Mirjana Radman-Funarić

UVOD U GOSPODARSKU STATISTIKU

Tko kaže da lažem?

u potpisu – Statistika

Požega, 2018. godina

Autor

Mirjana Radman-Funarić

Naslov:

UVOD U GOSPODARSKU STATISTIKU

Nakladnik:

Veleučilište u Požegi

Recenzenti:

prof. dr.sc. Zlatko Lacković

dr.sc. Katarina Potnik Galić, prof.v.š.

Lektor

Slavica Budić, prof.

Napomena:

Zahtjev sustava PDF viewer

ISBN 978-953-7744-39-7

Odluka Stručnog vijeća Veleučilišta u Požegi na 17. sjednici Stručnog vijeća, održanoj 14. lipnja 2018. godine

Sadržaj

1. TEMELJNI POJMOVI

1.1. Uvodna riječ o statistici

Predmet proučavanja statistike su masovne pojave. **Statistika** je znanstvena disciplina koja se bavi prikupljanjem, analizom i tumačenjem podataka masovnih pojava.[1] Svrha primjene statističkih metoda je donošenje suda o osobitosti promatranih pojava, ispitivanje različitih pretpostavki, procjena karakterističnih veličina, odabir statističkih modela koji generiraju pojave, predviđaju razine i stanja pojava. Statističke se metode primjenjuju u svim područjima stručne i znanstvene djelatnosti.[2]

U okviru metoda **deskriptivne statistike** zaključci se donose na temelju podataka cijele populacije. Ona obuhvaća postupke uređivanja, grupiranja, tabeliranja, grafičkog prikazivanja te izračunavanja različitih statističko-analitičkih veličina. Temelji se na potpunom obuhvatu statističkog skupa.

U okviru metoda **inferencijalne statistike** zaključci o populaciji donose se na temelju dijela jedinica statističkog skupa, odnosno na temelju podataka izabranih u uzorak. Pri tome iz populacije može biti izabran jedan ili više uzoraka. Inferencijalna statistika temelji se na teoriji vjerojatnosti te proučava metode kojima se pomoću dijela informacija (uzorka) donose zaključci o cjelini (populaciji).

Ovaj nastavni materijal obrađuje metode deskriptivne statistike.

[1] Horvat, J. i Mijoč, J. (2014), *Osnove statistike*, Drugo dopunjeno izdanje, Ljevak, Zagreb, str. 19.
[2] Šošić, I. (2006), Primijenjena statistika, 2. izmijenjeno izdanje, Školska knjiga, Zagreb, str. 3.

1.2. Statistički skup

Temeljni pojam u statistici – **statistički skup** – sastoji se od jedinica čija su svojstva predmet proučavanja statističkim metodama. Statistički kup mogu činiti npr. osobe, proizvodi, poslovni subjekti, vrijednosnice, novčane transakcije, prometne nezgode i dr.

Statistički skup potrebno je **definirati** pojmovno, prostorno i vremenski, npr. zaposleni djelatnici u Republici Hrvatskoj, stanje 31. prosinca 2017. godine.

Slika 1.1. Definiranje statističkog skupa

Opseg skupa uključuje broj njegovih jedinica. S obzirom na opseg statistički skupovi mogu biti:

- konačni statistički skup, npr. registrirana trgovačka društva na području Vukovarsko-srijemske županije u 2017. godini.
- beskonačni statistički skup, npr. broj ishoda bacanja novčića

Statistički skup može biti **realan**, ako njegove jedinice postoje u tekućem vremenu, npr. nezaposlene osobe u Istarskoj županiji u dobi od 18 do 30 godina, i **hipotetičan** (zamišljen) ako su njegovi elementi rezultati eksperimenta ili danog modela procesa.

Primjer 1.1. Realan i hipotetičan statistički skup
Ako istraživač ima namjeru ispitati učinkovitost nove metode proizvodnje kruha,statistički skup može definirati kao rezultat postignuća korištenja nove metode svih pekara na području Osječko-baranjske županije. Statistički skup je hipotetičan u slučaju da ne postoji skupina pekara koji bi koristili novu metodu proizvodnje kruha, a sastoji se od rezultata (učinkovitost) koji bi se postigli ako bi koristili novu metodu.

Podaci o varijabli za svaki element statističkog skupa tvore skup podataka koji se naziva **statističkom populacijom,** odnosno **osnovnim skupom**. Kao i statistički skup, populacija može biti realna ili hipotetična, a s obzirom na broj elemenata konačna i beskonačna.[3]
Podskup ili dio statističkog skupa, odnosno podaci o dijelu statističke populacije naziva se **uzorak**.

Primjer 1.2. Statistički skup:
Broj zaposlenih u Republici Hrvatskoj, na dan 31. ožujka 2015. godine, s podacima o visini bruto i neto plaće zaposlenih.

T: Statistički skup sastoji se od zaposlenih osoba u Republici Hrvatskoj, na dan 31. ožujka 2015. godine. Jedinica tog skupa je jedna zaposlena osoba (jedinica promatranja). Skup je realan i definiran (postoje propisi prema kojima se osoba smatra zaposlenom). Skup je prostorno određen (Republika Hrvatska). Skup je vremenski određen (31. ožujka 2015. godine). U skupu je 76.605 elemenata (zaposlenih osoba). Populacije su: 76.605 podataka o visini bruto plaće i 76.605 podataka o visini neto plaće.

Statističke metode omogućuju određivanje npr. prosječne plaće zaposlenih, odstupanje pojedinih zaposlenika u visini plaće od prosjeka populacije, struktura zaposlenika s obzirom na dob i slično. Navedeni pokazatelji zovu se

[3] Šošić, I. (2006), *Primijenjena statistika*, Školska knjiga, Zagreb, str. 3.

parametri jer su određeni pomoću podataka za svaki element statističkog skupa.

Ako se za donošenje suda o parametrima populacije koristi uzorak, riječ je o **procjeni parametara** koji u sebi sadrže i pogrešku jer je rezultat upotrebe dijela podataka.

1.3. Statističko obilježje ili statistička varijabla

Statističko obilježje ili **statistička varijabla** je svojstvo po kojem se jedinice statističkog skupa razlikuju ili jedne drugima nalikuju. Variranje svojstava statističkog obilježja nazivaju se **modaliteti obilježja**. Svako obilježje pojavljuje se u dva ili više modaliteta. Podjela varijabla često se temelji na upotrebi mjerne ljestvice.

Vrste statističke varijable:

1. Kvalitativna (kategorijska) varijabla je ona kod koje se vrijednosti varijable iskazuju pojmovno (riječima i slovnim oznakama), razvrstavaju se unutar kategorija prema nekoj karakteristici ili atributu. Dijele se na:
   a) Nominalnu varijablu koja nastaje prikupljanjem podataka na temelju nominalne ljestvice. U nominalne varijable se ubrajaju:
   - atributivna varijabla, npr. vrsta vlasništva (privatno, javno), pravni oblik udruživanja (trgovačko društvo, obrt)
   - geografska varijabla, npr. država uvoznik iz Republike Hrvatske (Austrija, Bosna i Hercegovina i dr.)

   b) Redoslijednu varijablu ili obilježje ranga koja nastaje prikupljanjem podataka na temelju ordinalne mjerne ljestvice, npr. stupanj stručne spreme (VSS, SSS), stupanj kvalitete, ocjena (od 1 do 10).
2. Kvantitativna varijabla je ona kod koje se vrijednosti varijable zapisuju brojevima, pri čemu su podaci izmjereni intervalnom ili odnosnom (omjernom) ljestvicom mjerenja:
   a) Numerička varijabla dijeli se na:
   - diskontinuiranu(diskretnu) varijablu koja se može izraziti isključivo cijelim brojem i koja može poprimiti konačan broj svojstava, npr. broj

registriranih obrta na području županije, broj zaposlenih u trgovačkom društvu, broj prodanih artikala

- kontinuiranu varijablu koja može poprimiti bilo koju vrijednosti unutar intervala, a može se izraziti cijelim i decimalnim brojem, npr. godine staža zaposlenika, ostvarena dobit u kunama, utrošena električna energija u kWh

b) Vremenska varijabla izražava se u jedinici vremena (godina, mjesec, sekunda)

Podaci o obilježjima (modalitetima) dobivaju se promatranjem ili provedbom statističkog pokusa.

Skup statističkih podataka od kojih se polazi u statističkoj analizi prikladno se može prikazati matricom podataka koja ima dimenzije $N \times K$, pri čemu je N broj jedinica, a K broj varijabli. U retku matrice (tablica1.2.) nalaze se podaci o *obliku svakog obilježja (varijable) jedne jedinice*, a stupci matrice sadrže podatke o *oblicima jednog obilježja (varijable) za svaku jedinicu*. Element d_{ij} matrice je modalitet *j*-te varijable *i*-te jedinice.

Tablica 1.2. Matrica prikupljenih podataka

	K / N	Oblik jednog obilježja za svaku jedinicu (j)		
		1	2	3
Oblik svakog obilježja jedne jedinice (i)	1	d_{11}	d_{12}	d_{13}
	2	d_{21}	d_{22}	d_{23}
	3	d_{31}	d_{23}	d_{33}
	4	d_{41}	d_{24}	d_{34}

	Podaci / Trgovačko društvo	Oblik jednog obilježja za svaku jedinicu (j)		
		Broj zaposlenih	Ostvarena dobit u kn	Oblik udruživanja
Oblik svakog obilježja jedne jedinice (i)	A	9	95.876	d.o.o.
	B	254	899.503	d.d.
	C	58	543.455	d.o.o.
	D	121	874.321	j.t.d.

U označenom retku nalaze se podaci o modalitetima svake varijable za trgovačko društvo B. Društvo B ima 254 zaposlenih, ostvarilo je dobit u iznosu od 899.503,00 kn i udruženo je u dioničko društvo. U označenom stupcu nalaze se podaci o modalitetima broja zaposlenih za svako trgovačko društvo.

1.4. Mjerenje i mjerne ljestvice

Mjerenje je postupak pridruživanja numeričkih i nenumeričkih oznaka jedinicama statističkih skupova na temelju određenog pravila. Temelji se na primjeni mjernih skala.

Mjerne ljestvice (skale)

1. **Nominalna**

Nominalnom razinom mjerenja (a = b), opisuju se atributivna i geografska obilježja u kojima nisu dopuštene računske operacije i u kojoj se modaliteti obilježja ne mogu uređivati redoslijedom, budući da neka kategorija ne može biti bolja od druge, npr. rod zaposlenika,

2. **Ordinalna**

Ordinalnom razinom mjerenja (a < b, a > b, a = b), jedinicama statističkih skupova pridružuju se slovne oznake, simboli ili brojevi prema intenzitetu mjernog svojstva, pri čemu nije moguće odrediti stupanj razlike između podataka, npr. stupanj stručne spreme, zakonom definirana veličina trgovačkog društva (mikro, mala, srednja i velika).

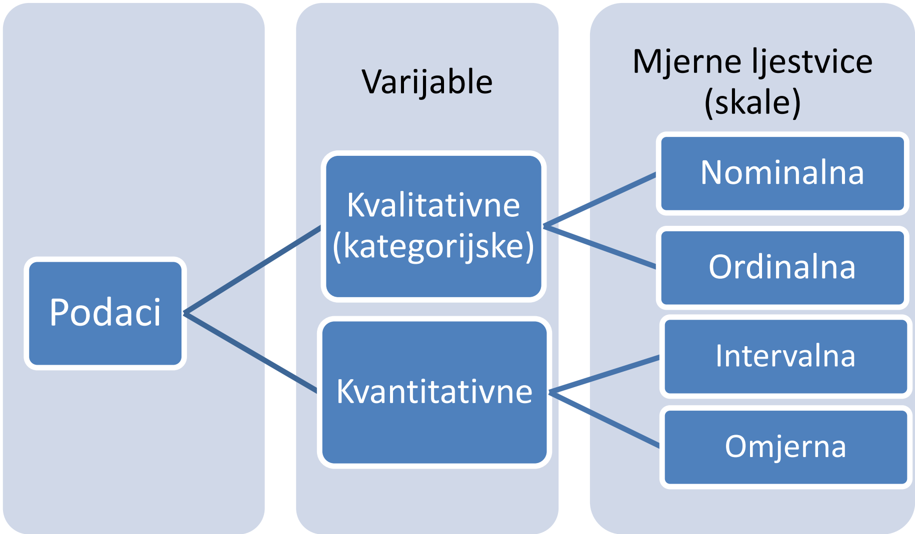
3. **Intervalna**

Intervalnom razinom mjerenja (a – b), jedinicama statističkih skupova pridružuju se brojevi sukladno intenzitetu mjernog svojstva, pri čemu jednake razlike brojeva na mjernoj ljestvici predstavljaju jednake razlike mjernog svojstva, odnosno ima definiranu mjernu jedinicu. Intervalna ljestvica nema dogovorenu apsolutnu nulu koja upućuje na odsutnost mjernog svojstva, nego ima dogovorno utvrđenu *relativnu nulu* koja označava stanje i intenzitet svojstva, npr. temperaturna ljestvica, nadmorska visina.

4. Omjerna (odnosna)

Omjernom razinom mjerenja (a/b), jedinicama statističkih skupova pridružuju se jednake razlike brojeva koje predstavljaju jednake razlike mjernog svojstva. Za ovu mjernu ljestvicu karakteristično je da ima definiranu mjernu jedinicu (*odnosna razina* pri čemu početna točka mjerenja daje značenje odnosu vrijednosti između podataka) i *apsolutnu nulu* koja označava nepostojanje svojstva, npr. plaća zaposlenika, broj lomova proizvoda pri transportu.

Slika 1.2. Statističke varijable i mjerne ljestvice

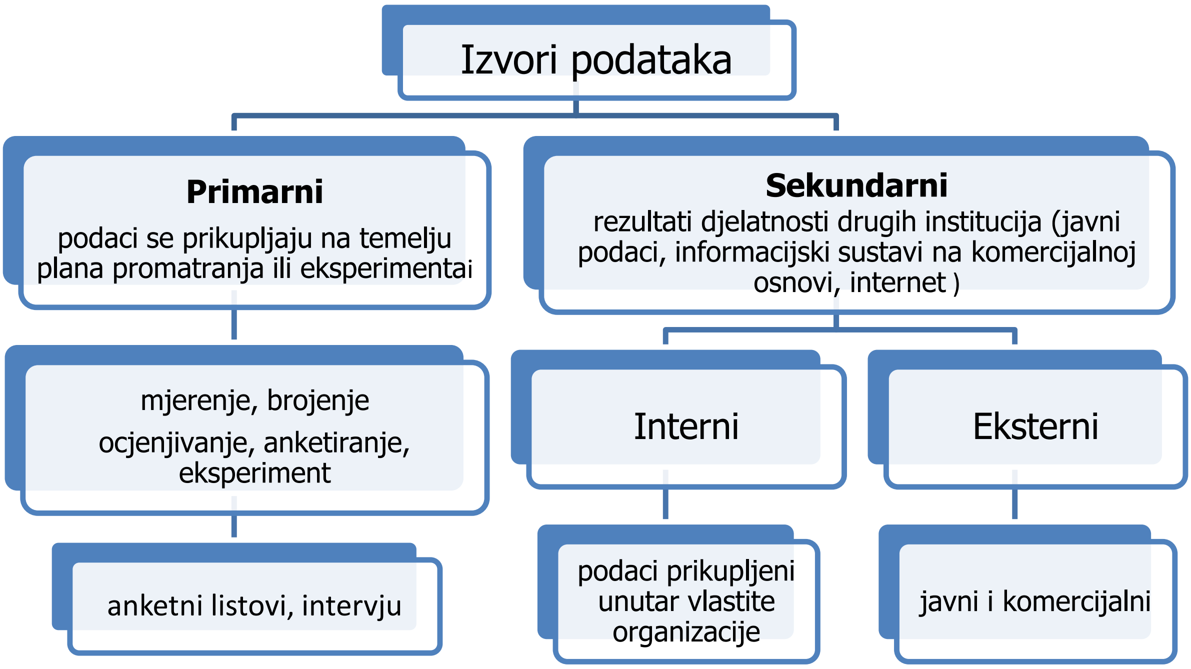


1.5. Izvori podataka

Izvori podataka mogu biti:

- **Primarni**, koji se prikupljaju u skladu s ciljem istraživanja. Prednost primarnih istraživanja je u tome što su rezultati prilagođeni zahtjevima istraživača, a nedostatak su visoki troškovi istraživanja.
- **Sekundarni** mogu biti **interni**, koji su prikupljeni ranije za neke druge potrebe i **eksterni**, koje prikupljaju razne institucije. Eksterni podaci mogu biti **javni** (npr. Državni zavod za statistiku Republike Hrvatske, banke, Hrvatska narodna banka, Hrvatski zavod za zapošljavanje, Hrvatska gospodarska komora) i **komercijalni** (npr. Financijska agencija – FINA, agencije za istraživanje tržišta). Prednost sekundarnih podataka je niska ili relativno niska cijena koja se plaća za sekundarne podatke, a nedostatak je što nisu uvijek prilagođeni vlastitim potrebama.

Slika 1.3. Izvori podataka



2. UREĐIVANJE PODATAKA

Kako bi nastali statistički nizovi, prikupljene podatke potrebno je urediti. Kako bi podaci bili pregledni, potrebno ih je grupirati u nizove. Mogu biti prikazani tablično i grafički.

2.1. Statistički nizovi

Uređivanjem podataka nastaju statistički nizovi.

Statistički niz grupiranih podataka je skup parova različitih obilježja s pripadajućim frekvencijama. Statistički niz nastaje grupiranjem statističkih podataka:

- **Nominalni niz** nastaje uređivanjem podataka o modalitetima nominalne varijable.
- **Redoslijedni niz** nastaje uređivanjem podataka o modalitetima rang varijable.
- **Numerički niz** nastaje uređenjem podataka koji predstavljaju vrijednosti numeričke varijable.
- **Vremenski niz** nastaje kronološkim nizanjem podataka o nekoj pojavi.

Broj jedinica statističkog skupa koje imaju ista obilježja naziva se ***frekvencijom.*** Frekvencijom se zapisuje učestalost pojavljivanja pojedinih modaliteta i označava se sa f. Zbroj frekvencija čini ***opseg skupa*** *(broj članova statističkog niza (N).*

$$0 < f < N \ldots\ldots \sum f = N$$

Grupiranje je postupak raščlanjivanja jedinica statističkog skupa prema modalitetima obilježja. Grupiranjem vrijednosti numeričkog obilježja nastaje distribucija frekvencija.

Distribucija frekvencija je razdioba jedinica statističkog skupa prema modalitetima jednog promatranog obilježja. Distribucija frekvencija se sastoji od parova razreda i pripadajućih frekvencija (x_i, f_i).

$(L_{i1} \le x_i \le L_{i2}, f_i)$, $i = 1,2,\ldots,k$

L_{i1} – donja granica i-tog razreda

L_{i2} – gornja granica i-tog razreda

f_i – frekvencija i-tog razreda

x_i– modalitet numeričkog obilježja i-tog razreda

Tablica 2.1. Distribucija frekvencija

	Obilježje (varijabla)	Frekvencija modaliteta obilježja	Distribucija frekvencija
	x_i	f_i	
Modaliteti obilježja	x_1	f_1	(x_1, f_1)
	x_2	f_2	(x_2, f_2)
	x_3	f_3	(x_3, f_3)
	…	…	…
	x_k	f_k	(x_k, f_k)
	Ukupno	$\sum f_i = N$	

Sturgersovo pravilo je pravilo o optimalnom broju i veličini razreda.

Za određivanje broja razreda *k* za *N* podataka koristi se izraz:

$k \approx 1 + 3{,}3 \log N$.

Određivanje širine (veličine) razredaΔx (čit. delta iks) ovisi o potrebama i zahtjevima istraživača. Razredi mogu biti jednaki i nejednaki. Ako su razredi jednaki, veličina im se aproksimativno određuje diobom raspona varijacija i broja razreda prema izrazu:

$$\Delta x = i_i \approx \frac{x_{\max} - x_{\min}}{k}$$

, pri čemu je:

Δx – širina razreda

$X_{\max}$ – najveća vrijednost numeričkog obilježja

$X_{\min}$ – najmanja vrijednost numeričkog obilježja

Primjer 2.1. Određivanje broja razreda i njihove veličine
Ako je istraživanjem obuhvaćeno 21 veliko trgovačko društvo i ako je utvrđeno da je najmanji broj zaposlenih u tim društvima 150, a najveći broj zaposlenih 190, tada je k = 5, a Δx = 8 zaposlenih.

Tablica 2.2. Granice razreda u distribuciji frekvencija

	Obilježje (varijabla)	Frekvencija modaliteta obilježja	Distribucija frekvencija
	x_i	f_i	
Modaliteti obilježja	$L_{11} – L_{12}$	f_1	(x_1, f_1)
	$L_{21} – L_{22}$	f_2	(x_2, f_2)
	$L_{31} – L_{32}$	f_3	(x_3, f_3)
	$L_{41} – L_{42}$	f_4	(x_4, f_4)
	$L_{51} – L_{52}$	f_5	(x_5, f_5)
	Ukupno	$\sum f_i = N$	

2.2. Statističke tablice

Statističke tablice koriste se radi lakšeg prikaza i razumijevanja prikupljenih podataka.

Tablica 2.3. Elementi statističke tablice

		Ime stupca				Zbirni stupac
		1	2	…	n	
Ime retka		Oznaka stupca	Oznake stupaca	Oznake stupaca	Oznaka stupca	Ukupno
Oznaka retka	1	a_{11}	a_{12}	…	a_{1n}	$\sum_{j=1}^{n} a_{1j}$
Oznake redaka	…	…	…	…	…	…
Oznaka retka	m	a_{m1}	a_{m2}	…	a_{mn}	$\sum_{j=1}^{n} a_{mj}$
Ukupno (Zbirni red)		$\sum_{i=1}^{m} a_{i1}$	$\sum_{i=1}^{m} a_{i2}$	…	$\sum_{i=1}^{m} a_{in}$	$\sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij}$

Izvor podataka:

Dodatne oznake:

- - nema pojave
- … ne raspolaže podatkom
- *0* *podatak manji od 0,5 upotrijebljene mjerne jedinice*
- *0,0* *podatak manji od 0,05 upotrijebljene mjerne jedinice*
- *0,00* *podatak manji od 0,005 upotrijebljene mjerne jedinice*
- *Ø* *prosjek*
- *( )* *nepotpun podatak*
- * *ispravljen podatak*
- *1)* *oznaka za napomenu ispod tablice*

Statističke tablice dijele se na:

- Jednostavne statističke tablice
- Skupne statističke tablice
- Kombinirane statističke tablice

Jednostavne statističke tablice prikazuju jedan statistički niz prema jednom obilježju.

Primjer 2.2. Jednostavna statistička tablica

Tablica 2.4. Poslovni subjekti u Republici Hrvatskoj, stanje 31. ožujka 2017. godine

Vrsta poslovnog subjekta	Registrirane pravne osobe
Trgovačka društva	190.021
Zadruge	4.254
Ustanove, tijela, udruge i organizacije[1)]	65.654
Ukupno	259.929

Izvor: Broj i struktura poslovnih subjekata u ožujku 2017.,URL

[1)]Uključena su tijela državne vlasti te tijela jedinica lokalne i područne (regionalne) samouprave

Skupne statističke tablice prikazuju dva ili više statističkih nizova grupiranih prema modalitetu jednog obilježja.

Primjer 2.3. Skupna statistička tablica

Tablica 2.5. Broj poduzetnika, broj zaposlenih, investicije i izvoz po regijama NUTS-2 razine u 2013. godini

Naziv	Broj poduzetnika	Broj zaposlenih	Investicije 2013.	Izvoz 2013.
Kontinentalna Hrvatska	61.993	603.821	29.061	73.080
Jadranska Hrvatska	39.198	227.107	8.429	24.086
Ukupno Republika Hrvatska	101.191	830.928	37.489	97.166

Izvor: Izvoz i investicije poduzetnika po regijama i županijama u 2013. godini, URL

Kombinirane statističke tablice prikazuju jedan statistički niz grupiran prema modalitetima dvaju ili više obilježja te zbog toga sadrži zbirni red i zbirni stupac.

Primjer 2.4. Kombinirana statistička tablica

Tablica 2.6. Zaštita industrijskog i intelektualnog vlasništva u razdoblju od 1997. do 2008. godine

NUTS II regija	Priznati patenti	Konsenzualni patenti	Žigovi	Industrijski dizajn	Ukupno
Sjeverozapadna Hrvatska	248	33	6.797	814	7.892
Panonska Hrvatska	42	7	778	114	941
Jadranska Hrvatska	160	14	1.390	418	1.982
Ukupno Republika Hrvatska	450	54	8.965	1.346	10.815

Izvor: Definiranje razvojnih prioriteta širih regija (NUTS 2) (2008), URL

2.3. Grafičko prikazivanje statističkih nizova[4]

Prednost grafičkog prikaza je u tome što se grafikonima mogu jasnije prikazati karakteristike promatrane pojave i međusobni odnos pojedinih podataka. Osnovna podjela grafikona je na površinske grafikone, linijske grafikone i kartograme.

[4]Kako napraviti grafikon u aplikaciji MS Excel detaljno je opisano: *Vrste grafikona*, Pearson Education, Inc., Dostupno na: http://pup.skole.hr/VodicHTML.aspx?xml=/datoteke/hr/hr/Student/PH/Tutorials/Understanding%20Chart%20Types/chart%20types.xml&locale=&app=&pageNumber=0

Dostupne vrste grafikona koje je moguće izraditi u aplikaciji MS Excel, Microsoft, *Available chart types*. Dostupno na: https://support.office.com/hr-hr/article/dostupne-vrste-grafikona-10b5a769-100d-4e41-9b0f-20df0544a683.

U **površinske grafikone** se ubrajaju: **histogram, stupci** koji mogu biti jednostavni, razdijeljeni, razdijeljeni-strukturni, dvostruki i višestruki, **površina kvadrata**, **strukturni krug**, **pravokutni dijagram**, **Tukeyev S-L** dijagram (dijagram stablo-list). Površinski grafikon je prikladan za prikazivanje nominalnih nizova (atributivni i geografski), ordinalnog niza, numeričkog niza i vremenskog intervalnog niza.

Najčešće korišteni **linijski grafikon** je **poligon frekvencija**. U linijskim grafikonima podaci se prikazuju linijama koje nastaju povezivanjem točaka (sjecište vrijednosti x i y). U linijske grafikone još se ubrajaju **kumulativni poligon frekvencija** i **dijagram rasipanja**.

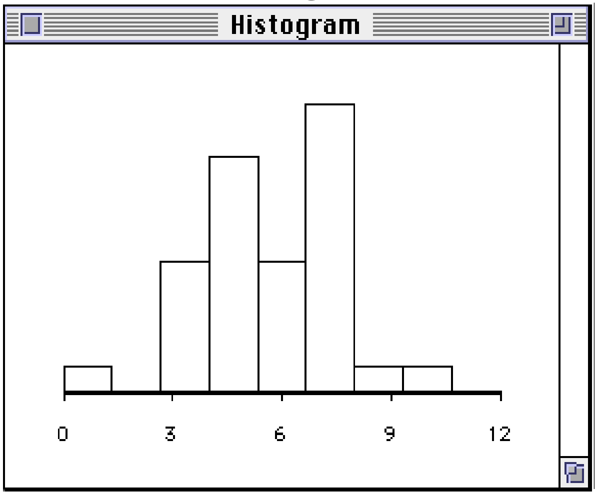
U **kartograme** se ubrajaju statistička karta, piktogram i dijagramska karta. Njima se prikazuju geografski nizovi. Geografski nizovi se prikazuju i svim oblicima površinskih grafikona kojima se prikazuju atributivni nizovi. **Kartogram** je geografska karta u kojoj su frekvencije predstavljene površinama geometrijskih likova ili odabranih znakova. Kartogram jasno prikazuje prostornu raspoređenost promatrane pojave.
Dijagramska karta koristi se u prikazivanju manjeg broja frekvencija geografskog niza**,** a frekvencije promatrane pojave prikazane su odgovarajućim površinama npr. pravokutnika, kvadrata, kruga. Način izračunavanja tih površina objašnjen je u nastavku. **Piktogram** se koristi za prikazivanje nizova s velikim brojem geografskih skupina, a frekvencije se prikazuju pomoću znakova tako da se odabrani znak u točno određenom omjeru ucrtava na površinu pripadajuće geografske skupine. Npr, ako jedna oznaka predstavlja 1000 zaposlenih za županiju u kojoj je zaposleno 50000 osoba, bit će ucrtano 50 točaka. **Statistička karta** se upotrebljava za prikazivanje nizova s velikim brojem geografskih skupina, a frekvencije se prikazuju sjenčanjem, crtanjem linija ili bojama različitog intenziteta. Prikladna je za prikazivanje relativnih brojeva, npr. postotak BDP *per capita* pojedine zemlje u odnosu na prosjek Europske unije.

U nastavku su prikazani često korišteni grafički prikazi[5].

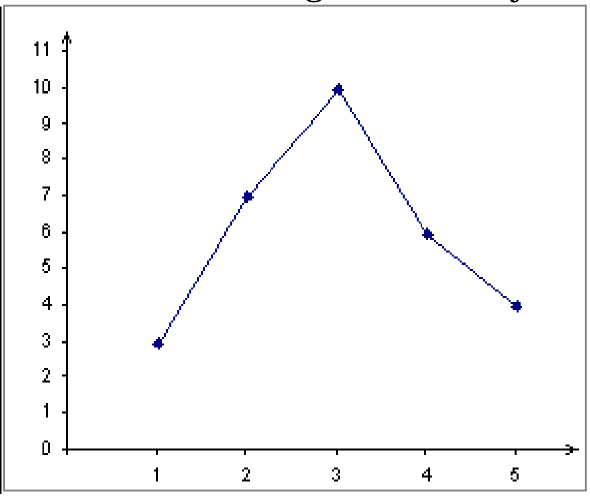
Pri izradi grafikona, na apscisu (os *x*) upisuju se vrijednosti obilježja, a na ordinatu (os *y*) vrijednosti frekvencija (apsolutne, relativne, korigirane).

Primjer 2.5. Najčešće korišteni grafički prikazi

Grafikon 2.1. Histogram



Grafikon 2.2. Poligon frekvencija

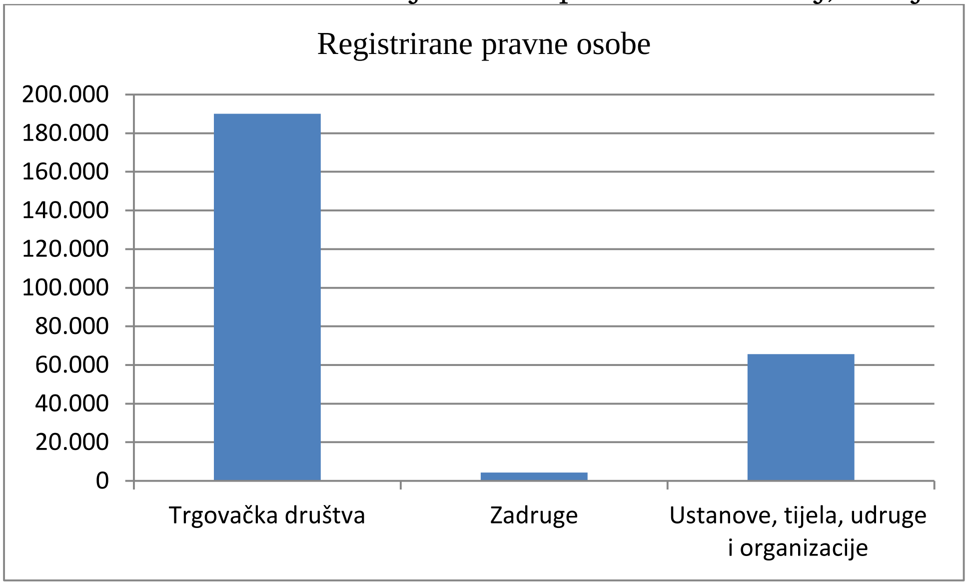


Podaci iz jednostavne tablice grafički se prikazuju **jednostavnim stupcima**. Podaci iz tablice 2.4. prikazani su grafikonom 2.3.

[5] Detaljnije o grafičkom prikazivanju statističkih podataka vidi: Horvat i Mijoč (2014), *Osnove statistike*, Drugo dopunjeno izdanje, Ljevak, Zagreb, str. 62 – 74.

Primjer 2.6. Grafički prikaz jednostavnim stupcima

Grafikon 2.3. Poslovni subjekti u Republici Hrvatskoj, stanje 31. ožujka 2017.

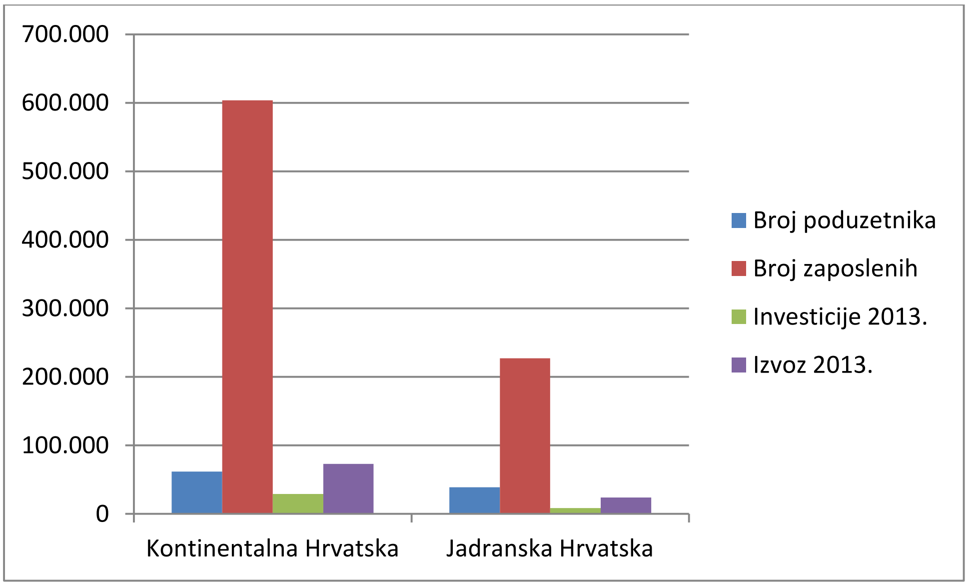


Izvor: Broj i struktura poslovnih subjekata u ožujku 2017.,URL

Podaci iz skupne statističke tablice grafički se prikazuju **dvostrukim ili višestrukim stupcima**. Podaci iz tablice 2.5. prikazani su grafikonom 2.4.

Primjer 2.7. Grafički prikaz višestrukim stupcima

Grafikon 2.4. Broj poduzetnika, broj zaposlenih, investicije i izvoz po regijama NUTS-2 razine u 2013. godini

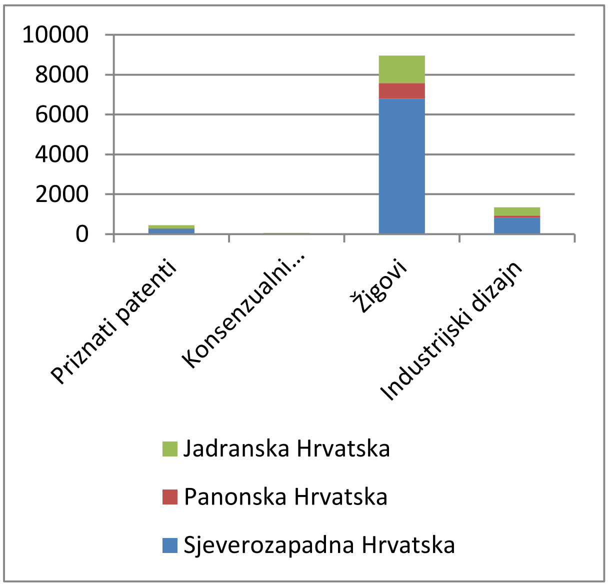


Izvor: Izvoz i investicije poduzetnika po regijama i županijama u 2013. godini, URL

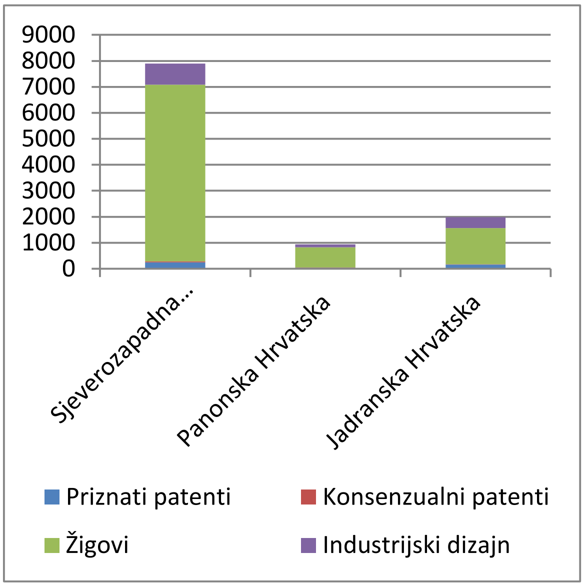
Podaci iz kombinirane statističke tablice grafički se prikazuju **razdijeljenim stupcima**. Podaci iz tablice 2.6. prikazani su grafikonima 2.5. i 2.6. Podaci iz kombinirane statističke tablice prikazana su na dva načina. U prvom primjeru je na os x upisana: vrsta intelektualnog vlasništva (atributivno obilježje), a u drugom primjeru je na os x upisano obilježje NUTS II regija Republike Hrvatske (geografsko obilježje).

Primjer 2.8. Grafički prikaz razdijeljenim stupcima

Grafikon 2.5. Zaštita industrijskog i intelektualnog vlasništva u razdoblju od 1997. do 2008. godine



Grafikon 2.6. Zaštita industrijskog i intelektualnog vlasništva u razdoblju od 1997. do 2008. godine



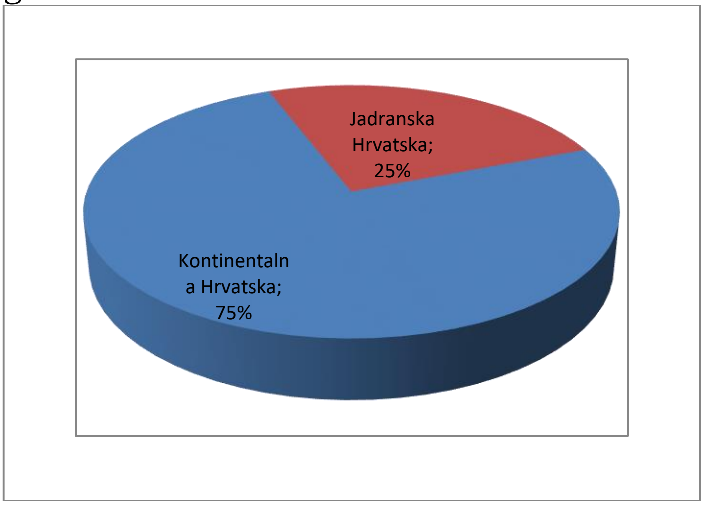
Izvor: Definiranje razvojnih prioriteta širih regija (NUTS 2) (2008), URL

Strukturnim stupcima i **strukturnim krugom** prikazuju se relativni brojevi strukture. Iako prikazuju strukturu brojčane vrijednosti na temelju kojih je grafikon napravljen, mogu se prikazati u relativnom i u apsolutnom broju (grafikon 2.7. i 2.8.) Svaki isječak u strukturnom krugu predstavlja broj stupnjeva kruga (dio/cjelina * 360°). U ovom primjeru broj stupnjeva kruga za izvoz Jadranske Hrvatske je 90° (0,25*360°).

$$p_i = \frac{f_i}{\sum_{i=1}^{k} f_i} * 360^\circ$$

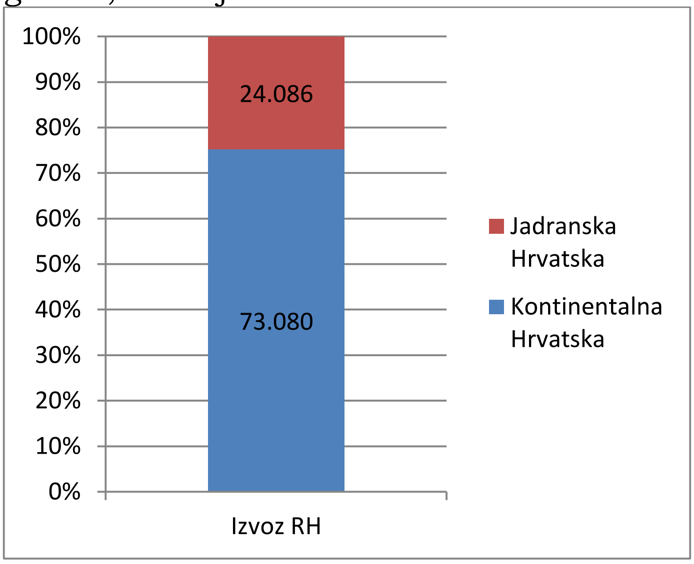
Primjer 2.9. Grafički prikaz strukturnim krugom i strukturnim stupcima

Grafikon 2.7. Izvoz Republike Hrvatske po regijama NUTS-2 razine u 2013. godini



Izvor: Izvoz i investicije poduzetnika po regijama i županijama u 2013. godini, URL

Grafikon 2.8. Izvoz Republike Hrvatske po regijama NUTS-2razine 2013. godine, u milijunima kuna



Izvor: Izvoz i investicije poduzetnika po regijama i županijama u 2013. godini, URL

Ukoliko je potrebno prikazati strukturu više nizova pomoću strukturnih krugova, veličina krugova određuje se prema izrazu za izračunavanje polumjera $r = \sqrt{P/\pi}$. P = $r^2\pi$.Pri tome P (površina kruga) predstavlja zbroj frekvencija svakog pojedinog niza. Na taj način omogućuje se međusobna usporedba nizova i usporedba unutar nizova.

Na isti način može se prikazati struktura pojave pomoću polukruga. Pri tome svaki isječak u **strukturnom polukrugu** predstavlja broj stupnjeva kruga (dio/cjelina * 180°).

