COMPOSITIO MATHEMATICA

PAUL LÉVY

Sur le développement en fraction continue d'un nombre choisi au hasard

Compositio Mathematica, tome 3 (1936), p. 286-303

<http://www.numdam.org/item?id=CM_1936__3__286_0>

Sur le développement en fraction continue d'un nombre choisi au hasard

par

Paul Lévy

Paris

§ 1. *Introduction.*

Dans ce travail, nous représenterons le développement en fraction continue d'un nombre compris entre 0 et 1 par les formules

$$(1)\quad x = 1/(a_1 + 1/(a_2 + \ldots + 1/(a_{n-1} + 1/x_n)\ldots)) = \frac{P_n x_n + P_{n-1}}{Q_n x_n + Q_{n-1}}.$$

Lorsque nous dirons que la variable x (ou une autre variable aléatoire) est choisie au hasard entre 0 et 1, il sera sous-entendu que des intervalles égaux sont également probables. Une propriété sera dite *presque sûre* si sa probabilité est 1, et *presque sûre asymptotiquement* (dans le cas d'une propriété dépendant de n) si sa probabilité tend vers 1 pour n infini. Nous dirons que deux expressions fonctions de x et de n sont *équivalentes modo bernoulliano* (par abréviation m.b.) lorsque, quel que soit ε positif, leur rapport est asymptotiquement presque sûrement compris entre $1 - \varepsilon$ et $1 + \varepsilon$, et qu'elles sont *presque sûrement équivalentes* lorsqu'il est presque sûr que ce rapport tend vers 1; en d'autres termes, il y a dans le premier cas *convergence en mesure*, et dans le second *convergence presque sûre*, de ce rapport vers 1.

Dans un travail récent paru dans ce recueil [1]), M. Khintchine a établi quelques résultats relatifs à l'ordre de grandeur, pour n très grand, d'expressions telles que

$$s_n = a_1^\alpha + a_2^\alpha + \ldots + a_n^\alpha .$$

Il résulte de différentes remarques (p. 363 et p. 381) que ce

[1]) Vol. 1 (1935), 361—382.

savant considère qu'il s'agit là de résultats très cachés qui ne pouvaient être obtenus que grâce aux progrès les plus récents de l'analyse. Cette opinion ne me semble qu'en partie exacte.

Les problèmes du genre de ceux considérés par M. Khintchine peuvent se répartir en trois groupes. Le premier comprend ceux dont la solution ne dépend que de l'ordre de grandeur des a_n, et n'est pas modifiée lorsqu'on multiplie ces nombres par des facteurs positifs bornés inférieurement et supérieurement. Les deux autres groupes comprennent les problèmes dont la résolution implique une évaluation asymptotique précise des a_n, mais dans ceux du second groupe, les grandes valeurs des a_n ne jouent qu'un rôle négligeable, tandis que dans ceux du troisième groupe, elles jouent un rôle prépondérant. Ainsi, l'évaluation asymptotique précise de s_n est un problème du second groupe si $\alpha < 1$ et du troisième groupe si $\alpha \geqq 1$.

Je montrerai d'abord au § 2 que les problèmes du premier groupe se ramènent bien simplement à l'étude de la somme de variables aléatoires indépendantes les unes des autres, c'est-à-dire à l'un des problèmes les plus classiques du calcul des probabilités. Ce fait est implicitement contenu dans le mémoire classique de M. Borel sur les probabilités dénombrables, et je le croyais connu depuis longtemps; le mémoire de M. Khintchine me fait penser qu'il est utile de l'indiquer explicitement. Les § 3 et 4 sont consacrés au rappel de résultats connus et à quelques compléments. Compte tenu de la remarque du § 2, il résulte *ipso facto* d'un de ces résultats que la série $\sum \frac{1}{s_n}$ de M. Khintchine, pour $\alpha = 1$, est presque sûrement divergente. C'est le résultat que M. Khintchine semble considérer comme le plus caché; quoique, pour une partie de la démonstration, j'utilise la méthode même qu'il a employée, ce résultat me semble au contraire le plus simple, parce qu'il se rattache au premier groupe.

Les problèmes du second groupe impliquent l'utilisation d'une formule énoncée par Gauss et démontrée seulement en 1928 par M. Kuzmin: [2])

$$\lim_{n \to \infty} \mathfrak{P}\{a_n \geqq p\} = \frac{1}{\log 2} \log \left(1 + \frac{1}{p}\right),$$

$\mathfrak{P}\{E\}$ désignant la probabilité d'un événement E; cette formule est exacte sous la seule condition que x dépende d'une loi de probabilité absolument continue.

---

[2]) Atti congr. intern. Bologna 1928, 6, 83.

J'ai complété ce résultat en 1929 en me plaçant au point de vue de la loi forte des grands nombres [3]). J'ai montré que, sauf pour des valeurs de x constituant un ensemble de mesure nulle, la fréquence de chacune des valeurs possibles dans la suite des a_n tend vers la probabilité résultant de la formule de Gauss, et établi un résultat analogue pour la suite des valeurs du rapport $\lambda_n = \frac{Q_{n-1}}{Q_n}$.

Il est clair que la première partie de cet énoncé donne la solution des problèmes du second groupe [4]). Je montrerai d'une manière précise au § 5 qu'il est équivalent au théorème de M. Khintchine relatif à ces problèmes.

Pour arriver au problème du troisième groupe, une extension de mon théorème de 1929 est nécessaire. Elle n'introduit d'ailleurs aucun principe nouveau, et est même plus simple que ce théorème lui-même si l'on se contente de rechercher des valeurs principales b.m. de s_n pour $\alpha = 1$. Cette somme n'a d'ailleurs pas de valeur principale presque sûre; on ne peut la borner supérieurement que par des fonctions croissant plus rapidement que celle trouvée pour sa valeur principale b.m.

