MBA χίος
ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΙΓΑΙΟΥ

ΣΤΑΤΙΣΤΙΚΗ ΚΑΙ ΘΕΩΡΙΑ ΑΠΟΦΑΣΕΩΝ

ΤΥΠΟΛΟΓΙΟ

ΧΙΟΣ, 2020

ΑΡΙΘΜΗΤΙΚΕΣ ΜΕΘΟΔΟΙ ΣΥΝΟΨΗΣ ΔΕΔΟΜΕΝΩΝ

ΑΡΙΘΜΗΤΙΚΗ ΠΕΡΙΓΡΑΦΗ ΔΕΔΟΜΕΝΩΝ

ΑΤΑΞΙΝΟΜΗΤΑ ΔΕΔΟΜΕΝΑ	ΤΑΞΙΝΟΜΗΜΕΝΑ ΔΕΔΟΜΕΝΑ

ΜΕΤΡΑ ΚΕΝΤΡΙΚΗΣ ΤΑΣΗΣ

Αριθμητικός Μέσος

$\overline{X} = \dfrac{\sum_{i=1}^{n} X_i}{n}$	$\overline{X} \cong \dfrac{\sum_{i=1}^{k} f_i\, m_i}{n}$ $\overline{X} \cong m_0 + \delta \dfrac{\sum_{i=1}^{\kappa} f_i\, \xi_i}{n}$

Διάμεσος

• Αν *n* περιττός M: Η τιμή της παρατήρησης στη θέση $\dfrac{n+1}{2}$ δηλαδή η $\left(X_{\frac{n+1}{2}}\right)$ • Αν *n* άρτιος $M=\dfrac{1}{2}\left(X_{\frac{n}{2}+1}+X_{\frac{n}{2}}\right)$	$M=L_M+\delta\dfrac{\left(\frac{n}{2}-F_{M-1}\right)}{f_M}$

Επικρατούσα Τιμή

$$T_0 = L_{T_0} + \delta \frac{\Delta_1}{\Delta_1 + \Delta_2}$$

ΜΕΤΡΑ ΣΧΕΤΙΚΗΣ ΘΕΣΗΣ

i – Τεταρτημόριο

Q_i: Η τιμή στη θέση $\frac{i(n+1)}{4}$ δηλαδή η $\left(X_{\frac{i(n+1)}{4}}\right)$ $Q_i = X_{\frac{i(n+1)}{4}} = X_{A_Q} + \Delta_Q\left(X_{A_Q+1} - X_{A_Q}\right)$ Όπου A_Q: Το ακέραιο μέρος του πηλίκου $\frac{i(n+1)}{4}$ Δ_Q: Το δεκαδικό μέρος του πηλίκου $\frac{i(n+1)}{4}$	$Q_i = L_{Q_i} + \delta \frac{\left(\frac{ni}{4} - F_{Q_i-1}\right)}{f_{Q_i}}$

i – Δεκατημόριο

D_i: Η τιμή στη θέση $\frac{i(n+1)}{10}$ δηλαδή η $\left(X_{\frac{i(n+1)}{10}}\right)$ $D_i = X_{\frac{i(n+1)}{10}} = X_{A_D} + \Delta_D\left(X_{A_D+1} - X_{A_D}\right)$ Όπου A_D: Το ακέραιο μέρος του πηλίκου $\frac{i(n+1)}{10}$ Δ_D: Το δεκαδικό μέρος του πηλίκου $\frac{i(n+1)}{10}$	$D_i = L_{D_i} + \delta \frac{\left(\frac{ni}{10} - F_{D_i-1}\right)}{f_{D_i}}$

i – Ποσοστιαίο Σημείο

P_i: Η τιμή στη θέση $\frac{i(n+1)}{100}$ δηλαδή η $\left(X_{\frac{i(n+1)}{100}}\right)$ $P_i = X_{\frac{i(n+1)}{100}} = X_{A_P} + \Delta_P\left(X_{A_P+1} - X_{A_P}\right)$ Όπου A_P: Το ακέραιο μέρος του πηλίκου $\frac{i(n+1)}{100}$ Δ_P: Το δεκαδικό μέρος του πηλίκου $\frac{i(n+1)}{100}$	$P_i = L_{P_i} + \delta \frac{\left(\frac{ni}{100} - F_{P_i-1}\right)}{f_{P_i}}$

ΜΕΤΡΑ ΔΙΑΣΠΟΡΑΣ

Εύρος

$$R = X_{max} - X_{min}$$

Ενδοτεταρτημοριακό Εύρος

$Q_3 - Q_1$	$Q_3 - Q_1$

Τεταρτημοριακή Απόκλιση

$Q = \frac{Q_3 - Q_1}{2}$	$Q = \frac{Q_3 - Q_1}{2}$

Διακύμανση

$S^2 = \dfrac{\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2}{n-1}$	$\widehat{S}^2 \cong \dfrac{\sum_{i=1}^{k} f_i\left(m_i - \overline{X}\right)^2}{n-1}$
$S^2 = \dfrac{\sum_{i=1}^{n} X_i^2}{n-1} - \dfrac{\left(\sum_{i=1}^{n} X_i\right)^2}{(n-1)n}$	$S^2 = \dfrac{\sum_{i=1}^{k} f_i m_i^2}{n-1} - \dfrac{\left(\sum_{i=1}^{k} f_i m_i\right)^2}{(n-1)n}$
$S^2 = \dfrac{\sum_{i=1}^{n} X_i^2 - n\overline{X}^2}{n-1}$	$S^2 = \dfrac{\sum_{i=1}^{k} f_i m_i^2 - n\overline{X}^2}{n-1}$
	$\widehat{S}^2 \cong \delta^2 \left\{ \dfrac{\sum_{i=1}^{k} f_i \xi_1^2}{n-1} - \dfrac{\left(\sum_{i=1}^{k} f_i \xi_1^2\right)^2}{(n-1)n} \right\}$

Τυπική Απόκλιση

$s = +\sqrt{s^2}$	$s = +\sqrt{s^2}$

ΜΕΤΡΑ ΣΧΕΤΙΚΗΣ ΔΙΑΣΠΟΡΑΣ

Συντελεστής Μεταβλητότητας

$C = \dfrac{s}{\overline{\overline{X}}} * 100$	$C = \dfrac{s}{\overline{\overline{X}}} * 100$

ΜΕΤΡΑ ΑΣΥΜΜΕΤΡΙΑΣ

$S_p = \dfrac{\overline{X} - T_0}{\widehat{s}}$	$S_p = \dfrac{\overline{X} - T_0}{\widehat{s}}$

$S_p = \dfrac{3\left(\overline{X} - M\right)}{\widehat{s}}$	$S_p = \dfrac{3\left(\overline{X} - M\right)}{\widehat{s}}$

$S_B = \dfrac{(Q_3 - M) - (M - Q_1)}{Q_3 - Q_1}$	$S_B = \dfrac{(Q_3 - M) - (M - Q_1)}{Q_3 - Q_1}$

