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

PAUL LÉVY

Propriétés asymptotiques des sommes de variables aléatoires indépendantes ou enchaînées

Journal de mathématiques pures et appliquées 9e série, tome 14, nº 1-4 (1935), p. 347-402.
<http://www.numdam.org/item?id=JMPA_1935_9_14_1-4_347_0>

Propriétés asymptotiques des sommes de variables aléatoires indépendantes ou enchaînées;

PAR M. PAUL LÉVY.

INTRODUCTION.

On sait depuis longtemps le rôle que jouent, dans l'étude d'une somme

$$S_n = x_1 + x_2 + \ldots + x_n,$$

de variables aléatoires indépendantes et à valeurs probables nulles, les grandes valeurs possibles des différents termes. En dehors du cas de convergence probable de Σx_n, seule l'existence de ces grandes valeurs peut empêcher la variable aléatoire S_n d'être à la limite du *type* de Gauss (¹), quand n augmente indéfiniment. Une étude systématique du rôle de ces grandes valeurs nous a conduit à des résultats nouveaux, dont le plus important est une condition *nécessaire et suffisante* pour que l'on obtienne à la limite le type de Gauss; les conditions indiquées jusqu'ici n'étaient que suffisantes; du moins on n'avait pas démontré qu'elles étaient nécessaires (²).

(¹) Nous disons que des lois de probabilité, ou des variables aléatoires, sont du même type, si l'on peut passer de l'une à l'autre par un changement de variable linéaire; dire que S_n est à la limite du type de Gauss équivaut donc à dire qu'on peut déterminer N et a_n en fonction de n de manière que $\frac{S_n}{N} - a_n$ dépende d'une loi tendant vers celle de Gauss; N est le *coefficient de réduction*.

(²) La condition du théorème VI de mon Mémoire du *Studia Mathematica*, t. III, est nécessaire et suffisante. J'avais seulement montré qu'elle était suffisante, et ajouté : « Il n'est peut-être pas impossible de donner dans cet ordre d'idées une condition nécessaire et suffisante ».

Il faut avant tout écarter les termes *non individuellement négligeables;* nous désignerons ainsi tout terme x_n qui, dans des cas de probabilité non négligeable, est trop grand pour être négligé dans l'étude de S_n. S'il existe une suite illimitée de termes non individuellement négligeables, il est clair que les variables S_{n-1} et S_n ne peuvent être simultanément à la limite du type de Gauss, pour les valeurs de n correspondant à ces termes, que si x_n est à la limite du type de Gauss. Comme il ne saurait à chaque instant exister qu'un nombre fini de termes non individuellement négligeables pour l'étude de S_n (pour une valeur donnée de n), cette remarque résout le problème en ce qui concerne le rôle de ces termes. On peut étudier aussi le problème de la convergence intermittente vers la loi de Gauss; il dépend d'un lemme que je n'ai pu démontrer et que j'indiquerai seulement à titre hypothétique; mais les résultats essentiels sont indépendants de cette hypothèse.

Ces grands termes étant écartés, on n'est pas sûr pour cela que le plus grand terme de S_n soit presque sûrement négligeable; on est seulement sûr que pour chaque valeur de ν au plus égale à n, la probabilité que x_ν soit négligeable (devant S_ν, donc *a fortiori* devant S_n, à cause du principe d'augmentation de la dispersion) est très faible. Mais l'addition des petites probabilités peut faire qu'il y ait une probabilité non négligeable, ou même voisine de l'unité, que le plus grand terme de S_n ne soit pas négligeable; seulement dans ce cas son rang sera inconnu, chaque valeur étant très peu probable.

Si le plus grand terme est presque sûrement négligeable [1], et cela d'autant plus exactement que n est plus grand, nous disons que l'*hypothèse des grands nombres est vérifiée*, ou que *la loi des grands nombres s'applique*. On sait que dans ces conditions S_n est à la limite du type de Gauss. Le résultat essentiel de ce travail, en ce qui concerne le cas des variables indépendantes, est que cette hypothèse

---

[1] Nous disons qu'une propriété *dépendant de n est presque sûre* si sa probabilité tend vers l'unité pour n infini. Une propriété de la suite des x_n, *ne dépendant d'aucune valeur particulière de n*, ne sera *presque sûre* que si sa probabilité est 1. Il s'agit donc dans le premier cas d'un énoncé du type bernoullien classique, et dans le second d'un énoncé se rattachant à la loi forte des grands nombres. Il ne sera question dans ce travail que d'énoncés du premier type.

est nécessaire pour que S_n soit à la limite du type de Gauss. En d'autres termes, si chaque terme de S_n est presque sûrement négligeable, *la condition nécessaire et suffisante pour la tendance vers la loi de Gauss est que le plus grand terme soit presque sûrement négligeable.*

Ces résultats sont exposés Chapitre III; le Chapitre II est consacré au cas simple où les différents termes x_n dépendent de la même loi de probabilité; il est facile dans ce cas d'éliminer n, et de donner à l'hypothèse des grands nombres une forme très simple ; en posant

$$Y = F(X) = \mathfrak{T}\{|x| > X\}, \qquad Z = \int_0^X x^2 \,|dF(x)|,$$

($\mathfrak{T}\{E\}$ désignant la probabilité d'un événement E), cette condition est que $\frac{X^2 Y}{Z}$ soit infiniment petit avec $\frac{1}{X}$. Le Chapitre II se termine par l'étude des circonstances possibles lorsque S_n n'est pas d'un type tendant vers celui de Gauss.

Le Chapitre I est un chapitre préliminaire, consacré à l'étude des *lois stables*, qui donnent l'exemple le plus simple de cas où la loi des grands nombres n'est pas applicable. Je rappelle qu'une loi est stable si, x_1 et x_2 étant indépendants l'un de l'autre et dépendant de cette loi, c_1 et c_2 étant des constantes positives quelconques, $c_1 x_1 + c_2 x_2$ est de la forme cx, c étant constant et x dépendant de la même loi que x_1 et x_2. La relation entre c_1, c_2 et c est nécessairement de la forme

$$c_1^\alpha + c_2^\alpha = c^\alpha \qquad (0 < \alpha \leqq 2),$$

α étant l'*exposant caractéristique* de la loi étudiée. Pour $\alpha = 2$, il n'y a pas d'autre loi stable que la loi de Gauss. Pour α compris entre zéro et 2, j'ai donné en 1924 la première théorie complète, et montré qu'il existe un type de loi symétrique L_α et un type dissymétrique $L_{\alpha,\beta}$ dépendant d'un second paramètre β (et se réduisant au premier pour $\beta = 0$). La démonstration était basée sur la définition asymptotique de ces lois en partant d'une loi simple de leur *domaine d'attraction.* J'ai donné depuis une définition constructive très simple de ces lois. Les deux méthodes utilisaient la notion de *fonction caractéristique.* L'objet du Chapitre I est de montrer que la définition constructive et la définition asymptotique peuvent être toutes les deux obtenues d'une manière tout à fait élémentaire, sans utiliser la fonction caractéristique.

Tout ce travail a d'ailleurs ce caractère élémentaire; on sait que le théorème limite classique du calcul des probabilités a été démontré d'une manière élémentaire par M. Lindeberg; c'est cette démonstration qu'il faut donc prendre comme point de départ si l'on veut rendre toute la théorie indépendante de la notion de fonction caractéristique [1].

Le Chapitre IV est consacré au cas des variables enchaînées. Il prend pour point de départ des résultats que j'ai énoncés dans une Note présentée le 1er octobre 1934 à l'Académie des Sciences et démontrés dans un Mémoire récemment publié [2]. Il s'agit, en prenant le langage commode de la théorie des jeux, du cas où l'enjeu est borné à chaque coup, et où, quoique la règle du jeu puisse dépendre des coups antérieurs, le jeu reste équitable. Je me suis proposé d'étudier, au lieu du gain S_n après n coups, le gain $\mathfrak{S}(t)$ au bout d'un temps t, en supposant la durée de chaque coup égale à la valeur probable du carré du gain. Dans ces conditions le théorème classique reste applicable : $\frac{\mathfrak{S}(t)}{t}$ dépend d'une loi tendant, pour t infini, vers celle de Gauss.

Dans le Chapitre IV du présent travail, je précise l'extension, déjà brièvement indiquée dans le précédent, de ce résultat au cas où les x_n ne sont pas bornés. La définition de t doit alors être légèrement modifiée, car la valeur probable de x_n^2 peut être infinie; la condition que le jeu soit équitable doit être remplacée par une condition de symétrie approchée en ce qui concerne les grandes valeurs du gain. Pour simplifier l'exposé, j'ai considéré surtout le cas de la symétrie parfaite, me bornant à indiquer dans des remarques l'extension au cas de la symétrie approchée. Dans ces conditions, j'ai pu étendre le résultat

---

[1] Du moins en ce qui concerne les petits termes. Pour les termes non individuellement négligeables, nous avons utilisé le fait que S_{n-1} et S_n ne peuvent être en même temps du type de Gauss que si x_n est de ce type. Quelque simple que soit sa démonstration, il est moins élémentaire, car il implique la définition de $\mathfrak{T}\{x_n < x\}$ à l'aide d'une équation intégrale.

[2] *Bulletin des Sciences Mathématiques*, mars-avril 1935.

fondamental du Chapitre III et montrer que, pour l'étude des sections à t constant, c'est-à-dire pour l'étude de $\mathfrak{S}(t)$ [mais non pour celle de S_n pour n donné], et s'il n'y a pas de terme non individuellement négligeable, l'hypothèse des grands nombres donne toujours la condition nécessaire et suffisante pour que $\mathfrak{S}(t)$ soit d'un type tendant vers celui de Gauss.

L'étude de S_n, si t n'est pas constant ou presque constant pour chaque valeur donnée de n, est plus difficile. Le Mémoire se termine par quelques remarques relatives à ce cas, et par le principe d'une méthode susceptible de s'appliquer à des cas où, le jeu n'étant plus équitable, le gain probable ne dépend que d'un petit nombre de termes précédant celui étudié.

DÉFINITIONS ET REMARQUES PRÉLIMINAIRES.

Nous désignerons par $\mathcal{E}\{x\}$ la valeur probable de x, et par $\sigma^2\{x\}$ celle de $(x - \mathcal{E}\{x\})^2$. Nous appellerons *dispersion* de la variable x, non la quantité $\sigma\{x\}$, mais une fonction $l = \varphi(\gamma)$ de la probabilité γ égale à la borne inférieure des nombres l' pour lesquels

$$\operatorname*{Max}_{-\infty < a < \infty} \mathcal{P}\{a < x < a + l'\} \geqq \gamma,$$

D'après l'inégalité de Tchebycheff, elle est au plus égale à $\frac{2\sigma}{\sqrt{1-\gamma}}$. Inversement, γ est la *concentration* (maxima) pour un intervalle de longueur l.

Quand nous parlons de l'ordre de grandeur d'une variable aléatoire, il s'agit de l'ordre de grandeur de sa dispersion. L'ordre de grandeur de S_n, ainsi défini, est souvent indépendant de γ; s'il en dépend, il est sous-entendu que nous prenons une valeur de γ arbitrairement voisine de l'unité, mais restant fixe quand n augmente indéfiniment. Un terme sera donc négligeable devant S_n s'il est négligeable, sauf dans des cas de probabilité arbitrairement petite, par rapport à la dispersion de S_n; il faut pour cela (d'après l'inégalité de Tchebycheff) qu'il soit négligeable devant $\sigma\{S_n\}$. Mais la réciproque n'est pas vraie; la considération de $\sigma\{S_n\}$, que cette quantité soit finie ou infinie, ne saurait suffire quand des valeurs de S_n très grandes et très peu probables interviennent dans son calcul pour une part non négligeable. C'est ce qui nous a obligé à utiliser la fonction de dispersion $\varphi(\gamma)$.

La dispersion ainsi définie a, comme l'expression σ, la propriété de ne pouvoir qu'augmenter par l'addition à la variable étudiée d'un terme indépendant d'elle; donc, inversement, la concentration ne peut que décroître. Cette remarque est importante pour la démonstration du théorème réciproque du théorème limite classique du calcul des probabilités. Le principe de cette démonstration est le suivant : on peut décomposer S_n en groupes de termes dont aucun n'est négligeable, de manière que, si l'hypothèse des grands nombres n'était pas vérifiée mais que chaque terme soit tout de même individuellement négligeable, la dispersion, pour les valeurs de γ assez voisines de l'unité, serait trop grande, pour un au moins des groupes de termes partiels, pour que leur somme dépende de la loi de Gauss.

CHAPITRE I.

DÉMONSTRATION ÉLÉMENTAIRE DE L'EXISTENCE DES LOIS STABLES.

1. Existence des lois stables dans le cas où $\alpha < 1$. — Considérons la portion $x > x_0 > 0$ de l'axe des x comme divisée en intervalles élémentaires dx, à chacun desquels on fait correspondre une variable aléatoire u, égale à x dans les cas de probabilité très petite

$$|dx^{-\alpha}| = \frac{\alpha\, dx}{x^{\alpha+1}} \qquad (\alpha > 0), \tag{1}$$

et nulle dans les autres cas. Ces variables sont indépendantes les unes des autres. D'après le principe des petites probabilités, on obtient ainsi, pour l'ensemble des valeurs non nulles de ces variables, lorsque la plus grande des probabilités correspondant aux intervalles élémentaires dx tend vers zéro, une loi limite bien déterminée. Pour tout intervalle $(x_1, x_2)(x_0 \leqq x_1 < x_2 \leqq \infty)$, le nombre n' des valeurs non nulles de u situées dans cet intervalle a pour valeur probable l'intégrale dans cet intervalle de la probabilité élémentaire (1), et dépend de la loi de Poisson. C'est une variable aléatoire dépendant d'une loi ainsi bien connue, et, n' étant supposé choisi d'après cette loi au cours d'une expérience préliminaire, on peut ensuite choisir chacune des n' valeurs u_ν de l'intervalle (x_1, x_2) par un tirage au sort indépendant des autres et donnant à chaque intervalle élémentaire une probabilité proportionnelle à l'expression (1). On reconstitue ainsi, pour l'ensemble des u, la loi définie d'abord autrement.

Appelons $\overline{S}$ la somme des variables u_ν. Le nombre des termes non nuls étant fini (quoiqu'on ne puisse lui assigner à l'avance aucune borne supérieure), il ne se pose aucune question de convergence. Seulement, les valeurs très grandes des u étant possibles, celles de $\overline{S}$ le seront aussi, et notamment la valeur probable

$$\mathcal{E}\{\overline{S}\} = \int_{x_0}^{\infty} x\,|\,dx^{-\alpha}\,| = \alpha \int_{x_0}^{\infty} x^{-\alpha}\,dx \tag{2}$$

sera infinie si $\alpha \leqq 1$.

Faisons maintenant tendre x_0 vers zéro ; cela revient, *si* $\alpha < 1$, *cas dont nous allons d'abord nous occuper*, à ajouter à $\overline{S}$ un terme, sûrement positif, et dont la valeur probable

$$\alpha \int_0^{x_0} x^{-\alpha}\,dx = \frac{\alpha}{1-\alpha}\,x_0^{1-\alpha}$$

est, si x_0 est assez petit, inférieur à un nombre arbitrairement petit ε^2 ; il est donc lui-même inférieur à ε, sauf dans des cas de probabilité inférieure à ε. Il en résulte que, quand x_0 tend vers zéro, $\overline{S}$ tend vers une limite S (sauf dans des cas dont la probabilité est nulle) et la loi dont dépend S est la limite de celle dont dépend $\overline{S}$.

Nous pouvons maintenant montrer bien simplement que *cette loi limite est une loi stable d'exposant caractéristique* α.

Soient en effet S′ et S″ deux variables aléatoires, indépendantes l'une de l'autre, dépendant de cette loi, et deux coefficients positifs c' et c''. Posons

$$c'^{\alpha} = C', \qquad c''^{\alpha} = C'', \qquad C' + C'' = C = c^{\alpha},$$

et proposons-nous d'étudier la somme $c'S' + c''S''$. Le premier terme est une variable aléatoire du même type que S′, à cela près que,

x étant remplacé par $\frac{x}{c'}$, la probabilité (1) se trouve multipliée par C′; la même remarque s'applique à $c''S''$. La somme étudiée apparaît donc comme la somme de deux séries de variables u'_ν et u''_ν telles que, pour chaque intervalle dx, les probabilités qu'un u'_ν ou un u''_ν soit dans cet intervalle sont respectivement

$$C'\,|\,dx^{-\alpha}\,| \quad \text{et} \quad C''\,|\,dx^{-\alpha}\,|.$$

D'après le principe des petites probabilités, on peut négliger la probabilité de l'existence d'un intervalle dx contenant à la fois un u'_ν et un u''_ν, et la somme étudiée peut être considérée comme la somme d'une seule série de variables u_ν, la probabilité que chaque intervalle dx en contienne une étant $C\,|\,dx^{-\alpha}\,|$. Cela revient à dire que cette somme est du type cS, S dépendant de la même loi que S′ et S″.

