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

FAKULTETA ZA MATEMATIKO IN FIZIKO

Oddelek za matematiko

Statistika

Teoretični izpit

12. julij 2017

Navodila

Pazljivo preberite besedilo naloge, preden se lotite reševanja. Nalog je 10, ocena pa je enaka navzgor zaokroženemu številu pravilnih odgovorov. Ko je ponujenih več možnosti, je lahko pravilnih odgovorov več. Ko ni ponujenih odgovorov, na kratko pojasnite vaš razmislek.

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

1. Naj bodo vektor $\mathbf{X} = (X_1, X_2, X_3, X_4)^T$ večrazsežen normalen z

$$E\,(\mathbf{X}) = 0 \qquad \text{in} \qquad \mathrm{var}\,(\mathbf{X}) = \begin{pmatrix} 1 & \rho & \rho & \rho \\ \rho & 1 & \rho & \rho \\ \rho & \rho & 1 & \rho \\ \rho & \rho & \rho & 1 \end{pmatrix}$$

za nek $\rho \in (-1, 1)$. Označite

$$\alpha = \frac{\rho}{1+\rho}\,.$$

Pokažite, da sta vektorja

$$\begin{pmatrix} X_2 - \alpha X_1 - \alpha X_4 \\ X_3 - \alpha X_1 - \alpha X_4 \end{pmatrix} \quad \text{in} \quad \begin{pmatrix} X_1 \\ X_4 \end{pmatrix}$$

neodvisna, in izračunajte

$$\mathrm{cov}\,(X_2, X_3 | X_1, X_4)\ .$$

2. Naj bodo $X_1, X_2, X_3, Y_1, Y_2, Y_3$ med sabo neodvisne, standardizirano normalne slučajne spremenljivke. Definirajte

$$U = \frac{(X_1Y_1 + X_2Y_2 + X_3Y_3)^2}{X_1^2 + X_2^2 + X_3^2}\,.$$

Navedite porazdelitev slučajne spremenljivke U.

Namig: pogojujte na ustrezne slučajne spremenljivke.

3. Naj bosta $\mathbf{H}$ in $\mathbf{K}$ simetrični idempotentni matriki s $\mathbf{HK} = \mathbf{KH}$ in naj bo $\mathbf{Z} \sim \mathrm{N}(0, \mathbf{I})$. Pokažite, da iz neodvisnosti kvadratnih form

$$\mathbf{Z}^T\mathbf{HZ} \qquad \text{in} \qquad \mathbf{Z}^T\mathbf{KZ}$$

sledi $\mathbf{HK} = 0$.

Kot znano privzemite, da lahko simetrični matriki, ki komutirata, diagonaliziramo z isto ortogonalno prehodno matriko $\mathbf{Q}$.

4. Naj bodo X, Y in Z diskretne slučajne spremenljivke. Predpostavljajte, da je vektor (X, Z) neodvisen od Y. Predpostavite, da je $E(|Z|) < \infty$. Utemeljite, da je
$$E(Z|X, Y) = E(Z|X)\,.$$

5. Naj bo $S_m \sim \mathrm{NegBin}(m, p)$. Za fiksen $x \in \mathbb{R}$ izrazite
$$\lim_{m\to\infty} P\left(S_m \geq \frac{m}{p} + x\sqrt{m}\right)$$
s porazdelitveno funkcijo $\Phi(z)$ standardizirane normalne porazdelitve.

   Namig: kaj je treba sešteti, da dobite S_m?

6. Populacija velikosti N je razdeljena na K stratumov velikosti $N_1, \ldots, N_K$. Izbrali bomo stratificiran vzorec, tako da iz vsakega stratuma izberemo enostavne slučajne vzorce velikosti $n_1, \ldots, n_K$. Naj bodo $\mu_1, \ldots, \mu_K$ populacijska povprečja po stratumih. Predlagajte nepristransko cenilko količine
$$\gamma = \sum_{k=1}^{K} \mu_k^2\,.$$

   Namig: če je σ_k^2 populacijska varianca v k-tem stratumu, je
$$\sigma_k^2 = \frac{1}{N_k}\sum_{i=1}^{N_k} y_{ki}^2 - \mu_k^2\,,$$
   kjer je y_{ki} vrednost spremenljivke za i-to enoto v k-tem stratumu.

