

---

# STATISTIKA

U FIZIČKOM OBRAZOVANJU I SPORTU

**Korelacija i linearna regresija**

---

# Sadržaj

---

- Korelacija
  - Izračunavanje korelacionog koeficijenta
  - Regresiona analiza
-

# KORELACIJA

## Korelacija

Izračunavanje  
korelacionog  
koeficijenta

Regresija

- **Brojna vrednost (koeficijent)** koji označava u kojoj meri su dve varijable međusobno povezane naziva se **KOEFICIJENT KORELACIJE**.

Može imati vrednosti

- od **-1.00 do +1.00**.

Koeficijent korelacije se ima svoju **brojnu vrednost** i **predznak**

Kada želimo da utvrdimo da li su dve pojave međusobno povezane – na primer, visina tela i dužina koraka – koristimo brojčani pokazatelj koji se naziva **koeficijent korelacije**.

Ovaj broj nam govori **u kojoj meri su dve varijable povezane** na osnovu podataka koje trenutno imamo. Pritom se ne zaključuje o uzroku ili promeni kroz vreme, već samo o **stepenu i smeru povezanosti u tom trenutku**.

Koeficijent može imati vrednosti od **-1.00 do +1.00**.

Pozitivne vrednosti (npr. +0.75) znače da osobe koje imaju višu vrednost jedne osobine (npr. visinu), u proseku imaju i višu vrednost druge (npr. duži korak).

Negativne vrednosti (npr. -0.60) govore da više vrednosti jedne osobine prate niže vrednosti druge – na primer, što više ponavljanja na testu snage, to manje vreme u trčanju na 20 metara.

Ako je koeficijent blizu nule, to znači da nema jasne povezanosti među posmatranim varijablama.

Dakle, svaki koeficijent korelacije ima dve važne osobine:

**brojnu vrednost**, koja pokazuje jačinu povezanosti, i **predznak**, koji pokazuje smer te povezanosti.

# KORELACIJA

## Korelacija

Izračunavanje  
korelacionog  
koeficijenta

Regresija

### Pozitivna korelacija:

- osobe koje imaju višu vrednost jedne osobine (npr. visinu), u proseku imaju i višu vrednost druge (npr. duži korak).

### Negativna korelacija:

- više vrednosti jedne osobine (npr. veća mehanička snaga nogu) prate niže vrednosti druge (vreme na sprintu 20 m)

Pozitivne vrednosti (npr.  $+0.75$ ) znače da osobe koje imaju višu vrednost jedne osobine (npr. visinu), u proseku imaju i višu vrednost druge (npr. duži korak).

Negativne vrednosti (npr.  $-0.60$ ) govore da više vrednosti jedne osobine prate niže vrednosti druge – na primer, što više ponavljanja na testu snage, to manje vreme u trčanju na 20 metara.

Ako je koeficijent blizu nule, to znači da nema jasne povezanosti među posmatranim varijablama.

# KORELACIJA

Korelacija

Izračunavanje  
korelacionog  
koeficijenta

Regresija

Iako koeficijent korelacije vrednuje stepen povezanosti između dve varijable,

- **Ne ukazuje na međusobnu uzročno-posledičnu vezu**

**Korelacija se koristi u različite svrhe**, a jedna od najčešćih je **utvrđivanje povezanosti između osobina**. Na primer, često se pokazuje da viši ljudi u proseku imaju i veću telesnu masu – to je tipičan primer pozitivne korelacije između visine i težine.

Korelacija se koristi i za **proveru validnosti merenja** – recimo, ako želimo da procenimo da li debljina kožnog nabora može da posluži kao zamena za direktno merenje telesne masti, proveravamo kolika je korelacija između ta dva pokazatelja.

Takođe, korelacija se koristi i za **predviđanje** – ako postoji dovoljno jaka povezanost između dve pojave, možemo koristiti vrednost jedne varijable za procenu druge. Na primer, školski uspeh bruća se ponekad pokušava predvideti na osnovu njihovih rezultata na prijemnim ispitima.

Međutim, iako koeficijent korelacije pokazuje **stepen povezanosti** između dve varijable, **ne ukazuje na uzročno-posledičnu vezu**.

Drugim rečima, to što su dve varijable povezane, **ne znači da promena jedne uzrokuje promenu druge**. U mnogim slučajevima može postojati **treći faktor** koji utiče na obe posmatrane pojave i stvara prividnu vezu među njima.

Zato je pri tumačenju rezultata veoma važno da ne izvodimo pre nagljene zaključke o uzrocima, već da korelaciju shvatimo kao **signal da među varijablama postoji neka veza** koju tek treba dublje istražiti.

# KORELACIJA

## Korelacija

Izračunavanje  
korelacionog  
koeficijenta

Regresija

Vizuelni prikaz koeficijenta korelacije može se predstaviti kao **tačkasti dijagram**.

- rezultati svakog ispitanika za dve varijable prikazuju tako što se jedna varijabla postavlja na X-osu, a druga na Y-osu.

Da bismo bolje razumeli odnos između dve varijable, često koristimo **grafički prikaz koji se naziva tačkasti dijagram**.

Na ovom dijagramu, **svaki ispitanik je predstavljen jednom tačkom**, čija je **horizontalna pozicija određena vrednošću prve varijable (X-os)**, a **vertikalna pozicija vrednošću druge varijable (Y-os)**.

