ANNALES DE L'I. H. P.

ALF GULDBERG

Les fonctions de fréquence discontinues et les séries statistiques

Annales de l'I. H. P., tome 3, n° 3 (1933), p. 229-278

<http://www.numdam.org/item?id=AIHP_1933__3_3_229_0>

Les fonctions de fréquence discontinues et les séries statistiques

PAR

ALF GULDBERG

Qu'il me soit permis tout d'abord, d'adresser mes bien sincères remerciements aux membres du Comité de Direction de l'Institut Henri Poincaré, qui m'ont fait le grand honneur de m'appeler en qualité de conférencier ; puis de rappeler ici la mémoire du grand savant sous l'égide duquel cet institut a été créé. Je tiens aussi à remercier M. FRÉCHET pour l'amabilité et l'estime qu'il a bien voulu me témoigner, et j'ose espérer que je ne détromperai pas la confiance que, les uns et les autres, ont mise en moi.

M. Émile BOREL écrit dans son livre *Eléments de la théorie des Probabilités*, p. 107 :

Le problème général de la statistique mathématique est le suivant :

Déterminer un système de tirages effectués dans des urnes de composition fixe, de telle manière que les résultats d'une série de tirages, interprétés à l'aide de coefficients fixes convenablement choisis, puissent conduire, avec une très grande vraisemblance, à un tableau identique au tableau des observations.

« Dans le cas où l'on peut y parvenir grâce à une seule urne, le problème considéré est dit *normal.* »

Je me propose dans ces conférences de faire quelques remarques sur ce problème, et de donner un très court aperçu de quelques unes des méthodes employées dans la théorie des observations, dont l'importance s'est accrue considérablement ces dernières années du fait de son application à la majorité des sciences physiques : biologie, économie politique, etc., c'est-à-dire à toutes les sciences fondées sur l'expérience.

Considérons d'abord les différentes manières de caractériser les séries d'observations. Soit une série d'observations d'un certain événement E :

$$(1) \qquad x_1, x_2, \cdots x_n,$$

où les $x_1, x_2, \cdots x_n$ sont par exemple les résultats de n mesures de la même grandeur z dans les mêmes conditions de précision, ou les résultats de n observations, par exemple la taille des hommes adultes dans une région donnée.

Pour fixer les idées, je donnerai ici quelques exemples tirés de différentes sciences.

Ex. 1 (1). — Classement des soldats tués par ruade de cheval dans l'armée prussienne, pendant la période 1875-1894, d'après les corps d'origine, corps d'armée, et corps de la Garde, désignés respectivement par I, II,... XV et G.

Désignation des corps	Années																			
	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94
I.....				2		3		2				1	1	1		2		3	1	
II.....				2		2			1	1			2	1	1			2		
III.....				1	1	1	2		2				1		1	2	1			
IV.....		1		1	1	1	1					1					1	1		
V.....					2	1			1			1		1	1	1	1	1	1	
VI.....			1		2			1	2	1	1	1	3	1	1	1		3		
VII.....	1		1				1		1				2			2	1		2	
VIII.....	1				1			1					1				1	1		1
IX.....						2	1	1	1	2	2	1	1		1	2		1		1
X.....			1	1		1		2							2	1	3		1	1
XI.....					2	4		1	3		1	1	1	1	2	1	3	1	3	
XIV.....	1	1	2	1	1	3		4		1		3	2	1		2	1	1		
XV.....		1						1		1	1				2	2				
G.....		2	2	1			1	1		3		2	1			1		1		1

(1) L. v. BORTKEWICZ : *Das Gesetz der kleinen Zahlen*, p. 23.

Les résultats de ce tableau sont résumés dans le suivant, où $F(x)$ désigne le nombre des soldats tués dans les différents corps d'armée, pendant les années 1875-1894.

Evènements par an	Fréquence des évènements observés
x	$F(x)$
0	144
1	91
2	32
3	11
4	2
	280

Ex. 2 (1). — Le tableau suivant donne le nombre des glandes de Müller de la patte droite antérieure de 2000 truies.

Nombre de glandes	Fréquence des glandes	Nombre de glandes	Fréquence des glandes
x	$F(x)$	x	$F(x)$
0	15	6	134
1	209	7	72
2	365	8	22
3	482	9	8
4	414	10	2
5	277		2000

Ex. 3. — Ce tableau donne le nombre de particules alpha rayonnées par le polonium dans un intervalle de temps donné, déterminé par MM. RUTHERFORD et GEIGER.

(1) DAVENPORT, *Statistical Methods*, p. 32.

Nombre de particules	Fréquence de ce nombre	Nombre de particules	Fréquence de ce nombre
x	$F(x)$	x	$F(x)$
0	57	8	45
1	203	9	29
2	383	10	10
3	525	11	4
4	532	12	0
5	408	13	1
6	273	14	1
7	139		2608

Chacune de ces séries peut être regardée comme une variable statistique x, avec la fréquence relative empirique $\frac{F(x)}{N}$, N étant le nombre des observations.

Le problème à résoudre est le suivant : Définir une variable aléatoire x et trouver la fonction de fréquence (loi de probabilité) $f(x)$, telle que $f(x)$ soit égale, ou au moins sensiblement égale, à la grandeur $\frac{F(x)}{N}$ fournie par l'expérience. Cela revient, dans l'esprit de M. BOREL, à déterminer une urne de composition fixe en boules blanches et noires, telle que les résultats d'une série de tirages de cette urne conduisent à un tableau de nombres de boules blanches tirées, égaux ou très approchés, des nombres fournis par l'expérience.

Avant de traiter ce problème, il faut d'abord examiner notre ensemble donné d'observations. Comme les observations $x_1, x_2 \cdots, x_n$ sont équivalentes, ces observations seront caractérisées par des fonctions symétriques.

1. — Les moments moyens autour de la moyenne arithmétique sont donnés par

$$m_r = \frac{1}{n} \sum_{1}^{n} (x_i - m_1)^r \qquad r \geqslant 2$$

et, en particulier, les moments par rapport à l'origine par

$$\sigma_r = \frac{\sum_1^n x_i^r}{n}.$$

On voit que les moments moyens m_r et les moments σ_r se correspondent.

Parmi ces moments, les plus importants sont la moyenne arithmétique

$$\sigma_1 = m_1 = \frac{x_1 + x_1 + \cdots + x_n}{n}.$$

et la moyenne quadratique

$$m_2 = \frac{(x_1 - m_1)^2 + (x_1 - m_1)^2 + \cdots (x_n - m_1)^2}{n}.$$

Les deux fonctions m et m_2 donnent une première caractéristique de notre ensemble. Elles ont aussi une grande importance dans la vie pratique. Je prends un exemple de la théorie des assurances sur la vie.

La prime unique qu'une personne d'âge x doit payer à une société d'assurances sur la vie, pour que la société paye à son décès un franc, est déterminée de la manière suivante. Les tables de mortalité de la société donnent, pour un groupe de l_x personnes d'âge x, les nombres d_x des personnes qui mourront à cet âge, les nombres d_{x+1} des personnes d'âge $x + 1$ qui mourront à l'âge $x + 1$, etc. Si l'on désigne par i le taux de l'intérêt, la valeur actuelle de 1 franc payable dans n années est v^n, où $v = \frac{1}{1+i}$. En supposant que toutes les l_x personnes achètent une telle assurance, les valeurs actuelles des sommes que la société doit payer (les sommes sont supposées payées à la fin de l'année de décès) seront

$$d_x v,\ d_{x+1} v^2 \cdots,\ d_\omega v^{\omega - x + 1}\ ; \qquad \text{et enfin} \qquad d_{\omega+1} = 0.$$

Dans ces conditions la moyenne arithmétique :

$$\frac{d_x v + d_{x+1} v^2 + \cdots + d_\omega v^{\omega - x + 1}}{l_x}$$

est la prime cherchée. On la désigne par A_x. Les rapports $\frac{d_x}{l_x}$, $\frac{d_{x+1}}{l_x}$ sont les fréquences relatives des observations $v, v^2, \cdots v^{\omega - x + 1}$.

La moyenne quadratique est

$$\frac{d_x(v - A_x)^2 + d_{x+1}(v^2 - A_x)^2 + \cdots + d_\omega(v^{\omega - x + 1} - A_x)^2}{l_x} =$$

$$= A^1{}_x - A^2{}_x = M(A_x),$$

où

$$A^1{}_x = \frac{d_x v^2 + d_{x+1} v^4 + \cdots + d_\omega v^{2(\omega - x + 1)}}{l_x}.$$

La racine carrée de cette moyenne quadratique est appelée le risque moyen.

Un exemple numérique, calculé d'après les tables de mortalité anglaise H^m, $i = 0{,}035$, donne

$$A_{70} = 0{,}74738, \qquad M(A_{70}) = 0{,}1374, \qquad \frac{M}{A} = 0{,}184.$$

La notion de risque moyen permet à la société de comparer ses assurances et d'examiner l'influence de leur grandeur et de leur répartition, questions importantes pour la réassurance, la stabilité et le plein. Un problème délicat constitue le calcul des réserves du risque et de la taxe.

Si les moments moyens d'ordres impairs m_{2r+1}, sont nuls, l'ensemble sera symétrique.

2. — *Les semi-invariants μ_r de* THIELE. — THIELE introduit les semi-invariants, par l'identité en t

$$e^{\mu_1 t + \frac{\mu_2}{2!} t^2 + \frac{\mu_3}{3!} t^3 + \cdots} = 1 + \frac{\sigma_1}{1} t + \frac{\sigma_2}{2!} t^2 + \cdots$$

Cette identité est toute formelle ; elle remplace le système d'équations linéaires auxquelles elle conduit. En prenant la dérivée logarithmique de notre identité on aura :

$$(\mu_1 + \frac{\mu_2}{1!} t + \frac{\mu_3}{2!} t^2 + \cdots)(1 + \frac{\sigma_1}{1} t + \frac{\sigma_2}{2!} t^2 + \cdots)$$

$$= \sigma_1 + \frac{\sigma_2}{1!} t + \frac{\sigma_3}{2!} t_2 + \cdots$$

En comparant les coefficients de t^r on a

$$\sigma_{r+1} = \mu_1 \sigma_r + C^1_r \mu_2 \sigma_{r-1} + C^2_r \mu_3 \sigma_{r-2} + \cdots$$

d'où l'on voit que les semi-invariants et les moments sont uniformément déterminés. On trouvera facilement la formule :

$$\mu_{r+1} = m_{r+1} - C_r^2 m_2 \mu_{r-2} \cdots - C_r^{r-1} m_{r-1} \mu_2 \qquad (r > 2).$$

On a, en particulier :

$$\mu_1 = \sigma_1 = m_1$$
$$\mu_2 = \sigma_2 - \sigma_1^2 = m_2$$
$$\mu_3 = \sigma_3 - 2\sigma_2\sigma_1 + 2\sigma_1^3 = m_3$$
$$\mu_4 = m_4 - 3m_2^2.$$

On voit que les expressions de μ_r ne contiennent que les moments jusqu'à l'ordre r. μ_r est donc fini si σ_r est fini.

Les semi-invariants d'ordre impair étant nuls, l'ensemble donné est symétrique.

La raison pour laquelle THIELE a introduit les semi-invariants, notion qui semble quelque peu artificielle, est la suivante : si les observations suivent la loi de GAUSS, tous les semi-invariants d'ordre $r > 2$ sont nuls. On a donc, comme me l'a d'ailleurs fait remarquer un ancien élève de THIELE, mon collègue M. HEEGAARD, un moyen d'examiner si un ensemble statistique peut se représenter par l'intermédiaire de la loi de GAUSS ou non.

Les semi-invariants ont cependant d'autres avantages. Quand on fait une transformation linéaire des observations

$$x_i = ax_i + b$$

c'est-à-dire, quand on change l'origine des observations et l'unité qui les mesure, les semi-invariants se transforment d'après les formules

$$\mu_1' = a\mu_1 + b$$
$$\mu_r' = a^r \mu_r \quad r > 1.$$

Les moments moyens se transforment d'une façon analogue.

Plus généralement, soit x_i est une fonction linéaire de m observations indépendantes $x_i^{(1)}$, $x_i^{(2)} \cdots x_i^{(n)}$

$$x_i = a_0 + a_1 x_i^{(1)} + a_2 x_i^{(2)} + \cdots a_m x_i^{(m)} i; \qquad i = 1, 2, \cdots n.$$

Les semi-invariants des x_i s'expriment au moyen des semi-invariants des $x_i^{(1)} \cdots x_i^{(n)}$

$$\mu_1(x) = a_0 + a_1 \mu_1(x^{(1)}) + \cdots + a_m \mu_1(x^{(r)})$$
$$\mu_r(x) = a_1^r \mu_r(x^{(1)}) + \cdots + a_m^r \mu_r(x^{(c)}).$$

Les moments moyens se transforment d'une manière plus compliquée.

Ces relations montrent la grande importance que présentent les semi-invariants pour l'étude d'un ensemble statistique.

On retrouvera d'une manière très simple plusieurs résultats importants de M. Paul LÉVY sur la composition des lois de probabilité et en particulier le beau théorème de M. d'OCAGNE.

3. — *Les moments factoriels de* MM. STEFFENSEN et SHEPPARD.

Les moments factoriels sont définis par la formule suivante :

$$\sigma_{(r)} = \frac{1}{n} \sum x_i(x_i - 1)(x_i - 2) \cdots (x_i - r + 1).$$

Comme pour les semi-invariants, on peut également définir les moments factoriels par une identité

$$\sigma_0 + \frac{\sigma_{(1)}}{1} t + \frac{\sigma_{(2)}}{1.2} t^2 + \cdots = \frac{1}{n} \sum (1 + t)^{x_i} \qquad |t| < 1$$

Les trois espèces de fonctions symétriques : les moments, les semi-invariants et les moments factoriels sont équivalents et complètement déterminés par les observations. Inversement, les observations sont complètement déterminées par ces fonctions symétriques.

