \text{NB}(p, r)$ , ahol  $0 < p < 1$  valószínűség és  $r$  pozitív egész). Mi az  $X$  eloszlása (lehetséges értékek, súlyfüggvény), várható értéke, szórásnégyzete? Milyen kapcsolat van a geometriai és a negatív binomiális eloszlás között?

- IV. 3 férfit és 3 nőt rangsorolnak egy vizsgán. Tfh. nincs két egyforma pontszám (azaz a  $6!$  lehetséges sorrend mindegyike egyformán valószínű). Legyen  $X$  a legjobb nő helyezése (pl.  $X = 1$  azt jelenti, hogy a legjobb versenyző egy nő volt). Mi az  $X$  eloszlása ill. várható értéke?

- V. Mi  $n$  kockadobás maximumának eloszlása? (hf: és a minimumé?)

- VI. 5-ször dobunk egy szabályos kockával. Legyen  $X$  a 6-os dobások száma.  $\mathbb{D}^2(X) = ?$

- VII. a) Háromszor dobunk egy hamis érmével, amin a fej valószínűsége  $\frac{3}{8}$ . Jelölje  $X$  a fej dobások számát. Számoljuk ki az  $X$  várható értékét és szórását!
- b) Háromszor dobunk egy hamis érmével, amin a fej valószínűsége  $\frac{3}{8}$ . Jelölje  $U$  azt a számot ahányszor sikerült az előző dobást megismételni (az  $U$  lehetséges értékei  $0, 1, 2$ ). Számoljuk ki  $U$  várható értékét és szórását!

- VIII. Thf. repülés közben a repülőgép motorjai egymástól (teljesen) függetlenül  $1 - p$  valószínűséggel hibásodnak meg.

Ha egy repülőgépnek a repüléshez a motorjainak legalább a felére van szükség, akkor milyen  $p$  értékre mondhatjuk, hogy legalább 95% valószínűséggel nem zuhan le az 5 motoros gép?

- IX. Egy sportlövő  $\frac{1}{7}$  valószínűséggel talál el egy léggömböt. Az ötödik találatig lő. Mennyi a lövések számának eloszlása ill. várható értéke? Mennyi "idő" telik el a lövések között?

- X.<sup>gyak</sup> Egy tóban lévő  $N$  halból 100 darab ponty. Egy bácsi addig horgászik, amíg ki nem fog egy pontyot. Mi a kifogott halainak az eloszlása, ha

- a) a halakat visszaengedi
- b) nem engedi vissza a halakat?

- XI.<sup>gyak</sup> **Hipergeometriai eloszlás.** Egy erdő  $N$  őze közül  $m$ -et megjelöltek. Az  $N$  őz közül most befogunk  $n$ -et, mindegyiket egyenlő valószínűséggel. Legyen  $X$  a megjelölt őzek száma a befogottak között. Ekkor mondjuk, hogy  $X$  hipergeometriai eloszlású. Írd fel  $X$  súlyfüggvényét! Nézz utána, hogy mi a várható értéke illetve a szórásnégyzete  $X$ -nek! (tipp: Balázs Márton jegyzetének 45. oldala)

XII. Egy bulvárlapban oldalanként várhatóan 0,2 nyomtatási hiba van. Mi a valószínűsége, hogy a következő oldalon

- a) 0
- b) 2 vagy annál több hiba van?

XIII. Egy államban az öngyilkossági ráta 1 öngyilkosság per 100.000 lakos per hónap. Vizsgáljuk meg az állam egy 300.000 lakosú városát!

- a) Mi a valószínűsége annak, hogy egy adott hónapban legalább 4 ember lesz öngyilkos?
- b) Mi a valószínűsége annak, hogy egy adott évben legalább 2 hónapban lesz legalább 4 öngyilkosság?
- c) Legyen a mostani hónap az 1., mi a valószínűsége annak, hogy először az  $i$ -edik hónapban lesz legalább 4 öngyilkosság? Várhatóan mikor lesz először legalább 4 öngyilkosság?

Milyen feltevésekkel éltünk?

XIV. Átlagosan hány mazsolának kell egy sütiben lennie, ha azt kívánjuk elérni, hogy egy véletlenszerűen választott sütiben legalább 0,99 valószínűséggel legyen (legalább 1) mazsola? (A mazsolák számát egy sütiben modellezhetjük Poisson eloszlással.)

