Algoritm

Dacă matricea A este nesingulară, pașii algoritmului sunt:

Pasul 1 Se inițializează matricea extinsă $[A, B]$:

$$[A^1, B^1] = \left(\begin{array}{cccc|c} a_{11}^1 & a_{12}^1 & \dots & a_{1n}^1 & b_1^1 \\ a_{21}^1 & a_{22}^1 & \dots & a_{2n}^1 & b_2^1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{n1}^1 & a_{n2}^1 & \dots & a_{nn}^1 & b_n^1 \end{array}\right)$$

Pasul 2 Se obțin zerouri pe prima coloană astfel:

Pasul 2.1 Dacă $a_{11}^1 = 0$ se schimbă $L_1 \leftrightarrow L_i$ unde $a_{i1}^1 \neq 0$ (există un astfel de $i > 1$ deoarece matricea A este nesingulară);

Pasul 2.2 $L_k \leftarrow L_k - \frac{a_{k1}^1}{a_{11}^1} L_1$ pentru $k = 2, 3, \dots, n$ și se obține

$$[A^2, B^2] = \left(\begin{array}{cccc|c} a_{11}^1 & a_{12}^1 & \dots & a_{1n}^1 & b_1^1 \\ 0 & a_{22}^2 & \dots & a_{2n}^2 & b_2^2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & a_{n2}^2 & \dots & a_{nn}^2 & b_n^2 \end{array}\right)$$

Pasul 3 Se obțin zerouri pe a doua coloană astfel:

Pasul 3.1 Dacă $a_{22}^2 = 0$ se schimbă $L_2 \leftrightarrow L_i$ unde $a_{i2}^2 \neq 0$ (există un astfel de $i > 2$ deoarece matricea A este nesingulară);

Pasul 3.2 $L_k \leftarrow L_k - \frac{a_{k2}^2}{a_{22}^2} L_2$ pentru $k = 3, \dots, n$ și se obține

$$[A^3, B^3] = \left(\begin{array}{ccccc|c} a_{11}^1 & a_{12}^1 & a_{13}^1 & \dots & a_{1n}^1 & b_1^1 \\ 0 & a_{22}^2 & a_{23}^2 & \dots & a_{2n}^2 & b_2^2 \\ 0 & 0 & a_{33}^3 & \dots & a_{3n}^3 & b_3^3 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & a_{n3}^3 & \dots & a_{nn}^3 & b_n^3 \end{array}\right)$$

Pasul n $[A^n, B^n] = \left(\begin{array}{cccc|c} a_{11}^1 & a_{12}^1 & \dots & a_{1n}^1 & b_1^1 \\ 0 & a_{22}^2 & \dots & a_{2n}^2 & b_2^2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & a_{nn}^n & b_n^n \end{array}\right)$

Sistemul având matricea triunghiulară superior obținută la ultimul pas se rezolvă prin substituție inversă.

Exemplu

Fie sistemul $\begin{cases} x_1 + 2x_2 + x_3 = 8 \\ 2x_1 + 3x_2 - 2x_3 = 2 \\ 3x_1 + 4x_2 + 4x_3 = 23 \end{cases}$. Efectuăm următoarele transformări elementare asupra matricei extinse:

$$\left(\begin{array}{ccc|c} 1 & 2 & 1 & 8 \\ 2 & 3 & -2 & 2 \\ 3 & 4 & 4 & 23 \end{array}\right) \begin{array}{c} L_2 - 2L_1 \\ \sim \\ L_3 - 3L_1 \end{array} \left(\begin{array}{ccc|c} 1 & 2 & 1 & 8 \\ 0 & -1 & -4 & -14 \\ 0 & -2 & 1 & -1 \end{array}\right) \begin{array}{c} L_3 - 2L_2 \\ \sim \end{array}$$

$$\left(\begin{array}{ccc|c} 1 & 2 & 1 & 8 \\ 0 & -1 & -4 & -14 \\ 0 & 0 & 9 & 27 \end{array}\right) \Rightarrow \begin{cases} x_1 + 2x_2 + x_3 = 8 \\ \quad -x_2 - 4x_3 = -14 \\ \quad\quad 9x_3 = 27 \end{cases} \Rightarrow x_3 = 3 \Rightarrow$$

$$-x_2 - 12 = -14 \Rightarrow x_2 = 2 \Rightarrow x_1 + 4 + 3 = 8 \Rightarrow x_1 = 1.$$

