Ime in priimek: ______________________________ Vpisna št: | | | | | | | | |

FAKULTETA ZA MATEMATIKO IN FIZIKO

Oddelek za matematiko

Statistika

Teoretični izpit

8. julij 2020

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 bo $\mathbf{X}$ tak slučajni vektor, da ima za vsako ortogonalno matriko $\mathbf{Q}$ vektor $\mathbf{QX}$ enako porazdelitev kot $\mathbf{X}$. Kakšne oblike je lahko matrika $\mathrm{var}(\mathbf{X})$?

2. Naj bo $\mathbf{Z} \sim \mathrm{N}(0, \mathbf{I})$. Kot znano privzemite, da je za simetrični matriki $\mathbf{A}$ in $\mathbf{B}$

$$\mathrm{cov}\left(\mathbf{Z}^T\mathbf{A}\mathbf{Z}, \mathbf{Z}^T\mathbf{B}\mathbf{Z}\right) = 2\mathrm{Sl}\left(\mathbf{A}\mathbf{B}\right) .$$

   Naj bosta $\mathbf{H}$ in $\mathbf{K}$ simetrični idempotentni matriki z $\mathbf{HK} = \mathbf{KH}$. Utemeljite, da iz neodvisnosti

$$U = \mathbf{Z}^T\mathbf{H}\mathbf{Z} \quad \text{in} \quad V = \mathbf{Z}^T\mathbf{K}\mathbf{Z}$$

   sledi $\mathbf{HK} = 0$.

3. Naj ima par (X, Y) bivariatno normalno porazdelitev in naj bo

$$\alpha = \frac{\mathrm{cov}(X, Y)}{\mathrm{var}(X)} .$$

   Utemeljite, da je za poljubno omejeno zvezno funkcijo f

$$E\left[f(Y)|X\right] = \psi(X) ,$$

   kjer je

$$\psi(x) = E\left[f\left(Y - \alpha X + \alpha x\right)\right] .$$

4. Naj bodo $Z_0, Z_1, \ldots, Z_n$ neodvisne standardizirano normalne in definirajte

$$W_k = Z_0 + Z_k$$

za $k = 1, 2, \ldots, n$. Izračunajte

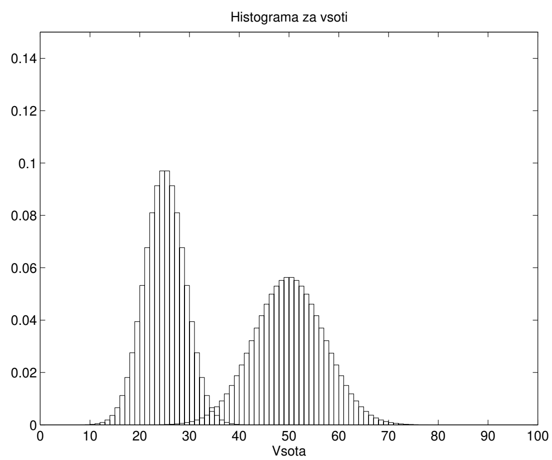
$$E\left(Z_0 | W_1, W_2, \ldots, W_n\right) .$$

Namig: velja $(\mathbf{I} + \mathbf{1}\mathbf{1}^T)^{-1} = \mathbf{I} + c\mathbf{1}\mathbf{1}^T$ *za primerni* c.

5. Na spodnji sliki sta histograma za porazdelitvi vsot 25 neodvisnih izbiranj iz ene od naslednjih dveh škatel:

   (i) $\boxed{0}\ \boxed{1}\ \boxed{2}$

   (ii) $\boxed{0}\ \boxed{1}\ \boxed{2}\ \boxed{3}\ \boxed{4}$



Sl. 1 Histograma za vsote neodvisnih izbiranj iz škatel (i) ali (ii).

Kateri od zgornjih dveh histogramov pripada izbiranjem iz škatle (i) in kateri izbiranjem iz škatle (ii)? Utemeljite odgovor.

