Παρασκευή 15/05/20

Παρατήρηση:

Αφού έχουμε κατασκευάσει δ.ε για την σ^2, μπορούμε να προσδιορίσουμε δ.ε για την τυπική απόκλιση σ. Έστω άγνωστη μέση τιμή μ, έστω τ.δ $X_1, \dots, X_n$ από πληθυσμό με κατανομή $N(\mu, \sigma^2)$ με άγνωστο μ.

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

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

Δ.Ε για τη διαφορά των μέσων τιμών δύο ανεξάρτητων κανονικών πληθυσμών.

Έστω μ_1 ο μέσος δείκτης νοημοσύνης του πληθυσμού των παιδιών από υψηλόμισθες οικογένειες και μ_2 ο μέσος δείκτης νοημοσύνης των παιδιών από χαμηλόμισθες οικογένειες. Κοινό χαρακτηριστικό των δύο πληθυσμών είναι ο δείκτης νοημοσύνης τους. Οποιαδήποτε σύγκριση των δεικτών νοημοσύνης των δύο πληθυσμών εκφράζεται με τη βοήθεια των μέσων τιμών μ_1, μ_2.

Ειδικότερα η διαφορά $\mu_1 - \mu_2$ αποτελεί μέτρο διαφορετικότητας της συμπεριφοράς των δύο πληθυσμών ως προς το κοινό χαρακτηριστικό που αντιπροσωπεύεται από τα μ_1, μ_2 κατά μέσο όρο

- Όταν $\mu_1 - \mu_2$ κοντά στο 0 σημαίνει ότι δεν υπάρχει διαφορά στους δείκτες νοημοσύνης.

- $\mu_1 - \mu_2 > 0$: κατά μέσο όρο τα παιδιά από υψηλόβαθμες οικογένειες έχουν υψηλότερο δείκτη νοημοσύνης

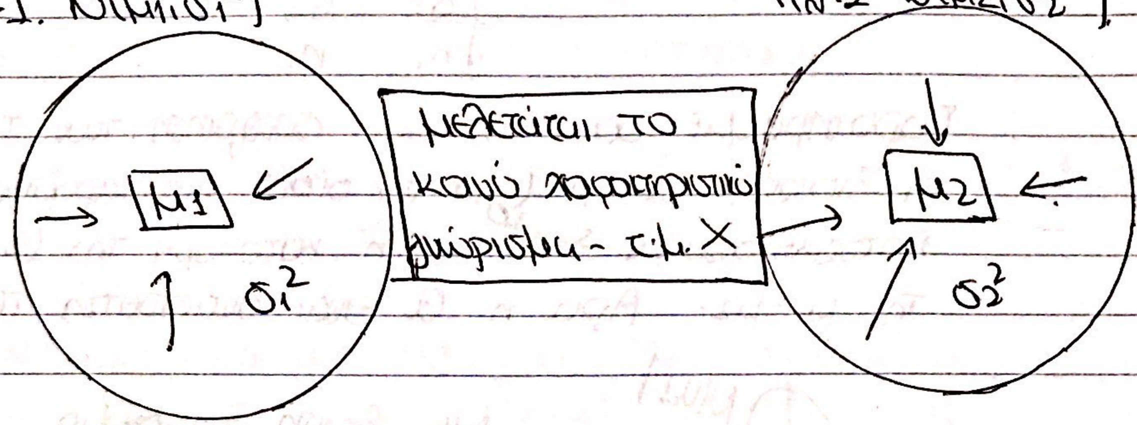
- $\mu_1 - \mu_2 < 0$: ακριβώς το αντίθετο από

Η προσέγγιση της διαφοράς $\mu_1 - \mu_2$ μπορεί να επιτευχθεί με τη βοήθεια στατιστικών μεθοδολογιών που στηρίζονται στις υποθέσεις

α) οι δύο πληθυσμοί περιγράφονται από κανονικές κατανομές $N(\mu_1, \sigma_1^2)$ και $N(\mu_2, \sigma^2)$

β) οι πληθυσμοί είναι ανεξάρτητοι που σημαίνει ότι η συμπεριφορά των μελών του ενός πληθυσμού δεν επηρεάζει την συμπεριφορά των μελών του άλλου πληθυσμού.

