δ.ε θ , $P(L(x_1,\dots,x_n) \le \theta \le U(x_1,\dots,x_n)) = 100(1-a)\%$ ← σ.ε

δ.ε μ , σ^2 γνωστό $\left(\bar{x} - z_{a/2}\frac{\sigma}{\sqrt{n}} , \bar{x} + z_{a/2}\frac{\sigma}{\sqrt{n}}\right)$

δ.ε μ, σ^2 άγνωστο.

<u>Παρατήρηση</u>

1. Το δ.ε για μ βασίζεται στο επαρκές στατιστικό $\bar{x}$ που επιπλέον είναι Ε.Μ.Π και Α.Ο.Ε.Δ
2. Το δ.ε εκφράζεται και μέσω της μεταβλητότητας σ^2 του πληθυσμού.

Έστω τ.δ $X_1,\dots,X_n$ από $N(\mu,\sigma^2)$

II) <u>δ.ε για την σ^2</u>

α. μ γνωστή.

• Κατασκευή αντιστρεπτής ποσότητας.

Τα $X_i \sim N(\mu,\sigma^2)$, $\forall i=1,\dots,n$

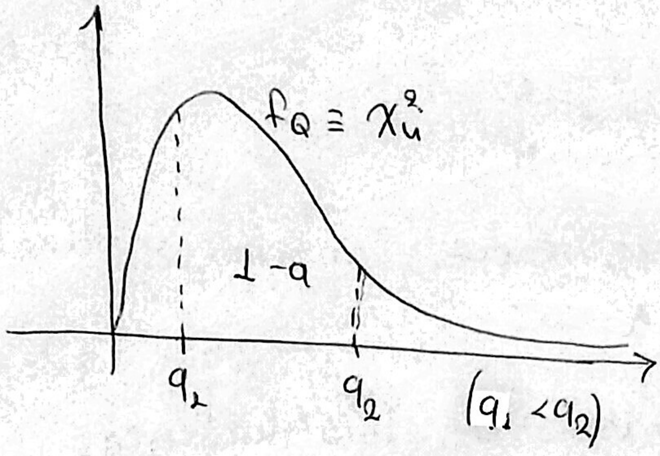
$$\frac{X_i-\mu}{\sigma} \sim N(0,1) \quad \forall i=1,\dots,n$$

$$\left(\frac{X_i-\mu}{\sigma}\right)^2 \sim X^2_1 \quad \forall i=1,\dots,n.$$

$$\sum_{i=1}^{n}\left(\frac{X_i-\mu}{\sigma}\right)^2 = \frac{\sum_{i=1}^{n}(X_i-\mu)^2}{\sigma^2} \sim X^2_{\sum_{i=1}^{n} 1} \equiv X^2_n$$

Παρατηρώ ότι η $Q = \frac{\sum_{i=1}^{n}(X_i-\mu)}{\sigma^2} \sim X_n^2$ είναι συνάρτηση του τ.δ. και της άγνωστης παραμέτρου. Άρα Q αντιστρεπτή.

$\frac{(n-1)S^2}{\sigma^2} \sim X_{n-1}^n$ (μια άλλη πιθανή Q)



Αφού Q αντιστρεπτή
$\exists\, q_1, q_2$ με $q_1 < q_2$:
$1-a = P(q_1 < Q < q_2) =$

$= P\left(q_1 < \frac{\Sigma(x_i-\mu)^2}{\sigma^2} < q_2\right) = P\left(\frac{\Sigma(x_i-\mu)^2}{q_2} < \sigma^2 < \frac{\Sigma(x_i-\mu)^2}{q_1}\right)$

Άρα το $\left(\frac{\Sigma(x_i-\mu)^2}{q_2}, \frac{\Sigma(x_i-\mu)^2}{q_1}\right)$ είναι ένα $100(1-a)\%$ δ.ε για σ^2

δ.ε ελαχίστου μήκους.

