

ESERCIZI DI INFERENZA STATISTICA E STUDIO DELLE ASSOCIAZIONI

ES 1

In un collettivo di 40 pazienti osservati, la media dei globuli bianchi era pari a $2.9 (\times 1000/\text{ml}^3)$ e la varianza era pari a 0.336. Fornire una stima puntuale e una stima intervallare mediante Intervallo di Confidenza al 90% del valore medio atteso dei globuli bianchi nella popolazione dei pazienti affetti dalla malattia.

ES 2

Un gruppo di 60 soggetti viene esaminato per la presenza una certa anomalia genetica, ed essa viene individuata in 14 persone. Qual è la stima della prevalenza di questa anomalia genetica? (Fornire un intervallo di stima al livello di confidenza del 95%)

ES 3

La distribuzione del Body Mass Index ($\text{BMI} = \text{peso (kg)} / [\text{altezza (m)}]^2$) in una certa popolazione è approssimativamente Normale, con deviazione standard pari a 2.2. In un gruppo di 26 individui estratti a caso da quella popolazione, la media del BMI è risultata pari a 24.2. Calcolare un stima puntuale e intervallare (IC95%) della media del BMI, e fornire la sua interpretazione.

Poi, assumendo che la media incognita sia proprio uguale a quella stimata (stima puntuale), fornire un range di valori del BMI tale che il 95% della popolazione sia compreso in quel range.

ES 4

Fra gli 80 pazienti trattati con un nuovo farmaco per una certa malattia, si osserva un valore medio dei globuli rossi piuttosto basso (pari a $3.4 \times 10^3/\text{ml}$), inferiore al valore medio di riferimento che per questo tipo di soggetti è pari a $5 \times 10^3/\text{ml}$. Sapendo che la deviazione standard del campione era pari a $4.5 \times 10^3/\text{ml}$, si vuole verificare se è corretto desumere che il trattamento può indurre un abbassamento dei globuli rossi. Impostare quindi il test di ipotesi e svolgerlo sia col metodo del p-value, sia col metodo delle regioni di rifiuto.

ES 5

Fra gli stessi 80 pazienti dell'esercizio 4, si osserva anche l'insorgere di dermatite in 25 di essi. La dermatite è in effetti un disturbo collaterale che si presenta nel 10% dei soggetti affetti dalla malattia. Si deve desumere che il trattamento modifica l'incidenza della dermatite? Effettuare un test di significatività a due code sia col metodo del p-value, sia col metodo delle regioni di rifiuto, al livello del 5%. Esprimere poi l'effetto del trattamento sulla probabilità di insorgenza della dermatite tramite il Risk Ratio.

ES 6

Un medico dietologo "inventa" un nuovo tipo di dieta (A) per facilitare la riduzione di peso. Decide di confrontarne l'efficacia rispetto al tipo di dieta che prescriveva precedentemente (B). Dispone dei risultati di due gruppi di pazienti, 39 trattati con la vecchia dieta B, che hanno perso in media 2.9 kg ($\text{std}=1.2\text{kg}$), e 34 trattati con la nuova dieta A, che hanno avuto una riduzione di peso media di 3.5 kg ($\text{std}=1.1\text{kg}$). Ipotezzando la Normalità della perdita di peso e l'uguaglianza delle varianze, effettuare un test di ipotesi al livello del 5% bilaterale. E' corretto affermare che la dieta A è migliore della dieta B? quali valutazioni sui due campioni bisognerebbe fare per essere maggiormente sicuri della conclusione?

ES 7

Un gruppo di 7 soggetti viene sottoposto a un trattamento per l'ipertensione. Si misura la pressione diastolica (minima) prima e dopo un certo numero di giorni di trattamento: i risultati sono in tabella. Si vuole valutare se il trattamento ha effetto: possiamo impostare un test fra quelli che conosciamo?