Površina kvadrata koristi se kod usporedbe manjeg broja frekvencija. Površina kvadrata *P* određena je stranicom *a*. Pri tome P predstavlja frekvenciju koja se želi prikazati kvadratom P = a^2, $a = \sqrt{P}$.

Primjer 2.10. Izračun za stranica za prikaz podataka površinom kvadrata

Ako se 61.993 poduzetnika po NUTS II regijama Republike Hrvatske(RH) u 2013. godini želi prikazati površinom kvadrata, u mjerilu 1 cm^2 = 200 poduzetnika, tada će broj poduzetnika u Kontinentalnoj Hrvatskoj biti prikazan kvadratnom stranicom 1,245 cm (a = 249 cm, 249/200), a u Jadranskoj Hrvatskoj kvadratnom stranicom 0,99 cm (a = 198 cm, 198/200).

Ukoliko je potrebno grafički prikazati manji broj podataka pogodan je dijagram s točkama i dijagram stablo-list (S-L dijagram).

Kod **dijagrama s točkama** broj pojavljivanja svakog modaliteta obilježja (frekvencija) označava se točkama.

Primjer 2.11. Dijagram s točkama

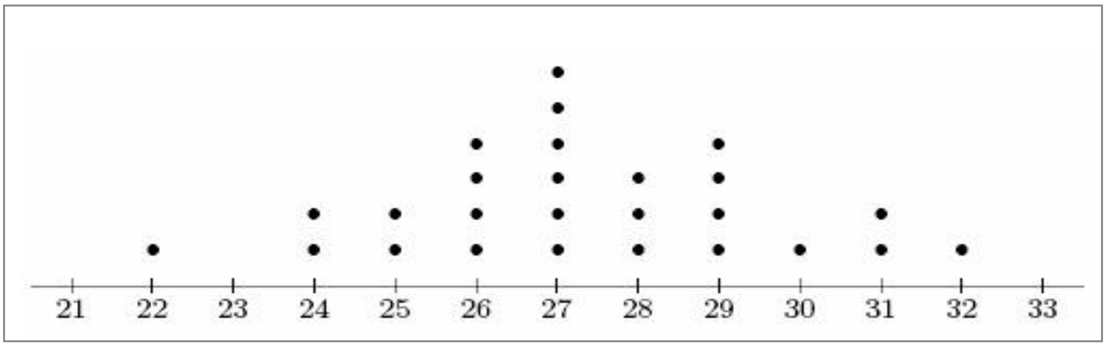
Broj posjetitelja seminara po danima održavanja:

x_i= 24 31 26 29 29 31 22 28 28 30 26 27 27 32
26 27 25 29 27 24 28 25 27 29 27 26

Podaci uređeni po veličini:

x_i = 22 24 24 25 25 26 26 26 26 27 27 27 27 27
28 28 28 29 29 29 29 30 31 31 32

Grafikon 2.9. Podaci prikazani grafičkim prikazom dijagram s točkama



Dijagram stablo-list (S-L dijagram), omogućuje uočavanje pojedinačnih vrijednosti modaliteta obilježja. Pri izradi dijagrama u stablo se upisuju desetice (bez ponavljanja), a u stupac list upisuju se jedinice.

Primjer 2.12. Dijagram stablo-list

Navršene godine staža zaposlenika u trgovačkom društvu A:

$x_i=$	22	2	1	5	8	15	14	15	15	15	27	29	33	26
	10	9	20	23	22	29	13	13	4	14	17	21	20	8

Grafikon 2.10. Podaci prikazani dijagram

Stablo	List	Frekvencije
0	1245889	7
1	03334455557	11
2	00122366799	11
3	3	1
		30

Broj prikazanih frekvencija u dijagramu prikazuje broj pojavljivanja individualne vrijednosti varijable. Iz dijagrama je vidljivo da je u trgovačkom društvu *A* zaposleno sedam zaposlenika koji imaju manje od navršenih 10 godina staža. U razredu od 10 do 20 godina nalazi se zaposlenik s najviše 17 navršenih godina staža, a u razredu od 30 godina i više, nalazi se zaposlenik koji ima 33 navršene godine staža.

3. RELATIVNI BROJEVI

Relativnim brojevima izračunava se relativni odnos dijelova u okviru cjeline ili relativnih odnos dviju ili više pojava koje predstavljaju relativne frekvencije. Relativne frekvencije izračunavaju se na temelju apsolutnih frekvencija i upravno su proporcionalne s apsolutnim. U sklopu deskriptivne statističke analize, relativni brojevi omogućuju osnovnu analizu podataka.

Vrste relativnih brojeva:

1. Relativni brojevi strukture
2. Relativni brojevi dinamike (indeksi)
3. Relativni brojevi koordinacije

3.1. Relativni brojevi strukture

Relativni brojevi strukture nastaju izračunavanjem strukture statističkog niza iskazane relativnim brojem. Relativne frekvencije strukture predstavljaju udio u cjelini, a izračunavaju se diobom i-te apsolutne frekvencije s ukupnim zbrojem apsolutnih frekvencija. Mogu se zapisati u obliku:

- proporcija $p_i = \frac{dio}{cjelina}$

$$p_i = \frac{f_i}{N} = \frac{f_i}{\sum_{i=1}^{k} f_i}, ...0 < p_i < 1 \xrightarrow{proporcije} \sum p_i = 1$$

- postotak $p_i = \frac{dio}{cjelina} * 100$

$$P_i = \frac{f_i}{N} \cdot 100 = \frac{f_i}{\sum_{i=1}^{k} f_i} \cdot 100, ...0 < P_i < 100 \xrightarrow{postoci} \sum P_i = 100$$

- promil (kada su dijelovi cjeline izrazito mali) $p_i = \frac{dio}{cjelina} * 1000$

$$P_i = \frac{f_i}{N} \cdot 1000 = \frac{f_i}{\sum_{i=1}^{k} f_i} \cdot 1000, ...0 < P_i < 100 \xrightarrow{promili} \sum P_i = 1000$$

Pokazatelji gospodarskih učinaka najčešće se prikazuju u postotcima. Tablični prikazi apsolutne vrijednosti mogu se preračunati u relativne tako da cjelinu predstavlja zbirna vrijednost iz kuta tablice, zbirni stupac ili zbirni redak (tablica 3.1.). Od toga potječu nazivi kutno 100, vodoravno 100 i okomito 100.

Kutno sto nastaje kada svaku apsolutnu vrijednost iz tablice stavimo u odnos prema ukupnoj statističkoj masi (N), odnosno zbrojem svih apsolutnih frekvencija i pomnožimo sa 100.

Vodoravno 100 nastaje kada svaku apsolutnu vrijednost iz reda tablice stavimo u odnos prema ukupnoj vrijednosti promatranog reda, dakle, prema vrijednostima iz zbirnog stupca. Izračunate proporcije pomnože se sa 100 te zbroj svih postotaka u retku iznosi 100, a nalaze se u zbirnom stupcu.

Okomito 100 nastaje kada svaku apsolutnu vrijednost iz stupca tablice stavimo u odnos prema ukupnoj vrijednosti promatranog stupca, dakle, prema vrijednostima iz zbirnog retka. Izračunate proporcije se pomnože sa 100 te zbroj svih postotaka u stupcu iznosi 100, a nalaze se u zbirnom retku.

Primjer 3.1. Izračun prema načelu kutno 100, vodoravno 100 i okomito 100

Tablica 3.1. Broj poslovnih subjekata zajedno sa subjektima u obrtu i slobodnim zanimanjima, stanje 31. ožujka 2017. godine, apsolutne vrijednosti i kutno 100

Apsolutne vrijednosti			
	Broj poslovnih subjekata		Ukupno
	Aktivni	Neaktivni	
Trgovačka društva	118.383	71.638	190.021
Zadruge	1.063	3.191	4.254
Ustanove, tijela, udruge i organizacije	29.606	36.048	65.654
Subjekti u obrtu i slobodnim zanimanjima	76.827	0	76.827
Ukupno	225.879	110.877	336.756

Kutno 100			
	Broj poslovnih subjekata		Ukupno
	Aktivni	Neaktivni	
Trgovačka društva	35,15	21,27	56,43
Zadruge	0,32	0,95	1,26
Ustanove, tijela, udruge i organiz.	8,79	10,7	19,5
Subjekti u obrtu i slobodnim zanimanjima	22,81	0	22,81
Ukupno	67,07	32,93	**100**

Izvor: Broj i struktura poslovnih subjekata u ožujku 2017.,URL

Tablica 3.2. Broj poslovnih subjekata zajedno sa subjektima u obrtu i slobodnim zanimanjima, stanje 31. ožujka 2017. godine, vodoravno 100 i okomito 100

Vodoravno 100			
	Broj poslovnih subjekata		Ukupno
	Aktivni	Neaktivni	
Trgovačka društva	62,3	37,7	**100**
Zadruge	24,99	75,01	**100**
Ustanove, tijela, udruge i organizacije	45,09	54,91	**100**
Subjekti u obrtu i slobodnim zanimanjima	100	0	**100**
Ukupno	67,07	32,93	**100**

Okomito 100			
	Broj poslovnih subjekata		Ukupno
	Aktivni	Neaktivni	
Trgovačka društva	52,41	64,61	56,43
Zadruge	0,47	2,88	1,26
Ustanove, tijela, udruge i organiz.	13,11	32,51	19,5
Subjekti u obrtu i slobodnim zanimanjima	34,01	0	22,81
Ukupno	**100**	**100**	**100**

Izvor: Broj i struktura poslovnih subjekata u ožujku 2017.,URL

3.2. Relativni brojevi dinamike

Relativni brojevi dinamike (indeksi) kvalitativnih nizova pokazuju odnos između stanja pojava promatranih prema kvalitativnom obilježju. Pokazuju koliko jedinica nekog modaliteta obilježja (brojnik),dolazi na 100 jedinica drugog modaliteta istog obilježja koji je izabran kao baza usporedbe i ne mijenja se (nazivnik).[6]

$$I_i = \frac{x_i}{x_1} * 100$$

I_i – individualni indeks

x_i – promatrani modalitet obilježja

x_1 – baza statističkog niza

$x_i > x_1 \rightarrow I_i > 100$

$x_i < x_1 \rightarrow I_i < 100$

$x_i = x_1 \rightarrow I_i = 100$

[6] Indeksi vremenskih nizova objašnjeni su u potpoglavlju 12.3.

Vrijednost izračunatog indeksa ne može biti negativan broj, za razliku od stope promjene koja će biti negativna ukoliko je brojnik manji od nazivnika, to jest ukoliko je vrijednost promatranog modaliteta obilježja manja od vrijednosti baze statističkog niza.

Stopa promjene izražava se u postotku i izračunava prema:

$$St_i = I_i - 100$$

Primjer 3.2. Izračun relativnih brojeva dinamike (indeksa) kvalitativnih podataka

Tablica 3.3. Bruto domaći izdatci za istraživanje i razvoj[7] prema sektorima u 2014.godini

Sektor	Bruto domaći izdaci (u 000 kuna)	Društvene znanosti = 100
Prirodne znanosti	550.493	231
Tehničke znanosti	908.779	381
Biomedicina i zdravstvo	515.683	216
Biotehničke znanosti	233.369	98
Društvene znanosti	238.536	**100**
Humanističke znanosti	145.422	60,96
Umjetničko područje	467	0,20
Interdisciplinarna područja znanosti	1.855	0,78

Izvor: Istraživanje i razvoj u 2014. (2015), URL

U primjeru su iznosi bruto domaćih izdataka za istraživanje i razvoj svakog znanstvenog sektora stavljeni u odnos s iznosom bruto domaćih izdataka u sektoru društvenih znanosti, čija vrijednost je određena kao baza usporedbe i zbog toga označena indeksom 100.

Indeks za bruto domaće izdatke istraživanja i razvoja u sektoru prirodnih znanosti iznosi:

$$I_i = \frac{550.493}{238.536} * 100 = 230,78$$

[7]Bruto domaći izdatci za istraživanje i razvoj su ukupni unutarnji izdatci za istraživanje i razvoj na području države u promatranoj kalendarskoj godini, a sastoje se od tekućih i investicijskih izdataka u bruto iznosima.

T: Bruto domaći izdatci istraživanja i razvoja u sektoru prirodnih znanosti u 2014. godini bili su za 130,78% veći (230,78 – 100 = 130,78) nego u sektoru društvenih znanosti.

Indeks za bruto domaće izdatke istraživanja i razvoja u sektoru biotehničkih znanosti iznosi:

$$I_i = \frac{233.369}{238.536} * 100 = 97,83$$

T: Bruto domaći izdaci istraživanja i razvoja u sektoru biotehničkih znanosti u 2014. godini bili su za 2,17% manji (97,83-100 = -2,17) nego u sektoru društvenih znanosti.

Indeks za bruto domaće izdatke istraživanja i razvoja u interdisciplinarna područja znanosti iznosi:

$$I_i = \frac{1.855}{238.536} * 100 = 0,78$$

T: Bruto domaći izdaci istraživanja i razvoja u sektoru interdisciplinarnih područja znanosti u 2014. godini bili su za 99,22% manji (0,78-100 = -99,22) nego u sektoru društvenih znanosti.

3.3. Relativni brojevi koordinacije

Relativni brojevi koordinacije(RBK) pokazuju odnos između dvije vrijednosti pojava koje su u smislenom odnosu. To su pokazatelji koji nastaju izračunavanjem omjera apsolutnih frekvencija dviju koordinirajućih veličina (npr. brutoplaća po zaposlenom, pokrivenost izvoza uvozom (robna razmjena s inozemstvom), BDP po jednom stanovniku (BDP *per capita*), broj stanovnika na 100 liječnika.

Ukoliko se zamjene brojnik i nazivnik dobije se recipročna vrijednost koja ima isto kvantitativno značenje.

$$RBK_1 = \frac{pojava_1}{pojava_2} \; RBK_2 = \frac{pojava_2}{pojava_1}$$

Primjer 3.3. Izračun relativnih brojeva koordinacije

Prema podacima DZS[8] u Hrvatskoj je u 2011. godini živjelo 4.284.889 stanovnika. Prema istom izvoru u Hrvatskoj je evidentirano 1.519.038 kućanstava.

RBK_1 = broj stanovnika/broj kućanstava

$$RBK_1 = \frac{4.284.889}{1.519.038} = 2{,}82 \, {}^{sta}\!/_{kuć}$$

T: U Hrvatskoj 2011. godine na svako kućanstvo otpadalo je 2,82 stanovnika, to jest prosječan broj stanovnika po kućanstvu je 2,82 stanovnika.

Ukoliko je brojnik znatno manji od nazivnika, poželjno je izračunati RBK pomnožiti s 10, 100, 1000, 10000 itd, radi lakšeg tumačenja rezultata.

RBK_2 = broj kućanstava/broj stanovnika · 10

$$RBK_2 = \frac{1.519.038}{4.284.889} = 0{,}3545 \cdot *10 = 3{,}55 = {}^{kuć}\!/_{10st}$$

T: U Hrvatskoj 2011. godine na 10 stanovnika dolazi 3,55 kućanstava.

Prema podacima DZS[9] iz Hrvatske je izvezeno robe u vrijednosti 104.348.016 tisuća kuna, a uvezeno je u Hrvatsku 162.681.147 tisuća kuna.

RBK_1 = izvoz/uvoz

$$RBK = \frac{104.348.016}{162.681.147} = 0{,}641$$

[8] Hrvatska u brojkama 2015, URL

[9] Robna razmjena Republike Hrvatske s inozemstvom − privremeni podaci od siječnja do prosinca 2017. i za siječanj 2018.,URL

T: Na svaku kunu izvezene robe u inozemstvo, u Hrvatsku je uvezeno robe u vrijednosti od 0,641 kn.

Navedeni izračun koristi se za iskazivanje uobičajenog pokazatelja pokrivenosti izvoza uvozom koji se iskazuje u %. Tako je u istoj publikaciji Državnog zavoda za statistiku objavljen navedeni pokazatelj i on iznosi 64,1%.

RBK_2 = uvoz/izvoz

$$RBK_2 = \frac{162.681.147}{104.348.016} = 1{,}559$$

T: Na svaku kunu uvezene robe u Hrvatsku, izvezeno je robe za 1,559 kuna.

Relativni broj koordinacije grafički se prikazuje površinskim grafikonom, tzv. Varzarovim znakom, pri čemu se vrijednost nazivnika nanosi na apscisu (os x), a relativni broj koordinacije na ordinatu (os y).[10]

Varzarovim znakom mogu se grafički prikazati i relativni brojevi strukture pri čemu se vrijednost cjeline nanosi na apscisu (os x), a relativni broj strukture na ordinatu (os y).

[10] Detaljnije o grafičkom prikazivanju Varzarovim znakom vidi: Horvat i Mijoč (2014), str. 84.

4. NUMERIČKI NIZOVI

Numerički nizovi nastaju uređenjem podataka koji predstavljaju vrijednosti numeričke varijable (obilježja). Način njihova uređivanja ovisi o tome jesu li podaci diskontinuirani ili kontinuirani.
Numerički niz, odnosno distribucija frekvencija sastoji se od parova (x_i, f_i), i=1,2,....,k. Radi numeričke analize potrebno je odrediti **granice razreda**, **veličinu razreda** i **sredinu razreda**.

Nominalne granice su one koje su zabilježene prikupljanjem i popisom podataka. Kod nominalnih granica donja granica *i*-tog razreda nije jednaka gornjoj granici prethodnog razreda ($L_{i1} \neq L_{i2-1}$). Takve granice se u statističkim postupcima preračunavaju u prave ili precizne. Prema tome, **prave ili precizne granice** su one kod kojih je donja granica tekućeg razreda jednaka gornjoj granici prethodnog razreda ($L_{i1} = L_{i2-1}$).

Prije statističke analize numeričkog niza potrebno je ispravno odrediti granice razreda ovisno o vrsti statističkog obilježja.

1. **Diskontinuirana (diskretna, prekidna) varijabla**

Diskontinuirano numeričko obilježje sadrži podatke koji su cjelobrojni, npr. broj stanovnika, broj djece, broj poslovnih subjekata, broj kućanstava.

Kod diskontinuirane numeričke varijable potrebno je konstruirati **precizne granice**, tako da se donje granice razreda povećavaju, a gornje granice smanjuju za polovicu razlike između donje granice tekućeg razreda i gornje granice prethodnog razreda. Najčešće je to polovica jedinične vrijednosti numeričkog obilježja.

Primjer 4.1. Preračunavanje nominalnih u precizne granice, određivanje sredine razreda i veličine razreda

Tablica 4.1. Zaposlenici u poslovnom subjektu Alfa prema broju djece

Broj djece	Broj zaposlenih	Precizne granice		Sredina razreda	Veličina razreda
x_i	f_i	L_{1i}	L_{2i}	x_i	i_i
0–1	14	0	1,5	0,75	1,5
2–4	29	1,5	4,5	3	3
5–7	12	4,5	7,5	6	3
8–9	2	7,5	9,5	8,5	2
Ukupno	57				

2. Kontinuirana (neprekidna) varijabla

Kontinuirano numeričko obilježje je ono koje sadrži podatke koji se mogu prikazati i cijelim i decimalnim brojem. Npr. godine staža zaposlenika, visina dobiti poduzeća, iznos investicija.

Ukoliko su kod kontinuirane numeričke varijable granice razreda prekinute, potrebno ih je „preračunati“ u prave, tako da je gornja granica tekućeg razreda jednaka donjoj granici sljedećeg razreda. Nominalne granice „preračunavaju“ se u **prave granice** da se gornja granica tekućeg razreda poveća za 1 (jedan), tako da bude jednaka donjoj granici sljedećeg razreda.

Primjer 4.2. Preračunavanje nominalnih u prave granice, određivanje sredine razreda, veličine razreda i korigiranih frekvencija

Tablica 4.2. Zaposlenici u poslovnom subjektu Alfa prema dobi

Dob zaposlenika	Broj zaposlenih	Prave granice		Sredina razreda	Veličina razreda	Korigirane frekvencije
x_i	f_i	L_{1i}	L_{2i}	x_i	i_i	fc_i
18 – 20	2	18	21	19,5	3	0,67
21 – 24	6	21	25	23,0	4	1,50
25 – 29	9	25	30	27,5	5	1,80
30 – 34	5	30	35	32,5	5	1,00
35 – 39	4	35	40	37,5	5	0,80
40 – 44	9	40	45	42,5	5	1,80
45 – 49	5	45	50	47,5	5	1,00
50 – 54	11	50	55	52,5	5	2,20
55 – 59	4	55	60	57,5	5	0,80
60 – 64	2	60	65	62,5	5	0,40
$\sum$	57					

Sredina razreda određuje se kao poluzbroj gornje i donje prave ili precizne granice razreda, prema: $x_i = \frac{L_{1i} + L_{2i}}{2} \quad i = 1,2,...,k$

Prvi i posljednji razred mogu biti **otvoreni razredi.** Vrijednosti donje granice prvog razreda i gornje granice posljednjeg razreda procjenjuju se i stavljaju u zagradu.

Veličine razreda određuju se prema: $i_i = L_{i2} - L_{i1} \quad i = 1,2,...,k$

Kada su razredi različitih veličina, potrebno je korigirati frekvencije. Frekvencije *i*-tog razreda korigiraju se dijeljenjem s veličinom *i*-tog razreda. Kada sve veličine razreda nemaju zajednički nazivnik,**korigirane frekvencije** dobiju se pomoću sljedećeg izraza:

$$fc_i = \frac{f_i}{i_i}$$

Navedeni izraz može se koristi prilikom korigiranja frekvencija neovisno o veličini razreda, ali pogodnije je da, ukoliko veličine razreda imaju zajednički nazivnik, skratiti ih te frekvencije podijeliti s dobivenim veličinama pomoću sljedećeg izraza.

$$fc_i = \frac{f_i}{\frac{i_i}{i_b}}$$

Primjer 4.3.Izračun korigiranih frekvencija skraćivanjem veličine razreda

Tablica 4.3. Mikro trgovačka društva prema ostvarenoj dobiti

Ostvarena dobit (u 000 kuna)	Broj mikro trgovačkih društava	Prave granice	Prave granice	Veličina razreda	Korigirane frekvencije	$i_b = 10$	Korigirane frekvencije
x_i	f_i	L_{1i}	L_{1i}	i_i	fc_i	i_i / i_b	$fc_i = f_i/(i_i / i_b)$
0 – 10	23	0	10	10	2,30	1	23
10 – 20	33	10	20	10	3,30	1	33
20 – 30	31	20	30	10	3,10	1	31
30 – 50	57	30	50	20	2,85	2	28,5
50 – 100	69	50	75	50	1,38	5	13,80
100 – 200	21	100	200	100	0,21	10	2,10
200 – 500	32	200	500	300	0,11	30	1,07
500 – 1.000	3	500	1000	500	0,01	50	0,06
1.000 – 5.000	1	1000	5000	5000	0,00	500	0,002
Ukupno	270						

U navedenom primjeru zajednički nazivnik i_b je 10. Veličine razreda su skraćene s 10, a potom su frekvencije podijeljene dobivenim veličinama. Ovako izračunate korigirane frekvencije omogućuje pregledan grafički prikaz distribucije frekvencija. Prilikom izrade histograma ili poligona frekvencija u koordinatni sustav se unose vrijednosti korigiranih frekvencija.

Kumulativni niz nastaje postupnim zbrajanjem apsolutnih ili relativnih frekvencija. Kumulativni niz može nastati zbrajanjem frekvencija, počevši od pripadajućih frekvencija najmanje vrijednosti obilježja, pri čemu nastaje kumulativni niz „manje od“ i zbrajanjem frekvencija, počevši od pripadajućih frekvencija najveće vrijednosti obilježja, pri čemu nastaje kumulativni niz „više od“.

Kumulativni niz pokazuje koliko je jedinica niza manje ili veće od određene vrijednosti obilježja. Grafički se prikazuje kumulantom. Kumulanta niza „manje od“ nastaje spajanjem parova točaka gornje granice razreda, a kumulanta niza „više od“ nastaje spajanjem parova točaka donje granice razreda.

Primjer 4.4. Izračunavanje kumulativnog niza „manje od“ i „više od“

Tablica 4.4. Zaposlenici u poslovnom subjektu Alfa prema dobi

Dob zaposlenika	Broj zaposlenih	Prave granice		Kumulativni niz	
x_i	f_i	L_{1i}	L_{2i}	"manje od"	"više od"
18 – 20	2	18	21	2	57
21 – 24	6	21	25	8	55
25 – 29	9	25	30	**17**	49
30 – 34	5	30	35	22	40
35 – 39	4	35	40	26	35
40 – 44	9	40	45	35	31
45 – 49	5	45	50	40	**22**
50 – 54	11	50	55	51	17
55 – 59	4	55	60	55	6
60 – 64	2	60	65	57	2
Ukupno	57				

Treći član kumulativnog niza „manje od“ pokazuje da 17 zaposlenih osoba u poslovnom subjektu Alfa ima manje od 30 godina (manje od gornje granice), a npr. sedmi član kumulativnog niza „više od“ pokazuje da 22 zaposlene osobe u poslovnom subjektu Alfa imaju više od 45 godina (više od donje granice).

Ako su podaci grupirani kumulativni niz je:

$$S(X \le x_i) = S(x_i) = \sum_{j=1}^{i} f_j \qquad i = 1,2,....., k$$

Ako se zbrajaju relativne frekvencije izražene kao proporcije nastaje empirijska distribucija frekvencije. Ona je kumulativni niz proporcija (relativnih frekvencija).

$$F(X \le x_i) = F(x_i) = \sum_{j=1}^{i} p_j \qquad i = 1,2,....., k$$

S(x)i je broj, a *F(x)i* proporcija jedinica koje imaju vrijednost manju od gornje granice i-tog razreda.

Primjer 4.5. Izračunavanje funkcije distribucije

Tablica 4.5. Zaposlenici u poslovnom subjektu Alfa prema broju djece

Dob zaposlenika	Broj zaposlenih	Prave granice		Kumulativni niz"manje od"	Struktura	Funkcija distribucije
x_i	f_i	L_{1i}	L_{2i}	$S(x_i)$	p_i	$F(x_i)$
18 – 20	2	18	21	2	0,0351	0,0351
21 – 24	6	21	25	8	0,1053	0,1404
25 – 29	9	25	30	**17**	0,1579	**0,2982**
30 – 34	5	30	35	22	0,0877	0,3860
35 – 39	4	35	40	26	0,0702	0,4561
40 – 44	9	40	45	35	0,1579	0,6140
45 – 49	5	45	50	40	0,0877	0,7018
50 – 54	11	50	55	51	0,1930	0,8947
55 – 59	4	55	60	55	0,0702	0,9649
60 – 64	2	60	65	57	0,0351	1,0000
Ukupno	57				1,0000	

Treći član funkcije distribucije $F(x_3)$ pokazuje da 29,82% zaposlenih osoba u poslovnom subjektu Alfa ima manje od 30 godina.

5. MJERE SREDIŠNJE TENDENCIJE (SREDNJE VRIJEDNOSTI)

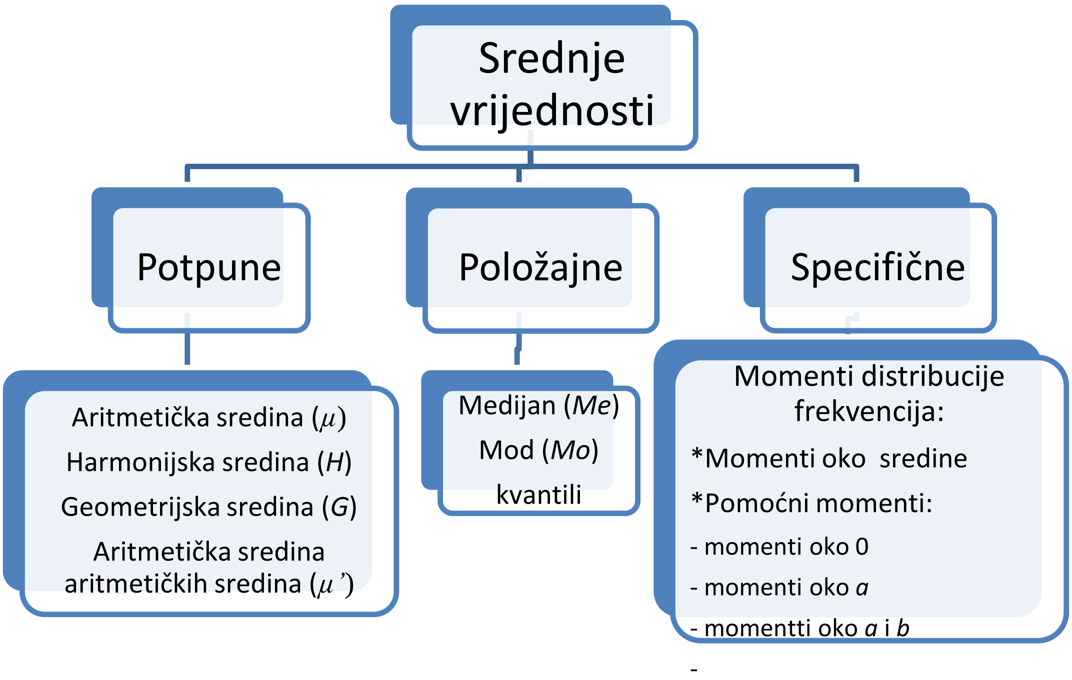
Mjerama središnje tendencije opisuje se srednja vrijednost promatrane pojave. Mjere središnje tendencije koje se računaju za populaciju nazivaju se **parametri**, a koje se računaju za uzorak nazivaju se **pokazatelji uzorka**. Dijele se na potpune, položajne i specifične.

Potpune srednje vrijednosti su: aritmetička sredina, geometrijska sredina, harmonijska sredina i aritmetička sredina aritmetičkih sredina. Pri izračunavanju potpunih srednjih vrijednosti sudjeluju svi članovi statističkog niza.

Položajne srednje vrijednosti su medijan, mod i kvantili. Pri izračunu položajnih srednjih vrijednosti sudjeluju samo one vrijednosti statističkog obilježja koje su određene položajem u statističkom nizu.

Specifične srednje vrijednosti su momenti numeričkog niza koje predstavljaju aritmetičke sredine odstupanja vrijednosti numeričke varijable od njezine aritmetičke sredine ili neke druge vrijednosti, računate na neku potenciju. Dijele se na momente oko sredine i pomoćne momente (pomoćni moment oko nule, pomoćni moment oko a i pomoćni moment oko a i b).

Slika 5.1. Vrste srednjih vrijednosti



5.1. Aritmetička sredina

Aritmetička sredina (engl. *mean, average*) je srednja vrijednost koja se najčešće primjenjuje u proučavanju pojava. Označava se simbolom μ (grčko slovo, čit. mi), ukoliko predstavlja srednju vrijednost populacije. Ukoliko označava srednju vrijednost uzorka, označava se $\overline{x}$ (čit. iks prosjek). Izračunava se dijeljenjem totala (zbroja vrijednosti numeričkog obilježja) s ukupnim brojem statističkih jedinica (ukupnom frekvencijom).

5.1.1. Jednostavna aritmetička sredina

Aritmetička sredina negrupiranih podataka izračunava se pomoću izraza za izračun **jednostavne aritmetičke sredine.**

$$\mu = \frac{\sum_{i=1}^{N} x_i}{N}$$

Primjer 5.1. Izračun jednostavne aritmetičke sredine

Utvrđeno je da je broj prodanih proizvoda B u 5 prodavaonica, u ponedjeljak bio 2, 12, 3, 4, 17.

$$\mu = \frac{2+12+3+4+17}{5} = 7{,}6$$

Prosječan broj prodanih proizvoda B u 5 prodavaonica, u ponedjeljak je bio 7,6 proizvoda.

5.1.2. Vagana aritmetička sredina

Kada su podaci zapisani u obliku distribucije frekvencija (grupirani podaci), izračunava se **vagana (ponderirana) aritmetička sredina.**

$$\mu = \frac{\sum_{i=1}^{k} f_i x_i}{\sum_{i=1}^{k} f_i}$$

Primjer 5.2. Izračun vagane aritmetičke sredine

Tablica 5.1. Zaposlenici u poslovnom subjektu Alfa prema broju djece

Broj djece	Broj zaposlenih	Precizne granice		Sredina razreda	$f_i x_i$	p_i	$p_i x_i$	$P_i x_i$
x_i	f_i	L_{1i}	L_{2i}	x_i				
0 – 1	14	0	1,5	0,75	10,50	0,25	0,1842	18,42
2 – 4	29	1,5	4,5	3,00	87,00	0,51	1,5263	152,63
5 – 7	12	4,5	7,5	6,00	72,00	0,21	1,2631	126,32
8 – 9	2	7,5	9,5	8,50	17,00	0,04	0,2982	29,82
Ukupno	57				186,50	1,00	3,2719	327,19

$$\mu = \frac{186{,}5}{57} = 3{,}27$$

U poslovnom subjektu Alfa zaposlenici u prosjeku imaju 3,27 djece.

Tablica 5.2. Zaposlenici u poslovnom subjektu Alfa prema dobi

Dob zaposlenika u godinama	Broj zaposlenih	Prave granice		Sredina razreda	$f_i x_i$	p_i	$p_i x_i$	$P_i x_i$
x_i	f_i	L_{1i}	L_{2i}	x_i				
18 – 20	2	18	21	19,5	39,00	0,0351	0,68	68,42
21 – 24	6	21	25	23,0	138,00	0,1053	2,42	242,11
25 – 29	9	25	30	27,5	247,50	0,1579	4,34	434,21
30 – 34	5	30	35	32,5	162,50	0,0877	2,85	285,09
35 – 39	4	35	40	37,5	150,00	0,0702	2,63	263,16
40 – 44	9	40	45	42,5	382,50	0,1579	6,71	671,05
45 – 49	5	45	50	47,5	237,50	0,0877	4,17	416,67
50 – 54	11	50	55	52,5	577,50	0,1930	10,13	1013,16
55 – 59	4	55	60	57,5	230,00	0,0702	4,04	403,51
60 – 64	2	60	65	62,5	125,00	0,0351	2,19	219,30
Ukupno	57				2.289,50	1,00	40,17	4.016,67

$$\mu = \frac{2.289{,}50}{57} = 40{,}17$$

T: Prosječna dob zaposlenika u poslovnom subjektu Alfa je 40,17 godina.

Total je zbroj vrijednosti numeričke varijable (u ovom primjeru 2.289,50), a aritmetička sredina je jednaki dio totala po jedinici (2.289,50/57).

Budući da su relativne frekvencije upravno proporcionalne apsolutnim, aritmetička sredina može se izračunati iz proporcija ili postotaka koristeći sljedeće izraze:

$$\mu = \frac{\sum_{i=1}^{k} p_i x_i}{1} \quad \text{ili} \quad \mu = \frac{\sum_{i=1}^{k} P_i x_i}{100}$$

Primjer 5.3. Izračun aritmetičke sredine iz proporcija i postotaka

Prema podacima iz tablice 5.1. aritmetička sredina je izračunata iz proporcija i postotaka i iznosi 3,27 djece. Rezultat je isti neovisno računamo li srednje vrijednosti iz apsolutnih i relativnih frekvencija.

$$\mu = \frac{3{,}27}{1} = 3{,}27 \quad \text{ili} \quad \mu = \frac{327{,}19}{100} = 3{,}27$$

Prema podacima iz tablice 5.2. aritmetička sredina izračunata iz relativnih frekvencija (postotaka) iznosi 40,17 godina.

$$\mu = \frac{4.016{,}67}{100} = 40{,}17$$

Za izračun aritmetičke sredine moguće je koristiti **kodiranje vrijednosti numeričke varijable**, tzv. **linearna transformacija**. Vrijednost numeričke varijable se smanjuje za *a* (neku od sredina razreda u nizu), te skraćuje za *b*, veličinu razreda, prema:

$$d_i = \left(\frac{x_i - a}{b} \right)$$

a i b – transformacijske konstante

$$\mu = a + b\overline{d} \qquad \text{ili} \qquad \mu = a + b\frac{\sum_{i=1}^{k} f_i d_i}{\sum_{i=1}^{k} f_i}$$

Linearnu transformaciju moguće je provesti i smanjivanjem vrijednosti varijable samo za vrijednost *a*, prema izrazu:

$$\mu = a + \frac{\sum_{i=1}^{k} f_i d_i}{\sum_{i=1}^{k} f_i} \qquad \text{gdje je} \qquad d_i = x_i - a$$

Primjer 5.4. Izračun aritmetičke sredine linearnom transformacijom

Tablica 5.3. Zaposlenici u poslovnom subjektu Alfa prema dobi

Dob zaposlenika u godinama (prave granice)		Broj zaposlenih	Sredine razreda	Veličine razreda	$d_i=(x_i–a)/b$	$f_i d_i$	$d_i= x_i–a$	$f_i d_i$
L_{1i}	L_{2i}	f_i	x_i	i_i				
18	21	2	19,5	3	-5,63	-11,250	-22,50	-45,00
21	25	6	23,0	4	-4,75	-28,500	-19,00	-114,00
25	30	9	27,5	5	-3,63	-32,625	-14,50	-130,50
30	35	5	32,5	5	-2,38	-11,875	-9,50	-47,50
35	40	4	37,5	5	-1,13	-4,500	-4,50	-18,00
40	45	9	42,5	5	0,13	1,125	0,50	4,50
45	50	5	47,5	5	1,38	6,875	5,50	27,50
50	55	11	52,5	5	2,63	28,875	10,50	115,50
55	60	4	57,5	5	3,88	15,500	15,50	62,00
60	65	2	62,5	5	5,13	10,250	20,50	41,00
		57			-4,38	-26,1250	-17,50	-104,50

$$\mu = 42 + 4 \cdot \frac{-26{,}125}{57} = 42 + 4 \cdot (-0{,}458333) = 40{,}17$$

ili

$$\mu = 42 + \frac{-104{,}50}{57} = 42 + (-1{,}8333) = 40{,}17$$

T: Prosječna dob zaposlenika u poslovnom subjektu Alfa je 40,17 godina.

Svojstva aritmetičke sredine omogućuju daljnje razumijevanje aritmetičke sredine i odstupanja od aritmetičke sredine. Aritmetička sredina ima pet svojstava:

1. svojstvo aritmetičke sredine

Zbroj vrijednosti odstupanja individualnih vrijednosti numeričke varijable od aritmetičke sredine jednak je nuli.

$$\sum (x_i - \mu) = 0$$

$$\sum f(x_i - \mu) = 0$$

Primjer 5.5. Dokaz 1.svojstva aritmetičke sredine

Tablica 5.4. Mikro trgovačka društva prema broju zaposlenih

Broj zaposlenih	Broj mikro trgovačkih društava	$f_i x_i$	$f_i(x_i-\mu)$	$f_i(x_i-\mu)^2$
x_i	f_i			
0	28	0	-105,05185	394,14
1	54	54	-148,60000	408,93
2	38	76	-66,57037	116,62
3	31	93	-23,30741	17,52
4	28	112	6,94815	1,72
5	25	125	31,20370	38,95
9	27	243	141,70000	743,66
7	12	84	38,97778	126,61
8	17	136	72,21852	306,79
9	10	90	52,48148	275,43
Ukupno	270	1013	0,00	2.430,37

$$\mu = \frac{1013}{270} = 3{,}7519$$

2. svojstvo aritmetičke sredine

Zbroj kvadrata odstupanja individualnih vrijednosti numeričke varijable od aritmetičke sredine jednak je minimumu.

$\sum(x_i-\mu)^2=\min$

$\sum f(x_i-\mu)^2=\min$

U tablici 5.2. zbroj kvadrata odstupanja iznosi 2.430,37, dok bi izračun odstupanja od bilo koje druge vrijednosti, veće ili manje od aritmetičke sredine, rezultirao većim brojem od 2.430,37.

3. svojstvo aritmetičke sredine

Aritmetička sredina uvijek se nalazi između najmanje i najveće vrijednosti numeričkog obilježja varijable x_i. Iz prethodnog primjera, 3,75 prosječno zaposlenih je manje od 0 i veće od 9 zaposlenih, odnosno 3,75 se nalazi između 0 i 9.

$x_{\min}<\mu<x_{\max}$ $0<3{,}75<9$

4. svojstvo aritmetičke sredine

Ako je vrijednost varijable x_i jednaka konstanti c, aritmetička sredina te varijable jednaka je konstanti c. Npr. ako šest poslovnih subjekata ima po pet zaposlenih, aritmetička sredina iznosit će 5 zaposlenih.

$x_1=x_2=x_3=....=x_n=c\Rightarrow\mu=c$ $5+5+5+5+5+5=30\Rightarrow 30/6=5$

5. svojstvo aritmetičke sredine

Aritmetička sredina sklona je *izdvojenicama* (ekstremima).

U tablici 4.3, u kojoj su prikazana mikro trgovačka društva prema ostvarenoj dobiti, izdvojenicu predstavlja trgovačko društvo koje je ostvarilo dobit od 1.000.000 do 5.000.000 kuna i znatno će povećati prosječnu vrijednost ostvarene dobiti navedenih trgovačkih društava. Zbog toga će aritmetička sredina izgubiti na svojoj reprezentativnosti.

