A március 19-i előadáshoz kapcsolódó gyakorlat feladatai

Nem kötelező házi feladat.

Legyen

$$\begin{array}{ll} \xi_{1,1}\dots, & \xi_{1,n_1} \\ \vdots & \vdots \\ \xi_{k,1}\dots, & \xi_{k,n_k} \\ \vdots & \vdots \end{array}$$

szériasorozat, azaz tegyük fel, hogy az egy sorban álló valószínűségi változók függetlenek. Tegyük fel továbbá, hogy e valószínűségi változók teljesítik a következő feltételeket:

1.) A $\xi_{k,j}$ valószínűségi változók nem negatív egész értékeket vesznek fel.

2.) $P(\xi_{k,j}=1)=\lambda_{k,j}$, $\lim\limits_{k\to\infty}\sum\limits_{j=1}^{n_k}\lambda_{k,j}=\lambda>0$.

3.) $\sup\limits_{1\le j\le n_k}\lambda_{k,j}\to 0$, ha $k\to\infty$, és $\sum\limits_{j=1}^{n_k}P\left(\xi_{k,j}\ge 2\right)\to 0$, ha $k\to\infty$.

Ekkor az $S_k=\sum\limits_{j=1}^{n_k}\xi_{k,j}$ valószínűségi változók eloszlásban konvergálnak a λ paraméterű Poisson eloszláshoz, ha $k\to\infty$, azaz $\lim\limits_{k\to\infty}P(S_k=l)=\dfrac{\lambda^l}{l!}e^{-\lambda}$ minden $l=0,1,2,\dots$ számra.

Segítség: Az eredmény egyszerűbb, az órán tárgyalt változatának a bizonyítása megfelelő változtatással alkalmazható. Adjunk jó aszimptotikus becslést a $\xi_{k,j}$ valószínűségi változók $g_{k,j}(x)=\sum\limits_{l=0}^{\infty}P(\xi_{k,j}=l)x^l$ generátorfüggvényére, $|x|<1$.

1. Egy tóban 3000 hal van. Véletlenül kihalásznak belőle 1000 darabot, és ezekre piros pöttyöt festenek és visszaengedik őket. Ezután ismét kifognak véletlenül 1000 halat. Mi annak a valószínűsége, hogy a kifogott halak között 100 megfestett van?

   Megoldás: $\dfrac{\binom{1000}{100}\binom{2000}{900}}{\binom{3000}{1000}}$, mert a keresett eloszlás hipergeometrikus. Azt kell kiszámítani, hogy ha 3000 objektumból (halból) két tipus van, az egyikből (a megjelölt halból) 1000, a másikból 2000, és kiválasztunk (visszatevés nélkül) 1000 objektumot, akkor mi a valószínűsége annak, hogy 100-at választunk az első és 900-at a második objektumból.

   Megjegyzés: A gyakorlatban előforduló kérdés ennek fordítottja. Elvégezzük a fenti kisérletet, összeszámláljuk a megfestett halakat, és ebből próbálunk következtetni a tóban levő halak számára. Hogyan érdemes ezt csinálni?

Ezt a problémát érdemes részletesebben megtárgyalni. Valójában, nem tudjuk, hogy hány hal van a tóban. De ennek megbecslése érdekében a következő eljárást alkalmazhatjuk:

Végezzünk két fogást, az első fogásban kifogott halakat jelöljuk meg. Ezután meg akarjuk állapítani mennyi hal lehet összesen a tóban. Ezt természetesen csak bizonyos (véletlentől függő) pontossággal tudjuk meghatározni. Az ilyen tipusú feladatok tipikusak a matematikai statisztikában, az ilyen problémák vizsgálatát nevezik becsléseleméletnek. Világos, hogy az, hogy 1000-nél alig több hal van, nem túl valószínű, mert akkor sokkal több megjelölt hal lenne a második fogásban. Az hogy rengeteg, mondjuk 1 000 000 hal lenne a tóban szintén nem túl valószínű, mert akkor sokkal kevesebb megjelölt hal lenne a második fogásban. A matematikai statisztikában kidolgoztak egy általános elvet a maximum likelihood módszernek nevezett eljárást, mely nagyon általános feltételek mellett nagyon jó módszert ad, és mely jelen esetben is alkalmazható. Tárgyaljuk meg ezt a módszert a jelen esetben. Tekintsünk kissé általánosabb esetet. Vezessük be a következő jelöléseket:

x a tóban lévő halak (ismeretlen) száma,
n az első fogásban kifogott (és megjelölt) halak száma,
r a második fogásban kifogott halak száma,
k a második fogásban kifogott előzöleg megjelölt halak száma. Annak valószínűsége, hogy adott (ismeretlen) x és n, r számok esetén pontosan k megjelölt halat fogunk ki

$$q_k(x,n,r)=\frac{\binom{n}{k}\binom{x-n}{r-k}}{\binom{x}{r}}.$$

Tekintsük az ismeretlen x szám (maximum likelihood) becslésének azt az x számot, melyre a $q_k(x,n,r)$ mennyiség (rögzített n, k és r számok mellett) maximális.

