STICHTING
MATHEMATISCH CENTRUM
2e BOERHAAVESTRAAT 49
AMSTERDAM

SP 1

D. v Dantzig

Sur le methode des fonctions generatrices

(Colloques internationaux du Centre national de
de la Recherche Scientifique, 13.
Le calcul de Probability et ses applications,
Lyon, 28.06-03.07.1948)

Overdruk uit
Calcul de Probability, Paris, Hermann, 1949
p 29-45)

Extrait
du
Calcul des Probabilit
(1949, pages 29 à 46.

D. v. Dantzig, Sur la méthode des fonctions génératrices, Colloques internationaux du centre national de la recherche scientifique, XIII, Le calcul des probabilités et ses applications, Lyon - 28 juin au 3 juillet 1948, Paris 1949.

SUR LA MÉTHODE DES FONCTIONS GÉNÉRATRICES

par D. VAN DANTZIG.

(Amsterdam.)

INTRODUCTION.

1. La méthode des fonctions génératrices est due en première ligne aux grands probabilistes français. Elle a été introduite comme une méthode générale par P. S. DE LAPLACE, après avoir été employée implicitement dans des cas spéciaux par A. DE MOIVRE. La méthode des fonctions caractéristiques, qui est intimement liée avec elle a été étudiée par A. L. CAUCHY et s'est renouvelée principalement dans les mains de M. Paul LÉVY. Beaucoup de résultats des plus beaux, obtenus en France de même qu'en d'autres pays doivent leur origine à ces méthodes. Un colloque organisé par les grands probabilistes français d'aujourd'hui, des descendants si dignes de leurs fameux précurseurs, peut donc être considéré comme la meilleure opportunité pour introduire publiquement une extension de cette méthode que j'ai développée peu à peu depuis 1940 et qui n'a été publiée encore que dans les cours cyclostylés que j'ai donnés depuis 1946 à l'université d'Amsterdam.

2. L'extension de la méthode des fonctions génératrices, que j'appellerai la *méthode des marques collectives,* peut être introduite d'une manière abstraite, ou bien d'une manière intuitive. D'après la première manière les variables figurantes dans les équations peuvent être considérées comme des symboles ou «marques» abstraits. Ici je choisirai la deuxième manière quoiqu'elle soit un peu moins générale, où les variables sont interprétées comme des probabilités, donc restreintes à des valeurs réelles $\geqslant 0$ et $\leqslant 1$.

§ 1. DÉFINITIONS DES MARQUES COLLECTIVES.

3. Considérons un système $\mathcal{A}$ d'éventualités (ou événements éventuels) [1] $\mathcal{A}_1, \mathcal{A}_2, \ldots$ Appellerons un tel système *complet* (sous un système déterminé de conditions $\mathcal{H}$, les «hypothèses»), lorsqu'il est certain qu'au moins une d'entre eux se produit; *exclusif* lorsqu'elles sont incompatibles deux à deux, et *catégorique* lorsqu'il est simultanément complet et exclusif. Un système catégorique sera appelé aussi brièvement une *catégorie.*

Soit $\mathcal{A} = \{\mathcal{A}_1, \mathcal{A}_2, \ldots\}$ une catégorie sous l'hypothèse $\mathcal{H}$ et soit $p_n = P\left[\frac{\mathcal{A}_n}{\mathcal{H}}\right]$ la probabilité d'$\mathcal{A}_n$ de manière que $\sum^n p_n = 1$. D'après la méthode de LAPLACE on fait correspondre avec le système des p_n la fonction

$$C(A) = \Sigma^n p_n A^n \tag{1}$$

où A est une variable arbitraire, telle que la série (si elle est infinie) converge. On peut généraliser cette méthode en faisant correspondre avec le système des p_n la fonction

$$C = C(A_1, A_2, \ldots) = \Sigma^n p_n A_n \tag{2}$$

où les A_n sont des variables, dépendantes ou non, telle que la série (2) converge. Ceci sera certainement le cas lorsque les A_n sont des nombres non-négatifs $\leqslant 1$. Les A_n seront appelées les *marques* des $\mathcal{A}_n$. On peut alors interpréter les A_n comme des probabilités, et bien de la manière suivante. Prenons une suite de loteries Λ_n où une éventualité déterminée $\mathcal{E}$ (la même pour tous les Λ_n) peut se produire ou non avec une probabilité déterminée $1 - A_n$ ou A_n respectivement, et supposons qu'$\mathcal{E}$ ne peut se produire d'aucune autre manière. Convenons de tirer un lot de Λ_n si et seulement si sous l'hypothèse $\mathcal{H}$ l'éventualité $\mathcal{A}_n$ se réalise. Alors C est la probabilité que $\mathcal{E}$ ne se réalise pas. Il est souvent aisé d'appeler l'éventualité $\mathcal{E}$ une (ou la) *catastrophe.* Alors C est la probabilité que la catastrophe $\mathcal{E}$ n'arrivera pas. La fonction $C = C(A) = C(A_1, A_2, \ldots)$ sera appelée la *marque collective* du système des $\mathcal{A}_n$ ou bien des p_n.

4. Considérons maintenant un système $\mathcal{A} = \{\mathcal{A}^{(1)}, \mathcal{A}^{(2)}, \ldots\}$ de catégories $\mathcal{A}^{(m)} = \{\mathcal{A}_1^{(m)}, \mathcal{A}_2^{(m)}, \ldots\}$, tous sous la même hypothèse $\mathcal{H}$. Les catégories $\mathcal{A}^{(m)}$ peuvent être dépendantes ou non. Soit

$$p_{n_1, n_2, \ldots} = P\left[\frac{\mathcal{A}_{n_1}^{(1)} \wedge \mathcal{A}_{n_2}^{(2)} \wedge \cdots}{\mathcal{H}}\right]$$

la probabilité que pour chaque m l'n_m-ième éventualité de la m-ième catégorie se réalise. Prenons pour chaque

[1] Je préfère le terme «éventualité» au terme «événement», puisque le premier exprime mieux qu'il s'agit d'un «événement» qui peut se produire ou non.

m et n_m une loterie $\Lambda_{n_m}^{(m)}$, et convenons de tirer pour chaque m un lot de $\Lambda_{n_m}^{(m)}$ lorsque dans la m-ième catégorie l'éventualité $\mathcal{A}_{n_m}^{(m)}$ se produit, et dans ce cas seulement. Supposons tous ces tirages mutuellement indépendants. Supposons d'ailleurs qu'un tel tirage mène avec une probabilité $1 - A_{n_m}^{(m)}$ à une «catastrophe» $\mathcal{E}^{(m)}$, dépendante de m seulement, qui ne peut se produire d'aucune autre manière. Alors la probabilité qu'aucune des catastrophes $\mathcal{E}^{(m)}$ n'arrive (toujours sous l'hypothèse $\mathcal{H}$) sera

$$(3) \qquad C = C(A_{n_m}^{(m)}) = \Sigma^{n_1} \Sigma^{n_2} \dots p_{n_1, n_2, \dots} \prod_m A_{n_m}^{(m)}.$$

Cette fonction sera appelée la *marque collective* du système de probabilités $p_{n_1, n_2, \dots}$ ou du système d'éventualités $\{\mathcal{A}_{n_m}^{(m)}\}$. Les $A_{n_m}^{(m)}$ seront appelées les *marques* des $\mathcal{A}_{n_m}^{(m)}$.

Récemment M. M. FRÉCHET dans ses belles recherches sur «Les probabilitées associées à un système d'événements compatibles et dépendants» a considéré des systèmes d'éventualités que nous dénoterons par $\mathcal{A}_1^{(1)}$, $\mathcal{A}_1^{(2)}$, ... L'éventualité complémentaire de $\mathcal{A}_1^{(m)}$, désignée par M. FRÉCHET par $E - A_m$ (où A_m correspond avec notre $\mathcal{A}_1^{(m)}$) étant dénotée par $\mathcal{A}_2^{(m)}$, les $\mathcal{A}^{(m)} = \{\mathcal{A}_1^{(m)}, \mathcal{A}_2^{(m)}\}$ sont des catégories consistant de deux éventualités. (Nous appellerons une telle catégorie une *alternative*.) En effet, M. FRÉCHET introduit une notation symbolique, *p. e.* $A_{\alpha_1} \dots A_{\alpha_r} (E - A_{\alpha_{r+1}}) \dots (E - A_{\alpha_m})$, (l. c. (1940) p. 26), correspondante avec notre $\prod_m A_{n_m}^{(m)}$.

5. Soit alors $\mathcal{H} = \{\mathcal{H}_1, \mathcal{H}_2, \dots\}$ un système exclusif d'hypothèses et soit $p_n = P\left[\dfrac{\mathcal{A}}{\mathcal{H}_n}\right]$ la probabilité d'une éventualité $\mathcal{A}$ sous l'hypothèse $\mathcal{H}_n$. Prenons une loterie Λ', dans laquelle les lots sont marqués de marques M_1, M_2, ..., formant un système exclusif, et soit X_n la probabilité que M_n sera tiré. Convenons maintenant de réaliser l'hypothèse $\mathcal{H}_n$, si et seulement si dans Λ' la marque M_n sera tiré. La probabilité qu'un tirage d'un lot de Λ' mènera à l'éventualité $\mathcal{A}$ est alors

$$(4) \qquad C = C(X) = C(X_1, \dots) = \Sigma^n X_n p_n.$$

Lorsque les X_n sont considérées comme des variables, indépendantes ou non, $C = C(X)$ sera appelée la *marque collective* du système de probabilités p_n; les X_n seront appelées les marques des $\mathcal{H}_n$.

Soit enfin $\mathcal{H} = \{\mathcal{H}^{(1)}, \mathcal{H}^{(2)}, \dots\}$ un système de systèmes exclusifs $\mathcal{H}^{(l)} = \{\mathcal{H}_1^{(l)}, \mathcal{H}_2^{(l)}, \dots\}$ d'hypothèses, et $\mathcal{A}$ un système de catégories comme dans le n° 4, et soit

$$p^{n_1, n_2, \dots}_{(k_1, k_2, \dots)} = P\left[\frac{\mathcal{A}_{n_1}^{(1)} \wedge \mathcal{A}_{n_2}^{(2)} \wedge \cdots}{\mathcal{H}_{k_1}^{(1)} \wedge \mathcal{H}_{k_2}^{(2)} \wedge \cdots}\right]$$

la probabilité que pour chaque m l'n_m-ième éventualité de la m-ième catégorie se présentera lorsque pour chaque l la k_l-ième hypothèse du l-ième système est réalisée.

Prenons un système de loteries $\Lambda_{n_m}^{(m)}$ comme dans le n° 4, et d'ailleurs un système de loteries $\Lambda'^{(l)}$ et dans chacune d'entre elles un système exclusif $M_1^{(l)}$, $M_2^{(l)}$, ... de marques, ayant les probabilités respectives $X_1^{(l)}$, $X_2^{(l)}$, ... Supposons que l'hypothèse $\mathcal{H}_{k_l}^{(l)}$ sera réalisée lorsque $M_{k_l}^{(l)}$ sera tirée dans $\Lambda'^{(l)}$. Alors la probabilité qu'aucune des «catastrophes» $\mathcal{E}^{(m)}$, considérées au n° 4 n'arrive est

$$(5) \qquad C = C(X, A) = \Sigma^{k_1, k_2, \dots} \Sigma^{n_1, n_2, \dots} \prod_l X_{k_l}^{(l)} \cdot p^{n_1, n_2, \dots}_{(k_1, k_2, \dots)} \prod_m A_{n_m}^{(m)}.$$

Les $A_{n_m}^{(m)}$ et les $X_{k_l}^{(l)}$ seront appelées des variables ou des marques de première et de seconde espèce respectivement; la fonction $C = C(X,A)$ sera dite la marque collective du système de probabilités $p^{n_1, \dots}_{(k_1, \dots)}$, ou aussi du système d'hypothèses et d'éventualités correspondantes.

6. On peut étendre les définitions au cas où les systèmes catégoriques et exclusifs ne sont pas dénombrables, mais sont des champs de probabilité (ou des sous-ensembles de ceux-ci) au sens de M. KOLMOGOROFF.

Pour pouvoir décrire ce cas général d'une manière nette, convenons de dénoter par f_λ la valeur qu'une fonction f, définie pour les éléments d'un ensemble Γ, prend sur l'élément λ, par F^Λ la valeur qu'une fonction additive (2) d'ensemble F, définie pour les sous-ensembles de Γ, prend sur l'ensemble Λ et par $(Ff)^\Lambda$ ou, plus explicitement par

$$\int_\Lambda F^{d\lambda} f_\lambda$$

la valeur, pourvu qu'elle existe, que l'intégrale de f par rapport à F prend sur l'ensemble Λ. Nous supposerons toujours et sans mention explicite que les ensembles Λ sont pris dans un corps fermé (3) de sous-ensembles, et que les fonctions f sont mesurables (4) sur ce corps. L'intégrale, définie comme

$$\lim_{\varepsilon \to 0} \Sigma^n F^{\Lambda_n \frown \Lambda} f_{\lambda_n}$$

où Λ_n est un système catégorique dénombrable de sous-ensembles de Γ tel que la variation $\sup_{\lambda \varepsilon \Lambda_n} f_\lambda - \inf_{\lambda \varepsilon \Lambda_n} f_\lambda$ de f sur Λ_n est $\leq \varepsilon$ pour chaque n, où $\frown$ dénote l'intersection et où $\lambda_n \varepsilon \Lambda_n \frown \Lambda$, existera en tout cas lorsqu'il existe un tel système pour au moins un $\varepsilon > 0$ tel que la somme

(2) Tout en omettant le prédicat «absolument» nous ne considérons que des fonctions d'ensemble dites «absolument additives».

(3) C'est-à-dire une famille contenant avec chaque sous-ensemble son complément, et avec chaque suite finie ou *infinie* de sous-ensembles son intersection (donc aussi son union).

(4) C'est-à-dire que les ensembles des λ où $\mathrm{Re} f_\lambda \leq x$ et $\mathrm{Im} f_\lambda \leq y$ appartiennent à ce corps pour chaque pair de nombres réels x et y, Re et Im dénotant la partie réelle et le coefficient de i de f_λ.

$\sum_n |F|^{\mathrm{A}_n \frown \mathrm{A}} \cdot |f_{\lambda_n}|$ est finie, $|F|^{\mathrm{A}} = \sup\limits_{\mathrm{M} \subset \mathrm{A}} F^{\mathrm{M}}$ étant la valeur que la différence des parties positives et négatives de F prend sur A.

