

Esercitazione 6 del corso di Statistica 2

Dott.ssa Paola Costantini

7 marzo 2012

Decisione	H_0 vera	H_0 falsa
H_0 è respinta	Errore di I tipo Probabilità = α	Decisione corretta Probabilità = $1 - \beta$
H_0 non è respinta	Decisione corretta Probabilità = $1 - \alpha$	Errore di II tipo Probabilità = β

Errore di I tipo: indica la probabilità di rifiutare l'ipotesi nulla quando è vera. E' indicata con α ed è anche denominata livello di significatività del test.

Errore di II tipo: indica la probabilità di accettare l'ipotesi nulla quando è falsa. E' indicata con β .

Esercizio n 1

Il consumo settimanale di gas, in decimetri cubi, per il riscaldamento di un'abitazione ha una distribuzione normale con scarto quadratico medio 35. In precedenza è stato osservato un consumo settimanale medio di gas pari a 140 decimetri cubi. Avendo coibentato le mura dell'appartamento, si ritiene che il consumo settimanale medio di gas sia, ora, 115 decimetri cubi. Per confrontare l'ipotesi nulla, che il consumo medio di gas sia rimasto invariato, con l'ipotesi alternativa, in base alla quale il consumo medio di gas è ridotto, si utilizza la seguente regola di decisione: l'ipotesi nulla è respinta se in un campione di 25 osservazioni la media campionaria è minore di 125 decimetri cubi.

- Calcolare la probabilità dell'errore di I tipo;
- Calcolare la probabilità dell'errore di II tipo;
- Calcolare quale sarebbe la probabilità dell'errore di I tipo se il valore critico fosse 128 decimetri cubi.
- Calcolare quale sarebbe la probabilità dell'errore di II tipo se il valore critico fosse 128 decimetri cubi.

Soluzione

a) La probabilità dell'errore di I tipo è la probabilità che \bar{X} assuma valori nella regione critica quando H_0 è vera ed è quindi la probabilità dell'evento $\bar{X} < 125$ calcolata sull'ipotesi nulla $\mu = 140$. Si ha:

$$\alpha = P(\bar{X} < 125 | H_0 : \mu = 140) = P\left(\frac{\bar{X} - 140}{35/\sqrt{25}} < \frac{125 - 140}{7} | H_0 : \mu = 140\right)$$

$$= P(Z < -2,14) = 1 - z(2,14) = 1 - 0,9838 = 0,0162$$

b) Sotto l'ipotesi alternativa la media campionaria ha una distribuzione Normale con media 115 decimetri cubi e scarto quadratico medio 35. La probabilità di commettere un errore di II tipo è la probabilità che media campionaria assuma valore nella regione di accettazione, $\bar{X} \geq 125$, quando è vera H_1 , cioè $\mu = 115$. Si ha:

$$\beta = P(\bar{X} \geq 125 | H_1 : \mu = 115) = P\left(\frac{\bar{X} - 115}{35/\sqrt{25}} \geq \frac{125 - 115}{7} | H_1 : \mu = 115\right)$$

$$= P(Z \geq 1,43) = 1 - z(1,43) = 1 - 0,9236 = 0,0764$$

$$c) \alpha = P(\bar{X} < 128 | H_0 : \mu = 140) = P\left(\frac{\bar{X} - 140}{35/\sqrt{25}} < \frac{128 - 140}{7} | H_0 : \mu = 140\right)$$

$$= P(Z < -1,71) = 1 - z(1,71) = 1 - 0,9838 = 0,0436$$

$$d) \beta = P(\bar{X} \geq 128 | H_1 : \mu = 115) = P\left(\frac{\bar{X} - 115}{35/\sqrt{25}} \geq \frac{128 - 115}{7} | H_1 : \mu = 115\right)$$

$$= P(Z \geq 1,87) = 1 - z(1,86) = 1 - 0,9686 = 0,0314$$

Con l'aumento della regione critica (da 125 a 128), la probabilità di commettere l'errore di II tipo diminuisce; ciò comporterà un aumento della probabilità di commettere un errore di I tipo.

Esercizio n. 2

In passato la lunghezza media delle pannocchie di grano è stata uguale a 27 cm con $\sigma^2 = 24$. Si vuole sottoporre a test l'ipotesi che le pannocchie di un determinato anno abbiano una lunghezza media diversa, sulla base di un campione di 20 elementi con un $\alpha = 0,04$.

