NÉV: .... NEPTUN-KÓD: .... SZAK: ....

ELŐADÓ: Pete Gábor GYAKVEZ.:.................................

Valószínűségszámítás Vizsga3, 2014. jan. 17.

Munkaidő: 100 perc. Kalkulátor nem használható.

Elm. 1. Legyen U egyenletes eloszlású a $[0, \pi]$ intervallumon, és $X = \sin(U)$.

(a) Írjuk le X eloszlásfüggvényét! **(5p)**

(b) Számoljuk ki X várható értékét, **(5p)**

(c) és szórásnégyzetét! **(5p)**

Megoldás: (a)

$$F_X(x) = \mathbb{P}[\sin U \le x] = \mathbb{P}[|U| \le \arcsin x] = \frac{2 \arcsin x}{\pi}, \text{ ahol } x \in [0, 1] .$$

A 2-es szorzót senki sem vette észre, pedig egy $[0, \pi]$ félkör ábra azonnal elárulta volna. Az meg sajnos kritikán aluli, hogy a többség szerint az $\arcsin(x)$ értelmezési tartománya $x \in [0, \pi]$.

Sokan a sűrűségfgvre nyomultak, nem tudom, pontosan miért. De azért leírom. $g(u) = \sin u$, $u \in [0, \pi]$, ez nem injektív, tehát vigyázni kell a szokásos formulával, nem létezik $g^{-1}(x)$:

$$f_X(x) = \sum_{u:g(u)=x} f_U(u)|g'(u)|^{-1} = \frac{1}{\pi} \sum_{u:g(u)=x} \frac{1}{|\cos u|} = \frac{2}{\pi} \frac{1}{\sqrt{1-x^2}} .$$

Utána ezt kéne integrálni, hogy az eloszlásfgvt megkapjuk.

(b) Én arra gondoltam, hogy

$$\mathbb{E}X = \mathbb{E}[\sin U] = \int_0^\pi \frac{1}{\pi} \sin u \, du = \frac{1}{\pi}[-\cos u]_{u=0}^{\pi} = \frac{2}{\pi} .$$

Ehelyett lehet $\int_0^1 \frac{2}{\pi} x \frac{1}{\sqrt{1-x^2}} \, dx$-szel vesződni, de az nehezebb.

(c) Persze $\mathbb{D}^2 X = \mathbb{E}[X^2] - (\mathbb{E}X)^2$, majd én azt csináltam volna, hogy:

$$\mathbb{E}[X^2] = \mathbb{E}[\sin^2 U] = \int_0^\pi \frac{1}{\pi} \sin^2 u \, du = \frac{1}{2}\frac{1}{\pi} \left(\int_0^\pi \sin^2 u \, du + \int_0^\pi \cos^2 u \, du \right) = \frac{1}{2} .$$

Elm. 2. (a) Mondd ki a Cauchy-Schwarz egyenlőtlenséget az $\mathbb{E}[XY]$ belsőszorzásra. Igazold a segítségével, hogy két változó $\rho(X, Y)$ korrelációs együtthatója mindig -1 és 1 között van. **(7p)**

(b) Adj egy példát, hogy $\rho(X, Y) = 0$, de X és Y nem függetlenek. **(5p)**

Megoldás: (a) Lásd jegyzet.

(b) Lásd pld jegyzet, de annyi csak a trükk, hogy (1) $\mathbb{E}X = 0$ áll bármilyen szimmetrikusra; (2) $\mathbb{E}[XY] = 0$ teljesül, ha $X \neq 0 \implies Y = 0$, vagy akkor is teljesül, ha XY szimmetrikus; (3) ne legyenek függetlenek.

Elm. 3. (a) Ha X és Y független valváltozók, hogyan kaphatjuk meg az $M_{X+Y}(t)$ momentumgeneráló függvényt az $M_X(t)$ és $M_Y(t)$ momentumgenerálók segítségével? (Bizonyítsd is.) **(5p)**

(b) Írd föl az $\mathsf{N}(\mu, \sigma^2)$ normális változó momentumgeneráló függvényét (bizonyítással). **(7p)**

(c) Az (a) és (b) részek segítségével igazold, hogy ha $X \sim \mathsf{N}(\mu_X, \sigma_X^2)$ és $Y \sim \mathsf{N}(\mu_Y, \sigma_Y^2)$ függetlenek, akkor $X + Y$ is normális. Mekkora a várható értéke és a szórásnégyzete? **(5p)**

Megoldás: Lásd jegyzet. A lényeg: momgenfgv definíciója, fgtlenség használata, normális sűrűségfgve, kitevő teljes négyzetté alakítása (ami standard normálisra könnyű, aztán változótrafó).

Gyak. 1. A Jó Kenyér pékség kalácsában szeletenként átlag 3, a Príma Pék kalácsában szeletenként átlag 2 mazsola van. Nagyjából minden másnap bemegyek valamelyik pékségbe, 1/2 valószínűséggel az egyikbe, 1/2-del a másikba, és veszek egy 16 szeletből álló egész kalácsot.

