Statistica

Capitolo 9

Stima: Ulteriori Argomenti

Obiettivi del Capitolo

Dopo aver completato il capitolo, sarete in grado di:

- Costruire intervalli di confidenza per la differenza tra le medie di due popolazioni dipendenti
- Costruire intervalli di confidenza per la differenza tra le medie di due popolazioni indipendenti (deviazioni standard note o non note)
- Calcolare i limiti dell'intervallo di confidenza per la differenza fra le proporzioni di due popolazioni indipendenti
- Costruire intervalli di confidenza per la varianza della popolazione
- Trovare specifici valori nella tavola della distribuzione chi-quadrato
- Determinare la dimensione del campione necessaria per stimare una media o una proporzione entro uno specifico margine di errore

Stima: Ulteriori Argomenti

Argomenti del Capitolo

- Medie delle Popolazioni: Campioni Dipendenti
- Medie delle Popolazioni: Campioni Indipendenti
- Proporzioni delle Popolazioni
- Varianza della Popolazione

Esempi:

- Stesso gruppo prima vs. dopo un trattamento
- Gruppo 1 vs. Gruppo 2 (indipendenti)
- Proporzione 1 vs. Proporzione 2
- Varianza di una distribuzione normale

Campioni Dipendenti

Campioni Dipendenti

Stima della differenza tra le medie di 2 popolazioni dipendenti

- Campioni appaiati
- Misure Ripetute (prima/dopo)
- Usiamo le differenze fra valori accoppiati:

$$\boxed{d_i = x_i - y_i}$$

- Elimina variazione da soggetto a soggetto
- Assunzioni:
  - Entrambe le popolazioni hanno distribuzione normale

Differenza tra due Medie

Campioni Dipendenti

La i[ma] differenza è d_i , dove

$$d_i = x_i - y_i$$

La stima puntuale della differenza tra le medie delle popolazioni è $\bar{d}$:

$$\bar{d} = \frac{\sum_{i=1}^{n} d_i}{n}$$

La deviazione standard campionaria è:

$$s_d = \sqrt{\frac{\sum_{i=1}^{n} (d_i - \bar{d})^2}{n-1}}$$

n è il numero di coppie nel campione

Intervallo di Confidenza per la Differenza tra due Medie

Campioni Dipendenti

L'intervallo di confidenza per la differenza tra due medie, μ_d , è

$$\bar{d} - t_{n-1,\alpha/2}\frac{s_d}{\sqrt{n}} < \mu_d < \bar{d} + t_{n-1,\alpha/2}\frac{s_d}{\sqrt{n}}$$

Dove

n = dimensione del campione (numero di coppie)

Intervallo di Confidenza per la Differenza tra due Medie

(continuazione)

Campioni Dipendenti

- Il margine di errore è

$$\text{ME} = \text{t}_{n-1,\alpha/2} \frac{s_d}{\sqrt{n}}$$

- $\text{t}_{n-1,\alpha/2}$ è il valore della distribuzione t di Student con (n – 1) gradi di libertà per il quale

$$\text{P}(\text{t}_{n-1} > \text{t}_{n-1,\alpha/2}) = \frac{\alpha}{2}$$

Esempio: Campioni Dipendenti

- Sei individui si iscrivono ad un programma di dimagrimento. Si raccolgono i seguenti dati:

Individuo	Peso: Prima (X)	Dopo (Y)	Differenze, d_i
1	136	125	11
2	205	195	10
3	157	150	7
4	138	140	- 2
5	175	165	10
6	166	160	6
			42

$$\bar{d} = \frac{\sum d_i}{n}$$

$$= 7.0$$

$$s_d = \sqrt{\frac{\sum (d_i - \bar{d})^2}{n-1}}$$

$$= 4.82$$

Esempio: Campioni Dipendenti

(continuazione)

- Per un livello di confidenza del 95%, l'appropriato valore di t è $t_{n\text{-}1,\alpha/2} = t_{5,.025} = 2.571$
- L'intervallo di confidenza al 95% per la differenza tra medie, μ_d , è

$$\bar{d} - t_{n-1,\alpha/2}\frac{s_d}{\sqrt{n}} < \mu_d < \bar{d} + t_{n-1,\alpha/2}\frac{s_d}{\sqrt{n}}$$

$$7 - (2.571)\frac{4.82}{\sqrt{6}} < \mu_d < 7 + (2.571)\frac{4.82}{\sqrt{6}}$$

$$\boxed{1.94 < \mu_d < 12.06}$$

Siccome l'intervallo non contiene il valore zero, sulla base di questi dati limitati possiamo essere confidenti al 95% che il programma di dimagrimento aiuti a perdere peso

