

PSP6075525 - Testing psicologico (matr. dispari)

Esame del 300123

Istruzioni

- Si avvii una nuova sessione di R (o RStudio).
- Si crei un nuovo script di R e lo si salvi come `cognome_nome.R`.
- Si effettui il download del file di dati dell'esame `dati_esame.Rdata` disponibile presso la pagina Moodle Esami del corso e lo si carichi nell'ambiente di lavoro di R.
- Si utilizzi il file `cognome_nome.R` per inserire il codice R utilizzato in risposta i quesiti d'esame. Attenzione: si ricorda di inserire il medesimo codice nel campo di risposta disponibile per ciascun quesito nel form Moodle dell'esame.
- Si invii il file `cognome_nome.R` mediante l'apposita funzione **Consegna codice R** presente nella pagina Moodle Esami del corso.
- Nota: la valutazione della prova sarà effettuata utilizzando primariamente il file `cognome_nome.R`. Si raccomanda pertanto la massima chiarezza nella scrittura delle risposte e la correttezza nel riportare i comandi e gli output di R per ciascun quesito d'esame.

Il file `data_exam.Rdata` contiene i dati relativi alla somministrazione del test AX001 ad un campione casuale di $n = 5000$ studenti frequentanti l'università di Teramo. Il test, somministrato per la valutazione delle abilità matematiche, è composto da $p = 15$ item rilevati su scale ordinali a 7 punti (livelli alti della scala indicano migliore performance matematica). Successivamente alla raccolta dei dati, le variabili osservate sono state adeguatamente quantificate mediante apposita procedura. L'obiettivo è quello di studiare la dimensionalità del test AX001 con particolare riferimento al numero e alla tipologia di dimensioni latenti che esso individua. Si importi il dataset in R e si risponda ai quesiti che seguono.

1. Si esegua una divisione a metà del dataset (50% di unità statistiche per la prima metà) e si utilizzi la prima metà del dataset per le analisi esplorative e la seconda metà per le analisi confermative. Successivamente si esegua un'analisi basata sul clustering gerarchico con metodo `ward.D2` e si individuino un numero congruo di raggruppamenti delle variabili osservate. Sulla base dell'analisi di raggruppamento si proponga un modello CFA. Nota: si suggerisce di impostare il seed di generazione casuale pari a `seedx=16001`.
2. Si adatti ai dati il modello CFA definito al punto precedente secondo la metrica ULI. Si commenti il risultato ottenuto anche alla luce dell'adattamento complessivo del modello ai dati. Nota: nell'interpretazione della soluzione fattoriale si utilizzino i coefficienti stimati standardizzati.
3. Si semplifichi il modello CFA adattato al punto precedente utilizzando il criterio $\hat{\lambda}_{jk} < 0.17$. Si commenti il risultato ottenuto.
4. Si consideri l'insieme delle variabili osservate utilizzate al punto precedente. Si definisca un modello CFA con $q = 1$ variabili latenti e lo si adatti ai dati secondo la metrica ULI. Successivamente si confronti il risultato ottenuto con quello del punto precedente e si scelga quale dei due modelli è da preferire.
5. Si consideri il modello unidimensionale definito al punto precedente e lo si migliori aggiungendo come parametri da stimare le covarianze di errore a coppia (ad esempio, sull'insieme $\{Y_1, Y_2, Y_3, Y_4\}$ si considerino solo le quantità $\text{COV}(Y_1, Y_2), \text{COV}(Y_2, Y_3), \text{COV}(Y_3, Y_4)$). Si confronti il nuovo modello con quello unidimensionale del punto precedente rispetto (i) al fit complessivo e (ii) all'errore di previsione tramite metodo Monte Carlo. Si scelga, con adeguata giustificazione, il modello finale. Nota: per il calcolo dell'errore di previsione si utilizzi `k=7` e `B=250`.