

ANNALES DE L'I. H. P.

B. HOSTINSKÝ

**Sur les probabilités relatives aux variables aléatoires
liées entre elles. Applications diverses**

Annales de l'I. H. P., tome 7, n° 2 (1937), p. 69-119

http://www.numdam.org/item?id=AIHP_1937__7_2_69_0

© Gauthier-Villars, 1937, tous droits réservés.

L'accès aux archives de la revue « Annales de l'I. H. P. » implique l'accord avec les conditions générales d'utilisation (<http://www.numdam.org/conditions>). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM

Article numérisé dans le cadre du programme
Numérisation de documents anciens mathématiques
<http://www.numdam.org/>

Sur les probabilités relatives aux variables aléatoires liées entre elles. Applications diverses

par

B. HOSTINSKÝ.

CHAPITRE I.

CHAINES DE MARKOFF SIMPLES A ÉLÉMENTS CONSTANTS.

1. **Hypothèses fondamentales.** — Supposons qu'on effectue une suite illimitée d'expériences et que le résultat de chacune d'elles consiste à amener un des r événements E_1, E_2, \dots, E_r , différents entre eux. Soient $0, 1, 2, \dots, n, \dots$ les numéros d'ordre de ces expériences; nous admettrons les hypothèses suivantes :

1° Si les résultats des autres expériences sont inconnus, il y aura une probabilité déterminée $p_k^{(n)}$ (*probabilité absolue*) pour que l'événement E_k se présente comme le résultat de la $n^{\text{ième}}$ expérience, et nous aurons

$$p_k^{(n)} \geq 0 \quad \text{pour } k = 1, 2, \dots, r; n = 0, 1, 2, \dots,$$

et

$$(1) \quad \sum_{k=1}^r p_k^{(n)} = 1 \quad \text{pour } n = 0, 1, 2, \dots$$

2° Étant donné que E_i est le résultat de la $n^{\text{ième}}$ expérience, il y a une probabilité déterminée p_{ik} (*probabilité de passage*) pour que E_k soit le résultat de la $(n+1)^{\text{ième}}$ expérience; cette probabilité ne dépend pas de n et elle ne change pas, si l'on connaît les résultats des autres expériences.

Nous aurons

$$(2) \quad p_{ik} \geq 0 \quad \text{pour } i, k = 1, 2, \dots, r$$

et

$$(3) \quad \sum_{p=1}^r p_{ip} = 1 \quad \text{pour } i = 1, 2, \dots, r.$$

Les probabilités des expériences successives seront alors liées en *chaîne simple*, à éléments constants p_{ik} .

Rappelons la *représentation géométrique* suivante : Soient A_1, A_2, \dots, A_r , r points fixes différents entre eux et M un point mobile dont la position actuelle dépend du résultat d'une certaine expérience; si les résultats des autres expériences sont inconnus, la probabilité pour que M se trouve, après la $n^{\text{ième}}$ expérience, au point A_k est égale à $p_k^{(n)}$. Étant donné que M se trouve, après la $n^{\text{ième}}$ expérience, au point A_i , la probabilité pour qu'il passe, à la suite de la $(n+1)^{\text{ième}}$ expérience, au point A_k , est égale à p_{ik} .

Attachons un nombre déterminé α_k à chaque événement

$$E_k \quad (k = 1, 2, \dots, r),$$

et soit $x^{(n)}$ une variable aléatoire attachée au résultat de la $n^{\text{ième}}$ expérience suivant la règle suivante : $x^{(n)} = \alpha_k$, si cette expérience a donné E_k comme résultat.

2. Formules générales. Les probabilités $P_{ik}^{(n)}$. Moyennes conditionnées.

— Les principes fondamentaux concernant l'addition et la multiplication des probabilités (Borel [1], Chap. I) ⁽¹⁾ donnent les relations suivantes entre $p_k^{(n)}$ et $p_k^{(n-1)}$:

$$(4) \quad p_k^{(n)} = \sum_{i=1}^r p_i^{(n-1)} p_{ik} \quad (k = 1, 2, \dots, r, n = 1, 2, 3, \dots).$$

Diminuons n d'une unité; la formule (4) donne alors

$$p_k^{(n)} = \sum_{i=1}^r p_i^{(n-2)} P_{ik}^{(2)},$$

⁽¹⁾ Les numéros entre crochets renvoient à la Bibliographie placée à la fin du texte.

avec

$$P_{ik}^{(2)} = \sum_{j=1}^r P_{ij} P_{jk};$$

et l'on trouve, en général,

$$(5) \quad p_k^{(m+n)} = \sum_{i=1}^r P_i^{(m)} P_{ik}^{(n)} \quad (m = 0, 1, 2, \dots, n = 1, 2, \dots),$$

où

$$(6) \quad P_{ik}^{(n)} = \sum_{\alpha=1}^r \sum_{\beta=1}^r \dots \sum_{\lambda=1}^r P_{i\alpha} P_{\alpha\beta} \dots P_{\lambda k};$$

il faut sommer au second membre par rapport aux $n - 1$ indices $\alpha, \beta, \dots, \lambda$. La formule (6) donne immédiatement l'interprétation suivante de $P_{ik}^{(n)}$: $P_{ik}^{(n)}$ est la probabilité pour que la $(m + n)$ ^{ème} expérience donne E_k , si la m ^{ème} a donné E_i ($m = 0, 1, 2, \dots$). En groupant convenablement les sommations dans (6), on trouve

$$(7) \quad P_{ik}^{(m+n)} = \sum_{j=1}^r P_{ij}^{(m)} P_{jk}^{(n)} \quad (m, n = 1, 2, \dots),$$

avec

$$P_{ik}^{(1)} = p_{ik}.$$

Écrivons encore les formules (5) pour $m = 0$:

$$(5 \text{ bis}) \quad p_k^{(n)} = \sum_{i=1}^r p_i^{(0)} P_{ik}^{(n)},$$

et la formulè (7) pour $m = 1$:

$$(7 \text{ bis}) \quad P_{ik}^{(n+1)} = \sum_{j=1}^r P_{ij}^{(1)} P_{jk}^{(n)}.$$

Il résulte de (6) que

$$(8) \quad \sum_{k=1}^r P_{ik}^{(n)} = 1 \quad (i = 1, 2, \dots, r; n = 1, 2, \dots).$$

Soit $\alpha^{(n)}$ la valeur moyenne de $x^{(n)}$, lorsque les résultats des expériences restent inconnus, et soit $\alpha_i^{(n)}$ la valeur moyenne de $x^{(n)}$, calculée avec la condition que $x^{(0)} = \alpha_i$. Nous avons

$$(9) \quad \alpha^{(n)} = \sum_{k=1}^r p_k^{(n)} \alpha_k$$

et

$$(10) \quad a_i^{(n)} = \sum_{k=1}^r P_{ik}^{(n)} a_k.$$

Remplaçons, dans (9), $P_k^{(n)}$ par l'expression (5 bis); il vient, si nous tenons compte de (10),

$$(11) \quad a^{(n)} = \sum_{k=1}^r \sum_{i=1}^r p_i^{(0)} P_{ik}^{(n)} a_k = \sum_{i=1}^r p_i^{(0)} a_i^{(n)}.$$

La formule (11) est un cas particulier de la formule sur les *moyennes conditionnées* : la valeur moyenne $a^{(n)}$ apparaît comme une fonction linéaire et homogène des moyennes $a_i^{(n)}$, les coefficients $p_i^{(n)}$ étant respectivement égaux aux probabilités des équations $x^{(0)} = \alpha_i$ (voir, pour la formule générale, Kolmogoroff [2], p. 46, Fréchet [3], p. 128).

La notion de moyenne conditionnée dont l'emploi n'est pas limité au cas particulier des chaînes que nous considérons ici, me semble être digne d'un intérêt particulier. Elle pourrait conduire, je pense, à des modifications profondes du Calcul des probabilités. En effet, si l'on arrive par un procédé quelconque à des valeurs moyennes [telles que, par exemple, $a^{(n)}$ et $a_i^{(n)}$ ($n = 1, 2, 3, \dots$), dans les équations (11)], les probabilités [$p_i^{(0)}$ pour $1, 2, \dots, r$] pourront être calculées comme solutions d'un système d'équations linéaires obtenu en prenant successivement $n = 1, 2, 3, \dots$ dans l'équation (11).

Ainsi, dans les applications aux problèmes réels, on pourrait regarder les moyennes comme des quantités données et en déduire les probabilités au moyen d'un système d'équations bilinéaires entre les valeurs moyennes et les probabilités.

Les formules (6) et (7) mettent en évidence la propriété suivante : si les p_{ik} sont les éléments d'une matrice P de degré r (1), les $P_{ik}^{(n)}$ seront les éléments de la $n^{\text{ième}}$ puissance de P (voir Hostinský [4], p. 15. Cet ouvrage [4] résume les principaux résultats de la théorie des chaînes jusqu'à 1931).

3. Les probabilités $p_k^{(n)}$ et $P_{ik}^{(n)}$ envisagées comme solutions d'équations aux différences finies. — *a.* L'équation (4) est une équation aux

(1) C'est-à-dire d'une matrice à r lignes et à r colonnes.

différences finies, si l'on y regarde les p_{ik} comme quantités connues et $p_k^{(n)}$ comme fonction inconnue des deux indices k et n ; (4) est du $(r - 1)^{\text{ième}}$ ordre par rapport à k , car elle contient les r quantités $p_1^{(n-1)}$, $p_2^{(n-1)}$, ..., $p_r^{(n-1)}$ et du premier ordre par rapport à n , car l'indice d'itération n n'y entre que par deux valeurs n et $n - 1$. D'une manière analogue, (7 bis) est une équation aux différences finies pour la fonction inconnue $P_{ik}^{(n)}$; si nous déterminons les $P_{ik}^{(n)}$ par une méthode quelconque, les $p_k^{(n)}$ seront également connues comme le montre la formule (5 bis); il faut toutefois que les $p_i^{(0)}$ ($i = 1, 2, \dots, r$) soient données.

Le problème principal consiste donc à déterminer les $P_{ik}^{(n)}$; nous supposons toujours que les p_{ik} sont données et qu'elles satisfont aux conditions (2) et (3). La formule (7 bis) permet de calculer les $P_{ik}^{(2)}$, $P_{ik}^{(3)}$, ... de proche en proche par voie de récurrence; pour obtenir une expression explicite de $P_{ik}^{(n)}$, on peut cependant employer la méthode suivante, analogue à celle qui sert à former les solutions des équations linéaires aux dérivées partielles, méthode qui tient compte de (7 bis) ainsi que de l'équation

$$(7 \text{ ter}) \quad P_{ik}^{(n)} = \sum_{j=1}^r p_{ij} P_{jk}^{(n-1)}.$$

Nous regardons $P_{ik}^{(n)}$ comme une fonction de trois indices i, k, n et nous essayons de satisfaire à (7 bis) en posant

$$(12) \quad P_{ik}^{(n)} = f_i g_k h^{(n)},$$

où chacun des trois facteurs du second membre ne dépend que d'une seule variable. La substitution du second membre de (12) à la place de $P_{ik}^{(n)}$ dans (7 bis) donne

$$f_i g_k h^{(n+1)} = \sum_{u=1}^r f_i g_u p_{uk} h^{(n)};$$

en supprimant le facteur f_i , nous trouvons

$$\frac{h^{(n+1)}}{h^{(n)}} = \frac{\sum_{u=1}^r g_u p_{uk}}{g_k}.$$

Le premier membre ne dépend pas de k et le second ne dépend pas de n ; les deux rapports sont donc égaux à une même constante s . Nous

aurons donc

$$(13) \quad s g_k - \sum_{u=1}^r p_{uk} g_u = 0 \quad (k = 1, 2, \dots, r).$$

Si une au moins des quantités g_k est différente de zéro (et ce cas est le seul qui nous intéresse), il faudra que

$$\begin{vmatrix} p_{11} - s & p_{12} & \dots & p_{1r} \\ p_{21} & p_{22} - s & \dots & p_{2r} \\ \dots & \dots & \dots & \dots \\ p_{r1} & p_{r2} & \dots & p_{rr} - s \end{vmatrix} = 0.$$

Les valeurs admissibles de s sont donc des racines de cette équation du $r^{\text{ième}}$ degré (*équation caractéristique*), parmi lesquelles se trouve l'unité, ainsi que le montrent les conditions (3).

Supposons que *toutes les racines soient simples* et désignons-les par $s_0 = 1, s_1, s_2, \dots, s_{r-1}$. Une racine s_j étant donnée, les rapports des quantités g_k seront déterminés par les équations (13) (avec s_j à la place de s). Écrivons φ_{kj} au lieu de g_k : les équations (13) seront remplacées par les suivantes :

$$(14) \quad \sum_{u=1}^r p_{uk} \varphi_{uj} - s_j \varphi_{kj} = 0 \quad (k = 1, 2, \dots, r).$$

Si nous écrivons ψ_{ij} au lieu de f_i , et si nous tenons compte de (7^{ter}), nous trouvons que

$$(15) \quad \sum_{u=1}^r p_{iu} \psi_{uj} - s_j \psi_{ij} = 0 \quad (i = 1, 2, \dots, r).$$

Il résulte de (3), (14) et (15) que

$$\sum_{i=1}^r \psi_{ij} \varphi_{ih} = 0 \quad \text{pour } j \neq h;$$

nous convenons de multiplier les ψ_{ij} par une constante telle que l'on ait

$$(16) \quad \sum_{i=1}^r \psi_{ij} \varphi_{ij} = 1 \quad (j = 0, 1, \dots, r-1).$$

Ainsi chaque racine s_j donne une solution de la forme (12); en faisant la somme par rapport à j , on trouve que

$$(17) \quad P_{ik}^{(n)} = \sum_{j=0}^{r-1} \psi_{ij} \varphi_{kj} s_j^n.$$

Cette formule a été donnée par M. Romanovsky en 1929 [5]; voir aussi [4], p. 22, [16], p. 8. En résumé, *si toutes les racines s_j de l'équation caractéristique sont simples, les équations (7 bis) et (7 ter) avec les conditions (3) admettent la solution (17) où les ψ_{ij} et φ_{kj} sont déterminées par les équations (14), (15) et (16).*

Remarquons que la formule (17) permet d'exprimer les $P_{ik}^{(n)}$ (même dans le cas où $n = 1$) en fonction des racines s_j .

b. Il serait intéressant d'obtenir une expression de $P_{ik}^{(n)}$ analogue à (17), même dans le cas le plus général où les racines sont multiples. Pour obtenir une telle formule on pourrait se servir de certains résultats obtenus par M. Perron [6]; M. Romanovsky [7] a donné différentes généralisations de la formule (17).

4. Cas semi-régulier et cas régulier. — *a.* Considérons la question suivante : $P_{ik}^{(n)}$ admet-il une limite déterminée quand n augmente indéfiniment ? L'existence de cette limite dépend surtout des valeurs absolues des racines s_i autres que $s_0 = 1$. Appelons, avec Fréchet, *cas semi-régulier* le cas où toutes les limites $\lim_{n \rightarrow \infty} P_{ik}^{(n)}$ ($i, k = 1, 2, \dots, r$) existent, mais où elles dépendent en général des deux indices i et k ; si toutes ces limites existent et si elles ne dépendent que du second indice k , nous aurons le *cas régulier*. Il résulte des travaux de MM. Kaucký [8], Konečný [9, 10] et de M. Fréchet [11] : 1° que la condition nécessaire et suffisante pour qu'on soit dans le cas semi-régulier est que la seule racine de module 1 de l'équation caractéristique soit précisément l'unité; 2° que la condition nécessaire et suffisante pour qu'on soit dans le cas régulier est que l'équation caractéristique n'ait pas d'autres racines de module 1 en dehors de l'unité et que, au surplus, l'unité soit une racine simple.

