Μη παραμετρική συμπερασματολογία με μεθόδους επαναδειγματοληψίας: το bootstrap

Μη Παραμετρική Στατιστική

Παναγιώτης Παπασταμούλης
Επίκουρος Καθηγητής
Τμήμα Στατιστικής ΟΠΑ

papastamoulis@aueb.gr

Τετάρτη, 22/04/2020

Περιεχόμενα

Μέθοδοι επαναδειγματοληψίας

- Το jackknife (Quenouille 1949, Tukey 1958) και το bootstrap (Efron, 1979) είναι μη παραμετρικές μέθοδοι για
  - εκτίμηση τυπικών σφαλμάτων συναρτησιακών
  - υπολογισμό διαστημάτων εμπιστοσύνης συναρτησιακών
- Υπολογιστικά, το jackknife είναι λιγότερο απαιτητικό, αλλά το bootstrap έχει περισσότερα πλεονεκτήματα, πχ:
  - εκτίμηση της δειγματικής κατανομής εκτιμητών συναρτησιακών
  - το bootstrap μπορεί να εφαρμοστεί σε περιπτώσεις που είναι γνωστό ότι το jackknife αποτυγχάνει (πχ: εκτίμηση ποσοστιαίων σημείων)

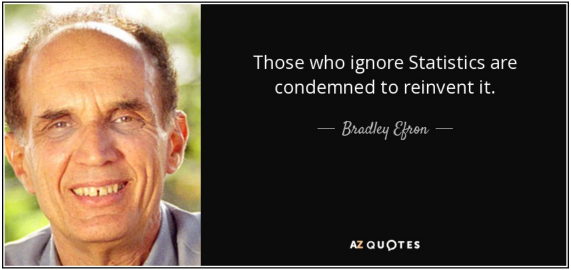
To Bootstrap

- *bootstrapping:* the process of completing a task without external help.
- E.g: pulling oneself out of a swamp by the straps of the boots.
- Rudolf Erich Raspe (1785): *Baron Munchausen's Narrative of his Marvellous Travels and Campaigns in Russia*

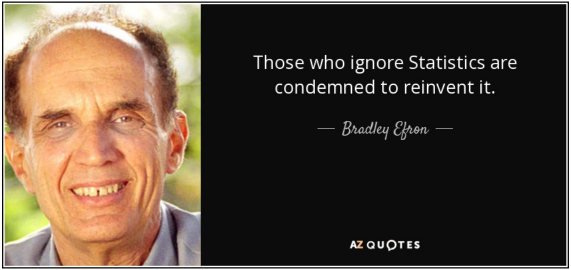
To Bootstrap

- *bootstrapping:* the process of completing a task without external help.
- E.g: pulling oneself out of a swamp by the straps of the boots.
- Rudolf Erich Raspe (1785): *Baron Munchausen's Narrative of his Marvellous Travels and Campaigns in Russia*

Bootstrap

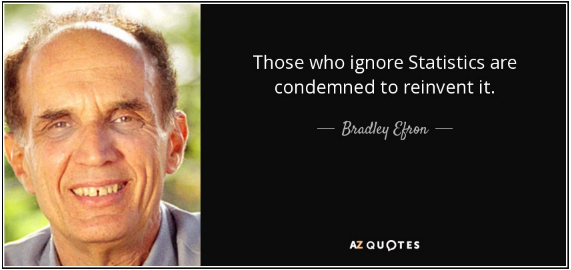

Those who ignore Statistics are condemned to reinvent it.
— Bradley Efron —
AZ QUOTES

Bootstrap



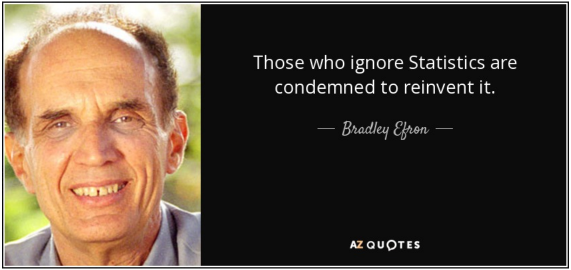
- Bradley Efron: εισήγαγε την τεχνική Bootstrap το 1979

Bootstrap



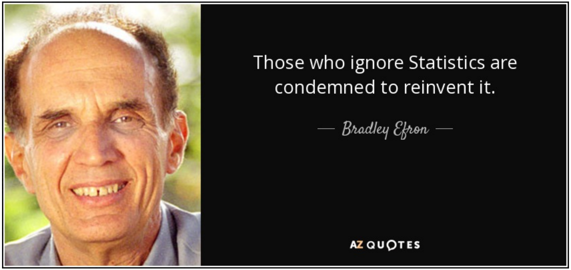
- Bradley Efron: εισήγαγε την τεχνική Bootstrap το 1979
- Το Bootstrap είναι μία από τις πρώτες υπολογιστικές μεθόδους της Στατιστικής

Bootstrap



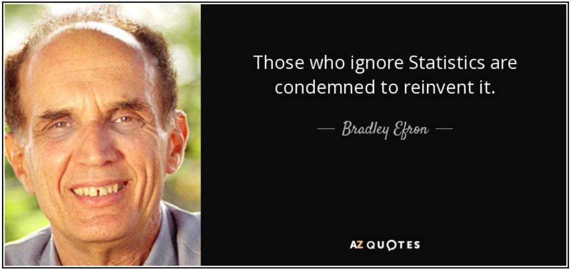
- Bradley Efron: εισήγαγε την τεχνική Bootstrap το 1979
- Το Bootstrap είναι μία από τις πρώτες υπολογιστικές μεθόδους της Στατιστικής
- Στην πράξη, απαιτείται προσομοίωση αν και υπάρχουν παραδείγματα όπου κάτι τέτοιο δεν είναι απαραίτητο

Bootstrap



- Bradley Efron: εισήγαγε την τεχνική Bootstrap το 1979
- Το Bootstrap είναι μία από τις πρώτες υπολογιστικές μεθόδους της Στατιστικής
- Στην πράξη, απαιτείται προσομοίωση αν και υπάρχουν παραδείγματα όπου κάτι τέτοιο δεν είναι απαραίτητο
- Βασίζεται στην υπόθεση ότι το παρατηρηθέν δείγμα είναι αντιπροσωπευτικό του πληθυσμού, ακριβέστερα: αρκεί η $\widehat{F}_n$ να είναι καλή εκτίμηση της F.

Bootstrap



- Bradley Efron: εισήγαγε την τεχνική Bootstrap το 1979
- Το Bootstrap είναι μία από τις πρώτες υπολογιστικές μεθόδους της Στατιστικής
- Στην πράξη, απαιτείται προσομοίωση αν και υπάρχουν παραδείγματα όπου κάτι τέτοιο δεν είναι απαραίτητο
- Βασίζεται στην υπόθεση ότι το παρατηρηθέν δείγμα είναι αντιπροσωπευτικό του πληθυσμού, ακριβέστερα: αρκεί η $\widehat{F}_n$ να είναι καλή εκτίμηση της F.
- Δεν δουλεύει πάντα: το συναρτησιακό πρέπει να είναι «ομαλό»

Bootstrap

- Στατιστική συμπερασματολογία: βασίζεται στην εκτίμηση κάποιας ποσότητας της δειγματικής κατανομής κάποιου εκτιμητή
- Παράδειγμα: το τυπικό σφάλμα είναι η εκτίμηση της διασποράς της δειγματικής κατανομής του εκτιμητή
  - Πάρε πολλά δείγματα από τον *πληθυσμό*
  - Υπολόγισε την εκτίμηση που προκύπτει σε κάθε δείγμα
  - Η κατανομή που προκύπτει είναι η *δειγματική κατανομή του εκτιμητή*

Bootstrap

- Στατιστική συμπερασματολογία: βασίζεται στην εκτίμηση κάποιας ποσότητας της δειγματικής κατανομής κάποιου εκτιμητή
- Παράδειγμα: το τυπικό σφάλμα είναι η εκτίμηση της διασποράς της δειγματικής κατανομής του εκτιμητή
  - Πάρε πολλά δείγματα από τον *πληθυσμό*
  - Υπολόγισε την εκτίμηση που προκύπτει σε κάθε δείγμα
  - Η κατανομή που προκύπτει είναι η *δειγματική κατανομή του εκτιμητή*
- Στην πράξη όμως έχουμε μόνο ένα δείγμα

Bootstrap

- Στατιστική συμπερασματολογία: βασίζεται στην εκτίμηση κάποιας ποσότητας της δειγματικής κατανομής κάποιου εκτιμητή
- Παράδειγμα: το τυπικό σφάλμα είναι η εκτίμηση της διασποράς της δειγματικής κατανομής του εκτιμητή
  - Πάρε πολλά δείγματα από τον *πληθυσμό*
  - Υπολόγισε την εκτίμηση που προκύπτει σε κάθε δείγμα
  - Η κατανομή που προκύπτει είναι η *δειγματική κατανομή του εκτιμητή*
- Στην πράξη όμως έχουμε μόνο ένα δείγμα
- Η ιδέα του bootstrap:
  1. Πάρε πολλά δείγματα από μία *εκτίμηση του πληθυσμού*
  2. Υπολόγισε την εκτίμηση που προκύπτει σε κάθε δείγμα
  3. Η κατανομή που προκύπτει είναι η *bootstrap κατανομή του εκτιμητή*

Bootstrap

- Στατιστική συμπερασματολογία: βασίζεται στην εκτίμηση κάποιας ποσότητας της δειγματικής κατανομής κάποιου εκτιμητή
- Παράδειγμα: το τυπικό σφάλμα είναι η εκτίμηση της διασποράς της δειγματικής κατανομής του εκτιμητή
    - Πάρε πολλά δείγματα από τον *πληθυσμό*
    - Υπολόγισε την εκτίμηση που προκύπτει σε κάθε δείγμα
    - Η κατανομή που προκύπτει είναι η *δειγματική κατανομή του εκτιμητή*
- Στην πράξη όμως έχουμε μόνο ένα δείγμα
- Η ιδέα του bootstrap:
    1. Πάρε πολλά δείγματα από μία *εκτίμηση του πληθυσμού*
    2. Υπολόγισε την εκτίμηση που προκύπτει σε κάθε δείγμα
    3. Η κατανομή που προκύπτει είναι η *bootstrap κατανομή του εκτιμητή*
- Το βήμα 1 τελικά ισοδυναμεί με δειγματοληψία με επανάθεση από το παρατηρηθέν δείγμα!