$l = \Sigma(x_i-\mu)^2\left(\frac{1}{q_1} - \frac{1}{q_2}\right)$

Αρκεί να ελαχιστοποιήσω (ως προς q_1, q_2) το l

ή το $l^* = \frac{1}{q_1} - \frac{1}{q_2}$ υπό την $1-a = P(q_1 < Q < q_2)$

ή $1-a = \int_{q_1}^{q_2} f_Q(q)\,dq$ (*)

ή $1-a = F_Q(q_2) - F_Q(q_1)$ (1), $Q \sim X_n^2$

$$\frac{d\ell^*}{dq_1} = -\frac{1}{q_1^2} - \frac{dq_2}{dq_1}\left(\frac{d}{dq_2}\left(\frac{1}{q_2}\right)\right) = -\frac{1}{q_1^2} + \frac{1}{q_2^2}\frac{dq_2}{dq_1}$$

$$\frac{d\ell^*}{dq_1} = 0 \Rightarrow -\frac{1}{q_1^2} + \frac{1}{q_2^2}\frac{dq_2}{dq_1} = 0 \Rightarrow \frac{1}{q_2^2}\frac{dq_2}{dq_1} = \frac{1}{q_1^2} \quad (2)$$

Από την (1)
$$\frac{d}{dq_1}(1-a) = \frac{d}{dq_1}F_Q(q_2) - \frac{d}{dq_1}F_Q(q_1) \Rightarrow$$

$$\Rightarrow 0 = \frac{dq_2}{dq_1}\frac{d}{dq_2}F_Q(q_2) - f_Q(q_1) \Rightarrow$$

$$\Rightarrow 0 = \frac{dq_2}{dq_1}f_Q(q_2) - f_Q(q_1) \Rightarrow \frac{dq_2}{dq_1} = \frac{f_Q(q_1)}{f_Q(q_2)} \quad (3)$$

Από (2) και (3)
$$\frac{d\ell^*}{dq_1} = 0 \Rightarrow \frac{1}{q_2^2}\frac{f_Q(q_1)}{f_Q(q_2)} = \frac{1}{q_1^2} \Rightarrow$$

$$\Rightarrow q_1^2 f_Q(q_1) = q_2^2 f_Q(q_2) \quad (4)$$

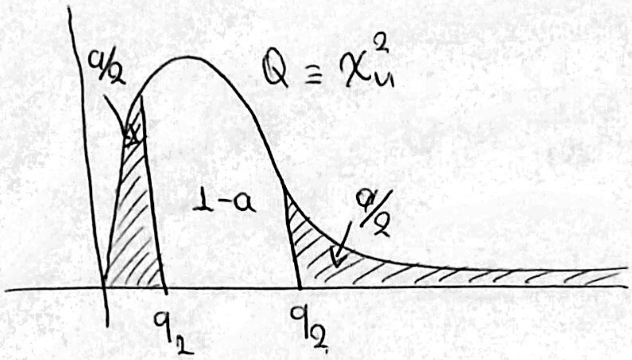
Επειδή η $Q \sim \chi^2_n \Rightarrow f_Q(q) = \frac{1}{\Gamma(\frac{n}{2})\, 2^{n/2}}\, q^{\frac{n}{2}-1} e^{-q/2}, \quad q>0$

Άρα η (*) γίνεται:
$$\int_{q_1}^{q_2} \frac{1}{\Gamma(\frac{n}{2})\, 2^{n/2}}\, q^{\frac{n}{2}-1} e^{-q/2}\, dq = 1-a \quad (5)$$

Η (4) και (5) συνιστούν σύστημα δύο εξισώσεων με δύο αγνώστους (τα q_1, q_2) που η λύση του οδηγεί στα q_1 και q_2 που (πιθανώς) ελαχιστοποιούν το ℓ^*.
Το σύστημα αυτό λύνεται μόνο με μεθόδους της αριθμητικής ανάλυσης.

Σε περιπτώσεις που δεν μπορούν τα q_1, q_2 να προσδιοριστούν αναλυτικά, καταφεύγουμε σε Δ.Ε <u>Ίσων Ουρών</u>

Δηλαδή Δ.Ε που τα q_1, q_2 είναι τέτοια ώστε να κόβουν ίσες ουρές από την κατανομή της Q



Άρα διαλέγω τα q_1, q_2 τέτοια ώστε:

$$P(Q \geq q_2) = \frac{a}{2} = P(Q \leq q_1) \quad (6)$$

Ερώτημα: Ποια τα q_1, q_2 που ικανοποιούν την (6)?