Differenza fra Due Medie

Medie delle popolazioni, campioni indipendenti

Obiettivo: Costruire un intervallo di confidenza per la differenza tra le medie di due popolazioni, $\mu_x - \mu_y$

- Diverse fonti di dati
  - Senza relazione
  - Indipendenti
    - Il campione selezionato da una popolazione non influenza il campione selezionato dall'altra popolazione
- La stima puntuale è la differenza fra le due medie campionarie:

$$\bar{x} - \bar{y}$$

Differenza fra Due Medie

(continuazione)

Medie delle popolazioni, campioni indipendenti

- σ_x^2 e σ_y^2 note → Intervallo di Confidenza usa $z_{\alpha/2}$
- σ_x^2 e σ_y^2 non note
  - σ_x^2 e σ_y^2 uguali
  - σ_x^2 e σ_y^2 diverse

  → Intervallo di confidenza usa un valore della distribuzione **t** di Student

σ_x^2 e σ_y^2 note

Medie delle popolazioni, campioni indipendenti

- σ_x^2 e σ_y^2 note *
- σ_x^2 e σ_y^2 non note

Assunzioni:

- campioni casuali e indipendenti
- entrambe le popolazioni hanno distribuzione normale
- varianze delle popolazioni note

σ_x^2 e σ_y^2 note

(continuazione)

Medie delle popolazioni, campioni indipendenti

- σ_x^2 e σ_y^2 note *
- σ_x^2 e σ_y^2 non note

Quando σ_x e σ_y sono note e entrambe le popolazioni hanno distribuzione normale, la varianza di $\overline{X} - \overline{Y}$ è

$$\sigma^2_{\overline{X}-\overline{Y}} = \frac{{\sigma_x}^2}{n_x} + \frac{{\sigma_y}^2}{n_y}$$

…e la variabile aleatoria

$$Z = \frac{(\overline{X} - \overline{Y}) - (\mu_X - \mu_Y)}{\sqrt{\frac{\sigma_x^2}{n_X} + \frac{\sigma_y^2}{n_Y}}}$$

ha distribuzione normale standard

Intervallo di Confidenza, σ_x^2 e σ_y^2 note

Medie delle popolazioni, campioni indipendenti

- σ_x^2 e σ_y^2 note *
- σ_x^2 e σ_y^2 non note

L'intervallo di confidenza per $\mu_x - \mu_y$ è :

$$(\bar{x} - \bar{y}) - z_{\alpha/2}\sqrt{\frac{\sigma_X^2}{n_x} + \frac{\sigma_Y^2}{n_y}} < \mu_X - \mu_Y < (\bar{x} - \bar{y}) + z_{\alpha/2}\sqrt{\frac{\sigma_X^2}{n_x} + \frac{\sigma_Y^2}{n_y}}$$

Intervallo di Confidenza, σ_x^2 e σ_y^2 non note ma Uguali

- Medie delle popolazioni, campioni indipendenti
  - σ_x^2 e σ_y^2 note
  - σ_x^2 e σ_y^2 non note
    - σ_x^2 e σ_y^2 uguali *
    - σ_x^2 e σ_y^2 diverse

Assunzioni:

- campioni casuali e indipendenti
- entrambe le popolazioni hanno distribuzione normale
- varianze delle popolazioni non note, ma uguali

Intervallo di Confidenza, σ_x^2 e σ_y^2 non note ma Uguali

(continuazione)

- Medie delle popolazioni, campioni indipendenti
  - σ_x^2 e σ_y^2 note
  - σ_x^2 e σ_y^2 non note
    - σ_x^2 e σ_y^2 uguali *
    - σ_x^2 e σ_y^2 diverse

Costruzione di stime per intervallo:

- assumiamo che le varianze delle popolazioni siano uguali, quindi usiamo le due varianze campionarie e le combiniamo per ottenere una stima per σ^2
- usiamo un valore di t con $(n_x + n_y - 2)$ gradi di libertà

Intervallo di Confidenza, σ_x^2 e σ_y^2 non note ma Uguali

(continuazione)

- Medie delle popolazioni, campioni indipendenti
  - σ_x^2 e σ_y^2 note
  - σ_x^2 e σ_y^2 non note
    - σ_x^2 e σ_y^2 uguali *
    - σ_x^2 e σ_y^2 diverse

