FUNZIONE CARATTERISTICA DI UN FENOMENO ALEATORIO.

In: « *Atti del Congresso Internazionale dei Matematici* ». Bologna, 3-10 settembre 1928. Bologna, Ed. Zanichelli, 1929, pp. 179-190

B. DE FINETTI (Roma - Italia)

FUNZIONE CARATTERISTICA DI UN FENOMENO ALEATORIO

§ 1. - Scopo di questa comunicazione è di mostrare come il metodo della funzione caratteristica ([1]), già così vantaggiosamente introdotto nella teoria delle variabili casuali, si presti pure assai utilmente allo studio dei fenomeni aleatori. Mostreremo quindi, in primo luogo, come si possa individuare completamente un fenomeno aleatorio mediante la sua funzione caratteristica, e accenneremo poi alle operazioni che ne fanno un potente strumento di calcolo. Un'esposizione completa si troverà in una memoria che sarà presentata quanto prima alla R. Accademia dei Lincei ([2]).

§ 2. - Un fenomeno di cui si può fare (o quanto meno si può concepire) un numero qualunque di prove lo diremo *fenomeno aleatorio* quando l'ordine in cui le prove favorevoli e sfavorevoli si alternano sia da attribuirsi al caso. Si esige cioè che tutte le $\binom{n}{h}$ successioni di n prove di cui h favorevoli, successioni che differiscono tra loro solo per l'ordine, abbiano uguale probabilità. Questa, in termini precisi, la proprietà caratteristica di quelli che abbiamo convenuto di definire fenomeni aleatori.

Sarà bene vedere con qualche esempio la portata di tale restrizione, e avere così un'idea chiara del campo di questa ricerca. Se si ha una moneta o un dado, e lo si lancia sempre allo stesso modo, non ci sarà nessun motivo, d'indole causale, nemmeno se della perfezione del pezzo non siamo sicuri, che possa influire sull'ordine in cui si alternano le prove favorevoli e sfavorevoli: l'ordine sarà dovuto al caso, e si ha quindi un fenomeno aleatorio, secondo la data definizione. Lo stesso si dica per il problema della roulette, per le estrazioni da un'urna scelta a sorte in una collezione nota, e tutti i casi consimili. Se invece si considera una successione di tiri al bersaglio di uno stesso tiratore, o la successione delle giornate piovose e non piovose, o delle giornate in cui il signore di rimpetto si rade la barba, tale condizione non si potrà ragionevolmente rite-

([1]) V. i Trattati di *Calcolo delle Probabilità* di G. CASTELNUOVO e di P. LÉVY.
([2]) Memorie della R. Acc. Naz. dei Lincei, S. 6ª, vol. IV, fasc. V.

nere verificata, perchè nel primo caso si può prevedere dapprima un progressivo addestramento del tiratore, e poi il sopravvenire della stanchezza, ciò che rende probabile un addensamento delle prove favorevoli nel periodo di forma migliore, perchè i giorni piovosi saranno riuniti in periodi di piovosità più o meno lunghi, senza parlare poi della periodicità stagionale, e perchè il signore di rimpetto si raderà sempre a intervalli più o meno regolari.

Per decidere, in pratica, se un certo fenomeno si possa considerare fenomeno aleatorio, oppure no, basta pensare se un'eventuale regolarità o altra singolarità riscontrata nell'ordine della successione si attribuirebbe al caso (e allora si ha un fenomeno aleatorio) o si potrebbe ritenere dovuta a qualche circostanza connessa al fenomeno, in modo da far pensare che anche in un'altra uguale serie di prove sia probabile si rinnovi.

§ 3. - Se di un certo fenomeno aleatorio si fanno n prove, il numero di quelle che risultano favorevoli è ovviamente una variabile casuale x_n capace di assumere soltanto i valori 0, 1,...., n; se indichiamo $\omega_h^{(n)}$ la probabilità che il fenomeno considerato si verifichi h volte su n prove, la variabile casuale x_n è caratterizzata dalle probabilità $\omega_0^{(n)}, \omega_1^{(n)}, \ldots, \omega_n^{(n)}$ colle quali può assumere i diversi valori possibili. Nel caso particolare e ben noto in cui il fenomeno abbia una probabilità costante p nota a priori, si sa che $\omega_h^{(n)} = \binom{n}{h} p^h (1-p)^{n-h}$, ma nel caso generale di cui ci occupiamo le $\omega_h^{(n)}$ potranno essere qualunque (a parte delle limitazioni imposte dalla natura stessa del problema, e che troveremo in seguito).

Un fenomeno aleatorio ci definirà dunque una successione di variabili casuali $x_1, x_2, \ldots, x_n, \ldots$, che è chiaro debbano risultare tra loro interdipendenti. Tale interdipendenza si traduce analiticamente in una relazione differenziale ricorrente che lega le loro funzioni caratteristiche $\psi_1, \psi_2, \ldots, \psi_n, \ldots$, e che costituisce la base di questa ricerca.

