

4ο online καθίσμα
 Εξαγωγή στην Στατιστική
 3/4/2020

(12)

Άσκηση 11^η: Αν X_1, \dots, X_n τ.δ. από $N(\mu, \sigma^2)$ να βρεθούν:

- (i) η κατανομή $X_i - \bar{X}$ για i σταθερό
- (ii) η κατανομή $n(\bar{X} - \mu)^2 / \sigma^2$
- (iii) η κρίση τιμή του $S^2 = \frac{\sum (X_i - \bar{X})^2}{n}$
- (iv) η κατανομή $\sum \frac{(X_i - \mu)^2}{\sigma^2}$

Λύση: (i) $X_i - \bar{X}$ θα πρέπει να γραφεί ως γραμμικός συνδυασμός ανεξάρτητων κανονικών

$$X_i - \bar{X} = X_i - \frac{1}{n} \sum_{i=1}^n X_i \Rightarrow X_i - \bar{X} = X_i - \left(\frac{1}{n} \dots \frac{1}{n} \right) \begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{pmatrix}$$

$$X_i - \bar{X} = \left(0 \dots 0 \underset{\substack{\uparrow \\ i\text{-οστή} \\ \text{θέση}}}{1} 0 \dots 0 \right) \begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix} - \left(\frac{1}{n} \dots \frac{1}{n} \right) \begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix}$$

Ανάλυση

$$X_i - \bar{X} = \left(-\frac{1}{n} \dots -\frac{1}{n} \underset{\substack{\uparrow \\ i\text{-οστή} \\ \text{θέση}}}{1 - \frac{1}{n}} -\frac{1}{n} \dots -\frac{1}{n} \right) \begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix}$$

Σημείωση $X_i - \bar{X} = \sum_{j=1}^n a_j X_j$ με $a_j = \begin{cases} -\frac{1}{n}, & j \neq i \\ 1 - \frac{1}{n}, & j = i \end{cases}$

Επομένως $X_i - \bar{X}$ ως γραμμικός συνδυασμός των $X_j, j=1, \dots, n$ ακολουθεί κανονική κατανομή με κρίση τιμή

$$-\frac{1}{n}(n-1) \cdot \mu + \left(1 - \frac{1}{n}\right) \mu = 0 \quad E(\sum a_j X_j) = \sum a_j E(X_j) = \mu \sum a_j$$

$$\text{Var}(\sum a_j X_j) = \sum \text{Var}(a_j X_j) = \sum a_j^2 \sigma^2$$

και διακύμανση $\sum_{j=1}^n a_j^2 \sigma^2 = \sigma^2 \left[\frac{1}{n^2}(n-1) + \left(1 - \frac{1}{n}\right)^2 \right] = \frac{\sigma^2(n-1)}{n}$

(ii) $\bar{X} \sim N(\mu, \sigma^2/n) \Rightarrow \frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1) \Rightarrow$

$$\frac{(\bar{X}-\mu)^2}{\sigma^2/n} \sim \chi^2_1$$

(iii) Είναι γνωστό ότι (υπάρχουν κι άλλοι τρόποι/αυτός είναι ο πιο εύκολος)

$$\frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{n-1} \quad \text{ή} \quad \frac{\cancel{(n-1)} \frac{1}{\cancel{n-1}} \sum_{i=1}^n (X_i - \bar{X})^2}{\sigma^2} \sim \chi^2_{n-1}$$

Άρα $\frac{\sum_{i=1}^n (X_i - \bar{X})^2}{\sigma^2} \sim \chi^2_{n-1}$ $E(\chi^2_n) = n$
 $Var(\chi^2_n) = 2n$

και $E\left[\frac{\sum_{i=1}^n (X_i - \bar{X})^2}{\sigma^2}\right] = (n-1)$

Άρα $E\left[\sum (X_i - \bar{X})^2\right] = (n-1)\sigma^2$

$$\Rightarrow E\left[\frac{\sum (X_i - \bar{X})^2}{n}\right] = \frac{n-1}{n} \sigma^2$$

(iv) $X_i \sim N(\mu, \sigma^2) \Rightarrow \frac{X_i - \mu}{\sigma} \sim N(0,1)$

$$\Rightarrow \frac{(X_i - \mu)^2}{\sigma^2} \sim \chi^2_1 \Rightarrow \sum_{i=1}^n \frac{(X_i - \mu)^2}{\sigma^2} \sim \chi^2_n$$