Na osnovu rasporeda tačaka možemo vizuelno proceniti da li postoji neka pravilnost – na primer, da li se tačke grupišu uz zamišljenu liniju, da li idu uzlazno, silazno, ili deluju nasumično rasute.

Tačkasti dijagram nam, dakle, omogućava da **brzo steknemo uvid u pravac i jačinu veze između**

**varijabli**, pre nego što se upustimo u izračunavanje koeficijenta korelacije.

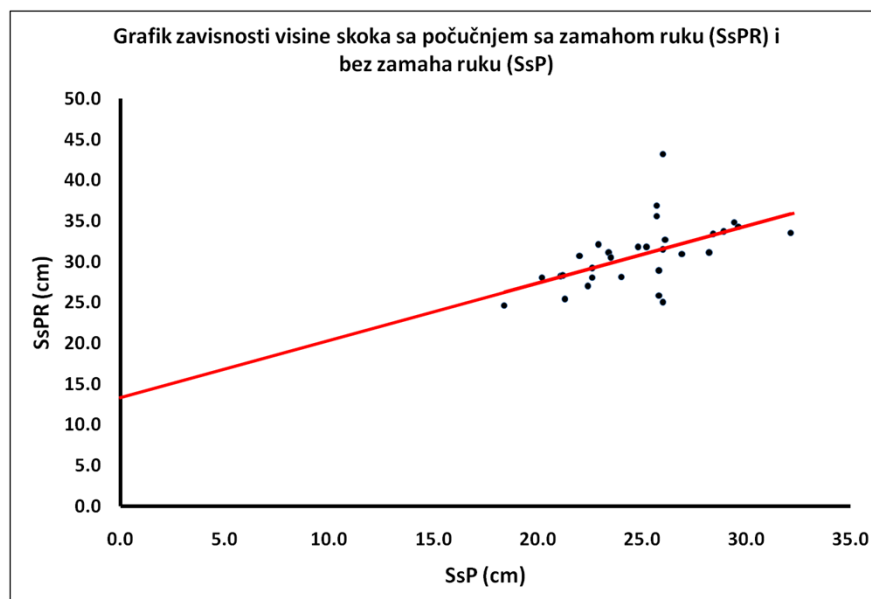
# KORELACIJA

## Korelacija

Izračunavanje  
korelacionog  
koeficijenta

Regresija

## Tačkasti dijagram (**Pozitivna korelacija**)



Na ovom primeru vidimo kako izgleda **pozitivna korelacija** prikazana na tačkastom dijagramu. Svaka zvezdica predstavlja rezultat jednog ispitanika — kombinaciju skoka iz čučnja sa zamahom ruku i skoka iz čučnja bez zamaha.

Jedna osa prikazuje rezultate **skoka iz čučnja bez zamaha ruku**, a druga **skoka iz čučnja sa zamahom ruku**. Možemo primetiti da kako raste rezultat u jednom testu, raste i u drugom — tačke se postepeno penju s leva na desno.

Kroz ovaj skup podataka povučena je **linija najbolje prilagođenosti** (best fit line), koja pokazuje opšti trend — bolje performanse u jednom testu prate bolje rezultate u drugom.

Ovakav obrazac rasporeda ukazuje na **snažnu pozitivnu povezanost** između ova dva testa eksplozivne snage donjih ekstremiteta.

Ipak, i ovde važi upozorenje — iako je veza očigledna, dijagram **ne govori ništa o uzroku**; on samo prikazuje povezanost između rezultata u datom trenutku.

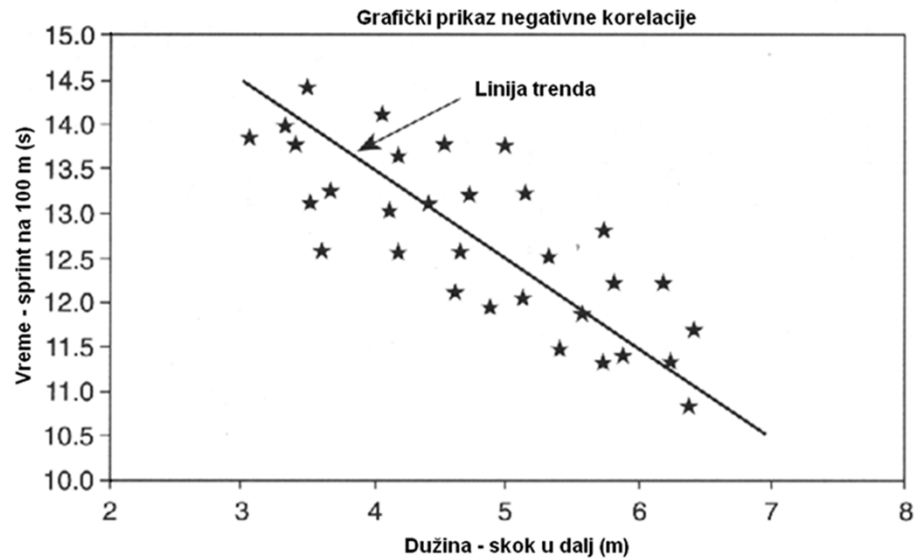
# KORELACIJA

## Korelacija

Izračunavanje  
korelacionog  
koeficijenta

Regresija

## Tačkasti dijagram (**Negativna korelacija**)



Za razliku od prethodnog primera, ovde vidimo **negativnu korelaciju**. Na tačkastom dijagramu prikazana je povezanost između **dužine skoka u dalj** i **vremena trčanja na 100 metara**.

