JEAN VILLE

Étude critique de la notion de collectif

Thèses de l'entre-deux-guerres, 1939

<http://www.numdam.org/item?id=THESE_1939__218__1_0>

Série A, 1852
N° D'ORDRE :
2720

THÈSES

PRÉSENTÉES

A LA FACULTÉ DES SCIENCES DE PARIS

POUR OBTENIR

LE GRADE DE DOCTEUR ÈS SCIENCES MATHÉMATIQUES

PAR

Jean VILLE

1re THÈSE. — ÉTUDE CRITIQUE DE LA NOTION DE COLLECTIF.

2e THÈSE. — LA TRANSFORMATION DE LAPLACE.

Soutenues le 9 Mars 1939, devant la Commission d'Examen.

MM.	E. BOREL	*Président.*
	M. FRÉCHET	*Examinateurs.*
	R. GARNIER	

PARIS
GAUTHIER-VILLARS, IMPRIMEUR-ÉDITEUR
LIBRAIRE DU BUREAU DES LONGITUDES, DE L'ÉCOLE POLYTECHNIQUE
Quai des Grands-Augustins, 55

1939

FACULTÉ DES SCIENCES DE L'UNIVERSITÉ DE PARIS.

MM.

Doyen honoraire..... M. Molliard.

Doyen............... C. Maurain, *Professeur*, Physique du Globe.

Professeurs honoraires

H. Lebesgue.	Freundler.	Vessiot.	Charles Fabry.
Émile Picard.	Auger.	P. Portier.	Léon Bertrand.
Léon Brillouin.	Blaise.	M. Molliard.	Wintrebert.
Guillet.	Dangeard.	L. Lapicque.	O. Duboscq.
Péchard.	Lespieau.	G. Bertrand.	Bohn.
	Marchis.	H. Abraham.	Rabaud.

PROFESSEURS

M. Caullery.....	T	Zoologie (Evolution des êtres organisés).
Émile Borel.....	T	Calcul des probabilités et Physique mathématique.
Jean Perrin.....	T	Chimie physique.
E. Cartan.......	T	Géométrie supérieure.
A. Cotton.....	T	Recherches physiques.
J. Drach.........	T	Analyse supérieure et Algèbre supérieure.
Charles Pérez...	T	Zoologie.
M. Guichard.....	T	Analyse et mesures chimiques.
Paul Montel.....	T	Théorie des fonctions et théorie des transformations.
L. Blaringhem...	T	Botanique.
G. Julia.........	T	Mécanique analytique et Mécanique céleste.
C. Mauguin......	T	Minéralogie.
A. Michel-Lévy..	T	Pétrographie.
H. Bénard.......	T	Mécanique expérimentale des fluides.
A. Denjoy.......	T	Application de l'analyse à la Géométrie.
L. Lutaud.......	T	Géographie physique et Géologie dynamique.
Eugène Bloch...	T	Physique.
G. Bruhat.......	T	Physique théorique et Physique Céleste.
E. Darmois......	T	Enseignement de Physique.
A. Debierne.....	T	Physique générale et Radioactivité.
A. Dufour.......	T	Physique (P. C. B.).
L. Dunoyer......		Optique appliquée.
A. Guilliermond.	T	Botanique.
M. Javillier.....	T	Chimie biologique.
Robert-Lévy.....	T	Physiologie comparée.
F. Picard........		Zoologie (Évolution des êtres organisés).
Henri Villat....	T	Mécanique des fluides et applications.
Ch. Jacob........	T	Géologie.
P. Pascal.......	T	Chimie générale.
M. Fréchet......	T	Calcul différentiel et Calcul intégral.
E. Esclangon...	T	Astronomie.
Mme Ramart-Lucas.	T	Chimie organique.
H. Béghin.......	T	Mécanique physique et expérimentale.
Foch............		Mécanique expérimentale des fluides.
Pauthenier......		Physique (P. C. B.).
De Broglie......	T	Théories physiques.
Chrétien........		Optique appliquée.
P. Job..........		Chimie générale.
Labrouste.......		Physique du Globe.
Prenant.........	T	Anatomie et Histologie comparées.
Villey..........		Mécanique physique et expérimentale.
Combes.........	T	Physiologie végétale.
Garnier.........	T	Mathématiques générales.
Pérès..........		Mécanique théorique des fluides.
Hackspill.......		Chimie (P. C. B.).
Laugier.........		Physiologie générale.
Toussaint.......	T	Technique Aéronautique.
M. Curie........		Physique (P.C.B.).
G. Ribaud.......		Hautes températures.
Chazy...........	T	Mécanique rationnelle.
Gault...........	T	Chimie (P. C. B.).
Croze...........		Recherches physiques.
Dupont..........	T	Théories chimiques.
Lanquine........	T	Géologie structurale et Géologie appliquée.
Valiron.........		Mathématiques générales.
Barrabé.........		Géologie structurale et Géologie appliquée.
Millot..........		Biologie animale (P. C. B.).
F. Perrin.......		Théories physiques.
Vavon...........		Chimie organique.
G. Darmois......		Calcul des Probabilités et Physique-Mathématique.
Chatton.........		Biologie maritime.
Aubel...........	T	Chimie biologique.
Jacques Bourcart		Géographie physique et Géologie dynamique.
Mme Joliot-Curie.		Physique générale et Radioactivité.
Plantefol.......		Biologie végétale (P.C.B.).
Cabannes........		Enseignement de Physique.
Grassé..........		Biologie animale (P.C.B.).
Prévost.........		Chimie (P.C.B.).
Bouligand.......		Mathématiques.

Secrétaire......................... A. Pacaud.

Secrétaire honoraire.............. D. Tombeck.

PREMIÈRE THÈSE.

ÉTUDE CRITIQUE

DE LA

NOTION DE COLLECTIF.

INTRODUCTION.

Le présent travail m'a été inspiré par une conférence tenue par M. Wald au Séminaire Mathématique dirigé par M. Karl Menger à l'Université de Vienne pendant le Semestre d'Hiver 1934-35. Cette conférence se trouve reproduite dans les Comptes rendus de ce Séminaire, que j'ai cités dans la Bibliographie qui fait suite à mon exposé. M. Wald étudiait, en les reprenant sous une forme un peu abstraite, mais d'une grande rigueur, les théories proposées par M. de Misès pour servir de fondement au Calcul des Probabilités. J'ai, à mon tour, examiné les résultats de M. Wald, et suis arrivé à des conclusions différentes des siennes, du point de vue logique, s'entend. Du point de vue mathématique, non seulement je reconnais la validité de ses résultats, mais encore je les utilise après les avoir convenablement généralisés et approfondis.

M. Fréchet, qui m'avait d'abord proposé un autre sujet, a bien voulu, sur ma demande, accepter l'étude de la théorie de M. de Misès comme sujet de Thèse; il m'a aidé de nombreuses suggestions, et les conversations que j'ai eues avec lui m'ont amené à préciser ma pensée sur quelques points essentiels et délicats. Je le remercie vivement pour l'aide qu'il m'a ainsi apportée.

Je remercie également M. E. Borel, de s'être intéressé à mes recherches, et d'avoir aplani toutes les difficultés matérielles relatives à l'impression. Je dois ajouter que M. E. Borel m'a fourni l'occasion d'exposer mes idées à une séance des ***Exposés et Discussions sur le Calcul des Probabilités*** qu'il dirige à l'Institut Henri Poincaré, et que la discussion qui s'en est suivie m'a été d'un grand profit.

J'ai utilisé certains résultats de M. Paul Lévy, qui, à cette occasion, a lu une partie du manuscrit; ses observations m'ont été précieuses.

Je n'ai pas la prétention d'avoir dit le dernier mot sur la délicate question des Collectifs; mais je crois que mon travail pourra rendre service a ceux qui voudront consacrer une étude à la définition de la probabilité comme limite d'une fréquence.

CHAPITRE I.

ÉTUDE DES FRÉQUENCES DES DIFFÉRENTES CONFIGURATIONS QUI SE PRÉSENTENT DANS UNE SUITE FORMÉE DE 0 ET DE 1.

SOMMAIRE. — 1. Définition des configurations et des fréquences. — 2. Relations entre les fréquences des différentes configurations. — 3. Degré d'indépendance des différentes fréquences. — 4. Construction d'une suite ou les fréquences des configurations de longueur donnée ont des valeurs données. — 5. Construction d'une suite ou les fréquences de toutes les configurations ont des valeurs données. — 6. Généralisation : suites dont les termes peuvent prendre k valeurs différentes.

Introduction. — Ce chapitre a pour but d'établir entre les fréquences que l'on observe dans une suite, des relations qui serviront dans les chapitres suivants.

1. Définitions des configurations et des fréquences. — Soit une suite *finie* de termes égaux à 0 ou à 1 :

$$a^{(1)},\quad a^{(2)},\quad \ldots,\; a^{(m)} \qquad (a^{(i)} = 0 \text{ ou } 1;\; i = 1, 2, \ldots, m).$$

Nous appellerons cette suite une *configuration de longueur m*, et la représenterons, en supprimant les parenthèses et les virgules pour abréger l'écriture, par la notation

$$(1) \qquad (a^1 a^2 \ldots a^m).$$

Nous aurons quelquefois recours à une notation encore plus condensée, empruntée à la numération binaire, en représentant la configuration (1) par la lettre u affectée de deux indices, l'indice supérieur étant égal à m, l'indice inférieur ayant pour valeur

$$(2) \qquad i = a^m + 2a^{m-1} + 4a^{m-2} + \ldots + 2^{m-1}a^1.$$

Nous écrirons dans ces conditions

$$(3) \qquad u_i^m = (a^1 a^2 \ldots a^m).$$

Soit maintenant une suite *infinie* de termes égaux à 0 ou à 1

$$(4) \qquad x = x^1 x^2 \ldots x^n \ldots \qquad (x^i = 0 \text{ ou } 1\,;\ i = 1, 2, \ldots, n, \ldots).$$

La configuration $(a^1 a^2 \ldots a^m)$ sera dite figurer dans x au rang i si les m égalités suivantes sont vérifiées

$$(5) \qquad a^1 = x^i, \qquad a^2 = x^{i+1}, \qquad \ldots, \qquad a^m = x^{i+m-1}.$$

Introduisons le nombre y^i (où i est un indice) défini par

$$(6) \qquad y^i = \prod_{j=1}^{m} (1 - |x^{i+j-1} - a^j|) \qquad (i = 1, 2, \ldots, n, \ldots),$$

y^i ne peut prendre que les valeurs 0 ou 1, et les m égalités (5) sont équivalentes à l'égalité $y^i = 1$.

Définition 1. — *La répétition r de la configuration $(a^1 a^2 \ldots a^m)$ dans les n premiers rangs de x est le nombre des rangs distincts, inférieurs ou égaux à $n - m + 1$ $(n \geqq m)$, où $(a^1 a^2 \ldots a^m)$ figure dans x; c'est aussi le nombre de fois où nous rencontrons $(a^1 a^2 \ldots a^m)$ en lisant les n premiers termes de x.*

La fréquence de la configuration $(a^1 a^2 \ldots a^m)$ dans la suite des n premiers termes de x est par définition le rapport $\frac{r}{n}$. Il est égal au nombre $\frac{1}{n} \sum_{i=1}^{n-m+1} y^i$, où y^i est défini par (6). *Nous représenterons cette fréquence par la notation*

$$[a^1 a^2 \ldots a^m]_n^x \qquad (n \geqq m).$$

Quand n augmente indéfiniment $[a^1 a^2 \ldots a^m]_n^x$ peut tendre vers une limite; dans ce cas nous poserons la définition :

Définition 2. — *Si le nombre* $\lim\limits_{n=\infty}[a^1 a^2 a^m]_n^x$ *existe, ce sera par définition la fréquence totale selon l'expression de M. Borel, ou parfois plus brièvement la fréquence de la configuration* $(a^1 a^2 \ldots a^m)$ *dans la suite* x. *Nous représenterons cette fréquence par la notation* $[a^1 a^2 \ldots a^m]^x$. C'est évidemment un nombre non négatif au plus égal à 1.

Nous emploierons quelquefois par la suite les notations abrégées $[u_i^m]_n^x$ et $[u_i^m]^x$ au lieu des notations introduites dans les définitions 1 et 2, u_i^m ayant le sens indiqué plus haut.

2. Relations entre les fréquences des différentes configurations. — Soient la suite (4), une configuration $(a^1 a^2 \ldots a^m)$ et les deux configurations $(a^1 a^2 \ldots a^m 0)$ et $(a^1 a^2 \ldots a^m 1)$ qui s'en déduisent par prolongement d'un terme. Désignons par y^i, y_0^i et y_1^i les nombres associés à ces configurations par les égalités (6). Nous avons

$$(7) \qquad y_0^i = (1 - |x^{i+m}|)y^i, \qquad y_1^i (1 - |x^{i+m} - 1|)y^i$$

et par conséquent $y_0^i + y_1^i = y^i$. Il en résulte que le nombre

$$[a^1 a^2 \ldots a^m]_n^x - [a^1 a^2 \ldots a^m 0]_n^x - [a^1 a^2 \ldots a^m 1]_n^x$$

ne peut être égal qu'à 0 ou à $\frac{1}{n}$. D'où la proposition suivante :

Théorème 1. — *Si deux des fréquences*

$$[a^1 a^2 \ldots a^m]^x, \quad [a^1 a^2 \ldots a^m 0]^x, \quad [a^1 a^2 \ldots a^m 1]^x$$

existent, la troisième existe aussi, et elles satisfont à la relation

$$(8) \qquad [a^1 a^2 \ldots a^m 0]^x + [a^1 a^2 \ldots a^m 1]^x = [a^1 a^2 \ldots a^m],$$

ou, *avec la notation abrégée*

$$[u_{2i}^{m+1}]^x + [u_{2i+1}^{m+1}]^x = [u_i^m]^x.$$

Considérons maintenant les trois configurations

$$(a^1 a^2 \ldots a^m), \quad (0 a^1 a^2 \ldots a^m), \quad (1 a^1 a^2 \ldots a^m),$$

que nous représentons par les notations

$$u_i^m, \quad u_i^{m+1}, \quad u_{i+\alpha_m}^{m+1} \qquad (\alpha_m = 2^m).$$

En désignant par y^i, ${}_0y^i$, ${}_1y^i$ les nombres associés à ces configurations par (6), nous avons

$$ {}_0y^{i-1} = (1 - x^{i-1})y^i, \qquad {}_1y^{i-1} = (1 - |x^{i-1} - 1|)y^i $$

et, par conséquent,

$$ {}_0y^i + {}_1y^{i-1} = y^i \qquad (i \geqq 2); $$

nous en déduisons comme ci-dessus la proposition.

Théorème 1'. — *Si deux des fréquences*

$$ [a^1 a^2 \ldots a^m]^x, \quad [0 a^1 a^2 \ldots a^m]^x, \quad [1 a^2 a^2 \ldots a^m]^x $$

existent, la troisième existe aussi, et elles satisfont à la relation

$$ [0 a^1 a^2 \ldots a^m]^x + [1 a^1 a^2 \ldots a^m]^x = [a^1 a^2 \ldots a^m], $$

ou, *en notation abrégée,*

$$ (8') \qquad [u_i^{m+1}]^x + [u_{i+\alpha_m}^{m+1}]^x = [u_i^m]^x \qquad (\alpha_m = 2^m). $$

3. **Degré d'indépendance des différentes fréquences.** — Soit une infinité de nombres $v_i^m (i < 2^m)$, *non négatifs*, donnés à l'avance. Posons-nous la question : *existe-t-il une suite x telle que, quels que soient i et m, on ait l'égalité*

$$ (9) \qquad [u_i^m]^x = v_i^m \qquad (i < 2^m;\ m = 1, 2, \ldots, n, \ldots). $$

Une première condition *nécessaire* est évidemment que $v_0^1 + v_1^1 = 1$. D'après les théorèmes 1 et 1', il faut en outre que nous ayons

$$ (10) \qquad v_i^m = v_{2i}^{m+1} + v_{2i+1}^{m+1} = v_i^{m+1} + v_{i+2^m}^{m+1}. $$

Nous verrons (Théorème 4, Remarques I et II, p. 17), que ces conditions sont suffisantes. Les relations (10) montrent que l'on ne peut pas choisir arbitrairement les v_i^m; jointes à la relation $v_0^1 + v_1^1 = 1$, elles permettent de calculer tous les v_i^m à partir de ceux d'entre eux qui sont de la forme v_{2i}^m avec $i < 2^{m-2} (m = 1, 2, \ldots, h, \ldots)$. Nous ne démontrerons pas ce dernier résultat, qui s'établit sans difficulté, et que nous n'utiliserons pas par la suite. Si nous ne considérons que les configurations de longueur au plus égale à m_0, nous voyons que, d'après les relations (10), nous pouvons calculer tous les v_i^m, où $m \leqq m_0$ à partir des 2^{m_0} valeurs $v_i^{m_0}$, qui sont liées elles-mêmes par 2^{m_0-1} relations; d'où le théorème :

Théorème 2. — *Parmi les fréquences des configurations de longueur au plus égale à m_0, il en existe 2^{m_0-1} indépendantes.*

La démonstration de ce théorème exigerait que l'on montre que les relations (10), où $m = m_0 - 1$. sont indépendantes, et que ce sont les seules à lier les $v_i^{m_0}$. Nous l'établirons par une autre voie (*voir* la Remarque après le Lemme 4, § 4, p. 14).

4. Construction d'une suite x où les fréquences des configurations de longueur donnée ont des valeurs données. — Soit un système de nombres non négatifs v_i^m ($i < 2^m$) satisfaisant à $v_0^1 + v_1^0 = 1$ et à la condition (10). Si pour une suite x et une valeur fixe de m nous avons

$$[u_i^m]^x = v_i^m \qquad (i = 0, 1, \ldots, 2^m - 1), \tag{11}$$

il résulte de (8), (8′) et (10) que pour $j \leqq m$ les relations

$$[u_i^j]^x = v_i^j \qquad (j \leqq m;\ i < 2^j) \tag{12}$$

sont vérifiées. Portons notre attention sur les nombres v_i^m, où m a une valeur donnée. Examinons d'abord le cas où ils sont tous différents de zéro (¹), et introduisons les 2^{m-1} nouvelles quantités λ_i définies par

$$\lambda_i \equiv \frac{v_{2i}^m}{v_{2i+1}^m} \qquad (0 < \lambda_i < \infty;\ i = 0, 1, 2, \ldots, 2^{m-1} - 1). \tag{13}$$

En posant $\alpha = 2^{m-1}$, nous voyons que les v_j^m et les λ_i sont liés par les relations

$$(\Sigma)\qquad \left\{\begin{array}{ll} v_{2i}^m = \lambda_i v_{2i+1}^m & (v_j^m > 0;\ 0 \leqq i < 2^{m-1}), \\ v_i^m + v_{i+\alpha}^m = v_{2i}^m + v_{2i+1}^m & (0 < \lambda_i < \infty;\ \alpha = 2^{m-1}), \\ \displaystyle\sum_{j=0}^{2\alpha-1} v_j^m = 1. & \end{array}\right.$$

(¹) Le cas où certains des v_i^m sont nuls est traité dans la Remarque I de la page 17.

Les λ_i étant ainsi déterminés par (Σ), considérons le système d'équations linéaires, aux inconnues v_j,

$$(\Sigma_1)\qquad \begin{cases} v_{2i} = \lambda_i v_{2i+1} & (0 \leqq i < 2^{m-1}), \\ v_i + v_{i+\alpha} = v_{2i} + v_{2i+1} & (\alpha = 2^{m-1}), \\ \displaystyle\sum_{j=0}^{2\alpha-1} v_j = 1. \end{cases}$$

Ce système admet la solution $\{v_j = v_j^m\}$. Supposons qu'il admette une autre solution $\{v_i = w_j\}$ avec $w_j \geqq 0$. Quel que soit μ, (Σ_1) admet alors la solution $\{v_j = \mu v_j^m + (1-\mu) w_j\}$. L'ensemble des valeurs de μ telles que, pour $j = 0, 1, \ldots, 2\alpha - 1$, nous ayons $\mu v_j^m + (1-\mu) w_j \geqq 0$ est un ensemble fermé, admettant le point $\mu = 1$ comme point intérieur, puisque les v_j^m sont différents de zéro; il existe certainement une valeur de μ telle que pour une valeur de j, on ait $\mu v_j^m + (1-\mu) w_j < 0$; nous en déduisons donc qu'il existe une valeur μ_0 de μ telle que $\mu_0 v_j^m + (1-\mu_0) w_j \geqq 0$, l'égalité ayant lieu pour une valeur au moins de l'indice j; soit j_0 une de ces valeurs, et soit $u_j = \mu_0 v_j^m + (1-\mu_0) w_j$.

Nous avons $u_{j_0} = 0$. Si j_0 est pair, soit $j_0 = 2k$, l'examen de Σ_1 montre que u_{2k+1}, u_k, $u_{k+\alpha}$ sont nuls, puisque les λ_i sont non nuls. Si $j_0 = 2k+1$, nous concluons à la nullité de u_{2k}, u_k, $u_{k+\alpha}$. En raisonnant sur u_k comme sur u_{j_0}, nous montrerons, en opérant de proche en proche, que $u_0 = 0$. De $u_j = 0$, nous déduisons la nullité de $u_{j+\alpha}$. De la nullité de u_{2j}, nous déduisons celle de u_j. En partant de $u_0 = 0$, nous obtiendrons ainsi le tableau des u_j nuls

$$\begin{tikzcd}
 & & & u_0 \arrow[dl] \\
 & & u_0 \arrow[rr] \arrow[dl] & & u_\alpha \arrow[dl] \\
 & u_0 \arrow[r] \arrow[dl] & u_\alpha \arrow[r] \arrow[dl] & u_{\alpha/2} \arrow[r] \arrow[dl] & u_{3\alpha/2} \arrow[dl] \\
u_0 \arrow[r] & u_\alpha . u_{\alpha/2} \arrow[r] & u_{3\alpha/2} . u_{\alpha/4} \arrow[r] & u_{5\alpha/4} . u_{3\alpha/4} \arrow[r] & u_{7\alpha/4}
\end{tikzcd}$$

Soit $\omega_j^i (0 < j \leqq 2^i)$ l'indice du terme de rang j dans la $(i+1)^e$ ligne de ce tableau. La loi de formation des ω_j^i est

$$\omega_1^0 = 0, \qquad \omega_{2j+1}^{i+1} = \frac{1}{2}\,\omega_j^i, \qquad \omega_{2j}^{i+1} = \omega_{2j-1}^{i+1} + \alpha \qquad (\alpha = 2^{m-1};\ i \leqq m).$$

Supposons démontré pour une valeur i que :

1° Quel que soit j, ω_j^i est divisible par 2^{m-i}.
2° Les 2^i nombres $\omega_j^i (j = 1, 2, \ldots, 2^i)$ sont distincts.
3° Le plus grand des nombres ω_j^i est $\omega_{2^i}^i = 2^m - 2^{m-i}$.

Les propriétés 1°, 2°, 3° sont encore vérifiées pour la valeur $i+1$. Pour 1°, c'est manifeste. Pour 2° : les nombres ω_{2j-1}^{i+1} sont distincts et les nombres ω_{2j}^{i+1} également; le plus grand des nombres ω_{2j-1}^{2i+1} est $\frac{1}{2}\omega_{2^i}^i = 2^{m-1} - 2^{m-i-1}$; le plus petit des nombres ω_{2j}^{i+1} est au moins égal à $\alpha = 2^{m-1}$; donc les nombres ω_j^{i+1} sont tous distincts. Enfin, 3°, le plus grand des ω_j^{i+1} est $\omega_{2^i}^i + \alpha = \omega_{i+1}^{i+1} = 2^m - 2^{m-i-1}$.

1°, 2°, 3° étant vérifiées quel que soit $i \leqq m$, en particulier, pour $i = m$, les 2^m nombres ω_j^m sont distincts, et le plus grand est $2^m - 1$. Ce sont des entiers non négatifs. Ce sont donc les entiers de 0 à $2^m - 1$, rangés dans un autre ordre que l'ordre naturel. Revenons à la signification des ω_j^i : nous concluons à la nullité de tous les u_j, ce qui est en contradiction avec $\sum_{j=0}^{2\alpha-1} u_j = 1$. D'où le lemme :

Lemme 1. — *Si les 2^m valeurs positives $v_{2_i}^m$ satisfont aux deux dernières relations du système (Σ), et si les 2^{m-1} valeurs λ_i leur sont associées par la première relation du système (Σ), le système d'équations (Σ_1) aux inconnues v_j admet comme unique solution à termes non négatifs la solution*

$$\{ v_j = v_j^m \}.$$

Nous mettons en évidence, dans l'énoncé du lemme 1, les valeurs v_i^m données; si nous considérons le système (Σ_1) seul, nous voyons qu'une conséquence immédiate du lemme 1 est le lemme :

Lemme 2. — *Le système (Σ_1), où les λ_i sont supposés positifs non nuls, admet au plus une solution en v_j à termes $\geqq 0$. De plus, cette solution satisfait à $v_j \neq 0 (j = 0, 1, \ldots, 2\alpha - 1)$.*

La deuxième partie du lemme résulte du fait que $v_{j_0} = 0$ entraîne $v_0 = 0$, ce qui entraîne la nullité de tous les v_j. La donnée des λ_i est donc équivalente à celle des v_j^m. La proposition suivante va nous permettre de faire un pas de plus dans la recherche de la suite x.

Lemme 3. — *m étant un nombre fixe, si les valeurs v_j^m satisfont aux conditions*

$$v_j^m > 0, \qquad v_i^m + v_{i+\alpha}^m = v_{2i}^m + v_{2i+1}^m \qquad (\alpha = 2^{m-1}), \qquad \sum_{j=0}^{2\alpha-1} v_j^m = 1,$$

toute suite x telle que, quel que soit i, on ait

$$\lim_{n=\infty} \frac{[u_{2i}^m]_n^x}{[u_{2i+1}^m]_n^x} = \frac{v_{2i}^m}{v_{2i+1}^m} \qquad (0 \leqq i < 2^{m-1}),$$

satisfait aux 2^m relations

$$\lim_{n=\infty} [u_j^m]_n^x = v_j^m \qquad (0 \leqq j < 2^m).$$

Démonstration. — Soit une suite x mettant le théorème en défaut. Il existe alors une valeur de j, soit j_0, et une suite de valeurs de n, soit $\{ n_k \}$, telles que

$$\lim_{k=\infty} [u_{j_0}^m]_{n_k}^x \neq v_{j_0}^m.$$

Soit x_k la suite périodique ayant pour période la suite finie des n_k premiers termes de x. Posons

$$(14) \qquad w_j^{(k)} = [u_j^m]^{x_k}, \qquad \mu_i^k = \frac{[u_{2i}^m]^{x_k}}{[u_{2i+1}^m]^{x_k}}, \qquad \lambda_i = \frac{v_{2i}^m}{v_{2i+1}^m}.$$

La configuration u_j^m figure autant de fois dans les n_k premiers termes de x que dans les n_k premiers termes de x, donc

$$[u_j^m]_{n_k}^{x_k} = [u_j^m]_{n_k}^x;$$

si nous considérons les $2n_k$ premiers termes de x_k. la configuration u_j^m figurera deux fois autant que dans les n_k premiers termes de x et peut-être encore en quelques-uns des rangs $n_k - m + 1$, $n_k - m + 2$, ..., n_k. D'où l'inégalité

$$(15) \qquad 0 \leqq 2n_k [u_j^m]_{2n_k}^{x_k} - 2n_k [u_j^m]_{n_k}^x \leqq m.$$

D'une manière générale, nous aurons

$$0 \leqq sn_k [u_j^m]_{sn_k}^{x_k} - sn_k [u_j^m]_{n_k}^{x} \leqq (s-1)m \qquad (s = 1, 2, \ldots, n, \ldots),$$

d'où, en faisant tendre s vers l'infini,

$$(16) \qquad 0 \leqq w_j^{(k)} - [u_j^m]_{n_k}^{x} \leqq \frac{m}{n_k}.$$

Tout μ_i^k sera donc de la forme

$$(17) \qquad \mu_i^k = \frac{n_k [u_{2i}^m]_{n_k}^{x} + \theta_0 m}{n_k [u_{2i+1}^m]_{n_k}^{x} + \theta_1 m} \qquad (0 \leqq \theta_0 \leqq 1,\ 0 \leqq \theta_1 \leqq 1).$$

La valeur de μ_i^k quand k tend vers l'infini dépend de la croissance de $n_k [u_{2i}^m]_{n_k}^{x}$ et de $n_k [u_{2i+1}^m]_{n_k}^{x}$.

Considérons d'une manière générale $n [u_j^m]_n^x$. Si cette quantité est bornée supérieurement, la configuration u_j^m ne figure qu'un nombre fini de fois dans x. Mais il résulte des hypothèses faites que si u_{2i}^m ne figure qu'un nombre fini de fois, il en est de même de u_{2i+1}^m, et réciproquement. Par ailleurs, nous avons vu au paragraphe 2 que

$$[u_i^m]_n^x - [u_{2i}^{m+1}]_n^x - [u_{2i+1}^{m+1}]_n^x \leqq \frac{1}{n}, \qquad [u_i^m]_n^x - [u_i^{m+1}]_n^x - [u_{i+\alpha}^{m+1}]_n^x \leqq \frac{1}{n},$$

d'où nous déduisons que

$$(18) \qquad | n [u_{2i}^m]_n^x + n [u_{2i+1}^m]_n^x - n [u_i^m]_n^x - n [u_{i+\alpha}^m]_n^x | \leqq 2.$$

Si donc l'une des configurations u_{2i}^m, u_{2i+1}^m ne figure dans x qu'un nombre fini de fois, il en est de même pour l'autre et pour les configurations u_i^m et $u_{i+\alpha}^m$. La détermination des indices j tels que u_j^m ne figure qu'un nombre fini de fois est la même que la détermination des nombres ω_j^i employés dans la démonstration du Théorème 3. Nous arrivons à la conclusion que chacune des configurations u_j^m ne figure qu'un nombre fini de fois dans x, ce qui est manifestement absurde. Donc, $n [u_j^m]_n^x$ n'est pas borné ; comme c'est une fonction non décroissante de n, nous avons

$$\lim_{n=\infty} n [u_j^m]_n^x = \infty,$$

ce qui, en nous reportant à (17), montre que

$$\lim_{k=\infty} \mu_i^k = \lambda_i.$$

A partir d'une certaine valeur de k, les μ_i^k sont donc positifs (non nuls) et finis. Les $w_j^{(k)}$ sont la solution, unique, du système (Σ_1) aux inconnues $v_j \geqq 0$ et aux paramètres μ_i^k (μ_i^k étant mis à la place de λ_i), tandis que les v_j^m sont la solution unique du système (Σ_1) avec les paramètres λ_i. Quand les λ_i sont positifs, les v_j, déterminés d'une manière univoque, sont des fonctions continues de λ_i, puisque le système Σ_1 est linéaire ; nous déduisons donc de (19) que

$$(20) \qquad \lim_{k=\infty} w_j^{(k)} = v_j^m ;$$

or, d'après (16),

$$\lim_{k=\infty} w_{j_0}^k = \lim_{k=\infty} [u_{j_0}^m]_{n_k}^x ;$$

nous avons supposé cette dernière limite différente de $v_{j_0}^m$: nous sommes amenés à contradiction, ce qui démontre le lemme **3**.

Grâce à cette dernière proposition, nous pouvons établir le Théorème suivant :

Théorème 3. — *m étant un nombre fixe, si les valeurs $v_j^m (0 \leqq j < 2^m)$ satisfont aux conditions*

$$v_j^m > 0, \qquad \sum_{j=0}^{2\alpha-1} v_j^m = 1, \qquad v_i + v_{i+\alpha} = v_{2i}^m + v_{2i+1}^m$$
$$(0 \leqq i < 2^{m-1} ;\ \alpha = 2^{m-1}),$$

on peut construire une suite x telle que dans cette suite

$$\lim_{n=\infty} [u_j^m]_n^x = v_j^m .$$

Démonstration. — Posons $\gamma_i = \dfrac{v_{2i}^m}{v_{2i+1}^m}$; construisons 2^m suites $y_0, \ldots, y_{\alpha-1}$, dont les termes sont égaux à 0 ou à 1

$$(21) \qquad y_j = y_j^1 y_j^2 \ldots y_j^n \ldots \qquad (j = 0, 1, 2, \ldots, \alpha - 1),$$

telles que dans ces suites on ait

$$(22) \qquad \frac{[0]^{y_j}}{[1]^{y_j}} = \gamma_j .$$

La construction de telles suites n'offre aucune difficulté.

Les $m - 1$ premiers termes de x seront arbitrairement choisis; ils formeront une configuration de longueur $m - 1$, soit la configulation $u_{j_0}^{m-1}$; nous prendrons pour x^m la valeur de $y_{j_0}^1$.

D'une manière générale, les n premiers termes de x ayant été choisis, soient

$$x^1 x^2 \dots x^n,$$

les termes $x^{n-m+2} x^{n-m+3} \dots x^n$ forment une configuration de longueur $m - 1$, soit $u_{i_n}^{m-1}$; le nombre de fois que cette configuration figure dans la suite finie $x^1, x^2, \dots, x^n$, à savoir :

$$n\left[u_{i_n}^{m-1}\right]_n^x,$$

est bien déterminé; soit k_n ce nombre. Nous choisirons pour x^{n+1} la valeur de $y_{i_n}^{k_n}$. Il reste à montrer que la suite ainsi construite satisfait bien aux conditions demandées.

Ne conservons, parmi les termes de la suite $x^1, x^2, \dots, x^{n+1}$, que les termes x^i tels que la configuration $x^{i-m+1} x^{i-m+2} \dots x^{i-1}$ soit identique à une configuration u_j^{m-1} donnée à l'avance. D'après la loi de formation de x, les termes conservés sont identiques, quant à la valeur et à l'ordre, aux $h_n^{(j)}$ premiers termes de la suite y_j, avec

$$h_n^{(j)} = n\left[u_j^{m-1}\right]_n^x.$$

Si donc, $h_n^{(j)}$ tend vers l'infini avec n, nous déduisons de la remarque précédente que nous aurons bien

$$\lim_{n=\infty} \frac{\left[u_{2j}^m\right]_n^x}{\left[u_{2j+1}^m\right]_n} = \gamma_j. \tag{23}$$

Si la condition (23) est vérifiée, quel que soit u_j^{m-1}, par application du théorème 4, nous voyons que x satisfait aux conditions demandées. Le seul point à démontrer est que, quel que soit la configuration de longueur $m - 1$ considérée, on a

$$\lim_{n=\infty} h_m^{(j)} = \infty. \tag{24}$$

Si, pour une configuration u_j^{m-1} la quantité $h_n^{(j)}$ est bornée, les configurations u_{2j}^m et u_{2j+1}^m ne figurent dans x qu'un nombre fini de fois. Inversement, si l'une des configurations u_{2j}^m ou u_{2j+1}^m ne figure dans x qu'un nombre fini de fois, il en est de même de l'autre; par application de l'inégalité (18) utilisée dans la démonstration du lemme 3, nous pouvons ajouter que u_i^m et $u_{m\,i}^{i+\alpha}$ ne figurent qu'un nombre fini de fois. Par un raisonnement identique à celui employé dans la démonstration du lemme 3, nous voyons que supposer un des $h_n^{(j)}$ borné amène à contradiction. Donc la relation (23) est vraie ponr tous les $h_n^{(j)}$, ce qui établit la proposition.

