

## Glossario per il lettore di un articolo scientifico. Parte I: la statistica descrittiva

Luisa Zanolla, Maria Stella Graziani

Azienda Ospedaliera Universitaria Integrata di Verona

### ABSTRACT

**Glossary for the reader of a scientific paper. Part I: descriptive statistics.** The medical literature uses statistics to describe results of observational studies, clinical trials and experimental studies and to test the underlying scientific hypothesis. The statistic language could, however, not be familiar to any reader; this can hamper the correct comprehension of the paper and the appropriate interpretation of results. These considerations represented the spur to formulate a glossary of the more common statistical terms. To facilitate the reading of the vast majority of articles, which are written in English, corresponding English terms are also reported. The present document represents the first part of the project, related to descriptive statistics.

### INTRODUZIONE

La letteratura medica usa la metodologia statistica per descrivere i risultati di studi osservazionali o di studi clinici o degli esperimenti scientifici in genere e saggiare le relative ipotesi scientifiche. Il linguaggio della statistica, tuttavia, non è a tutti familiare e tale carenza può rendere difficile la lettura di un lavoro scientifico o, peggio, non consentirne un'interpretazione corretta e critica. A partire da tali considerazioni si è voluto raccogliere un glossario dei termini statistici più di frequente utilizzati; per facilitare il lettore viene riportato per ciascuno anche il corrispondente termine inglese. Le voci considerate nel glossario, italiane e inglesi, sono elencate in Appendice con accanto le rispettive traduzioni. Iniziamo qui con la statistica descrittiva; seguiranno poi la statistica inferenziale e le meta-analisi.

### ASIMMETRIA ("SKEWNESS")

Indice di forma di una distribuzione: una distribuzione si definisce asimmetrica quando nessun valore può dividerla in due parti simmetriche. Si parla di **asimmetria positiva** se la coda più lunga è a destra, cioè vi sono relativamente pochi valori elevati; di **asimmetria negativa** se si ha una coda più lunga a sinistra, cioè vi sono relativamente pochi valori bassi.

Esempi di asimmetria positiva possono essere le distribuzioni nella popolazione generale dei valori dei marcatori di danno miocardico, quali troponine o peptide

natriuretico di tipo B (pochi soggetti con valori elevati); esempio di asimmetria negativa è la distribuzione della concentrazione di glucosio plasmatico in una popolazione a forte prevalenza di diabete (tante persone con valori elevati).

Vi sono numerose misure di asimmetria, tra cui il coefficiente di asimmetria  $\beta_1$  di Pearson, che è il rapporto tra il quadrato del momento centrale (vedi "Momento") di terzo ordine e il cubo del momento centrale di secondo ordine:

$$\beta_1 = \frac{m_3^2}{m_2^3}$$

E' utilizzato anche l'indice di asimmetria  $\gamma_1$  di Fisher:

$$\gamma_1 = \sqrt{\beta_1} = \sqrt{\frac{m_3^2}{m_2^3}} = \frac{m_3}{m_2^{3/2}}$$

Il valore di  $\gamma_1$  è <0 in caso di asimmetria negativa, >0 in caso di asimmetria positiva.

### CAMPIONE ("SAMPLE")

Un insieme di unità sperimentali estratte da una popolazione (vedi), secondo una procedura di tipo probabilistico.

Ad esempio, un campione della popolazione dei

Corrispondenza a: Maria Stella Graziani, Vicolo S Giovanni in Foro 5, 37121 Verona. Tel. 3475201101, E-mail mariastella@graziani.eu

Ricevuto: 26.05.2013

Revisionato: 09.08.2013

Accettato: 10.09.2013

pazienti diabetici può essere costituito dai pazienti diabetici che vivono in una data provincia o che sono seguiti da un dato Centro. In questi esempi, tuttavia, si tratta di campioni non casuali. Un campione casuale sarà estratto con tecniche affidate al caso.

