



STATISTIKA

GRADIVA ZA SEMINARSKE VAJE ZA 2. DEL PREDMETA

POLONA TOMINC

MAJA ROŽMAN



Univerzitetna založba
Univerze v Mariboru





Univerza v Mariboru

Ekonomsko-poslovna fakulteta

Statistika

Gradiva za seminarske vaje za 2. del predmeta

Avtorici

Polona Tominc

Maja Rožman

December 2021

Naslov <i>Title</i>	Statistika <i>Statistics</i>		
Podnaslov <i>Subtitle</i>	Gradiva za seminarske vaje za 2. del predmeta <i>Seminar Tutorials for the 2nd Part of the Course</i>		
Avtorici <i>Author</i>	Polona Tominc (Univerza v Mariboru, Ekonomsko-poslovna fakulteta) Maja Rožman (Univerza v Mariboru, Ekonomsko-poslovna fakulteta)		
Recenzija <i>Review</i>	Vesna Čančer (Univerza v Mariboru, Ekonomsko-poslovna fakulteta) Majda Bastič (Univerza v Mariboru, Ekonomsko-poslovna fakulteta)		
Jezikovni pregled <i>Language editing</i>	Alenka Plos		
Tehnični urednik <i>Technical editor</i>	Jan Perša (Univerza v Mariboru, Univerzitetna založba)		
Oblikovanje ovitka <i>Cover designer</i>	Jan Perša (Univerza v Mariboru, Univerzitetna založba)		
Grafične priloge <i>Graphic material</i>	Avtorici		
Založnik <i>Published by</i>	Univerza v Mariboru, Univerzitetna založba Slomškovo trg 15, 2000 Maribor, Slovenija https://press.um.si , zalozba@um.si		
Izdajatelj <i>Issued by</i>	Univerza v Mariboru, Ekonomsko-poslovna fakulteta Razlagova ulica 14, 2000 Maribor, Slovenija https://www.epf.um.si , epf@um.si		
Izdaja <i>Edition</i>	Prva izdaja	Izdano <i>Published at</i>	Maribor, december 2021
Vrsta publikacije <i>Publication type</i>	E-knjiga	Dostopno na <i>Available at</i>	https://press.um.si/index.php/ump/catalog/book/632

<p>CIP - Kataložni zapis o publikaciji Univerzitetna knjižnica Maribor</p> <p>311 (076) (0.034.2)</p> <p>TOMINC, Polona Statistika [Elektronski vir] : gradiva za seminarske vaje za 2. del predmeta / Polona Tominc, Maja Rožman. - 1. izd. - E-knjiga. - Maribor : Univerza v Mariboru, Univerzitetna založba, 2021</p> <p>Način dostopa (URL) : https://doi.org/10.18690/978-961-286-546-7 ISBN 978-961-286-546-7 COBISS.SI-ID 86875651</p>



© Univerza v Mariboru, Univerzitetna založba
/ University of Maribor, University Press

Besedilo / Text © Tominc in Rožman, 2021

To delo je objavljeno pod licenco Creative Commons Priznanje avtorstva-Nekomercialno-Deljenje pod enakimi pogoji 4.0 Mednarodna. / *This work is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International.*

Uporabnikom se dovoli reproduciranje, distribuiranje, dajanje v najem, javno priobčitev in predelavo avtorskega dela, če navedejo avtorja in širijo avtorsko delo/predelavo naprej pod istimi pogoji. Za nova dela, ki bodo nastala s predelavo, ni dovoljena komercialna uporaba.

Vsa gradiva tretjih oseb v tej knjigi so objavljena pod licenco Creative Commons, razen če to ni navedeno drugače. Če želite ponovno uporabiti gradivo tretjih oseb, ki ni zajeto v licenci Creative Commons, boste morali pridobiti dovoljenje neposredno od imetnika avtorskih pravic.

<https://creativecommons.org/licenses/by-nc-sa/4.0/>

ISBN 978-961-286-546-7 (pdf)

DOI <https://doi.org/10.18690/978-961-286-546-7>

Cena
Price Brezplačni izvod

Odgovorna oseba založnika
For publisher prof. dr. Zdravko Kačič,
rektor Univerze v Mariboru

Citiranje
Attribution Tominc, P. in Rožman, M. (2021). *Statistika: gradiva za seminarske vaje za 2. del predmeta*. Maribor: Univerzitetna založba.
doi: 10.18690/978-961-286-546-7

Kazalo

1	Uvod	1
2	Urejanje in prikazovanje podatkov.....	3
	Naloga 1	4
	Naloga 2	5
	Naloga 3	5
	Naloga 4	6
	Naloga 5	6
	Naloga 6	6
	Naloga 7	7
	Naloga 8	7
	Naloga 9	8
	Naloga 10	8
	Naloga 11	9
3	Deskriptivna statistika.....	11
	Naloga 12	12
	Naloga 13	13
	Naloga 14	13
	Naloga 15	14
	Naloga 16	14
	Naloga 17	15
	Naloga 18	15
	Naloga 19	15
4	Teoretične porazdelitve	17
	Naloga 20	18
	Naloga 21	18
	Naloga 22	18
5	Enostavna regresijska analiza	21
	Naloga 23	22
	Naloga 24	23
	Naloga 25	23
	Naloga 26	23
	Naloga 27	24
	Naloga 28	24
	Naloga 29	25
	Naloga 30	25

6	Osnove vzorčenja.....	27
	Naloga 31	28
	Naloga 32.....	28
	Naloga 33.....	29
	Naloga 34.....	29
	Naloga 35.....	29
	Naloga 36.....	29
	Naloga 37.....	30
	Naloga 38.....	30
	Naloga 39.....	30
	Naloga 40.....	31
	Naloga 41.....	31
	Naloga 42.....	31
	Naloga 43.....	31
	Naloga 44.....	32
	Naloga 45.....	32
	Naloga 46.....	32
	Naloga 47.....	32
	Naloga 48.....	33
7	Časovne vrste.....	35
	Naloga 49	36
	Naloga 50.....	36
	Naloga 51.....	37
	Naloga 52.....	37
	Naloga 53.....	38
	Naloga 54.....	38
	Naloga 55.....	39
	Naloga 56.....	39
	Naloga 57.....	40
	Naloga 58.....	40
	Naloga 59.....	40
	Naloga 60.....	41
8	Rešitve nalog	43
	Urejanje in prikazovanje podatkov.....	43
	Naloga 2b.....	43
	Naloga 2d.....	43
	Naloga 3	44
	Naloga 5b.....	45
	Naloga 6b.....	45
	Naloga 6c	46
	Naloga 7	46
	Naloga 9c	46
	Naloga 10c, 10d.....	47
	Naloga 11b, 11c	47
	Deskriptivna statistika.....	47
	Naloga 13a	47
	Naloga 13b.....	48
	Naloga 13c.....	48
	Naloga 13d.....	48
	Naloga 13e	48

Naloga 14b.....	49
Naloga 14c.....	49
Naloga 15a.....	49
Naloga 15b.....	50
Naloga 16c.....	50
Naloga 16d.....	50
Naloga 17a.....	50
Naloga 17b.....	50
Naloga 18a.....	51
Naloga 18b.....	51
Naloga 18c.....	51
Naloga 19b.....	51
Teoretične porazdelitve	52
Naloga 20a.....	52
Naloga 20b.....	53
Naloga 20c.....	53
Naloga 21a.....	53
Naloga 21b.....	54
Naloga 21c.....	54
Naloga 22a.....	54
Naloga 22b.....	54
Enostavna regresijska analiza	54
Naloga 23a.....	54
Naloga 23b.....	55
Naloga 23c.....	56
Naloga 23d.....	56
Naloga 23e.....	57
Naloga 23f.....	57
Naloga 23g.....	57
Naloga 24a.....	58
Naloga 26b.....	60
Naloga 26c.....	61
Naloga 27a.....	62
Naloga 27b.....	63
Naloga 28.....	63
Naloga 29.....	65
Naloga 30a.....	66
Naloga 30b.....	66
Osnove vzorčenja.....	67
Naloga 31.....	67
Naloga 32.....	67
Naloga 33.....	68
Naloga 34.....	69
Naloga 35.....	70
Naloga 36.....	71
Naloga 37.....	72
Naloga 38a.....	72
Naloga 38b.....	72
Naloga 38c.....	73
Naloga 38d.....	73
Naloga 38e.....	74
Naloga 39a.....	75
Naloga 39b.....	75

Naloga 40	76
Naloga 41	76
Naloga 42a	76
Naloga 42b.....	77
Naloga 43	77
Naloga 44	77
Naloga 45	78
Naloga 46	78
Naloga 47a	78
Naloga 47b.....	79
Naloga 48a	79
Naloga 48b.....	79
Časovne vrste	80
Naloga 50a	80
Naloga 50b.....	80
Naloga 51a, 51b, 51c	80
Naloga 52	80
Naloga 53a	81
Naloga 53b.....	81
Naloga 53c	81
Naloga 53d.....	82
Naloga 54a	82
Naloga 54b.....	82
Naloga 55a	83
Naloga 55b.....	83
Naloga 56a, 56b.....	83
Naloga 56c	83
Naloga 57a	83
Naloga 57b.....	84
Naloga 58a	84
Naloga 58c	84
Naloga 59a	85
Naloga 59b.....	85
Naloga 60b.....	85
Tabele	87
Ploščine $H(z)$ za standardizirano normalno porazdelitev	87
Kritične vrednosti za t porazdelitev.....	88
Literatura in viri.....	89

1 Uvod

Znanost o statistiki se ukvarja z zbiranjem, analizo, interpretacijo in predstavitvijo podatkov (Holmes idr., 2018; Ghauri idr., 2020).

Urejanje in opisovanje podatkov sodi v področje *opisne statistike*. Podatke lahko opisujemo na dva načina: grafično ali pa z izračunom določenih vrednosti, ki te podatke predstavljajo, na primer povprečja. Metode *inferenčne statistike* pa nam omogočajo na podlagi podatkov ustrezne kakovosti (verjetnostni, reprezentativni vzorec statistične množice) sklepati o značilnostih statistične množice kot celote, z določeno verjetnostjo. *Statistično sklepanje* uporablja koncept verjetnosti, da ugotovimo, kako prepričani smo lahko, da so naši sklepi pravilni.

Učinkovito statistično sklepanje temelji na ustreznih postopkih za pripravo podatkov in premišljenem pregledu podatkov. Cilj statistike pa vsekakor ni izvajanje številnih izračunov s pomočjo bolj ali manj zapletenih formul, temveč pridobivanje uporabnih informacij na osnovi pridobljenih podatkov. Izračune lahko opravimo s pomočjo kalkulatorja ali programskega orodja, vendar pa razumevanja uporabnosti informacij, ki jih pridobimo na osnovi statističnih izračunov, računalnik ne more nadomestiti. Zato je predmet ***Statistika (drugi del predmeta)***, za katerega je namenjena ta zbirka vaj¹, pomemben za to, da študenti razvijejo sposobnost razumevanja informacij v podatkih.

¹ Nekatere naloge so prirejene po zbirki nalog, ki se je uporabljala pri predmetu v preteklem obdobju (Leskovar Špacapan in Tominc, 2008).

Predmet Statistika (drugi del predmeta) vsebuje naslednje vsebinske sklope:

- a) prikazovanje podatkov v tabelah in grafih,
- b) relativna števila,
- c) srednje vrednosti, mere variabilnosti, asimetrije in sploščenosti,
- d) intervalno ocenjevanje vrednosti statističnih parametrov in osnove preizkušanja domnev o statističnih parametrih,
- e) osnove enostavne regresije,
- f) osnove analize in napovedovanja vrednosti v časovnih vrstah.

Študenti v okviru predmeta spoznajo uporabnost statističnih metod pri reševanju poslovnih problemov ter utrdijo in nadgradijo teoretično znanje na področju statističnih tehnik in metod, ki omogočajo spremeniti različne podatke v uporabne informacije za poslovno odločanje. Študenti osvojijo analitičen matematično statističen pristop k preučevanju poslovnih problemov, ki se sestoji iz naslednjih korakov:

- formulacija problema na statističen način,
- izbira ustrezne statistične metode,
- reševanje problema,
- interpretacija rezultatov v smislu možnih rešitev problema.

V tej zbirki nalog pri nekaterih primerih uporabljamo programsko orodje IBM SPSS.

2 Urejanje in prikazovanje podatkov

S pomočjo uporabe statističnih metod želimo običajno proučiti populacijo ali *statistično množico*. Statistično množico sestavljajo *statistične enote*, ki jih proučujemo po izbranih značilnostih. Za preučevanje statistične množice pogosto izberemo (verjetnostni) *vzorec*. Ideja vzorčenja je izbrati del (ali podskupino) statistične množice in proučiti ta del (vzorec), da pridobimo informacije o celotni statistični množici. Ker za pregled celotne statistične množice potrebujemo veliko časa in povzroča velike stroške, je vzorčenje zelo uporabna tehnika. Vzorec mora dobro opisovati značilnosti statistične množice, da je reprezentativen (Holmes idr., 2018).

Iz vzorčnih podatkov lahko izračunamo različne *vzorčne vrednosti statističnih parametrov* – *ocene statističnih parametrov* v statistični množici. *Statistični parameter* je številska značilnost celotne statistične množice.

Spremenljivka je značilnost, ki jo je mogoče določiti za vsako statistično enoto v statistični množici. *Spremenljivke so lahko številske ali opisne* (Sakaran in Bougie, 2016). Vrednosti številske spremenljivke so izražene z merskimi enotami, kot so na primer teža v kilogramih, čas v urah, dobiček v EUR. Vrednosti opisne spremenljivke se izražajo z besedami. Opisne (ali kategorične) spremenljivke uvrščajo statistično enoto v določeno kategorijo, na primer sedež podjetja v eni od 12 statističnih regij v Sloveniji. Pri številskih spremenljivkah lahko podatke opišemo z izračunom različnih številskih vrednosti (lahko bi izračunali povprečno vrednost), vendar pa tega ni smiselno delati z vrednostmi opisne spremenljivke (izračun »povprečne regije« za sedež podjetij v Sloveniji ni smiseln).

Najosnovnejšo obliko statističnih podatkov predstavljajo statistične vrste, ki jih delimo v tri skupine: *časovne*, *krajevne* in *stvarne statistične vrste* (Tominc in Kramberger, 2007). Tabelačno urejene podatke statistične vrste lahko grafično prikažemo z ustreznim grafom, ki še bolj nazorno prikazuje značilnosti statistične vrste. Vzemimo primer časovne statistične vrste, kjer so vrednosti spremenljivke urejene v nekih enakomernih časovnih presledkih. Najpomembnejša značilnost časovne statistične vrste, ki jo želimo običajno spoznati, je dinamika – spreminjanje vrednosti opazovane spremenljivke skozi čas. Za prikaz dinamike v časovni vrsti je tako zelo pogosto uporabljen grafični prikaz, ki ga imenujemo *linijski grafikon*.

Vrednosti številske spremenljivke lahko uredimo v razrede *frekvenčne porazdelitve*, ki jih grafično prikazujemo s *frekvenčnimi histogrami ali pologoni*, vrednosti opisne spremenljivke pa grupiramo v skupine na osnovi možnih vrednosti opisne spremenljivke (Tominc in Kramberger, 2007). Izraz *frekvenca* označuje število ponovitev vrednosti v podatkih (bodisi posameznih vrednosti spremenljivke bodisi opredeljenega intervala za vrednosti spremenljivke). Vsota vseh frekvenc predstavlja skupno število statističnih enot v opazovanem podatkovnem nizu.

Relativna frekvenca je razmerje (ali delež) med številom statističnih enot z določeno vrednostjo spremenljivke (ali vrednostjo spremenljivke v opredeljenem intervalu za vrednosti spremenljivke) in skupnim številom statističnih enot v podatkovnem nizu. Relativne frekvence so lahko izražene v deležu ali v odstotku. *Strukturo podatkovnega niza* (vzorca ali statistične množice) grafično najpogosteje prikazujemo s strukturnim stolpcem (Tominc in Kramberger, 2007).

Naloga 1

V nekem mestu, ki šteje 350 tisoč prebivalcev, je javno podjetje, ki izvaja potniški mestni promet, zbiralo informacije o potrebni spremembi voznih redov ter cene prevozov. V ta namen so v času od 1. do 15. maja lani zbrali odgovore 150 naključno izbranih prebivalcev o najvišji ceni (v EUR), ki so jo pripravljene odšteti za eno vožnjo z mestnim avtobusom.

- Opreделите statistično množico in vzorec.
- Opreделите statistično enoto.
- Opreделите statistično spremenljivko in njene značilnosti.

- d) Naštejte nekaj možnih statističnih parametrov pri analizi obravnavane spremenljivke.
- e) Ali so bili v tem primeru zbrani primarni ali sekundarni podatki? Opredelite razliko med obema viroma podatkov.

Naloga 2

V preglednici so podatki o številu trgovin neke trgovske verige, v šestih slovenskih regijah, v nekem časovnem obdobju ter podatki o višini investicij v trgovske objekte te trgovske verige (v d.e.):

Regija	Gorenjska	Goriška	Primorska	Koroška	Pomurska	Posavska	Zasavska
Št. trgovin	21	25	14	17	25	12	18
Investicije (v d.e.)	314	100	100	300	260	175	250

- a) Kako imenujemo statistični vrsti v preglednici?
- b) Statistično vrsto za število trgovin prikažite grafično.
- c) Navedite poljubni primer stvarne in krajevne statistične vrste.
- d) Za podatke o višini investicij po regijah prikažite relativno strukturo skupno porabljenih sredstev po regijah.

Naloga 3

Za analizirano podjetje imate na razpolago podatke o številu zaposlenih trgovskih potnikov na sredini meseca ter podatke o prodaji izdelka v prvih petih mesecih lanskega leta:

Mesec	I.	II.	III.	IV.	V.
Vrednost prodaje (v 10^3 EUR)	2.600	3.800	4.100	2.900	3.200
Število zaposlenih trgovskih potnikov	30	75	82	80	90

- a) Izračunajte povprečno prodajo na trgovskega potnika po mesecih.
- b) Izračunajte povprečno prodajo na trgovskega potnika v obravnavanem petmesečnem obdobju.

Naloga 4

Za neko fakulteto poznamo za pet študijskih programov podatke o številu vpisanih študentov v preteklem akademskem letu.

Študijski program	A	B	C	D	E
Število diplomantov	45	55	100	30	25

Prikažite relativno strukturo skupnega števila vpisanih študentov v teh pet študijskih programov fakultete.

Naloga 5

V preglednici so podatki o številu diplomantov po področjih študija in stopnji študija v neki državi v preteklem letu.

Področje študija	Dodiplomski študij	Magistrski študij
Kmetijske vede	71	206
Medicinske vede	243	268
Tehniške in naravoslovne vede	793	1.112
Družbene in humanistične vede	2.110	2.921

- Opreделите statistično enoto in statistični spremenljivki.
- Izpišite statistično vrsto diplomantov na dodiplomski stopnji po področjih študija. Kako imenujemo takšno statistično vrsto? Statistično vrsto prikažite grafično.
- Grafično prikažite relativno strukturo diplomantov magistrskega študija po področjih študija.