Soggetto	1	2	3	4	5	6	7
prima	118	100	108	97	120	102	104
dopo	100	105	89	78	96	97	88

ES 8

186 pazienti hanno fatto una terapia per una certa malattia; 122 hanno seguito la terapia A, gli altri 64 hanno seguito la terapia B. Nel gruppo A, hanno risposto 37 soggetti. Nel gruppo B, hanno risposto 32 soggetti. I dati sono riportati anche nella sottostante tabella. Calcolare le percentuali di risposta e il Risk Ratio assumendo come gruppo di confronto (baseline) il trattamento A.

Per valutare la significatività della differenza dovremmo usare dei test appositi; in alternativa, calcoliamo una misura dell'associazione fra Trattamento e Risposta sulla tabella doppia ed effettuiamo il test per l'ipotesi di assenza di relazione.

Trattamento	Risposta		Tot
	no	si	
A	85	37	122
B	32	32	64
tot	117	69	186

ES 9

Data la seguente serie di dati su Peso e Altezza di 8 pazienti, calcolare una misura dell'associazione, poi, individuare la retta che descrive la dipendenza del Peso dall'Altezza, e dire che Peso è predetto dal modello per un soggetto alto 165 cm. Osservando che il soggetto id=1 ha altezza 165 cm e peso osservato 55.2, commentare la congruenza fra modello e dati.

id	Altezza, cm	Peso, kg
1	165	55.2
2	157	48.1
3	168	65.6
4	178	72.0
5	171	73.5
6	182	77.9
7	182	80.3
8	153	49.4

ES 10

La seguente tabella riporta i risultati di un modello di regressione logistica (per outcome dicotomici), ossia l'effetto stimato di una serie di fattori sul rischio di progressione di malattia a breve termine. L'effetto è valutato in termini di odds ratio, con intervallo di confidenza al livello 95%, e l'ultima colonna riporta il p-value per la significatività dell'effetto. Per ciascuna delle seguenti domande, dire se la risposta fornita è Vera o Falsa, e se Falsa correggerla opportunamente.

Fattore	OR	95% CI		p-value
Trattamento B vs. A	0.47	0.31	0.72	0.000428
Precedente trattamento	1.44	0.88	2.36	0.142867
Microglob. Beta2 alte	2.36	1.35	4.12	0.002690
Anomalia DEL13	2.03	1.29	3.17	0.000518

- a) Il trattamento ha effetto sul rischio di progressione?
- Sì, ha un effetto fortemente significativo ($p < 0.001$), B dà più rischio di progressione del trattamento A

- b) Qual è l'effetto di un precedente trattamento?
 - L'aumento medio dell'odds di progressione è stimato pari a 44%, ma la stima è molto variabile, secondo questo modello potrebbe anche essere un fattore protettivo; questo fattore non è statisticamente significativo
- c) Qual è il fattore di rischio più importante?
 - E' la presenza dell'anomalia DEL13 ($p=0.0005$)
- d) Qual è l'effetto delle microglobuline Beta2 alte?
 - L'odds di progressione aumenta circa del 36% (l'effetto è molto significativo)

SOLUZIONI

ES 1

La stima puntuale della media della popolazione è semplicemente la media aritmetica del campione, qui 2.9; infatti, la media aritmetica campionaria ne è uno stimatore non distorto e con altre proprietà (efficienza, consistenza e altre).

La stima intervallare con IC90% si ottiene applicando la formula fornita – infatti, il campione è grande ($n=40$), quindi possiamo utilizzare la formula, stimando la varianza della popolazione dal campione, qui $s^2=0.336$ (e quindi $s=0.579655\approx 0.58$).

$$\left(\bar{x} - z_{\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}, \bar{x} + z_{\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}} \right) = \left(2.9 - 1.64 \cdot \frac{0.58}{\sqrt{40}}, 2.9 + 1.64 \cdot \frac{0.58}{\sqrt{40}} \right) = (2.75, 3.05)$$

(notare che utilizziamo il quantile 1.64, che lascia nella coda esterna una probabilità pari a 5%, e quindi nel complesso delle 2 code il 10%; solitamente si costruiscono IC al livello 5%, e quindi si utilizza il quantile 1.96, che lascia appunto nella coda esterna un'area pari a 2.5%, e nella somma delle due code una probabilità del 5%)

Quindi per un paziente di questo tipo ci aspettiamo in media un valore dei WBC compreso fra 2.75 e 3.05 (interpretazione clinica). Attenzione: questo non è il range dei valori dei WBC che ci possiamo attendere per qualsiasi soggetto! Si veda l'esercizio 3 – e si riproduca per esso il ragionamento per questo esercizio.

