

ANCORA SULL'ESTENSIONE ALLE CLASSI NUMERABILI DEL
TEOREMA DELLE PROBABILITÀ TOTALI

(In risposta alla seconda nota del Prof. Fréchet: *Sur l'extension du théorème des probabilités
totales au cas d'une suite infinie d'évènements*).

In: « *Rendiconti del Reale Istituto Lombardo di Scienze e Lettere* », serie II, vol. LXIII, fasc.
16-17. Pavia 1930, pp. 5-10.

SUR L'EXTENSION DU THÉORÈME
DES PROBABILITÉS TOTALES
AU CAS D'UNE SUITE INFINIE D'ÉVÈNEMENTS

Seconde note par MAURICE FRÉCHET (Paris)

(Adunanza del 20 novembre 1930)

Sunto. — Les exemples où cette extension conduit à une antinomie sont irrecevables, soit parce que les probabilités des événements considérées ne sont pas exprimables au moyen des nombres usuels, soit parce que les valeurs qui leurs sont attribuées ne peuvent se présenter dans la réalité.

Les questions soulevées par M. de Finetti dans ses notes récentes sont d'importance et on doit lui savoir gré d'avoir attiré l'attention sur quelques points peut être trop laissés dans l'ombre jusqu'ici.

1. M. de Finetti a donné des exemples simples montrant qu'on arrive à des antinomies en admettant que les probabilités de *tous* les événements fortuits peuvent être définies de façon à vérifier le principe des probabilités totales, non seulement sous sa forme *restreinte* (c'est à dire concernant seulement un nombre fini d'événements), mais encore sous sa forme *complète* (c'est à dire étendue à une suite dénombrable d'événements — toujours incompatibles —).

A vrai dire, ce fait était familier à tous les auteurs au courant de la théorie de la mesure et je l'avais déjà signalé explicitement, comme on peut le voir à la page 329 du Calcul des Probabilités de M. Paul Lévy. Mais les exemples de M. de Finetti ont l'avantage d'être directs et simples, sans emprunt à une théorie complexe.

La solution de la difficulté consiste, comme je l'ai indiqué dans une note précédente, à n'appliquer le principe *complet* des probabilités totales qu'à une classe déterminée d'événements fortuits que je propose d'appeler événements « probabilisables » par analogie avec le mot « mesurables ».

Dans sa dernière note, M. de Finetti reconnaît qu'on évite ainsi formellement la difficulté. Mais il soulève de nouvelles objections. J'espère donner à celles-ci, dans ce qui suit, une solution satisfaisante, mais je tiens à dire que ces objections sont d'un ordre si fondamental qu'on ne doit pas s'y dérober et que je sympathise avec l'état d'esprit qui les a provoquées.

2. M. de Finetti se demande d'abord si les événements non probabilisables au sens indiqué ci-dessus, ne sont pas des événements aussi intéressants que les autres, si l'on n'est pas dans la nécessité pratique de les considérer bien qu'ils donnent lieu à des raisonnements et à des formules moins simples.

L'exemple des ensembles non mesurables peut déjà nous faire prévoir le sens de la réponse : Tous les ensembles linéaires rencontrés pratiquement par les mathématiciens sont des ensembles mesurables. Quant aux physiciens, ils seraient sans doute unanimes à déclarer qu'ils se contentent d'ensembles linéaires beaucoup plus simples ; les ensembles formés chacun d'un nombre fini d'intervalles. Ainsi dans la répartition la plus simple — la répartition uniforme — des probabilités concernant la position d'un point sur une droite, les seuls événements pratiquement admissibles sont des événements « probabilisables ».

Mais M. de Finetti considère d'autres exemples. Est-il admissible, se demande-t-il, d'exclure la conception d'une infinité d'événements incompatibles qui seraient également probables, ou même simplement du même ordre de grandeur ? La même difficulté se présenterait si on édifiait une théorie de la mesure des ensembles de points situés non sur un segment fixe, mais sur la droite indéfinie. En donnant à la droite indéfinie la mesure 1, on pourrait attribuer la mesure $\frac{1}{2}$ à une