Monte Carlo εκτιμητής

- Έστω ότι μελετούμε χαρακτηριστικό από κάποιον πληθυσμό με συνάρτηση κατανομής F.
- Ενδιαφέρει η εκτίμηση κάποιου συναρτησιακού $T(F)$
  - Παράδειγμα: γραμμικό συναρτησιακό

$$T(F) = \int \alpha(x) \mathrm{d}F(x)$$

- Αν γνωρίζουμε την κατανομή F τότε μπορούμε να προσομοιώσουμε ένα δείγμα $(x_1^\star, \ldots, x_B^\star)$ B τιμών από αυτήν και να υπολογίσουμε τον Monte Carlo εκτιμητή

$$\widehat{T}_B(F) = \frac{1}{B} \sum_{b=1}^{B} \alpha(x_b^\star)$$

- Monte Carlo εκτίμηση της διασποράς του εκτιμητή

$$v_{\mathrm{mc}} = \frac{1}{B} \sum_{i=1}^{B} \left(\alpha(x_b^\star) - \widehat{T}_B(F) \right)^2$$

Bootstrap

- Τι γίνεται όταν η συνάρτηση κατανομής F είναι άγνωστη;
- Έστω ότι έχουμε παρατηρήσει ένα *τυχαίο δείγμα* $(X_1, \ldots, X_n)$ από την F.
- Γνωρίζουμε ότι η εμπειρική συνάρτηση κατανομής

$$\widehat{F}_n(x) = \frac{1}{n} \sum_{i=1}^{n} \mathrm{I}\{X_i \leqslant x\}.$$

  είναι ένας συνεπής εκτιμητής της F.
- Ο plug-in εκτιμητής είναι

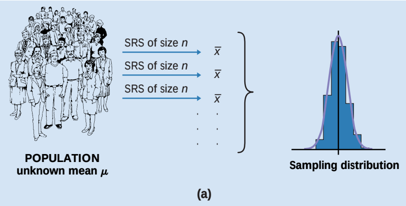
$$\widehat{\theta}_n := T(\widehat{F}_n).$$

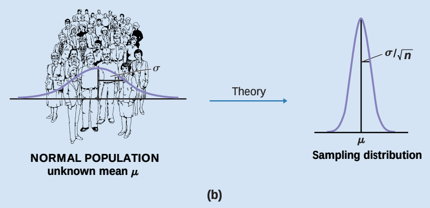
  - Ποιά είναι η διασπορά του εκτιμητή;
  - Γενικότερα: Ποιά είναι η δειγματική κατανομή του εκτιμητή;
- Η ιδέα του Bootstrap

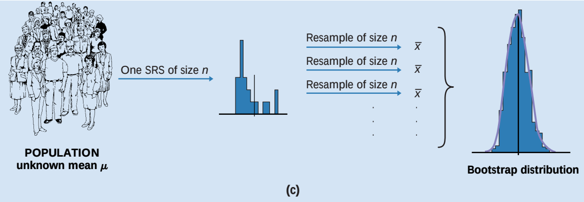
Μπορούμε να προσομοιώσουμε από την $\widehat{F}_n$ για να εκτιμήσουμε τη δειγματική κατανομή του plug-in εκτιμητή.

Προσομοίωση από την $\widehat{F}_n$

- Η $\widehat{F}_n$ είναι συνάρτηση κατανομής διακριτής τυχαίας μεταβλητής
- Το στήριγμα της είναι το σύνολο των n παρατηρηθέντων τιμών $x_1, \ldots, x_n$
- Κάθε μία από αυτές τις τιμές έχει πιθανότητα $\dfrac{1}{n}$
- Προσομοίωση από την $\widehat{F}_n \Leftrightarrow$ Τυχαία δειγματοληψία με επανάθεση από το παρατηρηθέν δείγμα.
- Για την εκτίμηση χαρακτηριστικών (τυπικό σφάλμα κλπ) του εκτιμητή $\widehat{\theta}_n := T(\widehat{F}_n)$ αρκεί να λάβουμε δείγματα (έστω B) με επανάθεση μεγέθους n.
- Για κάθε ένα προσομοιωμένο δείγμα υπολογίζουμε την τιμή του εκτιμητή $\widehat{\theta}_n$.







Παράδειγμα: $\widehat{\theta}_n = \overline{X}_n$

(a) «Πραγματικός» κόσμος: Αν συγκεντρώναμε «πολλά» δείγματα μεγέθους n από τον πληθυσμό και υπολογίζαμε για κάθε ένα την εκτίμηση $\overline{x}_n$ θα προέκυπτε η δειγματική κατανομή του $\overline{X}_n$

(b) Παραμετρικός κόσμος:

$$(X_1, \ldots, X_n) \overset{\text{iid}}{\sim} \mathcal{N}(\mu, \sigma^2) \Rightarrow$$
$$\overline{X}_n \sim \mathcal{N}\left(\mu, \frac{\sigma^2}{n}\right)$$

(c) (Μη παραμετρικός) Bootstrap κόσμος: Χρησιμοποιώ την πληροφορία που υπάρχει στο δείγμα ώστε να προσομοιώσω πολλά δείγματα μεγέθους n από ~~τον πληθυσμό~~ μία εκτίμηση του πληθυσμού και να εκτιμήσω την δειγματική κατανομή του στατιστικού από την bootstrap κατανομή.

Εικόνα από:
Hesterberg et al (2003). Bootstrap methods and permutation tests. Freeman and Company, New York.

Γενικός αλγόριθμός Bootstrap

- Δεδομένα: $(X_1, \ldots, X_n)$ τ.δ. από συνάρτηση κατανομής F
- Εκτιμητής: $\widehat{\theta}_n = \widehat{\theta}_n(X_1, \ldots, X_n)$ συναρτησιακού $\theta = T(F)$
- Στόχος: δειγματική κατανομή του $\widehat{\theta}_n$.

1 Για $b = 1, \ldots, B$

1.1 Κάνε δειγματοληψία με επανάθεση n τιμών από τα $(x_1, \ldots, x_n)$

$$(x^\star_{b,1}, \ldots, x^\star_{b,n})$$

(ισοδύναμει με προσομοίωση ένα δείγματος μεγέθους n από την $\widehat{F}_n$)

1.2 Υπολόγισε τον εκτιμητή για το συγκεκριμένο bootstrap δείγμα

$$\theta^\star_{n,b} := \widehat{\theta}_n(x^\star_{b,1}, \ldots, x^\star_{b,n})$$

2 Μέσω των $\{\theta^\star_{n,b}; b = 1, \ldots, B\}$ εκτιμάται η δειγματική κατανομή του $\widehat{\theta}_n = \widehat{\theta}_n(X_1, \ldots, X_n)$, η οποία μπορεί να χρησιμοποιηθεί για περαιτέρω συμπερασματολογία όπως:

- εκτίμηση τυπικού σφάλματος του εκτιμητή
- κατασκευή διαστημάτων εμπιστοσύνης

Bootstrap εκτίμηση τυπικού σφάλματος

- Έστω
$$\theta_n^\star := \frac{1}{B}\sum_{b=1}^{B}\theta_{n,b}^\star. \tag{1}$$
- bootstrap εκτίμηση της διασποράς του εκτιμητή
$$v_{\text{boot}} = \frac{1}{B}\sum_{b=1}^{B}\left(\theta_{n,b}^\star - \theta_n^\star\right)^2. \tag{2}$$
- Η bootstrap εκτίμηση του τυπικού σφάλματος του εκτιμητή είναι
$$\text{se}_{\text{boot}} = \sqrt{v_{\text{boot}}}. \tag{3}$$

Bootstrap εκτίμηση μεροληψίας

- Μεροληψία του $\widehat{\theta}_n$

$$\text{bias}(\widehat{\theta}_n) = \mathrm{E}\widehat{\theta}_n - \theta$$

- Bootstrap εκτίμηση της μεροληψίας

$$\widehat{\text{bias}}(\widehat{\theta}_n) = \theta_n^\star - \widehat{\theta}_n \tag{4}$$

όπου $\theta_n^\star$ ορίστηκε στην εξίσωση (1).