Soit maintenant pour chaque m $\mathcal{A}^{(m)} = \{\alpha_m\}$ un ensemble arbitraire (non nécessairement dénombrable) catégorique d'éventualités et $\{\mathrm{A}_m\}$ un corps fermé de sous-ensembles de $\mathcal{A}^{(m)}$. Soit d'ailleurs pour chaque l $\mathcal{H}^{(l)} = \{\eta_l\}$ un ensemble exclusif arbitraire d'hypothèses et $\{\mathrm{H}_l\}$ un corps formé de sous-ensembles de $\mathcal{H}^{(l)}$. Supposons donné pour chaque suite d'éléments η_1, η_2, avec $\eta_l \varepsilon \mathcal{H}^{(l)}$ et chaque suite d'ensembles $\mathrm{A}_1, \mathrm{A}_2, \ldots$ avec $\mathrm{A}_m \subset \mathcal{A}^{(m)}$ la probabilité :

$$p_{(\eta)}^{\mathrm{A}} = p_{(\eta_1, \eta_2, \ldots)}^{\mathrm{A}_1, \mathrm{A}_2, \ldots} = P\left[\frac{\mathrm{A}_1 \times \mathrm{A}_2 \times \ldots}{\eta_1 \times \eta_2 \times \ldots}\right]$$

que l'ensemble des hypothèses η_l (considéré comme un élément $\eta = \eta_1 \times \eta_2 \times \ldots$ de l'ensemble produit $\mathcal{H} = \mathcal{H}^{(1)} \times \mathcal{H}^{(2)} \times \ldots$) amène l'éventualité A qui est la conjonction des A_m (considérée comme un sous-ensemble $\mathrm{A} = \mathrm{A}_1 \times \mathrm{A}_2 \times \ldots$ de l'ensemble produit $\mathcal{A} = \mathcal{A}^{(1)} \times \mathcal{A}^{(2)} \times \ldots$). Par conséquence $0 \leq p_{(\eta)}^{\mathrm{A}} \leq 1$ pour chaque A et chaque η est une fonction (absolument) additive de A (et de chaque A_m), et $p_{(\eta)}^{\mathcal{A}} = 1$ pour chaque η. Dans le cas où les $\mathcal{A}^{(m)} = (\mathcal{A}_1^{(m)}, \mathcal{A}_2^{(m)})$ sont des alternatives, une éventualité A dans l'ensemble produit est introduite par M. FRÉCHET (l. c., p. 26) sous le nom de « fonction des événements $\mathrm{A}_1^{(m)}$ ».

Prenons maintenant sur chaque $\mathcal{A}^{(m)}$ une fonction arbitraire $A^{(m)}$, ≥ 0 et ≤ 1, et considérons la valeur $A_{\alpha_m}^{(m)}$ sur un élément $\alpha_m \varepsilon \mathcal{A}^{(m)}$, comme la probabilité que la réalisation de l'éventualité $\mathcal{A}_{\alpha_m}^{(m)}$ ne même pas à une « catastrophe » $\mathcal{E}^{(m)}$, indépendamment de ces probabilités pour d'autres valeurs de m, tellement que $\mathrm{A}_\alpha = \prod\limits_m A_{\alpha_m}^{(m)}$ est la probabilité que la catastrophe $\mathcal{E}$, consistant à l'arrivée d'au moins une des catastrophes $\mathcal{E}^{(m)}$, n'arrive pas (sous l'hypothèse η). Chaque $A^{(m)}$ est bornée, donc intégrable par rapport à $p_{(\eta_1, \eta_2, \ldots)}^{\mathrm{A}_1, \mathrm{A}_2, \ldots}$ pour chaque suite $\eta = (\eta_1, \eta_2, \ldots)$.

A_α aussi est intégrable par rapport à $p_{(\eta)}^{\mathrm{A}}$ lui-même. Nous nous restreindrons toujours au cas où les intégrations répétées sont échangeables et équivalentes aux intégrations multiples.

Alors la probabilité que l'ensemble d'hypothèses η ne mène pas à la catastrophe $\mathcal{E}$ est

$$(5) \qquad C_{(\eta)} = C_{(\eta)}(A) = \int p_{(\eta)}^{d\alpha} A_\alpha = \iint \ldots p_{(\eta)}^{d\alpha_1, d\alpha_2 \ldots} A_{\alpha_1}^{(1)} A_{\alpha_2}^{(2)} \ldots$$

Prenons d'ailleurs pour chaque l sur $\mathcal{H}^{(l)}$ une fonction d'ensemble additive arbitraire $X_{(l)}^{\mathrm{H}_l}$, ≥ 0 et ≤ 1 et considérons-la comme la probabilité que (sous une hypothèse générale $\mathfrak{S}$) l'hypothèse H_l sera réalisée. Ici aussi nous nous restreignons au cas où $X^{\mathrm{H}_1 \times \mathrm{H}_2 \times \ldots} = X_{(1)}^{\mathrm{H}_1} X_{(2)}^{\mathrm{H}_2} \ldots$ détermine une fonction (absolument) additive sur les sous-ensembles de $\mathcal{H}$ et où toutes les intégrations répétées sont échangeables et équivalentes aux intégrations multiples.

Alors la probabilité $C^{\mathrm{A}} = C^{\mathrm{A}}(X)$ qu'une éventualité $\mathrm{A} = \mathrm{A}_1 \times \mathrm{A}_2 \times \ldots$ arrivera (sous l'hypothèse générale $\mathfrak{S}$) est

$$(6) \qquad C^{\mathrm{A}} = \int X^{d\eta} p_{(\eta)}^{\mathrm{A}} = \iint \ldots X_{(1)}^{d\eta_1} X_{(2)}^{d\eta_2} \ldots p_{(\eta_1, \eta_2, \ldots)}^{\mathrm{A}_1, \mathrm{A}_2, \ldots}.$$

D'ailleurs la probabilité

$$C = C(X, A) = P\left[\frac{\neg \mathcal{E}}{\mathfrak{S}}\right] \text{[5]}$$

que (sous l'hypothèse générale $\mathfrak{S}$) la catastrophe $\mathcal{E}$ n'arrivera pas est

$$(7) \qquad C = \int X^{d\eta} C_{(\eta)} = \iint X^{d\eta} p_{(\eta)}^{d\alpha} A_\alpha = \iint \ldots \iint \ldots X_{(1)}^{d\eta_1} X_{(2)}^{d\eta_2} \ldots p_{(\eta_1, \eta_2, \ldots)}^{d\alpha_1, d\alpha_2, \ldots} A_{\alpha_1}^{(1)} A_{\alpha_2}^{(2)} \ldots$$

Les fonctions $C_{(\eta)}$ (pour η constant) C^{A} (pour A constant) et C seront appelées les « marques collectives » du système de probabilités $p_{(\eta)}^{\mathrm{A}}$; les valeurs $A_{\alpha_m}^{(m)}$ et A_α des « marques (ou variables) de première espèce », et les valeurs $X_{(l)}^{\mathrm{H}_l}$ et X^{H} des « marques (ou variables) de seconde espèce ». Evidemment on aura toujours (pour $0 \leq X^{\mathrm{H}} \leq 1$, $0 \leq A_\alpha \leq 1$) : $0 \leq C(X, A) \leq X^{\mathcal{H}} A_{\mathcal{A}} \leq 1$, où $X^{\mathcal{H}}$ est la valeur de X sur l'ensemble $\mathcal{H}$ tout entier, tandis que $A_{\mathcal{A}}$ est la borne supérieure des valeurs A_α sur tout $\mathcal{A}$. D'ailleurs C est une fonction(nelle) non décroissante de tous ses arguments.

7. La fonction $C(X, A)$ étant donnée pour toutes les X^{H} et A_α admises, les $p_{(\eta)}^{\mathrm{A}}$, $C_{(\eta)}$ et C^{A} sont déterminées. Posons en effet, les λ et Λ étant des éléments et des sous-ensembles d'un ensemble arbitraire :

$$(8)^{(6)} \qquad \left|\begin{matrix}\Lambda \\ \lambda\end{matrix}\right. = \begin{cases} 1 & \text{si } \lambda \varepsilon \Lambda \\ 0 & \text{si non.} \end{cases}$$

Évidemment pour Λ constant la fonction $|^{\Lambda}$ (prenant les valeurs $|^{\Lambda}_{\lambda}$), est une fonction mesurable (par rapport au corps fermé d'ensembles d'où les Λ sont pris) des éléments λ ou « fonction de première espèce ». Elle est souvent appelée la *fonction caractéristique* de l'ensemble Λ. De même pour λ constant la fonction $|_{\lambda}$ (prenant aussi les valeurs $|^{\Lambda}_{\lambda}$) est une fonction (absolument) additive d'ensemble ou « fonction de seconde espèce »; on pourrait l'appeler la *fonction caractéristique* de l'élément λ.

Pour $\Lambda = \Lambda_1 \times \Lambda_2 \times \ldots$, $\lambda = \lambda_1 \times \lambda_2 \times \ldots$ on a évidemment :

$$(9) \qquad \left|\begin{matrix}\Lambda \\ \lambda\end{matrix}\right. = \left|\begin{matrix}\Lambda_1 \\ \lambda_1\end{matrix}\right. \cdot \left|\begin{matrix}\Lambda_2 \\ \lambda_2\end{matrix}\right. \ldots$$

---

[5] $\neg \mathcal{E}$ dénote la négation de $\mathcal{E}$.

[6] En (8), (9), etc. la ligne verticale désigne un I.

D'après les définitions et l'équivalence supposée des intégrations multiples et répétées on trouve aussitôt :

$$(10)\qquad C_{(\eta)}(A) = C(|_\eta, A).$$

$$(11)\qquad C^{A}(X) = C(X, |^{A}).$$

$$(12)\qquad p^{A}_{(\eta)} = C(|_\eta, |^{A}).$$

D'ailleurs, en définissant les opérations $\left(\frac{\partial}{\partial A}\right)^{A}$ et $\left(\frac{\partial}{\partial X}\right)_\eta$ par :

$$(13)\qquad \left(\frac{\partial f}{\partial A}\right)^{A} = \left(\frac{\partial}{\partial A}\right)^{A} f(A) = \lim_{\varepsilon\to 0} \frac{f(A+\varepsilon\,|^{A}) - f(A)}{\varepsilon}.$$

$$(14)\qquad \left(\frac{\partial f}{\partial X}\right)_\eta = \left(\frac{\partial}{\partial X}\right)_\eta f(X) = \lim_{\varepsilon\to 0} \frac{f(X+\varepsilon\,|_\eta) - f(X)}{\varepsilon}.$$

pourvu que ces limites existent, on a aussi :

$$(15)\qquad C_{(\eta)}(A) = \left(\frac{\partial}{\partial X}\right)_\eta C(X, A).$$

$$(16)\qquad C^{A}(X) = \left(\frac{\partial}{\partial A}\right)^{A} C(X, A).$$

$$(17)\qquad p^{A}_{(\eta)} = \left(\frac{\partial}{\partial X}\right)_\eta \left(\frac{\partial}{\partial A}\right)^{A} C(X, A) = \left(\frac{\partial}{\partial A}\right)^{A} \left(\frac{\partial}{\partial X}\right)_\eta C(X, A).$$

Pour $A = A_1 \times \dots$ ou $\eta = \eta_1 \times \dots$ on aura aussi, en supposant fini le nombre de systèmes catégoriques ou exclusifs :

$$(18)\qquad \left(\frac{\partial}{\partial A}\right)^{A_1\times A_2\times\cdots} C = \left(\frac{\partial}{\partial A^{(1)}}\right)^{A_1} \left(\frac{\partial}{\partial A^{(2)}}\right)^{A_2} \dots C.$$

$$(19)\qquad \left(\frac{\partial}{\partial X}\right)_{\eta_1\times\eta_2\times\cdots} C = \left(\frac{\partial}{\partial X^{(1)}}\right)_{\eta_1} \left(\frac{\partial}{\partial X^{(2)}}\right)_{\eta_2} \dots C.$$

Sans rechercher les conditions pour que ces relations restent valides pour des systèmes infinis, nous ne considérons que des cas où elles sont remplies.

§ 2. Opérations sur les marques collectives.

8. Les opérations principales qu'on peut appliquer aux marques collectives sont : le *démarcage*, la *substitution* et le *tirage*. Nous les considérons dans cet ordre, en laissant de côté dans chaque cas particulier les variables non concernées (par exemple les variables de seconde espèce lorsqu'il s'agit des variables de première espèce seulement).

Lorsque, en considérant $C = C(A) = \int p^{d\alpha} A_\alpha$, on ne s'intéresse pas à ceux des p^{A} où A est un vrai sous-ensemble de A_0, on peut restreindre la variabilité des A_α, en prenant la même loterie pour tous les $\alpha\varepsilon A_0$, c'est-à-dire en substituant pour tous les A_α avec $\alpha\varepsilon A_0$ une même variable A_{A_0}. Ce procédé, qu'on peut étendre à un système catégorique d'ensembles A_0, et qu'on peut appliquer ou bien aux $A = A_1 \times A_2 \times \dots$, ou bien aux A_n d'un rang défini n sera appelé un *démarcage simple de première espèce*. Nous nous bornerons au cas où le nombre des $\mathcal{A}^{(m)}$ et celui des $\mathcal{H}_{(l)}$ est fini.

Généralement, soit donné une image univoque de chaque $\mathcal{A}^{(m)}$ sur une catégorie $\mathcal{B}^{(m)}$, tel qu'avec chaque $\alpha_m \varepsilon \mathcal{A}^{(m)}$ corresponde un $\beta_m = \beta_m(\alpha_m)\, \varepsilon \mathcal{B}^{(m)}$ et un seulement, l'original d'un $\beta_m \varepsilon \mathcal{B}^{(m)}$ ou d'un $B_m \subset \mathcal{B}^{(m)}$ étant un $A_{m,\beta_m} \subset \mathcal{A}^{(m)}$ ou un $A_{m,B_m} \subset \mathcal{A}^{(m)}$ respectivement. Posons $A^{(m)}_{\alpha_m} = B^{(m)}_{\beta_m}$ pour chaque $\alpha_m \varepsilon A_{m,\beta_m}$, et soit $C'(B)$ le résultat de ces substitutions en $C(A)$. Alors, en omettant les m :

$$(20)\qquad \left(\frac{\partial}{\partial A}\right)^{A_B} = \left(\frac{\partial}{\partial B}\right)^{B}.$$

$$(21)\qquad q^{B} = p^{A_B}.$$

et

$$(22)\qquad C'(B) = \int q^{d\beta} B_\beta = C(A(B)) = \int p^{d\alpha} A_\alpha.$$

Le procédé correspondant, effectué sur les marques de second espèce s'applique lorsque les $p_{(\eta)}$ pour tous les η pris dans un sous-ensemble H_0 sont égales. Dans ce cas on ne définira X^{H} que pour des ensembles H qui ou bien contiennent H_0, ou bien ne contiennent aucun élément de H_0.

