DELL'INFLUENZA DEL GOZZO
SULLE STATISTICHE DELLA STATURA

NOTA

del Dott. FERNANDO DE HELGUERO

È ben noto agli studiosi che le seriazioni relative ad un determinato carattere in un gruppo omogeneo di individui seguono spesso la legge normale espressa dall'equazione:

$$y = \frac{n}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-b}{\sigma}\right)^2},$$

dove b rappresenta il valore medio, σ misura la variabilità e dicesi *deviazione normale*, ed n è il numero degli individui misurati.

È pure noto però che non tutte le seriazioni seguono rigorosamente questa legge anche se il materiale è omogeneo. Io ritengo che in molti casi la abnormalità sia dovuta a qualche causa di perturbazione che favorisca alcune classi a danno di altre; io ritengo, cioè, che la legge di variazione sia fondamentalmente quella normale che non possa manifestarsi completamente per l'influenza di cause estranee.

Così, per esempio, considerando la variazione dell'indice perimetro-toracico/statura, poichè si sa che ad un valore basso di tale indice va connessa una gracilità generale ed una minor resistenza alle malattie, noi dovremo trovare nella seriazione un minor numero di indici bassi di quello che richiederebbe la legge normale. La causa perturbatrice sarebbe in tal caso la selezione che colpisce più gli indici bassi che gli alti e la seriazione risulterebbe abnormale. Nè questa è la sola ipotesi che possiamo fare per spiegare teoricamente la abnormalità, ma molte altre ancora.

Ora a me sembra che l'analisi statistica davanti ad una seriazione abnormale debba sempre cercare se è possibile dare una simile spiegazione ed in tal caso ricercare quale sarebbe stata la variazione normale ipotetica e quale l'azione della causa perturbatrice. In una comunicazione da me fatta al IV Congresso internazionale dei Matematici tenutosi a Roma nell'aprile 1908 io esposi i metodi matematici per fare tale ricerca partendo da due diverse ipotesi sulla natura della causa perturbatrice. Una ipotesi è che la perturbazione sia dovuta a selezione come nell'esempio sopra esposto, l'altra ipotesi è che una parte degli individui sia colpita da una causa di atrofia per cui essi non raggiungano il loro sviluppo normale e figurino perciò in una classe diversa da quella nella quale dovrebbero trovarsi.

Nella presente nota voglio considerare appunto un esempio di quest'ultimo genere di abnormalità e mostrare come possa adoperarsi il metodo statistico nello studio di un fatto biologico.

Diciamo perciò qualche cosa di più su questo genere di curve abnormali; eccone la storia.

Nel 1883 il dott. Ridolfo Livi studiando la statura degli Italiani (1) osservò che alcune seriazioni per esempio quella dei circondari di Aosta e di Sondrio, presentano un numero di basse stature superiore a quello che la legge normale avrebbe richiesto: egli attribuiva ciò al gozzo che è oltremodo diffuso fra quelle popolazioni ed impedisce in molti casi il normale sviluppo dell'organismo. Il dott. Livi si contentò però di studiare il fatto all'ingrosso, considerando solo le alte e le basse stature senza ricercare la legge normale ipotetica.

Nel 1894 il prof. A. Giard interpretò in un modo analogo la statistica delle chele nei maschi della *Forficula auricularia*. Si era visto che erano molto frequenti gli individui con chele lunghe e quelli con chele corte, mentre erano scarsi i maschi con chele di una lunghezza media. La seriazione presentava due massimi, era cioè

(1) *Archivio per l'Antropologia e l'Etnologia*, Vol. XIII, p. 243, Firenze, 1883.

(2) A. Giard, *Évolution des êtres organisés. Sur certains cas de dédoublement des courbes de Galton dus au parasitisme*, etc. Comptes Rendus de l'Académie des Sciences. Paris, 1894, Vol. 118, p. 870.

bimodale. Il prof. A. Giard spiegò questo fatto attribuendolo alla presenza di parassiti in alcuni degli individui che perciò subivano una parziale castrazione, con minore sviluppo delle chele che rappresentano un carattere sessuale secondario. Uno dei massimi della seriazione era perciò dovuto agli individui normali, l'altro a quelli colpiti dalla causa di perturbazione.

