

REVUE DE STATISTIQUE APPLIQUÉE

G. DELAITRE

Distributions poissonniennes liées

Revue de statistique appliquée, tome 11, n° 3 (1963), p. 83-106

http://www.numdam.org/item?id=RSA_1963__11_3_83_0

© Société française de statistique, 1963, 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/>

DISTRIBUTIONS POISSONNIENNES LIEES

G. DELAITRE
Ingénieur à la Cegedur

Il arrive dans certaines circonstances que deux caractères aléatoires soient distribués selon deux lois de Poisson, et que ces deux caractères aient tendance à varier simultanément.

Il en est ainsi, en particulier, lorsque les deux caractères aléatoires X et Y peuvent être considérés comme les effets de causes A et B , distribuées selon deux lois de Poisson indépendantes, la cause A entraînant certainement l'apparition du caractère X , et éventuellement celle du caractère Y avec une probabilité p , et de même, la cause B entraînant certainement l'apparition du caractère Y , et éventuellement celle du caractère X avec une probabilité p .

Bien souvent, les causes A et B ont des distributions difficilement accessibles, et la probabilité p est inconnue. La seule donnée de l'expérience est la distribution du couple X, Y .

L'étude des distributions à deux variables aléatoires correspondant à ce schéma permet d'évaluer les paramètres des distributions poissonniennes des causes et la probabilité de liaison p .

L'étude ci-dessous est exposée à partir du cas concret de certaines taches apparaissant sur des disques métalliques recuits en piles.

Un second exemple d'application concerne la distribution des commandes urgentes de deux produits, et conduit à la détermination des variations simultanées des stocks des deux produits.

Enfin le schéma indiqué peut se généraliser au cas de plus de deux variables, avec des probabilités de liaison différentes.

Les distributions de Poisson s'observent assez fréquemment dans les problèmes industriels.

On sait que ce résultat peut généralement s'interpréter par les hypothèses suivantes :

1/ Un événement E a une probabilité égale à λdS d'apparaître une fois dans un domaine élémentaire dS .

2/ La probabilité d'apparition de l'évènement E plus d'une fois dans le domaine dS est un infiniment petit d'ordre supérieur au premier.

Moyennant des conditions de continuité, on montre que la probabilité d'apparition de l'évènement E dans un domaine S suit une loi de Poisson de moyenne λS :

Assez souvent, on peut estimer que ces hypothèses sont valables et on est en droit de s'attendre alors à trouver des distributions poissonniennes. C'est notamment le cas de certains types de défauts sur des produits laminés.

Mais il peut arriver que les défauts soient causés par des "germes" engendrant un ou deux défauts : les distributions "visibles" sont celles des défauts alors que celles qu'il est intéressant de connaître sont celles des "germes".

I - EXPOSE DU PROBLEME SUR UN CAS CONCRET -

Des disques métalliques pour emboutissage sont découpés dans une bande laminée, puis recuits en piles dans un four.

Désignons par A la face supérieure de la bande en cours de laminage (qui correspondra dans notre cas à la face supérieure des disques dans les piles) et par B la face inférieure de la bande.

Dans les piles la face A d'un disque est en contact avec la face B du disque immédiatement supérieur.

Or on constate sur les faces des disques recuits un certain nombre de petites taches provenant du cracking de l'huile de laminage qui s'est amassée par capillarité autour des poussières (qu'on nous appellerons "germes") tombées sur la bande au cours du laminage.

L'expérience montre que le nombre de taches a tendance à varier simultanément sur les faces A et B en contact.

Un examen détaillé indique qu'un nombre appréciable de taches coïncident sur les faces A et B en contact.

Nous sommes ainsi conduits à admettre qu'un germe présent sur une face A ou B a une probabilité d'adhérer également à la face B ou A en contact dans la pile et de créer une tache sur cette face.

Les "germes" sont difficilement décelables sur les disques métalliques avant recuit, car ils sont de taille généralement bien inférieure à celle des taches. Seules les distributions des taches sur les faces A et B après recuit sont pratiquement accessibles, alors que ce qu'il est intéressant de connaître pour agir, ce sont les distributions des germes.

On pourrait penser utiliser les faces "libres" des disques extrémités de piles, non sujettes aux interactions, cependant diverses raisons conduisent à ne pas utiliser ces résultats (retombées possibles de suies dans le four, se comportant comme des "germes", etc.).

La première idée pour évaluer la probabilité p est d'observer simultanément les faces A et B en contact, mais cette façon de procéder conduit à une évaluation de p trop forte pour des questions d'appréciation personnelle : très souvent en effet un germe donne une tache légère et acceptable sur la face en contact ; mais les observateurs auxquels on a expliqué le mécanisme d'apparition des taches, ont tendance à compter comme défaut une tache très légère et acceptable, dès l'instant où elle coïncide avec une tache appréciable sur la face en contact.

De plus, cette méthode basée sur la fréquence des coïncidences nécessite des observations beaucoup plus longues que celles résultant d'un simple dénombrement.

Aussi, est-il préférable, une fois admis le schéma indiqué, d'étudier les distributions liées des taches sur les faces A et B en contact.

II - ETUDE THEORIQUE -

Nous admettrons que les germes sont distribués suivant des lois de Poisson de moyenne A pour les faces A et de moyenne B pour les faces B, ces deux distributions étant indépendantes.

Le schéma peut se résumer par le tableau suivant :

face A	germes existant initialement : x_1 Distribution $\mathcal{Q}(A)$	germes venant de la face B en contact : y_2	Total X
	↓	↑	
face B	germes venant de la face A en contact : y_1	germes existant initialement : x_2 Distribution $\mathcal{Q}(B)$	Y

La probabilité qu'a un germe de devenir commun aux deux faces est p . Le nombre de taches sur la face A est une variable aléatoire X, celui sur la face B une variable aléatoire Y.

$$X = x_1 + y_2$$

$$Y = x_2 + y_1$$

les variables aléatoires poissonniennes x_1 et x_2 étant indépendantes.

1 - Distributions marginales.

Le nombre de germes issus d'une face A et adhérant à la face B en contact est une variable aléatoire y_1 . Il est intuitif que y_1 est distribué suivant une loi poissonnienne de paramètre pA dans les hypothèses retenues : en effet, si nous nous rapportons à la génération des distributions poissonniennes rappelée précédemment, la probabilité d'apparition d'un germe sur une surface dS étant λdS (et l'apparition de plus d'un germe, étant d'ordre supérieur), la probabilité pour qu'un germe de dS vienne adhérer à une surface dS' en contact de dS est $p\lambda dS$ (la probabilité qu'il y en ait plus d'un est d'ordre supérieur).

S étant la surface d'un disque, et $A = \lambda S$, la variable aléatoire y_1 est donc distribuée suivant une loi de Poisson de moyenne $p\lambda S = pA$.

