Διάλεξη 5

Έστω $X_1, \dots, X_n$ τ.δ. από $\mathcal{N}(\mu, \sigma^2)$ όπου μ και σ^2 είναι άγνωστα.

Θα βρούμε την ε.μ.π. του $\theta = (\mu, \sigma^2)$.

$$L(\theta) = \prod_{i=1}^{n} f(x_i; \theta) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(x_i-\mu)^2}$$

$$= (2\pi\sigma^2)^{-\frac{n}{2}} e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2} \quad \left(\begin{array}{l}\text{συνάρτησης}\\ \text{πιθανοφάνειας}\end{array}\right).$$

$$\ell(\theta) = \log L(\theta) = -\frac{n}{2}\log(2\pi) - \frac{n}{2}\log\sigma^2 - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2$$

Διαδοχική Μεγιστοποίηση

Αν πάρουμε το σ^2 σταθερό $\Rightarrow$ ανάγεται στην εύρεση ε.μ.π. με το σ^2 γνωστό.

Η μεγιστοποίηση ως προς μ, ανάγεται σε ελαχιστοποίηση του $\sum_{i=1}^{n}(x_i-\mu)^2$.

$$f(\mu) = \sum_{i=1}^{n} (x_i - \mu)^2 = \sum_{i=1}^{n} (\mu - x_i)^2$$

$$f'(\mu) = 2 \sum_{i=1}^{n} (\mu - x_i) = 2\left(n\mu - \sum_{i=1}^{n} x_i\right)$$

$$f'(\mu) = 0 \iff \boxed{\mu^* = \bar{x}} \rightarrow$$ μοναδικό στάσιμο σημείο

• $f''(\mu) = 2n > 0$ (γνήσια κυρτή).

Ειδικά $f''(\mu^*) > 0 \Rightarrow$ Τοπικό ελάχιστο

$\Rightarrow$ λόγω μοναδ. στάσιμο σημείου έχουμε ολικό ελάχιστο $\Rightarrow \mu^* = \bar{x}$ είναι ολικό μέγιστο $\Longrightarrow$ + ανεξάρτητο του σ^2

$\Rightarrow \hat{\mu} = \bar{x}$ (βρήκαμε κατ' ευθείαν την ε.μ.π. του μ)

Τελικό το μέγιστο ως προς σ^2, επιτυγχάνεται με αντικατάσταση του $\mu = \bar{x}$ στην αρχική συνάρτηση και έχουμε μία μεγιστοποίηση μιας συνάρτησης μόνο ως προς σ^2 (μία παράμετρο)

$$\Rightarrow \max_{\mu,\sigma^2} \ell(\mu,\sigma^2) = \max_{\sigma^2} \max_{\mu} \ell(\mu,\sigma^2)$$

$$= \boxed{\max_{\sigma^2} \ell(\bar{x},\sigma^2)}$$

$$\ell(\bar{x},\sigma^2) = -\frac{n}{2}\log(2\pi) - \frac{n}{2}\log\sigma^2 - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\bar{x})^2.$$

Θέτουμε $\varphi = \sigma^2$

$$g(\varphi) = -\frac{n}{2}\log\varphi - \frac{1}{2\varphi}\sum_{i=1}^{n}(x_i - \bar{x})^2.$$

$$g'(\varphi) = -\frac{n}{2\varphi} + \frac{1}{2\varphi^2}\sum_{i=1}^{n}(x_i - \bar{x})^2$$

$$g'(\varphi) = 0 \iff \frac{n}{\cancel{2}\cancel{\varphi}} = \frac{1}{\cancel{2}\varphi^{\cancel{2}}}\sum(x_i - \bar{x})^2$$

$$\iff \varphi^* = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n}$$

$g''(\varphi^*) < 0$ (τοπικό μέγιστο + μον. στασ. σημ.)

$\Downarrow$

ολικό μέγιστο.

