Ime in priimek: ______________________ Vpisna št: | | | | | | | | |

FAKULTETA ZA MATEMATIKO IN FIZIKO

Oddelek za matematiko

Verjetnost

Teoretični izpit

19. februar 2024

Navodila

Pazljivo preberite besedilo naloge, preden se lotite reševanja. Nalog je 10, ocena pa je enaka navzgor zaokroženemu številu pravilnih odgovorov.

Naloga	Točke
1.	
2.	
3.	
4.	
5.	
6.	
7.	
8.	
9.	
10.	
Skupaj	

1. V nekem verjetnostnem prostoru velja, da lahko vsak dogodek A zapišemo kot unijo disjunktnih dogodkov $A', A'' \subset A$, za katera je $P(A') = P(A'') = \frac{1}{2}P(A)$. Naj bodo A, B in C dogodki. Utemeljite, da obstaja tak dogodek D, da bo veljalo

$$\begin{aligned}
&P(A \cap D) = \tfrac{1}{2}P(A)\\
&P(B \cap D) = \tfrac{1}{2}P(B)\\
&P(C \cap D) = \tfrac{1}{2}P(C)\\
&P\big((A \cup B) \cap D\big) = \tfrac{1}{2}P(A \cup B)\\
&P\big((A \cup C) \cap D\big) = \tfrac{1}{2}P(A \cup C)\\
&P\big((B \cup C) \cap D\big) = \tfrac{1}{2}P(B \cup C)\\
&P\big((A \cup B \cup C) \cap D\big) = \tfrac{1}{2}P(A \cup B \cup C)\,.
\end{aligned}$$

Rešitev: trditev velja v splošnejši obliki: če so $A_1, \ldots, A_n$ dogodki v tem prostoru, namreč obstaja tak dogodek $D \subset \cup_{k=1}^n A_k$, da velja

$$P\big((A_{i_1} \cup A_{i_2} \cup \cdots \cup A_{i_m}) \cap D\big) = \tfrac{1}{2}P(A_{i_1} \cup A_{i_2} \cup \cdots \cup A_{i_m})$$

za vsak nabor $1 \le i_1 < i_2 < \cdots < i_m \le n$. Oglejmo si dogodke oblike

$$C_p = A_1^* \cap A_2^* \cap \cdots \cap A_n^*\,,$$

kjer je $*$ ali nič ali c; indeks p predstavlja n-terico izbir za $*$ pri vseh množicah. Za različne nabore p so dogodki C_p disjunktni, po predpostavki pa za vsak nabor p obstaja tak dogodek $D_p \subset C_p$, da je $P(D_p) = \frac{1}{2}P(C_p)$. Naj bo $D = \cup_p D_p$. Vsako unijo $A_{i_1} \cup A_{i_2} \cup \cdots \cup A_{i_m}$ lahko napišemo kot disjunktno unijo dogodkov C_p za $p \in \mathcal{P}$, kjer je $\mathcal{P}$ množica vseh n-teric p, pri katerih pri vsaj eni od komponent z indeksi $i_1, i_2, \ldots, i_m$ velja izbira nič. Končno je

$$\begin{aligned}
P\big((A_{i_1} \cup A_{i_2} \cup \cdots \cup A_{i_m}) \cap D\big) &= P\big(\cup_{p\in\mathcal{P}}(C_p \cap D)\big)\\
&= P\big(\cup_{p\in\mathcal{P}} D_p\big)\\
&= \sum_{p\in\mathcal{P}} P(D_p)\\
&= \tfrac{1}{2}\sum_{p\in\mathcal{P}} P(C_p)\\
&= \tfrac{1}{2}P(A_{i_1} \cup A_{i_2} \cup \cdots \cup A_{i_m})\,.
\end{aligned}$$

2. Gostota vektorja (X, Y) je za $(x, y) \in \mathbb{R}^2$ oblike

$$f_{X,Y}(x, y) = f_X(x)\, g(y - ax)$$

za funkcijo g in konstanto a. Sta slučajni spremenljivki X in $Y - aX$ neodvisni?

Rešitev: označimo $Z = Y - aX$. Transformacijska formula nam da

$$f_{X,Z}(x, z) = f_X(x)\, g(z)\,.$$

Gostota na desni je produkt funkcije samo spremenljivke x in funkcije samo spremenljivke z, iz česar sledi neodvisnost.

