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

Univerza na Primorskem

FAMNIT (MA, MF)

Verjetnost

Računski izpit

26. junij 2023

Navodila

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

Question	a.	b.	c.	d.	Total
1.			●	●	
2.			●	●	
3.			●	●	
4.				●	
5.			●	●	
6.			●	●	
Total					

1. (20) V r škatel vržemo n kroglic. Meti so med seboj neodvisni in vsaka kroglica pristane v posamezni škatli z verjetnostjo $1/r$. Privzemimo, da je $n \geq 2r$. Za $i = 1, 2, \ldots, r$ naj bo A_i dogodek, da i-ta škatla vsebuje natanko dve kroglici.

a. (10) Za $i \leq r$ izračunajte verjetnost $P(A_1 \cap A_2 \cap \cdots \cap A_i)$.

Rešitev: najprej izberemo pare metov, pri katerih kroglica pristane v škatlah $1, 2, \ldots, i$. To lahko naredimo na

$$\binom{n}{2}\binom{n-2}{2}\cdots\binom{n-2(i-1)}{2} = \frac{n!}{2^i \cdot (n-2i)!}$$

načinov, pri čemer se držimo dogovora $0! = 1$. Preostalih $n - 2i$ kroglic pa mora pristati v preostalih $r - i$ škatlah. Zaradi neodvisnosti je

$$P(A_1 \cap A_2 \cap \cdots \cap A_i) = \frac{n!}{2^i \cdot (n-2i)!} \cdot \left(\frac{1}{r}\right)^{2i} \cdot \left(\frac{r-i}{r}\right)^{n-2i},$$

pri čemer se dogovorimo, da je $0^0 = 1$.

b. (10) Kolikšna je verjetnost, da nobena škatla ne bo vsebovala natančno dveh kroglic? Vsot in binomskih simbolov vam ni treba poenostavljati.

Rešitev: dogodek, da vsaj ena škatla vsebuje natanko dve kroglici, je $\cup_{i=1}^r A_i$. Po formuli za vključitve in izključitve je

$$P\left(\bigcup_{i=1}^r A_i\right) = \sum_{i=1}^r (-1)^{i-1}\binom{r}{i}\frac{n!}{2^i \cdot (n-2i)!} \cdot \left(\frac{1}{r}\right)^{2i} \cdot \left(\frac{r-i}{r}\right)^{n-2i},$$

iskana verjetnost pa je

$$1 - P\left(\bigcup_{i=1}^r A_i\right) = \sum_{i=0}^r (-1)^{i}\binom{r}{i}\frac{n!}{2^i \cdot (n-2i)!} \cdot \left(\frac{1}{r}\right)^{2i} \cdot \left(\frac{r-i}{r}\right)^{n-2i}.$$

2. (20) Iz posode, v kateri je sprva a belih in $b \geq 2$ črnih kroglic, na slepo vlečemo kroglice. Vsakič, ko izvlečemo kroglico, jo ne glede na njeno barvo zamenjamo z belo kroglico. Vlečenja so med seboj neodvisna. Naj bo X število izvlečenih kroglic do vključno prve črne, Y pa naj bo število kroglic med prvo in drugo izvlečeno črno, vključno z drugo, ne pa tudi s prvo črno kroglico.

a. (10) Poiščite skupno porazdelitev slučajnih spremenljivk X in Y.

Rešitev: možne vrednosti slučajnega vektorja (X, Y) so vsi celoštevilski pari (k, l) s $k, l \geq 1$. Dogodek $\{X = k, Y = l\}$ se zgodi, če najprej izvlečemo $k - 1$ belih kroglic, nato črno kroglico, nato $l - 1$ belih kroglic in nazadnje črno kroglico. Torej velja

$$P(X = k, Y = l) = \left(\frac{a}{a+b}\right)^{k-1} \cdot \frac{b}{a+b} \cdot \left(\frac{a+1}{a+b}\right)^{l-1} \cdot \frac{b-1}{a+b},$$

kar se poenostavi v

$$P(X = k, Y = l) = \frac{a^{k-1}(a+1)^{l-1}b(b-1)}{(a+b)^{k+l}}.$$

Z drugimi besedami, X in Y sta neodvisni z $X \sim \mathrm{Geom}\left(\frac{b}{a+b}\right)$ in $Y \sim \mathrm{Geom}\left(\frac{b+1}{a+b}\right)$.

b. (10) Poiščite porazdelitev slučajne spremenljivke $Z = X + Y$.

