

## Q&A Webinar "I modelli di Regressione lineare con IBM SPSS Statistics" del 20 aprile 2021

Veronica

Q: L'algoritmo stepwise da che modello parte?

A: Parte da un modello "nullo" (con nessuna esplicativa) e di step in step aggiunge o rimuove rispettivamente la variabile che meglio o che peggio si adatta a tale modello.

Anonymous

Q: Fino a che valore consideriamo deviazione standard e varianza accettabili?

A: Non esiste un limite; considerato che la deviazione standard assume valori maggiori o uguali a zero, più basso è il suo valore meglio è (NB: la deviazione standard ha la stessa unità di misura dell'outcome di interesse).

Giulia

Q: Esiste un numero di casi medio da dover considerare/avere per inserire nel modello altre variabili esplicative?

A: No. In genere deve valere la regola che il numero di casi/osservazioni deve essere maggiore al numero delle variabili.

Anonymous

Q: Per stimare l'assunto di normalità, è più funzionale osservare i valori di curtosi e asimmetria o lanciare un test statistico di normalità (es. Shapiro-Wilk...)?

A: Molto meglio osservare i valori di curtosi/asimmetria, oltre che media e mediana.

Giulia

Q: Qual è il cut-off di asimmetria e curtosi per considerare una curva normale?

A: Il loro valore di cut-off è zero.

Alessandro

Q: Visto che stiamo vedendo la verifica della correlazione tra out come e variabili, se l'out come fosse una variabile nominale invece che in scala si procede analogamente con la correlazione?

A: No, in quel caso NON si può applicare la regressione lineare ma si deve optare per altre tecniche statistiche, come la classificazione.

Maurizio

Q: Non mi è chiara l'utilità della regressione nell'esempio degli incendi: la variabile di out come non dovrebbe essere la probabilità che si inneschi un incendio in funzione delle variabili esplicative? Non capisco lo scopo di stimare l'entità dell'area che risulta poi incendiata

A: In quell'esempio era utile per fare previsione e monitorare le aree passibili di incendio. Il modello alternativo di cui parla potrebbe essere interessante con una regressione logistica, e abbinare i 2 risultati.

Veronica

Q: Se il test F del modello dovesse essere non significativo, dobbiamo rigettare l'intero modello e quindi non considerare le eventuali significatività che potrebbero comparire nella tabella dei coefficienti (se può succedere)?

A: Sì, esatto.

Giovanni

Q: Per variabile quantitativa si intende solo quantitativa continua o anche dicotomica?

A: L'outcome deve essere per forza quantitativo continuo, mentre le esplicative possono essere anche qualitative.

Roberto

Q: Ora (secondo esempio) sta usando forward, non ho capito che modello è stato usato nell'analisi precedente

A: Il modello "full" con tutte le esplicative.

Anonymous

Q: Fra i metodi stepwise, forward etc c'è un criterio di scelta?

A: In genere il forward e lo stepwise sono preferibili con un numero molto elevato di variabili esplicative (in genere superiore a 100).

Giuseppe

Q: Cosa si intende per residui, in questo caso?

A: La differenza tra il valore predetto dal modello ed il valore vero di quell'osservazione.

Anonymous

Q: Se tra le diverse covariate ho anche una variabile dicotomica come il sesso, posso comunque inserirla nella regressione multipla? Come? Devo specificarlo nell'inserimento?

A: Sì, le esplicative qualitative si possono inserire nel modello, in genere come variabili dummy.

Anonymous

Q: Se F risulta significativo insieme alla significatività delle variabili indipendenti predittive ma R quadrato aggiustato ha un valore minimo come ad esempio 0,09 si può accettare e descrivere l'analisi?

A: Nulla vieta di accettare l'analisi, basta sottolineare che quel modello in termini esplicativi è molto "povero".

Irene

Q: Quali sono i pro e contro dei diversi metodi: T test, forward, backward, stepwise?

A: In genere il forward e lo stepwise sono preferibili con un numero molto elevato di variabili esplicative (in genere superiore a 100).