demi-droite, la mesure $\frac{1}{3}$, à l'ensemble formé par une suite infinie de segments de longueurs égales à l séparés deux à deux par des intervalles de longueurs égales à $2l$; etc. Dans une telle théorie, la droite illimitée serait la réunion d'une suite de segments égaux consécutifs dont les mesures ne pourraient qu'être nulles et dont la somme ne pourrait être égale à l'unité. Il faudrait donc exclure la conception de mesures égales de ces segments ou les considérer comme non mesurables. [On sortirait d'embarras en considérant leurs mesures

comme des infiniments petits *actuels* ε par rapport à la mesure de la droite. En introduisant cette nouvelle espèce de nombres on pourrait admettre une égalité telle que $0 \times \infty = 1$ ou mieux $\varepsilon \cdot \omega = 1$. Une même manière de faire permettrait d'attribuer à des événements incompatibles en nombre infini, une égale probabilité ε , pourvu que ce nombre ε soit d'une autre espèce que les nombre ordinaires et soit considéré comme un infiniment petit actuel].

3. M. de Finetti fait ensuite très justement remarquer qu'on ne saurait éluder une difficulté de principe par une convention. Une fois posée la définition de la probabilité, d'une manière conforme à notre intuition, si cette définition permet d'attribuer une valeur à la probabilité d'un des événements classés comme non probabilisables, on n'a pas le droit d'exclure cet événement. La principale difficulté que soulève un essai de solution de cette question, réside dans le fait que, jusqu'à présent, aucune définition de la probabilité n'a obtenu une adhésion générale. Si on adopte le point de vue axiomatique, la solution est immédiate qui consiste à poser comme postulat le principe des probabilités totales sous sa forme complète. C'est, par exemple, ce que fait M. Paul Lévy dans son livre déjà cité. Mais sans se contenter de la commodité d'une convention arbitraire, cet auteur montre, page 330, et suivantes que cette convention est justifiée au point de vue concret.

Je voudrais signaler maintenant que si l'on adopte la définition empirique de la probabilité comme nous l'avons précisée M. Halbwachs et moi, dans notre petit livre ⁽¹⁾, on peut démontrer le principe *complet* des probabilités totales.

Convenons que la probabilité d'un événement fortuit est une grandeur physique dont la mesure expérimentale est la fréquence dans un nombreux groupe d'épreuves.

Alors, soient E_1, E_2, \dots des modalités incompatibles, (possédant chacune une probabilité déterminée) d'un événement fortuit E . Au cours de n épreuves, E_1 s'est produit r_1 fois, ... E_s s'est produit r_s fois, ... et E s'est produit $r_1 + r_2 + \dots + r_s + \dots = r$ fois. Les fréquences

$$f = \frac{r}{n}, f_1 = \frac{r_1}{n}, f_2 = \frac{r_2}{n}, \dots, f_s = \frac{r_s}{n}, \dots$$

(1) *Le Calcul des Probabilités à la portée de tous*, Dunod, Paris, 1924.

des évènements $E, E_1, E_2, \dots, E_s, \dots$ seront toujours telles que

$$(1) \quad f = f_1 + f_2 + \dots + f_s + \dots$$

Si on fait varier le groupe des n épreuves, on aura des valeurs différentes de f, f_1, f_2, \dots (1) mais on aura toujours l'égalité (1). Comme $f, f_1, f_2, \dots, f_s, \dots$ sont les mesures expérimentales des probabilités de $E, E_1, E_2, \dots, E_s, \dots$, nous sommes assez légitimement amenés à conclure que : le principe *complet* des probabilités totales est nécessairement vérifié dans la famille des évènements fortuits auxquels la définition empirique ci dessus attribue une probabilité.

4. J'arrive enfin à une objection de M. de Finetti qui me paraît d'une importance beaucoup moindre que les précédentes. Elle me semble relever du désir d'éviter quelque difficulté, désir que l'auteur semblait avoir banni quelques lignes auparavant. Si l'on adoptait le principe des probabilités totales sous sa forme complète, on ne pourrait plus, d'après M. de Finetti, dire que si une loi de probabilité tend vers une loi limite, cette loi limite est une loi de probabilité. Plus royaliste que le roi, je dirais, moi aussi, que si cela est incommode cela ne suffit pas pour le nier. Et d'ailleurs, cette incommodité subsisterait même en rejetant le principe complet des probabilités totales. Supposons, par exemple qu'un nombre aléatoire X_n ne puisse prendre que les valeurs $-n$ et $+n$, et avec les mêmes probabilités $\frac{1}{2}$. La fonction de probabilité totale de X_n tendra quand n croît indéfiniment vers une fonction limite. Mais cette fonction limite restant constamment égale à $\frac{1}{2}$ ne peut être une fonction de probabilité totale, du moins au sens ordinaire.