Rešitev: za $n \geq 2$ izračunamo

$$\begin{aligned}
P(Z = n) &= \sum_{k=1}^{n-1} P(X = k, Y = n - k) \\
&= \frac{b(b-1)}{(a+b)^n} \sum_{k=1}^{n-1} a^{k-1}(a+1)^{n-k-1} \\
&= \frac{b(b-1)(a+1)^{n-2}}{(a+b)^n} \sum_{k=1}^{n-1} \left(\frac{a}{a+1}\right)^{k-1} \\
&= \frac{b(b-1)(a+1)^{n-2}}{(a+b)^n} \cdot \frac{1 - \left(\frac{a}{a+1}\right)^{n-1}}{1 - \frac{a}{a+1}} \\
&= \frac{b(b-1)(a+1)^{n-2}}{(a+b)^n} \cdot \frac{(a+1)^{n-1} - a^{n-1}}{(a+1)^{n-2}} \\
&= \frac{b(b-1)}{(a+b)^n} \left((a+1)^{n-1} - a^{n-1}\right).
\end{aligned}$$

3. (20) Naj bosta X in Y neodvisni slučajni spremenljivki z

$$X \sim \Gamma(a, 1) \quad \text{in} \quad Y \sim \Gamma\left(a + \frac{1}{2}\,,\, 1\right).$$

Definirajmo

$$(U, V) = \left(2\sqrt{\frac{Y}{X}}\,,\, 2\sqrt{XY}\right).$$

a. (10) Izračunajte gostoto slučajnega vektorja (U, V).

Rešitev: preslikava

$$\Phi\left(2\sqrt{\frac{y}{x}}\,,\, 2\sqrt{xy}\right)$$

je bijektivna na $(0, \infty)^2$ *z inverzom*

$$\Phi^{-1}(u, v) = \left(\frac{v}{u}\,,\, \frac{uv}{4}\right).$$

Poleg tega imata preslikavi Φ *and* Φ^{-1} *obe zvezne parcialne odvode. Izračunajmo*

$$J_{\Phi^{-1}}(u, v) = \det\begin{pmatrix} -\frac{v}{u^2} & \frac{1}{u} \\ \frac{1}{4}v & \frac{1}{4}u \end{pmatrix} = -\frac{v}{2u}\,.$$

Po transformacijski formuli je

$$f_{U,V}(u, v) = \frac{1}{\Gamma(a)\,\Gamma\left(a + \frac{1}{2}\right)} \left(\frac{v}{u}\right)^{a-1} e^{-\frac{v}{u}} \left(\frac{uv}{4}\right)^{a-\frac{1}{2}} e^{-\frac{uv}{4}} \cdot \frac{v}{2u}$$

za vse $u, v > 0$*; drugje lahko postavimo* $f_{U,V}(u, v) = 0$*. Gostota se poenostavi v*

$$f_{U,V}(u, v) = \frac{1}{4^a\,\Gamma(a)\,\Gamma\left(a + \frac{1}{2}\right)}\, u^{-\frac{1}{2}} \cdot v^{2a-\frac{1}{2}} \cdot e^{-\frac{uv}{4} - \frac{v}{u}}\,.$$

b. (10) Določite porazdelitev slučajne spremenljivke V, pri čemer jo eksplicitno poimenujte.

