TEST STATISTICO

Un test statistico è una procedura per VERIFICARE IPOTESI sul parametro ϑ (la legge) di un campione aleatorio.

e.g. Bernoulli ①

Test per infezione: contiamo quanti segmenti di DNA vengono colorati dal reagente

Se ci sono n segmenti abbiamo esiti $x_1, \dots, x_n \in \{0,1\}$

Consideriamo un campione di variabili aleatorie Bernoulli (ϑ) $\vartheta \in [0,1]$

Vogliamo stimare ϑ:

Il paziente è infetto se almeno l'1% dei segmenti sono colorati

C'è un'influenza della casualità su come scelgo il campione. Bisogna valutarne la rilevanza.

e.g. Bernoulli ②

Controllo qualità: vogliamo confrontare la produzione di due stabilimenti.

Consideriamo misurazioni $x_1, \dots, x_n \in \{0,1\}$ (0 corretto, 1 difettoso)

e $y_1, \dots, y_m \in \{0,1\}$ (secondo stabilimento)

A prima vista potremmo stimare i parametri θ_1 e θ_2 di campioni indipendenti, $\theta_1, \theta_2 \in (0,1)$

$X_1, \dots, X_n$ e $Y_1, \dots, Y_m$ rispettivamente Bernoulli (θ_1) e Bernoulli (θ_2)

In effetti però siamo interessati al singolo parametro $\theta_1 - \theta_2 = \theta \in (-1,1)$

Dobbiamo decidere se accettare o meno l'ipotesi $\theta = \emptyset$

e.g. Errori di misurazioni fisiche: effettuiamo n misurazioni di una qualità fisica (lunghezza) ③

Abbiamo esiti $x_1, \dots, x_n \in \mathbb{R}$ che descriviamo con

un campione $X_1, \dots, X_n$ Gaussiano $N(m, \sigma^2)$

Dobbiamo decidere se accettare o meno l'ipotesi che m prende un certo valore proposto

Def

Consideriamo un campione $X_1, \dots, X_n$ con legge dipendente da un parametro $\theta \in \Theta \subseteq \mathbb{R}$,

assumiamo di avere una PARTIZIONE $\Theta = \Theta_0 \cup \Theta_1$, $\Theta_0 \cap \Theta_1 = \emptyset$

diciamo che le affermazioni

$H_0)$ $\theta \in \Theta_0$

$H_1)$ $\theta \in \Theta_1$

sono rispettivamente l'ipotesi NULLA e l'ipotesi ALTERNATIVA

↳ Quelle sotto a variazioni dei dati che dipendono dalla casualità dell'estrazione

C'è un'ASIMMETRIA tra H_0 e H_1 nella procedura che scegliamo:

decidiamo se accettare o meno H_0

Un test per H_0 è dato da un evento $C \subseteq \Omega$, "regione omega"

che se si realizza conduce al rifiuto di H_0:

1. Se i dati $x_1, \dots, x_n$ sono realizzati da $X_1(\omega), \dots, X_n(\omega)$ per un qualche $\omega \in C$
   RIFIUTIAMO H_0 (test NEGATIVO)
2. Se $(x_1, \dots, x_n) \neq (X_1(\omega), \dots, X_n(\omega))$ $\forall\, \omega \in C$ allora
   ACCETTIAMO H_0 (test POSITIVO)

Quindi: rifiutando H_0 in senso stretto diciamo di non avere prove a sufficienza e NON che abbiamo prove per H_1

E.g. le regioni critiche negli esempi precedenti possono essere scelte come:

① e ② $\{ \overline{X}_n > d \}$ con d a scelta

③ $\{ |\overline{X}_n - m_0| > d \}$ con d a scelta

Fissato $\alpha \in (0,1)$, diciamo che il test di regione critica C ha livello α

se $\forall\, \theta \in \Theta_0$, $\underbrace{P_\theta(C) \le \alpha}_{\text{Probabilità di rifiutare } H_0 \text{ quando } H_0 \text{ è verificata (falso positivo)}}$ (usualmente $\alpha = 0{,}05$ o $\alpha = 0{,}01$)

La **POTENZA** del test é la funzione $\Theta_1 \ni \theta \longmapsto \mathbb{P}_\theta(C) \in [0,1]$

Vogliamo fissare il livello (porre un limite superiore all'errore di prima specie)

e massimizzare la potenza (minimizzare al contempo l'errore di seconda specie)

Z - TEST (esempio ③)

Test per la media di $N(m, \sigma^2)$ con σ nota.