On peut même généraliser le procédé lorsqu'on se contente de connaître les valeurs moyens $p_{(H_0)}$ de tous les $p_{(\eta)}$ avec $\eta\varepsilon H_0$.

Généralement, soit donné une image univoque de chaque $\mathcal{H}^{(l)}$ sur un système exclusif $\mathcal{Z}^{(l)}$, tel qu'avec chaque $\eta_l \varepsilon \mathcal{H}^{(l)}$ corresponde un $\zeta_l \varepsilon \mathcal{Z}^{(l)}$ et un seul, l'original d'un $\zeta_l \,\varepsilon\, \mathcal{Z}^{(l)}$ ou d'un $Z_l \subset \mathcal{Z}^{(l)}$ étant un H_{l,ζ_l} ou $H_{l,Z_l} \subset \mathcal{H}^{(l)}$ respectivement. Soit pour chaque ζ_l $x^{H_l}_{\zeta_l}$ une fonction (absolument) additive et non négative sur les sous-ensembles H_l de H_{l,ζ_l} prenant la valeur 1 pour $H_l = H_{l,\zeta_l}$, et définissons :

$$(23)\qquad x^{H_l}_{\zeta_l} = x^{H_l \frown H\zeta_l}_{\zeta_l}.$$

pour un H_l arbitraire. Posons enfin :

$$(24)\qquad X^{H_l}_{(l)} = \int Y^{d\zeta_l}_{(l)} x^{H_l}_{\zeta_l}$$

les $Y^{Z_l}_{(l)}$ étant des variables de seconde espèce sur $\mathcal{Z}^{(l)}$, $\geqslant 0$ et $\leqslant 1$, et soit $C'(Y) = C(X(Y))$ le résultat de ces substitutions en $C(X)$. Alors, en omettant les l :

$$(25)\qquad \left(\frac{\partial}{\partial Y}\right)_\zeta = \int x^{d\eta}_\zeta \left(\frac{\partial}{\partial X}\right)_\eta$$

et

$$(26)\qquad C(X(Y)) = \int X^{d\eta} p_{(\eta)} = \iint Y^{d\zeta} x^{d\eta}_\zeta p_{(\eta)} = \int Y^{d\zeta} q_{(\zeta)}$$

où

$$(27)\qquad q_{(\zeta)} = \int x^{d\eta}_\zeta p_{(\eta)} = \int_{H\zeta} x^{d\eta}_\zeta p_{(\eta)}$$

sont des valeurs moyennes des $p_{(\eta)}$ sur H_ζ. Le procédé décrit par (23) — (27) sera appelé le *démarcage simple de seconde espèce*. Supposons que pour chaque (7) m il y a un parmi les sous-ensembles A_{m,β_m}, que nous dénoterons par $A_{m,0}$, dont on ne veut plus s'occuper. Alors on peut poser $B_0^{(m)} = 1$, c'est-à-dire on prendra la loterie correspondante tellement qu'elle ne mène en aucun cas à la catastrophe correspondante. Tandis que $C(A)$ et $C'(B)$ étaient encore des fonctions multilinéaires homogènes, elle devient maintenant inhomogène. En dénotant par α'_m les α_m n'appartenat pas à $A_{m,0}$ on aura :

$$C = \int\int \ldots p^{d\alpha'_1, d\alpha'_2, \cdots} A_{\alpha'_1} A_{\alpha'_2} \ldots + \int \ldots p^{A_{1,0}, d\alpha'_2, \cdots} A_{\alpha'_2} \ldots + \\ + \int \ldots p^{d\alpha'_1, A_{2,0}, \cdots} A_{\alpha'_1} \ldots + \ldots + p^{A_{1,0}, A_{2,0}, \cdots}.$$

De même, lorsqu'il y a parmi les H_{l,ζ_l} un p. e. $H_{l,0}$ dont on ne s'intéresse plus, on prendra $Y_{(l)}^{Z_l} = 0$ sur ζ_l. Alors les hypothèses appartenant aux $H_{l,0}$ n'apparaîtront plus du tout dans C. Ici la fonction multilinéaire reste homogène. Remarquons que la substitution équivaut à prendre $X_{(l)}^{H_{l,0}} = 0$, puisque $X_{(l)}^{H_l}$ étant $\geqslant 0$, celà amène que $X_{(l)}^{H_l} = 0$ pour chaque $H_l \subset H_{l,0}$.

Nous appellerons ce procédé le démarcage simple *complet* de première ou seconde espèce.

Montrons comment on peut représenter les fonctions génératrices des systèmes de probabilité introduits par M. Fréchet au moyen de notre $C\,(A_{\alpha_m}^{(m)})$. Soient les $\mathcal{A}^{(m)}$ des alternatives, comme dans la fin du numéro 4. Dénotons par $s_{m_1,\ldots,m_r}$ la probabilité que les éventualités $\mathcal{A}_1^{(m_1)}, \ldots, \mathcal{A}_1^{(m_r)}$, marquées $A_1^{(m_1)}, \ldots, A_1^{(m_r)}$ supposées toutes différentes se réalisent, sans considérer s'il en est ainsi quant aux $A_1^{(m)}$ pour $m \neq m_i$ $(1 \leqslant i \leqslant r)$, et par $q_{[m_1,\ldots m_r]}$ la probabilité que ces éventualités se réalisent sans qu'il en soit ainsi pour les $A_1^{(m)}$ avec $m \neq m_i$. Alors on voit aisément qu'en démarquant complètement les $A_2^{(m)}$ on obtient la fonction génératrice des $q_{[m_1,\ldots,m_r]}$. En omettant l'indice 1 on aura (cf. Fréchet, p. 31).

$$G\,(A^{(m)}) = C\,(A^{(m)},\ 1)$$
$$= \Sigma r\ \Sigma m_1, \ldots, m_r \quad \frac{1}{r!}\, q_{[m_1,\ldots,m_r]}\, A^{(m_1)} \ldots A^{(m_r)}.$$

Le facteur $\frac{1}{r!}$ doit être adjoint afin de pouvoir étendre la sommation indépendamment sur tous les $m_1, \ldots, m_r$. On pourrait aussi l'omettre et restreindre la sommation aux conditions $m_1 < m_2 < \ldots < m_r$. D'autre part on aura pour $\prod A_2^{(m)} \neq 0$:

$$C\,(A_1^{(m)},\ A_2^{(m)}) = \Pi A_2^{(m)} \cdot G\left(\frac{A_1^{(m)}}{A_2^{(m)}}\right)$$

c'est-à-dire $C\,(A_{n_m}^{(m)})$ se laisse représenter au moyen de la fonction génératrice $G(A^{(m)})$. La fonction obtenue par démarcage des $A^{(m)}$, voire $A^{(m)} = A$, est aussi introduite par M. Fréchet sous la dénotation $g\,(u)$.

La fonction génératrice des $s_{m_1,\ldots,m_r}$ devra être introduite avec des coefficients $(-1)^r/r!$. En dénotant les variables par B_m nous pouvons la présenter par :

$$H\,(B^{(m)}) = \Sigma r\ \Sigma m_1, \ldots, m_r \quad \frac{(-1)^r}{r!}\, s_{m_1,\ldots,m_r}\, B^{(m_1)} \ldots B^{(m_r)}.$$

Elle se laisse présenter au moyen des marques collectives C, comme suit (Cf. Fréchet, p. 32) :

$$H\,(B^{(m)}) = G\,(1 - B^{(m)}) = C\,(1 - B^{(m)},\ 1).$$

9. Supposons maintenant qu'il y ait des correspondances bi-univoques entre tous les $\mathcal{A}^{(m)}$, (8) de même qu'on peut faire parcourir tous les $\alpha_1, \alpha_2, \ldots$ un même ensemble B, les $A_1, A_2, \ldots$ parcourront les sous-ensembles de B. Supposons d'ailleurs qu'on ne s'intéresse qu'aux sous-ensembles auxquels les éventualités appartenant aux différentes catégories appartiennent, ***sans avoir regard au numéro d'ordre de cette catégorie***. C'est à dire qu'on ne considère au lieu de $p^{A_1, A_2, \cdots}$ que la somme de toutes les probabilités obtenues par permutations arbitraires des ensembles $A_1, A_2, \ldots$ entre eux. Dénotons cette somme par $p\,\{A_1, A_2, \ldots\}$.

On peut représenter le système de toutes ces probabilités en convenant de prendre une même loterie pour tous les m, c'est-à-dire en posant $A_\alpha^{(1)} = A_\alpha^{(2)} = \ldots = B_\alpha$. Par cette substitution $C(A^{(1)}, A^{(2)}, \ldots)$, qui est une fonction multilinéaire des $A_\alpha^{(1)}, \ldots$ passe (lorsque le nombre des catégories est fini) en un polynôme homogène $C(B)$ par rapport aux B_α. Les probabilités $p\,\{A\}$ ne peuvent plus être obtenues d'une manière simple par des substitutions du type (11), mais bien par des dérivations du type (16), (18). En effet, à part de la question de convergence dans le cas d'un nombre infini de catégories (dont nous ne nous occuperons pas ici), on aura :

$$(28) \qquad p^{\{A\}} = p^{\{A_1, A_2, \ldots\}} = \left(\frac{\partial}{\partial B}\right)^{A_1} \left(\frac{\partial}{\partial B}\right)^{A_2} \ldots C\,(B).$$

La combinaison de ce procédé avec celui décrit au numéro précédent sera appelé le *démarcage* (*général*) *de première espèce*.

Considérons un cas spécial. Considérons des catégories $\mathcal{A}^{(m)}$ arbitraires, et prenons de chaque catégorie un sous-ensemble arbitraire A_{m0}. Appliquons le démarcage simple par rapport à tous les $\mathcal{A}_{\alpha_m}^{(m)}$ avec $\alpha_m \varepsilon A_{m0}$, et complet par rapport à tous les autres $\mathcal{A}_{\alpha_m}^{(m)}$, c'est-à-dire mettons $A_{\alpha_m}^{(m)} = B^{(m)}$ pour $\alpha_m \varepsilon A_m$ et

$$A_{\alpha_m}^{(m)} = 1 \text{ pour } (\neg\ \alpha_m \varepsilon A_m).$$

(7) Le procédé s'applique de même lorsqu'il n'y a lieu que pour quelques-uns parmi les m.

(8) Il est possible d'appliquer le même procédé lorsqu'il y a des correspondances entre des *sous-ensembles* $\mathcal{A}'^{(m)}$ de *quelques-uns des* $\mathcal{A}^{(m)}$ en se bornant à ceux-ci.

Appliquons le démarcage général en omettant tous les indices m, c'est-à-dire :

$$(29) \qquad A^{(m)}_{\alpha_m} = \begin{cases} B & \alpha_m \varepsilon \mathrm{A}_{m0} \\ 1 & \alpha_m \varepsilon \mathrm{A}_{m0} \end{cases}$$

ce qui peut aussi s'écrire :

$$A^{(m)}_{\alpha_m} = (B - 1) \Big|^{\mathrm{A}_{m0}}_{\alpha_m} + 1.$$

Par ces substitutions $C(A)$ passe en une fonction analytique de B :

$$(30) \qquad C(A(B)) = \Sigma^r p_{[r]} B^r,$$

le coefficient $p_{[r]}$ dénotant la probabilité qu'il y a exactement r entre les catégories, parmi lesquels un événement se réalise qui appartient au sous-ensemble A_{m0}.

Nous nous passerons du procédé analogue qu'on peut appliquer aux marques de seconde espèce.

10. Nous ne discuterons pas complètement l'opération de substitution. Nous nous bornerons, au contraire, à quelques remarques. Évidemment on peut toujours substituer une marque collective pour une marque arbitraire de première espèce, puisque cela ne revient qu'à un choix spécial de la manière dont on détermine la probabilité d'une «catastrophe». Or, lorsqu'on veut éviter des démarcages nouveaux, il sera nécessaire en général de choisir un ensemble de variables *différent* pour chaque marque collective qu'on substitue pour une marque. Considérons par exemple le cas de marques collectives linéaires. Soit $C = \int p^{d\alpha} A_\alpha$. Lorsque nous voulons substituer pour A_α $C_\alpha = C_\alpha(B_\beta)$, on devra donc prendre pour les B_β des marques différentes pour différents α, c'est-à-dire on substituera $C_\alpha = C_\alpha(B_{\alpha\beta})$, où les β parcourront un ensemble dépendant d'α ou non. Pourvu alors que C_α ne soit pas constante (et, comme nous supposerons, dépende d'une manière continue de ses variables), on ne perd rien de généralité. En effet, en choisissant $B_{\alpha\beta} = B_\alpha$, indépendamment de β, C_α, devient une fonction continue, non constante et non décroissante de B_α; on peut donc choisir la valeur de B_α tellement que C_α prend la valeur A_α, qui peut parcourir un intervalle. Lorsque spécialement $C_\alpha(B_\beta) = \int q^{d\beta}_\alpha B_\beta$ aussi est linéaire (et homogène), on prendra $B_{\alpha\beta} = A_\alpha . C_\alpha(1)^{-1}$ où $C_\alpha(1)$ est la valeur que C_α prend lorsque $B_\beta = 1$ pour chaque β. Lorsque d'autre part :

$$C_\alpha(B_\beta) = q_\alpha + \int q^{d\beta}_\alpha B_\beta + \int q^{d\beta_1, d\beta_2}_\alpha B_{\beta_1} B_{\beta_2} + \dots$$

on peut obtenir pour C_α une valeur arbitraire A_α pourvu qu'elle soit $\geqslant C_\alpha(0) = q_\alpha$ et $\leqslant C_\alpha(1)$. Or, en démarquant d'abord $B_{\alpha\beta} = B_\beta$, on obtiendrait dans le premier cas $C = \int p^{d\alpha} \int q^{d\beta}_\alpha B_\beta = \int r^{d\beta} B_\beta$ avec $r^{\mathrm{B}} = \int p^{d\alpha} q^{\mathrm{B}}_\alpha$, et on ne pourrait pas en général regagner les p^{A}.