Si dimostra che al crescere di n la funzione

$$\psi_n\left(\frac{t}{n}\right) = \sum_0^n {}_h \omega_h^{(n)} e^{i \frac{h}{n} t}$$

tende uniformemente in ogni regione finita alla funzione intera

$$\psi(t) = \sum_0^\infty {}_h \omega_h^{(h)} \frac{i^h t^h}{h!},$$

che è quella appunto che si dirà per definizione « *funzione caratteristica del fenomeno aleatorio* ». Nota la ψ si ricavano tutte le ψ_n e di conseguenza tutte le $\omega_h^{(n)}$, e ciò giustifica bene la denominazione.

L'integrale da $-\infty$ a $+\infty$ della funzione $\dfrac{e^{it} - e^{-i\xi t}}{it} \psi(t)$ esiste sempre per ogni valore di ξ, ed è uguale rispettivamente a 0 e 2π per $\xi < 0$ e $\xi > 1$; di conseguenza esiste una variabile casuale di cui $\psi(t)$ è la funzione caratteristica, la corrispondente funzione di ripartizione è

$$\Phi(\xi)=\frac{1}{2\pi}\int_{-\infty}^{+\infty}\frac{e^{it}-e^{-i\xi t}}{it}\psi(t)dt,$$

ed è $\Phi(\xi)=0$ per $\xi<0$ e $\Phi(\xi)=1$ per $\xi>1$.

Da tali risultati scendono due teoremi importanti:

I. La probabilità che la frequenza su n prove sia compresa entro limiti assegnati ξ_1 e ξ_2 tende a $\Phi(\xi_2)-\Phi(\xi_1)$ al crescere di n;

II. La probabilità che tutte le frequenze dopo l'n^{ma} siano comprese fra limiti assegnati ξ_1 e ξ_2 tende a

$$\operatorname*{limd}_{\xi\to\xi_2}\Phi(\xi)-\operatorname*{lims}_{\xi\to\xi_1}\Phi(\xi)$$

al crescere di n.

Perchè, assegnata una $\psi(t)$, possa esistere un fenomeno aleatorio di cui essa sia funzione caratteristica occorre e basta che la corrispondente funzione di ripartizione $\Phi(\xi)$ (naturalmente reale e mai decrescente) sia nulla per $\xi<0$ e $=1$ per $\xi>1$.

§ 4. - Esponiamo succintamente i calcoli.

Tra le $\omega_h^{(n)}$ dovrà sussistere la relazione

$$\omega_k^{(m)}=\sum_{k}^{n-m+k}{}_h\,\omega_h^{(n)}\frac{\binom{h}{k}\binom{n-h}{m-k}}{\binom{n}{m}} \tag{1}$$

perchè $\binom{h}{k}\binom{n-h}{m-k}:\binom{n}{m}$ è la probabilità che su m prove, prese tra n di cui h favorevoli, le favorevoli siano k, quando tutte le combinazioni sono ugualmente probabili. In particolare (per $m=n-1$):

$$n\omega_k^{(n-1)}=(n-k)\omega_k^{(n)}+(k+1)\omega_{k+1}^{(n)} \tag{2}$$

e ponendo

$$\Omega_n(z)=\sum_{0}^{n}{}_h\,\omega_h^{(n)}z^h \tag{3}$$

tutte le (2) per $k=0, 1,...., n-1$ si riassumono nella relazione differenziale ricorrente

$$n\Omega_{n-1}(z)=n\Omega_n(z)+(1-z)D\Omega_n(z). \tag{4}$$

Derivando la (4) e ricavandone i valori delle successive derivate per $z=1$ si ha

$$\Omega_n(1+z)=\sum_{0}^{n}{}_h\binom{n}{h}\omega_h^{(h)}z^h \tag{5}$$

e si trova che $\Omega_n\left(1+\frac{z}{n}\right)$, quando $n \to \infty$, tende uniformemente a

$$\Omega(1+z) = \sum_0^\infty {}_h \omega_h^{(h)} \frac{z^h}{h!}. \tag{6}$$

La funzione caratteristica ψ_n è

$$\psi_n(t) = \Omega_n(e^{it}), \tag{7}$$

si dimostra che $\Omega\left(e^{\frac{t}{n}}\right)$ tende pure a $\Omega(1+t)$, e quindi la funzione caratteristica del fenomeno aleatorio $\left(\text{funzione limite cui tende uniformemente } \psi_n\left(\frac{t}{n}\right)\right)$ è

$$\psi(t) = \Omega(1+it). \tag{8}$$

§ 5. - Una prima conseguenza notevole: come mostrano le (8), (6), per n indefinitamente crescente, il momento m^{mo} di $\frac{x_n}{n}$, ossia il valor-probabile-limite della potenza m^{ma} della frequenza per un numero indefinitamente crescente di prove, tende a $\omega_m^{(m)}$, cioè al valore della probabilità che m prove siano tutte favorevoli.

