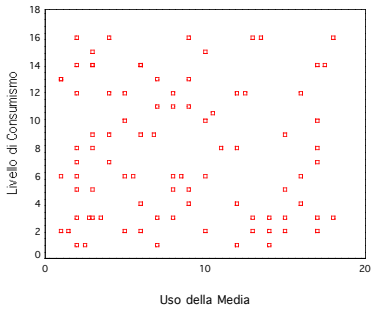
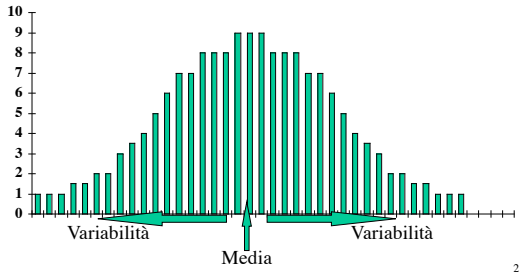


1

Variabilità

Dispersione rispetto alla Media



2

Regressione Lineare

Somma dei Quadrati Totale

$$\Sigma(Y - \overline{Y})^2$$

Variabilità totale di Y.

3

Regressione Lineare

Somma dei Quadrati di Regressione

$$\Sigma(Y' - \overline{Y})^2$$

Variabilità di Y spiegata da X.

4

Regressione Lineare

Somma dei Quadrati dei Residui

$$\Sigma(Y - Y')^2$$

Variabilità di Y Non spiegata da X.

5

Regressione Lineare

In Termini di Somme de Quadrati

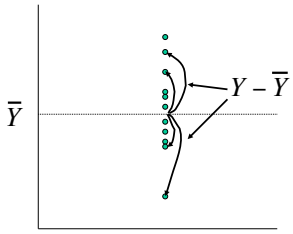
$$\Sigma(Y - \overline{Y})^2 = \Sigma(Y' - \overline{Y})^2 + \Sigma(Y - Y')^2$$

Totale Regressione Residui

6

Regressione Lineare

Somma dei Quadrati Totale



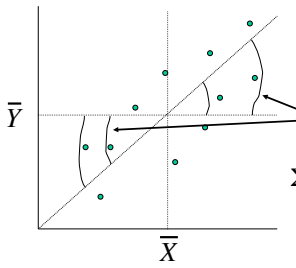
$$\Sigma(Y - \overline{Y})^2$$

Senza informazioni su X

7

Regressione Lineare

Somma dei Quadrati di Regressione

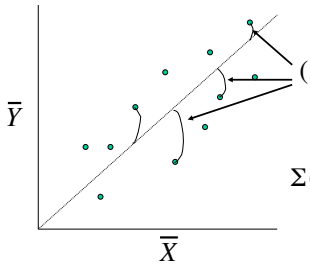


$$\Sigma(Y' - \overline{Y})^2$$

8

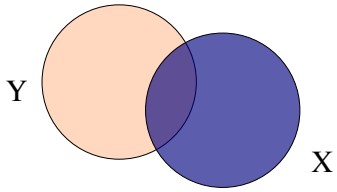
Regressione Lineare

Somma dei Quadrati Residui



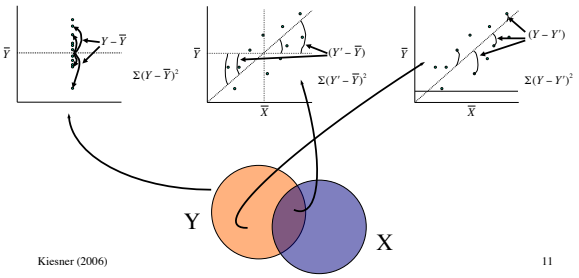
$$\Sigma(Y - Y')^2$$

9



$$\Sigma(Y-\overline{Y})^2 = \Sigma(Y'-\overline{Y})^2 + \Sigma(Y-Y')^2$$

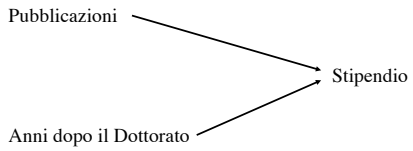
Totale Regressione Residui



Regressione Multipla

- Se vogliamo prevedere lo stipendio di un professore in base a
  - Numero di pubblicazioni
  - Anni trascorsi dopo il dottorato

