# SUR L'APPLICATION DU THÉOREME DE FUBINI AU CALCUL DES PROBABILITÉS

# SUR L'APPLICATION DU THÉOREME DE FUBINI AU CALCUL DES PROBABILITÉS

PAR

Paul Lévy (Paris).

---

1. — Considérons deux variables aléatoires $x$ et $t$, indépendantes l'une de l'autre et choisies au hasard entre 0 et 1 de manière qu'à chaque intervalle corresponde une probabilité égale à sa longueur. Le point $x$, $t$ d'abscisse $x$ et d'ordonnée $t$ se trouve choisi dans le carré unitaire S de manière que la notion de probabilité coïncide avec celle de mesure superficielle: la probabilité $\mathfrak{P}(\mathrm{E})$ que le point choisi appartienne à un ensemble E n'est bien définie que si cet ensemble est mesurable superficiellement et n'est autre que sa mesure. D'après le théorème de Fubini, on a dans ces conditions [1]

$$\mathfrak{P}(\mathrm{E}) = \mathfrak{M}\{\mathfrak{P}_x(\mathrm{E})\}, \tag{1}$$

---

[1] On sait que M. Fubini a établi en 1907 que, pourvu qu'une fonction $f(x, t)$ soit bornée et mesurable superficiellement, on peut calculer son intégrale dans S par deux intégrations successives, comme s'il s'agissait d'une fonction continue. Il suffit même que $f(x, t)$ soit sommable. Rappelons la démonstration très simple de la formule (1), relative au cas où $f(x, t)$ est la fonction égale à 1 pour les points de E et à zéro en dehors de cet ensemble.

Cette formule est évidente si l'ensemble E est composé d'un nombre fini de rectangles. Compte tenu du fait connu que, si la fonction mesurable bornée $\varphi_n(x)$ tend d'une manière monotone, quand $n$ croît, vers la fonction mesurable bornée $\varphi(x)$, on a

$$\lim_{n \to \infty} \int_0^1 \varphi_n(x)\, dx = \int_0^1 \varphi(x)\, dx ,$$

le résultat s'étend au cas d'un ensemble E′ formé par la réunion d'une infinité dénombrable de rectangles. La formule (2) du texte résulte alors de l'application de ce résultat à un ensemble mesurable contenant E et de mesure dépassant arbitrairement peu la mesure extérieure de E, et la formule (1), pour un ensemble mesurable, résulte de l'application de (2) à E et à son ensemble complémentaire.

$\mathcal{P}_x$ désignant la probabilité évaluée lorsque $x$ est connu, et $\mathcal{M}\{u\}$ désignant la valeur probable de la variable $u$.

Cette formule (1) subsiste si E est un événement dépendant du choix de deux variables aléatoires X et T dépendant de lois quelconques, indépendantes ou non. Ce cas se ramène en effet au précédent en liant chaque système de valeurs $\xi$ et $\tau$ de ces variables à deux variables auxiliaires $x$ et $y$ par les formules

$$x = \mathcal{P}(\mathrm{X} < \xi) \ , \qquad t = \mathcal{P}_\xi(\mathrm{T} < \tau) \ ;$$

sauf dans des cas de probabilité nulle, $\xi$ et $\eta$ sont des fonctions bien déterminées de $x$ et $y$.

2. — Dans le cas d'un ensemble non mesurable, la formule (1) subsiste seulement sous la forme

$$\mathcal{P}'(\mathrm{E}) \geqq \mathcal{M}'\{\mathcal{P}'_x(\mathrm{E})\} \ , \tag{2}$$

dans laquelle les accents indiquent qu'il s'agit de la mesure extérieure (ou, pour $\mathcal{M}'$, de l'intégrale supérieure). Mais on ne peut pas dans ce cas affirmer l'égalité des deux membres. Pour le montrer, nous allons indiquer l'exemple d'un cas où le second membre change lorsqu'on échange l'ordre des intégrations. On n'a d'ailleurs pas réussi jusqu'ici à nommer un ensemble non mesurable; nous n'y avons pas réussi plus que ceux qui l'ont essayé avant nous. C'est dire que notre exemple ne peut intéresser que les personnes qui admettent l'axiome de Zermelo.

La variable $x$ étant considérée par rapport au module 1, $\alpha$ étant un nombre irrationnel choisi une fois pour toutes, désignons par $\mathrm{F}_0$ l'ensemble des nombres $n\alpha$ ($n$ entier entre $-\infty$ et $+\infty$), par F un quelconque des ensembles déduits de $x$ par une translation, par $\mathrm{G}_0$ un ensemble ayant un élément commun et un seul avec chacun des F, et par $\mathrm{G}_n$ l'ensemble déduit de $\mathrm{G}_0$ par la translation $n\alpha$. Les $\mathrm{G}_n$ sont des ensembles disjoints, deux à deux superposables, et dont la réunion constitue tout l'intervalle (0, 1). Leur mesure ne pouvant évidemment être ni nulle, ni positive, ils ne sont pas mesurables. On sait que leur mesure intérieure est zéro et leur mesure extérieure un. Divisons, d'autre part, l'intervalle (0, 1) en une infinité dénombrables d'intervalles $\mathrm{H}_n$