C. Q. F. D.

Nous avons ainsi établi, pour tout α compris entre zéro et 1, l'existence d'une loi stable d'exposant caractéristique α. La variable S qui dépend de cette loi étant essentiellement positive, elle s'identifie avec la loi $L_{\alpha,-1}$ de mes travaux antérieurs.

Si maintenant, en conservant les notations précédentes, nous considérons la somme

$$S = c'S' - c''S'' \qquad (c' > 0,\ c'' > 0), \tag{3}$$

ce qui revient à dire que cette fois nous ajoutons des valeurs u_ν des deux signes, la probabilité qu'un des x_ν soit dans un intervalle dx étant $\frac{C'\alpha\,|\,dx\,|}{|\,x\,|^{\alpha+1}}$ si x est positif et $\frac{C''\alpha\,|\,dx\,|}{|\,x\,|^{\alpha+1}}$ si x est négatif, nous obtenons encore une loi stable d'exposant caractéristique α, mais comportant des valeurs des deux signes. Il résulte immédiatement de (3) que la loi dont dépend $\frac{S}{c}$ s'identifie avec la loi $L_{\alpha,\beta}$ de mes travaux antérieurs pour

$$\beta = \frac{C' - C''}{C' + C''} = \frac{c'^{\alpha} - c''^{\alpha}}{c^{\alpha}}.$$

Les cas limites, $\beta = \pm 1$, s'obtiennent pour $c'' = 0$ et pour $c' = 0$.

Revenons au cas $\beta = -1$ considéré d'abord. La variable S est essentiellement positive, et il est bien évident que toutes les valeurs positives sont effectivement possibles. Ainsi, x_0 étant assez petit, la somme des u_ν inférieurs à x_0 peut, comme nous l'avons vu, être négligée, et il y a une probabilité positive pour qu'il n'y ait qu'un u_ν supérieur à x_0, ses valeurs possibles se répartissant d'une manière continue de x_0 à l'infini, proportionnellement à l'expression (1), c'est-à-dire qu'à chaque intervalle dx correspond la probabilité $x_0^\alpha |dx^{-\alpha}|$.

La probabilité des grandes valeurs de S est d'ailleurs facile à préciser. Si X est assez grand, on peut négliger la probabilité que deux des u_ν soient supérieurs à X; la somme des u_ν inférieurs à X a pour valeur probable $\frac{\alpha}{1-\alpha} X^{1-\alpha}$, et est inférieure à X', sauf dans des cas de probabilité au plus égale à $\frac{\alpha}{1-\alpha} \frac{X^{1-\alpha}}{X'}$. La probabilité des grandes valeurs de S est donc asymptotiquement la même que celle des grandes valeurs des u_ν, c'est-à-dire que

$$\mathfrak{P}\{S > X'\} \sim X'^{-\alpha} \qquad (X' \to \infty).$$

Ce résultat s'étend sans difficulté à la somme (2) pour laquelle on a :

$$\mathfrak{P}\{S > c'X\} \sim \mathfrak{P}\{S < -c''X\} \sim X^{-\alpha}.$$

On peut montrer aussi que, pour toute loi stable, la densité de probabilité ne comporte qu'un seul maximum. Pour celle dont dépend $\frac{c'S' - c''S''}{c}$, l'abscisse de ce maximum est évidemment une fonction continue, décroissante et impaire de β. Il serait intéressant de préciser ses valeurs extrêmes, ainsi que les valeurs correspondantes du maximum, et de déterminer l'ordre de grandeur de la probabilité des petites valeurs de $|S|$ pour les cas extrêmes ($\beta = \pm 1$). Je n'ai pas résolu ces questions.

2. Cas où $\alpha > 1$. — La définition de $\overline{S}$ donnée dans le cas précédent subsiste, mais il n'est plus possible de faire tendre x_0 vers zéro; la somme étudiée augmenterait indéfiniment.

Substituons alors à u_ν la quantité $u_\nu - \mathcal{E}\{u_\nu\}$, et par suite à $\overline{S}$ la valeur $\overline{S} - \mathcal{E}\{\overline{S}\}$, la valeur probable $\mathcal{E}\{\overline{S}\}$, donnée par la formule (2), étant finie dans le cas qui nous occupe. Il s'agit de savoir si cette expression a une limite, c'est-à-dire si la série $\Sigma[u_\nu - \mathcal{E}\{u_\nu\}]$ est convergente. Comme il n'y a de difficulté que pour les petites valeurs, nous pouvons ne considérer que les u_ν inférieurs à une valeur fixe U. Dans ces conditions, on sait que, la probabilité de la convergence étant toujours 1 ou zéro, la condition pour qu'elle ait la valeur 1 est la convergence de la série

$$\Sigma\mathcal{E}\{[u_\nu - \mathcal{E}\{u_\nu\}]^2\} = \Sigma x^2[|dx^{-\alpha}| - |dx^{-\alpha}|^2],$$

c'est-à-dire, les intervalles élémentaires dx tendant vers zéro, celle de l'intégrale

$$\int_0^{\mathrm{U}} x^2 |dx^{-\alpha}| = \alpha \int_0^{\mathrm{U}} \frac{dx}{x^{\alpha-1}}.$$

Il faut donc que $\alpha < 2$. Non seulement, ce qui était à prévoir, le cas où $\alpha > 2$ doit être exclu, mais le cas où $\alpha = 2$, celui de la loi de Gauss, échappe absolument au mode de définition des autres lois stables que nous exposons ici. Au contraire, α étant compris entre 1 et 2, la limite S de $\overline{\mathrm{S}} - \mathcal{E}\{\overline{\mathrm{S}}\}$, x_0 tendant vers zéro, est bien définie et dépend d'une loi stable, d'exposant caractéristique α, qui sera la loi $\mathrm{L}_{\alpha,-1}$. Les autres lois $\mathrm{L}_{\alpha,\beta}$ s'obtiennent, comme dans le cas précédent, par application de la formule (3).

Les formules relatives aux probabilités des grandes valeurs de la variable, établies dans le cas précédent, subsistent; mais S est infini, et naturellement les valeurs possibles de S varient de $-\infty$ à $+\infty$, même dans les cas limites $\beta = \pm 1$; c'est peut-être la différence la plus remarquable avec le cas où $\alpha < 1$.

3. Remarques. — 1° Nous avons laissé de côté le cas $\alpha = 1$. La loi symétrique L'_1 s'obtient aisément par le procédé exposé au paragraphe **1**, mais en associant les valeurs de x deux à deux égales et opposées; dans ces conditions la somme des valeurs $\mathcal{E}\{u_\nu\}$ ainsi associées est nulle. La somme $\Sigma[u_\nu - \mathcal{E}\{u_\nu\}]$ étant convergente, d'après le paragraphe 2, Σu_ν l'est aussi; sa limite S (x_0 tendant vers zéro) dépend de la loi L_1.

Quant aux lois dissymétriques d'exposant caractéristique 1, elles se déduisent de L_1 par addition d'une constante à S.

2° Pour chaque intervalle très petit dx, il est indifférent de dire qu'on a peut-être plusieurs valeurs de u_ν situées dans cet intervalle, leur nombre dépendant de la loi de Poisson et ayant la valeur probable (1), ou qu'il y en a au plus une, la probabilité de son existence ayant la valeur (1). Le premier point de vue est exact, même pour un intervalle fini; le second comporte une erreur de dx^2 et ne devient exact que par le passage à la limite.

Si l'on introduit la fonction caractéristique, le premier point de vue s'exprime immédiatement, dans le cas $\alpha < 1$, $\beta = -1$, par la formule

$$\log \mathcal{E}\{e^{itS}\} = \int_0^\infty (e^{itx} - 1)\,|\,dx^{-\alpha}\,|,$$

et par des formules analogues dans les autres cas. C'est de ces formules que j'ai déduit la seconde démonstration de l'existence des lois stables, citée dans l'Introduction. Elles montrent bien la définition constructive de ces lois qui font partie d'un groupe dont la loi de Poisson, si $\alpha < 1$, est le seul élément fondamental; si $\alpha \geqq 1$ il faut introduire des constantes additives.

Ce qui précède constitue une démonstration, indépendante de la notion de fonction caractéristique, du fait que ce procédé de définition donne bien des lois stables.

3° Les raisonnements qui précèdent s'étendent sans difficulté à l'étude des lois semi-stables, obtenues en remplaçant la probabilité (1) par

$$\varphi(\log x)\,|\,dx^{-\alpha}\,|,$$

$\varphi(z)$ étant une fonction périodique, de période $\log q$. On ne retrouve la même probabilité, à un facteur constant près, par le changement de x en cx, que si c est une puissance de q, d'exposant entier. Alors la somme de p variables indépendantes les unes des autres et dépendant d'une telle loi dépendra sensiblement d'une loi du même type si p est voisin d'une puissance de q^α, et cela d'autant plus exactement que $\log p - \alpha h \log q$ (h désignant l'exposant de cette puissance) est plus petit. Quel que soit q, il existe donc des entiers p réalisant cette condition avec autant de précision que l'on veut.

On peut généraliser un peu, et remplacer la probabilité précédente pour chaque intervalle dx par

$$x^{-\alpha}\, d\Phi(\log x),$$

la fonction Φ étant non décroissante et telle que la différence $\Phi(\log qx) - \Phi(\log x)$ soit constante. On peut ainsi obtenir des lois discontinues pour les u; mais la loi dont dépend la somme S, c'est-à-dire la loi semi-stable étudiée, est toujours continue.

4. Formation asymptotique des lois stables. — Supposons d'abord α compris entre 0 et 1 et désignons par $\mathcal{L}_{\alpha,-1}$ la loi dont dépend une variable x, toujours supéreure à 1, chaque intervalle dx, entre 1 et ∞, ayant la probabilité (1). Soient $x_1, x_2, \ldots, x_n$ des variables indépendantes les unes des autres et dépendant de cette loi. Posons

$$S_n = x_1 + x_2 + \ldots + x_n = n^{\frac{1}{\alpha}} s_n. \tag{4}$$

Nous allons montrer que la loi dont dépend s_n tend, pour n infini, vers la loi $L_{\alpha,-1}$. C'est à l'aide de cette définition asymptotique de $L_{\alpha,-1}$ que j'ai donné autrefois la première démonstration de l'existence de cette loi stable. On va voir que cette démonstration peut, elle aussi, être rendue indépendante de la notion de fonction caractéristique.

La variable aléatoire s_n est une somme de n termes indépendants, qui varient de $x_0 = n^{-\frac{1}{\alpha}}$ à l'infini suivant la même loi que pour les n' termes non nuls u_ν considérés au paragraphe **1** et dont $\overline{S}$ est la somme. La seule différence entre $\overline{S}$ et s_n est donc que n', au lieu d'être égal à n, dépend de la loi de Poisson et a pour valeur probable n. Or on a

$$\sigma^2\{n'\} = \mathcal{E}\{(n'-n)^2\} = n,$$

de sorte que $|n'-n|$ est de l'ordre de grandeur de $\sqrt{n}$, et par suite l'erreur relative commise en confondant n et n' est infiniment petite pour n infini (donc pour x_0 infiniment petit) sauf dans des cas de probabilité infiniment petite [1].

Or, nous savons que, si nous multiplions par $C = c^\alpha$ le nombre de termes dont $\overline{S}$ est la somme, la variable S obtenue à la limite se trouve

---

[1] Ainsi, par exemple, $\mathfrak{L}\left\{|n'-n| \geqq n^{\frac{3}{4}}\right\} \leqq \frac{1}{\sqrt{n}}$. Il est d'ailleurs facile d'obtenir des résultats bien plus précis, car si l'on pose $n' - n = \xi\sqrt{n}$, ξ dépend d'une loi tendant vers celle de Gauss.

multipliée par c. Comme on ne peut qu'augmenter S_n en augmentant le nombre de termes, le remplacement de n par n', c'est-à-dire, quelque petit que soit ε, sa multiplication par un facteur compris à partir d'un certain moment entre $e^{-\alpha\varepsilon}$ et $e^{\alpha\varepsilon}$, ne peut que multiplier S par un facteur compris entre $e^{-\varepsilon}$ et e^{ε}. Il en résulte bien que la loi dont dépend s_n se confond à la limite avec celle dont dépend $\overline{S}$, et tend vers $L_{\alpha,-1}$.

C. Q. F. D.

Bien entendu, tandis qu'au paragraphe **1** $\overline{S}$ tendait vers S, ici la suite des s_n n'a pas de limite; il s'agit seulement de la loi dont dépend s_n, qui tend vers une limite.

L'extension du procédé précédent au cas des lois stables quelconques (la loi de Gauss étant seule exclue) se fait sans difficulté. Si $|\beta| < 1$ il faut, dans la définition de $\mathcal{L}_{\alpha,\beta}$, considérer des valeurs de x des deux signes. Si $\alpha > 1$ il faut, dans la définition de s_n, remplacer x_ν par $x_\nu - \mathcal{E}\{x_\nu\}$. On traite aussi sans difficulté le cas des lois semi-stables.

CHAPITRE II.

ÉTUDE DES SOMMES DE VARIABLES ALÉATOIRES DÉPENDANT DE LA MÊME LOI DE PROBABILITÉ.

5. Condition nécessaire et suffisante pour l'application du théorème limite sous la forme classique. — Conservons les notations de la formule (4), α étant maintenant égal à 2 et la loi $\mathcal{L}$ dont dépendent les variables x étant quelconque; nous supposons seulement, si $\mathcal{E}\{|x|\}$ est fini, qu'on ait annulé $\mathcal{E}\{x\}$ par un déplacement de l'origine.

Théorème I. — *La condition nécessaire et suffisante pour que s_n dépende d'une loi tendant, pour n infini, vers celle de Gauss* [1] *est que* $\mathcal{E}\{|x^2|\} = 1$.

Le fait que cette condition soit suffisante est classique; de même $\mathcal{E}\{x^2\}$ ne doit pas avoir une valeur finie autre que 1. Il reste à montrer

(1) Il est entendu que nous parlons de la loi de Gauss réduite. Si ξ dépend de cette loi, on a $\mathcal{E}\{\xi\} = 0$, $\mathcal{E}\{\xi^2\} = 1$.

que, si la loi dont dépend s_n tend vers celle de Gauss, $\mathcal{E}\{x^2\}$ ne peut pas être infini.

Posons $x_n = x'_n + x''_n$, x'_n étant égal à x_n si $|x_n|$ ne dépasse pas une certaine valeur X et à zéro dans le cas contraire. La probabilité ε de cette seconde circonstance tend vers zéro pour X infini; nous pouvons donc la supposer arbitrairement petite. Posons

$$S'_n = \sum_1^n x'_\nu, \qquad S''_n = \sum_1^n x''_\nu, \qquad m^2 = \sigma^2\{x'\} \quad (^1).$$

Le nombre des termes non nuls de la somme S'_n est $(1-\varepsilon)n$, avec une erreur dont la valeur quadratique moyenne est $\sqrt{\varepsilon(1-\varepsilon)n}$, et qui par suite peut être négligée à côté de n, sauf dans les cas de probabilité arbitrairement petite que l'on peut aussi négliger. Il en résulte, en posant

$$S'_n = s'_n m \sqrt{(1-\varepsilon)n}, \qquad S''_n = s''_n m \sqrt{(1-\varepsilon)n},$$

que s'_n dépend d'une loi tendant pour n infini vers celle de Gauss.

D'ailleurs, à la limite, s'_n est presque indépendant de s''_n; chacun des x'_ν non nul est en effet choisi suivant une loi déterminée entre $-X$ et $+X$; chacun des x''_ν non nul est de même choisi suivant une loi déterminée à l'extérieur de l'intervalle $(-X, +X)$. Quant au nombre de termes de chacune des deux sommes, nous avons vu que, s'il dépend du hasard, ses variations possibles n'ont qu'une influence négligeable sur s'_n qui est bien ainsi à la limite une variable aléatoire indépendante de s''_n (²). Le principe de l'augmentation de la dispersion s'applique alors à l'étude de la somme

$$s_n = m\sqrt{1-\varepsilon}(s'_n + s''_n).$$

(¹) Tous les x'_ν dépendant de la même loi, nous désignons par x' une variable quelconque dépendant de cette loi.

(²) En termes précis, l'erreur relative commise en assimilant s'_n à une telle variable tend vers zéro pour n infini. On peut aussi justifier la conclusion du texte en appliquant le principe de l'augmentation de la dispersion, étendu aux variables enchaînées, sous la forme que j'ai indiquée dans un travail récent (*Bull. Sciences math.*, 1935).

En liant donc les nombres α et h par la formule

$$\alpha = \frac{1}{\sqrt{2\pi}} \int_{-h}^{+h} e^{-\frac{x^2}{2}} dx = \varphi(h),$$

la concentration relative à la variable aléatoire s'_n et à l'intervalle $2h$ tendant vers α, on a *a fortiori* pour la variable s_n, pour n assez grand

$$\mathfrak{T}\{|s_n| \leqq hm\sqrt{1-\varepsilon}\} < \alpha + \varepsilon' = \varphi(h) + \varepsilon',$$

ε' étant, comme ε, un nombre arbitrairement petit. Or par hypothèse le premier membre tend vers $\varphi(hm\sqrt{1-\varepsilon})$; $\varphi(h)$ étant une fonction croissante, et ε et ε' étant aussi petits qu'on veut, on a $m \leqq 1$, et par suite

$$\sigma^2\{x\} = \lim_{X \to \infty} m^2 \leqq 1,$$

de sorte que $\mathcal{E}\{x^2\}$ est nécessairement fini.