Voici deux exemples simples (pour $r = 2$) :

$$1^\circ \quad p_{11} = p_{22} = 1, \quad p_{12} = p_{21} = 0.$$

Nous avons, pour n quelconque,

$$P_{11}^{(n)} = P_{22}^{(n)} = 1, \quad P_{12}^{(n)} = P_{21}^{(n)} = 0.$$

L'équation caractéristique admet $s = 1$ comme racine double, les limites existent (cas semi-régulier).

$$2^\circ \quad p_{11} = p_{22} = 0, \quad p_{12} = p_{21} = 1.$$

Nous avons

$$P_{11}^{(n)} = P_{22}^{(n)} = \frac{1 + (-1)^n}{2}, \quad P_{12}^{(n)} = P_{21}^{(n)} = \frac{1 + (-1)^{n+1}}{2};$$

l'équation caractéristique admet encore $s = 1$ comme racine double, mais les limites n'existent pas; ce cas n'est pas un cas semi-régulier.

J'ai montré ([4], p. 30) que si les limites $P_{ik}^{(n)}$ existent et si la condition

$$(18) \quad \sum_{i=1}^r p_{ik} = 1 \quad (k = 1, 2, \dots, r)$$

est vérifiée, toutes ces limites sont égales entre elles. M. Fréchet [11] a complété ce théorème en montrant que la condition nécessaire et suffisante pour que les $P_{ik}^{(n)}$ tendent vers une même limite, indépendante de i et de k , est que l'équation caractéristique n'ait pas de racine de module 1 autre que l'unité, que l'unité soit une racine simple et enfin que la condition (18) soit vérifiée.

b. Le cas régulier se présente, si toutes les probabilités p_{ik} sont positives. Avec cette hypothèse, Markoff a démontré en 1907 l'existence des limites $P_k = \lim_{n \rightarrow \infty} P_{ik}^{(n)}$ (voir [4], n° 6). *Le cas régulier se présente encore si les p_{ik} ne sont pas toutes positives, mais s'il existe un entier m tel que toutes les $P_{ik}^{(m)}$ soient positives; les limites P_k existent alors et elles sont toutes positives.*

A l'époque où j'ai démontré ce résultat (voir [12] et [4], p. 35), je ne connaissais pas un travail de Frobenius ([13], théorème IX) qui permet de le compléter. Une matrice P de degré r est dite décomposable, suivant Frobenius, s'il est possible de choisir r_1 indices de lignes et $r - r_1$ indices de colonnes, de telle sorte que : 1° dans ce groupe de r indices chacun des nombres 1, 2, ..., r apparaisse une seule fois; et que : 2° les éléments qui se trouvent à l'intersection d'une ligne choisie avec une colonne

choisie soient tous égaux à zéro. Il résulte de cette définition et des formules fondamentales (nous désignons par p_{ik} les éléments de la matrice)

$$P_{ik}^{(2)} = \sum_{j=1}^r p_{ij} p_{jk}, \quad P_{ik}^{(n+1)} = \sum_{j=1}^r P_{ij}^{(n)} p_{jk}$$

qu'à ces points d'intersection, les $P_{ik}^{(n)}$ sont égaux à zéro pour n quelconque. Nous ne considérons dans ce paragraphe que des *matrices indécomposables*.

Une telle matrice est dite *primitive*, si son équation caractéristique admet une seule racine de module maximum et si cette racine est simple.

Voici le résultat de Frobenius : *si les éléments de P ne sont pas négatifs, et si P est une matrice primitive, une des puissances P^α de P est positive* (c'est-à-dire les éléments de P^α sont tous positifs); inversement, *si l'une des puissances d'une matrice non négative P est positive, P est une matrice primitive*. Dans le cas où les éléments p_{ik} de P satisfont à la condition (3), le module maximum des racines de l'équation caractéristique est l'unité; si la matrice est primitive, la seule racine de module maximum est égale à l'unité.

c. Dans le cas régulier les limites P_k sont déterminées par les équations suivantes qui se déduisent de (7 bis) en faisant augmenter indéfiniment n :

$$(19) \quad P_k = \sum_{i=1}^r P_j p_{jk} \quad (k = 1, 2, \dots, r),$$

avec la condition supplémentaire [voir l'équation (8)]

$$(19 \text{ bis}) \quad \sum_{k=1}^r P_k = 1.$$

Les équations (19) admettent toujours une solution, même dans le cas non régulier, car leur déterminant est égal à zéro en vertu de (3). Dans le cas régulier, les P_k sont déterminées d'une manière univoque par (19) et (19 bis). Or, le théorème de Frobenius, rappelé plus haut, nous montre que dans le cas où la matrice des probabilités P_{ki} est indécomposable, non négative et primitive, les limites P_k existent; dans ce cas elles sont également déterminées d'une manière univoque par (19) et (19 bis).

d. Dans le cas où toutes les racines de l'équation caractéristique sont simples, on peut appliquer la formule (17).

Si, outre cela, la matrice des p_{ik} est primitive, les propriétés des quantités ψ_{ij} et φ_{kj} (voir n° 3) montrent que $\psi_{i0} = 1$, $\varphi_{k0} = P_k$; dans ce cas (17) se réduit donc à

$$(17 \text{ bis}) \quad P_{ik}^{(n)} = P_k + \sum_{j=1}^{r-1} \psi_{ij} \varphi_{kj} s_j^n.$$

Remarque. — M. Fréchet a montré ([14], p. 23) qu'avec l'hypothèse (3), si S est une racine de l'équation caractéristique, et si ω est le plus petit des nombres $p_{11}, p_{22}, \dots, p_{rr}$, on a la formule

$$|S - \omega| < 1 - \omega.$$

5. **Problème de deux urnes avec échange de boules.** — Considérons deux urnes A et B, renfermant e boules chacune; dans le nombre total $2e$ de boules il y en a autant de blanches que de noires. Supposons que l'on tire en même temps une boule de chaque urne, et qu'ensuite on mette dans chaque urne la boule extraite de l'autre.

Supposons que l'on répète cette opération un nombre quelconque de fois, en agitant à chaque fois les urnes pour bien mêler les boules et cherchons la probabilité pour qu'après n opérations, il y ait k boules blanches dans l'urne A (voir Laplace [15], p. 292).

Soit p_{ik} ($i, k = 0, 1, 2, \dots, e$) la probabilité pour que le nombre de boules blanches dans A soit égal à k après le tirage (c'est-à-dire après l'opération qui vient d'être décrite), étant donné qu'avant le tirage il était égal à i ; nous avons

$$(20) \quad p_{j,i-1} = \frac{j^2}{e^2}, \quad p_{jj} = 2 \frac{(e-j)j}{e^2}, \quad p_{j,j+1} = \frac{(e-j)^2}{e^2} \quad \text{pour } 0 < j < e$$

et

$$(20') \quad p_{00} = 0, \quad p_{01} = 1, \quad p_{e,e-1} = 1, \quad p_{ee} = 0;$$

la probabilité p_{ik} est égale à zéro si $|i - k| \geq 2$. Il résulte de ces formules que les conditions (3) sont vérifiées (avec $r = e + 1$), c'est-à-dire que

$$\sum_{k=0}^{e+1} p_{ik} = 1 \quad (i = 0, 1, 2, \dots, e).$$

Désignons maintenant (conformément aux notations introduites au n° 1) par $p_k^{(n)}$ la probabilité absolue pour qu'il y ait k boules blanches dans A, après le $n^{\text{ième}}$ tirage. Cette probabilité peut s'exprimer au moyen des probabilités $p_{k-1}^{(n-1)}$, $p_k^{(n-1)}$ et $p_{k+1}^{(n-1)}$. En effet, $p_k^{(n)}$ est la probabilité totale, égale à la somme de trois probabilités qui correspondent respectivement aux trois cas possibles pour lesquels le nombre de boules blanches dans A avant le $n^{\text{ième}}$ tirage est égal à $k - 1$, à k ou à $k + 1$. Nous avons donc

$$p_k^{(n)} = p_{k-1}^{(n-1)} p_{k-1,k} + p_k^{(n-1)} p_{kk} + p_{k+1}^{(n-1)} p_{k+1,k},$$

ce qui n'est autre chose que la formule (4). Si nous remplaçons les p_{ik} par les expressions (20), il vient

$$(21) \quad p_k^{(n)} = p_{k-1}^{(n-1)} \frac{(e-k+1)^2}{e^2} + 2 p_k^{(n-1)} \frac{(e-k)k}{e^2} + p_{k+1}^{(n-1)} \frac{(k+1)^2}{e^2},$$

équation déjà obtenue par Laplace. Si l'on connaît les probabilités $p_0^{(0)}$, $p_1^{(0)}$, ..., $p_e^{(0)}$ relatives aux différentes compositions de l'urne A avant le premier tirage, on pourra en déduire les $p_k^{(n)}$ de proche en proche. Si, en particulier, la composition initiale de A est donnée et si i est le nombre de boules blanches qu'elle renferme, nous aurons

$$p_i^{(0)} = 1, \quad p_j^{(0)} = 0 \quad \text{pour } j \neq i,$$

et

$$p_k^{(n)} = P_{ik}^{(n)} \quad \text{avec} \quad P_{ik}^{(1)} = p_{ik},$$

où $P_{ik}^{(n)}$ ($k = 0, 1, 2, \dots, e; n = 1, 2, 3, \dots$) est la probabilité pour que A renferme k boules blanches après le $n^{\text{ième}}$ tirage, lorsqu'on sait qu'il y en avait i avant le premier. L'équation (21) devient, en y substituant $P_{ik}^{(n)}$ à la place de $p_k^{(n)}$,

$$P_{ik}^{(n)} = P_{i,k-1}^{(n-1)} \frac{(e-k+1)^2}{e^2} + 2 P_{ik}^{(n-1)} \frac{(e-k)k}{e^2} + P_{i,k+1}^{(n-1)} \frac{(k+1)^2}{e^2},$$

ce qui n'est autre chose que l'équation (7 bis) où il faut écrire n au lieu de $n + 1$, et où les p_{ik} sont définies par les formules (20). Il est facile à démontrer que les $P_{ik}^{(n)}$ sont toutes positives à partir d'une valeur de l'indice n suffisamment élevée; donc (voir n° 4, b, c, d) la matrice de degré $e + 1$ des p_{ik} est primitive, la racine $s_0 = 1$ de l'équation séculaire est simple, et les $P_{ik}^{(n)}$ admettent, pour n infini, des valeurs limites P_k indépendantes de i . Les P_k sont déterminées par les équations (19) et

(19 bis), qui s'écrivent maintenant

$$P_k = \sum_{i=0}^e P_i p_{ik}, \quad \sum_{k=0}^e P_k = 1,$$

les p_{ik} étant données par (20). La résolution de ce système linéaire donne

$$P_k = \frac{\binom{e}{e-k} \binom{e}{k}}{\binom{2e}{e}}.$$

Le second membre de cette formule est égal à la probabilité pour qu'un groupe de e boules, choisies au hasard parmi les $2e$ boules (dont e sont noires et e blanches) contienne k boules blanches. Donc, *après un nombre infini de tirages successifs la probabilité pour que A renferme k boules blanches est égale à celle d'obtenir k boules blanches dans un groupe de e boules extrait d'une seule urne où se trouvent e boules blanches et e noires.* Ceci est le résultat qui a été obtenu par Laplace par une autre voie (voir mes Leçons [16], n° 21).

Une étude plus approfondie de ce problème exige la résolution de l'équation caractéristique. Soit $D_e(s)$ le premier membre de cette équation; $D_e(s)$ est un polynôme de degré $e + 1$ en s , si e est le nombre de boules dans A. Nous avons, les p_{ik} étant toujours déterminées par les formules (20),

$$D_e(s) = \begin{vmatrix} p_{00} - s & p_{01} & \dots & p_{0e} \\ p_{10} & p_{11} - s & \dots & p_{1e} \\ \dots & \dots & \dots & \dots \\ p_{e0} & p_{e1} & \dots & p_{ee} - s \end{vmatrix}.$$

Voici les résultats du calcul pour $2 \leq e \leq 6$:

$$D_2(s) = -(s-1) \left(s + \frac{1}{2} \right) s$$

$$D_3(s) = (s-1) \left(s + \frac{1}{9} \right) \left(s + \frac{1}{3} \right) \left(s - \frac{1}{3} \right)$$

$$D_4(s) = -(s-1) \left(s - \frac{1}{2} \right) \left(s + \frac{1}{4} \right) \left(s + \frac{1}{8} \right) \left(s - \frac{1}{8} \right)$$

$$D_5(s) = (s-1) \left(s - \frac{3}{5} \right) \left(s - \frac{7}{25} \right) \left(s + \frac{1}{5} \right) \left(s + \frac{3}{25} \right) \left(s - \frac{1}{25} \right)$$

$$D_6(s) = -(s-1) \left(s - \frac{2}{3} \right) \left(s - \frac{7}{18} \right) \left(s + \frac{1}{6} \right) \left(s - \frac{1}{6} \right) \left(s + \frac{1}{9} \right) s.$$

Si donc le nombre e ne dépasse pas six, toutes les racines de $D_e(s)=0$ sont simples, et l'on peut appliquer la formule de M. Romanovsky que nous écrirons sous la forme (17 bis), car la matrice des p_{ik} est primitive (1). Convenons de ranger les s_i suivant leurs valeurs absolues décroissantes, de sorte que

$$s_0 = 1, \quad 1 > s_1 \geq s_2 \geq \dots \geq s_e.$$

La formule

(17 bis)
$$P_{ik}^{(n)} = P_k + \sum_{j=1}^e \varphi_{kj} \psi_{ij} s_j^n$$

conduit aux relations suivantes

$$(22) \quad \begin{cases} \lim_{n \rightarrow \infty} P_{ik}^{(n)} = P_k \\ \lim_{n \rightarrow \infty} (P_{ik}^{(n)} - P_k - \varphi_{k1} \psi_{i1} s_1^n) : s_2^n = \varphi_{k2} \psi_{i2} \\ \lim_{n \rightarrow \infty} (P_{ik}^{(n)} - P_k - \varphi_{k1} \psi_{i1} s_1^n - \varphi_{k2} \psi_{i2} s_2^n) : s_3^n = \varphi_{k3} \psi_{i3} \\ \dots \dots \dots \end{cases}$$

Les expressions qui figurent dans les parenthèses s'obtiennent en retranchant successivement de $P_{ik}^{(n)} - P_k$ un, deux, trois, ... des premiers termes de la somme du second membre (17 bis); si plusieurs racines s_j ont des valeurs absolues égales, il faut retrancher tous les termes de cette somme qui correspondent à une même valeur absolue des racines s_j .

Les relations (22) (2) pourraient être vérifiées par la statistique de tirages de la manière suivante. Faisons un très grand nombre de tirages successifs et notons, après chaque tirage, la composition de l'urne A. Les quantités $P_k, \varphi_{kj}, \psi_{ij}, s_j$ qui entrent dans les formules (22) sont connues, elles se calculent comme nous l'avons indiqué au n° 4.

Pour obtenir une valeur approchée de $P_{ik}^{(n)}$, procédons comme suit : soit N_i le nombre des cas où il y avait i boules blanches dans A; et soit N_{ik} le nombre des cas qui vérifient la condition suivante : il y avait

(1) M. Potoček a remarqué que les racines de $D_e(s)=0$ sont exprimées par la formule

$$s_k = \frac{(e-k)^2 - k}{e^2} \quad (k = 0, 1, 2, \dots, e),$$

valable pour $2 \leq e \leq 6$, et peut-être dans le cas général.