ΠΡΟΣΟΧΗ

↓
Ασκήση 12^η: Αν X_1, X_2 τ.δ. από μια τυπική κανονική κατανομή $N(0,1)$ να βρεθούν οι κατανομές των στατιστικών συναρτήσεων

(i) $(X_2 - X_1) / \sqrt{2}$

(ii) $\frac{(X_1 + X_2)^2}{(X_2 - X_1)^2}$

(iii) $\frac{X_1 + X_2}{\sqrt{(X_2 - X_1)^2}}$ και $X_2^2 | X_1^2$

Λύση: (i) $X_1 \sim N(0,1)$ και $X_2 \sim N(0,1)$

Τότε $X_2 - X_1$ ως γραμμικός συνδυασμός ακολουθεί $N(0,2)$ (γιατί?)

Άρα $\frac{X_2 - X_1}{\sqrt{2}} \sim N(0,1)$ και αποδείχθηκε.
 $E(X_2 - X_1) = E(X_2) - E(X_1) = 0$
 $Var(X_2 - X_1) = Var(X_2) + Var(X_1) = 2$
επειδή X_2, X_1 ανεξάρτητες.

$(X_2 - X_1) = \sum_{i=1}^2 a_i X_i$
↓
 $a_1 = -1, a_2 = 1$

(ii) Όμοια με το (i) ερώτημα ($X_i \sim N(0,1)$)

$X_1 + X_2 \sim N(0,2)$ ως γραμμικός συνδυασμός κανονικών είναι κανονική

$\Rightarrow \frac{X_1 + X_2}{\sqrt{2}} \sim N(0,1)$

$\Rightarrow \frac{(X_1 + X_2)^2}{2} \sim X_1^2$

Οπότε και $\frac{(X_2 - X_1)^2}{2} \sim X_1^2$

Επομένως θα έλεγε κάποιος

$$\frac{\frac{(X_1 + X_2)^2}{2} \mid 1}{\frac{(X_2 - X_1)^2}{2} \mid 1} \sim F_{1,1}$$

(Εάν σταματούσα την άσκηση εδώ, δεν θα είπαμε όλα τα κομμάτια της άσκησης.)

(15)

Όπως αυτό ισχύει υπό μια ΒΑΣΙΚΗ ΥΠΟΘΕΣΗ. Ποια;
 $X_1 + X_2$, $X_2 - X_1$ ανεξάρτητες τ.μ. και τότε θα ισχύει.

Πώς θα το δείχνω αυτό;

Στα πλαίσια αυτού του μαθήματος που μας "λείπουν" διαφορές γνώσεις και αποτελέσματα θα δούμε έναν τρόπο.

Θα προσπαθήσουμε να γράψουμε το διάνυσμα:

$$\begin{pmatrix} X_1 + X_2 \\ X_2 - X_1 \end{pmatrix}_{2 \times 1} = A \begin{pmatrix} X_1 \\ X_2 \end{pmatrix}_{2 \times 1}$$

Εδώ είναι $A = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$

Τότε στον επόμενο πίνακα θα υπολογίσω τον πίνακα

$$A \begin{pmatrix} \text{Var } X_1 & 0 \\ 0 & \text{Var } X_2 \end{pmatrix} A^t$$

Εδώ είναι

$$\begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} \rightarrow \text{αν τα στοιχεία εκτός της διαγωνίου είναι 0 τότε ανεξάρτητα}$$

$$\left. \begin{array}{l} \text{(iii)} \quad X_1 + X_2 \sim N(0, 2) \\ \quad \quad X_2 - X_1 \sim N(0, 1) \end{array} \right\} \begin{array}{l} \text{ανεξάρτητεια} \\ \text{υπάρχει από (ii)} \end{array} \quad \left(\frac{X_2 - X_1}{\sqrt{2}} \right)^2 \sim \chi^2_1$$

$$\frac{\frac{X_1 + X_2}{\sqrt{2}}}{\sqrt{\frac{(X_2 - X_1)^2}{2} \mid 1}} \sim t_1$$

Επίσης $\frac{\chi^2_2 \mid 1}{\chi^2_1 \mid 1} \sim F_{1,1}$

(16)

