E. BOREL

Remarques sur certaines questions de probabilité

Bulletin de la S. M. F., tome 33 (1905), p. 123-128
<http://www.numdam.org/item?id=BSMF_1905__33__123_1>

REMARQUES SUR CERTAINES QUESTIONS DE PROBABILITÉ;

Par M. Émile Borel.

1. On sait que les questions de probabilité où interviennent des variables continues ne peuvent acquérir de sens qu'en vertu de *conventions* précises. Comme le fait observer Joseph Bertrand, si une variable x est assujettie à rester comprise entre o et 1, son carré x^2 est assujetti aux mêmes conditions et la probabilité pour que x soit compris entre o et $\frac{1}{2}$ est égale à la probabilité pour que x^2 soit compris entre o et $\frac{1}{4}$. Cela serait absurde (1) si l'on supposait à chacune de ces probabilités une valeur *intrinsèque*, c'est-à-dire définie objectivement d'une manière indépendante de toute convention.

(1) *Voir* Poincaré, *Calcul des probabilités*, 8e Leçon.

La convention la plus commode, au moins dans le cas où l'ensemble des valeurs possibles de la ou des variables est *borné,* consiste à regarder la probabilité comme proportionnelle à l'*étendue :* longueur, aire, volume, suivant qu'il y a une, deux, trois dimensions.

A la définition de la probabilité se rattache celle de la valeur moyenne; si, dans un intervalle x_0, x_1, on connaît les valeurs d'une fonction réelle $f(x)$, sa valeur moyenne est, par définition,

$$\frac{1}{x_1 - x_0} \int_{x_0}^{x_1} f(x)\, dx.$$

De même, la valeur moyenne d'une fonction de deux variables x et y définie à l'intérieur d'une aire S du plan de ces variables est, par définition,

$$\frac{\int\int_S f(x, y)\, dx\, dy}{\int\int_S dx\, dy}.$$

Si une fonction f est connue en tout point A d'une portion de surface Σ, sur laquelle est défini un système de coordonnées curvilignes u, v, on peut définir la valeur moyenne de f par l'expression

$$\frac{\int\int_\Sigma f(\mathrm{A})\, \varphi(u, v)\, du\, dv}{\int\int_\Sigma \varphi(u, v)\, du\, dv},$$

dans laquelle $\varphi(u, v)$ désigne une fonction positive arbitrairement choisie une fois pour toutes. On choisira généralement $\varphi(u, v)$ par la condition que l'intégrale qui figure au dénominateur représente l'aire de Σ; mais ce choix n'est nullement nécessaire; il pourrait être parfois plus commode de prendre $\varphi(u, v) = 1$, par exemple.

Tout ce qui précède est bien connu, mais les questions de probabilité ont donné lieu à tant de controverses de mots, provenant simplement d'un défaut d'entente sur les conventions de langage, qu'il ne m'a pas paru inutile de préciser les notions dont je me servirai. Je me bornerai d'ailleurs au cas *d'une dimension;* il n'y aurait aucune difficulté à étendre ce qui suit au cas général d'un nombre quelconque de dimensions.

2. *Une fonction $f(x)$ est définie pour toutes les valeurs de x comprises entre* 0 *et* 1; *elle est égale à* 1 *si x est commensurable et à* 0 *si x est incommensurable; quelle est sa valeur moyenne?*

Cette question est évidemment équivalente à la suivante :

On sait que x est compris entre 0 *et* 1; *quelle est la probabilité pour que x soit commensurable?*

La réponse évidente [1] aux deux questions précédentes est *zéro*. Cette réponse peut-elle se déduire des formules que nous avons rappelées? Il ne le semble pas, si l'on se borne aux définitions classiques de l'intégrale. Considérons, en effet, une fonction $f(x)$ égale à 0 pour x incommensurable et à 1 pour x commensurable et une fonction $F(x)$ égale à 1 pour x incommensurable et à 0 pour x commensurable; les intégrales

$$\int_0^1 f(x)\,dx, \qquad \int_0^1 F(x)\,dx$$

sont toutes deux dépourvues de sens; pour chacune d'elles l'intégrale inférieure, au sens de M. Darboux, est égale à *zéro* et l'intégrale supérieure égale à *un*. Donc, quelle que soit la convention que l'on adopte, si cette convention ne fait dépendre la valeur moyenne que des intégrales précédentes, calculées au sens classique, on devra attribuer la même valeur moyenne à $f(x)$ et à $F(x)$. Or, cette conclusion est absurde, car la valeur moyenne de $f(x)$ est 0 et la valeur moyenne de $F(x)$ est 1 [2].

