Univerza na Primorskem

FAMNIT, Oddelek za matematiko

Verjetnost

2. kolokvij

29. maj 2019

Ime in priimek: ______________________ Vpisna št: ☐☐☐☐☐☐☐☐

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 in matematični priročnik.

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

1. (20) V kupu 52 standardnih kart so 4 asi, 4 kralji in 4 dame. Karte dobro premešamo in jih začnemo deliti z vrha, dokler ne dobimo prvega asa. Naj bo X število kraljev in Y število dam, ki smo jih razdelili do prvega asa.

a. (10) Določite porazdelitev slučajnega vektorja (X, Y).

Namig: ali je treba gledati vseh 52 kart?

Rešitev: Slučajni vektor (X, Y) ima za vrednosti vse pare (k, l) z $0 \leq k, l \leq 4$. Zanimajo nas le relativni položaji asov, kraljev in dam. Če iz dobro premešanega kupa kart izločimo vse ostale karte, nam ostane naključna permutacija teh 12 kart. Za izračun verjetnosti $P(X = k, Y = l)$ moramo prešteti vse permutacije, ki imajo pred prvim asom k kraljev in l dam. Izbiramo najprej k kraljev, kar lahko naredimo na $\binom{4}{k}$ načinov, in l dam, kar lahko naredimo na $\binom{4}{l}$ načinov. Teh $k + l$ kart lahko poljubno permutiramo, za njimi pride eden od štirih asov, ostalih $12 - (k + l + 1)$ kart pa lahko spet poljubno permutiramo. Izmed vseh $12!$ permutacij je torej ugodnih

$$\binom{4}{k}\binom{4}{l} \cdot (k + l)! \cdot 4 \cdot (11 - k - l)!\,.$$

Sledi

$$\begin{aligned} P(X = k, Y = l) &= \frac{4\binom{4}{k}\binom{4}{l}\,(k + l)!\,(11 - k - l)!}{12!} \\ &= \frac{4 \cdot (4!)^2\,(k + l)!\,(11 - k - l)!}{12!\,k!\,l!\,(4 - k)!\,(4 - l)!}\,. \end{aligned}$$

b. (5) Sta X in Y neodvisni?

Rešitev: Za neodvisnost bi moralo za $k, l \in \{0, 1, 2, 3, 4\}$ na zalogi vrednosti (X, Y) veljati

$$P(X = k, Y = l) = f(k)\,g(l)$$

za neki funkciji $f, g\colon \{0, 1, 2, 3, 4\} \to \mathbb{R}$. Tedaj bi za $k, l \in \{0, 1, 2, 3, 4\}$ moralo veljati tudi

$$H(k, l) := (k + l)!\,(11 - k - l)! = \tilde{f}(k)\,\tilde{g}(l)\,,$$

potem pa bi morala biti npr. tudi matrika

$$\begin{bmatrix} H(0,0) & H(0,1) \\ H(1,0) & H(1,1) \end{bmatrix} = \begin{bmatrix} 11! & 10! \\ 10! & 2 \cdot 9! \end{bmatrix} = 9! \begin{bmatrix} 110 & 10 \\ 10 & 2 \end{bmatrix}$$

izrojena, to pa ni res. Slučajni spremenljivki X in Y sta torej odvisni.

c. (5) Določite porazdelitev vsote $X + Y$.

Rešitev: Razmišljamo kot v prvem delu naloge, le da preprosto združimo kralje in dame. Vseh permutacij je spet $12!$*, takih, kjer je* $X + Y = n$*, pa je*

$$\binom{8}{n} \cdot n! \cdot 4 \cdot (12 - n - 1)!\,,$$

torej za $n = 0, 1, \ldots, 8$ *velja*

$$P(X + Y = n) = \frac{4\binom{8}{n}\, n!\,(11 - n)!}{12!} = \frac{(11 - n)(10 - n)(9 - n)}{2970}\,.$$

2. (20) Pri *Poker testu* za kontrolo generatorjev slučajnih števil nastopi naslednji problem: imamo neodvisne, enako porazdeljene slučajne spremenljivke $\xi_1, \xi_2, \xi_3, \xi_4, \xi_5$ z enakomerno porazdelitvijo na množici $\{0, 1, \ldots, m-1\}$ za dan $m > 0$. Naj bo X število različnih števil v množici $\{\xi_1, \xi_2, \xi_3, \xi_4, \xi_5\}$. Primer: za množico $\{1, 2, 5, 2, 5\}$ dobimo $X = 3$.

a. (10) Izračunajte $E(X)$.