### CAMPO DI VARIAZIONE (“RANGE”)

Misura di dispersione (vedi) costituita dalla differenza tra il valore minimo e il valore massimo che una variabile può assumere.

Ad esempio, il peso minimo di un insieme di pazienti adulti con diabete è 45 kg, il massimo è 135 kg: il campo di variazione è  $90 \text{ kg} = 135 - 45$ .

E' evidente come questa misura di variabilità sia estremamente sensibile a valori estremi; un solo valore aberrante può farla aumentare.

Ad esempio, se all'insieme precedente di pazienti adulti con diabete si aggiunge un solo paziente che pesa 180 kg, il campo di variazione diviene  $135 \text{ kg} = 180 - 45$ .

A causa di questa sensibilità del campo di variazione ai valori estremi, come misura di dispersione viene spesso utilizzata in alternativa la Differenza interquartile (vedi).

### COEFFICIENTE DI VARIAZIONE (“COEFFICIENT OF VARIATION”)

È la DS (vedi) espressa come percentuale della media, ottenuta dividendo per il valore assoluto della media e moltiplicando per 100. A differenza della DS, che è espressa nella stessa unità di misura della variabile originale, il CV è adimensionale e consente di dare un giudizio sulla “grandezza” della DS di un gruppo di variabili, indipendentemente dalla scala con cui esse sono state misurate.

Ad esempio, se la media della statura è 172 cm in una data popolazione e la DS 5,2 cm, il CV è  $[5,2/172 \cdot 100] = 3,0\%$ , mentre se la media del peso corporeo è 75 kg e la DS 10,2 kg, il CV è  $[10,2/75 \cdot 100] = 13,6\%$ . Possiamo concludere che il peso corporeo ha una variabilità maggiore della statura.

### CURTOSI (“KURTOSIS”)

Indice di forma di una distribuzione; descrive l'appiattimento della curva di distribuzione. Si dice **leptocurtica** una distribuzione con eccesso di frequenze nelle classi centrali (più “appuntita”) e **platicurtica** quella con difetto di frequenze nelle classi centrali (più “piatta”).

La curtosi è misurata dal coefficiente di curtosi,  $\beta_2$ ; il coefficiente è il rapporto tra il momento centrale (vedi “Momento”) di quarto ordine e il quadrato del momento centrale di secondo ordine:

$$\beta_2 = \frac{m_4}{m_2^2} - 3$$

$\beta_2$  è una quantità adimensionale: se è  $>0$ , la curva è leptocurtica, se è  $<0$ , la curva è platicurtica. Per una

distribuzione gaussiana  $\beta_2$  è pari a 3.

### DEVIANZA (“SUM OF SQUARES”)

Misura di dispersione (vedi) costituita dalla somma dei quadrati degli scarti delle singole osservazioni dalla media aritmetica (vedi).

$$\sum_{i=1}^n (x_i - \bar{x})^2$$

Ad esempio, la media di 5, 11, 9, 13, 7 è 9. Il calcolo della devianza risulta:

$$\begin{aligned} SS &= (5-9)^2 + (11-9)^2 + (9-9)^2 + (13-9)^2 + (7-9)^2 = \\ &= 16+4+0+16+4 = 40 \end{aligned}$$

### DEVIAZIONE STANDARD (“STANDARD DEVIATION”)

Detta anche scarto quadratico medio, misura di dispersione (vedi) costituita dalla radice quadrata della varianza (vedi), cioè la radice quadrata della devianza (vedi) divisa per i gradi di libertà (vedi) (1, 2). Mentre la varianza è in unità di misura al quadrato rispetto a quella della variabile originale, la DS si riporta nella scala di misura della variabile originale, per cui è la misura di dispersione comunemente usata per la descrizione di una variabile

Ad esempio, la media di 5, 11, 9, 13, 7 è 9. Il calcolo della devianza risulta 40 (vedi esempio in “Devianza”) che diviso per i gradi di libertà  $(n-1)=4$  dà:  $40/4=10$ , la cui radice quadrata è 3,16.

