Esercizio 1
La durata della gestazione in donne sane è distribuita normalmente con media pari a 280 giorni e con deviazione standard 10 giorni.

a) Quale proporzione di donne incinte sane avrà una gestazione più lunga di una settimana rispetto al termine della gravidanza atteso?

b) Quali sono i valori di durata della gestazione che definiscono l'intervallo di riferimento per una donna sana? (considerate un intervallo di riferimento che esclude complessivamente il 5% dei valori estremi, sia piccoli che grandi)

Soluzione

La durata della gestazione in donne sane è distribuita normalmente con media pari a 280 giorni e con deviazione standard 10 giorni.

c) Quale proporzione di donne incinte sane avrà una gestazione più lunga di una settimana rispetto al termine della gravidanza atteso?

Y= Durata della gestazione in giorni   Y ~ N (280, 10)

P(Y>287)=?

$$Z = \frac{287-280}{10} = \frac{7}{10} = 0.70$$

$$P(Y > 287) = P(Z > 0.70) = 1 - P(Z < 0.7) = 1 - 0.758 = 0.242$$

d) Quali sono i valori di durata della gestazione che definiscono l'intervallo di riferimento per una donna sana? (considerate un intervallo di riferimento che esclude complessivamente il 5% dei valori estremi, sia piccoli che grandi)

$$P(y_1 < Y < y_2) = 0.95$$

$$y_1 ?\, y_2 ?$$

$$P(y_1 < Y < y_2) = P(z_1 < Z < z_2) = 0.95$$

$$z_1 ?\, z_2 ?$$

$$P(z_1 < Z < z_2) = 0.95 \Longrightarrow \begin{array}{ll} P(Z < z_1) = 0.025 & z_1 = -1.96 \\ P(Z > z_2) = 0.025 & z_2 = 1.96 \end{array}$$

$$z_1 = \frac{y_1 - 280}{10} \Rightarrow y_1 = z_1 * 10 + 280 = -19.6 + 280 = 260.4$$

$$z_2 = \frac{y_2 - 280}{10} \Rightarrow y_2 = z_2 * 10 + 280 = 19.6 + 280 = 289.6$$

$$P(y_1 < Y < y_2) = P(260.4 < Y < 289.6) = 0.95$$

Esercizio 2
Supponete che il peso di un maschio adulto sia distribuito normalmente con media pari a 70 kg e deviazione standard pari a 12 Kg.

a) Descrivete la popolazione di interesse. Da quali individui è costituita?

La popolazione di interesse sono tutti i maschi adulti.
Distribuzione peso ~N(70,12)

b) Calcolate il valore medio atteso e lo standard error della media campionaria del peso per un campione di 10 soggetti.

Prendiamo 10 soggetti dalla popolazione con distribuzione del peso X.
X~media campionaria

Il valore medio atteso è E(X) = 70
Standard error SE(X)= 12 /$\sqrt{10}$

c) Un vecchio ascensore ha una portata nominale di 750 kg. Se 10 uomini adulti salgono sull'ascensore qual è la probabilità che il loro peso complessivo ecceda la portata dell'ascensore?

Variabile Y=ΣX
Media: ΣX= Σ E[X]=10*70=700
Varianza=10^2 * Var[X]=100 *12^2/10 =10*12^2

Y~N(700, 12 x $\sqrt{10}$)
(750 – 700)/120 =

P(Y>750)=P(z>(750-700)/ 12 x $\sqrt{10}$)=P(Z>1.31)=
1-P(Z≤1.31)=1 -0.907 =0.093=9.3%

Oppure

Z=(75-70)/ (12 /$\sqrt{10}$)= 1.31
P(Y>=75)= P(Z>=1.31)
1-P(Z≤1.31)=1 -0.907 =0.093=9.3%

d) Se la distribuzione dei pesi non fosse normale il valore medio atteso e lo standard error calcolati al punto b) sarebbero ancora validi? E la probabilità calcolata in c)?

