TESTE PARAMETRICE

Notiuni generale

Modelul: $F_\theta = P_\theta \circ X^{-1}$ cu parametrul $\theta \in \Theta \subseteq R^k, k \geq 1$
Consideram familia

$$\{F_\theta, \theta \in \Theta\}$$

Pentru $\Theta_0 \subset \Theta$, o ipoteza statistica este o subfamilie

$$H : \{F_\theta, \theta \in \Theta_0\} \stackrel{notat}{=} \{\theta \in \Theta_0\}$$

Ipoteza alternativa lui H este subfamilia complementara

$$H_A : \{F_\theta, \theta \in \Theta - \Theta_0\} \stackrel{notat}{=} \{\theta \in \Theta - \Theta_0\}$$

Ipoteza H se numeste simpla daca Θ_0 se reduce la un singur punct, $\Theta_0 = \{\theta_0\}$.
Ipoteza H se numeste compusa daca $card\,(\Theta_0) > 1$.

Observatiile: $X_1, ..., X_n$, var. al. indep. id. rep (F_θ)
Spatiul de selectie $n-$dimensional $\left(S^n, \mathcal{S}^n, \bigotimes_{i=1}^{n} P_\theta \circ X_i^{-1}\right)$

Definitie:

O multime masurabila $B \in \mathcal{S}^n$ se numeste regiune critica pentru ipoteza $H : \{\theta \in \Theta_0\}$ daca i se ataseaza urmatoarea regula de decizie:

- $(X_1, ..., X_n)(\omega) = (x_1, ..., x_n) \in B \Longrightarrow$ respingem ipoteza $H : \{\theta \in \Theta_0\}$
- $(X_1, ..., X_n)(\omega) = (x_1, ..., x_n) \in B^C \Longrightarrow$ acceptam ipoteza $H : \{\theta \in \Theta_0\}$

A construi un test pentru ipoteza $H : \{\theta \in \Theta_0\}$ cu alternativa $H_A : \{\theta \in \Theta - \Theta_0\}$ revine la a defini o regiune critica B pentru H.

Fie ipotezele H,H_A si un test bazat pe o regiune critica B.Posibilele erori de decizie sunt:

- eroare de I tip: respingerea lui H cand H este adevarata
- eroare de II tip: acceptarea lui H cand H este falsa.

Probabilitatile de eroare sunt

$$\alpha(\theta) = P_\theta\left((X_1, ..., X_n) \in B\right) \quad \text{pentru } \theta \in \Theta_0$$
$$\beta(\theta) = P_\theta\left((X_1, ..., X_n) \in B^C\right) \quad \text{pentru } \theta \in \Theta - \Theta_0$$

Functia caracteristica operatoare a testului este

$$OC(\theta) = P_\theta\left((X_1, ..., X_n) \in B^C\right), \quad \theta \in \Theta$$

Puterea testului

$$\pi(\theta) = 1 - OC(\theta), \quad \theta \in \Theta$$

TESTE PENTRU IPOTEZE SIMPLE CU ALTERNATIVE SIMPLE

Pentru doua valori $\theta_0, \theta_1 \in \Theta$,$\theta_0 \neq \theta_1$ restrangem familia de repartitii la $\{F_\theta, \theta \in \{\theta_0, \theta_1\}\}$ si formulam ipotezele

$$H : \{\theta = \theta_0\}\,, \quad H_A : \{\theta = \theta_1\}$$

Pentru un test bazat pe regiunea critica B avem

$$\alpha = P_{\theta_0}\left((X_1, ..., X_n) \in B\right)$$
$$\beta = P_{\theta_1}\left((X_1, ..., X_n) \in B^C\right)$$

Observatie:
Daca $B = S^n$, avem $\alpha = 1$ si $\beta = 0$
Daca $B = \Phi$, avem $\alpha = 0$ si $\beta = 1.$

Strategia Neyman - Pearson de constructie a lui $B:$

- probabilitatea erorii de I tip se tine sub control;
- se cauta B^* care minimizeaza probabilitatea erorii de II tip.