E' largamente diffuso l'uso di presentare i risultati come: media  $\pm$  1 DS. Benché questo formato sia prevalente, alcune riviste, prevalentemente anglosassoni, utilizzano il formato media (deviazione standard).

### DIFFERENZA INTERQUARTILE (“INTERQUARTILE DISTANCE”)

Misura di dispersione costituita dall'intervallo tra il 25° e il 75° percentile (vedi “Percentile”). Viene utilizzata come misura di dispersione in alternativa al “Campo di variazione” (vedi).

### DISTRIBUZIONE DI FREQUENZA (“FREQUENCY DISTRIBUTION”)

Una rappresentazione nella quale al valore che una variabile può assumere viene associata la frequenza con la quale esso si presenta.

Nel caso di variabile discreta (vedi) si può tabulare la frequenza in corrispondenza di ogni valore della variabile. Nel caso di variabile continua (vedi) o di ampio intervallo di valori possibili, si suddivide l'intervallo di variabilità in intervalli (usualmente di ampiezza uguale) e si tabula la frequenza dei valori nell'intervallo in corrispondenza del suo valore centrale. La

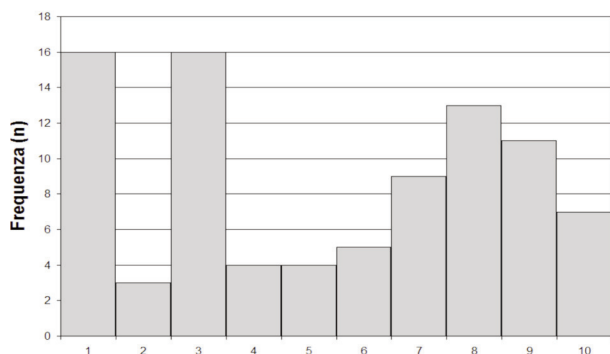


Figura 1

Rappresentazione grafica della distribuzione di frequenza: diagramma di frequenza.

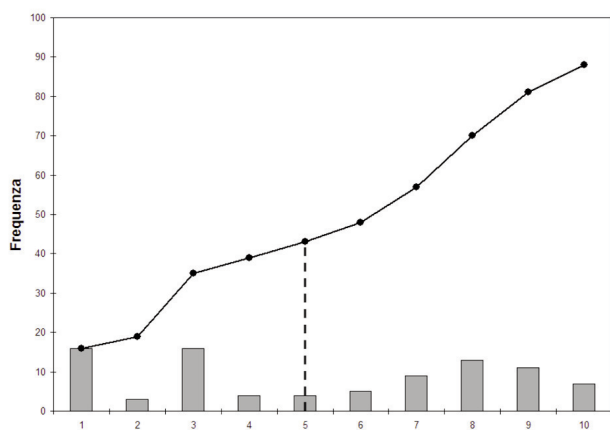


Figura 2

Rappresentazione grafica della frequenza cumulativa: le barre rappresentano la frequenza, la linea la frequenza cumulativa, cioè la somma dei valori fino a quel punto. La linea tratteggiata verticale in corrispondenza del valore 5 individua sull'asse y la frequenza con cui la variabile presenta un valore non superiore a 5.

rappresentazione grafica della distribuzione di frequenza è il **diagramma di frequenza** (Figura 1).

### ERRORE STANDARD DELLA MEDIA (“STANDARD ERROR OF THE MEAN”)

Misura di dispersione (vedi) della distribuzione della media campionaria (1). Viene usato nel calcolo dell'intervallo di confidenza della media.

Non di rado è usata in modo inappropriato in luogo della DS (vedi) in quanto il valore è più piccolo, essendo pari alla DS divisa per la radice quadrata del numero delle osservazioni:

$$s_x = \frac{s}{\sqrt{n}}$$

### FREQUENZA ASSOLUTA (“ABSOLUTE FREQUENCY”)

È il numero di volte in cui una modalità di un carattere compare in un insieme di osservazioni.

