

SUR L'INDÉPENDANCE DES VARIABLES ALÉATOIRES

Autor(en): **Kantorowicz, I.**

Objektyp: **Article**

Zeitschrift: **L'Enseignement Mathématique**

Band (Jahr): **29 (1930)**

Heft 1: **L'ENSEIGNEMENT MATHÉMATIQUE.**

PDF erstellt am: **20.09.2024**

Persistenter Link: <https://doi.org/10.5169/seals-23247>

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.

SUR L'INDÉPENDANCE DES VARIABLES ALÉATOIRES

PAR

I. KANTOROWICZ (Varsovie).

Introduction. — M. RAJCHMAN dans son Cours, professé à l'Université de Varsovie, envisageait le Calcul des probabilités comme un chapitre de la théorie générale des fonctions. Toute variable aléatoire y était considérée comme la fonction d'une ou de plusieurs variables indépendantes, dites variables de base ou variables apparentes¹. Dans cet ordre d'idées *l'espérance mathématique* d'une variable aléatoire se présente tout à fait naturellement sous la forme d'une intégrale définie (qui se réduit à une moyenne dans le cas courant des variables discontinues), tandis que la probabilité elle-même n'apparaît que beaucoup plus tard, comme une notion relativement compliquée. La définition classique de l'indépendance de deux ou de plusieurs variables aléatoires, intimement liée au théorème de la multiplication des probabilités, se prête mal à cette manière d'exposition et M. Rajchman a posé à ses auditeurs la question suivante: *définir l'indépendance de deux ou de plusieurs variables aléatoires en n'employant que la notion de l'espérance mathématique.*

Dans ce qui suit je donne une solution de ce problème, en établissant des conditions nécessaires et suffisantes de l'indépendance des variables aléatoires.

¹ Par exemple le gain d'un joueur en jeu de roulette (variable aléatoire) était considéré comme la fonction du numéro qui sort (variable de base).

1. ÉNONCÉS DES THÉORÈMES.

Je prouve notamment les théorèmes suivants:

I — Soit x, y, z, \dots des variables aléatoires prenant respectivement m, n, p, \dots valeurs distinctes. Dans ces conditions pour l'indépendance de ce système de variables, il faut et il suffit que l'on ait:

$$\|x^s y^t z^u \dots\| = \|x^s\| \cdot \|y^t\| \cdot \|z^u\| \dots \quad (1)$$

pour

$$s = 0, 1 \dots m - 1 ; \quad t = 0, 1, \dots n - 1 ; \quad u = 0, 1, \dots p - 1 ;$$

$\|x\|$ désigne ici l'espérance mathématique de la variable aléatoire x .

Dans le cas de deux variables aléatoires seulement l'équation (1) pour $s = 0$ et t quelconque (ainsi que évidemment pour $t = 0$ et s quelconque) est une tautologie. Dans le cas de trois variables les équations du système (1), correspondant respectivement à $s = 0, t = 0, u = 0$ expriment l'indépendance des variables x, y, z deux à deux.

Notre énoncé met ainsi bien en relief la différence de la notion de l'indépendance des variables aléatoires deux à deux et leur indépendance tout court.

II. — Si l'on supprime la restriction relative au nombre de valeurs, que prennent x, y, z, \dots on retombe sur le théorème connu¹: pour qu'ait lieu l'indépendance d'un système des variables aléatoires x, y, z, \dots il faut et il suffit que l'on ait:

$$\|x^m y^n z^p \dots\| = \|x^m\| \cdot \|y^n\| \cdot \|z^p\| \dots$$

pour toutes les valeurs entières positives ou nulles de m, n, p, \dots

Remarque. — Les lois de probabilité attachées aux variables x, y, z, \dots doivent être ici supposées absolument continues, sans quoi la définition classique de l'indépendance par l'application du théorème sur la multiplication des probabilités deviendrait inutilisable.

¹ Ce théorème peut-être n'a pas été énoncé explicitement, mais il est un corollaire à peu près immédiat du théorème classique des moments.