C. Q. F. D.

6. L'APPARITION PROGRESSIVE DES GRANDES VALEURS DES x_ν DANS L'ÉTUDE DE S_n. — Considérons d'abord un exemple simple : supposons que la loi $\mathcal{L}$ soit une loi symétrique, $|x|$ ayant comme valeurs possibles 1 et X, la probabilité de cette dernière valeur étant ε. Nous supposerons X très grand, et ε très petit; pour fixer les idées lions-les par la relation

$$\mathcal{E}\{x^2\} = 1 - \varepsilon + \varepsilon X^2 = 4.$$

Lorsqu'on étudie les premières sommes S_n, tant que le produit $n\varepsilon = \eta$ qui borne supérieurement la probabilité de la réalisation de la valeur X pour l'un au moins des termes $|x_\nu|$ est petit, tout se passe comme dans le jeu de pile ou face classique : le processus qui fait que $\frac{S_n}{\sqrt{n}}$ dépend d'une loi tendant vers celle de Gauss commence par fonctionner, et si ε est assez petit, il existe des valeurs de n pour lesquelles on obtient ainsi des lois très voisines de celle de Gauss. Mais, n croissant, l'influence de la valeur possible X cesse d'être négligeable. Au moment où par exemple η, sans être négligeable, est encore assez petit pour que l'on puisse négliger son carré, la loi dont dépend s_n différera peu de la loi continue pour laquelle la densité de probabilité est

$$\frac{1-\eta}{2}\varphi'(s)+\frac{\eta}{4}\varphi'(s+h)+\frac{\eta}{4}\varphi'(s-h), \qquad \left(h^2=\frac{3}{\eta}\right).$$

Cette densité a un maximum absolu pour $x=0$, et deux autres maxima d'abscisses voisines respectivement de $-h$ et $+h$ et d'ordonnées non négligeables. Quand n croît, le nombre de réalisations des valeurs $\pm X$ auquel il faut s'attendre augmente avec n. Le processus de convergence vers la loi de Gauss recommence alors, après une période intermédiaire pendant laquelle on aura eu des lois nettement différentes, et à une échelle double, puisque ce n'est pas $\frac{S_n}{\sqrt{n}}$, mais $\frac{S_n}{2\sqrt{n}}$, qui dépend d'une loi tendant finalement vers celle de Gauss.

Au lieu d'une seule grande valeur possible pour $|x|$, on peut en considérer plusieurs ou même une infinité de valeurs indéfiniment croissantes qui interviendront successivement dans l'étude de S_n. On sait que, si $\mathcal{E}\{x^2\}=m^2$ est fini, $\frac{S_n}{\sqrt{n}}$ dépend d'une loi tendant vers celle de Gauss. Le processus qui conduit à ce résultat sera seulement retardé si les grandes valeurs de x interviennent dans le calcul de m^2 pour une part non négligeable; alors elles n'interviennent dans S_n que tardivement, et pendant longtemps, leur intervention étant très peu probable, l'ordre de grandeur auquel il faut s'attendre pour S_n n'est qu'une fraction de celui indiqué par la formule limite.

Si au contraire $\mathcal{E}\{x^2\}$ est infini, le processus de convergence vers la loi de Gauss est constamment modifié par suite de l'apparition des grandes valeurs. D'après le paragraphe 5, S_n croît plus vite que $\sqrt{n}$; en termes précis, quelque grand que soit k, $\mathcal{Q}\{|S_n|\leq k\sqrt{n}\}$ tend vers zéro. Il est d'ailleurs possible que $\frac{S_n}{N}$, N étant alors un facteur croissant plus vite que $\sqrt{n}$, ou une expression de la forme $\left(\frac{S_n}{N}\right)-a_n$, dépendant d'une loi tendant vers celle de Gauss; nous exprimons ce fait en disant que la somme S_n, *après une réduction convenable*, dépend d'une telle loi, ou encore que S_n est d'un *type* tendant vers celui de Gauss. Il en sera ainsi, d'après le théorème fondamental du calcul des probabilités, si chaque fois qu'une grande valeur de x apparaît dans S_n, sa probabilité cessant d'être négligeable, elle est petite par rapport à la somme des autres; si, en d'autres termes, le plus grand terme de S_n est négligeable à côté de cette somme. Nous dirons dans ce cas que *la loi des grands nombres s'applique* (sous sa forme généralisée, le cas classique étant celui où N est proportionnel à $\sqrt{n}$).

Naturellement il peut arriver, suivant la rapidité et la régularité de l'apparition des grandes valeurs de x' (¹), que cette circonstance se produise pour toutes les valeurs très grandes de n, ou bien pour certaines valeurs indéfiniment croissantes, ou ne se produise jamais. Dans le second cas, on aura évidemment une convergence *intermittente* vers le type de la loi de Gauss. Dans le troisième cas, on peut trouver une loi limite, qui sera nécessairement une loi stable, ou ne trouver aucune loi limite (comme dans le cas des lois semi-stables). Le processus de formation des lois stables ou semi-stables, exposé au Chapitre I, donne une idée nette de l'effet produit par l'apparition des grandes valeurs, dans le cas simple où leur probabilité F(X) décroît comme $x^{-\alpha}$ $(0 < \alpha < 2)$.

7. Condition nécessaire pour la convergence vers la loi de Gauss. — Commençons par une remarque sur le *coefficient de réduction* N, tel que $\left(\frac{S_n}{N}\right) - a_n$ dépende d'une loi tendant vers celle de Gauss. Posons $N = \lambda(n)\sqrt{n}$, et introduisons un facteur positif constant k dont nous supposerons d'abord que c'est un nombre entier; S_{kn} est la somme de k termes indépendants de la forme S_n; on peut alors prendre comme coefficient de réduction correspondant à cette somme la valeur $N\sqrt{k} = \lambda(n)\sqrt{kn}$; on peut évidemment aussi prendre $\lambda(kn)\sqrt{kn}$, de sorte que, pour n infini, $\lambda(n)$ et $\lambda(kn)$ sont des infiniment grands équivalents. Ce résultat s'étend évidemment au cas où k est rationnel, puis à celui où k est un nombre positif quelconque. Par suite $\lambda(n)$ est une fonction à croissance lente, de la forme n^{ε}, $\varepsilon = \frac{\log \lambda(n)}{\log n}$ tendant vers zéro. D'une manière plus précise, en remplaçant au besoin $\lambda(n)$

(¹) Cette rapidité et cette régularité dépendront naturellement de la rapidité et de la régularité de la croissance de X en fonction $Y = \mathfrak{P}\{|x| > X\}$; c'est en effet de Y que dépend, pour chaque valeur de n, la probabilité de l'existence d'un terme de S_n de module $> X$.

par une fonction équivalente, on peut supposer $\lambda(n)$ défini même pour n non entier et tel que $\frac{d \log \lambda(n)}{d \log n}$ tende vers zéro; c'est une conséquence évidente de ce qu'une variation non très petite de $\log \lambda(n)$ n'est possible que si la variation de $\log n$ est très grande.

Posons maintenant

$$\mathfrak{L}\{|x| > \mathrm{X}\} = \mathrm{F}(x) = \mathrm{Y}. \tag{5}$$

Nous allons montrer que :

Lemme I. — *Une condition nécessaire pour la convergence vers la loi de Gauss est que* $\mathrm{X}^\alpha \mathrm{Y}$ *tende vers zéro pour tout* $\alpha < 2$.

La démonstration de ce lemme repose sur le lemme suivant :

Lemme II. — *Si* n *et* X *augmentent indéfiniment, et si* $n\mathrm{Y} = \eta$ *reste fini,* S'_n *et* S''_n *peuvent être considérés à la limite comme des variables aléatoires indépendantes* [1].

Soit p le nombre de termes non nuls de S''_n; sa valeur probable est η; les grandes valeurs de p sont donc très peu probables [2]. Or S'_n ne peut dépendre de S''_n que par l'intermédiaire de p; c'est en effet une somme de $n-p$ termes dépendant de la loi de probabilité déduite de la loi $\mathfrak{L}$ en multipliant par $\frac{1}{(1-\mathrm{Y})}$ les probabilités des valeurs de x comprises entre $-\mathrm{X}$ et $+\mathrm{X}$, et annulant celles des valeurs extérieures à cet intervalle. Si donc n est assez grand, l'erreur relative commise en considérant S'_n comme une somme de n termes de cette nature (au lieu de $n-p$), est très petite, sauf dans des cas très peu probables. A cette erreur près, S'_n, peut donc être assimilé à une variable aléatoire indépendante de S''_n, ce qui démontre le lemme II.

---

[1] En termes précis : en négligeant des cas de probabilité infiniment petite, S_n peut, avec une erreur relative aussi petite que l'on veut, être assimilé à une variable aléatoire indépendante de S_n. Ce résultat a déjà été établi paragraphe 5 dans le cas où X est constant.

[2] La probabilité $\mathfrak{L}\{p \geqq h\}$ est évidemment au plus $\frac{\eta}{h}$. La loi de Poisson donne une évaluation précise; elle est, pour h très grand, équivalente à $\frac{e^{-\eta}\eta^h}{h!}$.

Soit alors $\alpha < \beta < 2$; si $X^\alpha Y$ ne tend pas vers zéro, on peut trouver des valeurs de X aussi grandes qu'on veut et telles que $Y > X^{-\beta}$; X étant choisi, prenons $n \sim \frac{1}{Y}$ (donc $nY = \eta \sim 1$). La somme S''_n peut alors être nulle, ou supérieure à X en valeur absolue, aucune de ces circonstances n'ayant une probabilité très petite $\left(\text{la plus petite de ces probabilités tend, pour } n \text{ infini, vers } \frac{1}{e}\right)$. La dispersion de la variable aléatoire S''_n $\left(\text{pour une probabilité supérieure à } 1 - \frac{1}{e}\right)$ est donc supérieure à X, donc à $n^{\frac{1}{\beta}}$. D'après le lemme II, et le principe de l'accroissement de la dispersion, il en est de même de celle de S_n. Or, dans le cas de convergence vers la loi de Gauss, nous avons vu que la dispersion de S_n est proportionnelle à

$$N = \lambda(n)\sqrt{n} = n^{\varepsilon + \frac{1}{2}},$$

ε tendant vers zéro. Cette hypothèse n'est donc pas compatible avec la précédente, ce qui démontre le lemme I.

La propriété établie pour N et le lemme I s'étendent, sans autre changement que celui des exposants 2 et $\frac{1}{2}$, au cas de convergence vers des lois stables d'exposants caractéristiques inférieurs à 2. Les considérations des prochains paragraphes, reposant essentiellement sur la loi des grands nombres, ne seront naturellement pas sucesptibles d'être étendues aux lois stables autres que la loi de Gauss.

8. Condition nécessaire et suffisante pour l'application de la loi des grands nombres. — Le cas où $\mathcal{E}\{x^2\}$ est fini étant résolu, nous supposerons cette expression infinie. Alors x', x'_ν, S''_n étant toujours définis comme au paragraphe **5**, $\mathcal{E}\{x'^2\} = Z$ augmente indéfiniment avec X, et $(\mathcal{E}\{x'\})^2$, fini ou infini, est sûrement négligeable à côté de l'expression précédente [1], de sorte que $\sigma^2\{x'\}$ et Z sont des infiniment grands équivalents. En posant

$$\mathcal{P}\{|x| > X\} = F(X) = Y, \tag{5}$$

$$\sigma'^2_n = \sigma^2\{S'_n\} = n\sigma^2\{x'\} \sim nZ, \tag{6}$$

---

[1] Désignons en effet par ε la probabilité de l'inégalité $X' < |x|$, et par $\mathcal{E}'\{x'\}$ et $\mathcal{E}'\{x'^2\}$ les parties des sommes $\mathcal{E}\{x'\}$ et $\mathcal{E}\{x'^2\}$ correspondant aux

Z peut être considéré comme une fonction de Y, bien définie et continue même si pour certaines valeurs de X ces quantités Y et Z ont en même temps une variation brusque. En faisant dans ces cas une interpolation linéaire, on a toujours

$$X^2 = -\frac{dZ}{dY}. \tag{7}$$

La probabilité que S_n comprenne au moins un terme de module supérieur à X et soit par suite différent de S'_n est plus égale à $nY = \eta$, et n'en diffère que par des infiniment petits d'ordres supérieurs, de sorte qu'elle est négligeable en même temps que η. Pour qu'on puisse, en assimilant S_n et S'_n, écrire que

$$\frac{S_n - \mathcal{E}\{S'_n\}}{\sigma\{S'_n\}} = \frac{S_n}{\sqrt{n}\,\sigma\{x'\}} - \sqrt{n}\frac{\mathcal{E}\{x'\}}{\sigma\{x'\}} \tag{8}$$

dépend d'une loi peu différente de celle de Gauss, il suffit donc que d'une part η soit petit et que, d'autre part,

$$\frac{X^2}{\sigma^2\{S''_n\}} \sim \frac{X^2}{nZ} = \eta' \tag{9}$$

soit petit, condition qui permet de considérer S'_n comme une somme de termes très petits. Compte tenu de $nY = \eta$, ces conditions impliquent que

$$\eta\eta' = \frac{X^2 Y}{Z} = \frac{d \log Z}{d|\log Y|} \qquad (|\log Y| = -\log Y), \tag{10}$$

soit petit.

Nous allons démontrer que :

THÉORÈME II. — *La condition nécessaire et suffisante pour l'application de la loi des grands nombres est que l'expression* (10) *tende vers zéro pour* X *infini.*

---

valeurs de $|x'|$ supérieures à X'. L'inégalité de Schwarz donne

$$(\mathcal{E}'\{x'\}^2) \leqq \mathfrak{P}\{X' < |x'|\}\,\mathcal{E}'\{x'^2\} < \varepsilon\,\mathcal{E}\{x'^2\},$$

et le résultat énoncé se déduit de ce que, pour X' assez grand, mais fixe, et X indéfiniment croissant, ε est aussi petit qu'on veut, tandis que $\mathcal{E}'\{x'\}$ représente $\mathcal{E}\{x'\}$ avec une erreur finie.

THÉORÈME III. — *La condition nécessaire et suffisante pour que la loi des grands nombres s'applique à l'étude de* S_n *pour une suite de valeurs de* n *indéfiniment croissantes est que l'expression* (10) *n'admette aucune borne inférieure positive.*

Ces deux théorèmes correspondent aux deux cas de convergence simple, ou de convergence intermittente, vers la loi de Gauss. Nous reviendrons sur cette question au paragraphe **10**. Démontrons maintenant les résultats énoncés.

1° S'il existe des valeurs de X indéfiniment croissantes pour lesquelles l'expression (10) tende vers zéro, on peut la considérer comme le produit de deux facteurs η et η' tous deux arbitrairement petits et tels que $\frac{\eta}{Y} = n$ soit entier. On obtient ainsi une suite d'entiers n tels que la loi des grands nombres s'applique à S_n pour ces valeurs de n.

2° S'il existe des valeurs de X indéfiniment croissantes pour lesquelles l'expression (10) reste supérieure à un nombre fixe qu'on peut représenter par $2\varepsilon\varepsilon'$, en prenant pour n un entier compris entre $\frac{\varepsilon}{Y}$ et $\frac{2\varepsilon}{Y}$, on obtient une succession de valeurs de n, indéfiniment croissantes, et associées aux valeurs considérées de X de manière que les expressions nY et $\frac{X^2}{nZ}$ soient respectivement supérieures à ε et ε'. On peut d'ailleurs supposer ε, et par suite $\eta = nY$, petits. Alors, dans des cas de probabilité supérieure au nombre fixe ε, ou voisine de ce nombre, S_n diffère de S'_n par un terme ou un petit nombre de termes, supérieurs en valeur absolue à X, c'est-à-dire à une fraction non très petite de $\sigma\{S'_n\}$. On ne peut donc pas, à côté de S'_n, négliger ces quelques grands termes, et il ne saurait être question d'appliquer la loi des grands nombres.

3° Le théorème III est maintenant démontré; pour démontrer le théorème II, il reste à démontrer que, l'expression (10) tendant vers zéro, on peut appliquer la loi des grands nombres à S_n, non seulement comme dans le cas 1° ci-dessus, à certaines valeurs indéfiniment croissantes, mais à toutes les valeurs assez grandes de l'entier n. A cet effet, prenons pour η un nombre fixe arbitrairement petit, et associons à chaque entier n le nombre X tel que $nY = \eta$, ce qui est sûrement possible si la loi $\mathcal{L}$ est continue. Alors η' tend vers zéro avec $\frac{1}{n}$. Les deux nombres η et η' peuvent donc être rendus tous les deux arbitrairement petits, pour tout n assez grand, et la loi des grands nombres s'applique.

