REPUBLIKA SLOVENIJA
MINISTRSTVO ZA ŠOLSTVO IN ŠPORT

Naložba v vašo prihodnost
OPERACIJO DELNO FINANCIRA EVROPSKA UNIJA
Evropski socialni sklad

STATISTIKA

MARIKA ŠADL

Višješolski strokovni program: Ekonomist
Učbenik: Statistika (2.del predmeta Poslovna matematika s statistiko)
Gradivo za 1. letnik

Avtorica:

Marika Šadl, univ. dipl. ekon.
Ekonomska šola Murska Sobota
Višja strokovna šola

Strokovna recenzentka:
Mag. Majda Bukovnik, univ. dipl. ekon.

Lektorica:
Cvetka Mencigar Rituper, prof. slov. j.

CIP - Kataložni zapis o publikaciji
Narodna in univerzitetna knjižnica, Ljubljana

311.1(075.8)(0.034.2)

ŠADL, Marika
Statistika [Elektronski vir] : gradivo za 1. letnik / Marika Šadl. - El. knjiga. - Ljubljana : Zavod IRC, 2008. - (Višješolski strokovni program Ekonomist / Zavod IRC)

Način dostopa (URL): http://www.zavod-irc.si/docs/Skriti_dokumenti/Statistika-Sadl.pdf. - Projekt Impletum

ISBN 978-961-6820-54-7

249248768

Izdajatelj: Konzorcij višjih strokovnih šol za izvedbo projekta IMPLETUM
Založnik: Zavod IRC, Ljubljana.
Ljubljana, 2008

Strokovni svet RS za poklicno in strokovno izobraževanje je na svoji 120. seji dne 10. 12. 2009 na podlagi 26. člena Zakona o organizaciji in financiranju vzgoje in izobraževanja (Ur. l. RS, št. 16/07-ZOFVI-UPB5, 36/08 in 58/09) sprejel sklep št. 01301-6/2009 / 11-3 o potrditvi tega učbenika za uporabo v višješolskem izobraževanju.

Gradivo je sofinancirano iz sredstev projekta Impletum 'Uvajanje novih izobraževalnih programov na področju višjega strokovnega izobraževanja v obdobju 2008–11'.
Projekt oz. operacijo delno financira Evropska unija iz Evropskega socialnega sklada ter Ministrstvo RS za šolstvo in šport. Operacija se izvaja v okviru Operativnega programa razvoja človeških virov za obdobje 2007–2013, razvojne prioritete 'Razvoj človeških virov in vseživljenjskega učenja' in prednostne usmeritve 'Izboljšanje kakovosti in učinkovitosti sistemov izobraževanja in usposabljanja'.

KAZALO

Predgovor

Knjiga je namenjena študentom višjih strokovnih šol – program Ekonomist. Večina vas nekaj znanja že ima, saj prihajate iz srednjih šol, v katerih je statistika obvezen predmet. Nekateri pa se s predmetom srečujete prvič. Vendar vam beseda statistika, predvsem zaradi pomena zbiranja in objavljanja statističnih podatkov, ni neznana.

Pripravljena je kot študijsko gradivo in je usklajena s katalogom znanj za del predmeta poslovna matematika s statistiko – statistika. Omogoča vam, da razširite in poglobite srednješolsko znanje iz statistike.

Gradivo je razdeljeno na osem poglavij. Vsako je sestavljeno iz razlage, dopolnjene s primerom, kar vam bo olajšalo učenje in zagotovilo večjo uspešnost pri usvajanju znanja. Na koncu poglavja je dodan še primer za utrjevanje predelane snovi.

Študijsko gradivo dopolnjuje zbirka nalog, ki je samostojno gradivo in vam bo dobrodošel pripomoček pri učenju.

Strokovni pregled gradiva je opravila gospa mag. Majda Bukovnik, univ. dipl. ekon. Za vse pripombe in predloge sem ji neizmerno hvaležna, prav tako se zahvaljujem gospe Cvetki Mencigar Rituper, prof. slov. j., ki je popravila moje jezikovne napake.

Hvala tudi vsem kolegom, ki ste me vzpodbujali pri pisanju tega gradiva.

Marika Šadl

Murska Sobota, 2008

UVOD

Statistika je prvotno predstavljala številske podatke, ki so se nanašali na različne družbenoekonomske pojave. Kasneje je dobila širši pomen in postala znanstvena veda, ki se ukvarja z zbiranjem in urejanjem podatkov, razvojem metod za obdelavo in analizo podatkov ter predstavitvijo rezultatov statistične analize.

Predmet statističnega preučevanja so množični pojavi, torej pojavi, ki v času in prostoru nastopajo v velikem številu, ustrezajo določenim opredeljujočim pogojem in se razlikujejo v značilnostih, po katerih jih raziskujemo. O teh pojavih statistika najprej zbere številske in druge podatke ter jih preoblikuje v statistične parametre. To so značilnosti pojavov, ki jih lahko koristno uporabimo pri poslovnem odločanju. Bistvo statistike niso podatki v obliki številk, ampak iskanje pomena teh številk. Tako statistika z različnimi postopki in metodami ugotavlja značilnosti in zakonitosti preučevanih pojavov.

Podatki se lahko nanašajo na različne pojave; tako govorimo na primer o statistiki prebivalstva, statistiki šolstva, statistiki industrije, statistiki kmetijstva, statistiki trgovine, statistiki gostinstva in turizma, statistiki cen in življenjskih stroškov in podobno.

Na osnovi rezultatov statističnih preučevanj gospodarski in drugi subjekti sprejemajo poslovne in druge odločitve. Zato je še kako pomembno, da poznamo vsaj osnovne statistične metode in razlago nekaterih statističnih parametrov, ki kažejo na značilnosti določenih pojavov. Znano je namreč, da imajo statistično neizobraženi pogosto negativen odnos do statistike. Poglejmo naslednja primera:

- Povprečna bruto plača na zaposleno osebo v Sloveniji leta 2007 je bila 1.248,79 evra in je bila nominalno za 5,9 % večja kot povprečna plača leta 2006. Glede na izračunani odstotek bi morali biti prejemniki zadovoljni. In zakaj niso? V tem času so namreč naraščale tudi cene dobrin, z upoštevanjem teh izračunamo realno rast plač, ki je bila le 2,2 %. Seveda pa plače niso naraščale po enaki stopnji v vseh dejavnostih in vsi delavci ne prejemajo povprečne plače, ampak mnogi izmed njih precej nižjo.

- Trideset študentov je opravljalo izpit iz statistike na prvem roku in dvanajst izmed njih oziroma 40 % ga je opravilo z ocenami 9, 9, 8, 6, 6, 6, 7, 10, 7, 8, 7, 8; tako je povprečna ocena 7,6. Če pa upoštevamo tudi ocene tistih, ki izpita niso opravili; od teh jih je dvanajst dobilo oceno 5, trije oceno 4 in trije oceno 3; pa je povprečna ocena vseh, ki so pristopili k izpitu, 5,6. Vidimo, kako pomembno je pravilno opredeliti podatke, iz katerih računamo določene parametre, predvsem pa je pomembna pravilna razlaga izračunanih parametrov.

V obeh primerih je možna tudi zloraba statistike. Če želimo zaposlenim prikazati, da so njihove plače zelo porasle, bomo prikazali le nominalno rast plač, zamolčali pa bomo rast cen in njihov vpliv na kupno moč plač. Prav tako je povprečna ocena študentov pri izpitu, izračunana le iz pozitivnih ocen, precej višja kot tista, izračunana z upoštevanjem vseh ocen, vendar ne kaže prave slike izpitnih rezultatov. Pri povprečni oceni 5,6 bi bilo nujno analizirati razloge za slabe rezultate in tudi sprejeti ukrepe za izboljšanje. Pri povprečni oceni 7,6 pa to ne bi bilo potrebno.

1 TEMELJNI POJMI V STATISTIČNI ANALIZI

V tem poglavju boste usvojili osnovne statistične pojme, kot so: populacija, enota, spremenljivka, parameter. Pri preučevanju določenega pojava v praksi, na primer pri tržnih raziskavah, raziskavah javnega mnenja in podobnih, boste znali opredeliti populacijo, določiti enoto in spremenljivke oziroma značilnosti, po katerih boste enote opazovali oziroma zanje zbirali podatke.

1.1 POPULACIJA, ENOTA IN SPREMENLJIVKA

Vzemimo, da bomo preučevali trgovska podjetja v Sloveniji. Namen preučevanja je ugotoviti število trgovskih podjetij po lastništvu, velikosti (kot kriterij za velikost vzamemo število zaposlenih in promet v prvem polletju) in vrsti trgovinske dejavnosti.

Trgovska podjetja, ki bodo predmet preučevanja, moramo ustrezno opredeliti. Opredelimo jih po treh vidikih: krajevnem, časovnem in vsebinskem.

Primer 1.1:
Preučevali bomo trgovska podjetja po lastništvu, velikosti in vrsti trgovinske dejavnosti v Sloveniji 30. 6. 2008 .

Tako opredeljeno skupino pojavov imenujemo ***populacija***, ki je opredeljena:
- krajevno: *območje republike Slovenije;*
- časovno: *30. 6. 2008; in*
- vsebinsko: *trgovska podjetja po lastništvu, velikosti in vrsti trgovinske dejavnosti.*

Populacijo sestavljajo ***enote***, v našem primeru je enota posamezno trgovsko podjetje.

Populacijo preučujemo z namenom, da ugotovimo določene značilnosti. Najprej ugotavljamo značilnosti enot. Katere značilnosti oziroma lastnosti enot bomo preučevali, je odvisno od namena preučevanja. V našem primeru bodo te: vrsta lastnine, število zaposlenih, promet v prvem polletju leta 2008, vrsta trgovinske dejavnosti.

Lastnosti enot, po katerih jih opazujemo, imenujemo ***spremenljivke***.

Spremenljivka ima pri vsaki opazovani enoti neko ***vrednost***. Lastnina je lahko individualna, zasebna, delniška, zadružna. Število zaposlenih je lahko od npr. 3 do 1.000 in več, promet je lahko od 2 milijona do nekaj 100 milijonov evrov, trgovska podjetja se lahko ukvarjajo s trgovino na veliko, na drobno ali z zunanjo trgovino, nekatera tudi z vsemi vrstami.

1.1.1 Populacija in enota

Populacija je skupnost enot, ki jih opredelimo zato, da jih preučimo.

Enota je vsak posamezen element populacije, *npr. podjetje, gospodinjstvo, delavec, študent, rojstvo, prometna nesreča, gradnja cest, proizvodnja žita.*

Naštete enote razvrstimo glede na časovno opredelitev v tri skupine:

- **realne enote, ki dejansko obstajajo in jih opazujemo v danem trenutku**; *to so podjetje, gospodinjstvo, študent, dijak;*
- **dogodki, ki se v danem trenutku zgodijo in jih opazujemo v časovnem razmiku**, *kot npr. rojstvo, prometna nesreča;*
- **dogajanja, ki v času nastajajo in jih opazujemo v danem trenutku ali časovnem razmiku**, *kot npr. gradnja stanovanj, proizvodnja obutve, pridobivanje električne energije.*

1.1.2 Populacije glede na časovno opredelitev

Glede na navedeno časovno opredelitev enot delimo populacije na:

- **trenutne** ali momentne, *npr. trgovska podjetja v Sloveniji 30. 6. 2008*;
- **razmične** ali intervalne, *npr. nesreče pri delu v Sloveniji v letu 2007;*
- **korespondirajoče** populacije, *ki predstavljajo povezavo trenutnih in razmičnih populacij, npr. zaloga konec meseca = zaloga na začetku meseca + nabava – prodaja.*

1.1.3 Spremenljivka in vrste spremenljivk

Spremenljivke so lastnosti enot, ki jih opazujemo s statističnimi opazovanji. Spremenljivka ima različne vrednosti.

Vrednost spremenljivke je lahko izražena:
- s številom, npr. število zaposlenih, polletni promet;
- z opisom, npr. vrsta lastnine, vrsta trgovinske dejavnosti.

Glede na način izražanja vrednosti so spremenljivke:
- **številske**, ki so lahko:
  - **zvezne**, *kot npr. polletni promet, ki ima poljubno vrednost na razmiku od 3 milijone do 100 milijonov evrov;*
  - **diskretne**, *kot npr. število zaposlenih, ki lahko ima le določeno vrednost, ki je celo število (npr. 345, 122, 65 zaposlenih);*
- **opisne**, *kot je vrsta podjetij, vrsta trgovinske dejavnosti.*

Izjeme so nekatere spremenljivke, ki jih sicer izražamo s številkami, vendar jih štejemo kot opisne, npr. *matična številka občana, kakovost blaga* (označimo z vrednostmi: I, II, III).

Številske spremenljivke so torej tiste, iz katerih lahko izračunamo parametre, npr. vsoto in povprečno vrednost.

1.2 STATISTIČNI PARAMETRI

Posamezna trgovska podjetja opazujemo po določenih lastnostih, in sicer zato, da bi dobili podatke, na osnovi katerih lahko ugotovimo lastnosti vseh opazovanih trgovskih podjetij, torej celotne populacije. **Mero, s katero izražamo lastnost populacije, imenujemo parameter.**

Parametre dobimo:
- **s preštevanjem enot,** *npr. število trgovskih podjetij v Sloveniji;*
- **s seštevanjem vrednosti spremenljivk**, *npr. promet trgovskih podjetij v Sloveniji v 1. polletju leta 2008;*
- **z razvrščanjem enot v skupine glede na vrednost spremenljivke,** *npr. število trgovskih podjetij po obliki lastnine;*
- **z izračunavanjem ustreznih kazalcev,** *npr. če želimo ugotoviti povprečen promet na trgovsko podjetje v prvem polletju, seštejemo promet vseh trgovskih podjetij v tem obdobju in dobljeno vrednost delimo s številom podjetij.*

S preštevanjem enot, s seštevanjem vrednosti spremenljivk in razvrščanjem enot v skupine glede na vrednost dobimo **enostavne parametre**, z izračunavanjem ustreznih kazalcev, npr. relativnih števil, srednjih vrednosti, mer variabilnosti, pa **izvedene parametre**.

V tabeli 1.1 so razvidne posamezne spremenljivke in nekaj njihovih smiselnih vrednosti za trgovska podjetja.

Tabela 1.1: **Spremenljivke, njihove vrednosti in vrsta spremenljivk**

Spremenljivka	**Vrednosti spremenljivke**	**Vrsta spremenljivke**
število zaposlenih	3; 15; 87; 455; 1.000;	številska – diskretna
promet v 1. polletju 2008	515.316,85; 2.885.678,80; 24.789.543,75 EUR	številska – zvezna
vrsta podjetja	samostojni podjetnik, delniška družba, d.o.o.	opisna
vrsta trgovinske dejavnosti	trgovina na drobno, trgovina na debelo, zunanja trgovina	opisna

Spremenljivke so glede na vrednosti bolj ali manj variabilne, torej se njihove vrednosti od enote do enote bolj ali manj razlikujejo. Spremenljivka *vrsta trgovinske dejavnosti* ima samo tri vrednosti, *število zaposlenih* pa zelo veliko.

1.3 PRIMER ZA UTRJEVANJE

Primer 1.2:
Preučiti želimo študente prvih letnikov višjih strokovnih šol po programih študija, načinu študija, kraju študija, spolu, starosti ter uspehu pri poklicni oziroma splošni maturi na začetku študijskega leta 2008/09 v Sloveniji.

Statistično **enoto predstavlja vsak študent višje strokovne šole, vpisan**
- v študijskem letu 2008/09, kar predstavlja časovno opredelitev;
- v Sloveniji, kar predstavlja krajevno opredelitev;

- v katerikoli program višje strokovne šole, kar predstavlja vsebinsko opredelitev.

Populacijo predstavljajo vsi študentje prvih letnikov višjih strokovnih šol v Sloveniji na začetku študijskega leta 2008/09.

Spremenljivke, po katerih opazujemo enote in nekaj smiselnih vrednosti, so razvidne iz tabele 1.2.

Tabela 1.2: **Spremenljivke, njihove vrednosti in vrsta spremenljivke**

Spremenljivka	**Vrednosti spremenljivke**	**Vrsta spremenljivke**
program študija	Ekonomist, Poslovni sekretar, Logistično inženirstvo, Lesarstvo, Organizator socialne mreže	opisna
način študija	redni, izredni	opisna
kraj študija	Murska Sobota, Celje, Brežice, Slovenj Gradec, Maribor, Šentjur, Novo mesto	opisna
spol	moški, ženske	opisna
uspeh pri poklicni maturi	odličen, prav dober, dober, zadosten	opisna
uspeh pri splošni maturi (število točk)	od 10 do 34 točk	številska – diskretna
starost rednih študentov	od 18 do 25 let	številska – zvezna
starost izrednih študentov	od 18 do 45 let (predvidoma, saj starost ni omejena navzgor)	številska – zvezna

Ugotavljamo, da ima spremenljivka *starost izrednih študentov* širši razmik vrednosti kot spremenljivka *starost rednih študentov*. Starost je praviloma številska – zvezna spremenljivka, vendar jo lahko opredelimo tudi kot diskretno. V tem primeru bi jo opredelili kot starost v dopolnjenih letih.

Statistični parametri so lahko:
- število študentov prvega letnika vseh višjih strokovnih šol v Sloveniji na začetku študijskega leta 2008/09,
- število študentov po programih,
- število rednih in izrednih študentov,
- struktura študentov v % po načinu študija,
- struktura študentov v % po spolu,
- povprečna starost izrednih študentov.

Prvi trije parametri so enostavni, drugi trije pa izvedeni.

Pridobljeno znanje lahko utrdite in preverite z reševanjem nalog 1.1 in 1.2 v Zbirki vaj iz statistike.

2 ZBIRANJE, UREJANJE IN PRIKAZOVANJE PODATKOV

S pridobljenim znanjem boste lahko izvedli zbiranje podatkov o določenem pojavu, te podatke obdelali oziroma uredili ter prikazali v tabelah in grafikonih. Ugotovili boste, kako je na tej stopnji statističnega preučevanja potrebno obvladati osnovna računalniška orodja za tabelarno in grafično prikazovanje.

2.1 STATISTIČNO OPAZOVANJE

Za pojav, ki je predmet preučevanja, moramo najprej zbrati podatke. Zbiranje podatkov oziroma ***statistično opazovanje*** je eno od opravil statističnega raziskovanja. S tem lahko pričnemo, ko je:
- **znan namen preučevanja** in je
- **populacija točno in nedvoumno opredeljena** (krajevno, časovno in vsebinsko).

Zbrati moramo popolne in točne podatke, torej podatke za vse opazovane enote po predvidenih spremenljivkah, ki ustrezajo dejanskemu stanju. Zbrani podatki so osnova za statistično analizo, njihova kakovost je bistvena za kakovost končnega rezultata raziskovanja, s katerim želimo odkriti osnovne zakonitosti variiranja vrednosti spremenljivk opazovanih pojavov.

Statistična opazovanja so lahko:
- ***popolna***, z njimi zberemo podatke za vse enote preučevane populacije;
- ***delna***, z njimi zberemo podatke le za del enot.

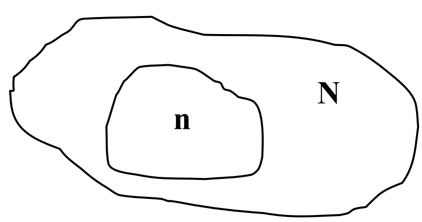
Popolno opazovanje sta ***popis realnih enot in sprotno spremljanje dogodkov***.

Najbolj znan je ***popis prebivalstva***, ki ga v Sloveniji izvajajo vsakih deset let. S tem dobijo podrobno in popolno sliko prebivalstva v določenem trenutku (podatki se nanašajo na 31. marec). Z zaporednimi popisi spremljajo gibanje prebivalstva.

Sprotno spremljanje dogodkov predstavlja evidenca rojstev, umrlih, prometnih nesreč in podobno.

Delno opazovanje imenujemo tudi vzorčenje. Iz populacije običajno slučajno izberemo poljubno število enot, to je ***slučajni vzorec***, za katere zberemo popolne in točne podatke. Parametre izračunamo za vzorec in jih posplošimo na celotno populacijo. Torej ne ugotavljamo dejanskih statističnih parametrov za celotno populacijo, ampak te le ocenimo.

N – število enot v populaciji
n – število enot v vzorcu



Delno opazovanje izvajamo, ko popolna slika pojava ni potrebna ali pa zaradi velikega števila enot ali drugih razlogov popolnega opazovanja ni mogoče izvesti. Npr. Statistični urad RS z opazovanjem določenega števila gospodinjstev ugotavlja gibanje življenjskih stroškov, agencije za raziskave javnega mnenja ocenjujejo rezultate izida volitev, gledanosti TV-programov, priljubljenost politikov, podjetja ocenjujejo svoje prodajne možnosti na trgu.

Pred izvajanjem statističnega opazovanja je smiselno predhodno preučiti statistične vire, ki se nanašajo na pojav, ki ga bomo preučevali. Morda so podatke o tem pojavu zbrali in obdelali že drugi in so že objavljeni ali urejeni v ustreznih evidencah. Tako nam bo potrebne podatke lahko posredoval Statistični urad Republike Slovenije ali katera druga inštitucija, ki je pooblaščena za statistična preučevanja s svojega delovnega področja. Te evidence imenujemo ***sekundarni vir statističnih podatkov.*** Če teh ni na razpolago, sami zberemo podatke; gre za ***primarni vir***:

- **s posrednim opazovanjem** po osebah, ki enoto poznajo in nam dajo zahtevane podatke; npr. podatke o številu zaposlenih in prometu v 1. polletju nam dajo trgovska podjetja;

- **z neposrednim opazovanjem**, ko sami ugotavljamo vrednost spremenljivk pri enotah; npr. učitelj športne vzgoje izmeri telesno višino in težo dijakov, meri čas teka na 400 in 800 metrov, dolžino skoka v daljino in podobno.

Oseba, ki izvaja opazovanje, zapisuje podatke oziroma odgovore v za to pripravljene ***vprašalnike***.

Na vprašalniku so v obliki vprašanj vse spremenljivke, po katerih bomo enote opazovali. Sestavljen mora biti tako, da čim lažje zberemo čim bolj kakovostne odgovore. Pri sestavljanju vprašanj moramo upoštevati sposobnosti in pripravljenost tistih, ki ga bodo izpolnjevali. Torej morajo biti vprašanja kratka in razumljiva najširšemu krogu ljudi. Če želimo dobiti pravilne odgovore, vprašanja ne smejo biti preveč osebna.

2.2 OBDELAVA OZIROMA UREJANJE PODATKOV

S statističnim opazovanjem zberemo več podatkov za vsako opazovano enoto populacije, tako je število podatkov za vse enote populacije pogosto zelo veliko. Tako zbrani podatki pa so nepregledni in iz njih ne moremo neposredno ugotoviti lastnosti preučevane populacije. Torej moramo podatke zapisati na način in v obliki, ki omogočata in olajšata prepoznavanje značilnosti populacije in izračunavanje parametrov z uporabo izbranih statističnih metod.

Podatke moramo urediti in prikazati v pregledni obliki. Pogosto zadošča že, da enote razvrstimo v skupine po vrednostih določene spremenljivke. V primeru, da ima spremenljivka le nekaj (manjše število) vrednosti, za vsako vrednost opredelimo skupino. Npr. študente razvrstimo po spolu v dve skupini (moški, ženske), prav tako po načinu študija (redni, izredni). Za spremenljivke, kot so *starost izrednih študentov*, *stalno prebivališče, mesečni prejemki*, ni mogoče opredeliti skupine za vsako vrednost, saj jih je preveč. Preveliko število skupin pomeni manjšo preglednost, pogosto pa tudi ne pridejo do izraza osnovne značilnosti preučevanega pojava. Zato v takih primerih v skupine združujemo sorodne vrednosti, npr. za spremenljivko starost 138 izrednih študentov v *Izobraževalnem središču Znanje* bi opredelili razrede, kot so razvidni iz tabele 2.1:

Tabela 2.1: **Skupine za spremenljivko starost izrednih študentov v *Izobraževalnem središču Znanje***

Starost izrednih študentov v dopolnjenih letih
do 20
od 21 do 25
od 26 do 30
od 31 do 35
od 36 do 40
od 41 do 45

Z razvrščanjem enot v skupine (razrede), ki smo jih opredelili za posamezne ali skupinske vrednosti določene spremenljivke, dobimo ***statistične vrste.***

2.2.1 Statistične vrste

Statistična vrsta nastaja v času raziskave in je običajno ***neurejena***. Če pa vrednosti uredimo po določenem kriteriju, je vrsta ***urejena.***

Neurejena statistična vrsta

Promet v prodajalnah trgovskega podjetja *Košarica* v 1. polletju 2008 v milijonih evrov:

18,5 23,4 12,6 22,7 31,2 15,6 21,7 26,9 14,2 11,1

Urejena statistična vrsta

Promet prodajaln uredimo po velikosti od najmanjše do največje vrednosti (lahko tudi obratno) in dobimo ***ranžirno vrsto***:

11,1 12,6 14,2 15,6 18,5 21,7 22,7 23,4 26,9 31,2

Z razvrščanjem enot po krajevni, časovni ali vsebinski (stvarni) opredelitvi opazovane populacije dobimo krajevne, časovne in stvarne vrste.

2.2.1.1 Časovne vrste

Časovne vrste so lahko:

- ***trenutne oziroma momentne***, če se podatki nanašajo na časovne trenutke;
- ***razmične oziroma intervalne***, če se podatki nanašajo na časovne razmike.

Trenutna časovna vrsta

Tabela 2.2: **Registrirane brezposelne osebe konec meseca v letu 2007 v Sloveniji**

Mesec	Število brezp. konec meseca
Januar	79.969
Februar	77.669
Marec	74.216
April	72.573
Maj	70.730
Junij	69.272
Julij	70.134
Avgust	68.539
September	66.658
Oktober	69.500
November	68.355
December	68.411

Vir: Pomembnejši statistični podatki o Sloveniji, letnik III, št. 3/2008

Razmična časovna vrsta

Tabela 2.3: **Prenočitve turistov v Sloveniji v letu 2007 po mesecih**

Mesec	Prenočitve turistov v 1000
Januar	472,3
Februar	497,2
Marec	502,1
April	567,6
Maj	618,6
Junij	811,1
Julij	1.226,2
Avgust	1.325,0
September	776,1
Oktober	572,5
November	440,7
December	445,2

Vir: Pomembnejši statistični podatki o Sloveniji, letnik III, št. 3/2008

2.2.1.2 Krajevna vrsta

Tabela 2.4***:*** **Bruto družbeni produkt v tekočih cenah v standardih kupne moči (PPS[1]) v Sloveniji in sosednjih državah v letu 2007**

Država	BDP v tekočih cenah
Slovenija	22.000
Avstrija	31.600
Hrvaška	13.900
Italija	25.200
Madžarska	15.700

Vir: www. http://epp.eurostat.ec.europa.eu (15. 8. 2008)

[1] *PPS – Purchasing Power Standard*

2.2.1.3 Stvarna vrsta

Tabela 2.5: **Zaposlene osebe pri pravnih osebah po stopnji izobrazbe 31. 12. 2006 v Sloveniji**

Stopnja izobrazbe	**Število oseb**
Nižja izobrazba	2.690
Visokokvalificirani	5.038
Kvalificirani	182.558
Polkvalificirani	15.284
Nekvalificirani	108.266
Neznano	1.977
Skupaj	**315.813**

Vir: http://www.stat.si/letopis/2007/ (18. 8. 2008)

2.3 PRIKAZOVANJE PODATKOV

Statistične vrste prikazujemo v ***tabelah*** in z ***grafikoni***. Tabele imajo to prednost, da lahko v njih prikažemo natančne podatke, razvrščene tudi po več spremenljivkah hkrati in so poljubno velike. Grafikoni pa so zelo pregledni in nazorni. Običajno pa se grafikon in tabela dopolnjujeta.

2.3.1 Prikazovanje podatkov v tabelah

Tabela je sestavljena iz vrstic in stolpcev, ki tvorijo polja, v katera vnašamo podatke. V glavi in čelu tabele so zapisane spremenljivke in njihove vrednosti. Prvi ali zadnji stolpec je zbirni stolpec, v katerem je seštevek podatkov vrstic. Prva ali zadnja vrstica je zbirna vrstica, v njej je seštevek podatkov v stolpcih.

Nad tabelo napišemo naslov, iz katerega je razvidno, katere podatke prikazuje, in mersko enoto, v kateri so podatki izraženi. Pod tabelo zapišemo vir, iz katerega smo črpali podatke.

Naslov tabele

	GLAVA TABELE		
Č			**s**
E			**t**
L		**v r s t i c a**	**o**
O		***polje***	**l**
			p
			e
	zbirna vrstica		**c**

Vir:

Za pojasnila k posameznim podatkom v tabeli uporabljamo naslednje oznake:

– (črtica v polju tabele) pojava ni
… (tri pike v polju tabele) pojav obstoja, vendar ni podatka zanj
0 (nična v polju tabele) podatek je manjši od 0,5 dane merske enote
0,0 (dve nični v polju tabele) podatek je manjši od 0,05 dane merske enote
∅ (prečrtan krog pred številom) podatek je izraz povprečja
( ) (število v oklepaju) nepopoln oziroma ne dovolj preverjen podatek

* (zvezdica nad številom) popravljen podatek
1) (številka desno nad podatkom) oznaka za opombo pod tabelo

Glede na to, po koliko spremenljivkah so razvrščeni podatki in koliko statističnih vrst tabele prikazujejo, jih delimo na:

- **enostavne ali enorazsežne**, ki prikazujejo eno samo statistično vrsto in so podatki razvrščeni samo po eni spremenljivki (npr. tabele 2.2, 2.3, 2.4 in 2.5);
- **sestavljene**, ki prikazujejo več istovrstnih statističnih vrst za več časovnih obdobij (npr. tabela 2.6);
- **kombinirane ali dvorazsežne**, ki prikazujejo podatke, razvrščene po dveh spremenljivkah.

2.3.2 Prikazovanje podatkov z grafikoni

Statistične vrste pogosto prikazujemo z grafikoni, saj na enostaven in pregleden način posredujejo osnovne informacije o preučevanem pojavu. Da bi te cilje dosegli, moramo grafikon ustrezno opremiti in narisati.

Za grafično prikazovanje uporabimo pozitivni del pravokotnega koordinatnega sistema, v katerem je abscisna os skala za vrednosti spremenljivke (leta, meseci), na ordinatni osi pa je merilo za podatke.

Časovne vrste lahko prikažemo s stolpci ali z linijskim grafikonom.

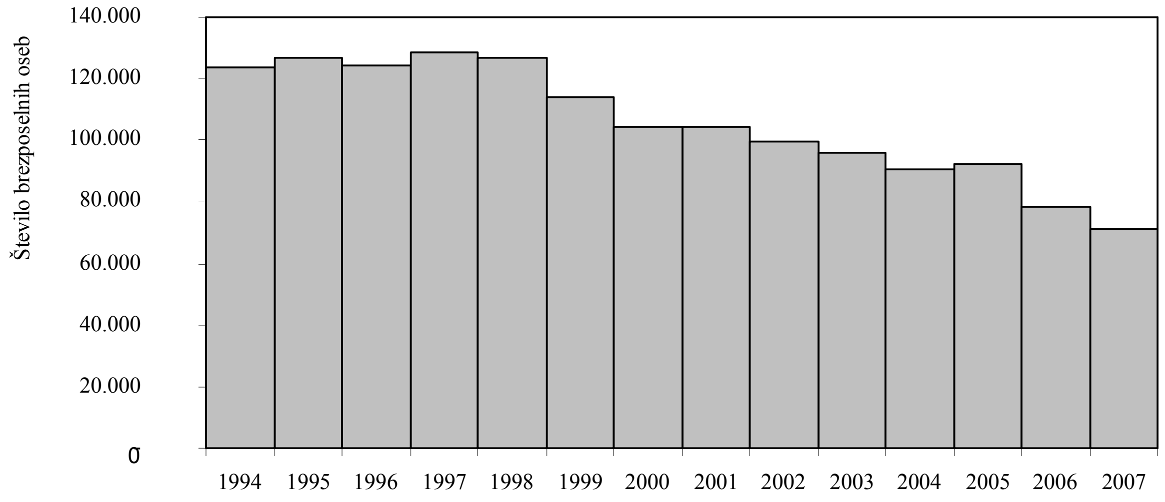
Prikažimo s stolpci registrirane brezposelne osebe v Sloveniji v letih 1994 do 2007, in sicer s preprostimi stolpci skupno število le-teh (slika 2.1), nato pa brezposelne osebe po spolu z razdeljenimi in še s sestavljenimi stolpci.

Tabela 2.6: **Registrirane brezposelne osebe po spolu v Sloveniji v letih od 1994 do 2007 (stanje 31.12)**

Leto	Moški	Ženske	Skupaj
1994	67.120	56.397	123.517
1995	66.875	59.884	126.759
1996	64.521	59.949	124.470
1997	65.660	62.912	128.572
1998	63.361	63.264	126.625
1999	56.445	57.903	114.348
2000	52.003	52.580	104.583
2001	51.353	52.963	104.316
2002	51.378	48.229	99.607
2003	45.669	50.324	95.993
2004	42.911	47.817	90.728
2005	42.877	49.698	92.575
2006	35.711	42.592	78.303
2007	32.192	39.144	71.336

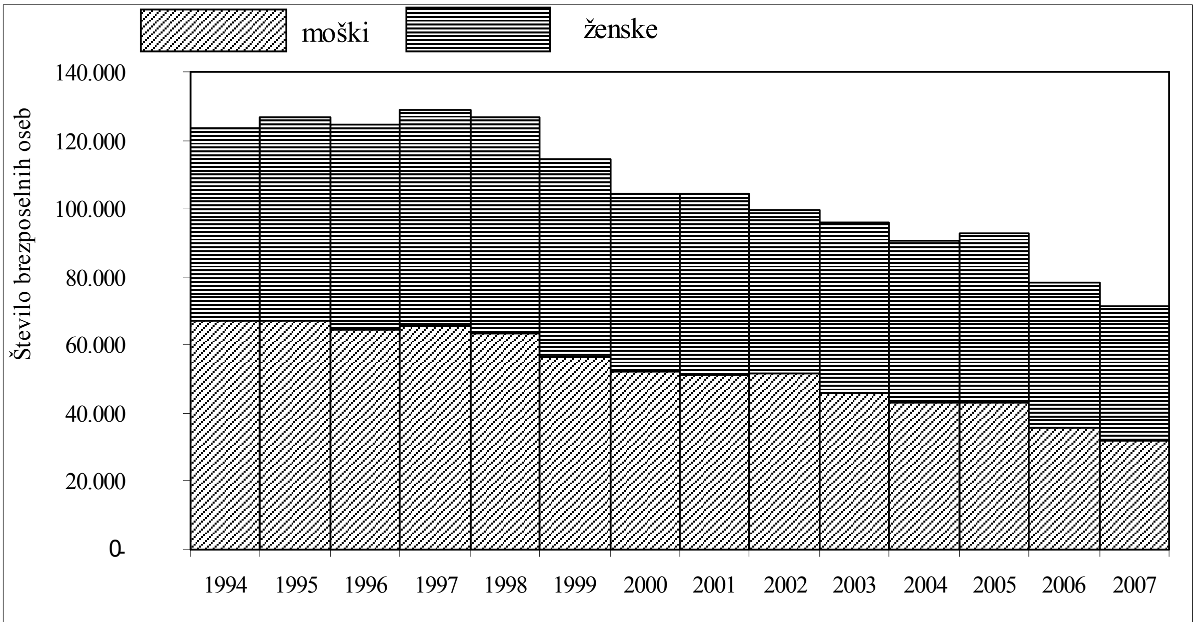
Vir: www.ess.gov.si (17. 8. 2008)

Na sliki 2.2 je grafični prikaz z razdeljenimi, na sliki 2.3 pa s sestavljenimi stolpci. S primerjavo grafikonov ugotovimo določene prednosti enega in drugega načina. Tako v prikazu z razdeljenimi stolpci ne moremo neposredno odčitati števila brezposelnih žensk, lahko pa odčitamo število vseh brezposelnih oseb po letih. V prikazu s sestavljenimi stolpci pa je lepo razvidno število brezposelnih moških in žensk, vendar ni razvidno število vseh brezposelnih oseb po letih.



