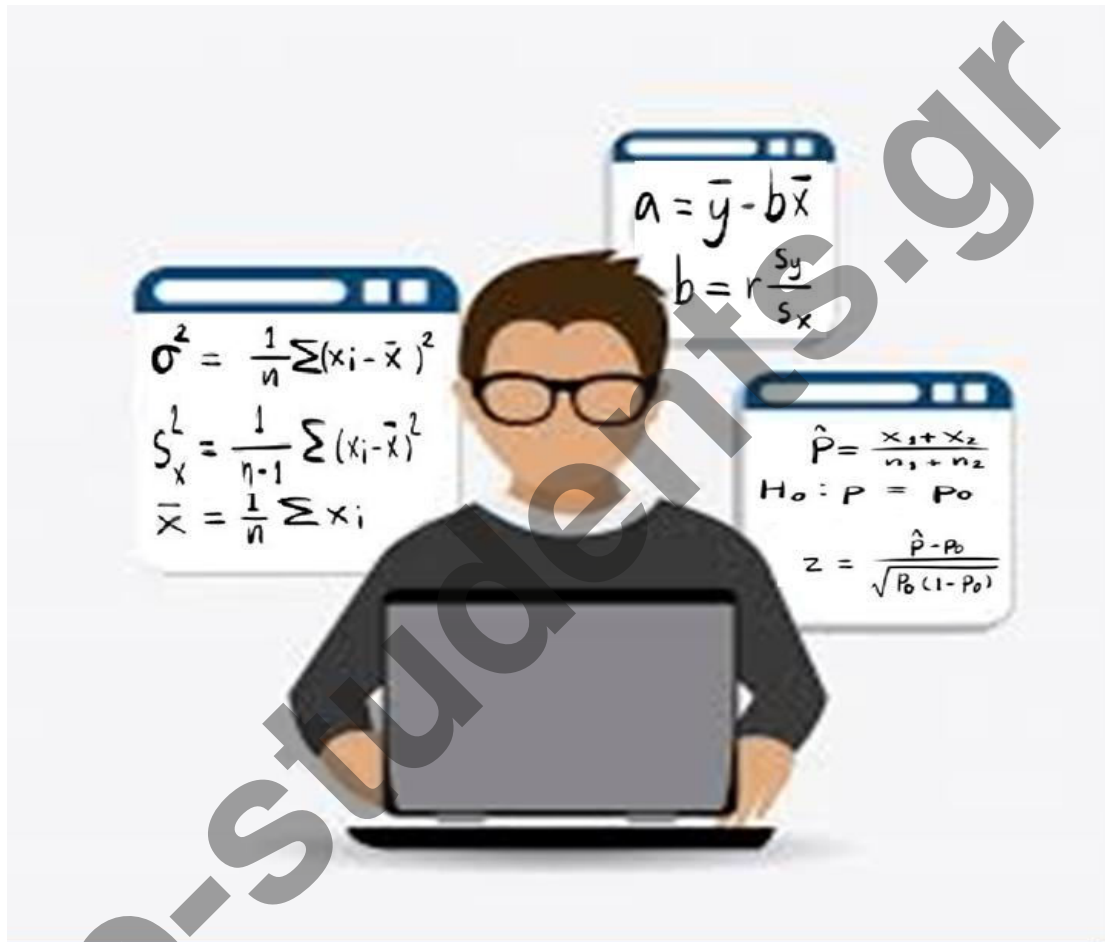


ΣΤΑΤΙΣΤΙΚΗ



Πληθυσμός

Είναι το σύνολο των δεδομένων το οποίο θέλουμε να μελετήσουμε.

Δείγμα

Ένα μέρος του πληθυσμού που επιλέγεται με τυχαίο τρόπο.

Παράμετροι Πληθυσμού

Η εξαγωγή όσο των δυνατόν ασφαλέστερων συμπερασμάτων για τις άγνωστες παραμέτρους του πληθυσμού όπως η μέση τιμή, η διασπορά, το ποσοστό των ατόμων με κάποιο χαρακτηριστικό και άλλες.