La varianza campionaria ponderata è:

$$s_p^2 = \frac{(n_x - 1)s_x^2 + (n_y - 1)s_y^2}{n_x + n_y - 2}$$

Intervallo di Confidenza, σ_x^2 e σ_y^2 non note ma Uguali

- σ_x^2 e σ_y^2 non note
  - σ_x^2 e σ_y^2 uguali *
  - σ_x^2 e σ_y^2 diverse

L'intervallo di confidenza per $\mu_1 - \mu_2$ è :

$$(\bar{x}-\bar{y}) - t_{n_x+n_y-2,\alpha/2}\sqrt{\frac{s_p^2}{n_x}+\frac{s_p^2}{n_y}} < \mu_X - \mu_Y < (\bar{x}-\bar{y}) + t_{n_x+n_y-2,\alpha/2}\sqrt{\frac{s_p^2}{n_x}+\frac{s_p^2}{n_y}}$$

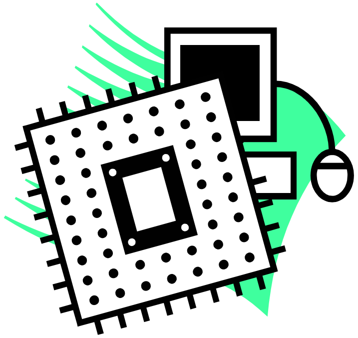
Dove

$$s_p^2 = \frac{(n_x-1)s_x^2 + (n_y-1)s_y^2}{n_x+n_y-2}$$

Esempio: Varianza Campionaria Ponderata

Esaminiamo la velocità di due processori. Calcoliamo un intervallo di confidenza per la differenza nella velocità della CPU. Registriamo i seguenti dati sulla velocità (in Mhz):

	CPU_x	CPU_y
Numero Esaminato	**17**	**14**
Media Campionaria	**3004**	**2538**
Dev St Campionaria	**74**	**56**



Assumiamo che entrambe le popolazioni abbiano distribuzione normale con stessa varianza e usiamo un livello di confidenza del 95%

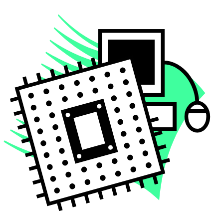
Calcolo Varianza Campionaria Ponderata

La varianza campionaria ponderata è:

$$s_p^2 = \frac{(n_x - 1)s_x{}^2 + (n_y - 1)s_y{}^2}{(n_x - 1) + (n_y - 1)} = \frac{(17-1)74^2 + (14-1)56^2}{(17-1) + (14-1)} = \boxed{4427.03}$$

Il valore t per un intervallo di confidenza al 95% è:

$$\boxed{t_{n_x + n_y - 2,\, \alpha/2} = t_{29,\, 0.025} = 2.045}$$



Limiti dell'Intervallo di Confidenza

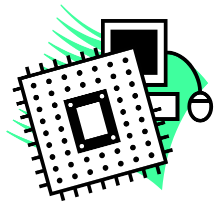
- L'intervallo di confidenza al 95% è

$$(\bar{x}-\bar{y}) - t_{n_x+n_y-2,\alpha/2}\sqrt{\frac{s_p^2}{n_x}+\frac{s_p^2}{n_y}} < \mu_X - \mu_Y < (\bar{x}-\bar{y}) + t_{n_x+n_y-2,\alpha/2}\sqrt{\frac{s_p^2}{n_x}+\frac{s_p^2}{n_y}}$$

$$(3004-2538)-(2.045)\sqrt{\frac{4427.03}{17}+\frac{4427.03}{14}} < \mu_X - \mu_Y < (3004-2538)+(2.045)\sqrt{\frac{4427.03}{17}+\frac{4427.03}{14}}$$

$$\boxed{416.89 < \mu_X - \mu_Y < 515.11}$$

Siamo confidenti al 95% che la differenza tra le velocità medie delle due CPU sia compresa fra 416.89 e 515.11 Mhz.



Intervallo di Confidenza, σ_x^2 e σ_y^2 non note e Diverse

Medie delle popolazioni, campioni indipendenti

- σ_x^2 e σ_y^2 note
- σ_x^2 e σ_y^2 non note
  - σ_x^2 e σ_y^2 uguali
  - σ_x^2 e σ_y^2 diverse *

Assunzioni:

- campioni casuali e indipendenti
- entrambe le popolazioni hanno distribuzione normale
- varianze delle popolazioni non note e diverse

Intervallo di Confidenza, σ_x^2 e σ_y^2 non note e Diverse

(continuazione)

Medie delle popolazioni, campioni indipendenti

- σ_x^2 e σ_y^2 note
- σ_x^2 e σ_y^2 non note
  - σ_x^2 e σ_y^2 uguali
  - σ_x^2 e σ_y^2 diverse *

