
ATTI ACCADEMIA NAZIONALE DEI LINCEI
CLASSE SCIENZE FISICHE MATEMATICHE NATURALI
RENDICONTI

ADRIANA BROGINI BRATTI

**Relazioni fra statistiche sufficienti in presenza di una
statistica sufficiente minimale**

*Atti della Accademia Nazionale dei Lincei. Classe di Scienze Fisiche,
Matematiche e Naturali. Rendiconti, Serie 8, Vol. 66 (1979), n.5, p. 377–382.*

Accademia Nazionale dei Lincei

<http://www.bdim.eu/item?id=RLINA_1979_8_66_5_377_0>

L'utilizzo e la stampa di questo documento digitale è consentito liberamente per motivi di ricerca e studio. Non è consentito l'utilizzo dello stesso per motivi commerciali. Tutte le copie di questo documento devono riportare questo avvertimento.

*Articolo digitalizzato nel quadro del programma
bdim (Biblioteca Digitale Italiana di Matematica)
SIMAI & UMI*

<http://www.bdim.eu/>

Calcolo delle probabilità. — *Relazioni fra statistiche sufficienti in presenza di una statistica sufficiente minimale.* Nota di ADRIANA BROGINI BRATTI (*), presentata (**) dal Socio G. SCORZA DRAGONI.

SUMMARY. — In this paper the A. finds out necessary and sufficient conditions for the existence, or non existence, of a minimal sufficient statistics between general probability spaces. The method used here comes out from the study of the relations, in the range, between sufficient statistics and from the study of the σ -algebras induced by them.

INTRODUZIONE

In (6) è proposto un metodo per la ricerca e la costruzione di una statistica minimale sufficiente rispetto ad un assegnato insieme di misure di probabilità; esso è basato sulla partizione indotta dal « rapporto di verisimiglianza ». Altri metodi si possono ritrovare in [9] e in [1].

In questa Nota si affronta un problema connesso all'esistenza di una statistica minimale sufficiente; precisamente: siano (X, S) e $(Y, \bar{T})$ spazi misurabili (per def. e richiami, cfr. il paragrafo successivo); sia M un assegnato insieme di misure finite su S e si supponga che $T_0: X \rightarrow Y$ sia una statistica sufficiente minimale (per il seguito, ogni statistica sarà supposta suriettiva: ciò al fine di poter applicare il Lemma 2, di [4]); se T_1 e T_2 sono statistiche sufficienti, studiare, in Y , le relazioni fra $T_1(x)$ e $T_2(x)$, $\forall x \in X$. Lo spunto per tale lavoro è ricavato da [4] nel quale il teorema fondamentale è il seguente:

TEOREMA 1). *Sia M un assegnato insieme di misure su S ; M sia dominato. Sia $T: X \rightarrow Y$ una statistica. Condizione necessaria e sufficiente al fine che T sia statistica sufficiente per M è che: M sia equivalente ad una misura e per ogni $n \in M$ la derivata di Radon-Nikodym di n rispetto alla misura cui è equivalente M sia una funzione misurabile anche rispetto alla σ -algebra $T^{-1}(\bar{T})$.*

In base al citato Lemma 2) di [4] il teorema di sopra implica che: se $M \equiv m$, i.e.: M è equivalente alla misura m , T è sufficiente se e solo se esiste una funzione $g_n: Y \rightarrow R$, R il corpo reale, tale che g_n è misurabile e $dn/dm = g_n \circ T$. Siano, allora $T_i: X \rightarrow Y$, $i = 1, 2$, due statistiche sufficienti; $\forall n \in M$ si deve avere: $dn/dm = g_{1,n} \circ T = g_{2,n} \circ T$; è dunque immediato il problema: supposto che T_1 sia statistica sufficiente e che T_2 sia solo una funzione misurabile, è vero che il legame: $g_1 \circ T_1 = g_2 \circ T_2$, con le g_i , $i = 1, 2$, misurabili, è condizione oltre ché necessaria anche sufficiente a rendere la T_2 sta-

(*) Indirizzo dell'A.: Istituto di Statistica, Via VIII Febbraio 1-35100 Padova.