III. — Si l'on suppose les variables aléatoires x, y, z, \dots non négatives, pour leur indépendance il faut et il suffit, que l'égalité suivante:

$$\|x^m y^n z^p \dots\| = \|x^m\| \cdot \|y^n\| \cdot \|z^p\| \dots$$

soit remplie pour m, n, p, \dots aussi grands que l'on veut.

Donc dans ce cas m, n, p, \dots n'ont pas à prendre toutes les valeurs entières, mais ils peuvent faire partie de n'importe quelle suite de valeurs réelles (entières ou non) indéfiniment croissantes.

2. — DÉMONSTRATIONS.

Nous nous bornerons à démontrer les théorèmes dans le cas de deux variables aléatoires. La démonstration dans le cas d'un nombre plus grand se fait absolument de la même manière et ne présente aucune difficulté nouvelle.

Démonstration du théorème I. — Appelons

p_k la probabilité de l'égalité $x = a_k$ ($k = 1, 2, \dots, m$)

q_l » » » $y = b_l$ ($l = 1, 2, \dots, n$)

$r(k, l)$ la probabilité pour que l'on ait en même temps $x = a_k$ et $y = b_l$.

Supposons que les a_k sont tous différents, ainsi que les b_l . Posons

$$\varepsilon(k, l) = r(k, l) - p_k q_l.$$

L'indépendance de x et de y s'exprime par le système d'égalités

$$\varepsilon(k, l) = 0. \quad (k = 1, 2, \dots, m; \quad l = 1, 2, \dots, n) \quad (2)$$

Pour toute valeur de s et de t on a évidemment les identités suivantes:

$$\|x^s\| = \sum_{k=1}^m p_k a_k^s, \quad \|y^t\| = \sum_{l=1}^n q_l b_l^t. \quad (3)$$

En particulier:

$$\|x^0\| = \sum_{k=1}^m p_k = 1, \quad \|y^0\| = \sum_{l=1}^n q_l = 1. \quad (4)$$

On a de même :

$$\begin{aligned} \|x^s y^t\| &= \sum_{l=1}^n \sum_{k=1}^m r(k, l) a_k^s b_l^t = \\ &= \sum_{l=1}^n \sum_{k=1}^m [\varepsilon(k, l) + p_k q_l] a_k^s b_l^t = \\ &= \sum_{l=1}^n \sum_{k=1}^m \varepsilon(k, l) a_k^s b_l^t + \|x^s\| \cdot \|y^t\| . \end{aligned} \quad (5)$$

$$\|x^0 y^t\| = \|y^t\| , \quad \|x^s y^0\| = \|x^s\| . \quad (6)$$

Si l'on rapproche les conditions (1) des identités (5) et (6), on trouve que le problème se ramène à étudier l'équivalence du système (2) au système d'équations linéaires suivant :

$$\sum_{l=1}^n \sum_{k=1}^m \varepsilon(k, l) a_k^s b_l^t = 0 . \quad (7)$$

$$(s = 0, 1, 2, \dots, m - 1 ; \quad t = 0, 1, 2, \dots, n - 1)$$

Il est bien évident que (2) entraîne (7).

Pour prouver la réciproque, posons :

$$z(l, s) = \sum_{k=1}^m \varepsilon(k, l) a_k^s . \quad (8)$$

Le système (7) prend alors la forme :

$$\sum_{l=1}^n b_l^t z(l, s) = 0 . \quad (t = 0, 1, \dots, n - 1) \quad (9)$$

Pour un s fixe, le système (9) peut être envisagé comme un système d'équations linéaires à n inconnues :

$$z(1, s) , \quad z(2, s) \dots \dots \quad z(n - 1, s) , \quad z(n, s) .$$

Le déterminant de ce système est évidemment le déterminant de Vandermonde égal à

$$\prod_{h \neq i} (b_h - b_i) .$$

Les b_l étant tous différents, ce déterminant est manifestement différent de zéro. Par conséquent on déduit de (9)

$$z(l, s) = 0 \quad (l = 1, 2, \dots, n; \quad s = 0, 1, \dots, m-1) \quad (10)$$

