VJEROJATNOST I STATISTIKA

Prvi ispitni rok – 24. lipnja 2024.

Zadatak 1. (*12 bodova*)

- (3 boda) Definiraj matematičko očekivanje diskretne slučajne varijable.
- (3 boda) Definiraj pojam nezavisnosti dvije diskretne slučajne varijable.
- (6 bodova) Dokaži tvrdnju: ako su X, Y dvije nezavisne diskretne slučajne varijable, tada

$$\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y).$$

Rješenje. Vidi predavanja.

VJEROJATNOST I STATISTIKA

Prvi ispitni rok – 24. lipnja 2024.

Zadatak 2. (*12 bodova*)

Broj dana na koje građevinski radnici ne mogu raditi zbog vrućina tijekom srpnja ima distribuciju

$$X \sim \begin{pmatrix} 5 & 6 & 7 & 8 & 9 & 10 & 11 \\ 0.03 & 0.15 & 0.18 & 0.25 & 0.19 & 0.16 & 0.04 \end{pmatrix}.$$

(a) (2 boda) Pronađite vjerojatnost da će idući srpanj biti "izgubljeno" (neradno) najviše tjedan dana rada. Za potrebe zadatka, pretpostavite da radnici rade svaki dan kada nije prevruće (uključujući i vikende).

(b) (4 boda) Izračunajte očekivani broj neradnih dana tijekom srpnja. Koliko u prosjeku taj broj varira?

(c) (3 boda) Na projektu radi 8 radnika i svakom radniku se dnevno računa plaća od 70 eura za svaki dan koji radi, a svaki neradni dan (uslijed vrućine) im se računa 20 eura. Ako će građevinska firma dobiti 25000 eura na kraju srpnja, koliki je očekivani profit?

(d) (3 boda) Koliko radnika trebaju imati tijekom srpnja ako bi htjeli da im očekivana zarada bude bar 15000 eura?

Rješenje.

(a) $\mathbb{P}(X \leq 7) = \mathbb{P}(X = 5) + \mathbb{P}(X = 6) + \mathbb{P}(X = 7) = 0.36$

(b) $\mathbb{E}\, X = 5 * 0.03 + 6 * 0.015 + ... + 11 * 0.04 = 8.06$

Računamo varijancu $\operatorname{Var} X = \mathbb{E}\,(X^2) - (\mathbb{E}\, X)^2 = (5^2 * 0.03 + 6^2 * 0.15 + ... + 11^2 * 0.04) - 8.06^2 = 2.2364$

Očekivani broj neradnih dana je 8.06, a on u prosjeku varira $\sqrt{2.2364} \approx 1.5$ dana.

(c) X je broj neradnih rana u srpnju, pa je onda $31 - X$ broj radnih dana u srpnju. Profit možemo modelirati kao slučajnu varijablu $Y = 25000 - 8(70(31 - X) + 20X)$ računamo $\mathbb{E}\,(25000 - 8(70(31 - X) + 20X)) = 25000 - 8(2170 - 70\mathbb{E}\, X + 20\mathbb{E}\, X) = 10864$

(d) Slično kao u prethodnom zadatku tražimo $\mathbb{E}\,(25000 - n(70(31 - X) + 20X)) \geq 15000$. Iz toga dobijemo $-1767n \geq -10000$, tj. $n \leq 5.66$. Dakle treba uzeti $n = 5$ radnika.

VJEROJATNOST I STATISTIKA

Prvi ispitni rok – 24. lipnja 2024.

Zadatak 3. (*12 bodova*) Jan igra nagradnu igru. Prvo baca simetričan novčić te zatim izvlači novčanice iz kutije u kojoj se nalaze 3 novčanice od 10 eura te 2 novčanice od 20 eura. Ako dobije pismo, smije izvlačiti dva puta (s time da nakon prvog izvlačenja ne vraća novčanicu u kutiju), a ako dobije glavu smije izvlačiti samo jednom. Neka X označava koliko puta Jan izvlači novac iz kutije, a Y neka bude novac koji je zaradio izvlačenjem iz kutije.

(a) (4 boda) Odredite distribuciju slučajnog vektora (X, Y).

(b) (2 boda) Koliko Jan očekuje da će izvući novaca u ovoj igri?

(c) (3 boda) Ako Jan mora platiti 20 eura da sudjeluje u igri, kolika je njegova očekivana zarada, odnosno gubitak, nakon 5 odigranih, nezavisnih, igri?

(d) (3 boda) Jesu li slučajne varijable X i Y korelirane? A nezavisne?

Rješenje.

