5. heti feladatsor, (statisztika. gyak.) 2012. március 28–29.

(1) Legyen $X_1, \ldots, X_n$ független minta $N(m, 1)$ eloszlásból. Adjunk 95%-os, illetve 98%-os megbízhatóságú kétoldali konfidenciaintervallumot m-re, ha
a) $n = 10,\ \sum_{i=1}^{n} X_i = 349$;
b) $n = 250,\ \sum_{i=1}^{n} X_i = 48,5$;
c) $n = 10$, és a minta a következő:

1,49	3,84	3,16	2,09	1,79	1,19	2,38	2,85	3,01	1,74

(2) Valaki azt állítja, hogy a klíma változik, és ezt azzal véli bizonyítottnak, hogy az elmúlt 10 évben 2-szer is volt jégeső, pedig korábban ez egyes évekre a jégeső valószínűsége a hivatalos adatok alapján csupán $0,1$ volt. Írjuk fel a hipotéziseket, a próbát és állapítsuk meg az elsőfajú hiba valószínűségét, valamint az erőfüggvényt a $p = 0,2$ pontban.

(3) Egy tejgyárban minden szállítás előtt megvizsgálják a 25 dkg-os túrók minőségét. Jelölje p annak valószínűségét, hogy egy csomag túró romlott. A $H_0 : p = 0.01$ nullhipotézist (ez még elfogadható a gyár számára) vizsgálják a $H_1 : p > 0.01$ ellenhipotézissel szemben. A gyár eljárása a következő: N csomagot bontanak fel és amennyiben legalább 2 köztük romlott, akkor újrahasznosítást rendelnek el.
a) Írjuk le a kísérletet (mintatér, kritikus tartomány, elfogadási tartomány, statisztikai próba)!
b) Milyen N-re lesz az elsőfajú hiba valószínűsége kisebb 5%-nál?
c) Írjuk fel az erőfüggvényt.

(4) Adott egy 4 elemű független minta: X_1, X_2, X_3, X_4. Két hipotézisünk van:
$H_0 : P(X_i = 2) = 1/4, P(X_i = 3) = 1/2, P(X_i = 10) = 1/4$
$H_1 : P(X_i = 2) = 1/4, P(X_i = 3) = 1/4, P(X_i = 10) = 1/2$
Határozzuk meg a valószínűséghányados-próbát 5 %-os elsőfajú hibavalószínűség mellett.

(5) Az alábbi minta 5 év április 18-án Budapesten mért napi középhőmérsékleteit tartalmazza. Ellenőrizzük a $H_0 : m = 17$ hipotézist $\alpha = 0,05$ elsőfajú hibavalószínűség mellett a $H_1 : m < 17$ alternatívával szemben, úgy, hogy
a) a korábbi tapasztalatok alapján tekintsük az értékek szórását 2-nek.
b) Ne használjunk a szórásra vonatkozóan előzetes információt.

középhőmérséklet (°C):	14,8	12,2	16,8	17,1	16,1

(6) Egy tejgyárban minden szállítás előtt megvizsgálják a 25 dkg-os túrók ólomtartalmát. Egy alkalommal a még megengedett szint %-ában a mérések a következők voltak: 98,5; 101,4; 99,5; 100,9; 99,7. A korábbi tapasztalatok alapján az ellenőr feltételezi, hogy 1 szórásúak. Elfogadható-e a $H_0 : m = 100$ nullhipotézis? Hogyan válasszuk meg a H_1 hipotézist? Mennyi lesz a p-érték (szignifikancia-szint)? Mennyi a próba erőfüggvényének értéke az $m = 102$ pontban? Hány elemű mintára van szükség, ha azt szeretnénk, hogy ez legalább 0,99 legyen?

(7) Az alábbi két minta 10 - egyforma képességűnek feltételezett - sportoló súlylökésben elért eredményeit tartalmazza. A sportolók két ötfős csoportban készültek az edzőtáborban. Edzéstervük ugyanaz volt, csak az étrendjük különbözött. 2 hét felkészülés után értékelték az eredményeket. Tételezzük fel, hogy a minták normális eloszlásból származnak.