Regressione Multipla



13

Regressione Multipla

Stipendio - Pubblicazioni r = ,461

Stipendio - Anni dopo il Dottorato r = ,618

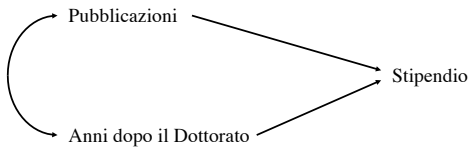
14

Regressione Multipla

- Dato che $b_{YX} = r\frac{sd_Y}{sd_X}$
- Usando le correlazioni possiamo calcolare *b* per le pubblicazioni e gli anni dopo il dottorato.
  - $b_{\text{pubblicazioni}}$ = \$566 per pubblicazione
  - $b_{\text{anni dopo il dottorato}}$ = \$520 per anno

15

Regressione Multipla



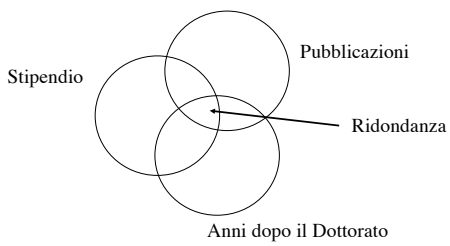
16

Regressione Multipla

Stipendio - Pubblicazioni	r = ,461
Stipendio - Anni dopo il Dottorato	r = ,618
Pubblicazioni - Anni dopo il Dottorato	r = ,683

17

Regressione Multipla



18

Regressione Multipla

$$Y = a + b_1 X_1 + b_2 X_2 + e$$

b_1, b_2 → Coefficienti di Regressione Parziale

19

Regressione Multipla

- Adesso abbiamo la stima dello Stipendio:

$$Y = 20{,}138 + (88) X_{pubb} + (479) X_{anni} + e$$

20

Regressione Multipla

- Nota la differenza fra le stime degli effetti:

$b_{\text{pubblicazioni}}$ = \$566 per pubblicazioni
$b_{\text{anni da dottorato}}$ = \$520 per anno

$b_{\text{pubblicazioni}}$ = \$88 per pubblicazioni
$b_{\text{anni da dottorato}}$ = \$479 per anno

21

Regressione Multipla

- C'è un problema, però, nell'interpretazione dei valori dei *b*.
- Siccome le due variabili indipendenti utilizzano scale diverse, non possiamo confrontare i *b* per sapere quale variabile indipendente ha un effetto relativamente più importante.

22

Regressione Multipla

$$z_Y = \beta_{Y1 \cdot 2} z_1 + \beta_{Y2 \cdot 1} z_2 + e$$

Soluzione standardizzata.

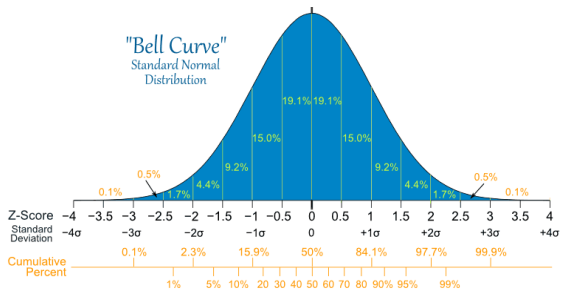
23

Regressione Multipla

$$z_Y = ,072 z_{pubb} + ,570 z_{anni} + e$$

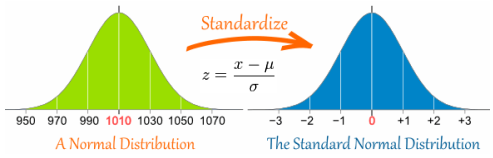
24

mathisfun.com



25

mathisfun.com



26

Regressione Multipla

r = misura dell'associazione fra due variabili

R = misura dell'associazione fra una variabile e una combinazione di variabili indipendenti.

r^2 = proporzione della varianza che ciascuna delle due variabili ha in comune con l'atra.