(2) La manière d'obtenir ces relations ne diffère pas de celle que l'on emploie pour résoudre une équation algébrique suivant la méthode de Graeffe.

i boules blanches dans A après le $m^{\text{ième}}$ tirage (m quelconque) et k boules blanches après le $(m+n)^{\text{ième}}$ tirage. Si N_i et n sont assez grands, le rapport $N_{ik} : N_i$ donne une valeur approchée de $P_{ik}^{(n)}$; en substituant ce rapport à la place de $P_{ik}^{(n)}$ dans la première relation (22), on devrait obtenir

$$\left(\frac{N_{ik}}{N_i} - P_k - \varphi_{k1} \psi_{i1} s_1^n \right) : s_2^n$$

à peu près égale à $\varphi_{k2} \psi_{i2}$, et ainsi de suite pour les autres relations (22).

6. Tirages de plusieurs urnes avec échange de boules et théorie de la diffusion. — Laplace a également étudié le problème suivant : m urnes sont disposées circulairement; elles contiennent des boules blanches et noires. On tire en même temps une boule de chaque urne et l'on met ensuite dans la seconde urne la boule extraite de la première, dans la troisième urne la boule extraite de la seconde . . . , et dans la première urne la boule extraite de la $m^{\text{ième}}$. Il s'agit de démontrer que le nombre moyen de boules blanches sera le même pour chaque urne après un nombre infini de tirages successifs; on suppose que chaque urne renferme le même nombre de boules.

Ce problème, ainsi que d'autres problèmes où les urnes seraient disposées de diverses manières, peuvent être considérés comme donnant des modèles du phénomène de la diffusion dans les fluides. En effet, lorsque, par exemple, une goutte d'eau colorée tombe dans un vase rempli d'eau pure, il se produit un mélange progressif spontané, les molécules de l'eau colorée tendant à se répartir uniformément dans le vase. A chaque instant ce mélange peut être regardé comme le résultat d'une permutation entre les molécules de l'eau colorée et de celles de l'eau pure. Chaque molécule peut, bien entendu, occuper une infinité de positions (dépendant de paramètres continus), et il y a toujours un espace libre inoccupé entre les diverses molécules. Supposons que le volume du vase soit divisé en petits volumes partiels v_1, v_2, v_3, \dots . Si certaines molécules quittent à un instant donné le volume v_1 , d'autres y entrent et cet échange peut être rapproché de l'échange des boules extraites de deux ou plusieurs urnes dans les problèmes de Laplace que nous venons d'examiner.

Admettons que le passage d'une molécule, laquelle partant d'un volume

v_i aboutit à un autre volume v_j adjacent à v_i , dure un certain temps (variable entre des limites finies); admettons également que tout passage d'un volume v_i à un volume v_k non adjacent à v_i ne puisse avoir lieu que de proche en proche par l'intermédiaire de volumes adjacents deux à deux. Dans ces conditions on peut affirmer que la vitesse de la diffusion sera une quantité finie.

Ce résultat est contraire de celui déduit de la théorie classique de la diffusion. Dans mes leçons de 1930 [(16), n° 22] j'ai rappelé différents exemples calculés par Smoluchowski. Bornons-nous ici de considérer le cas le plus simple de la diffusion sur l'axe illimité Ox en l'absence de toute force extérieure. La densité de probabilité $v(x)$ du passage du point $x = 0$ au point x en t secondes, est donnée par la formule de Gauss

$$v(x) = \frac{1}{2\sqrt{\pi t}} \cdot e^{-\frac{x^2}{4t}}$$

qui satisfait à l'équation de la diffusion

$$\frac{\partial v}{\partial t} = \frac{\partial^2 v}{\partial x^2};$$

donc la probabilité pour que la particule se trouve, à l'instant t ($t > 0$) entre x_1 et x_2 ($x_1 < x_2$) est égale à

$$f(x_1, x_2) = \int_{x_1}^{x_2} \frac{e^{-\frac{x^2}{4t}}}{2\sqrt{\pi t}} dx.$$

Choisissons t arbitrairement petit et x_1 arbitrairement grand : $f(x_1, x_2)$ sera toujours positive, ce qui veut dire que la vitesse de diffusion est infinie. La formule de Gauss, malgré toutes les applications importantes et utiles dont elle est susceptible, présente donc l'inconvénient précédent, qui est assez considérable.

Remarque sur l'emploi de la formule de Gauss dans la théorie des erreurs et dans la théorie cinétique des gaz. — D'après la théorie classique des erreurs d'observation, la probabilité pour qu'une erreur soit comprise entre x_1 et x_2 est égale à $f(x_1, x_2)$. On est conduit ainsi à attribuer des probabilités positives (quoique très petites) à des erreurs arbitrairement grandes. Or, des erreurs très grandes sont évidemment impossibles.

Considérons un gaz à température constante et uniforme, enfermé

dans un vase imperméable à la chaleur. La probabilité pour que la composante de la vitesse d'une molécule suivant une direction fixe soit comprise entre les limites x_1 et x_2 est encore donnée par la fonction $f(x_1, x_2)$, dans laquelle il faut remplacer t par une constante convenable. On admet par conséquent que des vitesses arbitrairement grandes possèdent des probabilités positives. Soit E l'énergie interne totale du gaz; elle est égale à la somme des énergies cinétiques qui proviennent du mouvement de translation des molécules (nous ne tenons pas compte des rotations des molécules). Si toutes les molécules sauf une étaient arrêtées et, si par conséquent, toute l'énergie du gaz se réduisait à l'énergie d'une seule molécule de masse m et de vitesse v , nous aurions

$$\frac{m v^2}{2} = E, \quad v^2 = \frac{2 E}{m}.$$

Le carré de la vitesse possède donc une limite supérieure $\frac{2 E}{m}$.

Si E est donnée, il est impossible que v^2 surpasse cette limite.

Il semble donc nécessaire de chercher d'autres formules pouvant remplacer avantageusement celle de Gauss, soit dans la théorie de la diffusion, soit dans la théorie des erreurs, soit dans la théorie cinétique des gaz.

7. Dispersion. — a . Revenons aux notations introduites plus haut (n^{os} 1 et 2) et supposons que nous soyons dans le *cas régulier*; les limites $\lim_{n \rightarrow \infty} P_{ik}^{(n)} = P_k$ existent et elles ne dépendent pas de i . La valeur moyenne $\alpha_i^{(n)}$ de la variable $x^{(n)}$ attachée à la $n^{\text{ième}}$ expérience (sous l'hypothèse que $x^{(0)} = \alpha_i$) admet la valeur limite a quand n augmente indéfiniment, et nous avons

$$a = \lim_{n \rightarrow \infty} \alpha_i^{(n)} = \sum_{k=1}^r P_k \alpha_k,$$

ce qui résulte immédiatement de la formule (10).

Nous désignerons par $E(x)$ la valeur moyenne de x . Markoff a démontré que la limite

$$\lim_{n \rightarrow \infty} E \frac{[x^{(1)} - a + x^{(2)} - a + \dots + x^{(n)} - a]^2}{n}$$

existe; il la désigne par $\frac{C}{2}$. Nous appellerons ce nombre *dispersion*. On

a donné différentes expressions de $\frac{C}{2}$ (voir [4], nos 11 et 12). Récemment M. Potoček [17] et M. Fréchet [11] ont donné à $\frac{C}{2}$ une forme particulièrement simple. Voici la formule

$$(23) \quad \frac{C}{2} = \sum_{k=1}^r (\alpha_k - \alpha)^2 P_k + 2 \sum_{i=1}^r \sum_{k=1}^r (\alpha_i - \alpha) (\alpha_k - \alpha) P_k s_{ki},$$

où

$$(24) \quad s_{ik} = \sum_{n=1}^{\infty} (P_{ik}^{(n)} - P_k).$$

Les quantités s_{ik} s'obtiennent, suivant M. Fréchet, par la résolution des systèmes linéaires suivants :

$$s_{ik} - (p_{ik} - P_k) = \sum_{j=1}^r s_{ij} p_{jk}$$

$$\sum_{i=1}^r s_{ij} = 0,$$

où $i, k = 1, 2, \dots, r$.

b. Rappelons les formules valables dans le cas où $r = 2$ (cas régulier). Les valeurs limites P_1 et P_2 sont données par les formules

$$P_1 = \frac{p_{21}}{1 - p_{11} + p_{21}}, \quad P_2 = \frac{p_{12}}{1 - p_{11} + p_{21}}.$$

L'équation caractéristique a comme racines

$$s_0 = 1, \quad s_1 = p_{11} - p_{21},$$

et l'on trouve (avec les notations du n° 3)

$$\varphi_{0k} = P_k, \quad \psi_{0k} = 1, \quad \varphi_{11} = 1, \quad \varphi_{21} = 1, \quad \psi_{11} = P_2, \quad \psi_{21} = -P_1,$$

de sorte que la formule (17 bis) du n° 4 devient

$$P_{ik}^{(n)} = P_k + (-1)^{i+k} P_{3-i} (p_{11} - p_{21})^n; \quad i, k = 1, 2.$$

La formule (24) donne

$$s_{ik} = \sum_{n=1}^{\infty} \varphi_{k1} \psi_{i1} (p_{11} - p_{21})^n = (-1)^{i+k} P_{3-i} \frac{p_{11} - p_{21}}{1 - p_{11} + p_{21}}$$

pour $i, k = 1, 2$.

Enfin, nous avons d'après (23)

$$\frac{C}{2} = \sum_{k=1}^2 (\alpha_k - a)^2 P_k + 2 \sum_{i=1}^2 \sum_{\kappa=1}^2 (-1)^{i+k} (\alpha_i - a) (\alpha_k - a) P_k P_{3-k} \frac{p_{11} - p_{21}}{1 - p_{11} + p_{21}}.$$

c. Supposons que $\alpha_1 = 1$, $\alpha_2 = 0$, d'où il résulte que $a = P_1$. Dans ce cas, la formule précédente donne

$$(23 \text{ bis}) \quad \frac{C}{2} = \lim_{n \rightarrow \infty} E \frac{(m - n P_1)^2}{n} = P_1(1 - P_1) \frac{1 + p_{11} - p_{21}}{1 - p_{11} + p_{21}},$$

où $m = x^{(1)} + x^{(2)} + \dots + x^{(n)}$ est le nombre d'expériences ayant amené le résultat E_1 , et n le nombre total d'expériences.

On peut résumer cette formule comme suit : On fait des expériences successives, chacune comportant deux éventualités possibles A_1 et A_2 ; on convient de dire qu'une expérience a réussi, si elle a donné A_1 comme résultat. La probabilité pour que la $N^{\text{ième}}$ expérience réussisse est égale à peu près à P_1 , si le nombre N est grand. Dans une série de n expériences, la valeur moyenne du nombre d'expériences réussies est à peu près égale à $n P_1$, si n est assez grand.

Le facteur $\frac{1 + p_{11} - p_{21}}{1 - p_{11} + p_{21}}$ qui figure dans (23 bis) est dû à la dépendance particulière qui se présente dans le cas d'une chaîne. Si les expériences étaient indépendantes l'une de l'autre, il n'y aurait pas de chaîne proprement dite; on aurait dans ce cas $p_{11} = p_{21} = P_1$, et (23 bis) se réduirait alors à

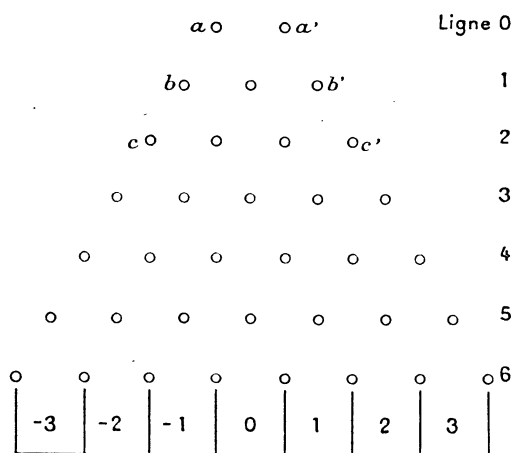
$$(23 \text{ ter}) \quad E \frac{(m - n P_1)^2}{n} = P_1(1 - P_1),$$

formule connue, valable pour toute valeur de n , de sorte que le signe \lim qui figure dans (23 bis) devient superflu.

8. Application à la théorie de l'appareil de Galton. — *a.* Une bille roule sur un plan incliné; elle passe entre des clous disposés de manière à former un réseau triangulaire régulier (*voir* la figure) ou un réseau composé de triangles isocèles. Les lignes aa' , bb' , ... sont horizontales. Le diamètre de la bille est plus petit que la distance de deux clous voisins; elle passe d'abord entre a et a' , elle traverse ensuite la ligne bb' , puis cc' ... Supposons que la ligne horizontale aa' ait l'indice zéro, que bb' ait l'indice 1, ..., et que l'indice de la dernière

ligne horizontale soit égal à un nombre pair $n = 2v$. Après avoir traversé la dernière ligne horizontale, la bille tombe dans un des $n + 1$ compartiments situés au-dessous de la $n^{\text{ième}}$ ligne des clous. Attachons le nombre zéro au compartiment du milieu; les numéros des compartiments qui sont à gauche (ou à droite) du compartiment zéro sont numérotés $-1, -2, \dots, -v$ (ou $1, 2, \dots, v$); (voir la figure où $n = 2, v = 6$). Si le nombre de passages à droite est égal à celui de passages à gauche ($= v$), la bille tombe dans le compartiment zéro. Si le nombre de passages à droite est $v \pm h$, celui de passages à gauche est égal à $v \mp h$, et la bille s'arrête au compartiment $\pm h$. On déduit expérimentalement la valeur moyenne de $\frac{h^2}{2v}$ de la façon suivante : on fait rouler un grand nombre N de billes; elles se répartissent de telle sorte qu'il y en ait N_h

Fig. 1.



dans le compartiment h . La valeur expérimentale de la moyenne cherchée sera

$$E \frac{h^2}{2v} = \frac{1}{2v \cdot N} \sum_{h=-v}^{+v} h^2 N_h.$$

Cherchons maintenant à obtenir par le calcul la valeur moyenne de $\frac{h^2}{2v}$.

b. Une première solution est la suivante : Admettons que, si la bille

tombe contre un clou, il y a toujours une probabilité constante $\frac{1}{2}$ pour qu'elle passe à droite et la même probabilité pour qu'elle passe à gauche. Regardons le passage à droite comme une expérience qui a réussi; si m est le nombre de ces passages et h le numéro du compartiment où la bille arrive, nous aurons

$$P_1 = \frac{1}{2}, \quad n P_1 = \nu, \quad h = m - \nu,$$

et la formule (23 *ter*) donne

$$E \frac{h^2}{2\nu} = \frac{1}{4}.$$

Or, les expériences dont nous parlerons plus loin n'ont pas confirmé ce résultat; la valeur expérimentale déterminée plus haut était très différente de $\frac{1}{4}$.

c. Voici la *seconde solution* : Admettons que, si la bille passe à droite, il y ait une probabilité p_{11} pour qu'elle passe encore à droite en traversant la ligne qui suit immédiatement, et que si elle a passé à gauche il y ait une probabilité p_{21} pour qu'elle passe ensuite à droite. Les symboles p_{22} et p_{12} ont une signification analogue, l'indice 1 (ou 2) se rapportant au passage à droite (ou à gauche). Nous supposons, à cause de la symétrie de l'appareil, que $p_{21} = p_{12}$. Il en résulte (*voir 7 b*) que

$$P_1 = P_2 = \frac{1}{2},$$

c'est-à-dire qu'après un grand nombre de passages successifs, la probabilité de passer à gauche sera égale à celle de passer à droite. Donc le nombre moyen de m sera égal à $P_i 2\nu = \nu$, et la limite de la valeur moyenne de

$$\frac{h^2}{n} = \frac{(m - \nu)^2}{2\nu}$$

sera cette fois déterminée par la formule (23 *bis*). L'emploi de cette formule suppose que le nombre de lignes horizontales $n = 2\nu$ est assez grand. Les probabilités p_{11} et p_{21} se déduisent par des observations statistiques sur le mouvement des billes. Par exemple, p_{11} sera égal au nombre des cas dans lesquels une bille a passé à droite et où ce passage a été immédiatement suivi par un autre à droite, ce nombre étant ensuite divisé par le nombre total de passages à droite.

