Esercitazione 5 del corso di Statistica 2
Prof. Domenico Vistocco

Alfonso Iodice D'Enza

May 30, 2007

1 Esercizio

Si consideri una popolazione caratterizzata dai numeri $2, 3, 6, 8, 11$. Si considerino tutti i possibili campioni di ampiezza $N = 2$, estratti senza ripetizione.

- Calcolare media e scarto quadratico medio della popolazione.
- Calcolare media della distribuzione della media campionaria.
- Calcolare l'*errore standard* della media campionaria.

1.1 Svolgimento

Bisogna in primo luogo elencare tutti i possibili campioni di ampiezza 2 che è possibile estrarre senza ripetizione da una popolazione di 5 elementi. Ricordando le regole del conteggio, il numero di possibili campioni estraibili è $5^2 = 25$. In particolare

(2,2)	(2,3)	(2,6)	(2,8)	(2,11)
(3,2)	(3,3)	(3,6)	(3,6)	(3,11)
(6,2)	(6,3)	(6,6)	(6,6)	(6,11)
(8,2)	(8,3)	(8,6)	(8,6)	(8,11)
(11,2)	(11,3)	(11,6)	(11,6)	(11,11)

2	2.5	4	5	6.5
2.5	3	4.5	5.5	7
4	4.5	6	7	8.5
5	5.5	7	8	9.5
6.5	7	8.5	9.5	11

Calcolando la media per ciascun campione si ottiene la distribuzione della media campionaria.

- Calcolare media e scarto quadratico medio della popolazione.

Calcoliamo i parametri della popolazione:

$$\mu = \frac{2+3+6+8+11}{5} = \frac{30}{5} = 6$$

$$\sigma^2 = \frac{(2-6)^2 + (3-6)^2 + (6-6)^2 + (8-6)^2 + (11-6)^2}{5} = \frac{54}{5} = 10.8$$

- Calcolare media della distribuzione della media campionaria.

Per ottenere la media della distribuzione media campionaria si effettua la media delle medie campionarie calcolate in precenza

$$\mu_{\bar{X}} = \frac{2+2.5+4+\ldots+9.5+11}{25} = \frac{150}{25} = 6$$

Si nota dunque che la media della distribuzione della media campionaria coincide con la media della popolazione, ovvero $\mu = \mu_{\bar{X}} = 6$.

- Calcolare l'*errore standard* della media.

L'errore standard della media campionaria $\sigma_{\bar{X}}$ si ottiene analogamente

$$\sigma_{\bar{X}} = \sqrt{\frac{(2-6)^2 + (2.5-6)^2 + (4-6)^2 + \ldots + (9.5-6)^2 + (11-6)^2}{25}} =$$
$$= \sqrt{\frac{135}{25}} = \sqrt{5.4} = 2.32$$

Il che conferma la relazione tra la varianza della popolazione e la varianza della distribuzione della media campionaria $\sigma^2_{\bar{X}} = \frac{\sigma}{N}$: infatti, essendo $N = 2$, taglia del campione, risulta essere $\frac{\sigma}{N} = \frac{10.8}{2} = 5.4$, che coincide con la varianza della distribuzione della media campionaria.

2 Esercizio

Si consideri la popolazione composta dagli elementi $2, 3, 6, 8, 11$. Si considerino tutti i campioni di taglia $N = 2$ estratti *senza ripetizione*.

- Calcolare media della distribuzione della media campionaria.
- Calcolare l'*errore standard* della media campionaria.

2.1 Svolgimento

Il numero di possibili campioni che possono essere estratti è dato da $\binom{N_p}{N} = \binom{5}{2} = \frac{5!}{2!(5-2)!} = 10$. I parametri della popolazione sono $\mu = 6$ e $\sigma^2 = 10.8$. I possibili campioni sono

(2,3)	(2,6)	(2,8)	(2,11)	(3,6)
(3,8)	(3,11)	(6,8)	(6,11)	(8,11)

Mentre le medie campionarie sono

2.5	4	5	6.5	4.5
5.5	7	7	8.5	9.5

- Calcolare media della distribuzione della media campionaria.

Da cui la media della distribuzione della media campionaria

$$\mu_{\bar{X}} = \frac{2.5 + 4 + \ldots + 8.5 + 9.5}{10} = \frac{60}{10} = 6$$

Anche in questo caso la media della distribuzione delle medie campionarie e la media della popolazione coincidono.

- Calcolare l'*errore standard* della media campionaria.

Per calcolare l'errore standard

$$\sigma_{\bar{X}} = \sqrt{\frac{(2.5-6)^2 + (4-6)^2 + \ldots + (8.5-6)^2 + (9.5-6)^2}{10}} = \\ = \sqrt{4.05} = 2.01$$

E' possibile dunque verificare la relazione tra la varianza della distribuzione della media campionaria e la varianza della popolazione

$$\sigma^2_{\bar{X}} = \frac{\sigma^2}{N}\left(\frac{N_p - N}{N_p - 1}\right) = \frac{10.8}{2}\left(\frac{5-2}{5-1}\right) = 10.8 \times 0.375 = \mathbf{4.05}$$

3 Esercizio

Si consideri l'esperimento consistente nel lancio di un dado 120 volte. Trovare la probabilità che esca testa

- tra il 40% e il 60% dei lanci;
- più di dei $\frac{5}{8}$ del totale dei lanci.