Τελικά $\hat{\theta} = (\overline{X}, M_2)$

$\hookrightarrow \frac{1}{n}\sum_{n=1}^{n}(X_i - \overline{X})^2$

(κεντρική δειγματική ροπή $2^{ης}$ τάξης)

ή δειγματική διασπορά

Παρατήρηση

- Το $1^{ο}$ πρόβλημα είναι η εύρεση της ε.μ.π. του μ όταν το σ^2 είναι γνωστό
- Το $2^{ο}$ πρόβλημα είναι η εύρεση της ε.μ.π. του σ^2 όταν το μ είναι γνωστό ($\mu = \overline{X}_n$),

$\hookrightarrow \hat{\sigma}^2 = \sum_{i=1}^{n}(x_i - \mu)^2 / n \xrightarrow{\mu = \overline{X}} \frac{\sum_{i=1}^{n}(X_i - \overline{X})^2}{n}$

Παρατηρήσεις για την ε.μ.π.

- Η ε.μ.π. μας δίνει πάντα εκτιμήτρια δηλ. σ.σ. με τιμές στον παραμετρικό χώρο. Σε αντίθεση με τη μέθοδο των ροπών.
- Η ε.μ.π. είναι πολύ δημοφιλής μέθοδος εύρεσης εκτιμητριών, διότι έχει καλές ασυμπτωτικές ιδιότητες. Για μικρό δείγμα δε συμπεριφέρεται κατ' ανάγκη καλά.

Σε κάποιο πρόβλημα ενδεχομένως με διαφορετικές μεθόδους να πάρουμε διαφορετικές εκτιμήτριες, π.χ. να διαφέρει η εκτιμήτρια ροπών από την ε.μ.π. και τίθεται ένα θέμα σύγκρισης εκτιμητριών.

Ερώτημα

Ποιές είναι οι ιδιότητες εκείνες που χαρακτηρίζουν μια καλή εκτιμήτρια ?

Σφάλμα Εκτιμήτρια

- Σφάλμα : $\hat{\theta} - \theta$
- Απόλυτο Σφάλμα : $|\hat{\theta} - \theta|$
- Τετραγωνικό Σφάλμα : $(\hat{\theta} - \theta)^2$

$\Downarrow$

bias

- Μέσο Σφάλμα (μεροληψία) $E_\theta(\hat{\theta}) - \theta =: b(\theta)$
- Μέσο Απόλυτο Σφάλμα : $E_\theta|\hat{\theta} - \theta|$ (ΜΑΣ)
- Μέσο Τετραγωνικό Σφάλμα : $E_\theta(\hat{\theta} - \theta)^2$ (ΜΤΣ)

MSE (Mean Square Error)

<u>Ορσ</u>: Μια εκτιμήτρια λέγεται <u>αμερόληπτη εκτιμήτρια</u> του θ (α.ε.) αν $b(\theta) = 0$ ή $E_\theta(\hat{\theta}) = \theta, \ \forall \theta \in \Theta$.

<u>Πρόταση</u> (τρόπος εύρεσης του ΜΤΣ)

$$\boxed{MT\Sigma_{\hat{\theta}}(\theta) = b^2_{\hat{\theta}}(\theta) + V_\theta(\hat{\theta})}$$

$$\| \quad E(\hat{\theta} - \theta)^2$$

↳ προσθαφαίρεση της μέσης τιμής

• Γενικά αν μπορούμε επιλέγουμε αμερόλ. εκτιμήτριες με το μικρότερο δυνατό ΜΤΣ.

$\boxed{\text{Παράδειγμα}}$ (Άσκηση).

Έστω $X_1, X_2, \ldots, X_n$ τ.δ. από $\underline{\mathcal{U}[0,\theta]}$

Ν.δ.ο. $\overline{\theta} = 2\overline{X}$ (ε.ρ.).

$\hat{\theta} = X_{(n)} = \max\{X_1, \ldots, X_n\}$. (ε.μ.π.)

για ε.μ.π.