Cette seconde solution a été proposée en 1934 et contrôlée expérimentalement par deux auteurs. M. Schulz [18] (qui travaillait avec un appareil à 60 lignes horizontales) et M. Münzner [19] ont trouvé que $p_{11} = 0,75$ environ, donc $p_{12} = p_{21} = 0,25$, et

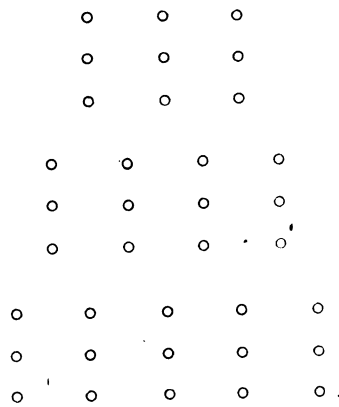
$$P_1(1 - P_1) \frac{1 + p_{11} - p_{21}}{1 - p_{11} + p_{21}} \cong \frac{3}{4}.$$

Si donc on fait rouler N billes et si N_h billes se trouvent en fin de compte dans le compartiment h , il faut que l'on ait approximativement

$$\frac{1}{2 \nu \cdot N} \sum_{h=-\nu}^{+\nu} N_h \cdot h^2 = \frac{3}{4}.$$

Les expériences faites par les auteurs cités ont confirmé ce résultat.

Fig. 2.



C'est donc bien la théorie des chaînes qu'il faut appliquer ici pour obtenir un résultat correct.

En même temps, M. Seitz et M. Hamacher-Odenhausen [20] ont proposé une modification de l'appareil de Galton, de façon à en réaliser un autre dans lequel les statistiques concernant le mouvement des billes correspondent à la première solution donnée plus haut. Pour cela ils remplacent chaque clou par trois autres (par un triplet) séparés par une distance de 5^{mm} , 5 (voir la figure). La bille, de diamètre 9^{mm} , 5, passe par un canal long de 11^{mm} et large de 10^{mm} ; elle perd ainsi la composante hori-

zontale de sa vitesse, ainsi que sa rotation. La répartition des billes entre les divers compartiments correspond dans ce cas à la première solution, en d'autres termes on a

$$\frac{1}{2 \nu N} \sum_{h=-\nu}^{+\nu} N_h h^2 = \frac{1}{4},$$

ce qui a été confirmé d'après le travail cité par des expériences effectuées avec un grand nombre N de boules (N = 9320, N = 9026).

9. Moyennes arithmétiques Π_{ik} . — a. Si les quantités $p_{ik}(i, k = 1, 2, \dots, r)$ vérifient les conditions (2) et (3), les limites

$$\Pi_{ik} = \lim_{n \rightarrow \infty} \frac{P_{ik}^{(1)} + P_{ik}^{(2)} + \dots + P_{ik}^{(n)}}{n}$$

existent toujours (voir Fréchet [1]), et nous avons

$$\sum_{k=1}^r \Pi_{ik} = 1 \quad (i = 1, 2, \dots, r).$$

Si la limite $\lim_{n \rightarrow \infty} P_{ik}^{(n)}$ existe, elle coïncide évidemment avec Π_{ik} . Si elle n'existe pas, le théorème de M. Fréchet que nous venons d'énoncer nous renseigne sur l'allure générale des oscillations de $P_{ik}^{(n)}$ lorsque n augmente. Ces oscillations doivent présenter une certaine régularité (voir, pour des indications plus précises, le travail cité de M. Fréchet), sans quoi la limite Π_{ik} n'existerait pas.

b. Les moyennes Π_{ik} donnent la solution des équations (19) avec la condition (19 bis), quelles que soient les p_{ik} pourvu qu'elles satisfassent aux conditions (2) et (3). En effet, si les a_s sont des constantes telles que

$$a_s \geq 0, \quad a_1 + a_2 + \dots + a_r = 1,$$

nous aurons

$$\sum_{j=1}^r \left[\sum_{s=1}^r a_s \frac{\sum_{m=1}^n P_{sj}^{(m)}}{n} \right] p_{jk} = \sum_{s=1}^r a_s \frac{\sum_{m=1}^n \sum_{j=1}^r P_{sj}^{(m)} p_{jk}}{n} = \sum_{s=1}^r a_s \frac{\sum_{m=1}^n P_{sk}^{(m+1)}}{n}.$$

Lorsque n augmente indéfiniment, ces formules donnent

$$\sum_{i=1}^r \left(\sum_{s=1}^r a_s \Pi_{sj} \right) P_{ik} = \sum_{s=1}^r a_s \Pi_{sk},$$

donc, si a_1, a_2, \dots, a_r sont des constantes non négatives dont la somme est égale à l'unité, l'expression $\sum_{s=1}^r a_s \Pi_{sk}$ substituée à P_k satisfait aux équations (19) et (19 bis) ⁽¹⁾; de plus, elle n'est pas négative. Si la limite $\lim_{n \rightarrow \infty} P_{ik}^{(n)} = P_k$ existe et si elle ne dépend pas de i , on aura $\Pi_{sk} = P_k$, et l'expression précédente ne contiendra aucune constante arbitraire.

CHAPITRE II.

CHAINES SIMPLES. GÉNÉRALISATIONS.

10. Chaîne simple à éléments variables. — Les probabilités absolues $p_k^{(n)}$ ($k = 1, 2, \dots, r, n = 1, 2, 3, \dots$) étant définies comme au n° 1, admettons, au lieu de l'hypothèse 2° que, si la $(n - 1)^{\text{ième}}$ expérience a donné E_i , il y a une probabilité déterminée $p_{ik}^{(n)}$ pour que la $n^{\text{ième}}$ expérience donne E_k . Les probabilités de passage $p_{ik}^{(n)}$ dépendent donc dans ce cas de i , de k et de n , et elles doivent satisfaire aux conditions

$$p_{ik}^{(n)} \geq 0 \quad (i, k = 1, 2, \dots, r; n = 1, 2, 3, \dots),$$

$$\sum_{k=1}^r p_{ik}^{(n)} = 1 \quad (i = 1, 2, \dots, r; n = 1, 2, 3, \dots).$$

L'ensemble de quantités $p_{ik}^{(n)}$ définit une chaîne simple à éléments variables et à r éventualités; on suppose (comme au n° 1) que les $p_{ik}^{(n)}$ ne changent pas, lorsqu'on connaît après coup les résultats d'autres expériences.

Donnons un exemple simple. Les probabilités $p_{ik}^{(n)}$ dépendant de n

(1) Cela résulte des formules contenues dans une lettre de M. Kolmogoroff (du 2 juillet 1936) et d'une lettre de M. Fréchet (du 10 octobre 1936); ces lettres se rapportent à un autre sujet (voir n° 20).

pourraient s'introduire dans l'étude de l'appareil de Galton (*voir* n° 8). Nous avons admis au n° 8c que la probabilité de passer à droite, si le passage précédent était à droite, ne dépend pas de l'indice de la ligne horizontale où passe la bille. Mais on peut introduire l'hypothèse plus générale suivante : la probabilité pour que la bille conserve sa direction dépend de la hauteur de la ligne horizontale. La vitesse du mouvement augmente en descendant; si donc on faisait les statistiques d'une manière plus détaillée, on trouverait probablement que la tendance de conserver la direction augmente en descendant (c'est-à-dire la valeur de $p_{11}^{(n)}$ augmente avec n , $p_{11}^{(n)}$ étant la probabilité pour que la bille passe à droite en traversant la $n^{\text{ième}}$ ligne, si elle a passé à droite en traversant la ligne précédente).

Revenons au cas général à r éventualités. Les probabilités relatives à plusieurs passages successifs dépendent ici de quatre indices i, k, m, n . En effet, la probabilité pour que la $(m+n)^{\text{ième}}$ expérience donne E_k si la $m^{\text{ième}}$ a donné E_i est égale à

$$P_{ik}^{(m,n)} = \sum_{\alpha, \beta, \dots, \lambda} p_{i\alpha}^{(m+1)} p_{\alpha\beta}^{(m+2)} \dots p_{\lambda k}^{(m+n)},$$

la somme étant étendue à $(n+1)$ indices $\alpha, \beta, \dots, \lambda$ dont chacun varie de 1 à r .

11. Chaîne simple à une infinité d'éventualités et à éléments constants.

— Supposons qu'une expérience admette comme résultat un des phénomènes en nombre infini E_1, E_2, \dots, E_n . Admettons les hypothèses du n° 1 (à cela près que $r = \infty$, de sorte que (3) est maintenant une série infinie); les quantités $p_{ik}(i, k = 1, 2, 3, \dots)$ en nombre infini définissent alors une *chaîne simple à une infinité d'éventualités*, et les formules relatives aux probabilités des expériences répétées s'écrivent

$$P_{ik}^{(n)} \geq 0, \quad \sum_{k=1}^{\infty} P_{ik}^{(n)} = 1,$$

$$P_{ik}^{(m+n)} = \sum_{j=1}^{\infty} P_{ij}^{(m)} P_{jk}^{(n)},$$

où $i, k, m, n = 1, 2, 3, \dots$ avec $P_{ik}^{(0)} = p_{ik}$.

C'est à M. Fortet [21] que nous devons le commencement de l'étude

de ces chaînes. Pour étudier le comportement de $P_{ik}^{(n)}$ lorsque n croît indéfiniment, Fortet introduit les bornes supérieures P_k^n et les bornes inférieures p_k^n des $P_{ik}^{(n)}$ lorsque i varie; on a

$$0 \leq p_k^{n-1} \leq p_k^n \leq P_{ik}^{(n)} \leq P_k^n \leq P_k^{n-1} \leq 1.$$

Donc, p_k^n a une limite p_k , P_k^n a une limite $P_k \geq p_k$, et la série $\sum_{k=1}^{\infty} P_k^n$ converge, sa somme étant au plus égale à 1; $\sum_{k=1}^{\infty} p_k$ converge également et l'on a

$$\sum_{\kappa} p_k = \lim_{n \rightarrow \infty} \sum_k p_k^n \quad \text{donc} \quad 0 \leq \sum_{\kappa} p_k \leq 1.$$

En partant de ces résultats, M. Fortet donne des théorèmes analogues à ceux que nous avons exposés au n° 4, et qui se rapportent au cas régulier et au cas semi-régulier.

Remarquons que le cas suivant peut encore se présenter : chaque expérience ne peut donner qu'un nombre limité de résultats; si une première expérience donne A_i , la seconde peut donner A_k , la troisième A_l , ..., de sorte que, si le nombre d'expériences successives augmente indéfiniment, l'ensemble de toutes les éventualités qui peuvent se présenter comme résultats d'une série illimitée d'expériences, sera infini. Dans un tel cas, la série $\sum_k p_{ik}$ se réduit à la somme d'un nombre fini de termes p_i , mais l'ensemble de ces termes dépend de i .

Nous donnons au n° 12 un exemple de ce genre, où toutefois les éléments de la chaîne dépendent de n .

Dans un travail que je n'ai pu encore analyser en détail, M. Kolmogoroff [22] expose des résultats remarquables sur les chaînes infinies. Il démontre, en particulier, que les limites

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{m=1}^n P_{ik}^{(m)}$$

existent toujours.

12. Chaîne simple à une infinité d'éventualités et à éléments variables.

— *Exemples* : La $(n - 1)^{\text{ième}}$ expérience ayant amené le résultat E_i , désignons par $p_{ik}^{(n)}$ la probabilité supposée fonction de n , d'obtenir le

résultat E_k dans la $n^{\text{ième}}$ expérience. Si les résultats possibles E_1, E_2, \dots sont en nombre infini, les $p_{ik}^{(n)}(i, k, n = 1, 2, 3, \dots)$ définiront une chaîne simple à une infinité d'éventualités et à éléments variables.

Les probabilités relatives aux tirages d'une urne dont on change la composition après chaque tirage peuvent servir comme exemple; le changement de composition de l'urne après le $n^{\text{ième}}$ tirage doit dépendre du résultat de ce tirage; le nombre d'éventualités étant fini pour un seul tirage, il ne devient infini que si le nombre de tirages successifs augmente indéfiniment.

Supposons qu'une urne contienne a boules blanches et b noires avant le premier tirage. Le $n^{\text{ième}}$ tirage consiste à extraire une boule de l'urne et à l'y déposer à nouveau en y ajoutant $f_1(n)$ blanches et $g_1(n)$ noires si la boule extraite est blanche ou $f_2(n)$ blanches et $g_2(n)$ noires si c'est une noire. Supposons que l'on ait

$$f_1(n) + g_1(n) = f_2(n) + g_2(n).$$

Le nombre total de boules contenues dans l'urne augmente ainsi, par le $n^{\text{ième}}$ tirage seul, de $f_1(n) + g_1(n)$ unités; il est donc égal, après le $n^{\text{ième}}$ tirage, à

$$a + b + \sum_{s=1}^n [f_1(s) + g_1(s)].$$

Soit i le nombre de boules blanches qui se trouvent dans l'urne après le $(n-1)^{\text{ième}}$ tirage, et soit $p_{ik}^{(n)}$ la probabilité pour qu'il y ait dans l'urne k boules blanches après le changement de composition qui résulte du $n^{\text{ième}}$ tirage. Les $p_{ik}^{(n)}$ satisfont à l'équation

$$(24) \quad p_{i1}^{(n)} + p_{i2}^{(n)} + \dots + p_{ik}^{(n)} + \dots = 1;$$

pour qu'une probabilité $p_{ik}^{(n)}$ soit différente de zéro, il faut : 1° que i soit au plus égal au maximum de boules blanches contenues dans l'urne après le $(n-1)^{\text{ième}}$ tirage; ce maximum est inférieur à

$$a + b + \sum_{s=1}^{n-1} [f_1(s) + g_1(s)];$$

2° que la différence $k - i$ soit égale à $f_1(n)$ ou à $f_2(n)$. Nous avons

pour ces deux cas :

$$(25) \quad \left\{ \begin{array}{l} P_{ik}^{(n)} = \frac{i}{a + b + \sum_{s=1}^{n-1} [f_1(s) + g_1(s)]} \quad \text{si } k - i = f_1(n), \\ P_{ik}^{(n)} = 1 - \frac{i}{a + b + \sum_{s=1}^{n-1} [f_1(s) + g_1(s)]} \quad \text{si } k - i = f_2(n). \end{array} \right.$$

Pour une valeur fixe de i la série qui figure au premier membre de (24) n'a que deux termes différents de zéro; ils sont donnés par les formules (25). Ces formules servent à étudier les probabilités relatives aux variations de composition de l'urne; on en peut déduire les probabilités relatives à l'extraction d'une boule blanche.

