

NOTE SULLE MISURE DI CONCENTRAZIONE

Domenico Suppa*

20/04/2011

Introduzione

Queste brevi note sulla misura della concentrazione trattano solo di alcuni degli indici maggiormente utilizzati, cercando di evidenziare i rapporti tra essi intercorrenti e, talvolta, l'equivalenza delle diverse misure proposte al di là del metodo di costruzione seguito per ottenerle.¹

1 L'indice di entropia

Uno degli approcci utilizzati per misurare la concentrazione delle modalità di un carattere in una popolazione è basato sulla teoria dell'informazione. Il ragionamento che porta alla costruzione assiomatica di questa misura è basato sull'assunto che il contributo informativo apportato dal verificarsi di un determinato evento è una funzione della probabilità associata all'evento in un senso specifico:² si ipotizza che l'informazione possa essere misurata dall'effetto sorpresa che il soggetto osservatore sperimenta nel momento in cui l'evento i si verifica, dato che ad esso è associata la probabilità p_i . Così se l'evento i ha una probabilità di

*Unina.

¹Non saranno trattati in questa sede molti altri importanti indici di concentrazione, per una disamina più completa e un riferimento bibliografico esteso si rinvia a [Piccolo(1998), pp. 119 e ss.]. Molti degli indici esposti, inoltre, presentano vantaggi e aspetti critici che ne consigliano l'uso in certe situazioni e non in altre, neppure questi aspetti saranno affrontati nel seguito se non in modo occasionale (come per l'indice di Herfindahl).

²V., [Baumgärtner(2004)]

verificarsi pari a 0.999, l'informazione derivante dal suo verificarsi (la sorpresa dell'osservatore) dovrebbe essere molto bassa e dovrebbe essere zero se si tratta di un evento certo (che si verifica con probabilità pari a 1). Viceversa, gli eventi che si verificano con probabilità molto bassa (tendente a zero) dovrebbero generare una misura molto elevata (tendente ad infinito). Gli assiomi posti a fondamento di questo tipo di misura e che devono essere soddisfatti da una funzione $h(p)$ in grado di rappresentarla, seguendo il ragionamento svolto sopra, sono i seguenti: ³

1. h è continua per $p \in]0, 1]$,
2. h è monotona: $p_1 < p_2 \Rightarrow h(p_1) > h(p_2)$,
3. h è additiva: $h(p_1 p_2) = h(p_1) + h(p_2)$.

Oltre a questi tre assunti molto generali si richiede che $h(p)$ sia non negativa e che, quindi, il suo codominio sia $[0, +\infty[$. Dall'analisi matematica è noto che esiste una sola funzione in grado di soddisfare i precedenti assiomi, essa è l'opposto della funzione logaritmo di p [v. Figura 1]:

$$h(p) := -\log p \quad (1)$$

In generale vi saranno n eventi osservabili e l'informazione attesa sarà data dalla media:

$$\bar{h} = \sum_{i=1}^n p_i h(p_i) = - \sum_{i=1}^n p_i \log p_i \quad (2)$$

Naturalmente ogni misura proporzionale ad \bar{h} soddisfa i postulati e la base dei logaritmi può essere qualunque numero positivo. Per ottenere una espressione normalizzata (compresa tra 0 e 1) di \bar{h} possiamo dividere quest'ultima per il suo massimo.⁴

La massimizzazione della funzione $\bar{h}(p_1, \dots, p_n)$ soggetta al vincolo

$$\sum_{i=1}^n p_i = 1 \quad (3)$$

richiede che siano soddisfatte le seguenti condizioni:

$$p_i = \bar{p} \quad \forall i = 1, \dots, n \quad (4)$$

³V. [Shannon(1948), Khinchin(1957)]

⁴ \bar{h} è infatti una funzione concava, essendo la somma di funzioni concave $-p_i \log p_i$ ognuna delle quali ha derivata prima pari a $-(1 + \log p_i)$ e derivata seconda $-\frac{1}{p_i} < 0$.