Άσκηση 13η: Αν X_1, X_2, X_3 ανεξάρτητες τ.β. $N(0,1)$ να υπολογιστεί
 συνάρτηση της α.β.κ. γνωστής κατανομής

(i) $P(X_1^2 + X_2^2 + X_3^2 < 1)$

(ii) η κατανομή της $2X_1^2 / (X_2^2 + X_3^2)$

Λύση: $X_1^2 + X_2^2 + X_3^2 \sim \chi_3^2$

Άρα

(i) $P(X_1^2 + X_2^2 + X_3^2 \leq 1) = F_{\chi_3^2}(1)$

(ii) $X_1^2 \sim \chi_1^2$
 $X_2^2 + X_3^2 \sim \chi_2^2$ } ανεξάρτητες

$$\frac{X_1^2 / 1}{(X_2^2 + X_3^2) / 2} \sim F_{1,2}$$

Άσκηση 14η: Έστω $X_i, i=1, \dots, 5$ ανεξάρτητες $N(0,1)$

(i) Να βρεθεί η μέση τιμή και τυπική απόκλιση

$$X_1^2 + \dots + X_5^2$$

(ii) Να βρεθεί η c έτσι ώστε

$$P(-c \leq 2X_5 / \sqrt{X_1^2 + \dots + X_4^2} \leq c) = 0.9$$

→ Ουβερπικό διάστημα.

Λύση: (i) $X_1^2 + \dots + X_5^2 \sim \chi_5^2$

$$E(X_n^2) = n, \quad \text{Var}(X_n^2) = 2n$$

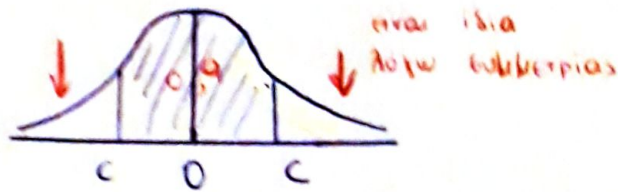
Άρα $E(X_1^2 + \dots + X_5^2) = 5, \quad \sigma = \sqrt{2 \cdot 5}$

(ii) $X_1^2 + \dots + X_4^2 \sim \chi_4^2$

$$X_5 \sim N(0,1)$$

Άρα $\frac{X_5}{\sqrt{(X_1^2 + \dots + X_4^2) / 4}} \sim t_4$ ή $\frac{2X_5}{\sqrt{X_1^2 + \dots + X_4^2}} \sim t_4$.

$$P(-c < t_4 < c) = 0.90 \quad (17)$$



$$2\alpha = 0.1$$

$$\alpha = 0.05$$

Άρα

$$P(t_4 \geq c) = 0.05$$

$$c = t_{0.05, 4} = 2.132$$

Άσκηση 15^η: Έστω S^2 η διακύμανση τ.δ. $n=10$ από $N(\mu=0, \sigma^2=4)$

Βρείτε την $\text{Var}(S^2)$. Σημειώση από κανονικό πληθυσμό.

$$\text{Λύση: } \frac{(n-1)S^2}{\sigma^2} \sim \chi_{n-1}^2 \Rightarrow 9 \frac{S^2}{4} \sim \chi_9^2 \quad \text{Var}(\chi_n^2) = 2n \Rightarrow$$

$$\text{Var}\left(9 \frac{S^2}{4}\right) = 2 \cdot 9 \Rightarrow \frac{9^2}{4^2} \text{Var} S^2 = 2 \cdot 9 \Rightarrow \text{Var} S^2 = \frac{32}{9}$$

Άσκηση 16^η: Να βρεθεί η κατανομή της $\sum X_i^2 / 6^2$ όταν

X_1, \dots, X_n τ.δ. $N(0, 6^2)$

$$\text{Λύση: } X_i \sim N(0, 6^2) \Rightarrow \frac{X_i}{6} \sim N(0, 1)$$

$$\Rightarrow \frac{X_i^2}{6^2} \sim \chi_1^2 \Rightarrow \sum_{i=1}^n \frac{X_i^2}{6^2} \sim \chi_n^2$$

Άσκηση 2 βελ. 311 (από το βιβλίο)

(1)

Άσκηση 17ⁿ: Δίνεται τ.δ. X_1, X_2, X_3, X_4 από πληθυσμό $N(0,1)$.