Les principes que j'utilise peuvent d'ailleurs servir à la solution de problèmes autres que ceux considérés par M. Khintchine. Citons à titre d'exemple ceux où interviennent des expressions de la forme

$$u_n = f(a_n, a_{n+1}, \ldots, a_{n+p})$$

Pour n très grand et p fixe, il résulte de mon mémoire de 1929 que $a_n, a_{n+1}, \ldots, a_{n+p}$ peuvent être considérés comme les premiers quotients incomplets du développement d'un nombre x_n choisi entre 0 et 1 d'après la loi de probabilité indiquée par Gauss, chaque intervalle dx_n ayant la probabilité $\frac{dx_n}{(1+x)\log 2}$. Sauf pour des valeurs de x constituant un ensemble de mesure nulle, chaque système de valeurs possibles de $a_{n+1}, a_{n+2}, \ldots, a_{n+p}$ est

---

[3]) Bull. Soc. Math. de France 57 (1929), 190 et suiv. En dehors du résultat indiqué dans le texte, j'ai indiqué dans ce mémoire des formules de récurrence qui semblent n'être pas connues par M. M. Kuzmin et Khintchine, et dont l'emploi conduit à cette conclusion que la convergence de la probabilité étudiée vers sa limite est plus rapide que les formules de M. Kuzmin ne le laissent penser.

[4]) Nous supposons ici qu'on ait préalablement établi qu'il s'agit d'un problème du second groupe; le problème de la distinction entre le deuxième groupe et le troisième est, du moins dans les cas simples considérés par M. Khintchine, un problème du premier groupe.

réalisé, quand n augmente indéfiniment, avec une fréquence tendant vers la probabilité théorique ainsi calculée. Ces remarques définissent la corrélation qui existe asymptotiquement entre a_n et a_{n+p} si p a une valeur finie, tandis que ces deux variables peuvent être considérées comme indépendantes si p est grand, quels que soient dans ce cas les ordres de grandeur relatifs de n et p, n pouvant même rester fini.

Il en résulte que l'on obtient sans peine, pour les sommes de termes tels que u_n, des théorèmes asymptotiques analogues à ceux établis par M. Khintchine dans le cas où chaque terme ne dépend que d'un seul quotient [5]).

Comme conclusion, il me semble que l'on peut dire que, d'une manière très générale, la difficulté spéciale à la théorie des fractions continues est résolue pour les questions de cette nature par la formule de Gauss et par mon théorème de 1929, ces problèmes étant ramenés par ce théorème à des problèmes (plus ou moins difficiles) relatifs à une suite de variables aléatoires indépendantes les unes des autres.

---

[5]) On peut appliquer aussi mes méthodes à l'étude d'expressions où interviennent, au lieu des a_n, soit les quotients complets x_n, soit les rapports $X_n = \dfrac{Q_{n+1}}{Q_n}$. Bien que x_n détermine tous les x_{n+p}, sa donnée à ε près est, si p est grand, sans influence appréciable sur la loi dont dépend x_{n+p}; d'une manière précise, la probabilité conditionnelle $\mathfrak{P}'\{x_{n+p} > \xi\}$, calculée quand on connaît x_n à ε près, diffère de sa valeur limite $\left[\log\left(1+\dfrac{1}{\xi}\right)\right] / \log 2$ d'une quantité inférieure à $a\varepsilon^{\alpha}k^{p}$ (a, α, et $k < 1$ étant des constantes absolues). D'après mon mémoire de 1929, les X_n dépendent asymptotiquement de la même loi que les x_n, et la corrélation entre X_n et X_{n+p} est la même qu'entre x_{n+p} et x_n, c'est-à-dire que, si l'on ne s'occupe que de calculs numériques approchés, elle est très faible si p est grand. Pourvu que la fonction $f(x)$ soit continue, tout ce que nous disons dans le texte des expressions $f(a_n)$ s'applique alors sans difficulté à $f(x_n)$ et à $f(X_n)$; il faut seulement remplacer la loi discontinue dont dépendent les a_n par la loi continue dont dépendent (asymptotiquement) les x_n et les X_n.

Ainsi, d'après cette loi, $\log x_n$ et $\log X_n$ ont asymptotiquement pour valeur probable $m = \dfrac{\pi^2}{12 \log 2}$. Il résulte alors des considérations qui précèdent qu'il y a presque sûrement convergence en moyenne arithmétique vers m de la suite des $\log x_n$ comme de celle des $\log X_n$. En d'autres termes il est presque sûr que $\sqrt[n]{x_1 x_2 \dots x_n}$ et $\sqrt[n]{Q_n}$ tendent vers e^m; il en est de même de $\sqrt[n]{P_n}$.

La présente Note, postérieure au reste de ce travail, a été écrite aprés réception d'une lettre de M. Khintchine me disant avoir démontré que $\sqrt[n]{Q_n}$ tend presque sûrement vers une constante absolue (dont il n'indiquait pas la valeur).

(Reçu le 25 juillet 1935.)

§ 2. *Le principe fondamental relatif aux problèmes du premier groupe.*

Nous désignerons par $\mathfrak{P}'$ des probabilités évaluées lorsqu'on connaît $a_1, a_2, \ldots, a_{n-1}$, c'est-à-dire lorsque x est choisi au hasard dans l'intervalle compris entre $\frac{P_n}{Q_n}$ et $\frac{P_n + P_{n-1}}{Q_n + Q_{n-1}}$ que l'on obtient en faisant varier x_n de 1 à ∞. On a

$$\mathfrak{P}'\{x_n > \xi\} = \frac{1+\lambda_n}{\xi + \lambda_n} \qquad \left(\xi > 1,\ \lambda_n = \frac{Q_{n-1}}{Q_n} < 1\right), \tag{2}$$

et l'on vérifie immédiatement que cette probabilité est toujours comprise entre $\frac{1}{\xi}$ et $\frac{2}{\xi}$ (car son produit par ξ varie avec ξ d'une manière monotone de 1 à $1 + \lambda_n < 2$). Ces limites étant indépendantes de $a_1, a_2, \ldots, a_{n-1}$, s'appliquent aussi à la probabilité *a priori*, x étant alors choisi au hasard entre 0 et 1, ainsi que dans le cas où l'on a des renseignements de nature absolument quelconque (certains ou probables) sur $a_1, a_2, \ldots, a_{n-1}$.