Definitii:
Fie ipoteza simpla $H : \{\theta = \theta_0\}$ cu alternativa simpla $H_A : \{\theta = \theta_1\}$. Fie $\alpha \in (0, 1)$ fixat (va fi numit "prag de semnificatie").
Familia regiunilor critice pentru H pentru care probabilitatea erorii de I tip este egala cu α este

$$\mathcal{C}_\alpha = \{B \in \mathcal{S}^n \mid P_{\theta_0}\left((X_1, ..., X_n) \in B\right) = \alpha\}$$

Multimea $B^* \in \mathcal{C}_\alpha$ se numeste cea mai buna regiune critica pentru H, la pragul de semnificatie α, daca pentru orice $B \in \mathcal{C}_\alpha$ are loc relatia

$$P_{\theta_1}\left((X_1, ..., X_n) \in (B^*)^C\right) \leq P_{\theta_1}\left((X_1, ..., X_n) \in B^C\right)$$

sau relatia echivalenta

$$P_{\theta_1}\left((X_1, ..., X_n) \in B^*\right) \geq P_{\theta_1}\left((X_1, ..., X_n) \in B\right)$$

In continuare vom construi o asemenea regiune critica.

Fie modelul $F_\theta = P_\theta \circ X^{-1}$,

$$F_\theta = \begin{cases} \sum\limits_x p(x;\theta) \cdot \delta_{\{x\}}, & \text{in caz discret} \\ \text{sau} & \\ f(x;\theta) \cdot l, & \text{in caz continuu} \end{cases}$$

Repartitia vectorului observatiilor $(X_1, ..., X_n)$ este fie discreta, data prin masele de probabilitate

$$P_\theta(X_1 = x_1, ..., X_n = x_n) = \prod_{i=1}^n p(x_i, \theta),$$

fie continua, data prin densitatea de repartite

$$f(x_1, ..., x_n; \theta) = \prod_{i=1}^n f(x_i, \theta).$$

Definitie:
Fie modelul $F_\theta = P_\theta \circ X^{-1}, \theta \in \{\theta_0, \theta_1\}$, ipotezele simple $H : \{\theta = \theta_0\}$, $H_A : \{\theta = \theta_1\}$ si datele statistice $(x_1, ..., x_n) = (X_1, ..., X_n)(\omega)$.
Numim raport al probabilitatilor functia

$$u_n(x_1, ..., x_n) = \begin{cases} \prod\limits_{i=1}^n \frac{p(x_i,\theta_1)}{p(x_i,\theta_0)}, & \text{in caz discret} \\ \text{sau} & \\ \prod\limits_{i=1}^n \frac{f(x_i,\theta_1)}{f(x_i,\theta_0)}, & \text{in caz continuu} \end{cases}$$

Teorema Neyman - Pearson

Fie modelul $F_\theta = f(x;\theta) \cdot l$, $\theta \in \Theta \subseteq R^k$, $k \geq 1$ si fie ipoteza simpla $H : \{\theta = \theta_0\}$ cu alternativa simpla $H_A : \{\theta = \theta_1\}$, $\theta_0 \neq \theta_1$. Fie $X_1, ..., X_n$ observatii independente, identic repartizate (F_θ) si fie $u_n(x_1, ..., x_n)$ raportul probabilitatilor corespunzator. Fie $\alpha \in (0,1)$ arbitrar fixat si fie $k_{1-\alpha}$ cuantila de rang $(1-\alpha)$ a repartitiei lui $u_n(X_1, ..., X_n)$ cand $\theta = \theta_0$, adica

$$P_{\theta_0}(u_n(X_1, ..., X_n) < k_{1-\alpha}) = 1 - \alpha.$$

Atunci multimea masurabila

$$\widetilde{B} = \{(x_1, ..., x_n) \mid u_n(x_1, ..., x_n) \geq k_{1-\alpha}\}$$

este cea mai buna regiune critica pentru H la pragul de semnificatie α (adica $\widetilde{B} = B^*$)

Demonstratie:
Avem $\widetilde{B} \in \mathcal{C}_\alpha$ pentru ca

$$P_{\theta_0}\left((X_1, ..., X_n) \in \widetilde{B}\right) = P_{\theta_0}\left(u_n\left(X_1, ..., X_n\right) \geq k_{1-\alpha}\right) = \alpha$$

