Υπολογιστικές μέθοδοι στην εκτίμηση παραμέτρων, στατιστικών ελέγχων

Δημήτρης Κουγιουμτζής

23 Οκτωβρίου 2019

Παραμετρική εκτίμηση και έλεγχος

Για την (παραμετρική) εκτίμηση παραμέτρων

1. Επιλέγουμε ένα στατιστικό, π.χ. $\bar{x}$ για την παράμετρο μ.

2. Επιλέγουμε κανονικοποίηση του στατιστικού ώστε να ακολουθεί κάποια γνωστή κατανομή, π.χ. $t = \frac{\bar{x}-\mu}{s/\sqrt{n}} \sim t_{n-1}$.

3. Με βάση τη γνωστή κατανομή υπολογίζουμε την κρίσιμη τιμή, π.χ. $t_{n-1,1-\alpha/2}$ για επίπεδο εμπιστοσύνης $(1-\alpha)\%$.

4. Υπολογίζουμε το διάστημε εμπιστοσύνης της παραμέτρου, π.χ. $(1-\alpha)\%$ δ.ε. για μ: $\bar{x} \pm t_{n-1,1-\alpha/2}\frac{s}{\sqrt{n}}$

Για τον (παραμετρικό) έλεγχο υπόθεσης:

1. Ορίζουμε την τιμή της παραμέτρου για έλεγχο, π.χ. H_0: $\mu = \mu_0$.

2. Επιλέγουμε ένα στατιστικό ..., π.χ. $\bar{x}$ για μ.

3. Επιλέγουμε κανονικοποίηση ..., π.χ. $t = \frac{\bar{x}-\mu}{s/\sqrt{n}} \sim t_{n-1}$.

4. Υπολογίζουμε το κανονικοποιημένο στατιστικό στο δείγμα, $\tilde{t} = \frac{\bar{x}-\mu_0}{s/\sqrt{n}}$

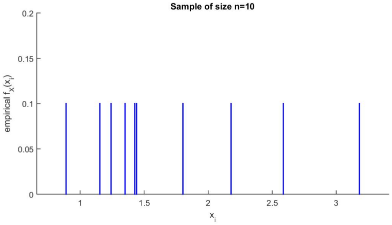
5. Κρίσιμη τιμή της κατανομής ..., π.χ. $t_{n-1,1-\alpha/2}$

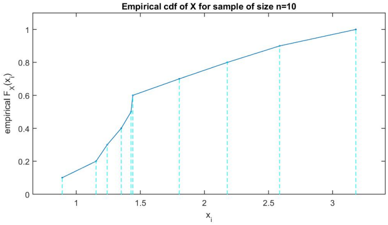
6. Ορίζουμε την απορριπτική περιοχή και αποφασίζουμε την απόρριψη ή όχι της H_0, π.χ. απόρριψη της H_0 αν $|\tilde{t}| > t_{n-1,1-\alpha/2}$.

Αν δε γνωρίζουμε την κατανομή του στατιστικού; π.χ. $\frac{\bar{x}-\mu}{s/\sqrt{n}} \sim ???$

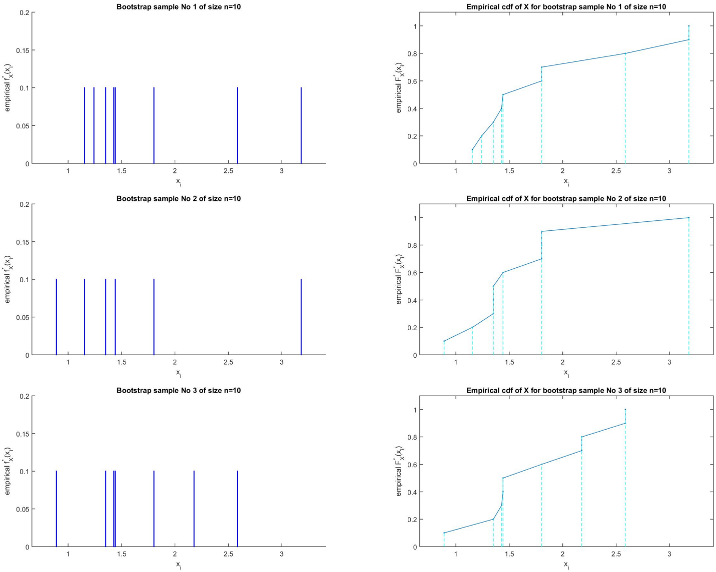
Οι **μέθοδοι επαναδειγματοληψίας** είναι υπολογιστικές μέθοδοι που μας επιτρέπουν να προσεγγίσουμε την κατανομή οποιουδήποτε στατιστικού με βάση μόνο τις παρατηρήσεις του δείγματος.