Ce résultat peut d'ailleurs se retrouver directement : la probabilité d'avoir sur une face B, y_1 germes venant de la face A en contact est :

$$\sum_{x_1=y_1}^{\infty} e^{-A} \frac{A^{x_1}}{x_1!} e^{-y_1} p^{y_1} (1-p)^{x_1-y_1}$$

On vérifie sans peine que cette expression est égale à $e^{-pA} \frac{(pA)^{y_1}}{y_1!}$

Le nombre de taches sur les faces B est une variable aléatoire Y

$$Y = x_2 + y_1$$

x_2 et y_1 étant des variables poissonniennes indépendantes de moyenne B et pA , leur somme est également une variable aléatoire poissonnienne de moyenne $B + pA$.

Ainsi la distribution marginale de Y nombre de taches sur les faces B est poissonnienne de paramètre $B + pA$.

La distribution marginale de X nombre de taches sur les faces A est également poissonnienne de paramètre $A + pB$.

2 - Probabilité du couple (X, Y).

La probabilité du couple (X, Y) est égale à la somme des probabilités de réalisation de x_1, y_1, x_2, y_2 tels que :

$$\begin{cases} X = x_1 + y_2 & y_1 \leq x_1 \\ Y = x_2 + y_1 & y_2 \leq x_2 \end{cases}$$

Cette probabilité est donc égale à :

$$\sum_{x_1, x_2} \frac{e^{-A} A^{x_1}}{x_1!} \cdot e^{-B} \frac{B^{x_2}}{x_2!} \cdot e^{y-x_2} e^{x-x_1} p^{x+y-(x_1+x_2)} (1-p)^{2(x_1+x_2)-(x+y)}$$

En posant, pour simplifier

$$\begin{cases} A + B = D \\ H = \frac{p}{1-p} \\ G = \frac{(1-p)^2}{p} \end{cases} \quad \text{donc} \quad \begin{cases} GH = (1-p) \\ GH^2 = p \end{cases}$$

Cette probabilité, s'écrit :

$$e^{-D} H^{x+y} \sum_{x_1, x_2} \frac{A^{x_1} B^{x_2} G^{x_1+x_2}}{(Y-x_2)! (x_1+x_2-Y)! (x-x_1)! (x_1+x_2-X)!}$$

La somme s'étendant au domaine de variation de x_1 et x_2 pour X et Y donnés.

Or d'après ce qui précède :

$$\begin{cases} 0 \leq X - x_1 \leq x_2 \\ 0 \leq Y - x_2 \leq x_1 \end{cases}$$

Donc :

$$\begin{cases} x_1 + x_2 \geq X \\ x_1 + x_2 \geq Y \end{cases}$$

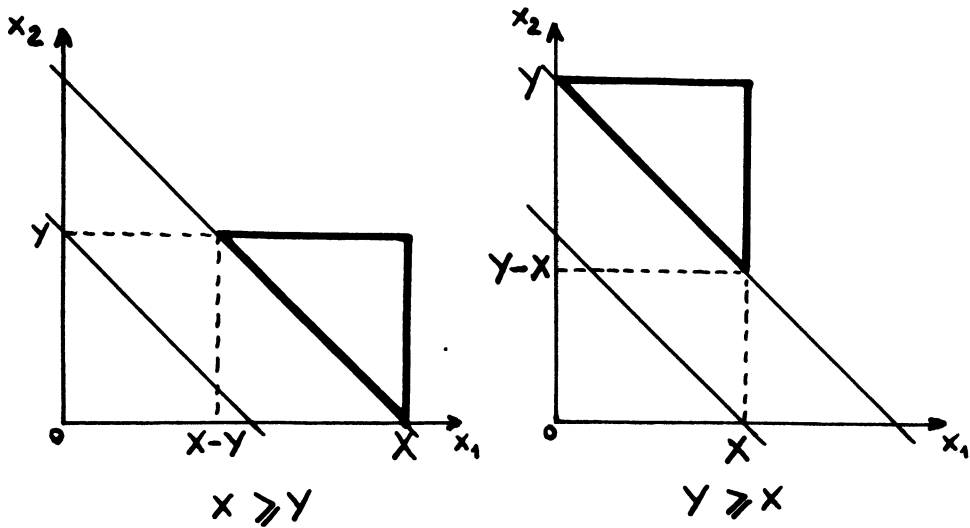


Figure 1 - Domaine de variation de x_1 et x_2 pour X et Y donnés.

a) Pour $X \geq Y$ nous voyons que x_1 peut varier dans l'intervalle $X - Y, X$

$$X - Y \leq x_1 \leq X$$

et x_2 varie de $X - x_1$ à Y

$$X - x_1 \leq x_2 \leq Y$$

b) Pour $Y > X$ nous voyons que x_1 peut varier de 0 à X :

$$0 \leq x_1 \leq X$$

et x_2 varie de $Y - x_1$ à Y :

$$Y - x_1 \leq x_2 \leq Y$$

En fait les deux cas sont symétriques et il nous suffit de calculer la probabilité du couple X, Y pour $Y \geq X$; la probabilité du couple X, Y pour $X \geq Y$ s'obtiendra en permutant X et Y et A et B dans l'expression de la probabilité précédente.

Plaçons nous dans le cas où $Y \geq X$.

La variable x_1 variant de 0 à X , nous pouvons poser :

$$x_2 = Y - x_1 + s \quad \text{avec} \quad 0 \leq s \leq x_1$$

La probabilité du couple X, Y est donc :

$$e^{-0} H^{x+y} (BG)^y \sum_{x_1=0}^x \sum_{s=0}^{x_1} \frac{A^{x_1}}{(X-x_1)!} \frac{(BG)^s}{(x_1-s)! s! (Y-X+s)!}$$

$$= e^{-0} H^{x+y} (BG)^y \sum_{x_1=0}^x \frac{A^{x_1}}{(X-x_1)!} \left[\frac{1}{x_1! 0! (Y-X)!} + \frac{BG}{(x_1-1)! 1! (Y-X+1)!} + \dots + \frac{(BG)^{x_1}}{0! x_1! (Y-X+x_1)!} \right]$$

La sommation par rapport à x_1 introduit des expressions proportionnelles au développement de $(1 + \frac{A}{B})^m = (\frac{D}{B})^m$, si bien que nous pouvons écrire :

$$f(X, Y) = e^{-0} H^{x+y} (BG)^y \left[\frac{1}{0! (Y-X)! X!} \left(\frac{D}{B}\right)^X + \frac{AG}{1! (Y-X+1)! (X-1)!} \left(\frac{D}{B}\right)^{X-1} + \dots + \frac{(AG)^X}{X! Y! 0!} \right]$$

pour $Y > X$

Expression que nous pouvons mettre sous la forme :

$$f(X, Y) = e^{-0} H^X (BGH)^Y B^{-X} \left[\frac{(AG)^X}{X! 0! Y!} + \frac{(AG)^{X-1} D}{(X-1)! 1! (Y-1)!} + \frac{(AG)^{X-2} D^2}{(X-2)! 2! (Y-2)!} + \dots + \frac{(AG)^0 D^X}{0! X! (Y-X)!} \right] \quad (1)$$

pour $Y \geq X$

Dans le cas où $X \geq Y$, il suffit comme nous l'avons indiqué de permuter X et Y et A et B dans la formule (1) :

$$f(X, Y) = e^{-0} H^Y (AGH)^X A^{-Y} \left[\frac{(BG)^Y}{Y! 0! X!} + \frac{(BG)^{Y-1} D}{(Y-1)! 1! (X-1)!} + \dots + \frac{(BG)^0 D^Y}{0! Y! (X-Y)!} \right] \quad (2)$$

pour $X \geq Y$

On vérifie que si $p = 0$ (soit $GH = 1$ et $GH^2 = 0$) la probabilité du couple (X, Y) se réduit à :

$$e^{-0} \frac{A^X B^Y}{X! Y!}$$

ce qui correspond à la fréquence du couple de 2 variables poissonniennes indépendantes X et Y .