Fie $B \in \mathcal{C}_\alpha$. Evaluam urmatoarea diferenta

$$P_{\theta_1}\left((X_1, ..., X_n) \in \widetilde{B}\right) - P_{\theta_1}\left((X_1, ..., X_n) \in B\right) =$$

$$P_{\theta_1}\left((X_1, ..., X_n) \in \widetilde{B} \cap B^C\right) - P_{\theta_1}\left((X_1, ..., X_n) \in \left(\widetilde{B}\right)^C \cap B\right) =$$

$$\int\limits_{\widetilde{B} \cap B^C} \prod_{i=1}^{n} f\left(x_i, \theta_1\right) dx_1 ... dx_n - \int\limits_{\left(\widetilde{B}\right)^C \cap B} \prod_{i=1}^{n} f\left(x_i, \theta_1\right) dx_1 ... dx_n =$$

$$\int\limits_{\widetilde{B} \cap B^C} u_n\left(x_1, ..., x_n\right) \prod_{i=1}^{n} f\left(x_i, \theta_0\right) dx_1 ... dx_n - \int\limits_{\left(\widetilde{B}\right)^C \cap B} u_n\left(x_1, ..., x_n\right) \prod_{i=1}^{n} f\left(x_i, \theta_0\right) dx_1 ... dx_n$$

Tinand cont de constructia lui $\widetilde{B}$ obtinem

$$P_{\theta_1}\left((X_1, ..., X_n) \in \widetilde{B}\right) - P_{\theta_1}\left((X_1, ..., X_n) \in B\right) \geq$$

$$\int\limits_{\widetilde{B} \cap B^C} k_{1-\alpha} \prod_{i=1}^{n} f\left(x_i, \theta_0\right) dx_1 ... dx_n - \int\limits_{\left(\widetilde{B}\right)^C \cap B} k_{1-\alpha} \prod_{i=1}^{n} f\left(x_i, \theta_0\right) dx_1 ... dx_n =$$

$$k_{1-\alpha}\left(P_{\theta_0}\left((X_1, ..., X_n) \in \widetilde{B} \cap B^C\right) - P_{\theta_0}\left((X_1, ..., X_n) \in \left(\widetilde{B}\right)^C \cap B\right)\right) =$$

$$k_{1-\alpha}\left(P_{\theta_0}\left((X_1, ..., X_n) \in \widetilde{B}\right) - P_{\theta_0}\left((X_1, ..., X_n) \in B\right)\right) = k_{1-\alpha}\left(\alpha - \alpha\right) = 0$$

Deci

$$P_{\theta_1}\left((X_1, ..., X_n) \in \widetilde{B}\right) \geq P_{\theta_1}\left((X_1, ..., X_n) \in B\right)$$

adica $\widetilde{B} = B^*$.

$\blacksquare$

In concluzie, FORMA celei mai bune regiuni critice este

$$\begin{aligned} B^* &= \left\{(x_1, ..., x_n) \mid u_n\left(x_1, ..., x_n\right) \geq k\right\} \\ &= \left\{(x_1, ..., x_n) \mid \ln u_n\left(x_1, ..., x_n\right) \geq c\right\} \end{aligned}$$

iar constanta k (respectiv c) se determina din conditia ca pragul de semnificatie sa fie α,

$$P_{\theta_0}\left(u_n\left(X_1,...,X_n\right)<k\right)=1-\alpha$$

O versiune a teoremei Neyman - Pearson se obtine imediat pentru cazul discret,

$$F_\theta=\sum_{x\in A}p\left(x;\theta\right)\cdot\delta_{\{x\}}.$$

TESTUL RAPORTULUI PROBABILITATILOR PENTRU $H:\{\theta=\theta_0\}$, $H_A:\{\theta=\theta_1\}$

(a) Constructia lui B^*

- se calculeaza $u_n\left(x_1,...,x_n\right)$
- se determina $k=k_{1-\alpha}$ (respectiv $c=c_{1-\alpha}$) asa incat $P_{\theta_0}\left(u_n\left(X_1,...,X_n\right)<k_{1-\alpha}\right)=1-\alpha$

(b) Aplicarea testului

- Se observa $\left(x_1,...,x_n\right)$
- Se calculeaza valoarea numerica a lui $u_n\left(x_1,...,x_n\right)$
- Regula de decizie:

$$\begin{array}{lcl} u_n\left(x_1,...,x_n\right)\geq k_{1-\alpha} & \Longrightarrow & \text{se respinge } H:\{\theta=\theta_0\} \\ u_n\left(x_1,...,x_n\right)<k_{1-\alpha} & \Longrightarrow & \text{se accepta } H:\{\theta=\theta_0\} \end{array}$$