Come si vede la seriazione è dello stesso tipo di quella considerata dal Livi, soltanto la più intensa perturbazione produce in questo caso la bimodalità che manca nelle curve della statura.

In generale io ho dato il nome di *curve di Giard* a tutte quelle curve di frequenza in cui una parte degli individui è colpita da una causa di atrofia per la quale non può raggiungere, per il carattere in esame, lo sviluppo normale. Non le ho chiamate del Livi perchè questo nome già si adopera per altre curve.

Per poter fare l'analisi matematica di tali curve io ho dovuto porre delle ipotesi che rendano più preciso il problema. Esse riguardano la legge di perturbazione e sono:

1° La probabilità di essere colpito è la stessa per ogni individuo, cioè essa è indipendente dal carattere in esame, la indico con h.

2° La causa perturbatrice produce lo stesso effetto su ogni individuo colpito, cioè se per il carattere in esame esso doveva giungere alla grandezza x arriverà soltanto alla grandezza $x - c$, con c indipendente da x.

La prima ipotesi mi sembra possa accettarsi senza difficoltà; per la seconda potremo ritenere il valore di c che troviamo coll'analisi come una media, anche se non vogliamo supporlo costante per tutti i colpiti.

In queste ipotesi è facile vedere che tutti gli individui si distribuiscono in due gruppi normali, uno (la cui media indico con b), formato dagli individui non colpiti dalla causa perturbatrice, l'altro di media $b - c$ formato dai colpiti; i due gruppi normali hanno la stessa deviazione normale (che indico con σ) e questo è molto notevole perchè a ciò si deve la relativa semplicità dell'analisi.

Nella citata comunicazione fatta al IV Congresso dei Matematici, io indico il procedimento per ricavare dalla seriazione empirica la variazione normale ipotetica, cioè la sua media b e la deviazione normale σ, e i parametri della perturbazione cioè h che dico *indice di morbilità*, e c *indice di atrofia*. Moltiplicando h per 100 ho la percentuale dei colpiti.

Per applicare il metodo ad un esempio biologico io non saprei trovare migliore seriazione di quella della statura nei paesi dove è notevole l'influenza del gozzo endemico, cioè una di quelle già considerate dal Livi. Prendiamo perciò la seriazione delle stature dei coscritti del circondario d'Aosta per il quinquennio 1855-59 e per evitare l'effetto della tendenza all'arrotondamento per cui predominano le stature espresse con multipli di 5, raggruppiamole in classi di cinque in cinque centimetri. Per esempio sommo le stature da 1.23 a 1.27, quelle da 1.28 a 1.32 e così via; ad ogni classe faccio corrispondere la sua statura media che viene espressa da un multiplo di 5. Ecco la variazione:

Classi	1.25	1.30	1.35	1.40	1.45	1.50	1.55	1.60	1.65	1.70	1.75	1.80	185
Frequenza	42	24	54	70	164	300	570	893	920	567	204	49	4

Sono in tutto 3861 reclute. La media è a cm. 160.29

La seriazione è nettamente asimmetrica come può osservarsi nel poligono empirico tracciato nella figura.

La curva normale che corrisponderebbe (1) a questa seriazione avrebbe $\sigma =$ cm. 9.805; i punti corrispondenti sono segnati con crocette nella figura; come si vede questa curva dà una pessima rappresentazione della statistica. Dobbiamo perciò concludere che la seriazione non segue la legge normale. Vediamo se può ottenersi una migliore rappresentazione con una *Curva di Giard.*

Per applicare il metodo devo ricavare dulla seriazione i primi quattro *momenti*. Prima di introdurli nei calcoli li ho però alquanto modificati per diminuire l'influenza enorme che avrebbe la classe di m. 1.25. Su queste modificazioni che ritengo sempre utili quando si adopera nella perequazione il metodo dei momenti dirò altrove.