Έστω $x=(x_1,\dots,x_n)$, με $x_i>0$.

$$L(\theta)=\prod_{i=1}^{n} f(x_i;\theta)=\prod_{i=1}^{n}\frac{1}{\theta}\cdot \mathbb{1}_{[0,\theta]}(x_i)$$

$$=\frac{1}{\theta^n}\cdot\prod_{i=1}^{n}\mathbb{1}_{[0,\theta]}(x_i)$$

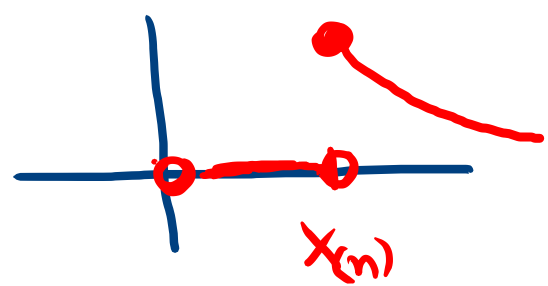
$$\prod_{i=1}^{n}\mathbb{1}_{[0,\theta]}(x_i)=1 \iff 0\le x_i\le\theta,\ 1\le i\le n$$

$$\iff x_i\le\theta,\ \forall\, 1\le i\le n$$

$$\iff X_{(n)}\le\theta$$

$\theta>0$

$$L(\theta)=\frac{1}{\theta^n}\mathbb{1}_{[X_{(n)},+\infty)}(\theta)$$



$$\Rightarrow \boxed{\hat{\theta} = X_{(n)}} \longrightarrow \text{κατανομή ?}$$

$$X_i \sim \underline{\mathcal{U}[0,\theta]} = \theta \cdot \mathcal{U}[0,1] \quad (\theta \cdot U_i)$$

$$\Rightarrow X_{(n)} = \max\{X_1, \ldots, X_n\} = \max\{\theta \cdot U_1, \ldots, \theta \cdot U_n\}$$

$$= \theta \cdot \max\{U_1, \ldots, U_n\} = \theta \cdot U_{(n)}$$

$$\Rightarrow X_{(n)} = \theta \cdot \boxed{U_{(n)}} \qquad \boxed{0 < u \leq 1}$$

$$\Big[\; P\left(U_{(n)} \leq u\right) = P\left(U_1 \leq u, \ldots, U_n \leq u\right) = \prod_{i=1}^{n} \underbrace{P(U_i \leq u)}_{F_U} = F_U^{n}(u)$$

$$F_U^{n}(u) = u^n$$

13

Βρείτε $E(U_{(n)})$, $V(U_{(n)})$.

$$\boxed{U_{(n)} \sim Beta(n, 1)}$$

$X \sim Beta(a, b)$, $a, b > 0$.

$$f_X(x) = \frac{1}{B(a,b)} \underbrace{x^{a-1}(1-x)^{b-1}}, \quad 0 < x < 1$$

$$B(a,b) = \int_0^1 x^{a-1}(1-x)^{b-1}\,dx$$

$E(U_{(n)}) = \frac{n}{n+1}$ διότι $E(Beta(a,b)) = \frac{a}{a+b}$.

(14)

$$E(X_{(n)}) = E(\theta \cdot U_{(n)}) =$$

$$\theta \cdot E(U_{(n)}) = \frac{n}{n+1}\theta < \theta, \ \forall \theta > 0$$

$\Rightarrow$ η $X_{(n)}$ είναι αρνητικά μεροληπτική

δηλ. κατά μέσο όρο υποεκτιμά το θ.

$\Rightarrow$ η $\tilde{\theta}_n = \frac{n+1}{n} X_{(n)} \Rightarrow E(\tilde{\theta}_n) = \theta, \ \forall \theta > 0$

και άρα είναι α.ε. του θ.

[να υπολογίσετε το ΜΤΣ όλων των εκτιμητριών $\bar{\theta}_n, \hat{\theta}_n, \tilde{\theta}_n$ (να συγκριθούν)]

Επιπλέον Ιδιότητες Εκτιμητριών

Ασυμπτωτική Συμπεριφορά :

π.χ. ασυμπτωτικά αμερόληπτη εκτιμήτρια :

$$\lim_{n\to\infty} b_{\hat\theta_n}(\theta) = 0$$

π.χ. $b_{X_{(n)}}(\theta) = E(X_{(n)}) - \theta = \frac{n}{n+1}\theta - \theta$

$= -\frac{\theta}{n+1} \xrightarrow[n\to\infty]{} 0 \Rightarrow$ είναι α.α.ε.

- συνέπεια εκτιμήτριας.

Μια εκτιμήτρια $\hat\theta_n$ του θ είναι συνεπής

αν $\hat\theta_n \xrightarrow{P} \theta$, $\forall \theta \in \Theta$.