ES 2

La stima puntuale della prevalenza nella popolazione non è altro che la proporzione, o percentuale, osservata nel campione, che è $14/60=23\%$. Siccome il campione è grande, per l'inferenza possiamo assimilare il comportamento della proporzione a quello della media aritmetica per un campione di osservazioni di un carattere quantitativo continuo. Quindi:

$$\left(\hat{\pi} - z_{\frac{\alpha}{2}} \cdot \sqrt{\frac{\hat{\pi}(1-\hat{\pi})}{n}}, \hat{\pi} + z_{\frac{\alpha}{2}} \cdot \sqrt{\frac{\hat{\pi}(1-\hat{\pi})}{n}} \right) = \left(0.23 - 1.96 \sqrt{\frac{0.23 \cdot 0.77}{60}}, 0.23 + 1.96 \sqrt{\frac{0.23 \cdot 0.77}{60}} \right) = (0.13, 0.34)$$

Dunque la prevalenza dell'anomalia nella popolazione è compresa fra il 13% e il 34%.

ES 3

Il testo dell'esercizio ci permette di assumere la Normalità e la conoscenza della varianza (ci dà la deviazione standard, che ne è la radice quadrata). Quindi sebbene il campione sia piccolo, possiamo usare la nostra formula dell'IC - la stima puntuale di μ è semplicemente 24.2.

$$\left(\bar{x} - z_{\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}, \bar{x} + z_{\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}} \right) = \left(24.2 - 1.96 \cdot \frac{2.2}{\sqrt{26}}, 24.2 + 1.96 \cdot \frac{2.2}{\sqrt{26}} \right) = (23.35, 24.05)$$

Dunque per un soggetto estratto a caso dalla stessa popolazione da cui proviene il campione, ci aspettiamo un BMI pari a 24.2, ovvero compreso fra 23.35 e 24.05.

Diciamo che questa è la previsione per l'individuo "medio", ma poi sappiamo che attorno all'individuo medio ve ne sono altri diversi, poichè il carattere (BMI) presenta una certa variabilità – espressa dalla deviazione standard della popolazione, 2.2. Per avere dei range di valori possibili per il BMI, facciamo quindi riferimento alle proprietà della Normale, di cui ora assumiamo di conoscere anche la media (avendola stimata nel campione).

Essendo la distribuzione del BMI Normale di media stimata 24.2 e di deviazione standard 2.2, per le proprietà della Normale sappiamo che il 95% dei valori del BMI sono compresi fra $\mu - 1.96 \cdot \sigma$ e $\mu + 1.96 \cdot \sigma$, ossia fra 19.89 e 28.51. Quest'ultimo è il range che ci aspettiamo per il 95% dei soggetti provenienti da quella popolazione.

(Un range ancora più ampio, di raggio 3σ attorno alla media, fornisce la quasi totalità – 99.7% - dei valori che possiamo attenderci: (17.6,30.8))

ES 4

Il test per rispondere a questo quesito va impostato in questo modo: l'ipotesi di base è la condizione 'neutra', di riferimento, e quindi che il valore μ della media dei globuli rossi della popolazione è 5 (tralasciamo l'unità di misura); l'ipotesi alternativa è che μ sia inferiore a 5.

Nota: il quesito focalizza sull'inferiorità, e quindi suggerisce di fare un test a una coda; tuttavia, non è opportuno fare test a una coda con l'usuale livello di significatività 5%: è preferibile fare il test a due code al livello 5%, ovvero, equivalentemente, fare il test a una coda con il livello alfa dimezzato al 2.5%: entrambe gli approcci portano a prendere come limite della regione di rifiuto il valore 1.96.

$$H_0: \mu=5 \quad \text{vs} \quad H_1: \mu<5$$

Si osservi che nella formulazione delle ipotesi non si utilizzano i valori calcolati nel campione!!