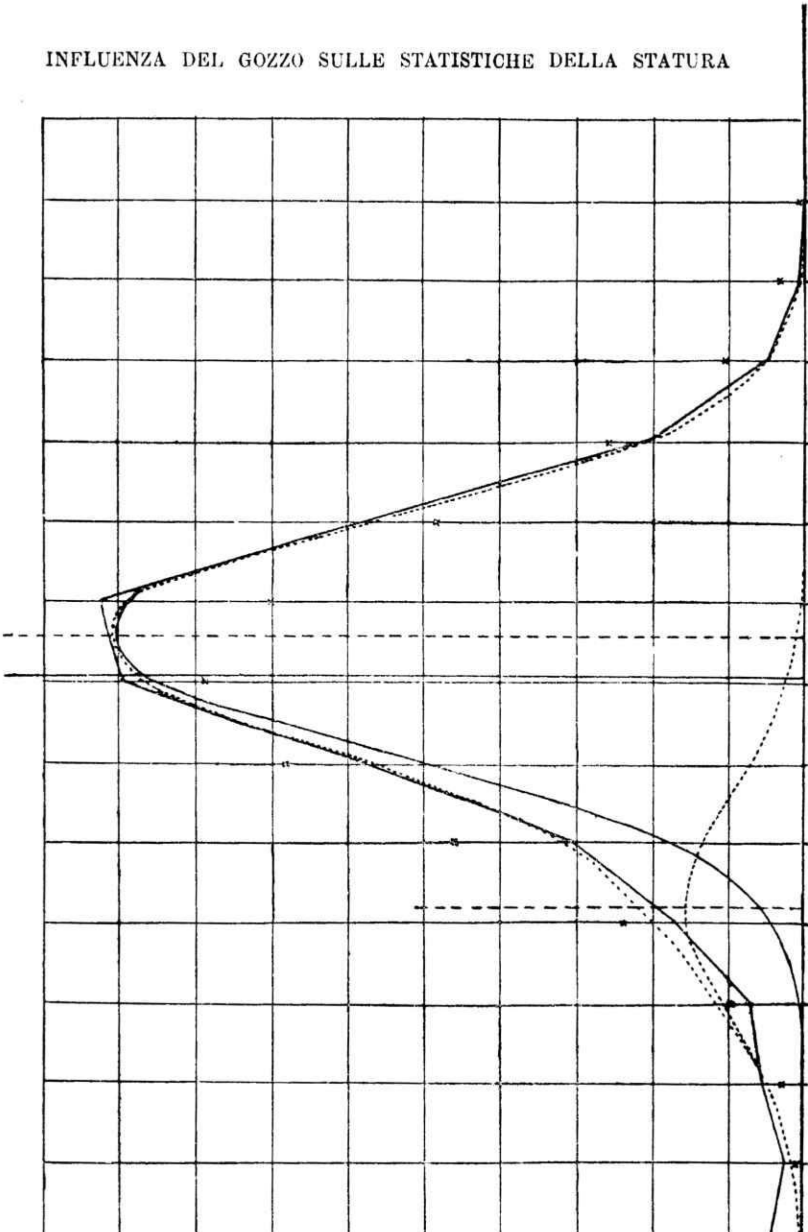
I momenti rispetto alla media (che con tali correzioni viene a 160.4725) come li introduco nei calcoli sono:

$$\mu_2 = 3.475, \quad \mu_3 = -5.05, \quad \mu_4 = 49.8395$$

Con questi momenti calcolo

$$\frac{1}{2}(\mu_4 - 3\,\mu_2^2) = 6.8063\;, \; \frac{1}{2}\,\mu_3^2 = 12.7512.$$

(1) Per il modo di trovare σ e di calcolare i *momenti* di una seriazione rimando alla memoria presentata a questa Società dal dott. Velio Zanolli ed al manuale del Davenport « Statistical methods with special reference to Biological Variation ». Wiley & Sons, New York, 1904.



Statura dei coscritti del circondario di Aosta delle leve 1855-1859.

