Ime in priimek: ______________________ Vpsina št: | | | | | | | | |

Univerza na Primorskem

FAMNIT (MA, MF)

Verjetnost

2. kolokvij

28. maj 2025

Navodila

Pazljivo preberite besedilo naloge, preden se lotite reševanja. Nalog je 6, s 5 rešenimi nalogami pa desežete 100%. Če jih rešite več, imate presežne točke v dobrem. Na razpolago imate 120 minut, dovoljena sredstva pa so A4 format list s formulami.

Naloga	a.	b.	c.	d.	Skupaj
1.			•	•	
2.			•	•	
3.			•	•	
4.			•	•	
5.			•	•	
6.			•	•	
Skupaj					

1. (20) Pri igri *Črni Peter* imamo 26 parov kart. Karti v vsakem paru sta enaki. Predpostavite, da imamo štiri igralce. Vseh 52 kart dobro premešamo in jih razdelimo igralcem, tako da vsak dobi 13 kart. Za $k = 1, 2, 3, 4$ naj bo X_k število parov, ki jih dobi igralec k.

a. (10) Konkretni primer para kart sta dva obuta mačka. Izračunajte verjetnost, da prvi igralec dobi točno ta par kart.

Rešitev: zaradi načina deljenja kart bo prvi igralec dobil naključni vzorec 13 *kart izmed vseh* 52 *kart. Takih naključnih vzorcev je* $\binom{52}{13}$. *Vzorcev, ki vsebujejo par obutih mačkov, pa je* $\binom{52-2}{13-2}$, *tako da je*

$$P(\textit{prvi igralec ima par obutih mačkov}) = \frac{\binom{50}{11}}{\binom{52}{13}} = \frac{1}{17}\,.$$

b. (10) Izračunajte $E(X_k)$.

Rešitev: pare oštevilčimo z $1, 2, \ldots, 26$ *in definiramo indikatorje*

$$I_{k,l} = \begin{cases} 1\,, & \textit{če igralec } k \textit{ ima } l\textit{-ti par} \\ 0 & \textit{sicer.} \end{cases}$$

Velja $X_k = I_{k,1} + \cdots + I_{k,26}$. *Iz prvega dela po simetriji sledi*

$$E(I_{k,l}) = P(I_{k,l} = 1) = \frac{1}{17}\,,$$

torej

$$E(X_k) = E(I_{k,1} + \cdots + I_{k,26}) = \frac{26}{17}\,.$$

2. (20) Nenegativni celoštevilski slučajni spremenljivki imata porazdelitev dano z

$$P(X=k,Y=l)=\frac{(a)_{k+l}}{2^{2k+2l+a}k!\cdot l!}$$

za $a>0$ in $k,l\geq 0$, kjer je

$$(a)_0=1 \qquad \text{in} \qquad (a)_n=a(a+1)\cdots(a+n-1)$$

Pochhammerjev simbol. Definirajte $Z=X+Y$.

a. (10) Izračunajte porazdelitev slučajne spremenljivke Z.

Rešitev: računamo

$$\begin{aligned}
P(Z=n) &= \sum_{k=0}^{n} P(X=k,Y=n-k) \\
&= \sum_{k=0}^{n} \frac{(a)_n}{2^{2n+a}k!\cdot(n-k)!} \\
&= \frac{(a)_n}{n!\cdot 2^{2n+a}} \sum_{k=0}^{n} \binom{n}{k} \\
&= \frac{(a)_n}{n!\cdot 2^{n+a}}\,.
\end{aligned}$$

b. (10) Izračunajte porazdelitev slučajne spremenljivke X.

Namig: po Newtonu je za $|x|<1$

$$(1-x)^{-a}=\sum_{l=0}^{\infty}\frac{(a)_l}{l!}x^l\,.$$

Kot znano upoštevajte

$$(a)_{k+l}=(a)_k\,(a+k)_l$$

Rešitev: po formuli za robno porazdelitev je

$$\begin{aligned}
P(X=k) &= \sum_{l=0}^{\infty} P(X=k, Y=l) \\
&= \sum_{l=0}^{\infty} \frac{(a)_{k+l}}{2^{2k+2l+a} k! \cdot l!} \\
&= \frac{1}{2^{2k+a} k!} \sum_{l=0}^{\infty} \frac{(a)_{k+l}}{2^{2l} l!} \\
&= \frac{1}{2^{2k+a} k!} \sum_{l=0}^{\infty} \frac{(a)_k (a+k)_l}{2^{2l} l!} \\
&= \frac{(a)_k}{2^{2k+a} k!} \left(1 - \frac{1}{4}\right)^{-a-k} \\
&= \frac{2^a (a)_k}{3^{k+a} k!} \,.
\end{aligned}$$