Costruzione di stime per intervallo:

- le varianze delle popolazioni sono diverse, quindi non ha senso usare una varianza ponderata
- usiamo un valore di t con ν gradi di libertà, dove

$$v = \frac{\left[\left(\frac{s_x^2}{n_x}\right) + \left(\frac{s_y^2}{n_y}\right)\right]^2}{\left(\frac{s_x^2}{n_x}\right)^2 / (n_x - 1) + \left(\frac{s_y^2}{n_y}\right)^2 / (n_y - 1)}$$

Intervallo di Confidenza, σ_x^2 e σ_y^2 non note e Diverse

σ_x^2 e σ_y^2 non note

- σ_x^2 e σ_y^2 uguali
- σ_x^2 e σ_y^2 diverse *

L'intervallo di confidenza per $\mu_1 - \mu_2$ è:

$$(\bar{x}-\bar{y}) - t_{\nu,\alpha/2}\sqrt{\frac{s_x^2}{n_x}+\frac{s_y^2}{n_y}} < \mu_X - \mu_Y < (\bar{x}-\bar{y}) + t_{\nu,\alpha/2}\sqrt{\frac{s_x^2}{n_x}+\frac{s_y^2}{n_y}}$$

Dove

$$\nu = \frac{\left[\left(\frac{s_x^2}{n_x}\right)+\left(\frac{s_y^2}{n_y}\right)\right]^2}{\left(\frac{s_x^2}{n_x}\right)^2/(n_x-1)+\left(\frac{s_y^2}{n_y}\right)^2/(n_y-1)}$$

Proporzioni di Due Popolazioni

Proporzioni delle Popolazioni

Obiettivo: Costruire un intervallo di confidenza per la differenza tra le proporzioni di due popolazioni,

$$p_x - p_y$$

Assunzioni:

Entrambi i campioni sono grandi (in generale almeno 40 osservazioni in ciascun campione)

La stima puntuale per la differenza è: $\hat{p}_x - \hat{p}_y$

Proporzioni di Due Popolazioni

(continuazione)

Proporzioni delle Popolazioni

- La variabile aleatoria

$$Z = \frac{(\hat{P}_x - \hat{P}_y) - (p_x - p_y)}{\sqrt{\dfrac{\hat{p}_x(1-\hat{p}_x)}{n_x} + \dfrac{\hat{p}_y(1-\hat{p}_y)}{n_y}}}$$

ha una distribuzione approssimativamente normale

Intervallo di Confidenza per la Differenza tra due Proporzioni

Proporzioni delle Popolazioni

L'intervallo di confidenza per $p_x - p_y$ è:

$$(\hat{p}_x - \hat{p}_y) \pm z_{\alpha/2}\sqrt{\frac{\hat{p}_x(1-\hat{p}_x)}{n_x} + \frac{\hat{p}_y(1-\hat{p}_y)}{n_y}}$$

Esempio:
Differenza tra due Proporzioni

Costruiamo un intervallo di confidenza al 90% per la differenza tra la proporzione di uomini e la proporzione di donne che hanno una laurea.

- In un campione casuale, 26 dei 50 uomini e 28 delle 40 donne hanno una laurea

Esempio: Differenza tra due Proporzioni

(continuazione)

Uomini: $$\hat{p}_x = \frac{26}{50} = 0.52$$

Donne: $$\hat{p}_y = \frac{28}{40} = 0.70$$

$$\sqrt{\frac{\hat{p}_x(1-\hat{p}_x)}{n_x} + \frac{\hat{p}_y(1-\hat{p}_y)}{n_y}} = \sqrt{\frac{0.52(0.48)}{50} + \frac{0.70(0.30)}{40}} = 0.1012$$

Per un livello di confidenza al 90%, $z_{\alpha/2} = 1.645$

Esempio: Differenza tra due Proporzioni

(continuazione)

I limiti di confidenza sono:

$$(\hat{p}_x - \hat{p}_y) \ \pm \ z_{\alpha/2} \sqrt{\frac{\hat{p}_x(1-\hat{p}_x)}{n_x} + \frac{\hat{p}_y(1-\hat{p}_y)}{n_y}}$$

$$= (.52 - .70) \ \pm \ 1.645\,(0.1012)$$

quindi l'intervallo di confidenza è

$$-0.3465 \ < \ p_x - p_y \ < \ -0.0135$$

Siccome questo intervallo non contiene il valore zero, siamo confidenti al 90% che le due proporzioni non siano uguali