- Προσοχή:
  - το $\widehat{\text{bias}}(\widehat{\theta}_n)$ είναι **εκτίμηση** της μεροληψίας του εκτιμητή
  - Μπορεί να είναι διαφορετικό από μηδέν ακόμα για αμερόληπτους εκτιμητές
  - Πάντα πρέπει να λαμβάνεται υπόψη και το τυπικό σφάλμα του εκτιμητή.

Bootstrap στον δειγματικό μέσο

Έστω τυχαίο δείγμα $n = 10$ παρατηρήσεων

- $\boldsymbol{x} = (12.8, 8.6, 5.8, 8.6, 12.8, 12.2, 14.5, 7.8, 6, 13.6)$
- τυπικό σφάλμα του $\overline{X}_n = \sqrt{\sigma^2/n}$
- $\bar{x} \approx 10.27$, (εκτιμηθέν[1]) τυπικό σφάλμα $\widehat{s.e.(\bar{x})} \approx 1.03$

Bootstrap δείγματα

	Δείγμα										Μέσος
1	6	8.6	5.8	14.5	13.6	13.6	14.5	12.8	13.6	14.5	11.75

Εκτίμηση τυπικού σφάλματος από τα Bootstrap δείγματα:

[1] εδώ χρησιμοποιήθηκε ο $S_n^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X}_n)^2$ ως εκτιμητής του σ^2

Bootstrap στον δειγματικό μέσο

Έστω τυχαίο δείγμα $n = 10$ παρατηρήσεων

- $\boldsymbol{x} = (12.8, 8.6, 5.8, 8.6, 12.8, 12.2, 14.5, 7.8, 6, 13.6)$
- τυπικό σφάλμα του $\overline{X}_n = \sqrt{\sigma^2/n}$
- $\bar{x} \approx 10.27$, (εκτιμηθέν[1]) τυπικό σφάλμα $\widehat{s.e.(\bar{x})} \approx 1.03$

Bootstrap δείγματα

	Δείγμα										Μέσος
1	6	8.6	5.8	14.5	13.6	13.6	14.5	12.8	13.6	14.5	11.75
2	12.8	12.8	6	8.6	8.6	7.8	8.6	6	6	7.8	8.50

Εκτίμηση τυπικού σφάλματος από τα Bootstrap δείγματα: 2.30

[1] εδώ χρησιμοποιήθηκε ο $S_n^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X}_n)^2$ ως εκτιμητής του σ^2

Bootstrap στον δειγματικό μέσο

Έστω τυχαίο δείγμα $n = 10$ παρατηρήσεων

- $\boldsymbol{x} = (12.8, 8.6, 5.8, 8.6, 12.8, 12.2, 14.5, 7.8, 6, 13.6)$
- τυπικό σφάλμα του $\overline{X}_n = \sqrt{\sigma^2/n}$
- $\bar{x} \approx 10.27$, (εκτιμηθέν[1]) τυπικό σφάλμα $\widehat{s.e.(\bar{x})} \approx 1.03$

Bootstrap δείγματα

	Δείγμα										Μέσος
1	6	8.6	5.8	14.5	13.6	13.6	14.5	12.8	13.6	14.5	11.75
2	12.8	12.8	6	8.6	8.6	7.8	8.6	6	6	7.8	8.50
3	6	12.8	8.6	14.5	8.6	8.6	8.6	7.8	14.5	13.6	10.36

Εκτίμηση τυπικού σφάλματος από τα Bootstrap δείγματα: 1.63

[1] εδώ χρησιμοποιήθηκε ο $S_n^2 = \frac{1}{n-1}\sum_{i=1}^n (X_i - \overline{X}_n)^2$ ως εκτιμητής του σ^2

Bootstrap στον δειγματικό μέσο

Έστω τυχαίο δείγμα $n = 10$ παρατηρήσεων

- $\boldsymbol{x} = (12.8, 8.6, 5.8, 8.6, 12.8, 12.2, 14.5, 7.8, 6, 13.6)$
- τυπικό σφάλμα του $\overline{X}_n = \sqrt{\sigma^2/n}$
- $\bar{x} \approx 10.27$, (εκτιμηθέν[1]) τυπικό σφάλμα $\widehat{s.e.(\bar{x})} \approx 1.03$

Bootstrap δείγματα

	Δείγμα										Μέσος
1	6	8.6	5.8	14.5	13.6	13.6	14.5	12.8	13.6	14.5	11.75
2	12.8	12.8	6	8.6	8.6	7.8	8.6	6	6	7.8	8.50
3	6	12.8	8.6	14.5	8.6	8.6	8.6	7.8	14.5	13.6	10.36
4	12.8	12.8	8.6	7.8	6	12.8	12.8	12.8	5.8	12.8	10.50

Εκτίμηση τυπικού σφάλματος από τα Bootstrap δείγματα: 1.34

[1] εδώ χρησιμοποιήθηκε ο $S_n^2 = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \overline{X}_n)^2$ ως εκτιμητής του σ^2

Bootstrap στον δειγματικό μέσο

Έστω τυχαίο δείγμα $n = 10$ παρατηρήσεων

- $\boldsymbol{x} = (12.8, 8.6, 5.8, 8.6, 12.8, 12.2, 14.5, 7.8, 6, 13.6)$
- τυπικό σφάλμα του $\overline{X}_n = \sqrt{\sigma^2/n}$
- $\bar{x} \approx 10.27$, (εκτιμηθέν[1]) τυπικό σφάλμα $\widehat{s.e.(\bar{x})} \approx 1.03$

Bootstrap δείγματα

	Δείγμα										Μέσος
1	6	8.6	5.8	14.5	13.6	13.6	14.5	12.8	13.6	14.5	11.75
2	12.8	12.8	6	8.6	8.6	7.8	8.6	6	6	7.8	8.50
3	6	12.8	8.6	14.5	8.6	8.6	8.6	7.8	14.5	13.6	10.36
4	12.8	12.8	8.6	7.8	6	12.8	12.8	12.8	5.8	12.8	10.50
5	7.8	12.8	6	8.6	5.8	13.6	12.8	12.2	6	14.5	10.01

Εκτίμηση τυπικού σφάλματος από τα Bootstrap δείγματα: 1.17

[1] εδώ χρησιμοποιήθηκε ο $S_n^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X}_n)^2$ ως εκτιμητής του σ^2

Bootstrap στον δειγματικό μέσο

Έστω τυχαίο δείγμα $n = 10$ παρατηρήσεων

- $\boldsymbol{x} = (12.8, 8.6, 5.8, 8.6, 12.8, 12.2, 14.5, 7.8, 6, 13.6)$
- τυπικό σφάλμα του $\overline{X}_n = \sqrt{\sigma^2/n}$
- $\bar{x} \approx 10.27$, (εκτιμηθέν[1]) τυπικό σφάλμα $\widehat{s.e.(\bar{x})} \approx 1.03$

Bootstrap δείγματα

	Δείγμα										Μέσος
1	6	8.6	5.8	14.5	13.6	13.6	14.5	12.8	13.6	14.5	11.75
2	12.8	12.8	6	8.6	8.6	7.8	8.6	6	6	7.8	8.50
3	6	12.8	8.6	14.5	8.6	8.6	8.6	7.8	14.5	13.6	10.36
4	12.8	12.8	8.6	7.8	6	12.8	12.8	12.8	5.8	12.8	10.50
5	7.8	12.8	6	8.6	5.8	13.6	12.8	12.2	6	14.5	10.01
6	12.8	12.2	12.8	12.8	8.6	8.6	8.6	12.8	7.8	5.8	10.28

Εκτίμηση τυπικού σφάλματος από τα Bootstrap δείγματα: 1.04

[1] εδώ χρησιμοποιήθηκε ο $S_n^2 = \frac{1}{n-1}\sum_{i=1}^n (X_i - \overline{X}_n)^2$ ως εκτιμητής του σ^2

Bootstrap στον δειγματικό μέσο

Έστω τυχαίο δείγμα $n = 10$ παρατηρήσεων

- $\boldsymbol{x} = (12.8, 8.6, 5.8, 8.6, 12.8, 12.2, 14.5, 7.8, 6, 13.6)$
- τυπικό σφάλμα του $\overline{X}_n = \sqrt{\sigma^2/n}$
- $\bar{x} \approx 10.27$, (εκτιμηθέν[1]) τυπικό σφάλμα $\widehat{s.e.(\bar{x})} \approx 1.03$

Bootstrap δείγματα

	Δείγμα										Μέσος
1	6	8.6	5.8	14.5	13.6	13.6	14.5	12.8	13.6	14.5	11.75
2	12.8	12.8	6	8.6	8.6	7.8	8.6	6	6	7.8	8.50
3	6	12.8	8.6	14.5	8.6	8.6	8.6	7.8	14.5	13.6	10.36
4	12.8	12.8	8.6	7.8	6	12.8	12.8	12.8	5.8	12.8	10.50
5	7.8	12.8	6	8.6	5.8	13.6	12.8	12.2	6	14.5	10.01
6	12.8	12.2	12.8	12.8	8.6	8.6	8.6	12.8	7.8	5.8	10.28
...	...	...	...	...	...	...	...	...	...	...	...

