IME IN PRIIMEK: ____________________ VPISNA ŠT: | | | | | | | | |

FAKULTETA ZA MATEMATIKO IN FIZIKO

ODDELEK ZA MATEMATIKO

STATISTIKA

TEORETIČNI IZPIT

6. julij 2018

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 Z_0, Z_1, Z_2, Z_3 neodvisne standardizirano normalne slučajne spremenljivke in $\theta \in (-1,1)$. Definirajmo nove slučajne spremenljivke

$$X = Z_1 + \theta Z_0 \,, Y = Z_2 + \theta Z_1 \quad \text{in} \quad W = Z_3 + \theta Z_2 \,.$$

Izračunajte $E(W|X,Y)$.

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

$$\operatorname{cov}\big(\mathbf{Z}^T\mathbf{A}\mathbf{Z}, \mathbf{Z}^T\mathbf{B}\mathbf{Z}\big) = 2\,\mathrm{Sl}\,(\mathbf{AB}) \ .$$

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

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

sledi $\mathbf{HK} = 0$.

3. Naj bodo X, Y in Z diskretne slučajne spremenljivke. Predpostavite, da je $E(|Z|) < \infty$. Pojasnite, zakaj je

$$E(Z|X,Y) = E(Z|X-Y, X+Y) \,.$$

4. Naj bo (X,Y,Z) slučajni vektor s komponentami, ki imajo končen drugi moment. Privzemite, da je $E(X) = E(Y) = E(Z) = 0$. Pokažite, da je

$$\operatorname{cov}\big(E\,(Z|X,Y)\,, Y - E\,(Y|X)\big) = E\big[E\,(ZY|X) - E\,(Z|X)\,E\,(Y|X)\big] \,.$$

5. Iz populacije velikosti N, ki je razdeljena na K stratumov velikosti $N_1, \ldots, N_K$, izberemo stratificiran vzorec. Velikosti enostavnih slučajnih vzorcev po stratumih naj bodo $n_1, \ldots, n_K$. Naj bosta μ in σ^2 populacijsko povprečje in populacijska varianca za celotno populacijo. Predlagajte nepristransko cenilko za parameter $\theta = \sigma^2 + \mu^2$ na podlagi izbranega stratificiranega vzorca.

6. Predpostavite, da podatki $x_1, \ldots, x_n$ nastanejo kot med sabo neodvisne, enako porazdeljene slučajne spremenljivke $X_1, \ldots, X_n$ z $X_k \sim \mathrm{Po}(\lambda)$ z $\lambda > 0$. Pokažite, da je

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

nepristranska cenilka parametra λ.

7. Recimo, da ocenjujemo parametre normalne porazdelitve na podlagi metode največjega verjetja. Če so $X_1, \ldots, X_n$ opazovane vrednosti, je cenilka parametra σ^2 po metodi največjega verjetja enaka

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

Vemo, da je $\sum_{k=1}^{n} (X_k - \bar{X})^2 \sim \sigma^2 \chi^2(n-1)$. Kako lahko na podlagi tega dejstva pojasnite, da je porazdelitev cenilke približno normalna?

8. Predpostavite, da so opazovane vrednosti pari $(x_1, y_1), \ldots, (x_n, y_n)$, za katere predpostavljamo, da so nastali kot med sabo neodvisni pari slučajnih spremenljivk $(X_1, Y_1), \ldots, (X_n, Y_n)$ z

$$\begin{pmatrix} X_k \\ Y_k \end{pmatrix} \sim \mathrm{N}\left(\boldsymbol{\mu}, \mathbf{I}\right) .$$

Preizkušamo domnevo H_0, da za nek $k \in \mathbb{R}$ velja $\boldsymbol{\mu} = k\mathbf{1}$, proti domnevi H_1, da tak k ne obstaja. Označimo

$$\mathbf{Z}_k = \begin{pmatrix} X_k \\ Y_k \end{pmatrix}$$

za $k = 1, 2, \ldots, n$ in $\bar{\mathbf{Z}} = \frac{1}{n}\sum_{k=1}^{n} \mathbf{Z}_k$. Za Wilksovo testno statistiko dobimo

$$\lambda = n\bar{\mathbf{Z}}^T \left(\mathbf{I} - \frac{1}{n}\mathbf{1}\mathbf{1}^T \right) \bar{\mathbf{Z}} \, .$$

Je Wilksova aproksimacija porazdelitve testne statistike, če H_0 drži, aproksimacija ali je točna? Navedite še stopinje prostosti in utemeljite odgovor.

9. Predpostavite regresijski model

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

kjer predpostavljamo

$$\boldsymbol{\epsilon} \sim \mathrm{N}\left(0, \sigma^2(\mathbf{I} + \mathbf{1}\mathbf{1}^T)\right) \, .$$

pri čemer je

$$\mathbf{X} = \begin{pmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{pmatrix}$$

in $\sum_{k=1}^{n} x_k = 0$. Matrika $\mathbf{X}$ naj ima poln rang. Označite $\hat{\boldsymbol{\beta}} = \begin{pmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \end{pmatrix}$. Utemeljite, da sta $\hat{\beta}_1$ in $\hat{\beta}_2$ neodvisni cenilki.

Namig: inverz matrike $\mathbf{I} + \mathbf{1}\mathbf{1}^T$ *je oblike* $\mathbf{I} + b\mathbf{1}\mathbf{1}^T$.

10. Privzemite linearni regresijski model

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

kjer je β skalarni parameter, $E(\boldsymbol{\epsilon}) = 0$ in $\mathrm{var}(\boldsymbol{\epsilon}) = \boldsymbol{\Sigma}$. Predpostavljajte, da je matrika $\boldsymbol{\Sigma}$ obrnljiva. Vektorji $\mathbf{Y}$, $\mathbf{x}$ in $\boldsymbol{\epsilon}$ so n-dimenzionalni. Privzemite, da za določen vektor $\mathbf{a}$ velja $\mathbf{a}^T\mathbf{x} = 1$ in še, da je

$$\mathrm{cov}(\boldsymbol{\epsilon}, \mathbf{a}^T\mathbf{Y}) = k\mathbf{x}$$

za neko konstanto k. Pokažite, da je $\hat{\beta} = \mathbf{a}^T \mathbf{Y}$ nepristranska, linearna cenilka parametra β z najmanjšo možno varianco.