Ad esempio, se su 80 persone si rileva che 4 hanno il gruppo sanguigno AB, 4 è la frequenza assoluta.

### FREQUENZA CUMULATIVA (“CUMULATIVE FREQUENCY DISTRIBUTION”)

È la frequenza con cui una variabile presenta un valore non superiore a una soglia definita (Figura 2).

### FREQUENZA RELATIVA (“RELATIVE FREQUENCY”)

È il rapporto tra la frequenza assoluta (vedi) di un valore di un carattere e il numero totale di osservazioni.

Ad esempio, se su 80 persone si rileva che 4 hanno il gruppo sanguigno AB, la frequenza relativa è  $4/80 = 0,05$ .

### GRADI DI LIBERTÀ (“DEGREES OF FREEDOM”)

È il numero degli scarti che risulta indipendente dagli altri scarti. Valga come esempio la media, per la quale per definizione la somma degli scarti è pari a zero: se sostituiamo a ogni valore la sua differenza dalla media, la somma del risultato è sempre zero. Tutti gli scarti sono indipendenti l'uno dall'altro, tranne l'ultimo, il cui valore dovrà essere necessariamente quel numero il cui scarto dalla media, sommato a tutti gli altri, dia zero. In questo caso, quindi, i gradi di libertà sono  $(n-1)$ . Come regola generale il numero di gradi di libertà è pari al numero degli scarti, diminuito del numero delle statistiche stimate sul campione e utilizzate per il calcolo degli scarti stessi.

Ad esempio, per il calcolo della devianza, il numero degli scarti è  $n$ ; per calcolarli è però necessario determinare la media, quindi gli scarti indipendenti risultano  $(n-1)$ .

### INTERVALLO DI CONFIDENZA (“CONFIDENCE INTERVAL”)

Calcolato per un parametro (vedi), rappresenta l'intervallo di valori intorno allo stimatore (vedi) del parametro che include il valore vero del parametro stesso (che rimane ignoto), a un dato livello di probabilità (3).

Ad esempio, se l'intervallo di confidenza a un dato livello di probabilità  $(1-\alpha)$ , quasi sempre il 95%, di una media è 7,5-18,5, ciò significa che tale intervallo ha una probabilità pari a  $(1-\alpha)$  di contenere il parametro (vedi) vero della popolazione.

La formula per il calcolo dell'intervallo di confidenza varia al variare dello stimatore considerato.

### MEDIA ARITMETICA (“MEAN”, “ARITHMETIC MEAN”)

Il più delle volte indicata solo come media, è la misura di tendenza centrale (vedi) più utilizzata ed è costituita dalla somma delle osservazioni divisa per il loro numero (2):

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{n}$$

La media di una variabile è spesso indicata con la lettera che descrive la variabile soprassegnata; ad es., la media della variabile  $x$  è indicata come  $\bar{x}$

Ad esempio, la media di 5, 11, 9, 13, 7 è:

$$\bar{x} = \frac{5+11+9+13+7}{5} = \frac{45}{5} = 9$$

### MEDIA ARMONICA (“HARMONIC MEAN”)

È il reciproco della media aritmetica del reciproco dei dati. La media armonica viene usata quando si voglia calcolare la media di una variabile ottenuta come reciproco di un'altra:

$$\bar{x} = \frac{n}{\sum_{i=1}^n \frac{1}{x_i}}$$

Un esempio può essere costituito dalla velocità di eritrosedimentazione (VES), che è misurata in mm/ora. Sia la VES di quattro pazienti 8, 11, 13, 60. La media armonica è:

$$\bar{x} = \frac{4}{\frac{1}{8} + \frac{1}{11} + \frac{1}{13} + \frac{1}{60}} = \frac{4}{0.309} = 12.9$$

Benché la media armonica sia regolarmente descritta nei testi di statistica, il suo utilizzo effettivo è molto raro.

