

REVUE DE STATISTIQUE APPLIQUÉE

H. BRENY

**Sur le caractère quasi-fortuit de la méthode de Daniels pour
l'échantillonnage des faisceaux de fibres parallèles**

Revue de statistique appliquée, tome 2, n° 2 (1954), p. 79-86

http://www.numdam.org/item?id=RSA_1954__2_2_79_0

© Société française de statistique, 1954, tous droits réservés.

L'accès aux archives de la revue « *Revue de statistique appliquée* » (<http://www.sfds.asso.fr/publicat/rsa.htm>) implique l'accord avec les conditions générales d'utilisation (<http://www.numdam.org/conditions>). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM

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

SUR LE CARACTÈRE QUASI-FORTUIT DE LA MÉTHODE DE DANIELS POUR L'ÉCHANTILLONNAGE DES FAISCEAUX DE FIBRES PARALLÈLES

par

H. BREN Y, Dr. Sc.

International Wool Secretariat ()*

La question de l'échantillonnage des populations de fibres fait l'objet d'une littérature déjà étendue (voir par exemple les références bibliographiques de Palmer et Daniels, J. Textile Inst. 38 (1947) (T 94). Un cas particulier important est celui des assemblées de fibres quasi-parallèles (rubans ou mèches) qui constituent les stades intermédiaires de la filature peignée ; dans ce cas, l'échantillonnage est susceptible d'une théorie rigoureuse basée sur des hypothèses vraisemblablement en bon accord avec la réalité. C'est dans ce sens que Daniels a proposé, en 1942, une méthode d'échantillonnage qu'il a appelée « cut-squaring » (ce que nous traduirons par « équarrissage coupé ») et qui consiste essentiellement à prendre, pour former l'échantillon, les fibres dont une extrémité se trouve entre deux sections parallèles du ruban considéré. La théorie de cet échantillonnage est ordinairement faite à partir de l'hypothèse d'uniformité longitudinale qui sert de base au modèle de Spencer-Smith et Todd et de Martindale ; on en déduit que les échantillons de Daniels sont équivalents à des échantillons fortuits.

Malheureusement, cette hypothèse d'uniformité est en contradiction grossière avec les faits dans le cas des rubans produits par les peigneuses rectilignes, éléments constitutifs (après étirage et doublage) des rubans sur lesquels se fait d'ordinaire l'échantillonnage. Nous avons donc procédé à une analyse plus précise, basée sur une hypothèse de distribution des têtes de fibre directement inspirée par le fonctionnement discontinu des peigneuses. Nous n'avons toutefois étudié que le cas de la longueur des fibres, propriété directement influencée par la non-uniformité du ruban ; pour les autres propriétés, leur corrélation avec la longueur est suffisamment faible pour rendre négligeable la correction de non-uniformité.

Mathématiquement parlant, l'article ci-dessous renferme notamment une démonstration rigoureuse de la convergence stricte (et non seulement de la quasi-convergence asymptotique) des séries employées pour représenter les moments du quotient de deux variables aléatoires ; elle n'est valable que pour le cas particulier considéré mais son principe paraît susceptible d'applications étendues.

La conclusion de cette étude est que les échantillons « cut-squaring » sont fidèles (la moyenne de la distribution de leur moyenne est celle de la population) mais plus dispersés que des échantillons fortuits ; fait qui peut avoir une certaine importance lorsqu'il s'agit de juger si un lot de ruban de peigné satisfait, en ce qui concerne la longueur, à telle ou telle spécification imposée a priori.

(*) Le texte de cette note résulte du remaniement d'un rapport présenté par l'auteur à la Commission Technique de la XXXII^e Conférence Lainière Internationale (Estoril, Lisbonne, Juin 1953). Les chiffres romains entre crochets renvoient à l'index bibliographique en fin de texte. Les variables aléatoires sont dénotées par des symboles en caractères gras.

I,1. — L'industrie textile fait grand usage, dans les stades de fabrication préliminaires à la filature, de « rubans » ou de « mèches » formés d'un nombre immense de fibres approximativement cylindriques et parallèles. Il est important, aussi bien du point de vue technique que du point de vue commercial, d'estimer les paramètres caractérisant statistiquement ces populations de fibres — les plus importants sont les longueur et diamètre moyens — et, dans ce but, d'extraire de ces populations des échantillons doués de propriétés précises.