2. Határozzuk meg a fenti feladatban a maximum likelihood becslést.
Megoldás: Némi számolás mutatja, hogy

$$\frac{q_k(x,n,r)}{q_k(x-1,n,r)}=\frac{x-n}{x-n-r+k}\cdot\frac{x-r}{x}=\frac{x^2-rx-nx+rn}{x^2-rx-nx+kx}.$$

Ez a tört kisebb mint egy, ha $rn<kx$, nagyobb mint egy, ha $rn>kx$. Ezért a becslés $rn=kx$, azaz $x=\frac{rn}{k}$, pontosabban az e számot közrefogó egész számok valamelyike. Valóban $x<\frac{rn}{k}$ esetében a $q_k(x,n,r)$ függvény (mint az x változó függvénye rögzített k, n és r paraméterekkel) monoton nő, $x>\frac{rn}{k}$ esetében pedig a $q_k(x,n,r)$ függvény monoton csökken.

3. Feldobunk egy pénzdarabot, mely p valószínűséggel esik a fej, $1-p$ valószínűséggel az írás oldalra kétszer egymástól függetlenül. Jelölje ξ azt a valószínűségi változót, mely a fejdobások számát adja meg. Adjuk meg a ξ valószínűségi változó $F(x)$ eloszlásfüggvényét.

Megoldás: A vizsgált $\xi(\omega)$ valószínűségi változó olyan, hogy $P(\xi=0)=(1-p)^2$, $P(\xi=1)=2p(1-p)$, $P(\xi=2)=p^2$. Innen $F(x)=0$, ha $x\le 0$, $F(x)=(1-p)^2$, ha $0<x\le 1$, $F(x)=(1-p)^2+2p(1-p)=1-p^2$, ha $1<x\le 2$, és $F(x)=1$, ha $2<x<\infty$.

4. Legyen adva egy $\xi(\omega)$ valószínűségi változó $F(x)$ eloszlásfüggvénye. Határozzuk meg ennek segítségével az $\{\omega\colon a<\xi(\omega)<b\}$ alakú események, $-\infty<a<b<\infty$, valószínűségét, valamint annak valószínűségét, hogy $\xi(\omega)$ valamilyen páros egész értéket vesz fel.
Megoldás:
$$\begin{aligned} P\left(\{\omega\colon a<\xi(\omega)<b\}\right) &= \lim_{n\to\infty} P\left(\left\{\omega\colon a+\frac{1}{n}\ge \xi(\omega)<b\right\}\right)\\ &= \lim_{n\to\infty}\left[F(b)-F\left(a+\frac{1}{n}\right)\right]. \end{aligned}$$
Egy tetszőleges u pontra $P(\xi=u)=\lim\limits_{n\to\infty}\left[F\left(u+\frac{1}{n}\right)-F(u)\right]$. Ezért annak valószínűsége, hogy a ξ valószínűségi változó valamely páros egész értéket vesz fel $\sum\limits_{k=-\infty}^{\infty}\lim\limits_{n\to\infty}\left[F\left(2k+\frac{1}{n}\right)-F(2k)\right]$.

5. Lássuk be, hogy létezik olyan diszkrét eloszlású ξ valószínűségi változó, melynek $F(x)=P(\xi<x)$ eloszlásfüggvényére igaz, hogy minden $-\infty<a<b<\infty$ számpárra $F(a)<F(b)$ szigorú egyenlőtlenséggel.
Megoldás: Egy lehetséges konstrukció a következő. Legyen $r_1,r_2,\dots,$ a racionális számok sorozata valamilyen sorrendben felsorolva. Definiáljunk olyan ξ valószínűségi változót, melyre $P(\xi=r_j)=2^{-j}$, $j=1,2,\dots$. Ekkor ξ diszkrét eloszlású valószínűségi változó, és ennek $F(x)$ eloszlására $F(b)-F(a)=P(a<\xi<b)>0$, mert $P(a<\xi<b)>P(\xi=r_j)>0$, ahol r_j egy az $[a,b]$ intervallum belsejében lévő racionális szám.