ΜΕΤΡΑ ΚΥΡΤΩΣΗΣ

$k = \dfrac{\dfrac{Q_3 - Q_1}{2}}{P_{90} - P_{10}}$	$k = \dfrac{\dfrac{Q_3 - Q_1}{2}}{P_{90} - P_{10}}$

$\beta_2 = \dfrac{\mu_4}{s^4} = \dfrac{\dfrac{\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^4}{n}}{s^4}$	$\beta_2 = \dfrac{\mu_4}{s^4} = \dfrac{\dfrac{\sum_{i=1}^{k} f_i\left(m_i - \overline{X}\right)^4}{n}}{s^4}$

ΜΕΤΡΑ ΣΥΓΚΕΝΤΡΩΣΗΣ

<u>Συντελεστής Gini</u>

$g = \dfrac{d}{2\overline{\overline{X}}}$ όπου: $d = \dfrac{\sum_i \sum_j \lvert d_{ij} \rvert}{n^2}$	$g = \dfrac{d}{2\overline{\overline{X}}}$ όπου: $d = \dfrac{2\delta}{n^2} \sum_{i=1}^{k} \left(n - \varphi_i\right) \varphi_i$ Φ_i η δεξιόστροφη αθροιστική συχνότητα της i τάξης

ΠΙΘΑΝΟΤΗΤΕΣ

Π Ι Θ Α Ν Ο Τ Η Τ Ε Σ

Δεσμευμένη Πιθανότητα

$$P(A_1|A_2) = \frac{P(A_1 \cap A_2)}{P(A_2)}, \quad P(A_2) > 0$$

$$P(A_2|A_1) = \frac{P(A_1 \cap A_2)}{P(A_1)}, \quad P(A_1) > 0$$

Θεώρημα της Ολικής Πιθανότητας

$$P(E) = \sum_{i=1}^{n} P(A_i)\, P(E|A_i)$$

Θεώρημα του Bayes

$$P(A_k|E) = \frac{P(A_k)P(E|A_k)}{\sum_{i=1}^{n} P(A_i)P(E|A_i)} = \frac{P(A_k)P(E|A_k)}{P(E)}$$

ΑΡΧΕΣ ΑΠΑΡΙΘΜΗΣΗΣ

Μεταθέσεις

$$P_n = n!$$

Επαναληπτικές Μεταθέσεις

$$\frac{n!}{n_1!\,n_2!\ldots n_k!}$$

Διατάξεις

$$\underset{n\ x}{P} = P(n,x) = \frac{n!}{(n-x)!}$$

Επαναληπτικές Διατάξεις

n^x

Συνδυασμοί

$$\underset{n\ x}{C} = C(n,x) = \frac{n!}{x!(n-x)!}$$

ΚΑΤΑΝΟΜΕΣ ΠΙΘΑΝΟΤΗΤΑΣ

	ΔΙΑΚΡΙΤΗ τυχαία μεταβλητή	**ΣΥΝΕΧΗΣ τυχαία μεταβλητή**

Μέτρα Κεντρικής Τάσης

Μέση Τιμή	$E(X)=\mu=$ $=\sum_{i=1}^{\infty} x_i P(x_i),$ $E(X)\langle\infty$	$E(X)=\mu=$ $=\int_{-\infty}^{+\infty} x f(x)dx,$ $E(X)\langle\infty$

Μέτρα Διασποράς

Διακύμανση	$V(X)=\sigma^2=E\left[X-E(X)\right]^2$ $=E(X-\mu)^2$ $=\sum_{i=1}^{n} x_i^2 P(x_i)-\mu^2$	$V(X)=\sigma^2=E\left[X-E(X)\right]^2$ $=E(X-\mu)^2$ $=\int_a^b x^2 f(x)dx-\mu^2$
Τυπική Απόκλιση	$\sigma=\sqrt{V(X)}$	$\sigma=\sqrt{V(X)}$

ΘΕΩΡΗΤΙΚΕΣ ΚΑΤΑΝΟΜΕΣ

ΘΕΩΡΗΤΙΚΕΣ ΚΑΤΑΝΟΜΕΣ

ΔΙΑΚΡΙΤΕΣ ΚΑΤΑΝΟΜΕΣ

Διωνυμική Κατανομή [X ~ B(n, p)]

$$P(X=k)=\binom{n}{k}p^k q^{n-k}=\frac{n!}{k!(n-k)!}p^k q^{n-k},$$

$$k=0,1,2,...,n \quad n=1,2,... \quad 0<p<1 \quad q=1-p$$

$$E(X)=np \qquad V(X)=npq$$

Κατανομή Poisson [X ~ P(λ)]

$$P(X=k)=\frac{e^{-\lambda}\lambda^k}{k!}, \quad k=0,1,2,... \quad \lambda>0$$

$E(X) = \lambda$ $\qquad$ $V(X) = \lambda$

ΣΥΝΕΧΕΙΣ ΚΑΤΑΝΟΜΕΣ

Κανονική Κατανομή [X ~ N(μ, σ²)]

$$f(x)=\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}, -\infty<x<+\infty$$

όπου $\pi = 3.1416$ & $e = 2.7183$

$-\infty < \mu < +\infty, \quad \sigma > 0$

$E(X) = \mu$ $\qquad$ $V(X) = \sigma^2$

Τυποποιημένη Κανονική Κατανομή $[\,Z=\frac{X-\mu}{\sigma} \sim N(0, 1)]$

$$f(z)=\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}z^2}, \quad -\infty<z<+\infty$$