5.1.3. Aritmetička sredina aritmetičkih sredina

Aritmetička sredina aritmetičkih sredina (μ')izračunava se na temelju već izračunatih aritmetičkih sredina. Aritmetička sredina aritmetičkih sredina je vagana aritmetička sredina pri čemu se aritmetičke sredine μ_1, μ_2, μ_3, ….,μ_k ponderiraju s brojem elemenata podskupova N_1, N_2, N_3, …,N_k, prema izrazu

$$\mu' = \frac{N_1 \cdot \mu_1 + N_2 \cdot \mu_2 + ... + N_k \cdot \mu_k}{N_1 + N_2 + ... + N_k} = \frac{\sum_{i=1}^{k} N_i \mu_i}{\sum_{i=1}^{k} N_i}.$$

Primjer 5.6. Izračun aritmetičke sredine aritmetičkih sredina

Tablica 5.5. Broj zaposlenih i prosječne brutoplaće po županijama Republike Hrvatske, na dan 31. ožujka 2015. godine

	Broj zaposlenih prema administrativnim izvorima	Prosječne bruto plaće zaposlenih	
Županije Republike Hrvatske	N_i	μ_i	$N_i \mu_i$
Zagrebačka	76.605	7.720	591.390.600
Krapinsko-zagorska	33.326	6.673	222.384.398
Sisačko-moslavačka	37.550	7.055	264.915.250
Karlovačka	33.786	7.223	244.036.278
Varaždinska	59.259	6.176	365.983.584
Koprivničko-križevačka	29.818	7.112	212.065.616
Bjelovarsko-bilogorska	28.683	6.407	183.771.981
Primorsko-goranska	100.844	7.863	792.936.372
Ličko-senjska	12.887	6.989	90.067.243
Virovitičko-podravska	18.294	6.291	115.087.554
Požeško-slavonska	16.754	6.584	110.308.336
Brodsko-posavska	31.973	6.753	215.913.669
Zadarska	44.133	7.350	324.377.550
Osječko-baranjska	76.556	7.077	541.786.812
Šibensko-kninska	27.706	7.092	196.490.952
Vukovarsko-srijemska	38.399	6.587	252.934.213
Splitsko-dalmatinska	128.528	7.301	938.382.928
Istarska	75.363	7.830	590.092.290
Dubrovačko-neretvanska	37.082	7.838	290.648.716
Međimurska	35.392	6.311	223.358.912
Grad Zagreb	389.888	9.742	3.798.288.896
Ukupno	1.332.826	149.974	10.565.222.150

Izvor: Zaposlenost i plaće, Stopa registrirane nezaposlenosti i zaposlenost, Državni zavod za statistiku,URL

$$\mu' = \frac{10.565.222.150}{1.332.826} = 7.926{,}93$$

Prosječna brutoplaća zaposlenih u svim županijama, odnosno zaposlenih u Republici Hrvatskoj je 7.926,93 kune[11].

5.2. Harmonijska sredina

Harmonijska sredina je recipročna vrijednost aritmetičke sredine izračunata iz recipročnih vrijednosti varijable x. Koristi se za izračunavanje[12]:

- sredine relativnih brojeva s jednakim brojnicima
- sredine relativnih brojeva koordinacije, kada su poznati brojnici relativnih brojeva, a nazivnici nepoznati
- prosječnog vremena za izradu jedinice proizvoda
- prosječnog vremena potrebnog za obrtaj kapitala
- izračunavanje prosječnog vremena prijeđene jedinice puta i slično.

Izrazi za izračunavanje harmonijske sredine negrupiranih i grupiranih podataka:

za negrupirane

$$H = \frac{N}{\sum_{i=1}^{N} \frac{1}{x_i}}$$

za grupirane podatke

$$H = \frac{\sum_{i=1}^{k} f_i}{\sum_{i=1}^{k} \frac{1}{x_i} f_i}$$

Harmonijska sredina u analizi poslovanja najčešće se koristi za izračunavanje proizvodnosti rada zaposlenih, izraženu u vremenu potrebnom za izradu proizvoda.

---

11 Prema metodološkim objašnjenjima Državnog zavoda za statistiku prosječne brutoplaće po zaposlenome izračunavaju se dijeljenjem ukupnih isplata s brojem zaposlenih koji su primili isplate, kao što je i ovdje prikazano. Međutim prema istim izvorima prosječna brutoplaća za Republiku Hrvatsku 2015. godine iznosi 7.978 kn. Zbog nedostatka dodatnih metodoloških objašnjenja nije moguće dati objašnjenje zbog čega je nastala razlika.

12 Horvat i Mijoč (2014, str. 120)

Primjer 5.7. Izračun harmonijske sredine

U mjesecu ožujku 2018. godine 24 djelatnika za izradu proizvoda A, B, C i D utrošili su 4320 sati rada

Tablica 5.6. Utrošeno vrijeme za izradu proizvoda A, B, C i D u ožujku 2018.

Proizvod	Utrošeno vrijeme za izradu proizvoda u satima	Broj osoba koje sudjeluju u izradi pojedinog proizvoda	
	x_i	f_i	f_i / x_i
A	1315	10	0,0076
B	1220	5	0,0041
C	580	2	0,0034
D	1205	7	0,0058
Ukupno	4320	24	0,0210

$$H = \frac{24}{0,0210} = 1.142,86$$

Za izradu proizvoda A, B, C i D,24 zaposlenika jeu ožujku 2018. godine u prosjeku potrošilo 1.142,86 sati rada po pojedinoj vrsti proizvoda.

Harmonijska sredina manja je od aritmetičke i geometrijske sredine.

$$H < G \leq \mu$$

5.3. Geometrijska sredina

Geometrijska sredina izračunava se kao *N*-ti korijen *N*-umnožaka vrijednosti numeričke varijable $G = \sqrt[N]{x_1 x_2 .... x_i .... x_N}$.

Za grupirane podatke geometrijska sredina dana je izrazom:

$$G = \sqrt[N]{x_1^{f^1} x_2^{f^2} .... x_i^{f^i} .... x_k^{f^k}}$$

Geometrijska i harmonijska sredina primjenjuju se relativno rijetko. Može se izračunati i pomoću logaritamskog izraza, pri čemu je logaritam geometrijske sredine numeričke varijable jednak aritmetičkoj sredini logaritama vrijednosti varijable:

za negrupirane

$$\log G = \frac{1}{N} \sum_{i=1}^{N} \log x_i$$

za grupirane podatke

$$\log G = \frac{1}{N} \sum_{i=1}^{k} f_i \log x_i \quad N = \sum_{i=1}^{k} f_i$$

Geometrijska sredina jednaka je aritmetičkoj sredini ukoliko su svi članovi numeričke varijable međusobno jednaki. U suprotnom je manja od aritmetičke sredine.

Primjer 5.8. Izračun geometrijske sredine

$$G = \sqrt[5]{2*12*3*4*17} = \sqrt[5]{4896} = 5{,}47$$

5,47<7,6

Ukoliko se izračunaju aritmetička sredina, harmonijska sredina i geometrijska sredina na temelju istih vrijednosti numeričke varijable. Geometrijska sredina primjerena je za izračunavanje distribucije frekvencija s asimetričnim rasporedom podataka.
Geometrijska sredina primjenjuje se u analizi vremenskih nizova (potpoglavlje 13.2.2.3.).

5.4. Mod

Mod je vrijednost varijable koja se najčešće javlja u statističkom nizu. To je vrijednost varijable oko koje se elementi statističkog skupa najgušće gomilaju. Budući da je mod vrijednost varijable s najvećom frekvencijom, naziva se i frekvencijska srednja vrijednost.

Mod se **određuje** kod negrupiranih nizova, grupiranih nizova koji prikazuju distribuciju prema kvalitativnom obilježju i grupiranih nizova prema diskontinuiranom obilježju kod kojih je veličina razreda 1. U svim drugim distribucijama frekvencija moda se izračunava (aproksimira).

Primjer 5.9. Određivanje moda za negrupirani statistički niz:
Poslovni subjekti prema broju službenih automobila:
2, 3, 1, 1, **2**, 3, **2**, **2**, 0, **2.**

Analizirani poslovni subjekti najčešće imaju dva automobila.

Primjer 5.10. Određivanje moda za podatke grupirane prema kvalitativnom obilježju

Tablica 5.7. Strani turisti u Republici Hrvatskoj prema vrsti smještaja 2014. godine

Vrsta smještaja	Broj stranih turista
Hoteli	**4.911.252**
Odmarališta	4.155.018
Kampovi	2.372.536
Ostali smještaj	184.155
Ukupno	11.622.961

Strani turisti pristigli u Hrvatsku 2014. godine, najčešće su birali hotelski smještaj. Mod je hotelski smještaj.

Primjer 5.11. Određivanje moda za grupirani statistički nizu kojem su podaci prikazani prema diskontinuiranom obilježju kod kojih je veličina razreda 1.

Tablica 5.8. Mikro trgovačka društva prema broju zaposlenih

Broj zaposlenih	Broj mikro trgovačkih društava
x_i	f_i
0	28
Mo = 1	**54**
2	38
3	31
4	28
5	25
9	27
7	12
8	17
9	10
Ukupno	270

Iz podataka je vidljivo da je mod 1, odnosno mikro trgovačka društva imaju najčešće jednu zaposlenu osobu.

Primjer 5.12. Izračunavanje moda za grupirani statistički niz(podaci grupirani u tablicu distribucije frekvencija),u kojem su podaci prikazani prema diskontinuiranom obilježju (veličina razreda veća od 1) ili u kojem su podaci prikazani prema kontinuiranom obilježju.

Tablica 5.9. Zaposlenici u poslovnom subjektu Alfa prema dobi

Dob zaposlenika	Broj zaposlenih	Prave granice		Veličina razreda	Korigirane frekvencije	
x_i	f_i	L_{1i}	L_{2i}	i_i	fc_i	
18 – 20	2	18	21	3	0,67	
21 – 24	6	21	25	4	1,50	
25 – 29	9	25	30	5	1,80	
30 – 34	5	30	35	5	1,00	
35 – 39	4	35	40	5	0,80	
40 – 44	9	40	45	5	1,80	
45 – 49	5	45	50	5	1,00	a
50 – 54	11	**50**	55	**5**	**2,20**	b
55 – 59	4	55	60	5	0,80	c
60 – 64	2	60	65	5	0,40	
∑	57					

Ukoliko su podaci grupirani prema numeričkom diskontinuiranom obilježju veličine razreda veće od 1 ili prema numeričkom kontinuiranom obilježju (tablica 5.9.), mod se ne određuje, nego se izračunava (aproksimira) pomoću sljedećeg izraza:

$$Mo = L_1 + \frac{(b-a)}{(b-a)+(b-c)} \cdot i$$

$$Mo = 50 + \frac{(2{,}2-1)}{(2{,}2-1)+(2{,}2-0{,}8)} \cdot 5 = 1{,}67 = 50 + \frac{1{,}2}{1{,}2+1{,}4} \cdot 5 = 52{,}31$$

Najčešća dob zaposlenika u poslovnom subjektu Alfa je 52,31 godine.

Modalni razred je razred s najvećom frekvencijom ili korigiranom frekvencijom (ukoliko su razredi različitih veličina).

b –apsolutna frekvencija (ili relativna frekvencija) ili korigirana frekvencija(ili korigirana relativna frekvencija) modalnog razreda a – frekvencija razreda ispred modalnog razreda

c – frekvencija razreda iza modalnog razreda

L_1 – donja granica modalnoga razreda

i –veličina modalnog razreda

Kada su razredi različitih veličina, potrebno je korigirati frekvencije. Radi boljeg grafičkog prikaza moda, ukoliko veličine razreda imaju zajednički nazivnik, potrebno ih je skratiti, a potom njima dijeliti frekvencije.

Mod dijeli distribuciju frekvencija na lijevu (rastuću-uzlaznu) i desnu (opadajuću-silaznu) stranu. Grafički se mod može odrediti kada se na krivulji distribucije frekvencija (poligon frekvencija) pronađe najveća frekvencija na ordinati (tjeme), iz kojeg se spušta okomica na apscisu, gdje se potom pročita vrijednost moda. Grafički se mod može precizno odrediti i pomoću histograma. Najviši pravokutnik histograma upućuje na najveću frekvenciju te se dijagonalno spoje vrhovi najvišeg stupca s visinom njima suprotnih stupaca. Iz sjecišta dijagonala povuče se okomica na apscisu te se s apscise očita vrijednost moda.[13]

Nedostaci moda su:

- ovisnost o načinu formiranja razreda
- njegova vrijednost nema smisla ako se distribucija približava pravokutnoj
- sporan je kod bimodalne ili multimodalne distribucije

Prednost moda je što kod distribucija s ekstremno malim ili velikim vrijednostima numeričke varijable medijan i aritmetička sredina imaju težnju njihovom približavanju, pri čemu će primicanje medijana biti značajno manje od primicanja aritmetičke sredine. Mod neće imati tu tendenciju jer ga određuje najveća frekvencija.

[13] Grafičko određivanje moda vidi: Horvat i Mijoč (2014, str. 138)

5.5. Medijan

Medijan je srednja položajna vrijednost koja numerički niz uređen po veličini dijeli na dva jednakobrojna dijela. 50% vrijednosti članova niza je manje od medijana, a 50% vrijednosti članova niza je veće od medijana.
Kod određivanja medijana potrebno je jedinice promatranog niza redoslijedno zapisati po vrijednosti od najmanje do najveće.

Određivanje medijana za negrupirani statistički niz razlikuje se po tomu ima li niz parni ili neparni broj članova.
U nizu s **neparnim brojem članova** (N/2≠*INT*)medijan je ona vrijednost statističkog niza koji se nalazi u središtu, $Me = x_r$,odnosno *r* označava redni broj pod kojim se nalazi vrijednost statističkog niza *x*, $r = INT\left(\frac{N}{2}\right) + 1$.

Primjer 5.13. Određivanje medijana u nizu s neparnim brojem članova
Poslovni subjekti prema broju službenih automobila:
2, 3, 1, 1, 2, 3, 2, 2, 0, 2, 2.
Redoslijedno zapisan niz: 0, 1, 1, 2, 2, **2**, 2, 2,2, 3, 3.
N/2 = 11/2 = 5,5 ≠ *INT*.

$$r = INT\left(\frac{11}{2}\right) + 1 = INT(5{,}5) + 1 = 6\ ,\quad Me = x_6 = 2$$

Medijan je šesti član niza i iznosi 2. Budući da je analiziran niz koji nema veliki broj članova, iz njega je vidljivo, a to pokazuje i rezultat, da polovica poslovnih subjekata ima 2 službena automobila i manje, a polovica ima 2 službena automobila i više.

Kod niza s **parnim brojem članova** (N/2 = *INT*),medijan predstavlja polovica središnjih dvaju članova niza, prema izrazu: $Me = \frac{x_r + x_{r+1}}{2}\ ,\ r = \frac{N}{2}$.

Primjer 5.14. Određivanje medijana u nizu s parnim brojem članova
Poslovni subjekti prema broju službenih automobila:
2, 3, 1, 1, 2, 3, 2, 2, 0, 2.
Redoslijedno zapisan niz: 0, 1, 1, 2, **2**, **2**, 2,2, 3, 3.

N/2 = 10/2 = 5 =*INT*.

$$Me = \frac{x_{\frac{N}{2}} + x_{\frac{N}{2}+1}}{2} = \frac{x_5 + x_6}{2} = \frac{2+2}{2} = 2$$

I u ovom primjeru polovica poslovnih subjekata ima 2 službena automobila i manje, a polovica ima 2 službena automobila i više.

Određivanje medijana za grupirani statistički nizu kojem su podaci grupirani u tablicu distribucije frekvencija prema diskontinuiranom obilježju kojih je veličina razreda 1, pojednostavljuje se korištenjem kumulativnog niza „manje od“. Medijan je jednak vrijednosti one varijable čija kumulativna frekvencija prva uključuje N/2.

Primjer 5.15. Određivanje medijana za grupirani statistički niz

Tablica 5.10. Mikro trgovačka društva prema broju zaposlenih

Broj zaposlenih	Broj mikro trgovačkih društava	Kumulativni niz "manje od"
x_i	f_i	
0	28	28
1	54	82
2	38	120
Me = 3	31	**151**
4	28	179
5	25	204
6	27	231
7	12	243
8	17	260
9	10	270
Ukupno	270	

N/2 = 270/2=135. Medijalni razred je onaj u kojem je kumulativna frekvencija prvi puta veća od N/2. U kumulativu 151 nalazi se i 135. član, 135. mikro trgovačko društvo.

Budući da je medijan broj zaposlenih osoba onog mikro trgovačkog društva koje se nalazi na 135. mjestu, medijan je 3 zaposlene osobe, što znači da 50% mikro trgovačkih društava ima tri zaposlena i manje, a 50% ima tri zaposlena i više.

Na isti način **određuje se medijan za grupirani ordinalni statistički niz**.

Primjer 5.16. Određivanje medijana za ordinalni statistički niz

Tablica 5.11. Zaposleni u poslovnom subjektu Alfa prema završenom obrazovanju

Završeno obrazovanje	Broj zaposlenih	Kumulativni niz "manje od"
x_i	f_i	
Bez završenog obrazovanja	1	1
Osnovna škola	7	8
Srednja škola	21	**29**
Prvostupnik	8	37
Diplomski studij	15	52
Specijalistički studij	3	55
Doktor znanosti	2	57
Ukupno	57	

N/2 = 57/2=28,5. **Medijalni razred** je onaj u kojem je kumulativna frekvencija prvi puta veća od N/2. U kumulativu 29 nalazi se zaposlenik koji ima srednjoškolsko obrazovanje.

Medijan je srednja škola, što znači da 50% zaposlenih ima srednjoškolsko obrazovanje i manje, a 50% ima obrazovanje veće od srednjoškolskog (30. zaposlenik je prvostupnik).

Izračunavanje medijana kod grupiranog niza u kojem su podaci prikazani prema diskontinuiranom obilježju (veličina razreda veća od 1) i u kojem su podaci prikazani prema kontinuiranom obilježju

Ukoliko su podaci grupirani u tablicu distribucija frekvencija prema numeričkom diskontinuiranom obilježju, s veličinom razreda većom od 1, ili prema numeričkom kontinuiranom obilježju, medijan se ne određuje nego se izračunava (aproksimira) pomoću sljedećeg izraza:

$$Me = L_1 + \frac{\frac{N}{2} - \sum f_i}{f_{med}} \cdot i$$

N/2 – središnja jedinica niza

L_1 – donja prava granica medijalnog razreda

N – zbroj apsolutnih ili relativnih frekvencija

$\sum f_i$ – frekvencija kumulativnog niza „manje od“ ispred medijalnog razreda (zbroj svih frekvencija do medijalnog razreda)

f_{med} – apsolutna ili relativna frekvencija medijalnog razreda

i – veličina medijalnog razreda

Primjer 5.17. Izračunavanje medijana za grupirani statistički niz

Tablica 5.12. Zaposleni u poslovnom subjektu Alfa prema broju djece

Broj djece	Broj zaposlenih	Precizne granice	Veličina razreda	Kumulativni niz "manje od"
x_i	f_i	x_i	i_i	
0–1	14	0-1,5	1,5	**14**
2–4	**29**	**1,5**-4,5	**3**	**43**
5–7	12	4,5-7,5	3	55
8–9	2	7,5-9,5	2	57
Ukupno	57			

N/2 = 28,5

$$Me = 1{,}5 + \frac{(28{,}5 - 14)}{29} \cdot 3 = 3$$

U poslovnom subjektu50% zaposlenih imaju troje djece i manje, a ostalih 50% imaju troje djece i više.

Primjer 5.18. Izračunavanje medijana za grupirani statistički niz

Tablica 5.13. Zaposleni u poslovnom subjektu Alfa prema dobi

Dob zaposlenika	Broj zaposlenih	Veličina razreda	Kumulativni niz "manje od"
x_i	f_i	i_i	
18 – 21	2	3	2
21 – 25	6	4	8
25 – 30	9	5	17
30 – 35	5	5	22
35 – 40	4	5	**26**
40 – 45	**9**	**5**	**35**
45 – 50	5	5	40
50 – 55	11	5	51
55 – 60	4	5	55
60 – 65	2	5	57
Ukupno	57		

N/2 = 28,5

$$Me = 40 + \frac{(28{,}5 - 26)}{9} \cdot 5 = 41{,}39$$

U poslovnom subjektu Alfa50% zaposlenih imaju 41,39 godina i manje, a ostalih 50% imaju 41,39 godina i više.

Medijan se grafički može odrediti pomoću ucrtanih kumulativnih nizova „manje od“ i „više od“ u koordinatni sustav tako da se dobiju kumulante. Na mjestu gdje se kumulante sijeku, povlači se okomica na apscisu. Na mjestu na kojem okomica sječe apscisu, može se očitati vrijednost medijana.[14]

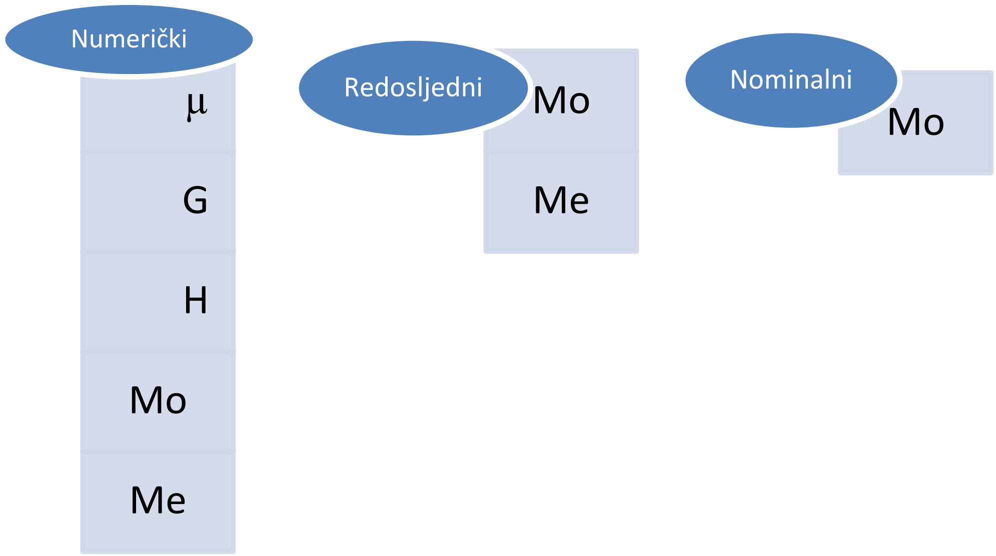
Prednost medijana u odnosu na aritmetičku sredinu je u tomu što je manje osjetljiv na izdvojenice, jer u njegovu izračunavanju ne sudjeluju svi članovi niza, već vrijednost središnje jedinice. Time je reprezentativnost medijana u izrazito asimetričnim distribucijama znatno veća nego li je reprezentativnost aritmetičke sredine.

5.6. Izbor srednjih vrijednosti

Izbor srednjih vrijednosti ovisi o vrsti statističkog niza.

[14] Grafičko određivanje medijana vidi: Horvat i Mijoč (2014, str. 130)

Slika 5.2. Mogućnosti izbora srednjih vrijednosti s obzirom na vrstu statističkog niza



Osim što izbor srednjih vrijednosti ovisi o vrsti niza (slika 5.2.) postoje i druge specifičnosti koje uvjetuju upotrebu neke od srednjih vrijednosti. Tako, ukoliko se u statističkom nizu nalaze vrijednosti manje od nule ili jednake nuli za njih se ne može izračunati geometrijska i harmonijska sredina.

5.7. Kvantili

Kvantili su vrijednosti numeričkog obilježja koje niz uređen po veličini dijele na q jednakobrojnih dijelova i izračunava se prema izrazu $\frac{iN}{q}$, gdje je:

i – broj kvantila

N – broj članova niza

q – red kvantila

Broj kvantila p zajedan je manji od reda q:

1. Medijan je kvantil prvog reda. Budući da je moguće izračunati samo jedan medijan (p= 1), medijan dijeli niz na dva jednaka dijela (q = 2).

2. Kvartili su kvantili trećeg reda. Budući da je moguće izračunati tri kvartila, među kojima je i medijan, (p = 3), kvartili dijele niz na četiri jednaka dijela (q = 4). Svaki pojedini kvartil dijeli niz na dva dijela.

   Q_1 – prvi ili donji kvartil

   Me – drugi kvartil ili medijan

   Q_3 – treći ili gornji kvartil
3. Decili su kvartili devetog reda. Budući da je moguće izračunati devet decila (p = 9), decili dijele niz na deset jednakih dijelova (q = 10).
4. Percentili su kvantili devedeset devetog reda. Moguće je izračunati 99 percentila (p = 99), te oni dijele niz na 100 jednakih dijelova (q = 100).

Medijan, osim što je kvantil prvog reda, ujedno je i drugi kvartil, peti decil i pedeseti percentil.

5.7.1. Kvartili

Određivanje kvartila za negrupirani statistički niz razlikuje se ovisno o tome je li niz djeljiv ili ne sa 4.

Kod niza s **brojem članova koji nije djeljiv s 4** ($N/4 \neq INT$) ili ($3N/4 \neq INT$),

donji kvartil se računa pomoću izraza $Q_1 = x_r$, pri čemu je, $r = INT\left(\frac{N}{4}\right)+1$, a

gornji kvartil se računa pomoću izraza $Q_3 = x_r$, pri čemu je, $r = INT\left(\frac{3N}{4}\right)+1$.

Primjer 5.19. Određivanje kvartila u nizu s brojem članova nedjeljivih s 4

Tablica 5.14. Zadnja cijena dionica KUNA-R-A u zadnjih 15 dana.

		Broj članova nije djeljiv s 4			Broj članova je djeljiv s 4	
Dan	Cijena dionice	N/4 ≠ INT; 3N/4 ≠ INT			N/4 = INT; 3N/4 ≠ INT	
		Redoslijedno zapisane cijene dionica			Redoslijedno zapisane cijene dionica (prvih 12 dana)	
1.	25,00	23,00			23,00	
2.	24,50	23,00			23,00	
3.	23,90	23,70			**23,70**	Q_1 =23,80
4.	23,00	**23,90**	Q_1		**23,90**	
5.	24,20	23,90			23,90	
6.	24,10	23,90			23,90	
7.	23,90	24,00			24,00	
8.	23,00	**24,00**	Q_2 = Me		24,00	
9.	23,90	24,00			**24,00**	Q_3 = 24,05
10.	24,00	24,10			**24,10**	
11.	24,10	24,10			24,10	
12.	23,70	**24,20**	Q_3		24,20	
13.	24,00	24,20				
14.	24,20	24,50				
15.	24,00	25,00				

N/4 = 15/4 = 3,75≠ *INT* *3*N/4 = 45/4 = 11,25 ≠ *INT*

$$r = INT\left(\frac{15}{4}\right)+1 = INT(3{,}75)+1 = 4\,,\quad r = INT\left(\frac{3\cdot 15}{4}\right)+1 = INT(11{,}25)+1 = 12$$

$$Q_1 = x_4 = 23{,}90 \qquad\qquad Q_3 = x_{12} = 24{,}20$$

U proteklih 15 dana dionica KUNA-R-A je četvrtinu (25%) dana završila trgovanje s cijenom 23,90 kn i manje, a tri četvrtine (75%) s cijenom 23,90 kn i više. U istom razdoblju,cijena dionice je u 75% dana završila trgovanje s cijenom 24,20 kn i manje, a u 25% dana s cijenom 24,20 kn i više.

Kod niza s **brojem članova koji je djeljiv s 4** (N/4 = *INT*) ili (3N/4 = *INT*) donji kvartil se računa pomoću izraza $Q_1 = \frac{x_r + x_{r+1}}{2}$, pri čemu je, $r = \frac{N}{4}$, a gornji kvartil se računa pomoću izraza $Q_3 = \frac{x_r + x_{r+1}}{2}$, pri čemu je, $r = \frac{3N}{4}$.

Primjer 5.20. Određivanje kvartila u nizu s brojem članova djeljivih s 4

Ukoliko se iz prethodnog primjera uzme trgovanje dionicom KUNA-R-A za prvih 12 dana, donji i gornji kvartil računaju se prema:

$$r = \frac{12}{4} = 3 = INT, \qquad r = \frac{3N}{4} = \frac{3 \cdot 12}{4} = 9 = INT$$

$$Q_1 = \frac{x_3 + x_4}{2} = \frac{23{,}70 + 23{,}90}{2} = 23{,}80 \qquad Q_3 = \frac{x_9 + x_{10}}{2} = \frac{24{,}00 + 24{,}10}{2} = 24{,}05$$

U 12 dana trgovanja dionicom KUNA-R-A,vrijednost dionice je u 25% slučajeva završila s cijenom 23,80 kuna i manje, a u 75% slučajeva s cijenom 23,80 kuna i više. U tom razdoblju je zadnja cijena u 75% dana završila s cijenom 24,05 kn i manje, a u 25% s cijenom od 24,05 kn i više.

Određivanje kvartila za grupirani statistički niz u kojem su podaci grupirani u tablicu distribucije frekvencija prema diskontinuiranom obilježju, kod kojih je veličina razreda 1, pojednostavljuje se korištenjem kumulativnog niza „manje od". Prvi kvartil (Q_1) je jednak vrijednosti one varijable čija kumulativna frekvencija prva uključuje N/4. Treći kvartil (Q_3) je jednak vrijednosti one varijable čija kumulativna frekvencija prva uključuje 3N/4.

Primjer 5.21. Određivanje kvartila grupiranog statističkog niza

Tablica 5.15. Mikro trgovačka društva prema broju zaposlenih

Broj zaposlenih	Broj mikro trgovačkih društava	Kumulativni niz "manje od"
x_i	f_i	
0	28	28
Q_1 = 1	54	**82**
2	38	120
3	31	151
4	28	179
Q_3 = 5	25	**204**
6	27	231
7	12	243
8	17	260
9	10	270
Ukupno	270	

N/4 = 270/4=67,5

Razred prvog kvartila je onaj u kojem je kumulativna frekvencija prvi puta veća od N/4. U kumulativu 82 nalazi se i 68. član, odnosno 68. mikro trgovačko društvo koje ima jednu zaposlenu osobu.

U 25% mikro trgovačkih društava broj zaposlenih je jedan i manje, a u 75% zaposlena je jedna osoba i više.

3N/4 = 3*270/4=202,5

Razred trećeg kvartila je onaj u kojem je kumulativna frekvencija prvi puta veća od 3N/4. U kumulativu 204 nalazi se i 203. mikro trgovačko društvo koje ima pet zaposlenih osoba.

U 75% mikro trgovačkih društava zaposleno je pet osoba i manje, a u 25% zaposleno je pet osoba i više.

Izračunavanje kvartila kod grupiranog niza u kojem su podaci prikazani prema diskontinuiranom obilježju (veličina razreda veća od 1) i u kojem su podaci prikazani prema kontinuiranom obilježju

Ukoliko su podaci grupirani u tablicu distribucija frekvencija prema numeričkom diskontinuiranom obilježju, s veličinom razreda većom od 1 ili prema numeričkom kontinuiranom obilježju, kvartili se ne određuju, nego se izračunavaju (aproksimiraju) pomoću sljedećih izraza:

$$Q_1 = L_1 + \frac{\frac{N}{4} - \sum f_1}{f_{Q_1}} \cdot i$$

N/4 – četvrtina elemenata niza

L_1 – donja granica razreda prvog kvartila

N – zbroj apsolutnih ili relativnih frekvencija

$\sum f_i$ – frekvencija kumulativnog niza „manje od“ ispred razreda prvog kvartila

f_{Q1}– apsolutna ili relativna frekvencija razreda prvog kvartila

i – veličina razreda prvog kvartila

$$Q_3 = L_1 + \frac{3\frac{N}{4} - \sum f_1}{f_{Q_3}} \cdot i$$

3N/4 – tri četvrtine elemenata niza

L_1 – donja granica razreda trećeg kvartila

N – zbroj apsolutnih ili relativnih frekvencija

$\sum f_i$ – frekvencija kumulativnog niza „manje od“ ispred razreda trećeg kvartila

f_{Q3}– apsolutna ili relativna frekvencija razreda trećeg kvartila

i – veličina razreda trećeg kvartila

Primjer 5.22. Izračunavanje kvartila

Tablica 5.16. Zaposleni u poslovnom subjektu Alfa prema broju djece

Broj djece	Broj zaposlenih	Precizne granice	Veličina razreda	Kumulativni niz "manje od"
x_i	f_i	x_i	i_i	
0–1	14	0-1,5	1,5	**14**
$Q_1,Q_3 \rightarrow$2–4	**29**	**1,5**-4,5	3	**43**
5–7	12	4,5-7,5	3	55
8–9	2	7,5-9,5	2	57
Ukupno	57			

N/4 = 14,25

$$Q_1 = 1{,}5 + \frac{(14{,}25 - 14)}{29} \cdot 3 = 1{,}53$$

25% zaposlenih u poslovnom subjektu Alfa imaju 1,53 djece i manje, a ostalih 75% imaju 1,53 djece i više.

3N/4 = 42,75

$$Q_3 = 1{,}5 + \frac{(42{,}75 - 14)}{29} \cdot 3 = 4{,}47$$

75% zaposlenih u poslovnom subjektu Alfa imaju 4,47 djece i manje, a ostalih 25% imaju 4,47 djece i više.

Tablica 5.17. Zaposleni u poslovnom subjektu Alfa prema dobi

Dob zaposlenika	Broj zaposlenih	Veličina razreda	Kumulativni niz "manje od"
x_i	f_i	i_i	
18 – 21	2	3	2
21 – 25	6	4	**8**
$Q_1 \rightarrow$ **25** – 30	**9**	**5**	**17**
30 – 35	5	5	22
35 – 40	4	5	26
40 – 45	9	5	35
45 – 50	5	5	**40**
$Q_3 \rightarrow$**50** – 55	**11**	**5**	**51**
55 – 60	4	5	55
60 – 65	2	5	57
Ukupno	57		

N/4 = 14,25

$$Q_1 = 25 + \frac{(14{,}25 - 8)}{9} \cdot 5 = 28{,}47$$

25% zaposlenih u poslovnom subjektu Alfa imaju 28,47 godina i manje, a ostalih 75% imaju 28,47 godina i više.

3N/4 = 42,75

$$Q_3 = 50 + \frac{(42{,}75 - 40)}{11} \cdot 5 = 51{,}25$$

75% zaposlenih u poslovnom subjektu Alfa imaju 51,25 godina i manje, a ostalih 25% imaju 51,25 godina i više.

5.7.2. Decili

Decili za negrupirane statističke nizove računaju se prema sljedećim izrazima:

Ako N nije djeljiv s 10

$$\frac{iN}{10} \neq INT \;\; D_i = x_r \;\; r = INT\left(\frac{iN}{10}\right) + 1 \;\; i = 1,2,3,....9$$

Ako je N djeljiv s 10

$$\frac{iN}{10} = INT \;\; D_i = \frac{x_r + x_{r+1}}{2} \;\; r = \frac{iN}{10} \;\; i = 1,2,3,....9$$

Decili za grupirane statističke nizove računaju se prema sljedećem izrazu:

$$D_i = L_1 + \frac{\frac{iN}{10} - \sum f_1}{f_{D_i}} \cdot i \quad i = 1,2,3,....9$$

$iN/10$ – desetina elemenata niza

L_1 – donja granica razreda *i*-tog decila

N – zbroj apsolutnih ili relativnih frekvencija

$\sum f_i$ – frekvencija kumulativnog niza „manje od“ ispred razreda *i*-tog decila

f_{Di}– apsolutna ili relativna frekvencija razreda *i*-tog decila

i – veličina razreda *i*-tog decila

5.7.3. Percentili

Percentili za negrupirane statističke nizove računaju se prema sljedećim izrazima:

Ako *N* nije djeljiv s 100

$$\frac{iN}{100} \neq INT \quad P_i = x_r \quad r = INT\left(\frac{iN}{100}\right) + 1 \quad i = 1,2,3,....99$$

Ako je *N* djeljiv s 100

$$\frac{iN}{100} = INT \quad P_i = \frac{x_r + x_{r+1}}{2} \quad r = \frac{iN}{100} \quad i = 1,2,3,....99$$

Percentili za grupirane statističke nizove računaju se prema sljedećem izrazu:

$$P_i = L_1 + \frac{\frac{iN}{100} - \sum f_1}{f_{P_i}} \cdot i \quad i = 1,2,3,....99$$

$iN/100$ – stotina elemenata niza

L_1 – donja granica razreda *i*-tog percentila

N – zbroj apsolutnih ili relativnih frekvencija

$\sum f_i$ – frekvencija kumulativnog niza „manje od“ ispred razreda *i*-tog percentila

f_{Pi} – apsolutna ili relativna frekvencija razreda *i*-tog percentila

i – veličina razreda *i*-tog percentila

Primjer 5.23. Izračunavanje decila i percentila

Tablica 5.18. Zaposleni u poslovnom subjektu Alfa prema dobi

	Dob zaposlenika	Broj zaposlenih	Veličina razreda	Kumulativni niz "manje od"
	x_i	f_i	i_i	
$P_2 \rightarrow$	**18** – 21	**2**	**3**	**2**
$P_{10} \rightarrow$	**21** – 25	**6**	**4**	**8**
$D_2 \rightarrow$	**25** – 30	**9**	**5**	17
	30 – 35	5	5	**22**
$P_{40}=D_4 \rightarrow$	**35** – 40	**4**	**5**	**26**
$Me=D_5 \rightarrow$	**40** – 45	**9**	**5**	35
	45 – 50	5	5	**40**
$P_{75}=Q_3 \rightarrow$	**50** – 55	**11**	**5**	**51**
$P_{90} \rightarrow$	**55** – 60	**4**	**5**	55
	60 – 65	2	5	57
	Ukupno	57		

Izračunavanje 2. decila:

iN/10 = 2·57/10 = 11,4

$$D_2 = 25 + \frac{(11{,}4 - 8)}{9} \cdot 5 = 26{,}89$$

20% zaposlenih u poslovnom subjektu Alfa imaju 26,89 godina i manje, a ostalih 80% imaju 26,89 godina i više.

Izračunavanje 5. decila:

Budući da se 5. decil nalazi na polovini statističkog niza njegova vrijednost mora biti jednaka vrijednosti medijana. Dokaz:

iN/10 = 5·57/10 = 28,5

$$D_5 = 40 + \frac{(28{,}5 - 26)}{9} \cdot 5 = 41{,}39$$

Tumačenje rezultata je isto kao kod tumačenja medijana: 50% zaposlenih u poslovnom subjektu Alfa imaju 41,39 godina i manje, a ostalih 50% imaju 41,39 godina i više.

Izračunavanje 10. percentila

$iN/100 = 10 \cdot 57/100 = 5{,}7$

$$P_{10} = 21 + \frac{(5{,}7 - 2)}{6} \cdot 4 = 23{,}46$$

10% zaposlenih u poslovnom subjektu Alfa imaju 23,46 godina i manje, a ostalih 90% imaju 23,46 godina i više.

Izračunavanje 90. percentila

$iN/100 = 90 \cdot 57/100 = 51{,}3$

$$P_{90} = 55 + \frac{(51{,}3 - 50)}{4} \cdot 5 = 56{,}63$$

90% zaposlenih u poslovnom subjektu Alfa imaju 56,63 godina i manje, a ostalih 10% imaju 56,63 godina i više.

Izračunavanje 40. percentila

Budući da se 40. percentil nalazi na četiri desetine statističkog niza njegova vrijednost mora biti jednaka vrijednosti 4. decila. Dokaz $P_{40} = D_4$.

$iN/100 = 40 \cdot 57/100 = 22{,}8$

$$P_{40} = 35 + \frac{(22{,}8 - 22)}{4} \cdot 5 = 36$$

Izračunavanje 75. percentila

Budući da se 75. percentil nalazi na tri četvrtine statističkog niza njegova vrijednost mora biti jednaka vrijednosti gornjeg kvartila. Dokaz $P_{75} = Q_3$.

$iN/100 = 75 \cdot 57/100 = 42{,}75$

$$P_{75} = 50 + \frac{(42{,}75 - 40)}{11} \cdot 5 = 51{,}25$$

75% zaposlenih u poslovnom subjektu Alfa imaju 51,25 godina i manje, a ostalih 25% imaju 51,25 godina i više.

Izračunavanje 2. percentila

$iN/100$ = 2·57/100 = 1,14

$$P_2 = 18 + \frac{(1{,}14 - 0)}{2} \cdot 3 = 19{,}71$$

2% zaposlenih u poslovnom subjektu Alfa imaju 19,71 godina i manje, a ostalih 98% imaju 19,71 godina i više.

5.8. Momenti

Momenti numeričkih nizova su specifične srednje vrijednosti koje predstavljaju aritmetičke sredine odstupanja vrijednosti numeričke varijable od njezine aritmetičke sredine. Dijele se na momente oko sredine (aritmetičke) i na pomoćne.

Momenti oko sredine računaju se pomoću izraza:

Za negrupirane nizove

$$\mu_r = \frac{\sum_{i=1}^{N} (x_i - \mu)^r}{N}$$

Za grupirane nizove

$$\mu_r = \frac{\sum_{i=1}^{k} f_i (x_i - \mu)^r}{N}$$

Pomoćni momenti su momenti oko nule, momenti oko a i momenti oko a uz b.