3. (20) Slučajne spremenljivke Z, S in X naj bodo neodvisne. Naj bo $Z \sim \exp(1/2)$, S naj ima gostoto

$$f_S(s) = \frac{1}{2\sqrt{\pi s^3}}\, e^{-\frac{1}{4s}}$$

za $s > 0$ in 0 sicer ter $X \sim \mathrm{N}(0,1)$. Kot znano privzemite, da za $a > 0$ in $b \geq 0$ velja

$$\int_0^\infty \frac{e^{-ay-\frac{b}{y}}}{\sqrt{y}}\,\mathrm{d}y = \frac{\sqrt{\pi}}{\sqrt{a}}\, e^{-2\sqrt{ab}}\,.$$

Definirajte

$$U = \frac{\sqrt{Z}}{S} X\,.$$

a. (10) Izračunajte gostoto slučajnega vektorja (Z, S, U) in iz nje izpeljite gostoto slučajnega vektorja (S, U).

Rešitev: definiramo preslikavo Φ s predpisom

$$\Phi(z, s, x) = \left(z, s, \sqrt{z}x/s\right)\,.$$

Na množici $(0,\infty)^2 \times (-\infty,\infty)$ je ta preslikava bijektivna in zvezno parcialno odvedljiva, prav tako je zvezno parcialno odvedljiv njen inverz. Računamo

$$\Phi^{-1}(z, s, u) = \left(z, s, su/\sqrt{z}\right)\,.$$

Računamo

$$J_{\Phi^{-1}}(z, s, u) = s/\sqrt{z}\,.$$

Gostota slučajnega vektorja (Z, S, U) je torej

$$f_{Z,S,U}(z, s, u) = \frac{1}{2}e^{-z/2} f_S(s) \frac{s}{\sqrt{2\pi z}}\, e^{-\frac{s^2u^2}{2z}}\,.$$

Gostoto slučajnega vektorja (S, U) dobimo kot robno gostoto. Računamo z uporabo integrala, navedenega na začetku

$$\begin{aligned}
f_{S,U}(s, u) &= \int_0^\infty f_{Z,S,U}(z, s, u) \\
&= \frac{s f_S(s)}{2\sqrt{2\pi}} \int_0^\infty \frac{1}{\sqrt{z}}\, e^{-z/2} e^{-\frac{s^2u^2}{2z}} \\
&= \frac{1}{2\sqrt{2\pi}}\, s f_S(s) \sqrt{2\pi}\, e^{-s|u|} \\
&= \frac{s}{2}\, f_S(s)\, e^{-s|u|} \\
&= \frac{1}{4\sqrt{\pi s}}\, e^{-\frac{1}{4s} - s|u|}\,.
\end{aligned}$$

b. (10) Izračunajte gostoto slučajne spremenljivke U.

Rešitev: z upoštevanjem prvega dela naloge dobimo

$$
\begin{aligned}
f_U(u) &= \int_0^\infty f_{S,U}(s,u)\,ds \\
&= \frac{1}{4\sqrt{\pi}} \int_0^\infty \frac{1}{\sqrt{s}}\, e^{-\frac{1}{4s}} e^{-s|u|}\, ds \\
&= \frac{1}{4\sqrt{\pi}} \cdot \frac{\sqrt{\pi}}{\sqrt{|u|}}\, e^{-\sqrt{|u|}} \\
&= \frac{1}{4\sqrt{|u|}}\, e^{-\sqrt{|u|}}\,.
\end{aligned}
$$

4. (20) Naj bodo $Y_1, Y_2, \ldots, Y_n$ slučajne spremenljivke z $\mathrm{var}(Y_k) = \sigma^2$ za vse $k = 1, 2, \ldots, n$ in $\mathrm{cov}(Y_k, Y_l) = \sigma^2 \rho$ za $k \neq l$; pri tem sta σ^2 in $\rho \in (0,1)$ konstanti. Označimo $\bar{Y} = \frac{1}{n}\sum_{k=1}^n Y_k$.

a. (10) Izračunajte $\mathrm{cov}(Y_k, \bar{Y})$ in $\mathrm{var}(\bar{Y})$.