($n$ étant toujours un entier de signe quelconque) dont le plus grand ait sa longueur inférieure à un nombre arbitrairement petit $\varepsilon$. Désignons par $E_n$ l'ensemble formé par les points du carré S dont l'abscisse appartient à $G_n$ et l'ordonnée à $H_n$, et par E l'ensemble formé par la réunion des $E_n$. On a évidemment, quels que soient $x$ et $t$,

$$\mathcal{T}_x(E) < \varepsilon \,, \qquad \mathcal{T}'_t(E) = 1 \;;$$

en effet $\mathcal{T}_x(E)$ est la mesure d'un $H_n$ et $\mathcal{T}'_t(E)$ est la mesure extérieure d'un $G_n$; on en déduit

$$\mathcal{M}'\{\mathcal{T}_x(E)\} < \varepsilon < \mathcal{M}\{\mathcal{T}'_t(E)\} = 1 \,, \qquad \text{c. q. f. d.}$$

3. — Théoriquement, l'hypothèse que E soit mesurable superficiellement est donc essentielle pour l'application du théorème de Fubini; même les personnes qui n'admettent pas l'axiome de Zermelo doivent en convenir. Pratiquement, on n'a jamais nommé d'ensemble non mesurable, et il y a évidemment à cela une raison assez profonde: les procédés dont on se sert pour définir un ensemble avec précision ne sont en général que des combinaisons d'opérations dont on sait une fois pour toutes qu'elles ne peuvent donner que des ensembles mesurables. Tant qu'on n'introduira pas en analyse un élément essentiellement nouveau, ou qu'on ne se contentera pas de parler d'un ensemble choisi parmi une infinité d'autres d'une manière que l'on ne précise pas, il est permis de penser que l'on ne nommera jamais que des ensembles mesurables, ce qui permet l'application du théorème de Fubini; ajoutons que même la vérification précise de ce résultat ne sera sans doute jamais difficile.

Cela s'applique naturellement en particulier à la théorie des probabilités dénombrables. C'est une remarque très simple, et qui pourtant permet de résoudre une difficulté qui m'avait d'abord arrêté, et qui, je crois, n'avait pas arrêté que moi. Aussi mérite-t-elle quelques développements.

4. — Un problème de probabilité dénombrable suppose essentiellement l'existence d'une suite de variables aléatoires, indé-

pendantes ou enchaînées, $x_1, x_2, \ldots, x_n, \ldots$; les lois dont elles dépendent sont considérées comme bien définies si l'on a, pour tout système d'inégalités de la forme

$$x_\nu < a_\nu \qquad (\nu = 1, 2, \ldots, n) \tag{3}$$

une probabilité bien définie. Rappelons que tous ces choix peuvent toujours se ramener au choix d'une variable $x$ entre 0 et 1 avec probabilité uniformément répartie dans cet intervalle, les $x_n$ étant des fonctions bien déterminées de $t$. D'autre part, des problèmes tels que la convergence ou la sommabilité de la suite des $x_n$, ou de la série $\Sigma\, x_n$, ou tels que la question de savoir si des inégalités du type (3) sont vérifiées une infinité de fois, ou un nombre fini donné de fois, ou bien d'autres problèmes de cette nature, conduisent toujours à des probabilités bien définies.

5. — Le problème dont nous allons nous occuper est celui de l'étude des fonctions $X(t)$ à accroissements aélatoires indépendants: si $t_0 < t_1 < t_2$, $X(t_1) - X(t_0)$ et $X(t_2) - X(t_1)$ doivent être des variables aléatoires indépendantes. La définition d'une loi de probabilité de cette nature se ramène à un problème de probabilité dénombrable. On peut en effet, au lieu de choisir $X(t_0)$, puis $X(t_1)$, puis $X(t_2)$, choisir $X(t_0)$, puis $X(t_2)$, puis $X(t_1)$ par des lois déduites des précédentes par les formules connues des probabilités composées et des probabilités *a posteriori*. La répétition de ce procédé d'interpolation permet de définir $X(t)$ pour des valeurs de $t$ constituant un ensemble E′ dénombrable partout dense dans un intervalle (0, 1). Tous ces choix peuvent se ramener au choix d'une variable unique $x$, comme nous l'avons dit au nº 4.