Momenti oko nule računaju se pomoću izraza:

Za negrupirane nizove

$$m_r = \frac{\sum_{i=1}^{N} (x_i - 0)^r}{N} = \frac{\sum_{i=1}^{N} x_i^r}{N}$$

Za grupirane nizove

$$m_r = \frac{\sum_{i=1}^{k} f_i x_i^r}{N}$$

Momenti oko *a* računaju se pomoću izraza:

Za negrupirane nizove

$$m_r^{'} = \frac{\sum_{i=1}^{N}(x_i - a)^r}{N} = \frac{\sum_{i=1}^{N} d_i^r}{N}$$

Za grupirane nizove

$$m_r^{'} = \frac{\sum_{i=1}^{k} f_i d_i^r}{N}$$

Momenti oko *a* uz *b*, računaju se pomoću izraza:

Za negrupirane nizove

$$m_r^{''} = \frac{\sum_{i=1}^{N}\left(\frac{x_i - a}{b}\right)^r}{N} = \frac{\sum_{i=1}^{N} d_i^{'r}}{N}$$

Za grupirane nizove

$$m_r^{''} = \frac{\sum_{i=1}^{k} f_i d_i^{'r}}{N}$$

Nulti moment oko sredine je konstanta i iznosi 1. $\mu_0 = \frac{\sum_{i=1}^{N}(x_i - \mu)^0}{N} = \frac{N}{N} = 1$

Prvi moment oko sredine je konstanta i iznosi 0. $\mu_1 = \frac{\sum_{i=1}^{N}(x_i - \mu)^1}{N} = \frac{0}{N} = 0$

Drugi moment oko sredine je varijanca. $\mu_2 = \frac{\sum_{i=1}^{N}(x_i - \mu)^2}{N} = \sigma^2$

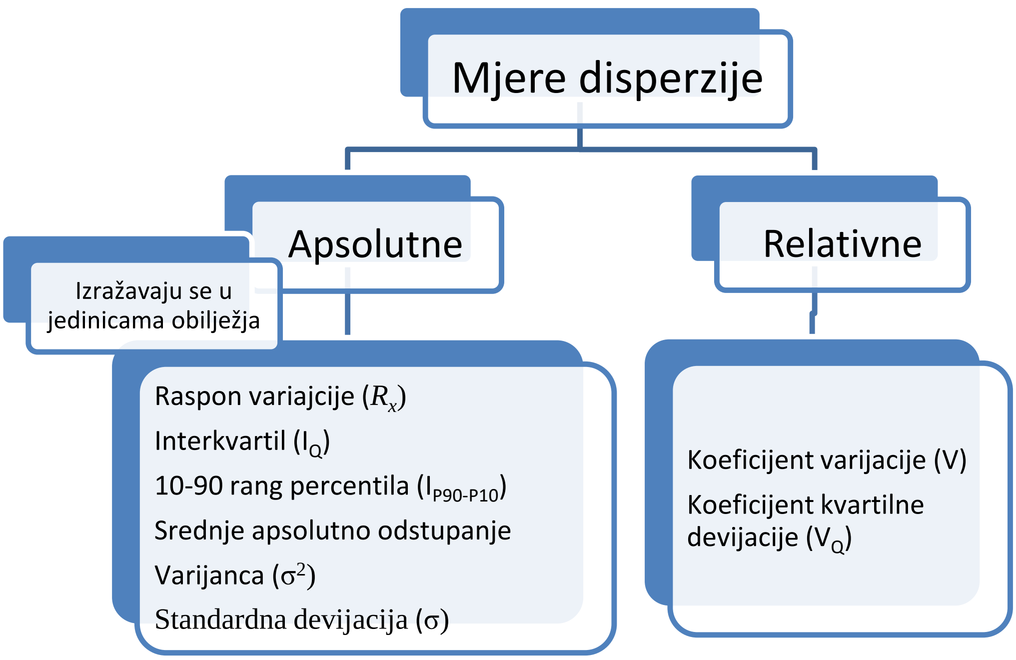
Pomoćnim momentima izračunavaju se momenti oko sredine. Primjer izračuna momenata oko sredine (drugi, treći i četvrti), prikazan je u poglavljima gdje je opisan način izračuna varijance (potpoglavlje 6.6.), mjere asimetrije α_3 (poglavlje 8.) i mjere zaobljenosti α_4 (poglavlje 9).

6. MJERE DISPERZIJE (RASPRŠENOSTI)

Individualne vrijednosti numeričkog obilježja različite su u većoj ili manjoj mjeri od mjera središnje tendencije (srednjih vrijednosti),koja je izračunata iz tih individualnih vrijednosti. Dva ili više promatranih nizova mogu imati jednaku aritmetičku sredinu, ali različit raspored jedinica niza, što je rezultat različite raspršenosti individualne vrijednosti obilježja i dovodi do različitih vrijednosti pokazatelja raspršenosti (tablica 6.1.).

Uz pojmove raspršenost i disperzija, za opisivanje odstupanja individualnih vrijednosti obilježja koristi se i pojam varijabilnost te se može reći da reprezentativnost srednje vrijednosti ovisi o stupnju varijabilnosti podataka. Mjere disperzije dijele se na apsolutne i relativne. Apsolutne se izražavaju u jedinicama obilježja, a relativne u proporciji ili postotku. Apsolutnim mjerama disperzije nije moguće usporediti raspršenost raznorodnih distribucija koje se mogu usporediti samo relativnim mjerama disperzije.

Slika 6.1. Podjela mjera disperzije



6.1. Raspon varijacije

Raspon varijacije (*R_x*) predstavlja razliku između najveće i najmanje vrijednosti. $R_x = x_{\max} - x_{\min}$

Ukoliko se raspon varijacije izračunava na temelju podataka prikazanih distribucijom frekvencija s razredima, aproksimira se kao razlika između gornje granice posljednjeg i donje granice prvog razreda ili kao razlika razrednih sredina posljednjeg i prvog razreda.

Nedostatak ove mjere disperzije je što u izračunu obuhvaća samo rubne vrijednosti niza te predstavlja grubu mjeru disperzije i zbog toga je nepouzdana.

Primjer 6.1. Tri promatrana niza s jednakom aritmetičkom sredinom, ali različitim rasporedom jedinica niza

Tablica 6.1. Iznos mirovine tri skupine umirovljenika u kunama

Red. broj	Iznos mirovine umirovljenika u skupini A	Iznos mirovine umirovljenika u skupini B	Iznos mirovine umirovljenika u skupini C
	$x_{(A)}$	$x_{(B)}$	$x_{(C)}$
1.	650	750	700
2.	1.150	900	980
3.	1.200	1.150	1.100
4.	1.850	1.900	1.620
5.	1.850	1.850	1.850
6.	1.850	1.850	1.850
7.	1.850	1.850	1.850
8.	2.300	2.600	2.270
9.	3.500	3.700	3.980
10.	4.900	4.850	5.100
11.	8.200	8.300	8.250
12.	9.700	9.300	9.450
Ukupno	39.000	39.000	39.000
μ	3.250	3.250	3.250
Mod	1.850	1.850	1.850
Me	1.850	1.850	1.850
Q_1	1.525	1.525	1.360
Q_3	4.200	4.275	4.540

U primjeru podataka o visini mirovina umirovljenika skupine A, B i C, unatoč istim srednjim vrijednostima, rezultat različite raspršenosti individualne vrijednosti obilježja dovodi do različitih vrijednosti pokazatelja raspršenosti. Raspon varijacije je:

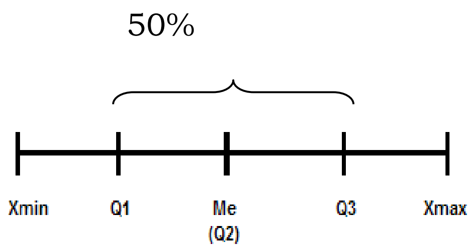
$R_{x(A)} = 9.700 - 650 = 9.050$ $R_{x(B)} = 9.300 - 750 = 8.550$ $R_{x(c)} = 9.450 - 700 = 8.750$

T: Raspon varijacije između najviše i najniže mirovine umirovljenika u skupini A je 9.050,00 kn. Raspon mirovina umirovljenika u skupini A, također je i najveći jer u skupini B iznosi 8.550,00 kn, a u skupini C iznosi 8.750,00 kn.

6.2. Interkvartil

Interkvartil je raspon varijacije središnjih 50% elemenata uređenog statističkog niza. Statistički niz mora biti uređen po redoslijedu vrijednosti od najmanje do najveće. Izražava se kao razlika gornjeg i donjeg kvartila, prema izrazu: $I_Q = Q_3 - Q_1$.

Slika 6.2. Interkvartil



Primjer 6.2. Izračun interkvartila

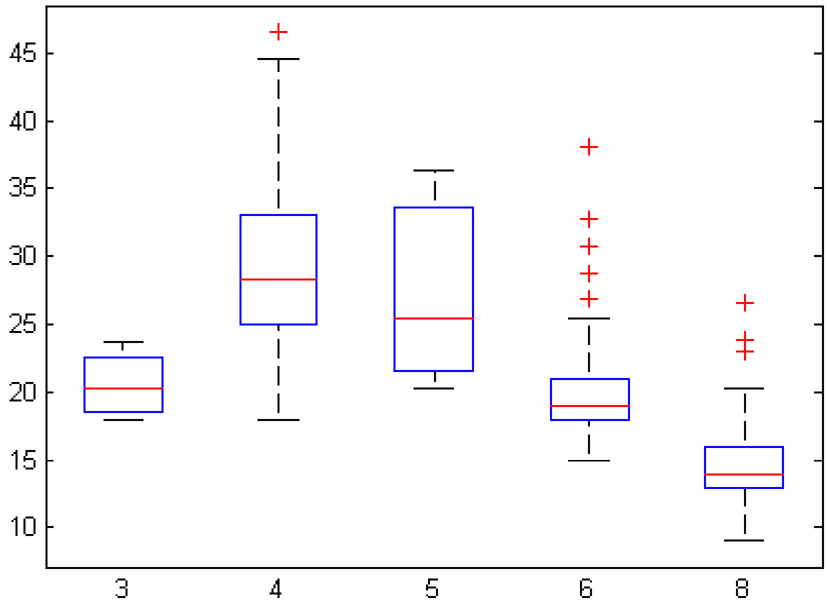
U primjeru podataka o visini mirovina umirovljenika A, B i C skupine (tablica 6.1.),raspon varijacije središnjih 50% elemenata statističkog niza je:

$I_{Q(A)} = 4.200 - 1.525 = 2.675$ $I_{Q(B)} = 4.275 - 1.525 = 2.750$ $I_{Q(C)} = 4.540 - 1.360 = 3.180$

T : Raspon varijacije središnjih 50% mirovina umirovljenika u skupini A je 3.675,00 kn. Raspon varijacije mirovina središnjih 50% umirovljenika u skupini C je najveći, 3.180,00 kn.

Odstupanja individualnih vrijednosti mogu se dobro prikazati grafičkim prikazom Box-plot dijagramom (B-P dijagram ili pravokutni dijagram). B-P dijagram za grafički prikaz koristi pet specifičnih vrijednosti niza: medijan, prvi i treći kvartil te najmanju i najveću vrijednost niza (engl. *five summary numbers* – 5's).

Grafikon 6.1. Pravokutni (B-P) dijagram



Na grafičkom prikazu vidljivi su raspon varijacije i interkvartilni raspon te se prosuđuju moguće asimetrije kao i pojave netipičnih vrijednosti (outliers).

Pomoću pravokutnog dijagrama lako se mogu uočiti izdvojenice. Razlikuju se umjerene i sumnjive izdvojenice (engl. *outliers*), koje se u grafikonu označavaju kružićem te ozbiljne (engl. *extreme outliers*) izdvojenice koje se označavaju zvjezdicom.[15]

Umjerene i sumnjive izdvojenice:

[15] Način izrade pravokutnog dijagrama vidi: Statistics How To: http://www.statisticshowto.com/how-to-find-a-five-number-summary-in-statistics/

- donja unutarnja međa Q_1-1,5 I_q
- gornja unutarnja međa Q_3+1,5 I_q

Ozbiljne izdvojenice:

- donja vanjska međa Q_1-3 I_q
- gornja vanjska međa Q_3+3 I_q

6.3. Koeficijent kvartilne devijacije

Koeficijent kvartilne devijacije je relativna mjera disperzije koja objašnjava disperziju središnjih 50% jedinica niza, a njegove vrijednosti mogu se očekivati u intervalu od 0 do 1. Njime se može uspoređivati stupanj disperzije raznorodnih nizova. Određuje se kao omjer interkvartila i zbroja kvartila:

$$V_Q = \frac{Q_3 - Q_1}{Q_3 + Q_1} \qquad 0 \le V_Q < 1$$

Primjer 6.3. Izračunavanje koeficijenta kvartilne devijacije (tablica 6.1.)

$$V_{Q(A)} = \frac{4.200 - 1.525}{4.200 + 1.525} = \frac{2.675}{5725} = 0,47$$

T: Raspon varijacije središnjih 50% mirovina umirovljenika u skupini A iznosi 0,47, što predstavlja veliko variranje središnjih 50% podataka.

6.4. 10-90 rang percentila

10-90 rang percentila je mjera koja predstavlja razliku između 90-og i 10-og percentila i govori o varijaciji središnjih 80% podataka. Ova mjera ne uzima u obzir 10% rubnih podataka na donjoj granici niza i 10% na gornjoj. Izračunava se prema izrazu:

$$I_{P90-P10} = P_{90} - P_{10}$$

Primjer 6.4. Izračunavanje 10-90 rang percentila (tablica 5.18.)

P_{10} = 23,46 godina, a P_{90}= 56,63 godina

$$I_{P90-P10} = 56{,}63 - 23{,}46 = 33{,}17$$

T: Raspon varijacije dobi središnjih 80% zaposlenika u poslovnom subjektu Alfa iznosi 33,17 godina.

Za mjerenje disperzije oko aritmetičke sredine koriste se apsolutne mjere: srednje apsolutno odstupanje, varijanca i standardna devijacija te relativna mjera koeficijent varijacije. To su potpune mjere disperzije jer ovise o vrijednosti svakog člana niza. Srednje apsolutno odstupanje računa se i kao mjera disperzije oko medijana.

6.5. Srednje apsolutno odstupanje

Srednje apsolutno odstupanje (MAD), od engl. *mean absolute deviation* je prosječno apsolutno odstupanje vrijednosti varijable od njezine aritmetičke sredine. Za izračunavanje disperzije koristi se i srednje apsolutno odstupanje oko medijana:

za negrupirane podatke

$$MAD = \frac{\sum_{i=1}^{N} |x_i - \mu|}{N}$$

za podatke prikazane distribucijom frekvencija

$$MAD = \frac{\sum_{i\_1}^{k} f_i |x_i - \mu|}{\sum_{i\_=1}^{k} f_i}$$

za negrupirane podatke

$$MAD_{Me} = \frac{\sum_{i=1}^{N} |x_i - Me|}{N}$$

za podatke prikazane distribucijom frekvencija

$$MAD_{Me} = \frac{\sum_{i\_1}^{k} f_i |x_i - Me|}{\sum_{i=1}^{k} f_i}$$

Za izračunavanje navedenih mjera disperzije korišteni su podaci o broju zaposlenih u poslovnom subjektu Alfa prema dobi.

Primjer 6.5. Izračunavanje mjera disperzije oko aritmetičke sredine

Tablica 6.2. Zaposleni u poslovnom subjektu Alfa prema dobi

Dob zaposlenika u godinama	Broj zaposlenih	Sredina razreda						
x_i	f_i	x_i	$f_i x_i$	$x_i - \mu$	$f_i \lvert x_i - \mu \rvert$	$f_i \lvert x_i - Me \rvert$	$f_i (x_i - \mu)$	$f(x_i - \mu)^2$
18 – 21	2	19,50	39,00	-20,67	41,33	43,78	-82,67	854,22
21 – 25	6	23,00	138,00	-17,17	103,00	110,34	-206,00	1.768,17
25 – 30	9	27,50	247,50	-12,67	114,00	125,01	-228,00	1.444,00
30 – 35	5	32,50	162,50	-7,67	38,33	44,45	-76,67	293,89
35 – 40	4	37,50	150,00	-2,67	10,67	15,56	-21,33	28,44
40 – 45	9	42,50	382,50	2,33	21,00	9,99	42,00	49,00
45 – 50	5	47,50	237,50	7,33	36,67	30,55	73,33	268,89
50 – 55	11	52,50	577,50	12,33	135,67	122,21	271,33	1.673,22
55 – 60	4	57,50	230,00	17,33	69,33	64,44	138,67	1.201,78
60 – 65	2	62,50	125,00	22,33	44,67	42,22	89,33	997,56
Ukupno	57		2.289,50		614,67	608,55	0,00	8.579,17

Primjer 6.6. Izračunavanje srednjeg apsolutnog odstupanja (MAD) oko aritmetičke sredine

$$\mu = 40{,}17 \quad MAD = \frac{614{,}67}{57} = 10{,}78$$

T: Prosječna dob zaposlenika u poslovnom subjektu Alfa je 40,17 godina. Prosječno apsolutno odstupanje od prosječne dobi zaposlenika u poslovnom subjektu Alfa iznosi 10,78 godina.

Primjer 6.7. Izračunavanje srednjeg apsolutnog odstupanja (MAD) oko medijana

$$Me = 41{,}39 \quad MAD_{med} = \frac{608{,}55}{57} = 10{,}68$$

T: Prosječno apsolutno odstupanje od medijana iznosi 10,68 godina.

6.6. Varijanca

Varijanca je aritmetička sredina kvadrata odstupanja vrijednosti numeričke varijable od njezine aritmetičke sredine. Označava se sa σ^2 (malo grčko slovo, čit. sigma na kvadrat). Izračunava se pomoću izraza:

za negrupirane podatke za podatke prikazane distribucijom frekvencija

$$\sigma^2 = \frac{\sum_{i=1}^{N}(x_i - \mu)^2}{N} \qquad \sigma^2 = \frac{\sum_{i=1}^{k} f(x_i - \mu)^2}{\sum_{i=1}^{k} f_i}$$

Primjer 6.8. Izračunavanje varijance

Prema podacima iz tablice 6.2. varijanca je:

$$\sigma^2 = \frac{8.579{,}17}{57} = 150{,}51$$

T: Dob zaposlenika u poslovnom subjektu Alfa prosječno kvadratno odstupa od prosječne dobi za 150,51 godinu.

Varijanca se može izračunati i upotrebom izvedenog oblika, tzv. skraćenog izraza:

$$\sigma^2 = \frac{\sum_{i=1}^{k} f_i x_i^2}{\sum_{i=1}^{k} f_i} - \frac{\sum_{i=1}^{k} f_i x_i}{\sum_{i=1}^{k} f_i} \quad \text{ili} \quad \sigma^2 = \frac{\sum_{i=1}^{k} f_i x_i^2}{\sum_{i=1}^{k} f_i} - \mu^2$$

Primjer 6.9. Izračunavanje mjera disperzije oko aritmetičke sredine skraćenim postupkom

Tablica 6.3. Zaposleni u poslovnom subjektu Alfa prema dobi

Dob zaposlenika u godinama	Broj zaposlenih	Sredina razreda		
x_i	f_i	x_i	$f_i x_i$	$f_i x_i^2$
18 – 21	2	19,50	39,00	760,50
21 – 25	6	23,00	138,00	3.174
25 – 30	9	27,50	247,50	6.806,25
30 – 35	5	32,50	162,50	5.281,25
35 – 40	4	37,50	150,00	5.625
40 – 45	9	42,50	382,50	16.256,25
45 – 50	5	47,50	237,50	11.281,25
50 – 55	11	52,50	577,50	30.318,75
55 – 60	4	57,50	230,00	13.225
60 – 65	2	62,50	125,00	7.812,50
Ukupno	57		2.289,50	100.540,75

$$\sigma^2 = \frac{100.540{,}75}{57} - 40{,}17^2 = 1.763{,}87 - 1.613{,}63 = 150{,}24$$ [16]

[16] Nastala razlika od 0,27 godina (150,51–150,24) posljedica je zaokruživanja.

6.7. Standardna devijacija

Budući da je (prema 1. svojstvu aritmetičke sredine), zbroj kvadrata odstupanja od aritmetičke sredina jednak nuli, daljnje računske operacije s nulom rezultirale bi nulom. Zbog toga su odstupanja kvadrirana, što otežava tumačenje. Stoga je iz varijance izračunat drugi korijen kako bi se dobio rezultat u prvom stupnju, koji predstavlja standardnu devijaciju.
Prema tome, **standardna devijacija** je pozitivni drugi korijen iz varijance i tumači se kao prosječno odstupanje vrijednosti numeričkog obilježja od njezine aritmetičke sredine. Predstavlja potpunu i apsolutnu mjeru disperzije.

$$\sigma = \sqrt{\sigma^2}$$

Primjer 6.10. Izračunavanje standardne devijacije

Prema podacima iz tablice 6.2. i 6.3. standardna devijacija je:

$$\sigma = \sqrt{150{,}51} = 12{,}27$$

T: Dob zaposlenika u poslovnom subjektu Alfa prosječno odstupa od prosječne dobi za 12,27 godina.
Nedostatak apsolutnih mjera disperzije jest u njihovoj isključivoj primjenjivosti za uspoređivanje disperzije istorodnih distribucija jer su izražene u jedinicama obilježja.

6.8. Koeficijent varijacije

Koeficijent varijacije je relativna mjera disperzije. Predstavlja omjer standardne devijacije i aritmetičke sredine pomnožen sa sto, prema izrazu:

$$V = \frac{\sigma}{\mu} \cdot 100$$

Varijabilnost, mjerena koeficijentom varijacije, veća od 30% smatra se velikom. Koeficijent varijacije veći od 100% pokazuje izraženu heterogenost podataka, to jest veliku disperziju. Tada je primjerenije upotrijebiti položajne srednje vrijednosti jer aritmetička sredina nije reprezentativan pokazatelj.

Primjer 6.11. Izračunavanje koeficijenta varijacije

Prema podacima iz tablice 6.2. i 6.3. koeficijent varijacije je:

$$V = \frac{12,27}{40,17} \cdot 100 = 30,55\%$$

T: Postotni omjer standardne devijacije i prosječne dobi zaposlenika u poslovnom subjektu Alfa iznosi 30,55%.

Relativne mjere disperzije (koeficijent varijacije i koeficijent kvartilne devijacije), daju podatak o tome **koja** distribucija ima veću disperziju u odnosu na neku drugu distribuciju koja ne mora biti mjerena istorodnim obilježjem. Međutim, ne daju podatak o tome **koliko** svaka pojedina vrijednost numeričkog obilježja odstupa od prosjeka, neovisno je li riječ o usporedbi istorodnih ili raznorodnih distribucija. Dakle, relativne mjere disperzije ne omogućuju usporedbu odstupanja originalnih vrijednosti obilježja od aritmetičke sredine u raznorodnim distribucijama frekvencija (npr. plaće u kunama i proizvodnja u komadima).

7. STANDARDIZIRANO OBILJEŽJE

Standardizirano obilježje u literaturi se nalazi i pod pojmovima standardizirana varijabla, Z-vrijednost, Z-test.
Ukoliko je potrebno usporediti dvije ili više distribucija izraženih različitim jedinicama obilježja (raznorodne distribucije), pogodno je odstupanja standardizirati. Odstupanja individualnih vrijednosti numeričkog obilježja od aritmetičke sredine u raznorodnim distribucijama frekvencija izračunavaju se standardiziranim (normaliziranim) obilježjem.

Standardizirano obilježje izračunava se tako da se odstupanja numeričke varijable od njezine aritmetičke sredine podijele sa standardnom devijacijom. Izračunata odstupanja izražena su u jedinicama standardnih devijacija.

$$z_i = \frac{x_i - \mu}{\sigma}$$

Aritmetička sredina standardiziranog obilježja jednaka je nuli, a standardna devijacija standardiziranog obilježja je jedan.

$$\bar{z} = 0 \qquad \sigma_z = 1$$

Primjer 7.1. Z-test
Iz primjera prikazanih u tablicama 5.1. i 5.2., moguće je izračunati odstupa li više od prosjeka zaposlenik u poslovnom subjektu Alfa koji ima 2 djece ili zaposlenik u dobi od 40 godina.

$$z_{djeca} = \frac{2 - 3{,}27}{2{,}03} = \frac{-1{,}27}{2{,}03} = -0{,}6256\sigma$$

$$z_{dob} = \frac{40 - 40{,}17}{12{,}27} = \frac{-0{,}17}{12{,}27} = -0{,}0139\sigma$$

T: Od prosjeka više odstupa zaposlenik koji ima dvoje djece, nego zaposlenik u dobi od 40 godina. Zaposlenik koji ima dvoje djece, udaljen je od prosjeka za -0,6256 standardnih devijacija, dok je zaposlenik u dobi od 40 godina udaljen od prosjeka za -0,0139 standardnih devijacija.

U navedenom primjeru standardizirane vrijednosti su izračunate iz iste populacije (zaposlenici poslovnog subjekta Alfa), a na isti način može se ispitati disperzija individualnih vrijednosti varijabli iz različitih populacija. Standardizirano obilježje korisno je u identificiranju *izdvojenica* u nizu. To su one vrijednosti koje od prosjeka odstupaju za više od +3σ i manje od -3σ. Numeričku varijablu moguće je standardizirati i primjenom sljedećih izraza:

$$z_i = \frac{x_i}{\sigma}, \qquad z_i = \frac{x_i}{x_{\max}}, \qquad z_i = \frac{x_i}{\mu}, \qquad z_i = \frac{x_i}{x_{\max} - x_{\min}}, \qquad z_i = \frac{x_i - x_{\min}}{x_{\max} - x_{\min}}$$

Odstupanja individualnih vrijednosti od prosjeka izražena u standardnim devijacijama predstavljaju površinu krivulje, mogu biti promatrana kao najmanja i interpretiraju se Čebiševljevim pravilom. Ako se želi izračunati proporcija mogućih odstupanja, primjenjuje se empirijsko pravilo.

7.1. Čebiševljevo pravilo

Čebiševljevo pravilo govori da je najmanja proporcija članova bilo kojeg niza,bez obzira na oblik distribucije, obuhvaćena intervalom $\mu \pm k\sigma, k > 1$ jednaka $1 - \frac{1}{k^2}$.

$$p = \left(1 - \frac{1}{k^2}\right)$$

k – broj standardnih devijacija (k ne mora biti cijeli broj)
p – proporcija članova niza obuhvaćena intervalom od k standardnih devijacija

Interval za određivanje najmanje proporcije članova glasi:

$$P(\mu - k\sigma < X < \mu + k\sigma) \geq 1 - \frac{1}{k^2}$$

Čebiševljevim pravilom određuje se površina onih podataka koji odstupaju od prosjeka za k standardnih devijacija. Njime se izračunava raspon varijacije u kojem se očekuje određeni dio podataka niza.

U pojasu $\mu \pm 2\sigma$ nalazi se najmanje 75% svih podataka.

U pojasu $\mu \pm 3\sigma$ nalazi se najmanje 88.89% svih podataka.

Primjer 7.2. Čebiševljevo pravilo

Prosječna cijena oranica u Republici Hrvatskoj iznosi 21.158,00 kn/ha, uz standardno odstupanje od 5.240,00 kn/ha. Potrebno je naći interval u kojem su zapisane cijene oranica po *ha* za najmanje 75% oranica u Hrvatskoj.

$$P(21.158 - 2 \cdot 5.240 < X < 21.158 + 2 \cdot 5.240) \geq 1 - \frac{1}{2^2}$$

$$P(10.678 < X < 31.638) \geq 0{,}75$$

T: Cijene oranica za najmanje 75% oranica u Republici Hrvatskoj kreću se u intervalu od 10.678,00 kn/ha do 31.638,00 kn/ha.

Primjer 7.3. Čebiševljevo pravilo

Mjerenjem je utvrđeno da je djelatnica banke tijekom jednog radnog tjedna prosječno obrađivala zahtjeve za odobrenje kredita klijenata u trajanju od 95 minuta. Prosječno, standardno, odstupanje od prosječne obrade kredita iznosilo je 13 minuta. Koristeći Čebiševljevo pravilo potrebno je procijeniti najmanji postotak zahtjeva kredita čija je obrada trajala od 60 do 120 minuta.

$$k = \frac{120 - 95}{13} = 1{,}923$$

$$p = \left(1 - \frac{1}{k^2}\right) = 1 - \frac{1}{1{,}923^2} = 1 - \frac{1}{3{,}698} = 1 - 0{,}270 = 0{,}73$$

T: Najmanje 73% zahtjeva za kredit djelatnica je obradila u trajanju od 60 do 120 minuta.

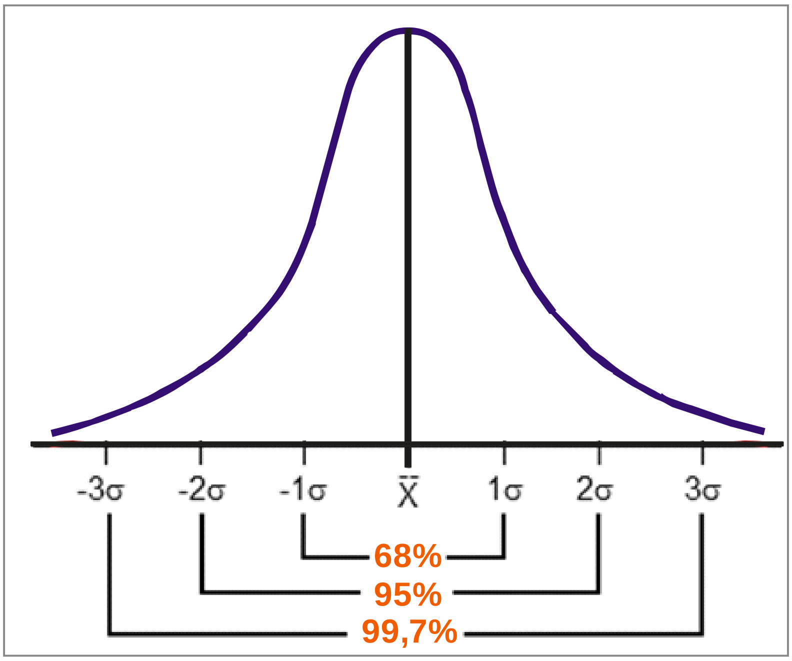
7.2. Empirijsko pravilo

Za razliku od Čebiševljevog pravila koje je primjenjivo na bilo koji oblik distribucije, za zvonolike (simetrične) distribucije, prilikom određivanja rasporeda podataka, moguće je primijeniti svojstva normalne distribucije (slika 7.1.):

- interval $\mu+1\sigma$ obuhvaća oko 68% podataka
- interval $\mu\pm2\sigma$ obuhvaća oko 95% podataka
- interval $\mu\pm3\sigma$ obuhvaća oko 99.73% podataka

Odstupanja od ±2σ nazivaju se *tipičnim* odstupanjima. To su odstupanja u kojima se očekuje najveći broj podataka.

Slika 7.1. Raspodjela podataka prema empirijskom pravilu



Primjer 7.4. Empirijsko pravilo

Mjerenjem je utvrđeno da je djelatnik tvrtke Pošta Expres tijekom radnog dana zaprimanje paketa prosječno obrađivao 7,2 minute. Prosječno, standardno, odstupanje od prosječnog trajanja zaprimanja paketa iznosilo je 1,8 minuta. Potrebno je utvrditi:

Koliko je najkraće i najduže trajalo zaprimanje paketa?

Nalazi li se zaprimanje paketa koje je trajalo 10 minuta u tipičnim odstupanjima?

Koji klijenti su dio izdvojenica s obzirom na trajanje zaprimanja njihovog paketa?

$\mu = 7{,}2 \qquad \sigma = 1{,}8$

$\mu - 1\sigma = 7{,}2 - 1{,}8 = 5{,}4 \qquad \mu + 1\sigma = 7{,}2 + 1{,}8 = 9$

$\mu - 2\sigma = 7{,}2 - 2 \cdot 1{,}8 = 3{,}6 \qquad \mu + 2\sigma = 7{,}2 + 2 \cdot 1{,}8 = 2 * 10{,}8$

$\mu - 3\sigma = 7{,}2 - 3 \cdot 1{,}8 = 1{,}8 \qquad \mu + 3\sigma = 7{,}2 + 3 \cdot 1{,}8 = 12{,}6$

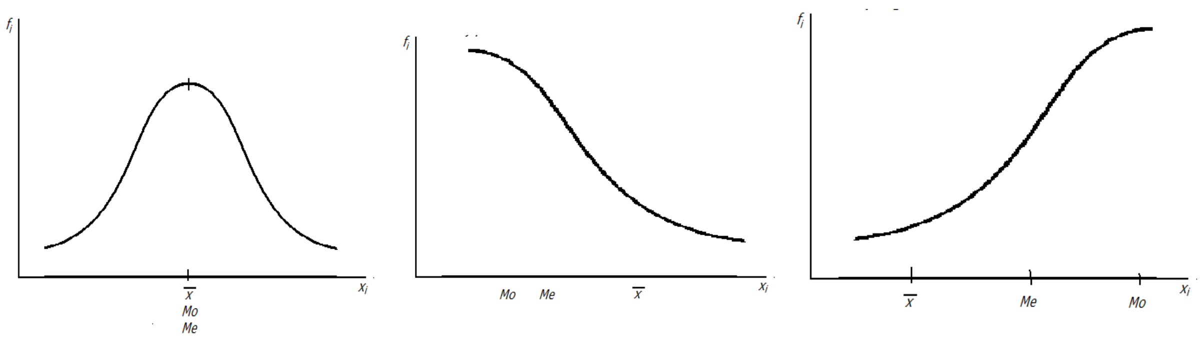
T: Prema empirijskom pravilu, gotovo svi paketi zaprimljeni su u trajanju od 1,8 do 12,6 minuta. Zaprimanje paketa koje je trajalo 10 minuta, naziva se *tipičnim*, odnosno odstupanje od prosjeka zaprimanja je *tipično*. Zaprimanja paketa koja traju kraće od 1,8 i duže od 12,6 minuta smatraju se *izdvojenicama*.

8. MJERE ASIMETRIJE

Mjerama asimetrije mjeri se način rasporeda podataka prema osi simetrije. Mjere asimetrije određuju odstupanja individualnih vrijednosti numeričkog obilježja od aritmetičke sredine, medijana, moda i drugih srednjih vrijednosti. Normalna distribucija Gaussova distribucija je simetrična (slika 8.1.a), u odnosu na normalnu distribucija može biti pozitivno asimetrična (slika 8.1.b), ili negativno asimetrična (slika 8.1.c).

Slika 8.1. Raspored podataka u distribuciji frekvencija i odnosi srednjih vrijednosti

a) simetrična distribucija b) pozitivno asimetrična c) negativno asimetrična
distribucija distribucija



Mo = Me = μ   Mo < Me <μ   Mo > Me >μ

U simetričnoj distribuciji aritmetička sredina, medijan i mod imaju jednake vrijednosti. U pozitivno asimetričnoj distribuciji mod ima najmanju vrijednost, a aritmetička sredina najveću, dok u negativno asimetričnoj distribuciji aritmetička sredina ima najmanju vrijednost. Vrijednost medijana je uvijek između moda i aritmetičke sredine.

Promatra li se asimetrija u odnosu na aritmetičku sredinu u simetričnom rasporedu, postoji jednako negativnih i pozitivnih odstupanja. Pozitivno asimetrična distribucija u kojoj ima više pozitivnih odstupanja još se naziva desnostrano asimetrična, a negativno asimetrična, u kojoj ima više negativnih odstupanja, naziva se i ljevostrano asimetrična.

Najčešće se koriste sljedeće mjere asimetrije: koeficijent asimetrije α_3 koji predstavlja potpunu mjeru asimetrije te nepotpune mjere: Pearsonove mjere i Bowleyeva mjera.

8.1. Koeficijent asimetrije Alfa 3 (α_3)

Alfa 3 izračunava se pomoću trećeg momenta oko sredine koji je pogodan zbog toga što, zbog neparne potencije, zadržava predznak odstupanja kako negativni, tako i pozitivni.

Treći moment oko sredine izračunava se pomoću izraza:

Za negrupirane nizove

$$\mu_3 = \frac{\sum_{i=1}^{N} (x_i - \mu)^3}{N}$$

Za grupirane nizove

$$\mu_3 = \frac{\sum_{i=1}^{k} f_i (x_i - \mu)^3}{N}$$

Alfa 3 predstavlja omjer trećeg momenta oko sredine i standardne devijacije na treću potenciju. Izračunava se pomoću izraza:

$$\alpha_3 = \frac{\mu_3}{\sigma^3}$$

simetrična distribucija – $\alpha_3=0$

pozitivno asimetrična – $\alpha_3>0$

negativno asimetrična – $\alpha_3<0$

Očekivani rezultat alfa 3 je u intervalu $\pm 2\sigma$.

Za daljnju analizu distribucije frekvencija potrebni su izračuni u radnoj tablici 8.1.

Tablica 8.1. Zaposleni u poslovnom subjektu Alfa prema dobi

Dob zaposlenika	Broj zaposlenih	Veličina razreda	Korigirane frekvencije	Brojnik trećeg momenta oko sredine	Brojnik četvrtog momenta oko sredine
x_i	f_i	i_i	$fc_i=f_i/i_i$	$(x_i-\mu)^3$	$(x_i-\mu)^4$
18 – 21	2	3	0,67	-17.653,93	364.847,80
21 – 25	6	4	1,50	-30.353,53	521.068,89
25 – 30	9	5	1,80	-18.290,67	231.681,78
30 – 35	5	5	1,00	-2.253,15	17.274,14
35 – 40	4	5	0,80	-75,85	202,27
40 – 45	9	5	1,80	114,33	266,78
45 – 50	5	5	a 1,00	1.971,85	14.460,25
Mo50 – 55	11	5	**b 2,20**	20.636,41	254.515,69
55 – 60	4	5	c 0,80	20.830,81	361.067,46
60 – 65	2	5	0,40	22.278,74	497.558,54
Ukupno	**57**			**-2.794,97**	**2.262.943,60**

Primjer 8.1. Izračunavanje mjere asimetrije α_3

$$\mu_3 = \frac{-2.794,97}{57} = -49,03 \qquad \alpha_3 = \frac{-49,03}{12,27^3} = \frac{-49,03}{1.847,28} = -0,03$$

Alfa 3 pokazuje da je distribucija frekvencija jako slabo negativno asimetrična. Obzirom na veličinu α_3 (-0,03), distribucija je gotovo simetrična.

Prema pravilu ako je Mo > Me >μ distribucija je negativno asimetrična: 52,31 > 41,39 > 40,17.

8.2. Pearsonove mjere asimetrije

Pearsonove mjere asimetrije (skewnes), temelje se na odnosu srednjih vrijednosti u distribucijama frekvencija. Definiraju se kao standardizirano odstupanje vrijednosti medijana ili moda od aritmetičke sredine.

U upotrebi su dvije Pearsonove mjere asimetrije: skewnes na temelju moda i skewnes na temelju medijana. Ako je distribucija jednomodalna, očekivani rezultat skewnesa je u intervalu ±3σ.

skewnes na temelju moda

$$S_k = \frac{(\mu - Mo)}{\sigma}$$

skewnes na temelju medijana

$$S_k = \frac{3(\mu - Me)}{\sigma}$$

simetrična distribucija – $S_k = 0$

pozitivno asimetrična – $S_k > 0$

negativno asimetrična – $S_k < 0$

Primjer 8.2. Izračunavanje mjere asimetrije S_k na temelju moda

$$Mo = 50 + \frac{(2{,}2 - 1)}{(2{,}2 - 1) + (2{,}2 - 0{,}8)} \cdot 5 = 50 + \frac{1{,}2}{1{,}2 + 1{,}4} \cdot 5 = 50 + 2{,}3077 = 52{,}31$$

Prema izračunatom:

$\mu = 40{,}17$

$\sigma = 12{,}27$

$$S_k = \frac{(40{,}17 - 52{,}31)}{12{,}27} = \frac{-12{,}14}{12{,}27} = -0{,}99$$

Primjer 8.3. Izračunavanje mjere asimetrije S_k na temelju medijana

Prema izračunatom:

$\mu = 40{,}17$

Me = 41,39

$\sigma = 12{,}27$

$$S_k = \frac{3(40{,}17 - 41{,}39)}{12{,}27} = \frac{-3{,}66}{12{,}27} = -0{,}30$$

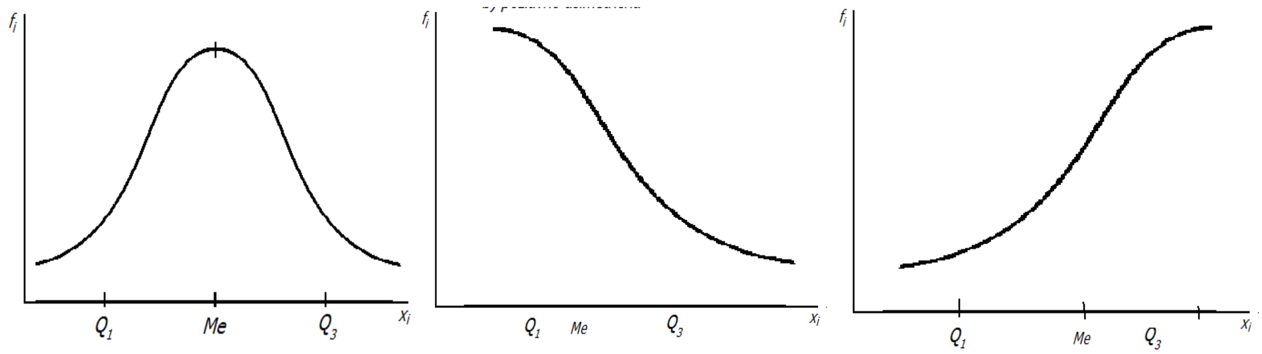
Pearsonove mjere asimetrije pokazuju blago negativno asimetričnu distribuciju.

8.3. Bowleyjeva mjera asimetrije

Bowleyjeva mjera asimetrije temelji se na odnosu medijana i kvartila. U simetričnim distribucijama, Bowleyeva mjera jednaka je 0, u pozitivno asimetričnim mjera je pozitivna, a u negativno asimetričnim je negativna. U simetričnim distribucijama, razlika između medijana i donjeg kvartila jednaka je razlici između gornjeg kvartila i medijana. U pozitivno asimetričnoj distribuciji, razlika između medijana i donjeg kvartila manja je od razlike između gornjeg kvartila i medijana, a u negativnoj, razlika između medijana i donjeg kvartila veća je od razlike između gornjeg kvartila i medijana. Očekivana vrijednost Bowleyjeve mjere asimetrije je u intervalu ±1σ.

Slika 8.2. Raspored podataka u distribuciji frekvencija i odnosi medijana, donjeg kvartila i gornjeg kvartila

a) simetrična distribucija b) pozitivno asimetrična distribucija c) negativno asimetrična distribucija



simetrična distribucija – $S_{kQ}=0$ $\quad Me - Q_1 = Q_3 - Me \quad Q_1 + Q_3 - 2Me = 0$

pozitivno asimetrična – $S_{kQ}>0$ $\quad Me - Q_1 < Q_3 - Me$

negativno asimetrična – $S_{kQ}<0$ $\quad Me - Q_1 > Q_3 - Me$

$$S_{kQ} = \frac{Q_1 + Q_3 - 2Me}{Q_3 - Q_1}$$

Primjer 8.4. Izračunavanje mjere asimetrije S_{kQ}

Prema izračunatom $Q_1 = 28{,}47 \; Q_3 = 51{,}25 \; Me = 41{,}39$

$$S_{kQ} = \frac{28{,}47 + 51{,}25 - 2 \cdot 41{,}39}{51{,}25 - 28{,}47} = \frac{-3{,}06}{22{,}78} = -0{,}13$$

Distribucija je vrlo blago negativno asimetrična.