En 1907, déjà Markoff, dans son premier travail sur les chaînes, a donné un exemple de tirages successifs pour montrer qu'il y a des cas dans lesquels la loi des grands nombres ne s'applique pas. L'exemple de Markoff [qui correspond, dans la notation que nous avons adoptée, à $a = b = 1, f_1(n) = f_2(n) = 1, g_1(n) = g_2(n) = 0$] est le suivant : Une urne contient deux boules, une blanche et une noire. On tire une boule et on la remet dans l'urne en y déposant en même temps une autre boule de la même couleur. Après n tirages successifs, il y a dans l'urne $2 + n$ boules; $n + 1$ cas sont possibles, le nombre total de boules blanches extraites dans ces n tirages pouvant être égal à l'un des nombres $0, 1, 2, \dots, n$. Tous ces cas ont la même probabilité $\frac{1}{n + 1}$.

Si l'on faisait les tirages à la manière ordinaire, en remettant toujours la boule extraite sans changer la composition de l'urne, les probabilités relatives aux différents cas (nombre de boules blanches extraites dans une série de n tirages = $0, 1, 2, \dots, n$) auraient un maximum, correspondant au nombre moyen de boules blanches extraites; la loi des grands nombres s'appliquerait.

(Voir à ce sujet ma Note [23], ainsi que les travaux de MM. Eggenberger-Pólya [24], et de MM. Onicescu-Mihoc [25].)

Le cas étudié d'abord par Markoff et par MM. Eggenberger-Pólya, où $f_1(n) = g_2(n)$ est une constante Δ et où $f_2(n) = g_1(n) = 0$, a formé ces dernières années l'objet de nombreux travaux. Un grand nombre de

questions concernant les probabilités et les valeurs moyennes qui se posent dans ce domaine ont été résolues par des formules simples.

Dans une note récente [27], M. Kaucký a montré comment on peut étendre à ce cas un résultat connu concernant la répétition des phénomènes indépendants. Whitworth ([26], p. 174) a donné la solution du problème suivant : On fait des tirages successifs d'une urne où se trouvent des boules, blanches et noires, en remettant toujours la boule extraite; quelle est la probabilité pour que, sur n tirages successifs, on obtienne, une fois au moins, k tirages successifs amenant chacun une boule blanche? Whitworth montre que cette probabilité est égale au coefficient de x^n dans le développement de la fonction génératrice

$$\frac{1 - \rho^k x^k}{1 - x + (1 - \rho)\rho^k x^{k+1}}$$

suivant les puissances de x ; ρ représente la probabilité de tirer une boule blanche de l'urne. Dans la note citée, M. Kaucký donne l'expression de la fonction génératrice qui permet de résoudre le problème de Whitworth sur les répétitions dans le cas où

$$f_1(n) = g_2(n) = \Delta, \quad f_2(n) = g_1(n) = 0;$$

cette expression contient une intégrale définie qui dépend de la composition initiale de l'urne, du nombre Δ , de n , de k et de la variable x .

13. Probabilités des phénomènes qui dépendent d'autres phénomènes liés en chaîne simple. — Considérons une série d'expériences (que nous nommons *expériences de première espèce*) où les probabilités sont liées en chaîne simple à r éventualités E_1, E_2, \dots, E_r et à éléments constants; nous conservons les notations introduites au n° 1. Soient F_1, F_2, \dots, F_s autres phénomènes qui se présentent comme résultats d'autres expériences (*expériences de seconde espèce*). Nous admettons l'hypothèse suivante : si, dans la suite d'expériences de première espèce, le phénomène E_i apparaît comme résultat de la $n^{\text{ième}}$ expérience, il y aura une probabilité ρ_{ik} pour que la $n^{\text{ième}}$ expérience de seconde espèce donne F_k . La probabilité pour que, dans la suite des résultats fournis par les expériences successives de seconde espèce, F_k se présente à la $n^{\text{ième}}$ place, est égale à $R_{ik}^{(n)}$ où

$$(26) \quad R_{ik}^{(n)} = \sum_{u=1}^r P_{iu}^{(n)} \rho_{uk} \quad (i, u = 1, 2, \dots, r, k = 1, 2, \dots, s);$$

nous supposons que E_i se présente comme résultat de l'expérience de première espèce d'indice zéro.

Supposons maintenant que le *cas régulier* se présente pour les expériences de première espèce (voir n° 4). Les $P_{ik}^{(n)}$ admettent alors, pour n infini, des limites P_k indépendantes de i ; par conséquent, les $R_{ik}^{(n)}$ tendent en même temps vers des valeurs limites R_k indépendantes de i , et la formule (26) donne

$$(26 \text{ bis}) \quad R_k = \lim_{n \rightarrow \infty} R_{ik}^{(n)} = \sum_{u=1}^r P_u \rho_{uk} \quad (k = 1, 2, \dots, s).$$

Les probabilités relatives aux résultats des expériences de seconde espèce ne sont pas liées entre elles en chaîne proprement dite.

Introduisons la représentation géométrique suivante : r points fixes A_1, A_2, \dots, A_r étant donnés (comme au n° 4), la position d'un point mobile M sera déterminée par le résultat des expériences du premier type, de façon que si l'expérience donne E_i , M se trouvera en A_i . De même, s points fixes B_1, B_2, \dots, B_s étant donnés, la position d'un autre point mobile N sera déterminée par les expériences du second type, N se trouvant en B_k lorsque l'expérience aura donné le résultat F_k .

Supposons qu'un observateur ne voie que le mouvement du point N , les positions successives du point M lui étant inconnues. Les changements de position de N seront régies par des lois statistiques différentes de celles qui gouvernent le point M ; ces dernières se déduisent directement des formules relatives à une chaîne simple.

Afin d'analyser de plus près cette différence, attachons à chaque position possible B_k ($k = 1, 2, \dots, s$) du point N un coefficient α_k , et désignons par $x^{(n)}$ la variable aléatoire attachée à N ; elle devient égale à α_k , si le point N vient en B_k après la $n^{\text{ième}}$ expérience de seconde espèce. La valeur moyenne de $x^{(n)}$ sera égale à (voir [26])

$$(27) \quad E(x^{(n)}) = \sum_{\kappa=1}^s R_{i\kappa}^{(n)} \alpha_\kappa = \sum_{u=1}^r \sum_{\kappa=1}^s P_{i\kappa}^{(n)} \rho_{u\kappa} \alpha_\kappa,$$

toujours dans l'hypothèse que M était en A_i avant la première expérience de première espèce. Après un nombre infini d'expériences successives, cette valeur moyenne est égale à

$$(27 \text{ bis}) \quad \alpha = \sum_{\kappa=1}^s R_\kappa \alpha_\kappa = \sum_{\kappa=1}^s \sum_{u=1}^r P_u \rho_{u\kappa} \alpha_\kappa.$$

Markoff [28] a démontré (toujours dans l'hypothèse que le mouvement de M correspond au cas régulier d'une chaîne simple) que la valeur moyenne suivante :

$$(28) \quad \frac{1}{n} [x^{(1)} - a + x^{(2)} - a + \dots + x^{(n)} - a]^2$$

tend vers une limite déterminée lorsque n augmente indéfiniment, et il a calculé cette valeur limite, dans quelques cas particuliers, au moyen des fonctions génératrices. Dans une note publiée en 1936 dans les *Comptes rendus* [29], j'ai fait l'observation suivante : pour simplifier l'expression que Markoff donne de la limite de (28), on pourrait appliquer la méthode qui consiste à remplacer la formule de Markoff, relative à la dispersion dans le cas régulier d'une chaîne simple, par une formule moins compliquée (voir n° 7). Peu de temps après cette observation, M. Doeblin m'a communiqué la solution suivante qui donne une expression indépendante des fonctions génératrices de la limite cherchée de (28).

Soit $P_{ijkl}^{(n)} = p_{ijkl}$ la probabilité pour que le couple de points M, N passe, par une seule transition, de la position A_i, B_j à A_k, B_l . Si $P_{ijkl}^{(n)}$ désigne la probabilité pour que le couple, supposé être d'abord en A_i, B_j , passe par n transitions successives en A_k, B_l , nous aurons

$$P_{ijkl}^{(n)} = P_{ik}^{(n)} \rho_{kl}$$

et

$$\lim_{n \rightarrow \infty} P_{ijkl}^{(n)} = P_k \rho_{kl} \quad \text{avec} \quad P_k = \lim_{n \rightarrow \infty} P_{ik}^{(n)}.$$

Les changements de position que subit le couple M, N sont liés en chaîne simple. Dans l'hypothèse que MN soit initialement en $A_i B_j$ la quantité aléatoire $x^{(n)}$ attachée à la $n^{\text{ième}}$ position du point N admet comme valeur moyenne la valeur $E(x^n)$ donnée par (27); et sa valeur limite pour n infini, égale à a , est donnée par (27 bis). La formule (23) du n° 7 peut être appliquée au cas du couple MN à condition d'y remplacer α_i par α_l , α_k par α_u , a par (27 bis), $P_{ik}^{(n)}$ par $P_{ik}^{(n)} \rho_{kl}$, P_k par $P_k \rho_{kl}$ et s_{ik} par

$$\sum_{n=1}^{\infty} (P_{ijku}^{(n)} - P_k \rho_{ku}) = \sum_{n=1}^{\infty} \rho_{ku} (P_{ik}^{(n)} - P_k);$$

il faut ensuite substituer aux sommations par rapport à i, k , quatre

autres sommations, i et k variant chacun de 1 à r , l et u de 1 à s . Nous arrivons ainsi à la formule suivante démontrée par M. Doeblin :

$$\begin{aligned}
 (29) \quad \frac{C}{2} &= \lim_{n \rightarrow \infty} E \frac{[x^{(1)} - a + x^{(2)} - a + \dots + x^{(n)} - a]^2}{n} \\
 &= \sum_{k=1}^r \sum_{l=1}^s (\alpha_l - a)^2 P_{k\rho_{kl}} \\
 &\quad + 2 \sum_{i=1}^r \sum_{k=1}^r \sum_{l=1}^s \sum_{u=1}^s (\alpha_k - a)(\alpha_l - a) P_{k\rho_{kl}\rho_{lu}} \sum_{n=1}^{\infty} (P_{ki}^{(n)} - P_i).
 \end{aligned}$$

14. Dispersion dans le problème de deux urnes avec échange de boules. Fluctuations. — α . Revenons au problème de deux urnes du n° 5. Supposons qu'il reste i boules blanches dans l'urne

$$A(i = 0, 1, \dots, e - 1)$$

après en avoir tiré une boule et avant d'y avoir remis la boule tirée de B; et soit j un nombre tel que $j = 1$ ou $j = 0$, suivant que la boule sortie de A est blanche ou noire.

$p_{ij;kl}$ désignera la probabilité pour : 1° qu'il y ait, en A, k boules blanches ($k = 0, 1, \dots, e - 1$) après l'échange des boules qui viennent d'être extraites de A et de B, après avoir tiré de nouveau une boule de A et avant d'y déposer la boule extraite de B, et 2° qu'à ce tirage de A il sorte une boule blanche, si $l = 1$, ou noire si $l = 0$.

Ces probabilités $p_{ij;kl}$ ($i, k = 0, 1, \dots, e - 1; j, l = 0, 1$) peuvent être regardées comme probabilités de passage définissant une chaîne simple; l'état du système est donné si l'on connaît, avant que la boule extraite de B soit déposée en A, la couleur de la boule extraite en A et le nombre de boules blanches qui y restent après ce tirage. Les probabilités $p_{ij;kl}$ sont faciles à calculer; les probabilités itérées $P_{ij,kl}^{(n)}$ déterminées par la formule

$$P_{ij,kl}^{(n+1)} = \sum_{s=0}^{e-1} \sum_{t=0}^1 P_{ij;s,t}^{(n)} p_{s,t;kl}$$

sont toutes positives, si n est suffisamment grand, et l'on a pour toute valeur de n

$$\begin{aligned}
 \sum_{k=0}^{e-1} \sum_{l=0}^1 P_{ij,kl}^{(n)} &= \sum_{k=0}^{e-1} \sum_{l=0}^1 p_{ij;kl} = 1, \\
 \lim_{n \rightarrow \infty} P_{ij,kl}^{(n)} &= P_{kl} \quad (k = 0, 1, \dots, e - 1; l = 0, 1).
 \end{aligned}$$

Pour calculer cette valeur limite nous ferons usage des probabilités $P_{ik}^{(n)}$ introduites au n° 5; nous avons

$$P_{k1} = \lim_{n \rightarrow \infty} P_{i, k+1}^{(n)} \cdot \frac{k+1}{e} = P_{k+1} \cdot \frac{k+1}{e},$$

$$P_{k0} = \lim_{n \rightarrow \infty} P_{i, k}^{(n)} \cdot \frac{e-k}{e} = P_k \cdot \frac{e-k}{e};$$

où (voir n° 5)

$$P_k = \binom{e}{e-k} \binom{e}{k} : \binom{2e}{e}.$$

Attachons à la boule tirée de A au $n^{\text{ième}}$ tirage la variable aléatoire $x^{(n)}$ égale à 1 ou 0, selon que la boule extraite est blanche ou noire. Il vient

$$\lim_{n \rightarrow \infty} E x^{(n)} = \sum_{k=0}^{e-1} P_{k+1} \frac{k+1}{e} = \frac{1}{2}.$$

Si m est le nombre de boules blanches extraites dans une série de n tirages, cherchons la dispersion

$$\frac{C}{2} = \lim_{n \rightarrow \infty} E \frac{\left[x^{(1)} + x^{(2)} + \dots + x^{(n)} - \frac{n}{2} \right]^2}{n} = \lim_{n \rightarrow \infty} E \frac{\left(m - \frac{n}{2} \right)^2}{n}.$$

Elle peut être calculée comme suit :

Si l'on y remplace dans (23) chaque somme simple par une somme double et chaque probabilité à un ou deux indices par une autre à deux ou quatre, on aura

$$(29 \text{ bis}) \quad \frac{C}{2} = \sum_{k,l} (\alpha_{kl} - \alpha)^2 P_{kl} + 2 \sum_{i,j} \sum_{k,l} (\alpha_{ij} - \alpha)(\alpha_{kl} - \alpha) P_{kl} s_{kl ij}$$

avec

$$s_{kl ij} = \sum_{n=1}^{\infty} (P_{kl ij}^{(n)} - P_{ij}).$$

Pour obtenir la dispersion cherchée, introduisons dans cette formule les probabilités $p_{ij,kl}$ définies plus haut, ainsi que les $P_{ij,kl}^{(n)}$ et les P_{kl} qui s'en déduisent; la quantité α sera remplacée par $\frac{1}{2}$, et $\alpha_{i1} = 1$, $\alpha_{i0} = 0$; les indices i et h varient de 0 à $e-1$, j et l de 0 à 1.