πλ.1 $N(\mu_1, \sigma_1^2)$ πλ.2 $N(\mu_2, \sigma_2^2)$



Θα κατασκευαστεί διάστημα εμπιστοσύνης για $\mu_1 - \mu_2$ όταν:

α) οι διακυμάνσεις σ_1^2 και σ_2^2 είναι γνωστές

β) όταν οι διακυμάνσεις σ_1^2 και σ_2^2 είναι άγνωστες αλλά ίσες $\sigma_1^2 = \sigma_2^2 = \sigma^2$

Δ.Ε. για $\mu_1 - \mu_2$ όταν σ_1^2, σ_2^2 γνωστές.

Έστω τ.δ $X_1, \dots X_{n_1}$ από $N(\mu_1, \sigma_1^2)$ και $Y_1, \dots, Y_{n_2}$ από $N(\mu_2, \sigma_2^2)$ με σ_1^2, σ_2^2 γνωστές.
Θεωρούμε ότι τα τ.δ είναι ανεξάρτητα.
Για την κατασκευή δ.ε του $\mu_1 - \mu_2$ θα πρέπει να στηριχτούμε στους εκτιμητές των μ_1 και μ_2 (ή στα επαρκή στατιστικά) που είναι $\bar{X}$ και $\bar{Y}$ αντίστοιχα.

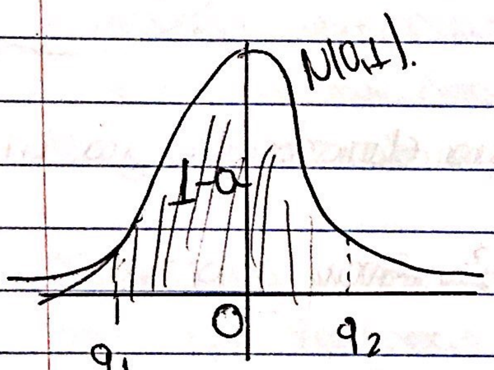
Γνωρίζουμε ότι $\bar{X} \sim N\left(\mu_1, \frac{\sigma_1^2}{n_1}\right)$ και $\bar{Y} \sim N\left(\mu_2, \frac{\sigma_2^2}{n_2}\right)$.

Και αφού τα τ.δ είναι ανεξάρτητα

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

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

Παρατηρούμε ότι η Q είναι συνάρτηση των τ.δ., περιέχει τη διαφορά $\mu_1-\mu_2$ (για την οποία ενδιαφερόμαστε να κατασκευάσουμε δ.ε) και η κατανομή της Q είναι ανεξάρτητη της $\mu_1-\mu_2$. Άρα η Q είναι αντιστρεπτή ποσότητα



Με βάση το σχήμα
Υπάρχουν q_1, q_2 $(-\infty<q_1<q_2<+\infty)$
τ.ω.
$P(q_1<Q<q_2)=1-a$

Επίσης $1-a=P(q_1<Q<q_2)$

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

$1-a = (\bar{X}-\bar{Y}) - q_2\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}} < \mu_1-\mu_2 < (\bar{X}-\bar{Y}) - q_1\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}$

Επομένως ένα διάστημα εμπιστοσύνης για $\mu_1-\mu_2$ με σ.ε $100(1-a)\%$ είναι

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

Για την εύρεση του δ.ε ελάχιστου μήκους αναζητούνται τα q_1, q_2 που ελαχιστοποιούν το μήκος

$$l = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}\,(q_2 - q_1)$$ υπό τον περιορισμό $\Phi(q_2) - \Phi(q_1) = 1 - a$.