$\mu, \sigma, \sigma^2, \dots$



<u>Πληθυσμός</u>	<u>Δείγμα</u>
N	n
μ	$\bar{x}$
σ^2	S^2
σ	S

● Δειγματική Μέση Τιμή $\bar{x}$



Σημειώσεις - Λυμένα Θέματα

e-students.gr

☎ 698 93 94 741

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{n}$$

$$\text{ή } \bar{x} = \frac{\sum f_i x_i}{n}$$

n : μέγεθος του δείγματος

x_i : παρατηρήσεις του δείγματος

f : συχνότητα

● Διασπορά S^2



Σημειώσεις - Λυμένα Θέματα

e-students.gr

☎ 698 93 94 741

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

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

● Τυπική απόκλιση S

$$S = \sqrt{S^2}$$

Κεντρικό Οριακό Θεώρημα (κ.ο.θ)

Έστω $X_1, X_2, \dots, X_n$ τυχαίο δείγμα ατόμων
υποτιμή με μέση τιμή μ και διασπορά σ^2

$$\frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \sim N(0,1) \text{ προσεγγιστικά}$$

ή

$$\frac{\sum X_i - n\mu}{\sigma\sqrt{n}} \sim N(0,1) \text{ προσεγγιστικά}$$

● Διάστημα Εμπιστοσύνης



Σημειώσεις - Λυμένα Θέματα

e-students.gr

© 698 93 94 741

Είναι το εύρος τιμών που πιστεύουμε
ότι περιλαμβάνει τη μέση τιμή της
αγνωστής παραμέτρου που μελετάμε

Διάστημα Εμπιστοσύνης

για μέση τιμή

για διακύμανση

για ποσοστό



Σημειώσεις - Λυμένα Θέματα

e-students.gr

© 698 93 94 741

● Διάστημα εμπιστοσύνης για τη μέση
τιμή μ του πληθυσμού



Σημειώσεις - Λυμένα θέματα

e-students.gr

☎ 698 93 94 741

→ Με γνωστή διασπορά σ

$$\bar{x} \pm z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$$

(διηλεκτρο)

→ Με άγνωστη διασπορά σ

$$\bar{x} \pm t_{n-1, \alpha/2} \frac{s}{\sqrt{n}}$$

$$s^2 = \frac{\sum (x_i - \bar{x})^2}{n-1}$$



Σημειώσεις - Λυμένα θέματα

e-students.gr

☎ 698 93 94 741

● Εφαρμογή

Σε ένα δείγμα μεγέθους 16 με μέση τιμή 45 και τυπική σφύραση 2,3 να κατασκευαστεί ένα 96% ΔΕ για την μέση τιμή πληθυσμού.

Λύση

$$\bar{x} \pm Z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$$
$$45 \pm 2,06 \cdot \frac{2,3}{\sqrt{16}}$$

$$(43,2, 46,18)$$

$$\bar{x} = 45$$

$$\sigma = 2,3$$

$$n = 16$$

$$96\% \text{ ΔΕ}$$

$$\rho = \frac{1+\alpha}{2} = \frac{1+0,96}{2}$$

$$\alpha = 0,98$$

πρώτο πίνακα
και βρούμε
 $Z_{\alpha/2} = 2,06$

● Εφαρμογή

Για τον προσδιορισμό της αντοχής ενός υλικού είχαν οι εξής μετρήσεις
20.2 21.7 18.1 20.6 22.2 19.8
17.9

Να κατασκευαστεί το 95% ΔΕ για την μέση τιμή της αντοχής του υλικού.

