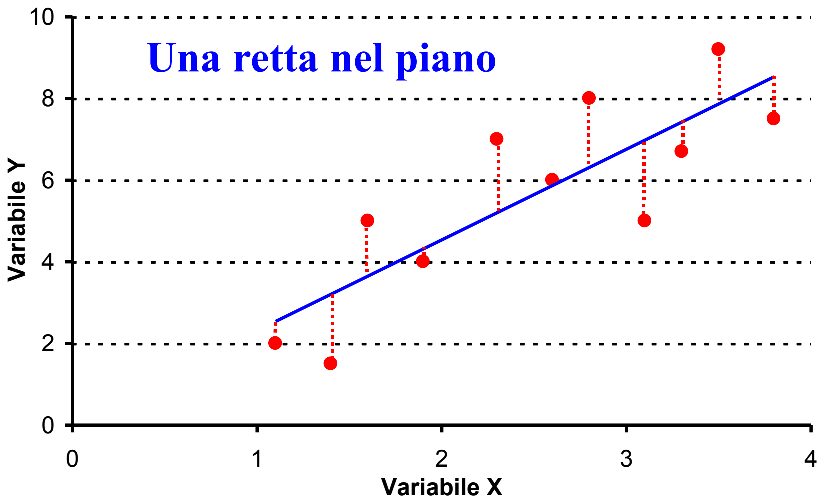
La regressione lineare multipla

- Prof. Giuseppe Verlato
- Sezione di Epidemiologia e Statistica Medica, Dipartimento di Medicina e Sanità Pubblica, Università degli Studi di Verona

Regressione lineare semplice

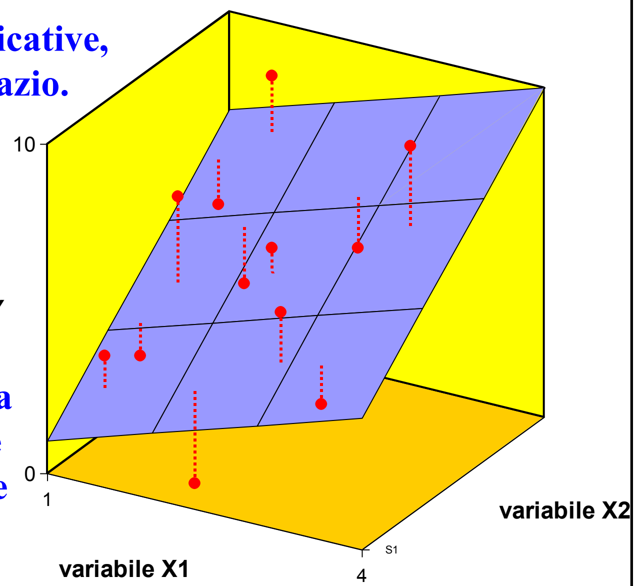
$$y = \beta_0 + \beta_1 x + \varepsilon$$



Regressione lineare multipla

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \varepsilon$$

Con 2 variabili esplicative, un piano nello spazio.



Con più variabili, la regressione lineare multipla può essere rappresentata nell'iperspazio

Regressione lineare multipla

Intercetta (corner, grand mean)

Variabile di risposta (dipendente, response variable)

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_{13} x_1 x_3 + \varepsilon$$

Termine di errore

Effetto principale

Termine di interazione

Coefficienti di regressione parziali, parametri ignoti del modello stimati sulla base dei dati disponibili

Variabili esplicative (predittive, covariate, indipendenti, explanatory)