Le théorème 3 va nous permettre d'énoncer un résultat relatif au système (Σ_1) de la page 6. La construction utilisée dans la démonstration du théorème 3 ne fait intervenir que les valeurs γ_i. Choisissons ces valeurs rationnelles positives; nous pouvons alors astreindre les suites y_j à être périodiques. Construisons la suite x par le procédé indiqué : il est facile de montrer que cette suite x sera elle-même périodique à partir d'un certain rang. Soit, en effet, M le plus petit commun multiple des longueurs des périodes des suites y_j; si, en deux endroits de x, soient aux rangs n_1 et n_2, nous avons

$$h_{n_1}^{(j)} \equiv h_{n_2}^{(j)} \pmod{\mathrm{M}} \qquad (j = 0, 1, \ldots, \alpha - 1 \,;\, \alpha = 2^{m-1}),$$

la suite x à partir du rang n_2 reproduira la suite x elle-même à partir du rang n_1, à condition que les configurations $(x^{n_1-m+2} \ldots x^{n_1})$ et $(x^{n_2-m+2} \ldots x^{n_2})$ soient identiques. Considérons l'ensemble d'une configuration de longueur $m - 1$ et d'un système de α valeurs $h^{(j)}$, deux valeurs $h^{(j)}$ n'étant considérées comme distinctes que si elles ne sont pas congruentes (mod M). Cet ensemble de $\alpha + 1$ éléments peut prendre au plus $\alpha \mathrm{M}^\alpha$ valeurs différentes, à savoir α pour la configuration, M pour chacune des α valeurs $h^{(j)}$. Donc, sur un segment de x contenant plus de $\alpha \mathrm{M}^\alpha$ termes, la coïncidence qui entraîne périodicité se présente au moins une fois.

La suite x étant périodique, les fréquences $[u_j^m]^x$ existent toutes. Elles constituent une solution du système (Σ_1) aux paramètres γ_i. Le système Σ_1 admet donc une solution telle que $v_j \geqq 0$ pour tout système de valeurs rationnelles positives des λ_i. Cette solution est unique d'après le lemme 2; elle est manifestement continue en $\lambda_0, \lambda_1, \ldots, \lambda_{\alpha-1}$. En tenant compte du lemme 2, nous pouvons donc énoncer le résultat :

Lemme 4. — *Le système Σ_1 de la page 7 admet, pour tout système de valeurs positives des λ_i, une solution en v_j qui satisfait à $v_j \geqq 0$. Cette solution est unique et satisfait de plus à la condition $v_j > 0$.*

Remarque. — Il résulte du théorème 3 et du lemme 4 que les fréquences des configurations de longueur m dépendent de 2^{m-1} paramètres, à savoir les λ_i.

Nous pouvons maintenant passer à l'étude du cas où les valeurs de toutes les fréquences sont données.

5. Construction d'une suite x, où les fréquences de toutes les configurations ont des valeurs données. — Soit maintenant un système $\{ v_j^m \}$ de nombres non négatifs, satisfaisant à $v_0^1 + v_1^1 = 1$ et aux égalités (10) (§ 3). m prend toutes les valeurs entières positives, et j les valeurs entières non négatives inférieures à 2^m. Nous allons prouver qu'il existe une suite x telle que

$$[u_i^m]^x = v_i^m,$$

quels que soient m et i et montrer comment la construire, la construction même établissant cette existence.

Soit $\{ \varepsilon_i \}$ une suite de nombres > 0, avec $\lim \varepsilon_i = 0$.

Soit une valeur de i; nous pouvons trouver des nombres w_j^m, *positifs* ($m \leqq i$, $j < 2^m$), tels que

$$(25)\quad w_0^1 + w_1^1 = 1, \quad | w_j^m - v_j^m | < \frac{\varepsilon_i}{3} \quad (m = 1, 2, \ldots, i;\ j = 0, 1, 2, \ldots, 2^m - 1)$$

et satisfaisant de plus aux égalités (10).

Posons, en effet, $w_j^m = (1 - \mu) v_j^m + \mu 2^{-m}$, où $0 \leqq \mu \leqq 1$. Les expressions sont $\geqq 0$; elles satisfont à (10) pour $\mu = 0$ et pour $\mu = 1$, donc pour toute valeur de μ comprise entre 0 et 1. Les différences $| w_j^m - v_j^m |$ sont alors fonctions continues de μ; nous pouvons donc trouver une valeur de $\mu > 0$ telle que (25) soit vérifiée. Pour cette valeur de $\mu > 0$, tous les w_j^m sont bien positifs.

D'après le théorème 3, il existe une suite x_i telle que $[u_j^m]^{x_i} = w_j^m$. D'où le *premier stade* de la construction de x :

A tout entier i, nous associons une suite x_i telle que

$$(26)\quad | [u_j^m]^{x_i} - v_j^m | < \frac{\varepsilon_i}{3} \qquad (m = 1, 2, \ldots, i;\ j = 0, 1, \ldots, 2^m - 1).$$

Associons à tout entier i un entier μ_i tel que pour $n \geqq \mu_i$ on ait

$$(27)\qquad | [u_j^m]_n^{x_i} - [u_j^m]^{x_i} | < \frac{\varepsilon_i}{3} \qquad (n \geqq \mu_i;\ m \leqq i;\ j < 2^m).$$

Soit N_1 le plus grand des nombres μ_1 et $\left[\frac{3\mu_2}{\varepsilon_1}\right]$ (¹). Soit S_1 la suite formée des N_1 premiers termes de x, à la suite desquels nous écrivons les termes de x_2. Nous allons montrer que pour $n \geqq N_1$, nous avons

$$(28) \qquad |[u_j^1]_n^{S_1} - v_j^1| < \varepsilon_1 \qquad (j = 0, 1).$$

Si, en effet, $N_1 \leqq n \leqq N_1 + \mu_2$, nous avons

$$(29) \qquad |[u_j^1]_n^{S_1} - [u_j^1]_{N_1}^{x_1}| \leqq \frac{n - N_1}{n} \leqq \frac{\mu_2}{N_1} \leqq \frac{\varepsilon_1}{3};$$

en tenant compte de $N_1 \geqq \mu_1$, de (26) et de (27), nous obtenons bien (28).

Si maintenant $n \geqq N_1 + \mu_2$, nous voyons que

$$n[u_j^1]_n^{S_1} = N_1[u_j^1]_{N_1}^{x_1} + (n - N_1)[u_j^1]_{n-N_1}^{x_2}$$

et par conséquent $[u_j^1]_n^{S_1}$ est compris entre les nombres $[u_j^1]_{n_1}^{x_1}$ et $[u_j^1]_{n-N_1}^{x_2}$, avec $n - N_1 \geqq \mu_2$, $N_1 \geqq \mu_1$. En tenant compte de (27), nous obtenons encore l'inégalité (28).

Nous allons former maintenant, à partir de S_1 et x_3 une suite S_2, d'une manière analogue à celle dont S_1 a été formée à partir de x_1 et x_2

Soit ν_1 un nombre tel que $\nu_1 \geqq N_1$, et que pour $n \geqq \nu_1$ on ait

$$(30) \qquad |[u_j^m]_n^{S_1} - [u_j^m]^{S_1}| \leqq \frac{\varepsilon_2}{3} \qquad (m = 1, 2;\ j = 0, 1, \ldots, 2^m - 1).$$

S_1 se confondant avec la suite x_2 si l'on supprime ses N_1 premiers termes, nous déduisons de (30) et de (26) que

$$(31) \qquad |[u_j^m]_n^{S_1} - v_j^m| \leqq \frac{2\varepsilon_2}{3} \qquad (m = 1, 2;\ n \geqq \nu_1).$$

Soit N_2 le plus grand des nombres ν_1 et $\left[\frac{3\mu_3}{\varepsilon_2}\right]$. La suite S_2 sera formée des N_2 premiers termes de S_1, à la suite desquels nous écrirons x_3. Pour $n \geqq N_2$, nous aurons

$$(32) \qquad |[u_j^m]_n^{S_2} - v_j^m| < \varepsilon_2 + \frac{1}{n} \qquad (m = 1, 2).$$

(¹) Nous représenterons par $[z]$ la valeur approchée à une unité près par excès du nombre z.

En effet, pour $N_1 \leqq n \leqq N_2 + \mu_3$, nous aurons

$$| [u_j^m]_n^{S_2} - [u_j^m]_{N_2}^{S_1} | \leqq \frac{n - N_2}{n} \leqq \frac{\mu_3}{N_2} \leqq \frac{\varepsilon_2}{3} \qquad (m = 1, 2),$$

ce qui, tenu compte de $N_2 \geqq \nu_1$, et de (31), nous donne (32).

Si maintenant $n \geqq N_2 + \mu_3$, nous avons

$$n [u_j^m]_n^{S_2} = N_2 [u_j^m]_{N_2}^{S_1} + (n - N_2) [u_j^m]_{n-N_2}^{x_3} + \theta,$$

où θ est un nombre égal à 0 ou à 1 (θ est égal à 1 si le N_2^e terme de S_1 et le premier terme de x_3 forment une configuration u_j^m; il faut pour cela que $m = 2$). Nous en déduisons que, à $\frac{1}{n}$ près, $[u_j^m]^{S_2}$ est compris entre $[u_j^m]_{N_2}^{S_1}$ et $[u_j^m]_{n-N_2}^{x_3}$, ce qui, tenu compte de $N_2 \geqq \nu_1$, $n - N_2 \geqq \mu_3$, de (31), de (26) et (27), nous donne encore (32).

D'une manière générale, supposons formée S_i, et déterminé N_i tel que pour $n \geqq N_i$ nous ayons

$$(33) \qquad | [u_j^m]^{S_i} - v_j^m | < \varepsilon_i + \frac{i-1}{n} \qquad (m = 1, 2, \ldots, i).$$

Supposons qu'à partir d'un certain rang, les termes de S_i soient ceux de x_{i+1}. Il existe un nombre ν_i tel que pour $n \geqq \nu_i$ on ait

$$| [u_j^m]_n^{S_i} - v_j^m | < \frac{2\varepsilon_{i+1}}{3} \qquad (m = 1, 2, \ldots, i+1).$$

Prenons pour N_{i+1} le plus grand des nombres N_i, ν_i, $\left[\frac{3\mu_{i+2}}{\varepsilon_{i+1}}\right]$. La suite S_{i+1} sera formée des N_{i+1} premiers termes de S_i, à la suite desquels nous écrirons x_{i+2}. Nous démontrerions comme plus haut que pour $n \geqq N_{i+1}$

$$(34) \qquad | [u_j^m]_n^{S_{i+1}} - v_j^m | < \varepsilon_{i+1} + \frac{i}{n} \qquad (m = 1, 2, \ldots, i+1).$$

S_{i+1} jouit des mêmes propriétés que S_i; le procédé de formation de ces suites S_i se poursuit donc indéfiniment. Nous prendrons pour x la suite limite de S_i quand i devient infiniment grand, c'est-à-dire la suite obtenue en écrivant à la suite les uns des autres les N_1 premiers termes de x_1, les $N_2 - N_1$ premiers termes de x_i, et ainsi de suite. Pour $N_i \leqq n \leqq N_{i+1}$, nous aurons alors

$$| [u_j^m]_n^x - v_j^m | < \varepsilon_i + \frac{i-1}{n} \qquad (m = 1, 2, \ldots, i;\ j < 2^m),$$

ce qui entraîne bien, quels que soient m et j,

$$\lim_{n=\infty} [u_j^m]_n^x = v_j^m.$$

Nous pouvons donc énoncer le théorème

THÉORÈME 4. — *Soit le système de nombres v_j^m, positifs ou nuls, satisfaisant aux relations* (10), *et à $v_0^1 + v_1^1$. Il existe une suite x telle que, quels que soient m et $j(< 2^m)$ on ait*

$$[u_j^m]^x = v_j^m.$$

Remarque I. — Si au lieu de donner tous les v_j^m ($m = 1, 2, 3, \ldots;$ $j < 2^m$) on se donne seulement ceux qui correspondent à une longueur donnée m_0, les relations (10) permettent de calculer les v_j^m où $m < m_0$. Procédons à la construction indiquée dans la démonstration du théoreme 4, en ne tenant pas compte des conditions où interviennent des longueurs de configuration $> m_0$. La construction est encore possible, et se fait exactement de la même manière, ce qui montre que le problème traité page 11 dans le cas où les v_i^m sont différents de zéro est soluble également si certains des v_i^m sont nuls.

Remarque II. — L'existence d'une suite x telle que les configurations s'y présentent avec des fréquences égales à des valeurs donnees, sous la seule condition que ces valeurs satisfassent aux relations (10) (§ 3), montre que ces relations sont les seules qui existent entre les fréquences des différentes configurations.

Nous terminerons ce paragraphe par l'énoncé d'une proposition qui nous sera utile par la suite.

THÉORÈME 5. — *Si, dans une suite x, m étant un entier donné, les limites*

$$\lim_{n=\infty} \frac{[u_{2i}^m]_n^x}{[u_{2i+1}^m]_n^x} \qquad (i = 0, 1, \ldots, 2^{m-1} - 1)$$

existent et sont finies positives, il en est de même des limites

$$\lim_{n=\infty} [u_j^m]_n^x \qquad (j = 0, 1, \ldots, 2^m - 1).$$

Démonstration. — Posons

$$\gamma_i = \lim_{n=\infty} \frac{[u_{2i}^m]_n^x}{[u_{2i+1}^m]_n^x}.$$

Les γ_i sont des nombres finis et positifs par hypothèse. D'après le lemme 4 (p. 13), le système (Σ_1) admet, pour les valeurs γ_i des paramètres, un système unique de solutions, soit

$$v_j^m (j = 0, 1, \ldots, 2^m - 1), \qquad \text{avec } v_j^m > 0, \qquad \sum_j v_j^m = 1.$$

Nous sommes dans les conditions d'application du lemme 3 (§ 4); donc nous avons

$$\lim_{n=\infty} [u_j^m]_n^x = v_j^m > 0 \qquad (0 \leqq j < 2^m).$$

6. Généralisation : suites dont les termes peuvent prendre k valeurs différentes.

Soit une suite

$$x = x^1 x^2 \ldots x^n \ldots,$$

dont les termes peuvent prendre k valeurs différentes, que nous désignerons par $1, 2, \ldots, k$. Les relations (8) et (8′) (p. 4) se généralisent facilement. Considérons une configuration $(a^1 a^2 \ldots a^m)$ $(a^j = 1, 2, \ldots, k)$ et sa fréquence dans la suite des n premiers termes de x, soit $[a^1 a^2 \ldots a^m]_n^x$ cette fréquence. Nous avons alors

$$[a^1 a^2 \ldots a^m]_n^x - \sum_{j=1}^{k} [a^1 a^2 \ldots a^m j]_n^x = 0 \quad \text{ou} \quad \frac{1}{n}$$

et

$$[a^1 a^2 \ldots a^m]_n^x - \sum_{j=1}^{k} [j a^1 a^2 \ldots a^m]_n^x = 0 \quad \text{ou} \quad \frac{1}{n}.$$

Si les fréquences totales des configurations considérées existent, nous aurons donc (en désignant par $[a^1 a^2 \ldots a^m]^x$ la limite, pour n infini, de $[a^1 a^2 \ldots a^m]_n^x$)

$$(1) \qquad [a^1 a^2 \ldots a^m]^x = \sum_{j=1}^{k} [a^1 a^2 \ldots a^m j]^x = \sum_{j=1}^{k} [j a^1 a^2 \ldots a^m]^x.$$

Bornons-nous, pour simplifier, aux configurations de longueur 1 (qui sont les termes 1, 2, ..., k eux-mêmes) et de longueur 2 :

$$[i]^x = \sum_{j=1}^{k} [ij]^x = \sum_{j=1}^{k} [ji]^x. \tag{2}$$

Nous allons donner une application simple des considérations précédentes. Soit une chaîne de Markoff constante à k éléments; supposons que nous soyons dans le cas régulier. Nous avons alors par exemple une suite de variables aléatoires

$$X_1, \quad X_2, \quad \ldots, \quad X_n, \quad \ldots \qquad (X_i = 1, 2, \ldots, k) \tag{3}$$

et un tableau carré $\|p_{ij}\|$ de probabilités positives, telles que

$$p_{ij} = \text{Pr.}\left\{ X_n = j, \text{ si } X_{n-1} = i \right\} > 0.$$

Les p_{ij} satisfont naturellement à

$$\sum_{j=1}^{k} p_{ij} = 1.$$

Si la probabilité $p_i^{(n)}$ que $X_n = i$ tend vers une limite quand n tend vers l'infini, soit p_i cette limite, les limites p_i satisfont au système

$$p_i = \sum_{j=1}^{k} p_j p_{ji} \qquad \text{et} \qquad \sum_{j=1}^{k} p_i = 1 \qquad (i = 1, 2, \ldots, k;\ p_i \geqq 0). \tag{4}$$

Dans l'hypothèse où tous les p_{ji} sont positifs, on montre que le système (4) aux inconnues p_i *a une solution unique* (Fréchet, **3**, p. 45). Considérons une valeur i quelconque ($1 \leqq i \leqq k$), et de la suite

$$x = x_1, \ x_2, \ \ldots, \ x_n, \ \ldots \tag{5}$$

des résultats obtenus en déterminant la suite (3), ne conservons que les termes x_n tels que $x_{n-1} = i$. La suite partielle ainsi obtenue peut être considérée comme une suite de résultats d'épreuves indépendantes; soit

$$x_{n_1}, \quad x_{n_2}, \quad \ldots, \quad x_{n_m}, \quad \ldots \tag{6}$$

cette suite. Nous avons

$$\text{Pr.}\left\{ X_{n_j} = s \right\} = p_{is} \qquad (s = 1, 2, \ldots, k).$$

Aux $n+1$ premiers termes de (5) correspondent les $n[i]_n^x$ premiers termes de (6), parmi lesquels les termes égaux à $1, 2, \ldots, k$ figurent respectivement un nombre de fois égal à

$$(n+1)[i1]_{n+1}^x, \quad (n+1)[i2]_{n+1}^x, \quad \ldots, \quad (n+1)[ik]_{n+1}^x.$$

Si donc $\lim\limits_{n=\infty} n[i]_n^x = \infty$, en appliquant la loi forte des grands nombres à la suite (6), nous établissons que, presque certainement,

$$(7) \qquad \lim_{n=\infty} \frac{n+1}{n} \frac{[ij]_x^n}{[i]_n^x} = p_{ij} \qquad \left(\text{si } \lim_{n=\infty} n[i]_n^x = \infty.\right)$$

Or, $\lim\limits_{n=\infty} n[i]_n^x = \infty$ signifie que dans (5) il y a une infinité de termes égaux à i; on montre facilement que dans le cas considéré cela a lieu pour $i=1, 2, \ldots, k$. En effet, cela a lieu au moins pour une valeur de i, soit $i=i_0$. En tenant compte de (7), on voit alors que toutes les configurations $(i_0 j)$ sont répétées une infinité de fois, et par conséquent tous les termes $j=1, 2, \ldots, k$. Ceci posé nous pouvons donc admettre (7) sans tenir compte de la restriction entre parenthèses.

Revenons aux p_i; supposons que l'on n'ait pas

$$(8) \qquad \lim_{n=\infty} [i]_n^x = p_i.$$

Il existe alors des quantités δ_i non toutes nulles, telles que l'on puisse trouver des valeurs de n telles que les quantités

$$(9) \qquad | [i]_n^x - p_i - \delta_i |$$

soient arbitrairement petites, et par conséquent, en tenant compte de (7), que les quantités

$$| [ij]_n^x - p_{ij}(p_i + \delta_i) |$$

soient arbitrairement petites. En sommant par rapport aux i, nous obtenons des expressions telles que

$$\left| [j]_n^x - \sum_{i=1}^{k} p_i p_{ij} - \sum_{i=1}^{k} p_{ij} \delta_i \right|,$$

que, par un choix convenable de n, nous pouvons rendre aussi petites que nous voulons, et en tenant compte de (9), nous aurons, avec telle approximation qui nous plaîra,

$$p_j + \delta_j \sim \sum_{i=1}^{k} (p_i + \delta_i) p_{ij},$$

ce qui montre que le systeme $p_j + \delta_j$ est solution du système (4). Nous sommes en contradiction avec le fait que le système (4) admet une solution unique. Donc (8) a lieu.

Nous concluons donc : dans le cas régulier, la convergence presque certaine dans une chaîne de Markoff est une conséquence *arithmétique* de la convergence presque certaine dans le cas de Bernoulli.

CHAPITRE II.

COLLECTIFS ET SÉLECTIONS.

SOMMAIRE. — 1. Définition des collectifs de M. de Misès. — 2. Définition des collectifs de M. Wald. — 3. Champ des événements probabilisables. — 4. Irrégularité de la suite constituant un collectif. — 5. Composition des sélections. — 6. Construction d'un collectif. — 7. Collectifs indépendants. — 8. Conséquences de la forme de l'irrégularité imposée par les sélections.

1. Définition des collectifs de M. de Misès. — La notion de collectif a été introduite par M. de Misès dans ses recherches sur les fondements du Calcul des Probabilités (de Misès : 1, 2, 3). Nous nous reporterons dans ce qui suit à son *Traité* (de Misès, 3), dont voici trois extraits que le lecteur retrouvera aux pages 9, 12 et 14 :

L'objet du Calcul des Probabilités est constitué par des événements collectifs ou par des processus de répétition ; ces derniers consistent en un très grand nombre d'éléments successifs (observations), dont chacun donne lieu à l'apparition d'un résultat (résultat d'observation) caractérisé par un nombre (ou un groupe de nombres) ; ces résultats se présentent en « succession irrégulière ». Les observations sont effectuées d'après des règles précises, conduisant à la détermination sans ambiguïté du résultat, ou, éventuellement, au rejet de l'observation.

La notion de « succession irrégulière » est imprécise. Voici comment M. de Misès la définit, à l'aide de la notion de sélection. Considérons le cas simple de l'alternative, c'est-à-dire : supposons que les observations ne puissent donner lieu qu'à l'apparition de deux résultats différents. La suite des résultats peut alors s'écrire sous la forme d'une suite de 0 et de 1. M. de Misès définit une sélection comme suit :

D'une suite indéfinie d'observations, nous extrayions ***une*** *suite partielle par* « *sélection* » *si nous nous donnons un procédé qui décide de l'appartenance ou de la non-appartenance de la* $n^{\text{ième}}$ *observation* ($n = 1, 2, \ldots$) *à la suite partielle. indépendamment du résultat de cette n observation, en tenant compte au plus de la connaissance des résultats des observations précédentes.*

Une fois cette définition posée, M. de Misès passe à la définition du *collectif le plus simple*, ou *alternative :*

Nous désignerons par collectif le plus simple, ou alternative, une suite infinie d'observations de même nature, dont les résultats peuvent à chaque fois être représentés par deux signes, soient « 0 » *et* « 1 », *si les conditions suivantes sont remplies :*

Première condition : n_0 *et* n_1 *étant le nombre de celles parmi les n premières observations qui ont fourni les résultats* « 0 » *et* « 1 » *respectivement, les deux limites suivantes existent :*

$$\lim_{n=\infty} \frac{n_0}{n} = p, \qquad \lim_{n=\infty} \frac{n_1}{n} = q.$$

Deuxième condition : *Si l'on extrait de la suite totale* ***une*** *suite partielle par* « *sélection* », *les limites correspondantes existent dans la suite partielle et gardent les mêmes valeurs :*

$$\lim_{n'=\infty} \frac{n'_0}{n'} = p, \qquad \lim_{n'=\infty} \frac{n'_1}{n'} = q.$$

Enfin, voici la définition de la *probabilité.*

Nous appellerons les valeurs limites p (et q) les probabilités d'apparition du résultat 0 (*et* 1) *à l'intérieur du collectif considéré.*

La théorie de M. de Misès a soulevé de nombreuses objections, que l'on peut ranger en deux catégories principales : les objections de la première catégorie s'attaquant à la contradiction *interne* de la théorie ; les objections de la deuxième catégorie reprochent à M. de Misès d'exclure de ses considérations les suites d'observations telles que la fréquence des résultats présentant une certaine caractéristique n'ait pas de limite, suites qui sont *logiquement possibles*. Nous n'insisterons pas sur ces dernières objections, qui sont d'ordre philosophique : une théorie non contradictoire ne peut être attaquée que sur le chapitre des vérifications expérimentales, et il est évident que l'on ne peut pas décider expérimentalement si les suites d'observations réelles sont ou non des collectifs.

Les objections de la première théorie ont été exprimées d'une manière très complète par M. A. Wald, qui a en même temps proposé une modification aux définitions de M. de Misès rendant sa théorie non contradictoire. Sur le cas simple de l'alternative, nous pouvons montrer en quoi la théorie de M. de Misès prête à la critique.

Étant donné une suite d'entiers croissants $\{n_i\}$, le procédé, qui, d'une suite

$$x = x^1 x^2 \ldots x^n \ldots \qquad (x^n = 0 \text{ ou } 1),$$

extrait la suite partielle

$$y = x^{n_1} x^{n_2} \ldots x^{n_m} \ldots,$$

est bien une sélection. Or, si nous considérons *toutes* les suites $\{n_i\}$ possibles, quelle que soit x contenant une infinité de 1 et une infinité de 0, il existera une suite $\{n_i\}$ telle que l'on ait identiquement

$$x^{n_i} = 1,$$

et une autre suite $\{m_i\}$ telle que l'on ait identiquement

$$x^{m_i} = 0.$$

On en conclut qu'il n'existe aucune suite satisfaisant à la définition du collectif, telle que l'a présentée M. de Misès avec $0 < p < 1$. Nous ne faisons qu'indiquer ce point sommairement en renvoyant à l'article de M. Wald; nous passons maintenant à la définition posée par ce dernier.

2. **Définition des collectifs de M. Wald.** — Nous considérons dans ce qui suit un ensemble M fini ou infini, dénombrable ou

non, de points m; chacun des points m représente un état d'un certain système. M représente alors l'ensemble des états possibles sous certaines hypothèses : nous l'appellerons *espace représentatif*.

Si, par exemple, il s'agit d'étudier le phénomène : apparition de pile ou face, l'ensemble M est constitué par deux points, correspondant à pile et face ; si l'on considère le phénomène : apparition d'un point donné quand on jette un dé, M est constitué par 6 points correspondant aux 6 cas possibles ; si l'on étudie un système dont l'état peut être caractérisé par un nombre réel (ce qui correspond à une variable aléatoire dans la théorie classique), M sera constitué par l'ensemble (ou un sous-ensemble) de l'ensemble des nombres réels.

La suite des résultats d'une suite d'observations portant sur le système considéré sera représentée par une suite de points de M :

$$\{m_i\} = m_1, m_2, \ldots, m_n, \ldots. \tag{1}$$

Dans les trois exemples cités ci-dessus, $\{m_i\}$ sera respectivement : une suite de P et F, une suite d'entiers pouvant prendre les valeurs 1 à 6, une suite de nombres réels (distincts ou non).

Un collectif sera une suite $\{m_i\}$ satisfaisant à certaines conditions que nous allons indiquer.

Il faut exprimer sous forme mathématique l'opération de sélection, qui consiste à extraire d'une suite donnée une suite partielle, le sort de chaque terme de la suite donnée (à savoir, être conservé ou rejeté) ne dépendant que des valeurs des termes précédents. Si un individu veut fixer à l'avance, sans savoir quelle est la valeur des termes de la suite qui va lui être présentée, la méthode d'après laquelle il opérera la sélection, il peut recourir au procédé suivant : il inscrira sur un carnet tous les débuts de suite logiquement possibles ; s'il s'agit, par exemple, d'une suite ne devant comprendre que des 0 et des 1, il aura une liste de débuts de suite de la forme suivante :

0; 1; 00; 01; 10; 11; 000; 001; 010; 011; ....

Nous n'insistons pas sur la manière dont il poursuivra cette liste, elle devra comprendre tous les arrangements avec répétition des termes 0 et 1. Sous chacun des termes de la liste ainsi dressée, qui correspondent chacun à un début de suite possible, il inscrira, arbitrairement, mais une fois pour toutes, un signe qui indiquera sa conduite; il pourra inscrire, par exemple, « oui » ou « non » sous chaque terme. « Oui » inscrit sous le terme 010 signifierait : « Si les trois premiers termes de la suite que l'on me présente sont 0, 1, 0 (dans cet ordre), je conserverai le terme suivant (le quatrième) ». « Non » signifierait qu'on rejette ce terme.

Il est plus commode, pratiquement, d'inscrire 1 et 0 au lieu de oui et non respectivement; en exprimant le procédé précédent sous une forme mathématique, nous obtenons pour la « fonction de sélection » la définition de M. Wald, qui suit :

Définition 1. — *Toute fonction $f_n(t_1, t_2, \ldots, t_n)$ qui associe à tout groupe ordonné de n points dans* M *une valeur numérique égale à* 0 *ou à* 1 *sera dite une fonction de sélection à n variables. Pour $n = 0$, f_0 sera une constante.*

Définition 2. — *Une sélection est une opération qui associe à toute suite $\{m_i\}$ une suite partielle $\{m_{i_j}\}$ par le procédé suivant :*

Soit $\{f_n\} = f_0, f_1, f_2, \ldots$ une suite infinie de fonctions de sélection, f_0 étant constante, $f_1, f_2, \ldots$ étant à une, deux, ... variables respectivement. Formons la suite

$$(2) \qquad \{y_i\} = y_1, y_2, \ldots, y_n, \ldots; \qquad y_i = f_{i-1}(m_1, m_2, \ldots, m_{i-1})$$

et soit la suite des rangs, dans $\{y_i\}$, des y_i égaux à 1

$$(3) \qquad \{i_j\} = i_1, i_2, \ldots, i_j, \ldots.$$

La suite $\{m_{i_j}\}$ sera dite extraite de $\{m_i\}$ par la sélection définie par la suite $\{f_n\}$.

Notation. — *Si nous désignons une sélection par une lettre, soit* S, *nous représenterons la suite, extraite de $\{m_i\}$ par la sélection* S, *par la notation* S$\{m_i\}$.

Exemple : Supposons que l'on étudie la probabilité d'amener un point donné avec un dé; si, de la suite $\{m_i\}$ des résultats, nous ne gardons que la suite des résultats tels que le résultat précédent soit 5, nous opérons une sélection. Dans ce cas,

$$\begin{aligned} f_n(m_1, m_2, \ldots, m_n) &= 0 \quad \text{si } m_n \neq 5, \\ \text{»} \qquad &= 1 \quad \text{si } m_n = 5. \end{aligned}$$

La définition que nous venons de donner de la sélection satisfait aux conditions exigées par M. de Misès, puisque la fonction dont la valeur décide de la présence de m_i dans $\{m_{ij}\}$ est $f_{i-1}(m_1, m_2, \ldots, m_{i-1})$. La définition que M. Wald introduit ensuite est celle de la probabilité d'un sous-ensemble de M; soit L un tel sous-ensemble : *la probabilité de* L *dans une suite* $\{\mu_i\}$ *formée de points de* M sera égale à la valeur limite pour n infini, si elle existe, de la fréquence parmi les n premiers points de $\{\mu_i\}$, de ceux qui appartiennent à L; si cette valeur limite n'existe pas, la probabilité de L dans $\{\mu_i\}$ ne sera pas définie.

Voici quelle est l'interprétation intuitive de la probabilité d'un sous-ensemble de M : le point courant m dans M définit les différents états possibles du système; donc l'expression « m appartient au sous-ensemble L » signifie : « le système prend un état appartenant à une certaine classe d'états possibles ». Reprenons l'exemple du dé, et considérons, dans M (formé des points 1, 2, 3, 4, 5, 6), le sous-ensemble (1, 3, 5). Il correspond à l'événement : « il apparaît un point impair ». L'expression « probabilité d'un sous-ensemble L » doit donc être interprétée comme « probabilité de l'événement : le système prend un état dont le point représentatif est un point de L ».

Sous la première forme modifiée par M. Wald [1], la définition du collectif devient alors :

Définition 3. — *Soient :*

a. $\mathfrak{S}$ *un système de sélections, contenant la sélection qui laisse toute suite inchangée.*

b. $\mathfrak{L}$ *un système de sous-ensembles de* M.

Avec ces données : une suite $\{m_i\}$ *de points de* M *sera par définition un collectif relativement à* $\mathfrak{S}$ *et à* $\mathfrak{L}$ *si, quels que soient la sélection* S *appartenant à* $\mathfrak{S}$, *et l'ensemble* L *appar-*

---

[1] La première modification apportée par M. Wald a pour but unique de préciser les définitions de M. de Misès, de manière à rendre leur discussion possible. La deuxième modification (p. 28), a pour but de les rendre non contradictoires.

tenant à $\mathcal{L}$, la probabilité de L *dans* S $\{m_i\}$ *est définie et égale à une fonction* P(L) *de* L *ne dépendant pas de* S (1).

S*i* $\{m_i\}$ *est un tel collectif*, P (L) *sera, par définition, la probabilité de* L *dans le collectif* $\{m_i\}$.

NOTATION. — *Une suite* $\{m_i\}$ *qui est un collectif relativement à* $\mathfrak{S}$ *et à* $\mathcal{L}$ *sera quelquefois représentée par une notation du type* K($\mathfrak{S}$, $\mathcal{L}$).

Le mérite de M. Wald est d'avoir restreint l'ensemble des sélections qui interviennent dans la définition; plus précisément, il suppose toujours $\mathfrak{S}$ dénombrable. Comme d'autres auteurs l'avaient fait avant lui, indépendamment de la notion de collectif, il a également restreint l'ensemble des ensembles pour lesquels la probabilité est définie. Il faut remarquer d'ailleurs que la théorie de M. de Misès ayant pour but les applications pratiques, la limitation des S et des L pouvait être sous-entendue, puisque seuls peuvent intervenir dans les applications une infinité dénombrable de S et de L; en d'autres termes : on ne procède jamais qu'à une infinité dénombrable de sélections, et l'on ne considère jamais qu'une infinité dénombrable d'événements. Pour $\mathfrak{S}$, nous retrouvons l'hypothèse de M. Wald; pour $\mathcal{L}$, nous allons voir que l'hypothèse de M. Wald est moins restrictive que celle qui consiste à supposer $\mathcal{L}$ dénombrable.