Le cas d'une discontinuité de la loi $\mathcal{L}$ ne soulève d'ailleurs aucune difficulté essentielle. Si par exemple la valeur X_0 de $|x|$ a une probabilité α, on peut, en posant $\alpha = \alpha' + \alpha''$, et en prenant $X = X_0$, considérer X_0 comme appartenant à l'intervalle $(-X, +X)$ dans des cas de probabilité α', et à l'intervalle extérieur dans des cas de probabilité α''. En choisissant convenablement α' entre o et α, on peut donner à Y n'importe quelle valeur entre $F(X_0 - \varepsilon)$ et $F(X_0 + \varepsilon)$. On peut aussi rendre la loi $\mathcal{L}$ continue par une modification très petite, et sans effet sur le calcul de $\mathcal{E}\{x'\}$ quand X n'est pas très voisin de X_0. L'erreur qui en résulte sur S_n est au plus de l'ordre du produit de $\sqrt{n}$ par un nombre très petit, et peut être négligée. La démonstration des théorèmes II et III est ainsi complète.

9. Application des théorèmes II et III. — Nous allons montrer que :

Théorème IV. — 1° *La condition que* $\mathcal{E}\{|x|^\alpha\}$ *soit fini pour tout exposant* α *compris entre* o *et* 2 *est nécessaire pour qu'il y ait convergence vers la loi de Gauss, et par suite pour que l'expression* (10) *tende vers zéro;* 2° *elle n'est pas suffisante;* 3° *la même condition est suffisante pour que l'expression* (10) *admette zéro comme limite inférieure (pour* X *infini) et que par suite il y ait au moins convergence intermittente vers la loi de Gauss;* 4° *elle n'est pas nécessaire.*

Le 1° est un corollaire évident du lemme I. Le 2° et le 4° résulteront dans la suite de ce paragraphe de deux exemples de convergence intermittente vers la loi de Gauss, dont l'un vérifie la condition indiquée, mais non l'autre.

Pour démontrer le 3°, supposons qu'il existe une constante positive a telle que

$$\frac{X^2 Y}{Z} = \frac{d \log Z}{|d \log Y|} > a,$$

et par suite, pour tout $X > X_0$,

$$Z > cY^{-a} \qquad (c = Z_0 Y_0^a),$$

et aussi, pour $\alpha = \frac{2}{(1+a)} < 2$,

$$X^\alpha |dY| > \left(\frac{aZ}{Y}\right)^{\frac{\alpha}{2}} |dY| > \frac{\lambda |dY|}{Y} \qquad \left[\lambda = (ac)^{\frac{\alpha}{2}}\right].$$

Il en résulte que, dans l'hypothèse indiquée, $\mathcal{E}\{|x|^\alpha\}$ est infini pour une valeur de α inférieure à 2. C. Q. F. D.

On voit que, pour les lois d'allure assez régulière pour que les expressions $\frac{X^2 Y}{Z}$ et $\frac{|\log Y|}{\log X}$ tendent respectivement vers des limites déterminées, la loi des grands nombres s'applique sous la forme classique, sous la forme généralisée, ou ne s'applique pas, suivant que cette dernière limite est supérieure, égale, ou inférieure à 2. On voit d'ailleurs aisément que des irrégularités de la fonction $F(X)$ que l'on pourrait faire disparaître en faisant une erreur finie sur $\log X$ seraient sans importance, et que pour former des exemples de convergence intermittente il faut des irrégularités très étendues sur l'axe des X.

Supposons, pour fixer les idées, que la loi $\mathcal{L}$ soit symétrique, et qu'en dehors d'un intervalle fini, $|x|$ n'ait comme valeurs possibles que les valeurs

$$(11) \qquad x_p = q^{p^2} \qquad (q > 1;\ p = p_0, p_0 + 1, \ldots),$$

chacune de ces valeurs ayant la probabilité

$$(12) \qquad \alpha_p = \lambda a^p x_p^{-2} \qquad (\lambda > 0,\ a > 1).$$

Alors, pour p très grand et X compris entre x_{p-1} et x_p, on a

$$(13) \qquad Y \sim \alpha_p, \qquad Z \sim \frac{\lambda a^p}{a-1}, \qquad \frac{Z}{Y} \sim \frac{x_p^2}{a-1},$$

de sorte que, quand X croît de x_{p-1} à x_p, $a\,\frac{X^2 Y}{Z}$ croît depuis q^{1-2p} jusqu'à 1; on est bien dans le cas de convergence intermittente vers la loi de Gauss [1]. Posons

$$n_p = a^{-p} x_p^2 = a^{-p} q^{2p^2}.$$

---

[1] En attendant la démonstration du théorème du paragraphe **10**, il faut entendre par là : le cas où la loi des grands nombres s'applique d'une manière intermittente.

Si n est grand par rapport à n_{p-1} mais petit par rapport à n_p, la valeur x_p a encore une probabilité négligeable, et $\frac{S_n}{N}$ $(N = \sqrt{nZ})$ dépend d'une loi peu différente de celle de Gauss (x_{p-1} étant négligeable à côté de N). Quand n atteint l'ordre de grandeur de n_p, le processus de convergence est modifié par l'apparition de la grande valeur x_p, qui est de l'ordre de grandeur atteint à ce moment par S'_n. On est dans les conditions indiquées paragraphe **6**, et l'on ne retrouvera des lois du type de celle de Gauss que quand n deviendra grand par rapport à n_p; les mêmes circonstances se reproduisent indéfiniment.

Si maintenant, conservant l'expression (12) de x_p, nous prenons

$$\alpha_p = \frac{\lambda a^p}{x_p}, \tag{12 bis}$$

les formules (13) sont remplacées par les suivantes :

$$Y \sim a_p, \qquad Z \sim \lambda a^{p-1} x_{p-1}, \qquad \frac{Z}{Y} \sim \frac{1}{a} x_{p-1} x_p, \tag{13 bis}$$

et $a \frac{X^2 Y}{Z}$ varie de $\frac{x_{p-1}}{x_p}$ à $\frac{x_p}{x_{p-1}}$, c'est-à-dire d'une valeur très petite à une valeur très grande. Nous avons le même processus de convergence intermittente; il y a toutefois une différence avec le cas précédent : quand la valeur x_p cesse d'avoir une probabilité négligeable, elle est encore grande par rapport à la somme des termes plus petits; quand le nombre des termes égaux à $\pm x_p$ deviendra grand, leur somme sera prépondérante, la somme des termes plus petits étant négligeable. Cela est lié au fait que Z soit équivalent à son plus grand terme.

On peut naturellement varier ces exemples de bien des manières. Dans ceux que nous venons d'indiquer, $\mathcal{E}\{|x|^\alpha\}$ est fini, pour le premier exemple, si $\alpha < 2$, et pour le second, seulement si $\alpha < 1$. Cette remarque complète la démonstration du théorème IV.

10. Le théorème réciproque.

Théorème V. — 1° *La tendance vers le type de Gauss n'est possible que si la loi des grands nombres est applicable* (le théorème II donne donc la condition nécessaire et suffisante pour que S_n soit asymptotiquement du type de Gauss).

2° *La tendance intermittente vers le type de Gauss n'est possible qu'en cas d'application intermittente de la loi des grands nombres* (dans les conditions du théorème III).

1° Supposons donc la loi des grands nombres non applicable. Il existe donc des valeurs de X indéfiniment croissantes et telles que, quel que soit le nombre n qu'on leur associe, on ait

$$\eta X^2 = n X^2 Y > a^2 \sigma_n'^2 \qquad (a \text{ const.}; \ \sigma_n'^2 \sim n Z). \tag{14}$$

Pour chacune de ces valeurs, définissons n de manière à donner à $nY = \eta$ une valeur fixe (à Y près) et très petite. D'après la loi des petits nombres, S''_n a o ou 1 terme non nul dans des cas dont les probabilités tendent vers $e^{-\eta}$ et $\eta e^{-\eta}$, nombres tous les deux supérieurs, par exemple, à $\frac{3\eta}{4}$. Comme $S''_n = 0$ dans le premier cas et $|S''_n| > X$ dans le second, la probabilité que S''_n soit dans un intervalle donné de longueur X (concentration pour la longueur X) est inférieure à $1 - \frac{3\eta}{4}$. D'ailleurs, d'après le lemme II, S'_n peut avec une erreur négligeable par rapport à σ'_n, donc à X (η étant petit, mais fixe), être considéré comme une variable aléatoire indépendante de S''_n. A cela près qu'il faut remplacer X par un nombre plus petit, par exemple $\frac{2X}{3}$, on peut donc appliquer le principe d'augmentation de la dispersion pour passer de S''_n à S_n. Donc, quel que soit l'intervalle d'étendue

$$\frac{2X}{3} = 2l\sigma'_n$$

que l'on considère, la probabilité que S_n soit extérieur à cet intervalle est au moins

$$\frac{3\eta}{4} > \frac{3a^2\sigma_n'^2}{4X^2} = \frac{a^2}{12l^2}.$$

Or, si S_n était d'un type très peu différent de celui de Gauss, il en serait à peu près de même de S'_n, qui ne diffère de S_n que dans des cas de probabilité η. Le coefficient de réduction, à une erreur relative près tendant vers zéro avec η, est donc $\leqq \sigma'_n$ (des valeurs très peu probables peuvent, si elles sont grandes, augmenter la valeur prévue pour $\sigma\{S'_n\}$, mais non la diminuer). La variable aléatoire $\frac{S_n}{\sigma_n}$ est donc de la forme $k\xi$, ξ dépendant d'une loi peu différente de celle de Gauss et k étant inférieur à une fonction de η qui tend vers un pour $\eta = 0$.

Or nous pouvons prendre η aussi petit que nous voulons, donc $l > \frac{a}{3\sqrt{\eta}}$ aussi grand que nous voulons. Le résultat obtenu est donc en contradiction avec le précédent, d'après lequel la probabilité des valeurs de $\frac{S_n}{\sigma'_n}$ extérieurs à n'importe quel intervalle d'étendue $2l$ tend vers zéro avec $\frac{1}{l}$ beaucoup moins rapidement que ne le comporte la loi de Gauss. Le 1° du théorème V est donc démontré.

2° Supposons que la loi des grands nombres ne soit même pas applicable par intermittence. Il existe donc un nombre a tel que l'équation (14) soit vérifiée pour tout X assez grand. Si alors la loi $\mathcal{L}$ est continue, le nombre n défini par $n\mathrm{Y} = \eta$ prend toutes les valeurs entières consécutives, et le raisonnement précédent s'appliquant à toutes les valeurs de n, il ne peut même pas y avoir de convergence intermittente vers le type de Gauss. Les discontinuités de la loi $\mathcal{L}$ sont d'ailleurs sans importance; le raisonnement par lequel nous avons terminé la démonstration du théorème II (§ 8, 3°) s'applique sans modification. Le théorème V est donc complètement démontré.

11. Le classement des termes de S_n par ordre de grandeurs décroissantes. — Désignons par $\xi_1, \xi_2, \ldots, \xi_n$ les nombres $x_1, x_2, \ldots, x_n$ rangés par ordre de modules décroissants. Le rôle que joue dans ce qui précède le plus grand des $|x_\nu|$ fait penser qu'il y a intérêt à étudier les variables aléatoires ξ_p, et considérer S_n comme leur somme.

Posons $y_p = \mathrm{F}\{|\xi_p|\}$. Les nombres $y_1, y_2, \ldots, y_n$ sont des variables aléatoires, choisies d'abord au hasard entre 0 et 1, puis rangées par ordre de grandeurs croissantes. On a évidemment

$$\mathfrak{P}\{y < y_p < y + dy\} = p\,C_n^p\, y^{p-1}(1-y)^{n-p}\,dy,$$

d'où l'on déduit aisément, à l'aide des intégrales eulériennes,

$$\mathcal{E}\{y_p\} = \frac{p}{n+1}, \qquad \sigma^2\{y_p\} = \frac{p(n-p+1)}{(n+1)^2(n+2)}.$$

Si, pour n infini, $p \sim \alpha n$, les expressions précédentes sont équivalentes respectivement à α et $\frac{\alpha(1-\alpha)}{n}$, comme cela se déduit aussi immédiatement du théorème de Bernoulli. Ce qui nous intéresse surtout est le cas où p reste fixe quand n augmente indéfiniment; on a alors

$$\mathcal{E}\{y_p\} \sim \frac{p}{n}, \qquad \sigma^2\{y_p\} \sim \frac{p}{n^2},$$
$$\mathfrak{L}\{ny_p < \eta\} \sim \frac{1}{(p-1)!}\int_0^\eta \eta^{p-1}e^{-\eta}\,d\eta,$$

de sorte que ny_p est à la limite une variable aléatoire de valeur probable p, d'écart quadratique moyen $\sqrt{p}$, l'écart réduit $\frac{ny_p - p}{\sqrt{p}}$ dépendant à la limite de la loi de Gauss (pour p très grand) [1]. Il y a d'ailleurs entre les y_p une corrélation positive : si y_p est connu, y_{p+q} doit être considéré comme la $q^{\text{ième}}$ de $n-p$ variables, choisies entre y_p et 1, et rangées par ordre de grandeur croissante. Le fait que chaque y_p diffère peu de sa valeur probable nous permet de négliger cette corrélation.

Supposons, pour fixer les idées, la loi étudiée symétrique; chaque ξ_p est ainsi une variable de module ρ_p connu en fonction de y_p, et de signe choisi au hasard. La somme S_n peut alors s'écrire

$$\pm\rho_1 \pm \rho_2 \pm \ldots \pm \rho_p \pm \ldots,$$

le nombre des termes augmentant avec n, de sorte qu'on peut la considérer comme une série. Si $\frac{\rho_p}{\rho_1}$ est d'un ordre de grandeur indépendant de n, S_n apparaît comme de l'ordre de grandeur de ρ_1 ou d'un ordre supérieur suivant que cette série a sa probabilité de convergence égale à 1 ou 0, c'est-à-dire suivant la convergence ou la divergence de la série $\Sigma\rho_p^2$. C'est dans le cas de la divergence de cette série que la loi des grands nombres s'applique.

Ainsi, pour $F(x) \sim x^{-\alpha}$, on a $\left(\text{en prenant } \frac{p}{n} \text{ comme valeur approchée de } y_p\right)$

$$\rho_p^\alpha \sim \frac{n}{p}, \qquad \frac{\rho_p}{\rho_1} \sim p^{-\frac{1}{\alpha}},$$

[1] Nous avons fait augmenter d'abord n, ensuite p. S'ils varient en même temps, la condition pour qu'on obtienne la loi de Gauss est que p et $n-p$ soient tous les deux très grands.

de sorte que la loi des grands nombres s'applique si $\Sigma p^{-\frac{2}{\alpha}}$ diverge, c'est-à-dire si $\alpha \geqq 2$, et dans ce cas seulement. On retrouve ainsi des résultats connus par un procédé intuitif qu'il nous a paru utile d'indiquer, l'application de la loi des grands nombres étant liée à la divergence de la série $\Sigma \rho_p^2$.

12. Remarques sur la loi des grands nombres. — Je voudrais insister ici sur la différence profonde entre la convergence vers la loi de Gauss, résultant de la loi des grands nombres, et la convergence vers les autres lois stables, résultant d'hypothèses précises sur la probabilité des grandes valeurs de la variable. On n'obtient à la limite une de ces lois stables que parce que la loi initiale lui ressemble suffisamment; il faut qu'elle soit telle que l'intervention progressive des grandes valeurs s'effectue, sans irrégularité trop grande, avec une rapidité déterminée. Dans le cas de la loi de Gauss, il suffit que cette rapidité soit bornée supérieurement, et en deçà de cette limite, n'importe quelle modification des lois de probabilité composantes (même si ces lois varient d'un terme à l'autre; dans ce cas il faut seulement s'assurer que la série Σx_n ne devient pas convergente) est sans importance.

On peut observer qu'une modification portant sur un intervalle fini n'a jamais d'importance; elle peut seulement, dans le cas classique où $\mathcal{E}\{x^2\}$ est fini, modifier le coefficient de réduction. Mais, en dehors du cas classique, la somme des valeurs de x comprises entre deux limites finies a une dispersion négligeable devant S_n, et cela n'a rien d'étonnant que leur détermination exacte soit sans importance. Dans le cas classique au contraire, cette somme est du même ordre de grandeur que S_n; il est très remarquable que malgré cela une modification quelconque des lois composantes (même si elles varient avec n) soit sans influence sur le type limite.