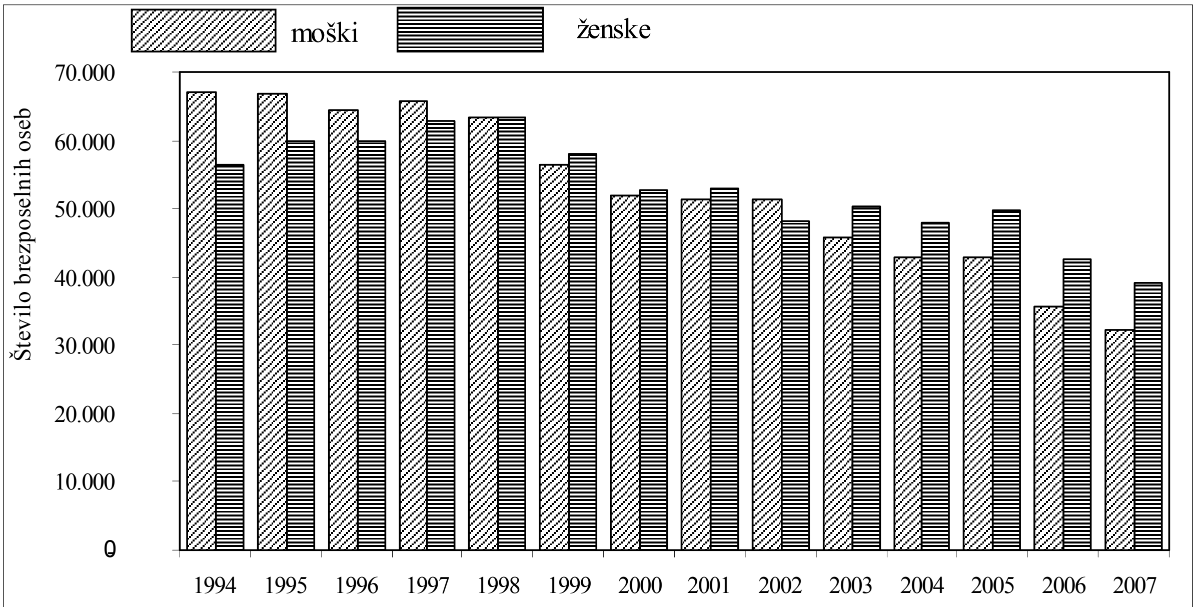
Slika 2.1: **Registrirane brezposelne osebe v letih od 1994 do 2007 v Sloveniji**

Vir: Tabela 2.6



Slika 2.2: **Registrirane brezposelne osebe po spolu v letih od 1994 do 2007 v Sloveniji**

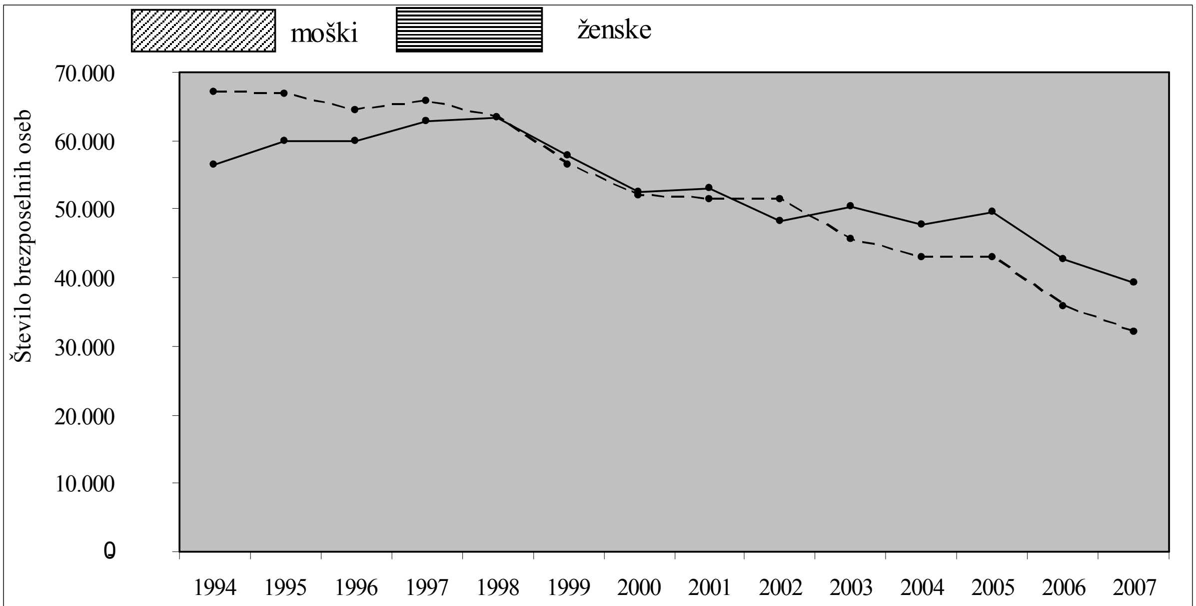
Vir: Tabela 2.6



Slika 2.3: **Registrirane brezposelne osebe po spolu v letih od 1994 do 2007 v Sloveniji**

Vir: Tabela 2.6

Prikažimo brezposelne osebe po spolu še z linijskim grafikonom (slika 2.4).



Slika 2.4: **Registrirane brezposelne osebe po spolu v letih od 1994 do 2007 v Sloveniji**

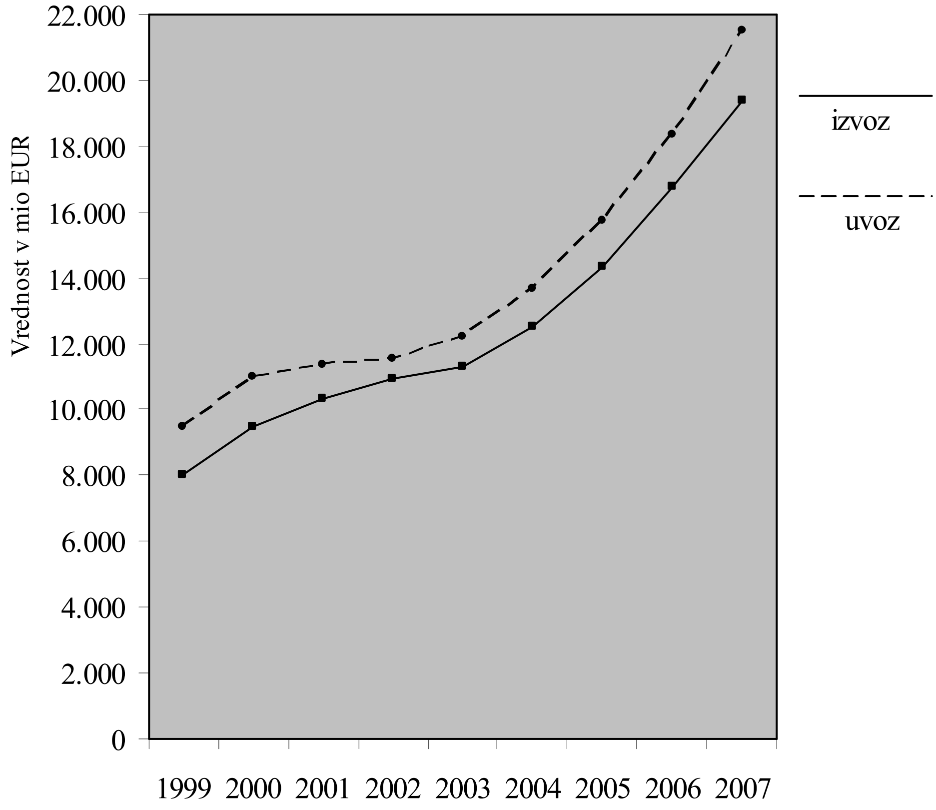
Vir: Tabela 2.6

Prikažimo z linijskim grafikonom še vrednost izvoza in uvoza Slovenije (slika 2.5).

Tabela 2.7: **Izvoz in uvoz Slovenije v letih od 1999 do 2007**

Leto	Vrednost izvoza v mio EUR	Vrednost uvoza v mio EUR
1999	8.030,8	9.477,6
2000	9.491,6	10.984,2
2001	10.346,8	11.344,5
2002	10.962,0	11.574,1
2003	11.285,0	12.238.9
2004	12.536,4	13.698,7
2005	14.314,5	15.728,2
2006	16.757,2	18.340,8
2007	19.385,2	21.486,7

Vir: Pomembnejši statistični podatki o Sloveniji, letnik III, št. 3/2008



Slika 2.5**: Izvoz in uvoz Slovenije v letih od 1999 do 2007**

Vir: Tabela 2.7

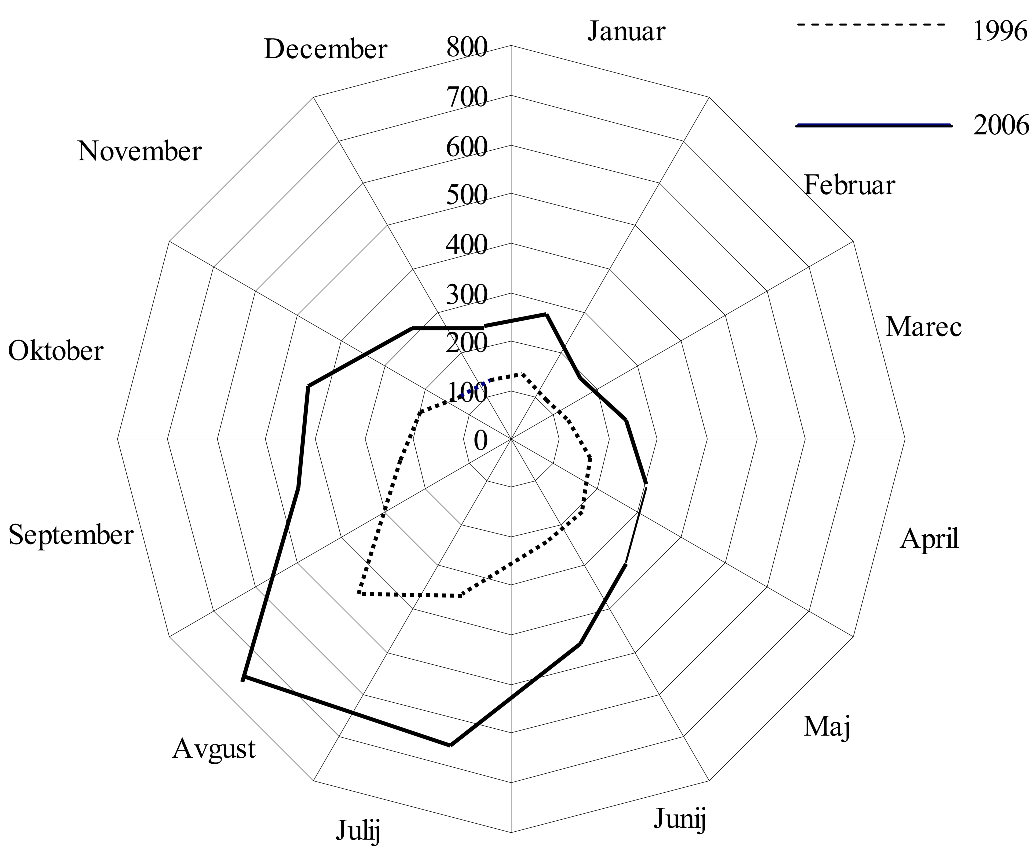
V linijskem grafikonu o številu brezposelnih oseb po spolu so točke, ki so povezane z daljicami, na koncu časovnih razmikov, v tem primeru let, saj so podatki dani za konec leta. Gre namreč za trenutno (momentno) časovno vrsto. V grafikonu o vrednosti izvoza in uvoza pa so točke na sredini časovnih razmikov. Gre za razmično (intervalno) časovno vrsto, pri teh pa točke vedno nanašamo nad sredine časovnih razmikov.

Časovne vrste z izrazitejšimi sezonskimi nihanji, ki so značilna predvsem za pojave v turizmu, prikazujemo s polarnim grafikonom (slika 2.6).

Tabela 2.8: **Prenočitve tujih turistov v Sloveniji v letih od 1996 in 2006 po mesecih**

Mesec	Število prenočitev v 1.000	
	1996	2006
Januar	140,1	271,5
Februar	119,4	195,8
Marec	141,5	240,3
April	180,6	305,6
Maj	214,8	355,1
Junij	240,2	462,3
Julij	356,3	658,5
Avgust	448,3	752,5
September	260,7	478,1
Oktober	191,5	328,6
November	126,1	209,5
December	131,1	227,7

Vir: Mesečni statistični pregled, marec 1997 in Pomembnejši statistični podatki o Sloveniji, letnik I, št. 1/2007



Slika 2.6: **Prenočitve turistov iz tujine v Sloveniji po mesecih v letih 1996 in 2006**

Vir: Tabela 2.8

Preučevanje vsakega pojava se prične s statističnim opazovanjem oziroma z zbiranjem podatkov o enotah populacije po določenih spremenljivkah. Od kakovosti zbranih podatkov je odvisna kakovost končnega rezultata preučevanja.

Zbrane podatke uredimo in prikažemo v tabelah ali grafikonih. Največkrat pa na oba načina, saj se lepo dopolnjujeta. Tako prikazani podatki so osnova za statistično analizo, s katero izračunamo ustrezne statistične parametre, ki jih bomo spoznali v naslednjih poglavjih. Za urejanje in prikazovanje podatkov v tem gradivu sem uporabila računalniško orodje Excel. Svetujem tudi vam, da tako preizkusite svoje znanje.

3 RELATIVNA ŠTEVILA

V tem poglavju se boste naučili izračunati in pojasniti parametre, kot so: strukture, koeficienti in indeksi. Spoznali boste pomen izračunanih relativnih števil v statistični analizi določenega pojava oziroma pri primerjavi več pojavov. Ugotovili boste, katero relativno število je smiselno uporabiti glede na namen preučevanja,.

3.1 POMEN IN VRSTE RELATIVNIH ŠTEVIL

Podatki nekega pojava, ki jih zberemo, uredimo in prikažemo v tabeli, so osnova za statistično analizo. Namen te je ugotoviti lastnosti preučevanega pojava. Zato zbrane podatke med seboj primerjamo. Primerjavo zelo poglobimo in razširimo ***z relativnimi števili***.

Veliko boljšo predstavo o spolu registriranih brezposelnih oseb v Sloveniji dobimo, če rečemo, da je bilo ob koncu leta 2007 med brezposelnimi 45,1 % moških in 54,9 % žensk, kakor pa če navedemo, da je bilo brezposelnih 32.192 moških in 39.144 žensk. Prav tako je bolje, če rečemo, da je bilo število brezposelnih oseb ob koncu leta 2007 za 8,9 % manjše v primerjavi s koncem leta 2006, kot pa, da jih je bilo 6.967 manj.

Relativno število iz dveh primerjanih podatkov dobimo tako, da ju delimo. Računamo ga za podatke, katerih primerjava je smiselna. Glede na podatke, iz katerih relativna števila računamo, ločimo:

- **strukture**, kadar primerjamo podatek, ki predstavlja del pojava, s podatkom, ki predstavlja celoten pojav; npr. od skupnega števila brezposelnih oseb v Sloveniji je bilo ob koncu leta 2007 54,9 % žensk;

- **statistične koeficiente**, kadar primerjamo raznovrstne podatke, ki so v vsebinski zvezi, npr. družbeni bruto produkt na prebivalca v Sloveniji je bil leta 2007 16.600 evrov;

- **indekse**, kadar primerjamo istovrstne podatke; npr. indeks spremembe števila registriranih brezposelnih oseb za konec leta 2007 je 91,1, če upoštevamo, da je število brezposelnih konec leta 2006 enako 100.

3.2 STRUKTURE

Uspeh študentov pri opravljanju izpitov prikažemo z odstotkom študentov, ki so izpit opravili, ne pa s številom študentov, saj v tem primeru ne bi mogli tega uspeha primerjati po predmetih, ker je lahko število študentov, ki so pristopili k izpitu določenega predmeta, različno. Prav tako bi bilo zaradi različnega števila prebivalstva po državah nesmiselno primerjati število aktivnega prebivalstva, ampak bomo primerjali odstotek aktivnega prebivalstva. Tako ravnamo vselej, kadar želimo nazorno in primerljivo izraziti del pojava v razmerju do celotnega pojava ali pa sestavo pojava.

Strukture izražamo v obliki ***strukturnih deležev, odstotkov ali odtisočkov***. Strukturni odstotek izračunamo tako, da del pojava primerjamo s celotnim, izračunano razmerje pa pomnožimo s 100. Če računamo strukturni odtisoček, izračunano razmerje pomnožimo s 1000.

- **strukturni delež:** $P_j = \frac{Y_j}{Y}$ *(3.1)*
- **strukturni odstotek:** $P_j\% = \frac{Y_j}{Y} \times 100$ *(3.2)*
- **strukturni odtisoček:** $P_j‰ = \frac{Y_j}{Y} \times 1.000$ *(3.3)*

3.2.1 Enorazsežna struktura

Enorazsežna struktura je izračunana za opazovani pojav, razčlenjen le po vrednostih ene spremenljivke.

Primer 3.1: Konec leta 2006 je bilo v Sloveniji 1 milijon 723 tisoč prebivalcev, starih 15 let in več, njihova šolska izobrazba pa je razvidna iz tabele 3.1.

Tabela 3.1: **Prebivalstvo, staro 15 let in več, po šolski izobrazbi v Sloveniji 31.12. 2006**

Šolska izobrazba	**Število prebivalcev v tisoč**	**Struktura v %**
Brez izobrazbe ali nepopolna osnovna izobrazba	80	4,6
Osnovna izobrazba	397	23,0
Nižja ali srednja poklicna izobrazba	430	25,0
Srednja strokovna ali splošna izobrazba	530	30,8
Višja strokovna, višješolska ali visoka strokovna izobrazba	143	8,3
Univerzitetna ali podiplomska (spec., magisterij, doktorat) izobrazba	143	8,3
Skupaj	1.723	100,0

Vir: Statistični letopis 2007

$$P\% = \frac{80}{1.723} \times 100 = 4,6\ \%,$$

kar pomeni, da je bilo v Sloveniji konec leta 2006 le še 4,6 % prebivalstva brez izobrazbe ali z nepopolno osnovno izobrazbo. Z izračunom ostalih odstotkov ugotovimo, da je bilo 23 % prebivalcev, starih 15 let in več, z osnovno izobrazbo, 25 % z nižjo ali srednjo poklicno izobrazbo, 30,8 % s srednjo strokovno ali splošno izobrazbo, 8,3 % z višjo strokovno, višješolsko ali visoko strokovno izobrazbo in 8,3 % z univerzitetno ali podiplomsko izobrazbo.

3.2.2 Dvorazsežna struktura

Dvorazsežna struktura je izračunana za pojav, razčlenjen po vrednostih dveh spremenljivk.

Primer 3.2: V Sloveniji je bilo konec leta 2006 pri pravnih osebah v gostinstvu zaposlenih 6.282 oseb. V tabeli 3.2 so razvrščene po spolu in stopnji strokovne izobrazbe.

Tabela 3.2: **Zaposlene osebe po stopnji izobrazbe in spolu pri pravnih osebah v gostinstvu v Sloveniji 31.12. 2006**

Spol	**Stopnja izobrazbe**			**Skupaj**
	Visoka in podiplomska	Višja	Srednja	
Moški	294	252	2.003	**2.549**
Ženske	593	424	2.716	**3.733**
Skupaj	**887**	**676**	**4.719**	**6.282**

Vir: Statistični letopis 2007

Iz podatkov v tabeli 3.2 lahko izračunamo tri strukture:

a) strukturo zaposlenih po spolu,
b) strukturo zaposlenih po stopnji izobrazbe ter
c) strukturo zaposlenih po spolu in stopnji izobrazbe hkrati.

a) Struktura zaposlenih v gostinstvu po spolu

Tabela 3.3: **Struktura zaposlenih oseb po spolu pri pravnih osebah v gostinstvu v Sloveniji 31.12. 2006**

Spol	**Stopnja izobrazbe**			**Skupaj**
	Visoka in podiplomska	Višja	Srednja	
Moški	33,1	37,3	42,4	40,6
Ženske	66,9	62,7	57,6	59,4
Skupaj	**100,0**	**100,0**	**100,0**	**100,0**

Vir: Tabela 3.2

V gostinstvu je bilo zaposlenih 40,6 % moških in 59,4 % žensk. Od zaposlenih z visoko izobrazbo je bilo 33,1 % moških in 66,9 % žensk, z višjo izobrazbo 37,3 % moških in 62,7 % žensk ter s srednjo izobrazbo 42,4 % moških in 57,6 % žensk.

Navedena sta dva primera, kako izračunamo posamezne odstotke:

Odstotek zaposlenih moških z visoko izobrazbo od vseh zaposlenih z visoko izobrazbo:

$$P_1\,\% = \frac{294}{887} \times 100 = 33{,}1\ \%$$

Odstotek zaposlenih žensk s srednjo izobrazbo od vseh zaposlenih s srednjo izobrazbo:

$$P_6\,\% = \frac{2.716}{4.719} \times 100 = 57{,}6\ \%$$

b) Struktura zaposlenih v gostinstvu po stopnji strokovne izobrazbe

Tabela 3.4: **Struktura zaposlenih po stopnji strokovne izobrazbe pri pravnih osebah v gostinstvu v Sloveniji 31.12. 2006**

Spol	Stopnja izobrazbe			Skupaj
	Visoka in podiplomska	Višja	Srednja	
Moški	11,5	9,9	78,6	**100,0**
Ženske	15,9	11,4	72,7	**100,0**
Skupaj	14,1	10,8	75,1	**100,0**

Vir: Tabela 3.2

Iz tabele 3.4 je razvidno, da je bilo zaposlenih 14,1 % oseb z visoko, 10,8 % oseb z višjo in 75,1 % oseb s srednjo izobrazbo. Pri ženskah so bili ti odstotki nekoliko ugodnejši kot pri moških, saj je bilo 15,9 % z visoko, 11,4 % z višjo in 72,7 % s srednjo izobrazbo. Zaposlenih moških pa je bilo z visoko izobrazbo 11,5 %, z višjo 9,9 % in s srednjo 78,6 %.

Navedena sta dva primera, kako izračunamo posamezne odstotke.

Odstotek zaposlenih moških z visoko izobrazbo od vseh zaposlenih moških:

$$P_1\% = \frac{294}{2.549} \times 100 = 11{,}5\ \%$$

Odstotek zaposlenih žensk z višjo izobrazbo od vseh zaposlenih žensk:

$$P_5\,\% = \frac{424}{3.733} \times 100 = 11{,}4\ \%$$

c) Struktura zaposlenih v gostinstvu po spolu in stopnji izobrazbe

Tabela 3.5: **Struktura zaposlenih po spolu in stopnji izobrazbe pri pravnih osebah v gostinstvu v Sloveniji 31.12. 2006**

Spol	Stopnja izobrazbe			Skupaj
	Visoka in podiplomska	Višja	Srednja	
Moški	4,7	4,0	31,9	40,6
Ženske	9,4	6,8	43,2	59,4
Skupaj	14,1	10,8	75,1	**100,0**

Vir: Tabela 3.2

Iz tabele 3.5 so razvidni odstotki zaposlenih po vrednostih obeh spremenljivk, spolu in stopnji izobrazbe. Tako je bilo od skupnega števila zaposlenih 4,7 % moških z visoko izobrazbo, 4,0 % z višjo in 31,9 % s srednjo, ter žensk z visoko izobrazbo 9,4 %, z višjo 6,8 % in s srednjo 43,2 %.

Navedena sta dva primera, kako izračunamo posamezne odstotke.

Odstotek zaposlenih moških z visoko izobrazbo od vseh zaposlenih:

$$P_1\,\% = \frac{294}{6.282} \times 100 = 4,7\ \%$$

Odstotek zaposlenih žensk s srednjo izobrazbo od vseh zaposlenih:

$$P_6\,\% = \frac{2.716}{6.282} \times 100 = 43,2\ \%$$

Katere od navedenih struktur bomo izračunali pri analizi določenega pojava, je odvisno od ciljev analize. Nesmiselno je izračunati vse strukture, saj se na nek način dopolnjujejo. Izberemo tisto, ki z vsebinskega vidika najbolje osvetljuje preučevani pojav.

3.2.3 Grafično prikazovanje struktur

3.2.3.1 Prikaz s strukturnim stolpcem

Enorazsežno strukturo prikažemo s strukturnim stolpcem tako, da se višina stolpca nanaša na celoto, torej 100 %, posamezni deli stolpca pa prikazujejo ustrezne deleže celote. Da bo stolpec ustrezno razdeljen na posamezne dele, postopno seštevamo strukturne odstotke, da dobimo delne vsote.

S strukturnimi stolpci prikazujemo strukture določenega pojava po letih ali kaki drugi časovni spremenljivki, npr. strukturo prebivalstva, starega 15 let in več, po izobrazbi prikažemo po letih popisa (slika 3.1).

Tabela 3.6: **Prebivalstvo, staro 15 let in več, po šolski izobrazbi, popisi 1971, 1981, 1991 in 2002**

Šolska izobrazba	**Leto popisa**			
	1971	1981	1991	2002
Brez izobrazbe	375.260	370.169	263.488	115.556
Osnovna šola	522.217	458.626	451.222	433.910
Srednja šola	317.261	487.937	652.292	899.341
Višja šola	16.668	40.136	69.509	84.044
Visoka in podiplomska izobrazba	25.810	44.469	65.240	131.018
Skupaj	1.257.216	1.401.337	1.501.751	1.663.869

Vir: Statistični letopis 2004, 112

Izračunamo strukturo prebivalstva po šolski izobrazbi za leta popisa in jih prikažemo grafično s strukturnimi stolpci.

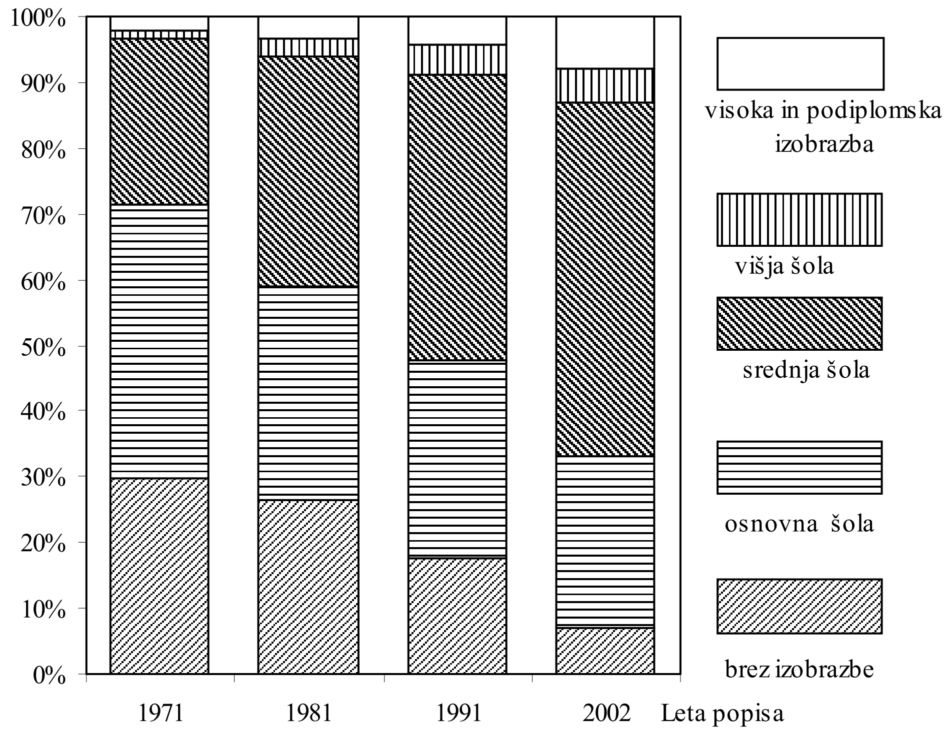
V tabeli 3.7 lahko vidimo bistveno spremembo v strukturi prebivalstva po šolski izobrazbi od popisa leta 1971 do popisa 2002. Tako je bilo ob popisu leta 1971 kar 29,9 % prebivalstva, starega
15 let in več, brez osnovne izobrazbe, leta 2002 pa le še 6,9 %. S srednješolsko

izobrazbo je bilo leta 1971 25,2 %, leta 2002 pa že 54,1 % prebivalstva. Te razlike so še bolj opazne na slikah 3.1, 3.5 in 3.6.

Tabela 3.7: **Struktura prebivalstva, starega 15 let in več, po izobrazbi v Sloveniji, popisi 1971, 1981, 1991 in 2002**

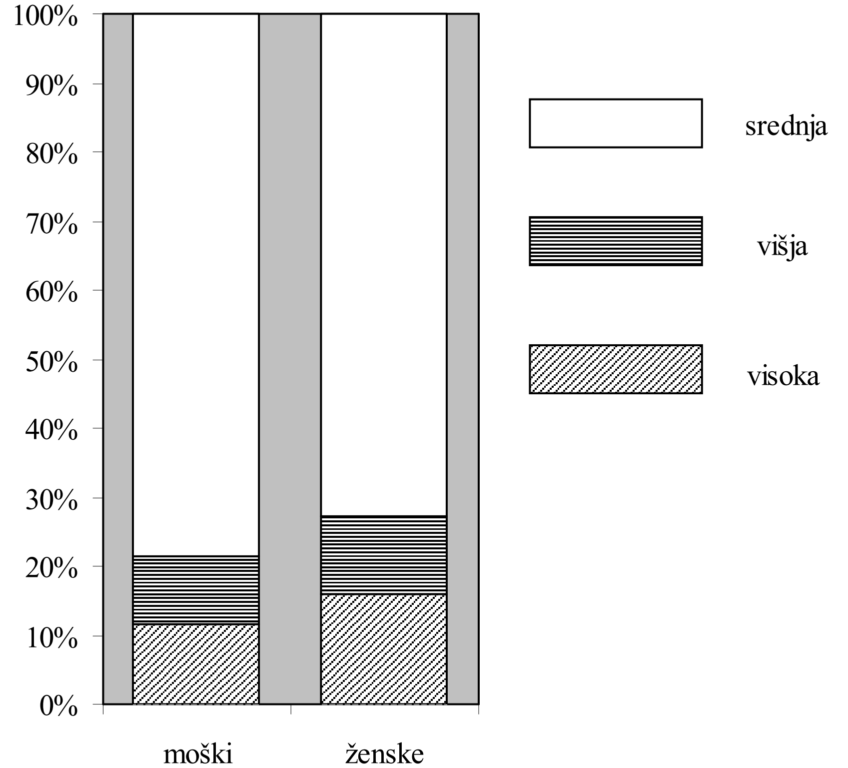
Šolska izobrazba	**Leto popisa**			
	1971	1981	1991	2002
Brez izobrazbe	29,9	26,4	17,6	6,9
Osnovna šola	41,5	32,7	30,0	26,1
Srednja šola	25,2	34,8	43,4	54,0
Višja šola	1,3	2,9	4,6	5,1
Visoka in podiplomska izobrazba	2,1	3,2	4,3	7,9
Skupaj	**100,0**	**100,0**	**100,0**	**100,0**

Vir: Tabela 3.6



Slika 3.1: **Struktura prebivalstva, starega 15 let in več, po izobrazbi, popisi 1971, 1981, 1991, 2002**

Vir: Tabela 3.6



Slika 3.2: **Struktura zaposlenih moških in žensk po stopnji izobrazbe pri pravnih osebah v gostinstvu 31. 12. 2006**

Vir: Tabela 3.2

S strukturnimi stolpci pa lahko primerjamo tudi strukturo dveh ali več pojavov po isti spremenljivki, npr. strukturo zaposlenih moških in žensk v gostinstvu po stopnji izobrazbe (slika 3.2).

3.2.3.2 Prikaz s strukturnim krogom

Za prikaz struktur s strukturnim krogom moramo strukturne odstotke preračunati v ločne stopinje, kar naredimo tako, da jih pomnožimo s 3,6. 360 stopinj namreč predstavlja celoto, to je 100 odstotkov.

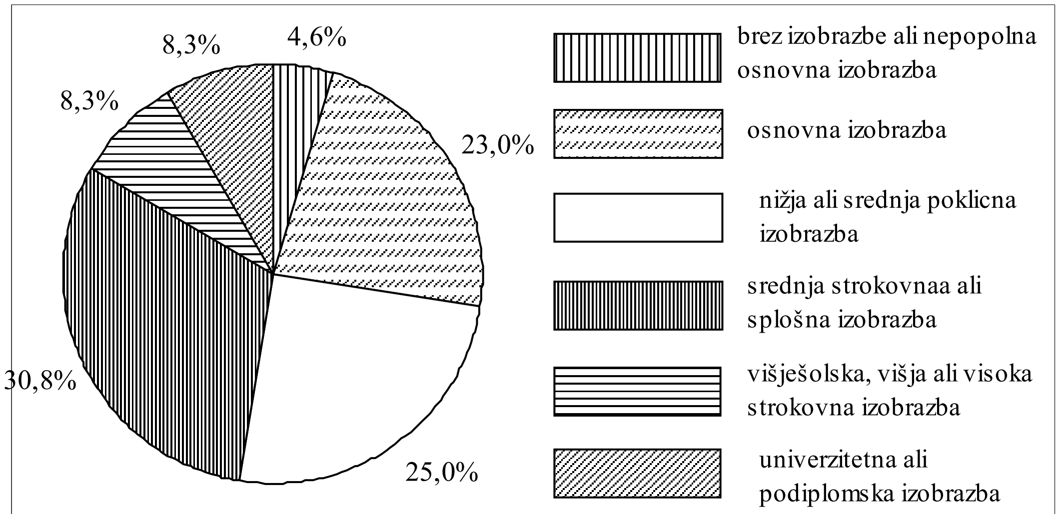
$$\boldsymbol{P^o = P\% \times 3{,}6} \quad (3.4)$$

S strukturnim krogom prikažimo strukturo prebivalstva, starega 15 let in več, po šolski izobrazbi konec leta 2006 (slika 3.3).

Tabela 3.8 **Struktura prebivalstva, starega 15 let in več, po izobrazbi v Sloveniji ob koncu leta 2006**

Šolska izobrazba	**Št. prebivalcev v tisoč**	**P %**	**P°**	**Delne vsote**
Brez izobrazbe ali nepopolna osnovna izobrazba	80	4,6	16	16
Osnovna izobrazba	397	23,0	83	99
Nižja ali srednja poklicna izobrazba	430	25,0	90	189
Srednja strokovna ali splošna izobrazba	530	30,8	111	300
Višja strokovna, višješolska ali visoka strokovna izobrazba	143	8,3	30	330
Univerzitetna ali podiplomska (spec., magisterij, doktorat) izobrazba	143	8,3	30	360
Skupaj	**1.723**	**100,0**	**360**	

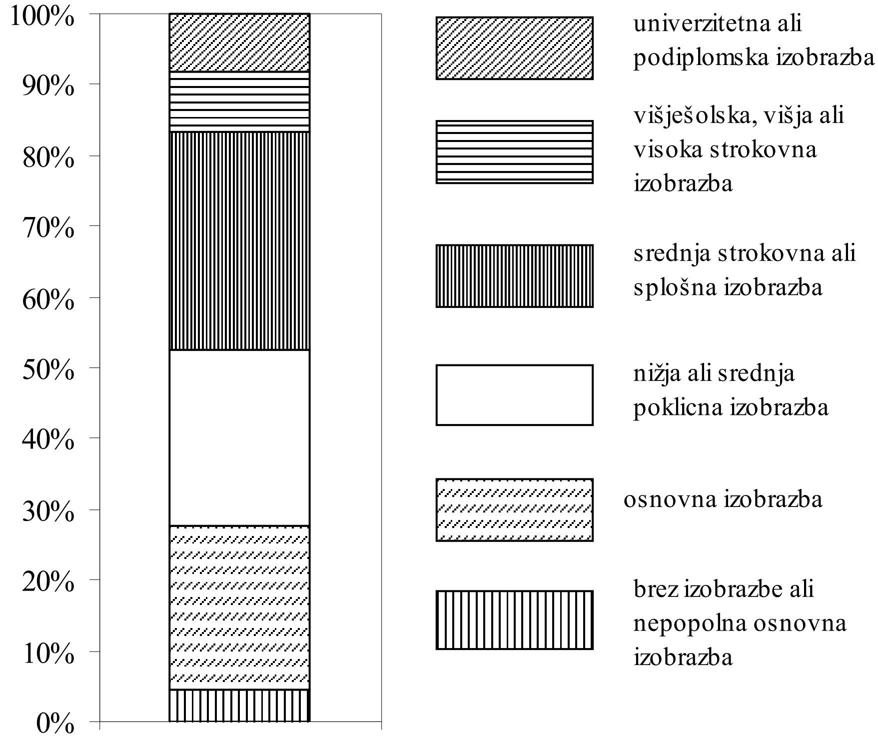
Vir: SURS, Statistični letopis 2007



Slika 3.3: **Struktura prebivalstva, starega 15 let in več, po izobrazbi konec leta 2006**

Vir: Tabela 3.8

Za primerjavo je struktura prebivalstva po šolski izobrazbi prikazana še s strukturnim stolpcem (slika 3.4).



Slika 3.4: **Struktura prebivalstva, starega 15 let in več, po izobrazbi konec leta 2006**

Vir: Tabela 3.8

3.2.3.3 Prikaz struktur z dvema krogoma in polkrogoma

Strukturo dveh pojavov lahko prikažemo tudi z dvema krogoma, pri tem lahko upoštevamo tudi velikost primerjanih pojavov. Ploščini krogov namreč prikažemo v sorazmerju z velikostjo pojavov. Poljubno izberemo polmer kroga r_A za pojav, ki ga označimo z Y_A, na osnovi tega izračunamo polmer kroga r_B za pojav Y_B po obrazcu:

$$\frac{r_B\pi}{r_A\pi} = \frac{Y_B}{Y_A}$$

$$r_B = r_A \times \sqrt{\frac{Y_B}{Y_A}} \qquad (3.5)$$

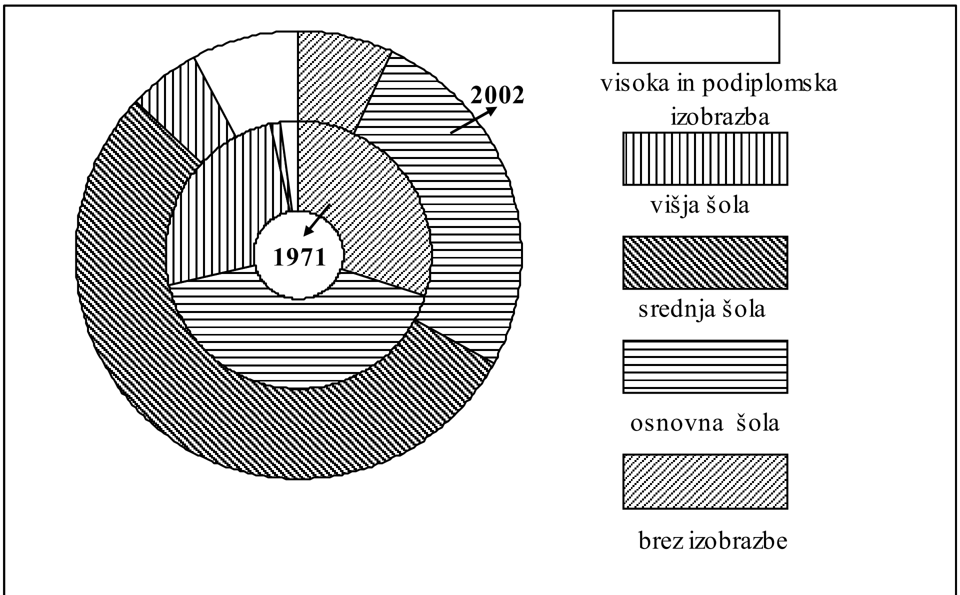
Primerjava je še nazornejša, če manjši krog narišemo koncentrično v večjega.

Namesto dveh krogov lahko narišemo polkroga, pri tem stopinje izračunamo tako, da odstotne deleže pomnožimo z 1,8.