1. diéta	14,8	12,2	16,8	17,1	16,1
2. diéta	18,0	12,1	17,2	17,7	17,0

a) Melyik diéta volt jobb, ha a dobások szórását 2-nek tekintjük?
b) Melyik diéta okozott nagyobb változékonyságot az eredményekben?
c) Ha nem ismerjük a szórást, akkor tekinthetjük-e valamelyik diétát jobbnak?

2012. április 19-ig beadható házi feladatok

(8) Egy HÖK-elnök számára az elfogadható, ha a hallgatók legfeljebb 72 százaléka utálja (ez a H_0 hipotézis). Az ennél nagyobb arány esetén (H_1 hipotézis) lemond posztjáról. Minden negyedév végén 10 hallgatót kérdez meg, és akkor mond le, ha a tízből legalább 8 diák utálja. Mekkora a próba terjedelme? Várhatóan hány negyedévet fog tevékenykedni az elnök, ha minden félévben a diákok 65 százaléka utálja?

(9) 8 fiú és 8 lány testmagasságára tekintsük a következő adatokat (centiméterben mérve).

Fiúk:	185	193	185	184	188	176
Lányok:	163	169	163	170	156	160

A saját testmagasságot a megfelelő adatsorhoz hozzátéve $\alpha = 0.01$ elsőfajú hibavalószínűség mellett milyen következtetés vonható le az adatokból: bizonyítható-e, hogy a fiúk magasságának várható értéke több a lányokénál? Mennyi a p-érték?

4. heti feladatsor, (statisztika. gyak.) 2012. március 7–8.

(1) Legyen $Z_1, Z_2, \ldots, Z_n$ független minta a következő diszkrét eloszlásból:

$$P(Z_i = 1) = c,\ P(Z_i = 2) = 3c,\ P(Z_i = 3) = 1 - 4c,\ i = 1, \ldots, n,$$

ahol c az ismeretlen paraméter.
a) Határozzuk meg c maximum likelihood becslését!
b) Határozzuk meg c momentum-módszerrel kapható becslését!
c) Írjunk fel egydimenziós elégséges statisztikát az ismeretlen paraméterre!
d) Adjunk torzítatlan becslést c-re Z_1 segítségével!
e) Az előzőben megadottnál konstruáljunk jobb becslést a Rao–Blackwell–Kolmogorov-tétel segítségével.

(2) Legyen a $Z_1, Z_2, \ldots, Z_n$ független minta $N(2m + 5, 1/d^2)$ eloszlású. Becsüljük meg az ismeretlen paramétereket momentum-módszerrel!

Megoldás.

A momentummódszerben az első két momentumot használjuk. A becslés tehát az az m, d pár lesz, amely mellett az első két momentum megegyezik a tapasztalati eloszlás első két momentumával. Ezek a mintaközép, illetve $\frac{1}{n}\sum_{i=1}^{n} Z_i^2$. Tehát:

$$\mathbb{E}_{m,d}(Z_1) = \frac{Z_1 + \ldots + Z_n}{n}; \qquad \mathbb{E}_{m,d}\left(Z_1^2\right) = \frac{Z_1^2 + \ldots + Z_n^2}{n};$$

$$2m + 5 = \overline{Z}; \qquad \frac{1}{d^2} + (2m + 5)^2 = \frac{Z_1^2 + \ldots + Z_n^2}{n};$$

$$m = \frac{\overline{Z} - 5}{2}; \qquad \frac{1}{d^2} = \frac{Z_1^2 + \ldots + Z_n^2}{n} - \overline{Z}^2 = s_n^2 \Rightarrow d = \frac{1}{s_n}.$$

