2. vizsga, 2017-01-09, megoldásokkal

1. **(a)** Adja meg a

i	0	1	2	3	...
p_i	p_0	p_1	p_2	p_3	...

diszkrét eloszlás második momentumának a definícióját!

(b) Felhasználva, hogy az $X = i$ esemény relatív gyakorisága – nagy kísérletszám esetén – közel van a p_i valószínűséghez, írja le annak a vázlatos bizonyítását , hogy sok kísérleti eredmény négyzetének az átlaga közel van a második momentumhoz! *(Matematikai levezetést kérünk, nem pedig számítógépes szimuláció elmesélését.)*

Megoldás **(a)** (5 pont)
A második momentum:

$$\mathbf{E}(X^2) = \sum_{i=0}^{\infty} i^2 p_i = \sum_{i=1}^{\infty} i^2 p_i$$

(b) (5 pont)
$X_1, \dots, X_N$ kísérletek. Legyen N_i, hogy ebből hányszor kaptuk az i eredményt. Tehát az i relatív gyakorisága $N_i/N \approx p_i$. Így

$$\frac{X_1^2 + \cdots + X_N^2}{N} = \frac{N_1}{N}\,1^2 + \cdots + \frac{N_i}{N}\,i^2 + \cdots \approx p_1 1^2 + \cdots + p_i i^2 + \cdots = \mathbf{E}(X^2)$$

2. Egy városban pár év óta elég sok trafipax-szal próbálják a száguldozást megfékezni. A taxi-központ az év végén minden egyes taxisra összesítette, hogy hányszor kapták el gyorshajtásért. Mi csak annyit tudunk, hogy 110 taxist egyszer, 55-öt kétszer. **(a)** Kb. hány taxi szaladgál a városban? **(b)** A taxisoknak kb. hány százaléka kapott egy vagy több büntetést gyorshajtásért? *(A használt modell jgosságát indokolni is kell!)*

Megoldás Minden egyes taxisra igaz, hogy sok gyorshajtásának mindegyikét egymástól függetlenül kis valószínűséggel veszik észre. Így az egyes taxisok tettenéréseinek X számát független Poisson változókkal modellezhetjük, ismeretlen λ paraméterrel
(2 pont).

N db taxi esetén:

$$110 \approx N\,\mathbf{P}(X = 1) = N e^{-\lambda} \lambda$$

illetve

$$55 \approx N\,\mathbf{P}(X = 2) = N e^{-\lambda} \lambda^2/2$$

(a) (5 pont)
A két egyenlet hányadosa $\lambda/2 = 1/2$-et ad, ahonnan $\lambda = 1$ és $N \approx 110e \approx 299$.

(b) (3 pont)
A tettenértek aránya $\mathbf{P}(X > 0) = 1 - e^{-1} \approx 63.2\%$.

3. 100 diák mindegyike, a többitől függetlenül 0.8 valószínűséggel adja be házi feladatát a megadott határidőig. **(a)** Mi a valószínűsége annak, hogy a diákoknak legalább a háromnegyede a megadott határidőig beadja a házi feladatát? *(A választ itt egy korrekt szumma alakjában kérjük.)* **(b)** A feltett kérdésre adjon közelítő választ normális eloszlás segítségével! *(A lentebbi táblázatot használja!)*

Megoldás **(a)** (4 pont)
Az időben leadott dolgozatok X száma Binom eloszlást követ, $n = 100$, $p = 0.8$ paraméterekkel.

$$\mathbf{P}(X \geq 75) = \sum_{k=75}^{100} \binom{100}{k} 0.8^k 0.2^{100-k}$$

(b) (6 pont)

$$\mathbf{P}(X \geq 75) = \mathbf{P}(X \geq 74.5) = \mathbf{P}\left(\frac{X - np}{\sqrt{np(1-p)}} \geq \frac{74.5 - 1000.8}{\sqrt{np(1-p)}}\right) \approx 1 - \Phi(-5.5/4) = \Phi(1.375) \approx 0.92$$

x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$
0,0	0,50	0,5	0,69	1,0	0,84	1,5	0,93	2,0	0,98	2,5	0,99
0,1	0,54	0,6	0,73	1,1	0,86	1,6	0,95	2,1	0,98	2,6	1,00
0,2	0,58	0,7	0,76	1,2	0,88	1,7	0,96	2,2	0,99		
0,3	0,62	0,8	0,79	1,3	0,90	1,8	0,96	2,3	0,99		
0,4	0,66	0,9	0,82	1,4	0,92	1,9	0,97	2,4	0,99		

4. Van egy piros és egy zöld lámpám. Élettartamaik függetlenek, és rendelkeznek az örökifjú tulajdonsággal. A piros élettartamának a várható értéke 200 nap, a zöldnek pedig a mediánja 200 nap. **(a)** *Az eloszlásfüggvények összevetésével* döntse el, melyik lámpa a jobb! **(b)** Mennyi a valószínűsége annak, hogy a piros élettartama nagyobb a zöldénél? *(Segítség: határozza meg az eseményteret, adja meg a kétdimenziós sűrűségfüggvény képletét, határozza meg a kedvező kimenetelek halmazát, stb.)*