Z dvema krogoma grafično prikažimo strukturo prebivalstva po šolski izobrazbi ob popisu 1971 in 2002. Pri tem je upoštevano število prebivalstva leta 1971 in 2002. Polmer kroga za število prebivalstva ob popisu 1971 je 3 cm, na osnovi tega je izračunan polmer kroga za število prebivalstva ob popisu 2002:

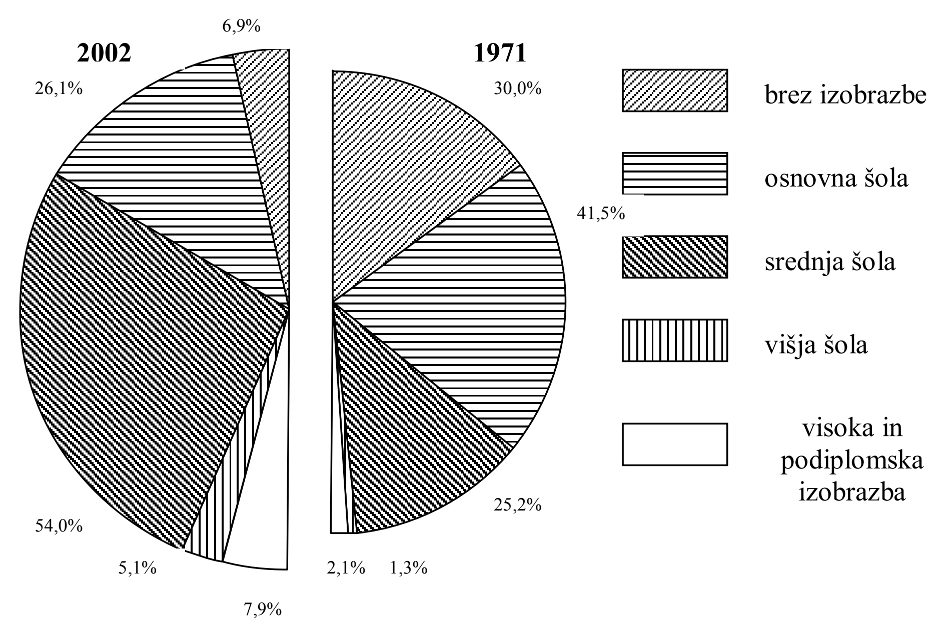
$$r_{2002} = r_{1971} \times \sqrt{\frac{Y_{2002}}{Y_{1971}}} = 3 \times \sqrt{\frac{1.663.869}{1.257.216}} = 3,5\ cm$$

Struktura prebivalstva po šolski izobrazbi je grafično prikazana še s polkrogoma (slika 3.6).



Slika 3.5: **Struktura prebivalstva, starega 15 let in več, po izobrazbi v Sloveniji ob popisih 1971 in 2002**

Vir: Tabela 3.7



Slika 3.1: **Struktura prebivalstva, starega 15 let in več, po izobrazbi, popisi 1971 in 2002**

Vir: Tabela 3.7

3.3 STATISTIČNI KOEFICIENTI

3.3.1 Pojem

Statistični koeficient je razmerje med dvema raznovrstnima podatkoma, katerih primerjava je vsebinsko smiselna. Če primerjamo število študentov s številom predavateljev, izraža koeficient število študentov na predavatelja, iz tega pa sklepamo na obremenjenost predavateljev in tudi na raven in kakovost višjega in visokega šolstva v določeni državi.

Podobno presojamo po koeficientu, ki kaže število zdravnikov na sto tisoč prebivalcev, raven zdravstvenega varstva. Po družbenem produktu na prebivalca sklepamo na gospodarsko razvitost, število avtomobilov na tisoč prebivalcev ali število telefonskih naročnikov na tisoč prebivalcev kažeta življenjsko raven prebivalstva.

Primerjana podatka se morata nanašati na isti časovni trenutek ali razmik. V teh primerih ni težav z izračunom koeficienta po obrazcu:

$$\boldsymbol{K = \frac{Y}{X}}, \tag{3.6}$$

kjer sta Y in X podatka, ki se nanašata na raznovrstna pojava in je Y podatek, ki je po vsebini v števcu koeficienta in ga primerjamo s podatkom X, ki je po vsebini v imenovalcu koeficienta. ***V imenovalcu koeficienta je podatek, na katerega primerjamo oziroma računamo koeficient.***

Primer 3.3: Iz podatkov o številu prodajaln in številu zaposlenih oseb v trgovini na drobno v Sloveniji konec leta 2006 bomo izračunali koeficient, ki kaže ***število zaposlenih oseb na prodajalno.***

Tabela 3.9: **Prodajalne in število zaposlenih oseb v trgovini na drobno ter izračunani koeficienti po vrsti trgovinske dejavnosti v Sloveniji konec leta 2006**

Vrsta trgovine	**Število prodajaln**	**Število zaposlenih**	**Število zaposlenih na prodajalno**
Trgovina z živili, s pijačami in tobakom	3.135	27.234	8,7
Trgovina z neživili	7.683	28.693	3,7
Trgovina z motornimi vozili in gorivi	1.322	4.427	3,3
SKUPAJ	12.140	60.354	5,0

Vir: Statistični letopis 2007

$$\text{Število zaposlenih na prodajalno v trgovini z živili, s pijačami in tobakom} = \frac{\text{število zaposlenih}}{\text{število prodajaln}} =$$

$$= \frac{27.234 \text{ zaposlenih}}{3.135 \text{ prodajaln}} = 8{,}7.$$

Število zaposlenih na prodajalno v trgovini z neživili = $\frac{\text{število zaposlenih}}{\text{število prodajaln}}$ =

$= \frac{28.693 \text{ zaposlenih}}{7.683 \text{ prodajaln}} = 3{,}7.$

Število zaposlenih na prodajalno v trgovini z motornimi vozili in gorivi = $\frac{\text{število zaposlenih}}{\text{število prodajaln}}$ =

$= \frac{4.427 \text{ zaposlenih}}{1.322 \text{ prodajaln}} = 3{,}3.$

Z izračunanimi koeficienti ugotovimo, da je bilo število zaposlenih na prodajalno največje, kar 8,7, v trgovini z živili, s pijačami in tobakom, v trgovini z neživili 3,7 in v trgovini z motornimi vozili in gorivi najmanj, samo 3,3 zaposlenega na prodajalno.

Statistični koeficienti so imenovana števila, saj z izračunano vrednostjo navedemo tudi merski enoti, v katerih merimo pojava, iz katerih je koeficient izračunan.

Izračunani koeficienti se nanašajo na eno enoto podatka v imenovalcu, npr. na prodajalno. V nekaterih primerih računamo koeficient na 10, 100, 1.000 ali 100.000 enot podatka v imenovalcu. Zato lahko zapišemo:

$$\boldsymbol{K = \frac{Y}{X} \times E} \qquad (3.7)$$

E – je 1, 10, 100, 1.000, 100.000, **odvisno od tega, na koliko enot podatka v imenovalcu računamo koeficient.** Tako npr. računamo število prebivalcev na zdravnika ali pa število zdravnikov na 100.000 prebivalcev.

Primer 3.4: Izračunajmo vrednost uvoza na 1.000 evrov izvoza v Sloveniji v letih 2001 do 2007.

Tabela 3.10: **Izvoz in uvoz v Sloveniji v letih od 2001 do 2007 in vrednost uvoza na 1.000 evrov izvoza**

Leto	Izvoz v mio EUR	Uvoz v mio EUR	Vrednost uvoza na 1.000 evrov izvoza
2001	10.346,8	11.344,5	1.096,43
2002	10.962,0	11.574,1	1.055,84
2003	11.285,0	12.238,9	1.084,53
2004	12.783,1	14.143,0	1.106,38
2005	14.314,5	15.728,2	1.098,76
2006	16.757,2	18.340,8	1.094,50
2007	19.385,2	21.486,7	1.108,41

Vir: Pomembnejši statistični podatki o Sloveniji, letnik III, št. 2/2008

Vrednost uvoza na 1.000 EUR izvoza 2001 = $\frac{\text{vrednost uvoza}}{\text{vrednost izvoza}} \times 1.000 =$

$= \frac{11.344{,}5 \text{ mio EUR}}{10.346{,}8 \text{ mio EUR}} \times 1.000 = 1.096{,}43 \text{ EUR}.$

Iz tabele vidimo, da je vsa leta vrednost uvoza na 1.000 evrov vrednosti izvoza večja od 1.000 evrov, kar za državo ni ugodno.

3.3.2 Primerjava podatka, ki se nanaša na časovni interval, s podatkom, ki se nanaša na časovni trenutek

Časovna opredelitev pojavov, iz katerih smo izračunali koeficiente, ki so v tabelah 3.9 in 3.10, je bila enaka. Tako se podatki o številu prodajaln in zaposlenih oseb v trgovini nanašajo na trenutek – konec leta, podatki o izvozu in uvozu pa na časovni razmik – npr. leto 2001. S primerjavo podatkov zato ni bilo težav.

Če pa se eden od podatkov, ki ju želimo primerjati, nanaša na časovni trenutek, drugi pa na časovni razmik, ju ni mogoče neposredno primerjati. Podatke, ki se nanašajo na časovne trenutke, moramo prilagoditi za primerjavo, tako da izračunamo povprečje, ki se mora nanašati na tisti časovni razmik, za katerega upoštevamo razmični podatek. Tako so v statističnih publikacijah pogosto že objavljeni podatki, ki izražajo povprečje, npr. število prebivalcev sredi leta, število brezposelnih oseb sredi meseca in podobno.

Tako je obrazec za izračun koeficienta: $$\boldsymbol{K = \frac{Y}{\overline{X}} \times E,} \qquad (3.8)$$

kjer je $\overline{X}$ izračunano povprečje podatkov, ki se nanašajo na časovne trenutke.

Povprečje za podatke, ki se nanašajo na časovne trenutke, izračunamo odvisno od tega, v katerem trenutku pojav opazujemo:

- ***če pojav opazujemo v sredini obdobij***, izračunamo povprečje iz N podatkov tako:

$$\boldsymbol{\overline{X} = \frac{1}{N}\left(X_1 + X_2 + \dots + X_N\right)} \qquad (3.9)$$

- ***če pojav opazujemo na začetku ali koncu obdobij***, izračunamo povprečje tako, da najprej izračunamo povprečja za posamezna obdobja, nato pa iz teh povprečij izračunamo povprečje za celotno opazovano obdobje:

$$\boldsymbol{\overline{X} = \frac{1}{N}\left(\overline{X}_1 + \overline{X}_2 + \dots + \overline{X}_N\right) = \left(\frac{X_0 + X_1}{2} + \frac{X_1 + X_2}{2} + \dots + \frac{X_{N-1} + X_N}{2}\right) =} \qquad (3.10)$$
$$\boldsymbol{= \frac{1}{N}\left(\frac{X_0}{2} + X_1 + X_2 + \dots + X_{N-1} + \frac{X_N}{2}\right)}$$

Primer 3.5: S primerjavo podatkov v tabeli št. 3.11 izračunajmo vrednost izvoza na prebivalca v Sloveniji v letih 2003 do 2007.

Ker se podatki o številu prebivalcev nanašajo na konec leta, podatki o vrednosti izvoza pa na časovni razmik enega leta, izračunamo število prebivalcev sredi leta:

$$\text{Število prebivalcev sredi leta 2004} = \frac{1.996.433 + 1.997.590}{2} = 1.997.011{,}5.$$

in nato vrednost izvoza na prebivalca:

$$\text{Vrednost izvoza na prebivalca v letu 2004} = \frac{\text{vrednost izvoza v letu 2004}}{\text{srednje število prebivalcev v letu 2004}} =$$

$$= \frac{8.505.359 \text{ tisoč EUR}}{1.997.011{,}5 \text{ prebivalcev}} = 4.259{,}04 \text{ EUR.}$$

$$\text{Vrednost izvoza na prebivalca v letu 2005} = \frac{\text{vrednost izvoza v letu 2005}}{\text{srednje število prebivalcev v letu 2005}} =$$

$$= \frac{9.770.474 \text{ tisoč EUR}}{2.000.477 \text{ prebivalcev}} = 4.884{,}07 \text{ EUR.}$$

itd.

Tabela 3.11: **Število prebivalcev in vrednost izvoza v Sloveniji v letih od 2003 do 2007 ter izračunani koeficient vrednost izvoza na prebivalca**

Leto	Vrednost izvoza v 1.000 EUR	Število prebivalcev konec leta	Število prebivalcev sredi leta	Vrednost izvoza v EUR na prebivalca
2003	-	1.996.433		-
2004	8.505.359	1.997.590	1.997.011,5	4.259,04
2005	9.770.474	2.003.358	2.000.477,0	4.884,07
2006	11.465.282	2.010.377	2.006.867,5	5713,02
2007	13.686.843	2.025.866	2.018.121,5	6.781,97

Vir: Pomembnejši statistični podatki o Sloveniji, letnik III; št. 2/2008

Ugotovimo, da vrednost izvoza na prebivalca po letih narašča.

3.3.3 Koeficient obračanja zalog

V trgovskih in drugih podjetjih je smiselno računati koeficient obračanja zalog kot razmerje med vrednostjo prodaje (v proizvodnih podjetjih tudi vrednost porabljenih surovin) in vrednostjo povprečne zaloge. Vrednost prodaje se nanaša na razmik, povprečno zalogo pa izračunamo iz podatkov, ki se nanašajo na opazovani trenutek (lahko je začetek ali konec meseca, ali pa je že izračunano mesečno povprečje). Izračunani koeficient, ki je neimenovano število, saj sta primerjana podatka izražena z vrednostjo, pove, kolikokrat je prodaja večja od zaloge. V poslovni praksi koeficient obračanja zalog razložimo bolj z vsebinskega vidika in pove, kolikokrat se zaloge obrnejo na določeno časovno enoto, npr. letno, mesečno.

$$K_{obračanja\ zalog} = \frac{promet}{zaloga} \times čas \qquad (3.11)$$

Primer 3.6: Izračunajmo koeficient obračanja zalog s podatki v tabeli 3.12.

Tabela 3.12: **Vrednost prodaje in zaloge ter podatki o številu prodajalcev v prodajalni z živili *Košarica* po mesecih za prvo polletje leta 2008**

Mesec	**Vrednost prodaje v 1.000 EUR**	**Vrednost zaloge v 1.000 EUR na začetku meseca**	**Srednje število prodajalcev**
Januar	278	129	4
Februar	335	131	5
Marec	356	117	5
April	384	95	5
Maj	333	107	5
Junij	362	98	6
Julij	-	106	-

Vir: Fiktivni podatki

Izračunajmo:

- ***povprečno mesečno vrednost prodaje na prodajalca v prvem polletju in***
- ***mesečni koeficient obračanja zalog v prvem polletju.***

Povprečna mesečna vrednost prodaje na prodajalca:

$$K_{\text{januar}-\text{junij}} = \frac{\text{povprečna mesečna vrednost prodaje}}{\text{povprečno mesečno število prodajalcev}} =$$

$$= \frac{341{,}33 \text{ tisoč EUR}}{5 \text{ prodajalcev}} = 68.266 \text{ EUR na prodajalca povprečno mesečno}$$

$$\text{Povprečna mesečna prodaja} = \frac{1}{6}\left(278+335+356+384+333+362\right) = \frac{2048}{6} = 341{,}33 \text{ tisoč EUR.}$$

$$\text{Povprečno mesečno število prodajalcev} = \frac{1}{6}\left(4+5+5+5+5+6\right) = \frac{30}{6} = 5.$$

Mesečni koeficient obračanja zalog

Izračunamo še povprečno mesečno zalogo (stanje zaloge se nanaša na začetek meseca, zato imamo podatek še za začetek julija, sicer ne bi mogli izračunati povprečje za mesec junij):

$$\text{Povprečna mesečna zaloga} = \frac{1}{6}\left(\frac{129}{2}+131+117+95+107+98+\frac{106}{2}\right) = \frac{665{,}5}{6} = 110{,}92 \text{ tisoč EUR.}$$

$$K_{\text{obračanja zalog}} = \frac{\text{povprečna mesečna vrednost prodaje}}{\text{povprečna mesečna vrednost zaloge}} \times \text{čas} = \frac{341{,}33 \text{ tisoč EUR}}{110{,}92 \text{ tisoč EUR}} \times 1 = 3{,}1$$

V prvem polletju se je zaloga povprečno mesečno obrnila 3,1-krat.

3.3.4 Recipročni koeficient

Nekatere koeficiente lahko izrazimo na dva načina. Govorimo lahko o številu študentov na predavatelja ali o številu predavateljev na sto študentov, o površini prodajnega prostora na prodajalno ali o številu prodajaln na tisoč m^2 prodajnega prostora. V takih primerih govorimo o recipročnih koeficientih. Če prvi koeficient izračunamo $\boldsymbol{K = \frac{Y}{X}}$,

velja za izračun recipročnega koeficienta naslednji obrazec: $\boldsymbol{K_{rec} = \frac{X}{Y} = \frac{1}{K}}$ (3.12)

V tabeli 3.13 so podatki o številu prodajaln in površini prodajnega prostora v trgovini na drobno za najmanj in najbolj razvito slovensko regijo. Izračunajmo:

- ***povprečno površino prodajnega prostora na prodajalno in***
- ***obratni koeficient – število prodajaln na 1.000 m^2 površine prodajnega prostora.***

Tabela 3.13: **Število prodajaln in prodajni prostor v trgovini na drobno v Pomurski in Osrednjeslovenski regiji 30. 6. 2006 ter izračunani koeficienti**

Regija	Število prodajaln	Prodajni prostor v m^2	Prodajni prostor na prodajalno v m^2	Št. prodajaln na 1.000 m^2
Pomurska	737	107.161	145,4	6,9
Osrednjeslovenska	3.005	584.398	194,5	5,1

Vir: Statistični letopis 2007

Izračunana sta koeficienta za Pomursko regijo:

$$\text{Površina prodajnega prostora na prodajalno} = \frac{\text{površina v m}^2}{\text{število prodajaln}} =$$

$$= \frac{107.161\ \text{m}^2}{737\ \text{prodajaln}} = 145{,}4\ \text{m}^2.$$

$$\text{Število prodajaln na 1.000 m}^2\text{površine} = \frac{\text{število prodajaln}}{\text{površina prodajnega prostora v m}^2} \times 1.000 =$$

$$= \frac{737\ \text{prodajaln}}{107.161\ \text{m}^2} \times 1.000 = 6{,}9.$$

Ali:

$$\text{Število prodajaln na 1.000 m}^2\ \text{površine} = \frac{1}{\text{K}_{\text{površina prodajnega prostora na prodajalno}}} \times 1.000 =$$

$$= \frac{1}{145{,}4\ \text{m}^2} \times 1.000 = 6{,}9.$$

Razlika v izračunanih koeficientih obeh regij je velika, saj je bil v Pomurski regiji prodajni prostor na prodajalno 145,4 m^2, v Osrednjeslovenski pa 194,5 m^2, v Pomurski regiji je bilo na 1.000 m^2 prodajnega prostora 6,9, v Osrednjeslovenski pa 5,1 prodajalne.

Obrnemo lahko tudi koeficient obračanja zalog. Če pri tem upoštevamo povprečno mesečno število delovnih dni (običajno 25 dni), nam recipročni koeficient kaže dolžino trajanja enega obrata v dnevih ali povprečni čas skladiščenja blaga.

$$K_{rec} = \frac{\text{povprečna mesečna vrednost zaloge}}{\text{povprečna mesečna vrednost prodaje}} \times 25 \text{ (dni)} = \frac{110{,}92 \text{ tisoč EUR}}{341{,}33 \text{ tisoč EUR}} \times 25 = 8{,}12 \text{ dneva}$$

Povprečni čas skladiščenja blaga ali dolžina enega obrata je 8,12 dneva.

3.4 INDEKSI

Indeksi so relativna števila, s katerimi analiziramo relativne spremembe med členi krajevne, časovne ali stvarne vrste. Računamo jih iz istovrstnih podatkov, torej podatkov, ki se nanašajo na isto spremenljivko. Indeks izračunamo tako, da primerjana podatka med seboj delimo in dobljeno razmerje pomnožimo s 100.

$$\boldsymbol{I_{j/0} = \frac{Y_j}{Y_0} \times 100} \tag{3.13}$$

Y_j – podatek, ki ga primerjamo s podatkom v imenovalcu

Y_0 – podatek, s katerim delimo podatek v števcu, imenujemo ga osnova ali baza

- **indeks ima vrednost večjo od 100 ⇒ $I_{j/0}$ > 100,** če je primerjani podatek Y_j večji od osnove Y_0
- **indeks ima vrednost enako 100 ⇒ $I_{j/0}$ = 100**, če sta primerjana podatka enaka
- **indeks ima vrednost manjšo od 100 ⇒ $I_{j/0}$ < 100**, če je primerjani podatek Y_j manjši od osnove Y_0

Indeks razložimo tako, da od indeksa odštejemo 100 in dobimo relativno razliko $\mathbf{D_{j/0}}$**,** ki jo izrazimo v odstotkih in je:

- **večja od 0, torej pozitivna, če je $I_{j/0}$ > 100;**
- **enaka 0, če je $I_{j/0}$ = 100;**
- **manjša od 0, torej negativna, če je $I_{j/0}$ < 100.**

Ko preučujemo spreminjanje pojava v času, relativno razliko imenujemo stopnja rasti in jo označimo s S_j.

Primer 3.7: Primerjajmo število prenočitev domačih turistov in turistov iz tujine v Sloveniji v mesecu decembru glede na november v letu 2007.

Domači turisti:

$$I_{december/november} = \frac{220.035}{202.501} \times 100 = 108{,}7$$

$$S_{december/november} = 108{,}7 - 100 = 8{,}7\ \%$$

Turisti iz tujine:

$$I_{december/november} = \frac{225.187}{238.211} \times 100 = 94{,}5$$

$$S_{december/november} = 94{,}5 - 100 = -5{,}5\ \%$$

Tabela 3.13: **Prenočitve domačih turistov in turistov iz tujine v Sloveniji v mesecu novembru in decembru 2007**

Mesec	**Prenočitve turistov**	
	Domači	**Iz tujine**
November	202.501	238.211
December	220.035	225.187
$\text{Indeks}_{dec/nov}$	108,7	94,5

Vir: Pomembnejši statistični podatki o Sloveniji, letnik III, št. 2/2008

V mesecu decembru leta 2007 je bilo 8,7 % več prenočitev domačih turistov in 5,5 % manj prenočitev turistov iz tujine kot v mesecu novembru istega leta.

3.4.1 Krajevni indeksi

Indeks, ki ga izračunamo iz dveh istovrstnih podatkov, ki se nanašata na dve geografski področji, je krajevni indeks.

Primer 3.8: Izračunajmo indekse za bruto družbeni produkt v Sloveniji in sosednjih državah za leto 2007 glede na povprečje držav članic EU.

Tabela 3.14***:*** **Bruto družbeni proizvod v tekočih cenah v standardih kupne moči (PPS[2])v Sloveniji in sosednjih državah v letu 2007 in izračunani indeksi**

Država	**BDP v tekočih cenah v EUR**	**Indeksi (EU = 100)**
Avstrija	31.600	127,4
Hrvaška	13.900	56,0
Italija	25.200	101,6
Madžarska	15.700	63,3
Slovenija	22.000	88,7
Povprečje EU	**24.800**	**100,0**

Vir: www.http://epp.eurostat.ec.europa.eu (15. 8. 2008)

$$I_{Avstrija/\,povprečje\,EU} = \frac{31.600}{24.800} \times 100 = 127,4 \qquad I_{Slovenija/\,povprečje\,EU} = \frac{22.000}{24.800} \times 100 = 88,7$$

$$S_{Avstrija/\,povprečje\,EU} = 127,4 - 100 = 27,4\ \% \qquad S_{Slovenija/\,povprečje\,EU} = 88,7 - 100 = -11,3\ \%$$

Pri izračunu indeksov smo za osnovo določili povprečni bruto družbeni produkt držav članic Evropske unije, saj nas zanima, kako družbeni produkt primerjanih držav odstopa od povprečja. Izračunani indeksi kažejo relativne razlike bruto družbenega produkta v Sloveniji in njenih sosednjih državah glede na povprečje v EU. Ugotovimo, da je bil le-ta v Sloveniji za 11,3 % pod povprečjem, v Avstriji za 27,4 % in v Italiji za 1,6 % nad povprečjem, v Madžarski za 36,7 % in v Hrvaški za 44 % pod povprečjem držav članic EU.

[2] PPS – Purchasing Power Standard

3.4.2 Časovni indeksi

S primerjavo dveh istovrstnih podatkov, ki se nanašata na različna časovna trenutka ali razmika, izračunamo časovni indeks.

Časovne indekse je smiselno računati za pojav, pri katerem spremljamo dinamiko oziroma gibanje v odvisnosti od časa za daljše časovno obdobje. Časovna vrsta sicer že sama prikazuje dinamiko pojava, z izračunanimi indeksi pa je ta še bolj nazorna.

Iz časovnih vrst lahko izračunamo:

- ***indekse s stalno osnovo***, ko vsak člen časovne vrste primerjamo s členom, ki smo ga določili za osnovo: $\boldsymbol{I_{j/0} = \frac{Y_j}{Y_0} \times 100}$ *(3.14)*

Y_j – vrednost j-tega člena v časovni vrsti za dano obdobje (dani trenutek ali razmik)
Y_o – vrednost člena, ki smo ga določili za osnovo

- ***verižne indekse***, ko vsak člen časovne vrste primerjamo s členom prejšnjega obdobja; osnova je torej predhodni člen: $\boldsymbol{V_j = \frac{Y_j}{Y_{j-1}} \times 100}$ *(3.15)*

Y_j – vrednost j-tega člena v časovni vrsti za dano obdobje
Y_{j-1} – vrednost člena za predhodno obdobje

Primer 3.9: Izračunajmo indekse s stalno osnovo in verižne indekse za prihode domačih turistov v Sloveniji v letih 2002 do 2007.

Tabela 3.15: **Prihodi domačih turistov v Sloveniji v letih od 2002 do 2007 ter izračunani indeksi s stalno osnovo (2004 = 100) in verižni indeksi**

Leto	Prihodi turistov v tisoč	Indeksi s stalno osnovo 2004 = 100 $I_{j/0}$	Verižni indeksi V_j
2002	860,4	102,1	-
2003	872,9	103,6	101,5
2004	842,4	**100,0**	96,5
2005	840,0	99,7	99,7
2006	867,2	102,9	103,2
2007	926,6	110,0	106,8

Vir: Pomembnejši statistični podatki o Sloveniji, letnik III, št. 2/2008

Indeksi s stalno osnovo 2004 = 100

$$I_{2002/2004} = \frac{Y_{2002}}{Y_{2004}} \times 100 = \frac{860,4}{842,4} \times 100 = 102,1 \rightarrow S_{2002/2004} = 102,1 - 100 = 2,1\,\%$$

$$I_{2003/2004} = \frac{Y_{2003}}{Y_{2004}} \times 100 = \frac{872,9}{842,4} \times 100 = 103,6 \rightarrow S_{2003/2004} = 103,6 - 100 = 3,6\,\%$$

$$I_{2005/2004} = \frac{Y_{2005}}{Y_{2004}} \times 100 = \frac{840,0}{842,4} \times 100 = 99,7 \rightarrow S_{2005/2004} = 99,7 - 100 = -0,3\ \%$$

itd.

V letu 2002 je bilo za 2,1 % in v letu 2003 za 3,6 % več prihodov domačih turistov kot leta 2004, leta 2005 pa za 0,3 % manj kot leta 2004. Leta 2006 je bilo za 2,9 % in leta 2007 za 10 % več prihodov domačih turistov kot leta 2004.

Verižni indeksi:

$$V_{2003} = \frac{Y_{2003}}{Y_{2002}} \times 100 = \frac{872,9}{860,4} \times 100 = 101,5 \rightarrow S_{2003} = 101,5 - 100 = 1,5\ \%$$

$$V_{2004} = \frac{Y_{2004}}{Y_{2003}} \times 100 = \frac{842,4}{872,9} \times 100 = 96,5 \rightarrow S_{2004} = 96,5 - 100 = -3,5\ \%$$

$$V_{2005} = \frac{Y_{2005}}{Y_{2004}} \times 100 = \frac{840,0}{842,4} \times 100 = 99,7 \rightarrow S_{2005} = 99,7 - 100 = -0,3\ \%$$

itd.

Leta 2003 je bilo za 1,5 % več prihodov domačih turistov kot leta 2002, leta 2004 za 3,5 % manj kot leta 2003, leta 2005 za 0,3 % manj kot leta 2004, leta 2006 za 3,2 % več kot leta 2005 in leta 2007 za 6,8 % več kot leta 2006.

Indeksi s stalno osnovo prikazuje spremembe členov v časovni vrsti z nekim karakterističnim členom, ki smo ga določili za osnovo. Verižni indeksi pa prikazujejo spremembe od člena do člena časovne vrste.

3.4.3 Preračunavanje indeksov

- ***Preračunavanje indeksov s stalno osnovo na novo osnovo***

je smiselno, kadar primerjamo dinamiko več pojavov, za katere so dane indeksne vrste z različno osnovo, nimamo pa osnovnih podatkov. Tako ne moremo izračunati indeksov iz osnovne časovne vrste.

V tem primeru indekse preračunamo na novo osnovo enostavno tako, da vsak indeks indeksne vrste delimo z indeksom tistega obdobja (meseca, leta), ki smo ga določili za novo osnovo.

Indekse s stalno osnovo za prihode domačih turistov v Sloveniji 2004 = 100 preračunamo na novo osnovo 2002 = 100.

$$I_{2003/2002} = \frac{I_{2003/2004}}{I_{2002/2004}} \times 100 = \frac{103,6}{102,1} \times 100 = 101,5 \rightarrow S_{2003/2002} = 101,5 - 100 = 1,5\ \%$$

$$I_{2004/2002} = \frac{I_{2004/2004}}{I_{2002/2004}} \times 100 = \frac{100,0}{102,1} \times 100 = 97,9 \rightarrow S_{2004/2002} = 97,9 - 100 = -2,1\ \%$$

itd.

Tabela 3.16: **Indeksi s stalno osnovo 2004 = 100 in preračunani indeksi 2002 = 100 za prihode domačih turistov v Sloveniji v letih od 2002 do 2007**

Leto	**Indeksi s stalno osnovo 2004 = 100** $I_{j/2004}$	**Indeksi s stalno osnovo 2002 = 100** $I_{j/2002}$
2002	102,1	**100,0**
2003	103,6	101,5
2004	**100,0**	97,9
2005	99,7	97,6
2006	102,9	100,8
2007	110,0	107,7

Vir: Tabela 3.15

Leta 2003 je bilo 1,5 % več, leta 2004 pa 2,1 % manj prihodov domačih turistov kot leta 2002.

- ***Preračunavanje verižnih indeksov v indekse s stalno osnovo***

Verižni indeksi kažejo relativne spremembe od člena do člena v časovni vrsti. Če imamo dano vrsto verižnih indeksov, ne pa tudi osnovnih podatkov, želimo pa prikazati relativne spremembe členov glede na nek karakteristični člen, verižne indekse preračunamo v indekse s stalno osnovo.

Verižni indeksi za prihode domačih turistov v Sloveniji so preračunani v indekse s stalno osnovo 2004 = 100.

Tabela 3.17: **Verižni indeksi in preračunani indeksi na stalno osnovo 2004 = 100 za prihode domačih turistov v Sloveniji v letih od 2002 do 2007**

Leto	**Verižni indeksi** V_j	**Indeksi s stalno osnovo 2004 = 100** $I_{j/2004}$
2002	-	102,1
2003	101,5	103,6
2004	96,5	**100,0**
2005	99,7	99,7
2006	103,2	102,9
2007	106,8	110,0

Postopek preračunavanja:

- ***indeks s stalno osnovo za obdobje pred izhodiščnim obdobjem izračunamo*** tako, da indeks s stalno osnovo izhodiščnega obdobja delimo z verižnim indeksom tega obdobja in izračunani količnik pomnožimo s 100; za ostala obdobja nadaljujemo s postopnim deljenjem:

$$I_{2003/2004} = \frac{I_{2004/2004}}{V_{2004}} \times 100 = \frac{100,0}{96,5} \times 100 = 103,6$$

$$I_{2002/2004} = \frac{I_{2003/2004}}{V_{2003}} \times 100 = \frac{103,6}{101,5} \times 100 = 102,1$$

- ***indeks s stalno osnovo za izhodiščnim obdobjem*** izračunamo tako, da indeks s stalno osnovo izhodiščnega leta pomnožimo z verižnim indeksom naslednjega obdobja in dobljeni rezultat delimo s 100; za ostala obdobja nadaljujemo s postopnim množenjem:

$$I_{2005/2004} = \frac{I_{2004/2004} \times V_{2005}}{100} = \frac{100,0 \times 99,7}{100} = 99,7$$
$$I_{2006/2004} = \frac{I_{2005/2004} \times V_{2006}}{100} = \frac{99,7 \times 103,2}{100} = 102,9$$
$$I_{2007/2004} = \frac{I_{2006/2004} \times V_{2007}}{100} = \frac{102,9 \times 106,8}{100} = 110,0$$

Postopek lahko poenostavimo tako, da verižne indekse spremenimo v koeficiente rasti, ki so prikazani v naslednji točki.

Pred izhodiščnim obdobjem:

$$I_{2003/2004} = \frac{I_{2004/2004}}{K_{2004}} = \frac{100,0}{0,965} = 103,6$$

Za izhodiščnim obdobjem:

$$I_{2005/2004} = I_{2004/2004} \times K_{2005} = 100,0 \times 0,997 = 99,7$$

3.4.4 Koeficienti in stopnje rasti

Za analizo dinamike časovnih vrst uporabljamo tudi

- ***koeficient rasti***, ki izraža razmerje med dvema zaporednima vrednostima členov v časovni vrsti in ga izračunamo: $\boldsymbol{K_j = \frac{Y_j}{Y_{j-1}}}$ *(3.16)*

  Tako kot verižni indeks tudi ta kaže spremembe od člena do člena v časovni vrsti, vendar v obliki koeficienta.

- ***stopnjo rasti***, ki kaže neposredno odstotno spremembo med dvema zaporednima vrednostima členov v časovni vrsti in jo izračunamo: $\boldsymbol{S_j = \frac{Y_j - Y_{j-1}}{Y_{j-1}} \times 100}$ *(3.17)*

Izračunani koeficienti in stopnje rasti za prihode domačih turistov v Sloveniji so razvidni iz tabele 3.18.

$$K_{2003} = \frac{872,9}{860,4} = 1,015 \text{ in iz tega: } S_{2003} = (1,015 - 1) \times 100 = 1,5\ \%$$
$$\mathrm{K}_{2004} = \frac{842,4}{872,9} = 0,965 \text{ in iz tega: } \mathrm{S}_{2004} = (0,965 - 1) \times 100 = -3,5\ \%$$

Stopnje rasti lahko izračunamo tudi neposredno:

$$S_{2003} = \frac{872{,}9 - 860{,}4}{860{,}4} \times 100 = 1{,}5\ \%$$

$$S_{2004} = \frac{842{,}4 - 872{,}9}{872{,}9} \times 100 = -3{,}5\ \%$$

Tabela 3.18: **Koeficienti in stopnje rasti za prihode domačih turistov v Sloveniji v letih 2002 do 2007**

Leto	Prihodi domačih turistov v tisoč	Koeficienti rasti K_j	Stopnje rasti S_j
2002	860,4	-	-
2003	872,9	1,015	1,5
2004	842,4	0,965	−3,5
2005	840,0	0,997	−0,3
2006	867,2	1,032	3,2
2007	926,6	1,068	6,8

3.4.5 Vrednosti posameznih kazalcev dinamike, ko pojav narašča, pada ali stagnira

Kazalec dinamike	Pojav narašča	Pojav pada	Pojav stagnira
Verižni indeks – V_j	$V_j > 100$	$V_j < 100$	$V_j = 100$
Koeficient rasti – K_j	$K_j > 1$	$K_j < 1$	$K_j = 1$
Stopnja rasti – S_j	$S_j > 0$	$S_j < 0$	$S_j = 0$

Povezava med kazalci dinamike:

$$S_j = \frac{y_j - y_{j-1}}{y_{j-1}} \times 100 = \frac{y_j}{y_{j-1}} \times 100 - \frac{y_{j-1}}{y_{j-1}} \times 100 = V_j - 100 \tag{3.19}$$

$$\begin{aligned} S_j &= V_j - 100 = 100\ K_j - 100 \text{ in} \\ S_j &= 100 \times (K_j - 1) \text{ in iz tega} \end{aligned} \tag{3.20}$$

$$K_j = \frac{S_j}{100} + 1 \tag{3.21}$$

3.4.6 Grafično prikazovanje indeksov

Indekse s stalno osnovo grafično prikažemo z linijskim grafikonom, kar je primerno predvsem takrat, ko prikazujemo indeksne vrste več pojavov, katerih dinamiko želimo primerjati.

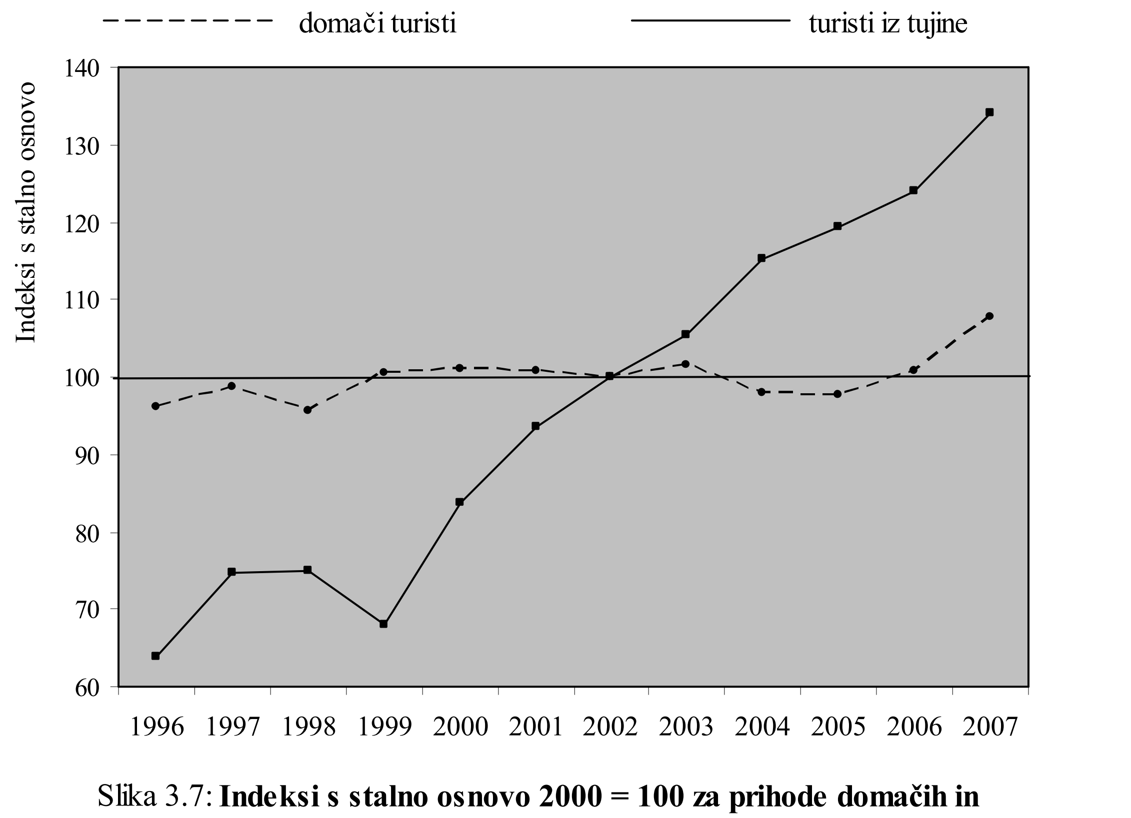
Na sliki 3.6 so prikazani indeksi s stalno osnovo za prihode domačih in tujih turistov v letih od 1996 do 2007, ki so izračunani v tabeli 3.19.

