

# Računarska statistika

Snježana Lubura Strunjak

Zagreb, 27. svibnja 2021.

# Bootstrap uzorkovanje za testiranje hipoteza

Opis metode:

Testira se  $H_0 : \mu_1 = \mu_2$ , gdje je  $X_1, \dots, X_m$  uzorak iz prve populacije, a  $Y_1, \dots, Y_n$  uzorak iz druge populacije. Pretpostavka je da su uzorci međusobno nezavisni.

- Ponavljano uzimanje bootstrap uzoraka iz  $(X_1 - \text{mean}X, X_2 - \text{mean}X, \dots, X_m - \text{mean}X), (Y_1 - \text{mean}Y, Y_2 - \text{mean}Y, \dots, Y_n - \text{mean}Y)$ ,
- na svakom bootstrap uzorku se računa pogodna test statistika, kao npr. Welchova  $t$  statistika:

$$t_W = \frac{\text{mean}X - \text{mean}Y}{\sqrt{\text{stderr}X^2 + \text{stderr}Y^2}}$$

- Bootstrap distribucija  $t_W$  konvergira sa vjerojatnošću 1 standardnoj normalnoj distribuciji (isto kao i  $t_W$  pod nultom hipotezom u "stvarnom svijetu")

## Primjer

Programi *Rjesenja primjera Bootstrap uzorkovanje za testiranje hipoteze(2uzorka).sas* i *Rjesenja primjera Bootstrap uzorkovanje za testiranje hipoteze(2uzorka) sa surveyselect.sas* se nalaze u folderu Primjeri na Merlinu.

Dane su ocjene iz statistike studenata koji kontinuirano studiraju na University of Washington (grupa=1,  $n=13$ ), te onih koji su prešli na University of Washington sa nekog drugog fakulteta (grupa=2,  $n=34$ ) (podaci za aproksimativan randomizacijski test). Bootstrap testom (metodom opisanom na prethodnom slide-u) treba testirati hipotezu  $H_0 : \mu_1 = \mu_2$ .

**Zadatak** Pročitajte članak *Bootstrap Boos.pdf* u folderu *Primjeri* na Merlinu. Nakon toga pokrenite program *Rjesenja primjera Bootstrap PCB Example 2 Boos.sas* iz foldera *Primjeri* na Merlinu. Uočite da se u programu koristi parametarski i neparametarski bootstrap. Pomoću njih dobijemo procjenu pogreške procjene tražene vjerojatnosti.

# Zadaća

9. zadaća: rok za predaju 10.6.

Zadaća se nalazi u folderu Zadaće na MERLINU.

UPUTE: Svaki zadatak iz zadaće mora biti u svom .sas programu. Sve .sas programe nazovite na način *prezime\_ime\_zad1.sas*, ako je npr. 1.zadatak u pitanju, itd. Sve što radite u zadaćama mora biti u obliku koda (možete koristiti sve dostupne materijale da dobijete tražene rezultate, ali sve mora biti napisano u obliku koda).