

PER IL CENTESIMO ANNIVERSARIO DELLA TEORIA DELLE RELAZIONI STATISTICHE DI CORRADO GINI

Nota del m.e. EUGENIO REGAZZINI (*) e del s.c. DONATO CIFARELLI (**)

(Adunanza del 5 febbraio 2015)

SUNTO. – A cento anni dalla pubblicazione dei primi lavori di Corrado Gini sulle relazioni statistiche e sulla misura della loro intensità, questa Nota si propone di chiarire il valore pionieristico di tali lavori e di ricordarne alcuni sviluppi, particolarmente significativi, in tema di metrizzazione di spazi di distribuzioni di probabilità e in tema di analisi della dipendenza monotona fra le componenti di un vettore aleatorio.

* * *

ABSTRACT. – This Note aims at highlighting Gini's contributions to definition and measurement of the intensity of a statistical relationship, on the occasion of the centennial anniversary of the publication of his early papers on that topic. The Note stresses the precursory value of those contributions and mentions some of their most significant developments apropos of metrization of spaces of probability distributions and analysis of monotone dependence.

1. Nel triennio 1914-1916 Corrado Gini tracciò le linee generali di una metodologia unitaria per lo studio di alcune proprietà, particolarmente significative, di relazioni statistiche che legano due caratteri comuni a tutte le unità di una popolazione. Cfr. [15]-[18]. La formulazione di una relazione statistica scaturisce, in generale, dalla combinazione di un complesso di conoscenze e opinioni *a priori* (o iniziali) con l'osservazione dei

(*) Dipartimento di Matematica "F. Casorati", Università degli Studi di Pavia, Italia. E-mail: eugenio.regazzini@unipv.it

(**) Università "L. Bocconi", Milano, Italia. E-mail: michele.cifarelli@unibocconi.it

due caratteri in ciascuna delle unità che compongono la popolazione o parte di essa. Quindi, la sua espressione più comune sarà nella forma di funzione aleatoria governata da una legge di probabilità che, in casi estremi, potrà anche ridursi ad una relazione rigida del tipo delle cosiddette *leggi deterministiche*. Il valore di leggi siffatte, che non intendono esprimere alcunché di necessario ed immutabile, è legato all'impiego che se ne può fare per formulare previsioni probabilistiche, le uniche ad essere dotate di senso effettivo. Come detto sopra, alla base di una legge statistica vi è generalmente anche un processo di osservazione in relazione al quale si pone il problema, non secondario, di come elaborare i dati, che da tale processo scaturiscono, in modo da estrarne l'informazione necessaria e sufficiente in vista di qualche obiettivo prefissato. Il contributo di Gini, che ci accingiamo a rievocare, è ascrivibile a quest'ultimo ordine di studi (detti di statistica descrittiva) nell'intento duplice di predisporre strumenti idonei per: l'analisi della monotonia globale della relazione statistica fra due caratteri ordinati (*concordanza* o *discordanza*, nella terminologia di Gini, a seconda che s'indaghi sul crescere o sul decrescere della relazione); lo studio della discrepanza (*connessione*, nel linguaggio di Gini) tra la distribuzione empiricamente accertata della coppia di caratteri e quella che si avrebbe se essi si combinassero fra loro secondo quanto prescritto dalla condizione d'*indipendenza in distribuzione* (condizione d'*indipendenza stocastica* col linguaggio delle probabilità). A commento della distinzione proposta, in [15] sostiene: «Se non vi è connessione tra le modalità di due caratteri, non vi può essere tra queste né concordanza, né discordanza; essendovi invece connessione, potrà esservi concordanza rispetto a tutte le modalità, oppure discordanza rispetto a tutte le modalità...L'esame della concordanza...appare dunque come una ricerca subordinata all'esame della connessione». Confrontando le analisi che il Gini fa dei due succitati temi, quella dedicata alla monotonia della relazione statistica sembra essere la più interessante, e bisognosa di aggiustamenti che ne rendano più soddisfacenti tutte le conclusioni. Infatti il programma del Gini, chiaro e trasparente nell'indicazione degli obiettivi, richiede una tipologia di strumenti tecnici e conoscenze i quali, ancorché elementari per un probabilista o statistico moderno, non potevano essere, un secolo fa, in possesso dell'Autore. Della definizione di uno di tali strumenti egli si occupa diffusamente nel primo dei lavori citati all'inizio, a proposito della misura della connessione. Si allude all'indice di *dissomiglianza* che, riscoperto anche fuori dal campo della statistica, e in forme più generali di quelle indicate dal Gini, è divenuto in anni recenti oggetto di

indagine approfondita soprattutto da parte di cultori dell'analisi matematica. Di questo parleremo nell'ultimo paragrafo, dopo aver diffusamente trattato dell'aspetto della monotonia di una relazione statistica e della sua quantificazione. In mezzo, tra questa presentazione introduttiva e l'inizio della descrizione dei lavori centenari da riportare all'attenzione, diamo alcune notizie sulla vita di Corrado Gini e sulla sua figura scientifica.

2. Nato a Motta di Livenza (Treviso) il 23 maggio 1884, si laurea in giurisprudenza a Bologna con una tesi dal titolo *Il sesso dal punto di vista statistico*, che consegue il premio Vittorio Emanuele per le scienze sociali e politiche (1907) alla stessa Università di Bologna. Notevolmente ampliata, la tesi è pubblicata nel 1908 in un volume di ben cinquecento pagine recante lo stesso titolo seguito dal sottotitolo, molto appropriato in questo caso: *Le leggi della produzione dei sessi*. Cfr. [12]. L'approfondimento del tema conduce il Nostro allo studio della probabilità e, segnatamente, a ripensarne il valore in relazione alle applicazioni statistiche, pervenendo ad una concezione radicalmente frequentista alla quale resterà fedele fino alla fine. Ne tratta in un certo numero di lavori a stampa e, soprattutto, nel congresso della Società Filosofica Italiana tenutosi a Parma nel 1907, di fronte ad eminenti esponenti della matematica e della fisica, sotto l'autorevole presidenza di Federigo Enriques (1871-1946). Come egli stesso ebbe a ricordare in occasione delle celebrazioni per il suo ottantesimo anniversario (v. [20]), fu proprio l'impressione suscitata durante i lavori di quel congresso ad appianargli la strada verso una rapida ascesa alla cattedra universitaria di statistica (Cagliari, 1910). All'atto della nomina, oltre ai lavori sulla probabilità, può vantare articoli di notevole rilevanza sull'analisi della variabilità e della concentrazione dei caratteri trasferibili (reddito, ricchezza, ad esempio), che, sviluppati negli anni del periodo cagliaritano (1909-1913), gli daranno duratura fama internazionale. Inoltre, nel 1911 pubblica un lavoro sul rapporto dei sessi alla nascita [13] nel quale fa uso di una originale metodologia che oggi si direbbe bayesiana-empirica, partendo, in modo pionieristico, da una distribuzione iniziale di tipo beta. Già questa scelta costituisce una notevole innovazione rispetto al punto di vista tradizionale sul problema delle probabilità inverse. Si trasferisce a Padova nel 1913, dove intraprende, tra altre cose, gli studi sulle relazioni statistiche e promuove di un attivo laboratorio di statistica. Nel 1925, all'apice della carriera accademica e di una prestigiosa notorietà internazionale, si trasferisce a Roma dove, oltre alla cattedra di Politica e Statistica Economica, è