Naloga 6

Razpolagamo s podatki o oceni zadovoljstva zaposlenih (na lestvici od 1 do 100) za 50 zaposlenih:

15 17 18 18 18 18 19 19 19 20
 23 24 25 25 26 27 29 29 29 29
 30 30 35 37 38 36 39 39 39 39
 41 45 42 46 46 48 48 48 46 49
 55 62 65 72 72 85 91 95 91 95

- Opreделите statistično množico, statistično enoto ter spremenljivko, njene značilnosti in zalogo vrednosti spremenljivke.
- Sestavite frekvenčno porazdelitev pri pogojih: $y_{1,\min} = y_{\min}$; $i = 10$, meje razredov so podane nezvezno.
- Frekvenčno porazdelitev ter kumulativno frekvenčno porazdelitev prikažite grafično.

Naloga 7

Za 300 študentov, ki smo jih opazovali glede na število ur študija (zvezna spremenljivka) za izpit iz predmeta Statistika, imate na razpolago naslednje podatke:

$$\begin{array}{llll}
 y_{\min} = 41 \text{ ur} & y_{\max} = 97 \text{ ur} & y_{1,\min} = 40 \text{ ur} & y_{6,\max} = 100 \text{ ur} \\
 r = 6 & i = 10 & \text{za } k = 1,2,\dots,r & \\
 F_1 = 25 & F_2 = 75 & F_3 = 175 & F_4 = 250 & F_5 = 290 & F_6 = 300
 \end{array}$$

- Na osnovi danih podatkov sestavite frekvenčno porazdelitev in jo grafično prikažite.
- Koliko odstotkov študentov je porabilo od 60–70 ur študija?
- Koliko odstotkov študentov je porabilo do 80 ur študija?

Naloga 8

Podatki za 500 kupcev glede na porabljen znesek za nakup v neki trgovini, na dan 31.12. preteklega leta, v eni od slovenskih regij so:

Znesek za nakup				Število kupcev
Od 1	do/pod	50		80
Od 50	do/pod	100		125
Od 100	do/pod	150		148
Od 150	do/pod	200		112
Od 200	do/pod	250		35

- Opreделите statistično enoto, statistično množico, statistično spremenljivko ter njene vrednosti.
- Prikažite relativno strukturo statistične množice s strukturnim stolpcem.

Naloga 9

Določiti želimo omejitveno hitrost na cesti skozi naselje. Ena izmed možnosti je ta, da je omejitvena hitrost tista, ki jo pri neomejeni hitrosti na cesti presega 15 % vozil.

- Opišite potek statistične raziskave. Opredelite statistično enoto in statistično spremenljivko za to raziskavo.
- Pri merjenju hitrosti smo dobili podatke v preglednici. Ali so podatki, prikazani na tak način, pregledni?
- Kako bi lahko podatke pregledneje prikazali?

Hitrost v km/h	Število osebnih vozil
41	0
42	1
44	0
46	1
48	1
50	1
52	2
54	6
56	7
58	9
60	1
62	2
64	8
66	7
68	3
70	4
72	5
74	2
76	3
78	2
80	3
82	0
84	1
86	1
88	0
90	0

Naloga 10

V osrednjeslovenski regiji je bilo na dan 31.12. v podjetjih neke panoge toliko zaposlenih, kot prikazujejo podatki v preglednici.

Število zaposlenih	Število podjetij
Od 11 do 40	15
Od 41 do 55	17
Od 56 do 70	23
Od 71 do 85	25
Od 86 do 100	17
SKUPAJ	97

- Opreделите statistično enoto, statistično množico in spremenljivko.
- Frekvenčno porazdelitev prikažite grafično.
- Izračunajte in pojasnite člene kumulativne frekvenčne porazdelitve.
- Izračunajte in pojasnite relativne frekvence ter grafično prikažite relativno strukturo statistične množice.

Naloga 11

Število vozil, ki pridejo na bencinski servis v eni uri, je podano v preglednici.

- Pojasnite statistično enoto, statistično množico, statistično spremenljivko in njene vrednosti.
- Izračunajte in pojasnite člene kumulativne frekvenčne porazdelitve in jo prikažite grafično.
- Izračunajte in pojasnite relativne frekvence ter grafično prikažite strukturo statistične množice glede na skupno število vozil po delovnih urah v posameznih razredih.

Število vozil	Število delovnih ur
Od 0 do 4	9
Od 5 do 9	15
Od 10 do 14	5
Od 15 do 19	6
Od 20 do 24	4
Od 25 do 29	2
Od 30 do 34	2

3 Deskriptivna statistika

V okviru opisne statistike običajno izračunavamo naslednje mere (Holmes idr., 2018):

- kvantili: to so mere, ki opredeljujejo položaj posamezne statistične enote glede na ostale po vrednosti obravnavane statistične spremenljivke. Najpomembnejši so kvartili, decili ter mediana;
- mere osrednje tendence: aritmetična sredina, modus, pa tudi že zgoraj omenjena mediana;
- mere variabilnosti: variacijski razmik, kvartilni in decilni razmik, varianca in standardni odklon ter koeficient variabilnosti;
- mere asimetrije in sploščenosti: koeficient asimetrije in koeficient sploščenosti.

Za izračun kvantilov je treba podatke razvrstiti po vrednosti obravnavane spremenljivke od najmanjših do največjih (*ranžirna vrsta* pri manjšem številu statističnih enot ali frekvenčna porazdelitev pri večjem številu statističnih enot). *Kvartili* (Q) razdelijo urejene podatke na četrtine (po 25 %), *decili* (D) pa na desetine (po 10 %). *Mediana* (Me) je enaka drugemu kvartilu, pa tudi petemu decilu.

Med merami osrednje tendence sta najpomembnejši *aritmetična sredina* (\bar{y} – če obravnavamo podatke statistične množice – ter \bar{Y} – če obravnavamo podatke vzorca) in že omenjena mediana. Aritmetično sredino vrednosti spremenljivke izračunamo, če vsoto vseh vrednosti spremenljivke (*total*) delimo s številom statističnih enot v podatkih (bodisi vzorca bodisi statistične množice).

Med merami variabilnosti je najenostavnejši *variacijski razmik* (VR), ki je razlika med največjo in najmanjšo vrednostjo spremenljivke. *Kvartilni razmik* je mera variabilnosti, ki označuje razpon vrednosti srednjih 50 % podatkov. To je razlika med *tretjim kvartilom* (Q_3) in *prvim kvartilom* (Q_1). Podobno je *decilni razmik* razlika med *devetim* (D_9) in *prvim decilom* (D_1) (srednjih 80 % podatkov).

Najpomembnejši meri variabilnosti sta *standardni odklon in varianca*. Standardni odklon predstavlja mero za skupno razpršenost vrednosti okoli aritmetične sredine v danem podatkovnem nizu.

Standardni odklon (σ) je vedno pozitiven (ali nič). Za izračun standardnega odklona moramo najprej izračunati varianco. Standardni odklon je kvadratni koren iz variance. Varianca je povprečni kvadratni odklon (razlika) vrednosti spremenljivke od njene aritmetične sredine. *Varianco pri podatkih statistične množice* (σ^2) izračunamo tako, da prej omenjeno vsoto kvadratov odklonov delimo s *številom statističnih enot v statistični množici* (N). Če so podatki iz vzorca in ne iz statistične množice, vsoto kvadratov odklonov delimo z $n - 1$, za eno manj kot *število statističnih enot v vzorcu* (n) (*nepristranska ocena variance* – s^2) (Tominc in Kramberger, 2007).

Koeficient variabilnosti, izražen v % ($KV\%$), pove, koliko odstotkov aritmetične sredine predstavlja standardni odklon za dani podatkovni niz.

Asimetrijo in sploščenost porazdelitve vrednosti v danem podatkovnem nizu lahko analitično proučujemo z izračunom *kazalca asimetrije na podlagi mediane* (KA_{Me}) ali na podlagi *modususa* (KA_{Mo}).

Naloga 12

Definirajte oziroma pojasnite naslednje pojme:

ranžirna vrsta, kvartili, decili, mediana, modus, aritmetična sredina, variacijski razmik, kvartilni oziroma decilni razmik, varianca, standardni odklon, koeficient variabilnosti, koeficient asimetrije, koeficient sploščenosti.

Naloga 13

V dveh organizacijah smo opazovali zaposlene glede na čas (v minutah), ki so ga porabili za izdelavo enega izdelka. Podatki so:

Organizacija A: $N = 12$ $y_i = 26, 38, 45, 22, 33, 29, 34, 41, 40, 39, 43, 30$ minut

Organizacija B: $N = 730$

Organizacija B	
Poraba časa v minutah	Število zaposlenih
Nad 22 do 26	76
Nad 26 do 30	123
Nad 30 do 34	235
Nad 34 do 38	162
Nad 38 do 42	98
Nad 42 do 46	36
Skupaj	730

Izračunajte in pojasnite:

- kvartilni razmik in mediano (za organizacijo A),
- povprečno porabljeni čas za en izdelek v organizaciji A in v organizaciji B,
- variacijski razmik za podatke organizacije A,
- varianco in standardni odklon za podatke organizacij A in B,
- koeficient variabilnosti za podatke organizacij A in B.

Naloga 14

Na nekem področju smo opazovali 260 sodnikov okrajnih sodišč po številu obravnavanih zadev v določenem časovnem razdobju. Podatki so v tabeli:

Število obravnavanih zadev	Število sodnikov
Od 31 do 60	35
Od 61 do 90	52
Od 91 do 120	74
Od 121 do 150	41
Od 151 do 180	32
Od 181 do 210	26
Skupaj	260

- Navedite statistično množico, enoto, spremenljivko in vrednosti spremenljivke.
- Grafično ocenite asimetrijo gornje porazdelitve.
- Izračunajte delež standardnega odklona v aritmetični sredini.

Naloga 15

V banki so analizirali višino vloženi sredstev na današnji dan (v 10^2 EUR) za devet varčevalcev. Podatki so v sledeči:

9 12 16 41 60 80 100 120 150

- Kolikšne zneske je vložilo 50 % varčevalcev, ki so vložili srednje velike zneske, in kolikšne zneske tistih 50 %, ki so vložili najmanj?
- V banki "Y" je višina povprečno vloženi sredstev 456 vlagateljev 7.530 EUR, varianca pa 7.022.500 EUR². V kateri banki se vlagatelji glede na vložena sredstva med seboj bolj razlikujejo?

Naloga 16

Na nekem cestnem odseku so v 45 zaporednih dneh prešteli takšno število osebnih vozil:

Število osebnih vozil	Število dni
Od 1 do 10	6
Od 11 do 20	14
Od 21 do 30	15
Od 31 do 40	5
Od 41 do 50	5
Skupaj	45

- Opređelite statistično množico, statistično enoto ter spremenljivko ter variacijski razmik.
- Grafično prikažite frekvenčno porazdelitev.
- Izračunajte in pojasnite aritmetično sredino.
- Izračunajte in pojasnite koeficient variabilnosti v odstotku.

Naloga 17

V petih državah je bil nekega leta odstotek odraslih, ki se ukvarjajo s podjetništvom, tak:

Država	Brazilija	Mehika	Kitajska	Irska	Kanada
% preb.	13,53	12,40	12,34	9,14	8,82

- Povprečni BDP na prebivalca v teh državah je 18.000 \$, standardni odklon pa znaša 35 % povprečnega BDP/preb. Pri kateri spremenljivki je variabilnost večja in kaj to pomeni?
- Izračunajte in pojasnite mediano.

Naloga 18

Za osem podjetij imamo podatke o odstotku zmanjšanja stroškov energije, ki so ga dosegli z uvedbo izboljššanega tehnološkega postopka: 2,5 %, 3,3 %, 5,7 %, 1,2 %, 5 %, 2 %, 2,4 % in 8 %.

- Opreделите vse srednje vrednosti.
- Kolikšen odstotek zmanjšanja stroškov energije so dosegla podjetja, ki ležijo v kvartilnem razmiku?
- Podjetja smo analizirali tudi glede na dosežen dobiček, kjer smo ugotovili, da standardni odklon predstavlja 75 % aritmetične sredine. Pri kateri spremenljivki je variabilnost večja in kaj to pomeni?

Naloga 19

Donosi 10-ih delnic (v d.e.) so bili v opazovanem obdobju sledeči:

10 40 50 50 71 82 800 850 1.000 1.100

- Opreделите statistično enoto, statistično spremenljivko in statistično množico.
- Kolikšne donose je doseglo 25 % najmanj donosnih delnic in kolikšne 25 % najbolj donosnih delnic?

4 Teoretične porazdelitve

V tem poglavju bomo kot najpomembnejšo teoretično porazdelitev obravnavali *normalno porazdelitev*, ki jo bomo uporabljali za dva namena v okviru tega dela predmeta: za računanje verjetnosti, da naključno izbrana statistična enota zavzame vrednost spremenljivke z določenega intervala, ter v postopkih regresijske analize (naslednje poglavje). Kasneje se bomo srečali še z drugo teoretično porazdelitvijo – *t-porazdelitev*, v poglavju o vzorčenju.

Normalna porazdelitev je definirana z dvema parametroma, aritmetično sredino (\bar{y}), velikokrat pa se za aritmetično sredino v statistični množici uporablja tudi oznaka μ) in standardnim odklonom (σ) (Holmes idr., 2018). Normalna porazdelitev je simetrična glede na aritmetično sredino, aritmetična sredina pa je po vrednosti hkrati tudi enaka mediani in modusu.

Sprememba standardnega odklona σ povzroči spremembo oblike normalne krivulje; krivulja postane bolj sploščena ali pa bolj koničasta, odvisno od σ . Sprememba aritmetične sredine pa povzroči, da se graf premakne v levo ali desno. To pomeni, da obstaja neskončno število normalnih porazdelitev. Ena od posebno pomembnih porazdelitev v statistiki se imenuje *standardizirana normalna porazdelitev*.

Standardizirana normalna porazdelitev je normalna porazdelitev standardiziranih vrednosti, imenovanih tudi z-vrednosti (vsaki vrednosti y spremenljivke, porazdeljeni po poljubni normalni porazdelitvi, je mogoče izračunati njeno standardizirano z-vrednost z upoštevanjem transformacijske enačbe) (Holmes idr., 2018).

Povprečje za standardizirano normalno porazdelitev je enako 0, standardni odklon in varianca pa sta enaka 1, kar zelo poenostavi izračunavanje verjetnosti. Ploščine pod krivuljo normalne porazdelitve (ki so grafično predstavljene verjetnosti) so za standardizirano normalno porazdelitev prikazane v tabeli v prilogi.

Naloga 20

V gospodinjstvih v neki družbi je čas uporabe družinskega računalnika v gospodinjstvu za igranje igrice porazdeljen po normalni porazdelitvi, z aritmetično sredino dve uri in standardnim odklonom 0,5 ure.

- Izračunajte verjetnost, da je v naključno izbranem gospodinjstvu družinski računalnik v uporabi za igranje igrice med 1,8 in 2,75 urami na dan.
- V koliko odstotkih gospodinjstev se družinski računalnik uporablja več kot 5 ur dnevno za igranje igrice?
- Koliko časa je družinski računalnik v uporabi za igranje igrice pri tistih 25 odstotkih gospodinjstev, kjer je ta čas najkrajši?

Naloga 21

V neki državi je starost uporabnikov pametnih telefonov (analizirali smo tiste, ki so stari 13 let ali več) porazdeljena približno po normalni porazdelitvi, s povprečno starostjo 36,9 let in standardnim odklonom 13,9 let.

- Izračunajte verjetnost, da je naključno izbrani uporabnik pametnega telefona v analizirani populaciji star med 23 in 64 let.
- Izračunajte verjetnost, da je naključno izbrani uporabnik pametnega telefona v analizirani populaciji star največ 50,8 let.
- Izračunajte, kolikšna je starost 30 % najstarejših uporabnikov pametnih telefonov v analizirani populaciji.

Naloga 22

Število minut pogovora na mesec pri vseh naročnikih mobilne telefonije nekega operaterja se porazdeljuje po normalni porazdelitvi, z aritmetično sredino 220 minut in standardnim odklonom 70 minut.

- a) Izračunajte, kolikšna je verjetnost, da bo naključni izbrani uporabnik imel na mesec več kot 195 minut pogovora.
- b) Koliko minut pogovora beležimo pri tistih 68,3 % naročnikov, ki ležijo v simetričnem intervalu okoli aritmetične sredine?

5 Enostavna regresijska analiza

O enostavni regresijski analizi govorimo takrat, kadar utemeljeno domnevamo, da neka spremenljivka (neodvisna spremenljivka, x) vpliva na odvisno spremenljivko (y). Analiza tega vpliva vključuje izračun linearne enačbe, ki ponazarja odvisnost spremenljivke y od spremenljivke x , ter izračun osnovnih kazalnikov te odvisnosti (kazalniki linearne korelacije in regresije). Grafični prikaz, ki ga uporabljamo pri enostavni regresijski analizi, se imenuje *razsevni grafikon*.

Linearno *regresijsko enačbo* enostavne regresije zapišemo v obliki $y = b_0 + b_1x$, kjer sta b_0 in b_1 konstanti. *Konstanta b_1 je regresijski koeficient*, ki opisuje (povprečno) spremembo, ki se pojavi pri vrednosti odvisne spremenljivke, ko se neodvisna spremenljivka spremeni za 1 enoto. Kaže tudi smer povezanosti, glede na predznak. *Regresijska konstanta b_0* pa opisuje povprečno vrednost spremenljivke y , takrat ko je neodvisna spremenljivka enaka 0 (če je to smiselno) (Holmes idr., 2018).

Med *kazalniki linearne korelacije in regresije* pri enostavni regresiji obravnavamo še naslednje kazalnike (Tominc in Kramberger, 2007):

- *korelacijski koeficient*, ki opredeljuje smer in jakost linearne povezanosti med spremenljivkama y in x ;
- *determinacijski koeficient*, ki opisuje delež pojasnjene variance spremenljivke y v skupni varianci odvisne spremenljivke;

- *standardno napako ocene odvisne spremenljivke*, ki kaže na vpliv drugih (neodvisnih) spremenljivk, ki jih nismo eksplicitno vključili v analizo ter velikost slučajnih vplivov.

Opisane rezultate regresijske analize lahko s pridom uporabimo pri napovedovanju vrednosti odvisne spremenljivke pri izbrani vrednosti spremenljivke x . *Točkovna ocena vrednosti spremenljivke y pri izbrani vrednosti spremenljivke x* je pridobljena tako, da vrednost spremenljivke x vstavimo v regresijsko enačbo. Pri *intervalni oceni vrednosti spremenljivke y pri izbrani vrednosti spremenljivke x* pa upoštevamo, da na odvisno spremenljivko vplivajo še druge spremenljivke in slučajni vplivi (Tebachnick in Fidel, 2013). Intervalna ocena pomeni, da z določeno stopnjo verjetnosti ocenimo, kakšno vrednost spremenljivke y lahko v povprečju pričakujemo pri izbrani vrednosti spremenljivke x , če upoštevamo tudi standardno napako ocene odvisne spremenljivke (Tominc in Kramberger, 2007).