Dokaz pokazatelja negativne asimetričnosti distribucije $Me - Q_1 > Q_3 - Me$

$41{,}39 - 28{,}47 > 51{,}25 - 41{,}39 \qquad 12{,}92 > 9{,}86$

9. MJERE ZAOBLJENOSTI

9.1. Koeficijent zaobljenosti α4

Koeficijent zaobljenosti α4 mjeri zaobljenost modalnog vrha distribucije. Izračunava se kao omjer četvrtog momenta oko sredine i standardne devijacije na četvrtu potenciju, pomoću izraza:

za negrupirane nizove — za grupirane nizove

$$\alpha_4 = \frac{\mu_4}{\sigma^4} \qquad \mu_4 = \frac{\sum_{i=1}^{N}(x_i-\mu)^4}{N} \qquad \mu_4 = \frac{\sum_{i=1}^{k} f_i(x_i-\mu)^4}{N}$$

$\alpha_4 = 3$ zaobljenost normalne distribucije

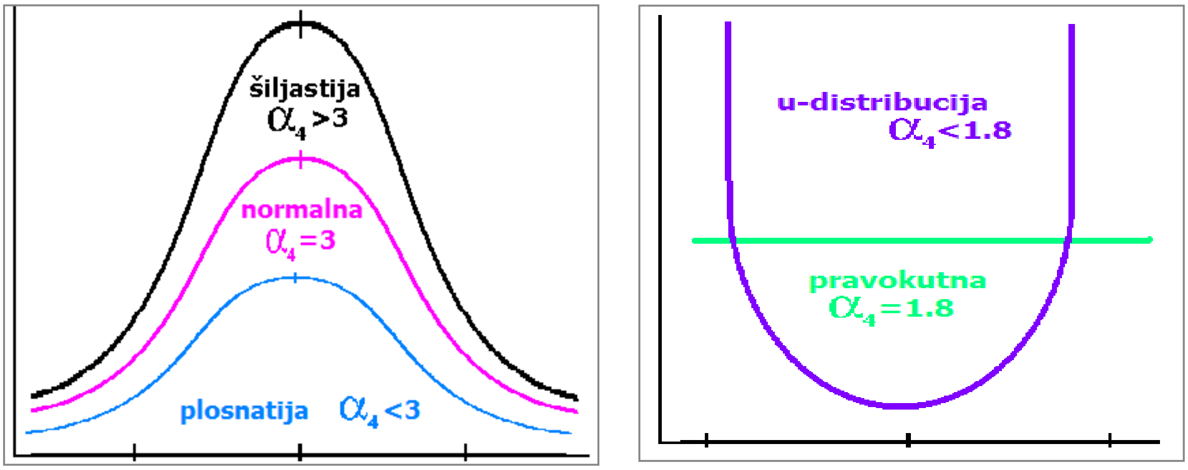
$\alpha_4 > 3$ šiljatija distribucija od normalne distribucije

$\alpha_4 < 3$ plosnatija distribucija od normalne distribucije

$\alpha_4 \approx 1.8$ pravokutna distribucija

$\alpha_4 < 1.8$ *U* distribucija

Slika 9.1. Izgled vrha distribucije s obzirom na vrijednost α4



Primjer 9.1. Izračunavanje mjere zaobljenosti α4

Prema podacima iz tablice 8.1.

$$\mu_4 = \frac{2.262.943,60}{57} = 39.700,76 \qquad \alpha_4 = \frac{39.700,76}{12,27^4} = \frac{39.700,76}{22.666,18} = 1,75$$

Alfa 4 pokazuje da je distribucija približno pravokutna s blagim izgledom *U* distribucije.

9.2. Koeficijent zaobljenosti Eksces

Ponekad se od mjere zaobljenosti alfa 4 oduzima tri, a rezultat je alternativna mjera zaobljenosti tzv. Eksces (engl. *Kurtosis*)., pomoću izraza:

$K = \alpha_4 - 3$

K = 0 zaobljenost normalne distribucije

K > 0 šiljatija distribucija od normalne distribucije

K < 0 plosnatija distribucija od normalne distribucije

Primjer 9.2. Izračunavanje mjere zaobljenosti K

Prema podacima iz primjera 9.1.:

$K = 1{,}75 - 3 = -1{,}25$

Budući da je K<0,to pokazuje da je distribucija plosnatija od normalne.

Rezultati pokazuju da je α_4 preciznija mjera zaobljenosti od K.

10. MJERE KONCENTRACIJE

Mjere koncentracije su mjere nejednakost statističkog niza. Njima se mjeri način rasporeda totala po jedinicama niza ili modaliteta statističkih varijabli. Kao apsolutne mjere koncentracije najčešće se koriste **koncentracijski omjeri i Herfindahlov indeks**, a kao relativna mjera (mjera nejednakosti), najčešće se koristi **Ginijev koeficijent koncentracije.**

10.1. Koncentracijski omjeri

Koncentracijski omjer reda *r* uređenog niza s *N* članova,određuje se tako da se zbroj r vrijednosti podijeli sa zbrojem *N* vrijednosti. Pri tome su podaci poredani od najvećeg prema najmanjem, u padajućem nizu.

$$C_r = \frac{x_1 + x_2 + ... + x_r}{\sum_{i=1}^{N} x_i} = \frac{\sum_{i=1}^{r} x_i}{\sum_{i=1}^{N} x_i} \qquad C_r = \frac{1}{N}\sum_{i=1}^{r}\frac{x_i}{\overline{x}} \qquad \frac{r}{N} \le C_r \le 1$$

N = broj vrijednosti, odnosno modaliteta statističke varijable

x_i - pojedinačne vrijednosti varijable

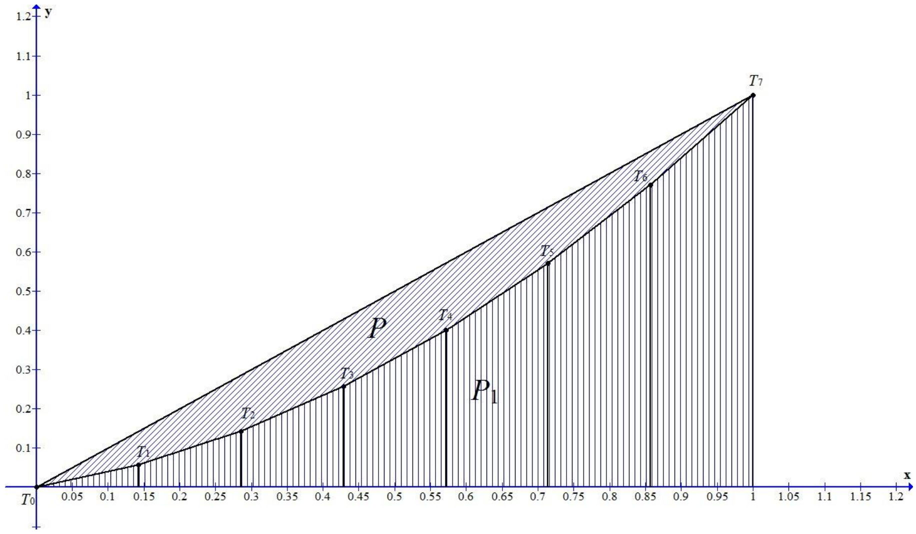
Ako se radi o *ravnomjernoj raspodjeli*, koncentracijski omjer poprima vrijednost od 1 do N. Ako su vrijednosti međusobno jednake, vrijednost omjera jednaka je 1/N,a ako se radi o *maksimalnoj raspodijeli*, koncentracijski omjer poprima vrijednost 1. U tom slučaju su sve vrijednosti, osim posljednje, jednake nuli.

Za analizu koncentracije koristi se grafički prikaz pomoću Lorenzove krivulje

1) na apscisi se nalazi aritmetičko mjerilo za kumulativni niz relativnih frekvencija
2) na ordinati se nalazi aritmetičko mjerilo za kumulativni niz proporcija podataka
3) prva točka ima koordinate (0,0); posljednja točka ima koordinate (1,1); koordinate ostalih točaka određene su vrijednostima članova kumulativnih nizova

4) u grafički prikaz ucrtava se pravac jednolike raspodjele, on prolazi točkama (0,0) i (1,1)

Slika 10.1. Primjer Lorenzove krivulje



Izvor: Kovačić, B., Opačić, R. i Marohnić, L., O Ginijevu koeficijentu koncentracije, *math e, Hrvatski matematički elektronički časopis,* https://hrcak.srce.hr/file/151776

Primjer 10.1. Izračunavanje koncentracijskih omjera

Predstavljeni su podaci o broju aktivnih trgovačkih društava prema veličini u Republici Hrvatskoj na dan 31. ožujka 2018. godine[17].

Tablica 10.1. Aktivna trgovačka društva u Republici Hrvatskoj prema veličini

Veličina aktivnih trgovačkih društava	Velika	Srednja	Mala	Mikro	UKUPNO
	332	1.184	8.804	99.924	110244
Trgovačka društva po veličini frekvencije	Mikro	Mala	Srednja	Velika	UKUPNO
	99.924	8.804	1.184	332	110.244
Kumulativ	99.924	108.728	109.912	110.244	
C_r	0,906389	0,986249	0,996988	1	
%	90,64	98,62	99,70	100	

Izvor: Registar poslovnih subjekata, Hrvatska gospodarska komora, URL

[17] U Registru poslovnih subjekata nije izričito naznačeno da su u evidenciju uključena samo trgovačka društva (ne svi poslovni subjekti), no nakon upita proslijeđenog nadležnima u Hrvatskoj gospodarskoj komori, potvrđeno je da su podaci preuzeti od Državnog zavoda za statistiku Republike Hrvatske te da Registar poslovnih subjekata obuhvaća trgovačka društva registrirana u Republici Hrvatskoj, a evidenciju o obrtima vodi Hrvatska obrtnička komora.

Primjer izračuna vrijednosti koncentracijskog omjera drugog reda: N=4, R=2

$$C_2 = \frac{99.924 + 8.804}{110.244} = \frac{108.728}{110.244} = 0{,}9862$$

T: Koncentracijski koeficijent drugog reda pokazuje da mikro i mala trgovačka društva predstavljaju 98,62% svih trgovačka društva u Republici Hrvatskoj na dan 31. ožujka 2018. godine

10.2. Herfindahl– Hirschmanov indeks

Herfindahl– Hirschmanov indeks (engl. *HH index* ili *HHI-score*) je apsolutni indeks koncentracije. Izračunava se pomoću izraza:

$$H = p_1^2 + p_2^2 + ... + p_i^2 + ... + p_N^2 = \left(\frac{x_1}{\sum_{i=1}^{N} x_i}\right)^2 + \left(\frac{x_2}{\sum_{i=1}^{N} x_i}\right)^2 + ... + \left(\frac{x_i}{\sum_{i=1}^{N} x_i}\right)^2 + ... + \left(\frac{x_N}{\sum_{i=1}^{N} x_i}\right)^2 = \sum_{i=1}^{N} p_i^2$$

ili

$$H = \frac{\sum_{i=1}^{N} x_i^2}{\left(\sum_{i=1}^{N} x_i\right)^2} \qquad \frac{1}{N} \le H \le 1$$

Može se izračunati i pomoću izraza:

$$H = \frac{1}{N}\left(v^2 + 1\right) = \frac{1}{N}\left(\frac{\sigma^2}{\mu^2} + 1\right)$$

μ– aritmetička sredina

σ – standardna devijacija

υ – koeficijent varijacije niza nepomnožen sa sto

Kada je razdioba totala po jedinicama niza ravnomjerna, indeks poprima najmanju vrijednost 1/N, a poprima najveću vrijednost kada je koncentracija maksimalna.

Primjer 10.2. Izračunavanje Herfindahl – Hirschmanova indeksa

Tablica 10.2. Aktivna trgovačka društva u Republici Hrvatskoj prema veličini

Trgovačka društva po veličini	Mikro	Mala	Srednja	Velika	UKUPNO (∑)
frekvencije x_i	99.924	8.804	1.184	332	110.244
$x_i/\sum x_i$	0,9064	0,0799	0,0107	0,0030	1,0000
$(x_i/\sum x_i)^2 = p_i^2$	0,8215	0,0064	0,0001	0,0000	0,8280
x_i^2	9.984.805.776	77.510.416	1.401.856	110.224	10.063.828.272
u %	90,64	7,99	1,07	0,30	100,00
u %^{2}	8.215,42	63,77	1,15	0,09	8.280,44

Izvor: Registar poslovnih subjekata, Hrvatska gospodarska komora, URL

Iz tablice se može očitati da je:

$$H = \sum_{i=1}^{N} p_i^2 = 0{,}828$$

ili izračunati:

$$H = \frac{\sum_{i=1}^{N} x_i^2}{\left(\sum_{i=1}^{N} x_i\right)^2} = \frac{10.063.828.272}{110.244^2} = 0{,}828$$

T: Prema Herfindahl – Hirschmanovom indeksu,s obzirom na veličinu trgovačkih društava, koncentracija je izrazito velika.

Za razliku od raspona od 0 do 1, ako se koriste cijeli postoci (što je u gospodarskoj praksi čest slučaj), indeks se kreće od 0 do 10.000 bodova. Npr. H od 0,75 jednak je 7.500 bodova. Tada je česta upotreba simbola HHI.

Interpretacija vrijednosti HH indeksa:

HHI< 1000 niska koncentracija

HHI = 1000 – 1800 srednja koncentracija

HHI > 1800 visoka koncentracija

$$HHI = 90{,}64^2 + 7{,}99^2 + 1{,}07^2 + 0{,}3^2 = 8.280{,}44$$

T: HHI > 1800 (8.280,44) pokazuje visoku koncentraciju s obzirom na veličinu trgovačkih društava.

Herfindahl – Hirschmanov indeks pokazuje kolika je koncentracija, to jest stupanj koncentracije, ali ne pokazuje kojem modalitetu pripisati koncentraciju. Kvadriranje vrijednosti po modalitetima pridaje veću važnost modalitetima sa većim udjelom te je uočljivije kojem modalitetu pripada veći ili manji udio.

T: Udio mikro trgovačkih društava u Republici Hrvatskoj ($90,64^2$=8.215,42) je izrazito velik u odnosu na ostala trgovačka društva razvrstana prema veličini.

10.3. Ginijev koeficijent

Ginijev koeficijent koncentracije temelji se na utvrđivanju površina između pravca jednolike raspodjele i Lorenzove krivulje. Što je koncentracija veća, to se Lorenzova krivulja više udaljuje od toga pravca. Njihova vrijednost kreće se u rasponu od o do 1. Što je ta vrijednost bliža nuli, vrijednosti niza su ravnomjernije raspoređene, odnosno na svaki modalitet otpada približno jednak udio u totalu. Obrnuto, što je ta vrijednost bliža jedinici, vrijednosti niza su neravnomjernije raspoređene, odnosno većina totala otpada na jedan modalitet.

Ginijev koeficijent se izračunava pomoću izraza:

$$G = \frac{2\sum_{i=1}^{N} i \cdot x_i}{N\sum_{i=1}^{N} x_i} - \frac{N+1}{N} \qquad \text{ili} \qquad G = 1 - \sum_{i=1}^{N} p_i \cdot \left(y_{i-1} + y_i\right)$$

x_i – pojedinačne vrijednosti varijable
i – redni broj podatka

G = 0 nema koncentracije

$0 < G \leq 0,25$	slaba koncentracija
$0.25 < G \leq 0,5$	umjerena koncentracija
$0.5 < G \leq 0.75$	visoka koncentracija
$0.75 < G \leq 1$	izrazito visoka koncentracija
$G = 1$	maksimalna koncentracija

Za izračun Ginijevog koeficijenta statistički niz mora biti sastavljen od nenegativnih vrijednosti ($x_i \geq 0$), najmanje jedna od njih je strogo veća od nule ($i \in [n]$ takav da je $x_i > 0$) i elementi niza moraju biti poredani od najmanjeg prema najvećem ($x_1 \leq x_2 \leq \ldots \leq x_n$).

Primjer 10.3. Izračunavanje Ginijeva koeficijenta

Tablica 10.3. Aktivna trgovačka društva u Republici Hrvatskoj prema veličini

Veličina trgovačkih društava	Boj aktivnih trgovačkih društava	Kumulativ broja aktivnih trgovačkih društava (podtotala)		Lorenzova krivulja			
	x_i		i	p_i	y_i	$i \cdot x_i$	$y_{i-1} + y_i$
Velika	332	332	1	0,25	0,0030	332	0,0030
Srednja	1.184	1.516	2	0,5	0,0138	2.368	0,0168
Mala	8.804	10.320	3	0,75	0,0936	26.412	0,1104
Mikro	99.924	110.244	4	1	1,0000	399.696	1,1104
UKUPNO	110.244					428.808	1,2405

Izvor: Registar poslovnih subjekata, Hrvatska gospodarska komora, URL

Svakom razredu pridružena je apsolutna frekvencija. Umjesto apsolutne frekvencije, mogu se razmatrati i relativne frekvencije koje označavaju udio broja dotičnog razreda u ukupnom broju svih članova niza.

y_i – kumulativ proporcija podtotala

p_i – kumulativ proporcija trgovačkih društava

$$G = \frac{2 * 428.808}{4 * 110.244} - \frac{4+1}{4} = \frac{857.616}{440.976} - \frac{5}{4} = 1,9448133 - 1,25 = 0,69$$ ili

$$G = 1 - \frac{1}{4}(1,2405) = 1 - 0,31 = 0,69$$

Za manje vrijednosti N ($N \leq 20$) koristi se **normirani Ginijev koeficijent,** prema izrazu:

$$G^* = \frac{G}{G_{\max}} = G \cdot \frac{N}{N-1} \qquad\qquad G^* = 0,69 \cdot \frac{4}{4-1} = 0,92$$

T: Iz podataka za Lorenzovu krivulju i dobivenih vrijednosti Ginijevog koeficijenta, može se zaključiti da su trgovačka društva po veličini neravnomjerno raspoređena. Lorenzova krivulja pokazuje da velika trgovačka društva čine samo 0,30% ukupnog broja aktivnih trgovačka društva, dok mikro trgovačka društva čine više od 90,64% ukupnog broja aktivnih, $(1 - 0,0936) * 100 = 90,64\%$.

Ginijev koeficijent koncentracije i normirani Ginijev koeficijent pokazuju da su aktivna trgovačka društva u Republici Hrvatskoj prema veličini izrazito neravnomjerno raspoređena.

11. KORELACIJSKA ANALIZA

Kovarijanca je polazna veličina za mjerenje jakosti i smjera linearne statističke povezanosti dviju pojava. Kovarijanca je aritmetička sredina umnožaka odstupanja vrijednosti varijable X od njezine aritmetičke sredine i vrijednosti varijable Y od njezine aritmetičke sredine. Kao deskriptivnostatistička veličina dana je izrazom:

$$\mu_{11} = \frac{1}{N}\sum_{i=1}^{N}\left(x_i - \bar{x}\right)\left(y_i - \bar{y}\right) = \operatorname{cov}(X, Y)$$

cov (x,y) = 0, ako su jednake sve vrijednosti barem jedne varijable

cov (x,y) > 0, ako postoji tendencija da iznadprosječne vrijednosti jedne varijable prate iznadprosječne vrijednosti druge varijable

cov (x,y) < 0, ako postoji tendencija da ispodprosječne vrijednosti jedne varijable prate ispodprosječne vrijednosti druge varijable

11.1.Koeficijent jednostavne linearne korelacije

Kovarijanca pokazuje postoji li između pojava kovarijacija, ali ne pokazuje stupanj kovarijacije. Kako bi se izračunao stupanj linearne statističke povezanosti računa se **kovarijanca standardiziranih vrijednosti** varijabli X i Y,tzv. **Pearsonov koeficijent korelacije *r*.** Računa se prema izrazu:

$$r = \frac{\mu_{11}}{\sigma_x \sigma_y} \qquad \sigma_x = \sqrt{\frac{1}{N}\sum_{i=1}^{2}\left(x_i - \bar{x}\right)^2} \qquad \sigma_y = \sqrt{\frac{1}{N}\sum_{i=1}^{2}\left(y_i - \bar{y}\right)^2}$$

Koeficijent korelacije može se izračunati i prema.

$$r = \frac{N\sum_{i=1}^{N} x_i y_i - \sum_{i=1}^{N} x_i \sum_{i=1}^{N} y_i}{\sqrt{\left(N\sum_{i=1}^{N} x^2 - \left(\sum_{i=1}^{N} x_i\right)^2\right) \cdot \left(N\sum_{i=1}^{N} y_i^2 - \left(\sum_{i=1}^{N} y_i\right)^2\right)}} \text{ ili } \qquad r = \frac{\sum xy - N\bar{x}\bar{y}}{\sqrt{\left(\sum x^2 - N\bar{x}^2\right)\cdot\left(\sum y^2 - N\bar{y}^2\right)}}$$

Korelacijskom analizom određuju se jakosti veze. Ako je:

r = 0, među pojavama ne postoji veza.

r = -1, veza je potpuna i negativnog smjera.

r = +1, veza je potpuna i pozitivnog smjera.

11.2. Spearmanov koeficijent korelacije ranga

Spearmanov koeficijent korelacije ranga (engl. *Spearman's rho*) mjeri smjer i jakost povezanosti dviju rang varijabli.

Spearmanov koeficijent korelacije ranga označava se s r_s ili ρ. Koristi se kada barem jedan od skupova podataka slijedi ordinalnu ljestvicu te kada raspodjela podataka značajno odstupa od normalne raspodjele. On ne postavlja uvjet linearnosti, simetričnosti i veličine uzorka. Računa se pomoću izraza:

$$r_s = 1 - \frac{6\sum_{i=1}^{n} d_i^2}{n^3 - n} \qquad d_i = r(x_i) - r(y_i) \qquad -1 \le r_s \le 1$$

d – razlika rangova varijabli

Numeričke varijable potrebno je transformirati u rang varijable tako da se najmanjoj vrijednosti varijable *x* pridruži rang 1, sljedećoj 2, sve do *N*. Isto je potrebno učiniti i kod varijable *y*.

Tumači se kao Pearsonov koeficijent korelacije *r*.

Ako se u nizu nalaze vrijednosti koje su jednake po rangu, takve rangove nazivamo *vezani rangovi*. Umjesto istog ranga pridružuje se rang koji je jednak prosjeku njihovih pozicija. Npr. ako iza ranga 6 slijede tri ista ranga (s jednakim rezultatima) svaki od njih zauzet će rang 8 (7+8+9 = 24/3 = 8). Sljedeći rang zauzet će rang 10.

U slučaju da u nizu postoje vezani rangovi potrebno je izračunati korigirani koeficijent korelacije ranga.

Za sve vezane rangove izračunava se korekturni faktor *cf* prema izrazu:

$$cf = \frac{m\left(m^2-1\right)}{12}$$

Ako su npr. dva rezultata vezana u isti rang korigirani koeficijent korelacije ranga računa se prema:

$$r_s korig. = \frac{\frac{N\left(N^2-1\right)}{6} - \sum_{i=1}^{N} d_i^2 - \sum cf_1 - \sum cf_2}{\sqrt{\left[\frac{N\left(N^2-1\right)}{6} - 2\sum cf_1\right] \cdot \left[\sqrt{\frac{N\left(N^2-1\right)}{6} - 2\sum cf_2}\right]}}$$

Ako se na istim podacima računa *r* i r_s rezultat će biti isti, ali tumačenje je različito, jer r_s ne uvažava relativni položaj pojedinog rezultata među drugim rezultatima (*z*-vrijednosti) nego uvažava razlike među rangovima.

12. REGRESIJSKA ANALIZA

Regresijska i korelacijska analiza ispituju ovisnost jedne varijable o drugoj varijabli ili o više drugih varijabli. Iako ni regresijska ni korelacijska analiza ne objašnjavaju uzročno-posljedičnu vezu između varijabli, nego ispituju postoji li povezanost između njih, regresijom je ipak moguće donositi predviđanja o jednoj varijabli na temelju poznatih podataka druge varijable. Tako je moguće ispitati povezanost između materijalnog standarda stanovništva i uvoza, BDP i izvoza, vrijednosti proizvodnje i broja zaposlenih, potrošnje i visine plaće.

Regresijski model je jednadžba ili skup jednadžbi s konačnim brojem parametara i varijabli. Može biti **linearni** i **nelinearni**.

Zavisna varijabla je varijabla *y*, varijabla čije se varijacije objašnjavaju.

Nezavisne varijable su varijable *x*, kojima se objašnjavaju varijacije zavisne varijable.

12.1. Model jednostavne linearne regresije

Model jednostavne linearne regresije izražava odnos među dvjema pojavama. Model sadrži jednu zavisnu i jednu nezavisnu varijablu. Primjenjuje se za varijable koje su u linearnom statističkom odnosu.

Opći oblik modela jednostavne regresije glasi: $y = f_{(x)} + e$

$f_{(x)}$ – funkcionalni dio modela

e – stohastička varijabla koja odražava nesistemske utjecaje na zavisnu varijablu

Deskripitvno statistička analiza modela jednostavne linearne regresije

Model jednostavne linearne regresije može biti deterministički i statistički.

Deterministički model zapisuje se izrazom: $y = A + Bx$

A – zavisna varijabla

B – nezavisna varijabla

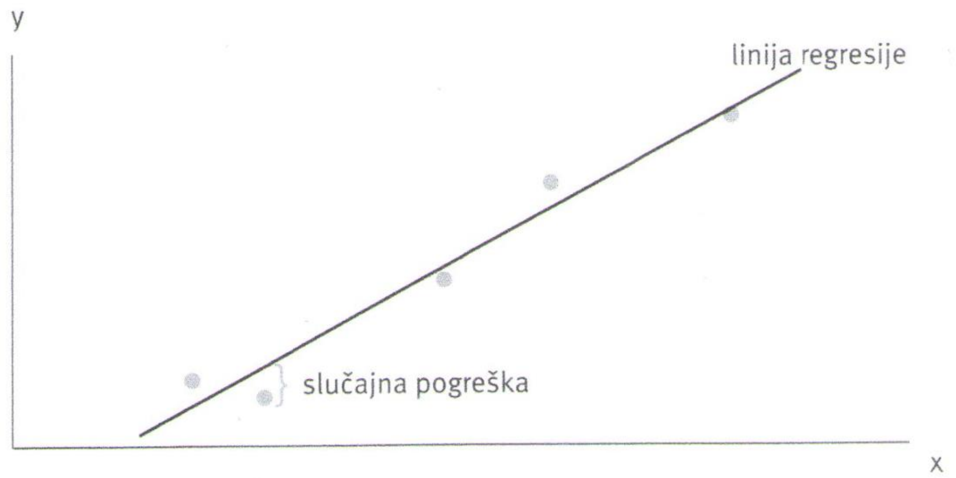
A,B – parametri populacije

U društvenim pojavama se nikada ili rijetko pojavljuje deterministička (funkcionalna) povezanost varijabla. Takva povezanost bi npr. bila kada bi određena plaća djelatnika uvijek rezultirala jednakom ušteđevinom. Društvene pojave uglavnom se mogu objasniti probabilističkim (statističkim) modelom. Pri tome jednakoj promjeni jedne varijable može odgovarati veća ili manja promjena druge varijable. Statistički model zapisuje se na sljedeći način: $y = A + Bx + u$.

u – slučajna pogreška

Slučajna pogreška obuhvaća varijable koje nisu uključene u postavljeni model,a utječu na zavisnu varijablu. Slučajnu pogrešku čini udaljenost originalnog para točaka varijable x i varijable y od linije regresije.

Slika 12.1. Slučajna pogreška



Izvor: Horvat i Mijoč (2014, str. 503)

Cilj regresijske analize je pronaći onaj regresijski pravac koji najbolje odgovara danim podacima, odnosno onaj regresijski pravac kojim su minimizirane udaljenosti svih upisanih točaka.

Primjenom metode najmanjih kvadrata, sve točke dijagrama trebale bi se nalaziti na najmanjoj udaljenosti od linije regresije. Manje udaljenosti točaka od linije regresije rezultiraju manjim zbrojem kvadrata pogreške te većom apsolutnom vrijednosti koeficijenta korelacije.

Primjenom metode najmanjih kvadrata, odnosno minimiziranjem zbroja kvadrata rezidualnih odstupanja, dobiju se izrazi za procjenu parametara *a* i *b*.

$$a = \bar{y} - b\bar{x} \qquad \bar{x} = \frac{\sum_{i=1}^{N} x_i}{N} \qquad \bar{y} = \frac{\sum_{i=1}^{N} y_i}{N}$$

$$b = \frac{\sum_{i=1}^{N} x_i y_i - N\bar{x}\bar{y}}{\sum_{i=1}^{N} x_i^2 - N\bar{x}^2}$$

Konstantni član *b* moguće je izračunati i pomoću izraza:

$$b = \frac{N\sum_{i=1}^{N} x_i y_i - \left(\sum_{i=1}^{N} x_i\right)\left(\sum_{i=1}^{N} y_i\right)}{N\sum_{i=1}^{N} x_i^2 - \left(\sum_{i=1}^{N} x_i\right)^2}$$

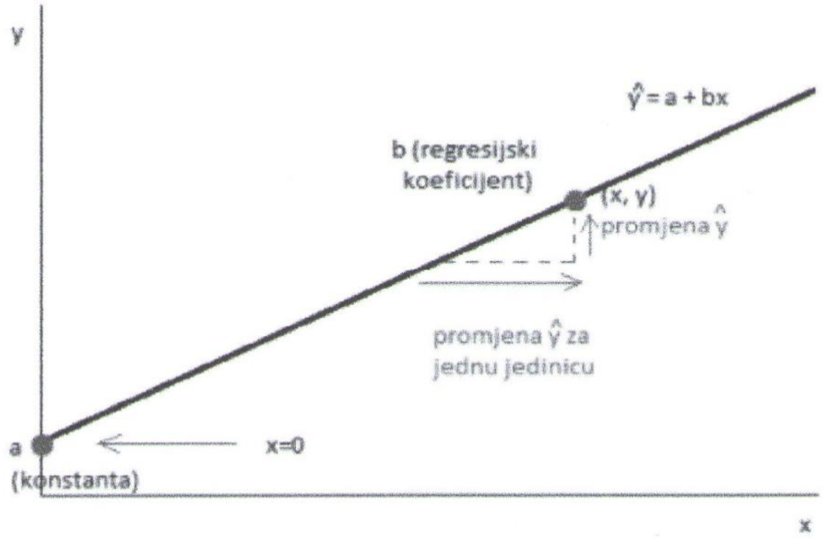
Konstantni član *a* označava odsječak na ordinati. Predstavlja vrijednost regresije u slučaju kada je nezavisna varijabla jednaka nuli. Ukoliko *a* rezultira negativnim predznakom, njegovo tumačenje je nesuvislo i treba ga zanemariti. [18] Negativan predznak parametra *a* ne mijenja tumačenje regresijske jednadžbe.

Regresijski koeficijent *b* pokazuje za koliko se linearno u prosijeku mijenja vrijednost zavisne varijable y ako se nezavisna varijabla poveća za jedan.
Pravac koji predstavlja liniju regresije ucrtava se na temelju procijenjenih parametara *a* i *b*.

---

[18] Takav regresijski model sličan je modelu koji ispituje odnos ponude i cijene, a koji je matematičkim rječnikom opisao Chaing (1994, str. 37) tako da daje odgovor i grafički prikaz za slučaj kada je presjek s okomitom osi negativan: "Kad pretpostavimo da je presjek negativan, postižemo da krivulja ponude ima pozitivan presjek s vodoravnom osi i da pritom zadovoljava ranije postavljeni uvjet da ne bude ponude dok cijena ne bude pozitivna i dovoljno visoka."

Slika 12.2. Linija regresije s procijenjenim parametrima



Izvor: Horvat i Mijoč (2014, str. 506)

Reprezentativnost regresije

Vrijednosti regresije $\hat{y}$ za dane stvarne vrijednosti nezavisne varijable x nazivaju se regresijskim vrijednostima. **Regresijske vrijednosti** određuju se tako da se u regresijsku jednadžbu redom uvrštavaju stvarne vrijednosti nezavisne varijable.

regresijske vrijednosti $\hat{y}_i = a + bx_i \qquad i = 1,2,\ldots,n$

Rezidualna odstupanja izračunavaju se tako da se od stvarne varijable y_i oduzme pripadajuća regresijska vrijednost $\hat{y}_i$. Rezidualna odstupanja su osnova za mjerenje reprezentativnosti modela regresije.

rezidualna odstupanja $u_i = y_i - \hat{y}_i$

ili u relativnom broju, rezidualna odstupanja se računaju prema:

$$u_{i,rel}=\frac{y_i-\hat{y}_i}{y_i}\cdot 100$$

Kako bi se utvrdila reprezentativnost regresije polazi se od jednadžbe analize varijance.

$$\sum_{i=1}^{N}\left(y_i-\bar{y}\right)^2=\sum_{i=1}^{N}\left(\hat{y}_i-\bar{y}\right)^2+\sum_{i=1}^{N}\left(y_i-\hat{y}_i\right)^2$$

Može se pisati i u sljedećem obliku: SS_{yy} = SSR+SSE

SS_{yy} – ukupan zbroj kvadrata

SSR – protumačeni zbroj kvadrata

SSE – rezidualni ili neprotumačeni zbroj kvadrata

ili razvijeni izraz

$$\sum_{i=1}^{N}y_i^{\,2}-N\bar{y}^2=\left[a\sum_{i=1}^{N}y_i+b\sum_{i=1}^{N}x_iy_i-N\bar{y}^2\right]+\left[\sum_{i=1}^{N}y_i^{\,2}-a\sum_{i=1}^{N}y_i-b\sum_{i=1}^{N}x_iy_i\right]$$

ukupni zbroj kvadrata $\sum_{i=1}^{N}\left(y_i-\bar{y}\right)^2$

protumačeni zbroj kvadrata $\sum_{i=1}^{N}\left(\hat{y}_i-\bar{y}\right)^2$

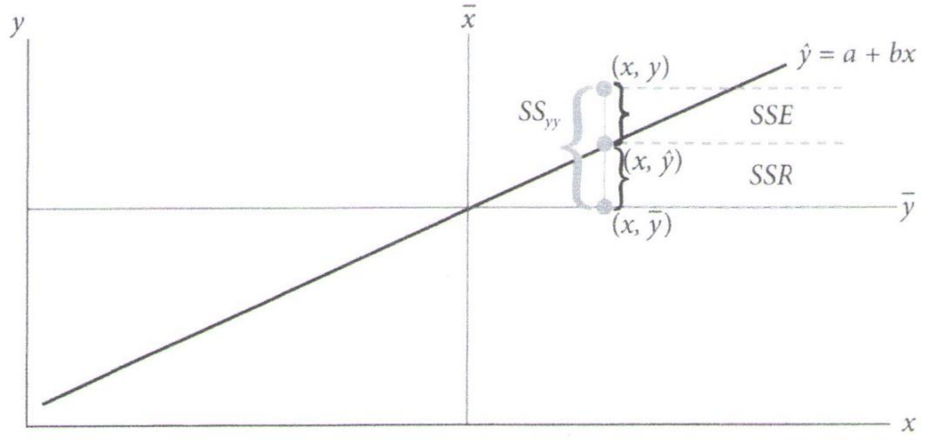
neprotumačeni zbroj kvadrata
ili zbroj kvadrata rezidualnih odstupanja $\sum_{i=1}^{N}\left(y_i-\hat{y}_i\right)^2$

y_i - empirijske ili stvarne vrijednosti zavisne varijable

$\bar{y}$ - aritmetička sredina zavisne varijable

$\hat{y}_i$ - regresijske vrijednosti

Slika 12.3. Ukupna, protumačena i neprotumačena odstupanja modela regresije



Izvor: Horvat i Mijoč (2014, str. 515)

Dijagram rasipanja posjeduje samo jedan pravac kojem je zbroj kvadrata pogreške najmanja.

Zbrojevi kvadrata omogućuju izračunavanje pokazatelja reprezentativnosti koji obuhvaćaju mjere disperzije: standardnu devijaciju i koeficijent varijacije regresije te standardniju (uobičajeniju) mjeru reprezentativnosti kod regresijskih modela, koeficijent determinacije.

Varijanca regresije predstavlja aritmetičku sredinu kvadrata rezidualnih odstupanja. Računa se prema izrazu:

$$\sigma_{\hat{y}}^{2} = \frac{\sum_{i=1}^{N} (y_i - \hat{y})^2}{N}$$

Standardna devijacija regresije interpretira se kao prosječno odstupanje empirijskih vrijednosti od regresijskih izraženih u apsolutnim jedinicama. Standardna devijacija regresije računa se prema izrazu:

$$\sigma_{\hat{y}} = \sqrt{\sigma_{\hat{y}}^{2}}$$

Može se računati i pomoću izraza:

$$\sigma_{\hat{Y}} = \sqrt{\frac{\sum_{i=1}^{N} y_i^2 - a\sum_{i=1}^{N} y_i - b\sum_{i=1}^{N} x_i y_i}{N}}$$

Koeficijent varijacije regresije je odstupanje empirijskih vrijednosti od regresijskih izraženo u relativnom broju.
Koeficijent varijacije regresije računa se prema izrazu: $V_{\hat{y}} = \frac{\sigma_{\hat{y}}}{\bar{y}} 100$

Koeficijent determinacije izračunava se pomoću izraza:

$$r^2 = \frac{a\sum_{i=1}^{N} y_i + b\sum_{i=1}^{N} x_i y_i - N\bar{y}^2}{\sum_{i=1}^{N} y_i^2 - N\bar{y}^2}, 0 \le r^2 \le 1 \qquad \text{ili} \qquad r^2 = \frac{\sum_{i=1}^{N} (\hat{y}_i - \hat{y})^2}{\sum_{i=1}^{N} (y_i - \bar{y})^2}, 0 \le r \le 1$$

Model regresije je reprezentativniji što se koeficijent determinacije više približava jedinici.

Drugi korijen iz koeficijenta linearne determinacije daje koeficijent linearne korelacije koji utvrđuje jakost i smjer veze, a predznak mu je isti kao i regresijskog koeficijenta *b*.

$$r = \pm\sqrt{r^2} \qquad -1 \le r \le 1$$

Tablica 12.1. Raspon koeficijenta determinacije i koeficijenta korelacije s tumačenjem

koeficijent determinacije	koeficijent linearne korelacije	tumačenje
r^2	r	
1	-1	potpuna negativna korelacija
1 – 0,49	-1 – -0,7	snažna negativna korelacija
0,49 – 0,09	-0,7 – -0,3	umjerena negativna korelacija
0,09 – 0	-0,3 – 0	slaba negativna korelacija
0	0	odsutnost korelacije
0,00 – 0,09	0 – 0,3	slaba pozitivna korelacija
0,09 – 0,49	0,3 – 0,7	umjerena pozitivna korelacija
0,49 – 1	0,7 – 1	snažna pozitivna korelacija
1	1	potpuna pozitivna korelacija

Primjer 12.1. Izračunavanje koeficijenta determinacije, koeficijenta korelacije i pokazatelja varijabilnosti, linearne regresijske jednadžbe – jednostavna linearna regresija

Slijedeći iskustva iz gospodarske prakse, pretpostavlja se da između varijabli ukupne bruto proizvodnje električne energije i izvoza električne energije u GWh,te između varijabli finalne potrošnje električne energije i uvoza električne energije,postoji linearna statistička veza. Podaci o proizvodnji, potrošnji, uvozu i izvozu električne energije odnose se na Republiku Hrvatsku u razdoblju od 2002. do 2010. godine. Kako bi se utvrdilo postoji li veza između navedenih varijabli i kojeg je smjera, primijenjen je model jednostavne linearne regresije. Parovi vrijednosti varijabli prikazani su u dvije tablice, gdje je x nezavisna, a y zavisna varijabla. Ukoliko bi se varijable zamijenile, rezultat bi se promijenio jedino u dijelu mjera disperzije. Nakon ispitanih odnosa između prikazanih varijabli, u tablici 11.3. dani su rezultati nereprezentativnog, a u tablici 11.4. reprezentativnog modela jednostavne linearne regresije.

Tablica 12.2. Ukupna bruto proizvodnja električne energije i izvoz električne energije Republike Hrvatske od 2002. do 2010. godine u GWh

Godina	Ukupna bruto proizvodnja el. energije u GWh	Izvoz el. energije u GWh				Regresijske vrijednosti
	x_i	y_i	x_i^2	$x_i y_i$	$y_i 2$	$\hat{y}_i = a + bx_i$
2002.	12.725	792	161.925.625	10.078.200	627.264	2.192,74
2003.	13.248	1.142	175.509.504	15.129.216	1.304.164	2.367,46
2004.	13.976	2.296	195.328.576	32.088.896	5.271.616	2.610,67
2005.	13.140	4.323	172.659.600	56.804.220	18.688.329	2.331,38
2006.	13.037	3.306	169.963.369	43.100.322	10.929.636	2.296,97
2007.	12.462	1.948	155.301.444	24.275.976	3.794.704	2.104,87
2008.	12.616	2.140	159.163.456	26.998.240	4.579.600	2.156,32
2009.	13.149	2.578	172.896.201	33.898.122	6.646.084	2.334,39
2010.	14.669	2.712	215.179.561	39.782.328	7.354.944	2.842,19
Ukupno	119.022	21.237	1.577.927.336	282.155.520	59.196.341	21.237,00
	Rezidualna odstupanja					
	$u_i = y - \hat{y}_i$	$x_i - \bar{x}$	$y_i - \bar{y}$	$(x_i - \bar{x})^2$	$(y_i - \bar{y})^2$	$(x_i - \bar{x}) \cdot (y_i - \bar{y})$
	-1.400,74	-499,67	-1.567,67	249.666,78	2.457.578,78	783.310,78
	-1.225,46	23,33	-1.217,67	544,44	1.482.712,11	-28.412,22
	-314,67	751,33	-63,67	564.501,78	4.053,44	-47.834,89
	1.991,62	-84,67	1.963,33	7.168,44	3.854.677,78	-166.228,89
	1.009,03	-187,67	946,33	35.218,78	895.546,78	-177.595,22
	-156,87	-762,67	-411,67	581.660,44	169.469,44	313.964,44
	-16,32	-608,67	-219,67	370.475,11	48.253,44	133.703,78
	243,61	-75,67	218,33	5.725,44	47.669,44	-16.520,56
	-130,19	1444,33	352,33	2.086.098,78	124.138,78	508.886,78
Ukupno	0,00	0,00	0,00	3.901.060,00	9.084.100,00	1.303.274,00

Rezultati	
$\bar{x}$	*13.224,67* GWh
$\bar{y}$	*2.359,67* GWh
a	*-2.057,37* GWh
b	*0,334* GWh
$\sigma_{\hat{y}}$	*980,29* GWh
$V_{\hat{y}}$	*41,54%*
r^2	*0,0479*
r	*0,2189*

Izvor: Statistički ljetopis Republike Hrvatske 2012., str. 303., URL

Rezultati ispod tablice pokazuju izračunate parametre koji opisuju model jednostavne linearne regresije.

$$\bar{x} = \frac{\sum x_i}{N} = \frac{119.022}{9} = 13.224{,}67$$

$$\bar{y} = \frac{\sum y_i}{N} = \frac{21.237}{9} = 2.359{,}67$$

Izračunati parametri a i b potrebni su za tumačenje regresijske jednadžbe.

$$b = \frac{\sum x_i y_i - n\bar{x}\bar{y}}{\sum x_i^2 - n\bar{x}^2} = \frac{282.155.520 - 9 \cdot 13.224{,}67 \cdot 2.359{,}67}{1.577.927.336 - 9 \cdot 13.224{,}67^2} = \frac{1.302.806{,}47}{3.900.266{,}52} = 0{,}334$$

$$a = \bar{y} - b\bar{x} = 2.359{,}67 - 0{,}334 \cdot 13.224{,}67 = -2.057{,}37$$

Linearna regresijska jednadžba glasi:

$$\hat{y} = a + bx$$

$$\hat{y} = -2.057{,}37 + 0{,}334x$$

T: a – konstantni član: nema smisleno tumačenje, budući da mu je predznak negativan.

b – regresijski koeficijent: ako se ukupna bruto proizvodnja električne energije poveća za 1 GWh tada će se izvoz električne energije povećati za 0,334 GWh.