Svaka tačka predstavlja uređeni par rezultata za jednog ispitanika – njegov rezultat na testu skoka i rezultat sprinta. Linija trenda pokazuje da kako raste dužina skoka u dalj, vreme potrebno za sprint se smanjuje.

Drugim rečima, sportisti koji dalje skaču obično imaju **kraće vreme trčanja** – što je i očekivano, jer su eksplozivnost i snaga donjih ekstremiteta zajednički faktori oba testa.

Ova negativna korelacija znači da **veće vrednosti jedne varijable prate manje vrednosti druge**. Ipak, kao i ranije, važi isto pravilo – **ova veza ne potvrđuje uzročno-posledični odnos**, već samo pokazuje da rezultati idu u suprotnim smerovima u datom skupu podataka.

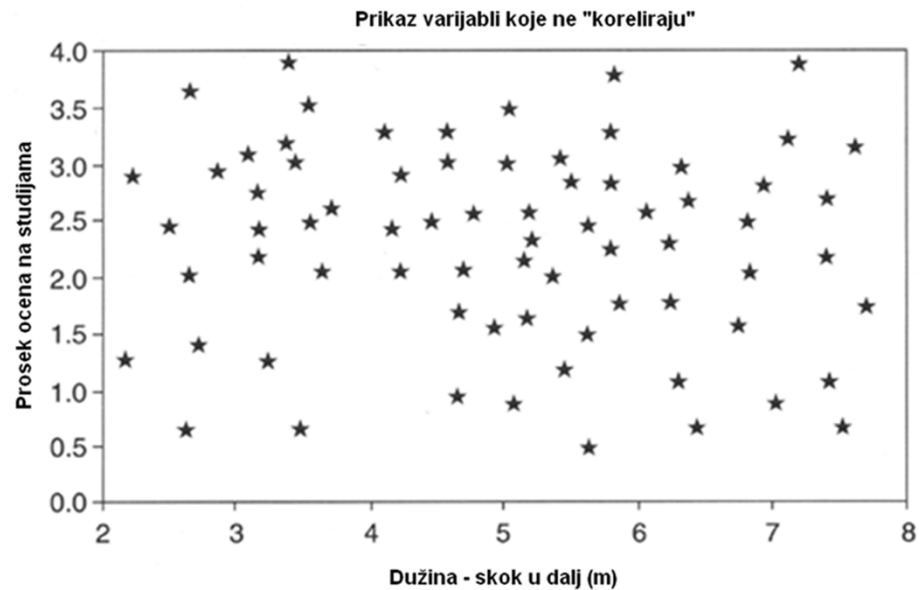
# KORELACIJA

## Korelacija

Izračunavanje  
korelacionog  
koeficijenta

Regresija

## Tačkasti dijagram (**Slaba povezanost**)



Konačno, vidimo situaciju kada **ne postoji korelacija** između dve varijable. Na tačkastom dijagramu prikazana je veza između **dužine skoka u dalj** i **prosečne ocene na studijama**.

Raspored tačaka deluje potpuno nasumično — nema jasnog trenda rasta niti opadanja. Tačke su razbacane bez pravca, što znači da **ne postoji konzistentna veza** između rezultata na skoku i akademskog uspeha.

U statističkom smislu, koeficijent korelacije bi ovde bio blizu nule, što označava **odsustvo značajne povezanosti**. Drugim rečima, sa ovakvim podacima **ne možemo predviđati ni ocene na osnovu skoka, ni skok na osnovu ocena**.

Ovo je važan podsetnik da ne postoji korelacija među svim varijablama, čak i kada se radi o istim ispitanicima. Zato je **analiza korelacije važan alat za proveru pretpostavki** — ne samo da bismo potvrdili postojanje veze, već i da bismo prepoznali kada je **nema**.

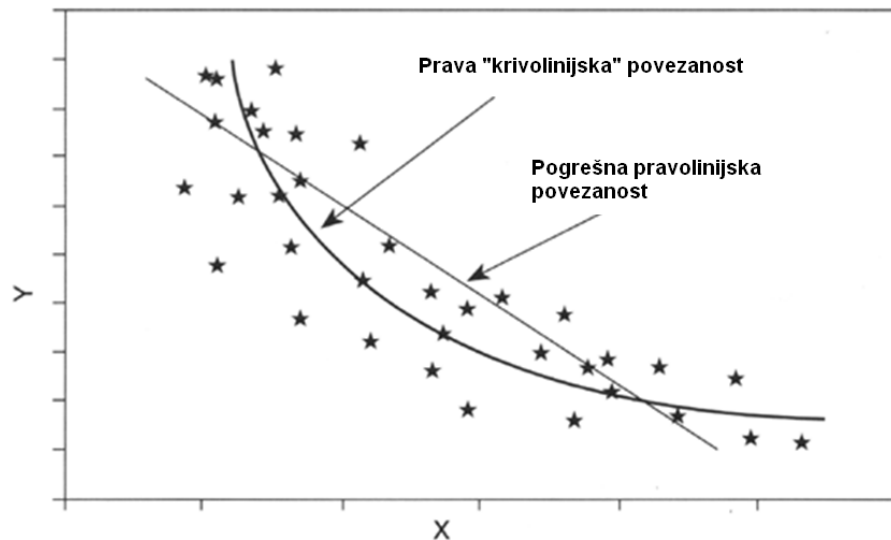
# KORELACIJA

## Korelacija

Izračunavanje  
korelacionog  
koeficijenta

Regresija

## Poređenje linearne i nelinearne povezanosti



Odnos između dve varijable **ne mora uvek biti pravolinijski (linearan)**. Iako se najčešće koristi **linijska korelacija**, postoje situacije kada je veza među varijablama **zakrivljena** – odnosno, **nelinearna**.

