

Dalla normalità alla diagnosi

Anteo Di Napoli¹, Michele Nichelatti², Francesco Franco³

¹ Comitato Tecnico-Scientifico RIDT, Roma

² Servizio di Biostatistica, Dipartimento di Ematologia e Oncologia, Ospedale Niguarda, Milano

³ Regione Lazio, Roma

From normal distribution to diagnostic test

Distribution curves of laboratory parameters in healthy and sick subjects are quite similar, leading to a “grey zone” referred to subjects that could be not correctly identified and classified. To identify a cut-off value, ROC (Receiver Operating Characteristics) curves can be applied; they represent the graphic relationship between sensibility (true positive) and specificity (false positive), depending on cut-off values variability. ROC curve subtends AUC (Area Under the Curve); AUC’s width is directly related to discerning power of any validated statistical test.

Keywords: Normal distribution, Sensibility, Specificity, ROC curves



Anteo Di Napoli

L’utilizzo del termine “normale” ai tempi del “politicamente corretto” può essere foriero di equivoci. Ma questo è, forse, il destino della statistica più in generale, troppo spesso considerata “la scienza che misura la media”. Il celeberrimo sonetto di Trilussa “La statistica” ha grandemente contribuito alla “cattiva fama” (appendice).

In realtà, la media da sola ha poco senso nel descrivere la distribuzione di una variabile, se non si fornisce anche una misura di variabilità. Del resto, una variabile che non varia sarebbe una costante!

Abbiamo già ampiamente parlato della più nota distribuzione di una variabile, la cosiddetta “normale”, detta anche “curva gaussiana”, dal nome del matematico tedesco Carl Friederich Gauss (1-3).

Come è noto, una distribuzione normale presenta una forma “a campana”: i valori più frequenti (più probabili) si collocano al centro e media, mediana e moda coincidono in valore; quelli meno frequenti (meno probabili) si collocano ai margini della campana, nelle cosiddette code della distribuzione.

Nuovamente, un’erronea interpretazione della definizione statistica di “non normale”, cioè meno probabile, in quanto meno frequente, ha quasi portato a identificarla con un’idea negativa di “aberrazione”. E, invece, basterebbe pensare che un genio sicuramente si collocherà nelle code di una gaussiana e che anche l’evoluzione stessa è la conseguenza della comparsa di una “variante della normalità”.

La grande fortuna che la “gaussiana” ha avuto nell’ambito delle scienze biologiche è legata al fatto che la variabilità insita in molti fenomeni naturali (e, quindi, nelle misure che la descrivono) segue, appunto, il modello teorico descritto dalla curva di distribuzione normale. Inoltre, quand’anche i valori di una variabile in una popolazione non si distribuissero normalmente, i valori delle medie di campioni estratti da tale popolazione, invece, lo sarebbero, che è quanto stabilisce il teorema del limite centrale (per comprenderlo, nelle sue linee essenziali, si consiglia di provare a utilizzare due applet pubblicate agli indirizzi web http://onlinestatbook.com/stat_sim/sampling_dist/ e <http://www.amstat.org/publications/jse/v6n3/applets/clt.html>).

A questo punto, è necessario fare un ulteriore passo in avanti. Ci si può accontentare della definizione statistica di normalità, limitandosi a considerare via via meno normali i valori che, essendo meno frequenti, sono meno probabili?

In medicina, per esempio, è comunemente accettato considerare come normali i valori compresi fra i percentili 2.5 e 97.5 della distribuzione di una popolazione definita come “sana”. Attenzione, però! Se la definizione di “sano” e quella di “normale” (in senso statistico) coincidessero, automaticamente tutti i soggetti con valori collocati nelle code di una distribuzione, al di sotto del 2.5% e al di sopra del 97.5%, dovrebbero essere considerati “ammalati”. Conseguentemente, utilizzando come riferimento i valori di “normalità” di un parametro, ad esempio di laboratorio, la prevalenza di una qual-

Accepted: October 28, 2015

Published online: November 19, 2015

Indirizzo per la corrispondenza:

Dr. Anteo Di Napoli

anteo.dinapoli@libero.it

Dr. Michele Nichelatti

michele.nichelatti@ospedaleniguarda.it

Dr. Francesco Franco

franco_francesco@fastwebnet.it

siasi patologia diagnosticata sarebbe sempre del 5%!

