

Esercizi sull'inferenziale



Se anche tu sei di quelli che..... i calcoli vanno bene per fare venire le coliche, prova a fare gli esercizi che seguono...

Finora la frequenza di episiotomie (taglio vaginale per facilitare l'espulsione della testa del feto) in corso di parto è stata del 12%.

Nel nostro ospedale, il 18 novembre hanno partorito 200 donne, il 17 hanno partorito 50 donne e il 16 hanno partorito 100 donne. Quali sono i limiti di controllo al 99,8%?

Giorno 18/11:

$$\sqrt{200 \times 0,12 \times 0,88} = 4,6 \cong 5 \text{ eventi intesi come deviazione standard}$$

Per il 18/11 vanno da 9 a 39 eventi (valore medio \pm 3 deviazioni standard)

Giorno 17/11:

$$\frac{12}{100} \times 50 = 6 \text{ eventi attesi in media}$$

$$\sqrt{50 \times 0,12 \times 0,88} = 2,29 \cong 2 \text{ eventi intesi come deviazione standard}$$

Per il 17/11 vanno da 0 a 12 eventi

Giorno 16/11

$$\frac{12}{100} \times 100 = 12$$

$$\sqrt{100 \times 0,12 \times 0,88} = 3,24$$

Per il giorno 16 vanno da 2 a 22 eventi

Finora il consumo di disinfettante in sala operatoria è stato di 18 litri al giorno, con deviazione standard di 2 litri. Quali sono i limiti di controllo?

Al 68,8%, 16 e 20 litri

Al 95,5%, 14 e 22 litri

Al 99,8% 12 e 24 litri

Finora il consumo di disinfettante in sala operatoria è stato di 18 litri al giorno, con deviazione standard di 2 litri.

E' possibile che un giorno se ne consumino 25 litri?

E' possibile che un giorno se ne consumino 10 litri?

E' possibile che un giorno se ne consumino 20 litri e un altro 17 litri?

E' possibile che un giorno se ne consumino 18 litri?

25 litri è possibile, ma altamente improbabile (<0,5%): bisogna controllare il perché è successo

10 litri è possibile, ma altamente improbabile(<0,05%): bisogna controllare il perché è successo

Un giorno 20 e il successivo 17 litri è quanto accade di solito: sono variazioni entro una deviazione standard in più o in meno dalla media: vi è il 68% di probabilità che accada.

18 litri è il consumo medio: ha una probabilità del 39% di ripetersi.

A metà dei vostri primi 8000 pazienti avete dato la dieta A e sono dimagriti nel 20% dei casi.

All'altra metà avete dato la dieta B sono dimagriti del 22%

Si può dire che è meglio la dieta B per la popolazione che voi seguite di solito?

No, confrontare le semplici frequenze medie comporta una possibilità di errore > 60%.

Si vuole decidere qualcosa che possa esser sbagliato al massimo nel 5% dei casi.

Il prossimo paziente, che probabilità ha di dimagrire con dieta A o con dieta B?

Con dieta A

$$0,20 \pm 1,96 \times \sqrt{1 \times 0,20 \times 0,80} =$$

$$= 0,20 \pm 1,96 \times 0,4 = 0,20 \pm 0,78$$

$$0 \Leftrightarrow 0,98$$

Con dieta B

$$0,22 \pm 1,96 \times \sqrt{1 \times 0,22 \times 0,78} =$$

$$= 0,20 \pm 1,96 \times 0,42 = 0,20 \pm 0,81$$

$$0 \Leftrightarrow 1$$

La probabilità di dimagrire varia dall'impossibile al certo in ambedue i casi. Le due diete sono equivalenti. Forse il diverso successo dipende dall'esser state assegnate alle persone più adatte.

A metà dei vostri primi 8000 pazienti avete dato la dieta A e sono dimagriti nel 20% dei casi. All'altra metà avete dato la dieta B sono dimagriti nel 22% dei casi: Non avete creduto al fatto che non vi sia una sostanziale differenza fra le due diete. Siete intenzionati a pubblicare un articolo per dimostrare che la dieta B è migliore della A. Ci riuscirete?