Mais, si l'on utilise la nouvelle définition de l'intégrale qui est due à M. Lebesgue [3], on reconnaît que chacune des fonctions $f(x)$ et $F(x)$ est intégrable au sens de M. Lebesgue ou, plus briè-

---

(1) Poincaré, *Calcul des probabilités*, p. 126.

(2) Si l'on n'admettait pas ce point, on pourrait observer que la somme des valeurs moyennes de $f(x)$ et de $F(x)$ doit être égale à la valeur moyenne de $f(x) + F(x)$, c'est-à-dire à 1.

(3) Lebesgue, *Leçons sur l'intégration et la recherche des fonctions primitives*.

vement, intégrable (L) et que leur intégrale (L) fournit la valeur correcte de la valeur moyenne ou de la probabilité cherchée. Les méthodes de M. Lebesgue permettent donc d'étudier des questions de probabilité qui paraissent inaccessibles par les procédés d'intégration classiques. D'ailleurs, dans les cas particuliers les plus simples, il suffira de se servir de la théorie des ensembles que j'avais appelés *mesurables* [1] et auxquels M. Lebesgue a donné le nom de *mesurables* (B); l'application de cette théorie des ensembles mesurables au calcul des probabilités a été, à ma connaissance, faite pour la première fois par M. Wiman [2].

3. Rappelons brièvement les principes de la théorie de la mesure des ensembles linéaires de points.

Par définition, *l'ensemble composé de tous les points d'un intervalle a pour mesure la longueur de cet intervalle;* ces ensembles sont les premiers que l'on sache mesurer; on leur donne le nom de *mesurables.* On étend ensuite la catégorie des ensembles mesurables à l'aide des conventions suivantes : 1° *la somme d'une infinité dénombrable d'ensembles mesurables deux à deux sans points communs est mesurable et a pour mesure la somme des mesures*; 2° *si un ensemble mesurable* E *contient tous les points d'un autre ensemble mesurable* F, *l'ensemble* E — F *formé de tous les points de* E *qui n'appartiennent pas à* F *est mesurable et a pour mesure la différence des mesures de* E *et de* F; 3° *Si un ensemble* E *est contenu dans un ensemble mesurable* A *de mesure* α *et contient un ensemble mesurable* B *de mesure* β, *nous dirons que la mesure de* E *est inférieure ou égale à* α *et supérieure ou égale à* β; 4° *Si l'on peut démontrer que la mesure d'un ensemble* E *est, quel que soit* ε, *inférieure ou égale à* $m + \varepsilon$ *et supérieure ou égale à* $m - \varepsilon$, *l'ensemble* E *est mesurable et a pour mesure* m [3].

---

(1) Borel, *Leçons sur la théorie des fonctions*, Chap. III.

(2) A. Wiman, *Ueber eine Wahrscheinlichkeitsaufgabe bei Kettenbruchentwicklungen* (*Kongl. Vetenskaps Akademien Förhandlingar*, 1900, p. 829) et *Bemerkungen über eine von Gylden aufgeworfene Wahrscheinlichkeitsfrage* (Lund, 1901, Håkan Ohlssons Boktryckeri).

(3) L'introduction explicite et générale de la convention 4° est due à M. Lebesgue; j'en avais fait usage dans un cas particulier, sans l'énoncer (*Leçons sur la théorie des fonctions*, p. 67).