Naloga 23

Za šest podjetij so podatki o investicijah v novo tehnologijo (v 10^6 evrov) ter ustvarjenim dobičkom (v 10^3 evrov) v letu X naslednji:

Podjetje	A	B	C	D	E	F
Investicije v tehnologijo (x)	115	130	140	149	160	171
Ustvarjen dobiček (y)	328	330	390	361	421	400

- S prikazom dvojic vrednosti opazovanih spremenljivk v razsevnem grafikonu določite obliko, smer in jakost odvisnosti med spremenljivkama.
- Z metodo najmanjših kvadratov izračunajte parametre regresijske premice, izračunano regresijsko premico vrišite v razsevni grafikon.
- Ocenite višino dobička za podjetje, ki bi investiralo v tehnologijo $x = 180$ (v 10^6 evrov), točkovna ocena.
- Izračunajte parameter, na osnovi katerega določite smer in jakost linearne korelacijske odvisnosti.
- Izračunajte delež pojasnjene variance v skupni varianci za odvisno spremenljivko.
- Izračunajte standardno napako ocene odvisne spremenljivke.
- Ob upoštevanju linearne korelacijske odvisnosti ocenite z verjetnostjo 95 % višino dobička pri $x = 180$ (v 10^6 evrov), intervalna ocena (upoštevajte, da je popravek h_1 enak 0).

Naloga 24

V banki "X" so ob različni obrestni meri (x) zabeležili naslednje zneske kratkoročnih oblik varčevanja (y):

Obrestna mera v %	(x)	3,4	4,2	5,6	6,4	8
Varčevanje (v d.e.)	(y)	123	165	197	234	258

- Ocenite z zanesljivostjo 95 % znesek kratkoročnih oblik varčevanja pri obrestni meri $x = 7\%$.
- Pojasnite vse kazalce linearne regresije in korelacije, ki ste jih izračunali pod a).

Naloga 25

Pojasnite:

- smer in jakost povezanosti med odvisno in neodvisno spremenljivko, če je $r_{xy} = -0,8$;
- determinacijski koeficient, če je $r_{xy}^2 = 0,90$;
- velikost standardne napake ocene (σ_{ey}), če je $r_{xy} = -1$.

Naloga 26

Domnevamo, da je znesek stroškov na delavca (v d.e.), ki ga podjetja namenjajo za okoljevarstvene dejavnosti, odvisen od velikosti podjetja (merjeno v številu zaposlenih). Za pet podjetij so podatki v preglednici.

Število zaposlenih	25	46	120	91	37
Znesek stroškov/delavca	15	17	35	30	25

- Narišite razsevni grafikon in ga pojasnite.
- Izračunajte in pojasnite vse kazalce linearne korelacije in regresije.
- Ocenite znesek stroškov za okoljevarstvene dejavnosti podjetja, če ima podjetje 180 zaposlenih. Upoštevajte $\alpha = 5\%$ (upoštevajte, da je popravek h_1 enak 0).

Naloga 27

V tabeli so podatki o tedenskem številu potrjenih primerov s SARS-CoV-2 glede na dve različni lokaciji prenosa okužbe, in sicer delovno mesto in trgovina. Zanima nas, ali okuženost oseb v trgovini vpliva na povečanje okuženosti oseb na delovnem mestu.

Podatki o tedenskem številu potrjenih primerov s SARS-CoV-2 na delovnem mestu in v trgovini:

Teden	Delovno mesto	Trgovina
1. (2020–47)	257	186
2. (2020–48)	260	216
3. (2020–49)	274	221
4. (2020–50)	269	266
5. (2020–51)	233	227
6. (2020–52)	178	156
7. (2020–53)	127	174

- Ocenite osnovne značilnosti povezanosti med spremenljivkama ter izračunajte in pojasnite vse kazalce linearne korelacije in regresije.
- Izpišite enačbo regresijske premice.

Naloga 28

V tabeli so podatki o zasedenosti sob v hotelih ter njihovih cenah:

Povprečni % zasedenosti	Cena v dolarjih
78,70	80,99
68,70	82,48
63,90	76,50
71,70	101,76
70,80	100,25
80,50	169,19
66,90	78,74

Izračunajte koeficiente linearne regresijske funkcije in ocenite, kolikšno ceno lahko pričakujemo za sobo, ki je povprečno zasedena 95 %-no (upoštevajte, da je popravek h_1 enak 0). Intervalno oceno napravite s 95%-no verjetnostjo.

Naloga 29

V preglednici so podatki o povprečni začetni plači diplomanta MBA programa (v d.e.) in o oceni kvalitete MBA programa (na lestvici od 1 do 100), ki jo je opravila zaposlitvena agencija. Ocenite osnovne značilnosti povezanosti med spremenljivkama ter izračunajte in pojasnite vse kazalce linearne korelacije in regresije.

Povprečna začetna plača	Ocena kvalitete programa
105,1	100
96	91
93,3	90
101	87
85,7	69
84,1	63
72,3	61
69,7	60

Naloga 30

Rezultati analize linearne korelacijske odvisnosti povprečne letne porabe izdelka A (y) od povprečne letne porabe izdelka B (x) s programom SPSS so v spodnji tabeli:

Rezultati linearne korelacije in regresije

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	0,840(a)	0,706	0,608	11,334

a Predictors: (Constant), povprečna poraba izdelka B

Mode 1		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta	B	Std. Error
1	(Constant)	-8,647	38,282		-0,226	0,836
	Povprečna poraba izdelka B	0,734	0,273	0,840	2,684	0,075

a Dependent Variable: povprečna poraba izdelka A

- Izpišite in vsebinsko pojasnite vse kazalce linearne korelacije in regresije.
- Ocenite povprečno letno porabo izdelka A (y) pri povprečni letni porabi izdelka B enaki $x = 180$. Izračunajte točkovno oceno. Za intervalno oceno upoštevajte verjetnost 95 % (ter da je popravek h_1 enak 0).

6 Osnove vzorčenja

Na podlagi vzorčnih podatkov verjetnostnega vzorca lahko opravimo postopek statističnega sklepanja, ko rezultat iz vzorca posplošimo na statistično množico. Ta del statistike se imenuje inferenčna statistika. Vzorčni podatki nam služijo za izračun ocene statističnega parametra v statistični množici. Zavedamo se, da ta *vzorčna oziroma točkovna ocena* najverjetneje ni natančna vrednost parametra v statistični množici, ampak blizu nje. Po izračunu točkovnih ocen izračunamo *intervalne ocene*, imenovane *intervali zaupanja*, z določeno verjetnostjo, ki je rečemo *stopnja zaupanja* (Heumann idr., 2016).

Postopek, ki ga uporabimo za izračun intervala zaupanja, je odvisen od želene stopnje zaupanja, od tega, ali razpolagamo z informacijami o porazdelitvi obravnavane spremenljivke v statistični množici (na primer znan standardni odklon v statistični množici) ter od vzorca in njegove velikost (Holmes idr., 2018).

V okviru tega dela predmeta Statistika se ukvarjamo z *intervali zaupanja za aritmetično sredino, strukturni odstotek in total*, v velikih in malih vzorcih.

Intervali zaupanja so eden od načinov za oceno statističnega parametra v statistični množici. Drug način statističnega sklepanja je *preverjanje hipoteze* o vrednosti določenega statističnega parametra. Na primer, prodajalec avtomobilov trdi, da njegov nov majhen tovornjak v povprečju porabi le 6 litrov na 100 prevoženih kilometrov; izobraževalna agencija trdi, da njen način poučevanja pomaga 90 % študentom, da dobijo oceno 8 ali

več; podjetje pravi, da ženske menedžerke v njihovem podjetju v povprečju zaslužijo 60.000 EUR na leto ipd.

Ta postopek se imenuje *preizkušanje hipotez*. Testiranje hipoteze vključuje zbiranje podatkov iz slučajnega vzorca in vrednotenje zbranih podatkov. Na podlagi analize podatkov se odločimo, ali obstaja dovolj dokazov za zavrnitev ničelne hipoteze (H_0) ali ne.

Ničelna hipoteza, H_0 , je trditev, da ni razlike med vzorčno aritmetično sredino (ali vzorčnim strukturnim odstotkom ali vzorčno oceno totala statistične množice) in aritmetično sredino (ali strukturnim odstotkom ali totalom) statistične množice. Z drugimi besedami, razlika je enaka 0 oziroma vzorčna ocena parametra in parameter v statistični množici sta enaka.

Alternativna ali raziskovalna hipoteza, H_1 , pa je trditev o vrednosti statističnega parametra v statistični množici, ki je nasprotna trditvi H_0 , in menimo, da drži takrat, ko H_0 ne moremo sprejeti oziroma ko jo zavrremo (Agresti in Finaly, 2009). Za to, da zavrremo ničelno hipotezo, običajno potrebujemo 90 ali več odstotno verjetnost, da je to pravilna odločitev (oziroma 10 ali manj odstotno verjetnost), da se zmotimo, če H_0 zavrremo, kar imenujemo *stopnja tveganja ali stopnja značilnosti preizkusa*. Običajno pri preizkušanju hipoteze H_0 uporabljamo kot najvišjo še dopustno 5 %-no stopnjo tveganja.

Naloga 31

V slučajnem vzorcu $n = 200$ kupcev je bila povprečna poraba izdelka 20,8 kosov izdelka v časovni enoti, nepristranska ocena variance pa 38,44 kosov². Izračunajte 95 %-ni interval zaupanja za povprečno porabo izdelka v osnovni statistični množici.

Naloga 32

Na zavodu za zaposlovanje, na katerem je prijavljenih 10.000 nezaposlenih oseb, so želeli ugotoviti odstotek (%) nezaposlenih, ki prejemajo plačila na osnovi »dela na črno«. Iz slučajnega vzorca 625 nezaposlenih so ugotovili, da 125 ljudi dela na črno. Upoštevajte stopnjo tveganja 0,20.

Naloga 33

V tabeli so podatki o številu delavcev v podjetjih neke panoge, $n = 20$, ki smo jih zajeli v slučajni vzorec.

Število delavcev	Število podjetij
Od 1 do 5	3
Od 6 do 10	4
Od 11 do 15	2
Od 16 do 20	1
Od 21 do 25	1
Od 26 do 30	5
Od 31 do 35	2
Od 36 do 40	2

Z 99 %-no verjetnostjo določite povprečno število zaposlenih delavcev v vseh podjetjih te gospodarske panoge.

Naloga 34

Ali lahko trdimo, da delavci izdelajo povprečno 100 izdelkov v časovni enoti, če smo za 50 slučajno izbranih delavcev ugotovili, da izdelajo povprečno 102 izdelka v časovni enoti in je nepristranska ocena variance 6,25? Upoštevajte stopnjo tveganja 0,05.

Naloga 35

S strojem proizvajamo nek izdelek. Stroj naj bi bil nastavljen tako, da je povprečna dolžina izdelkov 20 cm. Na osnovi slučajnega vzorca $n = 5$ izdelkov smo ugotovili, da je vzorčna aritmetična sredina 20,3 cm in $s = 0,2$ cm. Upoštevajte, da je $\alpha = 1$ %. Ali rezultati kažejo, da stroj ni pravilno nastavljen?

Naloga 36

Od 75 naključno izbranih kupcev nekega izdelka smo ugotovili, da jih je 55 zadovoljnih s kakovostjo. Na načilnostni ravni 0,10 preizkusite domnevo, da je tri četrtine vseh kupcev zadovoljnih s kakovostjo izdelka.

Naloga 37

V slučajnem vzorcu smo zajeli 10 študentov neke fakultete, za katere velja, da v povprečju porabijo 45 % svojega prostega časa za športne aktivnosti, standardni odklon pa je enak 25 %. Ali lahko trdimo, da študenti te fakultete porabijo v povprečju polovico svojega prostega časa za športne aktivnosti? Upoštevajte 1 %-no tveganje.

Naloga 38

V banki „X“ je bilo v slučajni vzorec izbranih 340 imetnikov vrednostnih papirjev, ki imajo v vrednostne papirje vložene naslednje zneske:

Znesek v 10 ² EUR	Število oseb
Od 10 do pod 20	45
od 20 do pod 40	132
od 40 do pod 80	92
od 80 do pod 120	41
od 120 do pod 200	22
od 200 do pod 500	8
Skupaj	340

- S 95 % verjetnostjo določite interval zaupanja za povprečni znesek, ki so ga v vrednostne papirje vložili imetniki vrednostnih papirjev pri tej banki.
- Če je v opazovani banki registriranih 3.648 lastnikov vrednostnih papirjev, ocenite z 90 % verjetnostjo skupni znesek vlog v vrednostne papirje pri tej banki.
- Na ravni značilnosti $\alpha = 0,10$ preverite domnevo, da je povprečen znesek, ki so ga lastniki vložili v vrednostne papirje, 7.500 EUR.
- Z 99 % verjetnostjo preizkusite domnevo, da je % imetnikov vrednostnih papirjev, ki imajo zneske vrednostnih papirjev od 40.10² do 120.10² EUR 35 % oz. največ 35 %.
- Z 80 % verjetnostjo preizkusite domnevo, da je skupni znesek vseh imetnikov vrednostnih papirjev 20.000.000 EUR.

Naloga 39

V knjižnem klubu so za devet naključno izbranih oseb zbrali podatke o višini letnih zneskov za nakup knjig. Zneski so:

600 800 1.200 900 400 800 700 500 400 EUR

- a) Z 80 % verjetnostjo ocenite povprečen letni znesek za nakup knjig.
- b) Z 10 % tveganjem preizkusite domnevo, da je skupni znesek za nakup knjig v opazovanem knjižnem klubu, ki ima skupno 2.860 članov, 1.800.000 EUR.

Naloga 40

Na ravni značilnosti $\alpha = 0,10$ preizkusite domnevo, da je povprečna višina denarnih sredstev 120 vlagateljev 5.000 EUR, če smo iz te skupine naključno izbrali 36 vlagateljev in zanje ugotovili, da je

$$\sum y_i = 187.200 \quad \text{in} \quad \sum (y_i - \bar{y})^2 = 28.000.000.$$

Naloga 41

Med 250 podjetji neke regije, ki smo jih zajeli v slučajni vzorec, je 145 takih, ki imajo do 100 zaposlenih. Ali lahko pri 5 %-ni stopnji značilnosti preizkusa trdimo, da ima polovica vseh podjetij te regije do 100 zaposlenih?

Naloga 42

V nekem podjetju so na osnovi naključnega vzorca 150 ljudi ugotovili, da je njihova povprečna zamuda na delo 12,5 minut, nepristranska ocena standardnega odklona pa znaša 5 minut.

- a) Izračunajte in pojasnite 85 %-ni interval zaupanja za povprečno zamujen čas vseh delavcev.
- b) Pri 1 %-ni stopnji značilnosti preizkusa preverite domnevo, da je povprečna zamuda delavcev v podjetju enaka 10 minut.

Naloga 43

V šestih naključno izbranih podjetjih neke regije je med zaposlenimi tak odstotek žensk: 23 %, 47 %, 5 %, 45 %, 65 % in 12 %. Izračunajte in pojasnite interval zaupanja za povprečen odstotek žensk med zaposlenimi v vseh podjetjih. Oceno napravite s 5 % tveganjem.

Naloga 44

Naključno smo anketirali 40 gospodinjstev. Skupni dohodek vseh 40 gospodinjstev skupaj je bil 2.750.000 d.e., nepristranska ocena standardnega odklona pa 460 d.e. S 95 %-no verjetnostjo ocenite povprečni dohodek na gospodinjstvo. Napravite tudi grafični prikaz.

Naloga 45

Anketirali smo sedem naključno izbranih študentov v nekem študentskem domu. Povprečni mesečni zaslužek sedmih anketiranih študentov je 4.500 d.e. in $s = 1.200$ d.e. S 5 %-nim tveganjem izračunajte intervalno oceno za povprečni zaslužek študentov v študentskem domu.

Naloga 46

Vsebnost nikotina v štirih cigaretah neke znamke, merjena v miligramih, je 21, 19, 23 in 19. Pri stopnji tveganja 5 % preizkusite domnevo, da je povprečna vsebnost nikotina v cigaretah opazovane znamke enaka 22 miligramom.

Naloga 47

V trgovini X so za naključno izbranih pet oseb zbrali podatke o višini celotnega nakupa na določen dan. Podatki so naslednji:

Kupec	Znesek v EUR
1.	186
2.	216
3.	221
4.	266
5.	227
Skupaj	1.116

- Določite 95 %-ni interval zaupanja za povprečno zapravljen znesek naključno izbranih oseb v trgovini X.
- Pri stopnji tveganja 10 % preizkusite domnevo, da je povprečno zapravljen znesek naključno izbranih oseb v trgovini X več kot 220 EUR.

Naloga 48

Raziskovalci so v slučajni vzorec zajeli $n = 7$ okuženih oseb s SARS-CoV-2. Zanimalo jih je, koliko dni so bile prisotne posledice slabega počutja s SARS-CoV-2 naključno izbranih oseb.

Podatki o slabem počutju v številu dni okuženih oseb s SARS-CoV-2 so v tabeli:

Oseba	Število dni
1	25
2	30
3	21
4	60
5	17
6	14
7	24

- Določite 80 %-ni interval zaupanja za povprečno število dni slabega počutja pri okuženih osebah s SARS-CoV-2.
- Pri stopnji tveganja 5 % preizkusite domnevo, da je povprečno število dni slabega počutja pri okuženih osebah s SARS-CoV-2 največ 30 dni.

7 Časovne vrste

Dinamiko v časovnih vrstah lahko analiziramo s pomočjo relativnih števil, ki jih imenujemo indeksi. Pri izračunu indeksnih števil medsebojno primerjamo dva ali več istovrstnih podatkov. Indeksna števila lahko izračunamo iz absolutnih vrednosti, pri primerjavi lahko uporabimo statistične koeficiente, strukturne odstotke ipd. (Mišić, 2019). Tukaj obravnavamo enostavne indekse, medtem ko z agregatnimi indeksi analiziramo relativne spremembe ekonomskih količin, ki so izražene vrednostno ali količinsko in jih na tem mestu ne obravnavamo (Tominc, 2016).

Ločimo:

- a) indekse s stalno osnovo,
- b) verižne indekse,
- c) koeficiente dinamike in stopnje rasti.

Indeksi s stalno osnovo izražajo posredno odstotno spremembo med vrednostmi členov statistične vrste, ki jih primerjamo z enim od členov, ki ga izberemo za osnovo ali bazo, medtem ko *verižni indeksi* izražajo posredno odstotno spremembo med dvema zaporednima členoma *časovne* statistične vrste (Tominc, 2016). Pri verižnih indeksih vrednost člena y_t v časovni statistični vrsti vedno primerjamo s predhodnim členom y_{t-1} .