XV.<sup>HF</sup> London központi kerületében a bekövetkező autóbalesetek száma száraz, napos időben  $\lambda = 10$  paraméterű Poisson-eloszlású, nedves időben  $\mu = 20$  Poisson-eloszlású. Kora novemberben Londonban  $p = 0,6$  valószínűséggel van ronda esős idő,  $q = 0,4$  valószínűséggel van verőfényes napsütés.

Azt olvastam a Times-ban, hogy múlt csütörtökön 17 autóbaleset volt London központi kerületében. Mit mondhatunk ezen információ birtokában arról, hogy múlt csütörtökön esett-e a eső? **(Beadható a házi feladatok között +5 pontért! Részletes indoklással!)**

XVI. A "Kocogj velünk!" mozgalom keretében futóversenyt rendeztek a Dunakanyarban. A pályát kullancsok fertőzött területen vezették át. A versenyzők közül 300-ban találtak 1, 75-ben pedig 2 kullancsot. Ennek alapján becsüljük meg kb. mennyien indultak a versenyen!

XVII. Az  $X$  valószínűségi változó eloszlásfüggvénye a következő:

$$F(x) = \begin{cases} 0, & x \leq 0 \\ x/2, & 0 < x \leq 1 \\ 2/3, & 1 < x \leq 2 \\ 11/12, & 2 < x \leq 3 \\ 1, & x > 3 \end{cases}$$

Ennek ismeretében adjunk választ a következő valószínűségekre:

$\mathbb{P}(X = 1), \mathbb{P}(X = 2), \mathbb{P}(X = \frac{1}{2}), \mathbb{P}(X < 3), \mathbb{P}(X \geq \frac{1}{2}), \mathbb{P}(2 \leq X < 4), \mathbb{P}(2 < X \leq 4)$

(Hf: Balázs Márton jegyzetben megnézni az 50-51-dik oldalt, meghallgatni hozzá az anyagot: 4.ea vége kb. 1:23:00-tól 5.ea kb. 14:45:00-ig, itt már megemlíti az AF eloszlásokat is, és mutat egy példát szinguláris vv-ra, ez utóbbi csak érdekesség, előbbi még nem lesz anyaga a zh-nak, de csupán a definíciót említi. Azt kell érteni, hogy mit fejez ki az eloszlásfüggvény!!!)

## 6-7-8.gyakorlat

2015.11.06. és 2015.11.13. és 2015.11.20.

I. Az  $X$  valószínűségi változó eloszlásfüggvénye a következő:

$$F(x) = \begin{cases} 0, & x \leq 0 \\ x/2, & 0 < x \leq 1 \\ 2/3, & 1 < x \leq 2 \\ 11/12, & 2 < x \leq 3 \\ 1, & x > 3 \end{cases}.$$

Ennek ismeretében adjunk választ a következő valószínűségekre:

$\mathbb{P}(X = \frac{1}{2})$ ,  $\mathbb{P}(X < 3)$ ,  $\mathbb{P}(X \leq 3)$ ,  $\mathbb{P}(X \geq \frac{1}{2})$ ,  $\mathbb{P}(2 \leq X < 4)$ ,  $\mathbb{P}(2 < X \leq 4)$

II. Egy rendszer a bekapcsolástól számított  $X$  hónapig működik. Határozzuk meg az  $X$  várható értékét, és annak a valószínűségét, hogy a rendszer legalább 5 hónapig működik, ha  $X$  sűrűségfüggvénye:

$$f(x) = \begin{cases} \frac{C}{(x+1)^2} & , x > 0 \\ 0 & , x \leq 0 \end{cases}$$

III. A buszok rendre minden óra egészkor, 15-kor, 30-kor és 45-kor indulnak a megállóból. Ha egyenletesen érkezem 7:00 és 7:30 között, akkor mi a valószínűsége, hogy

- a) 5 percnél kevesebbet várok?
- b) 8 percnél többet várok?
- c) ugyanez a két kérdés, ha 7:08 és 7:38 között érkezem egyenletesen!

IV. A "fény az éjszakában" típusú villanykörte élettartama exponenciális eloszlású. A gyártó mérése szerint a körték 95 százaléka bírja legalább egy évig. Mennyi időre vállalhat a gyártó garanciát a körték működésére, ha azt akarja, hogy a vevőknek legfeljebb 0,5 százaléka reklamáljon?

V. Reggel a földalatti szerelvények követési ideje exponenciális eloszlású 3 perc várható értékkel. Az egyik szerelvényt pont lekéstem.