On a, en particulier, pour les moments factoriels

$$\sigma_{(1)} = \sigma_1$$
$$\sigma_{(2)} = \sigma_2 - \sigma_1$$
$$\sigma_{(3)} = \sigma_3 - 3\sigma_2 + 2\sigma_2, \qquad \text{etc.}$$

Cela étant, supposons que dans n mesures, on tombe $\nu_n(x)$ fois sur l'observation X ; la grandeur $f(x)$ définie par

$$\lim_{n \to \infty} \frac{\nu_n(x)}{n} = f(x)$$

est appelée la probabilité de l'observation X.

A mon avis, cette définition de la notion probabilité est, en réalité, identique à la définition classique. Lorsque tous les cas sont également probables, la probabilité d'un événement est le rapport du nombre des cas favorables au nombre total des cas possibles. La notion de cas « également probables » ne peut avoir aucun autre sens que celui-ci : les différents cas se produisent constamment le *même* nombre de fois,

et de plus en plus nombreux à mesure qu'on multiplie les expériences. Toutes les notions que nous utilisons sont des abstractions déduites de nos expériences. Quand nous disons que dans un jeu de 52 cartes chacune des cartes a la même probabilité d'être tirée, ce jugement est fondé sur notre expérience, qui nous apprend que les différentes cartes apparaissent les mêmes nombres de fois lorsqu'on fait un nombre de tirages suffisamment grand. Un être, qui n'a jamais vu un jeu de cartes et qui n'a jamais fait une expérience sur des objets analogues, serait dans l'impossibilité absolue d'affirmer quoi que ce soit à ce sujet.

Soit X une variable aléatoire, avec la fonction de fréquence c'est-à-dire la loi de probabilité, $f(x)$: X prend les valeurs 0, 1, $2 \cdots k$ avec les probabilités $f(0)$, $f(1), \cdots f(k)$.

Les moments, les semi-invariants et les moments factoriels prennent alors une forme simple. On a, pour les moments :

$$\sigma_r = \sum_0^k x^r f(x)$$

et pour les moments moyens

$$m_r = \sum_0^k (x - m_1)^r f(x)$$

où

$$\sum_0^k f(x) = 1.$$

M. Paul LÉVY a montré comment on peut calculer très simplement les moments, en utilisant la fonction caractéristique :

$$\varphi(t) = \sum_0^k e^{ixt} f(x), \qquad i = \sqrt{-1}.$$

En effet,

$$\sum e^{ixt} f(x) = 1 + \frac{it}{1} \sigma_1 + \frac{(it)^2}{1 \cdot 2} \sigma_2 + \cdots$$

d'où

$$\varphi^{(p)}(o) = i^p \sigma_p$$

Soit, par exemple, X une variable aléatoire, ayant comme fonction de fréquence la fonction binomiale. L'expression

$$f(x) = C_k^x p^x (1 - p)^{k-x}$$

représente la probabilité pour que l'événement E avec la probabilité p se présente x fois, et que l'événement contraire se présente $(k - x)$ fois dans une suite de k épreuves. p et k caractérisent complètement $f(x)$. La fonction caractéristique est

$$\varphi(t) = \sum_0^k C_k^x p^x (1 - p)^{k-x} e^{ixt}$$

ou

$$\varphi(t) = \sum C_k^x (pe^{it})^x (1 - p)^{k-x} = (pe^{it} + 1 - p)^k.$$

On aura donc les moments, exprimés en fonction de p et k

$$\sigma_1 = \frac{\varphi'(o)}{i} = kp$$

$$\sigma_2 = -\varphi''(o) = k^2p^2 - kp^2 + kp, \qquad \text{etc.}$$

Considérons les semi-invariants. Ils sont définis par

$$e^{\mu_1 t + \frac{\mu_2}{2!} t^2 + \cdots} = \sum_0^k e^{xt} f(x).$$

Mais $\sum_0^k e^{xt} f(x)$ est justement la fonction caractéristique au sens de CAUCHY. Si nous remplaçons t par it dans notre équation, nous aurons :

$$e^{\mu_1 it + \frac{\mu_2}{1.2}(it)^2 + \cdots} = \sum_0^k e^{ixt} f(x)$$

Mais $\sum_0^k e^{ixt} f(x)$ est la fonction caractéristique $\varphi(x)$, au sens de M. LÉVY. On aura donc

$$e^{\mu_1 it + \frac{\mu_2}{1.2}(it)^2 + \cdots} = \varphi(t).$$

Posons $\log \varphi(t) = \psi(t)$ et appelons $\psi(t)$ la fonction caractéristique des semi-invariants ; on a

$$\psi(t) = \frac{\mu_1 t}{1} + \frac{\mu_2}{1 \cdot 2}(it)^2 + \cdots$$

Les semi-invariants se déterminent très simplement d'une façon analogue aux moments, par leur fonction caractéristique (1). On a

$$u_r = \frac{\psi^{(r)}(0)}{i^r}.$$

Nous avons déjà remarqué que les fonctions symétriques, les moments, les semi-invariants, les moments factoriels sont équivalents aux observations ; on peut donc substituer ces fonctions à notre fonction de fréquence. Inversement, la fonction de fréquence est déterminée lorsque les fonctions symétriques sont données. Pour les moments factoriels, on le voit facilement (2) ; on a :

$$\begin{aligned}
\sigma_{(1)} &= f(0) + f(1) + \cdots + f(k) \\
\sigma_{(2)} &= 1f(1) + 2f(2) + \cdots kf(k) \\
\sigma_{(3)} &= 2 \cdot 1\, f(2) + \cdots + k(k-1)f(k) \\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
\sigma_{(k)} &= k(k-1)\cdots(k-(k-1))f(k).
\end{aligned}$$

On voit directement qu'une solution existe ; on trouve

$$f(x) = \frac{1}{x!} \sum_{s=0}^{k-x} \frac{(-1)^s}{s!} \sigma_{(x+s)}.$$

Les fonctions symétriques remplacent complètement la fonction de fréquence, lorsque celle-ci est discontinue et finie. Quand le nombre des observations est infini, le problème a encore une solution sous certaines conditions.

Si x est une variable aléatoire avec la fonction de fréquence binomiale $f(x) = C_k^x p^x (1-p)^{k-x}$, la fonction caractéristique des semi-invariants sera

$$\psi(t) = k \log (pi^{it} + 1 - p)$$

(1) M. Paul Lévy utilisait en effet, la notion fonction caractéristique des semi-invariants, sans remarquer son rapport avec ceux-ci. M. Lévy a établi d'importantes propriétés de $\psi(t)$.

(2) Voir Steffensen, *Iagttagelselære*, p. 35.

d'où l'on tire

$$\mu_1 = kp$$
$$\mu_2 = kp(1 - p)$$
$$\mu_3 = kp(1 - p)(1 - 2p)$$
$$\mu_4 = kp(1 - p)(1 - 6p(1 - p)) \text{ etc.}$$

En éliminant les paramètres p et k entre les $(k + 1)$ premiers semi-invariants, on obtiendra les relations auxquelles doivent satisfaire les semi-invariants de la fonction binomiale. Des quatre premières relations on tire : $\mu_1 > \mu_2$, et on doit avoir, de plus

$$\frac{\mu_3\mu_1}{2\mu_2^2 - \mu_1 \cdot \mu_2} = 1$$

$$\frac{\mu_4 \cdot \mu_1}{\mu_1^2\mu_2 - 6\mu_1\mu_2^2 + 6\mu_2} = 1$$

Une série statistique dont les semi-invariants satisfont approximativement aux relations précédentes, doit donc être représentée d'une manière approchée par la fonction de fréquence binomiale (1).

Dans la théorie de la dispersion, que nous traiterons plus tard, on a utilisé en partie ces remarques en comparant entre eux les seconds semi-invariants de la série statistique et de la fonction binomiale.

Considérons la variable aléatoire X, ayant la fonction de fréquence de POISSON

$$f(x) = \frac{\lambda^x e^{-\lambda}}{x!},$$

qui est le cas limite de fonction binomiale pour $k \to \infty$, $p \to 0$, et $kp \to \lambda$.

La fonction de POISSON a été l'objet d'un grand nombre de recherches importantes. Je cite le mémoire classique de v. BORTKIEWICZ, *Gesetz der kleinen Zahlen*, où BORTKIEWICZ a donné un supplément important à la théorie de la dispersion. M. CHARLIER a examiné dans plusieurs mémoires les propriétés de la fonction de POISSON et montré sa grande maniabilité ; il l'a également employée comme fonction génératrice pour un développement en série d'une fonction arbitraire de répartition. Dans sa thèse de doctorat, M. JORGENSEN a généralisé les recherches de M. CHARLIER dans plusieurs directions et en a donné des applications à la théorie de la corrélation. ERLANG a montré l'utilité de la fonction dans la théorie des communications téléphoniques.

(1) Voir R. FRISCH : On the use of difference equations in the study of frequency distributions, *Metron*, t. X, N. 3, 1932, p. 56.

La fonction de POISSON a été appelée : « loi des événements rares » ou la loi-limite des petites probabilités. M. FRÉCHET écrit [1] : « Ce qui rend cette loi particulièrement précieuse, c'est que les valeurs de k et p n'y interviennent pas directement. Or, dans nombre des statistiques, on connaît le nombre de répétitions d'un événement assez rare, sans savoir d'une façon très précise sur quel groupe il porte. Par exemple, le nombre d'accidents mortels survenus dans certaines profession... ».

La fonction caractéristique des semi-invariants de la fonction de POISSON est

$$\psi(x) = \lambda e^{it} - \lambda = \lambda\left(\frac{it}{1} + \frac{(ix)}{1\cdot 2} + \cdots\right),$$

car on a

$$\psi(t) = \log \sum_0^\infty \frac{\lambda^x e^{-\lambda}}{x!} e^{ixt} = \log \sum_0^\infty \frac{(\lambda e^{it})^x}{x!} e^{-\lambda} = \log e^{\lambda e^{it} - \lambda}.$$

On trouvera :

$$\mu_1 = \mu_2 = \mu_3 = \cdots = \lambda.$$

Les semi-invariants de la fonction de POISSON sont constants $= \lambda$. (Théorème dû à V. BORTKIEWICZ). Grâce à cette propriété, on peut s'assurer si une série statistique est susceptible d'être représentée par la fonction de POISSON.

Considérons les moments factoriels. Ils sont définis par

$$\sum_0^k (1 + t)^x f(x) = 1 + \frac{\sigma_{(1)}}{1} t + \frac{\sigma_{(2)}}{1\cdot 2} t^2 + \cdots$$

Posons $\sum_0^k (1 + t)^x f(x) = \Theta(t)$ et appelons $\Theta(t)$ la fonction caractéristique des moments factoriels; on aura :

$$\sigma_{(r)} = \Theta^{(r)}(0).$$

La fonction caractéristique des moments factoriels de la fonction de fréquence binomiale, est

$$\theta(t) = \sum_0^k C_k^x p^x (1 - p)^{k-x}(1 + t)^x = \sum_0^k C_k^x (1 + t)\ (p)^x (1 - p)^{k-x}$$

1) FRÉCHET et HALBWACHS, *Le calcul des Probabilités à la Portée de tous*, p. 244.

ou

$$\theta(t) = (pt + 1)^k$$

d'où

$$\sigma_{(r)} = k(k-1)(k-2)\cdots(k-(r-1))p^r.$$

La fonction caractéristique des moments factoriels dans le cas où la fonction de fréquence est celle de Poisson, est

$$\theta(t) = \sum_0^\infty \frac{\lambda^x e^{-\lambda}}{x!}(1+t)^x = \sum \frac{(\lambda(1+t))^x}{x!} e^{-\lambda}$$

ou

$$\theta(t) = e^{\lambda t},$$

d'où

$$\sigma_{(r)} = \lambda^r.$$

Cherchons à appliquer ces remarques à nos séries statistiques.

Nous avons remarqué que tous les semi-invariants de la fonction de fréquence de Poisson sont constants. Une série statistique dont tous les semi-invariants sont constants peut donc être représentée par la fonction de fréquence de Poisson.

Considérons la série statistique de notre premier exemple :

x	0	1	2	3	4
$F(x)$	144	91	32	11	2

Formons les premiers semi-invariants ; ils ont pour valeur :

$$\mu_1 = 0{,}70, \quad \mu_2 = 0{,}76, \quad \mu_3 = 0{,}48.$$

Les trois premiers semi-invariants sont sensiblement égaux. Voyons si la fonction de fréquence de Poisson peut être employée pour représenter de notre série avec

$$\lambda = 0{,}65 = \frac{\mu_1 + \mu_2 + \mu_3}{3}.$$

Si l'on pose

$$H(x) = 280 \cdot \frac{0{,}65^x e^{-0{,}65}}{x!}$$

on aura le tableau suivant :

x :	0,	1,	2,	3,	4,	5
$H(x)$:	143,1,	92,1,	33,3,	8,9,	2,9,	0,6

On voit donc que la variable aléatoire x, avec la fonction de fréquence de Poisson

$$f(x) = \frac{\lambda^x e^{-\lambda}}{x!}, \qquad \lambda = 0{,}65$$

donne une bonne approximation de notre série.