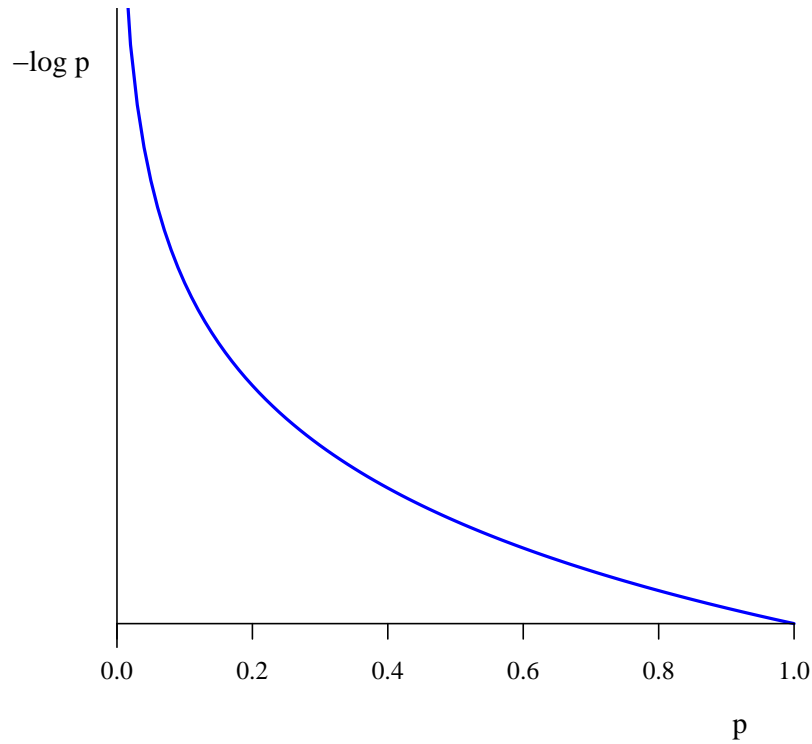


Figura 1: Funzione $h(p)$

che, sostituite nel vincolo (3), determinano il seguente risultato:

$$\sum_{i=1}^n p_i = \sum_{i=1}^n \bar{p} = n\bar{p} = 1 \Rightarrow \bar{p} = \frac{1}{n} \quad (5)$$

In definitiva, il massimo della funzione \bar{h} si ottiene quando tutte le probabilità p_i sono uguali a $\frac{1}{n}$, in questo caso, come è facile verificare, $\bar{h} = \log n$ (il minimo di \bar{h} tende a zero se p_i tende ad assumere valore 1 mentre le altre p_j tendono a 0, $\forall j \neq i$; ciò è vero per costruzione). Pertanto, la misura normalizzata di \bar{h} è data da:⁵

⁵Fissato un determinato valore di \bar{h}^* ci si può chiedere quale è il numero degli eventi equiprobabili in grado di produrre quello stesso valore \bar{h}^* , in tal modo si perviene all'indice di equivalenza numerica $N = e^{\bar{h}^*}$ [v. §4].

$$S = -\frac{\sum_{i=1}^n p_i \log p_i}{\log n} \quad (6)$$

Questo indice, nella versione normalizzata S oppure in quella non normalizzata \bar{h} , è noto come indice di *entropia* (spesso detto indice di Shannon, nell'ambito della teoria dell'informazione, o indice di Theil in economia).⁶ Con riferimento ad S , si può verificare che esso varia da un minimo di 0 (nel caso di massima concentrazione o perfetta disuguaglianza) ad un massimo di 1 (nel caso della concentrazione minima o perfetta uguaglianza).

2 La differenza semplice media

Un indice di *mutua variabilità* che può essere utilizzato per misurare la concentrazione di un fenomeno quantitativo è dato dalla *differenza semplice media*. Esso consiste in una media delle differenze assolute tra tutte le coppie delle modalità assunte da un carattere quantitativo.⁷ Siano $x_i, \forall i = 1, \dots, n$, tali modalità, l'indice (detto *differenza semplice media con ripetizione*) viene calcolato nel modo seguente:

$$\Delta_R = \frac{\sum_{i=1}^n \sum_{j=1}^n |x_i - x_j|}{n^2} \quad (7)$$

Se si scartano le differenze tra coppie di modalità per le quali $i = j$, il cui risultato è zero, allora tutti i confronti possibili non sono n^2 ma $n(n-1)$ (il numero delle disposizioni di n oggetti a coppie - di classe 2 - senza ripetizioni). Quindi, la *differenza semplice media senza ripetizione* è data da:

$$\Delta = \frac{\sum_{i \neq j=1}^n |x_i - x_j|}{n(n-1)} = \Delta_R \frac{n}{(n-1)} \quad (8)$$