Désignons alors par $\eta_1, \eta_2, \ldots, \eta_n, \ldots$ des variables aléatoires indépendantes les unes des autres et choisies au hasard entre 0 et 1; posons

$$y_n = \frac{1}{\eta_n} = \frac{\xi_n + \lambda_n}{1 + \lambda_n}, \tag{3}$$

a_n étant la partie entière de ξ_n et $\lambda_n = \frac{Q_{n-1}}{Q_n}$ étant par (1) une fonction bien définie de $a_1, a_2, \ldots, a_{n-1}$ ($\lambda_1 = 0$). Les a_n peuvent ainsi se calculer successivement quand on connaît les η_n, et comme

$$\mathfrak{P}'\{\xi_n > \xi\} = \mathfrak{P}'\left\{\eta_n < \frac{1+\lambda_n}{\xi + \lambda_n}\right\} = \frac{1+\lambda_n}{\xi + \lambda_n},$$

on trouve pour la suite des a_n exactement la même loi de probabilité que s'ils étaient déduits du développement en fraction continue d'un nombre x choisi au hasard entre 0 et 1. L'égalité établie plus haut, appliquée pour $\xi = \xi_n$, donne

$$\frac{a_n}{2} \leqq \frac{\xi_n}{2} < y_n \leqq \xi_n < a_n + 1,$$

de sorte que: *pour toutes les propriétés ne dépendant que de l'ordre de grandeur des a_n, on ne change rien en remplaçant la suite des a_n par la suite des y_n, $\eta_n = \frac{1}{y_n}$ étant une variable aléatoire indépendante de $\eta_1, \eta_2, \ldots, \eta_{n-1}$ et choisie au hasard entre* 0 *et* 1.

La probabilité de ces propriétés, et en particulier celle de la convergence de la série de M. Khintchine $\Sigma \frac{1}{s_n}$ (pour $\alpha = 1$) peuvent donc s'étudier sans que l'on ait rien d'autre à savoir sur les fractions continues. On est ramené au problème classique de l'addition de variables aléatoires indépendantes les unes des autres, dans le cas simple où elles dépendent de la même loi, et je pourrais presque me contenter de renvoyer le lecteur aux traités classiques et à mes travaux antérieurs sur ce sujet. Mais il me paraît préférable de rappeler explicitement les résultats les plus importants.

§ 3. *Etude des sommes du type* $S_n = f(y_1) + f(y_2) + \ldots + f(y_n)$.

Nous poserons

$$z = f(y), \quad m = \mathfrak{E}\{z\}, \quad \sigma^2 = \sigma^2\{z\} = \mathfrak{E}\{(z-m)^2\},$$

le symbole $\mathfrak{E}$ représentant la valeur probable. Nous supposerons la fonction $f(y)$ non décroissante, ce qui ne restreint en rien la nature des variables aléatoires $z_n = f(y_n)$ mais permet de donner une expression simple à des hypothèses sur les probabilités des grandes valeurs de ces variables.

1⁰. Dans le cas classique où σ est fini, il résulte de l'inégalité de Tchebycheff que, si m n'est pas nul, S_n est *modo bernoulliano* équivalent à nm, et de la loi du logarithme itéré que S_n est presque sûrement équivalent à nm. Nous allons montrer que ce résultat subsiste dans des conditions bien plus larges.

2⁰. Supposons d'abord $f(y) = y^\alpha$, avec $\frac{1}{2} < \alpha < 1$; m est fini et σ infini. C'est sur l'étude de S_n dans ce cas que je me suis appuyé depuis longtemps pour montrer l'existence des lois stables [6]). La variable réduite

$$\overline{S}_n = \frac{S_n - nm}{n^\alpha} \tag{4}$$

dépend d'une loi de probabilité tendant vers une limite, qui est la loi stable dont il s'agissait d'établir l'existence; il est donc très peu probable que $|\,\overline{S}_n\,|$ soit très grand, et cela suffit pour établir l'équivalence m.b. de S_n et de nm.

Leur équivalence presque sûre s'établit de la manière suivante:

---

[6]) Paul Lévy, Calcul des probabilités [1925], 258—263. J'ai d'autre part indiqué, dans un mémoire qui paraîtra prochainement dans le Journal de Mathématiques, une méthode élémentaire reposant sur l'étude des probabilités des grandes valeurs.

β étant un nombre choisi entre 1 et $\frac{1}{\alpha}$, on peut supposer que l'on a toujours $y_n < n^\beta$; sauf dans des cas dont la probalité est nulle, cette hypothèse n'est, en effet, en défaut qu'un nombre fini de fois, ce qui est sans influence sur la valeur principale de S_n. Compte tenu de cette hypothèse, la valeur probable m_n de y_n^α reste, pour n très grand, très peu différente de m, et celle de $(y_n^\alpha - m_n)^2$ prend une valeur finie, mais augmentant avec n et équivalente, à un facteur constant près, à $n^\gamma (\gamma = 2\alpha\beta - \beta < 1)$. On peut alors appliquer l'inégalité de Tchebycheff et de M. Kolmogoroff à l'étude de l'oscillation maxima de $S_n - nm$ quand n varie de $2^p + 1$ à 2^{p+1}; l'ordre de grandeur à prévoir est celui $2^{p\gamma'}$ $\left[\gamma' = \frac{1+\gamma}{2} < 1\right]$, c'est-à-dire que cette oscillation est inférieure à $2^{p\delta}$ (δ étant choisi entre γ' et 1), sauf dans des cas dont la probabilité totale est le terme d'une série convergente et qui, par suite, peuvent être négligés. Finalement, en ajoutant les oscillations de $S_n - nm$ dans les différents intervalles considérés, on trouve que, sauf dans des cas dont la probabilité est nulle, cette différence est à partir d'un certain moment inférieure à $n^{\alpha+\varepsilon}$, ε étant arbitrairement petit; l'équivalence presque sûre de S_n et nm en résulte, c.q.f.d [7]).

