QUATRE FAMILLES DISCRÈTES

par

Georges MATHERON

N–703

Juin 1981

CENTRE DE GÉOSTATISTIQUE

35, RUE SAINT–HONORÉ, 77305 FONTAINEBLEAU (France)

ECOLE DES MINES
DE PARIS

ÉCOLE NATIONALE SUPÉRIEURE DES MINES DE PARIS

QUATRE FAMILLES DISCRETES

G. MATHERON

Juin 1981

TABLE DES MATIERES

QUATRE FAMILLES DISCRETES

G. MATHERON

Juin 1981

INTRODUCTION

Le point de départ de ce travail est une réflexion sur le modèle proposé par Sichel pour représenter des distributions discrètes à queues très allongées. Du point de vue de la géostatistique, cependant, l'intérêt se concentre sur le problème du changement de support. Mais, s'agissant de lois discrètes, les techniques usuelles, fondées sur les anamorphoses, sont inopérantes : il n'existe pas, en effet, de transformation bijective croissante appliquant sur lui-même l'ensemble des entiers naturels. On est donc contraint de rechercher des lois discrètes multivariables admettant comme marginales des lois du type de Sichel (ou d'autres analogues). Or cela n'est pas facile du tout en général. Seul le cas particulier des lois discrètes <u>indéfiniment divisibles</u> se prête à peu près bien à ce genre de recherche. Or le modèle de Sichel (qui dépend de trois paramètres) contient bien un sous-modèle indéfiniment divisible, qui est du reste celui que l'on utilise le plus souvent à cause de ses bonnes propriétés et qui a fait l'objet de travaux très approfondis et d'un très haut intérêt (Sichel). Mais ce sous-modèle ne dépend plus que de deux paramètres, et cela se révèle insuffisant. D'où notre premier fil directeur : il s'agit de trouver un (ou plusieurs) modèles de lois indéfiniment divisibles, dépendant au moins de trois paramètres, et contenant comme cas particulier le sous-modèle bien éprouvé de Sichel.

Parmi ses bonnes propriétés, le modèle (général) de Sichel contient une <u>relation de récurrence</u> entre p_{n+2} , p_{n+1} et p_n , qui permet donc le calcul numérique rapide des probabilités p_n une fois connues les deux premières p_0 et p_1 . Cette relation de récurrence à trois termes provient du fait que la fonction génératrice G(s) de ce modèle vérifie une équation différentielle du second ordre d'un type très simple. D'où le second fil directeur : comme nous souhaitons que notre modèle de lois indéfiniment divisibles contienne lui aussi une relation de récurrence à trois termes, nous imposerons à G(s) de vérifier l'équation différentielle du second ordre la plus générale qui entraine l'existence de ces relations de récurrence. Comme nous le verrons cette équation différentielle générale est celle des fonctions hypergéométriques (celle de Sichel est un cas particulier dégénéré). D'où finalement le problème que nous nous proposons d'étudier : parmi les lois discrètes dont la fonction génératrice est du type hypergéométrique, trouver celles qui sont indéfiniment divisibles.

Parallèlement à cette recherche de lois discrètes, nous aborderons celles des lois continues (associées à des variables positives). Il s'agit en fait de la même recherche : de fait, si $\Phi(\lambda)$ est la transformée de Laplace d'une variable positive, la fonction

$$G(s) = \Phi(1-s)$$

est la fonction génératrice du mélange poissonien (Poisson compound) correspondant - et G est indéfiniment divisible dès que Φ possède cette propriété. La réciproque n'est pas vraie, en général : si G(s) est une fonction génératrice, en posant

$$\Phi(\lambda) = G(1-\lambda)$$

on n'obtient pas forcément une loi continue, sauf précisément si G est un mélange poissonien. Ainsi, pour chacune de nos familles de lois discrètes, nous rechercherons s'il s'agit de mélange poissonien et, le cas échéant, nous donnerons l'expression de la loi continue correspondante. En fait, sur les quatre familles que nous avons trouvées, deux seulement sont des mélanges poissoniens, les deux autres restant autonomes.

Dans un premier paragraphe, je rappelle le modèle de Sichel et ses principales propriétés. Je pose ensuite les équations différentielles correspondant au problème énoncé plus haut, et je fais l'inventaire des solutions possibles. Parmi celles-ci, je recherche ensuite celles aui correspondent effectivement à des lois : il apparaît qu'il en existe (outre le cas "rationnel") quatre familles, dépendant chacune de trois paramètres, dont j'examine rapidement les propriétés: le lecteur préssé pourra se contenter de lire les paragraphes 4 et 5 où sont présentées ces lois à queue allongée.

1 - LE MODELE DE SICHEL.