Lorsque C n'est pas linéaire on opère d'une même manière. En effet, soit par exemple $C = \iint p^{d\alpha_1, d\alpha_2} A^{(1)}_{\alpha_1} A^{(2)}_{\alpha_2}$ et supposons que α_1 et α_2 parcourent un même ensemble et que $p^{\mathrm{A}_1, \mathrm{A}_2} = p^{\mathrm{A}_2, \mathrm{A}_1}$, de manière qu'on ne perd rien de généralité en posant $A^{(1)}_\alpha = A^{(2)}_\alpha = A_\alpha$. Or, en substituant alors $A_\alpha = C_\alpha(B_\beta) = \iint q^{d\beta_1, d\beta_2}_\alpha B_{\beta_1} B_{\beta_2}$, on ne pourrait plus distinguer entre le paire d'éventualités $(B_{\beta_1} B_{\beta_2}, B_{\beta_3} B_{\beta_4})$ et le paire $(B_{\beta_1} B_{\beta_3}, B_{\beta_2} B_{\beta_4})$. Mais en substituant $A_\alpha = C_\alpha(B_{\alpha\beta})$ on ne rencontre pas cette difficulté, puisque les produits $B_{\alpha_1\beta_1} B_{\alpha_1\beta_2} B_{\alpha_2\beta_3} B_{\alpha_2\beta_4}$ et $B_{\alpha_1\beta_1} B_{\alpha_1\beta_3} B_{\alpha_2\beta_2} B_{\alpha_2\beta_4}$ sont différents excepté pour $\alpha_1 = \alpha_2$, un cas dans lequel il n'y a rien à perdre.

On peut considérer comme un cas spécial d'une telle substitution le remplacement d'une marque dépendante de plusieurs indices, par exemple $A_{\alpha\beta\gamma}$ par un produit, par exemple $A^{(1)}_\alpha A^{(2)}_\beta A^{(3)}_\gamma$. Cette substitution ne diminue en aucun degré la généralité de la représentation jusqu'à ce qu'on substitue la marque collective en une autre où elle figure d'une manière non linéaire. En effet, de :

$$C_0 = \Sigma^{\alpha, \beta, \gamma \dots} \int p^{d\alpha, d\beta, d\gamma, \dots} A^{(1)}_\alpha A^{(2)}_\beta A^{(3)}_\gamma \dots$$

on déduit :

$$p^{\mathrm{A}, \mathrm{B}, \Gamma \dots} = C_0(|^{\mathrm{A}}, |^{\mathrm{B}}, |^{\Gamma}, \dots),$$

donc :

$$C = \Sigma^{\alpha, \beta, \gamma, \dots} \int p^{d\alpha, d\beta, d\gamma, \dots} A_{\alpha\beta\gamma\dots} = \int C_0(|^{d\alpha}, |^{d\beta}, |^{d\gamma}, \dots) A_{\alpha\beta\gamma\dots}$$

Mais dans un produit, par exemple C_0^2, on ne pourrait plus distinguer par exemple entre $(A^{(1)}_\alpha A^{(2)}_\beta \dots)$ $(A^{(1)}_{\alpha'} A^{(2)}_{\beta'} \dots)$ et $(A^{(1)}_{\alpha'} A^{(2)}_\beta \dots)$ $(A^{(1)}_\alpha A^{(2)}_{\beta'} \dots)$. Le remplacement de $A_{\alpha\beta\dots}$ par $A^{(1)}_\alpha A^{(2)}_\beta \dots$ doit donc être considéré comme un *démarcage* spécial. Le remplacement réciproque, qu'on ne peut pas obtenir par une substitution simple, peut être considéré comme un *marcage* (ou *rémarcage*). Un cas encore plus spécial est celui, où une suite de marques A_n, par exemple $C = \sum p_n A_n$ est remplacée par une suite de puissances $C_0 = \sum p_n A^n$. Étant donné C, on obtient C_0 par une substitution simple : $A_n = A^n$. Mais étant donné C_0 on ne peut regagner C qu'ou bien par des différentiations :

$$C = \Sigma^n \frac{1}{n!} A_n \left[\left(\frac{\partial}{\partial A} \right)^n C_0(A) \right]_{A=1}$$

ou bien par une intégration :

$$C = \int C_0(A)\, dF(A)$$

ou $F(A)$ est choisie tellement que :

$$A_n = \int A^n dF(A).$$

Cette condition n'assujettit A_n qu'à des inégalités, de sorte que les A_n restent des variables indépendantes quoique dans un intervalle restreint. Quant aux marques de seconde espèce, on voudra substituer aux X^{H} des marques collectives C^{H}, dépendant d'une manière additive des ensembles H. Nous nous bornerons aux substitutions du type spécial :

$$C^{\mathrm{H}}(Y) = \iint\limits_{\mathrm{H}} Y^{d\eta, d\zeta} q_{\eta\zeta} = \iint Y^{d\eta, d\zeta} q_{\eta\zeta} \Big|_{\eta}^{\mathrm{H}}$$

où aussi :

$$C^{\mathrm{H}}(Y) = \int\limits_{\mathrm{H}} U^{d\eta} \int V_{\eta}^{d\zeta} q_{\eta\zeta},$$

c'est-à-dire $Y^{\mathrm{H,Z}} = \int_{\mathrm{H}} U^{d\eta} V_{\eta}^{\mathrm{Z}}$. En supposant que pour chaque η $q_{\eta\zeta}$ n'évanouit pas identiquement, donc qu'il y a une fonction $\zeta = \zeta(\eta)$ avec $q_{\eta,\zeta(\eta)} \neq 0$, on peut prendre $V_{\eta}^{\mathrm{Z}} = \Big|_{\zeta(\eta)}^{\mathrm{Z}} \cdot q_{\eta,\zeta(\eta)}^{-1}$, $U^{\mathrm{H}} = X^{\mathrm{H}}$ afin d'obtenir :

$$C^{\mathrm{H}}(Y) = \int X^{d\eta} \Big|_{\eta}^{\mathrm{H}} \int \Big|_{\zeta(\eta)}^{d\zeta} \cdot q_{\eta,\zeta(\eta)}^{-1} q_{\eta\zeta} = X^{\mathrm{H}}.$$

Le démarcage $B_{\alpha\beta} = B_{\beta}$ correspondrait ici avec $V_{\eta}^{\mathrm{Z}} = V^{\mathrm{Z}}$, donc avec $Y^{\mathrm{H,Z}} = U^{\mathrm{H}} V^{\mathrm{Z}}$, $C^{\mathrm{H}}(Y) = C^{\mathrm{H}}(U, V)$ $= \int_{\mathrm{H}} U^{d\eta} \int V^{d\zeta} q_{\eta\zeta}$. Seulement lorsqu'on pourrait trouver un V^{Z} indépendant de η, tel que $k_{\eta} = \int V^{d\zeta} q_{\eta\zeta} \neq 0$ pour chaque η, et ayant une inverse intégrable par rapport à X, on pourrait poser $U^{\mathrm{H}} = \int_{\mathrm{H}} X^{d\eta} k_{\eta}^{-1}$ et l'on obtiendrait alors aussi $C^{\mathrm{H}}(Y) = X^{\mathrm{H}}$.

11. Au moyen d'une suite de substitutions successives, on peut représenter aisément une suite d'expériments dépendants d'une manière quelconque des précédents.

Soit $p_{(\alpha_1, \dots \alpha_k)}^{\mathrm{A}_{k+1}}$ la probabilité que le $(k+1)$-ième expériment mène à une éventualité appartenante à $\mathcal{A}_{k+1}$, supposé que les expériments antérieurs ayent donné successivement les résultats $\alpha_1, \dots, \alpha_k$. La marque collective du $(k+1)$-ième expériment sous cette supposition sera alors :

$$(31) \qquad C_{(\alpha_1, \dots, \alpha_k)}(A_{\alpha_{k+1}}^{(k+1)}) = \int p_{(\alpha_1, \dots, \alpha_k)}^{d\alpha_{k+1}} A_{\alpha_{k+1}}^{(k+1)}.$$

L'ensemble parcouru par α_{k+1} peut dépendre de $\alpha_1, \dots, \alpha_k$, mais peut aussi être choisi indépendant d'eux, en joignant les ensembles pour tous les systèmes $\alpha_1, \dots \alpha_k$ possibles, et prenant $p_{(\alpha_1, \dots, \alpha_k)}^{\mathrm{A}_{k+1}}$ zéro hors du sous-ensemble correspondant avec les hypothèses $\alpha_1, \dots, \alpha_k$.

Remplaçant pour les k augmentants :

$A_{\alpha_{k+1}}^{(k+1)}$ dans $C_{(\alpha_1, \dots, \alpha_k)}$ par $A_{\alpha_{k+1}}^{(k+1)} \cdot C_{(\alpha_1, \dots, \alpha_{k+1})}(A_{\alpha_{k+2}}^{(k+2)})$,

on obtient successivement :

$$C_1 = C_1(A^{(1)}) = C(A_{\alpha_1}^{(1)}) = \int p^{d\alpha_1} A_{\alpha_1}^{(1)},$$

$$C_2 = C_2(A^{(1)}, A^{(2)}) = C(A_{\alpha_1}^{(1)} C_{(\alpha_1)}(A_{\alpha_2}^{(2)})) = \int p^{d\alpha_1} A_{\alpha_1}^{(1)} \int p_{(\alpha_1)}^{d\alpha_2} A_{\alpha_2}^{(2)},$$

$$C_3 = C_3(A^{(1)}, A^{(2)}, A^{(3)}) = C(A_{\alpha_1}^{(1)} C_{(\alpha_1)}(A_{\alpha_2}^{(2)} C_{(\alpha_1\alpha_2)}(A_{\alpha_3}^{(3)}))) = \\ = \int p^{d\alpha_1} A_{\alpha_1}^{(1)} \cdot \int p_{(\alpha_1)}^{d\alpha_2} A_{\alpha_2}^{(2)} \int p_{(\alpha_1\alpha_2)}^{d\alpha_3} A_{\alpha_3}^{(3)}.$$

etc., ou généralement :

$$(32) \qquad C_k = \int \dots \int p^{d\alpha_1, \dots, d\alpha_k} A_{\alpha_1}^{(1)} \dots A_{\alpha_k}^{(k)},$$

avec :

$$(33) \quad p^{\mathrm{A}_1, \dots, \mathrm{A}_k} = \int\limits_{\mathrm{A}_1} p^{d\alpha_1} \int\limits_{\mathrm{A}_2} p_{(\alpha_1)}^{d\alpha_2} \dots \int\limits_{\mathrm{A}_k} p_{(\alpha_1, \dots, \alpha_{k-1})}^{d\alpha_k} = \\ = \int \dots \int p^{d\alpha_1} p_{(\alpha_1)}^{d\alpha_2} \dots p_{(\alpha_1, \dots, \alpha_{k-1})}^{d\alpha_k} \Big|_{\alpha_1}^{\mathrm{A}_1} \dots \Big|_{\alpha_k}^{\mathrm{A}_k}.$$

c'est-à-dire la probabilité que les expériments mènent à des événements appartenants à $\mathrm{A}_1, \dots, \mathrm{A}_k$ successivement.

En décidant chaque fois avec une probabilité auxiliaire X si l'on fera l'expériment suivant ou non, on obtient la marque collective de la suite des expériments toute entière :

$$(34) \quad C = C(X, A^{(1)}, A^{(2)}, \dots) = \sum_{0}^{\infty} {}_k X^k (1 - X) C_k$$

où l'on posera $C_0 = 1$.

Remarquons encore que la forme (34) à marques factorisées peut être trop spéciale. Dans ce cas on la devra « remarquer » de la manière mentionnée au numéro 10.

12. On peut appliquer la représentation (34) à une foule de problèmes. Mentionnons, par exemple, le problème de la première épreuve favorable [9]. Démarquons d'abord pour chaque k complètement les marques des éventualités défavorables, c'est-à-dire mettons $A_{\alpha_k}^{(k)} = 1$ pour un α_k défavorable et mettons $A_{\alpha_k}^{(k)} = A_k$ pour un α_k favorable.

En posant alors $A_k = 0$, C_k deviendra la probabilité d'une suite de k expériments consécutifs (dès le commencement) étant pris tous et ayant tous un résultat défavorable. Donc $C_k - C_{k+1}$ devient la probabilité q_{k+1} que, $k+1$ expériments étant pris, le dernier est le premier ayant un résultat favorable. Or, en lotissant après chaque résultat défavorable avec une probabilité X constante et indépendante des résultats des lotissements précédents, qu'une catastrophe ne se présente pas, la

(9) Cf. M. Fréchet, l. c., p. 24.

probabilité qu'aucune catastrophe n'arrivera sera :

$$(35)\qquad \sum_0^\infty {}_k q_{k+1} X^{k+1} = \sum_0^\infty {}_k (C_k - C_{k+1})\, X^{k+1} =$$
$$= 1 - \sum_0^\infty {}_k X^k (1 - X)\, C_k = 1 - C.$$

D'une part, elle est la fonction génératrice ordinaire selon LAPLACE des q_{k+1} ; d'autre part, elle est le complément par rapport à 1 de la fonction obtenue de (34) au moyen des démarcages mentionnés.

De la même manière, on peut résoudre le problème de la r-ième épreuve favorable. Or, on peut obtenir ce résultat d'une manière directe, simultanément pour tous les r.

En effet, posons sous les conditions du commencement de n° 12 $A_k = A$. Alors C_k peut s'écrire :

$$C_k = \sum_0^k {}_k p_{kr} A^r,$$

où p_{kr} est la probabilité qu'on a obtenu précisément r épreuves favorables parmi les k (premiers) expériments. Donc (34) devient :

$$(36)\qquad C = \sum_0^\infty {}_k X^k (1 - X) \sum_0^k {}_r p_{kr} A^r.$$

Interprétons cette probabilité. Supposons que X soit la probabilité de jeter pile avec une pièce de monnaie, 1-X celle de jeter face. Dès qu'on jette face, le jeu finit. Toujours lorsqu'on jette pile, on joue le coup suivant du jeu principal, déterminé par (31); lorsqu'on obtient un résultat défavorable, disons $\mathcal{B}$, on recommence de jeter pile ou face. Or, lorsqu'on obtient un résultat favorable, disons on jette une seconde pièce de monnaie, avec une probabilité A d'obtenir, disons par distinction Pile et la probabilité 1-A d'obtenir Face. Ce dernier résultat sera la «catastrophe». Nous savons déjà que C est la probabilité qu'elle ne se présente pas.

