

VAJE 1: Osnovni pojmi statistike

Na računalniških vajah se za urejanje in prikazovanje statističnih podatkov uporabi statistični programski paket SPSS in podatkovna datoteka *podatki.sav*.

NALOGE:

1. Ugotovi, koliko statističnih enot in koliko statističnih spremenljivk vsebuje dana podatkovna datoteka. Kaj merijo statistične spremenljivke?
2. Izdelaj frekvenčno porazdelitev statističnih spremenljivk *Starost* in *KoličinaTD* ter ju prikaži s histogramom (uporabi postopek *Analyze - Descriptive Statistics - Frequencies - Charts - Histogram*).
3. Izdelaj frekvenčno porazdelitev statistične spremenljivke *Rtekočine* in jo prikaži s stolpčnim diagramom (uporabi postopek *Analyze - Descriptive Statistics - Frequencies - Charts - Bar charts*).
4. Izdelaj frekvenčno porazdelitev statistične spremenljivke *Šopz* in jo prikaži s stolpčnim diagramom za deleže (uporabi postopek *Analyze - Descriptive Statistics - Frequencies - Charts - Bar charts - Percentages*).
5. Izdelaj frekvenčno porazdelitev statističnih spremenljivk *Šport* in *Rasa* ter ju prikaži s strukturnim krogom (uporabi postopek *Analyze - Descriptive Statistics - Frequencies - Charts - Pie charts*).
6. Glede na lastnost merjenja, ki jo ima statistična spremenljivka, ugotovi, kakšne vrste so statistične spremenljivke *Starost*, *KoličinaTD*, *Šopz*, *Šport* in *Rasa*. Klasificiraj tudi omenjene spremenljivke glede na zalogo vrednosti, ki jo zavzame.
7. Izdelaj tabelo, v kateri bo prikazan frekvenčni in odstotkovni odnos spremenljivk *KoličinaPT* in *Šport* ter porazdelitev prikaži tudi s stolpčnim diagramom. Interpretiraj dane podatke (uporabi postopek *Analyze - Descriptive Statistics - Crosstabs* in v *Cells* označi *Percentages Total* ter označi tudi *Display clustered bar charts*).
8. Podobno kot pri prejšnji nalogi izdelaj tabelo, v kateri bo prikazana frekvenčna porazdelitev spremenljivk *KoličinaPT* in *Šport* glede na spremenljivko *Rasa*. Porazdelitev prikaži tudi s stolpčnim diagramom (uporabi postopek *Analyze*

- *Descriptive Statistics - Crosstabs - Row (KoličinaPT) - Column (Šport) - Layer (Rasa)*).

Teoretično ozadje

Statistika

Ohlapno lahko rečemo, da je statistika veda, ki kvantitativno proučuje množične pojave v naravi in družbi, se ukvarja z zbiranjem in razvrščanjem podatkov, pripravo statističnih eksperimentov, vzorčenjem, statističnim sklepanjem. Statistiko delimo na opisno in analitično. Opisna statistika obsega področji zbiranja in razvrščanja podatkov, medtem ko analitična statistika obsega predvsem področji vzorčenja in statističnega sklepanja (analiza rezultatov raziskovalnega dela). Statistika kot znanost je del matematike. Matematična statistika znanstveno proučuje in razvija statistična orodja, ki se potem uporabljajo v drugih znanostih in strokah npr. medicini, ekonomiji, družboslovju. Prav tako posamezne veje znanosti razvijajo svoje posebne statistične metode, ki jih terja obravnava specifičnih problemov stroke. Zaradi tega se v literaturi pogosto statistika deli na različne veje.

Statistična populacija in vzorec

V statistiki imamo opravka s končno ali neskončno množico G , ki jo imenujemo statistična populacija, na kateri opazujemo ali merimo neko merljivo količino X , ki jo imenujemo statistični znak. Statistično populacijo G sestavljajo statistične enote. Podmnožici H statistične populacije G rečemo vzorec. Vzorec je torej del populacije, ki je praviloma izbran za obravnavo določenih lastnosti te populacije. Npr. statistično populacijo G tvorijo vsi vpisani študenti na univerzi v Mariboru, ki so statistične enote. Vpisani študenti neke fakultete pa predstavljajo vzorec te populacije.

Statistični znaki in statistične spremenljivke

Statistični znak je torej vsaka merljiva lastnost, ki jo imajo enote statistične populacije. Vrednosti posameznih lastnosti se praviloma pri enotah razlikujejo, zato jim pravimo tudi statistične spremenljivke. Statistične spremenljivke se delijo na atributivne ali opisne, katerih vrednosti opisujemo z besedami, in numerične ali številske, katerih vrednosti so realna števila. Glede na lastnost merjenja, ki jo ima statistična spremenljivka, najdemo v literaturi tudi natančnejšo delitev spremenljivk na:

- *Nominalne* ali *imenske*, katere vrednosti lahko le razlikujemo med seboj, dve vrednosti sta enaki ali nista, npr. spol.
- *Ordinalne* ali *urejenostne*, katere vrednosti lahko razvrščamo po velikosti, torej lahko primerjamo katera vrednost je večja oz. manjša, npr. izobrazba.
- *Intervalne* ali *razmične*, kjer lahko primerjamo tudi razlike med vrednostmi, dve vrednosti sta lahko enaki ali pa je razlika med njima enaka nekemu številu, npr. telesna teža, starost.
- *Razmernostne*, primerjamo lahko razmerja med vrednostmi, dve vrednosti sta enaki ali pa je njun kvocient enak nekemu številu, npr. krvni pritisk.

Omenimo, da matematična statistika pozna delitev statističnih spremenljivk samo glede na zalogo vrednosti, ki jo zavzame statistična spremenljivka. Tako v matematiki spremenljivke delimo na:

- *diskretne*, ki imajo končno ali števno zalogo vrednosti, npr. spol, krvna skupina, izobrazba, učni uspeh;
- *zvezne*, katerih zaloga vrednosti je praviloma interval, npr. krvni pritisk, telesna teža, telesna višina.

Urejanje in prikazovanje statističnih podatkov

Statistične podatke uredimo s pomočjo tabel in grafičnih prikazov. Pri tem uporabljamo ranžirno vrsto, frekvenčno tabelo individualnih vrednosti, frekvenčno tabelo grupiranih vrednosti, kontingenčno tabelo, strukturne stolpce, strukturne kroge, histogram, poligon.

Literatura

- [1] D. Benkovič, Vaje iz biostatistike, Medicinska fakulteta Univerze v Mariboru.
- [2] R. Jamnik, Matematična statistika, Državna založba Slovenije, Ljubljana 1980.
- [3] J. Sagadin, Statistične metode za pedagoge, Obzorja, Maribor 2003.