3. Računalnik generira neodvisne slučajne spremenljivke $U_1, U_2, \ldots$ z $U_k \sim U(0,1)$ za $k = 1, 2, \ldots$. Kako bi generirali neodvisne pare $(X_1, Y_1), (X_2, Y_2), \ldots$ z gostoto

$$f(x,y) = \begin{cases} 2 & \text{za } 0 < x < y < 1, \\ 0 & \text{sicer?} \end{cases}$$

Rešitev:

Prvi način: *postavimo* $(X_k, Y_k) = \big(\min(U_{2k-1}, U_{2k}), \max(U_{2k-1}, U_{2k})\big)$ *za* $k = 1, 2, \ldots$ *Neodvisnost je zagotovljena, pravilnost porazdelitve pa utemeljimo tako, da za Jordanovo merljivo množico* $G \subset \{(x,y) \colon 0 < x < y < 1\}$ *opazimo, da je*

$$P((X_k, Y_k) \in G) = P((U_{2k-1}, U_{2k}) \in G) + P((U_{2k-1}, U_{2k}) \in \tilde{G}) = 2|G|\,,$$

kjer je $\tilde{G}$ *množica* G, *zrcaljena preko simetrale lihih kvadrantov,* $|G|$ *pa ploščina množice* G.

Drugi način: *preslikava*

$$\Phi(u_1, u_2) = (\sqrt{u_2}\,u_1, \sqrt{u_2})$$

preslika $(0,1)^2$ *bijektivno na* G. *Velja*

$$\Phi^{-1}(x,y) = \left(\frac{x}{y}, y^2\right)$$

z $J_{\Phi^{-1}}(x,y) = 2$. *Definiramo lahko torej*

$$(X_k, Y_k) = \left(\sqrt{U_{2k}}\,U_{2k-1}, \sqrt{U_{2k}}\right)\,.$$

Neodvisnost parov sledi iz neodvisnosti $U_1, U_2, \ldots$, *vsi pa imajo pravo gostoto. Kako zgornje uganemo? Par* (X_k, Y_k) *ima dano gostoto, kar pomeni, da je gostota* Y_k *enaka* $f_Y(y) = 2y$, *kar je točno gostota* $\sqrt{U_2}$. *Ko enkrat izberemo* y *koordinato, mora biti* X_k *pogojno enakomerno porazdeljen na* $(0, y)$, *kar izraža prva komponenta.*

4. Slučajni spremenljivki X in Y naj imata gostoto, dano z

$$f_{X,Y}(x,y) = \begin{cases} 2 & \text{za } x, y > 0 \text{ in } x + y < 1 \\ 0 & \text{sicer.} \end{cases}$$

Utemeljite, da imajo slučajne spremenljivke X, Y in $1 - X - Y$ enako porazdelitev.

Rešitev: slučajni spremenljivki X in Y imata enako porazdelitev po simetriji, gostota pa je $f_X(x) = 2(1-x)$. Označimo $Z = 1 - X - Y$. Računamo

$$P(Z \leq z) = P(1 - z \leq X + Y) = 1 - (1-z)^2$$

za $0 < z < 1$. Z odvajanjem sledi $f_Z(z) = 2(1-z)$.

5. Naj bodo spremenljivke U_1, U_2, Z_1, Z_2 neodvisne, z $U_i \sim \Gamma(1,1)$ in $Z_i \sim \mathrm{N}(0,1)$ za $i = 1, 2$. Vzemimo še $\theta \in \mathbb{R}$ in definirajmo

$$X_1 = \theta U_1 + \sqrt{U_1}\, Z_1 \qquad \text{ter} \qquad X_2 = \theta U_2 + \sqrt{U_2}\, Z_2 \,.$$

Utemeljite, da ima vsota $X_1 + X_2$ enako gostoto kot slučajna spremenljivka

$$X = \theta U + \sqrt{U}\, Z \,,$$

kjer sta U in Z neodvisni, $U \sim \Gamma(2,1)$ in $Z \sim \mathrm{N}(0,1)$.