Donc 1-C est la probabilité totale que la catastrophe se présente. Comment est-ce qu'elle peut se présenter? Le dernier résultat sera d'obtenir Face (prob. = 1-A.) Le résultat précédent doit avoir été une épreuve favorable, disons la r-ième ($r \geqslant 1$). Lorsqu'il y a eu k épreuves défavorables, la probabilité totale, quant au jeu (31), de ce résultat sera celle que nous cherchons, disons $q_{r,k+r}$. Or, auparavant on doit avoir obtenu r-1 fois Pile (prob. = A^{r-1}), après chaque épreuve favorable précédente, et $k + r$ fois pile (prob. = X^{k+r}), puisqu'on a joué $k + r$ fois, sans jamais obtenir face. Donc :

$$(37)\qquad 1 - C = \sum_1^\infty {}_r A^{r-1} (1 - A) \sum_0^\infty {}_k q_{r,k+r} X^{k+r}$$

sera la marque collective du système des $q_{r,k+1}$ avec $r \geqslant 1$. En substituant $A = 0$ en (37), c'est-à-dire en prenant le coefficient de $A^0(1-A)$, on retombe sur (35).

En incluant le cas $r = 0$ avec $q_{00} = 1$, $q_{0k} = 0$ pour $k > 0$, on obtient pour la marque collective, disons C^* du système de tous les $q_{r,k+r}$, y compris q_{0k} :

$$(37')\qquad C^* = \sum_0^\infty {}_r A^r (1 - A) \sum_r^\infty {}_n q_{rn} X^n =$$
$$= (1 - A) + A(1 - C) = 1 - AC.$$

Remarquons que les marques de première et seconde espèce y sont échangées, en comparaison avec (36).

Appliquons ce résultat à une suite d'événements indépendants. Soient q et $p = 1$-q les probabilités d'une épreuve favorable $\mathcal{A}$ et défavorable $\mathcal{B}$ respectivement.

Alors (32) devient :

$$(38)\qquad C_k = (pB + qA)^k,$$

c'est-à-dire la marque collective de la collection de BERNOULLI du rang k. La marque collective C devient :

$$(39)\qquad C = C(X; A, B) = \sum_0^\infty {}_k X^k (1 - X) C_k = \frac{1 - X}{1 - X(pA + qB)}.$$

La marque collective (ici la fonction génératrice) C^* du rang de la r-ième épreuve favorable est d'après (37) :

$$(40)\qquad C^* = \sum_0^\infty {}_r A^r (1 - A) \sum_r^\infty {}_n q_{rn} X^n =$$
$$= 1 - A \frac{1 - X}{1 - X(pA + q)} = \frac{(1 - A)(1 - qX)}{1 - pXA - qX}$$

donc :

$$\sum_r^\infty {}_n q_{rn} X^n = \left(\frac{pX}{1 - qX}\right)^r$$

ou :

$$(41)\qquad \sum_0^\infty {}_k q_{r,k+r} X^k = \left(\frac{1 - q}{1 - qX}\right)^r$$

d'où le résultat bien connu de PASCAL (pour $q = 1/2$), de MONTMORT et DE MOIVRE :

$$(42)\qquad q_{r,k+r} = \binom{k + r - 1}{k} p^r q^k.$$

Remarquons d'ailleurs qu'on peut obtenir (39) d'une manière directe. Considérons plus généralement un champs de probabilité p^A, représenté par $C_1 = \int p^{d\alpha} A_\alpha$. C_1 est la probabilité qu'un seul tirage n'amène pas de catastrophe. Supposons un nombre arbitraire de tirages faits sans qu'une catastrophe soit arrivée. Alors la probabilité d'une catastrophe doit encore être C. D'autre part elle consiste de la probabilité $1\text{-}X$ qu'on ne tire plus (auquel cas la catastrophe ne peut pas arriver conformément à ce qui est convenu) et de la probabilité X qu'on tire, ce qui amène la probabilité C_1 que la catastrophe n'arrive pas. Alors on a de nouveau la probabilité C qu'elle n'arrivera pas plus tard. Donc :

$$C = (1 - X) + XC_1C$$

ou :

$$(43) \qquad C = \frac{1 - X}{1 - X\int p_{d\alpha}A^{\alpha}}$$

ce qui généralise (39). Cette méthode directe peut être appliquée lorsqu'on réussit à introduire un système de probabilités auxiliaires (marques) tellement qu'une même situation se reproduit toujours de nouveau.

13. Passons maintenant à l'opération de *tirage*. Partons d'une boîte, contenant ν objets (éléments), dont pour chaque λ $(1 \leq \lambda \leq k)$ ν_λ possèdent une marque A_λ. Supposons que chaque objet possède une marque et une seulement, donc $\sum^\lambda \nu_\lambda = \nu$, $\nu_\lambda \geq 0$. En supposant égales les chances de tirer l'un quelconque de ces objets et zéro la chance de tirer quelqu'autre objet, la probabilité de tirer un élément à marque A_λ est ν_λ/ν.

Représentons le contenu de la boîte symboliquement par :

$$(44) \qquad C_0 = \prod_\lambda A_\lambda^{\nu_\lambda}.$$

On peut interpréter ce symbole comme ci-dessus, en supposant que chaque élément dans la boîte avant que les tirages commencent peut mener indépendamment de chaque autre à une «catastrophe», avec une probabilité $1\text{-}A_\lambda$; C_0 sera la probabilité qu'aucune catastrophe n'arrive.

Lorsqu'un A_λ est tiré, le symbole C_0 devient

$$C_0A_\lambda^{-1} = \frac{1}{\nu_\lambda}\frac{\partial}{\partial A_\lambda}C_0.$$

La probabilité de tirer un tel élément étant $\nu_\lambda\nu^{-1}$, le produit devient :

$$\nu_\lambda\nu^{-1}A_\lambda^{-1}C_0 = \frac{1}{\nu}\frac{\partial}{\partial A_\lambda}C_0.$$

Nous devons encore tenir compte : 1° de l'élément tiré lui-même; 2° des modifications qu'on peut introduire à la collection d'éléments dans la boîte. Quant à 1°, nous supposerons que l'élément tiré peut mener à une catastrophe avec une probabilité modifiée, disons $1\text{-}A'_\lambda$. La modification du contenu de la boîte peut être représentée par un opérateur Ω'_λ, opérant sur la collection restante $\frac{1}{\nu}\frac{\partial}{\partial A_\lambda}C_0$. Après le tirage la marque collective est donc devenue :

$$(45) \qquad C_1 = \nu^{-1}\sum^\lambda A'_\lambda\,\Omega'_\lambda\frac{\partial}{\partial A_\lambda}C_0$$

ou aussi $C_1 = \Omega' C_0$ avec :

$$(46) \qquad \Omega' = \nu^{-1}\sum^\lambda A'_\lambda\,\Omega'_\lambda\frac{\partial}{\partial A_\lambda}.$$

La forme de l'opérateur Ω'_λ sera déterminée par la nature de la modification. Lorsque par exemple on laisse la collection inaltérée, on prendra $\Omega'_\lambda = 1$, donc :

$$(46.\text{A}) \qquad \Omega' = \nu^{-1}\sum^\lambda A'_\lambda\frac{\partial}{\partial A_\lambda}.$$

Lorsque l'objet tiré est remis, on doit réinstituer le facteur perdu A_λ; on posera donc $\Omega'_\lambda = A_\lambda$ et :

$$(46.\text{B}) \qquad \Omega' = \nu^{-1}\sum^\lambda A'_\lambda A_\lambda\frac{\partial}{\partial A_\lambda} = \nu^{-1}\sum^\lambda A'_\lambda\frac{\partial}{\partial \ln A_\lambda}.$$

Lorsque l'objet tiré est remis, ensemble avec d_λ objets portants la même marque, $\nu_\lambda - 1$ doit passer en $\nu_\lambda + d_\lambda$, tandis que ν_μ pour $\mu \neq \nu$ doit rester le même. On aura donc $\Omega'_\lambda = A_\lambda^{d_\lambda+1}$, donc :

$$(46.\text{C}) \qquad \Omega' = \nu^{-1}\sum^\lambda A_\lambda A_\lambda^{d_\lambda+1}\frac{\partial}{\partial A_\lambda} = \sum^\lambda -\frac{d_\lambda}{\nu}A'_\lambda\frac{\partial}{\partial A_\lambda^{-d_\lambda}}.$$

Naturellement, on peut aussi considérer des modifications plus compliquées, où Ω'_λ ne peut pas être représentée par une multiplication simplement.

Supposons maintenant que le nombre des éléments dans la boîte après le tirage sera ν_1, indépendant de la marque de l'élément tiré (brièvement : de la marque tirée). Dans les deux premiers exemples cela sera le cas avec $\nu_1 = \nu - 1$ et $\nu_1 = \nu$ respectivement; dans la troisième exemple seulement si $d_1 = \ldots = d_k\ (= d)$, et bien avec $\nu_1 = \nu + d$ (tirage d'Eggenberger-Polya). (Dans ce cas C_1 sera un polynôme homogène de degré 1 par rapport aux A'_λ). Pour un deuxième tirage, on peut appliquer un tel opérateur comme Ω' à chaque terme de ce polynôme. En supposant qu'on peut prendre pour les Ω''_λ des opérateurs *linéaires*, on les peut appliquer à la marque collective C_1 toute entière. On obtiendra alors :

$$C_2 = \Omega_2 C_0 = \Omega''\Omega' C_0$$

avec :

$$\Omega'' = \frac{1}{\nu_1}\sum^\lambda A''_\lambda\,\Omega''_\lambda\frac{\partial}{\partial A'_\lambda}.$$

En continuant de cette manière, on obtiendra après n tirages :

$$(47) \qquad C_n = \Omega C_0 = \Omega^{(n)}\Omega^{(n-1)}\ldots\Omega' C_0.$$

$$(48) \qquad \Omega^{(h)} = \frac{1}{\nu_{h-1}}\sum^\lambda A_\lambda^{(h)}\Omega_\lambda^{(h)}\frac{\partial}{\partial A_\lambda}.$$

Plus généralement, on peut admettre que les $A_\lambda^{(h)}$ dépendent des marques déjà tirées, donc qu'on ait $A^{(h)}_{(\lambda_1,\ldots,\lambda_{h-1})\lambda_h}$

au lieu de $A_\lambda^{(h)}$. Dans ce cas la marque d'une suite de tirages $A_\lambda^{(1)}, \ldots, A_{\lambda_n}^{(n)}$ devient :

$$B^{(1,\ldots,n)}_{\lambda_1,\ldots,\lambda_n} = \prod_h A^{(h)}_{(\lambda_1,\ldots,\lambda_{h-1})\lambda_h}$$

Naturellement, on peut admettre aussi que les $\Omega_\lambda^{(h)}$ dépendent d'une même manière de $\lambda_1, \ldots, \lambda_{h-1}$.

Lorsqu'on ne considère que les nombres des marques tirées sans se soucier de leur succession, on peut démarquer les $A_\lambda^{(h)}$ par rapport à h : $A_\lambda^{(h)} = A'_\lambda$ (mais non $= A_\lambda$). Lorsque d'ailleurs $\Omega_\lambda^{(h)}$ est indépendante de h, on aura :

$$\Omega_n = \frac{1}{\prod\limits_{h=1}^{n} \nu_{h-1}} \left(\sum{}^\lambda A'_\lambda \Omega'_\lambda \frac{\partial}{\partial A_\lambda} \right)^n. \tag{49}$$

Dans les cas (46 A, B, C), le dernier pour $d_\lambda = d$, on aura successivement :

$$\Omega_n = \frac{1}{\nu^{!n}} \left(\sum{}^\lambda A'_\lambda \frac{\partial}{\partial A_\lambda} \right)^n. \tag{49.A}$$

$$\Omega_n = \frac{1}{\nu^{!n}} \left(\sum{}^\lambda A'_\lambda \frac{\partial}{\partial \ln A_\lambda} \right)^n. \tag{49.B}$$

$$\Omega_n = \frac{1}{\left(-\frac{\nu}{d}\right)^{!n}} \left(\sum{}^\lambda A'_\lambda \frac{\partial}{\partial A^{-d}} \right)^n \tag{49.C}$$

où nous avons posé :

$$x^{!n} = x(x-1)\ldots(x-n+1) \tag{50}$$

de manière que :

$$\left(-\frac{\nu}{d}\right)^{!n} = (-d)^{-n} \nu(\nu+d)\ldots(\nu+nd-d).$$

Remarquons encore que la forme (47) à marques factosirées $\prod_h A_{\lambda_h}^{(h)}$ peut être trop spéciale. Alors on pourra les «rémarquer» de la manière mentionnée au n° 10. Nous rencontrerons un exemple dans le paragraphe 3.

La méthode décrite jusqu'ici a encore le désavantage l'être restreinte aux polynômes homogènes. Cela n'est nécessitée que par les dénominateurs, par exemple ν en (46). Or, comme on a remplacé le numérateur ν_λ par la différentiation $\frac{\partial}{\partial A_\lambda}$, on peut de même remplacer le dénominateur ν par une intégration. Introduisons le symbôle :

$$S^{B_\lambda}_{A_\lambda} = S^{B_1}_{A_1} \ldots S^{B_k}_{A_k} \tag{51}$$

signifiant le remplacement de chaque A_λ par B_λ. Alors C_0 étant donné par (44), $\nu^{-1}\, C_0$ peut s'écrire :

$$\nu^{-1} C_0 = \int X^{-1} dX S^{A_\lambda X}_{A_\lambda} C_0 = \int_0^\infty dt S^{A_\lambda e^{-t}}_{A_\lambda} C_0. \tag{52}$$

Or, l'opérateur dans le second membre restant le même pour chaque degré ν, on peut l'appliquer à chaque terme d'un polynôme, homogène ou non. Pourvu que les opérateurs Ω'_λ restent linéaires, on aura donc généralement (47) avec :

$$\Omega^{(h)} = \sum{}^\lambda A_\lambda^{(h)} \Omega_\lambda^{(h)} \frac{\partial}{\partial A_\lambda} \int_0^1 X^{-1} dX S^{A_1 X}_{A_1} \ldots S^{A_n X}_{A_n}. \tag{53}$$

En particulier (49) devient :

$$\Omega_n = \left\{ \sum{}^\lambda A'_\lambda \Omega'_\lambda \frac{\partial}{\partial A_\lambda} \int_0^1 X^{-1} dX S^{A_1 X}_{A_1} \ldots S^{A_n X}_{A_n} \right\}^n. \tag{54}$$

Cet opérateur peut aussi être appliqué au cas (46 C) si d_λ est dépendant de ν, quoiqu'on ne voit pas encore qu'on puisse en tirer beaucoup de profit, à cause de la non-commutativité des termes.

§ 3. Application. Le problème des itérations.

14. La méthode des marques collectives introduite jusqu'ici n'est au fond qu'une élaboration de la méthode classique des fonctions génératrices. Les différences principales entre notre méthode et l'ancienne consistent en une interprétation probabilistique des variables et en une augmentation de leur nombre. La première différence n'est point du tout essentielle. On peut, mais on ne doit pas, interpréter les variables de la manière susdite. Cette interprétation a le désavantage de restreindre les variables à l'intervalle unité de l'axe réel, une restriction qui ne peut pas être maintenue toujours. Par contre elle nous procure l'avantage de pouvoir parfois déterminer la marque collective d'une manière directe, sans devoir recourir aux équations à différences finie, comme le fait Laplace. Un exemple d'une telle construction directe a été donné à la fin du n° 12 ; d'autres suivront ci-dessous. L'introduction des variables auxiliaires en grand nombre a le désavantage que les démarcages nécessaires ne sont pas déterminés uniquement par le problème donné; ils peuvent être choisis parfois de plusieurs manières, et leur choix dépend du degré de simplicité qu'ils admettent. C'est-à-dire qu'on retombe sur des difficultés d'un type malfamé dans l'instruction élémentaire, celui des lignes auxiliaires. D'autre part, elle a le grand avantage de permettre de poser le problème avec tous ses degrés de liberté ce qui admet la résolution simultanée d'un grand nombre de questions.

Nous montrerons ces avantages à l'exemple du problème des itérations.

15. Considérons ici seulement les itérations Bernoulliennes. Partons d'un système catégorique de marques, ayant les probabilités p_λ, $1 \leq \lambda \leq k$, $p_\lambda \geq 0$, $\sum p_\lambda = 1$. Dénotant les marques par A_λ, nous obtenons comme marque collective :

$$C_1 = \sum{}^\lambda p_\lambda A_\lambda. \tag{55}$$

Les A_λ correspondent aux A'_λ du n° 13, les A_λ originaux ayant été démarqués complètement, mais nous n'avons pas besoin de ces résultats-là.

Une suite de n tirages consécutifs peut être représentée par un produit $A^{(1)}_{\lambda_1} \ldots A^{(n)}_{\lambda_n}$.

Appelons une suite partielle $A^{(h+1)}_{\lambda_{h+1}} \ldots A^{(h+l)}_{\lambda_{h+l}}$ avec $\lambda_{h+1} = \ldots = \lambda_{h+l}$ (disons $= \lambda$) une *itération de longueur l et de marque* λ, λ servira à abréviation de A_λ, λl à abréviation de $I_{\lambda l}$; appelons la *complète,* lorsque ou bien $h = 0$ ou bien $\lambda \neq \lambda_h$, et ou bien $h + l = n$ ou bien $\lambda \neq \lambda_{h+l+1}$. Dans ce cas nous la dénoterons par la marque $I_{\lambda l}$. La suite considérée peut donc être représentée par :

$$I^{(1)}_{\lambda_1 l_1} \ldots I^{(r)}_{\lambda_r l_r}$$

avec

$$l_1 + \ldots + l_r = n,\ l_i \geq 1,\ (1 \leq i \leq r),\ \lambda_i \neq \lambda_{i+1} (1 \leq i \leq r-1).$$

Il suffit de soumettre les itérations à la condition d'être *semi-complète,* c'est-à-dire que ou bien $h + l = n$, ou bien $\lambda \neq \lambda_{h+l+1}$. Toutes les itérations étant semi-complètes, elles sont complètes aussi. Nous ne voulons nous occuper que du nombre d'itérations de chaque type, sans nous soucier de l'ordre de leur succession. Nous pourrons donc démarquer les indices supérieures, en posant $I^{(h)}_{\lambda l} = I_{\lambda l}$. Afin d'interpréter ces marques considérons les probabilités $A^{(h)}_{(\lambda_1, \ldots, \lambda_{h-1})\lambda_h}$, introduites au n° 13. Supposons-les choisies indépendantes des indices $\lambda_1, \ldots, \lambda_{g-1}$ toujours quand $\lambda_g = \lambda_{g+1} = \ldots = \lambda_{h-1} = \lambda \neq \lambda_{g-1}$. Alors le produit

$$\mathrm{B}^{(1,\ldots,m)}_{\lambda_1,\ldots,\lambda_m} = \prod_{h=1}^{m} A^{(h)}_{(\lambda_1,\ldots,\lambda_{h-1})\lambda_h}$$

devient un produit $I_{\lambda_1 l_1} \ldots I_{\lambda_r l_r}$ avec $\lambda_i \neq \lambda_{i+1}$, où :

$$I_{\lambda_k l_k} = \prod_{h=1}^{l_k} A^{(m_{k-1}+h)}_{(\lambda_k^{h-1})\lambda_h},$$

où $m_{k-1} = l_1 + \ldots + l_{k-1}$, tandis qu'une suite d'indices égaux a été représentée symboliquement par une puissance. On peut spécialiser les $A^{(h)}_{(\lambda_1,\ldots,\lambda_{h-1})\lambda_h}$ encore plus, pourvu qu'elles restent dépendantes du numéro d'ordre de l'itération dans laquelle λ_k paraît.

Il s'agit alors de déterminer la probabilité

$$P_{(n)m_{11},\ldots,m_{k1};\ldots;m_{1l},\ldots,m_{kl}} = P_{(n)m_{\lambda l}}$$

qu'une suite de n tirages consécutifs se résout en une combinaison, contenant pour chaque λ et chaque l précisément $m_{\lambda l}$ itérations $I_{\lambda l}$. Évidemment :

$$(56) \qquad \sum_{1}^{k} {}^{\lambda} \sum_{1}^{f} {}^{l}\, l m_{\lambda l} = n, \qquad m_{\lambda l} \geq 0$$

de sorte que f reste borné.

Supposons de nouveau qu'on joue à pile ou face, avec une probabilité T pour pile, et que chaque fois qu'on obtient pile ou tire un élément de la catégorie (qu'on remet alors) et que le jeu finit dès qu'on a jeté face pour la première fois.

La marque collective C du système des $P_{(n)m\lambda}$ sera alors :

$$(57) \qquad C = \sum_{0}^{\infty} {}^{n}\, T^n (1 - T) \sum P_{(n)m_{\lambda l}} \prod_{\lambda} \prod_{l} I_{\lambda l}^{m_{\lambda l}},$$

où la seconde somme s'étend sur toutes les valeurs de $m_{11}, \ldots, m_{kf}$, satisfaisant (56). Le problème consiste en déterminer C comme fonction de toutes ses variables, voir T et les $I_{\lambda l}$. Une fois cette fonction trouvée on peut la développer par rapport à toutes les variables qu'on voudra.

16. Ce problème peut être résolu complètement. A cette fin commençons par «rémarquer» les produits $\prod_h A^{(h)}_{\lambda_h}$ (cf. n° 10) en les remplaçant par $\mathrm{A}_{\lambda_1,\ldots,\lambda_n}$. Cette marque sera interprétée comme la probabilité qu'une suite de tirages $\lambda_1, \ldots, \lambda_n$ mènera à une catastrophe déterminée. Factorisons maintenant cette probabilité en un produit de probabilités indépendantes, interprétons $I_{\lambda l}$ comme la probabilité, à un instant arbitraire donné que, *lorsqu'*au moins l tirages seront faits encore et lorsqu'une itération semi-complète de marque λ et de longueur l suivra, une catastrophe déterminée n'arrivera pas. Toutes ces probabilités sont supposées mutuellement indépendantes et l'on suppose que la catastrophe ne peut pas avoir lieu par une autre cause. Alors C est simplement la probabilité, avant un jet arbitraire avec la monnaie, qu'aucune de ces catastrophes n'arrivera. En effet $T^n(1\text{-}T)$ est la probabilité qu'exactement n tirages seront faits, $P_{(n)m_{\lambda l}}$ celle qu'ils seront résolubles en le système d'itérations spécifié par les nombres $m_{\lambda l}$, et $\prod I_{\lambda l}^{m_{\lambda l}}$ celle que ce système n'aboutira pas en une catastrophe. Puisque la probabilité que la suite de tirages fusse infinie est zéro à cause de $T < 1$ (évidemment on ne peut pas admettre ici $T = 1$), le système d'éventualités est exclusif et complet par rapport à la possibilité d'aucune des catastrophes considérées, donc C est la probabilité totale d'aucune telle. Évaluons-là d'une autre manière.

$I_{\lambda l}$ est la probabilité de manque de catastrophe sous la condition : 1° qu'on tire au moins l fois; 2° qu'on obtient une itération de marque λ et de longueur l; 3° qu'elle soit semi-complète. La probabilité pour que 1° soit remplie est T^l, celle pour 2° est p_λ^l. La condition 3° sera remplie lorsqu'ou bien les tirages finissent à cause de l'apparition de face (prob. $1 - T$), ou bien les tirages continuent (prob. T), mais une autre marque apparaît (prob. $1 - p_\lambda$). Donc la probabilité de la condition 3° est $(1 - T) + T(1 - p_\lambda) = 1 - Tp_\lambda$. Donc la probabilité pour une $I_{\lambda l}$ sans catastrophe est $(Tp_\lambda)^l I_{\lambda l}$, et celle pour une itération $\mathcal{S}_\lambda$ semi-complète de marque λ et de longueur $\geqslant 1$ quelconque (sans catastrophe) est :

$$(58) \qquad S_\lambda = \sum_{1}^{\infty}{}^{l} (Tp_\lambda)^l I_{\lambda l} (1 - Tp_\lambda).$$

Soit d'ailleurs R_λ la probabilité d'une itération $\mathcal{R}_\lambda$ de marque λ et de longueur arbitraire ($\geqslant 1$), semi-complète ou non, sans catastrophe. Or $\mathcal{S}_\lambda$ peut arriver parce que d'abord $\mathcal{R}_\lambda$ arrive et qu'ou bien elle se révèle comme une semi-complète (prob. $1-Tp_\lambda$), ou bien une itération semi-complète de la même marque (sans catastrophe) suit (prob. S_λ). Donc :

$$(59) \qquad S_\lambda = R_\lambda (1 - Tp_\lambda) + R_\lambda S_\lambda$$

c'est-à-dire :

$$(60) \qquad R_\lambda = \frac{S_\lambda}{1 - Tp_\lambda + S_\lambda} = \frac{\varphi_\lambda}{1 + \varphi_\lambda}$$

où :

$$(61) \qquad \varphi_\lambda = \frac{S_\lambda}{1 - Tp_\lambda} = \sum_{1}^{\infty}{}^{l} (Tp_\lambda)^l I_{\lambda l}.$$

La relation (60), équivalente à :

$$S_\lambda = \frac{R_\lambda}{1 - Tp_\lambda - R_\lambda}$$

entraîne la relation entre les $I_{\lambda l}$ et les marques collectives $J_{\lambda l}$ des itérations complètes ou non, pour lesquelles :

$$R_\lambda = \sum_{1}^{\infty}{}^{l} (Tp_\lambda)^l (1 - Tp_\lambda) J_{\lambda l} :$$

$$I_{\lambda l} = \sum{}^{l_1, \ldots, l_k} J_{\lambda l_1} \ldots J_{\lambda l_k}, \quad J_{\lambda l} = \sum{}^{l_1, \ldots, l_k} I_{\lambda l_1} \ldots I_{\lambda l_k},$$

les sommations étant étendues à tous les $k, l_1, \ldots, l_k$ avec $l_i \geqslant 1$ et $\sum l_i = l$ $(i = 1, \ldots, k)$.

Contraire à R_λ, S_λ, etc. φ_λ n'est pas une probabilité.

Soit enfin R la probabilité d'une itération (sans catastrophe), semi-complète ou non, de marque et de longueur ($\geqslant 1$) quelconques. Donc :

$$(62) \qquad R = \sum{}^{\lambda} R_\lambda.$$

Une suite de tirages (en nombre $\geqslant 0$) sans catastrophe (prob. C) peut provenir ou bien d'un jet de face (prob. $1 - T$), ou bien d'une itération de longueur $\geqslant 1$ sans catastrophe (prob. R), suivie d'une suite nouvelle sans catastrophe (prob. C). Donc :

$$C = (1 - T) + RC$$

ou

$$(63) \qquad C = \frac{1 - T}{1 - R}.$$

En substituant (58), (60), (62) en (64) on obtient enfin la fonction cherchée (57) :

$$(64) \qquad \boxed{C = \frac{1 - T}{1 - \sum{}^{\lambda} \dfrac{\sum^{l} (Tp_\lambda)^l I_{\lambda l}}{1 + \sum^{l} (Tp_\lambda)^l I_{\lambda l}}}}.$$

17. De la formule (64) [ou aussi de (58), (60), (62), (64)] ensemble avec (57) on peut obtenir nombre d'applications. Remarquons d'abord que démarcage complet des $I_{\lambda l}$ donne naturellement :

$$(65) \qquad S_\lambda = Tp_\lambda, \; R_\lambda = Tp_\lambda, \; R = T, \; \varphi_\lambda = \frac{Tp_\lambda}{1 - Tp_\lambda},$$

$$1 + \varphi_\lambda = \frac{1}{1 - Tp_\lambda},$$

donc $C = 1$.

D'abord on obtient en développant (64) :

$$\frac{C}{1 - T} = \sum_{0}^{\infty}{}^{s} \left(\sum{}^{\lambda} \frac{\varphi_\lambda}{1 + \varphi_\lambda}\right)^s = \sum_{0}^{\infty}{}^{s_1} \cdots \sum_{0}^{\infty}{}^{s_k} \left(\sum{}^{\lambda} s_\lambda\right)! \prod_\lambda \frac{1}{s_\lambda!} \left(\frac{\varphi_\lambda}{1 + \varphi_\lambda}\right)^{s_\lambda} =$$

$$= \sum{}^{s_1, \ldots s_k, t_1, \ldots, t_k} \left(\sum{}^{\lambda} s_\lambda\right)! \prod_\lambda \frac{(-1)^{t_\lambda}}{s_\lambda!} \binom{s_\lambda + t_\lambda - 1}{t_\lambda} \varphi_\lambda^{s_\lambda + t_\lambda} =$$

$$= \sum{}^{s_\lambda \, m_{\lambda l}} \left(\sum{}^{\lambda} s_\lambda\right)! \prod_\lambda \left[\frac{(-1)^{n_\lambda - s_\lambda}}{s_\lambda!} \binom{n_\lambda - 1}{s_\lambda - 1} . m_\lambda \prod_l \frac{\{(Tp_\lambda)^l I_{\lambda l}\}^{m_{\lambda l}}}{m_{\lambda l}}\right]$$

où

$$\sum{}^{l} m_{\lambda l} = s_\lambda + t_\lambda = m_\lambda.$$

Donc, en comparant les coefficients avec ceux de (57) :

$$(66) \qquad P_{(n) m_{\lambda l}} = \sum_{1}^{m_1}{}^{s_1} \cdots \sum_{1}^{m_k}{}^{s_k} \left(\sum{}^{\lambda} s_\lambda\right)! \sum_{\lambda=1}^{h} \left[\frac{(-1)^{m_\lambda - s_\lambda} (m_\lambda - 1)! \, m_\lambda!}{s_\lambda! \, (s_\lambda - 1)! \, (m_\lambda - s_\lambda)! \prod_l m_{\lambda l}} \cdot p_\lambda^{n_\lambda}\right]$$

$$\left(m = \sum{}^{l} m_{\lambda l}, \; n = \sum{}^{l} l m_{\lambda l}\right).$$

Puisque $n = \sum^{\lambda} n_\lambda$ est donné et $l \geqslant 1$, seulement un nombre fini des valeurs de l apparaissent dans la formule.

