Εισαγωγή στη Στατιστική

28/3/2018.

Διάλεξη 10η

σταντάρ θέμα:

Εφαρμογές Απλής Στατιστικής Συμπερασματολογίας

▶ Συμπερασματολογία: για την μέση τιμή (μ) ενός πληθυσμού.

~~Παράδειγμα 4.2 (βιβλίο)~~

~~Ένας υγειονομικός σταθμός...~~

~~170, 175, 190, 198, 215, 185, 184, 207~~

~~210, 193, 196, 180 με $n=12$ και $N(\mu, \sigma^2)$~~

~~• Υπάρχει λόγος ανησυχίας; ($\mu \geq 200$)~~

~~Η παράμετρος που με ενδιαφέρει είναι το μ.~~

~~Απάντηση:~~

ΓΕΝΙΚΑ έστω ένα τ.δ $x_1, x_2, \dots, x_n$ από πληθ (μ, σ^2) και

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i \quad , \quad S^2 = \frac{1}{n-1}\sum_{1}^{n}(x_i - \bar{x})^2$$

• Μας ενδιαφέρει η παράμετρος μ.

I) Πληθυσμός Κανονικής, σ^2 = γνωστό

$N(\mu, \sigma^2 = \text{γνωστό})$

$$\hat{\mu} = \bar{x} \quad , \quad \sigma^2_{\bar{x}} = \frac{\sigma^2}{n} \quad , \quad \sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} \quad \Bigg\| \quad \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} \sim N(0,1)$$

διάστημα εμπιστοσύνης

Ένα $(1-a)100\%$ Δ.Ε για το μ:

$$\boxed{L = \bar{x} - z_{a/2}\frac{\sigma}{\sqrt{n}}}$$

$$\boxed{U = \bar{x} + z_{a/2}\frac{\sigma}{\sqrt{n}}}$$

▷ Για να ελέγξουμε τις υποθέσεις:

(i) $H_0: \mu \leq \mu_0$ V $H_a: \mu > \mu_0$ (με μ_0 γνωστό)

(ii) $H_0: \mu \geq \mu_0$ V $H_a: \mu < \mu_0$

(iii) $H_0: \mu = \mu_0$ V $H_a: \mu \neq \mu_0$

Χρησιμοποιούμε το στατιστικό

$Z = \dfrac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} \overset{H_0}{\sim} N(0,1)$ και κρίσιμες περιοχές μεγέθους α (επίπεδο σημαντικότητας).

⇝ Για τον (i) έλεγχο η κρίσιμη περιοχή είναι:

- $C = [z_a, \infty)$ ή $Z \geq Z_\alpha$

⇝ Για τον (ii) έλεγχο η κρίσιμη περιοχή είναι:

- $C = (-\infty, -Z_a)$ ή $Z \leq -Z_a$

⇝ Για τον (iii) έλεγχο η κρίσιμη περιοχή είναι:

- $C = (-\infty, -Z_{a/2}] \cup [Z_{a/2}, +\infty)$ ή $|Z| \geq Z_{a/2}$

όπου $\alpha = P($απόρριψη $H_0 \,/\,$ αληθ $H_0)$

II) Πληθυσμός κανονικής, σ^2 άγνωστο, $N(\mu, \sigma^2$ - άγνωστο$)$

$\hat{\mu} = \bar{X}$, $\sigma^2_{\bar{X}} = \dfrac{\sigma^2}{n}$ και $\hat{\sigma}^2_{\bar{X}} = \dfrac{S^2}{n}$

$E(S^2) = \sigma^2$, $\hat{\sigma}_{\bar{X}} = \dfrac{S}{\sqrt{n}}$

Είναι $\dfrac{\bar{X} - \mu}{S/\sqrt{n}} \sim t_{n-1}$ (κατανομή με n-1 βαθμούς ελευθερίας)

$\rightarrow P\left(-t_{a/2, n-1} \leq \dfrac{\bar{X} - \mu}{S/\sqrt{n}} \leq t_{a/2, n-1}\right) = 1 - a$

Είναι $(1-a)100\%$ ΔΕ για το μ:

$\boxed{L = \bar{X} - t_{a/2, n-1} \dfrac{S}{\sqrt{n}}}$

$\boxed{U = \bar{X} + t_{a/2, n-1} \dfrac{S}{\sqrt{n}}}$

1. Για να ελέγξουμε τις υποθέσεις (i), (ii), (iii)

Χρησιμοποιούμε το στατιστικό

$$t = \frac{\bar{X} - \mu_0}{S/\sqrt{n}} \overset{H_0}{\sim} t_{n-1}$$ και κρίσιμες περιοχές μεγέθους α

ποσοστό ρίσκου που θα έχω στο συμπέρασμα

→ Για τον (i) έλεγχο η κρίσιμη περιοχή είναι:

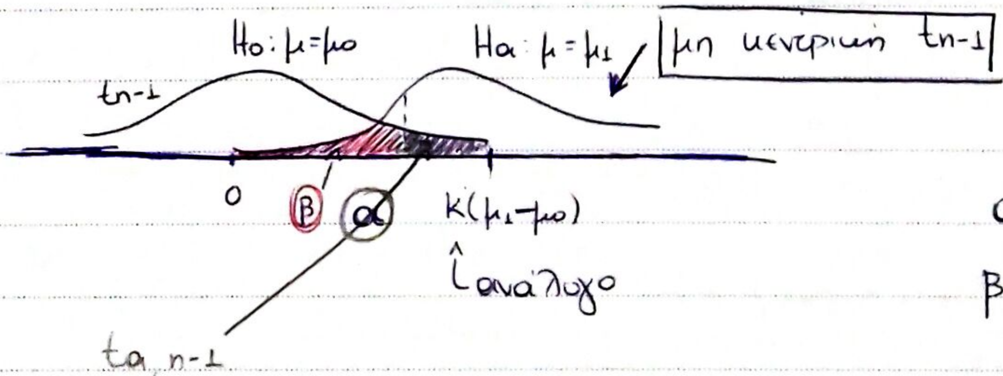
- $C = [t_{\alpha,n-1}, \infty)$ ή $t \geq t_{\alpha,n-1}$