Εκτίμηση τυπικού σφάλματος από τα Bootstrap δείγματα: 1.04

[1] εδώ χρησιμοποιήθηκε ο $S_n^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X}_n)^2$ ως εκτιμητής του σ^2

Bootstrap στον δειγματικό μέσο

Έστω τυχαίο δείγμα $n = 10$ παρατηρήσεων

- $\boldsymbol{x} = (12.8, 8.6, 5.8, 8.6, 12.8, 12.2, 14.5, 7.8, 6, 13.6)$
- τυπικό σφάλμα του $\overline{X}_n = \sqrt{\sigma^2/n}$
- $\bar{x} \approx 10.27$, (εκτιμηθέν[1]) τυπικό σφάλμα $\widehat{s.e.(\bar{x})} \approx 1.03$

Bootstrap δείγματα

	Δείγμα										Μέσος
1	6	8.6	5.8	14.5	13.6	13.6	14.5	12.8	13.6	14.5	11.75
2	12.8	12.8	6	8.6	8.6	7.8	8.6	6	6	7.8	8.50
3	6	12.8	8.6	14.5	8.6	8.6	8.6	7.8	14.5	13.6	10.36
4	12.8	12.8	8.6	7.8	6	12.8	12.8	12.8	5.8	12.8	10.50
5	7.8	12.8	6	8.6	5.8	13.6	12.8	12.2	6	14.5	10.01
6	12.8	12.2	12.8	12.8	8.6	8.6	8.6	12.8	7.8	5.8	10.28
...	...	...	...	...	...	...	...	...	...	...	...
10^3	5.8	5.8	13.6	5.8	12.8	12.2	13.6	8.6	8.6	7.8	9.46

Εκτίμηση τυπικού σφάλματος από τα Bootstrap δείγματα: 0.98

[1] εδώ χρησιμοποιήθηκε ο $S_n^2 = \frac{1}{n-1}\sum_{i=1}^n (X_i - \overline{X}_n)^2$ ως εκτιμητής του σ^2

Παιχνίδι

- Έστω το προηγούμενο τύχαιο δείγμα 10 παρατηρήσεων

$$\boldsymbol{x} = (12.8, 8.6, 5.8, 8.6, 12.8, 12.2, 14.5, 7.8, 6, 13.6)$$

- Ανοίξτε την R
- Προσομοιώστε $B = 1000$ bootstrap δείγματα για να εκτιμήσετε την δειγματική κατανομή του δειγματικού μέσου $\widehat{\theta}_n = \overline{X}_n$ και με βάση αυτό
  - Εκτιμήστε το τυπικό σφάλμα του $\widehat{\theta}_n$
  - Εκτιμήστε την μεροληψία του $\widehat{\theta}_n$

Παράδειγμα 2: Bootstrap στην κανονική κατανομή

- Έστω τυχαίο δείγμα $n = 20$ τιμών από $\mathcal{N}(\mu, \sigma^2)$
  19.15 19.43 18.20 21.41 24.18 23.27 18.92 18.79 19.59 19.38 21.91 19.22
  19.71 19.26 21.22 16.67 17.56 22.55 20.47 21.87
- Θα χρησιμοποιήσουμε bootstrap για να εκτιμήσουμε (μη παραμετρικά) τις δειγματικές κατανομές των

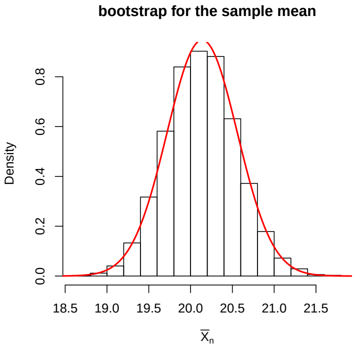
$$\overline{X}_n = \frac{1}{n} \sum_{i=1}^{n} X_i$$

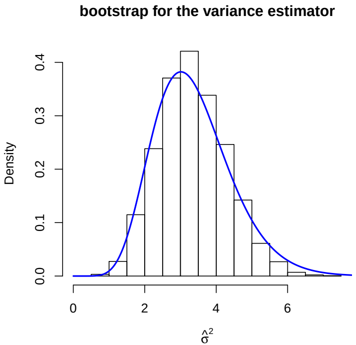
και

$$\widehat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^{n} \left(X_i - \overline{X}_n \right)^2$$

- και θα συγκρίνουμε με τις κατανομές που προκύπτουν κάνοντας χρήση γνωστών παραμετρικών αποτελεσμάτων.

Παράδειγμα 2: Bootstrap στην κανονική κατανομή





- Ιστογράμματα: $B = 10000$ bootstrap δείγματα
- Συναρτήσεις πυκνότητας των δειγματικών κατανομών[2,3]

$$\overline{X}_n \sim \mathcal{N}\left(\mu, \frac{\sigma^2}{n}\right), \qquad \widehat{\sigma}^2 \sim \mathcal{G}\left(\frac{n-1}{2}, \frac{n}{2\sigma^2}\right)$$

[2] Υπενθύμιση: $\frac{\sum_i (X_i - \overline{X}_n)^2}{\sigma^2} \sim \mathcal{X}^2_{n-1} \equiv \mathcal{G}((n-1)/2, 1/2)$ (shape - rate)

[3] Αντικαθιστώ τα μ και σ^2 με τις εκτιμήσεις $\overline{x}_n$ και $\widehat{\sigma}^2$

Παράδειγμα 2: Bootstrap στην κανονική κατανομή

- Εκτιμήσεις τυπικών σφαλμάτων και μεροληψίας

	$\overline{X}_n$		$\widehat{\sigma}^2$	
	$\widehat{\text{s.e.}}$	$\widehat{\text{bias}}$	$\widehat{\text{s.e.}}$	$\widehat{\text{bias}}$
jackknife	0.432	0	1.053	−0.1866
bootstrap	0.419	−0.0040	0.955	−0.1874
αναλυτικά	0.432	0	1.150	−0.1866

- bootstrap δείγματα μεγέθους $B = 10000$
- Οι αναλυτικές εκφράσεις για τα τυπικά σφάλματα είναι

$$s.e.(\overline{X}_n) = \sqrt{\frac{\sigma^2}{n}} \Rightarrow \widehat{s.e.}(\overline{X}_n) = \sqrt{\frac{S_n^2}{n}}$$

$$s.e.(\widehat{\sigma}^2) = \sqrt{\frac{2(n-1)}{n^2}\sigma^4} \Rightarrow \widehat{s.e.}(\widehat{\sigma}^2) = \sqrt{\frac{2(n-1)}{n^2}S_n^4}$$

- $\text{bias}(\widehat{\sigma}^2) = \text{E}\widehat{\sigma}^2 - \sigma^2 = (\frac{n-1}{n} - 1)\sigma^2 = -\frac{1}{n}\sigma^2 \Rightarrow \widehat{\text{bias}}(\widehat{\sigma}^2) = -\frac{1}{n}S_n^2$

Bootstrap Διαστήματα Εμπιστοσύνης

Bootstrap Διαστήματα Εμπιστοσύνης

- Έστω ότι ο $\widehat{\theta}_n$ (πχ $\widehat{\theta}_n = T(\widehat{F}_n)$) εκτιμά κάποια μονοδιάστατη ποσότητα θ
- c_α: (κάτω) α-ποσοστιαίο σημείο της κατανομής (έστω H) της $\widehat{\theta}_n - \theta$, $\alpha \in (0,1)$

$$\mathrm{P}\left(\widehat{\theta}_n - \theta \leqslant c_{\alpha/2}\right) = \frac{\alpha}{2} = \mathrm{P}\left(\widehat{\theta}_n - \theta \geqslant c_{1-\alpha/2}\right)$$

- $\widehat{\theta}_n - \theta \leqslant c_{\alpha/2} \Leftrightarrow \theta \geqslant \widehat{\theta}_n - c_{\alpha/2}$
- $\widehat{\theta}_n - \theta \geqslant c_{1-\alpha/2} \Leftrightarrow \theta \leqslant \widehat{\theta}_n - c_{1-\alpha/2}$
- Το τυχαίο διάστημα

$$\left[\widehat{\theta}_n - c_{1-\alpha/2}, \widehat{\theta}_n - c_{\alpha/2}\right] = \left[\widehat{\theta}_n - H^{-1}(1-\alpha/2), \widehat{\theta}_n - H^{-1}(\alpha/2)\right] \tag{5}$$

  είναι ένα $100(1-\alpha)\%$ Διάστημα Εμπιστοσύνης για το θ
- Αυτή η (ιδανική) περιγραφή του προβλήματος είναι σπανίως εφαρμόσιμη διότι η κατανομή του $\widehat{\theta}_n - \theta$ είναι άγνωστη
- Όλες οι μέθοδοι που θα περιγράψουμε έχουν σκοπό να προσεγγίσουν τα ποσοστημόρια της κατανομής του $\widehat{\theta}_n - \theta$.

