

STATISTIKA

ISPITNI ROK 23.02.2012.

ZADATAK 1: [8+9+8+5+10=40 bodova]Pretpostavimo da slučajne varijable Y_1, \dots, Y_n zadovoljavaju

$$Y_i = \beta x_i + \varepsilon_i, \quad i = 1, \dots, n,$$

gdje su $x_i, i = 1, \dots, n$ poznate konstante, β nepoznati parametar, $\varepsilon_i, i = 1, \dots, n$ nezavisne s $\mathcal{N}(0, \sigma^2)$ distribucijom, pri čemu σ^2 nepoznato.

- (a) Nađite dovoljnu statistiku za dvodimenzionalni parametar (β, σ^2) .
- (b) Nađite procjenitelja maksimalne vjerodostojnosti $\hat{\beta}$ nepoznatog parametra β te provjerite je li nepristran.
- (c) Odredite distribuciju od $\hat{\beta}$.
- (d) Pokažite da je procjenitelj

$$\hat{\beta}_1 = \frac{\sum_{i=1}^n Y_i}{\sum_{i=1}^n x_i},$$

nepristran za β .

- (e) Koji je od procjenitelja $\hat{\beta}$ i $\hat{\beta}_1$ efikasniji?

ZADATAK 2: [10 bodova]Neka je (X_1, \dots, X_n) jednostavan slučajni uzorak. Koji uvjet mora biti ispunjen za α da bi

$$V = \frac{1}{\alpha} \sum_{i=1}^{n-1} (X_{i+1} - X_i)^2$$

bio nepristran procjenitelj za varijancu ove populacije?

ZADATAK 3: [4+6+5+8+7=30 bodova]U tablici **Puerto** (iz paketa **BSDA**) nalaze se podaci o tjednom dohotku 50 portorikanskih obitelji koje žive u Miamiu.

- (a) Ukazuje li boxplot na postojanje outliera? Radi li se o natprosječno siromašnoj ili natprosječno bogatoj obitelji?
- (b) Nacrtajte histogram, QQ-plot za normalnu distribuciju, te graf uzoračke funkcije gustoće usporedno s grafom gustoće normalne distribucije čije parametre ćete procijeniti. Komentirajte možemo li na osnovu svake od tih slika naslutiti da je obilježje normalno distribuirano. Testirajte je li obilježje normalno distribuirano na razini značajnosti 0.01.
- (c) Ako iz podataka izbacite outlier koji ste uočili na boxplotu, je li obilježje normalno distribuirano na nivou značajnosti 0.05?
- (d) Nađite 95% interval pouzdanosti za prosječni dohodak. Smatra se da je dohodak na granici siromaštva 300\$ tjedno. Žive li portorikanske obitelji u prosjeku ispod granice siromaštva?
- (e) Poznati su i podaci od dohotku 10 slučajno odabranih američkih obitelji

312, 302, 331, 296, 298, 315, 320, 278, 333, 325.

Imaju li američke obitelji veći prosječni dohodak na razini značajnosti 0.05?

ZADATAK 4: [2+4+4=10 bodova]

Kako bi se utvrdilo povećava li neki lijek na intelektualne sposobnosti, odabrani su blizanci i podvrgnuti testu koji testira njihovu inteligenciju. Pri tome je jednom od blizanaca dan lijek. Podaci se nalaze u tablici **Twin** (iz paketa **BSDA**). U stupcu **TwinA** nalaze se rezultati testa za blizanca koji je uzeo lijek, a u stupcu **TwinB** su rezultati drugog blizanca.

- (a) Provjerite jesu li rezultati normalno distribuirani u oba promatrana slučaja te nacrtajte usporedno boxplot dvije skupine podataka različitom bojom. Komentirajte. Što bi se moglo reći na osnovu slike?
- (b) Povećava li lijek intelektualne sposobnosti, na razini značajnosti 0.05?
- (c) Povećava li lijek varijabilnost rezultata, na razini značajnosti 0.05?

ZADATAK 5: [10 bodova]

U tablici **Funeral** (iz paketa **BSDA**) nalaze se podaci o troškovima pogreba ispitanim u pogrebnim poduzećima po različitim regijama SAD-a. Troškovi su kategorizirani na manje, prosječne i velike. Je li trošak pogreba jednako distribuiran po svim regijama ($\alpha = 0.05$)?