chiamato dal Governo a riformare l'intero sistema statistico italiano. L'incarico ha per esito la costituzione dell'Istituto Centrale di Statistica, nel 1926, che Gini presiede fino al 1932, insieme ad altre importanti posizioni in organismi ufficiali nazionali e internazionali. A quel tempo, la cosiddetta scuola statistica italiana, con Gini ormai identificata, incomincia a chiudersi ai nuovi indirizzi che, provenendo dal Regno Unito, in pochi anni raggiungono posizione predominante in Europa, nella parti più avanzate dell'impero britannico e, soprattutto, negli Stati Uniti. All'origine della chiusura italiana si possono vedere motivazioni diverse tra loro. Da una parte vi è la giusta esigenza di salvaguardare taluni principi di logica statistica dei quali l'indirizzo anglosassone manifestamente non si cura. Dall'altra parte vi è però una questione di visione e competenza matematica, a sfavore degli statistici italiani, che facilita invece gli angloamericani nel dare soluzioni ardite e attraenti a problemi di ordine pratico avvertiti, con modalità non molto dissimili, sia da cultori di scienze naturali o sociali, sia da decisori nei campi dell'economia, della gestione e della tecnologia. Il divario si rende palese alla fine del secondo conflitto mondiale e la Facoltà di Statistica di Roma - presieduta dal Gini fin dalla fondazione (1936), con breve interruzione per il processo di epurazione a cui è sottoposto - cerca di farvi fronte con la chiamata di Giuseppe Pompilj (1913-1968) a ricoprire una cattedra di geometria. Il Pompilj, infatti, pur avendo alle spalle un'attività di ricerca prevalentemente nel campo della geometria algebrica sotto la guida di Enriques, si è appassionato alla statistica matematica d'impronta anglosassone grazie a letture spontanee che, da prigioniero di guerra, ha potuto compiere in un campo britannico dislocato in India. A Pompilj riesce, nonostante la prematura scomparsa, di dar vita ad un gruppo di ricerca statistica, opportunamente collegato ad ambienti scientifici internazionali, che, fra altre cose, si dedica anche allo sviluppo di antiche idee giniane, ma da un punto di vista e con un linguaggio più aggiornati di quanto non siano quelli della scuola statistica italiana tradizionale. Fra i risultati più significativi si possono citare quelli di Dall'Aglio sugli estremi di certi momenti definiti per gli elementi di una classe di Fréchet. Cfr. [7]. Da parte sua, Gini scrive numerosi saggi e libri anche nel dopoguerra, spaziando dalla statistica alla demografia, dalla sociologia all'economia ed alla politica economica. Si vedano il necrologio [4] e il volume [3] per una valutazione complessiva della produzione scientifica di questa fase. Per quanto riguarda la statistica, si limita a riproporre con stile e linguaggio immutati, e dunque ormai estranei alle nuove generazioni, alcuni dei lavori giovanili. Si spegne a Roma, improvvisamente, il 13 marzo 1965.

3. Ritornando ai lavori di statistica, si lascia ancora apprezzare l'originalità di certi contenuti, che, presentati con linguaggio moderno e strumenti matematici adeguati, potrebbero divenire sorgente di ispirazione per nuove ricerche e, soprattutto, stimolo a rivedere posizioni cristallizzate, da ormai troppo tempo accettate acriticamente. Con questo spirito diamo inizio alla descrizione dei contributi di Gini allo studio della monotonia di una relazione statistica.

Elemento primo di quasi tutte le considerazioni che verranno svolte è il concetto di *distribuzione* (di frequenza, generalmente) su un sottoinsieme Ω di \mathbb{R}^2 , inteso come insieme di tutte le coppie ordinate di numeri reali, munito di una topologia equivalente a quella euclidea. Più propriamente, Ω è da pensarsi come boreliano di \mathbb{R}^2 . La suddetta distribuzione di frequenze, φ' in simboli, è da pensarsi come funzione reale, numericamente additiva sulla classe $\mathcal{D} := \{\Omega \cap B : B \in \mathcal{B}(\mathbb{R}^2)\} :=$ classe di tutti i boreliani di \mathbb{R}^2 , e tale che $\varphi'(\Omega) = 1$. Spesso, φ' , appositamente definita in modo da specificare che le coppie (x, y) logicamente possibili sono tutte e soltanto quelle contenute in Ω , verrà sostituita con φ , definita su $\mathcal{B}(\mathbb{R}^2)$ da

$$\varphi(B) := \varphi'(B \cap \Omega) \quad (B \in \mathcal{B}(\mathbb{R}^2)).$$

Ovviamente, si ha $\varphi(\Omega^c) = 0$ dove, per un generico insieme A , con A^c si denota il complementare, rispetto a \mathbb{R}^2 nel nostro caso. È noto che una qualunque φ su $\mathcal{B}(\mathbb{R}^2)$ è completamente caratterizzata dalla *funzione di ripartizione* (f.r., per brevità), intesa come funzione reale definita su \mathbb{R}^2 da

$$F(x, y) := \varphi((-\infty, x] \times (-\infty, y]) \quad ((x, y) \in \mathbb{R}^2).$$

Dalla conoscenza di φ [F , rispettivamente] associata ad una data coppia di caratteri quantitativi si estrae subito quella delle distribuzioni [f.r., rispettivamente] marginali: φ_1 [F_1 , rispettivamente] per il primo carattere, φ_2 [F_2 , rispettivamente] per il secondo. Infatti,

$$\begin{aligned} \mathcal{B}(\mathbb{R}) \ni A' &\mapsto \varphi_1(A') = \varphi(A' \times \mathbb{R}), \\ R \ni x &\mapsto F_1(x) = \lim_{y \rightarrow +\infty} F(x, y); \\ \mathcal{B}(\mathbb{R}) \ni A'' &\mapsto \varphi_2(A'') = \varphi(\mathbb{R} \times A''), \\ R \ni y &\mapsto F_2(y) = \lim_{x \rightarrow +\infty} F(x, y). \end{aligned}$$

Punto fermo dell'impostazione giniana è che l'analisi della monotonia della relazione statistica fra due caratteri vada fatta nell'ambito delle distribuzioni bidimensionali con marginali uguali a quelle dei due caratteri presi in considerazione. In altri termini, Gini è consapevole del fatto che la possibilità di raggiungere il massimo assoluto di concordanza o discordanza dipende dalla coppia di marginali assegnate, che potrebbero anche non consentirlo. Perciò, la valutazione del grado di monotonia proprio di una specifica distribuzione su \mathbb{R}^2 con marginali assegnate andrà fatta in termini relativi con riferimento al massimo consentito da quelle marginali. Posta in questi termini, la questione suggerisce di procedere con l'ausilio della corrispondente classe di Fréchet, dal ben noto matematico francese Maurice R. Fréchet (1878-1973) che ne studiò le proprietà più idonee a risolvere anche i problemi di Gini: cfr. [11]. In realtà, la definizione di tali classi era già stata data da Wassily Hoeffding (1914-1991) in un lavoro risalente al 1940 (cfr. [21]) e prima ancora era stata chiaramente considerata in un lavoro poco noto (cfr. [8], Sezione 8) di Bruno de Finetti (1906-1985). Date le distribuzioni non degeneri φ_1 e φ_2 , su \mathbb{R} , la classe di Fréchet $\mathcal{F}(\varphi_1, \varphi_2)$ [o anche, $\mathcal{F}(F_1, F_2)$] da esse generata, nell'ordine dato, è l'insieme di tutte le distribuzioni su \mathbb{R}^2 che ammettono φ_1 come prima marginale e φ_2 come seconda marginale. Si noti che $\mathcal{F}(\varphi_1, \varphi_2) \neq \mathcal{F}(\varphi_2, \varphi_1)$ se $\varphi_1 \neq \varphi_2$. La classe è non vuota, infatti la distribuzione corrispondente all'indipendenza dei due caratteri (data da $\varphi_1 \otimes \varphi_2$) banalmente vi appartiene. Ma essa ne contiene altre due particolarmente importanti ai nostri fini, e cioè \underline{h} e \overline{h} generate, rispettivamente, dalle f.r.