Rešitev: iz bilinearnosti kovariance sledi

$$\begin{aligned}
\mathrm{cov}(Y_k, \bar{Y}) &= \frac{1}{n}\sum_{l=1}^n \mathrm{cov}(Y_k, Y_l) \\
&= \frac{1}{n}\left(\mathrm{cov}(Y_k, Y_k) + \sum_{l;\, l\neq k} \mathrm{cov}(Y_k, Y_l)\right) \\
&= \sigma^2\left(\frac{1}{n} + \frac{n-1}{n}\,\rho\right)
\end{aligned}$$

in nadalje:

$$\mathrm{var}(\bar{Y}) = \mathrm{cov}(\bar{Y}, \bar{Y}) = \frac{1}{n}\sum_{k=1}^n \mathrm{cov}(Y_k, \bar{Y}) = \sigma^2\left(\frac{1}{n} + \frac{n-1}{n}\,\rho\right) .$$

b. (10) Dokažite, da obstaja tak $c \in \mathbb{R}$, da je $\mathrm{cov}(Y_k - c\bar{Y}, Y_l - c\bar{Y}) = 0$ za vse $k \neq l$.

Rešitev: velja

$$\begin{aligned}
&\mathrm{cov}\big(Y_k - c\bar{Y},\, Y_l - c\bar{Y}\big) \\
&\qquad = \mathrm{cov}(Y_k, Y_l) - 2c \cdot \mathrm{cov}(Y_k, \bar{Y}) + c^2 \cdot \mathrm{var}(\bar{Y}) \\
&\qquad = \sigma^2\left[\rho - 2c\left(\frac{1}{n} + \frac{n-1}{n}\,\rho\right) + c^2\left(\frac{1}{n} + \frac{n-1}{n}\,\rho\right)\right] .
\end{aligned}$$

Dobili smo kvadratno enačbo z diskriminanto

$$D = 4\left(\frac{1}{n} + \frac{n-1}{n}\,\rho\right)^2 - 4\rho\left(\frac{1}{n} + \frac{n-1}{n}\,\rho\right) = \frac{4}{n^2}\big(1 + (n-1)\rho\big)(1-\rho)\,,$$

ki je vedno nenegativna, tako da enačba ima realni rešitvi.

5. (20) V televizijskem kvizu voditelj tekmovalcu postavlja vprašanja. Za vsak pravilen odgovor tekmovalec dobi točko. Tekmovalec izpade, ko napačno odgovori na dve zaporedni vprašanji.

Privzemite, da tekmovalec na vprašanja odgovarja pravilno z verjetnostjo $p \in (0, 1)$, neodvisno od pravilnosti odgovorov na ostala vprašanja. Naj bo X število vprašanj, ki jih bo dobil tekmovalec, Y pa število točk, ki jih bo dosegel.

a. (10) Definirajte

$$\begin{aligned} H_1 &= \{\text{tekmovalec odgovori pravilno na prvo vprašanje}\}\,, \\ H_2 &= \{\text{tekmovalec odgovori nepravilno na prvo in pravilno na drugo vprašanje}\}\,, \\ H_3 &= \{\text{tekmovalec odgovori nepravilno na prvi dve vprašanji}\}\,. \end{aligned}$$

Za $i = 1, 2, 3$ izrazite $E(X \mid H_i)$, $E(Y \mid H_i)$ z $E(X)$ in $E(Y)$.

Rešitev: če tekmovalec odgovori pravilno na prvo vprašanje, se zaradi neodvisnosti zgodba začne na novo. Pogojna porazdelitev slučajne spremenljivke X glede na H_1 se torej ujema z brezpogojno porazdelitvijo slučajne spremenljivke $X + 1$, pogojna porazdelitev slučajne spremenljivke Y glede na H_1 pa se ujema z brezpogojno porazdelitvijo slučajne spremenljivke $Y + 1$. Sledi

$$\begin{aligned} E(X \mid H_1) &= E(X + 1) = E(X) + 1\,, \\ E(Y \mid H_1) &= E(Y + 1) = E(Y) + 1\,, \end{aligned}$$

Podobno dobimo pogojno pričakovano vrednost glede na H_2, le da moramo X povečati za 2, Y pa za 1. Torej je

$$\begin{aligned} E(X \mid H_2) &= E(X + 2) = E(X) + 2\,, \\ E(Y \mid H_2) &= E(Y + 1) = E(Y) + 1\,, \end{aligned}$$

Končno je očitno

$$E(X \mid H_3) = 2\,, \quad E(Y \mid H_3) = 0\,.$$

b. (10) Izračunajte $E(X)$ in $E(Y)$.