Daniels, l'un des auteurs qui ont le plus étudié cette question de l'échantillonnage des faisceaux de fibres **I, II, II bis** a proposé l'emploi de la méthode suivante : on prend comme échantillon l'ensemble des fibres ayant une extrémité de sens déterminé (disons, pour fixer les idées, extrémité droite) intérieure à une zone comprise entre 2 sections parallèles du faisceau. Il a décrit les précautions à prendre pour que l'échantillonnage soit correct, et la technique ainsi mise au point a reçu le nom de « méthode d'équarrissage coupé » (« Cut-squaring method »).

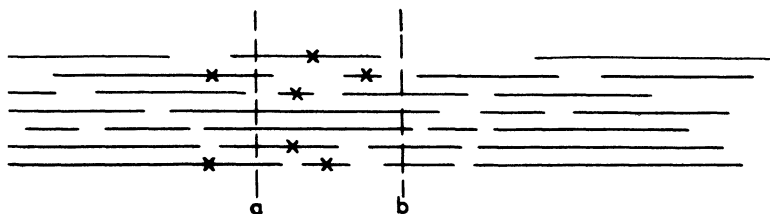


Figure 1. — La zone d'échantillonnage est comprise entre les sections a et b ; l'échantillon est formé des fibres marquées d'une croix.

Il est généralement admis que les échantillons ainsi obtenus sont équivalents à des échantillons fortuits.

Or, vu leur définition, ils ne sont vraiment tels, dans l'étude d'une certaine propriété des fibres que si la probabilité de présence de l'extrémité droite d'une fibre dans la zone d'échantillonnage est indépendante de la propriété considérée. On peut vraisemblablement admettre qu'il en est ainsi lorsque la propriété étudiée est (statistiquement) indépendante de la longueur des fibres. Pour la longueur elle-même, et pour des propriétés qui en dépendent, il n'en est rigoureusement ainsi dans tous les cas que si le faisceau de fibres est « idéal » (au sens de la théorie de Martindale) [**III, IV**] ou d'une théorie équivalente quant à l'uniformité de la distribution des têtes de fibre au long du faisceau. Nous nous proposons ici d'étudier théoriquement les propriétés de cette méthode d'échantillonnage sous des hypothèses plus larges que celles de Martindale. (Nos raisonnements pourraient aussi s'appliquer, tels quels ou après quelques modifications, à l'échantillonnage de phénomènes échelonnés dans le temps, par exemple : durées de conversations téléphoniques ; vie d'appareils en usage simultané, etc...)

I,2. — D'une manière précise, nous admettrons que le faisceau $\mathcal{F}$, est constitué de la manière suivante : il est formé par la juxtaposition de D faisceaux partiels $\mathcal{F}_i$ ($i = 1, \dots, D$; D est « l'indice de doublage »), formés eux-mêmes de fibres cylindriques parallèles, de longueurs aléatoires douées d'une fonction de répartition $F(l)$ commune aux D faisceaux $\mathcal{F}_i$ [nous admettrons l'existence d'une quantité finie l_m telle que $F(l_m - 0) = 1$] ; les fibres de $\mathcal{F}_i$ sont disposées indépendamment et au hasard le long de l'axe de ce faisceau de telle sorte que la probabilité qu'une tête de fibre (c'est-à-dire pour fixer les idées, son extrémité droite) soit entre les abscisses a et b soit égale à $(1/2L) \int_a^b p_i(t) dt$, $2L$ étant la longueur commune des D faisceaux $\mathcal{F}_i$ (cette longueur est supposée extrêmement grande, et $p_i(t)$ est une fonction positive, bornée, non nécessairement continue, mais telle que $(1/2L) \int_{-L}^L p_i(t) dt = 1$). Le faisceau $\mathcal{F}_i$ est alors entièrement caractérisé par la constante positive N_i telle que, pour toutes valeurs de a et b , $T_i(a, b) = N_i \int_a^b p_i(t) dt$ soit le nombre moyen de têtes de fibre de $\mathcal{F}_i$ situées dans l'intervalle (a, b) . On sait que, dans ces conditions, la distribution du nombre de têtes de fibres dans cet intervalle peut être considérée comme une distribution de Poisson de moyenne $T_i(a, b)$. (Pour une déduction rigoureuse de ces résultats, cfr [**V**].)

