

Lois de probabilité et polynomes d'Hermite.

Autor(en): **Mirimanoff, D.**

Objektyp: **Article**

Zeitschrift: **Commentarii Mathematici Helvetici**

Band (Jahr): **3 (1931)**

PDF erstellt am: **22.07.2024**

Persistenter Link: <https://doi.org/10.5169/seals-4689>

Nutzungsbedingungen

Die ETH-Bibliothek ist Anbieterin der digitalisierten Zeitschriften. Sie besitzt keine Urheberrechte an den Inhalten der Zeitschriften. Die Rechte liegen in der Regel bei den Herausgebern.

Die auf der Plattform e-periodica veröffentlichten Dokumente stehen für nicht-kommerzielle Zwecke in Lehre und Forschung sowie für die private Nutzung frei zur Verfügung. Einzelne Dateien oder Ausdrucke aus diesem Angebot können zusammen mit diesen Nutzungsbedingungen und den korrekten Herkunftsbezeichnungen weitergegeben werden.

Das Veröffentlichen von Bildern in Print- und Online-Publikationen ist nur mit vorheriger Genehmigung der Rechteinhaber erlaubt. Die systematische Speicherung von Teilen des elektronischen Angebots auf anderen Servern bedarf ebenfalls des schriftlichen Einverständnisses der Rechteinhaber.

Haftungsausschluss

Alle Angaben erfolgen ohne Gewähr für Vollständigkeit oder Richtigkeit. Es wird keine Haftung übernommen für Schäden durch die Verwendung von Informationen aus diesem Online-Angebot oder durch das Fehlen von Informationen. Dies gilt auch für Inhalte Dritter, die über dieses Angebot zugänglich sind.

Lois de probabilité et polynomes d'Hermite

par D. MIRIMANOFF, Genève

Introduction

On sait qu'il est commode, dans l'étude des problèmes relatifs au calcul des probabilités et à la statistique, de représenter une loi de probabilité à une variable x par des masses réparties d'une certaine manière le long de l'axe des x . Le cas le plus simple est celui où la variable aléatoire x ne prend qu'un nombre fini ou une infinité dénombrable de valeurs possibles x_i . Si l'on désigne par y_i les probabilités correspondantes, la loi de probabilité sera représentée par des masses égales à y_i concentrées aux points d'abscisses x_i , masses de première classe d'après M. P. Lévy ¹⁾.

Mais les masses peuvent aussi être réparties sur l'axe des x avec une densité qui est une certaine fonction $f(x)$ de la variable aléatoire x , fonction sommable, jamais négative. On a alors des masses de deuxième classe. Ordinairement la fonction $f(x)$ est continue, mais dans ce travail nous aurons surtout à envisager des fonctions $f(x)$ ayant des discontinuités de première espèce.

M. P. Lévy distingue encore des masses de troisième classe; ce sont des masses réparties sur un ensemble de mesure nulle, sans qu'aucun point contienne de masse finie. Nous laisserons ce cas de côté. Nous laisserons aussi de côté le cas mixte où les masses des deux premières classes se présentent à la fois.

Si maintenant on porte en ordonnée les masses y_i ou la densité $f(x)$, on obtient dans le cas des masses de première classe un ensemble de points (x_i, y_i) et dans le cas des masses de deuxième classe une courbe $y = f(x)$. Dans le premier cas l'ordonnée y est une fonction de x égale à y_i pour $x = x_i$ et nulle partout ailleurs. Dans le second on suppose parfois que y s'annule en dehors d'un certain intervalle.

Si l'on transporte l'origine au centre de gravité des masses, la variable x devient l'écart, qu'on désigne souvent par l . Alors le moment des masses par rapport à cette origine est nul.

¹⁾ *Paul Lévy*, Calcul des Probabilités, Gauthier-Villars, Paris, 1925, p. 142.

Nous aurons à envisager des moments d'ordre quelconque. Si les masses sont de première classe, les moments d'ordre k par rapport au centre de gravité sont les sommes

$$(1) \quad \mu_k = \sum_i x_i^k y_i$$

$\mu_0 = 1$, $\mu_1 = 0$ et μ_2 est le moment d'inertie.

Si les masses sont de deuxième classe, les moments, que je désignerai par m_k , sont donnés par les intégrales

$$(2) \quad m_k = \int_{-\infty}^{\infty} x^k f(x) dx.$$

Je rappelle encore qu'on simplifie souvent l'étude d'une loi de probabilité en introduisant l'écart réduit t lié à l'écart x par la relation $x = at$, où a est une constante convenablement choisie.

La plupart des statisticiens modernes posent $a = \sqrt{\mu_2} = \sigma$, où σ est la dispersion, ou l'écart moyen quadratique. Mais il est plus commode pour le but que j'ai en vue, de poser, comme on le faisait autrefois,

$$a = \sqrt{2\mu_2} = \sigma\sqrt{2}, \text{ d'où } t = \frac{x}{\sigma\sqrt{2}}.$$

Rappelons enfin une propriété très simple des séries d'Hermite. Supposons qu'une fonction $f(x)$ soit développable suivant les polynômes d'Hermite ou suivant ces polynômes multipliés par $e^{-\alpha x^2}$, où α est une constante positive. Il suffit pour cela que $f(x)$ vérifie des conditions très larges indiquées par M. M. Plancherel et Rotach²⁾. Il suffit en particulier que $f(x)$ soit nulle en dehors d'un certain intervalle (A, B) et qu'à l'intérieur de cet intervalle elle n'ait que des discontinuités de première espèce en nombre fini et un nombre fini de maxima et de minima³⁾. On pourra écrire alors, en posant $\alpha = 1$

$$(3) \quad f(x) = \frac{e^{-x^2}}{\sqrt{\pi}} \sum_{n=0}^{\infty} c_n H_n(x)$$

²⁾ W. Rotach, Reihenentwicklungen einer willkürlichen Funktion nach Hermite'schen und Laguerre'schen Polynomen, Dissertation, Zürich, 1925.

M. Plancherel et W. Rotach, Sur les valeurs asymptotiques des polynômes d'Hermite, Commentarii Math. Helv., t. 1, 1929, p. 227.

³⁾ H. Galbrun, Sur un développement d'une fonction à variable réelle en séries de polynômes, Bull. Soc. Math. de Fr. t. 41, 1912, p. 24.

où $H_n(x)$ sont les polynômes d'Hermite définis par la relation

$$(4) \quad H_n(x) = e^{x^2} \frac{d^n}{dx^n} (e^{-x^2}).$$

(Ici encore je m'écarte des notations modernes).

On voit immédiatement, et c'est là la propriété que je tenais à rappeler, que les coefficients c_n sont des fonctions linéaires des moments de $f(x)$. En effet on peut écrire symboliquement

$$(5) \quad c_n = \frac{1}{2^n n!} H_n(m).$$

en convenant de remplacer dans $H_n(m)$ les puissances m^k ($k = 0, 1, \dots$) par les moments m_k .

1. Séries d'Hermite et expressions approchées

Ceci rappelé, supposons qu'on se trouve en présence d'une répartition de deuxième classe et que la fonction $f(x)$ soit développable en série d'Hermite. En arrêtant le développement à l'un de ses premiers termes, on aura une expression approchée de $f(x)$ et, chose essentielle, pour la calculer, il suffira de connaître les premiers moments de $f(x)$.

Mais supposons maintenant que la loi de probabilité soit de première classe, ce qui se présente par exemple dans le cas des épreuves répétées de Bernoulli. La fonction y qui se réduit à y_i pour $x = x_i$ et qui est nulle pour tout $x \neq x_i$ ne saurait être développée en série d'Hermite. Comment allons-nous procéder alors pour avoir des expressions approchées des y_i ?

La première idée qui vient à l'esprit est d'envisager la série qui s'obtient de la série d'Hermite (3) en y remplaçant les moments m par les moments μ des masses y_i , ce qui revient à traiter y comme une fonction de deuxième classe dont les moments m seraient égaux aux μ . On obtient ainsi les séries A de Charlier. Mais ces séries divergent pour tout $x = x_i$, comme l'a montré récemment M. Mosè Jacob⁴⁾. Malgré cela ses premiers termes fournissent parfois une bonne approximation pour y .

⁴⁾ *Mosè Jacob*: Sullo Sviluppo di una funzione di ripartizione in serie di polinomi di Hermite, Giorn. Ist. ital. Attuari, anno II, p. 100, 1931.

Mais on peut aborder le problème par un côté différent. On peut chercher à développer en série d'Hermite non pas la fonction y , ce qui est impossible, mais une fonction de deuxième classe $f(x)$ qui se réduit à y_i pour $x = x_i$. Une fonction de cette sorte, et il en existe une infinité, sera représentée par une courbe passant par les points (x_i, y_i) . La série d'Hermite correspondante fournira donc les valeurs y_i pour $x = x_i$ et en l'arrêtant à l'un des premiers termes on aura des expressions approchées des y_i . Bien entendu, les coefficients de ces séries, et c'est là un inconvénient, s'expriment à l'aide des moments m de $f(x)$, qui sont des intégrales dont le calcul n'est pas toujours facile. On choisira alors parmi les fonctions f celle dont les moments se calculent d'une manière particulièrement simple et, si possible, à partir des moments μ de y .

Supposons que la variable aléatoire x ne prenne qu'un nombre fini de valeurs x_i ($i = 1, 2, \dots, \nu$). On pourrait alors, avec M. Kameda⁵⁾ envisager une courbe crénelée composée d'une part de traits horizontaux de même longueur passant par les points (x_i, y_i) et d'autre part de segments de l'axe des x .

Supposons en particulier, et c'est le cas qui se présente d'habitude en statistique, que les points x_i soient équidistants. Sans nuire à la généralité, on peut supposer que les intervalles (x_i, x_{i+1}) soient tous égaux à l'unité.

Dans ce cas la courbe de M. Kameda peut être remplacée par la courbe en escalier montant ou descendant que M. Pearson a appelée histogramme, composée de ν traits horizontaux de même longueur égale à l'unité dont les points (x_i, y_i) occupent les milieux.

La fonction $f(x)$ est donc constante et égale à y_i le long de l'intervalle $(x_i - \frac{1}{2}, x_i + \frac{1}{2})$. De plus f est nulle en dehors de l'intervalle $(x_1 - \frac{1}{2}, x_\nu + \frac{1}{2})$. Quant aux valeurs de f aux points de discontinuité $x_i \pm \frac{1}{2}$, elles seront par définition les moyennes arithmétiques des valeurs $f(x_i \pm \frac{1}{2} + 0)$ et $f(x_i \pm \frac{1}{2} - 0)$. La fonction f ainsi définie n'a donc que des discontinuités de première espèce en nombre fini et est développable par conséquent en série d'Hermite. Les coefficients c_n de cette série peuvent être obtenus, comme l'a montré M. Kameda, à partir de la fonction caractéristique, mais il est plus simple, au point de vue où je me place, de procéder d'une manière plus directe.

⁵⁾ *T. Kameda*: Theory of generating functions and its application to the theory of Probability, Journ. of the Fac. Sc. Imper. Univ. of Tokyo, vol. 1. p. 1 (1925).

Il est évident d'abord que les moments de $f(x)$ sont des fonctions linéaires des moments μ .

En effet

$$m_k = \int_{x_1 - \frac{1}{2}}^{x_1 + \frac{1}{2}} x^k f(x) dx = \sum_i \int_{x_i - \frac{1}{2}}^{x_i + \frac{1}{2}} x^k y_i dx$$

$$= \frac{1}{k+1} \sum_i \left\{ (x_i + \frac{1}{2})^{k+1} - (x_i - \frac{1}{2})^{k+1} \right\} y_i = \mu_k + \frac{k(k-1)}{3! 2^2} \mu_{k-2} + \dots$$

qu'on peut écrire symboliquement

$$(6) \quad m_k = \frac{1}{k+1} \left\{ (\mu + \frac{1}{2})^{k+1} - (\mu - \frac{1}{2})^{k+1} \right\}.$$

En particulier

$$m_0 = \mu_0 = 1; \quad m_1 = \mu_1 = 0; \quad m_2 = \mu_2 + \frac{1}{12}.$$

Introduisons l'écart réduit $t = \frac{x}{a}$, où $a = \sqrt{2\mu_2}$. La probabilité y devient une fonction de t ; les points $t_i = \frac{x_i}{a}$ sont encore équidistants, mais l'intervalle constant $t_{i+1} - t_i$ devient égal à $\frac{1}{a}$. Notre graphique a donc subi une contraction (si $a > 1$) dans le sens horizontal. La fonction f regardée comme fonction de t ne définit plus une loi de probabilité, puisque $\int_{-\infty}^{\infty} f dt = \frac{1}{a} \neq 1$ (si $a \neq 1$). Mais multiplions f par a , le graphique subira dans le sens vertical une dilatation égale à la contraction horizontale et la fonction transformée af définira une loi des probabilités élémentaires. C'est cette fonction af que nous chercherons à développer en série d'Hermite.

Posons

$$(7) \quad af = \frac{e^{-t^2}}{\sqrt{\pi}} \sum_{n=0}^{\infty} c_n H_n(t)$$

d'où

$$(8) \quad y_i = \frac{e^{-t^2}}{a\sqrt{\pi}} \sum_{n=0}^{\infty} c_n H_n(t).$$

Les coefficients c_n sont encore des fonctions linéaires des μ , puisque

$$(9) \quad c_n = \frac{1}{2^n \cdot n!} H_n \left(\frac{\mu}{a} \right).$$

Mais pour avoir l'expression finale de c_n à l'aide des moments μ , il vaut mieux partir de la formule

$$c_n = \frac{a}{2^n \cdot n!} \int_{-\infty}^{\infty} f H_n(t) dt$$

dont le second membre se calcule à l'aide de la relation

$$\int_{\alpha}^{\beta} H_n(t) dt = -\frac{1}{2n+2} \{ H_{n+1}(\beta) - H_{n+1}(\alpha) \}.$$

On trouve ainsi l'expression symbolique suivante

$$(10) \quad c_n = -\frac{a}{2^n \cdot n! (2n+2)} \left\{ H_{n+1} \left(\frac{\mu + \frac{1}{2}}{a} \right) - H_{n+1} \left(\frac{\mu - \frac{1}{2}}{a} \right) \right\}.$$

En particulier

$$c_0 = 1; \quad c_1 = 0; \quad c_2 = \frac{\mu_2}{2a^2} - \frac{1}{4} + \frac{1}{24a^2} = \frac{1}{48a^2}.$$

Mais on peut donner à l'expression (10) une forme différente qui nous sera utile.

Je ferai remarquer d'abord qu'en vertu de la formule de Taylor

$$H_{n+1}(t+h) - H_{n+1}(t-h) = 2 \left\{ h H'_{n+1}(t) + \frac{h^3}{3!} H^{(3)}_{n+1}(t) + \dots \right\}$$

et que d'autre part

$$H_{n+1}^{(2k+1)}(t) = -2^{2k+1} (n+1)n \dots (n+1-2k) H_{n-2k}(t).$$

On en tire en posant $t = \frac{\mu}{a}$ et $h = \frac{1}{2a}$

$$(11) \quad c_n = \frac{1}{2^n \cdot n!} \left\{ H_n \left(\frac{\mu}{a} \right) + \frac{n!}{3!(n-2)!} \frac{1}{a^2} H_{n-2} \left(\frac{\mu}{a} \right) + \dots \right\}.$$

On pourra donc, à partir des moments μ calculer les coefficients c_n et former ainsi la série d'Hermite (8) qui convergera quel que soit t .

Au lieu de l'histogramme on pourrait aussi, avec M. Kameda, envisager la ligne polygonale dont les sommets sont les points x_i, y_i et les extrémités les points $x_i - 1, x_i + 1$ de l'axe des x . Les coefficients de la série d'Hermite correspondante se calculent d'une manière analogue.

2. Quelques problèmes connexes

On peut à partir de la série (8) obtenir une expression différente de la probabilité y_i .

Je commencerai par faire remarquer que

$$y_i = \int_{x_i - \frac{1}{2}}^{x_i + \frac{1}{2}} f(x) dx = a \int_{t_i - h}^{t_i + h} f(x) dt$$

où $h = \frac{1}{2a}$.

D'autre part

$$\int_{t_i - h}^{t_i + h} e^{-t^2} H_n(t) dt = \left[e^{-t^2} H_{n-1}(t) \right]_{t_i - h}^{t_i + h}.$$

On trouve donc en intégrant (8) entre les limites $t_i - h$ et $t_i + h$

$$(12) \quad y_i = \frac{1}{\sqrt{\pi}} \int_{t_i - h}^{t_i + h} e^{-t^2} dt + \frac{1}{\sqrt{\pi}} \sum_{n=2}^{\infty} c_n \left[e^{-t^2} H_{n-1}(t) \right]_{t_i - h}^{t_i + h}.$$

Si l'on néglige les termes qui suivent l'intégrale, on obtient l'expression approchée de y_i qui, dans le cas des épreuves répétées, a été proposée récemment par M. de Montessus de Ballore.

Soit maintenant P_k^l la probabilité pour que t soit compris au sens large entre deux limites données t_k et t_l . Comme $P_k^l = \sum_{i=k}^{i=l} y_i$, on trouve, en vertu de (12)

$$(13) \quad P_k^l = \frac{1}{\sqrt{\pi}} \int_{t_k-h}^{t_l+h} e^{-t^2} dt + \frac{1}{\sqrt{\pi}} \sum_{n=2}^{\infty} c_n \left[e^{-t^2} H_{n-1}(t) \right]_{t_k-h}^{t_l+h}.$$

En arrêtant le développement au premier terme, on aura, dans le cas des épreuves répétées, la formule approchée de J. Eggenberger.

Des rapprochements intéressants peuvent être faits entre la formule (13) et des expressions analogues indiquées dans quelques travaux récents. Je m'en occuperai dans un autre paragraphe de ce travail.

3. Les solutions histogrammiques et les séries de Charlier

L'expression (11) des coefficients c_n que nous venons de donner va nous permettre de faire un rapprochement entre les séries (8) construites à partir d'un histogramme et les séries A de Charlier. Rappelons que les séries de Charlier s'obtiennent des séries d'Hermite représentant les fonctions de deuxième classe en remplaçant les moments m par les moments μ . Elles s'obtiennent par exemple des séries (8) en posant

$$c_n = \frac{1}{2^n \cdot n!} H_n \left(\frac{\mu}{a} \right).$$

Il suffit donc pour avoir la série de Charlier de supprimer dans l'expression (11) du coefficient c_n tous les termes qui suivent le premier. Or pour a suffisamment grand le c_n de Charlier diffère peu du c_n de la série (8).

Il n'est donc pas étonnant que les séries de Charlier fournissent parfois des formules approchées assez commodes, lorsqu'on les arrête à l'un de leurs premiers termes, bien qu'elles divergent pour tout $t = t_i$. D'autre part on voit qu'il suffit d'ajouter aux coefficients de Charlier des termes supplémentaires donnés par la formule (11) pour obtenir une série convergente quel que soit t . Ce sont ces termes supplémentaires qui assurent la convergence de nos séries (8), de même que la convergence des produits canoniques de Weierstrass est assurée par certains facteurs exponentiels.

4. Les épreuves répétées

Envisageons une série d'épreuves vérifiant les conditions de Jacques Bernoulli. Soient s le nombre de ces épreuves, p la probabilité cons-

tante de l'événement attendu, q celle de l'événement contraire, x l'écart. Les valeurs possibles de cet écart sont

$$-sp, -sp + 1, \dots, sq.$$

La probabilité y_i pour que l'écart ait la valeur x_i est donnée par la formule

$$y_i = \frac{s!}{(sp + x_i)!(sq - x_i)!} p^{sp+x_i} q^{sq-x_i}.$$

On connaît l'expression des moments μ .

En particulier

$$\mu_0 = 1; \quad \mu_1 = 0; \quad \mu_2 = \sigma^2 = spq; \quad \mu_3 = (q-p)\sigma^2.$$

On sait que μ_n est un polynome en σ^2 de degré $\frac{n}{2}$, si n est pair, et de degré $\frac{n-1}{2}$, si n est impair et $p \neq q$.

Dans la formule (8) posons $a = \sigma\sqrt{2}$. Il vient

$$(14) \quad y = \frac{e^{-t^2}}{\sigma\sqrt{2\pi}} \left\{ 1 + \sum_{n=2}^{\infty} c_n H_n(t) \right\}$$

où

$$c_2 = \frac{1}{48\sigma^2}, \quad c_3 = \frac{p-q}{12\sigma\sqrt{2}}, \quad \dots$$

Demandons-nous si l'on peut, à partir de cette série, obtenir les expressions asymptotiques de y , lorsqu'on suppose que le nombre des épreuves s et par conséquent la dispersion σ augmentent indéfiniment.

Cherchons à développer le second membre de (14) suivant les puissances croissantes de $\frac{1}{\sigma}$. Dans le cas qui nous occupe les termes de ce développement peuvent être calculés facilement de proche en proche. En effet les coefficients c_n sont des polynomes en $\frac{1}{\sigma}$ et l'on peut montrer à l'aide des propriétés que j'ai établies dans un travail antérieur⁶⁾,

⁶⁾ Commentarii Math. Helv., vol. 1, p. 15.

qu'une puissance quelconque $\frac{1}{\sigma^k}$ ne figure que dans un nombre fini des premiers coefficients c_n et qu'à partir d'un certain rang tous les termes des polynomes c_n contiennent $\frac{1}{\sigma}$ à une puissance supérieure. On pourra donc écrire

$$(15) \quad y = \frac{e^{-t^2}}{\sigma \sqrt{2\pi}} \left\{ 1 + \frac{\gamma_1}{\sigma} + \dots + \frac{\gamma_k}{\sigma^k} + R\left(\frac{1}{\sigma}, t\right) \right\}$$

les coefficients $\gamma_1, \gamma_2, \dots$ étant des fonctions linéaires de certains polynomes d'Hermite et $R\left(\frac{1}{\sigma}, t\right)$ une série convergente quels que soient σ et t . De plus tous les termes de R contiennent $\frac{1}{\sigma}$ à une puissance supérieure à la $k^{\text{ième}}$.

Posons

$$(16) \quad S_k = \frac{e^{-t^2}}{\sigma \sqrt{2\pi}} \left\{ 1 + \frac{\gamma_1}{\sigma} + \dots + \frac{\gamma_{k-1}}{\sigma^{k-1}} \right\}.$$

Il vient

$$\sigma^{k+1} (y - S_k) = \frac{e^{-t^2}}{\sqrt{2\pi}} \left\{ \gamma_k + \sigma^k R\left(\frac{1}{\sigma}, t\right) \right\}.$$

Supposons maintenant que le nombre des épreuves s et par conséquent la dispersion σ augmentent indéfiniment. Comme tous les termes de $\sigma^k R$ s'annulent pour $\sigma = \infty$, il pourrait sembler à première vue que

$$\lim_{\sigma=\infty} \sigma^k R\left(\frac{1}{\sigma}, t\right) = 0$$

et que par conséquent

$$(17) \quad \sigma^{k+1} (y - S_k) \sim \frac{e^{-t^2}}{\sqrt{2\pi}} \gamma_k.$$

Il pourrait sembler en d'autres termes que la série

$$\frac{e^{-t^2}}{\sigma \sqrt{2\pi}} \left(1 + \sum_{n=1}^{\infty} \frac{\gamma_n}{\sigma^n} \right)$$

représente asymptotiquement la fonction y .

Mais cette conclusion serait inexacte.

Nous connaissons en effet une série qui fournit les expressions asymptotiques de y lorsqu'on suppose qu'en tendant vers l'infini la variable s parcourt une suite de valeurs entières convenablement choisie⁷⁾. C'est la série que j'ai appelée la série de Laplace et dont j'ai donné la structure dans mon mémoire cité des Commentarii.

Elle s'écrit

$$\frac{1}{\sigma \sqrt{2\pi}} \left(1 + \frac{P_1(t)}{\sigma} + \frac{P_2(t)}{\sigma^2} + \dots \right),$$

où $P_1(t)$, $P_2(t)$, ... sont des fonctions linéaires des polynomes d'Hermite.

Or si $P_1(t) = \gamma_1(t)$, l'égalité $P_k(t) = \gamma_k(t)$ n'a pas lieu pour $k > 1$. Par exemple pour $k = 2$

$$\gamma_2(t) = P_2(t) + \frac{1}{48} H_2(t).$$

On voit donc que la limite de $\sigma^k R\left(\frac{1}{\sigma}, t\right)$ pour $\sigma \rightarrow \infty$ n'est pas égale à la somme des limites de ses termes. Cela tient à ce que cette série n'est uniformément convergente par rapport à $\frac{1}{\sigma}$ dans aucun intervalle fermé aboutissant à $\frac{1}{\sigma} = 0$.

Avons-nous du moins le droit de conclure de (14) que y ou $y \sigma \sqrt{2}$ tend vers la loi de Gauss? Je me borne à poser cette question.

5. Etude du cas général

Jusqu'à présent nous avons supposé que les points x_i étaient équidistants et en nombre fini. Nous chercherons maintenant à nous affranchir de ces restrictions.

Soit x_i un ensemble de points sans point d'accumulation à distance finie, rangés par ordre de grandeurs croissantes, l'indice i parcourant

⁷⁾ Il suffit que cette suite soit choisie de telle sorte qu'à partir d'un certain rang le point envisagé t ne soit un point de discontinuité pour aucune des fonctions histogrammiques correspondantes. Ce choix peut être fait d'une infinité de manières différentes.

dans le cas général la suite des nombres entiers négatifs et positifs de $-\infty$ à $+\infty$. Si les intervalles (x_i, x_{i+1}) séparant deux points voisins ont une borne inférieure > 0 , on pourra construire une courbe crénelée de M. Kameda telle que $f(x_i)$ soit égale à y_i quel que soit i . Mais si cette borne est nulle, un même trait pourra recouvrir plusieurs points x_i et la fonction $f(x_i)$ pourra se réduire à une somme d'un certain nombre des y_i . Cela tient à ce que les traits horizontaux sont par hypothèse de longueur égale. Mais cette restriction n'a rien d'essentiel au point de vue où nous nous plaçons dans cette étude.

Marquons sur l'axe des x une seconde série de points ξ_i , le point ξ_i étant situé à l'intérieur de l'intervalle (x_{i-1}, x_i) au sens étroit. Si l'ensemble x_i a un premier élément x_1 , on marquera encore un point $\xi_1 < x_1$ et s'il a un dernier élément x_v , on marquera un point $\xi_{v+1} > x_v$. Chacun des intervalles $\delta_i = (\xi_i, \xi_{i+1})$ ne contiendra qu'un seul point de l'ensemble: le point x_i .

Posons

$$f(x) = \frac{y_i}{\delta_i} \text{ dans } \xi_i < x < \xi_{i+1}$$

$$\text{et } f(\xi_i) = \frac{1}{2} \left(\frac{y_{i-1}}{\delta_{i-1}} + \frac{y_i}{\delta_i} \right).$$

La fonction $f(x)$, qui définit une loi des probabilités élémentaires, sera représentée par un histogramme irrégulier.

Supposons que $f(x)$ soit développable en série d'Hermite.

Posons

$$f(x) = \frac{e^{-x^2}}{\sqrt{\pi}} \sum_{n=0}^{\infty} c_n H_n(x)$$

où

$$c_n = \frac{1}{2^n \cdot n!} \int_{-\infty}^{\infty} f(x) H_n(x) dx.$$

On a

$$\begin{aligned} \int_{-\infty}^{\infty} f(x) H_n(x) dx &= \sum_i \int_{\xi_i}^{\xi_{i+1}} f(x) H_n(x) dx = \sum_i \frac{y_i}{\delta_i} \int_{\xi_i}^{\xi_{i+1}} H_n(x) dx \\ &= \sum_i \frac{y_i}{(2n+2)\delta_i} \left\{ H_{n+1}(x_i + \delta_i'') - H_{n+1}(x_i - \delta_i') \right\} \\ &= \sum_i y_i \left\{ H_n(x_i) - n(\delta_i'' - \delta_i') H_{n-1}(x_i) + \dots \right\} \end{aligned}$$

en posant

$$\delta_i' = x_i - \xi_i, \quad \delta_i'' = \xi_{i+1} - x_i,$$

et par conséquent

$$c_n = \frac{1}{2^n \cdot n!} H_n(u) + R_n,$$

R_n pouvant être calculé à partir des y_i et des δ_i', δ_i'' . Ici encore la présence de R_n assure la convergence de la série d'Hermite et lorsque R_n est suffisamment petit, les premiers termes de la série de Charlier fournissent une expression approchée de y_i .

6. Lois des probabilités totales (fonctions de répartition)

Dans le cas des masses de première classe envisagé dans le n° précédent, la fonction de répartition, que je désignerai par $Y(x)$, est représentée par une courbe en escalier. Lorsque x n'est pas un point de l'ensemble x_i , on a $Y(x) = \sum y_i$, la somme étant étendue à toutes les masses réparties dans l'intervalle $(-\infty, x)$. Nous conviendrons que les x_i sont des points de discontinuité réguliers de première espèce de Y . Dans le cas des masses de deuxième classe la fonction de répartition, que je désignerai par $F(x)$, est donnée par l'intégrale

$$F(x) = \int_{-\infty}^x f(x) dx.$$

Supposons que $f(x)$ soit une solution histogrammique développable en une série d'Hermite. Supposons en d'autres termes qu'on ait

$$(18) \quad f(x) = \frac{e^{-x^2}}{\sqrt{\pi}} \sum_{n=0}^{\infty} c_n H_n(x)$$

où

$$c_n = \frac{1}{2^n \cdot n!} H_n(m).$$

J'écrirai $c_n(m)$ pour mettre en évidence qu'il s'agit de masses de deuxième classe.

On aura

$$(19) \quad F(x) = \frac{1}{\sqrt{\pi}} \int_{-\infty}^x e^{-x^2} dx + \frac{e^{-x^2}}{\sqrt{\pi}} \sum_{n=1}^{\infty} c_{n+1}(m) H_n(x).$$

Envisageons maintenant la série de Charlier

$$\frac{e^{-x^2}}{\sqrt{\pi}} \sum_{n=0}^{\infty} c_n(\mu) H_n(x)$$

où

$$c_n(\mu) = \frac{1}{2^n \cdot n!} H_n(\mu).$$

Nous savons qu'elle diverge pour $x = x_i$. Mais intégrons-la terme à terme de $-\infty$ à x . Chose curieuse, si l'intégrale de Stieltjes

$$\int_{-\infty}^{\infty} e^{\frac{x^2}{2}} dY(x)$$

qui se réduit ici⁸⁾ à la somme $\sum_i e^{\frac{x_i^2}{2}} y_i$, converge, on obtient, en vertu d'un théorème de M. Cramér⁹⁾, une série convergente qui représente la fonction de répartition $Y(x)$. On aura donc

$$Y(x) = \frac{1}{\sqrt{\pi}} \int_{-\infty}^x e^{-x^2} dx + \frac{e^{-x^2}}{\sqrt{\pi}} \sum_{n=1}^{\infty} c_{n+1}(\mu) H_n(x).$$

Je montrerai que cet important théorème, que M. Cramér a donné sans démonstration, peut être établi à partir des développements (18) et (19).

Supposons d'abord que le nombre des valeurs x_i soit fini. Nous savons alors que les séries (18) et (19) convergent.

⁸⁾ C'est à cause des notations un peu différentes que j'écris $e^{\frac{x^2}{2}}$ au lieu de $e^{\frac{x^2}{4}}$.

⁹⁾ *H. Cramér*: On the composition of elementar errors, Skand. Aktuarietidskrift, 1928, p. 65.

Posons

$$(20) \quad F(x) - Y(x) = \eta(x).$$

Envisageons un intervalle (ξ_i, ξ_{i+1}) .

Dans cet intervalle $F(x)$ est représentée par un segment de droite qui va du point P_i de coordonnées $\xi_i, y_1 + \dots + y_{i-1}$ au point P_{i+1} de coordonnées $\xi_{i+1}, y_1 + \dots + y_i$. Quant à la partie correspondante de la courbe $Y(x)$, elle relie aussi le point P_i au point P_{i+1} , mais elle se compose de deux traits horizontaux avec un point de discontinuité en x_i . Dans l'intervalle (ξ_i, ξ_{i+1}) la fonction $\eta(x)$ est donc représentée par deux segments de droites parallèles à $P_i P_{i+1}$, le premier partant de ξ_i , le second aboutissant à ξ_{i+1} , ayant respectivement pour projections δ_i' et δ_i'' .

La fonction $\eta(x)$ étant nulle en dehors d'un intervalle fini, elle est développable en une série d'Hermite et l'on pourra écrire

$$(21) \quad \eta(x) = \frac{e^{-x^2}}{\sqrt{\pi}} \sum_{n=0}^{\infty} \gamma_n H_n(x).$$

Calculons les coefficients γ_n .

On a

$$\gamma_n = \frac{1}{2^n \cdot n!} \int_{-\infty}^{\infty} \eta(x) H_n(x) dx = \frac{1}{2^n \cdot n!} \sum_i \int_{x_i - \delta_i'}^{x_i + \delta_i''} \eta(x) H_n(x) dx.$$

Or

$$\int_{x_i - \delta_i'}^{x_i + \delta_i''} \eta(x) H_n(x) dx = \mathcal{F}_1 + \mathcal{F}_2,$$

en posant

$$\mathcal{F}_1 = \int_{x_i - \delta_i'}^{x_i} \eta(x) H_n(x) dx, \quad \mathcal{F}_2 = \int_{x_i}^{x_i + \delta_i''} \eta(x) H_n(x) dx.$$

Mais dans \mathcal{F}_1 la fonction $\eta(x) = \frac{y_i}{\delta_i} \left\{ x - (x_i - \delta_i') \right\}$

et dans \mathcal{F}_2 , $\eta(x) = \frac{y_i}{\delta_i} \left\{ x - (x_i + \delta_i'') \right\}$.