Rešitev: dogodki H_1, H_2 in H_3 tvorijo popoln sistem dogodkov (particijo verjetnostnega prostora). Če vpeljemo $q = 1 - p$, iz prvega dela naloge in formule za popolno pričakovano vrednost dobimo

$$\begin{aligned} E(X) &= P(H_1)\,E(X \mid H_1) + P(H_2)\,E(X \mid H_2) + P(H_3)\,E(X \mid H_3) \\ &= (1 - q)\big(E(X) + 1\big) + q(1 - q)\big(E(X) + 2\big) + 2q^2\,. \end{aligned}$$

Rešimo enačbo in dobimo

$$E(X) = \frac{1+q}{q^2}\,.$$

Podobno dobimo tudi

$$\begin{aligned} E(Y) &= (1-q)\big(E(Y)+1\big) + q(1-q)\big(E(Y)+1\big)\,, \\ E(Y) &= \frac{1-q^2}{q^2}\,. \end{aligned}$$

6. (20) Zavarovalnica izda 240.000 polic obveznega avtomobilskega zavarovanja. Premija je enaka €252. Povprečje dejansko izplačanih zahtevkov v preteklem letu je bilo €1.240, standardni odklon dejansko izplačanih zahtevkov pa je bil €1.830.

a. (10) Matematik, ki računa tveganja, najprej privzame, da bo zahtevkov 20%, torej 48.000. Zahtevki so med seboj neodvisne, enako porazdeljene slučajne spremenljivke s povprečjem €1240 in standardnim odklonom €1830. Ocenite verjetnost, da zbrane premije ne bodo dovolj za izplačilo vseh zahtevkov.

Rešitev: označimo $n = 48.000$ *in posamezne zahtevke z* $X_1, X_2, \ldots, X_n$. *Zavarovalnica od vseh zavarovancev zbere* $C = 240.000 \times 252$ €. *Označimo* $S_n = X_1 + X_2 + \cdots + X_n$. *Zanima nas verjetnost* $P(S_n > C)$. *Računamo s pomočjo CLI.*

$$\begin{aligned} P(S_n > C) &= P\left(\frac{S_n - E(S_n)}{\sqrt{\mathrm{var}(S_n)}} > \frac{C - E(S_n)}{\sqrt{\mathrm{var}(S_n)}}\right) \\ &\approx P\left(Z > \frac{C - E(S_n)}{\sqrt{\mathrm{var}(S_n)}}\right). \end{aligned}$$

Vemo, da je $E(S_n) = nE(X_1)$ *in* $\mathrm{var}(S_n) = n \cdot \mathrm{var}(X_1)$. *Sledi*

$$\frac{C - E(S_n)}{\sqrt{\mathrm{var}(S_n)}} = 2,39\,.$$

Iz tabele preberemo, da je približek verjetnosti $P(Z > 2,39) \doteq 0,008$.

b. (10) Matematik v zavarovalnici se zaveda, da ne more predvideti točnega števila zahtevkov, ki jih bo zavarovalnica morala izplačati v naslednjem letu. Zato spremeni svojo škatlo tako, da vanjo doda listke z ničlami in sicer štirikrat toliko, kolikor je prvotno listkov. Povprečje škatle se tako spremeni v €248, standardni odklon pa na €956,97. V tem primeru je potrebno računati, kot da lističe izbiramo 240.000-krat. Ocenite zdaj verjetnost, da zbrane premije ne bodo dovolj za izplačilo vseh zahtevkov.

Rešitev: računamo podbno kot v a., le da označimo $n = 240.000$ *in uporabimo novo povprečje in nov standardni odklon. Dobimo*

$$\begin{aligned} P(S_n > C) &= P\left(\frac{S_n - E(S_n)}{\sqrt{\mathrm{var}(S_n)}} > \frac{C - E(S_n)}{\sqrt{\mathrm{var}(S_n)}}\right) \\ &\approx P\left(Z > \frac{C - E(S_n)}{\sqrt{\mathrm{var}(S_n)}}\right), \end{aligned}$$

pri čemer je

$$\frac{C - E(S_n)}{\sqrt{\mathrm{var}(S_n)}} = 2,04\,.$$

Iz tabele preberemo, da je verjetnost približno $P(Z > 2,04) \doteq 0,021$.