JOURNAL
DE
MATHÉMATIQUES
PURES ET APPLIQUÉES
FONDÉ EN 1836 ET PUBLIÉ JUSQU'EN 1874
PAR JOSEPH LIOUVILLE

PAUL LÉVY

La loi forte grands nombres pour les variables aléatoires enchaînées

Journal de mathématiques pures et appliquées 9^e série, tome 15 (1936), p. 11-24.
<http://www.numdam.org/item?id=JMPA_1936_9_15__11_0>

La loi forte des grands nombres pour les variables aléatoires enchaînées;

PAR PAUL LÉVY.

1. Je prie le Maître auquel ce volume est consacré de m'excuser si je ne puis lui offrir que l'hommage d'un travail portant sur une des rares branches de l'analyse au développement desquelles il n'ait pas contribué. J'ai toujours constaté que je diminuais singulièrement l'efficacité de mon travail en ne suivant pas la pente naturelle de mon esprit, qui m'a conduit, depuis plusieurs années, à m'occuper de calcul des probabilités; j'ai en conséquence pensé qu'il valait mieux ne pas forcer mon talent, et prier M. Goursat d'accepter l'hommage d'un travail qui complète mes travaux récents sur les suites de variables aléatoires enchaînées.

L'idée qui est à la base de ces travaux, que j'ai indiquée pour la première fois en 1929 à propos d'une application à l'étude des fractions continues (1), est que la plupart des théorèmes relatifs aux suites de variables aléatoires indépendantes les unes des autres peuvent s'étendre à une suite de variables enchaînées

$$u_1, \quad u_2, \quad \ldots, \quad u_n, \quad \ldots,$$

si l'on prend soin d'introduire, pour chacune de ces variables u_n, non sa loi de probabilité *a priori*, mais sa loi *a posteriori*, celle dont elle dépend lorsqu'on connaît $u_1, u_2, \ldots, u_{n-1}$, et qui, en pratique, caractérise les conditions de l'expérience qui doit déterminer u_n. Il est bien

(1) *Bulletin de la Société mathématique de France*, t. 57, 1929, p. 190.

connu que, sans cette précaution, l'extension des théorèmes asymptotiques les plus simples est impossible; elle devient au contraire facile grâce à l'introduction de ces lois *a posteriori*.

L'application la plus simple de cette remarque conduit à penser que l'on a, sous des conditions assez peu restrictives, une bonne évaluation de la somme

$$S_n = u_1 + u_2 + \ldots + u_n,$$

en remplaçant chaque terme u_ν, non par $\mathcal{E}\{u_\nu\}$, mais par $\mathcal{E}_{\nu-1}\{u_\nu\}$ (1). Sans doute objectera-t-on que la valeur approchée ainsi obtenue est une variable aléatoire, et n'a pas la valeur pratique d'une évaluation *a priori*. Mais en calcul des probabilités, du moins dans une théorie générale, on ne peut que préciser la relation probable entre la loi de probabilité et le résultat de l'expérience, entre la cause et l'effet; les énoncés obtenus pourront seulement conduire à des conclusions plus précises dans les applications particulières où l'on précisera la manière dont les conditions de chaque expérience dépendent des résultats des précédentes. L'application déjà mentionnée à l'étude des fractions continues suffit à montrer l'intérêt de cette méthode.

Les considérations qui précèdent conduisent à étudier l'erreur commise sur S_n en remplaçant chaque terme u_ν par $\mathcal{E}_{\nu-1}\{u_\nu\}$, ou, ce qui revient au même, en changeant de notations et en désignant par u_n ce qui a été jusqu'ici désigné par $u_n - \mathcal{E}_{n-1}\{u_n\}$, d'étudier les sommes de termes u_n dans le cas où est vérifiée l'hypothèse

$$(\mathcal{C}) \qquad \mathcal{E}_{n-1}\{u_n\} = 0.$$

J'ai étudié quelques conséquences de cette hypothèse dans un Mémoire récent (2), dont le résultat principal est l'extension à ces variables du théorème classique indiquant le rôle de la loi de Gauss. Après avoir rappelé ce résultat, j'aborderai la question qui est l'objet propre du présent travail : l'extension aux suites de variables aléatoires enchaînées de la loi forte des grands nombres et du théorème plus

(1) $\mathcal{E}\{x\}$ et $\mathcal{E}_n\{x\}$ désignent respectivement les valeurs probables d'une variable x, *a priori*, lorsqu'on ne connaît aucun des u_n et, *a posteriori*, lorsque $u_1, u_2, \ldots, u_n$ sont connus.