3⁰. L'extension au cas où, pour y infini, $f(y)$ est infini mais de la forme $O(y^\alpha)$ avec $\alpha < 1$, tandis que $f(1) > -\infty$, est immédiate. Choisissons un nombre α' entre α et 1, puis Y assez grand pour que, ε et ε' étant arbitrairement petits, on ait

$$\int_Y^\infty \frac{f(y)dy}{y^2} < \varepsilon,$$
$$0 < f(y) < \varepsilon' y^{\alpha'} \qquad (\text{pour } y < Y).$$

Désignons par z_n' une variable égale à $f(y_n)$ si $y_n \leqq Y$ et à 0 si $y_n > Y$. On a

$$S_n = \sum_1^n z_\nu' + \theta_n \varepsilon' \sum_1^n y_\nu^{\alpha'} \qquad (0 \leqq \theta_n < 1).$$

L'application des résultats des 1⁰ et 2⁰ du présent § aux sommes

---

[7]) Le principe du raisonnement précédent est tiré de mon mémoire de Studia Mathematica 3 (1931), 119—155. Si j'ai dû le reproduire ici, c'est que je n'avais énoncé le th. IX, qu'il faudrait appliquer ici, que dans le cas où $\alpha > 1$ (avec les notations du présent mémoire, qui ne sont pas celles du mémoire cité; pour passer de l'un à l'autre il faut changer α en $\frac{1}{\alpha}$).

$\sum_1^n z'_\nu$ et $\sum_1^n y_\nu^{\alpha'}$, compte tenu de ce que $\mathfrak{E}\{z'_\nu\}$ est compris entre m et $m-\varepsilon$, montre que ces résultats s'étendent au cas qui nous occupe. On traite de même le cas où, pour $y-1$ très petit, $f(y)$ est infini, par valeurs négatives, et d'ordre au plus égal à α. On a ainsi le résultat suivant:

Si la fonction $f(y)$ est non décroissante; si, α étant < 1, $(y-1)^\alpha$ $f(y)$ est borné pour $y-1$ très petit; si, de même $y^{-\alpha} f(y)$ est borné pour y infini; si enfin la valeur probable m de $f(y)$ n'est pas nulle: alors il est presque sûr que S_n a pour n infini la valeur principale nm.

4⁰. Soit maintenant $f(y) = y^\alpha$ avec $\alpha > 1$ (nous reviendrons ensuite sur le cas où $\alpha = 1$); m est infini. La variable réduite à considérer est ici $\frac{S_n}{n^\alpha}$. C'est cette variable qui, d'après les résultats exposés dans mon livre, dépend d'une loi de probabilité tendant vers une limite déterminée; ses valeurs possibles vont de zéro à l'infini, les valeurs très petites et très grandes étant très peu probables. Donc: *S_n est modo bernoulliano de l'ordre de grandeur de n^α, mais n'a pas de valeur principale.*

D'autre part, si la fonction $\lambda(n)$ croît régulièrement, il est presque sûr que la condition

$$S_n = o[\lambda^\alpha(n)]$$

est vérifiée si la série $\sum \frac{1}{\lambda(n)}$ est convergente, et que, dans le cas contraire, même la condition moins restrictive

$$S_n = O[\lambda^\alpha(n)]$$

n'est pas vérifiée. (Cf. loc. cit [7])).

Ces résultats s'étendent au cas où $f(y)$ est seulement de l'ordre de grandeur de y^α (toujours avec $\alpha > 1$). Ils s'étendent aussi *ipso facto* aux sommes $\sum_1^n a_\nu^\alpha$ relatives aux fractions continues.

On remarque que l'on obtient pour l'ordre de grandeur de S_n des résultats d'autant moins précis que la fonction $f(y)$ est plus rapidement croissante. Cela était à prévoir car, dans le cas d'une croissance rapide, une grande valeur de y fortuitement réalisée a une grande influence. L'étude du cas où $f(y) = y$ va confirmer cette remarque.

5⁰. *Cas où* $f(y) = y$, $S_n = y_1 + y_2 + \dots + y_n$.

Commençons par une remarque simple. Soit Y le plus grand des nombres $y_1, y_2, \dots, y_n$. Il résulte de la formule évidente

$$\mathfrak{P}\{Y < kn\} = \left(1 - \frac{1}{kn}\right)^n \to e^{-k} \qquad (n \to \infty),$$

que Y est de l'ordre de la grandeur de n, les valeurs beaucoup plus petites et beaucoup plus grandes étant très peu probables. La probabilité de l'existence de p valeurs supérieures à n dans la suite $y_1, y_2, \ldots, y_n$ est d'autre part

$$C_n^p \left(\frac{1}{n}\right)^p \left(1 - \frac{1}{n}\right)^{n-p} \backsim \frac{1}{ep!} \qquad (n \to \infty)$$

(c'est la loi de Poisson), de sorte que la probabilité que le nombre de ces valeurs soit égal ou supérieur à p a pour n infini une limite très petite si p est grand. Ces valeurs peuvent donc être négligées pour l'étude de la valeur principale m.b. de la somme $S_n = y_1 + y_2 + \ldots + y_n$ qui, comme nous allons le voir, est de l'ordre de grandeur de $n \log n$.