Θυμάμαι τον ορισμό των αντίστροφων εκατοστιαίων σημείων χ^2_n

Το $\chi^2_{n,a}$ ορίζεται $P(\chi^2_n \geq \chi^2_{n,a}) = a$

Άρα από $P(Q \geq q_2) = \frac{a}{2} \Rightarrow q_2 \equiv \chi^2_{n,\frac{a}{2}}$

$$P(Q \leq q_1) = \frac{a}{2} \Rightarrow 1 - P(Q \geq q_1) = \frac{a}{2} \Rightarrow P(Q \geq q_1) = 1 - \frac{a}{2} \Rightarrow$$

$$\Rightarrow q_1 \equiv \chi^2_{n,1-\frac{a}{2}}$$

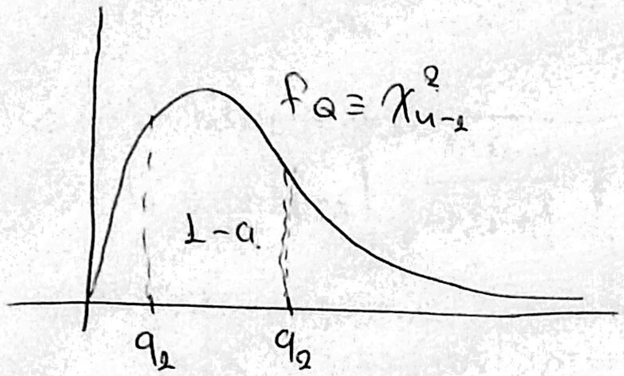
Το $100(1-a)\%$ Δ.Ε. ίσων ουρών για το σ^2 όταν μ γνωστό είναι

$$\left(\frac{\sum_{i=1}^{n}(x_i-\mu)^2}{\chi^2_{n,\frac{a}{2}}} , \frac{\sum_{i=1}^{n}(x_i-\mu)^2}{\chi^2_{n,1-\frac{a}{2}}} \right)$$

β. μ άγνωστο.

Αντιστρεπτή: $Q = \frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{n-1}$

Αφού Q αντιστρεπτή $\exists\, q_1, q_2 \ (q_1 < q_2)$ τέτοια ώστε



$$1-a = P(q_1 < Q < q_2) =$$

$$= P\left(q_1 < \frac{(n-1)S^2}{\sigma^2} < q_2\right) =$$

$$= P\left(\frac{(n-1)S^2}{q_2} < \sigma^2 < \frac{(n-1)S^2}{q_1}\right)$$

Από τον ορισμό του δ.ε, ένα δ.ε για σ^2 είναι

$$\left(\frac{(n-1)S^2}{q_2}, \frac{(n-1)S^2}{q_1}\right)$$

Η ελαχιστοποίηση του μήκους δεν οδηγεί σε αναλυτική λύση για τα q_1, q_2

Καταφεύγουμε σε δ.ε ίσων ουρών $\begin{cases} q_1 = \chi^2_{n-1, 1-\frac{a}{2}} \\ q_2 = \chi^2_{n-1, a/2} \end{cases}$

Τελικά το $100(1-a^2)\%$ δ.ε ίσων ουρών για σ^2 είναι:

$$\left(\frac{(n-1)S^2}{\chi^2_{n-1, \frac{a}{2}}}, \frac{(n-1)S^2}{\chi^2_{n-1, 1-\frac{a}{2}}}\right.$$

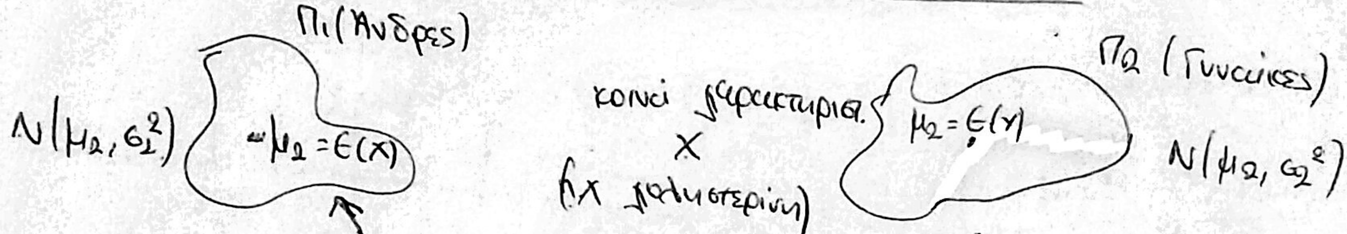
Ποιο το δ.ε για σ^2?

