Test d'Ipotesi – Esercizi

Esercizio 1. Un ricercatore intende saggiare, con livello di significatività del 5%, l'affermazione di una ditta farmaceutica secondo la quale il tempo che intercorre tra l'assunzione di un farmaco e la manifestazione dei primi effetti è al più di 4 minuti.

A questo scopo considera un campione casuale di 100 pazienti e trova che in media il tempo necessario per riscontrare efficacia nel farmaco è di 4 minuti e 6 secondi, con scarto quadratico medio $s = 0.6$ minuti.

Quali sono le conclusioni del test? Cambia qualcosa se i dati sperimentali hanno scarto quadratico medio $s = 0.64$ minuti?

Test d'Ipotesi a Una Coda

- Se si vuole un livello di significatività del 5%, si avranno elementi per rigettare l'ipotesi quando $\overline{x}$ si lascia *a sinistra* almeno il 95% dei dati. Questo accade quando $\overline{x}$ non appartiene all'intervallo $(-\infty\,,\,\mu+1.64\,\frac{s}{\sqrt{n}}]$, entro il quale cade il 95% dei dati.
- Se si vuole un livello di significatività dell'1%, l'ipotesi diventa rigettabile (con un margine di rischio dell'1%) quando $\overline{x}$ non appartiene all'intervallo $(-\infty\,,\,\mu+2.33\,\frac{s}{\sqrt{n}}]$, entro il quale cade il 99% dei dati.

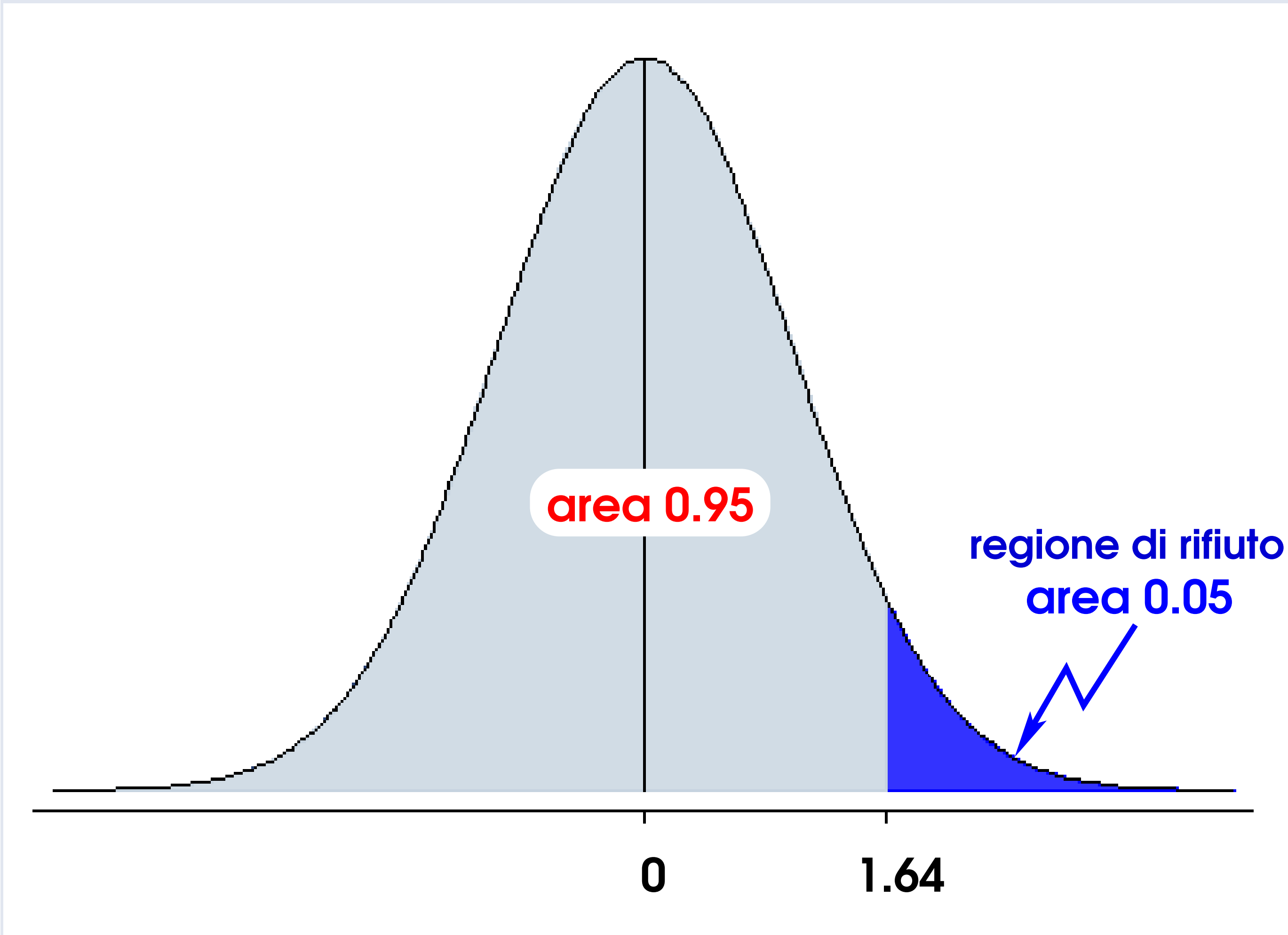
n è la numerosità del campione, s è la deviazione standard campionaria.