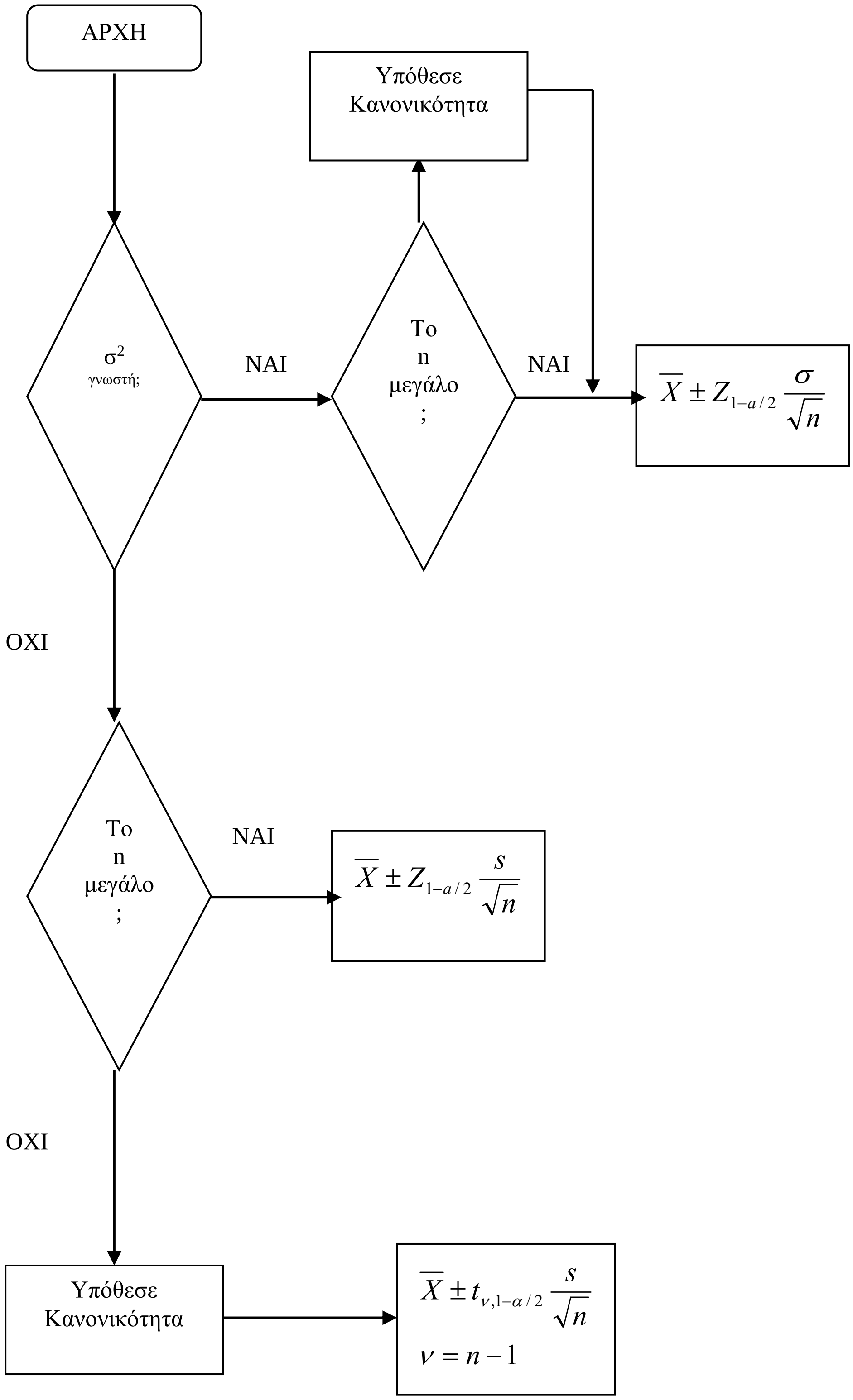
$E(X) = 0$ $\qquad$ $V(X) = 1.$

$$E(X)=\frac{\alpha}{\alpha+\beta} \qquad V(X)=\frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta-1)}.$$

