

Άσκηση 1 (Jackknife εκτίμηση του συντελεστή συσχέτισης). Έστω $X_1, \dots, X_n \in \mathbb{R}^2$ ανεξάρτητες και ισόνομες παρατηρήσεις από διμεταβλητή κανονική κατανομή με διάνυσμα μέσης τιμής

$$\mu = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

και πίνακα συνδιακύμανσης

$$\Sigma = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}.$$

Υποθέτουμε ότι $n = 30$ παρατηρήσεις παράγονται.

1. Έστω T_n ο δειγματικός συντελεστής συσχέτισης Pearson που υπολογίζεται από ολόκληρο το δείγμα. Να υπολογιστεί η τιμή του.
2. Για κάθε $i = 1, \dots, n$, ο leave-one-out εκτιμητής ορίζεται ως

$$T(-i) = \text{ο συντελεστής συσχέτισης από το δείγμα χωρίς την } i\text{-οστή παρατήρηση.}$$

Να υπολογιστούν όλα τα $T(-i)$.

3. Ορίστε τις jackknife ψευδο-τιμές ως

$$\tilde{T}_i = nT_n - (n-1)T(-i), \quad i = 1, \dots, n.$$

Να υπολογίσετε τον jackknife εκτιμητή του συντελεστή συσχέτισης

$$T_{\text{jack}} = \frac{1}{n} \sum_{i=1}^n \tilde{T}_i.$$

4. Να υπολογίσετε τη jackknife εκτίμηση της μεροληψίας του T_n

$$b_{\text{jack}} = (n-1)(\bar{T}_n - T_n),$$

όπου $\bar{T}_n = \frac{1}{n} \sum_{i=1}^n T(-i)$.

Να επαληθευτεί ότι

$$T_{\text{jack}} = T_n - b_{\text{jack}}.$$

5. Να υπολογίσετε τη jackknife εκτίμηση του τυπικού σφάλματος

$$\hat{s}_{\text{ejack}} = \sqrt{\tilde{s}_n^2/n}.$$

όπου

$$\tilde{s}_n^2 = \text{δειγματική διασπορά των } \{\tilde{T}_1, \dots, \tilde{T}_n\} = \frac{1}{n-1} \sum_{i=1}^n (\tilde{T}_i - T_{\text{jack}})^2$$

6. Να κατασκευαστεί διάστημα εμπιστοσύνης 95% για τον πραγματικό συντελεστή συσχέτισης:

$$T_{\text{jack}} \pm t_{n-1, 0.025} \cdot \hat{s}_{\text{ejack}}.$$