Statistica

Monday, 15 May 2023 12:45

1) La concentrazione di PCB nel latte materno ha approssimativamente una distribuizione
   Normale con m, o^2 incogniti
   Abbiamo 20 individui e la nostra media empirica = 5.8 e una deviazione standard empirica = 5.085
   a. IC al 95% per m
      Noi abbiamo un IC per la media di una popolazione normale con varianza incognita.
      (Tutti i passaggi che ha fatto la prof)
      $\overline{X_n} \sim N\left(m, \frac{o^2}{n}\right) \Rightarrow \frac{\overline{X_N} - n}{\sqrt{\frac{o^2}{n}}} \sim N(0,1)$
      Detto questo
      $\frac{(n-1)}{o^2} * S_n^2 \sim X^2(n-1)$
      Quindi ora abbiamo tuttti i dati
      Noi sappiamo che
      $\frac{\overline{X_n} - m}{\sqrt{\frac{S_n^2}{n}}} \sim t(n-1)$
      E questa è la nostra formula che ci serve
      Ora grazie a formule
      $P\left(\left|\frac{\overline{X_n} - m}{\sqrt{\frac{S_n^2}{n}}}\right| < t\right) = 1 - \alpha$
      E svolgendo il tutto
      (negli esercizi passeremo a questo)
      $\left(\overline{X_n} \pm t_{n-1,\frac{\alpha}{2}} * \frac{S_n}{\sqrt{n}}\right)$
      Ora calcoliamo $\alpha \rightarrow 100 * (1-\alpha) = 95$
      $\alpha = 0.05$
      $n = 20$
      $\overline{x_n} = 5.8$
      $S_n = 5.085$
      Ora si sostituisce

Ora si sostituisce
$t_{n-1,\frac{\alpha}{2}} = t_{19,0.025} = 2.09$
$$\left(58 \pm 2.09 * \frac{5.085}{\sqrt{20}}\right) \simeq (3.12, 8.18)$$

b. Estremo superiore di confidenza al 95% per m
$$\overline{x_n} + t_{n-1,\alpha} * \frac{S_m}{\sqrt{n}}$$
$\alpha = 0.05$
$\overline{x_n} = 5.8$
$S_n = 5.085$
$n = 20$
$t_{19,0.05} = 1.73$
Ed ora sostituendo
$$5.8 + 1.73 * \frac{5.085}{\sqrt{20}} \simeq 7.77$$

c. Come cambiano le risposte se $o^2 = 25$
Ora noi abbiamo o^2 nota e m no
Quindi le formule cambiano
$$\frac{\overline{X_n} - m}{\sqrt{\frac{o^2}{n}}} \sim N(0,1)$$
Detto questo
$$P\left(\left|\frac{\overline{X_n} - m}{\sqrt{\frac{o^2}{n}}}\right| < z\right) = 1 - \alpha$$
Siccome ~N
$2\phi(z) - 1 = 1 - \alpha$
$$\phi = 1 - \frac{\alpha}{2}$$
E quindi esce la formula
$$\left(\overline{X_n} \pm z_{\frac{\alpha}{2}} * \frac{o}{\sqrt{n}}\right)$$
$n = 20$
$\overline{x_n} = 20$
$o = \sqrt{o^2} = 5$
$\alpha = 0.05$
$z_{\frac{\alpha}{2}} = z_{0.025} = 1.96$
Sostituendo

$$\left(5.8 \pm 1.96 * \frac{5}{\sqrt{20}}\right) \simeq (3.609, 7.991)$$

2) Dato il seguente campione
1.75, 2.25, 1.9, 2.3, 2.1, 1.7
Proveniente da una legge normale
   a. Fornire una stima puntuale della varianza usando uno stimatore non distorto
      Iniziamo a prendere lo stimatore della varianza
      $$S_n^2 = \frac{1}{n-1}\sum (X_i - \overline{X_n})^2$$
      N=6
      $\overline{x_n} = 2$
      $$S_n^2 = \frac{1}{5} * (\dots)$$
   b. Calcolare IC al 99% della varianza
      Si usa la formula
      $$\frac{n-1}{o^2} S_n^2 \sim X^2(n-1)$$
      Intervallo:
      $$\left(\frac{n-1}{X^2_{n-1,\frac{\alpha}{2}}} * S_n^2, \frac{n-1}{X^2_{n-1,1-\frac{\alpha}{2}}} * S_n^2\right)$$
      $n = 6$
      $\alpha = 0.01$
      $S_n^2 = 0.065$
      $X^2_{5,0.05} = 16.75$
      $X^2_{0.995} = 0.41$
      $(0.019, 0.79)$
   c. Rispondere ad a. b. se la media è nota e pari a m=2
      Qui si utilizza la seguente formula
      $$\overline{S_n^2} = \frac{1}{n}\sum (x_i - m)^2$$
      $$\frac{n}{o^2} * \overline{S_n^2} \sim X^2(n)$$
      E quindi IC
      $$\left(\frac{n * \overline{S_n^2}}{X^2_{n,\frac{\alpha}{2}}}, \frac{n * \overline{S_n^2}}{X^2_{n,1-\frac{\alpha}{2}}}\right)$$

3)

3)

Domanda 6. Per un campione casuale estratto da una popolazione normale con media μ e varianza σ^2 entrambi incognite, l'intervallo bilatero per la media μ a livello di confidenza del 99% è $(-0.12, 0.77)$. Con gli stessi dati, calcoliamo un intervallo di confidenza bilatero al livello del 95%. Quale tra questi è un possibile risultato?

