

VIŠESTRUKA LINEARNA REGRESIJA



PREDAVANJE BR.14

REGRESIONA ANALIZA

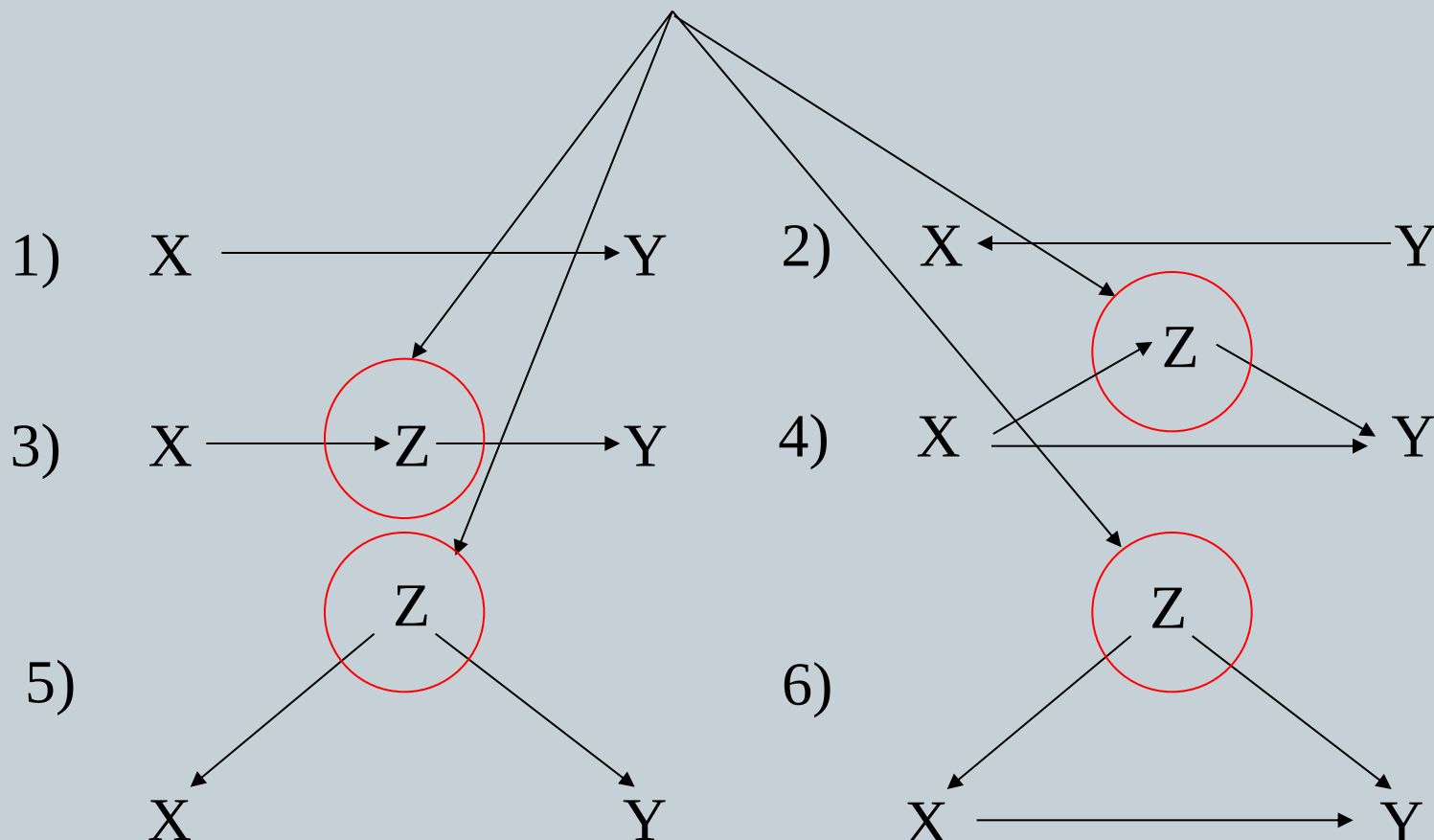


- Predviđanje jedne numeričke karakteristike preko druge numeričke karakteristike
- Ispitivanje zavisnosti jednog parametra od drugog
- Jedna pojava – jedan prediktor – jedan uzrok?
- Većina pojava uslovljena sa više faktora!