Risolvo poi l'equazione

$$p_2^3 + 6.8063\, p_2 + 12.7512 = 0,$$

e trovo con successive approssimazioni

$$p_2 = -1.437\,;$$

e poi $$p_1 = -\frac{\mu_2}{p_2} = +2.4182.$$

Le due medie dei due gruppi sono allora date dalle radici della

$$b^2 - 2.4182\, b - 1.537 = 0,$$

che dà

$$b = +0.4938.$$
$$b_1 = -2.9120.$$

(Ho cambiato i segni perchè μ_3 è negativo).

Ricordando la posizione del baricentro ed il valore dell'unità delle ascisse trovo le medie

$$b = \text{cm. } 162.94,$$
$$b_1 = b\text{-}c = \text{cm. } 145.9125,$$

per cui $c = 17.0275$.

La deviazione normale è data da $\sigma = \sqrt{p_2 + \mu_2}$, onde $\sigma = 1.1276$ e in cm. $\sigma =$ cm. 7.138.

Gli altri parametri si trovano colle formule

$$y_1 = -\frac{b_1\, \mu_0}{b\text{-}b_1}\,,\quad y_2 = \frac{b\, \mu_0}{b\text{-}b_1}\,,\quad h = \frac{y_2}{\mu_0}\,,$$

dove μ_0 è l'area 3861. Si trova

$$y_1 = 3301.20\,,\quad y_2 = 559.80\,,\quad h = .145.$$

L'equazione della curva perturbata è:

$$y = \frac{3301.20}{7.139\sqrt{2\pi}}\, e^{-\frac{1}{2}\left(\frac{x-162.94}{7.138}\right)^2} + \frac{559.80}{7.138\sqrt{2\pi}}\, e^{-\frac{1}{2}\left(\frac{x-145.9125}{7.138}\right)^2}.$$

La curva perturbata insieme colle due componenti è tracciata nella figura. In essa una delle componenti è punteggiatà l'altra è continua; di entrambe è segnata solo quella parte nella quale sono distinte dalla curva perturbata che è punteggiata.

Ecco i valori calcolati per le diverse classi:

Classi	1.25	1.30	1.35	1.40	1.45	1.50	1.55	1.60	1.65	1.70	1.75	1.80	1.85	1.90
Frequenze	2	13	49	117	194	311	565	870	890	567	221	53	8	1

Come si vede dalla figura la rappresentazione è buona e questo ci è confermato dal calcolo della differenza delle aree che è del 5.70 per cento rispetto all'area del poligono empirico. Poichè la rappresentazione è soddisfaciente possiamo trarre delle conclusioni:

La seriazione della statura dei coscritti del circondario di Aosta per il quinquennio 1855-59 è bene rappresentata da una *curva di Giard* che suppone una variazione normale colla media a cm. 162.94 e la deviazione normale di cm. 7.138.

La perturbazione colpisce il 14.50 per cento degli individui e la statura media dei colpiti è di cm. 145.9125 con una differenza dalla statura media normale di cm. 17.0275.

Questo ci dice l'analisi statistica della seriazione; discutiamone i risultati.

Intanto se confrontiamo la media della seriazione empirica (160.29) e quella della variazione normale ipotetica (162.94) con quelle dei circondari vicini a quello di Aosta, vediamo che la seconda è infinitamente più probabile della prima. Infatti in tutti i circondari del Piemonte la media oscilla da un minimo di 162.20 per Pallanza ad un massimo di 163.75 per Alessandria, eccettuato Pinerolo con 161.57 che però presenta anch'esso, come Aosta, un grande numero di riformati per gozzo.

Anche la deviazione normale ipotetica 7.137 è molto più probabile di quella attuale di cm. 9.805 molto superiore a quella degli altri circondari del Piemonte. Nè deve ritenersi il numero 7.138 eccessivamente piccolo perchè altri circondari danno cifre anche minori; per es., Alessandria ha $\sigma = 6.08$.

