Capitolo 6

Statistica

108. Tra tutte le "parole" ottenute prendendo una sola volta 4 delle 6 lettere A, B, C, D, E ed F, la percentuale di quelle che *non* contengono vocali è

A. 7% circa

B. 67% circa

C. 33% circa

D. 40%

E. 75%

Statistica; conteggi.

Una *parola* di 4 lettere è una sequenza (ordinata, da sinistra verso destra) di 4 caselle:

$$\square\square\square\square$$

Il numero totale di parole ottenute prendendo una sola volta 4 delle 6 lettere A, B, C, D, E ed F è dato da

$$N = 6 \times 5 \times 4 \times 3$$

Infatti, abbiamo 6 modi di scegliere la prima lettera:

$$A\square\square\square,\quad B\square\square\square,\quad C\square\square\square,\quad D\square\square\square,\quad E\square\square\square,\quad F\square\square\square$$

e per ciascuna di queste scelte 5 modi di scegliere la seconda:

$$\begin{array}{lllll} AB\square\square, & AC\square\square, & AD\square\square, & AE\square\square, & AF\square\square \\ BA\square\square, & BC\square\square, & BD\square\square, & BE\square\square, & BF\square\square \end{array}$$

ecc.

e poi ancora 4 modi di scegliere la terza, e infine 3 modi di scegliere la quarta.

Le lettere che *non* sono vocali, tra queste 6, sono 4. Quindi le parole che *non* contengono vocali sono

$$n = 4 \times 3 \times 2 \times 1$$

in quanto abbiamo 4 modi di scegliere la prima lettera, ..., e 1 modo solo di scegliere la quarta.

Il rapporto tra parole non contenenti vocali e tutte le parole è quindi:

$$\frac{n}{N} = \frac{4 \times 3 \times 2 \times 1}{6 \times 5 \times 4 \times 3} = \frac{1}{15}$$

Dobbiamo confrontare il numero $\frac{1}{15}$ con le 5 percentuali proposte dalle risposte, che tradotte in frequenze relative (cioè in numeri tra 0 e 1) sono

A. $0,07$

B. $0,67$

C. $0,33$

D. $0,40$

E. $0,75$

Poiché $\frac{1}{15} < \frac{1}{10} = 0,1$ tutte le risposte tranne la A restano escluse, indicando valori più alti. Scegliamo quindi la risposta A.

Controlliamo che effettivamente

$$\frac{1}{15} \simeq 0,07 = \frac{7}{100}$$

Infatti, $7 \times 15 = 105 \simeq 100$. (Notiamo che la risposta A diceva 7% *circa*.)

☞ Il modo in cui abbiamo conteggiato il numero di parole (totale o senza vocali), al di là dei termini tecnici del calcolo combinatorio (disposizioni, permutazioni, ...) rientra nel principio generale con cui si conteggia il numero totale di scelte in uno *schema ad albero*.

> "se al passo 1 abbiamo n_1 opzioni, e per ciascuna di queste al passo 2 abbiamo n_2 opzioni, ..., e per ciascuna di queste all'ultimo passo k abbiamo n_k opzioni, il numero totale di scelte è il prodotto $n_1 \times n_2 \times \cdots \times n_k$"

Questo è un modo di ragionare utile per ricostruire varie situazioni combinatorie.

⚓ *Conteggio del numero totale di scelte in uno schema ad albero; traduzione di percentuali in frequenze relative e viceversa.*

---

109. Per trasmettere segnali Aldo issa 5 bandierine (3 gialle e 2 blu) su di un'asta. Quanti segnali diversi può ottenere Aldo?

A. 6

B. 5

C. 10

D. 25

E. 20

Statistica; conteggi.

Possiamo ragionare così: su 5 "posti" che ha l'asta, dobbiamo sceglierne esattamente 2 in cui mettere le bandierine blu (negli altri posti andranno allora le bandierine gialle).

