Sur le calcul numérique des moments ordinaires et des moments composés d'une distribution statistique.

Par **Ragnar Frisch,** Oslo (Norvège).

§ 1. Introduction.

Considérons une distribution statistique (empirique ou théorique)

$$\begin{pmatrix} x_1 & x_2 & \dots & x_n \\ P_1 & P_2 & \dots & P_n \end{pmatrix}$$

x_ν ($\nu = 1, 2 \dots n$) désignant les valeurs que peut assumer une variable fortuite une-dimensionelle x, et P_ν ($\nu = 1, 2 \dots n$) désignant les fréquences observées (absolues ou relatives) ou bien les probabilités des valeurs x_ν.

Les moments ordinaires m_h ($h = 0, 1 \dots \infty$) de cette distribution, pris autour d'un point quelconque a sont définis

$$m_h = m_h(x-a) = \sum_{\nu=1}^{n} (x_\nu - a)^h P_\nu .$$

D'une façon analogue les moments factoriels $m_{[k]}$ pris autour d'un autre point quelconque b sont définis

$$m_{[k]} = m_{[k]}(x-b) = \sum_{\nu=1}^{n} (x_\nu - b)^{[k]} P_\nu$$

où

$$(x_\nu - b)^{[k]} = (x_\nu - b)(x_\nu - b - 1) \cdots (x_\nu - b - k + 1) \qquad (k \geqq 1)$$

$$(x_\nu - b)^{[0]} = 1 .$$

Les moments factoriels sont, comme il est bien connu, introduits par M. STEFFENSEN.[1]

Dans ce qui suit nous allons considérer des *moments composés* $m_{h[k]}$ $(h, k = 0, 1 \cdots \infty)$ définis ainsi

$$m_{h[k]} = \sum_{\nu=1}^{n} (x_\nu - a)^h (x_\nu - b)^{[k]} P_\nu .$$

On voit que les moments ordinaires et factoriels sont des cas spéciaux des moments composés

$$m_h = m_{h[0]} \qquad m_{[k]} = m_{0[k]} .$$

Tous ces moments ainsi que les coëfficients de Fourier de la distribution considérée et d'autres paramètres encore peuvent du reste être considérés comme des cas particuliers de la notion plus générale de fonction génératrice (fonction caractéristique).

En effet, soit $f(x, t)$ une fonction des deux variables x et t. Alors la fonction génératrice $\varphi(t)$ définie vis-à-vis de f et relative à la distribution considérée est

$$\varphi(t) = \sum_{\nu=1}^{n} f(x_\nu, t) P_\nu .$$

On voit par exemple que les moments ordinaires ne sont autre chose que les valeurs que prend pour $t = 0, 1 \ldots$ la fonction génératrice $\varphi(t)$ définie vis-à-vis de $f(x, t) = (x-a)^t$, et une interprétation analogue est possible pour les moments factoriels, les moments composés, les coëfficients de Fourier etc. mais nous n'y insistons pas.

Dans le cas où la distribution considérée est par parties continues on n'a qu'à remplacer le signe Σ par le signe $\int$ pris dans le sens de STIELTJES pour avoir la définition des moments, des coëfficients de Fourier et de la fonction génératrice. Pour fixer les idées nous considérons une distribution

[1] Skandinavisk Aktuarietidskrift 6 (1923) p. 73 et 7 (1924) p. 151.

discrète $\begin{pmatrix} x_1 & x_2 \ldots & x_n \\ P_1 & P_2 \ldots & P_n \end{pmatrix}$, mais les formules recurrentes que nous allons développer sont exactes aussi dans le cas d'une distribution par parties continues.

Les moments factoriels jouent un rôle important dans la théorie des distributions statistiques. En ce qui concerne les distributions théoriques il est bien connu que très souvent — par exemple pour la distribution binomiale et la distribution hypergéométrique — les moments factoriels peuvent être exprimés d'une façon simple à l'aide des paramètres qui caractérisent le schéma de réalisation. En ce qui concerne les distributions empiriques équidistantes, les moments factoriels peuvent être calculés par une sommation répétée, ce qui comporte un très grand avantage, surtout si l'on dispose d'une machine à additionner.

