

Statističke metode u biologiji

Statistika

- grana primijenjene matematike koja se bavi analizom podataka
- poznavanja određenih svojstava uzorka donosi zaključke o svojstvima populacije

Osnovne statističke vrijednosti

1. Mjere centralne tendencije

a) Srednja vrijednost (aritmetička sredina)

Osnovna formula za izračunavanje aritmetičke sredine glasi:

ARITMETIČKA SREDINA =SUMA SVIH REZULTATA / BROJ REZULTATA

$$\bar{X} = \Sigma X / N$$

Osnovne statističke vrijednosti

1. Mjere centralne tendencije

b) Medijan - vrijednost koja se u nizu rezultata, poredanih po veličini, nalazi točno u sredini

- prednost medijana pred aritmetičkom sredinom je u tome što na nju ne utječe vrijednost pojedinih rezultata, pa prema tome jedan vrlo ekstremni rezultat neće promijeniti vrijednost medijana, koja je uvjetovana samo brojem rezultata

Osnovne statističke vrijednosti

2. Mjere raspršenja podataka

a) Varijanca

Varijanca (s) predstavlja mjeru raspršenja rezultata oko aritmetičke sredine, odnosno prosječnu kvadriranu udaljenost svih rezultata od aritmetičke sredine.

$$s = \sum (X - \bar{X})^2 / N$$

Osnovne statističke vrijednosti

2. Mjere raspršenja podataka

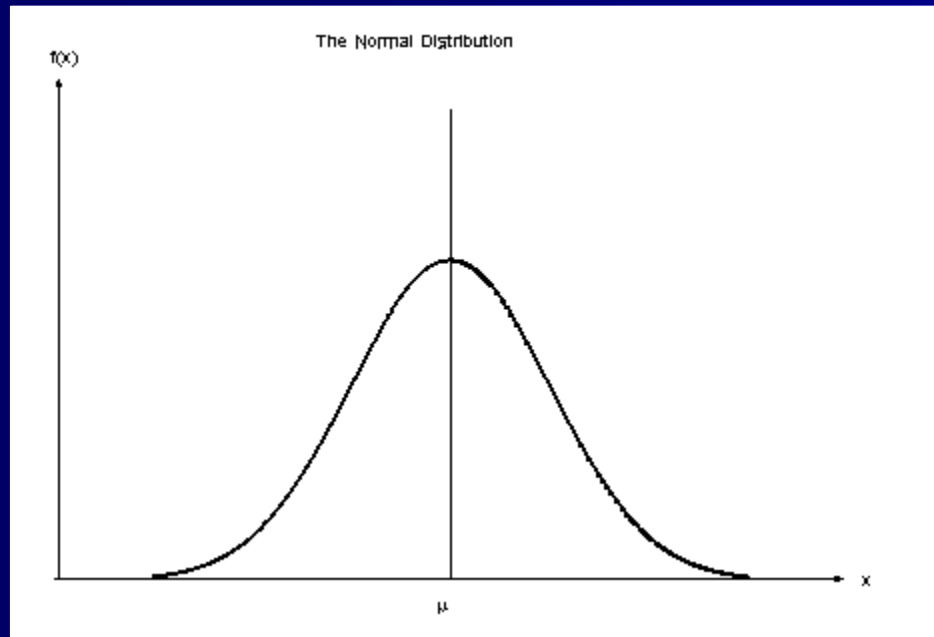
b) Standardna devijacija

Drugi korijen iz varijance. $SD = \sqrt{s}$

SD je mjera koja se najčešće koristi za prikaz varijabiliteta rezultata.

Parametrijske i neparametrijske metode u statistici

Za korištenje parametrijskih metoda, podaci moraju imati normalnu raspodjelu koja odgovara unimodalnoj, Gaussovoj krivulji.



Gaussova krivulja:

- u potpunosti je definirana aritmetičkom sredinom i standardnom devijacijom
- aritmetička sredina određuje središte, a standardna devijacija širinu
- unimodalna je i simetrična

Parametrijske i neparametrijske metode u statistici

1. Parametrijske metode :

a) t- test

- testiranje statistički značajne razlike između dvije aritmetičke sredine

b) analiza varijance (ANOVA)

- usporedba više uzoraka pri čemu svaki uzorak predstavlja osnovni skup (populaciju)

- provjeravaju se promjene aritmetičkih sredina uzoraka

Parametrijske i neparametrijske metode u statistici

Ako podaci nemaju unimodalnu raspodjelu, koristimo

2. Neparametrijske metode:

- a) Mann-Whitney U test - ekvivalent t-testu
 - testiranje statistički značajne razlike između medijana od dvije grupe podataka