Λύση

$$n = 7$$

$$\bar{x} = \frac{\sum x_i}{n} = \frac{20,2 + 21,7 + 18,1 + 20,6 + 22,2 + 19,8 + 17,9}{7}$$

$$\bar{x} = 20,07$$

Πειραματική διασπορά $S^2 = \frac{1}{n-1} \sum (x_i - \bar{x})^2$

$$S^2 = 2,694$$

$$S = 1,64$$

$$\bar{x} \pm t_{n-1, \alpha/2} \frac{S}{\sqrt{n}}$$

$$t_{n-1, \alpha/2} = ?$$

$$20,07 \pm 2,447 \frac{1,64}{\sqrt{7}}$$

$$0 = \frac{1 + 0,95}{2}$$

$$0 = 0,975$$

$$(18,55, 21,59)$$

Πίνακας $t(1-0,975)$
 $t_{0,975} = 2,447$

● Διαστήματα εμπιστοσύνης για την
διαφορά των μέσων $\mu_1 - \mu_2$ δύο
ανεξάρτητων πληθυσμών



Σημειώσεις - Λυμένα θέματα

e-students.gr

© 698 93 94 741

→ Με γνωστές διασπορές

$$\bar{x} - \bar{y} \pm Z_{\alpha/2} \sqrt{\frac{\sigma_x^2}{n} + \frac{\sigma_y^2}{m}}$$

→ Με άγνωστες διασπορές $\sigma_1^2 = \sigma_2^2 = \sigma^2$

$$\bar{x} - \bar{y} \pm t_{n+m-2, \alpha/2} \cdot S_0 \sqrt{\frac{1}{n} + \frac{1}{m}}$$

όπου $S_0^2 = \frac{(n-1)S_x^2 + (m-1)S_y^2}{n+m-2}$

→ Με άγνωστες διασπορές $\sigma_1^2 \neq \sigma_2^2$

$$\bar{x} - \bar{y} \pm t_{V, \alpha/2} \sqrt{\frac{S_x^2}{n} + \frac{S_y^2}{m}}$$

όπου $V = \frac{\left(\frac{S_x^2}{n} + \frac{S_y^2}{m}\right)^2}{\frac{(S_x^2/n)^2}{n-1} + \frac{(S_y^2/m)^2}{m-1}}$



Σημειώσεις - Λυμένα θέματα

e-students.gr

© 698 93 94 741

● Διαστήματα Εμπιστοσύνης για την
διασπορά και την τυμική απόκλιση
ενός πληθυσμού

$$\left[\frac{\sqrt{n-1}S^2}{\sqrt{\chi^2_{n-1, \alpha/2}}}, \frac{\sqrt{n-1}S^2}{\sqrt{\chi^2_{n-1, 1-\alpha/2}}} \right]$$

● Διαστήματα Εμπιστοσύνης για τον λόγο $\frac{\sigma_x^2}{\sigma_y^2}$
των διασπορών δύο ανεξάρτητων
πληθυσμών

$$\left[\frac{1}{F_{n-1, m-1, \alpha/2}} \cdot \frac{S_x^2}{S_y^2}, \frac{1}{F_{n-1, m-1, 1-\alpha/2}} \cdot \frac{S_x^2}{S_y^2} \right]$$

● Διάστημα Εμπιστοσύνης για το Ποσοστό
 p ενός πληθυσμού



Εμπειρίες - Αμείβοντάς σε

e-students.gr

☎ 698 93 94 741

$$\hat{p} \pm z_{\alpha/2} \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \quad \text{όπου } \hat{p} = \frac{x}{n}$$

● Διαστήματα Εμπιστοσύνης για την Διαφορά
των Ποσοστών $p_1 - p_2$ δύο ανεξάρτητων
πληθυσμών

$$\hat{p}_1 - \hat{p}_2 \pm z_{\alpha/2} \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n} + \frac{\hat{p}_2(1-\hat{p}_2)}{m}}$$

όπου $\hat{p}_1 = \frac{x_1}{n}$
 $\hat{p}_2 = \frac{x_2}{m}$



Εμπειρίες - Αμείβοντάς σε

e-students.gr

☎ 698 93 94 741

Εφαρμογή



Επικοινωνία - Λογμένα Θέματα

e-students.gr

© 698 93 94 741

Από τα προϊόντα μηχανής τριανούρε τυχαίο δείγμα 125. Ανάμεσα σε αυτά υπάρχουν 7 ελαττωματικά. Να βρεθεί το 98% διάστημα εμπιστοσύνης για το ποσοστό των ελαττωματικών προϊόντων που παράγονται από την συγκεκριμένη μηχανή