Quanti sono i modi in cui scegliere 2 oggetti (i posti blu) su 5 (i posti disponibili)? Sono

$$\binom{5}{2} = \frac{5 \times 4}{2} = 10$$

Questo è chiaro se si conoscono i *coefficienti binomiali* e il concetto di *combinazioni semplici*; altrimenti il risultato si può ottenere ragionando secondo uno schema ad albero (v. commento al quesito n° 108): abbiamo 5 modi di scegliere il primo posto (b = blu)

$$b\square\square\square\square, \quad \square b\square\square\square, \quad \square\square b\square\square, \quad \square\square\square b\square, \quad \square\square\square\square b$$

e per ciascuno di questi ne abbiamo 4 per scegliere il secondo

$$\begin{array}{llll} bb\square\square\square, & b\square b\square\square, & b\square\square b\square, & b\square\square\square b \\ bb\square\square\square, & \square bb\square\square, & \square b\square b\square, & \square b\square\square b \\ \text{ecc.} & & & \end{array}$$

quindi in tutto 20; a questo punto però dobbiamo dividere per 2, perché – ad esempio – la scelta "1° e 2° posto" dà lo stesso risultato della scelta "2° e 1° posto" (come mostra chiaramente lo schema di sopra).

In definitiva, la risposta esatta è la C.

Combinazioni e coefficienti binomiali.

110. L'età media dei partecipanti ad una festa è di 24 anni. Se l'età media degli uomini è 28 anni e quella delle donne è 18 anni, qual è il rapporto tra il numero degli uomini e quello delle donne?

A. $\frac{3}{2}$

B. 2

C. $\frac{14}{9}$

D. $\frac{4}{3}$

E. $\frac{9}{14}$

Statistica; media.

Indichiamo con u e d il numero totale di uomini e donne, rispettivamente: si chiede di calcolare il rapporto u/d.

L'età media dei partecipanti alla festa non è semplicemente la media aritmetica tra l'età media degli uomini e l'età media delle donne, perché i due gruppi non sono ugualmente numerosi (come si vede dalle risposte); occorre invece fare una *media ponderata*

$$\text{età media dei partecipanti} = \frac{u \times 28 + d \times 18}{u + d}$$

Imponiamo che questa uguagli 24, come affermato nel quesito, e risolviamo

$$\begin{aligned} 24 &= \frac{u \times 28 + d \times 18}{u + d} \\ 24u + 24d &= 28u + 18d \\ 6d &= 4u \\ \frac{u}{d} &= \frac{6}{4} = \frac{3}{2} \end{aligned}$$

Il rapporto $\frac{u}{d}$ vale $\frac{3}{2}$, quindi la risposta esatta è la A.

Chi non avesse familiarità con il concetto di media ponderata potrebbe arrivarci così. Partiamo da

$$\begin{aligned} \text{età media dei partecipanti} &= \frac{\text{somma età dei partecipanti}}{\text{n° totale dei partecipanti}} \\ &= \frac{\text{somma età degli uomini} + \text{somma età delle donne}}{u + d} \end{aligned}$$

Ma

$$\text{età media degli uomini} = \frac{\text{somma età degli uomini}}{\text{n° totale degli uomini}}$$

e quindi

$$\text{somma età degli uomini} = 28 \times u$$

Analogamente

$$\text{somma età delle donne} = 18 \times d$$

Si ottiene così la formula della media ponderata, già vista prima.

Calcolo di medie e medie ponderate.

---

111. Ammettiamo che nella popolazione degli studenti del Politecnico i maschi (M) siano il doppio delle femmine (F). Allora, se si sorteggiano un primo numero di matricola e poi un secondo (potendo uno stesso numero di matricola essere sorteggiato due volte), le frequenze statistiche degli abbinamenti MM, MF, e FF sono

A. $\text{MM} = \frac{4}{9}$; $\text{MF} = \frac{4}{9}$; $\text{FF} = \frac{1}{9}$

B. tutte uguali

C. $\text{MM} = \frac{3}{6}$; $\text{MF} = \frac{2}{6}$; $\text{FF} = \frac{1}{6}$

D. $\text{MM} = \frac{5}{9}$; $\text{MF} = \frac{3}{9}$; $\text{FF} = \frac{1}{9}$

E. $\text{MM} = \frac{6}{9}$; $\text{MF} = \frac{2}{9}$; $\text{FF} = \frac{1}{9}$

Statistica; frequenze relative.