Riguardo all'azione della causa perturbatrice osservo che la nostra curva teorica supporrebbe che i coscritti colpiti costituissero un gruppo normale simmetrico rispetto alla media 145.91. Però per le stature inferiori a cm. 150 la curva teorica si allontana abbastanza sensibilmente dal poligono empirico, prima per eccesso e poi per difetto, onde possiamo ritenere che la variazione dei colpiti dalla causa perturbatrice non sia rigorosamente normale, ma piuttosto asimmetrica con una più lenta diminuzione dalla parte dei nani.

In ogni modo se non rigorosamente almeno all'ingrosso la nostra analisi ci ha rivelato la natura e la intensità della abnormalità: ci dice che essa è prodotta da un gruppo di individui formato da circa il 14.50 per cento di tutte le reclute, che ha una media intorno a 146 cm.

Naturalmente che questa causa perturbatrice sia proprio il gozzo non ce lo può rivelare l'analisi, per quanto accurata di una sola statistica. Tale è l'opinione del dott. Livi, ed anche la mia, come mostra il titolo di questa nota, ma per verificarla rigorosamente bisognerebbe confrontare le divergenze dalla normalità delle seriazioni dei differenti circondari colle corrispondonti cifre dei riformati per gozzo. Un tale lavoro con molto materiale, accuratamente studiato, sta preparando il dott. Luigi Munaron già noto per le sue ricerche cliniche sul decorso del gozzo e sulla patogenesi del cretinismo endemico. Solo da uno studio sintetico, complessivo, come quello che egli sta preparando potrà venire la conferma decisiva dell'ipotesi molto probabile del dott. Livi.

Questo non posso farlo io coll'analisi di una sola seriazione. Solo voglio evitare una obbiezione che potrebbe farsi alla interpretazione del gozzo come la causa perturbatrice rivelata dalla mia analisi. Si potrebbe obbiettare che il numero dei gozzuti del circondario di Aosta è del 32-40 per cento secondo i calcoli dei diversi autori, mentre le mie ricerche condurrebbero ad una percentuale di anormali del 14.50 per cento.

È però ovvio che non in tutti i colpiti da gozzo endemico deve verificarsi un minore sviluppo, ma solo in quelli che vengono gravemente colpiti durante la fase del più rapido accrescimento e questi costituiscono solo una parte del totale dei gozzuti.

Tornando ora al metodo statistico del qnale questa nota vuol essere solo un esempio espositivo si noti il procedimento generale che è quello che si adopera comunemente in tutte le ricerche che si fanno con metodi matematici di fenomeni complicati come sono quelli biologici:

Per precisare il problema si pongono delle ipotesi se non indiscutibili almeno verosimilmente approsimate, poi mediante i metodi matematici si traggono le conclusioni che si confrontano coi dati empirici.

Se esse concordano entro i limiti degli errori di osservazione, le ipotesi restano con ciò rigorosamente verificate; se invece si notano delle discrepanze non accidentali ne dobbiamo dedurre che le ipotesi poste non sono rigorosamente esatte ma solo approssimate, e coi risultati del confronto possiamo modificarle convenientemente.

Nel problema ora trattato le ipotesi poste erano le seguenti:

1° che la variazione ipotetica fosse normale;

2° che la morbilità fosse indipendente dalla statura;

3° che l'effetto prodotto dalla causa perturbatrice fosse indipendente dalla statura.

La prima ipotesi è stata verificata da molte altre osservazioni statistiche. La generale concordanza della curva calcolata col poligono empirico ci dice che anche le altre due ipotesi sono con grande approssimazione vere; se però vogliamo spiegare la divergenza notata per le stature inferiori a m. 1.50 dobbiamo dire che:

o la morbilità è un po' maggiore per gli individui di bassa statura;

o la diminuzione di sviluppo per la causa perturbatrice è un po' maggiore per gli individui di bassa statura.

Se pensiamo che molte volte una statura molto al di sotto della media può esser dovuta a cause patologiche per cui sia connessa con una gracilità generale ed una minor resistenza alle malattie, queste spiegazioni ci appaiono bene accettabili.

In ogni modo i valori trovati per h e per c devon considerarsi come medie dalle quali gli indici di morbilità e di atrofia, che cambiano leggermente da classe a classe, di pochissimo si allontanano.

Asti, aprile 1908.