Namig: izrazite X z indikatorji

$$I_k = \begin{cases} 1 & \text{če se število } k \text{ pojavi v množici } \{\xi_1, \xi_2, \xi_3, \xi_4, \xi_5\} \\ 0 & \text{sicer.} \end{cases}$$

Rešitev:

Prvi način: *Za $k = 0, 1, \ldots, m-1$ definirajmo indikatorje*

$$I_k = \begin{cases} 1 & \text{če se število } k \text{ pojavi v množici } \{\xi_1, \xi_2, \xi_3, \xi_4, \xi_5\} \\ 0 & \text{sicer.} \end{cases}$$

Zapišemo lahko $X = I_0 + \cdots + I_{m-1}$. Zaradi enakomerne porazdelitve imajo vsi indikatorji enako porazdelitev. Računamo

$$P(I_k = 0) = \left(\frac{m-1}{m}\right)^5, \qquad P(I_k = 1) = 1 - \left(\frac{m-1}{m}\right)^5.$$

Sledi

$$E(X) = m\left[1 - \left(\frac{m-1}{m}\right)^5\right].$$

Drugi način: *določimo porazdelitev slučajne spremenljivke X. Velja namreč*

$$P(X = k) = S(5, k)\,\frac{m(m-1)\cdots(m-k+1)}{m^5},$$

kjer je $S(n, k)$ Stirlingovo število druge vrste, t. j. število neurejenih razbitij n-elementne množice na k nepraznih množic. Velja $S(5, 1) = 1$, $S(5, 2) = 15$, $S(5, 3) = 25$, $S(5, 4) = 10$ in $S(5, 5) = 1$. Sledi:

$$\begin{aligned} E(X) &= \sum_{k=1}^{5} k\,P(X = k) \\ &= \frac{1}{m^4}\Big(1 + 30(m-1) + 75(m-1)(m-2) + 40(m-1)(m-2)(m-3) \\ &\qquad\qquad + 5(m-1)(m-2)(m-3)(m-4)\Big) \\ &= \frac{5m^4 - 50m^3 + 175m^2 - 250m + 120}{m^4}. \end{aligned}$$

Krajši račun pokaže, da pride isto kot pri prvem načinu.

b. (10) Izračunajte $\mathrm{var}(X)$.

Namig: uporabite $P(I_k = 1, I_l = 1) = 1 - P\big(\{I_k = 0\} \cup \{I_l = 0\}\big)$.

Rešitev:

Prvi način: *uporabimo* $\mathrm{var}(X) = E(X^2) - \big(E(X)\big)^2$ *in računamo*

$$\begin{aligned} E(X^2) &= \sum_{k=0}^{m-1} E(I_k^2) + \sum_{\substack{0 \le k,l < m \\ k \ne l}} E(I_k I_l) \\ &= \sum_{k=0}^{m-1} P(I_k = 1) + \sum_{\substack{0 \le k,l < m \\ k \ne l}} P(I_k = 1, I_l = 1)\,. \end{aligned}$$

Pri drugem členu uporabimo namig:

$$\begin{aligned} P(I_k = 1, I_l = 1) &= 1 - P\big(\{I_k = 0\} \cup \{I_l = 0\}\big) \\ &= 1 - P(I_k = 0) - P(I_l = 0) + P(I_k = 0, I_l = 0)\,. \end{aligned}$$

Podobno kot pri točki a. za $k \ne l$ *izračunamo*

$$P(I_k = 0, I_l = 0) = \left(\frac{m-2}{m}\right)^5 .$$

Sledi

$$P(I_k = 1, I_l = 1) = 1 - 2\left(\frac{m-1}{m}\right)^5 + \left(\frac{m-2}{m}\right)^5 .$$

