alcuni esercizi
(unità e-f)

guido masarotto

24 maggio 2004

---

1. La seguente tabella di contingenza mostra come 319 studenti universitari di varie università si distribuiscono sulla base delle due variabili "Tipo di maturità" e "Numero di esami superati durante il 1° anno".

	Esami superati		
Maturità	$0-1$	$2-5$	>5
Classica	10	67	31
Scientifica	4	52	36
Altre	14	65	40

(a) Analizzare la dipendenza esistente tra le due variabili con gli strumenti a voi noti.

(b) Calcolare una stima della probabilità che, a prescindere dalla maturità di provenienza, uno studente faccia più di 5 esami. Calcolare inoltre un intervallo di confidenza per la medesima probabilità con un livello di copertura uguale a 0,95.

Schema di soluzione.

(a) Completiamo la tabella di contingenza con le distribuzioni marginali e calcoliamo le frequenze relative condizionate "di riga" (visto che quello che interessa è capire se il tipo di maturità influenza l'esito) e le frequenze attese in ipotesi di indipendenza.

— Tabella completata con le distribuzioni marginali

	Esami superati			
Maturità	$0-1$	$2-5$	>5	totale
Classica	10	67	31	108
Scientifica	4	52	36	92
Altre	14	65	40	119
totale	28	184	107	319

— Distribuzioni del numero di esami superati: marginale e condizionate al tipo di maturità (frequenze relative)

	Esami superati			
Maturità	$0-1$	$2-5$	>5	totale
Classica	0,09	0,62	0,29	1
Scientifica	0,04	0,57	0,39	1
Altre	0,12	0,55	0,34	1
totale	0,09	0,58	0,34	1

Commento: Gli studenti provenienti dallo Scientifico inclusi nel campione sembrano essere "andati" leggermente meglio degli altri (solo un 4% nella categoria $0-1$ esami contro una frequenza marginale del 9%, 39% nella categoria "con più di 5 esami" dove la frequenza marginale è del 34%). Gli studenti provenienti dal Classico vanno leggermente meglio di quelli provenienti da "altre" maturità per quanto riguarda la proporzione di studenti con più di 1 esame ma peggio per quanto concerne la proporzione di studenti "ottimi" (più di 5 esami).

— Frequenze attese in ipotesi di indipendenza

	Esami superati			
Maturità	$0-1$	$2-5$	>5	totale
Classica	9,48	62,29	36,23	108
Scientifica	8,08	53,07	30,86	92
Altre	10,45	68,64	39,92	119
totale	28	184	107	319

Commento. I commenti sono gli stessi di quelli fatti guardando alla precedente tabella. Ad esempio, guardando alla riga degli studenti provenienti dallo Scientifico vediamo che, se ci fosse stata indipendenza tra maturità e risultato al 1° anno, (i) ci saremmo aspettati di trovare 8 studenti con 1 o meno esami mentre ne abbiamo osservati solo 4 (cioè la metà) e (ii) viceversa, abbiamo trovato 36 studenti con più di 5 esami nel campione quando ce ne saremmo attesi 30/31.

La dipendenza appena descritta nel campione è una "manifestazione" di una situazione di dipendenza anche nella popolazione da cui il campione è stato estratto? Oppure la dipendenza osservata può essere più banalmente un semplice effetto del caso? Proviamo, per rispondere a queste domande, a considerare il seguente sistema d'ipotesi

$$\begin{cases} H_0 : \begin{pmatrix}\text{tra "tipo di maturità" e "numero di esami} \\ \text{al 1° anno" esiste indipendenza}\end{pmatrix} \\ H_1 : (\text{qualche forma di dipendenza esiste}) \end{cases}$$

La popolazione di riferimento di interesse è ovviamente qualcosa del tipo "tutti gli studenti universitari al primo anno dell'anno in cui è stata condotta l'indagine". Calcoliamo innanzitutto l'X^2 di Pearson

$$X^2 = \frac{(10-9{,}48)^2}{9{,}48} + \cdots + \frac{(40-39{,}92)^2}{39{,}92} = 5{,}48$$