Valorile probabilitatilor de eroare:
Prin constructie,

$$P_{\theta_0}\left(\left(X_1,...,X_n\right)\in B^*\right)=\alpha$$

In virtutea teoremei Neyman - Pearson,

$$\beta=\beta_{\min}=P_{\theta_1}\left(u_n\left(x_1,...,x_n\right)<k_{1-\alpha}\right)$$

APLICATIA 1
T.R.P. pentru modelul $B(1;\theta)$, $\theta \in (0,1)$

$$F_\theta = \sum_{x=0}^{1} \theta^x (1-\theta)^{1-x} \cdot \delta_{\{x\}},\ \theta \in (0,1)$$

Consideram $0 < \theta_0 < \theta_1 < 1$ si ipotezele $H : \{\theta = \theta_0\}$, $H_A : \{\theta = \theta_1\}$.

$$u_n(x_1, ..., x_n) = \prod_{i=1}^{n} \frac{\theta_1^{x_i}(1-\theta_1)^{1-x_i}}{\theta_0^{x_i}(1-\theta_0)^{1-x_i}} = \left(\frac{\theta_1}{\theta_0}\right)^{\sum_{i=1}^n x_i} \left(\frac{1-\theta_1}{1-\theta_0}\right)^{n-\sum_{i=1}^n x_i}$$

$$\ln u_n(x_1, ..., x_n) = \sum_{i=1}^{n} x_i \cdot \ln \frac{\theta_1(1-\theta_0)}{\theta_0(1-\theta_1)} + n \ln \frac{1-\theta_1}{1-\theta_0}$$

Pentru $\alpha \in (0,1)$ arbitar fixat, forma celei mai bune regiuni critice pentru H la pragul de semnificatie α este

$$B^* = \{\ln u_n(x_1, ..., x_n) \geq c\} = \left\{\sum_{i=1}^{n} x_i \geq C\right\}$$

unde

$$C = \frac{1}{\ln \frac{\theta_1(1-\theta_0)}{\theta_0(1-\theta_1)}} \left(c - n \ln \frac{1-\theta_1}{1-\theta_0}\right)$$

Determinam constanta C asa incat $B^* \in \mathcal{C}_\alpha$.

Pentru $\theta = \theta_0$, repartitia variabilei aleatoare $\sum_{i=1}^n X_i$ este binomiala, $B(n;\theta_0)$.

Fie $C_{1-\alpha}$ cuantila de rang $(1-\alpha)$ a acestei repartitii,

$$\sum_{\substack{y=0 \\ y<C_{1-\alpha}}}^{n} C_n^y \theta_0^y (1-\theta_0)^{n-y} \leq 1-\alpha$$

$$\sum_{\substack{y=0 \\ y\leq C_{1-\alpha}}}^{n} C_n^y \theta_0^y (1-\theta_0)^{n-y} \geq 1-\alpha$$

Rezulta

$$B^* = \left\{(x_1, ..., x_n) \mid \sum_{i=1}^{n} x_i \geq C_{1-\alpha}\right\}$$

si avem

$$P_{\theta_0}\left((X_1,...,X_n)\in B^*\right)=P_{\theta_0}\left(\sum_{i=1}^n X_i\geq C_{1-\alpha}\right)=1-P_{\theta_0}\left(\sum_{i=1}^n X_i<C_{1-\alpha}\right)\geq\alpha$$

$$P_{\theta_0}\left(\sum_{i=1}^n X_i>C_{1-\alpha}\right)=1-P_{\theta_0}\left(\sum_{i=1}^n X_i\leq C_{1-\alpha}\right)\leq\alpha$$

$$\beta_{\min}=P_{\theta_1}\left((X_1,...,X_n)\in(B^*)^C\right)=\sum_{\substack{y=0\\ y<C_{1-\alpha}}}^n C_n^y\theta_1^y\left(1-\theta_1\right)^{n-y}$$

APLICATIA 2
T.R.P. pentru modelul $N(\theta,1)$, $\theta\in R$

$$f(x;\theta)=\frac{1}{\sqrt{2\pi}}\exp\left\{-\frac{1}{2}(x-\theta)^2\right\},\ x\in R;\ \theta\in R$$