Na slici vidimo da su podaci raspoređeni tako da bolje odgovaraju **krivolinijskoj liniji trenda**, dok jednostavna prava ne uspeva da precizno prikaže obrazac u podacima.

Ukoliko bismo koristili **pravolinijsku korelaciju u ovakvom slučaju**, mogli bismo doći do **pogrešnog zaključka** – da je veza slaba ili čak nepostojeća, iako je zapravo vrlo izražena, ali u nelinearnom obliku.

Zato je važno uvek vizualizovati podatke pre izračunavanja korelacije, kako bismo uočili **da li je veza zaista linearna** ili postoji **krivolinijski (nelinearan) obrazac**, koji zahteva drugačiji statistički pristup.

# IZRAČUNAVANJE KORELACIONOG KOEFICIJENTA

Korelacija

Izračunavanje  
korelacionog  
koeficijenta

Regresija

## DEFINICIJA KORELACIONOG KOEFICIJENTA

$$r = \frac{\sum [(z_x)(z_y)]}{N}$$

$Z_x$  i  $Z_y$  predstavljaju Z vrednosti odgovarajućih varijabli

- Prvo izračunamo srednje vrednosti standardne devijacije za obe varijable.
- Zatim svaki pojedinačan rezultat (za obe varijable) pretvorimo u Z vrednost. Pomnožimo odgovarajuće Z vrednosti.
- Izračunamo sumu svih proizvoda pa je podelimo sa brojem parova ( $N$ )

Do sada smo se bavili **grafičkim prikazima** korelacije i upoznali se sa različitim oblicima povezanosti među varijablama — pozitivnim, negativnim, nepostojećim, pa čak i nelinearnim.

Ali kako tu povezanost možemo izraziti **brojem**, tj. **izračunati koeficijent korelacije**?

Jedan od najčešće korišćenih načina je preko tzv. **standardizovanih vrednosti** – Z vrednosti.

U ovoj formuli,  $Z_x$  i  $Z_y$  predstavljaju **Z vrednosti pojedinačnih rezultata** za obe varijable.

To znači da svaki rezultat prethodno pretvaramo u standardni oblik (oduzimanjem srednje vrednosti i deljenjem sa standardnom devijacijom), čime sve podatke dovodimo na zajedničku meru.

Zatim množenjem odgovarajućih Z vrednosti i sabiranjem svih proizvoda, dobijamo meru povezanosti – a deljenjem s brojem parova dobijamo **koeficijent korelacije ( $r$ )**.

Ukoliko je  $r$  blizu +1 – veza je jaka i pozitivna, ako je blizu -1 – jaka i negativna, a ako je blizu nule – nema povezanosti.

U nastavku ćemo videti **konkretan primer kako se ovo računa** korak po korak.

# IZRAČUNAVANJE KORELACIONOG KOEFICIJENTA

Korelacija

Izračunavanje  
korelacionog  
koeficijenta

Regresija

## DEFINICIJA KORELACIONOG KOEFICIJENTA

### Primer računanja

| $x$ | $y$ |
|-----|-----|
| 2   | 1   |
| 3   | 2   |
| 5   | 3   |
| 5   | 4   |

Here is the example from the textbook. We have a  $x$  and an  $y$  variables, with their respective values.

# IZRAČUNAVANJE KORELACIONOG KOEFICIJENTA

Korelacija

Izračunavanje  
korelacionog  
koeficijenta

Regresija

## DEFINICIJA KORELACIONOG KOEFICIJENTA

### Primer računanja

| x | y |
|---|---|
| 2 | 1 |
| 3 | 2 |
| 5 | 3 |
| 5 | 4 |

Da bismo videli kako sve to izgleda u praksi, posmatrajmo konkretan primer.

U tabeli su dati rezultati za dve varijable:

varijabla x (npr. broj ponavljanja u testu snage)

varijabla y (npr. broj uspešnih pokušaja u preciznosti)

# IZRAČUNAVANJE KORELACIONOG KOEFICIJENTA

Korelacija

Izračunavanje  
korelacionog  
koeficijenta

Regresija

## DEFINICIJA KORELACIONOG KOEFICIJENTA

### Primer računanja

|    | A                  | B  | C              | D    | E                  | F    | G     | H        | I | J |
|----|--------------------|----|----------------|------|--------------------|------|-------|----------|---|---|
| 1  |                    |    |                |      |                    |      |       |          |   |   |
| 2  |                    |    |                |      |                    |      |       |          |   |   |
| 3  |                    |    |                |      |                    |      |       |          |   |   |
| 4  |                    |    |                |      |                    |      |       |          |   |   |
| 5  |                    |    |                |      |                    |      |       |          |   |   |
| 6  |                    |    |                |      |                    |      |       |          |   |   |
| 7  | Srednje vrednosti: |    | 3.75           | 2.5  |                    |      |       |          |   |   |
| 8  | St. devijacije:    |    | 1.30           | 1.12 |                    |      |       |          |   |   |
| 9  |                    |    |                |      |                    |      |       |          |   |   |
| 10 | $\Sigma x =$       | 15 | $\Sigma x^2 =$ | 63   | $\Sigma Z_x Z_y =$ | 3.79 | $r =$ | 0.946729 |   |   |
| 11 | $\Sigma y =$       | 10 | $\Sigma y^2 =$ | 30   |                    |      | $r =$ | 0.946729 |   |   |
| 12 |                    |    |                |      |                    |      |       |          |   |   |

Prvo su za svaku varijablu izračunate srednje vrednosti i standardne devijacije. Zatim su pojedinačni rezultati pretvoreni u **Z vrednosti** – što znači da su standardizovani.