- a) Mi a valószínűsége, hogy legalább 5 percet várnom kell a következőre?
- b) Már 4 perce várok. Mennyi a valószínűsége, hogy még további 6 percet fogok várni?

VI. a) Legyen  $X$  3 várható értékű, 2 szórású normális eloszlású valószínűségi változó. Mennyi a valószínűsége, hogy pozitív?

- b) Egy normális valószínűségi változó szórása 2. Mennyi a várható értéke, ha annak valószínűsége, hogy a vv. negatív 0,8438.

VII. Tegyük fel, hogy a 25 éves fiatalok magassága centiméterben mérve normális eloszlású,  $\mu = 182$  és  $\sigma^2 = 169$  paraméterekkel.

- a) A 25 éves fiatalok hány százaléka magasabb 2 méternél?
- b) A kétméteres klub tagjainak hány százaléka magasabb 2 méter 10 centinél?

VIII. Milyen eloszlású két független, különböző paraméterű Poisson eloszlású valószínűségi változó összege? (Hf: Na és háromé?)

IX. Milyen eloszlású két független, azonos paraméterű Exponenciális valószínűségi változó összege? (Hf: Nézz utána az ún. Gamma eloszlásnak!)

- X. Határozzuk meg a binomiális és a negatív binomiális valószínűségi változók várható értékét, kihasználva a várható érték linearitását!

- XI. Legyen

$$X = \begin{cases} 1 & , 1/3 \text{ vsz.} \\ 0 & , 1/3 \text{ vsz.} \\ -1 & , 1/3 \text{ vsz.} \end{cases} \quad Y = \begin{cases} 1 & , \text{ha } X = 0 \\ 0 & , \text{ha } X \neq 0 \end{cases}$$

Határozzuk meg a két valószínűségi változó kovarianciáját. Független  $X$  és  $Y$ ?

- XII. Határozzuk meg a binomiális és a negatív binomiális valószínűségi változók szórásnégyzetét!

- XIII. **Kalapos feladat:**  $n$  ember színházba megy, mindenki leteszi kalapját a ruhatárba. Összekeverednek a kalapok, véletlenszerűen visszaosztják a kalapokat az  $n$  embernek. Jelölje  $X$  azon emberek számát, akik saját kalapjukban mennek haza. Határozd meg  $X$  várható értékét és szórásnégyzetét. (*Tipp: legyen  $X_i$  az  $i$ -edik emberhez tartozó indikátor,  $X_i = 1$  pontosan akkor, ha az  $i$ -edik ember a saját kalapjában megy haza.*)

- XIV. Legyen  $X_1, X_2, \dots, X_n$  faevv-k sorozata. Ezen minta alapján adjunk "jó tippet" a mögöttes eloszlás várható értékére és szórásnégyzetére.

*Útmutatás:*

Mutassuk meg, hogy a mintaátlagnak és a korrigált empirikus szórásnégyzetnek nevezett statisztikák konzisztens becslést adnak. Győződjünk meg róla, hogy ezen becslések ("tippek") szórásnégyzete nagy mintaelemszám mellett "kicsi". (Az  $S^2$  szórásnégyzetének kiszámítása házi feladat.)

Mintaátlag:  $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$

Korrigált empirikus szórásnégyzet:  $S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$

- XV. Egy gyárban ketyeréket gyártanak, várhatóan napi 50-et.

- Adjunk korlátot arra a valószínűségre, hogy holnap legalább 75 ketyere készül!
- Mit tudunk mondani arról a valószínűségről, hogy holnap 40 és 60 között lesz a termelés, ha a szórásnégyzet 25?

- XVI. Legyen  $X$  egyenletes a  $(0, 10)$  intervallumon. Mit tudunk mondani arról a valószínűségről, hogy  $X$  jobban eltér a várható értékétől, mint 4?

- XVII. Legyen  $X$  normális eloszlású. Mit tudunk mondani arról a valószínűségről, hogy  $X$  jobban eltér a várhatóértékétől, mint a szórás kétszerese?

- XVIII. Legyen  $X_1, X_2, \dots, X_{10}$  fae  $E(0, 1)$  vv-k. Határozzuk meg a  $\left\{ \sum_{i=1}^{10} X_i > 6 \right\}$  esemény valószínűségét!

- XIX. Jelölje  $X$  egy tanfolyamra beiratkozók számát, melyről feltételezzük, hogy Poisson eloszlású 100 várható értékkel. Mekkora annak a valószínűsége, hogy többen iratkoznak be a tanfolyamra, mint 120?