Namig: kot znano lahko privzamete, da za poljubna $\alpha, \beta > 0$ *velja*

$$\int_0^\infty \frac{1}{\sqrt{s}}\, e^{-\alpha s - \frac{\beta}{s}}\, ds = \sqrt{\frac{\pi}{\alpha}}\, e^{-2\sqrt{\alpha\beta}}\,.$$

Rešitev: za $v > 0$ *izračunajmo*

$$\begin{aligned} f_V(v) &= \frac{v^{2a-\frac{1}{2}}}{4^a\,\Gamma(a)\,\Gamma\left(a + \frac{1}{2}\right)} \int_0^\infty \frac{1}{\sqrt{u}}\, e^{-\frac{uv}{4} - \frac{v}{u}}\, du \\ &= \frac{v^{2a-\frac{1}{2}}}{4^a\,\Gamma(a)\,\Gamma\left(a + \frac{1}{2}\right)} \cdot \sqrt{\frac{4\pi}{v}}\, e^{-v} \\ &= \frac{\sqrt{\pi}}{2^{2a-1}\,\Gamma(a)\,\Gamma\left(a + \frac{1}{2}\right)} \cdot v^{2a-1} e^{-v}\,, \end{aligned}$$

medtem ko lahko za $v < 0$ *seveda postavimo* $f_V(v) = 0$. *Konstantni faktor lahko ignoriramo in razberemo, da je* $V \sim \Gamma(2a, 1)$.

Opomba: s primerjavo konstant pri gostoti pa dobimo Legendrovo podvojitveno formulo

$$\Gamma(2a) = \frac{1}{\sqrt{\pi}}\, 2^{2a-1}\, \Gamma(a)\, \Gamma\left(a + \frac{1}{2}\right).$$

4. (20) Iz posode, v kateri je sprva a belih in b črnih kroglic, na slepo vlečemo kroglice. Vsakič, ko izvlečemo kroglico, jo ne glede na njeno barvo zamenjamo z belo kroglico. Vlečenja so med seboj neodvisna. Naj bo $X_{a,b}$ število izvlečenih kroglic do vključno zadnje črne. Označimo $e_{a,b} = E(X_{a,b})$ in $v_{a,b} = \mathrm{var}(X_{a,b})$.

a. (5) Naj bo Z število izvlečenih kroglic do vključno prve črne. Dokažite, da sta Z in $X_{a,b} - Z$ neodvisni ter da ima $X_{a,b} - Z$ isto porazdelitev kot $X_{a+1,b-1}$.

Rešitev: ko izvlečemo prvo črno kroglico, v posodi ostane še $a + 1$ belih in $b - 1$ črnih kroglic. Na dogodku $\{Z = k\}$ je $X_{a,b} - Z = X_{a,b} - k$ število preostalih izvlečenih kroglic do vključno prve črne. Zaradi neodvisnosti je se pogojna porazdelitev slučajne spremenljivke $X_{a,b}-Z$ glede na $\{X = k\}$ ujema s porazdelitvijo slučajne spremenljivke $X_{a+1,b-1}$, ne glede na k. To pa pomeni, da sta $X_{a,b} - Z$ in Z neodvisni, $X_{a,b} - Z$ pa ima isto porazdelitev kot $X_{a+1,b-1}$.

b. (10) Izračunajte $e_{a,b}$. Dovolj je, da rešitev zapišete kot vsoto.

Rešitev: pišimo

$$e_{a,b} = E(X_{a,b}) = E(Z) + E(X_{a,b} - Z)\,.$$

Ker je $Z \sim \mathrm{Geom}(b/(a + b))$, velja

$$E(Z) = \frac{a+b}{b}\,,$$

drugo pričakovano vrednost pa lahko dobimo iz točke a. Zgornja formula tako dobi obliko

$$e_{a,b} = \frac{a+b}{b} + e_{a+1,b-1}\,.$$

Z iteriranjem dobimo

$$e_{a,b} = \frac{a+b}{b} + \frac{a+b}{b-1} + \cdots + \frac{a+b}{2} + e_{a+b-1,1}\,.$$