Αν $Y_1 = \frac{X_1 + X_2}{\sqrt{(X_1 - X_2)^2}}$, $Y_2 = \frac{\sum X_i^2}{X_1^2 + X_2^2}$ και $Y_3 = X_4 - \bar{X}$

(i) Να βρεθούν οι σταθερές c_1, c_2 έτσι ώστε $P(Y_1 \leq c_1) = 0.95$, $P(Y_2 \geq c_2) = 0.95$

(ii) Να υπολογιστεί η $P(Y_3 \geq \sqrt{3})$

Λύση: X_1, X_2, X_3, X_4 τ.δ. $N(0,1)$

Άρα $X_1 + X_2 \sim N(0,2)$

Επίσης $X_1 - X_2 \sim N(0,2)$

Σχεφτόμαι ότι $\frac{X_1 + X_2}{\sqrt{2}} \sim N(0,1)$ $Var(X_1 + X_2) = Var X_1 + Var X_2$

$X_1 + X_2$ ως χρ. συνδυασμός ανεξάρτητων μεταβλητών από κανονικό πληθυσμό ακολουθεί κανονική κατανομή

$E(X_1 + X_2) = E(X_1) + E(X_2) = 0$

και $\frac{X_1 - X_2}{\sqrt{2}} \sim N(0,1) \Rightarrow \frac{(X_1 - X_2)^2}{2} \sim \chi_1^2$

Επομένως: $\frac{X_1 + X_2}{\sqrt{2}} \sim N(0,1)$
 $\frac{\frac{(X_1 - X_2)^2}{2}}{\frac{(X_1 + X_2)^2}{2}} \sim t_1$
 $\rightarrow \chi_1^2$

Είναι αυτό σωστό ;
Για να το πω αυτό θα πρέπει να εξασφαλίσω ότι $X_1 + X_2, X_1 - X_2$ ανεξάρτητες

Ισχύει ; Αν ναι υπό ποια προϋπόθεση. (όπως στην άσκηση 12 ερωτήρια (ii))
Πρέπει ν.σ.ο. $X_1 + X_2, X_1 - X_2$ ανεξάρτητες. Είναι

$\begin{pmatrix} X_1 + X_2 \\ X_1 - X_2 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} X_1 \\ X_2 \end{pmatrix}$

Τότε: $\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$

Άρα είναι ανεξάρτητες, οπότε ούτως
 $Y_1 \sim t_1$

(i) $X_3 \sim N(0,1) \Rightarrow X_3^2 \sim \chi_1^2$

Ακόμη $X_1^2 + X_2^2 \sim \chi_2^2$ διότι $X_1, X_2 \sim N(0,1)$

Επομένως $\frac{X_3^2 / 1}{(X_1^2 + X_2^2) / 2} = \frac{X_3^2}{X_1^2 + X_2^2} \sim F_{1,2}$

(ii) $Y_3 = X_4 - \bar{X} \Rightarrow Y_3 = X_4 - \left(\frac{1}{4} \sum_{i=1}^4 X_i \right) \Rightarrow$

$Y_3 = X_4 - \frac{1}{4} (X_1 + X_2 + X_3 + X_4) \Rightarrow Y_3 = -\frac{1}{4} X_1 - \frac{1}{4} X_2 - \frac{1}{4} X_3 + \frac{3}{4} X_4$

Επομένως η Y_3 ως γραμμικός συνδυασμός X_i ακολουθεί κανονική κατανομή με μέση τιμή γιατί $X_i \sim N(0,1)$

$\mu = -\frac{1}{4}\mu_1 - \frac{1}{4}\mu_2 - \frac{1}{4}\mu_3 + \frac{3}{4}\mu_4 \xrightarrow{\text{όπου}} \mu = 0$

$\sigma^2 = \frac{1}{16} \cdot 1 + \frac{1}{16} \cdot 1 + \frac{1}{16} \cdot 1 + \frac{9}{16} \cdot 1 \Rightarrow \sigma^2 = \frac{12}{16} = \frac{3}{4}$

Άρα $Y_3 \sim N(0, 3/4)$ ή $Z = \frac{Y_3}{\sqrt{3/4}} \sim N(0,1)$

Είναι :