Ai teoremi enunciati nel § 3 si arriva traducendo i risultati ottenuti per le funzioni caratteristiche in quelli corrispondenti per le funzioni di ripartizione: indicando $\Phi_n(x)$ la funzione di ripartizione di x_n, si ha

$$\lim_{n=\infty} \Phi_n(n\xi) = \Phi(\xi). \tag{9}$$

L'altro teorema, relativo alla probabilità che tutte le frequenze da un certo n in poi appartengano a un dato intervallo, si può considerare come l'estensione al caso di un qualunque fenomeno aleatorio del teorema relativo alle prove ripetute indipendenti e con probabilità costante, che è un caso particolare del noto teorema di CANTELLI (¹). Anche la dimostrazione si può ricondurre a quella del caso trattato da CANTELLI.

§ 6. - Le (6), (5), (7) provano chiaramente l'asserto che la ψ basta a determinare completamente tutte le ψ_n. Altrettanto si potrà dire della Φ, perchè da essa si ha

$$\psi(t) = \int_0^1 e^{it\xi} d\Phi = e^{it} - it \int_0^1 e^{it\xi} \Phi(\xi) d\xi \tag{10}$$

$$\Omega_n(1+z) = \int_0^1 (1+z\xi)^n d\Phi = (1+z)^n - nz \int_0^1 (1+z\xi)^{n-1} \Phi(\xi) d\xi \tag{11}$$

$$\omega_h^{(n)} = \binom{n}{h} \int_0^1 \xi^h (1-\xi)^{n-h} d\Phi = \binom{n}{h} \int_0^1 (h-n\xi) \xi^{h-1} (1-\xi)^{n-h-1} \Phi(\xi) d\xi. \tag{12}$$

(¹) F. P. CANTELLI: *Sulla probabilità come limite della frequenza*. Rend. R. Acc. dei Lincei, serie V, vol. XXVI, gennaio 1917.

§ 7. - Consideriamo due casi di particolare importanza, che serviranno anche utilmente come esempi.

Nel caso ben noto in cui la probabilità di un fenomeno è nota a priori e uguale a p,

$$\omega_h^{(n)} = \binom{n}{h} p^h (1-p)^{n-h}, \qquad \omega_h^{(h)} = p^h,$$

$$\Omega_n(1+z) = (1+pz)^n, \qquad \Omega(1+z) = \lim_{n=\infty} \left(1+\frac{pz}{n}\right)^n = e^{pz},$$

$$\psi(t) = e^{ipt},$$

o anche direttamente (dalla (7))

$$\psi(t) = \sum_0^\infty {}_h p^h \frac{i^h t^h}{h!} = e^{ipt}, \qquad \Phi(\xi) = 0,\ {}^1/_2,\ 1, \text{ a seconda che } \xi < p,\ \xi = p,\ \xi > p.$$

Al crescere di n la probabilità che la frequenza sia compresa in un intorno $p \pm \varepsilon$ di p tende all'unità; inversamente da tale ipotesi scende che $\psi(t) = e^{ipt}$, e cioè che il fenomeno ha probabilità costante e uguale a p in tutte le prove. Come casi particolari, per $p=0$, $p=1$, si ha $\psi(t)=1$, $\psi(t)=e^{it}$.

Nel caso in cui tutte le frequenze siano ugualmente probabili si avrà

$$\omega_0^{(n)} = \omega_1^{(n)} = \dots = \omega_n^{(n)} = \frac{1}{n+1}, \qquad \omega_h^{(h)} = \frac{1}{h+1},$$

$$\Omega_n(z) = \frac{1}{n+1}(1+z+z^2+\dots+z^n) = \frac{1}{n+1} \cdot \frac{1-z^{n+1}}{1-z},$$

$$\psi_n(t) = \frac{1}{n+1} \cdot \frac{1-e^{it(n+1)}}{1-e^{it}},$$

$$\Omega(1+z) = \lim_{n-\infty} \frac{1}{n+1} \cdot \frac{\left(1+\frac{z}{n}\right)^{n+1} - 1}{\frac{z}{n}} = \frac{e^z - 1}{z},$$

$$\psi(t) = \frac{e^{it}-1}{it} = \sum_0^\infty {}_h \frac{t^h}{(h+1)!},$$

$$\Phi_n(h) = \frac{h+{}^1/_2}{n+1} \quad (h=0, 1, \dots, n), \qquad \Phi(\xi) = \lim_{n-\infty} \Phi_n(n\xi) = \xi.$$

Al crescere di n la probabilità che la frequenza sia compresa in un intervallo ξ_1, ξ_2 tende a $\xi_2 - \xi_1$; inversamente da tale ipotesi scende

$$\psi(t) = \int_0^1 e^{it\xi} d\xi = \left[\frac{e^{it\xi}}{it}\right]_0^1 = \frac{e^{it}-1}{it}, \qquad \omega_h^{(n)} = \frac{1}{n+1},$$

e il fenomeno ha quindi ugualmente probabili tutte le diverse frequenze su n prove.