I,3. — Les hypothèses ci-dessus peuvent, pensons-nous, être considérées comme une bonne approximation de la réalité dans le cas des rubans fournis par les peigneuses RECTILIGNES ; $p_i(t)$

est alors une fonction en créneaux, nulle pour la plus grande partie des valeurs de t . Il est d'autre part bien évident que, si les hypothèses que nous venons d'énoncer sont satisfaites, la méthode de Daniels appliquée à des TETES de fibre fournit un échantillon fortuit ; il n'en est plus ainsi quand on applique cette méthode à des PIEDS de fibre ; en effet, la probabilité qu'une fibre de longueur l du faisceau $\mathcal{F}_i$ ait son pied dans l'intervalle (a, b) est évidemment :

$$(1/2L) \int_{a+l}^{b+l} p_i(t) dt = (1/2L) \int_a^b p_i(t+l) dt$$

et dépend de l (à moins toutefois que $p(t) = 1$, auquel cas le faisceau $\mathcal{F}_i$ est idéal au sens de Martindale). En pratique, on ignore totalement si l'échantillon que l'on prélève sur un ruban de peigné est un échantillon de têtes ou de pieds de fibres (ce qui revient à dire qu'au vu d'un morceau d'un tel ruban, il n'est pas possible de préciser dans quel sens celui-ci a été formé par la peigneuse). Il semble donc raisonnable d'attribuer à chacun de ces cas la probabilité $1/2$.

1.4. — Cela étant, on peut, à propos de la méthode considérée, adopter 2 points de vue légèrement différents. On peut en effet considérer que, non seulement la largeur z de la zone d'échantillonnage, mais aussi l'abscisse a du début de cette zone [qui coïncide ainsi avec l'intervalle $(a, a+z)$] sont des paramètres non aléatoires. Dans ce cas, la distribution de la moyenne l_a des longueurs des fibres de l'échantillon varie avec a et z , et sa moyenne n'est, en général, pas égale à $\bar{l} = \int_0^\infty l dF(l)$.

On peut aussi admettre — et ce point de vue, plus réaliste, sera le nôtre — que l'abscisse a du début de la zone d'échantillonnage est une variable aléatoire distribuée rectangulairement entre $-L$ et $+L$; alors il est à supposer que la valeur moyenne de l_a par rapport à tous les échantillons possibles est bien égale à $\bar{l}$ [autrement dit, la méthode d'équarrissage coupé, considérée de ce point de vue, est fidèle (« unbiaisé »)] ; l'échantillonnage n'en est pas pour cela équivalent à un échantillonnage fortuit : sa dispersion, notamment, est plus grande que celle de ce dernier.

2.1. — La fonction de répartition des longueurs dans un échantillon de pieds de fibre est évidemment la probabilité conditionnelle qu'une fibre ait une longueur $\leq l$ sous l'hypothèse qu'elle appartient à l'échantillon ; en appliquant le théorème des probabilités composées, on voit immédiatement que cette distribution est, pour le faisceau $\mathcal{F}_i$:

$$G_{a,i}(l) = \left[\int_0^l dF(u) \int_a^{a+z} p_i(t+u) dt \right] / \left[\int_0^\infty dF(u) \int_a^{a+z} p_i(t+u) dt \right]$$

2.2. — Si, d'autre part, on tient compte du caractère aléatoire de a on trouve, par le même raisonnement, la fonction de distribution :

$$\int_0^l dF(u) \int_{-L}^{L-z} da \int_a^{a+z} p_i(t+u) dt / \int_0^\infty dF(u) \int_{-L}^{L-z} da \int_a^{a+z} p_i(t+u) dt$$

Abstraction faite de termes dus aux « effets d'extrémité », qui sont de l'ordre de $1/L$, et par conséquent négligeables, cette fonction n'est autre que $F(l)$. Ceci ne signifie nullement que l'échantillonnage par équarrissage coupé soit fortuit, mais seulement que, si l'on constitue un échantillon en effectuant POUR CHAQUE FIBRE UN NOUVEAU choix aléatoire de a , on obtient un échantillon fortuit ; mais, dans la méthode de Daniels, ce n'est pas ainsi que l'on procède, on n'effectue qu'un seul choix aléatoire de a . Il nous faudra donc établir la distribution conditionnelle, a étant fixé, de tel paramètre qui nous intéresse, et étudier ensuite le cas où a peut varier d'échantillon à échantillon.

2.3. — Nous désignerons par $E_{a,i}$ l'opérateur « valeur moyenne par rapport à la répartition $G_{a,i}(l)$: $E_{a,i} f(l) = \int_0^\infty f(l) dG_{a,i}(l)$ par $\mathcal{E}$ l'opérateur « valeur moyenne par rapport à toutes les dispositions possibles des fibres » et par $\mathcal{E}(P)$ l'opérateur « valeur moyenne par rapport aux dispositions de fibres ayant la propriété P », et nous poserons :

$$E_{a,i}(l) = \lambda_{a,i} ; \quad E_{a,i}^2 l^2 = m_{a,i} ; \quad m_{a,i} - \lambda_{a,i}^2 = \sigma_{a,i}^2$$