(1) Dans la série (1), les termes sont nécessairement nuls à partir d'un certain rang (sans quoi r serait infini), mais ce rang varie avec le groupe d'épreuves.

ANCORA SULL' ESTENSIONE
ALLE CLASSI NUMERABILI DEL TEOREMA
DELLE PROBABILITÀ TOTALI

Nota del dott. BRUNO de FINETTI

(Adunanza del 20 novembre 1930)

Sunto. — Si risponde a delle nuove obiezioni del FRÉCHET sull'argomento.

Ho ricevuto dal Prof. FRÉCHET la precedente nota che continua la discussione sulla legittimità dell'estensione alle classi numerabili del teorema delle probabilità totali, discussione che ebbe origine dalla mia nota « Sui passaggi al limite nel calcolo delle probabilità », pubblicata in questi Rend. (1), dove pure apparvero in seguito le prime obiezioni del FRÉCHET e la mia risposta (2). Di questa dice il FRÉCHET che « elle contient un certain nombre de remarques avec lesquelles je ne suis pas d'accord, mais qui cependant sont d'un grand intérêt. Je crois utile de les commenter, et j'ai rédigé à cet effet la note ci jointe ». Debbo essere assai grato al FRÉCHET d'essersi ancora interessato alle mie idee su quell'argomento, e sono lieto di rilevare nella Sua nota una disposizione d'animo molto più vicina alla mia di quanto dapprima sembrasse. Nè mi spiace che in qualche punto l'illustre A. abbia male interpretato il mio pensiero, chè avrò modo così di precisarlo meglio, rimuovendo la possibilità di erronee interpretazioni cui, come vedo, la precedente esposizione può prestarsi (3).

(1) Vol. LXIII, Fasc. II-V, 1930.

(2) id. id., Vol. LXIII, Fasc. XI-XV.

(3) Queste discussioni e il punto di vista che vi sostengo verranno indubbiamente assai chiarite conoscendo il mio modo di concepire la probabilità e di impostarne i principi. Un'esposizione succinta dei fondamenti del calcolo delle probabilità ispirata alle mie vedute si ha

1. (1) Io avrei dimostrato, secondo il FRÉCHET, che *non si può*, nell'attribuire a tutti gli eventi una probabilità, fare in modo che il teorema delle probabilità totali valga anche nelle classi numerabili di eventi incompatibili. Vi sarebbe una categoria di eventi dove tale estensione è valida, mentre per gli altri non lo sarebbe più. La mia opinione è invece molto diversa: io dico che è *sempre possibile* rispettare il principio *completo* delle probabilità totali, come chiama il FRÉCHET la proprietà precedente, ma che sono accettabili anche le leggi di probabilità che soddisfano soltanto il principio *ristretto*. Si può sempre attribuire cioè ad ogni evento E una probabilità $P(E)$ in modo che per un'infinità numerabile di eventi incompatibili E_1, E_2, \dots sia sempre

$$P(E_1 + E_2 + \dots) = P(E_1) + P(E_2) + \dots$$

ma è accettabile anche una legge di probabilità $P(E)$ che non soddisfi la condizione precedente, ma soltanto la

$$P(E_1 + \dots + E_n) = P(E_1) + \dots + P(E_n)$$

per n finito. Nessuna distinzione dunque fra *eventi* o *classi di eventi* per cui la proprietà sussiste o non sussiste, ma fra *leggi di probabilità* che, in un certo campo (per una certa classe d'eventi), la soddisfano o non la soddisfano.