Per pubblicare una verità che valga in senso assoluto, bisogna considerare che le nostre frequenze: 20 o 22% che siano, si sono comunque verificate su un campione di popolazione. Casualmente i due campioni hanno la stessa numerosità. Calcoliamo i limiti di confidenza al 95%

Con dieta A

$$0,20 \pm 1,96 \times \sqrt{\frac{0,20 \times 0,80}{4000}} =$$
$$= 0,20 \pm 1,96 \times 0,0063 = 0,20 \pm 0,012$$
$$LC_{95} = 0,188 \Leftrightarrow 0,212$$

Con dieta B

$$0,22 \pm 1,96 \times \sqrt{\frac{0,22 \times 0,78}{4000}} =$$
$$= 0,22 \pm 1,96 \times 0,0065 = 0,22 \pm 0,013$$
$$LC_{95} = 0,207 \Leftrightarrow 0,2213$$

Con la dieta A il 95% delle persone dimagrisce in percentuale variabile fra il 18,8 e il 21,2%

Con la dieta B il 95% delle persone dimagrisce in percentuale variabile fra il 20,7 e il 22,1%

La differenza fra le due diete non è significativa, perché $20,7 < 21,2$

Tra i vostri primi 8000 pazienti messi a dieta avete avuto un dimagrimento fino alla perfetta forma fisica nel 30% dei casi.

Se fra i prossimi 100 ne dimagriscono 27 va bene?

Se fra i prossimi 100 ne dimagriscono 82 va bene?

Se fra i prossimi 100 ne dimagriscono 12 va bene?

Se fra i prossimi 100 ne dimagriscono 2 va bene?

In media si prevede che diventino in perfetta forma fisica

30 persone, con deviazione standard = $\sqrt{100 \times 0,3 \times 0,7} = 4,58 \cong 5$ persone

oscillazioni da 25 a 35 persone si verificano con probabilità del 68,8%

oscillazioni da 20 a 40 persone si verificano con probabilità del 95,5%

oscillazioni da 15 a 45 persone si verificano con probabilità del 99,8%

E' prevedibilissimo che ne dimagriscano 27 anziché 30

È strano che ne dimagriscano 82, >45, bisogna verificare il perchè

È strano che ne dimagriscano 12, <15, bisogna verificare il perchè

È stranissimo che ne dimagriscano soltanto 2, bisogna verificare il perchè

E' stata messa a punto una nuova tecnica di indagine di laboratorio per il dosaggio della glicemia. Per verificare quali sono i "valori normali" con la nuova tecnica dosiamo la concentrazione di glucosio nel sangue su un campione di 300.000 persone sicuramente non diabetiche, provenienti da tutte le regioni di Italia, con età adulta variabile da 14 a 90 anni, maschi e femmine: insomma un campione rappresentativo della popolazione italiana.

Le misure rilevate portano a un valore medio di 80 e a un errore standard di 7. Quali saranno i "valori normali" da indicare nel foglietto degli esiti?

Si calcolano i limiti di confidenza al 95% : 7 è l'errore standard.

$$7 \times 1,96 = 13,72$$

limite inferiore: $80 - 13,72 = 66,28$, circa 66

limite superiore: $80 + 13,72 = 93,72$, circa 94

V.N. (66-94), si verificano nel 95% degli individui non diabetici

Un paziente ha un valore di uricemia (concentrazione di acido urico nel sangue) di 7,9. I valori normali sono per i maschi da 4 a 6. Il paziente è sicuramente malato? Se no, che probabilità ha di essere malato e che probabilità ha di esser sano?

No, non è sicuramente malato per un solo esame alterato. La probabilità di esser sano, però è di gran lunga inferiore al 2,5%

Il 95% dei neonati femmine alla nascita pesa dai 3 ai 4,5 Kg. Partorire una bambina di 6 Kg. è impossibile? Se non è impossibile, che probabilità c'è di partorire una bambina di 6 Kg.?

No, non è impossibile, ma la probabilità è di gran lunga inferiore al 2,5%

Con una certa dieta A, su 400 persone ne sono dimagrite 300 (75%). Con la Dieta B, su 500 persone ne sono dimagrite 400. (80%) vi è una differenza statisticamente significativa fra le due diete?

