Διάλεξη 10η

17/05/2019

# ΣΤΑΤΙΣΤΙΚΗ ΣΥΜΠΕΡΑΣΜΑΤΟΛΟΓΙΑ.

(*) Μαθήματα: Αν κλείσει το πανεπ. στις 24/05 τότε το μάθημα θα γίνει στις 22/05 και ώρα 18:00 ενώ την αμέσως επόμενη εβδομάδα (συνέδριο) θα γίνει στις 29/05 και ώρα 18:00.

---

Έστω $X_1, \ldots, X_n$ από $f(x,\theta)$ τότε:

$$Q = -2 \sum \log F_X(x_i) \sim \chi^2_{2n}$$

$$Q = -2 \sum \log(1 - F_X(x_i)) \sim \chi^2_{2n}$$, είναι αντιστρέψιμες.

<u>Παράδειγμα</u>: Έστω τ.δ. $X_1, \ldots, X_n$ από $U(0,\theta)$, $\theta > 0$. Να βρεθεί το δ.ε. ίσων ουρών Β.Ε. $100(1-\alpha)\%$, $0 < \alpha < 1$.

Αν $X \sim U(\alpha, \beta)$ : $F_X(x) = \begin{cases} 0, & x < \alpha \\ \dfrac{x-\alpha}{\beta-\alpha}, & \alpha < x < \beta \\ 1, & x \geq \beta. \end{cases}$

Στην περίπτωση μας : $F(x) = \begin{cases} 0, & x \leq 0 \\ \dfrac{x}{\theta}, & 0 < x < \theta. \\ 1, & x \geq \theta. \end{cases}$

Θεωρώ ως αντιστρέψιμη ποσότητα την:

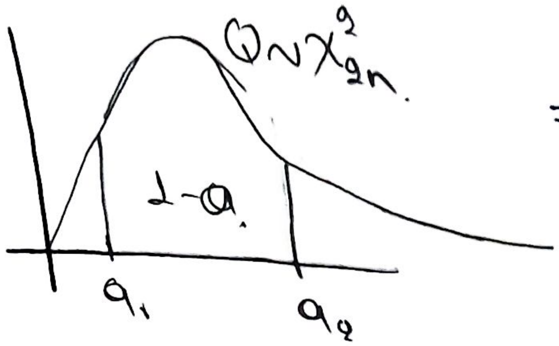
$$Q = -2 \sum_{i=1}^{n} \log F(x_i) \sim \chi^2_{2n}.$$

$$= -2 \sum_{i=1}^{n} \log \frac{x_i}{\theta} = -2 \sum (\log x_i - \log \theta)$$

$$= -2 \sum \log x_i + 2n \log \theta \sim \chi^2_{2n}$$

Αφού η $Q$ αντιστρέψιμη, $\exists\, q_1, q_2$ $(0<q_1<q_2<\infty)$ τ.ω.:

$1-\alpha = P(q_1 < Q < q_2) = P(q_1 < -2\sum \log x_i + 2n\log\theta < q_2) =$




$= P\left(q_1 + 2\sum \log x_i < 2n\log\theta < q_2 + 2\sum \log x_i\right) =$

$= P\left(e^{\frac{q_1 + 2\sum \log x_i}{2n}} < \theta < e^{\frac{q_2 + 2\sum \log x_i}{2n}}\right)$

Άρα από ορισμό του δ.ε. ένα δ.ε. για την $\theta$ με β.ε. $100(1-\alpha)\%$ είναι το $\left(e^{\frac{q_1 + 2\sum \log x_i}{2n}}, e^{\frac{q_2 + 2\sum \log x_i}{2n}}\right)$

Για δ.ε. ίσων ουρών επιλέγω τα $q_1, q_2$ έτσι ώστε

$P(Q > q_2) = \alpha/2$ και $P(Q < q_1) = \alpha/2$.

