Esercizio standard di Matematica e Statistica applicate alla Biologia

1. Sia $\mathcal{C}_N(X) = \{x_1, .., x_N\}$ un campione di dati di numerosità N corrispondenti alle realizzazioni di una v.a. gaussiana di valore atteso μ incognito e varianza σ^2 nota. Dato il numero reale μ_0, si supponga di voler testare l'ipotesi nulla
$$\mathbf{H}_0 : \mu = \mu_0$$
contro l'alternativa
$$\mathbf{H}_1 : \mu \neq \mu_0 \ .$$
   (a) Si verifichi l'ipotesi per un livello di confidenza α.
   (b) Si calcoli il p-dei dati *(p-value)* p e si indichi per quali valori del livello di confidenza α i dati del campione supportano l'ipotesi nulla.

   Risp. La media campionaria dei dati risulta
$$\bar{x} = \frac{1}{N} \sum_{i=1}^{N} x_i \ .$$
   (a) i. Se Z è una v.a. con distribuzione di probabilità normale standard $\mathbb{P}$,
$$\alpha = \mathbb{P}\left(|Z| > z_{\frac{\alpha}{2}}\right) = 2\mathbb{P}\left(Z > z_{\frac{\alpha}{2}}\right) \ ,$$
   quindi l'ipotesi è da accettarsi se
$$\sqrt{N}\frac{|\bar{x} - \mu_0|}{\sigma} \leq z_{\frac{\alpha}{2}} \ .$$
   ii.
$$p = \mathbb{P}\left(|Z| > \sqrt{N}\frac{|\bar{x} - \mu_0|}{\sigma}\right) = 2\mathbb{P}\left(Z > \sqrt{N}\frac{|\bar{x} - \mu_0|}{\sigma}\right)$$
   e i dati di $\mathcal{C}_N(X)$ supportano l'ipotesi nulla per tutti i livelli di confidenza $\alpha < p$.

2. Sia $\mathcal{C}_N(X) = \{x_1, .., x_N\}$ un campione di dati di numerosità N corrispondenti alle realizzazioni di una v.a. gaussiana di valore atteso μ e varianza

σ^2 incogniti. Dato il numero reale μ_0, si supponga di voler testare l'ipotesi nulla

$$\mathbf{H}_0 : \mu = \mu_0$$

contro l'alternativa

$$\mathbf{H}_1 : \mu \neq \mu_0 \ .$$

(a) Si verifichi l'ipotesi per un livello di confidenza α.

(b) Si calcoli il p-dei dati *(p-value)* p e si indichi per quali valori del livello di confidenza α i dati del campione supportano l'ipotesi nulla.

Risp. La media, la varianza e la deviazione standard campionarie dei dati risultano rispettivamente

$$\begin{aligned} \bar{x} &= \frac{1}{N}\sum_{i=1}^{N} x_i \ , \\ s_X^2 &= \frac{1}{N-1}\sum_{i=1}^{N}(x_i - \bar{x})^2 \ , \ s_X = \sqrt{s_X^2}. \end{aligned}$$

(a) i. Se T è una v.a. con distribuzione di probabilità di Student t con $N-1$ gradi di libertà $\mathbb{P}_{N-1}$,

$$\alpha = \mathbb{P}_{N-1}\left(|T| > t_{\frac{\alpha}{2},N-1}\right) = 2\mathbb{P}_{N-1}\left(T > t_{\frac{\alpha}{2},N-1}\right) \ ,$$

quindi l'ipotesi è da accettarsi se

$$\sqrt{N}\frac{|\bar{x} - \mu_0|}{s_X} \leq t_{\frac{\alpha}{2},N-1} \ .$$

ii.

$$p = \mathbb{P}_{N-1}\left(|T| > \sqrt{N}\frac{|\bar{x} - \mu_0|}{s_X}\right) = 2\mathbb{P}_{N-1}\left(T > \sqrt{N}\frac{|\bar{x} - \mu_0|}{s_X}\right)$$

e i dati di $\mathcal{C}_N(X)$ supportano l'ipotesi nulla per tutti i livelli di confidenza $\alpha < p$.

3. Sia $\mathcal{C}_N(X) = \{x_1, .., x_N\}$ un campione di dati di numerosità N corrispondenti alle realizzazioni di una v.a. gaussiana di valore atteso μ e varianza σ^2 incogniti. Dato il numero reale positivo σ_0, si supponga di voler testare l'ipotesi nulla

$$\mathbf{H}_0 : \sigma^2 \leq \sigma_0^2$$

contro l'alternativa

$$\mathbf{H}_1 : \sigma^2 > \sigma_0^2 \ .$$

(a) Si verifichi l'ipotesi per un livello di confidenza α.

(b) Si calcoli il p-dei dati *(p-value)* p e si indichi per quali valori del livello di confidenza α i dati del campione supportano l'ipotesi nulla.

Risp. La media e la varianza campionarie dei dati risultano rispettivamente

$$\begin{aligned} \bar{x} &= \frac{1}{N}\sum_{i=1}^{N} x_i \ , \\ s_X^2 &= \frac{1}{N-1}\sum_{i=1}^{N}(x_i-\bar{x})^2 \ . \end{aligned}$$

i. Se Y è una v.a. con distribuzione di probabilità χ^2 con $N-1$ gradi di libertà $\mathbb{P}_{N-1}$,

$$\alpha = \mathbb{P}_{N-1}\left(Y > \chi^2_{\alpha,N-1}\right) \ ,$$

quindi l'ipotesi è da accettarsi se

$$(N-1)\,\frac{s_X^2}{\sigma_0^2} \le \chi^2_{\alpha,N-1} \ .$$

ii.

$$p = \mathbb{P}_{N-1}\left(Y > (N-1)\,\frac{s_X^2}{\sigma_0^2}\right)$$

e i dati di $\mathcal{C}_N(X)$ supportano l'ipotesi nulla per tutti i livelli di confidenza $\alpha < p$.