APPLICAZIONI della REGRESSIONE LINEARE MULTIPLA -1

1) Valutare simultaneamente l'influenza su una variabile di risposta di molte variabili esplicative

variabile di risposta — variabili in studio

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \varepsilon$$

FEV_1 — pacchetti-anno — anni in miniera

2) Valutare l'influenza di una variabile (esplicativa) su un'altra variabile (di risposta) controllando per possibili **confondenti**

variabile di risposta — variabile in studio — confondente

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \varepsilon$$

FEV_1 — pacchetti-anno — età

APPLICAZIONI della REGRESSIONE LINEARE MULTIPLA -2

3) Valutare se l'effetto di una variabile esplicativa viene modificato da un'altra variabile **(modificatore d'effetto)**

variabile di risposta — variabili in studio — termine d'interazione

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2 + \varepsilon$$

frequenza cardiaca — tono vagale — tono simpatico — interazione vago-simpatica

4) Fare delle predizioni: in base al sesso, all'età e alla statura di una determinata persona posso stabilire qual è il suo FEV_1 teorico

variabile di risposta — variabili predittive

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \varepsilon$$

FEV_1 — età — statura

Il peso (Y) dipende dalla statura (X_1), dall'età (X2), dall'introito calorico (X3)

$$E(y) = \hat{y} = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_{13} x_1 x_3$$

E(y) = valore atteso (media) del peso degli individui che hanno quella determinata statura, età, introito calorico

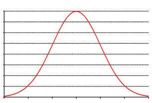
$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_{13} x_1 x_3 + \varepsilon$$

y = peso di un determinato individuo, che dipende dalla statura, età, introito calorico (parte sistematica del modello), ma anche da altre caratteristiche individuali (**ε**, parte probabilistica)

Regressione lineare multipla

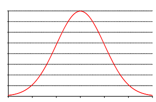
$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_{13} x_1 x_3 + \varepsilon$$

Variabile di risposta (dipendente)



Predittore lineare, parte deterministica del modello, senza variabilità casuale

Termine di errore, parte probabilistica



L'errore, e quindi la variabile di risposta, si distribuisce NORMALMENTE

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_{13} x_1 x_3 + \varepsilon$$

La funzione legame (link-function), che unisce la variabile dipendente al predittore lineare, è l'identità

I MODELLI LINEARI GENERALIZZATI si differenziano per la distribuzione dell'errore (error function) e per la funzione legame (link function)

REGRESSIONE LINEARE MULTIPLA

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_{13} x_1 x_3 + \varepsilon$$

La funzione legame (link-function) è l'IDENTITA'

L'errore segue la distribuzione NORMALE

MODELLO DI REGRESSIONE LOGISTICA

$$Log\,[y/(1-y)] = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_{13} x_1 x_3 + \varepsilon$$

La funzione legame (link-function) è il LOGIT [LOG(ODDS)]

L'errore segue la distribuzione BINOMIALE

MODELLO LOG-LINEARE

$$Log(y) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_{13} x_1 x_3 + \varepsilon$$

La funzione legame (link-function) è il LOGARITMO

L'errore segue la distribuzione di POISSON

1) Regressione lineare semplice

$y = \beta_0 + \beta_1 x + \varepsilon$, *in cui X ed Y sono variabili quantitative*

2) Regressione lineare multipla

$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \varepsilon$, *in cui X ed Y sono variabili quantitative*

3) Analisi della varianza (ANOVA)

$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \varepsilon$, *in cui Y quantitativa, X qualitative*

4) Analisi della covarianza (ANCOVA)

$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \varepsilon$, *in cui Y quantitativa, X qualitative e quantitative*

$\Downarrow$

Sono tutti riconducibili ad un unico modello lineare generalizzato, in cui:

la funzione legame (link-function) è l'IDENTITA'

l'errore segue la distribuzione NORMALE

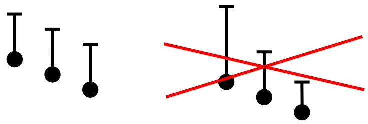
Regressione lineare multipla

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_{13} x_1 x_3 + \varepsilon$$

ASSUNZIONI

1) Il valore atteso degli errori E(ε) deve essere pari a ZERO

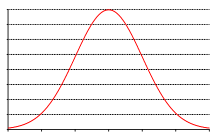
2) OMOSCEDASTICITA' (La varianza degli errori rimane costante)