Désignons maintenant par $y'_1 = \frac{1}{\eta'_1}, \ldots, y'_n = \frac{1}{\eta'_n}$, les nombres $y_1, y_2, \ldots, y_n$, rangés par ordre de grandeurs décroissantes; on voit aisément que

$$\mathfrak{P}\{\eta < \eta'_p < \eta + d\eta\} = \frac{n!}{(p-1)!\,(n-p)!}\,\eta^{p-1}(1-\eta)^{n-p}\,d\eta,$$

et par suite

$$\mathfrak{E}\{\eta'_p\} = \frac{p}{n+1}, \quad \mathfrak{E}\{\eta'^2_p\} = \frac{p(p+1)}{(n+1)(n+2)},$$

$$\sigma^2\{\eta'_p\} = \frac{p(n-p+1)}{(n+1)^2(n+2)} < \frac{1}{p}\left(\mathfrak{E}\{\eta'_p\}\right)^2.$$

Si p est grand, il résulte alors de l'inégalité de Tchebycheff que $\mathfrak{E}\{\eta'_p\}$ donne d'une manière assez précise l'ordre de grandeur de η'_p. L'erreur due aux premiers termes qui sont évalués avec moins de précision étant négligeable, cela nous permet de prévoir que

$$S_n = y_1 + y_2 + \ldots + y_n = y'_1 + y'_2 + \ldots + y'_n$$

est de l'ordre de grandeur de

$$(n+1)\left(1 + \frac{1}{2} + \ldots + \frac{1}{n}\right) \backsim n \log n.$$

Si les y_ν sont indépendants, les y'_ν ne le sont pas; il y a entre eux une corrélation positive. Pour limiter l'erreur commise dans l'évaluation précédente, nous pouvons partir de la formule

$$\mathfrak{E}\left\{\left|\eta'_p - \frac{p}{n+1}\right|\right\} \leqq \sigma\{\eta'_p\} < \frac{1}{\sqrt{p}}\,\frac{p}{n+1}.$$

L'erreur relative est donc très petite, si p est grand; l'erreur relative sur y'_p est la même que sur η'_p. Appliquant alors à la somme de termes positifs

$$y'_{p+1} + y'_{p+2} + \cdots + y'_n$$

(p étant grand, mais indépendant de n) la remarque que l'erreur relative sur la somme est au plus égale à la plus grande des erreurs relatives, et en se souvenant que l'erreur sur les premiers termes est sans influence sur l'ensemble, nous voyons que l'évaluation précédente de S_n comporte, si n est assez grand, une erreur relative arbitrairement petite. Donc: *bernoulliano modo, la valeur principale de S_n est $n \log n$.*

Mais cette fois il n'y a pas de valeur principale presque sûre. Ainsi la probabilité que y_n dépasse à lui seul $2n \log n$ est $\frac{1}{2n \log n}$, et il résulte de la divergence de la série $\Sigma \frac{1}{n \log n}$ que y_n, et par suite S_n, sont presque sûrement une infinité de fois supérieurs à $2n \log n$.

Pour avoir une borne supérieure presque sûre de S_n, il faut introduire une fonction $\varphi(n)$ telle que la série $\Sigma \frac{1}{\varphi(n)}$ soit convergente. Alors on peut presque sûrement trouver un N tel que $y_n < \varphi(n)$ pour tout $n > N$; il en sera alors de même du plus grand terme de S_n à partir du moment où $\varphi(n)$ dépasse la plus grande des valeurs $y_1, y_2, \ldots, y_N$. Ce terme est donc presque sûrement, pour n assez grand, inférieur à $\varphi(n)$. Pourvu que cette fonction soit d'allure assez régulière, elle dépasse aussi la somme des autres termes de S_n, et la limitation obtenue pour le plus grand terme s'applique à S_n; la démonstration précise résulte d'un raisonnement analogue à ceux développés au 2⁰ ci-dessus, et dans mon mémoire de 1931 (loc. cit. [7])).

6⁰. *Deuxième méthode pour l'étude du cas où $f(y) = y$.*

Nous avons observé que $y'_1 = Y$ est *bernoulliano modo* de l'ordre de grandeur de n. Considérons maintenant la loi de probabilité conditionnelle dont dépend S_n lorsque Y est connu. En écrivant $\mathfrak{E}'$ et σ' au lieu de $\mathfrak{E}$ et σ pour indiquer qu'il s'agit de cette loi, on trouve aisément

$$\mathfrak{E}'\{S_n\} = Y + \frac{(n-1)Y}{Y-1} \log Y, \tag{5}$$

$$\sigma'\{S_n\} = \frac{Y}{Y-1} \sqrt{(n-1)(y-1-\log Y)}. \tag{6}$$

L'inégalité de Tchebycheff

$$\mathfrak{P}\{|S_n - \mathfrak{E}'\{S_n\}| > c\,\sigma'\{S_n\}\} < \frac{1}{c^2},$$

étant vraie quel que soit Y pour la loi conditionnelle considérée est vraie aussi pour la loi de probabilité a priori. En négligeant deux fois de suite des probabilités très petites, nous pouvons donc dire d'abord que Y est de l'ordre de n, donc que $\mathfrak{E}'\{S_n\}$ est équivalent à $n \log n$, ensuite que $S_n - \mathfrak{E}'\{S_n\}$ est $O(\sigma'\{S_n\})$, donc $O(n)$. Donc, *toujours bernoulliano modo, la valeur principale de* S_n *est* $n \log n$, *et l'on peut même préciser que* $S_n - n \log n$ *est de l'ordre de grandeur de* n.

7⁰. *Troisième méthode.*

Négligeant toujours les valeurs des y_ν supérieures à n, nous diviserons l'intervalle conservé $(1, n)$ en intervalles séparés par des nombres en progression géométrique; nous désignerons un quelconque de ces intervalles par $(n^\beta, n^{\beta+\delta\beta})$. Nous supposerons que $\delta\beta \log n$ tende vers zéro avec $\frac{1}{n}$, mais moins rapidement que n'importe quelle puissance de $\frac{1}{n}$; dans ces conditions, on peut négliger l'erreur commise en remplaçant n'importe quelle valeur de l'intervalle $(n^\beta, n^{\beta+\delta\beta})$ par n^β, et on peut prendre comme valeur approchée du nombre des y_ν situés dans cet intervalle, sa valeur probable $n^{1-\beta}\,\delta\beta \log n$; si $\beta < 1$, cette valeur étant très grande, on sait que la valeur probable de l'erreur prise en valeur absolue est relativement très petite. L'erreur relative sur la somme étant au plus égale à la plus grande des erreurs relatives aux différents termes, on a, avec une erreur probable (en valeur absolue) relativement très petite, pour la somme des termes de S_n ne dépassant pas n^α $(\alpha < 1)$ la valeur approchée

$$\sum_{\beta<\alpha} n^\beta n^{1-\beta} \log n \;\delta\beta \backsim n \log n \int_0^\alpha d\beta = \alpha n \log n.$$

Une étude spéciale des valeurs comprises entre $n^{1-\varepsilon}$ et n est alors nécessaire pour que cette troisième méthode se suffise à elle-même. Mais la comparaison des résultats obtenus par cette méthode et par les précédentes nous montre que, si ε est assez petit, on peut avec une erreur relative arbitrairement petite (b.m.) négliger dans S_n les termes supérieurs à $n^{1-\varepsilon}$. Comme c'est une propriété ne dépendant que de l'ordre de grandeur des y_n, elle s'étend ipso facto à l'étude de la somme $a_1 + a_2 + \ldots + a_n$.