$$\begin{aligned} \underline{H}(x, y) &:= (F_1(x) + F_2(y) - 1) \vee 0 && \text{e} \\ \overline{H}(x, y) &:= F_1(x) \wedge F_2(y) && ((x, y) \in \mathbb{R}^2). \end{aligned}$$

Per incominciare, esse rappresentano gli elementi minimo e massimo di $\mathcal{F}(\varphi_1, \varphi_2)$, nel senso che, per ogni (x, y) in \mathbb{R}^2 , si ha

$$\underline{H}(x, y) \leq F(x, y) \leq \overline{H}(x, y).$$

Questo fatto si ricollega all'idea di *ordinamento* (parziale) di *concordanza* (o *discordanza*) fra due distribuzioni aventi marginali uguali, che Gini mostra di avere ben chiaro in mente. Si tratta di questo: se $\varphi^{(1)}$ e $\varphi^{(2)}$ sono elementi distinti di $\mathcal{F}(\varphi_1, \varphi_2)$ le cui f.r. $F^{(1)}$ e, rispettivamente, $F^{(2)}$ soddisfano

$$F^{(1)}(x, y) \leq F^{(2)}(x, y)$$

per ogni (x, y) di \mathbb{R}^2 , essendo necessariamente diverse per qualche coppia, allora vale anche $\varphi^{(1)}((x, +\infty) \times (y, +\infty)) \leq \varphi^{(2)}((x, +\infty) \times (y, +\infty))$ per ogni (x, y) in \mathbb{R}^2 , etc., e ciò indica inequivocabilmente che $\varphi^{(2)}$ si concentra, più di quanto non faccia $\varphi^{(1)}$, nelle regioni di \mathbb{R}^2 in cui grandi (piccoli, rispettivamente) valori del primo carattere si associano a grandi (piccoli, rispettivamente) valori del secondo. In modo diverso, tale circostanza si esprime dicendo che $\varphi^{(2)}$ è *più concordante* di $\varphi^{(1)}$, o che $\varphi^{(1)}$ è *più discordante* di $\varphi^{(2)}$. Applicato questo ordinamento a $\mathcal{F}(\varphi_1, \varphi_2)$, si conclude che, rispetto ad esso, \bar{h} è l'elemento più concordante e \underline{h} l'elemento più discordante tra tutti gli elementi di $\mathcal{F}(\varphi_1, \varphi_2)$. Lo studio del supporto (topologico) di questi elementi estremi chiarisce ulteriormente in che senso essi siano associabili alle relazioni statistiche monotone più strette compatibili con una data coppia di marginali. Si ricorda che il supporto topologico di una distribuzione è inteso come insieme (chiuso) di tutti i punti in ogni intorno dei quali essa depona massa positiva; in simboli, il supporto di μ sarà indicato con $S(\mu)$. Ritornando alla generica φ in $\mathcal{F}(\varphi_1, \varphi_2)$, ad ogni x e y nel supporto di φ_1 e, rispettivamente, di φ_2 si associano gli insiemi

$\mathcal{L}_x(\varphi) :=$ insieme delle coppie (x, v) contenute nel supporto di φ ,

$\mathcal{L}^y(\varphi) :=$ insieme delle coppie (u, y) appartenenti al supporto di φ .

Nel caso degli elementi estremi, Fréchet pone in evidenza le proprietà seguenti: $\mathcal{L}_x(\bar{h})$, $\mathcal{L}^y(\bar{h})$, $\mathcal{L}_x(\underline{h})$ e $\mathcal{L}^y(\underline{h})$ sono limitati; inoltre, se x_1, x_2 sono nel supporto di φ_1 e y_1, y_2 in quello di φ_2 e soddisfano le disuguaglianze $x_1 < x_2, y_1 < y_2$, si ha

$$(1) \quad \begin{aligned} \sup \mathcal{L}_{x_1}(\bar{h}) &\leq \inf \mathcal{L}_{x_2}(\bar{h}) \\ \sup \mathcal{L}^{y_1}(\bar{h}) &\leq \inf \mathcal{L}^{y_2}(\bar{h}) \end{aligned}$$

e

$$(2) \quad \begin{aligned} \sup \mathcal{L}_{x_2}(\underline{h}) &\leq \inf \mathcal{L}_{x_1}(\underline{h}) \\ \sup \mathcal{L}^{y_2}(\underline{h}) &\leq \inf \mathcal{L}^{y_1}(\underline{h}). \end{aligned}$$

Di conseguenza, i supporti di \bar{h} e di \underline{h} si possono inscrivere in due curve l'una non decrescente e l'altra non crescente, generalmente non riconducibili a grafici di funzioni in senso proprio. Chiariamo. Se φ_1 e φ_2 sono uguali, la *dipendenza funzionale* nei casi estremi si presenta, come pure

quando F_1 e F_2 hanno supporto convesso e sono ivi continue, ma, fuori da queste situazioni, anche gli estremi della classe di Fréchet potrebbero associare ad una determinazione del primo carattere più determinazioni del secondo, etc., dando luogo ad un oggetto che Fréchet propone di chiamare *funzione statisticamente monotona* (f.s.m., per brevità), non decrescente nel caso di \bar{h} e non crescente in quello di \underline{h} . Sempre Fréchet ne suggerisce immagini geometriche la cui scelta, in larga parte convenzionale, non ha comunque ripercussioni sulla parte sostanziale dei ragionamenti. Ricorrendo alla rappresentazione cartesiana, si pongono in ascissa le determinazioni contenute nel supporto del primo carattere e in ordinata quelle del secondo. La rappresentazione geometrica del supporto di ciascuna delle distribuzioni estreme non presenta problemi in corrispondenza alle coppie (x, y) in cui x e y sono punti *interni* del supporto di φ_1 e, rispettivamente, del supporto di φ_2 . Altrimenti, esisteranno almeno due punti, in almeno uno dei suddetti supporti, che risultano essere estremi di un intervallo aperto privo di punti dello stesso supporto. Per fissare le idee, verifichino questa circostanza i punti x_1 e x_2 con $x_1 < x_2$ contenuti nel supporto di φ_1 . Allora, a proposito della rappresentazione del supporto di \bar{h} si congiungono con un segmento rettilineo (non decrescente, per le (1)) i punti $(x_1, \sup \mathcal{L}_{x_1}(\bar{h}))$ e $(x_2, \inf \mathcal{L}_{x_2}(\bar{h}))$. Inoltre, in corrispondenza ad ogni x nel supporto di φ_1 , tale che $\inf \mathcal{L}_x(\bar{h}) < \sup \mathcal{L}_x(\bar{h})$, si congiungono con un segmento rettilineo verticale gli estremi $(x, \inf \mathcal{L}_x(\bar{h}))$ e $(x, \sup \mathcal{L}_x(\bar{h}))$. Si ottiene così un insieme “continuo” di punti disposti in modo da far pensare ad una curva non decrescente $\bar{\Gamma}$ (cfr. Figura 1) da interpretarsi come diagramma della f.s.m. non decrescente corrispondente al massimo della classe di Fréchet $\mathcal{F}(\varphi_1, \varphi_2)$ ¹.