On trouve ainsi :

$$\frac{C}{2} = \frac{1}{12}, \frac{1}{10}, \frac{3}{28} \text{ respectivement pour } e = 2, 3, 4; \text{ ce résultat est en}$$

accord avec la formule suivante

$$\frac{C}{2} = \frac{1}{4} \frac{e-1}{2e-1};$$

cette formule a été donnée par Markoff, en 1912, d'abord sans démonstration (1); dans un travail publié en 1915, il a démontré la formule plus générale suivante :

$$\frac{C}{2} = p(1-p) \left\{ 1 - \frac{2(a+b)c+d}{(a+b+c+d)(a+b+c+d-1)} \right\}$$

avec

$$p = \frac{a+c}{a+b+c+d}$$

où a (ou c) est le nombre de boules blanches dans la première (ou seconde) urne et b (ou d) le nombre de boules noires dans la première (ou la seconde) urne. On retrouve la formule précédente en posant $a+b=c+d=e$, voir [28].

b. L'exemple précédent nous indique de quelle manière on doit appliquer la théorie des variables aléatoires liées en chaîne à l'étude des fluctuations. Supposons qu'un gaz soit *enfermé* dans un vase et cherchons le nombre de molécules du gaz qui se trouvent dans une partie v_1 de ce vase. Ce nombre varie au cours du temps (on dit : *il y a fluctuation*), et il faut évaluer le carré de la différence entre le nombre réel de molécules dans v_1 et entre le nombre moyen. Si l'on ne s'occupe pas du mécanisme du mouvement des molécules (qui est cependant la vraie cause des fluctuations), on se contente de la solution suivante donnée par les for-

(1) Remarquons que la formule (29) du n° 13 sert à calculer la dispersion du nombre total m de boules blanches extraites de A dans les conditions suivantes : On tire de A une boule U_1 et on la remet dans A; ensuite on opère le tirage avec échange de boules comme au n° 5; puis on tire de A une boule U_2 , on l'y remet et l'on opère le tirage avec échange de boules, et ainsi de suite. La dispersion relative au nombre de boules blanches qui se trouvent dans la série U_1, U_2, \dots, U_n s'obtient en substituant (29)

$$\alpha_1 = 1, \quad \alpha_2 = 0, \quad C_{11} = \frac{i}{e}, \quad C_{12} = \frac{e-i}{e}, \quad \alpha = \frac{1}{2}.$$

On trouve ainsi

$$\frac{C}{2} = \frac{1}{4}, \frac{3}{10}, \frac{9}{28} \text{ respectivement pour } e = 2, 3, 4.$$

On voit que ces valeurs diffèrent de celles qu'on trouve dans le cas du problème considéré dans le texte (n° 14a).

mules classiques du Calcul des probabilités; on distribue un certain nombre de molécules entre des volumes égaux v_1, v_2, \dots ; il y a toujours un nombre moyen de molécules qui se trouvent dans un quelconque de ces volumes v_i , et l'on calcule la valeur moyenne du carré de la différence entre le nombre réel et entre ce nombre moyen, comme si la distribution actuelle n'avait aucune influence sur la distribution à un instant postérieur. Si, au contraire, on tient compte de cette influence, si l'on adopte une loi du mélange progressif, les valeurs moyennes en question doivent être calculées par des formules plus compliquées. La différence entre ces deux points de vue est assez grande. Nous avons vu, dans l'exemple précédent, combien la valeur de $\frac{C}{2}$, calculée au moyen de (29), diffère de celle qui est donnée par les formules classiques; en établissant la formule (29) nous avons tenu compte du mécanisme particulier qui fait changer le nombre de boules contenues dans A à chaque tirage, tandis que l'emploi des formules classiques ne tient pas compte de l'influence que peut avoir l'extraction d'une boule blanche dans un tirage déterminé sur l'extraction d'une boule de même couleur au tirage suivant.

CHAPITRE III.

CHAINES MULTIPLES.

SUITES STATIONNAIRES DE VARIABLES ALÉATOIRES.

15. Définition générale de la chaîne multiple à éléments variables. —

a. Soit

$$(30) \quad \dots, x^{(-n)}, x^{(-n+1)}, \dots, x^{(0)}, x^{(1)}, x^{(2)}, \dots, x^{(n)}, \dots$$

une suite illimitée de quantités aléatoires variables. Nous supposons :

1° que chaque variable $X^{(n)}$ prend une des valeurs $\alpha_1, \alpha_2, \dots, \alpha_r$ et qu'il y ait une probabilité déterminée $p_k^{(n)}$ (*probabilité absolue*) pour que l'on ait $X^{(n)} = \alpha_k$, les valeurs des autres variables restant inconnues; nous avons pour n quelconque

$$p_k^{(n)} \geq 0, \quad p_1^{(n)} + p_2^{(n)} + \dots + p_r^{(n)} = 1;$$

2° qu'il y ait une probabilité déterminée $p_{jk}^{(n)}$ (*probabilité à deux*

indices) pour que $X^{(n)} = \alpha_k$, si l'on sait que $X^{(n-1)} = \alpha_j$, les valeurs des autres variables étant inconnues; nous avons, pour $j, k = 1, 2, \dots, r$ et pour n quelconque

$$p_{jk}^{(n)} \geq 0, \quad p_{j_1}^{(n)} + p_{j_2}^{(n)} + \dots + p_{j_r}^{(n)} = 1;$$

3° qu'il y ait une probabilité déterminée $p_{ijk}^{(n)}$ (*probabilité à trois indices*) pour que l'on ait $X^{(n)} = \alpha_k$, si l'on sait que $X^{(n-2)} = \alpha_i$, $X^{(n-1)} = \alpha_j$, les valeurs que prennent les autres variables étant inconnues; nous avons pour $i, j, k = 1, 2, \dots, r$ et pour n quelconque

$$p_{ijk}^{(n)} \geq 0, \quad p_{ij_1}^{(n)} + p_{ij_2}^{(n)} + \dots + p_{ij_r}^{(n)} = 1;$$

4° Nous admettons, en général, l'existence des *probabilités à un nombre quelconque d'indices* $p_{\alpha\beta\dots\lambda k}^{(n)}$, le dernier indice (inférieur) k se rapportant à la valeur α_k que l'on attend comme valeur de $X^{(n)}$ résultant de la $n^{\text{ième}}$ expérience, tandis que les indices précédents $\alpha, \beta, \dots, \lambda$ marquent les valeurs données des variables qui précèdent $X^{(n)}$ immédiatement (1).

Ces probabilités sont liées entre elles par les relations suivantes [analogues de la formule (4) du n° 2] :

$$\begin{aligned} p_k^{(n+1)} &= \sum_{j=1}^r p_j^{(n)} h_{jk}^{(n+1)}, \\ p_k^{(n+2)} &= \sum_{i=1}^r \sum_{j=1}^r p_i^{(n)} p_{ij}^{(n+1)} p_{jk}^{(n+2)} = \sum_{i=1}^r \sum_{j=1}^r p_i^{(n)} p_{ij}^{(n+1)} p_{ijk}^{(n+2)}, \\ p_k^{(n+3)} &= \sum_{h=1}^r \sum_{i=1}^r \sum_{j=1}^r p_h^{(n)} p_{hi}^{(n+1)} p_{ij}^{(n+2)} p_{jk}^{(n+3)} \\ &= \sum_{h=1}^r \sum_{i=1}^r \sum_{j=1}^r p_h^{(n)} p_{hi}^{(n+1)} p_{ij}^{(n+2)} p_{hijk}^{(n+3)} = \dots, \end{aligned}$$

et ainsi de suite.

Une suite (3o) de variables aléatoires étant donnée et les hypothèses précédentes étant admises, plusieurs cas peuvent se présenter.

(1) Conformément au n° 1 on peut donner les définitions plus générales suivantes relatives à une suite illimitée d'événements : $p_i^{(n)}$ serait la probabilité pour que l'événement E_i se présente à la $n^{\text{ième}}$ expérience; $p_{in}^{(n)}$ serait la probabilité pour que la $n^{\text{ième}}$ donne E_k , si la $(n-1)^{\text{ième}}$ a donné E_i , et ainsi de suite.

Premier cas. — Les probabilités à deux indices $p_{jk}^{(n)}$ ne dépendent pas du premier indice j ; les probabilités à trois indices $p_{ijk}^{(n)}$ ne dépendent pas ni de i , ni de j , et en général les probabilités à un nombre quelconque d'indices ne dépendent que de l'indice supérieur n et du dernier indice inférieur. Nous avons, dans ce cas,

$$p_k^{(n)} = p_{jk}^{(n)} = p_{ijk}^{(n)} = \dots \quad (i, k = 1, 2, \dots, r).$$

Donc, la probabilité $p_{\alpha\beta\dots k}^{(n)}$ se confond avec la probabilité absolue $p_k^{(n)}$; la suite (30) est alors composée de *quantités aléatoires indépendantes*.

Second cas. — Les probabilités $p_{jk}^{(n)}$ dépendent réellement des deux indices j et k et de n ; en général, les probabilités à un nombre quelconque d'indices ne dépendent que des deux derniers indices inférieurs et de n . C'est le cas d'une *chaîne simple* à éléments variables.

Troisième cas. — Si les $p_{ijk}^{(n)}$ dépendent réellement de i , de j et de k et si les probabilités à un nombre quelconque d'indices plus grand que trois ne dépendent que des trois derniers, nous avons une *chaîne d'ordre deux* (à éléments variables en général; si les probabilités ne dépendent pas de l'indice supérieur, nous avons une chaîne d'ordre deux à éléments constants).

Cas d'une chaîne d'ordre σ . — Les probabilités à $\sigma + 1$ indices (inférieurs) dépendent réellement de tous ces indices, et les probabilités à un plus grand nombre d'indices ne dépendent que des $\sigma + 1$ derniers. Cela veut dire que la probabilité pour que $x^{(n)}$ prenne une valeur donnée dépend des valeurs de $x^{(n-\sigma)}$, $x^{(n-\sigma+1)}$, \dots , $x^{(n-1)}$; elle ne change pas si les valeurs des autres variables sont connues.

Cas général. — Une suite telle que (30), satisfaisant aux conditions énoncées au début de ce paragraphe, n'est pas une chaîne d'ordre fini; en général, quel que soit le nombre entier σ , les probabilités à σ indices dépendent de tous ces indices.

b. Le schéma suivant permet d'obtenir une suite de variables aléatoires formant une chaîne d'ordre σ avec les probabilités $p_{i_1, i_2, \dots, i_\sigma, k}^{(n)}$ données *a priori* ⁽¹⁾; chacun des indices $i_1, i_2, \dots, i_\sigma, k$ peut prendre une valeur

⁽¹⁾ Ces probabilités doivent être des nombres rationnels.

quelconque de 1 à r . Les valeurs de $x^{(n)}$ s'obtiennent par des tirages successifs d'urnes différentes dont le nombre total (pour le $n^{\text{ième}}$ tirage) est égal à σ^r . Il y a en somme r espèces de boules dans ces urnes; chaque boule porte un numéro (de 1 à r). Chaque urne est marquée par un groupe de σ numéros $i_1, i_2, \dots, i_\sigma$ (chacun des i_p varie de 1 à r), et la proportion de boules de différentes espèces qu'on y dépose est déterminée par la condition que la probabilité d'extraire une boule marquée k soit égale au nombre donné $p_{i_1 i_2 \dots i_\sigma k}^{(n)}$. Les tirages se font d'après la règle suivante : Si l'on a tiré, au $(n - \sigma)^{\text{ième}}$, $(n - \sigma + 1)^{\text{ième}}$, \dots , $(n - 1)^{\text{ième}}$ tirage des boules marquées respectivement $i_1, i_2, \dots, i_\sigma$, on devra extraire, au $n^{\text{ième}}$ tirage, des boules de l'urne marquée par le groupe $(i_1, i_2, \dots, i_\sigma)$. Ainsi, si les éléments de la chaîne sont variables, σ^r urnes de compositions données (variables avec n) devront être préparées avant de faire le $n^{\text{ième}}$ tirage⁽²⁾; dans le cas d'une chaîne d'ordre σ à éléments constants, la composition de ces σ urnes ne change pas avec n ; il faut remettre toujours dans l'urne la boule qui a été extraite.

c. Nous nous proposons dans ce Chapitre : 1° de rappeler la définition de la suite stationnaire considérée par MM. Khintchine, Slutsky et Bernstein; 2° de communiquer deux exemples simples donnés par M. Khintchine, et 3° de considérer les relations entre la notion de chaîne à éléments constants et entre celle de suite stationnaire.

Remarque sur le principe ergodique. — On peut se demander si, dans le cas d'une chaîne d'ordre supérieur, les probabilités de passage après n expériences consécutives admettent des limites pour n infini ou si la différence entre deux probabilités de ce genre, relatives à la même valeur de n , tend vers zéro pour n infini. Ces questions ont été étudiées par M. Mihoc [30] et par M. Ostenc [31].

16. Suite stationnaire. — Supposons que toutes les hypothèses sur les probabilités à un nombre quelconque d'indices, énoncées au début du n° 15 soient vérifiées, mais que toutes ces probabilités soient indépendantes de l'indice n ; nous écrirons donc $p_k, p_{jk}, p_{ijk}, \dots$, au lieu de $p_k^{(n)}, p_{jk}^{(n)}, p_{ijk}^{(n)}, \dots$. On dit alors que la suite (30) est *stationnaire*.

⁽²⁾ Ou, ce qui revient au même, une seule urne devra être préparée, sa composition dépendant toutefois des résultats donnés par les σ tirages qui précèdent.

Cette définition a été donnée en 1932 par M. Khintchine [32] qui a étudié de pareilles suites. M. Slutsky [33] et M. Bernstein [34] s'en sont également occupés.

Une suite stationnaire étant donnée, il y a une probabilité p_k , indépendante de n , pour que l'on ait $x^{(n)} = \alpha_k$, si les résultats des autres expériences restent inconnus; il y a une probabilité p_{ik} , indépendante de n , pour que l'on ait $x^{(n)} = \alpha_k$, étant donné que $x^{(n-1)} = \alpha_i$, les valeurs des autres variables étant inconnues; et ainsi de suite.

Remarquons que, même dans le cas d'une chaîne simple à éléments constants, les probabilités $p_k^{(n)}$ dépendent de n ; nous reviendrons au n° 18 sur ce point important.

17. Deux exemples indiqués par M. Khintchine. — En étudiant la notion de suite stationnaire j'ai demandé à M. Khintchine de donner des exemples simples de telles suites. Il a bien voulu m'en indiquer les deux suivants :

Premier exemple. — Considérons le mouvement uniforme d'un point sur une circonférence et définissons la probabilité pour que la position de ce point satisfasse à une certaine condition, comme la mesure relative de l'ensemble de positions satisfaisant à cette condition. Admettons que la période du mouvement soit égale à une seconde, et que l'on enregistre la position du point mobile après $\frac{1}{3}$ de seconde. Soit $x^{(n)}$ la variable aléatoire qui prend la valeur 1 ou 0, selon que le mobile se trouve au moment du $n^{\text{ième}}$ enregistrement sur la demi-circonférence de droite ou de gauche.

Nous emploierons la notation du n° 15 α (en supprimant, dans les symboles des probabilités, l'indice n). Chaque variable $x^{(n)}$ peut prendre les valeurs 1 ou 0; les indices 1 et 0 caractérisent les probabilités correspondantes (nous avons $r = 2$). Si l'on ignore le résultat des enregistrements, les probabilités absolues des deux alternatives auxquelles peut conduire l'expérience sont égales :

$$p_0 = p_1 = \frac{1}{2}.$$

Si l'on sait qu'un des enregistrements a donné $x^{(n)} = 1$, les résultats des autres étant inconnus, la probabilité pour que l'enregistrement

suisant donne $x^{(n+1)} = 1$ est égale à $\frac{1}{3}$, car ce sont seulement les points situés sur un tiers de la demi-circonférence de droite qui peuvent donner 1 suivi également de 1.