Sono 99% sicuro di prendere la media quindi becco un intervallo grande
Quindi al 95% sicuramente l'intervallo è più piccolo
E l'unico più piccolo è
(-0.08, 0.73)

4) $E\left(\frac{1}{2}\right)$
$\alpha = 2.5\%$ rifiuto
Quindi c'è evidenza empirica contro la distribuizione esponenziale di parametro 1/2
Aka i dati sono in contraddizione tra di loro

5) $\mathrm{n} = 100$

Domanda 8. Si vuole verificare l'ipotesi che un certo medicinale abbassi il livello di colesterolo nel sangue. Si misura il livello di colesterolo a 100 persone prima e dopo la cura con il medicinale. Quale test usereste per verificare l'efficacia del farmaco?

La z non si può applicare siccome dobbiamo stimare tutto
Il chi quadra no siccome è uno su 2 parametri
Quindi test t sulla differenza tra le medie di 2 campioni normali accoppiati

6) $H_0: m \geq 2$
E noi diciamo che $T(x_1, \dots, x_n) < 1$
E poi troviamo che
$T(x_1, \dots, x_n) = 0.8$
Ed è stato trovato che il vero valore di $m = 2.5$

Quindi
Abbiamo trovato che h_0 dice la verità siccome $m = 2.5$
Però dobbiamo rifiutare h_0 siccome abbiamo trovato che $T < 1$
Nota: si rifiuta sempre se è vero
Quindi si commette un errore di prima specie siccome rifiuto qualcosa che è vero

7) $\alpha = 1 - 0.98 = 0.02$
$\sigma = 1$

$$2 * z_{\frac{\alpha}{2}} * \frac{\sigma}{\sqrt{n}} = z_{0.01} * \frac{1}{10} * 2 = z_{0.01} * \frac{1}{5}$$

8) $n = 300$
20

$\frac{20}{300}$

a. $X \sim Be\left(\frac{20}{300}\right)$

$\alpha = 1 - 0.95 = 0.05$

$$\left(\bar{x} \pm z_{\frac{\alpha}{2}}\sqrt{\frac{\overline{x_n}(1-\overline{x_n})}{n}}\right)$$

$z_{1-\frac{\alpha}{2}} = z_{1-0.025} = z_{0.975} = 1.96$

$$\left(\frac{2}{30} \pm 1.96\sqrt{\frac{\frac{2}{30}\left(1-\frac{2}{30}\right)}{300}}\right) = (0.039, 0.095)$$

b. $1.96\sqrt{\frac{\frac{2}{30}\left(1-\frac{2}{30}\right)}{n}} * \frac{1}{2} < 0.02$

$$1.96\sqrt{\frac{\frac{2}{30}\left(1-\frac{2}{30}\right)}{n}} < 0.01$$

$$\sqrt{\frac{\frac{2}{30} * \frac{28}{30}}{n}} < \frac{0.01}{1.96}$$

$$\frac{\frac{1}{15} * \frac{14}{15}}{n} < \left(\frac{0.01}{1.96}\right)^2$$

$$\frac{1}{15} * \frac{14}{15} < \left(\frac{0.01}{1.96}\right)^2 * n$$

$$\frac{1}{15} * \frac{14}{15} * \left(\frac{1.96}{0.01}\right)^2 < n$$

$2390 < n$

c. $p_0 = 0.05$

$h_0: p = p_0$

$h_1: p \neq p_0$

$$\left|\frac{\bar{x}_n - p_0}{\sqrt{p_0(1-p_0)}}\sqrt{n}\right| > z_{\frac{\alpha}{2}}$$

$$\left|\frac{\frac{2}{30} - 0.05}{\sqrt{0.05 * 0.95}}\sqrt{300}\right| > 1.96$$

$1.3 > 1.96$

Falso quindi accetto h_0

9) $X \sim N(100, 20)$

9) $X \sim N(100, 20)$
$m = 100$
$\sigma^2 = 20$
$\bar{x} = 105$

a. Intervallo confidenza guadagno 95%
   Si presuppone che $\sigma^2 = 20$
   $\sigma = \sqrt{20}$
   $\alpha = 1 - \frac{0.05}{2} = 0.975$
   $z_{0.975} = 1.96$
   $\left(105 \pm 1.96 * \frac{\sqrt{20}}{\sqrt{10}}\right)$
   (105 ± 2.77)
   $(102.23, 107.77)$
b. F
   $1.96 * \frac{\sqrt{20}}{\sqrt{n}} < 2$
   $\frac{20}{n} < 1.04$
   $20 < n * 1.04$
   $\approx n > 20$
c. h_0: $m > 100$
   $h_1: m < 100$
   $\frac{\overline{x_n} - m_0}{\sigma}\sqrt{n} < -z_\alpha$
   $\frac{105 - 100}{\sqrt{20}}\sqrt{10} < -z_{0.05}$
   $3.5 < -\cdots$
   Quindi non si può rifiutare h_0
   Quindi c'è guadagno