Bootstrap Διαστήματα Εμπιστοσύνης

1. Κανονικά
2. Βασικά (Basic/Pivotal)
3. Ποσοστιαίων σημείων (Percentile)
4. Studentized
5. adjusted percentile (Bias corrected)

Κανονικά Bootstrap Διαστήματα Εμπιστοσύνης (I)

- Υποθέτουμε ότι η κατανομή του $\widehat{\theta}_n - \theta$ προσεγγίζεται από μία κανονική κατανομή $\mathcal{N}(0, v^2)$
- Αν η διασπορά ήταν γνωστή, τα άκρα του Δ.Ε. θα ήταν

$$\left[\widehat{\theta}_n - z_{\alpha/2} v, \widehat{\theta}_n + z_{\alpha/2} v\right]$$

όπου $z_{\alpha/2}$ το **άνω**-$\alpha/2$ ποσοστιαίο σημείο της $\mathcal{N}(0,1)$

- Στην πράξη η v είναι άγνωστη
- Εκτιμάται από την Bootstrap εκτίμηση του τυπικού σφάλματος (3).
- Οπότε το

$$\left[\widehat{\theta}_n - z_{\alpha/2} \mathrm{se}_{\mathrm{boot}}, \widehat{\theta}_n + z_{\alpha/2} \mathrm{se}_{\mathrm{boot}}\right]$$

είναι ένα $100(1-\alpha)\%$ Bootstrap Διάστημα Εμπιστοσύνης για το θ βασισμένο στην Κανονική προσέγγιση.

Κανονικά Bootstrap Διαστήματα Εμπιστοσύνης (II)

- Το προηγούμενο όμως δεν λαμβάνει υπόψη την μεροληψία του $\widehat{\theta}_n$
- Υποθέτουμε ότι η κατανομή του $\widehat{\theta}_n - \theta$ προσεγγίζεται από μία κανονική κατανομή $\mathcal{N}(\beta, v^2)$
- β: μεροληψία του $\widehat{\theta}_n$
- Αν η μεροληψία και η διασπορά ήταν γνωστές, τα άκρα του Δ.Ε. θα ήταν
$$\left[\widehat{\theta}_n - \beta - z_{\alpha/2}v, \widehat{\theta}_n - \beta + z_{\alpha/2}v\right]$$
- Η μεροληψία είναι επίσης άγνωστη (όπως και η διασπορά)
- Χρησιμοποιούμε την εξίσωση (4) για να την εκτιμήσουμε μέσω bootstrap
$$\text{bias}(\widehat{\theta}_n) = \theta_n^\star - \widehat{\theta}_n$$
- Οπότε το
$$\left[2\widehat{\theta}_n - \theta_n^\star - z_{\alpha/2}\text{se}_{\text{boot}}, 2\widehat{\theta}_n - \theta_n^\star + z_{\alpha/2}\text{se}_{\text{boot}}\right] \tag{6}$$
είναι ένα $100(1-\alpha)\%$ Bootstrap Διάστημα Εμπιστοσύνης για το θ βασισμένο στην Κανονική προσέγγιση, με διόρθωση μεροληψίας.

Βασικά Bootstrap ΔΕ (Ι)

- Ενδιαφέρει η συμπερασματολογία για την συνάρτηση κατανομής του

$$\widehat{\theta}_n - \theta$$

- Χρησιμοποιούμε τις προσομοιωμένες τιμές

$$\{\theta^{\star}_{n,b} - \widehat{\theta}_n; b = 1, \ldots, B\}$$

  για να προσεγγίσουμε την κατανομή του $\widehat{\theta}_n - \theta$.
- Έστω η πιθανότητα

$$H(u) = \mathrm{P}(\widehat{\theta}_n - \theta \leqslant u)$$

- Εκτίμηση

$$\widehat{H}_n(u) = \frac{\#\{\theta^{\star}_{n,b} - \widehat{\theta}_n \leqslant u\}}{B} = \frac{1}{B}\sum_{b=1}^{B} \mathrm{I}\{\theta^{\star}_{n,b} - \widehat{\theta}_n \leqslant u\}$$

- Χρησιμοποιούμε την εμπειρική συνάρτηση κατανομής $\widehat{H}_n(u)$ για να εκτιμήσουμε την πραγματική $H(u)$.

Βασικά Bootstrap ΔΕ (II)

- Έστω $c_\alpha = H^{-1}(\alpha)$ το κάτω α-ποσοστιαίο σημείο της H
- Χρησιμοποιούμε τον εκτιμητή αντικατάστασης

$$\widehat{c}_\alpha = \widehat{H}_n^{-1}(\alpha) = \inf\{u : \widehat{H}_n(u) \geqslant \alpha\}$$

- Οπότε αρκεί να βρω το bootstrap α-ποσοστιαίο σημείο των προσομοιωμένων τιμών

$$\{\theta^\star_{n,b} - \widehat{\theta}_n; b = 1, \ldots, B\}.$$

το οποίο είναι ίσο με

$$\widehat{c}_\alpha = \theta^\star_{[n,j_\alpha]} - \widehat{\theta}_n \tag{7}$$

$$\text{όπου } j_\alpha \in \{1, 2, \ldots, B\} : \frac{j_\alpha - 1}{B} < \alpha \leqslant \frac{j_\alpha}{B}$$

$$= \lfloor B\alpha + 0.5 \rfloor \tag{8}$$

και με $\theta^\star_{[n,1]} \leqslant \theta^\star_{[n,2]} \leqslant \ldots \leqslant \theta^\star_{[n,B]}$ συμβολίζουμε το διατεταγμένο bootstrap δείγμα.

- Πχ στην R: `quantile(`$\{\theta^\star_{n,b} - \widehat{\theta}_n; b = 1, \ldots, B\}$`, type = 1)`.

Βασικά Bootstrap ΔΕ (III)

- Συνεπώς το $100(1-\alpha)\%$ βασικό Bootstrap ΔΕ είναι

$$\left[\widehat{\theta}_n - \widehat{H}_B^{-1}(1-\alpha/2), \widehat{\theta}_n - \widehat{H}_B^{-1}(\alpha/2)\right] \tag{9}$$

- Αντιπαραθέστε την παραπάνω έκφραση με την (5)
- Αντικαθιστώντας τις (7) και (8) στην (9) προκύπτει ότι μία ισοδύναμη έκφραση για το $100(1-\alpha)\%$ βασικό Bootstrap διάστημα εμπιστοσύνης είναι

$$\left[2\widehat{\theta}_n - \theta^\star_{[n,\lfloor B(1-\alpha/2)+0.5\rfloor]}, 2\widehat{\theta}_n - \theta^\star_{[n,\lfloor B\alpha/2+0.5\rfloor]}\right], \tag{10}$$

  όπου
  - $\theta^\star_{[n,1]} \leqslant \theta^\star_{[n,2]} \leqslant \ldots \leqslant \theta^\star_{[n,B]}$ το διατεταγμένο bootstrap δείγμα.
  - $\lfloor y \rfloor$ το ακέραιο μέρος του y.

Bootstrap ΔΕ ποσοστιαίων σημείων

- Έστω G η συνάρτηση κατανομής των $\theta_n^\star$
- Η οποία εκτιμάται από την εμπειρική bootstrap κατανομή $\widehat{G}_n^\star$
- Το $100(1-\alpha)\%$ Bootstrap ΔΕ ποσοστιαίων σημείων (Bootstrap percentile CI) ορίζεται ως

$$\left[\widehat{G}_n^{\star-1}(\alpha/2), \widehat{G}_n^{\star-1}(1-\alpha/2)\right] = \left[\theta^\star_{[n,\lfloor B\alpha/2+0.5\rfloor]}, \theta^\star_{[n,\lfloor B(1-\alpha/2)+0.5\rfloor]}\right] \tag{11}$$

- Παρατηρήσεις
  - Η παραπάνω διαδικασία είναι η πιο άμεση από όλες: αρκεί να υπολογίσουμε τα κατάλληλα ποσοστιαία σημεία του Bootstrap δείγματος.
  - Στην πραγματικότητα, η μεθοδολογία αυτή στηρίζεται στην υπόθεση ότι υπάρχει ένας μετασχηματισμός m τέτοιος ώστε

$$m(\widehat{\theta}_n) \sim \mathcal{N}\left(m(\theta), c^2\right)$$

  - Απόδειξη: σελίδες 62-63 σημείωσεων Ιωαννίδη.
  - Η γνώση του μετασχηματισμού αυτού δεν παίζει ρόλο!