Entrambi gli indici, Δ e Δ_R , sono nulli se il carattere assume sempre la medesima modalità $\bar{x} = x_i, \forall i = 1, \dots, n$ (zero sarà quindi il valore minimo dell'indice associato, in tal caso, alla minima concentrazione o, se si vuole, alla massima uguaglianza). D'altra parte, se il carattere fosse concentrato in un'unica modalità, cioè:⁸

$$\exists k \in \{1, \dots, n\} \quad \text{per il quale} \quad x_k = n\bar{x} = \sum_{i=1}^n x_i$$

⁶V. [Theil(1967)]

⁷Supponiamo, inoltre, che il carattere assuma solo modalità non negative.

⁸Con \bar{x} si è indicata la media aritmetica.

allora si avrebbe la massima concentrazione possibile. In questo caso tutti gli addendi di Δ (e Δ_R) si annullerebbero, tranne quelli nei quali rientra il valore x_k .⁹ Per il massimo di Δ si ottiene:

$$\Delta_{max} = \frac{2}{n(n-1)} \sum_{i=1}^{n-1} x_k = \frac{2(n-1)x_k}{n(n-1)} = 2\bar{x} \quad (9)$$

Quindi, una misura normalizzata della *differenza semplice media* è data da:

$$\Delta_{norm} = \frac{\Delta}{2\bar{x}} \quad (10)$$

Tale indice normalizzato varia tra 0 (minima concentrazione o perfetta uguaglianza) e 1 (massima concentrazione o perfetta disuguaglianza).

3 La spezzata di Lorenz: il rapporto di concentrazione

La curva di concentrazione (o spezzata di Lorenz¹⁰) è ottenuta mettendo in relazione la quota cumulata del carattere (indicata con Q_i) con la quota cumulata della popolazione (indicata con P_i) ad essa relativa.

Per costruire il grafico della Figura 2 è necessario dapprima ordinare le modalità del carattere in modo crescente $x_{(1)} \leq \dots \leq x_{(i)} \leq \dots \leq x_{(n)}$, dove il pedice tra parentesi è riferito alle modalità ordinate, facendo in modo che anche le unità di rilevazione dalla popolazione, corrispondenti alle diverse modalità, seguano lo stesso ordinamento. Quindi, si procede al calcolo delle quote cumulate P_i e Q_i :

$$P_i = \frac{i}{n} \quad \forall i = 1, \dots, n \quad (11)$$

$$Q_i = \frac{\sum_{j=1}^i x_{(j)}}{\sum_{i=1}^n x_i} = \frac{\sum_{j=1}^i x_{(j)}}{n\bar{x}} \quad \forall i = 1, \dots, n \quad (12)$$

La spezzata della Figura 2 evidenzia con immediatezza quale quota del totale del carattere è imputabile dalla quota della popolazione che lo possiede in misura inferiore rispetto al resto dell'intera popolazione. In particolare, quanto più distante è la linea spezzata dalla retta a 45°, tanto maggiore è la concentrazione del

⁹Le differenze che è necessario considerare nel caso della massima concentrazione sono $n-1$ del tipo $|x_i - x_k|, \forall i \neq k$, più $n-1$ differenze del tipo $|x_k - x_i|, \forall i \neq k$, essendo tutte le altre nulle. Il risultato è conseguenza del fatto che si è posto $x_i = 0, \forall i \neq k$.

¹⁰Introdotta dall'economista americano Otto Max Lorenz agli inizi del 1900 per lo studio della distribuzione dei redditi.