Λύση

$$\hat{p} = \frac{7}{125} = 0,056$$



Επικοινωνία - Λογμένα Θέματα

e-students.gr

© 698 93 94 741

98% Δ.Ε. $\rightarrow p = \frac{1+0,98}{2} = 0,99$ πάλι στον πίνακα Z
 $Z_{\frac{\alpha}{2}} = 2,33$

$$\hat{p} \pm Z_{\frac{\alpha}{2}} \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

$$0,056 \pm 2,33 \sqrt{\frac{0,056(1-0,056)}{125}}$$

$$(0,036, 0,076)$$

P-value



Σημειώσεις - Λυμένα θέματα

e-students.gr

☎ 698 93 94 741

Παρατηρούμενο επίπεδο σημαντικότητας (ή απλά σημαντικότητα) ενός ελέγχου, ονομάζεται η πιθανότητα να παρατηρηθεί μια τιμή του στατιστικού τόσο ή και περισσότερο αξιωματική που έδωσε το δείγμα, δεδομένου ότι η H_0 είναι αληθινή.

Η H_0 απορρίπτεται εάν η παρατηρούμενη σημαντικότητα (p-value) είναι μικρότερη ή ίση με το α .



Σημειώσεις - Λυμένα θέματα

e-students.gr

☎ 698 93 94 741

Έλεγχος Υποθέσεων



Σημειώσεις - Λυμένα Θέματα

e-students.gr

© 698 93 94 741

Η μηδενική υπόθεση που συμβολίζεται με H_0 είναι ένας ισχυρισμός αναφορικά με μία ή περισσότερες παραμέτρους του πληθυσμού. Ο ισχυρισμός αυτός θεωρείται αληθής έως ότου συγκεντρώσουμε επαρκή στατιστικά στοιχεία για το αντίθετο.

Η εναλλακτική υπόθεση συμβολίζεται με H_1 και είναι ο ισχυρισμός που περιλαμβάνει όλες τις καταστάσεις που δεν καλύπτει η μηδενική υπόθεση.



Σημειώσεις - Λυμένα Θέματα

e-students.gr

© 698 93 94 741

Έλεγχος Υποθ. $\begin{cases} \text{για την μέση τιμή} \\ \text{για την διασπορά} \\ \text{για το προσδό} \end{cases}$

● Έλεγχος υπόθεσης για το μέσο μ ενός πληθυσμού



→ Με γνωστή διασπορά σ

$$Z = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}$$

• $H_0: \mu = \mu_0$
• $H_1: \mu \neq \mu_0$

Αν $|Z| > Z_{\alpha/2}$

Απορρ. την H_0

• $H_0: \mu = \mu_0$
• $H_1: \mu > \mu_0$

Αν $Z > Z_{\alpha}$

Απορρ. την H_0

• $H_0: \mu = \mu_0$
• $H_1: \mu < \mu_0$

Αν $Z < -Z_{\alpha}$

Απορρ. την H_0

→ Με άγνωστη διασπορά σ

$$t = \frac{\bar{X} - \mu_0}{s/\sqrt{n}}$$



• $H_0: \mu = \mu_0$
• $H_1: \mu \neq \mu_0$

Αν $|t| > t_{\alpha/2, n-1}$

Απορρ. την H_0

• $H_0: \mu = \mu_0$
• $H_1: \mu > \mu_0$

Αν $t > t_{\alpha, n-1}$

Απορρ. την H_0

• $H_0: \mu = \mu_0$
• $H_1: \mu < \mu_0$

Αν $t < -t_{\alpha, n-1}$

Απορρ. την H_0

● Εφαρμογή

Κανονικός πληθυσμός έχει μέση τιμή 20 και διασπορά 7. Ένα δείγμα μεγέθους 25 έχει μέση τιμή 21,5. Να γίνει έλεγχος υποθέσεων για το μέσο με σ.σ $\alpha = 0,04$

