
LV3 – STATISTIČKO ZAKLJUČIVANJE 2 – ZADACI

Opis zadatka (za oba zadatka)

1. Koristiti podatke iz danih tablica
2. Napisati izvještaj o dobivenim rezultatima i grafički ih prikazati.
3. Objasniti što ste koristili i zašto. Ako ste koristili ugrađene python funkcije zašto ste baš njih odabrali?
4. Prikazati deskriptivnu statistiku i usporediti srednje vrijednosti i odgovarajuće parametre.

Napomena

- Za teorijsku osnovu koristiti prezentacije i predloške s predavanja i predložak s LV2
- Na temelju odgovora na sljedeća pitanja možemo možemo odlučiti koji ćemo test koristiti:
 1. Jesu li vrijednosti koje ispitujemo kvalitativni ili kvantitativni?
 2. Ako su kvantitativni, prati li uzorak normalnu distribuciju?
 3. Proučavamo li jedan uzorak ili uspoređujemo dva ili više?
 4. Jesu li uzorci koje proučavamo zavisni ili nezavisni?
 5. Što se testira? Testiranje:
 - (A) Razlike (srednje vrijednosti, proporcije, medijan)
 - (B) Povezanosti (korelacija, regresija)
 - (C) Distribucije

Literatura:

1. <https://medium.com/@byanalytixlabs/parametric-vs-non-parametric-test-which-one-to-use-for-hypothesis-testing-2aa940c92c2b>
2. <https://www.healthknowledge.org.uk/public-health-textbook/research-methods/1b-statistical-methods/parametric-nonparametric-tests>
3. <https://www.kaggle.com/code/saurav9786/statistics-for-data-scientists>

Zadatak 1 Zaključivanje na malom uzorku

Proučava se preciznost mjernog instrumenta mjerenjem masa dvaju različitih listova papira. Ponavljajući mjerenja više puta dobiveni su prikazani podaci u excel tablici. Pretpostavite da uzorci dolaze iz normalne distribucije.

Papir 1	Papir 2
3.481	3.258
3.477	3.254
3.47	3.256
3.448	3.249
3.472	3.241
3.47	3.254
3.485	3.247
3.464	3.257
3.477	3.239
3.475	3.25
3.472	3.258
3.473	3.239
3.472	3.245
3.47	3.24
3.474	3.254

- Na razini značajnosti od 5% testirajte jesu li prosječne mase dva različita papira jednake.
Objasniti što se testira u problemu i zašto ste koristili odgovarajuća funkciju za izračunavanje.
- Ponovite prethodni test na razini značajnosti od 1%
- Usporedite svoje odgovore iz dijelova (b) i (c) te objasnite zašto su isti ili različiti.
- Objasnite kako se na pitanja iz dijelova (b) i (c) može odgovoriti pomoću intervala pouzdanosti.

Zadatak 2 Parametarski/neparametarski testovi

U istraživanju su sudjelovali ispitanici koji su prošli specifični program/intervenciju za poboljšanje vještina. Prikupljeni su rezultati testova *prije* (Pre) i *poslije* (Post) sudjelovanja u programu.

1. Analiza podataka:

- Učitajte priloženi skup podataka iz excel tablice.
- Provjerite osnovne statistike rezultata prije i poslije (medijan, kvartile, raspon, itd.) – deskriptivna statistika

2. Odabir metode:

- Odredite jesu li podaci normalno distribuirani.
- Koristiti odgovarajuću metodu kako bi odgovorili na pitanje postoji li razlika između skupina podataka (npr. prije i poslije intervencije).

3. Vizualizacija podataka:

- Prikazujte podatke koristeći boxplot ili histogram razlika.

4. Interpretacija rezultata:

- Napišite kratki izvještaj s odgovorom na pitanje: *Postoji li značajna razlika između rezultata prije i poslije?*