Esercitazione 5 del corso di Statistica 2

Dott.ssa Paola Costantini

18 Maggio 2012

Esercizio n 1

Una grande banca vuole stimare l'ammontare medio di denaro dovuto dai clienti in mora (vale a dire in ritardo nei pagamenti più di due mesi). La banca decide di selezionare casualmente 100 tra questi clienti, su cui si osserva un ammontare medio di € 2310. Osservazioni fatte negli anni precedenti fanno ritenere plausibile un valore $\sigma = 89$.
Si costruisca un intervallo di confidenza per l'ammontare medio di denaro dovuto dai clienti in mora.

1) Al 90%
2) Al 95%
3) Al 99%
4) Si desidera stimare il parametro μ (ammontare medio di denaro dovuto dai clienti in mora) accettando un errore di 5€. Determinare l'ampiezza campionaria a tal fine usando un livello di confidenza $\alpha = 0{,}10$.

SVOLGIMENTO

N=100
Per risolvere i tre quesiti può essere utile riassumere nella seguente tabella i coefficienti di confidenza con i rispettivi valori per la Z più comunemente utilizzati:

100(1-α)%	α	α/2	Zα/2
90%	0,10	0,050	1,645
95%	0,05	0,025	1,960
99%	0,01	0,005	2,576

$$P\left(-Z_{\alpha/2} \le \frac{\bar{x}-\mu_{\bar{x}}}{\sigma_{\bar{x}}} \le +Z_{\alpha/2} \right) = (1-\alpha)$$

$$P\left(\bar{x}-Z_{\alpha/2}\sigma_{\bar{x}} \le \frac{\bar{x}-\mu_{\bar{x}}}{\sigma_{\bar{x}}} \le \bar{x}+Z_{\alpha/2}\sigma_{\bar{x}} \right) = (1-\alpha)$$

$$P\left(\bar{x}-Z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \le \frac{\bar{x}-\mu_{\bar{x}}}{\sigma_{\bar{x}}} \le \bar{x}+Z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \right) = (1-\alpha)$$

Sostituendo i valori dell'esercizio si ha:
1) $\alpha = 0{,}10 \quad 100 \times (1-\alpha) = 0{,}90$

$$P\left(2310-1{,}645\frac{89}{\sqrt{100}}\leq \mu_{\bar{x}}\leq 2310+1{,}645\frac{89}{\sqrt{100}}\right)=0{,}90$$

$$P(2310-14{,}64\leq \mu_{\bar{x}}\leq 2310+14{,}64)=0{,}90$$

$$P(2295{,}36\leq \mu_{\bar{x}}\leq 2324{,}64)=0{,}90$$

2) $\alpha = 0{,}05 \quad 100\times(1-\alpha)=0{,}95$

$$P\left(2310-1{,}96\frac{89}{\sqrt{100}}\leq \mu_{\bar{x}}\leq 2310+1{,}96\frac{89}{\sqrt{100}}\right)=0{,}95$$

$$P(2310-17{,}44\leq \mu_{\bar{x}}\leq 2310+17{,}44)=0{,}95$$

$$P(2292{,}56\leq \mu_{\bar{x}}\leq 2327{,}44)=0{,}95$$

3) $\alpha = 0{,}01 \quad 100\times(1-\alpha)=0{,}99$

$$P\left(2310-2{,}576\frac{89}{\sqrt{100}}\leq \mu_{\bar{x}}\leq 2310+2{,}576\frac{89}{\sqrt{100}}\right)=0{,}99$$

$$P(2310-22{,}92\leq \mu_{\bar{x}}\leq 2310+22{,}92)=0{,}99$$

$$P(2287{,}08\leq \mu_{\bar{x}}\leq 2332{,}92)=0{,}99$$

4) $P\left(\bar{x}-Z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\leq \frac{\bar{x}-\mu_{\bar{x}}}{\sigma_{\bar{x}}}\leq \bar{x}+Z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right)=(1-\alpha)$

In questo caso si vuole calcolare la numerosità campionaria in grado di assicurare un livello di precisione nella stima fissato. Risolvendo rispetto a n si ha:

$$Z_{\alpha/2}\frac{\sigma}{\sqrt{n}}=\varepsilon \rightarrow Z_{\alpha/2}\sigma=\varepsilon\sqrt{n}\rightarrow \sqrt{n}=\frac{Z_{\alpha/2}\sigma}{\varepsilon}\rightarrow n=\frac{Z^2_{\alpha/2}\sigma^2}{\varepsilon^2}$$

Sostituendo i valori, ricordando che per un intervallo al)0% la Z vale 1,96, si ha:

$$n=\frac{Z^2_{\alpha/2}\sigma^2}{\varepsilon^2}=\frac{1{,}96^2\times 89^2}{5^2}=858$$

Esercizio 2

Si ripeta l'esercizio 1 nell'ipotesi di non avere informazioni plausibili sulla varianza campionaria della popolazione. Si usi a tal fine il valore di s=89 misurato sul campione di 100 clienti considerato.