A maximumlikelihood-módszerhez a likelihoodfüggvény:

$$f(m, d) = \prod_{i=1}^{n}\left[\frac{d}{\sqrt{2\pi}}\exp\left(-\frac{(Z_i - (2m + 5))^2}{2/d^2}\right)\right] =$$

$$= \left(\frac{d}{\sqrt{2\pi}}\right)^n \exp\left(-\frac{d^2}{2}\left(\sum_{i=1}^{n}(Z_i - (2m + 5))^2\right)\right)$$

(A négyzetet kibontva lehet látni, hogy $\left(\sum_{i=1}^{n} Z_i, \sum_{i=1}^{n} Z_i^2\right)$ elégséges statisztika.) A loglikelihood-függvény maximumhelyét keressük, ez megengedhető a logaritmusfüggvény szigorú monoton növekesése miatt.

$$l(m, d) = \log f(m, d) = -\frac{n}{2}\log(2\pi) + n\log d - \frac{d^2}{2}\left(\sum_{i=1}^{n}(Z_i - (2m + 5))^2\right).$$

Rögzített d mellett ez akkor maximális, ha $\sum_{i=1}^{n}(Z_i - (2m + 5))^2$ minimális. Ismert, vagy deriválással könnyen ellenőrizhető, hogy ez $2m + 5 = \overline{Z}$ esetén érhető el. Ilyenkor a négyzetösszeg éppen a tapasztalati szórásnégyzet n-szerese. Vagyis $\hat{m} = \frac{\overline{Z} - 5}{2}$, és a továbbiakban

$$l(\hat{m}, d) = -\frac{n}{2}\log(2\pi) + n\log d - \frac{d^2}{2}n \cdot s_n^2$$

maximumhelyét keressük d-ben. d szerint deriválva:

$$\frac{\partial}{\partial d} l(\hat{m}, d) = \frac{n}{d} - d \cdot n \cdot s_n^2.$$

Ez pontosan akkor nulla, ha $d = \frac{1}{s_n}$. Másrészt l folytonosan differenciálható, $\lim_{d \to 0+} l(\hat{m}, d) = \lim_{d \to \infty} l(\hat{m}, d) = -\infty$. Ezekből következik, hogy $d = \frac{1}{s_n}$ a maximumhely.

Tehát a maximumlikelihood-becslés: $\left(\hat{m}, \hat{d}\right) = \left(\frac{\overline{Z} - 5}{2}, \frac{1}{s_n}\right)$.

(3) $X_1, X_2, \ldots, X_n$ független minta $b > 0$ paraméterű Poisson-eloszlásból. Mi a b paraméter maximum likelihood becslése?

Megoldás.

A Poisson-eloszlás diszkrét eloszlás, méghozzá $P(\xi = k) = \frac{b^k}{k!}e^{-b}$ $(k = 0, 1 \ldots)$, ha ξ b paraméterű Poisson-eloszlású valószínűségi változó. Ha ξ független a mintától, akkor segítségével a likelihood-függvény így írható:

$$f_b(X_1, \ldots, X_n) = \prod_{i=1}^{n} P_b(\xi = X_i) = \prod_{i=1}^{n}\left[\frac{b^{X_i}}{X_i!}e^{-b}\right] =$$

$$= \left(\prod_{i=1}^{n}\frac{1}{X_i!}\right) b^{\sum_{i=1}^{n} X_i} e^{-nb}.$$

Észrevehetjük, hogy a mintaelemek összege elégséges statisztika. Azt a b-t keressük, melyre a likelihood-függvény maximális. A likelihood-függvény most mindenképpen pozitív, és a logaritmusfüggvény szigorúan monoton növő, így elegendő a log-likelihood függvény maximumhelyét megkeresni:

$$l(b) = \log f_b(X_1, \ldots, X_n) = \log\left(\prod_{i=1}^{n}\frac{1}{X_i!}\right) + \sum_{i=1}^{n} X_i \log b - nb.$$

Ennek b szerinti deriváltja: $\partial l(b) = \sum_{i=1}^n X_i \cdot \frac{1}{b} - n$. Könnyen látható, hogy ez pontosan akkor nulla, ha $b = \sum_{i=1}^n X_i \cdot \frac{1}{n} = \overline{X}$. Sőt $b < \overline{X}$ esetén a derivált pozitív, míg $b > \overline{X}$ esetén negatív, így a log-likelihood függvénynek valóban maximuma van a $b = \overline{X}$ helyen. Tehát a $\hat{b} = \overline{X}$ becslés a paraméter maximum likelihood becslése.