Tenendo conto delle (2), in modo analogo si può costruire la curva $\underline{\Gamma}$ (cfr. Figura 2) della f.s.m. non crescente associata all'elemento minimo \underline{h} della stessa classe.

Le curve $\bar{\Gamma}$ e $\underline{\Gamma}$ contengono i supporti di \bar{h} e, rispettivamente, \underline{h} , in modo che la scelta di tali costruzioni, fra tante altre possibili, non influisca sulle conclusioni di eventuali ragionamenti che dipendessero dai supporti effettivi.

¹Come spiegato diffusamente nel Capitolo 2 di [30], si tratta del grafico del *sub-differenziale* di una funzione convessa

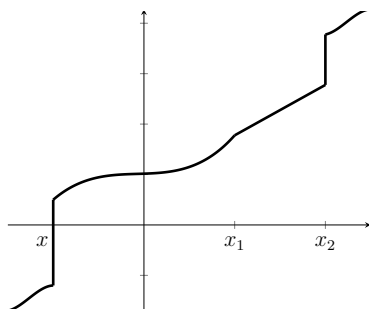


Figura 1 — La curva $\bar{\Gamma}$. In ascissa si legga il supporto di φ_1 , in ordinata quello di φ_2 . Diagramma della f.s.m. associata ad \bar{h} , in cui (x_1, x_2) non contiene punti del supporto di φ_1 .

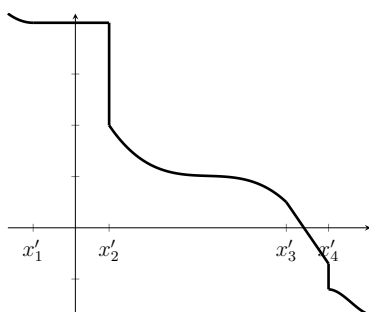


Figura 2 — La curva $\underline{\Gamma}$. In ascissa si legga il supporto di φ_1 , in ordinata quello di φ_2 . Diagramma della f.s.m. associata ad \underline{h} , in cui (x'_1, x'_2) e (x'_3, x'_4) non contengono punti del supporto di φ_1 .

4. Si hanno ora gli elementi per definire un indice sommario, idoneo a segnalare se un generico elemento φ di $\mathcal{F}(\varphi_1, \varphi_2)$ sia vicino ad \bar{h} [\underline{h} , rispettivamente] e, di conseguenza, i due caratteri aventi φ per distribuzione si possano pensare legati statisticamente in senso monotono non decrescente [non crescente, rispettivamente]. Il problema si riduce in parte a come valutare questa vicinanza. L'idea che proporremo per risolverlo è della stessa natura di quella originaria di Gini, ma se ne distacca sotto il profilo della realizzazione tecnica. Noi, infatti, ci baseremo esclusivamente sulle proprietà della classe di Fréchet, evitando di introdurre, ad

esempio, il concetto di *contrario* di un carattere, che non sembra idoneo alla scopo fuori dai casi in cui la f.s.m. è (almeno approssimativamente) lineare. Partiamo dunque dalla valutazione dell'allontanamento di φ da \underline{h} . Lo "scarto" di un punto $P \equiv (x, y)$ appartenente a $S(\varphi_1) \times S(\varphi_2)$ da \underline{h} si può ragionevolmente apprezzare considerando

$$\sigma_g(P, \underline{\Gamma}) := \inf\{g(d(P, Q)) : Q \in \mathcal{L}^y(\underline{\Gamma})\} \\ + \inf\{g(d(P, Q)) : Q \in \mathcal{L}_x(\overline{\Gamma})\}$$

dove: g è una funzione definita su \mathbb{R} , ivi continua, non decrescente e convessa tale che $g(x) = 0$ per $x \leq 0$ e $g(x) > 0$ per $x > 0$; d denota la distanza euclidea su \mathbb{R}^2 . La ragione per cui σ_g è definita come somma di due diversi tipi di "scarto" risiede nell'opportunità di tener conto complessivamente degli scarti (generalmente diversi) connessi a ciascuna delle due interpretazioni causali che si potrebbero dare di una relazione statistica. Per quanto concerne la scelta di g , si può dire che essa andrebbe fatta in relazione alle conseguenze (quando si sappiano valutare) dello scarto "puro" corrispondente all'identità $g(x) = x$ per $x > 0$. Ad esempio, nella teoria della correlazione è frequente l'uso di $g(x) = x^2$ per $x > 0$. Tuttavia, una misura utile della discrepanza di g da \underline{h} dovrà tenere conto, in maniera riassuntiva, di tutti gli scarti $\sigma_g(P, \underline{\Gamma})$ al variare di P , e del fatto che il punto P è distribuito secondo φ . Quindi, un indicatore conveniente della proprietà considerata è dato da

$$(3) \quad c_g(\varphi) := \int_{S(\varphi)} \sigma_g(P, \underline{\Gamma}) \varphi(dP) \quad (\varphi \in \mathcal{F}(\varphi_1, \varphi_2)).$$

L'integrale è ben definito, grazie alle precisate proprietà di g e alla definizione di σ_g , ma potrebbe essere infinito. È fondamentale, poi, osservare che $P \mapsto \sigma_g(P, \underline{\Gamma})$ è continua e *quasi monotona* sull'interno D dell'involucro convesso di $S(\varphi_1) \times S(\varphi_2)$, nel senso che se $(x_i, y_i) \in D$ per $i = 1, 2$ con $x_1 < x_2$ e $y_1 < y_2$, allora

$$\sigma_g((x_2, y_2), \underline{\Gamma}) - \sigma_g((x_2, y_1), \underline{\Gamma}) \\ - \sigma_g((x_1, y_2), \underline{\Gamma}) + \sigma_g((x_1, y_1), \underline{\Gamma}) \geq 0.$$

Cfr. [5].

Perciò, grazie al Teorema 1 in [2], si può enunciare la

PROPOSITION 1. *Se, oltre alle condizioni già specificate per g , vale che $c_g(\bar{h})$ è finita, allora per ogni coppia di elementi $\varphi^{(1)}$ e $\varphi^{(2)}$ di $\mathcal{F}(\varphi_1, \varphi_2)$ tali che $\varphi^{(2)}$ sia più concordante di $\varphi^{(1)}$, si ha $c_g(\varphi^{(1)}) \leq c_g(\varphi^{(2)}) \leq c_g(\bar{h})$, e il massimo $c_g(\bar{h})$ è raggiunto solo in presenza di \bar{h} .*