(2) *Bulletin des Sciences mathématiques*, t. 59, 1935, p. 84 à 96 et 109 à 128. Ce Mémoire sera désigné dans la suite par l'abréviation *var. ench.*

précis connu sous le nom de loi du logarithme itéré. Cette extension n'étant possible qu'en introduisant $\mathcal{E}_{n-1}\{u_n\}$, et non $\mathcal{E}\{u_n\}$, il suffit de considérer le cas où la condition $(\mathcal{C})$ est vérifiée. La démonstration de la loi du logarithme itéré repose essentiellement sur le rôle de la loi de Gauss, mais pour la loi forte des grands nombres, qui en est un corollaire évident, je donnerai un démonstration indépendante de la loi de Gauss (¹).

2. Rappel de résultats connus. — Il est d'abord à peu près évident que la condition $(\mathcal{C})$ suffit pour l'extension des formules connues relatives aux moments du second ordre. En posant

$$\mu_n^2 = \mathcal{E}_{n-1}\{u_n^2\}, \qquad m_n^2 = \mathcal{E}\{u_n^2\} = \mathcal{E}\{\mu_n^2\}, \qquad \mathrm{M}_n^2 = \mathcal{E}\{\mathrm{S}_n^2\},$$

on a, si la condition $(\mathcal{C})$ est vérifiée,

$$\mathrm{M}_n^2 = m_1^2 + m_2^2 + \ldots + m_n^2 \quad (^2),$$

et, compte tenu de l'inégalité de Tchebycheff, cette formule permet immédiatement l'extension de la loi des grands nombres sous sa forme classique : les valeurs de S_n grandes par rapport à M_n sont très peu probables.

La quantité qui va jouer un rôle essentiel dans la suite est, non M_n^2, mais la somme

$$\sigma_n^2 = \mu_1^2 + \mu_2^2 + \ldots + \mu_n^2,$$

qui est une variable aléatoire, et dont nous supposerons essentielle-

---

(¹) Ces résultats justifient l'énoncé que j'avais indiqué sans démonstration dans mon Mémoire de 1929 cité plus haut. Si je m'étais contenté d'un énoncé sans démonstration, c'était en partie pour ne pas interrompre par une trop longue digression un travail consacré à l'étude des fractions continues, en partie parce que, n'étant pas sûr d'avoir lu tous les travaux publiés sur la loi forte des grands nombres, je pensais qu'un résultat si simple pouvait être connu; je suis arrivé depuis à la conclusion qu'il était nouveau, et je ne crois pas que sa démonstration ait été publiée jusqu'ici.

Le théorème fondamental du présent travail a été énoncé, avec un résumé de sa démonstration, dans une Note présentée le 26 août 1935 à l'Académie des Sciences (*Comptes rendus*, **201**, p. 493).

(²) Ce résultat a été obtenu par M. Cantelli (antérieurement à mon Mémoire de 1935).

ment que, sauf peut-être dans des cas dont la probabilité est nulle, elle augmente indéfiniment avec n; c'est ce que nous appellerons la *condition* $(\mathcal{C}'')$ (1). Pour permettre l'introduction d'une variable continue, nous considérerons, non seulement les sommes S_n, mais les expressions intermédiaires

$$S_n + ku_{n+1} \qquad (0 \leqq k < 1),$$

et poserons

$$\sigma_n^2 + k^2 \mu_{n+1}^2 = t.$$

C'est cette expression t que nous prendrons pour variable, et non n. A une valeur donnée de t correspondent des valeurs de n et k qui sont des variables aléatoires, mais toujours déterminées sans ambiguïté, et c'est en fonction de ce paramètre t que j'ai, dans le Mémoire cité (*var. ench.*, p. 113 à 117), étudié l'expression

$$S_n + ku_{n+1} = X(t) = x(t)\sqrt{t},$$

$x(t)$ étant la variable réduite : sa valeur quadratique moyenne est 1.