Kada pomnožimo Z vrednosti odgovarajućih parova ( $Z_x \times Z_y$ ) i saberemo sve te proizvode, dobijamo ukupnu sumu, koja se potom **podeli sa brojem parova (N = 4)**.

Na kraju dobijamo koeficijent korelacije:

**$r = 0.946729$** , što ukazuje na **vrlo jaku pozitivnu korelaciju** između posmatranih varijabli.

Ovaj broj nam jasno govori da osobe koje imaju više rezultate na jednoj varijabli, u pravilu imaju i više rezultate na drugoj – što se i vizuelno može potvrditi ako se podaci prikažu na tačkastom dijagramu.

# IZRAČUNAVANJE KORELACIONOG KOEFICIJENTA

Korelacija

Izračunavanje  
korelacionog  
koeficijenta

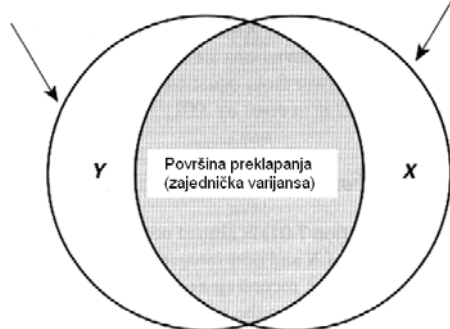
Regresija

## Procena jačine korelacije

Figure 7.5 Venov dijagram za bivarijantnu relaciju

Zavisna varijabla (ona koja  
čije se vrednosti  
predviđaju)

Nezavisna varijabla  
(prediktorska)



**Koeficijent determinacije ( $r^2$ )**

**Površina preklapanja ( $r^2$ )**

- predstavlja **veličinu varijanse zavisno promenljive** može se objasniti varijansom nezavisne, varijable

Copyright © 2012 by William J. Vincent and Joseph P. Weir

Kada posmatramo dve varijable, koeficijent korelacije  $r$  nam govori **da li su te dve pojave međusobno povezane i koliko snažno**.

Intuitivno, znamo sledeće:

Ako je  $r$  blizu  $+1$ , znači da postoji **vrlo jaka pozitivna povezanost** – kako raste jedna varijabla, raste i druga.

Ako je  $r$  blizu  $-1$ , to je znak **vrlo jake negativne povezanosti** – kada jedna raste, druga opada.

Ako je  $r$  blizu  $0$ , to znači da **nema uočljive veze** među njima.

Dakle, što je  $r$  bliži jedinici po apsolutnoj vrednosti (bilo  $+1$  ili  $-1$ ), to je **veza među varijablama jača i predviđanje pouzdanije**.

Ali i dalje –  $r$  je samo **pokazatelj pravca i jačine veze**.

Da bismo dobili **jasniju predstavu o tome koliko X pomaže da razumemo Y**, koristimo **koeficijent determinacije** –  $R^2$ .

Zamislimo da varijabla  $Y$  predstavlja nešto što želimo da predvidimo – recimo **rezultat na testu izdržljivosti**, a varijabla  $X$  nešto što koristimo kao prediktor – na primer, **broj sati treninga nedeljno**.

Korelacija ( $r$ ) nam govori **da li postoji povezanost između njih**, a **koeficijent determinacije ( $R^2$ )** nam kaže **koliko ta povezanost pomaže u objašnjenju rezultata na Y**.

Na ovom dijagramu, kružnici predstavljaju **ukupnu varijabilnost** svake

varijable.

Kružnica **Y** je sve što utiče na rezultat izdržljivosti.

Kružnica **X** je sve što se tiče sati treninga.

**Preklapanje prikazuje zajednički deo – ono što X može da objasni u Y.**

U ovom primeru,  $R^2 = 0.64$ , što znači da **64% varijacija u rezultatima testa može da se objasni razlikama u broju sati treninga**. Drugim rečima, X (trening) dosta dobro objašnjava Y (rezultat).

Ali preostalih **36% ( $1 - R^2$ )** je deo koji ne možemo objasniti varijablom X. To može biti posledica nekih drugih faktora – genetike, sna, ishrane, motivacije, itd.

**Zato što je veći  $R^2$ , to više znamo o Y kad znamo X – i tim je preciznije predviđanje.**

# IZRAČUNAVANJE KORELACIONOG KOEFICIJENTA

Korelacija

Izračunavanje  
korelacionog  
koeficijenta

Regresija

## Procena jačine korelacije

Za predviđanja, apsolutna vrednost  $r < 0.7$  ( $R^2=0.49$ ), može uzrokovati neprihvatljivo “velike” greške, posebno ako su Standardne Devijacije (SD) obe veličine i X & Y velike.