Intervallo di Confidenza per la Varianza della Popolazione

Varianza della Popolazione

- Obiettivo: costruire un intervallo di confidenza per la varianza della popolazione

- L'intervallo di confidenza è basato sulla varianza campionaria, S^2

- Assumiamo: la popolazione ha distribuzione normale

Intervallo di Confidenza per la Varianza della Popolazione

(continuazione)

Varianza della Popolazione

La variabile aleatoria

$$\chi^2_{n-1} = \frac{(n-1)S^2}{\sigma^2}$$

ha una distribuzione chi-quadrato con (n – 1) gradi di libertà

Con $\chi^2_{n-1,\alpha}$ si indica il valore per il quale

$$P(\chi^2_{n-1} > \chi^2_{n-1,\alpha}) = \alpha$$

Intervallo di Confidenza per la Varianza della Popolazione

(continuazione)

Varianza della Popolazione

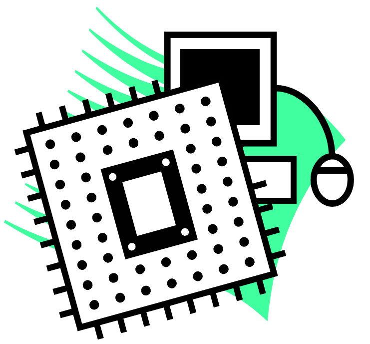
L'intervallo di confidenza all'(1 - α)% per la varianza della popolazione è

$$\frac{(n-1)s^2}{\chi^2_{n-1,\,\alpha/2}} < \sigma^2 < \frac{(n-1)s^2}{\chi^2_{n-1,\,1-\alpha/2}}$$

Esempio: Varianza della Popolazione

Stiamo controllando la velocità di un processore. Registriamo i seguenti dati (in Mhz):

	CPU_x
Dimensione campione	**17**
Media campionaria	**3004**
Dev St campionaria	**74**



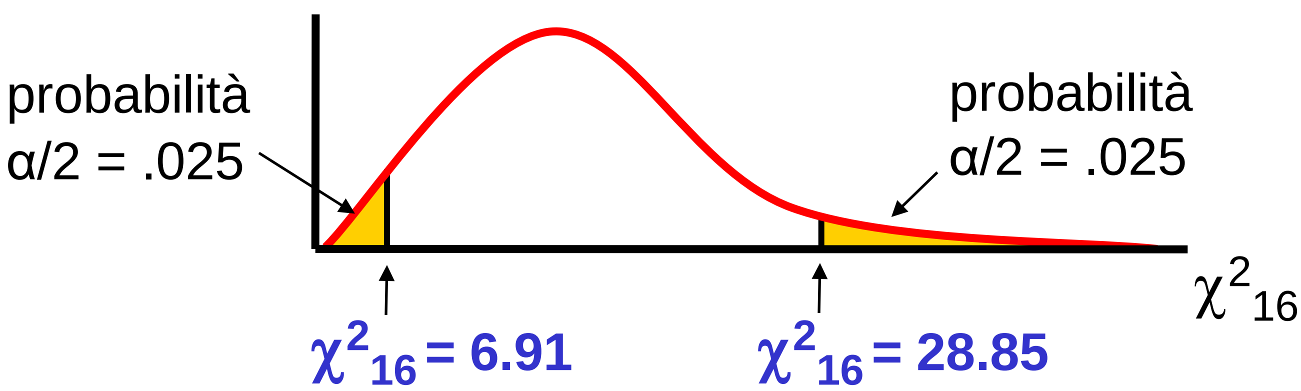
Assumiamo che la popolazione abbia distribuzione normale. Determiniamo l'intervallo di confidenza al 95% per σ_x^2

Trovare i Valori Chi-Quadrato

- n = 17 quindi la distribuzione chi-quadrato ha (n – 1) = 16 gradi di libertà
- α = 0.05, quindi usiamo i valori della chi-quadrato con area 0.025 in ciascuna coda:

$$\chi^2_{n-1,\,\alpha/2} = \chi^2_{16,\,0.025} = 28.85$$

$$\chi^2_{n-1,\,1-\alpha/2} = \chi^2_{16,\,0.975} = 6.91$$



Limiti dell'Intervallo di Confidenza

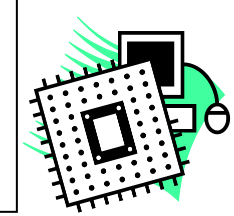
- L'intervallo di confidenza al 95% è

$$\frac{(n-1)s^2}{\chi^2_{n-1,\,\alpha/2}} < \sigma^2 < \frac{(n-1)s^2}{\chi^2_{n-1,\,1-\alpha/2}}$$

