Testiranje na podlagi razmerja verjetij

Okvir

Proučujemo slučajni eksperiment X, ki ima množico dopustnih porazdelitev z gostotami

$$f(\_;\vartheta),\ \text{kjer } \vartheta \in \Theta \subset \mathbb{R}^m.$$

Preizkušamo

$$\vartheta \in H_0 \text{ proti } \vartheta \in A,\ \text{kjer } A = \Theta \setminus H_0.$$

Ukvarjamo se z vzorcem dane velikosti n.

Funkcija verjetja

Funkcija verjetja $L\colon \mathbb{R}^n \times \Theta \to \mathbb{R}$:

$$L(x_1, x_2, \ldots, x_n; \vartheta) = f(x_1;\vartheta) \cdot f(x_2;\vartheta) \cdot \cdots \cdot f(x_n;\vartheta).$$

Razmerje verjetij ...

... za (ničelno) hipotezo H_0 parametričnega prostora Θ je

$$\lambda\colon \mathbb{R}^n \to [0,1],\ \ \lambda(x_1, \ldots, x_n) = \frac{\sup_{\vartheta \in H_0} L(x_1, x_2, \ldots, x_n; \vartheta)}{\sup_{\vartheta \in \Theta} L(x_1, x_2, \ldots, x_n; \vartheta)}.$$

Testiranje na podlagi razmerja verjetij

Razmerje verjetij ...

... za (ničelno) hipotezo H_0 parametričnega prostora Θ je

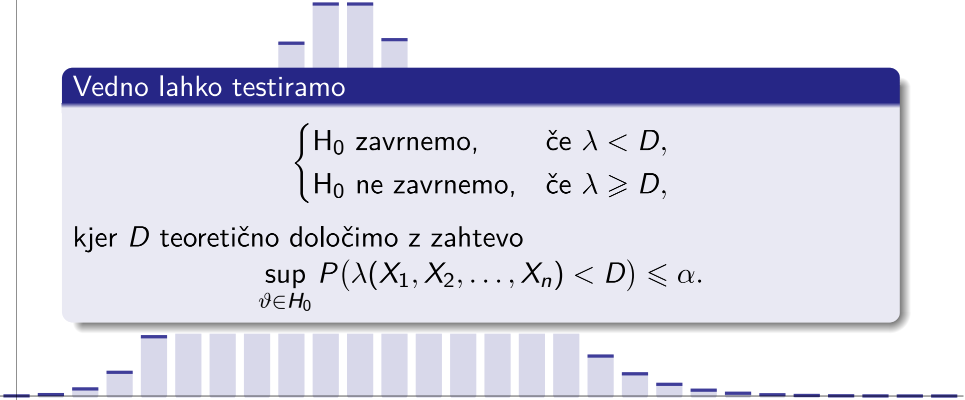
$$\lambda\colon \mathbb{R}^n \to [0,1], \quad \lambda(x_1,\ldots,x_n) = \frac{\sup_{\vartheta\in H_0} L(x_1,x_2,\ldots,x_n;\vartheta)}{\sup_{\vartheta\in\Theta} L(x_1,x_2,\ldots,x_n;\vartheta)}.$$

Kje je dosežen $\sup_{\vartheta\in\Theta} L(x_1,x_2,\ldots,x_n;\vartheta)$?
Pri cenilki največjega verjetja $\hat{\vartheta}$!

Filozofija verjetja

- Če je dejanski parameter $\vartheta \in H_0$, bo $\hat{\vartheta}$ „blizu“ H_0.

  Torej bo $\lambda(X_1,X_2,\ldots,X_n)$ blizu 1.
- Če je dejanski parameter ϑ „daleč stran“ od H_0, bo $\lambda(X_1,\ldots,X_n)$ blizu 0.



Vedno lahko testiramo

$$\begin{cases} \mathrm{H}_0 \text{ zavrnemo}, & \text{če } \lambda < D, \\ \mathrm{H}_0 \text{ ne zavrnemo}, & \text{če } \lambda \geqslant D, \end{cases}$$

kjer D teoretično določimo z zahtevo

$$\sup_{\vartheta \in H_0} P(\lambda(X_1, X_2, \ldots, X_n) < D) \leqslant \alpha.$$

- Parametrični prostor: $\Theta = \mathbb{R} \times (0, \infty)$.
- Ničelna hipoteza: $H_0 = \{\mu_0\} \times (0, \infty)$.
- Dopustne gostote: $f(x; \mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/(2\sigma^2)}$
- Funkcija verjetja za vzorec velikosti n:

$$L(x_1, x_2, \ldots, x_n; \mu, \sigma) = (\sigma^2 2\pi)^{-n/2} e^{-\sum_{i=1}^n (x_i-\mu)^2/(2\sigma^2)}$$

Cenilka največjega verjetja?

$$\hat{\mu} = \bar{x}, \qquad \hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^n (x_i - \bar{x})^2.$$

$$\Longrightarrow \ L(x_1, \ldots, x_n; \hat{\mu}, \hat{\sigma}) = (\hat{\sigma}^2 2\pi)^{-n/2} e^{-n/2} = \left(\frac{2\pi e}{n}\sum_{i=1}^n (x_i - \bar{x})^2\right)^{-\frac{n}{2}}$$

