

Analisi Statistica dei Dati: foglio di lavoro sulla valutazione Monte Carlo dell'errore di stima

Luca La Rocca

23 ottobre 2018

Dati simulati

1. Create una matrice `simuData` con $n = 5$ righe e $M = 10000$ colonne, usando la funzione `rnorm` in modo che ogni colonna contenga un campione casuale dalla popolazione normale standard.
2. Calcolate la varianza empirica corretta S_n^2 per ciascun campione, applicando la funzione `var` alle colonne della matrice `simuData` con l'aiuto della funzione `apply`, quindi salvate il vettore ottenuto come primo elemento `varEst$s` di una lista `varEst`.
3. Calcolate, per ciascun campione, lo stimatore di massima verosimiglianza $D_n^2 = S_n^2(n-1)/n$ e l'ulteriore stimatore $U_n^2 = S_n^2(n-1)/(n+1)$, quindi salvate i risultati ottenuti come ulteriori elementi `varEst$d` e `varEst$u` della lista `varEst`.
4. Scrivete una funzione `bias` che, prendendo come argomenti un vettore `t` di stime e uno valore `xi` del parametro di interesse, restituisca in output la distorsione di `t` rispetto a `xi`.
5. Scrivete una funzione `rmse` che, prendendo come argomenti un vettore `t` di stime e uno valore `xi` del parametro di interesse, restituisca in output la radice dell'errore quadratico medio di `t` rispetto a `xi`.
6. Confrontate S_n^2 , D_n^2 e U_n^2 , come stimatori per σ^2 , in termini di distorsione ed errore quadratico medio, applicando le funzioni `bias` e `rmse` agli elementi della lista `varEst` con l'aiuto della funzione `lapply`, verificando in questo modo quanto sapete sulle loro prestazioni.
7. Analogamente, avvalendovi di una lista `sdEst`, confrontate S_n , D_n e U_n come stimatori di σ e valutate quale sia preferibile per stimare σ .
8. Realizzate, nei due casi, usando la funzione `boxplot`, un grafico degli errori di stima relativi ai tre stimatori (confrontandone l'intera distribuzione).