6. Populacija velikosti $2N$ je razdeljena na stratuma velikosti N. Vemo, da sta populacijski varianci v obeh stratumih enaki τ^2, vendar ne vemo kakšni, nič pa ne vemo o populacijskih povprečjih. V vsakem stratumu izberemo enostavni slučajni vzorec velikosti n, pri čemer predpostavljamo $n \leq N$. Kot znano

privzemite, da je populacijska varianca σ^2 enaka

$$\sigma^2 = \tau^2 + \left(\frac{\mu_1 - \mu_2}{2}\right)^2 ,$$

kjer sta μ_1 in μ_2 populacijski povprečji za prvi in drugi stratum. Utemeljite, da je v primeru, ko je

$$\frac{1}{2}(N-1)(\mu_1 - \mu_2)^2 > \tau^2$$

bolje izbrati zgoraj opisano stratificirano vzorčenje s pripadajočo nepristransko cenilko populacijskega povprečja za celotno populacijo kot izbrati enostavni slučajni vzorec velikosti $2n$ iz celotne populacije in populacijsko povprečje oceniti z vzorčnim povprečjem. Večjo točnost razumemo v smislu manjše standardne napake.

7. Naj bodo $X_1, X_2, \ldots, X_n$ neodvisne in enako porazdeljene s končnimi variancami. Pokažite, da je

$$\hat{\sigma}^2 = \frac{1}{n-1} \sum_{k=1}^{n} (X_k - \bar{X})^2$$

nepristranska cenilka variance porazdelitve X_1.

8. Predpostavljamo, da so opazovane vrednosti vzorec med sabo neodvisnih, normalno porazdeljenih slučajnih spremenljivk $X_1, \ldots, X_n$ s parametroma μ in $\sigma^2 = 1$. Preverjamo domnevo

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

Domnevo zavrnemo, če je $|\bar{X}| \geq c_\alpha$, kjer je $\alpha \in (0,1)$ in za c_α velja $\Phi(\sqrt{n}c_\alpha) = 1 - \frac{\alpha}{2}$. Utemeljite, da je za vse $\mu \neq 0$ moč testa strogo večja od α.

9. Predpostavimo naslednji regresijski model:

$$\begin{aligned} Y_{i1} &= \beta x_{i1} + \epsilon_i \\ Y_{i2} &= \beta x_{i2} + \eta_i \end{aligned}$$

za $i = 1, 2, \ldots, n$. Predpostavljamo, da so pari $(\epsilon_1, \eta_1), \ldots, (\epsilon_n, \eta_n)$ med sabo neodvisni in enako porazdeljeni z $E(\epsilon_i) = E(\eta_i) = 0$. Kaj mora veljati za konstante $a_1, \ldots, a_n$ in $b_1, \ldots, b_n$, da bo

$$\hat{\beta} = \sum_{i=1}^{n} (a_i Y_{i1} + b_i Y_{i2})$$

nepristranska cenilka parametra β?

10. Označite z $\mathbf{1}_r$ vektor dimenzije r, ki ima vse komponente enake 1. Naj bo $m < n$ in naj za $n \times m$ matriko $\mathbf{X}$ velja

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}\,,$$

pri čemer je $\boldsymbol{\epsilon} \sim \mathrm{N}(\mathbf{0}, \sigma^2\boldsymbol{\Sigma})$ za neznan parameter σ^2 in

$$\boldsymbol{\Sigma} = \mathbf{I} + \rho \mathbf{1}_n \mathbf{1}_n^T$$

za $\rho \in (0,1)$. Matrika $\mathbf{X}$ ima poln rang. Privzemite, da je $\mathbf{X}^T\mathbf{X} = \mathbf{I}$ in $\mathbf{X}^T\mathbf{1}_n = a\mathbf{1}_m$ za znano konstanto a. Naj bo $\hat{\boldsymbol{\beta}} = \mathbf{X}^T\mathbf{Y}$ cenilka $\boldsymbol{\beta}$ po metodi najmanjših kvadratov in naj bo $\hat{\boldsymbol{\epsilon}} = \mathbf{Y} - \mathbf{X}\hat{\boldsymbol{\beta}}$. Pokažite, da je

$$\hat{\sigma}^2 = \frac{1}{n - m + n\rho - a^2\rho m} \sum_{k=1}^{n} \hat{\epsilon}_i^2$$

nepristranska cenilka σ^2.

Namig: pomaga, če opazite, da je $\mathrm{Sl}\left(\mathbf{X}\mathbf{1}_m\mathbf{1}_n^T\right)$ *vsota vseh elementov matrike* $\mathbf{X}$ *ter da je* $\mathbf{X}\mathbf{X}^T$ *idempotentna.*