(a) Mi a valószínűsége, hogy van mazsola az első szelet kalácsomban? **(4 p)**

(b) Volt benne mazsola. Mi a valószínűsége, hogy a Jó Kenyér pékségtől van a kalács? (**5p)**

(c) Ha az első szeletben volt mazsola, mi a valószínűsége, hogy a másodikban legalább kettő is lesz? **(7p)**

Megoldás: (a) $(1-e^{-2})/2+(1-e^{-3})/2$.

(b) Bayes tétel: $\frac{(1-e^{-3})/2}{(1-e^{-2})/2+(1-e^{-3})/2}$.
Volt, aki úgy értette a „volt benne mazsola” mondatot, hogy az nem az előtte, hanem az utána lévő mondatra vonatkozik, azaz a 16 szeletes kalácsban együtt volt mazsola. Ekkor az eredmény $\frac{(1-e^{-48})/2}{(1-e^{-32})/2+(1-e^{-48})/2}$ lenne, hiszen 16 szeletben a Poisson folyamat tulajdonságai miatt a mazsolák száma $\mathrm{Poi}(16\cdot 3)$ illetve $\mathrm{Poi}(16\cdot 2)$.

(c) A két esemény nem független, hiszen az első szeletbeli mazsolák száma információt ad arról, hogy melyik pékségből jött a kalács:

$$\begin{aligned}\mathbb{P}[X_2\geq 2\,|\,X_1\geq 1] &= \mathbb{P}[X_2\geq 2\,|\,J,\ X_1\geq 1]\,\mathbb{P}[J\,|\,X_1\geq 1]+\mathbb{P}[X_2\geq 2\,|\,P,\ X_1\geq 1]\,\mathbb{P}[P\,|\,X_1\geq 1]\\ &= \mathbb{P}[X_2\geq 2\,|\,J]\,\mathbb{P}[J\,|\,X_1\geq 1]+\mathbb{P}[X_2\geq 2\,|\,P]\,\mathbb{P}[P\,|\,X_1\geq 1]\\ &= (1-e^{-3}-3e^{-3})(\text{(b) rész})+(1-e^{-2}-2e^{-2})(1-\text{(b) rész})\,.\end{aligned}$$

Gyak. 2. Az előző feladatban leírt körülményekkel:

(a) Ha minden kalácsból négy szeletet eszek meg (kettőt frissen, kettőt másnap, a többit a család), átlagosan hány mazsolát kapok én, mekkora szórással? *(Segítség: a* $\mathrm{Poi}(\lambda)$ *szórása* $\sqrt{\lambda}$*, és ne feledkezzünk el a feltételes szórásnégyzet formuláról.)* **(3+9 p)**

(b) Ha egy év alatt 100-szor veszek kalácsot, mi a valószínűsége, hogy 1070-nél is több mazsolát fogok megenni? *(Kellhet, hogy* $\sqrt{3.5}\approx 1.87$*,* $\sqrt{3.55}\approx 1.88$*.)* **(7p)**

Megoldás: Sajnos itt elemi szövegértési problémái voltak sokaknak, a 4, 8, 12, 16, 32, 48, 100 számok elég sajátos kombinációit eredményezve.

(a) A feltételes szórásnégyzet formulát akkor használjuk, amikor két forrásból jön a véletlenség; jelen esetben, hogy melyik pékségből jön a kalács, és hogy az ezáltal megadott Poisson folyamat mennyi mazsolát is eredményez. Tehát X legyen a mazsolák száma 4 szelet kalácsban, ez érdekel minket, Y pedig az, hogy honnan jött a kalács. Világos, hogy $\mathbb{E}[X\,|\,Y=J]=12$, $\mathbb{E}[X\,|\,Y=P]=8$, $\mathbb{D}^2[X\,|\,Y=J]=12$, $\mathbb{D}^2[X\,|\,Y=P]=8$.

$$\mathbb{E}[X]=\mathbb{E}[\mathbb{E}[X|Y]]=12\,\mathbb{P}[Y=J]+8\,\mathbb{P}[Y=P]=10\,.$$

Persze ez a 10, ez józan paraszti ésszel is nyilvánvaló, a toronyszabály csak egy elegáns absztrakt leírási módja ennek.

$$\begin{aligned}\mathbb{D}^2[X] &= \mathbb{E}[\mathbb{D}^2[X|Y]]+\mathbb{D}^2[\mathbb{E}[X|Y]]\\ &= 12\,\mathbb{P}[Y=J]+8\,\mathbb{P}[Y=P]+\mathbb{D}^2[\text{1/2-del }8,\ \text{1/2-del }12]\\ &= 10+4\,.\end{aligned}$$

Tehát a szórás $\sqrt{14}$.