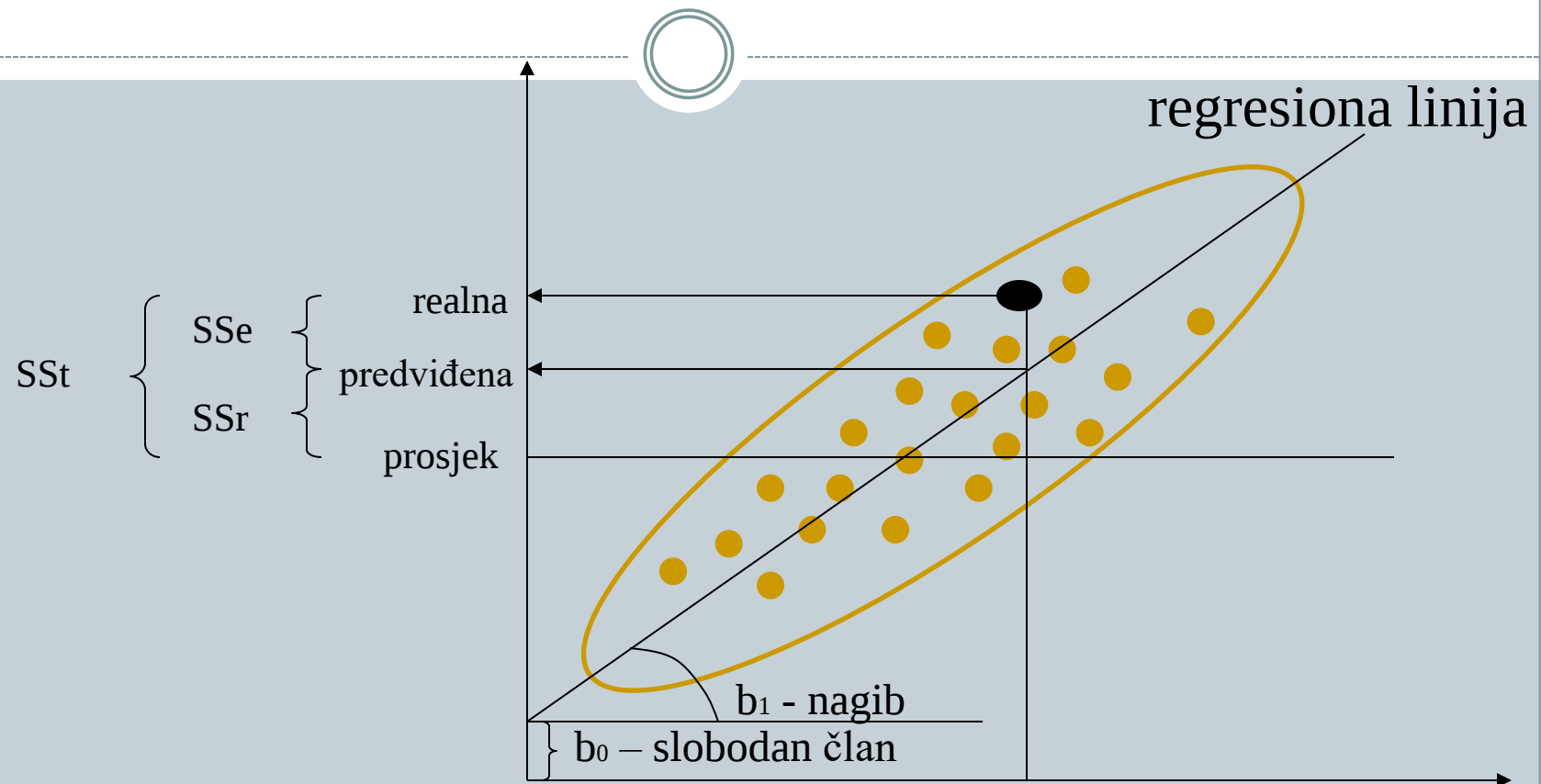
Podsjetnik o mogućim korelacijama između X i Y?