30	29	16	19	25	23	18	17	29	30
29	30	23	27	22	16	28	24	26	30

$$\bar{X} = \frac{\sum_{i=1}^n x_i}{n} = 24,55$$

IPOTESI	$H_0 : \mu = 27$ $H_1 : \mu \neq 27$
STATISTICA TEST	$X_{\text{test}} = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}}$

VALORI CRITICI	$\pm Z_{\alpha/2} \rightarrow \pm Z_{0,02} = \pm 2,05375$
REGOLA DI DECISIONE	$-2,05375 \leq v_{test} \leq +2,05375$
VALORE TEST	$V_{test} = \frac{24,55 - 27}{4,9\sqrt{20}} = -2,24$
DECISIONE	$-2,24 < -2,05375$ si rifiuta l'ipotesi H_0

Esercizio n 3

Una ditta produttrice di batterie per cellulari pubblicizza i propri prodotti garantendo una durata media di 18 ore con uno scarto quadratico medio di 0,5. Poiché ha ricevuto parecchi reclami da parte dei clienti che sostengono che la durata è inferiore, la ditta decide di effettuare una prova di durata su un campione di 10 batterie, ottenendo un tempo medio di accensione di 17,7 ore.

- a) sulla base di tale risultato come può la ditta verificare la validità della sua affermazione riguardante la durata media garantita.

Svolgimento

Si richiede di effettuare un test sulla media conoscendo la varianza del carattere osservato, con un livello di significatività del 5%.

IPOTESI	$H_0 : \mu = 18$ $H_1 : \mu < 18$
STATISTICA TEST	$X_{test} = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}}$
VALORI CRITICI	$z = -2,33$
REGOLA DI DECISIONE	$-1 \leq v_{test} \leq +1$
VALORE TEST	$V_{test} = \frac{17,7 - 18}{0,5\sqrt{10}} = -1,896$
DECISIONE	$-2,33 < -1,645$ si rifiuta l'ipotesi H_0

Esercizio n 3

Sia dato un campione di famiglie secondo l'indice di affollamento (n. medio di componenti per stanza):

Indice di affollamento	n. famiglie	c_i	$c_i n_i$	$(c_i - \bar{x})^2 n_i$
[0-0,6]	6	0,3	1,8	5,48
]0,6-1,2]	18	0,9	16,2	2,28
]1,2-2]	10	1,6	16	1,18
]2-3]	7	2,5	17,5	10,83
	41		51,5	19,77

Provare l'ipotesi che nella popolazione l'indice di affollamento sia pari a 1.

IPOTESI	$H_0 = \mu = 1$ $H_1 = \mu \neq 1$
STATISTICA TEST	$t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}$
VALORI CRITICI	$t = \frac{1,256 - 1}{0,7/\sqrt{41}} = \frac{0,256}{0,109} = 2,35$
REGOLA DI DECISIONE	Accetto H_0 se $V_{test} > -1,645$
VALORE TEST	$V_{test} = \frac{17,7 - 18}{0,5\sqrt{10}} = -1,896$
DECISIONE	$-2,33 < -1,645$ si rifiuta l'ipotesi H_0

Esercizio 4

Un economista del Ministero degli Esteri desidera verificare se gli accordi di negoziazione tra Italia e Giappone siano rispettati. In particolare egli sospetta che i produttori giapponesi fissino un prezzo più basso per i prodotti venduti sul mercato italiano rispetto a quello usato sul mercato interno, ostacolando al contempo le importazioni di prodotti italiani con forti ostacoli di tipo burocratico. Si interessa in particolare al mercato dell'auto e vuole testare l'ipotesi che prezzi più alti siano applicati in Giappone rispetto all'Italia per le autovetture di produzione giapponese. Esamina a tal fine due campioni relativi a pratiche di acquisto di tali autovetture nello stesso periodo di tempo (50 per il mercato italiano e 30 per il mercato giapponese). Convertendo i prezzi di vendita in Giappone usando il cambio corrente Yen/Euro, ottiene i risultati elencati nella seguente tabella:

	ITALIA	GIAPPONE
AMPIEZZA CAMPIONE	50	30
MEDIA CAMPIONARIA	€6545	€7243

Siano inoltre noti i seguenti valori per le rispettive popolazioni di riferimento

	ITALIA	GIAPPONE
DEVIAZIONE STANDARD	€989	€843

- a) Costruire un test d'ipotesi usando un livello di significatività $\alpha=0,05$

I due campioni sono selezionati in maniera indipendente dalle due popolazioni. Le rispettive ampiezze campionarie n_1 e n_2 sono sufficientemente grandi affinché sia \bar{x}_1 che \bar{x}_2 siano distribuite approssimativamente come normali.

SVOLGIMENTO

IPOTESI	$H_0 : \mu_1 - \mu_2 = 0$ $H_1 : \mu_1 - \mu_2 < 0$
STATISTICA TEST	$\frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sigma_{(\bar{x}_1 - \bar{x}_2)}} = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \approx Z$
VALORI CRITICI	$\alpha = 0,05 \rightarrow Z_\alpha = -1,645$
REGOLA DI DECISIONE	Accetto H_0 se $V_{\text{test}} > -1,645$
VALORE TEST	$= \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} = \frac{(16545 - 17243)}{\sqrt{\frac{1989^2}{50} + \frac{1843^2}{30}}} = -1,59$
DECISIONE	$-1,59 < -1,645$ si accetta l'ipotesi H_0

Esercizio 5

Al fine di verificare le spese trimestrali per la tenuta di c/c bancari, la Banca d'Italia esamina le spese di tenuta conto praticate da 8 banche italiane riscontrando una spesa trimestrale media di 32 Euro con una varianza campionaria pari a 1.8.

Allo stesso tempo, procede ad effettuare lo stesso tipo di rilevazione presso 10 banche straniere operanti in Italia, rilevando una spesa trimestrale media di 32.9 Euro con una varianza campionaria pari a 1.7.

Si vuole verificare se le spese trimestrali medie di tenuta conto sono le stesse presso i due tipi di banche.

SVOLGIMENTO

Si richiede di effettuare un test sulla differenza tra medie non conoscendo le varianze del carattere osservato nelle due popolazioni. In tal caso, per poter effettuare il test, è necessario ipotizzare che le due varianze siano uguali ($\sigma_1^2 = \sigma_2^2 = \sigma^2$).

Le ipotesi da sottoporre a verifica sono le seguenti:

$$H_0 : \mu_1 - \mu_2 = 0$$

$$H_1 : \mu_1 - \mu_2 \neq 0$$

Per poter effettuare il test bisogna anche ipotizzare che il carattere osservato abbia una distribuzione normale presso le due popolazioni.

La statistica test da considerare è la seguente:

$$T = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{n_1 s_1^2 + n_2 s_2^2}{n_1 + n_2 - 2} \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}} \approx t_{(n_1 + n_2 - 2)}$$

che ha una distribuzione t di Student con $n_1 + n_2 - 2$.

Fissato un livello di significatività $\alpha = 0.05$ è possibile determinare le regioni di accettazione e di rifiuto:

$$\text{Regione di accettazione: } \{-t_{\alpha/2} < t_i < t_{1-\alpha/2}\}; \quad \text{Regione di rifiuto: } \{t_i < -t_{\alpha/2}\} \cup \{t_i > t_{1-\alpha/2}\}.$$

Dai dati a disposizione risulta:

$$n_1 = 8; \quad n_2 = 10; \quad \bar{x}_1 = 32; \quad \bar{x}_2 = 32.9; \quad s_1^2 = 1.8 \quad s_2^2 = 1.7.$$

Il test è a due code; i valori critici della distribuzione t di Studente per questo livello di significatività sono $t_{0.025} = -2.12$ e $t_{0.975} = 2.12$.

La regola di decisione consiste nel rifiutare l'ipotesi nulla se il valore campionario della statistica test è inferiore a -2.12 o superiore a 2.12.

$$\text{Il valore campionario della statistica test è: } t = \frac{32 - 32.9}{\sqrt{\frac{8(1.8) + 10(1.7)}{8 + 10 - 2} \left(\frac{1}{8} + \frac{1}{10} \right)}} = -1.35$$

Al livello di significatività del 5% l'ipotesi nulla viene accettata, concludendo che le spese trimestrali medie di tenuta conto sono le stesse presso i due tipi di banche.

Esercizio 5

Utilizzo della laurea \ Tipo di contratto	Stabile	Atipico	Inserimento/ formazione	Senza contratto	Tot
In misura elevata	5	3	1	0	9
In misura ridotta	2	5	1	0	8
Per niente	2	4	0	2	8
totale	9	12	2	2	25

Verificare ad un livello di significatività del 90% se i due caratteri sono indipendenti

IPOTESI	$H_0 = \sum_i \sum_j \frac{(n_{ij} - \hat{n}_{ij})^2}{\hat{n}_{ij}} = 0$
	$H_1 = \sum_i \sum_j \frac{(n_{ij} - \hat{n}_{ij})^2}{\hat{n}_{ij}} > 0$
STATISTICA TEST	$\sum_i \sum_j \frac{(n_{ij} - \hat{n}_{ij})^2}{\hat{n}_{ij}}$
VALORI CRITICI	$\chi^2_{0,1,6} \quad 6=(r-1)*(c-1) = 10,645$
REGOLA DI DECISIONE	Accetto H_0 se $V_{\text{test}} < 10,645$
VALORE TEST	$\chi^2 = = \mathbf{7,4}$
DECISIONE	7,4 < 10,645 si accetta l'ipotesi H_0

Accetto l'ipotesi nulla, quindi i due caratteri sono indipendenti. Il valore test è diverso da zero, ma non significativamente diverso da zero.

.