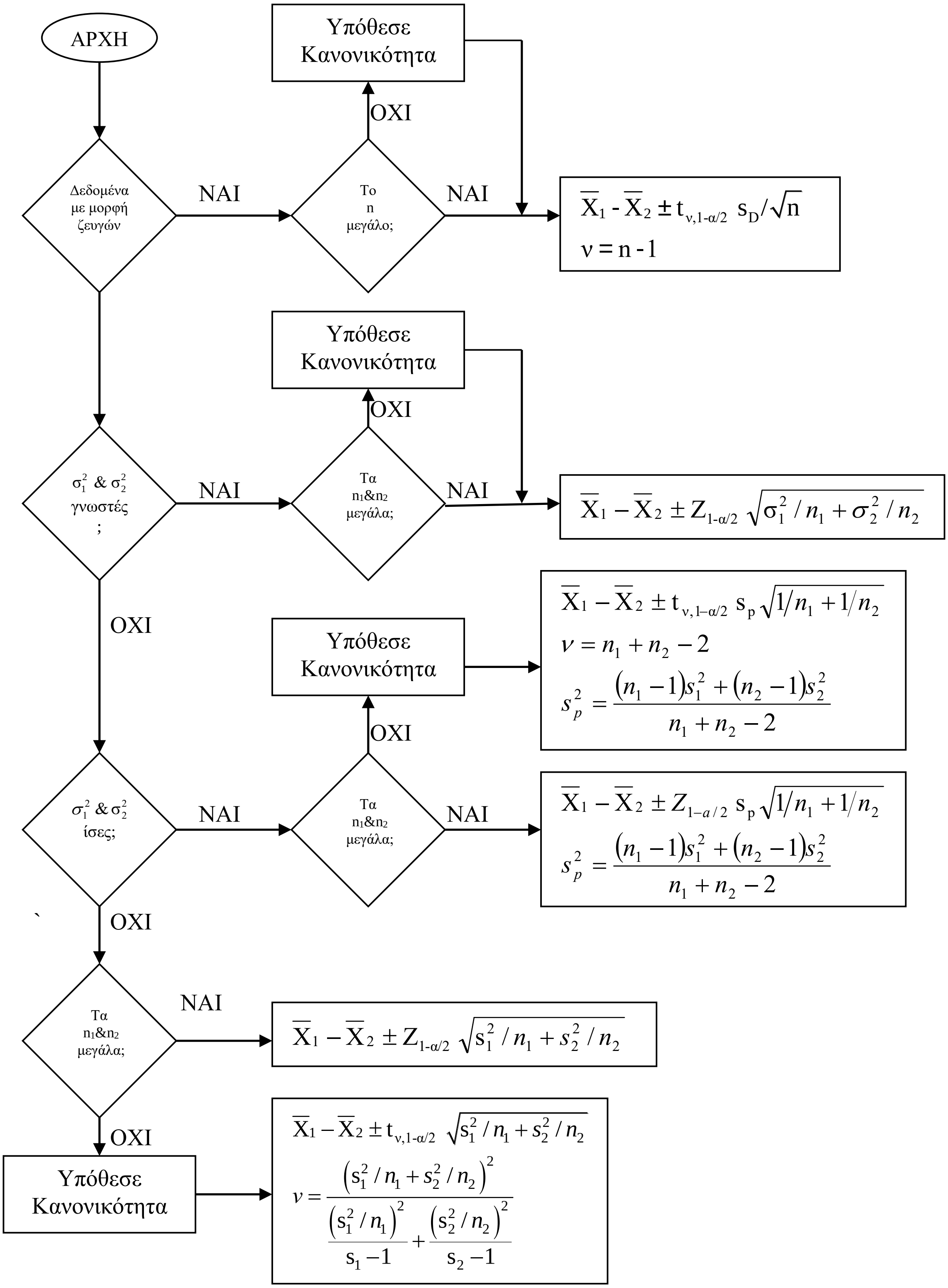
ΔΙΑΣΤΗΜΑΤΑ ΕΜΠΙΣΤΟΣΥΝΗΣ

ΕΚΤΙΜΗΣΗ ΠΑΡΑΜΕΤΡΩΝ ΠΛΗΘΥΣΜΟΥ ΜΕ ΔΙΑΣΤΗΜΑΤΑ ΕΜΠΙΣΤΟΣΥΝΗΣ

100 (1-α)% ΔΙΑΣΤΗΜΑΤΑ ΕΜΠΙΣΤΟΣΥΝΗΣ ΓΙΑ ΜΕΣΟΥΣ ΠΛΗΘΥΣΜΩΝ

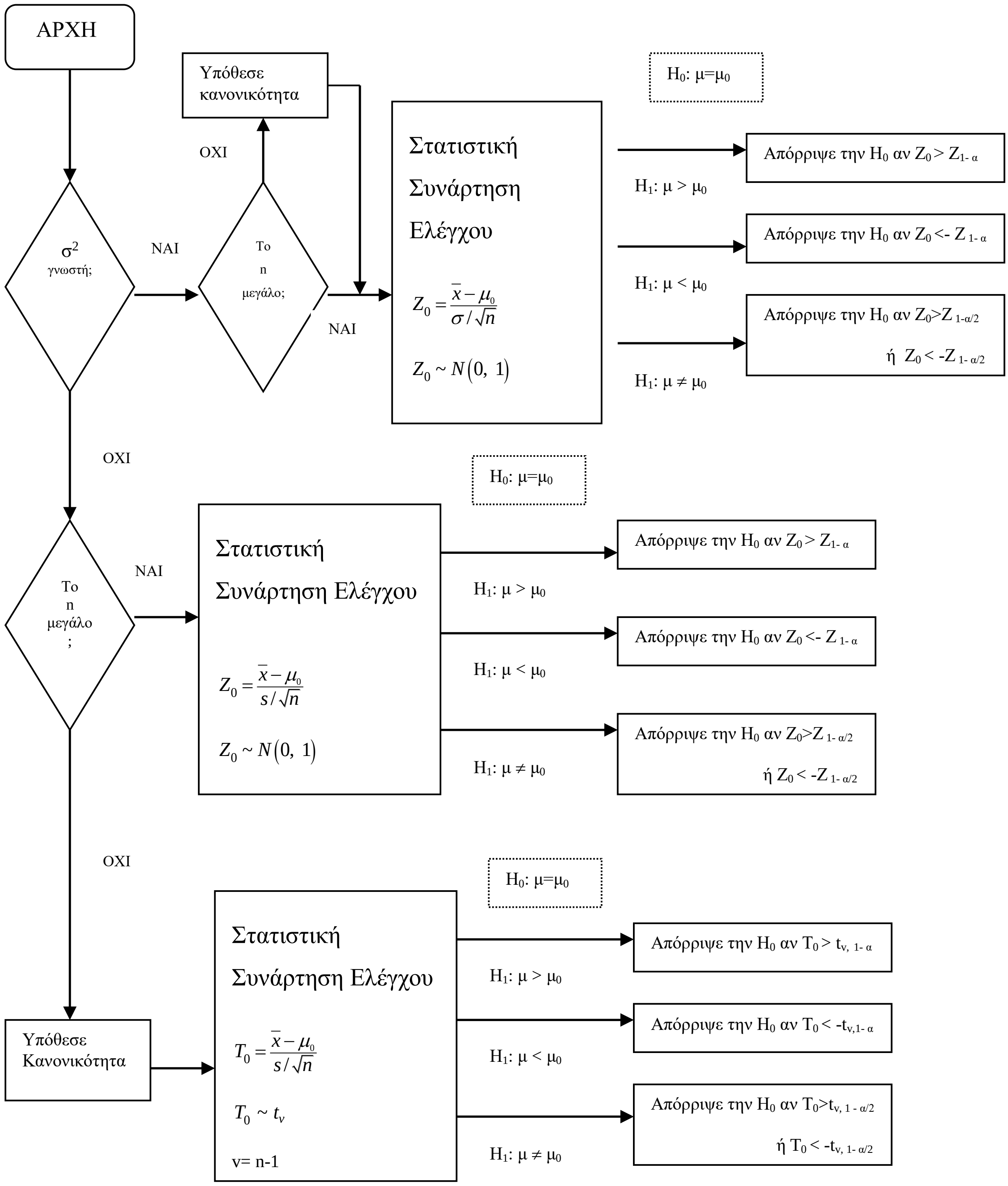


100 (1-α)% ΔΙΑΣΤΗΜΑΤΑ ΕΜΠΙΣΤΟΣΥΝΗΣ ΓΙΑ ΤΗ ΔΙΑΦΟΡΑ ΤΩΝ ΜΕΣΩΝ ΠΛΗΘΥΣΜΩΝ

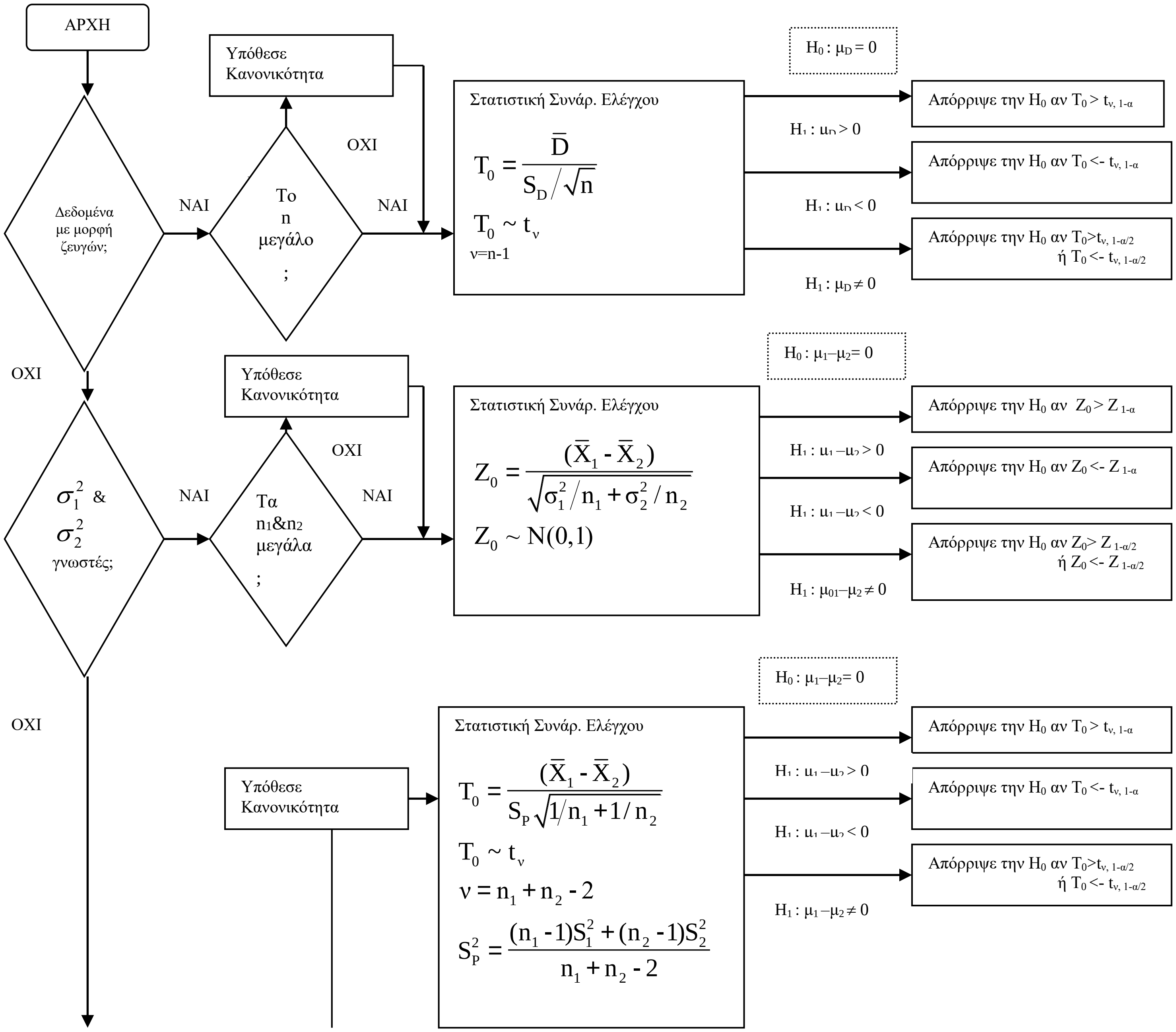


ΕΛΕΓΧΟΙ ΥΠΟΘΕΣΕΩΝ

ΠΑΡΑΜΕΤΡΙΚΟΙ ΕΛΕΓΧΟΙ ΣΤΑΤΙΣΤΙΚΩΝ ΥΠΟΘΕΣΕΩΝ ΓΙΑ ΜΕΣΟΥΣ ΠΛΗΘΥΣΜΩΝ



ΕΛΕΓΧΟΙ ΥΠΟΘΕΣΕΩΝ ΓΙΑ ΤΗ ΔΙΑΦΟΡΑ ΜΕΣΩΝ ΠΛΗΘΥΣΜΩΝ



σ_1^2 & σ_2^2 ίσες;

ΝΑΙ → Τα n_1&n_2 μεγάλα ;

ΟΧΙ

ΝΑΙ → Στατιστική Συνάρ. Ελέγχου

$$Z_0 = \frac{(\bar{X}_1 - \bar{X}_2)}{S_P\sqrt{1/n_1 + 1/n_2}}$$

$$Z_0 \sim N(0,1)$$

$$S_P^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}$$

$H_0: \mu_1 - \mu_2 = 0$

$H_1: \mu_1 - \mu_2 > 0$ → Απόρριψε την H_0 αν $Z_0 > Z_{1-\alpha}$

$H_1: \mu_1 - \mu_2 < 0$ → Απόρριψε την H_0 αν $Z_0 < - Z_{1-\alpha}$

$H_1: \mu_1 - \mu_2 \neq 0$ → Απόρριψε την H_0 αν $Z_0 > Z_{1-\alpha/2}$ ή $Z_0 < - Z_{1-\alpha/2}$

ΟΧΙ

Τα n_1&n_2 μεγάλα;

ΝΑΙ → Στατιστική Συνάρ. Ελέγχου

$$Z_0 = \frac{(\bar{X}_1 - \bar{X}_2)}{\sqrt{S_1^2/n_1 + S_2^2/n_2}}$$

$$Z_0 \sim N(0,1)$$

$H_0: \mu_1 - \mu_2 = 0$

$H_1: \mu_1 - \mu_2 > 0$ → Απόρριψε την H_0 αν $Z_0 > Z_{1-\alpha}$

