STATISTICĂ - LABORATOR 6

Teste parametrice pentru verificarea ipotezelor statistice

Presupunem că se studiază o caracteristica X, relativă la o populaţie. Fie

- $m = E(X) =$ media caracteristicii X,
- $\sigma^2 = Var(X) =$ dispersia,
- $\sigma = \sqrt{Var(X)} =$ abaterea standard.

Se consideră o selecţie repetată de volum n.

Fie $X_1, \ldots, X_n$ variabilele de selecţie şi $x_1, \ldots, x_n$ datele de selecţie.

Se definesc statisticile:

- media de selecţie ($\boxed{\text{mean}}$):

$$\overline{X} = \frac{1}{n} \sum_{i=1}^{n} X_i \tag{1}$$

- dispersia de selecţie ($\boxed{\text{var}}$):

$$s^2 = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \overline{X})^2 \tag{2}$$

- abaterea standard de selecţie ($\boxed{\text{std}}$):

$$s = \sqrt{s^2} \tag{3}$$

Testul Z privind media teoretică

Aplicabilitate: <u>Selecţie de volum mare ($n \geq 30$) sau provenind dintr-o distribuţie normală şi σ cunoscut</u>

Relativ la media teoretică $m = E\left(X\right)$ verificăm ipoteza nulă:

$H_0 : m = m_0$ cu una din alternativele:

$H_1 : m < m_0$ **(test Z la stânga)**

$H_1 : m > m_0$ **(test Z la dreapta)**

$H_1 : m \neq m_0$ **(test Z bilateral)**

Observaţie: $H_0 \cap H_1 = \varnothing$ (disjuncte).

Statistica de test care se utilizează este:

$$Z = \frac{\overline{X} - m}{\frac{\sigma}{\sqrt{n}}}, \tag{4}$$

care urmează legea normală $N\left(0, 1\right)$.

Se dă un nivel de semnificaţie $\alpha \in (0, 1)$.

Se defineşte regiunea U de respingere a ipotezei nule H_0, prin:

- $(-\infty, z_\alpha]$ (test Z la stânga)
- $[z_{1-\alpha}, \infty)$ (test Z la dreapta)
- $\left(-\infty, -z_{1-\frac{\alpha}{2}}\right] \cup \left[z_{1-\frac{\alpha}{2}}, +\infty\right)$ (test Z bilateral)

unde cuantilele z_α, $z_{1-\alpha}$ şi $z_{1-\frac{\alpha}{2}}$ se referă la legea $N(0,1)$. ($\boxed{\text{norminv}}$)

Se calculează valoarea z_0 a statisticii de test Z, atunci când ipoteza H_0 este adevărată:

$$z_0 = \frac{\overline{x} - m_0}{\frac{\sigma}{\sqrt{n}}} \tag{5}$$

Luarea deciziei folosind regiunea U:
dacă $z_0 \in U$ se respinge ipoteza nulă H_0 (în acest caz se acceptă alternativa H_1)
dacă $z_0 \notin U$ se acceptă ipoteza nulă H_0

Funcţia Matlab `ztest`

Sistemul Matlab, prin *Statistics toolbox*, dispune de funcţia `ztest`, cu aplicabilitate la testul Z. Apelarea acestei funcţii se face prin:

```
[h,P,ci,zval] = ztest(x,m0,sigma,alpha,tail);
```

Această instrucţiune efectuează testul Z asupra datelor conţinute în vectorul `x`, folosind nivelul de semnificaţie `alpha`.

Parametrul `tail` specifică una din cele 3 alternative, care conduc la testul unilateral stânga (`tail=-1`), unilateral dreapta (`tail=1`) şi bilateral (`tail=0`).

Dacă `h=1`, atunci ipoteza nulă H_0 va fi respinsă, iar dacă `h=0`, ipoteza nulă H_0 va fi acceptată.

Apelul permite obţinerea valorii critice `P`, a valorii `zval` a statisticii de test Z, precum şi a intervalului de încredere pentru media teoretică, corespunzător probabilităţii de încredere `1-alpha`, obţinut în vectorul cu două componente `ci`.

Testul T privind media teoretică

Aplicabilitate: Selecţie de volum mare ($n \geq 30$) sau provenind dintr-o distribuţie normală şi σ necunoscut

Relativ la media teoretică $m = E(X)$ verificăm ipoteza nulă:

$H_0 : m = m_0$ cu una din alternativele:

$H_1 : m < m_0$ **(test T la stânga)**
$H_1 : m > m_0$ **(test T la dreapta)**
$H_1 : m \neq m_0$ **(test T bilateral)**

Deoarece σ este necunoscut, utilizăm abaterea standard de selecţie ($\boxed{\text{std}}$):

$$s = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2} \tag{6}$$

Statistica de test care se utilizează este

$$T = \frac{\overline{X} - m}{\frac{s}{\sqrt{n}}}, \tag{7}$$

care urmează legea $T(n-1)$.

Se dă un nivel de semnificaţie $\alpha \in (0,1)$.