με Φ την α.σ.κ της $N(0,1)$ κατανομής

$$\frac{dl}{dq_1} = 0 \Rightarrow \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}\left(\frac{dq_2}{dq_1} - 1\right) = 0 \Rightarrow \frac{dq_2}{dq_1} = 1. \quad (1)$$

Παραγωγίζοντας ως προς q_1 τον περιορισμό παίρνουμε ότι $\frac{dq_2}{dq_1} = \frac{f_Z(q_1)}{f_Z(q_2)}$, $Z \sim N(0,1)$

Λαμβάνοντας υπόψη την (1) προκύπτει ότι $f_Z(q_1) = f_Z(q_2)$ $\Rightarrow q_2 = -q_1$ λόγω της συμμετρίας της σ.π.π. f_Z της $Z \sim N(0,1)$ κατανομής

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

από την οποία προκύπτει $q_2 = z_{\frac{a}{2}}$ και $q_1 = q_2 = -z_{\frac{a}{2}}$

Συμπερασματικά: Το $100(1-a)\%$ δ.ε για την $\mu_1 - \mu_2$ είναι

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

Δ.Ε για $\mu_1-\mu_2$ όταν σ_1^2, σ_2^2 άγνωστες αλλά ίσες $\sigma_1^2=\sigma_2^2=\sigma^2$

Έστω τ.δ $X_1, \dots X_{n_1}$ από $N(\mu_1, \sigma_1^2)$ και $Y_1 \dots Y_{n_2}$ από $N(\mu_2, \sigma_2^2)$ με σ_1^2, σ_2^2 άγνωστες αλλά ίσες με κοινή τιμή σ^2, άγνωστη. Υποθέτουμε ότι τ.δ ανεξάρτητα.

$$\bar{X} \sim N\left(\mu_1, \frac{\sigma^2}{n_1}\right) \quad ; \quad \bar{Y} \sim N\left(\mu_2, \frac{\sigma^2}{n_2}\right)$$

Λόγω ανεξαρτησίας

$$\bar{X}-\bar{Y} \sim N\left(\mu_1-\mu_2, \sigma^2\left(\frac{1}{n_1}+\frac{1}{n_2}\right)\right)$$ από όπου παίρνουμε

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

Η κατανομή της παραπάνω ποσότητας είναι ανεξ. της $\mu_1-\mu_2$ όμως δεν μπορεί να χρησιμοποιηθεί ως αντίστροφη ποσότητα διότι περιέχει την κοινή αλλά άγνωστη σ^2. Θα πρέπει με κάποιον τρόπο να απαλοιφεί η σ^2.

Γνωρίζουμε ότι $\frac{(n_1-1)S_1^2}{\sigma^2} \sim X^2_{n_1-1}$ και $\frac{(n_2-1)S_2^2}{\sigma^2} \sim X^2_{n_2-1}$

Λόγω της ανεξαρτησίας των δειγμάτων:

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

Συγκρίνοντας τις (1) και (2) παρατηρούμε ότι αν διαιρεθούν απαλοίφεται η άγνωστη σ^2 και η κατανομή του πηλίκου είναι κοντά στην t-κατανομή.

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

Από όποψη κατανομής για την $Q \equiv \dfrac{N(0,1)}{\sqrt{\dfrac{X^2_{n_1+n_2-2}}{n_1+n_2-2}}} \sim t_{n_1+n_2-2}$ (3)

$$Q = \frac{\bar{X}-\bar{Y}-(\mu_1-\mu_2)}{\sqrt{\dfrac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}}\sqrt{\dfrac{1}{n_1}+\dfrac{1}{n_2}}} \quad (4)$$

Από τις (3), (4) προκύπτει ότι η Q είναι αντιστρεπτή ποσότητα.
Αν αξιοποιηθεί η Q της (4) όπως στην περίπτωση δ.ε για την μ, κανονικού πληθυσμού $N(\mu, \sigma^2)$, με σ^2 άγνωστη, καταλήγουμε ότι το $100(1-a)\%$ δ.ε για την $\mu_1-\mu_2$ είναι

$$\left((\bar{X}-\bar{Y}) - t_{n_1+n_2-2,\frac{a}{2}}\sqrt{\left(\frac{1}{n_1}+\frac{1}{n_2}\right)S_p^2},\ (\bar{X}-\bar{Y}) + t_{n_1+n_2-2,\frac{a}{2}}\sqrt{\left(\frac{1}{n_1}+\frac{1}{n_2}\right)S_p^2}\right)$$

με $S_p^2 = \dfrac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}$.