Velja

$$X_{a+b-1,1} \sim \mathrm{Geom}\left(\frac{1}{a+b}\right)$$

in zato

$$e_{a+b-1,1} = E(X_{a+b-1,1}) = a + b\,.$$

Končno je

$$e_{a,b} = (a+b)\sum_{k=1}^{b} \frac{1}{k}\,.$$

c. (5) Naj bo $v_{a,b} = \mathrm{var}(X_{a,b})$. Pokažite, da je

$$v_{a,b} = \frac{a(a+b)}{b^2} + v_{a+1,b-1}\,,$$

in izračunajte $v_{a,b}$. Ponovno je dovolj, da rešite zapišete kot vsoto.

Rešitev: ker je varianca vsote neodvisnih slučajnih spremenljivk vsota njihovih varianc, velja

$$v_{a,b} = \mathrm{var}(Z) + \mathrm{var}(X_{a,b} - Z) = \frac{a(a+b)}{b^2} + v_{a+1,b-1}\,.$$

Z iteriranjem dobimo

$$v_{a,b} = \frac{a(a+b)}{b^2} + \frac{(a+1)(a+b)}{(b-1)^2} + \cdots + \frac{(a+b-2)(a+b)}{2^2} + v_{a+b-1,1}\,.$$

Spomnimo se, kako je porazdeljena slučajna spremenljivka $X_{a+b-1,1}$, kar nam da

$$v_{a+b-1,1} = \mathrm{var}(X_{a+b-1,1}) = (a+b)(a+b-1)\,.$$

Končno je

$$v_{a,b} = (a+b)\sum_{k=1}^{b} \frac{a+b-k}{k^2} = (a+b)^2 \sum_{k=1}^{b} \frac{1}{k^2} - (a+b)\sum_{k=1}^{b} \frac{1}{k}\,.$$

5. (20) Naj bo $Z_0, Z_1, \ldots$ proces razvejanja. Slučajno število Y potomcev vsakega posameznika naj ima porazdelitev

$$P(Y = k) = 2^{-(k+1)}$$

za $k = 0, 1, \ldots$

a. (10) Z matematično indukcijo pokažite, da je rodovna funkcija $G_n(s)$ spremenljivke Z_n enaka

$$G_n(s) = \frac{n - (n-1)s}{n + 1 - ns} .$$

Rešitev: Najprej izračunamo

$$G(s) = G_1(s) = \sum_{k=0}^{\infty} 2^{-(k+1)} s^k = \frac{1}{2 - s} ,$$

torej trditev drži za $n = 1$. Predpostavimo zdaj, da trditev drži za neki n. Velja

$$\begin{aligned} G_{n+1}(s) &= G_n(G_1(s)) \\ &= \frac{n - (n-1)G(s)}{n + 1 - nG(s)} \\ &= \frac{n - (n-1)\frac{1}{2-s}}{n + 1 - n\frac{1}{2-s}} \\ &= \frac{n(2-s) - (n-1)}{(n+1)(2-s) - n} \\ &= \frac{(n+1) - ns}{n + 2 - (n+1)s} , \end{aligned}$$

s čimer je indukcijski korak zaključen.

b. (10) Izračunajte $E(Y)$, $P(Z_n = 0)$ in $\eta = P(\text{proces izumre})$. Kako se to ujema s teorijo?

Rešitev: Vemo, da je $E(Y) = G'(1)$. Velja

$$G'(s) = \frac{1}{(2-s)^2} ,$$

torej je $E(Y) = 1$. Vemo tudi, da je $P(Z_n = 0) = G_n(0)$, in iz točke a. dobimo

$$P(Z_n = 0) = \frac{n}{n+1} .$$

Od tod sledi

$$P(\textit{proces izumre}) = \lim_{n \to \infty} P(Z_n = 0) = 1 ,$$

kar lahko izračunamo tudi kot prvo rešitev enačbe $\frac{1}{2-s} = s$ na intervalu $[0, 1]$. Ta enačba pa se prevede na kvadratno enačbo $s^2 - 2s + 1 = 0$, ki ima edino rešitev $s = 1$.