$\rightarrow P(\chi^2_{2n} > q_2) = \alpha/2$

$\rightsquigarrow q_2 = \chi^2_{2n,\alpha/2}$

$\rightarrow P(\chi^2_{2n} < q_1) = \alpha/2 \Rightarrow 1 - P(\chi^2_{2n} \geq q_1) = \frac{\alpha}{2}$

$\Rightarrow P(\chi^2_{2n} \geq q_1) = 1 - \frac{\alpha}{2} \rightsquigarrow$

$q_1 = \chi^2_{2n, 1-\frac{\alpha}{2}}$

Παρατήρηση: Στην $U(0,\theta)$ επαρκές είναι το $X_{(n)} = \max X_i$. Παρατηρώ ότι το δ.ε. που βασίζεται στην $Q = -2\log\sum F(x_i)$ δεν στηρίζεται στο επαρκές $X_{(n)}$, αλλά στο $\log\sum F(x_i)$ ή στο $\log\sum x_i$. Άρα το δ.ε. που κατασκευάσαμε είναι μη αποδεκτό αφού δεν στηρίζεται στο επαρκές $X_{(n)}$.

(*) Προσπαθώ να βρω καλύτερη λύση στο προηγούμενο παράδειγμα.

**Παράδειγμα:** Έστω τ.δ. $X_1,\dots,X_n$ από $U(0,\theta)$, $\theta>0$.

Έστω $X_{(n)}=\max_i X_i$ (το οποίο είναι επαρκές).

α) Νδο η $Q=\frac{X_{(n)}}{\theta}$ είναι αντιστρεπτή.

β) Να βρεθεί δ.ε. ίσων ουρών για θ.

γ) Να βρεθεί δ.ε. ελαχ. μήκους για την θ.

**Λύση**

α) Έστω $Y=X_{(n)}$, τότε $f_Y(y)=n[F(y)]^{n-1}f(y)$, $0<y<\theta$.

$=n\left(\frac{y}{\theta}\right)^{n-1}\frac{1}{\theta} \Rightarrow f_Y(y)=n\frac{y^{n-1}}{\theta^n}$, $0<y<\theta$.

Αρκεί να βρω την κατανομή της τ.μ. $Q=Y/\theta$.

Θεωρώ τον μετασχηματισμό $q=h(y)=y/\theta$. Για $0<y<\theta$ το $0<q<1$. Η $h$ είναι "1-1" άρα $\exists$ η $h^{-1}$ και είναι

$y=h^{-1}(q)=\theta q$.

$\frac{dh^{-1}(q)}{dq}=\theta\neq 0$. και συνεχής.

Άρα $f_Q(q)=f_Y(h^{-1}(q))\left|\frac{dh^{-1}(q)}{dq}\right|=n\frac{(\cancel{\theta}q)^{n-1}}{\theta^n}\cancel{|\theta|}$

$\Rightarrow f_Q(q)=nq^{n-1}$, $0<q<1$.

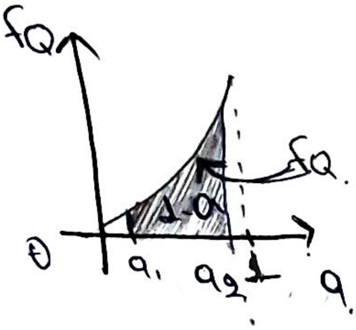
$\Rightarrow$ Q είναι αντιστρεπτή.

Έχω λοιπόν $f_Q(q) = nq^{n-1}, \ 0<q<1$ ή

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

B) Αφού $Q = \frac{X_{(n)}}{\theta}$ είναι αντιστρεπτή υπάρχουν $q_1, q_2$ με

$0<q_1<q_2<1$ τ.ω $1-a = P(q_1 < Q < q_2)$




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

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

Είναι $100(1-a)\%$ δ.ε. για την $\theta$.

Είναι $\left(\frac{X_{(n)}}{q_2}, \frac{X_{(n)}}{q_1}\right)$.