Une différence analogue existe entre la convergence vers le type de la loi de Gauss, dans le cas où $\mathcal{E}\{x^2\}$ est infini, et la convergence vers celui d'une autre loi stable. Supposons ici essentiellement que tous les x_n (sauf peut-être dans un intervalle fini) dépendent de la même loi, symétrique pour fixer les idées, et telle que $F(X) = X^{-\alpha} (0 < \alpha \leqq 2)$. Divisons l'intervalle $(1, \infty)$ en une infinité d'intervalles séparés par les nombres $X_n = q^n$ ($q > 1$), et dans chacun de ces intervalles effectuons une modification absolument quelconque de la répartition de la probabilité. Si $\alpha = 2$, cette modification est sans effet sur le type limite, qui est celui de Gauss; seul le coefficient de réduction peut être changé, si les modifications considérées modifient la contribution de chaque intervalle partiel aux moments des deux premiers ordres; encore la modification de ce coefficient est-elle limitée; le rapport de ses valeurs extrêmes est q. Si au contraire $\alpha < 2$, les modifications relatives aux différents intervalles sont des causes perturbatrices dont les effets agissent successivement, quand n croît, et pour chacune d'elles il y aura un moment où son effet n'est pas négligeable; chacune agira individuellement. Si les modifications se reproduisent périodiquement, on aura alors la tendance vers un ensemble de lois semi-stables associées (nous entendons par là une loi semi-stable, et ses puissances, dont les fonctions caractéristiques sont les puissances de celle de la loi initiale). S'il n'y a pas de périodicité dans la cause, il n'y aura pas de probabilité dans l'effet; donc, en général, la loi dont dépend $\frac{S_n}{N}$ (avec $N^\alpha = n$) varie avec n d'une manière tout à fait irrégulière, le seul caractère constant étant l'ordre de grandeur des probabilités des grandes valeurs, qui sera toujours $X^{-\alpha}$, à un facteur près borné inférieurement et supérieurement.

Ce caractère provient de ce que, chaque intervalle (X_n, X_{n+1}) ayant une étendue finie sur l'échelle logarithmique, les modifications considérées ne changent pas l'ordre de grandeur de S_n. Des modifications portant sur des intervalles plus étendus auront des effets plus profonds que nous examinerons au paragraphe **13**.

13. Cas ou la loi des grands nombres ne s'applique pas. — Nous nous proposons de présenter ici quelques remarques montrant la variété des circonstances possibles dans ce cas. Cela est à prévoir, puisque les valeurs de X correspondant à des valeurs de plus en plus petites de Y sont autant de causes dont les effets se manifestent successivement quand n croît. Pour préciser cette remarque, il faut avoir présent à l'esprit le fait suivant : quand pendant un certain temps Y varie proportionnellement à une puissance de $\frac{1}{X}$ d'exposant $\alpha < 2$ (avec une proportion déterminée entre les valeurs positives et les valeurs négatives de la variable), la convergence vers une loi stable d'exposant caractéristique α tend à se réaliser; l'effet de cette tendance sera appréciable si la variation de $\log X$ n'est pas petite depuis que la cause indiquée a commencé à se manifester; on arrivera à un type de loi très voisin du type limite si la cause subsiste assez longtemps pour que la variation de $\log X$ soit grande [1]. Les mêmes circonstances se présentent pour $\alpha \geqq 2$, avec ces différences que le signe de x importe peu, et que la tendance à la limite résultera du fonctionnement de la loi des grands nombres, quand n augmente, plus que de l'apparition de nouvelles valeurs très grandes de x.

Posons maintenant $Y = X^{-\beta}$, β étant variable; pour simplifier un peu la discussion, supposons la loi $\mathcal{L}$ symétrique, et étudions différentes circonstances possibles.

1° Il est bien clair que si β varie assez lentement, la convergence vers la loi L_β aura le temps de se manifester pour chaque valeur de β, et que $\frac{S_n}{N}$ (N étant convenablement déterminé) dépendra d'une loi infiniment voisine de la loi variable L_β (cette notation désignant la loi de Gauss si $\beta \geqq 2$). La condition nécessaire et suffisante pour qu'il en soit ainsi est que, quand $\beta < 2$, à toute variation finie de $\log X$ corresponde une variation relative de β petite par rapport à $\frac{1}{\log X}\cdot$ C'est en effet la condition pour que dans un invervalle où $\log X$ varie d'une quantité finie a, Y soit de la forme $cX^{-\alpha}$, α étant constant et c peu variable. S'il en est ainsi d'autant plus exactement que X est plus grand, c'est naturellement une propriété qui reste vraie pour des valeurs arbitrairement grandes de a, de sorte que c restera constant assez longtemps pour que l'on obtienne un type de loi très voisin de L_α.

2° Supposons maintenant seulement la variation de α assez lente pour qu'à toute variation finie de $\log X$ corresponde une variation petite de β; mais quel que soit l'exposant α, la variation de $c = YX^\alpha$ ne sera pas faible. Il y aura seulement *convergence grossière* vers L_α,

---

(1) D'après le Chapitre I, la loi stable s'obtient en effet par l'addition de variables toutes du même ordre de grandeur; les autres peuvent être négligées.

produisant cet effet que les grandes valeurs de $\frac{S_n}{N}$ ($N^\alpha = cn$) auront des probabilités de l'ordre de grandeur indiqué par cette loi. Si des variations de cette nature ont un caractère périodique, c étant (pendant assez longtemps) une fonction périodique de $\log X$, on peut en particulier, si $\alpha < 2$, obtenir des lois très peu différentes des lois semi-stables d'exposant caractéristique α.

3° Au sujet des variations plus rapides de β, observons que, Y étant monotone, on a nécessairement

$$X \log X \frac{d\beta}{dX} \geqq -\beta, \tag{14}$$

l'égalité correspondant au cas où Y reste constant. Au contraire, si β croît, rien ne limite la rapidité de sa croissance, une variation brusque étant même possible. Naturellement, si ces variations ont le caractère d'oscillations assez rapides pour que la variation relative de β soit petite, soit par rapport à $\frac{1}{\log X}$, soit par rapport à 1, les conclusions des deux cas précédents subsistent. Si au contraire β varie entre des valeurs distinctes α et α', on ne doit pas s'attendre à trouver à aucun moment de loi d'un type voisin de L_β (pour β compris entre α et α').

Ainsi, supposons que $|x|$ ait comme seules valeurs possibles les valeurs x_p définies par

$$\log x_p = \left(\frac{\alpha'}{\alpha}\right)^p \qquad (0 < \alpha < \alpha' < 2),$$

et que, de x_{p-1} à x_p, Y ait la valeur

$$Y_p = x_p^{-\alpha} = x_{p-1}^{-\alpha'}.$$

Chaque valeur x_p est alors très grande par rapport à la précédente, et, si nY_{p+1} est petit mais non nY_p, on peut négliger $x_1, x_2, \ldots, x_{p-1}$, et considérer S_n comme la somme d'un certain nombre de termes égaux à $\pm x_p$, ce nombre dépendant de la loi de Poisson et ayant pour valeur probable nY_p. Ce nombre devenant grand avant que nY_{p+1} cesse d'être petit, la loi des grands nombres s'applique à ce moment. On est donc dans un cas de convergence intermittente vers la loi de

Gauss [1], et l'on n'observe à aucun moment aucun caractère des lois stables d'exposant caractéristique compris entre α et α' (bien que β varie de α à α').

Naturellement, une variation monotone de β conduira nécessairement à une valeur limite, et à la convergence vers la loi stable correspondante, sauf si cette valeur est zéro, car il n'y a pas de loi stable d'exposant caractéristique nul. Nous allons maintenant examiner ce cas.

4° Supposons donc que α tende vers zéro. A une variation relative finie de Y correspondra donc une variation très grande de $\log X$. Les valeurs de X dont la probabilité cesse d'être négligeable quand n croît augmentent alors si vite que le plus grand terme l'emporte toujours sur la somme des autres. La loi dont dépend S_n est donc assimilable à la limite à celle dont dépend le plus grand terme, et se déduit des remarques du paragraphe **12**. Les valeurs possibles de y_1 se répartissent autour de la valeur probable $\frac{1}{n}$, une probabilité positive correspondant à une variation finie de $\log y_1$, donc à une variation très grande de $\log X$. Quel que soit N, la probabilité que S_n soit de l'ordre de grandeur de N tendra vers zéro, et S_n ne pourra qu'être très grand ou très petit par rapport à N; on pourra seulement s'arranger pour donner la probabilité que l'on veut à chacune de ces circonstances.

Il n'y a donc pas, dans ce cas, de type de loi limite, à moins de considérer qu'il y en a un lorsque, en prenant pour N une fonction de n à croissance suffisamment rapide, on constate que $\frac{S_n}{N}$ tend vers zéro. Mais cela est toujours possible, et lorsque nous disons qu'il y a un type de loi limite, cela veut dire un type de loi donnant à la variable étudiée aux moins deux valeurs possibles, finies et distinctes.

14. APPLICATION DE LA NOTION D'ENSEMBLE COMPACT. — Des fonctions

---

(1) Cela se vérifie d'ailleurs par application du théorème III; quand X varie de x_p à x_{p+1}, $\frac{X^2 Y}{Z}$ varie entre des limites équivalentes, pour p très grand, à $x_p^{\alpha-\alpha'}$ et à $\frac{x_{p+1}}{x_p}$, c'est-à-dire d'une valeur très petite à une valeur très grande.

qui varient d'une manière monotone de o à 1, et sont comprises entre deux fonctions particulières remplissant cette condition, constituent un ensemble compact. En appliquant cette remarque aux fonctions $\mathfrak{P}\{x \leq \xi\}$, et en ne considérant pas comme distinctes deux lois déduites l'une de l'autre par l'addition d'une constante à la variable x, on voit que : *des lois de probabilité pour lesquelles, pour tout* $\gamma < 1$, *la dispersion liée à la probabilité* γ *est bornée supérieurement, forment un ensemble compact.*

Pour chercher si des types de lois forment un ensemble compact, il faut d'abord réduire ces lois, par un changement d'unité, de manière qu'après réduction la dispersion pour tout $\gamma < 1$ soit bornée supérieurement; ce n'est qu'à cette condition qu'on peut considérer la valeur infinie comme ayant une probabilité nulle. D'après ce qui précède, un ensemble comportant une infinité de types de lois admet toujours un élément d'accumulation. Mais la question qui se pose est de savoir s'il en admet un autre que le type *dégénéré* pour lequel une seule valeur de la variable est possible. Or, compte tenu de la discussion faite paragraphe **13**, on voit aisément dans quel cas les types des lois dont dépendent les S_n tendent vers ce type dégénéré; c'est lorsque $\frac{\Delta \log X}{\Delta |\log Y|}$ n'est borné supérieurement que pour des intervalles $\Delta |\log Y|$ eux-mêmes bornés supérieurement. Dans tous les autres cas, on peut trouver des nombres X' arbitrairement grands et tels que, en posant $Y' = F(X') = \frac{1}{n'}$, $\log |Y'|$ soit le milieu d'un intervalle arbitrairement grand et auquel correspondent pour X des valeurs de l'ordre de grandeur de X'. Le plus grand terme de S_n a alors sa dispersion (pour γ quelconque entre o et 1) de l'ordre de grandeur de X', et la somme S_n, elle-même, ou bien est au plus de cet ordre de grandeur $\left(\text{et alors } \frac{S_{n'}}{X'} \text{ dépend d'une loi non dégénérée}\right)$, ou bien (si elle est plus grande) dépend d'une loi d'un type voisin de celui de Gauss. De toute façon, le cas de dégénérescence est exclu. Donc, *en dehors du cas étudié à la fin du paragraphe* **13** (celui où α tend vers zéro), *il y a un type d'accumulation non dégénéré*, même si l'on n'est pas dans un cas de convergence (constante ou intermittente) vers celui de Gauss.

Si d'ailleurs il y a un type d'accumulation qui ne soit pas un type stable, il y en a une infinité. Si en effet, pour une suite de valeurs n_p de n, on obtient des lois dont les types tendent vers celui d'une loi $\mathcal{L}'$, pour des valeurs n'_p telles que $\frac{n'_p}{n_p}$ ait une limite k, on obtient des lois dont les types tendent vers celui de $\mathcal{L}'^k$, en désignant ainsi la loi dont la fonction caractéristique s'obtient en élevant celle de $\mathcal{L}'$ à la puissance k. Il en résulte que $\mathcal{L}'$ appartient nécessairement au groupe des lois étudiées par MM. B. de Finetti, Kolmogoroff, et moi-même [1], pour lesquelles la loi $\mathcal{L}'^k$ est définie quelque petit que soit l'exposant positif k. Il se déduit de deux lois infinitésimales, celle de Gauss et celle de Poisson. On peut définir les lois de ce groupe, dans le cas symétrique, par la formule

$$(15) \qquad \log \mathcal{E}\{e^{izx}\} = -a\frac{z^2}{2} + \int_0^\infty (\cos zu - 1)\, d\varphi(u),$$

la fonction $\varphi(u)$ étant non décroissante, finie pour u infini, peut-être infinie pour $u=0$, mais telle que $u^2 d\varphi(u)$ soit intégrable (de 0 à 1). Dans le cas dissymétrique, on peut écrire

$$(16) \qquad \log \mathcal{E}\{e^{izx}\} = -a\frac{z^2}{2} + biz + \int_{-\infty}^{+\infty} \left(e^{izu} - 1 - \frac{izu}{1+u^2}\right) d\varphi(u),$$

le terme $\frac{-izu}{1+u^2}$ pouvant être remplacé, soit par zéro, soit par $-izu$, si dans ces conditions l'intégrale conserve un sens.

On peut d'ailleurs, au point de vue qui nous occupe, arriver à un résultat plus complet et n'utilisant pas la notion d'ensemble compact. Plaçons-nous d'abord dans le cas où S_n est de l'ordre de grandeur de son plus grand terme, et soit N un coefficient de réduction possible, c'est-à-dire que ce plus grand terme est de l'ordre de grandeur de N dans des cas de probabilité positive, et n'est d'un ordre supérieur que dans des cas très peu probables. Alors $\frac{S_n}{N}$ est une somme de termes du type étudié Chapitre I : la somme des termes ayant une valeur donnée

---

[1] *Voir* Paul Lévy, *Sur les intégrales dont les éléments sont des variables aléatoires indépendantes* (*Annali della R. Scuola Normale Superiore di Pisa*, 2e série, vol. III, 12, 1934, p. 1-30).

dépend de la loi de Poisson, et S_n dépend d'une loi du type (30), le terme gaussien $\frac{-az^2}{2}$ n'existant pas si la somme des termes u_ν petits par rapport à N peut être négligée. Si d'ailleurs S_n est d'un ordre supérieur à celui de son plus grand terme, cette somme est du type de Gauss. Dans tous les cas : on peut définir N, a, b et $\varphi(u)$ en fonction de n de manière que la loi dont dépend $\frac{S_n}{N}$ et la loi définie par la formule (16) soient infiniment peu différentes (pour n infini). Cette dernière peut être d'un type dégénéré, si le second membre se réduit au terme biz. Mais ce cas ne se produit que dans le cas étudié paragraphe **13** où α tend vers zéro, ou dans celui où α n'a pas de borne inférieure positive. *Dans tous les autres cas*, S_n *est, pour tout n assez grand, d'un type infiniment peu différent d'un type variable avec n, non dégénéré, et rentrant dans ceux définis par la formule* (16).

CHAPITRE III.

ÉTUDE DES SOMMES DE VARIABLES ALÉATOIRES NON ENCHAINÉES DÉPENDANT DE LOIS VARIANT D'UNE MANIÈRE QUELCONQUE.

15. Sur une hypothèse concernant la loi de Gauss.

LEMME (HYPOTHÉTIQUE) III. — *Si la somme $z = x + y$ de deux variables aléatoires indépendantes est du type de Gauss, il en est de même de chacun des termes.*

J'ai déjà indiqué ailleurs la vraisemblance de cette hypothèse, qui nous servira dans la suite pour des remarques accessoires. Les résultats fondamentaux seront indépendants de cette hypothèse. Présentons d'abord quelques remarques à son sujet.

1° Si le lemme est en défaut, on peut supposer que ni x ni y ne sont de la forme $a\xi + x_1$, ξ et x_1 étant indépendants et ξ dépendant de la loi de Gauss. On sait en effet que si x est de cette forme, il l'est d'une infinité de manières, dont l'une, bien déterminée, rend a maximum; alors x_1 n'admet plus de décomposition de la forme indiquée. Posant ainsi $x = a\xi + x_1$, $y = b\eta + y_1$, on obtient deux nouvelles variables x_1 et y_1, ayant la propriété indiquée, et dont la somme est du type de Gauss.

2° Si z dépend de la loi de Gauss, on peut poser

$$z = a_1 z_1 + a_2 z_2 + \ldots + a_n z_n,$$

les z_i dépendant de la même loi, et les a_i étant des coefficients (réels) et tels que $\Sigma a_i^2 = 1$. Si le lemme est en défaut, chacun des z_i peut être remplacé par $x_i + y_i$, chacun de ces termes n'étant pas gaussien, et la somme $\Sigma a_i x_i$ a la même propriété que x. Comme la loi dont dépend x n'est pas stable (puisque $\mathcal{E}\{x^2\}$ est fini et que ce n'est pas la loi de Gauss), ni même semi-stable, on obtient ainsi un groupe de lois ayant la propriété d'être des diviseurs de la loi de Gauss et qui dépend d'un nombre illimité de paramètres distincts.

Pour les autres lois stables, on est effectivement conduit de cette manière à un groupe constitué en partant d'une loi infinitésimale unique, celle de Poisson. Mais il n'y a pas d'autre loi infinitésimale que celle de Poisson, à l'aide de laquelle on ne peut pas reconstituer la loi de Gauss, et la loi de Gauss elle-même. Si donc le lemme III était faux, les diviseurs de la loi de Gauss devraient constituer un groupe ayant comme éléments constitutifs une ou plusieurs lois non infinitésimales. Du point de vue formel de la théorie des groupes, cela ne conduit à aucune contradiction. Cette hypothèse me semble toutefois peu vraisemblable, et ce sentiment, joint à celui que les conséquences du lemme sont très vraisemblables, me porte à croire à son exactitude [1].