Koeficient dinamike izraža odnos med dvema zaporednima členoma osnovne statistične vrste, *stopnja rasti* pa pokaže neposredno odstotno spremembo med dvema zaporednima členoma osnovne statistične vrste – pove nam, za koliko odstotkov se vrednost člena y_t razlikuje od vrednosti člena y_{t-1} (Tominc, 2016).

Povprečno vrednost v časovni statistični vrsti opredelimo s *povprečno stopnjo rasti*, ki se izraža v % (na primer, povprečna stopnja rasti plač v Sloveniji v zadnjih 10 mesecih, povprečna stopnja rasti košarice izbranih cen v zadnjem letu ipd.). Povprečno stopnjo rasti v časovni statistični vrsti izračunamo bodisi iz *povprečnega koeficienta dinamike* ali *povprečnega verižnega indeksa*. Pri tem izhajamo iz postopka izračuna *geometrijske sredine* (in ne aritmetične sredine!). Povprečno stopnjo rasti opazovane spremenljivke v preteklem obdobju lahko uporabimo tudi za ocenjevanje vrednosti spremenljivke v prihodnjih časovnih enotah – *napovedovanje vrednosti spremenljivke v prihodnjih časovnih enotah*.

Prvi korak pri analizi katerekoli časovne vrste je običajno grafična predstavitev podatkov – o grafičnem prikazu časovne statistične vrste smo govorili v prvem poglavju. Iz grafičnega prikaza lahko ocenimo, katere komponente ima časovna vrsta: *osnovno smer razvoja – trend, sezonsko ali periodično komponento* ter *slučajne vplive* (Brockwell in Davis, 2016). Tukaj bomo obravnavali komponento trenda, ki jo analitično opisujemo s *trendno funkcijo* (omejili se bomo predvsem na linearno funkcijo) ter na sezonsko komponento, ki jo opisujemo s *sezonskimi ali periodičnimi indeksi*.

Naloga 49

Definirajte oziroma pojasnite naslednje pojme: indeksi, vrste indeksov, koeficient dinamike, stopnje rasti, statistični koeficienti.

Naloga 50

Število in naravno gibanje prebivalstva na opazovanem področju v obdobju zadnjih sedmih let je opisano v tabeli – (vsi podatki so v 10^4 ljudi):

Leto	Srednje štev. preb. (30.6.)	Živorojeni	Umrli
1	198	1,94	1,93
2	198	1,89	1,89
3	199	1,87	1,86
4	198	1,82	1,89
5	198	1,78	1,90
6	198	1,75	1,88
7	199	1,81	1,85

- a) Analizirajte relativne spremembe v številu živorojenih v opazovanem obdobju:
- glede na leto 3,
 - od leta do leta.
- b) Izračunajte koeficiente rodnosti (natalitete) in umrljivosti (mortalitete). Komentirajte rezultate.

Naloga 51

Letne stopnje rasti pridelave mesa v kg na prebivalca na opazovanem področju v obdobju zadnjih devetih let so bile:

Leto	1	2	3	4	5	6	7	8	9
S _t %	–	+18,0	–2,1	+2,7	+2,5	+2,5	–5,0	+7,0	–5,8

- a) Pojasnite največjo pozitivno in največjo negativno stopnjo rasti opazovanega pojava.
- b) Zapišite in pojasnite relativne spremembe v pridelavi mesa v opazovanih letih:
- z vrsto verižnih indeksov,
 - z vrsto letnih koeficientov dinamike,
 - z vrsto indeksov s stalno osnovo v letu 4.
- c) Če je bila pridelava mesa na prebivalca v letu 4 enaka 84,5 kg, izračunajte pridelavo mesa na prebivalca v preostalih letih.

Naloga 52

V opazovani organizaciji so podatki o številu zaposlenih in vrednosti proizvodnje po mesecih zapisani v naslednji časovni vrsti:

Mesec	I	II	III	IV	V
Vrednost proizvodnje v 10 ⁶ EUR	652	730	840	752	–
Štev. zaposlenih na začetku meseca	214	240	226	208	200

Kolikšna je bila povprečna mesečna vrednost proizvodnje v opazovani organizaciji na 10 zaposlenih?

Naloga 53

Za opazovano organizacijo imate na razpolago podatke o povprečni plači zaposlenih v prvih petih mesecih leta:

Mesec	I	II	III	IV	V
Povprečna mesečna plača zaposlenih (v d.e.)	2.600	2.800	3.100	2.900	3.200
Mesečna stopnja rasti vrednosti prodaje S_t %	–	+15	–10	+8	–12

- Izračunajte in pojasnite, kako so se relativno spreminjale plače glede na zadnji opazovani mesec.
- Izračunajte in pojasnite, kako se je spreminjala vrednost prodaje glede na prvi opazovani mesec.
- Ali so se v opazovanem podjetju v povprečju bolj dvigovale plače ali vrednost prodaje?
- Napovejte povprečno mesečno plačo zaposlenih v oktobru, upoštevajoč povprečno stopnjo rasti plač v prvih petih opazovanih mesecih.

Naloga 54

V preglednici so podatki o številu promocijskih akcij novih izdelkov nekega podjetja, po regijah, v nekem časovnem obdobju ter podatki o porabljenih denarnih sredstvih (v d.e.) za te aktivnosti:

Regija	A	B	C	D	E	F	G
Št. promocijskih akcij	21	25	14	17	25	12	18
Porabljena denarna sredstva (v d.e.)	2.314	3.100	2.100	3.000	2.600	1.750	2.950

- Izračunajte porabljena denarna sredstva na promocijsko akcijo po posameznih regijah in povprečno porabljena denarna sredstva na promocijsko akcijo za vse regije skupaj.
- Izračunajte, v kateri regiji je vrednost porabljenih denarnih sredstev v primerjavi z regijo G najmanjša in koliko odstotkov porabljenih sredstev v tej regiji predstavljajo porabljena sredstva v ostalih regijah.

Naloga 55

V preglednici so podatki o številu izdanih gradbenih dovoljenj na nekem območju:

Leto	1	2	3	4	5	6	7
Št. izdanih gradb. dovoljenj	115	224	118	400	350	320	300

- Za koliko odstotkov se je število izdanih gradbenih dovoljenj v letu 4 razlikovalo od števila v predhodnem letu? Kako imenujemo izračunano vrednost?
- Koliko odstotkov števila izdanih gradbenih dovoljenj iz leta 1 predstavlja število izdanih dovoljenj v letu 7? Kako imenujemo izračunano vrednost?
- Statistično vrsto prikažite grafično.
- V letu 6 je bilo med izdanimi gradbenimi dovoljenji 175 dovoljenj za enodružinske hiše, 45 dovoljenj za gradnjo dvojčka, 63 dovoljenj za manjše bloke in 37 dovoljenj za poslovne objekte. Grafično prikažite absolutno in relativno strukturo izdanih gradbenih dovoljenj v letu 6.

Naloga 56

V preglednici so podatki o stopnji rasti plač zaposlenih oseb v nekem podjetju.

Leto	1	2	3	4
S _t %	/	+18,4	+11,7	-9,2

- Analizirajte, kako so se relativno spreminjale plače glede na leto 2.
- Izračunajte, kolikšne so bile povprečne plače zaposlenih v obravnavanih letih, če je bila plača leta 3 enaka 170,5 d.e.
- Izračunajte in pojasnite povprečno stopnjo rasti plač v obravnavani 4-letni časovni vrsti.
- Napovejte povprečno raven plač zaposlenih v letu 7, če upoštevate povprečno stopnjo rasti v opazovanem obdobju.

Naloga 57

Na opazovanem področju smo zabeležili naslednje število nočitev tujih gostov:

Leto	I–IV	V–VIII	IX–XII	Skupaj
1	1.200	2.500	1.000	4.700
2	1.100	2.600	1.200	4.900
3	1.000	2.800	1.100	4.900
4	1.300	2.500	1.200	5.000
5	1.200	2.700	1.300	5.200

- Letne podatke, ki prikazujejo skupno letno število nočitev, narišite v linijskem grafikonu, s prostoročno metodo določite osnovno smer razvoja pojava; z analitično metodo določite parametre funkcije; ocenite število nočitev v prvem prihodnjem letu.
- Izračunajte sezonske indekse, upoštevajte oceno za število nočitev v prihodnjem letu na osnovi funkcije trenda (rezultat naloge a) in izračunane sezonske indekse ter predvidite število nočitev po sezonah v prihodnjem letu.

Naloga 58

V podjetju, ki se ukvarja s popravilom strojev, načrtujejo stroške v zvezi z njihovim delom. Na osnovi četrletnih podatkov za 4 zaporedna leta je podjetje izračunalo funkcijo trenda za stroške ($t = \text{leto}$), in sicer:

$$\hat{Y} = 58,5 + 2,1t$$

- Ocenite stroške za prvo prihodnje leto.
- Sezonski indeks za zadnje četrletje znaša $SI_4 = 225,5\%$. Kaj ta indeks vsebinsko pomeni?
- Ocenite stroške popravil v zadnjem četrletju prvega prihodnjega leta.

Naloga 59

V preglednici so podatki o številu izdanih gradbenih dovoljenj na nekem območju v zadnjih sedmih letih:

Leto	1	2	3	4	5	6	7
Št. izdanih gradb. dovoljenj	115	120	118	125	130	125	145

- a) Časovno vrsto prikažite grafično in izračunajte koeficiente ustrezne funkcije trenda.
- b) Kolikšno število izdanih gradbenih dovoljenj lahko pričakujemo leta 8, če upoštevamo osnovno smer razvoja pojava?

Naloga 60

Na osnovi letnih podatkov iz zadnjih petih let so v nekem podjetju izračunali enačbo funkcije trenda za število novo zaposlenih po letih:

$$\hat{y} = 35 \cdot t^{0,2}$$

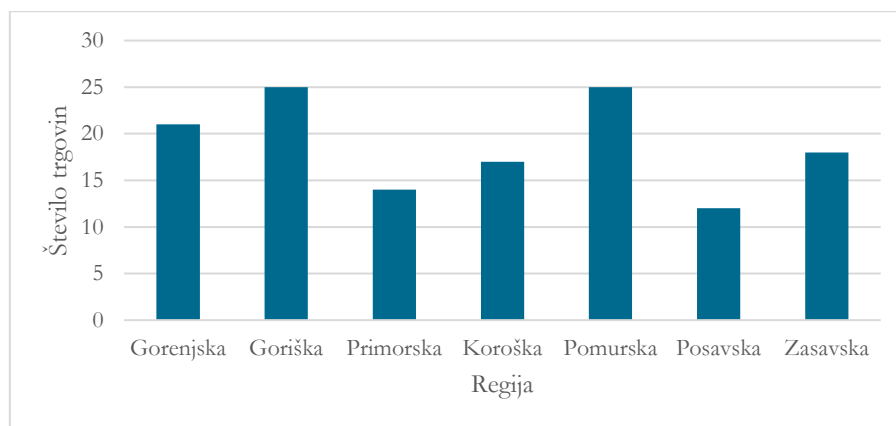
- a) Kako imenujemo zgornjo funkcijo?
- b) Napovejte število novo zaposlenih v šestem zaporednem letu po četrletjih, če so izračunani sezonski indeksi $SI_1 = 65$, $SI_2 = 155$, $SI_3 = 160$ in $SI_4 = 20$. Kaj vsebinsko pomeni sezonski indeks za zadnje četrletje?

8 Rešitve nalog

V tem poglavju prikazujemo rešitve izbranih nalog.

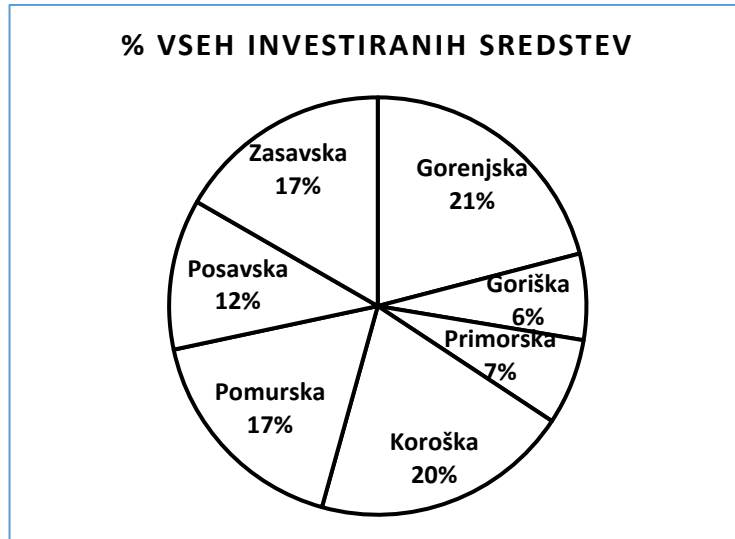
Urejanje in prikazovanje podatkov

Naloga 2b

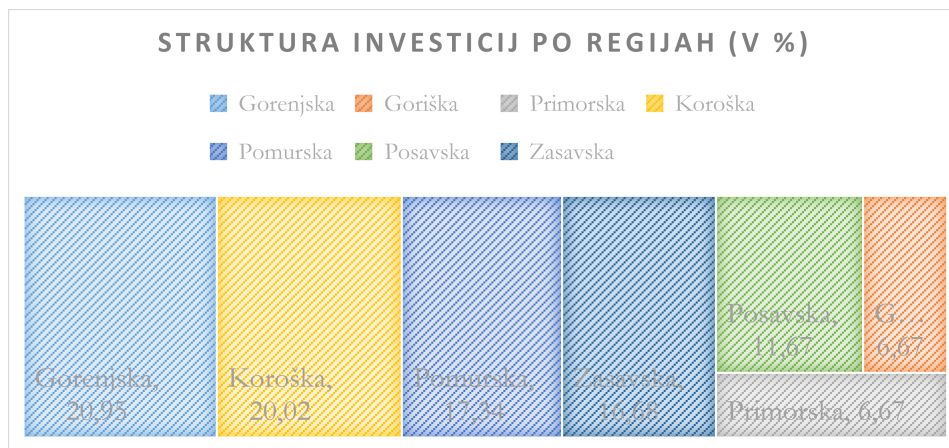


Naloga 2d

V tem primeru smo narisali strukturalni krog ter strukturalni pravokotnik, oboje s programom Excel.



ali

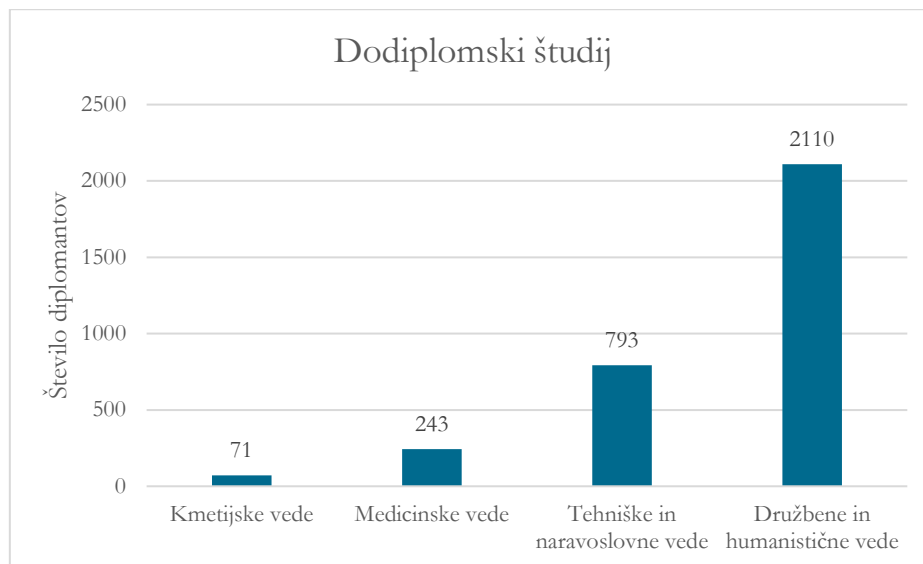


Naloga 3

Mesec	I.	II.	III.	IV.	V.	Skupaj
Vrednost prodaje (v 10 ³ EUR)	2.600	3.800	4.100	2.900	3.200	16.600
Število zaposlenih trgovskih potnikov	30	75	82	80	90	357
Vrednost prodaje (v 10 ³ EUR) na trgovskega potnika	86,67	50,67	50	36,25	35,56	46,50 – povprečna vrednost prodaje na trgovskega potnika v obdobju petih mesecev

Naloga 5b

Prikaz stvarne statistične vrste s stolpci.

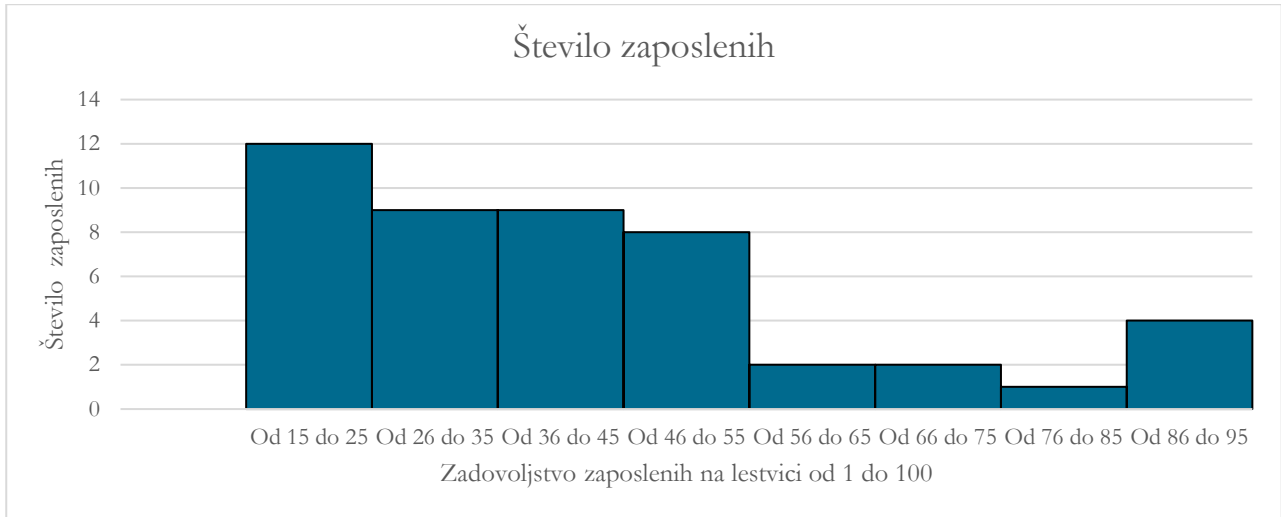


Naloga 6b

Zadovoljstvo	f_k	F_k
Od 15 do 25	12	12
Od 26 do 35	9	21
Od 36 do 45	9	30
Od 46 do 55	8	38
Od 56 do 65	2	40
Od 66 do 75	2	42
Od 76 do 85	1	43
Od 86 do 95	4	44

Naloga 6c

Histogram frekvenčne porazdelitve



Naloga 7

Število ur študija	f_k	F_k
Od 40 do pod 50	25	25
Od 50 do pod 60	50	75
Od 60 do pod 70	100	175
Od 70 do pod 80	75	250
Od 80 do pod 90	40	290
Od 90 do pod 100	10	300
Skupaj	300	

Naloga 9c

Hitrost v km/h	Število osebnih vozil
Od 41 do 50	4
Od 51 do 60	25
Od 61 do 70	24
Od 71 do 80	15
Od 81 do 90	2
Skupaj	70

Naloga 10c, 10d

Število zaposlenih	f_k	F_k	$f\%_{0k}$	$F\%_{0k}$
Od 11 do 40	15	15	15,5	15,5
Od 41 do 55	17	32	17,5	33
Od 56 do 70	23	55	23,7	56,7
Od 71 do 85	25	80	25,8	82,5
Od 86 do 100	17	97	17,5	100 %
Skupaj	97		100 %	

Naloga 11b, 11c

Število vozil	f_k	F_k	$f\%_{0k}$
Od 0 do 4	9	9	19,6
Od 5 do 9	15	24	32,7
Od 10 do 14	8	32	17,4
Od 15 do 19	6	38	13
Od 20 do 24	4	42	8,7
Od 25 do 29	2	44	4,3
Od 30 do 34	2	46	4,3
Skupaj	46		100 %

Deskriptivna statistika

Naloga 13a

Podatke najprej uredimo v ranžirno vrsto:

y_i	22	26	29	30	33	34	38	39	40	41	43	45
R_i	1	2	3	4	5	6	7	8	9	10	11	12

Organizacija A: $Q = Q_3 - Q_1 = 41 \text{ minut} - 29 \text{ minut} = 12 \text{ minut}$
 $Me = 36 \text{ minut}$

50 % zaposlenih na sredini ranžirne vrste se glede na potreben čas za izdelavo izdelka med seboj razlikuje za največ 12 minut; 50 % zaposlenih na sredini ranžirne vrste je porabilo več kot 12 pa do največ 41 minut za izdelavo izdelka.