Considérons notre exemple 2 :

x :	0,	1,	2,	3,	4,	5,	6,	7,	8,	9,	10
F(x) :	15,	209,	365,	482,	414,	277,	134,	72,	22,	8,	2.

On trouve ici :

$$\mu_1 = 3{,}5 \quad \mu_2 = 2{,}8 \quad \mu_3 = 2{,}4 \quad \mu_4 = 1{,}4$$

valeurs qui sont assez différentes. Si l'on substitue ces valeurs de μ_1, μ_2, μ_3, μ_4, dans les deux relations entre les semi-invariants de la fonction de fréquence binomiale, on trouvera les valeurs de 1,2 et 1,7 au lieu de 1. On doit donc supposer que la fonction de fréquence binomiale $f(x)$ donnera une approximation satisfaisante. On trouve en effet, en posant $H(x) = 2000\, f(x)$:

x :	0.	1,	2,	3,	4,	5,	6,	7,	8,	9,	10 (1).
H(x) :	41,	177,	304,	468,	423,	286,	150,	62,	21,	6,	1

Nous avons cherché jusqu'ici à caractériser les observations par des fonctions symétriques, mais nous tenons à faire remarquer qu'il existe d'autres modes de représentation. On peut distinguer entre l'école anglaise, représentée par M. KARL PEARSON, et l'école continentale, représentée par MM. GRAM, BRUNS et CHARLIER.

Disons d'abord quelques mots sur l'école de M. K. PEARSON, sur laquelle M. E. BOREL appelle l'attention (*l. c.* p. 168). M. PEARSON part du problème suivant :

Une urne contient un nombre k de boules blanches et un nombre h de boules noires. On fait m tirages *sans* remettre dans l'urne les boules tirées. On cherche la probabilité $f(x)$ d'extraire x boules blanches. On a

$$f(x) = \frac{C_k^x\, C_h^{m-x}}{C_{h+k}^m},$$

(1) Voir STEFFENSEN, *Iagttagelseslœre*, p. 122.

d'où l'on déduit l'équation aux différences finies de la fonction $f(x)$

$$f(x+1) = \frac{(k-x)(m-x)}{(x+1)(h-m+1+x)} f(x).$$

On appelle $f(x)$ la fonction de fréquence hypergéométrique.

M. PEARSON remplace cette équation aux différences finies (1) par l'équation différentielle

$$(a) \qquad \frac{1}{y}\frac{dy}{dx} = \frac{x-c}{c_0 + c_1 x + c_2 x^2}.$$

Il substitue ainsi à la fonction hypergéométrique discontinue une fonction continue, et aux entiers h, k, m des constantes arbitraires. Comme M. STEFFENSEN l'a indiqué, il court donc le danger de déformer complètement sa fonction. Les différentes valeurs des zéros de l'équation

$$c_0 + c_1 x + c_2 x^2 = 0$$

donnent une série de types de courbes. M. PEARSON déclare que ces types de courbes renferment la majorité de courbes de répartition : « I might almost say, without exception in a wide range of economic, physical, biometric and actuarial data » (2).

M. PEARSON détermine les constantes de l'équation différentielle (a) par les moments de la fonction y de la manière suivante :

Il écrit l'équation (a) sous la forme

$$(c_0 + c_1 x + c_2 x^2)\, dy = y(x-c)\, dx.$$

Multipliant cette équation par x^{n-1} et intégrant, on a :

$$\int x^{n-1}(c_0 + c_1 x + c_2 x^2) dy = \int y(x-c) x^{n-1} dx,$$

ou, en intégrant par parties

$$x^{n-1}(c_0 + c_1 x + c_2 x^2) y - \int y((n-1)c_0 x^{n-2} + n c_1 x^{n-1} + (n+1) c_2 x^n)\, dx$$
$$= \int y x^n\, dx - c \int y x^{n-1} dx.$$

En remarquant que l'expression $x^{n-1}(c_0 + c_1 x + c_2 x^2)$ disparaît aux limites et en posant

$$M_n = \int y x^n\, dx,$$

(1) Voir ELDERTON, *Frequency-curves and Correlation*, p. 37.
(2) *On the general theory of skew correlation and non linear regression*, p. 9.

on aura :

$$M_n[1 + (n+1)c_2] + M_{n-1}[nc_1 - c] + M_{n-2}[n-1]c_0 = 0 \qquad (A).$$

Cette relation est valable aussi pour $n = 1$, si l'on pose par définition $M_{-1} = 0$. En posant $n = 1, 2, 3, 4$ on détermine c, c_0, c_1, c_2 par les cinq moments m_0, m_1, m_2, m_3, m_4, ou $m_0 = 1$. En plaçant l'origine du système au centre de gravité de notre courbe, on a $m_1 = 0$; on trouve donc :

$$c_0 = -\frac{M_2(4M_2M_4 - 3M_3^2)}{10M_2M_4 - 18M_3^2 - 12M_2^3},$$

$$c = c_1 = -\frac{M_3(M_4 + 3M_2^2)}{10M_2M_4 - 18M_3^2 - 12M_3^2},$$

$$c_2 = -\frac{2M_2M_4 - 6M_3^2 - 3M_3^2}{10M_2M_4 - 18M_3^2 - 12M_3^2}.$$

Pour distinguer les différents types des courbes, M. PEARSON discute les valeurs de

$$k = \frac{c_1^2}{4c_0c_2} = \frac{\beta_1(\beta_2 + 3)^2}{4(2\beta_2 - 3\beta_1 - 6)(4\beta_2 - 3\beta_1)},$$

où

$$\beta_1 = \frac{M_3^2}{M_2^3}, \quad \beta_2 = \frac{M_4}{M^2}.$$

M. PEARSON utilise β_1 comme une mesure de l'asymétrie de la courbe et appelle $\beta_2 + 3 = \varepsilon$ l'excès de la courbe. Je cite M. ELDERTON, *Frequency curves and correlation*, p. 50 :

« Another matter with which it seems advisable to deal here is connected with the criterion k. This may have any value from $-\infty$ to $+\infty$, and from the following diagram it will be seen how the types cover all the possible values of the criterion and do not overlap. »

$k = -\infty$		$k = 0$		$k = 1$		$k = \infty$
	k ngateive		$0 > k < 1$		$k > 1$	
	Type I		Type IV		Type VI	
Type III		Type VII when $\beta_2 = 3$ Type II when β_2 not $= 3$.		Type V		Type III

Quand on sait *a priori* qu'un ensemble statistique peut être représenté par une des courbes de M. PEARSON, k indique quel type doit être employé.

Les conditions nécessaires auxquelles doivent satisfaire les moments d'un ensemble statistique donné pour qu'il puisse être représenté par une des courbes de M. PEARSON, s'obtiennent en effet très simplement en substituant les valeurs trouvées c_0, c_1, c_2 de c, dans la relation (A).

On trouve

$$(B) \cdots M_n[(8 - 2n)M_2M_4 - (12 - 6n)M_2^3 - (9 - 3n)M_3^2] -$$
$$- M_{n-1}(n - 1)[M_3M_4 + 3M_3M_3^2] - M_{n-2}(n - 1)[4M_3^2M_4 - 3M_2M_3^2] = 0$$

On voit au moyen de cette formule, que $M_3 = 0$, a pour conséquence que tous les moments *impairs* sont nuls. $\beta_1 = 0$ est donc, en effet, une mesure de la symétrie.

Comme exemple, prenons l'équation des courbes du type II, qui s'écrit

$$y = y_0 \left(1 - \frac{x_2}{\alpha_2}\right)^\nu.$$

Les moments sont

$$M_{2n+1} = 0, \; M_{2n} = \frac{1.3.5 \cdots (2n - 1)\alpha^{2n}}{(2\nu + 3)(2\nu + 5) \cdots (2\nu + 2n + 1)}$$

qui satisfont à la relation (B), ainsi qu'on le voit facilement.

Le critère de M. PEARSON se réduit pour la courbe du type VII, qui est la courbe de Gauss

$$y = y_0 e^{\frac{xz}{a}}$$

à

$$M_3 = 0, \; M_4 - 3M_2 = 0.$$

En substituant ces valeurs dans la relation (B), on aura :

$$M_{2n+1} = 0, \; M_{2n} = (2n - 1)M_{2n-2},$$

qui sont en effet les critères nécessaires et suffisants pour qu'une fonction arbitraire soit la fonction de Gauss.

Comme un exemple numérique de la représentation d'un ensemble statistique par une courbe de M. PEARSON, je cite le tableau :

x :	17	22,	27,	32,	37,	42,	47,	52,	57,	62,	67,	72,	77,	82,	87.
$H(x)$:	34	145,	156,	145,	123,	103,	86,	71,	55,	37,	21,	13,	7,	3,	1

Ce tableau peut être représenté par la courbe du type I

$$y = 149.47\left[1 + \frac{x}{1.99638}\right]^{0.409833}\left[1 - \frac{x}{13.52728}\right]^{2.776978}$$

qui donne le tableau :

x :	17,	22,	27,	32,	37,	42,	47,	52,	57,	62,	67,	72,	77,	82, 87.
y :	46,	138,	149,	142,	127,	108,	88,	69,	51,	36,	24,	14,	7,	3, 1.

On voit qu'on a une bonne représentation de notre série, mais on n'a aucune espèce de renseignement sur son origine et on ne peut rattacher à rien la nature et la grandeur des constantes de notre formule.

L'école continentale part d'autres considérations, et fait intervenir deux fonctions qui dérivent de la loi de fréquence binomiale :

$$f(x) = (c_k^x)p^x(1 - p)^{k-x}.$$

Ce sont 1° : la fonction de Gauss :

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

pour $k \to \infty$ et 2° : la fonction de Poisson

$$f(x) = \frac{\lambda^x e^{-\lambda}}{x!},$$

pour $k \to \infty$, $p \to 0$, mais $kp \to \lambda$.

Il y a plusieurs ensembles statistiques qui peuvent être représentés d'une manière approchée par ces fonctions ; il est donc assez naturel d'employer ces deux fonctions comme fonctions génératrices pour des ensembles statistiques quelconques. GRAM, THIELE, BRUNS et CHARLIER ont employé la fonction de Gauss comme fonction génératrice.

Soit $f(x)$ une fonction arbitraire, : posons

$$f(x) = c_0\varphi(x) + c_1\varphi'(x) + c_2\varphi''(x) + \cdots \tag{b}$$

avec

$$\varphi(x) = e^{-x^2}.$$

On trouve

$$\varphi^{(n)}(x) = N_n(x).\varphi(x).$$

Les $N_n(x)$ sont des polynômes d'HERMITE. Les $N_n(x)$ satisfont à la relation :

$$N_{n+1} + 2xN_n + 2nN_{n-1} = 0.$$

On a, en particulier :

$$N_0 = 1$$
$$N_1 = -2x$$
$$N_2 = 4x^2 - 2, \qquad \text{etc.}$$

et des relations d'orthogonalité :

$$\int_{-\infty}^{\infty} N_n(x)\varphi^{(m)}(x)dx = 0, \qquad n \neq m$$

$$\int_{-\infty}^{\infty} N_n(x)\varphi^{(n)}(x)dx = 2^n n!\sqrt{\pi}.$$

Les coefficients c_n se déterminent donc d'une manière tout à fait analogue à celle des coefficients d'une série de FOURIER. On trouve :

$$c_n = \frac{1}{2^n n!\sqrt{\pi}} \int_{-\infty}^{\infty} N_n(x)f(x)dx.$$

Puisque les $N_n(x)$ sont des polynômes en x on voit que les coefficients c_n se déterminent comme des fonctions des moments de notre fonction $f(x)$.

La série (b) est appelé la série de BRUN série Φ, ou, par M. CHARLIER, une série A. Elle est employée pour représenter analytiquement une fonction de répartition statistique. Dans les applications, on ne considère que les premiers termes de la série, en admettant implicitement qu'on est conduit de cette façon à une assez bonne approximation.

On a cependant critiqué assez sérieusement l'emploi de cette série, aussi bien du point de vue mathématique que du point de vue statistique. Je cite M. LINDEBERG :

« Eine analytische Darstellung einer Verteilung kann nur dann ein grösseres Interesse darbieten, wenn daraus irgendwelche Schlüsse auf die Entstehungsbedingungen der Verteilung gezogen werden können. Aus den Bemerkungen, die u. a. Steffensen (1) über die Bruns'sche Reihe macht, geht aber hervor, dass von einer Darstellung vermittels derselben in dieser Hinsicht sehr wenig zu erwarten ist. Es wird u. a. bemerkt, dass der hieraus erhaltene analytische Ausdruck

(1) *Matematisk Iagttagelseslære*, København, 1923.

manchmal auch negative Werte annimmt. Eine solche Funktion kann kein Wahrscheinlichkeitsgesetz darstellen, und deshalb kann sie niemals irgendwelchen theoretischen Hypothesen über die Entstehungsweise der Verteilung entsprechen. Wenn der Ausdruck nirgends negativ wird, kann zwar nicht mit Sicherheit behauptet werden, dass eine theoretische Deutung unmöglich sei, wenn man den stetigen Zusammenhang mit den Fällen bedenkt, wo eine solche Deutung sicher unmöglich ist, dürfte es aber nicht zu kühn sein, jeden Versuch in dieser Richtung als hoffnungslos zu bezeichnen ».

M. CRAMÉR dans son important mémoire (1) sur ce sujet arrive à la même conclusion. Il traite particulièrement les propriétés asymptotiques de la série, c'est-à-dire il examine si les deux premiers termes donnent une meilleure approximation que le terme principal, si les trois premiers termes donnent une meilleure approximation que les deux premiers, etc. M. CRAMÉR dit (p. 28 *l. c.*) : « There are cases where no asymptotic expansion exists. Thus we must conclude that the problem calls for fresh mathematical investigation. »

Le problème de la convergence de cette série a été traité par plusieurs mathématiciens distingués : MM. GALBRUN, PHRAGMÉN, KAMEDA, VON MISES, CRAMÉR, etc.