Il test sulla media è il t-test, possiamo usare la nostra formula che vale in casi limitati, poiché abbiamo un campione molto grande ($n=80$); la statistica è:

$$t = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} = \frac{3.4 - 5}{4.5/\sqrt{80}} = -3.18$$

Osserviamo che il valore della statistica test è molto lontano dal valore che ci attenderemmo nel caso valga l'ipotesi nulla, che è zero (assenza di differenza fra la media della nostra popolazione e il valore di riferimento 5). Esso è al di sotto della soglia -1.96, quindi il test rifiuta l'ipotesi nulla al livello del 2.5% (test unilaterale; ovvero al livello del 5% per il test bilaterale).

Per il calcolo del p-value, impostando il test come unilaterale, dobbiamo calcolare il valore nella coda inferiore al di sotto di -3.18. Siccome la tavola non ci permette di leggere le aree per i valori negativi, calcoliamo l'area che ci interessa come:

$$p = 1 - \Phi(3.18) = 1 - 0.999 = 0.001$$

(Se avessimo impostato il test a due code, questo valore andrebbe moltiplicato per 2: $p=0.002$)

ES 5

Con questo esercizio ripetiamo praticamente il precedente, infatti con questo grande campione l'inferenza sulla percentuale π di casi che soffrono di dermatite (carattere dicotomico) è svolta con gli stessi metodi usati per una media di un carattere continuo.

$$H_0: \pi = 0.10 \quad \text{vs} \quad H_1: \pi \neq 0.10$$

(Rimarchiamo che nel formulare le ipotesi non utilizziamo i valori osservati sui campioni; qui il test ha una ipotesi alternativa bilaterale perché così ci prescrive l'esercizio; i test dovrebbero comunque sempre essere bilaterali, o, se proprio si può escludere a priori che una differenza possa andare in uno dei due sensi, comunque la significatività deve essere valutata con riferimento a valori alfa più piccoli (dimezzati) rispetto al test a due code – come è stato fatto nell'ES 4)

Dal campione si ottiene la stima puntuale $\hat{\pi} = 0.31$, il cui errore standard è $\sqrt{\frac{\hat{\pi}(1-\hat{\pi})}{n}} = 0.052$; la statistica test è pari a 4.10. Questo valore è ancora più lontano dal valore 0 atteso sotto H_0 di quanto non fosse nell'esercizio precedente. Il test rigetta l'ipotesi nulla al livello di confidenza del 5%, e comunque i dati mostrano una differenza fortemente significativa tra il valore osservato nel campione è quello di riferimento ($p < 0.0001$ – è una notazione adatta per quando i calcoli – che comunque sono arrotondati – danno $p=0$). Secondo questo studio, il trattamento aumenta l'incidenza della dermatite. Il RR che esprime l'effetto del trattamento confrontando il rischio di chi fa il trattamento con il rischio di chi non lo fa è $RR = 0.31/0.1 \approx 3$: il trattamento triplica il rischio di dermatite.

ES 6

Impostiamo il test fissando anzitutto le ipotesi di base e alternativa:

$$H_0: \delta = \mu_A - \mu_B = 0 \quad \text{vs} \quad H_1: \delta = \mu_A - \mu_B \neq 0$$

Questi due gruppi sono indipendenti in quanto i soggetti del gruppo A non sono gli stessi del gruppo B, e non condividono con essi fattori comuni (come potrebbe accadere se si trattasse di gruppi familiari, ad esempio di fratelli – che pur non vivendo insieme condividerebbero fattori genetici e fattori ambientali dell'infanzia – o coniugi – che condividerebbero fattori ambientali, abitudini etc)