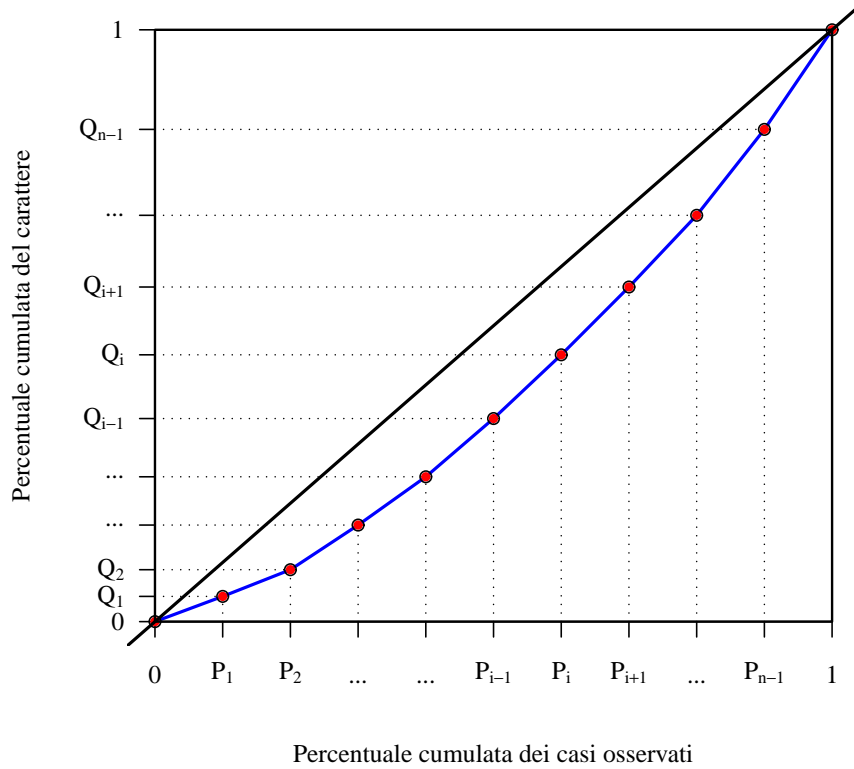


Figura 2: Spezzata di Lorenz

carattere nelle unità più ricche della popolazione (quindi, tanto minore è l'uguaglianza e maggiore è la disuguaglianza). Pertanto, la bisettrice degli assi cartesiani rappresenta la perfetta uguaglianza: il *benchmark* del grado di disuguaglianza. Queste considerazioni conducono alla costruzione del *rapporto di concentrazione di Gini*, G , che è una media ponderata delle differenze normalizzate $\frac{(P_i - Q_i)}{P_i}$:¹¹

$$G = \frac{\sum_{i=1}^{n-1} \left(\frac{P_i - Q_i}{P_i} \right) P_i}{\sum_{i=1}^{n-1} P_i} = \frac{\sum_{i=1}^{n-1} (P_i - Q_i)}{\sum_{i=1}^{n-1} P_i} \quad (13)$$

Si può verificare che la (13) costituisce una misura normalizzata in quanto nel caso di concentrazione nulla (o perfetta uguaglianza) essa assume valore zero

¹¹Notiamo che le differenze $P_i - Q_i$ sono le distanze orizzontali della spezzata dalla retta a 45°. V. [Gini(1914)]

(perché in tal caso le differenze $P_i - Q_i$ sono tutte nulle), mentre assume valore 1 nel caso della concentrazione massima (perfetta disuguaglianza) perché, in questo caso, le Q_i sarebbero tutte nulle per ogni $i \leq n-1$ e il numeratore risulterebbe uguale al denominatore, $\sum_{i=1}^{n-1} P_i$, del rapporto.

Inoltre, la (13) può essere semplificata, dato che:¹² $\sum_{i=1}^{n-1} P_i = \sum_{i=1}^{n-1} \frac{i}{n} = \frac{n-1}{2}$, ottenendo:

$$G = 1 - \frac{\sum_{i=1}^{n-1} Q_i}{\sum_{i=1}^{n-1} P_i} = 1 - \frac{2}{n-1} \sum_{i=1}^{n-1} Q_i \quad (14)$$

