Στατιστική Συμπερασματολογία → Εκτίμηση Παραμέτρων ή Εκτιμητική (Parameter Estimation)
→ Έλεγχος Στατιστικών Υποθέσεων (Statistical Testing)

<u>ΠΑ</u> (Εκλογές)

A, B

p = ποσοστό του A $\rightsquigarrow p \cong \frac{x}{n}$

$n = 100$

X = ψηφοφόροι του A $\rightsquigarrow X = 70$

Χρησιμοποιούμε Διωνυμική κατανομή : $X \sim B(n=100, p)$

Αν είναι $p > \frac{1}{2}$ τότε το A κερδίζει τις εκλογές

Εκτίμηση Παραμέτρων → Σημειοεκτιμητική ή Εκτίμηση σε σημείο (Point Estimation)
→ Διαστήματος Εμπιστοσύνης ή Εκτίμηση σε Διάστημα (Interval Estimation)

<u>ΠΑ</u>

τ.δ. $X_1, \dots, X_n$ με μέση τιμή μ

$\bar{x} \cong \mu$

↳ Εκτιμήτρια συνάρτηση ή Εκτιμητής

$L(X_1, \dots, X_n)$: L κάτω άκρο

$U(X_1, \dots, X_n)$: U άνω άκρο

<u>Εκτίμηση σε σημείο</u>

Για παράμετρο θ: εκτιμήτρια συνάρτησης, $\hat{\theta}$ ή $T(X_1, \dots, X_n)$

1) Πως βρίσκονται οι εκτιμήτριες συναρτήσεις

$E(X) = \mu$ και $m_1 = \bar{x} \implies \hat{\mu} = \bar{x}$

2) Πως αξιολογούμε (κρ. αξιολόγησης) → Αμεροληψία
→ Ελάχιστη Διακύμανση

- Ένας εκτιμητής θ θα λέγεται αμερόληπτος αν $\forall \theta$: $E(\hat{\theta}) = \theta$

Στο δείγμα: $X_1, \dots, X_{10}, X_{11}, \dots, X_{20}$ με $\bar{X}_{10} = \frac{1}{10}\sum_{i=1}^{10} X_i$, $\hat{\mu} = \bar{X}_{10}$

έχουμε $E(\bar{X}_{10}) = \mu$ και $Var(\bar{X}_{10}) = \frac{\sigma^2}{10}$

$\bar{X}_{20} = \frac{1}{20}\sum_{i=1}^{20} X_i$, $\hat{\mu} = \bar{X}_{20}$ έχουμε $E(\bar{X}_{20}) = \mu$ (Αμερόληπτος Εκτιμητής)

και $Var(\bar{X}_{20}) = \frac{\sigma^2}{20}$ (Μικρότερη Διακύμανση)

$E(S^2) = \sigma^2$ με $\hat{\sigma}^2 = S^2$

Πχ

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

Δείγματα: (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

$\bar{x}=\bar{x}$:	3	3,5	4	4,5	5
$P(\bar{X}=x)$:	$\frac{1}{9}$	$\frac{2}{9}$	$\frac{3}{9}$	$\frac{2}{9}$	$\frac{1}{9}$

με $E(\bar{x})=4(=\mu)$ και $\sigma^2_{\bar{x}}=\frac{1}{3}\left(=\frac{\sigma^2}{n}\right)$

Διαστήματα Εμπιστοσύνης

Έστω $X_1,\dots,X_n$ τ.δ. από πληθυσμό (θ) κ $L=L(X_1,\dots,X_n)$ κ $U=U(X_1,\dots,X_n)$ ως κάτω και άνω άκρα ενός διαστήματος τιμών

Θα μπορούσαμε να χρησιμοποιούμε το $[L,U]$ με τιμές $[l,u]$ για να εκτιμήσουμε το θ αρκεί το $[L,U]$ να έχει προκύψει με τις ιδιότητες.

1. "Ελάχιστο μήκος"
2. Να περιέχει την "αληθινή" τιμή του θ ένα μεγάλο ποσοστό φορών.

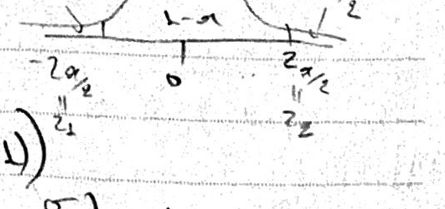
Βαθμός εμπιστοσύνης $=P(L\leq\theta\leq U)=1-\alpha$, όπου $0<\alpha<1$ είναι συνήθως 0,01 ή 0,05. Προκύπτει έτσι το $(1-\alpha)100\%$ διάστημα εμπιστοσύνης (Δ.Ε.) για το θ

Πχ

$X_1,\dots,X_n$ τ.δ. από $N(\mu,\sigma^2=\text{γνωστή})$

$(1-\alpha)100\%$ Δ.Ε. για το μ;

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



Άρα $P\left(z_1\leq\frac{\bar{x}-\mu}{\sigma/\sqrt{n}}\leq z_2\right)=1-\alpha$ (σε $N(0,1)$)

$\Rightarrow\dots\Rightarrow P\left(\bar{x}-z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\leq\mu\leq\bar{x}+z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right)=1-\alpha$

$L=\bar{x}-z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$ και $U=\bar{x}+z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$

$\hat{\mu}=\bar{x}$, $\hat{\sigma}^2=S^2$, $\frac{\bar{x}-\mu}{S/\sqrt{n}}\sim t_{n-1}$ άρα $P\left(-t_{\alpha/2,n-1}\leq\frac{\bar{x}-\mu}{S/\sqrt{n}}\leq t_{\alpha/2,n-1}\right)=1-\alpha$

$\Rightarrow P\left(\underbrace{\bar{x}-t_{\alpha/2,n-1}\frac{S}{\sqrt{n}}}_{L(x_1,\dots,x_n)}\leq\mu\leq\underbrace{\bar{x}+t_{\alpha/2,n-1}\frac{S}{\sqrt{n}}}_{U(x_1,\dots,x_n)}\right)$

Εφαρμογή

Πχ. 4.1

7, 6, 0, 8, -3, -4, 0, 1, -3, 1, 2, 7, 0, 6, -6, -5, -1, 6, -2, 4

(Πριν και μετά)

$n=20$ και $N(\mu, \sigma^2=25)$, 95% Δ.Ε. για το μ.

$1-\alpha=0{,}95 \Rightarrow \alpha=0{,}05$, $\frac{\alpha}{2}=0{,}025$ $z_{\alpha/2}=1{,}96$

$\bar{x}=0{,}6$, $\sigma=5$ $n=20$

$L=0{,}6-1{,}96\frac{5}{\sqrt{20}}=-1{,}59$

$U=0{,}6+1{,}96\frac{5}{\sqrt{20}}=2{,}79$

95% Δ.Ε. για το μ: $[-1.59, 2.79]$