(a) Računamo,

$X \setminus Y$	10	20	30	40	$\sum$
2	0	$\frac{1}{2}\cdot\frac{3}{5}\cdot\frac{2}{4}$	$\frac{1}{2}\cdot\frac{3}{5}\cdot\frac{2}{4}+\frac{1}{2}\cdot\frac{2}{5}\cdot\frac{3}{4}$	$\frac{1}{2}\cdot\frac{2}{5}\cdot\frac{1}{4}$	$\frac{1}{2}$
1	$\frac{1}{2}\cdot\frac{3}{5}$	$\frac{1}{2}\cdot\frac{2}{5}$	0	0	$\frac{1}{2}$
$\sum$	$\frac{6}{20}$	$\frac{7}{20}$	$\frac{6}{20}$	$\frac{1}{20}$	1

(b) Jan je izvukao Y novaca, pa je očekivano da će izvući

$$\mathbb{E}[Y] = 10 \cdot \frac{6}{20} + 20 \cdot \frac{7}{20} + 30 \cdot \frac{6}{20} + 40 \cdot \frac{1}{20} = 21 \text{ eur.}$$

(c) Očekivana zarada/gubitak u jednoj igri je $\mathbb{E}[Y - 20]$, pa nakon 5 igara očekujemo da će Jan zaraditi

$$\mathbb{E}[5(Y - 20)] = 5\mathbb{E}[Y] - 100 = 5 \cdot 21 - 100 = 5 \text{ eura.}$$

(d) Slučajne varijable X i Y nisu nezavisne. Primjerice, iz tablice distribucije čitamo

$$\mathbb{P}(X = 1, Y = 40) = 0 \neq \frac{1}{2} \cdot \frac{1}{20} = \mathbb{P}(X = 1)\mathbb{P}(Y = 40).$$

Računamo

$$\begin{aligned}
\mathbb{E}[X] &= 2 \cdot \frac{1}{2} + 1 \cdot \frac{1}{2} = \frac{3}{2}, \\
\mathbb{E}[Y] &= 21, \\
\mathbb{E}[XY] &= 2 \cdot 20 \cdot \frac{3}{20} + 2 \cdot 30 \cdot \frac{6}{20} + 2 \cdot 40 \cdot \frac{1}{20} + 1 \cdot 10 \cdot \frac{3}{10} + 1 \cdot 20 \cdot \frac{1}{5} = \\
&= 6 + 18 + 4 + 3 + 4 = 35.
\end{aligned}$$

Budući da $\mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y] = 35 - \frac{63}{2} \neq 0$, X i Y nisu nekorelirane.

VJEROJATNOST I STATISTIKA

Prvi ispitni rok – 24. lipnja 2024.

Zadatak 4. (*13 bodova*) Pretpostavimo da vrijeme završetka nekog maratona prati normalnu razdiobu s očekivanjem 250 (minuta) i standardnom devijacijom 49.

(a) (3 boda) Odredite vjerojatnost da će slučajno odabrani natjecatelj istrčati maraton za manje od 4 sata.

(b) (4 boda) Odredite kada bi trebalo zatvoriti stazu da budemo sigurni s vjerojatnosti 99.5% da su svi završili, tj. nađite vrijeme takvo da je 99.5% natjecatelja gotovo do tog trena.

Sponzor natjecanja je proizvođač napitka A koji se proizvode s okusom limuna i s okusom naranče. U tvornici stroj radi na način da s vjerojatnošću $\frac{2}{3}$ proizvodi piće s okusom limuna, a $\frac{1}{3}$ s okusom naranče.

(c) (3 boda) Ako je 684 natjecatelja i svaki će prije utrke dobiti jedan napitak A, izračunajte približnu vjerojatnost da je najviše 200 natjecatelja dobilo napitak s okusom naranče.

(d) (3 boda) Pronađite najmanji broj napitaka A koje proizvođač mora isporučiti na utrku da bi bio bar 95% siguran da ima bar 250 napitaka od naranče.

Rješenje.

(a) Označimo sa X vrijeme potrebno da slučajno odabrana osoba završi maraton $\sim N(250, 49^2)$. Trebamo izračunati $\mathbb{P}(X < 240) = \mathbb{P}(\frac{X-250}{49} < \frac{240-250}{49}) = \Phi(-0.20) = 1 - \Phi(0.20) = 1 - 0.5793 = 0.4207$

(b) Tražimo x takav da $\mathbb{P}(X \leq x) \geq 0.995$, tj.

$$\mathbb{P}(\frac{X-250}{49} \leq \frac{x-250}{49}) \geq 0.995$$

S obzirom da $\frac{X-250}{49}$ prati $N(0,1)$ razdiobu, u tablici distribucije od $N(0,1)$ tražimo z takav da je $\mathbb{P}(Z \leq z) \geq 0.995$ (pri čemu $Z \sim N(0,1)$). Vidimo $z = 2.58$, dakle $\frac{x-250}{49} = 2.58$ i $x = 376.42$. Stazu možemo zatvoriti nakon 6 sati i 16.42 min.