Curva Gaussiana

aree sottese dalla curva gaussiana
sull' intervallo [μ , μ + z σ]

z	0.00	0.01	0.02	0.03	0.04	0.05
0.00	0.0000	0.0040	0.0080	0.0120	0.0160	0.0199
0.10	0.0398	0.0438	0.0478	0.0517	0.0557	0.0596
0.20	0.0793	0.0832	0.0871	0.0910	0.0948	0.0987
0.30	0.1179	0.1217	0.1255	0.1293	0.1331	0.1368
0.40	0.1554	0.1591	0.1628	0.1664	0.1700	0.1736
0.50	0.1915	0.1950	0.1985	0.2019	0.2054	0.2088
0.60	0.2257	0.2291	0.2324	0.2357	0.2389	0.2422
0.70	0.2580	0.2611	0.2642	0.2673	0.2704	0.2734
0.80	0.2881	0.2910	0.2939	0.2967	0.2995	0.3023
0.90	0.3159	0.3186	0.3212	0.3238	0.3264	0.3289
1.00	0.3413	0.3438	0.3461	0.3485	0.3508	0.3531
1.10	0.3643	0.3665	0.3686	0.3708	0.3729	0.3749
1.20	0.3849	0.3869	0.3888	0.3907	0.3925	0.3944
1.30	0.4032	0.4049	0.4066	0.4082	0.4099	0.4115
1.40	0.4192	0.4207	0.4222	0.4236	0.4251	0.4265
1.50	0.4332	0.4345	0.4357	0.4370	0.4382	0.4394
1.60	0.4452	0.4463	0.4474	0.4484	0.4495	0.4505
1.70	0.4554	0.4564	0.4573	0.4582	0.4591	0.4599
1.80	0.4641	0.4649	0.4656	0.4664	0.4671	0.4678
1.90	0.4713	0.4719	0.4726	0.4732	0.4738	0.4744
2.00	0.4772	0.4778	0.4783	0.4788	0.4793	0.4798
2.10	0.4821	0.4826	0.4830	0.4834	0.4838	0.4842
2.20	0.4861	0.4864	0.4868	0.4871	0.4875	0.4878
2.30	0.4893	0.4896	0.4898	0.4901	0.4904	0.4906
2.40	0.4918	0.4920	0.4922	0.4925	0.4927	0.4929
2.50	0.4938	0.4940	0.4941	0.4943	0.4945	0.4946
2.60	0.4953	0.4955	0.4956	0.4957	0.4959	0.4960
2.70	0.4965	0.4966	0.4967	0.4968	0.4969	0.4970
2.80	0.4974	0.4975	0.4976	0.4977	0.4977	0.4978
2.90	0.4981	0.4982	0.4982	0.4983	0.4984	0.4984
3.00	0.4987	0.4987	0.4987	0.4988	0.4988	0.4989

Test a Una Coda



Test d'Ipotesi – Esercizi

Convertiamo il tempo medio campionario di reazione del farmaco in forma decimale: $\overline{x} = 4$ minuti e 6 secondi $= 4.1$ minuti.

Ipotesi zero: $\mu \leq 4$ minuti. Si tratta di un *test a una coda*.

Avremo ragione di dubitare dell'affermazione della ditta solo se si misurano, come in questo caso, tempi di reazione maggiori al limite superiore dichiarato dal produttore.

Calcoliamo il valore di u che risolve l'equazione:

$$\overline{x} - \mu = 0.1 = u \frac{s}{\sqrt{n}} = 0.06\, u \quad \Rightarrow \quad u = 1.67$$

Dunque, $\overline{x}$ si trova fuori dall'intervallo $\left(-\infty\,, \mu + 1.64 \frac{s}{\sqrt{n}}\right]$, entro il quale cade il 95% dei dati.

Al livello di significatività del 5%, l'ipotesi zero è da respingere.

Test d'Ipotesi – Esercizi

Se $s = 0.64$,

$$\overline{x} - \mu = 0.1 = u\,\frac{s}{\sqrt{n}} = 0.064\,u \qquad \Rightarrow \qquad u = 1.56$$

Dunque, $\overline{x}$ cade nell'intervallo $\left(-\infty\,,\ \mu + 1.64\,\frac{s}{\sqrt{n}}\right]$.

Al livello di significatività richiesto l'ipotesi non è rigettabile.

L'esempio evidenzia l'importanza dell'accuratezza della verifica sperimentale. Lo scarto quadratico medio misura infatti la dispersione dei dati attorno al valor medio e indica il grado di imprecisione nelle misure. Le conclusioni tratte dall'esame dei dati sono difformi nei due casi.

Esercizio 2. Si vuole sottoporre a verifica la seguente affermazione: la spesa media per le vacanze degli italiani è inferiore a 800 Euro a testa. A questo scopo si considera un campione di 100 italiani e si osserva che la spesa media per le vacanze di questo campione è stata di 808 Euro a testa con uno scarto quadratico medio $s = 40$ Euro.

Dopo aver precisato se il test debba essere a una o due code, trarre le conclusioni se il livello di significatività è del 5%. Cosa cambia se il livello di significatività del test è dell'1%?

Esercizio 3. Si vuole sottoporre a verifica la seguente affermazione: il peso medio degli abitanti adulti di una certa nazione è $\mu = 72\,\mathrm{Kg}$. A questo scopo si considera un campione casuale di 100 individui, che vengono pesati. Si ottiene un peso medio campionario $\bar{x} = 73.8\,\mathrm{Kg}$ con deviazione standard campionaria $s = 8\,\mathrm{Kg}$.

Dopo aver precisato se il test debba essere a una o due code, trarre le conclusioni se il livello di significatività è del 5%. Cosa cambia se il livello di significatività del test è dell'1%? E se il campione fosse stato composto da 400 individui?