D'autre part, à l'aide des polynomes B de M. NÖRLUND[1] les moments ordinaires pris autour d'un point quelconque peuvent être exprimés comme des formes linéaires en des moments factoriels pris autour d'un autre point quelconque.[2]

Au point de vue théorique l'introduction des polynomes B comporte des avantages très réels pour ce passage des moments factoriels aux moments ordinaires. Au contraire le calcul numérique des moments ordinaires d'ordre supérieur à l'aide des polynomes B n'est pas très rapide si on se tient à la forme explicite dans laquelle les polynomes B sont définis.

L'objet de la présente note est de montrer comment on peut abréger le calcul numérique en se servant de certaines formules récurrentes élémentaires. Dans ces formules les moments composés jouent un certain rôle. Nous ferons aussi une remarque relative au procédé déjà connu[3] de calcul par sommation répétée.

[1] Differenzenrechnung. Berlin 1924, 6. Kap.

[2] Voir à ce sujet mon article »Semi-invariants et moments etc.», Skrifter utg. av Det Norske Videnskapsak. i Oslo II. 1926, No. 3 p. 17.

[3] Voir par exemple STEFFENSEN: Interpolationslære, Kbhvn. 1925 p. 99—104. Il y a une certaine analogie entre la disposition du schéma de M. Steffensen p. 103 et celle du schéma ci-après. Pourtant les deux méthodes ne sont pas identiques.

§ 2. Distributions quelconques (équidistantes ou non).

Dans ce paragraphe nous supposons que les moments factoriels $m_{[k]} = m_{0[k]}$ pris autour d'un point b de la distribution considérée (empirique ou théorique) soient connus numériquement. Pour en tirer les valeurs numériques $m_{[h]} = m_{h[0]}$ des moments ordinaires d'ordre supérieur, pris autour d'un autre point a, on peut utiliser une des methodes que voici.

Première méthode.

Posons pour abréger $\delta = b - a$.

Soit $B_{h[k]} \begin{pmatrix} h = 0, 1 \cdots \infty \\ k = 0, 1 \cdots h \end{pmatrix}$ le système de nombres définis par la formule recurrente

(1) $$B_{h[k]} = B_{h-1, [k-1]} + (k + \delta) B_{h-1, [k]}$$

avec les conditions initiales

$$B_{h[0]} = \delta^h \qquad B_{h[h]} = 1 .$$

L'expression explicite de $B_{h[k]}$ est

$$k!\; B_{h[k]} = \varDelta^k \delta^h = \sum_{i=0}^{k} (-)^{k+i} \binom{k}{i} (i + \delta)^h .$$

$B_{h[k]}$ n'est autre chose que la valeur du polynome $\binom{h}{k} B_{h-k}^{(-k)}(\delta)$ dans la notation de M. Nörlund.

On en tire que les nombres $B_{h[k]}$ satisfont à la formule de controle

(2) $$\sum_{k=0}^{h} h^{[k]} B_{h[k]} = (h + \delta)^h .$$

D'une façon plus générale on a

$$\sum_{k=0}^{h} x^{[k]} B_{h[k]} = (x + \delta)^h \qquad (x \text{ quelconque})$$

d'où pour $x = 1, 2 \ldots$

(3) $$k!\; B_{h[k]} = (k + \delta)^h - \sum_{i=0}^{k-1} k^{[i]} B_{h[i]} \qquad (k = 1, 2 \ldots h)$$

$$B_{h[0]} = \delta^h .$$

La valeur numérique des quantités $B_{h[k]}$ peut donc être calculée par un schéma très simple. Numérotant par exemple les lignes $h = 0, 1 \ldots$ et les colonnes $k = 0, 1 \ldots$ on inscrit dans la colonne $k = 0$ les quantités $1, \delta, \delta^2 \ldots$ dans la diagonale les quantités $1, 1, 1 \ldots$ et dans une colonne à part (pour le contrôle) les quantités $1, (1 + \delta), (2 + \delta)^2 \ldots$ On calcule alors les quantités $B_{h[k]}$ ligne par ligne de droite à gauche par (1). Chaque ligne est contrôlée par (2). On pousse le calcul jusqu'à la valeur de h qui est égale à l'ordre du moment ordinaire le plus élevé dont on a besoin.