$H_1: \mu_1 - \mu_2 < 0$ → Απόρριψε την H_0 αν $Z_0 < - Z_{1-\alpha}$

$H_1: \mu_1 - \mu_2 \neq 0$ → Απόρριψε την H_0 αν $Z_0 > Z_{1-\alpha/2}$ ή $Z_0 < - Z_{1-\alpha/2}$

ΟΧΙ

Υπόθεσε Κανονικότητα → Στατιστική Συνάρ. Ελέγχου

$$T_0 = \frac{(\bar{X}_1 - \bar{X}_2)}{\sqrt{S_1^2/n_1 + S_2^2/n_2}}$$

$$T_0 \sim t_\nu$$

$$\nu \approx \frac{(S_1^2/n_1 + S_2^2/n_2)^2}{\frac{(S_1^2/n_1)^2}{n_1 - 1} + \frac{(S_2^2/n_2)^2}{n_2 - 1}}$$

$H_0: \mu_1 - \mu_2 = 0$

$H_1: \mu_1 - \mu_2 > 0$ → Απόρριψε την H_0 αν $T_0 > t_{\nu,\,1-\alpha}$

$H_1: \mu_1 - \mu_2 < 0$ → Απόρριψε την H_0 αν $T_0 < - t_{\nu,\,1-\alpha}$

$H_1: \mu_1 - \mu_2 \neq 0$ → Απόρριψε την H_0 αν $T_0 > t_{\nu,\,1-\alpha/2}$ ή $T_0 < - t_{\nu,\,1-\alpha/2}$

ΑΝΑΛΥΣΗ ΔΙΑΚΥΜΑΝΣΗΣ

ΑΝΑΛΥΣΗ ΔΙΑΚΥΜΑΝΣΗΣ ΚΑΤΑ ΕΝΑΝ ΠΑΡΑΓΟΝΤΑ

Πίνακας Ανάλυσης Διακύμανσης

Πηγή Μεταβλητότητας	Αθροίσματα Τετραγώνων S S	Βαθμοί Ελευθερίας D F	Μέσα Τετραγωνικά Σφάλματα M S	F (κάτω από την H_0)
Μεταξύ αγωγών	SST_r	k-1	MSTr=SSTr/k-1	F_o=MSTr/MSE
Εντός των αγωγών (σφάλμα)	SSE	n-k	MSE=SSE/n-k	
Σύνολο	SST	n-1		

$$Y_{i.} = \sum_{j=1}^{n_i} Y_{ij}$$

$$Y_{..} = \sum_{i=1}^{k} Y_{i.} = \sum_{i=1}^{k}\sum_{j=1}^{n_i} Y_{ij}$$

$$\overline{Y}_{i.} = Y_{i.} / n_i$$

$$\overline{Y}_{..} = Y_{..} / n = \frac{\sum_{i=1}^{k}\sum_{j=1}^{n} Y_{ij}}{n}$$

