Μαθημα 7ο 22/3/2017

Στατιστικη

Στατιστικη Συμπερασματολογια (Statistical Inference)

Εκτιμητική ή Εκτιμηση παραμετρων (Estimation of Parameter)

Ελεγχος Στατιστικων Υποθεσεων (Testing Statistical Hypothesis)

Σημειο Εκτιμητική ή Εκτιμηση σε σημειο (Point Estimation)

Εκτιμηση σε διαστημα ή Διαστημα Εμπιστοσυνης (Internal Estimation)

Παραδειγμα:

(Με Εκτιμητικη)

A, B υποψηφιοι. Μας ενδιαφερει η $p=P(A)$

X= αριθμος υποστηρικτων του A, $n=100$ (μεγεθος)

$X=x$ (θα παρει μια συγκεκριμενη τιμη) ή $p \approx \frac{x}{n}$ το ποσοστο που θα ψηφισει τον A

Αγνωστο p, το προσεγγιζω ετσι!

$X \sim B(n=100, p)$ → παραμετρος

Ενα ερωτημα που θελουμε να απαντησουμε ειναι το εξης:

Μηπως $p > \frac{1}{2}$; {το ερωτημα αυτο αφορα τον ελεγχο στατιστικων παραμετρων}

! Με την εκτιμητική προσδιορίζω μια ποσότητα από το δείγμα όταν πάρω πληροφορίες και έτσι προσεγγίζω την p με μια τιμή (εκτίμηση σε σημείο)

Σημείο Εκτίμησης

Έστω $x_1, \dots, x_n$ τ.δ από κάποιο πληθυσμό με άγνωστη μέση τιμή $\mu = E(x)$ (εμείς θέλουμε να μάθουμε για αυτό το μ του πληθυσμού)

Έστω ότι επιλέγω το $\bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i$

Το $\bar{x}$ είναι για εμάς μια εκτιμήτρια συνάρτηση με τιμή μετά το πείραμα, $\bar{x}$ που θα είναι για εμάς ο εκτιμητής του μ

⊛ Πρόβλημα 1:

Πώς βρίσκονται οι εκτιμήτριες συναρτήσεις; (μέθοδος ροπών)

($E(x) = \mu = \bar{x} = m$ εξισώνω)

Πρόβλημα 2:

Πώς αξιολογούνται;

(Βασικό κριτήριο) Αμεροληψία

Ένας εκτιμητής $\hat{\theta}$ του θ θα λέγεται αμερόληπτος αν η $E(\hat{\theta}) = \theta \;\; \forall \theta \in \Theta$ ή $\forall \theta$

↳ Ένας οποιοσδήποτε παραμετρικός χώρος!

⊛ Εκτιμητης $\bar{x}$

$\hat{\mu} = \bar{x}$ ή αλλη γραφη μεσα απο 6.6

$= T(x_1, \ldots, x_n)$

$= S(x_1, \ldots, x_n)$

Παραδειγμα :

Εστω $x_1, x_2, \ldots, x_{30}, \ldots, x_{40}$ τ.δ απο πληθυσμο με αγνωστη παραμετρο μ. Με ενδιαφερει να εκτιμησω το $\hat{\mu} =$;

Θα μπορουσα να παρω $\hat{\mu} = \bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i$

Παιρνω ως εκτιμητη το $\hat{\mu} = \bar{x}_{30}$ ή $\hat{\mu} = \bar{x}_{40}$

Ποιο απο τα δυο ομως βολευει;

Επιλεγω το $\hat{\mu} = \bar{x}_{40}$ γιατι χρησιμοποιω περισσοτερες πληροφοριες εφοσον το παιρνω οσο πιο κοντα μπορω στο συνολικο μεγεθος και δευτερος λογος η διακυμανση του ειναι μικροτερη

Γενικα : $Var(\bar{x}) = \frac{\sigma^2}{n}$

$Var(\bar{x}_{30}) = \frac{\sigma^2}{30}$, $Var(\bar{x}_{40}) = \frac{\sigma^2}{40}$ (σ^2 ιδιο)

!! Επιλεγω απο τους αμεροληπτους αυτον με τη μικροτερη διακυμανση.