Avec cette limitation, la notion de collectif n'est plus contradictoire; en effet, M. Wald a démontré dans son article le théorème suivant :

HYPOTHÈSES : *a* : M *est un ensemble de puissance quelconque.*

b : $\mathcal{K}$ *est un corps dénombrable de sous-ensembles de* M, *contenant* M *lui-même.*

c : μ *est une fonction additive d'ensemble, non négative, définie sur les éléments de* K, *avec* $\mu(\text{M}) = 1$.

d : $\mathcal{L}$ *est le système des sous-ensembles* L *de* M *mesurables par le procédé de Jordan, à l'aide de la fonction* μ *et des éléments de* K; *soit* μ(L) *cette mesure.*

e : $\mathfrak{S}$ *est un système dénombrable de sélections.*

---

(1) Il faut ajouter que cette condition n'est exigée que si S $\{m_i\}$ est une suite infinie. Si S $\{m_i\}$ est finie, cette dernière suite n'est astreinte à aucune condition.

Conclusion. — *Il existe une infinité de collectifs* $K(\mathfrak{S}, \mathcal{L})$, *tels que la probabilité de* L *dans chacun d'eux soit* $\mu(L)$.

La notion de collectif est ainsi *relativisée*, c'est-à-dire : on ne définit pas un collectif d'une manière absolue, mais, au contraire, la définition fait intervenir des données, qui jouent pour ainsi dire le rôle de paramètres ; ces données sont les sélections du système dénombrable $\mathfrak{S}$. D'après le théorème que nous venons d'énoncer, cette notion de collectif relativisée est non contradictoire. Nous allons examiner quelle est l'importance des restrictions apportées par M. Wald ; elles sont de deux sortes : les premières portent sur le champ de définition de la mesure $\mu(L)$, ou, pour employer une expression analogue à celle en usage dans la théorie de la mesure, sur le *champ des événements probabilisables* (Fréchet, 5) ; les secondes portent sur le système des sélections, c'est-à-dire sur la définition de l'*irrégularité* de la suite constituant le collectif.

3. Champ des événements probabilisables. — Si M est fini, il ne se présente aucune difficulté : on peut prendre pour $\mathcal{L}$ l'ensemble de tous les sous-ensembles de M. Si M est infini, il peut être nécessaire de spécifier l'ensemble $\mathcal{L}$ qui figure dans la définition 3. M. Wald a démontré (Wald, 1, p. 45), que si M est dénombrable, on peut prendre pour $\mathcal{L}$ le système de tous les sous-ensembles de M, à condition que la somme des probabilités des différents points de M soit égale à 1.

Quant au cas où M n'est pas dénombrable, nous nous bornerons au cas où M est l'ensemble des nombres réels. Il résulte des démonstrations de M. Wald, que si l'on prend pour $\mathcal{L}$ le système de *tous* les sous-ensembles de M, il n'existe aucun collectif relativement à $\mathcal{L}$, et cela quel que soit $\mathfrak{S}$. C'est pourquoi M. Wald se borne aux ensembles mesurables au sens de Jordan (Wald, 1 ; th. III, p. 46).

Mais cette restriction ne doit pas être considérée comme faite uniquement en vue des applications pratiques. Elle est imposée par la nature de la définition de la probabilité à l'aide d'un collectif. Nous allons montrer qu'il n'est pas nécessaire de supposer que $\mathcal{L}$ contient *tous* les sous-ensembles de M pour conclure à l'inexistence d'un collectif ; nous utiliserons pour cela (p. 32), le théorème suivant, qui est un peu plus précis qu'un théorème énoncé par M. Copeland (Copeland [2], p. 339, lignes 2 et 3).

THÉORÈME 1. — *Si* M *est l'ensemble des nombres réels* (*ou des points sur un axe*) *et* $\mathfrak{L}$ *le système des* G_δ $\left(\textit{ensembles de la forme } \prod_{i=1}^{\infty} O_i \textit{ où } \{O_i\} \textit{ est une suite d'ensembles ouverts}\right)$, *il n'existe aucun collectif, relativement à* $\mathfrak{L}$, *qui ne présente la particularité suivante : il existe parmi les ensembles constituant* $\mathfrak{L}$ *une infinité dénombrable d'ensembles réduits chacun à un point, tels que la somme de leurs probabilités dans le collectif soit égale à* 1.

Nous allons essayer de montrer quel est le sens intuitif de cette proposition. Dans le calcul des probabilités classique, si l'on considere une variable aléatoire X telle que, par exemple,

$$(\alpha) \qquad \text{Pr.}\{X < x\} = x \qquad (0 \leqq x \leqq 1),$$

tout événement de la forme $X \varepsilon L$, où L est un ensemble mesurable, au sens de Lebesgue, a une probabilité définie. Dans la théorie des collectifs, si je suppose que tout événement $X \varepsilon L$, où L est un G_δ, a une probabilité définie, je suis obligé de supposer que la fonction de répartition de X est une fonction se réduisant à une *fonction de sauts*. (Pour la définition précise d'une telle fonction, consulter, par exemple, FRÉCHET, *Recherches théoriques modernes sur le Calcul des Probabilités*, Premier Livre, page 271.)

Le résultat de M. Copeland était, avec les mêmes hypothèses que les nôtres, que la fonction de répartition de X ne pouvait être absolument continue. Passons maintenant à la démonstration du théorème 1.

Démonstration. — Remarquons tout d'abord qu'un ensemble réduit à un seul point et un ensemble dénombrable sont des G_δ.

Soit $\{m_i\}$ une suite de points; supposons qu'elle mette le théoreme en défaut. Rangeons les points de $\{m_i\}$ en une suite $\{a_j\}$ de points *distincts*. Si nous désignons par $f_j^{(n)}$ la fréquence de a_j dans les n premiers termes de $\{ m_i \}$ la limite

$$(2) \qquad p_j = \lim_{n=\infty} f_j^{(n)}$$

existe par hypothèse, quel que soit j, et par ailleurs $\sum_{j=1}^{\infty} p_j < 1$. Soit M_1 l'ensemble des points a_j. Puisque $\Sigma p_j < 1$, d'après le résultat de M. Wald, déjà cité (Wald, 1, p. 45), il n'existe aucun collectif, relativement au système $\mathcal{L}_1$ de tous les sous-ensembles de M_1, qui attribue à a_j la probabilité p_j. En particulier, $\{ m_i \}$ n'est pas un tel collectif. En supposant que le système des sélections est réduit à la seule sélection laissant toute suite inchangée, ce que nous pouvons faire puisque le résultat de M. Wald est indépendant des sélections choisies, nous déduirons de ce que $\{ m_i \}$ n'est pas un collectif relativement à $\mathcal{L}_1$ l'existence d'un sous-ensemble A de M_1 tel que la limite

$$(3) \qquad \lim_{n=\infty} \frac{1}{n} \sum_{i=1}^{n} \varphi_A(m_i) \qquad (\varphi_A = \text{fonction caractéristique de A})$$

n'existe pas. Revenons à M : A qui est un G_δ devrait avoir par hypothèse une probabilité bien définie, c'est-à-dire que la limite (3) devrait exister. Il y a contradiction, ce qui démontre le théorème 1.

Si donc un collectif existe relativement à $\mathcal{L}$, système des G_δ (ou, *a fortiori*, système contenant le système de G_δ), il existe un ensemble dénombrable D de points a_j figurant chacun dans le collectif avec une fréquence p_j, la somme Σ_{p_j} étant égale à 1. Tous les collectifs de même nature que le collectif considéré, c'est-à-dire attribuant aux G_δ les mêmes probabilités, ne contiendront donc que des points de D et des points dont l'ensemble forme un ensemble de probabilité nulle. En fait, la loi de probabilité ne sera définie que sur D. En dehors de ce cas, nous voyons bien qu'il n'existe pas de collectifs relativement à un système contenant les G_δ dans M. Or, s'il n'était pas gênant de restreindre l'ensemble de tous les sous-ensembles de E, il est gênant de ne pouvoir considérer tous les G_δ, ou de se limiter au cas exceptionnel où la fonction de la répartition de la probabilité se réduit à une fonction de sauts. Cette exclusion entraîne la conséquence suivante :

Le champ des événements probabilisables dans la théorie du collectif défini par M. Wald est plus restreint que le champ des événements probabilisables dans la théorie classique.

Si, en effet, X est une variable aléatoire qui a une fonction de répartition, la probabilité pour X d'appartenir à un G_δ est toujours bien définie dans la théorie classique.

4. **Irrégularité de la suite constituant un collectif.** — Nous ne traiterons dans ce qui suit que le cas de l'alternative (p. 23). Avec les définitions de M. Wald, dans le cas de l'alternative, $\mathcal{L}$ se réduit à quatre éléments (à savoir : le sous-ensemble vide, le sous-ensemble formé du point o, celui formé du point 1, et celui constitué par les points o et 1. qui coïncide avec M lui-même) et peut être sous-entendu. Nous caractériserons un collectif, dans ce cas simple, par $\mathfrak{S}$ et par la fréquence p d'un résultat, soit l'apparition de 1 si le collectif est une suite de o et de 1 : *nous dirons que le collectif ainsi défini est un collectif du type* K($\mathfrak{S}$; p), ou plus brièvement : un collectif K($\mathfrak{S}$; p) (1). Nous désignerons quelquefois, dans ce qui suit, par K($\mathfrak{S}$; p) l'*ensemble* des collectifs de fréquence totale p relativement au système $\mathfrak{S}$ (dénombrable) de sélections. De sorte que les phrases « x est un collectif K($\mathfrak{S}$; p) » et « x est un élément de l'ensemble K($\mathfrak{S}$; p) » seront équivalentes.

En quoi consiste l'irrégularité imposée par l'intervention de $\mathfrak{S}$? Soit la suite, où o et 1 alternent régulièrement

$$a = x^1 x^2 \ldots x^n \ldots = 1010 \ldots \tag{4}$$

La fréquence des 1 dans les n premiers termes a pour limite $\frac{1}{2}$ quand n tend vers l'infini. Considérons la sélection S_0 définie par la suite $\{f_i\}$, où les f_i sont des constantes, avec $f_i = \frac{1}{2}[1 + (-1)^{n+1}]$. La suite $\{f_i\}$ s'écrit

$$\{f_i\} = f_0, f_1, \ldots = 0, 1, 0, 1, \ldots \tag{5}$$

et, par conséquent, la suite $S_0 a$ ne contient plus le terme 1. La suite *a n'est pas* un collectif $K\left(\mathfrak{S}; \frac{1}{2}\right)$ si $\mathfrak{S}$ contient la sélection S_0;

(1) « K($\mathfrak{S}$; p) » joue ainsi le rôle d'une épithète. On dit un collectif K($\mathfrak{S}$; p) au lieu d'un collectif relativement à $\mathfrak{S}$, de fréquence p, comme on dirait une fonction I^α au lieu de : une fonction dont la puissance α est intégrable.

ou encore : en astreignant une suite à être un collectif d'un type convenablement choisi, *nous pouvons lui interdire d'avoir la forme trop régulière* (4). C'est parce que nous pouvons ainsi astreindre les suites à ne pas présenter certaines régularités que, dans la définition du collectif par M. de Misès, la condition qui fait intervenir les sélections est appelée *condition* (ou *axiome*) *d'irrégularité* (Regellosigkeitsaxiom).

Essayons de préciser cette notion : excluons les suites qui ne contiennent qu'un nombre fini de o, et représentons toute suite

$$x = x^1 x^2 \dots x^n \dots \qquad (x^i = 0 \text{ ou } 1) \tag{6}$$

par le point du segment (o. 1) qui a pour abscisse

$$\frac{x^1}{2} + \frac{x^2}{4} + \frac{x^3}{8} + \dots$$

et que nous appellerons le point x. La correspondance entre le point x et la suite est biunivoque. Dire que x présente une régularité d'un certain type, c'est dire que le point x appartient à un certain ensemble A du segment (o, 1). Désignons l'ensemble des points représentant des collectifs relativement à $\mathfrak{S}$, où la fréquence de 1 est p, par la même notation que l'ensemble de ces collectifs eux-mêmes, c'est-à-dire par $K(\mathfrak{S}; p)$. Dire que nous pouvons astreindre x à ne pas présenter la régularité du type considéré, au moyen d'un système de sélections $\mathfrak{S}$. c'est dire que l'ensemble A et l'ensemble $K(\mathfrak{S}; p)$ n'ont pas de point commun. Le problème de l'étude des collectifs relatifs à $\mathfrak{S}$ et de fréquence p se ramène donc à l'étude des ensembles $K(\mathfrak{S}; p)$. Nous utiliserons dans cette étude un résultat bien connu de la théorie moderne des épreuves répétées, qui est le suivant :

Si l'on procède à une infinité d'épreuves indépendantes, l'épreuve de rang i déterminant une variable aléatoire X_i *avec*

$$\Pr.\{X_i = 1\} = p, \qquad \Pr.\{X_i = 0\} = q \qquad (p + q = 1;\ 0 < p < 1)$$

la probabilité (au sens de la théorie classique modernisée) que la fréquence totale de la valeur 1 *dans la suite existe et soit égale à p est égale à un.* (*voir*, par exemple : Fréchet, [2], p. 239).

Soit maintenant une sélection S_0, et soit $\{Y_j\}$ la suite obtenue en appliquant S_0 à $\{X_i\}$. La suite $\{Y_j\}$ est une suite de variables aléatoires; de même le nombre N_n des termes de $\{Y_j\}$ qui correspondent aux n premiers termes de $\{X_i\}$ est une variable aléatoire. Nous avons alors la proposition :

Si l'on a $\Pr.\{\lim_{n=\infty} N_n = \infty\}$, *la fréquence de la valeur* 1 *dans* $\{Y_j\}$ *existe et est égale à* p, *avec une probabilité égale à* 1.

Démonstration. — Considérons les N_n premiers termes de $\{Y_j\}$, correspondant aux n premiers de $\{X_i\}$, et posons

$$(7)\qquad \begin{cases} \alpha_n = Y_1 + Y_2 + \ldots + Y_{N_n}, \\ \beta_n = N_n - \alpha_n. \end{cases}$$

Considérons la fonction

$$(8)\qquad \frac{\alpha_n!\,\beta_n!}{\alpha_n + \beta_n + 1!}\, p^{-\alpha_n} q^{-\beta_n} \qquad (0! = 1).$$

Cette fonction est une fonction des n variables aléatoires X_1, $X_2, \ldots, X_n$. Nous la désignerons par $s_n(X_1, X_2, \ldots, X_n)$. S_0 est définie par une suite $\{f_i\}$; supposons que $f_0 = 1$; dans ce cas

$$N_1 = 1, \qquad Y_1 = X_1, \qquad s_1(X_1) = \frac{X_1}{p}.$$

La valeur moyenne de $s_1(X_1)$ est donc égale à 1

$$(9)\qquad \mathfrak{M}\{s_1(X_1)\} = 1.$$

Si $f_0 = 0$, nous avons $\alpha_1 = \beta_1 = 0$, donc (9) reste vraie.

Considérons maintenant

$$s_{n-1}(X_1, X_2, \ldots, X_{n-1}) \quad \text{et} \quad f_{n-1}(X_1, X_2, \ldots, X_{n-1}).$$

Si $f_{n-1}(X_1, X_2, \ldots, X_{n-1}) = 0$, le terme X_n ne figure pas dans $\{Y_j\}$; donc $N_n = N_{n-1}$ et par conséquent

$$(10)\qquad \begin{cases} \alpha_n = \alpha_{n-1}, \qquad \beta_n = \beta_{n-1}, \\ s_n(X_1, X_2, \ldots, X_n) = s_{n-1}(X_1, X_2, \ldots, X_{n-1}). \end{cases}$$

Si $f_{n-1}(X_1, X_2, \ldots, X_{n-1}) = 1$, X_n figure dans $\{Y_j\}$,

$$N_n = N_{n-1} + 1, \qquad Y_{N_n} = X_n$$

et par conséquent

$$(11)\quad \begin{cases} \alpha_n = \alpha_{n-1} + X_n, \\ \beta_n = N_n - \alpha_n = \beta_{n-1} + 1 - X_n, \\ s_n = \begin{cases} \dfrac{\alpha_{n-1}+1}{N_{n-1}+2}\,\dfrac{s_{n-1}}{p}, & \text{si } X_n = 1, \\ \dfrac{\beta_{n-1}+1}{N_{n-1}+2}\,\dfrac{s_{n-1}}{q}, & \text{si } X_n = 0. \end{cases} \end{cases}$$

Soit $\mathfrak{M}_{n-1}\{s_n\}$ la valeur moyenne de s_n quand $X_1, X_2, \ldots, X_{n-1}$ sont connues. Nous déduisons de (10) et (11) que, dans tous les cas,

$$(12)\qquad \mathfrak{M}_{n-1}\{s_n\} = s_{n-1}.$$

Si nous nous reportons à (9), nous déduisons de (12) que la valeur moyenne de s_n est égale à 1 quel que soit n. Soit maintenant λ un nombre positif quelconque. A la suite de variables aléatoires $\{s_i\}$, substituons une suite $\{\sigma_i\}$ ainsi définie :

$$(13)\quad \begin{cases} \sigma_1 = s_1. \\ \sigma_n = \begin{cases} s_n & \text{si } \sigma_{n-1} \leqq \lambda, \\ \sigma_{n-1} & \text{si } \sigma_{n-1} > \lambda. \end{cases} \end{cases}$$

On vérifie sans peine que

$$(14)\qquad \mathfrak{M}\{\sigma_1\} = 1, \qquad \mathfrak{M}_{n-1}\{\sigma_n\} = \sigma_{n-1}$$

et, par suite,

$$(15)\qquad \mathfrak{M}\{\sigma_n\} = 1.$$

σ_n étant une quantité non négative, nous avons donc, quel que soit n,

$$(16)\qquad \text{Pr.}\{\sigma_n > \lambda\} \leqq \frac{1}{\lambda}.$$

La probabilité que nous venons d'écrire est une fonction croissante de n, puisque l'événement $\{\sigma_n > \lambda\}$ implique l'événement $\{\sigma_{n+1} > \lambda\}$.

Considérons l'événement qui se produit si la borne supérieure de σ_n, quand n varie de 1 à l'infini, est plus grande que λ, événement que nous noterons

$$\{\text{B sup. } \sigma_n > \lambda\}.$$

Si cet événement se produit, il existe au moins une valeur de n

à partir de laquelle on a constamment $\{\sigma_n > \lambda\}$; par conséquent, en tenant compte de (16), nous avons

$$\mathrm{Pr.}\left\{\mathrm{B\ sup.}\ \sigma_n > \lambda\right\} \leqq \frac{1}{\lambda}.$$

Nous avons donc, en remarquant que, d'après (13) l'événement $\{\mathrm{B\ sup.}\ s_n > \lambda\}$ implique l'événement $\{\mathrm{B\ sup.}\ \sigma_n > \lambda\}$:

(17) $$\mathrm{Pr.}\left\{\mathrm{B\ sup.}\ s_n > \lambda\right\} \leqq \frac{1}{\lambda}.$$

Or

$$\mathrm{Pr.}\left\{\mathrm{B\ sup.}\ s_n = \infty\right\} \leqq \mathrm{Pr.}\left\{\mathrm{B\ sup.}\ s_n > \lambda\right\}$$

quel que soit λ. Donc nous avons

(18) $$\mathrm{Pr.}\left\{\mathrm{B\ sup.}\ s_n = \infty\right\} = 0$$

La quantité $s = \frac{\alpha_n!\,\beta_n!}{\alpha_n + \beta_n + 1!} p^{-\alpha_n} q^{-\beta_n}$ est donc bornée, sauf dans des cas infiniment peu probables. Supposons que $\alpha_n + \beta_n \to \infty$. Si nous montrons que si $\frac{\alpha_n}{\alpha_n + \beta_n}$ ne tend pas vers p, s_n n'est pas bornée, nous déduirons de (18) que

$$\mathrm{Pr.}\left\{\frac{\alpha_n}{\alpha_n + \beta_n} = p\right\} = 1.$$

En modifiant légèrement les notations, nous sommes ramenés à montrer que si, dans l'expression

$$w(x, y) = \frac{x!\,y!}{x + y + 1!} p^{-x} q^{-y},$$

$x + y$ tend vers l'infini sans que $\frac{x + y}{x}$ tende vers p, w n'est pas borné supérieurement.

Supposons d'abord que $\frac{x}{x + y}$ tende vers $\varpi \neq p$, et posons

$$\frac{x}{x + y} = \alpha, \qquad \frac{y}{x + y} = \beta, \qquad x + y = n.$$

Nous pouvons mettre w sous la forme

$$w(x, y) = \int_0^1 \left(\frac{t}{p}\right)^x \left(\frac{1 - t}{q}\right)^y dt = \int_0^1 e^{n\xi}\, dt$$

avec

$$\xi = \alpha \operatorname{Log} \frac{t}{p} - \beta \operatorname{Log} \frac{1 - t}{q}.$$

Par hypothèse, quand n tend vers l'infini, α tend vers ϖ, et β

vers $1-\varpi$; donc $\xi(t)$ tend vers la fonction

$$\xi_0(t) = \varpi \operatorname{Log} \frac{t}{p} + (1-\varpi) \operatorname{Log} \frac{1-t}{q}.$$

Nous avons

$$\xi'_0(t) = \frac{\varpi}{t} - \frac{1-\varpi}{1-t};$$

$\xi_0(t)$ est maximum pour $t=\varpi$, la valeur du maximum étant

$$\xi_0(\varpi) = \varpi \operatorname{Log} \frac{\varpi}{p} + (1-\varpi) \operatorname{Log} \frac{1-\varpi}{q}.$$

Considérons l'expression ci-dessus comme une fonction de ϖ; sa dérivée par rapport à ϖ est

$$\operatorname{Log} \frac{\varpi}{p} - \operatorname{Log} \frac{1-\varpi}{q}.$$

$\xi_0(\varpi)$ est minimum pour $\varpi = p$, la valeur du minimum étant o; donc $\xi_0(\varpi) > 0$ puisque nous avons supposé $\varpi \neq p$. $\xi_0(\varpi)$ étant > 0, et, α tendant vers ϖ, il existe un nombre positif δ, un nombre positif τ et un nombre entier n_0 tel que pour $n > n_0$, on ait $\xi(t) > \delta$ dans l'intervalle $\varpi - \tau$, $\varpi + \tau$.

Nous en déduisons que pour $n > n_0$, nous aurons

$$w(x, y) > \int_{\varpi-\tau}^{\varpi+\tau} e^{n\xi}\, dt > 2\tau\, e^{n\delta},$$

et, par conséquent, si $\alpha \to \varpi$,

$$\lim_{n=\infty} w(x, y) = \infty.$$

Supposons maintenant que α ne tende pas vers p; il existe alors un nombre ϖ différent de p, et une suite $\{n_i\}$ de valeurs de n telles que, si l'on désigne par $\{\alpha_i\}$ la suite des valeurs correspondantes de α, on ait

$$\lim_{i=\infty} \alpha_i = \varpi.$$

Nous en déduisons, tenu compte du résultat précédent, que, lorsque n tend vers l'infini, nous avons

$$\text{borne supérieure } w(x, y) = \infty.$$

La proposition que nous venons de démontrer peut s'énoncer ainsi :

THÉORÈME 2. — *Étant donné un système $\mathfrak{S}$, dénombrable, de* *sélections, si l'on procède à une suite de tirages indépendants pouvant donner lieu à l'apparition d'un événement* E *avec une probabilité (au sens classique) constante et égale à p, il y a une probabilité (au sens classique) égale à* 1 *pour que le résultat de cette suite de tirages soit un collectif relativement à* $\mathfrak{S}$, *collectif dans lequel la probabilité (au sens de la théorie des collectifs) de l'événement* E *est égale à p.*

Pour éviter les confusions possibles, employons le langage de la mesure : soit p un nombre compris entre 0 et 1 ; partageons le segment (0, 1) en deux parties égales, et attribuons la mesure $q = 1 - p$ au segment $\left(0, \frac{1}{2}\right)$, la mesure p au segment $\left(\frac{1}{2}, 1\right)$; partageons ces segments eux-mêmes en deux parties égales, et ainsi de suite : si nous partageons en deux parties égales un segment auquel nous avons attribué une mesure $p^\alpha q^\beta$, nous attribuerons à la moitié de gauche la mesure $p^\alpha q^{\beta+1}$, à la moitié de droite la mesure $p^{\alpha+1} q^\beta$. Nous définissons ainsi une fonction d'ensemble additive, $\geqq 0$, définie sur tous les segments de la forme $[k 2^{-n}, (k+1) 2^{-n}]$. Il existe alors une fonction d'ensemble $\geqq 0$, complètement additive, définie sur tous les ensembles boréliens, qui coïncide avec la fonction d'ensemble que nous venons d'introduire sur tous les segments de la forme $[k 2^{-n}, (k+1) 2^{-n})$. C'est cette dernière fonction d'ensemble que nous appellerons la p-mesure.

Reportons-nous à la représentation des suites par des points du segment (0, 1) (p. 33); on montre sans peine que la p-mesure s'interprète comme une probabilité au sens classique modernisé, à savoir : la p-mesure d'un ensemble est la probabilité que la suite des épreuves (de probabilité constante p) fournisse une suite dont le point représentatif appartienne à l'ensemble. Ceci posé nous voyons que le théorème 2 s'exprime sous la forme :

Théorème 2'. — *Quel que soit le système dénombrable* $\mathfrak{S}$ *de sélections, la p-mesure de l'ensemble* K ($\mathfrak{S}$; p) (défini p. 49) *est égale à* 1.

Si nous revenons à la question de l'irrégularité, nous voyons que nous pouvons énoncer une condition nécessaire pour qu'une suite puisse être astreinte à ne pas présenter une particularité donnée :

THÉORÈME 3. — *Un nombre p étant donné, une condition nécessaire pour que l'on puisse astreindre par une condition d'irrégularité fondée sur la notion de sélection, c'est-à-dire par une condition de la forme* $x \varepsilon \mathrm{K}(\mathfrak{S}; p)$, *une suite x à ne pas appartenir à un ensemble représenté sur le segment* (0, 1) *par un ensemble* A, *est que la p-mesure de* A *soit nulle.*

La réciproque de ce théorème se présente naturellement à l'esprit; mais elle est fausse. En effet :

THÉORÈME 4. — *Soit donnée une valeur p quelconque comprise entre* 0 *et* 1*; on peut lui associer un ensemble* G, *de p-mesure nulle, jouissant de la propriété suivante :*

Quel que soit le système dénombrable $\mathfrak{S}$, *les ensembles* $\mathrm{K}(\mathfrak{S}; p)$ *et* G *ont au moins un point commun.*

Le théorème 4 est une conséquence du théorème suivant :

THÉORÈME 4'. — *Étant donné : un système dénombrable* $\mathfrak{S}$ *de sélections, un nombre* $p(0 < p < 1)$ *et une fonction* $\varphi(\mu) > 0$ *continue, croissante, telle que*

$$\lim_{\mu=\infty} \frac{1}{\mu}\varphi(\mu) = 0 \qquad \text{et} \qquad \lim_{\mu=\infty} \varphi(\mu) = \infty,$$

mais à part cela, à croissance arbitrairement lente :

il existe un collectif $\mathrm{K}(\mathfrak{S}; p)$ *tel que : si l'on considère la fréquence* $\frac{\nu}{\mu}$ *des* 1 *dans les* μ *premiers termes d'une suite infinie extraite du collectif par une sélection du système* $\mathfrak{S}$, *il existe deux nombres positifs* α, β *(dépendant de la sélection choisie, mais indépendants de* μ*) tels que l'on ait, quel que soit* μ :

$$-\frac{\alpha}{\mu} \leqq \frac{\nu}{\mu} - p < \frac{\alpha}{\mu} + \beta\frac{\varphi(\mu)}{\mu}. \tag{19}$$

Montrons d'abord que le théorème 4' implique le théorème 4.

Si nous considérons une suite d'épreuves indépendantes où un événement peut se produire avec une probabilité p, il est bien connu (*voir*, par exemple, Lévy [1], p. 258 et suivantes), que, *si f_μ représente la fréquence des succès dans les μ premières épreuves, il y a une probabilité égale à* 1 *pour que, quel que soit le nombre positif* δ, *l'inégalité*

$$|f_\mu - p| > \frac{\delta}{\sqrt{\mu}} \tag{20}$$

se produise une infinité de fois. Si nous posons donc, dans l'énoncé du théorème 4', $\varphi(\mu) = \mu^\varepsilon \left(0 < \varepsilon < \frac{1}{2}\right)$, et que nous désignions par G l'ensemble des points x correspondant aux suites x pour lesquelles l'inégalité (20) (où nous faisons, par exemple, $\delta = 1$) n'a lieu qu'un nombre fini de fois, nous constatons que G est de p-mesure 0, et a au moins un point commun avec l'ensemble K($\mathfrak{S}$; p), quel que soit $\mathfrak{S}$, ce qui démontre le théorème 4.

Quant au théorème 4', nous le démontrerons en construisant un collectif K($\mathfrak{S}$; p) satisfaisant aux conclusions du théorème. Nous avons besoin pour cette construction de poser quelques définitions préliminaires :

5. **Composition des sélections.** — Définition 4. — *Étant donné deux sélections* A_1 *et* A_2, *définies par les suites de fonctions* $\{f_i^1\}$ *et* $\{f_i^2\}$ *respectivement* (Déf. 2, p. 42), *la sélection* A *définie par la suite* $\{g_i\}$ *où*

$$g_i = 1 - (1 - f_i^1)(1 - f_i^2)$$

sera appelée somme des sélections A_1 *et* A_2 *et nous écrirons*

$$A = A_1 + A_2.$$

La sélection A est la sélection qui choisit un terme si A_1 *ou* A_2 le choisissent, et dans ce cas seulement. De la définition précédente, on déduit sans peine la définition de la sélection

$$A_1 + A_2 + \ldots + A_n \quad \text{et celle de} \quad \sum_{i=1}^{\infty} A_i.$$

Définition 5. — A_1 *et* A_2 *étant deux sélections définies par* $\{f_i^1\}$ *et* $\{f_i^2\}$ *respectivement, la sélection* B *définie par* $\{g_i\}$ *où*

$$g_i = f_i^1 f_i^2$$

sera appelée produit intérieur de A_1 *et* A_2, *et nous écrirons*

$$B = A_1 A_2.$$

B est la sélection qui choisit un terme si A_1 *et* A_2 le choisissent simultanément, et dans ce cas seulement. Les opérations de la somme et du produit intérieur sont commutatives et associatives.

Définition 6. — A_1 *étant une sélection définie par* $\{f_i^1\}$ *la sélection* C *définie par* $\{g_i\}$ *où*

$$g_i = 1 - f_i^1$$

sera dite la sélection complémentaire de A_1 *et nous écrirons*

$$C = 1 - A_1.$$

C est la sélection qui choisit un terme si A_1 ne le choisit pas, et dans ce cas seulement.

Définition 7. — A_1 *étant une sélection et* a *un nombre égal à* 0 *ou à* 1, *nous désignerons par* aA_1 *la sélection identique à* A_1 *si* $a = 1$, *et à* $1 - A_1$ *si* $a = 0$.

Les opérations introduites par les définitions 4 à 7 constituent dans leur ensemble ce que nous appellerons la *composition* des sélections. Nous allons définir encore une opération sur les sélections.

Définition 8. — A_1 *étant une sélection définie par* $\{f_i^1\}$ *et* m *un entier positif, la sélection* E *définie par* $\{g_i\}$ *où*

$$g_i = f_i^1, \qquad \text{si } \sum_{j=0}^{i} f_j^1 \leqq m,$$

$$g_i = 0, \qquad \text{si } \sum_{j=0}^{i} f_j^1 > m$$

sera appelée segment de longueur m *de la sélection* A_1 *et nous écrirons*

$$E = A_1^{(m)}.$$

La sélection E est celle qui ne choisit un terme que si ce terme est choisi par A_1 et si A_1 a choisi moins de m termes de rangs antérieurs.

6. Construction d'un collectif $K(\mathfrak{S};\, p)$. — Passons à la construction d'un collectif qui satisfasse aux conclusions du théorème 4'. Soient donc un nombre p, un système $\mathfrak{S}$ dénombrable de sélections A_1, A_2, ..., A_n, ... et une fonction $\varphi(\mu)$ satisfaisant à l'hypothèse. Nous allons définir de nouvelles sélections, les étudier, construire une certaine suite en nous appuyant sur leurs propriétés, démontrer que cette suite est un collectif relativement aux nouvelles sélections, puis que c'est un collectif relativement à $\mathfrak{S}$.

1° *Composition des sélections* $A_1, A_2, \ldots, A_n$. — L'équation $2^\alpha = \varphi(\mu)$ ou

$$\alpha \operatorname{Log} 2 = \operatorname{Log} \varphi(\mu) \tag{21}$$

définit une fonction inverse $\mu = \psi(\alpha)$; soit $\{m_i\}$ une suite d'entiers telle que

$$m_i > \psi(i+1) - \psi(i). \tag{22}$$

Nous associerons à toute suite finie $a^1 a^2 \ldots a^m$ de 0 et de 1 deux sélections $B_{a^1 a^2 \ldots a^m}$ et $C_{a^1 a^2 \ldots a^n}$ par les opérations suivantes :

$$B_{a^1} = a^1 A_1, \qquad C_{a^1} = B_{a^1}^{(m_1)},$$
$$B_{a^1 a^2} = (1 - C_{a^1})\, a^1 A_1 . a^2 A_2, \qquad C_{a^1 a^2} = B_{a^1 a^2}^{(m_2)}$$

et, d'une manière générale,

$$(23)\quad \begin{cases} B_{a^1 a^2 \ldots a^{n-1} a^n} = (1 - C_{a_1})(1 - C_{a^1 a^2}) \ldots (1 - C_{a^1 a^1 \ldots a^{n-1}})\, a^1 A_1 . a^2 A_2 \ldots a^n A_n \\ C_{a^1 a^2 \ldots a^n} \quad = B_{a^1 a^2 \ldots a^n}^{(m_n)}. \end{cases}$$

Les sélections B sont les nouvelles sélections annoncées. Les sélections C en sont les segments.

2° *Propriétés des sélections* B *et* C. — Soit une suite *quelconque* de 0 et de 1

$$x = x^1 x^2 \ldots x^i \ldots \qquad (x^i = 0 \text{ ou } 1).$$

Portons notre attention sur un terme, soit x^m. *Il ne peut exister deux sélections distinctes* $C_{a^1 a^2 \ldots a^n}$ et $C_{\alpha^1 \alpha^2 \ldots \alpha^\nu}$ *qui choisissent* x^m *en même temps*. Supposons en effet qu'il n'en soit pas ainsi, et considérons ces deux sélections. S'il existe un rang i tel que $a^i \neq \alpha^i$, soit, pour fixer les idées, $a^i = 1$ et $\alpha^i = 0$, nous voyons que, d'après la première des relations (23), x^m devrait être choisi par A_i, puisqu'il est choisi par $C_{a^1 a^2 \ldots a^n}$, et par $1 - A_i$, puisqu'il est choisi par $C_{\alpha^1 \alpha^2 \ldots \alpha^\nu}$, ce qui est contradictoire, d'après la définition 7.