### MEDIA GEOMETRICA (“GEOMETRIC MEAN”)

È la radice  $n$ -esima del prodotto degli  $n$  valori assunti da una variabile.

$$\sqrt[n]{\prod_{i=1}^n x_i}$$

La media geometrica viene usata per le grandezze biologiche misurate in scala di diluizione a raddoppio, quali i titoli anticorpali del siero o i dosaggi farmacologici. Siano, ad esempio, i titoli anticorpali di cinque pazienti 1:40, 1:40, 1:80, 1:80, 1:320. La media geometrica è:

$$\sqrt[5]{40 \cdot 40 \cdot 80 \cdot 80 \cdot 320} = 80$$

### MEDIANA (“MEDIAN”)

È la misura di tendenza centrale (vedi) che si ottiene disponendo le osservazioni in ordine crescente (o decrescente): la mediana è l'osservazione centrale della sequenza ordinata. Se il numero di osservazioni è pari, la mediana è la media aritmetica delle due osservazioni centrali. La mediana corrisponde al 50° percentile.

Ad esempio, per ottenere la mediana di 5, 11, 9, 13, 7 si ordinano i dati in ordine crescente:

5, 7, **9**, 11, 13

Il valore centrale, corrispondente a 9, è la mediana. Nel caso si abbia un numero pari di osservazioni:

5, 7, **9**, **11**, 13, 17

la mediana è la media dei due valori centrali, 9 e 11, quindi corrisponde a 10.

### MISURA DI DISPERSIONE (“MEASURE OF VARIATION”)

Valore numerico che fornisce informazioni sulla variabilità dei dati. Esempi di misure di dispersione sono: campo di variazione, differenza interquartile, varianza, DS.

### MISURA DI TENDENZA CENTRALE (“MEASURE OF CENTRAL TENDENCY”)

Valore numerico che fornisce informazioni sulla posizione centrale di una distribuzione di frequenza e quindi esprime una caratteristica comune a tutti i dati che contribuiscono a determinarne il valore. Le misure comunemente utilizzate di tendenza centrale sono la media aritmetica, la mediana e la moda.

### MODA (“MODE”)

Misura di tendenza centrale (vedi) costituita dalla modalità di una variabile che si presenta nella distribuzione con maggiore frequenza (Figura 3).

Una distribuzione che presenti una sola moda viene detta **unimodale**; quando le mode siano due, come nella Figura 3 (in corrispondenza dei valori 3 e 8), si parla di distribuzione **bimodale**.

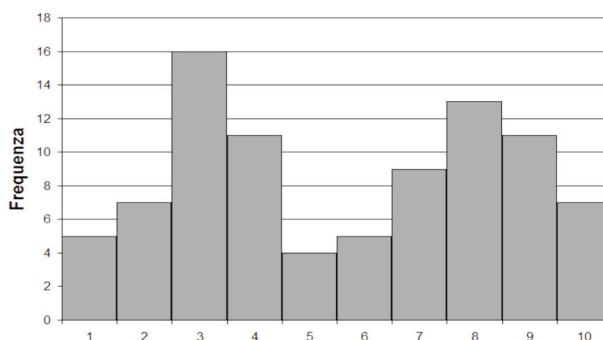


Figura 3

Distribuzione di frequenza nella quale la modalità della variabile che si presenta con maggiore frequenza è il 3, che pertanto rappresenta la “moda” di questa distribuzione.