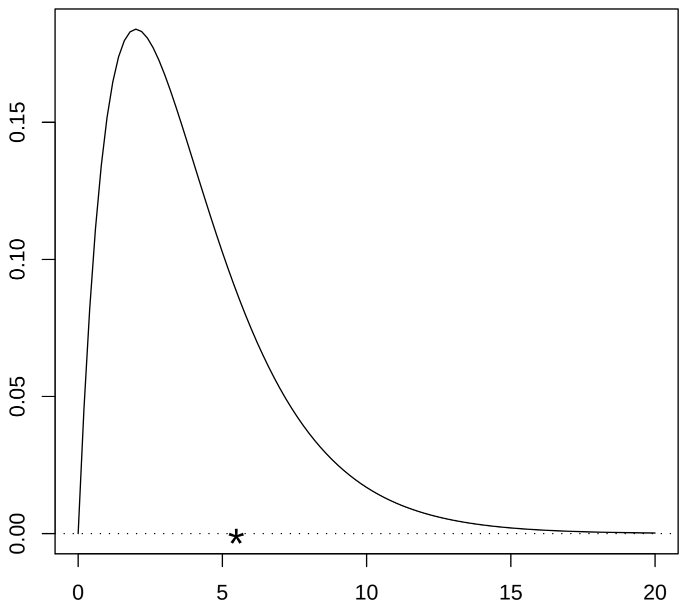

Figura 1: Densità di un χ^2 con 4 gradi di libertà. La stellina sull'asse delle ascisse indica il valore calcolato di X^2.

Questo valore sotto H_0 si distribuisce (approssimativamente) come un χ^2 di Pearson con 4 gradi di libertà[1]. Possiamo discutere i risultati in varie maniere alternative ma nella sostanza (e nelle conclusioni) equivalenti:

(i) da una tabella dei quantili di un χ^2 troviamo

$$\begin{aligned} \text{quantile-0,75 di un } \chi^2 \text{ con 4 gradi di libertà} &= 5{,}39 \\ \text{quantile-0,90 di un } \chi^2 \text{ con 4 gradi di libertà} &= 7{,}78 \end{aligned}$$

quindi il valore osservato (5,48) non è particolarmente grande rispetto ai valori attesi sotto l'ipotesi di indipendenza (si veda anche la figura 1).

(ii) dato che per questo test "lontano da H_0" equivale a "lontano da 0", dalle informazioni di prima sui quantili, troviamo

$$0{,}10 < (\text{livello di significatività osservato}) < 0{,}25.$$

[1] la più piccola della frequenze attese è circa 8 quindi possiamo fare riferimento a questo risultato asintotico.

Questi valori sono sufficientemente grandi per non "permetterci" di dubitare dell'ipotesi nulla [2].

(iii) pensando ad un approccio "accetto/rifiuto" dovremmo confrontare il valore osservato di X^2 con il quantili $1-\alpha$ (con α piccolo e scelto dall'utilizzatore) di un χ^2 con 4 gradi di libertà, e "rifiutare" l'ipotesi nulla se il valore calcolato dai dati è maggiore del percentile prescelto. Valori usuali per α e i relativi quantili $1-\alpha$ sono

$$\begin{aligned} \alpha = 0{,}10 &\Leftrightarrow \text{quantile} - 0{,}90 = 7{,}78 \\ \alpha = 0{,}05 &\Leftrightarrow \text{quantile} - 0{,}95 = 9{,}49 \\ \alpha = 0{,}01 &\Leftrightarrow \text{quantile} - 0{,}99 = 13{,}28 \end{aligned}$$

Il valore osservato è più piccolo di tutti questi quantili. Quindi, anche utilizzando questo approccio "accettiamo" H_0.

In conclusione: la dipendenza osservata nel campione di 319 studenti tra esito all'università e tipo di scuola superiore di provenienza potrebbe essere dovuta semplicemente al caso; sulla base di questi dati, non abbiamo elementi per pensare che nella popolazione di studenti da cui sono stati estratti gli studenti il "numero di esami superati" dipenda dalla "scuola media di provenienza".

