

---

## Laboratorijska vježba 2: Statističko zaključivanje

---

### Ciljevi vježbe

- studenti će uvježbati osnovne tehnike parametarskog i neparametarsko zaključivanja.
- Primijeniti različite metode zaključivanja na jednoj i dvjema varijablama.
- Provesti testiranje hipoteza o distribucijama i srednjim vrijednostima, kao i razumjeti razliku između parametarskih i neparametarskih metoda.

### Python biblioteka:

Biblioteka **scipy.stats** je dio SciPy paketa koji pruža alate za statističku analizu. Neke od glavnih funkcionalnosti uključuju:

- **Distribucije:** scipy.stats omogućava rad s različitim statističkim distribucijama, poput normalne (stats.norm), t-distribucije (stats.t), i mnogih drugih. Te distribucije možete koristiti za generiranje slučajnih brojeva, računanje gustoća vjerojatnosti, i kvantila.
- **Testiranje hipoteza:** pruža funkcije za statističke testove, uključujući:
  - **t-test:** testiranje razlika između srednjih vrijednosti
    - `stats.ttest_1samp()`: Ovaj test koristi se za testiranje da li se srednja vrijednost uzorka značajno razlikuje od neke pretpostavljene vrijednosti.  
**Primjer:** Provjera je li prosječna ocjena značajno različita od očekivanih 75 bodova.
      - Ovaj test ne koristiti za testiranje, jer ne podržava sve uvjete za izračunavanje kada je  $n < 30$  (provjeriti u dokumentaciji metode i isprobati ručno!)
      - `stats.ttest_ind()`: Koristi se za usporedbu srednjih vrijednosti dviju nezavisnih skupina. Primjer: Usporedba rezultata dviju grupa koje su koristile različite metode učenja.
      - `stats.ttest_rel()`: Koristi se za **paired t-test** (parni t-test), odnosno usporedbu srednjih vrijednosti zavisnih uzoraka (npr. mjerjenja prije i poslije tretmana na istim ispitanicima). Ovaj test se koristi za usporedbu dviju zavisnih uzoraka, odnosno kada imate dva skupa podataka koji su povezani, kao što su mjerjenja prije i poslije na istim subjektima ili parovi podataka koji dolaze iz iste populacije.
  - **z-test:** ne postoji direktna funkcija za **z-test** nego se koristi `stats.norm` (za izračunavanje **z-vrijednosti** i **p-vrijednosti**. **z-test** se koristi kada su **varijance poznate** i veliki uzorak ( $n > 30$ ))
    - npr: `stats.norm.ppf` za izračun kritične vrijednosti z-distribucije

- **Mann-Whitney U test** (`stats.mannwhitneyu`): Neparametrijski test za usporedbu rangova dviju skupina. Koristi se kada podaci **nisu normalno distribuirani** ili su **ordinalni**.
  - **Shapiro-Wilk test** (`stats.shapiro`) koristi se za provjeru **normalnosti podataka**. Ako je **p-vrijednost** manja od 0.05, podaci **nisu normalno distribuirani**.
  - **Kolmogorov-Smirnov test** (`stats.kstest`) koristi se za usporedbu distribucije uzorka s referentnom distribucijom (najčešće normalnom). Može se koristiti i za provjeru normalnosti.
  - **Analiza varijance (ANOVA)** - ANOVA se koristi za usporedbu srednjih vrijednosti **više od dvije grupe**. Pomaže nam utvrditi postoji li statistički značajna razlika između grupa, ali ne i koja grupa se razlikuje. Nakon značajnog rezultata ANOVA testa, koristi se **post-hoc analiza** (npr. Tukeyjev test) za utvrđivanje specifičnih razlika.
- **Procjene intervala povjerenja:** Funkcije za izračun intervala povjerenja koristeći različite distribucije, npr. `stats.t.interval` za t-distribuciju.
  - **Statističke funkcije:** `scipy.stats` pruža razne funkcije za izračunavanje statističkih vrijednosti poput srednje vrijednosti, standardne devijacije, moda, i drugih.