16

Π.χ. ο δειγματικός μέσος για τη μέση τιμή.

- ο δ.μ. $\overline{X}_n$ είναι εκτιμήτρια του $\mu = g(\theta)$

$E(\overline{X}_n) = E(X_1) = \mu$ είναι α.ε. του μ.

$$MT\Sigma_{\overline{X}_n}(\theta) = \underbrace{b^2(\theta)}_{=0} + V(\overline{X}_n) = \frac{\sigma^2(\theta)}{n} \xrightarrow[n\to\infty]{} 0$$

- ο δ.μ. $\overline{X}_n$ είναι συνεπής. $(\forall \mu)$

$\overline{X}_n \xrightarrow{P} E(X_1) = \mu \Longrightarrow$ ο $\overline{X}_n$ είναι συνεπής εκτιμήτρια του μ.

Πρόταση

Αν $MT\Sigma'_{\hat{\theta}_{(n)}}(\theta) \xrightarrow[n\to\infty]{} 0 \Rightarrow$

η $\hat{\theta}_n$ είναι συνεπής εκτιμήτρια του θ.

Αποδ:

$$E\left[(\hat{\theta}_n-\theta)^2\right] = MT\Sigma'_{\hat{\theta}_n}(\theta).$$

θετική >0

$$\left[P(X > a) \leq \frac{E(X)}{a}\right]$$

ανισότητα Markov

Έστω $\varepsilon > 0$.

$$P\left(|\hat{\theta}_n-\theta| > \varepsilon\right) = P\left(\underbrace{(\hat{\theta}_n-\theta)^2}_{X} > \underbrace{\varepsilon^2}_{a}\right)$$

$$\leq \frac{E(\hat{\theta}_n-\theta)^2}{\varepsilon^2} \longrightarrow 0 \Rightarrow$$ η $\hat{\theta}_n$ είναι συνεπής.

Διάστημα Εμπιστοσύνης

Μας δίνει έναν άλλο τρόπο να εκτιμάμε μέσω διαστήματος και όχι σημεία, δίνοντας μας μια μεγαλύτερη εξασφάλιση ότι η άγνωστη παράμετρος θα βρίσκεται μέσα στο διάστημα που εμείς προτείναμε με μεγάλη πιθανότητα.

Η πιο δημοφιλής μέθοδος κατασκευής διαστηματικών εκτιμητριών $I(X)$

↳ Interval

είναι μέσω διαστημάτων εμπιστοσύνης CI (Confidence Interval) δεδομένου συντελεστή εμπιστοσύνης $1-\alpha$ (coefficient).

↳ μικρο.

$1-\alpha \in \{0.9, 0.95, 0.99, 0.999\}$.

$$\begin{array}{cccc} \downarrow & \downarrow & \downarrow & \downarrow \\ 90\% & 95\% & 99\% & 99.9\% \\ \alpha = 0.1 & \alpha = 0.05 & \alpha = 0.01 & \alpha = 0.001 \end{array}$$

Ορσ: Λέμε διάστημα εμπιστοσύνης με συντελεστή (βαθμό) εμπιστοσύνης $1-\alpha$, κάθε τυχαίο διάστημα (διαστηματική εκτιμήτρια) $I(X)$:

$$P_\theta\left(\underbrace{I(X) \ni \theta}\right) = 1-\alpha$$

ή

$$P_\theta\left(\theta \in \underbrace{I(X)}\right) = 1-\alpha, \quad \forall \theta \in \Theta.$$

→ Υπάρχουν πολλά $I(X)$ που θα μπορούσαμε να βρούμε και πολλοί τρόποι να τα βρούμε.

Η πιο συνηθισμένη μέθοδος είναι μέσω pivot, δηλ. μιας ποσότητας οδηγού (αντιστρεπτή ποσότητα) και την οποία θα δούμε σε ένα απλό παράδειγμα.

Παράδειγμα

Έστω $X_1, \ldots, X_n$ τ.δ. $N(\mu, \sigma^2)$, όπου το σ^2 είναι γνωστό και θέλουμε να βρούμε $(1-\alpha)$- Δ.Ε. για το μ.

