# INTERVALE DE INCREDERE SI TESTE
# PENTRU PARAMETRII REPARTITIEI NORMALE
$N\left(\mu,\sigma^2\right)$

## Auxiliar: Repartitii de lucru deduse din repartitia normala ("CHI patrat", "Student", "Fisher")

(a) Repartitia "CHI patrat" cu $r$ grade de libertate $\left(\chi^2\left(r\right)\right)$ a fost introdusa la capitolul "Estimarea parametrilor"

Definitie

Repartitia $Gamma\left(\frac{r}{2},2\right)$, cu $r\in N^*$ se numeste repartitia CHI Patrat cu $r$ grade de libertate, avand densitatea de repartitie

$$f\left(y\right)=\frac{1}{2^{r/2}\Gamma\left(\frac{r}{2}\right)}y^{\frac{r}{2}-1}\exp\left(-\frac{y}{2}\right),\quad y\geq 0$$

$$M\left(Y\right)=r$$
$$D^2\left(Y\right)=2r$$

Proprietate

Fie $X_1,...,X_r$ variabile aleatoare independente, identic repartizate Normal $N\left(0,1\right)$. Atunci

$$Y=\sum_{i=1}^{r}X_i^2$$

este repartizata $\chi^2\left(r\right)$.

(b) Repartitia Student cu $r$ grade de libertate $\left(t\left(r\right)\right)$

Definitie:
Spunem ca o variabila aleatoare $Z$ este repartizata $t\left(r\right)$ daca are densitatea de repartie

$$f\left(z\right)=\frac{\Gamma\left(\frac{r+1}{2}\right)}{\sqrt{r\pi}\Gamma\left(\frac{r}{2}\right)}\left(1+\frac{z^2}{r}\right)^{-(r+1)/2},\quad z\in R$$

Observatii

- Pentru $r = 1$, repartitia $t(1)$ se numeste "repartitia Cauchy" si pentru aceasta nu exista $M(X)$.

$$M(|Z|) = \frac{2}{\pi}\int_0^{\infty} \frac{z}{1+z^2}dz = \frac{1}{\pi}\cdot \lim_{b\to\infty} \ln\left(1+b^2\right) = \infty$$

- Pentru $r = 2$, repartitia $t(2)$ are $M(Z) = 0$, iar $M\left(Z^2\right)$ nu exista.
- Pentru $r > 2$, repartitia $t(r)$ are

$$M(Z) = 0$$
$$D^2(Z) = \frac{r}{r-2}$$

Proprietate

Fie $X$ si $Y$ variabile aleatoare independente, cu $X \sim N(0,1)$ si $Y \sim \chi^2(r)$. Atunci variabila aleatoare

$$Z = \frac{X}{\sqrt{\frac{1}{r}Y}}$$

are repartitia $t(r)$.

Demonstratie:

$$f_{(X,Y)}(x,y) = f_X(x)\cdot f_Y(y) =$$

$$= \frac{1}{2^{(r+1)/2}\sqrt{\pi}\Gamma\left(\frac{r}{2}\right)} y^{\frac{r}{2}-1}\exp\left\{-\frac{x^2}{2}-\frac{y}{2}\right\},\quad x\in R,\ y\geq 0$$

Consideram schimbarea de variabila

$$\left\{\begin{array}{c} z = \frac{x}{\sqrt{\frac{1}{r}y}} \\ y = y \end{array}\right.,\ z\in R,\quad y\geq 0$$

respectiv transformarea inversa

$$\left\{\begin{array}{c} x = z\sqrt{\frac{1}{r}y} \\ y = y \end{array}\right.$$

de Jacobian $\sqrt{y}/\sqrt{r}$. Atunci densitatea de repartite a vectorului aleator $(Z,Y)$ este

$$f_{(Z,Y)}(z,y)=\frac{1}{2^{(r+1)/2}\sqrt{\pi}\Gamma\left(\frac{r}{2}\right)}y^{\frac{r}{2}-1}\exp\left\{-\frac{z^2\cdot y}{2r}-\frac{y}{2}\right\}\cdot\frac{\sqrt{y}}{\sqrt{r}},\quad z\in R,\ y\geq 0$$

Densitatea marginala a lui $Z$ este

$$f_Z(z)=\int_0^\infty f_{(Z,Y)}(z,y)\,dy=$$

$$=\frac{1}{\sqrt{r\pi}\Gamma\left(\frac{r}{2}\right)}\cdot\frac{1}{2^{(r+1)/2}}\int_0^\infty y^{\frac{r+1}{2}-1}\exp\left\{-\frac{y}{2}\left(1+\frac{z^2}{r}\right)\right\}dy$$

Cu schimbarea de variabila

$$t=\frac{y}{2}\left(1+\frac{z^2}{r}\right)$$

obtinem

$$f_Z(z)=\frac{1}{\sqrt{r\pi}\Gamma\left(\frac{r}{2}\right)}\cdot\Gamma\left(\frac{r+1}{2}\right)\left(1+\frac{z^2}{r}\right)^{-(r+1)/2},\ z\in R$$

■

(c) Repartitia Fisher cu $(r_1,r_2)$ grade de libertate $(\mathcal{F}(r_1,r_2))$

Definitie:
Spunem ca o variabila aleatoare $Z$ este repartizata $\mathcal{F}(r_1,r_2)$ daca are densitatea de repartie

$$f(z)=\left(\frac{r_1}{r_2}\right)^{r_1/2}\frac{\Gamma\left(\frac{r_1+r_2}{2}\right)}{\Gamma\left(\frac{r_1}{2}\right)\Gamma\left(\frac{r_2}{2}\right)}\cdot z^{\frac{r_1}{2}-1}\left(1+\frac{r_1}{r_2}z\right)^{-(r_1+r_2)/2},\quad z\geq 0$$

Proprietate

Fie $X$ si $Y$ variabile aleatoare independente, cu $X\sim\chi^2(r_1)$ si $Y\sim\chi^2(r_2)$. Atunci variabila aleatoare

$$Z=\frac{X}{r_1}\Big/\frac{Y}{r_2}$$

are repartita $\mathcal{F}(r_1,r_2)$.