A l'aide de (3) on peut du reste aussi si l'on veut calculer les $B_{h[k]}$ sur une ligne quelconque indépendamment des $B_{h[k]}$ sur les lignes précédentes. Au besoin la formule (3) peut être employée pour contrôler définitivement chaque quantité sur la dernière ligne du tableau calculé à l'aide de (1).

Les quantités $B_{h[k]}$ étant calculées, les moments ordinaires m_h pris autour du point a se calculent à l'aide des moments factoriels $m_{[k]}$ pris autour du point $b = a + \delta$ ainsi

(4) $$m_h = \sum_{k=0}^{h} B_{h[k]} m_{[k]} .$$

Dans le cas particulier $a = b$ donc $\delta = 0$, le calcul se simplifie considérablement. Dans ce cas les nombres $B_{h[k]}$ se confondent avec les nombres α_{hk} de Tschuprow et les nombres K_h^k et $[1_k]^{h-k}$ de M. d'Ocagne. On trouvera dans le tableau 1 ci-après les premiers nombres[1] $\alpha_{hk} = (B_{h[k]})_{\delta=0}$. Ces nombres sont les coëfficients qui permettent de développer la puissance x^h suivant les factorielles.

Tableau 1.

α_{hk}		$k=1$	2	3	4	5	6	7
		$x^{[1]}$	$x^{[2]}$	$x^{[3]}$	$x^{[4]}$	$x^{[5]}$	$x^{[6]}$	$x^{[7]}$
$h=1$	$x^1=$	1						
2	$x^2=$	1	1					
3	$x^3=$	1	3	1				
4	$x^4=$	1	7	6	1			
5	$x^5=$	1	15	25	10	1		
6	$x^6=$	1	31	90	65	15	1	
7	$x^7=$	1	63	301	350	140	21	1

Deuxième méthode.

D'après la définition du moment composé on a

$$\begin{aligned} m_{h[k]} &= \Sigma(x_\nu - a)^h (x_\nu - b)^{[k]} P_\nu = \\ &= \Sigma(x_\nu - a)^h (x_\nu - b)^{[k-1]} (x_\nu - b - k + 1) P_\nu = \\ &= \Sigma(x_\nu - a)^{h+1} (x_\nu - b)^{[k-1]} P_\nu + \\ &+ (a - b - k + 1)\, \Sigma(x_\nu - a)^h (x_\nu - b)^{[k-1]} P_\nu = \\ &= m_{h+1,[k-1]} + (a - b - k + 1) m_{h[k-1]} \end{aligned}$$

d'où

(5) $$m_{h[k]} = m_{h-1,[k+1]} + (k + \delta) m_{h-1,[k]}.$$

Cette formule peut être utilisée pour un calcul récurrent, directement en les moments eux-mêmes. On évite ainsi des calculs analogues aux calculs de la formule (4).

[1] Dans le Am. Journal of Mathem. 13 (1891) p. 152 on trouve un tableau qui s'étend jusqu'à $h=k=12$.

Le schéma de calcul sera analogue au schéma de la première méthode. On numérote les lignes $h=0, 1 \ldots$ et les colonnes $k=0, 1 \ldots$ On inscrit dans la ligne $h=0$ les moments factoriels donnés et l'on calcule les quantités $m_{h[k]}$ diagonale par diagonale (ou ligne par ligne) de droite à gauche (voir l'exemple numérique ci-après). La dernière quantité calculée sur chaque diagonale (ligne) est un moment ordinaire. Le calcul étant mis à fin, les moments ordinaires se trouvent dans la colonne $k=0$.