Poberemo skupaj in z uporabo rezultata iz prejšnje točke dobimo

$$\begin{aligned} \mathrm{var}(X) &= m\left[1 - \left(\frac{m-1}{m}\right)^5\right] + m(m-1)\left[1 - 2\left(\frac{m-1}{m}\right)^5 + \left(\frac{m-2}{m}\right)^5\right] \\ &\quad - m^2\left[1 - 2\left(\frac{m-1}{m}\right)^5 + \left(\frac{m-1}{m}\right)^{10}\right] \\ &= m\left(\frac{m-1}{m}\right)^5 + m(m-1)\left(\frac{m-2}{m}\right)^5 - m^2\left(\frac{m-1}{m}\right)^{10} . \end{aligned}$$

Drugi način: *varianco zapišemo kot vsoto varianc in kovarianc:*

$$\mathrm{var}(X) = \sum_{k=0}^{m-1} \mathrm{var}(I_k) + \sum_{\substack{0 \le k,l < m \\ k \ne l}} \mathrm{cov}(I_k, I_l)\,.$$

Zaradi simetrije so enake vse variance in vse kovariance. Vemo, da je

$$\operatorname{var}(I_k) = P(I_k = 0)\,P(I_k = 1) = \left(\frac{m-1}{m}\right)^5 \left[1 - \left(\frac{m-1}{m}\right)^5\right]$$

in

$$\operatorname{cov}(I_k, I_l) = P(I_k = 1, I_l = 1) - P(I_k = 1)\,P(I_l = 1)\,.$$

Tako kot pri prvem načinu za $k \neq l$ izračunamo

$$P(I_k = 1, I_l = 1) = 1 - 2\left(\frac{m-1}{m}\right)^5 + \left(\frac{m-2}{m}\right)^5\,,$$

torej je

$$\begin{aligned}\operatorname{cov}(I_k, I_l) &= P(I_k = 1, I_l = 1) - P(I_k = 1)\,P(I_l = 1)\\ &= \left(\frac{m-2}{m}\right)^5 - \left(\frac{m-1}{m}\right)^{10}.\end{aligned}$$

Poberemo skupaj in dobimo

$$\begin{aligned}\operatorname{var}(X) &= m\left(\frac{m-1}{m}\right)^5 \left[1 - \left(\frac{m-1}{m}\right)^5\right]\\ &\qquad + m(m-1)\left[\left(\frac{m-2}{m}\right)^5 - \left(\frac{m-1}{m}\right)^{10}\right]\\ &= m\left(\frac{m-1}{m}\right)^5 + m(m-1)\left(\frac{m-2}{m}\right)^5 - m^2\left(\frac{m-1}{m}\right)^{10},\end{aligned}$$

kar je isto kot pri prvem načinu.

Tretji način*: uporabimo eksplicitno porazdelitev za izračun:*

$$\begin{aligned}E(X^2) &= \sum_{k=1}^{5} k^2\,P(X = k)\\ &= \frac{1}{m^4}\Big(1 + 60(m-1) + 225(m-1)(m-2) + 160(m-1)(m-2)(m-3)\\ &\qquad + 25(m-1)(m-2)(m-3)(m-4)\Big)\\ &= \frac{25m^4 - 90m^3 + 140m^2 - 105m + 31}{m^4}\,.\end{aligned}$$

Krajši račun pokaže, da pride isto kot pri prvem načinu. Tako kot pri slednjem tudi nadaljujemo.

3. (20) Naj bosta U in V neodvisni in $U, V \sim \exp(1)$. Definirajte $X = U - V$.

a. (10) Naj bo $X = U - V$. Izračunajte gosototo X.

