

Άσκηση 1. Θεωρήστε το αρχείο δεδομένων `nerve_data.txt` που είναι διαθέσιμο στο `e-class` (κατηγορία Έγγραφα/datasets), το οποίο καταγράφει $n = 799$ χρόνους αναμονής μεταξύ διαδοχικών συσπάσεων νευρικής ίνας (πηγή: Cox and Lewis, 1966). Να υπολογιστούν 95% bootstrap διαστήματα εμπιστοσύνης (κανονικά, βασικά και ποσοστιαίων σημείων) για τα τεταρτημόρια της κατανομής των παρατηρήσεων ($q_p, p = 0.25, 0.5, 0.75$).

Άσκηση 2. Τα δεδομένα του Πίνακα 1 περιέχουν $n = 14$ μετρήσεις του ποσοστού επιβίωσης αρουραίων (`surv`) αφού εκτέθηκαν σε διάφορα επίπεδα ραδιενέργειας (`dose`).

1. Εκτιμήστε το απλό γραμμικό μοντέλο

$$\log(\text{surv}) = \beta_0 + \beta_1 \times \text{dose}.$$

2. Κάντε **bootstrap στα κατάλοιπα του γραμμικού μοντέλου** του προηγούμενου ερωτήματος, για να προσομοιώσετε $B = 1000$ τιμές των $(\hat{\beta}_0, \hat{\beta}_1)$ και κατασκευάστε το διάγραμμα διασποράς και τα ιστογράμματα των προσομοιωμένων τιμών.
3. Να κατασκευαστούν 95% bootstrap διαστήματα εμπιστοσύνης (κανονικά, βασικά και ποσοστιαίων σημείων) για το β_1 .

	dose	surv
1	117.50	44.00
2	117.50	55.00
3	235.00	16.00
4	235.00	13.00
5	470.00	4.00
6	470.00	1.96
7	470.00	6.12
8	705.00	0.50
9	705.00	0.32
10	940.00	0.11
11	940.00	0.01
12	940.00	0.02
13	1410.00	0.70
14	1410.00	0.01

Πίνακας 1: Δεδομένα από: Efron, B. (1988). Computer-intensive methods in statistical regression. *SIAM Review*, **30**, 421-449.

Άσκηση 3. Τα δεδομένα στον Πίνακα 2 αφορούν την καταθλιπτική επίδραση δύο ειδών ναρκωτικών σε δείγμα 20 clubbers, όπου σε 10 δόθηκε MDMA και στους υπόλοιπους αλκοόλ. Να εκτιμηθεί μέσω bootstrap το p-value του ελέγχου υπόθεσης

$$H_0 : \mu_1 = \mu_2 \quad \text{έναντι} \quad H_1 : \mu_1 > \mu_2$$

όπου μ_1 και μ_2 η μέση επίδραση για το MDMA και αλκοόλ, αντίστοιχα.

Ναρκοτικό	Καταθλιπτική Επίδραση
MDMA	28, 35, 35, 24, 39, 32, 27, 29, 36, 35
Αλκοόλ	5, 6, 30, 8, 9, 7, 6, 17, 3, 10

Πίνακας 2: Καταθλιπτική επίδραση ναρκωτικών.