konfundirajuća varijabla



Podsjetnik o predikciji



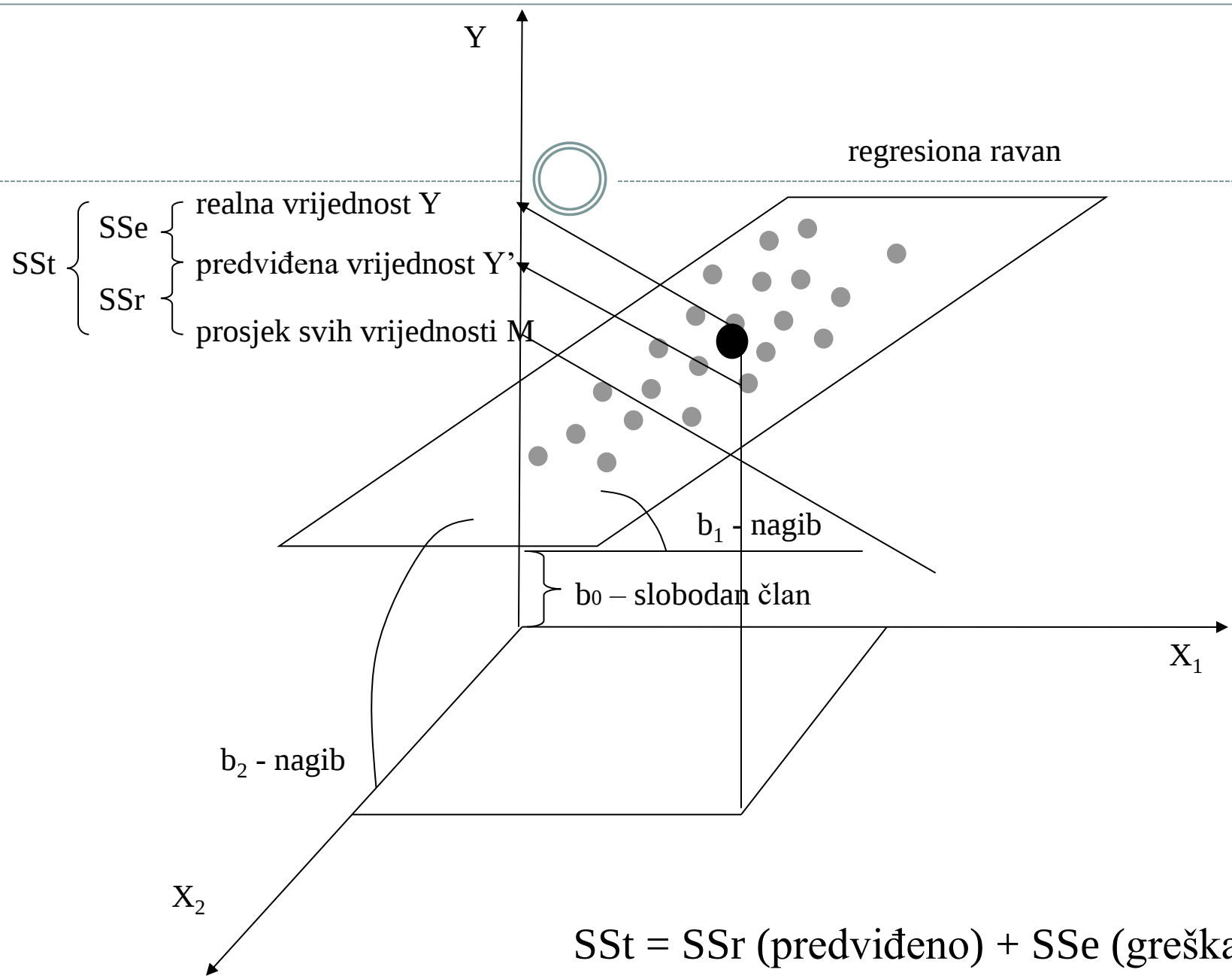
$$Y = b_1 * X + b_0$$

$$SSt = SSr \text{ (predviđeno)} + SSe \text{ (greška)}$$

Višestruka regresija



- Korelacioni multivarijantni nacrt
- Dva i više prediktora (numeričkih)
- Jedan kriterijum (numerički)
- Naći takvu kombinaciju prediktora koja daje najbolje predviđanje kriterijuma
- Prednosti:
 - Ispitivanje predviđanja jedne pojave preko više prediktora istovremeno
 - Bolja predikcija zavisne pojave
 - Međuzavisnost prediktora?



$$SSt = SSr \text{ (predviđeno)} + SSe \text{ (greška)}$$

$$Y = b_1 * X_1 + b_2 * X_2 + b_0$$

$$SS_t = SS_r + SS_e$$

$$R^2 = \frac{SS_r}{SS_t}$$



- R^2 – višestruki koeficijent determinacije
 - globalni (glavni) efekat