Consideriamo ipotesi $H_0)\ m = m_0$

$H_1)\ m \neq m_0$

Cerchiamo la regione critica nella forma $C = \{|\bar{X}_n - m_0| > d\}$ e determiniamo d imponendo livello α e massimizzando la potenza

$\mathbb{P}_{m_0}(|\bar{X}_n - m_0| > d) \leq \alpha$ Prendiamo C il più grande possibile, quindi $= \alpha$

$$\alpha = \mathbb{P}_{m_0}(|\bar{X}_n - m_0| > d) = \mathbb{P}_{m_0}\left(\left|\sqrt{n}\,\frac{\bar{X}_n - m_0}{\sigma}\right| > \frac{\sqrt{n}\,d}{\sigma}\right) =$$

$$= 1 - \mathbb{P}\left(|z_n| < \frac{\sqrt{n}\,d}{\sigma}\right) = 1 - \left(2\Phi\left(\frac{\sqrt{n}\,d}{\sigma}\right) - 1\right) = 2\left(1 - \Phi\left(\frac{\sqrt{n}\,d}{\sigma}\right)\right)$$

$$\Downarrow$$

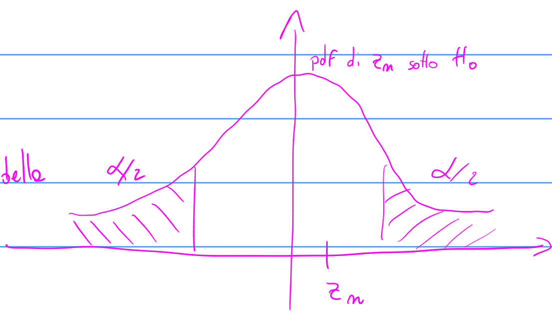
$$\Phi\left(\frac{\sqrt{n}\,d}{\sigma}\right) = 1 - \frac{\alpha}{2} \to d = \frac{\sigma}{\sqrt{n}}\, q_{1-\frac{\alpha}{2}}$$

Per effettuare il test:

- Fissare $\alpha = 0{,}05\%$
- Consideriamo gli esiti $x_1, \dots, x_n$
- Verifichiamo se é avvenuto l'evento $C = \{|\bar{X}_n - m_0| > \frac{\sigma}{\sqrt{n}}\, q_{1-\frac{\alpha}{2}}\}$

  ovvero se $|\bar{x}_n - m_0| > \frac{\sigma}{\sqrt{n}}\, q_{1-\frac{\alpha}{2}}$

  cioé, posto $z_n = \sqrt{n}\,\frac{\bar{x}_n - m_0}{\sigma}$ vale $|z_n| > q_{1-\frac{\alpha}{2}}$



IDEA: $\sqrt{n}\,\frac{\bar{X}_n - m_0}{\sigma}$ é $N(0,1)$ solo se $m = m_0$, ovvero se l'esito z_n della misurazione é verosimile sotto l'ipotesi nulla

P-VALUE

<u>Oss</u>: la scelta di α é arbitrario, peró:

① se α é molto piccolo, non avró quasi mai errore di prima specie ma C sará piccola e sará dunque facilissimo errore di seconda specie

② Viceversa se α non é piccolo abbastanza non sto controllando l'errore di prima specie

<u>DEF</u>: se abbiamo determinato una regione critica $C(\alpha)$ di livello α (per un'ipotesi nulla H_0) $\sigma \in (H_0)$)

dati gli esiti $X_1 \dots X_n$ dell'esperimento

il <u>P-VALUE</u> é il numero $\bar{\alpha}$ tale che:

- se $\alpha < \bar{\alpha}$ il test accetta H_0
- se $\alpha > \bar{\alpha}$ il test rifiuta H_0

<u>IDEA</u>: il p-value é la probabilitá calcolata sotto H_0 di ottenere esiti piú estremi di quelli misurati

Nel caso del test Z cerchiamo

$$\bar{\alpha} = \mathbb{P}_{m_0}\left(\left| \sqrt{n}\, \frac{\bar{X}_n - m_0}{\sigma} \right| > \left| \sqrt{n}\, \frac{\bar{x}_n - m_0}{\sigma} \right| \right) =$$

$$= \mathbb{P}\left(|Z_n| > |z_n| \right) = 2\left(1 - \Phi(|z_n|)\right)$$

<u>Solitamente</u>: si considera H_0 poco plausibile se il p-value é piccolo ($\lesssim 0.1$)