Opšta pretpostavka (**jedna od mogućih skala**):

- apsolutna vrednost  $r \geq 0.9$  – “visoka”
- apsolutna vrednost  $r = 0.7 - 0.8$  – “umerena”
- apsolutna vrednost  $r = 0.5 - 0.7$  – “niska”
- Vrednosti  $r \leq 0.5$  daju vrednost  $R^2 = 0.25$ , ili 25%, što “nije dovoljno za pouzdano predviđanje.”

Do sada smo govorili o tome **kako se korelacija računa** i šta ona matematički predstavlja.

Ali sada se postavlja praktično pitanje:

**Koliko velika vrednost koeficijenta korelacije je „dovoljno dobra“ da bi bila korisna?**

Ovo je posebno važno kada želimo da koristimo jednu varijablu **za predviđanje** druge. Na primer, da li na osnovu broja ponavljanja možemo predvideti rezultat na testu izdržljivosti?

U tom slučaju, **što je veća apsolutna vrednost koeficijenta  $r$** , to su varijable **jače povezane** i predikcija **pouzdanija**.

Jedna od najčešće korišćenih interpretacionih skala kaže:

Ako je  $|r| \geq 0.9$ , veza se smatra **visokom**

Ako je  $|r|$  između **0.7 i 0.8**, veza je **umerena**

Ako je  $|r|$  između **0.5 i 0.7**, veza je **slaba**

Ako je  $|r| \leq 0.5$ , veza je **vrlo slaba ili neupotrebljiva** za predikciju

U donjoj zoni vidimo praktičnu posledicu:

Ako je  $r = 0.5$ , tada je  $R^2 = 0.25$ , što znači da **samo 25% varijanse u jednoj varijabli možemo objasniti drugom**.

U praksi, to obično **nije dovoljno za validno predviđanje**, naročito ako su obe varijable vrlo promenljive (imaju veliku standardnu devijaciju).

Zato u predikcijama često tražimo da korelacija bude **najmanje 0.7**, kako bi

greška bila prihvatljiva.

# IZRAČUNAVANJE KORELACIONOG KOEFICIJENTA

Korelacija

Izračunavanje  
korelacionog  
koeficijenta

Regresija

## Procena jačine korelacije

- Upotrebljivost koeficijenta korelacije određeno je njegovom veličinom i da li je ta vrednost **statistički značajna**

### *Veličina uzorka:*

Broj parova (N) je presudan za procenu kolika je verovatnoća da dobijeni odnos nije nastao slučajno.

- Kada je N mali, r mora biti velik da bi bio statistički značajan.
- Kada je N veliki, čak i male vrednosti r mogu biti značajne.

Do sada smo govorili o tome **koliko je važna vrednost koeficijenta korelacije (r)**. Ali, jednako važno je i **koliko imamo podataka** – odnosno, **koliki je uzorak**.

Broj posmatranih parova (oznaka N) igra **ključnu ulogu u tumačenju značajnosti korelacije**.

Ako imamo **mali broj ispitanika**, slučajnosti lako mogu da naruše rezultate — zato u tom slučaju **r mora biti veoma visok** da bismo mogli da kažemo da je veza „stvarna“, a ne posledica slučajnosti.

S druge strane, ako imamo **veliki uzorak**, čak i **manje vrednosti r mogu biti statistički značajne** – jer imamo dovoljno podataka da budemo sigurni da veza postoji, iako je slaba.

Zato je važno da prilikom interpretacije r uvek uzmemo u obzir:

**Koliki je uzorak?**

**Da li je korelacija statistički značajna?**

**I, što je još važnije – da li je i praktično značajna?**

U sportu i praksi, često nam ne znači mnogo ako je neka korelacija „samo statistički značajna“ – važno je da ima **dovoljnu jačinu i smisao za donošenje odluka**.

# REGRESIJA

Korelacija

Izračunavanje  
korelacionog  
koeficijenta

Regresija

Kada je korelacija između dve varijable dovoljno “velika” (visoka),

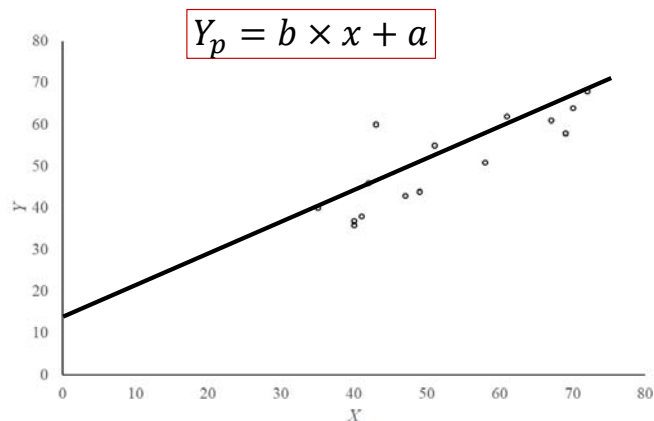
- na osnovu rezultata varijable **X** (nezavisno promenljive) možemo predvideti rezultat varijable **Y** (zavisno promenljive).