3° Si le lemme est vrai, on peut le généraliser de la manière suivante : *x_n et y_n étant des variables devenant indépendantes pour n infini, si $x_n + y_n$ dépend d'une loi d'un type tendant vers celui de Gauss, il en est de même de chaque terme* (à moins qu'un des termes ne soit à la limite négligeable à côté de l'autre, ou du moins que cela ait lieu pour une suite partielle de valeurs de n).

---

[1] Il est en outre facile de vérifier l'exactitude d'énoncés particuliers tels que le suivant : si x dépend de la loi du jeu de pile ou face, on ne peut pas déterminer la loi dont dépend y de manière que $x + y$ dépende à un facteur constant près de la loi de Gauss.

Par hypothèse, il existe une variable y'_n indépendante de x_n et dépendant d'une loi infiniment peu différente de celle dont dépend y_n (¹). Alors $x_n + y'_n$ dépend d'une loi dont on peut supposer, par une réduction convenable, qu'elle tend vers la loi de Gauss réduite. Les lois dont dépendent les x_n ont alors leur dispersion bornée par celle de la loi de Gauss; elles forment un ensemble compact. Si les x_n ne deviennent pas très petits, et s'ils ne dépendent pas d'une loi tendant vers le type de Gauss, il y a nécessairement un type d'accumulation différent. Le lemme montrant que cela n'est pas possible, s'il est vrai, la généralisation indiquée est vraie aussi.

S'il en est ainsi, l'application de ce lemme à la somme $S'_n + S''_n$ donne une démonstration simplifiée du théorème V; il n'y a qu'à remarquer que la variable S''_n, nulle dans des cas de probabilité peu différente de $e^{-\gamma}$, est d'un type très différent de celui de Gauss. Il suffirait même d'avoir démontré que la somme z ne peut pas être du type de Gauss si l'un des termes dépend d'une loi discontinue.

16. Étude des termes non individuellement négligeables. — Nous supposerons maintenant que chaque terme x_n dépende d'une loi $\mathcal{L}^{(n)}$, qui varie avec n; les x_n restent indépendants les uns des autres. On sait que la probabilité de la convergence de Σx_n, nécessairement nulle dans le cas où $\mathcal{L}^{(n)}$ ne dépend pas de n, peut ici être égale à 0 ou 1. Si elle est égale à 1 (cas de convergence), chacun des termes est du même ordre que la somme; si donc le lemme III est vrai, la somme ne peut donc être du type de Gauss que si chaque terme est de ce type; si le lemme III est faux, cette conclusion ne subsiste évidemment pas, comme le montre l'exemple d'une série $\Sigma a_n(x_n + y_n)$ analogue à la somme étudiée (§ **15**, 2°).

Plaçons-nous maintenant dans le cas de divergence; on sait que dans ce cas la dispersion de S_n augmente indéfiniment, et n'est jamais décroissante. Pour qu'un terme soit négligeable, il est donc nécessaire et suffisant qu'il le soit devant la dispersion de S_n (ou de S_{n-1}); en termes précis, quelque petit que soit ε, la dispersion de u_n pour la pro-

---

(¹) Cette indépendance approchée est celle déjà considérée dans le lemme II; il y a lieu de remarquer qu'elle ne fait pas jouer le même rôle aux deux termes.

babilité $1-\varepsilon$ devra, pour n infini, être infiniment petite par rapport à celle de S_n [1]. Si d'ailleurs, tant pour u_n que pour S_n, on est sûr que les valeurs de chaque signe ont des probabilités supérieures à un nombre fixe, cette condition se simplifie; il n'y a qu'à écrire que, pour ε' arbitrairement petit, on a

$$\lim_{n \to \infty} \mathcal{P}\{|u_n| > \varepsilon' |S_n|\} = 0. \tag{17}$$

Pour plus de généralité, prenons la première condition, et désignons par λ_n le rapport des deux dispersions (celle de u_n au numérateur). Si λ_n ne tend pas vers zéro, il existe des suites de termes non individuellement négligeables, pour lesquels λ_n reste supérieur à un nombre fixe λ. Pour que S_{n-1} et S_n soient en même temps à la limite du type de Gauss, il est évidemment nécessaire que x_n, pour les valeurs de n correspondant aux termes non négligeables, soit à la limite de ce type. Supposons cette condition vérifiée; les termes $x_\nu (\nu < n)$ non négligeables devant S_n sont en nombre fini, et tous de type tendant vers celui de Gauss, puisqu'un terme de rang fini est négligeable devant S_n et que par suite le plus petit des ν augmente indéfiniment avec n; la somme de ces termes est donc à la limite du type de Gauss.

Le problème est donc résolu si la somme de ces termes, comme dans le processus de formation des lois stables autres que celle de Gauss, définit S_n à une erreur près relativement aussi petite que l'on veut. Dans le cas contraire (que cela se produise pour toutes les valeurs de n ou seulement pour certaines valeurs de n), il est ramené à l'étude de la somme des termes individuellement négligeables.

Au sujet du rôle des grands termes, observons encore que, si le lemme III est faux, la somme $\Sigma a^n(x_n + y_n)$ (notations du paragraphe **15**, 2°, et $a > 1$) donne un exemple de convergence intermit-

---

(1) Le type limite de S_n s'obtient toujours en prenant un coefficient de réduction de l'ordre de grandeur de la dispersion de S_n pour une probabilité $\gamma = 1 - \varepsilon$ voisine de l'unité. Mais ce type ne peut être celui de Gauss que si cet ordre de grandeur est indépendant de γ. Pour la recherche des cas où le type limite est celui de Gauss, on peut donc déterminer la dispersion de S_n pour une probabilité fixe quelconque; mais celle de u_n doit toujours être calculée pour une probabilité arbitrairement voisine de l'unité.

tente avec une rapidité des oscillations entre les lois tendant vers le type de Gauss et celles qui en diffèrent, qui n'est pas possible si le lemme est vrai; dans ce cas chaque oscillation correspond nécessairement à un intervalle (n, n') tel que S_n soit une fraction négligeable de $S_{n'}$.

17. Condition d'application de la loi des grands nombres. — Nous supposons maintenant chaque terme d'indice très grand individuellement négligeable [1]. Nous introduirons un nombre X qui jouera le même rôle qu'au Chapitre II, et définirons de la même manière x'_n, x''_n, S'_n, S''_n. La condition nécessaire et suffisante pour l'application de la loi des grands nombres à S_n est toujours qu'on puisse déterminer X de manière à vérifier à la fois les deux conditions suivantes : 1° X^2 est petit par rapport à

$$\sigma'^2_n = \sigma^2\{s'_n\} = \sum_1^n \sigma^2\{x'^2_\nu\};$$

2° La probabilité de $S''_n \neq 0$, c'est-à-dire la probabilité de l'existence dans S_n de termes de modules dépassant X, est très petite.

Cette probabilité sera ici sensiblement, non nY, mais

$$\eta = \sum_1^n \mathfrak{P}\{|x_\nu| > X\} = \sum_1^n Y_\nu,$$

l'erreur étant de l'ordre de η^2. Comme il n'y a pas de termes individuellement non négligeables, les lois des petites probabilités sont toujours applicables; le nombre des termes de S''_n dépend de la loi de Poisson et a pour valeur probable η.

La condition pour l'application de la loi des grands nombres est alors en termes précis, que, quelque petits que soient ε et ε', on puisse, pour

---

(1) Si, pour tout ε, la dispersion de u_n est négligeable par rapport à celle de S_n, on peut, en ajoutant au besoin une constante à chaque u_n, supposer que u_n soit lui-même négligeable (sauf dans des cas très peu probables). Cette constante peut d'ailleurs être choisie indépendamment de ε; il suffit de la choisir de manière que chacune des inégalités $u_n \geqq 0$ et $u_n \leqq 0$ ait une probabilité $\geqq \frac{1}{2}$. Nous supposons cette réduction effectuée.

tout n assez grand, déterminer X *de manière que l'on ait à la fois*

$$X \leqq \varepsilon' \sigma'_n, \qquad \eta \leqq \varepsilon. \tag{18}$$

Si cette condition n'est vérifiée que pour une suite partielle de valeurs de n indéfiniment croissantes, la loi des grands nombres ne s'applique que d'une manière intermittente. En effet, pour toute suite de valeurs pour lesquelles, pour un système particulier de déterminations de ε et ε', on ne pourra pas déterminer X de manière à vérifier à la fois les deux inégalités (18), la loi des grands nombres n'est pas applicable, car il pourra y avoir dans S_n des termes (de rangs non connus d'avance) non négligeables devant la dispersion de cette somme.

Naturellement, si la loi des grands nombres s'applique, il y a convergence vers le type de loi de Gauss. Les principaux résultats du Chapitre II subsistent donc, à cela près que η et σ'^2_n n'étant plus proportionnels à n, il n'est pas possible en général d'éliminer cette variable et de ramener la discussion à l'étude d'une fonction simple de X. Observons que, si le nombre X associé à chaque valeur de n est indéterminé dans un intervalle assez large, le nombre σ'_n, qui est le coefficient de réduction tel que $\frac{S_n}{\sigma'_n} - a_n$ dépende d'une loi tendant vers celle de Gauss, est au contraire déterminé, du moins à une erreur relative près infiniment petite pour n infini.

18. Le théorème réciproque.

Théorème VI. — *S'il n'y a pas de termes individuellement non négligeables, on ne peut obtenir de convergence constante vers la loi de Gauss que dans le cas où la loi des grands nombres s'applique.*

Supposons en effet que, pour une suite de valeurs de n indéfiniment croissantes, on ne puisse pas déterminer X de manière à vérifier à la fois les deux inégalités (18). On peut alors le déterminer de manière à vérifier à la fois les deux inégalités inverses

$$\eta \geqq \varepsilon, \qquad \frac{X}{\sigma'_n} = \sqrt{\eta'} > \varepsilon'. \tag{19}$$

Cela est bien évident, puisque, ou bien l'on a $\eta = \varepsilon$ pour au moins une valeur de X, ou bien, quand X croît, η franchit par un saut brusque la valeur ε, et à ce moment $\frac{X}{\sigma'_n}$ varie dans le même sens que η; si donc les inégalités (18) ne sont pas vérifiées après le moment considéré, c'est que les inégalités (19) l'étaient avant. Nous pouvons donc supposer ces inégalités vérifiées.

D'ailleurs le lemme II s'applique toujours; du moment que chaque terme est individuellement négligeable, l'existence d'un terme ou d'un nombre fini de termes de S''_n, supprimant le nombre correspondant de termes de S'_n, est sans effet appréciable, et la dispersion de S_n, pour une probabilité telle que $1 - \frac{\eta}{2}$, est au moins égale à celle de S''_n, donc à $X > \varepsilon'\sigma'_n$. Si ε' est assez grand et si η n'est pas trop petit, ce résultat peut être incompatible avec l'hypothèse que $\frac{S_n}{\sigma'_n}$ dépende d'une loi tendant vers celle de Gauss, et il ne peut être question (si η est assez petit) d'un coefficient de réduction qui ne soit pas voisin de σ'_n [1]. Si l'incompatibilité n'apparaît pas ainsi, elle apparaîtra sûrement si l'on peut remplacer η et η' respectivement par $\frac{\eta}{p}$ et $p\eta'$, p étant un entier assez grand, fonction seulement de η et η' (ou de ε et ε'), mais indépendant de n [2]. Nous allons montrer que cela est possible.

Chacun des termes étant individuellement négligeable, la somme σ'^2_ν varie d'une manière presque continue avec ν, et l'on peut répartir les n premiers termes de S_n en p groupes dont chacun donnera à la somme σ'^2_n une contribution égale à $\frac{\sigma'^2_n}{p}$, et cela d'autant plus exactement que n est plus grand. Pour un au moins de ces groupes (allant du terme de rang $n'+1$ à celui de rang n''), sa contribution à la somme η sera $\geqq \frac{\eta}{p}$, de sorte que η' et η se trouvent remplacés respectivement par $p\eta'$ et par un nombre supérieur à $\frac{\eta}{p}$. On est donc sûr

---

(1) S'il était plus petit, on n'échapperait pas à la contradiction signalée; il ne peut pas être plus grand pour la raison indiquée à propos du théorème V.

(2) Car pour la loi de Gauss la probabilité correspondant à l'intervalle $(-p\eta', p\eta')$ tend vers l'unité, quand p croît, beaucoup plus rapidement que $1 - \frac{\eta}{2p}$.

que $S_{n''} - S_{n'}$ ne peut pas dépendre d'une loi d'un type tendant, pour n infini, vers celui de Gauss.

Or le groupe considéré, puisque chaque terme est individuellement négligeable, comprend un grand nombre de termes, de sorte que n'' croît indéfiniment avec n. Si $n' = 0$, le théorème est démontré. Si $n' > 0$, n' et n'' croissent tous les deux indéfiniment; d'ailleurs $S_{n''} - S_{n'}$, puisque $\sigma\{S'_{n''} - S'_{n'}\}$ diffère peu de $\frac{\sigma'_n}{\sqrt{p}}$, est une fraction non négligeable de σ'_n, donc *a fortiori* de $\sigma\{S'_{n''}\}$, et ne peut être négligé dans l'étude de $S_{n''}$. Cette expression dépendant d'une loi d'un type qui ne tend pas vers celui de Gauss, il est impossible que $S_{n'}$ et $S_{n''}$ dépendent tous les deux de lois de types tendant vers celui de Gauss [1]. Comme n' et n'' augmentent indéfiniment, nous avons établi l'impossibilité de la convergence vers le type de la loi de Gauss.

CHAPITRE IV.

ÉTUDE DES SOMMES DE VARIABLES ALÉATOIRES ENCHAINÉES.

19. Sur un principe général dans l'étude des probabilités en chaine. — Nous supposerons maintenant que la loi dont dépend x_n soit fonction de $x_1, x_2, \ldots, x_{n-1}$. Nous désignerons respectivement par $\mathfrak{T}_\nu$ et $\mathcal{E}_\nu$ des probabilités et des valeurs probables appréciées lorsque l'on connaît $x_1, x_2, \ldots, x_\nu$; $\mathfrak{T}$ et $\mathcal{E}$ correspondront à des évaluations *a priori*; les notations $\mathcal{L}^{(n)}$ et $\mathcal{L}^{(n)}_{n-1}$ désigneront alors respectivement les lois dont dépend x_n *a priori*, et au moment de l'expérience déterminant cette variable.

Soit A une propriété de la suite des x_n; soit

$$(20) \qquad \alpha = \Phi\{\mathcal{L}_1, \mathcal{L}_2, \ldots, \mathcal{L}_n, \ldots\}$$

[1] On remarque que nous n'avons pas utilisé le lemme III; mais, s'il est exact, il permet de simplifier le raisonnement et de ne pas s'occuper de la loi dont dépend $S_{n'}$. L'extension au cas de la convergence intermittente, que nous n'avons pas traitée dans le texte, est immédiate si le lemme III est vrai, mais semble assez difficile à établir sans utiliser ce lemme.

sa probabilité, évaluée dans le cas des variables indépendantes en fonction des lois $\mathcal{L}_n$. La question suivante se pose tout naturellement si l'on passe au cas des variables en chaînes : a-t-on nécessairement

$$(21) \qquad \mathfrak{T}\{A\} = \mathcal{E}\{\alpha\} = \mathcal{E}\left\{\Phi\{\mathcal{L}^{(1)}, \mathcal{L}_1^{(2)}, \ldots, \mathcal{L}_{n-1}^{(n)}\}\right\},$$

et, dans les cas importants où α ne prend que les valeurs o et 1, peut-on dire que (A' étant la propriété contraire de A)

$$(22) \qquad \mathfrak{T}\{A, \alpha = 0\} = \mathfrak{T}\{A', \alpha = 1\} = 0 ?$$

En d'autres termes, B étant dans ce cas la propriété de la suite des lois $\mathcal{L}_n$ qui dans le cas des variables indépendantes entraîne $\mathfrak{T}\{A\} = 1$, peut-on, la propriété contraire B' entraînant $\mathfrak{T}\{A\} = 0$, dire que la liaison entre A et B subsiste dans le cas des probabilités en chaîne, l'un n'étant réalisé sans l'autre que dans des cas dont la probabilité est nulle ?

Contrairement à ce qu'on pourrait penser, la réponse n'est pas toujours affirmative. Prenons pour A la propriété suivante : il arrive une infinité de fois que $x_{n-1} = x_n = 1$; si les variables sont indépendantes, sa probabilité est toujours o ou 1, et l'on est dans l'un ou l'autre cas suivant que la série

$$\Sigma \gamma_{n+1}\gamma_n \qquad [\gamma_n = \mathfrak{T}\{x_n = 1\}]$$

est convergente ou divergente; la propriété B est la divergence de cette série. Supposons maintenant les variables enchaînées; γ_n doit maintenant désigner $\mathfrak{T}_{n-1}\{x_n = 1\}$. La liaison entre A et B subsiste-t-elle? Pour voir qu'il n'en est pas toujours ainsi, il suffit de considérer le cas d'un jeu de pile ou face, d'enjeu égal tantôt à 1, tantôt à 2, et changeant chaque fois que x_{n-1} a été positif. Il est donc impossible que deux valeurs consécutives de x_n soient égales à $+1$; A n'est jamais réalisé.

