Statistički testovi

Ante Sosa

16. kolovoza 2019.

Sadržaj

1 Uvod

Ovaj projekt je napravljenu svrhu zadaće iz kolegija Matematički Softver.

2 Testiranje statističkih hipoteza

2.1 Statistička hipoteza

Promatramo statističko obilježje X. Statistička hipoteza je bilo koja pretpostavka o razdiobi od X. Kažemo da je statistička hipoteza jednostavna ukoliko jednoznačno određuje razdiobu od X. U suprotnom kažemo da je složena.

Primjer 1. Složena hipoteza
$H_1 : X$ ima normalnu razdiobu

Primjer 2. Jednostavna hipoteza
$H_2 : X \sim N(170, 64)$

3 Statistički testovi

3.1 Konstrukcija testa

Konstrukcija testa sastoji se od određivanja testne statistike na osnovi čijih vrijednosti se donose odluke, i (slike) kritičnog područja koji je skup onih mogućih vrijednosti testne statistike za koje se odbacuje H_0 u korist H_1. Takav skup također zovemo **kritičnim područjem** (Izvod optimalnog (tj. uniformno najjačeg testa).)[2]

3.2 Pogreške

Primjer 3. Želimo odrediti jeli točna hipoteza: $H_0 : \mu < 0$
Preciznije, želimo na osnovi realizacije slučajnog uzorka za X donijeti odluku hoćemo li odbaciti ili ne odbaciti tu hipotezu.

Postupak donošenja odluke o odbacivanju ili ne odbacivanju statističke hipoteze zove se testiranje statističkih hipoteza.

Budući da sve odluke bazirane na uzorcima iz populacije nisu 100% pouzdane, ni odluka statističkog testa nije 100% pouzdana. Dakle, može se dogoditi da je zaključak testa pogrešan.

	Zaključak	
Točno je	ne odbaciti H_0	odbaciti H_0
H_0	✓	pogrešno!(I)
H_1	pogrešno!(II)	✓

Pogreška koju činimo kada odbacujemo H_0, a ona je istinita, je **pogreška prve vrste**. Pogreška koju činimo kada ne odbacujemo H_0, a istinita je H_1, je **pogreška druge vrste**.

Test će u potpunosti biti sproveden ako možemo procjeniti vjerojatnosti mogućih pogrešaka u zaključku testa. Razumno je zahtjevati test kojemu se mogu kontrolirati vjerojatnosti obiju pogrešaka. To nije moguće jer smanjivanjem vjerojatnosti pogreške prve vrste povećava se vjerojatnost pogreške druge vrste i obratno.[2]

3.3 Razina značajnosti

S druge strane, u velikoj većini slučajeva moguće je za zadanu **razinu značajnosti testa** $\alpha(\alpha \in \langle a, b \rangle)$ među testovima kojima vjerojatnost pogreške prve vrste ne prelazi broj α naći (konstruirati) test s najmanjom vjerojatnosti pogreške druge vrste. Neka je

$X_1, X_2, \ldots, X_n$ slučajni uzorak za X i

$$X := (X_1, X_2, \ldots, X_n).$$

Tada su realizacije $x = (x_1, x_2, \ldots, x_n)$ tog uzorka elementi $\mathbb{R}^n$.

Definicija 4 (Test)**.** Skup **Test** (hipoteze H_0 u odnosu na alternativu H_1) je preslikavanje $\tau : R^n \to \{0, 1\}$.

Interpretacija 5. *Ako je za realizaciju x uzorka X $\tau(x) = 1$, tada odbacujemo H_0 u korist H_1, a ako je $\tau(x) = 0$, tada ne odbacujemo H_0 u korist H_1.*

Tada je

$$C := \tau^{-1}(1) = \{x \in R^n : \tau(x) = 1\}$$

područje realizacija uzoraka za koje se H_0 odbacuje u korist H_1. C se naziva **kritično područje** za test τ.