Testiranje na podlagi razmerja verjetij

Normalni model $N(\mu, \sigma)$, testiramo $\mu = \mu_0$ proti alternativi

Torej: $$\sup_{(\mu,\sigma)\in\Theta} L = \left(\frac{2\pi e}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2\right)^{-n/2}$$

Podobno

$$\sup_{(\mu,\sigma)\in\{\mu_0\}\times(0,\infty)} L = \left(\frac{2\pi e}{n}\sum_{i=1}^{n}(x_i - \mu_0)^2\right)^{-n/2}$$

Razmerje verjetij $\lambda(x_1, x_2, \ldots, x_n) =$

$$= \left(\frac{\sum_{i=1}^n (x_i - \mu_0)^2}{\sum_{i=1}^n (x_i - \bar{x})^2}\right)^{-\frac{n}{2}} = \left(\frac{\sum_{i=1}^n (x_i - \bar{x})^2 + n(\bar{x} - \mu_0)^2}{\sum_{i=1}^n (x_i - \bar{x})^2}\right)^{-\frac{n}{2}}$$

$$= \left(1 + \frac{(\bar{x} - \mu_0)^2}{\frac{1}{n}\sum_{i=1}^n (x_i - \bar{x})^2}\right)^{-\frac{n}{2}}$$

Testiranje na podlagi razmerja verjetij

Normalni model $N(\mu, \sigma)$, testiramo $\mu = \mu_0$ proti alternativi

Test

$$\begin{cases} H_0 \text{ zavrnemo,} & \text{če } \left(1 + \frac{(\bar{x}-\mu_0)^2}{\frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^2}\right)^{-\frac{n}{2}} < D, \\ H_0 \text{ ne zavrnemo,} & \text{če } \left(1 + \frac{(\bar{x}-\mu_0)^2}{\frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^2}\right)^{-\frac{n}{2}} \geqslant D \end{cases}$$

je ekvivalentnen standardnemu t-testu hipoteze $\mu = \mu_0$ proti alternativi.

Diskretna slučajna spremenljivka X

Naj bo $\operatorname{im} X = \{\xi_0, \xi_1, \ldots, \xi_m\}$ in $p_j = P(X = \xi_j)$.

Parametrični prostor

Simpleks $\left\{(p_0, p_1, p_2, \ldots, p_m) \,|\, p_j > 0,\ \sum_{j=0}^m p_i = 1\right\}$.

Porazdelitvena gostota enega eksperimenta

$$f(x; p_0, \ldots, p_m) = \sum_{j=0}^{m} \mathbb{1}_{\xi_j}(x) \cdot p_j.$$

Funkcija verjetja za vzorec velikosti n

$$L(x_1, \ldots, x_n; p_0, \ldots, p_m) = p_0^{T_0} p_1^{T_1} \ldots p_m^{T_m}.$$

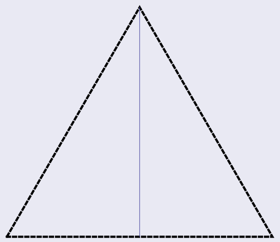
Tu je $T_j(x_1, \ldots, x_n) = \sum_{i=1}^n \mathbb{1}_{\xi_j}(x_i)$.

Testiranje čisto določene porazdelitve - prilagoditveni test

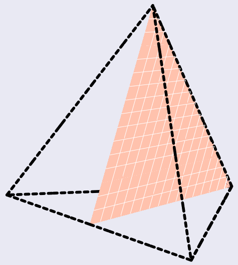
$$H_0 = \{(\pi_0, \dots, \pi_m)\}\,.$$

Na primer: igralna kocka je poštena; $\mathbf{p} = (\frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6})$.

$p_0 = p_1,\ m = 2$



$p_0 = p_1,\ m = 3$



Nastopanje ničelnih frekvenc

Privzemimo, da je $T_0 = 0$. Tedaj je izraz $L(\mathbf{x};\mathbf{p}) = p_1^{T_1} \dots p_m^{T_m}$ očitno maksimiziran pri $\hat{p}_0 = 0$.

Privzemimo $\forall j: \ T_j \neq 0$

1. Kot običajno, maksimiziramo $\log L = T_0 \log(p_0) + \dots T_m \log(p_m)$.
2. Ker gre za problem z vezjo $\sum_j p_j = 1$, nastavimo $$G = T_0 \log(p_0) + \dots + T_m \log(p_m) - \mu(\textstyle\sum_j p_j - 1).$$
3. Ekstremala za G je ena sama: $\mu = n$ in $\hat{p}_j = \frac{T_j}{n}$.
4. Izkaže se, da gre za maksimum.