Demonstratie

$$f_{(X,Y)}(x,y) = f_X(x) \cdot f_Y(y) =$$

$$= \frac{1}{2^{(r_1+r_2)/2} \cdot \Gamma\left(\frac{r_1}{2}\right)\Gamma\left(\frac{r_2}{2}\right)} \cdot x^{\frac{r_1}{2}-1} \cdot y^{\frac{r_2}{2}-1} \exp\left(-\frac{x}{2} - \frac{y}{2}\right), \quad x, y \geq 0$$

Consideram schimbarea de variabila

$$\begin{cases} z = \frac{r_2}{r_1} \cdot \frac{x}{y} \\ y = y \end{cases}, \ z \geq 0, \quad y \geq 0$$

respectiv transformarea inversa

$$\begin{cases} x = \frac{r_1}{r_2} yz \\ y = y \end{cases}$$

de Jacobian $r_1 y / r_2$. Atunci densitatea de repartite a vectorului aleator $(Z, Y)$ este

$$f_{(Z,Y)}(z,y) = \frac{1}{2^{(r_1+r_2)/2} \cdot \Gamma\left(\frac{r_1}{2}\right)\Gamma\left(\frac{r_2}{2}\right)} \left(\frac{r_1}{r_2}\right)^{r_1/2} z^{\frac{r_1}{2}-1} y^{\frac{r_1+r_2}{2}-1} \exp\left(-\frac{y}{2}\left(1 + \frac{r_1}{r_2} z\right)\right),$$

$$z \geq 0, \ y \geq 0$$

Densitatea marginala a lui $Z$ este

$$f_Z(z) = \int_0^\infty f_{(Z,Y)}(z,y)\, dy =$$

$$= \left(\frac{r_1}{r_2}\right)^{r_1/2} \frac{1}{\Gamma\left(\frac{r_1}{2}\right)\Gamma\left(\frac{r_2}{2}\right)} \cdot z^{\frac{r_1}{2}-1} \cdot \frac{1}{2^{(r_1+r_2)/2}} \int_0^\infty y^{\frac{r_1+r_2}{2}-1} \exp\left(-\frac{y}{2}\left(1 + \frac{r_1}{r_2} z\right)\right) dy$$

Cu schimbarea de variabila

$$t = \frac{y}{2}\left(1 + \frac{r_1}{r_2} z\right)$$

obtinem

$$f_Z(z) = \left(\frac{r_1}{r_2}\right)^{r_1/2} \frac{\Gamma\left(\frac{r_1+r_2}{2}\right)}{\Gamma\left(\frac{r_1}{2}\right)\Gamma\left(\frac{r_2}{2}\right)} \cdot z^{\frac{r_1}{2}-1} \left(1 + \frac{r_1}{r_2} z\right)^{-(r_1+r_2)/2}, \quad z \geq 0$$

■

INTERVALE DE ESTIMARE (DE INCREDERE)

Definitie

Fie modelul $F_\theta = P_\theta \circ X^{-1}$ cu $\theta \in \Theta \subseteq R$ si fie $X_1, ..., X_n$ variabile aleatore independente, identic repartizate $(F_\theta)$. Fie $\alpha \in (0,1)$ si functiile $A_\alpha, B_\alpha : S^n \longrightarrow R$ cu proprietatile:

$i)$ $A_\alpha, B_\alpha$ sunt masurabile si

$$A_\alpha(x_1, ..., x_n) \leq B_\alpha(x_1, ..., x_n) \;\; \forall (x_1, ..., x_n) \in S^n,$$

$ii)$ are loc relatia

$$P_\theta(A_\alpha(X_1, ..., X_n) \leq \theta \leq B_\alpha(X_1, ..., X_n)) = 1 - \alpha$$

Atunci, pentru datele statistice $(x_1, ..., x_n)$, intervalul

$$C_{n;1-\alpha}(x_1, ..., x_n) = [A_\alpha(x_1, ..., x_n) \;\; , \;\; B_\alpha(x_1, ..., x_n)]$$

se numeste interval de estimare pentru $\theta$, cu coeficientul de incredere $(1-\alpha)$ (sau interval de incredere pentru $\theta$).

Propozitie

Fie modelul $F_\theta = P_\theta \circ X^{-1}$ cu $\theta \in \Theta \subseteq R$ si fie $X_1, ..., X_n$ variabile aleatore independente, identic repartizate $(F_\theta)$. Presupunem ca exista o functie

$$g : S^n \times \Theta \longrightarrow R$$

cu urmatoarele proprietati:

- $g((x_1, ..., x_n), \cdot)$ continua si strict monotona ca functie in $\theta$, $\forall (x_1, ..., x_n)$
- $g(\cdot, \theta)$ masurabila ca functie in $(x_1, ..., x_n)$, $\forall \theta$ si variabila aleatoare $g((X_1, ..., X_n), \theta)$ are repartitia independenta de $\theta$ (o notam $G$).

Atunci, pentru orice $\alpha \in (0,1)$ arbitrar fixat, exista $C_{n;1-\alpha}(x_1, ..., x_n)$ interval de incredere pentru $\theta$.

Demonstratie:

Fie $\alpha \in (0,1)$ si $\theta \in \Theta$ arbitrari, fixati. Fie $a(\alpha), b(\alpha)$ doua cuantile ale repartitiei $G = P_\theta \circ g^{-1}$ asa incat

$$P_\theta(a(\alpha) \leq g((X_1, ..., X_n), \theta) \leq b(\alpha)) = G(b) - G(a) = 1 - \alpha$$

Rezolvand doua inegalitati in $\theta$, putem scrie

$$\{\omega \mid a(\alpha) \leq g((X_1, ..., X_n)(\omega), \theta) \leq b(\alpha)\}$$
$$= \{\omega \mid A_\alpha(X_1, ..., X_n)(\omega) \leq \theta \leq B_\alpha(X_1, ..., X_n)(\omega)\}$$

Rezulta ca

$$C_{n;1-\alpha}(x_1, ..., x_n) = [A_\alpha(x_1, ..., x_n)\ ,\ B_\alpha(x_1, ..., x_n)]$$

este un interval de estimare pentru $\theta$ cu coeficient de incredere $(1-\alpha)$.

■

Comentariu:

Cuantilele $a(\alpha), b(\alpha)$ nu sunt unic determinate prin conditia $G(b) - G(a) = 1 - \alpha$, deci nici intervalul de incredere nu este unic. Este de interes sa construim cel mai scurt interval de estimare cu coeficient de incredere dat.

## TESTE BAZATE PE INTERVALE DE INCREDERE

## PENTRU IPOTEZA SIMPLA CU ALTERNATIVA COMPUSA

$$H : \{\theta = \theta_0\},\ H_A : \{\theta \neq \theta_0\}$$

Ne plasam in conditiile propozitiei anterioare, care asigura existenta unui interval de incredere pentru $\theta$.