(b) 100-szor történik függetlenül az (a)-beli kísérlet, amiben $\mu = 10$, $\sigma = \sqrt{14}$. Tehát a CHT-t használhatjuk:

$$\begin{aligned}\mathbb{P}[X_1 + \cdots + X_{100} > 1070] &= \mathbb{P}\left[\frac{X_1 + \cdots + X_{100} - 100\mu}{10\sigma} > \frac{1070.5 - 100\mu}{10\sigma}\right] \\ &\approx \mathbb{P}\left[\mathsf{N}(0,1) > \frac{7.05}{\sqrt{14}}\right] \\ &\approx 1 - \Phi(\sqrt{3.55})\,.\end{aligned}$$

Gyak. 3. Legyenek X és Y független $\mathsf{Exp}(\lambda)$ eloszlású valváltozók.

(a) Határozd meg $(X, X+Y)$ együttes sűrűségfüggvényét. **(7p)**

(b) Határozd meg $X+Y$ sűrűségfüggvényét. **(7p)**

(c) Föltéve, hogy $X+Y = z$, valamilyen fix $z > 0$-ra, milyen eloszlású az X? **(7p)**

(Figyeljünk a határokra!)

Megoldás: (a) Az (X, Y) együttes sűrűségfgvényét tudjuk, függetlenség miatt: $f_{X,Y}(x,y) = \lambda^2 e^{-\lambda(x+y)}$, ahol $(x,y) \in [0,\infty)^2$. Namármost, egy szép lineáris leképezéssel megkaphatjuk $(X, X+Y)$-t:

$$g(X,Y) := \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} X \\ Y \end{pmatrix} = \begin{pmatrix} X \\ X+Y \end{pmatrix}.$$

Tehát a standard módszerrel:

$$\begin{aligned}f_{X,X+Y}(x,z) &= f_{X,Y}(g^{-1}(x,z)) \cdot |\mathsf{Jac}_g(g^{-1}(x,z))|^{-1} \\ &= f_{X,Y}(x, z-x) \cdot 1^{-1} = \lambda^2 e^{-\lambda z}\,.\end{aligned}$$

Első ránézésre mintha nem függene x-től, pedig de: csak a $0 \le x \le z$ síkrészen van ez.

(b) Az egyik lehetőség az előző együttesből integrálással megkapni a marginálist:

$$f_{X+Y}(z) = \int_0^z \lambda^2 e^{-\lambda z}\, dx = z\lambda^2 e^{-\lambda z}\,.$$

Egy másik lehetőség, amihez nem kell az (a), konvolúcióval (integrálási határokra, milyen változó szerint integrálunk, tessék figyelni):

$$f_{X+Y}(z) = \int_0^z \lambda e^{-\lambda x} \lambda e^{-\lambda(z-x)}\, dx = z\lambda^2 e^{-\lambda z}\,.$$

Egy harmadik lehetőség észrevenni, hogy egy Poisson(λ) pontfolyamat második pontjáról van szó, így, $N_{a,b}$-vel jelölve az érkezések számát (a,b)-ben,

$$\begin{aligned}f_{X+Y}(z)\, dz \approx \mathbb{P}[P_2 \in (z, z+dz)] &= \mathbb{P}[N_{0,z} = 1] \cdot \mathbb{P}[N_{z,z+dz} = 1] \\ &= \mathbb{P}[\mathsf{Poi}(\lambda z) = 1] \cdot \mathbb{P}[\mathsf{Poi}(\lambda dz) = 1] \\ &= \lambda z e^{-\lambda z} \cdot \lambda dz e^{-\lambda dz} \approx z\, dz\, \lambda^2 e^{-\lambda z}\,.\end{aligned}$$

(c) Ha valaki megcsinálta az (a) részt, a (b) nem is kell:

$$f_{X\,|\,X+Y}(x|z) = \frac{f_{X,X+Y}(x,z)}{f_{X+Y}(z)} = \frac{z\lambda^2 e^{-\lambda z} \mathbf{1}_{\{0 \le x \le z\}}}{f_{X+Y}(z)} = c_z \mathbf{1}_{\{0 \le x \le z\}}\,.$$

Ez a sűrűségfgv bizony konstans c_z a $[0,z]$ intervallumon, máshol nulla, szóval ez az egyenletes eloszlás $[0,z]$-n, és persze csak $c_z = 1/z$ lehet. (Amit amúgy (b)-ből is látunk.)

Egy másik lehetőség megint a Poisson(λ) pontfolyamatot használni:

$$\begin{aligned}\mathbb{P}\big[P_1 \in (x, x+dx) \,\big|\, P_2 \in (z, z+dz)\big] &= \frac{\mathbb{P}[N_{0,x} = 0] \cdot \mathbb{P}[N_{x,x+dx} = 1] \cdot \mathbb{P}[N_{x+dx,z} = 0] \cdot \mathbb{P}[N_{z,z+dz} = 1]}{\mathbb{P}[N_{0,z} = 1] \cdot \mathbb{P}[N_{z,z+dz} = 1]} \\ &= \frac{\lambda e^{-\lambda x} \cdot \lambda dx e^{-\lambda dx} \cdot e^{-\lambda(z-x)}}{\lambda z e^{-\lambda z}} = \frac{dx}{z}\,.\end{aligned}$$