(b) Senza tener conto della scuola di provenienza, abbiamo che $y = 107$ studenti tra gli $n = 319$ considerati hanno durante il primo anno superato più di 5 esami. Possiamo pensare di essere nel contesto di un campionamento binomiale[3] ovvero che

$$y \sim Bi(n, \vartheta)$$

dove ϑ è la probabilità di superare più di 5 esami al primo anno.

La stima di ϑ è

$$\hat{\vartheta} = \frac{y}{n} = \frac{107}{309} \approx 0{,}33.$$

Un intervallo di confidenza (probabilità che includa il vero ϑ approssimativamente uguale a $1-\alpha = 0{,}95$) è

$$\hat{\vartheta} \pm z_{1-\alpha/2}\sqrt{\frac{\hat{\vartheta}(1-\hat{\vartheta})}{n}} = 0{,}33 \pm 1{,}96\sqrt{\frac{0{,}33(1-0{,}33}{319}} \approx 0{,}33 \pm 0{,}05 = [0{,}28 - 0{,}38].$$

---

2. Tra gli effetti collaterali di un certo farmaco c'è il bruciore di stomaco. Durante uno studio, a 200 pazienti sono state somministrate pillole in cui il farmaco era mescolato con una sostanza tampone A, ad altri 100 pazienti pillole comprendenti la sostanza tampone B ed, infine, ad ulteriori 250 pazienti pillole

[2] il valore calcolato con R è

```
> 1-pchisq(5.48,4)
[1] 0.2414931
```

[3] il campionamento in situazioni di questo tipo è sempre "senza reinserimento" ma la popolazione di riferimento è composta da qualche centinaio di migliaia di studenti.

comprendenti la sostanza tampone C. La percentuale di pazienti che hanno manifestato bruciori di stomaco è stata rispettivamente 12% nel gruppo che aveva ricevuto A, 17% nel gruppo B e 14% nel gruppo C.

(a) Dire, utilizzando un appropriato test statistico, se le differenze osservate tra i tre gruppi possono essere attribuite al caso.

(b) Costruire un grafico che mostri per ogni sostanza tampone sia la stima che un intervallo di confidenza al 90% per la probabilità di avere dei bruciori.

Schema di soluzione.

(a) Dalle informazioni fornite nel testo del problema possiamo ricostruire la seguente tabella di contingenza

	sostanza tampone			
bruciore	A	B	C	totale
no	176	83	215	474
si	24	17	35	76
totale	200	100	250	550

La prima domanda chiede di verificare se, sulla base dei dati, ci aspettiamo che ci sia una differente capacità delle tre sostanze di "tamponare" gli effetti indesiderati del farmaco considerato. Differenti "probabilità di tamponare" equivalgono ovviamente alla dipendenza tra "somministrazione della sostanza tampone" e "presenza di bruciore". Possiamo perciò utilizzare il test X^2 di Pearson, calcolato a partire dalla tabella di sopra, per rispondere alla domanda:

(i) frequenze attese in ipotesi di indipendenza

	sostanza tampone			
bruciore	A	B	C	totale
no	172,36	86,18	215,45	474
si	27,64	13,82	34,55	76
totale	200	100	250	550

(i) il valore della statistica test è

$$X^2 = \frac{(176 - 172{,}36)^2}{172{,}36} + \cdots + \frac{(35 - 34{,}55)^2}{34{,}55} = 1{,}41;$$

(i) il valore osservato di X^2 va confrontato con i "valori prevedibili" per un χ^2 di Pearson con 2 gradi di libertà;

(i) in una tabella dei percentili di questa distribuzione troviamo che

$$1{,}41 \approx (\text{mediana di un } \chi^2 \text{ con 2 gradi di libertà}) = 1{,}39;$$

il valore osservato è perciò più o meno al centro dei valori che ci aspettiamo di osservare quando non esiste dipendenza[4].

In conclusione: i dati suggeriscono che non ci sono differenze di efficacia tra le tre sostanze tampone.

[4]alternativamente ed equivalentemente discutere il risultato utilizzando il livello di significatività osservato o via "accetto/rifiuto".