SST=SSTr + SSE

$$SST = \sum_{i=1}^{k}\sum_{j=1}^{n_i}\left(Y_{ij} - \overline{Y}_{..}\right)^2 = \sum_{i=1}^{k}\sum_{j=1}^{n_i} Y_{ij}^2 - \frac{Y_{..}^2}{n}$$

$$SSTr = \sum_{i=1}^{k} n_i \left(\overline{Y}_{i.} - \overline{Y}_{..}\right)^2 = \sum_{i=1}^{k} \frac{Y_{i.}^2}{n_i} - \frac{Y_{..}^2}{n}$$

$$SSE = \sum_{i=1}^{k}\sum_{j=1}^{n_i}\left(Y_{ij} - \overline{Y}_{i.}\right)^2 = \sum_{i=1}^{k}\sum_{j=1}^{n_i} Y_{ij}^2 - \sum_{i=1}^{k}\frac{Y_{i.}^2}{n_i}$$

$$MSTr = \frac{SSTr}{k - 1}$$

$$MSE = \frac{SSE}{n - k}$$

<u>Έλεγχος της υπόθεσης ότι όλες οι αγωγές ασκούν την ίδια επίδραση:</u>

H_0: Όλοι οι μέσοι είναι ίσοι ($\mu_{1.} = \mu_{2.} = \ldots = \mu_{\kappa.}$)

H_1: Δύο τουλάχιστον μέσοι διαφέρουν ($\mu_{i.} \neq \mu_{j.}$, για $i \neq j$, i,j=1,2,…,k)

Συνάρτηση Ελέγχου: $F_o = \dfrac{MSTr}{MSE} \sim F_{k-1,n-k}$

Κριτήριο Ελέγχου: Απορρίπτουμε την H_0 αν: $F_o > F_{k-1,\, n-k,\, 1-\alpha}$

Ανάλυση Διακύμανσης Κατά Δυο Παράγοντες

<u>Ανάλυση Διακύμανσης κατά δύο Παράγοντες με μία παρατήρηση ανά κυψελίδα</u>

Πίνακας Ανάλυσης Διακύμανσης κατά Δύο Παράγοντες

με Μια Παρατήρηση ανά Κυψελίδα

Πηγή Μεταβλητότητας	SS	DF	MS	F (κάτω από την H_0)
Μεταξύ Αγωγών	SSTr	k-1	SSTr/(k-1)=MSTr	MSTr/MSE=F_{Tr}
Μεταξύ Blocks	SSBl	b-1	SSBl/(b-1)=MSBl	MSBl/MSE=F_{Bl}
Σφάλμα	SSE	kb-(k+b-1) = = (b-1)(k-1)	SSE/(b-1)(k-1)=MSE	
Σύνολο	SST	kb-1		

$$Y_{i.} = \sum_{r=1}^{b} Y_{ir} \qquad Y_{.r} = \sum_{i=1}^{k} Y_{ir}$$

$$\overline{Y}_{i.} = Y_{i.} / b \qquad \overline{Y_{.r}} = Y_{.r} / k$$

$$Y.. = \sum_{r=1}^{b} Y_{.r} = \sum_{i=1}^{k} Y_{i.} = \sum_{i=1}^{k}\sum_{r=1}^{b} Y_{ir}$$

$$\overline{Y}.. = Y.. / kb$$

SST = SSTr + SSBl + SSE

$$SST = \sum_{i=1}^{k}\sum_{r=1}^{b}\left(Y_{ir} - \overline{Y}..\right)^2 = \sum_{i=1}^{k}\sum_{r=1}^{b} Y_{ir}^2 - \frac{(Y..)^2}{kb}$$

$$SSTr = \sum_{i=1}^{k}\sum_{r=1}^{b}\left(Y_{ir} - \overline{Y}..\right)^2 = b\sum_{i=1}^{k}\left(\overline{Y}_{i.} - \overline{Y}..\right)^2 = \frac{\sum_i \left(Y_{i.}\right)^2}{b} - \frac{(Y..)^2}{kb}$$

$$SSBl = \sum_{i=1}^{k}\sum_{r=1}^{b}\left(\overline{Y}_{ir} - \overline{Y}..\right)^2 = k\sum_{r=1}^{b}(\overline{Y}_{ir} - \overline{Y}..)^2 = \frac{\sum_r (Y_{.r})^2}{k} - \frac{(Y..)^2}{kb}$$

$$SSE = \sum_{i=1}^{k}\sum_{r=1}^{b} Y_{ir}^2 - \frac{\sum_i \left(Y_{i.}\right)^2}{b} - \frac{\sum_r \left(Y_{.r}\right)^2}{k} + \frac{\left(Y..\right)^2}{kb}$$

$$MSTr = \frac{SSTr}{k - 1}$$

$$MSBl = \frac{SSBl}{b - 1}$$

$$MSE = \frac{SSE}{(k - 1)(b - 1)}$$

Έλεγχος της υπόθεσης ότι όλες οι αγωγές ασκούν την ίδια επίδραση:

H_0: Όλοι οι μέσοι είναι ίσοι ($\mu_{1.} = \mu_{2.} = \dots = \mu_{\kappa.}$)
H_1: Δύο τουλάχιστον από τους μέσους διαφέρουν ($\mu_{i.} \neq \mu_{j.}$, για $i \neq j$, i,j=1,2,…,k)

Συνάρτηση Ελέγχου: $F_o = \dfrac{MSTr}{MSE} \sim F_{k-1,(b-1)(k-1)}$

Κριτήριο Ελέγχου: Απορρίπτουμε την H_0 αν: $F_o > F_{k-1, (b-1)(k-1), 1-\alpha}$

Έλεγχος της υπόθεσης ότι όλα τα blocks ακούν την ίδια επίδραση:

H_0: Όλοι οι μέσοι είναι ίσοι ($\mu_{.1} = \mu_{.2} = \dots \mu_{.\kappa}$)
H_1: Δύο τουλάχιστον από τους μέσους διαφέρουν ($\mu_{.i} \neq \mu_{.j}$, για $i \neq j$, i,j=1,2,…,k)

Συνάρτηση Ελέγχου: $F_o = \dfrac{MSBl}{MSE} \sim F_{b-1,(b-1)(k-1)}$

Κριτήριο Ελέγχου: Απορρίπτουμε την H_0 αν: $F_0 > F_{b-1, (b-1)(k-1), 1-\alpha}$

Ανάλυση Διακύμανσης κατά δύο Παράγοντες με m παρατηρήσεις ανά κυψελίδα

Πίνακας Ανάλυσης Διακύμανσης

κατά δύο Παράγοντες με m Παρατηρήσεις ανά Κυψελίδα

Πηγή	DF	SS	MS	F
Μεταξύ Αγωγών	k-1	SSTr	MSTr=SSTr/k-1	F_{Tr}=MSTr/MSE
Μεταξύ Blocks	b-1	SSBl	MSBl=SSBl/b-1	F_{Bl}=MSBl/MSE
Αλληλοεπίδραση	(k-1)(b-1)	SSI	MSI=SSI/(k-1)(b-1)	F_I=MSI/MSE
Σφάλμα	kb(m-1)	SSE	MSE=SSE/kb(m-1)	
Σύνολο	kbm - 1	SST		

$$SST = \sum_{i=1}^{k}\sum_{r=1}^{b}\sum_{j=1}^{m}\left(Y_{irj} - \overline{Y}_{..}\right)^2 = \sum_{i=1}^{k}\sum_{r=1}^{b}\sum_{j=1}^{m} Y_{irj}^2 - \frac{\left(Y_{...}^2\right)}{kbm}$$