La dernière quantité $m_{h[0]}$ calculée sur chaque diagonale peut être contrôlée par la formule

$$m_{h[0]} = m_{0[h]} + \sum_{k=0}^{h-1} k\, m_{h-1-k,[k]} + \delta \sum_{k=0}^{h-1} m_{h-1-k,[k]}.$$

Dans le cas particulier $a=b$ on a aussi la formule de contrôle

$$m_{h[0]} = m_{0[1]} + \sum_{i=0}^{h-2} m_{i[2]}$$

comme on le voit en étendant la sommation $\sum\limits_{i=0}^{h-2}$ à la formule de définition de $m_{i[2]}$.

Comme un dernier contrôle on doit calculer indépendamment par la méthode des formules (3) et (4) le moment ordinaire le plus élevé dont on a besoin.

Il va de soi que l'on peut toujours, même dans le cas $a \neq b$ utiliser le tableau 1. Pour cela on décompose le calcul en deux parties en calculant d'abord à l'aide du tableau 1 les moments ordinaires pris autour du même point que les moments factoriels donnés, pour revenir ensuite aux moments ordinaires pris autour du point voulu, à l'aide de la formule binomiale. C'est là le procédé habituel pour passer des moments factoriels aux moments ordinaires. Ou bien on peut d'abord transformer les moments factoriels jusqu'au point voulu à l'aide de la formule binomiale (puisque les factorielles satisfont aussi à la formule binomiale). On passe ensuite aux moments ordinaires à l'aide du tableau 1. Pourtant dans le cas ou $a \neq b$ nous croyons qu'il est plus rapide d'employer une des deux méthodes ci-dessus où on a en outre la possibilité de contrôler le résultat par étapes.

§ 3. **Distributions equidistantes.**

Dans ce paragraphe nous supposons que les fréquences P_ν soient connues numériquement.

Dans le cas d'une distribution équidistante on peut toujours s'arranger de façon à ce que $x_\nu = \nu$. Désignons les sommes répétées descendantes et ascendantes de la suite $P_1 P_2 \ldots P_n$ ainsi

$$\bar{S}_k(\nu) = \sum_{i=1}^{\nu} \bar{S}_{k-1}(i)$$

$$\underline{S}_k(\nu) = \sum_{i=0}^{n-\nu} \underline{S}_{k-1}(n-i)$$

$$\bar{S}_0(\nu) = \underline{S}_0(\nu) = P_\nu .$$

Alors on démontre facilement les formules

$$\bar{S}_k(\nu) = \sum_{i=0}^{\nu-1} \binom{k-1+i}{k-1} P_{\nu-i} = \sum_{i=1}^{\nu} \binom{k-1+\nu-i}{k-1} P_i \qquad (k \geqq 1)$$

$$\underline{S}_k(\nu) = \sum_{i=0}^{n-\nu} \binom{k-1+i}{k-1} P_{\nu+i} = \sum_{i=\nu}^{n} \binom{k-1-\nu+i}{k-1} P_i . \qquad (k \geqq 1)$$

D'autre part soit b un quelconque des nombres $2, 3 \ldots n$. Le moment factoriel d'ordre k $(\geqq 0)$ pris autour de b est par définition

$$m_{[k]} = \sum_{i=1}^{n} (i-b)^{[k]} P_i = (-)^k k! \sum_{i=1}^{b-1} \binom{k+(b-1)-i}{k} P_i +$$

$$+ k! \sum_{i=b+k}^{n} \binom{i-b}{k} P_i .$$

Donc

$$(6) \qquad \frac{m_{[k]}}{k!} = \underline{S}_{k+1}(k+b) + (-)^k \bar{S}_{k+1}(b-1) . \qquad (k \geqq 0)$$