Παρατήρηση:
Ανάλογο δε μπορεί να κατασκευαστεί για σ_1^2, σ_2^2

$\frac{(n_1-1)S_1^2}{\sigma_1^2} \sim \chi^2_{n-1}$ $\frac{(n_2-1)S_2^2}{\sigma_2^2} \sim \chi^2_{n_2-1}$

λόγω ανεξ. $\frac{S_1^2}{\sigma_1^2} \sim \frac{\chi^2_{n-1}}{n-1}$ $\frac{S_2^2}{\sigma_2^2} \sim \frac{\chi^2_{n_2-1}}{n_2-1}$

$Q = \frac{\sigma_1^2 S_2^2}{\sigma_2^2 S_1^2} \sim f_{n_2-1, n_1-1}$

$1-a = P(q_1 < Q < q_2) = P\left(q_1 \frac{S_1^2}{S_2^2} < \frac{\sigma_1^2}{\sigma_2^2} < q_2 \frac{S_1^2}{S_2^2}\right)$

όπου $q_1 = f_{n_2-1, n_1-1, 1-\frac{a}{2}}$, $q_2 = f_{n_2-1, n_1-1, \frac{a}{2}}$

Κατασκευή Αντίστρεπτης Ποσότητας

Πρόταση 1:

α) Έστω τ.μ. X με συνεχή κατανομή F_θ, $\theta \in \Theta$
Τότε η τ.μ. $Y = F_\theta(X)$ έχει ομοιόμορφη κατανομή $U(0,1)$ για $\theta \in \Theta \subset \mathbb{R}$.

β) Η τυχαία μεταβλητή $Z = -2\log F_\theta(X)$ έχει κατανομή χ^2_2.

$\boxed{\text{Αποδ}}$

Θεωρούμε τον μετασχηματισμό $z = h(y) = -2\log y$.
Επειδή $y \in (0,1)$ το $z > 0$. Η h είναι 1-1 και
$y = h^{-1}(z) = e^{-z/2}$, $z > 0$

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

Επομένως η σ.π.π της Z είναι

$f_Z(z) = f_Y(h^{-1}(z)) \left| \frac{dh^{-1}(z)}{dz} \right| , \quad z>0$

$f_Y(y)=1, \ y\in(0,1)$ αφού $F_\theta(X) \sim U(0,1)$

$f_Z(z) = \frac{1}{2} e^{-\frac{z}{2}} , \ z>0$

Η σ.π.π της X_n^2 είναι $f_{X_n^2}(x) = \frac{1}{\Gamma(\frac{n}{2}) 2^{n/2}} x^{\frac{n}{2}-1} e^{-\frac{x}{2}}, x>0$

για $n=2$ συμπίπτει με την f_Z.

Άρα $Z = -2\log F_\theta(X) = -2\log Y \sim X_2^2$.

δ) Η τ.μ $W = -2\log(1-F_\theta(X))$ έχει κατανομή X_2^2

Πρόταση 2:

Έστω τ.δ $X_1, \dots X_n$ από μια συνεχή κατανομή $F_\theta, \theta\in\Theta$

Τότε η τ.μ

α) $Q = -2\sum_{i=1}^{n} \log F_\theta(X_i)$ έχει κατανομή X_{2n}^2

β) $Q^* = -2\sum_{i=1}^{n} \log(1-F_\theta(X_i))$ έχει κατανομή X_{2n}^2.

και έτσι είναι αντιστρεπτές ποσότητες.

Αποδ α)

$Q = -2\sum_{i=1}^{n} \log F_\theta(X_i) = \sum_{i=1}^{n} (-2\log F_\theta(X_i)) = \sum_{i=1}^{n} X_2^2 \underset{\text{ανεξ.}}{\overset{X_2^2}{=}} X^2_{\sum_{i=1}^{n} 2} = X_{2n}^2$

Αν τ.δ από συνεχή κατανομή για να βρούμε αντίστροφη ποσότητα προσδιορίζουμε την α.σ.κ F_θ του πληθυσμού και μέσω αυτής τις Q, Q^* της πρότ. 2.