Metoda dreptunghiurilor

- Pentru o rețea de puncte echidistante notăm cu $y_0 = f(x_0)$, $y_1 = f(x_1), \dots y_n = f(x_n)$ valorile funcției f în punctele rețelei
- Pe fiecare interval $[x_k, x_{k+1}]$ aproximăm funcția $f(x)$ cu y_k, respectiv y_{k+1}
- Se obține o aproximare a ariei domeniului D_k cu $y_k h$, respectiv $y_{k+1} h$
- Se obțin aproximările

$$\int_a^b f(x)dx = \mathcal{A}(D) \simeq h(y_0 + y_1 + y_2 + \dots + y_{n-1}) \quad (1)$$

$$\int_a^b f(x)dx = \mathcal{A}(D) \simeq h(y_1 + y_2 + y_3 + \dots + y_n) \quad (2)$$

- Eroarea

$$|I - \mathcal{A}(D)| < \frac{(b-a)^2}{n} M, \text{ unde } M = \sup\{|f'(x)|; x \in [a, b]\}$$

Exemplu

Să se aproximeze $\int_0^1 \frac{1}{1+3x} dx$ folosind metoda dreptunghiurilor.

Avem $a = 0,\ b = 1,\ n = 4,\ h = \frac{b-a}{n} = \frac{1}{4} = 0.25$

$$\begin{aligned} x_0 &= a = 0 \Rightarrow y_0 = f(0) = 1 \\ x_1 &= a + h = 0.25 \Rightarrow y_1 = f(0.25) = 0.5714 \\ x_2 &= a + 2h = 0.5 \Rightarrow y_2 = f(0.5) = 0.4 \\ x_3 &= a + 3h = 0.75 \Rightarrow y_3 = f(0.75) = 0.3077 \\ x_4 &= a + 4h = b = 1 \Rightarrow y_4 = f(1) = 0.25 \end{aligned}$$

$$\int_0^1 \frac{1}{1+3x} dx \simeq h(y_0 + y_1 + y_2 + y_3) = 0.569775$$

$$\int_0^1 \frac{1}{1+3x} dx \simeq h(y_1 + y_2 + y_3 + y_4) = 0.382275$$

Valoarea exactă este $\int_0^1 \frac{1}{1+3x} dx = \frac{1}{3} \ln(1+3x) \Big|_0^1 = 0.4621$

Metoda trapezelor

- Considerăm din nou o rețea echidistantă, domeniile D_k și notăm punctele de pe graficul funcției f corespunzătoare rețelei alese cu $M_0(x_0, y_0)$, $M_1(x_1, y_1)$, $\dots M_n(x_n, y_n)$.
- Aproximăm aria domeniului D_k prin aria trapezului determinat de punctele M_k, M_{k+1} și punctele de pe Ox corespunzătoare lui x_k și x_{k+1}:

$$\mathcal{A}(D_k) \simeq \frac{h(y_k + y_{k+1})}{2}$$

- Se obține aproximarea

$$\int_a^b f(x)dx = \mathcal{A}(D) \simeq \frac{h}{2}(y_0 + 2y_1 + 2y_2 + \dots + 2y_{n-1} + y_n) \quad (3)$$

- Eroarea

$$|I - \mathcal{A}(D)| < \frac{(b-a)^3}{12n^2} M, \text{ unde } M = \sup\{|f''(x)|; x \in [a, b]\}$$

Exemplu

Să se aproximeze $\int_0^1 \frac{1}{1+3x} dx$ folosind metoda trapezelor.

Avem $a = 0,\ b = 1,\ n = 4,\ h = \frac{b-a}{n} = \frac{1}{4} = 0.25$

$$\begin{aligned} x_0 &= a = 0 \Rightarrow y_0 = f(0) = 1 \\ x_1 &= a + h = 0.25 \Rightarrow y_1 = f(0.25) = 0.5714 \\ x_2 &= a + 2h = 0.5 \Rightarrow y_2 = f(0.5) = 0.4 \\ x_3 &= a + 3h = 0.75 \Rightarrow y_3 = f(0.75) = 0.3077 \\ x_4 &= a + 4h = b = 1 \Rightarrow y_4 = f(1) = 0.25 \end{aligned}$$

$$\int_0^1 \frac{1}{1+3x} dx \simeq \frac{h}{2}(y_0 + 2y_1 + 2y_2 + 2y_3 + y_4) = 0.476$$