**Detaljnije pogledati na:** <https://docs.scipy.org/doc/scipy/tutorial/stats.html>

## Parametarsko zaključivanje – intervali povjerenja i testiranje hipoteza

- Paziti koliki je uzorak i je li standardna devijacija populacije poznata ili ne
  - Izračunavanje intervala povjerenja koristeći Python, na primjer: za srednju vrijednost populacije kada je varijanca nepoznata (koristeći t-distribuciju):
1. **primjer:** za srednju vrijednost populacije kada je varijanca nepoznata (koristeći t-distribuciju) – usporediti s formulama iz ViS-a

```
import numpy as np
import scipy.stats as stats

# Podaci o uzorku
podaci = [22, 25, 20, 23, 24, 19, 21]
n = len(podaci) # Veličina uzorka (PAZITI! Manji je od 30)
srednja_vrijednost = np.mean(podaci) # Srednja vrijednost uzorka
standardna_devijacija = np.std(podaci, ddof=1) # Procjena standardne
devijacije uzorka (uzorak manji od 30)

# Razina pouzdanosti (npr. 95%)
```

```

razina_pouzdanosti = 0.95
alfa = 1 - razina_pouzdanosti
t_vrijednost = stats.t.ppf(1 - alfa / 2, df=n-1) # Kritična vrijednost t-distribucije za zadani stupanj slobode

# Izračun margine pogreške
margina_pogreske = t_vrijednost * (standardna_devijacija / np.sqrt(n-1))
#PAZITI, FORMULE!

# Interval povjerenja
donja_granica = srednja_vrijednost - margina_pogreske
gornja_granica = srednja_vrijednost + margina_pogreske

print(f"95% interval povjerenja: ({donja_granica:.2f}, {gornja_granica:.2f})")

```

RJ: 95% interval povjerenja: (19.84, 24.16)

## 2. Primjer: Primjer intervalne procjene i testiranja hipoteza

- Koristeći Python i biblioteku scipy, odrediti interval povjerenja za srednju vrijednost populacije i provesti testiranje hipoteze koristeći z-test (poznata varijanca, veliki uzorak), odnosno t-test (nepoznata varijanca, mali uzorak).

**Zadatak:** (a) Na temelju danog uzorka podataka izračunati interval povjerenja:

- (A) za interval povjerenja s pouzdanošću od 95%, (rješenje u prošlom primjeru) i (B) za interval povjerenja s pouzdanošću od 99%.
- (B) Rješenje: 99% interval povjerenja: (18.73, 25.27)  
Objasniti zašto je došlo do ove razlike u intervalima i koja je interpretacija!

(b) Na temelju danog uzorka testirati hipotezu uz pouzdanost od 95% da je prosječna vrijednost jednaka 20.

(c) Na temelju danog uzorka testirati hipotezu uz pouzdanost od 95% da je prosječna vrijednost **veća od 20**.

**Rješenje (b) i (c):**  $t=2.26778$ ,

(b)  $p=0.06386 > 0.05$  (nemamo dovoljno dokaza da je prosječna vrijednost različita od 20) (c)  $p=0.03193$  (odbacujemo nultu hipotezu. Prosječna vrijednost je st.značajno veća od 20).

Primjer koda:

```

import numpy as np

# Podaci uzorka
podaci = [22, 25, 20, 23, 24, 19, 21]
sredina_populacije = 20

```

```

n = len(podaci)

# Izračun osnovnih statistika
srednja_vrijednost_uzorka = np.mean(podaci)
standardna_devijacija_uzorka = np.std(podaci, ddof=1)

# Izračun standardne pogreške, uz podjelu s n-1
standardna_pogreska = standardna_devijacija_uzorka / np.sqrt(n - 1)

# Izračun t-statistike
t_stat = (srednja_vrijednost_uzorka - sredina_populacije) /
standardna_pogreska

print(f"t-statistika: {t_stat}")
p_vrijednost = (1 - stats.t.cdf(t_stat, df=6))
print(f"p_vrijednosti: {p_vrijednost}")

if p_vrijednost < 0.05:
    print("Odbacujemo nul hipotezu: prosječna vrijednost je statistički
značajno veća od 20.")
else:
    print("Ne odbacujemo nul hipotezu: nema dovoljno dokaza da je prosječna
vrijednost veća od 20.")