**b- nagib r. prave**

$$b = \frac{r \sigma_Y}{\sigma_X}$$

**a- odsečak na y osi**

$$a = \bar{Y} - \frac{r \sigma_Y}{\sigma_X} \bar{X} = \bar{Y} - b \bar{X}$$



## Regresija: Predviđanje na osnovu poznate veze

Ako znamo da postoji **dovoljno jaka korelacija** između dve promenljive, možemo da idemo korak dalje — i da **predviđamo** vrednosti jedne promenljive (**Y**) na osnovu poznatih vrednosti druge (**X**). Taj postupak zovemo **regresija**.

Osnovna **regresiona jednačina** ima oblik:

$Y_p = b \times x + a$  gde je:

**Y<sub>p</sub>** – predviđena vrednost zavisne promenljive (**Y**)

**xxx** – poznata vrednost nezavisne promenljive (**X**)

**bbb** – **nagib prave** (kaže nam za koliko se promeni **Y** kada **X** poraste za 1)

**aaa** – **odsečak na Y-osi**, tj. vrednost **Y** kada je **X** = 0

Na ovom slajdu su prikazane i formule za izračunavanje nagiba (**b**) i odsečka (**a**) na osnovu korelacije i standardnih devijacija **X** i **Y**.

Primer sa grafikona pokazuje da se većina tačaka dobro uklapa uz pravu — što znači da je veza snažna, pa je i predviđanje dovoljno precizno.

# REGRESIJA

Korelacija

Izračunavanje  
korelacionog  
koeficijenta

Regresija

Kada je korelacija između dve varijable dovoljno “velika” (visoka),

- na osnovu rezultata varijable  $X$  (nezavisno promenljive) možemo predvideti rezultat varijable  $Y$  (zavisno promenljive).

Primer:

- Dva testa mišićne izdržljivosti (15 srednjoškolaca-rvača)
- **Test 1 ( $X$ )**: Broj sklekova u dva minuta
- **Test 2 ( $Y$ )**: Vreme izdržaja 50 % maksimalne jačine stiska šake (do otkaza)

$$r=0.845 (0.59 \div 0.95)$$

Regresija se zasniva na ideji da, ako je veza dovoljno snažna, onda **na osnovu poznate vrednosti nezavisne varijable ( $X$ )** možemo **proceniti očekivanu vrednost zavisne varijable ( $Y$ )**.

U ovom primeru posmatramo dva testa mišićne izdržljivosti kod grupe srednjoškolaca – rvača:

**Test 1 ( $X$ )**: Broj skleka za 2 minuta

**Test 2 ( $Y$ )**: Vreme izdržaja pri 50% maksimalne jačine stiska šake (do otkaza)

Dobijeni koeficijent korelacije je  **$r = 0.845$** , što ukazuje na **visoku povezanost** između rezultata ova dva testa.

# REGRESIJA

Korelacija

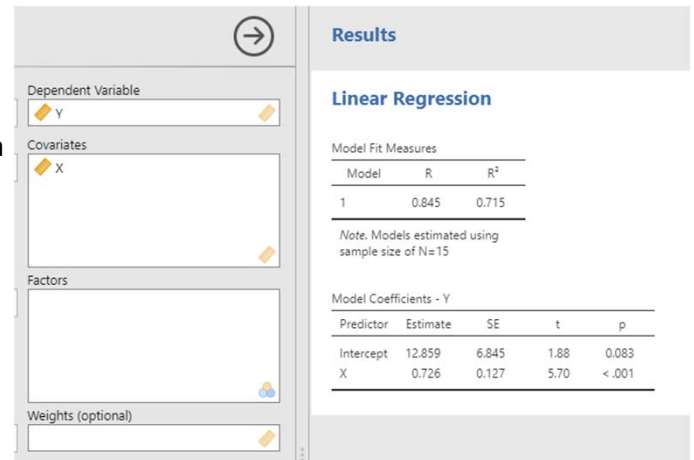
Izračunavanje  
korelacionog  
koeficijenta

Regresija

Primer:

- Dva testa mišićne izdržljivosti (15 srednjoškolaca-rvača)
- **Test 1 (X)**: Broj sklekova u dva minuta
- **Test 2 (Y)**: Vreme izdržaja 50 % maksimalne jačine stiska šake (do otkaza)

$$r=0.845 (0.59 \div 0.95)$$



$$Y_p = 0.726 \times x + 12.86$$

Na ovom primeru pokazano je kako možemo **izračunati i koristiti regresionu jednačinu** za predviđanje.

Imamo dva testa mišićne izdržljivosti, sprovedena na grupi od 15 srednjoškolaca – rvača:

**Test 1 (X)**: Broj skleka u 2 minuta

**Test 2 (Y)**: Vreme izdržaja pri 50% maksimalne jačine stiska šake

Na osnovu dobijenih rezultata i visoke korelacije (**r = 0.845**), možemo formirati **regresionu jednačinu** oblika:

$$Y_p = b \cdot x + a$$

U ovom primeru:

Nagib (**b**) je približno **0.72**, što znači da se za svaki dodatni sklek, očekivano vreme izdržaja poveća za oko 0.72 sekunde

Odsečak na y-osi (**a**) je oko **13 sekundi** – to bi bilo predviđeno vreme izdržaja kada bi broj skleka bio nula