50 % zaposlenih je porabilo do 36 minut za izdelavo izdelka, 50 % zaposlenih pa več kot toliko.

Naloga 13b

Povprečno porabljen čas za izdelavo enega izdelka:

$$\text{Organizacija A:} \quad \bar{y} = \frac{420}{12} = 35 \text{ minut}$$

$$\text{Organizacija B:} \quad \bar{y} = \frac{24124}{730} = 33,05 \text{ minut}$$

Naloga 13c

$$\text{Organizacija A:} \quad \text{VR} = 45 - 22 = 23 \text{ minut}$$

Zaposleni se glede na porabljen čas za izdelavo izdelka med seboj razlikujejo za največ 23 minut.

Naloga 13d

Varianca in standardni odklon

$$\text{Organizacija A:} \quad \sigma^2 = \frac{15266}{9} - 35^2 = 47,17 \text{ minut}^2$$

$$\sigma = 6,87 \text{ minut}$$

$$\text{Organizacija B:} \quad \sigma^2 = \frac{817296}{730} - 33,05^2 = 27,51 \text{ minut}^2$$

$$\sigma = 5,25 \text{ minut}$$

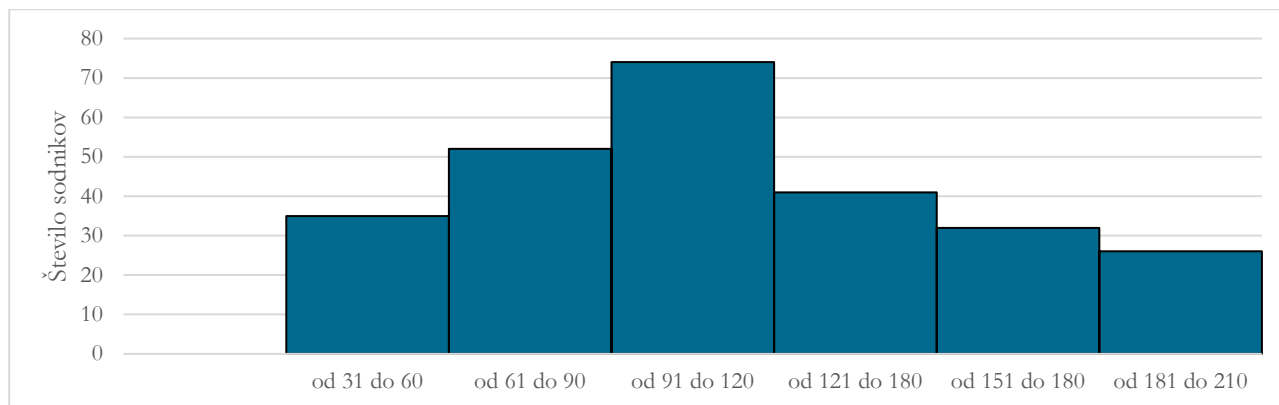
Naloga 13e

$$\text{Organizacija A:} \quad \text{KV \%} = \frac{6,87}{35} 100 = 19,6 \%$$

$$\text{Organizacija B:} \quad \text{KV \%} = \frac{5,25}{33,05} 100 = 15,9 \%$$

Naloga 14b

Porazdelitev je asimetrična v desno stran.



Naloga 14c

$$\bar{y} = \frac{29260}{260} = 112,54 \text{ zadev}$$

$$\sigma^2 = \frac{3815495}{260} - 112,54^2 = 2.010,08 \text{ zadev}^2$$

$$\sigma = 44,83 \text{ zadev}$$

$$\text{KV}\% = 39,83 \%$$

Delež standardnega odklona v aritmetični sredini znaša 39,83 %.

Naloga 15a

Ranžirna vrsta:

y_i 9 12 16 41 60 80 100 120 150

Q1: mesto v ranžirni vrsti: 2,5. Vrednost: $(12 + 16)/2 = 14$ (v 10^2 EUR)

Q3: mesto v ranžirni vrsti: 7,5. Vrednost: $(100 + 120)/2 = 110$ (v 10^2 EUR)

50 % vlagateljev, ki so vložili srednje velike zneske, je vložilo zneske v višini več kot 14 pa do največ 110 (v 10^2 EUR).

$$\text{Me} = 60$$

Vlagatelji, ki spadajo med 50 % tistih, ki so vložili najmanjše zneske, so vložili zneske v višini do 60 (v 10^2 EUR).

Naloga 15b

$$\text{Banka X: } KV \% = \frac{47,90}{65,33} 100 = 73,3 \%$$

$$\text{Banka Y: } KV \% = \frac{2650}{7530} 100 = 35,2 \%$$

Vlagatelji v banki X se med seboj (in od povprečja) bolj razlikujejo kot pa vlagatelji v banki Y.

Naloga 16c

$$\bar{y} = 23,06 \text{ vozil}$$

Naloga 16d

$$\sigma = 11,58 \text{ zadev}$$

$$KV \% = 50,2 \%$$

Naloga 17a

$$KV \frac{BDP}{\text{preb}} \% = 35 \%$$

$$KV \%_{\text{preb.-podj.}} \% = \frac{1,9}{11,25} 100 = 16,9 \%$$

Naloga 17b

Podatke uredimo v ranžirno vrsto:

Me = 12,34 % ljudi, ki se ukvarjajo s podjetništvom.

Naloga 18a

$Me = 2,925 \%$ zmanjšanja stroškov energije

Mo – ga ni mogoče opredeliti, saj se vsaka vrednost spremenljivke pojavlja samo enkrat.

$\bar{y} = 3,7625 \%$ zmanjšanja stroškov energije

Naloga 18b

$Q1 = 2 \%$ zmanjšanja stroškov energije

$Q3 = 5,7 \%$ zmanjšanja stroškov energije

$Q = 3,7 \%$ točk

Podjetja, ki ležijo v kvartilnem razmiku, so dosegla $\%$ zmanjšanja stroškov energije, ki je večji od 2% do največ $5,7 \%$, kar pomeni, da se glede na te spremenljivke podjetja razlikujejo za največ $3,7 \%$ -nih točk.

Naloga 18c

$KV_{\text{dobiček}} \% = 75 \%$

$KV_{\% \text{ zmanjšanja str.en.}} \% = \frac{2,14}{3,7625} 100 = 56,9 \%$

Naloga 19b

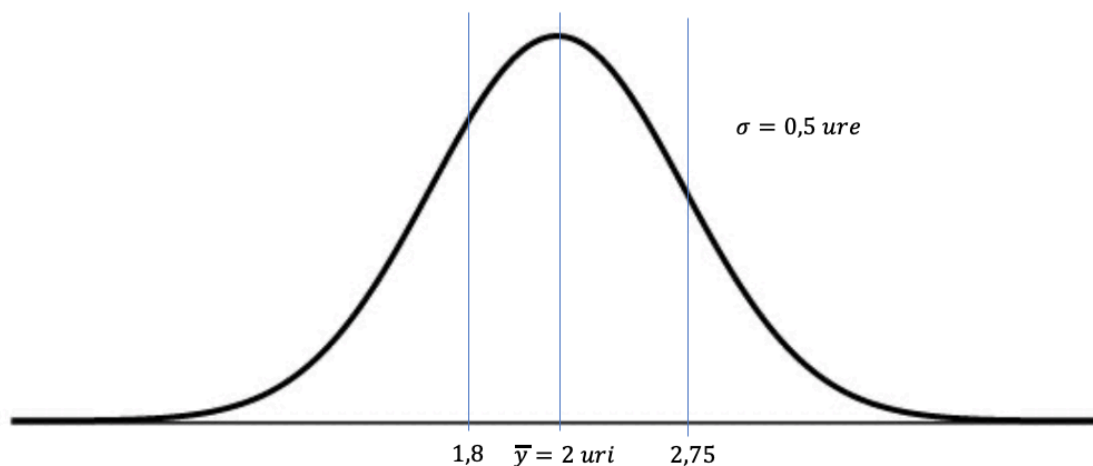
Podatke uredimo v ranžirno vrsto.

$Q1 = 50$ d.e.

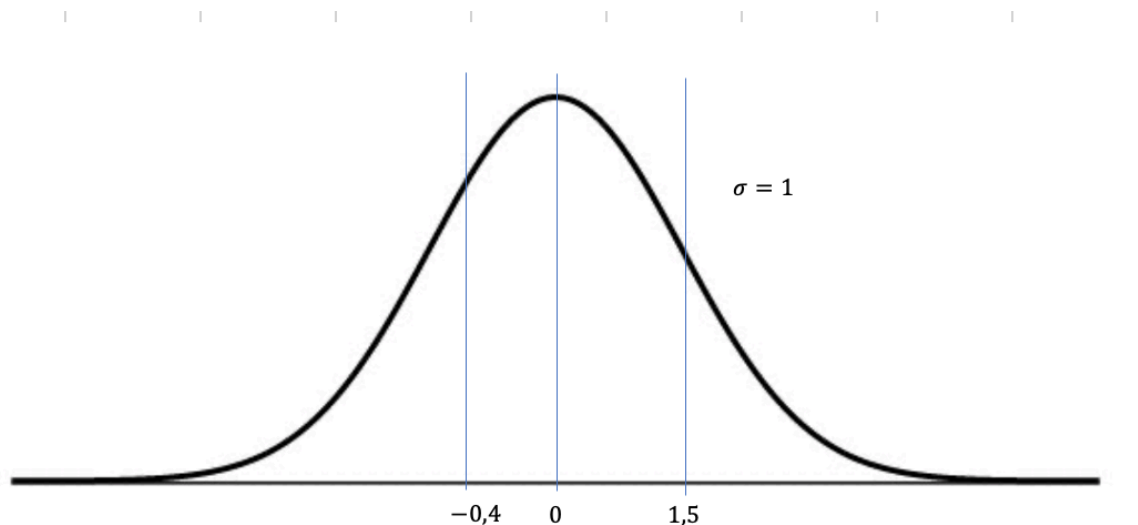
$Q3 = 850$ d.e.

Teoretične porazdelitve

Naloga 20a



Standardizirana normalna porazdelitev



$P(1,8 \text{ ure} < y < 2,75 \text{ ure}) = H(-0,4) + H(1,5) = 0,1554 + 0,4332 = 0,5886 = 58,86 \%$
Vrednost $H(z_i)$ odčitamo iz tabele: Ploščine $H(z)$ za standardizirano normalno porazdelitev.

Verjetnost, da je v naključno izbranem gospodinjstvu družinski računalnik v uporabi za igranje igrice med 1,8 in 2,75 urami na dan, je enaka 58,86 %.

Naloga 20b

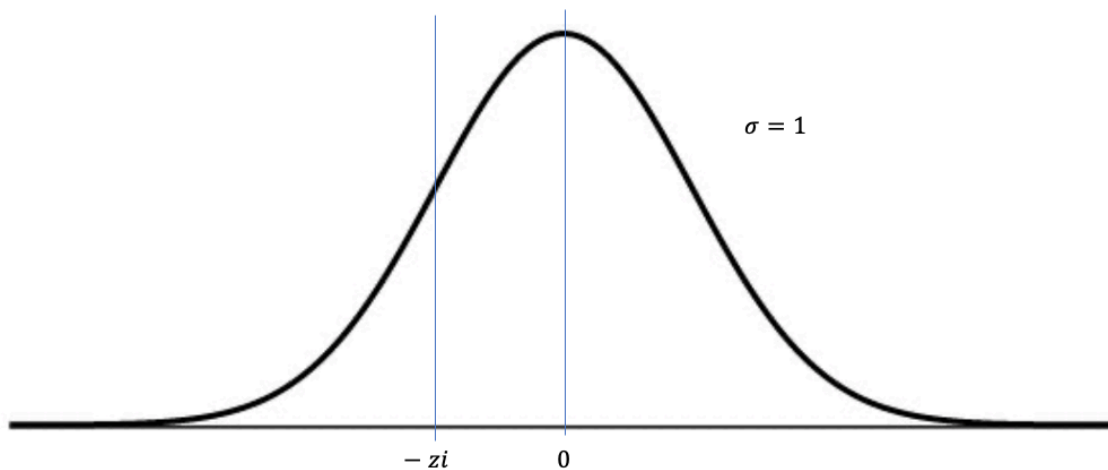
$$z_i = 6; H(6) = 0,5$$

$$P(y > 5) = 0,5 - H(6) = 0,5 - 0,5 = 0$$

Naloga 20c

V tem primeru je podana verjetnost:

$$P(y < z_i) = 0,25$$



$$H(z_i) = 0,25$$

Iz tabele Ploščine $H(z)$ za standardizirano normalno porazdelitev odčitamo vrednost z_i :

$$H(-0,67) = 0,2486 = 0,25$$

$$y_i = -0,67 \cdot 0,5 + 2 = 1,7 \text{ ure}$$

V 25 % gospodinjstev, kjer je čas uporabe družinskega računalnika za igranje igrvic najkrajši, je ta čas do 1,7 ure dnevno.

Naloga 21a

$$P(23 \text{ let} < y < 64 \text{ let}) = H(-1,00) + H(1,94) = 0,3413 + 0,4738 = 0,8151 = 81,51 \%$$

Naloga 21b

$$P(y > 50,8 \text{ let}) = 0,5 - H(1,00) = 0,5 - 0,3413 = 0,1587 = 15,87 \%$$

Naloga 21c

$$P(y < y_i) = 30 \%$$

$$H(0,52) = 0,1985 = 0,20$$

$$y_i = -0,52 \cdot 13,9 + 36,9 = 44,13 \text{ let}$$

Naloga 22a

$$P(y > 195 \text{ minut}) = 0,5 + H(-0,36) = 0,5 + 0,1406 = 0,6406 = 64,06 \%$$

Naloga 22b

Ta delež naročnikov pokriva interval $(\bar{y} \pm \sigma)$ oziroma 220 ± 70 . To pomeni, da pri tistih 68,3 % naročnikih, ki ležijo v simetričnem intervalu okoli aritmetične sredine, beležimo od 150 minut pa do 290 minut pogovora na mesec.

Enostavna regresijska analiza

Naloga 23a

Oblika: Linearna oblika

Smer med spremenljivkama: Pozitivna smer, kar pomeni, da z naraščanjem investicij v tehnologijo (x) v povprečju narašča ustvarjen dobiček (y).

Jakost povezanosti med spremenljivkama: obstaja močna povezanost med odvisno (ustvarjen dobiček) in neodvisno spremenljivko (investicije v tehnologijo).

Naloga 23b

Izračun parametrov regresijske premice:

Preden se lotimo izračuna regresijskih koeficientov b_0 in b_1 , izračunamo:

$$\sum x_i = 115 + 130 + 140 + 149 + 160 + 171 = 865$$

$$\sum x_i^2 = 115^2 + 130^2 + 140^2 + 149^2 + 160^2 + 171^2 = 126.767$$

$$\sum y_i = 328 + 330 + 390 + 361 + 421 + 400 = 2.230$$

$$\sum y_i^2 = 328^2 + 330^2 + 390^2 + 361^2 + 421^2 + 400^2 = 836.146$$

$$\sum x_i \cdot y_i = (115 \cdot 328) + (130 \cdot 330) + (140 \cdot 390) + \dots + (171 \cdot 400) = 324.769$$

$$\bar{x} = \frac{865}{6} = 144,1667$$

$$\bar{y} = \frac{2230}{6} = 371,6667$$

Oba regresijska koeficienta izračunamo po formuli:

$$b_1 = \frac{\frac{1}{n-1} \sum (x_i - \bar{x}) \cdot (y_i - \bar{y})}{\frac{1}{n-1} \sum x_i^2 - \bar{x}^2} = \frac{c_{xy}}{s_x^2} = \frac{(\sum x_i \cdot y_i) - n\bar{x}\bar{y}}{(n-1) \cdot s_x^2} = \frac{(\sum x_i \cdot y_i) - n\bar{x}\bar{y}}{(\sum x_i^2) - n\bar{x}^2}$$

$$b_0 = \bar{y} - b_1 \cdot \bar{x}$$

$$b_1 = \frac{(324769) - 6 \cdot 144,1667 \cdot 371,6667}{(126767) - 6 \cdot 144,1667^2} = 1,5887$$

$$b_0 = 371,6667 - 1,5887 \cdot 144,1667 = 142,6291$$

Izračunana regresijska koeficienta vstavimo v enačbo regresijske premice:

$$\hat{y}_i = b_0 + b_1 x_i$$

$$\hat{y} = 142,6291 + 1,5887 \cdot x_i$$

Pomen regresijskega koeficienta b_0 : Pri investicijah v tehnologijo $x = 0$ lahko v povprečju pričakujemo, da bo ustvarjen dobiček podjetja 142,6291 (v 10^3 EUR).

Pomen regresijskega koeficienta b_1 : Če se investicije v tehnologijo (x) povečajo za eno enoto (v 10^6 EUR), se ustvarjen dobiček (y) v povprečju poveča za 1,5887 (v 10^3 EUR).