Splošna teorija za primer, ko je $E(Y) = 1$ in $P(Z_1 = 0) > 0$, napoveduje, da bo rodbina izumrla z verjetnostjo 1; namesto da je $P(Z_1 = 0) > 0$, pa lahko opazimo tudi, da je $G''(1) > 0$. Na ta način se zgornji rezultat ujema s splošno teorijo.

6. (20) Zdolgočaseni statistik je n-krat z vračanjem izbiral lističe iz spodnjih škatel. Števila na izbranih lističih je označil z $X_1, X_2, \ldots, X_n$, njihovo vsoto pa z S_n.

(i) $\boxed{-1}\ \boxed{0}\ \boxed{1}$

(ii) $\boxed{-1}\ \boxed{0}\ \boxed{0}\ \boxed{0}\ \boxed{0}\ \boxed{0}\ \boxed{0}\ \boxed{0}\ \boxed{1}$

a. (10) Statistik je izračunal

$$P(-30 \le S_{1000} \le 30) \approx 0{,}96\,.$$

Za katero od škatel je računal verjetnosti? Utemeljite odgovor.

Rešitev: Za prvo škatlo dobimo $\operatorname{var}(X_1) = 2/3$, *za drugo pa* $\operatorname{var}(X_1) = 2/9$. *Če uporabimo centralni limitni izrek in vzamemo* $Z \sim N(0,1)$, *za prvo škatlo izračunamo*

$$\begin{aligned}
P(-30 \le S_{1000} \le 30) &= P\left(-\frac{30}{\sqrt{1000}\sqrt{2/3}} \le \frac{S_{1000}}{\sqrt{\operatorname{var}(S_{1000})}} \le \frac{30}{\sqrt{1000}\sqrt{2/3}}\right) \\
&= P\left(-1{,}16 \le \frac{S_{1000}}{\sqrt{\operatorname{var}(S_{1000})}} \le 1{,}16\right) \\
\textit{(CLI)} \qquad &\approx P\big(-1{,}16 \le Z \le 1{,}16\big) \\
&\doteq 0{,}75\,.
\end{aligned}$$

Za drugo škatlo računamo podobno. Dobimo

$$\begin{aligned}
P\left(-30 \le S_{1000} \le 30\right) &= P\left(-\frac{30}{\sqrt{1000}\sqrt{2/9}} \le \frac{S_{1000}}{\sqrt{\operatorname{var}(S_{1000})}} \le \frac{30}{\sqrt{1000}\sqrt{2/9}}\right) \\
&= P\left(-2{,}01 \le \frac{S_{1000}}{\sqrt{\operatorname{var}(S_{1000})}} \le 2{,}01\right) \\
\textit{(CLI)} \qquad &\approx P\big(-2{,}01 \le Z \le 2{,}01\big) \\
&\doteq 0{,}96\,.
\end{aligned}$$

Statistik je izbiral iz druge, večje škatle.

b. (10) Statistik je izračunal $P(S_{100} = 0) \approx 0{,}049$. Katero škatlo je obravnaval?

Rešitev: Spet uporabimo centralni limitni izrek. Za prvo škatlo lahko približno ocenimo

$$\begin{aligned}
P(S_{100} = 0) &= P\left(-\frac{1}{2} \le S_{100} \le \frac{1}{2}\right) \\
&= P\left(-\frac{1}{20\sqrt{2/3}} \le \frac{S_{100}}{\sqrt{\operatorname{var}(S_{100})}} \le \frac{1}{20\sqrt{2/3}}\right) \\
\textit{(CLI)} \qquad &\approx P\big(-0{,}061 \le Z \le 0{,}061\big) \\
&\doteq 0{,}049\,.
\end{aligned}$$

S povsem podobnim postopkom izračunamo, da za večjo škatlo velja $P(S_{100} = 0) \approx 0{,}085$. *Statistik je obravnaval prvo škatlo.*