Il faut naturellement introduire une condition limitant $|u_n|$, sans laquelle on ne pourrait même pas affirmer que la condition $(\mathcal{C}'')$ entraîne la divergence presque sûre de Σu_n. Nous supposerons tous les termes de $X(t)$ majorés par une expression que nous représenterons par $\varphi(t)\sqrt{t}$; c'est ce que nous appellerons la condition $(\mathcal{C}')$ (2). Si $\varphi(t)$

---

(1) Pour conserver les notations de mon précédent travail, je réserve la notation $(\mathcal{C}')$ pour la condition qui limitera supérieurement $|u_n|$.

(2) On peut, sans rien restreindre, supposer la fonction $\Phi(t) = \varphi(t)\sqrt{t}$ non décroissante. La condition $(\mathcal{C}')$, si elle est vérifiée pour t quelconque, entraîne en effet la condition analogue relative à la fonction $\operatorname{Min}\Phi(\tau)$ (pour $\tau \geqq t$). Cela revient alors au même de dire qu'elle majore le plus grand terme de $X(t)$, ou de dire qu'elle majore le dernier terme.

Au point de vue des résultats asymptotiques que nous établirons, il résulte de principes aujourd'hui bien connus qu'ils subsistent si, au lieu de la condition $(\mathcal{C}')$, on fait intervenir seulement l'hypothèse moins restrictive que la probabilité

$$\mathcal{P}_{n-1}\{|u_n| > \sigma_n \varphi(\sigma_n^2)\}$$

est le terme général d'une série presque sûrement convergente ($\mathcal{P}_{n-1}$ étant une probabilité conditionnelle, évaluée lorsqu'on connaît $u_1, \ldots, u_{n-1}$). En effet dans cette hypothèse, la condition $(\mathcal{C}')$ est vérifiée à partir d'un certain moment, sauf dans des cas de probabilité arbitrairement petite.

est borné, et même sous la condition moins restrictive que l'intégrale $\int \frac{dt}{t\varphi^2(t)}$ soit divergente, on voit aisément que cette condition, jointe aux précédentes, entraîne la divergence presque sûre de Σu_n.

Le résultat fondamental établi dans mon Mémoire cité (*var. ench.*, p. 117), et qui nous servira de point de départ, est le suivant : si les conditions $(\mathcal{C})$, $(\mathcal{C}')$ et $(\mathcal{C}'')$ sont vérifiées, on a

$$(1) \qquad |\mathcal{P}\{x(t)+\lambda z<\xi\}-\mathcal{P}\{y+\lambda z<\xi\}|<\frac{K}{\lambda^3}\varphi(t).$$

Dans cette formule, y est une variable aléatoire dépendant de la loi de Gauss réduite; z est une variable aléatoire indépendante à la fois des u_n, donc de $x(t)$, et de y; nous supposerons pour la suite

$$|z|\leq 1 \qquad \text{et} \qquad \mathcal{E}\{z\}=0.$$

Le coefficient K dépend de la loi de probabilité dont dépend z; λ est positif. Dans mon Mémoire cité, je n'avais pas introduit le facteur λ; mais, K ayant la valeur

$$K = K' \operatorname{Max}\left|\frac{d^3\mathcal{P}\{z<\zeta\}}{d\zeta^3}\right|$$

(K′ étant une constante absolue; la fonction $\mathcal{P}\{z<\zeta\}$ et ses deux premières dérivées sont supposées continues), il est évident qu'on peut remplacer à la fois z par λz et K par $\frac{K}{\lambda^3}$.

Il résulte de (1) que, si $\varphi(t)$ tend vers zéro avec $\frac{1}{t}$, la loi dont dépend $x(t)$ tend vers celle de Gauss; on peut en effet rendre à la fois λ et $\frac{\varphi(t)}{\lambda^3}$ aussi petits que l'on veut.

3. Énoncé du théorème fondamental.

Théorème. — *Si, la fonction $\varphi(t)$ étant monotone et telle que l'intégrale*

$$(2) \qquad \int_1^\infty \varphi(t)\frac{dt}{t}$$

soit finie, les conditions $(\mathcal{C})$, $(\mathcal{C}')$ et $(\mathcal{C}'')$ sont vérifiées, la probabilité

qu'il existe des valeurs de t arbitrairement grandes et telles que

$$x(t) > c\sqrt{2\log\log t} \tag{3}$$

est 0 *si* $c > 1$ *et* 1 *si* $c < 1$.