$P(Y_3 \geq \sqrt{3}) = P\left(\frac{Y_3}{\sqrt{3/4}} \geq \frac{\sqrt{3}}{\sqrt{3/4}}\right) = P(Z \geq 2) = 0,5 - 0,477$
(0,5 - P(0 ≤ Z ≤ 2))

Άσκηση 3 ελίδα 311-312

Άσκηση 18η: Έστω X_1, \dots, X_6 τυχαία σήματα από $N(0,1)$ και

$Z_1 = \frac{X_1 + X_2}{2}, Z_2 = \frac{X_3 + \dots + X_6}{4}, Z_3 = \frac{\sqrt{2} (X_1 - X_2)}{\sqrt{(X_3 - X_4)^2 + (X_5 - X_6)^2}}$

$Z_4 = \left[(X_1 - X_6)^2 + (X_2 - X_5)^2 + (X_3 - X_4)^2 \right] / 2$

α) να βρεθούν οι κατανομές των 6.6 :

(i) $\frac{Z_1 + Z_2}{2}$

(ii) $2Z_1^2 + 4Z_2^2$

(iii) $X_3 - Z_2$

(iv) Z_3

(v) Z_4

β) Να βρεθούν οι σταθμοί c_1 και c_2 έτσι ώστε

(i) $P(Z_3 \geq -c_1) = 0.99$

(ii) $P(Z_4 \leq c_2) = 0.99$

Λύση: $Z_1 = \frac{X_1 + X_2}{2}$ είναι γραμμικός συνδυασμός κανονικών από ακο-

λουθη κανονική. με μέση τιμή $\mu = \frac{1}{2} \cdot 0 + \frac{1}{2} \cdot 0 = 0$ και

$\sigma^2 = \frac{1}{4} \cdot 1 + \frac{1}{4} \cdot 1 = \frac{1}{2}$

Άρα $Z \sim N(0, 1/2)$

Επίσης $Z_2 = \frac{X_3 + \dots + X_6}{4}$ με το ίδιο σκεπτικό (ή ως μέσος τεσσα-

ρων το πλήθος κανονικών)

$Z_2 \sim N(0, 1/4)$

Επομένως

$\frac{Z_1 + Z_2}{2}$ (αθροισμα ανεξαρτήτων
τ.κ. από το $Z_1 \rightarrow X_1, X_2$
 $Z_2 \rightarrow X_3, \dots, X_6$)

ως γραμμικός συνδυασμός κανονικών ακολουθη κανονική κατανομή με

μέση τιμή

$\frac{1}{2} \cdot 0 + \frac{1}{2} \cdot 0 = 0$

και διακύμανση

$\frac{1}{4} \cdot \frac{1}{2} + \frac{1}{4} \cdot \frac{1}{4} = \frac{1}{4} \left(\frac{1}{2} + \frac{1}{4} \right) = \frac{1}{4} \cdot \frac{3}{4} = \frac{3}{16}$

Άρα $\frac{Z_1 + Z_2}{2} \sim N(0, 3/16)$

(ii) $Z_1 \sim N(0, 1/2) \Rightarrow \frac{Z_1}{\sqrt{1/2}} \sim N(0, 1) \Rightarrow 2Z_1^2 \sim \chi_1^2$

Άρα $2Z_1^2 + 4Z_2^2 \sim \chi_2^2$ (Z_1, Z_2 ανεξάρτητες) $Z_2 \sim N(0, 1/4) \Rightarrow \frac{Z_2}{\sqrt{1/4}} \sim N(0, 1) \Rightarrow 4Z_2^2 \sim \chi_1^2$

(iii) $X_3 - Z_2 = X_3 - \frac{1}{4}(X_3 + X_4 + X_5 + X_6)$ Επίτηθεν προς ωρίμαν

(4)

$$X_3 - Z_2 = \frac{3}{4} X_3 - \frac{1}{4} X_4 - \frac{1}{4} X_5 - \frac{1}{4} X_6$$

Είναι επομένως γραμμικός συνδυασμός κανονικών \rightarrow άρα κανονική

Μέση τιμή: $\frac{3}{4} \cdot 0 - \frac{1}{4} \cdot 0 - \frac{1}{4} \cdot 0 - \frac{1}{4} \cdot 0 = 0$