- XX. Egy csillagász távolságot mér, mérései faevv sorozatnak tekinthető. A mérések várható értéke megegyezik az objektum valódi  $d$  távolságával, a szórás 2 fényév. Hányszor mérjen a csillagász, hogy az átlag 95 százalékos valószínűséggel közelebb legyen a valódi távolsághoz, mint  $\pm 0,5$ ? Mennyire jár rosszul a csillagász, ha nem ismeri a CHT-t, csak a Csebisev-egyenlőtlenséget?

## 9-10.gyakorlat

2015.11.27. és 2015.12.04.

Felhasználom Móri F. Tamás, Szeidl László, Zempléni András: Matematikai statisztika példatárjának elméleti összefoglalóit és a 2/14, 3/6,13,15,19-es példáit.

### Becslésemélet

A matematikai statisztika egyik alapfeladata a véletlen jelenségeket leíró valószínűségeloszlás ismeretlen paraméterének (paramétereinek) becslése. A rendelkezésre álló információt a minta biztosítja, amely tipikusan független azonos eloszlású valószínűségű változók  $\mathbf{x} = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n$  realizációinak, megfigyeléseink sorozata (ezek az egymástól függetlenül, azonos körülmények között elvégzett  $n$  kísérlet eredményei), ez alapján adhatjuk meg a  $\vartheta$  paraméter (mely vektorértékű is lehet) becslését, amely a minta valamely  $T(\mathbf{x})$  függvénye.

Talán a leggyakoribb becslési módszer a *maximum likelihood elv*, amely azon alapul, hogy azt a paraméterértéket választjuk becslült értéknek, amely mellett az adott minta realizációjának a valószínűsége a legnagyobb. A becslés kiszámításához bevezetjük a likelihood-függvényt ( $L_{\vartheta}(\mathbf{x})$ ), ami a realizáció bekövetkezésének valószínűségét fejezi ki (együttes eloszlások). Független azonos eloszlású esetben a likelihood-függvény megegyezik a súly-ill. sűrűségfüggvények szorzatával. Célunk ezt a függvényt maximalizálni a paraméter(ek)ben. Becslésünket gyakran a loglikelihood-függvény paraméter szerinti deriváltjának gyöke adja (ezt az egyenletet nevezzük likelihood-egyenletnek).

Legyen  $X_1, X_2, \dots, X_n$  fae vv-k sorozata, mely diszkrét (D) vagy abszolút folytonos (AF) eloszlásból származik, akkor a likelihood-függvény:

$$L_{\vartheta}(\mathbf{x}) = \begin{cases} \prod_{i=1}^n p_{\vartheta}(x_i) & \text{D eset} \\ \prod_{i=1}^n f_{\vartheta}(x_i) & \text{AF eset} \end{cases}$$

I. Határozd meg a következő nevezetes eloszlások ismeretlen paramétereinek ML-becslését!

- a)  $Poi(\lambda)$
- b)  $Geom(p)$
- c)  $B(m, p)$ , ahol  $m$  adott
- d)  $Exp(\lambda)$
- e)  $U(a, b)$
- f)  $N(\mu, \sigma_0^2)$ , ahol  $\sigma_0$  adott
- g)  $N(\mu, \sigma^2)$ , ahol  $\sigma$  is ismeretlen

II. Azt feltételezik, hogy egy fiúkból álló osztály testmagassága normális eloszlású. Adjunk ML-becslést a normális eloszlás paramétereire a 173, 169, 180, 178, 176, 190 megfigyelés sorozat esetén!

III. A családok jövedelmét egy olyan skálán mérjük, ahol  $X = 1$  a létminimumnak felel meg. Feltételezzük, hogy a jövedelem eloszlása az

$$f(x) = \frac{\vartheta}{x^{\vartheta+1}} \mathbb{I}(x \geq 1)$$

sűrűségfüggvénnyel adható meg (ez az ún. *Pareto-eloszlás*). Adjunk maximum likelihood becslést  $\vartheta$ -ra, ha 10 véletlenszerűen választott család jövedelme

1, 53, 2, 76, 19, 65, 4, 16, 7, 31, 1, 21, 254, 2, 5, 43, 1, 12, 1, 63.