3) INDIPENDENZA degli errori

se le provette tra un esame e l'altro non vengono lavate adeguatamente, una determinazione risente della determinazione precedente

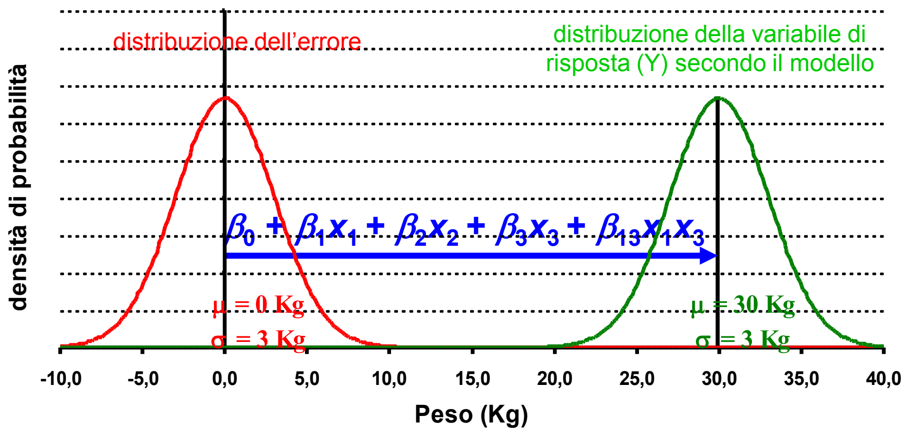
4) Distribuzione NORMALE degli errori



Regressione lineare multipla: ASSUNZIONI

1) Il valore atteso degli errori E(ε) deve essere pari a ZERO

4) Gli errori si distribuiscono normalmente



Regressione lineare multipla

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_{13} x_1 x_3 + \varepsilon$$

ASSUNZIONI

2) OMOSCEDASTICITA'

3) INDIPENDENZA degli errori → Metodo dei minimi quadrati

4) Distribuzione NORMALE degli errori → per fare inferenza

Metodi di ottimizzazione

per trovare il modello che meglio si adatta ai dati

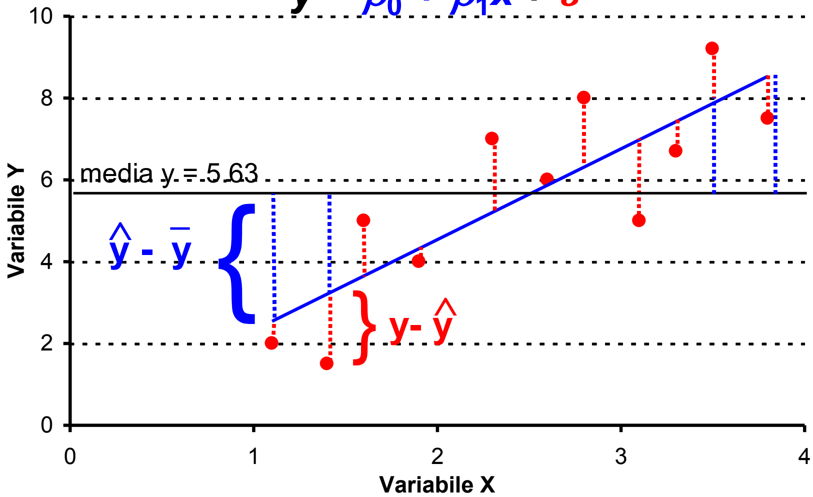
Metodo dei minimi quadrati (least-square method)

Necessita dell'omoscedasticità.