$$E_{a,i} \exp(lt) = \varphi_i(t; a) = 1 + \lambda_{a,i} t + m_{a,i} t^2 / 2 + \dots$$

$$T_{a,i} = N_i \int_0^\infty dF(u) \int_a^{a+z} p_i(t+u) dt.$$

La distribution du nombre $n_i(a)$ des fibres de $\mathcal{F}_i$ ayant leur pied entre a et $a+z$ est, comme nous l'avons vu, pratiquement identique à la distribution de Poisson de fonction génératrice :

$$\exp. [(x-1) T_{a,i}] = \sum_{k=0}^\infty P_{k,i} x^k \quad \text{avec} \quad P_{k,i} = [\exp(-T_{a,i})] T_{a,i}^k / k!.$$

2.4. — Nous désignerons alors par $l_{j,i}$ [$j = 1, \dots, n_i(a)$] les longueurs — aléatoires — des $n_i(a)$ fibres de $\mathcal{F}_i$ ayant leur pied entre a et $a + z$, poserons :

$$S_i(a) = \sum_1^{n_i(a)} l_{j,i}; \quad n(a) = \sum_1^D n_i(a); \quad S(a) = \sum_1^D S_i(a).$$

et désignerons par $n^*(a)$ et $S^*(a)$ les variables aléatoires $n(a)$ et $S(a)$ prises conditionnellement à l'hypothèse $n(a) \neq 0$. La méthode de Daniels conduit alors à prendre comme estimateur de $\bar{l}$ la variable aléatoire $\hat{q}(a) = S^*(a)/n^*(a)$.

3.1. — Nous noterons $\Omega_i(s, t; a)$ la fonction génératrice des moments de la distribution conjointe de $n_i(a)$ et $S_i(a)$ (a étant considéré ici simplement comme un paramètre). On a :

$$\begin{aligned} \Omega_i(s, t; a) &= \mathcal{E} \exp [n_i(a)s + S_i(a)t] \\ &= \sum_0^\infty P_K \mathcal{E} \left\{ \exp (Ks + t \sum_{j=1}^{K_i} l_{j,i}) \mid n_i(a) = K \right\} \\ &= \exp \left\{ [e^s \Phi_i(t; a) - 1] T_{a,i} \right\}. \end{aligned}$$

3.2. — Pour déduire de là l'expression de la fonction génératrice des moments de $n(a)$ et $S(a)$ il est nécessaire d'expliquer les relations mutuelles des fonctions $p_i(t)$; nous admettrons que chacune d'elles se déduit d'une même fonction $p(t)$, périodique de période $2L$, en la soumettant à une translation d'amplitude aléatoire τ_i : $p_i(t) = p(t + \tau_i)$, ces amplitudes étant distribuées rectangulairement entre $-L$ et $+L$, par définition, mutuellement indépendantes (1). Il revient au même d'admettre que chacune des paires de variables $n_i(a)$, $S_i(a)$ (pour $i = 1, \dots, D$) est obtenue par un nouveau choix aléatoire de a .

Dès lors, dénotant par la suppression de l'indice i le remplacement dans les formules ci-dessus* de $p_i(t)$ par $p(t)$, et tenant compte du caractère périodique de $p(t)$, on aura, pour la fonction génératrice des moments de la distribution conjointe de $n(a)$ et $S(a)$ l'expression suivante :

$$\Omega(s, t) = \left\{ (1/2L) \int_{-L}^{+L} \Omega(s, t; a) da \right\}^D \quad (3)$$

$n(a)$ est évidemment une variable aléatoire à valeurs entières ≥ 0 et il résulte de (3) que l'on a :

$$P_0 \equiv P_T [n(a) = 0] = \left[(1/2L) \int_{-L}^{+L} \exp(-T_a) da \right]^D$$

Si donc on pose : $1 - P_0 = 1/\rho$ la fonction génératrice des moments de la distribution conjointe de $n^*(a)$ et $S^*(a)$ sera :

$$\Omega^*(s, t) = \rho \Omega(s, t) + 1 - \rho. \quad (4)$$

3.3. — Nous poserons :

$$\Omega(s, t; a) = \sum_0^\infty \sum_0^\infty \alpha_{i,K}(a) s^i t^K / (i! K!)$$

On aura donc en particulier :

$$\begin{aligned} \alpha_{1,0}(a) &= T_a & \alpha_{0,1}(a) &= T_a \lambda_a \\ \alpha_{2,0}(a) &= T_a + T_a^2 & \alpha_{0,2}(a) &= T_a m_a + T_a^2 \lambda_a^2 \\ \alpha_{1,1}(a) &= \lambda_a T_a + \lambda_a T_a^2 \end{aligned}$$