Si $x^{(n)} = 1$, la probabilité pour que l'on ait $x^{(n+1)} = 0$ est égale à $\frac{2}{3}$. Nous trouvons ainsi

$$p_{11} = \frac{1}{3}, \quad p_{10} = \frac{2}{3}, \quad p_{01} = \frac{2}{3}, \quad p_{00} = \frac{1}{3}.$$

Si l'on sait que $x^{(n)} = 1$ et que $x^{(n+1)} = 1$, la probabilité p_{111} de la relation $x^{(n+2)} = 1$ sera égale à zéro. On détermine facilement les probabilités à trois indices; voici les résultats :

$$p_{111} = p_{000} = 0, \quad p_{110} = p_{001} = 1,$$

$$p_{101} = p_{100} = p_{011} = p_{010} = \frac{1}{2}.$$

Écrivons encore les probabilités à quatre indices. Si l'on sait, par exemple, que trois enregistrements consécutifs ont donné successivement 1, 0, 1, on sera certain que le suivant donnera nécessairement 1, de sorte que $p_{1010} = 0$. On trouve ainsi que six parmi les seize probabilités à quatre indices, sont égales à l'unité

$$p_{0110} = p_{1001} = p_{1101} = p_{1011} = p_{0100} = p_{0010} = 1,$$

toutes les autres étant nulles.

Ce ne sont donc que les probabilités à un, à deux et quelques-unes à trois indices qui sont différentes de zéro et de l'unité. Si le nombre q est égal à quatre ou supérieur à quatre, chaque probabilité à q indices est égale soit à zéro soit à l'unité. Si l'on ne tenait pas compte de ces probabilités à q indices, dont chacune est égale à zéro ou à l'unité, on pourrait considérer ce cas comme une chaîne d'ordre deux. Mais cela ne correspondrait pas à la définition donnée au n° 15 a (troisième cas); en effet, si l'on sait, par exemple, que $x^{(n)} = 1$, $x^{(n+1)} = 0$, la probabilité de l'équation $x^{(n+2)} = 1$ est $p_{101} = \frac{1}{2}$; si l'on sait, de plus, que $x^{(n-1)} = 0$ la probabilité de l'équation $x^{(n+2)} = 1$ est égale à $p_{0101} = 0$.

Second exemple. — Les conditions relatives au mouvement du point sur la circonférence étant les mêmes que dans l'exemple précédent, on

suppose que l'intervalle de temps entre deux enregistrements consécutifs soit mesuré par un nombre irrationnel.

Si l'on connaît un nombre q de résultats qui correspondent à q enregistrements consécutifs, le résultat de l'enregistrement R qui suit immédiatement, ne peut pas être prévu, car la position initiale du point est inconnue. Pour chacun des deux cas qui peuvent se présenter comme résultat de R , il y a une probabilité déterminée en vertu précisément de notre hypothèse fondamentale, mais, quel que soit le nombre q , ces probabilités dépendent en général de toute la suite des q nombres donnés par les q enregistrements consécutifs qui précèdent R . Il n'y a pas de nombre q tel que chacune des probabilités à q indices soit égale à zéro ou à l'unité. Cette suite stationnaire ne peut pas être considérée comme une chaîne d'ordre fini.

18. Relations entre les chaînes et les suites stationnaires. — Posons-nous le problème suivant : sous quelles conditions peut-on considérer une chaîne d'ordre fini comme une suite stationnaire ? Si σ est l'ordre de la chaîne, il faut d'abord que les probabilités à 2, 3, ..., σ indices soient indépendantes de l'indice n , mais il faut encore que les probabilités à un seul indice (probabilités absolues) ne dépendent pas de n . Nous avons vu, au n° 2, équation (5 bis), que, déjà dans le cas d'une chaîne simple à éléments constants, la probabilité absolue $p_k^{(n)}$ dépend en général de n , et cela a lieu encore dans le cas d'une chaîne d'ordre σ à éléments constants. Nous nous bornons ici à considérer le cas d'une chaîne simple ($\sigma = 1$). Il est possible de choisir les probabilités absolues, de telle manière que $p_k^{(n)}$ ne dépende pas de n . Il faut, pour cela, que

$$p_k^{(n+1)} = p_k^{(n)} = \dots = p_k^{(0)};$$

en remplaçant $p_k^{(n+1)}$ et $p_k^{(n)}$ par l'expression (5 bis), nous obtenons

$$(31) \quad \sum_{i=1}^r p_i^{(0)} P_{ik}^{(n)} = \sum_{i=1}^r p_i^{(0)} P_{ik}^{(n+1)}.$$

Or, on peut satisfaire à ces équations en prenant

$$(32) \quad p_i^{(n)} = p_i^{(0)} = \sum_{s=1}^r \alpha_s \Pi_{si}$$

pour n quelconque, où a_s ($s = 1, 2, \dots, r$) sont des constantes arbitraires satisfaisant aux conditions

$$(33) \quad a_s \geq 0, \quad a_1 + a_2 + \dots + a_r = 1,$$

et où les Π_{si} sont les limites des moyennes arithmétiques (voir n° 9 et n° 20). Si l'on substitue l'expression (32) à $p_i^{(0)}$ dans le premier membre de (31), ce dernier devient égal à

$$\sum_{i=1}^r \sum_{s=1}^r a_s \Pi_{si} P_{ik}^{(n)} = \sum_{s=1}^r a_s \sum_{i=1}^r \Pi_{si} P_{ik}^{(n)} = \sum_{s=1}^r a_s \Pi_{sk},$$

et l'on trouve que le second membre de (31) se réduit à la même valeur. D'ailleurs, les conditions (33) montrent que les $p_i^{(0)}$, définies par (32), sont des quantités non négatives, et que l'on a

$$\sum_{i=1}^r p_i^{(0)} = \sum_{i=1}^r \sum_{s=1}^r a_s \Pi_{si} = \sum_{s=1}^r a_s = 1.$$

En résumé, *une chaîne simple à éléments constants devient une suite stationnaire à condition de choisir les probabilités absolues $p_i^{(n)} = p_i^{(0)}$ d'après la formule (32) où les constantes a_i satisfont à (33).*

Nous voyons ainsi d'abord que, en général, la chaîne simple à éléments constants, définie au n° 1, ne donne pas lieu à une suite stationnaire $x^{(1)}, x^{(2)}, \dots$, de variables aléatoires. Ce n'est que par le choix particulier, défini par la formule (32), d'une probabilité absolue $p_k^{(n)}$ indépendante de n , que cette suite devient stationnaire. La notion de suite stationnaire est plus générale que celle d'une chaîne à éléments constants d'ordre fini, en ce sens qu'une suite stationnaire ne se réduit pas nécessairement à une chaîne d'ordre fini (voir n° 15 a).

CHAPITRE IV.

COEFFICIENTS DE CORRÉLATION. CHAINES INVERSES.

19. Coefficient de corrélation entre deux termes d'une chaîne simple.

— *a.* Commençons par donner la définition générale du coefficient de corrélation entre deux variables aléatoires x et y . Si $\alpha_1, \alpha_2, \dots, \alpha_r$ sont les valeurs que peut prendre x et si p_1, p_2, \dots, p_r sont les probabilités correspondantes, la valeur moyenne de x sera donnée par la formule

$$E(x) = \sum_{i=1}^r p_i \alpha_i.$$

De même, si $\gamma_i (i=1, 2, \dots, r)$ sont les valeurs que peut prendre y , et si q_i sont les probabilités correspondantes, on aura

$$E(y) = \sum_{i=1}^r q_i \gamma_i.$$

Soit c_{ik} la probabilité pour que l'on ait à la fois $x = \alpha_i$ et $y = \gamma_k$. La valeur moyenne du produit xy sera égale alors à

$$E(xy) = \sum_{i=1}^r \sum_{k=1}^r c_{ik} \alpha_i \gamma_k.$$

Le coefficient de corrélation R entre x et y est défini par la formule suivante :

$$(34) \quad R = \frac{E\{[x - E(x)][y - E(y)]\}}{\sqrt{E[x - E(x)]^2 E[y - E(y)]^2}}.$$

Pour mettre R sous une forme plus simple, rappelons que, si a, b, c, \dots sont des constantes,

$$E(ax + by + cz + \dots) = aE(x) + bE(y) + cE(z) + \dots$$

Nous pouvons donc écrire

$$\begin{aligned} E\{[x - E(x)][y - E(y)]\} &= E(xy) - E(x)E(y), \\ E[x - E(x)]^2 &= E(x^2) - [E(x)]^2, \\ E[y - E(y)]^2 &= E(y^2) - [E(y)]^2, \end{aligned}$$

de sorte que la formule (34) devient

$$(34 \text{ bis}) \quad R = \frac{E(x, y) - E(x)E(y)}{\sqrt{\{E(x^2) - [E(x)]^2\} \{E(y^2) - [E(y)]^2\}}}.$$

b. Introduisons maintenant deux variables auxiliaires ξ et η par les équations

$$\xi = \frac{x - E(x)}{\sqrt{E(x^2) - [E(x)]^2}}, \quad \eta = \frac{y - E(y)}{\sqrt{E(y^2) - [E(y)]^2}}.$$

On trouve que

$$(35) \quad E(\xi) = 0, \quad E(\eta) = 0, \quad E(\xi^2) = 1, \quad E(\eta^2) = 1,$$

et la formule (24 bis) prend la forme simplifiée

$$(36) \quad R = E(\xi, \eta).$$

En général, si deux variables aléatoires ξ et η satisfont aux conditions (35), le coefficient de corrélation entre ξ et η est égal à la valeur moyenne de leur produit.

Remarquons que le même nombre R est en même temps coefficient de corrélation entre x et y et entre les variables auxiliaires ξ et η .

c. Soient maintenant x la variable $x^{(m)}$ attachée au $m^{\text{ième}}$ terme d'une chaîne simple à éléments constants, et y la variable $x^{(n)}$ attachée au $n^{\text{ième}}$ terme de la même chaîne ($m < n$). Nous avons, toujours avec les notations des nos 1 et 2,

$$E(x) = E(x^{(m)}) = \sum_{i=1}^r p_i^{(m)} \alpha_i, \quad E(x^2) = \sum_{i=1}^r p_i^{(m)} \alpha_i^2,$$

$$E(y) = E(x^{(n)}) = \sum_{i=1}^r p_i^{(n)} \alpha_i, \quad E(y^2) = \sum_{i=1}^r p_i^{(n)} \alpha_i^2.$$

Soit $\beta_j^{(m)}$ la valeur que prend la variable auxiliaire ξ introduite plus haut lorsque $x = x^{(m)}$ est égal à α_j ; et soit $\beta_j^{(n)}$ la valeur de la variable auxiliaire η , lorsque $y = x^{(n)}$ prend la valeur α_j . Nous avons, pour tout indice m :

$$\beta_j^{(m)} = \frac{\alpha_j - \sum_{i=1}^r p_i^{(m)} \alpha_i}{\sqrt{\sum_{i=1}^r p_i^{(m)} \alpha_i^2 - \left[\sum_{i=1}^r p_i^{(m)} \alpha_i \right]^2}}.$$

La probabilité de l'égalité $\xi = \beta_j$ est évidemment égale à $p_j^{(m)}$; la probabilité pour que l'on ait $\eta = \beta_j$, étant donné que $\xi = \beta_i$, est égale à $P_{ij}^{(n-m)}$. Donc le coefficient de corrélation $R_{m,n}$ entre $x^{(m)}$ et $x^{(n)}$ est exprimé selon (36) par la formule

$$R_{m,n} = E(\xi \cdot \eta) = \sum_{i=1}^r \sum_{j=1}^r p_i^{(m)} P_{ij}^{(n-m)} \beta_i^{(m)} \beta_j^{(n)}.$$

Le coefficient de corrélation entre les variables aléatoires $x^{(m)}$ et $x^{(n)}$ attachées à deux termes d'une chaîne simple à éléments constants dépend de m et de n .

Il se réduirait à une fonction de la seule différence $n - m$, si l'on avait

$$p_i^{(0)} = p_i^{(1)} = p_i^{(2)} = \dots = p_i^{(n)} = \dots,$$

pour $i = 1, 2, \dots, r$. Or, ces conditions sont les mêmes que celles que nous avons trouvées au n° 18. Il suffit, pour les satisfaire, de prendre

$$p_i^{(n)} = p_i^{(0)} = \sum_{s=1}^r a_s \Pi_{si}, \quad a_s \geq 0, \quad \sum_{s=1}^r a_s = 1.$$

Avec cette hypothèse, les $\beta_j^{(m)}$ ne dépendent plus de m ; nous écrirons alors β_j au lieu de $\beta_j^{(m)}$.

Le coefficient de corrélation s'écrit alors

$$(37) \quad R_{m,n} = \sum_{i=1}^r \sum_{j=1}^r \sum_{s=1}^r a_s \Pi_{si} P_{ij}^{(n-m)} \beta_i \beta_j,$$

et nous pouvons dire : *si les probabilités absolues $p_i^{(n)}$ sont indépendantes de n , le coefficient de corrélation entre $x^{(m)}$ et $x^{(n)}$ est une fonction de la seule différence $n - m$.*

20. Chaînes inverses. — *a.* Supposons qu'une chaîne simple à éléments constants soit définie au moyen des probabilités $p_{ik}(i, k = 1, 2, \dots, r)$, et soit, comme au n° 1, $p_i^{(n)}$ la probabilité absolue de l'équation $x^{(n)} = \alpha_i$. Introduisons la *probabilité de passage inverse* $q_{ik}(i, k = 1, 2, \dots, r)$ pour que l'on ait $x^{(n)} = \alpha_i$, étant donné que $x^{(n+1)} = \alpha_k$. Nous désignerons par $Q_{ik}^{(n)}$ la probabilité pour que $x^{(m)} = \alpha_i$, si l'on sait que $[x^{(m+n)}] = \alpha_k$, et nous poserons $Q_{ik}^{(1)} = q_{ik}$.

Pour déduire les relations entre p_{ik} et q_{ik} , évaluons de deux manières différentes la probabilité pour que l'on ait $x^{(n)} = \alpha_i$ et $x^{(n+1)} = \alpha_k$. D'une part, cette probabilité est égale à $p_i^{(n)} p_{ik}$, d'autre part elle est égale à $q_{ik} p_k^{(n+1)}$, donc

$$p_i^{(n)} p_{ik} = q_{ik} p_k^{(n+1)}.$$

Dans un travail publié en 1936, en collaboration avec J. Potoček [35] nous nous sommes posé la question suivante : trouver les conditions pour que la chaîne inverse soit simple, c'est-à-dire pour que les q_{ik} ne dépendent que de i et de k . Or, la formule précédente montre que

$$q_{ik} = \frac{p_i^{(n)}}{p_k^{(n+1)}} p_{ik}$$

dépend en général de n . Pour que q_{ik} ne dépende pas de n , il faut et il suffit que le rapport $p_i^{(n)} : p_k^{(n+1)}$ soit indépendant de n . Si les limites $\lim_{n \rightarrow \infty} P_{ik}^{(n)} = P_{ik}$ existent, on peut prendre

$$p_i^{(m)} = P_{si}, \quad p_k^{(m+1)} = P_{sk},$$

s étant un indice arbitraire.

b. Nous avons communiqué le résultat précédent à M. Fréchet, et à M. Kolmogoroff qui travaillait en même temps sur le même sujet et qui est arrivé à des résultats équivalents [36]; voir aussi [22]. Les deux lettres que nous avons reçues en réponse (mentionnées au n° 9, remarque en bas de la page) contiennent une solution plus générale, obtenue au moyen des formules (32) du n° 18. Nous pouvons prendre

$$(38) \quad q_{ik} = \frac{\sum_{s=1}^r a_s \Pi_{si}}{\sum_{s=1}^r a_s \Pi_{sk}} p_{ik}, \quad a_s \geq 0, \quad \sum_{s=1}^r a_s = 1.$$

En résumé, si $X^{(1)}, X^{(2)}, \dots$ sont des variables aléatoires liées en chaîne simple à éléments constants, la probabilité q_{ik} pour que l'on ait $x^{(n)} = \alpha_i$, si l'on sait que $x^{(n+1)} = \alpha_k$ dépend en général de n . Pour qu'elle soit indépendante de n , on n'a qu'à choisir les probabilités absolues $p_i^{(n)}$ d'après la formule (32)

$$p_i^{(n)} = \sum_{s=1}^r a_s \Pi_{si}, \quad a_s \geq 0, \quad \sum_{s=1}^r a_s = 1.$$

21. Sur les applications des chaînes en Physique. — *a.* L'introduction des variables en chaîne ou, en général, des variables liées entre elles revient à généraliser la notion classique de loi physique. Du point de vue de la physique mathématique classique l'évolution d'un système physique est théoriquement déterminée par des lois (exprimées par des équations différentielles) et par les conditions initiales. Lorsqu'on adopte le point de vue probabiliste, on admet qu'il y a, par exemple, une probabilité pour qu'un point mobile se trouve dans une position donnée, et qu'il y a une probabilité pour qu'il passe au cours d'un temps donné d'une position à une autre.