Pornim de la relatia

$$P_{\theta_0}(a(\alpha) \leq g((X_1, ..., X_n), \theta_0) \leq b(\alpha)) = 1 - \alpha$$

Alegem REGIUNEA DE ACCEPTARE a ipotezei $H : \{\theta = \theta_0\}$ la pragul de semnificatie $\alpha$

$$A_{n;1-\alpha}(\theta_0) = \{(x_1, ..., x_n) \mid a \leq g((x_1, ..., x_n), \theta_0) \leq b\}$$

si REGIUNEA CRITICA pentru $H : \{\theta = \theta_0\}$ la pragul de semnificatie $\alpha$

$$B = A^C_{n;1-\alpha}(\theta_0)$$

Probabilitatea erorii de I tip este egala cu $\alpha$,

$$P_{\theta_0}((X_1, ..., X_n) \in B) = 1 - (1 - \alpha) = \alpha$$

Functia caracteristica operatoare a testului bazat pe aceasta regiune critica este

$$OC(\theta) = P_\theta\left((X_1, ..., X_n) \in A_{n;1-\alpha}(\theta_0)\right)$$

## APLICATIA 1

### Interval de incredere si testul "$z$" pentru media unei repartii normale cu dispersie cunoascuta

Modelul: $P_\mu \circ X^{-1} = N(\mu, \sigma^2)$, $\sigma^2$ cunoscut, $\mu \in R$
Observatii: $X_1, ..., X_n$ v.i.i.r. $N(\mu, \sigma^2)$

$$\overline{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$$
$$\frac{\sqrt{n}\left(\overline{X} - \mu\right)}{\sigma} \sim N(0,1)$$

Functia

$$g((x_1, ..., x_n); \mu) = \frac{\sqrt{n}(\overline{x} - \mu)}{\sigma}$$

indeplineste conditiile din constructiile anterioare.

Pentru $\alpha \in (0,1)$ fixat, fie $a, b$ doua cuantile ale repartitiei $N(0,1)$ asa incat

$$P_\mu\left(a \leq \frac{\sqrt{n}\left(\overline{X} - \mu\right)}{\sigma} \leq b\right) = 1 - \alpha$$

$$\left\{a \leq \frac{\sqrt{n}(\overline{x} - \mu)}{\sigma} \leq b\right\} = \left\{\overline{x} - b\frac{\sigma}{\sqrt{n}} \leq \mu \leq \overline{x} - a\frac{\sigma}{\sqrt{n}}\right\}$$

$$C_{n;1-\alpha}(x_1, ..., x_n) = \left[\overline{x} - b\frac{\sigma}{\sqrt{n}}\ ,\ \overline{x} - a\frac{\sigma}{\sqrt{n}}\right]$$

Lungimea acestui interval de incredere este

$$l = \frac{\sigma}{\sqrt{n}}(b - a)$$

Determinam acum cel mai scurt interval de incredere pentru $\mu$,cu coeficientul de incredere $(1 - \alpha)$.

Utilizand faptul ca $b = b(a)$, conditiile

$$\begin{cases} F_{N(0,1)}(b) - F_{N(0,1)}(a) = 1-\alpha \\ \min\left\{\frac{\sigma}{\sqrt{n}}(b-a)\right\} \end{cases}$$

conduc la

$$\begin{cases} f_{N(0,1)}(b)\cdot\frac{db}{da} - f_{N(0,1)}(a) = 0 \\ \frac{db}{da} - 1 = 0 \end{cases},$$

de unde obtinem

$$f_{N(0,1)}(b) = f_{N(0,1)}(a)$$

Rezulta

$$b = z_{1-\frac{\alpha}{2}},\ a = -z_{1-\frac{\alpha}{2}}$$

si deci cel mai scurt interval de incredere este

$$C^*_{n;1-\alpha}(x_1,...,x_n) = \left[\overline{x} - z_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}\ ,\ \overline{x} + z_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}\right]$$

Consideram acum ipoteza $H : \{\mu = \mu_0\}$ cu alternativa $H_A : \{\mu \neq \mu_0\}$

$$P_{\mu_0}\left(-z_{1-\frac{\alpha}{2}} \leq \frac{\sqrt{n}\left(\overline{X}-\mu_0\right)}{\sigma} \leq z_{1-\frac{\alpha}{2}}\right) = 1-\alpha$$

$$A_{n;1-\alpha}(\mu_0) = \left\{(x_1,...,x_n) \mid -z_{1-\frac{\alpha}{2}} \leq \frac{\sqrt{n}\left(\overline{x}-\mu_0\right)}{\sigma} \leq z_{1-\frac{\alpha}{2}}\right\}$$
$$= \left\{\mu_0 - z_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}} \leq \overline{x} \leq \mu_0 + z_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}\right\}$$

Testul "$z$" se bazeaza pe regiunea critica

$$B = A^C_{n;1-\alpha}(\mu_0)$$

$$P_{\mu_0}((X_1,...,X_n)\in B) = \alpha$$

$$OC(\mu) = P_\mu\left(-z_{1-\frac{\alpha}{2}} \leq \frac{\sqrt{n}\left(\overline{X}-\mu_0\right)}{\sigma} \leq z_{1-\frac{\alpha}{2}}\right) =$$

$$= P_\mu\left(-z_{1-\frac{\alpha}{2}} \leq \frac{\sqrt{n}\left(\overline{X}-\mu\right)}{\sigma} + \frac{\sqrt{n}\left(\mu-\mu_0\right)}{\sigma} \leq z_{1-\frac{\alpha}{2}}\right) =$$

$$= F_{N(0,1)}\left(z_{1-\frac{\alpha}{2}} - \frac{\sqrt{n}\left(\mu-\mu_0\right)}{\sigma}\right) - F_{N(0,1)}\left(-z_{1-\frac{\alpha}{2}} - \frac{\sqrt{n}\left(\mu-\mu_0\right)}{\sigma}\right)$$

# APLICATIA 2

## Interval de incredere si testul "$t$" pentru media unei repartii normale cu dispersie necunoascuta

Modelul: $P_\mu \circ X^{-1} = N\left(\mu, \sigma^2\right)$, $\sigma^2$ necunoscut, $\mu \in R$
Observatii: $X_1, ..., X_n$ v.i.i.r. $N\left(\mu, \sigma^2\right)$
La "estimarea parametrilor" s-a demonstrat:

Proprietate

Fie $X_1, ..., X_n$ variabile aleatoare independente, identic repartizate $N\left(\mu, \sigma^2\right)$ si fie E.V.M.

$$\widehat{\mu}_{VM} = \overline{X}$$
$$\widehat{\sigma^2}_{VM} = \frac{1}{n}\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2$$

Atunci

$$\widehat{\mu}_{VM} = \overline{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right),$$
$$\frac{n}{\sigma^2} \cdot \widehat{\sigma^2}_{VM} \sim \chi^2\left(n-1\right)$$

si cele doua componente ale E.V.M. sunt independente.