Nous poserons aussi : $\alpha_{i,K} = (1/2L) \int_{-L}^{+L} \alpha_{i,K}(a) da$

Dans ces conditions, on aura :

$$\Omega^*(s, t) = \left[\rho \sum_1^\infty \sum_1^\infty \alpha_{i,K} s^i t^K / (i! K!) + 1 - \rho \right]^D$$

3.4. — Afin d'étudier les moments de $\hat{q}(a)$, nous poserons :

$$\begin{aligned} X &= n^*(a) - \rho \alpha_{1,0} D & Y &= S^*(a) - \rho \alpha_{0,1} D \\ \Phi^*(s, t) &= \exp(-\rho \alpha_{1,0} D s - \rho \alpha_{0,1} D t) \Omega^*(s, t) \\ &= \sum_0^\infty \sum_0^\infty \mu_{i,K} s^i t^K / (i! K!) \end{aligned}$$

(*) L'opérateur $\mathcal{E}$ tient compte du caractère aléatoire de ces éléments.

On aura donc :

$$\begin{aligned}\mathcal{E} x^r y^s &= \mu_{r,s} \\ \mu_{2,0} &= \rho D [\alpha_{2,0} - \rho \alpha_{1,0}^2] \\ \mu_{1,1} &= \rho D [\alpha_{1,1} - \rho \alpha_{1,0} \alpha_{0,1}] \\ \mu_{0,2} &= \rho D [\alpha_{0,2} - \rho \alpha_{0,1}^2]\end{aligned}$$

3.5. — Il est tout d'abord évident que, si l'on ne tient compte du caractère aléatoire ni de a ni des T_i , la régression de $\hat{q}(a)$ sur $n^*(a)$ est égale à λ_a . On a par conséquent en désignant par $2L^*$ la longueur totale des intervalles I_K où $\mathcal{E} n(a) > 0$:

$$\mathcal{E} \hat{q}(a) = (1/2 L^*) \int_{\Sigma I_K} \lambda_a da. \quad (6)$$

3.61. — On a ensuite, en posant : $\alpha_{0,1} / \alpha_{1,0} = \beta$:

$$\left. \begin{aligned} [\hat{q}(a)]^2 &= [\beta + \gamma / \rho D \alpha_{1,0}]^2 [1 + X / \rho D \alpha_{1,0}]^{-2} \\ &= \beta^2 \sum_{k=0}^{\infty} [(-1)^k (K+1) X^k / (\rho D \alpha_{1,0})^k] \\ &\quad + [2\beta / \rho D \alpha_{1,0}] \sum_{k=0}^{\infty} [(-1)^k (K+1) Y X^k / (\rho D \alpha_{1,0})^k] \\ &\quad + [1 / (\rho D \alpha_{1,0})^2] \sum_{k=0}^{\infty} [(-1)^k (K+1) Y^2 X^k / (\rho D \alpha_{1,0})^k] \end{aligned} \right\} \quad (7)$$

3.62. — Les séries (7) divergent si $|X| > \rho D \alpha_{1,0}$, on ne peut donc pas leur appliquer terme à terme l'opérateur $\mathcal{E}$. Toutefois, l'on sait (cfr [VI] p. 120) que l'on peut définir des coefficients C_{iK} jouissant des propriétés suivantes :

a) Si $f(z) = \sum_{n=0}^{\infty} a_n z^n$ est une fonction holomorphe au voisinage de $z = 0$, la série de polynômes $\sum_{i=0}^{\infty} \sum_{K=0}^i C_{iK} a_K z^K$ converge uniformément dans tout domaine fermé intérieur à l'étoile rectiligne d'holomorphie, centrée en $z = 0$, de la fonction $f(z)$.

b) Quels que soient k et p ($p \geq k$) on a :

$$0 \leq \sum_{i=k}^p C_{iK} \leq 1 \text{ et } \lim_{n \rightarrow \infty} \sum_{i=k}^n C_{iK} = 1$$

3.63. — Considérons alors l'une des séries (7), par exemple la première, appelant A sa somme, on a : $A = \sum_{i=0}^{\infty} \sum_{K=0}^i (-1)^K (K+1) C_{iK} X^K / (\rho D \alpha_{1,0})^K$ et, vu la convergence uniforme de cette série vers sa limite et le caractère borné de celle-ci, on peut lui appliquer terme à terme l'opérateur $\mathcal{E}$, ce qui donne :

$$\mathcal{E} A = \sum_{i=0}^{\infty} \sum_{K=0}^i (-1)^K (K+1) C_{iK} \mu_{K,0} / (\rho D \alpha_{1,0})^K \equiv \sum_{i=0}^{\infty} \sum_{K=0}^i A_{iK}$$