# REGRESIJA

Korelacija

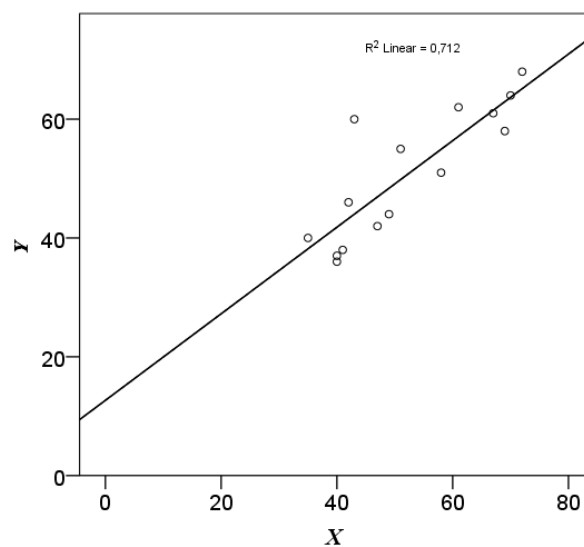
Izračunavanje  
korelacionog  
koeficijenta

Regresija

Primer:

- Dva testa mišićne izdržljivosti (15 srednjoškolaca-rvača)
- **Test 1 (X)**: Broj sklekova u dva minuta
- **Test 2 (Y)**: Vreme izdržaja 50 % maksimalne jačine stiska šake (do otkaza)

$$r=0.845 \quad (0.59 \div 0.95)$$



# REGRESIJA

Korelacija

Izračunavanje  
korelacionog  
koeficijenta

Regresija

Primer:

- Dva testa mišićne izdržljivosti (15 srednjoškolaca-rvača)
- **Test 1 (X)**: Broj sklekova u dva minuta
- **Test 2 (Y)**: Vreme izdržaja 50 % maksimalne jačine stiska šake (do otkaza)

$$r=0.845 (0.59 \div 0.95)$$

Model Fit Measures

| Model | R     | R <sup>2</sup> |
|-------|-------|----------------|
| 1     | 0.845 | 0.715          |

Note: Models estimated using sample size of N=15

Model Coefficients - Y

| Predictor | Estimate | SE    | t    | p      |
|-----------|----------|-------|------|--------|
| Intercept | 12.859   | 6.845 | 1.88 | 0.083  |
| X         | 0.726    | 0.127 | 5.70 | < .001 |

$$Y_p = 0.726 \times x + 12.86$$

Ako je učenik koji je uradio **43 sklekova (X = 43)**

$$Y_p = 0.726 \text{ s} \times 43 + 12.86 \text{ s}$$

$$Y_p = 44.1 \text{ s} \approx 44 \text{ s}$$

Predviđeno vreme bi bilo **44 sekunde**

Za ilustraciju je prikazan **odabrani podatak**:

Učenik je uradio **43 sklekova (X = 43)**

Prema regresionoj jednačini, **predviđeno vreme (Y<sub>p</sub>)** za 43 sklekova je **44.1 sekundi**

# REGRESIJA

Korelacija

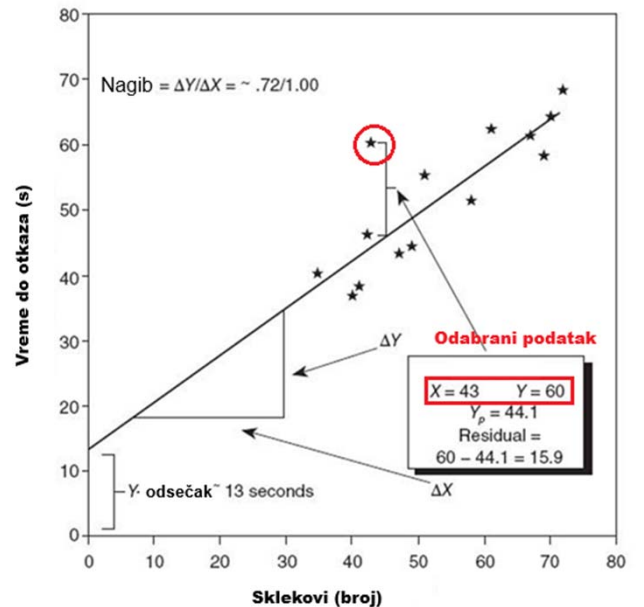
Izračunavanje  
korelacionog  
koeficijenta

Regresija

Primer:

- Dva testa mišićne izdržljivosti (15 srednjoškolaca-rvača)
- **Test 1 (X)**: Broj sklekova u dva minuta
- **Test 2 (Y)**: Vreme izdržaja 50 % maksimalne jačine stiska šake (do otkaza)

$$r=0.845 \ (0.59 \div 0.95)$$



Prema regresionoj jednačini, **predviđeno vreme ( $Y_p$ )** za 43 sklekova je **44.1 sekundi**

Razlika između stvarne i predviđene vrednosti naziva se **rezidual**:

$$\text{Rezidual} = Y - Y_p = 60 - 44.1 = 15.9$$

Rezidual nam pokazuje **koliko je stvarni rezultat odstupio od onoga što je model predvideo**.

U idealnom slučaju, svi bi reziduali bili blizu nule – ali u praksi, oni ukazuju na **grešku modela** i pomažu da procenimo **tačnost predviđanja**.

# REGRESIJA

---

## Litaratura:

- Vincent, W. J., & Weir, J. P. (2012). Statistics in kinesiology. Champaign: Human Kinetics.

## Literatura

Vincent, W. J., & Weir, J. P. (2012). *Statistics in kinesiology* (4th ed.). Human Kinetics.