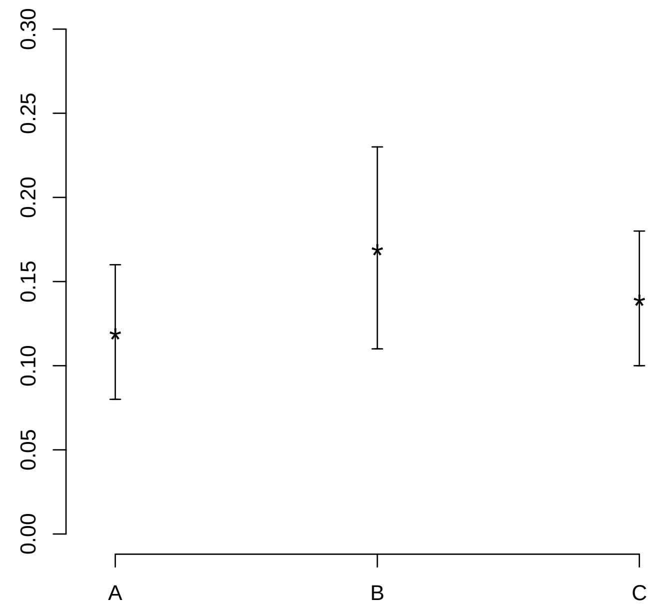

Figura 2: Tre sostanze tampone: intervalli di confidenza per le proporzioni di pazienti con bruciori

(b) Calcoliamo gli intervalli di confidenza facendo riferimento al procedimento visto per una distribuzione binomiale

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

osservando che in questo caso le stime ($\hat{\vartheta}$) e le numerosità (n) sono indicate separatamente per le tre sostanze tampone nel testo e che, sempre nel testo, è richiesto di porre

$$1 - \alpha = 0{,}9 \Rightarrow 1 - \alpha/2 = 0{,}95 \Rightarrow z_{1-\alpha/2} = 1{,}645.$$

Calcolo degli intevalli di confidenza:

sostanza A. $0{,}12 \pm 1{,}645\sqrt{0{,}12(1-0{,}12)/200} = [0{,}08 - 0{,}16]$;

sostanza B. $0{,}17 \pm 1{,}645\sqrt{0{,}17(1-0{,}17)/100} = [0{,}11 - 0{,}23]$;

sostanza C. $0{,}14 \pm 1{,}645\sqrt{0{,}14(1-0{,}14)/250} = [0{,}10 - 0{,}18]$;

Il grafico richiesto diventa quindi "qualcosa" simile a quello della figura b. Si osservi come gli intevalli di confidenza siano sovrapposti tra di loro. Questo di nuovo è una indicazione che i dati osservati nello studio sono compatibili con una efficacia uguale delle tre sostanze.

---

3. Una macchina dovrebbe produrre pezzi il cui peso nominale è di 50g. La media e la varianza dei pesi di 16 pezzi sono risultate uguali rispettivamente a 44g e 25g^2. Sapendo che gli scostamenti dei pesi effettivi dal peso medio si distribuiscono normalmente, dire se la macchina in questione è stata ben tarata. Rispondere sia mediante utilizzando un test statistico che mediante il calcolo di un intervallo di confidenza (diciamo al 99%).
[Nota: la varianza è stata calcolata "dividendo per $n-1$".]
Schema di soluzione. Siamo ovviamente nel *reame* del test t ad un campione. Il sistema di ipotesi da verificare è

$$\begin{cases} H_0 : \text{(la macchina produce pezzi mediamente di peso 50g)} \\ H_1 : (H_0 \text{ è falsa}). \end{cases}$$

Le osservazioni disponibili sono i pesi di $n = 16$ pezzi. I singoli pesi non li conosciamo ma ne conosciamo, e per il resto questo sarà sufficiente, la media e la varianza campionaria, indichiamoli con $\overline{y}$ e s^2.
L'usuale statistica test è

$$t_{oss} = \frac{\overline{y} - 50}{\sqrt{s^2/n}} = \frac{44-50}{\sqrt{25/16}} \approx -4{,}8.$$