En rapprochant ceci de la définition (8), on est amené au système d'équations en $\varepsilon(k, l)$:

$$\sum_{k=1}^m \varepsilon(k, l) a_k^s = 0 \quad (s = 0, 1, \dots, m-1) \quad (11)$$

Le déterminant de ce système est encore un déterminant de Vandermonde représenté cette fois par

$$\prod_{h \neq i} (a_h - a_i)$$

et, par conséquent, différent de zéro. On a donc pour toute valeur de k et de l :

$$\varepsilon(k, l) = 0 \quad \text{c. q. f. d.}$$

Démonstration du théorème II. — Soient respectivement

$$\int_{-\infty}^a f(x) dx; \quad \int_{-\infty}^b \varphi(y) dy \quad \text{et} \quad \int_{-\infty}^b \int_{-\infty}^a \psi(x, y) dx dy$$

les probabilités des inégalités:

$$x \leq a, \quad y \leq b$$

et du système de deux inégalités simultanées:

$$x \leq a \quad \text{et} \quad y \leq b.$$

Il s'agit d'établir l'équivalence de l'égalité

$$\psi(x, y) = f(x) \varphi(y) \quad (12)$$

et du système:

$$\int_{-\infty}^{+\infty} x^m f(x) dx \int_{-\infty}^{+\infty} y^n \varphi(y) dy = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} x^m y^n \psi(x, y) dx dy \quad (13)$$

$$(m = 0, 1, 2, 3, \dots; \quad n = 0, 1, 2, 3, \dots)$$

Pour $m, n = 0$ l'équation (13) est une tautologie.

Le passage de (12) à (13) est immédiat. Il ne s'agit donc que du passage inverse. Pour l'effectuer, posons:

$$\varepsilon(x, y) = \psi(x, y) - f(x) \varphi(y) . \quad (14)$$

L'équation (13) revêt alors la forme suivante

$$\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} x^m y^n \varepsilon(x, y) dx dy = \int_{-\infty}^{+\infty} y^n dy \int_{-\infty}^{+\infty} x^m \varepsilon(x, y) dx = 0 .$$

En appliquant le théorème des moments¹ à la fonction suivante de y :

$$\int_{-\infty}^{+\infty} x^m \varepsilon(x, y) dx .$$

on trouve, quel que soit y :

$$\int_{-\infty}^{+\infty} x^m \varepsilon(x, y) dx = 0 , \quad (m = 0, 1, 2, 3, \dots)$$

Ce qui nous donne, par une nouvelle application du théorème des moments:

$$\varepsilon(x, y) \equiv 0 \quad \text{c.q.f.d.}$$

Démonstration du théorème III. — Lemme I (fondamental).
Soit $A(x)$ une fonction continue croissante (au sens strict) et non négative dans l'intervalle (a, b) et soit $f(x)$ une fonction continue dans cet intervalle. Nous affirmons que l'égalité

$$\int_a^b [A(x)]^n f(x) dx = 0 \quad (15)$$

supposée remplie pour n aussi grand que l'on veut, entraîne comme conséquence:

$$f(x) = 0 \quad \text{pour} \quad a \leq x \leq b . \quad (16)$$

¹ A savoir que toute fonction continue à moments tous nuls est forcément égale à zéro.

Démonstration. Supposons, par impossible, que l'équation (16) n'ait pas lieu pour tous les points de l'intervalle (a, b) . Nous distinguerons deux cas: 1° $f(b) \neq 0$, 2° $f(b) = 0$.