Riferendoci all'esempio di un punto P scelto a caso nell'intervallo $(0,1)$, e in cui gli eventi che si considerano sono costituiti dall'appartenenza di P a un dato aggregato, l'opinione attribuitami dal FRÉCHET sarebbe questa: che non si può attribuire ad *ogni* aggregato una probabilità in modo da soddisfare la proprietà additiva completa, perchè un risultato analogo, dovuto al VITALI, sussiste per la *misura*. La difficoltà si eviterebbe limitandosi agli eventi che si potrebbero dire « *probabilizzabili* », corrispondenti agli aggregati *misurabili*: in tal

in « *Problemi determinati e indeterminati nel calcolo delle probabilità* », Rend. Lincei, nov. 1930; uno sviluppo più ampio e approfondito si ha nella memoria « *Sul significato soggettivo della probabilità* » che apparirà in « *Fundamenta Mathematicae* », Vol. XVII; una breve nota filosofica è « *Fondamenti logici del ragionamento probabilistico* », Boll. U. M. I., dic. 1930, mentre una critica filosofica completa è svolta in « *Probabilismo. Saggio critico sulla teoria delle probabilità e sul valore della scienza* », che sarà pubblicato nella rivista filosofica « *Logos* ».

(1) Rispondo al Prof. FRÉCHET punto per punto, usando la stessa numerazione.

modo *si sarebbe sicuri* di non incontrare dubbi. Il disaccordo fra il FRÉCHET e me starebbe in questo punto: nel contentarsi di restringersi a questo caso più semplice, il solo interessante in pratica, o nel volersi occupare anche dei casi meno trattabili. Io dico invece che esistono infinite leggi di probabilità estendibili a *tutti* gli aggregati e soddisfacenti la proprietà additiva *completa* (cfr. LEVY, *Calcul des Probabilités*, p. 330) (il teorema di VITALI esclude soltanto che una di queste leggi possa attribuire sempre probabilità uguali ad aggregati sovrapponibili); per rispetto a ciascuna di queste leggi di probabilità tutti gli aggregati danno luogo ad eventi probabilizzabili. Per evitare l'inconveniente non è dunque necessario limitare il campo degli eventi considerabili, ma, e questo è peggio, ciò non sarebbe neppure sufficiente. Se consideriamo la legge di probabilità in cui i valori possibili sono

$$1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \dots, \frac{1}{n}, \dots,$$

e sono ugualmente probabili, si ha un'antinomia anche limitandosi ai soli intervalli: i detti valori dividono infatti l'intervallo (0,1), che ha probabilità = 1, in un'infinità numerabile di intervalli che hanno evidentemente, secondo questa legge, probabilità nulla. Basta questo esempio semplicissimo a mostrare che la questione non riguarda soltanto dei casi complicati che è meglio trascurare, e trascurati i quali non abbiamo più a preoccuparcene: essa si riferisce collo stesso valore a tutti i casi dove il numero delle alternative è infinito. E si potrà accettare la mia opinione: che soltanto la proprietà additiva *ristretta* è necessaria, o l'opinione opposta, che esige la proprietà additiva *completa*, ma in nessun modo si può eludere la questione restringendosi a un campo più ristretto, a meno che non sia una classe finita.

Ciò potrebbe accadere sotto qualche particolare ipotesi restrittiva, come quelle del problema della misura, ma è ovvio che si tratterebbe allora di una proprietà accidentale di questa o quella particolare legge di probabilità, e non di un teorema di calcolo delle probabilità, sia pure soggetto a certe limitazioni. Ma mi preme aggiungere che per comprendere bene la questione è raccomandabile di liberarsi dagli esempi troppo artificiosi suggeriti dalla teoria della misura, dove è troppo facile lasciarsi suggestionare da esteriori analogie: le obiezioni che ho sollevato mi si presentarono trattando problemi con-

creti, relativi alla teoria delle funzioni a incremento aleatorio (1). Bastino due esempi. Sotto condizioni che è inutile qui rammentare, si può dimostrare che, comunque piccolo sia ε e comunque grande sia M , è sempre nulla la probabilità che, per una funzione f determinata da cause accidentali, esista in un intervallo assegnato almeno un intervallo di lunghezza $> \varepsilon$ in cui sia soddisfatta la condizione di LIPSCHITZ

$$|f(x_2) - f(x_1)| < M |x_2 - x_1|;$$

se volesse la proprietà additiva completa sarebbe lecito passare al limite ($M = \infty$, $\varepsilon = 0$), e dire che è nulla la probabilità che f ammetta derivate superiore e inferiore finite anche in un sol punto dell'intervallo. Per la proprietà conglomerativa (2), ecco quest'altro esempio. Supposto, sotto determinate condizioni che qui non interessano, che per un determinato valore x' dell'intervallo $(0, x)$ la funzione f assuma il valore