Plusieurs de ceux qui ont donné des exemples numériques sont arrivés au même résultat. V. MISES : *Wahrscheinlichkeitsrechnung* p. 264, écrit : Zusammenfassend wird man sagen dürfen, dass der wesentliche, Verlauf einer Häufigkeitsverteilung schon durch die *ersten Glieder* unseren beiden Beispielen, schon durch das *erste* alleinigermassen wiedergegeben wird. Andrerseits müsste man in der Rechung sehr weit gehen, wenn der Konvergenz der Reihensummen gegen den rechtwinkligen Verlauf der V-Linie in Erscheinung treten soll. Daran besteht aber wohl auch kein Interesse.

Dans le même ordre d'idées CZUBER s'exprime d'une manière ana logue : *Wahrscheinlichkeitsrechnung*, t. I, p. 431 :

Indessen zeigt sich, dass die vorstehende Reihe auch schon durch $\Phi[v]$ selbst, die Verteilungsfunktion also durch die Gausssche Exponentialfunktion fast mit dem gleichen Erfolg dargestellt werden kann, so dass hier die Ψ Reihe bei der angewendeten Abkürzung nahezu *keinen* Vorteil bringt.

(1) *Skandinavisk Aktuarietidsskrift*, 1928. p. 22.

M. CHARLIER choisit la fonction de Poisson

$$\Psi(x) = \frac{\lambda^x e^{-\lambda}}{x!},$$

comme fonction génératrice et donne un développement, en faisant usage des différences de $\psi(x)$. Il pose

$$f(x) = \Psi(x) + c_1\Psi_1(x) + c_2\Psi_2(x) + \cdots$$

où

$$\begin{aligned} \Psi_1(x) &= -\Psi(x) + \Psi(x-1) \\ \Psi_2(x) &= -\Psi_1(x) + \Psi_1(x-1) \qquad (c) \\ &\cdots\cdots\cdots\cdots\cdots\cdots \\ \Psi_v(x) &= -\Psi_{v-1}(x) + \Psi_{v-1}(x-1). \end{aligned}$$

M. CHARLIER appelle cette série une série B.

Mlle POLLACZEK-GEIRINGER (1) a montré comment on peut écrire un développement analogue au développement de la série de Bruns et a indiqué les critères de convergence.

Il s'agit ici d'une fonction discontinue ou, comme on dit souvent : d'une fonction de répartition arithmétique. Des relations (c) on tire

$$\Psi_v(x) = q_v(x)\Psi(x) \qquad q_0(x) = 1.$$

Les $q_v(x)$ sont les polynômes du degré v qui jouent ici le même rôle que les polynômes d'Hermite.

On trouve la formule de récurrence

$$q_{v+1}(x) = q_v(x) + \frac{x}{\lambda} q_v(x-1)$$

en remarquant que

$$\Psi(x-1) = \frac{x}{\lambda}\Psi(x)$$

et en substituant

$$\Psi_r(x) = q_v(x)\Psi(x)$$

dans les relations (c).

(1) *Skandinavish Ahtuarietidsskrift*, 1928, p. 98. Voir aussi v. MISES, *Wahrscheinlichkeitsrechnung*, p. 265.

En particulier, on a

$$q_0(x) = 1$$
$$q_1(x) = \frac{x}{\lambda} - 1$$
$$q_2(x) = \frac{x(x-1)}{\lambda^2} - 2\frac{x}{\lambda} + 1$$
$$\cdots\cdots\cdots\cdots\cdots\cdots$$
$$q_v(x) = \frac{v!}{\lambda^v} \cdot \sum_{t=0}^{v} \binom{x}{r-t} \cdot \frac{(-\lambda)^t}{t!}$$

On a de plus, les relations d'orthogonalité

$$\sum x q_\mu(x) \Psi_v(x) = 0, \quad \mu \neq v \qquad \mu, v = 0, 1, 2 \cdots$$
$$\sum x q_v(x) \Psi_v(x) = \frac{v!}{\lambda_v}, \quad \mu = v.$$

Au moyen de ces formules d'orthogonalité, on trouve, par une méthode analogue à celle suivie pour le calcul des cœfficients de la série de Bruns, les cœfficients c_v :

$$c_v = \frac{\lambda^v}{v!} \sum x q_v(x) f(x) \qquad v = 1, 2, \cdots$$

Les $q_v(x)$ étant des polynômes en x, les cœfficients c_v sont des fonctions des $(\nu + 1)$ premiers moments de notre fonction $f(x)\lambda. = M_0 M_1, \cdots M_r$ avec $M_t = \Sigma x^t f(x)$,

Une fonction de répartition arithmétique *bornée*, c'est-à-dire une fonction, qui n'a que des valeurs différentes de zéro pour $x = 0,1,2 \cdots k$, peut toujours être développée en une série de cette forme. Quand la fonction n'est pas bornée, Mlle POLLACZEK-GERINGER (*l. c.*) a établi les critères qui fixent la possibilité de ce développement.

Nous avons utilisé jusqu'ici deux cas limites de la fonction binomiale : la fonction de GAUSS et la fonction de POISSON ; il est donc naturel de choisir maintenant d'autres fonctions de fréquence. On peut choisir par exemple, pour une fonction de répartition arithmétique bornée la fonction binomiale $f(x) = C_k^x p^x (1-p)^{k-x}$ comme fonction génératrice.

On pose :

$$F(x) = f(x) + c_1 \Delta f(x) + \cdots + c_k \Delta^k f(x)$$

où

$$\Delta f(x) = f(x-1) - f(x) = f_1(x)$$
$$\Delta^2 f(x) = \Delta f_1(x) = f_1(x-1) - f_1(x) = f_2(x)$$
$$\Delta^3 f(x) = \Delta^2 f_2(x) = f_2(x-1) - f_2(x), \qquad \text{etc.}$$

On trouve :

$$\Delta^r f(x) = q_v(x) f(x).$$

Les $q_v(x)$ sont ici des fonctions *rationnelles* de x. Elles satisfont à la formule de récurrence :

$$q_r(x) = -q_{r-1}(x) + \frac{1-p}{p} \frac{x}{k+1-x} q_{r-1}(x-1).$$

En particulier, on a :

$$q_0(x) = 1$$

$$q_1(x) = \frac{1-p}{p} \frac{x}{k+1-x} - 1$$

$$q_2(x) = \left(\frac{1-p}{p}\right)^2 \frac{x(x-1)}{(k+1-x)(k+2-x)} - 2\frac{1-p}{p}\frac{x}{k+1-x} - 1.$$

Les coefficients c_k se déterminent par des valeurs données de F(0), F(1),···F(k).

Tous ces modes de représentation des séries statistiques n'ont que des intérêts secondaires ; tous prêtent le flanc à des objections, en raison de leurs propriétés asymptotiques.

Le problème essentiel est de découvrir l'origine de la série statistique donnée. Si les résultats des tirages de boules blanches, d'une urne de composition fixe en blanches et noires, nous fournissaient des chiffres comparables à ceux donnés par l'expérience, on pourrait dire que l'on s'est engagé dans la bonne voie. L'actuaire français DORMOYS, et plus tard, le statisticien allemand LEXIS ont développé à cet effet un procédé connu sous le nom de la théorie de la dispersion.

On peut dire (1), que cette théorie est basée sur la notion de la stabilité (ou dispersion) normale. On dit que la stabilité d'une série de nombres statistiques est normale si la répartition de ces nombres autour de leur moyenne correspond sensiblement à un tableau des nombres de boules blanches tirées dans une série d'épreuves où la probabilité d'extraction d'une boule blanche a une valeur constante. Si les oscillations sont plus grandes, la dispersion est dite hypernormale (cas de LEXIS), si les oscillations sont plus petites la dispersion est dite hyponormale (cas de POISSON). Comme critère de distinction de la dispersion normale, on introduit la grandeur du coefficient de divergence Q.

1) Voir TSCHUPROW : *Nordisk Statistisk Tidsskrift*, 1922, t. I, p. 340.

Si, *à priori*, il existe une probabilité constante p, on pose

$$Q^2 = \frac{\frac{1}{r}\sum_{i=1}^{i=r}[p_i - p]^2}{\frac{1}{n}p(1-p)},$$

où r est le nombre des séries, chaque série comprenant n épreuves. Le nombre total des épreuves est rn. La fréquence de l'événement considéré dans la $i^{\text{ième}}$ série est désignée par $p_i = \frac{m_i}{n}$, où m_i est le nombre des n épreuves de la $i^{\text{ième}}$ série qui sont favorables à l'événement.

Si la probabilité *à priori* p n'est pas connue, on la remplace par la moyenne arithmétique p_0 de toutes les rn épreuves

$$p_0 = \frac{m_1 + m_2 + \cdots + m_r}{rn};$$

on pose donc

$$(c) \qquad {}^{(2)}Q^2 = \frac{\frac{1}{r-1}\sum_{1}^{r}(p_i - p_0)^2}{\frac{1}{n}p_0(1-p_0)}.$$

On dit par définition que la distribution est normale si Q est sensiblement égal à 1. Si $Q > 1$, la distribution est hypernormale, et si $Q < 1$ la distribution est hyponormale.

von BORTKIEWICZ a remarqué une petite inexactitude dans la formule (*c*), donnée par LEXIS. La formule doit être écrite sous la forme suivante :

$${}^{(r)}Q^s = \frac{\frac{1}{r-1}\sum(p_i' - p_0)^2}{\frac{r}{nr-1}p_0(1-p_0)}.$$

V. BORTKIEWICZ remarque que dans le cas de la dispersion normale ce n'est pas la grandeur ${}^{(r)}Q$ qui est égale à 1, mais en réalité la valeur probable de ${}^{(r)}Q^2$, c'est-à-dire $E[{}^{(r)}Q^2]$. Il a également généralisé ces considérations aux nombres moyens statistiques arbitraires.

TSCHUPROW a démontré que la valeur probable du dénominateur et la valeur probable du numérateur de ${}^{(r)}Q^2$ sont égales. On ne

peut pas cependant en conclure que la valeur probable de $E[{}^{(r)}Q^2]$ soit égale à 1. En général, on n'a pas

$$E\frac{n}{r} = \frac{En}{Er}.$$

Cependant, TSCHUPROW, a démontré (1) que cette égalité est vraie pour ${}^{(r)}Q^2$.

M. FRÉCHET a montré que dans plusieurs exemples numériques cités dans les traités comme réalisant un bon accord avec la théorie de la dispersion, la répartition des fréquences numériques donne une très mauvaise approximation par rapport au polynôme binomial théorique.

TSCHUPROW a publié en 1922 un mémoire important (2) intitulé *Ist die normale Stabilität empirisch nachweisbar* (*l. c.*, p. 379), dont la conclusion est la suivante : Auf die Titel dieser Abhandlung gestellte Frage lautet also die Antwort : insofern man auf die von LEXIS ausgearbeitete Methode der Berechnung des Divergenzkoeffizienten angewiesen ist, lässt sich der normale Charakter der Stabilität einer Reihe statistischer Zahlen empirisch nicht nachweisen.

Et plus loin (*l. c.*, p. 389) : Ich sehe überhaupt keinen Weg, auf welchem man, ohne über die durch das Experiment gelieferten empirischen Angaben hinauszugehen, feststellen kann, ob eine vorliegende Reihe von meinetwegen aus einer geschlossenen Urne gezogenen Nummern, in der Weise erhalten worden ist, dass die gezogenen Nummern vor der nächsten Ziehung in die Urne zurückgelegt wurden oder nicht, mit anderen Worten, ob der betreffende Reihe das stochastische Schema der normalen, oder der übernormalen, oder der unternormalen, Stabilität zu Grunde liegt. Ich wage freilich nicht zu behaupten, dass die Aufgabe unlösbar sei. Der Zweck der vorliegenden Untersuchung beschränkt sich darauf, zu zeigen, dass die Mittel, deren man sich zur Lösung bedient, — namentlich, die Berechnung des Divergenzkoeffizienten Q^2 — untauglich sind. Ob hier ein Fall von « ignorabimus » oder von « ignoramus » vorliegt, mag vorläufig dahingestellt sein.

Je vais essayer de traiter ces questions d'une autre manière.

(1) *Bulletin de l'Académie des Sciences*, Pétrograd, 1916. Voir aussi MARKOFF *l. c.* et TSCHUPROW, *Skandinavisk Aktuarietidsskrift*, 1918, p. 223.

(2) *Nordisk Statistisktidsskrift*, 1922, t. I, p. 370.

Jusqu'ici nous avons toujours raisonné sur la forme finie de notre fonction de fréquence. Or, dans nos séries statistiques les fréquences sont données pour chaque valeur de l'élément variable. Nous avons toujours des valeurs distinctes de H(0), H(1)···H(x). N'est-il pas possible de comparer ces valeurs données avec les valeurs théoriques $f(0)$, $f(1)\cdots$ d'une fonction de fréquence convenablement choisie ? On aboutit au résultat en remarquant que, si nous basons nos raisonnements non plus sur la forme finie de notre fonction de fréquence mais sur l'équation aux différences finies qui la caractérise, tout se réduit précisément à examiner la relation liant $f(x)$ à $f(x+1)$ [1].

La théorie des fonctions de fréquence pose quatre problèmes assez importants à savoir :

1. Le calcul numérique de la fonction.

2. La manière de procéder si l'on veut substituer à la fonction discontinue une fonction continue, qui pour des valeurs entières de la variable prenne justement les mêmes valeurs que la première.

