

Statistički i grafički prikaz bivarijacionih podataka: zadaci za vežbe**Priprema za čas:**

Fajl koji ćemo koristiti u 1., 2. i 5. zadatku je **CY08MSP_STU_QQQ_SRB.sav**. U pitanju je, kao i na prethodnim vežbama, baza sa PISA 2022 testiranja čija je jedinica posmatranja učenik, a ne zemlja, a koja ima 613744 redova i 1278 varijabli. Bazu smo prilagodili za čas skrativši je na 6413 redova, odnosno podatke samo za zemlju Srbiju.

Sve informacije o testiranju, uzorku, varijablama, zadacima i ostalim materijalima PISA 2022 nalaze se na sledećem linku: <https://www.oecd.org/en/data/datasets/pisa-2022-database.html>.

Zadatak 1.

U originalnoj PISA bazi i bazi **CY08MSP_STU_QQQ_SRB.sav** nalaze se, između ostalog, podaci o čitalačkoj pismenosti (varijabla **PV1READ**) i matematičkoj pismenosti (varijabla **PV1MATH**).

- Grafički prikazati vezu (**Scatterplot**) između čitalačke pismenosti i matematičke pismenosti (varijabla **PV1READ** ide na X-osu, a **varijabla PV1MATH na Y-osu**);
- Korišćenjem komande **Analyze/Correlate/Bivariate Correlation** izračunati kovarijansu između matematičke i čitalačke pismenosti;
- Izračunati koeficijent linearne korelacije između matematičke i čitalačke pismenosti;

Šta zaključujemo o povezanostima?

Zadatak 2.

Fajl je isti kao za zadatak 1.

- U **Compute** komandi izračunati pointbiserijsku korelaciju između matematičke pismenosti (varijabla **PV1MATH**) i pola učenika (varijabla **ST004D01T**) korišćenjem obrasca:

$$r_{pb} = \frac{M_{x2y} - M_{x1y}}{S_y} * \sqrt{\frac{p1 * p2}{N}}$$

U obrascu je r_{pb} oznaka za pointbiserijsku korelaciju, X je pol, Y matematička pismenost, 1 i 2 su kodirane kategorije (dakle, proveriti Values za pol u Variable View), M je aritmetička sredina, S je standardna devijacija, p je proporcija, N je broj ispitanika.

- Izračunati koeficijent linearne korelacije (**Analyze/Correlate/Bivariate Correlation**) između matematičke pismenosti (varijabla **PV1MATH**) i pola učenika (varijabla **ST004D01T**) i uporedite ga sa pointbiserijskom korelacijom;

Šta zaključujemo o povezanostima?

Zadatak 3.

U varijabli **faksbar1roditelj** za svakog učenika dati su podaci o tome da li je bar jedan od dva roditelja univerzitetski obrazovan:

0 - NE, nijedan;

1 - DA, bar jedan.

U varijabli **matanivo2ivise** da li su postigli nivo funkcionalne pismenosti na testu iz matematike:

0 - nivo 1, odnosno funkcionalno nepismenij;

1 - matematička pismenost na nivoima od 2 do 6.

- Tabelom predstaviti bivarijatnu raspodelu varijabli **faksbar1roditelj** i **matanivo2ivise** i izračunati mere asocijacije ovih dveju varijabli (C koeficijent i Kramerov V koeficijent=fi koeficijent u ovom slučaju).

Za prikaz raspodele koristi se komanda **Analyze→Descriptive Statistics→Crosstabs**. Varijablu **faksbar1roditelj** stavite u redove, a varijablu **matanivo2ivise** u kolone tabele kontingencije. Ista komanda se koristi i za računanje mera asocijacije tako što se štikliraju traženi statistici u podmeniju **Statistics**.

- Izračunati šanse da učenik postigne funkcionalnu pismenost ukoliko mu je bar jedan roditelj visoko obrazovan;
- Izračunati šanse da učenik postigne funkcionalnu pismenost ukoliko mu nijedan roditelj nije visoko obrazovan;
- Izračunati količnik šansi da postignu funkcionalnu pismenost u odnosu da ne postignu funkcionalnu pismenost za učenike kojima nijedan roditelj nije visoko obrazovan, u odnosu na one kojima je bar jedan roditelj visoko obrazovan.

Zadatak 4.

U ovom i sledećem zadatku ne koristimo PISA bazu tako da je možete ugasiti. U fajlu **nlin.sav** (takođe skinuti sa mudla) nalaze se podaci o motivisanosti za obavljanje jednog zadatka (varijabla **motiv**) i o uspešnosti u obavljanju tog zadatka (varijabla **uspeh**).

- Grafički prikazati vezu između ovih varijabli (varijabla **motiv** ide na X osu, a varijabla **uspeh** na Y osu);
- Izračunati kovarijansu između motivisanosti i uspeha;
- Izračunati koeficijent linearne korelacije između motivisanosti i uspeha;

Šta zaključujemo o povezanostima?

Zadatak 5.

- Napraviti varijablu **motiv2 = -6.82+4.13*motiv -0.19* motiv**2**;
- Izračunati koeficijent linearne korelacije između varijable **motiv2** i uspeha;

Šta zaključujemo o povezanostima?