7. Predpostavite, da so opazovane vrednosti $x_1, \ldots, x_n$ nastale kot med sabo neodvisne, enako porazdeljene z $X_k \sim U(0, \theta)$, $k = 1, 2, \ldots, n$ in $\theta > 0$. Za dan $\alpha \in (0, 1)$ navedite eksakten interval zaupanja za parameter θ pri stopnji tveganja α.

   Namig: interval je lahko oblike $(\max(X_1, \ldots, X_n), a \max(X_1, \ldots, X_n))$ *za določen* $a > 1$.

8. Privzemite, da so opazovane vrednosti $x_1, \ldots, x_n$ nastale kot med sabo neodvisne, enako porazdeljene slučajne spremenljivke $X_1, \ldots, X_n$ z $X_k \sim \mathrm{N}(\mu, 1)$ za $k = 1, 2, \ldots, n$. Domnevo

   $$H_0\colon \mu = 0 \qquad \text{proti} \qquad H_1\colon \mu \neq 0$$

   preizkusimo s testno statistiko

   $$Z = \sqrt{n}\,\bar{X}\,.$$

   Domnevo H_0 zavrnemo, če je $|Z| > z_\alpha$, kjer izberemo z_α v skladu z izbrano stopnjo tveganja $\alpha \in (0, 1)$. Pokažite, da je za poljuben $\mu \neq 0$ moč testa strogo večja od α.

   Namig: za fiksen $a > 0$ *funkcija* $F(x) = \Phi(x + a) - \Phi(x - a)$ *v* $x = 0$ *doseže strogi maksimum.*

9. Predpostavite regresijski model oblike

   $$Y_k = \alpha + \beta x_k + \epsilon_k$$

   za $k = 1, 2, \ldots, n$, kjer je $E(X_k) = 0$, $\mathrm{var}(\epsilon_k) = \sigma^2$ in $\mathrm{cov}(\epsilon_k, \epsilon_l) = 0$ za $k \neq l$. Za nov x_{k+1} je napoved Y_{k+1} na podlagi regresijskega modela enaka

   $$\hat{Y}_{k+1} = \hat{\alpha} + \hat{\beta} x_{k+1}\,,$$

   kjer sta $\hat{\alpha}$ in $\hat{\beta}$ cenilki po metodi najmanjših kvadratov na podlagi prvih n opazovanj. Privzemamo $\mathrm{cov}(\epsilon_k, \epsilon_{n+1}) = 0$ za $k = 1, 2, \ldots, n$. Izračunajte

   $$E\left[\left(Y_{n+1} - \hat{Y}_{n+1}\right)^2\right]$$

eksplicitno.

Namig: prevedite na računanje varianc.

10. Predpostavite regresijski model oblike

$$Y_k = \alpha + \beta x_k + \epsilon_k$$

za $k = 1, 2, \ldots, n$, kjer je $E(X_k) = 0$, $\mathrm{var}(\epsilon_k) = \sigma^2$ in $\mathrm{cov}(\epsilon_k, \epsilon_l) = 0$ za $k \neq l$. Definirajte

$$\hat{\epsilon}_k = Y_k - \hat{\alpha} - \hat{\beta} x_k$$

za $k = 1, 2, \ldots, n$, pri čemer sta $\hat{\alpha}$ in $\hat{\beta}$ cenilki po metodi najmanjših kvadratov. Za fiksne $a_1, \ldots, a_n$ izračunajte

$$E\left[\left(\sum_{k=1}^{n} a_k \hat{\epsilon}_k\right)^2\right]$$

čim bolj eksplicitno.

Namig: računajte matrično.