Dieta A

$$1,96 \times \sqrt{\frac{0,75 \times 0,25}{400}} = 1,96 \times 0,0216 = 0,0425$$

Dieta B

$$1,96 \times \sqrt{\frac{0,8 \times 0,2}{500}} = 1,96 \times 0,018 = 0,035$$

79,25%-70,75%, limiti di confidenza al 95% del dimagrimento con dieta A
83,5% - 76,5%, limiti di confidenza al 95% del dimagrimento con dieta B
Non vi è differenza s.s. fra le due diete perché 79% > 76%

Le persone con epatite cronica hanno livelli di transaminasi superiori a 45. I livelli "normali" nelle persone sane vanno da 12 a 45. E' impossibile che una persona con 46 di transaminasi sia sana? Se no, che probabilità ha di esser sana?

La probabilità di esser sana è inferiore al 2,5%

Finora nel reparto XY si sono consumati in media 58 Kg. di cotone idrofilo al mese, con deviazione standard di 8 Kg. Ritenete impossibile che il prossimo mese se ne consumino 78 Kg.? Perché?

Non è impossibile, perché siamo dentro le tre deviazioni standard

Da un campione di 1600 donne otteniamo questi risultati con la terapia A: durata media del travaglio ore 9. La deviazione standard di questo campione è di minuti 45. Quali sono i limiti di confidenza al 95%?

Su un campione di 2500 donne abbiamo sperimentato nuove tecniche di assistenza al travaglio, con la terapia B la durata media del travaglio è di ore 8 e la deviazione standard di questo campione è di 2 ore. Quali sono i limiti di confidenza al 95%?

Le due terapie sono diverse in modo statisticamente significativo?

Terapia A

$$\frac{45}{\sqrt{1600}} = \frac{45}{40} = 1,125 = e.s._A$$

$$540' \pm 1,96 \times 1,125' = 540' \pm 2,205'$$

$$LC_{95\%} = 537,795' \Leftrightarrow 542,205'$$

Terapia B

$$\frac{120}{\sqrt{2500}} = \frac{120}{50} = 2,4 = e.s._B$$

$$480' \pm 1,96 \times 2,4' = 480' \pm 4,704'$$

$$LC_{95\%} = 475,296' \Leftrightarrow 484,704'$$

La differenza è s.s., perché i due intervalli, che comprendono il 95% delle persone trattate, non si sovrappongono mai

La durata media della gravidanza è 40 settimane, ma di solito si danno come normali durate da 39 a 41 settimane. Questi sono i limiti di confidenza al 95% calcolati su un numero tanto ampio di donne da esser quasi infinito.

L'errore standard, quale sarà?

$$LCS_{95} = 41 = 40 + 1,96 e.s.$$

$$e.s. = \frac{41 - 40}{1,96} = 0,51$$

Finora avete utilizzato una vostra dieta su 101 pazienti. Il dimagrimento medio è stato Kg. 8, con d.s di Kg. 2. Nel vostro ospedale sinora su 101 000 pazienti si era ottenuto il dimagrimento di Kg 6 con deviazione standard di Kg. 1

Vi viene il dubbio che la vostra dieta sia migliore.

1. Dire che la vostra è migliore perché fa dimagrire di 8 Kg., contro i soliti 6, con che sicurezza lo si può affermare?

Veramente molto scarsa: del 39%, con possibilità di errore del 61%

Finora avete utilizzato una vostra dieta su 101 pazienti. Il dimagrimento medio è stato Kg. 8, con d.s di Kg. 2

Nel vostro ospedale sinora su 101 000 pazienti si era ottenuto il dimagrimento di Kg 6 con deviazione standard di Kg. 1

Vi viene il dubbio che la vostra dieta sia migliore.

1. Come vi comportate per fare un'affermazione che possa esser valida al 95%, con possibilità di errore del 5%?
2. E' una bomba la vostra dieta?

1. Intanto bisogna ricavare l'errore standard da una deviazione standard che era calcolata come universale (l'universo dei vostri 101 pazienti)

$$\sigma = 2Kg = \sqrt{\frac{\text{devianza}}{101}}$$

si ricava che la devianza=404Kg.²
Calcoliamo ora l'errore standard

$$e.s. = \frac{\sqrt{\frac{404Kg^2}{101-1}}}{\sqrt{101}} = \frac{2,01Kg.}{10,05} = 0,2Kg.$$

Moltiplichiamo 0,2 Kg. per 1,96 e otteniamo che al 95% di probabilità i vostri pazienti dimagriscono da 8,039 a 7,961 (limiti di confidenza al 95%)

Adesso bisogna fare lo stesso calcolo con i 101000 pazienti dell'ospedale

$$\sigma = 1Kg = \sqrt{\frac{\text{devianza}}{101000}}$$

si ricava che la devianza è 101000 Kg.²
Calcoliamo ora l'errore standard

$$e.s. = \frac{\sqrt{\frac{101000Kg^2}{101000-1}}}{\sqrt{101000}} = \frac{1,000Kg.}{317,80} = 0,0032Kg.$$

Moltiplichiamo 0,0032 Kg. per 1,96 e otteniamo che al 95% di probabilità i pazienti dell'ospedale dimagriscono da 5,994 a 6,006 Kg. (limiti di confidenza al 95%).