Viene utilizzato per i modelli lineari generalizzati in cui la funzione legame (link function) è l'identità: Regressione lineare semplice, Regressione lineare multipla, Analisi della varianza, Analisi della covarianza

SCOMPOSIZIONE DELLA DEVIANZA nella Regressione lineare semplice - 1

$$y = \beta_0 + \beta_1 x + \varepsilon$$



$$(y - \bar{y}) = (\hat{y} - \bar{y}) + (y - \hat{y})$$

SCOMPOSIZIONE DELLA DEVIANZA nella Regressione lineare semplice - 2

Variabilità totale — **Variabilità residua**

$$(y - \bar{y}) = (\hat{y} - \bar{y}) + (y - \hat{y})$$

Variabilità spiegata dalla regressione

Si può dimostrare che:

Devianza totale, SST — **Devianza residua, SSE**

$$\Sigma(y - \bar{y})^2 = \Sigma(\hat{y} - \bar{y})^2 + \Sigma(y - \hat{y})^2$$

Devianza spiegata dalla regressione, SSR

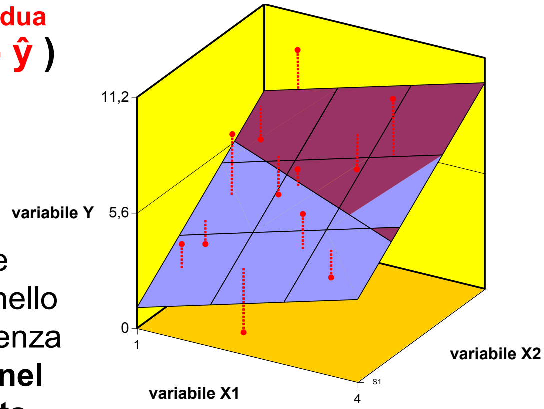
Regressione lineare multipla

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \varepsilon$$

Variabilità totale **Variabilità residua**

$$(y - \bar{y}) = (\hat{y} - \bar{y}) + (y - \hat{y})$$

Variabilità spiegata dalla regressione



La scomposizione delle devianza viene effettuata nello stesso modo: l'unica differenza è che **y atteso ($\hat{y}$) giace nel piano e non su una retta**

Metodo dei Minimi Quadrati

- Criterio dei minimi quadrati:

$$\min \Sigma(y_i - \hat{y}_i)^2$$

dove :

y_i = valore <u>osservato</u> della variabile dipendente per la *i*-esima osservazione

$\hat{y}_i$ = valore <u>stimato</u> della variabile dipendente per la *i*-esima osservazione.

Esempio sulla REGRESSIONE LINEARE MULTIPLA-1

Il peso alla nascita dipende da (regressione) ed è correlato con (correlazione) l'età gestazionale e la statura del neonato?

Modello ipotizzato: **Peso** $= \beta_0 + \beta_1$ **Statura** $+ \beta_2$ **Età gest.** $+ \varepsilon$

ANALISI GLOBALE DEL MODELLO, basata sulla SCOMPOSIZIONE DELLA DEVIANZA

IPOTESI NULLA: Tutte le variabili predittive sono irrilevanti.

$H_0: \beta_1 = \beta_2 = 0$

Esempio sulla REGRESSIONE LINEARE MULTIPLA-2

SCOMPOSIZIONE DELLA DEVIANZA

Fonte di variabilità	Gradi di libertà	Devianza	Varianza	Statistica-test
Regressione	p-1	$SSR = \Sigma(\hat{y} - \bar{y})^2$	MSR=SSR/(p-1)	F = MSR/MSE
Residua	n-p	$SSE = \Sigma(y - \hat{y})^2$	MSE=SSE/(n-p)	con (p-1) e (n-p)
TOTALE	n-1	$SST = \Sigma(y - \bar{y})^2$		gradi di libertà
Regressione	2	11 073 128	5 536 564	44,81
Residua	60–2-1 = 57	7 042 277	123 549	con 2 e 57 g.l.
TOTALE	60 – 1 = 59	18 115 405		P<0,001

p = parametri del modello ($\beta_0, \beta_1, \beta_2$)
SSR, SSE, SST = Somma di quadrati (Sum of Squares) spiegata dalla regressione, residua e totale
MSR, MSE = Varianza (Mean Square) spiegata dalla regressione o residua - MSE = Errore quadratico medio

Esempio sulla REGRESSIONE LINEARE MULTIPLA-3

ANALISI GLOBALE DEL MODELLO, basata sulla SCOMPOSIZIONE DELLA DEVIANZA

Coefficiente di determinazione

R^2 = SSR / SST = 11 073 128 / 18 115 405 = 0,611

Il 61,1% della variabilità nel peso neonatale è spiegata dalla correlazione con l'età gestazionale e con la statura.