(4) Határozzuk meg az ismeretlen paraméter maximum likelihood becslését, ha a minta
a) Pascal-eloszlású;
b) binomiális eloszlású (a rend ismert);
c) $(a, 1)$ intervallumon egyenletes eloszlású.

2012. március 22-ig beadható házi feladatok

(5) Egy ezerfős közvélemény-kutatásnál 650-en válaszolták azt, hogy szeretik a kutyákat. Becsüljük meg maximumlikelihood– és momentummódszerrel a tényleges arányt! Hogyan módosul ez a becslés, ha már tudjuk, hogy a kutyákat szeretők 10%-a hazudik, a kutyákat nem szeretőknél pedig 20% ez az arány?

(6) Legyen $X_1, X_2, \ldots, X_n$ Pascal-eloszlású minta.
a) Adjunk X_1 függvényeként torzítatlan becslést $p(1-p)$-re!
b) Az a)-ban megadottnál konstruáljunk jobb becslést a Rao–Blackwell–Kolmogorov-tétel segítségével!

3. heti feladatsor, (statisztika. gyak.) 2012. február 29–március 1.

(1) Legyen $X_1, X_2, \ldots, X_n$ független minta valamely véges szórású eloszlásból, és tekintsük az $T(X_1, X_2, \ldots, X_n) = a_1X_1 + a_2X_2 + \ldots + a_nX_n$ alakú lineáris becsléseket, ahol $a_1, \ldots, a_n$ valós számok. Feltéve, hogy a becslésünk torzítatlan a várható értékre, mely $a_1, \ldots, a_n$ együtthatókra lesz minimális a $T(X_1, X_2, \ldots, X_n)$ mennyiség szórásnégyzete?

Megoldás. $E_\vartheta(a_1X_1 + \ldots + a_nX_n) = (a_1 + \ldots + a_n)E_\vartheta(X_1)$, így a becslés pontosan akkor torzítatlan, ha az együtthatók összege 1.

A függetlenség miatt $D^2_\vartheta(a_1X_1 + \ldots + a_nX_n) = (a_1^2 + \ldots + a_n^2)D^2_\vartheta(X_1)$, így a szórásnégyzet akkor minimális, ha $a_1^2 + \ldots + a_n^2$ minimális. Nemnegatív számokra a számtani és négyzetes közepek közti egyenlőtlenség szerint ez akkor áll fenn, ha $a_1 = \ldots = a_n = 1/n$, azaz T a mintaátlag. Ennél kisebb szórásnégyzet negatív együtthatókkal sem érhető el: $a > 0$ esetén $a^2 + b^2 > (b-a)^2$.

(2) Torzítatlan, illetve konzisztens becslés-e a mintaátlag reciproka geometriai eloszlás paraméterére?

Megoldás. Nem torzítatlan: $E_p(\frac{n}{X_1+\ldots+X_n}) > \frac{1}{E_p(\overline{X})} = p$. Konzisztens: a nagy számok gyenge törvénye szerint $(X_1 + \ldots + X_n)/n$ sztochasztikusan konvegál $E_p(X_1) = 1/p$-hez, így az átlag reciproka sztochasztikusan konvergál ennek reciprokához, p-hez $n \to \infty$ esetén minden p-re.

(3) Adjunk meg torzítatlan becslést a $[0, \vartheta]$ intervallumon egyenletes eloszlás paraméterére a) mintaátlag; b) a maximum; c) a minimum segítségével. Számítsuk ki a becslések szórását is. Melyik hatásosabb? Melyik konzisztens?