Verižne indekse prikažemo s stolpci. V grafikonu lahko odčitamo stopnje rasti. N sliki 3.7 so prikazani verižni indeksi za prihode domačih turistov v Sloveniji v letih 1997 do 2007, ki so izračunani v tabeli 3.20.

Tabela 3.19: **Indeksi s stalno osnovo za prihode domačih in tujih turistov v Sloveniji v letih od 1996 do 2007**

Leto	Prihodi domačih turistov		Prihodi turistov iz tujine	
	Število v tisoč	Indeksi 2002 = 100	Število v tisoč	Indeksi 2002 = 100
1996	826	96,0	832	63,9
1997	849	98,7	974	74,8
1998	822	95,6	977	75,0
1999	865	100,6	884	67,9
2000	868	100,9	1090	83,7
2001	867	100,8	1219	93,6
2002	860	**100,0**	1302	**100,0**
2003	873	101,5	1373	105,5
2004	842	97,9	1499	115,1
2005	840	97,7	1555	119,4
2006	867	100,8	1615	124,0
2007	927	107,8	1746	134,1

Vir: Statistični letopis 2005, 433 in Pomembnejši statistični podatki o Sloveniji, št. 2/2008



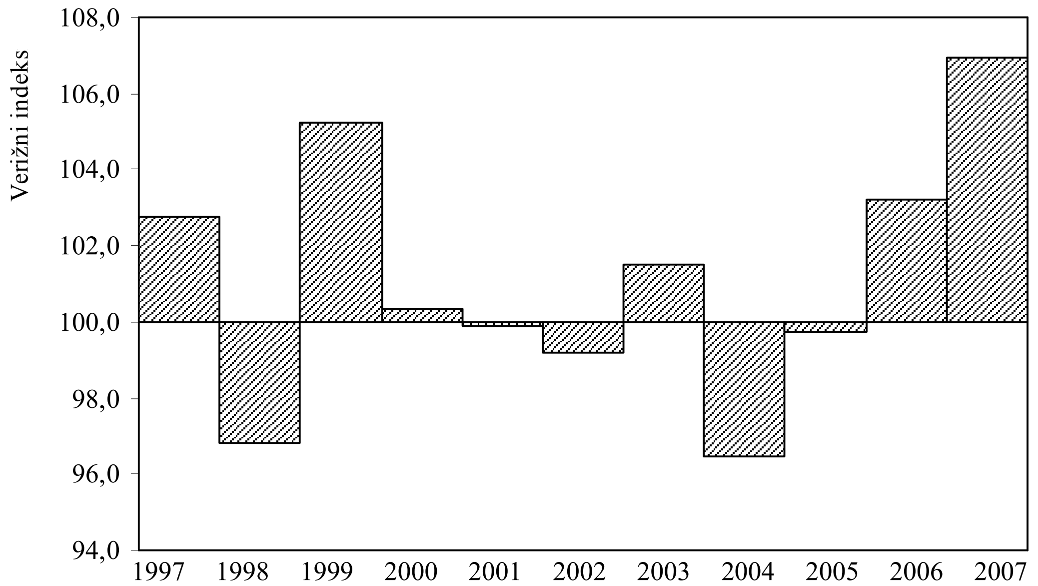
Slika 3.7: **Indeksi s stalno osnovo 2000 = 100 za prihode domačih in tujih turistov v Slovenijo v letih od 1996 do 2007**

Vir: Tabela 3.19

Tabela 3.20: **Verižni indeksi za prihode domačih turistov v Sloveniji v letih od 1997 do 2007**

Leto	Prihodi domačih turistov	
	Število v tisoč	**Verižni indeksi**
1996	826	-
1997	849	102,8
1998	822	96,8
1999	865	105,2
2000	868	100,3
2001	867	99,9
2002	860	99,2
2003	873	101,5
2004	842	96,4
2005	840	99,8
2006	867	103,2
2007	927	106,9

Vir: Statistični letopis 2005, 433 in Pomembnejši statistični podatki o Sloveniji, letnik III, št. 2/2008



Slika 3.8: **Verižni indeksi za prihode domačih turistov v Sloveniji v letih od 1997 do 2007**

Vir: Tabela 3.20

3.4.7 Računanje vrednosti členov časovne vrste z danimi indeksi, če je dana vrednost samo enega člena vrste

Dinamiko določenega pojava pogosto prikažemo le z indeksno vrsto, na pa tudi z osnovno časovno vrsto. V tem primeru lahko izračunamo člene časovne vrste, če poznamo vrednost vsaj enega člena. Način računanja je odvisen od danih indeksov; ali so dani indeksi s stalno osnovo ali verižni indeksi.

- **Dani so indeksi s stalno osnovo**

Za prihode domačih turistov v Sloveniji v letih od 2002 do 2007 imamo dane indekse s stalno osnovo 2002=100 in število prihodov v letu 2002, ko jih je bilo 860 tisoč.

Tabela 3.21: **Indeksi s stalno osnovo za prihode domačih turistov v Sloveniji v letih od 2002 do 2007 in število prihodov po letih**

Leto	**Indeksi s stalno osnovo 2002 = 100**	**Število turistov 1.000**
2002	100,0	**860**
2003	101,5	873
2004	97,9	842
2005	97,6	839
2006	100,8	867
2007	107,7	926

$$I_{j/0} = \frac{Y_j}{Y_0} \times 100 \text{ in iz tega: } \boldsymbol{Y_j = \frac{Y_0 \times I_{j/0}}{100}} \tag{3.22}$$

$$\textit{Število prihodov turistov v letu } 2003 = \frac{Y_{2002} \times I_{2003/2002}}{100} = \frac{860 \times 101{,}5}{100} = 873.$$

$$\textit{Število prihodov turistov v letu } 2004 = \frac{Y_{2002} \times I_{2004/2002}}{100} = \frac{860 \times 97{,}9}{100} = 842.$$

itd.

- **Dani so verižni indeksi**

  Za prihode domačih turistov v Sloveniji imamo dane verižne indekse in število prihodov turistov v letu 2002, ko jih je bilo 860 tisoč.

$$V_j = \frac{Y_j}{Y_{j-1}} \times 100 \text{ in iz tega: } \boldsymbol{Y_j = \frac{Y_{j-1} \times V_j}{100}} \tag{3.23}$$

Tabela 3.22: **Verižni indeksi za število prihodov domačih turistov v Sloveniji v letih od 2002 do 2007 in izračunano število prihodov turistov po letih**

Leto	**Verižni indeksi**	**Število turistov v 1.000**
2002	-	**860**
2003	101,5	873
2004	96,5	842
2005	99,8	840
2006	103,2	867
2007	106,9	927

$$\textit{Število prihodov turistov v letu } 2003 = \frac{Y_{2002} \times V_{2003}}{100} = \frac{860 \times 101{,}5}{100} = 873.$$

$$\textit{Število prihodov turistov v letu } 2004 = \frac{Y_{2003} \times V_{2004}}{100} = \frac{873 \times 96,5}{100} = 842.$$

$$\textit{Število prihodov turistov v letu } 2005 = \frac{Y_{2004} \times V_{2005}}{100} = \frac{842 \times 99,8}{100} = 840.$$

itd.

Lahko pa izračun poenostavimo in vrednosti množimo s koeficienti rasti. Tako odpade deljenje s 100:

$$\textit{Število prihodov turistov v letu } 2003 = Y_{2002} \times K_{2003} = 860 \times 1,015 = 873.$$

$$\textit{Število prihodov turistov v letu } 2004 = Y_{2003} \times K_{2004} = 873 \times 0,965 = 842.$$

$$\textit{Število prihodov turistov v letu } 2005 = Y_{2004} \times K_{2005} = 842 \times 0,998 = 840.$$

itd.

- **Dani so indeksi s stalno osnovo in vrednost člena, vendar se osnova indeksov in vrednost danega člena ne nanašata na isti časovni razmik**

Za prihode domačih turistov v Sloveniji imamo dane indekse s stalno osnovo 2004=100 in število turistov v letu 2002, ko jih je bilo 860 tisoč.

Tabela 3.23: **Indeksi s stalno osnovo za prihode domačih turistov 2004 = 100 v Sloveniji v letih od 2002 do 2007 in izračunano število turistov po letih**

Leto	Indeksi s stalno osnovo 2004 = 100	Prihodi turistov v 1.000
2002	102,1	**860**
2003	103,6	873
2004	**100,0**	842
2005	99,7	840
2006	102,9	867
2007	110,0	927

V tem primeru moramo indekse najprej preračunati na novo osnovo 2002 = 100, nato pa izračunati manjkajoče člene vrste. Lahko pa vrednosti členov izračunamo neposredno, kot je razvidno:

$$\textit{Število prihodov turistov v letu } 2003 = \frac{I_{2003/2004}}{I_{2002/2004}} \times Y_{2002} = \frac{103,6}{102,1} \times 860 = 873.$$

$$\textit{Število prihodov turistov v letu } 2004 = \frac{I_{2004/2004}}{I_{2002/2004}} \times Y_{2002} = \frac{100,0}{102,1} \times 860 = 842.$$

$$\textit{Število prihodov turistov v letu } 2005 = \frac{I_{2005/2004}}{I_{2002/2004}} \times Y_{2002} = \frac{99,7}{102,1} \times 860 = 840 \text{ itd.}$$

- **Dani so verižni indeksi in vrednost kateregakoli člena v časovni vrsti, vendar ne izhodiščnega**

Za prihode domačih turistov v Sloveniji imamo dane verižne indekse in število prihodov turistov v letu 2004, ko jih je bilo 842 tisoč.

Tabela 3.24: **Verižni indeksi za število prihodov domačih turistov v Sloveniji v letih od 2002 do 2007 in izračunano število prihodov turistov po letih**

Leto	Verižni indeksi	Število turistov v 1.000
2002	-	860
2003	101,5	873
2004	96,5	**842**
2005	99,8	840
2006	103,2	867
2007	106,9	927

V tem primeru moramo verižne indekse preračunati v indekse s stalno osnovo 2004 = 100 in s temi izračunati manjkajoče člene vrste oziroma vrednosti. Lahko pa te izračunamo neposredno, kot je razvidno:

- **pred osnovo:**

$$\textit{Število prihodov turistov v letu } 2003 = \frac{Y_{2004}}{V_{2004}} \times 100 = \frac{842}{96,5} \times 100 = 873.$$

$$\textit{Število prihodov turistov v letu } 2002 = \frac{Y_{2003}}{V_{2003}} \times 100 = \frac{873}{101,5} \times 100 = 860.$$

- **za osnovo:**

$$\textit{Število prihodov turistov v letu } 2005 = \frac{Y_{2004} \times V_{2005}}{100} = \frac{842 \times 99,8}{100} = 840.$$

$$\textit{Število prihodov turistov v letu } 2006 = \frac{Y_{2005} \times V_{2006}}{100} = \frac{840 \times 103,2}{100} = 867.$$

$$\textit{Število prihodov turistov v letu } 2007 = \frac{Y_{2006} \times V_{2007}}{100} = \frac{867 \times 106,9}{100} = 927.$$

Lahko pa izračun poenostavimo in verižne indekse spremenimo v koeficiente rasti. V tem primeru odpade množenje oziroma deljenje s 100:

$$\textit{Število prihodov turistov v letu } 2003 = \frac{Y_{2004}}{K_{2004}} = \frac{842}{0,965} = 873.$$

$$\textit{Število prihodov turistov v letu } 2005 = Y_{2004} \times K_{2005} = 842 \times 0,998 = 840.$$

3.5 RAZLIKA IN RELATIVNA RAZLIKA ZA RELATIVNA ŠTEVILA

Relativna števila so v glavnem neimenovana števila, kot so strukture, indeksi in stopnje, lahko pa so tudi imenovana, kot so statistični koeficienti. S primerjavo dveh neimenovanih relativnih števil ugotovljeno razliko izrazimo v točkah, s primerjavo koeficientov pa razliko izrazimo v isti merski enoti kot primerjani relativni števili.

V primeru 3.2 smo izračunali, da je bilo konec leta 2006 med moškimi, zaposlenimi pri pravnih osebah v gostinstvu 11,5 % z visoko izobrazbo, med zaposlenimi ženskami pa je bil ta odstotek 15,9. Izračunana razlika med strukturnima odstotkoma je:

$D_{M/Ž} = 15{,}9 - 11{,}5 = 4{,}4$ *odstotne točke*

Razlika je izražena v odstotnih točkah, kar pomeni, da je bil odstotek zaposlenih žensk z visoko izobrazbo za 4,4 odstotne točke višji od odstotka zaposlenih moških.

Za navedena strukturna odstotka lahko izračunamo tudi relativno razliko:

$$D_{M/Ž}\,\% = \frac{15{,}9 - 11{,}5}{11{,}5} \times 100 = 38{,}3\ \%$$

Odstotek zaposlenih žensk z visoko izobrazbo je bil za 38,3 % višji od odstotka zaposlenih moških z visoko izobrazbo.

S primerjavo dveh indeksov ugotovljeno razliko izrazimo z indeksnimi točkami, relativno razliko pa v odstotkih. V primeru 3.7 smo izračunali, da je indeks za mesec december 2007 v primerjavi z novembrom za prihode domačih turistov 108,7, za prihode turistov iz tujine pa 94,5. Izračunana razlika med indeksoma je:

$$D_{prihodi\ domačih\ turistov/\ prihodi\ turistov\ iz\ tujine} = 108{,}7 - 94{,}5 = 14{,}2\ \textit{indeksne točke}$$

Indeks prihodov domačih turistov je v opazovanem obdobju za 14,2 indeksne točke večji od indeksa prihodov turistov iz tujine.

Za navedena indeksa pa lahko izračunamo tudi relativno razliko:

$$D_{domači/iz\ tujine}\,\% = \frac{108{,}7 - 94{,}5}{94{,}5} \times 100 = 15\ \%$$

Indeks prihodov domačih turistov je v opazovanem obdobju za 15 % večji od indeksa prihodov turistov iz tujine.

V primeru 3.3 smo izračunali, da je povprečno število zaposlenih na prodajalno v trgovini z živili, s pijačami in tobakom 8,7, v trgovini z neživili pa 3,7.

Izračunana razlika med koeficientoma: $D_{trg.\ z\ živili,s\ pijačami\ in\ tobakom/trg.\ z\ neživili} = 8{,}7 - 3{,}7 = 5,$

kar pomeni, da je število zaposlenih na prodajalno v trgovini z živili, s pijačami in tobakom za 5 večje kot v trgovini z neživili.

In še relativna razlika: $D_{trg.\ z\ živili,\ s\ pijačami\ in\ tobakom/trg.\ z\ neživili}\,\% = \dfrac{8{,}7 - 3{,}7}{3{,}7} \times 100 = 135{,}1\ \%$

Število zaposlenih na prodajalno je v trgovini z živili, s pijačami in tobakom za 135,1 % večje kot v trgovini z neživili.

Relativna števila so najenostavnejši statistični parametri, kar ste lahko ugotovili, ko smo jih računali. Zelo pomembno pa je, da pri analizi določenih podatkov izberemo najustreznejši parameter, torej relativno število, ki nam v konkretnem primeru najbolje prikaže značilnosti preučevanega pojava. Seveda je izbira odvisna tudi od namena preučevanja. V primeru, ko želimo prikazati sestavo pojava, bomo izračunali ***strukturna števila****, ko analiziramo dinamiko pojava v nekem obdobju, bomo izračunali* ***indekse,*** *ko pa primerjamo raznovrstne podatke, ki so časovno in krajevno enako opredeljeni, računamo* ***statistične koeficiente****. Izračunane parametre dopolnimo z grafičnimi prikazi in tako še izboljšamo analizo.*

Za utrjevanje in samostojno preverjanje so v Zbirki vaj iz statistike pripravljene naloge od 2.1 do 2.21. Kar pogumno se jih lotite!

4 FREKVENČNE PORAZDELITVE

Urejanje podatkov je pri statističnem preučevanju zelo pomembno, saj se pogosto šele pri urejenih podatkih pokažejo določene značilnosti pojava. To poglavje obravnava način urejanja podatkov v frekvenčne porazdelitve. Primer, ki mu boste sledili, vam bo v pomoč, ko boste morali sami sestaviti frekvenčno porazdelitev za množico podatkov, zbranih za določeno številsko spremenljivko, in ugotavljati značilnosti preučevane populacije. Ugotovili boste, da se določene značilnosti pojava pokažejo šele v urejenih vrstah, poudarimo jih pa še z grafičnimi prikazi.

Frekvenčna porazdelitev je razvrstitev enot v skupine po vrednostih neke spremenljivke opazovanega pojava. Tako ločimo frekvenčne porazdelitve za opisne in številske spremenljivke.

Če ima spremenljivka veliko število vrednosti, združujemo v skupine več (sorodnih) vrednosti, saj s tem izboljšamo preglednost podatkov, predvsem pa lažje sklepamo o značilnostih preučevanega pojava.

Primer opisne porazdelitve prikazuje tabela 4.1.

Tabela 4.1: **Študenti na 1.000 prebivalcev in delež študentov med prebivalci, starimi od 19 do 26 let, po statističnih regijah v Sloveniji na začetku štud. leta 2006/07**

Regija	**Študenti na 1.000 prebivalcev**	**Odstotek študentov med preb., starimi 19–26 let**
Pomurska	44,4	40,1
Podravska	50,4	47,2
Koroška	57,3	49,7
Savinjska	58,3	50,2
Zasavska	53,6	46,9
Spodnjeposavska	54,5	48,7
Jugovzhodna Slovenija	58,6	49,4
Osrednjeslovenska	62,3	58,0
Gorenjska	60,1	52,9
Notranjsko-Kraška	60,7	53,5
Goriška	59,8	56,6
Obalno-Kraška	54,0	50,9
Slovenija	**57,1**	**51,6**

Vir: Slovenske regije v številkah, 2008

Primer številske porazdelitve prikazuje tabela 4.2.

Tabela 4.2: **Družine po številu otrok v Sloveniji ob popisu 2002**

Število otrok	**Število družin**
1	208.018
2	181.865
3	32.137
4	4.845
5 in več	1.438
Skupaj	428.303

Vir: Statistični letopis 2003

V tem primeru so enote razvrščene za posamezne vrednosti, le v zadnjem razredu je možnih več vrednosti. Ker pa imajo številske spremenljivke največkrat veliko število vrednosti, za te opredelimo razrede. Ti morajo biti opredeljeni tako, da lahko vsako enoto razvrstimo le v eno skupino. S tem je izpolnjen pogoj enolične opredelitve. Zato razrede razmejimo med seboj tako, da določimo spodnjo in zgornjo mejo razreda. Pri določanju mej razredov pa moramo upoštevati vrsto številske spremenljivke:

- ***Opredelitev razredov za vrednosti zvezne spremenljivke***

Opazovali so študente po oddaljenosti od doma do kraja študija. Najmanjša oddaljenost je bila nekaj čez 1 kilometer, največja pa nekaj manj kot 30 kilometrov. Upoštevajmo, da bi zajeli v razred vrednosti v razmiku 5 km.

Razrede lahko v tem primeru opredelimo na dva načina:

Oddaljenost v km	**ali**	**Oddaljenost v km**
od 0 do pod 5		nad 0 do 5
od 5 do pod 10		nad 5 do 10
od 10 do pod 15		nad 10 do 15
itd.		itd.

Pri prvem načinu opredelitve razredov dobimo **naslednje meje in sredine razredov**:

Oddaljenost v km	Širina razreda d_j	Spodnja meja $y_{j,min}$	Zgornja meja $y_{j,max}$	Sredina razreda y_j
od 0 do pod 5	5	0	5	2,5
od 5 do pod 10	5	5	10	7,5
od 10 do pod 15	5	10	15	12,5
itd				

Tudi če bi upoštevali drugi način opredelitve razredov, bi dobili enake vrednosti. Razlika med navedenima opredelitvama je le v tem, v kateri razred vključimo mejne vrednosti. Pri prvem načinu opredelitve je študent, ki je oddaljen od kraja študija 5 km, v drugem razredu, po drugem načinu pa v prvem. Torej po prvem načinu opredelitve vrednosti, ki so enake zgornji meji, niso vključene v ta razred, ampak v naslednji, po drugem načinu opredelitve pa vrednosti, ki so enake spodnji meji, niso vključene v ta razred, ampak v predhodni.

- ***Opredelitev razredov za vrednosti diskretne spremenljivke***

Opazovali so študente po številu točk, doseženih pri izpitu. Upoštevajmo, da je bilo najmanjše število točk 22 (od 80 možnih) in širina razreda 10.

Število točk	d_j	$y_{j,\min}$	$y_{j,\max}$	y_j
21 do 30	10	20,5	30,5	25,5
31 do 40	10	30,5	40,5	35,5
41 do 50	10	40,5	50,5	45,5
itd				

V primerih, da velja: $y_{j,\min} = y_{j-1,\max}$,

so meje razredov podane zvezno, sicer pa bomo zveznost zagotovili s ***popravkom za zveznost*** **(**Artenjak, 1997, 38):

$$\boldsymbol{\Delta y_j = \frac{y_{j,min} - y_{j-1,max}}{2}} \quad \textit{za } j=1,\ 2,\ 3, \ldots ..r, \qquad (4.1)$$

ki ga od spodnjih mej odštevamo, k zgornjim pa prištevamo.

V porazdelitvi 4.2 zadnji razred nima zgornje meje. V nekaterih porazdelitvah pa prvi razred nima spodnje meje. Take razrede imenujemo odprti razredi in jih opredelimo takrat, kadar so najmanjše ali največje vrednosti spremenljivke zelo razpršene. Zaradi odprtih razredov pogostokrat ne moremo izračunati vseh parametrov, npr. aritmetične sredine in variance, saj ne moremo določiti sredine razreda. Pomagamo si tako, da za sredino odprtega razreda vzamemo domnevno vrednost.

4.1 SESTAVLJANJE FREKVENČNE PORAZDELITVE

Z opazovanjem zberemo množico podatkov, ki so nepregledni in pogostokrat neustrezni za neposredno statistično analizo. Zato moramo podatke urediti. Najenostavneje jih uredimo tako, da enote razvrstimo v razrede. Osnovni namen razvrstitve enot v razrede je doseči boljšo preglednost enot. Sestavljanje frekvenčne porazdelitve pa je smiselno tudi zato, ker se pri tako urejenih podatkih pokažejo določene značilnosti pojava.

Za populacijo z N številom enot, ki imajo različne vrednosti spremenljivke, sestavimo frekvenčno porazdelitev po naslednjem postopku:

- **poiščemo najmanjšo in največjo vrednost** $y_{\min}$ **in** $y_{\max}$ **;**

- **v razmiku od** $y_{\min}$ **do** $y_{\max}$ **določimo najprimernejše število razredov**. Pri tem lahko uporabimo Sturgesovo pravilo, po katerem je približno število razredov:

$$\boldsymbol{r \approx 1 + 3{,}32 \log N} \qquad (4.2)$$

Število razredov je odvisno od števila enot v populaciji; več kot jih je, več razredov opredelimo.

Za srednje velike populacije naj bo to število med 8 in 16 (Košmelj, 1997, 115). Preveliko število razredov zmanjša preglednost, do izraza pa pridejo tudi slučajni vplivi pojava. Pri manjšem številu razredov je preglednost boljša, vendar se zabrišejo tudi osnovne značilnosti pojava;

- **izračunamo širino razredov:** $\boldsymbol{d_j = \frac{y_{max} - y_{min}}{r}}$ *(4.3)*

  Ta bo v vseh razredih enaka, če se vrednosti spremenljivke ne razlikujejo preveč. Če se vrednosti zelo razlikujejo, predvsem če se vrednosti manjšega števila enot zelo razlikujejo od vrednosti večine enot v populaciji, oblikujemo razrede z neenakimi širinami;

- **določimo meje razredov**, pri tem ni nujno, da se spodnja meja prvega razreda in zgornja meja zadnjega razreda ujemata z najmanjšo in največjo vrednostjo;

- **enote razvrstimo v razrede in jih preštejemo, tako dobimo frekvence** $\boldsymbol{f_j}$.

Sestavimo frekvenčno porazdelitev za naslednji primer:

Primer 4.1: Za 80 prodajaln trgovskega podjetja *Preskrba* imamo podatke o vrednosti prodaje v letu 2007. Podatki so v milijonih evrov, zaokroženi na eno decimalno mesto.

Opazovana spremenljivka je vrednost prodaje. Najmanjša vrednost je 10,2 milijona evrov, največja pa 23,8 milijona evrov. Razmik med najmanjšo in največjo vrednostjo ni velik, tudi število enot ni veliko, zato se odločimo za enako široke razrede. Frekvenčno porazdelitev sestavimo s črtkanjem.

20,2	12,9	18,8	18,5	13,6	16,6	17,1	15,5
12,7	23,8	13,8	15,7	16,3	14,4	13,1	17,5
11,0	20,6	12,5	15,6	18,5	12,2	13,0	15,8
14,3	11,7	17,9	16,2	12,8	17,7	18,3	13,5
14,9	22,9	17,3	14,2	15,3	15,1	14,2	16,3
12,8	15,7	15,1	19,1	15,7	17,1	14,8	13,5
11,4	15,8	11,8	10,8	13,4	17,3	14,2	15,6
11,8	19,1	19,8	16,9	15,5	19,7	11,3	14,1
10,2	16,3	17,7	16,2	12,4	14,1	18,8	12,2
12,4	10,5	21,8	16,2	15,3	19,3	16,2	14,2

S Sturgesovim pravilom izračunamo približno število razredov:

$r = 1 + 3{,}32 \log N = 1 + 3{,}32 \log 80 = 7{,}3 \approx 7$

in še širino razreda: $d_j = \frac{y_{\max} - y_{\min}}{r} = \frac{23{,}8 - 10{,}2}{7} = 1{,}94\ mio\ EUR,$ kar ustrezno zaokrožimo na 2.

Tabela 4.3: **Frekvenčna porazdelitev vrednosti prodaje za 80 prodajaln trgovskega podjetja *Preskrba* v letu 2007**

Vrednost prodaje v mio EUR	**Število prodajaln**	**Število prodajaln f_j**
od 10 do pod 12	~~////~~ ////	9
od 12 do pod 14	~~////~~ ~~////~~ ~~////~~ /	16
od 14 do pod 16	~~////~~ ~~////~~ ~~////~~ ~~////~~ ///	23
od 16 do pod 18	~~////~~ ~~////~~ ~~////~~ //	17
od 18 do pod 20	~~////~~ ~~////~~	10
od 20 do pod 22	///	3
od 22 do pod 24	//	2
Skupaj		**80**

Vir: Primer 4.1

Iz frekvenčne porazdelitve je razvidno, da nobena prodajalna ni prodala manj kot za 10 in nobena več kot za 24 milijonov evrov blaga in da je največ, kar 23 prodajaln prodalo blago v vrednosti od 14 do pod 16 milijonov evrov. Tako je v tretjem razredu največja frekvenca, torej največja gostitev pojava.

Z upoštevanjem manjšega števila razredov in večje širine razredov (npr. 5 milijonov) bi bila slika pojava preveč groba. Nasprotno pa bi z upoštevanjem manjše širine (npr. 1 milijon) bilo večje število razredov in s tem slabša preglednost. Pri določanju širine je smiselno izdelati več različic in se odločiti za tisto, pri kateri lastnosti pojava pridejo najbolj do izraza.

4.2 OPIS FREKVENČNE PORAZDELITVE

Pri analizi frekvenčne porazdelitve moramo poznati naslednje pojme:

- **frekvenca je število enot v posameznem razredu;**

- **relativna frekvenca izraža delež enot v posameznem razredu** in jo izračunamo kot razmerje med frekvenco in številom enot v populaciji. Po opredelitvi je enaka strukturnemu deležu:

$$f_j^{\circ} = \frac{f_j}{N} \tag{4.4}$$

  Vsota relativnih frekvenc je 1. Zapišemo: $\sum_{j=1}^{r} f_j^{\circ} = 1;$

- **kumulativa frekvenc**, ki jo dobimo s postopnim seštevanjem frekvenc, tako da je kumulativa v prvem razredu enaka frekvenci v tem razredu: $F_1 = f_1$

  Kumulativne frekvence v naslednjih razredih pa izračunamo po obrazcu:

$$F_j = F_{j-1} + f_j \tag{4.5}$$

  Tako izračunana kumulativna frekvenca za j-ti razred pove število enot z vrednostjo, enako največ zgornji meji tega razreda;

- **kumulativa relativnih frekvenc**, ki jo izračunamo podobno kot kumulativo frekvenc:

$$\boldsymbol{F_j^{\circ} = F_{j-1}^{\circ} + f_j^{\circ}} \qquad (4.6)$$

in pove delež enot z vrednostjo enako največ zgornji meji j-tega razreda.

Iz tabele 4.4 razložimo vse izračunane kazalce v tretjem razredu:

$f_3 = 23$	v 23 prodajalnah je bila vrednost prodaje od 14 do manj kot 16 milijonov evrov;
$f_3^{\circ} = 0{,}2875$	v 28,75 odstotkih prodajaln je bila vrednost prodaje od 14 do manj kot 16 milijonov evrov;
$F_3 = 48$	v 48 prodajalnah je bila vrednost prodaje do manj kot 16 milijonov evrov;
$F_3^{\circ} = 0{,}6000$	v 60,0 odstotkih prodajaln je bila vrednost prodaje do manj kot 16 milijonov evrov.

Tabela 4.4: **Frekvenčna porazdelitev vrednosti prodaje za 80 prodajaln trgovskega podjetja *Preskrba* v letu 2007 z izračunanimi relativnimi frekvencami ter kumulativo absolutnih in relativnih frekvenc**

Vrednost prodaje v mio EUR	**Število prodajaln f_j**	**Delež prodajaln – relativna frekvenca f_j°**	**Kumulativa frekvenc F_j**	**Kumulativa relativnih frekvenc – F_j°**
od 10 do pod 12	9	0,1125	9	0,1125
od 12 do pod 14	16	0,2000	25	0,3125
od 14 do pod 16	23	0,2875	48	0,6000
od 16 do pod 18	17	0,2125	65	0,8125
od 18 do pod 20	10	0,1250	75	0,9375
od 20 do pod 22	3	0,0375	78	0,9750
od 22 do pod 24	2	0,0250	80	1,0000
Skupaj	**80**	**1,0000**		

4.3 GRAFIČNO PRIKAZOVANJE FREKVENČNIH PORAZDELITEV

Frekvenčne porazdelitve grafično prikazujemo s stolpci ali z linijskim grafikonom, in sicer:

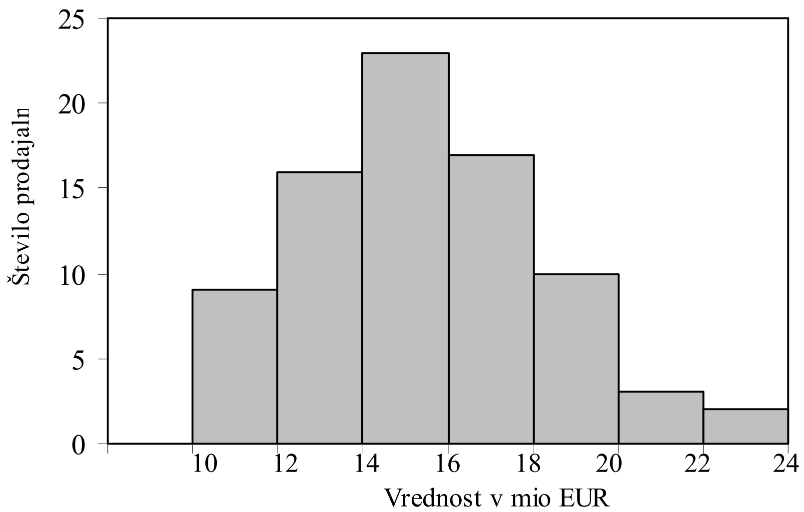
- prikaz s stolpci imenujemo ***histogram,*** pri njem je širina stolpca enaka širini razreda, višina pa številu enot v posameznem razredu;

- prikaz z linijskim grafikonom imenujemo ***poligon*** in ga narišemo tako, da nad sredine razredov nanašamo točke, ki jih povežemo z daljicami. Ker morata biti ploščini histograma in poligona enaki, moramo pri risanju poligona upoštevati še en razred pred prvim in en razred za zadnjim, ki imata frekvenco nič.

4.3.1 Grafično prikazovanje frekvenčnih porazdelitev z enakimi širinami razredov

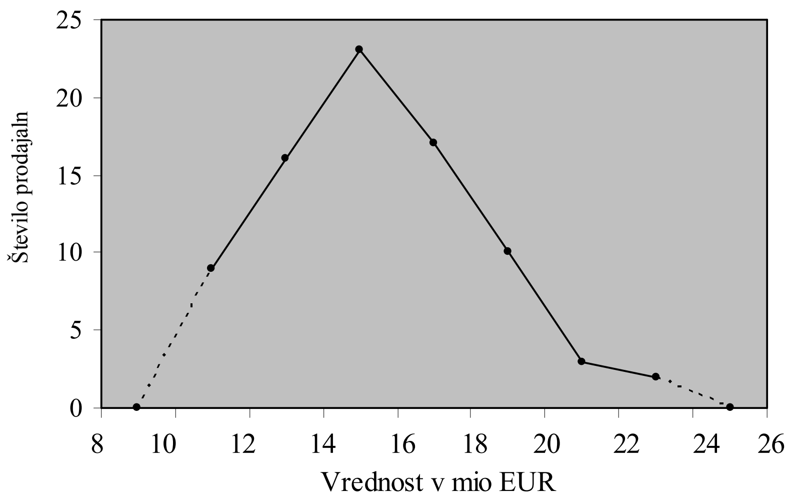
Predstavili bomo grafični prikaz vrednosti prodaje 80 prodajaln trgovskega podjetja *Preskrba* v letu 2007 s histogramom (slika 4.1) in poligonom (slika 4.2).

Iz obeh grafikonov lahko še bolj kot iz frekvenčne porazdelitve neposredno opazimo, na katerem intervalu so vrednosti spremenljivke pogostejše. Ugotovimo, da je največ prodajaln, kar 23, imelo od 14 do pod 16 milijonov evrov prodaje, torej je v tem intervalu gostitev opazovanega pojava največja.



Slika 4.1: **Histogram frekvenčne porazdelitve vrednosti prodaje za 80 prodajaln trgovskega podjetja *Preskrba* v letu 2007**

Vir: Tabela 4.4



Slika 4.2: **Poligon frekvenčne porazdelitve vrednosti prodaje za 80 prodajaln trgovskega podjetja *Preskrba* v letu 2007**

Vir: Tabela 4.4

4.3.2 Grafično prikazovanje frekvenčnih porazdelitev z neenakimi širinami razredov

Pri frekvenčnih porazdelitvah z neenakimi širinami razredov je višina stolpca enaka ***gostoti frekvence***, ki jo izračunamo po obrazcu:

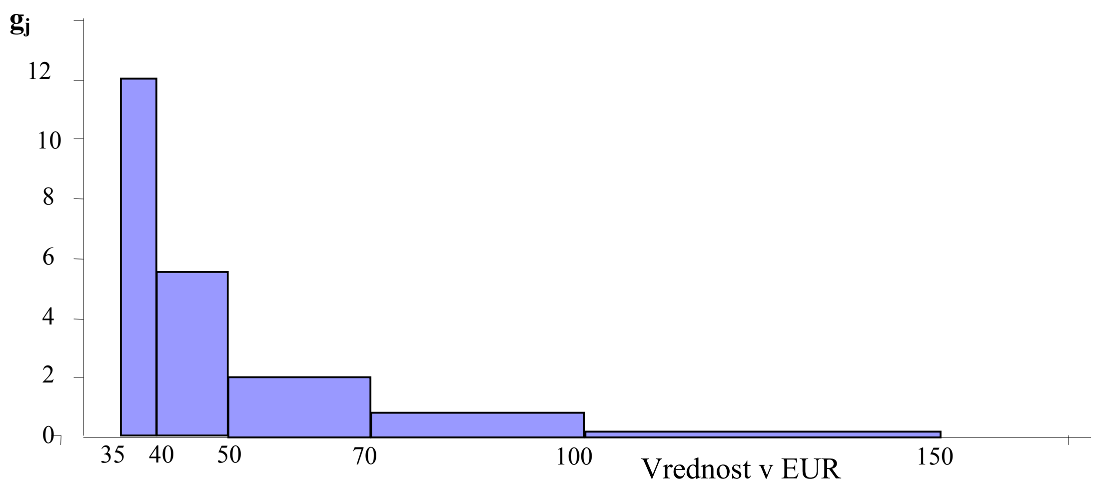
$$\boldsymbol{g_j} = \frac{\boldsymbol{f_j}}{\boldsymbol{d_j}} \quad \text{za } j = 1,2,3,....r \qquad (4.7)$$

Na abscisni osi moramo upoštevati različne širine razredov.