Pourtant il arrivera une infinité de fois que deux enjeux consécutifs soient égaux à 1, donc que $\gamma_{n-1}\gamma_n = \frac{1}{4}$, et B est vérifié, sauf dans des cas de probabilité nulle.

L'explication de ce paradoxe est bien simple. Désignons par E_n l'événement $x_{n-1} = x_n = 1$. Quelqu'un qui ne connaît que la succession des lois $\mathcal{L}_{n-1}^{(n)}$ croit qu'on a risqué une infinité de fois la réalisation de E_n, alors qu'il n'en est rien.

Nous pouvons, en nous inspirant de cette remarque, indiquer dans quelles conditions on peut appliquer sans risque d'erreur les formules (21) et (22). Nous supposons que la propriété A, pour une suite finie ou infinie d'expériences, exprime une propriété du nombre de réalisations d'un événement E_n positif ou nul, fini ou infini. Cet événement E_n dépend essentiellement de x_n et peut aussi dépendre des expériences antérieures. *La condition essentielle pour l'application des formules* (21) *et* (22) *est que, pour l'évaluation de* α, *on ne fasse intervenir que les probabilités* $\alpha_n = \mathcal{P}_{n-1}\{E_n\}$, *évaluées pour chaque* E_n *au moment de la* $n^{\text{ième}}$ *expérience*, sans tenir compte des évaluations antérieures qui peuvent être devenues fausses, si E_n dépend, non seulement de x_n, mais aussi de $x_1, x_2, \ldots, x_{n-1}$.

Ainsi, la probabilité de la réalisation de E_n, pour deux valeurs n et $n' > n$ de l'indice, est

$$\mathcal{E}\left\{\alpha_n \mathcal{E}_n\{\alpha_{n'}\}\right\} = \mathcal{E}\left\{\mathcal{E}_n\{\alpha_n \alpha_{n'}\}\right\} = \mathcal{E}\{\alpha_n \alpha_{n'}\},$$

et comme la probabilité de A se calcule dans tous les cas par des applications répétées de la formule des probabilités composées et de celle des probabilités totales, on voit bien que l'expression (20) de cette probabilité s'étend sous la condition indiquée ci-dessus au cas des variables enchaînées à condition seulement de remplacer l'expression obtenue

$$\varphi(\alpha_1, \alpha_2, \ldots, \alpha_n, \ldots)$$

qui est devenu une variable aléatoire, par sa valeur probable. La formule (21) est bien établie; la formule (22) s'établit aussi sans difficulté; nous n'aurons dans la suite besoin que de la première de ces formules.

20. Cas des variables enchaînées a des moyennes presque indépendantes. — Les auteurs des recherches antérieures sur les variables enchaînées ont toujours fait jouer un grand rôle à une condition de presque indépendance, que M. S. Bernstein, notamment, exprime en faisant intervenir les moments des trois premiers ordres. Dans un travail récent, cité déjà dans l'introduction, j'ai montré qu'on peut obtenir des résultats simples en ne considérant que les moments des deux premiers ordres, et en ne faisant intervenir que le premier dans la condition de presque indépendance, qui prend la forme

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

elle exprime que le jeu est toujours équitable.

Supposons d'abord les x_n bornés dans leur ensemble, et posons

$$\sigma_n^2 = \sum_1^n \mathcal{E}_{\nu-1}\{x_\nu^2\}.$$

Supposons que pour chaque chaîne, sauf dans des cas de probabilité nulle, σ_n^2 devienne infini avec n; cela revient à supposer nulle la probabilité de la convergence de Σx_n. Faisons, pour chaque chaîne, augmenter n jusqu'à ce que σ_n^2 ait une valeur donnée t; en raison de l'hypothèse que les x_n sont bornés, on peut, si t est assez grand, supposer σ_n^2 égal à t avec une erreur relative aussi petite que l'on veut. En arrêtant chaque chaîne à la valeur n ainsi obtenue, on obtient une *section à t constant; n* devient une variable aléatoire, fonction du paramètre t, et, au lieu d'étudier S_n pour chaque valeur donnée de n, il s'agit, en posant $S_n = \mathfrak{S}(t)$, d'étudier $\mathfrak{S}(t)$ pour t donné. J'ai montré que $\frac{\mathfrak{S}(t)}{\sqrt{t}}$ *dépend d'une loi qui tend, pour t infini, vers celle de Gauss.*

Supposons maintenant les x_n non bornés dans leur ensemble; donnons-nous X assez grand et définissons toujours x'_ν par les conditions $x'_\nu = 0$ ou x_ν suivant que $|x_\nu| > X$ ou non. Posons

$$\sigma_n'^2 = \sum_1^n \mathcal{E}_{\nu-1}\left\{(x_\nu - \mathcal{E}_{\nu-1}\{x_\nu\})^2\right\},$$

$$\eta_n = \sum_1^n \mathcal{P}_{\nu-1}\{|x_\nu| > X\}.$$

Pour chaque chaîne, faisons toujours augmenter n jusqu'à ce que $\sigma_n'^2$ atteigne une valeur donnée t; η_n aura alors une valeur η, et si l'on a

$$(23) \qquad \mathcal{P}\{\eta > \varepsilon\} < \gamma,$$

ε et γ étant simultanément très petits, la probabilité α de l'existence, dans la somme $\mathfrak{S}(t)$, d'au moins un terme de module dépassant X sera au plus égale à $\gamma + \varepsilon$ (au second ordre près) et pourra être négligée [1]. L'étude de $\mathcal{S}(t) = S_n$ se ramène alors de celle de $S'_n = \Sigma x'_\nu$. Si d'ailleurs il n'y a pas de terme non individuellement négligeable (et si ε et γ sont petits il ne peut y en avoir que pour des chaînes très peu probables) σ'^2_n varie d'une manière presque continue et peut être rendu égal à t avec une erreur relative d'autant plus petite que t est plus grand. L'expression section à t constant conserve donc un sens précis, et pour chaque chaîne, sauf dans des cas très peu probables, σ'_n est bien le nombre que le Chapitre III conduit à associer à la suite des lois $L^{(n)}_{n-1}$ [2]. Cela reste vrai, quand t croît, jusqu'au moment où ε et γ ne peuvent plus être pris simultanément très petits; pour aller plus loin, il faut alors prendre une valeur plus grande de X.

En posant alors $X = \eta' \sqrt{t}$, la condition pour que $\mathcal{S}(t)$ puisse ainsi être considéré comme une somme de termes très petits est que ε, η' et γ soient en même temps très petits. En termes précis, à tout t assez grand, on devra pouvoir faire correspondre une valeur X telle que ε, ε' et γ étant arbitrairement petits, on ait en même temps $\eta' < \varepsilon'$ et l'inégalité (23) [3].

Si cette condition est vérifiée, le résultat rappelé tout à l'heure s'applique à la suite des variables

$$y_\nu = x'_\nu - \mathcal{E}_{n-1}\{x'_\nu\}$$

qui vérifient la condition $(\mathcal{C})$ et la condition $|y_\nu| < 2X$ (X étant

---

(1) On peut donc mesurer la valeur de l'approximation commise en confondant S_n et S'_n par l'expression $\delta = \text{Min}\,[\varepsilon + \mathfrak{T}\{\eta > \varepsilon\}]$. On peut d'ailleurs, au lieu des deux quantités distinctes ε et γ, n'introduire qu'un seul nombre ε qui bornera supérieurement, soit α, soit $\mathcal{E}\{\bar{\eta}\}$, $\bar{\eta}$ désignant η si $\eta \leqq 1$ et 1 si $\eta > 1$; α et $\mathcal{E}\{\bar{\eta}\}$ sont du même ordre de grandeur.

(2) Toutefois cela n'est pas vrai, pour chaque chaîne prolongée indéfiniment, au sens de la loi forte des grands nombres. Nous n'étudions que des probabilités pour une valeur de t, très grande, et considérée indépendamment des autres.

(3) Cette hypothèse est un peu moins restrictive que celle que chaque chaîne possible de lois $L^{(n)}_{n-1}$ vérifie la condition du paragraphe **17**, sauf pour un ensemble de chaînes correspondant à une probabilité totale nulle. Il y a entre ces deux conditions la même différence qu'entre le théorème de Bernoulli et la loi forte des grands nombres.

petit par rapport à $\sigma'_n = \sqrt{t}$). Pour qu'il s'applique à la suite des x'_n, et par suite à celle des x_n, il suffit que chacune des lois $L_{n-1}^{(n)}$ soit symétrique; car dans ce cas $y_\nu = x'_\nu$. Nous appellerons cette condition *condition de symétrie* ou *condition* ($\mathcal{C}_1$). Nous avons ainsi établi le résultat suivant.

THÉORÈME VII. — *Si chaque variable x_n (quelles que soient les valeurs supposées connues des précédentes) dépend d'une loi symétrique, et si pour tout t assez grand on peut déterminer* X *de manière à rendre en même temps* ε, η' *et* γ *arbitrairement petits, alors* $\frac{\mathcal{S}(t)}{\sqrt{t}}$ *dépend d'une loi tendant, pour t infini, vers celle de Gauss* (¹).

Bien entendu dans cet énoncé, au lieu d'introduire l'inégalité (23) et les nombres ε et γ, on peut introduire un seul nombre ε qui devra borner supérieurement α; donc il s'agira de rendre α et η' simultanément très petits.

La condition de symétrie est d'ailleurs plus restrictive qu'il n'est nécessaire. En effet :

1° On peut modifier la loi $\mathcal{L}_{n-1}^{(n)}$ pour les valeurs de x dépassant en module un nombre grand $\overline{X}_n$ (pouvant dépendre de $x_1, x_2, \ldots, x_{n-1}$). Si la série $\Sigma \mathfrak{T}_{n-1} \{ |x_n| > \overline{X}_n \}$ a la probabilité de sa convergence égale à l'unité, cette modification est évidemment sans effet sur la loi limite.

2° On peut modifier la loi $\mathcal{L}_{n-1}^{(n)}$ pour les valeurs de x inférieures en module à un nombre assez petit $\overline{X}'_n$. Pourvu que $\overline{X}'_n$ soit petit par rapport à σ'_n, on peut prendre $\overline{X} > \overline{X}'_n$, et pourvu qu'en outre cette modification soit sans effet sur la valeur probable, on peut remplacer la loi $\mathcal{L}_{n-1}^{(n)}$ par une loi symétrique sans changer ni $\mathcal{E}_{n-1}\{x'_n\}$ ni $\mathcal{E}_{n-1}\{x'^2_n\}$, et, asymptotiquement, cette modification est sans effet sur la loi dont dépend $\frac{\mathcal{S}(t)}{\sqrt{t}}$.

3° Pour les valeurs de x de modules compris entre $\overline{X}'_n$ et $\overline{X}_n$, on peut se contenter d'une symétrie approchée. On remplacera par exemple

---

(¹) Il est entendu que la probabilité de la convergence de Σx_n est nulle. Cela résulte implicitement de ce que, par hypothèse, t augmente indéfiniment.

l'intervalle $(-X, +X)$ par un intervalle dissymétrique $(-X_n, +X'_n)$ déterminé de manière que $\mathcal{E}_{n-1}\{x'_n\}$ soit nulle. Pourvu que X_n et X'_n, vérifiant cette relation, puissent être pris petits par rapport à σ'_n, et tels que la somme qui remplace η soit petite, le théorème VII s'applique toujours. Si en outre les valeurs de x_n supérieures à X'_n et inférieures à $-X_n$ ont des probabilités dont le rapport ne soit ni très petit ni très grand, les raisonnements qui nous conduiront tout à l'heure au théorème réciproque ne subiront aucun changement essentiel.

Ces remarques permettent de définir une *condition* $(\mathcal{C}_2)$ ou *condition de symétrie approchée*, beaucoup plus large que $(\mathcal{C}_1)$, et qui peut remplacer $(\mathcal{C}_1)$ dans l'énoncé du théorème fondamental. L'essentiel est que la variable x'_ν, déduite de x_ν en remplaçant par zéro les grandes valeurs de x_ν vérifie, en plus de la condition fondamentale du théorème VII, la condition que la somme

$$\sum_1^n \mathcal{E}_{n-1}\{x'_\nu\}$$

soit, sauf dans des cas très peu probables, petite par rapport à $\sigma'_n = \sqrt{t}$.

Enfin c'est une remarque banale qu'on peut élargir encore la condition $(\mathcal{C}_2)$; pourvu que la somme précédente diffère, non de zéro, mais d'une certaine fonction $\omega(t)$ d'une quantité presque toujours petite par rapport à $\sqrt{t}$, le théorème VII s'applique en remplaçant $\mathfrak{S}(t)$ par $\mathfrak{S}(t) - \sqrt{t}$.

Pour simplifier l'exposé, dans la démonstration du théorème réciproque, nous n'introduirons que la condition $(\mathcal{C}_1)$, et non la condition plus large $(\mathcal{C}_2)$. Mentionnons simplement qu'on peut simplifier encore beaucoup plus en introduisant la *symétrie parfaite;* non seulement chaque loi $\mathcal{L}_{n-1}^{(n)}$ serait symétrique, mais elle ne dépendrait que des modules $|x_1|, |x_2|, \ldots, |x_{n-1}|$.

21. Le théorème réciproque.

Théorème VIII. — *Si la condition de symétrie est vérifiée et s'il n'y a pas de terme non individuellement négligeable, la deuxième condition de l'énoncé du théorème VII est nécessaire et suffisante pour que* $\frac{\mathfrak{S}(t)}{\sqrt{t}}$ *dépende d'une loi tendant, pour t infini, vers celle de Gauss.*

Il n'y a plus qu'à démontrer que cette condition est nécessaire. Supposons donc qu'elle ne soit pas vérifiée. Donnons-nous X assez grand, et faisons croître t jusqu'au moment où α [probabilité que le plus grand terme de $\mathfrak{S}(t)$, en valeur absolue, soit $> X$] a atteint ou dépassé une valeur donnée ε; dire que la condition considérée n'est pas réalisée, c'est dire qu'à ce moment, au moins pour certaines valeurs, arbitrairement grandes, de X, $\eta' = \frac{X}{\sqrt{t}}$ dépasse une valeur déterminée ε'; en outre, chaque terme étant par hypothèse individuellement négligeable, α varie lentement, tant que η' n'est pas très petit, de sorte qu'à l'instant considéré on a $\varepsilon \leqq \alpha < 2\varepsilon$. En d'autres termes, pour des valeurs de ε et ε' assez petites, mais fixes, il existera des systèmes de valeurs de X et t, indéfiniment croissantes, pour lesquelles $\eta' > \varepsilon'$ et $\varepsilon \leqq \alpha < 2\varepsilon$.

Dans ces conditions, cherchons d'abord une borne supérieure de $\mathfrak{P}\left\{ S'_n > \frac{X}{2} \right\}$. Désignons par β cette probabilité, et par m'_n et $-m''_n$ les moyennes (compte tenu des probabilités) des valeurs de S'_n respectivement supérieures au plus égales à $\frac{X}{2}$. On a

$$\beta m'_n = (1-\beta) m''_n, \qquad m'_n > \frac{X}{2}, \qquad m''_n > 0,$$
$$t = \sigma'^2_n \geqq \beta m'^2_n + (1-\beta) m''^2_n > \beta \frac{X^2}{4}.$$

On en déduit

$$\mathfrak{P}\left\{ S_n \geqq \frac{X}{2} \right\} \leqq \alpha + \beta < 2\varepsilon + \frac{4t}{X^2}. \tag{24}$$

Divisons maintenant l'intervalle $(0, t)$ en un certain nombre p de parties égales; p sera pris assez grand, mais indépendant de t; pour t croissant, comme il n'y a pas de terme individuellement non négligeable, à chacune des valeurs $\frac{ht}{p}$ $(h = 1, 2, \ldots, p)$ correspond une section pour laquelle σ'^2_n a la valeur considérée avec une erreur relative arbitrairement petite. Il y a alors au moins un intervalle partiel pour lequel la probabilité α_h de l'existence d'un terme x_ν de module $> X$ est $\geqq \frac{\varepsilon}{p}$ (car $\Sigma\alpha_h \geqq \alpha$). Nous désignerons un tel intervalle par (t, t'').