Si può dimostrare che il rapporto di concentrazione G coincide con la *differenza semplice media normalizzata* [v. §2, form. (10)]. Ricordando la (12), si può scrivere:

$$\begin{aligned} \sum_{i=1}^{n-1} Q_i &= \frac{1}{n\bar{x}} \sum_{i=1}^{n-1} \left(\sum_{j=1}^i x_{(j)} \right) = \\ &= \frac{1}{n\bar{x}} [(n-1)x_{(1)} + (n-2)x_{(2)} + \dots + x_{(n-1)} + 0x_{(n)}] = \\ &= \frac{1}{n\bar{x}} \sum_{i=1}^n (n-i)x_{(i)} = \frac{1}{n\bar{x}} \left[\sum_{i=1}^n nx_{(i)} - \sum_{i=1}^n ix_{(i)} \right] = \\ &= n - \frac{1}{n\bar{x}} \sum_{i=1}^n ix_{(i)} \end{aligned} \quad (15)$$

e sostituendo questo valore nella (14), si ottiene:

$$\begin{aligned} G &= 1 - \frac{2}{n-1} \left(n - \frac{1}{n\bar{x}} \sum_{i=1}^n ix_{(i)} \right) = \\ &= -\frac{n+1}{n-1} + \frac{2}{n(n-1)\bar{x}} \sum_{i=1}^n ix_{(i)} \end{aligned} \quad (16)$$

Considerando, infine, che la *differenza semplice media normalizzata* può essere scritta come:

¹²Ricodando che $\sum_{i=1}^n i = \frac{n(n+1)}{2}$.

$$\begin{aligned}
\Delta &= \frac{\sum_{i \neq j=1}^n |x_i - x_j|}{n(n-1)2\bar{x}} = \frac{2}{n(n-1)2\bar{x}} \sum_{i=1}^n \sum_{j=1}^i (x_{(i)} - x_{(j)}) = \\
&= \frac{1}{n(n-1)\bar{x}} \left[\left(\sum_{i=1}^n ix_{(i)} \right) - \sum_{i=1}^n \sum_{j=1}^i x_{(j)} \right] = \\
&= \frac{\left\{ \left(\sum_{i=1}^n ix_{(i)} \right) - \left[(x_{(1)}) + (x_{(1)} + x_{(2)}) + \dots + (x_{(1)} + \dots + x_{(n)}) \right] \right\}}{n(n-1)\bar{x}} = \\
&= \frac{1}{n(n-1)\bar{x}} \left[\left(\sum_{i=1}^n ix_{(i)} \right) - \sum_{i=1}^n (n-i+1)x_{(i)} \right] = \frac{2\sum_{i=1}^n ix_{(i)}}{n(n-1)\bar{x}} - \frac{n+1}{n-1}
\end{aligned} \tag{17}$$

Quindi, grazie alla (16) ed alla (17), risulta verificato che $G = \Delta$.

Partendo dalla *spezzata di Lorenz*, si può costruire una misura di concentrazione normalizzata, rapportando l'area compresa tra la retta a 45° e la spezzata con il suo valore massimo (corrispondente alla situazione di massima concentrazione) che è $\frac{n-1}{2n}$ (come si evince immediatamente dalla Figura 3).¹³

L'indice in questione può essere calcolato, seguendo semplici considerazioni geometriche, nel modo seguente:¹⁴

$$\begin{aligned}
&\frac{2n}{n-1} \left\{ \frac{1}{2} - \frac{1}{2} \sum_{i=0}^{n-1} (Q_i + Q_{i+1})(P_{i+1} - P_i) \right\} = \\
&= \frac{n}{n-1} \left(1 - \frac{1}{n} - \frac{2}{n} \sum_{i=1}^{n-1} Q_i \right) = 1 - \frac{2}{n-1} \sum_{i=1}^{n-1} Q_i
\end{aligned} \tag{18}$$

Il risultato della (18) è il *coefficiente di concentrazione di Gini*.

4 L'indice di Herfindahl: una generalizzazione

L'indice di concentrazione di Herfindahl è definito dalla seguente espressione:

$$H = \sum_{i=1}^n \left(\frac{x_i}{\sum_{i=1}^n x_i} \right)^2 \tag{19}$$

¹³Molto spesso per effettuare la normalizzazione si utilizza come denominatore del rapporto il valore $\frac{1}{2}$, anziché $\frac{n-1}{2n}$, questa però è solo un'approssimazione che migliora al crescere del numero delle osservazioni n , perché in tal caso $\frac{n-1}{2n}$ tende ad assumere il valore $\frac{1}{2}$.