Indichiamo con m la frequenza relativa dei maschi e con f quella delle femmine. Dire che i maschi sono il doppio delle femmine, significa dire che sono i 2/3 del totale; quindi

$$m = \frac{2}{3}, \qquad f = \frac{1}{3}$$

Se si estraggono due numeri di matricola a caso, la frequenza relativa con cui si estraggono 2 maschi è data da

$$\text{frequenza relativa delle coppie } MM = \frac{\text{n° totale delle coppie } MM}{\text{n° totale delle coppie}}$$

Detto N è il numero complessivo degli studenti (maschi + femmine), la totalità dei possibili abbinamenti di studenti è N^2 e la totalità delle coppie MM è $\left(\frac{2}{3}N\right)^2$ di modo che

$$\text{frequenza relativa delle coppie } MM = \frac{\left(\frac{2}{3}N\right)^2}{N^2} = \frac{4}{9}$$

Analogamente, la frequenza relativa con cui si estraggono 2 femmine vale

$$\text{frequenza relativa delle coppie } FF = \frac{\left(\frac{1}{3}N\right)^2}{N^2} = \frac{1}{9}$$

Per differenza, la frequenza relativa con cui si estraggono un maschio e una femmina sarà

$$1 - \left(\frac{4}{9} + \frac{1}{9}\right) = \frac{4}{9}$$

Quindi la risposta esatta è la A.

☞ Per il calcolo del numero totale e dei numeri parziali di abbinamenti bisogna tenere conto delle modalità di sorteggio: così, nel nostro caso, i possibili abbinamenti sono in totale N^2 perché la *prima* matricola può essere estratta in N modi diversi, e anche la *seconda* matricola può essere estratta in N modi diversi (si parla in questo caso di *due estrazioni con reimmissione*). Se invece, ad es., il sorteggio avvenisse *in blocco* (cioè si estraggono in un solo colpo due matricole, per forza fra loro differenti), il totale dei possibili abbinamenti sarebbe $\frac{N(N-1)}{2}$ perché la prima matricola può essere estratta in N modi diversi, la seconda matricola in $N-1$ modi diversi, e poi bisogna dividere per 2 perché l'ordine di estrazione non conta.

⚓ *Frequenza relativa; calcolo della frequenza relativa di coppie di un certo tipo in una popolazione.*

---

112. La seguente tabella rappresenta la distribuzione dei redditi annuali (in migliaia di euro) di una certa collettività di persone

reddito	≤ 10	≤ 20	≤ 30	≤ 50	> 50
% di persone	28%	47%	73%	94%	6%

Se ne deduce che

A. le persone con reddito inferiore a 10 000€ sono meno di quelle che hanno un reddito superiore a 30 000€

B. le persone con reddito inferiore a 20 000€ sono tante quante quelle che hanno un reddito compreso fra 20 000€ e 50 000€

C. il 60% delle persone ha un reddito inferiore a 25 000€

D. il 20% delle persone ha un reddito superiore a 40 000€

E. nessuno ha un reddito di 5 000€

Statistica; interpretazione di tabelle di percentuali.

Si tratta di interpretare bene quanto scritto nella tabella. Occorre capire che le colonne sono tra loro *cumulative*: ad esempio, chi ha reddito ≤ 10 ha anche reddito ≤ 20, quindi il 28% di popolazione che compare nella prima colonna è una parte del 47% di popolazione che compare nella seconda. Ne segue che, ad esempio, per sapere quale percentuale ha reddito compreso tra 10 e 20 (cioè ≤ 20 *ma non* ≤ 10) dovremo calcolare la differenza $47\% - 28\% = 19\%$.

Con questo chiarimento iniziale, riscriviamo le 5 risposte calcolando, in base alla tabella, le percentuali delle classi a cui si riferiscono.

A. Le persone con reddito inferiore a 10 000€ $(= 28\%)$ sono meno di quelle che hanno un reddito superiore a 30 000€ $(= 100\% - 73\% = 17\%)$: Falso.