Regresijske vrijednosti izračunate su tako da su u regresijsku jednadžbu uvrštene stvarne vrijednosti nezavisne varijable. Rezultati su prikazani u tablici.

T: Pri proizvodnji električne energije od 13.976GWh očekivani izvoz električne energije je 2.367,46GWh.

Rezidualna odstupanja izračunata se tako da su od stvarne varijable y_i oduzete pripadajuće regresijske vrijednosti $\hat{y}_i$.

T: Na razini proizvodnje 13.976 GWh električne energije prema regresiji izvoz iznosi 2.610,67 GWh, a zabilježen izvoz na toj razini proizvodnje, koja je ostvarena 2004. godine, iznosi 2.296 GWh. Prema regresiji za 2004. godinu precijenjen je izvoz električne energije za 314,67GWh ili za 2,25%.

$$u_{i,rel} = \frac{y_i - \hat{y}_i}{y_i} \cdot 100$$

$$u_{1,rel} = \frac{-314{,}67}{13.976} \cdot 100 = -2{,}25\%$$

T: Kao što prikazuju rezultati ispod tablice 12.2. vrijednosti pokazatelja varijabilnosti;varijanca, standardna devijacija *σ* i varijanca *V*, pokazuju izraženo prosječno i postotno odstupanje stvarnih vrijednosti od regresijskih.

T: Prosječno odstupanje empirijskih vrijednosti prometa od regresijskih vrijednosti iznosi 980,29 GWh ili u relativnom iznosu 41,54%. Varijabilnost podataka unutar modela je velika.

Jednadžba analize varijance:

$$9.084.100 = 435.400{,}41 + 8.648.699{,}51$$
$$SS_{yy} = SSR + SSE$$

Koeficijent determinacije iznosi $r^2 = \frac{SSR}{SS_{yy}} = \frac{435.400{,}41}{9.084.100} = 0{,}0479$

T: Vrijednost koeficijenta determinacije r^2koji je blizu 0, pokazuje da je proporcija protumačenog zbroja kvadrata u ukupnom zbroju kvadrata odstupanja mala. Zbog toga je model prikazan u tablici 12.3. **nereprezentativan.**

Koeficijent linearne korelacije $r = \sqrt{0{,}0479} = 0{,}2189$

ili

$$r = \frac{\sum xy - N\bar{x}\bar{y}}{\sqrt{\left(\sum x^2 - N\bar{x}^2\right)\cdot\left(\sum y^2 - N\bar{y}^2\right)}}$$

$$r = \frac{282.155.520 - 9\cdot 13.224{,}67\cdot 2.359{,}67}{\sqrt{\left(1.577.927.336 - 9\cdot 13.224{,}67^2\right)\cdot\left(59.196.341 - 9\cdot 2.359{,}67^2\right)}} \frac{282.155.520 - 280.852.713{,}53}{\sqrt{\left(1.577.927.336 - 1.574.027.069{,}48\right)\cdot\left(59.196.341 - 50.112.382{,}58\right)}}$$

$$r = \frac{1.302.806{,}47}{\sqrt{\left(3.900.266{,}52\right)\cdot\left(9.083.958{,}42\right)}} = 0{,}2189$$

ili

$$r = \frac{\frac{1.303.274}{9}}{\sqrt{\frac{3.901.060}{9}} \cdot \sqrt{\frac{9.084.100}{9}}} = \frac{144.808{,}22}{\sqrt{433.451{,}11} \cdot \sqrt{1.009.344{,}44}} = \frac{144.808{,}22}{658{,}37 \cdot 1.004{,}66} = 0{,}2189$$

T: Linearnom regresijskom vezom protumačeno je 21,89% svih odstupanja. Na temelju toga može se zaključiti da model nije reprezentativan. Izračunati Pearsonov koeficijent korelacije (0,2189)pokazuje pozitivnu, ali slabu korelaciju između varijable x i y, odnosno između ukupne bruto proizvodnje električne energije i izvoza električne energije.

Tablica 12.3. Finalna potrošnja električne energije i uvoz električne energije Republike Hrvatske od 2002. do 2010. godine u GWh

Godina	Finalna potrošnja el. energije u GWh	Uvoz el. energije u GWh				Regresijske vrijednosti
	x_i	y_i	x_i^2	$x_i y_i$	$y_i 2$	$\hat{y}_i = a + bx_i$
2002.	13.132	3.923	172.449.424	51.516.836	15.389.929	4.515,74
2003.	13.416	4.499	179.989.056	60.358.584	20.241.001	4.855,06
2004.	14.163	5.339	200.590.569	75.616.257	28.504.921	5.747,57
2005.	14.923	8.802	222.695.929	131.352.246	77.475.204	6.655,62
2006.	15.512	8.374	240.622.144	129.897.488	70.123.876	7.359,35
2007.	15.831	7.926	250.620.561	125.476.506	62.821.476	7.740,49
2008.	16.545	8.249	273.737.025	136.479.705	68.046.001	8.593,58
2009.	15.915	7.651	253.287.225	121.765.665	58.537.801	7.840,86
2010.	16.248	6.784	263.997.504	110.226.432	46.022.656	8.238,72
Ukupno	**135.685**	**61.547**	**2.057.989.437**	**942.689.719**	**447.162.865**	**61.547,00**
	Rezidualna odstupanja					
	$u_i = y - \hat{y}_i$	$x_i - \overline{x}$	$y_i - \bar{y}$	$(x_i - \overline{x})^2$	$(y_i - \bar{y})^2$	$(x_i - \overline{x}) \cdot (y_i - \bar{y})$
	-592,74	-1.944,11	-2.915,56	3.779.568,01	8.500.464,20	5.668.163,95
	-356,06	-1.660,11	-2.339,56	2.755.968,90	5.473.520,20	3.883.922,17
	-408,57	-913,11	-1.499,56	833.771,90	2.248.666,86	1.369.260,84
	2.146,38	-153,11	1.963,44	23.443,01	3.855.114,09	-300.625,16
	1.014,65	435,89	1.535,44	189.999,12	2.357.589,64	669.283,17
	185,51	754,89	1.087,44	569.857,23	1.182.535,42	820.899,73
	-344,58	1.468,89	1.410,44	2.157.634,57	1.989.353,53	2.071.786,17
	-189,86	838,89	812,44	703.734,57	660.065,98	681.550,62
	-1.454,72	1.171,89	-54,56	1.373.323,57	2.976,31	-63.933,05
Ukupno	0,00	0,00	0,00	12.387.300,89	26.270.286,22	14.800.308,44

Rezultati	
$\overline{x}$	15.076,11 GWh
$\bar{y}$	6.838,56 GWh
a	*-11.174,33*
b	*1,195* GWh
$\sigma_{\hat{y}}$	*976,78* GWh
$V_{\hat{y}}$	*14,28%*
r^2	*0,6731*
r	*0,8204*

Izvor: Statistički ljetopis Republike Hrvatske 2012., str. 303., URL

T: Kao što je prikazano u tablici niske vrijednosti pokazatelja varijabilnosti; standardna devijacija *σ* i varijanca *V*, pokazuju da je prosječno odstupanje (978,78 GWh) i postotno odstupanje (14,25%) stvarnih vrijednosti od regresijskih malo.

T: Vrijednost koeficijenta determinacije r^2koji je bliže 1 nego 0 pokazuje da je proporcija protumačenog zbroja kvadrata u ukupnom zbroju kvadrata odstupanja velika. Zbog toga je model u tablici 12.3. **reprezentativan**. Koeficijent determinacije 0,6731 pokazuje da je 67,31% varijacija količine uvoza električne energije rezultat varijacija finalne potrošnje, a 32,69% varijacija u količini uvoza rezultat je pogreške modela.

T: Pearsonov koeficijent korelacije (0,8204) pokazuje jaku i pozitivnu korelaciju između finalne potrošnja električne energije i uvoza električne energije.
Model regresije $\bar{y} = -11.174{,}33 + 1{,}195x$
pokazuje da s povećanjem 1 GWh finalne potrošnje električne energije uvoz električne energije raste za 1,195 GWh.

Primjer 12.2. Razdvajanje troškova metodom najmanjih kvadrata
Metoda najmanjih kvadrata često je korištena u poslovnoj praksi. Tako je prikladna za razdvajanje troškova na stalne (fiksne) i promjenjive (varijabilne).

U tablici 12.4. dani su podaci o proizvodnji u obiteljskom poljoprivrednom gospodarstvu (OPG) Gama i pripadajućim ukupnim troškovima na određenoj razini proizvodnje. Potrebno je odrediti koliki su stalni, a koliki promjenjivi troškovi te predvidjeti ukupni trošak pri promjeni opsega proizvodnje od 400 tona proizvoda. To omogućuje model jednostavne linearne regresije.

Tablica 12.4. Razina proizvodnje OPG-a Gama s pripadajućim ukupnim troškovima

Razdoblje	Proizvodnja Q, u t	Troškovi T, u 000 kn	Q*T	Q^2
	x_i	y_i	$x_i y_i$	x_i^2
1.	254	538	136.652	64.516
2.	278	583	162.074	77.284
3.	305	602	183.610	93.025
4.	362	725	262.450	131.044
5.	384	739	283.776	147.456
6.	395	804	317.580	156.025
7.	365	758	276.670	133.225
8.	310	641	198.710	96.100
9.	420	759	318.780	176.400
10.	361	684	246.924	130.321
Ukupno	3.434	6.833	2.387.226	1.205.396

Izvor: Radman-Funarić i Babler (2013)

Ukupni trošak u OPG-a Gama predstavlja zbroj ukupnog stalnog i ukupnog promjenjivog troška, a izračunava se pomoću jednadžbe pravca regresije, prema izrazu: $T = Ts + Tp$, gdje je:

T – ukupni trošak

Ts – stalni trošak

Tp – promjenjivi trošak

Stalni dio ukupnih troškova predstavlja odsječak na ordinati pravca regresije i izračunava se prema izrazu:

$$Ts = \frac{\sum_{i=1}^{N} x_i^2 \sum_{i=1}^{N} y_i - \sum_{i=1}^{N} x_i \sum_{i=1}^{N} x_i y_i}{N \sum_{i=1}^{N} x_i^2 - (\sum_{i=1}^{N} y_i)^2} = a$$

Promjenjivi (prosječni proporcionalni) trošak (*tp*) predstavlja nagib, odnosno koeficijent smjera pravca regresije i izražava se prema izrazu:

$$tp = \frac{N \sum_{i=1}^{N} x_i y_i - \left(\sum_{i=1}^{N} x_i\right)\left(\sum_{i=1}^{N} y_i\right)}{N \sum_{i=1}^{N} x_i^2 - \left(\sum_{i=1}^{N} x_i\right)^2} = b$$

U radnoj tablici 12.4. prikazan je postupak razdvajanja troškova na temelju podataka o opsegu proizvodnje i troškovima. Izračunat je umnožak opsega proizvodnje i troškova za sva razdoblja te je kvadriran opseg proizvodnje svakog razdoblja. Dijagramom rasipanja prikazanom u grafikonu 12.1., utvrđeno je da između opsega proizvodnje i utvrđenih troškova postoji linearna pozitivna korelacija.

Ukupni stalni trošak iznosi 148.074 kuna, prema:

$$Ts = \frac{1.205.396 * 6.833 - 3.434 * 2.387.226}{10 * 1.205.396 - 3.434^2} = \frac{38.736.784}{261.604} = 148{,}074$$

Ukupni promjenjivi trošak je razlika između ukupnog troška i ukupnog stalnog troška: $Tp = T - Ts$

Prosjek ukupnog troška za navedena razdoblja izračunava se dijeljenjem ukupnog troška svih razdoblja s brojem promatranih razdoblja i iznosi 683.300 kn ($ø_T$= 6.833.000 kn/10).

$$Tp = 683.300 - 148.074 = 535.226 kn$$

Prosječni opseg proizvodnje izračunava se dijeljenjem opsega proizvodnje Q s brojem promatranih razdoblja ($ø_Q$ = 3434 tona/10 promatranih razdoblja) te iznosi 343,4 tone

Prosječni promjenjivi trošak za navedena razdoblja izračunava se dijeljenjem ukupnog promjenjivog troška s prosječnim opsegom proizvodnje ($ø_{Tp}$ = 535.226 kn/343,4 tone) i iznosi 1.558,61 kn po toni.

Isti iznos prosječnog promjenjivog troška dobije se prema prikazanom izrazu:

$$tp = \frac{10 \cdot 2.387.226 - 3.434 \cdot 6.833}{10 \cdot 1.205.396 - 3434^2} = \frac{407.738}{261.604} = 1{,}55861$$

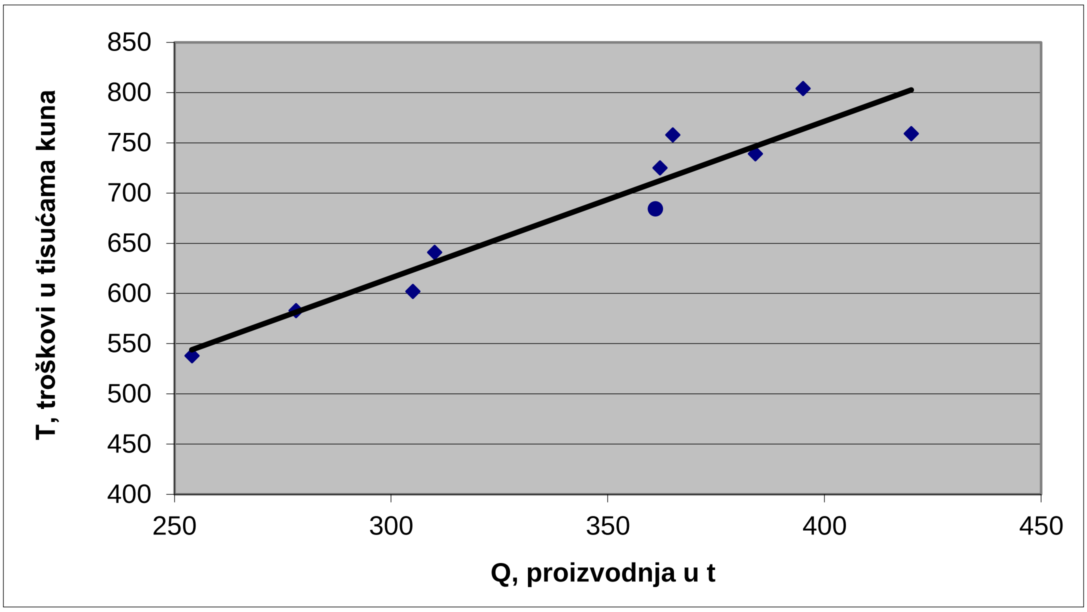
Dobiveni podaci o stalnom i prosječnom promjenjivom trošku omogućuju postavljanje opće jednadžbe ukupnih troškova u OPG-u Gama koja glasi:

$$T = 148.074 + 1.558{,}61x$$

Jednadžba omogućuje predviđanje ukupnog troška pri promjeni opsega proizvodnje. Tako se predviđa da će pri razini proizvodnje od 400 tona proizvoda, ukupni trošak iznositi771.673,19 kn.

$$T = 148.229{,}19 + 1.558{,}61 * 400 = 771.673{,}19 kn$$

Grafikon 12.1. Razina proizvodnje OPG-a Gama s pripadajućim ukupnim troškovima



Izvor: Radman-Funarić i Babler (2013)

12.2. Model višestruke linearne regresije

Modelom višestruke regresije (multiple regresije)izražava se ovisnost jedne varijable o više drugih varijabli.

Analiza modela višestruke linearne regresije
Prvi korak u analizi modela višestruke linearne regresije sastoji se u utvrđivanju njegova oblika te svojstava varijabli i parametara, a zatim se izračunavaju parametri koji omogućuju tumačenje modela.

Opći linearni regresijski model osnovnog skupa (populacije) ima sljedeći oblik:

$$y_i = \alpha + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_j X_j + \ldots + \beta_K X_K + e$$

Za izračunavanje parametra najčešće se koristi metoda najmanjeg kvadrata.
α – konstantni član – predstavlja vrijednost regresije kada su nezavisne varijable jednake 0.
β_j – regresijski koeficijent – pokazuje za koliko se u prosjeku mijenja zavisna varijabla ako se nezavisna varijabla x_j poveća za jedan, uz uvjet da ostale nezavisne varijable ostanu nepromijenjene.
e – slučajna varijabla – pogreška relacije

Varijanca, standardna devijacija i koeficijent varijacije omogućuju prosudbu stupnja disperzije oko regresijske ravnine. U sklopu deskriptivne statistike dobivaju se pomoću izraza:

$$\sigma_{\hat{y}}^{\;2} = \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{n}$$

$$\sigma_{\hat{Y}} = \sqrt{\frac{\sum_{i=1}^{n} y_i^2 - \alpha \sum_{i=1}^{n} y_i - \beta \sum_{i=1}^{n} x_{i1} y_i - \ldots - \beta_k \sum_{i=1}^{n} x_{iK} y_i}{n}} \qquad \text{ili} \qquad \sigma_{\hat{y}} = \sqrt{\sigma_{\hat{y}}^2}$$

$$V_{\hat{y}} = \frac{\sigma_{\hat{y}}}{\overline{y}} \cdot 100$$

Primjer 12.3. Izračunavanje jednadžbe višestruke linearne regresije

Model višestruke linearne regresije proveden je kako bi se provjerilo utječe li ukupna bruto proizvodnja i finalna potrošnja električne energije, kao nezavisne varijable, na uvoz električne energije kao zavisnu varijablu. Za analizu su poslužili podaci iz tablice 12.5.

Tablica 12.5. Ukupna bruto proizvodnja, finalna potrošnja i uvoz električne energije Republike Hrvatske od 2002. do 2010. godine u GWh

Godina	Ukupna bruto proizvodnja električne energije u GWh	Finalna potrošnja električne energije u GWh	Uvoz električne energije u GWh
	x_i	x_i	y_i
2002.	12.725	13.132	3.923
2003.	13.248	13.416	4.499
2004.	13.976	14.163	5.339
2005.	13.140	14.923	8.802
2006.	13.037	15.512	8.374
2007.	12.462	15.831	7.926
2008.	12.616	16.545	8.249
2009.	13.149	15.915	7.651
2010.	14.669	16.248	6.784

Izvor: Statistički ljetopis Republike Hrvatske 2012., Državni zavod za statistiku, str. 303.

Pri analizi polazi se od linearnog regresijskog modela osnovnog skupa za n vrijednosti:

$$y_i = \alpha + \beta_1 x_{i1} + \beta_2 x_{i2} + ...\beta_j x_{ij} + ... + \beta_K x_{iK} + e_i, i = 1,2,...., n$$

Procjene parametara a, β_1 i β_2 rade se metodom najmanjih kvadrata, to jest procjena parametara rješenje je sustava od tri linearne jednadžbe s tri nepoznanice, budući da su u analizi dvije nezavisne varijable, K=2.

Rješenjem jednadžbe dobiveni su ovi parametri:

a=3.240,05 β_1 = -0,6197 β_2=1,2216

Regresijski model s izračunatim parametrima glasi:

$$\hat{y} = 3.240{,}05 - 0{,}6197x_1 + 1{,}2216x_2$$

Regresijska jednadžba, a time i procjene parametara isti su analiziraju li se podaci sa stajališta deskriptivne ili inferencijalne statistike.

T: Dobiveni rezultat pokazuje da se uvoz električne energije smanjuje za 0,6197 GWh, ako se ukupna bruto proizvodnja električne energije u Hrvatskoj poveća za 1Gwh, uz nepromijenjenu finalnu potrošnju. Uvoz električne energije povećava se za 1,2216 GWh, ako se finalna potrošnja električne energije poveća za 1GWh, uz nepromijenjenu proizvodnju.
Konstantni član je 3.240,05 i nema smisleno tumačenje.

13. VREMENSKI NIZOVI

Vremenski niz je skup kronološki uređenih istovrsnih podataka. Podaci u vremenskom nizu analiziraju se u uzastopnim vremenskim intervalima ili istim vremenskim razmacima između trenutaka u kojima je pojava mjerena. Vremenski niz nastaje grupiranjem statističkih jedinica prema vremenskom obilježju. Namjena analize vremenskih nizova je promatrati vremenski razvoj pojava, tražiti njihove zakonitosti i predviđati njihov razvoj.

U analizi vremenskih nizova vrijeme je nezavisna varijabla x, a veličina promatrane pojave zavisna varijabla y. Zadatak dinamičke analize je ispitati promjene pojava kao funkciju vremena, odnosno otkriti pravilnosti i zakonitosti koje se očituju u varijaciji pojava tijekom vremena.

$$y = f(t)$$

U vremenskom nizu statističke jedinice označavaju se s Y_i, $i = 1,\ldots, N$. Vremenska jedinica može biti: 1 godina, 1 sekunda, 1 dan, 1 sat, 1 petogodišnje razdoblje. Razdoblje je širi pojam od vremenske jedinice i predstavlja skup vremenskih jedinica, npr. razdoblje od 2000. do 2005. godine odnosi se na 2000., 2001., do uključivo 2005. godine.

Frekvencije vremenskog niza nužno moraju biti usporedive, što znači da se tijekom čitavog promatranog razdoblja pojmovna i prostorna definicija ne smije mijenjati.

Vremenski nizovi mogu biti intervalni i trenutačni.

Intervalni vremenski niz nastaje zbrajanjem vrijednosti pojave tijekom promatranog vremenskog intervala, npr. godišnji prihod poduzeća, godišnja količina proizvedenih proizvoda, mjesečne obveze poreza, ostvarena dobit po kvartalima, dnevni promet u maloprodaji. Analiza vremenskih nizova moguća je jedino uz jednake vremenske intervale promatranja. Ako su vremenska razdoblja različita, potrebno je korigirati frekvencije kako bi nastali podaci za jednaka razdoblja.

Intervalni vremenski niz ima svojstvo kumulativnosti, jer svaki kumulativ u kumulativnom nizu ima smisleno značenje.

Trenutačni vremenski niz nastaje kronološkim uređivanjem vrijednosti koje predstavljaju stanja pojave u odabranim vremenskim trenutcima, npr. stanje novčanih sredstava na deviznom računu, broj nezaposlenih na dan 30. lipnja 2017. godine, cijena dionice na kraju dana. Trenutačni vremenski niz nema svojstvo kumulativnosti, jer kumulativ u kumulativnom nizu nema smisleno značenje, odnosno zbroj frekvencija mjerenih u određenim trenucima ne predstavlja vrijednost koju je moguće objasniti.

13.1.Grafičko prikazivanje vremenskih nizova

Intervalni nizovi prikazuju se *površinskim i linijskim grafikonima*, a trenutačni nizovi samo *linijskim grafikonima.* Prikazuju se u pravokutnom koordinatnom sustavu s aritmetičkim mjerilima na osima.

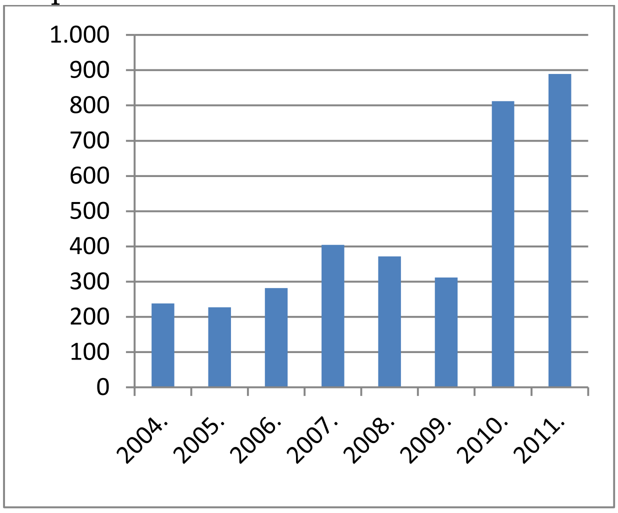
Na apscisu se unosi mjerilo za varijablu vremena, a na ordinatu frekvencije vremenskog niza.
Ako se ne raspolaže podacima za dio razdoblja, kod linijskih grafikona je radi preglednosti moguće izostaviti dio mjerila na osi apscisa (okomiti prekid grafikona) ili na osi ordinata (vodoravni prekid grafikona).

Površinski grafikon je vrsta grafičkog prikaza u kojem se stupac podiže iznad baze određene vremenske jedinice do visine koja je određena frekvencijom vremenskog niza.

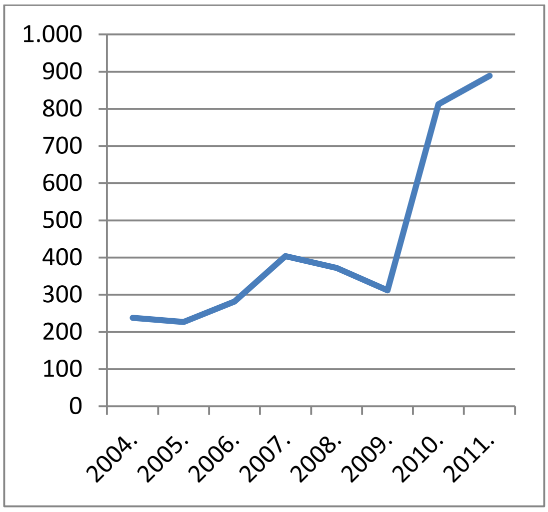
Linijski grafikon je vrsta grafičkog prikaza koji predstavlja liniju dobivenu spajanjem točaka pri čemu je svaka točka na grafikonu podignuta nad sredinu vremenske jedinice (ako je vremenski niz intervalni), odnosno iznad onog mjesta na apscisi koje se odnosi na trenutak kada je pojava snimljena (ako je vremenski niz trenutačni), do visine koja je određena frekvencijom vremenskog niza.

Primjer 13.1. Prikaz intervalnog vremenskog niza površinskim i linijskim grafikonom

Grafikon 13.1. Površinski grafikon – dijagram jednostavnih uspravnih stupaca



Grafikon 13.2. Linijski grafikon

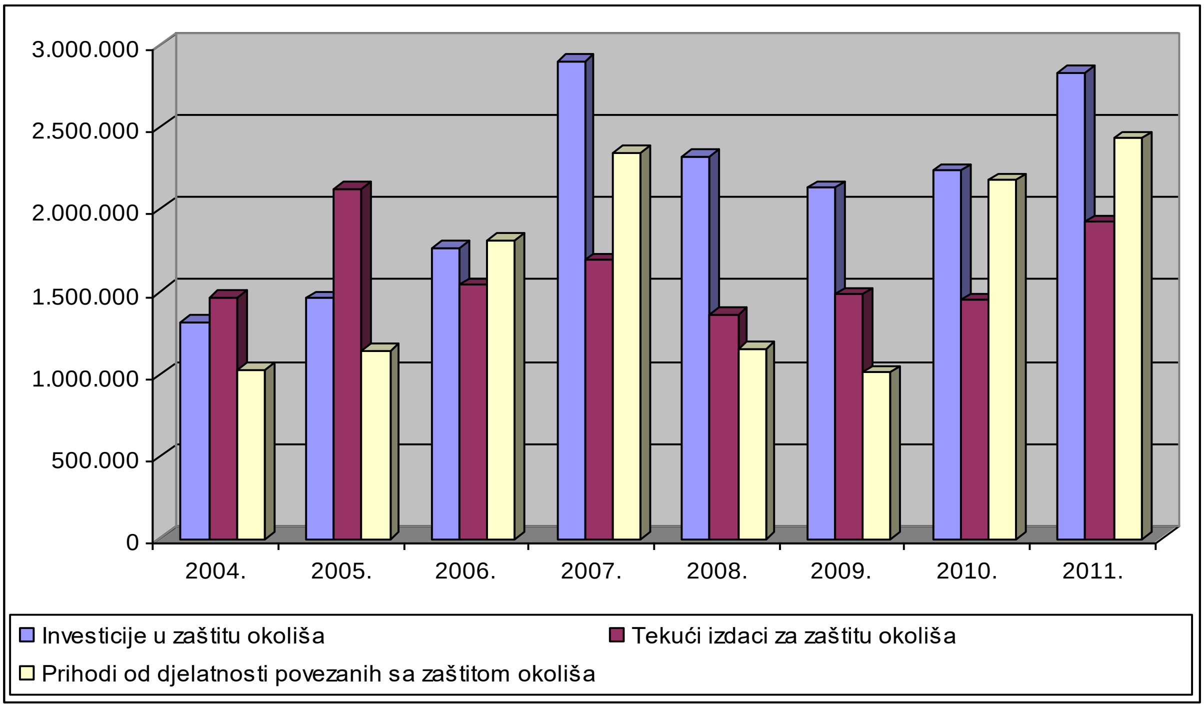


Polulogaritamski grafikon koristi se ako se na istom grafikonu uspoređuju raznorodni podaci (nizovi izraženi u različitim mjernim jedinicama). To je grafikon sa aritmetičkim mjerilom na osi apscisa, a logaritamskim na osi ordinata.

Pri grafičkoj usporedbi više vremenskih nizova na istom grafikonu, uglavnom se koriste zajedničke koordinatne osi koje se moraju odnositi na ista vremenska razdoblja i biti izražena u istim mjernim jedinicama na približno istoj vrijednosnoj razini. Površinski grafikoni koji se koriste za prikazivanje više vremenskih nizova su **grafikon višestrukih stupaca** (grafikon 13.3.) i **grafikon strukturnih stupaca** (grafikon 13.4.).

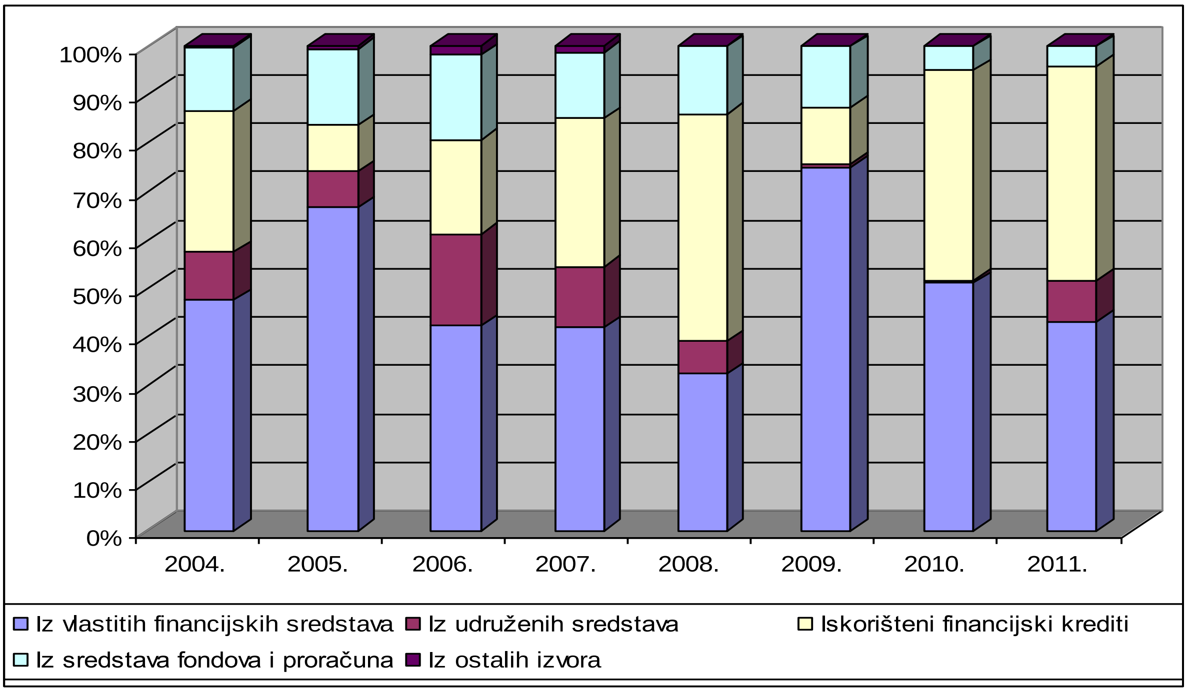
Primjer 13.2. Grafikon višestrukih stupaca i grafikon strukturnih stupaca kojima su prikazani vremenski nizovi

Grafikon 13.3. Investicije u zaštitu okoliša, tekući izdaci za zaštitu okoliša te prihodi od djelatnosti povezanih sa zaštitom okoliša, u razdoblju od 2004. do 2011. godine, u tisućama kuna



Izvor: Investicije u zaštitu okoliša, *Priopćenje*, Državni zavod za statistiku, 2004., 2005., 2006., 2007., 2008., 2009., 2010. i 2011.

Grafikon 13.4. Struktura izvora sredstava za investicije u integrirane tehnologije, u %

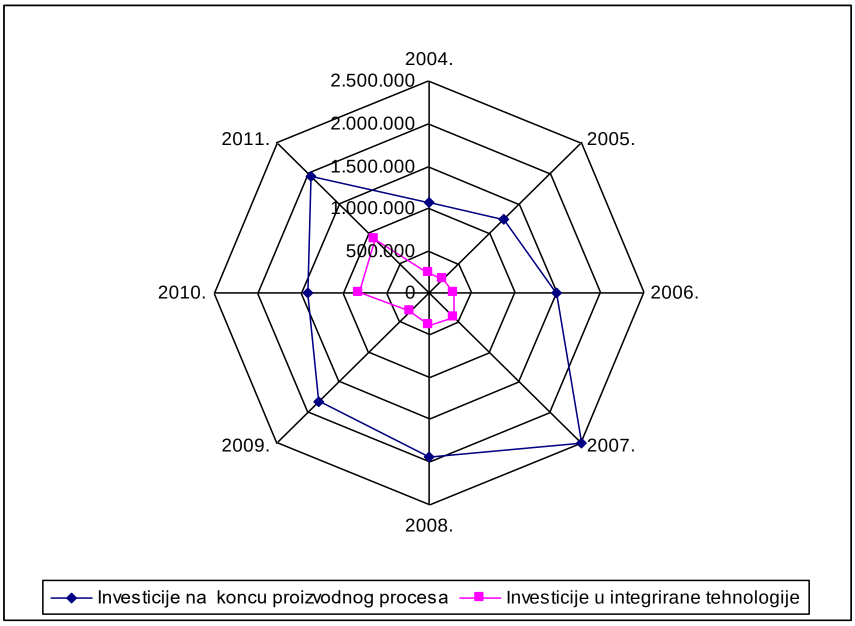


Investicije u zaštitu okoliša, 2004., 2005., 2006., 2007., 2008., 2009., 2010. i 2011., prema Radman-Funarić i Kurtagić (2013)

Vremenski niz može se prikladno prikazati i polarnim dijagramom. Polarni dijagram ima specifičan oblik i podsjeća na paukovu mrežu. Naziva se i radar-dijagram (engl. *radar chart*). To je vrsta linijskoga dijagrama, gdje se krivulja ne ucrtava u pravokutni, već u polarni koordinatni sustav (grafikon 13.5)

Primjer 13.3. Vremenski niz prikazan polarnim dijagramom

Grafikon 13.5. Iznos sredstava za investicije u razdoblju od 2004. do 2011. godine, u tisućama kuna



Investicije u zaštitu okoliša, 2004., 2005., 2006., 2007., 2008., 2009., 2010. i 2011., prema Radman-Funarić i Kurtagić (2013)

13.2. Pokazatelji dinamike vremenskih nizova

Pokazatelji dinamike vremenskih nizova dijele se na apsolutne i relativne. Apsolutnim pokazateljima analizira se dinamika istorodnih vremenskih nizova, a relativnim pokazateljima dinamika raznorodnih vremenskih nizova.

13.2.1. Apsolutni pokazatelji dinamike vremenskih nizova

Pojedinačna apsolutna stopa promjene pokazuje apsolutnu vrijednosnu promjenu pojave u odnosu na prethodno razdoblje ili trenutak, ili u odnosu na definirano bazno razdoblje ili trenutak.

Ukoliko se utvrđuje apsolutna promjena pojave u odnosu na prethodno razdoblje, računa se prema izrazu:

$$\Delta y_t = y_t - y_{t-1} \qquad t = 1,2,\ldots,n$$

Δ– operator diferencije

y – vremenski niz

t – vremensko razdoblje

y_t – pojava promatran u t razdoblju

Moguće je izračunati n-1 pojedinačnih apsolutnih mjera promjene.

Ukoliko se utvrđuje apsolutna promjena pojave u odnosu na bazno razdoblje, računa se prema izrazu:

$$\Delta y_t = y_t - y_b \qquad t = 1,2,...,n$$

b – bazno vremensko razdoblje

y_b – pojava promatran u baznom razdoblju

Bazno razdoblje moguće je izabrati proizvoljno. Vrijednosti u vremenskom nizu moguće je usporediti i s prosječnom vrijednosti ili prognoziranom vrijednosti niza.

Primjer 13.4. Izračunavanje apsolutnih stopa promjene

Tablica 13.1. Investicije u zaštitu okoliša u razdoblju od 2004. do 2011. godine

Godina	Investicije u zaštitu okoliša (u 000 kuna)	Pojedinačna apsolutna stopa promjene Δyt	
x	y	y_t-y_{t-1}	y_t-$y_ø$ 1)
2004.	1.311.648	–	-806.826
2005.	1.460.596	148.948	-657.878
2006.	1.766.631	306.035	-351.843
2007.	2.900.798	1.134.167	782.324
2008.	2.316.500	-584.298	198.026
2009.	2.130.107	-186.393	11.633
2010.	2.232.283	102.176	113.809
2011.	2.829.231	596.948	710.757

Investicije u zaštitu okoliša, 2004., 2005., 2006., 2007., 2008., 2009., 2010. i 2011., prema Radman-Funarić i Kurtagić (2013)

1) $y_ø$ – prosječne investicije u zaštitu okoliša u razdoblju od 2004. do 2005. godine iznose 2.118.474.000 kuna.

T: Investicije u zaštitu okoliša 2005. godine veće su za 148.948.000 kuna u odnosu na prethodno razdoblje, odnosno u odnosu na 2004. godinu.

T: Investicije u zaštitu okoliša 2005. godine manje su za 657.878.000 kuna u odnosu na prosječne investicije u zaštitu okoliša u razdoblju od 2004. do 2011. godine.

13.2.2. Relativni pokazatelji dinamike vremenskih nizova

Relativni pokazatelji dinamike primjenjuju se u analizi raznorodnih vremenskih nizova. Relativni pokazatelji dinamike su koeficijent dinamike, verižni indeks, pojedinačna stopa promjene te prosječna stopa promjene.

13.2.2.1. Koeficijent dinamike

Koeficijent dinamike vremenskog niza je kvocijent uzastopnih frekvencija promatrane pojave:

$$v_t = \frac{y_t}{y_{t-1}}$$

Koeficijent dinamike uvijek je pozitivan broj. Ukoliko je koeficijent dinamike veći od 1, pokazuje povećanje pojave u odnosu na prethodno razdoblje, a ukoliko je manji od 1, pokazuje smanjenje pojave.

Primjer 13.5. Izračun koeficijenta dinamike

Tablica 13.2. Tekući izdaci za zaštitu okoliša u razdoblju od 2004. do 2011. godine, u milijunima kuna

Godina	2004.	2005.	2006.	2007.	2008.	2009.	2010.	2011.
Tekući izdaci za zaštitu okoliša	1.467	2.123	1.539	1.689	1.363	1.483	1.447	1.921
v_t	-	1,447	0,725	1,097	0,807	1,088	0,976	1,328

Investicije u zaštitu okoliša, 2004., 2005., 2006., 2007., 2008., 2009., 2010. i 2011., prema Radman-Funarić i Kurtagić (2013)

13.2.2.2. Pojedinačna stopa promjene

Pojedinačna stopa promjene primjenjuje se u analizi raznorodnih vremenskih nizova, a izračunava se pomoću izraza:

$$S_t = \frac{y_t - y_{t-1}}{y_{t-1}} \cdot 100 \qquad S_t = \frac{\Delta y_t}{y_{t-1}} \cdot 100 \qquad S_t = I_t - 100$$

ili ili

Kao i pojedinačna apsolutna stopa promjene,pojedinačna stopa promjene može biti jednaka nuli, manja ili veća od nule. Prikazuje se stupcima (ili linijskim grafikonom) pri čemu visine stupaca predstavljaju relativnu promjenu prema prethodnom razdoblju.