**Tabella 1***Alcune corrispondenze fra parametro e stimatore*

Parametro (valore vero)	Simbolo	Stimatore (valore campionario)	Simbolo
Media della popolazione	$\mu$	Media campionaria	$\bar{x}$
Varianza	$\sigma^2$	Varianza campionaria	$s^2$
Deviazione standard	$\sigma$	Deviazione standard campionaria	$s$

**MOMENTO (“MOMENT”)**

I momenti dalla media aritmetica di ordine  $k$  sono definiti come:

$$m_k = \frac{1}{n} \cdot \sum_{i=1}^n (x_i - \bar{x})^k$$

per  $k=2$  il momento di secondo ordine corrisponde alla varianza.

I momenti di secondo, terzo e quarto ordine sono utilizzati per il calcolo del coefficiente di asimmetria (vedi “Asimmetria”) e di curtosi (vedi).

Ad esempio, la media di 7, 10, 9, 11, 18 è 11. Il corrispondente momento di terzo ordine è:

$$m_3 = \frac{1}{5} \cdot [(7-11)^3 + (10-11)^3 + (9-11)^3 + (11-11)^3 + (18-11)^3] = \\ = \frac{1}{5} \cdot 270 = 54$$

**PARAMETRO (“PARAMETER”)**

Caratteristica tipica e costante di una data popolazione (vedi) che ha, in genere, valore ignoto. I parametri sono anche detti costanti o valori veri per distinguerli dai valori campionari che sono invece variabili.

I parametri sono usualmente indicati con lettere greche, mentre gli stimatori corrispondenti (vedi “Stimatore”) sono indicati con le lettere latine. Alcuni esempi sono forniti nella Tabella 1.

**PERCENTILE (“PERCENTILE”)**

Disponendo le osservazioni in ordine crescente il percentile  $p$ -esimo è quel valore che lascia al di sotto di sé il  $p\%$  delle osservazioni e al di sopra il  $(100-p)\%$  (4).

Ad esempio, il 95° percentile lascia al di sotto di sé il 95% delle osservazioni ordinate e al di sopra il 5%.

**POPOLAZIONE (“POPULATION”)**

L'insieme esaustivo di unità sperimentali sulle quali si intende studiare un fenomeno. Nelle scienze biologiche si tratta spesso di un insieme potenzialmente infinito di individui.

Si immagini di voler studiare una variabile biologica nei pazienti diabetici; quale sia la definizione, ovviamente precisa e univoca, che si attribuisce alla diagnosi di diabete, la popolazione nella sua totalità non è accessibile

al ricercatore, che ricorrerà a un campione (vedi).

**PROPORZIONE (“PROPORTION”)**

In una variabile qualitativa o ordinale, o in una variabile quantitativa discreta, il rapporto tra il numero di volte in cui la variabile assume un dato valore e il numero totale di osservazioni.

Ad esempio, se su 30 donatori, 3 hanno il gruppo sanguigno B, la proporzione di donatori con gruppo sanguigno B è  $3/90 = 0,033$ .

**QUARTILI (“QUARTILES”)**

Il 25° e il 75° percentile (vedi “Percentile”)

**RANGHI (“RANKS”)**

Trasformazione dei dati che si applica ai dati ordinati, dal minore al maggiore (o viceversa): a ciascuno di essi viene assegnato un numero di rango, cioè al primo viene assegnato il rango 1, al secondo il rango 2, ecc.

Se la serie di dati è costituita da 5, 11, 9, 13, 7, 17 i ranghi si assegnano come:

Variabile:	5	7	9	11	13	17
Rango:	1	2	3	4	5	6

In caso di valori coincidenti, si assegna a tali valori un rango pari alla media dei rispettivi valori di rango.

Se la serie di dati è costituita da 5, 11, 11, 17, 7, 11 i ranghi si assegnano come:

Variabile:	5	7	11	11	11	17
Rango:	1	2	4	4	4	6

Cioè i ranghi corrispondenti a 3, 4, 5 vengono sostituiti dalla loro media, cioè 4.

**SCARTO QUADRATICO MEDIO (“STANDARD DEVIATION”)**

Vedi “Deviazione standard”.