Nous allons d'abord, au paragraphe **4**, considérer seulement les valeurs de t appartenant à une progression géométrique de raison $q > 1$; cette restriction ne pouvant que diminuer la probabilité étudiée, et la probabilité ainsi diminuée étant 0 si $c > 1$ et 1 si $c < 1$, la seconde partie du théorème énoncé se trouvera ainsi démontrée, mais non la première. Après avoir indiqué au paragraphe **5** un résultat un peu plus précis concernant cette seconde partie, nous compléterons au paragraphe **6** la démonstration de la première partie par l'étude du maximum de $X(t)$ dans chacun des intervalles limités par les valeurs de t considérées d'abord.

Les résultats établis pour $X(t)$ s'appliqueront naturellement à $-X(t)$, et par suite aussi à $|X(t)|$.

4. ÉTUDE DE $X(t)$ POUR LES VALEURS $t_p = q^p$ DE t. — La fonction $\varphi(t)$ étant monotone, la convergence de l'intégrale (2) entraîne celle de la série $\Sigma\varphi(t_p)$ $(p = 1, 2, \ldots)$. Il résulte alors de (1) que, en posant

$$\alpha_p = \mathcal{P}\{x(t_p) > c'\sqrt{2\log\log t_p} - z\},$$
$$\beta_p = \mathcal{P}\{\quad y \quad > c'\sqrt{2\log\log t_p} - z\},$$

les deux séries $\Sigma\alpha_p$ et $\Sigma\beta_p$ sont de la même nature, c'est-à-dire, d'après les propriétés connues de la loi de Gauss, et puisque $|z| \leqq 1$, convergentes si $c' > 1$ et divergentes si $c' < 1$. Lorsque $c > 1$, en choisissant c' entre 1 et c, on voit qu'il est presque sûr que l'on aura à partir d'un certain moment

$$x(t_p) \leqq c'\sqrt{2\log\log t_p} - z \leqq c\sqrt{2\log\log t_p}.$$

Au contraire, dans le cas où $c' < 1$, le fait que les probabilités α_p ne soient pas indépendantes les unes des autres ne permet de tirer aucune conclusion de la divergence de la série $\Sigma\alpha_p$.

Pour traiter ce cas, il faut appliquer le lemme II de mon Mémoire cité (*var. ench.*, p. 91) qui étend au cas des variables enchaînées,

conformément au principe rappelé au début du présent travail, le résultat connu, de M. Borel, relatif aux variables indépendantes. D'après ce lemme, E_p désignant un événement lié à la détermination de $X(t_p)$, et $\mathcal{P}'\{E_p\}$ indiquant sa probabilité évaluée lorsqu'on connaît tous les termes de $X(t_{p-1})$, il est presque sûr que le nombre de valeurs de p pour lesquelles E_p est réalisé, et la somme $\Sigma\mathcal{P}'\{E_p\}$, sont en même temps finis ou en même temps infinis. Si donc la série $\Sigma\mathcal{P}'\{E_p\}$ diverge, sûrement ou presque sûrement, il est presque sûr que E_p se trouve réalisé pour une infinité de valeurs de p.

$X(t_{p-1})$ étant compris entre deux sommes consécutives S_n et S_{n+1} (pour simplifier l'écriture, nous écrivons n, et non n_{p-1}, notation qui serait plus logique), on a

$$X(t_{p-1}) = S_n + ku_{n+1}$$

et, lorsque tous les termes de $X(t_{p-1})$ sont connus, S_{n+1} est connu, et

$$X(t_p) - \mathcal{E}'\{X(t_p)\} = X(t_p) - S_{n+1} \tag{4}$$

est une somme de termes u_ν vérifiant la condition $(\mathcal{C})$, ne dépassant pas en valeur absolue $\varphi(t_p)\sqrt{t_p}$, et tels que

$$\Sigma\mathcal{E}_{\nu-1}\{u_\nu^2\} = \mathcal{E}'\{[X(t_p) - S_{n+1}]^2\} = t_p - t_{p-1} - (1-k^2)\mu_{n+1}^2,$$

c'est-à-dire

$$\Sigma\mathcal{E}_{\nu-1}\{u_\nu^2\} = t_p(1-\eta_p) \qquad \left[\eta_p = \frac{1}{q} + (1-k^2)\frac{\mu_{n+1}^2}{t_p}\right]. \tag{5}$$