Metoda parabolelor (Simpson)

- Considerăm o rețea echidistantă cu un număr par de noduri $n = 2m$ și punctele de pe graficul funcției f corespunzătoare rețelei alese $M_0(x_0, y_0)$, $M_1(x_1, y_1)$, $\dots M_n(x_n, y_n)$.
- Pe fiecare interval $[x_{2k}, x_{2k+2}]$ aproximăm funcția f cu o funcție de gradul 2 al cărei grafic trece prin punctele $M_{2k}, M_{2k+1}, M_{2k+2}$. Avem:

$$\mathcal{A}(D_k) \simeq \frac{h}{3}(y_{2k} + 4y_{2k+1} + y_{2k+2})$$

- Se obține aproximarea

$$\int_a^b f(x)dx \simeq \frac{h}{3}(y_0 + 4y_1 + 2y_2 + 4y_3 + 2y_4 + \dots + 2y_{2m-2} + 4y_{2m-1} + y_{2m}) \quad (4)$$

- Eroarea

$$|I - \mathcal{A}(D)| < \frac{(b-a)^5}{180n^4} M, \text{ unde } M = \sup\{|f^{(4)}(x)|; x \in [a, b]\}$$

Exemplu

Să se aproximeze $\int_0^1 \frac{1}{1+3x} dx$ folosind metoda lui Simpson.

Avem $a = 0,\ b = 1,\ n = 4,\ h = \frac{b-a}{n} = \frac{1}{4} = 0.25$

$$\begin{aligned} x_0 &= a = 0 \Rightarrow y_0 = f(0) = 1 \\ x_1 &= a + h = 0.25 \Rightarrow y_1 = f(0.25) = 0.5714 \\ x_2 &= a + 2h = 0.5 \Rightarrow y_2 = f(0.5) = 0.4 \\ x_3 &= a + 3h = 0.75 \Rightarrow y_3 = f(0.75) = 0.3077 \\ x_4 &= a + 4h = b = 1 \Rightarrow y_4 = f(1) = 0.25 \end{aligned}$$

$$\int_0^1 \frac{1}{1+3x} dx \simeq \frac{h}{3}(y_0 + 4y_1 + 2y_2 + 4y_3 + y_4) = 0.4639$$

Interval de încredere pentru medie (σ necunoscut)

Intervalul de încredere pentru media unei variabile aleatoare X având repartiția $N(m,\sigma^2)$ cu $m \in \mathbb{R}$ necunoscut și $\sigma^2 > 0$ necunoscut este de forma $\left(\bar{X} - t_{n-1,1-\frac{\alpha}{2}} \cdot \frac{s}{\sqrt{n}}, \bar{X} + t_{n-1,1-\frac{\alpha}{2}} \cdot \frac{s}{\sqrt{n}}\right)$, cu

$$P\left(\bar{X} - t_{n-1,1-\frac{\alpha}{2}} \cdot \frac{s}{\sqrt{n}} < m < \bar{X} + t_{n-1,1-\frac{\alpha}{2}} \cdot \frac{s}{\sqrt{n}}\right) = 1-\alpha$$

unde $\bar{X} = \frac{X_1+X_2+\cdots+X_n}{n}$ este media de selecție, $s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i-\bar{X})^2$ este estimația nedeplasată a lui σ^2, $\alpha \in (0,1)$ este nivelul de semnificație iar $t_{n-1,1-\frac{\alpha}{2}}$ este cuantila de ordin $1-\frac{\alpha}{2}$ a repartiției Student cu $n-1$ grade de libertate:

$$t_{n-1,1-\frac{\alpha}{2}} = F_{n-1}^{-1}\left(1-\frac{\alpha}{2}\right)$$

unde F_{n-1} este funcția de repartiție Student cu $n-1$ grade de libertate.

Exemplu

S-a efectuat un studiu privind cantitatea de apă minerală dintr-un număr de 8 butelii și s-au obținut următoarele rezultate (în litri): 4.96, 4.90, 4.98, 5, 4.94, 5, 5.02, 4.92. Să se determine un interval de încredere 95% pentru volumul mediu al buteliilor de apă presupunând că volumul are o distribuție normală.