§ 8. - Passiamo alle operazioni sulle funzioni caratteristiche.

Come osservazione generale possiamo dire che tutte le operazioni che incontreremo sono distributive, a meno (ove occorra) di un fattore moltiplicativo che fa assumere il valore 1 alla $\psi(t)$ per $t=0$, come necessariamente deve aversi.

Introducendo l'operatore U:

$$Uf(t) = \frac{f(t)}{f(0)}$$

possiamo dire che le operazioni che ci si presenteranno sono prodotti del tipo UF con F distributiva.

Poichè la funzione di ripartizione $\Phi(\xi)$ è funzione lineare biunivoca della funzione caratteristica $\psi(t)$, ad ogni operazione distributiva sulla ψ corrisponderà la trasformata che opera su Φ.

Due operazioni utili per semplificare le notazioni sono P_n (leggere: «*polinomio ennesimo*») che applicato a ψ produce Ω_n [1], e $\begin{bmatrix} h \\ n \end{bmatrix}$ che applicato a ψ produce $\omega_h^{(n)}$. Si possono definire in generale ponendo: se

$$f(t) = \sum_0^\infty {}_h a_h \frac{i^h t^h}{h!},$$

(13) $$\mathrm{P}_n f(t) = \sum_0^n {}_h \binom{n}{h} a_h t^h = \sum_0^n {}_h (1+t)^h \left\{ \begin{bmatrix} h \\ n \end{bmatrix} f \right\}.$$

Si osservi in particolare che $\begin{bmatrix} n \\ n \end{bmatrix} f = a_n$, e che la $\psi_n(t)$ è $\mathrm{P}_n \psi(e^{it} - 1)$.

§ 9. - Sia $\psi(t)$ la funzione caratteristica di un certo fenomeno aleatorio. La funzione caratteristica del fenomeno contrario è

(14) $$K\psi(t) = e^{it}\psi(-t).$$

Infatti dire che su n prove quelle favorevoli sono h, equivale a dire che sono $n-h$ quelle favorevoli all'evento contrario, ciò che si esprime

(15) $$\begin{bmatrix} h \\ n \end{bmatrix} K = \begin{bmatrix} n-h \\ n \end{bmatrix},$$

e dà $\left(\text{scrivendo } \omega_h^{(n)} = \begin{bmatrix} n \\ h \end{bmatrix} \psi\right)$:

$$\mathrm{P}_n K\psi(z-1) = \omega_n^{(n)} + \omega_{n-1}^{(n)} z + \dots + \omega_0^{(n)} z^n =$$
$$= z^n \left\{ \omega_0^{(n)} + \omega_1^{(n)} \frac{1}{z} + \dots + \omega_n^{(n)} \frac{1}{z^n} \right\} = z^n \mathrm{P}_n \psi\left(\frac{1}{z} - 1\right),$$
$$\mathrm{P}_n K\psi(e^{it} - 1) = e^{nit} \mathrm{P}_n \psi(e^{-it} - 1)$$

e

$$K\psi(t) = \lim_{n=\infty} \mathrm{P}_n K\psi\left(e^{i\frac{t}{n}} - 1\right) = \lim_{n=\infty} \left(e^{i\frac{t}{n}}\right)^n \lim_{n=\infty} \mathrm{P}_n \psi\left(e^{-i\frac{t}{n}} - 1\right) = e^{it}\psi(-t).$$

Essendo in particolare $\begin{bmatrix} n \\ n \end{bmatrix} K = \begin{bmatrix} 0 \\ n \end{bmatrix}$:

(16) $$K\psi(t) = 1 + \omega_0^{(1)} it - \omega_0^{(2)} \frac{t^2}{2!} - \omega_0^{(3)} i \frac{t^3}{3!} + \dots + \omega_0^{(n)} i^n \frac{t^n}{n!} + \dots$$

---

[1] Precisamente la funzione $\Omega_n(1+z)$, non la $\Omega_n(z)$.

Se Φ è la funzione di ripartizione corrispondente a ψ, si trova che la funzione di ripartizione $K_\Phi \Phi$ corrispondente a $K\psi$ è

$$K_\Phi \Phi(\xi) = 1 - \Phi(1-\xi) \tag{17}$$

come era del resto intuitivo: la probabilità che al limite la frequenza d'un evento sia inferiore a ξ è uguale a quella che la frequenza dell'evento contrario sia superiore a $1-\xi$.