## Hipotézisvizsgálat

A matematikai statisztika másik alapfeladata a számunkra érdekes, előzetesen megfogalmazott hipotézis,  $H_0$  ellenőrzése ( $H_0 : \vartheta \in \theta_0$ ; ahol a  $\theta$  paraméterteret két diszjunkt részre bontottuk:  $\theta = \theta_0 \cup \theta_1$ ). Ezt általában egy ún. próbastatisztika alapján tesszük meg, kijelölve azon megfigyelés-értékek halmazát, amelyekre elfogadjuk ( $\chi_e$ ), illetve elutasítjuk ( $\chi_k$ ) a  $H_0$  hipotézist. A célunk az, hogy ezt az alap nullhipotézist csak akkor vessük el, ha arra nyomós indokunk van. Így előre megadott  $\alpha$  (tipikusan 0,05 vagy még kisebb) szintre állíthatjuk be a próba  $\alpha$  terjedelmét, ami az elsőfajú hiba valószínűsége (elsőfajú hibát akkor követünk el, ha elutasítjuk  $H_0$ -t, pedig igaz).

**Paraméteres próbák:** Adott valamilyen konkrét eloszlás, amelynek valamely paraméterére állítunk fel hipotézist.

A legfontosabb ilyen eset, amikor a mintaelemek normális eloszlásúak. A továbbiakban áttekintünk a **normális eloszlás várható értékére vonatkozó egymintás próbákat**:

| u-próba                                                                                                                                                                                                      | t-próba                                                                                                                                                                                                                     |
|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| $X_1, X_2, \dots, X_n \sim N(\mu, \sigma_0^2)$                                                                                                                                                               | $X_1, X_2, \dots, X_n \sim N(\mu, \sigma^2)$                                                                                                                                                                                |
| a) $H_0 : \mu = \mu_0, H_1 : \mu \neq \mu_0$<br>b) $H_0 : \mu = \mu_0, H_1 : \mu < \mu_0$<br>c) $H_0 : \mu = \mu_0, H_1 : \mu > \mu_0$                                                                       | a) $H_0 : \mu = \mu_0, H_1 : \mu \neq \mu_0$<br>b) $H_0 : \mu = \mu_0, H_1 : \mu < \mu_0$<br>c) $H_0 : \mu = \mu_0, H_1 : \mu > \mu_0$                                                                                      |
| $u(\mathbf{x}) = \frac{\bar{X} - \mu_0}{\sigma_0} \sqrt{n} \overset{H_0}{\sim} N(0, 1)$                                                                                                                      | $t(\mathbf{x}) = \frac{\bar{X} - \mu_0}{S_n^*} \sqrt{n} \overset{H_0}{\sim} t(n-1)$                                                                                                                                         |
| kritikus tartományok<br>a) $\chi_k = \{\mathbf{x} :  u(\mathbf{x})  > u_{\alpha/2}\}$<br>b) $\chi_k = \{\mathbf{x} : u(\mathbf{x}) < -u_\alpha\}$<br>c) $\chi_k = \{\mathbf{x} : u(\mathbf{x}) > u_\alpha\}$ | kritikus tartományok<br>a) $\chi_k = \{\mathbf{x} :  t(\mathbf{x})  > t_{\alpha/2}(n-1)\}$<br>b) $\chi_k = \{\mathbf{x} : u(\mathbf{x}) < -t_\alpha(n-1)\}$<br>c) $\chi_k = \{\mathbf{x} : u(\mathbf{x}) > t_\alpha(n-1)\}$ |

IV. Az alábbi minta 5 -egyforma képességűnek feltételezett - sportoló súlylökésben elért eredményeit tartalmazza (tegyük fel, hogy az adatok normális eloszlásból származnak). Az első dobás előtt az edző büszkén állította, hogy tanítványai átlagosan legalább 17 métert dobnak, amit a klub igazgatója kétségbe vont. Úgy döntött, hogy csak akkor hosszabbítja meg az edző szerződését, ha a  $H_0 : m = 17$  hipotézis  $\alpha = 0,05$  elsőfajú hibavalószínűség mellett elfogadható a  $H_1 : m < 17$  alternatívával szemben.