La vostra dieta è molto più efficace di quella dell'ospedale basta guardare la differenza riportando i valori su assi cartesiani:



Con la vostra dieta finora sono dimagriti il 72% dei pazienti (ne avete visti 100 in tutto).

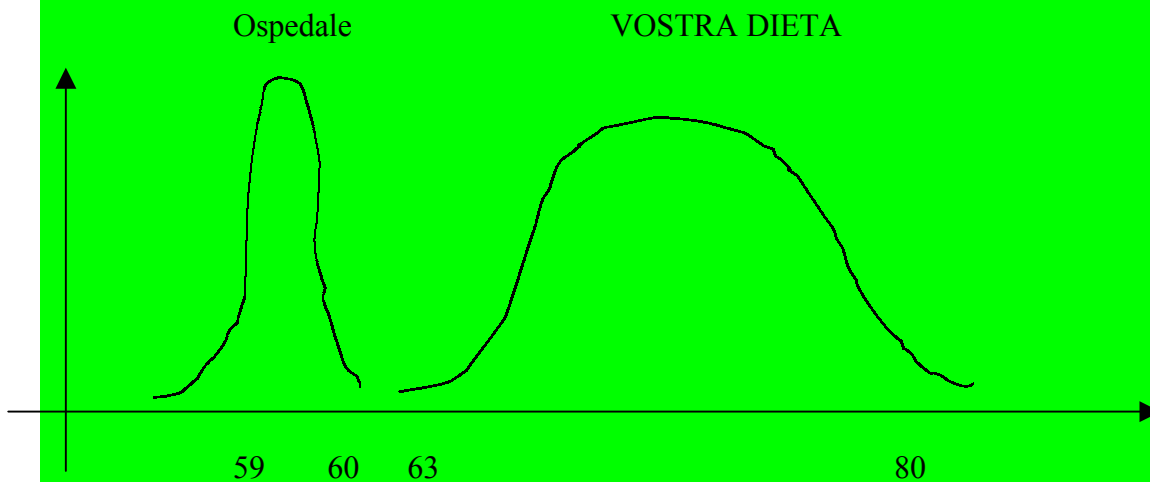
In ospedale, in genere dimagriscono il 60% dei pazienti (ne han messi a dieta 120 000 finora)
Qual è la dieta migliore, salvo errore 5%?

Ecco i vostri risultati, con confidenza del 95%

$$0,72 = p_s$$
$$\sqrt{\frac{0,72 \times 0,28}{100}} = 0,0449 = e.s._p$$
$$0,0449 \times 1,96 = 0,088$$
$$LC_{95} = 0,632 \Leftrightarrow 0,808$$
$$63,2\% \Leftrightarrow 80,8\%$$

ecco quelli dell'Ospedale

$$0,60 = p_s$$
$$\sqrt{\frac{0,60 \times 0,40}{120000}} = 0,0014 = e.s._p$$
$$0,0014 \times 1,96 = 0,00277$$
$$LC_{95} = 0,597 \Leftrightarrow 0,60277$$
$$59,7\% \Leftrightarrow 60,2\%$$



La vostra dieta è un pochino meglio, perché $63 > 60$. Certo non è una differenza strepitosa, ma è significativa

Finora con la vostra dieta sono dimagriti il 72% dei pazienti. Vi chiedete come costruire carte di controllo per i vostri prossimi 200 pazienti

$$\frac{72}{100} \times 200 = 144 = n\pi$$

$$\sigma_{\pi} = \sqrt{200 \times 0,72 \times 0,28} = 6,34 \cong 6$$

Le linee della carta saranno:

162 _____
156 _____
150 _____
144 _____
138 _____
132 _____
126 _____

Come sarà la carta di controllo dell'ospedale per i prossimi 200 pazienti?