Par contre si $p = 1$ la probabilité est nulle pour tout couple $(X, Y \neq X)$ et est égale à $e^{-0} \frac{D^X}{X!}$ pour le couple $(X, Y = X)$.

La figure n° 2 donne l'aspect de la distribution du couple X, Y dans un cas particulier.

3 - Fonctions caractéristiques de la distribution à deux variables aléatoires X, Y .

La première fonction caractéristique de la distribution est :

$$\varphi(u, v) = \sum_{x=0}^{\infty} \sum_{y=0}^{\infty} f(X, Y) U^x V^y$$

en posant $U = e^{iu}$ et $V = e^{iv}$.

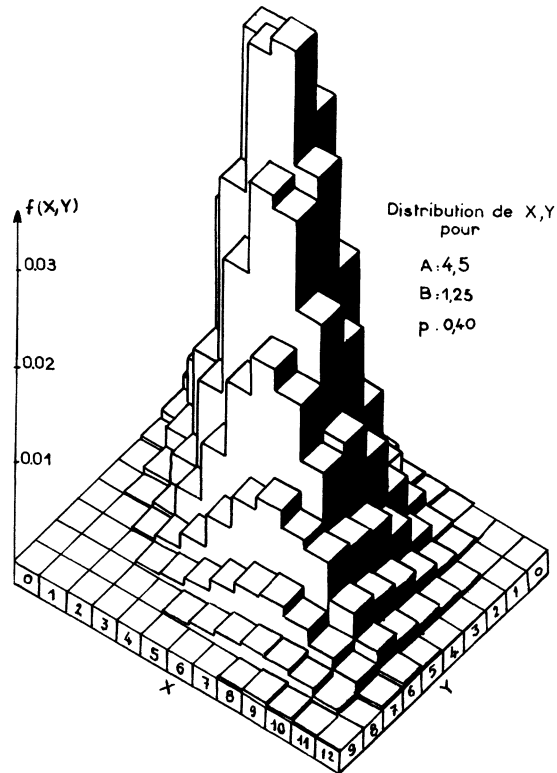


Figure 2 -

On peut voir que pour X donné la sommation par rapport à Y conduit, d'après les formules (1) et (2) à :

$$e^{-D} H^X U^X e^{BGHV} \frac{(AG + DGHV)^X}{X!}$$

La sommation par rapport à X donne alors :

$$e^{-D} e^{BGHV} e^{AGHU + DGH^2 UV}$$

En remplaçant GH par $(1 - p)$ et GH^2 par p , nous obtenons :

$$\varphi(u, v) = \exp \{ (1 - p)(AU + BV) + (A + B)(pUV - 1) \} \quad (3)$$

La seconde fonction caractéristique de la distribution de X, Y est :

$$\psi(u, v) = (1 - p)(AU + BV) + (A + B)(pUV - 1) \quad (4)$$

On vérifie que pour $u = v = 0$, soit $U = V = 1$

$$\varphi(0, 0) = 1$$

4 - Fonction caractéristique de la distribution de Y pour X donné.

D'après les résultats précédents, la première fonction caractéristique de la distribution de Y pour X donné est :

$$\varphi_Y(v) = \frac{e^{-D} H^X e^{BGHV} \frac{(AG + DGHV)^X}{X!}}{e^{-D} H^X e^{BGH} \frac{(AG + DGH)^X}{X!}}$$

Soit :

$$\varphi_Y(v) = e^{B(1-p)(v-1)} \left(\frac{A(1-p) + (A+B)pV}{A+Bp} \right)^X \quad (5)$$

La seconde fonction caractéristique de Y pour X donné est donc :

$$\psi_Y(v) = B(1-p)(V-1) + X \log \frac{A(1-p) + (A+B)pV}{A+Bp} \quad (6)$$

De même nous obtiendrions $\varphi_X(u)$ et $\psi_X(u)$ en changeant X en Y et A en B.

5 - Régression de Y en fonction de X.

La formule (6), où $V = e^{iv}$ nous montre que les cumulants de la distribution de Y sont des fonctions linéaires de X.

$$K_r = \alpha_r + \beta_r X$$

En particulier la moyenne de Y en fonction de X est :

$$K_1 = B(1-p) + \frac{(A+B)pX}{A+Bp} \quad (7)$$

La régression de Y en fonction de X est donc linéaire.

(On vérifiera sans peine que $\frac{(A+B)p}{A+Bp}$, coefficient de X dans l'expression de K_1 est égal à :

$$\rho \frac{\sigma_Y}{\sigma_X} = \frac{\mu_{11}}{(\sigma_X)^2}$$

ρ étant le coefficient de corrélation de X, Y et μ_{11} la covariance de X, Y).

La variance de Y en fonction de X est donnée par :

$$K_2 = B(1-p) + \frac{p(1-p)(A+B)A}{(A+Bp)^2} X \quad (8)$$

Pour $X = 0$ la distribution de Y a pour seconde fonction caractéristique :

$$\psi_0(v) = B(1-p)(V-1),$$

fonction caractéristique d'une loi de Poisson de moyenne $B(1 - p)$; ce résultat s'interprète aisément.

Des résultats analogues sont évidemment obtenus pour la distribution de X en fonction de Y .

6 - Normalité asymptotique de la distribution de Y .

Signalons, bien que ce résultat ait peu d'intérêt pratique, que lorsque X augmente indéfiniment, la distribution de Y pour X donné tend vers une loi normale.

Introduisons en effet la variable réduite :

$$y = \frac{Y - \left[B(1 - p) + \frac{(A + B)p}{A + Bp} X \right]}{\sqrt{\frac{B(1 - p) + \frac{A(A + B)p(1 - p)}{(A + Bp)^2}}{X}}}$$

Cette transformation classique nous donne la seconde fonction caractéristique $\psi'_x(v)$ de y .

Il est alors facile de voir, par un développement limité que $\psi'_x(v)$ tend vers la valeur $-\frac{v^2}{2}$ lorsque X augmente indéfiniment :

$$\psi'_{\infty}(v) = -\frac{v^2}{2}$$

la variable réduite y a une distribution qui tend donc vers la normale lorsque X augmente indéfiniment.