Questo valore deve essere confrontato con i valori "previsti" da una t di Student con $n - 1 = 16 - 1 = 15$ gradi di libertà. Il valore osservato della statistica è, se considerato in valore assoluto, più grande dei quantili di questa distribuzione usualmente considerati come "soglie di accettazione" per il test. Infatti, se indichiamo con $t_{g;p}$ il quantile-p di una t con g gradi di libertà, troviamo

$$\begin{array}{lllll} \alpha = 0{,}1 & \Leftrightarrow & 1-\alpha/2 = 0{,}95 & \Leftrightarrow & t_{15;0{,}95} = 1{,}75 \\ \alpha = 0{,}05 & \Leftrightarrow & 1-\alpha/2 = 0{,}975 & \Leftrightarrow & t_{15;0{,}975} = 2{,}13 \\ \alpha = 0{,}01 & \Leftrightarrow & 1-\alpha/2 = 0{,}995 & \Leftrightarrow & t_{15;0{,}995} = 2{,}95. \end{array}$$

Utilizzando l'ultimo valore possiamo anche immediatamente dire che il livello di significatività osservato è minore di 0,01.
In conclusione: il valore osservato della statistica test è caduto in una regione dove non ci attendiamo che cada se è vera H_0. Quindi possiamo concludere che i dati suggeriscono che la macchina non sia stata ben tarata.
Alla stessa conclusione possiamo arrivare utilizzando un intervallo di confidenza per la media dei pesi dei pezzi prodotti dalla macchina quando è tarata come ora. L'intervallo (con una copertura pari a $1 - \alpha = 0{,}99$) è

$$\overline{y} \pm t_{n;1-\alpha/2}\sqrt{\frac{s^2}{n}} = 2{,}95\sqrt{\frac{25}{16}} \approx [40{,}3 - 47{,}7].$$

L'intervallo di confidenza non include il peso nominale di 50g. Quindi questo peso "non sembra plausibile" sulla base dei dati.

---

4. E' possibile dimostrare che se $y_1, \ldots, y_n$ sono n determinazioni indipendenti ed identicamente distribuite di una variabile casuale $N(\mu, \sigma^2)$ allora

$$\frac{(n-1)s^2}{\sigma^2} \sim \chi^2 \text{ con } n-1 \text{ gradi di libertà}$$

dove s^2 indica la varianza campionaria.
(a) Utilizzare questo risultato per costruire un test per verificare il sistema di ipotesi

$$\begin{cases} H_0 : \sigma^2 = \sigma_0^2 \\ H_1 : \sigma^2 \neq \sigma_0^2 \end{cases}$$

dove σ_0^2 indica un valore preassegnato per la varianza.
(b) Utilizzare il test costruito al punto precedente per verificare se è credibile che la varianza della distribuzione che ha generato i 5 spessori considerati nell'unità B sia 0,01.

Schema di soluzione.
(a) Si consideri la statistica

$$\xi = \frac{(n-1)s^2}{\sigma_0^2}.$$

Se H_0 è "vera" ci aspettiamo

$$s^2 \approx \sigma_0^2 \Leftrightarrow \frac{s^2}{\sigma_0^2} \approx 1,$$

e quindi che ξ "cada" intorno a $n-1$. Viceversa se H_0 è falsa ci aspettiamo che la distribuzione di ξ si sposti verso valori più [grandi,piccoli] di $n-1$ a seconda se σ^2 sia più [grande,piccolo] di σ_0^2. Possiamo quindi usare ξ come statistica test. Per quanto enunciato nel testo la distribuzione di ξ sotto l'ipotesi nulla è un χ^2 con $n-1$ gradi di libertà. Il valore calcolato dai dati di ξ dovrà quindi essere confrontato con i valori "attesi" per questa distribuzione. Ad esempio, volendo utilizzare un approccio del tipo accetto/rifiuto possiamo pensare di utilizzare la procedura stilizzata nella figura 3. Infatti, la procedura descritta garantisce che

$$\begin{aligned} &P(\text{accettare } H_0 \text{ quando } H_0 \text{ è vera}) = \\ &\quad = P(h \leqslant \xi \leqslant k \text{ quando } H_0 \text{ è vera}) = \\ &\quad = P(h \leqslant (\chi^2 \text{ con } n-1 \text{ gradi di libertà}) \leqslant k) = 1-\alpha. \end{aligned}$$