3.64. — On sait (cfr [VII] p. 242) que, dans cette série double, on pourra intervertir l'ordre des sommations sous la condition que : $\lim_{m \rightarrow \infty} \left\{ \sum_{i=0}^{\infty} \sum_{K=m}^{\infty} A_{iK} \right\} = 0$

Or, la fonction génératrice des quantités $\mu_{j,K}$ (j fixé ; $K = 0, 1, \dots$) est une fonction entière, comme on le voit d'après la formule (5) ; il en résulte que la série $\sum_{K=0}^{\infty} \mu_{j,K} z^K$ est toujours absolument convergente et que, par conséquent, si k est suffisamment grand, on a : $|\mu_{0,k}| / (\rho D \alpha_{1,0})^k < 1/2^k$

On a donc, puisque $0 \leq \sum_{i=0}^p C_{iK} \leq 1$:

$$|\sum_{i=0}^p \sum_{K=m}^n A_{iK}| \leq \sum_{K=m}^n (K+1) 2^{-K} \sum_{i=0}^p C_{iK} \leq \sum_{K=m}^n (K+1) 2^{-K}.$$

Donc

$$|\sum_{i=0}^{\infty} \sum_{K=m}^n A_{iK}| \leq \lim_{p \rightarrow \infty} \lim_{n \rightarrow \infty} \sum_{K=m}^n (K+1) 2^{-K} = \sum_{K=m}^{\infty} (K+1) 2^{-K}.$$

et par conséquent : $\lim_{m \rightarrow \infty} \sum_{i=0}^{\infty} \sum_{K=m}^{\infty} A_{iK} = 0$. L'interversion des sommations est donc légitime, et on a :

$$\mathcal{E} A = \sum_{K=0}^{\infty} \sum_{i=K}^{\infty} A_{iK} = \sum_{K=0}^{\infty} (-1)^K (K+1) \mu_{0,K} / (\rho D \alpha_{1,0})^K$$

3.65. — On peut raisonner de même pour les 2 autres séries (7) et on trouve finalement :

$$\left. \begin{aligned} \mathcal{E} [\hat{q}(a)]^2 &= \beta^2 \sum_{k=0}^{\infty} (-1)^k (K+1) \mu_{K,0} / (\rho D \alpha_{1,0})^K \\ &\quad + 2\beta \sum_{k=0}^{\infty} (-1)^k (K+1) \mu_{K,1} / (\rho D \alpha_{1,0})^{K+1} \\ &\quad + \sum_{k=0}^{\infty} (-1)^k (K+1) \mu_{K,2} / (\rho D \alpha_{2,0})^{K+2} \end{aligned} \right\} \quad (8)$$

ce qui est d'ailleurs le résultat que l'on obtiendrait en appliquant ξ terme à terme aux séries (6). On tire de (8) :

$$\xi [\hat{q}(a)]^2 = \beta^2 + \left\{ 3\mu_{2,0} \beta^2 - 4\mu_{1,1} \beta + \mu_{0,2} \right\} / (\rho D \alpha_{1,0})^2 + 0 (1/D^2 \alpha_{1,0}^2) \quad (9)$$

($D \alpha_{1,0}$) est ordinairement de l'ordre de 500 et les termes non écrits dans (9) sont négligeables.

4,1. — Afin de mettre en nombres les formules ci-dessus, nous ferons les hypothèses suivantes, empruntées à la pratique du peignage de la laine, système continental :

4,11. — $F(l)$ sera représentée par une distribution rectangulaire, de moyenne $\bar{l} = 50$ mm et de coefficient de variation 46 % en écrivant :

$$dF(l) = \begin{cases} 1/80 \, dl & \text{pour } 10 \text{ mm} \leq l \leq 90 \text{ mm.} \\ 0 & \text{ailleurs.} \end{cases}$$

Cette représentation, plutôt grossière, est cependant suffisante pour notre but.

4,12. — La fonction $p(t)$ relative au ruban sortant de la peigneuse peut, vu le fonctionnement intermittent de cette machine, être considérée, dans un cas moyen, comme périodique, de période environ égale à 40 mm, et nulle sur les 4, 5 de cette période. Dans ce même cas moyen, on peut admettre que le ruban de peigné est obtenu par un doublage de 48 ($D = 48$) accompagné d'un étirage (total) de 36. Si on admet que cet étirage est parfait, on est conduit à prendre (t étant mesuré en mm) :

$$\begin{cases} p(t) = \begin{cases} 0 & 0 \leq t \leq 1.200 \\ 5 & 1.200 \leq t \leq 1.500 \end{cases} \\ p(t + K.500) = p(t) & K = \pm 1, \pm 2, \dots \end{cases}$$

4,13. — Remarquons en passant qu'avec une telle fonction $p(t)$ on peut admettre que les amplitudes τ_i des translations aléatoires imprimées aux faisceaux $\mathcal{F}_i$ sont distribuées rectangulairement entre 0 et 1500 mm sans rien changer aux résultats obtenus.