Il confronto fra questi due gruppi che hanno numerosità ampie e provengono da distribuzioni con la stessa varianza (come afferma il testo dell'esercizio) può essere fatto applicando le formule del t-test:

$$s = \sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}} = \sqrt{\frac{(34 - 1)1.1^2 + (39 - 1)1.2^2}{34 + 39 - 2}} = 1.156$$

$$t = \frac{\bar{y}_1 - \bar{y}_2}{s \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} = \frac{3.5 - 2.9}{1.156 \sqrt{\frac{1}{34} + \frac{1}{39}}} = 2.22$$

Applicando il metodo della regione rifiuto, possiamo rigettare l'ipotesi di base di assenza di differenza fra la media di A e la media di B al livello di significatività del 5% (il valore soglia per il test bilaterale è 1.96).

Volendo valutare la significatività statistica calcolando il p-value, andiamo sulle tavole e in corrispondenza di 2.22 leggiamo 0.987; dunque l'area in una delle due code esterne è $1 - 0.987 = 0.013$, e la probabilità complessiva di andare in una delle due code esterne (il p-value del test bilaterale) è 0.026.

Dunque la differenza osservata fra le due medie risulta abbastanza significativa, i dati supportano l'ipotesi che la dieta A sia più efficace in termini di riduzione di peso. Questa associazione “statistica” può indicare un nesso di causalità SE i due gruppi sono simili per composizione – rispetto a tutte le caratteristiche potenzialmente influenti sulla riduzione del peso, sesso, età, attività fisica etc – ovvero in assenza di fattori di confondimento, e se anche le metodiche della misurazione sono le stesse nel gruppo A e nel gruppo B (assenza di bias da osservazione o altra forma di distorsione).

ES 7

Il (primo) problema di questo esercizio di confronto fra due gruppi è che essi (“prima” e “dopo”) sono appaiati, ovvero non si tratta di campioni indipendenti: pertanto non si può procedere come

nell'ES 6. Potremmo lavorare sulle differenze Dopo-Prima, e quindi su un campione di 7 osservazioni di questo nuovo carattere, che chiameremo Riduzione, e andare a testare le ipotesi:

$$H_0: \delta = 0 \quad \text{vs} \quad H_1: \delta \neq 0$$

con il t-test per campioni singoli (come nell'ES 4). L'ulteriore problema è che il campione è molto piccolo e non conosciamo nemmeno la sua varianza, quindi non possiamo usare il t-test nella sua forma più semplice.

Nel seguito quindi svolgiamo l'esercizio utilizzando le tavole della distribuzione T di Student (*nota: esse sono state distribuite agli Studenti, ma non sono contenute nel Formulario, quindi questo svolgimento viene proposto come completamento del corso, in quanto tale è oggetto d'esame all'orale, ma non allo scritto*)

I dati sono:

Soggetto	1	2	3	4	5	6	7
riduzione	18	-5	19	19	24	5	16

quindi la media osservata è 13.71, la deviazione standard osservata è 10.095, e la statistica test è:

$$t = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} = \frac{13.71 - 0}{10.095/\sqrt{7}} = 3.59$$

La regione di rifiuto per un test di livello alfa pari a 5% (bilaterale) è costituita dalle due code esterne dei valori minori e maggiori di $\pm$ il valore soglia che troviamo sulla tavola della distribuzione T con 6 gradi di libertà, che vale 2.447. La nostra statistica test cade nella regione di rifiuto, quindi H_0 è rigettata al livello del 5%: i dati supportano l'ipotesi che il trattamento induca una riduzione della pressione.

ES 8

Le percentuali di risposta si calcolano guardando ai profili-riga, ovvero alle distribuzioni della Risposta condizionate ai due diversi livelli del carattere Trattamento: $p_A = 37/122 = 0.3$, $p_B = 32/64 = 0.5$. Il Risk Ratio è $0.5/0.3 = 1.65$: la probabilità di risposta con B aumenta del 65% rispetto ad A.