Rešitev:

Prvi način. *Zapišemo*

$$X_1 + X_2 = \theta(U_1 + U_2) + \sqrt{U_1 + U_2}\left(\sqrt{\frac{U_1}{U_1 + U_2}}\, Z_1 + \sqrt{\frac{U_2}{U_1 + U_2}}\, Z_2\right) .$$

Vemo, da je kvocient $U_1/(U_1 + U_2)$ neodvisen od vsote $U_1 + U_2$, zato je vsota v oklepaju neodvisna od $U_1 + U_2 \sim \Gamma(2,1)$. Poleg tega je vsota v oklepaju pogojno na $\{U_1 = u_1, U_2 = u_2\}$ standardizirano normalna in posledično tudi brezpogojno standardizirano normalna ter neodvisna od vsote $U_1 + U_2$. Trditev sledi.

Drugi način. *Naj bo g gostota porazdelitve $\Gamma(1,1)$, ϕ pa naj bo gostota standardne normalne porazdelitve. Po predpostavki ima slučajni vektor (U_1, U_2, Z_1, Z_2) gostoto $f_{U_1,U_2,Z_1,Z_2}(u_1, u_2, z_1, z_2) = g(u_1)\, g(u_2)\, \phi(z_1)\, \phi(z_2)$. Preslikava $\Phi\colon\ (0,\infty)^2 \times \mathbb{R}^2 \to (0,\infty)^2 \times \mathbb{R}^2$, definirana po predpisu*

$$\Phi(u_1, u_2, z_1, z_2) = \big(u_1,\ u_2,\ \theta u_1 + \sqrt{u_1}\, z_1,\ \theta u_2 + \sqrt{u_2}\, z_2\big) \,,$$

je difeomorfizem z inverzom

$$\Phi^{-1}(u_1, u_2, x_1, x_2) = \left(u_1,\ u_2,\ \frac{x_1 - \theta u_1}{\sqrt{u_1}},\ \frac{x_2 - \theta u_2}{\sqrt{u_2}}\right) ,$$

ki ima Jacobijevo determinanto $1/\sqrt{u_1 u_2}$. Sledi, da ima slučajni vektor (U_1, U_2, X_1, X_2) gostoto

$$f_{U_1,U_2,X_1,X_2}(u_1, u_2, x_1, x_2) = \frac{1}{\sqrt{u_1 u_2}}\, g(u_1)\, g(u_2)\, \phi\!\left(\frac{x_1 - \theta u_1}{\sqrt{u_1}}\right) \phi\!\left(\frac{x_2 - \theta u_2}{\sqrt{u_2}}\right) .$$

Iz nje dobimo pogojno gostoto

$$f_{X_1,X_2 \mid U_1,U_2}(x_1, x_2 \mid u_1, u_2) = \frac{1}{\sqrt{u_1}}\, \phi\!\left(\frac{x_1 - \theta u_1}{\sqrt{u_1}}\right) \cdot \frac{1}{\sqrt{u_2}}\, \phi\!\left(\frac{x_2 - \theta u_2}{\sqrt{u_2}}\right) ,$$

ki pove, da sta X_1 in X_2 pogojno na $U_1 = u_1, U_2 = u_2$ neodvisni z $X_1 \mid U_1 = u_1, U_2 = u_2 \sim \mathrm{N}(\theta u_1, u_1)$ in $X_2 \mid U_1 = u_1, U_2 = u_2 \sim \mathrm{N}(\theta u_2, u_2)$. Če *definiramo* $X := X_1 + X_2$, *je torej* $X \mid U_1 = u_1, U_2 = u_2 \sim \mathrm{N}\big(\theta(u_1+u_2), u_1+u_2\big)$, *kar pomeni, da ima slučajni vektor* (U_1, U_2, X) *gostoto*

$$f_{U_1,U_2,X}(u_1, u_2, x) = \frac{1}{\sqrt{u_1 + u_2}}\, g(u_1)\, g(u_2)\, \phi\!\left(\frac{x - \theta(u_1 + u_2)}{\sqrt{u_1 + u_2}}\right) .$$