Βελτιωμένες μέθοδοι κατασκευής Bootstrap ΔΕ

- Studentized bootstrap ΔΕ
  - Βελτιώνουν τα κανονικά bootstrap ΔΕ
  - Είναι της μορφής
$$[\widehat{\theta}_n + \zeta_{\alpha/2}\mathrm{se}_{\mathrm{boot}}, \widehat{\theta}_n + \zeta_{1-\alpha/2}\mathrm{se}_{\mathrm{boot}}]$$
  - με ζ_α συμβολίζουμε το **κάτω** α-ποσοστιαίο σημείο της κατανομής των
$$\frac{\theta^\star_{n,b} - \widehat{\theta}_n}{\mathrm{se}(\theta^\star_{n,b})}.$$
  - Πλεονέκτημα: πιο ακριβής μεθοδολογία
  - Μειονέκτημα: χρειάζεται εκτίμηση του $\mathrm{se}(\theta^\star_{n,b})$
    - Jackknife
    - Bootstrap

  για κάθε bootstrap επανάληψη $b = 1, \ldots, B$.
- adjusted percentile (Bias Corrected - BC_α)
  - Βελτιώνουν τα Bootstrap ΔΕ ποσοστιαίων σημείων
  - λαμβάνοντας υπόψη διορθώσεις μεροληψίας και ασυμμετρίας
  - Δες Davison, A.C. and Hinkley, D.V. (1997), ενότητα 5.3.2

Παράδειγμα 2: Παράμετροι κανονικής κατανομής

- Προσομοιωμένο τυχαίο δείγμα $n = 20$ παρατηρήσεων από $\mathcal{N}(\mu = 20, \sigma^2 = 4)$
- 95% (ατομικά) διαστήματα εμπιστοσύνης

	μ_{low}	μ_{up}	σ^2_{low}	σ^2_{up}
exact	19.23	21.04	2.16	7.96
jackknife (normal)	19.29	20.99	1.67	5.80
bootstrap (normal)	19.33	20.96	1.87	5.59
bootstrap (basic)	19.30	20.94	1.76	5.45
bootstrap (percentile)	19.34	20.97	1.64	5.33
bootstrap (studentized)	19.08	21.03	-0.48	5.02
bootstrap (BC_α)	19.38	21.02	2.10	6.24

- $B = 10^4$ bootstrap επαναλήψεις
- Τα studentized bootstrap διαστήματα υπολογίστηκαν με διπλό bootstrap 50 επαναλήψεων (για κάθε bootstrap επανάληψη).
- Υπενθύμιση: $100(1-\alpha)\%$ ακριβή διαστήματα εμπιστοσύνης ίσων ουρών
  - Μέση τιμή: $\left[\overline{X}_n - t_{n-1;\alpha/2} S_n/\sqrt{n}, \overline{X}_n + t_{n-1;\alpha/2} S_n/\sqrt{n}\right]$
  - Διασπορά: $\left[\dfrac{(n-1)S_n^2}{\mathcal{X}^2_{n-1;\alpha/2}}, \dfrac{(n-1)S_n^2}{\mathcal{X}^2_{n-1;1-\alpha/2}}\right]$

Πιθανότητα κάλυψης: μελέτη Monte Carlo

- Προσομοιώνω $(X_1, \ldots, X_n)$ από $\mathcal{N}(\mu, \sigma^2)$
- Υπολογίζω τα προηγούμενα διαστήματα εμπιστοσύνης με συντελεστή εμπιστοσύνης 95%
- Εξετάζω αν οι πραγματικές τιμές των παραμέτρων μ και σ^2 ανήκουν ή όχι εντός του εκάστοτε ΔΕ
- Επαναλαμβάνω τη διαδικασία αυτή πολλές φορές (M) ώστε να λάβω την Monte Carlo εκτίμηση των αριστερών και δεξιών πιθανοτήτων αστοχίας:

$$\widehat{\mathrm{P}}(\theta_{\text{low}} > \theta) = \frac{\text{αριθμός φορών που } \theta_{\text{low}} > \theta}{M}$$

$$\widehat{\mathrm{P}}(\theta_{\text{up}} < \theta) = \frac{\text{αριθμός φορών που } \theta_{\text{up}} < \theta}{M}$$

- Φυσικά, για τα ακριβή 95% ΔΕ εμπιστοσύνης ίσων ουρών γνωρίζουμε ότι $\mathrm{P}(\theta_{\text{low}} > \theta) = \mathrm{P}(\theta_{\text{up}} < \theta) = 0.025$.

Πιθανότητα κάλυψης: μελέτη Monte Carlo

Αριστερή και δεξιά πιθανότητα αστοχίας για $\theta = \mu$

	$n = 50$		$n = 100$		$n = 200$	
	μ_{low}	μ_{up}	μ_{low}	μ_{up}	μ_{low}	μ_{up}
exact	0.026	0.025	0.026	0.026	0.026	0.024
jackknife	0.032	0.029	0.027	0.028	0.027	0.025
bootstrap normal	0.032	0.030	0.028	0.028	0.028	0.025
bootstrap basic	0.032	0.030	0.029	0.028	0.028	0.025
bootstrap percentile	0.033	0.030	0.028	0.028	0.028	0.025
bootstrap BCa	0.033	0.030	0.028	0.028	0.028	0.024
bootstrap studentized	0.025	0.022	0.023	0.023	0.024	0.022

Πιθανότητα κάλυψης: μελέτη Monte Carlo

Αριστερή και δεξιά πιθανότητα αστοχίας για $\theta = \sigma^2$

	$n = 50$		$n = 100$		$n = 200$	
	σ^2_{low}	σ^2_{up}	σ^2_{low}	σ^2_{up}	σ^2_{low}	σ^2_{up}
exact	0.023	0.025	0.022	0.026	0.025	0.023
jackknife	0.009	0.078	0.008	0.057	0.014	0.044
bootstrap normal	0.011	0.085	0.009	0.061	0.015	0.046
bootstrap basic	0.008	0.098	0.007	0.070	0.012	0.054
bootstrap percentile	0.007	0.095	0.007	0.064	0.014	0.048
bootstrap BCa	0.026	0.048	0.024	0.038	0.026	0.026
bootstrap studentized	0.000	0.127	0.001	0.092	0.003	0.069

`boot` R package

- Περιέχει συναρτήσεις για παραμετρικό και μη παραμετρικό Bootstrap
  - `boot()`
- Υπολογισμός Bootstrap διαστημάτων εμπιστοσύνης
  - `boot.ci()`
- Angelo Canty and Brian Ripley (2019). `boot`: Bootstrap R (S-Plus) Functions. R package version 1.3-24.
- https://CRAN.R-project.org/package=boot
- Στα εργαστηριακά μαθήματα θα το χρησιμοποιούμε για να συγκρίνουμε τα αποτελέσματά μας

Παράδειγμα 3: Bootstrap στην παλινδρομηση

$$Y_i = \alpha + \beta x_i + e_i, \quad i = 1, \ldots, n$$

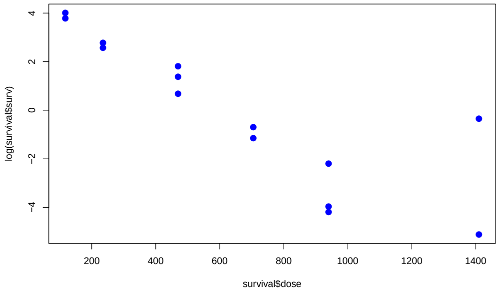
όπου e_i τυχαίο δείγμα από κάποια κατανομή (όχι απαραίτητα κανονική).

- Δύο προσεγγίσεις
  1. Δειγματοληψία με επανάθεση από τα ζεύγη $(x_1, Y_1), (x_2, Y_2), \ldots, (x_n, Y_n)$
     - Στη συνέχεια υπολογίζουμε τις εκτιμήσεις που μας ενδιαφέρουν (πχ τους εκτιμητές ελαχιστών τετραγώνων) για κάθε bootstrap δείγμα.
  2. Δειγματοληψία με επανάθεση στα (εκτιμηθέντα) κατάλοιπα $\varepsilon_i = Y_i - \widehat{\alpha} - \widehat{\beta} x_i$, $i = 1, \ldots, n$, όπου $\widehat{\alpha}$ και $\widehat{\beta}$ είναι οι εκτιμητές ελαχιστών τετραγώνων.
     - Τα bootstrap δείγματα υπολογίζονται ως

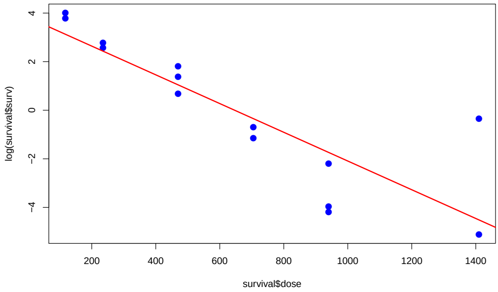
       $$Y_i^\star = \widehat{\alpha} + \widehat{\beta} x_i + \varepsilon_i^\star$$

       όπου $\varepsilon_i^\star$ δείγματα bootstrap από τα ε_i, $i = 1, \ldots, n$
     - Στη συνέχεια υπολογίζουμε τις ποσότητες που μας ενδιαφέρουν για κάθε bootstrap δείγμα
- Στην πράξη και οι δύο μέθοδοι δίνουν παρόμοια αποτελέσματα
- Η 2η μέθοδος είναι θεωρητικά πιο σωστή (γιατί;)

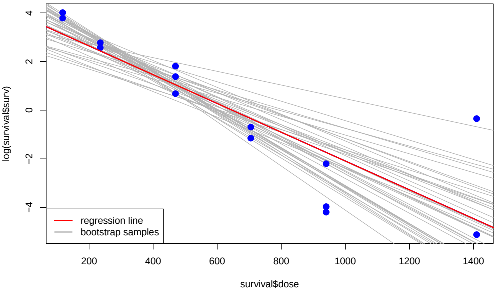
Παράδειγμα 3: Bootstrap στην παλινδρομηση