Si un tel i n'existe pas, les sélections ne sont distinctes que si $\nu \neq n$, soit $\nu > n$, et les sélections sont de la forme

$$C_{a^1 a_2 \ldots a^n}, \qquad C_{a^1 a^2 \ldots a^n \alpha^{n+1} \ldots \alpha^\nu}.$$

Toujours, d'après la première des relations (23), nous voyons qu'un terme choisi par $C_{a^1 a^2 \ldots a^n}$ ne peut l'être par $B_{a^1 a^2 \ldots a^n \alpha^{n+1} \ldots \alpha^\nu}$ ni *a fortiori* par $C_{a^1 a^2 \ldots a^n \alpha^{n+1} \ldots \alpha^\nu}$. D'ou contradiction.

Inversement, tout terme de x est choisi par une des sélections $C_{a^1 a^2 \ldots a^n}$. Supposons en effet qu'il n'en soit pas ainsi, et soit x^m un terme qui ne soit choisi par aucune de ces sélections. Quel que soit n, x^m est choisi par A_n ou par $1 - A_n$; soit $a^n A_n$ celle de ces deux dernières sélections qui choisit x^m; considérons la suite $\{ a^n \}$ et la suite correspondante :

$$(24) \qquad C_{a^1}, \quad C_{a^1 a^2}, \quad \ldots . \quad C_{a^1 a^2 \ldots a^n}, \quad \ldots .$$

D'après la première des relations (23), x^m n'étant choisi par aucune des C, est choisi par toutes les sélections $B_{a^1 a^2 \ldots a^n}$ correspondant aux sélections de la suite (24). D'après la seconde relation (23), chacune des $C_{a^1 a^2 \ldots a^n}$. ne choisissant pas x^m, devrait choisir m_n termes parmi les m premiers termes de x, ce qui conduit à contradiction, puisque deux quelconques de ces sélections ne peuvent choisir le même terme et que $\sum_{n=1}^{\infty} m_n = \infty$. Nous venons donc de démontrer que x^m est choisi par une et une seule des sélections de la suite (24), soit $C_{a^1 a^1 \ldots a^n}$; remarquons qu'il est choisi par les sélections $B_{a^1}, B_{a^1 a^2}, \ldots, B_{a^1 a^2 \ldots a^n}$; n'étant pas choisi par $C_{a^1}, C_{a^1 a^2}, \ldots, C_{a^1 a^2 \ldots a^{n-1}}$, nous en déduisons que chacune de ces dernières sélections a déjà choisi $m_1, m_2, \ldots, m_{n-1}$ termes dans x respectivement.

En résumé : *Si nous désignons par* $C_{a^1 a^2 \ldots a^n}(x_m)$ *la suite extraite des m premiers termes de x par la sélection* $C_{a^1 a^2 \ldots a^n}$:

Tout terme, parmi les m premiers de x, appartient à une suite $C_{a^1 a^2 \ldots a^n}(x_m)$ *et à une seule.*

Si la suite $C_{a^1a^2\ldots a^n}(x_m)$ *contient au moins un terme, chacune des suites* $C_{a^1a^2\ldots a^i}(x_m)$ $(i < n)$ *contient exactement* m_i *termes* ($a^1a^2\ldots a^i$ *étant un segment initial de* $a^1a^2\ldots a^n$).

3° *Construction d'une certaine suite x.* — Nous formerons, à partir des sélections B, C, et du nombre p, une suite x en procédant comme suit :

Nous prendrons $x^1 = 1$. D'une manière générale, supposons que l'on ait pu déterminer $x^1, x^2, \ldots, x^m$ de manière que *dans toute suite* $C_{a^1a^2\ldots a^n}(x_m)$ *non vide, le nombre de termes égaux à* 1 *soit égal à la valeur approchée à une unité près par excès du produit par p du nombre de termes contenu dans* $C_{a^1a^2\ldots a^n}(x_m)$. Quelle que soit la valeur de x^{m+1}, nous pouvons déterminer à quelle suite $C_{a^1a^2\ldots a^n}(x_{m+1})$ ce terme appartiendra. Les indices inférieurs de cette suite sont fonctions de $x^1, x^2, \ldots, x^m$. Désignons leur ensemble par (m). Donc, $C_{(m)}$ est celle des sélections $C_{a^1a^2\ldots a^n}$ telle que $C_{(m)}(x_{m+1})$ contienne x^{m+1}. Si $C_{(m)}(x_m)$ est vide, x^{m+1} est le premier terme de $C_{(m)}(x_{m+1})$; nous prendrons $x^{m+1} = 1$. Si $C_{(m)}(x_m)$ contient μ termes, dont ν égaux à 1, nous avons par hypothèse :

$$p\mu \leqq \nu < p\mu + 1.$$

Nous avons alors

$$p(\mu+1) \leqq \nu < p(\mu+1) + 1 \qquad \text{ou} \qquad p(\mu+1) \leqq \nu + 1 < p(\mu+1) + 1.$$

Nous prendrons $x^{m+1} = 0$ dans le premier cas, $x^{m+1} = 1$ dans le second cas. Toutes les suites $C_{a^1a^2\ldots a^n}(x_{m+1})$ jouiront alors de la même propriété que les suites $C_{a^1a^2\ldots a^n}(x_m)$.

Le processus de formation de x se poursuit indéfiniment.

4° *La suite x est un collectif relativement aux sélections* B. — Soit $B_{a^1a^2\ldots a^n}$ une quelconque des sélections B; les a^i étant ainsi choisis, et restant fixes, ainsi que n, au cours de ce § 4°, soit $B_{a^1a^1\ldots a^n}(x_m)$ la suite qu'elle déduit de la suite des m premiers termes de x. Les termes d'une suite de la forme $C_{a^1a^2\ldots a^n\alpha^1\alpha^2\ldots\alpha^s}(x_m)$ appartiennent tous à $B_{a^1a^2\ldots a^n}(x_m)$ et inversement, tout terme de cette dernière suite appartient à une suite et une seule de cette forme. Soit μ le nombre de termes de $B_{a^1a^2\ldots a^n}(x_m)$, parmi lesquels ν égaux à 1 ; soit s_0 la plus grande parmi les valeurs de s pour lesquelles il existe une suite $C_{a^1a^2\ldots a^n\alpha^1\ldots\alpha^s}(x_m)$ non vide; soit $C_{a^1a^2\ldots a^n\beta^1\beta^2\ldots\beta^{s_0}}(x_m)$ une de ces suites; les suites

$$C_{a^1a^2\ldots a^n}(x_m), \quad C_{a^1a^2\ldots a^n\beta^1}(x_m), \quad \ldots, \quad C_{a^1a^2\ldots a^n\beta^1\ldots\beta^{s_0-1}}(x_m)$$

contiennent respectivement m_n, m_{n+1}, ... m_{n+s_0-1} termes; d'où la première relation

$$(25)\qquad \mu \geqq m_n + m_{n+1} + \ldots + m_{n+s_0-1}.$$

Toute suite $C_{a^1a^2\ldots a^n\alpha^1\ldots\alpha^{s_0+1}}(x_m)$ étant vide, et les suites $C_{a^1a^2\ldots a^n\alpha^1\ldots\alpha^s}(x_m)$ étant au nombre de 2^s, nous en déduisons que

$$(26)\qquad \mu \leqq m_n + 2m_{n+1} + \ldots + 2^{n+s_0} m_{n+s_0}.$$

Au cas où $s_0 = 0$, les relations (25) et (26) deviennent $\mu \geqq 0$ et $\mu \leqq m_n$ respectivement.

Soit r le nombre des suites $C_{a^1a^2\ldots a^n\alpha^1\alpha^2\ldots\alpha^s}(x_m)$ non vides (les a^i restant fixes, les α^j variant); de la définition de s_0, nous déduisons

$$(27)\qquad r \leqq 1 + 2 + \ldots + 2^{s_0} < 2^{s_0+1}.$$

Si ces suites non vides contiennent respectivement $\mu_1, \mu_2, \ldots, \mu_r$ termes, dont $\nu_1, \nu_2, \ldots, \nu_r$ égaux à 1 respectivement, nous avons

$$(28)\qquad \begin{cases} \mu_1 + \mu_2 + \ldots + \mu_r = \mu, \\ \nu_1 + \nu_2 + \ldots + \nu_r = \nu. \end{cases}$$

Par ailleurs, de par la manière dont la suite x a été construite, nous avons

$$(29)\qquad p\mu_j \leqq \nu_j < p\mu_j + 1 \qquad (j = 1, 2, \ldots, r),$$

d'où, en faisant la somme,

$$(30)\qquad p\mu \leqq \nu < p\mu + r < p\mu + 2^{s_0+1}.$$

Nous aurons alors, d'après (22) et (25),

$$\mu + \psi(n) > \psi(n + s_0),$$

c'est-à-dire, d'après la définition de ψ,

$$2^{n+s_0} < \varphi[\mu + \psi(n)] = \varphi(\mu + \rho) \qquad [\rho = \psi(n)]$$

et, en tenant compte de (30), la sélection $B_{a^1a^2\ldots a^n}$ restant la même, nous avons :

$$(31)\qquad 0 \leqq \frac{\nu}{\mu} - p < \frac{\varphi(\mu + \rho)}{\mu 2^{n-1}}.$$

Si donc, $\lim\limits_{m=\infty} \mu = \infty$, tandis que n, donc ρ, restent fixes; nous voyons que :

$$\lim_{m=\infty} \frac{\nu}{\mu} = p.$$

5° *La suite x satisfait aux conclusions du théorème 4'.* — Soit A_n une quelconque des sélections du système $\mathfrak{S}$; désignons par $A_n(x_m)$ la suite extraite par A_n de la suite des m premiers termes de x. Les termes de $A_n(x_m)$ appartiennent :

a. soit à des suites $C_{a^1a^2\ldots a^s}(x_m)\,(s<n)$;
b. soit à des suites $B_{a^1a^2\ldots a^{n-1}1}$.

Si $A_n(x)$, suite extraite de x par A_n, est infinie, nous pourrons négliger les termes du type *a*, qui sont en nombre fini, soit N.

Nous écrirons donc

$$A_n = D + \sum B_{a^1a^2\ldots a^{n-1}1}, \tag{32}$$

où D est une sélection choisissant N termes; la sommation est étendue à toutes les valeurs des indices a^1, a^2, ..., a^{n-1} (Définition 4). Soient μ_0, μ_1, ..., μ_s les nombres de termes contenus dans $D(x_m)$, et dans les suites de la forme $B_{a^1a^2\ldots a^{n-1}1}(x_m)$ supposées numérotées de 1 à $s\,(s=2^{n-1})$; soient ν_0, ν_1, ..., ν_s les nombres des termes égaux à 1 dans ces suites respectivement. Si $A_n(x_m)$ contient μ termes, dont ν égaux à 1, nous aurons, en tenant compte de (31),

$$\begin{cases} \mu = \mu_0 + \mu_1 + \ldots + \mu_s & (\mu_0 = N), \\ \nu = \nu_0 + \nu_1 + \ldots + \nu_s & (0 \leqq \nu_0 \leqq \mu_0 = N); \end{cases}$$

$$\begin{cases} p\mu_i \leqq \nu_i < p\mu_i + 2^{1-n}\varphi(\mu_i + \rho) \quad [\rho = \psi(n)], \; (i = 1, 2, \ldots, s), \\ pN \leqq \nu_0 + pN \leqq pN + N, \end{cases}$$

d'où par sommation,

$$p\mu \leqq \nu + pN < p\mu + N + 2^{1-n}\sum_{i=1}^{s}\varphi(\mu_i + \rho),$$

ou encore, puisque φ est une fonction croissante,

$$-\frac{pN}{\mu} \leqq \frac{\nu}{\mu} - p < \frac{qN}{\mu} + \frac{\varphi(\mu+\rho)}{\mu} \qquad (q = 1 - p), \tag{33}$$

c'est-à-dire que, si $\lim\limits_{m=\infty} \mu = \infty$,

$$\lim_{m=\infty} \frac{\nu}{\mu} = p.$$

x est donc bien un collectif $K(\mathfrak{S}; p)$. En outre, (19) est vérifiée en faisant dans (33)

$$\alpha = \mathrm{N} \qquad \text{et} \qquad \beta = \underset{0<\mu<\infty}{\text{Borne sup.}} \frac{\varphi(\mu+\rho)}{\varphi(\mu)}.$$

Le théorème 4' est bien démontré.

Remarque. — Considérons la fréquence des 1 dans la suite x que nous venons de construire. Tout terme, parmi les m premiers de x, appartient à une suite $C_{a^1a^2\ldots a^n}(x_m)$. Or, chacune de ces dernières suites, par construction, contient un nombre de 1 égal à la valeur approchée par excès, à une unité près, du nombre de ses termes. Il en résulte que la fréquence des 1 parmi les m premiers termes de x est toujours $\geqq p$: *on peut, quels que soient $\mathfrak{S}$ et p, construire un collectif où la fréquence tende vers sa limite d'une manière unilatérale.*

7. **Collectifs indépendants.** — Soit à définir non plus la probabilité d'un seul événement, mais les probabilités de deux événements ; si nous procédons à une suite infinie d'épreuves, pouvant donner lieu à l'apparition de ces deux événements, soient E et F, nous pouvons représenter les résultats par une suite de groupes de deux termes :

$$(x, y) = (x^1, y^1), (x^2, y^2), \ldots, (x^n, y^n), \ldots, \tag{35}$$

$x^i = 1$ signifiant que E s'est produit à la i^{ieme} preuve, $x^i = 0$ représentant non E — $y^i = 1$, F — et $y^i = 0$, non F. La suite (x, y), dont les termes sont des points d'un espace formé des quatre points $(0, 0)$, (0.1), $(1, 0)$, $(1, 1)$, doit être un collectif pour qu'on puisse parler de probabilité au sens de M. de Misès. Considérons maintenant les deux suites

$$\left\{\begin{array}{l} x = x^1 x^2 \ldots x^n \ldots, \\ y = y^1 y^2 \ldots y^n \ldots. \end{array}\right. \tag{36}$$

Nous pouvons les considérer comme des collectifs. D'une manière plus précise, et en nous référant aux définitions de M. Wald :

Étant donné deux systèmes dénombrables $\mathfrak{S}_1$ et $\mathfrak{S}_2$ de sélections applicables à des suites formées de 0 *et de* 1, *il existe un système dénombrable $\mathfrak{S}$ de sélections applicables à des suites de la forme* (35), *tel que si (x, y) est un collectif relativement à $\mathfrak{S}$, x et y sont des collectifs relativement à $\mathfrak{S}_1$ et $\mathfrak{S}_2$ respectivement.*

Soit, en effet, A une sélection du système $\mathfrak{S}_1$; soit A′ la sélection, applicable à (x, y), qui choisit (x^n, y^n), connaissant $(x^1, y^1), \ldots, (x^{n-1}, y^{n-1})$ si A choisit x^n connaissant les valeurs $x^1, x^2, \ldots, x^{n-1}$ et dans ce cas seulement. De toute sélection A_1 de $\mathfrak{S}$, nous déduisons une sélection A′; soit $\mathfrak{S}'_1$ l'ensemble des A′; définissons de même $\mathfrak{S}'_2$, il suffit de choisir $\mathfrak{S}$ égal à la somme de $\mathfrak{S}'_1$ et de $\mathfrak{S}'_2$ pour satisfaire à la proposition ci-dessus énoncée.

Dans la terminologie de M. de Misès, deux collectifs x et y tels que la suite (x, y) soit un collectif sont dits *associables* (verbindbar) (de Misès, [3], p. 92). Dans la terminologie de M. Wald. cette notion n'a de sens que si l'on se donne $\mathfrak{S}$:

Définition 9. — *Deux collectifs x et y sont dits associables relativement à un système $\mathfrak{S}$ de sélections si la suite (x, y) est un collectif relativement à $\mathfrak{S}$* (1).

La définition (9) a été introduite dans le cas où x et y sont des alternatives, mais s'étend sans peine au cas où les collectifs sont définis d'une manière quelconque. Il suffit de relativiser à un système dénombrable de sélections les définitions de M. de Misès.

Soient maintenant deux collectifs (alternatives) x et y associables relativement à $\mathfrak{S}$; extrayons de x une suite partielle en ne conservant dans x que les termes x^i tels que $y^i = 1$; soit

$$x_1 = x^{i_1} x^{i_2} \ldots x^{i_n} \ldots \tag{37}$$

la suite ainsi obtenue. Si la limite

$$\lim_{n=\infty} \frac{1}{n} \sum_{j=1}^{n} x^{i_j} \tag{38}$$

existe et est égale à $\lim\limits_{n=\infty} \frac{1}{n} \sum\limits_{j=1}^{n} x^j$, les collectifs x et y seront dits *indépendants* (de Misès, [3], p. 93). La définition de l'indépendance est symétrique; en effet, dans la suite (37) figurent $\sum\limits_{j=1}^{m} y^j$ termes parmi les m premiers de x, de sorte que la limite (38) peut

(1) Cette définition ne figure pas dans les publications de M. Wald.

s'écrire

$$\lim_{m=\infty}\left[\sum_{j=1}^{m} x^j y^j : \sum_{j=1}^{m} y^j\right] = \left[\lim_{m=\infty}\frac{1}{m}\sum_{j=1}^{m} x^j y^j\right] : \left[\lim_{m=\infty}\frac{1}{m}\sum_{i=1}^{m} y^j\right],$$

d'où l'on déduit, si les collectifs sont indépendants :

$$\lim_{m=\infty}\frac{1}{m}\sum_{j=1}^{m} x^j y^j = \lim_{m=\infty}\sum_{j=1}^{m} x^j \times \lim_{m=\infty}\sum_{j=1}^{m} y^j = p_1 \times p_2$$

et, par suite,

$$\lim_{m=\infty}\frac{1}{m}\sum_{i=1}^{m} x^j(1-y^j) = p_1(1-p_2),$$

$$\lim_{m=\infty}\sum_{j=1}^{m}(1-x^j)y^j = (1-p_1)p_2,$$

....

Passons au cas de plus de deux alternatives : soient les n collectifs (alternatives)

$$(39)\quad \left\{\begin{matrix} x_1 = x_1^1\, x_1^2 \ldots x_1^m \ldots \\ x_2 = x_2^1\, x_2^2 \ldots x_2^m \ldots \\ \ldots\ldots\ldots\ldots\ldots\ldots \\ x_n = x_n^1 x_n^2 \ldots x_n^m \ldots \end{matrix}\right\} \qquad \lim_{m=\infty}\sum_{i=1}^{m} x_i^j = p_i \quad (i = 1, 2, \ldots, n),$$

relatifs à $\mathfrak{S}_1, \mathfrak{S}_2, \ldots, \mathfrak{S}_n$. La suite $(x_1^m, x_2^m, \ldots, x_n^m)$ peut prendre 2^n formes différentes, soient $m_1, m_2, \ldots. m_\alpha$ $(\alpha = 2^n)$, d'où, en désignant par m^i la $i^{\text{ième}}$ colonne de (39), la suite

$$(40)\quad \{m^i\} = m^1, m^2, \ldots, m^i, \ldots \qquad (m^i = m_1, \text{ ou } m_2, \ldots, \text{ ou } m_\alpha).$$

Définition 10. — *Soit $\mathfrak{S}$ un système de sélections relatif à une suite telle que $\{m^i\}$; les collectifs $x_1, x_2, \ldots, x_n$ seront associables relativement à $\mathfrak{S}$ si $\{m^i\}$ est un collectif relativement à $\mathfrak{S}$. Ils seront indépendants si nous avons de plus*

$$(41)\quad \left\{\begin{aligned} &\lim_{m=\infty}\frac{1}{m}\sum_{j=1}^{m} x_1^j x_2^j \ldots x_n^j = p_1 p_2 \ldots p_n, \\ &\lim_{m=\infty}\frac{1}{m}\sum_{j=1}^{m}(1-x_1^j)x_2^j \ldots x_n^j = (1-p_1)p_2 \ldots p_n, \\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots, \end{aligned}\right.$$

l'expression sous le signe Σ *prenant* 2^n *formes différentes, en remplaçant de toutes les manières possibles certains des* x_s^j *par* $1 - x_s^j$ $(s = 1, 2, \ldots, n)$.

Passons maintenant au cas où les collectifs considérés sont en infinité dénombrable; soient les collectifs $x_1, x_2, \ldots, x_n, \ldots$ de fréquences $p_1, p_2, \ldots, p_n, \ldots$.

Définition 11. — *Soient* $\mathfrak{S}_1, \mathfrak{S}_2, \ldots, \mathfrak{S}_n, \ldots$ *une suite de systèmes de sélections,* $\mathfrak{S}_i$ *étant relative à un collectif de la forme* $\{m_i^j\}$ *où* $\{m_i^j\}$ *est une suite de i termes égaux à* 0 *ou à* 1 ; *les collectifs* $x_1, x_2, \ldots, x_n, \ldots$ *seront associables relativement à la suite* $\{\mathfrak{S}_i\}$ *si quel que soit n, les n collectifs* $x_1, \ldots, x_n$ *sont associables relativement à* $\mathfrak{S}_n$. *Ils seront indépendants si quel que soit n, les n collectifs* $x_1, x_2, \ldots, x_n$ *sont indépendants.*

Dans le cas où $p_1 = p_2 = \ldots = p_n = \ldots$ nous allons voir que l'on peut satisfaire à la définition (11) en imposant des conditions plus restrictives, mais plus commodes. Considérons le tableau infini $\|x_i^j\|$ des termes des collectifs considérés, et écrivons ces termes en les rangeant en une suite unique

$$\xi = x_1^1 x_1^2 x_2^1 x_1^3 x_2^2 x_3^1 x_1^4 x_2^3 x_3^2 x_4^1 \ldots$$

D'une manière précise, nous aurons

$$\xi = \xi^1 \xi^2 \ldots \xi^n \ldots, \tag{42}$$

le terme x_i^j occupant dans ξ le rang $\dfrac{(i+j-2)(i+j-1)}{2} + i$. La correspondance entre ξ et le tableau des x_i^j est biunivoque, ainsi que l'on s'en assure facilement. Remarquons que dans ξ, x_i^{j+1} est à droite de x_i^j et que x_{i+1}^j est à droite de x_i^j. Étant donné un collectif x_n, il existe une sélection A qui, appliquée à ξ, donne x_n; de même étant donnée une sélection B appliquée à x_n, il existe une sélection B′ telle que $B'(\xi) = B(x_n)$ quelle que soit la valeur des x_i^j.

Considérons maintenant une sélection du système $\mathfrak{S}_i$, soit S_i, qui choisit certaines colonnes de i termes tenu compte des formes des colonnes précédentes : puisque x_h^{j+1} est à droite des x_h^j, il existe une sélection qui, appliquée à ξ, ne conserve dans ξ que les termes appartenant à des colonnes conservées par S_i, et cela toujours quelle que soit la valeur des termes x_i^j. Soit S_i' cette sélection. Si ξ

est de fréquence p, nous pouvons astreindre la suite $S'_i(\xi)$, à l'aide d'un nombre fini de sélections, à jouir de la propriété suivante, si elle est un collectif relativement à ces sélections : « si l'on partage $S'_i(\xi)$ en tranches de longueur i, toute tranche contenant α fois 0 et β fois 1 intervient avec la fréquence $p^\alpha q^\beta (\alpha + \beta = i ; p + q = 1)$. » Il suffit de considérer la sélection qui choisit les termes de rangs 1, $i+1$, $2i+1$, $3i+1$; ...; si $S'_i(\xi)$ est un collectif relativement à cette sélection, les tranches qui commencent par 1 ont la fréquence p. On considère ensuite la sélection qui choisit les termes de rang 2, $i+2$, $2i+2$, ... précédés d'un 1 : les tranches qui commencent par 11 ont alors la fréquence p^2; puis la sélection qui choisit les termes de rang 2, $i+2$, $2i+2$, ... précédés d'un 0 : les tranches qui commencent par 01 ont alors la fréquence qp, et ainsi de suite.

En remarquant que les tranches de i termes dans $s'_i(\xi)$ correspondent à des colonnes de i termes dans $\| x^j_i \|$, et en observant qu'une sélection appliquée à $S'_i(\xi)$ peut être considérée comme une sélection appliquée à ξ, nous voyons que dans l'hypothèse

$$p_1 = p_2 = \ldots = p_n = \ldots = p,$$

nous pouvons énoncer la proposition :

Quels que soient : les systèmes dénombrables de sélections définissant les collectifs x_n; les systèmes dénombrables $\mathfrak{S}_i$ définissant l'associabilité :

Il existe un système dénombrable $\mathfrak{S}$ tel que si ξ est un collectif $K(\mathfrak{S}, p)$, *en remontant de ξ au tableau $\| x^j_i \|$ les suites formant les lignes de ce tableau soient des collectifs satisfaisant aux conditions imposées, pour chaque collectif séparément, pour l'associabilité, et pour l'indépendance.*

8. **Conséquences de la forme de l'irrégularité imposée par les sélections.** — Les théorèmes de la théorie classique modernisée concernant les probabilités dénombrables de M. E. Borel ne peuvent être tous démontrés dans la théorie des collectifs telle qu'elle est exposée ci-dessus.

Nous allons le montrer sur un exemple. Considérons le théorème suivant (p. 40).

La probabilité que dans une suite infinie d'épreuves indépendantes de même probabilité p, l'inégalité

$$(43) \qquad \left|\frac{\nu}{\mu} - p\right| > \frac{1}{\sqrt{\mu}}$$

ait lieu une infinité de fois est égale à 1 ($\nu =$ nombre de succès parmi les μ premiers épreuves).

Pour démontrer ce théorème dans la théorie des collectifs, il faut former un collectif dont chaque élément soit une suite : nous devons donc considérer une infinité de collectifs indépendants, de même fréquence. Quelles conditions pouvons-nous faire intervenir dans la définition de ces collectifs et de leur indépendance? Seules des conditions de la nature de celles qui figurent dans la définition 11. du moins en l'état actuel de la théorie. Or, de telles conditions peuvent être mises sous une forme *plus restrictive* en astreignant ξ (p. 50) à être un collectif $K(\mathfrak{S}; p)$. Remarquons que $\mathfrak{S}$ contient les sélections qui extraient de ξ les collectifs x_n eux-mêmes. Reportons-nous maintenant au théorème IV'; quel que soit $\mathfrak{S}$, il existera une suite ξ telle que dans toute suite x_n qui s'en déduit en passant de ξ au tableau $\|x^j\|$ (ce qui peut être considéré, nous l'avons vu, comme le résultat d'une sélection) l'inégalité (43) n'ait lieu qu'un nombre fini de fois.

A fortiori, quelles que soient les conditions imposées pour l'indépendance, il existera des collectifs indépendants tels que dans chacun d'eux, l'inégalité (43) n'ait lieu qu'un nombre fini de fois.

Par ailleurs, il existe également des collectifs satisfaisant aux conditions d'indépendance données et tels que pour chacun d'eux l'inégalité (43) ait lieu une infinité de fois. Il suffit pour le voir de remarquer que si l'on détermine les suites $x_1, x_2, \ldots, x_n, \ldots$ en procédant par tirage au sort, il y a une probabilité égale à 1 pour que les relations d'indépendance soient satisfaites, et que pour chacune de ces suites (43) ait lieu une infinité de fois.

Nous pouvons donc conclure :

THÉORÈME V. — *Soient donnés :*

1° *une valeur p quelconque comprise entre* 0 *et* 1;

2° *un système* $\mathcal{R}$ *quelconque (mais dénombrable) de conditions d'associabilité et d'indépendance au sens de la définition* 11;

3° *la propriété* P *qui consiste pour une suite en ce que l'inégalité* (43) *n'a lieu qu'un nombre fini de fois.*

Dans ces conditions, il existe entre autres :

α. une suite de collectifs de fréquence totale p, *indépendants relativement aux conditions d'indépendance* $\mathcal{R}$, *suite de collectifs dans laquelle la fréquence totale de ceux qui présentent la propriété* P *est égale à* 1;

β. une suite de collectifs de fréquence totale p, *indépendants relativement aux conditions d'indépendance* $\mathcal{R}$, *suite de collectifs dans laquelle la fréquence totale de ceux qui présentent la propriété* P *est égale à* 0.

On ne peut donc pas songer à démontrer le théorème classique, envisagé ci-dessus, au moyen de la théorie des collectifs, telle qu'elle a été précisée par M. Wald. Il faut d'ailleurs reconnaître que des théorèmes de cette nature ([1]) n'ont jamais été envisagés dans la théorie classique avant la théorie des probabilités dénombrables de M. E. Borel, ni, après, par les auteurs qui définissent la probabilité par la fréquence totale, et que si, comme tel semble être le but de ces derniers, ils n'ont en vue que les applications, il n'y a aucun inconvénient à ne pas démontrer ces propositions.

([1]) Où intervient le principe d'additivité complète des probabilités.

CHAPITRE III.

LES SUITES INDIFFÉRENTES.

SOMMAIRE. — 1. Suites indifférentes. — 2. Sélections bornées. — 3. Composition des sélections bornées. — 4. Suites de Bernoulli. — 5. Sélections laissant invariante l'indifférence d'une suite.

Les *suites indifférentes* ont été introduites, en vue de fournir une représentation d'une suite prise au hasard, par différents auteurs, en particulier : MM. Popper, sous le nom de Nachwirkungsfreie Folgen; Reichenbach, sous le nom de Normale Folgen; Copeland, sous le nom de Admissible Numbers; enfin, elles constituent un cas particulier des collectifs de de Misès-Wald.

1. Suites indifférentes. — Soit $x = x^1\, x^2 \dots x^n \dots$ une suite formée de 0 et de 1, représentant les résultats d'une suite d'épreuves indépendantes et faites dans les mêmes conditions, M. Popper ([1], p. 106) considère qu'une suite donnée, telle que x, fournit une représentation satisfaisante d'une suite fournie par le hasard si :

a. La fréquence des 1 *dans* x *existe* (p. 3). *Soit* p *cette fréquence.*

b. Quelle que soit la configuration $(a^1\, a^2 \dots a^m)$ (p. 2) *on extrait de* x *une suite partielle en ne conservant que les termes qui sont précédés de la configuration* $(a^1\, a^2 \dots a^m)$, *la fréquence des* 1 *dans cette suite existe et a même valeur que dans* x.

M. Popper appelle de telles suites des *suites indifférentes de tout ordre*. Nous dirons simplement des *suites indifférentes*.

Avec les notations introduites dans le Chapitre I nous obtenons la définition :

DÉFINITION 1. — *Une suite* $x = \{x^i\}$ ($x^i = 0$ ou 1), *sera dite indifférente si* :

a. la fréquence des 1 *existe, soit* $[1]^x = p$;

b. quelle que soit la configuration $(a^1 a^2 \ldots a^m)$, *la relation suivante a lieu*

$$(1) \qquad \lim_{n=\infty} \frac{[a^1 a^2 \ldots a^m 1]_n^x}{[a^1 a^2 \ldots a^m]_n^x} = p.$$

Nous supposerons essentiellement dans ce qui suit que $0 < p < 1$. Nous déduisons de la définition 1 que si X est une suite indifférente où $[1]^x = p$ (nous dirons une *suite indifférente de fréquence* p), la limite $\lim_{n=\infty} [0]_n^x$ existe et a pour valeur $q = 1 - p$. D'une manière générale, supposons démontré que :

$$(2) \qquad \lim_{n=\infty} [a^1 a^2 \ldots a^m]_n^x = p^\nu q^{n-\nu}, \qquad \text{où } \nu = a^1 + a^2 + \ldots + a^m.$$

Nous déduisons de (1) que

$$\lim_{n=\infty} [a^1 a^2 \ldots a^m 1]_n^x = p^{\nu+1} q^{(m+1)-(\nu+1)}$$

et, par suite,

$$\lim_{n=\infty} [a^1 a^2 \ldots a^m 0]_n^x = p^\nu q^{m-\nu} - p^{\nu+1} q^{m-\nu} = p^\nu q^{m+1-\nu},$$

ce qui montre que (2) est vraie pour $m + 1$. Donc (2) a lieu pour tout m. Réciproquement, si (2) a lieu pour tout m, on en déduit immédiatement (1). La définition suivante est donc équivalente à la définition 1 :

DÉFINITION 2. — *Une suite* $x = \{x^i\}$ *est dite indifférente de fréquence p si, quelle que soit la configuration* $(a^1 a^2 \ldots a^m)$, *nous avons*

$$[a^1 a^2 \ldots a^m]^x = p^\nu q^{m-\nu} \qquad (\nu = a^1 + a^2 + \ldots + a^m).$$

Remarquons :

1° que $p^\nu q^{m-\nu}$ est la probabilité de voir apparaître la configuration $(a^1 a^2 \ldots a^m)$ si la probabilité de $\{x^i = 1\}$ est égale à p, et si $x^1, x^2, \ldots, x^m$ sont des variables aléatoires indépendantes;

2° que la définition 1 montre que les suites indifférentes sont un cas particulier des collectifs de de Misès-Wald, puisque ne conserver dans x que les termes précédés d'une certaine configuration, c'est opérer une sélection. Nous allons préciser ce dernier point.

2. Sélections bornées. — Une sélection (p. 14) est définie par une suite $\{f_i\}$ de fonctions, f_i étant à i arguments. Nous dirons qu'une sélection est bornée s'il existe un entier n tel que les fonctions f_i ne sont fonctions que des valeurs de n termes au plus précédant le terme à choisir. D'une manière plus précise :

Définition 3. — *Une sélection, définie par la suite $\{f_i\}$ est dite bornée et à n arguments s'il existe une fonction $\varphi(t_1, t_2, \ldots, t_n)$ à n arguments, ne pouvant prendre que les valeurs* 0 *et* 1, *telles que, pour $i \geqq n$, on ait*

$$f_i(x^1, x^2, \ldots, x^i) \equiv \varphi(x^{i-n+1}, x^{i-n+2}, \ldots, x^i), \tag{3}$$

de sorte que f_i est indépendante de i et de $x^1, x^2, \ldots, x^{i-n}$.

Pour $i < n$, f_i, dans une sélection bornée à n arguments, n'est soumise a aucune condition. Nous supposerons $f_i = 0$ pour $i < n$, ce qui ne supprime qu'un nombre fini de termes dans la suite obtenue en procédant à la sélection. Les sélections bornées à n arguments sont en nombre fini, à savoir le nombre de fonctions φ distinctes, qui est 2^{2^n}. *Le système $\mathfrak{A}$ des sélections bornées est donc dénombrable. Tout collectif* $K(\mathfrak{A}; p)$ est manifestement une *suite indifférente.* Nous allons montrer que *toute suite indifférente de fréquence p* est un *collectif* $K(\mathfrak{A}; p)$, et, qui de plus est, qu'en appliquant à une suite indifférente de fréquence p une sélection quelconque du système $\mathfrak{A}$, on obtient une suite indifférente.

3. Composition des sélections bornées. — Nous avons défini (p. 26), étant données deux sélections A et B, la somme $A + B$, le produit intérieur $A.B$, le complémentaire $1 - A$. Nous allons définir une autre opération : le *produit extérieur* (cette notion a été introduite par M. Wald).

Définition 4. — *Étant données deux sélections* A *et* B, *nous appellerons produit extérieur de* A *par* B *la sélection* C *qui de toute suite* x *extrait la suite* $C(x) = B[A(x)]$ *que nous noterons* $C = B \times A$.