 - stepen predviđanja vrijednosti zavisne promjenljive preko svih prediktora zajedno
 - Npr. ako je $R^2 = 0.9$, to znači da 90 % zavisne varijable (Y) možemo predvidjeti (objasniti) na osnovu svih prediktora zajedno
 - Determinacija = kvadrirana korelacija
- I dalje je moguće odrediti pojedinačne doprinose
 - Npr r_1^2 (doprinos prvog prediktora) i r_2^2 (doprinos drugog prediktora)

Korekcija determinacije



- Koeficijent višestruke determinacije izražen preko pojedinačnih korelacija

$$R^2 = b_1 r_{yx1} + b_2 r_{yx2} + \dots + b_n r_{yxn}$$

- Koeficijent višestruke determinacije precjenjuje populacijsku vrijednost
 - ✦ Precjenjivanje raste sa porastom broja prediktora
- Korigovani koeficijent determinacije:

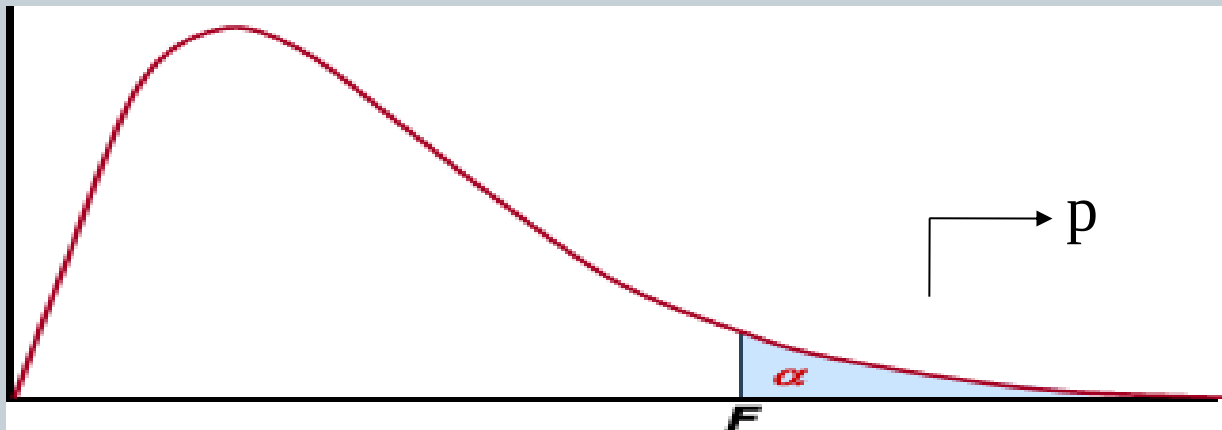
$$R_{kor}^2 = 1 - \frac{n-1}{n-m-1} (1 - R^2)$$