Tabela 4.5: **Frekvenčna porazdelitev vrednosti opravljenih 190 storitev podjetja *Prah***

Vrednost v EUR	**Število storitev** f_j	**Širina razreda** d_j	**Gostota razreda** g_j
nad 35 do 40	60	5	12,00
nad 40 do 50	55	10	5,50
nad 50 do 70	40	20	2,00
nad 70 do 100	25	30	0,83
nad 100 do 150	10	50	0,20
Skupaj	**190**		

$$g_1 = \frac{f_1}{d_1} = \frac{60}{5} = 12 \text{ in } g_2 = \frac{f_2}{d_2} = \frac{55}{10} = 5{,}5 \text{ itd.}$$

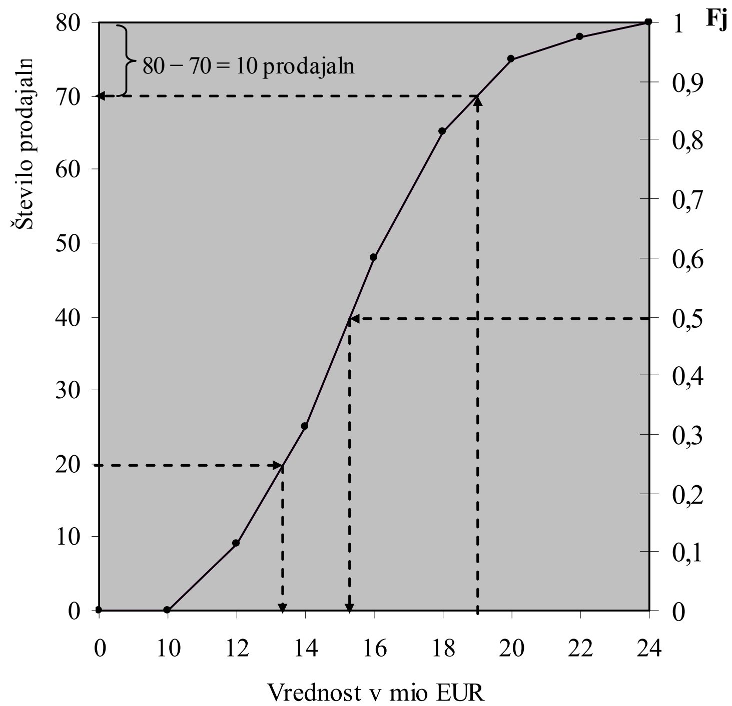


Slika 4.3: **Histogram porazdelitve vrednosti opravljenih 190 storitev podjetja *Prah***

4.3.3 Grafični prikaz kumulative frekvenc

Kumulative frekvenc prikazujemo z linijskim grafikonom, in sicer tako, da na abscisno os nanašamo vrednosti spremenljivke, to so širine razredov, na ordinatno os pa kumulative frekvenc. Kumulative frekvenc, ki nam povedo, koliko enot ima vrednosti do zgornje meje razreda, nanašamo kot točke nad zgornje meje razredov. Točke povežemo z daljicami. V istem grafikonu običajno prikazujemo tudi kumulative relativnih frekvenc. V tem primeru je skala za kumulativo relativnih frekvenc na desni strani grafikona, kot je razvidno tudi v sliki 4.4.

Iz grafikona lahko odčitamo, da je
- v 40 prodajalnah ali polovici prodajaln bila vrednost prodaje približno 15,3 milijona evrov,
- 20 prodajaln imelo do 13,5 milijona evrov vrednosti prodaje,
- več kot 19 milijonov evrov prodaje imelo le 10 prodajaln.



Slika 4.4: **Kumulativa frekvenčne porazdelitve vrednosti prodaje za 80 prodajaln trgovskega podjetja *Preskrba* v letu 2007**

Vir: Tabela 4.4

Frekvenčno porazdelitev je urejena vrsta največkrat za številske spremenljivke z velikim številom vrednosti. Te razvrstimo v razrede, ki jih oblikujemo tako, da bodo značilnosti preučevanih podatkov kar najbolj opazni.

Osnovna značilnost frekvenčne porazdelitve je frekvenca, to je število enot v razredu. S postopnim prištevanjem frekvenc pa dobimo kumulativo frekvenc, ki kaže število vrednosti do največ zgornje meje določenega razreda. Analizo frekvenčnih porazdelitev največkrat dopolnimo z grafičnim prikazom (poligon in histogram) ter prikazom kumulative frekvenc, z ocenami želenih vrednosti. Znanje o frekvenčnih porazdelitvah utrdite z nalogami od 3.1 do 3. 5 v Zbirki vaj iz statistike.

5 KVANTILI

Vrednosti, ki smo jih za dano spremenljivko ocenjevali v grafikonu kumulative frekvenc, lahko izračunamo. Prav tako lahko izračunamo položaj določene vrednosti, če poznamo podatek o tem, kolikšen delež vrednosti je manjših oziroma večjih od dane vrednosti. To nam omogočajo kvantili, ki jih obravnava to poglavje. Ugotovili boste, kako uporabno je znanje o kvantilih. Ste pomislili, kako drugačne bi bile izpitne ocene, če predavateljica ne bi vnaprej določila kriterija ocenjevanja, ampak šele, ko bi popravila vse naloge? Glede na rezultate bi se na primer odločila, da mora izpit opraviti 80 % študentov.

Pri statističnem preučevanju nas pogosto zanima položaj posamezne enote med ostalimi enotami v populaciji. Za opredeljevanje položaja posamezne enote uporabljamo naslednja parametra:

- ***rang R*** (določa, na katerem mestu v urejeni vrsti je posamezna enota) pove, koliko enot ima manjše in koliko večje vrednosti od izbrane enote. Ima vse vrednosti od 1 do N in zato lastnost diskretne spremenljivke;

- ***kvantilni rang P*** (položaj posamezne enote določa relativno) pove, koliko odstotkov enot ima manjše in koliko večje vrednosti od izbrane enote. Ima vse vrednosti na razmiku od 0 do 1 in lastnost zvezne spremenljivke.

  Vrednosti ranga so cela števila, kvantilnega ranga pa vse vrednosti med 0 in 1. Tako je povezava njunih vrednosti opredeljena z enačbama:

  $$\boldsymbol{R = N \times P + 0{,}5} \tag{5.1}$$

  $$\boldsymbol{P = \frac{R - 0{,}5}{N}} \tag{5.2}$$

  Vrednost 0,5 je popravek za zveznost. Tako velja:

- ***kvantilnemu rangu P = 0 ustreza rang R = 0,5 in***
- ***kvantilnemu rangu P = 1 ustreza rang R = N + 0,5***

Vrednosti, ki ustrezajo določenemu kvantilnemu rangu, so ***kvantili.*** Med njimi največkrat računamo:

- ***kvartile:***

  – ***prvi kvartil – Q_1 s kvantilnim rangom P = 0,25*** je vrednost, od katere ima 25 % enot manjše ali kvečjemu enake vrednost, 75 % enot pa večje vrednosti;

- ***drugi kvartil – Q_2 s kvantilnim rangom P = 0,50*** je vrednost, od katere ima 50 % enot manjše ali kvečjemu enake vrednosti, 50 % pa večje vrednosti;

- ***tretji kvartil – Q_3 s kvantilnim rangom P = 0,75*** je vrednost, od katere ima 75 % enot manjše ali kvečjemu enake vrednosti, 25 % pa večje vrednosti.

- ***decile:***

  - ***prvi decil – D_1 $P(D_1) = 0{,}10$;***
  - ***drugi decil – D_2 $P(D_2) = 0{,}20$;***
  - ***peti decil – D_5 $P(D_5) = 0{,}50$;***
  do
  - ***deveti decil – D_9 $P(D_9) = 0{,}90$*** je vrednost, od katere ima 90 % enot manjše ali kvečjemu enake vrednosti, 10 % pa večje.

- ***centile:***

  - ***prvi centil – C_{01} $P(C_{01}) = 0{,}01$;***
  - ***triintrideseti centil – C_{33} $P(C_{33}) = 0{,}33$*** je vrednost, od katere ima 33 % enot manjše ali kvečjemu enake vrednosti, 67 % enot pa večje vrednosti;
  - ***petdeseti centil – C_{50} $P(C_{50}) = 0{,}50$;***
  - ***devtindevetdeseti centil – C_{99} $P(C_{99}) = 0{,}99$.***

Iz navedenega ste že ugotovili, da je $Q_2 = D_5 = C_{50}$ in predstavlja vrednost spremenljivke, ki razdeli populacijo na dve polovici, tako da ima ena polovica enot manjše ali kvečjemu enake, druga polovica enot pa večje vrednosti. Ta vrednost je mediana ali sredinska vrednost populacije.

Kvantili so vrednosti, ki ustrezajo določenemu kvantilnemu rangu. Glede na to lahko rešujemo dve vrsti problemov:

- **za znano vrednost oziroma kvantil bomo priredili ustrezni kvantilni rang,**

- **k znanemu kvantilnemu rangu bomo priredili ustrezajočo vrednost oziroma kvantil.**

5.1 KVANTILI IN KVANTILNI RANGI IZ RANŽIRNE VRSTE

Kadar opazujemo manjše število enot po vrednostih določene spremenljivke, zbrane vrednosti uredimo po velikosti od najmanjše do največje v ranžirno vrsto. V ranžirni vrsti ima vsaka vrednost ustrezajoči rang, ki določa njen položaj.

Vzemimo primer, da je 14 študentov pisalo test iz angleškega jezika in doseglo naslednje število točk:

24 31 29 57 53 45 47 50 38 33 35 27 43 48 od 60 možnih.

Ranžirno vrsto lahko prikažemo grafično, tako da na abscisno os nanašamo vrednosti spremenljivke, ordinatna os pa je skala za range. Na desni strani dodamo še skalo za kvantilne range.

1. Določanje kvantilnega ranga P danemu kvantilu y

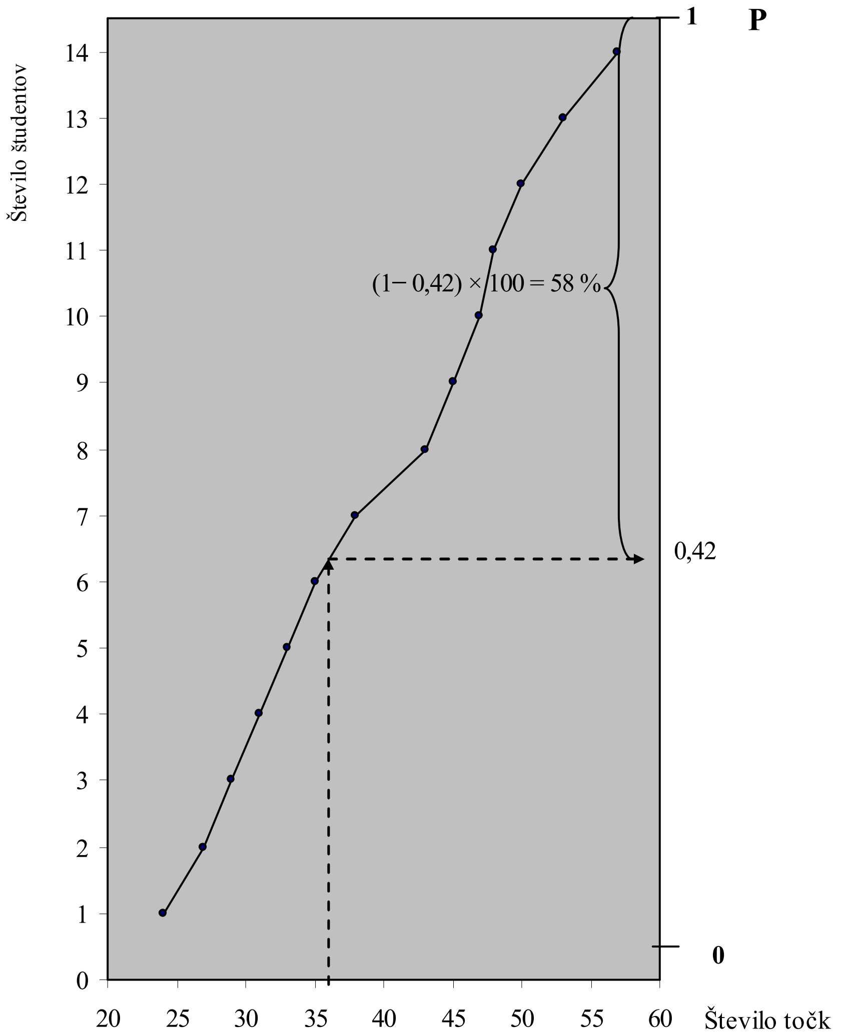
Ugotoviti želimo odstotek dijakov, ki so test opravili, če je bilo za to potrebnih 36 točk.

Podatke uredimo v ranžirno vrsto.

Ranžirna vrsta študentov po rezultatih testa

y	24	27	29	31	33	35	38	43	45	47	48	50	53	57
R	1	2	3	4	5	6	7	8	9	10	11	12	13	14

Slika 5.1 prikazuje ranžirno vrsto z grafičnima ocenama.



Slika 5.1: Ranžirna vrsta rezultatov pisnega preverjanja iz angleškega jezika

Kvantilni rang za dani kvantil izračunamo po naslednjem postopku:

- **Kvantil *y* je znan:**

  $\boldsymbol{y} = \mathbf{36}$ **točk**

- V ranžirni vrsti poiščemo položaj dane vrednosti tako, da ustreza neenačbi:

  $\boldsymbol{y_0 < y < y_1}$ **in vrednostim členov v neenačbi priredimo ustrezajoče range** $\boldsymbol{R_0 < R < R_1}$

  $y_0 = 35 < y = 36 < y_1 = 38$
  $R_0 = 6 < R < R_1 = 7$

- **Izračunamo rang *R* za vrednost** y po obrazcu:

$$\boldsymbol{R = R_0 + \frac{y - y_0}{y_1 - y_0} \times (R_1 - R_0)} \qquad (5.3)$$

$$R = 6 + \frac{36-35}{38-35} \times (7-6) = 6{,}33$$

- Na osnovi povezave med rangom in kvantilnim rangom **izračunamo kvantilni rang *P:***

$$\boldsymbol{P = \frac{R - 0{,}5}{N}} = \frac{6{,}33 - 0{,}5}{14} = 0{,}416$$

- **Razložimo dobljeni rezultat**:

  $(1 - 0{,}416) \times 100 = 58{,}4\ \%$: 41,6 % študentov je zbralo manj ali kvečjemu 36 točk, 58,4 % pa več. Torej je test opravilo 58,4 % študentov.

2. Določanje kvantila y danemu kvantilnemu rangu P

Ugotoviti želimo, koliko točk bi zadoščalo za pozitivno oceno, če mora test opraviti 80 % študentov.

Vrednost, ki ustreza danemu kvantilnemu rangu, izračunamo po naslednjem postopku:

- **Kvantilni rang je znan:**

  $\boldsymbol{P = 0{,}20}$
  Iskana vrednost je **drugi decil – D_2.**

- Na osnovi povezave med rangom in kvantilnim rangom **za iskano vrednost izračunamo rang *R*:**
  $\boldsymbol{R = N \times P + 0{,}5}$
  $R = 14 \times 0{,}20 + 0{,}5 = 3{,}3$

- Izračunani rang ustreza neenačbi:

$\boldsymbol{R_0 \leq R < R_1}$

$R_0 = 3 < R = 3{,}3 < R_1 = 4$

Rangom v neenačbi priredimo njihove vrednosti tako, da ustrezajo neenačbi:

$\boldsymbol{y_0 \leq y < y_1}$

$y_0 = 29 < y < y_1 = 31$

- **Izračunamo vrednost kvantila** $\boldsymbol{y}$ na osnovi obrazca:

$$\boldsymbol{y = y_0 + \frac{R - R_0}{R_1 - R_0} \times (y_1 - y_0)} \qquad (5.4)$$

$$y = 29 + \frac{3{,}3 - 3}{4 - 3} \times (31 - 29) = 29{,}6 \text{ točke}$$

- **Razložimo rezultat**:

Pozitivno ocenjeni bi bili vsi študentje, ki bi zbrali 29,6 ali, če zaokrožimo, 30 in več točk.

5.2 KVANTILI IN KVANTILNI RANGI IZ FREKVENČNE PORAZDELITVE

Iz porazdelitve prav tako kot iz ranžirne vrste določamo znanemu kvantilu ustrezajoči rang in znanemu kvantilnemu rangu ustrezajoči kvantil. Ranžirno vrsto nam nadomestijo razredi, range pa kumulativa frekvenc.

1. Določanje kvantilnega ranga P znanemu kvantilu y

Uporabili bomo frekvenčno porazdelitev prodajaln po vrednosti prodaje v letu 1999.

Tabela 5.1: **Frekvenčna porazdelitev vrednosti prodaje za 80 prodajaln trgovskega podjetja *Preskrba* v letu 2007**

Vrednost prodaje v mio EUR	**Število prodajaln** f_j	**Kumulativa frekvenc** F_j
od 10 do pod 12	9	9
od 12 do pod 14	16	25
od 14 do pod 16	23	48
od 16 do pod 18	17	65
od 18 do pod 20	10	75
od 20 do pod 22	3	78
od 22 do pod 24	2	80
Skupaj	**80**	

Ugotoviti želimo odstotek prodajaln, katerih vrednost prodaje je bila večja od 19 milijonov evrov.

Kvantilni rang za dani kvantil izračunamo po naslednjem postopku:

- **Kvantil y je znan:**

  y **= 19 mio EUR**

- V frekvenčni porazdelitvi določimo razred, v katerem je dani kvantil, **to je kvantilni razred**. V našem primeru je to peti razred (j = 5). **Dana vrednost namreč ustreza neenačbi**:

  $\boldsymbol{y_{j,min} < y < y_{j,max}}$
  $y_{5,\min} = 18 < y = 19 < y_{5,\max} = 20$

  Vrednostim členov v neenačbi priredimo ustrezajoče range, ki jih v frekvenčni porazdelitvi nadomestijo kumulative frekvenc:

  $\boldsymbol{F_{j-1} < R \leq F_j}$
  $F_1 = 65 < R < F_2 = 75$

- Za dano vrednost ***y*** **izračunamo rang *R*** po obrazcu:

$$\boldsymbol{R = F_{j-1} + f_j \times \frac{y - y_{j,min}}{d_j} =} \tag{5.5}$$

$$= 65 + 10 \times \frac{19 - 18}{2} = 70$$

- Na osnovi povezave med rangom in kvantilnim rangom **izračunamo kvantilni rang *P*:**

$$\boldsymbol{P = \frac{R - 0,5}{N}}$$

$$P_{y=19} = \frac{70 - 0,5}{80} = 0,869$$

- **Razložimo dobljeni rezultat**:

  $(1 - 0,869) \times 100 = 13,1$ %: 86,9 % prodajaln je imelo manjšo, 13,1 % prodajaln pa večjo vrednost prodaje kot 19 milijonov evrov.

*2. **Določanje kvantila y_p danemu kvantilnemu rangu P***

Ugotoviti želimo, kolikšno vrednost prodaje je imela prodajalna, od katere je le ena četrtina prodajaln imela manjšo vrednost.

Željeno vrednost, torej kvantil, ki ustreza danemu kvantilnemu rangu, izračunamo po naslednjem postopku:

- **Kvantilni rang *P* je znan:**

  $\boldsymbol{P = 0,25}$

Iskana vrednost je Q_1 (prvi kvartil)

- Na osnovi povezave med rangom in kvantilnim rangom za iskano vrednost **izračunamo rang R**:

$$\boldsymbol{R = N \times P + 0,5}$$
$$R = 80 \times 0,25 + 0,5 = 20,5$$

- Na osnovi izračunanega ranga **določimo razred, v katerem je iskana vrednost – kvantilni razred**.

Določimo ga tako, da v kumulativni vrsti ugotovimo položaj vrednosti z izračunanim rangom, ki ustreza neenačbi: $\boldsymbol{F_{j-1} < R \leq F_j}$

$F_1 = 9 < R = 20,5 < F_2 = 25$ Kvantilni razred je drugi razred ($j = 2$), v katerem so vse vrednosti z rangom nad 9 do 25, kar zapišemo v obliki neenačbe:

$$\boldsymbol{y_{j,min} < y < y_{j,max}}$$
$$y_{2,\min} = 12 < y_{P=0,25} = Q_1 < y_{2,\max} = 14$$

- **Izračunamo kvantil po obrazcu**:

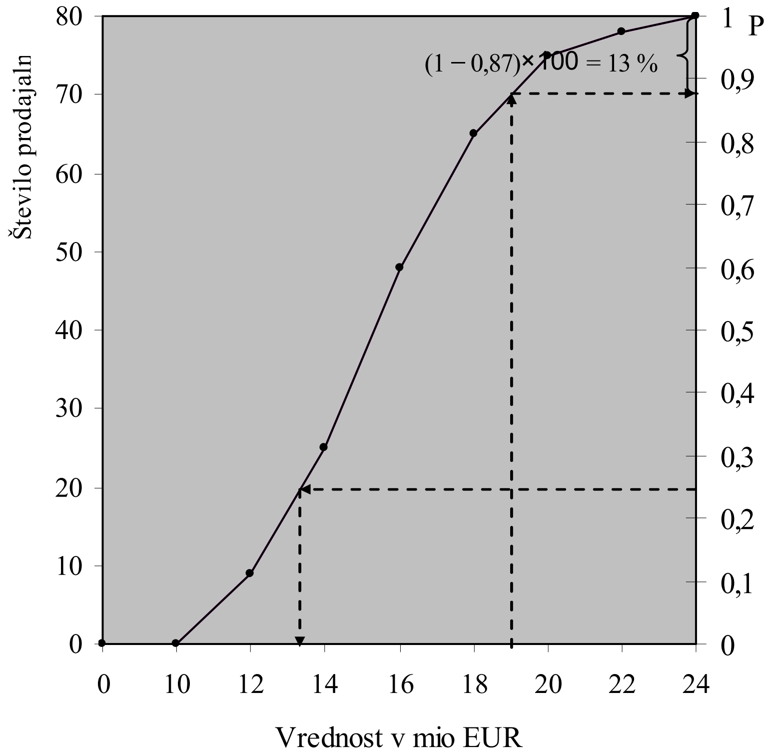
$$\boldsymbol{y = y_{j,min} + d_j \times \frac{R - F_{j-1}}{f_j}} \qquad (5.6)$$

$$y_{P=0,25} = Q_1 = 12 + 2 \times \frac{20,5 - 9}{16} = 13,4 \text{ mio EUR}$$

- **Razložimo rezultat**:

V prodajalni, od katere je le 25 % prodajaln imelo manjšo vrednost prodaje, je bila vrednost prodaje 13,4 milijona evrov.

Oba izračuna primerjajmo še z grafično oceno. Grafično prikažemo kumulativo frekvenc tako, da na desni strani dodamo še drugo os za kvantilni rang.



Slika 5.2: **Kumulativa frekvenčne porazdelitve vrednosti prodaje za 80 prodajaln trgovskega podjetja *Preskrba* v letu 2007**

Z usvojenim znanjem tega poglavja lahko rešujete dve vrsti problemov:

1. *znani vrednosti oziroma kvantilu (y) priredite ustrezni kvantilni rang (P) in*
2. *k znanemu kvantilnemu rangu (P) priredite ustrezajočo vrednost (y).*

Z reševanjem nalog od 4.1 do 4.9 v Zbirki vaj iz statistike se boste prepričali, da je to znanje zelo uporabno v praksi.

6 SREDNJE VREDNOSTI

Namen statistične analize je ugotoviti značilnosti preučevanega pojava. Te značilnosti pogosto ugotovimo z izračunanimi srednjimi vrednostmi. Spoznali in računali boste vse srednje vrednosti in, kar je najpomembnejše, pri analizi določenega pojava oziroma danih podatkih se boste znali odločiti, katera je najprimernejša pri analizi določenega pojava.

Pri opisovanju številskih spremenljivk smo videli, da so njihove vrednosti pri posameznih opazovanih enotah različne; pravimo, da vrednosti spremenljivke variirajo od enote do enote. Pri opisovanju značilnosti preučevane populacije praviloma ne navajamo vseh enot in njihovih vrednosti, ampak določimo ali izračunamo le nekaj vrednosti, ki pa morajo biti karseda dober predstavnik vseh opazovanih enot. Take vrednosti imenujemo srednje vrednosti. Tako kot smo pri frekvenčnih porazdelitvah sredino razreda šteli za reprezentančno vrednost vseh enot tega razreda, tako je srednja vrednost statistični parameter in kot taka reprezentančna vrednost vseh enot preučevane populacije.

V statistični analizi ločimo več srednjih vrednosti. Vsaka od njih na svoj način odkriva tisto, kar je najbolj značilno za opazovane enote. Največkrat uporabljane so:

- ***mediana***
- ***modus***
- ***aritmetična sredina***
- ***harmonična sredina***
- ***geometrijska sredina***

6.1 MEDIANA

Mediana ali središčna vrednost je srednja vrednost, ki jo opredelimo kot vrednost, ki ustreza kvantilnemu rangu $P = 0{,}50$ in je enaka drugemu kvartilu oziroma petemu decilu. To pomeni, da ima polovica enot manjše, polovica enot preučevane populacije pa večje vrednosti od vrednosti, ki je enaka mediani. Prednost mediane je v tem, da jo lahko določimo tudi v primerih, ko poznamo vrednosti le okoli sredinske vrednosti v ranžirni vrsti in ko imamo v frekvenčni porazdelitvi prvi ali zadnji razred odprt. Njena pomanjkljivost pa je v tem, da je neobčutljiva za spremembe najmanjših in največjih vrednosti v ranžirni vrsti.

Ker je mediana kvantil s kvantilnim rangom $P = 0{,}50$, postopek računanja že poznamo.

6.1.1 Mediana iz ranžirne vrste

Za skupino študentov, ki so pisali test iz angleškega jezika, izračunajmo tisto število točk, od katerega je polovica študentov zbrala manj, polovica pa več točk.

Ranžirna vrsta študentov po številu točk pisnega preverjanja iz angleškega jezika

y	24	27	29	31	33	35	38	43	45	47	48	50	53	57
R	1	2	3	4	5	6	7	8	9	10	11	12	13	14

- Znan je kvantilni rang:

  P = 0,50

- Na osnovi povazave med rangom in kvantilnim rangom izračunamo **rang *R***:

  $R = N \times P + 0,5 = 14 \times 0,50 + 0,5 = 7,5$

- Izračunani rang ustreza neenačbi:
  $R_0 \leq R < R_1 \quad \rightarrow \quad R_0 = 7 < R = 7,5 < R_1 = 8$

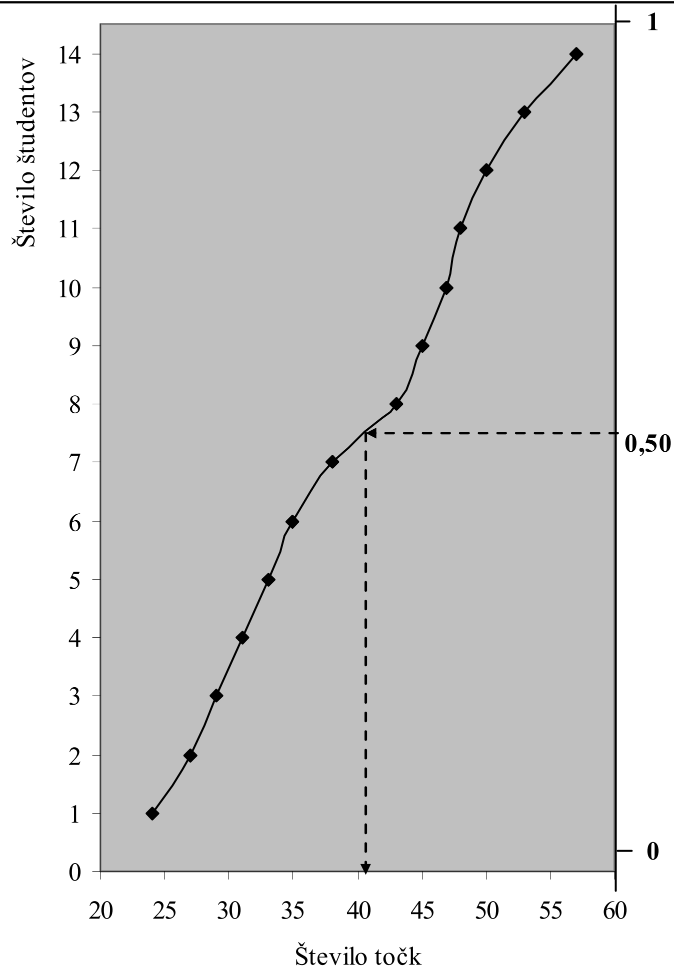
- Rangom v nenačbi priredimo vrednosti členov iz ranžirne vrste tako, da ustrezajo neenačbi:
  $y_0 \leq y < y_1 \quad \rightarrow \quad y_0 = 38 < y < y_1 = 43$

- Vrednost izračunamo po obrazcu:

$$\boldsymbol{y_{P=0,50} = Me = y_0 + \frac{R - R_0}{R_1 - R_0} \times (y_1 - y_0)} \qquad (6.1)$$

$$y_{P=0,50} = Me = 38 + \frac{7,5-7}{8-7} \times (43-38) = 40,5$$

Polovica študentov je pri pisnem preverjanju zbrala manj, polovica pa več kot 40,5 točke. V grafikonu ranžirne vrste pa lahko mediano ocenimo tudi grafično (slika 6.1).



Slika 6.1: **Ranžirna vrsta rezultatov pisnega preverjanja iz angleškega jezika z oceno mediane**

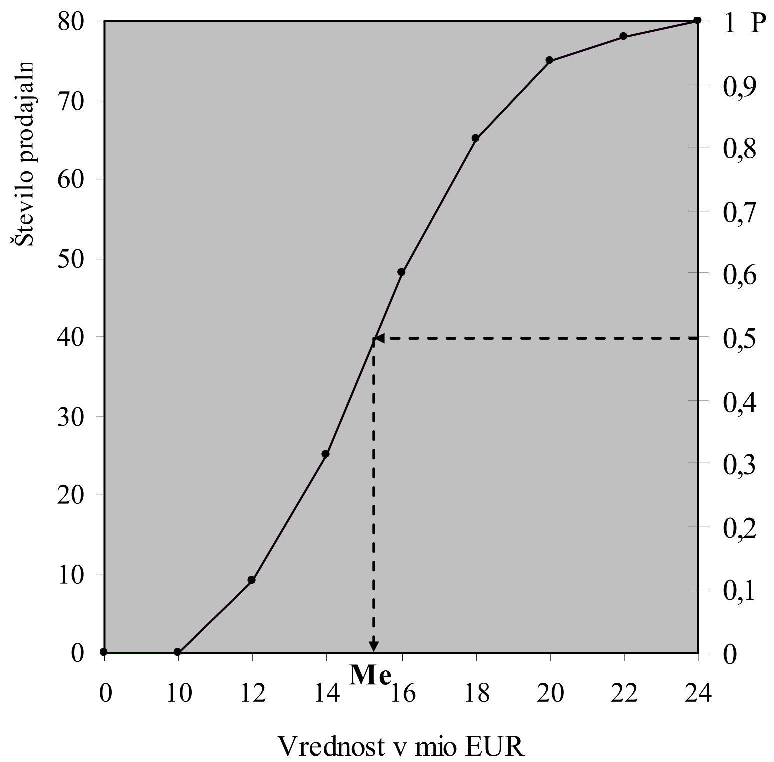
6.1.2 Mediana iz frekvenčne porazdelitve

Iz frekvenčne porazdelitve prodajaln po vrednosti prodaje izračunajmo tisto vrednost prodaje v milijonih evrov, od katere je polovica prodajaln imela manjšo, polovica pa večjo vrednost.

Tabela 6.1: **Frekvenčna porazdelitev vrednosti prodaje za 80 prodajaln trgovskega podjetja *Preskrba* v letu 2007**

Vrednost prodaje v mio EUR	**Število prodajaln f_j**	**Kumulativa frekvenc F_j**
od 10 do pod 12	9	9
od 12 do pod 14	16	25
od 14 do pod 16	23	48
od 16 do pod 18	17	65
od 18 do pod 20	10	75
od 20 do pod 22	3	78
od 22 do pod 24	2	80
Skupaj	**80**	

V grafičnem prikazu kumulative frekvenčne porazdelitve lahko mediano ocenimo grafično (slika 6.2).



Slika 6.2: **Kumulativa frekvenčne porazdelitve vrednosti prodaje za 80 prodajaln trgovskega podjetja *Preskrba* v letu 2007 z oceno mediane**

Vir: Tabela 6.1

- ***P = 0,50***
- Izračunamo rang: $R = N \times P + 0,5 = 80 \times 0,50 + 0,5 = 40,5$
- Določimo medialni razred: $F_{j-1} < R \leq F_j \quad \rightarrow \quad F_2 = 25 < R = 40,5 < F_3 = 48$
  Tretji razred (*j = 3*) je medialni razred.
- Mediana je v tretjem razredu: $y_{3,\min} = 14 < Me < y_{3,\max} = 16$
- Izračunamo jo po obrazcu:
  $$\boldsymbol{Me = y_{j,min} + d_j \times \frac{R - F_{j-1}}{f_j}} = 14 + 2 \times \frac{40,5 - 25}{23} = 15,35 \text{ mio EUR}$$

Polovica prodajaln je imela manjšo, polovica pa večjo vrednost prodaje od 15,35 milijona evrov.

6.2 MODUS

Modus je najpogostejša vrednost, torej vrednost, ki se najpogosteje pojavlja med opazovanimi vrednostmi.

6.2.1 Modus iz posameznih vrednosti

Modus iz posameznih vrednosti določimo tako, da pogledamo, katera vrednost se največkrat pojavlja med opazovanimi vrednostmi.

Primer 6.1: Določimo modus v primeru opazovanja 16 gostinskih obratov podjetja *Turist* po številu zaposlenih:

5 7 8 4 5 7 7 6 3 4 9 10 11 4 7 8

Ugotovimo, da je najpogostejše število zaposlenih 7, saj imajo to število zaposlenih kar štirje od 16 obratov.

6.2.2 Modus iz frekvenčne porazdelitve

Modus je srednja vrednost, ki kaže na gostitev pojava. Zato pri izračunavanju izhajamo iz razreda, v katerem je frekvenca največja. Prav zato lahko modus izračunamo le iz frekvenčnih porazdelitev z enakimi širinami razredov. Če širine razredov niso enake, po frekvencah ni mogoče presoditi, v katerem razredu je gostitev vrednosti največja.

Izračunajmo modus za frekvenčno porazdelitev vrednosti prodaje za 80 prodajaln:

- **določimo modalni razred,** to je razred z največjo frekvenco, v frekvenčni porazdelitvi vrednosti prodaje za 80 prodajaln trgovskega podjetja *Preskrba* je to tretji razred, v katerem je $f_3 = 23$;
- **modus izračunamo po obrazcu:** $$\boldsymbol{Mo = y_{j,min} + d_j \times \frac{f_j - f_{j-1}}{2f_j - f_{j-1} - f_{j+1}}}, \qquad (6.3)$$

kjer je**:**

$y_{j,\min}$ – spodnja meja modalnega razreda

d_j – širina razreda

f_j – frekvenca modalnega razreda

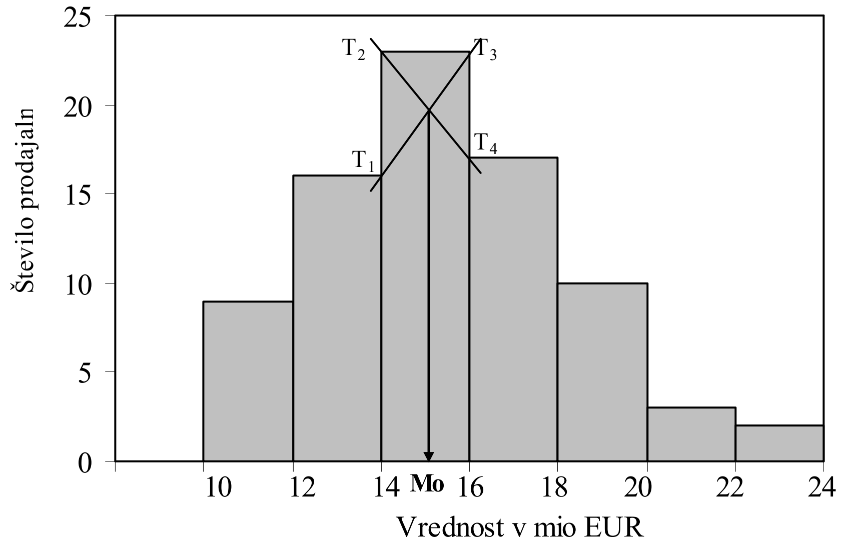
f_{j-1} – frekvenca pred modalnim razredom

f_{j+1} – frekvenca za modalnim razredom

$$Mo = 14 + 2 \times \frac{23 - 16}{2 \times 23 - 16 - 17} = 15{,}1 \text{ mio EUR}$$

Najpogostejša vrednost prodaje v teh 80 prodajalnah je bila 15,1 milijona evrov.

Modus grafično ocenimo v histogramu frekvenčne porazdelitve. Modalni razred je razred od 14 do pod 16 milijonov evrov. Skozi točki T_1 in T_3 ter točki T_2 in T_4 vrišemo premici p_1 in p_2 ter njuno presečišče preslikamo na abscisno os (slika 6.3).



Slika 6.3: **Histogram frekvenčne porazdelitve vrednosti prodaje za 80 prodajaln trgovskega podjetja *Preskrba* v letu 2007 z oceno modusa**

6.3 ARITMETIČNA SREDINA

Aritmetična sredina je najbolj znana in uporabljana srednja vrednost. V praksi jo poznamo pod izrazom povprečje. To je tista vrednost, ki jo izračunamo, če vsoto posameznih vrednosti številske spremenljivke y_j delimo s številom enot N. Če za aritmetično sredino uporabimo simbol M, za vsoto vrednosti spremenljivke y_j pa Y, zapišemo:

$$\boldsymbol{M = \frac{Y}{N} \qquad M = \frac{\sum_{i=1}^{N} y_i}{N}} \qquad (6.4)$$

Tako izračunano aritmetično sredino imenujemo navadna aritmetična sredina. Zanjo velja:

$$\boldsymbol{M \times N = Y} \qquad (6.5)$$

Iz zapisanega sklepamo: če bi bile vrednosti spremenljivke pri vseh opazovanih enotah enake, bi bile njihove vrednosti enake aritmetični sredini.

6.3.1 Aritmetična sredina iz posameznih vrednosti

Za izračun aritmetične sredine iz posameznih vrednosti uporabimo naslednji obrazec:

$$\bar{y} = M = \frac{y_1 + y_2 + \dots\dots + y_N}{N} = \frac{1}{N}\left(y_1 + y_2 + \dots\dots + y_N\right) = \frac{1}{N}\sum_{i=1}^{N} y_i \qquad (6.6)$$

Tako izračunano aritmetično sredino imenujemo ***navadna aritmetična sredina***.

Izračunajmo aritmetično sredino za število točk 14 študentov, ki so pisali test iz angleškega jezika:

$$\bar{y} = M = \frac{y_1 + y_2 + \dots + y_N}{N} = \frac{1}{14}\left(24 + 31 + 29 + \dots + 48\right) = \frac{560}{14} = 40 \text{ točk}$$

Skupno število točk, ki so jih zbrali vsi študenti, smo delili s številom študentov.

Ob predpostavki, da bi vsi študentje pri pisnem preverjanju zbrali enako število točk oziroma če bi bilo število točk enakomerno porazdeljeno na vse študente, bi bilo to število enako aritmetični sredini.