Για δ.ε. ίσων ουρών αναζητώ $q_1, q_2$ τ.ω

$$P(Q>q_2) = \frac{a}{2} = P(Q<q_1)$$

Αρχικά: $\frac{a}{2} = P(Q>q_2) = \int_{q_2}^{1} f_Q(q)\,dq = \int_{q_2}^{1} nq^{n-1}\,dq = n\frac{q^n}{n}\Big|_{q_2}^{1} = 1-q_2^n$

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

Έπειτα: $\frac{a}{2} = P(Q<q_1) = \int_0^{q_1} f_Q(q)\,dq = \int_0^{q_1} nq^{n-1}\,dq = n\cdot\frac{q^n}{n}\Big|_0^{q_1} = q_1^n$

$\Rightarrow q_1 = \sqrt[n]{a/2}$

Το δ.ε. ίσων ουρών: $\left(\frac{X_{(n)}}{\sqrt[n]{1-\frac{a}{2}}}, \frac{X_{(n)}}{\sqrt[n]{\frac{a}{2}}}\right)$

Για την εύρεση δ.ε. ελαχ. μήκους, αναζητώ τα $q_1, q_2$ $(0<q_1<q_2<1)$ τ.ω. να ελαχιστοποιούν το μήκος:

$l = X_{(n)}\left(\frac{1}{q_1} - \frac{1}{q_2}\right)$ υπό τον περιορισμό (π) ότι

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

υπό: $1-a = P(q_1 < Q < q_2) = \int_{q_1}^{q_2} f(q)\,dq = \int_{q_1}^{q_2} n q^{n-1}\,dq$

$= q_2^n - q_1^n$.

Άρα ισοδύναμα ελαχ την $q_2^n - q_1^n = 1-a$

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

Αλλά $\frac{d}{dq_1}(q_2^n - q_1^n) = \frac{1}{q_1}(1-a) = 0 \Rightarrow$

$$\frac{d}{dq_2}(q_2^n)\frac{dq_2}{dq_1} - nq_1^{n-1} = 0 \Rightarrow \frac{dq_2}{dq_1} = \frac{q_1^{n-1}}{q_2^{n-1}} \quad : (2)$$

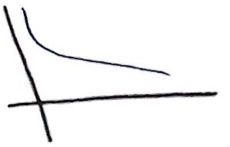
Από (1) κ' (2): $\frac{dl^*}{dq_1} = -\frac{1}{q_1^2} + \frac{1}{q_2^2}\frac{q_1^{n-1}}{q_2^{n-1}}$

$\frac{dl^*}{dq_1} = 0 \Rightarrow \dots\dots \Rightarrow q_1 = q_2$ Δεν μπορεί να ισχ. αφού εξ ορισμού $q_1 < q_2$.

Όμως $\frac{dl^*}{dq_1} = -\frac{1}{q_1^2} + \frac{1}{q_2^2}\frac{q_1^{n-1}}{q_2^{n-1}} < 0$, επειδή $q_1 < q_2$.

Άρα $l^*$ φθίνουσα ως προς $q_1$.

Άρα η $l^*$



παίρνει την μικρότερη τιμή της, στην μεγαλύτερη τιμή που μπορεί να πάρει το $q$.

Από περιορισμό $q_2^n - q_1^n = 1-a \Rightarrow q_1^n = q_2^n - 1 + a$ $\Big\}\Rightarrow$
Αλλά $q_2 < 1$.

$\Rightarrow q_1^n \le 1^n - 1 + a \Rightarrow q_1 \le \sqrt[n]{a}$. Άρα η $\ell^*$ ελαχ ως προς $q_1$ για $q_1 = \sqrt[n]{a}$. Για $q_1 = \sqrt[n]{a}$ το $q_2$ είναι $q_2^n = 1 - a + q_1^n =$

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

---

-Τελευταίο μέρος Στ. Συμπ-.