Pour $b = 0, 1$ on a simplement

$$\frac{m_{[k]}}{k!} = \underline{S}_{k+1}(k+b) . \qquad (k \geqq 0)$$

Pour calculer les moments factoriels pris autour de b, b étant un quelconque des nombres $0, 1 \ldots n$ on peut donc simplement diviser la suite $P_1 P_2 \ldots P_n$ en deux parties par une coupure à P_b et calculer les sommes répétées descendantes $\bar{S}_{k+1}(\nu)$ jusqu'à $\nu = b-1$ et les sommes répétées ascendantes $\underline{S}_{k+1}(\nu)$ jusqu'à $\nu = k+b$. Les sommes $\bar{S}$ à utiliser pour le calcul des moments factoriels se trouveront alors sur une ligne horizontale, tandis que les sommes $\underline{S}$ dont on a besoin se trouveront sur une diagonale descendante vers la droite. (Voir l'exemple numérique ci-après.)

Les moments factoriels étant calculés, on peut passer aux moments ordinaires par une des deux méthodes du § 2. Si les moments ordinaires doivent être pris autour d'un des points $0, 1 \ldots n$ (ce qui est par exemple le cas si on cherche la valeur des moments ordinaires pris autour d'une »moyenne provisoire«) il sera commode de choisir la coupure (b) en ce point. Le passage des moments factoriels (calculés à l'aide des sommes $\bar{S}$ et $\underline{S}$) aux moments ordinaires s'effectue alors simplement au moyen du tableau 1 ou bien par la formule (5) où l'on fait $\delta = 0$. Si les moments ordinaires doivent être pris autour d'un point à abcisse fractionnaire, on choisit la coupure (b) en un point voisin, à abcisse entière.

Voici un exemple simplifié pour montrer la disposition du schéma de calcul.

		$\bar{S}_1$	$\bar{S}_2$	$\bar{S}_3$	$\bar{S}_4$
$x_\nu = \nu$	P_ν	$k=1$	2	3	4
1	2	2	2	2	2
2	4	6	8	10	12
3	11	17	25	35	47
4	17	$\boxed{34}$	$\boxed{59}$	$\boxed{94}$	$\boxed{141}$
$\boxed{b=5}$	20	$\boxed{66}$			
6	19	46	$\boxed{88}$		
7	16	27	42	$\boxed{62}$	
8	8	11	15	20	$\boxed{26}$
9	2	3	4	5	6
10	1	1	1	1	1
		$\underline{S}_1$	$\underline{S}_2$	$\underline{S}_3$	$\underline{S}_4$

$$m_{[0]} = 0!\,(66 + 34) = 100$$
$$m_{[1]} = 1!\,(88 - 59) = 29$$
$$m_{[2]} = 2!\,(62 + 94) = 312$$
$$m_{[3]} = 3!\,(26 - 141) = -690.$$

Le passage aux moments ordinaires pris autour de $a = 4{,}5$ peut être effectué par la deuxième méthode (formule (5)) ainsi

$\delta = b - a = 0{,}5$

$m_{h\,[k]}$	$k=0$	1	2	3
$h=0$	100	29	312	-690
1	79	355,5	90	
2	395	623,25		
3	820,75			

ou bien par la première méthode (formules (1) et (4)) ainsi

$\delta = b - a = 0{,}5$

$B_{h\,[k]}$	$k=0$	1	2	3
$h=0$	$\delta^0 = 1$			
1	$\delta^1 = 0{,}5$	1		
2	$\delta^2 = 0{,}25$	2	1	
3	$\delta^3 = 0{,}125$	3,25	4,5	1

$$m_0 = 1 \,.\, 100 = 100$$
$$m_1 = 0{,}5 \,.\, 100 + 1 \,.\, 29 = 79$$
$$m_2 = 0{,}25 \,.\, 100 + 2 \,.\, 29 + 1 \,.\, 312 = 395$$
$$m_3 = 0{,}125 \,.\, 100 + 3{,}25 \,.\, 29 + 4{,}5 \,.\, 312 - 1 \,.\, 690 = 820{,}75\,.$$

Les méthodes précédentes sont tout-à-fait élémentaires, mais nous croyons qu'elles peuvent être de quelque utilité quand on a besoin de calculer la valeur numérique des moments d'ordre supérieur.