B. Le persone con reddito inferiore a 20 000€ $(= 47\%)$ sono tante quante quelle che hanno un reddito compreso fra 20 000€ e 50 000€ $(= 94\% - 47\% = 47\%)$: Vero.

C. Il 60% delle persone ha un reddito inferiore a 25 000€: la tabella non consente di dedurlo (il valore 25 000€ non compare in tabella), quindi la risposta C è falsa.

D. Il 20% delle persone ha un reddito superiore a 40 000€: la tabella non consente di dedurlo (il valore 40 000€ non compare in tabella), quindi la risposta D è falsa.

E. Nessuno ha un reddito di 5 000€: la tabella non consente di dedurlo (il valore 5 000€ non compare in tabella), quindi la risposta E è falsa.

La risposta esatta è la B.

Percentuali assolute e cumulative di classi.

---

113. Consideriamo la serie di 4 numeri:

$$S_1 = \{0; 2; 4; 10\}$$

e la serie S_2, pure di 4 numeri, ottenuta moltiplicando per 3 ciascun elemento di S_1. Siano v_1 la varianza dei numeri di S_1 e v_2 la varianza dei numeri di S_2. Allora

A. $v_2 = v_1^3$

B. $v_2 = 3v_1$

C. $v_2 = 9v_1$

D. $v_2 = v_1 + 9$

E. $v_2 = v_1 + 3$

Statistica; varianza.

La *varianza* di un insieme di numeri è la media dei quadrati delle differenze tra i singoli numeri e la loro media. Nel nostro caso la serie S_1 ha media

$$m_1 = \frac{0 + 2 + 4 + 10}{4} = 4$$

e quindi la sua varianza vale

$$\begin{aligned} v_1 &= \frac{1}{4}\left\{(0-4)^2 + (2-4)^2 + (4-4)^2 + (10-4)^2\right\} \\ &= \frac{16 + 4 + 0 + 36}{4} = \frac{56}{4} = 14 \end{aligned}$$

La serie S_2 si ottiene moltiplicando per 3 i numeri di S_1

$$S_2 = \{0; 6; 12; 30\}$$

La sua media è

$$m_2 = \frac{0 + 6 + 12 + 30}{4} = 12$$

e la sua varianza

$$\begin{aligned} v_2 &= \frac{1}{4}\left\{(0-12)^2+(6-12)^2+(12-12)^2+(30-12)^2\right\} \\ &= \frac{144+36+0+324}{4} = \frac{504}{4} = 126 \end{aligned}$$

Si constata subito che $126 = 9 \times 12$, e quindi la risposta esatta è la C.

Calcolo di medie e varianze.

---

114. Quanti sono i sottoinsiemi di 3 elementi contenuti in un dato insieme di 6 elementi?

A. 2

B. 9

C. 20

D. 40

E. 120

Statistica; conteggi.

Dato un insieme X avente 6 elementi, in quanti modi possiamo sceglierne 3? (Il numero dei modi in cui si può fare questa scelta sarà uguale al numero dei sottoinsiemi di X aventi 3 elementi). Chi ha studiato (e ricorda!) un po' di calcolo combinatorio, riconoscerà nel quesito una situazione tipica: il numero dei sottoinsiemi formati da 3 degli elementi di un insieme che ne contiene 6 in tutto, si calcola con il *coefficiente binomiale*

$$\binom{6}{3} = \frac{6 \times 5 \times 4}{3 \times 2} = 20$$

Chi non ha studiato (o non ricorda) queste cose, può arrivare alla risposta esatta ragionando secondo uno "schema ad albero" (v. quesito n° 108). Il primo elemento può essere scelto in 6 modi diversi; il secondo elemento può essere allora scelto tra i 5 rimanenti, e infine il terzo tra i 4 rimanenti. Complessivamente, questo dà