On trouverait bien entendu un résultat analogue pour la distribution de X lorsque Y croît indéfiniment.

7 - Forme limite de la distribution du couple (X, Y) .

La seconde fonction caractéristique du couple (X, Y) est donnée par la formule (4).

Les distributions marginales de X et Y étant poissonniennes de paramètres respectivement $A + pB$ et $B + pA$, introduisons les variables réduites

$$\xi = \frac{X - (A + pB)}{\sqrt{A + pB}}$$

$$\zeta = \frac{Y - (B + pA)}{\sqrt{B + pA}}$$

La seconde fonction caractéristique $\psi'(u, v)$ du couple ξ, ζ , est :

$$\psi'(u, v) = -i\sqrt{A + pB}u - i\sqrt{B + pA}v + (1 - p) [Ae^{iu'} + Be^{iv'}] + A + B [pe^{i(u'+v')} - 1]$$

$$\left\{ \begin{array}{l} u' = \frac{u}{\sqrt{A + pB}} \\ v' = \frac{v}{\sqrt{B + pA}} \end{array} \right.$$

Il est alors facile de voir par des développements limités, que, si ρ étant différent de zéro, A ou B augmentent indéfiniment, $\psi'(u, v)$ tend vers :

$$\psi(u, v) = -\frac{1}{2} \left[u^2 + 2 \frac{\rho(A+B)}{\sqrt{(A+\rho B)(B+\rho A)}} uv + v^2 \right]$$

Seconde fonction caractéristique d'une loi normale à deux variables réduites ayant pour coefficient de corrélation :

$$\rho = \frac{\rho(A+B)}{\sqrt{(A+\rho B)(B+\rho A)}}$$

Or, A et B représentent, dans les hypothèses faites, les nombres moyens de germes par face de disque. On peut donc augmenter simultanément A et B en groupant les disques n à n . Pour un groupement de n disques, les nombres de germes pour n faces seront distribués suivant des lois de Poisson de moyennes nA et nB .

En prenant une valeur de n assez élevée, à apprécier dans chaque cas particulier, nous pouvons donc admettre que le nombre de taches sur chacune des deux faces d'un groupe de n disques est distribué sensiblement suivant une loi normale à deux variables lorsque n est assez élevé.

Il est clair que nous devons trouver le même résultat si nous considérons les moyennes \bar{X} et \bar{Y} sur n disques, lorsque n augmente indéfiniment.

La fonction caractéristique des variables réduites :

$$\left\{ \begin{array}{l} \frac{\bar{X} - (A + \rho B)}{\sqrt{\frac{A + \rho B}{n}}} \\ \frac{\bar{Y} - (B + \rho A)}{\sqrt{\frac{B + \rho A}{n}}} \end{array} \right.$$

tend vers

$$\psi(u, v) = -\frac{1}{2} \left[u^2 + 2 \frac{\rho(A+B)}{\sqrt{(A+\rho B)(B+\rho A)}} uv + v^2 \right]$$

Le coefficient de corrélation ρ de \bar{X} , \bar{Y} étant nul en même temps que ρ , on peut tester l'hypothèse $\rho = 0$ en utilisant l'approximation normale en groupant les disques par n unités.

8 - Moments de la distribution de X, Y.

Les moments de la distribution du couple X, Y peuvent s'obtenir par le développement de la fonction caractéristique $\psi(u, v)$.

Les premiers moments peuvent d'ailleurs se déterminer directement sans difficulté, en particulier la covariance μ_{11} :

$$\mu_{11} = m_{11} - m_{10} \cdot m_{01}$$

Or :

$$m_{11} = E(X Y) = E [(x_1 + y_2) (x_2 + y_1)]$$

En observant que x_1, x_2 d'une part, y_1, y_2 d'autre part, sont indépendants entre eux, nous pouvons écrire :

$$m_{11} = E(x_1 y_1) + E(x_2 y_2) + E(x_1) E(x_2) + E(y_1) E(y_2)$$

Or, on peut voir simplement que :

$$E(x_1 y_1) = p(A + A^2) \text{ et } E(x_2 y_2) = p(B + B^2)$$

si bien que :

$$\mu_{11} = p(A + B)$$

Ainsi :

$$\left\{ \begin{array}{l} m_{10} = A + Bp \\ m_{01} = B + Ap \\ \mu_{11} = p(A + B) \end{array} \right.$$

Nous déterminerons A, B, p en ajustant les moments précédents aux moments empiriques correspondants.

III - APPLICATIONS -

1/ L'examen de 600 disques métalliques recuits par piles de 6 a permis d'établir le tableau de résultats suivants sur les 500 faces A et B en contact :

Tableau n° 1

Nombre de taches par face : X et Y

Faces A X \ Faces B Y	Faces B					Total
	0	1	2	3	4	
0	186	24	2			212
1	38	120	14	1		173
2	4	26	49	4	1	84
3			9	14	2	25
4				2	4	6
Total	228	170	74	21	7	500

L'ajustement des moments m_{10} , m_{01} , et μ_{11} conduit aux estimations suivantes :

$$\left\{ \begin{array}{l} A = 0,6085 \\ B = 0,3633 \\ p = 0,7472 \end{array} \right.$$

Les fréquences correspondant aux paramètres ci-dessus ont été calculées par les formules (1) et (2), ainsi que l'effectif théorique correspondant et celui des distributions marginales.

Nous avons pu dresser le tableau suivant.

Tableau n° 2

Effectif théorique

Faces A X \ Faces B Y	0	1	2	3	4 et plus			
0	189,19 186	17,37 24	6,32 10					
1	29,10 38	140,05 120				12,73 14		
2						21,34 26	51,82 49	4,67 4
3								7,82 9
4 et plus	6,81 6							

L'effectif observé correspondant a été rappelé sous l'effectif calculé. Le test de χ^2 appliqué à ce tableau donne :

$$\chi^2 = 12,05 \text{ pour } 8 \text{ degrés de liberté.}$$

Comme $\chi_{0,10}^2(8) = 13,36$, nous pouvons admettre que l'ajustement est satisfaisant, la valeur trouvée ayant une probabilité d'être dépassée supérieure à 0,10 dans l'hypothèse nulle.

2/ Un atelier de tubes sanitaires produit deux sections de tubes très courantes : tubes 12×14 (diamètres intérieur et extérieur en mm) et tubes 10×12 que nous appellerons respectivement tubes A et B.

Le délai normal de fabrication de ces tubes est de 4 semaines, et les commandes dont le délai de livraison est inférieur à cette valeur sont prélevées sur stock.

Sur une période d'observation de 188 jours ouvrables, les commandes journalières à prélever sur stock se répartissent suivant le tableau ci-après.

Tableau n° 3

Tubes B

Tubes A	Nbre de commandes reçues par jour	0	1	2	3	Total
	0	105	16	1		122
1	19	34	3	1	57	
2	1	4	1		6	
3				3	3	
Total	125	54	5	4	188	

Or, il arrive fréquemment que des clients passent des commandes simultanées de tubes A et B. Et on peut penser que certaines commandes "pilotes" de tubes A ou B ont une probabilité p d'entraîner une commande de tubes B ou A.

Nous connaissons les distributions des commandes résultantes des tubes A et B. Dans le cas du tableau n° 3, on peut voir que les distributions marginales des commandes de tubes A et B sont ajustables à des distributions de Poisson.

Admettons que les tubes A comportent certaines commandes "pilotes" distribuées suivant une loi de Poisson de moyenne A , et que les tubes B comportent également certaines commandes "pilotes" distribuées suivant une loi de Poisson de moyenne B , les deux distributions étant indépendantes, mais telles qu'une commande "pilote" de A ou B ait une probabilité p d'entraîner une commande de B ou A.

Nous obtenons un schéma analogue à celui de l'interaction des faces A et B en contact.

Dans le cas du tableau n° 3, l'ajustement des moments m_{10} , m_{01} et μ_{11} conduit aux estimations suivantes :

$$\left\{ \begin{array}{l} A = 0,2858 \\ B = 0,2650 \\ p = 0,4874 \end{array} \right.$$

estimations qui permettent d'établir le tableau n° 4 donnant l'effectif théorique correspondant.

Tableau n° 4

Tubes B

		Tubes B			
		0	1	2	3 et plus
Tubes A	Nbre de commandes reçues par jour				
	0	108,38	14,72	1	
	1	15,88	31,30	4,10	1,55
	2	1,16	4,43	4,51	
	3 et plus	0,07	0,27	0,64	

Les résultats de ce tableau sont très voisins du tableau n° 3, et en appliquant le test de χ^2 aux divers groupes de cases encadrés, nous obtenons $\chi^2 = 1,87$, résultat très satisfaisant pour 3 degrés de liberté (7 groupes, 3 paramètres déduits de la distribution empirique).

Ainsi nous sommes en droit d'admettre que tout se passe comme si chaque ensemble de tubes A et B comportaient des commandes "pilotes" distribuées selon 2 lois de Poisson indépendantes, une commande "pilote" d'un produit ayant une probabilité p d'entraîner une commande de l'autre produit. La fonction caractéristique du nombre de commandes reçues dans un temps donné sera de la forme (3).

Ce résultat peut être utilisé dans l'étude des variations simultanées des stocks de deux produits.

En effet on peut en général considérer les variations de stock d'un produit comme résultant d'un nombre aléatoire de prélèvements, d'importance aléatoire⁽¹⁾.

On sait que si le nombre de prélèvements dans un temps donné est distribué suivant une loi de Poisson de moyenne λ , et si l'importance des prélèvements individuels est caractérisée par leur poids x , de fonction caractéristique $\varphi_x(t)$, en admettant l'indépendance de x et de la fréquence des prélèvements, la seconde fonction caractéristique du tonnage total prélevé X pendant le temps considéré est :

$$\psi_x(t) = \lambda [\varphi_x(t) - 1]$$

Dans le cas de 2 produits A et B, dont le nombre de prélèvements sur stock, dans un temps donné, résulte du mécanisme ci-dessus, nous admettrons également que le poids du prélèvement d'un produit est indépendant de la fréquence des prélèvements de ce produit.

Par contre, on constate généralement que pour les commandes portant à la fois sur des produits A et B, les poids x et y commandés ne sont pas indépendants. Admettons que l'étude des commandes portant

 (1) Voir par exemple l'article : Somme d'un nombre aléatoire de termes aléatoires. Application aux problèmes de stockage par B. Roy. (Revue de Statistique Appliquée, 1960, Vol. VIII, n° 1).

à la fois sur les produits A et B, permette de déterminer la fonction de distribution $G(x, y)$ des poids x de A et y de B, dont la première fonction caractéristique est $\varphi_{x,y}(u, v)$ (nous admettons implicitement que la fonction de distribution des poids x des commandes portant uniquement sur le produit A est $G(x, \infty)$, et que de même, la fonction de distribution des poids y des commandes portant uniquement sur le produit B est $G(\infty, y)$, hypothèses à vérifier dans chaque cas particulier).

Soit $\psi_{x,y}(u, v)$ la seconde fonction caractéristique des tonnages totaux X et Y des produits A et B à prélever sur stock dans un temps donné.

Si nous nous intéressons uniquement à la distribution marginale de X, dans les hypothèses faites, sa seconde fonction caractéristique sera, d'après ce qui précède :

$$\psi_{x,y}(u, 0) = (A + pB) [\varphi_{x,y}(u, 0) - 1] \quad (9)$$

De même pour la distribution marginale de Y :

$$\psi_{x,y}(0, v) = (B + pA) [\varphi_{x,y}(0, v) - 1] \quad (10)$$

Soit alors la fonction $\psi(u, v)$ définie comme suit :

$$\psi(u, v) = (1 - p) [A \varphi_{x,y}(u, 0) + B \varphi_{x,y}(0, v)] + (A + B) [p\varphi_{x,y}(u, v) - 1] \quad (11)$$

Posons :

$$\psi_{x,y}(u, v) = \psi(u, v) + \Theta(u, v) \quad (12)$$

En faisant $v = 0$ dans l'équation (12), nous devons trouver une expression en v identique à (9) quel que soit v .

En remarquant que $\varphi_{x,y}(0, 0) = 1$, nous trouvons :

$$\Theta(u, 0) = 0$$

De même :

$$\Theta(0, v) = 0$$

Donc :

$$\Theta(u, v) = 0$$

et la seconde fonction caractéristique de X, Y tonnages totaux à prélever sur stock en un temps donné est :

$$\psi_{x,y}(u, v) = \psi(u, v)$$

Cette formule permet d'avoir les cumulants et les moments successifs des poids totaux à prélever sur stocks de deux produits A et B liés à la fois par la fréquence de leurs prélèvements (suivant le schéma indiqué) et par l'importance des prélèvements.

En particulier, d'après (12) la covariance de X, Y est égale à :

$$\text{Cov} [X, Y] = (A + B)p E(x y)$$

E(x y) étant l'espérance mathématique de x y et

$$\text{Var} [X] = (A + Bp) E [x^2]$$

$$\text{Var} [Y] = (B + pA) E [y^2]$$

Si bien que le coefficient de corrélation entre X et Y est :

$$\frac{(A + B)p}{\sqrt{(A + Bp)(B + pA)}} \frac{E(x y)}{\sqrt{E [x^2] E [y^2]}}$$

le premier terme est le coefficient de corrélation entre les nombres des prélèvements sur stock.

Dans le cas particulier étudié ci-dessus, les poids de commandes x et y sont distribués selon deux lois lognormales, log x et log y étant distribués suivant une loi normale à deux variables de moyennes m_1 et m_2 , d'écart types σ_1 et σ_2 et de coefficient de corrélation ρ .