3. La détermination des moments d'une fonction de fréquence donnée.

Nous avons déjà vu comment la détermination des moments d'une fonction de fréquence peut être réduite à la détermination de sa fonction caractéristique. Mais la détermination de la fonction caractéristique est très souvent plus difficile que la détermination des moments. M. STEFFENSEN a insisté sur le fait que nous sommes forcés de borner nos recherches aux fonctions de fréquence dont les moments se déterminent d'une manière assez simple. Si notre fonction de fréquence satisfait à une équation aux différences finies de la forme

$$f(x+1) = k.\frac{(x-a_1)(x-a_2)\cdots(x-a_n)}{(x-b_1)(x-b_2)\cdots(x-b_n)}f(x),$$

les fonctions de fréquence les plus simples satisferont à une équation aux différences finies de cette même forme. Dans ce cas, on peut établir facilement une formule de récurrence pour les moments et même pour les moments incomplets de la fonction de fréquence donnée.

4. Une série statistique étant donnée, chercher une fonction de fréquence qui en donne une représentation approchée et établir,

(1) Voir R. FRISCH, *loc. cit.*, *Metron*, p. 35.

s'il est possible, les critères nécessaires et suffisants pour qu'une fonction déterminée remplisse les conditions requises.

M. STEFFENSEN, qui a fait des recherches importantes sur les fonctions de fréquence, a remarqué que nous ne pouvons pas décider, dans un cas donné, si une fonction de fréquence déterminée peut être utilisée ou non. TSCHUPROW, comme nous l'avons déjà cité, écrit que la théorie de dispersion ne donne pas ces critères.

Nous traiterons d'abord ces questions pour la fonction de Poisson

$$f(x) = \frac{\lambda^x e^{-\lambda}}{x!}.$$

M. CHARLIER est le premier, à ma connaissance, qui ait proposé de trouver une fonction continue, identique à la fonction de Poisson, pour toutes les valeurs entières et positives de la variable x. M. CHARLIER introduit la fonction :

$$\psi(x) = \frac{e^{-\lambda}}{\pi}\int_0^\infty e^{\lambda\cos t}\cos[\lambda\sin t - xt]dt,$$

et donne pour $\Psi_\lambda(x)$ la série :

$$\Psi_\lambda(x) = \frac{e^{-\lambda}\sin\pi x}{\pi}\left[\frac{1}{x} - \frac{\lambda}{1!(x-1)!} + \frac{\lambda^2}{2!(x-2)} + \cdots + (-1^n)\frac{\lambda^n}{n!(x-n)} + \cdots\right]$$

Pour des valeurs positives et entières de x, on a

$$\Psi_\lambda(x) = \frac{\lambda^x e^{-\lambda}}{x!}$$

et la fonction $\Psi_\lambda(x)$ satisfait, comme a montré M. JORGENSEN, à l'équation aux différences finies :

$$F(x+1) = \frac{\lambda}{x+1}F(x) - e^{-\lambda}\frac{\sin\pi x}{\pi}.$$

Quand on cherche à introduire une fonction continue qui interpole le mieux possible une fonction discontinue, et qui, pour des valeurs entières et positives, soit identique à la fonction discontinue donnée, il semble assez naturel d'exiger que la fonction continue satisfasse à la même équation aux différences finies que la fonction discontinue donnée. On emploie, par exemple, ce principe quand on veut interpoler $x!$ par $\Gamma(1+x)$.

L'équation aux différences finies de la fonction de Poisson est

$$f(x+1) = \frac{\lambda}{x+1} f(x).$$

Une solution continue est :

$$f(x) = \frac{e^{-\lambda}\lambda^x}{\Gamma(1+x)}$$

qui pour des valeurs entières et positives de x est identique à la fonction de fréquence de Poisson. En remarquant que

$$\frac{1}{\Gamma(1+x)} = e^{Cx} \prod_1^\infty \left(1+\frac{x}{n}\right) e^{-\frac{x}{n}}$$

on a pour $f(x)$ l'expression

$$f(x) = e^{(c+\log\lambda)x-\lambda} \prod \left(1+\frac{x}{n}\right) e^{-\frac{x}{n}}.$$

$f(x)$ est une fonction entière de genre un, avec les zéros $-1, -2, -\cdots$ On a pour $f(x)$ les formules suivantes

$$f(x)\cdot f(-x) = e^{-2\lambda} \frac{\sin \pi x}{\pi x};$$

$$\prod_0^{n-1} f\left[\frac{x-s}{n}\right] = \frac{n^{x+\frac{1}{2}}}{(2\pi\lambda)^{\frac{n-1}{2}}} e^{-(n-1)\lambda} f(x).$$

La démonstration de ces formules est assez facile. On voit par exemple pour la première formule que le produit $f(x)\cdot f(-x)$ a les zéros ± 1, $\pm 2, \cdots$ La fonction $\frac{\sin \pi x}{\pi x}$ a aussi les zéros $\pm 1, \pm 2, \cdots$ On a donc

$$f(x)\cdot f(-x) = e^{ax+b} \frac{\sin \pi x}{\pi x}$$

Mais $f(x)\,f(-x)$ et $\frac{\sin \pi x}{\pi x}$ sont toutes les deux des fonctions paires, donc $a = 0$. De plus

$$f(0) = e^{-\lambda}, \qquad \left(\frac{\sin \pi x}{\pi x}\right)_{x \to 0} = 1 \qquad \text{ou} \qquad e^b = e^{-2\lambda}.$$

Au moyen de ces formules et de l'équation aux différences finies,

$$f(x+1) = \frac{\lambda}{x+1} f(x)$$

on peut serrer autant qu'on veut l'intervalle d'approximation de notre fonction.

Considérons maintenant le problème relatif à la détermination des moments de la fonction de Poisson.

Les moments d'ordre r de $f(x)$ sont

$$\sigma_r = \sum_0^\infty x^r f(x)$$

et avec, en particulier,

$$\sigma_1 = \sum_0^\infty x f(x).$$

Écrivons l'équation aux différences finies de la fonction de fréquence de Poisson de la manière suivante

$$(x + 1)f(x + 1) = f(x)$$

et prenons la somme pour x variant de 0 à ∞ :

$$\sum_0^\infty (x + 1)f(x + 1) = \lambda \sum_0^\infty f(x .$$

En remarquant que

$$\sum_0^\infty (x + 1)f(x + 1) = \sum_0^\infty x f(x) = \sigma_1,$$

on a immédiatement

$$\sigma_1 = \lambda.$$

Pour calculer les autres moments, multiplions notre équation par $(x + 1)^n$

$$(x + 1)^{n+1} f(x + 1) = \lambda x^n f(x) + \lambda \binom{1}{n} x^{(n-1)} f(x) + \cdots + \lambda f(x),$$

et faisons la somme, x variant de zéro à l'infini ; nous obtenons la formule de récurrence suivante pour σ_{n+1} :

$$\sigma_{n+1} = \lambda \sigma_n + \lambda \binom{1}{n} \sigma_{n-1} + \cdots + \lambda.$$

En particulier, on aura $\sigma_1 = \lambda$, $\sigma_2 = \lambda + \lambda^2$.

Passons aux moments incomplets.

M. FRISCH caractérise ce problème de la manière suivante (1) : « Dans l'analyse directe des problèmes relatifs aux moments *incomplets* de certaines distributions importantes, par exemple la distribution binomiale et la distribution hypergéométrique, on rencontre de très grandes difficultés qui empêchent dans bien des cas d'obtenir des expressions exactes. » Désignons le moment moyen incomplet d'ordre r par ${}_t\mu_r$, où

$$ {}_t\mu_r = \sum_{x=t}^{\infty} (x - \lambda)^r f(x) $$

et les moments incomplets d'ordre r autour de l'origine par

$$ {}_t\sigma_r = \sum_{x=t}^{\infty} x^r f(x). $$

Nous partirons de nouveau de l'équation aux différences finies de notre fonction de Poisson sous la forme :

$$ (x + 1)f(x + 1) = \lambda f(x). $$

En multipliant cette équation par $(x + 1)^n$, en faisant la somme pour x variant de t à ∞, et en remarquant que

$$ \sum_{x=t}^{\infty} (x + 1)^n f(x) = \sum_{x=t}^{\infty} x^n f(x) - t^n f(t) $$

on a la formule de récurrence :

$$ {}_t\sigma_{n+1} = \lambda_t\sigma_n + \lambda\binom{t}{1}{}_t\sigma_{n-1} + \cdots \lambda_t\sigma_o + t^{n+1}f(t) $$

En particulier

$$ {}_t\sigma_1 = \lambda_t\sigma_o + tf(t) $$

et

$$ {}_t\mu_1 = \sum_{x=t} (x - \lambda)f(x) = tf(t), $$

formule que nous retrouverons plus tard comme cas de limite de la fonction binomiale.

La formule donnant ${}_t\sigma_0$ est

$$ {}_t\sigma_o = \sum_{x=t}^{\infty} f(x) = f(t)\left[1 + \frac{\lambda}{t+1} + \frac{\lambda^2}{(t+1)(t+2)} + \cdots\right] $$

(1) RAGNAR FRISCH : *Det norske Videnskab. Akademi*, II, 1926, N° 3.

ou, si l'on veut

$$_{t}\sigma_{0} = tf(t)\int_{0}^{1}(1-x)^{t-1}e^{\lambda x}dx.$$

Pour obtenir les conditions qu'une série statistique doit remplir pour qu'elle puisse être représentée par la fonction de Poisson, écrivons l'équation aux différences finies

$$f(x+1) = \frac{\lambda}{x+1}f(x)$$

sous la forme

$$\frac{f(x+1)}{f(x)}(x+1) = \lambda,$$

où λ est le premier semi-invariant. On voit que pour toutes les valeurs de $x = 0, 1, 2, 3, \cdots$ il faut que l'expression $\frac{f(x+1)}{f(x)}(x+1)$ soit constante et égale à λ. On remarquera que λ est déterminé si l'on connaît deux termes consécutifs de notre fonction.

Soit une série statistique $(x\mathrm{H}(x))$

$$\begin{array}{l} x : 0, \ 1, \ 2, \ 3, \cdots \\ \mathrm{H}(x) : \mathrm{H}(0), \ \mathrm{H}(1), \cdots \end{array}$$

On forme le premier semi-invariant

$$\mu_1 = \frac{\sum_{0} x\mathrm{H}(x)}{\sum_{0}\mathrm{H}(x)} = \lambda.$$

Si pour toutes les valeurs de $x = 0, 1, 2 \cdots$ l'expression $\frac{\mathrm{H}(x+1)}{\mathrm{H}(x)}(x+1)$ est sensiblement égale à λ, on doit en conclure que la série donnée peut être représentée approximativement par la fonction de Poisson. On peut obtenir une valeur approchée de λ pour une valeur quelconque de

$$\frac{\mathrm{H}(x+1)}{\mathrm{H}(x)}(x+1).$$

Examinons quelques exemples.

1. Considérons d'abord un exemple donné par von BORTKIEWICZ :

$$\begin{array}{l} x : \ 0, \quad 1, \quad 2, \quad 3, \quad 4, \\ \mathrm{H}(x) : 109, \ 65, \ 22, \ 3, \ 1, \end{array} \qquad \sum \mathrm{H}(x) = 200.$$

On trouve $\mu_1 = 0{,}61$. Désignons l'expression $\frac{H(x+1)}{H(x)}(x+1)$ par $\Psi(x)$ et calculons $\Psi(x)$ pour $x = 0, 1, 2, 3, 4$. On trouve

$$\Psi(0) = 0{,}60, \quad \Psi(1) = 0{,}68, \quad \Psi(2) = 0{,}41, \quad \Psi(3) = 1{,}3.$$

On doit attendre une assez bonne approximation si l'on utilise la fonction de Poisson $f(x) = \frac{\lambda^x e^{-\lambda}}{x!}$, avec $\lambda = 0{,}61$. On trouve :

x :	0,	1,	2,	3,	4,
200 $f(x)$:	108,7,	66,3,	20,2,	4,1,	0,6

2. Nombre de suicides des personnes assurées dans une société d'assurances sur la vie en Norvège, au cours de la période 1900-1929. Soit $H(x)$ le nombre des années dans lesquelles on a observé x suicides. Par exemple $H(1) = 6$, signifie que, parmi les 30 années considérées, il y en a eu 6, au cours de chacune desquelles on a noté un suicide.

x :	0,	1,	2,	3,	4,	5,	6
$H(x)$:	3,	6,	8,	7,	3,	2,	1

On trouve $\mu_1 = \lambda = 2{,}6$ et $\Psi(0) = 2$, $\Psi(1) = 2{,}5$, $\Psi(2) = 2{,}6$, $\Psi(3) = 1{,}7$, $\Psi(4) = 3{,}3$, $\Psi(5) = 3{,}0$.

L'approximation par la fonction de Poisson $f(x) = \frac{\lambda^x e^{-\lambda}}{x!}$ avec $\lambda = 2{,}6$, en négligeant les décimales est :

x :	0,	1,	2,	3,	4,	5,	6
30 $f(x)$:	2,	6,	7,	6,	4,	2,	1

3. Le nombre (x) de particules en alpha rayonnées par le polonium dans un intervalle donné, déterminé par MM. RUTHERFORD et GEIGER :

x :	0,	1,	2,	3,	4,	5,	6,	7,	8,	9,	10,	11,	12,	13,	14
$H(x)$:	57,	203,	383,	525,	532,	408,	273,	139,	45,	27,	10,	4,	0,	1,	1

On trouve $\mu_1 = 3{,}88 = \lambda$. Désignons l'expression $\frac{H(x+1)}{H(x)}(x+1) : \lambda$ par $\alpha(x)$, et appelons $\alpha(x)$, les cœfficients d'approximation. Les $\alpha(x)$ pour

$x = 0, 1, 2, \cdots$ doivent être sensiblement égaux à 1, si notre série est représentée convenablement par la fonction de Poisson.