Primjer 13.6. Izračun pojedinačne stope promjene

Tablica 13.3. Tekući izdatci za zaštitu okoliša u razdoblju od 2004. do 2011. godine, u milijunima kuna

Godina	2004.	2005.	2006.	2007.	2008.	2009.	2010.	2011.
Tekući izdatci za zaštitu okoliša	1.467	2.123	1.539	1.689	1.363	1.483	1.447	1.921
Δv_t	-	656	-584	150	-326	120	-36	1,328
s_t		44,72	-27,51	9,75	-19,30	8,80	-2,43	0,09

Investicije u zaštitu okoliša, 2004., 2005., 2006., 2007., 2008., 2009., 2010. i 2011., prema Radman-Funarić i Kurtagić (2013)

T: Tekući izdaci za zaštitu okoliša u 2006. godini manji su za 27,51% u odnosu na izdatke u 2005. godini, a 2007. veći za 9,75% u odnosu na izdatke u 2006. godini.

13.2.2.3. Prosječna stopa promjene

Prosječna stopa promjene izračunava se pomoću **geometrijske sredine** koeficijenta dinamike. Koeficijenti dinamike su uvijek pozivni brojevi, za razliku od pojedinačnih stopa promjene. Računa se pomoću izraza:

$$G = \sqrt[n-1]{v_2 \cdot v_3 \cdot .... \cdot v_n} \quad G = \sqrt[n-1]{\frac{y_2}{y_1} \cdot \frac{y_3}{y_2} \cdot ... \cdot \frac{y_n}{y_{n-1}}} = \sqrt[n-1]{\frac{y_n}{y_1}} \quad \log G = \frac{\log y_n - \log y_1}{n-1}$$

Prosječna stopa promjene izračunava se pomoću izraza:

$$\bar{s} = (G-1) \cdot 100 \qquad \bar{s} = \left(\sqrt[n-1]{\frac{y_n}{y_1}} - 1 \right) \cdot 100$$

Umjesto originalnih frekvencija vremenskog niza moguće je prosječnu stopu promjene izračunati pomoću indeksa na stalnoj bazi, neovisno o baznom razdoblju za koje su izračunati, jer su bazni indeksi upravno proporcionalni originalnim frekvencijama vremenskog niza.

Primjer 13.7. Izračunavanje prosječne stope promjene prema podacima iz tablice 13.3.

$$G = \left(\sqrt[8-1]{\frac{1921}{1467}} = 1{,}03927 \right)$$

$$\bar{s} = \left(\sqrt[8-1]{\frac{1921}{1467}} - 1 \right) \cdot 100 = \left(\sqrt[7]{1{,}309475} - 1 \right) * 100 = \left(1{,}03927 - 1 \right) \cdot 100 = 3{,}93\%$$

T: Tekući izdaci za zaštitu okoliša u razdoblju od 2004. do 2011. godine u prosjeku su se povećavali 3,93% godišnje.

U ovom primjeru uočljiv je nedostatak prosječne stope promjene. Prosječna stopa promjene je pogodna za primjenu u vremenskim nizovima koji nemaju oscilacije u frekvencijama tijekom analiziranog vremenskog razdoblja, što u ovom primjeru nije slučaj.

Pomoću geometrijske sredine moguće je prognozirati razvoj pojave u budućim razdobljima za nizove koji se ponašaju po geometrijskoj progresiji (npr. globalni porast stanovništva) ili digresiji (npr. regionalni pad broja stanovnika).

Geometrijsku sredinu je potrebno izračunati kao u prethodnom primjeru koristeći originalne vrijednosti vremenskog niza. Nakon toga se izračunava razina pojave u budućem razdoblju, uz pretpostavku nepromijenjenih uvjeta, prema izrazu:

$$x = \frac{\log y_p}{\log G}$$

x – razdoblje za koje se računa pretpostavljena veličina pojave y_p

y_p – pretpostavljena veličina pojave

Primjer 13.8. Prognoziranje kada će se tekući izdaci za zaštitu okoliša u Hrvatskoj utrostručiti

$$\log G = \log 1{,}03927 = 0{,}016728$$

ili

$$\log G = \frac{\log y_n - \log y_1}{n-1} = \frac{\log 1921 - \log 1467}{7} = \frac{3{,}2835 - 3{,}1664}{7} = 0{,}0167$$

$$x = \frac{\log 3}{0{,}0167} = \frac{0{,}4771}{0{,}0167} = 28{,}57$$

T: Ako zanemarimo da je u kretanju tekućih izdataka za zaštitu okoliša u razdoblju od 2004. do 2011. godine bilo oscilacija, možemo pretpostaviti da će se oni utrostručiti za 28,57 godina. Budući da je prognostički podatak izračunat prema podacima do 2011. godine to utrostručenje može se očekivati 2039. godine.

Primjer 13.9. Prognoziranje vrijednosti tekućih izdataka za 28,57 godina

Ukoliko se prognozira vrijednost pojave u nekoj vremenskoj točki u budućnosti koristi se izraz:

$$F_{n+\tau} = y_n \cdot G^{\tau}$$

$F_{n+\tau}$ – očekivana vrijednost pojave nakon τ razdoblja
τ (čit. tau) – broj razdoblja ili vremenskih točaka unaprijed za koje se prognozira vrijednost pojave
n – duljina vremenskog niza

$$F_{2011.+28{,}57g} = 1.921 \cdot 1{,}03927^{28{,}57} = 1.921 \cdot 3{,}00559 = 5.773{,}74$$

T: Ako se vrijednost tekućih izdataka za zaštitu okoliša nastave kretati po istoj prosječnoj stopi, za 28,57 godina (2039. godine)iznositi će 5.773,74 milijuna kuna. Kao što je opisano u primjeru 13.8. vrijednost izdataka će se utrostručiti (1.921 · 3 = 5.763) za 28,57 godina. [19]

$$F_{2011.+10} = 1.921 \cdot 1{,}03927^{10} = 1.921 \cdot 1{,}46989 = 2.823{,}66$$

T: Ako se vrijednost tekućih izdataka za zaštitu okoliša nastave kretati po istoj prosječnoj stopi, za 10 godina (2021. godine) iznositi će 2.823,66 milijuna kuna.

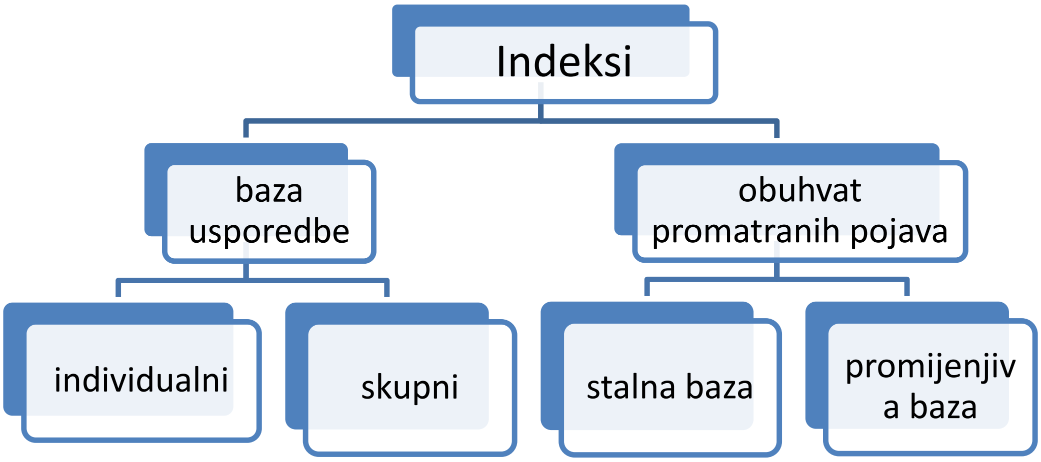
13.3. Indeksi

Dinamika promjene gospodarskih i društvenih pojava često se prikazuje indeksima. Tako se i u javnim publikacijama kretanje gospodarskih pojava najčešće prikazuje indeksima. **Indeksi** su relativni brojevi dinamike koji pokazuju odnos između stanja jedne te iste pojave ili skupine pojava na različitim mjestima ili u različitim vremenskim razdobljima. Pokazuju koliko jedinica pojave u jednom razdoblju dolazi na 100 jedinica pojave u drugom razdoblju.

Indeksi ne mogu biti negativni brojevi. Dijele se s obzirom na obuhvat podataka i s obzirom na bazu usporedbe, kako je prikazano u slici 12.1.

[19] Razlika u iznosu od 13,74 (5.773,74 – 5.763) nastala je kao posljedica korištenja različitih metoda. Budući da razliku predstavlja 13,74 milijuna kuna, razlika nije zanemariva.

Slika 13.1. Vrste indeksa



13.3.1. Individualni indeksi

Individualni indeksi prate dinamiku jedne pojave u više uzastopnih vremenskih razdoblja.

13.3.2. Bazni indeksi (indeksi na stalnoj bazi)

Baznim indeksima izražava se dinamika jedne pojave, pri čemu je jedno od promatranih razdoblja baza usporedbe. Nazivaju se i indeksi na stalnoj bazi. Kao bazno razdoblje preporučuje se ono koje nema ekstremne vrijednosti. Kao baza usporedbe mogu se koristiti i vrijednosti izvan niza ili prosjek vrijednosti niza.
Zbog same definicije indeksa najčešće se označava sa 100, npr. 2015. = 100. Izračunavaju se prema sljedećem izrazu:

$$I_t = \frac{y_t}{y_b} \cdot 100$$

I_t– individualni bazni indeks

y_b –vrijednost pojave u baznom razdoblju

Pri tomu vrijede relacije:

$Y_t > Y_b \rightarrow I_t > 100$

$Y_t < Y_b \rightarrow I_t < 100$

$Y_t = Y_b \rightarrow I_t = 100$

Iz izračunatih indeksa radi lakšeg tumačenja, osobito radi lakšeg objašnjavanja vrijednosti indeksa široj javnosti, koristi se pojedinačna stopa promjene (opisano u potpoglavlju 13.2.2.2.). Računa se prema izrazu:

$$S_t = I_t - 100$$

Primjer 13.10. Izračun baznih indeksa i stopa promjene

Tablica 13.4. Kretanje investicija u zaštitu okoliša od 2004. do 2011. godine

	Investicije u zaštitu okoliša		
Godina	Originalne vrijednosti y_t(u 000 000 kn)	Bazni indeks I_t	Stopa promjene u % s_t
2004.	1.312	100	–
2005.	1.461	111,36	11,36
2006.	1.767	134,68	34,68
2007.	2.901	221,11	121,11
2008.	2.317	176,60	76,60
2009.	2.130	162,35	62,35
2010.	2.232	170,12	70,12
2011.	2.829	215,63	115,63

Izvor: Investicije u zaštitu okoliša, 2004., 2005., 2006., 2007., 2008., 2009., 2010. i 2011., prema Radman-Funarić i Kurtagić (2013)

$$I_{06} = \frac{1.767}{1.312} \cdot 100 = 134{,}68 \qquad I_{11} = \frac{2.829}{1.312} \cdot 100 = 215{,}63$$

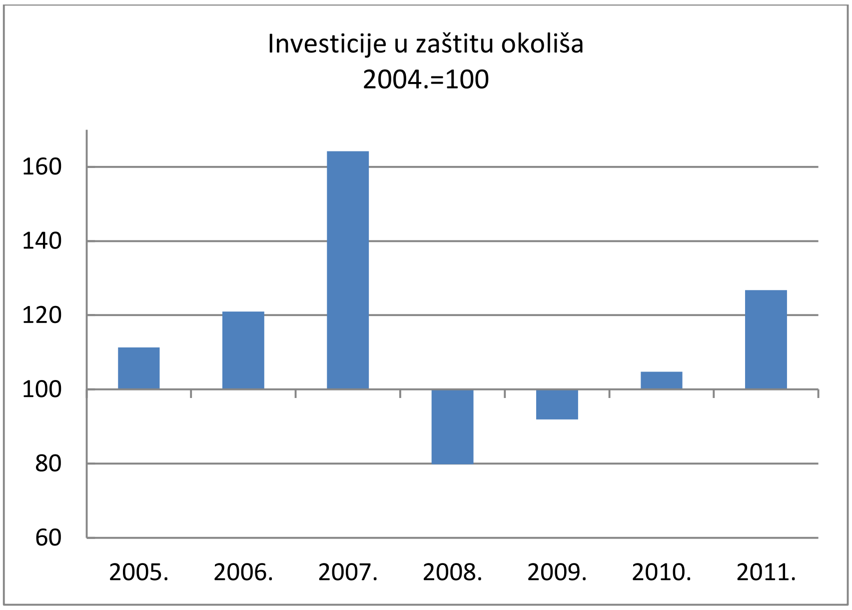
T: Investicije u zaštitu okoliša 2006. godine bile su veće za 34,68% nego li što su bile 2004. godine. Na svakih investiranih 100 milijuna kuna u zaštitu okoliša u 2004. godini, 2006. godine je uloženo134,68 milijuna kuna.

Grafički se bazni indeksi mogu prikazati površinskim i linijskim grafikonom[20].

[20] Grafičko prikazivanje baznih indeksa, vidi: Horvat i Mijoč (2014), str. 556-558.

Primjer 13.11. Površinski grafikon baznih indeksa

Grafikon 13.6. Kretanje investicija u zaštitu okoliša od 2004. do 2011. godine – bazni indeksi



Izvor: Investicije u zaštitu okoliša, 2004., 2005., 2006., 2007., 2008., 2009., 2010. i 2011., prema Radman-Funarić i Kurtagić (2013)

13.3.3. Verižni indeksi (indeksi na promjenjivoj bazi)

Verižni indeksi nazivaju se i indeksi na promjenjivoj bazi ili lančani indeksi jer im se baze mijenjaju iz razdoblja u razdoblje. Prate promjenu stanja pojave u uzastopnim razdobljima, to jest prilikom izračuna verižnih indeksa promatrane vrijednosti neke pojave stavljaju se u odnos prema vrijednostima pojave iz prethodnog razdoblja. Izračunavaju se prema izrazu:

$$V_t = \frac{y_t}{y_{t-1}} \cdot 100$$

Pri tomu vrijede relacije:

$Y_t > Y_{t-1} \rightarrow V_t > 100$

$Y_t < Y_{t-1} \rightarrow V_t < 100$

$Y_t = Y_{t-1} \rightarrow V_t = 100$

Primjer 13.12. Izračun verižnih indeksa i stopa promjene

Tablica 13.5. Kretanje investicija u zaštitu okoliša, tekućih izdataka za zaštitu okoliša i prihoda od djelatnosti povezanih sa zaštitom okoliša od 2004. do 2011. godine, u milijunima kuna

	Investicije u zaštitu okoliša			Tekući izdatci za zaštitu okoliša			Prihodi od djelatnosti povezanih s zaštitom okoliša		
Godina	Originalne vrijednosti y_t	Verižni indeks V_t	S_t u %	Originalne vrijednosti y_t	Verižni indeks V_t	S_t u %	Originalne vrijednosti y_t	Verižni indeks V_t	S_t u %
2004.	1.312	–	–	1.467	–	–	1.030	–	–
2005.	1.461	111,36	11,36	2.123	144,72	44,72	1.144	111,07	11,07
2006.	1.767	120,94	20,94	1.539	72,49	-27,51	1.809	158,13	58,13
2007.	2.901	164,18	64,18	1.689	109,75	9,75	2.340	129,35	29,35
2008.	2.317	79,87	-20,13	1.363	80,70	-19,30	1.153	49,27	-50,73
2009.	2.130	91,93	-8,07	1.483	108,80	8,80	1.010	87,60	-12,40
2010.	2.232	104,79	4,79	1.447	97,57	-2,43	2.175	215,35	115,35
2011.	2.829	126,75	26,75	1.921	132,76	32,76	2.428	111,63	11,63

Izvor: Investicije u zaštitu okoliša, 2004., 2005., 2006., 2007., 2008., 2009., 2010. i 2011., prema Radman-Funarić i Kurtagić (2013)

$$V_6 = \frac{1.767}{1.461} \cdot 100 = 120{,}94 \qquad V_{11} = \frac{2.829}{2.232} \cdot 100 = 126{,}75$$

T: Investicije u zaštitu okoliša 2011. godine bile su veće za 26,75% nego li što su bile prethodne 2010. godine. Na svakih investiranih 100 milijuna kuna u zaštitu okoliša 2010. godine, 2011. godine je uloženo 126,75 milijuna kuna.

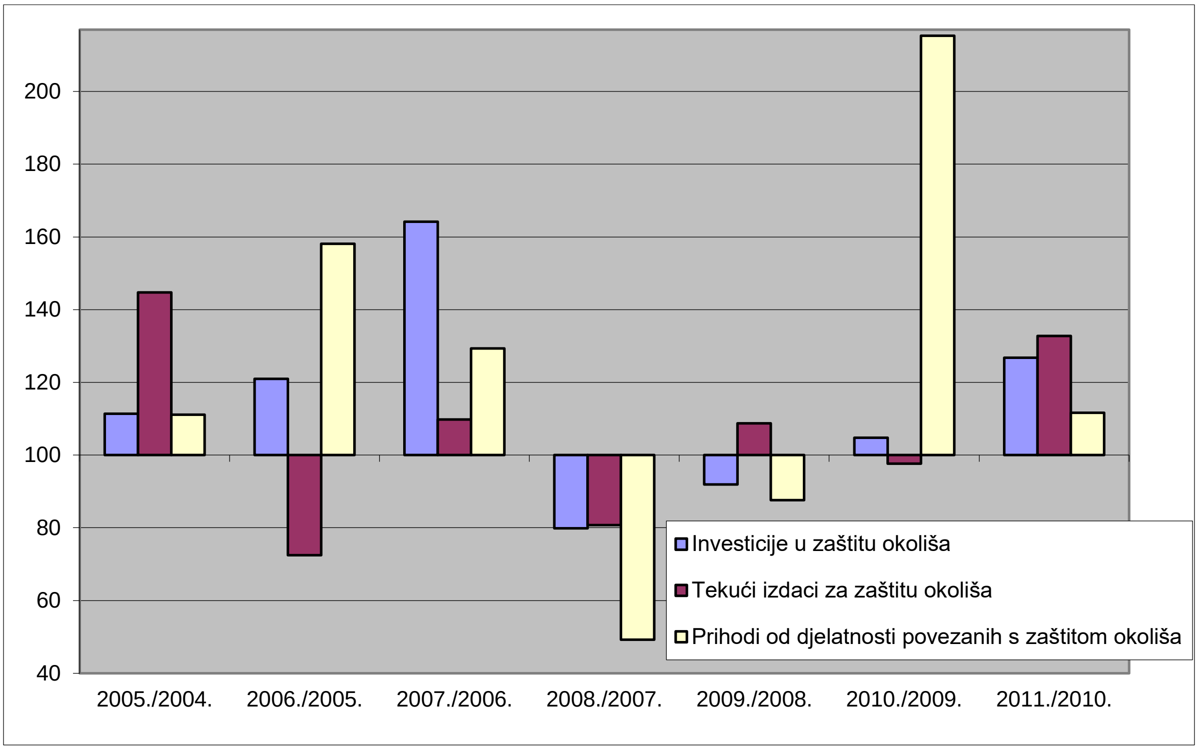
Verižni indeksi mogu se prikazivati površinskim grafikonom i linijskim grafikonom.[21]

Grafikonom 13.7. predstavljeni su verižni indeksi prikazani u tablici 13.5. (tri niza).

[21] Grafičko prikazivanje indeksa na promjenjivoj bazi, vidi: Horvat i Mijoč (2014) str. 561.

Primjer 13.13. Površinski grafikon verižnih indeksa

Grafikon 13.7. Verižni indeksi kretanja investicija u zaštitu okoliša, tekućih izdataka za zaštitu okoliša i prihoda od djelatnosti povezanih sa zaštitom okoliša– verižni indeksi



Izvor: Investicije u zaštitu okoliša, 2004., 2005., 2006., 2007., 2008., 2009., 2010. i 2011., prema Radman-Funarić i Kurtagić (2013)

13.3.3. Preračunavanje indeksa

Preračunavanje baznih indeksa u verižne

U prilikama kada nisu dostupne originalne vrijednosti numeričkog niza, a dostupni su izračunati bazni indeksi, verižni indeksi mogu se preračunati iz baznih indeksa stavljanjem u odnos baznog indeksa promatranog razdoblja prema vrijednosti baznog indeksa prethodnog vremenskog razdoblja.

Preračunavanje se vrši pomoću sljedećeg izraza:

$$V_t = \frac{I_t}{I_{t-a}} \cdot 100$$

Primjer 13.14. Preračunavanje baznih indeksa u verižne

Tablica 13.6. Kretanje investicija u zaštitu okoliša od 2004. do 2011. godine

Godina	Bazni indeks I_t	Postupak izračuna	Verižni indeks V_t
2004.	100		–
2005.	111,36	111/100*100	111,36
2006.	134,68	134,68/111,36*100	120,94
2007.	221,11	221,11/134,68*100	164,18
2008.	176,60	176,6/221,11*100	79,87
2009.	162,35	162,35/176,6*100	91,93
2010.	170,12	170,12/162,35*100	104,79
2011.	215,63	215,63/170,12*100	126,75

$$V_{07} = \frac{I_{07}}{I_{06}} \cdot 100 \qquad V_{07} = \frac{221{,}11}{134{,}68} \cdot 100 = 164{,}18$$

T: Investicije u zaštitu okoliša 2007. godine bile su veće za 64,18% (164,18 – 100) nego li što su bile prethodne 2006. godine. Na svakih investiranih 100 milijuna kuna u zaštitu okoliša 2006. godine, 2007. godine je uloženo 164,18 milijuna kuna.

Preračunavanje verižnih indeksa u bazne

Ukoliko su na raspolaganju verižni indeksi, a za potrebe analize pojave potrebni su i bazni, moguće ih je izračunati iz verižnih.

Preračunavanje verižnih indeksa u bazne, ako je baza usporedbe na početku vremenskog niza, vrši se tako da se za prvo razdoblje upiše vrijednost indeksa 100, a sljedeće vrijednosti se kronološki preračunavaju pomoću izraza:

$$I_t = \frac{I_{t-1} \cdot V_t}{100}$$

Primjer 13.15. Preračunavanje verižnih indeksa u bazne, ako je baza usporedbe na početku vremenskog niza

Tablica 13.7. Kretanje investicija u zaštitu okoliša od 2004. do 2011. godine

Godina	Verižni indeks V_t	Postupak izračuna	Bazni indeks I_t
2004.	–		100
2005.	111,36	100*111,36/100	111,36
2006.	120,94	111,36*120,94/100	134,68
2007.	164,18	134,68*164,18/100	221,11
2008.	79,87	221,11*79,87/100	176,60
2009.	91,93	176,6*91,93/100	162,35
2010.	104,79	162,35*104,79/100	170,12
2011.	126,75	170,12*126,75/100	215,63

$$I_{05} = \frac{100 \cdot 111{,}36}{100} = 111{,}36 \quad I_{06} = \frac{111{,}36 \cdot 120{,}94}{100} = 134{,}68 \quad I_{06} = \frac{134{,}68 \cdot 164{,}18}{100} = 221{,}11$$

T: Investicije u zaštitu okoliša 2006. godine bile su veće za 121,11% (221,11-100) nego li što su bile bazne 2004. godine. Na svakih investiranih 100 milijuna kuna u zaštitu okoliša 2004. godine, 2006. godine je uloženo 221,11 milijuna kuna.

Preračunavanje verižnih indeksa u bazne, ako je baza usporedbe na kraju vremenskog niza, obavlja se tako da se za zadnje razdoblje upiše vrijednost indeksa 100, a sljedeće vrijednosti se za svako prethodno razdoblje preračunavaju pomoću izraza:

$$I_{t-1} = \frac{I_t}{V_t} \cdot 100 \qquad \text{ili} \qquad I_t = \frac{I_{t+1}}{V_{t+1}} \cdot 100$$

Primjer 13.16. Preračunavanje verižnih indeksa u bazne, ako je baza usporedbe na kraju vremenskog niza

Tablica 13.8. Kretanje investicija u zaštitu okoliša od 2004. do 2011. godine

Godina	Verižni indeks Vt	Postupak izračuna	Bazni indeks It
2004.	–	51,64/111,36*100	46,38
2005.	111,36	62,46/120,94*100	51,64
2006.	120,94	102,55/164,18*100	62,46
2007.	164,18	81,9/79,87*100	102,55
2008.	79,87	75,29/91,93*100	81,90
2009.	91,93	78,9/104,79*100	75,29
2010.	104,79	100/126,75*100	78,90
2011.	126,75		100,00

$$I_{10} = \frac{I_{11}}{V_{12}} \cdot 100 \qquad I_{10} = \frac{100}{126{,}75} \cdot 100 = 78{,}90 \qquad I_{9} = \frac{78{,}9}{104{,}79} \cdot 100 = 75{,}29$$

T: Investicije u zaštitu okoliša 2010. godine bile su manje za 21,1% (78,9 – 100) nego li što su bile bazne 2011. godine. Na svakih investiranih 100 milijuna kuna u zaštitu okoliša 2011. godine, 2010. godine je uloženo 78,9 milijuna kuna.

Ukoliko je potrebno verižne indekse preračunati u bazne s bazom usporedbe u sredini vremenskog niza, indeksi za razdoblje prije baze usporedbe računaju se prema izrazu:

$$I_{t-1} = \frac{I_t}{V_t} \cdot 100$$

a indeksi za razdoblje poslije baze usporedbe računaju se prema izrazu:

$$I_t = \frac{I_{t-1} \cdot V_t}{100}$$

13.3.4. Skupni indeksi

Skupni indeksi prate intenzitet dinamike dviju ili više pojava u vremenu. Njima se utvrđuju varijacije heterogene skupine pojava na različitim mjestima, odnosno u različitim vremenskim intervalima. To su relativni brojevi kojima se mjere relativne promjene skupine pojava u vremenu.

Zajedničke varijacije više vremenskih nizova izražavaju se metodom agregiranja tako da se stavi u odnos zbroj podataka svih obuhvaćenih vremenskih nizova prema zbroju podataka poznatog razdoblja. Kako bi se provela analiza heterogene skupine pojava (izražene različitim mjernim sustavima: litre, komadi, grami), potrebno je osigurati usporedivost frekvencija vremenskog niza, što je omogućeno vrijednosnim načinom izražavanja. Npr. vrijednost proizvoda ili usluge izražava se umnoškom količina i cijena.

Pri metodi agregiranja koriste se simboli:
za bazno razdoblje:
p_0 – cijena, q_0 – količina, p_0q_0 – vrijednost
za izvještajno razdoblje:
P_1 – cijena, q_1 – količina, p_1q_1 – vrijednost

Individualni indeksi količina i cijena:

$$I_{i(q)} = \frac{q_{i1}}{q_{i0}} \cdot 100 \qquad I_{i(p)} = \frac{p_{i1}}{p_{i0}} \cdot 100 \qquad i = 1,2,\ldots,k$$

$I_{i(q)}$ – individualni indeks količina
q_{i1} – količina izvještajnog (tekućeg) razdoblja
q_{i0} – količina baznog razdoblja
$I_{i(p)}$ – individualni indeks cijena
p_{i0} – cijene baznog razdoblja
p_{i1} – cijene izvještajnog (tekućeg) razdoblja

k – količina entiteta – za skupinu od *k* entiteta moguće je izračunati onoliko individualnih indeksa koliko ima promatranih entiteta *k*.

Primjer 13.17. Izračun individualnih indeksa

Tablica 13.9. Cijene voća u srpnju 2016. i 2017. godine

PROIZVOD	Cijene u srpnju 2016.	Cijene u srpnju 2017.	Količine u srpnju 2016.	Količine u srpnju 2017.	Individualni indeks količina	Individualni indeks cijena
	p_{i0}	p_{i1}	q_{i0}	q_{i1}	$I_{i(q)}$	$I_{i(p)}$
Jabuke 1 kg	3,50	4,50	350	320	91,43	128,57
Kruške 1 kg	18,50	22,00	130	150	115,38	118,92
Banane 1kg	5,50	6,20	460	440	95,65	112,73
Naranče 1kg	7,20	8,50	250	280	112,00	118,06

k = 4, pa je moguće izračunati četiri individualna indeksa količina i četiri individualna indeksa cijena.

Primjer izračuna za prvi entitet individualnih indeksa količina i cijena:

$$I_{jabuke(q)} = \frac{320}{350} \cdot 100 = 91{,}43 \qquad I_{jabuke(p)} = \frac{4{,}5}{3{,}5} \cdot 100 = 128{,}57$$

T: U izvještajnom razdoblju manja je potrošnja jabuka (za 8,57%) i banana (za 4,35%), dok je porasla potrošnja krušaka (za 15,38%) i naranči (za 12%). Cijene svih navedenih vrsta voća rasle su u izvještajnom razdoblju, a najviše je porasla cijena jabuka, za 28,57%.

Skupni indeksi izračunavaju se kao vagana aritmetička sredina ili vagana harmonijska sredina individualnih indeksa.

Skupni indeksi koji za pondere koriste vrijednosti baznog razdoblja nazivaju se Laspeyresovi skupni indeksi (autor Laspeyres). Skupni indeksi koji za pondere koriste vrijednosti izvještajnog razdoblja nazivaju se Paascheovi skupni indeksi (autor Paasche).

13.3.4.1. Skupni indeksi količina

Skupni indeksi količina pokazuju relativnu promjenu vrijednosti svih promatranih entiteta u vremenu ili prostoru, izazvanu promjenama količina.

Laspeyresov skupni indeks količina računa se kao vagana aritmetička sredina individualnih indeksa količina. Kao ponder koristi se vrijednost baznog razdoblja u apsolutnom izrazu ($q_0 \cdot p_0$). Računa se i izravno, kao **agregatni indeks količina**:

$$Q_{(01)}(P_0) = \frac{\sum_{i=1}^{k} \frac{q_{i1}}{q_{i0}} \cdot q_{i0} p_{i0}}{\sum_{i=1}^{k} q_{i0} p_{i0}} \cdot 100 \qquad Q_{(01)}(P_0) = \frac{\sum_{i=1}^{k} q_{i1} p_{i0}}{\sum_{i=1}^{k} q_{i0} p_{i0}} \cdot 100$$

Vrijednost indeksa je isti, neovisno o korištenom izrazu.

Laspeyresov skupni indeks količina mjeri promjene u vrijednostima za istu količinu dobara i usluga u tekućem razdoblju, u usporedbi s baznim razdobljem.

Paascheov skupni indeks količina računa se kao vagana harmonijska sredina individualnih indeksa količina. Kao ponder koristi se vrijednost tekućeg (izvještajnog) razdoblja ($q_1 \cdot p_1$)u apsolutnom izrazu. Računa se i izravno, kao **agregatni indeks količina**:

$$Q_{(01)}(P_1) = \frac{\sum_{i=1}^{k} q_{i1} p_{i1}}{\sum_{i=1}^{k} \frac{q_{i0}}{q_{i1}} q_{i1} p_{i1}} \cdot 100 \qquad Q_{(01)}(P_1) = \frac{\sum_{i=1}^{k} q_{i1} p_{i1}}{\sum_{i=1}^{k} q_{i0} p_{i1}} \cdot 100$$

Vrijednost indeksa je isti, neovisno o korištenom izrazu.

Paascheov skupni indeks količina mjeri promjene u količinama uz nepromijenjene cijene izvještajnog razdoblja. Razlike u vrijednostima Laspeyresova i Paascheova skupnog indeksa količina nastaje zbog upotrebe različitih pondera. U Laspeyresovu indeksu to su vrijednosti baznog razdoblja, a u Paascheovu indeksu to su vrijednosti izvještajnog razdoblja.

Srednji indeks količina nastao je zbog pokušaja korigiranja prethodna dva načina izračuna indeksa količina (jer se uvijek, iako neznatno, razlikuju zbog različitih cijena u izvještajnom i baznom razdoblju).

Prema **Fisherovoj idealnoj formuli**,srednji indeks količina računa se kao geometrijska sredina Laspeyresova i Paascheova agregatnog indeksa količina:

$$Q_{01} = \sqrt{Q_{01}(P_0) \cdot Q_{01}(P_1)}$$

Q_{01}– srednji indeks količina

Primjer 13.18. Izračun skupnih indeksa količina

Tablica 13.10. Cijene voća u srpnju 2016. i 2017. godine

PROIZVOD	Cijene u srpnju 2016.	Cijene u srpnju 2017.	Količine u srpnju 2016.	Količine u srpnju 2017.	$q_{i1}p_{i0}$	$q_{i0}p_{i0}$	$q_{i1}p_{i1}$	$q_{i0}p_{i1}$
	p_{i0}	p_{i1}	q_{i0}	q_{i1}				
Jabuke 1 kg	3,50	4,50	350	320	1.120	1.225	1.440	1.575
Kruške 1 kg	18,50	22,00	130	150	2.775	2.405	3.300	2.860
Banane 1kg	5,50	6,20	460	440	2.420	2.530	2.728	2.852
Naranče 1kg	7,20	8,50	250	280	2.016	1.800	2.380	2.125
Ukupno					8.331	7.960	9.848	9.412

Laspeyresov skupni indeks količina

$$Q_{(01)}(P_0) = \frac{\sum_{i=1}^{k} q_{i1} p_{i0}}{\sum_{i=1}^{k} q_{i0} p_{i0}} \cdot 100 = \frac{8.331}{7.960} \cdot 100 = 104{,}66$$

T: Prodaja navedenih vrsta voća u srpnju 2017. godine porasla je prosječno za 4,66% u odnosu na prodaju u srpnju 2016. godine. Brojnik indeksa 8.331 pokazuje koliko u kunama iznosi vrijednost navedenih količina voća u srpnju 2017. godine, ali po cijenama iz srpnja 2016. godine.

Paascheov skupni indeks količina

$$Q_{(01)}(P_1) = \frac{\sum_{i=1}^{k} q_{i1} p_{i1}}{\sum_{i=1}^{k} q_{i0} p_{i1}} \cdot 100 = \frac{9.848}{9.412} \cdot 100 = 104{,}63$$

U brojniku je vrijednost količina izvještajnog razdoblja, a u nazivniku vrijednost količina baznog razdoblja po cijenama iz tekućeg razdoblja.

T: Prodaja navedenih vrsta voća u srpnju 2017. godine, po cijenama iz tog razdoblja, povećala se prosječno za 4,63% u odnosu na prodaju u srpnju 2016. godine zbog promijenjenih prodajnih količina.

Fisherov srednji indeks količina

$$Q_{01} = \sqrt{Q_{01}(P_0) \cdot Q_{01}(P_1)} = \sqrt{104{,}66 \cdot 104{,}63} = 104{,}645$$

T: Prodaja voća u srpnju 2017. godine porasla je prosječno za 4,65% u odnosu na prodaju u srpnju 2016. godine.

13.3.4.2. Skupni indeksi cijena

Skupni indeksi cijena pokazuju relativnu promjenu vrijednosti svih promatranih entiteta u vremenu ili prostoru, izazvanu promjenama cijena.

Laspeyresov skupni indeks cijena računa se kao vagana aritmetička sredina ili vagana harmonijska sredina individualnih indeksa cijena. Kao ponder koristi se vrijednost baznog razdoblja u apsolutnom izrazu ($p_0 \cdot q_0$). Računa se i izravno, kao **agregatni indeks cijena**:

$$P_{(01)}(Q_0) = \frac{\sum_{i=1}^{k} \frac{p_{i1}}{p_{i0}} \cdot p_{i0} q_{i0}}{\sum_{i=1}^{k} p_{i0} q_{i0}} \cdot 100 \qquad P_{(01)}(Q_0) = \frac{\sum_{i=1}^{k} p_{i1} q_{i0}}{\sum_{i=1}^{k} p_{i0} q_{i0}} \cdot 100$$

Laspeyresov skupni indeks cijena mjeri promjene u vrijednostima za istu količinu dobara i usluga u tekućem razdoblju, u usporedbi s baznim razdobljem. Pri tome q_0 upućuje da su cijene ponderirane količinama baznog razdoblja.

Paascheov skupni indeks cijena računa se kao vagana harmonijska sredina individualnih indeksa količina. Kao ponder koristi se vrijednost tekućeg (izvještajnog) razdoblja (q_1p_1)u apsolutnom izrazu. Računa se i izravno kao **agregatni indeks cijena**:

$$P_{(01)}(Q_1)=\frac{\sum_{i=1}^{k}p_{i1}q_{i1}}{\sum_{i=1}^{k}\frac{p_{i0}}{p_{i1}}p_{i1}q_{i1}}\cdot 100 \qquad P_{(01)}(Q_1)=\frac{\sum_{i=1}^{k}p_{i1}q_{i1}}{\sum_{i=1}^{k}p_{i0}q_{i1}}\cdot 100$$

$1/I_{(p)} = p_0/p_1 \cdot 100$ – recipročna vrijednost individualnog indeksa cijena

Paascheov skupni indeks cijena mjeri promjene u cijenama uzimajući u obzir količine u izvještajnom razdoblju koje služe kao ponder.

Srednji indeks cijena nastao je radi korigiranja prethodna dva načina izračuna indeksa cijena.

Prema **Fisherovoj idealnoj formuli**,srednji indeks cijena računa se kao geometrijska sredina Laspeyresova i Paascheova agregatnog indeksa cijena:

$$P_{01}=\sqrt{P_{01}(Q_0)\cdot P_{01}(Q_1)}$$

P_{01}– srednji indeks cijena

Primjer 13.19. Izračun skupnih indeksa cijena

Prema izračunatim brojnicima i nazivnicima iz tablice 12.9. proizlazi:

Laspeyresov skupni indeks cijena

$$P_{(01)}(Q_0)=\frac{\sum_{i=1}^{k}p_{i1}q_{i0}}{\sum_{i=1}^{k}p_{i0}q_{i0}}\cdot 100=\frac{9.412}{7.960}\cdot 100=118{,}24$$

T: Za nabavu voća u količinama iz srpnja 2016. godine, potrebno je u srpnju 2017. godine izdvojiti 9.412,00 kn, dok je za istu količinu u srpnju 2016. godine bilo potrebno izdvojiti 7.960,00 kn. Dakle, cijene količina voća iz srpnja

2016. godine porasle su u srpnju 2017. godine u prosjeku za 18,24%, u odnosu na cijene iz srpnja 2016. godine.

Paascheov skupni indeks cijena

$$P_{(01)}(Q_1) = \frac{\sum_{i=1}^{k} p_{i1} q_{i1}}{\sum_{i=1}^{k} p_{i0} q_{i1}} \cdot 100 = \frac{9.848}{8.331} \cdot 100 = 118{,}21$$

U brojniku je vrijednost cijena izvještajnog razdoblja, a u nazivniku vrijednost cijena baznog razdoblja, po količinama iz tekućeg razdoblja.

T: Za nabavu voća u količinama iz srpnja 2017. godine, potrebno je u srpnju 2017. godine izdvojiti 9.848,00 kn, dok je za istu količinu u srpnju 2016. godine bilo potrebno izdvojiti 8.331,00 kn. Dakle, cijene količina voća iz srpnja 2017. godine porasle su u srpnju 2017. godine u prosjeku za 18,21%, u odnosu na cijene iz srpnja 2016. godine.

Fisherov srednji indeks cijena

$$P_{01} = \sqrt{P_{01}(Q_0) \cdot P_{01}(Q_1)} = \sqrt{118{,}24 \cdot 118{,}21} = 118{,}23$$

T: Cijena voća u srpnju 2017. godine porasla je prosječno za 118,23% u odnosu na cijene u srpnju 2016. godine.

13.3.4.3. Skupni indeksi vrijednosti

Skupni indeks vrijednosti mjeri promjene vrijednosti skupine pojava. Računa s kao omjer vrijednosti izvještajnog razdoblja i vrijednosti baznog razdoblja, prema izrazu:

$$V_{01} = \frac{\sum_{i=1}^{k} q_{i1} p_{i1}}{\sum_{i=1}^{k} q_{i0} p_{i0}} \cdot 100$$

Skupni indeks vrijednosti ne pokazuje koliko su na rezultat djelovale promjene cijena, a koliko promjene količina.

Moguće ga je izračunati i kao umnožak Laspeyresova skupnog indeksa cijena i Paascheova skupnog indeksa količina

$$V_{01} = \left(P_{01}(Q_0) \cdot Q_{01}(P_1)\right)/100$$

te kao umnožak Laspeyresova skupnog indeksa količina i Paascheova skupnog indeksa cijena

$$V_{01} = \left(Q_{01}(P_0) \cdot P_{01}(Q_1)\right)/100$$

Primjer 13.20. Izračun skupnog indeksa vrijednosti

Prema izračunatim brojnicima i nazivnicima iz tablice 12.9. te izračunatih skupnih indeksa količina i cijena, izračunat je skupni indeks vrijednosti.

$$V_{01} = \frac{\sum_{i=1}^{k} q_{i1} p_{i1}}{\sum_{i=1}^{k} q_{i0} p_{i0}} \cdot 100 = \frac{9.848}{7.960} \cdot 100 = 123{,}72$$

$$V_{01} = \left(P_{01}(Q_0) \cdot Q_{01}(P_1)\right)/100) = 118{,}24 \bullet 104{,}63/100 = 123{,}72$$

$$V_{01} = \left(Q_{01}(P_0) \cdot P_{01}(Q_1)\right)/100 = 104{,}66 \cdot 118{,}21/100 = 123{,}72$$

T: Vrijednost izdataka za navedeno voće u srpnju 2017. godine povećala se u prosjeku za 23,72% u odnosu na srpanj 2016. godine.

Pokazatelje gospodarskih kretanja moguće je izračunati pomoću skupnih indeksa. Njima se može izračunati obujam industrijske proizvodnje, indeks troškova života, promjene cijena u trgovini na malo, promet na burzama. U Hrvatskoj se od 2004. godini izračunava **indeks potrošačkih cijena** (engl. *Consumer price index* – CPI). Njegov izračun je u skladu s metodologijom Međunarodne organizacije rada (ILO) i statističkog ureda Europske unije (EUROSTAT).

Budući da indeks potrošačkih cijena pokazuje promjene na razini cijena dobara i usluga, koje tijekom određenog vremena nabavlja stanovništvo radi osobne potrošnje, ujedno predstavlja i jedinstvenu opću mjeru inflacije u Hrvatskoj. Indeks potrošačkih cijena zamjenjuje dotadašnji izračun indeksa cijena na malo, indeksa troškova života i indeksa ugostiteljskih usluga. Izračun indeksa potrošačkih cijena računa se prema formuli Laspeyresova tipa[22].