Ces conventions permettent de construire de proche en proche des ensembles mesurables, en nombre infini, dont on se trouve connaître la mesure par leur construction même; elles ne donnent pas immédiatement une méthode simple pour le calcul de la mesure d'un ensemble donné, au cas où il est mesurable; une telle méthode résulte des recherches de M. Lebesgue. Mais, pour les applications que j'ai en vue, il m'a paru préférable de rester à mon point de vue primitif; celui de M. Lebesgue doit, dans d'autres questions, être choisi de préférence.

Rappelons que le point essentiel, dans la théorie de la mesure, consiste à démontrer que les définitions et conventions précédentes ne conduisent jamais à des contradictions. Il est inutile de reproduire ici cette démonstration.

4. Considérons une variable x assujettie à rester comprise entre o et 1; soient E un ensemble mesurable de points compris entre o et 1 et m la mesure de E. Par définition, la probabilité pour que x appartienne à E est m et la probabilité pour que x n'appartienne pas à E est $1 - m$.

Si E est l'ensemble des valeurs commensurables, on trouve immédiatement $m = 0$, ce qui donne les résultats énoncés plus haut.

Traitons un cas un peu plus compliqué.

Nous dirons qu'un nombre α compris entre o et 1 est commensurable aux unités près du $n^{\text{ième}}$ ordre s'il existe deux entiers premiers entre eux p et q, tels que l'on ait

$$(1)\qquad \left|\frac{p}{q} - \alpha\right| < \frac{1}{q^n}, \qquad q \geqq 2, \qquad p < q.$$

Quelle est la probabilité pour que x, assujetti à être compris entre o *et* 1, *soit commensurable aux unités près du $n^{\text{ième}}$ ordre?*

A chaque fraction irréductible $\frac{p}{q}$, faisons correspondre l'intervalle

$$E_{p,q}^{(n)}: \qquad \frac{p}{q} - \frac{1}{q^n}, \qquad \frac{p}{q} + \frac{1}{q^n};$$

sa longueur est $\frac{2}{q^n}$; la longueur des intervalles analogues pour une valeur donnée de q est

$$\frac{2\varphi(q)}{q^n},$$

en désignant par $\varphi(q)$ le nombre des nombres premiers à q et inférieurs à q.

L'ensemble $E^{(n)}$ des points distincts contenus dans les ensembles $E^{(n)}_{p,q}$ est évidemment mesurable; sa mesure e_n est inférieure à m_n en posant

$$m_n = 2\sum_{q=2}^{\infty} \frac{\varphi(q)}{q^n}.$$

La série est convergente pour $n > 2$; si $n = 2$, on sait, en effet, que tout nombre α satisfait d'une infinité de manières à l'inégalité (1). En supposant n au moins égal à 3, la série m_n est convergente et a une valeur visiblement inférieure à 1, car on a

$$\varphi(q) \leqq q - 1$$

et

$$2\sum_{2}^{\infty} \frac{q-1}{q^3} = \frac{\pi^2}{3}\left(1 - \frac{\pi}{4}\right) < 1.$$

Si l'on convient, lorsqu'un nombre incommensurable α vérifie l'inégalité (1) pour plusieurs systèmes de valeurs de p et de q, de lui donner un coefficient égal à ce nombre de fois [1], nous pourrons dire que *la probabilité pour que x soit commensurable aux unités près du $n^{\text{ième}}$ ordre est précisément* m_n; mais, si l'on ne fait pas cette convention, la probabilité est $e_n < m_n$; il est aisé de calculer e_n avec autant d'approximation que l'on veut; il n'est peut-être pas possible de l'exprimer en termes finis au moyen de nombres algébriques ou de nombres transcendants usuels. La probabilité pour que *x ne soit pas commensurable aux unités près du $n^{\text{ième}}$ ordre* est $1 - e_n$.

5. Dans une prochaine Note, j'indiquerai des applications des remarques précédentes aux problèmes les plus simples qui se posent dans les recherches de Mécanique et de Physique mathématique où intervient le calcul des probabilités.

---

(1) On est ainsi conduit à donner à certains α un coefficient infini; le fait que la probabilité reste finie doit être signalé.