On trouve :

$$\begin{array}{lllll} \alpha(0) = 0{,}92, & \alpha(1) = 0{,}97, & \alpha(2) = 1{,}05, & \alpha(3) = 1{,}05, & \alpha(4) = 0{,}94, \\ \alpha(5) = 1{,}04, & \alpha(6) = 0{,}92, & \alpha(7) = 0{,}67, & \alpha(8) = 1{,}4, & \alpha(9) = 0{,}96, \\ \alpha(10) = 1{,}14, & \alpha(11) = 0, & \alpha(12) = \infty, & \alpha(13) = 3{,}62. & \end{array}$$

L'approximation par la fonction de Poisson $f(x) = \frac{e^{-\lambda}\lambda^x}{x!}$ avec $\lambda = 3{,}88$ est :

x :	0,	1,	2,	3	4,	5,	6,	7,	8,	9,	10,	11,	12,	13,	14,
2.608 $f(x)$:	53,	205,	400,	520,	507,	396,	257,	143,	70,	30,	12,	4,	1,	0,	0,

La fonction de Poisson peut être appliquée, même si l'on ne connaît qu'une petite partie d'une série statistique, parce que λ peut être déterminé à partir d'une seule valeur de $\frac{H(x+1)}{H(x)}(x+1)$.

Examinons maintenant quelques autres fonctions de fréquence importantes.

La fonction de fréquence binomiale :

$$f(x) = C_k^x p^x (1-p)^{k-x} \qquad 0 \leqq x \leqq k,$$

représente la probabilité pourque l'événement fortuit E avec la probabilité p se présente x fois, l'événement contraire avec la probabilité $1-p$ se présentant alors $k-x$ fois dans une suite de k épreuves. Notre fonction $f(x)$ satisfait à l'équation aux différences finies

$$f(x+1) = \frac{k}{1-p}\,\frac{k-x}{x+1}\,f(x);$$

si $k \to \infty$, $p \to 0$ *et* $pk \to \lambda$ on à l'équation aux différences finies

$$f(x+1) = \frac{\lambda}{x+1}\,f(x),$$

c'est-à-dire l'équation aux différences finies de Poisson.

Déterminons d'abord les moments d'ordre r autour de l'origine

$$\sigma_r = \sum_0^k x^r f(x).$$

Nous donnerons une formule de récurrence pour σ_r. Ecrivons notre équation aux différences finies sous la forme :

$$(1 - p)(x + 1)f(x + 1) = pkf(x) - pxf(x).$$

En multipliant notre équation par $(x + 1)^n$, nous aurons :

$$(1 - p)(x + 1)^{n+1}f(x + 1) = pkx^n f(x) + C_n^1 pkx^{n-1}f(x) + \cdots + pkf(x) - px^{n+1}f(x) - C_n^1 px^n f(x) - \cdots - pxf(x).$$

Faisons la somme, x variant de 0 à k, et remarquons que

$$\sum_0^k (x + 1)^{n+1}f(x = 1) = \sum_0^k x^{n+1}f(x) = \sigma_{n+1}, \qquad f(k + 1) = 0,$$

et

$$\sum_0^k f(x) = 1,$$

on a

$$(1 - p)\sigma_{n+1} = pk\sigma_n + C_n^1 pk\sigma_{n-1} + \cdots + pk - p\sigma_{n+1} - C_n^1 p\sigma_n - \cdots - p\sigma_1 ;$$

ou la formule de récurrence de σ_{n+1} :

$$\sigma_{n+1} = pk\sigma_n + C_n^1 p\sigma_{n-1} + \cdots + pk - pC_n^1\sigma_n - pC_2^n\sigma_{n-1} - \cdots - p\sigma_1$$

En particulier, on a

$$\sigma_1 = pk$$
$$\sigma_2 = pk\ \sigma_1 + pk - p\sigma_1 = p^2k^2 - p^2k + pk.$$

Nous voyons que si nous posons $pk = \lambda$ et $p = 0$ dans notre formule de récurrence, nous retrouvons la formule pour les moments d'ordre $n + 1$ de la fonction de fréquence de Poisson.

Abordons la détermination des moments *incomplets* de la fonction binomiale.

Je cite encore M. R. FRISCH : « Avant de terminer, je me permettrai de dire quelques mots sur les moments incomplets de la distribution binomiale.

Le moment moyen incomplet d'ordre h de cette distribution est défini par

$$\mu_h = \sum_{r=t}^{s} (v - sp)^h Pv \qquad Pv = C_x^v p^v q^{s-v}$$

où la limite inférieure de sommation t est un quelconque des nombres 0, 1, $\cdots$ s. Le problème de trouver une expression pour ces moments est naturellement beaucoup plus complexe que le problème analogue pour les moments complets. Si l'on se borne à considérer des expressions algébriques, on ne connaît actuellement que l'expression d'un seul moment incomplet, c'est l'expression du moment incomplet d'ordre un. Pour ce moment, on a

$$_t\mu_1 = sC_{s-1}^{t-1} p^t q^{s-t+1}$$

Pour la démonstration de cette formule, je renvoie à un article publié dans le *Skandinavisk Aktuarietidsskrift* (1) ».

Désignons le moment *incomplet* d'ordre n autour de l'origine par $_t\sigma_n$; nous aurons le moment moyen incomplet d'ordre n :

$$_t\mu_n = \sum_{x=t}^{x=k} (x - kp)^n f(x) = \sum_{x=t}^{x=k} x^n f(x) - C_n^1 kp \sum_{x=t}^{x=k} x^{n-1} f(x) + \cdots$$
$$(-1)^n (kp)^n \sum_{x=t}^{x=k} f(x) = {}_t\sigma_n - C_n^1 kp\, {}_t\sigma_{n-1} + \cdots (-1)^n (kp)^n \sigma_0,$$

où

$$_t\sigma_n = \sum_{x=t}^{x=k} x^n f(x).$$

Mais on obtiendra facilement d'une manière analogue, une formule de récurrence pour $_t\sigma_n$. Passons à l'équation aux différences finies de la fonction binomiale sous la forme :

$$(1 - p)(x + 1) f(x + 1) = pkf(x) - pxf(x).$$

en multipliant l'équation par $(x + 1)^n$ on a

$$(1 - p)(x + 1)^{n+1} f(x + 1) = pkx^n f(x) + C_n^1 pkx^{n-1} f(x) + \cdots$$
$$+ pkf(x) - px^{n+1} f(x) - C_n^1 px^n f(x) - \cdots - pxf(x).$$

Sommons par rapport à x qui varie de t à k, et remarquons que

$$\sum_{x=t}^{x=k} x^{n+1} f(x) = \sum_{x=t}^{x=k} (x + 1)^{n+1} f(x + 1) + t^{n+1} f(t)$$

(1) *Matematikerkongressen i Köbenhavn*, 1925 p. 372. Voir aussi *Comptes Rendus* (Paris). Séance du 17 nov. 1924.

nous avons :

$$(1-p)\,{}_t\sigma_{n-1} - (1-p)t^{n+1}f(t) = p k_t\sigma_n + C_n^1 p k_t\sigma_{n-1} + \cdots + p k_t\sigma_0 - p_t\sigma_{n+1} - C_n^1 p_t\sigma_n - \cdots - p_t\sigma_1$$

ou

$${}_t\sigma_{n+1} = (1-p)t^{n+1}f(t) + p k_t\sigma_n + C_n^1 p k_t\sigma_{n-1} + \cdots + p k_t\sigma_0 - p C_n^1\,{}_t\sigma_n - p C_n^2\,{}_t\sigma_{n-1} - \cdots - p_t\sigma_1.$$

En particulier, pour $n = 0$, on a

$${}_t\sigma_1 = p k_t\sigma_0 + (1-p)t f(t)$$

ou

$${}_t\mu_1 = {}_t\sigma_1 - p k_t\sigma_0 = \sum_{x=t}^{x=k} (x - kp) f(x) = (1-p)t f(t),$$

qui est la formule obtenue par M. FRISCH.

Pour ${}_t\sigma_0$ on a, en désignant par $F(\alpha, \beta, \gamma, x)$ la série hypergéométrique

$${}_t\sigma_0 = \sum_{x=t}^{x=k} f(x) = f(t)\left[1 + \frac{k-t}{t+1}\cdot\frac{p}{1-p} + \frac{(k-t)(k-t-1)}{(t+1)(t+2)}\left(\frac{p^2}{1-p}\right) + \cdots\right]$$

ou

$${}_t\sigma_0 = f(t)F\left(1, t-k, t+1, \frac{p}{1-p}\right) = f(t)(1-p)F(1, k+t, t+1, p)$$

ou

$${}_t\sigma_0 = f(t)(1-p)t\int_0^1 (1-x)^{t-1}(1-px)^{-k-1}dx, \,(1)$$

En vue de trouver les critères caractéristiques d'une série statistique qui peut être représentée par la fonction binomiale, nous écrirons l'équation aux différences finies

$$f(x+1) = \frac{p}{1-p}\,\frac{k-x}{x+1}\,f(x)$$

(1) V. Poul Qvale : *Skand. Aktuarieitdsskrift*, 1932 p. 202.

d'une autre manière. Nous avons déterminé les deux premières semi-invariants μ_1 et μ_2 à l'aide de p et k

$$\mu_1 = kp.$$
$$\mu_2 = kp(1 - p),$$

d'où l'on tire

$$p = \frac{\mu_1 - \mu_2}{\mu_1}, \; k = \frac{\mu_1^2}{\mu_1 - \mu_2}.$$

Introduisons ces valeurs de p et k dans notre équation aux différences finies; elle prend alors la forme suivante

$$\frac{f(x+1)}{f(x)}(x+1) + \frac{\mu_1 - \mu_2}{\mu_2} \cdot x = \frac{\mu_1^2}{\mu_2}.$$

Cette équation montre qu'il suffit de connaître les valeurs de $f(x)$ pour 3 valeurs successives de x pour déterminer μ_1 et μ_2. La série $f(0)$, $f(1), \cdots f(k)$ est donc déterminée par trois termes successifs.

Désignons l'expression

$$\frac{f(x+1)}{f(x)}(x+1) + \frac{\mu_1 - \mu_2}{\mu_2} \cdot x$$

par $\Psi(x)$. $\Psi(x)$ est constant et égal à $\frac{\mu_1^2}{\mu_2}$ pour toutes les valeurs $x = 0, 1, 2, \cdots k$; l'expression

$$\frac{\Psi(x)}{\frac{\mu_1^2}{\mu_2}} \equiv \alpha(x)$$

sera donc égale à 1 pour toutes les valeurs de x.

Soit une série statistique $[x, H(x)]$, $x = 0, 1, 2, \cdots k$. Les deux premiers semi-invariants de cette série sont :

$$\mu_1 = \frac{\sum_0^k xH(x)}{\sum_0^k H(x)}$$

$$\mu_2 = \frac{\sum_0^k (x - \mu_1)^2 H(x)}{\sum_0^k H(x)}.$$

Formons l'expression :

$$\frac{H(x+1)}{H(x)}(x+1)+\frac{\mu_1-\mu_2}{\mu_2}x=\Psi(x).$$

Si $\Psi(x)$ est sensiblement constant pour toutes les valeurs $x=0,1,\cdots k$

$$\frac{\mu_1^2}{\mu_2},\qquad \text{ou}\qquad \alpha(x)=\frac{\Psi(x)}{\frac{\mu_1^2}{\mu_2}}$$

le cœfficient d'approximation sera sensiblement égal à 1 ; nous en concluons que notre série peut être représentée d'une manière approchée par la fonction de fréquence binomiale $f(x)=C_k^x p^x(1-p)^{k-x}$ de la variable x.

Lorsqu'on ne connaît qu'un petit nombre de termes d'une série statistique, on peut, comme on l'a déjà remarqué, déterminer des valeurs approchées de μ_1 et μ_2. Nous voyons que si $\mu_1=\mu_2$, nous retrouvons la fonction dePoisson.

Donnons quelques exemples.

Nous avons déjà considéré la série statistique donnant le nombre de glandes de Müller de 2.000 truies

x :	0,	1,	2,	3,	4,	5,	6,	7,	8,	9,	10
$H(x)$:	15,	209,	365,	482,	414,	277,	134,	72,	22,	8,	2,

Nous avons ici

$$\mu_1=3{,}5,\ \mu_2=2{,}8.$$

Les cœfficients d'approximation prennent les valeurs

$$\alpha(0)=4,\ \alpha(1)=0{,}9,\ \alpha(2)=1{,}1,\ \alpha(3)=0{,}97,\ \alpha(4)=0{,}93,\ \alpha(5)=0{,}86,$$
$$\alpha(6)=1{,}1,\ \alpha(7)=0{,}77,\ \alpha(8)=0{,}97,\ \alpha(9)=0{,}7.$$

L'approximation par la formule binomiale est :

x :	0,	1,	2,	3,	4,	5,	6,	7,	8,	9,	10
2000 $f(x)$:	41,	177,	364,	468,	423,	286,	150,	62,	21,	6,	1.