(**) Nella seduta del 12 maggio 1979.

tistica sufficiente? È chiaro che la risposta a tale problema è, in generale, negativa come si vede supponendo che le g_i siano ambedue nulle.

Un più approfondito studio delle relazioni fra statistiche sufficienti conduce ai risultati compendati nei Teoremi 2) e 3) del paragrafo 2) che danno, rispettivamente, un criterio di non esistenza ed uno di esistenza per le statistiche sufficienti minimali.

1) *Notazioni e richiami.* Sia X un insieme e sia S una σ -algebra di parti di X . Sia M un assegnato insieme di misure finite (l'ipotesi di finitezza non sarebbe necessaria per il seguito; la si assumerà avendo presente che le misure in oggetto siano misure di probabilità), su S dominato da una misura finita m' , i.e.: $M \ll m'$, (se $m'(E) = 0$, $\forall E \in S$, allora $n(E) = 0$, $\forall n \in M$). Da [4], Lemma 7), si ricava: se M è un insieme dominato, esiste una misura finita, m , alla quale M è equivalente, $M \equiv m$, i.e.: $M \ll m$ e se $n(E) = 0 \forall n \in M$, allora è anche $m(E) = 0$. Per il Teorema di Radon-Nikodym, [3], ogni misura n in M è derivabile rispetto alla misura m e si ha: $dn/dm = f: X \rightarrow \mathbb{R}$, con f integrabile su X rispetto a m . Siano E ed E' elementi in S ; sia k_E la funzione caratteristica di E ; se $n \in M$, si consideri la misura n' così definita:

$$n'(E') = \int_{E'} k_E dn = n(E \cap E')$$

n' è dominata da n ; ciò implica: se $(Y, \bar{T})$ è un secondo spazio misurabile e se $T: X \rightarrow Y$ è una funzione misurabile, la misura:

$$(n'T)^{-1}: \bar{T} \rightarrow \mathbb{R}, (n'T)^{-1}(F) = n'(T^{-1}(F))$$

è dominata, su $\bar{T}$, dalla misura $(nT)^{-1}$. Allora: sempre in base al teorema di Radon-Nikodym, esiste la derivata di $(n'T)^{-1}$ rispetto a $(nT)^{-1}$; se è: $g = d(n'T)^{-1}/d(nT)^{-1}$, risulta:

$$\forall F \in \bar{T}, (n'T)^{-1}(F) = \int d(nT)^{-1} g$$

Quanto sopra permette, seguendo [4], di dare la

DEFINIZIONE 1). $\forall E \in S$, la n -probabilità di E , dato $T(x)$, è $g(T(x))^{(1)}$. (Per il seguito, adottando un linguaggio probabilistico, la n -probabilità di E dato $T(x)$ si dirà la probabilità condizionata di E dato $T(x)$ e si indicherà con il simbolo: $p_n(E/T(x))$).

DEFINIZIONE 2). Se $T: X \rightarrow Y$ è misurabile, T si dirà una statistica.

DEFINIZIONE 3). Se $T: X \rightarrow Y$ è una statistica ed M è un insieme di misure su S , T si dirà sufficiente rispetto a M se e solo se: $\forall E \in S$ esiste una

(1) L'espressione n -probabilità è qui usata in modo del tutto formale, visto che non è detto che sia $n(X) = 1$.

funzione misurabile, $p_E: Y \rightarrow R$, tale che: $p_n(E/T(x)) = p_E(Tx)$, modulo $(nT)^{-1}$, $\forall n \in M$ ⁽²⁾.

La definizione 3) implica che la probabilità condizionata $p_n(E/T(x))$ è la stessa per ogni $n \in M$.

Per comodità di esposizione, d'ora in poi: l'insieme delle statistiche sufficienti di X su Y , rispetto ad un assegnato insieme di misure M su S , si indicherà con il simbolo $S(X, Y; M)$. Infine: se $T: X \rightarrow Y$ è un'applicazione, il sottoinsieme $T^{-1}(T(x))$, $\forall x \in X$, rappresenta « la curva di livello » di T passante per x , i.e.: l'insieme degli $x' \in X$ tali che $T(x') = T(x)$.