K è operazione distributiva reversibile involutoria:

$$K(\psi' + \psi'') = K\psi' + K\psi'', \qquad KK\psi = \psi, \qquad K^{-1}\psi = K\psi.$$

§ 10. - La funzione caratteristica nell'ipotesi che la prima prova sia favorevole è

$$R\psi = UD\psi, \tag{18}$$

nell'ipotesi che la prima prova sia sfavorevole è

$$S\psi = U(i-D)\psi, \tag{19}$$

più in generale, dopo r prove favorevoli ed s sfavorevoli diviene

$$R^r S^s \psi = UD^r(1-D)^s\psi. \tag{20}$$

La probabilità $\omega_n^{(n)}$ che le prime n prove siano tutte favorevoli è infatti uguale al prodotto della probabilità $\omega_1^{(1)}$ che la prima prova sia favorevole per la probabilità che, verificata quest'ipotesi, lo siano le prime $n-1$ prove successive, probabilità che è $\begin{bmatrix} n-1 \\ n-1 \end{bmatrix} R\psi$. Il coefficente n^{mo} dello sviluppo di $R\psi$ è quindi il coefficente $(n+1)^{\text{mo}}$, $\omega_{n+1}^{(n+1)}$, dello sviluppo di ψ diviso per il primo,

$$\omega_1^{(1)} = -iD\psi(0),$$

e quindi

$$R\psi(t) = \frac{1}{\omega_1^{(1)}} \left\{ \omega_1^{(1)} + \omega_2^{(2)} \frac{it}{1!} - \omega_3^{(3)} \frac{t^2}{2!} + \dots + \omega_{n+1}^{(n+1)} \frac{i^n t^n}{n!} + \dots \right\} = \frac{-iD\psi(t)}{-iD\psi(0)} = UD\psi(t).$$

Per dimostrare che $S = U(i-D)$ conviene partire dall'osservazione che S è ovviamente la trasformata di R mediante K: $S = KRK$. Dalla (14):

$$DK\psi(t) = e^{it}[i\psi(-t) - D\psi(-t)];$$

$$KDK\psi(t) = e^{it}\{e^{-it}[i\psi(t) - D\psi(t)]\} = (i-D)\psi(t); \qquad KDK = i - D;$$

e per facili proprietà di U:

$$S = KRK = UKDK = U(i-D),$$

da cui

$$S\psi(t) = \frac{i\psi(t) - D\psi(t)}{i - D\psi(0)} = \frac{1}{1-\omega_1^{(1)}} \sum_0^\infty {}_h (\omega_h^{(h)} - \omega_{h+1}^{(h+1)}) \frac{i^h t^h}{h!}.$$

Le operazioni R ed S sono permutabili:

$$RS = SR,$$

e quindi

$$R^r S^s = S^s R^r = S^{s_1} R^{r_1} S^{s_2} R^{r_2} \ldots \quad (s_1 + s_2 + \ldots = s, \; r_1 + r_2 + \ldots = r).$$

Ciò prova che dopo r prove favorevoli ed s sfavorevoli, indipendentemente dall'ordine in cui esse si alternano, la funzione caratteristica diviene $R^r S^s \psi$.

La probabilità che un fenomeno aleatorio che ha la funzione caratteristica ψ si verifichi h volte su n prove successive nell'ipotesi che delle prime $r+1$ prove quelle favorevoli siano r e s le sfavorevoli è data dalla formula

$$\begin{bmatrix} h \\ n \end{bmatrix} R^r S^s \psi = \frac{\binom{h+r}{r}\binom{n-h+s}{s}}{\binom{n+r+s}{n}} \cdot \frac{\begin{bmatrix} h+r \\ n+r+s \end{bmatrix} \psi}{\begin{bmatrix} r \\ r+s \end{bmatrix} \psi}. \tag{21}$$

§ 11. - Indichiamo R_Φ, S_Φ, le operazioni che trasformano la funzione di ripartizione corrispondente a ψ in quella corrispondente a $R\psi$ o a $S\psi$.