Naloga 23c

$$x = 180 \text{ (v } 10^6 \text{ EUR)}$$

$$\hat{y} = 142,6291 + 1,5887 \cdot x_i$$

$$\hat{y}_{x=180} = 142,6291 + 1,5887 \cdot 180$$

$$\hat{y}_{x=180} = 428,649 \text{ (v } 10^3 \text{ EUR)}$$

Naloga 23d

Parameter, na osnovi katerega določimo smer in jakost linearne korelacijske odvisnosti, je korelacijski koeficient.

Izračunamo ga po formuli:

$$r_{xy} = \frac{(\sum x_i \cdot y_i) - n\bar{x}\bar{y}}{(n-1) \cdot s_x \cdot s_y}$$

Izračun s_x in s_y :

$$s_x^2 = \frac{1}{n-1} \cdot \sum (x_i - \bar{x})^2$$

$$s_x^2 = \frac{1}{5} \cdot [(115-144,1667)^2 + (130-144,1667)^2 + (140-144,1667)^2 + \dots + (171-144,1667)^2]$$

$$= 412,5667$$

$$s_x = \sqrt{412,5667} = 20,3117$$

$$s_y^2 = \frac{1}{n-1} \cdot \sum (y_i - \bar{y})^2$$

$$s_y^2 = \frac{1}{5} \cdot [(328-371,6667)^2 + (330-371,6667)^2 + (390-371,6667)^2 + \dots + (400-371,6667)^2]$$

$$= 1465,8667$$

$$s_y = \sqrt{1465,8667} = 38,2866$$

$$(\sum x_i \cdot y_i) - n \cdot \bar{x} \cdot \bar{y} = 324769 - 6 \cdot 144,1667 \cdot 371,6667 = 3277,2302$$

$$r_{xy} = \frac{3277,2302}{5 \cdot 20,3117 \cdot 38,2866} = 0,8428$$

Na osnovi rezultata ($r_{xy} = 0,8428$) vidimo, da obstaja močna povezanost med odvisno (ustvarjen dobiček) in neodvisno spremenljivko (investicije v tehnologijo). Smer povezanosti je pozitivna.

Naloga 23e

Izračunati moramo delež pojasnjene variance v skupni varianci za odvisno spremenljivko (determinacijski koeficient):

$$r_{xy}^2 = 0,8428^2 = 0,7103 \text{ oz. } 71,03 \%$$

Delež pojasnjene variance v skupni varianci za odvisno spremenljivko znaša 71,03 %.

Naloga 23f

Standardno napako ocene odvisne spremenljivke izračunamo po formuli:

$$s_{y,x} = \sqrt{\frac{(\sum y_i^2) - b_0(\sum y_i) - b_1(\sum x_i \cdot y_i)}{n - 2}}$$

$$s_{y,x} = \sqrt{\frac{836146 - (142,6291 \cdot 2230) - (1,5887 \cdot 324769)}{6 - 2}} = 23,036$$

Standardna napaka ocene odvisne spremenljivke je različna od 0, kar pomeni, da na ustvarjen dobiček (odvisna spremenljivka) poleg investicij v tehnologijo (neodvisna spremenljivka) vplivajo še druge spremenljivke in slučajni vplivi.

Naloga 23g

Intervalno oceno izračunamo po formuli:

$$\hat{y}_{x_0} \pm t_{n-2; \alpha/2} \cdot s_{y,x} \cdot \sqrt{1 + h_1}$$

$$\hat{y}_x = 428,649 \text{ (izračunano pri nalogi c)}$$

$$s_{yx} = 23,036 \text{ (izračunano pri nalogi f)}$$

Upoštevamo, da je pri $\gamma = 95 \%$, $\alpha = 5 \%$.

Izračun: $t_{n-2;\alpha/2} = t_{4;0,025} = 2,776$ (gledamo tabelo: Kritične vrednosti za t porazdelitev)
Upoštevamo, da je popravek h_1 enak 0.

Izračunamo intervalno oceno:

$$P(428,649 - 2,776 \cdot 23,036 \cdot 1 < y_{x=180} < 428,649 + 2,776 \cdot 23,036 \cdot 1) = 95 \%$$

$$P(364,701 < y_{x=180} < 492,597) = 95 \%$$

Pri investicijah v tehnologijo $x = 180$ (v 10^6 EUR) bo ustvarjen dobiček podjetja med 364,701 EUR in 492,597 (v 10^3 EUR), kar trdimo s 95 % verjetnostjo.

Naloga 24a

$$\Sigma x_i = 27,6$$

$$\Sigma x_i^2 = 165,52$$

$$\Sigma y_i = 977$$

$$\Sigma y_i^2 = 202.483$$

$$\Sigma x_i \cdot y_i = 5.776$$

$$\bar{x} = \frac{27,6}{5} = 5,52$$

$$\bar{y} = \frac{977}{5} = 195,4$$

Izračun regresijskih koeficientov:

$$b_1 = \frac{\frac{1}{n-1} \sum (x_i - \bar{x}) \cdot (y_i - \bar{y})}{\frac{1}{n-1} \sum x_i^2 - \bar{x}^2} = \frac{c_{xy}}{s_x^2} = \frac{(\sum x_i \cdot y_i) - n\bar{x}\bar{y}}{(n-1) \cdot s_x^2} = \frac{(\sum x_i \cdot y_i) - n\bar{x}\bar{y}}{(\sum x_i^2) - n\bar{x}^2}$$

$$b_0 = \bar{y} - b_1 \cdot \bar{x}$$

$$b_1 = \frac{(5776) - 5 \cdot 5,52 \cdot 195,4}{(165,52) - 5 \cdot 5,52^2} = 29,083$$

Pomen regresijskega koeficienta b_1 : Če se obrestna mera (x) poveča za eno enoto, se znesek kratkoročnih oblik varčevanja (y) v povprečju poveča za 29,083 d.e.

$$b_0 = 195,4 - 29,083 \cdot 5,52 = 34,862$$

Pomen regresijskega koeficienta b_0 : Pri obrestni meri $x = 0$ bi v povprečju znesek kratkoročnih oblik varčevanja znašal 34,862 d.e.

Izračunana regresijska koeficienta vstavimo v enačbo regresijske premice:

$$\hat{y}_i = b_0 + b_1 x_i$$

$$\hat{y} = 34,862 + 29,083 \cdot x_i$$

$$x = 7$$

$$\hat{y}_{x=7} = 34,862 + 29,083 \cdot x_i$$

$$\hat{y}_{x=7} = 34,862 + 29,083 \cdot 7$$

$$\hat{y}_{x=7} = 238,443$$

Standardno napako ocene odvisne spremenljivke izračunamo po formuli:

$$s_{y \cdot x} = \sqrt{\frac{(\sum y_i^2) - b_0(\sum y_i) - b_1(\sum x_i \cdot y_i)}{n - 2}}$$

$$s_{yx} = \sqrt{\frac{202483 - (34,862 \cdot 977) - (29,083 \cdot 5776)}{5-2}} = \sqrt{\frac{439,418}{3}} = 12,103$$

Standardna napaka ocene odvisne spremenljivke je različna od 0, kar pomeni, da na znesek kratkoročnih oblik varčevanja (odvisna spremenljivka) poleg različne obrestne mere (neodvisna spremenljivka) vplivajo še druge spremenljivke in slučajni vplivi.

Izračun intervalne ocene:

$$\hat{y}_{x=7} = 238,443$$

$$s_{yx} = 12,103$$

Pri $\gamma = 95\%$ je $\alpha = 5\%$.

Izračun: $t_{n-2;\alpha/2} = t_{3;0,025} = 3,182$ (gledamo tabelo: Kritične vrednosti za t porazdelitev)

Vrednost za h_1 izračunamo po formuli:

$$h_1 = \frac{1}{n} + \frac{(x_0 - \bar{x})^2}{\sum_i (x_i - \bar{x})^2} = \frac{1}{n} + \frac{(x_0 - \bar{x})^2}{\sum_i (x_i)^2 - (\sum x_i)^2 / n}$$

$$h_1 = \frac{1}{5} + \frac{(7 - 5,52)^2}{165,52 - 27,6^2/5} = \frac{1}{5} + 0,1663 = 0,3663$$

Izračunane vrednosti vstavimo v enačbo:

$$\hat{y}_{x_0} \pm t_{n-2; \alpha/2} \cdot s_{y \cdot x} \cdot \sqrt{1 + h_1}$$

$$238,443 \pm 3,182 \cdot 12,103 \cdot \sqrt{1,3663}$$

Intervalna ocena:

$$P(238,443 - 3,182 \cdot 12,103 \cdot 1,169 < y_{x=7} < 238,443 + 3,182 \cdot 12,103 \cdot 1,169) = 95 \%$$

$$P(193,423 < y_{x=7} < 283,463) = 95 \%$$

S 95%-no verjetnostjo ocenjujemo, da bo pri obrestni meri $x = 7 \%$ znesek kratkoročnih oblik varčevanja med 193,423 d.e. in 283,463 d.e.

Naloga 26b

$$\Sigma x_i = 319$$

$$\Sigma x_i^2 = 26791$$

$$\Sigma y_i = 122$$

$$\Sigma y_i^2 = 3264$$

$$\Sigma x_i \cdot y_i = 9012$$

$$\bar{x} = 63,8$$

$$\bar{y} = 24,4$$

Izračun regresijskih koeficientov:

$$b_1 = \frac{(9012) - 5 \cdot 63,8 \cdot 24,4}{(26791) - 5 \cdot 63,8^2} = 0,191$$

$$b_0 = 24,4 - 0,191 \cdot 63,8 = 12,214$$

Izračunana regresijska koeficienta vstavimo v enačbo regresijske premice:

$$\hat{y} = 12,214 + 0,191 \cdot x_i$$

Korelacijski koeficient izračunamo po formuli:

$$r_{xy} = \frac{(\sum x_i \cdot y_i) - n\bar{x}\bar{y}}{(n-1) \cdot s_x \cdot s_y}$$

Izračun s_x in s_y :

$$s_x^2 = \frac{1}{n-1} \cdot \sum (x_i - \bar{x})^2$$

$$s_x^2 = \frac{1}{4} \cdot [(25-63,8)^2 + (46-63,8)^2 + (120-63,8)^2 + (91-63,8)^2 + (37-63,8)^2] = 1609,7$$

$$s_x = \sqrt{1609,7} = 40,121$$

$$s_y^2 = \frac{1}{n-1} \cdot \sum (y_i - \bar{y})^2$$

$$s_y^2 = \frac{1}{4} \cdot [(15-24,4)^2 + (17-24,4)^2 + (35-24,4)^2 + (30-24,4)^2 + (25-24,4)^2] = 71,8$$

$$s_y = \sqrt{71,8} = 8,473$$

$$(\sum x_i \cdot y_i) - n \cdot \bar{x} \cdot \bar{y} = 9012 - 5 \cdot 63,8 \cdot 24,4 = 1228,4$$

$$r_{xy} = \frac{1228,4}{4 \cdot 40,121 \cdot 8,473} = 0,903$$

Determinacijski koeficient:

$$r_{xy}^2 = 0,903^2 = 0,815 \text{ oz. } 81,5 \%$$

Standardno napako ocene odvisne spremenljivke izračunamo po formuli:

$$s_{y \cdot x} = \sqrt{\frac{(\sum y_i^2) - b_0(\sum y_i) - b_1(\sum x_i \cdot y_i)}{n-2}}$$

$$s_{yx} = \sqrt{\frac{3264 - (12,214 \cdot 122) - (0,191 \cdot 9012)}{5-2}} = 4,187$$

Naloga 26c

$$x = 180$$

$$\hat{y} = 12,214 + 0,191 \cdot x_i$$

$$\hat{y}_{x=180} = 12,214 + 0,191 \cdot 180$$

$$\hat{y}_{x=180} = 46,594$$

Izračun intervalne ocene:

$$\hat{y}_{x=180} = 46,594$$

$$s_{y,x} = 4,187$$

$$\alpha = 5 \%$$

Izračun: $t_{n-2;\alpha/2} = t_{3;0,025} = 3,182$ (gledamo tabelo: Kritične vrednosti za t porazdelitev)

Upoštevamo, da je popravek h_1 enak 0.

Intervalna ocena:

$$P(46,594 - 3,182 \cdot 4,187 \cdot 1 < y_{x=180} < 46,594 + 3,182 \cdot 4,187 \cdot 1) = 95 \%$$

$$P(33,271 < y_{x=180} < 59,917) = 95 \%$$

Naloga 27a

$$n = 7$$

$$\Sigma x_i = 1.446$$

$$\Sigma x_i^2 = 306.990$$

$$\Sigma y_i = 1.598$$

$$\Sigma y_i^2 = 383.188$$

$$\Sigma x_i \cdot y_i = 338.827$$

$$\bar{x} = 206,571$$

$$\bar{y} = 228,286$$

Izračun regresijskih koeficientov:

$$b_1 = \frac{(338827) - 7 \cdot 206,571 \cdot 228,286}{(306990) - 7 \cdot 206,571^2} = 1,052$$

$$b_0 = 228,286 - 1,052 \cdot 206,571 = 10,973$$

Korelacijski koeficient izračunamo po formuli:

$$r_{xy} = \frac{(\Sigma x_i \cdot y_i) - n\bar{x}\bar{y}}{(n-1) \cdot s_x \cdot s_y}$$

Izračun s_x in s_y :

$$s_x^2 = \frac{1}{n-1} \cdot \sum (x_i - \bar{x})^2$$

$$s_x^2 = \frac{1}{6} \cdot [(186-206,571)^2 + (216-206,571)^2 + \dots + (174-206,571)^2] = 1.381,286$$

$$s_x = \sqrt{1381,286} = 37,166$$

$$s_y^2 = \frac{1}{n-1} \cdot \sum (y_i - \bar{y})^2$$

$$s_y^2 = \frac{1}{6} \cdot [(257-228,286)^2 + (260-228,286)^2 + \dots + (127-228,286)^2] = 3.064,571$$

$$s_y = \sqrt{3064,571} = 55,359$$

$$(\sum x_i \cdot y_i) - n \cdot \bar{x} \cdot \bar{y} = 338.827 - 7 \cdot 206,571 \cdot 228,286 = 8.726,129$$

$$r_{xy} = \frac{8726,129}{6 \cdot 37,166 \cdot 55,359} = 0,707$$

Determinacijski koeficient:

$$r_{xy}^2 = 0,707^2 = 49,98 \%$$

Naloga 27b

Izračunana regresijska koeficienta vstavimo v enačbo regresijske premice:

$$\hat{y} = 10,973 + 1,052 \cdot x_i$$

Naloga 28

Povprečni % zasedenosti je neodvisna spremenljivka x .

Cena v dolarjih je odvisna spremenljivka y .

$$\sum x_i = 501,2$$

$$\sum x_i^2 = 36.105,98$$

$$\sum y_i = 689,91$$

$$\sum y_i^2 = 74.444,984$$

$$\sum x_i \cdot y_i = 50.210,032$$

$$\bar{x} = 71,6$$

$$\bar{y} = 98,559$$

Izračun regresijskih koeficientov:

$$b_1 = \frac{(50210,032) - 7 \cdot 71,6 \cdot 98,559}{(36105,98) - 7 \cdot 71,6^2} = 3,691$$

$$b_0 = 98,559 - 3,691 \cdot 71,6 = -165,717$$

Izračunana regresijska koeficienta vstavimo v enačbo regresijske premice:

$$\hat{y} = -165,717 + 3,691 \cdot x_i$$

$$\hat{y}_{x=95} = -165,717 + 3,691 \cdot 95$$

$$\hat{y}_{x=95} = 516,362$$

Standardna napaka ocene odvisne spremenljivke:

$$s_{yx} = \sqrt{\frac{74444,984 - (-165,717 \cdot 689,91) - (3,691 \cdot 50210,032)}{7-2}} = 26,266$$

Izračun intervalne ocene:

$$\hat{y}_{x=95} = 516,362$$

$$s_{yx} = 26,266$$

$$\alpha = 5 \%$$

Izračun: $t_{n-2;\alpha/2} = t_{5;0,025} = 2,571$ (gledamo tabelo: Kritične vrednosti za t porazdelitev)

Upoštevamo, da je popravek h_1 enak 0.

$$516,362 \pm 2,571 \cdot 26,266 \cdot 1$$

Izračun intervalne ocene:

$$P(516,362 - 2,571 \cdot 26,266 \cdot 1 < y_{x=95} < 516,362 + 2,571 \cdot 26,266 \cdot 1) = 95 \%$$

$$P(448,832 < y_{x=95} < 583,892) = 95 \%$$

Naloga 29

$$\Sigma x_i = 621$$

$$\Sigma x_i^2 = 50.001$$

$$\Sigma y_i = 707,2$$

$$\Sigma y_i^2 = 63.670,58$$

$$\Sigma x_i \cdot y_i = 56.233,9$$

$$\bar{x} = 77,625$$

$$\bar{y} = 88,4$$

Izračun regresijskih koeficientov:

$$b_1 = \frac{(56233,9) - 8 \cdot 77,625 \cdot 88,4}{(50001) - 8 \cdot 77,625^2} = 0,745$$

$$b_0 = 88,4 - 0,745 \cdot 77,625 = 30,569$$

Izračunana regresijska koeficienta vstavimo v enačbo regresijske premice:

$$\hat{y} = 30,569 + 0,745 \cdot x_i$$

Izračun s_x in s_y :

$$s_x^2 = \frac{1}{n-1} \cdot \Sigma (x_i - \bar{x})^2$$

$$s_x^2 = \frac{1}{7} \cdot [(100-77,625)^2 + (91-77,625)^2 + \dots + (60-77,625)^2] = 256,554$$

$$s_x = 16,017$$

$$s_y^2 = \frac{1}{n-1} \cdot \Sigma (y_i - \bar{y})^2$$

$$s_y^2 = \frac{1}{7} \cdot [(105,1-88,4)^2 + \dots + (96-88,4)^2] = 164,871$$

$$s_y = 12,840$$

$$(\Sigma x_i \cdot y_i) - n \cdot \bar{x} \cdot \bar{y} = 56.233,9 - 8 \cdot 77,625 \cdot 88,4 = 1337,5$$

$$r_{xy} = \frac{1337,5}{7 \cdot 16,017 \cdot 12,840} = 0,929$$

Determinacijski koeficient:

$$r_{xy}^2 = 0,929^2 = 0,863 \text{ oz. } 86,3 \%$$

Naloga 30a

Korelacijski koeficient: $r_{xy} = 0,840$

Determinacijski koeficient: $r_{xy}^2 = 0,706$

Standardna napaka ocene odvisne spremenljivke: $s_{yx} = 11,334$

$$b_0 = -8,647$$

$$b_1 = 0,734$$

Naloga 30b

$$x = 180$$

Izračunana regresijska koeficienta vstavimo v enačbo regresijske premice:

$$\hat{y} = -8,647 + 0,734 \cdot x_i$$

$$\hat{y}_{x=180} = 59,182 + 0,126 \cdot 180$$

$$\hat{y}_{x=180} = 81,862$$

$$\hat{y}_{x=180} = 81,862$$

$$s_{yx} = 11,334$$

$$\alpha = 5 \%$$

$$n = 12 \text{ (podatki za povprečno letno porabo izdelka A)}$$

Izračun: $t_{n-2;\alpha/2} = t_{10;0,025} = 2,228$ (gledamo tabelo: Kritične vrednosti za t porazdelitev)

Intervalna ocena:

$$P(81,862 - 2,228 \cdot 11,334 \cdot 1 < y_{x=180} < 81,862 + 2,228 \cdot 11,334 \cdot 1) = 95 \%$$

$$P(56,609 < y_{x=180} < 107,114) = 95 \%$$

Osnove vzorčenja

Naloga 31

Dvostranski interval za povprečno porabo izdelka v osnovni statistični množici s 95 %-no verjetnostjo:

Imamo velik vzorec, ker je $n = 200$, ter dvostransko intervalno ocenjevanje aritmetične sredine:

$$\gamma = 95\%, \alpha = 5\%, z = \pm 1,96$$

Izračunamo standardno napako ocene aritmetične sredine:

$$se_{\bar{y}} = \frac{\sigma}{\sqrt{n}} \approx \frac{s}{\sqrt{n}}$$

$$se_{\bar{y}} = 6,2 / \sqrt{200} = 0,438$$

$$P(\bar{Y} - z \cdot se_{\bar{y}} < \bar{y} < \bar{Y} + z \cdot se_{\bar{y}}) = \gamma$$

$$P(20,8 - 1,96 \cdot 0,438 < \bar{y} < 20,8 + 1,96 \cdot 0,438) = 95 \%$$

$$P(19,941 < \bar{y} < 21,658) = 95 \%$$

S 95 %-no verjetnostjo ocenjujemo, da je povprečna poraba izdelka v osnovni statistični množici med 19,941 in 21,658 enot.