Populacijska razdioba: $X \sim f(x|\theta), \theta \in \Theta$ Vjerodostojnost od $\theta : L(\theta|x) = \prod_{i=1}^{n} f(x_i|\theta)$. Preslikavanje $\gamma : \Theta \to [0, 1]$ definirano sa:

$$\gamma(\Theta) := E_\theta[\tau(X)] = P_\theta(XinC) = \int_C L(\theta|x)dx \tag{1}$$

(1) se zove **jakost testa** τ. Interpretacija. Ukoliko je θ_1 vrijednost parametra za koju je H_1 istinito, jakost testa $\gamma(\theta_1)$ je sposobnost testa da odbaci H_0 ako je H_0 neistinita hipoteza.

Neka su:

$$H_0 : \theta \in \Theta_0$$
$$H_1 : \theta \in \Theta_1$$

Preslikavanje $\alpha : \Theta_0 \to [0, 1]$ definirano sa:

$$\alpha(\theta) := \gamma(\theta) = P_\theta(X \in C)$$

je vjerojatnost pogreške prve vrste.

$$\alpha_\tau := \sup_{\theta \in \Theta_0} \alpha(\theta)$$

je **značajnost** testa τ. Kažemo da test ima razinu značajnosti α ukoliko mu je značajnost manja ili jednaka α.

$$H_0 : \theta \in \Theta_0$$
$$H_1 : \theta \in \Theta_1$$

Preslikavanje $\beta : \Theta_1 \to [0, 1]$ definirano sa:

$$\beta(\theta) := 1 - \gamma(\theta) = P_\theta(X \notin C)$$

je vjerojatnost pogreške druge vrste.[1]

Definicija 6 (Uniformno najjači test)**.** Kažemo da je test **uniformno najjači** ako za svaki drugi test τ' takav da je $\alpha_{\tau'} \leq \alpha_{\tau'}$, vrijedi da je $\gamma(\theta) \leq \gamma(\theta)$ za sve θ.

4 t-test

T-test je jedan od najpoznatijih statističkih postupaka, osnovan je na Studentovoj ili t razdiobi. Odnosi se na testiranje statističke značajnosti razlike između dvije aritmetičke sredine. Dobivena se razlika između obje aritmetičke sredine podijeli standardnom pogreškom te razlike.

Neka je $X_1, \ldots, X_n$ slučajni uzorak za $X \sim N(\mu, \sigma^2)$. Želimo testirati hipoteze o parametru $\mu(\sigma^2$ je nepoznat). Razlikujemo tri slučaja:

$$\begin{array}{ccc} (I) & (II) & (III) \\ H_0 : \mu = \mu_0 & H_0 : \mu = \mu_0 & H_0 : \mu = \mu_0 \\ H_1 : \mu > \mu_0 & H_1 : \mu < \mu_0 & H_1 : \mu \neq \mu_0 \end{array} \tag{2}$$

Alternative u (I) i (II) su jednostrane $\to$ jednostrani testovi. Dok je alternativa u (III) dvostrana $\to$ dvostrani test.

U sva tri slučaja, testna statistika je jednaka:

$$T = \frac{\bar{X} - \mu_0}{S}\sqrt{n} \overset{H_0}{\sim} t(n-1)$$

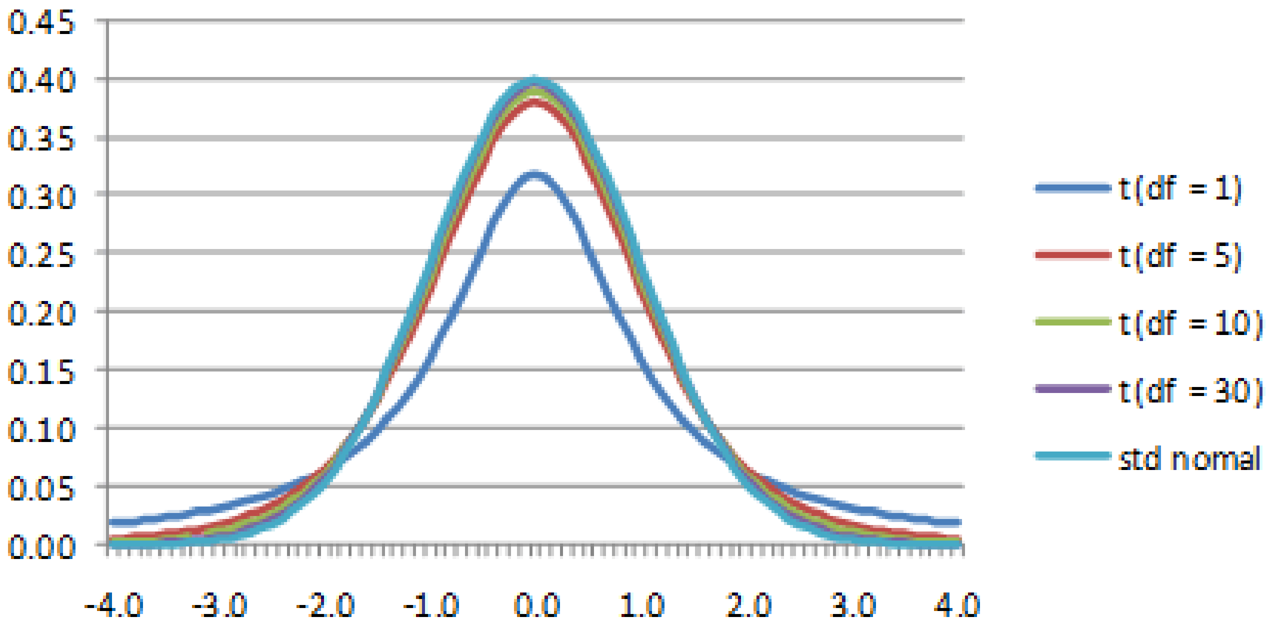
Kritična područja se razlikuju. Neka je α zadana razina značajnosti.

$$\begin{array}{rr} (I) & (II) \\ P(T \geq t_\alpha(n-1)|H_0) = \alpha & P(T \leq -t_\alpha(n-1)|H_0) = \alpha \\ \Rightarrow \text{kritično područje:} & \Rightarrow \text{kritično područje:} \\ [t_\alpha(n-1), +\infty\rangle & \langle -\infty,\ -t_\alpha(n-1)] \end{array}$$

$$(III)$$
$$P(|T| \geq t_{\frac{\alpha}{2}}(n-1)|H_0) = \alpha$$
$$\Rightarrow \text{kritično područje:}$$
$$\langle -\infty,\ -t_{\frac{\alpha}{2}}(n-1)] \bigcup [t_{\frac{\alpha}{2}}(n-1), +\infty\rangle$$

T-test je osnovan na Studentovoj t razdiobi.

Slika 1: Studentova t distribucija



4.1 Usporedba očekivanja dviju normalno distribuiranih populacija

Pretpostavimo da mjerimo isto statističko obilježje X, ali u dvije različite populacije. Nadalje, pretpostavimo da:

- u obje populacije X je normalno distribuirana varijabla
- s jednakim (populacijskim) varijancama.

Neka su:
X_1 = vrijednost varijable X na slučajno odabranoj jedinki iz populacije 1

X_2 = vrijednost varijable X na slučajno odabranoj jedinki iz populacije 1
Pretpostavke na $X \Rightarrow \quad X_1 \sim N(\mu_1, \sigma^2)$, $X_2 \sim N(\mu_2, \sigma^2)$

Neka su:

$X_{1_1}, X_{1_2}, \ldots, X_{1_{n_1}}$ slučajni uzorak za X_1 duljine n_1

$X_{2_1}, X_{2_2}, \ldots, X_{2_{n_1}}$ slučajni uzorak za X_2 duljine n_2

i ta dva uzorka su međusobno nezavisna.