$$f(x') = \varepsilon \quad (\varepsilon > 0)$$

scende che è $> a$ la probabilità che sia $f(x) = \varepsilon$. Si può concludere che subordinatamente all'ipotesi che *esista* un valore x' ($0 \leq x' \leq x$) per cui $f(x') = \varepsilon$, è $> a$ la probabilità che sia $f(x) = \varepsilon$? Si potrebbe allora concludere che, detta b la probabilità che sia $f(x) \geq \varepsilon$ la probabilità che sia

$$f(x) = \varepsilon \text{ è } > a b$$

quando $f(0) = 0$ ed f sia necessariamente continua, perchè se $f(x) \geq \varepsilon$ deve infatti esistere allora almeno un valore

$$x' \quad (0 \leq x' \leq x)$$

per cui $f(x') = \varepsilon$. Tali conclusioni sarebbero molto importanti, ma dopo averle pubblicate mi vennero dei dubbi, che la critica dei passaggi al limite e della proprietà conglomerativa confermarono. Quale degli argomenti oppostimi potrebbe togliere di mezzo questi dubbi? Anche ammesso che esistano eventi « probabilizzabili » per cui ogni dubbio sia infondato, come farei a riconoscere in casi pratici come questi quali essi sono?

(1) Cfr. « *Sulle funzioni a incremento aleatorio* », « *Sulla possibilità di valori eccezionali per una legge di incrementi aleatori* », « *Integrazione delle funzioni a incremento aleatorio* », Rend. Lincei, 1929, II sem.; v. anche la conferenza « *Le leggi differenziali e la rinunzia al determinismo* », Seminario matematico di Roma, 5 aprile 1930.

(2) Cfr. « *Sulla proprietà conglomerativa delle probabilità subordinate* ». Questi Rend., Vol. LXIII, Fasc. VI-X, 1930.

3. Il FRÉCHET sembra riconoscere la possibilità di pensare un'infinità numerabile di casi possibili ugualmente probabili, ammettendo che la loro probabilità sia un « infinitesimo attuale ». Ma questa opinione coincide perfettamente colla mia, come si può vedere dalla mia nota « *Sulle probabilità numerabili e geometriche* » (1) dove ho appunto definito in generale il rapporto di due probabilità, che risulta finito se le due probabilità sono *commensurabili*, e cioè finite o infinitesime dello stesso ordine, risulta nullo o infinito se non lo sono.

Ma con questa ammissione (mi sembra che il FRÉCHET non l'abbia rilevato) si viene appunto ad abbandonare il principio completo delle probabilità totali, perchè la somma di n casi ha sempre probabilità infinitesima, e nessun passaggio al limite conduce quindi alla probabilità $= 1$ della somma di tutti i casi possibili.

3. Il punto più significativo della questione è però, almeno per me, nella necessità di dare una *dimostrazione* delle proprietà che si vogliono affermare. Dal punto di vista logico-formale, perchè sia lecito stabilire *per convenzione* che una proposizione sia vera è necessario che vi sia in essa almeno un termine che non è stato ancora nominalmente definito, e quindi che la frase non abbia ancora un significato empirico o logico ben determinato. Dire che sia questo il nostro caso, vorrebbe dire che si tratta il calcolo delle probabilità come una vana ed astratta esercitazione aritmetica.

Il FRÉCHET, pur dicendo di *non vedere la necessità* di tale dimostrazione (per i motivi criticati nel n. 1), osserva giustamente che la difficoltà principale che si incontra è la mancanza di una definizione e anzi di una concezione generalmente accettata della probabilità. È una difficoltà però assai più esteriore che intrinseca, e basterebbe che ogni Autore pensasse a dare una dimostrazione a modo suo, come cerca di fare il FRÉCHET, appoggiandosi al suo ben noto e apprezzato punto di vista. Mi propongo di riprendere la sua dimostrazione per far vedere che, precisandola meglio, vi si trova un ottimo argomento per chiarire e sostenere la mia tesi.