(c) Označimo sa Y slučajnu varijablu koja označava koliko je napitaka s okusom naranče podijeljeno. Ona ima $\sim B(684, \frac{1}{3})$ razdiobu. Koristit ćemo CGT da izračunamo približnu vjerojatnost

$$\mathbb{P}(Y \leq 200) = \mathbb{P}\left(\frac{Y - 684*1/3}{\sqrt{684*1/3*2/3}} \leq \frac{200-684*1/3}{\sqrt{684*1/3*2/3}}\right) \approx \Phi\left(\frac{200-684*1/3}{\sqrt{684*1/3*2/3}}\right)$$
$$= \Phi(-2.27) = 1 - \Phi(2.27) = 1 - 0.9884 = 0.0116$$

(d) Neka je Z broj napitaka koje će proizvođač donijeti na utrku, tada je $Z \sim B(n, \frac{1}{3})$. Tražimo n takav da vrijedi $\mathbb{P}(Z \geq 250) \geq 0.95$, a lijevu stranu zbog CGT-a možemo procijeniti kao

$$\mathbb{P}\left(\frac{Z - n/3}{\sqrt{n*1/3*2/3}} \geq \frac{250-n/3}{\sqrt{n*1/3*2/3}}\right) \approx 1 - \Phi\left(\frac{250-n/3}{\sqrt{n*1/3*2/3}}\right)$$

Da bi ta vjerojatnost bila veća od 0.95, iz tablice čitamo da mora vrijediti $\frac{n/3-250}{\sqrt{2n/9}} \geq 1.65$. Riješimo kvadratnu jednadžbu (u terminima $\sqrt{n}$) i dobijemo $\sqrt{n} \geq 28.57$, tj. $n \geq 816.68$. Proizvođač mora uzeti bar 817 napitaka.

VJEROJATNOST I STATISTIKA

Prvi ispitni rok – 24. lipnja 2024.

Zadatak 5. (*11 bodova*) Godina je 2124. i ljudi se rađaju s raznim supermoćima (npr. putovanje kroz vrijeme, teleportacija, super snaga, razgovor sa životinjama, ...). Svaku moć čovjek pri rođenju dobiva nezavisno od drugih supermoći. Znamo da 10% ljudi na cijelom planetu ima moć letenja. Uzeli smo slučajan uzorak od 1000 ljudi i ustanovili da njih 121 posjeduje moć nevidljivosti.

(a) (6 boda) Na razini značajnosti od 1% testirajte je li moć nevidljivosti češća od letenja.

(b) (5 boda) Ako bi na razini značajnosti od 3% htjeli odbaciti pretpostavku da je nevidljivost jednako česta kao letenje (i zaključiti da je češća), koliko bi najmanje, od ispitanih 1000 ljudi, trebalo imati moć nevidljivosti?

Rješenje.

(a) Koristimo asimptotski Z–test. Neka je X slučajna varijabla koja označava moć nevidljivosti. Tada je $X \sim B(p)$, gdje nam p nije poznat. Imamo slučajan uzorak $X_1, \dots, X_{1000}$ za X te na temelju njega opaženo $\overline{x}_{1000} = \frac{121}{1000}$. Testiramo hipoteze

$$H_0 : p = 0.1$$
$$H_1 : p > 0.1,$$

budući da je vjerojatnost da osoba ima sposobnost letenja $p_0 = 10\% = 0.1$. Imamo $n = 1000$, $\alpha = 0.01$, pa je $z_\alpha = 2.33$ te $C_\alpha = [2.33, \infty)$. Računamo realizaciju testne statistike $Z = \frac{\overline{X}_n - p_0}{\sqrt{p_0(1-p_0)}}\sqrt{n} \overset{H_0}{\sim} AN(0,1)$,

$$z = \frac{\overline{x}_{1000} - 0.1}{\sqrt{0.1 \cdot 0.9}} \cdot \sqrt{1000} = 2.21 \notin C_\alpha,$$

pa na razini značajnosti od 1% ne odbacujemo nultu hipotezu da je nevidljivost jednako česta kao i letenje.

(b) Sada imamo $n = 1000$, $\alpha = 0.03$, $C_\alpha = [z_\alpha, \infty) = [1.89, \infty)$. Dakle, tražimo najmanji broj ljudi x takav da

$$z = \frac{\frac{x}{1000} - 0.1}{\sqrt{0.1 \cdot 0.9}}\sqrt{1000} \geq 1.89.$$

Rješavanjem gornje nejednakosti dobivamo $x \geq 117.93$, odnosno barem 118 ljudi bi trebalo imati moć nevidljivosti.