$E(s^2) = \sigma^2$

$\hat{\sigma}^2 = s^2$ (αμεροληπτος εκτιμητης)

$E(\hat{\theta}-\theta)^2 = Var\,\hat{\theta}$

Παραδειγμα (απο προηγουμενο μαθημα):

Πληθυσμος $x: 3, 4, 5$

$p_x: \frac{1}{3}, \frac{1}{3}, \frac{1}{3}$

$P(X=x) = \frac{1}{3}$ για $x = 3, 4, 5$, $\mu = E(x) = 4$, $\sigma^2 = \frac{2}{3}$

Παιρνω δειγμα $n=2$ με επανοθεση

Δυνατες περιπτωσεις:

	(3,3)	(3,4)	(3,5)	(4,3)	(4,4)	(4,5)	(5,3)	(5,4)	(5,5)
$\bar{x}$:	3	3.5	4	3.5	4	4.5	4	4.5	5

$\bar{x}$:	3	3.5	4	4.5	5
$p_{\bar{x}}$:	$\frac{1}{9}$	$\frac{2}{9}$	$\frac{3}{9}$	$\frac{2}{9}$	$\frac{1}{9}$

$\mu_{\bar{x}} = E(\bar{x}) = 4$, $\sigma^2_{\bar{x}} = \frac{\sigma^2}{n} = \frac{\sigma^2}{2} = \frac{1}{3}$

Αν μ αγνωστος

Επιλεγω για την εκτιμηση του $\hat{\mu} = \bar{x}$ δειγμ. μεσος

Αναλογα σε ποια διαδα θα πεσω θα χρησιμοποιησω αυτην

για να μαθω το $\hat{\mu}$, την αντιστοιχη τιμη του $\bar{x}$

$\hat{\mu} = E(x)$ $(= n)$ αμεροληπτος

Διαστηματα Εμπιστοσυνης :

Εστω $x_1, x_2, \dots, x_n$ τ.δ απο πληθυσμο (θ), θα μπορουσαμε να χρησιμοποιησουμε τις 6.6

$L = L(x_1, \dots, x_n)$ με τιμες $\ell = \ell(x_1, \dots, x_n)$

$U = U(x_1, \dots, x_n)$ $u = u(x_1, \dots, x_n)$

για να εκτιμησουμε το θ στο διαστημα τιμων (L, U) με την προϋποθεση οτι:

① να εχει οσο γινεται μικροτερο μηκος

② να περιεχει το "αγνωστο" θ ενα μεγαλο ποσοστο φορων

μεγαλο ποσοστο φορων $\equiv (1-a) \cdot 100\%$ για ενα $(1-a)\,100\%$ διαστημα εμπιστοσυνης

Για το θ ισχυει η σχεση:

$$P(L < \theta < U) = 1-a$$

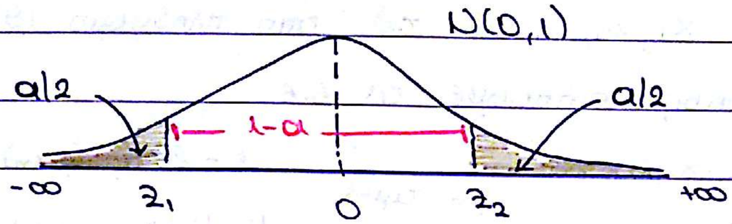
παραδειγμα :

Να βρεθει ενα $(1-a)\,100\%$ Δ.Ε (Διαστημα Εμπιστοσυνης) για το μ ενος κανονικου πληθυσμου $N(\mu, \sigma^2 = \text{γνωστο})$ οπου $x_1, \dots, x_n$ τ.δ απο τον πληθυσμο

απαντηση :

Ξεκινω με εναν εκτιμητη του $\hat{\mu} = \bar{x}$

$$\bar{x} \sim N\left(\mu, \frac{\sigma^2}{n}\right) \rightsquigarrow \frac{\bar{x}-\mu}{\sigma / \sqrt{n}} \sim N(0,1)$$ ακριβες αποτελεσμα



$z_1 = -z_{\frac{a}{2}}$, $z_2 = z_{\frac{a}{2}}$

θα βρω δυο τιμες z_1, z_2 που να ικανοποιουν την σχεση

$$P\left(\underbrace{z_1}_{L} \le \frac{\bar{x}-\mu}{\sigma / \sqrt{n}} \le \underbrace{z_2}_{U}\right) = 1-a$$

$$\Rightarrow P\left(-z_{\frac{a}{2}} \le \frac{\bar{x}-\mu}{\sigma / \sqrt{n}} \le z_{\frac{a}{2}}\right) = 1-a$$

$$\Rightarrow P\left(-z_{\frac{a}{2}} \cdot \frac{\sigma}{\sqrt{n}} - \bar{x} \le -\mu \le z_{\frac{a}{2}} \cdot \frac{\sigma}{\sqrt{n}} - \bar{x}\right) = 1-a$$