Pour les valeurs de $t$ n'appartenant pas à E′, $X(t)$ doit pouvoir être défini par un passage à la limite. Mais il y a une difficulté, provenant de ce que dans le cas le plus général $X(t)$ peut comprendre une partie indépendante du hasard susceptible d'avoir des discontinuités tout à fait quelconques. Il peut aussi exister des points où $X(t)$ augmente brusquement d'une quantité qui est une variable aléatoire. Mais j'ai montré antérieurement [1] que

[1] Paul Lévy, Sur les intégrales dont les éléments sont des variables aléatoires indépendantes. *Annali della R. Scuola Normale Superiore di Pisa*, série II, vol. III (1934), p. 343.

ces points ne peuvent que constituer une infinité dénombrable; je rappelle simplement que la démonstration est basée sur le fait que ces accroissements brusques (augmentés de constantes convenables) sont des variables indépendantes dont la somme a une dispersion finie, c'est-à-dire que les valeurs de $X(1) - X(0)$, en négligeant des cas très peu probables, est dans un intervalle fini. Ces discontinuités s'éliminent facilement et le cas général se ramène à celui où pour toute valeur donnée de $t$, il y a une probabilité unité que $X(t)$ soit continu.

J'avais alors, dans mon mémoire cité, raisonné de la manière suivante: si l'on choisit d'abord $t$, puis $X(t)$ (ou ce qui revient au même la variable auxiliaire $x$), il y a une probabilité unité que $X(t)$ soit continu au point choisi; ces deux choix étant indépendants, le résultat subsiste si on intervertit leur ordre; donc, sauf dans des cas de probabilité nulle, $X(t)$ est presque partout continu.

C'était appliquer le théorème de Fubini sans vérifier la légitimité de cette application. A une observation de M. Kolmogoroff sur ce point, j'ai déjà répondu [1] en montrant indirectement l'exactitude du résultat. Mais, d'après la remarque du n° 3 ci-dessus, il doit être facile de le vérifier directement en montrant que l'ensemble des points $t$, $x$ du carré S tels que la fonction $X(t)$ associée à $x$ soit continue en $t$ est mesurable superficiellement. C'est ce que je vais établir maintenant.

Donnons-nous deux nombres arbitrairement petits $\varepsilon$ et $\eta$, et désignons par $f_p(t)$ et $F_p(t)$ les valeurs extrêmes de $X(t)$ pour celles des valeurs $t_0, t_1, \ldots, t_p$ qui sont comprises entre $t - \eta$ et $t + \eta$; la probabilité

$$\varphi_p(t) = \mathfrak{P}[\,|F_p(t) - f_p(t)| \leq \varepsilon]$$

est par hypothèse bien déterminée, et, étant constante dans chacun des intervalles séparés par les points $t_i \pm \eta$, a une moyenne $\alpha_p$ qui est la mesure d'un ensemble $E_p$ dans le carré S. Or $E_p$ est contenu dans $E_{p+1}$, de sorte que pour $p$ infini on obtient un ensemble E, mesurable et de mesure $\alpha = \lim \alpha_p$, qui est l'ensemble des points $t$, $x$ pour lesquels l'oscillation de $X(t)$ entre

---

[1] Même recueil que le mémoire cité plus haut, vol. IV (1935), p. 217.

$t - \eta$ et $t + \eta$ ne dépasse pas $\varepsilon$. En faisant tendre $\eta$ vers zéro, on obtient un ensemble mesurable formé par la réunion de ceux qui correspondent aux valeurs positives de $\eta$ et à la même valeur de $\varepsilon$; c'est l'ensemble des points $t$, $x$ pour lesquels les valeurs limites de $X(t')$ quand $t'$ tend vers $t$ différent au plus de $\varepsilon$. Faisant enfin tendre $\varepsilon$ vers zéro, et prenant la partie commune à tous les ensembles précédents, on obtient l'ensemble des points $t$, $x$ pour lesquels la fonction $X(t')$ est continue pour $t' = t$; il est mesurable, c.q.f.d.

Cette mesure est donc nulle, c'est-à-dire qu'il est presque sûr que la fonction $X(t)$ est presque partout continue. Il est même presque sûr qu'elle n'admet pas d'autres discontinuités qu'une infinité dénombrable de points de discontinuité de première espèce; je ne puis que renvoyer pour ce point à mes mémoires cités plus haut, sa démonstration, assez délicate, ne me semblant pas pouvoir de la même manière être simplifiée par l'application du théorème de Fubini.

---

## SUR LA MESURE DES GRANDEURS [1]

PAR

Henri LEBESGUE, Membre de l'Institut (Paris).

---

### VI. — GRANDEURS MESURABLES.

84. — Le programme de la première classe de l'Enseignement secondaire, la classe de Sixième, comporte un chapitre: Mesure des grandeurs, notion de fraction. Le programme de la dernière classe de l'Enseignement secondaire, celle de Mathématiques, prévoit ce même chapitre: Mesure des grandeurs. D'une classe à l'autre le point de vue devrait être très différent et à cause de l'âge des élèves et parce qu'il devrait s'agir de notion pratique en Sixième, de notion abstraite en Mathématiques.

---

[1] Voir *L'Enseignement mathématique*, XXXI[e] année, p. 173-206. — XXXII[e] année, p. 23-51. — XXXIII[e] année, p. 22-48; p. 177-213.