Naloga 32

$$n = 625 \text{ (velik vzorec)}$$

$$n_a = 125$$

$$\alpha = 20 \% \rightarrow z = \pm 1,28$$

Izračun strukturnega odstotka iz vzorca:

$$p = 100 \frac{n_a}{n}$$

$$p = \frac{125}{625} \cdot 100 = 20 \%$$

Izračun standardne napake ocene strukturnega odstotka:

$$SE_{\pi} = \sqrt{\frac{p(100-p)}{n}}$$

$$se_{\pi} = \sqrt{\frac{20(100-20)}{625}} = 1,6$$

$$P(p - z \cdot se_{\pi} < \pi < p + z \cdot se_{\pi}) = \gamma$$

$$P(20 - 1,28 \cdot 1,6 < \pi < 20 + 1,28 \cdot 1,6) = 80 \%$$

$$P(17,95 < \pi < 22,05) = 80 \%$$

Z 80 %-no verjetnostjo ocenjujemo, da je odstotek (%) nezaposlenih, ki prejemajo plačila na osnovi »dela na črno«, med 17,95 % in 22,05 %.

Naloga 33

Število delavcev	Število podjetij (f_k)	Sredina razreda (y_k)	$(y_k - \bar{Y})^2$
Od 1 do 5	3	3	272,25
Od 6 do 10	4	8	132,25
Od 11 do 15	2	13	42,25
Od 16 do 20	1	18	2,25
Od 21 do 25	1	23	12,25
Od 26 do 30	5	28	72,25
Od 31 do 35	2	33	182,25
Od 36 do 40	2	38	342,25
Skupaj	20		

Uporabimo formulo za aritmetično sredino iz frekvenčne porazdelitve:

$$\bar{y} = \frac{1}{N} \sum_{k=1}^r f_k y_k$$

$$\bar{Y} = \frac{1}{20} \cdot 390 = 19,5 \text{ EUR}$$

Uporabimo formulo za vzorčno varianco iz frekvenčne porazdelitve:

$$s^2 = \frac{1}{n-1} \sum_{k=1}^r f_k (y_k - \bar{y})^2$$

$$s^2 = \frac{1}{19} \cdot 2855 = 150,26 \text{ EUR}^2$$

Standardni odklon:

$$s = \sqrt{150,26} = 12,26 \text{ EUR}$$

Nato izračunamo standardno napako ocene aritmetične sredine:

$$se_{\bar{y}} = \frac{\sigma}{\sqrt{n}} \approx \frac{s}{\sqrt{n}}$$

$$se_{\bar{y}} = 12,26 / \sqrt{20} = 2,74$$

Imamo mali vzorec, ker je $n = 20$, ter dvostransko intervalno ocenjevanje aritmetične sredine:

$$P(\bar{Y} - t_{n-1; \alpha/2} \cdot se_{\bar{y}} < \bar{y} < \bar{Y} + t_{n-1; \alpha/2} \cdot se_{\bar{y}}) = \gamma$$

$$P(19,5 - 2,861 \cdot 2,74 < \bar{y} < 19,5 + 2,861 \cdot 2,74) = 99 \%$$

$$P(11,661 < \bar{y} < 27,339) = 99 \%$$

Z 99 %-no verjetnostjo ocenjujemo, da je povprečno število zaposlenih delavcev v vseh podjetjih te gospodarske panoge med 11,661 in 27,339 zaposlenih delavcev.

$$t_{n-1; \alpha/2} = t_{19; 0,005} = 2,861 \text{ (gledamo tabelo: Kritična vrednost za t porazdelitev)}$$

$$\alpha/2 = 0,01/2 = 0,005$$

$$(\gamma = 99 \% \rightarrow \alpha = 1\%)$$

Naloga 34

$$n = 50$$

$$\bar{Y} = 102$$

$$s^2 = 6,25 \rightarrow s = 2,5$$

$$se_{\bar{y}} = \frac{\sigma}{\sqrt{n}} \approx \frac{s}{\sqrt{n}}$$

$$se_{\bar{y}} = 2,5 / \sqrt{50} = 0,35$$

$$\alpha = 5\% \rightarrow z = \pm 1,96$$

Zastavimo ničelno in raziskovalno domnevo:

$$H_0: \bar{y}_D = 100 \text{ EUR}$$

$$H_1: \bar{y}_D \neq 100 \text{ EUR}$$

Izračunamo z po formuli (izračun testne vrednosti pri preizkušanju domneve o aritmetični sredini – z -test):

$$z = \frac{\bar{Y} - \bar{y}_D}{SE_{\bar{y}}}$$

$$z = \frac{102 - 100}{0,35} = 5,71 \text{ (vidimo, da dobljeno število ne pade v interval } \pm 1,96\text{)}$$

Na osnovi rezultata vidimo, da delavci ne izdelajo povprečno 100 izdelkov v določeni časovni enoti, tako da zavrnemo domnevo H_0 in sprejmemo raziskovalno domnevo H_1 .

Naloga 35

$$\alpha = 1\% \rightarrow t_{n-1; \alpha/2} = t_{4; 0,005} = 4,604 \text{ (gledamo tabelo: Kritična vrednost za } t \text{ porazdelitev)}$$

Zastavimo ničelno in raziskovalno domnevo:

$$H_0: \bar{y}_D = 20 \text{ cm}$$

$$H_1: \bar{y}_D \neq 20 \text{ cm}$$

$$se_{\bar{y}} = \frac{\sigma}{\sqrt{n}} \approx \frac{s}{\sqrt{n}}$$

$$se_{\bar{y}} = 0,2 / \sqrt{5} = 0,089$$

Izračunamo t po formuli (izračun testne vrednosti pri preizkušanju domneve o aritmetični sredini – t -test):

$$t = \frac{\bar{Y} - \bar{y}_D}{SE_{\bar{y}}}$$

$$t = \frac{20,3 - 20}{0,089} = 3,37 \text{ (vidimo, da dobljeno število pade v interval } \pm 4,604\text{)}$$

Sprejmemo domnevo H_0 .

Naloga 36

$$n = 75 \quad (\text{velik vzorec})$$

$$n_a = 55$$

$$\alpha = 10 \% \rightarrow z = \pm 1,645$$

$\frac{3}{4}$ vseh kupcev zadovoljnih s kakovostjo izdelka: 75 %

Zastavimo ničelno in raziskovalno domnevo:

$$H_0: \pi_D = 75 \%$$

$$H_1: \pi_D \neq 75 \%$$

Izračun strukturnega odstotka iz vzorca:

$$p = 100 \frac{n_a}{n}$$

$$p = 73,33 \%$$

Izračun standardne napake ocene strukturnega odstotka:

$$SE_{\pi} = \sqrt{\frac{p(100-p)}{n}}$$

$$se_{\pi} = 5,106$$

Izračun testne vrednosti pri preizkušanju domneve o strukturnem deležu – z-test:

$$z_{izr} = \frac{p - \pi_D}{se_{\pi}}$$

$$z = -0,327$$

Sprejmemo domnevo H₀.

Naloga 37

$$\alpha = 1 \% \rightarrow t_{n-1; \alpha/2} = t_{9; 0,005} = 3,250$$

$$se_{\bar{y}} = 7,906$$

$$t = -0,632$$

Sprejmemo domnevo H₀.

Naloga 38a

Velik vzorec; dvostransko intervalno ocenjevanje aritmetične sredine

$$\bar{Y} = 60,515$$

$$s = 58,704$$

$$se_{\bar{y}} = 3,184$$

$$\gamma = 95 \% \rightarrow \alpha = 5 \% \rightarrow z = \pm 1,96$$

$$P(60,515 - 1,96 \cdot 3,184 < \bar{y} < 60,515 + 1,96 \cdot 3,184) = 95 \%$$

$$P(54,274 < \bar{y} < 66,756) = 95 \%$$

Naloga 38b

$$N = 3.648$$

$$\bar{Y} = 60,515$$

$$SE_{\bar{y}} = 3,184$$

$$\text{Ocena totala iz vzorca: } \hat{Y} = N \cdot \bar{Y} = 3.648 \cdot 60,515 = 220.758,72$$

$$\text{Standardna napaka ocene totala: } SE_Y = N \cdot SE_{\bar{y}} = 3.648 \cdot 3,184 = 11615,232$$

Kritična vrednost spremenljivke $z = \pm 1,645$

$$P(220.758,72 - 1,645 \cdot 11.615,232 < Y < 220.758,72 + 1,645 \cdot 11.615,232) = 90 \%$$

$$P(201.651,663 < Y < 239.865,777) = 90 \%$$

Naloga 38c

Velik vzorec; dvostransko preizkušanje domneve o aritmetični sredini

$$H_0: \bar{y}_D = 7.500$$

$$H_1: \bar{y}_D \neq 7.500$$

$$\bar{Y} = 60,515$$

$$SE_{\bar{y}} = 3,184$$

Kritična vrednost spremenljivke $z = \pm 1,645$ ($\alpha = 0,10$)

$$z = -4,549$$

Sprejmemo domnevo H_1 .

Naloga 38d

Velik vzorec; dvostransko preizkušanje domneve o strukturnem deležu:

Zastavimo ničelno in raziskovalno domnevo:

$$H_0: \pi_D = 35 \%$$

$$H_1: \pi_D \neq 35 \%$$

Kritična vrednost spremenljivke $z = \pm 2,58$ ($\alpha = 1\%$)

$$n = 340$$

$$n_a = 92 + 41 = 133$$

$$p = 39,12$$

$$SE_{\pi} = 2,65$$

$$z = 1,55$$

Sprejmemo domnevo H_0 .

Izračun za največ 35 % (enostransko preizkušanje domneve):

$$H_0: \pi_D \leq 35 \%$$

$$H_1: \pi_D > 35 \%$$

$$n = 340$$

$$n_a = 92 + 41 = 133$$

$$p = 39,12$$

$$SE\pi = 2,65$$

Kritična vrednost spremenljivke $z = + 2,33$

$$z = 1,55$$

Sprejmemo domnevo H_0 .

Naloga 38e

Velik vzorec; dvostransko preizkušanje domneve o totalu

$$H_0: \bar{y}_D = 20.000.000 \text{ EUR}$$

$$H_1: \bar{y}_D \neq 20.000.000 \text{ EUR}$$

$$N = 3648$$

$$\bar{Y} = 60,515$$

$$s = 58,704$$

$$SE_{\bar{y}} = 3,184$$

Ocena totala iz vzorca: $\bar{Y} = 220.758,72$

Standardna napaka ocene totala: $SE_Y = 11.615,232$

Kritična vrednost spremenljivke z : $z = 1,28$ ($\gamma = 80\% \rightarrow \alpha = 20\%$)

Sprejmemo domnevo H_1 .

Naloga 39a

Mali vzorec; dvostransko ocenjevanje aritmetične sredine

$$\gamma = 80 \% \rightarrow \alpha = 20 \%$$

$$\bar{y} = 700$$

$$s = 259,808$$

$$se_{\bar{y}} = 86,60$$

Kritična vrednost spremenljivke $t_{8;0,10} = 1,379$

$$P(700 - 1,397 \cdot 86,60 < \bar{y} < 700 + 1,397 \cdot 86,60) = 80 \%$$

$$P(579,02 < \bar{y} < 820,98) = 80 \%$$

Naloga 39b

$$\bar{Y} = 700$$

$$s = 259,808$$

$$se_{\bar{y}} = 86,60$$

$$N = 2.860$$

Ocena totala iz vzorca: $\hat{Y} = 2.860 \cdot 700 = 2.002.000$

$$SE_Y = 2.860 \cdot 86,60 = 247.676$$

Mali vzorec; dvostransko preizkušanje domneve o totalu

$$H_0: \bar{y}_D = 1.800.000 \text{ EUR}$$

$$H_1: \bar{y}_D \neq 1.800.000 \text{ EUR}$$

Kritična vrednost spremenljivke t: $t_{8;0,05} = \pm 1,86$ ($\alpha = 0,10$)

$$t = 0,816$$

Sprejmemo domnevo H_0 .

Naloga 40

$$H_0: \bar{y}_D = 5.000 \text{ EUR}$$

$$H_1: \bar{y}_D \neq 5.000 \text{ EUR}$$

$$\bar{y} = 5.200$$

$$se_{\bar{y}} = 149,07$$

$$s = 894,43$$

Kritična vrednost spremenljivke $z = \pm 1,645$

$$z = 1,34$$

Sprejmemo domnevo H_0 .

Naloga 41

$$H_0: \pi_D = 50 \%$$

$$H_1: \pi_D \neq 50 \%$$

Kritična vrednost spremenljivke $z = \pm 1,96$ ($\alpha = 5 \%$)

$$n = 250$$

$$n_a = 145$$

$$p = 58$$

$$se_{\pi} = 3,12$$

$$z = 2,56$$

Sprejmemo domnevo H_1 .

Naloga 42a

$$\bar{Y} = 12,5 \text{ minut}$$

$$s = 5 \text{ minut}$$

$$se_{\bar{y}} = 0,408$$

$$\gamma = 85 \% \rightarrow \alpha = 15 \% \rightarrow z = 1,44$$

$$P(12,5 - 1,44 \cdot 0,408 < \bar{y} < 12,5 + 1,44 \cdot 0,408) = 85 \%$$

$$P(11,91 < \bar{y} < 13,09) = 85 \%$$

Naloga 42b

$$\bar{Y} = 12,5 \text{ minut}$$

$$se_{\bar{y}} = 0,408$$

$$H_0: \bar{y}_D = 10 \text{ minut}$$

$$H_1: \bar{y}_D \neq 10 \text{ minut}$$

Kritična vrednost spremenljivke $z = \pm 2,58$ ($\alpha = 1 \%$)

$$z = 6,13$$

Sprejmemo domnevo H_1 .

Naloga 43

$$n = 6 \text{ (mali vzorec)}$$

$$\alpha = 5 \%$$

$$\bar{Y} = 32,83$$

$$s = 23,19$$

$$se_{\bar{y}} = 9,467$$

$$t_{5;0,025} = 2,571 \text{ } (\alpha = 0,05)$$

$$P(32,83 - 2,571 \cdot 9,467 < \bar{y} < 32,83 + 2,571 \cdot 9,467) = 95 \%$$

$$P(8,490 < \bar{y} < 57,169) = 95 \%$$

Naloga 44

$$\bar{Y} = 68.750$$

$$s = 460$$

$$\gamma = 95 \% \rightarrow \alpha = 5 \% \rightarrow z = \pm 1,96$$

$$se_{\bar{y}} = 72,73$$

$$P(68.750 - 1,96 \cdot 72,73 < \bar{y} < 68.750 + 1,96 \cdot 72,73) = 95 \%$$

$$P(68.607,75 < \bar{y} < 68.892,55) = 95 \%$$

Naloga 45

$$n = 7 \text{ (mali vzorec)}$$

$$\alpha = 5 \%$$

$$\bar{Y} = 4.500$$

$$s = 1.200$$

$$se_{\bar{y}} = 453,56$$

$$t_{6;0,025} = 2,447 \text{ } (\alpha = 0,05)$$

$$P(4.500 - 2,447 \cdot 453,56 < \bar{y} < 4.500 + 2,447 \cdot 453,56) = 95 \%$$

$$P(3.390,14 < \bar{y} < 5.609,86) = 95 \%$$

Naloga 46

$$\bar{Y} = 20,5$$

$$s = 1,9$$

$$se_{\bar{y}} = 0,95$$

$$H_0: \bar{y}_D = 22 \text{ miligramov}$$

$$H_1: \bar{y}_D \neq 22 \text{ miligramov}$$

$$\text{Kritična vrednost spremenljivke } t: t_{3;0,025} = \pm 3,182 \text{ } (\alpha = 0,05)$$

$$t = -1,58$$

Sprejmemo domnevo H_0 .

Naloga 47a

$$\bar{Y} = 223,2$$

$$s = 28,665$$

$$se_{\bar{y}} = 12,819$$

$$z = \pm 1,96$$

$$P(223,2 - 1,96 \cdot 12,819 < \bar{y} < 223,2 + 1,96 \cdot 12,819) = 95 \%$$

$$P(198,075 < \bar{y} < 248,325) = 95 \%$$

Naloga 47b

$$H_0: \bar{y}_D \geq 220$$

$$H_1: \bar{y}_D < 220$$

Kritična vrednost spremenljivke $z = -1,28$ ($\alpha = 10 \%$)

$$z = 0,249$$

Sprejmemo domnevo H_0 .

Naloga 48a

$$n = 7 \text{ (mali vzorec)}$$

$$\gamma = 80 \% \rightarrow \alpha = 20 \% \rightarrow t_{6;0,10} = 1,440$$

$$\bar{Y} = 27,286$$

$$s = 37,622$$

$$se_{\bar{y}} = 14,219$$

$$P(27,286 - 1,440 \cdot 14,219 < \bar{y} < 27,286 + 1,440 \cdot 14,219) = 80 \%$$

$$P(6,811 < \bar{y} < 47,761) = 80 \%$$

Naloga 48b

Enostransko preizkušanje domneve:

$$\alpha = 5 \% \rightarrow t_{n-1; \alpha} = t_{6;0,05} = +1,943 \text{ (Kritična vrednost spremenljivke } t.)$$

$$H_0: \bar{y}_D \leq 30$$

$$H_1: \bar{y}_D > 30$$

$$se_{\bar{y}} = 14,219$$

$$t = -0,191$$

Sprejmemo domnevo H_0 .