8⁰. *Nature de la série* $\sum \frac{1}{S_n}$.

Le résultat ainsi obtenu de trois manières différentes suffit pour établir la divergence presque sûre de la série $\sum \frac{1}{S_n}$; nous renvoyons le lecteur au mémoire de M. Khintchine, dans lequel cette partie du raisonnement semble aussi simple que possible.

Ce résultat sera peut-être rendu plus intuitif par les remarques suivantes: Pour une suite des η_n choisie au hasard, sauf dans des cas dont la probabilité est nulle, si l'on définit u_n par les conditions

$$\begin{aligned} u_n &= 0 \quad \text{si } S_n < 2n \log n \\ u_n &= 1 \quad \text{si } S_n \geqq 2n \log n, \end{aligned}$$

la suite des u_n aura les caractères suivants: la valeur 1 sera rare, mais, sa réalisation étant due à l'existence d'une grande valeur de y_n, pouvant dépasser $n \log n \log \log n$, on aura un grand nombre de valeurs de n consécutives pour lesquelles $u_n = 1$; ces séries partielles seront séparées en général par des séries plus longues encore pour lesquelles $u_n = 0$. Alors la moyenne arithmétique

$$\frac{1}{n}(u_1 + u_2 + \ldots + u_n)$$

oscille presque toujours entre 0 et 1. tandis qu'une moyenne telle que

$$\frac{1}{\log n}\left(u_1 + \frac{u_2}{2} + \ldots + \frac{u_n}{n}\right)$$

tend presque toujours vers zéro. On voit ainsi dans quel sens on peut dire que S_n est presque toujours inférieur à $2n \log n$, et cet énoncé ainsi précisé entraîne la divergence presque sûre de $\sum \frac{1}{S_n}$.

9⁰. *Généralisation des résultats précédents.*

Revenons au cas où la fonction $f(y)$ n'est plus égale à y, et demandons-nous si S_n a une valeur principale *modo bernoulliano.* Si la valeur probable m de $f(y)$ est finie, c'est le problème classique de la loi des grands nombres, et si $f(y)$ est de l'ordre de grandeur de y^α avec $\alpha > 1$, il ne peut pas y avoir de valeur principale en raison du rôle que peut jouer une grande valeur réalisée par hasard. Plaçons-nous dans le cas où m est infini, et où $f(y) = y\lambda(y)$, $\frac{d \log \lambda(y)}{d \log y}$ étant infiniment petit et de signe constant (donc $\lambda(y)$ lentement variable).

Appliquons, par exemple, la seconde méthode ci-dessus. Les formules (5) et (6) deviennent

$$\mathfrak{E}'\{S_n\} = f(Y) + (n-1)\frac{Y}{Y-1}\mu(Y),$$

$$\sigma'^2\{S_n\} = (n-1)\left[\frac{Y}{Y-1}\int_1^Y \lambda^2(y)dy - \frac{Y^2}{(Y-1)^2}\mu^2(Y)\right],$$

en posant

$$\mu(Y) = \int_1^Y \frac{\lambda(y)}{y}\,dy .$$

Comme m est infini, $\mu(Y)$ augmente indéfiniment avec Y.

Le premier terme de $\mathfrak{E}'\{S_n\}$ est de l'ordre de grandeur de $f(n) = n\,\lambda(n)$, et l'incertitude sur ce terme, quand on ne connaît pas Y, est du même ordre de grandeur; d'après l'expression de $\sigma'\{S_n\}$, l'incertitude sur S_n quand on connaît Y est aussi du même ordre. Le second terme de $\mathfrak{E}'\{S_n\}$ a au contraire une valeur principale déterminée $n\ \mu(n)$, et finalement la condition pour que l'ensemble, et par suite S_n, ait une valeur principale, est que le premier terme de $\mathfrak{E}'\{S_n\}$ soit négligeable devant le second, c'est-à-dire que $\lambda(n) = o[\mu(n)]$.

Or, en posant $\log y = \eta$, on a $\lambda(y) = \frac{d\mu(y)}{d\eta}$, et l'hypothèse faite sur $\lambda(y)$ s'écrit $\frac{d^2\mu(y)}{d\eta^2} = o\left[\frac{d\mu(y)}{d\eta}\right]$; elle entraîne $\frac{d\mu(y)}{d\eta} = o\,[\mu(y)]$ c'est-à-dire $\lambda(n) = o\,[\mu(n)]$. Donc, dans les conditions indiquées, S_n a une valeur principale *bernoulliano modo*.

On voit aisément aussi que dans les mêmes conditions il n'y a pas de valeur principale presque sûre; pour m fini, c'est le problème de la *loi forte des grands nombres*. Si l'on prend des fonctions de croissance assez régulières et de plus en plus rapides, $\frac{\log f(y)}{\log y}$ ayant par suite de l'hypothèse de régularité une limite déterminée α, c'est lorsque m devient infini que S_n cesse d'avoir une valeur principale presque sûre, et lorsque α dépasse la valeur 1 que S_n cesse d'avoir une valeur principale *bernoulliano modo*.