Dans ces conditions le calcul donne :

$$E(x y) = \exp \left\{ m_1 + m_2 + \frac{\sigma_1^2}{2} + \frac{\sigma_2^2}{2} + \rho \sigma_1 \sigma_2 \right\}$$

$$E [x^2] = e^{2m_1 + 2\sigma_1^2}$$

$$E [y^2] = e^{2m_2 + 2\sigma_2^2}$$

Le coefficient de corrélation entre X et Y est alors :

$$\frac{(A + B)p}{\sqrt{(A + Bp)(B + pA)}} e^{-\frac{1}{2}(\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2)}$$

Dans le cas des tubes sanitaires, l'expérience a montré que les logarithmes décimaux des poids x et y avaient le même écart-type :

$$\sigma_1' = \sigma_2' = 0,5069$$

et que le coefficient de corrélation entre ces logarithmes était :

$$\rho = 0,822$$

Le coefficient de corrélation entre les tonnages X et Y prélevés sur stock en 4 semaines est alors égal à :

$$0,6555 \times 0,9553 = 0,6262$$

IV - CONCLUSIONS -

Les deux exemples indiqués, pris dans des domaines très différents, permettent de tirer plusieurs conclusions pratiques :

1/ le premier exemple portant sur certaines taches apparaissant au recuit en piles de disques métalliques montre qu'un examen rapide du problème aurait conduit à admettre que les résultats sur les deux faces A et B étaient très comparables, alors qu'en réalité, les causes des taches sont nettement plus fréquentes sur la face A (correspondant à la face supérieure de la bande en cours de laminage).

Cette différence entre faces est masquée par la probabilité d'interaction entre faces au recuit, probabilité estimée à 0,747, donc très appréciable.

Ce résultat a permis de localiser le problème et de prendre des mesures pour diminuer le risque de pollution de la bande en cours de laminage.

2/ Le second exemple portant sur les variations de stocks de deux produits montre que dans les commandes (où les prélèvements) sur stock des deux produits, tout se passe comme si, pour chaque produit, il existait des commandes "pilotes" entraînant une commande de l'autre produit avec une probabilité p . Il n'est pas nécessaire de définir ce qu'est une commande "pilote" pour un produit, mais seulement de constater que le schéma adopté conduit à une distribution théorique s'ajustant bien avec la distribution observée.

Lorsque les distributions réelles à deux variables aléatoires peuvent entrer dans le schéma donné, on peut étudier, d'une façon relativement simple, les variations simultanées des stocks des deux produits, en admettant que ces variations de stocks résultent de prélèvements aléatoires (liés suivant le processus envisagé) de tonnages aléatoires (liés pour les deux produits), prélèvements et tonnages étant indépendants.

V - GENERALISATION -

Nous avons considéré dans cette étude que la probabilité d'interaction des "germes" était la même dans le sens face A \longrightarrow face B que dans le sens face B \longrightarrow face A.

Dans certains cas, il faut envisager des probabilités d'interaction différentes.

Changeons les notations pour rendre les résultats plus "symétriques" :

1/ Soit une première série de germes A_1 distribués suivant une loi de Poisson de paramètre A_1 , et une seconde série de germes A_2 indépendante de la première et distribuée suivant une loi de Poisson de paramètre A_2 .

Les probabilités d'interaction seront p_{12} , p_{21} telles que :

$$\begin{array}{ccc} \text{Série } A_1 & \xrightarrow{p_{12}} & \text{Série } A_2 \\ \text{Evènement résultant } X_1 & \xleftarrow{p_{21}} & \text{Evènement résultant } X_2 \end{array}$$

Il est clair que les distributions de chacun des évènements résultants X_1 et X_2 seront poissonniennes, de paramètre

$$A_1 + p_{21} A_2 \text{ et } p_{12} A_1 + A_2$$

L'expression de la probabilité du couple (X_1, X_2) sera plus compliquée que les formules (1) et (2) ; par contre la seconde fonction caractéristique du couple (X_1, X_2) , soit $\psi(t_1, t_2)$ est facile à obtenir, sans passer par l'expression de la probabilité du couple (X_1, X_2) .

En effet, les distributions marginales de X_1 et X_2 sont poissonniennes. En posant $e^{it_1} = U_1$ et $e^{it_2} = U_2$, la seconde fonction caractéristique $\psi(U_1, U_2)$ est telle que :

$$\psi(U_1, 1) = (A_1 + p_{21} A_2) (U_1 - 1)$$

(seconde fonction caractéristique de la distribution marginale de X_1).

De même, pour la distribution marginale de X_2 :

$$\psi(1, U_2) = (p_{12} A_1 + A_2) (U_2 - 1)$$

Il en résulte que la seconde fonction caractéristique $\psi(U_1, U_2)$ est nécessairement linéaire par rapport à U_1 et par rapport à U_2 ,

$$\psi(U_1, U_2) = a_1 U_1 + a_2 U_2 + b U_1 U_2 + c$$

De plus :

$$\psi(1, 1) = 0,$$

donc :

$$a_1 + a_2 + b + c = 0$$

Le coefficient b est égal au cumulante K_{11} :

$$b = K_{11} = \mu_{11}$$

En remarquant que :

$$X_1 = x_1 + y_2$$

$$X_2 = x_2 + y_1$$

et que x_1 et x_2 d'une part et y_1, y_2 d'autre part, sont indépendants, un calcul simple donne :

$$b = p_{12} A_1 + p_{21} A_2$$

La distribution marginale de X_1 donne $a_1 + b = m_{10} = A_1 + p_{21} A_2$

d'où :

$$a_1 = q_{12} A_1 \quad \text{avec} \quad q_{12} = 1 - p_{12}$$

de même :

$$a_2 = q_{21} A_2$$

Finalement :

$$\psi(U_1, U_2) = q_{12} A_1 U_1 + q_{21} A_2 U_2 + (p_{12} A_1 + p_{21} A_2) U_1 U_2 - (A_1 + A_2) \quad (13)$$

2/ Considérons une usine produisant 3 produits A_1, A_2, A_3 , les commandes urgentes étant livrées sur stock.

On constate qu'un bon nombre de commandes comportent à la fois 2 ou 3 des produits.

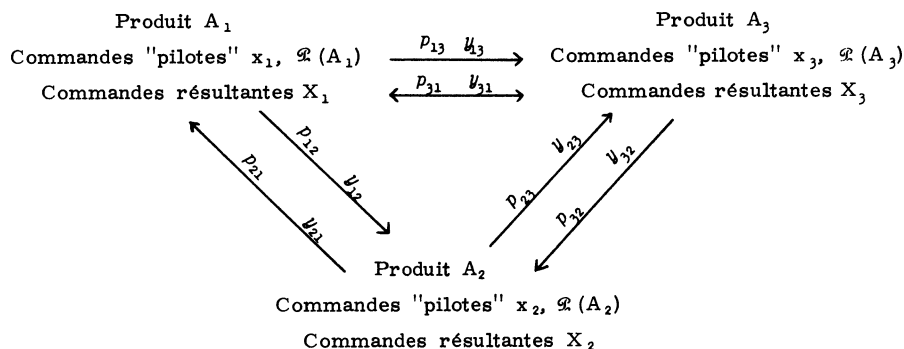
On peut considérer que certaines commandes d'un produit A_i ont une probabilité p_{ij} d'entraîner également une commande d'un produit A_j .