Rešitev: Definirajmo

$$\Phi(u, v) = (u - v, v) .$$

Iz

$$\Phi^{-1}(x, v) = (x + v, v)$$

sledi, da je

$$J_{\Phi^{-1}}(u, x) = 1 .$$

Torej je

$$f_{X,V}(x, v) = f_U(x + v) f_V(v) .$$

Zaradi simetrije bo $f_X(x)$ soda funkcija. Predpostavimo $x \geq 0$ in računamo

$$\begin{aligned} f_X(x) &= \int_0^\infty e^{-x-v} e^{-v} dv \\ &= e^{-x} \int_0^\infty e^{-2v} dv \\ &= \frac{1}{2} e^{-x} . \end{aligned}$$

Sledi

$$f_X(x) = \frac{1}{2} e^{-|x|}$$

za $x \in \mathbb{R}$.

b. (10) Naj bosta X in Y neodvisni in naj imata obe enako porazdelitev kot $U - V$. Izračunajte porazdelitev $Z = X - Y$.

Rešitev: Naj bodo $U, V, \tilde{U}, \tilde{V}$ enako porazdeljene in neodvisne. Naj bo $X = U - V$ in $Y = \tilde{U} - \tilde{V}$. Zapišemo lahko

$$X - Y = \left(U + \tilde{V}\right) - \left(V + \tilde{U}\right) .$$

Vemo, da je $U + \tilde{V} \sim \Gamma(2, 1)$ in enako za vsoto v drugem oklepaju. Zaradi *simetrije bo gosotota $f_Z(z)$ soda funkcija. Računamo kot v prvem delu za $z \geq 0$*

$$\begin{aligned} f_Z(z) &= \int_0^\infty (z+x)e^{-z-x} x e^{-x} dx \\ &= z e^{-z} \int_0^\infty x e^{-2x} dx + e^{-z} \int_0^\infty x^2 e^{-2x} dx \\ &= \frac{1}{4} z e^{-z} + \frac{1}{4} e^{-z} \\ &= \frac{1}{4}(z+1)e^{-z}\,. \end{aligned}$$

Končno sledi

$$f_Z(z) = \frac{1}{4}(|z|+1)e^{-|z|}$$

za $z \in \mathbb{R}$.

4. (20) Predpostavljajte, da so meti kovanca med sabo neodvisni, vsakič pa je verjetnost za grb enaka p. Označimo $q := 1 - p$. Naj bo W_n število potrebnih metov vključno z zadnjim, dokler prvič ne bomo videli n zaporednih grbov.

a. (10) Utemeljite, da velja

$$P\left(W_n = k+1 | W_{n-1} = k\right) = p$$

in za $l > k+1$

$$P(W_n = l | W_{n-1} = k) = qP(W_n = l - k - 1)\,.$$

Rešitev: Če na k-tem metu prvič dobimo $n-1$ grbov zapovrstjo, se lahko zgodi dvoje: na naslednjem metu dobimo grb in je $W_n = k+1$ ali na naslednjem metu dobimo številko in se "čakanje" na n grbov zapovrstjo začne znova. Zgornji enačbi sta matematični zapis tega dejstva.

b. (10) Pokažite, da je

$$E\left(W_n | W_{n-1} = k\right) = k + 1 + qE(W_n)$$

in izračunajte $E(W_n)$.

Rešitev: Iz pogojnih verjetnosti sledi, da je

$$\begin{aligned}
E\left(W_n | W_{n-1} = k\right) &= p(k+1) + \sum_{l=k+1+n}^{\infty} qlP(W_n = l - k - 1) \\
&= p(k+1) + q\sum_{m=n}^{\infty}(m+k+1)P(W_n = m) \\
&= p(k+1) + q(k+1) + qE(W_n) \\
&= k + 1 + qE(W_n)\,.
\end{aligned}$$

Obe strani zadnje enačbe pomnožimo z $P(W_{n-1} = k)$ in seštejemo. Sledi

$$\begin{aligned}
E(W_n) &= \sum_{k=n-1}^{\infty} E(W_n | W_{n-1} = k)P(W_{n-1} = k) \\
&= \sum_{k=n-1}^{\infty} \left(k + 1 + qE(W_n)\right) P(W_{n-1} = k) \\
&= E(W_{n-1}) + 1 + qE(W_n)\,.
\end{aligned}$$

Sledi

$$E(W_n) = \frac{1}{p} + \frac{1}{p}E(W_{n-1})\,.$$

Če upoštevamo $E(W_1) = p^{-1}$*, sledi*

$$E(W_n) = \sum_{k=1}^{n} \frac{1}{p^k} = \frac{1}{p^n} \, \frac{1-p^n}{1-p} \, .$$

5. (20) Dan je proces razvejanja, pri katerem ima vsak posameznik z verjetnostjo 1/4 dva potomca, z verjetnostjo 3/4 pa je brez potomcev. Števila potomcev vseh posameznikov so neodvisna.

a. (5) Dokažite, da tak proces skoraj gotovo izumre.

