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

FAKULTETA ZA MATEMATIKO IN FIZIKO

Oddelek za matematiko

Statistika

Teoretični izpit

19. junij 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 $\mathbf{X}$ slučajni vektor dimenzije n z $E(\mathbf{X}) = a\mathbf{1}$ in $\mathrm{var}(\mathbf{X}) = \mathbf{I} + b\mathbf{1}\mathbf{1}^T$. Izračunajte
$$E\left[\mathbf{X}^T\left(\mathbf{I} - \frac{1}{n}\mathbf{1}\mathbf{1}^T\right)\mathbf{X}\right] .$$

2. Naj ima par (X, Y) dvorazsežno normalno porazdelitev z $E(X) = E(Y) = 0$ in $\mathrm{var}(X) = \mathrm{var}(Y) = 1$ ter $\mathrm{cov}(X, Y) = \rho \in (-1, 1)$. Kot znano privzemite, da je $E(X^4) = 3$. Izračunajte $\mathrm{cov}(X, Y - \rho X)$. Izračunajte
$$E\left(X^2Y^2\right) = E\left(X^2(Y - \rho X + \rho X)^2\right) .$$

3. Naj bo $X \sim \mathrm{N}\left(0, \mathbf{H}\right)$, kjer je
$$\mathbf{H} = \mathbf{I} - \frac{1}{n}\mathbf{1}\mathbf{1}^T$$
idempotentna. Navedite porazdelitev $W = X_1^2 + \cdots + X_n^2$.

4. Naj bodo Z_0, Z_1, Z_2 neodvisne standardizirano normalne in definirajte
$$W_k = Z_0 + Z_k$$
za $k = 1, 2$. Izračunajte
$$E\left(Z_0 | W_1, W_2\right) .$$

5. Naj bodo $X_1, X_2, \dots$ med sabo neodvisne slučajne spremenljivke z $E(X_i) = 0$ za $i = 1, 2, \dots$. Označimo $S_n = X_1 + \cdots + X_n$ in naj bo $Z \sim \mathrm{N}(0,1)$. Naj bo f trikrat zvezno odvedljiva funkcija na $\mathbb{R}$ z $|f|, |f'|, |f''|, |f'''| \le M < \infty$. Označimo $\sigma_n^2 = \mathrm{var}(S_n)$. Dokaz centralnega limitnega izreka temelji na neenačbi

$$\left| E\left[f\left(\frac{S_n}{\sigma_n}\right)\right] - E\left[f(Z)\right] \right| \le \frac{1}{6}\left(1 + \sqrt{\frac{8}{\pi}}\right) M\big(E(|X_1|^3/\sigma_n^3)) + \ldots + E(|X_n|^3/\sigma_n^3)\big)\,.$$

Ali centralni limitni izrek še vedno velja, če predpostavljamo neodvisnost spremenljivk $X_1, X_2, \ldots$, predpostavko, da so $X_1, X_2, \dots$ enako porazdeljene, pa nadomestimo s predpostavko

$$\frac{n}{\sigma_n^3} \to 0\,,$$

ko $n \to \infty$ in $E\left(|X_i|^3\right) \le C < \infty$ za $i = 1, 2, \dots$. Na kratko obrazložite.

6. Iz populacije velikosti N izberemo verjetnostni vzorec velikosti n. Naj bo I_k indikator dogodka, da bo k-ta enota v populaciji izbrana v vzorec za $k = 1, 2, \dots, N$. Privzemite, da imajo vsi pari (I_k, I_l) za $k \neq l$ enako porazdelitev. Naj bo $\bar{X}$ vzorčno povprečje. Izračunajte $\mathrm{var}(\bar{X})$.

   Namig: kaj je $I_1 + I_2 + \cdots + I_N$?

7. Predpostavite, da so opazovane vrednosti $x_1, x_2, \dots, x_n$ vzorec iz normalne porazdelitve $\mathrm{N}(\mu, 1)$. Naj bo $\tilde{\mu}$ nepristranska cenilka parametra μ. To pomeni, da je $\tilde{\mu}$ funkcija $X_1, \dots, X_n$ in velja

$$\mu = \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} \tilde{\mu}(\mathbf{x}) e^{-\frac{1}{2}\sum_{k=1}^n (x_k - \mu)^2} d\mathbf{x}$$

za vsak μ. Odvajajte obe strani po μ in privzemite, da lahko odvajate pod integralskim znakom. Sklepajte, da je

$$\mathrm{cov}\left(\tilde{\mu}(X_1, \dots, X_n), \sum_{k=1}^n X_k\right) = 1\,.$$

8. Predpostavljamo, da so opazovane vrednosti nastale kot neodvisne, enako porazdeljene slučajne spremenljivke $X_1, X_2, \ldots, X_n$ s porazdelitvijo $\mathrm{N}(\mu, 1)$. Testirati želimo domnevo

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

Po metodi kvocienta verjetja dobimo

$$\lambda = n\bar{x}^2$$

z $\bar{x} = \frac{1}{n}\sum_{k=1}^{n} x_k$. Navedite najprej aproksimativno porazdelitev testne statistike po Wilksovem izreku, če H_0 drži. Je ta aproksimativna porazdelitev tudi eksaktna? Obrazložite odgovor.

9. Privzemite regresijski model

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

za $k = 1, 2, \ldots, n$, kjer so $\epsilon_1, \ldots, \epsilon_n$ nekorelirani $E(\epsilon_k) = 0$ in $\mathrm{var}(\epsilon_k) = \sigma^2$ za vse $k = 1, 2, \ldots, n$. Privzemite, da je $x_k > 0$ za vse $k = 1, 2, \ldots, n$. Oglejte si spodnje linearne cenilke parametra β. Katera od njih ima najmanjšo varianco? Utemeljite vaš odgovor

$$\begin{aligned} \hat{\beta}_1 &= \frac{\sum_{k=1}^{n} x_k Y_k}{\sum_{k=1}^{n} x_k^2} \\ \hat{\beta}_2 &= \frac{1}{n}\sum_{k=1}^{n} \frac{Y_k}{x_k} \\ \hat{\beta}_3 &= \frac{\sum_{k=1}^{n} Y_k}{\sum_{k=1}^{n} x_k} \end{aligned}$$

Namig: na vprašanje lahko odgovorite brez eksplicitnega računanja varianc.

10. Opazimo

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

pri čemer je $E(\boldsymbol{\epsilon}) = 0$ in $\mathrm{var}(\boldsymbol{\epsilon}) = \sigma^2\mathrm{I}$, $\mathbf{V}$ pa je znana $n \times n$ obrnljiva matrika. Predpostavljamo, da ima znana matrika $\mathbf{X}$ poln rang enak $m < n$. Zapišite formulo za nepristransko cenilko $\hat{\sigma}^2$ parametra σ^2.