Consideram $\theta_0<\theta_1$ si ipotezele $H:\{\theta=\theta_0\}$, $H_A:\{\theta=\theta_1\}$.

$$u_n(x_1,...,x_n)=\prod_{i=1}^n\frac{\frac{1}{\sqrt{2\pi}}\exp\left\{-\frac{1}{2}(x_i-\theta_1)^2\right\}}{\frac{1}{\sqrt{2\pi}}\exp\left\{-\frac{1}{2}(x_i-\theta_0)^2\right\}}$$
$$=\exp\left\{(\theta_1-\theta_0)\sum_{i=1}^n x_i-\frac{n}{2}\left(\theta_1^2-\theta_0^2\right)\right\}$$

$$\ln u_n(x_1,...,x_n)=(\theta_1-\theta_0)\sum_{i=1}^n x_i-\frac{n}{2}\left(\theta_1^2-\theta_0^2\right)$$

Pentru $\alpha\in(0,1)$ arbitar fixat, forma celei mai bune regiuni critice pentru H la pragul de semnificatie α este

$$B^*=\{\ln u_n(x_1,...,x_n)\geq c\}=\left\{\sum_{i=1}^n x_i\geq C\right\}$$

unde

$$C=\frac{1}{\theta_1-\theta_0}\left(c+\frac{n}{2}\left(\theta_1^2-\theta_0^2\right)\right)$$

Determinam constanta C asa incat $B^*\in\mathcal{C}_\alpha$.

Pentru $\theta = \theta_0$, repartitia variabilei aleatoare $\sum_{i=1}^{n} X_i$ este normala, $N(n\theta_0, n)$. Rezulta

$$\frac{1}{\sqrt{n}}\left(\sum_{i=1}^{n} X_i - n\theta_0\right) \ \sim \ N(0,1)$$

Pentru determinarea constantei C impunem conditia

$$P_{\theta_0}\left(\frac{1}{\sqrt{n}}\left(\sum_{i=1}^{n} X_i - n\theta_0\right) < \frac{1}{\sqrt{n}}(C - n\theta_0)\right) = 1 - \alpha$$

Fie $z_{1-\alpha}$ cuantila de rang $(1-\alpha)$ a repartitiei $N(0,1)$. Rezulta

$$\frac{1}{\sqrt{n}}(C - n\theta_0) = z_{1-\alpha},$$

$$C = \sqrt{n} z_{1-\alpha} + n\theta_0$$

Cea mai buna regiune critica la pragul de semnificatie α este

$$B^* = \left\{\sum_{i=1}^{n} x_i \geq \sqrt{n} z_{1-\alpha} + n\theta_0\right\} = \left\{\overline{x} \geq \theta_0 + \frac{1}{\sqrt{n}} z_{1-\alpha}\right\}$$

si probabilitatile de eroare sunt

$$P_{\theta_0}\left((X_1, ..., X_n) \in B^*\right) = P_{\theta_0}\left(\sum_{i=1}^{n} X_i \geq \sqrt{n} z_{1-\alpha} + n\theta_0\right) = \alpha$$

$$\beta_{\min} = P_{\theta_1}\left(\sum_{i=1}^{n} X_i < \sqrt{n} z_{1-\alpha} + n\theta_0\right) =$$

$$= P_{\theta_1}\left(\frac{\sum_{i=1}^{n} X_i - n\theta_1}{\sqrt{n}} < \frac{\sqrt{n} z_{1-\alpha} + n\theta_0 - n\theta_1}{\sqrt{n}}\right) =$$

$$= P_{\theta_1}\left(\frac{\sum_{i=1}^{n} X_i - n\theta_1}{\sqrt{n}} < z_{1-\alpha} - \sqrt{n}(\theta_1 - \theta_0)\right) = F_{N(0,1)}\left(z_{1-\alpha} - \sqrt{n}(\theta_1 - \theta_0)\right)$$

TESTE PENTRU IPOTEZE SIMPLE CU ALTERNATIVE COMPUSE

Fie modelul $F_\theta = P_\theta \circ X^{-1}$, $\theta \in \Theta \subseteq R^k$, $k \geq 1$ si fie $\theta^0 = (\theta_1^0, ..., \theta_k^0)' \in \Theta$.

