Esercitazione 8 del corso di Statistica 2

Prof. Domenico Vistocco

Dott.ssa Paola Costantini

16 Giugno 2008

Decisione	H_0 vera	H_0 falsa
H_0 è respinta	**Errore di I tipo** Probabilità = α	Decisione corretta Probabilità = 1 - β
H_0 non è respinta	Decisione corretta Probabilità = 1 - α	**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 par a 140 decimetri cubi. Avendo coibendato 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.

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

d) 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 $\overline{X}$ assuma valori nella regione critica quando H_0 è vera ed è quindi la probabilità dell'evento $\overline{X} < 115$ calcolata sull'ipotesi nulla $\mu = 140$. Si ha:

$$\alpha = P(\overline{X} \in R.C.|H_0 : \mu = 140) = P(\overline{X} < 125|H_0 : \mu = 140) = P\left(\frac{\overline{X} - 140}{35/\sqrt{25}} < \frac{125 - 140}{7}\middle|H_0 : \mu = 140\right)$$

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

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, $\overline{X} \geq 125$, quando è vera H_1, cioè $\mu = 115$. Si ha:

$$\beta = P(\overline{X} \in R.A.|H_1 : \mu = 115) = P(\overline{X} \geq 125|H_1 : \mu = 115) = P\left(\frac{\overline{X} - 115}{35/\sqrt{25}} \geq \frac{125 - 115}{7}\middle|H_1 : \mu = 115\right)$$

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

c) $\alpha = P(\overline{X} \in R.C.|H_0 : \mu = 140) = P(\overline{X} < 128|H_0 : \mu = 140) = P\left(\frac{\overline{X} - 140}{35/\sqrt{25}} < \frac{128 - 140}{7}\middle|H_0 : \mu = 140\right)$

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

d) $\beta = P(\overline{X} \in R.A.|H_1 : \mu = 115) = P(\overline{X} \geq 128|H_1 : \mu = 115) = P\left(\frac{\overline{X} - 115}{35/\sqrt{25}} \geq \frac{128 - 115}{7}\middle|H_1 : \mu = 115\right)$

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

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

Un'azienda produttrice di distributori automatici di bibite intende installare un distributore in un ufficio pubblico. Gli addetti all'ufficio sostengono che in media, col vecchio distributore, si vendevano 35 bibite al giorno, mentre l'azienda addetta all'installazione del distributore ha valutato una vendita media di 25 bibite al giorno. Per decidere sull'installazione del distributore, l'azienda lo mette in prova per 16 giorni. Determinare il valore:

a) il valore critico $\overline{x}^*$ legato alla regola di decisione assunta, posto $\alpha = 0{,}05$;
b) β, facendo riferimento al valore critico calcolato in a).

Soluzione

La regola di decisione è:

$\overline{x} > \overline{x}^* \rightarrow$ non si rifiuta H_0

$\overline{x} \leq \overline{x}^* \rightarrow$ si rifiuta H_0

Dalla probabilità α data, è possibile desumere il valore $\overline{x}^*$:

$$0{,}05 = P(\bar{x} \le \bar{x}^* / \mu_x = 35) = P\left(z \le \frac{\bar{x} - 35}{10\sqrt{16}} \right)$$

Essendo α=0,05 deve essere P(z < z^*) =0,05, da cui, riguardando la coda sinistra della variabile casuale, z^*= -1,645

Pertanto si giunge all'uguaglianza

$$\frac{\bar{x}^* - 35}{10/4} = -1{,}645$$, da cui

$$\bar{x}^* = -1{,}645 \cdot \left(\frac{10}{4} \right) + 35 = 30{,}8875$$

b) La probabilità dell'errore di II tipo è

$$\beta = P(\bar{X} \in R.A. | H_1 : \mu = 25) = P(\bar{X} \ge 30{,}8875 | H_1 : \mu = 25) = P\left(\frac{\bar{X} - 25}{10/\sqrt{16}} \ge \frac{30{,}8875 - 25}{10/4} \middle| H_1 : \mu = 25 \right)$$

$$= P(Z \ge 2{,}35) = 1 - z(2{,}35) = 1 - 0{,}9906 = 0{,}0094$$

Esercizio n 3 Test unidirezionale sulla proporzione

Un dirigente di un'azienda che eroga servizi di pubblica utilità, con uffici dislocati su tutto il territorio nazionale, afferma che almeno il 90% degli uffici aperti al pubblico sono attrezzati per l'accesso delle persone disabili. In un campione di 230 uffici, 202 risultano attrezzati per l'accesso delle persone disabili. Verificare, al livello di significatività del 5%, l'affermazione del dirigente.