θ, Q, Q^* είναι συναρτήσεις του τ.δ. και δεν εξαρτώνται οι κατανομές τους από την παράμετρο άρα είναι αντίστροφες.
Αν η F_θ είναι της μορφής 1-(μια ποσότητα) χρησιμοποιείται η Q^* αλλιώς η Q.

Παράδειγμα Βήτα Κατανομή

Έστω τ.δ $X_1, X_2, \dots X_n$ από Beta$(\theta, 1)$ και σ.π.π
$f(x,\theta) = \theta x^{\theta-1}$, $0<x<1$ και $\theta>0$
Να κατασκευαστεί διάστημα ίσων ουρών για τη θ με Β.Ε $100(1-a)\%$

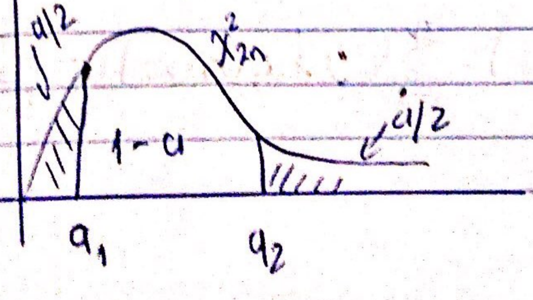
Πρώτα βρίσκουμε την α.σ.κ.

$$F_\theta(x) = \int_{-\infty}^{x} f_X(t)\,dt = \begin{cases} 0, & x<0 \\ \int_0^x \theta t^{\theta-1}dt = x^\theta, & 0<x<1 \\ 1, & x \geq 1. \end{cases}$$

Εφαρμόζοντας το (α) της πρότασης 2 έχουμε

$$Q = -2\sum_{i=1}^{n} \log F_\theta(X_i) = -2\sum_{i=1}^{n} \log X_i^\theta = -2\theta \sum_{i=1}^{n} \log X_i$$

Η $Q \sim X^2_{2n}$ είναι αντίστροφη ποσότητα



με βάση το σχήμα υπάρχουν q_1, q_2 $(0<q_1<q_2<+\infty)$ τ.ω

$$P(q_1<Q<q_2) = 1-a = P\left(q_1 < -2\theta \sum_{i=1}^{n} \log X_i < q_2\right)$$

$$\underset{\log X_i<0}{\overset{0<X_i<1}{=}} P\left(\frac{-q_1}{2\sum_{i=1}^{n}\log X_i} < \theta < -\frac{q_2}{2\sum_{i=1}^{n}\log X_i}\right)$$

Άρα το δ.ε για την θ είναι το:

$$\left(\frac{-q_1}{2\sum_{i=1}^{n}\log X_i}, \frac{-q_2}{2\sum_{i=1}^{n}\log X_i}\right)$$

Για τη κατασκευή δ.ε ίσων ουρών τα q_1 και q_2 επιλ. ώστε $P(Q\geq q_2) = \frac{a}{2} = P(Q\leq q_1)$

$$P(Q\geq q_2) = \frac{a}{2} \Rightarrow P(X^2_{2n}\geq q_2) = \frac{a}{2}, \quad q_2 = X^2_{2n, \frac{a}{2}}$$

$$P(Q\leq q_1) = \frac{a}{2} \Rightarrow P(X^2_{2n}\leq q_1) = \frac{a}{2} \Rightarrow P(X^2_{2n}\geq q_1) = 1-a/2$$

$$\Rightarrow q_1 = X^2_{2n, 1-\frac{a}{2}}$$

Έτσι το $100(1-a)\%$ δ.ε ίσων ουρών για τη θ είναι

$$\left(-\frac{X^2_{2n, 1-\frac{a}{2}}}{2\sum_{i=1}^{n}\log X_i}, -\frac{X^2_{2n, \frac{a}{2}}}{2\sum_{i=1}^{n}\log X_i}\right)$$

Παρατήρηση:

a) αν αναζητούσαμε δ.ε ελαχίστου μήκους τα q_1, q_2 δεν θα μπορούσαν να προσδιορισθούν σε κλειστή μορφή θ) Το επαρκές στατιστικό της Beta(θ,1) είναι το $\prod_{i=1}^{n} X_i$ (Neyman-Fisher)

Το παραπάνω διάστημα ίσων ουρών είναι συνάρτηση του επαρκούς στατιστικού.