Questo paradosso, se fosse ancora necessario, conferma quanto sia errato considerare sinonimi “sano” e “normale” e, complementariamente, “malato” e “anormale”.

È pur vero che la “normalità” è definita a partire dalla distribuzione del valore di una variabile tra i soggetti “sani” di una popolazione, a loro volta identificati come individui senza la malattia. In termini pratici, un medico necessita di una definizione operativa di valore “normale” per una variabile di laboratorio dalla quale far discendere un comportamento verso un paziente. Per le considerazioni succitate, si è sostituito il concetto di valore normale con quello di valore di riferimento, riferito a soggetti con caratteristiche confrontabili con quelle del soggetto a cui il risultato di laboratorio si riferisce, al fine di stabilire dei valori al di fuori dei quali si presume che vi sia un’anormalità con un limite di errore accettabile.

In una situazione ideale, le curve di distribuzione dei valori dei sani e dei malati, rispetto a un parametro di laboratorio, non dovrebbero sovrapporsi, essendo, in tal caso, le due popolazioni del tutto differenti. Nella realtà, tale situazione non si verifica mai e la “campana” dei sani e quella dei malati risultano essere più o meno sovrapposte. Tale area di sovrapposizione rappresenta una “zona grigia” che si riferisce alla quota di soggetti che potrebbero non essere correttamente classificati come sani o malati rispetto al valore di riferimento del parametro assunto come soglia di discriminazione (Fig. 1).

Per esempio, secondo la classificazione del JNC 7 (*Joint National Committee on Prevention, Detection, Evaluation and Treatment of High Blood Pressure*), si considerano normali una pressione sistolica inferiore a 120 mmHg e una pressione diastolica inferiore a 80 mmHg, mentre, al di sopra, dei 140 mmHg di massima o dei 90 mmHg di minima si è ipertesi (4, 5).

Fissando un valore soglia di 140 mmHg per la pressione sistolica e di 90 mmHg per quella diastolica, potrà accadere che si classificheranno come ipertese persone con valori più elevati, che, invece, sono sane; viceversa, non si considererebbero ipertese persone con valori inferiori a quella soglia e che, invece, lo sono.

Si tratta, evidentemente, di un esempio “scolastico”, poiché nessun medico dovrebbe basare la propria diagnosi di ipertensione esclusivamente sui valori misurati con lo sfigmomanometro. Volevamo, qui, sottolineare gli effetti che la scelta di un certo valore soglia determina sulla classificazione dei soggetti come sani o malati. Le conseguenze sono intuibili e ben note, anche alle case farmaceutiche...

Se il mercato di un dato farmaco si satura, una strategia possibile è incrementare la platea di pazienti che, teoricamente, potrebbero trarre beneficio dall’esserne trattati. Pertanto, abbassare i valori soglia di un parametro diagnostico oltre i quali un soggetto si definisce malato determina una ridefinizione diagnostica, classificando come “non normali” persone che in precedenza erano considerate assolutamente sane. È un fenomeno che, negli ultimi due decenni, ha riguardato, per esempio, i valori di “normalità” della pressione arteriosa, della glicemia e della colesterolemia. In quest’ultimo caso, poiché

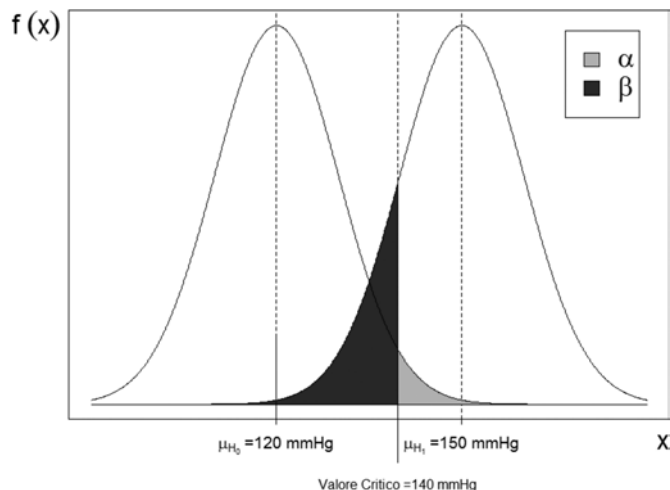


Fig. 1 - Probabilità di errori α (falsi positivi) e β (falsi negativi).

non si è individuato un valore soglia oltre il quale si ha un incremento del rischio di morte coronarica, si è introdotto il concetto di “valore desiderato”, fissato per l’adulto a 200 mg/dL.