Označimo sa: $\bar{X}_1, S_1^2$ i $\bar{X}_2, S_2^2$ aritmetičke sredine i uzoračke varijance uzorka za X_1 i X_2.

Želimo testirati gornja tri slučaja: (2) .

U sva tri slučaja, testna statistika je jednaka:

$$T = \frac{\bar{X}_1 - \bar{X}_2}{S_d} \cdot \frac{1}{\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \overset{H_0}{\sim} t(n_1 + n_2 - 2)$$

gdje je S_d procjenitelj (zajedničke) standardne devijacije σ na osnovi oba uzorka:

$$S_d = \sqrt{\frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}}$$

(S_d^2 je nepristrani procjenitelj varijance σ^2)

(1)

$$P(T \geq t_\alpha(n_1 + n_2 - 2)|H_0) = \alpha$$

$\Rightarrow$ kritično područje:

$$[t_\alpha(n_1 + n_2 - 2), +\infty\rangle$$

(2)

$$P(T \leq -t_\alpha(n_1 + n_2 - 2)|H_0) = \alpha$$

$\Rightarrow$ kritično područje:

$$\langle -\infty,\ -t_\alpha(n_1 + n_2 - 2)]$$

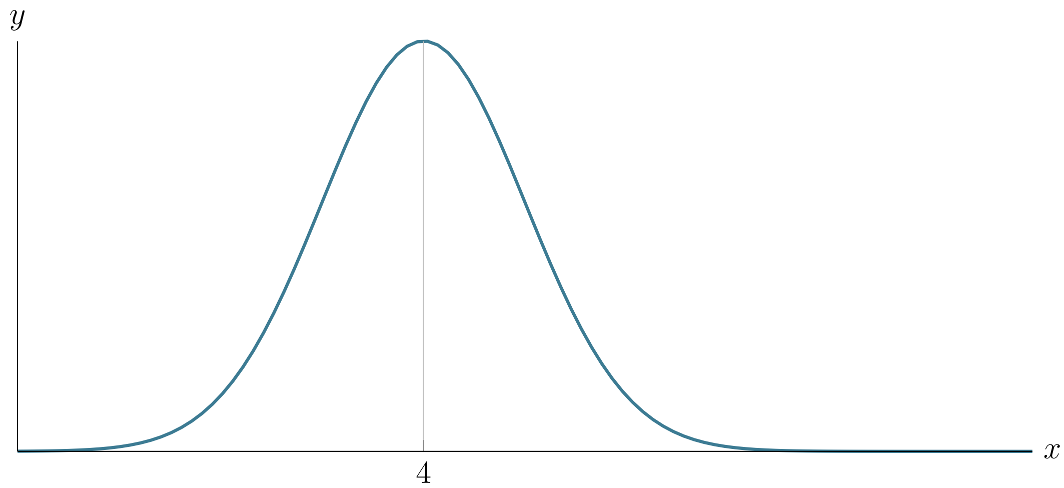
(3)

$$P(|T| \geq t_{\frac{\alpha}{2}}(n_1 + n_2 - 2)|H_0) = \alpha$$

$\Rightarrow$ kritično područje:

$$\langle -\infty,\ -t_{\frac{\alpha}{2}}(n_1 + n_2 - 2)] \bigcup [t_{\frac{\alpha}{2}}(n_1 + n_2 - 2), +\infty\rangle$$

Slika 2: Crtež t distribucije



Literatura

[1] Friedman, Jerome, Trevor Hastie, and Robert Tibshirani(2001), The elements of statistical learning Vol. 1. New York: Springer series in statistics,

[2] Slajdovi kolegija Statistika, `https://web.math.pmf.unizg.hr/nastava/stat/index.php?sadrzaj=predavanja.php`

[3] T-test using Python and Numpy, `https://towardsdatascience.com/inferential-statistics-series-t-test-using-numpy-2718f8f9bf2f`

Popis slika