Constructie:

$$S^2 = \frac{n}{n-1}\widehat{\sigma^2}_{VM}$$

$$\frac{\sqrt{n}\left(\overline{X} - \mu\right)}{\sigma} \sim N\left(0,1\right)$$
$$\frac{n-1}{\sigma^2} \cdot S^2 \sim \chi^2\left(n-1\right)$$

independenta

$$Z = \frac{\sqrt{n}\left(\overline{X} - \mu\right)}{\sigma} \Bigg/ \sqrt{\frac{1}{n-1}\frac{n-1}{\sigma^2} \cdot S^2} = \frac{\sqrt{n}\left(\overline{X} - \mu\right)}{S} \ \sim \ t\left(n-1\right)$$

Functia

$$g\left(\left(x_1, ..., x_n\right); \mu\right) = \frac{\sqrt{n}\left(\overline{x} - \mu\right)}{s}$$

indeplineste conditiile din constructiile anterioare.

Pentru $\alpha \in (0,1)$ fixat, fie $a, b$ doua cuantile ale repartitiei $t(n-1)$ asa incat

$$P_\mu\left(a \leq \frac{\sqrt{n}\left(\overline{X}-\mu\right)}{S} \leq b\right) = 1-\alpha$$

$$\left\{a \leq \frac{\sqrt{n}\left(\overline{x}-\mu\right)}{s} \leq b\right\} = \left\{\overline{x} - b\frac{s}{\sqrt{n}} \leq \mu \leq \overline{x} - a\frac{s}{\sqrt{n}}\right\}$$

$$C_{n;1-\alpha}\left(x_1, ..., x_n\right) = \left[\overline{x} - b\frac{s}{\sqrt{n}}\ ,\ \overline{x} - a\frac{s}{\sqrt{n}}\right]$$

Lungimea acestui interval de incredere este

$$l = \frac{s}{\sqrt{n}}\left(b-a\right)$$

Determinam acum cel mai scurt interval de incredere pentru $\mu$,cu coeficientul de incredere $(1-\alpha)$.

Utilizand faptul ca $b = b(a)$, conditiile

$$\left\{\begin{array}{c} F_{t(n-1)}\left(b\right) - F_{t(n-1)}\left(a\right) = 1-\alpha \\ \min\left\{\frac{s}{\sqrt{n}}\left(b-a\right)\right\} \end{array}\right.$$

conduc la

$$\left\{\begin{array}{c} f_{t(n-1)}\left(b\right)\cdot\frac{db}{da} - f_{t(n-1)}\left(a\right) = 0 \\ \frac{db}{da} - 1 = 0 \end{array}\right. ,$$

de unde obtinem

$$f_{t(n-1)}\left(b\right) = f_{t(n-1)}\left(a\right)$$

Rezulta

$$b = t_{n-1;1-\frac{\alpha}{2}},\ a = -t_{n-1;1-\frac{\alpha}{2}}$$

si deci cel mai scurt interval de incredere este

$$C^*_{n;1-\alpha}\left(x_1, ..., x_n\right) = \left[\overline{x} - t_{n-1;1-\frac{\alpha}{2}}\frac{s}{\sqrt{n}}\ ,\ \overline{x} + t_{n-1;1-\frac{\alpha}{2}}\frac{s}{\sqrt{n}}\right]$$

Consideram acum ipoteza $H : \{\mu = \mu_0\}$ cu alternativa $H_A : \{\mu \neq \mu_0\}$

$$P_{\mu_0}\left(-t_{n-1;1-\frac{\alpha}{2}} \leq \frac{\sqrt{n}\left(\overline{X}-\mu_0\right)}{S} \leq t_{n-1;1-\frac{\alpha}{2}}\right) = 1-\alpha$$

$$A_{n;1-\alpha}(\mu_0) = \left\{ (x_1, ..., x_n) \mid -t_{n-1;1-\frac{\alpha}{2}} \leq \frac{\sqrt{n}(\overline{x} - \mu_0)}{s} \leq t_{n-1;1-\frac{\alpha}{2}} \right\}$$
$$= \left\{ \mu_0 - t_{n-1;1-\frac{\alpha}{2}} \frac{s}{\sqrt{n}} \leq \overline{x} \leq \mu_0 + t_{n-1;1-\frac{\alpha}{2}} \frac{s}{\sqrt{n}} \right\}$$

Testul "$t$" se bazeaza pe regiunea critica

$$B = A^C_{n;1-\alpha}(\mu_0)$$

$$P_{\mu_0}((X_1, ..., X_n) \in B) = \alpha$$

$$OC(\mu) = P_\mu \left( -t_{n-1;1-\frac{\alpha}{2}} \leq \frac{\sqrt{n}(\overline{X} - \mu_0)}{S} \leq t_{n-1;1-\frac{\alpha}{2}} \right) =$$

$$= P_\mu \left( -t_{n-1;1-\frac{\alpha}{2}} \leq \frac{\sqrt{n}(\overline{X} - \mu)}{S} + \frac{\sqrt{n}(\mu - \mu_0)}{S} \leq t_{n-1;1-\frac{\alpha}{2}} \right) =$$

$$= F_{t(n-1)} \left( t_{n-1;1-\frac{\alpha}{2}} - \frac{\sqrt{n}(\mu - \mu_0)}{s} \right) - F_{t(n-1)} \left( -t_{n-1;1-\frac{\alpha}{2}} - \frac{\sqrt{n}(\mu - \mu_0)}{s} \right)$$

Functia din R: t.test(x,...)

t.test(x, alternative = c("two.sided", "less", "greater"), mu = 0, conf.level = 0.95, ...)

Arguments

x a numeric vector of data values.

alternative a character string specifying the alternative hypothesis, must be one of "two.sided" (default), "greater" or "less".

mu a number indicating the true value of the mean

conf.level confidence level of the interval.