On voit facilement qu'appliquer à x la sélection A, ce qui donne la suite $A(x)$, puis appliquer à $A(x)$ la sélection B, ce qui donne la suite $(B \times A)(x)$, équivaut bien à appliquer une sélection unique à x. Remarquons que le produit extérieur n'est pas commutatif.

Nous allons voir comment, si A et B sont des sélections bornées, on peut exprimer la sélection $A \times B$.

Soit $(a^1 a^2 \dots a^m)$ une configuration quelconque. Désignons par $\{a^1 a^2 \dots a^m\}$ la sélection qui choisit x^n si les termes $x^{n-m} x^{n-m+1} \dots x^{n-1}$ forment la configuration $(a^1 a^2 \dots a^m)$ et dans ce cas seulement. Considérons maintenant la sélection $A \times B$, et soit n un entier quelconque; parmi les 2^n configurations $(a^1 a^2 \dots a^n)$ possibles, rangeons dans une même catégorie celles qui jouissent de la propriété suivante :

Si $x^1 = a^1$, $x^2 = a^2$, $\dots$, $x^n = a^n$, le terme x^{n+1} appartient à la suite $(A \times B)(x)$.

Appelons $(\alpha^1 \alpha^2 \dots \alpha^n)$ les configurations de cette catégorie, et soit $\Sigma\{\alpha^1 \alpha^2 \dots \alpha^n\}$ la somme des sélections associées à ces configurations. Tous les termes de la suite $[\Sigma\{\alpha^1 \alpha^2 \dots \alpha^n\}](x)$ appartiennent à $(A \times B)(x)$, et réciproquement, si x^ν est choisi par $A \times B$, il sera choisi par une des sélections $\{\alpha^1 \alpha^2 \dots \alpha^{\nu-1}\}$. Nous pouvons donc exprimer $A \times B$ par la double somme infinie

$$A \times B = \sum_n \sum \{\alpha^1 \alpha^2 \dots \alpha^n\}. \tag{4}$$

Les sélections qui figurent dans le second membre de (4) peuvent ne pas être toutes distinctes. Si en effet nous trouvons dans ce second membre deux termes de la forme

$$\{\alpha^1 \alpha^2 \dots \alpha^n \alpha^{n+1} \dots \alpha^{n+m}\},$$
$$\{\alpha^{n+1} \dots \alpha^{n+m}\},$$

tout terme de x choisi par la première sélection le sera par la seconde; nous pouvons donc supprimer la première sélection.

Opérons dans le second membre de (4) cette suppression de termes superflus, d'une manière méthodique : soit s la longueur minimum des termes figurant dans (4); nous supprimerons tous les termes de longueur $s+1$ qui peuvent être obtenus en prolongeant à gauche un terme de longueur s figurant dans (4). Nous supprimons ensuite tout terme de longueur $s+2$ qui peut être obtenu en prolongeant à gauche un terme de longueur s ou $s+1$ figurant déjà dans la suite, et ainsi de suite. Cette réduction étant faite, nous obtenons pour $A \times B$ une expression

$$A \times B = \sum_{n=s} \{\beta^1 \beta^2 \ldots \beta^n\}. \tag{5}$$

Par exemple, le développement de $\{1\} \times \{0\}$ sera

$$\{1\} \times \{0\} = \{010\} + \{0110\} + \{01110\} + \ldots.$$

Les sélections $\{\beta^1 \beta^2 \ldots \beta^n\}$ sont disjointes, en ce sens que les suites $\{\beta^1 \beta^2 \ldots \beta^n\}(x)$ sont disjointes. Considérons les sélections A et B elles-mêmes, qui sont bornées par hypothèse; supposons-les à h et k arguments respectivement. La sélection $1 - A$ n'est pas exactement une sélection bornée, puisqu'elle choisit les h premiers termes de toute suite, et que nous avons supposé qu'une sélection bornée à n arguments ne pouvait choisir des termes qu'à partir du $n+1^e$. Nous appellerons $\overline{A}$ la sélection qui ne choisit aucun terme de rang $\leq h$, et qui, au delà, coïncide avec $1 - A$. De même nous définirons $\overline{B}$. $\overline{A}$ et $\overline{B}$ étant bornées, les sélections $A \times \overline{B}$, $\overline{A} \times B$, $\overline{A} \times \overline{B}$ sont susceptibles de représentations telles que (5) : nous les noterons

$$\left\{ \begin{array}{ll} A \times B = \sum \{u_i\}, & A \times \overline{B} = \sum \{v_i\}, \\ \overline{A} \times B = \sum \{w_i\}, & \overline{A} \times \overline{B} = \sum \{t_i\}, \end{array} \right. \tag{6}$$

$\{u_i\}$, $\{v_i\}$, $\{w_i\}$, $\{t_i\}$ représentant des sélections de la forme $\{a^1 a^2 \ldots a^m\}$.

Considérons la sélection

$$\Gamma = A \times B + A \times \overline{B} + \overline{A} \times B + \overline{A} \times \overline{B}.$$

La suite $\Gamma(x)$ ne contient pas tous les termes de x mais les contient sûrement tous à partir du rang $k+2h+1$. En effet, parmi les $k+2h+1$ premiers termes, il y en a $2h+1$ qui appartiennent à la suite $(B+\overline{B})(x)$; donc l'une des suites $B(x)$ ou $\overline{B}(x)$ en contient au moins $h+1$; soit $B'(x)$ cette suite, l'une des suites $(A \times B')(x)$ ou $(\overline{A} \times B')(x)$ contiendra le terme de rang $h+1$ de $B'(x)$, donc un terme de rang $\leqq k+2h+1$ de x.

Considérons maintenant le système $\mathfrak{A}$ de toutes les sélections bornées, et soit x une suite indifférente. Étant donnée une configuration $(\rho^1\rho^2\ldots\rho^s)$, elle figure dans x avec une fréquence égale au nombre

$$(7) \qquad [\rho^1\rho^2\ldots\rho^s] = p^\nu q^{s-\nu} \qquad \left(\nu = \sum_1^s \rho^i\right).$$

Revenons à la formule (6), où A et B sont deux éléments de $\mathfrak{A}$. Étant donnée une sélection $\{u_i\}$, où $u_i = (\sigma^1\sigma^2\ldots\sigma^s)$, nous lui associons le nombre

$$(8) \qquad [u_i] = [\sigma^1\sigma^2\ldots\sigma^s].$$

Nous allons montrer d'abord que :

$$(9) \qquad \sum_i [u_i] + \sum_i [v_i] + \sum_i [w_i] + \sum_i [t_i] = 1.$$

Soit n un entier $> k+2h+1$; considérons les sélections

$$(10) \qquad \begin{cases} (A \times B)' = \sum_i{}' \{u_i\}, & (A \times \overline{B})' = \sum_i{}' \{v_i\}, \\ (\overline{A} \times B)' = \sum_i{}' \{w_i\}, & (\overline{A} \times \overline{B})' = \sum_i{}' \{t_i\}, \end{cases}$$

les sommations étant étendues aux configurations de longueur $< n$. La sélection

$$\Gamma' = (A \times B)' + (\overline{A} \times B)' + (A \times \overline{B})' + (\overline{A} \times \overline{B})'$$

choisit dans x tous les termes depuis le rang $k+2h+1$ jusqu'au rang n, peut-être quelques termes de rang $\leqq k+2h$ et enfin des termes de rang $> n$. n restant fixe, soit N un entier $> n$; considérons les N premiers termes de x. Si nous désignons par r le quotient entier de N par n, on voit, en découpant x en tranches de longueur n, que la sélection T′ choisit, parmi les N premiers termes de x, des termes en nombre égal au moins à $r(n-k-2h)$. Soit f_N la fréquence, parmi les N premiers termes de x, de ceux choisis par T′, nous avons

$$f_N \geqq \frac{r}{N}(n-k-2h) > \frac{N-n}{N}\,\frac{n-k-2h}{n}. \tag{11}$$

Mais nous pouvons exprimer directement f_N à l'aide des fréquences des configurations (u_i), (v_i), ... En effet, nous avons

$$f_N = \left[\sum{}' [u_i]^x_{N-1} + \sum{}' [v_i]^x_{N-1} + \sum{}' [w_i]^x_{N-1} + \sum{}' [t_i]^x_{N-1}\right] \frac{N-1}{N}. \tag{12}$$

Faisons tendre N vers l'infini; il vient, en tenant compte de (11),

$$\sum{}' [u_i] + \sum{}' [v_i] + \sum{}' [w_i] + \sum{}' [t_i] \geqq \frac{n-k-2h}{n}. \tag{13}$$

Par ailleurs la somme (finie) qui figure dans (13) ne peut jamais être > 1, puisque les sélections $A \times B$, $A \times \bar{B}$, $\bar{A} \times B$, $\bar{A} \times \bar{B}$ ne choisissent jamais en même temps un même terme. Donc nous avons bien, en faisant tendre n vers l'infini,

$$\sum [u_i] + \sum [v_i] + \sum [w_i] + \sum [t_i] = 1. \qquad \text{C. Q. F. D.} \tag{14}$$

T choisit tous les termes de x à partir du rang $k + 2h + 1$; donc

$$\sum [u_i]^x_n + \sum [v_i]^x_n + \sum [w_i]^x_n + \sum [t_i]^x_n \geqq \frac{n-k-2h}{n}. \tag{15}$$

Faisons tendre n vers l'infini; nous déduisons de (14) et (15) que

$$\lim_{n=\infty} \sum [u_i]^x_n = \sum \lim_{n=\infty} [u_i]^x_n = \sum [u_i].$$

La conclusion des considérations précédentes est la suivante : *A et B étant deux sélections bornées, il existe une suite* (u_1), (u_2), ..., (u_n), ... *de configurations telles que si l'on désigne par* $\{u_i\}$ *la sélection qui choisit tout terme précédé de la configuration* (u_i), *on ait*

$$A \times B = \{u_1\} + \{u_2\} + \ldots + \{u_n\} + \ldots. \tag{17}$$

La fréquence des termes choisis par $A \times B$ *dans une suite* *indifférente x existe et est égale à la somme des fréquences des termes choisis par $\{u_1\}, \{u_2\}, \ldots, \{u_n\}, \ldots$, respectivement.*

Soit maintenant la suite $(A \times B)(x)$ extraite d'une suite indifférente x; cherchons si le signe 1 a dans $(A \times B)(x)$ une limite bien définie. La suite partielle extraite de la suite des n premiers termes de x contient $(n-1)\sum_i [u_i]_{n-1}^x$ termes, et le nombre 1 y figure $n\sum_i [u_i 1]_n^x$ fois, en désignant par $(u_i 1)$ la configuration obtenue en ajoutant un 1 à droite de la configuration (u_i). Or, nous avons

$$(n-1)\sum_i [u_i]_{n-1}^x = n\sum_i [u_i 0]_n^x + n\sum_i [u_i 1]_n^x.$$

Divisons par n et faisons tendre n vers l'infini; de (16), du fait que x est indifférente, et que

$$\sum_i [u_i] = \sum_i [u_i 0] + \sum_i [u_i 1],$$

nous déduirons que

$$\lim_{n=\infty} \sum_i [u_i 1]_n^x = \sum_i [u_i 1] = p\sum_i [u_i].$$

Donc la fréquence des 1 dans $(A \times B)(x)$ existe et est égale à p.

Nous voyons ainsi que *toute suite indifférente x est un collectif relativement au système dénombrable des sélections de la forme* $A \times B$, où A et B sont deux sélections bornées. Nous en déduirons que :

Toute suite indifférente x de fréquence p est un collectif $K(\mathfrak{A}; p)$.

Étant donnée une sélection bornée A *et une suite indifférente x, la suite* $A(x)$ *est encore une suite indifférente.*

La propriété d'être indifférente reste donc invariante par application d'une sélection bornée.

4. Suites de Bernoulli. — Soit x une suite de 0 et de 1. Partageons-la en tranches de même longueur, soit en tranches de longueur m. Étant donnée une configuration $(a^1 a^2 \ldots a^m)$, nous allons étudier sa fréquence dans la suite des tranches. Pour cela, de la suite

$$x = x^1 x^2 \ldots x^n,$$

nous déduisons une suite

$$y = y^1 y^2 \ldots y^k \ldots \qquad (y^k = 0 \text{ ou } 1)$$

de la manière suivante

$$\begin{array}{ll} y^k = 1 & \text{si } x^{m(k-1)+1} = a^1,\ x^{m(k-1)+2} = a^2,\ \ldots,\ x^{mk} = a^m; \\ y^k = 0 & \text{s'il existe un } j\ (0 < j \leqq m) \text{ avec } x^{(k-1)+j} \neq a^j. \end{array}$$

De sorte que la fréquence, dans la suite des tranches de longueur m, de la configuration $(a^1 a^2 \ldots a^m)$ sera égale (si elle existe) à la fréquence de 1 dans la suite y. Désignons par $| a^1 a^2 \ldots a^m |^x$ cette fréquence.

Si x était obtenue par un tirage au sort, avec Pr. $\{ x^i = 1 \} = p$, la fréquence de $(a^1 a^2 \ldots a^m)$ dans la suite des tranches serait presque certainement égale à $p^\nu q^{m-\nu}$ $(\nu = \Sigma a_i)$. Nous sommes donc amené à étudier les suites telles que :

$$(18) \qquad | a^1 a^2 \ldots a^m |^x = p^\nu q^{m-\nu} \qquad \left(\nu = \sum a^i,\ q = 1 - p\right).$$

De telles suites, où (18) a lieu pour toutes les configurations et toutes les valeurs de m, seront dites *suites de Bernoulli de fréquence* p (Reichenbach, p. 302). Nous allons montrer que *les notions de suite indifférente et de suite de Bernoulli sont équivalentes*. Nous allons établir pour cela un lemme :

Soient donnés une configuration $(a^1 a^2 \ldots a^m)$, un nombre $\varepsilon > 0$, un entier $N > m$, un nombre $p (0 < p < 1)$. Considérons la suite de variables aléatoires indépendantes

$$X_1,\ X_2,\ \ldots,\ X_N; \qquad X_i = 0 \text{ ou } 1; \qquad \text{Pr.} \{ X_i = 1 \} = p.$$

Soit f_N la fréquence de la configuration $(a^1 a^2 \ldots a^m)$ dans la suite des tranches

$$X_1,\ X_2,\ \ldots,\ X_m;\quad X_{m+1},\ \ldots,\ X_{2m};\quad \ldots;\quad X_{(k-1)m+1},\ \ldots,\ X_{km}$$
$$[km \leqq N < (k+1)m]$$

et P_N la probabilité que

$$|f_N - p^\nu q^{m-\nu}| < \varepsilon.$$

Nous pouvons considérer chaque tranche de m épreuves comme une épreuve, la probabilité de voir apparaître $(a^1 a^2 \ldots a^m)$ étant $p^\nu q^{m-\nu}$. Donc nous avons

$$\lim_{N=\infty} P_N = 1.$$

Soit maintenant $(b^1 b^2 \ldots b^N)$ une configuration de longueur N. Partageons-la en tranches de longueur m, et soit $A_N^{(0)}$ le nombre de fois que $(a^1 a^2 \ldots a^m)$ figure dans cette suite finie de tranches

$$b^1 b^2 \ldots b^m;\quad b^{m+1} \ldots b^{2m};\quad \ldots;\quad b^{(k-1)m+1} \ldots b^{km}$$
$$[km \leqq N < (k+1)m].$$

Supprimons les $m-1$ premiers termes de $b^1 b^2 \ldots b^N$ et partageons la suite ainsi obtenue en tranches; $(a^1 a^2 \ldots a^m)$ figurera $A_N^{(1)}$ fois dans

$$b^m b^{m+1} \ldots b^{2m-1};\quad b^{2m} \ldots b^{2m-1};\quad \ldots;\quad b^{k_1 m} \ldots b^{k_1 m+m-1}$$
$$[k_1 m \leqq N - m + 1 < (k_1+1)m].$$

D'une manière générale, supprimons les $m-i$ premiers termes $(i < m)$; $(a^1 a^2 \ldots a^m)$ figurera $A_N^{(i)}$ fois dans la suite

$$b^{m-i+1} \ldots b^{2m-i};\quad \ldots;\quad b^{k_i m-i+1} \ldots b^{(k_i+1)m-i}$$
$$[k_i m \leqq N - m + i < (k_i+1)m].$$

Revenons à la suite $X_1, X_2, \ldots, X_N$ de variables aléatoires. Si nous comptons dans $(X_1 X_2 \ldots X_N)$ la configuration $(a^1 a^2 \ldots a^m)$ comme nous venons de le faire dans $(b^1 b^2 \ldots b^N)$, $A_N^{(0)}$, $A_N^{(1)}$, $\ldots$, $A_N^{(m-1)}$ sont des variables aléatoires. Soit $P_N^{(i)}$ la probabilité suivante

$$P_N^{(i)} = \mathrm{Pr.}\left\{ \left| \frac{m}{N} A_N^{(i)} - p^\nu q^{m-\nu} \right| < \varepsilon \right\} \qquad (0 \leqq i < m).$$

Nous aurons encore

$$\lim_{N=\infty} P_N^{(i)} = 1.$$

Soit maintenant $\varphi(b^1, b^2, \ldots, b^N)$ une fonction à N arguments, égale à 1 pour toute configuration telle que, pour tout i

$$(19) \qquad \left| \frac{m}{N} A_N^{(i)} - p^\nu q^{m-\nu} \right| < \varepsilon \qquad (0 \leqq i < m)$$

et égale à 0 dans le cas contraire. Si nous considérons les variables aléatoires $X_j (j = 1, 2, \ldots, N)$, la probabilité que les m inégalités (19) aient lieu est égale à la valeur moyenne

$$\mathfrak{M}\{\varphi(X_1, X_2, \ldots, X_N)\},$$

d'où

$$(20) \qquad \mathfrak{M}\{\varphi(X_1, X_2, \ldots, X_N)\} \geqq 1 - \sum_{i=0}^{m-1} (1 - P_N^{(i)}).$$

Posons

$$[b^1 b^2 \ldots b^N] = p^\rho q^{N-\rho} \qquad \left(\rho = \sum b^j\right).$$

La relation (20) peut s'écrire

$$(20') \qquad \sum [b^1 b^2 \ldots b^N] \varphi(b^1, b^2, \ldots, b^N) \geqq 1 - \sum_{i=0}^{m-1} (1 - P_N^{(i)}),$$

d'où le lemme

Lemme 1. — *Soient : une configuration $(a^1 a^2 \ldots a^m)$, un nombre $\varepsilon > 0$, un entier $N > m$, une suite indifférente x de fréquence p. Étant donnée une configuration $(b^1 b^2 \ldots b^N)$ de longueur* N, *appelons :*

$[b^1 b^2 \ldots b^N]$ *la fréquence de $(b^1 b^2 \ldots b^N)$ dans x;*

$A_N^{(i)}$ *le nombre de fois que $(a^1 a^2 \ldots a^m)$ figure dans la suite $b^{m-i+1}, \ldots, b^N$ partagée en tranches de longueur m autant de fois qu'il est possible;*

$A_N^{(0)}$ *le nombre de fois que $(a^1 a^2 \ldots a^m)$ figure dans la suite $b^1 b^2 \ldots b^N$ partagée en tranches de longueur m autant de fois qu'il est possible : $\varphi(b^1 b^2, \ldots, b^N)$ une fonction égale à 0 ou 1, prenant la valeur 1 si les m inégalités $\left|\frac{m}{N} \cdot A_N^{(i)} - p^\nu q^{m-\nu}\right| < \varepsilon$ ont lieu $(\nu = \Sigma a^j)$.*

Dans ces conditions nous avons

$$\lim_{N=\infty} \sum [b^1 b^2 \ldots b^N] \varphi(b^1, b^2, \ldots, b^N) = 1,$$

la sommation étant étendue à toutes les configurations de longueur N.

Considérons maintenant les n premiers termes de la suite indifférente x, et partageons-la en tranches de m termes. Soit N un nombre entier, nous supposerons $m < N < n$. ε étant donné, ainsi qu'une configuration $(a^1 a^2 \ldots a^m)$, considérons la suite de configurations

$$(x^1 x^2 \ldots x^N); \quad (x^2 x^3 \ldots x^{N+1}); \quad \ldots; \quad (x^{n-N+1} \ldots x^n)$$

et la suite de fonctions

$$\varphi(x^1, x^2, \ldots, x^N); \quad \varphi(x^2, x^3, \ldots, x^{N+1}); \quad \ldots; \quad \varphi(x^{n-N+1}, \ldots, x^n).$$

Considérons l'une d'elles, et supposons qu'elle soit égale à 1

$$\varphi(x^{k+1}, x^{k+2}, \ldots, x^{k+N}) = 1 \qquad (0 \leqq k \leqq n - N).$$

D'après la définition de φ, nous en déduisons que la suite $x^{k+1} \ldots x^{k+N}$ empiète sur au moins $\frac{N}{m}(p^\nu q^{m-\nu} - \varepsilon)$ tranches, parmi celles en lesquelles on a découpé x, qui sont identiques à $(a^1 a^2 \ldots a^m)$. Réciproquement, considérons une tranche de longueur m quelconque; parmi les suites $x^{k+1}, \ldots, x^{k+N}$ il en existe au plus $N - m + 1$ qui empiètent sur elle. D'où l'évaluation suivante du nombre K_n de tranches de longueur m identiques à $(a^1 a^2 \ldots a^m)$ dans le partage de la suite des n premiers termes de x

$$K_n \geqq \frac{N}{m(N - m + 1)}(p^\nu q^{m-\nu} - \varepsilon) \sum_{k=0}^{n-N} \varphi(x^{k+1}, \ldots, x^{k+N}). \tag{21}$$

Parmi les n premiers termes de x, la configuration $(b^1 b^2 \ldots b^N)$ figure $n[b^1 b^2 \ldots b]_n^x$ fois; par ailleurs le nombre de tranches de longueur m est au plus égal à $\frac{n}{m}$; d'où pour la fréquence F_n de la tranche $a^1 a^2 \ldots a^m$ parmi celles obtenues en découpant $x^1 x^2 \ldots x^n$ en tranches de longueur m, l'inégalité

$$F_n \geqq \frac{N}{N - m + 1}(p^\nu q^{m-\nu} - \varepsilon) \sum [b^1 b^2 \ldots b^N]_n^x \varphi(b^1, b^2, \ldots, b^N).$$

Faisons tendre n, puis N, vers l'infini, nous obtenons, d'après le lemme 1

$$\varliminf_{n=\infty} F_n \geqq p^\nu q^{m-\nu} - \varepsilon$$

et, ε étant arbitraire,

$$(22) \qquad \varliminf_{n=\infty} F_n \geqq p^\nu p^{m-\nu},$$

(22) étant vraie quelle que soit la configuration $(a^1 a^2 \ldots a^m)$ considérée, nous en déduisons que $|a^1 a^2 \ldots a^m|$ existe et est égale à $p^\nu q^{m-\nu}$. Donc :

Lemme 2. — *Toute suite indifférente est une suite de Bernoulli.*

La réciproque de ce lemme est aisée à démontrer. Supposons en effet que x soit une suite de Bernoulli. Soit une configuration $(a^1 a^2 \ldots a^m)$ et une configuration $(b^1 b^2 \ldots b^N)$; introduisons la fonction φ définie dans le lemme 1. Si $\varphi(b^1, b^2, \ldots, b^N) = 1$, la configuration $(a^1 a^2 \ldots a^m)$ figurera dans $(b^1 b^2 \ldots b^N)$ [au sens primitif introduit dans le premier chapitre, c'est-à-dire sans partager $(b^1 b^2 \ldots b^N)$ en tranches] un nombre de fois égal à

$$A_N^{(0)} + A_N^{(1)} + \ldots + A_N^{(m-1)},$$

c'est-à-dire au moins égal à $N(p^\nu q^{m-\nu} - \varepsilon)$. Ceci étant posé, partageons la suite x en tranches *de longueur* N, et considérons la suite des fonctions

$$\varphi(x^1, x^2, \ldots, x^N), \quad \varphi(x^{N+1}, \ldots, x^{2N}), \quad \ldots.$$

Si nous considérons les n premiers termes de x; ils contiendront k tranches de longueur N $[kN \leqq n < (k+1)N]$, et la configuration $(a^1 a^2 \ldots a^m)$ y figurera au moins

$$N(p^\nu q^{m-\nu} - \varepsilon) \sum_{i=0}^{k-1} \varphi(x^{Ni+1}, \ldots, x^{N(i+1)}) \text{ fois};$$

d'où la relation

$$(23) \qquad [a^1 a^2 \ldots a^m]_n^x \geqq \frac{N}{n}(p^\nu q^{m-\nu} - \varepsilon) \sum_{i=0}^{k-1} \varphi(x^{Ni+1}, \ldots, x^{N(i+1)}).$$

Soit $|b^1 b^2 \ldots b^N|_k^x$ la fréquence de $(b^1 b^2 \ldots b^N)$ dans les k premières tranches de longueur N. (23) peut alors s'écrire

$$[a^1 a^2 \ldots a^m]_n^x \geqq \frac{kN}{n}(p^\nu q^{m-\nu} - \varepsilon) \sum |b^1 b^2 \ldots b^N|_k^x \varphi(b^1, b^2, \ldots, b^N).$$

Faisons tendre n vers l'infini; $|b^1 b^2 \ldots b^N|_k^x$ a alors pour limite $p^\rho q^{N-\rho}$ $(\rho = \Sigma b^i)$, ce que nous pouvons désigner par $[b^1 b^2 \ldots b^N]$, et nous avons alors

$$\varliminf_{n=\infty} [a^1 a^2 \ldots a^m]_n^x \geqq (p^\nu q^{m-\nu} - \varepsilon) \sum [b^1 b^2 \ldots b^N] \varphi(b^1, b^2, \ldots, b^N)$$

et, en faisant tendre N vers l'infini, et tenant compte de (20′) et de $\lim\limits_{n=\infty} P_N^{(i)} = 1$, puis du fait que ε a été arbitrairement choisi

$$\varliminf_{n=\infty} [a^1 a^2 \ldots a^m]_n^x \geqq p^\nu q^{m-\nu}, \tag{24}$$

(24) ayant lieu quelle que soit $(a^1 a^2 \ldots a^m)$, nous en déduisons que $[a^1 a^2 \ldots a^m]_n^x$ existe et a pour valeur $p^\nu q^{m-\nu}$. Donc x est une suite indifférente.

La conclusion de ce paragraphe est donc la suivante :

Les notions de suite indifférente et de suite de Bernoulli sont équivalentes.

M. Copeland a montré (Copeland, I; p. 550) que les notions de suite de Bernoulli et de nombres admissibles étaient équivalentes; il y a donc identité entre les nombres admissibles et les suites indifférentes.

5. **Sélections laissant invariante l'indifférence d'une suite.** — Soit une sélection S; nous disons qu'elle *laisse invariante l'indifférence* d'une suite, si. *quelle que soit la suite indifférente x, la suite $S(x)$ est encore une suite indifférente.* Remarquons que si S_1 et S_2 sont deux telles sélections, il en est de même pour leur produit extérieur $S_1 \times S_2$ (p. 47). Nous avons vu (p. 53) que les sélections bornées laissent invariante l'indifférence. Nous allons signaler une autre classe de sélections jouissant de la même propriété : ce sont les sélections arithmétiquement périodiques, dont voici la définition.

Définition 5. — *Une sélection* S, *définie par une suite de fonctions* $\{f_i\}$ *est arithmétiquement périodique si, d'une manière générale,* f_i *n'est fonction que de l'entier* i, *et s'il existe un entier* k (*la période*) *tel que* $f_{i+k} = f_i$ *quel que soit* i.

Soit S une telle sélection, de période k. Elle est également de période hk, où h est un entier quelconque. Soit x une suite indifférente de fréquence p; partageons-la en tranches de longueur k. Le nombre de termes de $S(x)$ provenant d'une tranche quelconque de x est un nombre fixe, indépendant de la tranche choisie, soit u cet entier. Si nous partageons $S(x)$ en tranches de longueur u, puisque x est une suite de Bernoulli, nous avons, dans $S(x)$

$$|a^1 a^2 \dots a^u| = p^\nu q^{u-\nu} \qquad \left(\nu = \sum_1^u a^i\right)$$

pour toute configuration de longueur u.

De même, pour tout multiple de u, nous aurons, dans $S(x)$,

$$(25) \qquad |a^1 a^2 \dots a^{hu}| = p^{\nu'} q^{hu-\nu'} \qquad \left(\nu' = \sum_1^{hu} a^i\right).$$

Nous ne sommes pas encore sûr que $S(x)$ soit une suite de Bernoulli; mais si nous examinons la démonstration de la réciproque du lemme 2 (p. 57), nous voyons qu'il n'est pas besoin de considérer toutes les valeurs de N; il suffit de supposer que la relation

$$\lim_{k=\infty} |b^1 b^2 \dots b^N|_k = p^\rho q^{N-\rho} \qquad \left(\rho = \sum b^j\right)$$

a lieu pour une infinité de valeurs de N. Dans le cas qui nous occupe, d'après (25), nous pouvons reproduire la démonstration de la réciproque du lemme 2 en prenant pour N les valeurs de la forme hu, où h est un entier quelconque. Donc $S(x)$ est une suite indifférente. Par conséquent :

La propriété, pour une suite, d'être indifférente, reste invariante quand on procède soit à une sélection bornée, soit à une sélection arithmétiquement périodique, soit à une série finie de sélections successives des types précédents.

Cette invariance n'a été signalée, à notre connaissance, par aucun des auteurs ayant traité de la notion de suite indifférente à titre principal ou accessoire.

La notion de suite indifférente fournit donc une représentation assez parfaite de l'irrégularité exigée par M. de Misès. Mais, naturellement, les remarques faites dans le chapitre précédent (p. 53) sur les collectifs de de Misès-Wald s'appliquent intégralement aux suites indifférentes.

CHAPITRE IV.

CRITÈRES D'IRRÉGULARITÉ FONDÉS SUR LA NOTION DE MARTINGALE.

Nous allons dans ce chapitre proposer une nouvelle manière d'exprimer l'axiome d'irrégularité de M. de Misès, qui reste conforme à son idée générale citée page 22, mais non équivalente à la traduction mathématique qu'il en a donnée.

1. Rappel des propriétés des conditions d'irrégularité proposées par M. de Misès. — Considérons uniquement le cas de l'alternative (p. 23). Si nous considérons une suite x où la fréquence des 1 est égale à $p (0 < p < 1)$, une condition d'irrégularité, selon M. de Misès sera, avec la notation de M. de Wald, de la forme (p. 32)

(A). « x est un collectif $K(\mathfrak{S}; p)$ ».

Si $\mathfrak{S}$ n'est pas dénombrable, on n'est pas sûr qu'une telle condition n'est pas impossible à réaliser. Si l'on prend pour $\mathfrak{S}$ l'ensemble total de tous les S possibles, comme l'autorisaient les premières définitions de M. de Misès, la condition est impossible. Mais M. de Misès a accepté la modification de M. Wald, qui prend pour $\mathfrak{S}$ un système dénombrable. Si $\mathfrak{S}$ est dénombrable, la portée d'une telle condition est définie par les théorèmes 3 et 4 (p. 39) du Chapitre II; on peut la résumer mathématiquement en ces mots :

Tout ensemble $K(\mathfrak{S}; p)$ *est de p-mesure égale à* 1; *mais quel que soit p, il existe au moins un ensemble de p-mesure* *nulle tel que, quel que soit* $\mathfrak{S}$, *il ait un point commun avec* K($\mathfrak{S}$; p).

En langage de la théorie classique modernisée, la remarque précédente signifie : quels que soient $\mathfrak{S}$ et p, si l'on procède à une suite d'épreuves indépendantes répétées où la probabilité d'un événement est p, on obtient presque certainement un collectif K($\mathfrak{S}$; p). Mais, il existe, quel que soit p, au moins une propriété de la suite des résultats des épreuves considérées, propriété qui presque certainement n'a pas lieu, qui est possédée quel que soit $\mathfrak{S}$ par un au moins des collectifs K($\mathfrak{S}$, p).

Or la seconde partie de cette remarque exprime une propriété que beaucoup considéreront comme un défaut de la notion de collectif. Par exemple, il en résulte que dans un collectif la fréquence n'est pas astreinte à tendre vers sa limite en oscillant, bien que dans la théorie classique la probabilité qu'une telle oscillation ait lieu soit égale à 1. Nous allons rendre plus stricte la définition de l'irrégularité, de manière à éviter cette objection. En termes précis :

Nous allons voir que l'on peut donner des conditions d'irrégularité (plus strictes que celle de M. Wald), formant un système $\mathcal{C}$ de conditions C telles que si l'on désigne également par C l'ensemble des points x tels que la suite x satisfasse à la condition C (p. 33), on ait la propriété :

(Γ). *Tout ensemble complémentaire d'un ensemble* C *peut être enfermé dans un ensemble de p-mesure nulle, et tout ensemble de p-mesure nulle peut être enfermé dans un complémentaire d'ensemble* C.

Dans le langage de la théorie classique modernisée, la propriété Γ, relative au système $\mathcal{C}$ de conditions C, s'interprète : Quel que soit C appartenant à $\mathcal{C}$, il y a une probabilité égale à 1 pour que C soit satisfaite par une suite d'épreuves indépendantes; quelle que soit la propriété de suite P que l'on considère, si cette propriété P a une probabilité nulle d'être possédée par une suite d'épreuves indépendantes, il existe dans $\mathcal{C}$ une condition C telle que toute suite qui possède la propriété C ne possède pas la propriété P.

Il y a une manière immédiate de résoudre la question, c'est de prendre pour les C tous les ensembles de p-mesure égale à 1, ou même seulement les F_σ ([1]) de p-mesure égale à 1. Mais ces conditions d'irrégularité ne présentent plus aucun caractère intuitif. Nous allons voir que l'on peut trouver des conditions équivalentes, plus naturelles, en poussant jusqu'au bout l'idée générale du système de jeu posé par M. de Misès, mais qu'il a rétrécie en la traduisant sous la forme (A) de la page 69, dont nous venons de signaler un inconvénient.

2. Système de jeu ou martingale. — Pour introduire son axiome d'irrégularité, M. de Misès donne l'explication suivante (de Misès, [3], p. 4) :

« *Ce deuxième axiome, qui doit exprimer le « caractère aléatoire » des phénomènes considérés, est l'axiome d'irrégularité ou « principe de l'impossibilité de trouver un système de jeu ». Sa position en face du monde réel est la même que celle du principe de l'impossibilité du mouvement perpétuel.* »

La notion de « système de jeu » est une notion qui a besoin d'être précisée au même titre que celle de sélection; nous procéderons dans le même esprit que M. Wald (p. 26).

Toujours dans le cas de l'alternative, soit A un joueur qui joue contre un banquier, en étant libre de varier ses mises, avec la convention suivante : si A mise la somme λ sur l'événement $\{x^n = 1\}$, et que cet événement se produise, il reçoit la somme $\frac{\lambda}{p}$; s'il joue la somme λ sur l'événement $\{x^n = 0\}$ et que cet événement se produise, il touche la somme $\frac{\lambda}{q}$ $(q = 1 - p)$. Supposons que A possède avant le début du jeu, une somme égale à 1. Si après le $(n-1)^e$ coup il possède une somme s_{n-1}, il ne peut que jouer une partie ou la totalité, de cette somme, sur $x^n = 0$ ou $x^n = 1$, ou ne pas miser du tout. Sa conduite peut donc être définie par deux fonctions $\lambda_n(x^1, x^2, \ldots, x^n)$, $\mu_n(x^1, x^2, \ldots, x^n)$, non négatives, de somme inférieure ou égale à 1, ayant la signification suivante :

([1]) Rappelons qu'un F_σ est un ensemble défini comme somme d'une infinité dénombrable d'ensembles fermés.

Quel que soit n, A joue les sommes

$$\begin{matrix} \lambda_{n-1} s_{n-1} & \text{sur } x^n = 1 & (\lambda_{n-1} \geqq 0) \\ \mu_{n-1} s_{n-1} & \text{sur } x^n = 0 & (\mu_{n-1} \geqq 0) \end{matrix} \quad (\lambda_{n-1} + \mu_{n-1} \leqq 1).$$

s_{n-1} étant la somme dont il dispose après le n^e coup. s_n, λ_n, μ_n sont des fonctions de $x^1, x^2, \ldots x^n$; pour $n = 0$, s_0, λ_0, μ_0 sont des constantes, avec $s_0 = 1$.

Remarquons que si les fonctions λ_n et μ_n sont données, on peut en déduire les fonctions s_n. En effet, d'après la signification de λ_n, μ_n, et la règle du jeu, nous avons :

$$(1) \qquad \left\{ \begin{aligned} s_n &= \frac{\lambda_{n-1}}{p} s_{n-1} + s_{n-1}(1 - \lambda_{n-1} - \mu_{n-1}), & \text{si } x^n = 1, \\ s_n &= \frac{\mu_{n-1}}{q} s_{n-1} + s_{n-1}(1 - \lambda_{n-1} - \mu_{n-1}), & \text{si } x^n = 0, \end{aligned} \right.$$

ce qui, joint à $s_0 = 1$, détermine la fonction s_n quel que soit n. Explicitons les arguments des fonctions dans (1) : nous obtenons

$$(1\ bis) \qquad \left\{ \begin{aligned} & s_n(x^1, x^2, \ldots, x^{n-1}, 1) \\ & \quad = \frac{\lambda_{n-1}}{p} s_{n-1}(x^1, x^2, \ldots, x^{n-1}) \\ & \qquad + s_{n-1}(x^1, x^2, \ldots, x^{n-1})[1 - \lambda_{n-1} - \mu_{n-1}], \\ & s_n(x^1, x^2, \ldots, x^{n-1}, 0) \\ & \quad = \frac{\mu_{n-1}}{q} s_{n-1}(x^1, x^2, \ldots, x^{n-1}) \\ & \qquad + s_{n-1}(x^1, x^2, \ldots, x^{n-1})[1 - \lambda_{n-1} - \mu_{n-1}], \end{aligned} \right.$$

d'où la relation

$$(2) \qquad \begin{aligned} & s_{n-1}(x^1, x^2, \ldots, x^{n-1}) \\ & \quad = p\, s_n(x^1, x^2, \ldots, x^{n-1}, 1) + q\, s_n(x^1, x^2, \ldots, x^{n-1}, 0). \end{aligned}$$

Réciproquement, si des fonctions non négatives s_n satisfont aux égalités (2), on peut trouver deux suites de fonctions λ_n et μ_n non négatives, de somme $\leqq 1$, satisfaisant à (1 *bis*). Le système de jeu de A est donc suffisamment défini par la donnée de la suite de fonctions $\{ s_n \}$. Les conditions (2) dépendent de p; pour obtenir des conditions ne dépendant pas de p, posons

$$(3) \qquad \sigma_n(x^1, x^2, \ldots, x^n) = p^{\nu} q^{n-\nu} s_n(x^1, x^2, \ldots, x^n) \qquad \left(\nu = \sum_{i=1}^{n} x^i \right).$$

Les fonctions σ_n, non négatives, satisfont alors à

$$(5) \quad \begin{cases} \sigma_0 = 1, \\ \sigma_{n-1}(x^1, x^2, \ldots, x^{n-1}) \\ \quad = \sigma_n(x^1, x^2, \ldots, x^{n-1}, 0) + \sigma_n(x^1, x^2, \ldots, x^{n-1}, 0). \end{cases}$$

Nous désignerons par σ une suite $\{\sigma_n\}$ de telles fonctions, et appellerons martingale le système de jeu correspondant. *Si le joueur* A *adopte la manière de jouer définie par* σ, la proposition

« A ne gagne pas indéfiniment »

se traduit par l'inégalité

$$(5) \qquad \underset{n=1,2,3,\ldots}{\text{B sup.}}\ s_n(x^1, x^2, \ldots, x^n) < \infty.$$

(la notation $< \infty$ étant une notation abrégée pour indiquer qu'un nombre est fini).

L'inégalité (5) est une condition imposée à la suite x^1, x^2, $\ldots$, x^n, ... La borne supérieure de $s_n(x^1, x^2, x^u, \ldots)$ peut varier d'une suite à l'autre.

3. Élargissement à l'aide de la notion de martingale de l'axiome d'irrégularité au sens de de Misès-Wald. — Il ressort de l'exposé de M. de Misès que la notion de sélection et d'invariance de la fréquence était destinée dans son esprit à exprimer l'impossibilité de gagner indéfiniment à un jeu équitable : « *Le fait que l'on ne peut modifier ses chances par un choix systématique des coups* [*sur lesquels on mise*] *compte parmi les conceptions les plus essentielles qui sont, à nos yeux, indissolublement liées à la notion de* « *hasard* » *et de jeu de hasard* ». (de Misès, [3], p. 11, *in fine*). Il y a là une restriction très importante; M. de Misès ne cherche pas à traduire d'une manière tout à fait générale la conduite du joueur qui cherche à modifier ses chances. Il ne parle que d'un *choix systématique des coups*, et ne cherche pas à exprimer le fait que l'on peut également chercher à modifier ses chances par une *répartition systématique des mises*. Nous verrons (p. 67) que la restriction faite par M. de Misès l'empêche d'atteindre pleinement, à notre avis du moins, le but qu'il semble s'être proposé en écrivant le passage cité. Nous allons voir ce qu'il en est, du moins en définissant les collectifs comme l'a fait M. Wald (p. 27), pour être sûr de l'existence de ces collectifs.

Soit une suite x telle que pour i égal à un nombre premier, on ait $x^i = 1$. Il est facile de concevoir que A puisse au jeu décrit plus haut, gagner indéfiniment s'il possède ce renseignement sur la suite x : il n'aura qu'à jouer sur les coups de rang premier, en misant sur 1, et s'abstenir aux autres coups. Mais la question se présente en fait d'une manière beaucoup plus compliquée; il est exigé, par exemple, pour que x soit un collectif, que dans la suite des coups de rangs premiers, la fréquence des 1 ait pour limite p. Supposons que A sache qu'il n'en sera pas ainsi, mais *rien de plus;* d'une manière précise, soit

$$(6) \qquad i_1, \quad i_2, \quad \ldots, \quad i_n, \quad \ldots,$$

la suite des nombres premiers et soit f_ν la fréquence des 1 dans la suite $x^{i_1}, x^{i_2}, \ldots, x^{i_\nu}$. Si A sait seulement que

$$\overline{\lim_{\nu=\infty}} f_\nu \neq \underline{\lim_{\nu=\infty}} f_\nu \qquad \text{ou que} \qquad \lim_{\nu=\infty} f_\nu \neq p,$$

est-il possible qu'il puisse utiliser ce renseignement pour gagner indéfiniment? Ce n'est pas évident *a priori*; nous verrons qu'il en est bien ainsi (p. 76).

Faisons abstraction des suites ne contenant qu'un nombre fini de 1 [1]; une suite s étant donnée (§ 2), et p étant un nombre positif inférieur à 1, nous en déduirons une suite de fonctions s_n d'après les égalités (3). Appelons $H(\sigma; p)$ l'ensemble des points du segment (0,1) correspondant à des suites x (p. 33) satisfaisant à l'inégalité (5). Nous avons alors :

THÉORÈME 1. — *A tout ensemble* G *de* p*-mesure nulle nous pouvons associer une suite* σ *telle que le complémentaire de* $H(\sigma; p)$ *contienne* G.

Démonstration. — Soit $F(x)$ la p-mesure du segment $(0, x)$. La p-mesure d'un ensemble M sera alors

$$\mu(M) = \int_M dF(x).$$

---

[1] Pour lesquelles il est aisé de concevoir un système de jeu permettant de gagner indéfiniment.

Soit G un ensemble de p-mesure nulle. Quel que soit n, nous pouvons trouver un ensemble G_n, formé de la somme d'une infinité dénombrable de segments ouverts, contenant G, et tel que $\mu(G_n) = 2^{-2n}$.

Soit $f_n(x)$ une fonction nulle en tout point du complémentaire de G_n, et égale à 2^n sur G^n ; nous aurons alors

$$(7) \qquad \int f_n(x)\, d\,\mathrm{F}(x) = 2^{-n},$$

et si nous posons

$$f(x) = \sum_1^{\infty} f_n(x),$$

nous aurons

$$\int f(x)\, d\,\mathrm{F}(x) = 1.$$

Soit $I_{x^1x^2\ldots x^n}$ l'intervalle ensemble des points représentant les suites dont les n premiers termes sont $x^1, x^2, \ldots, x^n$. Considérons les fonctions $\sigma_n(x^1, x^2, \ldots, x^n)$ définies par

$$(8) \qquad \sigma_n(x^1, x^2, \ldots, x^n) = \int_{I_{x^1x^2\ldots x^n}} f(x)\, d\,\mathrm{F}(x), \qquad \sigma_0 = \int d\mathrm{F}(x).$$

Elles sont $\geqq 0$, et $\sigma_0 = 1$. Les relations (4) sont manifestement vérifiées. Considérons l'ensemble $H(\sigma; p)$ correspondant (σ désignant la suite $\{\sigma_n\}$. Soit x_0 un point de G et soit

$$x_0^1, \quad x_0^2, \quad \ldots, \quad x_0^n, \quad \ldots$$

la suite correspondante.

Quel que soit n, il existe un segment ouvert I_n parmi ceux constituant G_n, qui contient x_0 ; il existe donc une valeur de m, soit m_n, telle que le segment $I_{x_0^1 x_0^2 \ldots x_0^{m_n}}$ soit intérieur à I_n ; soit μ_0 la p-mesure de $I_{x_0^1 \ldots x_0^{m_n}}$.

Puisque sur le segment I_n nous avons $f_n(x) = 2^n$, nous aurons d'après (8)

$$(9) \qquad \sigma_{m_n}(x_0^1, x_0^2, \ldots, x_0^{m_n}) \geqq \mu_0\, 2^n.$$

Soit ν le nombre de 1 figurant dans les m_n premiers termes de x_0 ; d'après la définition de la p-mesure, nous aurons

$$\mu_0 = p^{\nu} q^{m_n - \nu},$$

et par conséquent, d'après (3) et (9),

$$s_{m_n} = p^{-\nu} q^{\nu - m_n} \sigma_{m_n} = \frac{1}{\mu_0} \sigma_{m_n} \geqq 2^n,$$

ce qui montre que x_0 appartient au complémentaire de $H(\sigma; p)$. Le corollaire suivant est immédiat, puisque tout ensemble $K(\mathfrak{S}; p)$ est de p-mesure égale à 1.

Corollaire. — *A tout système de sélections* $\mathfrak{S}$ *et à tout nombre* p *nous pouvons associer une suite* σ *telle que l'ensemble* $H(\sigma; p)$ *soit intérieur à l'ensemble* $K(\mathfrak{S}; p)$.

Ce corollaire montre en particulier que si A a pour unique renseignement sur la suite x que ce n'est pas un collectif $K(\mathfrak{S}; p)$ ($\mathfrak{S}$ étant un système connu), il peut diriger son jeu de manière à gagner indéfiniment. C'est ce que nous avions annoncé plus haut (p. 74).

Les conditions d'irrégularité de M. Wald peuvent donc être exprimées au moyen de la notion de système de jeu ou « martingale ». Mais la réciproque n'est pas vraie; d'après les théorèmes 4 (p. 39) et 1 (p. 74) il existe des $H(\sigma; p)$ que l'on ne peut enfermer dans aucun $K(\mathfrak{S}; p)$. En d'autres termes : il existe pour un joueur des manières de diriger son jeu telles que, quelle que soit la façon de définir un collectif au sens de M. Wald (c'est-à-dire quel que soit le choix du système dénombrable $\mathfrak{S}$ de sélections) on n'est pas sûr, si l'on sait que la suite des résultats sera un collectif, que le joueur ne gagnera pas indéfiniment.

4. Conditions d'irrégularité fondées sur la notion de la martingale. — Soient les conditions d'irrégularité de la forme

« x appartient à $H(\sigma; p)$ ».

D'après le théorème 1 que nous venons de démontrer, ces conditions satisfont à la deuxième partie de la condition Γ (p. 70). Quand à la première, elle résulte du théorème :

Théorème 2. — *Quel que soit* σ, *l'ensemble* $H(\sigma; p)$ *est de* p*-mesure égale à* 1.

Ce théorème est un cas particulier d'un théorème que nous démontrerons plus tard (p. 86). Nous ne le démontrons pas ici pour ne pas faire double emploi.

Nous parvenons donc à la conclusion : on peut définir l'irrégularité d'une suite d'une manière plus complète, au sens de la condition Γ de la page 70, qu'en utilisant les sélections, en poussant plus loin les conséquences du « principe de l'impossibilité d'un système de jeu » indiqué par M. de Misès, mais au prix, il faut le reconnaître, d'une complication plus grande des définitions.

En résumé : la condition primitive de M. de Misès, trop stricte, est impossible à réaliser; la condition de M. Wald, moins stricte, est possible à réaliser, mais il y a certaines propriétés, de probabilité nulle au sens classique modernisé, que l'on ne peut exclure, quelle que soit la forme donnée à la condition de M. Wald.

La condition d'irrégularité par la martingale échappe à cet inconvénient de la condition de M. Wald (qu'elle contient d'ailleurs comme cas particulier). Nous citons pour mémoire la condition d'irrégularité imaginée par les auteurs de la théorie des suites indifférentes; c'est un cas particulier de la condition de M. Wald, et elle est entachée du même défaut que cette dernière.

Mais la condition d'irrégularité par la martingale est relative; elle suppose un choix préalable des propriétés (de probabilité nulle) à exclure. Si, dans un certain sens, elle résout la question de l'irrégularité plus complètement que la condition de M. Wald, elle ne parvient pas à donner un modèle arithmétique d'une suite présentant *tous* les caractères d'une suite prise au hasard; ce dernier problème est considéré par nous comme insoluble, et nous nous soumettons sur ce point à l'opinion de nombreux mathématiciens, parmi lesquels MM. E. Borel, Fréchet, P. Lévy.

Par ailleurs, quelle que soit la manière dont on définit l'irrégularité, tous les inconvénients de la définition de la probabilité par la fréquence subsistent. Nous allons maintenant montrer qu'indépendamment du rôle qu'elle peut jouer dans la discussion de la théorie du collectif, la notion de martingale peut être utilisée, en restant dans le cadre de la théorie classique modernisée, pour simplifier les démonstrations de théorèmes connus ou même pour obtenir des résultats nouveaux. Ceci fera l'objet du chapitre suivant.

CHAPITRE V.

APPLICATIONS DE LA NOTION DE MARTINGALE.

SOMMAIRE. — Première section : Jeu discontinu. 1. Conditions d'équitabilité, — 2. Théorème de la ruine des joueurs. — 3. Étude du comportement de la fréquence dans le cas de Bernoulli.

Deuxième section : Jeu continu. — 1. Jeu continu équitable. — 2. Formule de la ruine des Joueurs. — 3. Remarque sur les probabilités conditionnelles. — 4. Cas de Chapmann-Kolmogoroff. — 5. Borne supérieure stochastique. — 6. Recherche de fonctions limites. — 7. Cas d'une fonction à accroissements indépendants obéissant à une loi de Gauss. — 8. Plus grande valeur.

Les fondements de toutes les sciences resteront toujours controversés. Quoi qu'on pense de l'utilité des martingales pour élucider, comme nous avons tenté de le faire, les difficultés d'une définition de l'irrégularité, nous nous proposons de montrer ici que cette notion peut également servir des buts mathématiques précis. Nous allons dans ce chapitre passer en revue un certain nombre de questions classiques; nous associerons à chaque problème un jeu équitable hypothétique, et nous étudierons l'espérance mathématique correspondante. Nos considérations reviennent au fond à traiter dans les cas étudiés, le problème de la ruine des joueurs, mais d'un point de vue nouveau : la probabilité de ruine ne sera pas évaluée pour elle-même, mais pour servir à établir certaines propositions sur la limitation des écarts.

PREMIÈRE SECTION.

JEU DISCONTINU.

1. **Conditions d'équitabilité.** — Soit une suite de variables aléatoires

$$X = X_1, X_2, \dots, X_n, \dots \qquad (-\infty < X < +\infty).$$

Quel que soit n, on suppose connue la loi de probabilité de X_1, $X_2, \ldots, X_n$, c'est-à-dire la fonction

$$(1) \qquad F_n(x_1, x_2, \ldots x_n) = \Pr.\{X_1 < x_1, X_2 < x_2, \ldots, X_n < x_n\}.$$

Nous avons évidemment quel que soit n :

$$(2) \qquad F_n(x_1, x_2, \ldots, x_n) = F_{n+1}(x_1, x_2, \ldots, x_n, +\infty),$$

puisque les x_n ne peuvent prendre que des valeurs finies.

Considérons le jeu suivant : un joueur A mise une somme égale à 1, contre la promesse de recevoir la somme $s_1(X_1)$ après la détermination de X_1, ou $s_2(X_1, X_2)$ après la détermination de X_1 et X_2, ou ..., ou $s_n(X_1, X_2, \ldots, X_n)$ après la détermination de $X_1, X_2, \ldots, X_{n-1}$ et X_n, ou ..., étant entendu que A est libre d'interrompre le jeu quand il lui plaît, c'est-à-dire de toucher après un coup quelconque, la somme convenue et en garder la possession définitive. A quelles conditions un tel jeu est-il équitable ?

La réponse intuitive à cette question est la suivante : après chaque coup, l'espérance mathématique du joueur, relative aux coups qui vont suivre, est égale à la valeur de la somme qu'il possède à l'instant considéré, et cela *indépendamment de la conduite qu'il se propose de tenir*, en s'inspirant par exemple des résultats de coups déjà joués. Nous allons traduire cette condition en termes mathématiques.

Par définition, les fonctions $s_1(x_1)$, $s_2(x_1, x_2)$, ... sont essentiellement non négatives. Nous ajoutons la fonction constante $s_0 = 1$ pour unifier les notations.

$s_1(x_1)$ est évidemment soumise à

$$(3) \qquad \int s_1(x_1)\, d\,F_1(x_1) = 1.$$

Pour s_2, la question se complique un peu; faisons une hypothèse sur la conduite de A après le premier coup : supposons que A décide de continuer le jeu si X_1 prend une valeur appartenant à un ensemble E_1 de l'axe des x et de s'arrêter si X_1 appartient au complémentaire $\overline{E}_1$. L'espérance mathématique de A est alors, au début de la partie en supposant qu'il s'arrête après le deuxième coup :

$$(4)\qquad \int_{\overline{E_1}} s_1(x_1)\,d\,F_1(x_1) + \int_{E_1} s_2(x_1,\,x_2)\,d\,F_2(x_1,\,x_2).$$

Cette espérance doit être égale à 1, quel que soit E_1.

Si E_1 est constitué par l'ensemble des valeurs de x supérieures ou égales à une valeur donnée τ, nous aurons

$$(5)\qquad \int_{x_1<\tau} s_1(x_1)\,d\,F_1(x_1) + \int_{x_1\geqq\tau} s_2(x_1,\,x_2)\,d\,F_2(x_1,\,x_2) = 1.$$

Considérons le deuxième terme du premier membre de (5). C'est la valeur moyenne de la fonction $\sigma_2(X_1,\,X_2)$ égale à $s_2(X_1,\,X_2)$ pour $X_1 \geqq \tau$ et à o pour $X_1 < \tau$. Supposons que cette valeur moyenne existe quel que soit τ; nous pouvons alors l'évaluer d'une autre manière, en utilisant la notion de *probabilité conditionnelle* introduite par M. P. Lévy (P. Lévy, [1], p. 68). Supposons que l'on sache que $x_1 \leqq X_1 < x_1 + h$; quelle est la probabilité que $X_2 < x_2$? La probabilité que $x_1 \leqq X_1 < x_1 + h$ et que $X_2 < x_2$ est égale à

$$F_2(x_1 + h,\,x_2) - F_2(x_1,\,x_2).$$

En désignant par ϖ la probabilité cherchée, nous devrons avoir, par application du principe des probabilités composées,

$$[F_1(x_1 + h) - F_1(x_1)]\varpi = F_2(x_1 + h,\,x_2) - F(x_1,\,x_2),$$

d'où

$$\varpi = \frac{F_2(x_1 + h,\,x_2) - F_2(x_1,\,x_2)}{F_1(x_1 + h) - F_1(x_1)}.$$

Si l'on sait que $X_1 = x_1$, et que l'on cherche la probabilité que $X_2 < x_2$, on est amené à chercher la limite du rapport ci-dessus pour $h = 0$. Cette limite peut ne pas être définie pour toutes les valeurs de x_1 et x_2, d'où une certaine indécision dans la définition de la probabilité cherchée. Nous adopterons une definition équivalente, appelons avec M. P. Lévy, *fonction de répartition conditionnelle de* X_2 quand X_1 est connu, la fonction $\Phi_2(X_1,\,x_2)$, solution de l'équation intégrale

$$(6)\qquad F_2(x_1,\,x_2) = \int_{t<x_1} \Phi_2(t,\,x_2)\,d\,F_1(t).$$

Cette fonction n'est définie pour chaque valeur de x_2 qu'à un ensemble, de valeurs de X_1, sur lequel $\int dF_1 = 0$, près. La valeur moyenne de $\sigma_2(X_1, X_2)$ quand X_1 est connu est alors, par définition, la variable aléatoire

$$\mathcal{M}_{X_1}\{\sigma_2\} = \int_{x_2} \sigma_2(X_1, x_2)\, d\Phi_2(X_1; x_2). \tag{7}$$

On démontre alors, par application du théorème de Fubini (P. Lévy, [1], p. 72) que la valeur moyenne $\mathcal{M}\{\sigma_2\}$ de $\sigma_2(X_1, X_2)$ a pour expression

$$\mathcal{M}\{\sigma_2\} = \mathcal{M}\left[\mathcal{M}_{X_1}\{\sigma_2\}\right],$$

ce qui nous permet d'écrire, au lieu de (5),

$$\int_{x_1<\tau} s_1\, dF_1 + \int_{x_1 \geqq \tau}\left[\int_{x_2} s_2\, d\Phi_2(x_1; x_2)\right] dF_1 = 1. \tag{8}$$

Pour $\tau = -\infty$ nous obtenons $\mathcal{M}\{\sigma_2(X_1, X_2)\} = 1$. En tenant compte de cette dernière égalité, (8) devient

$$\int_{x_1<\tau} s_1(x_1)\, dF(x_1) = \int_{x_1<\tau}\left[\int_{x_2} s_2(x_1, x_2)\, d\Phi_2(x_1; x_2)\right] dF_1(x_1).$$

Ce qui exige que l'on ait, sauf peut-être sur un ensemble sur lequel $\int dF_1 = 0$, l'égalité

$$s_1(x_1) = \int_{x_2} s_2(x_1, x_2)\, d\Phi_2(x_1; x_2), \tag{9}$$

que nous écrivons plus simplement

$$s_1(x_1) = \mathcal{M}_{x_1}\{s_2(x_1, X_2)\}. \tag{10}$$

Si de même $\Phi_n(X_1, X_2, \ldots, X_{n-1}; x_n)$ est la fonction de répartition conditionnelle de X_n quand on connaît $X_1, \ldots, X_{n-1}$, c'est-à-dire une solution de l'équation

$$\begin{aligned}(11)\quad & F_n(x_1, x_2, \ldots, x_n) \\ & = \int_{t_1<x_1} \cdots \int_{t_{n-1}<x_{n-1}} \Phi_n(t_1, t_2, \ldots, t_{n-1}; x_n)\, dF_{n-1}(t_1, t_2, \ldots, t_{n-1}),\end{aligned}$$

et que nous considérions la valeur moyenne conditionnelle de $s_n(X_1, X_2, \ldots, X_n)$, c'est-à-dire la variable aléatoire

$$\mathfrak{M}_{X_1, X_2, \ldots, X_{n-1}}\{s_n\} = \int_{x_n} s_n(X_1, X_2, \ldots, X_{n-1}; x_n)\, d\,\Phi_n(X_1, \ldots, X_{n-1}; x_n),$$

nous devons avoir

$$(12) \quad \mathfrak{M}_{x_1, x_2, \ldots, x_{n-1}}\{s_n(x_1, x_2, \ldots, x_{n-1}; X_n)\} = s_{n-1}(x_1, x_2, \ldots, x_{n-1}).$$

Inversement, supposons (12) vérifiée pour toutes les valeurs de n. Soit l'espace E à une infinité de coordonnées $x_1, x_2, \ldots, x_n, \ldots$. La manière dont A arrêtera le jeu (en supposant qu'elle ne dépend pas d'éléments aléatoires) sera caractérisée par une suite d'ensembles disjoints $E_1, E_2, \ldots, E_n, \ldots$ dans E, ayant la signification suivante : E_n, d'une manière générale, a une fonction caractéristique ne dépendant que de $x_1, x_2, \ldots, x_n$; si le point $X_1, X_2, \ldots, X_n, \ldots$ appartient à E_n, A se retire du jeu après le $n^{\text{ième}}$ coup. L'espérance mathématique est donc représentée par la série

$$(13) \qquad \mathcal{E} = \int_{E_1} s_1\, dF_1 + \int_{E_2} s_2\, dF_2 + \ldots + \int_{E_n} s_n\, dF_n + \ldots.$$

Soit S_n la somme des n premiers termes de cette série. En tenant compte de (12) et en transformant le n^{ieme} terme d'après la formule générale (P. Lévy, [1], p. 73)

$$\mathfrak{M}\{\varphi(X, Y)\} = \mathfrak{M}\{\mathfrak{M}_X\, \varphi(x, Y)\},$$

où $Y = X_n$ et où X est une variable aléatoire représentant l'ensemble des variables $X_1, X_2, \ldots, X_{n-1}$, nous trouvons

$$S_n = \int_{E_1} s_1\, dF_1 + \ldots + \int_{E_{n-1}+E_n} s_{n-1}\, dF_{n-1},$$

d'où, par itération de ce procédé,

$$S_n = \int_{E_{(n)}} s_1\, dF_1, \qquad E_{(n)} = E_1 + E_2 + \ldots + E_n.$$

Nous devons supposer, pour que l'espérance mathématique de A soit déterminée que $E_1 + E_2 + \ldots + E_n + \ldots = E$ ou du moins que la probabilité que la partie s'arrête soit égale à 1. Dans ces conditions nous avons

$$\mathcal{E} = \lim_{n=\infty} S_n = \mathfrak{M}\{s_1\} = 1.$$

Le jeu est bien équitable, quels que soient les $E_1, E_2, \ldots, E_n, \ldots$.

Dans ces conditions nous poserons la définition suivante :

DÉFINITION 1. — *Soit* $X_1, X_2, \ldots, X_n, \ldots$ *une suite de variables aléatoires, telle que les probabilités*

$$\text{Pr.}\{X_1 < x_1, X_2 < x_2, \ldots, X_n < x_n\} \qquad (n = 1, 2, 3, \ldots)$$

soient bien définies et que les X_i *ne puissent prendre que des valeurs finies.*

Soit une suite de fonctions $s_0, s_1(x_1), s_2(x_1, x_2), \ldots$ *non négatives telles que*

$$(14)\quad \begin{cases} s_0 = 1, \\ \mathfrak{M}_{x_1, x_2, \ldots, x_{n-1}}\{s_n(x_1, x_2, \ldots, x_{n-1}, X_n)\} = s_{n-1}(x_1, x_2, \ldots, x_{n-1}), \end{cases}$$

où $\mathfrak{M}_X\{Y\}$ *représente d'une manière générale la valeur moyenne conditionnelle de la variable* Y *quand on connaît la position du point aléatoire* X, *au sens indiqué par M. P. Lévy.*

Dans ces conditions, nous dirons que la suite $\{s_n\}$ *définit une martingale ou un jeu équitable.*

2. Théorème de la ruine des joueurs. — Le problème de la ruine des joueurs est le suivant : Soient deux joueurs A et B disposant respectivement des sommes a et b. Ils jouent à un jeu équitable, on demande la probabilité que l'un d'eux, soit A, se ruine, c'est-à-dire perde la somme qu'il possédait au début et les gains qu'il aurait pu réaliser. Soient p_1, p_2 et q les probabilités que A se ruine, que B se ruine, et que la partie se termine sans que ni A ni B se ruinent (si la partie peut être indéfiniment prolongée, q sera justement la probabilité que la partie se prolonge indéfiniment). Si A se ruine, B touchera la somme a; le jeu étant équitable, nous aurons, en écrivant que son espérance mathématique est égale à sa mise initiale

$$b = p_1(a + b) + b' \qquad (b' \geqq 0),$$

le terme b' correspondant à l'éventualité de probabilité q. Nous avons donc

$$p_1 \leqq \frac{b}{a+b} \qquad \text{et, de même,} \qquad p_2 \leqq \frac{a}{a+b}.$$

En particulier, si les conditions du jeu sont telles que $q = 0$, nous aurons

$$p_1 = \frac{b}{a+b}, \qquad p_2 = \frac{a}{a+b}.$$

Le raisonnement que nous venons de faire n'entre pas dans les détails du jeu auquel se livrent A et B. Dans les traités classiques, tels que celui de J. Bertrand, on fait des hypothèses plus précises; on suppose par exemple que la partie est une succession de coups semblables, avec des enjeux égaux, et l'on démontre les formules ci-dessus. Si, comme c'est généralement le cas, $q = 0$, et si l'on suppose de plus que B représente le public, c'est-à-dire si l'on suppose que b est pratiquement infini, on obtient la conclusion $p_1 = 0$: A est certain de se ruiner.

Supposons maintenant que A possède une somme égale à l'unité. et qu'il veuille jouer jusqu'à ce qu'il soit en possession d'une somme $\lambda > 1$. On peut supposer qu'il joue contre B qui possède la somme $\lambda - 1$. La probabilité que A arrive à ses fins est p_2. Elle est au plus égale à $\frac{1}{\lambda}$.

Les raisonnements classiques portent sur des cas particuliers, et le raisonnement général que nous avons présenté peut paraître sommaire. Nous allons exprimer et démontrer d'une manière plus précise, dans le cas du jeu discontinu décrit dans le paragraphe précédent, le résultat énoncé en dernier lieu, ce qui nous donne :

Théorème 1. — *Soit $\{s_n\}$ une suite définissant une martingale et λ un nombre > 1. Nous avons l'inégalité*

$$\text{(15)} \qquad \Pr.\left\{ \text{B.}\sup_{(n)}.\, s_n(X_1, X_2, \ldots, X_n) \geqq \lambda \right\} \leqq \frac{1}{\lambda}.$$

La démonstration est immédiate. Considérons l'expression (13), et faisons $E_n = E$; nous déduisons de la démonstration qui suit que, quel que soit n,

$$\text{(16)} \qquad \mathfrak{M}\{s_n(X_1, X_2, \ldots, X_n)\} = 1.$$

Considérons maintenant la suite de fonctions $\sigma_n(x_1, x_2, \ldots, x_n)$ non négatives, ainsi définies

$$(17)\begin{cases} \sigma_n(x_1, x_2, \ldots, x_{n-1}, x_n) = \quad s_n(x_1, \ldots, x_n) & \text{si } \sigma_{n-1}(x_1, \ldots, x_{n-1}) \leqq \lambda, \\ \sigma_n(x_1, x_2, \ldots, x_{n-1}, x_n) = \sigma_{n-1}(x_1, \ldots, x_{n-1}) & \text{si } \sigma_{n-1}(x_1, \ldots, x_{n-1}) > \lambda. \end{cases}$$

Il suffit de se reporter à la définition (1) pour voir que la suite $\{\sigma_n\}$ définit une martingale. Donc, d'après (16) $\mathfrak{M}\{\sigma_n\} = 1$ et, puisque $\sigma_n \geqq 0$

$$\text{Pr.}\left\{\sigma_n \geqq \lambda\right\} \leqq \frac{1}{\lambda} \qquad (\lambda > 1).$$

En reprenant le raisonnement utilisé pour démontrer l'inégalité (17) de la page 36, on obtient sans peine la démonstration du théorème 1.

3. Étude du comportement de la fréquence dans le cas de Bernoulli. — Supposons que les variables X_i soient indépendantes et que

$$(18) \qquad \text{Pr.}\left\{X_i = 1\right\} = p, \qquad \text{Pr.}\left\{X_i = 0\right\} = 0 \qquad (p + q = 1).$$

Les conditions (14) deviennent alors

$$(14') \begin{cases} s_0 = 1, \\ s_{n-1}(x_1, x_2, \ldots, x_{n-1}) \\ \qquad = p\, s_n(x_1, x_2, \ldots, x_{n-1}, 1) + q\, s_n(x_1, x_2, \ldots, x_{n-1}, 0), \end{cases}$$

elles sont identiques aux conditions (3) et (4) du Chapitre IV.

On a, quel que soit λ,

$$\text{Pr.}\left\{\text{B. sup. } s_n = \infty\right\} < \text{Pr.}\left\{\text{B. sup. } s_n \geqq \lambda\right\} \leqq \frac{1}{\lambda},$$

d'où

$$\text{Pr.}\left\{\text{B. sup. } s_n = \infty\right\} = 0,$$

ce qui fournit la démonstration du théorème 2, p. 76.