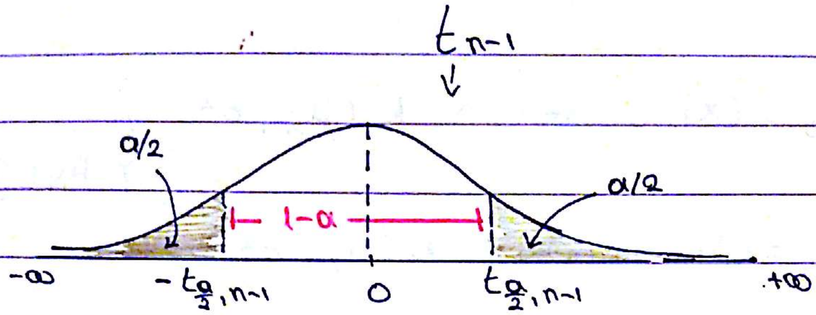
$$\Rightarrow P\left(\bar{x} - z_{\frac{a}{2}} \cdot \frac{\sigma}{\sqrt{n}} \le \mu \le \bar{x} + z_{\frac{a}{2}} \cdot \frac{\sigma}{\sqrt{n}}\right) = 1-a$$

Δηλαδη $L = \bar{x} - z_{\frac{a}{2}} \cdot \frac{\sigma}{\sqrt{n}}$, $U = \bar{x} + z_{\frac{a}{2}} \cdot \frac{\sigma}{\sqrt{n}}$

! Αν εχω προσδιορισει το α
παω στον πινακα και τα βρισκω

$t_{n-1} \sim \frac{\bar{x}-\mu}{s / \sqrt{n}}$ (ομοια με πριν τα βρισκω)

$$P\left(-t_{\frac{a}{2}, n-1} \le \frac{\bar{x}-\mu}{s / \sqrt{n}} \le t_{\frac{a}{2}, n-1}\right) = 1-a$$



$\rightarrow \quad L = \bar{x} - t_{\frac{a}{2},n-1} \cdot \frac{s}{\sqrt{n}}$

$U = \bar{x} + t_{\frac{a}{2},n-1} \cdot \frac{s}{\sqrt{n}}$

Παράδειγμα 4.1 | Βιβλίο Λουκά :

$n = 20$ άτομα (μέγεθος)

$\sigma^2 = 25$, 95% Δ.Ε για το μ

$x_i \sim N(\mu, \sigma^2 = 25)$, x_i διάφορες

απάντηση :

Σύμφωνα με τα προηγούμενα

$$\left.\begin{aligned} L &= \bar{x} - z_{\frac{a}{2}} \cdot \frac{\sigma}{\sqrt{n}} \\ U &= \bar{x} + z_{\frac{a}{2}} \cdot \frac{\sigma}{\sqrt{n}} \end{aligned}\right\} \rightarrow \begin{aligned} \bar{x} &= 0.6 \\ \sigma &= 5 \\ a &= 0.05 \end{aligned} \quad \text{Άρα} \quad \left.\begin{aligned} L &= \bar{x} - z_{0.025} \cdot \frac{\sigma}{\sqrt{n}} \\ U &= \bar{x} + z_{0.025} \cdot \frac{\sigma}{\sqrt{n}} \end{aligned}\right\}$$

$\rightarrow \quad (\ell, u) = (-1.59, 2.79)$

$\underbrace{P(-1.59 \le \mu \le 2.79) \neq 0.95 \ (0 \text{ ή } 1)}_{\text{Δεν μπορώ να το γράψω αυτό.}}$

⊛ Τα παίρνει "έτοιμο" από το βιβλίο συνεπώς δες το βιβλίο του...

Ασκηση 4.11 Βιβλίο Λαιτα :

$\underset{\sim}{X} = (X_1, \dots, X_n) \sim N(\mu_X, \sigma^2)$

Ανεξαρτητες

$\underset{\sim}{Y} = (Y_1, \dots, Y_n) \sim N(\mu_Y, \sigma^2)$

1) ΝΔΟ $\forall\, \omega \in [0,1]$ η σ.σ $\eta = U(X, Y) = \omega S^2_X + (1-\omega) S^2_Y$ είναι αμερόληπτος εκτιμητής του σ^2

$E(\eta) = \omega E(S^2_X) + (1-\omega) E(S^2_Y) = \omega \sigma^2 + (1-\omega)\sigma^2 = \sigma^2$

2) Για ποιο ω η $var\,\eta \to min$;

$Var\,\eta = \omega^2 var(S^2_X) + (1-\omega)^2 var(S^2_Y) = g(\omega), \dots$