Λύση

Πληθ.	Δείγμα
$\mu_0 = 20$	$\bar{x} = 21,5$
$\sigma^2 = 7$	$n = 25$
$\sigma = \sqrt{7}$	

$$Z = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}} = \frac{21,5 - 20}{\sqrt{7} / 5} = 2,83$$

$$H_0: \mu = 20$$

$$H_1: \mu \neq 20$$

$$Z_{\alpha/2} = ?$$

$$\gamma = 1 - \alpha = 0,96$$

$$\rho = \frac{1 + 0,96}{2} = 0,98 \xrightarrow{\text{ΠΙΝΑΚΑΣ Z}} Z_{\alpha/2} = 2,06$$

$$|Z| > Z_{\alpha/2} \quad \text{Απορρίπτεται η } H_0$$

● Έλεγχος υπόθεσης για την διαφορά
μέσων $\mu_1 - \mu_2$ δύο ανεξάρτητων πληθυσμών
με γνωστές διασπορές σ_1^2 και σ_2^2

$$Z = \frac{\bar{X} - \bar{Y} - \delta}{\sqrt{\sigma_1^2/n + \sigma_2^2/m}}$$

$$\delta = \mu_1 - \mu_2$$

• $H_0: \mu_1 = \mu_2$
 $H_1: \mu_1 \neq \mu_2$

• Αν $|Z| > Z_{\alpha/2}$

Απορρ. την H_0

• $H_0: \mu_1 = \mu_2$
 $H_1: \mu_1 > \mu_2$

• Αν $Z > Z_\alpha$

Απορρ. την H_0

• $H_0: \mu_1 = \mu_2$
 $H_1: \mu_1 < \mu_2$

• Αν $Z < -Z_\alpha$

Απορρ. την H_0

● Έλεγχος υπόθεσης για την διαφορά
μέσων $\mu_1 - \mu_2$ δύο ανεξάρτητων πληθυσμών
με άγνωστες διασπορές αλλά ίσες

$$t = \frac{\bar{X} - \bar{Y} - \delta}{S_p \sqrt{1/n + 1/m}}$$

$$S_p^2 = \frac{(n-1)S_1^2 + (m-1)S_2^2}{n+m-2}$$

• $H_0: \mu_1 = \mu_2$
 $H_1: \mu_1 \neq \mu_2$

• Αν $|t| > t_{n+m-2, \alpha/2}$
 Απορρ. την H_0

• $H_0: \mu_1 = \mu_2$
 $H_1: \mu_1 > \mu_2$

• Αν $t > t_{n+m-2, \alpha}$
 Απορρ. την H_0

• $H_0: \mu_1 = \mu_2$
 $H_1: \mu_1 < \mu_2$

• Αν $t < t_{n+m-2, \alpha}$
 Απορρ. την H_0

● Έλεγχος υπόθεσης για την διαφορά
μέσων $\mu_1 - \mu_2$ δύο ανεξαρτήτων πλη-
θυσιών με άγνωστες διασπορές αλλά
όμοιες

$$t = \frac{\bar{x} - \bar{y} - \delta}{\sqrt{S_1^2/n + S_2^2/m}}$$



• $H_0: \mu_1 = \mu_2$
 $H_1: \mu_1 \neq \mu_2$

• $H_0: \mu_1 = \mu_2$
 $H_1: \mu_1 > \mu_2$

• $H_0: \mu_1 = \mu_2$
 $H_1: \mu_1 < \mu_2$

Αν $|t| > t_{1-\alpha/2}$

Αν $t > t_{1-\alpha}$

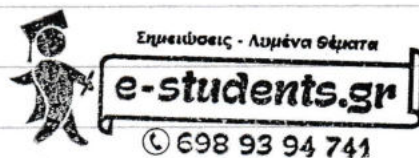
Αν $t < -t_{1-\alpha}$

Απορ. την H_0

Απορ. την H_0

Απορ. την H_0

όπου
$$v = \frac{\left(\frac{S_1^2}{n} + \frac{S_2^2}{m} \right)^2}{\frac{(S_1^2/n)^2}{n-1} + \frac{(S_2^2/m)^2}{m-1}}$$