¹⁴Tenendo presente la Figura 2 e che $Q_0 = P_0 = 0$, $Q_n = 1$ e $P_{i+1} - P_i = \frac{i+1-i}{n} = \frac{1}{n}$.

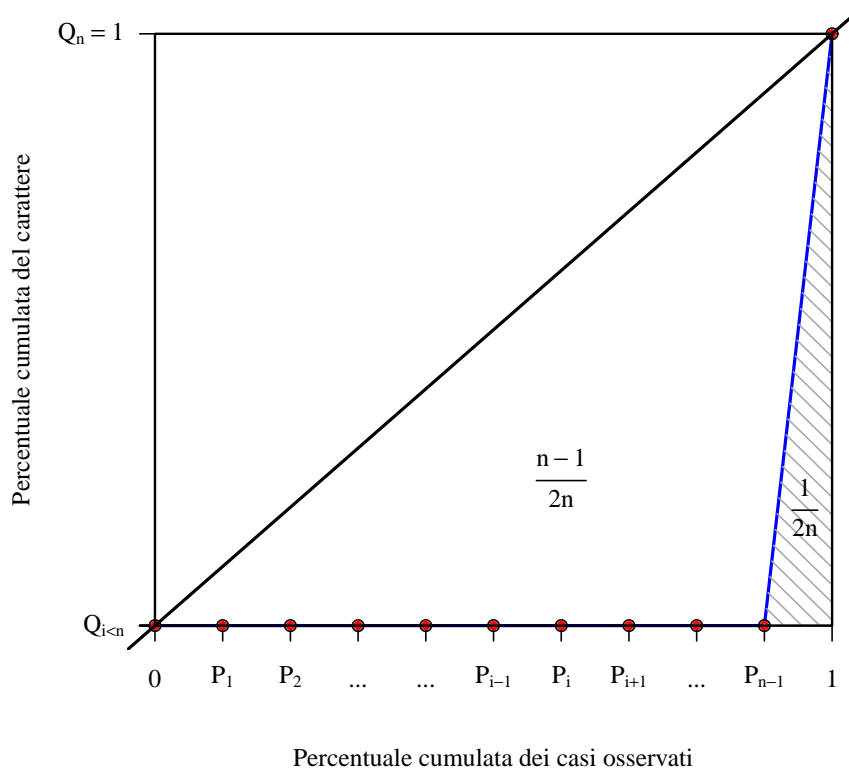


Figura 3: Spezzata di Lorenz nel caso di concentrazione massima

dove x_i indica il valore della modalità del carattere relativa all'osservazione i -esima. Posto $s_i = \frac{x_i}{\sum_{i=1}^n x_i}$, generalmente per H viene utilizzata la notazione più sintetica: $H = \sum_{i=1}^n s_i^2$. L'indice varia tra $\frac{1}{n}$ (nel caso di minima concentrazione o perfetta uguaglianza: $s_i = \frac{1}{n}, \forall i = 1, \dots, n$) e 1 (nel caso di massima concentrazione o perfetta disuguaglianza: se $s_k = 1$ per una sola osservazione e zero per tutte le altre).¹⁵ Inoltre, si può verificare immediatamente che valgono le seguenti uguaglianze, ordinatamente per la somma e per la media degli s_i :

$$\sum_{i=1}^n s_i = 1 \tag{20}$$

¹⁵Quindi una versione normalizzata dell'indice di Herfindahl è costituita dal rapporto $\frac{nH-1}{n-1} = \frac{CV^2}{n-1}$, dove l'ultimo rapporto è il *Coefficiente di variazione normalizzato*, v. successiva (24).