10) $n_0 = 150$
$vita_0 = 1400$
$\sigma_0 = 120$

$n_1 = 100$
$vita_1 = 1200$
$\sigma_1 = 80$

a. Stima varianza campionaria combinata
   $S_p^2 = \frac{(n_x - 1)S_x^2 + (n_y - 1)S_y^2}{n_x + n_y - 2} = \frac{(150 - 1)120^2 + (100 - 1)80^2}{150 + 100 - 2}$

$$= \frac{2145600 + 633600}{248} = 11206.45161$$

$S_p = 105.86$

b. $\alpha = 0.05$

$h_0: m_x = m_y$

$h_1: m_x \neq m_y$

$$\left| \frac{\bar{x} - \bar{y}}{s_p \sqrt{\frac{1}{n_x} + \frac{1}{n_y}}} \right| > t_{n_x + n_y - 2, \frac{\alpha}{2}}$$

$$\left| \frac{1400 - 1200}{105.86 \sqrt{\frac{1}{150} + \frac{1}{100}}} \right| > t_{248,0.025}$$

$t_{248,0.025} = 1.97$

Rifiuto

c. $\ldots > t_{200,0.0005}$

$\ldots > 1.972$

Quindi si rifiuto ancora

11) N=100

$\sigma^2 = 1$

$\alpha = 1 - 0.95 = 0.05$

Vogliamo media con ampiezza sia la metà di quella del primo intervallo desiderato

$$-z_{0.05} * \frac{1}{\sqrt{n}} * \frac{1}{2} \to z_{0.05} * \frac{1}{2\sqrt{n}} = z_{0.05} * \frac{1}{\sqrt{4n}} \to n = 400$$

12) $n = 625$

$p = 6$

a. $\bar{p} = \frac{6}{625}$

$$\left(\overline{x_n} \pm z_{\frac{\alpha}{2}} \sqrt{\frac{\overline{x_n}(1 - \overline{x_n})}{n}} \right)$$

$$\left(\frac{6}{625} \pm \right)$$

$\alpha = 1 - 0.9 = 0.1$

$z_{1-\frac{0.1}{2}} = z_{0.95}$

$\phi(z_{0.945}) = 1.645$

$$\left(\frac{6}{625} \pm 1.645\sqrt{\frac{\frac{6}{625}\left(1-\frac{6}{625}\right)}{625}}\right) = (0.0032, 0.016)$$

b. $$\left(-\infty, \overline{x_n} + z_\alpha\sqrt{\frac{\frac{6}{625}\left(1-\frac{6}{625}\right)}{625}}\right)$$

$$z_\alpha = z_{0.9} = \phi(z_{0.9}) = 1.285$$

$$\left(-\infty, \frac{6}{625} + 1.285\sqrt{\frac{\frac{6}{625}\left(1-\frac{6}{625}\right)}{625}}\right) = (-\infty, 0.01461)$$

c. Trovare minima ampiezza n affinchè ampiezza = 0.01

$$1.645\sqrt{\frac{\frac{6}{625}\left(1-\frac{6}{625}\right)}{n}} < 0.005$$

$$\sqrt{\frac{\frac{6}{625}\left(1-\frac{6}{625}\right)}{n}} < \frac{0.005}{1.645}$$

$$\frac{\frac{6}{625}\left(1-\frac{6}{625}\right)}{n} < 0.038^2$$

$$\frac{\frac{6}{625}\left(1-\frac{6}{625}\right)}{0.038^2} < n$$

$$n = 1024$$

Non comprendiamo ciò che la prof ha fatto

13) $Bi\left(100, \frac{1}{2}\right)$

$P(X > 60)$

$$\frac{60 - 100 * \frac{1}{2}}{\sqrt{100 * \frac{1}{2} * \left(1-\frac{1}{2}\right)}} = 2$$

Quindi il risultato sarebbe

$\phi(2)$

Però siccome siamo col >

$1 - \phi(2)$

14) F

$$T_1 = \frac{X_1 + \dots + X_n}{\quad} \qquad T_2 = \frac{X_1 + \dots + X_{3n}}{\quad}$$

n , $3n$.

Si estrae campione $X_1, \ldots, X_{3n}$

Allora
Sono tutti e due non distorit siccome sono delle medie
Ed è preferibile T2 siccome è più vicina al campione

15) F

Domanda 7. In un test di ipotesi, l'ipotesi nulla NON viene rifiutata a livello di significatività del 4%. Quale delle seguenti affermazioni è sicuramente vera?

- ■ Il p-value del test è maggiore di 0.04
- ☐ Il p-value del test è compreso tra 0.04 e 0.05.
- ☐ L'ipotesi nulla NON viene rifiutata a livello di significatività del 5%.
- ☐ L'ipotesi nulla viene rifiutata a livello di significatività del 5%.

Questo esercizio si risolve così: