BULLETIN DE LA S. M. F.

J. HAAG

Sur une question de probabilité

Bulletin de la S. M. F., tome 51 (1923), p. 219-223.

<http://www.numdam.org/item?id=BSMF_1923__51__219_1>

SUR UNE QUESTION DE PROBABILITÉ ;

PAR M. J. HAAG.

1. M. Borel a démontré (*Introduction géométrique à quelques théories physiques*, note V) que, si des points M sont distribués au hasard sur une droite indéfinie, à raison d'un point en moyenne par unité de longueur, la probabilité pour qu'il y en ait n dans un intervalle de longueur x est $e^{-x}\frac{x^n}{n!}\cdot$ Je me propose d'étendre cette formule au cas où *la densité des points le long de la droite n'est pas uniforme*.

2. Soit $f(x)$ la densité au point A, d'abscisse x. Cela veut dire que la probabilité pour qu'il y ait au moins un point M dans le segment $(x, x+dx)$ est $f(x)\,dx$.

Il est à peu près évident que la probabilité pour qu'il y en ait plus d'un est du second ordre par rapport à dx. Démontrons-le néanmoins.

Si p est son ordre infinitésimal, elle est de la forme $A(dx)^p$, A étant une certaine fonction de x. Considérons alors, de part et d'autre de A, deux intervalles infiniment petits du premier ordre ε et ε'. Pour qu'il y ait au moins deux points dans l'intervalle total $\varepsilon+\varepsilon'$, il faut qu'il n'y en ait pas dans ε et qu'il y en ait au moins deux dans ε'; ou bien qu'il n'y en ait pas dans ε' et qu'il y en ait au moins deux dans ε; ou enfin qu'il y en ait au moins un dans ε et au moins un dans ε'. On a donc

$$A(\varepsilon+\varepsilon')^p = A(\varepsilon^p+\varepsilon'^p)+f^2\varepsilon\varepsilon'.$$

Ceci exige $p=2$ et l'on a ensuite $A=\frac{f^2}{2}\cdot$

La probabilité pour qu'il y ait juste un point dans dx est, au second ordre près, $f(x)\,dx$.

La probabilité pour qu'il n'y en ait aucun est $1-f(x)\,dx$.

3. Cela posé, nous commencerons par calculer la probabilité P_0 pour qu'il n'y ait aucun point sur le segment (o, x).

Pour qu'il n'y ait aucun point sur le segment $(o, x+dx)$, il faut et il suffit qu'il n'y en ait pas sur le segment (o, x), ni sur le segment $(x, x+dx)$. Dès lors, si nous admettons, comme nous l'avons d'ailleurs fait implicitement tout à l'heure, que *les probabilités relatives à des segments différents sont indépendantes*, nous avons

$$P_0 + dP_0 = P_0[1 - f(x)\,dx]$$

ou

$$\frac{dP_0}{P_0} = -f(x)\,dx;$$

d'où

$$P_0 = e^{-\int_0^x f(x)\,dx} = e^{-I}, \tag{1}$$

en posant

$$I = \int_0^x f(x)\,dx.$$

4. Cherchons maintenant la probabilité P_n pour qu'il y ait n points sur le segment (o, x). Il peut y avoir n points sur le segment $(o, x+dx)$ de trois manières différentes dont chacune exclut les deux autres :

1° Il y a n points sur (o, x) et zéro sur $(x, x+dx)$; la probabilité correspondante est $P_n[1 - f(x)\,dx]$;

2° Il y a $n-1$ points sur (o, x) et 1 point sur $(x, x+dx)$; la probabilité correspondante est $P_{n-1} f(x)\,dx$;

3° Il y a $n-p$ points sur (o, x) et p points sur $(x, x+dx)$, p étant plus grand que 1; la probabilité correspondante est infiniment petite du second ordre au moins par rapport à dx.

On a donc, au second ordre près,

$$P_n + dP_n = P_n[1 - f(x)\,dx] + P_{n-1} f(x)\,dx$$

ou

$$\frac{dP_n}{dx} + P_n f(x) = P_{n-1} f(x). \tag{2}$$

Posons

$$P_n = e^{-I} Q_n.$$

L'équation (2) devient

$$\frac{dQ_n}{dx} = Q_{n-1} f(x). \tag{3}$$

Comme nous savons que $Q_0 = 1$ et que Q_n doit s'annuler avec x, dès que $n \geqq 1$, on peut calculer, par des quadratures successives, toutes les fonctions Q_n. Ces quadratures s'effectuent aisément, en intégrant par parties. Mais, il est plus simple de vérifier *a posteriori* que la fonction

$$Q_n = \frac{I^n}{n!}$$

satisfait à l'équation (3) et aux conditions ci-dessus.