Βασιζόμαστε στο δ.ε για σ^2

$$1-\alpha = P\left(\frac{(n-1)S^2}{\chi^2_{n-1,\frac{\alpha}{2}}} < \sigma^2 < \frac{(n-1)S^2}{\chi_{n-1,1-\frac{\alpha}{2}}}\right) \Rightarrow$$

$$\Rightarrow 1-\alpha = P\left(\sqrt{\frac{(n-1)S^2}{\chi^2_{n-1,\frac{\alpha}{2}}}} < \sigma < \sqrt{\frac{(n-1)S^2}{\chi_{n-1,1-\frac{\alpha}{2}}}}\right)$$

Δύο πληθυσμοί - Δύο ανεξάρτητα δείγματα.



Έστω τ.δ $X_1,\dots X_{n_1}$ και $Y_1,\dots,Y_{n_2}$

Διαστήματα Εμπιστοσύνης για την διαφορά των μέσων τιμών δύο ανεξάρτητων κανονικών πληθυσμών.

Έστω τ.δ $X_1,\dots,X_{n_1}$ από πληθυσμό $N(\mu_1,\sigma_1^2)$

Έστω τ.δ $Y_1,\dots,Y_{n_2}$ από πληθυσμό $N(\mu_2,\sigma_2^2)$

Μαθηματική Αναγκαιότητα: Υποθέτω ότι τα δύο τ.δ ανεξάρτητα

1) Ζητώ τ.δ για $\mu_1-\mu_2$ όταν σ_1^2, σ_2^2 γνωστές.

- Κατασκευή αντιστρέψιμης: Αφού $X_1,\dots,X_{n_1} \sim N(\mu_1,\sigma_1^2)$

  Τότε $\bar{X} \sim N\left(\mu_1, \frac{\sigma_1^2}{n_1}\right)$

  Αφού $Y_1,\dots,Y_{n_2} \sim N(\mu_2,\sigma_2^2) \longrightarrow \bar{Y} \sim N\left(\mu_2, \frac{\sigma_2^2}{n_2}\right)$

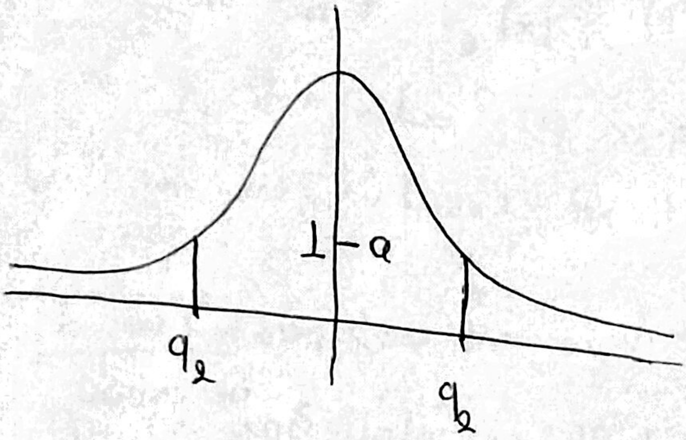
Επειδή τα τ.δ ανεξάρτητα ($\bar{X}, \bar{Y}$ ανεξάρτητα) τότε

$$\bar{X} - \bar{Y} \sim N\left(\mu_1 - \mu_2, \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}\right)$$

Θεωρώ $Q = \dfrac{\bar{X} - \bar{Y} - (\mu_1 - \mu_2)}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}} \sim N(0,1)$

Άρα Q αντιστρεπτή

Αφού Q αντιστρεπτή $\exists\ q_1, q_2\ (q_1 < q_2)$ τέτοια ώστε



$$1-\alpha = P(q_1 < Q < q_2) =$$

$$= P\left(q_1 < \frac{\bar{X} - \bar{Y} - (\mu_1 - \mu_2)}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}} < q_2\right)$$

$$= P\left((\bar{X} - \bar{Y}) + q_1\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} < \mu_1 - \mu_2 < (\bar{X} - \bar{Y}) + q_2\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}\right)$$