DEFINIZIONE 4). Se $T_0 \in S(X, Y; M)$, T_0 si dirà la statistica sufficiente minimale se e solo se per ogni altra $T \in S(X, Y; M)$ si ha: la curva di livello di T passante per $x \in X$, è contenuta nella curva di livello della T_0 passante per il medesimo x , $\forall x \in X$.

2) In questo paragrafo si dimostreranno criteri di esistenza (o non), per le statistiche minimali sufficienti; essi sono il risultato da due lemmi preliminari che riducono lo studio delle statistiche sufficienti allo studio di certe applicazioni di X sullo spazio delle curve di livello delle statistiche sufficienti. Lo strumento usato per questo studio è il Teorema 1) dell'introduzione; per compendiare le idee, poi, nel confronto fra statistiche sufficienti, una delle quali minimale, si scriverà un diagramma riassuntivo dei risultati. Sia $T \in S(X, Y; M)$; si indichi con $k_T(x)$ la curva di livello di T passante per $x \in X$ e con $K_T(X)$ l'insieme di tutte le curve di livello della T . Si consideri la funzione $k_T: X \rightarrow K_T(X)$, $x \rightarrow k_T(x)$; essa si fattorizza, in modo naturale, attraverso l'isomorfismo di insiemi $\hat{T}$, $\hat{T}: K_T(X) \rightarrow Y$, definito dall'equazione: $\hat{T}(k_T(x)) = T(x)$, (i.e.: $T = \hat{T} \circ k_T$). Su $K_T(X)$, attraverso $\hat{T}$, si ponga la σ -algebra $\hat{T}^{-1}(\bar{T})$.

LEMMA 1). $T \in S(X, Y; M)$ se e solo se $k_T \in S(X, K_T(X); M)$, quando $K_T(X)$ è dotato della σ -algebra $\hat{T}^{-1}(\bar{T})$.

Dimostrazione. Poiché si suppone che $M \equiv m$, per ogni n in M sia $f_n: X \rightarrow R$ la derivata, di Radon-Nikodym, di n rispetto a $m: f_n = dn/dm$. In base al Teorema 1) citato nell'introduzione sia $g_n: Y \rightarrow R$ tale che: $f_n = g_n \circ T$; inoltre, $\hat{g}_n: K_T(X) \rightarrow R$ sia la mappa così definita: $\hat{g}_n(k_T(x)) = f_n(x)$. Verifichiamo che $\hat{g}_n$ è ben definita; infatti: se $k_T(x) = k_T(x')$, è $T(x) = T(x')$ e, quindi: $\hat{g}_n(k_T(x)) = f_n(T(x)) = g_n(T(x)) = g_n(T(x')) = f_n(x')$ che è, per definizione di $\hat{g}_n$, eguale a $\hat{g}_n(k_T(x'))$. Ovvio che è: $f_n = \hat{g}_n \circ k_T$, e, quindi $\hat{g}_n = f_n \circ k_n^{-1}$ ovvero $\hat{g}_n^{-1} = k_T \circ f_n^{-1}$.

Se B è un boreliano in R , risulta: $\hat{g}_n^{-1}(B) = k_T(f_n^{-1}(B)) = k_T(T^{-1}(F))$, per qualche F in $\bar{T}$. Poiché: $T = \hat{T} \circ k_T$, dalle eguaglianze di sopra si trae

(2) i.e.: a meno di insiemi di $(nT)^{-1}$ misura nulla.

$k_T(T^{-1}(F)) = k_T(k_T^{-1}(T^{-1}(F))) = \hat{T}^{-1}(F)$ e quest'ultimo insieme è della σ -algebra posta su $K_T(X)$. Ciò dimostra che la $\hat{g}_n: K_T(X) \rightarrow R$ è misurabile. Allora, sempre in base a quel teorema, poiché $f_n = \hat{g}_n \circ K_T$, $\forall n \in M$, la funzione $k_T: X \rightarrow K_T(X)$ è statistica sufficiente.