Θεωρούμε το δείγμα $\mathbf{x} = \{x_1, \dots, x_n\}$ της τ.μ. X με άγνωστη κατανομή με ασκ $F_X(x)$.
Θέλουμε να εκτιμήσουμε την παράμετρο θ της κατανομής F και έχουμε ένα στατιστικό $\hat{\theta} = g(\mathbf{x})$, π.χ. $\bar{x} = \sum_{i=1}^{n} x_i$.
Η εμπειρική κατανομή με ασκ $\hat{F}_X(x)$ δίνεται απλά με το να ορίσουμε πιθανότητα $1/n$ σε κάθε παρατήρηση x_i, $x_i \in \mathbf{x}$.





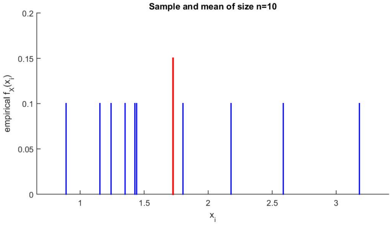
Το **δείγμα bootstrap $\mathbf{x}^*$** είναι ένα δείγμα με n παρατηρήσεις, όπου η κάθε παρατήρηση επιλέγεται τυχαία και με επανάθεση από τις παρατηρήσεις του $\mathbf{x} = \{x_1, \ldots, x_n\}$.

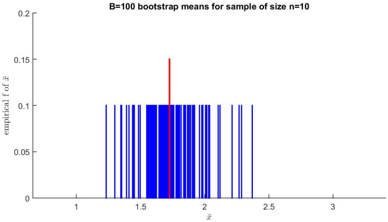


Η **επανάληψη bootstrap του** $\hat{\theta}$ είναι $\hat{\theta}^* = g(\mathbf{x}^*)$, π.χ. $\bar{x}^* = \sum_{i=1}^{n} x_i^*$.
Για B δείγματα bootstrap έχουμε B επαναλήψεις bootstrap του στατιστικού $\hat{\theta}$

$$\{\hat{\theta}^{*1}, \hat{\theta}^{*2}, \ldots, \hat{\theta}^{*B}\}$$

Παράδειγμα: $\theta = \mu$ και $\hat{\theta} = \bar{x}$





Bootstrap εκτίμηση του τυπικού σφάλματος εκτιμητή

Το τυπικό σφάλμα (standard error) $\mathrm{se}(\hat{\theta}) = \sigma_{\hat{\theta}}$ είναι η τυπική απόκλιση του εκτιμητή $\hat{\theta}$

Για $\theta = \mu$, $\mathrm{se}(\bar{x}) = \sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$ και η εκτίμηση του $\widehat{\mathrm{se}}(\bar{x}) = s_{\bar{x}} = \frac{s}{\sqrt{n}}$

Γενικά το $\mathrm{se}(\hat{\theta}) = \sigma_{\hat{\theta}}$ δεν είναι γνωστό.

... μπορούμε όμως να το εκτιμήσουμε με τη μέθοδο bootstrap.

Bootstrap εκτίμηση $\widehat{\mathrm{se}}(\hat{\theta})$

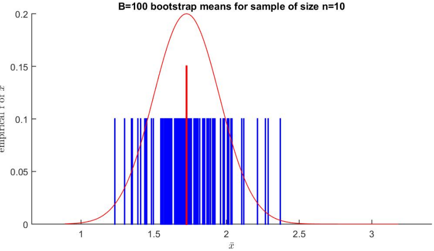
1. Επιλέγουμε B δείγματα bootstrap $\mathbf{x}^{*1}, \dots, \mathbf{x}^{*B}$.
2. Υπολογίζουμε B επαναλήψεις bootstrap του στατιστικού $\hat{\theta}$
$$\{\hat{\theta}^{*1}, \hat{\theta}^{*2}, \dots, \hat{\theta}^{*B}\}$$
3. Εκτιμούμε το τυπικό σφάλμα $\mathrm{se}(\hat{\theta})$ από την τυπική απόκλιση του $\hat{\theta}$ στις B επαναλήψεις
$$\widehat{\mathrm{se}}_B(\hat{\theta}) = \sqrt{\frac{1}{B-1} \sum_{b=1}^{B} \left(\hat{\theta}^{*b} - \bar{\theta}^{*}\right)^2}$$
όπου $\bar{\theta}^{*} = \frac{1}{B} \sum_{b=1}^{B} \hat{\theta}^{*b}$