$\boxed{\text{Μέθοδος κατασκευής}}$

Β1 Σκεφτόμαστε μια εκτιμήτρια του μ. εδώ το $\overline{X}$ εκτιμά το μ, και βρίσκουμε την κατανομή του, εδώ $\overline{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$

23

Ορσ: Λέμε ποσότητα οδηγό μια τ.μ. της μορφής $Q(\dot{X}, \dot{\mu})$

(γενικά δεν είναι σ.σ)

που έχει κατανομή ανεξάρτητη της παραμ. μ.

(γενικά αυτής που θέλαμε να φτιάξουμε το Δ.Ε.)

$$\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right) \Rightarrow Z = \frac{\dot{\bar{X}} - \dot{\mu}}{\frac{\sigma}{\sqrt{n}}} \sim N(0,1)$$

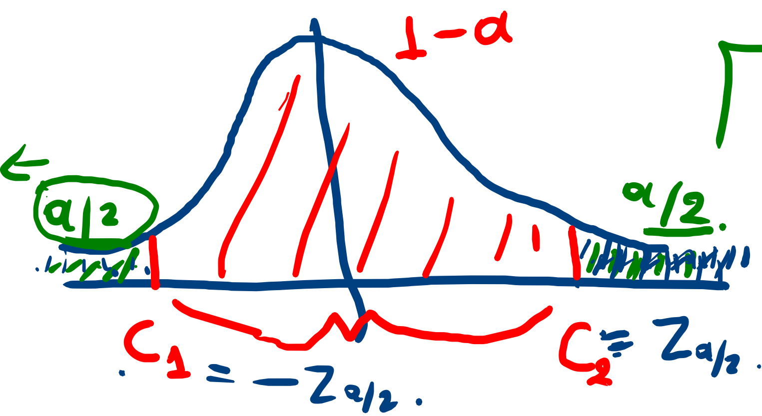
τυποποίηση

$\sigma \to$ γνωστό

Β2 εύρεση της ποσότητας οδηγού

B3 Βρίσκουμε σταθερές c_1 και c_2: 24

$$\boxed{P(c_1 \le Z \le c_2) = 1-\alpha}$$



→ με ίσες ουρές

⇓

→ συμμετρικά Δ.Ε.

(σε αυτήν την περίπτωση που έχουμε συμμετρική κατανομή)

$$\left.\begin{aligned} c_1 &: \Phi(c_1) = \frac{\alpha}{2} \\ c_2 &: P(Z > c_2) = \frac{\alpha}{2} \end{aligned}\right| \iff \left.\begin{aligned} P(Z > c_1) &= 1 - \frac{\alpha}{2} \\ P(Z > c_2) &= \frac{\alpha}{2}. \end{aligned}\right\} \iff$$

$$c_1 = z_{1-\frac{a}{2}} = -z_{\frac{a}{2}} \longrightarrow$$ άνω $\left(1-\frac{a}{2}\right)$-ποσοστημόριο της $N(0,1)$

$$c_2 = z_{\frac{a}{2}} \longrightarrow$$ άνω $\frac{a}{2}$-ποσοστημόριο της $N(0,1)$

Β4 Βρίσκαμε το Δ.Ε., λύνοντας τις ανισότητες ως προς μ.

$$c_1 \le Z \le c_2 \iff$$

$$-z_{a/2} \le \frac{\overline{X}-\mu}{\sigma/\sqrt{n}} \le z_{a/2} \iff$$

$$\boxed{\overline{X} - z_{a/2}\frac{\sigma}{\sqrt{n}} \le \mu \le \overline{X} + z_{a/2}\frac{\sigma}{\sqrt{n}}}$$

26

Τελικά

$$I^{\mu}_{1-a}(X) = \left[\overline{X} - Z_{a/2}\frac{\sigma}{\sqrt{n}}, \overline{X} + Z_{a/2}\frac{\sigma}{\sqrt{n}}\right]$$

$$= \overline{X} \pm Z_{a/2}\frac{\sigma}{\sqrt{n}}$$

Εκτιμήτρια   συμμ.   καταν. (ποσοστ. σημείο) $\sigma_{\overline{X}}$

$$P\left(\mu \in I^{\mu}_{1-a}(X)\right) = 1-a, \quad \forall \mu$$