(b) Le misure da considerare sono

$$14{,}33 \quad 14{,}19 \quad 14{,}39 \quad 14{,}43 \quad 14{,}17.$$

In questo caso, n vale 5,

$$\sum_{i=1}^{n} y_i = 71{,}51 \qquad \sum_{i=1}^{n} y_i^2 = 1022{,}791$$

scegliere α

↓

determinare
$h = \text{quantile-}(\alpha/2)$ e
$k = \text{quantile-}(1-\alpha/2)$
di un χ^2 con $n-1$ gradi di libertà

↓

calcolare $\xi = \dfrac{(n-1)s^2}{\sigma_0^2}$

↓

verificare se
$h \leqslant \xi \leqslant k$

se la risposta è si → accettare H_0

se la risposta è no → rifiutare H_0

Figura 3: Test sulla varianza di una distribuzione normale

$$s^2 = \frac{1}{n-1}\left(\sum_{i=1}^{n} y_i^2 - \frac{1}{n}\left(\sum_{i=1}^{n} y_i\right)^2\right) = \frac{1}{4}\left(1022{,}791 - \frac{71{,}51^2}{5}\right) = 0{,}0137$$

e quindi

$$\xi = \frac{(n-1)s^2}{\sigma_0^2} = \frac{4 \times 0{,}0137}{0{,}01} = 5{,}48.$$

Scegliendo $\alpha = 0{,}05$ troviamo[5]

$$\begin{aligned} h &= \text{quantile-}0{,}025 \text{ di un } \chi^2 \text{ con 4 gradi di libertà} = 0{,}48 \\ h &= \text{quantile-}0{,}975 \text{ di un } \chi^2 \text{ con 4 gradi di libertà} = 11{,}14. \end{aligned}$$

Il valore di ξ calcolato dai dati cade ampiamente all'interno dell'intervallo definito da questi due valori. Concludiamo quindi a favore di H_0.

---

5. La sordità (di origine genetica) è purtroppo diffusa tra le gatte bianche europee. Uno zoologo sostiene che il carattere è più diffuso tra le gatte bianche con gli occhi arancioni che tra le gatte bianche con gli occhi celesti. A prova della sua affermazione riporta la seguente tabella contingenza

[5] la tabella dei quantili di di un χ^2 riportata nei lucidi è "troppo sintetica" per contenere questi valori. Usando R

```
> qchisq(0.025,4)
[1] 0.4844186
> qchisq(0.975,4)
[1] 11.14329
```


	colore degli occhi	
sorda	arancioni	celesti
si	17	11
no	8	14

(a) Vi sembra condivisibile l'affermazione dello zoologo?
(b) Determinare, separatamente per i due gruppi, un intervallo di confidenza al 95% per la frazione (nella popolazione di tutte le gatte bianche da cui le gatte utilizzate nello studio provengono) sorde.

Schema di soluzione. In sintesi:
(a) le differenze delle percentuali nel campione di gatte sorde tra i due colori sono molto grandi (68% tra le gatte con occhi arancioni, 44% tra le gatte con occhi celesti); se si calcola l'X^2 (= 2,92) e si confronta il valore trovato con la distribuzione attesa sotto H_0 (χ^2 con 1 grado di libertà) si conclude a favore di una dubbiosa accettazione (o di un dubbioso rifiuto) dell'ipotesi che la sordidà non dipenda dal colore degli occhi ($p - value \approx 0{,}08$ ad esempio); le differenze richiamate prima potrebbero perciò anche essere dovute al caso; probabilmente il nostro zoologo dovrebbe ripetere l'esperimento "guardando gli occhi a più gatte" (la numerosità è piccola e quindi gli errori attesi sono grandi);
(b) calcoliamo gli intervalli di confidenza facendo riferimento alla distribuzione binomiale; il risultato è per le gatte con gli occhi arancioni $\approx [0{,}48-0{,}83]$ e per le gatte gli occhi celesti $\approx [0{,}27-0{,}63]$; il commento è simile a prima: i dati indicano che le gatte con occhi celesti potrebbero essere meno esposte alla sordità; però non sono sufficienti per escludere la possibilità che la proporzione di gatte sorde nei due gruppi sia la stessa (ad esempio, i due intervalli di confidenza presentano dei valori comuni); questo è in parte legato alla scarsa numerosità campionaria (si guardi l'ampiezza degli intervalli di confidenza).