6.3.2 Aritmetična sredina iz frekvenčne porazdelitve

V frekvenčni porazdelitvi so enote razvrščene v razrede. Za vsak razred poznamo število enot oziroma frekvenco, ne poznamo pa posameznih vrednosti. Vemo pa, da so njihove vrednosti v razmiku med spodnjo in zgornjo mejo tega razreda. Tako ne moremo uporabiti obrazca, ki smo ga zapisali za aritmetično sredino iz posameznih vrednosti. Namesto teh pri izračunu uporabimo sredino razreda kot reprezentančno vrednost vseh enot razreda.

Aritmetično sredino izračunamo po naslednjem postopku:

- v vsakem razredu izračunamo **sredino razreda:** $y_j = \frac{y_{j,min} + y_{j.max}}{2}$;
- **sredine razredov pomnožimo s frekvencami razredov in dobimo produkte** $f_j y_j$;
- **produkte** $f_j y_j$ **seštejemo in delimo s številom enot N:** $M = \frac{1}{N}\sum_{j=1}^{k} y_j f_j$, (6.7)

kjer je:

k – število razredov
y_j – sredina v j-tem razredu
f_j – frekvenca v j-tem razredu

$$M = \frac{1}{80}\left[(9\times 11) + (16\times 13) + (23\times 15) + (17\times 17) + (10\times 19) + (3\times 21) + (2\times 23)\right] =$$

$$= \frac{1.240}{80} = 15{,}5 \text{ mio EUR}$$

Povprečna prodaja 80 prodajaln trgovskega podjetja *Preskrba* v letu 2007 je bila 15,5 milijona evrov.

Ali:

Tabela 6.2: **Frekvenčna porazdelitev vrednosti prodaje za 80 prodajaln trgovskega podjetja *Preskrba* v letu 2007 z izračuni za aritmetično sredino**

Vrednost prodaje v mio EUR	**Število prodajaln** f_j	y_j	$f_j y_j$
od 10 do pod 12	9	11	99
od 12 do pod 14	16	13	208
od 14 do pod 16	23	15	345
od 16 do pod 18	17	17	289
od 18 do pod 20	10	19	190
od 20 do pod 22	3	21	63
od 22 do pod 24	2	23	46
Skupaj	**80**		**1.240**

Aritmetično sredino, izračunano iz frekvenčne porazdelitve, ***imenujemo tehtana (ponderirana) aritmetična sredina***. Frekvence v posameznem razredu, ki jih imenujemo uteži ali ponderji, množimo s sredinami razredov, ki predstavljajo vrednosti. Te vrednosti moramo upoštevati tolikokrat, kolikor je enot v razredu. Izračun aritmetične sredine na osnovi zmnožkov sredin in frekvenc v tabeli 6.2 (4. stolpec) končamo kot pri prejšnjem načinu.

Tako izračunana aritmetična sredina se nekoliko razlikuje od tiste, ki jo izračunamo iz posameznih vrednosti in je le približna vrednost za aritmetično sredino, ki bi jo dobili z izračunom iz posameznih vrednosti.

Iz frekvenčne porazdelitve izračunana aritmetična sredina je 15,5 milijona evrov, kar je približna vrednost, a se skoraj ne razlikuje od aritmetične sredine, izračunane iz posameznih vrednosti, ki je 15,51 milijona evrov. Aritmetična sredina, izračunana iz frekvenčne porazdelitve, sodi v osrednji razred, ki vključuje vrednosti od 14 do pod 16 milijonov evrov. V tem primeru je aritmetična sredina kar dober predstavnik vseh opazovanih vrednosti, saj se vrednost prodaje večine prodajaln ne razlikuje dosti od 15,5 milijona evrov.

$$M = \frac{1}{N}(y_1 + y_2 + ... + y_N) =$$

$$= \frac{1}{80}(20,2 + 12,9 + 18,8 + ... + 19,3 + 16,2 + 14,2) = 15,51 \text{ mio EUR}$$

Primer 6.2: Izračunajmo tehtano aritmetično sredino tudi v naslednjem primeru.

V prodajalni s sadjem in zelenjavo so določenega dne prodali:
34 kilogramov pomaranč po ceni 2,5 evra za kilogram,
65 kilogramov pomaranč po ceni 2,7 evra za kilogram,
32 kilogramov pomaranč po ceni 3,2 evra za kilogram.

Izračunajmo povprečno ceno prodanih pomaranč:

$$\text{Povprečna cena} = \frac{\text{skupna vrednost prodanih pomaranč}}{\text{prodana količina}} =$$

$$= \frac{(34 \times 2,5) + (65 \times 2,7) + (32 \times 3,2)}{34 + 65 + 32} = \frac{85+175,5+102,4}{131} = \frac{362,9}{131} = 2,77 \text{ EUR za kg.}$$

6.4 HARMONIČNA SREDINA

V statistični analizi uporabljamo harmonično sredino predvsem pri izračunavanju povprečij iz relativnih števil. Tako je v določenih primerih le s harmonično sredino mogoče izračunati povprečni strukturni odstotek in povprečni statistični koeficient.

Enostavno harmonično sredino za N vrednosti spremenljivke *y* izračunamo po obrazcu:

$$\boldsymbol{H = \frac{N}{\frac{1}{y_1} + \frac{1}{y_2} + \ldots + \frac{1}{y_N}} = \frac{N}{\sum_{i=1}^{N} \frac{1}{y_i}}} \qquad (6.8)$$

Primer 6.3: Za zgled vzemimo, da tri šivilje opravljajo enako delovno operacijo, ki jo opravijo v času 10, 15 in 12 minut. Izračunajmo povprečni čas s harmonično sredino:

$$H = \frac{3}{\frac{1}{10} + \frac{1}{15} + \frac{1}{12}} = \frac{3}{\frac{15}{60}} = \frac{180}{15} = 12 \text{ minut}$$

Povprečni čas za opravljeno delovno operacijo je 12 minut.

Z aritmetično sredino pa:

$$M = \bar{y} = \frac{10 + 15 + 12}{3} = 12,33 \text{ minute}$$

Število skupno opravljenih delovnih operacij npr. v 180 minutah mora biti enako na osnovi posamične in povprečne učinkovitosti. Ugotovimo lahko, da prva šivilja v 60 minutah opravi 6, druga 4 in tretja 5 delovnih operacij, torej vse tri skupaj 15. Enak rezultat dobimo z izračunom harmonične sredine.

Z uporabo aritmetične sredine pa ugotavljamo, da bi bilo opravljenih v 180 minutah 14,6 delovnih operacij, kar kaže, da aritmetična sredina napačno oceni povprečni čas.

Tehtano harmonično sredino uporabimo v primerih, ko se posamezne vrednosti pojavljajo z različno pogostnostjo frekvenc. Obrazec za tehtano harmonično sredino zapišemo:

$$\boldsymbol{H = \frac{f_1 + f_2 + \ldots + f_k}{\frac{f_1}{y_1} + \frac{f_2}{y_2} + \ldots + \frac{f_k}{y_k}} = \frac{\sum_{j=1}^{k} f_j}{\sum_{j=1}^{k} \frac{f_j}{y_j}}} \qquad (6.9)$$

6.4.1 Uporaba tehtane aritmetične oziroma tehtane harmonične sredine za izračun povprečja

Študentje so pogosto v dvomih, ali naj za izračun povprečja uporabijo tehtano aritmetično ali tehtano harmonično sredino. Teh dvomov ne boste imeli, če si boste obrazec za izračun naloge nastavili z besedami. Katero sredino uporabiti, je namreč odvisno od podatkov, ki jih imate na voljo. Iz nastavljenega obrazca boste tudi ugotovili, katerih podatkov nimate in kako manjkajoče podatke izračunati. Naslednji primeri naj vam pomagajo odpraviti dvome glede uporabe ustrezne sredine za izračun povprečja.

6.4.1.1 Izračun povprečja iz osnovnih podatkov

- Primer 6.4: **Na tržnici *Lepa pot* so štiri branjevke na dan 5. 6. 2008 prodajale jagode po različni ceni. Izračunajmo povprečno ceno prodanih jagod, če so znani podatki o prodanih količinah in cenah za kilogram.**

Tabela 6.3: **Prodane količine jagod in cene na dan 5. 6. 2008**

Branjevka	**Prodana količina v kg**	**Cena za kg v EUR**
B1	35	4,80
B2	38	5,10
B3	42	4,90
B4	23	5,20

Vir: Fiktivni podatki

Ker nimamo podatka o izkupičku posamezne branjevke, ga izračunamo tako, da pomnožimo prodano količino s ceno za kilogram. Izkupičke vseh štirih branjevk seštejemo in delimo s skupno prodano količino.

Za izračun povprečja smo uporabili tehtano aritmetično sredino. Teže ali ponderji so v tem primeru prodane količine.

$$\text{Povprečna cena} = \frac{\text{skupna vrednost prodaje v EUR}}{\text{prodana količina}} =$$

$$= \frac{(35\times 4,80)+(38\times 5,10)+(42\times 4,90)+(23\times 5,20)}{35+38+42+23} = \frac{687,2}{138} = 4,98 \text{ EUR za kg.}$$

- Primer 6.5: **Izračunajmo povprečno ceno prodanih jagod na dan 5. 6. 2008 na tržnici *Lepa pot*, če so za štiri branjevke znani podatki o izkupičku od prodaje in ceni za kilogram.**

Tabela 6.4: **Izkupiček od prodaje jagod in cena za kg na dan 5. 6. 2008**

Branjevka	**Izkupiček od prodaje**	**Cena za kg v EUR**
B1	168,00	4,80
B2	193,80	5,10
B3	205,80	4,90
B4	119,60	5,20

Vir: Fiktivni podatki

Med podatki nimamo podatkov o prodani količini posamezne branjevke. Izračunamo jih tako, da izkupiček od prodaje delimo s ceno. Uporabimo tehtano harmonično sredino. Teže ali ponderji so v tem primeru izkupički posameznih branjevk.

$$\text{Povprečna cena} = \frac{\text{skupna vrednost prodaje v EUR}}{\text{prodana količina}} =$$

$$= \frac{168,00+193,80+205,80+119,60}{\frac{168,00}{4,8}+\frac{193,80}{5,1}+\frac{205,80}{4,9}+\frac{119,60}{5,2}} = \frac{687,20}{138} = 4,98 \text{ EUR za kg.}$$

6.4.1.2 Izračun povprečja iz odstotnih števil

- Primer 6.6: **Izračunajmo povprečni odstotek izdelkov z napako treh serij proizvodnje, če so znani podatki o številu izdelkov v seriji in odstotku izdelkov z napako za tri serije proizvodnje.**

Tabela 6.5**: Število izdelkov v seriji in odstotek izdelkov z napako po serijah proizvodnje**

Serija	**Število izdelkov v seriji**	**Odstotek izdelkov z napako**
A	3.815	4,3
B	2.933	2,8
C	4.812	3,1

Vir: Fiktivni podatki

Nimamo podatkov o številu izdelkov z napako v posamezni seriji. Izračunamo ga tako, da skupno število izdelkov množimo z odstotkom izdelkov z napako v vsaki seriji in dobljeno število delimo s 100. Za izračun povprečja vseh serij uporabimo tehtano aritmetično sredino. Teže ali ponderji so podatki o skupnem številu izdelkov.

$$\text{Povprečni odstotek izdelkov z napako} = \frac{\text{število izdelkov z napako}}{\text{število vseh izdelkov v seriji}} \times 100 =$$

$$= \frac{3.815\times\frac{4,3}{100}+2.933\times\frac{2,8}{100}+4.812\times\frac{3,1}{100}}{3.815+2.933+4.812}\times 100 = 3,42\ \%.$$

- Primer 6.7: **Izračunajmo povprečni odstotek izdelkov z napako, če so znani podatki o številu izdelkov z napako in odstotku izdelkov z napako za tri serije proizvodnje.**

Tabela 6.6: **Število izdelkov z napako in odstotek izdelkov z napako po serijah proizvodnje**

Serija	**Število izdelkov z napako v seriji**	**Odstotek izdelkov z napako**
A	164	4,3
B	82	2,8
C	149	3,1

Vir: Fiktivni podatki

Nimamo podatkov o skupnem številu izdelkov v seriji. Za posamezno serijo to število izračunamo tako, da število izdelkov z napako v seriji delimo z odstotkom izdelkov z napako in to pomnožimo s 100. Za izračun povprečja vseh treh serij uporabimo tehtano harmonično sredino. Ponderji so podatki o številu izdelkov z napako.

$$\text{Povprečni odstotek izdelkov z napako} = \frac{\text{število izdelkov z napako}}{\text{število vseh izdelkov v seriji}} \times 100 =$$

$$= \frac{164+82+149}{\frac{164}{4,3}\times 100 + \frac{82}{2,8}\times 100 + \frac{149}{3,1}\times 100} \times 100 = 3,42\ \%.$$

Dobili smo enak rezultat, čeprav se vmesni rezultati nekoliko razlikujejo od rezultatov v gornjem izračunu. Razlog je zaokroževanje podatkov.

6.4.1.3 Izračun povprečja iz statističnih koeficientov

- Primer 6.8: **Izračunajmo povprečno število prodajalcev na prodajalno v trgovskem podjetju Preskrba v letu 2008, če so znani podatki o številu prodajaln in povprečnem številu prodajalcev na prodajalno za tri vrste prodajaln.**

Tabela 6.7: **Število prodajaln in povprečno število prodajalcev na prodajalno v trgovskem podjetju *Preskrba* 30. 6. 2008 po vrsti prodajaln**

Vrsta prodajalne	**Število prodajaln**	**Število prodajalcev na prodajalno**
Prodajalne z živili	83	6,2
Prodajalne s tekstilom in konfekcijo	27	4,8
Prodajalne z obutvijo in usnjenimi izdelki	14	3,7

Vir: Fiktivni podatki

Nimamo podatkov o številu vseh prodajalcev. Število prodajalcev v posamezni vrsti prodajalne izračunamo tako, da število prodajaln pomnožimo s številom prodajalcev na prodajalno. Povprečje za vse vrste prodajaln izračunamo s tehtano aritmetično sredino, v kateri so ponderji podatki o številu prodajaln.

$$\text{Povprečno število prodajalcev na prodajalno} = \frac{\text{število vseh prodajalcev}}{\text{število prodajaln}} =$$

$$= \frac{(83\times 6,2)+(27\times 4,8)+(14\times 3,7)}{83+27+14} = \frac{514,6+129,6+51,8}{124} = \frac{696}{124} = 5,6.$$

- Primer 6.9: **Izračunajmo povprečno število prodajalcev na prodajalno, če so znani podatki o številu vseh prodajalcev in številu prodajalcev na prodajalno za tri vrste prodajaln.**

Tabela 6.8: **Število vseh prodajalcev in število prodajalcev na prodajalno v trgovskem podjetju *Preskrba* 30. 6. 2008 po vrsti prodajaln**

Vrsta trgovine	Število prodajalcev	Število prodajalcev na prodajalno
Prodajalne z živili	515	6,2
Prodajalne s tekstilom in konfekcijo	130	4,8
Prodajalne z obutvijo in usnjenimi izdelki	51	3,7

Vir: Fiktivni podatki

Nimamo podatkov o številu prodajaln za posamezne vrste trgovine na drobno. Število prodajaln izračunamo tako, da število zaposlenih oseb delimo s številom zaposlenih oseb na prodajalno. Povprečno število zaposlenih oseb na prodajalno za vse trgovine na drobno izračunamo s tehtano harmonično sredino.

$$\text{Povprečno število prodajalcev na prodajalno} = \frac{\text{število vseh prodajalcev}}{\text{število prodajaln}} =$$

$$= \frac{515+130+51}{\frac{515}{6,2}+\frac{130}{4,8}+\frac{51}{3,7}} = \frac{696}{83,1+27,1+13,8} = \frac{696}{124} = 5,6.$$

6.4.1.4 Izračun povprečnega koeficienta obračanja zalog

- Primer 6.10: **Izračunajmo mesečni koeficient obračanja zalog, če so znani podatki o povprečni mesečni zalogi in koeficientu obračanja zalog za mesec junij v letu 2008 za tri vrste surovin.**

Tabela 6.9: **Povprečna zaloga in koeficient obračanja zalog za tri vrste surovin**

Vrsta surovin	Povprečna zaloga v 1.000 EUR	Koeficient obračanja zalog
A	5,4	2,1
B	6,9	1,8
C	8,7	4,3

Vir: Fiktivni podatki

$$\text{Koeficient obračanja zalog} = \frac{\text{poraba}}{\text{zaloga}} \times \text{čas}$$

Nimamo podatkov o porabi posamezne vrste surovin. Porabo posamezne vrste surovin izračunamo tako, da zalogo pomnožimo s koeficientom obračanja zalog. Za izračun povprečnega koeficienta obračanja zalog vseh vrst surovin uporabimo tehtano aritmetično sredino, v kateri so ponderji podatki o povprečni zalogi. Za čas vzamemo 1, saj gre za mesečni koeficient.

$$\text{Povprečni koeficient obračanja zalog} = \frac{\text{poraba}}{\text{zaloga}} \times \text{čas} =$$

$$= \frac{(5,4\times2,1)+(6,9\times1,8)+(8,7\times4,3)}{5,4+6,9+8,7} = \frac{11,34+12,42+37,41}{21} = \frac{61,17 \text{ tisoč EUR}}{21 \text{ tisoč EUR}} = 2,9.$$

Povprečno mesečno so se zaloge obrnile 2,9-krat.

- Primer 6.11: **Izračunajmo povprečni koeficient obračanja zalog, če so znani podatki o mesečni porabi in koeficientu obračanja zalog za mesec junij 2008 za tri vrste surovin.**

Tabela 6.10: **Poraba in koeficient obračanja zalog za tri vrste surovin**

Vrsta surovin	**Poraba v 1.000 EUR**	**Koeficient obračanja zalog**
A	11,3	2,1
B	12,4	1,8
C	37,4	4,3

Vir: Fiktivni podatki

Ni Koeficient obračanja zalog $= \dfrac{\text{poraba}}{\text{zaloga}} \times \text{čas}$ mamo podatkov o stanju zaloge za posamezne vrste surovin. Izračunamo jih tako, da porabo delimo s koeficientom obračanja zalog. Povprečni koeficient obračanja zalog izračunamo s tehtano harmonično sredino, v kateri za ponderje vzamemo podatke o porabi surovin.

$$\text{Povprečni koeficient obračanja zalog} = \frac{\text{poraba}}{\text{zaloga}} \times \text{čas} =$$

$$= \frac{11,3+12,4+37,4}{\frac{11,3}{2,1}+\frac{12,4}{1,8}+\frac{37,4}{4,3}} = \frac{61,1}{5,4+6,9+8,7} = \frac{61,1 \text{ tisoč EUR}}{21 \text{ tisoč EUR}} = 2,9.$$

6.5 GEOMETRIJSKA SREDINA

Geometrijska sredina za N pozitivnih vrednosti je enaka N-temu korenu iz produkta teh vrednosti:

$$\boldsymbol{G = \sqrt[N]{y_1 \times y_2 \times y_3 \times ..... \times y_N}} \qquad (6.10)$$

Geometrijsko sredino uporabljamo predvsem pri analizi časovnih vrst, ko želimo za preteklo obdobje poznati ***povprečno stopnjo rasti, povprečni koeficient rasti ali povprečni verižni indeks.*** Na osnovi teh lahko nato predvidimo razvoj pojava v prihodnosti, seveda ob predpostavki, da se razmere ne bodo spremenile.

6.5.1 Izračunavanje povprečnega koeficienta rasti

Koeficient rasti kaže relativno spremembo med dvema zaporednima časovnima trenutkoma ali intervaloma. Tako lahko problem ponazorimo podobno kot pri obrestno obrestnem računu: podatek za pojav v obdobju N, torej Y_N, izrazimo z začetnim stanjem Y_0 pomnoženim s koeficienti rasti od obdobja do obdobja:

$$\boldsymbol{Y_N = Y_0 \times K_1 \times K_2 \times ..... \times K_N} \qquad (6.11)$$

Enako oziroma stalno relativno spremembo iz obdobja v obdobje kaže povprečni koeficient rasti, tako lahko zapišemo:

$$\boldsymbol{Y_N = Y_0 \times \overline{K}^N} \tag{6.12}$$

in iz tega: $\boldsymbol{\overline{K} = \sqrt[N]{\frac{Y_N}{Y_0}}}$ *(6.13)*

Povprečni koeficient rasti lahko z geometrijsko sredino izračunamo tudi iz koeficientov rasti za posamezna obdobja:

$$\boldsymbol{\overline{K} = \sqrt[N]{K_1 \times K_2 \times K_3 \times \ldots \times K_N}} \tag{6.14}$$

Povprečni koeficient rasti je N-ti koren iz produkta koeficientov, pri čemer je N število koeficientov.

Izračunajmo povprečni koeficient rasti na primeru števila brezposelnih oseb po anketi v Sloveniji v letih 2002 do 2008.

Tabela 6.12: **Brezposelne osebe po anketi v Sloveniji v letih od 2002 do 2008, koeficienti rasti, indeksi s stalno osnovo 2002 = 100 in 2004 = 100**

Leto	Število brezposelnih po anketi v tisoč	Koeficienti rasti	Indeksi s stalno osnovo 2002 = 100	Indeksi s stalno osnovo 2004 = 100
2002	62	-	100,0	96,9
2003	64	1,032	103,2	100,0
2004	64	1,000	103,2	100,0
2005	67	1,047	108,1	104,7
2006	61	0,910	98,4	95,3
2007	50	0,820	80,6	78,1
2008	52	1,040	83,9	81,3

Vir: Pomembnejši statistični podatki o Sloveniji, letnik III, št. 2/2008 in Statistične informacije

- **Izračunajmo najprej povprečni koeficient rasti iz osnovnih podatkov** (za število obdobij N ne štejemo izhodiščnega leta, torej velja N = 6):

$\boldsymbol{\overline{K}} = \sqrt[N]{\frac{\boldsymbol{Y_N}}{\boldsymbol{Y_0}}} = \sqrt[6]{\frac{52}{62}} = 0,971$ in povprečna stopnja rasti:

$$\overline{S} = (\overline{K} - 1) \times 100 = (0,971 - 1) \times 100 = -2,9\ \%$$

Število brezposelnih oseb po anketi se je v obdobju od leta 2002 do 2008 povprečno letno zmanjševalo za 2,9 %.

- **Izračunajmo povprečni koeficient rasti še iz koeficientov rasti**:

$$\boldsymbol{\overline{K} = \sqrt[N]{K_1 \times K_2 \times K_3 \times \ldots \times K_N}} = \sqrt[6]{1,032 \times 1,000 \times 1,047 \times 0,910 \times 0,833 \times 1,040} = 0,971$$

Število brezposelnih oseb po anketi se je v obdobju od leta 2002 do 2008 povprečno letno zmanjševalo za 2,9 %.

- **Izračunajmo povprečni koeficient rasti še iz indeksov s stalno osnovo**. Izračunamo ga kot iz osnovnih podatkov:

  − najprej iz indeksov s stalno osnovo 2002 = 100 $\boldsymbol{\overline{K} = \sqrt[N]{\frac{I_{N/0}}{I_{1/0}}}} = \sqrt[6]{\frac{83,9}{100,0}} = 0,971$

  − in še indeksov s stalno osnovo 2004 = 100

  $$\overline{K} = \sqrt[6]{\frac{81,3}{96,9}} = 0,971$$

  in iz tega povprečno stopnjo rasti:

  $$\overline{S} = (\overline{K} - 1) \times 100 = (0,971 - 1) \times 100 = -2,9\ \%$$

V obeh primerih dobimo enak rezultat, torej različna osnova v indeksni vrsti ne vpliva na rezultat. Povprečno letno se je število brezposelnih po anketi zmanjševalo za 2,9 %.

6.5.2 Izračunavanje povprečnega verižnega indeksa

- **neposredno iz verižnih indeksov,** in sicer je N-ti koren produkta iz verižnih indeksov:

$$\boldsymbol{\overline{V} = \sqrt[N]{V_1 \times V_2 \times V_3 \times ..... \times V_N}} \qquad (6.15)$$

- **iz povprečnega koeficienta rasti:**

$$\boldsymbol{\overline{V} = 100 \times \overline{K}} \qquad (6.16)$$

- **iz podatkov, ki se nanašata na začetno in končno stanje za opazovani pojav:**

$$\boldsymbol{\overline{V} = \sqrt[N]{\frac{y_N}{Y_0}} \times 100} \qquad (6.17)$$

6.5.3 Izračunavanje povprečne stopnje rasti

- Povprečne stopnje rasti ne moremo izračunati neposredno iz stopenj rasti za posamezna obdobja. Le-te so v nekaterih obdobjih lahko enake nič ali celo negativne, tako bi dobili nesmiseln rezultat. Zato stopnje rasti spremenimo v verižne indekse ali koeficiente rasti,

- izračunamo povprečni verižni indeks ali povprečni koeficient rasti, iz tega pa povprečno stopnjo rasti.

  Izračunajmo povprečno letno stopnjo rasti iz stopenj rasti za število brezposelnih po anketi v Sloveniji v letih od 2002 do 2008.

Tabela 6.13: **Stopnje rasti in verižni indeksi za število brezposelnih oseb po anketi v Sloveniji v letih od 2002 do 2008**

Leto	Stopnje rasti S_j	Verižni indeksi V_j
2002	-	-
2003	3,2	103,2
2004	0,0	100,0
2005	4,7	104,7
2006	–9,0	91,0
2007	–16,7	83,3
2008	4,0	104,0

- Stopnje rasti spremenimo v verižne indekse:

$$V_{2003} = S_{2003} + 100 = 3,2 + 100,0 = 103,2$$
$$V_{2004} = S_{2004} + 100 = 0,0 + 100,0 = 100,0 \text{ itd.}$$

- Izračunamo povprečni verižni indeks:

$$\overline{V} = \sqrt[N]{V_{2003} \times V_{2004} \times \ldots \times V_{2008}} = \sqrt[6]{103,2 \times 100,0 \times 104,7 \times 91,0 \times 83,3 \times 104,0} = 97,1$$

- Povprečni verižni indeks pretvorimo v povprečno stopnjo rasti:

$$\overline{S} = \overline{V} - 100,0 = 97,1 - 100,0 = -2,9\ \%$$

Število brezposelnih oseb po anketi se je v obdobju od leta 2002 do 2008 povprečno letno zmanjševalo za 2,9 %.

ALI:

- Stopnje rasti spremenimo v koeficiente rasti:

$$K_j = \frac{S_j}{100} + 1$$
$$K_{2003} = \frac{S_{2003}}{100} + 1 = \frac{3,2}{100} + 1 = 1,032 \text{ (zapisani v tabeli 6.12)}$$

- Izračunamo povprečni koeficient rasti:

$$\overline{K} = \sqrt[N]{K_1 \times K_2 \times K_3 \times \ldots \times K_N} = \sqrt[6]{1,032 \times 1,000 \times 1,031 \times .... \times 1,041} = 0,971$$

- Iz povprečnega koeficienta rasti dobimo povprečno stopnjo rasti:

$$\overline{S} = (\overline{K} - 1) \times 100 = (0,971 - 1) \times 100 = -2,9\ \%$$

Na osnovi povprečnega koeficienta rasti lahko ocenimo pojav v prihodnosti, seveda s predpostavko, da bodo tudi v prihodnosti enake razmere.

Ocenimo število brezposelnih oseb po anketi v Sloveniji v letu 2011, če je bilo 52 tisoč brezposelnih v letu 2008:

$$Y_{2011} = Y_{2008} \times \overline{K}^3 = 52 \times 0{,}971^3 = 47{,}6 \text{ tisoč}$$

Ob predpostavki, da se bo število brezposelnih po anketi zmanjševalo po enaki povprečni letni stopnji kot v letih 2002 do 2008, ocenjujemo, da bo leta 2011 v Sloveniji 47,6 tisoč brezposelnih oseb.

Srednje vrednosti so:

- *mediana kot sredinska vrednost, od katere ima polovica enot manjše, polovica pa večje vrednosti;*
- *modus kot najpogostejša vrednost in jo računamo le iz frekvenčnih porazdelitev z enakimi širinami razredov;*
- *aritmetična sredina kot povprečna vrednost in jo najpogosteje računamo;*
- *harmonična sredina prav tako povprečna vrednost in*
- *geometrijska sredina, s katero računamo povprečja iz indeksov in koeficientov rasti.*

Pri statistični analizi določenega pojava praviloma ne računamo vseh, ampak tisto oziroma tiste, s katerimi najbolje osvetlimo njegove značilnosti.

V Zbirki vaj iz statistike so za utrjevanje znanja iz srednjih vrednosti pripravljene naloge od 5.1 do 5.17.

7 MERE VARIABILNOSTI, ASIMETRIJE IN SPLOŠČENOSTI

Do sedaj smo že spoznali, da imajo nekatere spremenljivke veliko vrednosti, druge manj. Veliko vrednosti imajo predvsem številske spremenljivke. Običajno se te vrednosti veliko ne razlikujejo od izračunanega povprečja, ki smo ga spoznali v prejšnjem poglavju. Analizo spremenljivk lahko dopolnimo z merami variabilnosti, s katerimi ugotavljamo značilnosti porazdeljevanja posameznih vrednosti spremenljivk. Poleg mer variabilnosti, od katerih sta najpomembnejši varianca in standardni odklon, boste spoznali in računali še mere asimetrije in sploščenosti, s katerimi ugotavljamo podobnost preučevane populacije s teoretično normalno populacijo.

7.1 MERE VARIABILNOSTI

Srednje vrednosti so statistični parametri, ki jih največkrat uporabljamo v statistični analizi, le-to pa dopolnimo z merami variabilnosti. Če poznamo povprečno vrednost prodaje 80 prodajaln, nas zanima vsaj še podatek, ki pove razpon med najmanjšo in največjo vrednostjo prodaje. Pri tem upoštevamo le dva podatka, zato ta parameter ni dobra mera variabilnosti. Boljši so tisti, ki jih izračunamo iz vseh vrednosti spremenljivke.

Z merami variabilnosti torej ugotavljamo, kako se vrednosti spremenljivke med seboj razlikujejo. Večje kot so razlike med njimi, večja je variabilnost. Na splošno jih delimo na:

- ***absolutne mere variabilnosti,*** kot so: variacijski razmik, kvartilni razmik, decilni razmik, povprečni absolutni odklon od aritmetične sredine in mediane ter varianca kot teoretično najpomembnejša mera variabilnosti;
- ***relativne mere variabilnosti,*** kot je koeficient variabilnosti**,** ki je edini primeren za primerjavo variabilnosti različnih pojavov.

7.1.1 Variacijski razmik

Variacijski razmik je najenostavnejša mera variabilnosti, ki jo izračunamo kot razliko med največjo in najmanjšo vrednostjo spremenljivke:

$$\boldsymbol{VR = y_{max} - y_{min}} \qquad (7.1)$$

Izračunamo ga lahko iz posameznih vrednosti, ne pa iz frekvenčnih porazdelitev, saj v teh posamezne vrednosti niso razvidne.

Iz podatkov o vrednosti prodaje 80 prodajaln lahko ugotovimo, da je bila najmanjša vrednost prodaje 10,2 milijona evrov, največja pa 23,8 milijona evrov. Prodajalna z največjo vrednostjo prodaje ima kar za 13,6 milijona evrov večjo vrednost prodaje kot tista z najmanjšo vrednostjo. Če to primerjamo s povprečno vrednostjo prodaje in količnik pomnožimo s 100, ugotovimo, da razpon med najmanjšo in največjo vrednostjo predstavlja 87,7 % povprečne vrednosti prodaje:

$$VR = y_{\max} - y_{\min} = 23,8 - 10,2 = 13,6$$

$$100 \times \frac{VR}{M} = 100 \times \frac{13,6}{15,5} = 87,7\ \%$$

Ko smo enote razvrstili v razrede frekvenčne porazdelitve, najmanjša in največja vrednost nista več razvidni in variacijskega razmika ne moremo izračunati.

7.1.2 Kvartilni in decilni razmik

Kvartilni razmik je absolutna mera variabilnosti, ki predstavlja ***razliko med tretjim in prvim kvartilom:***

$$\boldsymbol{Q = Q_3 - Q_1} \qquad (7.2)$$

Izračunajmo ga na primeru rezultatov pisnega preverjanja iz angleškega jezika za 14 študentov.

Zapišimo ranžirno vrsto:

R	1	2	3	**4**	5	6	7	8	9	10	**11**	12	13	14
y_i	24	27	29	**31**	33	35	38	43	45	47	**48**	50	53	57

$P = 0,25 \ \ in \ \ R = N \times P + 0,5 = 14 \times 0,25 + 0,5 = 4$

$P = 0,75 \ \ in \ \ R = N \times P + 0,75 = 14 \times 0,75 + 0,5 = 11$

$$\boldsymbol{Q = Q_3 - Q_1 = 48 - 31 = 17}$$

Prvi kvartil je vrednost četrtega člena, tretji pa vrednost enajstega člena v ranžirni vrsti. Razlika je 17, kar pomeni, da je bila pri 50 % srednje uspešnih študentov največja razlika pri pisnem preverjanju iz angleškega jezika 17 točk.

Decilni razmik je absolutna mera variabilnosti, ki predstavlja ***razliko med devetim in prvim decilom:***

$$\boldsymbol{D = D_9 - D_1} \qquad (7.3)$$

Iz ranžirne vrste izračunajmo prvi in deveti decil:

$P = 0,10$ in $R = 14 \times 0,10 + 0,5 = 1,9$

$$D_1 = 24 + \frac{1,9 - 1}{2 - 1} \times (27 - 24) = 26,7 \text{ točke}$$

$P = 0,90$ in $R = 14 \times 0,90 + 0,5 = 13,1$

$$D_9 = 53 + \frac{13,1 - 13}{14 - 13} \times (57 - 53) = 53,4 \text{ točke}$$

D = D$_9$ – D$_1$ = 53,4 – 26,7 = 26,7 točke

Prvi decil je 26,7 točke, deveti pa 53,4 točke, razlika, torej decilni razmik je 26,7 točke, kar pomeni, da je bila pri 80 % srednje uspešnih študentov največja razlika 26,7 točke.

Kvartilni in decilni razmik lahko izračunamo tudi za frekvenčno porazdelitev vrednosti prodaje 80 prodajaln trgovskega podjetja ***Preskrba***.

Zapišimo frekvenčno porazdelitev in izračunajmo prvi in tretji kvartil ter prvi in deveti decil.

Tabela 7.1: **Frekvenčna porazdelitev vrednosti prodaje za 80 prodajaln trgovskega podjetja *Preskrba* v letu 2007**

Vrednost prodaje v mio EUR	**Število prodajaln** f_j	**Kumulativa frekvenc** F_j
od 10 do pod 12	9	9
od 12 do pod 14	16	25
od 14 do pod 16	23	48
od 16 do pod 18	17	65
od 18 do pod 20	10	75
od 20 do pod 22	3	78
od 22 do pod 24	2	80
Skupaj	**80**	

Kvartilni razmik

Izračunajmo prvi in tretji kvartil:

$P = 0,25$ *in* $R = N \times P + 0,5 = 80 \times 0,25 + 0,5 = 20,5$; prvi kvartil je v drugem razredu: $j = 2$

$F_1 = 9 < R = 20,5 < F_2 = 25$

$y_{2,\min} = 12 < Q_1 = y_{P=0,25} < y_{2,\max} = 14$

$$y = Q_1 = y_{j,\min} + d_j \times \frac{R - F_{j-1}}{f_j} = 12 + 2 \times \frac{20,5 - 9}{16} = 13,44 \text{ mio EUR}$$

$P = 0,75$ *in* $R = N \times P + 0,75 = 80 \times 0,75 + 0,5 = 60,5$; tretji kvartil je v četrtem razredu: $j = 4$

$$Q_3 = 16 + 2 \times \frac{60,5 - 48}{17} = 17,47 \text{ mio EUR}$$

Q = Q$_3$ – Q$_1$ = 17,47 – 13,44 = 4,03 mio EUR

Prvi kvartil je 13,44 milijona evrov, tretji kvartil pa 17,47 milijona evrov. Razlika, torej kvartilni razmik je 4,03 milijona evrov, kar pomeni, da je bila pri 50 % prodajaln s srednje veliko vrednostjo prodaje največja razlika v prodaji 4,03 milijona evrov.

Decilni razmik

Izračunajmo prvi in deveti decil:

$P = 0{,}10$ in $R = N \times P + 0{,}5 = 80 \times 0{,}10 + 0{,}5 = 8{,}5$; prvi decil je v prvem razredu: $j = 1$

$F_0 = 0 < R = 8{,}5 < F_1 = 9$

$y_{1,\min} = 10 < D_1 = y_{P=0,10} < y_{1,\max} = 12$

$$D_1 = 10 + 2 \times \frac{8{,}5 - 0}{9} = 11{,}89 \text{ mio EUR}$$

$P = 0{,}90$ in $R = N \times P + 0{,}5 = 80 \times 0{,}90 + 0{,}5 = 72{,}5$; deveti decil je v petem razredu in: $j = 5$

$$D_9 = 18 + 2 \times \frac{72{,}5 - 65}{10} = 19{,}5 \text{ mio EUR}$$

$$\boldsymbol{D = D_9 - D_1 = 19{,}5 - 11{,}89 = 7{,}61 \text{ mio EUR}}$$

Deveti decil je 19,5 milijona evrov, prvi pa 11,89 milijona evrov, razlika, torej decilni razmik je 7,61 milijona evrov, kar pomeni, da je bila pri 80 % prodajaln s srednje veliko vrednostjo največja razlika vrednosti prodaje 7,61 milijona evrov.

7.1.3 Povprečni absolutni odklon od aritmetične sredine in mediane

Za mero variabilnosti lahko uporabimo **povprečni absolutni odklon od aritmetične sredine**:

$$\boldsymbol{AD_M = \frac{1}{N} \sum_{i=1}^{N} |y_i - M|} \qquad (7.4)$$

in povprečni absolutni odklon od mediane:

$$\boldsymbol{AD_{Me} = \frac{1}{N} \sum_{i=1}^{N} |y_i - Me|} \qquad (7.5)$$

Za rezultate pisnega preverjanja iz angleškega jezika za 14 študentov, izražene v številu točk

y_i: 24 31 29 57 53 45 47 50 38 33 35 27 43 48

je povprečni absolutni odklon od aritmetične sredine:

$$M = \frac{1}{N} \sum_{i=1}^{N} y_i = \frac{1}{14} 560 = 40 \text{ točk}$$

$$\boldsymbol{AD_M = \frac{1}{14}\left(|24 - 40| + |27 - 40| + ... + |57 - 40|\right) = \frac{1}{14} \times 126 = 9 \text{ točk}}$$

Aritmetična sredina je 40 točk, povprečni absolutni odklon od aritmetične sredine pa 9 točk.