Cette quantité η_p est d'ailleurs arbitrairement petite si $q = \frac{t_p}{t_{p-1}}$ est assez grand, et si p est aussi assez grand; car, $\varphi(t_p)$ tendant vers zéro, le plus grand terme de $X(t_p)$, et *a fortiori* μ_{n+1} sont $o(\sqrt{t_p})$. Dans ces conditions on peut appliquer l'inégalité (1) en remplaçant $x(t)\sqrt{t}$ par l'expression (4), le premier symbole $\mathcal{P}$ par $\mathcal{P}'$, et t par l'expression (5). Au second membre, pour $\lambda = 1$, on trouve $\frac{K\varphi(t_p)}{\sqrt{1-\eta_p}}$, qui est le terme général d'une série convergente. Les séries

$$\Sigma\mathcal{P}'\left\{\frac{X(t_p) - S_{n+1}}{\sqrt{t_p(1-\eta_p)}} > c'\sqrt{2\log\log t_p} - z\right\}, \qquad \Sigma\mathcal{P}\{y > c'\sqrt{2\log\log t_p} - z\}$$

sont donc de la même nature, c'est-à-dire divergentes, si $c < 1$ et si c'

est choisi entre c et 1. Compte tenu du lemme que nous venons de rappeler il est alors presque sûr que l'on a une infinité de fois

$$\frac{X(t_p) - S_{n+1}}{\sqrt{t_p(1-\eta_p)}} > c'\sqrt{2\log\log t_p} - z, \tag{6}$$

et *a fortiori*

$$\frac{X(t_p)}{\sqrt{t_p}} > c\sqrt{2\log\log t_p}. \tag{7}$$

En effet, en négligeant η_p et z, on commet, si q et p sont assez grands, des erreurs relatives arbitrairement petites; il en est de même si l'on néglige

$$S_{n+1} = X(t_{p-1}) + (1-k)\,u_{n+1},$$

le premier terme étant, d'après le résultat établi au début du présent paragraphe, presque sûrement inférieur à partir d'un certain moment à $2\sqrt{t_{p-1}\log\log t_{p-1}}$, et le second étant inférieur à $\varphi(t_p)\sqrt{t_p}$. Ces erreurs peuvent être compensées par une modification très petite du coefficient numérique. Si donc l'on a pris q assez grand, (6) entraîne (7) pour les grandes valeurs de p; il est donc presque sûr que l'inégalité (7) est, comme l'inégalité (6), vérifiée une infinité de fois. C. Q. F. D.

L'hypothèse que q soit très grand, qui nous a permis de négliger t_{p-1} devant t_p et $X(t_{p-1})$ devant $\sqrt{t_p\log\log t_p}$, n'est d'ailleurs pas essentielle pour la conclusion qui reste vraie *a fortiori* si, en intercalant des termes entre les différentes puissances de q, on remplace la progression de raison q par une progression de raison $\sqrt{q}$, ou $q^{\frac{1}{h}}$ (h étant entier). Elle est donc établie pour tout $q > 1$.

5. Indication d'un résultat plus précis. — On peut obtenir un tel résultat en prenant pour les t_p une suite de valeurs croissant un peu plus rapidement qu'une progression géométrique. Le résultat le plus précis que l'on puisse déduire des hypothèses $(\mathcal{C})$, $(\mathcal{C}')$ et $(\mathcal{C}'')$, *du moins par le mode de raisonnement employé*, s'obtient en prenant

$$t_p = e^{p\sqrt{\log p}}, \tag{8}$$

et en appliquant la formule (1), pour

$$\lambda = \lambda_p = k(\log p)^{-\frac{1}{6}}$$

(k étant un facteur constant arbitrairement petit) à l'étude de la probabilité

$$(9) \qquad \mathcal{P}'\left\{\frac{X(t_p)-S_{n+1}}{\sqrt{t_p(1-\eta_p)}} > f(t_p)-\lambda_p z\right\}.$$

Au second membre de la formule (1), on a dans ces conditions

$$\frac{K}{\lambda_p^3}\frac{\varphi(t_p)}{\sqrt{1-\eta_p}} \sim \frac{K}{k^3}\varphi(P)\frac{d\log P}{dp} \qquad (P = e^{p\sqrt{\log p}}),$$

ce qui est le terme d'une série convergente. L'expression (9) est alors le terme général d'une série de même nature que la série

$$\Sigma\mathcal{P}\{y > f(t_p)-\lambda_p z\},$$

et par suite divergente, si l'on prend

$$f(t) = \sqrt{2\log\log t} - \varepsilon(\log\log t)^{-\frac{1}{6}},$$

avec $\varepsilon > k$, ce qui n'empêche pas ε d'être arbitrairement petit. Il est donc presque sûr que l'on a une infinité de fois

$$\frac{X(t_p)-S_{n+1}}{\sqrt{t_p(1-\eta_p)}} > \sqrt{2\log\log t_p} - \varepsilon(\log\log t_p)^{-\frac{1}{6}}.$$