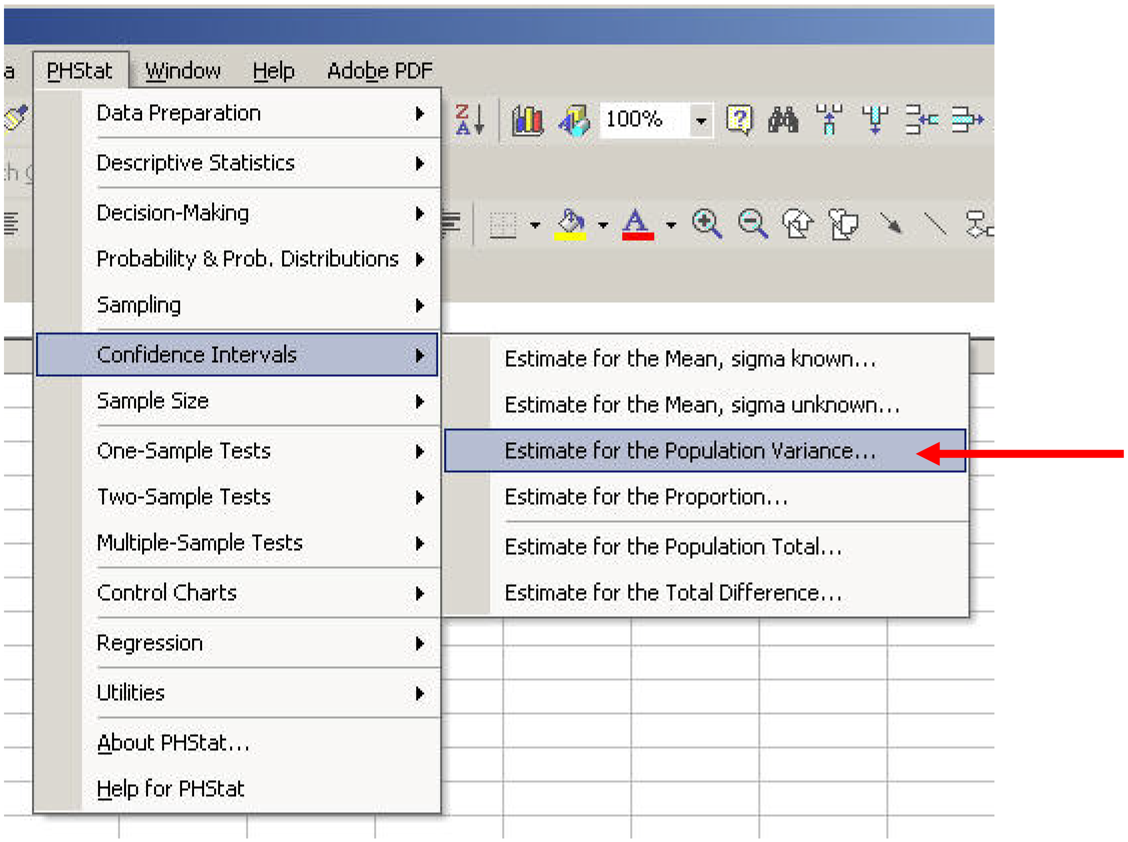
$$\frac{(17-1)(74)^2}{28.85} < \sigma^2 < \frac{(17-1)(74)^2}{6.91}$$

$$\boxed{3037 < \sigma^2 < 12680}$$

In termini di deviazione standard: siamo confidenti al 95% che la deviazione standard della popolazione delle velocità delle CPU sia compresa fra 55.11 e 112.61 Mhz

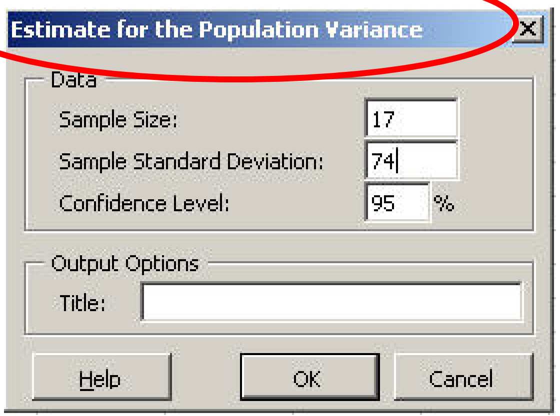


Output PHStat



Output PHStat

(continuazione)