6. Ha ξ és η két független Poisson eloszlású valószínűségi változó $\lambda>0$ és $\mu>0$ paraméterekkel, azaz $P(\xi=k)=\frac{\lambda^k}{k!}e^{-\lambda}$, $P(\eta=k)=\frac{\mu^k}{k!}e^{-\mu}$, $k=0,1,2,\dots,$ akkor $\xi+\eta$ Poisson eloszlású valószínűségi változó $\lambda+\mu$ paraméterrel.
Megoldás:
$$\begin{aligned} P(\xi+\eta=k) &= \sum_{j=0}^{k} P(\xi=j,\eta=k-j)=\sum_{j=0}^{k}P(\xi=j)P(\eta=k-j)\\ &= \sum_{j=0}^{k}\frac{\lambda^j}{j!}e^{-\lambda}\frac{\mu^{k-j}}{(k-j)!}e^{-\mu}=\frac{e^{-(\lambda+\mu)}}{k!}\sum_{j=0}^{k}\frac{k!}{j!(k-j)!}\lambda^j\mu^{k-j}\\ &= \frac{e^{-(\lambda+\mu)}}{k!}(\lambda+\mu)^k. \end{aligned}$$

7. Legyen adva k darab urna, és ezekbe dobjunk be véletlen ξ számú golyót, ahol ξ Poisson eloszlású valószínűségi változó $\lambda>0$ paraméterrel. Legyenek az egyes

dobások eredményei egymástól és a ξ valószínűségi változótól függetlenek. Tegyük fel továbbá, hogy minden egyes dobásnál a golyó az j-ik urnába $p_j \geq 0$ valószínűséggel esik, $j = 1, \ldots, k$, $\sum_{j=1}^{k} p_j = 1$. Jelölje η_j a j-ik urnába eső golyók számát. Ekkor az η_j, $j = 1, \ldots, k$ valószínűségi változók függetlenek, és η_j Poisson eloszlású λp_j paraméterrel, $j = 1, \ldots, k$.
Megoldás:

$$
\begin{aligned}
P(\eta_1 = l_1, \ldots, \eta_k = l_k) &= P(\xi = l_1 + \cdots + l_k) \frac{(l_1 + \cdots + l_k)!}{l_1! \cdots l_k!} p_1^{l_1} \cdots p_k^{l_k} \\
&= \frac{\lambda^{(l_1 + \cdots + l_k)}}{l_1! \cdots l_k!} p_1^{l_1} \cdots p_k^{l_k} e^{-\lambda} \\
&= \frac{\lambda^{l_1} \cdots \lambda^{l_k}}{l_1! \cdots l_k!} p_1^{l_1} \cdots p_k^{l_k} e^{-\lambda(p_1 + \cdots + p_k)} = \prod_{j=1}^{k} \frac{(\lambda p_j)^{l_j}}{l_j!} e^{-\lambda p_j}
\end{aligned}
$$

tetszőleges $l_1 \geq 0$, ... , $l_k \geq 0$ egész számokra. Innen adódik az állítás.

8. Egy pontot ledobunk az egységintervallumba úgy, hogy annak valószínűsége, hogy ez a pont valamely $[a, b] \subset [-1, 1]$ intervallumba esik $\frac{b-a}{2}$. Legyen ξ az a valószínűségi változó, amelyik megadja, hogy hova esett ez a pont. Adjuk meg a ξ^2 valószínűségi változó eloszás és sűrűségfüggvényét.
Megoldás: $P(\xi^2 < x) = 0$, ha ≤ 0, $P(\xi^2 < x) = P(-\sqrt{x} < \xi < \sqrt{x})$, ha ≥ 0, ahonnan $P(\xi^2 < x) = \frac{\sqrt{x} - (-\sqrt{x})}{2} = \sqrt{x}$, ha $0 \leq x \leq 1$, $P(\xi^2 < x) = 1$, ha $x \geq 1$. Innen ξ^2 eloszlásfüggvénye $F(x) = 0$, ha $x \leq 0$, $F(x) = \sqrt{x}$, ha $0 \leq x \leq 1$, $F(x) = 1$, ha $x \geq 1$, $F(x) = 1$, ha $x \geq 1$. A ξ^2 valószínűségi változó sűrűségfüggvénye pedig $f(x) = \frac{dF(x)}{dx}$, $f(x) = 0$, ha $x \leq 0$ vagy $x \geq 1$, $f(x) = \frac{1}{2} x^{-1/2}$, ha $0 \leq x \leq 1$.