Viceversa: se $k_T \in S(X, K_T(X); M)$ allora è: $f_n = \hat{g}_n \circ k_T$ con la $\hat{g}_n: K_T(X) \rightarrow R$ misurabile. Se si definisce $g_n: Y = T(X) \rightarrow R$ attraverso l'equazione $g_n(T(x)) = f_n(x)$ si dimostra, come sopra, che la g_n è ben definita e misurabile.

Il lemma è completamente dimostrato.

Siano T e T_0 in $S(X, Y; M)$; T_0 sia la statistica minimale. Sia $L: K_T(X) \rightarrow K_{T_0}(X)$ così definita: $L(k_T(x)) = k_{T_0}(x)$. L è ben definita per il fatto che T_0 è minimale. Su $K_T(X)$ si ponga la σ -algebra fornita da $L^{-1}(k_{T_0}(T_0^{-1}(\bar{T})))$; che si tratti di una σ -algebra deriva dal fatto che: $k_{T_0}T_0^{-1}(\bar{T}) = \hat{T}^{-1}(\bar{T})$ e, quindi, la σ -algebra di sopra è l'immagine inversa, mediante L^{-1} , della σ -algebra $\hat{T}_0^{-1}(\bar{T})$. Sia F in $\bar{T}$; si ha: $L^{-1}(k_{T_0}(T_0^{-1}(F))) = k_T(T_0^{-1}(F))$. Infatti: se $x \in T_0^{-1}(F)$, a destra dell'eguaglianza di sopra compare il termine $k_T(x)$ che è un elemento del sottoinsieme $L^{-1}(k_{T_0}(x))$; ciò dimostra l'inclusione da destra verso sinistra.

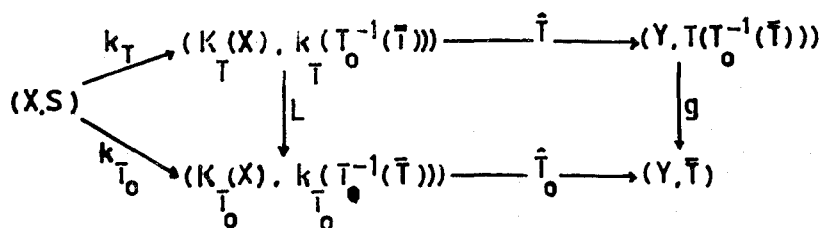
Viceversa: se $k_T(x) \in L^{-1}(k_{T_0}(T^{-1}(F)))$, allora: $k_{T_0}(x) \in k_{T_0}(T_0^{-1}(F))$ e, quindi: esiste x' in $T_0^{-1}(F)$ tale che: $k_{T_0}(x') = k_{T_0}(x)$. Ciò significa che $x \in T_0^{-1}(T_0(x')) \subseteq T_0^{-1}(F)$, ovvero: $k_T(x) \in k_T(T_0^{-1}(F))$. Ne discende che: su $K_T(X)$ la σ -algebra $L^{-1}(k_{T_0}(T_0^{-1}(\bar{T})))$ coincide con la σ -algebra $k_T(T_0^{-1}(\bar{T}))$.

LEMMA 2. La funzione $k_T: X \rightarrow K_T(X)$ è misurabile non appena $K_T(X)$ ha la σ -algebra $k_T(T_0^{-1}(\bar{T}))$.

Dimostrazione.

Se $F \in \bar{T}$, $T_0^{-1}(F) \subseteq k_0^{-1}(k_T(T^{-1}(F)))$; sostituendo a $k_T \hat{T}^{-1} \circ T$ si ottiene: $k_T^{-1}(\hat{T}_0^{-1}(T(T_0^{-1}(F)))) = T^{-1}(T(T_0^{-1}(F)))$. Per il fatto che T_0 è la minimale si ha: $T^{-1}(T(T_0^{-1}(F))) \subseteq T_0^{-1}(T_0(T_0^{-1}(F))) = T_0^{-1}(F)$; e, in definitiva: $k_T^{-1}(k_T(T_0^{-1}(F))) = T_0^{-1}(F)$. Ciò dimostra il lemma.

