

Άσκηση 1. Θεωρήστε το αρχείο δεδομένων `nerve_data.txt` που είναι διαθέσιμο στο e-class (κατηγορία Έγγραφα/Υλικό Παπασταμούλη/datasets), το οποίο καταγράφει $n = 799$ χρόνους αναμονής μεταξύ διαδοχικών συσπάσεων νευρικής ίνας (πηγή: Cox and Lewis, 1966).

1. Υπολογίστε την εμπειρική συνάρτηση κατανομής μέσω της εντολής `ecdf()`.
2. Απεικονίστε σε ένα διάγραμμα την εμπειρική συνάρτηση κατανομής μέσω της αντίστοιχης μεθόδου `plot()`.
3. Υπολογίστε την 95% ζώνη εμπιστοσύνης DKW για τη συνάρτηση κατανομής των δεδομένων και προσθέστε την στο διάγραμμα του προηγούμενου ερωτήματος.
4. Για $x = 0.1, 0.2, \dots, 1.2$: υπολογίστε τις σημειακές εκτιμήσεις της $F(x)$, τα όρια της 95% ζώνης εμπιστοσύνης DKW και το 95% ασυμπτωτικό διαστήμα εμπιστοσύνης Wald ίσων ουρών για την $F(x)$. Παρουσιάστε τα αποτελέσματα αυτά σε έναν πίνακα.
5. Υπολογίστε τη στατιστική συνάρτηση Kolmogorov-Smirnov ανάμεσα στην εμπειρική συνάρτηση κατανομής και τη συνάρτηση κατανομής της $\mathcal{G}(1.1, 1 + y)$ (shape - rate parameterization), όπου y το τελευταίο ψηφίο του αριθμού μητρώου σας.
6. Ελέγξτε την υπόθεση

$$H_0 : F(x) = F_0(x), \forall x \in \mathbb{R}, \quad \text{VS} \quad H_1 : \exists x : F(x) \neq F_0(x)$$

όπου F_0 η συνάρτηση κατανομής της $\mathcal{G}(1.1, 1 + y)$ (shape - rate parameterization), σε επίπεδο σημαντικότητας $\alpha = 0.05$, όπου y το τελευταίο ψηφίο του αριθμού μητρώου σας. Για το σκοπό αυτό χρησιμοποιήστε

- (α) τον έλεγχο Kolmogorov-Smirnov.
 - (β) τη ζώνη εμπιστοσύνης DKW.
7. Υπολογίστε τους εκτιμητές αντικατάστασης για τη μέση τιμή, διασπορά και λοξότητα της κατανομής.
 8. Εκτιμήστε σημειακά και με ένα 95% διάστημα εμπιστοσύνης τα τεταρτημόρια της κατανομής των παρατηρήσεων ($q_p, p = 0.25, 0.5, 0.75$).

Αναφορές

Cox, David Roxbee, and Peter AW Lewis. "The statistical analysis of series of events." (1966).