Les $m_{\lambda l}$ étant les exposants des $I_{\lambda l}$, leurs moments factoriaux sont obtenus par dérivations successives par rapport aux $I_{\lambda l}$, suivies d'un démarcage complet de ces marques.

Or :

$$\frac{\partial C}{\partial I_{\lambda l}} = \frac{1-T}{(1-R)^2}\frac{\partial R}{\partial I_{\lambda l}} = \frac{1-T}{(1-R)^2}\frac{1}{(1+\varphi_\lambda)^2}\frac{\partial \varphi_\lambda}{\partial I_{\lambda l}} = \frac{(1-T)(Tp_\lambda)^l}{(1-R)^2(1+\varphi_\lambda)^2}. \tag{67}$$

$$\frac{\partial^2 C}{\partial I_{\lambda l}\,\partial I_{\mu m}} = \frac{2(1-T)(Tp_\lambda)^l(Tp_\mu)^m}{(1-R)^3(1+\varphi_\lambda)^2(1+\varphi_\mu)^2}; \quad (\lambda \neq \mu). \tag{68}$$

$$\frac{\partial^2 C}{(\partial I_{\lambda l})\,\partial I_{\lambda m}} = \frac{2(1-T)(Tp_\lambda)^l(Tp_\lambda)^m}{(1-R)^3(1+\varphi_\lambda)^4} - \frac{2(1-T)(Tp_\lambda)^l(Tp_\lambda)^m}{(1-R)^2(1+\varphi_\lambda)^3} = \frac{2(1-T)(Tp_\lambda)^{l+m}}{(1-R)^3(1+\varphi_\lambda)^4}((1+\varphi_\lambda)R-\varphi_\lambda). \tag{69}$$

Par démarcage complet on obtient tout de suite au moyen de (66), $\mathcal{E}$ dénotant l'expectation totale :

$$\mathcal{E}\, m_{\lambda l} = \left(\frac{\partial C}{\partial I_{\lambda l}}\right)_1 = \frac{(Tp_\lambda)^l}{1-T}(1-Tp_\lambda)^2.$$

$$\mathcal{E}\, m_{\lambda l}m_{\mu m} = \left(\frac{\partial^2 C}{\partial I_{\lambda l}\,\partial I_{\mu m}}\right) = \frac{2(Tp_\lambda)^l(Tp_\mu)^m(1-Tp_\lambda)^2(1-Tp_\mu)^2}{(1-T)^2}; \quad (\lambda \neq \mu).$$

$$\mathcal{E}\, m_{\lambda l}m_{\lambda m} = \frac{2(Tp_\lambda)^{l+m}(1-Tp_\lambda)^3\,T(1-p_\lambda)}{(1-T)^2}; \quad (l \neq m).$$

La dernière expression pour $l = m$ donne $\mathcal{E}m_{\lambda l}^{[2]}$ au lieu de $\mathcal{E}m_{\lambda l}^2$. Les expectations conditionnelles pour un n donné, dénotées par $\mathcal{E}_{(n)}$ sont obtenues simplement par développement, par rapport à T, des membres gauches, *divisés par* $1-T$. En effet, à cause de $C = \sum T^n(1-T)C_{(n)}$ on a aussi :

$\frac{\mathcal{E}x}{1-T} = \sum T^n\,\mathcal{E}_{(n)}\,x$ pour chaque variable stochastique x.

On obtient, en supposant n suffisamment grand (voire $n \geqslant l+1$ en 71.1 et $n \geqslant l+m+2$ dans les autres cas), et en posant :

$$q_\lambda = 1 - p_\lambda : \tag{70}$$

$$\left(\frac{1-Tp_\lambda}{1-T}\right)^2 = \left(\frac{1-p_\lambda}{1-T}+p_\lambda\right)^2 = p_\lambda^2 + 2p_\lambda q_\lambda \sum T^m + q_\lambda^2 \sum (m+1)\,T^m$$

$$\left(\frac{1-Tp_\lambda}{1-T}\right)^3 = p_\lambda^3 + 3p_\lambda^2 q_\lambda \sum T^n + 3p_\lambda q_\lambda^2 \sum (n+1)\,T^n + q_\lambda^3 \sum \frac{(n+1)(n+2)}{2}\,T^n$$

donc :

$$\mathcal{E}_{(n)}m_{\lambda l} = \alpha = p_\lambda^l\,\{2p_\lambda q_\lambda + (n-l+1)\cdot q_\lambda^2\} = p_\lambda^l q_\lambda\,\{2 + (n-l-1)\,q_\lambda\}. \tag{71.1}$$

$$\mathcal{E}_{(n)}m_{\lambda l}m_{\lambda m} - \delta_{lm}\mathcal{E}_{(n)}m_{\lambda l} = p_\lambda^{l+m}q_\lambda^2\,\{6p_\lambda^2 + 6p_\lambda q_\lambda(n-l-m) + q_\lambda^2(n-l-m)(n-l-m+1)\} \tag{71.2}$$

donc :

$$\begin{aligned}\mathcal{E}_{(n)}(m_{\lambda l}-\mathcal{E}_{(n)}m_{\lambda l})(m_{\lambda m}-\mathcal{E}_{(n)}m_{\lambda m}) - \delta_{lm}\mathcal{E}_{(n)}m_{\lambda l} &= \\ = -p_\lambda^{l+m}q_\lambda^2\,[q_\lambda^2\,\{n(l+m+1)-(l^2+lm+m^2)+1\} &- 2p_\lambda q_\lambda\{n-(l+m)\} - 2p_\lambda^2].\end{aligned} \tag{71.3}$$

Pour l'écart quadratique $\sigma^2 = \sigma_{\lambda l}^2$ on trouve donc :

$$\sigma^2 - \alpha = -p_\lambda^{2l}q_\lambda^2\,[q_\lambda^2\,\{n(2l+1)-3l^2+1\} - 2p_\lambda q_\lambda(n-2l) - 2p_\lambda^2]. \tag{71.4}$$

De (71.3) et (71.4) on peut déduire aisément le coefficient de corrélation entre $m_{\lambda l}$ et $m_{\lambda m}$. D'une manière analogue on trouvera le coefficient de corrélation entre $m_{\lambda l}$ et $m_{\mu m}$ avec $\lambda \neq \mu$.

Lorsqu'on ne veut considérer que les itérations d'un seul type λl, on démarquera les autres : $I_{\mu m} = 1$, excepté pour $\mu = \lambda$ et $m = l$. Alors $S_\mu = R_\mu = Tp_\mu$ pour $\mu \neq \lambda$, tandis que :

$$S_\lambda = Tp_\lambda - (1-Tp_\lambda)\,J_{\lambda l}, \quad J_{\lambda l} = (Tp_\lambda)^l(1-I_{\lambda l}). \tag{72}$$

$$R_\lambda = Tp_\lambda - \frac{(1-Tp_\lambda)^2 J_{\lambda l}}{1-J_{\lambda l}(1-Tp_\lambda)} \tag{73}$$

et :

$$C = \frac{1-T}{1-R} = \frac{1}{1 + \frac{1}{1-T}\,\frac{(1-Tp_\lambda)^2 J_{\lambda l}}{1 - J_{\lambda l}(1-Tp_\lambda)}} = \frac{(1-T)\{1 - J_{\lambda l}\cdot(1-Tp_\lambda)\}}{1 - T + J_{\lambda l}\cdot Tq_\lambda\,(1-Tp_\lambda)}. \tag{74}$$

En développant on trouve, en omettant les indices λ et l :

$$\frac{C}{1-T} = \{1 - J(1-Tp)\}\sum_0^\infty{}^m \frac{(-JTq)^m}{1-T}\left(\frac{1-Tp}{1-T}\right)^m =$$

$$= \{1 - J(1-Tp)\}\sum_0^\infty{}^m \sum_0^m{}^h \sum_0^\infty{}^g \binom{m}{h}\binom{h+g}{g}(-J)^m (Tq)^{m+h}T^g =$$

$$= \{1 - (1-Tp)(Tp)^l(1-I)\}\sum{}^{m,h,g}\sum{}^f \binom{m}{h}\binom{h+g}{g}\binom{m}{f}(-1)^{m+f}p^{lm}q^{m+h}I^f T^{ml+m+h+g}.$$

Donc la probabilité qu'il y ait précisément f itérations du type λl dans une suite de n tirages est approximativement égale à :

$$\sum_0^\infty{}^m \sum_0^m{}^h (-1)^{m+f}\binom{m}{h}\binom{m}{f}\binom{n-ml-m}{h}p^{lm}q^{m+h} \tag{75}$$

où nous avons négligé le terme $J(1\text{-}Tp) \leqslant (Tp)^l$ par rapport à 1.

Pour trouver la distribution de n'importe quel type on démarquera (incomplètement!) tous les $I_{\lambda l}$: $I_{\lambda l} = I$. On obtient :

$$\varphi_\lambda = \frac{Tp_\lambda I}{1 - Tp_\lambda}$$

$$R_\lambda = \frac{Tp_\lambda I}{1 - Tp_\lambda(1-I)} = Tp_\lambda I\left(1 + \frac{Tp_\lambda(1-I)}{1 - Tp_\lambda(1-I)}\right)$$

$$\frac{C}{1-T} = \frac{1}{1 - TI - I(1-I)T^2\sum{}^\lambda \frac{p_\lambda^2}{1 - Tp_\lambda(1-I)}}. \tag{76}$$

En nous bornant, pour simplifier, au cas $p_1 = \ldots = p_k = p$, le développement devient :

$$\frac{C}{1-T} = \frac{1}{1 - TI - \frac{I(1-I)T^2p}{1 - Tp(1-I)}} = \frac{1 - Tp(1-I)}{1 - T(p+qI)} \tag{77}$$

$$= \{1 - Tp + TpI\}\cdot\sum_0^\infty{}^m T^m \sum_0^m{}^h \binom{m}{h}p^{m-h}q^h I^h.$$

Donc la probabilité qu'une suite de n tirages contienne précisément r itérations est :

$$\binom{n}{r}p^{n-r}q^r - \binom{n-1}{r}p^{n-r}q^r + \binom{n-1}{r-1}p^{n-r+1}q^{r-1} = \binom{n-1}{r-1}p^{n-r}q^{r-1}, \tag{78}$$

formule bien connue pour $k = 2$ $\left(p = q = \frac{1}{2}\right)$. La distribution de la longueur moyenne des itérations formant une suite, c'est-à-dire de $\frac{n}{r}$ ne peut être déterminée qu'approximativement. Elle a été considérée par L. von Bortkiewicz.

Pour le cas $k = 2$, $p_1 = p$ arbitraire, $p_2 = q = 1 - p_1$, (76) devient :

$$\frac{C}{I-T} = 1 + \frac{TI - 2pqT^2I(1-I)}{I - T + T^2pq(1-I^2)} = \tag{79}$$

$$= 1 + TI(1 - 2pqI + 2pqTI)\sum_0^\infty{}^s \frac{(pqT^2I^2)^s}{(1-pT)^{s+1}(1-qT)^{s+1}} =$$

$$= 1 + TI(1 - 2pqT + 2pqTI)\sum_0^\infty{}^s \sum_0^\infty{}^k \sum_0^\infty{}^l \binom{s+k}{s}\binom{s+l}{s}(pT)^{s+k}(qT)^{s+l}I^{2s}$$

d'où la probabilité $P_{(n)r}$ qu'une suite de n tirages contienne précisément r itérations :

$$P_{(n+2),\,2(s+1)} = 2pq\sum_0^{n-2s}{}^k \binom{s+k}{s}\binom{n-s-k}{s}p^{s+k}q^{n-s-k};\ n \geqslant 2s \geqslant 0) \tag{80}$$

$$P_{(n+1),\,2s+1} = \sum_0^{n-2s}{}^k \binom{s+k}{s}\binom{n-s-k}{s}p^{s+k}q^{n-s-k} - 2pq\sum_0^{n-2s-1}{}^k \binom{s+k}{s}\binom{n-1-s-k}{s}p^{s+k}q^{n-1-s-k} =$$

$$= \frac{1}{2pq}P_{(n+2),\,2(s+1)} - P_{(n+1),\,2(s+1)};\ (n \geqslant 2s \geqslant 0).$$

18. Au moyen de la méthode expliquée au n° 12 nous pouvons déterminer la valeur moyenne du nombre de tirages, nécessaire afin d'obtenir un système d'itérations de types choisis arbitrairement. Dénotons la marque collective de ce système par A et la valeur moyenne choisie par $E_{(A)} n$. A cause de (37′) on aura :

$$(81) \qquad \left(\frac{\partial C^*}{\partial T}\right)_{T=1} = \sum_{0}^{\infty}{}^{r} A^r (1 - A)\, \mathcal{E}_{(A^r)}\, n = - A \left(\frac{\partial C}{\partial T}\right)_{T=1}$$

$\frac{C}{1-T}$ étant développée par rapport aux puissances de T et de A. Évidemment $\mathcal{E}_{(A^r)}\, n = 0$ pour $r = 0$; en divisant par $A(1 - A)$ et substituant $A = 0$ on obtient :

$$(82) \qquad \mathcal{E}_{(A)}\, n = - \left(\frac{\partial C}{\partial T}\right)_{T=1,\, A=0}$$

relation qu'on aurait pu obtenir aussi de (35). Or :

$$\left(-\frac{\partial C}{\partial T}\right)_{T=1} = \left(\frac{1}{1-R}\right)_{T=1}$$

puisque $C = 0$ pour $t = 1$, le dénominateur étant > 0 dès qu'au moins un $I_{\lambda l}$ n'est pas démarqué complètement. Donc :

$$(83) \qquad \mathcal{E}_{(A)}\, n = \left(\frac{1}{1-R}\right)_{T=1,\, A=0}$$

Considérons par exemple $A = I_{\lambda l}^{r}\, I_{\lambda m}^{s}$ avec $l \neq m$. On substitue dès le commencement $T = 1$ et l'on obtient avec $p = p_\lambda$, $q = 1 - p$, $I_{\lambda l} = I_l$, $I_{\lambda m} = I_m$, tous les autres I étant $= 1$:

$$\varphi_\lambda = \frac{p}{q} - p^l - p^m + I_l p^l + I_m p^m,$$

$$\frac{1}{1-R} = \frac{1}{q} \frac{\frac{1}{q} - p^l - p^m + I_l p^l + I_m p^m}{p^l + p^m - (I_l p^l + I p^m)} =$$

$$= \frac{1}{q} \left\{ \left(\frac{1}{q} - p^l - p^m\right) + (I_l p^l + I_m p^m) \right\} \sum_{0}^{\infty} \frac{(I_l p^l + I_m p^m)^h}{(p^l + p)^{h+1}}.$$