On peut admettre que chaque produit comporte des commandes "pilotes" entraînant la commande d'un autre produit.

Or la distribution du nombre de commandes urgentes reçues dans un délai donné pour un produit A_i s'ajuste en général assez bien à une distribution de Poisson (tout au moins si la période d'observation correspond à une activité économique constante).

On peut alors admettre que le produit A_i comporte des commandes "pilotes" urgentes dont le nombre dans un temps donné suit une loi de Poisson $\mathcal{P}(A_i)$ de paramètre A_i , les commandes "pilotes" des produits A_i, A_j étant indépendantes.

Nous aurons ainsi le schéma suivant :



Les distributions $\mathcal{P}(A_1), \mathcal{P}(A_2), \mathcal{P}(A_3)$ étant indépendantes.

Les commandes "résultantes" sont telles que :

$$\begin{cases} X_1 = y_1 + y_{21} + y_{31} \\ X_2 = y_{12} + x_2 + y_{32} \\ X_3 = y_{13} + y_{23} + x_3 \end{cases} \quad (14)$$

D'après tout ce qui précède, il est évident que la distribution de X_i nombre de commandes résultantes du produit A_i est poissonnienne.

Si

$$\|P\| = \begin{vmatrix} 1 & p_{12} & p_{13} \\ p_{21} & 1 & p_{23} \\ p_{31} & p_{32} & 1 \end{vmatrix}$$

est la matrice des probabilités d'interaction des commandes "pilotes"

$$\|A\| = \begin{vmatrix} A_1 \\ A_2 \\ A_3 \end{vmatrix}$$

la matrice des paramètres des distributions indépendantes des commandes "pilotes" et

$$\|m\| = \begin{vmatrix} m_{100} \\ m_{010} \\ m_{001} \end{vmatrix}$$

la matrice des paramètres des distributions poissonniennes des commandes résultantes.

Nous aurons :

$$\|m\| = \|P\| \|A\| \quad (15)$$

Posons :

$$e^{i t_k x_k} = U_k \quad (K \text{ de } 1 \text{ à } 3)$$

La seconde fonction caractéristique du point (X_1, X_2, X_3) sera

$$\psi(U_1, U_2, U_3).$$

La distribution marginale de X_i étant poissonnienne, sa seconde fonction caractéristique, obtenue en faisant $U_j = 1$ (pour $j \neq i$) dans $\psi(U_1, U_2, U_3)$ sera linéaire en U_i .

Il en résulte que $\psi(U_1, U_2, U_3)$ est linéaire par rapport à chacun des U_i , $\psi(U_1, U_2, U_3)$ est donc nécessairement de forme :

$$\psi(U_1, U_2, U_3) = a_1 U_1 + a_2 U_2 + a_3 U_3 + b_1 U_2 U_3 + b_2 U_1 U_3 + b_3 U_1 U_2 + c U_1 U_2 U_3 + d$$

avec évidemment :

$$\psi(1, 1, 1) = a_1 + a_2 + a_3 + b_1 + b_2 + b_3 + c + d = 0 \quad (16)$$

Les 8 coefficients inconnus sont déterminés, en remarquant que :

$$c = K_{111} = \mu_{111}$$

et que les distributions marginales des couples (X_1, X_2) , (X_2, X_3) et (X_1, X_3) donnent :

$$\begin{cases} b_1 + c = K_{011} = \mu_{011} \\ b_2 + c = K_{101} = \mu_{101} \\ b_3 + c = K_{110} = \mu_{110} \end{cases}$$

et que les distributions marginales de X_1, X_2, X_3 donnent :

$$\begin{cases} a_1 + b_2 + b_3 + c = m_{100} \\ a_2 + b_1 + b_3 + c = m_{010} \\ a_3 + b_1 + b_2 + c = m_{001} \end{cases}$$

Le coefficient d est alors déterminé par l'équation (16).

Le calcul des différents moments centrés ne présente pas de difficulté en partant des expressions (14) et en observant que les variables n'ayant pas le même premier indice sont indépendantes.

On trouve :

$$\begin{cases} c = p_{12} p_{13} A_1 + p_{21} p_{23} A_2 + p_{31} p_{32} A_3 \\ b_1 = p_{23} q_{21} A_2 + p_{32} q_{31} A_3 \\ b_2 = p_{13} q_{12} A_1 + p_{31} q_{32} A_3 \quad (\text{avec } q_{ij} = 1 - p_{ij}) \\ b_3 = p_{12} q_{13} A_1 + p_{21} q_{23} A_2 \\ a_1 = q_{12} q_{13} A_1 \\ a_2 = q_{21} q_{23} A_2 \\ a_3 = q_{31} q_{32} A_3 \\ d = - (A_1 + A_2 + A_3) \end{cases} \quad (17)$$

On vérifie aisément que si on annule les coefficients p et A contenant l'indice i , on retombe sur une distribution à deux variables dont la seconde fonction caractéristique est donnée par (13).

Admettons, comme dans le second exemple, que pour les commandes portant à la fois sur les 3 produits, les poids commandés z_1, z_2, z_3 aient une fonction de distribution $G(z_1, z_2, z_3)$ de fonction caractéristique $\varphi(t_1, t_2, t_3)$ (en admettant implicitement que les poids z_1 des commandes portant uniquement sur le produit A_1 ont une fonction de distribution $G(z_1, \infty, \infty)$, et que les poids z_1 et z_2 des commandes portant uniquement sur les produits A_1 et A_2 ont une fonction de distribution $g(z_1, z_2, \infty)$, et que les hypothèses similaires soient remplies pour les autres combinaisons de produits).

Soit $\psi_{z_1, z_2, z_3}(t_1, t_2, t_3)$ la seconde fonction caractéristique de la distribution des tonnages totaux Z_1, Z_2, Z_3 des produits A_1, A_2, A_3 à prélever sur stock dans un temps donné.

D'après ce qui précède la seconde fonction caractéristique de la distribution marginale de Z_1 sera :

$$\psi_{z_1, z_2, z_3}(t_1, 0, 0) = (A_1 + p_{21} A_2 + p_{31} A_3) \varphi(t_1, 0, 0) - 1 \quad (18)$$

On obtiendrait des expressions analogues pour les secondes fonctions caractéristiques des distributions marginales de Z_2 et Z_3 .