```

**Napomena:** Probati riješiti zadatak sa `stats.ttest_1samp()` i usporediti rješenja. Pogledati u dokumentaciju kako koriste ovu metodu i što izračunavaju i kako.

## Neparametarsko zaključivanje – razlika između promatranih uzoraka

- **Mann-Whitney U test:** Koristiti za usporedbu dviju nezavisnih grupa kada podaci nisu normalno distribuirani. Alternativa t-testu za zavisne uzorke
- **Wilcoxon rang-sum test:** Koristiti za usporedbu zavisnih uzoraka, odnosno za usporedbu rezultata istih ispitanika u različitim uvjetima.
- **Kruskal-Wallis test:** Koristiti za usporedbu više nezavisnih grupa. Proširenje Mann-Whitney U testa za više od dvije grupe, kada želimo usporediti više uzoraka. **Kruskal-Wallis test** je neparametrijska alternativa ANOVA testu i koristi se kada podaci nisu normalno distribuirani ili kada postoji sumnja na heterogenost varijance između grupa. Umjesto srednjih vrijednosti, Kruskal-Wallis koristi rangove podataka za usporedbu više grupa, što ga čini robusnijim na odstupanja od normalnosti.

## 1. Primjer:

```
from scipy.stats import wilcoxon

prije = [78, 85, 82, 90, 88]
poslije = [80, 87, 85, 92, 90]

stat, p = wilcoxon(prije, poslije)
print(f"P-vrijednost (Wilcoxon): {p}")
if p < 0.05:
    print("Postoji značajna razlika između mjerenja prije i poslije.")
else:
    print("Ne postoji značajna razlika između mjerenja prije i poslije.")
```

## 2. Primjer

```
from scipy.stats import kruskal

grupa_A = [78, 85, 82, 90, 88]
grupa_B = [72, 79, 80, 68, 75]
grupa_C = [85, 89, 90, 92, 88]

stat, p = kruskal(grupa_A, grupa_B, grupa_C)
print(f"P-vrijednost (Kruskal-Wallis): {p}")
if p < 0.05:
    print("Postoji značajna razlika između grupa.")
else:
    print("Ne postoji značajna razlika između grupa.")
```

Testiranje hipoteza o distribucijama, normalnosti i razlike između srednjih vrijednosti

## Testiranje hipoteza o distribuciji podataka

- **Definicija distribucije:** Distribucija podataka opisuje kako su podaci raspoređeni i koja je njihova frekvencija.
- **Cilj testiranja distribucije:** Utvrditi prati li uzorak određenu distribuciju, npr. normalnu distribuciju.
- **Primjer testa:** (*pogledati predavanje!*)
  - **Chi-square test (Hi-kvadrat)** za ispitivanje slaganja podataka s teorijskom distribucijom.

```
from scipy.stats import chisquare

# Primjer podataka
observed_values = [25, 30, 45]
expected_values = [33, 33, 34]
```

```
# Chi-square test
chi2_stat, p_value = chisquare(observed_values, f_exp=expected_values)
print(f"Chi-square statistika: {chi2_stat}, P-vrijednost: {p_value}")
```

## Testiranje normalnosti podataka

- **Cilj:** Utvrditi prati li uzorak **normalnu distribuciju**, što je ključna prepostavka za mnoge parametarske statističke testove.
- **Metode:**
  - **Shapiro-Wilk test i Kolmogorov-Smirnov test.**
  - Ako je **p-vrijednost** manja od 0.05, odbacujemo hipotezu da podaci prate normalnu distribuciju.