Le raisonnement se termine comme au paragraphe 4. On peut négliger au premier membre

$$\eta_p \sim \frac{t_{p-1}}{t_p} \sim e^{-\sqrt{\log p}} = o\left[(\log\log t_p)^{-\frac{2}{3}}\right],$$

l'erreur relative ayant moins d'importance qu'une variation très petite de ε n'en a au second membre. On peut de même négliger S_{n+1}, expression peu différente de $X(t_p)$, et par suite [le premier résultat du paragraphe 4 s'appliquant sans difficulté lorsque t_p a la valeur (8)] presque sûrement inférieure, pour p assez grand, à

$$2\sqrt{t_{p-1}\log\log t_{p-1}} = o\left[\sqrt{t_p}(\log\log t_p)^{-\frac{2}{3}}\right].$$

Finalement : *les hypothèses étant celles du théorème fondamental, ε étant un nombre positif arbitrairement petit, il est presque sûr que l'on a une infinité de fois, même en ne considérant que les valeurs de t de la*

forme (8),

$$x(t) > \sqrt{2\log\log t} - \varepsilon(\log\log t)^{-\frac{1}{n}}. \tag{10}$$

6. ÉTUDE DE $x(t)$ DANS LES INTERVALLES (t_{p-1}, t_p). — Il est entendu que t_p a de nouveau la valeur q^p; q sera ici un nombre qu'il y a intérêt à rendre voisin de 1, de manière à pouvoir négliger $t_p - t_{p-1}$ devant t_p. Il s'agit d'étudier le maximum M_p de $X(t)$ dans l'intervalle (t_{p-1}, t_p) et de montrer que, c' étant choisi entre 1 et $c<1$, on a presque sûrement à partir d'une certaine valeur de p

$$M_p \leqq c' T_p \qquad (T_p = \sqrt{2 t_p \log\log t_p}). \tag{11}$$

Il en résultera bien, si $q-1$ est assez petit, qu'on a presque sûrement, à partir d'une certaine valeur de t,

$$X(t) \leqq c\sqrt{2\, t \log\log t}.$$

Pour borner M_p, nous utiliserons un principe de raisonnement bien connu, que MM. Khintchine, Kolmogoroff et moi avons souvent appliqué depuis quelques années (la priorité, pour ce genre de raisonnements, semble appartenir à Désiré André) : si, avec une probabilité non négligeable, $X(t)$ pouvait dépasser dans l'intervalle (t_{p-1}, t_p) la valeur $c' T_p$, il y aurait peu de chances, cette circonstance se produisant, que la valeur finale $X(t_p)$ soit inférieure, sinon à $c' T_p$, du moins à $c'' T_p$, c'' étant choisi entre 1 et c'. Contrairement au résultat obtenu paragraphe **4**, on ne pourrait pas négliger la probabilité de l'inégalité

$$X(t_p) \geqq c'' T_p.$$

On peut préciser le raisonnement de plusieurs manières; l'application de la formule (1) à $X(t_p) - X(t)$ permettrait d'éviter l'introduction d'une nouvelle constante c''. Nous allons indiquer un raisonnement plus élémentaire, n'utilisant que l'inégalité de Tchebycheff.

Plaçons-nous dans l'hypothèse

$$M_p > c' T_p \qquad (c' > 1). \tag{12}$$

Comme nous pouvons écarter l'hypothèse que $X(t_{p-1})$ et $X(t_p)$ dépassent $c' T_p$, l'hypothèse (12) implique l'existence, dans l'intervalle considéré, d'au moins une valeur de t correspondant à un entier n (ici

encore nous écrivons pour simplifier n au lieu de n_p) et telle que

$$X(t) = S_n > c' T_p.$$

S'il y en a plusieurs, nous choisirons la plus petite. On a

$$\mathcal{E}_n \{ [X(t_p) - S_n]^2 \} = t_p - t < t_p - t_{p-1},$$

et par suite, $\mathcal{E}'$ désignant une valeur probable évaluée dans l'hypothèse (12), mais sans connaître n, ni t, ni S_n,

$$\mathcal{E}' \{ [X(t_p) - S_n]^2 \} < t_p - t_{p-1} = \frac{q-1}{q} t_p \quad (^1).$$