4,14. — Nous prendrons aussi $N_i = 13 \text{ mm}^{-1}$, ce qui correspond, pour le faisceau total $\mathcal{F}$, à un nombre moyen de fibres en section droite : $T = D. N_i. \bar{l} = 31200$, c'est-à-dire, si on admet pour les fibres d'un diamètre R. C. M. (1) de 25μ , à un titre de 20 ktx. (2). Enfin, nous prendrons $z = 1 \text{ mm}$, valeur qui correspond à la pratique usuelle de cette méthode d'échantillonnage.

4,2. — On peut alors, par application des formules des §§ 2,1 et 2,3, obtenir les expressions analytiques de T_a et de $G_a(l)$; elles sont très compliquées ; toutefois, il est aisé d'obtenir leurs valeurs numériques pour un réseau arbitrairement dense de valeurs de a ; on obtient de cette façon les valeurs suivantes :

$$\begin{array}{lll} \alpha_{1,0} = 13,000 & \alpha_{0,1} = 649,88 & \beta = 49,991. \\ \alpha_{2,0} = 783,78 & \alpha_{0,2} = 1990,21 & \alpha_{1,1} = 39\,186,49. \end{array}$$

d'où :

$$\mu_{2,0} = 614,076 \, D ; \quad \mu_{1,1} = 30\,737,8 \, D ; \quad \mu_{0,2} = 1567\,864 \, D \quad (10)$$

On a donc en vertu de (6) :

$$E \hat{q}(a) = 50,00 \quad (11a)$$

et en vertu de (9) et (10) :

$$[E \hat{q}(a)]^2 = 2\,502,22 \quad (11b)$$

5,1. — Les valeurs (11) se rapportent évidemment à un échantillon de pieds de fibre. Pour un échantillon de têtes de fibre, on peut recommencer les raisonnements ci-dessus, en prenant toutefois :

$$\begin{aligned} T_a &= (N/80) \int_{10}^{90} du \int_a^{a+u} p(t) \, dt \\ \lambda_a &= (1/80) \int_{10}^{90} l \, dl = 50 ; \quad m_a = (1/80) \int_{10}^{90} l^2 \, dl = 9\,100/3. \end{aligned}$$

(1) C'est-à-dire « racine du carré moyen », c'est en fait le diamètre déduit de l'aire moyenne.

(2) Grammes par mètre.

On trouve alors que le caractère aléatoire de $n^*(a)$, qui est évidemment sans influence sur $\xi \hat{q}(a)$, n'influence $\xi[\hat{q}(a)]^2$ que par des termes de deuxième ordre en $(D \alpha_{1,0})^{-1}$, c'est-à-dire que, à cette approximation, on a :

$$\xi \hat{q}(a) = 50 \quad \xi [\hat{q}(a)]^2 = 2500 + \sigma_a^2 / \rho D \alpha = 2500,85. \quad (12)$$

5.2. — Désignant alors par ξ' l'opérateur de valeur moyenne tenant compte de tous les choix aléatoires possibles, y compris celui de la définition de l'échantillon, on aura :

$$\begin{aligned} \xi' \hat{q}(a) &= 50, \\ \xi' [\hat{q}(a)]^2 &= (2502,22 + 2500,85) / 2 = 2501,54 \\ \text{Donc : } \xi' [\hat{q}(a) - \xi' \hat{q}(a)]^2 &= \text{var} [\hat{q}(a)] = 1,54 \end{aligned} \quad (13)$$

A titre de comparaison, notons que, dans un faisceau uniforme au sens de Martindale, basé sur la même distribution des longueurs de fibre et la même valeur de $\xi n(a) = 625$, les formules (12) seraient valables pour des échantillons tant de pieds que de têtes de fibre. On aurait donc :

$$\text{var} [\hat{q}(a)] = 0,85$$

5.3. — Le rapport des écarts types déduits de (13) et (14) vaut 1,35. Cette valeur, attachée évidemment au cas particulier que nous venons d'étudier, peut cependant, pensons-nous, être considérée comme typique ; s'il en est ainsi et si les hypothèses que nous avons faites sont suffisamment proches de la réalité, on peut conclure que la méthode d'échantillonnage de Daniels est fidèle, mais douée d'une dispersion d'environ 35 % plus forte que celle d'un authentique échantillonnage fortuit (1).