	A	B	C	D	E
1	**Confidence Interval Estimate for the Population Variance**				
2					
3	**Data**				
4	**Sample Size**	**17**			
5	**Sample Standard Deviation**	**74**			
6	**Confidence Level**	**95%**			
7					
8	Intermediate Calculations				
9	Degrees of Freedom	16			
10	Sum of Squares	87616			
11	Single Tail Area	0.025			
12	Lower Chi-Square Value	6.907664			
13	Upper Chi-Square Value	28.84532			
14					
15	**Results**				
16	**Interval Lower Limit for Variance**	**3037.442**			
17	**Interval Upper Limit for Variance**	**12683.88**			
18					
19	**Interval Lower Limit for Standard Deviation**	**55.11299**			
20	**Interval Upper Limit for Standard Deviation**	**112.6227**			
21					
22	*Assumption:*				
23	**Population from which sample was drawn has an approximate normal distribution.**				
24					

Input

Output

Determinazione dell’Ampiezza Campionaria

Determinazione Ampiezza Campionaria

- **Per la Media**
- **Per la Proporzione**

Margine di Errore

- È possibile calcolare l'ampiezza campionaria necessaria per garantire un desiderato margine di errore (ME), con un prefissato livello di confidenza $(1 - \alpha)$

- Il margine di errore è anche chiamato errore di campionamento
  - l'ammontare di imprecisione nella stima del parametro della popolazione
  - l'ammontare aggiunto e sottratto dalla stima puntuale per formare l'intervallo di confidenza

Determinazione dell'Ampiezza Campionaria

Determinazione Ampiezza Campionaria

Per la Media

$$\bar{x} \pm z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$$

Margine di Errore
(errore di campionamento)

$$ME = z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$$

Determinazione dell'Ampiezza Campionaria

(continuazione)

Determinazione Ampiezza Campionaria

Per la Media

$$\mathrm{ME} = z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$$

Risolvere rispetto ad n per ottenere

$$n = \frac{z_{\alpha/2}^{2}\ \sigma^{2}}{\mathrm{ME}^{2}}$$

Determinazione dell'Ampiezza Campionaria

(continuazione)

- Per determinare l'ampiezza campionaria richiesta per la media, dobbiamo conoscere:
  - il livello di confidenza (1 - α) richiesto, che determina il valore $z_{\alpha/2}$
  - il margine di errore accettabile (errore di campionamento), ME
  - la deviazione standard, σ

Esempio: Determinazione dell'Ampiezza Campionaria

Se σ = 45, quale ampiezza campionaria è necessaria per stimare la media entro ± 5 con un livello di confidenza pari al 90%?

$$n = \frac{z_{\alpha/2}^2 \, \sigma^2}{ME^2} = \frac{(1.645)^2 (45)^2}{5^2} = 219.19$$

Quindi l'ampiezza campionaria necessaria è ***n* = 220**

(Sempre arrotondare per eccesso)

Determinazione dell'Ampiezza Campionaria

Determinazione Ampiezza Campionaria

Per la Proporzione

$$\hat{p} \pm z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

$$\text{ME} = z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

Margine di Errore
(errore di campionamento)

Determinazione dell'Ampiezza Campionaria

(continuazione)

Determinazione Ampiezza Campionaria

Per la Proporzione

$$\text{ME} = z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

$\hat{p}(1-\hat{p})$ non può essere maggiore di 0.25, quando $\hat{p}$ = 0.5

Sostituire 0.25 in $\hat{p}(1-\hat{p})$ e risolvere ottenendo:

$$n = \frac{0.25\,z_{\alpha/2}^2}{\text{ME}^2}$$

Determinazione dell'Ampiezza Campionaria

(continuazione)

- La proporzione campionaria e la proporzione della popolazione, $\hat{p}$ e p, sono generalmente non note (siccome il campione non è ancora stato estratto)
- $p(1 – p) = 0.25$ genera il valore massimo per il margine di errore (quindi garantisce che l'ampiezza campionaria ottenuta determini il livello di confidenza richiesto)
- Per determinare l'ampiezza campionaria necessaria per la proporzione, dobbiamo conoscere:
  - Il livello di confidenza richiesto, $(1 - \alpha)$, che determina il valore critico $z_{\alpha/2}$
  - L'errore di campionamento (margine di errore) accettabile, ME
  - La stima $p(1 – p) = 0.25$

Esempio: Determinazione dell'Ampiezza Campionaria

Qual è l'ampiezza campionaria necessaria per stimare la vera proporzione di difetti (in una popolazione molto numerosa) entro 3%, con livello di confidenza pari al 95%?