Rezolvare:

- Avem $n=8$, $\bar{X} = \frac{4.96+4.90+4.98+5+4.94+5+5.02+4.92}{8} = 4.965$
- Estimația nedeplasată a dispersiei $s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i-\bar{X})^2 = 0.0018$
- Cum nivelul de încredere este $1-\alpha = 0.95$, găsim nivelul de semnificație $\alpha = 0.05$, de unde $1-\frac{\alpha}{2} = 0.975$.
- Cuantila corespunzătoare acestei valori din repartiția Student cu $n-1=7$ grade de libertate este $t_{n-1,1-\frac{\alpha}{2}} = 2.3646$
- Marja de eroare
$$t_{n-1,1-\frac{\alpha}{2}} \cdot \frac{s}{\sqrt{n}} = 0.0355$$
- Intervalul de încredere 95% este $(4.9295, 5.0005)$.

Interval de încredere pentru dispersie

Intervalul de încredere pentru dispersia unei variabile aleatoare X având repartiția $N(m,\sigma^2)$ cu $m \in \mathbb{R}$ necunoscut și $\sigma^2 > 0$ necunoscut este de forma

$$\left(\frac{(n-1)\cdot s^2}{\chi^2_{n-1,1-\frac{\alpha}{2}}}, \frac{(n-1)\cdot s^2}{\chi^2_{n-1,\frac{\alpha}{2}}}\right)$$

unde $s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i-\bar{X})^2$ este estimația nedeplasată a lui σ^2, $\alpha \in (0,1)$ este nivelul de semnificație, iar $\chi^2_{n-1,1-\frac{\alpha}{2}}$ și $\chi^2_{n-1,\frac{\alpha}{2}}$ sunt cuantilele de ordin $1-\frac{\alpha}{2}$ și $\frac{\alpha}{2}$ ale repartiției χ^2 cu $n-1$ grade de libertate.

Exemplu: Grosimea unui tip de sticlă are o repartiție normală. S-a efectuat un studiu pe 5 bucăți și s-au obținut următoarele rezultate (în mm): 4.8, 5, 4.6, 4.4, 5.2. Să se determine un interval de încredere 90% pentru abaterea medie pătratică de la grosimea medie. **R**: $(0.2053, 0.7502)$

Pentru testarea unei ipoteze statistice se parcurg următoarele etape:

1. Se alege ipoteza nulă H_0 și ipoteza alternativă H_a
2. Se alege un criteriu de testare $C(X,n)$
3. Se alege un nivel de semnificație α
4. Se determină regiunea de respingere W cu proprietatea că
$$P(C(X,n) \in W | H_0 \text{ este adevărată}) = \alpha \text{ (sau } \le \alpha)$$
5. Se face o selecție de volum n, obținându-se valorile $x_1, x_2, \ldots, x_n$ ale variabilelor aleatoare de selecție $X_1, X_2, \ldots, X_n$ corespunzătoare caracteristicii studiate
6. Se calculează $C(x_1, x_2, \ldots, x_n)$:
   - Dacă $C(x_1, x_2, \ldots, x_n) \in \bar{W}$ atunci H_0 este acceptată
   - Dacă $C(x_1, x_2, \ldots, x_n) \in W$ atunci H_0 este respinsă (H_a este acceptată)

Test bilateral pentru medie (σ cunoscut)

Fie X o variabilă aleatoare teoretică cu repartiție normală $N(m,\sigma^2)$ cu media necunoscută și m_0 o estimație a acestei medii. Se face o selecție de volum n, obținându-se variabilele aleatoare de selecție $X_1, X_2, \ldots, X_n$.

1. Dorim să verificăm ipoteza
$$H_0: \ m = m_0$$
față de alternativa
$$H_a: \ m \ne m_0.$$
2. Alegem criteriul de testare
$$C(X,n) = \frac{\bar{X}-m_0}{\frac{\sigma}{\sqrt{n}}}$$
deoarece dacă H_0 este adevărată, atunci $C(X,n)$ are repartiția complet specificată $N(0,1)$.
3. Alegem un nivel de semnificație α (de obicei 0.1, 0.05 sau 0.01).