Παράδειγμα 3: Bootstrap στην παλινδρομηση

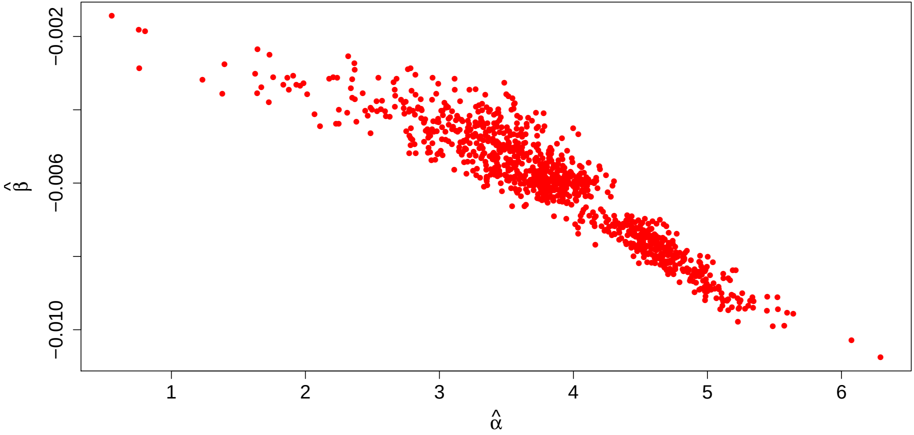


Παράδειγμα 3: Bootstrap στην παλινδρομηση



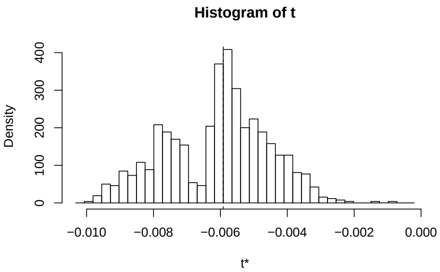
Παράδειγμα 3: Bootstrap στην παλινδρομηση

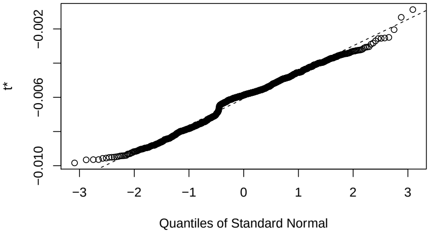
Bootstrap εκτίμηση της από κοινού κατανομής των εκτιμητών ελαχίστων τετραγώνων



Παράδειγμα 3: Bootstrap στην παλινδρομηση

Εκτιμηθείσα κατανομή του $\widehat{\beta}$





Bootstrap Διαστήματα Εμπιστοσύνης για το β

Level	Normal	Basic	Percentile	BCa
95%	$(-0.0088, -0.0026)$	$(-0.0085, -0.0026)$	$(-0.0093, -0.0033)$	$(-0.0087, -0.0026)$

Bootstrap έλεγχοι υποθέσεων

Bootstrap έλεγχοι υποθέσεων

- Έστω ότι θέλουμε να ελέγξουμε μία υπόθεση της μορφής

$$H_0 : \theta = \theta_0 \quad \text{έναντι της} \quad H_1 : \theta \neq \theta_0$$

- Έστω ότι έχουμε στη διάθεση μας μία ελεγχοσυνάρτηση $T_n = T(X_1, \ldots, X_n)$
- Η κατανομή του πληθυσμού είναι άγνωστη
- Η κατανομή της T_n υπό την ισχύ της μηδενικής υπόθεσης είναι επίσης άγνωστη
- Μπορούμε να χρησιμοποιήσουμε Bootstrap για να εκτιμήσουμε το p-value του ελέγχου.
- Εναλλακτικά, μπορούμε να χρησιμοποιήσουμε και bootstrap διαστήματα εμπιστοσύνης.

Εκτίμηση p-value μέσω bootstrap

- Το p-value του ελέγχου

$$\mathrm{p-value} = P(T_n \geqslant t | H_0)$$

όπου t η παρατηρηθείσα τιμή της T.

- Έστω ότι μπορούμε να προσομοιώσουμε δείγμα από την $\widehat{F}_n$ **υπό την ισχύ της** H_0

$$(\widetilde{x}^{\star}_{b,1}, \widetilde{x}^{\star}_{b,2}, \ldots, \widetilde{x}^{\star}_{b,n})$$

(για $b = 1, \ldots, B$)

- Κατόπιν υπολογίζουμε την T_n για κάθε $b = 1, \ldots, B$:

$$T^{\star}_{n,b} := T_n(\widetilde{x}^{\star}_{b,1}, \widetilde{x}^{\star}_{b,2}, \ldots, \widetilde{x}^{\star}_{b,n})$$

- Bootstrap εκτίμηση του p-value

$$\mathrm{p-value}_{\mathrm{boot}} = \frac{1}{B} \sum_{b=1}^{B} \mathrm{I}\{T^{\star}_{n,b} \geqslant t\}$$

Παράδειγμα 1

- Έστω οι 10 παρατηρήσεις
  $-0.89, -0.47, 0.05, 0.155, 0.279, 0.775, 1.0016, 1.23, 1.89, 1.96$
- Να ελεγχθεί η υπόθεση ότι ο μέσος του πληθυσμού είναι 1 έναντι της εναλλακτικής ότι διαφέρει.
- $H_0 : \mu = 1$ έναντι της $H_1 : \mu \neq 1$
- Μία λογική ελεγχοσυνάρτηση είναι η $T = |\overline{X} - 1|$
- «Μεγάλες» τιμές της T: απόρριψη της H_0
- Η H_0 υποθέτει μέση τιμή 1
- Η παρατηρηθείσα τιμή του δειγματικού μέσου ειναι $\overline{x} = 0.598$
- Αν πάρουμε bootstrap δείγματα από την $\widehat{F}_n$ δεν προσομοιώνουμε από την H_0

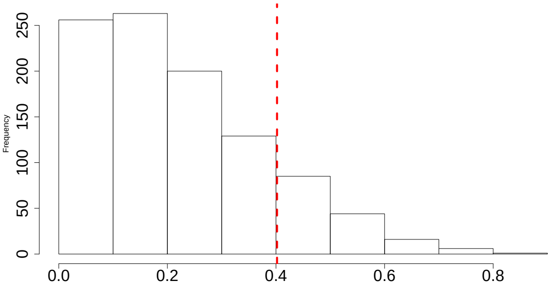
Παράδειγμα 1: μετασχηματισμός

- Λύση: Μετασχηματίζουμε τα δεδομένα ώστε να ικανοποιείται η υπόθεση ότι η μέση τιμή είναι 1
    - Σε κάθε παρατήρηση προσθέτουμε 0.402:

$$\widetilde{X}_i = X_i + 0.402$$

    - Η εμπειρική συνάρτηση κατανομής των $\widetilde{X}_i$ μπορεί να χρησιμοποιηθεί για προσομοίωση υπό την H_0
    - Ενώ διορθώνουμε τη μέση τιμή, άλλα χαρακτηριστικά όπως η διασπορά, ασυμμετρία παραμένουν ίδια.

Παράδειγμα 1: bootstrap



- Ιστόγραμμα $B = 1000$ bootstrap τιμών της $T_n = \left|\overline{\widetilde{X}}_n - 1\right|$
- Ιδιαίτερα ασύμμετρη κατανομή
- Κόκκινη γραμμή: παρατηρηθείσα τιμή της T_n: $t = 0.402$
- Bootstrap εκτίμηση του p-value

$$\mathrm{p-value_{boot}} = \frac{1}{B}\sum_{b=1}^{B} \mathrm{I}\{T^{\star}_{n,b} \geqslant t\} = 0.15$$

- Δεν απορρίπτουμε την H_0 στα συνήθη επίπεδα σημαντικότητας

Παράδειγμα 2: σύγκριση μέσων τιμών 2 ανεξάρτητων πληθυσμών

- Έστω $\boldsymbol{X} = (X_1, \ldots, X_n)$ τυχαίο δείγμα από κατανομή με μέση τιμή μ_1
- Έστω $\boldsymbol{Y} = (Y_1, \ldots, Y_m)$ τυχαίο δείγμα από κατανομή με μέση τιμή μ_2
- X_i, Y_j είναι ανεξάρτητα για κάθε i, j
- Ενδιαφέρει ο έλεγχος της

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

- Η H_0 υπαγορεύει ισότητα μέσων άλλα όχι πχ ισότητα διασπορών

Παράδειγμα 2: ελεγχοσυνάρτηση

- Μία στατιστική συνάρτηση που είναι κατάλληλη για τον παραπάνω έλεγχο είναι η Welch t-statistic

$$T(\boldsymbol{X}, \boldsymbol{Y}) = \frac{\overline{X}_n - \overline{Y}_m}{\sqrt{\frac{S_1^2}{n} + \frac{S_2^2}{m}}} \tag{12}$$

- $\overline{X}_n = \frac{1}{n} \sum_{i=1}^{n} X_i$
- $\overline{Y}_m = \frac{1}{m} \sum_{i=1}^{m} Y_i$
- $S_1^2 = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \overline{X}_n)^2$
- $S_2^2 = \frac{1}{m-1} \sum_{i=1}^{m} (Y_i - \overline{Y}_m)^2$
- Υπενθύμιση: **Αν** τα δείγματα είναι κανονικά (και ανεξάρτητα) η κατανομή της $T(\boldsymbol{X}, \boldsymbol{Y})$ υπό την H_0 προσεγγίζεται από μία κατανομή t.
- Μέσω bootstrap δεν χρειάζεται να υποθέσουμε κανονικότητα.