Megoldás **(a)** (4 pont)
Ha X λ paraméterű eloszlást követ, akkor

$$P(X < x) = 1 - e^{-\lambda x} \quad \text{és} \quad \mathbf{E}(X) = 1/\lambda$$

A második állítás alapján:

$$1/\lambda_{\text{piros}} = 200, \quad \text{így} \quad \lambda_{\text{piros}} = \frac{1}{200}, \quad \text{ezért} \quad F_{\text{piros}}(x) = 1 - e^{-\frac{1}{200}x}$$

Az első első állítás alapján:

$$e^{-\lambda_{\text{zöld}}200} = 1/2, \quad \text{így} \quad \lambda_{\text{zöld}} = \frac{\ln 2}{200}, \quad \text{ezért} \quad \text{zöld}(x) = 1 - e^{-\frac{\ln(2)}{200}x}$$

tehát piros lámpa esetén annak a valószínűsége, hogy x előtt kiég nagyobb, mint zöld lámpa esetén (minden x-re), tehát a zöld lámpa jobb.

Mivel $\ln(2) < 1$, ezért $\lambda_{\text{zöld}} < \lambda_{\text{piros}}$, így a zöld lámpa élettartartamának nagyobb a várható értéke és a mediánja is, mint a piros lámpa élettartamának. Sőt, tetszőleges x időtartamra $F_{\text{piros}}(x) > F_{\text{zöld}}(x)$. A zöld lámpa jobb.

(b) (6 pont)
Legyen $X =$ a piros lámpa élettartama, $Y =$ a zöld lámpa élettartama. Az (X, Y) kétdimenziós valószínűségi változó sűrűségfüggvénye szorzat alakú:

$$f(x, y) = \lambda_{\text{piros}} e^{-\lambda_{\text{piros}} x} \lambda_{\text{zöld}} e^{-\lambda_{\text{zöld}} y}$$

ahol (x, y) a pozitív síknegyedben van.

$$\mathbf{P}(X > Y) = \int_0^\infty \int_y^\infty f(x, y)\, dx\, dy = \int_0^\infty \lambda_{\text{zöld}} e^{-\lambda_{\text{zöld}} y} \left(\int_y^\infty \lambda_{\text{piros}} e^{-\lambda_{\text{piros}} x}\, dx \right) dy$$

$$= \int_0^\infty \lambda_{\text{zöld}} e^{-\lambda_{\text{zöld}} y} e^{-\lambda_{\text{piros}} y} \, dy$$

$$= \frac{\lambda_{\text{zöld}}}{\lambda_{\text{zöld}} + \lambda_{\text{piros}}}$$

Jelen esetben ez $\frac{\ln 2}{\ln 2+1}$.

5. *(Folytatás)* Kísérleti eredményekre támaszkodva magyarázza el, hogy hogy mit jelent az a tény, hogy **(a)** a piros lámpa élettartamának a várható értéke 200 nap! **(b)** a zöld lámpa élettartamának a mediánja 200 nap!

Megoldás **(a)** (5 pont)
Sok (N db) kísérletet végzünk, az eredményeket $X_1, \ldots, X_N$-nel jelöljük. Az élettartamok átlaga nagy kísérletszám esetén körülbelül 200.

(b) (5 pont)
Azon kísérletek aránya, melyekben az élettartam 200 napnál nagyobb, körülbelül 1/2.

6. Az (X, Y) kétdimenziós valószínűségi változó sűrűségfüggvénye

$$f(x,y) = \frac{2x}{y} \quad (0 < x < 1,\ x < y < \frac{1}{x})$$

(a) Határozza meg az $f_{2|1}(y|x)$ feltételes sűrűségfüggvény képletét! **(b)** Milyen $y = k(x)$ függvénnyel tippeljünk X-ből Y-ra, ha azt szeretnénk, hogy a hiba négyzetének a várható értéke minimális legyen? *(Határozza meg a $k(x)$ függvény képletét!)*

Megoldás **(a)** (5 pont)
X sűrűségfüggvénye:

$$f_1(x) = \int_x^{1/x} \frac{2x}{y} \, dy = 2x[\ln y]_x^{1/x} = 4x \ln \frac{1}{x} \quad \text{ha} 0 < x < 1$$

Így a feltételes sűrűségfüggvény:

$$f_{2|1}(y|x) = \frac{f(x,y)}{f_1(x)} = \frac{1}{2y\ln(1/x)}, \qquad \text{ha } x < y < \frac{1}{x}$$

(b) (5 pont)
A hiba négyzetének várható értékét a feltételes várható érték minimalizálja, tehát

$$k(x) = \int_x^{1/x} y \, f_{2|1}(y|x) \, dy = \int_x^{1/x} \frac{y}{2y\ln(1/x)} \, dy = \left(\frac{1}{x} - x\right) \frac{1}{2\ln(1/x)}$$