4. Se determină $z =$ cuantila de ordin $1-\frac{\alpha}{2}$ a repartiției normale standard, corespunzătoare nivelului de semnificație ales α. Avem
$$P(|C(X,n)| > z | m = m_0) = \alpha$$
deci regiunea de respingere este
$$W = (-\infty, -z) \cup (z, \infty).$$
5. Se face o selecție de volum n, obținându-se valorile $x_1, x_2, \ldots, x_n$ și se calculează
$$\bar{X} = \frac{x_1+x_2+\cdots+x_n}{n}.$$
6. Dacă $\frac{\bar{X}-m_0}{\frac{\sigma}{\sqrt{n}}} \in [-z, z]$ atunci H_0 este acceptată.
Dacă $\frac{\bar{X}-m_0}{\frac{\sigma}{\sqrt{n}}} \in (-\infty, -z) \cup (z, \infty)$ atunci H_0 este respinsă și se acceptă H_a.

Exemplu

La o fabrică patronul asigură că durata medie de funcționare a unui produs este de 800 ore cu abaterea medie pătratică de 50 ore. Pe un eșantion de 20 bucăți s-a obținut că durata medie de funcționare este 790 ore. Presupunând că durata de funcționare este normal distribuită să se verifice cu nivelul de semnificație de 4% dacă afirmația patronului este adevărată.

1. ipotezele $H_0: \ m = 800, \qquad H_a: m \ne 800$
2. criteriul de testare $\frac{\bar{X}-800}{\frac{50}{\sqrt{20}}}$
3. nivelul de semnificație $\alpha = 0.04$
4. cuantila de ordin $1-\frac{\alpha}{2} = 0.98$ este $z = 2.05$
5. pentru $\bar{X} = 790$ avem $\frac{\bar{X}-800}{\frac{50}{\sqrt{20}}} = -0.894$
6. ipoteza H_0 se acceptă deoarece $-0.894 \in [-2.05, 2.05]$

Teste pentru proporție

Considerăm problema testării ipotezei că proporția de indivizi dintr-o populație care au o anumită caracteristică are o valoare specificată. Considerăm variabila aleatoare teoretică Bernoulli $X: \begin{pmatrix} 0 & 1 \\ 1-p & p \end{pmatrix}$, unde 1 înseamnă că individul are caracteristica respectivă. Media de selecție corespunzătoare $\hat{p} = \frac{X_1+X_2+\cdots+X_n}{n}$ este un estimator nedeplasat pentru proporția p și pentru valori mari ale lui n are repartiția $N\left(p, \frac{pq}{n}\right)$ unde $q = 1-p$.

1. Se formulează ipotezele $H_0: \ p = p_0, \qquad H_a: \ p \ne p_0$
2. Se alege criteriul de testare $C(X,n) = \frac{\hat{p}-p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$
3. Se alege un nivel de semnificație α
4. Se determină $z =$ cuantila de ordin $1-\frac{\alpha}{2}$ a repartiției normale standard, corespunzătoare nivelului de semnificație ales α
5. Se află estimația punctuală a proporției $\hat{p}$ pe un eșantion de volum n
6. Dacă $\frac{\hat{p}-p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} \in [-z, z]$ atunci H_0 este acceptată.
Dacă $\frac{\hat{p}-p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} \in (-\infty, -z) \cup (z, \infty)$ atunci H_0 este respinsă.

Exemplu

Se studiază proporția produselor defecte dintr-un proces tehnologic. Pe un eșantion de 300 produse testate s-au obținut 13 defecte. Aceste date ne permit să afirmăm cu un nivel de semnificație $\alpha = 0.05$ că 5% din produse sunt defecte?

1. ipotezele $H_0: \ p = 0.05, \qquad H_a: p \ne 0.05$
2. criteriul de testare $\frac{\hat{p}-0.05}{\sqrt{\frac{0.05\cdot 0.95}{300}}}$
3. nivelul de semnificație $\alpha = 0.05$
4. cuantila de ordin $1-\frac{\alpha}{2} = 0.975$ este $z = 1.96$
5. pentru $\hat{p} = \frac{13}{300}$ avem $\frac{\hat{p}-0.05}{\sqrt{\frac{0.05\cdot 0.95}{13}}} = -0.5298$
6. ipoteza H_0 se acceptă deoarece $-0.5298 \in [-1.96, 1.96]$