$$\bar{s} = \frac{1}{n} \sum_{i=1}^n s_i = \frac{1}{n} \quad (21)$$

Grazie alla (20) e alla (21), si può dimostrare che l'indice di Herfindal è determinato tanto dalla dispersione del carattere rispetto alla media della popolazione (o nel campione), quanto dalla numerosità di quest'ultima (o di quest'ultimo). Considerato che la varianza delle quote s_i è data da:

$$\sigma_{s_i}^2 = \frac{1}{n} \sum_{i=1}^n (s_i - \bar{s})^2 = \frac{1}{n} \sum_{i=1}^n \left(s_i^2 - \frac{2}{n} s_i + \frac{1}{n^2} \right) = \frac{1}{n} \left(H - \frac{1}{n} \right) \quad (22)$$

otteniamo:

$$H = \sum_{i=1}^n s_i^2 = n\sigma_{s_i}^2 + \frac{1}{n} \quad (23)$$

Ora, dato che $n\sigma_{s_i}^2$ è la *devianza*, l'indice aumenta al crescere della dispersione e si riduce al crescere di n . Inoltre, la (23) può essere riscritta esplicitando la relazione tra H e il *coefficiente di variazione* $CV = \frac{\sigma_{s_i}}{\bar{s}}$:

$$H = n\sigma_{s_i}^2 + \frac{1}{n} = \frac{1}{n} \left(\frac{\sigma_{s_i}^2}{\frac{1}{n^2}} + 1 \right) = \frac{CV^2 + 1}{n} \quad (24)$$

Una proprietà rilevante dell'indice di Herfindahl è che esso corrisponde alla pendenza media normalizzata dei segmenti che compongono la *spezzata di concentrazione di Lorenz* [v. §3]. Infatti, utilizzando la notazione introdotta in questo paragrafo, la pendenza media della spezzata di Lorenz è data dalla formula:

$$\sum_{i=1}^n \frac{s_i}{\frac{1}{n}} s_i = n \sum_{i=1}^n s_i s_i = nH \quad (25)$$

il valore di nH varia tra 1 (concentrazione minima) a n (concentrazione massima), per cui segue l'interpretazione di H come pendenza media normalizzata della curva di concentrazione (con campo di variazione $[\frac{1}{n}, 1]$):

$$H = \frac{1}{n} \sum_{i=1}^n \frac{s_i}{\frac{1}{n}} s_i \quad (26)$$

Coerentemente con questo risultato, il reciproco di H può essere interpretato come *indice di equivalenza numerica*: dato il valore di H , ci si chiede, in tal caso, quale sarebbe il numero di osservazioni N che, in condizioni di minima concentrazione (o perfetta uguaglianza: $H = \frac{1}{n}$) darebbe luogo allo stesso valore

di H . Tale ragionamento può essere generalizzato rispetto ad una intera classe di misure di concentrazione in funzione di (dipendenti da) un parametro α :¹⁶

$$R(s_1, \dots, s_n; \alpha) = \sum_{i=1}^n s_i^\alpha \quad (27)$$

Nel caso di assenza di concentrazione (o di perfetta uguaglianza o, ancora, di concentrazione minima) risulta:

$$R(\alpha) = \sum_{i=1}^n s_i^\alpha = \frac{n}{n^\alpha} = n^{1-\alpha} \quad (28)$$

per cui assegnato un certo valore ad R (calcolato) è possibile ricavare un *indice di equivalenza numerica* (teorico) nel modo seguente:

$$N(\alpha) = R^{\frac{1}{1-\alpha}} = \left(\sum_{i=1}^n s_i^\alpha \right)^{\frac{1}{1-\alpha}} \quad (29)$$

Per $\alpha = 2$ la formula precedente è uguale al reciproco dell'indice di Herfindahl, per $\alpha = 1$ invece l'indice non è definito ma si può ancora ottenere l'indice di equivalenza numerica per $\alpha \rightarrow 1$. Prendendo il logaritmo di ambo i membri della (29):

$$\log N(\alpha) = \frac{1}{1-\alpha} \log \sum_{i=1}^n s_i^\alpha \quad (30)$$

Il limite della funzione $\log n(\alpha)$ per $\alpha \rightarrow 1$, essendo di forma indeterminata, può essere calcolato applicando le regole di l'Hôpital:

$$\lim_{\alpha \rightarrow 1} \log N(\alpha) = \lim_{\alpha \rightarrow 1} -\frac{1}{\sum_{i=1}^n s_i^\alpha} \sum_{i=1}^n s_i^\alpha \log s_i = -\sum_{i=1}^n s_i \log s_i \quad (31)$$

Quindi, per α che tende a 1, l'*indice di equivalenza numerica* è:

$$N = e^{-\sum_{i=1}^n s_i \log s_i} \quad (32)$$

giustificando così le considerazioni svolte in nota alla fine del §1 (e questa è un'argomentazione costruttiva per l'*indice di entropia* alternativa, e speculare, alla derivazione dagli assiomi proposti nel §1).