Nous allons essayer de trouver une forme commode pour s_n. Cherchons si nous pouvons mettre s_n sous la forme d'une fonction de deux variables x et y égales aux répétitions de 1 et 0 dans la suite $X_1, X_2, \ldots, X_n$, c'est-à-dire sous la forme

$$s_n = s(x, y) \qquad \text{avec } x = \sum_1^n X_i,\ y = n - x.$$

Les conditions (14) deviennent

$$(14'') \quad s(0, 0) = 1, \qquad s(x, y) = p\, s(x+1, y) + q\, s(x, y+1) \qquad (s \geqq 0)$$

et, en posant $\sigma(x, y) = p^x q^y s(x, y)$,

$$(19) \qquad \sigma(x, y) = \sigma(x+1, y) + \sigma(x, y+1).$$

Les relations (19) permettent de calculer $\sigma(x, y)$ à partir de la suite des valeurs

$$\sigma(0, y) = \nu_y \qquad (y = 0, 1, 2, \ldots).$$

En effet, nous tirons de (19)

$$\sigma(x+1, y) = -[\sigma(x, y+1) - \sigma(x, y)] = -\Delta\,\sigma(x, y),$$

en désignant par $\Delta f(x, y)$ d'une manière générale, la différence première $f(x, y+1) - f(x, y)$. De proche en proche, nous obtenons donc

$$\begin{aligned} \sigma(x, y) &= -\Delta\,\sigma(x-1, y), \\ \sigma(x-1, y) &= -\Delta\,\sigma(x-2, y), \\ &\ldots\ldots\ldots\ldots\ldots\ldots, \\ \sigma(2, y) &= -\Delta\,\sigma(1, y), \\ \sigma(1, y) &= -\Delta\nu_y \end{aligned}$$

et, par conséquent,

$$(20) \qquad \sigma(x, y) = (-1)^x \Delta^x \nu_y.$$

Inversement, si une suite ν_y de nombres positifs ou nuls est telle que $\nu_0 = 1$ et que les expressions $(-1)^x \Delta^x \nu_y$ soient toutes positives ou nulles, en posant

$$s(x, y) = (-1)^x p^{-x} q^{-y} \Delta^x \nu_y,$$

on satisfait aux conditions (14″).

Le problème dépend donc de la suite arbitraire $\{\nu_y\}$ assujettie à ce que toutes les expressions (20) soient $\geqq 0$. M. de Finetti, en poursuivant un but différent (de Finetti, [1], p. 28-37) est arrivé à la même condition (20); il a trouvé la solution générale qui est de la forme

$$(21) \qquad \sigma(x, y) = \int_0^1 t^x (1-t)^y \, d\,F(t),$$

$F(t)$ étant une fonction non décroissante, telle que $F(0)=0$, $F(1)=1$. D'où la forme générale de $s(x, y)$

$$s(x, y) = p^{-x}q^{-y}\int_0^1 t^x(1-t)^y\, dF(t), \tag{22}$$

ce qui nous donne d'après (15)

$$\Pr.\left\{ \underset{x+y=1,2,\ldots,n,\ldots}{\text{B. sup.}}\ p^{-x}q^{-y}\int_0^1 t^x(1-t)^y\, dF(t) \geqq \lambda \right\} \leqq \frac{1}{\lambda}. \tag{23}$$

Nous allons utiliser l'inégalité (23) pour démontrer une proposition qui constitue une partie d'un théorème, dû à M. Kolmogoroff, et cité (sans démonstration) par M. P. Lévy ([1], p. 266).

Soit x la répétition d'un événement dans n épreuves indépendantes de probabilité constante p; on se propose d'étudier la probabilité que l'inégalité

$$|x - np| > \varphi(n)\sqrt{2npq}, \tag{24}$$

où $\varphi(n)$ est une fonction > 0 de l'entier n, ait lieu pour une infinité de valeurs de n. si l'on procède à une suite infinie d'épreuves.

Le résultat de M. Kolmogoroff est le suivant : en supposant que $\varphi(n)$ est la suite des valeurs prises par une fonction *croissante* $\varphi(\theta)$, de la variable continue t, pour les valeurs entières de θ, *la condition nécessaire et suffisante pour que la probabilité d'une infinité de réalisations de* (24) *soit nulle est que l'intégrale*

$$\int \varphi\, e^{-\varphi^2}\frac{d\theta}{\theta}$$

soit convergente pour $\theta = +\infty$.

Les valeurs de $\varphi(t)$ pour les valeurs entières de t intervenant seules, on peut aussi considérer la série de terme général $\varphi(n)e^{-\varphi^2(n)}\frac{1}{n}$. Cette série est convergente par exemple pour

$$\varphi(n) = (1+\varepsilon)\sqrt{\operatorname{Log}\operatorname{Log} n} \qquad (\varepsilon > 0).$$

Faisons les hypothèses suivantes :

1° $\varphi(n)$ *est une fonction croissante de n.*

2° *La série de terme général* $\frac{1}{n}\varphi(n)\,e^{-\varphi^2(n)}$ *est convergente.*

$\varphi(n)$ étant croissante, la limite $\lim\limits_{n=\infty}\varphi(n)$ existe. Nous allons voir qu'elle est infinie : si en effet elle était finie, la série considérée aurait, pour les grandes valeurs de n, ses termes comparables à ceux de $\sum\frac{1}{n}$, il y aurait contradiction.

Par ailleurs, la fonction $u\,e^{-u^2}$ est une fonction décroissante de u pour $u > \frac{1}{\sqrt{2}}$; donc le terme général de la série considérée décroît, quand n augmente, à partir d'une certaine valeur, finie, de n. Nous en déduisons que si l'on interpole la fonction $\varphi(n)$, définie pour les valeurs entières de la variable, par une fonction $\varphi(\theta)$ croissante de la variable continue θ, la convergence de la série entraîne celle de l'intégrale $\int_1^\infty \varphi\,e^{-\varphi^2}\frac{d\theta}{\theta}$ (et réciproquement).

Puisque $\varphi\,e^{-\varphi^2}$ décroît pour les grandes valeurs de θ, nous avons

$$\int_1^{\theta_0}\varphi\,e^{-\varphi^2}\frac{d\theta}{\theta} > \varphi(\theta_0)\,e^{-\varphi^2(\theta_0)}\operatorname{Log}\theta_0 - \rho, \tag{25}$$

où ρ est un nombre indépendant de θ_0. Nous en déduisons que pour toutes les valeurs de θ supérieures à une certaine valeur finie, nous avons

$$\varphi(\theta) > \sqrt{\operatorname{Log}\operatorname{Log}\theta}. \tag{26}$$

Si en effet (26) n'avait pas lieu pour une suite de valeurs de θ tendant vers l'infini, nous déduirions de (25) que l'intégrale diverge.

Remarquons que si l'on prend $\varphi(n) = (1+\varepsilon)\sqrt{\operatorname{Log}\operatorname{Log} n}$, les conditions de l'énoncé sont satisfaites. De plus, pour cette fonction $\varphi(n)$ particulière, l'expression $\sqrt{n}\,e^{-\varphi^2(n)}$ est une fonction croissante à partir d'une certaine valeur de u, et infiniment grande avec n. Nous ajouterons, pour une raison de commodité, aux hypothèses 1° et 2° l'hypothèse supplémentaire, qui, nous venons de le voir, n'est pas contradictoire avec les hypothèses déjà faites :

3° *La fontion* $\sqrt{n}\,e^{-\varphi^2(n)}$ *est, pour n suffisamment grand, une fonction croissante de n, infiniment grande avec n.*

La proposition que nous allons démontrer est la suivante :

Si les hypothèses 1°, 2°, 3° *sont réalisées, la probabilité que l'inégalité* (24) *ait lieu pour une infinité de valeurs de n est nulle.*

Les hypothèses 1°, 2°, 3° étant faites sur $\varphi(n)$, nous allons chercher à interpoler $\varphi(n)$ au moyen d'une fonction $\varphi(\theta)$, de la variable continue θ, positive, croissante et telle que $\sqrt{\theta}\, e^{-\varphi^2(\theta)}$ soit croissante pour θ suffisamment grand. Considérons l'intervalle $(n,\ n+1)$ pour θ, et soit $u(\theta)$ une fonction de θ continue et dérivable. Nous avons

$$\frac{d}{d\theta}\sqrt{\theta}\, e^{-u^2} = \frac{e^{-u^2}}{2\sqrt{\theta}}\left[1 - 4u\theta\frac{du}{d\theta}\right].$$

Considérons l'équation différentielle

$$\frac{du}{d\theta} = \frac{1}{4u\theta}.$$

Sur toute courbe intégrale, $\sqrt{\theta}\, e^{-u^2}$ est une constante; donc, d'après l'hypothèse 3°, si nous désignons par $u_1(\theta)$ l'intégrale définie par la condition $u_1(n) = \varphi(n)$, nous aurons

$$\sqrt{n+1}\, e^{-u_1^2(n+1)} < \sqrt{n+1}\, e^{-\varphi^2(n+1)}. \tag{27}$$

Considérons maintenant l'équation différentielle

$$\frac{du}{d\theta} = \frac{k}{4u\theta} \qquad (k = \text{const.})$$

et soit u_k l'intégrale telle que $u_k(n) = \varphi(n)$. u_0 est une constante, donc

$$\sqrt{n+1}\, e^{-u_0^2(n+1)} > \sqrt{n+1}\, e^{-\varphi^2(n+1)}. \tag{28}$$

En comparant (27) et (28) et en observant que si $u_k(n) = \varphi(n)$, $u_k(n+1)$ est une fonction continue de k, nous voyons qu'il existe une valeur de k comprise entre 0 et 1 telle que

$$\sqrt{n+1}\, e^{-u_k^2(n+1)} = \sqrt{n+1}\, e^{-\varphi^2(n+1)}.$$

Par ailleurs

$$\frac{d}{d\theta}\sqrt{\theta}\, e^{-u_k^2} = \frac{e^{-u_k^2}}{2\sqrt{\theta}}(1-k) > 0.$$

La fonction $u_k(\theta)$ déterminée par le procédé ci-dessus réalise l'interpolation cherchée. En opérant de même pour toute valeur de n, nous obtiendrons une fonction $\varphi(\theta)$ croissante telle que $\sqrt{\theta}\, e^{-\varphi^2(\theta)}$ soit croissante pour les valeurs de θ suffisamment grandes, telle que $\int_1^\infty \varphi\, e^{-\varphi^2} \frac{d\theta}{\theta}$ soit finie, et coïncidant avec la fonction $\varphi(n)$ donnée pour θ entier. Dans chaque intervalle de la forme $(n, n+1)$, nous avons, nous l'avons vu,

$$0 < \frac{d\varphi}{d\theta} < \frac{1}{4\varphi\theta}.$$

Or, si nous considérons la fonction $\frac{\varphi}{\sqrt{\theta}}$, nous avons

$$\frac{d}{d\theta}\frac{\varphi}{\sqrt{\theta}} = \frac{1}{\sqrt{\theta}}\left[\frac{d\varphi}{d\theta} - \frac{\varphi}{2\theta}\right] < \frac{1}{2\theta\sqrt{\theta}}\left[\frac{1}{2\varphi} - \varphi\right].$$

Pour θ suffisamment grand, $\frac{\varphi}{\sqrt{\theta}}$ est donc une fonction décroissante. On voit facilement qu'elle tend vers zéro quand θ tend vers l'infini.

Considérons maintenant la fonction de θ définie par

$$z(\theta) = \theta p + \varphi(\theta)\sqrt{2\theta pq} \tag{29}$$

et posons

$$\alpha(\theta) = \frac{z(\theta)}{\theta}.$$

Il existe une valeur de θ, soit θ_0, telle que, pour $\theta > \theta_0$, $\alpha(\theta)$ soit une fonction décroissante de θ, comprise entre p et 1. De plus, nous avons

$$\lim_{\theta=\infty} \alpha(\theta) = p, \qquad \lim_{\theta=\infty} \sqrt{\theta}[\alpha(\theta) - p] = \infty. \tag{30}$$

Supposons que nous ayons déterminé une fonction $s(x, y)$ définie par une expression de la forme (22), où x et y sont $\geqq 0$ (entiers ou non), et telle que

$$s(x+h, y-h) > s(x, y) \qquad (h > 0), \tag{31}$$

$$\lim_{\theta=\infty} s[z(\theta), \theta - z(\theta)] = \infty. \tag{32}$$

Soit E_1 l'événement qui a lieu s'il existe une suite infinie de valeurs de n, soit $\{n_i\}$, telle qu'en désignant par $\{x_i\}$ la suite des répétitions correspondantes, on ait, pour tout i,

$$x_i - n_i p > \varphi(n_i)\sqrt{2 n_i pq}.$$

E_1 implique que pour tout i

$$s(x_i, y_i) \geqq s[z(n_i), n_i - z(n_i)] \qquad (y_i = n_i - x_i)$$

d'après (31). Si donc E_1 se réalise, nous aurons, en désignant par α la répétition u correspondant à n

$$\underset{n=1,2,\ldots,ad\ \text{inf.}}{\text{B. sup.}}\ s(x, y) = \infty \qquad (y = n - x)$$

et, par conséquent, d'après (23), la *probabilité de* E_1 *sera nulle*. On traiterait de même l'événement E_2 correspondant aux inégalités

$$x_i - n_i p < -\varphi(n_i)\sqrt{2 n_i pq}.$$

Notre méthode de démonstration se ramène donc à la recherche d'une fonction $s(x, y)$ *satisfaisant à* (31) *et* (32).

Nous prendrons pour $s(x, y)$ une fonction de la forme

$$(33) \qquad s(x, y) = p^{-x} q^{-y} \int_p^1 t^x (1-t)^y f(t)\, dt,$$

$f(t)$ étant une fonction non négative, telle que

$$\int_p^1 f(t)\, dt = 1.$$

Un calcul simple montre que (31) est vérifiée. Passons à (32), et posons

$$w(\theta) = s[z(\theta), \theta + z(\theta)].$$

Astreignons $f(t)$ à être décroissante; nous aurons alors

$$w(\theta) \geqq p^{-z} q^{z-\theta} f(\alpha) \int_p^{\alpha} t^z (1-t)^{\theta - z}\, dt \qquad \left(\alpha = \frac{z}{\theta}\right).$$

Posons

$$\xi(t) = \alpha \operatorname{Log} t + (1-\alpha)\operatorname{Log}(1-t) - \alpha \operatorname{Log} p - (1-\alpha)\operatorname{Log} q;$$

nous aurons

$$\varpi(\theta) \geqq f(\alpha) \int_p^\alpha e^{\theta\xi}\, dt.$$

Quand $\theta \to \infty$, $\alpha \to p$ et $\theta\xi \to \infty$. Nous avons au second membre une forme indéterminée ; nous allons lever l'indétermination en développant ξ par la formule de Taylor

$$\xi(t) = \xi(\alpha) + \frac{t-\alpha}{1}\xi'(\alpha) + \frac{(t-\alpha)^2}{1.2}\xi''(\alpha) + \frac{(t-\alpha)^3}{1.2.3}\xi'''(\beta),$$

où β est une valeur comprise entre α et t. On constate que $\xi^{\text{IV}}(t) < 0$; on en déduit que dans l'intervalle $p < t < \alpha$

$$\xi(t) \geqq \xi(\alpha) + \frac{t-\alpha}{1}\xi'(\alpha) + \frac{(t-\alpha)^2}{1.2}\xi''(\alpha) + \frac{(t-\alpha)^3}{1,2.3}\xi'''(p),$$

ou encore, en explicitant

$$\xi(t) \geqq \alpha \operatorname{Log}\frac{\alpha}{p} + (1-\alpha)\operatorname{Log}\frac{1-\alpha}{q} - \frac{1+\eta}{2\alpha(1-\alpha)}(t-\alpha)^2$$

avec

$$\eta = \frac{\alpha - t}{6}\left\{2\alpha(1-\alpha)\left[\frac{2\alpha}{p^3} - \frac{2(1-\alpha)}{q^3}\right]\right\}.$$

Quand α varie de p à 1, l'accolade a, en valeur absolue, une borne supérieure bien déterminée, qui ne dépend que de p et q ; il existe donc une constante k telle que

$$|\eta| \leqq k(\alpha - t) < k(\alpha - p),$$

d'où

$$\xi(t) \geqq \alpha \operatorname{Log}\frac{\alpha}{p} + (1-\alpha)\operatorname{Log}\frac{1-\alpha}{q} - \frac{1+k(\alpha-p)}{2\alpha(1-\alpha)}(t-\alpha)^2.$$

Opérons de même pour

$$\tau(\alpha) = \alpha \operatorname{Log}\frac{\alpha}{p} + (1-\alpha)\operatorname{Log}\frac{1-\alpha}{q}.$$

De $\tau^{\text{IV}}(\alpha) > 0$, nous déduisons

$$\tau(\alpha) \geqq \tau(p) + (\alpha - p)\tau'(p) + \frac{(\alpha-p)^2}{2}\tau''(p) + \frac{(\alpha-p)^3}{6}\tau'''(p),$$

d'où, en explicitant

$$\tau(\alpha) \geqq (1+\eta_1)\frac{(\alpha-p)^2}{2pq}$$

avec

$$\eta_1 = \frac{pq}{3}(\alpha - p)\left(\frac{1}{q^2} - \frac{1}{p^2}\right) = k_1(\alpha - p).$$

Nous obtenons en définitive pour $\xi(t)$

$$\xi(t) \geqq [1 + k_1(\alpha - p)]\frac{(\alpha - p)^2}{2pq} - [1 + k(\alpha - p)]\frac{(t - \alpha)^2}{2\alpha(1 - \alpha)} = \xi_1(t),$$

où k et k_1 sont des constantes. Donc,

$$\varpi(\theta) \geqq f(\alpha)\int_p^\alpha e^{\theta\xi_1}\,dt.$$

Faisons le changement de variable

$$t = \alpha - \frac{u}{\sqrt{\theta}}\sqrt{\frac{2\alpha(1 - \alpha)}{1 + k(\alpha - p)}};$$

nous obtenons

$$\varpi(\theta) \geqq f(\alpha)\frac{e^{\varphi^2[1 + k_1(\alpha - p)]}}{\sqrt{\theta}}\sqrt{\frac{2\alpha(1 - \alpha)}{1 + k(\alpha - p)}}\int_0^l e^{-u^2}\,du,$$

avec

$$l = \sqrt{\frac{1 + k(\alpha - p)}{2\alpha(1 - \alpha)}}\,\varphi\sqrt{2pq}.$$

Quand θ tend vers l'infini, α tend vers p, l tend vers l'infini, et $\varphi^2(\alpha - p)$, qui est comparable à $\frac{\varphi^3}{\sqrt{\theta}}$, tend vers zéro (ainsi qu'on s'en assure facilement en tenant compte de ce que $\sqrt{\theta}\,e^{-\varphi^2}$ tend vers l'infini). Pour assurer $\lim\limits_{\theta=\infty} \varpi(\theta) = \infty$, il suffira d'assurer la condition

$$\lim_{\theta=\infty} f(\alpha)\frac{e^{\varphi^2}}{\sqrt{\theta}} = \infty. \tag{34}$$

Considérons le système

$$f_1(t) = \sqrt{\theta}\,e^{-\varphi^2(\theta)}, \qquad t = p + \varphi(\theta)\sqrt{\frac{2pq}{\theta}}. \tag{35}$$

Il existe une valeur de θ, soit θ_0, telle que pour $\theta > \theta_0$, les seconds membres soient des fonctions respectivement croissante et décroissante de θ, et que

$$p + \varphi(\theta_0)\sqrt{\frac{2pq}{\theta_0}} \leqq 1.$$

Posons

$$t_0 = p + \varphi(\theta_0)\sqrt{\frac{2pq}{\theta_0}}.$$

Le système (35) définit, par élimination de θ, la fonction $f_1(t)$, positive, décroissante, définie dans l'intervalle $(p < t \leqq t_0)$. Nous avons, identiquement,

$$\frac{1}{\sqrt{\theta}} f_1(\alpha)\, e^{\varphi^2(\theta)} = 1.$$

Évaluons maintenant l'intégrale

$$\int_p^{t_0} f_1(t)\, dt,$$

pour savoir si elle converge, bien que $\lim\limits_{t=p} f(t) = \infty$.

Considérons un intervalle $(p+\varepsilon,\ t_0)$ pour t, et soient ω et θ_0 les valeurs de θ associées par (35) aux extrémités de cet intervalle. Prenons θ pour nouvelle variable d'intégration, nous aurons

$$\int_{p+\varepsilon}^{t_0} f_1(t)\, dt = -\sqrt{2pq}\int_{\varphi(t_0)}^{\varphi(\omega)} e^{-\varphi^2}\, d\varphi + \sqrt{\frac{pq}{2}}\int_{t_0}^{\omega} \varphi\, e^{-\varphi^2}\frac{d\theta}{\theta},$$

ce qui montre que l'intégrale de $f_1(t)$ est bien convergente dans le cas présent. $f_1(t)$ n'est défini que dans l'intervalle $(p,\ t_0)$; pour $t_0 \leqq t \leqq 1$, nous poserons par exemple

$$f_1(t) = \frac{1-t}{1-t_0} f_1(t_0).$$

Puisque $\int_p^1 f_1(t)\, dt$ a une valeur finie, on peut donc trouver une fonction $f(t)$ décroissante, positive, telle que

$$\int_p^1 f(t)\, dt = 1, \qquad \text{et} \qquad \lim_{t=p}\frac{f(t)}{f_1(t)} = \infty \quad (^1).$$

---

(1) En procédant à un changement de variable linéaire, cela revient à dire :

Cette fonction satisfait à (31); la proposition que nous avons en vue est démontrée.

DEUXIÈME SECTION.

JEU CONTINU.

Au lieu d'une suite *discrète* $\{X_i\}$ de variables aléatoires, considérons une suite *continue* $\{X_t\}$, où t est un nombre réel que nous supposons $\geqq 0$. Il revient au même de considérer l'espace E_0 des fonctions réelles d'une variable t; nous dirons, au lieu de la suite $\{X_t\}$ la fonction $X(t) = X$, ou le point X de E_0. Une fonction de répartition sera alors une fonction $\geqq 0$, additive au sens complet, définie sur certains sous-ensembles de E_0 (Kolmogoroff, [1]). Soit $P(L)$ cette fonction, nous supposons naturellement $P(E_0) = 1$.

1. Jeu continu équitable. — Généralisons les définitions de la page 83. Au lieu de la suite $\{s_i\}$ de fonctions $s_n(X_1, X_2, \ldots, X_n)$, nous aurons une suite $\{S_\tau\}$ continue $(\tau \geqq 0)$ de fonctionnelles $S_\tau(X)$, c'est-à-dire de fonctions du point dans E_0 telles que

(*a*) $$S_0(X) = 1;$$

(*b*) $$S_\tau(X) \geqq 0;$$

(*c*) $S_\tau(X)$ *ne dépend que des valeurs prises par* $X(t)$ *pour* $t \leqq \tau$;

(*d*) *La valeur moyenne de* $S_{\tau_1+\tau_2}(X)$ $(\tau_2 > 0)$, *quand on sait que* $X(t) \equiv X_0(t)$ *pour* $t \leqq \tau_1$ *est égale à* $S_{\tau_1}(X_0)$.

étant donnée une fonction $f(t)$ croissante positive telle que

$$\int_0^1 f(t)\,dt = 1, \qquad \lim_{t=\infty} f(t) = \infty,$$

on peut trouver une fonction $F(t)$, satisfaisant aux mêmes conditions, et telle que $\lim\limits_{t=\infty} \dfrac{F(t)}{f(t)} = \infty$. Il suffit pour cela de prendre par exemple

$$F(t) = f(t) f\left[\int_0^t f(u)\,du\right].$$

Les conditions (c) et (d) ont besoin d'être précisées dans le cas continu. La condition (c) peut se mettre sous la forme

(c') *Soit* $E^{(\tau)}$ [*l'ensemble des points* $X^{(\tau)}$ *de* E_0 *correspondant à des fonctions nulles sur l'intervalle fermé* $(0, \tau)$. *Alors, quels que soient* τ, X, $X^{(\tau)}$ *nous devons avoir*

$$S_\tau(X) = S_\tau(X + X^{(\tau)}).$$

La condition (d) fait intervenir la notion de valeur moyenne conditionnelle, que nous rattacherons à celle de ***probabilité conditionnelle*** telle qu'elle a été définie par Doob [Doob, [1], p. 123]; cette dernière notion est la généralisation pour un espace fonctionnel, de la notion de probabilité conditionnelle (due à M. P. Lévy), que nous avons déjà utilisée (p. 71).

Soit M_τ un ensemble dans E_0, tel que sa fonction caractéristique soit une fonction de X ne tenant compte que des valeurs de $X(t)$ pour $t \leqq \tau$; en d'autres termes, si X est un point de M_τ, il en est de même de tous les points $X + X^{(\tau)}$. Considérons la probabilité pour X d'appartenir à M_τ et à un ensemble Λ, c'est-à-dire la quantité

$$P(\Lambda . M_\tau). \tag{1}$$

La probabilité conditionnelle d'appartenir à Λ si X appartient à M_τ sera, si $P(M_\tau) > 0$,

$$P(\Lambda / M_\tau) = P(\Lambda . M_\tau) : P(M_\tau). \tag{2}$$

La probabilité conditionnelle d'appartenir à Λ si $X(t) \equiv X_0(t)$ pour $t \leqq \tau$ sera, *par définition*, une fonction $P_{\tau,\Lambda}(X_0)$ de l'ensemble Λ et du point X_0 satisfaisant, quel que soit M_τ, à

$$P(\Lambda . M_\tau) = \int_{M_\tau} P_{\tau,\Lambda}(X)\, dP, \tag{3}$$

la variable d'intégration étant X. Nous utilisons la notion d'intégrale au sens de M. Fréchet (Fréchet, [1]). L'existence de $P_{\tau,\Lambda}(X_0)$ résulte d'une démonstration de M. Nikodym (Nikodym, [1], p. 168-179).

Étant donnée une fonction de point $f(X)$, *sa valeur moyenne conditionnelle, sachant que* $X(t) = X_0(t)$ *pour* $t \leqq \tau$ *sera par* *définition*

$$(4) \qquad \mathfrak{M}_{\tau, X_0}\{f(X)\} = \int_{E_0} f(X)\, d P_{\tau, \Lambda}(X_0),$$

la variable d'intégration étant X, considérée comme point de Λ dans $dP_{\tau,\Lambda}(X_0)$. La condition (d) devient alors

$$(5) \qquad S_{\tau_1}'(X_0) = \mathfrak{M}_{\tau_1}\{ S_{\tau_1+\tau_2}(X) \}.$$

$S_\tau(X)$ représente alors la somme que le joueur A, qui a misé une somme égale à l'unité, recevra s'il arrête le jeu à « l'instant » τ et si le hasard a fourni la fonction $X(t)$.

2. Formule de la ruine des joueurs. — On pourrait se demander s'il est légitime de généraliser l'inégalité (15) du paragraphe **2** de la première section de ce chapitre par l'inégalité

$$(6) \qquad \Pr.\left\{ \underset{0<\tau<\infty}{\text{B. sup.}}\, S_\tau(X) \geqq \lambda \right\} \leqq \frac{1}{\lambda},$$

que nous appellerons formule de la ruine des joueurs dans le cas du jeu continu. *Cette formule n'est pas toujours vraie*, même si les conditions (a)-(d) sont vérifiées.

Considérons une fonction aléatoire $X(t)$ à accroissements indépendants. La variable aléatoire $X(t+\theta) - X(t)\ (\theta > 0)$ admet alors une fonction de répartition indépendante des valeurs prises par $X(t)$ pour les abscisses $\leqq t$. Soit dans E_0 la fonction additive P qui définit la loi de probabilité de $X(t)$. Considérons les abscisses t_0 et $t_0+\theta$, une ordonnée y, et soit Λ_y l'ensemble défini dans E_0 par l'inégalité :

$$(1) \qquad X(t_0+\theta) - X(t_0) < y.$$

Soit $F(y, t_0, \theta)$ la fonction de répartition de $X(t+\theta) - X(t)$; soit M_{t_0} un sous-ensemble de E_0 tel que la fonction caractéristique de M_{t_0} ne dépende que des valeurs de $X(t)$ pour $t \leqq t_0$. Si $X(t_0+\theta) - X(t_0)$ ne dépend pas des valeurs de $X(t)$ pour $t \leqq t_0$, nous devons avoir

$$(2) \qquad P(M_{t_0}.\Lambda_y) = P(M_{t_0})\, F(y, t_0, \theta).$$

Réciproquement, si entre $P(M_{t_0}, \Lambda_y)$ et $P(M_{t_0})$ quels que soient t_0, M_{t_0}, θ, y existe une relation de la forme (2), X sera dite à accroissements successifs indépendants, et (2) nous donnera la loi de répartition de $X(t_0+\theta) - X(t_0)$.

Considérons maintenant la fonction aléatoire $X_1(t)$ définie comme suit : soit dans E_0 l'ensemble E_1 des points représentant des courbes $X_1(t)$ telles que

$X_1(t)$ ne peut prendre que les valeurs 0 ou 1 ;

$X_1(t) = 1$ a une racine et une seule, comprise entre 0 et 1.

Associons à tout point X_1 de E_1 l'abscisse de la racine de $X_1(t) = 1$, soit $\theta(X_1)$ la fonction de X_1 ainsi définie. Définissons la fonction additive P dans E_0 par

$$P(E_1) = 1,$$
$$P(E_u) = u \qquad (0 \leqq u \leqq 1),$$

où E_u est le sous-ensemble de E formé par les points X_1 tels que $\theta(X_1) < u$. Considérons la fonction $F(y, t, \theta)$ telle que

$$F(y, t, \theta) = 1, \qquad \text{pour} \quad y > 0;$$
$$F(y, t, \theta) = 0, \qquad \text{pour} \quad y \leqq 0.$$

L'équation (2) est vérifiée, quels que soient M_{t_0} et y. Nous devons donc considérer $X(t)$ comme une fonction aléatoire à accroissements aléatoires indépendants, telle que

$$\text{(3)} \qquad \text{Pr.}\left\{X_1(t) = X_1(t+\theta)\right\} = 1.$$

Soit maintenant la fonction $s(t, x)$, de la forme

$$\text{(4)} \qquad s(t, x) = x + 1.$$

Sachant que $X(t) = x$, nous avons, pour valeur moyenne conditionnelle de $s[t+\theta, X(t+\theta)]$, d'après (3), la valeur $s(t, x)$. Donc les conditions (*a*), (*b*), (*c*), (*d*), (p. 95) sont vérifiées. Nous devrions avoir, d'après (6) (p. 97) ;

$$\text{Pr.}\left\{\text{B. sup. } s[t, X_1(t)] \leqq 2\right\} \leqq \frac{1}{2};$$

or, nous avons manifestement

$$\text{Pr.}\left\{\text{B. sup. } s[t, X_1(t)] = 2\right\} = 1.$$

La formule (6) *est mise en défaut.*

3. **Remarque sur les probabilités conditionnelles.** — L'exemple traité dans le paragraphe précédent est extrêmement simple, et permet de mettre en évidence une singularité intéressante. En définissant l'indépendance des accroissements successifs par (2), nous sommes arrivés à la conclusion qu'ils étaient indépendants dans l'exemple choisi. Or, on est tenté de dire : deux accroissements successifs de $X_1(0)$ ne peuvent pas être considérés comme indépendants, puisque, si nous savons que $X_1(t)=1$, nous sommes *sûr* que $X_1(t+\theta)=0$, ce qui semble contredire l'affirmation : « si $X_1(t)$ est connu, la probabilité conditionnelle que $X_1(t+\theta)=X_1(t)$ est égale à 1 ».

Cette objection est intéressante, mais n'est pas fondée si l'on se place dans l'axiomatique de Kolmogoroff. En effet, la seule définition admise jusqu'ici, à notre connaissance du moins, pour la probabilité conditionnelle, est une définition indirecte : la probabilité que l'événement E se produise si la variable aléatoire Y a la valeur y est une fonction $F(y)$ telle que l'on ait, si

$$G(y)=\Pr.\{Y<y\},$$
$$\Pr.\{E \text{ et } Y<y\}=\int^{y-0} F(y)\,d\,G(y).$$

Si l'on connaît

$$\Phi(y)=\Pr.\{E \text{ et } Y<y\},$$

et $G(y)$, F est définie par une équation intégrale

$$\Phi(y)=\int^{y-0} F(y)\,d\,G(y).$$

On peut donc modifier arbitrairement $F(y)$ sur un ensemble de valeurs de y sur lequel $\int dG(y)=0$. En particulier, soit y_0 une valeur de continuité de $G(y)$; considérons l'événement E_0 ainsi défini

$$\begin{cases} \text{si } Y\neq y_0, & E_0 \text{ est identique à } E; \\ \text{si } Y=y_0, & E_0 \text{ se produit certainement}; \end{cases}$$

et la fonction $F_0(y)$ égale à $F(y)$ pour $y\neq y_0$, à 0 pour $y=y_0$.

Nous avons alors

$$\int^{y-0} F_0(y)\, d\,G(y) = \int^{y-0} F(y)\, d\,G(y),$$

et de même

$$\text{Pr.}\left\{ E \text{ et } Y < y \right\} = \text{Pr.}\left\{ E_0 \text{ et } Y < y \right\}.$$

Donc $F_0(y)$ peut être considéré comme probabilité conditionnelle de E_0 quand $Y = y$, bien que pour $Y = y_0$, $F_0(y_0) = 0$, et que $\{ Y = y_0 \}$ implique logiquement E_0.

En revenant à $X_1(t)$, nous voyons donc qu'en définissant l'indépendance par (2), nous pouvons considérer cette fonction comme étant à accroissements indépendants presque certainement nuls, en considérant deux variables comme indépendantes lorsque la fonction de répartition conditionnelle de l'une d'entre elles, au sens indiqué plus haut (p. 80), est indépendante de la valeur prise par l'autre; c'est-à-dire, avec les notations de la page 71 : X_1 et X_2 sont indépendantes lorsque $\Phi_2(X_1\,;\, x_2)$ ne dépend pas de X_1, ou encore lorsqu'on peut poser

$$\Phi_2(X_1\,;\, x_2) = \varphi(x_2).$$

M. Paul Lévy est d'avis de distinguer entre l'indépendance ci-dessus, qu'il propose d'appeler *indépendance presque certaine,* et l'*indépendance proprement dite*, définie directement (P. Lévy, [1], p. 3).

Notre discussion se place uniquement sur le terrain de M. Kolmogoroff; la conclusion peut être différente en suivant le point de vue de M. Paul Lévy.

Les remarques précédentes montrent qu'il est impossible, étant donnée une fonction aléatoire à accroissements indépendants, de démontrer qu'elle est presque certainement continue, si l'on ne se donne rien d'autre que la loi de répartition des accroissements. Considérons en effet $X_1(t)$, la fonction aléatoire définie dans le paragraphe précédent, et soit $X(t)$ une fonction aléatoire quelconque à accroissements indépendants. Supposons $X_1(t)$ indépendante de $X(t)$: la fonction aléatoire $X(t) + X_1(t)$ sera à accroissements indépendants, et ses accroissements suivront la même loi que ceux de $X(t)$; mais $X(t) + X_1(t)$ est presque certainement discontinue puisque $X_1(t)$ admet un point de discontinuité.

Le seul résultat que l'on puisse, sans hypothese supplémentaire, établir sur la continuité de $X(t)$ est qu'elle est *équivalente stochastiquement* (p. 92) à une fonction continue. C'est ce qu'établit en particulier M. Paul Lévy pour les courbes aléatoires à accroissements indépendants suivant la loi de Gauss (P. Lévy, [1], p. 172).

Nous allons abandonner ces questions de principe, et appliquer la formule de la ruine des joueurs à des cas plus concrets.