L'inégalité de Tchebycheff donne donc

$$\mathcal{P}' \{ | X(t_p) - S_n | \geq (c' - c'') T_p \} \leq \frac{q-1}{2q(c'-c'')^2 \log\log t_p},$$

et, pour p assez grand, cette probabilité est inférieure à un nombre arbitrairement petit ε. La probabilité de l'inégalité inverse dépasse donc $1-\varepsilon$; on en déduit

$$\mathcal{P} \{ X(t_p) > c'' T_p \} > (1-\varepsilon) \mathcal{P} \{ M_p > c' T_p \}.$$

La convergence de la série

$$\Sigma \mathcal{P} \{ X(t_p) > c'' T_p \},$$

établie au paragraphe 4 pour tout $c'' > 1$, entraîne donc celle de

$$\Sigma \mathcal{P} \{ M_p > c' T_p \}.$$

Il donc presque sûr qu'il existe une valeur de p à partir de laquelle l'inégalité (11) est constamment vérifiée. C. Q. F. D.

7. La loi forte des grands nombres. — Considérons une suite d'expériences enchaînées, susceptibles de réaliser des événements $E_1, E_2, \ldots,$

---

(1) Cela résulte de ce que les symboles $\mathcal{E}'\{..\}$ et $\mathcal{E}'\{\mathcal{E}_n\{..\}\}$ sont équivalents, d'après le théorème de Fubini, qui s'applique sans difficulté toutes les fois que l'on est en présence de lois de probabilité bien définies, et qu'en conséquence on n'est jamais conduit, par l'emploi d'un nombre fini ou d'une infinité dénombrable d'inégalités, qu'à considérer des ensembles mesurables et à y intégrer des fonctions bien définies, sauf dans des cas de probabilité nulle, et mesurables.

E_n, ...; désignons par

$$\alpha_n = \mathcal{P}_{n-1}\{E_n\}$$

la probabilité de E_n lorsque les résultats des $n-1$ premières expériences sont connus. Désignons par ν_n une variable égale à 1 si E_n est réalisé et à zéro dans le cas contraire, et posons $u_n = \nu_n - \alpha_n$.

Cette variable u_n vérifie les conditions $(\mathcal{C})$ et $(\mathcal{C}')$; on remarque même que $|u_n| \leqq 1$; on peut donc remplacer $\varphi(t)\sqrt{t}$ par 1. Supposons de plus la condition $(\mathcal{C}'')$ vérifiée, c'est-à-dire que la série

$$\Sigma \mu_n^2 = \Sigma \alpha_n (1 - \alpha_n),$$

et par suite les séries $\Sigma\alpha_n$ et $\Sigma(1-\alpha_n)$, sont presque sûrement divergentes. D'après le lemme II de mon Mémoire déjà cité, il est donc presque sûr que chacune des deux éventualités, $u_n = 0$ et $u_n = 1$, se produit une infinité de fois.

Le théorème fondamental nous montre alors que, pour tout $c > 1$, il existe presque sûrement une valeur N telle que, pour $n > N$, on ait

$$S_n < c\sqrt{2\sigma_n \log\log\sigma_n} < c\sqrt{2A_n \log\log A_n} = o(A_n),$$

en posant $A_n = \alpha_1 + \alpha_2 + \ldots + \alpha_n$. Donc : *si σ_n augmente indéfiniment avec n, il est presque sûr que le nombre*

$$V_n = \nu_1 + \nu_2 + \ldots + \nu_n = A_n + S_n$$

est un infiniment grand équivalent à A_n.