$$6 \times 5 \times 4 \qquad \text{scelte}$$

Così facendo, però, abbiamo contato più volte lo stesso sottoinsieme: infatti, indicando ad esempio con a, b, c tre elementi di X, è chiaro che le scelte $\{a, b, c\}$, $\{a, c, b\}$, $\{b, a, c\}$, ... individuano lo stesso insieme. Precisamente, in quanti modi possiamo *permutare* tra loro 3 elementi (cioè scambiarli di posto tra loro)? Per rispondere, ragioniamo al solito modo: al primo posto possiamo mettere un elemento qualsiasi scelto fra questi 3, al secondo posto uno scelto tra i 2 rimanenti, al terzo posto mettiamo l'oggetto rimasto; in totale le permutazioni di 3 oggetti sono 3×2. Quindi il numero di sottoinsiemi formati da 3 degli oggetti di X si ottiene dividendo il numero $6 \times 5 \times 4$ (scelte di 3 oggetti disposti in un certo ordine) per il numero 3×2 di modi in cui possiamo scambiare di posto i 3 oggetti scelti senza cambiare il sottoinsieme. Ritroviamo come risultato

$$\frac{6 \times 5 \times 4}{3 \times 2} = 20$$

Pertanto la risposta esatta è la C.

⚓ *Combinazioni e coefficienti binomiali; permutazioni.*

---

115. Se si lancia una moneta regolare, le possibilità che esca Testa oppure Croce sono le medesime: si dice allora che la frequenza con cui esce Testa (oppure Croce) è 1/2. Lanciando tre volte una moneta regolare, con quale frequenza usciranno 2 Teste e 1 Croce?

A. $\frac{1}{8}$

B. $\frac{3}{8}$

C. $\frac{1}{2}$

D. $\frac{1}{3}$

E. $\frac{2}{3}$

Statistica; conteggi, frequenze relative.

I possibili risultati di 3 lanci si possono rappresentare come sequenze di 3 esiti T (Testa) o C (Croce). Il numero totale di sequenze di questo tipo è $2 \times 2 \times 2 = 8$ (ad ognuno dei 3 lanci abbiamo 2 scelte), come si vede anche costruendo uno schema ad albero:

$$
\begin{array}{cccccccc}
 & & \text{T} & & & & \text{C} & \\
 & \swarrow & & \searrow & & \swarrow & & \searrow \\
 & \text{TT} & & \text{TC} & & \text{CT} & & \text{CC} \\
 \swarrow & \searrow & \swarrow & \searrow & \swarrow & \searrow & \swarrow & \searrow \\
\text{TTT} & \boxed{\text{TTC}} & \boxed{\text{TCT}} & \text{TCC} & \boxed{\text{CTT}} & \text{CTC} & \text{CCT} & \text{CCC}
\end{array}
$$

Di queste 8 sequenze, solo le tre evidenziate contengono 2 Teste e 1 Croce. Quindi la frequenza con cui usciranno 2 Teste e 1 Croce è data dal rapporto $\frac{3}{8}$.

La risposta esatta è la B.

Il quesito fa leva su due diversi ingredienti. Il primo è la capacità di *conteggiare* il numero di sequenze di un certo tipo (calcolo combinatorio, ragionamento su "schemi ad albero"); il secondo è invece la nozione di *frequenza* (relativa) di un certo tipo di esito, come rapporto tra numero di casi di un certo tipo e numero totale dei casi. Questo è un concetto di *statistica* elementare, che corrisponde anche al concetto di *probabilità* elementare come rapporto tra casi favorevoli e casi possibili. Avere i primi rudimenti di statistica o di probabilità, quindi, aiuta a individuare il ragionamento corretto.

Conteggio del numero di casi in uno schema ad albero; calcolo di frequenze relative.

---

116. Aldo ha superato 9 esami universitari; la media dei suoi voti (espressi in trentesimi) è 24. Qual è il voto *minimo* che Aldo dovrà prendere nel decimo esame affinché la sua media diventi almeno $24,5$?

A. 26

B. 27

C. 28

D. 29

E. 30

Statistica; media.