Un exemple d'une nature tout à fait différente est le suivant : Le livre *Thesaurus logarithmorum* de VEGA contient dans sa première partie le logarithme des nombres de 1000 à 11000 avec dix chiffres décimaux. Dans chacune des pages du livre, les logarithmes sont rangés

en 5 colonnes, et chaque colonne contient 60 lignes. BRUNS a examiné combien de fois dans les premières 1000 colonnes une ligne finit par un zéro.

BRUNS donne le tableau suivant ; x est le nombre des zéros, $H(x)$, la fréquence de x :

x	: 1,	2,	3,	4,	5,	6,	7,	8,	9,	10.	11,	12,	13,	14
$H(x)$	: 6,	36,	78,	149,	161,	183,	134,	114,	74,	34,	19,	10,	0,	2

Le hasard est ici exclu. On peut cependant se poser, comme le fait BRUNS, le problème suivant : peut-on comparer notre série à une variable aléatoire $(x, f(x))$, où $f(x)$ est la fonction de fréquence binomiale variable $f(x) = C_k^x p^x (1-p)^{1-x}$? Formons

$$\frac{H(x+1)}{H(x)}(x+1) + \frac{\mu_1 - \mu_2}{\mu_2} x = \frac{\mu_1^2}{\mu_2}$$

et posons

$$\alpha(x) = \frac{\Psi(x)}{\frac{\mu_1^2}{\mu_2}}.$$

On trouve

$$\mu_1 = 6,02,\ \mu_2 = 4,95,\ \mu_3 = 4,27$$

Si notre série était susceptible d'être représentée par la variable aléatoire citée, nous aurions les $\alpha(x)$ sensiblement égaux à 1. On trouve

$\alpha(1) = 1,7$, $\alpha(2) = 0,96$, $\alpha(3) = 1,14$, $\alpha(4) = 0,88$, $\alpha(5) = 1,1$, $\alpha(6) = 0,9$, $\alpha(7) = 1,14$, $\alpha(8) = 1,1$, $\alpha(9) = 1,1$, $\alpha(10) = 1,1$, $\alpha(11) = 1,2$, $\alpha(12) = \infty$, $\alpha(13) = 0$.

A part les deux dernières valeurs, les α oscillent autour de 1. Pour trouver les valeurs correspondantes de la fonction binomiale, nous nous bornerons, pour avoir la valeur la plus probable, à employer la probabilité approchée $\frac{1}{\sqrt{2\pi\mu_2}} = 0,180$. La valeur correspondante de x est déterminée par

$$\mu_1 - \frac{\mu_2}{\mu_1} < x < (\mu_1 + 1) - \frac{\mu_2}{\mu_1}$$

où

$$6,02 - 0,83 < x < 7,02 - 0,83.$$

Les éléments $f(x+1)\ f(x+2)\cdots$ se calculent en partant de $f(6)$ au moyen de l'équation aux différences finies

$$f(x+1) = \left(\frac{\mu_1^2}{\mu_2} - x\,\frac{\mu_1 - \mu_2}{\mu_2}\right)\frac{f(x)}{x+1}.$$

Nous trouvons ainsi :

x	: 1,	2,	3,	4,	5,	6,	7,	8,	9,	10,	11,	12,	13,	14
$1000\,f(x)$	: 1,	47,	93,	140,	177,	180,	149,	103,	60,	31,	14,	5,	2,	0, 5

$$\Sigma\ 1000 f(x) = 1002{,}5.$$

BRUNS a choisi la probabilité $\frac{1}{10}$ pour un zéro. Il a donc $\mu_1 = 6$, $\mu_2 = 5{,}4$ et avec ces chiffres il trouve le tableau

x	: 0,	1,	2,	3,	4,	5,	6,	7,
$1000\,f(x)$	: 1,8,	12,	39,$_3$,	84,$_3$,	133,$_6$,	166,$_2$,	169,$_3$,	145,$_1$
x	: 8,	9,	10,	11,	12,	13,	14	
$1000\,f(x)$	: 106,$_3$,	68,$_2$,	38,$_6$,	19,$_6$,	10,	3,$_6$,	1,4	

Les premiers semi-invariants μ_1 des deux fonctions binomiales sont pour ainsi dire égaux, mais les seconds semi-invariants μ_2 sont en première approximation égaux au second semi-invariant de la série statistique donnée et le second semi-invariant de la seconde approximation est un peu plus grand ($5{,}4 > 4{,}95$) que celui du μ_2 de la série statistique. La différence entre les deux approximations n'a pas importance. Peut-être pourrait-on conclure que, μ_2 étant le même pour la série statistique et la première fonction binomiale, le zéro soit un peu favorisé dans notre série.

Considérons la fonction de fréquence de PASCAL :

$$f(x) = C_{k+x}^{k} p^k q^x$$

avec

$$q = 1 - p,\ (x \geqq 0),$$

qui n'est autre que le $(k+1)^{e}$ terme du développement de $(p+q)^{k+x}$, où p et q représentent respectivement les probabilités de l'événement favorable et de l'événement défavorable. Les expressions C_{k+x}^{k}, C_{k+x+1}^{k} sont les nombres du triangle arithmétique de PASCAL.

$$\sigma_r = \sum_{0}^{\infty} x^r f(x).$$

On voit facilement que l'expression $C_{k+x+1}^{k} p^k q^{x+1}$ est liée à celle de $f(x)$ par l'équation aux différences finies

$$f(x+1) = q\,\frac{k+1+x}{x+1}\,f(x).$$

On voit que si l'on pose $q = \frac{\lambda}{k}$ et $k \to \infty$ on retrouve l'équation aux différences finies de Poisson.

Nous déterminerons d'abord les moments.

Ecrivons notre équation aux différences finies sous la forme

$$(x+1)f(x+1) = q(k+1)f(x) + qxf(x).$$

En multipliant cette équation par $(x+1)^n$, on aura

$$\begin{aligned}(x+1)^{n+1}f(x+1) = {}& \\ & q(k+1)x^n f(x) + q(k+1)\binom{n}{1}x^{n-1}f(x) + \cdots + q(k+1)f(x) + qx^{n+1}f(x) \\ & + q\binom{n}{1}x^n f(x) + \cdots + qxf(x).\end{aligned}$$

Sommons, x variant de x à l'infini ; nous aurons :

$$\begin{aligned}\sigma_{n+1} = q(k+1)\sigma_n + q(k+1)C_n^1\sigma_{n-1} + \cdots + q(k+1) + q\sigma_{n+1} \\ + qC_n^1\sigma_n + \cdots + q\sigma_1.\end{aligned}$$

ou

$$\begin{aligned}(1-q)\sigma_{n+1} = q(k+1)\sigma_n + q(k+1)C_n^1\sigma_{n-1} + \cdots q(k+1) \\ + qC_n^1\sigma_n + \cdots + q\sigma_1.\end{aligned}$$

En particulier, nous aurons

$$\sigma_1 = q\,\frac{k+1}{1-q}$$

$$\sigma_2 = \frac{q^2(k+1)(k+2)}{(1-q)^2} + \frac{q(k+1)}{1-q},$$

d'où, pour les deux premiers semi-invariants :

$$\mu_1 = q\,\frac{k+1}{1-q}$$

$$\mu_2 = q\,\frac{(k+1)}{(1-q)^2}.$$

On voit ici que

$$\mu_2 > \mu_1.$$

Pour déterminer les moments *incomplets*, nous suivrons une méthode analogue à celle présentée au cours de cette étude, et nous aurons

$$(1-q)_t\sigma_{n+1} = t^{n+1}f(t) + q(k+1)_t\sigma_n + q(k+1)C_{n1}{}_t\sigma_{n-1} + qC_1^n{}_t\sigma_n + \cdots + q\sigma_1.$$

En particulier, on aura

$$_t\sigma_1 = \frac{q(k+1)}{1-q}\,{}_t\sigma_0 + \frac{1}{1-q}\,tf(t)$$

et

$$_t\mu_1 = \sum_{x=t}^{\infty}\left(x - \frac{q(\lambda+1)}{1-q}f(x)\right) = \frac{tf(t)}{1-q}dt.$$

Pour $q = 0$ nous aurons la formule correspondante à celle de Poisson.

Pour le moment d'ordre zéro, on a

$$_t\sigma_0 = f(t)\left[1 + \frac{k+t+1}{t+1}q + \frac{(k+t+1)(k+1+2)}{(t+1)(t+2)}q^2 + \cdots\right]$$

ou

$$_t\sigma_0 = f(t)\mathrm{F}(1,\ k+t+1,\ t+1,\ q),$$

ou, enfin

$$_t\sigma_0 = f(t)t\int_0^1 (1-x)^{t-1}(1-qx)^{-k-t-1}dx.$$

Si $q \to 0$, $k \to \infty$ et $q\,k \to \lambda$ on tombe sur la forme correspondante à la fonction de Poisson soit $tf(t)\int_0^1 (1-x)^{t-1}\,e^{\lambda x}$.

Les conditions pour qu'une série statistique puisse être représentée par une fonction de Pascal, s'établissent de la manière suivante : Exprimons les constantes q et k dans l'équation aux différences finies de la fonction de Pascal au moyen des premier ssemi-invariants. Nous pouvons donc mettre notre équation aux différences finies sous la forme :

$$\frac{f(x+1)}{f(x)}(x+1) + \frac{\mu_1-\mu_2}{\mu_2}x = \frac{\mu_1^2}{\mu_2}.$$

On voit que l'équation aux différences finies pour les trois fonctions de fréquence, la fonction binomiale et celles de Poisson et de Pascal, ont la même forme, mais qu'on a respectivement $\mu_1 \gtreqless \mu_2$. L'expression

$$\Psi(x) = \frac{f(x+1)}{f(x)}(x+1) + \frac{\mu_1-\mu_2}{\mu_2}x$$

est constante et égale à $\frac{\mu_1^2}{\mu_2}$, ou $\varkappa(x) = \frac{\Psi(x)}{\frac{\mu_1^2}{\mu_2}}$ est égale à 1 pour toutes les valeurs de x.

Pour qu'une série statistique $(x, H(x))$, $x = 0, 1, 2\ldots$ puisse être représentée d'une manière approchée par une de ces trois fonctions de fréquence : la binomiale, celle de Poisson ou celle de Pascal il faut que les

$$\varkappa(x) = \frac{\frac{H(x+1)}{H(x)}(x+1) + \frac{\mu_1 - \mu_2}{\mu_2} x}{\frac{\mu_1}{\mu_2}}$$

soient sensiblement égaux à 1 pour les valeurs $x = 0, 1, 2, \cdots$ avec, respectivement $\mu_1 \gtreqless \mu_2$.

Considérons l'exemple suivant donné par v. BORTKIEWICZ, et concernant le nombre des suicides d'enfants en Prusse dans les années 1869-1893.

Soit $H(x)$ le nombre des années durant lesquelles on a observé x suicides. Les résultats des observations sont donnés dans le tableau suivant :

x :	0,	1,	2,	3,	4,	5,	6
$H(x)$:	4,	8,	5,	3,	4,	0,	1

On trouve ici

$$\mu_1 = 1{,}96, \quad \mu_2 = 2{,}46$$
$$\mu_2 > \mu_1.$$

Nous utilisons en ce cas la fonction de Pascal et nous trouvons pour les $\varkappa(x)$

$$\varkappa(0) = 1{,}3, \quad \varkappa(1) = 0{,}7, \quad \varkappa(2) = 0{,}9, \quad \varkappa(3) = 0{,}3,$$
$$\varkappa(4) = 0{,}1, \quad \varkappa(5) = \infty, \quad \varkappa(6) = 0{,}8.$$

Les fréquences théoriques sont :

x :	0,	1,	2,	3,	4,	5.	6
$25f(x)$:	4,	6,2,	5,5,	3,6,	2,	0,9,	0,4

v. BORTKIEWICZ a, de son côté, fait état de la fonction de Poisson, car μ_1 n'est pas sensiblement différent de μ_2. Il trouve le tableau :

x :	0,	1,	2,	3,	4,	5,	6
$25f(x)$:	3,4,	6,8,	6,8,	4,5,	2,2,	0,9.	0,3

La fonction de Pascal donne donc une meilleure approximation que la fonction de Poisson.

Considérons enfin la fonction de fréquence hypergéométrique

$$f(x) = \frac{C_k^x C_h^{m-x}}{C_{k+h}^m};$$

$f(x)$ est la probabilité de tirer x boules blanches d'une urne qui contient k boules blanches et h boules noires, en faisant m tirages sans remettre dans l'urne les boules tirées.

La fonction $f(x)$ satisfait à l'équation aux différences finies

$$f(x+1) = \frac{(k-x)(m-x)}{(x+1)(h-m+1+x)} f(x) \qquad (m \geqq x \geqq 0).$$

En divisant, dans la fraction

$$\frac{(k-x)(m-x)}{(x+1)(h-m+1+x)},$$

le numérateur et le dénominateur par $(h+k)$, et en posant

$$\frac{k}{h+k} = p, \quad \frac{h}{h+k} = 1-p \text{ et } h+k \to \infty$$

on a la fonction binomiale comme cas limite.