I lemmi 1) e 2) appena dimostrati si possono riassumere nel diagramma:



Si osservi che:

1) $T(T_0^{-1}(\bar{T}))$ è σ -algebra su Y poiché essa è l'immagine, mediante l'isomorfismo $\hat{T}$, della σ -algebra $k_T(T_0^{-1}(\bar{T}))$;

2) la funzione $g: Y \rightarrow Y$ si intende definita in questo modo: $g(T(x)) = T_0(x)$; che sia ben definita è immediato; inoltre: essa risulta essere bi-misurabile nel senso che: essa è misurabile e che $\forall F \in \bar{T}$ si ha $g(T(T_0^{-1}(F)))$ è in $\bar{T}$. Ciò è conseguenza del fatto che: $g = \hat{T}_0 L_0 \hat{T}_0$ e $\hat{T}$ e $\hat{T}_0$ sono isomorfismi; inoltre L è essa stessa bi-misurabile.

Il diagramma di sopra permette il seguente

COROLLARIO 1). *Se $S(X, Y; M)$ ammette statistica minimale sufficiente, allora: se T_1 e T_2 sono in $S(X, Y; M)$ esistono due applicazioni, g_i , di Y su Y tali che: $g_i(Y, T_i(T_0^{-1}(\bar{T}))) \rightarrow (Y, \bar{T})$ è bi-misurabile, ed è: $g_1(T_1(x)) = g_2(T_2(x))$.*

Per la dimostrazione è sufficiente sostituire nel diagramma, alla g , rispettivamente la g_1 e la g_2 ed osservare che: $g_i(T_i(x)) = T_0(x)$, $i = 1, 2$.

Il diagramma di sopra fornisce anche il

TEOREMA 2). *Le seguenti proposizioni, a) e b) sono equivalenti:*

- a) $S(X, Y; M)$ non ammette statistica sufficiente minimale;
- b) $\forall T_1$ in $S(X, Y; M)$ esiste una $T_2 \in S(X, Y; M)$ tale che la corrispondenza: $T_2(x) \rightarrow T_1(x)$, di Y su Y , non è una funzione.

Dimostrazione.

a) implica b). Se b) fosse falsa, esisterebbe $T_0 \in S(X, Y; M)$ tale che per ogni altra $T \in S(X, Y; M)$ la corrispondenza: $T(x) \rightarrow T_0(x)$ è una funzione di Y su Y . Ciò implicherebbe: se $T(x') = T(x)$, anche $T_0(x') = T_0(x)$.

Ovvio che T_0 risulta la minimale.

b) implica a). Se a) fosse falsa, la statistica minimale T_0 renderebbe falsa la proposizione in b).

Il teorema che seguirà dà una condizione necessaria e sufficiente per l'esistenza di una statistica sufficiente minimale.

TEOREMA 3). *Le seguenti proposizioni, a) e b), sono equivalenti:*

- a) $S(X, Y; M)$ ha statistica sufficiente minimale;
- b) per ogni $T_i \in S(X, Y; M)$ esiste una funzione $f_i: Y \rightarrow Y$ tale che:
 - I^o) $T_i: X \rightarrow Y$ è misurabile quando Y è dotato σ -algebra $f_i^{-1}(\bar{T})$;
 - II^o) le funzioni $f_i \circ T_i: X \rightarrow Y$ sono indipendenti da i .

Dimostrazione.

a) implica b). Sia $f_i: Y \rightarrow Y$ così definita: $f_i(T_i(x)) = T_0(x)$, se T_0 rappresenta la statistica sufficiente minimale. Dal diagramma, le f_i sono rappresentate dalla g ; la σ -algebra $f_i^{-1}(\bar{T})$ risulta essere la $T_i(T_0^{-1}(\bar{T}))$ e, quindi, la $T_i: X \rightarrow Y$ è misurabile rispetto alla $f_i^{-1}(\bar{T})$. Poiché, inoltre, è: $f_i(T_i(x)) = T_0(x)$, l'implicazione è dimostrata.

b) implica a). Per ogni x in X si ponga: $T_0(x) = f_i \circ T_i(x)$. In base a b), T_0 è indipendente da i e risulta essere misurabile perchè è la composta di funzioni misurabili.

Sia $T_i(x') = T_i(x)$; dalla definizione di T_0 , risulta pure: $T_0(x') = T_0(x)$, ciò implica che le curve di livello di ogni T_i sono contenute nelle curve di livello T_0 . Allora, per concludere, basterà dimostrare che la T_0 è statistica sufficiente.