Rešitev: Rodovna funkcija števila potomcev je $G(s) = \frac{3}{4} + \frac{1}{4}s^2$ *in enačba* $G(s) = s$ *ima dve rešitvi,* $s = 1$ *in* $s = 3$. *Število* 1 *je torej edina negibna točka funkcije* G *na intervalu* $[0, 1]$, *torej je to tudi verjetnost, da proces izumre.*

b. (5) Naj bo Z_1 število potomcev v prvi generaciji, N naj bo število *vseh* predstavnikov tega procesa, H pa naj bo rodovna funkcija tega števila. Za $k \in \{0, 2\}$ izrazite $E(s^N \mid Z_1 = k)$ s $H(s)$.

Rešitev: Če je $Z_1 = 0$, *je* $N = 1$, *zato je* $E(s^N \mid Z_1 = 0) = s$. *Pogojno na* $Z_1 = 2$ *pa je* N *porazdeljena enako kot skupno število predstavnikov dveh neodvisnih procesov, porazdeljenih enako kot dani proces brezpogojno, povečano za 1. Torej je* $E(s^N \mid Z_1 = 2) = s\big(H(s)\big)^2$.

c. (10) Izračunajte porazdelitev števila vseh predstavnikov tega procesa, t. j. za vsak n iz zaloge vrednosti izračunajte $P(N = n)$.

Pomoč: $(1 + x)^m = \sum_{k=0}^{\infty} \binom{m}{k} x^k$.

Rešitev: Po formuli za popolno pričakovano vrednost je

$$\begin{aligned} H(s) &= E(s^N) \\ &= P(Z_1 = 0)\, E\big(s^N \mid Z_1 = 0\big) + P(Z_1 = 2)\, E\big(s^N \mid Z_1 = 2\big) \\ &= \frac{s}{4}\left[3 + \big(H(s)\big)^2\right]. \end{aligned}$$

Rešimo na $H(s)$ *in dobimo*

$$H(s) = \frac{2}{s}\left(1 \pm \sqrt{1 - \tfrac{3}{4}s^2}\,\right),$$

kar je rodovna funkcija neke nenegativne celoštevilske slučajne spremenljivke, le če vzamemo negativni koren. Torej je

$$H(s) = 2\sum_{k=1}^{\infty} (-1)^{k-1} \binom{1/2}{k} \frac{3^k}{4^k}\, s^{2k-1}.$$

Slučajna spremenljivka N *torej zavzame vrednosti le v množici pozitivnih lihih števil in za* $n = 2k - 1$ *je:*

$$P(N = n) = 2(-1)^{k-1}\binom{1/2}{k}\frac{3^k}{4^k} = -2\,\frac{3^k}{4^k}\,\frac{\left(-\frac{1}{2}\right)_k}{k!} = \frac{3^k}{4^k}\,\frac{\left(\frac{1}{2}\right)_{k-1}}{k!} = 2\,\frac{3^k}{4^k}\,\frac{(n-2)!!}{(n+1)!!},$$

kjer je $(x)_r := x(x+1)(x+2)\cdots(x+r-1)$ *Pochhammerjev simbol.*

6. (20) V škatli so čudežne kroglice bele in rdeče barve. Prvih je B in drugih R. Iz škatle naključno izberemo n kroglic, tako da so vse izbire n kroglic enako verjetne. Ker pa so kroglice čudežne, takoj po izbiri in preden jih pogledamo, z verjetnostjo $1/4$ spremenijo barvo v drugo neodvisno ena od druge in neodvisno od postopka izbire. Naj bo M_1 število belih kroglic med izbranimi in M_2 število rdečih med izbranimi. Naj bo N_1 število belih po spreminjanju barv in N_2 število rdečih po spreminjanju barv.

a. (5) Izračunajte $E(N_1|M_1 = m_1, M_2 = m_2)$.