και διακύμανση: $\left(\frac{3}{4}\right)^2 \cdot 1 + \left(\frac{1}{4}\right)^2 \cdot 1 + \left(\frac{1}{4}\right)^2 + \left(\frac{1}{4}\right)^2 \cdot 1$

Επομένως $X_3 - Z_2 \sim N(0, 1^2/16 = 3/4)$

(iv) $X_1 - X_2 \sim N(0, 2)$ (γιατί;) $X_1 \sim N(0,1)$ γραμμικός συνδυασμός κανονικών

$X_2 \sim N(0,1)$ ανεξάρτητες

$\Rightarrow \frac{X_1 - X_2}{\sqrt{2}} \sim N(0,1) \Rightarrow (X_1 - X_2) \frac{1}{\sqrt{2}} \sim N(0,1)$

Παρόμοια: $\frac{X_3 - X_4}{\sqrt{2}} \sim N(0,1) \Rightarrow \frac{(X_3 - X_4)^2}{2} \sim \chi^2_1$

$\frac{(X_5 - X_6)^2}{2} \sim \chi^2_1$ } ανεξάρτητες

και άρα $\frac{(X_3 - X_4)^2 + (X_5 - X_6)^2}{2} \sim \chi^2_2$

Επομένως: $\frac{\frac{X_1 - X_2}{\sqrt{2}}}{\sqrt{\frac{(X_3 - X_4)^2 + (X_5 - X_6)^2}{2}}} \sim t_2$

χ^2_2 \uparrow χ^2_2 \uparrow b.e.

ή $Z_3 = \frac{\sqrt{2} (X_1 - X_2)}{\sqrt{(X_3 - X_4)^2 + (X_5 - X_6)^2}} \sim t_2$

Ισχύουν όλες οι υποθέσεις; Ναι

$(X_1 - X_2), (X_3 - X_4), (X_5 - X_6)$ ανεξάρτητες

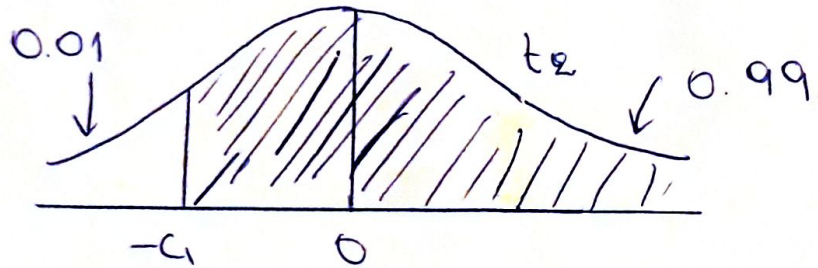
5

$$(v) X_1 - X_6 \sim N(0, 2) \Rightarrow \frac{X_1 - X_6}{\sqrt{2}} \sim N(0, 1) \Rightarrow \frac{(X_1 - X_6)^2}{2} \sim \chi_1^2$$

Όμοια $\frac{(X_2 - X_5)^2}{2} \sim \chi_1^2$ και $\frac{(X_3 - X_4)^2}{2} \sim \chi_1^2$

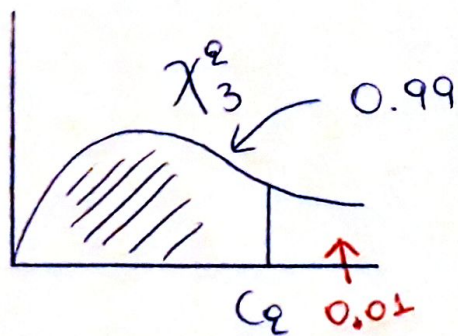
Επόμεως $Z_4 \sim \chi_3^2$ (εξήγησον ;)

b) $Z_3 \sim t_2$, $P(Z_3 \geq -c_1) = 0.99$



Επόμεως $c_1 = t_{0.01, 2} = 6.965$ (βλ. 325 η ΤΥΠΟΛΟΓΙΟ)

$Z_4 \sim \chi_3^2$, $P(Z_4 \leq c_2) = 0.99 \Rightarrow$



$$P(Z_4 \geq c_2) = 0.01$$

$$\Rightarrow c_2 = \chi_{0.01, 3}^2 = 11.345 \text{ (βλ. 324 η ΤΥΠΟΛΟΓΙΟ)}$$