R^2 = proporzione della varianza della variabile dipendente in Comune con l'insieme delle variabili indipendenti.

27

Regressione Multipla

R^2

- Può avere valori fra 0 e 1 (non può essere negativo).

28

Regressione Multipla

$$R_{Y \cdot 12} = r_{Y\hat{Y}_{12}}$$

29

Regressione Multipla

$$R_{Y \cdot 12} = r_{Y\hat{Y}_{12}}$$

Non può essere più piccola del valore assoluto della correlazione più grande fra Y e le variabili indipendenti.

30

Regressione Multipla

- Test di Significatività di $R^2_{Y\cdot 12}$

$$F = \frac{SQ_{reg}/k}{SQ_{res}/(n-k-1)} = \frac{MS_{reg}}{MS_{res}}$$

$$df_{num} = k$$
$$df_{den} = n-k-1$$

31

Regressione Multipla

Test Condizionale degli effetti di X_1 e X_2.
Vogliamo sapere se X_1 e X_2 hanno effetti unici.

$$F = \frac{SQ_{Y1\cdot 2}/k}{SQ_{res}/(n-k-1)}$$

$$F = \frac{SQ_{Y2\cdot 1}/k}{SQ_{res}/(n-k-1)}$$

32

Regressione Multipla

$SQ_{reg} = 799{,}07$
$SQ_{res} = 416{,}93$
$n = 20$
$k = 2$
$SQ_{Y1} = 595{,}84$
$SQ_{Y2} = 437{,}76$

33

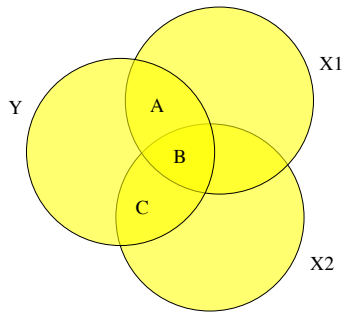
Regressione Multipla

Test di Significatività di $R^2_{Y\cdot12}$

$$F = \frac{SQ_{reg}/k}{SQ_{res}/(n-k-1)} = \frac{799{,}07/2}{416{,}93/17} = 16{,}29$$

$p < 0{,}01$

34



35

Regressione Multipla

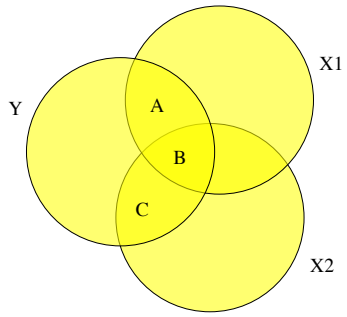
Test Condizionale degli effetti di X_1.

$$SQ_{reg} - SQ_{Y2} = SQ_{Y1\cdot2}$$
$$799{,}07 - 437{,}76 = 361{,}31$$

$$F = \frac{SQ_{Y1\cdot2}/k}{SQ_{res}/(n-k-1)} = \frac{361{,}31}{24{,}53} = 14{,}73$$

$p < 0{,}01$

36



37

Regressione Multipla

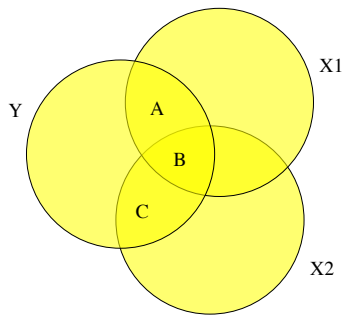
Test Condizionale degli effetti di X_2.

$$SQ_{reg} - SQ_{Y1} = SQ_{Y2 \cdot 1}$$
$$799{,}07 - 595{,}84 = 203{,}23$$

$$F = \frac{SQ_{Y2 \cdot 1} / k}{SQ_{res} / (n - k - 1)} = \frac{203{,}23}{24{,}53} = 8{,}28$$

$p < 0{,}05$

38



39