Parametrijske i neparametrijske metode u statistici

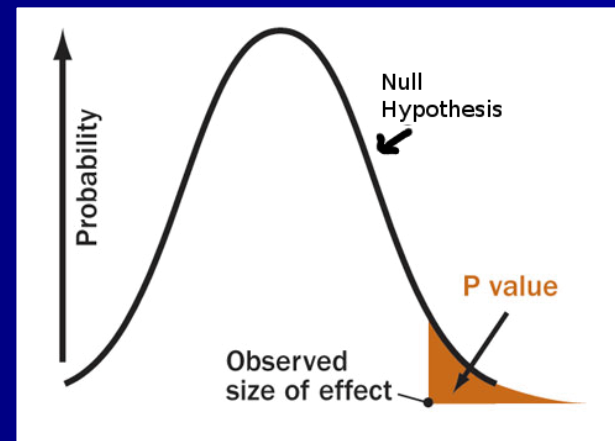
Ako podaci nemaju unimodalnu raspodjelu, koristimo

2. Neparametrijske metode:

- b) Kruskal Wallis test
 - ekvivalent analizi varijance
 - testiranje statistički značajne razlike između medijana više od dvije grupe podataka

p vrijednost

- p vrijednost nam ukazuje da li se proučavani skupovi podataka razlikuju ili ne u odnosu na našu hipotezu
- Razina statističke značajnosti ovisi o svojstvu koje analiziramo i uvjetima istraživanja
- Najčešća razina značajnosti iznosi 0,05 (5%) pa ukoliko p iznosi manje od 0,05 ($p < 0,05$) razlike proglašavamo statistički značajnim
- Suprotno navedenom ako je $p > 0,05$ razlike proglašavamo statistički neznačajnim



Programski paketi

Microsoft Office Excel

Statistica

Primer

Canoco

R

Biološka raznolikost

Raznolikost vrsta na pojedinom staništu može se definirati brojem različitih vrsta, brojnošću i strukturom. Navedene komponente uključene su u različite indekse:

Shannonov indeks

Simpsonov indeks

Sörensenov indeks (0-1)

Margalefov indeks

Pielouov indeks (0-1)

Biološka raznolikost

Raznolikost vrsta na pojedinom staništu može se definirati brojem različitih vrsta, brojnošću i strukturom. Navedene komponente uključene su u različite indekse:

Shannonov indeks (H')

H – indeks raznolikosti

n_i – broj jedinki jedne vrste

N – ukupan broj svih jedinki

p_i – vjerojatnost pojavljivanja za vrstu i

$$H' = - \sum p_i / \ln p_i$$

$$p_i = n_i / N$$

Biološka raznolikost

Raznolikost vrsta na pojedinom staništu može se definirati brojem različitih vrsta, brojnošću i strukturom. Navedene komponente uključene su u različite indekse:

Simpsonov indeks (D)

$$D = \sum p_i^2$$

- jednostavan, koristi se kao alternativa ostalim indeksima koji se uglavnom teško izračunavaju

$$p_i = n_i/N$$

n_i – broj jedinki jedne vrste
 N – ukupan broj svih jedinki

Biološka raznolikost

Raznolikost vrsta na pojedinom staništu može se definirati brojem različitih vrsta, brojnošću i strukturom. Navedene komponente uključene su u različite indekse:

Sörensenov indeks ili kvocijent sličnosti

- uzima u obzir kvalitativnu zastupljenost vrsta
- vrijednosti kvocijenta sličnosti izražavaju se u postocima
- što je veći QS to je struktura uspoređivanih zajednica sličnija

QS – kvocijent sličnosti

a – broj vrsta zajednice A

b – broj vrsta zajednice B

c – broj zajedničkih vrsta

$$QS = (2c/a+b) * 100$$

Biološka raznolikost

Raznolikost vrsta na pojedinom staništu može se definirati brojem različitih vrsta, brojnošću i strukturom. Navedene komponente uključene su u različite indekse:

Margalefov indeks bogatstva vrsta (d)

$$d = (S - 1) / \ln N$$

S = ukupan broj vrsta

N = ukupan broj jedinki

Biološka raznolikost

Raznolikost vrsta na pojedinom staništu može se definirati brojem različitih vrsta, brojnošću i strukturom. Navedene komponente uključene su u različite indekse:

Pielouov indeks ujednačenosti (J')

$$J' = H' / \ln S$$

H' = Shannonov indeks

S = ukupan broj vrsta u uzorku