Supponiamo, per metterci dal punto di vista del FRÉCHET, che si possa fare un numero indefinito di « prove », in ciascuna

(1) Questi Rend., Vol. LXI, Fasc. XVI-XX, 1928.

delle quali si presenti uno e uno solo degli eventi (1) $E_1, E_2, \dots, E_s, \dots$ costituenti un'infinità numerabile. Siano $f_1^{(n)}, f_2^{(n)}, \dots, f_s^{(n)}, \dots$ le frequenze su n prove, avremo

$$f_1^{(n)} + f_2^{(n)} + \dots + f_s^{(n)} + \dots = 1$$

qualunque sia n , e le probabilità $p_1, p_2, \dots, p_s, \dots$ sarebbero, per n molto grande, sempre più prossime ad $f_1^{(n)}, f_2^{(n)}, \dots, f_s^{(n)}, \dots$. Potremo dire, matematicamente, che sono il limite delle frequenze (2) quando $n \rightarrow \infty$; ma allora è noto che non si può concludere che $p_1 + p_2 + \dots + p_s + \dots = 1$.

Ma è meglio lasciare le considerazioni astrattamente matematiche, e vedere il comportamento delle frequenze che è prevedibile quando sia $p_1 = p_2 = \dots = p_s = \dots = 0$ (nel caso, ad es., che le infinite alternative siano egualmente probabili). Su n prove è praticamente certo (probabilità = 1) che si verificano n eventi tutti distinti. Nella serie delle frequenze $f_1^{(n)}, f_2^{(n)}, \dots, f_s^{(n)}, \dots$ possiamo dire (con certezza pratica) che ve ne saranno n uguali ad $\frac{1}{n}$ e le altre saranno tutte nulle.

Pur di prendere n sufficientemente grande, possiamo fare in modo che tutte le frequenze siano (con certezza pratica) minori di un ε comunque assegnato (basta che $n > \frac{1}{\varepsilon}$): è assurdo dunque supporre che, al crescere di n , le frequenze tendano a dei limiti tali che $p_1 + p_2 + \dots + p_s + \dots = 1$, e quindi non nulli.

Se nella trattazione di questo problema si ammettessero i passaggi al limite che ho criticati, si dovrebbe concludere che vi è probabilità = 1 che in una successione infinita di prove nessun evento si presenti più di una volta; allora sarebbe certo (praticamente) che tutte le frequenze tendono a zero in senso matematico. Noi ci contentiamo di affermarlo per n assegnato ma comunque grande, il che è sufficiente.

4. L'ultima obiezione ha anche per me un valore del tutto secondario, come ha ragione di giudicare il FRÉCHET.

(1) Secondo la terminologia cui mi attengo, dovrei usare, in questo senso, la parola « fenomeno », ed ogni singola prova sarebbe un « evento ».

(2) Il FRÉCHET mi fa osservare che parlare del *limite* in senso matematico non corrisponde esattamente al suo punto di vista; ciò non infirma peraltro le considerazioni del capoverso successivo.

Mettendo in luce che adottando il mio punto di vista si rimuove un inconveniente formale, non cercavo con ciò di giustificarlo, ma solo, tutt' al più, di attenuare le prevenzioni che suscita una verità troppo scomoda. Tuttavia, non mi sembra che l' inconveniente sia proprio puramente formale: dire che il limite di una legge di probabilità è una legge di probabilità mi sembra una nozione intuitiva che, sebbene non abbia un vero e proprio valore logico, avrei qualche difficoltà ad abbandonare.

Con un suo esempio, il FRÉCHET vorrebbe provare che, pur adottando il mio punto di vista, tale proprietà non sussisterebbe, ma l' assurdo cui giunge è un assurdo soltanto quando si ammetta la proprietà additiva completa, che è quella appunto che vogliamo negare. È del resto ovvio che la proprietà additiva ristretta sussiste al limite: la somma di un numero finito di termini ha infatti sempre come limite la somma dei limiti (1). È solo quando si tratta di una serie infinita che la proprietà non sussiste, ed è perciò che la proprietà additiva completa, passando al limite, può cadere in difetto. Come appunto avevo rilevato.

(1) Cf. la nota già cit. sui « *Problemi determinati e indeterminati nel calcolo delle probabilità* ».