**STIMATORE (“ESTIMATOR”)**

La stima del parametro (vedi “Parametro”) effettuata sul campione e come tale soggetta a fluttuazione casuale di campionamento. Esempi di stimatori campionari sono la media campionaria (vedi “Media”) o la proporzione (vedi).

### UNITÀ SPERIMENTALI (“EXPERIMENTAL UNITS”)

Una singola unità di un insieme di entità su cui si conduce uno studio. In una larga maggioranza di studi coincide con una persona, ma l'unità sperimentale può essere costituita anche da un occhio, da un arto, da un'area di cute.

### VARIABILE BINARIA (“BINARY VARIABLE”)

Variabile qualitativa (vedi) caratterizzata dalla presenza o assenza di una certa proprietà.

Ad esempio, risultato positivo/negativo alla ricerca del sangue occulto nelle feci o alla ricerca della proteina di Bence Jones nelle urine.

### VARIABILE CATEGORICA (“CATEGORICAL VARIABLE”)

Vedi “Variabile qualitativa”.

### VARIABILE CONTINUA (“CONTINUOUS VARIABLE”)

Variabile quantitativa (vedi) che può teoricamente assumere tutti i valori dell'asse reale, compatibilmente con il significato biologico della variabile.

Ad esempio, statura, colesterolo plasmatico, peso. La statura, tuttavia, non può assumere un valore di 2 cm.

### VARIABILE DISCRETA (“DISCRETE VARIABLE”)

Variabile quantitativa (vedi) che può assumere solo valori interi.

Ad esempio, numero di figli, numero di ricoveri ospedalieri.

### VARIABILE NOMINALE (“NOMINAL VARIABLE”)

Vedi “Variabile qualitativa”.

### VARIABILE ORDINALE (“ORDINAL VARIABLE”)

Variabile qualitativa (vedi) in cui è possibile definire un ordine della classificazione, benché questo non rappresenti la grandezza della variabile studiata. Per i valori assunti, oltre alle relazioni di uguaglianza e disuguaglianza, è possibile stabilire relazioni di “maggiore” e “minore”.

Ad esempio, peggiorato-invariato-migliorato; peggiorato è inferiore a invariato, ma non è possibile asserire che la distanza tra peggiorato e invariato sia la stessa che esiste tra invariato e migliorato. Altro esempio è la positività di nitriti nell'esame urine, che viene usualmente refertata come +; ++; +++. Certamente una concentrazione classificata ++ è più elevata di una classificata +, ma non è detto che una concentrazione di nitriti classificata come ++, sia il doppio di quella classificata +.

### VARIABILE QUALITATIVA (“QUALITATIVE VARIABLE”)

Detta anche variabile categorica o variabile nominale, è una variabile che si può esprimere utilizzando un attributo. Per i valori assunti valgono solo le relazioni di uguaglianza e disuguaglianza.

Ad esempio, gruppi sanguigni.

### VARIABILE QUANTITATIVA (“QUANTITATIVE VARIABLE”)

Una variabile che si può esprimere utilizzando una misura. Può essere continua (vedi “Variabile continua”) o discreta (vedi “Variabile discreta”).

### VARIANZA (“MEAN SQUARES”)

Misura di variabilità costituita dalla devianza divisa per il numero di gradi di libertà.

Ad esempio, la media di 5, 11, 9, 13, 7 è 9. Il calcolo della devianza risulta 40 (vedi “Devianza”); la varianza risulta:

$$s^2 = \frac{40}{4} = 10$$

La sua unità di misura è il quadrato dell'unità di misura della variabile originale, per cui nel descrivere la distribuzione di una variabile si preferisce usare la DS (vedi “Deviazione standard”), che è la radice quadrata della varianza e, come tale, è espressa nell'unità di misura della variabile originale.

### CONFLITTO DI INTERESSI

Nessuno.