Razmerje verjetij za $H_0 = \{(\pi_0, \ldots, \pi_m)\}$

$$\lambda_n(\mathbf{x}) = \frac{\pi_0^{T_0}\pi_1^{T_1}\ldots\pi_m^{T_m}}{(T_0/n)^{T_0}(T_1/n)^{T_1}\ldots(T_m/n)^{T_m}} = \prod_{j=0}^{m}\left(\frac{n\pi_j}{T_j}\right)^{T_j}.$$

Test

$$\begin{cases} \mathrm{H}_0 \text{ zavrnemo}, & \text{če } \prod_{j=0}^{m}\left(\frac{n\pi_j}{T_j}\right)^{T_j} < D, \\ \mathrm{H}_0 \text{ ne zavrnemo}, & \text{če } \prod_{j=0}^{m}\left(\frac{n\pi_j}{T_j}\right)^{T_j} \geqslant D, \end{cases}$$

kjer D teoretično določimo z zahtevo (vzamemo največjo možnost):

$$\sup_{\vartheta\in H_0} P(\lambda(X_1, X_2, \ldots, X_n) < D) \leqslant \alpha.$$

To pomeni $P\left(\prod_{j=0}^{m}\left(\frac{n\pi_j}{T_j}\right)^{T_j} < D\right) \leqslant \alpha$, kjer ima vektor $(T_0, T_1, \ldots, T_m)$ multinomsko porazdelitev $(n; \pi_0, \pi_1, \ldots, \pi_m)$.

Vedno lahko testiramo

$$\begin{cases} \mathrm{H}_0 \text{ zavrnemo}, & \text{če } \lambda < D, \\ \mathrm{H}_0 \text{ ne zavrnemo}, & \text{če } \lambda \geqslant D, \end{cases}$$

kjer D teoretično določimo z zahtevo

$$\sup_{\vartheta \in H_0} P(\lambda(X_1, X_2, \ldots, X_n) < D) \leqslant \alpha.$$

Problem?

V praksi je določitev konstante D tipično neizvedljiva.

Kaj je glavna vrlina testiranja na podlagi razmerja verjetij?

Obnašanje pri velikih vzorcih.

Izrek

1. Naj veljajo primerni regularnostni privzetki za funkcije $f(\_;\vartheta)$.
2. Naj bo Θ „lepa“ podmnožica v $\mathbb{R}^m$ (odprta množica ali „odprta“ večrazsežna ploskev).
3. Naj bo H_0 zaprta množica v Θ.

Če dejanski parameter θ pripada H_0, velja:

$$-2\log\lambda_n(X_1,\ldots,X_n) \xrightarrow[n\to\infty]{d} \chi^2_{\dim\Theta-\dim H_0}.$$

Asimptotični test

$$\begin{cases} \mathrm{H}_0 \text{ zavrnemo}, & \text{če } -2\log\lambda > C, \\ \mathrm{H}_0 \text{ ne zavrnemo}, & \text{če } -2\log\lambda \leqslant C. \end{cases}$$

Za C vzamemo zgornji percentil $\chi^2_{\dim\Theta-\dim H_0;\alpha}$.
Test je približne velikosti α.

1. $\lambda_n(\mathbf{x}) = \frac{\pi_0^{T_0}\pi_1^{T_1}\ldots\pi_m^{T_m}}{(T_0/n)^{T_0}(T_1/n)^{T_1}\ldots(T_m/n)^{T_m}} = \prod_{j=0}^{m}\left(\frac{n\pi_j}{T_j}\right)^{T_j}$.
2. $-2\log\lambda_n(\mathbf{x}) = \sum_{0\leqslant j\leqslant m;\ T_j\neq 0} 2T_j\big(\log(T_j) - \log(n\pi_j)\big)$.

Veliki vzorci

1. Če je dejanski parameter enak $(\pi_0, \ldots, \pi_m)$, velja $-2\log\lambda_n \xrightarrow{d} \chi^2_m$.
2. H_0 zavrnemo, če $-2\log\lambda_n > \chi^2_{m;\alpha}$, kjer $P(\chi^2_m > \chi^2_{m;\alpha}) = \alpha$.

Hi-kvadrat prilagoditveni test

Če je dejanski parameter enak $(\pi_0, \ldots, \pi_m)$, se limitna porazdelitev testnih statistik $-2\log\lambda_n$ ujema z limitno porazdelitvijo statistik:

- $\sum_{j=0}^{m} \frac{(T_j - n\pi_j)^2}{T_j}$ in
- $\sum_{j=0}^{m} \frac{(T_j - n\pi_j)^2}{n\pi_j}$.

Tu je T_j opažena frekvenca, $n\pi_j$ pa (pri H_0) pričakovana frekvenca.

V praksi tako testiramo $H_0 : (p_0, \ldots, p_m) = (\pi_0, \ldots, \pi_m)$ proti $(p_0, \ldots, p_m) \neq (\pi_0, \ldots, \pi_m)$:

$$\begin{cases} H_0 \text{ zavrnemo,} & \text{če } \sum_{j=0}^{m} \frac{(T_j - n\pi_j)^2}{n\pi_j} > \chi^2_{m;\alpha}, \\ H_0 \text{ ne zavrnemo,} & \text{če } \sum_{j=0}^{m} \frac{(T_j - n\pi_j)^2}{n\pi_j} \leqslant \chi^2_{m;\alpha}. \end{cases}$$

Test je približne velikosti α.