La proposizione afferma che $c(\cdot)$ è monotona non decrescente rispetto all'ordinamento della concordanza in una assegnata classe di Fréchet. La proprietà è fondamentale ai fini della costruzione di un indicatore sommario della monotonia della relazione statistica prodotta da un elemento qualunque di detta classe. La costruzione, per essere completata, richiede che si guardi anche all'allontanamento da $\bar{\Gamma}$. A questo proposito, se si definisce $\sigma_g(P, \bar{\Gamma})$ allo stesso modo di $\sigma_g(P, \underline{\Gamma})$, prendendo come riferimento la f.s.m. non decrescente al posto della non crescente, si perviene alla media

$$(4) \quad d_g(\varphi) := \int_{S(\varphi)} \sigma_g(P, \bar{\Gamma}) \varphi(dP) \quad (\varphi \in \mathcal{F}(\varphi_1, \varphi_2))$$

per la quale vale la

PROPOSITION 2. *Se, oltre alle condizioni per g , vale che $d_g(\underline{h}) < +\infty$, allora per la stessa coppia $\varphi^{(1)}$ e $\varphi^{(2)}$ della Proposizione 1 si ha $d_g(\underline{h}) \geq d_g(\varphi^{(1)}) \geq d_g(\varphi^{(2)})$, e il massimo $d_g(\underline{h})$ è raggiunto solo in presenza di \underline{h} .*

Si direbbe quindi che d_g è utile alla costruzione sopra accennata quando si guarda all'allontanamento dalla relazione di massima concordanza compatibile con la coppia di marginali assegnate. Acquista un senso ben preciso anche la differenza

$$\Delta_g(\varphi) := c_g(\varphi) - d_g(\varphi)$$

per ogni $\varphi \in \mathcal{F}(\varphi_1, \varphi_2)$ se sono finite $c_g(\bar{h})$ e $d_g(\underline{h})$. Infatti, alla luce delle precedenti proposizioni, si può dire che se $\Delta_g(\varphi)$ ha valore non lontano da $c_g(\bar{h})$ [$-d_g(\underline{h})$, rispettivamente], φ giustifica l'esistenza di una relazione statistica crescente [decrescente, rispettivamente] fra due caratteri distribuiti secondo φ . Se $\Delta_g(\varphi)$ assume il valore zero si dice che i due caratteri sono vicendevolmente *indifferenti*, occupando la loro distribuzione una posizione intermedia fra le estreme rispetto al problema della mo-

notonia delle relazioni statistiche². Conclusioni analoghe sarebbe lecito trarre nel caso in cui $|\Delta_g(\varphi)|$ fosse piccolo, ma rimarrebbero comunque due zone di valori facilmente individuabili in cui la sola indicazione fornita da Δ non è sufficiente. In questi ultimi casi, essa andrebbe integrata in un acconcio ragionamento di statistica inferenziale, il cui completamento esula però dagli scopi di questa relazione. Ci limitiamo ad osservare che un'indagine approfondita sulla questione sarebbe nuova. Si osserva infine che il criterio basato su $\Delta_g(\cdot)$ sopra descritto ha notevoli punti di contatto coi criteri α , β e γ del Gini, ove l'intervento del concetto di carattere contrario venga evitato col ricorso all'elemento minimo della classe di Fréchet, e al posto di ragionare in termini d'indipendenza in distribuzione si ragioni in termini di indifferenza, concetti più idonei rispetto agli obiettivi della ricerca sulla monotonia.

5. L'indicazione fornita da un indice generalmente migliora in espressività quando è possibile normalizzarlo col valore massimo che esso raggiunge in corrispondenza all'estremo della proprietà per la cui "misura" viene proposto. La realizzazione di questo punto, ben presente nel programma originario del Gini, era all'epoca causa di difficoltà di calcolo e di concetto. Le seconde si possono dire ben superate grazie alla conoscenza delle classi di Fréchet e relative proprietà. In un tentativo di semplificazione dei calcoli richiesti per la determinazione dei massimi, Tommaso Salvemini (1906-1994) propose nel 1939 un metodo basato su certe tabelle, che egli stesso ideò e chiamò di *cograduazione* e *contrograduazione*. Cfr. [26]. Altro non sono che una rappresentazione delle distribuzioni \bar{h} e, rispettivamente, \underline{h} nel caso in cui Ω è finito. Anche lo studio già citato di Dall'Aglio, ben più generale del precedente, va visto come un passo decisivo verso la risoluzione dello stesso problema. Il quale si può ritenere ora risolto in base alle Proposizioni 1 e 2, che, a loro volta, sono corollari di risultati del citato articolo [2] di Stamatis Cambanis (1943-1995), Gordon Simons e William Stout. Seguendo la terminologia di Gini, si dice di *omofilia* ogni indice ottenuto dalla normalizzazione di $\Delta(\cdot)$. Più precisamente:

²La questione della caratterizzazione della situazione d'indifferenza meriterebbe di essere approfondita. Ad esempio, andrebbe considerata, in alternativa, anche la definizione: una distribuzione $\bar{\varphi}$ in $\Gamma(\varphi_1, \varphi_2)$ realizza l'indifferenza reciproca dei due caratteri X e Y se $\sigma_g(P, \underline{\Gamma})$ e $\sigma_g(P, \bar{\Gamma})$ hanno la stessa distribuzione quando (X, Y) si distribuisce secondo $\bar{\varphi}$. Ovviamente, questa definizione è più restrittiva di quella data nel testo.

Siano φ_1, φ_2 distribuzioni su \mathbb{R} e sia $g: \mathbb{R}_+ \rightarrow \mathbb{R}$ una funzione continua, strettamente crescente, convessa e nulla nell'origine, tali che $c_g(\bar{h})$ e $d_g(\underline{h})$ risultino finiti; allora, si dice indice di omofilia il funzionale

$$(5) \quad \mathcal{F}(\varphi_1, \varphi_2) \ni \varphi \mapsto O_g(\varphi) = \frac{\Delta_g(\varphi)}{c_g(\bar{h})} \mathbb{1}_{(\Delta_g(\varphi) > 0)} + \frac{\Delta_g(\varphi)}{d_g(\underline{h})} \mathbb{1}_{(\Delta_g(\varphi) < 0)}.$$

Esso è sempre diverso da zero a meno che φ costituisca un caso d'indifferenza fra i due caratteri; il suo valore assoluto non supera 1; vale 1 solo se $\varphi = \bar{h}$ e, quindi, i due caratteri sono "legati" da una f.s.m. non decrescente; vale -1 solo se $\varphi = \underline{h}$ e, quindi, i due caratteri sono "legati" da una f.s.m. non crescente.

Gini propone poi di riservare il termine di *indici di correlazione* a quegli indici ottenuti normalizzando $\Delta(\cdot)$ col massimo raggiungibile sull'intera classe delle distribuzioni su \mathbb{R}^2 . Sarebbe dunque sbagliato usare indici simili per valutare l'esistenza di una relazione monotona qualunque fra i caratteri. Potrebbe infatti accadere che, pur esistendo, non possa esser segnalata da un indice che raggiunge il valore estremo solo in corrispondenza ad uno specifico tipo di relazione statistica non realizzabile dalle distribuzioni effettive dei due caratteri. Questa considerazione prelude ad una doverosa discussione sul valore del celebre *coefficiente di correlazione lineare* (Bravais, Pearson) in rapporto al corrispondente indice di omofilia.