Ici on doit après développement substituer zéro pour tous les termes contenant le facteur $A = I_l^r\, I_m^s$ et 1 pour tous les autres ; le résultat sera $\mathcal{E}_{(A)}\, n$. Pour $s = 0$, donc $A = I_{\lambda l}^{r}$ le résultat devient simplement :

$$\frac{1}{1-R} = \frac{1}{q} \frac{\left(\frac{1}{q} - p^l\right) + I_l p^l}{p^l (1 - I)} = \frac{1}{q} \left\{ \left(\frac{1}{q p^l} - 1\right) + I_l \right\} \sum_{0}^{\infty} I_l^n$$

donc :

$$(84) \qquad \mathcal{E}_{(I_l^r)}\, n = \frac{1 - qp}{q^2 p^l}\, r + \frac{1}{q}(r - 1) = \frac{r}{q^2 p^l} - \frac{1}{q}.$$

On aurait aussi pu faire usage de (81) avec $A = I_l$, ce qui donnerait :

$$\sum I_l^r\, \mathcal{E}_{(I_l^r)}\, n = + \frac{I_l}{1 - I_l} \cdot \frac{1}{1-R} = \left(\frac{1}{q^2 p^l} - \frac{1}{q}\right) \frac{I_l}{(1 - I_l)^2} + \frac{1}{q} \frac{I_l^2}{(1 - I_l)^2}$$

$$= \left(\frac{1}{q^2 p^l} - \frac{1}{q}\right) \sum_{1}^{\infty}{}^{r}\, r I_l^r + \frac{1}{q} \sum_{2}^{\infty}{}^{r} (r - 1) I_l^r$$

donnant le même résultat (84).

19. Donnons enfin, en le généralisant, la solution d'un problème posé par M. Fréchet dans son livre mentionné ci-dessus, page 105. Puisque pour un l suffisamment grand (et $p_\lambda < 1$) $p_\lambda^{2l} << p_\lambda^{l}$, on conclut de (71.1) et (71.4) que $\sigma^2 - \alpha << \alpha$. Ce fait, prouvé par lui pour un cas très peu différent, lui a mené à conjecturer que la distribution de $I_{\lambda l}$ s'approchera, pour l et n augmentant indéfiniment tel que $n p_\lambda^l q_\lambda^2 = \alpha$, de la distribution de Poisson avec la moyenne α. Nous démontrerons qu'il en est ainsi, et même que les distributions des $I_{\lambda l}$ avec les l suffisamment grands tendent simultanément et uniformément à des distributions de Poisson mutuellement indépendantes.

Remarquons d'abord que la marque collective de la distribution de Poisson est :

$$\sum_{0}^{\infty}{}^{n}\, e^{-\alpha} \frac{\alpha^n}{n!} A^n = e^{-\alpha(1-A)}.$$

On le voit aussi en partant de la distribution de Bernoulli avec $p = \frac{\alpha}{n}$, l'autre marque étant démarquée, et passant à la limite pour $n \longrightarrow \infty$:

$$\lim_{n \to \infty} \left\{ \frac{\alpha}{n} A + \left(1 - \frac{\alpha}{n}\right) \right\}^n = \lim_{n \to \infty} \left\{ 1 - \frac{\alpha}{n}(1 - A) \right\}^n = e^{-\alpha(1-A)}$$

Soit $\max\limits_{\lambda} p_\lambda = \alpha < 1$ et $1 < \beta < \frac{1}{\alpha}$.

Posons :

$$P_\lambda = \beta p_\lambda \text{ et } Q_\lambda = 1 - P_\lambda.$$

Soit maintenant pour chaque λ $(1 \leqslant \lambda \leqslant k)$ $l_\lambda \geqslant 1$ si grand que :

$$(86) \qquad \frac{(l_\lambda + 1) P_\lambda^{l_\lambda}}{Q_\lambda^4} \leqslant \varepsilon < \min\left(\frac{1}{16k}, \frac{\beta - 1}{k}\right).$$

Démarquons complètement tous les $I_{\lambda l}$ avec $l \leqslant l_\lambda - 1$. Alors :

$$\varphi_\lambda = \sum_{1}^{\infty}{}^{l} (T p_\lambda)^l I_{\lambda l} = \frac{T p_\lambda}{1 - T p_\lambda} - \sum_{l_\lambda}^{\infty} (T p_\lambda)^l (1 - I_{\lambda l}) =$$

$$= \frac{T p_\lambda}{1 - T p_\lambda} - \gamma_\lambda(T)$$

avec :

$$(87)\qquad \gamma_\lambda(T)=\sum_{l_\lambda}^{\infty}{}_l (Tp_\lambda)^l(1-I_{\lambda l}).$$

En supposant $0 \leqslant I_{\lambda l} \leqslant 1$, et en admettant pour T des valeurs réelles ou complexes avec $|T| \leqslant \beta$, de sorte que $|T|\, p_\lambda \leqslant P_\lambda < 1$, on aura :

$$|\gamma_\lambda(T)| \leqslant \sum_{l_\lambda}^{\infty}{}_l P_\lambda^l = \frac{P_\lambda^{l_\lambda}}{Q_\lambda} \leqslant \frac{\varepsilon Q_\lambda^3}{l_\lambda+1}$$

et

$$|\gamma_\lambda(T)(1-Tp_\lambda)| \leqslant |\gamma_\lambda(T)|(1+P_\lambda) \leqslant \varepsilon\frac{Q_\lambda^2}{l_\lambda+1};$$

d'ailleurs :

$$|\gamma'_\lambda(T)| \leqslant \sum_{l_\lambda}^{\infty}{}_l\, lP_\lambda^l = P_\lambda^l\left(\frac{P_\lambda}{Q_\lambda^2}+\frac{l_\lambda}{Q_\lambda}\right) \leqslant \frac{\varepsilon Q_\lambda^2(P_\lambda+l_\lambda Q_\lambda)}{l_\lambda+1} \leqslant \varepsilon Q_\lambda^2.$$

Alors :

$$R_\lambda=\frac{\varphi_\lambda}{1+\varphi_\lambda}=Tp_\lambda-\frac{\gamma_\lambda(T)(1-Tp_\lambda)^2}{1-\gamma_\lambda(T)(1-Tp_\lambda)}$$

avec :

$$|R_\lambda - Tp_\lambda| = \left|\frac{\gamma_\lambda(T)(1-Tp_\lambda)^2}{1-\gamma_\lambda(T)(1-Tp_\lambda)}\right| \leqslant \frac{|\gamma_\lambda(T)|(1+P_\lambda)^2}{1-|\gamma_\lambda(T)|(1+P_\lambda)}$$
$$\leqslant \varepsilon\frac{Q_\lambda^3(1+P_\lambda)^2}{l_\lambda+1-\varepsilon Q_\lambda^2} \leqslant \frac{\varepsilon Q_\lambda}{l_\lambda}.$$

Donc :

$$R = T - \gamma(T) \text{ avec } \gamma(T)=\sum{}_\lambda \frac{\gamma_\lambda(T)(1-Tp_\lambda)^2}{1-\gamma_\lambda(T)(1-Tp_\lambda)}$$
$$|\gamma(T)| \leqslant \varepsilon\sum{}_\lambda\frac{Q_\lambda}{l_\lambda} \leqslant k\varepsilon$$

et

$$\gamma'(T)=\sum{}_\lambda \frac{(1-Tp_\lambda)^2\{\gamma'_\lambda(T)+p_\lambda\gamma_\lambda(T)^2\}-2p_\lambda\gamma_\lambda(T)(1-Tp_\lambda)}{\{1-\gamma_\lambda(T)(1-Tp_\lambda)\}^2}$$

$$\therefore |\gamma'(T)| \leqslant \sum{}_\lambda \frac{\frac{1}{Q_\lambda^2}\left(\varepsilon Q_\lambda^2+p_\lambda\varepsilon^2\frac{Q_\lambda^3}{l_\lambda+1}\right)+2p_\lambda\varepsilon\frac{Q_\lambda^2}{l_\lambda+1}}{\left(1-\frac{\varepsilon Q_\lambda^2}{l_\lambda+1}\right)^2} \leqslant$$

$$\leqslant 4\varepsilon.\sum\left(1+\varepsilon\frac{p_\lambda Q_\lambda^2}{l_\lambda+1}+2\frac{p_\lambda Q_\lambda^2}{l_\lambda+1}\right) \leqslant 16k\varepsilon < 1.$$

$$(88)\qquad \frac{C}{1-T}=\frac{1}{1-R}=\frac{1}{1-T+\gamma(T)}=\sum_0^\infty{}_n T^nC_n$$

et l'on aura :

$$(89)\qquad C_n \approx \frac{1}{|T|^{n+1}(1-\gamma'(T))},$$

T étant la racine la plus petite en valeur absolue de l'équation :

$$(90)\qquad 1-T+\gamma(T)=0.$$

Or, pour un T avec $|T| \leqslant \beta$ on a $|\gamma(T)| < k\varepsilon$, donc si un tel T satisfait à (90) la racine à valeur absolue minimale est celle qui pour $\varepsilon \longrightarrow 0$ tend vers 1. On peut donc appliquer le développement de Lagrange :

$$(91)\qquad f(T)=f(1)+\sum_1^\infty{}_k \frac{1}{k!}\left[\left(\frac{d}{dT}\right)^{k-1} f'(T)\gamma(T)^k\right]_{T=1}$$

pour trouver le membre droit de (87), ou plutôt son logarithme. En particulier, en posant $\gamma^{(k)}=\gamma^{(k)}(1)$:

$$T=1+\sum_1^\infty \frac{1}{k!}\left[\left(\frac{d}{dT}\right)^{k-1}\gamma(T)^k\right]_{T=1}=1+\gamma+\gamma\gamma'+\left(\gamma\gamma'^2+\frac{1}{2}\gamma^2\gamma''\right)+\left(\gamma\gamma'^3+\frac{3}{2}\gamma^2\gamma'\gamma''+\frac{1}{6}\gamma^3\gamma'''\right)+\dots$$

Puisque, d'après (85), toutes les dérivées de $\gamma(T)$ ont des coefficients $\geqslant 0$, T est un nombre réel et positif, > 1, mais, pour un ε assez petit on aura $T \leqslant 1+2\gamma < \beta$. Les autres racines de (88) ayant la valeur absolue $|T| > \beta$ on peut conclure que la faute relative faite en (87) est au plus de l'ordre de :

$$\left(\frac{1+2\gamma}{\beta}\right)^n.$$

D'ailleurs :

$$\ln(1-\gamma'(T))=\ln(1-\gamma')-\sum_1^\infty\frac{1}{k!}\left[\left(\frac{d}{dt}\right)^{k-1}\frac{\gamma(T)^k\gamma''(T)}{1-\gamma'(T)}\right]_{T=1}=-\gamma'-\frac{1}{2}(\gamma'^2+\gamma\gamma'')-\frac{1}{6}(2\gamma'^3+5\gamma\gamma'\gamma''+\gamma^2\gamma''')-\dots$$

On a donc d'après (87) :

$$\ln C_n \approx (n+1)\ln T+\ln(1-\gamma'(T))=\{(n+1)\gamma-\gamma'\}+\left\{(n+1)\left(\gamma\gamma'-\frac{1}{2}\gamma^2\right)-\frac{1}{2}(\gamma'^2+\gamma\gamma'')\right\}+\dots$$

Or :

$$\gamma_\lambda = \gamma_\lambda(1) = \sum_{l_\lambda}^{\infty} p_\lambda^l (1 - I_{\lambda l})$$

$$\gamma'_\lambda = \gamma'_\lambda(1) = \sum_{l_\lambda}^{\infty} l p_\lambda^l (1 - I_{\lambda l})$$

$$\gamma = \sum\lambda \frac{\gamma'_\lambda q_\lambda^2}{1 - \gamma_\lambda q_\lambda} \approx \sum\lambda\, \gamma_\lambda q_\lambda^2 + \sum\lambda\, \gamma_\lambda^2 q_\lambda^4$$

$$\gamma' \approx \sum\lambda \left\{ \frac{\gamma'_\lambda q_\lambda^2 - 2\gamma_\lambda p_\lambda q_\lambda}{1 - \gamma_\lambda q_\lambda} + \frac{\gamma_\lambda q_\lambda^2 (\gamma'_\lambda q_\lambda - \gamma_\lambda p_\lambda)}{(1 - \gamma_\lambda q_\lambda)^2} \right\} \approx$$

$$\approx \sum\lambda (\gamma'_\lambda q_\lambda^2 - 2\gamma_\lambda p_\lambda q_\lambda) + \sum\lambda\, \gamma_\lambda q_\lambda^2 (2\gamma'_\lambda q_\lambda - 3\gamma_\lambda p_\lambda).$$

Le terme principal de $\ln C_n$ étant $-(n+1)\gamma + \gamma'$ nous obtenons donc :

$$(90) \quad C_n \approx e^{-\sum\lambda \sum_{l_\lambda}^{\infty} l (1 - I_{\lambda l}) \{ (n+1) p_\lambda^l q_\lambda^2 - (l q_\lambda - 2p_\lambda) p_\lambda^l q_\lambda \}} =$$

$$= e^{-\sum\lambda \sum_{l_\lambda}^{\infty} l (1 - I_{\lambda l}) \{ (n+1-l) p_\lambda^l q_\lambda^2 + 2 p_\lambda^{l+1} q_\lambda \}}$$

ce qui est d'accord avec (71.1) et ce qui prouve qu'en première approximation les $I_{\lambda l}$ avec $l \geqslant l_\lambda$ sont distribués mutuellement indépendamment selon POISSON avec $\alpha_{\lambda l} = n p_\lambda^l q_\lambda$. En particulier on y trouve démontrée la conjecture de M. FRÉCHET. Il est d'ailleurs facile de calculer la seconde approximation.