Corespunzător celor trei alternative, definim regiunea U de respingere a ipotezei nule H_0, prin:

- $(-\infty, t_\alpha]$ (test T la stânga)
- $[t_{1-\alpha}, \infty)$ (test T la dreapta)
- $\left(-\infty, -t_{1-\frac{\alpha}{2}}\right] \cup \left[t_{1-\frac{\alpha}{2}}, +\infty\right)$ (test T bilateral)

unde cuantilele t_α, $t_{1-\alpha}$ şi $t_{1-\frac{\alpha}{2}}$ se referă la legea $T\,(n-1)$. ($\boxed{\text{tinv}}$)

Se calculează valoarea t_0 a statisticii de test T, atunci când ipoteza nulă H_0 este adevărată:

$$t_0 = \frac{\overline{x} - m_0}{\frac{s}{\sqrt{n}}}. \tag{8}$$

<u>Luarea deciziei</u> folosind regiunea U:

dacă $t_0 \in U$ se respinge ipoteza nulă H_0 (în acest caz se acceptă alternativa H_1)

dacă $t_0 \notin U$ se acceptă ipoteza nulă H_0

Funcţia Matlab `ttest`

Sistemul Matlab, prin *Statistics toolbox*, dispune de funcţia `ttest`, cu aplicabilitate la testul T. Apelarea acestei funcţii se face prin:

```
[h,P,ci,stats] = ttest(x,m0,alpha,tail);
```

În urma executării acestei instrucţiuni, se efectuează testul T asupra datelor conţinute în vectorul `x`, folosind nivelul de semnificaţie `alpha`.

Parametrul `tail` specifică una din cele 3 alternative, care conduc la testul unilateral stânga (`tail=-1`), unilateral dreapta (`tail=1`) şi bilateral (`tail=0`).

Dacă `h=1`, atunci ipoteza nulă H_0 va fi respinsă, respectiv dacă `h=0`, ipoteza nulă H_0 va fi acceptată.

Apelul permite obţinerea valorii critice `P`, precum şi a intervalului de încredere pentru media teoretică, corespunzător probabilităţii de încredere `1-alpha`, obţinut în vectorul cu două componente `ci`.

Parametrul `stats` are următoarele câmpuri:

`tstat` – valoarea statisticii de test
`df` – numărul gradelor de libertate ale testului
`sd` – abaterea standard de selecţie s.

De exemplu, pentru returnarea valorii statisticii de test, se utilizează instrucţiunea:

```
t0 = stats.tstat
```

Testul χ^2 privind dispersia teoretică a legii normale

Fie caracteristica $X \sim N(m, \sigma)$, cu dispersia teoretică $\sigma^2 = Var(X)$ necunoscută.
Referitor la dispersia teoretică, verificăm ipoteza nulă:

$H_0 : \sigma^2 = \sigma_0^2$ cu una din alternativele:

$H_1 : \sigma^2 < \sigma_0^2$ **(testul χ^2 la stânga)**

$H_1 : \sigma^2 > \sigma_0^2$ **(testul χ^2 la dreapta)**

$H_1 : \sigma^2 \neq \sigma_0^2$ **(testul χ^2 bilateral)**

Statistica de test care se utilizează este:

$$V = \frac{(n-1)\, s^2}{\sigma^2},$$

care urmează legea $\chi^2(n-1)$.

Obs: s^2 este dispersia de selecţie definită în formula (2). ($\boxed{\text{var}}$)

Se dă un nivel de semnificaţie $\alpha \in (0, 1)$.

Se defineşte regiunea U de respingere a ipotezei nule H_0, astfel:

- $U = (-\infty, h_\alpha]$ pentru testul χ^2 la stânga
- $U = [h_{1-\alpha}, \infty)$ pentru testul χ^2 la dreapta
- $U = (-\infty, h_{\frac{\alpha}{2}}] \cup [h_{1-\frac{\alpha}{2}}, \infty)$ pentru testul χ^2 bilateral

unde cuantilele h_α, $h_{1-\alpha}$, $h_{\frac{\alpha}{2}}$ şi $h_{1-\frac{\alpha}{2}}$ se referă la legea $\chi^2(n-1)$. $\boxed{\text{chi2inv}}$

Se calculează valoarea v_0 a statisticii V, atunci când ipoteza H_0 este adevărată, adică $\sigma^2 = \sigma_0^2$:

$$v_0 = \frac{(n-1)\, s^2}{\sigma_0^2}.$$

Luarea deciziei folosind regiunea U:
dacă $v_0 \in U$, ipoteza nulă H_0 este respinsă
dacă $v_0 \notin U$, ipoteza nulă H_0 este acceptată.

Funcţia Matlab `vartest`

Sistemul Matlab, prin *Statistics toolbox*, dispune de funcţia `vartest`, cu aplicabilitate la testul χ^2 privind dispersia teoretică a legii normale. Apelarea acestei funcţii se face prin:

```
[h,P,ci,stats] = vartest(x,σ0^2,alpha,tail);
```

În urma executării acestei instrucţiuni, se efectuează testul χ^2 asupra datelor conţinute în vectorul `x`, folosind nivelul de semnificaţie `alpha`.

Parametrul `tail` specifică una din cele 3 alternative, care conduc la testul unilateral stânga (`tail=-1`), unilateral dreapta (`tail=1`) şi bilateral (`tail=0`).

Dacă `h=1`, atunci ipoteza nulă H_0 va fi respinsă, iar dacă `h=0`, ipoteza nulă H_0 va fi acceptată.

Apelul permite obţinerea valorii critice `P`, precum şi a intervalului de încredere pentru dispersia teoretică, corespunzător probabilităţii de încredere `1-alpha`, obţinut în vectorul cu două componente `ci`.

Parametrul `stats` are următoarele câmpuri:

`chisqstat` – valoarea statisticii de test

`df` – numărul gradelor de libertate ale testului.

De exemplu, pentru returnarea valorii statisticii de test, se utilizează instrucţiunea:

```
v0 = stats.chisqstat
```