¹⁶V. [Hannah and Kay(1977)]

Sommario degli indici di concentrazione trattati

Nella seguente tabella si riportano le versioni normalizzate degli indici esposti nei precedenti paragrafi.

<i>Indice</i>	<i>Formula</i>	<i>Intervallo</i>	<i>Minim. Conc.</i>	<i>Mass. Conc.</i>
S	$-\frac{\sum_{i=1}^n p_i \log p_i}{\log n}$	$[0, 1]$	1	0
Δ_{norm}	$\frac{1}{2\bar{x}} \frac{\sum_{i \neq j=1}^n x_i - x_j }{n(n-1)}$	$[0, 1]$	0	1
G	$1 - \frac{2}{n-1} \sum_{i=1}^{n-1} Q_i$	$[0, 1]$	0	1
H	$\sum_{i=1}^n \left(\frac{x_i}{\sum_{i=1}^n x_i} \right)^2$	$[\frac{1}{n}, 1]$	$\frac{1}{n}$	1
$R(\alpha > 1)$	$\sum_{i=1}^n s_i^\alpha$	$[n^{1-\alpha}, 1]$	$n^{1-\alpha}$	1
$R(0 < \alpha < 1)$	$\sum_{i=1}^n s_i^\alpha$	$[1, n^{1-\alpha}]$	$n^{1-\alpha}$	1
$N(0 < \alpha \neq 1)$	$(\sum_{i=1}^n s_i^\alpha)^{\frac{1}{1-\alpha}}$	$[1, n]$	n	1
$N(\alpha \rightarrow 1)$	$e^{-\sum_{i=1}^n s_i \log s_i}$	$[1, n]$	n	1
$N(\alpha = 2)$	$N = \frac{1}{H}$	$[1, n]$	n	1

Tabella 1: Tabella degli indici di concentrazione

Legenda della tabella Tabella 1: S indice di entropia, Δ_{norm} differenza semplice media normalizzata, G indice di Gini, H indice di Herfindahl, $R(\alpha)$ indice generale di concentrazione, $N(\alpha)$ indice di equivalenza numerica.

Riferimenti bibliografici

- [Baumgärtner(2004)] S. Baumgärtner. *Diversity as a potential for surprise*. University of Heidelberg, Germany, 2004.
- [Gini(1914)] C. Gini. Sulla misura della concentrazione e della variabilità dei caratteri. *Atti del R. Istituto Veneto di Scienze, Lettere e Arti*, 1914.
- [Hannah and Kay(1977)] L. Hannah and J. Kay. *Concentration in Modern Industry. Theory, measurement and the UK experience*. Macmillan, London, 1977.
- [Khinchin(1957)] A. I. Khinchin. *Mathematical Foundations of Information Theory*. Dover, New York, 1957.
- [Piccolo(1998)] D. Piccolo. *Statistica*. Il Mulino, Bologna, 1998.
- [Shannon(1948)] C. E. Shannon. A mathematical theory of communication. *Bell System Technical Journal*, 1948.
- [Theil(1967)] H. Theil. *Economics and Information Theory*. Rand McNally and Company, Chicago, 1967.

Indice

Introduzione	1
1 L'indice di entropia	1
2 La differenza semplice media	4
3 La spezzata di Lorenz: il rapporto di concentrazione	5
4 L'indice di Herfindahl: una generalizzazione	8
Sommario degli indici di concentrazione trattati	12
Riferimenti bibliografici	13

Elenco delle figure

1	Funzione $h(p)$	3
2	Spezzata di Lorenz	6
3	Spezzata di Lorenz nel caso di concentrazione massima	9

Elenco delle tabelle

1	Tabella degli indici di concentrazione	12
---	--	----