C'est *la loi forte des grands nombres.*

On peut en donner une démonstration indépendante du rôle de la loi de Gauss. Elle utilise l'inégalité connue de M. Kolmogoroff, ou plutôt l'extension de cette inégalité aux suites de variables enchaînées vérifiant la condition $(\mathcal{C})$ (*voir* la démonstration *var. ench.*, p. 94). D'après cette inégalité, on a

$$\mathcal{P}\left\{\operatorname*{Max}_{\tau \leqq t} |X(\tau)| \geqq c\sqrt{t}\right\} \leqq \frac{1}{c^2}. \tag{13}$$

Appliquons cette formule pour $t = t_p = q^p$, $c = q^{\alpha p}$ $(q > 1, \alpha > 0)$. La probabilité considérée étant le terme général d'une série convergente, il existe presque sûrement une valeur P de p telle que,

pour $p > P$, on ait

$$\operatorname*{Max}_{\tau < t_p} |X(\tau)| < t_p^{\frac{1}{2}+\alpha},$$

et par suite, pour $t > q^{P-1}$,

$$|X(t)| < (qt)^{\frac{1}{2}+\alpha},$$

α étant d'ailleurs arbitrairement petit. Comme cette expression est $o(t)$, elle suffit pour la démonstration de la loi forte des grands nombres.

Il y a lieu de remarquer que ce raisonnement subsiste sans modification si l'hypothèse $(\mathcal{C}'')$ n'est pas réalisée, $\operatorname*{Max}_{\tau \leq t} |X(\tau)|$ désignant alors le maximum de $|X(\tau)|$ pour les valeurs de τ au plus égales à t et effectivement atteintes ou dépassées par σ_n^2. L'hypothèse $(\mathcal{C}')$ n'intervient pas non plus.

On peut aussi, lorsque $(\mathcal{C})$ et $(\mathcal{C}')$ sont vérifiées, faire abstraction de $(\mathcal{C}'')$ pour la démonstration de la loi du logarithme itéré, du moins pour la partie de l'énoncé relative au cas où $c > 1$. Le moyen le plus simple d'y arriver est, sans modifier la suite des u_n d'indices finis, de la prolonger par des u_n d'indices transfinis, de manière que n'importe quelle valeur de t soit sûrement atteinte ou dépassée par σ_n^2 pour une valeur finie ou transfinie de n; nous pouvons d'ailleurs, pour chaque u_n d'indice transfini, nous donner une loi déterminée, indépendante des u_ν d'indices inférieurs à n, par exemple celle du jeu de pile ou face avec enjeu unité. Alors σ_n^2 augmente indéfiniment au plus tard lorsque n tend vers 2ω, et il n'y a à considérer que des nombres transfinis de la forme $n = \omega + p$ (p étant fini). La variable aléatoire $X(t)$ est dans ces conditions toujours bien définie; on la modifie d'ailleurs aussi peu qu'on veut, sauf dans des cas de probabilité arbitrairement petite, en faisant abstraction des u_n d'indices finis supérieurs à un nombre suffisamment grand, et comme on est ainsi ramené au cas où $X(t)$ ne comprend qu'un nombre fini de termes, la formule (1) s'applique, ainsi que la loi du logarithme itéré, qui en résulte. Or on ne peut que diminuer les chances de trouver pour t arbitrairement grand des valeurs de $X(t)$ définies et supérieures à $c\sqrt{2t \log\log t}$ si l'on supprime les indices transfinis, de sorte que la somme $X(t)$ se trouve dans certains cas cesser d'être définie, sans être modifiée dans

les cas où elle reste définie; la probabilité de l'existence de ces grandes valeurs (pour t arbitrairement grand) est donc nulle pour tout $c > 1$.

D'autre part, d'après le Théorème I de mon Mémoire cité (*var. ench.*, p. 96), la convergence de $\Sigma\mu_n^2$ entraîne presque sûrement celle de Σu_n; dans le cas où $u_n = v_n - \alpha_n$, $S_n = V_n - A_n$, V_n et A_n différent donc d'une quantité tendant vers une limite. On en déduit un nouvel énoncé de la loi des grands nombres, indépendant de toute hypothèse sur σ_n^2 : *il est presque sûr que* V_n *et* A_n, *ou bien restent tous deux finis, ou bien sont des infiniment grands équivalents* (1).

(1) Cet énoncé comprend celui dont j'ai parlé plus haut [p. 13 (1)], que j'avais indiqué sans démonstration en 1929. A cette époque je ne connaissais pas l'inégalité de M. Kolmogoroff. Mais j'avais pu m'en passer en utilisant l'inégalité de Tchebycheff, qui permet de borner supérieurement $X(t_p)$, avec une probabilité aussi grande que l'on veut, et le principe de raisonnement qui m'a servi au paragraphe **6** du présent Mémoire, qui permet d'en déduire une borne supérieure, aussi très probablement exacte, de M_p.