Esempio: Determinazione dell'Ampiezza Campionaria

(continuazione)

Soluzione:

Per un livello di confidenza pari al 95%, usiamo $z_{0.025} = 1.96$

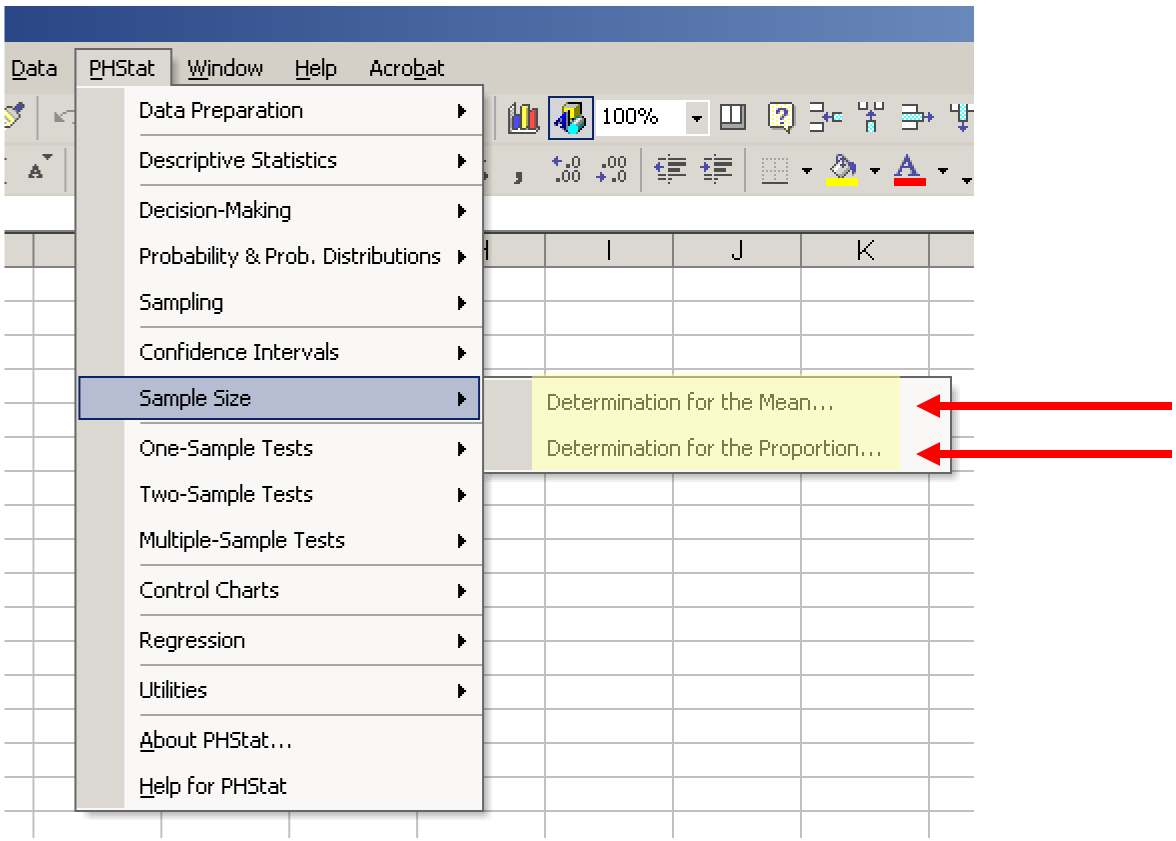
ME = 0.03

Stima $p(1 - p) = 0.25$

$$n = \frac{0.25\, z_{\alpha/2}^2}{ME^2} = \frac{(0.25)(1.96)^2}{(0.03)^2} = \boxed{1067.11}$$

Quindi $n = 1068$

Determinazione dell'Ampiezza Campionaria con PHStat



Riepilogo del Capitolo

- Confrontati due campioni dipendenti
  - Costruiti intervalli di confidenza per la differenza tra le medie
- Confrontati due campioni indipendenti
  - Costruiti intervalli di confidenza per la differenza fra le medie, varianze delle popolazioni note, usando Z
  - Costruiti intervalli di confidenza per la differenza fra le medie, varianze delle popolazioni non note, usando t
  - Costruiti intervalli di confidenza per la differenza fra le proporzioni di due popolazioni
- Costruiti intervalli di confidenza per la varianza di una popolazione distribuita normalmente con l'utilizzo della distribuzione chi-quadrato
- Determinata l'ampiezza campionaria necessaria per garantire un livello di confidenza e un margine di errore prefissati.