Soit alors la fonction $\psi(t_1, t_2, t_3)$ définie comme suit :

$$\begin{aligned} \psi(t_1, t_2, t_3) = & a_1 \varphi(t_1, 0, 0) + a_2 \varphi(0, t_2, 0) + a_3 \varphi(0, 0, t_3) \\ & + b_1 \varphi(0, t_2, t_3) + b_2 \varphi(t_1, 0, t_3) + \varphi(t_1, t_2, 0) \\ & + c \varphi(t_1, t_2, t_3) + d \end{aligned} \quad (19)$$

où les coefficients sont ceux définis par les équations (17).

Nous pouvons poser :

$$\psi_{z_1, z_2, z_3}(t_1, t_2, t_3) = \psi(t_1, t_2, t_3) + \Theta(t_1, t_2, t_3)$$

En identifiant alors $\psi_{z_1, z_2, z_3}(t_1, 0, 0)$ à l'expression (18) et en opérant de même pour les variables Z_2 et Z_3 , nous trouverions que

$$\Theta(t_1, t_2, t_3) \equiv 0$$

Par conséquent :

$$\psi_{z_1, z_2, z_3}(t_1, t_2, t_3) = \psi(t_1, t_2, t_3)$$

VI - REMARQUES -

Les généralisations précédentes ne sont utilisables pratiquement, que si l'on connaît a priori les différents paramètres A_i et p_{ij} . En effet, si nous généralisons le processus au cas de n variables, il faut introduire n paramètres A_i et $2 C_n^2$ paramètres p_{ij} , donc au total n^2 paramètres.

Or, une seconde fonction caractéristique de n variables aléatoires X_1, \dots, X_n linéaire par rapport à chacun des U_i dépend, dans le cas général, de :

$$C_n^1 + C_n^2 + C_n^3 + \dots + C_n^n = 2^n - 1 \text{ coefficients indépendants.}$$

Ceci étant, abstraction faite de la complexité très rapidement croissante des calculs lorsque n augmente, pour $n \leq 4$ le nombre de paramètres A_i et p_{ij} est supérieur au nombre de coefficients indépendants entrant dans la fonction de distribution de X_1, \dots, X_n . Ainsi si des valeurs données de A_i et p_{ij} déterminent des coefficients de seconde fonction caractéristique uniques, inversement la connaissance de ces coefficients ne permet pas de déterminer en général une solution unique pour les A_i et p_{ij} .

Pour $n \leq 4$, la distribution de X_1, \dots, X_n ne permet d'avoir une solution unique (correspondant bien entendu à $A_i > 0$ et $0 < p_{ij} < 1$) que si on ramène le nombre de paramètres indépendants à $2^n - 1$ au moyen d'hypothèses supplémentaires.

C'est ce que nous avons fait implicitement au début de cet exposé, en admettant que la probabilité d'interaction p est la même des faces A aux faces B et des faces B aux faces A : c'est alors une hypothèse logique entraînée par des conditions de symétrie évidentes, et correspondant à une réalité physique.

Il n'y a plus alors que 3 paramètres indépendants à déterminer pour une distribution dans laquelle entrent 3 coefficients indépendants : la solution trouvée, lorsqu'elle convient est unique.

Par contre, dans le cas des commandes sur stock de deux produits, les probabilités d'interaction p_{12} et p_{21} n'ont aucune raison d'être égales ; nous avons alors 4 paramètres à déterminer d'après la connaissance de 3 coefficients.

L'introduction des commandes "pilotes" fictives est alors un simple artifice permettant de décrire aisément la distribution des deux caractères. Il existe alors une infinité de systèmes de commandes "pilotes". Par contre les distributions poissonniennes de ces commandes "pilotes" sont déterminées d'une façon unique (si elles existent) dès qu'on impose à ces commandes une condition supplémentaire, par exemple d'avoir des probabilités d'interaction égales $p_{12} = p_{21} = p$.

Dans le cas de deux variables aléatoires X_1, X_2 dont les distributions marginales sont poissonniennes, et dont la seconde fonction caractéristique a par conséquent la forme :

$$\psi(U_1, U_2) = a_1 U_1 + a_2 U_2 + b U_1 U_2 - (a_1 + a_2 + b) \quad (20)$$

On pourrait voir que la connaissance de a_1, a_2, b entraîne l'existence d'une infinité de systèmes $A_1 > 0, A_2 > 0, 0 \leq p_{12} \leq 1, 0 < p_{21} \leq 1$ dès que :

$$\begin{cases} m_{10} > \mu_{11} > 0 \\ m_{01} > \mu_{11} > 0 \end{cases} \quad (21)$$

m_{10} et m_{01} étant les moyennes des distributions marginales de X_1, X_2 et μ_{11} leur covariance (ce qui entraîne que a_1, a_2 et b soient positifs).

Lorsque les conditions (21) sont réalisées, il existe un seul système de valeurs $A_1 = A, A_2 = B$ et $p_{12} = p_{21} = p$. La fréquence théorique du couple X, Y peut alors être calculée, en particulier, par les formules (1) et (2).

En reprenant l'analyse faite initialement, on pourrait voir également, en faisant par exemple $p_{12} = 0$, que la fonction de fréquence du couple X, Y peut s'exprimer en fonction des moments m_{10}, m_{01}, μ_{11} suivant la formule :

$$f(X, Y) = e^{-(a_1 + a_2 + b)} a_1^X a_2^Y \sum_{k=0}^{\min(X, Y)} \frac{\omega^k}{(X - k)! k! (Y - k)!} \quad (22)$$

avec

$$\begin{cases} \omega = \frac{b}{a_1 a_2} \\ a_1 = m_{10} - \mu_{11} \\ a_2 = m_{01} - \mu_{11} \\ b = \mu_{11} \end{cases} \quad \text{et} \quad \begin{cases} K = X \text{ si } X < Y \\ K = Y \text{ si } Y < X \end{cases}$$

formule d'emploi plus commode que (1) ou (2) et qui a l'avantage de faire apparaître directement les moyennes m_{10} et m_{01} de X et Y et la covariance $(X, Y) = b = \mu_{11}$.

L'application de la formule (22) aux tableaux n° 1 et n° 3 avec

$$m_{10} = \frac{\sum X}{n}, m_{01} = \frac{\sum Y}{n} \text{ et } \mu_{11} = \frac{\sum XY}{n} - \frac{(\sum X)(\sum Y)}{n^2}$$

aurait permis de calculer les fréquences théoriques figurant aux tableaux n° 2 et n° 4.

L'étude développée à partir d'un cas très précis apparaît ainsi comme un mode de génération très particulier des distributions à deux variables aléatoires X, Y dont la seconde fonction caractéristique à la forme (20) les coefficients a_1 , a_2 et b étant positifs.

Il va sans dire, d'après ce qui précède que les résultats obtenus en (I - 4, 5, 6, 7) sont valables, dans le cas des distributions à deux variables X, Y, quels que soient A_1 , A_2 , p_{12} , p_{21} . En particulier, la seconde fonction caractéristique de Y pour X donné est :

$$\psi_Y(V) = (m_{01} - \mu_{11}) (V - 1) + X \log \left[1 + \frac{\mu_{11}}{m_{10}} (V - 1) \right]$$