Παράδειγμα 2: μετασχηματισμός

- Για να εφαρμόσουμε bootstrap πρέπει να μετασχηματίσουμε τα δεδομένα έτσι ώστε να ικανοποιείται η H_0 (ίδια μέση τιμή)
- Έστω

$$\widetilde{X}_i = X_i - \overline{X}_n + \overline{Z}, \quad i = 1, \ldots, n$$
$$\widetilde{Y}_j = Y_j - \overline{Y}_M + \overline{Z}, \quad j = 1, \ldots, m$$
$$\text{όπου} \quad \overline{Z} = \frac{1}{n+m}\left(\sum_{i=1}^{n} X_i + \sum_{j=1}^{m} Y_j\right)$$

- Προφανώς:

$$\mathrm{E}\widetilde{X}_i = \mathrm{E}\widetilde{Y}_j = \frac{n}{n+m}\mu_1 + \frac{m}{n+m}\mu_2$$

για κάθε $i = 1, \ldots, n$ και $j = 1, \ldots, m$

Παράδειγμα 2: αλγόριθμος bootstrap

(1) Θέσε $\widetilde{x}_i = x_i - \overline{x}_n + \overline{z}$ και $\widetilde{y}_j = y_j - \overline{y}_m + \overline{z}$ για $i = 1, \ldots, n$, $j = 1, \ldots, m$

(2) Για $b = 1, \ldots, B$

(2.1) Πάρε δείγμα

$$\widetilde{\boldsymbol{x}}_b^\star = (\widetilde{x}_{b,1}^\star, \ldots, \widetilde{x}_{b,n}^\star)$$

με επανάθεση από τα $(\widetilde{x}_1, \ldots, \widetilde{x}_n)$

(2.2) Πάρε δείγμα

$$\widetilde{\boldsymbol{y}}_b^\star = (\widetilde{y}_{b,1}^\star, \ldots, \widetilde{y}_{b,m}^\star)$$

με επανάθεση από τα $(\widetilde{y}_1, \ldots, \widetilde{y}_m)$

(2.3) Υπολόγισε την τιμή της ελεγχοσυνάρτησης Welch (12)

$$T_{n,b}^\star := T(\widetilde{\boldsymbol{x}}_b^\star, \widetilde{\boldsymbol{y}}_b^\star)$$

(3) Υπολόγισε την bootstrap εκτίμηση του p-value

$$\mathrm{p-value}_{\mathrm{boot}} = \frac{1}{B} \sum_{b=1}^{B} \mathrm{I}\{T_{n,b}^\star \geqslant t\}$$

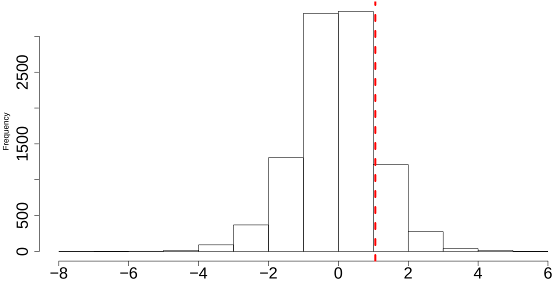
Παράδειγμα 2: Εφαρμογή

- 16 ποντίκια συμμετείχαν σε ένα πείραμα όπου σε 7 από αυτά δόθηκε ένα καινούργιο φάρμακο και στα υπόλοιπα δόθηκε ένα ήδη υπάρχον φάρμακο.
- Μελέτη επιβίωσης σε μέρες:

Φάρμακο	Μέρες επιβίωσης
Νέο	$94, 197, 16, 38, 99, 141, 23$
Παλιό	$52, 104, 146, 10, 51, 30, 40, 27, 46$

- Μεγαλώνει το νέο φάρμακο την επιβίωση;
- $H_0 : \mu_1 = \mu_2$ έναντι της $H_1 : \mu_1 > \mu_2$
- $n = 7$ και $m = 9$
- $\overline{x}_n = 86.86$, $\overline{y}_m = 56.22$
- $S_1^2 = 4457.81$, $S_2^2 = 1804.194$
- $T(\boldsymbol{x}, \boldsymbol{y}) = 1.058$

Παράδειγμα 2: bootstrap



- Ιστόγραμμα $B = 10000$ bootstrap τιμών της $T_n = \dfrac{\overline{\widetilde{X}}_n - \overline{\widetilde{Y}}_m}{\sqrt{\dfrac{\widetilde{S}_1^2}{n} + \dfrac{\widetilde{S}_2^2}{m}}}$
- Κόκκινη γραμμή: παρατηρηθείσα τιμή της T_n: $t = 1.058$

Μέθοδος	bootstrap	t-test	Mann-Whitney
p-value	0.146	0.158	0.340

- Δεν απορρίπτουμε την H_0 στα συνήθη επίπεδα σημαντικότητας

Pivotal ελεγχοσυναρτήσεις

- Η επιλογή της ελεγχοσυνάρτησης για τη διεξαγωγή ενός bootstrap ελέγχου παίζει ρόλο
- Ιδανικά, θα πρέπει η κατανομή της ελεγχοσυνάρτησης (που μας είναι άγνωστη) να μην εξαρτάται από άγνωστες παραμέτρους
- Γενικά, μέσω bootstrap δεν είναι εύκολο να εξακριβώσουμε αν η ελεγχοσυνάρτηση που χρησιμοποιούμε έχει τη συγκεκριμένη ιδιότητα
- Μία ελεγχοσυνάρτηση που έχει την ιδιότητα ότι η κατανομή της (υπό την H_0) δεν εξαρτάται από άγνωστες παραμέτρους ονομάζεται ποσότητα οδηγός (Pivotal)

Pivotal ελεγχοσυναρτήσεις

- Η επιλογή της ελεγχοσυνάρτησης για τη διεξαγωγή ενός bootstrap ελέγχου παίζει ρόλο
- Ιδανικά, θα πρέπει η κατανομή της ελεγχοσυνάρτησης (που μας είναι άγνωστη) να μην εξαρτάται από άγνωστες παραμέτρους
- Γενικά, μέσω bootstrap δεν είναι εύκολο να εξακριβώσουμε αν η ελεγχοσυνάρτηση που χρησιμοποιούμε έχει τη συγκεκριμένη ιδιότητα
- Μία ελεγχοσυνάρτηση που έχει την ιδιότητα ότι η κατανομή της (υπό την H_0) δεν εξαρτάται από άγνωστες παραμέτρους ονομάζεται ποσότητα οδηγός (Pivotal)
- Παράδειγμα: έλεγχος για τη μέση τιμή πληθυσμού $H_0 : \mu = \mu_0$
    - Υπό την H_0, η ασυμπτωτική κατανομή του $T_1 = \overline{X}_n - \mu_0$ είναι η $\mathcal{N}(0, \sigma^2/n)$
    - Αντίθετα η ασυμπτωτική κατανομή της $T_2 = \sqrt{n}\frac{\overline{X}_n - \mu_0}{S_n}$ είναι η $\mathcal{N}(0, 1)$
    - Η T_2 είναι pivotal ελεγχοσυνάρτηση, σε αντίθεση με την T_1.

Πότε το bootstrap αποτυγχάνει

- Μικρά σύνολα δεδομένων (η $\widehat{F}_n$ δεν είναι καλή προσέγγιση της F)
- Προβλήματα μεγάλων διαστάσεων: καθώς η διάσταση των μεταβλητών αυξάνει, η $\widehat{F}_n$ δεν αποτελεί καλή προσέγγιση της F για πεπερασμένο n.
- Άπειρες ροπές (π.χ: κατανομή t_1)
- Εξαρτημένες παρατηρήσεις (π.χ: χρονολογικές σειρές, χωρικά προβλήματα)
- Εκτίμηση ακραίων τιμών (π.χ: εκτίμηση του 99.99% ποσοστιαίου σημείου, ή το $\max X_i$).
- Πιο γενικά, σε όλες τις περιπτώσεις που το συναρτησιακό δεν είναι «ομαλό». Υπάρχει πληθώρα θεωρητικών αποτελεσμάτων που συνδέουν τη συμπεριφορά του bootstrap με το πόσο ομαλό θεωρείται ένα συναρτησιακό (π.χ: Hadamard derivative)

Βιβλιογραφία

Ιωαννίδης, Ε.
Μη παραμετρική Στατιστική.
Πανεπιστημιακές σημειώσεις ΟΠΑ

Καρλής, Δ.
Υπολογιστική Στατιστική.
Πανεπιστημιακές σημειώσεις ΟΠΑ, 2004

Davison, A.C. and Hinkley, D.V.
Bootstrap methods and their application.
Cambridge University Press, 1997

Wasserman, L.
All of Non-Parametric Statistics.
Springer, 2006