Megoldás. $2\overline{X}$, $\frac{n+1}{n}X_n^*$, $(n+1)X_n^*$. Szórásnégyzetek: $\vartheta^2/3n$, $\frac{\vartheta^2}{n(n+2)}$, $\frac{\vartheta^2 n}{(n+2)}$.

(4) A paraméter mely függvényeit lehet torzítatlanul becsülni a) indikátor b) geometriai c) Poisson-eloszlásból származó minta esetén a mintaátlag és a tapasztalati szórásnégyzet segítségével?

Igaz-e, hogy a Poisson-eloszlás esetén $\overline{X}^2$ torzítatlan becslése a paraméter négyzetének?

(5) Legyen $X_1, \ldots, X_n$ független, b paraméterű Poisson-eloszlású minta ($b > 0$). Számítsuk ki a Fisher-információt! Adjunk elégséges statisztikát! Mi történik $a > 0$ paraméterű exponenciális minta esetén?

Megoldás. a) Ha $X_1, \ldots, X_n$ független, b paraméterű Poisson-eloszlású minta, akkor

$$\mathbb{P}(X_1 = k_1, \ldots, X_n = k_n) = \prod_{i=1}^n \left(\frac{b^{k_i}}{k_i!}e^{-b}\right) = \\ = \frac{b^{\sum_{i=1}^n k_i}}{\prod_{i=1}^n k_i!}e^{-nb} = \frac{1}{\prod_{i=1}^n k_i!}b^{\sum_{i=1}^n k_i}e^{-nb}.$$

A faktorizációs tétel alapján látható, hogy a mintaelemek összege elégséges statisztika.

Tehát a likelihood, majd a loglikelihood-függvény:

$$f(b) = \frac{b^{\sum_{i=1}^n X_i}}{\prod_{i=1}^n X_i!}e^{-nb};$$

$$l(b) = \log f(b) = \left(\sum_{i=1}^n X_i\right)\log b - \log\left(\prod_{i=1}^n X_i!\right) - nb.$$

Ennek b szerinti deriváltja:

$$\frac{\partial}{\partial b}l(b) = \frac{\sum_{i=1}^n X_i}{b} - n.$$

Ha X_i-k független, Poisson(b)-eloszlásúak, akkor az összegük is Poisson-eloszlású, és paramétere nb. A Poisson-eloszlás várható értéke és a szórásnégyzete is a saját paramétere. Tehát az összeg várható értéke és szórásnégyzete is nb. Ez alapján a Fisher-információ:

$$I(b) = E\left(\frac{\partial}{\partial b}l(b)\right)^2 = D^2\left(\frac{\partial}{\partial b}l(b)\right) + \left[E\left(\frac{\partial}{\partial b}l(b)\right)\right]^2 = \\ = D^2\left(\frac{\sum_{i=1}^n X_i}{b} - n\right) + \left[E\left(\frac{\sum_{i=1}^n X_i}{b} - n\right)\right]^2 = \\ = \frac{1}{b^2}D^2\left(\sum_{i=1}^n X_i\right) + \left[\frac{1}{b}E\left(\sum_{i=1}^n X_i\right) - n\right]^2 = \\ = \frac{nb}{b^2} + \left(\frac{1}{b}nb - n\right)^2 = \frac{n}{b}.$$

Ez tehát a minta Fisher-információja. Vegyük észre, hogy az n elemű minta Fisher-információja n-szerese az egyelemű mintáénak ($n = 1$ helyettesítéssel).

b) Ha $X_1, \ldots, X_n$ független, a paraméterű exponenciális eloszlású minta, akkor X_i sűrűségfüggénye a t_i helyen $a\exp(-at_i)$, ha $t_i > 0$, és 0 különben.