Άρα $100(1-\alpha)\%$ δ.ε για $\mu_1 - \mu_2$

$$\left((\bar{X} - \bar{Y}) + q_1\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}},\ (\bar{X} - \bar{Y}) + q_2\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}\right)$$

Μετά από γνωστή διαδικασία το μήκος ελαχιστοποιείται για $q_2 = -q_1$

οπότε $q_1 = -z_{\alpha/2}$, $q_2 = z_{\alpha/2}$

2) Ζητώ δ.ε για $\mu_1-\mu_2$ όταν σ_1^2, σ_2^2 άγνωστες αλλά υποθέτω ότι είναι ίσες μεταξύ τους $\sigma_1^2=\sigma_2^2=\sigma^2$

Κατασκευή Αντιστρεπτής: $\bar{X} \sim N\left(\mu_1, \frac{\sigma^2}{n_1}\right)$ και $\bar{Y} \sim N\left(\mu_2, \frac{\sigma^2}{n_2}\right)$

Λόγω της ανεξαρτησίας των δειγμάτων $\bar{X}-\bar{Y} \sim N\left(\mu_1-\mu_2, \frac{\sigma^2}{n_1}+\frac{\sigma^2}{n_2}\right) \equiv$

$\equiv N\left(\mu_1-\mu_2, \sigma^2\left(\frac{1}{n_1}+\frac{1}{n_2}\right)\right)$

Άρα $\dfrac{\bar{X}-\bar{Y}-(\mu_1-\mu_2)}{\sigma\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}} \sim N(0,1)$

$$\left.\begin{array}{l}\text{Το } \dfrac{(n_1-1)S_1^2}{\sigma^2} \sim \chi^2_{n_1-1} \\ \text{Το } \dfrac{(n_2-1)S_2^2}{\sigma^2} \sim \chi^2_{n_2-1}\end{array}\right\} \Rightarrow \text{Λόγω της ανεξαρτησίας των δειγμάτων}$$

$$\frac{(n_1-1)S_1^2+(n_2-1)S_2^2}{\sigma^2} \sim \chi^2_{n_1+n_2-2}$$

Θεωρώ την $Q = \dfrac{\dfrac{\bar{X}-\bar{Y}-(\mu_1-\mu_2)}{\sigma\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}}{\sqrt{\dfrac{\frac{(n_1-1)S_1^2+(n_2-1)S_2^2}{\sigma^2}}{n_1+n_2-2}}} \Rightarrow$

$$\Rightarrow \begin{cases} \text{αλγεβρικά} & Q = \dfrac{\bar{X}-\bar{Y}-(\mu_1-\mu_2)}{\sqrt{\frac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}}\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}} \\ \text{πιθανοθεωρητικά} & Q = \dfrac{N(0,1)}{\sqrt{\frac{\chi^2_{n_1+n_2-2}}{n_1+n_2-2}}} \overset{\text{ορ}}{\sim} t_{n_1+n_2-2} \end{cases}$$

Άρα η Q αντιστρεπτή

Ερώτημα: Πως μπορώ να κατασκευάζω γενικά αντιστρεπτές ποσότητες?

ΠΡΟΤΑΣΗ 1

Έστω τ.μ X με συνεχή κατανομή F_θ, $\theta \in \Theta \subseteq \mathbb{R}$

α. Τότε η τ.μ $Y = F_\theta(x) \sim U(0,1)$, $\forall \theta \in \Theta$

β. Η τ.μ $Z = -2 \log F_\theta(x) \sim \chi^2_2$

γ. Η τ.μ. $W = -2 \log(1 - F_\theta(x)) \sim \chi^2_2$

Απόδειξη

α. Το αποδείξαμε στο 1^o μάθημα

β. $Z = -2 \log F_\theta(x)$, $Y = F_\theta(x) \sim U(0,1)$

$Z = -2 \log Y$ όπου $Y \sim U(0,1)$

Θεωρώ τον μετασχηματισμό $Z = h(y) = -2 \log y$, $y \in (0,1)$

Άρα $Z > 0$

Ο h είναι 1-1 : $y = h^{-1}(z) = e^{-z/2}$, $z > 0$

$$\frac{dh^{-1}(z)}{dz} = -\frac{1}{2} e^{-z/2}$$ συνεχής και $\neq 0$ $\forall z > 0$