R (coefficiente di correlazione multipla) = $\sqrt{R^2}$ = 0,782

Nella correlazione semplice:

Pendenza negativa — pendenza positiva

-1 — **0** — **r** — **+1**

Poiché non si può attribuire alcun significato alla direzione di una correlazione multipla con più variabili predittive

0 — **R** — **+1**

Esempio sulla REGRESSIONE LINEARE MULTIPLA-4

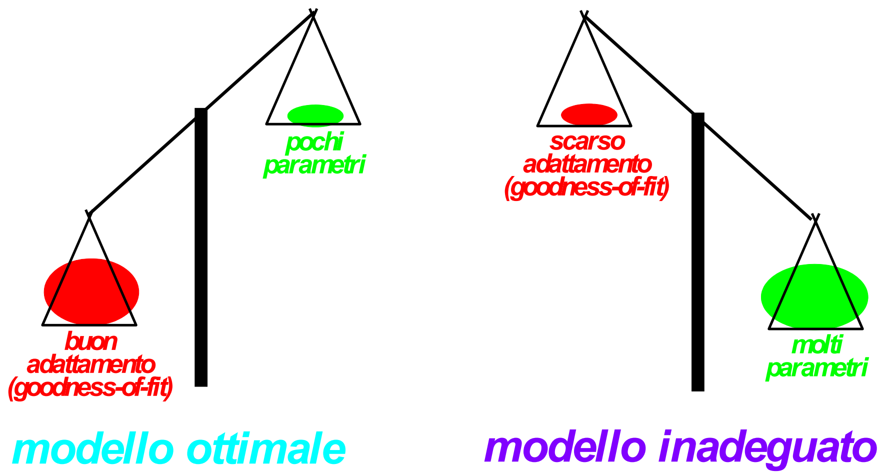
La significatività di R^2 (SSR / SST) si valuta con il test F (MSR/MSE) descritto in precedenza.

R^2 in genere aumenta quando nel modello si introducono nuove variabili, non può essere utilizzato per confrontare modelli con un numero diverso di variabili.

Il valore assunto da R^2 può essere corretto per tener conto del probabile contributo di ogni variabile inclusa, sottraendo il valore atteso in assenza di correlazione.

$$R^2_a = R^2 - (1- R^2)\,(p-1) / (n-p) =$$
$$= 0{,}611 - (1-0{,}611)\ 2 / 57 = 0{,}611 - 0{,}013 = 0{,}598$$

Anche con questo aggiustamento, R^2 non misura in modo soddisfacente la bontà dell'adattamento della regressione interpolata.



Esempio sulla REGRESSIONE LINEARE MULTIPLA-5

ANALISI dei SINGOLI PARAMETRI del MODELLO

Le singole variabili esplicative esercitano un'influenza significativa sulla variabile di risposta?

Test d'ipotesi: $H_0: \beta_i = 0$ $\quad$ b_i = statistica campionaria

$H_1: \beta_i \neq 0$ $\quad$ β_i = parametro ignoto

$$t = \frac{b_i}{ES_b}$$

Sotto H_0, la statistica test segue la distribuzione t di Student con i gradi di libertà della varianza residua (N-p)

	Coefficiente $b = \hat{\beta}$	ES_b	Test $t = b/ES_b$	Significatività
Intercetta	-4010,8	643,8		
Statura	44,3	9,0	4,909	P<0,001
Età gestazionale	118,3	19,2	6,152	P<0,001

Analisi della varianza per eliminare più di una variabile

A volte si vuole testare se la variabilità sia significativamente influenzata dalla soppressione di un gruppo di variabili predittive:

ad esempio, in uno studio sulla resistenza alla fatica fisica può interessare valutare l'effetto di 3 variabili antropometriche (altezza, peso e circonferenza toracica) prese in blocco.