*Primjer zadatka:*

Studentima dati uzorak podataka i zatražiti ih da koriste **Shapiro-Wilk test** kako bi provjerili prati li uzorak normalnu distribuciju.

```
from scipy.stats import shapiro

# Primjer podataka
podaci = [22.5, 23.1, 21.9, 20.7, 25.3]

# Shapiro-Wilk test
stat, p_value = shapiro(podaci)
print(f"Shapiro-Wilk statistika: {stat}, P-vrijednost: {p_value}")
```

## Testiranje razlika između srednjih vrijednosti

- **Parametarski testovi:**
  - **t-test za zavisne uzorke:** Koristi se kada želimo usporediti srednje vrijednosti istih ispitanika prije i poslije intervencije.
  - **t-test za nezavisne uzorke:** Koristi se kada uspoređujemo srednje vrijednosti dviju različitih grupa.
- **Neparametrijski testovi:**
  - **Mann-Whitney U test:** Koristi se za usporedbu medijana dviju grupa kada podaci nisu normalno distribuirani.
- **One-way ANOVA** (jednofaktorska ANOVA) je statistički test koji se koristi za usporedbu **srednjih vrijednosti više od dvije nezavisne grupe**. Cilj je utvrditi postoji li statistički značajna razlika između srednjih vrijednosti grupa na temelju jednog faktora ili nezavisne varijable.
- **Ključne značajke One-way ANOVA:**

- **Jedan faktor:** Koristi se za usporedbu grupa koje su definirane jednom nezavisnom varijablu (npr. različite metode učenja, različiti lijekovi, itd.).
  - **Usporedba više od dvije grupe:** Za razliku od t-testa, koji može usporediti samo dvije grupe, ANOVA se koristi za usporedbu tri ili više grupa istovremeno.
  - **Pretpostavke:** One-way ANOVA prepostavlja da su podaci normalno distribuirani unutar svake grupe i da su varijance među grupama jednake (homogenost varijance).
- **Primjer kada koristiti One-way ANOVA** - Želimo usporediti prosječne rezultate studenata koji su koristili tri različite metode učenja kako bismo utvrdili koja je metoda najučinkovitija. **Ako test pokazuje da postoji statistički značajna razlika**, to znači da barem jedna grupa ima drugačiju srednju vrijednost, ali ne govori nam koja točno grupa se razlikuje. Za daljnje analize koristi se **post-hoc test** (npr. Tukeyjev test) kako bi se utvrdile specifične razlike.
- **Kruskal-Wallis test** je neparametrijska alternativa One-way ANOVA, koja se koristi kada podaci nisu normalno distribuirani ili kada varijance među grupama nisu jednake.
- **Two-way ANOVA:** Koristi se kada imamo **dvije nezavisne varijable** (faktore) i želimo vidjeti kako oni zajedno utječu na zavisnu varijablu. Ova metoda također omogućuje ispitivanje interakcija između dvaju faktora.

**Primjer:** Želite testirati utjecaj i spola i različitih metoda učenja na akademski uspjeh. Faktor 1 je metoda učenja (npr. tri različite metode), a faktor 2 je spol (muškarci i žene). Two-way ANOVA vam pomaže utvrditi postoji li učinak metoda učenja, postoji li učinak spola i postoji li interakcija između metoda učenja i spola.

- **Repeated Measures ANOVA:** Koristi se kada se **isti ispitanici** mjere više puta pod različitim uvjetima (npr. različita vremena ili različiti tretmani). Ovo je slično t-testu za zavisne uzorke, ali se koristi kada postoje više od dvije mjere.

**Primjer:** Ako provodite test s istim ispitanicima u tri različita vremena (npr. prije tretmana, odmah nakon tretmana i mjesec dana nakon tretmana), koristite Repeated Measures ANOVA za utvrđivanje postoji li promjena između tih mjeranja.

- **Multivariate ANOVA (MANOVA):** Proširenje ANOVA gdje postoji **više zavisnih varijabli**. MANOVA analizira utjecaj nezavisnih varijabli na više zavisnih varijabli istovremeno.