Ne propunem sa testam ipoteza simpla

$$H : \{\theta = \theta_0\}$$

cu alternativa compusa

$$H_A : \{\theta \in \Theta - \{\theta_0\}\} = \{\theta \neq \theta_0\}.$$

Fie sirul observatiilor independente, identic repartizate $(X_1, X_2.....)$ si, pentru primele n observatii, notam cu $L(x_1, ..., x_n; \theta)$ functia de verosimilitate.

$$L(x_1, ..., x_n; \theta) = \begin{cases} \prod_{i=1}^{n} p(x_i; \theta), & \text{in caz discret} \\ \prod_{i=1}^{n} f(x_i; \theta), & \text{in caz continuu} \end{cases}$$

In conditii de regularitate pentru L ca functie in θ, scriem sistemul de verosimilitate maxima

$$\frac{\partial \ln L}{\partial \theta_i} = 0, \quad i = 1, ..., k$$

Notam cu $\widehat{\theta}_{VM}(X_1, ..., X_n)$ estimatorul de verosimilitate maxima, determinat pentru selectii $n-$dimensionale.

Numim raport al verosimilitatilor functia

$$\Lambda(x_1, ..., x_n) = \frac{L(x_1, ..., x_n; \theta_0)}{L\left(x_1, ..., x_n; \widehat{\theta}_{VM}(x_1, ..., x_n)\right)}$$

TEOREMA (cazul $k = 1$)

Fie $\{X_n, n \geq 1\}$ un sir de variabile aleatoare independente, identic repartizate $F_\theta = P_\theta \circ X^{-1}$, $\theta \in \Theta \subseteq R$ si fie $\theta_0 \in \Theta$ valoarea adevarata a parametrului. Presupunem verificate urmatoarele conditii:

1. Θ este un interval deschis al lui R;

2. F_θ admite densitatea de repartitie $f(x;\theta)$ si $\{x \mid f(x;\theta) > 0\}$ este independenta de θ;

3. Exista o vecinatate V a lui θ_0 asa incat pentru orice $\theta \in V$ avem:

- functia $f(x;\theta)$ este de trei ori derivabila in raport cu θ oricare ar fi x si derivatele sunt integrabile;

- exista functiile G_1, G_2 si $H(\cdot, \theta)$ integrabile pe R asa incat

$$\left|\frac{\partial f(x;\theta)}{\partial \theta}\right| < G_1(x)$$

$$\left|\frac{\partial^2 f(x;\theta)}{\partial \theta^2}\right| < G_2(x)$$

$$\left|\frac{\partial^3 f(x;\theta)}{\partial \theta^3}\right| < H(x,\theta)$$

$$\int_R H(x,\theta) f(x;\theta)\, dx < K$$

unde K este o constanta independenta de θ;

- exista "informatia Fisher"

$$M_\theta \left(\frac{\partial f(X;\theta)}{\partial \theta}\right)^2 \overset{notat}{=} i_1(\theta)$$
$$0 < i_1(\theta) < \infty$$

Atunci, cu o probabilitate tinzand la 1, ecuatia de verosimilitate maxima

$$\frac{\partial \ln L}{\partial \theta} = 0$$

are o solutie $\widehat{\theta}_n(x_1, ..., x_n)$ asa incat au loc urmatoarele convergente pentru $n \longrightarrow \infty$:

$$\widehat{\theta}_n(X_1, ..., X_n) \overset{P_{\theta_0}}{\longrightarrow} \theta_0$$

$$\sqrt{n}\left(\widehat{\theta}_n(X_1, ..., X_n) - \theta_0\right) \overset{repartitie}{\longrightarrow} Y \sim N\left(0; \frac{1}{i_1(\theta_0)}\right)$$

$$-2\ln \Lambda(X_1, ..., X_n) \overset{repartitie}{\longrightarrow} Z \sim \chi^2(1)$$

(rezultatul va fi reluat la cursul de "Capitole de statistica matematica" de la Master)

Pentru demonstratie:
Craiu Virgil, Paunescu Virgil, "Elemente de statistica matematica cu aplicatii", Editura Mondo - Ec, 1998