Non è questa la sede per aprire un dibattito sull’argomento, meno che mai quella per approfondire il ruolo avuto anche dalle società scientifiche su queste scelte, talora in condizione di potenziale conflitto di interessi. Qui si vuole solo ribadire al lettore di porsi sempre con un atteggiamento critico nei confronti delle “ultimissime” definizioni diagnostiche basate sull’aggiornamento dei valori soglia di un parametro utilizzato come *test* diagnostico.

Esiste una tecnica statistica rappresentata dalle cosiddette curve ROC (*Receiver Operating Characteristics*), che consente di individuare il valore soglia di un *test* diagnostico, se questo è espresso da un parametro definito come variabile continua. Tale valore, noto anche come “*best cut-off*”, è quello dove si ottiene l’equilibrio ottimale tra sensibilità e specificità e, quindi, il maggiore potere predittivo possibile per quel *test* (6).

Si ricorda che la sensibilità rappresenta la probabilità che il *test* risulti positivo quando la condizione definita è presente ed esprime la proporzione di soggetti con la condizione definita che risultano positivi al *test*; è data dal rapporto tra veri positivi e il totale dei soggetti con la condizione; il suo complemento è la proporzione di falsi negativi, definiti come la frazione di soggetti con la condizione che risultano negativi al *test*.

La specificità, invece, rappresenta la probabilità che il *test* risulti negativo quando la condizione definita non è presente ed esprime la proporzione di soggetti senza la condizione definita che risultano negativi al *test*; è data dal rapporto tra veri negativi e il totale dei soggetti senza la condizione; il suo complemento è la proporzione di falsi positivi, definiti come la frazione di soggetti senza la condizione che risultano positivi al *test*.

La relazione tra *test* diagnostico e stato del paziente è riassunta nella Tabella I.

Tornando alla curva ROC, essa rappresenta la relazione grafica tra la sensibilità (veri positivi), posta sull’asse delle

TABELLA I - Test diagnostico e condizione del soggetto

		Malati	Non Malati	Totale
TEST	Positivi	a	b	(a + b)
	Negativi	c	d	(c + d)
	Totale	(a + c)	(b + d)	(a + b + c + d)

		Malati	Non Malati	Totale
TEST	Positivi	VP	FP	(VP + FP)
	Negativi	FN	VN	(FN + VN)
	Totale	(VP + FN)	(FP + VN)	(VP + FP + FN + VN)

VP = veri positivi; FP = falsi positivi; VN = veri negativi; FN = falsi negativi

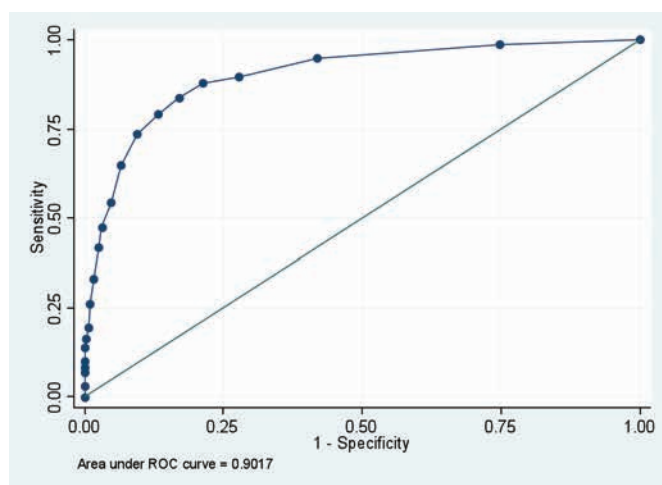


Fig. 2 - Curva ROC di mortalità intra-ospedaliera per punteggio CRIB. Esempio tratto da uno studio nel quale si voleva misurare il potere predittivo della mortalità intra-ospedaliera di un indice di gravità di neonati prematuri (punteggio CRIB).