Παράδειγμα (Ομοιόμορφη Κατανομή):

Έστω τ.δ $X_1, \dots X_n$ από $U(0,\theta)$, $0<x<\theta$, $\theta>0$
Να κατασκευασθεί δ.ε ίσων ουρών για τη θ.

Η α.σ.κ. είναι $F_\theta(x) = \begin{cases} 0 & , x<0 \\ x/\theta & , 0<x<\theta \\ 1 & , x \geq \theta \end{cases}$

Η αντιστρ. ποσοτ. είναι $Q = -2\sum_{i=1}^{n} \log F_\theta(X_i) =$

$= -2\sum_{i=1}^{n} \log X_i + 2n\log\theta \sim \chi^2_{2n}$

$\exists\, q_1, q_2$ με $0<q_1<q_2<+\infty$ τ.ω

$$1-a = P(q_1 < Q < q_2) = P\left(q_1 < -2\sum_{i=1}^{n} \log X_i + 2n\log\theta < q_2\right)$$

$$= P\left(e^{\frac{q_1}{2n}}\left(\prod_{i=1}^{n} X_i\right)^{1/n} < \theta < e^{\frac{q_2}{2n}}\left(\prod_{i=1}^{n} X_i\right)^{1/n}\right)$$

Έτσι για $q_1 = \chi^2_{2n,1-\frac{a}{2}}$, $q_2 = \chi^2_{2n,\frac{a}{2}}$ το δ.ε ίσων ουρών είναι

$$\left(e^{\frac{q_1}{2n}}\left(\prod_{i=1}^{n} X_i\right)^{1/n}, e^{\frac{q_2}{2n}}\left(\prod_{i=1}^{n} X_i\right)^{1/n}\right)$$

Scanned with CamScanner

Παρατήρηση: Το δ.ε ίσων ουρών δεν είναι συνάρτηση του επαρκούς στατιστικού

Παράδειγμα (ομοιόμορφη κατανομή):

Έστω τ.δ $X_1, \dots X_n$ από πληθυσμό με κατανομή $U(0,\theta)$ και έστω $X_{(n)} = \max\{X_1, \dots X_n\}$

α) Να δειχθεί ότι $X_{(n)}/\theta$ είναι αντιστρέψιμη ποσότητα
β) Να βρεθεί το $100(1-a)\%$ δ.ε ίσων ουρών για τη θ.
γ) Να βρεθεί το $100(1-a)\%$ δ.ε ελαχίστου μήκους για τη θ.

α) Έστω $Y = X_{(n)}$. Η σ.π.π της Y είναι

$$f_Y(y) = n\,[F_X(y)]^{n-1} f_X(y) = \begin{cases} \frac{n}{\theta^n} y^{n-1}, & 0<y<\theta \\ 0, & \text{αλλού} \end{cases}$$

$Q = \frac{X_{(n)}}{\theta}$ η σ.π.π της Q είναι.

$$f_Q(q) = \begin{cases} n q^{n-1}, & 0<q<1 \\ 0, & \text{αλλού} \end{cases}$$ (αλλαγή μεταβλητών).

Η Q είναι συνάρτηση του τ.δ, μέσω της $X_{(n)}$ είναι συνάρτηση της θ και η κατανομή της δεν εξαρτάται από το θ. Άρα Q αντιστρέψιμη ποσότητα.