Bornons-nous à considérer le mouvement d'un seul point dont les positions possibles sont A_1, A_2, \dots , et supposons qu'on observe sa position au bout de chaque seconde. Si le mouvement est *stationnaire*, les probabilités absolues $p_i^{(n)}$ pour que le point se trouve, au bout de n secondes, en A_i , sont indépendantes de n ; de même les probabilités de passage ne dépendent pas de n . Ces cas correspondent soit à une chaîne à éléments constants où les probabilités absolues ont été choisies selon (32), soit à une suite stationnaire générale (chaîne d'ordre infini).

Si les positions successives du point mobile sont liées en chaîne simple sans que les probabilités absolues vérifient la condition (32), le mouvement ne sera pas stationnaire, mais il pourra tendre vers un état stationnaire. Par exemple, dans le cas régulier les $p_{ik}^{(n)}$ tendent, si n (qui est proportionnel au temps) augmente indéfiniment, vers des limites P_k , et l'on peut étudier comment elles deviennent asymptotiquement indépendantes de n .

L'étude du cas le plus général d'un système non stationnaire conduit à considérer les chaînes à éléments variables; ce cas correspond aux mouvements d'un système subissant l'effet de forces qui changent au cours du temps. Quoique l'étude de ce cas général puisse, sans doute, donner des résultats importants, il semble que l'étude du cas non stationnaire à éléments constants (forces extérieures constantes ou peu variables) soit actuellement plus nécessaire, car il est préférable d'examiner différentes questions d'abord dans les cas plus simples avant de traiter le cas le plus général.

Nous avons, au cours de ces Leçons, considéré les moyennes arithmétiques Π_{ik} . Ces moyennes, interprétées du point de vue physique,

expriment des propriétés relatives à l'ensemble de toutes positions du point mobile. Il faut distinguer ces propriétés des autres qui nous renseignent plus précisément sur la nature des mouvements et que nous n'avons pas considérées ici : Il s'agit, par exemple, de savoir comment les différentes positions possibles se partagent en groupes tels que le point mobile ne puisse plus sortir d'un tel groupe, s'il y est entré. Voir, pour ces problèmes, les travaux de MM. Hadamard [37], Mises [38], Romanovsky [7] et Kolmogoroff [22], ainsi que les travaux non encore publiés de M. Doeblin, voir le résumé [40].

b. Enfin, je voudrais dire quelques mots sur la *réversibilité des lois physiques*; je considère le mouvement d'un point qui correspond au cas régulier d'une chaîne simple à éléments constants. Il s'agit d'éclaircir, au moyen du Calcul des probabilités, la notion si difficile de réversibilité. Quelques auteurs ont introduit l'hypothèse $p_{ik} = p_{ki}$ qui correspond, selon leur manière de voir, à la réversibilité des phénomènes physiques; la probabilité de passer d'un état A_i à un autre A_k est égale à celle de passer de A_k à A_i . Il résulte de cette hypothèse que, si les probabilités absolues ont été choisies d'une manière convenable (voir n° 20), on aura $p_{ik} = p_{ki} = q_{ik} = q_{ki}$.

Du point de vue physique ce cas correspond, à condition d'introduire les variables continues, au phénomène de diffusion en l'absence de forces extérieures.

Mais on peut poser le problème de la réversibilité d'une autre manière qui correspond, dans le cas des variables continues, à la diffusion sous l'influence de forces extérieures (constantes ou peu variables). Dans un travail récent [39], M. Kolmogoroff traite le problème de la diffusion dans un espace à un nombre quelconque de dimensions; il suppose que la densité de probabilité de passage satisfait à une équation aux dérivées partielles du type parabolique. La densité de probabilité $p(A, B, t)$ de passage de A à B en t secondes n'est pas symétrique en A et B; q étant la densité de probabilité inverse, M. Kolmogoroff pose comme condition

$$(39) \quad q(B, A, t) = p(A, B, t).$$

Il en résulte que *la probabilité absolue* pour que le point mobile soit d'abord en A et ensuite, après t secondes, en B, *est une fonction*

symétrique de A et de B. Le résultat principal et très remarquable de ce travail [39] consiste à démontrer que la possibilité de trouver une telle fonction $p(A, B, t)$ dépend de l'existence du potentiel des forces extérieures agissant sur le fluide.

Pour donner un exemple simple où la condition [39] se trouve vérifiée, je rappelle les exemples sur la diffusion linéaire calculés par Smoluchowski (*voir mon Cours de 1930, [16], p. 51*). Si, par exemple, $\Phi(A, B, t)$ désigne la densité de probabilité de passage sous l'influence de la pesanteur d'un grain dans une colonne verticale d'eau de A en B en t secondes, nous avons

$$\lim_{t \rightarrow \infty} \Phi(A, B, t) = P(A).$$

La fonction $P(A)$ définit une distribution stationnaire des grains. Regardons $P(A)$ comme probabilité absolue pour que le grain se trouve en A; la probabilité absolue pour qu'il se trouve à l'époque $t = 0$ en A, et pour qu'il se trouve en B après t secondes est égale à $P(A) \Phi(A, B, t)$. La formule de Smoluchowski montre que ce produit est une fonction symétrique de A et de B; il est évident que le potentiel existe ici, car il n'y a ici qu'une force extérieure à considérer. Si $\Psi(A, B, t)$ est la densité de probabilité pour le passage inverse de B à A ⁽¹⁾, il résulte de (39), où il faut remplacer respectivement p et q par Φ et Ψ , que

$$P(A) \Phi(A, B, t) = \Psi(A, B, t) P(B);$$

cette formule sert à déterminer $\Psi(A, B, t)$.

Revenons encore au cas discontinu; le problème de la recherche des chaînes continues telles que (39) soit valable revient ici à la détermination d'une chaîne simple à éléments constants telle que

$$(40) \quad q_{ki} = p_{ik}.$$

Une solution rappelée déjà est immédiate dans le cas symétrique où $p_{ki} = p_{ik}$. Il faudrait chercher si l'on peut trouver d'autres chaînes qui satisfont à (40).

(1) C'est-à-dire la densité de probabilité pour que, B étant la position donnée du point à l'époque $t (t > 0)$, A soit sa position à l'époque $t = 0$.

BIBLIOGRAPHIE.

1. BOREL (E.). — *Principes et formules classiques du Calcul des probabilités* (*Traité du Calcul des probabilités et de ses applications*, t. 1, Rédigé par R. Lagrange, Paris, 1925).
2. KOLMOGOROFF (A.). — *Grundbegriffe der Wahrscheinlichkeitsrechnung*, Berlin, 1933.
3. FRÉCHET (M.). — *Recherches théoriques modernes sur la Théorie des Probabilités; Livre I* (*Traité du Calcul des Probabilités et de ses applications*, t. 1, fasc. 3, Paris, 1937).
4. HOSTINSKÝ (B.). — *Méthodes générales du Calcul des Probabilités* (*Mémorial des Sciences mathématiques*, fasc. 52, Paris, 1931).
5. ROMANOVSKY (V.). — *Sur les chaînes de Markoff* (*C. R. de l'Académie de l'U. R. S. S.*, 1929, A, n° 9, p. 203-208).
6. FERRON (O.). — *Zur Theorie der Matrizes* (*Math. Ann.*, Bd. 64, 1907, p. 248-263).
7. ROMANOVSKY (V.). — *Recherches sur les chaînes de Markoff* (*Acta Mathematica*, t. 66, 1936, p. 147-251).
8. KAUCKÝ (J.). — *Quelques remarques sur les chaînes de Markoff* (en tchèque, résumé en français) (*Publications de la Faculté des Sciences de l'Université Masaryk*, n° 131, Brno, 1930).
9. KONEČNÝ (M.). — *Sur la théorie des chaînes de Markoff* (en tchèque, résumé en français) (*Ibid.*, n° 147, Brno, 1931).
10. KONEČNÝ (M.). — *Trois théorèmes sur la limite des transformations itérées* (*Ibid.*, n° 163, Brno, 1932).
11. FRÉCHET (M.). — *Compléments à la théorie des Probabilités discontinues, en chaîne* (*Annali della R. Scuola Normale Superiore di Pisa, Scienze Fis. e Matem.*, série II, vol. II, 1933, p. 131-164).
12. HOSTINSKÝ (B.). — *Complément à la Note sur les Probabilités relatives aux transformations répétées* (*C. R. Acad. Sc.*, t. 186, 1928, p. 487-489); *Sur les transformations itérées des variables aléatoires* (*Publications de la Faculté des Sciences de l'Université Masaryk*, n° 93, Brno, 1928).
13. FROBENIUS (G.). — *Ueber Matrizen aus nicht negativen Elementen* (*Sitzungsberichte der Akad. der Wissenschaften zu Berlin*, 1912, p. 456-477).
14. FRÉCHET (M.). — *Comportement asymptotique des solutions d'un système d'équations linéaires et homogènes aux différences finies du premier ordre à coefficients constants* (*Publications de la Faculté des Sciences de l'Université Masaryk*, n° 178, Brno, 1933).
15. LAPLACE (P.-S.). — *Théorie analytique des Probabilités* (*Œuvres de L.*, t. VII, Paris, 1886).
16. HOSTINSKÝ (B.). — *Application du Calcul des Probabilités à la théorie du mouvement brownien* (*Annales de l'Inst. H. Poincaré*, vol. 3, 1932, p. 1-74).

17. POTOČEK (J.). — *Sur la dispersion dans la théorie des chaînes de Markoff* (en tchèque, résumé en français) (*Publications de la Faculté des Sciences de l'Université Masaryk*, n° 154, Brno, 1932).
18. SCHULZ (G.). — *Zur Theorie des Galtonschen Brettes* (*Zeitschrift für Physik*, Bd. 92, 1934, p. 747-754).
19. MÜNZNER (H.). — *Ueber eine spezielle Markoffsche Kette am Galtonbrett* (*Zeitschrift für angewandte Mathematik und Mechanik*, Bd. 14, 1934, p. 343-346).
20. SEITZ (W.) und HAMACHER-ODENHAUSEN (K.). — *Untersuchungen über das Galtonbrett* (*Physikalische Zeitschrift*, Bd. 35, 1934, p. 530-532).
21. FORTET (R.). — *Sur les probabilités en chaîne* (*C. R. Acad. Sc.*, t. 201, 16 juillet 1935, p. 184-186; t. 202, 20 avril 1936, p. 1362-1364).
22. KOLMOGOROFF (A.). — *Anfangsgründe der Theorie der Markoffschen Ketten mit unendlich vielen möglichen Zuständen* (*Matemat. Sbornik*, Moscou, Nouvelle série I (43), 1936, p. 607-610].
23. HOSTINSKÝ (B.). — *Sur les probabilités en chaîne* (*C. R. Acad. Sc.* t. 202, 23 mars 1936, p. 1000-1002).
24. EGGENBERGER (F.) et PÓLYA (M.). — *Ueber die Statistik verketteter Vorgänge* (*Zeitschrift für angewandte Mathem. und Mechanik*, Bd. 3, 1923, p. 279-289).
25. ONICESCU (O.) et MIHOC (G.). — *Sur les chaînes de variables statistiques* (*Bull. des Sciences mathém.*, 2^e série, t. 59, 1935, p. 174-192).
26. WHITWORTH (W.-A.). — *Choice and Chance*, Cambridge 1878.
27. KAUCKÝ (J.). — *Le problème des itérations dans un cas des probabilités dépendantes* (*C. R. Acad. Sc.*, t. 202, 2 mars 1936, p. 722-724).
28. MARKOFF (A.-A.). — *Sur les épreuves liées en chaîne par les événements laissés sans observation* (en russe) (*Bull. de l'Acad. Impér. des Sciences de Saint-Petersbourg*, 6^e série, t. 6, 1912, p. 551-572); *Sur l'application de la méthode des espérances mathématiques aux sommes liées* (*Ibid.*, t. 9, 1915, p. 1453-1484).
29. HOSTINSKÝ (B.). — *Sur les mouvements qui dépendent du hasard* (*C. R. Acad. Sc.*, t. 202, 22 juin 1936, p. 2029-2031).
30. MIHOC (G.). — *Sur les chaînes multiples discontinues* (*Ibid.*, t. 198, 18 juin 1934, p. 2135-2137).
31. OSTENG (E.). — *Sur le principe ergodique dans les chaînes de Markoff à éléments variables* (*Ibid.*, t. 199, 16 juillet 1934, p. 175-176).
32. KHINTCHINE (A.). — *Sulle successioni stazionarie di eventi* (*Giornale dell'Istituto italiano degli attuari*, Anno III, 1932, p. 267-272); *Korrelationstheorie der stationären stochastischen Prozesse* (*Mathem. Ann.*, Bd. 109, 1934, p. 604-615).
33. SLUTSKY (E.). — *Alcune applicazioni dei coefficienti di Fourier all'analisi delle funzioni aleatorie stazionarie* (*Giornale dell'Istituto italiano degli attuari*, t. 5, 1934, p. 435-482).
34. BERNSTEIN (S.). — *Sur les liaisons entre les grandeurs aléatoires* (*Verhandlungen der internat. Mathematiker-Kongresses*, Zürich, 1932, Bd. I, p. 288-309).
35. HOSTINSKÝ (B.) et POTOČEK (J.). — *Chaînes de Markoff inverses* (*Bulletin international de l'Académie des Sciences de Bohême*, Prague, 1935, p. 64-67).
36. KOLMOGOROFF (A.). — *Zur Theorie der Markoffschen Ketten* (*Math. Ann.*, Bd. 112, 1935, p. 155-160).

37. HADAMARD (J.). — *Sur le battage des cartes et ses relations avec la mécanique statistique* (*Atti del Congresso internaz. dei Matematici*, Bologne, 1928, t. V, p. 133-140).
 38. MISES (R. v.). — *Wahrscheinlichkeitsrechnung*, Leipzig, Berlin, 1931 (IV Abschnitt).
 39. KOLMOGOROFF (A.). — *Zur Umkehrbarkeit der statistischen Naturgesetze* (*Math. Ann.*, Bd. 113, 1937, p. 766-772).
 40. DOEBLIN (W.). — *Le cas discontinu des probabilités en chaîne* (*Publications de la Fac. des Sc. de l'Univ. Masaryk*, n° 236, Brno, 1937).
-