Fonte di variabilità	Gradi libertà	Devianza	Varianza	Statistica-test
Regressione 1	p-1	SSR_1		
Regressione 2	p-1-Δp	SSR_2		
Regr.1 – Regr.2	Δp	$\Delta SSR = SSR_1 - SSR_2$	$MSR = \Delta SSR / \Delta p$	F = MSR/MSE con Δp e $(n-p)$ gradi di libertà
Residua Regr1	n-p	$SSE = \Sigma(y-\hat{y})^2$	$MSE = SSE/(n-p)$	
TOTALE	n-1	$SST = \Sigma(y-\bar{y})^2$		

Correlazione parziale - 1

Il coefficiente di correlazione lineare tra 2 variabili (r_{12}) rispecchia anche eventuali associazioni tra queste variabili ed un eventuale **confondente**. Ad esempio:

Il coefficiente di correlazione parziale è il coefficiente di correlazione tra due variabili, ottenuto tenendo costante il valore di una terza variabile.

$$r_{12.3} = \frac{r_{12} - r_{13} * r_{23}}{\sqrt{(1 - r_{13}^2)(1 - r_{23}^2)}}$$

Correlazione parziale - 2

$$r_{12.3} = \frac{r_{12} - r_{13} * r_{23}}{\sqrt{(1 - r_{13}^2)(1 - r_{23}^2)}}$$

Test d'ipotesi: $H_0: \rho_{12.3} = 0$ $\quad$ $r_{12.3}$ = statistica campionaria

$H_1: \rho_{12.3} \neq 0$ $\quad$ $\rho_{12.3}$ = parametro ignoto

$$t = \frac{r_{12.3}}{\sqrt{1 - r_{12.3}^2}} * \sqrt{n-3}$$

Sotto H_0, la statistica test segue la distribuzione t di Student con n-3 gradi di libertà (i gradi di libertà della varianza residua).

Misura di variabilità	Formula	Gradi libertà
devianza	$\Sigma(y-\bar{y})^2$	n-1
codevianza	$\Sigma(x-\bar{x})(y-\bar{y})$	n-2
SSE (regr.lineare semplice)	$\Sigma(y-\hat{y})^2 = \Sigma(y-b_0-b_1x)^2$	n-2
SSE (regr.lineare multipla)	$\Sigma(y-\hat{y})^2=\Sigma(y-b_0-b_1x-b_2x)^2$	n-3

I gradi di libertà sono sempre pari ad n meno il numero di parametri stimati.

SSE = devianza residua

SELEZIONE DELLE VARIABILI in un MODELLO MULTIVARIATO

1) Procedure automatiche (fishing)

good for prediction, not for explanation

a) procedura step-up (ingresso progressivo)

b) procedura step-down (eliminazione regressiva)

c) procedura stepwise

d) selezione del miglior sottoinsieme

2) Scelta basata su quesiti scientifici

il computer (una “sausage machine”) non può sostituire il cervello umano

David Clayton, Michael Hills: Statistical methods in epidemiology. Oxford Science Publication; Oxford ‘94

Procedure automatiche (fishing)

1) Procedura ad ingresso progressivo (Step-up, forward, bottom-up, modello marginale)

a) Il computer calcola tutte le regressioni con una sola variabile predittiva e sceglie quella con la maggore devianza spiegata dalla regressione (SSR).

b) Alla prima variabile introdotta nel modello vengono affiancate ad una ad una tutte le altre variabili e vengono calcolate le regressioni corrispondenti. Viene scelta come seconda variabile del modello quella che incrementa maggiormente la devianza spiegata (SSR).

c) La procedura ciclica prosegue, mantenendo allo stadio successivo tutte le variabili selezionate allo stadio precedente.

d) Quando l’incremento della SSR diventa modesto, la procedura si arresta.