Exponenciális eloszlásból származó mintánál az $X_1, \ldots, X_n$ mintaelemek pozitívak. Tehát a likelihood, majd a loglikelihood-függvény:

$$f(\underline{X}, a) = \prod_{i=1}^n f_a(X_i) = \prod_{i=1}^n a\exp(-aX_i) = a^n\exp(-a\sum_{i=1}^n X_i);$$

A faktorizációs tétel alapján látható, hogy a mintaelemek összege elégséges statisztika.

$$l(a) = \log f(\underline{X}, a) = n\log a - a\left(\sum_{i=1}^n X_i\right).$$

Ennek a szerinti deriváltja:

$$\frac{\partial}{\partial a}l(a) = \frac{n}{a} - \sum_{i=1}^n X_i.$$

Az exponenciális eloszlás várható értéke $\frac{1}{a}$, szórásnégyzete $\frac{1}{a^2}$. Ezért ha X_i-k független, $\exp(a)$-eloszlásúak, akkor az összegük várható értéke $\frac{n}{a}$, szórásnégyzete $\frac{n}{a^2}$. Konstans n/a hozzáadásával a szórásnégyzet nem változik. Ezek alapján a Fisher-információ:

$$I(a) = E\left(\frac{\partial}{\partial a}l(a)\right)^2 = D^2\left(\frac{\partial}{\partial a}l(a)\right) + \left[E\left(\frac{\partial}{\partial a}l(a)\right)\right]^2 = \\ = D^2\left(\sum_{i=1}^n X_i\right) + 0^2 = \frac{n}{a^2}.$$

Ez tehát a minta Fisher-információja. Vegyük észre, hogy az n elemű minta Fisher-információja n-szerese az egyelemű mintáénak ($n = 1$ helyettesítéssel).

2012. március 14-ig beadható házi feladatok

(6) Legyen $X_1, \ldots, X_n$ független minta $Bin(4, p)$-ből, továbbá $Y_1, \ldots, Y_n$ független minta $Bin(6, p)$-ből, és tegyük fel, hogy a két minta egymástól is független. Milyen a és b számokra lesz $a\overline{X} + b\overline{Y}$ a p paraméter torzítatlan becslése? Ezen számpárok közül melyikre lesz a becslés szórása minimális?

(7) Mutassuk meg, hogy exponenciális eloszlású minta esetén
a) $n \cdot \min(X_1, \ldots, X_n)$ torzítatlan, de nem konzisztens a várható értékre.
b) $1/\overline{X}$ nem torzítatlan, de konzisztens a paraméterre.

2. heti feladatsor, (statisztika. gyak.) 2012. február 22-23.

(1) Tegyük fel, hogy február 17-e középhőmérséklete Budapesten az elmúlt 10 évben így alakult (Celsius-fokban):
2,0 8,5 1,6 -4,5 3,3 7,9 0,2 -1,6 -2,2 5,6
Számítsuk ki és ábrázoljuk a középhőmérséklet sűrűségfüggvényének Parzen–Rosenblatt-becslését, ha $h = 0,25$ és a magfüggvény:

$$k(x) = \begin{cases} 1/2, & \text{ha } |x| < 1 \\ 0 & \text{különben.} \end{cases}$$

Megoldás. $\hat{f}(x) = \frac{1}{nh}\sum_{i=1}^{n} k\left(\frac{x-x_i}{h}\right) = \frac{1}{5}|\{: i|x_i - x| < 0,25\}|$.

(2) Mi a mintatér az alábbi esetekben:
a) Három kockadobás.
b) Öten addig lottóznak hetente egy-egy szelvénnyel, míg legalább két találatuk nem lesz. A megfigyelés, hogy kinek hányadik héten sikerül ez először.
c) Húsz napon át feljegyezzük egy diák felkelésének időpontját.
Megoldás. a) $\{1, 2, 3, 4, 5, 6\}^3$; b) $\{1, 2, 3, \ldots\}^5$; c) $[0, 24)^2 0$.