ordinate, e il complemento a 1 della specificità (falsi positivi), posta sull'asse delle ascisse, al variare del valore della soglia stabilita per quel test; nel grafico della curva, per tutti i possibili valori soglia del punteggio, vengono riportati i corrispondenti valori della proporzione di veri e di falsi positivi. Congiungendo i punti relativi a tali valori si costruisce la curva ROC, che sottintende un'area detta AUC (Area Under the Curve), la cui estensione è una misura del potere discriminante del test in questione (maggiore al crescere dell'AUC, quanto più la curva ROC si avvicina al vertice del grafico).

Un esempio è presentato nella Figura 2.

Il test ideale, quando sia la specificità che la sensibilità sono pari a 1, equivale a un'area sottintesa dalla curva ROC uguale a 1, espressione del massimo potere discriminante del test. Il test con il minimo potere discriminante, che non aggiunge, cioè, alcuna informazione diagnostica, equivale a un'area delimitata dalla diagonale che congiunge il valore di sensibilità 0 e di specificità 1 con il valore di sensibilità 1 e di specificità 0; in

questo caso, l'area sottintesa dalla curva ROC è uguale a 0.5.

Secondo la classificazione proposta da John A. Swets si considera poco informativo un test con AUC compresa tra 0.5 e 0.7, moderatamente informativo un test con AUC compresa tra 0.7 e 0.9 e molto informativo un test con AUC superiore a 0.9.

In conclusione, è indiscutibile l'importanza che assume per un medico l'ausilio di un test nel porre o nell'escludere una diagnosi; tuttavia, la valutazione del test deve sempre tenere conto dei limiti insiti in qualsiasi misura e non dovrebbe mai essere considerato l'unico parametro da considerare nella scelta, da condividere possibilmente con il paziente, del percorso diagnostico e terapeutico da intraprendere.

Disclosures

Financial support: No financial support was received for this submission.

Conflict of interest: The authors have no conflict of interest.

Bibliografia

- Nichelatti M, Nordio M, Maggiore U, Postorino M, Limido A, Di Napoli A. La distribuzione normale - Parte I: Introduzione. *Giornale di Tecniche Nefrologiche & Dialitiche* 2014;26(1):78-80.
- Nichelatti M, Nordio M, Maggiore U, Postorino M, Limido A, Di Napoli A. La distribuzione normale - Parte II: Un esempio pratico. *Giornale di Tecniche Nefrologiche & Dialitiche* 2014;26(2):196-8.
- Nichelatti M, Di Napoli A, Nordio M, Maggiore U, Postorino M, Limido A. La distribuzione normale - Parte III: Una semplice introduzione al Teorema del Limite Centrale. *Giornale di Tecniche Nefrologiche & Dialitiche* 2014;26(3):290-3.
- Ministero della Salute. Ipertensione arteriosa. http://www.salute.gov.it/portale/salute/p1_5.jsp?lingua=italiano&id=18&area=Malattie_cardiovascolari.
- U.S. Department of Health and Human Services. Prevention, Detection, Evaluation, and Treatment of High Blood Pressure. The Seventh Report. NIH Publication No. 04-5230, August 2004.
- Hanley JA, Mc Neil BJ. A method of comparing the area under the receiver operating characteristics curves derived from the same cases. *Radiology* 1983;148:831-43.

Appendice

LA STATISTICA - Trilussa

*Sai ched'è la statistica? È na' cosa
che serve pe fà un conto in generale
de la gente che nasce, che sta male,
che more, che va in carcere e che spòsa.
Ma pè me la statistica curiosa
è dove c'entra la percentuale,
pè via che, lì, la media è sempre eguale
puro co' la persona bisognosa.
Me spiego: da li conti che se fanno
seconno le statistiche d'adesso
risurta che te tocca un pollo all'anno:
e, se nun entra nelle spese tue,
t'entra ne la statistica lo stesso
perch'è c'è un antro che ne magna due.*