## Ključne razlike:

- **One-way ANOVA:** Koristi se za ispitivanje razlika između više grupa temeljenih na jednom faktoru.
- **Two-way ANOVA:** Ispituje učinak dvaju faktora i njihovu interakciju.
- **Repeated Measures ANOVA:** Koristi se kada se ispitanici mjere više puta.
- **MANOVA:** Proširuje ANOVA na više zavisnih varijabli.

```

from scipy.stats import ttest_ind, mannwhitneyu
import numpy as np

# Primjer podataka
grupa_A = [78, 85, 82, 90, 88]
grupa_B = [72, 79, 80, 68, 75]

# t-test za nezavisne uzorke
stat, p = ttest_ind(grupa_A, grupa_B)
print(f"P-vrijednost (t-test): {p}")
if p < 0.05:
    print("Postoji značajna razlika između skupina (t-test).")
else:
    print("Ne postoji značajna razlika između skupina (t-test).")

# Mann-Whitney U test
stat, p = mannwhitneyu(grupa_A, grupa_B)
print(f"P-vrijednost (Mann-Whitney): {p}")
if p < 0.05:
    print("Postoji značajna razlika između skupina (Mann-Whitney).")
else:
    print("Ne postoji značajna razlika između skupina (Mann-Whitney).")

```

### Primjer ANOVA I post-hoc test

```

import pandas as pd
import numpy as np
import scipy.stats as stats
from statsmodels.stats.multicomp import pairwise_tukeyhsd
import matplotlib.pyplot as plt

# Kreiranje DataFrame-a s podacima temeljenim na primjeru s linka
data = {'Plant': ['A', 'A', 'A', 'B', 'B', 'B', 'C', 'C', 'C'],
        'Weight': [4.17, 5.18, 5.18, 6.11, 5.00, 4.92, 6.31, 5.64, 5.14]}
df = pd.DataFrame(data)

# ANOVA analiza koristeći scipy.stats
f_stat, p_value = stats.f_oneway(df[df['Plant'] == 'A']['Weight'],
                                   df[df['Plant'] == 'B']['Weight'],
                                   df[df['Plant'] == 'C']['Weight'])
print(f"F-statistika: {f_stat}, P-vrijednost: {p_value}")

# Objasnjenje rezultata ANOVA testa
if p_value < 0.05:
    print("P-vrijednost je manja od 0.05, što znači da postoji statistički značajna razlika između barem dvije grupe.")
else:
    print("P-vrijednost je veća od 0.05, što znači da ne postoji statistički značajna razlika između grupa.")

```

```

# Post-hoc test (Tukeyjev HSD test)
tukey_result = pairwise_tukeyhsd(endog=df['Weight'], groups=df['Plant'],
alpha=0.05)
print("Tukey post-hoc test:", tukey_result)

# Objasnjenje rezultata Tukeyjevog testa
print("Objasnjenje rezultata Tukeyjevog testa:")
for res in tukey_result.summary().data[1:]:
    group1, group2, mean_diff, p_adj, lower, upper, reject = res
    if reject:
        print(f"Grupa {group1} se statistički značajno razlikuje od grupe
{group2} (P-vrijednost prilagođena: {p_adj}).")
    else:
        print(f"Grupa {group1} se ne razlikuje statistički značajno od grupe
{group2} (P-vrijednost prilagođena: {p_adj}).")

# Prikaz rezultata post-hoc testa
tukey_result.plot_simultaneous(figsize=(8, 6))
plt.show()

# Objasnjenje rezultata:
# ANOVA analiza:
# - F-statistika: Mjeri omjer varijance između grupa i varijance unutar
grupa. Veća vrijednost F-statistike ukazuje na veće razlike između grupa u
odnosu na varijabilnost unutar grupa.
# - P-vrijednost: Ako je p-vrijednost manja od razine značajnosti (npr.
0.05), možemo zaključiti da postoji statistički značajna razlika između barem
dvije grupe.
#
# Tukeyjev HSD post-hoc test:
# - Ako ANOVA test pokaže značajnu razliku, Tukeyjev test pomaže u
određivanju koja od grupa se međusobno razlikuje.
# - Rezultati post-hoc testa pokazuju parove grupa i jesu li razlike između
njih statistički značajne.
# - Grafički prikaz rezultata post-hoc testa omogućava vizualizaciju
intervala povjerenja za razlike između grupa.