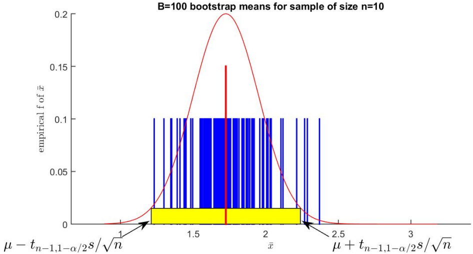
Εκτίμηση διαστήματος εμπιστοσύνης με ποσοστιαία bootstrap

Για B δείγματα bootstrap έχουμε B επαναλήψεις bootstrap του στατιστικού $\hat{\theta}$

$$\{\hat{\theta}^{*1}, \hat{\theta}^{*2}, \ldots, \hat{\theta}^{*B}\}$$

Η παραμετρική εκτίμηση διαστήματος εμπιστοσύνης υποθέτει γνωστή κατανομή για το στατιστικό $\hat{\theta}$, π.χ. κατανομή student





$(1-\alpha)\%$ δ.ε. για μ: $\bar{x} \pm t_{n-1,1-\alpha/2} \frac{s}{\sqrt{n}}$

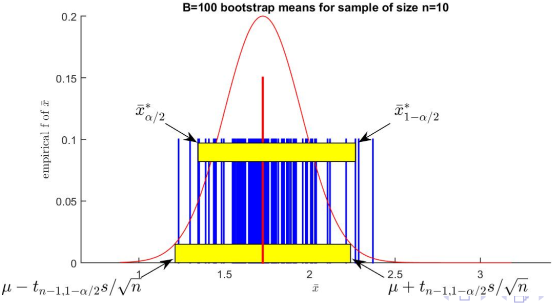
δ.ε. για μ από B επαναλήψεις bootstrap του στατιστικού $\hat{\theta}$ $\boxed{?}$

Για B δείγματα bootstrap έχουμε B επαναλήψεις bootstrap του στατιστικού $\hat{\theta}$, $\{\hat{\theta}^{*1}, \hat{\theta}^{*2}, \dots, \hat{\theta}^{*B}\}$

Οι τιμές $\hat{\theta}$, $\{\hat{\theta}^{*1}, \hat{\theta}^{*2}, \dots, \hat{\theta}^{*B}\}$ σχηματίζουν την εμπειρική κατανομή του $\hat{\theta}$.

Οι ουρές της κατανομής δίνονται από τα $\alpha/2$ και $1-\alpha/2$ ποσοστιαία σημεία $\hat{\theta}^*_{\alpha/2}$ και $\hat{\theta}^*_{1-\alpha/2}$ του δείγματος των $\{\hat{\theta}^{*1}, \hat{\theta}^{*2}, \dots, \hat{\theta}^{*B}\}$

Το διάστημα $[\hat{\theta}^*_{\alpha/2}, \hat{\theta}^*_{1-\alpha/2}]$ είναι το $(1-\alpha)\%$ δ.ε. για θ με τη μέθοδο των ποσοστιαίων bootstrap.



Ορισμός ποσοστιαίων σημείων από τα $\{\hat{\theta}^{*1}, \hat{\theta}^{*2}, \dots, \hat{\theta}^{*B}\}$

- Το $\alpha/2$% ποσοστιαίο σημείο είναι το σημείο στη θέση $k = [(B+1)\alpha/2]$ (ακέραιο μέρος) (θεωρώντας αύξουσα σειρά για τα $\hat{\theta}^{*i}$).
- Το $(1-\alpha/2)$% ποσοστιαίο σημείο είναι το $B+1-k$ σημείο.