§ 4. *Remarques.*

1°. Le cas où $f(y) = y$ peut se traiter par la méthode que nous avons employée pour le cas où $f(y) = y^\alpha$ ($\alpha \neq 1$). Je n'avais pas traité ce cas dans mon livre, où je me proposais d'établir l'existence des lois stables, question qui se trouvait résolue d'une manière élémentaire pour $\alpha = 1$; cela m'a empêché de remarquer une circonstance assez curieuse, que je vais indiquer ici: *pour* $\alpha = 1$, *l'étude de* S_n, *ou plutôt de la variable réduite correspondante*

$$\overline{S}_n = \frac{y_1 + y_2 + \cdots + y_n}{n} - \log n$$

ne conduit pas à la limite à une loi dissymétrique stable mais à une loi ayant la propriété suivante: si S' et S'' sont des variables aléatoires indépendantes l'une de l'autre et qui dépendent de cette loi, il en est de même, non de leur demi-somme, mais de

$$S = \frac{S' + S''}{2} - \log 2 .$$

La vérification, par la méthode des fonctions caractéristiques exposée dans mon livre, est immédiate. La fonction caractéristique d'une des variables aléatoires y est

$$\varphi(t) = \mathfrak{E}\{e^{ity}\} = \int_1^\infty e^{ity} \frac{dy}{y^2},$$

d'où, par le changement de $|t|y$ en y, on déduit

$$\varphi(t) - 1 = \int_{|t|}^\infty [\,|t|(\cos y - 1) + it \sin y]\, \frac{dy}{y^2}$$

$$= -\frac{\pi}{2}|t| - it\log|t| + Ait + o(t),$$

avec

$$A = \lim_{|t|\to 0}\left(\log|t| + \int_{|t|}^\infty \frac{\sin y}{y^2}\, dy\right).$$

La fonction caractéristique $\Phi_n(t)$ liée à $\overline{S}_n$ se déduit de $\varphi(t)$ par la formule

$$\log \Phi_n(t) = \log \mathfrak{E}\{e^{it\overline{S}_n}\} = n \log \varphi\left(\frac{t}{n}\right) - it \log n ,$$

de sorte que, d'après l'expression trouvée pour $\varphi(t)$, on a

$$\log \Phi_n(t) = -\frac{\pi}{2}|t| - it\log|t| + Ait + no\left(\frac{t}{n}\right).$$

Le dernier terme tendant uniformément vers zéro dans tout intervalle fini, on trouve bien à la limite la loi pour laquelle

$$\log \Phi(t) = \log \mathfrak{E}\{e^{itS}\} = -\frac{\pi}{2}|t| - it\log|t| + Ait. \tag{7}$$

La propriété annoncée se vérifie immédiatement d'après cette expression. Cette vérification est d'ailleurs inutile; d'après l'expression même de $\overline{S}_n$, si cette variable dépend d'une loi ayant une limite, cette loi limite a nécessairement la propriété indiquée.

2°. On peut obtenir la loi stable symétrique d'exposant 1, c'est-à-dire la loi de Cauchy, par un procédé analogue à celui employé dans ce qui précède. Il n'y a qu'à considérer la somme alternée

$$S'_n = \sum_1^\infty (-1)^\nu y_\nu = n\overline{S}'_n$$

et la variable réduite $\overline{S}'_n$ dépend à la limite de la loi de Cauchy. S'_n est donc ici, *bernoulliano modo*, de l'ordre de grandeur de n, et n'a pas de valeur principale.

La série $\sum \frac{1}{S'_n}$, malgré les changements de signes, est presque sûrement divergente; le signe est, en effet, le même pour des groupes de termes d'étendue finie (en général ni très petite ni très grande) sur l'échelle logarithmique.

La nature alternée de la somme S'_n fait que l'extension du résultat précédent aux sommes $\sum_1^n (-1)^\nu a_\nu$ liées aux fractions continues nécessite quelques précautions; il n'y a pas de difficulté essentielle, l'ordre de grandeur des sommes étudiées étant déterminé par un petit nombre de valeurs très grandes des a_n, presque indépendantes les unes des autres parce qu'elles se trouvent réalisées pour des quotients éloignés les uns des autres.

§ 5. *Les problèmes du second groupe.*

Le théorème de M. Khintchine est le suivant: *Si $f(y)$ est une fonction toujours inférieure en valeur absolue à cy^α* ($\alpha < 1$), *il est presque sûr que*

$$\lim_{n\to\infty} \frac{1}{n}[f(a_1) + \dots + f(a_n)] = \sum_1^\infty \frac{f(p)}{\log 2} \log\frac{(p+1)^2}{p(p+2)}.$$

Mon théorème de 1929 résulte de ce théorème; il suffit de supposer $f(y)$ égal à zéro pour $y < p$ et à un pour $y \geqq p$. Inversement, supposons mon théorème établi et posons

$$\frac{1}{n}[f(a_1) + \dots + f(a_n)] = S'_n + S''_n,$$

S'_n provenant des termes $f(a_\nu)$ pour lesquels $a_\nu < 2p$, et S''_n de ceux pour lesquels $a_\nu \geqq 2p$. Mon théorème donne immédiatement la limite presque sûre de S'_n; d'autre part, S''_n ne peut, d'après le principe du § 2, qu'être majoré si l'on remplace $f(a_\nu)$ par $c(2y_\nu)^\alpha$, et les y_ν ainsi considérés sont $\geqq p$. D'après le 3° du § 3, on a presque sûrement, à partir d'un certain moment

$$S''_n < C\sum_p^\infty h^\alpha \log\frac{(h+1)^2}{h(h+2)} \backsim \frac{C}{(1-\alpha)p^{1-\alpha}} \quad \left(C > \frac{2^\alpha c}{\log 2}\right).$$

Il suffit alors de prendre p assez grand pour obtenir le théorème de M. Khintchine.