Se la distribuzione originaria dei valori non fosse normale, il valor medio atteso e lo standard error calcolati in b) sarebbero ancora validi. Tuttavia, le assunzioni di distribuzione richieste per poter applicare la distribuzione media campionaria e lo z test non sarebbero soddisfatte.

Esercizio 3

La valutazione della capacità di percorrere camminando una certa distanza in un dato tempo rappresenta una misura rapida ed economica della performance individuale e una componente importante della qualità della vita. Tale variabile infatti riflette la capacità di svolgere le normali attività quotidiane o, di converso, il grado di limitazione funzionale del soggetto.

In un campione di 46 soggetti affetti da broncopneumopatia cronico ostruttiva (BPCO) di grado moderato, la media e la deviazione standard della distanza percorsa in 6 minuti erano rispettivamente 532.6 m e 78.2 m.

a) Calcolate l'IC al 95% della vera media della distanza percorsa in 6 minuti nella popolazione dei soggetti affetti da BPCO di grado moderato .

La distribuzione si distribuisce normalmente con distribuzione media campionaria. La stima della deviazione standard è $78.2/\sqrt{n}$ cioè $78.2/\sqrt{46}$

$532.6 \pm 1.96 * 78.2/\sqrt{46}$ IC95%= [510.0;555.2]

b) Assumendo che la media della distanza percorsa in 6 minuti nella popolazione dei soggetti con sintomi cronici di tosse/catarro sia 584.4, ritenete che i soggetti con BPCO di grado moderato percorrano mediamente una distanza diversa da quella dei soggetti che riportano sintomi cronici di tosse/catarro?

I soggetti con BPCO di grado moderato percorrono in media una distanza inferiore a quella dei soggetti con sintomi cronici di tosse/catarro in quanto l'intervallo di confidenza non contiene la media della distanza percorsa dalla popolazione di soggetti con sintomi cronici di tosse/catarro.

Esercizio 4

In uno studio sulla mortalità di lavoratori di materiale isolante americani e canadesi furono osservate 26 morti per cancro in un campione di 556 lavoratori. Sulla base di studi effettuati su individui della stessa età ma estratti dalla popolazione generale di americani e canadesi, le morti attese per cancro nel campione avrebbero dovuto essere 14.4.

a) Calcolate l'intervallo di confidenza al 95% per la proporzione di morti per cancro nella popolazione dei lavoratori di materiale isolante canadesi e americani.

b) I risultati ottenuti ci permettono di concludere che la popolazione dei lavoratori è effettivamente più a rischio di morte per cancro rispetto alla popolazione generale?

Soluzione

- π = 14.4/556 = 0.026
- p = 26/556 = 0.047
- ES(p) = $\sqrt{[p(1-p) / n]}$ = 0.0089

n>30:

IC 95%: $$p \pm z_{0.05/2} \cdot \sqrt{\frac{p(1-p)}{n}}$$

$0.047 \pm 1.96 \cdot 0.0089$; (0.03; 0.064)

Dato che l'intervallo di confidenza del rischio di morte del campione estratto non comprende il valore del rischio di morte della popolazione generale, possiamo affermare che i lavoratori di materiale isolante americani e canadesi hanno una proporzione di morti per cancro diversa (e più elevata) rispetto alla popolazione generale.

Esercizio 5
Il livello di colesterolo medio di 10 bambini i cui padri sono morti per malattie cardiache è 200 mg/dl con una deviazione standard (stimata) pari a 50 mg/dl.
Determinate l'intervallo di confidenza (I.C.) al 95% per la media del livello di colesterolo nella popolazione dei bambini i cui padri sono morti per malattie cardiache.

Soluzione

INTERVALLO DI CONFIDENZA AL 95%:

L'intervallo di confidenza al 95% per la media del livello di colesterolo dei bambini i cui padri sono morti per malattie cardiache è:

$$\overline{x} \pm t_{0.05;9} \cdot s/\sqrt{n} \rightarrow 200 \pm 2.2622 \cdot 50/\sqrt{10}$$

200 ± 35.78 mg/dl (164.2 mg/dl; 235.8 mg/dl)