Soluzione

Sia $X \sim Ber(\pi)$ la v.c. che descrive la popolazione; X assume valore 1 se gli uffici sono attrezzati per le persone disabili. Le ipotesi da sottoporre a test sono:

$H_0 : \pi \ge 0{,}9$ vs $H_1 : \pi < 0{,}9$

In corrispondenza del livello di significatività α = 0,05 si ha $z_{0,05} =$ 1,645 in quanto ϕ(0,05) = 1-0,05 =0,95, che sulle tavole è uguale a 1,645.

$$R.A.: \frac{\hat{p} - \pi_0}{\sqrt{\pi_0(1 - \pi_0)/n}} > -1.645$$

$$R.C.: \frac{\hat{p} - \pi_0}{\sqrt{\pi_0(1 - \pi_0)/n}} < -1.645$$

Dato che la proporzione campionaria è $\hat{p} = \dfrac{202}{230} = 0{,}88$, il valore osservato della statistica test risulta:

$$\frac{\hat{p} - \pi_0}{\sqrt{\pi_0(1 - \pi_0)/n}} = \frac{0{,}88 - 0{,}90}{\sqrt{0{,}90 * 0{,}10/230}} = \frac{-0{,}02}{0{,}01978} = -1{,}011$$

Il valore -1,011 si trova nella regione di accettazione; quindi, al livello di significatività del 5% non vi è motivo di dubitare che il dirigente affermi il vero.

Esercizio n 4 Test per la varianza

Un'azienda produttrice di tubi di metallo riscontra una varianza dei diametri dei tubi prodotti pari a 96 mmq. A seguito dell'introduzione di un nuovo processo, si estrae un campione di 30 tubi per verificare se c'è stato un decremento significativo nel diametro e si riscontra che la varianza campionaria corretta è 85mmq.
Decidere, ad un livello di significatività $\alpha = 0{,}05$, se il decremento nella varianza dei diametri è significativo, supponendo che la distribuzione dei diametri sia normale.

Soluzione

Le ipotesi a confronto sono:

$H_0 : \sigma^2 = 96\,\text{mmq}$

$H_0 : \sigma^2 < 96\,\text{mmq}$

Il test da utilizzare si distribuisce come una v.c χ^2 con g = 30 -1 =29 gradi di libertà. Il valore della statistica test è:

$$t_n = \frac{(n-1)\hat{\sigma}^2}{\sigma_0^2} \sim \chi^2_{n-1,\alpha} = \frac{(30-1)85}{96} = 25{,}677$$

Per $\alpha = 0{,}05$ e 29 gradi di libertà, dalla tavola risulta $\chi^2_{0,950;29} = 17{,}71$

Essendo $t_n > \chi^2_{0,950;29}$ si rifiuta H_0, e si conclude che il decremento nella vacanza è significativo.

Esercizio n 5 (test per la μ con σ^2 non nota)

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.

$$H_0 = \mu = 1$$
$$H_1 = \mu \neq 1$$

Per il test si sceglie un livello di significatività del 10%.

$\bar{x} =$ 51,5/41= 1,256

Se σ^2 è sconosciuto, una sua stima corretta è $s^2 = \frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n-1} = \frac{19{,}77}{40} = 0{,}49; s = 0{,}7$

Allora, sotto H_0 la quantità $t = \frac{\bar{x} - \mu_0}{s / \sqrt{n}}$ è una determinazione di una v.c. t di Student con (n-1) gradi di libertà.

$$t = \frac{1{,}256 - 1}{0{,}7 / \sqrt{41}} = \frac{0{,}256}{0{,}109} = 2{,}35$$

Dato il livello di significatività α=0,10, per determinare la regola di decisione occorre individuare il valore critico $t_{40,0,05}$ tale che una v.c. t_{40} assuma valori maggiori con probabilità 0,05. Dalle tavole risulta $t_{40,0,05}$ =1,697.

Il valore della statistica test è risultato 2,25, pertanto H_0 è respinta cioè, concludiamo che, al livello di significatività del 10%, l'indice di affollamento non sia pari ad 1.