Soit ν, pour chaque chaîne contenant un ou plusieurs de ces termes, le rang du premier d'entre eux. D'après l'hypothèse de symétrie, il y a une chance sur deux pour que x_ν ait le signe de $S_{\nu-1} - \mathfrak{S}(t')$ et que par suite

$$|S_\nu - \mathfrak{S}(t')| > X.$$

La formule (24) limitant la probabilité que $\mathfrak{S}(t'') - S_\nu$ ait le signe contraire du précédent et soit en module $\geqq \frac{X}{2}$, il vient

$$\mathfrak{P}\left\{|\mathfrak{S}(t'') - \mathfrak{S}(t')| > \frac{X}{2}\right\} > \frac{\alpha_h}{2}\left(1 - \frac{4t}{pX^2} - 2\varepsilon\right) \geqq \frac{\varepsilon}{2p}\left(1 - \frac{4t}{pX^2} - 2\varepsilon\right),$$

d'où, comme $X \geqq \varepsilon' \sqrt{t}$,

$$\mathfrak{P}\left\{\xi > \frac{\varepsilon'}{2}\sqrt{p}\right\} > \frac{\varepsilon}{2p}\left(1 - \frac{4}{p\varepsilon'^2} - 2\varepsilon\right) \qquad \left[\mathfrak{S}(t'') - \mathfrak{S}(t') = \xi\sqrt{\frac{t}{p}}\right].$$

Or, γ, ε et ε' étant fixes, on peut prendre p arbitrairement grand. Cette formule donnant pour les grandes valeurs de ξ des probabilités qui ne décroissent pas assez vite, la loi dont dépend ξ ne peut pas être rendue arbitrairement voisine de celle de Gauss, comme cela serait le cas, pour n'importe lequel des intervalles partiels $\left[\frac{ht}{p}, (h+i)\frac{t}{p}\right]$, si $\frac{\mathfrak{S}(t)}{\sqrt{t}}$ dépendait d'une loi tendant, pour t infini, vers celle de Gauss. Il ne peut d'ailleurs être question, si ε est petit, d'un coefficient de réduction qui ne soit pas voisin de $\sqrt{t}$; s'il est plus petit, la même contradiction subsiste, et il ne peut pas être plus grand pour la raison indiquée à propos du théorème V. Le théorème VIII est donc démontré.

22. La somme d'un nombre donné de termes : Exemples et remarques.

1° Les théorèmes précédents s'appliquent naturellement à l'étude de S_n pour une valeur donnée de n si σ'_n est constant (à une erreur près relativement très faible). Tel sera le cas notamment si, les x_n étant bornés, $\mathcal{E}_{n-1}\{x_n\}$ et $\mathcal{E}_{n-1}\{x_n^2\}$ ne dépend que de n.

Il y a au contraire un cas important où une extension de ces théorèmes montre que $\frac{S_n}{\sigma_n}$ ne peut pas dépendre asymptotiquement de la loi de Gauss; c'est celui où, en posant $S_n = \sigma_n \xi_n$, ξ_n dépend asymptotiquement de la loi de Gauss et est (à peu près) indépendant de σ_n. La formule

$$\mathcal{E}\{S_n^p\} = \mathcal{E}\{\sigma_n^p\}\,\mathcal{E}\{\xi_n^p\} \qquad (p \text{ entier})$$

montre que S_n ne peut dépendre de la loi de Gauss (à un facteur constant près), que si les moments successifs $\mathcal{E}\{\sigma_n^p\}$ varient en progression géométrique, ce qui n'est le cas que si σ_n n'est pas une variable aléatoire.

Je rappelle (renvoyant pour plus de détail à mon travail antérieur) que cette circonstance se produit si σ_n est à peu près indépendant de $\mathfrak{S}_{n-1}$. Tel sera le cas (en supposant assurée l'application de la loi des grands nombres), dans l'hypothèse de la *symétrie parfaite*, où chaque loi $\mathcal{L}_{n-1}^{(n)}$, non seulement est symétrique, mais ne dépend que de $|x_1|$, $|x_2|, \ldots, |x_{n-1}|$; elle pourrait même sans inconvénient dépendre des signes d'un nombre fini de termes précédant x_n (qui ne constituent qu'une fraction négligeable de S_{n-1}).

Dans ces cas, la condition que σ_n soit constant (pour chaque n et à une erreur près relativement très petite si n est grand) est nécessaire et suffisante pour que S_n soit à la limite du type de Gauss.

2° Indiquons un exemple de cas où la somme de deux termes non gaussiens conduit à une somme gaussienne; d'après cet exemple notre lemme hypothétique III ne s'applique en tout cas pas aux variables enchaînées, même dans l'hypothèse de la symétrie parfaite. Supposons x égal à 0, a, ou $-a$, avec les probabilités $1-\alpha$, $\frac{\alpha}{2}$, $\frac{\alpha}{2}$: y dépendra d'une loi continue, symétrique, de densité $f(y)$ ou $g(y)$ suivant que $|x|$ sera nul ou non. Pour que la somme x, y dépende de la loi de Gauss, il suffit que

$$(1-\alpha)\,f(x) + \alpha\,\frac{g(x+a)+g(x-a)}{2} = \frac{1}{\sqrt{2\pi}}\,e^{-\frac{x^2}{2}},$$

condition facile à réaliser; il n'y a qu'à prendre $g(x)$ pair et au plus égal à

$$\frac{1}{\alpha}\sqrt{\frac{2}{\pi}}\,e^{-(|x|+a)^2},$$

et en déduire $f(x)$. On peut d'ailleurs rendre à volonté les moments

$$\int_{-\infty}^{+\infty} y^2 f(y)\,dy, \qquad \int_{-\infty}^{+\infty} y^2 g(y)\,dy$$

égaux ou différents, de sorte que la loi de Gauss peut se trouver obtenue, soit pour une section à t constant, soit pour une section à t variable.

3° Voici un autre exemple, utile pour la suite: il n'y aura plus symétrie, mais le jeu sera équitable (ce qui permet le calcul de $\sigma\{S\}$ par les mêmes formules que si les variables étaient indépendantes [1]). Donnons-nous une suite de nombres positifs λ_n, et, pour $n > 1$, prenons x_n tel que $S_n = \pm(1+\lambda_n)S_{n-1}$, c'est-à-dire que x_n a les valeurs possibles $\lambda_n S_{n-1}$ et $-(2+\lambda_n)S_{n-1}$, leurs probabilités étant proportionnelles à $2+\lambda_n$ et λ_n $\left(\text{pour } \lambda_n \text{ très petit, la probabilité de la seconde est donc} \sim \frac{\lambda_n}{2}\right)$. On a donc sûrement

$$|S_n| = |x|_1(1+\lambda_1)(1+\lambda_2)\ldots(1+\lambda_n),$$

et pour les signes, la loi est symétrique si celle de x_1 l'est, et tend à le devenir, quelle que soit la loi initiale $\mathcal{L}_1$, si la série $\Sigma\lambda_n$ est divergente.

Supposons alors que x_1 dépende de la loi de Gauss, et que pour n indéfiniment croissant, λ_n tend vers zéro, et que la série $\Sigma\lambda_n$ soit divergente. Dans ces conditions, tous les S_n sont du type de Gauss, le coefficient de réduction étant

$$P_n = (1+\lambda_1)(1+\lambda_2)\ldots(1+\lambda_n).$$

Mais $\sigma_n^2 = 1 + x_1^2(P_n^2 - 1)$ dépend de x_1^2; la section $n = \text{const.}$ n'est pas une section à t constant. D'autre part la somme $\Sigma\lambda_\nu$ ne peut être négligée que pour un groupe de termes pour lequel la variation relative de P_n est aussi négligeable, et est grande si cette variation est grande. Si donc, pour chaque terme de S_n pris individuellement, la probabilité qu'il dépasse $\varepsilon_i|S_n|$ peut être négligée, celle que le plus grand terme de S_n dépasse cette valeur tend au contraire vers l'unité. Nous avons ainsi obtenu un exemple de tendance vers le type de loi de Gauss, en *écartant en même temps* deux des hypothèses du théo-

---

[1] Nous entendons par là que $\sigma^2\{S_n\} = \sum_1^n \sigma^2\{x_\nu\}$, $\sigma^2\{x_\nu^2\}$ étant la valeur probable de $\sigma_{\nu-1}^2\{x_\nu\}$; il ne faut pas confondre σ et $\sigma_{\nu-1}$.

rème VII (constance de σ_n pour la section étudiée, et hypothèse des grands nombres).

4° Considérons une suite de variables gaussiennes indépendantes x_n (sans coefficients, c'est-à-dire que $\mathcal{E}\{x_n^2\} = 1$; à chacune d'elles associons une variable y_n, dépendant d'elle comme x_n dépendait de S_{n-1} dans l'exemple précédent, et étudions la somme

$$S_n = x_1 + y_1 + \ldots + x_n + y_n.$$

Toutes les sommes S_n, aussi bien que les sommes intermédiaires $S_{n-1} + x_n$, sont du type de Gauss. On a ici

$$\sigma_n^2 = \sum_1^n (1 + \lambda_\nu)^2 x_\nu^2,$$

de sorte que σ_n^2 est variable; mais sa valeur probable est $\sum\limits_1^n (1 + \lambda_\nu)^2$, et l'inégalité de Tchebycheff montre aisément que l'erreur commise en remplaçant σ_n par cette valeur est relativement très faible, si le plus grand des λ_ν^2 est négligeable devant leur somme. C'est précisément la condition pour que les grandes valeurs des y_ν soient négligeables, de sorte que non seulement chaque terme de S_n est individuellement négligeable, mais l'hypothèse des grands nombres est vérifiée. Dans ce schéma qui conduit à la loi de Gauss, il faut donc conserver ou abandonner à la fois les deux hypothèses du théorème VII.

5° Voici un autre exemple, dans lequel nous définirons x_n d'après deux formules différentes, qui alterneront à chaque retour au zéro de la somme S_{n-1}; nous appellerons *série* l'ensemble des termes consécutifs définis d'après une même formule. La première sera la règle du jeu de pile ou face ordinaire, avec enjeu constamment égal à l'unité; la seconde sera celle du jeu de pile ou face avec enjeu doublé chaque fois (à partir du troisième coup de la série), de manière à augmenter les chances d'un retour rapide au zéro; le premier enjeu A sera terminé, soit en fonction du nombre $2p - 1$ des séries précédant celle étudiée, soit en fonction du nombre total de leurs termes. Le nombre probable des termes d'une série est 3 avec une seconde formule; avec la première, il est infini (bien que la probabilité d'une série infinie soit nulle). De ce fait, à l'aide des principes connus on déduit aisément la prépondérance du premier procédé; pour une grande valeur de n, la probabilité que n soit choisi d'après la seconde formule est sensiblement $3\sqrt{\frac{2}{\pi n}}$, et $\frac{S_n}{\sqrt{n}}$ dépend d'une loi tendant vers celle de Gauss. Mais le choix des A, s'ils ne sont pas bornés, influe d'une part sur la valeur de σ_n, d'autre part sur la probabilité de l'existence d'un grand terme dans S_n. Ici encore, une étude sans difficulté montre que si les A sont assez grands pour quc l'hypothèse des grands nombres ne soit pas vérifiée, on ne peut pas sans une erreur non négligeable remplaçer σ_n par sa valeur probable. La conclusion est la même que pour l'exemple précédent.

6° Voici enfin un exemple de convergence vers le type de loi de Gauss d'un caractère beaucoup plus général. Soit d'abord une variable aléatoire x; désignons par $f(x)dx$ la probabilité qu'elle soit dans un intervalle dx (avec une erreur aussi petite qu'on veut sur la probabilité totale, la probabilité élémentaire est de cette forme). Soit $g(x)$ la plus petite des fonctions $f(x)$ et

$$f_1(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}},$$

et soit $1 - \alpha$ son intégrale (de $-\infty$ à $+\infty$). Lions x et x' par la relation

$$\int_{-\infty}^{x} [f(\xi) - g(\xi)]\, d\xi = \int_{-\infty}^{x'} [f_1(\xi) - g(\xi)]\, d\xi,$$

et définissons y par les conditions suivantes : si $f(x) = g(x)$, $y = 0$; si $f(x) > g(x)$, $|y|$ a les valeurs 0 et $|x' - x|$ avec des probabilités proportionnelles à $g(x)$ et $f(x) - g(x)$; les deux signes de y sont d'ailleurs également probables. On voit aisément, en étudiant $x + y$, que le nombre α qui définit (ou borne supérieurement) l'écart de la loi dont dépend x et de la loi de Gauss, est rendu deux fois plus petit (au moins) par l'addition à x du terme y. En recommençant, on obtient une convergence rapide vers la loi de Gauss.

Faisons alors alterner le procédé précédent avec un autre, absolument quelconque, que nous interromprons pour la $p^{\text{ième}}$ fois lorsque l'on aura

$$\delta_n = \mathrm{Max}\left[\left|\mathfrak{P}\{S_n < ax\} - \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-\frac{u^2}{2}}\, du\right|\right] > \alpha_p,$$

a étant d'ailleurs, soit déterminé de manière à rendre δ_n aussi petit que possible, soit égal à $\sigma\{S_n\}$; les α_p sont des nombres donnés tendant vers zéro. Quel que soit l'autre procédé, pourvu seulement que chaque terme soit individuellement négligeable, il est clair que l'alternance indiquée forcera la convergence vers la loi de Gauss; ce sera en principe une convergence avec oscillations, et l'on peut régler l'amplitude des oscillations en modifiant convenablement la règle suivant laquelle on fait alterner les deux procédés.

Nous ne pouvons pas déterminer σ_n sans faire d'hypothèse sur le second procédé; mais nous pouvons étudier la part des termes choisis d'après le premier procédé, tant en ce qui concerne la somme σ_n qu'en ce qui concerne la probabilité de l'existence de grands termes. La conclusion de cette étude, que nous nous contentons d'indiquer, est la même que pour les exemples précédents : si l'hypothèse des grands nombres est vérifiée, la part du hasard est négligeable pour les termes de σ_n^2 que nous considérons; dans le cas contraire, le rôle prépondérant de quelques termes fait que les variations possibles de σ_n^2 ne peuvent plus être négligées.

Ces exemples nous conduisent à formuler *l'hypothèse suivante* : *si la condition* ($\mathcal{C}$), *pour le cas des variables bornées, et une condition de symétrie approchée, pour le cas des variables non bornées, sont vérifiées, et si chaque terme est individuellement négligeable, on ne peut obtenir asymptotiquement la loi de Gauss que si les deux autres conditions du théorème* VII (section à t constant et possibilité de négliger le plus grand terme de S_n) *sont toutes les deux vraies ou toutes les deux fausses.* Cet énoncé comprend deux propositions, dont l'une est le théorème VIII (complété par les remarques finales du § **20**); il s'agit de savoir si l'autre est exacte.

23. Principe d'une méthode pour l'étude des variables enchainées les plus générales. — Nous supposerons seulement tous les $\mathcal{E}_\nu\{x_n\}$ bien définis (sauf peut-être dans des cas dont la probabilité *a priori* est nulle) et $\mathcal{E}\{x_n\} = 0$, cette dernière hypothèse n'étant pas restrictive (il n'est pas question ici de section à t constant; nous nous proposons d'étudier S_n). Posons

$$y_{\nu,n} = \mathcal{E}_\nu\{x_n\} - \mathcal{E}_{\nu-1}\{x_n\} \qquad [\nu \leqq n;\ \mathcal{E}_n\{x_n\} = x_n],$$
$$Y_{\nu,n} = y_{\nu,\nu} + {}_{\nu,\nu+1} + \ldots + y_{\nu,n}.$$

Ces expressions représentent respectivement les parties de x_n et de S_n qui cessent de devenir aléatoires au moment de la détermination de x_ν. Chacune d'elles, avant cette détermination, mais après celle de $x_{\nu-1}$, a sa valeur probable nulle. On peut donc étudier la somme

$$S_n = Y_{1,n} + Y_{2,n} + \ldots + Y_{n,n}$$

par application des méthodes précédemment exposées. Le fait que $Y_{\nu,n}$ dépende de n n'empêche pas d'appliquer les résultats obtenus à l'étude de S_n pour chaque valeur de n,

Ainsi, en supposant, les $Y_{\nu,n}$ bornés dans leur ensemble, on pourra définir la quantité

$$\sigma_n^2 = \sum_1^n \mathcal{E}_{\nu-1}\{Y_{\nu,n}^2\}$$

qui jouera le même rôle que dans l'exposé précédent. Seulement il n'est pas sûr qu'elle croisse avec n, et l'on ne pourra pas définir les sections à t constant.

Un cas d'application important de cette méthode est celui où $y_{\nu,n}$ décroît rapidement quand $n - \nu$ augmente; $Y_{\nu,n}$ tend alors rapidement vers une limite Y_ν. Chaque expérience, outre sa répercussion immédiate sur x_ν, a bien un effet retardé, mais agissant seulement sur un petit nombre de termes et atteignant vite sa valeur limite. On pourra donc, sans grande erreur, remplacer S_n par une expression de la forme

$$Y_1 + Y_2 + \ldots + Y_{n-p} + Y_{n-p+1,n} + \ldots + Y_{n,n},$$

et si p est assez petit pour que les p derniers termes puissent être négligés, on pourra obtenir des résultats simples.

Je me borne ici à cette brève indication; j'espère revenir sur cette question, et indiquer des applications de cette méthode. Malgré le rôle que joue une hypothèse de presque indépendance de x_ν et x_n si $n - \nu$ est grand, le principe est assez différent de celui que M. S. Bernstein a appliqué à l'étude de cas de ce genre, et semble pouvoir donner des résultats nouveaux.