EXTENSIA TEOREMEI in cazul $k > 1$ (parametrul θ este un vector $k-$dimensional) ofera pentru comportamentul asimptotic al raportului de verosimilitati urmatoarea concluzie:

$$-2\ln\Lambda\left(X_1,...,X_n\right) \overset{repartitie}{\longrightarrow} Z \ \sim \ \chi^2\left(k\right)$$

TESTUL RAPORTULUI DE VEROSIMILITATI PENTRU $H:\{\theta=\theta_0\}$, $H_A:\{\theta\neq\theta_9\}$

Algoritm:

- se observa $\left(x_1,...,x_n\right);$
- se calculeaza valorile $\widehat{\theta}_{VM}\left(x_1,...,x_n\right)$ si $\Lambda\left(x_1,...,x_n\right);$
- pentru $\alpha\in(0,1)$ arbitrar fixat, fie $h_{k;1-\alpha}$ cuantila de rang $(1-\alpha)$a repartitiei χ^2cu k grade de libertate. Daca

  $$-2\ln\Lambda\left(x_1,...,x_n\right)\geq h_{k;1-\alpha}$$

  decidem sa respingem ipoteza $H:\{\theta=\theta_0\}$.

Observatii:
Asimptotic, probabilitatea erorii de I tip (respingerea ipotezei H cand H este adevarata) este egala cu α.
Acesta este un test general, cāci repartitia limita a lui $-2\ln\Lambda\left(X_1,...,X_n\right)$ este independenta de model.

APLICATIE

T.R.V pentru modelul $N\left(\mu,\sigma^2\right),\ \theta=\left(\mu,\sigma^2\right)\in R\times(0,\infty)$

$$H:\left\{\theta=\left(\mu_0,\sigma_0^2\right)\right\},\quad H_A:\left\{\theta\neq\left(\mu_0,\sigma_0^2\right)\right\}$$

Functia de verosimilitate este

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

Reamintim ca E.V.M. pentru parametrii repartitiei normale sunt

$$\widehat{\mu}_{VM}\left(X_1,...,X_n\right)=\overline{X}=\frac{1}{n}\sum_{i=1}^{n}X_i$$

$$\widehat{\sigma^2}_{VM}\left(X_1,...,X_n\right)=\frac{1}{n}\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2$$

Raportul de verosimilitati este

$$\Lambda\left(x_1,...,x_n\right)=\frac{L\left(x_1,...,x_n;\mu_0,\sigma_0^2\right)}{L\left(x_1,...,x_n;\widehat{\mu}_{VM},\widehat{\sigma^2}_{VM}\right)}=$$

$$=\frac{\left(2\pi\sigma_0^2\right)^{-n/2}\exp\left\{-\frac{1}{2\sigma_0^2}\sum_{i=1}^{n}\left(x_i-\mu_0\right)^2\right\}}{\left(2\pi\widehat{\sigma^2}_{VM}\right)^{-n/2}\exp\left\{-\frac{1}{2\widehat{\sigma^2}_{VM}}\sum_{i=1}^{n}\left(x_i-\overline{x}\right)^2\right\}}=$$

$$=\left(\frac{\sigma_0^2}{\widehat{\sigma^2}_{VM}}\right)^{-n/2}\exp\left\{-\frac{1}{2\sigma_0^2}\sum_{i=1}^{n}\left(x_i-\mu_0\right)^2+\frac{n}{2}\right\}$$

$$\begin{aligned}&-2\ln\Lambda\left(x_1,...,x_n\right)\\&=n\ln\left(\frac{\sigma_0^2}{\widehat{\sigma^2}_{VM}}\right)+\sum_{i=1}^{n}\left(\frac{x_i-\mu_0}{\sigma_0}\right)^2-n\end{aligned}$$

Repartitia limita a lui $-2\ln\Lambda\left(X_1,...,X_n\right)$ pentru $n\to\infty$ este repartitia $\chi^2\left(2\right)$.

Pentru $\alpha\in(0,1)$ arbitrar fixat, fie $h_{2;1-\alpha}$ cuantila de rang $(1-\alpha)$ a repartitiei χ^2 cu 2 grade de libertate. Daca

$$-2\ln\Lambda\left(x_1,...,x_n\right)\geq h_{2;1-\alpha}$$

decidem sa respingem ipoteza $H:\left\{\theta=\left(\mu_0,\sigma_0^2\right)\right\}$.