Rešitev: Bele kroglice naključno spreminjajo barvo. Če jih je m_1, bo pričakovano število belih po spreminjanju barv enako $3m_1/4$. Od m_2 črnih, pa bo belih postalo $m_2/4$. Sledi

$$E(N_1|M_1 = m_1, M_2 = m_2) = \frac{3m_1}{4} + \frac{m_2}{4}\,.$$

b. (5) Izračunajte $E(N_1N_2|M_1 = m_1, M_2 = m_2)$.

Namig: kaj je $\operatorname{cov}(N_1, N_2|M_1 = m_1, M_2 = m_2)$.

Rešitev: Ker čudežne kroglice spreminjajo barvo neodvisno ena od druge in neodvisno od postopka izbire, je

$$\begin{aligned}
&\operatorname{cov}(N_1, N_2|M_1 = m_1, M_2 = m_2)\\
&= E(N_1N_2|M_1 = m_1, M_2 = m_2)\\
&\qquad -E(N_1|M_1 = m_1, M_2 = m_2)E(N_2|M_1 = m_1, M_2 = m_2)\\
&= 0\,.
\end{aligned}$$

Sledi, da je

$$E(N_1N_2|M_1 = m_1, M_2 = m_2) = \left(\frac{3m_1}{4} + \frac{m_2}{4}\right)\left(\frac{m_1}{4} + \frac{3m_2}{4}\right)\,.$$

c. (10) Pokažite, da je

$$\operatorname{cov}(N_1, N_2) = -\frac{\operatorname{var}(M_1)}{4}\,.$$

Rešitev: Velja $M_1 \sim \mathrm{HiperGeom}(n, B, N)$.

$$\begin{aligned}
E(N_1) &= \sum_{\substack{k,l \\ k+l=n}} E(N_1 | M_1 = m_1, M_2 = m_2) P(M_1 = m_1, M_2 = m_2) \\
&= \sum_{\substack{k,l \\ k+l=n}} \left(\frac{3m_1}{4} + \frac{m_2}{4}\right) P(M_1 = m_1, M_2 = m_2) \\
&= \frac{3}{4} E(M_1) + \frac{1}{4} E(M_2) \\
&= \frac{3nB}{4N} + \frac{nR}{4N} \\
&= \frac{n(3B+R)}{4N}
\end{aligned}$$

Po drugi strani, upoštevajoč $M_1 + M_2 = n$, *velja*

$$\mathrm{var}(M_1) = \mathrm{var}(M_2) \qquad \textit{in} \qquad \mathrm{cov}(M_1, M_2) = -\mathrm{var}(M_1)\,.$$

Računamo

$$\begin{aligned}
&E(N_1 N_2) \\
&\quad = \sum_{\substack{k,l \\ k+l=n}} E(N_1 N_2 | M_1 = m_1, M_2 = m_2) P(M_1 = m_1, M_2 = m_2) \\
&\quad = \sum_{\substack{k,l \\ k+l=n}} \left(\frac{3m_1}{4} + \frac{m_2}{4}\right)\left(\frac{m_1}{4} + \frac{3m_2}{4}\right) P(M_1 = m_1, M_2 = m_2) \\
&\quad = \frac{1}{16} E\left(3M_1^2 + 3M_2^2 + 10 M_1 M_2\right) \\
&\quad = \frac{1}{16}\left(6\mathrm{var}(M_1) + 3E(M_1)^2 + 3E(M_2)^2 - 10\mathrm{var}(M_1) + 10E(M_1)E(M_2)\right) \\
&\quad = -\frac{1}{4}\mathrm{var}(M_1) + \frac{3n^2B^2}{16N^2} + \frac{3n^2R^2}{16N^2} + \frac{10n^2BR}{16N^2} \\
&\quad = -\frac{1}{4}\mathrm{var}(M_1) + \frac{n^2(3B+R)(B+3R)}{16N^2} \\
&\quad = -\frac{1}{4}\mathrm{var}(M_1) + E(N_1)E(N_2)\,.
\end{aligned}$$

Enakost sledi.