Il vient alors, après quelques transformations très simples,

$$\gamma_n = \frac{1}{2^{n+1} \cdot (n+1)!} \sum_i y_i \left\{ \frac{1}{\delta_i} \int_{x_i - \delta_i'}^{x_i + \delta_i''} H_{n+1}(x) dx - H_{n+1}(x_i) \right\}.$$

D'autre part le coefficient $c_{n+1}(m)$ dans (19) est

$$\begin{aligned} c_{n+1}(m) &= \frac{1}{2^{n+1} \cdot (n+1)!} \int_{-\infty}^{\infty} f(x) H_{n+1}(x) dx \\ &= \frac{1}{2^{n+1} \cdot (n+1)!} \sum_i \frac{y_i}{\delta_i} \int_{x_i - \delta_i'}^{x_i + \delta_i''} H_{n+1}(x) dx, \end{aligned}$$

d'où

$$\begin{aligned} c_{n+1}(m) - \gamma_n &= \frac{1}{2^{n+1} \cdot (n+1)!} \sum_i y_i H_{n+1}(x_i) \\ &= \frac{1}{2^{n+1} \cdot (n+1)!} H_{n+1}(u) = c_{n+1}(u). \end{aligned}$$

On a donc bien, en vertu de (20)

$$Y(x) = \frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} e^{-x^2} dx + \frac{e^{-x^2}}{\sqrt{\pi}} \sum_{n=1}^{\infty} c_{n+1}(u) H_n(x)$$

et le théorème de M. Cramér est établi dans le cas où le nombre des x_i est fini.

Supposons maintenant que l'ensemble x_i soit infini. Je me bornerai au cas envisagé dans le n° précédent.

Supposons avec M. Cramér que la somme

$$(22) \quad \sum_i e^{\frac{x_i^2}{2}} y_i$$

converge.

Pour établir le théorème de M. Cramér il suffit de montrer que les séries (18), (19) et (21) convergent.

Or, en vertu d'un critère indiqué par M. Cramér¹⁰⁾, les séries (18) et (19) convergent sûrement, si l'intégrale

$$(23) \quad \int_{-\infty}^{\infty} e^{-\frac{x^2}{2}} f(x) dx$$

converge.

Je montrerai qu'on peut toujours choisir les ξ_i de telle façon que cette condition soit vérifiée.

En effet pour $\xi_i > 1$

$$\int_{\xi_i}^{\xi_{i+1}} e^{-\frac{x^2}{2}} f(x) dx < \frac{y_i}{\delta_i} e^{-\frac{(x_i + \delta_i'')^2}{2}} \cdot \delta_i < y_i e^{-\frac{x_i^2}{2}} \cdot e^{\frac{3}{2}},$$

si $\delta_i'' < \frac{1}{x_i}$.

Par conséquent la convergence de (22) entraîne celle de $\int_0^{\infty} e^{-\frac{x^2}{2}} f(x) dx$.

Conclusion analogue pour $\int_{-\infty}^0 e^{-\frac{x^2}{2}} f(x) dx$, si l'on suppose $\delta_i' < \frac{1}{|x_i|}$ pour $\xi_{i+1} < -1$.

Quant à la série (21), sa convergence s'établit très simplement à partir des critères de M. Rotach¹¹⁾. Pour montrer que ces critères s'appliquent à la fonction $\eta(x)$, il suffit de faire remarquer que dans tout intervalle (ξ_i, ξ_{i+1}) ne comprenant pas l'origine à son intérieur, $\eta(x)$ vérifie l'inégalité $\left| \frac{\eta(x)}{x} \right| < \frac{y_i}{\delta_i}$.

7. Valeur pratique des solutions histogrammiques

Quelle est la valeur pratique des séries convergentes qu'on obtient à partir de l'histogramme?

Il est difficile, dans l'état actuel de la science, de donner une réponse précise à cette question.

¹⁰⁾ H. Cramér. On some classes of series used in mathematical statistics, C. R. du 6^e congrès des mathématiciens scandinaves, 1926, p. 407.

¹¹⁾ Loc. cit. p. 6.

Tant qu'on n'aura pas trouvé un critère commode permettant d'évaluer les limites de l'erreur qu'on commet en arrêtant ces séries à l'un de leurs premiers termes, cette valeur restera discutable. La même remarque s'applique du reste à la plupart des expressions approchées dont on se sert en statistique.

On se contente d'habitude, pour montrer l'excellence de certaines formules approchées, de donner quelques exemples numériques où ces formules fournissent une bonne approximation. Mais les exemples particuliers peuvent donner une idée tout-à-fait fautive du degré d'approximation de ces formules. Aucune loi ne se dégage des exemples particuliers. C'est ainsi que dans le cas des séries d'Hermite les valeurs approchées peuvent osciller très irrégulièrement autour de la valeur vraie. L'oscillation embrasse plusieurs termes de la série et rien ne nous permet jusqu'à présent de prévoir l'allure de ce mouvement oscillatoire.

On voit donc combien il serait important de trouver une expression approchée commode de l'erreur commise en arrêtant une série d'Hermite à l'un de ses termes. Serait-il possible de résoudre ce problème à l'aide des méthodes de M. M. Plancherel et Rotach? En tout cas la solution ne me paraît pas bien lointaine.

Je ne sais si les remarques qui précèdent sont entièrement nouvelles, mais j'ai pensé qu'elles pouvaient présenter quelque intérêt, étant donné l'importance des problèmes auxquels elles touchent.

(Reçu le 14 juillet 1931)