---

6. Per capire se l'ordine alla nascità tra due gemelli dipende dal peso, per dieci coppie di gemelli è stata calcolata la differenza (in Kg).

$$D = \begin{pmatrix}\text{peso alla nascità pri-}\\ \text{mo gemello nato}\end{pmatrix} - \begin{pmatrix}\text{peso alla nascità se-}\\ \text{condo gemello nato}\end{pmatrix}$$

La media e la varianza (calcolata dividendo per $n-1$) delle 10 differenze ottenute valgono rispettivamente 0,15 e 1,21. Supponendo che sia possibile assumere che la distribuzione di D sia normale, verificare l'ipotesi che l'ordine di nascità non è influenzato dal peso del gemello (condurre il test sia utilizzando un livello di significatività prefissato uguale a 0,90 che calcolando e commentando il livello di significatività osservato).

Schema di soluzione. La situazione è molto simile a quanto visto nell'unità in cui è stato presentato il test t ad un campione (dati sulle ore di *extra-sonno*). L'ipotesi nulla sotto verifica è

$$H_0: \text{l'ordine alla nascita non è influenzata dal peso}$$

che, "traduciamo" in

$$H_0 : \text{la media della distribuzione normale "che genera" D è nulla.}$$

La statistica test è

$$t_{oss} = \frac{\sqrt{n}\overline{y}}{s} = \frac{\sqrt{10} \times 0{,}15}{\sqrt{1{,}21}} \approx 0{,}43.$$

La distribuzione sotto H_0 è una t di Student con 9 gradi di libertà.
Volendo procedere con un approccio del tipo accetto/rifiuto con $1 - \alpha = 0{,}9$ dobbiamo per prima cosa determinare il percentile $0{,}95 = 1 - \alpha/2$ di questa distribuzione. Da una tabella dei quantili troviamo che vale 1,83. Poichè

$$-1{,}83 \leqslant t_{oss} \leqslant 1{,}83$$

concludiamo favore di H_0.
Il livello di significatività osservato in questo caso vale

$$2 \times P(|t \text{ con 9 gradi di libertà}| > 1{,}83).$$

Dalle tabelle troviamo che è maggiore di 0,5. Nessun "sospetto" quindi su H_0.

---

7. Il test di Shapiro-Wilk è stato applicato a due insiemi di dati numerici univariati, chiamiamoli l'insieme A e quello B. Il livello di significatività osservato è risultato $\approx 0{,}65$ per l'insieme A e $< 10^{-5}$ per l'insieme B.
(a) disegnare, separatamente per i due gruppi, dei *normal probability plot* compatibili con l'informazione data;
(b) disegnare, su di uno stesso piano cartesiano dei diagrammi a scatola con baffi compatibili con le informazioni formite.
Schema di soluzione. Il test di Shapiro-Wilk verifica l'ipotesi che la distribuzione di un certo insieme di dati sia normale di media e varianza arbitrarie. Quindi, l'informazione che viene data nel testo del problema è:
(i) il gruppo A ha una distribuzione normale (o almeno così sembra dai dati);
(ii) la distribuzione dei dati del gruppo B non è normale (o almeno così sembra dai dati).
I due grafici devono quindi riflettere questa informazione. Un esempio è mostrato nelle figure -. I due insiemi di dati tra l'altro danno luogo ai $p\text{-}value$ menzionati nel testo del problema. La non normalità nel gruppo B si nota dalla non linearità del *normal probability plot* e dalla asimmetria del diagramma a scatola con baffi.