Έλεγχος υπόθεσης για το πρσοστό p ενός πληθυσμού

$$Z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$$

$$\hat{p} = \frac{x}{N}$$

$$\begin{aligned} & \cdot H_0: p = p_0 \\ & H_1: p \neq p_0 \end{aligned}$$

$$\begin{aligned} & \cdot H_0: p = p_0 \\ & H_1: p > p_0 \end{aligned}$$

$$\begin{aligned} & \cdot H_0: p = p_0 \\ & H_1: p < p_0 \end{aligned}$$

$$\text{Αν } |Z| > Z_{\alpha/2}$$

$$\text{Αν } Z > Z_{\alpha}$$

$$\text{Αν } Z < -Z_{\alpha}$$

Απορ. την H_0

Απορ. την H_0

Απορ. την H_0

Έλεγχος υπόθεσης για την διαφορά

πρσοστών p_1, p_2 δύο ανεξάρτητων πληθυσμών

$$Z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n} + \frac{1}{m}\right)}}$$

$$\hat{p} = \frac{x+y}{n+m}$$

$$\begin{aligned} & \cdot H_0: p_1 = p_2 \\ & H_1: p_1 \neq p_2 \end{aligned}$$

$$\begin{aligned} & \cdot H_0: p_1 = p_2 \\ & H_1: p_1 > p_2 \end{aligned}$$

$$\begin{aligned} & \cdot H_0: p_1 = p_2 \\ & H_1: p_1 < p_2 \end{aligned}$$

$$\text{Αν } |Z| > Z_{\alpha/2}$$

$$\text{Αν } Z > Z_{\alpha}$$

$$\text{Αν } Z < -Z_{\alpha}$$

Απορ. την H_0

Απορ. την H_0

Απορ. την H_0



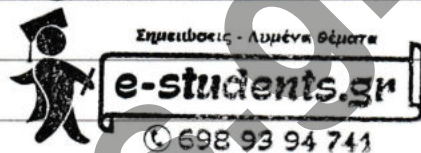
Εμπειροίς · Λυμένα θέματα

e-students.gr

© 698 93 94 741

Εφαρμογή

Είναι γνωστό ότι σε μια πόλη το ποσοστό των συνταξιούχων είναι 22%. Επιλέχουμε τυχαία 260 κατοίκους της πόλης και διαπιστώνουμε ότι ανάμεσα σε αυτούς υπάρχουν 62 συνταξιούχοι. Με σταθμην σημαντικότητας $\alpha=0,02$ είναι αληθές ότι οι συνταξιούχοι αποτελούν το 22% των κατοίκων της πόλης αυτής



Λύση

$$p_0 = 0,22$$

$$\begin{matrix} n=260 \\ x=62 \end{matrix} \Rightarrow \hat{p} = \frac{62}{260} = 0,238$$

$$H_0: p = 0,22$$

$$H_1: p \neq 0,22$$

$$Z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} = 0,7 \Rightarrow Z = 0,7$$