Se nel diagramma si pone T_i al posto di T , in base al lemma 2) risulta che k_{T_i} è misurabile e dal lemma 1), poichè $T_i S(X, Y; M)$, risulta che k_{T_i} è in $S(X, K_{T_i}(X); M)$. D'altra parte, $L: K_{T_i}(X) \rightarrow K_{T_0}(X)$ è bi-misurabile; allora, per dimostrare che la T_0 è statistica sufficiente, basta dimostrare che se $T_1: X \rightarrow Y$ e $T_2: Y \rightarrow Z$ sono tali che: I°) $T_1 S(X, Y; M)$; II°) T_2 è bi-misurabile, segue che $T_2 \circ T_1: X \rightarrow Z$, è statistica sufficiente rispetto a M . (Nel caso in esame: $T_1 = k_{T_i}$ ed è $T_2 = L$; allora $k_{T_0} = L \circ k_{T_i}$, risulta statistica sufficiente e quindi lo è la stessa T_0).

Dimostrazione.

Sarà usata la stessa tecnica del lemma 2) di [4]. Se $n \in M$ e se $f_n = d_n/d_m$, esiste $g_n: Y \rightarrow R$, con $f_n = g_n \circ T_1$ e g_n misurabile.

Sia $z_0 \in Z$ e sia Y_0 il sottoinsieme di Y così determinato: $Y_0 = T_0^{-1}(z_0)$. Poichè T_2 è bi-misurabile, risulta: ponendo $E = \{y \in Y: g_n(y) = g_n(y_0)\}$, con y_0 in Y_0 , $E = g_n^{-1}(g_n(y_0)) = T_2^{-1}(H)$, per qualche H nella σ -algebra di Z . Dal momento che: $y_0 \in E$, $T_2(y_0) = z_0 \in H$; allora: $Y_0 = T_0^{-1}(z_0) \cap T_2^{-1}(H) = E$, ciò significa che la g_n è costante sugli elementi di Y_0 . L'equazione: $h_n(z_0) = g_n(y_0)$, con $y_0 \in T_2^{-1}(z_0)$, definisce quindi univocamente la funzione h_n . Dalla definizione di h_n , si ha: $h_n(T_2(y)) = g_n(y)$ e ciò dà anche la misurabilità della h_n .

In conclusione: se $n \in M$ e se $f_n = d_n/d_m = g_n$, $T_1 = h_n \circ T_2 \circ T_1$ che dà, sempre in base al teorema di Halmos-Savage, la sufficienza della statistica $T_2 \circ T_1$.

BIBLIOGRAFIA

- [1] R. R. BAHADUR (1954) - *Sufficiency and statistical decision functions*, «Ann. Math. Statist.», 25, 423-462.
- [2] R. A. FISHER (1922) - *On the mathematical foundation of theoretical statistics*, «Philos. trans. Roy. Soc.», A, 222, 309-368.
- [3] P. R. HALMOS (1950) - *Measure theory*, Springer Verlag.
- [4] P. R. HALMOS and L. J. SAVAGE (1950) - *Application of the Radon Nikodym theorem to the theory of sufficient statistics*, «Ann. Math. Statist.», 20, 225-241.
- [5] E. L. LEHMANN (1959) - *Testing statistical hypotheses*, «J. Wiley and Sons».
- [6] E. L. LEHMANN and H. SCHEFFÉ (1950) - *Completeness, similar regions and unbiased estimation*, Part. I, «Sankhya», 10, 305-340.
- [7] J. NEYMAN (1935) - *Su un teorema concernente le cosiddette statistiche sufficienti*, «Inst. Ital. Att. Giorn.», 6, 320-334.
- [8] S. S. WILKS (1962) - *Mathematical Statistics*, J. Wiley and Sons.
- [9] S. ZACKS - *The theory of statistical inference*, J. Wiley and Sons.
- [10] A. ZANELLA (1975-1976) - *Appunti di statistica metodologica*, Dis. della Fac. di Sc. Stat. Dem. ed Att., Padova.