Sappiamo che

$$24 = \text{media dei primi 9 esami} = \frac{\text{somma dei punteggi dei primi 9 esami}}{9}$$

da cui ricaviamo che

$$\text{somma dei punteggi dei primi 9 esami} = 24 \times 9 = 216$$

Quindi, se indichiamo con x il voto del decimo esame, la somma dei punteggi dei primi 10 esami sarà $216 + x$, e

$$\text{media sui primi 10 esami} = \frac{\text{somma dei punteggi dei primi 10 esami}}{10} = \frac{216 + x}{10}$$

Se vogliamo che questa media sia $\geq 24,5$ dovrà essere

$$\begin{aligned} \frac{216 + x}{10} &\geq 24,5 \\ 216 + x &\geq 245 \\ x &\geq 245 - 216 = 29 \end{aligned}$$

Aldo quindi deve prendere almeno 29. Sarà bene che si dia da fare! Per intanto, la risposta esatta è la D.

Calcolo di medie.

---

117. Il diagramma seguente riporta i dati ufficiali (in valori assoluti) delle assunzioni annuali di ingegneri in Italia nel periodo 2001-2005. Prendendo come base di riferimento (= 100%) il 2001, dal diagramma si può ricavare che le percentuali di assunzioni negli anni successivi sono

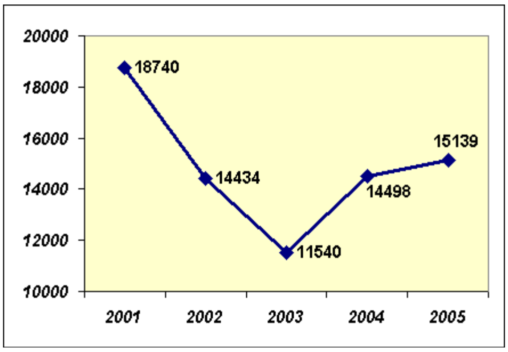
A. $2002 = 77,4\%$; $2003 = 61,6\%$; $2004 = 77,0\%$; $2005 = 80,8\%$

B. $2002 = 61,6\%$; $2003 = 77,0\%$; $2004 = 80,8\%$; $2005 = 77,4\%$

C. $2002 = 77,0\%$; $2003 = 61,6\%$; $2004 = 80,8\%$; $2005 = 77,4\%$

D. $2002 = 77,4\%$; $2003 = 61,6\%$; $2004 = 77,0\%$; $2005 = 80,8\%$

E. $2002 = 77,0\%$; $2003 = 61,6\%$; $2004 = 77,4\%$; $2005 = 80,8\%$



Statistica; interpretazione di diagrammi.

Il quesito chiede di trasformare i valori assoluti riportati nel diagramma in percentuali. I valori assoluti sono

2001	2002	2003	2004	2005
18 740	14 434	11 540	14 498	15 139

e quindi i valori percentuali (posto 100 il valore del 2001) si ottengono con il calcolo

2001	2002	2003	2004	2005
100	$\frac{14\,434}{18\,740} \times 100$	$\frac{11\,540}{18\,740} \times 100$	$\frac{14\,498}{18\,740} \times 100$	$\frac{15\,139}{18\,740} \times 100$

Senza una calcolatrice, però, il calcolo esatto non è agevole, per cui conviene cercare la risposta esatta tra quelle proposte con un confronto più qualitativo che quantitativo.
Cominciamo a notare che l'andamento nel passare da un anno al successivo è: diminuzione; diminuzione; aumento; aumento. Questo è vero per A, D, E, ma non per B e C (in entrambe, l'ultimo valore è più basso del penultimo). Quindi scartiamo B e C.
Confrontiamo i valori assoluti dei 5 anni con le percentuali fornite dalle risposte A, D, E:

	2001	2002	2003	2004	2005
	18 740	14 434	11 540	14 498	15 139
A	100	77,4	61,6	77,0	80,8
D	100	77,4	61,6	77,0	80,8
E	100	77,0	61,6	77,4	80,8

Osserviamo i valori degli anni 2002 e 2004: in valore assoluto, il 2004 ha un valore più alto del 2002, quindi questo dev'essere vero anche in percentuale. Le risposte A e D contraddicono questo fatto, e vanno scartate. Per esclusione, la risposta esatta è la E.

Calcolo di percentuali.