Časovne vrste

Naloga 50a

Leto	Živorojeni	$I_{t/3}$ v %	V_t
1	1,94	103,7	/
2	1,89	101,1	97,4
3	1,87	100	98,9
4	1,82	97,3	97,3
5	1,78	95,2	97,8
6	1,75	93,6	98,3
7	1,81	96,8	103,4

Naloga 50b

Leto	Srednje štev. preb. (30.6.)	Živorojeni	Koeficient natalitete	Umrli	Koeficient mortalitete
1	198	1,94	9,8	1,93	9,7
2	198	1,89	9,5	1,89	9,5
3	199	1,87	9,4	1,86	9,4
4	198	1,82	9,2	1,89	9,5
5	198	1,78	9,0	1,90	9,6
6	198	1,75	8,8	1,88	8,8
7	199	1,81	9,1	1,85	9,3

Koeficient natalitete meri število živorojenih na 1000 prebivalcev, koeficient mortalitete pa število mrtvih na 1000 prebivalcev.

Naloga 51a, 51b, 51c

Leto	1	2	3	4	5	6	7	8	9
$S_t\%$	/	+18,0	-2,1	+2,7	+2,5	+2,5	-5,0	+7,0	-5,8
V_t	/	118	97,9	102,7	102,5	102,5	95	107	94,2
K_t	/	1,18	0,979	1,027	1,025	1,025	0,95	1,07	0,942
$I_{t/4}$	84,3	99,5	97,4	100	102,5	105,1	99,8	106,8	100,6
Y_t (v kg)	71,2	84	82,3	84,5	86,6	88,8	84,3	90,2	85,0

Naloga 52

Mesec	I	II	III	IV	V
Vrednost proizvodnje v 10^6 EUR	652	730	840	752	-
Štev. zaposlenih na začetku meseca	214	240	226	208	200
Štev. zaposlenih na sredini meseca	227	233	217	204	-

Za izračun tega statističnega koeficienta je potrebno najprej uskladiti različni podatek (vrednost proizvodnje) ter trenutni podatek (število zaposlenih), tako da opredelimo število zaposlenih na sredini meseca (izračunamo aritmetično sredino števila zaposlenih na začetku meseca in na koncu meseca (oziroma na začetku sledečega meseca), kar je prikazano v 3. vrstici tabele.

Povprečni statistični koeficient – povprečna vrednost proizvodnje na 10 zaposlenih – za opazovano obdobje je enak (v 10^6 EUR):

$$K = \frac{652 + 730 + 840 + 752}{227 + 233 + 217 + 204} 10 = 33,76$$

Naloga 53a

Mesec	I.	II.	III.	IV.	V.
Povprečna mesečna plača zaposlenih (v d.e.)	2.600	2.800	3.100	2.900	3.200
Povprečna mesečna plača zaposlenih – $I_{t/V}$ (v %)	81,25	87,5	96,9	90,6	100

Naloga 53b

Mesec	I.	II.	III.	IV.	V.
Mesečna stopnja rasti vrednosti prodaje S_t %	–	+15	–10	+8	–12
Mesečna stopnja rasti vrednosti prodaje K_t	–	1,15	0,90	1,08	0,88
Mesečna stopnja rasti vrednosti prodaje – $I_{t/I}$ (v %)	100	115	103,5	111,8	98,4

Naloga 53c

Povprečna mesečna plača zaposlenih

$$K = \sqrt[4]{\frac{3200}{2600}} = 1,053$$

$$S = +5,3 \%$$

Vrednost prodaje:

$$K = \sqrt[4]{\frac{98,4}{100}} = 0,9960$$

$$S = -0,4 \%$$

V opazovanem obdobju so povprečne plače naraščale povprečno za 5,3 % na mesec, vrednost proizvodnje pa se je v povprečju zmanjševala za 0,4 % mesečno.

Naloga 53d

Napoved povprečne mesečne plače za oktober (v d.e.):

$$Y_t = 3200 * 1,053^5 = 4.142,8$$

Naloga 54a

Regija	A	B	C	D	E	F	G
Št. promocijskih akcij	21	25	14	17	25	12	18
Porabljen denarna sredstva (v d.e.)	2.314	3.100	2.100	3.000	2.600	1.750	2.950
Porabljen denarna sredstva (v d.e.) na promocijsko akcijo	110,19	124	150	176,47	104	145,83	163,89

Povprečno porabljen denarna sredstva na promocijsko akcijo za vse regije skupaj (v d.e.):

$$K = \frac{17814}{132} = 134,95$$

Naloga 54b

Regija	A	B	C	D	E	F	G
Porabljen denarna sredstva (v d.e.)	2.314	3.100	2.100	3.000	2.600	1.750	2.950
Porabljen denarna sredstva (v d.e.) v primerjavi z regijo G (v %)	78,44	105,08	71,19	101,69	88,14	59,32	100

V primerjavi z regijo G je najmanjša vrednost porabljenih denarnih sredstev dosežena v regiji F.

Naloga 55a

Verižni indeks za 4. časovno enoto, v %:

$$V_4 = \frac{400}{118} 100 = 338,98$$

Naloga 55b

Indeks z osnovo 1, za 7. časovno enoto, v %

$$I_{7/1} = \frac{300}{115} 100 = 260,87$$

Naloga 56a, 56b

Leto	1	2	3	4
S_t (v %)	/	+18,4	+11,7	-9,2
K_t	/	1,184	1,117	0,908
$I_{t/2}$ (v %)	84,46	100	111,7	101,42
Y_t	128,92	152,64	170,5	154,81

Naloga 56c

$$K = \sqrt[3]{\frac{101,42}{84,46}} = 1,063$$

Napoved za 7. časovno enoto, v d.e.:

$$Y_7 = 154,81 * 1,063^3 = 185,90$$

Naloga 57a

Sistem normalnih enačb za izračun funkcije trenda:

$$5a + 15b = 24.700$$

$$15a + 55b = 75.200$$

Funkcija trenda: $\hat{Y} = 4610 + 110t$

Napoved za število nočitev za 6. leto: $\hat{Y}_t = 4610 + 110 * 6 = 5.270$

Naloga 57b

Leto	I-IV	V-VIII	IX-XII	SKUPAJ
1	1.200	2.500	1.000	4.700
2	1.100	2.600	1.200	4.900
3	1.000	2.800	1.100	4.900
4	1.300	2.500	1.200	5.000
5	1.200	2.700	1.300	5.200
Skupaj	5.800	1.3100	5.800	24.700

Povprečno 4-mesečno število nočitev v obdobju petih let: $24700/3 = 8.233,3$ nočitev.

$$SI1 = \frac{5800}{8233,3} 100 = 70,4 \%$$

$$SI2 = \frac{13100}{8233,3} 100 = 159,2 \%$$

$$SI3 = \frac{5800}{8233,3} 100 = 70,4 \%$$

Povprečno 4-mesečno število nočitev v 6. letu: $5270/3 = 1.756,6$.

Napoved po 4-mesečjih za 6. leto:

$$\hat{Y}_{I-IV;6.letu} = 0,704 * 1.756,6 = 1.236,7$$

$$\hat{Y}_{V-VIII;6.letu} = 1,592 * 1.756,6 = 2.796,6$$

$$\hat{Y}_{IX-XII;6.letu} = 0,704 * 1.756,6 = 1.236,7$$

Naloga 58a

Ocena stroškov za prihodnje leto, v d.e.

$$\hat{Y}_5 = 58,5 + 2,1 * 5 = 69$$

Naloga 58c

Povprečni četrletni stroški za 5. leto: $69/4 = 17,25$ d.e.

Ocena stroškov za 4. četrletje, 5. leto, v d.e.:

$$\hat{Y}_{4,\text{četrletje},5.\text{leto}} = 2,255 * 17,25 = 38,90$$

Naloga 59a

$$\hat{Y}_t = 109,43 + 4 * t$$

Naloga 59b

Napoved za 8. leto, za število izdanih gradbenih dovoljenj:

$$\hat{Y}_t = 109,43 + 4 * 8 = 141,43$$

Naloga 60b

Napoved za število novo zaposlenih za 6. leto: $\hat{Y}_6 = 35 * 6^{0,2} = 50,08$

Napoved po četrletjih za število novo zaposlenih za 6. leto

Povprečna četrletna vrednost za število novo zaposlenih za 6. leto: $50,08/4 = 12,52$

$$\hat{Y}_{1,\text{četrletje},6.\text{leto}} = 0,65 * 12,52 = 8,1$$

$$\hat{Y}_{2,\text{četrletje},6.\text{leto}} = 1,55 * 12,52 = 19,4$$

$$\hat{Y}_{3,\text{četrletje},6.\text{leto}} = 1,60 * 12,52 = 20,0$$

$$\hat{Y}_{4,\text{četrletje},6.\text{leto}} = 0,20 * 12,52 = 2,5$$

Tabele

Ploščine $H(z)$ za standardizirano normalno porazdelitev

$$z = \frac{y - \bar{y}}{\sigma}$$

z	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0,0	0,0000	0,0040	0,0080	0,0120	0,0160	0,0199	0,0239	0,0279	0,0319	0,0359
0,1	0,0398	0,0438	0,0478	0,0517	0,0557	0,0596	0,0636	0,0675	0,0714	0,0753
0,2	0,0793	0,0832	0,0871	0,0910	0,0948	0,0987	0,1026	0,1064	0,1103	0,1141
0,3	0,1179	0,1217	0,1255	0,1293	0,1331	0,1368	0,1406	0,1443	0,1480	0,1517
0,4	0,1554	0,1591	0,1628	0,1664	0,1700	0,1736	0,1772	0,1808	0,1844	0,1879
0,5	0,1915	0,1950	0,1985	0,2019	0,2054	0,2088	0,2123	0,2157	0,2190	0,2224
0,6	0,2257	0,2291	0,2324	0,2357	0,2389	0,2422	0,2454	0,2486	0,2518	0,2549
0,7	0,2580	0,2612	0,2642	0,2673	0,2704	0,2734	0,2764	0,2794	0,2823	0,2852
0,8	0,2881	0,2910	0,2939	0,2967	0,2995	0,3023	0,3051	0,3078	0,3106	0,3133
0,9	0,3159	0,3186	0,3212	0,3238	0,3264	0,3289	0,3315	0,3340	0,3365	0,3389
1,0	0,3413	0,3438	0,3461	0,3485	0,3508	0,3531	0,3554	0,3577	0,3599	0,3621
1,1	0,3643	0,3665	0,3686	0,3708	0,3729	0,3749	0,3770	0,3790	0,3810	0,3830
1,2	0,3849	0,3869	0,3888	0,3907	0,3925	0,3944	0,3962	0,3980	0,3997	0,4014
1,3	0,4032	0,4049	0,4066	0,4082	0,4099	0,4115	0,4131	0,4147	0,4162	0,4177
1,4	0,4192	0,4207	0,4222	0,4236	0,4251	0,4265	0,4279	0,4292	0,4306	0,4319
1,5	0,4332	0,4345	0,4357	0,4370	0,4382	0,4394	0,4406	0,4418	0,4429	0,4441
1,6	0,4452	0,4463	0,4474	0,4484	0,4495	0,4505	0,4515	0,4525	0,4535	0,4545
1,7	0,4554	0,4564	0,4573	0,4582	0,4591	0,4599	0,4608	0,4616	0,4625	0,4633
1,8	0,4641	0,4649	0,4656	0,4664	0,4671	0,4678	0,4686	0,4693	0,4699	0,4706
1,9	0,4713	0,4719	0,4726	0,4732	0,4738	0,4744	0,4750	0,4756	0,4761	0,4767
2,0	0,4772	0,4778	0,4783	0,4788	0,4793	0,4798	0,4803	0,4808	0,4812	0,4817
2,1	0,4821	0,4826	0,4830	0,4834	0,4838	0,4842	0,4846	0,4850	0,4854	0,4857
2,2	0,4861	0,4864	0,4868	0,4871	0,4875	0,4878	0,4881	0,4884	0,4887	0,4890
2,3	0,4893	0,4896	0,4898	0,4901	0,4904	0,4906	0,4909	0,4911	0,4913	0,4916
2,4	0,4918	0,4920	0,4922	0,4925	0,4927	0,4929	0,4931	0,4932	0,4934	0,4936
2,5	0,4938	0,4940	0,4941	0,4943	0,4945	0,4946	0,4948	0,4949	0,4951	0,4952
2,6	0,4953	0,4955	0,4956	0,4957	0,4959	0,4960	0,4961	0,4962	0,4963	0,4964
2,7	0,4965	0,4966	0,4967	0,4968	0,4969	0,4970	0,4971	0,4972	0,4973	0,4974
2,8	0,4974	0,4975	0,4976	0,4977	0,4977	0,4978	0,4979	0,4979	0,4980	0,4981
2,9	0,4981	0,4982	0,4983	0,4983	0,4984	0,4984	0,4985	0,4985	0,4986	0,4986
3,0	0,4987									
3,5	0,4997									
4,0	0,4999									

Primer: Za $z = 1,96$ iz preglednice odčitamo površino 0,4750. (Vir: Artenjak, 2000)

Kritične vrednosti za t porazdelitev

(Vir: Artenjak, 2000)

$$t = \frac{\bar{Y} - \bar{Y}_D}{SE_{\bar{Y}}}$$

Prostostne stopinje	0,10	0,05	0,025	0,01	0,005
1	3,078	6,314	12,706	31,821	63,657
2	1,886	2,920	4,303	6,965	9,925
3	1,638	2,353	3,182	4,541	5,841
4	1,533	2,132	2,776	3,747	4,604
5	1,476	2,015	2,571	3,365	4,032
6	1,440	1,943	2,447	3,143	3,707
7	1,415	1,895	2,365	2,998	3,499
8	1,397	1,860	2,306	2,896	2,355
9	1,383	1,833	2,262	2,821	3,250
10	1,372	1,812	2,228	2,764	3,169
11	1,363	1,796	2,201	2,718	3,106
12	1,356	1,782	2,179	2,681	3,055
13	1,350	1,771	2,160	2,650	3,012
14	1,345	1,761	2,145	2,624	2,977
15	1,341	1,753	2,131	2,602	2,947
16	1,337	1,746	2,120	2,583	2,921
17	1,333	1,740	2,110	2,567	2,898
18	1,330	1,734	2,101	2,552	2,878
19	1,328	1,729	2,093	2,539	2,861
20	1,325	1,725	2,086	2,528	2,845
21	1,323	1,721	2,080	2,518	2,831
22	1,321	1,717	2,074	2,508	2,819
23	1,319	1,714	2,069	2,500	2,807
24	1,318	1,711	2,064	2,492	2,797
25	1,316	1,708	2,060	2,485	2,787
26	1,315	1,706	2,056	2,479	2,779
27	1,314	1,703	2,052	2,473	2,771
28	1,313	1,701	2,048	2,467	2,763
29	1,311	1,699	2,045	2,462	2,756
30	1,310	1,697	2,042	2,457	2,750
40	1,303	1,684	2,021	2,423	2,704
60	1,296	1,671	2,000	2,390	2,660
120	1,289	1,658	1,980	2,358	2,617
	1,282	1,645	1,960	2,326	2,576

Literatura in viri

- Artenjak, J. (2000). *Poslovna statistika*. Maribor: Ekonomsko – poslovna fakulteta.
- Agresti, A., Finlay, B. (2009). *Statistical methods for the Social sciences*. Boston: Pearson.
- Brockwell, P. J., Davis, R. A. (2016). *Introduction to Time Series and Forecasting*, Springer Nature.
- Ghauri, P., Grønhaug, K., Strange, R. (2020). *Research Methods in Business Studies*. United Kingdom: University Printing House.
- Heumann, C., Schomaker, M., Shalabh, S. (2016). *Introduction to Statistics and Data Analysis*. Singapore: Springer.
- Holmes, A., Illowsky, B., Dean, S. (2018). *Introductory Business Statistics*. Houston: Rice University.
- Leskovar Špacapan, G., Tominc, P. (2008). *Učno gradivo pri predmetu Statistika, naloge za delo na vajah in samostojno delo*, interno gradivo. Maribor: EPF.
- Mišić, E. (2019). *Indeksna števila*. RS: Statistični urad.
- Sakaran, U., Bougie, R. (2016). *Research Methods for Business, A Skill-Building Approach*, 7th Ed. United Kingdom: John Wiley & Sons.
- Tabachnick, B. G., Fidell, L. S. (2013). *Using multivariate statistics*. Boston: Pearson.
- Tominc, P. (2016). Statistika (2. del predmeta). *Učno gradivo pri predmetu Statistika (2. del predmeta)*, interno gradivo. Maribor: EPF.
- Tominc, P., Kramberger, T. (2007). *Statistične metode v logistiki*. Celje: UM Fakulteta za logistiko.

STATISTIKA:

GRADIVA ZA SEMINARSKE VAJE ZA

2. DEL PREDMETA

POLONA TOMINC IN MAJA ROŽMAN

Univerza v Mariboru, Ekonomsko-poslovna fakulteta, Maribor, Slovenija.

E-pošta: polona.tominc@um.si. maja.rozman@um.si

Povzetek Predmet Statistika (2. del predmeta), za katerega je namenjena ta zbirka vaj, je pomemben za to, da študenti razvijejo sposobnost razumevanja informacij v podatkih. Gradivo obravnava naslednje vsebinske sklope: prikazovanje podatkov v tabelah in grafih, relativna števila, mere centralne tendence, mere variabilnosti, asimetrije in sploščenosti, intervalno ocenjevanje vrednosti statističnih parametrov in osnove preizkušanja domnev o statističnih parametrih, osnove enostavne regresije ter osnove analize in napovedovanja vrednosti v časovnih vrstah.

Ključne besede:

urejanje in prikazovanje podatkov, deskriptivna statistika, enostavna regresijska analiza, osnove vzorčenja, časovne vrste

STATISTICS: SEMINAR TUTORIALS FOR THE 2ND PART OF THE COURSE

POLONA TOMINC & MAJA ROŽMAN

University of Maribor, Faculty of Economics and Business, Maribor, Slovenia.

E-mail: polona.tominc@um.si. maja.rozman@um.si

Keywords:
data
visualisation,
descriptive
statistics,
simple
regression
analysis,
basics of
sampling,
time
series

Abstract The course Statistics (the second part of the course), for which this collection of exercises/tutorials is intended, is important for students to develop the ability to understand the information in data. These tutorials contain the following chapters: data visualisation using tables and graphs, relative numbers, measures of central tendency, measures of dispersion, skewness and kurtosis, confidence intervals for statistical parameters and the basis of hypothesis testing of statistical parameters, basics of simple regression and basics of time series analysis and forecasting.







Univerza v Mariboru

Ekonomsko-poslovna fakulteta