Για $\theta = \mu$ το παραμετρικό δ.ε. $\bar{x} \pm t_{n-1,1-\alpha/2}\frac{s}{\sqrt{n}}$ και το δ.ε. ποσοστιαίων bootstrap $[\hat{\theta}^*_{\alpha/2}, \hat{\theta}^*_{1-\alpha/2}]$ γενικά συμφωνούν και συγκλίνουν για $n \to \infty$. Για άλλα στατιστικά όχι απαραίτητα. Ποιο από τα δύο θα επιλέξουμε;

Υπάρχουν και άλλα δ.ε. με bootstrap:

- Studentized bootstrap ή bootstrap-t διορθώνει τις κρίσιμες τιμές για $\alpha/2$ και $1-\alpha/2$ στο παραμετρικό δ.ε. με αυτές από επαναλήψεις bootstrap.
- Bias corrected and accelerated (BCa) bootstrap που διορθώνει τη μεροληψία και λοξότητα στην κατανομή bootstrap του στατιστικού.
- Άλλα ...

Δ.ε. διαφοράς μέσων τιμών, παραμετρικό και bootstrap

Έστω ότι έχουμε δείγμα $\{x_1, \ldots, x_n\}$ της τ.μ. X και $\{y_1, \ldots, y_m\}$ της τ.μ. Y.

Θέλουμε να ελέγξουμε αν οι μέσες τιμές των X και Y είναι ίσες.

Παραμετρικό $(1-\alpha)\%$ δ.ε. για $\mu_X - \mu_Y$

Εκτίμηση κοινής (pooled) διασποράς: $s_p^2 = \frac{(n-1)s_X^2+(m-1)s_Y^2}{n+m-2}$

Στατιστικό παραμετρικού δ.ε.: $\hat{\theta} \equiv t \equiv \frac{(\bar{x}-\bar{y})-(\mu_X-\mu_Y)}{s_p\sqrt{\frac{1}{n}+\frac{1}{m}}} \sim t_{n+m-2}$

$$(1-\alpha)\% \text{ δ.ε.: } (\bar{x}-\bar{y}) \pm t_{n+m-2,1-\alpha/2}\, s_p\sqrt{\frac{1}{n}+\frac{1}{m}}$$

Bootstrap $(1-\alpha)\%$ δ.ε. για $\mu_X - \mu_Y$

1. Επιλέγουμε B δείγματα bootstrap για τη X: $\{\mathbf{x}^{*1}, \ldots, \mathbf{x}^{*B}\}$ και για τη Y: $\{\mathbf{y}^{*1}, \ldots, \mathbf{y}^{*B}\}$
2. Υπολογίζουμε B επαναλήψεις bootstrap του στατιστικού $\hat{\theta} \equiv \bar{x} - \bar{y}$, $\{\bar{x}^{*1} - \bar{y}^{*1}, \ldots, \bar{x}^{*B} - \bar{y}^{*B}\}$
3. Τα ποσοστιαία σημεία $(\bar{x}^* - \bar{y}^*)_{\alpha/2}$ και $(\bar{x}^* - \bar{y}^*)_{1-\alpha/2}$ ορίζουν το $(1-\alpha)\%$ δ.ε. για $\mu_X - \mu_Y$ με τη μέθοδο των ποσοστιαίων bootstrap.

Έλεγχος ισότητας μέσων τιμών: παραμετρικός, bootstrap και τυχαιοποίησης

Έστω ότι έχουμε δείγμα $\{x_1, \ldots, x_n\}$ της τ.μ. X και $\{y_1, \ldots, y_m\}$ της τ.μ. Y.
Θέλουμε να ελέγξουμε αν οι μέσες τιμές των X και Y είναι ίσες.

Παραμετρικός έλεγχος υπόθεσης $\mathrm{H}_0 : \mu_X - \mu_Y$

Το κατάλληλο στατιστικό προκύπτει από την κανονικοποίηση του εκτιμητή της παραμέτρου $\theta = \mu_X - \mu_Y$:
$\hat{\theta} \equiv t \equiv \frac{(\bar{x}-\bar{y})-(\mu_X-\mu_Y)}{s_p\sqrt{\frac{1}{n}+\frac{1}{m}}} \sim \mathrm{t}_{n+m-2}$

Για την H_0 το στατιστικό είναι $t \equiv \frac{\bar{x}-\bar{y}}{s_p\sqrt{\frac{1}{n}+\frac{1}{m}}} \sim \mathrm{t}_{n+m-2}$

Η τιμή του στατιστικού στο δείγμα: $\tilde{t} \equiv \frac{\bar{x}-\bar{y}}{s_p\sqrt{\frac{1}{n}+\frac{1}{m}}} \sim \mathrm{t}_{n+m-2}$, με αντικατάσταση των τιμών $\bar{x}$, $\bar{y}$, s_p, n και m από το δείγμα των X και Y.