Procedure automatiche (fishing)

2) Procedura ad eliminazione regressiva (step-down, backward, top-down, modello condizionale)

a) Il computer calcola la regressione su tutte le *p* variabili predittive e scarta la meno significativa.

b) Il computer ricalcola la regressione sulle *p*-1 variabili rimanenti.

c) La procedura si arresta quando tutti i coefficienti di regressione rimasti sono significativi.

Procedure automatiche (fishing)

3) Stepwise

E' un compromesso tra i due metodi precedenti, le variabili vengono sia introdotte nel modello, sia rimosse.

a) Le variabili più significative vengono introdotte nel modello secondo la procedura step-up.

b) Tuttavia dopo l'inclusione di una nuova variabile, si rivaluta il contributo di ogni variabile, e se la variabile meno significativa fornisce un contributo insufficiente sulla base di un criterio prestabilito, essa viene eliminata.

c) Pertanto può succedere che una variabile venga dapprima inclusa nel modello e successivamente eliminata, perché altre variabili, introdotte in un secondo momento, l'hanno resa superflua.

d) In genere il criterio di inclusione è più rigido, più conservativo rispetto al criterio di esclusione. Ad esempio, una variabile può essere inclusa soltanto se il suo coefficiente di regressione parziale è significativo al livello 5% ed eliminata se non risulta più significativo al livello 10%.

Procedure automatiche (fishing)

Le procedure step-up, step-down e stepwise possono portare a risultati diversi, a scegliere variabili diverse. Inoltre, possono non selezionare la migliore regressione possibile sulla base dell' R^2_a (R^2 corretto).

3) Selezione del miglior sottoinsieme

Un algoritmo computerizzato include nel modello il 'migliore' sottoinsieme di variabili sulla base dell' R^2_a , che tiene conto sia della bontà di adattamento (rapporto tra devianza spiegata e devianza totale) che della parsimonia del modello (numero di parametri).

Scelta basata su quesiti scientifici

Il computer (una "sausage machine") non può sostituire il cervello del ricercatore esperto in un settore

1) Usare il rasoio di Occam (Occam's razor)

A parità di ogni altra condizione, adottare sempre il modello più semplice

2) Non inserire troppe variabili nel modello

dovrebbero esserci almeno 10 osservazioni per ogni variabile esplicativa;

anche con molte osservazioni non si dovrebbero introdurre nel modello più di 2-3 variabili esplicative (explanatory) e 5-6 variabili di confondimento (confounders)

3) Non inserire nel modello variabili correlate fra loro

ad esempio, la pressione diastolica e la pressione sistolica sono collineari

4) Non fidarsi solo della significatività statistica

significatività statistica ≠ significatività clinica

5) Non inserire il termine di interazione senza i corrispondenti effetti principali

6) Usare le procedure automatiche solo se non ci sono informazioni disponibili su un determinato problema

Le seguenti diapositive non fanno parte del programma d'esame

Regressione lineare multipla

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_{13} x_1 x_3 + \varepsilon$$

ASSUNZIONI

2) OMOSCEDASTICITA' (La varianza degli errori rimane costante)

3) INDIPENDENZA degli errori

La matrice di varianza-covarianza (varianza di un vettore) del vettore degli errori ha le varianze (σ^2) uguali e le covarianze pari a zero