→ Για τον (ii) έλεγχο η κρίσιμη περιοχή είναι:

- $C = (-\infty, -t_{\alpha,n-1})$ ή $t \leq -t_{\alpha,n-1}$

→ Για τον (iii) έλεγχο η κρίσιμη περιοχή είναι:

- $|t| \geq t_{\alpha/2,n-1}$.



$\alpha = P($απορρ $H_0 /$ αληθ $H_0)$

$\beta = P($δεχομαι $H_0 /$ αληθ $H_0)$

● Παραδειγμα 4.2 (βιβλιο)

170, 175, 190, 198, 215, 185, 184, 207,

210, 193, 196, 180

$n=12$, $N(\mu, \sigma^2)$

· Υπαρχει λογος ανησυχιας; ($\mu \geq 200$)

<u>Λυση</u> (Παιρνουμε 12 δειγματα)

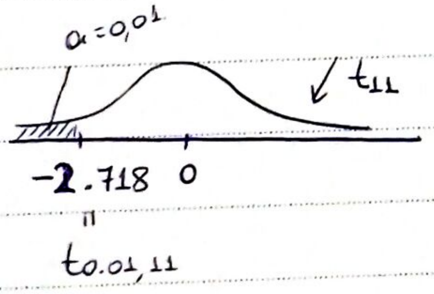
Ειμαστε σε κανονική κατανομή, με αγνωστη διακυμανση.

Αρα:

$$t = \frac{\bar{x} - \mu_0}{S/\sqrt{n}}$$, κρισιμη περιοχη: $t \leq -t_{\alpha, n-1} \left(= -t_{0.01, 11} = -2.718\right)$

πηγαινετε να καυτεψητε (επιτρεπεται)

το $H_0 : \mu \geq \mu_0 (=200)$ V $H_a : \mu < 200$, $\alpha = 0.01$.



Απο τα δεδομένα ο δειγματικός μέσος θα είναι:

$\bar{X} = 191.9$

$S^2 = 196.841$

$S = 14.03$

Αρα $$t = \frac{191.9 - 200}{14.03/\sqrt{12}} = -2.197$$

όμως $-2.197 \not< -2.718$ δεν απορρ H_0

$(1-\alpha)100\%$ ΔΕ για το μ : $\bar{X} \pm t_{\alpha/2, n-1} \frac{S}{\sqrt{n}}$, $\alpha = 0,05$

$t_{0.005, 11} = 2.201$

95% ΔΕ για το μ : [182.99, 200.81]

έχει πιασει και το 200

Κ.Ο.Θ (κεντρικο ολικό θεωρημα)

III) Μη Κανονικός Πληθυσμός, n μεγάλο (Κ.Ο.Θ)

$$\hat{\mu} = \bar{x}, \quad \hat{\sigma}^2_{\bar{x}} = \frac{S^2}{n}, \quad \hat{\sigma}_{\bar{x}} = \frac{S}{\sqrt{n}}$$

$$\bar{X} \overset{\text{προσεγ}}{\sim} N\left(\mu, \frac{S^2}{n}\right) \Rightarrow \frac{\bar{x}-\mu}{S/\sqrt{n}} \overset{\text{προσεγγ}}{\sim} N(0,1).$$

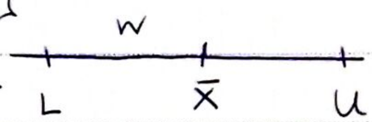
$(1-\alpha)100\%$ ΔΕ για το μ: $\boxed{L = \bar{x} - Z_{\alpha/2}\frac{S}{\sqrt{n}}}$

$$\boxed{U = \bar{x} + Z_{\alpha/2}\frac{S}{\sqrt{n}}}$$

▷ Για τον έλεγχο των υποθέσεων (i) (ii) (iii)

Χρησιμοποιούμε το στατιστικό

$$Z = \frac{\bar{x}-\mu_0}{S/\sqrt{n}} \underset{\text{προσεγ}}{\overset{H_0}{\sim}} N(0,1)$$ και κρίσιμες περιοχές



$$U - L = 2w = 2\, t_{\alpha/2, n-1}\frac{S}{\sqrt{n}}$$

$$n = \frac{t^2_{\alpha/2, n-1} S^2}{w^2} \approx \frac{Z^2_{\alpha/2}\sigma^2}{w^2}$$

$D_i = X_i - Y_i$
μετα πριν

με τ.δ $D_1, \ldots, D_n$, $D_i \sim N(\mu, \sigma^2 = \text{αγνωστη})$

τοτε εφαρμο τη t.

Παραδειγμα 4.3 (βιβλιο)
↳ στηρίζεται στα σημερινα)

Εστω $\bar{X} = 6$ και $S^2 = 64$ με $n = 25$
$\downarrow S = 8$

$H_0: \mu = 0$ $H_a: \mu > 0$

Η κρισιμη περιοχή $t \geq t_{0,05,24}$ $(= t_{\alpha, n-1})$ $(= 1.711)$

$$t = \frac{\bar{X}_0 - 0}{\sqrt{64}/\sqrt{25}} = 3.75$$

Όμως $3.75 > 1.711$ συνεπως απορρ H_0.