Pour déterminer les moments, écrivons l'équation aux différences finies sous la forme :

$$\begin{aligned}&((h-m)+x+1)(x+1)f(x+1)\\ &= k \cdot mf(x) - (k+m)xf(x) + x^2 f(x).\end{aligned}$$

En multipliant cette équation par $(x+1)^n$ on a :

$$\begin{aligned}&(h-m)(1+x)^{n+1}f(x+1) + (1+x)^{n+2}f(x+1)\\ &= k \cdot mx^n f(x) + kmC_n^1 x^{n-1}f(x) + \cdots + kmf(x)\\ &\quad - (k+m)x^{n+1}f(x) - (k+m)C_n^1 x^n f(x) - \cdots - (k+m)xf(x)\\ &\quad + x^{n+2}(x) + Cf_n^1 x^{n+1}f(x) + \cdots + x^2 f(x).\end{aligned}$$

Faisant la somme, x variant de 0 à m, on aura :

$$\begin{aligned}&(h-m)\sigma_{n+1} + \sigma_{n+2} = km\sigma_n + kmC_n^2\sigma_{n-1} + \cdots + km - (k+m)\sigma_{n+1}\\ &\quad - (k+m)C_n^1\sigma_n - \cdots - (k-m)\sigma_1 + \sigma_{n+2} + C_n^1\sigma_{n+1} + \cdots + \sigma_2;\end{aligned}$$

ou

$$\begin{aligned}(h+k-\mathrm{C}_n^1)\sigma_{n+1} &= k\cdot m\sigma_n + k\cdot m\mathrm{C}_n^1\sigma_{n-1} + \cdots + k\cdot m \\ &- \mathrm{C}_n^1(k+m)\sigma_n - \mathrm{C}_n^2(k+m)\sigma_{n-1} - \cdots - (k+m)\sigma_1 \\ &+ \mathrm{C}_n^2\sigma_n + \mathrm{C}_n^3\sigma_{n-1} + \cdots + \sigma_2.\end{aligned}$$

Pour $n = 0$, nous avons

$$\sigma_1 = \frac{k\cdot m\cdot}{k+h}$$

Pour $n = 1$, nous avons

$$(h+k-1)\sigma_2 = (km-(k+m))\sigma_1 + k\cdot m$$

ou

$$\sigma_2 = \frac{km(km+h-m)}{(h+k)(h+k-1}$$

Les moments incomplets se déterminent d'une manière tout à fait analogue, en remarquant que

$${}_t\sigma_n = \sum_{x=t}^{m} x^n f(x) = \sum_{x=t}^{m} (x+1)^n f(x+1) + t^n f(t).$$

On trouve

$$\begin{aligned}&(h-m)({}_t\sigma_{n+1} - t^{n+1}f(t)) + ({}_t\sigma_{n+2} - t^{n+2}f(t)) \\ &= k\cdot m_t\sigma_n + km\mathrm{C}_n^1\cdot{}_t\sigma_{n-1} + \cdots + km_t\sigma_0 - (k+m)_t\sigma_{n+1} \\ &- (k+m)\mathrm{C}_n^1{}_t\sigma_n - \cdots - (k+m)_t\sigma_1 \\ &+ {}_t\sigma_{n+2} + \mathrm{C}_n^1{}_t\sigma_{n+1} + \cdots + {}_t\sigma_2.\end{aligned}$$

En particulier

$$(h+k)_t\sigma_1 = km_t\sigma_0 + t^2f(t) + (h-m)tf(t)$$

et

$${}_t\mu_1 = \sum_{x=t}^{m}\left(x - m\frac{k}{k+h}\right)f(x) = \frac{1}{k+h}[t^2f(t) + (h-m) + tf(t)],$$

formule qui, pour $h+k\to\infty$, $\frac{h}{h+k}\to 1-p$, donne la forme correspondante de la fonction binomiale.

Les conditions que doit remplir une série statistique donnée pour être représentée par la fonction hypergéométrique, s'obtiennent en intro-

duisant dans notre équation aux différences finies à la place des constantes h, k, m les moments factoriels ou les semi-invariants.

Ecrivons notre équation aux différences finies sous la forme

$$\frac{f(x+1)}{f(x)}(x+1)(h-m)+\frac{f(x+1)}{f(x)}(x+1)^2+(k+m)x-x^2=k\cdot m,$$

et exprimons les grandeurs $(h-m)$, $(k+m)$, km au moyen des moments factoriels

$$\sigma_{(1)}=m\frac{k}{k+h}$$

$$\sigma_{(2)}=m\frac{k}{k+h}\cdot(m-1)\frac{k-1}{h+k-1}$$

$$\sigma_{(3)}=m\frac{k}{k+h}\cdot(m-1)\frac{k-1}{h+k-1}\cdot(m-2)\frac{k-2}{h+k-2},$$

On trouve

$$h-m=\frac{2\sigma_{(1)}\sigma_{(2)}+\sigma_{(1)}\sigma_{(3)}+2\sigma^2_{(2)}-3\sigma^2_{(1)}\sigma_{(2)}-2\sigma^2_{(1)}\sigma_{(3)}+\sigma_{(1)}\sigma^2_{(2)}+\sigma_{(2)}\sigma_{(3)}}{\sigma^2_{(1)}\sigma_{(2)}+\sigma_{(1)}\sigma_{(3)}-2\sigma^2_{(2)}}$$

$$k+m=\frac{3\sigma^2_{(1)}\sigma_{(2)}+2\sigma^2_{(1)}\sigma_{3}-\sigma_{(1)}\sigma^2_{(2)}-4\sigma^2_{(2)}-\sigma_{(2)}\sigma_{(3)}+\sigma_{(1)}\sigma_{(3)}}{\sigma^2_{(1)}\sigma_{(2)}+\sigma_{(1)}\sigma_{(3)}-2\sigma^2_{(2)}}$$

$$km=2\frac{\sigma^2_{(1)}\sigma_{(2)}+\sigma^2_{(1)}\sigma_{(3)}-\sigma_{(1)}\sigma^2_{(2)}}{\sigma^2_{(1)}\sigma_{(2)}+\sigma_{(1)}\sigma_{(3)}-2\sigma^2_{(2)}}$$

En introduisant ces valeurs dans notre équation aux différences finies on aura :

$$\begin{aligned}&\frac{f(x+1)}{f(x)}(x+1)\Big[2\sigma_{(1)}\sigma_{(2)}+\sigma_{(1)}\sigma_{(3)}+2\sigma^2_{(2)}-3\sigma^2_{(1)}\sigma_{(2)}-2\sigma^2_{(1)}\sigma_{(3)}\\&\qquad+\sigma_{(1)}\sigma^2_{(2)}+\sigma_{(2)}\sigma_{(3)}\Big]\\&+\frac{f(x+1)}{f(x)}(x+1)^2\Big[\sigma^2_{(1)}\sigma_{(2)}+\sigma_{(1)}\sigma_{(3)}-2\sigma^2_{(2)}\Big]\\&+x\Big[3\sigma^2_{(1)}\sigma_{(2)}+2\sigma^2_{(1)}\sigma_{(3)}-\sigma_{(1)}\sigma^2_{(2)}-4\sigma^2_{(2)}-\sigma_{(2)}\sigma_{(3)}+\sigma_{(1)}\sigma_{(3)}\Big]\\&-x^2\Big[\sigma^2_{(1)}\sigma_{(2)}+\sigma_{(1)}\sigma_{(3)}-2\sigma^2_{(2)}\Big]=2\Big(\sigma^2_{(1)}\sigma_{(2)}+\sigma^2_{(1)}\sigma_{(3)}-\sigma_{(1)}\sigma^2_{(2)}\Big)\Big]\end{aligned}$$

En remplaçant les moments factoriels par les moments, cette équation prend la forme :

$$\begin{aligned}
&\frac{f(x+1)}{f(x)}(x+1)\Big[\Big(\sigma_1^2-\sigma_2\Big)\Big(\sigma_2-\sigma_3\Big)-\sigma_1\Big(\sigma_1\sigma_3-\sigma_2^2\Big)\Big]\\
&+\frac{f(x+1)}{f(x)}(x+1)^2\Big[\Big(\sigma_1\sigma_3-\sigma_2^2\Big)-\Big(\sigma_1^2-\sigma_2\Big)\Big(\sigma_1-\sigma_2\Big)\Big]\\
&+x\Big[\Big(\sigma_1+\sigma_2-\sigma_2^2\Big)\Big(\sigma_1\sigma_3\Big)-\Big(\sigma_1^2-\sigma_2\Big)\Big(\sigma_2-\sigma_3\Big)\Big]\\
&-x^2\Big[\Big(\sigma_1\sigma_2-\sigma_2^2\Big)-\Big(\sigma_1^2-\sigma_2\Big)\Big(\sigma_1-\sigma_2\Big)\Big]=2\sigma_1\Big(\sigma_1\sigma_3-\sigma_2^2\Big).
\end{aligned}$$

On peut aussi, en faisant intervenir les semi-invariants au lieu des moments, mettre l'équation aux différences finies sous la forme :

$$\begin{aligned}
&\frac{f(x+1)}{f(x)}(x+1)[(\mu_1^2-\mu_2)(\mu_2-\mu_3)+2\mu_1\mu_2(2\mu_2-\mu_1)]\\
&+\frac{f(x+1)}{f(x)}(x+1)^2[\mu_1\mu_2+\mu_1\mu_3-2\mu_2^2]\\
&+x[(\mu_1^2-\mu_2)(\mu_2+\mu_3)+2\mu_1(\mu_1\mu_2-2\mu_2^2+\mu_3)]\\
&-x^2[\mu_1\mu_2+\mu_1\mu_3-2\mu_2^2]\\
&=2\mu_1(\mu_1^2\mu_2+\mu_1\mu_3-\mu_2^2).
\end{aligned}$$

Dans le cas simple où $h = k$, c'est-à-dire dans le cas de la fonction hypergéométrique symétrique, on aura

$$\mu_1=\frac{m}{2}$$

$$\mu_2=\frac{m}{4}\cdot\frac{2h-m}{2h-1}$$

En remplaçant m et h par μ_1 et μ_2, l'équation s'écrit :

$$\begin{aligned}
&\frac{f(x+1)}{f(x)}(x+1)\Big[\frac{\mu_1-\mu_1^2+4\mu_1\mu_2-3\mu_2}{\mu_1-2\mu_2}+x\Big]+\\
&+\frac{3\mu_1^2-4\mu_1\mu_2}{\mu_1-2\mu_2}-\mu_2 x-x^2=\frac{2\mu_1^3-3\mu_1\mu_2}{\mu_1-2\mu_2}.
\end{aligned}$$

En désignant par $\psi(x)$ le second membre on voit que $\psi(x)$ est constant et $\alpha(x)$ est égal à 1, pour toutes les valeurs de $x = 0, 1, 2 \cdots m$.

Donc, si l'on a une série statistique $(x, H(x))$ et que l'on forme l'expression correspondante $\alpha(x)$ et si les $\alpha(x)$ sont sensiblement égaux

à 1 pour toutes les valeurs de x, on doit s'attendre à ce que la série statistique puisse être représentée d'une manière approximative par la fonction hypergéométrique.

Comme exemple, nous traiterons la série statistique donnée par M. KARL PEARSON :

x :	0,	1,	2,	3,	4,	5,	6,	7,	8
H(x) :	215,	1724,	5262,	7440,	6371,	2950,	852,	166,	20

H(x) désigne le nombre de fois que l'avant-main dans le jeu de whist a reçu x atouts (triomphes) dans 25.000 données.

La probabilité pour que l'avant-main reçoive x atouts est égale, comme on sait à :

$$f(x) = \frac{C_{13}^{x}\, C_{39}^{13-x}}{C_{52}^{13}}.$$

Notre série peut donc être représentée approximativement au moyen de la fonction hypergéométrique. On trouve :

$$\mu_1 = 3{,}15,\quad \mu_2 = 1.66,\quad \mu_3 = 0{,}39,$$

et

$$\alpha(0) = 1{,}02,\ \alpha(1.) = 1{,}02,\ \alpha(2) = 0{,}97,$$
$$\alpha(3) = 1{,}02,\ \alpha(4) = 0{,}99,\ \alpha(5) = 1{,}02,\ \alpha(6) = 1{,}02,\ \alpha(7) = 1{,}0,\ \alpha(8) = 0{,}83.$$

Les valeurs approchées sont ;

x :	0,	1,	2,	3,	4,	5,	6,	7,	8,	9
$25000f(x)$:	220,	1732,	5163,	1649,	6264,	2990,	861,	154,	18,	1

Il n'est point nécessaire d'approximer la fonction de fréquence théorique, lorsque l'on s'est rendu compte que les $\alpha(x)$ oscillent sensiblement autour de 1. En fait, on désire connaître, l'approximation fournie par les $\alpha(x)$ (c'est-à-dire l'élément important est constitué par les oscillations peu marquées des $\alpha(x)$ autour de 1).

Rappelons-nous la remarque de TSCHUPROW :

« Je ne vois aucune manière de décider si les numéros d'une série donnée de rn numéros, tirés d'une urne fermée, ont été remis dans l'urne après le tirage ou non. Je n'ose vraiment pas prétendre que le problème soit irrésoluble. Pour le moment je ne sais pas s'il faut dire : « ignorabimus » ou « ignoramus ».

Cependant on voit que, d'après les remarques faites, la série donnée dans ces deux cas doit satisfaire à deux équations aux différences finies qui sont *essentiellement* différentes.

Donc nous oserons dire : « *scimus* ».

Certainement, nous n'avons pas résolu le problème général de la statistique mathématique posé par M. BOREL ; nous avons cherché uniquement à donner quelques indications qui pourront être utiles pour des recherches futures.

M. EINSTEIN a dit : « Dieu ne joue pas aux dés ».

Cependant, lorsqu'on s'occupe de ces problèmes, on a le sentiment vague, difficile à analyser et à définir, que chaque fois que nous réussissons à mettre en parallèle les résultats des observations avec les lois du hasard, il nous est permis d'en conclure que nous ne sommes pas très éloignés d'une interprétation satisfaisante des phénomènes statistiques étudiés.

Conférences faites à l'Institut Henri Poincaré en Avril 1932.

Manuscrit reçu le 20 Juin 1932.