§ 6. *Les problèmes du troisième groupe.*

Ces problèmes sont liés à une extension de mon théorème de 1929. D'après ce théorème, *sauf pour des nombres x constituant un ensemble de mesure nulle,* la fréquence des valeurs inférieures à λ $(0<\lambda<1)$ dans la suite des λ_n tend vers $G(\lambda)=\frac{\log(1+\lambda)}{\log 2}$ ce qui donne immédiatement la limite de $\frac{1}{n}\sum_1^n f(\lambda_\nu)$, si $f(\lambda)$ est une fonction continue de λ. On a en particulier

$$(8)\quad \lim_{n\to\infty}\frac{1}{n}\left[(1+\lambda_1)+(1+\lambda_2)+\ldots+(1+\lambda_n)\right]=$$
$$=\int_0^1(1+\lambda)\,dG(1+\lambda)=\frac{1}{\log 2}.$$

Or, choisir x, c'est choisir successivement les a_ν; a_ν étant la partie entière de x, la probabilité de chaque valeur possible résulte de la formule (2), et pour l'étude des valeurs vérifiant la condition

$$(9)\qquad f(n)<a_\nu<c_nf(n)\qquad (c_n>1,\ f(n)\to\infty),$$

on peut négliger λ_n au dénominateur. La probabilité α_ν de cette condition est donc équivalente à

$$\left(1-\frac{1}{c_n}\right)\frac{1+\lambda_\nu}{f(n)},$$

et pour la probabilité totale, ν variant de 1 à n, on a *presque sûrement*

$$A_n=\alpha_1+\alpha_2+\ldots+\alpha_n\backsim\left(1-\frac{1}{c_n}\right)\frac{n}{f(n)\log 2}=N'.$$

Or, N' est une valeur approchée du nombre N des valeurs de ν au plus égales à n pour lesquelles l'égalité (9) est vérifiée. *Sous la seule condition que N' augmente indéfiniment avec n, N et N' sont équivalents m.b. et l'écart probable $\mathfrak{E}\{|N-N'|\}$, inférieur a l'écart quadratique moyen, est relativement très petit;* cela résulte de l'expression connue de cet écart, dont on sait qu'elle s'applique à l'étude d'une somme de variables aléatoires enchaînées, si, pour la loi de probabilité conditionnelle dont dépend chaque terme quand on connaît les précédents, sa valeur probable est nulle.

On peut aussi montrer que, sous la même condition que N'

augmente indéfiniment, il y a équivalence presque sûre entre N et N'. Mais c'est un fait moins élémentaire et dont nous n'aurons pas besoin. L'énoncé indiqué d'abord suffit, sinon pour tous les problèmes du troisième groupe, du moins pour retrouver le résultat de M. Khintchine relatif à la somme

$$s_n = a_1 + a_2 + \ldots + a_n .$$

La condition que N' augmente indéfiniment est, en effet, vérifiée si l'on prend $f(n) = n^\beta$ $(\beta < 1)$ et $c_n = n^{\delta\beta}$, $\delta\beta$ vérifiant les conditions indiquées § 3, 7°. Il n'y a alors rien à changer, si ce n'est l'introduction du facteur $\frac{1}{\log 2}$, au raisonnement fait à cet endroit, et l'on voit que la partie de s_n provenant des termes inférieurs à n^α $(\alpha < 1)$, a pour valeur principale m.b.

$$\frac{\alpha n \log n}{\log 2} .$$

La remarque finale du § 3, 7° nous apprenant qu'on peut négliger les autres termes si $1 - \alpha$ est assez petit, nous obtenons finalement la formule de M. Khintchine

$$\lim_{n \to \infty} \mathfrak{P} \left\{ \left| \frac{(a_1 + a_2 + \ldots + a_n) \log 2}{n \log n} - 1 \right| < \varepsilon \right\} = 1 .$$

Naturellement, le rôle joué par les grandes valeurs des a_ν empêche, de même que pour la somme $y_1 + y_2 + \ldots + y_n$ étudiée au § 3, qu'on puisse trouver pour $a_1 + a_2 + \ldots + a_n$ une valeur principale presque sûre. Le résultat indiqué à la fin du § 3, 5°, au sujet de la recherche d'une borne supérieure presque sûre de la première de ces sommes s'applique d'ailleurs à la seconde, car les grandes valeurs de cette somme proviennent des grandes valeurs des a_n, et ces valeurs sont assez espacées pour pouvoir être considérées comme indépendantes les unes des autres. Par suite, si $\varphi(n)$ désigne une fonction croissante d'allure assez régulière, la probabilité qu'il existe un N tel que pour tout $n > N$ on ait

$$a_1 + a_2 + \ldots + a_n < \varphi(n)$$

est 0 ou 1 suivant que la série $\sum \frac{1}{\varphi(n)}$ est divergente ou convergente; sauf la restriction relative à la régularité de $\varphi(n)$ l'énoncé est le même que s'il s'agissait de l'inégalité $a_n < \varphi(n)$ étudiée par M. Borel. Si l'on supprime cette restriction, le résultat énoncé ne reste pas exact, mais la probabilité étudiée reste égale à 0 ou 1, toute valeur intermédiaire étant exclue; cela se déduit aisément du principe du § 2 et de résultats connus sur les sommes de variables aléatoires indépendantes.

(Reçu le 13 juillet 1935.)

---

[8]) Ayant communiqué le présent travail à M. Khintchine, j'ai le plaisir, grâce à sa réponse, de pouvoir mentionner son accord sur les points suivants: l'étude de sa série $\sum \frac{1}{(a_1 + \dots + a_n)}$ est en effet plus simple qu'il ne l'avait pensé d'abord les autres problèmes traités dans son travail cité deviennent faciles lorsque l'on connaît, d'une part la formule de Gauss, d'autre part quelques théorèmes récents de calcul des probabilités.

Au sujet de ma démonstration de cette formule, je rappelle qu'elle est indépendante de celle de M. Kuzmin, communiquée oralement au congrès de Bologne. mais publiée seulement après la mienne. D'ailleurs nos méthodes sont très différentes; la mienne présente l'avantage de conduire à une évaluation plus précise de la rapidité de la convergence vers la limite indiquée par Gauss.

(Reçu le 10 octobre 1935.)