Partendo dall'espressione di $D\psi(t)$ che si ottiene derivando la (10) si ricava

$$R_\Phi \Phi(\xi) = \frac{\xi\Phi(\xi) - \int_0^\xi \Phi(\lambda)d\lambda}{1 - \int_0^1 \Phi(\lambda)d\lambda} = \frac{\int_0^{\Phi(\xi)} \xi d\Phi}{\int_0^1 \xi d\Phi}; \tag{22}$$

analogamente si trova

$$S_\Phi \Phi(\xi) = \frac{\int_0^{\Phi(\xi)} (1-\xi)d\Phi}{\int_0^1 (1-\xi)d\Phi}, \tag{23}$$

e in generale

$$R_\Phi^r S_\Phi^s \Phi(\xi) = \frac{\int_0^{\Phi\xi} \xi^r (1-\xi)^s d\Phi}{\int_0^1 \xi^r (1-\xi)^s d\Phi}. \tag{24}$$

A tali risultati possiamo dare una forma più espressiva. Pel teorema della media:

$$[R_\Phi^r S_\Phi^s \Phi]_{\xi_1}^{\xi_2} = \frac{\bar{\xi}^r (1-\bar{\xi})^s}{\int_0^1 \xi^r (1-\xi)^s d\Phi} [\Phi]_{\xi_1}^{\xi_2} \tag{25}$$

con $\xi_1 \leq \bar{\xi} \leq \xi_2$ e indicando $[\Phi]_{\xi_1}^{\xi_2} = \Phi(\xi_2) - \Phi(\xi_1)$.

Si deduce da questa formula un teorema asintotico assai notevole nel campo delle relazioni fra probabilità e frequenza. Se la frequenza su un numero sufficentemente grande di prove è f, la funzione caratteristica tende a quella stessa che varrebbe se il fenomeno avesse probabilità nota a priori uguale ad f, a meno che in un intorno di f la Φ sia costante. In termini precisi: se per nessun $\varepsilon > 0$ è

$$\Phi\left(\frac{r}{r+s} - \varepsilon\right) = \Phi\left(\frac{r}{r+s} + \varepsilon\right), \quad \text{è} \quad \lim_{n\to\infty} (R^r_\Phi S^s_\Phi)^n \Phi(\xi) = 0,\ {}^1/_2,\ 1$$

a seconda che $\xi < f$, $\xi = f$, $\xi > f$, e conseguentemente

$$\lim_{n\to\infty} (R^r S^s)^n \psi(t) = e^{ift}, \tag{26}$$

e la tendenza al limite è uniforme in ogni regione finita (CASTELNUOVO, Op. cit., vol. II, p. 198).

Quindi, qualunque sia la natura di un fenomeno aleatorio (purchè la sua Φ soddisfaccia la detta restrizione), la probabilità che, nell'ipotesi che sulle prime n prove la frequenza osservata sia f, in quelle successive la frequenza tenda a un limite che differisca da f per più di un dato ε può rendersi piccola quanto si vuole pur di prendere n sufficentemente grande.

§ 12. - Come esercizio, applichiamo i risultati trovati alle funzioni caratteristiche considerate nel § 8.

Se $\psi(t) = e^{ipt}$, si ha

$$K\psi(t) = e^{i(1-p)t}, \qquad R\psi(t) = S\psi(t) = R^r S^s \psi(t) = e^{ipt}.$$

Non si hanno altre funzioni caratteristiche che rimangono invariate conoscendo l'esito di una prova: se $R\psi = \psi$ o $S\psi = \psi$ scende che $\psi(t) = e^{ipt}$ $(0 \leqslant p \leqslant 1)$.

Se $\psi(t) = \dfrac{e^{it} - 1}{it}$:

$$K\psi(t) = \psi(t) = \frac{e^{it} - 1}{it};$$

$$R\psi(t) = \frac{2}{t^2}(e^{it} - ite^{it} - 1); \qquad S\psi(t) = \frac{2}{t^2}(1 + it - e^{it});$$

$$\begin{bmatrix} h \\ n \end{bmatrix} R^r S^s \psi = \frac{\binom{h+r}{r}\binom{n-h+s}{s}}{\binom{n+r+s}{n}} \cdot \frac{n+r+s+1}{r+s+1}. \tag{27}$$

In particolare la probabilità che l'$(r+s+1)^{\text{ma}}$ prova sia favorevole, dopo r prove favorevoli ed s sfavorevoli è

$$\begin{bmatrix} 1 \\ 1 \end{bmatrix} R^r S^s \psi = \frac{r+1}{r+s+2}. \tag{28}$$

È questa la formula di cui si fece uso e anche abuso nella teoria della probabilità a posteriori. Essa è rigorosamente esatta quando sia $\psi(t) = \dfrac{e^{it} - 1}{it}$, ma è valida solo in questo caso specialissimo.

§ 13. - Due altri problemi meritano un cenno.

Se si hanno due fenomeni (analogamente per tre o più), indipendenti tra loro, le cui funzioni caratteristiche sono

$$\psi'(t)=\sum_0^\infty{}_h a_h \frac{i^h t^h}{h!}, \qquad \psi''(t)=\sum_0^\infty{}_h b_h \frac{i^h t^h}{h!},$$

il loro coverificarsi è un fenomeno aleatorio la cui funzione caratteristica è

$$\psi(t)=\sum_0^\infty{}_h a_h b_h \frac{i^h t^h}{h!}. \tag{29}$$

Infatti a_h è la probabilità che il primo fenomeno si verifichi sempre su h prove, b_h l'analoga per il secondo, e di conseguenza $a_h b_h$ è la probabilità che su h prove si verifichino sempre entrambi.