Nous avons donc, en définitive,

$$P_n = e^{-I} \frac{I^n}{n!}. \tag{4}$$

C'est bien la formule de M. Borel, *la longueur x de l'intervalle devant être simplement remplacée par la valeur probable du nombre de points situés dans cet intervalle.* Cette valeur probable est, en effet,

$$\sum_{n=1}^{\infty} n P_n = I e^{-I} \sum_{n=1}^{\infty} \frac{I^{n-1}}{(n-1)!} = I.$$

5. Voici maintenant une autre méthode, qui a l'avantage de s'appliquer à des *points répartis dans un espace à un nombre quelconque de dimensions*. Pour simplifier, nous supposerons seulement trois dimensions et nous appellerons $f(x, y, z)$ la densité au point $A(x, y, z)$, ce qui signifie que la probabilité pour qu'il y ait au moins un point dans un volume dV entourant A est $f(x, y, z)\, dV$.

Appelons encore P_n la probabilité pour qu'il y ait juste n points dans un volume donné V et commençons toujours par calculer P_0.

Partageons V en volumes infiniment petits $v_1, v_2, \ldots, v_m$. La probabilité pour qu'il n'y ait aucun point dans aucun de ces volumes est

$$(1 - f_1 v_1)(1 - f_2 v_2) \ldots (1 - f_m v_m).$$

Son logarithme est, au second ordre près,

$$-(f_1 v_1 + f_2 v_2 + \ldots + f_m v_m).$$

Lorsque les volumes élémentaires tendent vers zéro, la somme entre parenthèses tend vers l'intégrale triple

$$(5) \qquad \mathrm{I} = \int\int\int_{\mathrm{V}} f(x, y, z)\, d\mathrm{V}.$$

Il s'ensuit que

$$(6) \qquad \mathrm{P}_0 = e^{-\mathrm{I}}.$$

6. Cherchons maintenant la relation entre P_n et P_{n-1}. Parmi toutes les répartitions possibles de points M dans l'espace, appelons répartitions de la classe C_i celles pour lesquelles il y a un point dans le volume élémentaire v_i et $n - 1$ points dans la portion de V extérieure à v_i. Appelons répartition de la classe C toute répartition pour laquelle il y a n points dans V.

Si une répartition est de la classe C, elle appartient à la fois à n classes C_i, sauf si plusieurs points se trouvaient être dans un même volume élémentaire. Mais, cette dernière circonstance a une probabilité infiniment petite du second ordre au moins, si les volumes v_i sont infiniment petits du premier ordre. Il suit de là que l'on a, au second ordre près,

$$(7) \qquad n\mathrm{P}_n = p_1 + p_2 + \ldots + p_m,$$

en appelant p_i la probabilité pour qu'une répartition appartienne à la classe C_i. Or,

$$(8) \qquad p_i = f_i v_i \mathrm{P}_{n-1},$$

en remplaçant la probabilité pour qu'il y ait $n - 1$ points dans $\mathrm{V} - v_i$, par la probabilité pour qu'il y ait $n - 1$ points dans V, ce qui n'entraîne qu'une erreur du second ordre sur p_i. Portant (8) dans (7), il vient, toujours au second ordre près,

$$n\mathrm{P}_n = \mathrm{P}_{n-1}(f_1 v_1 + f_2 v_2 + \ldots + f_m v_m).$$

A la limite, quand les volumes élémentaires tendent vers zéro, on a

$$n\mathrm{P}_n = \mathrm{P}_{n-1}\mathrm{I}.$$

Telle est la relation de récurrence cherchée. On en déduit immédiatement

$$(9) \qquad \mathrm{P}_n = \mathrm{P}_0 \frac{\mathrm{I}^n}{n!} = e^{-\mathrm{I}} \frac{\mathrm{I}^n}{n!}.$$

On retrouve encore une fois la formule de M. Borel, où I désigne toujours la valeur probable du nombre de points situés dans V.

7. A titre de vérification, on peut faire le raisonnement suivant : Partageons V en m volumes quelconques $v_1, v_2, \ldots, v_m$. Pour qu'une répartition appartienne à la classe C, il faut et il suffit qu'elle ait a_1 points dans V_1, a_2 points dans $v_2, \ldots, a_m$ points dans v_m, avec

$$a_1 + a_2 + \ldots + a_m = n. \tag{10}$$

Or, la probabilité d'une telle répartition est

$$e^{-I_1}\frac{I_1^{a_1}}{a_1!}e^{-I_2}\frac{I_2^{a_2}}{a_2!}\cdots e^{-I_m}\frac{I_m^{a_m}}{a_m!} = e^{-I}\frac{I_1^{a_1}I_2^{a_2}\ldots I_m^{a_m}}{a_1!\,a_2!\ldots a_m!},$$

en posant

$$I = I_1 + I_2 + \ldots + I_m.$$

La probabilité totale pour qu'une répartition appartienne à la classe C est

$$P_n = \Sigma e^{-I}\frac{I_1^{a_1}I_2^{a_2}\ldots I_m^{a_m}}{a_1!\,a_2!\ldots a_m!}$$

la sommation étant étendue à toutes les valeurs possibles des a_i satisfaisant à (10). En comparant avec (9), on obtient la formule bien connue qui donne le développement de $(I_1 + I_2 + \ldots + I_m)^n$.