6. È ben noto che il coefficiente di correlazione lineare per due caratteri distribuiti secondo φ con marginali non degeneri φ_1 e φ_2 è definito, quando la varianza di ciascuno di essi è finita, dalla formula

$$\rho(\varphi) = \frac{\text{Cov}(\varphi)}{\sigma_1 \cdot \sigma_2}$$

dove $\text{Cov}(\cdot)$ denota la *covarianza* e σ_i la radice aritmetica della *varianza* determinata da φ_i ($i = 1, 2$). Il suo valore assoluto non supera 1, e vale 1 solo se ciascun carattere è trasformazione affine dell'altro. Ovviamente, non è detto che le marginali φ_1, φ_2 consentano a $|\text{Cov}(\varphi)|$, neppure nei casi estremi, di raggiungere $\sigma_1 \cdot \sigma_2$ al variare di φ in $\mathcal{F}(\varphi_1, \varphi_2)$ e, quindi, $\rho(\cdot)$ non può in generale essere interpretato come indice di omofilia.

Sarebbe quindi interessante trovare condizioni in presenza delle quali ρ possa assolvere al compito di indice di omofilia su $\mathcal{F}(\varphi_1, \varphi_2)$. La prima è che $\bar{\Gamma}$ e $\underline{\Gamma}$ siano rette inclinate, circostanza necessaria e sufficiente affinché le situazioni estreme della correlazione lineare siano raggiungibili su $\mathcal{F}(\varphi_1, \varphi_2)$. Come mostrato in [5], ciò si può presentare se e solo se esistono parametri $a > 0$, $\alpha = -a$, b e β in \mathbb{R} tali che

$$(6) \quad F_2(x) = F_1\left(\frac{x-b}{a}\right) = 1 - F_1\left(\frac{x-\beta}{\alpha} - 0\right)$$

e F_1 ha momento secondo finito. La specificità delle condizioni (6) è sufficiente a mettere in giusta luce la differenza fra analisi della correlazione lineare e analisi della concordanza/discordanza secondo Gini. Conviene comunque completare il ragionamento dimostrando che l'uguaglianza $\rho(F) = O_g(\varphi)$ con $g(x) = x^2$, $x > 0$, vale per ogni $\varphi \in \mathcal{F}(\varphi_1, \varphi_2)$ se F_1 e F_2 soddisfano (6). Dapprima, si osserva che la seconda di (6) implica $a \int_{\mathbb{R}} x dF_1(x) + b = \alpha \int_{\mathbb{R}} x dF_1(x) + \beta$, la quale, insieme alla prima porge $\int_{\mathbb{R}} x dF_2(x) = a \int_{\mathbb{R}} x dF_1(x) + b$ e, per $k = 1, 2$, $\int_{\mathbb{R}} (ax+b)^k dF_1(x) = \int_{\mathbb{R}} (\alpha x + \beta)^k dF_1(x)$ e $\sigma_2^2 = a^2 \sigma_1^2$. Inoltre, i punti (x, y) di $\bar{\Gamma}$ [$\underline{\Gamma}$, rispettivamente] soddisfano $y = ax + b$ [$y = \alpha x + \beta$, rispettivamente] e, quindi,

$$\begin{aligned} \sigma_g(P, \underline{\Gamma}) &= \left| y - \alpha x - \beta \right|^2 + \left| x - \frac{y - \beta}{\alpha} \right|^2 \\ \sigma_g(P, \bar{\Gamma}) &= \left| y - ax - b \right|^2 + \left| x - \frac{y - b}{a} \right|^2. \end{aligned}$$

Da quanto precede segue

$$\Delta_g(\varphi) = 4a \left(1 + \frac{1}{a^2}\right) \text{Cov}(\varphi)$$

e

$$c_g(\bar{h}) = 4a^2 \left(1 + \frac{1}{a^2}\right) \sigma_1^2 = d_g(\underline{h})$$

sicché

$$O_g(\varphi) = \frac{4a \text{Cov}(\varphi)}{4a\sigma_1 a \sigma_1} = \frac{\text{Cov}(\varphi)}{\sigma_1 \cdot \sigma_2}.$$

In conclusione; Se $g(x) = x^2$ per $x \geq 0$ e se F_1 ha momento secondo finito, allora $O_g(\varphi) = \rho(\varphi)$ vale per ogni φ in $\mathcal{F}(\varphi_1, \varphi_2)$ se e solo se F_1 e

F_2 soddisfano (6).

Un caso notevole in cui tutte queste condizioni sono soddisfatte si ha se φ_1 e φ_2 sono *gaussiane*.

7. Le stesse condizioni sono verificate in almeno altre due situazioni rilevanti. La prima si ha quando delle n unità di una popolazione statistica o di un campione vengono considerati due caratteri (eventualmente qualitativi) ordinati, le cui modalità vengono sostituite – in accordo all'ordine prestabilito – con *gradi* (o *ranghi*) da 1 a n . Il caso più semplice da descrivere è quello in cui la stessa modalità si presenta in corrispondenza ad una sola unità (assenza di ripetizioni). Ne discende che l' i -esima unità si presenta – con frequenza = $1/n$ – come caratterizzata dalla coppia (R_i, S_i) essendo R_i [S_i , rispettivamente] il grado della modalità del primo [secondo, rispettivamente] carattere. È facile vedere che in questa situazione l'indice di omofilia (5) si riduce al generico *indice di cograduazione* G_g [terminologia proposta estendendo quella adottata da Gini in [14] nello studio del caso particolare in cui $g(x) = x$ per $x \geq 0$]:

$$(7) \quad G_g(F) = O_g(F) = \frac{\sum_{i=1}^n \left\{ g\left(\left|\frac{n+1-R_i-S_i}{n}\right|\right) - g\left(\left|\frac{R_i-S_i}{n}\right|\right) \right\}}{\sum_{i=1}^n \left\{ g\left(\left|\frac{n+1-R_i-S_i}{n}\right|\right) \right\}}$$

dove $F \in \mathcal{F}(\varphi_1, \varphi_2)$ con $F_1 = F_2 = k/n$ sull'intervallo $[k, k+1)$ per $k = 1, \dots, n-1$. Come caso particolare rilevante della (7), oltre al già citato indice di Gini, va segnalato il *coefficiente di correlazione per ranghi* dovuto a Charles Spearman (1863-1945) corrispondente alla specificazione $g(x) = x^2$ per $x \geq 0$. Cfr. [28]. La (5) può essere anche usata – senza alcun accorgimento particolare – nel caso di marginali F_i ($i = 1, 2$) con supporto contenuto in $\{1, \dots, n\}$ e codominio incluso in $\{0, 1/n, \dots, 1\}$.

La seconda delle due situazioni nominate all'inizio del paragrafo si ricollega ad una idealizzazione che consente di estendere la nozione di grado alle modalità di un carattere reale retto da una generica funzione di frequenza cumulata, diciamo G . In accordo alle definizioni accolte nella prima parte, se n continua ad indicare il numero delle unità esaminate, per grado corrispondente ad x si intenderà il numero (non necessariamente