## APLICATIA 3

### Interval de incredere si testul "CHI patrat" pentru dispersia unei repartii normale cu medie cunoascuta

Modelul: $P_\mu \circ X^{-1} = N\left(\mu, \sigma^2\right)$, $\mu$ cunoscut, $\sigma^2 \in (0, \infty)$
Observatii: $X_1, ..., X_n$ v.i.i.r. $N\left(\mu, \sigma^2\right)$. Variabilele aleatoare

$$\frac{X_i - \mu}{\sigma}, \ i = 1, ..., n$$

sunt i.i.r. $N(0,1)$. Rezulta ca

$$\frac{1}{\sigma^2} \sum_{i=1}^{n} \left(X_i - \mu\right)^2 \sim \chi^2(n).$$

Functia

$$g\left(\left(x_1, ..., x_n\right); \sigma^2\right) = \frac{1}{\sigma^2} \sum_{i=1}^{n} \left(x_i - \mu\right)^2$$

indeplineste conditiile din constructiile anterioare.

Pentru $\alpha \in (0,1)$ fixat, fie $0 < a < b$ doua cuantile ale repartitiei $\chi^2(n)$ asa incat

$$P_{\sigma^2}\left(a \leq \frac{1}{\sigma^2} \sum_{i=1}^{n} \left(X_i - \mu\right)^2 \leq b\right) = 1 - \alpha$$

$$\left\{a \leq \frac{1}{\sigma^2} \sum_{i=1}^{n} \left(x_i - \mu\right)^2 \leq b\right\} = \left\{\frac{1}{b} \sum_{i=1}^{n} \left(x_i - \mu\right)^2 \leq \sigma^2 \leq \frac{1}{a} \sum_{i=1}^{n} \left(x_i - \mu\right)^2\right\}$$

$$C_{n;1-\alpha}\left(x_1, ..., x_n\right) = \left[\frac{1}{b} \sum_{i=1}^{n} \left(x_i - \mu\right)^2 \ , \ \frac{1}{a} \sum_{i=1}^{n} \left(x_i - \mu\right)^2\right]$$

Lungimea acestui interval de incredere este

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

Cautam cel mai scurt interval de incredere pentru $\sigma^2$,cu coeficientul de incredere $(1 - \alpha)$.
Utilizand faptul ca $b = b(a)$, conditiile

$$\begin{cases} F_{\chi^2(n)}(b) - F_{\chi^2(n)}(a) = 1 - \alpha \\ \min\left\{\sum\limits_{i=1}^{n} \left(x_i - \mu\right)^2 \left(\frac{1}{b} - \frac{1}{a}\right)\right\} \end{cases}$$

conduc la

$$\begin{cases} f_{\chi^2(n)}(b) \cdot \frac{db}{da} - f_{\chi^2(n)}(a) = 0 \\ -\frac{1}{b^2} \cdot \frac{db}{da} + \frac{1}{a^2} = 0 \end{cases},$$

de unde rezulta

$$b^2 \cdot f_{\chi^2(n)}(b) = a^2 \cdot f_{\chi^2(n)}(a)$$

Aceasta ecuatie nu are o solutie analitica explicita, deci nu putem obtine forma explicita a celui mai scurt interval de incredere pentru $\sigma^2$, cu coeficientul de incredere $(1-\alpha)$.

Prin CONVENTIE, lucram cu

$$C_{n;1-\alpha}(x_1, ..., x_n) = \left[\frac{1}{h_{n;1-\frac{\alpha}{2}}} \sum_{i=1}^{n} (x_i - \mu)^2 \ , \ \frac{1}{h_{n;\frac{\alpha}{2}}} \sum_{i=1}^{n} (x_i - \mu)^2\right],$$

unde $h_{n;\frac{\alpha}{2}}$ si $h_{n;1-\frac{\alpha}{2}}$ sunt cuantile ale repartitiei $\chi^2(n)$.

Consideram acum ipoteza $H : \{\sigma^2 = \sigma_0^2\}$ cu alternativa $H_A : \{\sigma^2 \neq \sigma_0^2\}$

$$P_{\sigma_0^2}\left(h_{n;\frac{\alpha}{2}} \leq \frac{1}{\sigma_0^2} \sum_{i=1}^{n} (X_i - \mu)^2 \leq h_{n;1-\frac{\alpha}{2}}\right) = 1 - \alpha$$

$$\begin{aligned} A_{n;1-\alpha}\left(\sigma_0^2\right) &= \left\{(x_1, ..., x_n) \mid h_{n;\frac{\alpha}{2}} \leq \frac{1}{\sigma_0^2} \sum_{i=1}^{n} (x_i - \mu)^2 \leq h_{n;1-\frac{\alpha}{2}}\right\} \\ &= \left\{\sigma_0^2 \cdot h_{n;\frac{\alpha}{2}} \leq \sum_{i=1}^{n} (x_i - \mu)^2 \leq \sigma_0^2 \cdot h_{n;1-\frac{\alpha}{2}}\right\} \end{aligned}$$

Testul "*CHI patrat*" se bazeaza pe regiunea critica

$$B = A_{n;1-\alpha}^C\left(\sigma_0^2\right)$$

$$P_{\sigma_0^2}\left((X_1, ..., X_n) \in B\right) = \alpha$$

$$\begin{aligned} OC\left(\sigma^2\right) &= P_{\sigma^2}\left(h_{n;\frac{\alpha}{2}} \leq \frac{1}{\sigma_0^2} \sum_{i=1}^{n} (X_i - \mu)^2 \leq h_{n;1-\frac{\alpha}{2}}\right) = \\ &= P_{\sigma^2}\left(h_{n;\frac{\alpha}{2}} \cdot \frac{\sigma_0^2}{\sigma^2} \leq \frac{1}{\sigma^2} \sum_{i=1}^{n} (X_i - \mu)^2 \leq h_{n;1-\frac{\alpha}{2}} \cdot \frac{\sigma_0^2}{\sigma^2}\right) = \\ &= F_{\chi^2(n)}\left(h_{n;1-\frac{\alpha}{2}} \cdot \frac{\sigma_0^2}{\sigma^2}\right) - F_{\chi^2(n)}\left(h_{n;\frac{\alpha}{2}} \cdot \frac{\sigma_0^2}{\sigma^2}\right) \end{aligned}$$

# APLICATIA 4

## Interval de incredere si testul "CHI patrat" pentru dispersia unei repartii normale cu medie necunoscuta

Modelul: $P_\mu \circ X^{-1} = N\left(\mu, \sigma^2\right), \mu \in R$ necunoscut, $\sigma^2 \in (0, \infty)$
Observatii: $X_1, ..., X_n$ v.i.i.r. $N\left(\mu, \sigma^2\right)$. Am demonstrat ca

$$\frac{1}{\sigma^2} \sum_{i=1}^{n} \left(X_i - \overline{X}\right)^2 \sim \chi^2 (n-1).$$

Functia

$$g\left((x_1, ..., x_n); \sigma^2\right) = \frac{1}{\sigma^2} \sum_{i=1}^{n} (x_i - \overline{x})^2 = \frac{(n-1) \cdot s^2}{\sigma^2}$$

indeplineste conditiile din constructiile anterioare.