$$\frac{72}{100} \times 12000 = 72000 = n\pi$$

$$\sigma_{\pi} = \sqrt{200 \times 0,72 \times 0,28} = 6,34 \cong 6$$

72018 _____
72012 _____
72006 _____
72000 _____
71994 _____
71988 _____
71982 _____

Abbiamo valutato 3 diverse diete in pazienti diabetici. Risultati del calo glicemico ottenuto:

DIETA A	DIETA B	DIETA C
$\mu=125$	$\mu=100$	$\mu=110$
$\sigma=5$	$\sigma=6$	$\sigma=3$
N= 120	N=110	N=150

Quale dieta è la migliore, se una è migliore delle altre?

DIETA A

$$\sigma = 5 = \sqrt{\frac{dev_A}{N}}$$

$$25 = \frac{dev_A}{120}$$

$$dev_A = 120 \times 25 = 3000$$

$$es_A = \frac{\sqrt{\frac{3000}{119}}}{\sqrt{120}} = \frac{5,021}{10,95} = 0,46$$

DIETA B

$$\sigma = 6 = \sqrt{\frac{dev_B}{110}}$$

$$36 = \frac{dev_B}{110}$$

$$dev_B = 110 \times 36 = 3960$$

$$es_B = \frac{\sqrt{\frac{3960}{109}}}{\sqrt{110}} = \frac{6,03}{10,49} = 0,57$$

DIETA C

$$\sigma = 3 = \sqrt{\frac{dev_C}{150}}$$

$$9 = \frac{dev_C}{150}$$

$$dev_C = 150 \times 9 = 1350$$

$$es_C = \frac{\sqrt{\frac{1350}{149}}}{\sqrt{150}} = \frac{3,01}{12,25} = 0,25$$

$$\begin{aligned} \text{A, media stimata} = 125, \text{ LC al } 95\% & \quad 125 - (1,96 \times 0,46) = 124,098 \\ & \quad 125 + (1,96 \times 0,46) = 125,9026 \end{aligned}$$

$$\begin{aligned} \text{B, media stimata} = 100, \text{ LC al } 95\% & \quad 100 - (1,96 \times 0,57) = 98,88 \\ & \quad 100 + (1,96 \times 0,57) = 101,12 \end{aligned}$$

$$\begin{aligned} \text{C, media stimata} = 110, \text{ LC al } 95\% & \quad 110 - (1,96 \times 0,25) = 109,5 \\ & \quad 110 + (1,96 \times 0,25) = 110,49 \end{aligned}$$

Le tre diete sono significativamente diverse, la più efficace è la A, segue la C, per ultima la B

Dobbiamo studiare una popolazione di obesi così costituita:

	Uomini	Donne
30-39 anni	154	695
40-49 anni	687	254
50-59 anni	300	450
60-69 anni	200	500

Volendo fare un campionamento stratificato di tipo probabilistico,

- Quante donne ultra sessantenni occorre estrarre?
- Quanti uomini sotto i 40 anni occorre estrarre?
- Quante persone fra i 40 e i 49 occorre estrarre?

Il 10% di 500, cioè 50 donne ultra sessantenni

Il 10% di 154, cioè 15 o 16 uomini sotto i 40 anni

Il 10% di 687, cioè 69 uomini e il 10% di 254, cioè 25 o 26 donne, quindi 94 o 95 persone sotto i 40 anni, divise tra i due sessi nel modo predetto.

Non possiamo eseguire a tutte queste persone un dosaggio della glicemia, per pochi fondi. Si decide di fare un campionamento randomizzato stratificato al 20%. Quante persone dovremo sorteggiare e sottoporre a glicemia?

	Uomini	Donne
30-39 anni	16	70
40-49 anni	69	26
50-59 anni	30	45
60-69 anni	20	50

Occorre, cella per cella, cioè strato per strato, calcolare il 20%. In tutto saranno 53 persone.

	Uomini	Uomini campione	Donne	Donne campione
30-39 anni	16	3	70	2
40-49 anni	69	14	26	5
50-59 anni	30	6	45	9
60-69 anni	20	4	50	10
totale		27		26

Se, dalla tua peculiare esperienza lavorativa, emergessero problematiche di carattere generale, che vorresti risolvere, scrivi a questo indirizzo:

elbalcos@alma.unibo.it

comunque una risposta e la tua esperienza potrà esser utile ad altri. Grazie