

Statistica inferenziale

Test di ipotesi

Bari, 7 Gennaio 2008

Metodi non parametrici

I test che abbiamo finora trattato, con eccezione dei test chi-quadrato, si basavano su qualche assunzione circa la densità di probabilità dei dati.

Con il termine "metodi non parametrici" o "distribution free" intendiamo tutte le tecniche che non fanno alcuna assunzione circa la densità delle variabili osservate.

Il test del segno

Vogliamo studiare la quantità di energia consumata a riposo da pazienti affetti da fibrosi cistica. Misuriamo il consumo di energia in soggetti con la patologia ed in soggetti sani appaiati per alcune caratteristiche.

C	D	S	Δ	segno
1	1153	996	157	+
2	1132	1080	52	+
3	1165	1182	-17	-
4	1460	1452	8	+
5	1634	1162	472	+
6	1493	1619	-126	-
7	1358	1140	218	+
8	1453	1123	330	+
9	1185	1113	72	+
10	1824	1463	361	+
11	1793	1632	161	+
12	1930	1614	316	+
13	2075	1836	239	+

Razionale

Se il consumo di energia a riposo tra soggetti sani e malati non cambia (H_0) allora $\mathbb{P}\{\Delta \leq 0\} = \mathbb{P}\{\Delta > 0\}$ o, equivalentemente, che la mediana della v.a. Δ é 0.

In questa ipotesi, detto P il numero di segni $+$, allora P é $B(n, \frac{1}{2})$ dove $n = 13$ é il numero di coppie considerate.

Osservazione. Se avessimo trovato coppie con $\Delta = 0$, tali coppie sarebbero state escluse dall'analisi e quindi n sarebbe stato ridotto di conseguenza.

Test

Se (H_0) é vera, allora $\mathbb{E}(P) = \frac{13}{2} = 6.5$, ossia il numero di segni $+$ che ci aspettiamo di trovare in media é 6 oppure 7.

Noi abbiamo $P = 11$. Calcoliamo:

$$\mathbb{P}\{P \geq 11\} = \binom{13}{11} \frac{1^{13}}{2} + \binom{13}{12} \frac{1^{13}}{2} + \binom{13}{13} \frac{1^{13}}{2} = 0.0112$$

Poiché abbiamo misurato un $p\text{-value} = 0.0112$ allora possiamo concludere che il consumo a riposo di energia é maggiore nei soggetti affetti dalla patologia rispetto ai sani.