Απορριπτική περιοχή: $R = \{|\tilde{t}| > t_{n+m-2,1-\alpha/2}\}$

Bootstrap έλεγχος υπόθεσης $H_0 : \mu_X - \mu_Y$

1. Επιλέγουμε (με επανάθεση) B δείγματα μεγέθους $n+m$ από το κοινό δείγμα $\{x_1, \ldots, x_n, y_1, \ldots, y_n\}$. Οι πρώτες n τιμές είναι για X και οι υπόλοιπες m για Y:
$$\{[\mathbf{x}\ \mathbf{y}]^{*1}, \ldots, [\mathbf{x}\ \mathbf{y}]^{*B}\}$$
2. Υπολογίζουμε B επαναλήψεις bootstrap του στατιστικού $\hat{\theta} \equiv \bar{x} - \bar{y}$
$$\{(\bar{x} - \bar{y})^{*1}, \ldots, (\bar{x} - \bar{y})^{*B}\}$$
3. Βρίσκουμε τη θέση (rank) r του στατιστικού $\bar{x} - \bar{y}$ του αρχικού δείγματος στη λίστα των $B+1$ τιμών του στατιστικού
$$\{\bar{x} - \bar{y}, (\bar{x} - \bar{y})^{*1}, \ldots, (\bar{x} - \bar{y})^{*B}\}$$
όταν μπαίνουν σε αύξουσα σειρά.
   - Αν $r < (B+1)\alpha/2$ ή $r > (B+1)(1-\alpha/2)$, απορρίπτεται η H_0 (τα μ_X και μ_Y διαφέρουν).
   - Αν $(B+1)\alpha/2 \leq r \leq (B+1)(1-\alpha/2)$, δεν απορρίπτεται η H_0 (τα μ_X και μ_Y δε διαφέρουν).

Σύμφωνα με την $H_0 : \mu_X - \mu_Y$, θεωρούμε πως οι X και Y έχουν την ίδια κατανομή.

Έλεγχος τυχαίας αντιμετάθεσης για $H_0 : \mu_X - \mu_Y$

1. Επιλέγουμε (χωρίς επανάθεση) B δείγματα μεγέθους $n+m$ από τυχαία αντιμετάθεση των τιμών στο κοινό δείγμα $\{x_1, \ldots, x_n, y_1, \ldots, y_n\}$. Οι πρώτες n τιμές είναι για X και οι υπόλοιπες m για Y:
$$\{[\mathbf{x}\ \mathbf{y}]^{*1}, \ldots, [\mathbf{x}\ \mathbf{y}]^{*B}\}$$
2. Υπολογίζουμε B επαναλήψεις τυχαίας αντιμετάθεσης του στατιστικού $\hat{\theta} \equiv \bar{x} - \bar{y}$
$$\{(\bar{x} - \bar{y})^{*1}, \ldots, (\bar{x} - \bar{y})^{*B}\}$$
3. Βρίσκουμε τη θέση (rank) r του στατιστικού $\bar{x} - \bar{y}$ του αρχικού δείγματος στη λίστα των $B+1$ τιμών του στατιστικού
$$\{\bar{x} - \bar{y}, (\bar{x} - \bar{y})^{*1}, \ldots, (\bar{x} - \bar{y})^{*B}\}$$
όταν μπαίνουν σε αύξουσα σειρά.
   - Αν $r < (B+1)\alpha/2$ ή $r > (B+1)(1-\alpha/2)$, απορρίπτεται η H_0 (τα μ_X και μ_Y διαφέρουν).
   - Αν $(B+1)\alpha/2 \leq r \leq (B+1)(1-\alpha/2)$, δεν απορρίπτεται η H_0 (τα μ_X και μ_Y δε διαφέρουν).