$$SSTr = mb\sum_{i=1}^{k}\left(\overline{Y}_{i..} - \overline{Y}_{...}\right)^2 = \frac{\sum_i \left(Y_{i..}\right)^2}{mb} - \frac{\left(Y_{...}\right)^2}{kmb}$$

$$SSBl = mk\sum_{r=1}^{b}\left(\overline{Y}_{.r.} - \overline{Y}_{...}\right)^2 = \frac{\sum_r \left(Y_{.r.}\right)^2}{mk} - \frac{\left(Y_{...}\right)^2}{kbm}$$

$$SSI = m\sum_{i=1}^{k}\sum_{j=1}^{b}\left(\overline{Y}_{ir.} - \overline{Y}_{i..} - \overline{Y}_{...}\right)^2 = \frac{1}{m}\sum_{i=1}^{k}\sum_{r=1}^{b} Y_{ir.}^2 - \frac{1}{mk}\sum_{r=1}^{b} Y_{.r.}^2 - \frac{1}{mb}\sum_{i=1}^{k} Y_{i..}^2 + \frac{\left(Y_{...}\right)^2}{kbm}$$

$$SSE = \sum_{i=1}^{k}\sum_{r=1}^{b}\sum_{j=1}^{m} Y_{irj}^2 - \sum_{i=1}^{k}\sum_{j=1}^{b}\frac{\left(Y_{ir.}\right)^2}{m} = \sum_{i=1}^{k}\sum_{r=1}^{b}\sum_{j=1}^{m}\left(Y_{irj} - \overline{Y}_{ir.}\right)^2 = Y_{irj}^2 - \frac{\left(Y_{ir.}\right)^2}{m}$$

$$MSTr = \frac{SSTr}{k-1}$$

$$MSBl = \frac{SSBl}{b-1}$$

$$MSI = \frac{SSI}{(b-1)(k-1)}$$

$$MSE = \frac{SSE}{kb(m-1)}$$

<u>Έλεγχος της υπόθεσης ότι όλες οι αγωγές ασκούν την ίδια επίδραση:</u>

H_0: Όλοι οι μέσοι είναι ίσοι ($\mu_{1..} = \mu_{2..} = \ldots = \mu_{\kappa..}$)
H_1: Δύο τουλάχιστον από τους μέσους διαφέρουν ($\mu_{i..} \neq \mu_{j..}$, για $i \neq j$,
$i,j = 1,2,\ldots,k$)

Συνάρτηση Ελέγχου: $F_{Tr} = \dfrac{MSTr}{MSE}$

Κριτήριο Ελέγχου: Απορρίπτουμε την H_0 αν: $F_{Tr} > F_{k-1,\ kb(m-1),\ 1-\alpha}$

<u>Έλεγχος της υπόθεσης ότι όλα τα blocks ασκούν την ίδια επίδραση:</u>

H_0: Όλοι οι μέσοι είναι ίσοι ($\mu_{.1.} = \mu_{.2.} = \ldots = \mu_{.\kappa.}$)
H_1: Δύο τουλάχιστον από τους μέσους διαφέρουν ($\mu_{.i.} \neq \mu_{.j.}$, για $i \neq j$, i,j=1,2,…,k)

Συνάρτηση Ελέγχου: $F_{Bl} = \dfrac{MSBl}{MSE}$

Κριτήριο Ελέγχου: Απορρίπτουμε την H_0 αν: $F_{Bl} > F_{b-1,\ kb(m-1),\ 1-\alpha}$

<u>Έλεγχος της υπόθεσης ότι δεν υπάρχει αλληλεπίδραση αγωγών και blocks:</u>

H_0: Δεν υπάρχει αλληλεπίδραση αγωγών και blocks
H_1: Υπάρχει αλληλεπίδραση αγωγών και blocks

Συνάρτηση Ελέγχου: $F_I = \dfrac{MSI}{MSE}$

Κριτήριο Ελέγχου: Απορρίπτουμε την H_0 αν: $F_I > F_{(k-1)(b-1),\ kb(m-1),\ 1-\alpha}$

ΓΡΑΜΜΙΚΑ ΜΟΝΤΕΛΑ

ΑΠΛΗ ΓΡΑΜΜΙΚΗ ΠΑΛΙΝΔΡΟΜΗΣΗ

Εκτίμηση των παραμέτρων του μοντέλου της απλής γραμμικής παλινδρόμησης:

$$Y = \alpha + \beta X + \varepsilon$$
$$\hat{Y} = \hat{\alpha} + \hat{\beta} X$$
$$\hat{y} = a + bx$$

$$a = \frac{\left(\sum X_i^2\right)\left(\sum Y_i\right) - \left(\sum X_i\right)\left(\sum X_i Y_i\right)}{n\left(\sum X_i^2\right) - \left(\sum X_i\right)^2} = \bar{Y} - b\bar{X} = \bar{Y} - \frac{S_{XY}}{S_{XX}}\bar{X}$$

$$b = \frac{n\left(\sum X_i Y_i\right) - \left(\sum X_i\right)\left(\sum Y_i\right)}{n\left(\sum X_i^2\right) - \left(\sum X_i\right)^2} = \frac{\sum_i \left(X_i - \bar{X}\right)\left(Y_i - \bar{Y}\right)}{\sum_i \left(X_i - \bar{X}\right)^2} = \frac{S_{XY}}{S_{XX}}$$

όπου:

$$\bar{Y} = \frac{1}{n}\sum_{i=1}^{n} Y_i \quad \text{και} \quad \bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$$

$$S_{XX} = \sum_i \left(X_i - \bar{X}\right)^2 = \sum_i X_i^2 - \frac{\left(\sum X_i\right)^2}{n}$$

$$S_{YY} = \sum_i \left(Y_i - \bar{Y}\right)^2 = \sum_i Y_i^2 - \frac{\left(\sum Y_i\right)^2}{n}$$

$$S_{XY} = \sum_i \left(X_i - \bar{X}\right)\left(Y_i - \bar{Y}\right) = \sum_i \left(X_i - \bar{X}\right) Y_i = \sum_i X_i Y_i - \frac{\left(\sum X_i\right)\left(\sum Y_i\right)}{n}$$

Στατιστική Συμπερασματολογία στην Απλή Γραμμική Παλινδρόμηση

Εκτίμηση της Διακύμανσης των τυχαίων σφαλμάτων

$$Y_i = \alpha + \beta X_i + \varepsilon_i$$

$$\hat{Y} = \hat{\alpha} + \hat{\beta} X$$

$$y_i = a + bx_i$$

$$\hat{\sigma}^2 \equiv S^2{}_{Y/X} = \frac{\sum\left(Y_i - \hat{Y}_i\right)^2}{n-2} = \frac{1}{n-2}\left(S_{YY} - \frac{S^2{}_{XY}}{S_{XX}}\right)$$

Διαστήματα Εμπιστοσύνης για τις παραμέτρους του μοντέλου της Απλής Γραμμικής Παλινδρόμησης