Povprečni absolutni odklon od mediane:

$$P = 0,50 \text{ in } R = 14 \times 0,50 + 0,5 = 7,5$$

$$Me = 38 + \frac{7,5-7}{8-7} \times (43-38) = 40,5 \text{ točke}$$

$$\boldsymbol{AD_{Me} = \frac{1}{14}\left(|24-40,5| + |27-40,5| + ... + |57-40,5|\right) = \frac{1}{14} \times 116 = 8,3 \text{ točke}}$$

Mediana je 40,5 točke, povprečni absolutni odklon od mediane pa je 8,3 točke.

Pri izračunavanju povprečnega absolutnega odklona iz frekvenčne porazdelitve moramo upoštevati še frekvence razredov, sredine razredov pa nadomestijo posamezne vrednosti. Obrazec torej priredimo.

Povprečni absolutni odklon od aritmetične sredine:

$$\boldsymbol{AD_M = \frac{1}{N}\sum_{j=1}^{k} f_j \left|y_j - M\right|} \tag{7.6}$$

Povprečni absolutni odklon od mediane:

$$\boldsymbol{AD_{Me} = \frac{1}{N}\sum_{j=1}^{k} f_j \left|y_j - Me\right|} \tag{7.7}$$

Za 80 prodajaln po vrednosti prodaje v letu 2007 je povprečni absolutni odklon od aritmetične sredine:

$$\boldsymbol{AD_M = \frac{1}{80}\left(9 \times |11-15,5| + 16 \times |13-15,5| + ... + 2 \times |23-15,5|\right) =}$$

$$\boldsymbol{= \frac{1}{80} \times 184 = 2,3 \text{ mio EUR}}$$

Povprečni absolutni odklon od mediane:

$$\boldsymbol{AD_{Me} = \frac{1}{80}\left(9 \times |11-15,35| + 16 \times |13-15,35| + ... + 2 \times |23-15,35|\right) =}$$

$$\boldsymbol{= \frac{1}{80} \times 181,6 = 2,27 \text{ mio EUR}}$$

Povprečni absolutni odklon od aritmetične sredine je 2,3 milijona evrov, povprečni absolutni odklon od mediane pa 2,27 milijona evrov.

7.1.4 Varianca in standardni odklon

Varianca, ki ima podobno osnovo kot povprečni absolutni odklon od aritmetične sredine, je najpomembnejša mera variabilnosti. Pri izračunu variance upoštevamo razlike (odklone) opazovanih vrednosti številske spremenljivke Y od aritmetične sredine. Kaže razlike posameznih vrednosti od aritmetične sredine. Ker imajo nekatere enote večje, druge pa manjše vrednosti od aritmetične sredine, je vsota teh razlik enaka 0. Zato pri izračunu variance te razlike kvadriramo. ***Varianco tako opredelimo kot povprečje kvadratov odklonov vrednosti številske spremenljivke od aritmetične sredine*** (Košmelj, 1994, 171).

7.1.4.1 Varianca iz posameznih vrednosti

- **Za vsako vrednost izračunamo razliko od aritmetične sredine**:

  $y_i - M$

- **Razlike kvadriramo**:

  $(y_i - M)^2$

- **Kvadrate razlik seštejemo in delimo s številom enot.** Varianco tako izračunamo po obrazcu:

$$VAR = \sigma^2 = \frac{1}{N}\sum_{i=1}^{N}(y_i - M)^2 \tag{7.8}$$

- Zaradi enostavnejšega izračunavanja pogosteje uporabljamo **obrazec, ki je izveden iz zgoraj navedenega:**

$$VAR = \sigma^2 = \frac{1}{N}\sum_{i=1}^{N} y_i^2 - M^2 \tag{7.9}$$

  (aritmetična sredina M ali $\bar{y}$)

Izračunajmo varianco in standardni odklon za primer rezultatov pisnega preverjanja iz angleškega jezika za 14 študentov:

Število točk – y_i: 24 31 29 57 53 45 47 50 38 33 35 27 43 48

$$VAR = \sigma^2 = \frac{1}{N}\sum_{i=1}^{N}(y_i - M)^2 = \frac{1}{14}\times\left[(24-40)^2 + (31-40)^2 + ... + (48-40)^2\right] =$$

$$= \frac{1}{14}\times 1410 = 100{,}7\,točke^2$$

in še po drugem obrazcu:

$$VAR = \sigma^2 = \frac{1}{N}\sum_{i=1}^{N} y_i^2 - M^2 = \left[\frac{1}{14}\times\left(24^2 + 31^2 + ... + 48^2\right)\right] - 40^2 =$$

$$= \left(\frac{1}{14}\times 23810\right) - 1600 = 100{,}7\,točke^2$$

Varianca je osnovna mera variabilnosti, ki je neposredno, zaradi neprimerne enote mere (kvadrat osnovne enote), ne uporabljamo veliko. Običajno pa jo uporabljamo posredno, in sicer tako, da izračunamo kvadratni koren iz variance. To mero variabilnosti imenujemo ***standardni odklon*** in je izražen v istih merskih enotah kot opazovana številska spremenljivka.

$$\boldsymbol{SD = \sigma = \sqrt{\sigma^2}}$$

$$\boldsymbol{\sigma = \sqrt{100{,}7\ točke^2} = 10{,}03\ točke}$$

Varianca rezultatov preverjanja iz angleškega jezika je 100,7 točke^2, standardni odklon pa 10,03 točke.

7.1.4.2 Varianca iz frekvenčne porazdelitve

- V frekvenčni porazdelitvi posamezne vrednosti nadomestijo sredine razredov, tako **izračunamo razlike sredin razredov od aritmetične sredine**:

  $y_j - M$

- **Razlike v posameznih razredih kvadriramo in pomnožimo s frekvencami** teh razredov:

  $(y_j - M)^2 f_j$

- **Produkte kvadratov razlik in frekvenc seštejemo in delimo s številom enot**. Varianco tako izračunamo po obrazcu:

$$VAR = \sigma^2 = \frac{1}{N}\sum_{j=1}^{k} f_j (y_j - M)^2 \qquad (7.11)$$

- Podobno kot iz posameznih vrednosti izračunamo tudi iz frekvenčnih porazdelitev varianco po enostavnejšem obrazcu:

$$VAR = \sigma^2 = \frac{1}{N}\sum_{j=1}^{k} f_j y_j^2 - M^2 \qquad (7.12)$$

Izračunajmo varianco iz frekvenčne porazdelitve vrednosti prodaje 80 prodajaln v letu 2007.

Tabela 7.1: **Frekvenčna porazdelitev vrednosti prodaje za 80 prodajaln trgovskega podjetja *Preskrba* v letu 2007 z izračuni za varianco**

Vrednost prodaje v mio EUR	**Št. prodajaln f_j**	y_j	$y_j - M$	$(y_j - M)^2$	$(y_j - M)^2 f_j$
od 10 do pod 12	9	11	–4,5	20,25	182,25
od 12 do pod 14	16	13	–2,5	6,25	100,00
od 14 do pod 16	23	15	–0,5	0,25	5,75
od 16 do pod 18	17	17	1,5	2,25	38,25
od 18 do pod 20	10	19	3,5	12,25	122,50
od 20 do pod 22	3	21	5,5	30,25	90,75
od 22 do pod 24	2	23	7,5	56,25	112,50
Skupaj	**80**				**652,00**

$$VAR = \sigma^2 = \frac{1}{N}\sum_{j=1}^{k} f_j (y_j - M)^2 = \frac{1}{80} \times 652 = 8{,}15\,(mio\ EUR)^2$$

Pri frekvenčnih porazdelitvah z enakimi širinami razredov izračunamo popravek variance ali **Sheppardov popravek**:

$$\sigma_{cor}^2 = \sigma^2 - \frac{d^2}{12} \qquad (7.13)$$

$$\sigma^2 = 8,15 - \frac{2^2}{12} = 7,82 \ (mio\ EUR)^2$$

Varianca je 7,82 (milijona evrov)^2.

Standardni odklon:

$$\sigma = SD = \sqrt{\sigma^2} = \sqrt{7,82\ mio\ EUR^2} = 2,8\ mio\ EUR$$

Standardni odklon je 2,8 milijona evrov.

Izračunajmo varianco še po drugem obrazcu:

$$VAR = \frac{1}{N}\sum_{j=1}^{k} f_j y_j^2 - M^2 = \left(\frac{1}{80} \times 19872,00\right) - 15,5^2 = 8,15\,(mio\ EUR)^2$$

$$VAR_{cor} = 8,15 - \frac{2^2}{12} = 7,82\ (mio\ EUR)^2$$

Standardni odklon: $\sigma = SD = \sqrt{\sigma^2} = \sqrt{7,82\ mio\ EUR^2} = 2,8\ mio\ EUR$

Po tem obrazcu izračunana varianca se ne razlikuje od variance, izračunane po obrazcu 7.12.

Tabela 7.2 **Frekvenčna porazdelitev vrednosti prodaje za 80 prodajaln trgovskega podjetja *Preskrba* v letu 2007 z izračuni za varianco**

Vrednost prodaje v mio EUR	Št. prodajaln f_j	y_j	y_j^2	$f_j y_j^2$
od 10 do pod 12	9	11	121	1.089
od 12 do pod 14	16	13	169	2.704
od 14 do pod 16	23	15	225	5.175
od 16 do pod 18	17	17	289	4.913
od 18 do pod 20	10	19	361	3.610
od 20 do pod 22	3	21	441	1.323
od 22 do pod 24	2	23	529	1.058
Skupaj	**80**			**19.872**

7.1.5 Koeficient variabilnosti

Koeficient variabilnosti je relativna mera variabilnosti, ki jo izrazimo kot razmerje med standardnim odklonom in aritmetično sredino. Kadar rezultat izrazimo v odstotku, izračunani količnik pomnožimo s 100.

$$\boldsymbol{KV}\ \% = \frac{\boldsymbol{\sigma}}{\boldsymbol{M}} \times \boldsymbol{100} \tag{7.14}$$

Koeficient variabilnosti primerja standardni odklon z aritmetično sredino. Večji kot je, večja je variabilnost pojava. Ker je izražen v odstotku, je primerna mera variabilnosti, s katero moremo primerjati variabilnost različnih pojavov.

Izračunani koeficient variabilnosti za rezultate pisnega preverjanja za 14 študentov je:

$$KV\ \% = \frac{\boldsymbol{10{,}03}}{\boldsymbol{40}} \times \boldsymbol{100 = 25{,}08}\ \%$$

Standardni odklon predstavlja 25,08 % aritmetične sredine.

S primerjavo izračunanega koeficienta variabilnosti s koeficienti variabilnosti za rezultate drugih pisnih preverjanj bi lahko sklepali na stopnjo variiranja rezultatov pisnega preverjanja iz angleškega jezika.

Izračunani koeficient variabilnosti za vrednost prodaje 80 prodajaln je:

$$\boldsymbol{KV}\ \% = \frac{\boldsymbol{2{,}8}}{\boldsymbol{15{,}5}} \times \boldsymbol{100 = 18{,}06}\ \%$$

Standardni odklon predstavlja 18,06 % aritmetične sredine.

Pomen standardnega odklona bomo podrobneje opredelili, ko bomo obravnavali značilnosti teoretične normalne porazdelitve.

7.2 MERE ASIMETRIJE

Za frekvenčne porazdelitve je značilno, da so lahko:

- **simetrične**, za katere velja, da so vse tri srednje vrednosti enake, torej: $\boldsymbol{Mo = Me = M}$**;**
- **asimetrične v desno:** $\boldsymbol{Mo < Me < M}$**;**
- **asimetrične v levo:** $\boldsymbol{M < Me < Mo}$**.**

Natančnejši odgovor glede stopnje asimetrije dobimo z izračunom koeficienta asimetrije, ki ga lahko izračunamo na osnovi modusa ali mediane.

- Koeficient asimetrije na osnovi modusa:

$$\boldsymbol{KA_{Mo}} = \frac{\boldsymbol{M - Mo}}{\boldsymbol{\sigma}} \tag{7.15}$$

- Koeficient asimetrije na osnovi mediane:

$$KA_{Me} = \frac{3(M - Me)}{\sigma} \tag{7.16}$$

Na osnovi izračunanega koeficienta asimetrije sklepamo o smeri in jakosti asimetrije. Če velja:

KA_{Me}, $KA_{Mo} < 0$, je porazdelitev asimetrična v levo;

(razmerje med srednjimi vrednostmi: $M < Me < Mo$)

KA_{Me}, $KA_{Mo} > 0$, je porazdelitev asimetrična v desno;

(razmerje med srednjimi vrednostmi: $Mo < Me < M$)

KA_{Me}, $KA_{Mo} = 0$, je porazdelitev simetrična.

(razmerje med srednjimi vrednostmi: $Mo = Me = M$)

Bolj kot izračunani koeficient odstopa od 0, večja je jakost asimetrije. Pri tem upoštevamo absolutno vrednost. Teoretični meji za vrednost koeficientov asimetrije sta –3 in +3.

Asimetričnost porazdelitve je razvidna iz njenega grafičnega prikaza, vendar le smer, jakost moramo izračunati.

Izračunajmo koeficienta asimetrije za porazdelitev vrednosti prodaje 80 prodajaln trgovskega podjetja *Preskrba.*

V poglavju o srednjih vrednostih in varianci smo izračunali:

$M = 15{,}5$ mio EUR $\quad Mo = 15{,}1$ mio EUR

$Me = 15{,}35$ mio EUR $\quad \sigma = 2{,}8$ mio EUR

$$KA_{Mo} = \frac{M - Mo}{\sigma} = \frac{15{,}5 - 15{,}1}{2{,}8} = 0{,}143$$

$$KA_{Me} = \frac{3(M - Me)}{\sigma} = \frac{3(15{,}5 - 15{,}35)}{2{,}8} = 0{,}161$$

$$KA_{Mo}, KA_{Me} > 0$$

Porazdelitev je zelo rahlo asimetrična v desno, kar lahko sklepamo že iz razmerja med srednjimi vrednostmi: $Mo = 15{,}1 < Me = 15{,}35 < M = 15{,}5$.

7.3 MERA SPLOŠČENOSTI

Za analizo sploščenosti uporabljamo koeficient sploščenosti:

$$KS = 1{,}9 \times \frac{Q_3 - Q_1}{D_9 - D_1} \tag{7.17}$$

Na osnovi izračunanega koeficienta sklepamo o sploščenosti porazdelitve. Če velja:

- **$KS > 1$, je porazdelitev sploščena;**
- **$KS < 1$, je porazdelitev koničasta;**
- **$KS = 1$, pri normalni porazdelitvi.**

Izračunajmo koeficient sploščenosti za porazdelitev vrednosti prodaje 80 prodajaln.

V točki 7.1.2 smo izračunali:

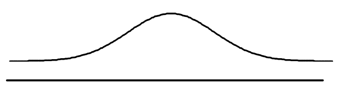
$Q_1 = 13{,}44$ mio EUR
$Q_3 = 17{,}47$ mio EUR
$D_1 = 11{,}89$ mio EUR
$D_9 = 19{,}5$ mio EUR

Izračunajmo koeficient sploščenosti: $KS = 1{,}9 \times \frac{Q_3 - Q_1}{D_9 - D_1} = 1{,}9 \times \frac{17{,}47 - 13{,}44}{19{,}5 - 11{,}89} = 1{,}006$

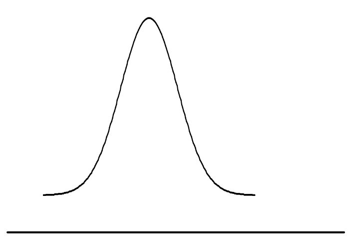
$KS = 1{,}006 > 1$, kar pomeni, da je frekvenčna porazdelitev zelo rahlo sploščena.

7.4 OBLIKE FREKVENČNIH PORAZDELITEV

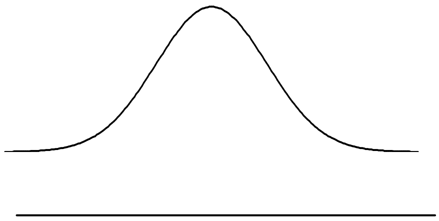
Na slikah A, B, C, D, E, in F je prikazanih nekaj tipičnih frekvenčnih porazdelitev.



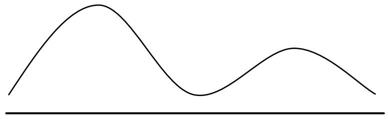
A) sploščena



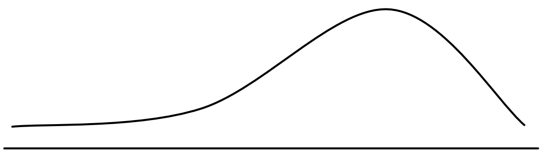
B) koničasta



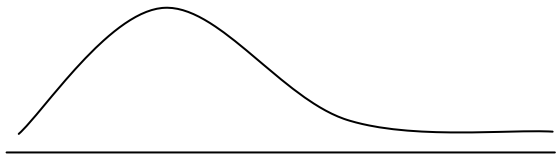
C) unimodalna – simetrična



D) bimodalna



E) asimetrična v levo



F) asimetrična v desno

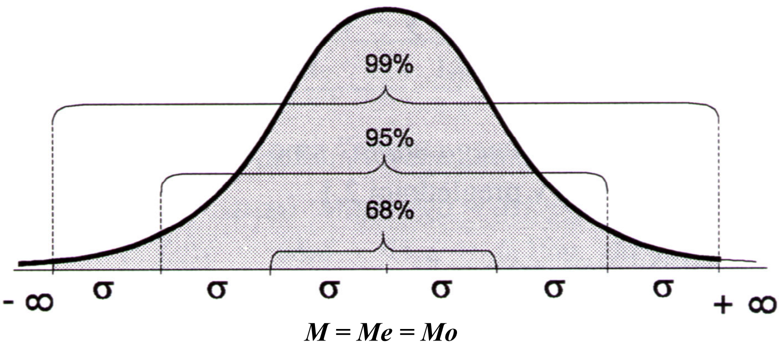
7.5 LASTNOSTI NORMALNE PORAZDELITVE

Porazdelitve, ki jih oblikujemo na osnovi empiričnih podatkov, imenujemo stvarne ali empirične porazdelitve. V nasprotju s temi poznamo porazdelitve, ki jih tvorimo na osnovi teoretičnih predpostavk. Take porazdelitve so teoretične porazdelitve. Za te že vnaprej poznamo nekatere statistične parametre, kot so aritmetična sredina, modus, mediana in standardni odklon. Od teoretičnih porazdelitev bomo opredelili lastnosti normalne porazdelitve.

Normalno porazdelitev (Gaussovo porazdelitev) lahko na osnovi parametrov, ki smo jih spoznali v prejšnjih poglavjih, opredelimo z naslednjimi značilnostmi (Artenjak, 1997, 111):

- **V razmiku:**
  - $M - \sigma$ ***do*** $M + \sigma$ **se nahaja 68,3 % vseh vrednosti spremenljivke;**
  - $M - 2\sigma$ ***do*** $M + 2\sigma$ **se nahaja 95,4 % vseh vrednosti spremenljivke;**
  - $M - 3\sigma$ ***do*** $M + 3\sigma$ **se nahaja 99,7 % vseh vrednosti spremenljivke.**
- $M = Me = Mo$
- $KA_{Mo}, KA_{Me} = 0$
- $KS = 1$
- $VR = y_{max} - y_{min} \cong 6\sigma$

Grafični prikaz normalne porazdelitve:



Slika 7.1: **Normalna porazdelitev spremenljivke** ***y***

Podobnost določene preučevane porazdelitve z normalno po prvi lastnosti ugotavljamo po naslednjem postopku (Artenjak,1997, 111**):**

- **izračunamo aritmetično sredino in standardni odklon (*M in σ*);**
- **določimo vrednosti:** $y_1 = M - k \times \sigma$ **in** $y_2 = M + k \times \sigma$ (*k* označuje število standardnih odklonov);
- **za določeni vrednosti izračunamo kvantilna ranga**, torej $P(y_1)$ in $P(y_2)$;
- **izračunamo razliko med kvantilnima rangoma** $P = P(y_2) - P(y_1)$ **in pomnožimo s 100, tako dobimo odstotek enot, ki imajo vrednosti v določenem razmiku;**
- **izračunani odstotek primerjamo s teoretično vrednostjo in sklepamo o podobnosti empirične z normalno porazdelitvijo.**

Postopek izračuna opravimo za porazdelitev vrednosti prodaje za 80 prodajaln trgovskega podjetja ***Preskrba*** in preverimo, ali je porazdelitev podobna teoretični normalni porazdelitvi, torej izračunajmo ali ima 68,3 % prodajaln vrednost prodaje v razmiku $M - \sigma$ do $M + \sigma$.

Tabela 7.4: **Frekvenčna porazdelitev vrednosti prodaje za 80 prodajaln trgovskega podjetja *Preskrba* v letu 2007 s kumulativo**

Vrednost prodaje v mio EUR	**Št. prodajaln** f_j	**Kumulativa frekvenc** F_j
od 10 do pod 12	9	9
od 12 do pod 14	16	25
od 14 do pod 16	23	48
od 16 do pod 18	17	65
od 18 do pod 20	10	75
od 20 do pod 22	3	78
od 22 do pod 24	2	80
Skupaj	**80**	

Aritmetično sredino in standardni odklon imamo že izračunana:

- $M = 15,5$ mio EUR in $\sigma = 2,8$ mio EUR

- izračunamo vrednosti y_1 in y_2:

$y_1 = M - \sigma = 15,5 - 2,8 = 12,7$ *(j = 2)*

$P_{y=12,7} = ?$ $\quad P = \dfrac{R_y - 0,5}{N}$

$$R = F_{j-1} + f_j \frac{y - y_{j,\min}}{d_j}$$

$$R_{y=12,7} = 9 + 16 \times \frac{12,7 - 12}{2} = 14,6$$

in $P_{y=12,7} = \dfrac{14,6 - 0,5}{80} = 0,1763$

$y_2 = M + \sigma = 15,5 + 2,8 = 18,3$ *(j = 5)*

$P_{y=18,3} = ?$ $\quad P = \dfrac{R - 0,5}{N}$

$$R = F_{j-1} + f_j \frac{y - y_{j,\min}}{d_j}$$

$$R_{y=18,3} = 65 + 10 \times \frac{18,3 - 18}{2} = 66,5$$

in $P_{y=18,3} = \dfrac{66,5 - 0,5}{80} = 0,825$

- ***(P**$_{y=18,3}$ **– P**$_{y=12,7}$ **) × 100 = (0,825 – 0,1763) × 100 = 64,87 %***

V razmiku $M - \sigma$ do $M + \sigma$ se v teoretični porazdelitvi nahaja 68,3 % vseh vrednosti, v porazdelitvi vrednosti prodaje za 80 prodajaln pa 64,87 %, iz česar sklepamo, da po tej lastnosti ni podobna teoretični normalni porazdelitvi.

7.6 PRIMER ZA UTRJEVANJE

Izračunajmo vse parametre, ki smo jih spoznali v poglavjih o kvantilih, srednjih vrednostih ter merah variabilnosti, asimetrije in sploščenosti, za frekvenčno porazdelitev površin prodajnega prostora prodajaln. Izračunane parametre primerjajmo s parametri, ki smo jih izračunali za frekvenčno porazdelitev vrednosti prodaje teh prodajaln.

Tabela 7.5: **Frekvenčna porazdelitev površine prodajnega prostora za 80 prodajaln**

Površina prodajnega prostora v m^2	**Št. prodajaln f_j**	**Kumulativa fr. F_j**	**$f_j y_j$**
od 40 do pod 60	6	6	300
od 60 do pod 80	9	15	630
od 80 do pod 100	10	25	900
od 100 do pod 120	15	40	1.650
od 120 do pod 140	18	58	2.340
od 140 do pod 160	12	70	1.800
od 160 do pod 180	7	77	1.190
od 180 do pod 200	3	80	570
Skupaj	**80**		**9.380**

Izračunajmo najprej srednje vrednosti:

- **aritmetično sredino *(M):***

$$M = \frac{1}{N}\sum_{j=1}^{k} y_j f_j = \frac{1}{80} \times 9.380 = 117,25 \text{ m}^2$$

- **modus – najpogostejšo površino *(Mo)*:**

$$Mo = y_{j,\min} + d_j \times \frac{f_j - f_{j-1}}{2\times f_j - f_{j-1} - f_{j+1}} =$$

$$= 120 + 20 \times \frac{18-15}{2\times 18 - 15 - 12} = 126,67 \text{ m}^2$$

- **mediano** – tisto vrednost, od katere je polovica enot imela manjšo ali kvečjemu enako, polovica enot pa večjo vrednost *(Me):*

$$P = 0,50$$

$$R_p = 80 \times 0,50 + 0,5 = 40,5$$

$$Me = y_{j,\min} + d_j \times \frac{R - F_{j-1}}{f_j} = 120 + 20 \times \frac{40,5-40}{18} = 120,56 \text{ m}^2$$

Povprečna površina prodajaln je 117,25 m^2, najpogostejša površina teh prodajaln je 126,67 m^2. Polovica prodajaln je imela manjšo ali kvečjemu enako, polovica pa večjo površino kot 120,56 m^2.

Tabela 7.6: **Frekvenčna porazdelitev površine prodajnega prostora za 80 prodajaln**

Površina v m^2	f_j	y_j	$f_j\ y_j$	$y_j - M$	$f_j\ (y_j - M)^2$
od 40 do pod 60	6	50	15.000	–67,25	27.135,38
od 60 do pod 80	9	70	44.100	–47,25	20.093,06
od 80 do pod 100	10	90	81.000	–27,25	7.425,625
od 100 do pod 120	15	110	181.500	–7,25	788,437
od 120 do pod 140	18	130	304.200	12,75	2.926,125
od 140 do pod 160	12	150	270.000	32,75	12.870,75
od 160 do pod 180	7	170	202.300	52,75	19.477,94
od 180 do pod 200	3	190	108.300	72,75	15.877,69
Skupaj	**80**		**1.206.400**		**106.595,00**

- **Izračunajmo varianco in standardni odklon:**

$$VAR = \sigma^2 = \frac{1}{N}\sum_{j=1}^{k} f_j (y_j - M)^2 = \frac{1}{80} \times 106.595 = 1.332,44 \left(\text{m}^2\right)^2$$

$$ali: VAR = \frac{1}{N}\sum_{j=1}^{k} f_j y_j^2 - M^2 = \frac{1}{80} \times 1.206.400 - 117,25^2 = 1.332,44 \left(\text{m}^2\right)^2$$

$$VAR_{cor} = VAR - \frac{d^2}{12} = 1.332,44 - \frac{20^2}{12} = 1.299,11\left(\text{m}^2\right)^2$$

$$SD = \sigma = \sqrt{VAR} = \sqrt{1299,11} = 36,04\ \text{m}^2$$

in še koeficient variabilnosti: $KV\ \% = \frac{\sigma}{M} \times 100 = \frac{36,04}{117,25} \times 100 = 30,74\ \%$

Varianca je 1.291,11(m²)². Standardni odklon je 36,04 m². Koeficient variabilnosti je 30,74 %, kar pomeni, da standardni odklon predstavlja 30,74 % aritmetične sredine. Če variabilnost te spremenljivke primerjamo z variabilnostjo spremenljivke vrednost prodaje, ugotovimo, da je variabilnost za spremenljivko vrednost prodaje manjša, saj standardni odklon predstavlja le 18,06 % aritmetične sredine.

- **Izračunajmo še koeficient asimetrije in koeficient sploščenosti:**

Koeficient asimetrije:

$$KA_{Mo} = \frac{M - Mo}{\sigma} = \frac{117,25 - 125,67}{36,04} = -0,2613$$

$$KA_{Me} = \frac{3(M - Me)}{\sigma} = \frac{3(117,25 - 120,56}{36,04} = -0,2756$$

Porazdelitev prodajaln po površini prodajnega prostora je malo asimetrična v levo, saj sta izračunana koeficienta negativna. Asimetričnost porazdelitve je vidna tudi iz grafičnega prikaza s histogramom (slika 7.2).

Koeficient sploščenosti:

Izračunati moramo prvi in tretji kvartil ter prvi in deveti decil:

$P = 0,25\ in\ R = N \times P + 0,5 = 80 \times 0,25 + 0,5 = 20,5$; prvi kvartil je v tretjem razredu: $j = 3$

$$F_2 = 15 < R = 20,5 < F_3 = 25$$
$$y_{3,\min} = 80 < Q_1 = y_{P=0,25} < y_{3,\max} = 100$$

$$y = Q_1 = y_{j,\min} + d_j \times 0 \frac{R - F_{j-1}}{f_j} = 80 + 20 \times \frac{20,5 - 15}{10} = 91\,\text{m}^2$$

$P = 0,75\ in\ R = N \times P + 0,5 = 80 \times 0,75 + 0,5 = 60,5$; tretji kvartil je v šestem razredu: $j = 6$

$$Q_3 = 140 + 20 \times \frac{60,5 - 58}{12} = 144,17\,\text{m}^2$$

$P = 0,10$ in $R = N \times P + 0,5 = 80 \times 0,10 + 0,5 = 8,5$; prvi decil je v drugem razredu: $j = 2$

$F_1 = 6 < R = 8,5 < F_2 = 15$

$y_{2,\min} = 60 < D_1 = y_{P=0,10} < y_{2,\max} = 80$

$$D_1 = 60 + 20 \times \frac{8,5-6}{9} = 65,56 \text{ m}^2$$

$P = 0,90$ in $R = N \times P + 0,5 = 80 \times 0,90 + 0,5 = 72,5;$ deveti decil je v sedmem razredu: $j = 7$

$$D_9 = 160 + 20 \times \frac{72,5-70}{7} = 167,14 \text{ m}^2$$

$$KS = 1,9 \times \frac{Q_3 - Q_1}{D_9 - D_1} = 1,9 \times \frac{144,17-91}{167,14-65,56} = 0,995$$

Izračunani koeficient sploščenosti je manjši od ena, torej je porazdelitev bolj koničasta.

- **Izračunajmo še delež prodajaln v razmiku $M-\sigma$ do $M+\sigma$.in ga primerjajmo z deležem, ki velja za teoretično normalno porazdelitev:**

$y_1 = M - \sigma = 117,25 - 36,04 = 81,21$ *(j = 3)* $y_2 = M + \sigma = 117,25 + 36,04 = 153,29$ *(j = 6)*

$P_{y=81,21} = ?$ $P = \dfrac{R_y - 0,5}{N}$

$$R = F_{j-1} + f_j \frac{y - y_{j,\min}}{d_j}$$

$$R_{y=81,21} = 15 + 10 \times \frac{81,21-80}{20} = 15,605$$

in $P_{y=81,21} = \dfrac{15,605-0,5}{80} = 0,1888$

$P_{y=153,29} = ?$ $P = \dfrac{R_y - 0,5}{N}$

$$R = F_{j-1} + f_j \frac{y - y_{j,\min}}{d_j}$$

$$R_{y=153,29} = 58 + 12 \times \frac{153,29-140}{20} = 65,974$$

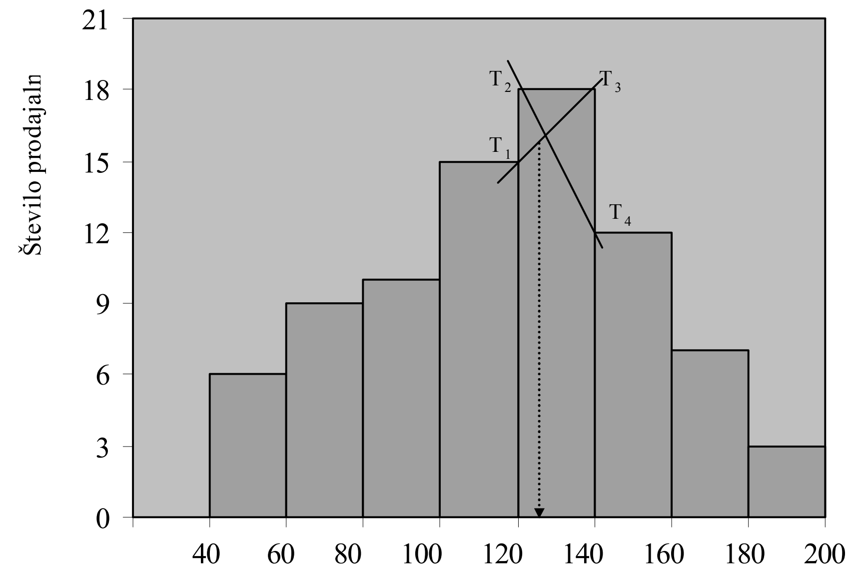
in $P_{y=153,29} = \dfrac{65,974-0,5}{80} = 0,8184$

$(P_{y=153,29} - P_{y=81,21}) \times 100 = (0,8184 - 0,1888) \times 100 = 62,96\ \%$

V razmiku $M-\sigma$ do $M+\sigma$ se v teoretični normalni porazdelitvi nahaja 68,3 % vseh enot, v porazdelitvi površine prodajnega prostora 80 prodajaln pa 62,96 %, iz česar sklepamo, da ni podobna teoretični normalni porazdelitvi.

Na sliki 7.2 je histogram frekvenčne porazdelitve z oceno modusa, na sliki 7.3 pa kumulativa frekvenčne porazdelitve z oceno razmika $M-\sigma$ **do** $M+\sigma$**.**

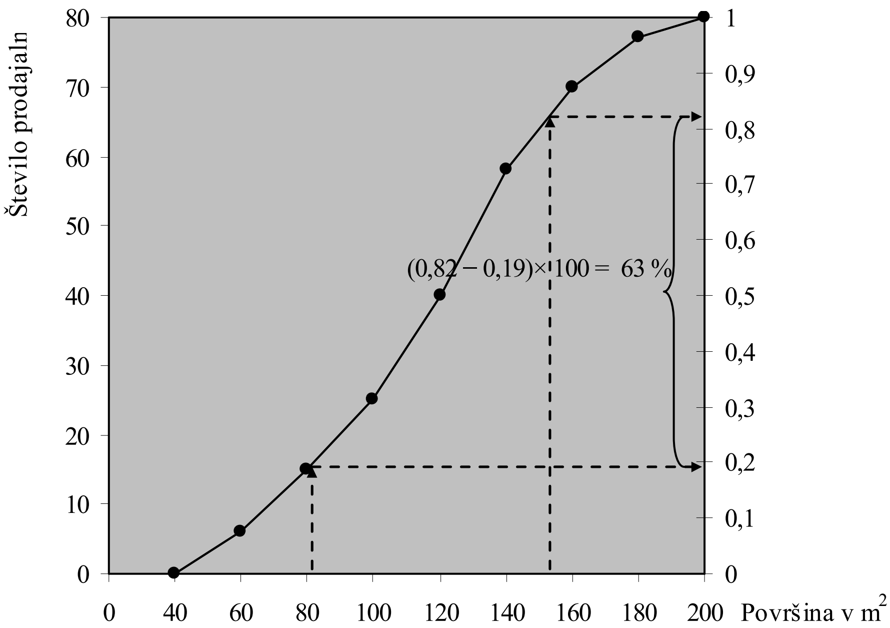
Na primeru 7.6 smo ponovili snov 6. in 7. poglavja ter dela 5. poglavja. Tako ste ugotovili, kako pri statistični analizi izračunavamo različne statistične parametre, da ugotovimo značilnosti preučevanega pojava. Gotovo pa ste se prepričali tudi o praktični uporabi obravnavanih statističnih metod. Za boljšo *nazornost izračunane parametre dopolnimo še z grafičnimi prikazi. Reševanje nalog od 6.1 do 6.11 v Zbirki vaj iz statistike bo za utrjevanje znanja zelo koristno.*



Slika 7.2: **Frekvenčna porazdelitev površine prodajnega prostora za 80 prodajaln trgovskega podjetja *Preskrba***

Vir: Tabela 7.6

Delež vrednosti v razmiku $M-\sigma$ do $M+\sigma$ ocenimo grafično v sliki 7.3.



Slika 7.3: **Kumulativa frekvenčne porazdelitve površine prodajnega prostora za 80 prodajaln**

Vir: Tabela 7.6

8 ANALIZA ČASOVNIH VRST

Spreminjanje pojavov v času smo spremljali s kazalci dinamike (poglavje 3), kot so indeksi, koeficienti in stopnje rasti. V tem poglavju bomo analizo časovnih vrst dopolnili, saj z analizo spreminjanja lastnosti v preteklosti lahko napovedujemo njihov razvoj v prihodnosti. Spoznali bomo trend in sezonske indekse, predvsem pa njihov pomen pri napovedovanju gospodarskih in drugih pojavov, ki se spreminjajo v odvisnosti od časa.

Ekonomski in drugi pojavi se običajno spreminjajo s časom, tako se na primer družbeni proizvod na prebivalca iz leta v leto povečuje, povečuje se tudi poraba energije na prebivalca, kot rezultat boljše zdravstvene oskrbe prebivalstva, napredka medicine in boljših življenjskih pogojev pa se zmanjšuje umrljivost prebivalstva v razvitih državah. Spremembe so rezultat različnih dejavnikov, ki vplivajo na pojav, in so opazne, če za določen pojav zberemo in uredimo podatke za zaporedne časovne trenutke ali intervale. Tako dobimo časovno vrsto, ki prikazuje spremembe pojava v odvisnosti od časa, prikazuje torej dinamiko pojava. Preučevanje te pa je zelo pomembno, saj nam poznavanje spreminjanja pojava v preteklosti omogoča napovedovanje razvoja v prihodnosti. To seveda ni povsem zanesljivo, saj so spremembe rezultat različnih dejavnikov v preteklosti in ni povsem gotovo, da bodo tudi v prihodnosti delovali isti dejavniki v enaki meri. Zato je napovedovanje pogosto tvegano, predvsem za bolj oddaljeno prihodnost.

Spreminjanje pojavov v času preučujemo z zelo različnimi metodami. Za enostavno analizo časovne vrste uporabljamo enostavne kazalce dinamike, kot so indeksi s stalno osnovo, verižni indeksi, koeficienti rasti, stopnje rasti in še povprečni koeficient rasti in povprečna stopnja rasti; vse smo že spoznali. Grobo predstavo o razvoju časovne vrste dobimo z njenim grafičnim prikazom, ta je največkrat linijski grafikon. Grafični prikaz v veliki meri vpliva na izbor nadaljnjih metod analize. Nekatere od teh, predvsem tiste, ki temelje na razčlenitvi časovne vrste na njene sestavine, bomo spoznali v tem poglavju.