intero) $n \cdot G(x)$. Quando G è f.r. continua su \mathbb{R} e ξ è un numero aleatorio con f.r. G , è ben noto che il grado $n \cdot G(x)$ ha distribuzione (continua) uniforme su $[0, n]$. Quindi, al posto della $\mathcal{F}(F_1, F_2)$, nell'ipotesi che F_1 e F_2 siano continue, si va a considerare la classe $\mathcal{CF}(F_1, F_2)$ di tutte le f.r. di $(F_1(\xi_1), F_2(\xi_2))$ al variare della legge di (ξ_1, ξ_2) in $\mathcal{F}(F_1, F_2)$. Ciascun elemento (inteso come f.r.) di $\mathcal{CF}(F_1, F_2)$ è, in accordo ad un linguaggio ormai consolidato, una *copula*: quella ottenuta dalla f.r. F verrà indicata con $C = C_F$. Per le cose già dette, le marginali di C_F sono uniformi su $[0, 1]$, sicché riescono soddisfatte le condizioni di cui si parla all'inizio del paragrafo. Va anche detto che ogni indice di omofilia definito su $\mathcal{CF}(F_1, F_2)$ è coerente con l'ordinamento di concordanza/discordanza introdotto nel terzo paragrafo. Cfr. [5], a giustificazione di ciò che molti Autori hanno fatto finora definendo, come indici di concordanza/discordanza tra due caratteri distribuiti rispettivamente secondo F_1 e F_2 , opportuni funzionali su $\mathcal{CF}(F_1, F_2)$. In genere, in tal modo si ottengono espressioni più semplici. Ad esempio, l'indice di omofilia (5) su $\mathcal{CF}(F_1, F_2)$ ha la forma

$$O_g(C) = \frac{\int_{(0,1)^2} \{g(|y+x-1|) - g(|y-x|)\} dC(x, y)}{\int_0^1 g(x) dx}$$

per ogni $C \in \mathcal{CF}(F_1, F_2)$. L'impostazione consistente nel definire indici di concordanza per elementi di $\mathcal{F}(F_1, F_2)$ tramite funzionali su $\mathcal{CF}(F_1, F_2)$ è stata seguita, ad es., in [27], [6], [22] e in [25] contenente una trattazione sistematica delle copule e, quindi, indicazioni tecniche per estendere il suddetto espediente anche a situazioni diverse da quelle particolari accennate in questo paragrafo. Con tale impostazione, la cui manifestazione più elementare è costituita dalla sostituzione delle modalità effettive con gradi, fa comunque perdere qualcosa rispetto a quella che opera direttamente su $\mathcal{F}(F_1, F_2)$ con indici di omofilia. In particolare, se i caratteri hanno modalità reali il passaggio a ranghi o, più in generale, a copule, non tiene conto della forma effettiva delle situazioni estreme. Comunque si vogliano considerare queste osservazioni a favore del punto di vista completo del Gini, a conclusione della trattazione si può dire asseverato il fatto che tale punto di vista consente di abbracciare in maniera unitaria l'aspetto della monotonia di una relazione statistica e le proposte di indici atti a quantificarlo.

8. Come detto nel primo paragrafo, la definizione dell'indice di disso-

miglianza è strettamente legata allo studio della connessione di due caratteri, che Gini intendeva sviluppare attraverso qualche misura della “discrepanza” fra la distribuzione di uno dei due caratteri e la distribuzione *condizionata* dello stesso data una generica determinazione dell’altro [distribuzione *parziale*, nel suo linguaggio]. Quindi, l’indice di dissomiglianza viene proposto allo scopo di dare un senso preciso alla desiderata misura della discrepanza fra le suddette distribuzioni. Tale indice, ritrovato successivamente e indipendentemente da altri Autori al di fuori del campo strettamente statistico, si può considerare contributo duraturo, ancorché fortuito, del Gini ad un campo inconsueto per uno statistico d’inizio Novecento: il capitolo della (moderna) teoria delle probabilità riguardante l’analisi degli spazi di misure di probabilità e della convergenza di successioni da essi estratte. La definizione segue un’idea guida ben precisa che, col linguaggio più consono delle probabilità, si potrebbe descrivere con le parole della breve *Note B* di Paul Lévy (1886-1971) contenuta in [10]. Il linguaggio delle probabilità consente anche un più facile raccordo con gli impieghi più recenti del concetto di dissomiglianza. Dati due elementi aleatori X e Y distribuiti, rispettivamente, secondo ν_1 e ν_2 nello stesso spazio metrico S con distanza d_S , la “diversità” degli stessi non è generalmente apprezzabile sulla base della conoscenza delle sole marginali ν_1 e ν_2 . Ogni misura accettabile della diversità richiede la conoscenza della distribuzione del numero aleatorio $d_S(X, Y)$ o, perlomeno, di una sua media ottenuta attraverso la legge congiunta del vettore aleatorio (X, Y) . L’altro aspetto caratterizzante la definizione giniana è la considerazione che, per estrarre da una misura della diversità di due elementi aleatori quanto sia dovuto soltanto alla dissomiglianza delle leggi marginali ν_1 e ν_2 , si debba ricorrere all’estremo inferiore della suddetta media al variare della legge di (X, Y) nella classe di tutti gli accoppiamenti possibili di tali marginali [idea di classe di Fréchet *in nuce*]. Vi si scorge lo stesso programma scientifico che culminerà, verso la fine degli anni sessanta in un celebre teorema di Volker Strassen [29]. Nonostante la ristrettezza degli strumenti matematici a sua disposizione, Gini definisce l’indice di dissomiglianza con sufficiente precisione in base alla suddetta proprietà estremante e ricerca il minimo in casi molto particolari di distribuzioni con supporto finito sull’asse reale [$d_S(X, Y) = |X - Y|$]. Ritornando all’ambito più generale, in cui si richiede soltanto che S sia spazio metrico *separabile*, dato $p \geq 1$, l’indice di dissomiglianza $G_p(\nu_1, \nu_2)$ di ordine p

tra ν_1 e ν_2 è quindi definito da
(8)

$$G_p(\nu_1, \nu_2) = \left(\inf \left\{ \int_{S^2} d(x, y)^p \nu(dxdy) : \nu \in \mathcal{F}(\nu_1, \nu_2) \right\} \right)^{1/p}$$

purché siano finiti gli integrali $\int_S d(x, y_0)^p \nu_1(dx)$ e $\int_S d(x_0, y)^p \nu_2(dy)$ per qualche scelta di x_0 e y_0 in S . Si può dimostrare che G_p è una distanza sullo spazio $[S]_p$ delle misure di probabilità, sui boreliani di S , che soddisfano la suddetta condizione di finitezza per i momenti marginali di d . Oggidì, (8) con $p = 1$ è quasi universalmente nota come *distanza di Wasserstein*, con riferimento a [32]. La natura di distanza per l'indice di dissomiglianza tra distribuzioni sull'asse reale era però già stata messa in evidenza, ad es., da Giampiero Landenna nel 1956. Cfr. [23].

Un'importante generalizzazione dell'indice di dissomiglianza, motivato da questioni di ottimizzazione di specifici processi economici, è dovuta al matematico sovietico Leonid V. Kantorovič (1912-1986). L'attività di questo vigoroso e originale studioso, rimasta per molto tempo sconosciuta, anche per ragioni di segretezza militare, alla comunità scientifica, riguardò per una parte significativa la ripartizione ottimale di risorse e nel 1975 gli valse il premio Nobel per l'economia [condiviso con Tjalling C. Koopmaus (1910-1985)]. A lui si deve l'invenzione della programmazione lineare, verso la fine degli anni '30 del secolo scorso e, per quanto riguarda la presente relazione, lo studio, a partire dall'inizio degli anni '40, di problemi come il seguente, che riducono indirettamente la distanza fra due misure di probabilità al costo ottimale di trasporto dell'una sull'altra: «Data una funzione di danno (costo) $c: S^2 \rightarrow [0, +\infty)$ e due qualunque misure di probabilità ν_1 e ν_2 , ci si prefigge di appurare l'esistenza e, in caso affermativo, di determinare le distribuzioni in $\mathcal{F}(\nu_1, \nu_2)$ per cui risulta minimo il costo medio $\int_{S^2} c(x, y) \nu(dxdy)$ al variare di ν in $\mathcal{F}(\nu_1, \nu_2)$ ». Se si prende $c = d_S^p$, il problema si riduce a quello formulato da Gini. Fra i contributi più rilevanti di Kantorovič alla materia si deve ricordare un teorema di dualità risalente al 1958, che ha giocato un ruolo molto importante in applicazioni all'analisi matematica delle distanze di cui stiamo parlando. Per questi si vedano [9] e i più recenti trattati [1] e [31]. A Kantorovič va riconosciuto anche il merito di aver messo in luce la connessione del suo problema col più classico problema del trasporto ottimo di Gaspard Monge (1746-1818). Si veda [24], dove Monge pone e studia la questione del cammino ottimale, rispetto al costo di trasporto, che le particelle di un materiale da costruzione dovrebbero compiere da un luogo determinato di estrazione verso una parte determinata di una