4. Cas de Chapmann-Kolmogoroff. [$S_\tau(X)$ **ne dépendant que de** $X(\tau)$]. — Nous allons généraliser l'étude du cas de Bernoulli (p. 87), en nous plaçant dans les hypothèses suivantes :

I. $S_\tau(X)$ ne dépend que de $X(\tau)$. En d'autres termes, il existe une fonction $s(t, x)$ telle que

$$S_\tau(X) = s[\tau, X(\tau)]. \tag{7}$$

II. $X(t_2) - X(t_1)$, $(t_2 > t_1)$ ne dépend pas des valeurs de $X(t)$ pour $t < t_1$. De sorte que nous avons une relation de la forme

$$\text{Pr.}\left\{X(t_2) < x_2/X(t_1) = x_1\right\} = F(x_1, x_2, t_1, t_2) \qquad (t_1 < t_2). \tag{8}$$

La fonction F satisfait alors à l'équation de Chapmann-Kolmogoroff (Hostinsky, [1], p. 59).

$$\left\{\begin{aligned} &F(x_1, x_3, t_1, t_3) = \int_{x_2=-\infty}^{x_2=+\infty} F(x_2, x_3, t_2, t_3)\, d\,F(x_1, x_2, t_1, t_2) \\ &\qquad\qquad (t_1 < t_2 < t_3) \end{aligned}\right. \tag{9}$$

et les conditions (a), (b), (c), (d), imposées à $S_\tau(X)$, se traduisent pour $s(t, x)$ par les conditions

$$(a') \qquad s(0, x) = 1;$$

$$(b') \qquad s(t, x) \geqq 0 \qquad (t \geqq 0);$$

$$(c') \qquad \text{(vérifiée d'elle-même)};$$

$$(d') \qquad s(t_1, x_0) = \int s(t_2, x)\, d_x F(x_0, x, t_1, t_2) \qquad (t_1 < t_2),$$

et enfin la formule de la ruine des joueurs devient

$$\text{Pr.}\left\{\text{B. sup.}\, s[t, X(t)] \geqq \lambda\right\} \leqq \frac{1}{\lambda}. \tag{10}$$

5. **Borne supérieure stochastique.** — Même dans le cas du paragraphe précédent, l'inégalité (10), nous l'avons vu, peut être mise en défaut (p. 98). Nous allons la remplacer par une inégalité moins stricte, et faire une hypothèse supplémentaire. X étant déterminée, la fonction de t

$$Y(t) = s[t, X(t)] \tag{11}$$

peut être considérée comme obtenue à partir de $X(t)$ par une transformation $Y(t) = T[X(t)]$. Nous supposerons que la fonction aléatoire $Y(t) = T[X(t)]$ est une fonction stochastiquement continue (Slutsky, [1]), c'est-à-dire que, quels que soient t, $\varepsilon > 0$, $\eta > 0$, il existe $\tau > 0$ tel que pour $|t' - t| < \tau$, on ait

$$\text{Pr.}\left\{ |Y(t') - Y(t)| < \varepsilon \right\} > 1 - \eta. \tag{12}$$

Dans ces conditions, soit $\{t_i\}$ une suite dénombrable partout dense sur l'axe des t; la variable aléatoire

$$Z = \text{B. sup. } Y(t_i)$$

est une variable aléatoire dont la fonction de répartition, si elle existe, est indépendante de la suite dense que l'on a choisie. En effet, soit $\overline{Y}(t)$ la fonction de t :

$$\overline{Y}(t) = \overline{\lim_{t_i \to t}}\, Y(t_i).$$

Slutsky a montré que $P\{Y(t) = \overline{Y}(t)\} = 1$ quel que soit t. Si donc nous considérons une autre suite $\{t'_i\}$ et la variable Z' correspondante, nous aurons

$$\text{Pr.}\left\{ Z' = \text{B. sup. } \overline{Y}(t) \right\} = 1,$$

et par conséquent

$$\text{Pr.}\left\{ Z' \geqq Z \right\} = 1.$$

On démontrerait de même que $P\{Z \geqq Z'\} = 1$, d'où la propriété annoncée. Nous appellerons cette variable Z la borne supérieure stochastique de $Y(t)$ et écrirons

$$Z = \text{B. sup.}_{st}\, Y(t).$$

Nous allons montrer que

$$\text{Pr.}\left\{\text{B. sup.}_{st}\, \text{Y}(t) \geqq \lambda\right\} \leqq \frac{1}{\lambda},$$

ce qui équivaut, une suite dense $\{t_i\}$ étant choisie, à

$$\lim_{n=\infty} \text{Pr.}\left\{\underset{i \leqq n}{\text{B. sup.}}\, s[t_i, \text{X}(t_i)] \geqq \lambda\right\} \geqq \frac{1}{\lambda}. \tag{13}$$

Rangeons par ordre de grandeur croissante les valeurs $t_1, t_2, \ldots, t_n$; soit $\theta_1, \theta_2, \ldots, \theta_n$ la suite obtenue. Posons

$$\text{X}(\theta_i) = \text{X}_i, \qquad s_i(\text{X}_i) = s[\theta_i, \text{X}(\theta_i)].$$

Les fonctions

$$s_1(\text{X}_1), \quad s_2(\text{X}_2), \quad \ldots, \quad s_n(\text{X}_n)$$

satisfont aux conditions (p. 83) du jeu discontinu, relativement aux variables aléatoires $\text{X}_1, \text{X}_2, \ldots, \text{X}_n$. Nous avons donc

$$\text{Pr.}\left\{\text{B. sup.}\, s_i \geqq \lambda\right\} \leqq \frac{1}{\lambda},$$

d'où nous déduisons immédiatement l'inégalité (13). *La formule de la ruine des joueurs devient donc vraie dans le cas de Chapmann-Kolmogoroff, en introduisant la notion de borne supérieure stochastique* et l'hypothèse de Slutsky, et s'écrit

$$\text{Pr.}\left\{\text{B. sup.}_{st}\, s[t, \text{X}(t)] \geqq \lambda\right\} \leqq \frac{1}{\lambda}. \tag{14}$$

6. Recherche de fonctions limites. — Poursuivant toujours l'analogie avec le problème traité dans le cas de Bernoulli, soit $\varphi(t)$ une fonction non aléatoire. Supposons que nous ayons trouvé une fonction $s(t, x)$ satisfaisant aux conditions (a'), (b'), (c'), (d'), et telle que, en outre

$$\left\{\begin{array}{ll} s(t, x') \geqq s(t, x''), & \text{pour } \; x' > x'', \\ \lim\limits_{t=\infty} s[t, \varphi(t)] = \infty. & \end{array}\right. \tag{15}$$

Nous ne supposons pas que $\text{Y}(t) = s[t, \text{X}(t)]$ est stochastiquement continue. Soit maintenant $\{t_i\}$ une suite dénombrable partout dense; soit t_0 une valeur de t, désignons par λ_0 la borne inférieure de $s[t_i, \varphi(t_i)]$ pour $t_i \geqq t_0$. Soient $\theta_1, \theta_2, \ldots, \theta_n$ les n premiers termes de $\{t_i\}$ qui sont $\geqq t_0$; supposons-les rangés par ordre de grandeur croissante. Si nous posons

$$X_1 = X(\theta_1), \qquad X_2 = X(\theta_2), \qquad \ldots, \qquad X_n = X(\theta_n);$$
$$s_0 = 1; \qquad s_1(x_1) = s(\theta_1, x_1), \ \ldots, \ s_n(x_n) = s(\theta_n, x_n),$$

nous vérifierons facilement que cette suite $s_0, s_1, \ldots, s_n$ définit un jeu équitable discontinu (p. 83). Donc, nous avons

$$\Pr.\left\{ \underset{i=1,2,\ldots,n}{\text{B. sup.}}\ s[\theta_i, X(\theta_i)] \geqq \lambda_0 \right\} \leqq \frac{1}{\lambda_0}$$

et, en faisant tendre n vers l'infini,

$$\Pr.\left\{ \underset{t_i \geqq t_0}{\text{B. sup.}}\ s[t_i, X(t_i)] \geqq \lambda_0 \right\} \leqq \frac{1}{\lambda_0}. \tag{16}$$

Reportons-nous aux conditions (15). Nous déduirons de (16) que

$$\Pr.\left\{ X(t_i) \leqq \varphi(t_i) \quad \text{pour} \quad t_i \geqq t_0 \right\} \geqq 1 - \frac{1}{\lambda_0},$$

ou que, si nous désignons par T la borne supérieure des valeurs de t_i pour lesquelles $X(t_i) > \varphi(t_i)$, nous avons

$$\Pr.\left\{ T \leqq t_0 \right\} \geqq 1 - \frac{1}{\lambda_0},$$
$$\Pr.\left\{ T < \infty \right\} \geqq 1.$$

T est presque certainement fini. $\varphi(t)$ *sera alors appelée une fonction limite supérieure stochastique de* $X(t)$.

7. Cas où $X(t+\theta) - X(t)$ est une variable obéissant à la loi de Gauss. — Quand la fonction $F(x_1, x_2, t_1, t_2)$ est de la forme $F(x_2 - x_1, t_1, t_2)$, la fonction $X(t)$ est dite une intégrale à éléments aléatoires indépendants ou fonction aléatoire à accroissements indépendants (P. Lévy). Si la fonction $X(t)$ aléatoire est de plus continue (avec probabilité égale à 1) $F(x_2 - x_1, t_1, t_2)$ est une fonction de repartition suivant la loi de Gauss, que l'on peut mettre, en faisant au besoin un changement de variables pour t, sous la forme $F(x_2 - x_1, t_2 - t_1)$, avec

$$F(u, v) = \frac{1}{\sqrt{2\pi v}} \int^{u} e^{-\frac{u^2}{2v}}\, du. \tag{18}$$

Réciproquement, si $F(x_1, x_2, t_1, t_2)$ a la forme ci-dessus indiquée, la courbe aléatoire $X(t)$ est équivalente stochastiquement à une fonction continue, c'est-à-dire qu'il existe une fonction aléatoire $Y(t)$, déterminée par la même épreuve qui détermine $X(t)$, continue, telle que, quel que soit t,

$$\Pr\{X(t) = Y(t)\} = 1.$$

Les résultats que nous venons de rappeler sont dus à M. Paul Lévy (P. Lévy, [1], p. 158-173) et montrent l'importance de la place tenue, parmi les intégrales à éléments aléatoires indépendants, par celles où $F(x_2 - x_1, t_1, t_2)$ est la fonction de répartition d'une loi de Gauss.

Cherchons à déterminer une fonction limite stochastique (p. 104) par le procédé du paragraphe 5 quand F a la forme (18). La fonction $s(t, x)$ doit satisfaire à l'équation

$$s(t, x) = \frac{1}{\sqrt{2\pi\theta}} \int s(t+\theta, x+y)\, e^{-\frac{y^2}{2\theta}}\, dy, \tag{19}$$

qui exprime la condition (d_1') (p. 101).

Nous allons essayer de trouver une solution de cette équation en raisonnant par analogie avec ce qui a été obtenu pour le cas de Bernoulli discontinu. Nous rappelons que nous avions trouvé

$$s(x, y) = p^{-x} q^{-y} \int_0^1 t^x (1-t)^y\, dF(t);$$

x étant la répétition de l'événement de probabilité p. $s(x, y)$ est le quotient de deux quantités, que nous pouvons écrire

$$\frac{x+y!}{x!\,y!} \int_0^1 t^x (1-t)^y\, dF(t), \qquad \frac{x+y!}{x!\,y!} p^x q^y,$$

en mettant en évidence que la seconde est la probabilité d'obtenir les répétitions x et y, tandis que la première est la valeur moyenne d'une probabilité du même genre. Ceci nous conduit à essayer, pour l'équation (19), une solution de la forme

$$(20) \qquad s(t, x) = \frac{1}{\sqrt{2\pi t}} \int e^{-\frac{(x-\lambda t)^2}{2t}} dF(\lambda) : \frac{1}{\sqrt{2\pi t}} e^{-\frac{x^2}{2t}}$$
$$= \int e^{\lambda x - \lambda^2 \frac{t}{2}} dF(\lambda)$$

[F(λ) est une fonction de répartition]. Portons l'expression (20) dans le second membre de (19), nous obtenons

$$\frac{1}{\sqrt{2\pi\theta}} \int dy \int e^{\lambda(x+y) - \lambda^2 \frac{t+\theta}{2} - \frac{y^2}{2\theta}} dF(\lambda) \qquad (\theta > 0).$$

La fonction sous le signe $\int$ étant non négative, nous pouvons intervertir l'ordre des intégrations, ce qui montre que (19) est vérifiée. Il reste à satisfaire à (15), et en particulier, en remplaçant, pour une raison de commodité, $\varphi(t)$ par $\varphi(t)\sqrt{2t}$, à

$$(21) \qquad \lim_{t=\infty} s[t, \varphi(t)\sqrt{2t}] = \infty.$$

Prenons pour $F(\lambda)$ une fonction dérivable, soit $F'(\lambda) = f(\lambda)$, et choisissons f nulle pour $\lambda \leqq 0$, décroissante pour $\lambda > 0$. Nous aurons alors

$$s(t, x) = \int_0^\infty e^{\lambda x - \lambda^2 \frac{t}{2}} f(\lambda)\, d\lambda, \qquad \int_0^\infty f(\lambda)\, d\lambda = 1,$$

et la première des conditions (15) est vérifiée. Passons à (21). Nous avons, pour $x > 0$,

$$s(t, x) \geqq \int_0^{\frac{x}{t}} e^{\lambda x - \lambda^2 \frac{t}{2}} f(\lambda)\, d\lambda \geqq f\left(\frac{x}{t}\right) \int_0^{\frac{x}{t}} e^{\lambda x - \lambda^2 \frac{t}{2}} dt$$
$$= \sqrt{\frac{2}{t}} \cdot e^{\frac{x^2}{2t}} f\left(\frac{x}{t}\right) \int_0^{\frac{x}{\sqrt{2t}}} e^{-u^2} du.$$

Or, il est bien connu qu'on peut trouver une constante k telle que pour $\alpha > 0$

$$\int_\alpha^\infty e^{-n^2} du < k\, e^{-\alpha^2},$$

d'où pour $s(t, x)$ l'inégalité

$$s(t, x) > \sqrt{\frac{\pi}{t}} f\left(\frac{x}{t}\right) e^{\frac{x^2}{2t}} - k\sqrt{\frac{2}{t}} f\left(\frac{x}{t}\right).$$

Le raisonnement se poursuit à partir de cet instant comme dans le cas de Bernoulli (p. 90). Si la fonction croissante $\varphi(t)$ rend convergente l'intégrale $\int \varphi e^{-\varphi^2} \frac{dt}{t}$ et si $\sqrt{t}e^{-\varphi^2(t)}$, est pour t suffisamment grand, une fonction croissante, on peut trouver $f(\lambda)$ telle que

$$\int f(\lambda) d\lambda = 1$$

et

$$\lim_{t=\infty} s\left[t, \varphi(t)\sqrt{2t}\right] = \infty.$$

Donc : *toute fonction* $\varphi(t)$ *croissante, rendant convergente l'intégrale* $\int \varphi e^{-\varphi^2} \frac{dt}{t}$ *et telle que* $\sqrt{t}e^{-\varphi^2}$ *soit croissante pour* t *suffisamment grand, est une fonction limite stochastique* (p. 104) *pour la courbe aléatoire* $\frac{X(t)}{\sqrt{2t}}$, *où* $X(t)$ *est une courbe aléatoire dont la loi de probabilité est définie par* (18).

Nous démontrons ainsi en ne faisant appel qu'à des considérations élémentaires de probabilité un théorème dû à Petrowsky; la démonstration de ce dernier utilise la théorie de l'équation de la chaleur (Petrowsky, [1],

8. Plus grande valeur. — Dans certaines études on s'intéresse particulièrement aux plus grandes valeurs prises par une suite de variables aléatoires. Nous ne traiterons qu'un cas schématique, et chercherons à trouver une fonction limite inférieure de la plus grande; ce sera un résultat nouveau.

Soient

$$X_1, \quad X_2, \quad \ldots, \quad X_n, \quad \ldots$$

des variables aléatoires indépendantes, ayant toutes même fonction de répartition. Considérons la suite des maxima de la suite $\{X_i\}$. C'est une suite de variables enchaînées

$$Y_1, \quad Y_2, \quad \ldots, \quad Y_n, \quad \ldots,$$

où

$$Y_i = \max\{X_1, X_2, \ldots, X_i\}.$$

Soit $F_i(y)$ la fonction

$$F_i(y) = \Pr.\{Y_i < y\},$$

connaissant $Y_1, Y_2, \ldots, Y_{i-1}$. Si nous supposons que

$$\Pr.\{X_i < x\} = x \qquad (0 \leqq x \leqq 1),$$

nous aurons

$$(1) \qquad F_i(y) = \begin{cases} y, & \text{pour } y > Y_{i-1}; \\ 0, & \text{pour } y \leqq Y_{i-1}. \end{cases}$$

Remplaçons la suite discrète $\{Y_i\}$ par une suite continue, c'est-à-dire par une fonction aléatoire $Y(t)$, telle que pour t entier, on ait

$$\Pr.\{Y(n) = Y_n\} = 1.$$

Remarquons que $F_i(y)$ ne dépend que de Y_{i-1}; si nous calculons la fonction de répartition de Y_i connaissant Y_{i-k} seul, nous obtenons une fonction nulle pour $y \leqq Y_{i-k}$ et égale à y^k pour $y > Y_{i-k}$. Nous sommes donc amené, pour $Y(t)$, à faire l'hypothèse, compatible avec la précédente :

$$\Pr.\{Y(t+\theta) < y / Y(t) = u\} = F(y, u, \theta) \qquad (\theta > 0),$$

avec

$$(2) \qquad F(y, u, \theta) = \begin{cases} y^\theta, & \text{pour } y > u; \\ 0, & \text{pour } y \leqq u. \end{cases}$$

Avec les notations de la page 88, nous avons

$$F(x_1, x_2, t_1, t_2) = \begin{cases} x_2^{t_2 - t_1}, & \text{pour } x_2 > x_1; \\ 0, & \text{pour } x_2 \leqq x_1. \end{cases}$$

L'équation (9) est vérifiée, ainsi que l'on s'en assure par simple substitution. Cherchons une solution de l'équation (d') (p. 101), qui s'écrit ici

$$s(t_1, y_0) = \int s(t_2, y)\, d_y F(y, y_0, t_2 - t_1) \qquad (t_r > t_1)$$

ou, en posant $t_1 = t$, $t_2 - t_1 = \theta$, $y_0 = y$, $y = u$, et en explicitant :

$$(3) \qquad s(t, y) = y^\theta s(t+\theta, y) + \theta \int_y^1 s(t+\theta, u) u^{\theta-1}\, du.$$

$s(t, y)$ doit être $\geqq 0$; de plus, nous devons avoir $s(0, 0) = 1$, ce qui donne

$$(4) \qquad t\int_0^1 s(t, y) y^{t-1}\, dt = 1.$$

Supposons s dérivable, et différentions (3) par rapport à y; il vient

$$(5) \qquad d_y\, s(t, y) = y^{\theta}\, d_y\, s(t+\theta, y).$$

Nous en déduisons que si s satisfait à une équation aux dérivées partielles, cette dernière sera de la forme

$$\frac{\partial s}{\partial y} = y^{-t}\, \rho(y),$$

ce qui nous amène à chercher pour s une fonction de la forme

$$(6) \qquad s(t, y) = \int_y^1 \rho(u) u^{-t}\, du.$$

Nous prendrons $\rho(u) \geqq 0$ pour être sûr que $s \geqq 0$; on vérifie sans peine que la solution (6) satisfait bien à (3); quant à (4), nous en tirons

$$\int_0^1 \rho\, du = 1.$$

Cherchons maintenant une fonction $\varphi(t)$ telle que l'on ait

$$Y(t) \geqq \varphi(t),$$

à partir d'une valeur de t presque certainement finie. Remarquons que $s(t, y)$ définie par (6) est une fonction décroissante de y. En raisonnant comme au paragraphe 5 (p. 102), nous verrons que si nous pouvons déterminer $s(t, y)$ telle que la fonction $s[t, \varphi(t)]$ tende vers l'infini avec t, il sera démontré que, si nous nous donnons une suite partout dense $\{t_\iota\}$, la borne supérieure des valeurs de t_ι pour lesquelles (7) n'a pas lieu est presque certainement finie. Remarquons que nous n'avons besoin ici que de considérer les valeurs entières de t; donc le résultat précédent sera satisfaisant [1].

---

[1] Par ailleurs, si $\varphi(t)$ est croissante, comme $Y(t)$ est non décroissante, $Y(t_\iota) \geqq \varphi(t_\iota)$ pour $t_\iota > t_0$ implique $Y(t) \geqq \varphi(t)$ pour $t > t_0$.

Nous sommes donc ramenés, φ étant donné, à chercher $\rho \geqq 0$, tel que $\int \rho \, du = 1$, et que $\lim\limits_{t=\infty} s[t, \varphi(t)] = \infty$, s étant définie par (6).

Pour trouver à quelle condition doit satisfaire φ pour que ρ puisse être ainsi définie, nous allons raisonner d'abord d'une manière peu rigoureuse; nous démontrerons ensuite rigoureusement la proposition à laquelle nous serons amenés. Si t est grand, u^{-t} décroît très rapidement; nous poserons

$$(8) \qquad s(t, y) \sim \rho(y) \int_y^1 u^{-t} \, dt = \frac{\rho(y)}{t-1}\left[\frac{1}{y^{t-1}} - 1\right].$$

Faisons $y = \varphi(t)$ dans le second membre, nous obtenons une fonction $f(t)$

$$(9) \qquad f(t) = \frac{\rho(\varphi)}{t-1}\left[\frac{1}{\varphi^{t-1}} - 1\right] \sim s[t, \varphi(t)].$$

Supposons φ continue croissante, variant de 0 à 1 quand t varie de 0 à ∞, et soit $t = t(\varphi)$ la fonction inverse de $\varphi(t)$. Posons

$$f[t(\varphi)] = \psi(\varphi), \qquad \text{d'où} \quad f(t) = \psi[\varphi(t)].$$

Il faut que

$$\lim_{t=\infty} \psi[\varphi(t)] = \infty,$$

ou

$$\lim_{\varphi=1} \psi(\varphi) = \infty.$$

Donnons-nous alors une fonction croissante $\psi(\varphi)$, telle que $\psi(\varphi) \to \infty$ quand $\varphi \to 1$, et déterminons ρ. Nous tirons de (9)

$$(10) \qquad \rho(\varphi) = \frac{(t-1) f(t) \varphi^{t-1}}{1 - \varphi^{t-1}}.$$

Si nous posons $t = t(\varphi)$ dans le second membre, $\rho(\varphi)$ est déterminé, et nous avons

$$(11) \qquad \rho(\varphi) = \frac{[t(\varphi) - 1]\, \psi(\varphi) \varphi^{t(\varphi)-1}}{1 - \varphi^{t(\varphi)-1}}.$$

Il faut que $\int_0^1 \rho \, du = 1$. φ variant de 0 à 1, intégrons le second membre de (11), par rapport à φ. Négligeons le facteur ψ, qui croît lentement quand $\varphi \to 1$. Il vient

$$\rho\, d\varphi \sim \frac{(t-1)\varphi^{t-1}}{1-\varphi^{t-1}}\, d\varphi = \rho_1\, d\varphi \qquad [t = t(\varphi)].$$

Supposons que le coefficient de $d\varphi$ soit une fonction croissante de φ, et intégrons par parties entre o et 1, en posant

$$u = \operatorname{Log} \frac{1}{1-\varphi^{1-t}}.$$

Il vient

$$du = \frac{(t-1)\varphi^{t-2}\, d\varphi + \varphi^{t-1} \operatorname{Log}\varphi\, dt}{1-\varphi^{t-1}},$$

d'où, en supposant φ^t monotone,

$$\begin{aligned}\int_0^1 \rho_1\, d\varphi &= \int_0^\infty \rho \frac{du}{dt}\, dt - \int_{t=0}^\infty \frac{\varphi^t \operatorname{Log}\varphi}{1-\varphi^{t-1}}\, dt \\ &= [\varphi u]_{t=0}^{t=\infty} - \int_{t=0}^\infty u\, d\varphi + \int_0^\infty \frac{\varphi^t}{1-\varphi^{t-1}} \operatorname{Log}\frac{1}{\varphi}\, dt.\end{aligned}$$

$\int \rho_1\, d\varphi$ sera sûrement convergente si $\lim\limits_{t=\infty} \varphi^t < 1$ et si l'intégrale $\int \varphi^t \log \frac{1}{\varphi}\, dt$ converge. Remarquons que cette dernière condition entraîne la première, puisqu'elle implique que $\lim\limits_{t=\infty} \varphi^t = 0$, si φ^t est une fonction monotone à partir d'une certaine valeur de t.

Nous obtenons ainsi, par un raisonnement intuitif, mais peu rigoureux, la proposition suivante, que nous démontrerons plus loin :

Soit $\varphi(t)$ une fonction croissante continue, ainsi que la fonction $t\varphi^t$, telle que φ^t soit monotone et que $\int \varphi^t \log \frac{1}{\varphi}\, dt$ soit convergente. Soit $Y(t)$ la fonction aléatoire définie par (2) (p. 108); l'inégalité $Y \geqq \varphi(t)$ aura lieu constamment à partir d'une valeur de t presque certainement finie.

Démonstration. — Supposons que $\lim\limits_{t=\infty} \varphi(t) = 1$, et $\varphi(0) = 0$, ce qui ne restreint pas la généralité de l'hypothèse; soit $t(\varphi)$ la fonction inverse de la fonction $\varphi(t)$, et considérons la fonction

$$\rho_1(\varphi) = t(\varphi)\, \varphi^{t(\varphi)}.$$

C'est une fonction croissante, puisque $t\varphi^t$ est une fonction croissante de t; considérons l'intégrale

$$\int_0^1 \rho_1(\varphi)\,d\varphi$$

et faisons-y le changement de variable $\varphi = \varphi(t)$. Elle devient

$$\int_0^\infty t\varphi^t \frac{d\varphi}{dt}\,dt = \int_0^\infty \varphi^{t+1} \operatorname{Log}\frac{1}{\varphi}\,dt + \int_0^\infty \varphi\,\frac{d\varphi^t}{dt}\,dt.$$

La première intégrale est convergente puisque $\int_0^\infty \varphi^t \operatorname{Log}\frac{1}{\varphi}\,dt$ converge. La seconde converge évidemment. Donc $\int \rho_1\,d\varphi$ a une valeur finie. Nous pouvons prendre $\rho(u)$ croissante telle que

$$\int_0^1 \rho(u)\,du = 1, \qquad \lim_{u=1} \frac{\rho(u)}{\rho_1(u)} = \infty \qquad [\rho(u) \geqq 0].$$

Si nous considérons maintenant

$$s(t, y) = \int_y^1 u^{-t}\,\rho(u)\,du,$$

nous aurons, puisque $\rho(u)$ est croissante,

$$s(t, y) > \rho(y) \int_y^1 u^{-t}\,du = \frac{\rho(y)}{t-1}\left[\frac{1}{y^{t-1}} - 1\right],$$

$$s[t, \varphi(t)] > \frac{\rho(\varphi)}{\rho_1(\varphi)}\,\frac{\rho_1(\varphi)}{t-1}\left[\frac{1}{\varphi^{t-1}} - 1\right] = \frac{\rho(\varphi)}{\rho_1(\varphi)}\,\frac{t}{t-1}\,\varphi(1-\varphi^{t-1})$$

et, quand $t \to \infty$,

$$s[t, \varphi(t)] \to \infty;$$

en effet, on déduit de la convergence de $\int \varphi^t \operatorname{Log}\frac{1}{\varphi}$ que la fonction (supposée monotone) φ^t a pour limite o quand $t \to \infty$.

Nous pouvons considérer, au lieu d'une inégalité de la forme $Y > \varphi$, une inégalité de la forme

$$1 - Y < \frac{\psi(t)}{t},$$

ce qui nous donnera, en posant $\varphi = 1 - \frac{\psi}{t}$, une condition suffisante pour que $\psi(t)$ soit une fonction limite supérieure stochastique de $t(1 - Y)$. Dans $\int \varphi^t \log \frac{1}{\varphi} dt$, $\log \frac{1}{\varphi}$ est comparable à $\frac{\psi}{t}$ et φ^t à $e^{-\psi}$, ce qui donne pour ψ la condition :

L'intégrale $\int_0^\infty \psi e^{-\psi} \frac{dt}{t}$ *doit être convergente.*

D'une manière plus précise, évaluons le rapport des éléments différentiels, il a pour logarithme :

$$\begin{aligned}
&\operatorname{Log}\left(\psi\, e^{-\psi} \frac{1}{t}\right) - \operatorname{Log}\left(\varphi^t \operatorname{Log} \frac{1}{\varphi}\right) \\
&= \operatorname{Log}\psi - \psi - \operatorname{Log} t - t \operatorname{Log}\left(1 - \frac{\psi}{t}\right) - \operatorname{Log}\operatorname{Log} \frac{1}{1 - \frac{\psi}{t}} \\
&= \frac{\psi^2}{2t} + \frac{\psi^3}{3t^2} + \ldots - \operatorname{Log}\left(1 + \frac{\psi}{2t} + \frac{\psi^2}{3t^2} + \ldots\right).
\end{aligned}$$

Si nous supposons que $\lim\limits_{t=\infty} \frac{\psi^2}{t} = \lim\limits_{t=\infty} \frac{\psi}{t} = 0$, les éléments différentiels sont équivalents pour les grandes valeurs de t. Comme nous cherchons à prendre pour ψ une fonction qui croisse le plus lentement possible, les conditions $\frac{\psi^2}{t} \to 0$, $\frac{\psi}{t} \to 0$ ne sont pas gênantes. On voit sans peine que ψ doit croître plus rapidement que $\operatorname{Log}\operatorname{Log} t$ et peut croître plus lentement que $\operatorname{Log} t$. Il faut naturellement s'assurer que les conditions secondaires imposées à φ sont vérifiées (monotonie de φ et φ', croissance de $t\varphi'$).

INDICATIONS BIBLIOGRAPHIQUES.

BERTRAND (J.). — 1. Calcul des Probabilités (Paris, 1888).

BOREL (E.). — Traité du Calcul des Probabilités et de ses applications :
1. Tome 1, Fascicule 1, *Principes et formules classiques.*
2. Tome 4, Fascicule 3, *Valeur pratique et philosophie des probabilités.*

COPELAND (A. H.). — 1. Admissible Numbers in the Theory of Probability (*Am. J. of Math.*, vol. 50, 1928).
— 2. Consistency of the conditions determining Kollectivs (*Trans. of the Am. Math. Soc.*, vol. 42).

DOOB (J. L.). — 1. Stochastic processes depending of a continuous parameter (*Trans. of the Am. Math. Soc.*, vol. 42).

FINETTI (B. DE). — 1. La prévision, ses lois logiques, ses sources subjectives (*Annales de l'Institut Henri Poincaré*, 1937).

FRÉCHET (M.). — 1. Sur l'intégrale d'une fonctionnelle étendue à un ensemble abstrait (*Bull. de la Soc. Math. de France*, t. 43).
— Traité du Calcul des Probabilités, publié sous la direction de M. E. Borel *Recherches théoriques modernes :*
2. Livre I. — *Variables aléatoires.*
3. Livre II. — *Probabilités en chaîne.*
— 4. L'analyse générale et la question des fondements (Institut international de Coopération intellectuelle, *Réunion d'études sur les fondements et la méthode dans les sciences mathématiques*, Zurich, 6-9 décembre 1938).

KOLMOGOROFF (A.). — 1. Grundbegriffe der Wahrscheinlichkeitsrechnung (*Ergebnisse der Mathematik und ihre Grenzgebiette*, vol. 2, n° 3).

LÉVY (P.). — 1. Théorie de l'addition des variables aléatoires (*Monographies des probabilités*, Gauthier-Villars, 1937).
— 2. Calcul des Probabilités (Gauthier-Villars, 1925).

MISÈS (R. DE). — 1. Grundlagen der Wahrscheinlichkeitsrechnung (*Math. Zeitschrift*, vol. 5).
— 2. Wahrscheinlichkeit, Statistik und Wahrheit (Berlin, Springer, 1928). (Une seconde édition de cet Ouvrage a paru, que le lecteur pourra consulter.)
— 3. Wahrscheinlichkeitsrechnung und seine Anwendung (Leipzig und Wien, Deuticke, 1931).
— 4. Sul concetto di probabilità fondato sul limite di frequenze relative (*Giornale dell' Instituto Italiano degli Attuari*, 1936).
— 5. Théorie des probabilités. Fondements et applications (*Annales de l'Inst. Henri Poincaré*, 1932).

NICODYM. — 1. Sur une généralisation des intégrales de M. J. Radon (*Fundamenta Mathematicæ*, t. 15).

PETROWSKY. — 1. Zur ersten Randwertaufgabe der Warmeleitungsgleichung (Anhang) (*Compositio Mathematicæ*. 1, p. 168-179).

POPPER. — 1. Logik der Forschung (Julius Springer, Wien, 1935).

REICHENBACH (H.). — 1. Wahrscheinlichkeitslehre (Leyde, 1935).
— 2. Les fondements logiques du Calcul des Probabilités (*Annales de l'Inst. Henri Poincaré*, 1937).

SLUTSKY (E). — 1. Qualche proposizione relativa alla teorie delle funzione aleatorie (*Giornale dell' Instituto Italiano degli Attuari*, 1937).

WALD (A.). — 1. Die Widerspruchfreiheit des Kollectivbegriffes (*Ergebnisse eines Mathematischen Kolloquiums*, *Heft* 8, Wien, 1937).

Conférences internationales des sciences mathématiques, organisées à l'Université de Genève et publiées par les soins de M. R. Wavre. — Colloque consacré à la théorie des probabilités et présidé par M. Maurice Fréchet (Hermann, Paris, 1938).

PREMIÈRE PARTIE :

M. FRÉCHET. — *Les principaux courants dans l'évolution récente des recherches sur le Calcul des probabilités.*

DEUXIÈME PARTIE. — *Les fondements du Calcul des probabiltés :*

M. W. FELLER. — *Sur les axiomatiques du Calcul des probabilités et leurs relations avec les expériences.*

M. M. FRÉCHET. — *Exposé et discussion de quelques recherches récentes sur les fondements du Calcul des probabilités.*

M. R. DE MISÈS. — *Quelques remarques sur les fondements du Calcul des probabilités.*

M. J. F. STEFFENSEN. — *Fréquence et probabilité.*

M. A. WALD. — *Die Widerspruchfreiheit des Kollestivbegriffres.*

HUITIÈME PARTIE :

DE FINETTI. — *Résumé des conférences et des discussions au cours du colloque de 1937.*

Vu et approuvé :

Paris, le 22 octobre 1938.

LE DOYEN DE LA FACULTÉ DES SCIENCES,

CH. MAURAIN.

Vu et permis d'imprimer :

Paris, le 22 octobre 1938.

LE RECTEUR DE L'ACADÉMIE DE PARIS,

G. ROUSSY.