8.1 DEJAVNIKI SPREMINJANJA POJAVOV IN SESTAVINE ČASOVNIH VRST

Analiza časovnih vrst temelji na ugotavljanju dejavnikov, ki povzročajo različna kratkoročna ali dolgoročna nihanja, in poskuša ovrednotiti njihov pomen. V ta namen časovno vrsto razčlenimo na naslednje sestavine:

- ***trend***, ki prikazuje osnovno smer razvoja in ga opazimo v daljšem časovnem obdobju, npr. 10 ali več let. Pojavi, ki nimajo trenda, so redki. Večina pojavov ima naraščajoči trend, npr. število prihodov turistov v Slovenijo, bruto družbeni produkt na prebivalca; nekateri pojavi imajo padajoči trend, npr. število brezposelnih oseb v Sloveniji;

- ***periodična nihanja***, ki se pojavljajo na krajša, enako dolga časovna obdobja, npr. vsak dan, mesec ali leto. To so npr. večja poraba vode v gospodinjstvih po 15. uri, večji promet na slovenskih cestah ob koncu tedna, večje število turistov ob morju v poletnih mesecih. Nihanja, ki se pojavljajo vsako leto, imenujemo sezonska nihanja;

- ***ciklična nihanja,*** to so dolgoročna nihanja okoli trenda, ki so bolj ali manj pravilna, vendar njihova dolžina in oblika nista stalni kot pri periodičnih nihanjih. Značilna so za ekonomske pojave. Tako obdobjem recesije sledijo obdobja gospodarske prosperitete;

- ***iregularna nihanja***, ki so:
  - *slučajna*, povzročajo le manjše odklone od osnovne smeri razvoja, čeprav so vedno prisotna in nastanejo iz neznanih in nepomembnih vzrokov;
  - *nihanja kot posledica posebnih, enkratnih dejavnikov*, kot so naravne in druge katastrofe.

8.2 DOLOČANJE TRENDA

Trend je smer razvoja, ki je opazna na daljša časovna obdobja iz osnovnih podatkov časovne vrste, še bolj pa iz njenega grafičnega prikaza. Ta je tudi osnova za določanje trenda. Od metod za določanje trenda bomo spoznali le dve:

- prostoročno metodo in
- metodo najmanjših kvadratov

8.2.1 Prostoročna metoda za določanje trenda

Prostoročna metoda za določanje trenda je najenostavnejša. Osnova za določanje trenda je grafični prikaz časovne vrste z linijskim grafikonom. Nato vrišemo linijo, ki se časovni vrsti dobro prilega in poteka med točkami, ki se nanašajo na podatke o pojavu za posamezne časovne razmike ali trenutke. Trend v vsakem primeru poteka med dejanskimi vrednostmi, tako da se osnovna časovna vrsta od trenda odklanja navzgor in navzdol. Linija je lahko premica ali krivulja, ki jo vrišemo po lastni presoji, zato je tako določanje trenda subjektivno in manj točno, kot je določanje trenda z analitičnimi metodami.

Prostoročno določanje trenda je tudi osnova za izbiro ustrezne krivulje trenda in njej ustrezne enačbe trenda, katere parametre določimo z analitičnimi metodami.

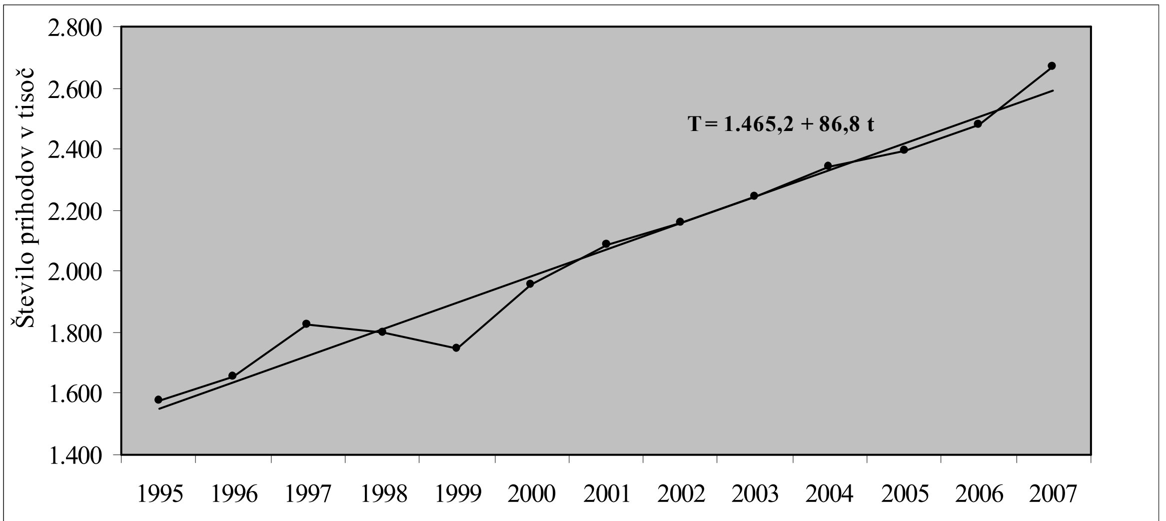
Primer 8.1: Vzemimo primer prihodov turistov v Sloveniji v letih od 1995 do 2007 (tabela 8.1).

Iz grafičnega prikaza (slika 8.1) je še bolj kot iz osnovne časovne vrste opazno, da število prihodov turistov iz leta v leto narašča. To naraščanje sicer ni enakomerno, saj se v nekaterih letih število prihodov celo zmanjša, kljub temu pa je osnovna časovna vrsta za opazovano obdobje naraščajoča. Trend, ki ga ponazarja vrisana premica, je naraščajoč.

Tabela 8.1: **Prihodi turistov v Slovenijo v letih od 1995 do 2007**

Leto	**Število prihodov turistov v 1.000**
1995	1.577
1996	1.658
1997	1.823
1998	1.799
1999	1.750
2000	1.957
2001	2.086
2002	2.162
2003	2.246
2004	2.341
2005	2.395
2006	2.482
2007	2.672

Vir: Statistični letopis 2005, Pomembnejši statistični podatki o Sloveniji, letnik III, št. 2/2008



Slika 8.1: **Prihodi turistov v Slovenijo v letih od 1995 do 2007**

Vir: Tabela 8.1

8.2.2 Analitična metoda določanja trenda

Med analitičnimi metodami, s katerimi določamo enačbo trenda, največkrat uporabljamo metodo najmanjših kvadratov. Pri tej metodi je postavljena zahteva, da je vsota kvadratov odklonov izračunanih vrednosti trenda T_t od vrednosti časovne vrste Y_t minimalna:

$$\sum_{t=1}^{N}\left(Y_t - T_t\right)^2 = min \qquad (8.1)$$

Najvažnejša je izbira pravilnega tipa funkcije, ki naj predstavlja trend, saj je ena od zahtev te metode, da se trend kot smer razvoja čim bolje prilega časovni vrsti. Po sliki časovne vrste, poznavanju pojava, ki ga vrsta prikazuje, in lastnosti funkcij, ki pridejo v poštev, izberemo najustreznejši tip funkcije. Najenostavnejša, s katero pa v mnogih primerih zadovoljivo opišemo smer razvoja, je premica. Premica opisuje linearni trend. Poleg tega bomo spoznali še parabolični trend.

8.2.2.1 Linearni trend

Parametre linearnega trenda bomo izračunali za časovno vrsto prihodov turistov.

Enačba premice*:* ***y = a + bx***

Enačba linearnega trenda: ***T = a + bt,*** *(8.2)*
kjer je:
a – konstanta
b – smerni koeficient
t – čas

Namesto letnic ali drugih oznak časa uporabimo zaporedne številke meritev *t* =1,2,3….

Parametra a in b izračunamo iz sistema normalnih enačb:

$$\sum_{t=1}^{N} Y_t = aN + b\sum_{t=1}^{N} t_t$$
$$\sum_{t=1}^{N} Y_t t_t = a\sum_{t=1}^{N} t_t + b\sum_{t=1}^{N} t_t^2 \qquad (8.3)$$

$$26.948 = 13a + 91b / \times(-7) \qquad -188.636 = -91a - 637b$$
$$204.437 = 91a + 819b \qquad 204.437 = 91a + 819b$$
$$15.801 = 182b$$
$$b = 86,8$$

in a:
$$26.948 = 13a + 91 \times 86,8$$
$$a = 1.465,2$$

Z rešitvijo enačb dobimo: ***a = 1.465,2***
b = 86,8

Kako dobimo zmnožke in seštevke le-teh za izračun parametrov po zgoraj navedenih enačbah, je razvidno iz tabele 8.2.

Zapišimo enačbo linearnega trenda: $T = 1.465,2 + 86,8\,t$ in izračunajmo vrednosti trenda:

$$T_{t=1} = 1.465,2 + 86,8 \times 1 = 1.552,0$$
$$T_{t=2} = 1.465,2 + 86,8 \times 2 = 1.638,8 \text{ itd.}$$

Izračunane vrednosti trenda so razvidne iz tabele 8.2. Vidimo, da se ne razlikujejo veliko od dejanskih vrednosti.

Tabela 8.2: **Prihodi turistov v Slovenijo v letih od 1995 do 2007, podatki za izračun parametrov trenda in vrednosti trenda**

Leto	**Prihodi v 1000 Y_t**	**Čas t_t**	$Y_t t_t$	t_t^2	T_t
1995	1.577	1	1.577	1	1.552
1996	1.658	2	3.316	4	1.639
1997	1.823	3	5.469	9	1.726
1998	1.799	4	7.196	16	1.812
1999	1.750	5	8.750	25	1.899
2000	1.957	6	11.742	36	1.986
2001	2.086	7	14.602	49	2073
2002	2.162	8	17.296	64	2.160
2003	2.246	9	20.214	81	2.246
2004	2.341	10	23.410	100	2.333
2005	2.395	11	26.345	121	2.420
2006	2.482	12	29.784	144	2.507
2007	2.672	13	34.736	169	2.594
	26.948	**91**	**204.437**	**819**	

Parametra a in b lahko izračunamo tudi neposredno s splošno rešitvijo sistema enačb:

$$b = \frac{\frac{1}{N}\sum_{t=1}^{N} Y_t t_t - \overline{Y}\,\overline{t}}{\frac{1}{N}\sum_{t=1}^{N} t_t^2 - \overline{t}^2} \qquad (8.4)$$

$$a = \overline{Y} - b\overline{t} \qquad (8.5)$$

$$\overline{t} = \frac{91}{13} = 7 \qquad \overline{Y} = \frac{26.948}{13} = 2.072,9$$

$$b = \frac{\frac{1}{13} \times 204.437 - 7 \times 2.072,9}{\frac{1}{13} 819 - 7^2} = 86,8$$

$$a = 2.072,9 - 7 \times 86,8 = 1.465,2$$

$$\underline{T = 1.465,2 + 86,8\, t}$$

Napovedovanje na osnovi linearnega trenda

Na osnovi trendne enačbe, katere parametra smo izračunali iz sistema normalnih enačb, lahko napovedujemo (predvidimo) razvoj časovne vrste v prihodnosti. Spremenljivko *t,* torej čas, določimo tako, da nadaljujemo s številom, ki sledi. V primeru časovne vrste prihodov turistov v Slovenijo velja za leto 2008 t = 14, za leto 2009 t = 15 itd.

Na osnovi enačbe trenda predvidimo število prihodov turistov v letu 2010:

$$\boldsymbol{T_{t=16} = 1.465,2\ tisoč + 86,8 \times 16 = 2.854\ tisoč}$$

Kako zanesljivo je napovedovanje?

Na kakovost napovedovanja vpliva izbira ustrezne funkcije za izračun parametrov enačbe trenda in način, kako se krivulja izbrane funkcije prilega časovni vrsti. Manjši kot so odkloni osnovne časovne vrste od trenda, zanesljivejše je napovedovanje. Očitno je tudi, da so kratkoročna napovedovanja pojava boljša kot dolgoročna. Verjetnost, da se dejavniki, ki vplivajo na pojav, spremenijo, je v kratkem obdobju manjša kot v dolgem obdobju.

Primerjava napovedi o številu prihodov turistov na osnovi trenda in povprečnega koeficienta rasti

Napovedujemo lahko tudi s povprečnim koeficientom rasti, kar smo spoznali pri srednjih vrednostih. Ocenimo število prihodov turistov za leto 2010 na osnovi povprečnega koeficienta rasti, ki ga izračunamo za obdobje 1995 do 2007:

$\overline{K} = \sqrt[12]{\frac{2.672}{1.577}} = 1,045$, kar pomeni, da je število prihodov turistov v Slovenijo v tem obdobju naraščalo povprečno letno za 4,5 odstotka.

Ocena za leto 2010:

$$\boldsymbol{Y_{2010} = Y_{2007} \times \overline{K}^3 = 2.672 \times 1,045^3 = 3.049\ tisoč}$$

Po tej oceni bo prišlo leta 2010 v Slovenijo 3.049 tisoč turistov.

8.2.2.2 Parabolični trend

V primerih, ko se časovni vrsti prilega krivulja, katere funkcija je parabola 2. stopnje, je enačba trenda:

$$\boldsymbol{T = a + bt + ct^2} \tag{8.6}$$

Parametre a, b in c izračunamo iz sistema normalnih enačb:

$$\begin{aligned} \sum_{t=1}^{N} Y_t &= aN + b\sum_{t=1}^{N} t_t + c\sum_{t=1}^{N} t_t^2 \\ \sum_{t=1}^{N} Y_t t_t &= a\sum_{t=1}^{N} t_t + b\sum_{t=1}^{N} t_t^2 + c\sum_{t=1}^{N} t_t^3 \\ \sum_{t=1}^{N} Y_t t_t^2 &= a\sum_{t=1}^{N} t_t^2 + b\sum_{t=1}^{N} t_t^3 + c\sum_{t=1}^{N} t_t^4 \end{aligned} \tag{8.7}$$

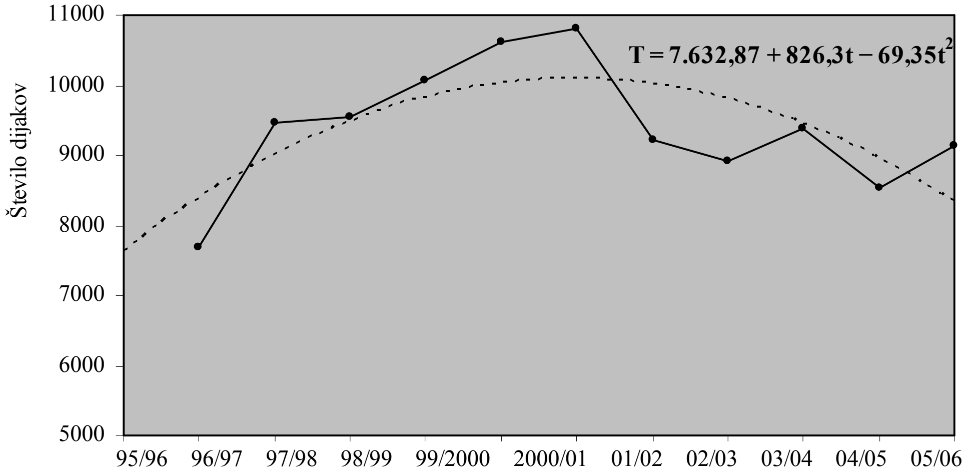
Parametre paraboličnega trenda bomo izračunali za časovno vrsto dijakov, ki so zaključili srednješolsko poklicno in strokovno izobraževanje v Sloveniji v šolskih letih 1995/96 do 2005/06. Podatki so v tabeli 8.3.

Tabela 8.3: **Dijaki, ki so zaključili srednješolsko poklicno in strokovno izobraževanje v Sloveniji v šolskih letih od 1995/96 do 2005/06**

Leto	Število dijakov (konec šolskega leta)
1995/96	7.677
1996/97	9.465
1997/98	9.561
1998/99	10.066
1999/2000	10.627
2000/01	10.801
2001/02	9.231
2002/03	8.922
2003/04	9.388
2004/05	8.538
2005/06	9.129

Vir: Statistični letopis 2007, Statistične informacije

Časovno vrsto najprej prikažemo grafično in vrišemo trend.



Slika 8.2: **Dijaki, ki so zaključili srednješolsko izobraževanje tehniških in strokovnih programov v šolskih letih od 1995/96 do 2005/06 v Sloveniji**

Iz podatkov v tabeli 8.3, še bolje pa iz slike 8.2 lahko ugotovimo, da je število dijakov, ki so zaključili srednješolsko poklicno in strokovno izobraževanje, do konca šolskega leta 2000/01 naraščalo, po tem letu pa se je zmanjševalo oziroma nekoliko poraslo v šolskih letih 2003/04 in 2005/06. Splošna smer pa je vendarle padajoča in jo najbolje ponazorimo s parabolo druge stopnje.

V tabeli 8. 4 so podatki za izračun paraboličnega trenda.

Tabela 8.4: **Dijaki, ki so zaključili srednješolsko poklicno in strokovno izobraževanje v Sloveniji v šolskih letih od 1995/96 do 2005/06, podatki za izračun parametrov funkcije trenda in vrednosti trenda**

Šolsko leto	Število dijakov Y_t	t_t	t_t^2	t_t^3	t_t^4	$Y_t t_t$	$Y_t t_t^2$	T_t
94/95	7.677	1	1	1	1	7.677	7.677	8.389,7
95/96	9.465	2	4	8	16	18.930	37.860	9.008,0
96/97	9.561	3	9	27	81	28.683	86.049	9.487,5
97/98	10.066	4	16	64	256	40.264	161.056	9.828,4
98/00	10.627	5	25	125	625	53.135	265.675	10.030,5
00/01	10.801	6	36	216	1.296	64.806	388.836	10.094,0
01/02	9.231	7	49	343	2.401	64.617	452.319	10.018,7
02/03	8.922	8	64	512	4.096	71.376	571.008	9.804,8
03/04	9.388	9	81	729	6.561	84.492	760.428	9.452,1
04/05	8.538	10	100	1.000	10.000	85.380	853.800	8.960,8
05/06	9.129	11	121	1.331	14.641	100.419	1.104.609	8.330,7
	103.405	66	506	4.356	39.974	619.779	4.689.317	

$$103.405 = 11a + 66b + 506c$$
$$619.779 = 66a + 506b + 4.356c$$
$$4.689.317 = 506a + 4.356b + 39.974c$$

(Navodilo: Najprej rešite prvi dve enačbi,in sicer tako da prvo množite z –6, nato pa prvo in tretjo, tako da prvo množite z –46). Ostaneta dve enačbi:

$$-651 = 110b + 1.320c$$
$$-67.313 = 1.320b + 16.698c$$

(Prvo enačbo pomnožite z –12 in z rešitvijo izračunamo parameter c = –69,35).

Z rešitvijo enačb dobimo: $\boldsymbol{a = 7.632,87}$, $\boldsymbol{b = 826,26}$, $\boldsymbol{c = -69,35}$

Zapišimo enačbo trenda: $\boldsymbol{T_t = 7.632,87 + 826,26\,t - 69,35\,t^2}$

Z vstavljanjem vrednosti za spremenljivko *t* izračunamo vrednosti trenda. V tem primeru so razlike med dejanskimi vrednostmi in trendom večje, saj je že iz grafikona opazno, da se

krivulja trenda osnovni časovni vrsti najbolje ne prilega. Napovedovanje je v tem primeru precej tvegano.

Predvidimo število dijakov, ki bodo zaključili srednješolsko poklicno in strokovno izobraževanje v Sloveniji konec šolskega leta 2009/10:

$$\boldsymbol{T_{2009/10(t=15)} = 7.632,9 + 826,26 \times 15 - (69,35 \times 15^2) = 5.608}$$

8.3 DOLOČANJE SEZONSKE SESTAVINE V ČASOVNIH VRSTAH

Za številne ekonomske pojave so značilna nihanja, ki se ponavljajo na krajša, enako dolga časovna obdobja. Tako se poraba vode in električne energije ponavlja po urah v dnevu, cestni promet po dnevih v tednu, prenočitve gostov, prodaja piva, poraba kurilnega olja po mesecih. Ta nihanja imenujemo periodična nihanja. Periodična nihanja, ki se ponavljajo na leto, imenujemo sezonska nihanja. Ta so značilna predvsem za pojave v turizmu; število turistov in prenočitev se poveča v poletnih mesecih.

Časovno obdobje, v katerem se nihanje ponavlja, imenujemo perioda. Ta je razdeljena na več krajših obdobij. Tako je perioda eno leto, če analiziramo sezonska nihanja. Perioda pa je razdeljena na več krajših obdobij, kot so trimesečja ali meseci. Pri tem lahko uporabimo naslednje oznake:

t (t=1,2,3....N) – perioda (npr. leto)
N – število vseh opazovanih period

p (p=1,2,3...P) – obdobje znotraj periode (npr. trimesečje)
P – število obdobij znotraj periode

Če opazujemo prenočitve domačih turistov v Sloveniji po trimesečjih za tri leta, je število obdobij znotraj period 4 (leto ima 4 trimesečja), število period je 3, saj opazujemo pojav za 3 leta.

Postopek določanja sezonske sestavine je odvisen od značilnosti časovne vrste oziroma od tega, kakšni dejavniki vplivajo na opazovani pojav. Če je perioda zelo kratko časovno obdobje, npr. dan, teden, v taki časovni vrsti ne pridejo do izraza trend in drugi dolgoročni dejavniki. So pa ti prisotni, če je perioda daljše časovno obdobje, npr. leto, opazovani podatki pa se nanašajo na mesece ali trimesečja.

Najbolj preprosta metoda za analizo periodične sestavine je ***metoda vsot,*** ki predpostavlja, da so v časovni vrsti poleg periodičnih nihanj prisotna le še slučajna nihanja.

8.3.1 Postopek za izračun sezonskega indeksa

Y_{tp} je oznaka za splošni člen časovne vrste, kjer t pomeni časovno obdobje znotraj periode p; v časovni vrsti so poleg periodičnih nihanj $\left(Y_p\right)$ praviloma še slučajna nihanja $\left(E_t\right)$. Model take časovne vrste zapišemo: $\boldsymbol{Y_{tp} = Y_p + E_t}$

- Iz opazovanih vrednosti po obdobjih *(t)* znotraj periode *(p),* torej Y_{tp}, **izračunamo vsote**

$$S_p: \sum_{t=1}^{N} \boldsymbol{Y_{tp} = S_p} \tag{8.8}$$

  Z izračunom vsot po obdobjih se učinki slučajnih nihanj izravnajo.

- Izračunamo še **povprečje vsot.** Vsote S_p seštejemo in delimo s številom obdobij znotraj periode:

$$\overline{S} = \frac{\sum_{p=1}^{P} S_p}{P} \tag{8.9}$$

- **Periodični indeks** izračunamo tako, da vsote za obdobja znotraj periode delimo s povprečjem vsot in količnik pomnožimo s 100: $I_p = \frac{S_p}{\overline{S}} \times 100$ *(8. 10)*

  $I_p > 100$ periodični indeks je večji od 100, če so učinki periodičnih nihanj močnejši od povprečja
  $I_p < 100$ periodični indeks je manjši od 100, če so učinki periodičnih nihanj manjši od povprečja

Vsota izračunanih periodičnih indeksov mora biti enaka produktu: $100 \times P$ (P je število obdobij znotraj periode).

$$\sum_{p=1}^{P} I_p = 100 \times P \tag{8.11}$$

Primer 8.3: Analizirajmo periodično nihanje za prenočitve domačih turistov v Sloveniji po četrtletjih za tri leta, in sicer za leta 2005, 2006 in 2007.

Tabela 8.5: **Prenočitve turistov v Sloveniji po četrtletjih za leta 2005, 2006 in 2007, vsote po četrtletjih in periodični indeksi**

Leto	**Četrtletje (število v tisoč)**			
	1.	**2.**	**3.**	**4.**
2005	1.359,1	1.805,4	3.101,5	1.294,4
2006	1.404,2	1.814,3	3.114,8	1.383,6
2007	1.471,5	1.997,4	3.327,3	1.458,4
Vsote po četrtletjih S_p	4.234,8	5.617,1	9.543,6	4.136,4
Periodični indeksi I_p	**72,0**	**95,5**	**162,2**	**70,3**

Vir: Mesečni statistični pregled, št. 2006 in Pomembnejši podatki o Sloveniji, letnik II, št. 3/2007 in letnik III, št. 2/2008

Vsota četrtletnih vsot: $\sum_{p=1}^{P} S_p = 4.234,8 + 5.617,1 + 9.543,6 + 4.136,4 = 23.531,9$

Povprečje vsot: $\overline{S} = \frac{\sum_{p=1}^{P} S_p}{P} = \frac{23.531,9}{4} = 5.882,975$

Sezonski indeksi:

$I_{1.\check{c}etr.} = \frac{4.234,8}{5.882,975} \times 100 = 72,0$ $I_{3.\check{c}etr.} = \frac{9.543,6}{5.882,975} \times 100 = 162,2$

$$I_{2.četr.} = \frac{5.617,1}{5.882,975} \times 100 = 95,5 \qquad I_{3.četr.} = \frac{4.136,4}{5.882,975} \times 100 = 70,3$$

Število prenočitev domačih turistov je v prvem četrtletju 28 %, v drugem pa 4,5 % pod povprečjem, v tretjem četrtletju je 62,2 % nad povprečjem, v četrtem pa spet pod povprečjem, in sicer 29,7 %. Torej gre za tipično časovno vrsto s sezonskim značajem, saj je v poletnih mesecih (julij, avgust, september) sezonski indeks visoko nad povprečjem.

Vsota sezonskih indeksov je 400:

$$\sum_{p=1}^{P} I_p = 100 \times P = 100 \times 4 = 400 \rightarrow \sum_{p=1}^{P} I_p = 72,0 + 95,5 + 162,2 + 70,3 = 400$$

8.3.2 Napovedovanje s periodičnimi indeksi

Predpostavimo, da bo v letu 2010 10 milijonov prenočitev turistov. Na osnovi periodičnih indeksov razdelimo prenočitve po četrtletjih:

$$\text{Povprečno četrtletno število prenočitev} = \frac{10.000.000}{4} = 2.500.000$$

$$\text{1.četrtletje} = 2.500.000 \times \frac{72,0}{100} = 1.800.000$$

$$\text{2.četrtletje} = 2.500.000 \times \frac{95,5}{100} = 2.387.500$$

$$\text{3.četrtletje} = 2.500.000 \times \frac{162,2}{100} = 4.055.000$$

$$\text{4.četrtletje} = 2.500.000 \times \frac{70,3}{100} = 1.757.500$$

Ob predpostavki, da bo ostal vpliv sezonskih nihanj nespremenjen, bo v letu 2010 pri 10 milijonih prenočitev turistov v prvem četrtletju 1.800.000, v drugem 2.387.500, v tretjem 4.055.000 in v četrtem četrtletju 1.757.500 prenočitev turistov.

V tem poglavju ste spoznali, da ima časovna vrsta, za katero imamo zbrane podatke za daljše časovno obdobje, neko osnovno smer razvoja. Ta osnovno smer razvoja je trend, ki ga lahko določamo z različnimi metodami. Najenostavnejša je prostoročna, pri kateri v linijski grafikon časovne vrste vrišemo linijo, ki se ji najbolje prilega. V mnogih primerih smer razvoja zadovoljivo opiše premica. V tem primeru je enačba premice osnova za izračun parametrov trenda, ki jih lahko izračunamo iz sistema normalnih enačb ali neposredno s splošno rešitvijo le-teh.

Na osnovi enačbe trenda lahko napovedujemo razvoj časovne vrste v prihodnosti. Seveda je pri tem prisotna določena stopnja tveganja. Ta je manjša, čim manjši so odkloni časovne vrste od trenda in čim krajše je obdobje, za katero napovedujemo.

Za ponovitev in utrditev snovi tega poglavja rešujte naloge od 7.1 do 7.12. Želim vam veliko vztrajnosti.

POJMOVNIK

aritmetične sredina – srednja vrednost, ki izraža povprečje, 72–80

ciklična nihanja – dolgoročna nihanja okoli trenda, 104

decilni razmik – absolutna mera variabilnosti, ki predstavlja razliko med devetim in prvim decilom, 86–88

dogajanja – enote, ki v času nastaja in jih opazujemo v danem trenutku ali časovnem razmiku, 6

dogodki – enote, ki se v danem trenutku zgodijo in jih opazujemo v časovnem razmiku, 6

enota – vsak posamezen pojav populacije, 5–6

frekvenca – število enot v določenem razredu frekvenčne porazdelitve, 53–54

frekvenčna porazdelitev – statistična vrsta, ki jo dobimo z razvrščanjem enot v skupine po vrednostih katerekoli spremenljivke, 50–58

geometrijska sredina – srednja vrednost, ki jo uporabimo za izračun povprečja iz indeksov in koeficientov rasti, 80–84

gostota frekvence – razmerje med frekvenco in širino razreda, 57

grafikon – grafični način prikazovanja podatkov; stolpci, linijski, 14–18

harmonična sredina – srednja vrednost, ki jo največkrat uporabljamo za izračun povprečja iz struktur in statističnih koeficientov, 75–80

histogram – grafični prikaz frekvenčne porazdelitve s stolpci, 56

indeksi – relativna števila, s katerimi analiziramo relativne spremembe med členi krajevne, časovne ali stvarne vrste, 36–46

indeksi s stalno osnovo–osnovni kazalci dinamike, ki v časovni vrsti kažejo spremembe členov v primerjavi z nekim karakterističnim členom, 38–39

iregularna nihanja – slučajna enkratna nihanja okoli trenda, 105

koeficient asimetrije – koeficient, na osnovi katerega sklepamo o smeri in jakosti asimetrije frekvenčne porazdelitve, 93–94

koeficienti rasti – osnovni kazalci dinamike pojava, ki v obliki koeficienta kažejo spremembe od člena do člena v časovni vrsti, 41–42

koeficient sploščenosti – koeficient, na osnovi katerega sklepamo o sploščenosti frekvenčne porazdelitve, 95

koeficient variabilnosti – relativna mera variabilnosti, ki izraža delež standardnega odklona v aritmetični sredini, 92–93

kumulativna frekvenca – število enot, ki imajo vrednost manjšo od neke vrednosti, npr. zgornje meje določenega razreda, 54–55

kumulativna relativna frekvenca – delež enot, ki imajo vrednost manjšo od neke vrednosti, npr. zgornje meje določenega razreda, 5

kvantili – vrednosti, ki ustrezajo določenemu kvantilnemu rangu; lahko so kvartili, decili ali centili, 59–66

kvantilni rang – relativen položaj posamezne enote v urejeni vrsti, 59

kvartilni razmik – absolutna mera variabilnosti, ki predstavlja razliko med tretjim in prvim kvartilom, 86–87

mediana – srednja vrednost, od katere ima polovica enot preučevane populacije manjše ali kvečjemu enake, polovica enot pa večje vrednosti od vrednosti, ki je enaka mediani, 67–70

modus – najpogostejša vrednost; vrednost, ki jo ima največ enot preučevane populacije, 71–72

normalna frekvenčna porazdelitev – je teoretična frekvenčna porazdelitev, sestavljena na osnovi teoretičnih predpostavk in vnaprej znanih nekaterih statističnih parametrov, 96–98

obdelava podatkov – urejanje podatkov v preglednejšo obliko, 10–11

opisne spremenljivke – spremenljivke, ki imajo vrednosti izražene z besedo (npr.spol, stopnja izobrazbe, kraj študija), 6–8

parametri – značilnosti populacije; lahko so enostavni ali izvedeni, 7–8

periodična nihanja – nihanja v časovni vrsti, ki se pojavljajo na krajša, enako dolga časovna obdobja, 104, 111–113

poligon – grafični prikaz frekvenčne porazdelitve z linijskim grafikonom, 56–57

populacija- skupnost enot, ki jih opredelimo zato, da jih preučimo, 5

povprečni absolutni odklon od aritmetične sredine – absolutna mera variabilnosti, ki pove, za koliko se v povprečju posamezne vrednosti odklanjajo od aritmetične sredine, 86–87

povprečni absolutni odklon od mediane – absolutna mera variabilnosti, ki pove, za koliko se posamezne vrednosti v povprečju odklanjajo od mediane, 88–89

realne enote – enote, ki dejansko obstajajo in jih opazujemo v danem trenutku, 5–6

relativna frekvenca – delež enot v določenem razredu frekvenčne porazdelitve, 54

rang – določa, na katerem mestu v urejeni vrsti je posamezna enota, 59

ranžirna vrsta – statistična vrsta, v kateri so enote razvrščene po velikosti njihovih vrednosti, 60–61

Sheppardov popravek – uporabimo pri izračunu variance iz frekvenčnih porazdelitev z enako širokimi razredi, 92

sezonski indeksi – indeksi, s katerimi merimo sezonska nihanja v časovni vrsti, 111–113

spremenljivka – značilnost enote, ki jo opazujemo, 5–8

srednja vrednost – statistični parameter in je kot tak reprezentančna vrednost vseh enot preučevane populacije, 67

standardni odklon – kvadratni koren iz variance, 90–92

statistična vrsta – enote, razvrščene v skupine po vrednostih ali skupinah vrednosti določene spremenljivke; lahko je neurejena in urejena; časovna, krajevna in stvarna, 11–13

statistika – veda, ki se ukvarja z zbiranjem in urejanjem podatkov, razvojem metod za obdelavo in analizo podatkov ter predstavitvijo rezultatov statistične analize, 4

statistični koeficient – relativno število, izračunano iz dveh raznovrstnih podatkov, ki sta enako opredeljena in je njuna primerjava smiselna, 29–36

statistično opazovanje – zbiranje podatkov o enotah; lahko je popolno in delno; neposredno in posredno, 9–10

stopnje rasti – kažejo relativne spremembe (v odstotku) od člena do člena v časovni vrsti, 41–42

struktura – relativno število, izračunano iz dveh podatkov, od katerih podatek v števcu predstavlja del pojava, podatek v imenovalcu pa celoten pojav, 20–29

Sturgesovo pravilo – pravilo, ki ga uporabimo za določanje najprimernejšega števila razredov frekvenčne porazdelitve, 52

številske spremenljivke – spremenljivke, ki imajo vrednosti izražene s števili (npr. starost, vrednost prodaje); lahko so zvezne in diskretne, 6–8

tabela (preglednica) – pregleden način prikazovanja podatkov; lahko je enostavna, sestavljena in kombinirana, 13–14

trend – osnovna smer razvoja časovne vrste, 103–111

variacijski razmik – najenostavnejša mera variabilnosti; razlika med največjo in najmanjšo vrednostjo preučevanih enot, 85–86

varianca – povprečje kvadratov odklonov vrednosti številske spremenljivke od aritmetične sredine, 89–92

verižni indeks – osnovni kazalec dinamike pojava, ki v časovni vrsti kaže spremembe od člena do člena, 39–40

Literatura in viri

Artenjak, J. *Poslovna statistika*. Maribor: Ekonomsko-poslovna fakulteta, 1997.
Blejec**,** M. *Statistične metode za ekonomiste*. 2. predelana in razširjena izdaja. Ljubljana: Ekonomska fakulteta, 1973.
Košmelj, B. *Statistika*. Ljubljana: Državna založba Slovenije, 1994.
Košmelj, B., in drugi. *Statistični terminološki slovar*. Ljubljana: Statistično društvo Slovenije, in Društvo matematikov, fizikov in astronomov Slovenije, 1993.
Kristan, A. *Statistika I*. Maribor: Ekonomsko-poslovna fakulteta, 1989.
Šadl, M. *Statistika za komercialiste*. Murska Sobota: Ekonomska šola Murska Sobota, Višja strokovna šola, 2006.
Šadl, M. *Statistika za srednje šole*. Celovec: Mohorjeva družba, 2007.
Šadl, M. *Zbirka vaj iz statistike*. Murska Sobota: Ekonomska šola Murska Sobota, Višja strokovna šola, 2007.
Šadl, M. *Zbirka nalog iz statistike*. Celovec: Mohorjeva družba, 2005.

Mesečni statistični pregled, marec 1997**.** Ljubljana: Statistični urad Republike Slovenije, 1997.
Mesečni statistični pregled, februar 2004. Ljubljana: Statistični urad Republike Slovenije, 2004.
Mesečni statistični pregled, februar 2005. Ljubljana: Statistični urad Republike Slovenije, 2005.
Statistični letopis 2003. Ljubljana: Statistični urad Republike Slovenije, 2004.
Statistični letopis 2004. Ljubljana: Statistični urad Republike Slovenije, 2005.
Statistični letopis 2005. Ljubljana: Statistični urad Republike Slovenije, 2006.
Statistični letopis 2007. Ljubljana: Statistični urad Republike Slovenije, 2008.
Pomembnejši statistični podatki o Sloveniji, letnik II, št. 1/2007. Ljubljana: Statistični urad Republike Slovenije, 2007.
Pomembnejši statistični podatki o Sloveniji, letnik II, št. 3/2007. Ljubljana: Statistični urad Republike Slovenije, 2007.
Pomembnejši statistični podatki o Sloveniji, letnik III, št. 3/2008. Ljubljana: Statistični urad Republike Slovenije, 2008.
Pomembnejši statistični podatki o Sloveniji, letnik III, št. 2/2008. Ljubljana: Statistični urad Republike Slovenije, 2008.
Slovenske regije v številkah. Ljubljana: Statistični urad Republike Slovenije, 2008.
Statistični portret Slovenije v EU 2008. Ljubljana: Statistični urad Republike Slovenije, 2008.

Zavod Republike Slovenije za zaposlovanje. Statistične informacije (online). 2008. (Citirano 18. 8. 2008). Dostopno na naslovu: www.ess.gov.si
Eurostat: Euro indikatoren. (online). 2008. (Citirano 15. 8. 2008). Dostopno na naslovu: www.http://epp.eurostat.ec.europa.eu
Statistični urad Republike Slovenije: Statistični letopis 2007: Zaposlene osebe pri pravnih osebah po stopnji strokovne izobrazbe (online). 2008. (citirano 18. 8. 2008). Dostopno na naslovu: shttp://www.stat.si/letopis/2007

Projekt **Impletum**
Uvajanje novih izobraževalnih programov na področju višjega strokovnega izobraževanja v obdobju 2008–11

Konzorcijski partnerji:

Operacijo delno financira Evropska unija iz Evropskega socialnega sklada ter Ministrstvo RS za šolstvo in šport. Operacija se izvaja v okviru Operativnega programa razvoja človeških virov za obdobje 2007-2013, razvojne prioritete 'Razvoj človeških virov in vseživljenjskega učenja' in prednostne usmeritve 'Izboljšanje kakovosti in učinkovitosti sistemov izobraževanja in usposabljanja'