costruzione. Nell'arco temporale, che dalla fine del '700 arriva ai giorni nostri, si colloca, a buon diritto, anche la proposta fatta un secolo fa dal Gini. Come risultato degli avanzamenti e dei ritorni che si sono avvicendati lungo quell'arco, la distanza determinata dall'indice di dissomiglianza è spesso richiamata, oltre che col nome di Wasserstein, anche con quelli di Monge e Kantorovič.

9. Alla fine di questa relazione viene spontaneo esprimere ammirazione per il lavoro di Gini sulla teoria delle relazioni statistiche, soprattutto per originalità e profondità di ragionamento e per la notevole abilità di arrivare a conclusioni significative, pur nella sostanziale mancanza di mezzi tecnici adeguati. In particolare, colpisce positivamente la chiarezza con cui in tale lavoro vengono comunque precisate le idee di fondo sottostanti le definizioni dei concetti principali come, ad esempio, quello di ordinamento di concordanza/discordanza e di indice di dissomiglianza.

RIFERIMENTI BIBLIOGRAFICI

- [1] AMBROSIO, L., GIGLI, N. and SAVARÉ, G.. (2005) *Gradient Flows in Metric Spaces and in the Space of Probability Measures* Birkhäuser.
- [2] CAMBANIS, S., SIMONS, G. and STOUT, W.. (1976) Inequalities for $\mathcal{E}k(X, Y)$ when the marginals are fixed. *Z. Wahrscheinlichkeitstheorie verw. Gebiete* **36** 285-294.
- [3] CASSATA, F.. (2006) *Il Fascismo Razionale (Corrado Gini fra Scienza e Politica)*. Carocci, Roma.
- [4] CASTELLANO, V.. (1965) Corrado Gini: a memoir. *Metron* **24** 3-84.
- [5] CIFARELLI, D.M. and REGAZZINI, E.. (1990) Some contributions to the theory of monotone dependence. *CNR-IAM* **90** 17.
- [6] CONTI, P.. (1993) On some descriptive aspects of measures of monotone dependence. *Metron* **51** 43-60.
- [7] DALL'AGLIO, G.. (1956) Sugli estremi dei momenti delle funzioni di ripartizione doppia. *Annali della Scuola Normale Superiore di Pisa* **10** 35-74.
- [8] DE FINETTI, B.. (1937) A proposito di "correlazione". *Supplemento Statistico ai Nuovi Problemi di Politica, Storia ed Economia* **3** 41-57.
- [9] DUDLEY, R.M.. (2002) *Real Analysis and Probability*. Cambridge University Press, Cambridge.
- [10] FRECHÉT, M.. (1937) *Généralités sur les Probabilités (Premier Livre). Variables Aléatoires*. Gauthier-Villors, Paris.

-
- [11] FRECHÉT, M.. (1951) Sur les tableaux de corrélation dont les marges sont données. *Ann. Univ. Lyon(A)* **14** 53-77.
- [12] GINI, C.. (1908) *Il Sesso dal Punto di Vista Statistico. Le Leggi della Produzione dei Sessi*. Sandron, Milano.
- [13] GINI, C.. (1911) Considerazioni sulle probabilità a posteriori e applicazioni al rapporto dei sessi nelle nascite umane (con un'appendice di L. Galvani). *Studi Economico-Giuridici dell'Università di Cagliari* **3** 133-171.
- [14] GINI, C.. (1914) *Di Una Misura delle Relazioni tra le Graduatorie di Due Caratteri: l'Indice di Cograduazione*. Cecchini, Roma.
- [15] GINI, C.. (1914) Di una misura della dissomiglianza tra due gruppi di quantità e delle sue applicazioni allo studio delle relazioni statistiche. *Atti del R. Istituto Veneto di Scienze, Lettere e Arti* **74** 185-213.
- [16] GINI, C.. (1914) Indici di omofilia e di rassomiglianza e loro relazioni col coefficiente di correlazione e con indici di attrazione. *Atti del R. Istituto Veneto di Scienze, Lettere e Arti* **74** 583-610.
- [17] GINI, C.. (1915) Nuovi contributi alla teoria delle relazioni statistiche. *Atti del R. Istituto Veneto di Scienze, Lettere e Arti* **74** 1903-1942.
- [18] GINI, C.. (1915) Sul criterio di concordanza tra due caratteri. *Atti del R. Istituto Veneto di Scienze, Lettere e Arti* **75** 309-331.
- [19] GINI, C.. (1916) Indici di concordanza. *Atti del R. Istituto Veneto di Scienze, Lettere e Arti* **75** 1419-1461.
- [20] GINI, C.. (1964) Le applicazioni induttive del calcolo delle probabilità. *Rivista di Politica Economica* **54** 1059-1083.
- [21] HOEFFDING, W.. (1940) Masstaloivariate korrelations theorie. *Sebr. Math. Inst. Univ. Berlin* **5** 181-233.
- [22] KRUSKAL, W.H.. (1958) Ordinal measures of association. *Journal of the American Statistical Association* **53** 814-861.
- [23] LANDENNA, G.. (1956) La dissomiglianza. *Statistica* **16** 21-57.
- [24] MONGE, G.. (1781) Mémoire sur la théorie des déblais et des remblais. *Mémoires de l'Académie Royale des Sciences de Paris* 666-705.
- [25] NELSEN, R.B.. (2006) *An Introduction to Copulas* (2nd ed.) Springer, New York.
- [26] SALVEMINI, T.. (1939) Sugli indici di omofilia. *Atti della I Riunione Scientifica della Società Italiana di Statistica in Supplemento Statistico ai Nuovi Problemi di Politica, Storia ed Economia* **5** 105-115.
- [27] SCHWEIZER, B. and WOLFF, E.F.. (1981) On nonparametric measures of dependence for random variables. *Ann. Statist.* **9** 879-885.
- [28] SPEARMANN, C.. (1906) A footrule for measuring correlation. *British Journal of Psychology* **2** 89.
- [29] STRASSEN, V.. (1965) The existence of probability measures with given marginals. *Ann. Math. Statist.* **36** 423-439.

-
- [30] VILLANI, C.. (2003) *Topics in Optimal Transportation*. Graduate Studies in Mathematics (v.58) AMS, Providence.
 - [31] VILLANI, C.. (2009) *Optimal Transport, Old and New*. Springer, Berlin.
 - [32] WASSERSTEIN, L.. (1969) Markov processes over denumerable products of spaces describing large systems of automata. *Problems of Information Transmission* **5** 47-52.