In particolare se $\psi''(t)=e^{ipt}$ (fenomeno a probabilità nota p) si ha $\psi(t)=\psi'(pt)$; se

$$\psi''(t)=\frac{e^{it}-1}{it}$$

è

$$\psi(t)=\frac{1}{t}\int_0^t \psi'(t)dt.$$

Ad esempio per

$$\psi'(t)=e^{ipt}, \qquad \psi''(t)=\frac{e^{it}-1}{it}: \qquad \psi(t)=\frac{e^{ipt}-1}{ipt}.$$

§ 14. - Se un fenomeno può dipendere da diverse cause (incompatibili) che hanno rispettivamente le probabilità $\lambda_1, \lambda_2, \ldots, \lambda_m$, e nelle diverse ipotesi il fenomeno ha rispettivamente le funzioni caratteristiche $\psi^{(1)}(t), \psi^{(2)}(t), \ldots, \psi^{(m)}(t)$, la funzione caratteristica del fenomeno aleatorio è

$$\psi(t)=\lambda_1\psi^{(1)}(t)+\lambda_2\psi^{(2)}(t)+\ldots+\lambda_m\psi^{(m)}(t). \tag{30}$$

Su questo teorema si può fondare in modo formalmente impeccabile la teoria delle probabilità delle ipotesi. L'esempio classico cui si applica tale risultato è quello delle estrazioni da un'urna che è stata scelta a caso in una collezione nota. Se sappiamo che le percentuali di palle nere possono essere $p_1, p_2, \ldots, p_m$ colle probabilità $\lambda_1, \lambda_2, \ldots, \lambda_m$, la funzione caratteristica sarà

$$\psi(t)=\lambda_1 e^{ip_1 t}+\lambda_2 e^{ip_2 t}+\ldots+\lambda_m e^{i\lambda_m t}.$$

Un esempio che, pur riferendosi, per fissare le idee, alle estrazioni da un'urna, meglio si avvicina al tipo dei problemi che si possono presentare in pratica, è quest'altro. Abbiamo un'urna A contenente n palle tra bianche e nere, che vi sono state immerse da un individuo il quale aveva a sua disposizione $N=cn$ palle (c intero maggiore di 1), di cui $H=ch$ bianche e $K=ck$ nere ($H+K=N$, ossia $h+k=n$). Di tutte le ipotesi possibili riteniamo che soltanto le due seguenti possano essersi verificate: *a*) le n palle sono state scelte a caso fra le N disponibili; *b*) l'individuo che preparò l'urna ebbe cura di scegliere le n palle in modo da conservare le percentuali di palle bianche e nere (e quindi prendendo h palle bianche e k nere). Conosciamo ancora la probabilità che hanno queste due ipotesi: siano α e β. Il fenomeno consistente nell'estrazione di una palla bianca dall'urna A ha la funzione caratteristica

$$\psi(t) = \alpha \cdot \psi^{(\alpha)}(t) + \beta \cdot \psi^{(\beta)}(t)$$

ove

$$\psi^{(\beta)}(t) = e^{i\frac{h}{n}t},$$

e

$$\psi^{(\alpha)}(t) = \frac{1}{\binom{N}{n}} \sum_{0}^{n}{}_l \binom{H}{l}\binom{N-H}{n-l} e^{i\frac{l}{n}t}$$

$\left(\text{essendo } \binom{H}{l}\binom{N-H}{n-l} : \binom{N}{n}\right.$ la probabilità che delle n palle estratte a sorte e immesse nell'urna l siano bianche). Dopo $r+s$ estrazioni di cui r diedero palle bianche, la funzione caratteristica è

$$R^r S^s \psi(t) = \frac{1}{\left[\begin{matrix} r \\ r+s \end{matrix}\right]_\psi} \left\{ \alpha \cdot \left[\begin{matrix} r \\ r+s \end{matrix}\right]_{\psi^{(\alpha)}} \cdot R^r S^s \psi^{(\alpha)}(t) + \beta \cdot \left[\begin{matrix} r \\ r+s \end{matrix}\right]_{\psi^{(\beta)}} \cdot R^r S^s \psi^{(\beta)}(t) \right\}.$$

Derivando si ha, per $t=0$, la probabilità di ottenere, all'$(r+s+1)^{ma}$ estrazione, una palla bianca (determinazione di una probabilità a posteriori); la probabilità che le palle siano state scelte a caso (ipotesi *a*) dopo r estrazioni di palle bianche e s di palle nere è

$$\alpha \cdot \frac{\left[\begin{matrix} r \\ r+s \end{matrix}\right]_{\psi^{(\alpha)}}}{\left[\begin{matrix} r \\ r+s \end{matrix}\right]_{\psi}}$$

(determinazione di una probabilità delle ipotesi).