Preslikava $\Phi\colon\ (0,\infty)^2 \times \mathbb{R}^2 \to (0,\infty)^2 \times \mathbb{R}^2$, *definirana po predpisu*

$$\Psi(u_1, u_2, x) = \left(u_1,\ u_2,\ \frac{x_1 - \theta(u_1 + u_2)}{\sqrt{u_1 + u_2}}\right) ,$$

je difeomorfizem z inverzom

$$\Psi^{-1}(u_1, u_2, z) = \big(u_1,\ u_2,\ \theta(u_1 + u_2) + \sqrt{u_1 + u_2}\, z\big) ,$$

ki ima Jacobijevo determinanto $\sqrt{u_1 + u_2}$. *Če torej definiramo* $U := U_1 + U_2$ *in* $Z := (X - \theta U)/\sqrt{U}$, *je slučajni vektor* (U_1, U_2, Z) *spet po transformacijski formuli porazdeljen zvezno z gostoto* $f_{U_1,U_2,Z}(u_1, u_2, z) = g(u_1)\, g(u_2)\, \phi(z)$. *Torej so slučajne spremenljivke* U_1, U_2 *in* Z *neodvisne,* Z *pa je porazdeljena standardizirano normalno. Potem pa sta neodvisni tudi* U *in* Z, *vemo, da je* $U \sim \Gamma(2,1)$, *velja pa tudi* $X_1 + X_2 = X = \theta U + \sqrt{U}\, Z$.

6. Naj bo X celoštevilska nenegativna slučajna spremenljivka in naj bo $E(X) = a$ in $\operatorname{var}(X) = b$. Izračunajte

$$\sum_{k=1}^{\infty} 2k\, P(X \geq k)\,.$$

Rešitev: računamo

$$\begin{aligned}
\sum_{k=1}^{\infty} 2k\, P(X \geq k) &= \sum_{k=1}^{\infty} 2k \sum_{l=k}^{\infty} P(X = l) \\
&= \sum_{l=1}^{\infty} P(X = l) \sum_{k=1}^{l} 2k \\
&= \sum_{l=1}^{\infty} l(l+1)\, P(X = l) \\
&= E\big(X(X+1)\big) \\
&= b + a^2 + a\,.
\end{aligned}$$

7. Naj bodo $X_1, X_2, \ldots, X_n$ neodvisne, enako porazdeljene slučajne spremenljivke. Definirajte $\bar{X} = \frac{1}{n}\sum_{k=1}^{n} X_k$ in označite $\sigma^2 = \mathrm{var}(X_1)$. Izračunajte

$$E\left[\sum_{k=1}^{n}\left(X_k - \bar{X}\right)^2\right] .$$

Rešitev: iz linearnosti dobimo, da je $E(X_k) = E(\bar{X})$ za vsak $k = 1, 2, \ldots, n$. Računamo z upoštevanjem neodvisnosti in bilinearnosti kovarianc ter dejstva, da je kovarianca neodvisnih spremenljivk enaka 0. *Dobimo*

$$\begin{aligned} E\left[\left(X_k - \bar{X}\right)^2\right] &= \mathrm{var}(X_k - \bar{X}) \\ &= \mathrm{var}(X_k) - 2\,\mathrm{cov}(X_k, \bar{X}) + \mathrm{var}(\bar{X}) \\ &= \sigma^2 - \frac{2\sigma^2}{n} + \frac{\sigma^2}{n} \\ &= \frac{\sigma^2(n-1)}{n} . \end{aligned}$$

Končno sledi

$$E\left[\sum_{k=1}^{n}\left(X_k - \bar{X}\right)^2\right] = (n-1)\sigma^2 .$$

8. Naj bodo X, Y, Z diskretne s končno mnogo vrednostmi ter pričakovano vrednostjo 0 in varianco 1. Naj velja

$$E(Y \mid X = x, Z = z) = ax + bz$$

za konstanti a in b. Izrazite ti dve konstanti s $\operatorname{cov}(X, Y)$, $\operatorname{cov}(Y, Z)$ in $\operatorname{cov}(X, Z)$.

Namig: izračunajte $E(XY)$ in $E(ZY)$ po formuli za popolno pričakovano vrednost.

Rešitev: po formuli za popolno pričakovano vrednost je

$$\begin{aligned}
\operatorname{cov}(X, Y) &= E(XY) \\
&= \sum_{x,z} E(XY \mid X = x, Z = z)\, P(X = x, Z = y) \\
&= \sum_{x,z} x(ax + bz)\, P(X = x, Z = z) \\
&= a\, E(X^2) + b\, E(XZ) \\
&= a + b \operatorname{cov}(X, Z)\,.
\end{aligned}$$