5.4. — Rappelons que cette conclusion s'applique à la population des échantillons définis par les opérations suivantes :

5.41. — D faisceaux de fibres parallèles $\mathcal{F}_i$ sont construits par disposition aléatoire de fibres cylindriques, parallèles, de répartition de longueurs $F(l)$, avec une densité de têtes de fibres $p(t)$, périodique de période $2L$. (La longueur de ces faisceaux est au moins $4L$; [abscisses entre $-2L$ et $2L$] ; mais si $p(t)$ est périodique, avec une période $2\omega < 2L$, on peut se contenter d'une longueur $2L + 2\omega$ [abscisses entre $-L - \omega$ et $L + \omega$]). Ces faisceaux sont, par construction, mutuellement indépendants.

5.42. — Le faisceau $\mathcal{F}_i$ est soumis à une translation d'amplitude aléatoire τ_i (distribuée rectangulairement entre $-L$ et L , ou entre $-\omega$ et $+\omega$ si $p(t)$ admet 2ω comme période). Les portions de ces faisceaux situées entre les abscisses $-L$ et $+L$ sont juxtaposées (doublage) pour former le faisceau $\mathcal{F}$.

5.43. — Une section de $\mathcal{F}$, d'abscisse a , est choisie au hasard (ce choix ne change d'ailleurs rien à la distribution des variables aléatoires étudiées) ; il est décidé, également au hasard (probabilités égales à $1/2$) si l'échantillon sera formé des fibres ayant leur tête ou leur pied entre a et $a + z$.

5.44. — Si l'échantillon ainsi « obtenu » ne comporte aucune fibre, il n'en est tenu aucun compte (en d'autres termes, les distributions étudiées sont conditionnelles à $n(a) \neq 0$).

5.45. — On prend comme estimateur de $\bar{l}$ le quotient de la somme des longueurs des fibres de l'échantillon par le nombre de ces fibres.

5.46. — Si donc, on considérait divers échantillons issus d'un même faisceau $\mathcal{F}$ — autrement dit : divers échantillons ayant en commun les opérations 5.41 et 5.42 ci-dessus — les résultats obtenus jusqu'ici ne seraient plus valables. La théorie des distributions relatives à une telle méthode n'est pas bien compliquée ; si l'on suppose, par exemple, que ces p échantillons sont pris dans $\mathcal{F}$ après le choix aléatoire des abscisses $a_1, \dots, a_p$, ($a_i < a_{i+1}$) sous la condition que :

$$\min_{2 \leq k \leq p} (a_k - a_{k-1} + z) \geq l_m$$

on devra simplement remplacer la formule (3) par la suivante :

$$\Omega(s_1, t_1; \dots; s_p, t_p) = \left\{ \prod_{k=1}^p \left[(1/2L) \int_{-L}^{+L} \Omega(s_k, t_k; a_k) da_k \right] \right\}^D$$

Toutefois, la mise en nombres de cette formule est d'une invraisemblable complication, et nous ne l'avons pas entreprise.

(1) Le fait que, dans le mode opératoire décrit par Daniels, l'échantillon est compris entre 2 sections quelque peu inclinées sur l'axe du faisceau ne change pratiquement rien à cette conclusion.

Nous tenons, en terminant cette note, à remercier l'International Wool Secretariat, dont les subventions ont rendu possibles les recherches dont elle est un résultat partiel, de même que M. F. Monfort, Directeur du Laboratoire Peltzer et Fils (Verviers), où ces recherches ont eu lieu.

INDEX BIBLIOGRAPHIQUE

- I. — PALMER R. C. and H. E. DANIELS, *Journal of the Textile Institute*, **38** (1947), T. 94.
- II. — DANIELS H. E., *ibid.* **33** (1942), T. 137.
- II bis. — DANIELS H. E., J. ROY, *Stat. Soc. (Suppl.)*, **5** (1938), 89.
- III. — MARTINDALE J. G., *Journal of the Textile Institute*, **36** (1945), T. 35.
- IV. — BRENY H., *Appl. Scient. Res.* **3** (1953), 433.
- V — BRENY H., *Bull. Soc. Royale des Sc. de Liège*, 1953, 405.
- VI. — BOREL E., *Leçons sur les fonctions de variables réelles et les développements en série de polynômes*, Paris, (Gauthiers-Villars), 1905.
- VII. — KNOPP K., *Theory and Applications of infinite series*, 2 d. English edition, London (Blackie and Son), 1951.