(3) Adjunk torzítatlan becslést a valószínűségszámítás vizsga bukási arányára, ha 300-ból 100-an buktak meg. Adjunk a becslés szórására felső becslést.
Megoldás. A bukások relatív gyakorisága torzítatlan: $1/3$.

(4) n elemű $a > 0$ paraméterű exponenciális minta esetén adjunk torzítatlan becslést $1/a$-ra és $\exp(-3a)$-ra.
Megoldás. X_1 torzítatlan $1/a$-ra: $E_a(X_1) = 1/a$. $I(X_1 > 3)$ torzítatlan $\exp(-3a)$-re: $E_a(I(X_1 > 3)) = P_a(X > 3) = \exp(-3a)$.

(5) n elemű $b > 0$ paraméterű Poisson-eloszlású minta esetén adjunk torzítatlan becslést b-re, b^2-re és $\exp(-b)$-re.
Megoldás. $X_1, X_1^2 - X_1, I(X_1 = 0)$, mert $E_b(X_1) = b$, $E_b(X_1^2) = D_b^2(X_1) + [E_b(X_1)]^2 = b + b^2$, és $Pb(X_1 = 0) = \exp(-b)$.

1. heti feladatsor, (statisztika. gyak.) 2012. február 15.

(1) Legyen X standard normális eloszlású valószínűségi változó, mennyi annak valószínűsége, hogy
$2X + 3 < 2$?
Megoldás. $P(2X + 3 < 2) = P(2X < -1) = P(X < -1/2) = \Phi(-1/2) = 1 - \Phi(1/2) \approx 1 - 0,6915 = 0,3085$.

(2) n elemű független mintára határozzuk meg a tapasztalati közép és a tapasztalati szórásnégyzet várható értékét!
Megoldás. $E(X_1)$, $\frac{n-1}{n}D^2(X_1)$

(3) 16 darab független, 10 várható értékű és 25 szórásnégyzetű normális eloszlás átlaga milyen valószínűséggel nagyobb 9-nél?
Megoldás. $\overline{X} \sim N(10, \frac{25}{16})$; $P(\overline{X} > 9) = 1 - \Phi(-\frac{4}{5}) = \Phi(4/5) \approx 0,7881$.

(4) Tekintsünk n elemű exponenciális mintát, azaz legyenek $X_1, X_2, \ldots, X_n$ független, azonos λ paraméterű exponenciális eloszlású valószínűségi változók. Határozzuk meg a legkisebb mintaelem eloszlását!
Megoldás. $\exp(\lambda n)$

(5) Határozzuk meg a mintaközepet, tapasztalati szórásnégyzetet és szórást, a harmadik tapasztalati momentumot, a szórási együtthatót és a tapasztalati eloszlásfüggvényt az alábbi mintákon:

a) 1, 3, 0, 1;

b) négy kockadobás;

c) fogyasztói árindex:

1993	1994	1995	1996	1997	1998	1999	2000
122,5	118,8	128,2	123,6	118,3	114,3	110,0	109,8

2001	2002	2003	2004	2005	2006	2007	2008
109,2	105,3	104,7	106,8	103,6	103,9	108,0	106,1

Megoldás. a) $\overline{X} = 1,25, s_n = 1,0897, s_n^2 = 1,1875, \hat{X}^3 = 7,25$
c) $\overline{X} = 112,068, s_n = 7,6127, s_n^2 = 57,95, \hat{X}^3 = 1427320$

(6) Dobjunk fel háromszor egy dobókockát.

a) Határozzuk meg az alapstatisztikákat: mintaközép, korrigált tapasztalati szórásnégyzet és szórás, kvartilisek, rendezett minta.

b) Határozzuk meg az F_n tapasztalati eloszlásfüggvényt, és a $\sup_x |F(x) - F_n(x)|$ statisztika értékét, ahol F az $\{1, 2, 3, 4, 5, 6\}$ pontokon egyenletes eloszlás eloszlásfüggvénye.