Per fare un'applicazione numerica: se l'urna contiene 6 palle scelte fra 12 che si avevano a disposizione, di cui 4 bianche e 8 nere, e si ritiene $\alpha = \frac{2}{3}$, $\beta = \frac{1}{3}$:

$$\psi^{(\alpha)}(t) = \frac{1}{33}\left\{1 + 8e^{\frac{it}{6}} + 15e^{\frac{it}{3}} + 8e^{\frac{it}{2}} + e^{\frac{2it}{3}}\right\}, \qquad \psi^{(\beta)}(t) = e^{\frac{it}{3}},$$

$$\psi(t) = \frac{1}{99}\left\{2 + 16e^{\frac{it}{6}} + 63e^{\frac{it}{3}} + 16e^{\frac{it}{2}} + 2e^{\frac{2it}{3}}\right\}.$$

Dopo $r+s$ estrazioni, di cui r diedero palle bianche, la probabilità dell'ipotesi *b*) (scelta non a caso) è

$$\frac{33.2^r.4^s}{16.5^s + 63.2^r.4^s + 16.3^{r+s} + 2.4^r.2^s}.$$

Dopo 6 estrazioni l'ipotesi *b*) ha le probabilità 0,088.353 se si ebbero $r=6$ palle bianche, 0,176.707 se $r=5$, 0,279.365 se $r=4$, 0,359.918 se $r=3$, 0,389.812 se $r=2$, 0,353.947 se $r=1$, e finalmente 0,260,018 se $r=0$, ossia se si estrassero solo palle nere. Al crescere del numero delle prove, la probabilità dell'ipotesi b) tende rispettivamente a $\frac{33}{63}=0{,}523.810$, $\frac{33}{79}=0{,}517.722$, o a zero, a seconda che la frequenza è compresa fra

$$(\log {}^5/_4)/(\log {}^5/_2)=0{,}243.529 \quad \text{e} \quad (\log {}^4/_3)/(\log 2)=0{,}415.037,$$

è uguale a uno di questi due limiti, o è esterna.

§ 15. - Il problema delle probabilità a posteriori consiste nel cercare di determinare la probabilità di un fenomeno aleatorio in base all'osservazione della frequenza effettivamente constatata in un certo numero di prove. In base a quanto s'è visto: *un problema di probabilità a posteriori è pienamente determinato se e solo se si riferisce a un fenomeno noto* (di cui è nota la funzione caratteristica).

Il problema delle probabilità delle ipotesi consiste nel ricercare la probabilità di un'ipotesi o *causa* cui un fenomeno aleatorio si possa far risalire, in base sempre all'osservazione della frequenza con cui il fenomeno s'è verificato in un certo numero di prove. E possiamo concludere: *un problema di probabilità delle ipotesi è pienamente determinato quando e solo quando è noto il fenomeno* (è data la sua funzione caratteristica), *è nota la precisa influenza dell'ipotesi* (è data la funzione caratteristica del fenomeno subordinatamente all'ipotesi), *e è nota la probabilità a priori dell'ipotesi stessa.*

Altrimenti questi due problemi non hanno senso.

Il teorema del § 11 contiene tutto quello che vi può essere di preciso in un tentativo d'inversione del teorema asintotico di Bernoulli. Dopo un numero indefinitamente crescente di prove, la probabilità di un fenomeno aleatorio tende a divenire uguale alla frequenza (colla restrizione detta a suo luogo). Ma la convergenza non è uniforme per tutte le funzioni caratteristiche, e quindi, per quanto grande sia il numero delle prove già eseguite, non è possibile dedurre che la probabilità sia anche approssimativamente uguale alla frequenza senza conoscere quale fosse a priori la funzione caratteristica del fenomeno. Possiamo dire però che col crescere del numero delle prove divengono sempre meno restrittive le condizioni che si debbono supporre verificate dalla funzione caratteristica del fenomeno aleatorio perchè ne consegua che l'uguaglianza approssimativa tra probabilità e frequenza sussista.

Queste conclusioni e questi esempi possono chiarire — in modo preciso nei due casi trattati delle probabilità a posteriori e delle probabilità delle cause, come spirito anche nel caso generale — l'influenza che sulla valutazione di una probabilità esercitano i dati dell'esperienza [1].

---

[1] Per una discussione più esauriente rimando a *Probabilismo. Saggio critico sulla teoria delle probabilità e sul valore della scienza.* Biblioteca di Filosofia diretta da A. Aliotta, Perrella ed., Napoli, 1931 (L. 5); cfr. specialmente il n.° 22.