Άρα $f_Z(z) = f_Y(h^{-1}(z)) \left| \frac{dh^{-1}(z)}{dz} \right|$ (1)

Αφού $Y \sim U(0,1)$ $\quad f_Y(y) = 1 \quad \forall y \in (0,1)$

Άρα ① $f_Z(z) = \frac{1}{2} e^{-z/2}$

$$f_{\chi^2_n}(x) = \frac{1}{\Gamma(\frac{n}{2}) 2^{n/2}} x^{n/2-1} e^{-x/2}$$

$\Gamma(1) = (1-1)! = 1$

Για $n=2$ $\quad f_{\chi^2_2}(x) = \frac{1}{\Gamma(1) 2} x^0 e^{-x/2} = \frac{1}{2} e^{-x/2}$ ②

Από ① και ② η $Z = -2\log Y \sim \chi^2_2$

Θεωρώ τον μετασχηματισμό $W = u(y) = -2\log(1-y)$ $\quad , y \in (0,1)$

Άρα $w > 0$

<u>ΠΡΟΤΑΣΗ 2</u>

Έστω τ.δ $X_1, \dots, X_n$ από συνεχή κατανομή F_θ, $\theta \in \Theta \subseteq \mathbb{R}$. Τότε

α. Η τ.μ $Q = -2 \sum_{i=1}^{n} \log F_\theta(x_i) \sim \chi^2_{2n}$

β. Η τ.μ. $Q = -2 \sum_{i=1}^{n} \log(1 - F_\theta(x_i)) \sim \chi^2_{2n}$

$\boxed{\text{Απόδειξη}}$

α. $Q = \sum_{i=1}^{n} (-2 \log F_\theta(x_i)) \xlongequal{\text{Πρόταση 1β.}} \sum_{i=1}^{n} \chi^2_2 \ \underset{\text{αφού } X_i \text{ ανεξ}}{\overset{\chi^2_2 \text{ ανεξάρτ}}{=\!=\!=}}$

$= \chi^2_{\sum_{i=1}^{n} 2} \equiv \chi^2_{2n}$

Αφού $Y \sim U(0,1)$ $f_Y(y) = 1$ $\forall y \in (0,1)$

Άρα ① $f_Z(z) = \frac{1}{2} e^{-z/2}$

$$f_{\chi^2_\nu}(x) = \frac{1}{\Gamma(\frac{\nu}{2}) 2^{\nu/2}} x^{\nu/2 - 1} e^{-x/2}$$

$\Gamma(1) = (1-1)! = 1$

Για $\nu = 2$ $f_{\chi^2_2}(x) = \frac{1}{\Gamma(1) 2} x^0 e^{-x/2} = \frac{1}{2} e^{-x/2}$ ②

Από ① και ② η $Z = -2\log Y \sim \chi^2_2$

Θεωρώ τον μετασχηματισμό $W = h(y) = -2\log(1-y)$, $y \in (0,1)$

Άρα $W > 0$

<u>ΠΡΟΤΑΣΗ 2</u>

Έστω τ.δ $X_1, \dots, X_n$ από συνεχή κατανομή F_θ, $\theta \in \Theta \subseteq \mathbb{R}$. Τότε

a. Η τ.μ $Q = -2 \sum_{i=1}^{n} \log F_\theta(x_i) \sim \chi^2_{2n}$

b. Η τ.μ. $Q = -2 \sum_{i=1}^{n} \log(1 - F_\theta(x_i)) \sim \chi^2_{2n}$

Απόδειξη

a. $Q = \sum_{i=1}^{n} (-2 \log F_\theta(x_i)) \xrightarrow{\text{Πρόταση 1 b.}} \sum_{i=1}^{n} \chi^2_2 \; \frac{\chi^2_2 \text{ ανεξαρτ}}{\text{αφού } x_i \text{ ανεξ}}$

$= \chi^2_{\sum_{i=1}^{n} 2} = \chi^2_{2n}$