$$\alpha = 0,02 \Rightarrow \gamma = 1 - \alpha = 0,98$$

$$\rho = \frac{1 + 0,98}{2} = 0,99$$

$$Z_{\alpha/2} = 2,33 \quad |Z| < Z_{\alpha/2} \text{ ΔΕΥΤΗ Η } H_0$$

● Έλεγχος υπόθεσης για τη διασπορά

σ^2 ενός πληθυσμού

$$\chi^2 = \frac{(n-1)s^2}{\sigma_0^2}$$



Σημειώσεις - Λυμένα θέματα

e-students.gr

☎ 698 93 94 741

• $H_0: \sigma^2 = \sigma_0^2$
• $H_1: \sigma^2 \neq \sigma_0^2$

• $H_0: \sigma^2 = \sigma_0^2$
• $H_1: \sigma^2 > \sigma_0^2$

• $H_0: \sigma^2 = \sigma_0^2$
• $H_1: \sigma^2 < \sigma_0^2$

• Αν $\chi^2 > \chi_{n-1, 1-\alpha/2}^2$
ή $\chi^2 < \chi_{n-1, 1-\alpha/2}^2$

Απορρ. την H_0

• Αν $\chi^2 > \chi_{n-1, \alpha}^2$

Απορρ. την H_0

• Αν $\chi^2 < \chi_{n-1, 1-\alpha}^2$

Απορρ. την H_0

● Έλεγχος υπόθεσης για το λόγο των διασπορών σ_2^2/σ_1^2 δύο ανεξάρτητων

πληθυσμών

$$f = \frac{s_1^2}{s_2^2} \cdot \lambda_0$$



Σημειώσεις - Λυμένα θέματα

e-students.gr

☎ 698 93 94 741

• $H_0: \sigma_2^2/\sigma_1^2 = \lambda_0$

• $H_0: \sigma_2^2/\sigma_1^2 = \lambda_0$

• $H_0: \sigma_2^2/\sigma_1^2 = \lambda_0$

• $H_1: \sigma_2^2/\sigma_1^2 \neq \lambda_0$

• $H_1: \sigma_2^2/\sigma_1^2 < \lambda_0$

• $H_1: \sigma_2^2/\sigma_1^2 > \lambda_0$

• Αν $f < F_{n-1, m-1, 1-\alpha/2}$
ή $f > F_{n-1, m-1, \alpha/2}$

Απορρ. την H_0

• Αν $f > F_{n-1, m-1, \alpha}$

Απορρ. την H_0

• Αν $f < F_{n-1, m-1, 1-\alpha}$

Απορρ. την H_0

Έλεγχος χ^2



Σημειώσεις - Λυμένα Θέματα

e-students.gr

© 698 93 94 741

Ως στατιστική συνάρτηση έλεγχου (έλεγχοςυνάρτηση) χρησιμοποιείται η στατιστική συνάρτηση χ^2 και ορίζεται από τον τύπο

$$\chi^2 = \sum_{i=1}^k \frac{(O_i - E_i)^2}{E_i}$$

όπου $O_1, O_2, \dots, O_k$ οι παρατηρηθείσες συχνότητες

$E_1, E_2, \dots, E_k$ αναμενόμενες συχνότητες

$$\text{ώ}' E_i = N \cdot p_i$$



Σημειώσεις - Λυμένα Θέματα

e-students.gr

© 698 93 94 741

Σε επίπεδο σημαντικότητας α , η μηδενική υπόθεση απορρίπτεται

$$\text{όταν } \chi^2 = \sum_{i=1}^k \frac{(O_i - E_i)^2}{E_i} \geq \chi_{k-1, \alpha}^2$$

Τάρι ρίχνεται 120 φορές, και η πλευρά με $i = 1, 2, \dots, 6$ και εμφανίζεται 25, 17, 15, 23, 24 και 16 φορές αντίστοιχα. Χρησιμοποιώντας τα δεδομένα αυτά, τι συμπέρασμα βγάλετε για την τιμότητα του τάρου σε επίπεδο σημαντικότητας 5%.

Λύση

$$H_0: p = \frac{1}{6}$$

$$H_1: p \neq \frac{1}{6}$$

$$\chi^2 = \sum_{i=1}^6 \frac{(n_i - E_i)^2}{E_i}$$

$$= \frac{1}{120 \cdot \frac{1}{6}} \left[(25 - \frac{120}{6})^2 + \dots \right] = 5$$

$$\chi^2_{5, 0.05} = 11,1$$

$$\Rightarrow \chi^2 < \chi^2_{5, 0.05}$$

Άρα δεχόμαστε την H_0