Značajnost koeficijenta višestruke determinacije



$$f = \frac{R^2 / m}{(1 - R^2) / (n - m - 1)}$$

- Fišerov F statistik
- Snidikorova distribucija uzorkovanja



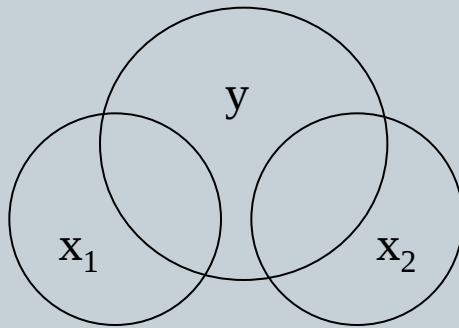
H_0 – nema predviđanja; $p < 0.05$ – odbacujemo H_0

Odnos pojedinačnih i višestrukih efekata



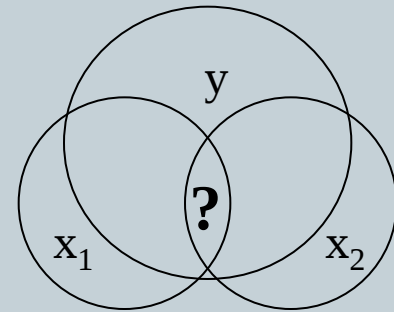
- Prediktori nisu međusobno korelisani

$$r_{x_1x_2} = 0$$



$$R^2 = r_1^2 + r_2^2$$

- Prediktori jesu međusobno korelisani $r_{x_1x_2} \neq 0$



$$R^2 \neq r_1^2 + r_2^2$$

Regresioni koeficijenti



- “b” koeficijenti - doprinos svakog pojedinačnog prediktora
- “ β ” - standardizovani koeficijenti (moguće upoređivanje doprinosa prediktora)

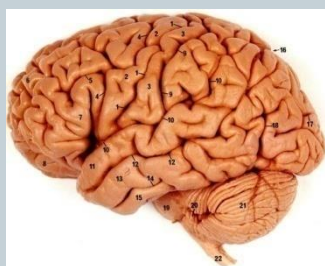
$$y = b_1 \cdot x_1 + b_2 \cdot x_2 + b_3 \cdot x_3 \dots + b_0$$

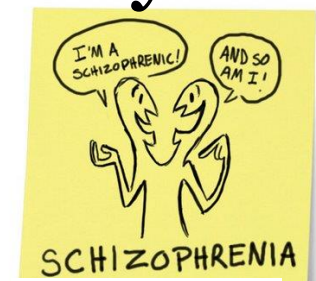
- Linearna kombinacija
- Koliko prediktora – toliko elemenata jednačine (pondera) + intercept

Regresija kao recept*

- Prediktori – sastojci
- Ponderi određuju količinu sastojaka


$$b_0 + b_1 * x_1 + b_2 * x_2 \dots = y$$



$$b_1 *$$

$$b_2 *$$


Parcijalne i semiparcijalne korelacije

- Povezanosti svakog pojedinačnog prediktora sa kriterijumom
 - Parcijalne - korelacija prediktora sa kriterijumom kada su vrijednosti ostalih prediktora konstantne
 - Semiparcijalne – dodatni doprinos prediktora nakon predikcije na osnovu ostalih prediktora

semiparcijalna
korelacija x_1 i y