3.1 Svolgimento

- tra il 40% e il 60% dei lanci;

È possibile determinare tale probabilità in due modi diversi. Il primo modo consiste nell'utilizzare l'approssimazione della binomiale alla normale. In particolare, la variabile casuale numero di teste in 120 lanci si distribuisce come una binomiale di parametri $(n = 120, p = \frac{1}{2})$.
Il numero di successi corrispondenti al 40% del totale di 120 prove è 48, mentre il 60% corrisponde a 72. Poichè la v.c. numero di teste è discreta, si è interessati alla probabilità che essa assuma valori compresi nell'interallo $[47.5, 72.5]$. Il valore atteso del numero di teste in 120 prove è $\mu = np = 120 \times \frac{1}{2} = 60$ mentre lo scarto quadratico medio è $\sigma = \sqrt{np(1-p)} = 120 \times \frac{1}{2}\frac{1}{2} = 5.48$. Standardizzando i valori

$$z_1 = \frac{47.5 - 60}{5.48} = -2.28 \text{ e } z_2 = \frac{72.5 - 60}{5.48} = 2.28$$

Dalla simmetria della normale, segue che $P(-2.28 <= X <= 2.28) = 2 \times P(0 <= X <= 2.28) = 2 \times 0.4887 = 0.9774$.

Il secondo metodo consiste nel ricorrere alla distribuzione delle proporzioni campionarie, la cui media è data da $\mu_p = p$ e $\sigma_p = \sqrt{\frac{p(1-p)}{N}}$. Rispetto ai dati del problema $\mu_p = p = 0.5$, mentre lo scarto quadratico medio è $\sigma_p = \sqrt{\frac{0.25}{120}} = 0.0456$.
Standardizzando il problema rispetto ai parametri della distribuzione campionaria

$$z_1 = \frac{0.4 - 0.5}{0.0456} = -2.19 \text{ e } z_2 = \frac{0.6 - 0.5}{0.0456} = 2.19$$

La probabilità corrispondente è $P(-2.19 <= X <= 2.19) = 2 \times P(0 <= X <= 2.19) = 2 \times 0.4857 = 0.9714$. Il risultato non coincide con quanto ottenuto utilizzando l'approssimazione della binomiale alla normale. Questo è dovuto al fatto che bisogna tenere conto che la proporzione è una variabile discreta. Per tenere conto di questo bisogna sottrarre $(1/2N) = 1/(2 \times 120) = 0.00417$ a 0.4 ed aggiungere la stessa quantità a 0.6. Pertanto gli estremi dell'intervallo di valori standardizzati è

$$z_1 = \frac{0.4 - 0.00417 - 0.5}{0.0456} = -2.28 \text{ e } z_2 = \frac{0.6 + 0.00417 - 0.5}{0.0456} = 2.28$$

Che porta alla coincidenza tra i risultati nei due metodi.

- più di dei $\frac{5}{8}$ del totale dei lanci.

La proporzione corrispondente a $\frac{5}{8} = 0.6250$. Standardizzando il problema tenendo conto del fatto che la proporzione è una variabile discreta si ottiene

$$z = \frac{0.6250 - 0.00417 - 0.5}{0.0456} = 2.65$$

La probabilità corrispondente è $P(z >= 2.65) = 1 - P(z < 2.65)$=1-0.996=0.004

4 Esercizio

Si considerino due imprese concorrenti che producono lampadine. La fabbrica A produce lampadine con una durata media di 1400 ore ed una deviazione standard di 200 ore. La fabbrica B produce lampadine con una vita media di 1200 ore e una deviazione standard di 100 ore. Si consideri di analizzare un lotto di 125 lampadine di tipo A ed uno di tipo B.

- Qual è la probabilità che le lampadine nel lotto A abbiano una durata media superiore di 160 ore di quelle nel lotto B?
- Qual è la probabilità che le lampadine nel lotto A abbiano una durata media superiore di 250 ore di quelle nel lotto B?

4.1 Svolgimento

Indicando con $\bar{X}_A$ e $\bar{X}_B$ le medie di durata nei campioni A e B si ha che

$$\mu_{\bar{X}_A-\bar{X}_B} = \mu_{\bar{X}_A} - \mu_{\bar{X}_B} = 1400 - 1200 = 200$$

$$\sigma_{\bar{X}_A-\bar{X}_B} = \sqrt{\frac{\sigma_{\bar{X}_A}}{N_A} + \frac{\sigma_{\bar{X}_B}}{N_B}} = \sqrt{\frac{(200)^2}{125} + \frac{(100)}{125}} = 20$$

La standardizzazione della variabile differenze medie è

$$z = \frac{(\bar{X}_A - \bar{X}_B) - \mu_{\bar{X}_A-\bar{X}_B}}{\sigma_{\bar{X}_A-\bar{X}_B}} = \frac{(\bar{X}_A - \bar{X}_B) - 200}{20}$$

distrubuita (quasi) normalmente.

- Qual è la probabilità che le lampadine nel lotto A abbiano una durata media superiore di 160 ore di quelle nel lotto B?

  $$z = \frac{(\bar{X}_A - \bar{X}_B) - \mu_{\bar{X}_A-\bar{X}_B}}{\sigma_{\bar{X}_A-\bar{X}_B}} = \frac{160 - 200}{20} = -2$$

  La probabilità corrispondente è $P(Z >= -2) = P(Z <= 2) = 0.9772$

- Qual è la probabilità che le lampadine nel lotto A abbiano una durata media superiore di 250 ore di quelle nel lotto B?

  $$z = \frac{(\bar{X}_A - \bar{X}_B) - \mu_{\bar{X}_A-\bar{X}_B}}{\sigma_{\bar{X}_A-\bar{X}_B}} = \frac{250 - 200}{20} = 2.5$$

  La probabilità corrispondente è $P(Z >= 2.5) = 1 - P(Z <= 2.5)$=0.0062.