### BIBLIOGRAFIA

1. Altman DG, Bland JM. Standard deviations and standard errors. *Br Med J* 2005;331:903.
2. Mean and standard deviation. In: Swinscow TDV. *Statistics at square one*. BMJ Publishing Group 1997. Consultabile nel sito: <http://www.bmj.com/about-bmj/resources-readers/publications/statistics-square-one/2-mean-and-standard-deviation>.
3. Davies HTO, Cromble IK. What are confidence intervals and p-values? Consultabile nel sito: [http://www.whatseries.co.uk/whatis/pdfs/What\\_are\\_conf\\_int.pdf](http://www.whatseries.co.uk/whatis/pdfs/What_are_conf_int.pdf).
4. Altman DG, Bland JM. Statistics notes: Quartiles, quintiles, centiles, and other quantiles. *Br Med J* 1994;309:996.



## APPENDICE

*Termini considerati, ordinati alfabeticamente secondo la lingua italiana (a sinistra) e inglese (a destra)*

Italiano	Inglese	Inglese	Italiano
Asimmetria (indice di)	Skewness	Absolute frequency	Frequenza assoluta
Campione	Sample	Binary variable	Variabile binaria
Campo di variazione	Range	Categorical variable	Variabile categorica
Coefficiente di variazione	Coefficient of variation	Coefficient of variation	Coefficiente di variazione
Curtosi	Kurtosis	Confidence interval	Intervallo di confidenza
Devianza	Sum of squares	Continuous variable	Variabile continua
Deviazione standard	Standard deviation	Cumulative frequency distribution	Frequenza cumulativa
Differenza interquartile	Interquartile distance	Degrees of freedom	Gradi di libertà
Distribuzione di frequenza	Frequency distribution	Discrete variable	Variabile discreta
Errore standard della media	Standard error of the mean	Estimator	Stimatore
Frequenza assoluta	Absolute frequency	Experimental units	Unità sperimentali
Frequenza cumulativa	Cumulative frequency distribution	Frequency distribution	Distribuzione di frequenza
Frequenza relativa	Relative frequency	Geometric mean	Media geometrica
Gradi di libertà	Degrees of freedom	Harmonic mean	Media armonica
Intervallo di confidenza	Confidence interval	Interquartile distance	Differenza interquartile
Media aritmetica	Mean, arithmetic mean	Kurtosis	Curtosi
Media armonica	Harmonic mean	Mean, arithmetic mean	Media aritmetica
Media geometrica	Geometric mean	Mean squares	Varianza
Mediana	Median	Measure of central tendency	Misura di tendenza centrale
Misura di dispersione	Measure of variation	Measure of variation	Misura di dispersione
Misura di tendenza centrale	Measure of central tendency	Median	Mediana
Moda	Mode	Mode	Moda
Momento	Moment	Moment	Momento
Parametro	Parameter	Nominal variable	Variabile nominale
Percentile	Percentile	Ordinal variable	Variabile ordinale
Popolazione	Population	Parameter	Parametro
Proporzione	Proportion	Percentile	Percentile
Quartili	Quartiles	Population	Popolazione
Ranghi	Ranks	Proportion	Proporzione
Scarto quadratico medio	Standard deviation	Qualitative variable	Variabile qualitativa
Stimatore	Estimator	Quantitative variable	Variabile quantitativa
Unità sperimentali	Experimental units	Quartiles	Quartili
Variabile binaria	Binary variable	Range	Campo di variazione
Variabile categorica	Categorical variable	Ranks	Ranghi
Variabile continua	Continuous variable	Relative frequency	Frequenza relativa
Variabile discreta	Discrete variable	Sample	Campione
Variabile nominale	Nominal variable	Skewness	Asimmetria (indice di)
Variabile ordinale	Ordinal variable	Standard deviation	Deviazione standard, scarto quadratico medio
Variabile qualitativa	Qualitative variable	Standard error of the mean	Errore standard della media
Variabile quantitativa	Quantitative variable	Sum of squares	Devianza