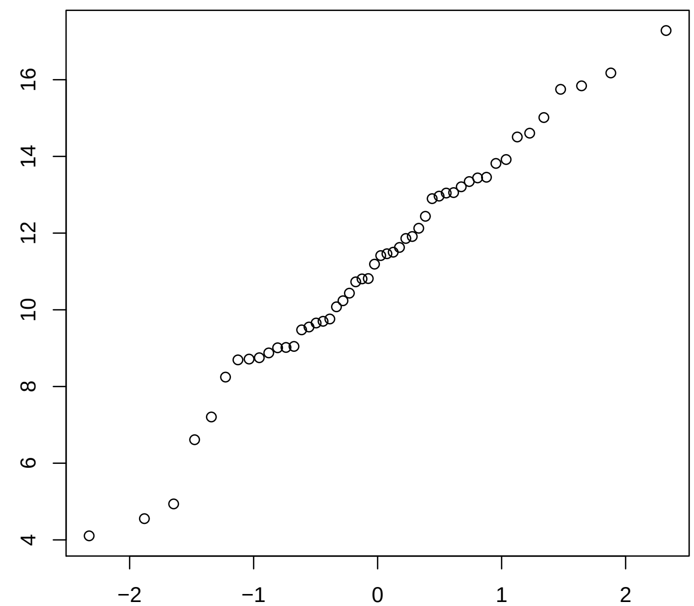

Figura 4: *Normal probability plot* di un ipotetico gruppo A

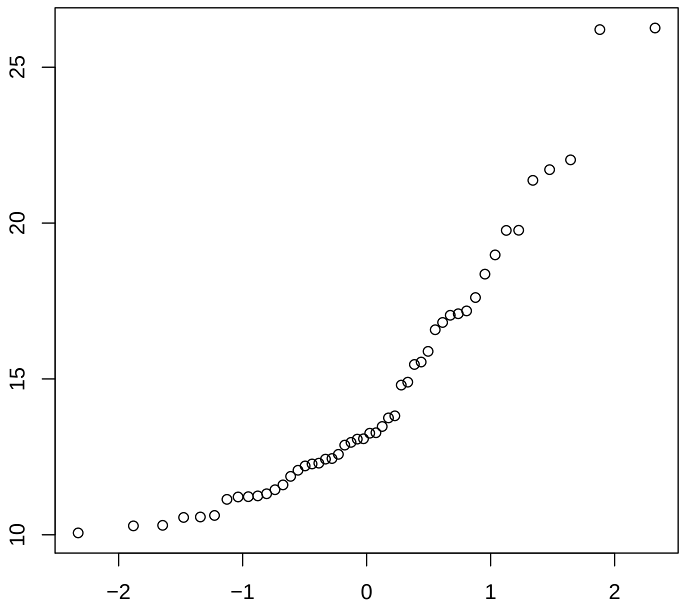

Figura 5: *Normal probability plot* di un ipotetico gruppo B

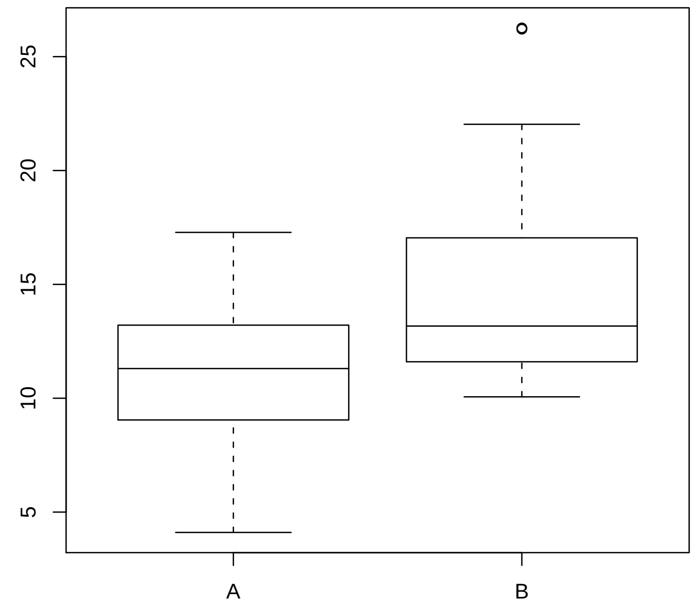

Figura 6: Diagramma a scatola con baffi dei dati del gruppo A e B