σ^2	0	0	0	0
0	σ^2	0	0	0
0	0	σ^2	0	0
0	0	0	σ^2	0
0	0	0	0	σ^2

NOTAZIONE MATRICIALE DI UNA REGRESSIONE LINEARE MULTIPLA-1

Soggetto 1	$y_1 = \beta_0 + \beta_1 x_{11} + \beta_2 x_{21} + \beta_3 x_{31} + \beta_{13} x_{11} x_{31} + \varepsilon_1$
Soggetto 2	$y_2 = \beta_0 + \beta_1 x_{12} + \beta_2 x_{22} + \beta_3 x_{32} + \beta_{13} x_{12} x_{32} + \varepsilon_2$
Soggetto 3	$y_3 = \beta_0 + \beta_1 x_{13} + \beta_2 x_{23} + \beta_3 x_{33} + \beta_{13} x_{13} x_{33} + \varepsilon_3$
Soggetto 4	$y_4 = \beta_0 + \beta_1 x_{14} + \beta_2 x_{24} + \beta_3 x_{34} + \beta_{13} x_{14} x_{34} + \varepsilon_4$
Soggetto 5	$y_5 = \beta_0 + \beta_1 x_{15} + \beta_2 x_{25} + \beta_3 x_{35} + \beta_{13} x_{15} x_{35} + \varepsilon_5$
Soggetto 6	$y_6 = \beta_0 + \beta_1 x_{16} + \beta_2 x_{26} + \beta_3 x_{36} + \beta_{13} x_{16} x_{36} + \varepsilon_6$
Soggetto 7	$y_7 = \beta_0 + \beta_1 x_{17} + \beta_2 x_{27} + \beta_3 x_{37} + \beta_{13} x_{17} x_{37} + \varepsilon_7$
Soggetto 8	$y_8 = \beta_0 + \beta_1 x_{18} + \beta_2 x_{28} + \beta_3 x_{38} + \beta_{13} x_{18} x_{38} + \varepsilon_8$

…………………………………………………………………………

NOTAZIONE MATRICIALE DI UNA REGRESSIONE LINEARE MULTIPLA- 2

$$
\begin{matrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \\ y_6 \\ y_7 \\ y_8 \\ \dots \end{matrix}
=
\begin{matrix} 1 & x_{11} & x_{21} & x_{31} & x_{11}x_{31} \\ 1 & x_{12} & x_{22} & x_{32} & x_{12}x_{32} \\ 1 & x_{13} & x_{23} & x_{33} & x_{13}x_{33} \\ 1 & x_{14} & x_{24} & x_{34} & x_{14}x_{34} \\ 1 & x_{15} & x_{25} & x_{35} & x_{15}x_{35} \\ 1 & x_{16} & x_{26} & x_{36} & x_{16}x_{36} \\ 1 & x_{17} & x_{27} & x_{37} & x_{17}x_{37} \\ 1 & x_{18} & x_{28} & x_{38} & x_{18}x_{38} \\ \dots & & & & \end{matrix}
*
\begin{matrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_3 \\ \beta_{13} \end{matrix}
+
\begin{matrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \varepsilon_4 \\ \varepsilon_5 \\ \varepsilon_6 \\ \varepsilon_7 \\ \varepsilon_8 \\ \dots \end{matrix}
$$

$$\underline{y} = \underline{\underline{X}} * \underline{\beta} + \underline{\varepsilon}$$

N.B. Per moltiplicare una matrice per un'altra matrice, si moltiplica ogni riga della I matrice per ogni colonna della II matrice.

PRODOTTO DI UNA MATRICE PER UN VETTORE

Esempio: Calcolo della capacità vitale attesa nei maschi

Capacità vitale (l) = – 4,34 + altezza(m) * 5,76 - età(anni) *0,026

Matrice dei dati * **Vettore delle costanti** = **risultato**

		m	anni
Tony	1	1,80	24
Bepi	1	1,82	46
Gigi	1	1,60	43
Piero	1	1,70	32
Fabio	1	1,75	57
………..	………………		

$$* \begin{matrix} -4,34 \\ 5,76 \\ -0,026 \end{matrix} =$$

1*(-4,34) + 1,80*5,76 + 24*(-0,026)
1*(-4,34) + 1,82*5,76 + 46*(-0,026)
1*(-4,34) + 1,60*5,76 + 43*(-0,026)
1*(-4,34) + 1,70*5,76 + 32*(-0,026)
1*(-4,34) + 1,75*5,76 + 57*(-0,026)
……………..

$$\underline{E(y)} \equiv \hat{y} = \underline{\underline{X}} * \underline{\beta}$$

N.B. Per moltiplicare una matrice per un'altra matrice, si moltiplica ogni riga della I matrice per ogni colonna della II matrice.