

Facoltà di Economia
Corso di Metodi Statistici per la Finanza
Esercitazione n. 4

inviare svolgimento a bart@stat.unipg.it entro venerdì 18-11-2011

Utilizzando il software R:

1. reperire dei dati reali (con almeno 3 covariate) su cui applicare il modello di regressione lineare multipla e applicare la funzione *lm* di R per ottenere le stime dei parametri, l'indice di determinazione multiplo e l'indice di determinazione aggiustato, commentando i risultati ottenuti;
2. con riferimento al modello stimato al punto precedente, illustrare come, sulla base del test F, sia possibile verificare l'ipotesi che i coefficienti di regressione per due covariate siano congiuntamente pari a 0;
3. stimare e commentare i parametri di un modello di regressione su dati reali in cui, per una data variabile esplicativa, siano presenti sia una covariata quantitativa che una qualitativa (con più di due categorie), inserendo quest'ultima tramite delle variabili dummy;
4. con riferimento al modello stimato al punto precedente, tramite un opportuno test F verificare l'ipotesi che la variabile qualitativa non abbia un effetto significativo sulla variabile risposta.
5. reperire dei dati reali (con almeno 5 covariate di cui, possibilmente, almeno una qualitativa da inserire tramite dummy) su cui applicare il metodo stepwise backward per la selezione delle covariate significative;
6. ripetere la procedura di cui al punto precedente utilizzando il metodo stepwise forward e confrontare il modello selezionato tramite il procedimento backward di cui al punto precedente;
7. illustrare, sulla base di dati reali, come si accerta la presenza di multicollinearità e quali effetti potrebbe avere;
8. reperire dei dati reali in cui siano presenti più variabili esplicative e più variabili risposta e illustrare, sulla base di questi dati, come viene stimato il modello di regressione multivariata sia per quanto riguarda i coefficienti di regressione che la matrice di varianza-covarianza dei termini di errore;
9. tramite un esempio reale illustrare l'utilizzo dei modelli logit e probit per variabili risposta binarie.