Pentru $\alpha \in (0,1)$ fixat, fie $h_{n-1;\frac{\alpha}{2}}$ si $h_{n-1;1-\frac{\alpha}{2}}$ cuantile ale repartitiei $\chi^2 (n-1)$. Ca si in Aplicatia 3, obtinem

$$C_{n;1-\alpha} (x_1, ..., x_n) = \left[ \frac{(n-1) \cdot s^2}{h_{n-1;1-\frac{\alpha}{2}}} , \frac{(n-1) \cdot s^2}{h_{n-1;\frac{\alpha}{2}}} \right]$$

Consideram acum ipoteza $H : \{\sigma^2 = \sigma_0^2\}$ cu alternativa $H_A : \{\sigma^2 \neq \sigma_0^2\}$

$$P_{\sigma_0^2} \left( h_{n-1;\frac{\alpha}{2}} \leq \frac{1}{\sigma_0^2} \sum_{i=1}^{n} \left(X_i - \overline{X}\right)^2 \leq h_{n-1;1-\frac{\alpha}{2}} \right) = 1 - \alpha$$

$$\begin{aligned} A_{n;1-\alpha}\left(\sigma_0^2\right) &= \left\{ (x_1, ..., x_n) \mid h_{n-1;\frac{\alpha}{2}} \leq \frac{(n-1) \cdot s^2}{\sigma_0^2} \leq h_{n-1;1-\frac{\alpha}{2}} \right\} \\ &= \left\{ \sigma_0^2 \cdot \frac{h_{n-1;\frac{\alpha}{2}}}{n-1} \leq s^2 \leq \sigma_0^2 \cdot \frac{h_{n-1;1-\frac{\alpha}{2}}}{n-1} \right\} \end{aligned}$$

Testul "*CHI patrat*" se bazeaza pe regiunea critica

$$B = A^C_{n;1-\alpha}\left(\sigma_0^2\right)$$

$$P_{\sigma_0^2} ((X_1, ..., X_n) \in B) = \alpha$$

$$
\begin{aligned}
OC\left(\sigma^2\right) &= P_{\sigma^2}\left(h_{n-1;\frac{\alpha}{2}} \leq \frac{1}{\sigma_0^2}\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2 \leq h_{n-1;1-\frac{\alpha}{2}}\right) = \\
&= P_{\sigma^2}\left(h_{n-1;\frac{\alpha}{2}}\cdot\frac{\sigma_0^2}{\sigma^2} \leq \frac{1}{\sigma^2}\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2 \leq h_{n-1;1-\frac{\alpha}{2}}\cdot\frac{\sigma_0^2}{\sigma^2}\right) = \\
&= F_{\chi^2(n-1)}\left(h_{n-1;1-\frac{\alpha}{2}}\cdot\frac{\sigma_0^2}{\sigma^2}\right) - F_{\chi^2(n-1)}\left(h_{n-1;\frac{\alpha}{2}}\cdot\frac{\sigma_0^2}{\sigma^2}\right)
\end{aligned}
$$

## APLICATIA 5

## TESTUL FISHER PENTRU DREAPTA DE REGRESIE

La capitolul "Regresie" am stabilit urmatoarele rezultate:

- Variabila aleatoare
$$SS_{resid} = \sum_{i=1}^{n} \left( X_i - \widehat{a} - \widehat{b} y_i \right)^2$$
are proprietatea
$$\frac{1}{\sigma_x^2 (1-\rho^2)} \cdot SS_{resid} \quad \sim \quad \chi^2 (n-2)$$

- Daca $b = 0$, atunci
$$\frac{1}{\sigma_x^2 (1-\rho^2)} \cdot SS_{regresie} \sim \chi^2 (1)$$
$$\frac{1}{\sigma_x^2 (1-\rho^2)} \cdot SS_{total} \sim \chi^2 (n-1)$$
iar variabilele $\frac{1}{\sigma_x^2(1-\rho^2)} \cdot SS_{regresie}$ si $\frac{1}{\sigma_x^2(1-\rho^2)} \cdot SS_{resid}$ sunt independente.

Formulam ipoteza $H : \{b = 0\}$ cu alternativa $H_A : \{b \neq 0\}$.
Daca $H$ este adevarata, atunci variabila aleatoare

$$Z = \frac{1}{\sigma_x^2 (1-\rho^2)} \cdot SS_{regresie} \Big/ \frac{1}{n-2} \cdot \frac{1}{\sigma_x^2 (1-\rho^2)} \cdot SS_{resid} \overset{notat}{=} \frac{\overline{SS_{regresie}}}{\overline{SS_{resid}}}$$

are o repartitie Fisher cu $(1, n-2)$ grade de libertate.
Pentru $\alpha \in (0,1)$ arbitrar fixat, fie $f_{(1,n-2);1-\alpha}$ cuantila de rang $(1-\alpha)$ a repartitiei Fisher cu $(1, n-2)$ grade de libertate.

TESTUL FISHER: Regiunea critica pentru $H : \{b = 0\}$ este
$$B = \left\{ \frac{\overline{SS_{regresie}}}{\overline{SS_{resid}}} \geq f_{(1,n-2);1-\alpha} \right\}$$

$$P_{(b=0)}\left(\frac{\overline{SS_{regresie}}}{\overline{SS_{resid}}} \geq f_{(1,n-2);1-\alpha}\right) = \alpha$$

Acest test este implementat in functia "*anova*" din R.