Indeks kupovne moći pokazuje kupovnu moć novca koja je za nepromijenjene količine dobara i usluga obrnuto proporcionalna povećanju nabavnih cijena.

Indeks troškova života (potrošačke kupovne moći) računa se prema izrazu:

$$I_{01} = \frac{1}{P_{0t}(q_0)} \cdot 100$$

I_{0t}– indeks potrošačke kupovne moći u razdoblju *t* prema baznom razdoblju 0

$P_{0t}(q_0)$ – indeks troškova života kojem je bazno razdoblje 0

Primjer 13.21. Izračunavanje indeksa potrošačke kupovne moći

$$I_{01} = \frac{1}{102{,}8} \cdot 100 = 0{,}9728$$

Ako je indeks troškova života u 2017. godini u odnosu na 2016, godinu 102,8, indeks potrošačke kupovne moći iznosi 0,9728 što znači da se kupovna moć

[22] O metodologiji izračuna indeksa potrošačkih cijena vidi: Indeksi cijena, URL: https://www.hnb.hr/statistika/statisticki-podaci/odabrane-nefinancijske-statistike/indeksi-cijena, te njihovoj strukturi: Indeksi potrošačkih cijena, URL: https://www.dzs.hr/Hrv_Eng/publication/2017/13-01-01_06_2017.htm
i EUROSTAT URL:
http://ec.europa.eu/eurostat/cache/metadata/EN/prc_hicp_esms_hr.htm

kune 2017. godine u odnosu na 2016. godinu smanjila za 0,027 ili 2,7%. U 2017. godini jedna kuna vrijedi 0,973 kn prema jednoj kuni u 2016. godini.

Burzovni indeks je pokazatelj relativne promjene vrijednosti odabranih dionica, obveznica i drugih vrijednosnih papira koje kotiraju na burzi. Indeks se temelji na podacima o cijenama i količinama po kojima se trgovalo vrijednosnim papirom. Burzovni indeks računa se kao skupni indeks vrijednosti, aritmetičke ili geometrijske sredine zaključnih cijena. Računa se prema izrazu:

$$V_{0t} = \frac{\sum_{i=1}^{k} p_{it} q_{it}}{\sum_{i=1}^{k} p_{io} q_{i0}}$$

13.4. Srednje vrijednosti vremenskog niza

Srednja vrijednost vremenskog niza je frekvencija vremenskog niza koja reprezentira sve njegove pojedinačne vrijednosti,a izražena je u onim jedinicama mjere u kojima su zadane frekvencije.

Upotreba pojedine srednje vrijednosti ovisi o tendenciji kretanja frekvencija i vrsti vremenskog niza. Za izračun srednjih vrijednosti vremenskog niza koristi se aritmetička sredina intervalnog vremenskog niza, kronološka sredina trenutačnog vremenskog niza, geometrijska sredina koja se upotrebljava u analizi intervalnog i trenutačnog vremenskog niza.

Tablica 13.11. Srednje vrijednosti vremenskog niza

Vrsta vremenskog niza	Oscilirajući	Tendencija rasta ili pada
INTERVALNI	Aritmetička sredina intervalnog niza	Geometrijska sredina
TRENUTAČNI	Kronološka sredina	

Aritmetička sredina intervalnog niza izračunava se prema izrazu za jednostavnu aritmetičku sredinu. Pokazuje npr. prosječan broj putnika godišnje prevezen u gradskom i zračnom prijevozu.

To je prosječna vrijednost frekvencije po jednoj zadanoj vremenskoj jedinici, a koristi se za oscilirajuće intervalne vremenske nizove. Aritmetička sredina niza y (prosječna ordinata), računa se prema izrazu:

$$\overline{Y} = \frac{\sum_{t=1}^{k} Y_t}{N}$$

Kronološka sredina je prosječno stanje pojave trenutačnog vremenskog niza u vremenskom trenutku. Pri izračunavanju kronološke sredine svaka frekvencija niza množi se vremenskim ponderom odnosno s dužinom razdoblja za koje se pretpostavlja da ima frekvenciju na razini koju pokazuje pojava u definiranom trenutku vremena, prema izrazu:

$$\overline{Y} = \frac{\sum_{i=1}^{k} p_t Y_t}{\sum_{t=1}^{k} p_i}$$

$\bar{y}$– kronološka sredina

p_t – ponder

Pri određivanju pondera uzima se pretpostavka da je razina pojave u trenutku snimanja *t* ista do polovine intervala *t*+1, odnosno do polovine intervala *t*-1, s tim da ti intervali mogu biti međusobno jednaki ili različiti.

Primjer 13.22. Izračunavanje kronološke sredine

Tablica 13.12. Broj nesilica na obiteljskom gospodarstvu KOKA

Stanje na dan	Broj nesilica Y_t	Vremenski interval od prethodnog mjerenja	p_t	$p_t\ Y_t$
31.12.2016.	5.422		1,5	8.133,00
31.03.2017.	6.224	3 mjeseca	2,0	12.448,00
31.04.2017.	7.934	1 mjesec	2,0	15.868,00
31.07.2017.	9.681	3 mjeseca	2,5	24.202,50
30.09.2017.	9.802	2 mjeseca	1,0	9.802,00
Ukupno	39.063		9,0	70.453,50

Ponder za prvo razdoblje izračunat je 0+3 = 3/2= 1,5, za drugo razdoblje 3+1 = 4/2 = 2, itd.

$$\overline{Y} = \frac{70.453,5}{9} = 7.828,17$$

T: Prosječno stanje nesilica na obiteljskom gospodarstvu od 31. prosinca 2016. do 30. rujna 2017. godine je 7.828 nesilica.

Ako se vrijednosti trenutačnog vremenskog niza promatraju u jednako udaljenim točkama vremena, ekvidistantno (na primjer, svaki mjesec ili svaku godinu), kronološka sredina se računa na sljedeći način:

$$\overline{Y} = \frac{\frac{Y_1 + Y_k}{2} + \sum_{t=2}^{k-1} Y_t}{k-1}$$

Y_1 – prva vrijednost vremenskog niza

Y_k – zadnja vrijednost vremenskog niza

Geometrijska sredina je srednja vrijednost koja se izračunava za intervalne ili trenutačne nizove koji pokazuju tendenciju rasta ili pada. Predstavlja veličinu kojom se određuje veličina rasta ili pada za promatranu vremensku jedinicu. Geometrijska sredina je srednja vrijednost koja predstavlja prosječnu stopu kretanja promatranog vremenskog niza i koja se odnosi na promatranu vremensku jedinicu.

Koristi se u:

- u analizi vremenskih nizova negrupiranih i grupiranih podataka (prosječan tempo promjene)

- kao srednja vrijednost numeričkih nizova za nizove sa asimetričnim rasporedom podataka

U pravilu se koristi za nizove koji prikazuju pojave koje se ponašaju približno po geometrijskoj progresiji, kao npr. priraštaj pučanstva, ekonomska ulaganja, investicije, štedni ulozi.

Izračunavanje geometrijske sredine opisano je u potpoglavlju 13.2.2.3.

13.5. Trend

Ovisno o vrsti čimbenika koji djeluju na neku pojavu, vremenski niz čine sljedeće komponente:

- komponenta trenda koja predstavlja osnovnu tendenciju razvoja pojave u vremenu. Izražava se funkcijom vremena.
- sezonska komponenta koja se očituje u sustavnim kretanjima pojave unutar jedne godine. Često je prisutna u gospodarskim pojavama (turistička sezona u ljetnim ili zimskim mjesecima, proizvodnja u vrijeme dozrijevanja žitarica i dr).
- ciklička komponenta očituje se kada se vremenska pojava obnavlja na približno isti način (redovito ili periodično) dugoročnije od sezonske. Javlja se u nepravilnim intervalima. Sastoji se od uzlazne faze ili ekspanzije i silazne faze ili recesije, što je svojstveno poslovnim ciklusima.
- slučajna (stohastička) komponenta proizlazi iz utjecaja koji nisu predvidivi, ili nisu na vrijeme predvidivi. Slična je slučajnoj pogreški u u regresijskom modelu.

Vremenski niz zapisuje se općim oblikom aditivnog modela:

$$Y = T + C + S + e$$

Y – promatrana pojava

T– vrijednost komponente trenda

C – vrijednost cikličke komponente

S – vrijednost sezonske komponente

e – vrijednost slučajne komponente

U kratkom vremenskom razdoblju trend i ciklička komponenta se ne razdvajaju pa se model može pisati:

$$Y = T + S + e$$

T– trend i ciklička komponenta

Opći oblik multiplikativnog modela:

$$Y = T \cdot I_s \cdot I_\varepsilon$$

I_s – sezonska komponenta

I_ε– slučajna komponenta

a u logaritamskom obliku:

$$\log Y = \log T \cdot \log I_s \cdot \log I_\varepsilon$$

S obzirom na funkciju vremena, trend može biti *linearni i eksponencijalni.*

13.5.1. Linearni trend

Linearni trend koristi se za opisivanje kretanja dugoročno linearnih pojava u vremenu.

Model linearnog trenda identičan je modelu jednostavne linearne regresije, u kojemu je vrijeme nezavisna varijabla. Prikladan je kada se vremenska pojava mijenja od razdoblja do razdoblja za približno isti apsolutni iznos, odnosno kada su prve diferencije približno konstantne.

Jednadžba linearnog trenda glasi:

$$yc_t = a + bx_t$$

yc – zavisna (trend) varijabla

x – oznaka za vrijeme

a – vrijednost trenda u ishodištu

b – koeficijent smjera pravca

Jednadžba linearnog trenda izračunava se metodom najmanjih kvadrata (parametrijska metoda).

Parametrom *a* procjenjuje se vrijednost varijable *y* (trend vrijednosti), kada linija trenda presijeca os y (x = 0). Parametar *b* naziva se koeficijent smjera linije trenda. On određuje smjer i nagib (intenzitet) trenda, odnosno pokazuje za koliko se prosječno linearno mijenja pojava *y* za svaku promjenu jedne jedinice vremena (varijabla *x*).

Jednadžba linearnog trenda može imati ishodište na početku vremenskog niza, što se u praksi i najčešće koristi, ili u sredini vremenskog niza.

Parametar *b* izračunava se prema izrazima:

$$b = \frac{\sum_{t=1}^{N} x_t y_t - \sum_{t=1}^{N} y_t}{\sum_{t=1}^{N} x_t^2 - \bar{x}\sum_{t=1}^{N} x_t} \quad \text{ili} \quad b = \frac{\sum_{t=1}^{N} x_t y_t - N \cdot \bar{x} \cdot \bar{y}}{\sum_{t=1}^{N} x_t^2 - N \cdot \bar{x}^2}$$

Parametar *a* izračunava se prema izrazu:

$$a = \bar{y} - b\bar{x} \qquad \bar{x} = \frac{\sum_{t=1}^{N} x_t}{N} \qquad \bar{y} = \frac{\sum_{t=1}^{N} y_t}{N}$$

Ukoliko se ishodište postavi u sredinu vremenskog niza, parametri *a* i *b* računaju se prema skraćenim izrazima. Izrazi postaju skraćeni jer drugi dio izraza ima vrijednost nula, budući da je zbroj vrijednosti varijable *x* jednak nuli, a time je i aritmetička sredina jednaka nuli.

$$b = \frac{\sum_{t=1}^{N} x_t y_t}{\sum_{t=1}^{N} x_t^2} \qquad a = \bar{y}$$

Primjer 13.23. Izračunavanje jednadžbe linearnog trenda s ishodištem na početku vremenskog niza

Tablica 13.13. Dobit u trgovačkom društvu Beta

Godina	Ostvarena dobit (u 000 kuna)	x_t	$x_t\ y_t$	x_t^2	$Yc = a+bx_t$	Rezidualna odstupanja	$(y_t - yc)^2$
x_t	y_t					y_t - yc	
2011.	1.864	0	0	0	1.746,61	117,39	13.781,083
2012.	1.625	1	1.625	1	1.841,50	-216,50	46.872,250
2013.	2.128	2	4.256	4	1.936,39	191,61	36.713,297
2014.	2.005	3	6.015	9	2.031,29	-26,29	690,939
2015.	1.957	4	7.828	16	2.126,18	-169,18	28.621,389
2016.	2.250	5	11.250	25	2.221,07	28,93	836,862
2017.	2.390	6	14.340	36	2.315,96	74,04	5.481,287
Ukupno	**14.219**	**21**	**45.314**	**91**	**14.219,00**	**0**	**132.997,107**

$$\bar{x} = \frac{21}{7} = 3 \qquad \bar{y} = \frac{14.219}{7} = 2.031{,}29$$

$$b = \frac{45.314 - 7 \cdot 3 \cdot 2.031{,}29}{91 - 7 \cdot 3^2} = \frac{45.314 - 42.657{,}09}{28} = 94{,}89$$

$$a = 2.031{,}29 - 94{,}89 \cdot 3 = 1.746{,}62$$

Rješenje jednadžbe linearnog trenda:

$yc = 1.746{,}62 + 94{,}89x$
ishodište: 2011. (30.6.)
jedinica za x: 1 godina
jedinica za y: 1000 kn

T: Prema trendu u razdoblju od 2011. do 2017. godine, dobit trgovačkog društva Beta 2011. godine (30.6.) iznosila je 1.746.620 kuna s prosječnim godišnjim porastom od 94.890 kuna.

Nakon zapisane jednadžbe izračunate su trend vrijednosti, tako da su u jednadžbu uvrštene pripadajuće vrijednosti x_t. Zbroj trend vrijednosti jednaka je zbroju originalnih vrijednosti niza y.

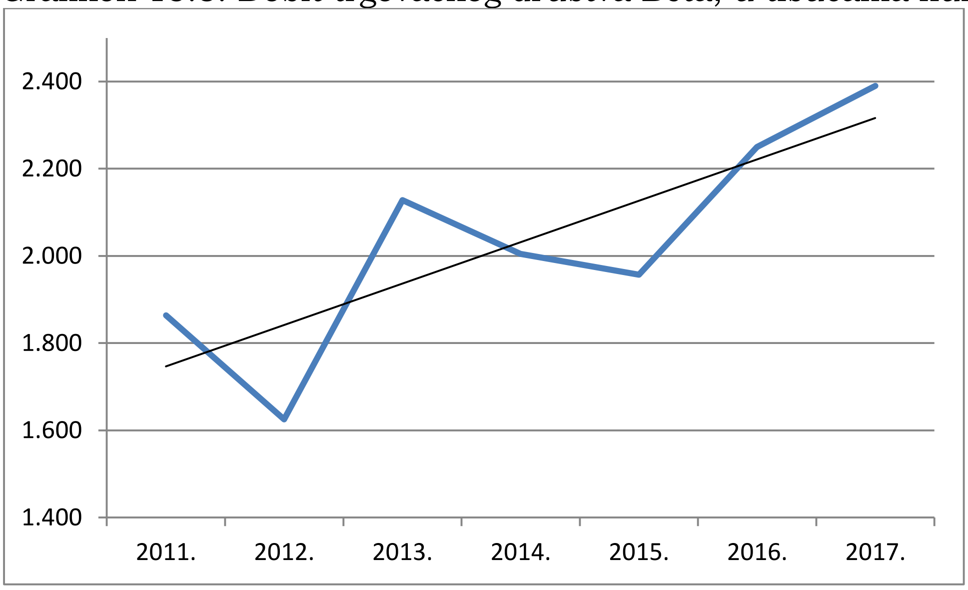
T: npr. prema trendu dobit trgovačkog društva Beta 2013. godine iznosila je 1.936.390 kn.

Varijabla vrijeme može poprimiti prvu vrijednost 1, odnosno ishodište na početku vremenskog niza može poprimiti vrijednost 1 umjesto 0. Trend vrijednosti će poprimiti iste vrijednosti.

Rezidualna odstupanja nastala su kao razlika između originalnih vrijednosti y i trend vrijednosti. Zbroj tih razlika uvijek je nula.

Primjer 13.24. Grafički prikaz linearnog trenda

Grafikon 13.8. Dobit trgovačkog društva Beta, u tisućama kuna



Primjer 13.25. Izračunavanje jednadžbe linearnog trenda s ishodištem u sredini vremenskog niza

Tablica 13.14. Dobit u trgovačkom društvu Beta

Godina	Ostvarena dobit (u 000 kuna)	x_t	$x_t\ y_t$	x_t^2	$yc_t = a+bx_t$	Rezidualna odstupanja	$(y_t - yc_t)^2$
x_t	y_t					y_t - yc_t	
2011.	1.864	-3	-5.592	9	1.746,61	117,39	13.781,083
2012.	1.625	-2	-3.250	4	1.841,50	-216,50	46.872,250
2013.	2.128	-1	-2.128	1	1.936,39	191,61	36.713,297
2014.	2.005	0	0	0	2.031,29	-26,29	690,939
2015.	1.957	1	1.957	1	2.126,18	-169,18	28.621,389
2016.	2.250	2	4.500	4	2.221,07	28,93	836,862
2017.	2.390	3	7.170	9	2.315,96	74,04	5.481,287
Ukupno	**14.219**	**0**	**2.657**	**28**	**14.219,00**	**0**	**132.997,107**

$$b = \frac{2.657}{28} = 94{,}89 \qquad a = \bar{y} = \frac{14.219}{7} = 2.031{,}29$$

Rješenje jednadžbe linearnog trenda s ishodištem u sredini vremenskog niza:

$yc = 2.031{,}29 + 94{,}89x$
ishodište: 2014. (30.6.)
jedinica za x: 1 godina
jedinica za y: 1000 kn

T: Prema trendu u razdoblju od 2011. do 2017. godine, dobit trgovačkog društva Beta 2014. godine (30.6.) iznosila je 2.031.290 kuna s prosječnim godišnjim porastom od 94.890 kuna.

Rezultat je isti kao u prethodnom primjeru, no budući da je promijenjeno ishodište vremenskog niza, jednadžba ima različitu vrijednost parametra *a*, koji pokazuje vrijednost trenda u ishodištu. Parametar *b* nije promijenjen jer je izračunat iz istog niza podataka i pokazuje za koliko se dobit trgovačkog društva mijenjala u prosjeku godišnje. Budući da su trend vrijednosti iste, nisu se promijenile ni vrijednosti rezidualnih odstupanja.

Reprezentativnost modela trenda ispituje se pomoću varijance, standardne devijacije i koeficijenta varijacije. Radno pravilo je da se varijabilnost, mjerena koeficijentom varijabilnosti, veća od 30% smatra velikom, a reprezentativnost prosjeka slabom.

Standardna devijacija i koeficijent varijacije trenda računa se prema izrazu:

$$\hat{\sigma} = \sqrt{\frac{\sum_{i=1}^{N}(y_t - yc_t)^2}{N}} \qquad V_{yc} = \frac{\hat{\sigma}}{\bar{y}} \cdot 100$$

Primjer 13.26. Izračun standardne devijacije i koeficijenta varijacije trenda

$$\hat{\sigma} = \sqrt{\frac{132.997{,}11}{7}} = \sqrt{18.999{,}59} = 137{,}84$$

$$V_{yc} = \frac{137{,}84}{2.031{,}29} \cdot 100 = 6{,}79\%$$

T: Standardna devijacija i koeficijent varijacije trenda pokazuju da je dobit u trgovačkom društvu Beta vrlo malo varirala u razdoblju od 2011. do 2017. godine, u prosjeku za 137.000 kuna, ili 6,79%. To pokazuje da je reprezentativnost modela trenda visoka te je linearni trend primjeren model za analiziranje prikazanog niza podataka.

Ukoliko se analizira trend vrijednosti vremenskog niza s parnim brojem razdoblja s ishodištem u sredini niza vrijednosti, x_t poprimaju vrijednosti polovice jedinice vremenskog razdoblja.

Primjer 13.27. Izračunavanje jednadžbe linearnog trenda s ishodištem u sredini vremenskog niza (paran broj razdoblja)

Tablica 13.15. Prihod u trgovačkom društvu Beta

Godina	Ostvaren prihod (u 000 kuna)	x_t	$x_t\, y_t$	x_t^2	$yc_t = a + bx_t$	Rezidualna odstupanja	$(y_t - yc_t)^2$
x_t	y_t					y_t - yc_t	
2009.	3.125	-7	-21.875	49	2.553,67	571,33	326.421,778
2010.	1.988	-5	-9.940	25	2.496,90	-508,90	258.984,057
2011.	2.544	-3	-7.632	9	2.440,14	103,86	10.786,306
2012.	1.999	-1	-1.999	1	2.383,38	-384,38	147.748,717
2013.	2.347	1	2.347	1	2.326,62	20,38	415,383
2014.	2.305	3	6.915	9	2.269,86	35,14	1.235,020
2015.	2.154	5	10.770	25	2.213,10	-59,10	3.492,247
2016.	2.378	7	16.646	49	2.156,33	221,67	49.136,111
Ukupno	**18.840**	**0**	**-4.768**	**168**	**18.840,00**	**0**	**798.219,619**

$$\hat{\sigma} = \sqrt{\frac{798.219,62}{8}} = \sqrt{99.777,45} = 315,88$$

$$V_{yc} = \frac{315,88}{2.355,00} \cdot 100 = 13,41\%$$

T: Model linearnog trenda je reprezentativan.

Standardna devijacija i koeficijent varijacije trenda pokazuju da je prihod u trgovačkom društvu Beta varirao u razdoblju od 2009. do 2016. godine, u prosjeku za 315.880 kuna, ili 13,41%.

$$b = \frac{-4.768}{168} = -28,38 \qquad a = \bar{y} = \frac{18.840}{8} = 2.355,00$$

Rješenje jednadžbe linearnog trenda s ishodištem u sredini vremenskog niza:

$yc = 2.355,00 - 28,38x$
ishodište: 2012. (31.12.)
jedinica za x: 1/2 godine
jedinica za y: 1000 kn

T: Prema trendu u razdoblju od 2009. do 2016. godine, prihod trgovačkog društva Beta 2012. godine (31.12.) iznosio je 2.355.000 kuna s prosječnim polugodišnjim padom od 28.380 kuna.

Primjer 13.28. Premještanje ishodišta na početak vremenskog niza

Premještanjem ishodišta na početak vremenskog niza rješenje jednadžbe linearnog trenda glasi:

$yc = 2.553{,}67 - 28{,}38x$
ishodište: 2009. (30.06.)
jedinica za x: 1/2 godine
jedinica za y: 1000 kn

T: Prema trendu u razdoblju od 2009. do 2016. godine, prihod trgovačkog društva Beta 2009. godine (30.06.) iznosilo je 2.553.670 kuna s prosječnim polugodišnjim padom od 28.380 kuna.

Primjer 13.29. Preračunavanjem polugodišnje jednadžbe u godišnju

Preračunavanjem polugodišnje jednadžbe u godišnju rješenje jednadžbe linearnog trenda glasi:

$yc = 2.553{,}67 - 56{,}76x$
ishodište: 2009. (30.06.)
jedinica za x: 1 godina
jedinica za y: 1000 kn

T: Prema trendu u razdoblju od 2009. do 2016. godine, prihod trgovačkog društva Beta 2009. godine (30.06.) iznosilo je 2.553.670 kuna s prosječnim godišnjim padom od 56.760 kuna.

13.5.2. Eksponencijalni trend

Eksponencijalni trend je vrsta krivolinijskog trenda. Prikladan je kada se vremenska pojava mijenja od razdoblja do razdoblja na približno istoj razini, odnosno kada su verižni indeksi približno konstantni. Prema toj informaciji odlučit će se za primjenu eksponencijalnog trenda.

Model jednostavnog eksponencijalnog trenda

$y_t = \alpha \cdot \beta^{x_t} \cdot \varepsilon_t$ Ili $\log y_t = \log \alpha + \log \beta x_t + \log \cdot \varepsilon_t$

ε–rezidualna komponenta (malo grčko slovo, čit. epsilon)

Logaritamskom transformacijom model eksponencijalnog trenda svodi se na model linearnog trenda.

13.6. Pomični prosjeci

Vremenski nizovi često pokazuju komponentu neregularnog ponašanja uvjetovanu različitim čimbenicima koji iskazuju obrasce ponašanja koji izgledaju prilično nepredvidivi. Neregularne komponente mogu biti izražene u tolikoj mjeri da zamagljuju bilo koje regularnosti koje se u njima pojavljuju te onemogućuju grafički prikaz izraženih odstupanja. Svaka velika neregularna komponenta u određenom vremenskom trenutku prouzročiti će manji učinak ako je uprosječena s njezinim neposrednim susjedima. U tome pomaže **metoda pomičnih prosjeka** kojom se izglađuju originalni podaci zamjenom članova vremenskog niza nizom aritmetičkih sredina, tako da se kreira novi niz pomicanjem aritmetičkih sredina kroz vremenski niz. Metoda pomičnih prosjeka je neparametrijska metoda za utvrđivanje trenda.

Svrha izračuna pomičnih prosjeka uočljivija je kod dužeg vremenskog niza. Poteškoću predstavlja nedostatak vrijednosti pomičnih prosjeka na početku i na kraju serije.

Pomični prosjeci su aritmetičke sredine *M* uzastopnih vrijednosti članova vremenskog niza ($M<n$). Niz pomičnih prosjeka čini izvedeni niz koji ima manji stupanj varijabilnosti u usporedbi s izvornim nizom. Stoga se pomičnim prosjecima izglađuje vremenska serija. Računanje pomičnih prosjeka može se provesti uzimanjem jednostavnog pomičnog prosjeka od (2m+1) točaka, s centom u dotičnoj točki, odnosno opažanje y_t zamjenjuje se njegovim i prosjekom njegovih susjeda. Ako je broj članova pomičnog prosjeka neparan (M=2m+1), pomični prosjeci računaju se prema izrazu:

$$y_t^* = \frac{1}{M}\sum_{s=-m}^{m} y_{t+s} \qquad t = m+1, m+2, ..., n-m$$

Kada je broj članova pomičnog prosjeka paran (M=2m) provodi se postupak centriranja. Centriranje se može provesti određivanjem jednostavnih pomičnih prosjeka od prethodnih pomičnih prosjeka od po dva člana ili izravno prema izrazu:

$$y_t^* = \frac{1}{M}\left[\frac{1}{2}y_{t-m} + \sum_{s=-(m-1)}^{m-1} y_{t+s} + \frac{1}{2}y_{t+m}\right] \qquad t = m+1, m+2, ..., n-m$$

m–broj uzastopnih vrijednosti članova niza

y_t^*– vrijednosti pomičnih prosjeka

y_t– originalna vrijednost vremenskih razdoblja

t–promatrano razdoblje za koje se računa pomični prosjek

n –broj članova vremenskog niza koji sudjeluju u računanju pomičnih prosjeka

Primjer 13.30. Rezultati izračuna pomičnih prosjeka

Proizvodnja primarne energije u Republici Hrvatskoj izražene u PJ (petajoule) kretala se prema podacima prikazanim u tablici 13.13.

Tablica 13.16. Proizvodnja primarne energije u Republici Hrvatskoj 2003 – 2010.

		Proizvodnja primarne energije u PJ	Pomični prosjek M = 3	Pomični prosjek M = 5	Pomični prosjek M = 4	Centrirani pomični prosjek M = 4
	x	y_t	$y_t(3)$	$y_t(5)$	$y_t(4)$	
y_1	2003.	148,00	*	*	*	*
y_2	2004.	161,29	155,86	*	*	*
y_3	2005.	158,30	164,14	161,91	160,11	162,75
y_4	2006.	172,84	166,76	165,59	165,39	166,03
y_5	2007.	169,14	169,46	167,30	166,67	168,11
y_6	2008.	166,40	168,46	170,57	169,56	169,78
y_7	2009.	169,84	170,30	*	170,01	*
y_8	2010.	174,65	*	*	*	*

Izvor: Nacionalni akcijski plan energetske učinkovitosti RH za razdoblje do kraja 2013., prema Pisker i Radman-Funarić (2013)

U proizvodnji primarne energije prisutne su oscilacije u razdoblju od 2003. do 2010. Na temelju navedenih podataka u nastavku je prikazan izračun pomičnih prosjeka kako bi se izgladila vremenska serija y_t^*.

$$y_{2004}(3) = \frac{148{,}00 + 161{,}29 + 158{,}3}{3} = 155{,}86 PJ$$

$$y_{2009}(3) = \frac{166{,}4 + 169{,}84 + 174{,}65}{3} = 170{,}30 PJ$$

$$y_{2005}(5) = \frac{148{,}00 + 161{,}29 + 158{,}3 + 172{,}84 + 169{,}14}{5} = 161{,}91 PJ$$

$$y_{2008}(5) = \frac{172{,}84 + 169{,}14 + 166{,}40 + 169{,}84 + 174{,}65}{5} = 170{,}57 PJ$$

Prilikom računanja pomičnih prosjeka četiriju točaka (parni broj) prvi član je prosjek prva četiri opažanja prema:

$$Y_{2005}(4)=\frac{148{,}00+161{,}29+158{,}3+172{,}84}{4}=160{,}11PJ$$

Drugi član niza izračunat je prema:

$$y_{2006}(4)=\frac{161{,}29+158{,}3+172{,}84+169{,}14}{4}=165{,}39PJ$$

Zadnji član je prosjek zadnja četiri opažanja prema:

$$y_{2009}(4)=\frac{169{,}14+166{,}40+169{,}84+174{,}65}{4}=170{,}01PJ$$

Pomični prosjeci niza od četiri točke (M = 4) su centrirani računanjem prosjeka susjednih parova prema:

$$y_{2005}(4)=\frac{160{,}11+165{,}39}{2}=162{,}75PJ$$

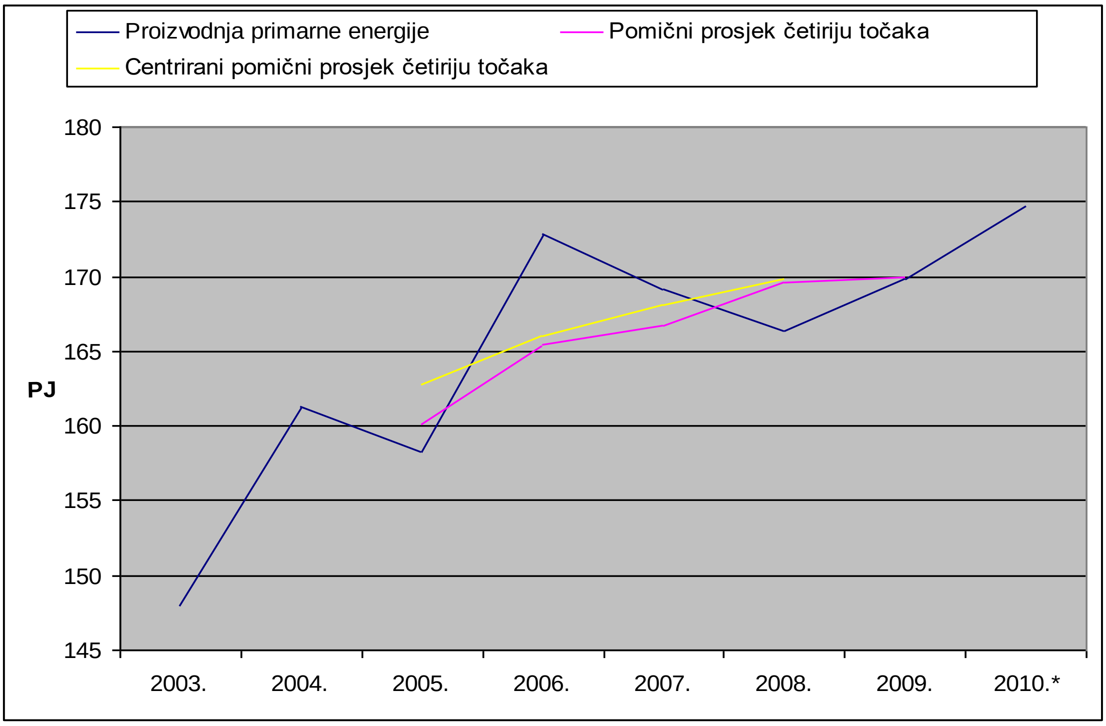
Centrirani pomični prosjek izračunat izravnim putem:

$$y_{2005}(4)=\frac{1}{4}\left[\frac{1}{2}\cdot 148+161{,}29+158{,}3+172{,}84+\frac{1}{2}\cdot 169{,}14\right]=\frac{74+161{,}29+158{,}3+172{,}84+84{,}57}{4}=162{,}75PJ$$

Izračunata vrijednost predstavlja centrirani pomični prosjek koji odgovara trećem opažanju izvornog niza.

U grafikonu 13.9. prikazan je izvorni niz, niz pomičnih prosjeka četiri točke i niz centriranih pomičnih prosjeka u kojem je vidljivo da je neregularna komponenta izglađena, a izglađenim nizom dominira uzlazni trend.

Grafikon 13.9. Izvorni niz, niz pomičnih prosjeka i niz centriranih pomičnih prosjeka



Izvor: Nacionalni akcijski plan energetske učinkovitosti RH za razdoblje do kraja 2013., preuzeto od Pisker i Radman-Funarić (2013)

13.7. Sezonski utjecaji

Na neke pojave djeluju čimbenici u istom vremenskom razmaku unutar jedne godine te izazivaju sezonske oscilacije. Zbog toga se odstupanja izazvana sezonskim oscilacijama trebaju kvantificirati.

Postupkom **desezoniranja vremenskog niza** utvrđuju se sezonski utjecaji i uklanjaju iz pojave. Na taj način se dobije slika o pojavi bez utjecaja sezone.

Postupak desezoniranja:

Izračunaju se pomični prosjeci za dvanaest mjeseci ili četiri tromjesečja na prethodno opisan način. Zbroje se vrijednosti pojave i podjele s 12 (ako su zabilježene mjesečne vrijednosti niza) ili s 4 (ako su zabilježene kvartalne vrijednosti). Potom se izračunaju centrirani pomični presjeci računanjem prosjeka susjednih parova. Prvi pomični total centrira se na sredinu razdoblja za koje se računa, npr. ako je izračunat za razdoblje od siječnja od prosinca bit će centriran između lipnja i srpnja, itd. Centrirani pomični prosjeci predstavljaju trend vrijednosti promatranog niza.

Zajednički indeks sezonskog i rezidualnog utjecaja računa se dijeljenjem originalnih vrijednosti vremenskog niza s izračunatim vrijednostima centriranih pomičnih prosjeka te množenjem sa 100.

Polazeći od multiplikativnog modela $Y = T \cdot I_s \cdot I_\varepsilon$

$$I_s \cdot I_\varepsilon = \frac{Y}{T}$$

T – trend i ciklička komponenta

I_s – sezonska komponenta

I_ε– rezidualna komponenta

sezonski i rezidualni indeks računa se prema:

$$I_t(I_s I_\varepsilon) = \frac{y_t}{y_t^*} \cdot 100$$

$I_t(I_s I_\varepsilon)$– sezonski i rezidualni indeks

y_t^* – vrijednosti centriranih pomičnih prosjeka

y_t – originalni podaci vremenskog niza

Kako bi se izračunao sezonski indeks (bez utjecaja slučajne komponente) potrebno je izračunati prosjek sezonskih i rezidualnih indeksa $øI_t(I_s I_\varepsilon)$, npr. prosjek indeksa za siječanj 2013., 2014., 2015. i 2016. godine ili prosjek indeksa za prvo tromjesečje 2013., 2014., 2015. i 2016. godine, te ih pomnožiti s korektivnim faktorom:

korektivni faktor = 1200/suma prosjek indeksa (za mjesečni niz u razdoblju od jedne godine)

ili

korektivni faktor = 400/suma prosjek indeksa (za tromjesečni niz u razdoblju od jedne godine)

Sezonski indeks:

$I_t(I_S) = \o I_t(I_S I_\varepsilon) \cdot korektivni\ faktor$

Literatura

1. Available chart types, Microsoft. URL: https://support.office.com/hr-hr/article/dostupne-vrste-grafikona-10b5a769-100d-4e41-9b0f-20df0544a683 [pristup: 15. veljače 2018.]

2. Broj i struktura poslovnih subjekata u ožujku 2017. *Priopćenje*, 9. svibnja 2017., godina LIV., broj 11.1.1/1., Državni zavod za statistiku Republike Hrvatske, Zagreb. URL: 2017.https://www.dzs.hr/Hrv_Eng/publication/2017/11-01-01_01_2017.htm [pristup: 5. siječnja 2018.]

3. Chaing, A.C. (1994), Osnovne metode matematičke ekonomije, MATE d.o.o. Zagreb, Treće izdanje, Naziv originala: Fundamental Methods of Mathematical Economics, McGraw-Hill, Inc., 1967.

4. Definiranje razvojnih prioriteta širih regija (NUTS 2), Ekonomski fakultet Sveučilišta u Zagrebu, 27. rujna 2008.URL:vladimir-cavrak.from.hr/wp-content/uploads/2015/09/Glavni_sažetak_v11a.pdf[pristup: 3. siječnja 2018.]

5. Harmonised index of consumer prices (HICP), EUROSTAT. URL: http://ec.europa.eu/eurostat/cache/metadata/EN/prc_hicp_esms_hr.htm [pristup: 12. ožujka 2018.]

6. Horvat, J. i Mijoč, J. (2014), *Osnove statistike*, Drugo dopunjeno izdanje, Ljevak, Zagreb

7. Hrvatska u brojkama 2015, Državni zavod za statistiku Republike Hrvatske, Zagreb, 2015. URL: https://www.dzs.hr/Hrv_Eng/CroInFig/croinfig_2015.pdf[pristup: 20. veljače 2018.]

8. Indeksi cijena, Hrvatska narodna banka. URL: https://www.hnb.hr/statistika/statisticki-podaci/odabrane-nefinancijske-statistike/indeksi-cijena [pristup: 12. ožujka 2018.]

9. Indeksi potrošačkih cijena u lipnju 2017.(2017) *Priopćenje*, Državni zavod za statistiku Republike Hrvatske, 17. srpnja 2017., godina LIV., broj 13.1.1/6., Zagreb.URL: https://www.dzs.hr/Hrv_Eng/publication/2017/13-01-01_06_2017.htm [pristup: 12. ožujka 2018.]

10. Istraživanje i razvoj u 2014., *Priopćenje*, Državni zavod za statistiku Republike Hrvatske, 30. listopada 2015., godina LII. , broj 8.2.1.,Zagreb.URL: https://www.dzs.hr/Hrv_Eng/publication/2015/08-02-01_01_2015.htm [pristup:12. ožujka 2018.]

11. Izvoz i investicije poduzetnika po regijama i županijama u 2013. godini, FINA, URL: www.fina.hr/fgs.axd?id=15461, [pristup: 12. ožujka 2018.]

12. Kovačić, B., Opačić, R. i Marohnić, L., O Ginijevu koeficijentu koncentracije, *math e, Hrvatski matematički elektronički časopis,* URL: https://hrcak.srce.hr/file/151776 [pristup: 17. veljače 2018.]

13. Nacionalni akcijski plan energetske učinkovitosti RH za razdoblje do kraja 2013., Ministarstvo gospodarstva. URL: http://www.mingo.hr/default.aspx?id=3251 [pristup: 15. veljače 2018.]

14. Pisker, B. iRadman-Funarić, M. (2013), Beyond the Inconsistency of the Croatian Primary Energy Production Data: A Method to Improve its Statistical Analysis, u: Jurevičienė, D. (ur.), Zbornik radova *3rd International Scientific Conference, "Whither Our economies" Conference Proceedings,* Vilnius: Mykolas Romeris University, Faculty of Economics and Finance Managemen, Vilnius, Litva, October 24-25, 2013., str. 61-68.

15. Radman-Funarić, M., Babler, V.(2013),Primjena metode najmanjih kvadrata u razdvajanju troškova, u: Šutić, B. (ur.), *1. interdisciplinarna znanstveno-stručna konferencija s međunarodnim sudjelovanjem "Održivi razvoj ruralnih krajeva"* Gospić: Veleučilište "Nikola Tesla", 269-278.

16. Radman-Funarić, M. i Kurtagić, E. (2013), Statistička analiza podataka o investicijama u zaštitu okoliša u Republici Hrvatskoj u: Čala, I. (ur.) *Maintenance 2013, Zbornik radova 19. međunarodnog savjetovanja HDO* - Hrvatsko društvo održavatelja, Croatian Maintenance Society, Šibenik, 3.-5. lipnja 2013., str. 179-186.

17. Registar poslovnih subjekata, Hrvatska gospodarska komora, URL: http://www1.biznet.hr/HgkWeb/do/extlogon;jsessionid=900DB5975375CB2C9EBAAF50FD2F0577 [pristup: 15. veljače 2018.]

18. Robna razmjena Republike Hrvatske s inozemstvom − privremeni podaci od siječnja do prosinca 2017. i za siječanj 2018. *Priopćenje*, Državni zavod za statistiku Republike Hrvatske, Zagreb, godina LIV., broj 4.2.1/12., 8. ožujka 2018.URL: https://www.dzs.hr/Hrv_Eng/publication/2017/04-02-01_12_2017.htm [pristup: 20. veljače 2018.]

19. Statistics How To, URL: http://www.statisticshowto.com/how-to-find-a-five-number-summary-in-statistics/ [pristup: 25. veljače 2018.]

20. Statistički ljetopis Republike Hrvatske 2012., Državni zavod za statistiku Republike Hrvatske, URL: http://www.dzs.hr/Hrv_Eng/ljetopis/2012/sljh2012.pdf [pristup: 10. siječnja 2018.]

21. Šošić, I. (2006), *Primijenjena statistika*, Školska knjiga, Zagreb

22. *Vrste grafikona*, Pearson Education, Inc., URL: http://pup.skole.hr/VodicHTML.aspx?xml=/datoteke/hr/hr/Student/PH/T

utorials/Understanding%20Chart%20Types/chart%20types.xml&locale=&app=&pageNumber=0 [pristup: 15. veljače 2018.]

23. Zaposlenost i plaće, Stopa registrirane nezaposlenosti i zaposlenost, Državni zavod za statistiku Republike Hrvatske, Zagreb. URL: https://www.dzs.hr/Hrv_Eng/Pokazatelji/Zaposlenost%20i%20place.xlsx [pristup: 20. veljače 2018.]