```

Primjer 2, možete pronaći na linku:

<https://www.kaggle.com/code/haseebwar07/hypothesis-testing-and-anova-post-hoc-with-python>

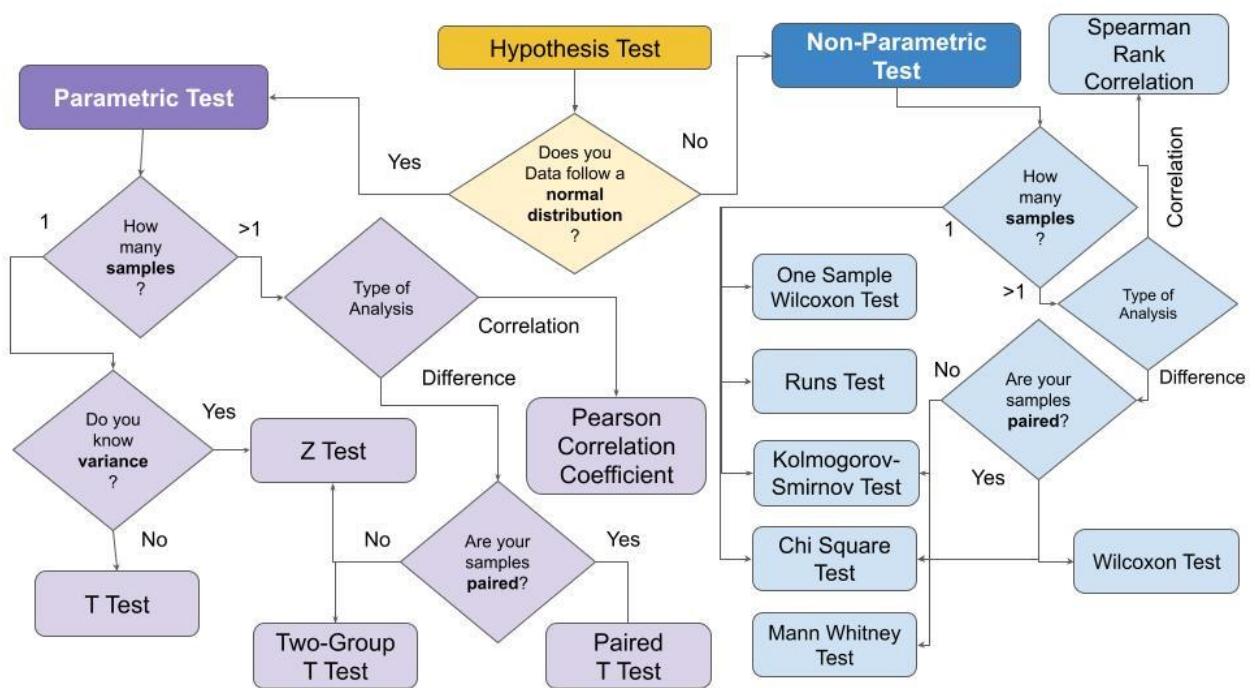


Figure 1(izvor: Testing Your Hypotheses: A Practical Guide to Parametric and Non-Parametric Tests in Quantitative Research Design )

## Classification of Hypothesis Test

Type of Test	Level of Measure	Sample characteristics				Correlation	
		One Sample	Two Samples		K samples (i.e. >2)		
			Independent	Dependent	Independent	Dependent	
Parametric Test	Interval or Ratio	Z-test or t test	Independent sample t-test	Paired sample t-test	One way ANOVA	Repeated measure ANOVA	Pearson's r test
Non-Parametric Test	Categorical or nominal	Chi square test	Chi square test	McNemar test	Chi square test	Cochran's Q Test	Spearman (ρ) test
	Rank or ordinal	Chi square test	Mann-Whitney u-test	Wilcoxon signed-rank test	Kruskal Wallis test	Friedman's test	

Figure 2(izvor: <https://www.lismcqspractice.com/2020/12/parametric-test-vs-non-parametric-test.html> )