Διάστημα Εμπιστοσύνης για το β

$$\hat{\beta} \pm t_{n-2,1-a/2} S_{\hat{\beta}}$$

$$S_{\hat{\beta}} = S_{Y/X} / \sqrt{S_{XX}}$$

Διάστημα Εμπιστοσύνης για το α

$$\widehat{\alpha} \pm t_{n-2,1-a/2} S_{\hat{\alpha}}$$

$$S_{\hat{\alpha}} = S_{Y/X}\sqrt{\frac{1}{n} + \frac{\bar{X}^2}{S_{XX}}}$$

Διάστημα Εμπιστοσύνης για την $E\left(Y / X_t\right)$

$$\widehat{Y} \pm t_{n-2,1-a/2} S_{\hat{Y}_t / X_t}$$

$$S_{\hat{Y}_t / X_t} = S_{Y/X_t}\sqrt{\left\{\frac{1}{n} + \frac{(X_t - \bar{X})^2}{S_{XX}}\right\}}$$

Διάστημα Πρόβλεψης

$$\widehat{Y}_k \pm t_{n-2,1-a/2} S_{\hat{Y}_k / X_k}$$

$$S_{\hat{Y}_k / x_k} = S_{Y/X} \sqrt{\left\{1 + \frac{1}{n} + \frac{\left(X_k - \bar{X}\right)^2}{S_{XX}}\right\}}.$$

Έλεγχοι Στατιστικών Υποθέσεων για τις παραμέτρους του μοντέλου:

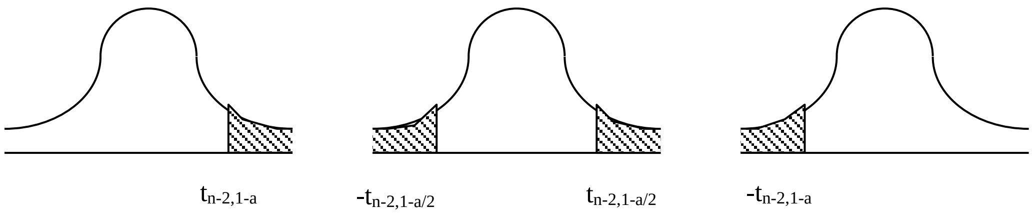
Έλεγχοι Υποθέσεων για το β

$H_o : \beta = \beta_o$	$H_o : \beta = \beta_o$	$H_o : \beta = \beta_o$
$H_1 : \beta > \beta o$	$H_1 : \beta \neq \beta o$	$H_1 : \beta < \beta o$

Στατιστική Συνάρτηση Ελέγχου: $$T_o = \frac{\widehat{\beta} - \beta_o}{S_{Y/X} \Big/ \sqrt{S_{XX}}} \sim t_{n-2}$$

Κριτήριο Απόρριψης: Απορρίπτουμε την Ho αν:

$T_0 > t_{n-2,1-a}$	$T_0 > t_{n-2,1-a/2}$ ή $T_0 < -t_{n-2,1-a/2}$	$T_0 < -t_{n-2,1-\alpha}$



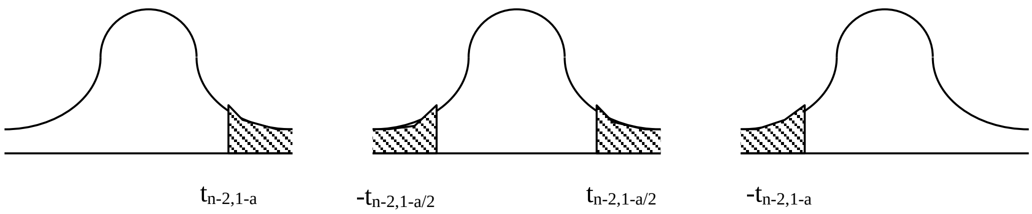
Έλεγχοι Υποθέσεων για το α

$H_o : \alpha = \alpha_o$	$H_o : \alpha = \alpha_o$	$H_o : \alpha = \alpha_o$
$H_1 : \alpha > \alpha_o$	$H_1 : \alpha \neq \alpha_o$	$H_1 : \alpha < \alpha_o$

Στατιστική Συνάρτηση Ελέγχου: $$T_o = \frac{\widehat{\alpha} - \alpha_o}{S_{Y/X} \sqrt{\frac{1}{n} + \frac{\bar{X}^2}{S_{XX}}}} \sim t_{n-2}$$

Κριτήριο Απόρριψης: Απορρίπτουμε την Ho αν:

$T_0 > t_{n-2,1-a}$	$T_0 > t_{n-2,1-a/2}$ ή $T_0 < -t_{n-2,1-a/2}$	$T_0 < -t_{n-2,1-\alpha}$



ΣΥΣΧΕΤΙΣΗ

Συντελεστής Προσδιορισμού

$$R^2=\frac{\sum\left(\hat{Y}_i-\bar{Y}\right)^2}{\sum\left(Y_i-\bar{Y}\right)^2}=\frac{b^2\sum\left(X_i-\bar{X}\right)^2}{S_{yy}}=b^2\frac{S_{xx}}{S_{yy}}=b\frac{S_{xy}}{S_{yy}}=\frac{S_{xy}^2}{S_{xx}S_{yy}}$$

$$R^2{}_{adj}=1-\left[\left(1-R^2\right)\frac{n-1}{n-2}\right]$$

Συντελεστής Συσχέτισης

$$\rho = \frac{Cov(X,Y)}{\sigma_X \sigma_Y}$$

$$\hat{\rho} = \frac{\sum (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum (X_i - \bar{X})^2 (Y_i - \bar{Y})^2}}$$

$$r = \frac{\sum (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum (X_i - \bar{X})^2}\sqrt{\sum (Y_i - \bar{Y})^2}} = \frac{\sum (X_i Y_i - n\bar{X}\bar{Y})}{\sqrt{\left(\sum X_i^2 - n\bar{X}^2\right)}\sqrt{\left(\sum Y_i^2 - n\bar{Y}^2\right)}} = \frac{S_{XY}}{\sqrt{S_{XX}}\sqrt{S_{YY}}}$$

Έλεγχοι Υποθέσεων για το Συντελεστή Συσχέτισης

Ακριβής Έλεγχος

$H_o : \rho = \rho_o$	$H_o : \rho = \rho_o$	$H_o : \rho = \rho_o$
$H_1 : \rho > \rho_o$	$H_1 : \rho \neq \rho_o$	$H_1 : \rho < \rho_o$

Στατιστική Συνάρτηση Ελέγχου: $T_o = r\sqrt{\frac{n-2}{1-r^2}} \sim t_{n-2}$

Κριτήριο Απόρριψης: Απορρίπτουμε την Ho αν:

$t_0 > t_{n-2,1-a}$	$t_0 > t_{n-2,1-a/2}$ ή $t_0 < -t_{n-2,1-a/2}$	$t_0 < -t_{n-2,1-\alpha}$