- Hogyan döntött az igazgató, ha a korábbi tapasztalataik alapján a dobások szórást 2-nek tekintették?
- Változott volna-e a helyzet, ha nem tekintik a szórást ismertnek?
- Az előzőek alapján az igazgató végül is még egy esélyt adott az edzőnek. Ő az első kísérlet után mindenkinek elmagyarázta, hogy mire kellene odafigyelni a jobb eredmény érdekében. Segített-e az "edzés"?
- Végül is mi legyen az edző sorsa?

|                  |      |      |      |      |      |  |
|------------------|------|------|------|------|------|--|
| első eredmény    | 14,8 | 12,2 | 16,8 | 17,1 | 16,1 |  |
| második eredmény | 18,8 | 12,1 | 17,2 | 17,7 | 17,0 |  |

**Nemparaméteres próbák** A következőkben arról az esetről lesz szó, amikor nem tudjuk néhány valós paraméterrel megadni az összes szóba jöhető eloszlást. Gyakran maga az eloszlás az ismeretlen, és erre vonatkozik a nullhipotézis. A leggyakrabban használt próba a  $\chi^2$ -próba, amely az  $A_1, A_2, \dots, A_r$  teljes eseményrendszerre vonatkozó  $H_0 : \mathbb{P}(A_i) = p_i, i = 1, \dots, r$  hipotézis ellenőrzésére alkalmazható. A próbastatisztika:

$$\chi^2 = \sum_{i=1}^r \frac{(\nu_i - np_i)^2}{np_i},$$

ahol  $\nu_i$  az  $A_i$  esemény gyakorisága. A  $\chi^2$  statisztika  $r-1$  szabadságfokú  $\chi^2$ -eloszlással közelíthető, innen tudunk kritikus értéket szerezni.

A  $\chi^2$ -próbák alkalmazhatók például: illeszkedésvizsgálatra, homogenitásvizsgálatra, függetlenségvizsgálatra.

**Illeszkedésvizsgálat**  $H_0$  : A minta egy adott eloszlásból származik.

*Diszkrét eset:*

Tfh, hogy az eloszlás értékkészlete véges:  $X = x_1, x_2, \dots, x_r$

$$H_0 : \mathbb{P}(X = x_i) = p_i, i = 1, \dots, r$$

A próbastatisztika a fenti  $\chi^2$ , ahol  $\nu_i$  az  $\{X = x_i\}$  megfigyelések gyakorisága.

$$\chi_k = \{(\nu_1, \dots, \nu_r) : \chi^2 \geq \chi_\varepsilon^2(r-1)\}$$

Ha az értékkészlet megszámlálhatóan végtelen, akkor egyes értékeket (akár végtelen sokat) össze kell vonni.

*Folytonos eset:*

Az értékkészletet fel kell osztani, úgy, hogy minden intervallumba essen legalább 3 megfigyelés. Ekkor  $r$  jelöli az intervallumok számát,  $\nu_i$  az  $i$ -edik intervallumban lévő megfigyelések számát,  $p_i$  pedig  $H_0$  fennállása esetén az  $i$ -edik intervallumba esés elméleti valószínűségét. Ezekkel az értékekkel kell kiszámolni a  $\chi^2$  statisztika értékét, és összehasonlítani az  $r-1$  szabadságfokú  $\chi^2$  eloszlás  $1-\varepsilon$  kvantilisével.

*Becsüléses eset*ben a szabadságfokokat a becsült paraméterek számával csökkenteni kell.

V. Tekintsük az alábbi megfigyelés-sort (az adatokat az egyszerűség kedvéért nagyság szerint rendezve adjuk meg): 0,03, 0,05, 0,08, 0,08, 0,11, 0,12, 0,14, 0,19, 0,23, 0,24, 0,31, 0,43, 0,45, 0,56, 0,61, 0,67, 0,69, 0,78, 0,85, 0,88.  $\alpha = 0,05$  elsőfajú hibavalószínűség mellett döntsünk arról a hipotézisről, hogy a minta a  $[0, 1]$  intervallumon egyenletes eloszlásból származik.

VI. Döntsünk arról a hipotézisről, hogy az alábbi minta -ahol csak azt adjuk meg, hogy az egyes intervallumokba hány elem esik- normális eloszlásból származik. A minta a átlaga 2,2, korrigált tapasztalati szórásnégyzete 2,25.

|               |     |     |     |     |
|---------------|-----|-----|-----|-----|
| intervallumok | 0-1 | 1-2 | 2-3 | 3-4 |
| gyakoriságok  | 4   | 15  | 19  | 12  |

VII. Döntsünk arról a hipotézisről, hogy az alábbi minta Pascal-eloszlásból származik!  
*Mj: A geometriai másik elnevezése a Pascal.*

|              |     |    |   |   |
|--------------|-----|----|---|---|
| értékek      | 1   | 2  | 3 | 4 |
| gyakoriságok | 160 | 32 | 6 | 2 |