Commençons par montrer l'impossibilité du cas 2°. Dans ce cas il est possible de déterminer les points c et d tels que l'on ait:

$$\begin{aligned} f(x) &= 0 & \text{pour} & & c \leq x \leq b \\ f(x) &\neq 0 & \text{pour} & & d \leq x < c ; \end{aligned}$$

la fonction $f(x)$ garde le signe constant dans (d, c) . Sans restreindre la généralité, nous pouvons supposer que ce signe est positif [dans le cas contraire nous aurions pris $-f(x)$ au lieu de $f(x)$]. Soit α le point où $f(x)$ atteint son maximum relatif à (d, c) . Posons

$$f(\alpha) = 2\varepsilon .$$

Il existe alors un intervalle (g, h) ¹ tel, que l'on ait:

$$f(x) \geq \varepsilon \quad \text{pour} \quad g \leq x \leq h . \quad (17)$$

En rapprochant les formules (15) et (17), en posant

$$B_n(x) = [A(x)]^n f(x) ,$$

on trouve

$$\int_a^d B_n(x) dx + \int_d^g B_n(x) dx + \int_g^h B_n(x) dx + \int_h^b B_n(x) dx = 0 . \quad (18)$$

Puisque les quantités

$$\int_d^g B_n(x) dx \quad \text{et} \quad \int_h^b B_n(x) dx$$

sont positives, on peut écrire:

$$\int_a^d B_n(x) dx + \int_g^h B_n(x) dx < 0 . \quad (19)$$

¹ On a évidemment:

$$a \leq d < g < \alpha < h < c \leq b .$$

On a évidemment

$$\int_g^h B_n(x) dx = \int_g^h [A(x)]^n f(x) dx > \varepsilon \int_g^h [A(x)]^n dx > \varepsilon [A(g)]^n (h - g) . \quad (20)$$

En désignant par M le maximum absolu de $f(x)$ ¹, on a :

$$\left| \int_a^d B_n(x) dx \right| = \left| \int_a^d [A(x)]^n f(x) dx \right| < M \cdot [A(d)]^n (d - a) . \quad (21)$$

En rapprochant les inégalités (19), (20) et (21), on trouve

$$\varepsilon [A(g)]^n (h - g) < M [A(d)]^n (d - a) ,$$

c'est-à-dire

$$\left[\frac{A(g)}{A(d)} \right]^n < \frac{M}{\varepsilon} \cdot \frac{d - a}{h - g} . \quad (22)$$

Puisque, par hypothèse, on a

$$d < g , \quad \text{et, par suite} \quad \frac{A(g)}{A(d)} > 1 ,$$

le premier membre de l'inégalité (22) tend vers l'infini avec n , tandis que le second membre reste constant, ce qui est manifestement absurde. Notre démonstration de l'absurdité de l'hypothèse 2° est ainsi achevée. En ce qui concerne l'absurdité de l'hypothèse 1°, elle est encore plus manifeste. Pour l'établir on n'a qu'à reprendre le raisonnement de tout à l'heure en y introduisant quelques simplifications.

Il suffit de supposer la non-existence du point c et de l'égalité $h = c$.

Lemme II (analogue au théorème des moments). — *Toute fonction continue $f(x)$, remplissant pour n aussi grand que l'on veut, l'équation des moments*

$$\int_0^{+\infty} x^n f(x) dx = 0 \quad (23)$$

est nulle identiquement.

¹ C'est-à-dire le maximum relatif à (a, b) tout entier.

Remarque. — Pour éviter le malentendu possible nous rappelons encore une fois que n parcourt une suite indéfiniment croissante de quantités réelles, positives, d'ailleurs absolument quelconque.

Démonstration. Posons $x = \text{tang } z$.

L'équation (23) devient alors

$$\int_{z=0}^{z=\frac{\pi}{2}} (\text{tang } z)^n \frac{f(\text{tang } z)}{\cos^2 z} dz = 0 .$$

Appliquant le lemme I aux fonctions: $\text{tang } z$ (fonction croissante) et $\frac{f(\text{tang } z)}{\cos^2 z}$ (fonction continue), on trouve

$$f(x) \equiv 0 \quad \text{c. q. f. d.}$$

Pour démontrer le théorème III nous n'avons qu'à répéter la démonstration du théorème II. Les seuls changements à y introduire sont le remplacement de la suite de tous les nombres naturels par une suite croissante réelle arbitrairement donnée et la mise du lemme II à la place du théorème des moments.

Varsovie, 25 mai 1929.