Testul Fisher prezentat aici este echivalent cu un test "*t*", bazat pe urmatoarele fapte:

$$\frac{\widehat{b}}{\sqrt{\frac{\sigma_x^2(1-\rho^2)}{\sum\limits_{i=1}^{n}(y_i-\overline{y})^2}}} \sim N(0,1)$$

$$SS_{regresie} = \left(\widehat{b}\right)^2 \sum_{i=1}^{n} (y_i - \overline{y})^2$$

$$\frac{1}{\sigma_x^2(1-\rho^2)} \cdot SS_{resid} \quad \sim \quad \chi^2(n-2)$$

$SS_{regresie}$ si $SS_{resid}$ sunt variabile aleatoare independente, ceea ce implica $\widehat{b}$ si $SS_{resid}$ sunt variabile aleatoare independente. Atunci

$$\frac{\widehat{b}}{\sqrt{\frac{\sigma_x^2(1-\rho^2)}{\sum\limits_{i=1}^{n}(y_i-\overline{y})^2}}} \Bigg/ \sqrt{\frac{1}{n-2} \cdot \frac{1}{\sigma_x^2(1-\rho^2)} \cdot SS_{resid}} \sim t(n-2)$$

TESTUL "t": Regiunea critica pentru $H : \{b = 0\}$ la pragul de semnificatie $\alpha$ este

$$B = \left\{ \frac{\widehat{b} \cdot \sqrt{(n-2)\sum\limits_{i=1}^{n}(y_i-\overline{y})^2}}{\sqrt{SS_{resid}}} \geq t_{n-2;1-\alpha} \right\},$$

unde $t_{n-2;1-\alpha}$ este cuantila de rang $(1-\alpha)$ a repartitiei $t(n-2)$.

Si acest test este implementat in functia "*anova*" din R.

# APLICATIA 6

## COMPARAREA TRATAMENTELOR

## (COMPARAREA PARAMETRILOR A DOUA REPARTITII NORMALE)

PROBLEMA DE BIOSTATISTICA:

- Caracteristica de interes care este investigata poate fi modelata printr-o variabila aleatoare cu repartitie normala $N(\mu, \sigma^2)$(ex: nivelul colesterolului, nivelul tensiunii arteriale sistolice, nivelul hemoglobinei, etc.)
- Exista doua tratamente posibile $T_1$si $T_2$. Eventual $T_1 =$"tratament" si $T_2 =$"placebo".
- Se considera doua loturi independente, formate din pacienti suferind de aceeasi boala, selectati in mod independent dintr-o populatie bine definita (ex: barbati, din mediul urban, in varsta 40 - 50 ani, supraponderali).
- Pacientilor din primul lot li se administreaza $T_1$si celor din al doilea lot li se administreaza $T_2$.Experimentul este "blind", adica pacientii nu stiu ca primesc tratamente diferite.
- Se doreste identificarea situatiei in care se obtin raspunsuri diferite la cele doua tratamente.

Model: $T_1 = X_1 \sim N(\mu_1, \sigma_1^2)\,;\, T_2 = X_2 \sim N(\mu_2, \sigma_2^2)\,,\, X_1, X_2$ variabile aleatoare independente

Observatii:

$X_{11}, X_{12}, ..., X_{1n}$ $v.a.i.i.r.N(\mu_1, \sigma_1^2)$

$X_{21}, X_{22}, ..., X_{2m}$ $v.a.i.i.r.N(\mu_2, \sigma_2^2)$

$\{X_{11}, X_{12}, ..., X_{1n}\}, \{X_{21}, X_{22}, ..., X_{2m}\}$ familii independente

Ipoteze ce urmeaza a fi testate:

$$H_1 : \{\sigma_1^2 = \sigma_2^2\}\ ,\ H_{1A} : \{\sigma_1^2 \neq \sigma_2^2\}$$
$$H_2 : \{\mu_1 = \mu_2\}\ ,\ H_{2A} : \{\mu_1 \neq \mu_2\}$$

Reamintim proprietatile E.V.M. pentru parametrii repartitiei normale:

$$\overline{X_1} = \frac{1}{n}\sum_{j=1}^{n} X_{1j} \ \sim \ N\left(\mu_1, \frac{\sigma_1^2}{n}\right)$$

$$S_1^2 = \frac{1}{n-1}\sum_{j=1}^{n}\left(X_{1j} - \overline{X_1}\right)^2 \ ; \ \frac{n-1}{\sigma_1^2}\cdot S_1^2 \ \sim \ \chi^2\left(n-1\right)$$

$$\overline{X_1}, \ \frac{n-1}{\sigma_1^2}\cdot S_1^2 \ \text{ independente}$$

$$\overline{X_2} = \frac{1}{m}\sum_{j=1}^{m} X_{2j} \ \sim \ N\left(\mu_2, \frac{\sigma_2^2}{m}\right)$$

$$S_2^2 = \frac{1}{m-1}\sum_{j=1}^{m}\left(X_{2j} - \overline{X_2}\right)^2 \ ; \ \frac{m-1}{\sigma_2^2}\cdot S_2^2 \ \sim \ \chi^2\left(m-1\right)$$

$$\overline{X_2}, \ \frac{m-1}{\sigma_2^2}\cdot S_2^2 \ \text{ independente}$$

$(a)$ Testul Fisher de comparare a dispersiilor,
$H_1 : \left\{\sigma_1^2 = \sigma_2^2\right\}$ , $H_{1A} : \left\{\sigma_1^2 \neq \sigma_2^2\right\}$

Folosind asociativitatea independentei, avem

$$\frac{1}{n-1}\cdot\frac{n-1}{\sigma_1^2}\cdot S_1^2 \Big/ \frac{1}{m-1}\cdot\frac{m-1}{\sigma_2^2}\cdot S_2^2 \ = \frac{\sigma_2^2}{\sigma_1^2}\cdot\frac{S_1^2}{S_2^2} \ \sim \ \mathcal{F}\left(n-1, m-1\right)$$

Reparametrizam si rescriem ipotezele $H_1, H_{1A}$ :

$$\gamma = \frac{\sigma_2^2}{\sigma_1^2}$$

$$H_1 : \{\gamma = 1\} \ , \ H_{1A} : \{\gamma \neq 1\}$$

Daca ipoteza $H_1$ este adevarata, atunci $S_1^2/S_2^2 \ \sim \ \mathcal{F}\left(n-1, m-1\right)$.
Pentru $\alpha \in (0,1)$ arbitrar fixat, fie $f_{1,\alpha}$ si $f_{2,\alpha}$ cuantile ale repartitiei $\mathcal{F}\left(n-1, m-1\right)$, cu proprietatea

$$F_{\mathcal{F}(n-1,m-1)}\left(f_{2,\alpha}\right) - F_{\mathcal{F}(n-1,m-1)}\left(f_{1,\alpha}\right) = 1 - \alpha$$

Facem observatia ca aceasta relatie determina unic cuantilele pentru ca

$$Z \sim \mathcal{F}(n-1, m-1) \implies \frac{1}{Z} \sim \mathcal{F}(m-1, n-1)$$

deci avem si

$$F_{\mathcal{F}(m-1,n-1)}\left(\frac{1}{f_{1,\alpha}}\right) - F_{\mathcal{F}(m-1,n-1)}\left(\frac{1}{f_{2,\alpha}}\right) = 1-\alpha.$$