▶ Έστω τ.δ. $X_1,\dots,X_n$ από $f(x,\theta)$, $\theta \in \Theta$, $\Theta \subseteq \mathbb{R}^n$.

Συμπερασματολογία για την $\theta$:

(I) Εκτίμηση σε σημείο: Στατιστική συνάρτηση
$T = T(X_1,\dots,X_n)$ τ.ω $T(X_1,\dots,X_n) \simeq \theta$, υπό κάποια κριτήρια.

(II) Εκτίμηση σε διάστημα: Δύο στατιστικές συναρτήσεις
$L = L(X_1,\dots,X_n)$, $U = U(X_1,\dots,X_n)$ έτσι ώστε
$1 - a = P(L < \theta < U)$.

(III) Έλεγχοι Στατιστικών υποθέσεων - Στατιστικά τεστ:

Ένας παιδίατρος μελετάει:

Γνωρίζω σε $\mu \simeq 40$. αλλά βλέπει ότι $\mu > 40$.

Έτσι ο παιδίατρος διερωτάται:
Ισχύει $\mu = 40$ ή $\mu > 40$;


πληθυσμός παιδιών ηλικ. 10 ετών
↓
ενδιαφέρει τ.μ. X - Βάρος
$E(X) = \mu$.
$Var(X) = \sigma^2$.

Καταλήγω λοιπόν στις υποθέσεις: $H_0: \mu = 40$ έναντι $H_a: \mu > 40$.
$(H_0: \mu = 40 \lor H_a: \mu > 40)$

(*) Ο έλεγχος Στατιστικών υποθέσεων εφοδιάζει με μεθοδολογίες που επιτρέπουν ότι θα ελέγξουμε 2 στατιστικές υποθέσεις και να δεχτούμε ή να απορρίψουμε κάποια από αυτές με δεδομένη πιθανότητα σφάλματος.

Στατιστικό Τεστ: Μεθοδολογία για αποδοχή ή απόρριψη

## Στοιχεία Στατιστικού Τεστ:

① Στατιστικές υποθέσεις: Έστω τ.δ. $x_1, \dots, x_n$ από πληθυσμό $f(x,\theta), \theta \in \Theta$. Όταν λέμε στατιστικές υποθέσεις: εννοούμε Απόψεις - Εικασίες - Υποθέσεις που διατυπώνονται στην παράμετρο θ.

π.χ

$H_0: \theta = \theta_0$ ($\theta_0$: γνωστό), $H_a: \begin{matrix} \theta > \theta_0 \\ \theta < \theta_0 \\ \theta \neq \theta_0 \end{matrix}$

Μια υπόθεση λέγεται απλή αν καθορίζει πλήρως την κατανομή $f(x,\theta)$, ενώ μια υπόθεση λέγεται σύνθετη αν δεν καθορίζει πλήρως την κατανομή $f(x,\theta)$.

π.χ: Έστω $f(x,\theta) = \theta e^{-\theta x}, x > 0, \theta > 0$

H

$H: \theta = 5 \leftarrow$ απλή

$H: \theta > 5 \leftarrow$ σύνθετη

## 2) Στατιστική Συνάρτηση του Τεστ:

Εννοούμε μια συνάρτηση του τ.δ. $X_1,\dots,X_n$ και πιθανώς της παραμέτρου (όπως αυτή καθορίζεται από την $H_0$) δηλαδή ΣΣΤ: $T=T(X_1,\dots,X_n,\theta)$

(συνήθως συνάρτηση επαρκούς, ΑΟΕΔ, ΕΜΠ)

## 3) Τρόπο Λήψης Απόφασης:

Κρίσιμη περιοχή: $C=\{x: T(x)\in A\}=\{x: T(x) \gtrless k\}$ κρισ. σημείου

**Κανόνα Λήψης απόφασης:** Αν η τιμή της ΣΣΤ περιλαμβάνεται στην Κ.Π. τότε απορρίπτω την $H_0$ αλλιώς δεν μπορώ να απορ. την $H_0$.