β) $Q = X_{(n)}/\theta$. Ένα δ.ε για τη θ με β.ε $100(1-a)\%$

$$1-a = P(q_1<Q<q_2) = P\left(q_1 < \frac{X_{(n)}}{\theta} < q_2\right) = P\left(\frac{X_{(n)}}{q_2} < \theta < \frac{X_{(n)}}{q_1}\right)$$

Άρα ένα $100(1-a)\%$ Δ.Ε για την Θ είναι το: $\left(\frac{X_{(n)}}{q_2}, \frac{X_{(n)}}{q_1}\right)$

Το Δ.Ε ίσων ουρών για την Θ προκύπτει για q_1, q_2 τέτοια ώστε

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

$$P(Q \leq q_1) = \int_0^{q_1} f_Q(q)\,dq = \int_0^{q_1} n q^{n-1}\,dq = q_1^n$$

$$P(Q \geq q_2) = \int_{q_2}^{1} f_Q(q)\,dq = \int_{q_2}^{1} n q^{n-1}\,dq = 1-q_2^n$$

$$\Rightarrow q_1^n = \frac{a}{2} \Rightarrow q_1 = \sqrt[n]{\frac{a}{2}}.$$

$$q_2^n = 1-\frac{a}{2} \Rightarrow q_2 = \sqrt[n]{1-(a/2)}$$

Άρα το $100(1-a)\%$ Δ.Ε ίσων ουρών για την Θ είναι $\frac{X_{(n)}}{\sqrt[n]{1-(a/2)}}, \frac{X_{(n)}}{\sqrt[n]{a/2}}$.

γ) Εύρεση Δ.Ε ελαχίστου μήκους:

$$\left(\frac{X_{(n)}}{q_2}, \frac{X_{(n)}}{q_1}\right) \qquad l = X_{(n)}\left(\frac{1}{q_1} - \frac{1}{q_2}\right)$$

Ζητούνται τα q_1, q_2 που ελαχιστοποιούν το l υπό τον περιορισμό $P(q_1 \leq Q \leq q_2) = 1-a$

Ο περιορισμός γίνεται

$$P(q_1 \le Q \le q_2) = \int_{q_1}^{q_2} f_Q(q)\,dq = \int_{q_1}^{q_2} n q^{n-1}\,dq = q_2^n - q_1^n$$

$$\Rightarrow q_2^n - q_1^n = 1-a$$

Αρα ζητούνται τα q_1, q_2 που ελαχιστοποιούν το $l^* = \frac{1}{q_1} - \frac{1}{q_2}$ υπο τον περιορισμό $q_2^n - q_1^n = 1-a$.

Παραγωγίζοντας τον περιορισμό ως προς q_1

$$n q_2^{n-1} \frac{dq_2}{dq_1} - n q_1^{n-1} = 0 \Rightarrow \frac{dq_2}{dq_1} = \left(\frac{q_1}{q_2}\right)^{n-1}$$

$$\frac{dl^*}{dq_1} = -\frac{1}{q_1^2} + \frac{1}{q_2^2}\frac{dq_2}{dq_1} = -\frac{1}{q_1^2} + \frac{1}{q_2^2}\left(\frac{q_1}{q_2}\right)^{n-1} < 0 \quad (q_1 < q_2)$$

Αρα l^* ειναι φθίνουσα συνάρτηση του q_1.
Αξιοποιώντας τον περιορισμό:

$$q_1^n = q_2^n - 1 + a \le 1 - 1 + a = a \Rightarrow q_1 \le \sqrt[n]{a}.$$

Αρα το ελάχιστο της l^* ως προς q_1 είναι για $q_1 = \sqrt[n]{a}$
παίρνουμε ότι $q_2^n - (\sqrt[n]{a})^n = 1-a \Rightarrow q_2 = 1$

Αρα το $100(1-a)\%$ δ.ε. για την θ είναι

$$\left(X_{(n)}, \frac{X_{(n)}}{\sqrt[n]{a}}\right)$$

Παρατήρηση:

Εισαγωγή αντιστρεπτής ποσότητα που βασίζεται στο επαρκές στατιστικό.

Ξεκινάμε με το επαρκές στατιστικό και προσπαθούμε με αλλαγή μεταβλητών να δημιουργήσουμε μια συνάρτηση του που θα περιέχει την άγνωστη παράμετρο θ και που η κατανομή της δεν θα εξαρτάται από το θ οπότε θα είναι αντιστρεπτή ποσότητα που θα οδηγεί σε δ.Ε που θα εξαρτάται από το επαρκές στατιστικό.