Regiunea de acceptare a ipotezei $H_1 : \{\gamma = 1\}$ este

$$A_{n,m;1-\alpha}(\gamma = 1) = \left\{(x_{11}, ..., x_{1n}, x_{21}, ..., x_{2m}) \mid f_{1,\alpha} \leq \frac{s_1^2}{s_2^2} \leq f_{2,\alpha}\right\}$$

iar regiunea critica este $B = A^C_{n,m;1-\alpha}(\gamma = 1)$. Probabilitatea erorii de I tip este

$$P_{(\gamma=1)}((X_{11}, ..., X_{1n}, X_{21}, ..., X_{2m}) \in B) = \alpha$$

si functia caracteristica operatoare a testului este

$$OC(\gamma) = P_\gamma\left(f_{1,\alpha} \leq \frac{S_1^2}{S_2^2} \leq f_{2,\alpha}\right) = P_\gamma\left(\gamma \cdot f_{1,\alpha} \leq \gamma \cdot \frac{S_1^2}{S_2^2} \leq \gamma \cdot f_{2,\alpha}\right) =$$
$$= F_{\mathcal{F}(n-1,m-1)}(\gamma \cdot f_{2,\alpha}) - F_{\mathcal{F}(n-1,m-1)}(\gamma \cdot f_{1,\alpha})$$

Functia din R: var.test(x,y,...)

var.test(x, y, ratio = 1, alternative = c("two.sided", "less", "greater"), conf.level = 0.95, ...)

Arguments

x, y numeric vectors of data values, or fitted linear model objects (inheriting from class "lm").

ratio the hypothesized ratio of the population variances of x and y.

alternative a character string specifying the alternative hypothesis, must be one of "two.sided" (default), "greater" or "less".

conf.level confidence level for the returned confidence interval.

(*b*) Testul "$t$" de comparare a mediilor,
$H_2 : \{\mu_1 = \mu_2\}$ , $H_{2A} : \{\mu_1 \neq \mu_2\}$

Presupunem ca s-a acceptat ipoteza de egalitate a dispersiilor, $H_1 : \{\sigma_1^2 = \sigma_2^2\}$. Rezulta:

$$\overline{X_1} \sim N\left(\mu_1, \frac{\sigma^2}{n}\right)$$
$$\overline{X_2} \sim N\left(\mu_2, \frac{\sigma^2}{m}\right)$$

Folosind independenta, avem

$$\overline{X_1} - \overline{X_2} \sim N\left(\mu_1 - \mu_2, \sigma^2\left(\frac{1}{n} + \frac{1}{m}\right)\right)$$

Pe de alta parte,

$$\frac{1}{\sigma^2}\left((n-1)S_1^2 + (m-1)S_2^2\right) \sim \chi^2(n+m-2)$$

Folosind asociativitatea independentei,

$$\frac{\left(\overline{X_1} - \overline{X_2}\right) - (\mu_1 - \mu_2)}{\sqrt{\sigma^2\left(\frac{1}{n} + \frac{1}{m}\right)}} \Bigg/ \sqrt{\frac{1}{n+m-2} \cdot \frac{1}{\sigma^2}\left((n-1)S_1^2 + (m-1)S_2^2\right)} \sim t(n+m-2)$$

Reparametrizam si rescriem ipotezele $H_2, H_{2A}$ :

$$\delta = \mu_1 - \mu_2$$

$$H_2 : \{\delta = 0\}\ ,\ H_{2A} : \{\delta \neq 0\}$$

Daca ipoteza $H_2$ este adevarata, atunci

$$Z = \frac{\overline{X_1} - \overline{X_2}}{\sqrt{\frac{1}{n+m-2}\left(\frac{1}{n} + \frac{1}{m}\right)\left((n-1)S_1^2 + (m-1)S_2^2\right)}} \sim t(n+m-2)$$

Pentru $\alpha \in (0,1)$ arbitrar fixat, fie $t_{n+m-2;1-\alpha/2}$ cuantila de rang $\left(1 - \frac{\alpha}{2}\right)$ a repartitiei $t(n+m-2)$.

Regiunea de acceptare a ipotezei $H_2$ este

$$A_{n,m;1-\alpha}(\delta = 0) = \left\{(x_{11}, ..., x_{1n}, x_{21}, ..., x_{2n}) \mid -t_{n+m-2;1-\alpha/2} \leq z \leq t_{n+m-2;1-\alpha/2}\right\}$$

Regiunea critica pentru $H_2$, la pragul de semnificatie $\alpha$ este

$$B = A^C_{n,m;1-\alpha}(\delta = 0)$$

cu probabilitatea de eroare de tip I

$$P_{(\delta=0)}\left((X_{11},...,X_{1n},X_{21},...,X_{2m})\in B\right)=\alpha$$

si functia caracteristica operatoare

$$OC\left(\delta\right)=P_{\delta}\left(-t_{n+m-2;1-\alpha/2}\leq Z\leq t_{n+m-2;1-\alpha/2}\right)=$$

$$F_{t(n+m-2)}\left(t_{n+m-2;1-\alpha/2}-\delta\Bigg/\sqrt{\frac{1}{n+m-2}\left(\frac{1}{n}+\frac{1}{m}\right)\left((n-1)\,s_1^2+(m-1)\,s_2^2\right)}\right)-$$
$$-F_{t(n+m-2)}\left(-t_{n+m-2;1-\alpha/2}-\delta\Bigg/\sqrt{\frac{1}{n+m-2}\left(\frac{1}{n}+\frac{1}{m}\right)\left((n-1)\,s_1^2+(m-1)\,s_2^2\right)}\right)$$

Functia din R: t.test(x,y,....)

t.test(x, y = NULL, alternative = c("two.sided", "less", "greater"), mu = 0, paired = FALSE, var.equal = FALSE, conf.level = 0.95, ...)

Arguments

x  a numeric vector of data values.

y  an optional numeric vector data values.

alternative  a character string specifying the alternative hypothesis, must be one of "two.sided" (default), "greater" or "less".

mu  a number indicating the difference in means (if you are performing a two sample test).

paired  a logical indicating whether you want a paired t-test.

var.equal  a logical variable indicating whether to treat the two variances as being equal. If TRUE then the pooled variance is used to estimate the variance. Otherwise the Welch approximation to the degrees of freedom is used.

conf.level  confidence level of the interval.