La misura di associazione e il test richiesto sono quelli del Chi-quadrato. L'ipotesi di base H_0 è che non vi è associazione fra Trattamento e Risposta, l'ipotesi alternativa è che vi sia associazione, ossia in questo caso che vi sia una differenza fra le percentuali di risposta dei due gruppi A e B nella popolazione da cui essi provengono.

Il chi-Quadrato si basa su una misura di distanza fra le frequenze che abbiamo osservato e quelle che osserveremmo se i due campioni si comportassero come secondo l'ipotesi nulla:

freq. attese:

	Risposta	
Trattamento	no	si
A	76.74	45.26
B	40.26	23.74

differenze (addendi della statistica test):

	Risposta	
Trattamento	no	si
A	0.889	1.507
B	1.694	2.872

Chi-quadrato=6.962, superiore al valore soglia che per un test al livello del 5% avendo 1 grado di libertà (tabella 2x2) è 3.841. Quindi il test rigetta l'ipotesi nulla di assenza di differenza al livello di

significatività del 5%. Osserviamo che anche scegliendo un livello di significatività più basso pari a 0.01 – e quindi un test più prudente, che richiede maggiore “forza” dell’evidenza contraria ad H0 per poterla rigettare – abbiamo comunque il rifiuto di H0 (il valore soglia è 6.635).

ES 9

La misura di associazione che possiamo calcolare è il coefficiente di correlazione lineare: (usiamo le ‘formule rapide’ della covarianza e delle deviazioni standard)

	Altezza, X	Peso, Y	X*Y	X^2	Y^2
	165	55.2	9108	27225	3047.04
	157	48.1	7551.7	24649	2313.61
	168	65.6	11020.8	28224	4303.36
	178	72	12816	31684	5184
	171	73.5	12568.5	29241	5402.25
	182	77.9	14177.8	33124	6068.41
	182	80.3	14614.6	33124	6448.09
	153	49.4	7558.2	23409	2440.36
somma	1356	522	89415.6	230680	35207.12
somma/n	169.5	65.25	11176.95	28835	4400.89
			- media^2 * n/(n-1) sqrt	104.75 119.7143 10.9414	143.3275 163.8029 12.79855

Covarianza = $11176.95 - 169.5 \cdot 65.25 = 117.065$

Coeff. di Correlazione $r = 117.065 / (10.94 \cdot 12.80) = 0.836$

Il valore di r indica una forte associazione positiva fra i due caratteri.

Calcoliamo ora la pendenza e l’intercetta della retta di regressione:

$\text{beta} = 11176.95 / 119.7143 = 0.978$

$\text{intercetta} = 65.25 - 0.978 \cdot 169.5 = -100.513$

Il valore di peso predetto dal modello di regressione così stimato per un soggetto di altezza 165 è quindi:

$y^* = -100.513 + 0.978 \cdot 165 = 60.84$

Il modello ha sovrastimato il peso del soggetto id=1 (ma l’ordine di grandezza è abbastanza coerente; se fosse venuto molto diverso, avremmo doverosamente ricontrollato i calcoli ...)

ES 10

- Il trattamento ha effetto sul rischio di progressione?
 - Sì, ha un effetto fortemente significativo ($p < 0.001$), B dà meno rischio di progressione del trattamento A. Infatti il OR è inferiore a 1, il che indica che rispetto al baseline (gruppo di riferimento, qui A) il fattore è protettivo.
- Qual è l’effetto di un precedente trattamento?
 - L’aumento medio dell’odds di progressione è stimato pari a 44%, ma la stima è molto variabile, secondo questo modello potrebbe anche essere un fattore protettivo; questo fattore non è statisticamente significativo (questa risposta è sostanzialmente corretta)
- Qual è il fattore di rischio più importante?
 - E’ ~~la presenza dell’anomalia DEL13~~ ($p = 0.0005$) l’innalzamento delle microglobuline beta2 (poiché presenta l’OR più alto rispetto all’anomalia DEL13)
- Qual è l’effetto delle microglobuline Beta2 alte?

- L'odds di progressione aumenta circa del 36% (l'effetto è molto significativo) (questa risposta è sostanzialmente corretta)