SVOLGIMENTO

N = 100 $\bar{x} = 2310$ s = 89

In questo caso σ non è noto e invece di usare la distribuzione Z si dovrebbe usare la distribuzione t. Il campione in questione è però di numerosità elevata e quindi non ci sono sensibili variazioni, come si può notare dalle seguenti tabelle che riportano i valoridella distribuzione Z e della distribuzione t per i valori di α più comunemente utilizzati:

$$P\left(-t_{n-1},_{\alpha/2} \le \frac{\bar{x}-\mu_{\bar{x}}}{s_{\bar{x}}} \le +t_{n-1},_{\alpha/2}\right) = (1-\alpha)$$

$$P\left(\bar{x}-t_{n-1},_{\alpha/2}\, s_{\bar{x}} \le \frac{\bar{x}-\mu_{\bar{x}}}{\sigma_{\bar{x}}} \le \bar{x}+t_{n-1},_{\alpha/2}\, s_{\bar{x}}\right) = (1-\alpha)$$

$$P\left(\bar{x}-t_{n-1},_{\alpha/2} \frac{s}{\sqrt{n}} \le \frac{\bar{x}-\mu_{\bar{x}}}{\sigma_{\bar{x}}} \le \bar{x}+t_{n-1,\alpha/2} \frac{s}{\sqrt{n}}\right) = (1-\alpha)$$

Sostituendo i valori dell'esercizio si ha:

1) α = 0,10 $100 \times (1-\alpha) = 0{,}90$

$$P\left(2310 - 1{,}660\frac{89}{\sqrt{100}} \le \mu_{\bar{x}} \le 2310 + 1{,}660\frac{89}{\sqrt{100}}\right) = 0{,}90$$

$$P(2310 - 14{,}78 \le \mu_{\bar{x}} \le 2310 + 14{,}78) = 0{,}90$$

$$P(2295{,}22 \le \mu_{\bar{x}} \le 2324{,}78) = 0{,}90$$

2) α = 0,05 $100 \times (1-\alpha) = 0{,}95$

$$P\left(2310 - 1{,}984\frac{89}{\sqrt{100}} \le \mu_{\bar{x}} \le 2310 + 1{,}984\frac{89}{\sqrt{100}}\right) = 0{,}95$$

$$P(2310 - 17{,}66 \le \mu_{\bar{x}} \le 2310 + 17{,}66) = 0{,}95$$

$$P(2292{,}34 \le \mu_{\bar{x}} \le 2327{,}66) = 0{,}95$$

3) $\alpha = 0{,}01 \quad 100 \times (1-\alpha) = 0{,}99$

$$P\left(2310 - 2{,}626\frac{89}{\sqrt{100}} \leq \mu_{\bar{x}} \leq 2310 + 2{,}626\frac{89}{\sqrt{100}}\right) = 0{,}99$$

$$P(2310 - 23{,}37 \leq \mu_{\bar{x}} \leq 2310 + 23{,}37) = 0{,}99$$

$$P(2286{,}63 \leq \mu_{\bar{x}} \leq 2333{,}37) = 0{,}99$$

4) $$P\left(\bar{x} - t_{n-1,\alpha/2}\frac{s}{\sqrt{n}} \leq \mu_{\bar{x}} \leq \bar{x} + t_{n-1\alpha/2}\frac{s}{\sqrt{n}}\right) = (1-\alpha)$$

In questo caso si vuole calcolare la numerosità campionaria in grado di assicurare un livello di precisione nella stima fissato. Risolvendo rispetto a n si ha:

$$t_{n-1,\alpha/2}\frac{s}{\sqrt{n}} = \varepsilon \rightarrow t_{n-1,\alpha/2}\, s = \varepsilon\sqrt{n} \rightarrow \sqrt{n} = \frac{t_{n-1,\alpha/2}\, s}{\varepsilon} \rightarrow n = \frac{t^2_{n-1\alpha/2} s^2}{\varepsilon^2}$$

Sostituendo i valori, ricordando che per un intervallo al 90% la t vale 1,984, si ha:

$$n = \frac{t^2_{\alpha/2} s^2}{\varepsilon^2} = \frac{1{,}984^2 \times 89^2}{5^2} = 874$$

Esercizio n 3

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.

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

Soluzione

La regola di decisione è:

$$\bar{x} > \bar{x}^* \rightarrow \text{non si rifiuta } H_0$$
$$\bar{x} \leq \bar{x}^* \rightarrow \text{si rifiuta } H_0$$

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

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

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

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

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

b) La probabilità dell'errore di II tipo è

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

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

ice di affollamento non sia pari ad 1.