Podobno dobimo

$$\operatorname{cov}(Y, Z) = a \operatorname{cov}(X, Z) + b\,.$$

Iz obeh enačb sledi

$$a = \frac{\operatorname{cov}(X, Y) - \operatorname{cov}(X, Z) \operatorname{cov}(Y, Z)}{1 - \operatorname{cov}(X, Z)^2}$$

in

$$b = \frac{\operatorname{cov}(Y, Z) - \operatorname{cov}(X, Y) \operatorname{cov}(X, Z)}{1 - \operatorname{cov}(X, Z)^2}\,.$$

9. Naj bosta X, Y diskretni slučajni spremenljivki s končno mnogo vrednostmi in $\mathrm{var}(Y) = a$. Naj bo

$$E(Y \mid X = x) = f(x)$$

in

$$\mathrm{var}(Y \mid X = x) = g(x)$$

za funkciji f, g. Izračunajte

$$E\big(g(X)\big) + \mathrm{var}\big(f(X)\big)\,.$$

Rešitev: računamo

$$\begin{aligned}
E\big(g(X)\big) &= \sum_x g(x)\, P(X = x) \\
&= \sum_x \big(E(Y^2 \mid X = x) - E(Y \mid X = x)^2\big)\, P(X = x) \\
&= E(Y^2) - \sum_x E(Y \mid X = x)^2\, P(X = x)\,.
\end{aligned}$$

Po drugi strani je

$$\begin{aligned}
&\mathrm{var}\big(f(X)\big) \\
&\quad = E\big(f(X)^2\big) - E\big(f(X)\big)^2 \\
&\quad = \sum_x f(x)^2\, P(X = x) - \left(\sum_x f(x)\, P(X = x)\right)^2 \\
&\quad = \sum_x E(Y \mid X = x)^2\, P(X = x) - \left(\sum_x E(Y \mid X = x)\, P(X = x)\right)^2 \\
&\quad = \sum_x E(Y \mid X = x)^2\, P(X = x) - E(Y)^2
\end{aligned}$$

Skupaj je

$$E\big(g(X)\big) + \mathrm{var}\big(f(X)\big) = E(Y^2) - E(Y)^2 = \mathrm{var}(Y) = a\,.$$

10. Naj bo $Z_0, Z_1, \ldots$ proces razvejanja. Predpostavite, da je $E(Z_n^2) < \infty$ za vse $n \geq 0$. Utemeljite zvezo

$$\operatorname{var}(Z_{m+n}) = E(Z_m)\operatorname{var}(Z_n) + E(Z_n)^2 \operatorname{var}(Z_m)$$

Rešitev: označimo rodovno funkcijo slučajne spremenljivke Z_n z G_n. Ker je $G_n = G \circ G \circ \cdots \circ G$ in je kompozitum asociativen, sledi $G_{m+n}(s) = G_m(G_n(s))$. Odvajajmo, pošljemo $s \uparrow 1$ in upoštevamo, da velja tudi $G_n(s) \uparrow 1$. Dobimo

$$E(Z_{m+n}) = \lim_{s \to 1} G'_n(G_m(s))\, G'_m(s) = E(Z_n)\, E(Z_m)$$

in

$$E\big(Z_{m+n}(Z_{m+n} - 1)\big) = \lim_{s \to 1} \left(G''_m\big(G_n(s)\big) \left(G'_n(s)\right)^2 + G'_m\big(G_n(s)\big)\, G''_n(s) \right).$$

Limito prepišemo v

$$E\big(Z_{m+n}(Z_{m+n} - 1)\big) = E\big(Z_m(Z_m - 1)\big)\, E(Z_n)^2 + E(Z_m)\, E\big(Z_n(Z_n - 1)\big).$$

Iz tega sledi

$$\begin{aligned}
\operatorname{var}(Z_{m+n}) &= E(Z_{m+n}^2) - E(Z_{m+n})^2 \\
&= \big(E(Z_m^2) - E(Z_m)\big)\, E(Z_n)^2 + E(Z_m)\big(E(Z_n^2) - E(Z_n)\big) \\
&\qquad + E(Z_m)\, E(Z_n) - E(Z_m)^2\, E(Z_n)^2 \\
&= E(Z_m)\big(E(Z_n^2) - E(Z_n)^2\big) + E(Z_n)^2 \big(E(Z_m^2) - E(Z_m)^2\big) \\
&= E(Z_m)\operatorname{var}(Z_n) + E(Z_n)^2 \operatorname{var}(Z_m).
\end{aligned}$$