Le modèle général de Sichel est un mélange poissonien constitué à partir d'une variable <u>positive</u> X admettant une densité de la forme :

$$f(x) = A \frac{e^{-ax - \frac{b}{x}}}{x^{1+\gamma}} \tag{1-1}$$

(a et b sont des constantes > 0, γ peut être positif ounnon, A est une constante de normalisation). Au point de départ des calculs auxquels se prête ce modèle figure la représentation de Sonine-Schläfli des fonctions de Bessel modifiées de seconde espèce :

$$(c\rho)^{\mu} K_{-\mu}(c\rho) = \frac{1}{2} c^{2\mu} \int_0^{\infty} e^{-\frac{u\rho^2}{2} - \frac{c^2}{2u}} \frac{du}{u^{\mu+1}}$$

On en tire l'expression de la constante A de (1-1)

$$\frac{1}{A} = 2 \left(\frac{a}{b}\right)^{\gamma/2} K_{-\gamma} \left(2\sqrt{aB}\right)$$

et, plus généralement, la transformée de Laplace $\Phi(\lambda)$ de la densité f(x) :

$$\Phi(\lambda) = A\ 2 \left(\frac{a+\lambda}{B}\right)^{\gamma/2} K_{-\gamma} \left(2\sqrt{b(a+\lambda)}\right)$$

Le mélange poissonien correspondant admet la fonction génératrice :

$$G(s) = \sum p_n s^n = \Phi(1-s)$$

qui s'exprime donc, elle aussi, à l'aide de fonctions de Bessel. Les probabilités p_n elles-mêmes sont données par :

$$p_n = \frac{1}{n!} \int_0^\infty x^n e^{-x} f(x)\, dx$$

A l'aide de la même formule de Sonine-Schäfli, il vient :

$$p_n = \frac{2A}{n!} \left(\frac{a}{b}\right)^{\gamma - \frac{n}{2}} K_{n-\gamma} (2\sqrt{ab})$$

Les propriétés classiques des fonctions de Bessel permettent ensuite d'établir les relations de récurrence entre p_n, et l'équation différentielle correspondante que vérifie G(s), et qui est de la forme

$$(1\text{-}2) \qquad (a_0 + a_1 s)\, G'' + a_2 G' + a_3 G = 0$$

(a_0, a_1, a_2, a_3 sont des constantes convenables).

Nous verrons ci-dessous que les seules solutions de (1-2) qui soient associées à des lois indéfiniment divisibles correspondent au cas

$$\gamma = \frac{1}{2}$$

C'est ce cas qui va surtout retenir notre attention.

Le sous-Modèle I.D. ($\gamma = \frac{1}{2}$)

Pour $\gamma = \frac{1}{2}$, la transformée de Laplace de la loi (1-1) prend une forme exponentielle très simple :

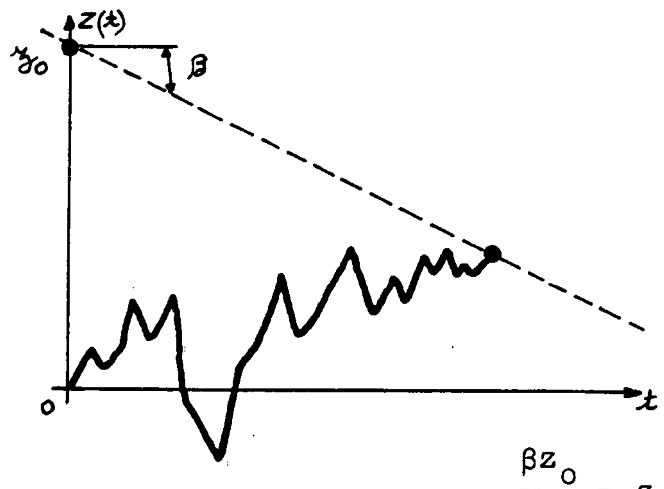
$$(1\text{-}3) \qquad \Phi(\lambda) = e^{2\sqrt{ab} - 2\sqrt{b(a+\lambda)}}$$

Pour a = 0, il s'agit donc de la loi stable

$$\Phi(\lambda) = e^{-c\sqrt{\lambda}}$$

loi qui n'admet pas de moyenne finie. Le facteur exponentiel e^{-ax} a justement pour but d'accélérer la décroissance de f(x), pour assurer l'existence d'une moyenne et d'une variance : mais ces lois de Sichel restent très proches de la loi stable a = 0, et cela explique leur queue très allongée.

La loi (1-3) peut être rattachée au mouvement brownien. Soit, en effet, Z(t) un mouvement brownien d'espérance nulle et de variance ct : l'instant aléatoire T où ce processus atteint pour la première fois la droite d'équation



$$z = z_0 - \beta t$$

(z_0 et β, constantes positives données) admet la transformée de Laplace

$$E(e^{-\lambda T}) = e^{\frac{\beta z_0}{c} - z_0\sqrt{\frac{2\lambda}{c} + \frac{\beta^2}{c^2}}}$$

et la densité de probabilité

$$f(t) = \frac{1}{\sqrt{2\pi c}} \frac{z_0}{t^{3/2}} e^{\beta\frac{z_0}{c} - \frac{z_0^2}{2ct} - \frac{\beta^2 t}{2c}}$$

Il s'agit bien, aux notations près, de la loi (1-3) pour $\gamma = \frac{1}{2}$.

Passons au mélange poissonien : nous obtenons sa fonction génératrice G(s) en changeant λ en 1-s dans (1-3). En adoptant les notations commodes introduites par Sichel, on trouve une expression de la forme :

(1-4) $$G(s) = e^{\alpha(\sqrt{1-\theta} - \sqrt{1-\theta s})}$$

($\alpha > 0$, $0 < \theta \leq 1$: le cas $\theta = 1$ correspond à $a = 0$, c'est-à-dire au cas de la loi stable sans moyenne finie). Comme moyenne et variance, on trouve :

$$\begin{cases} m = \dfrac{\alpha\theta}{2\sqrt{1-\theta}} \\ \sigma^2 = m + \dfrac{\alpha\,\theta^2}{4\,(1-\theta)^{3/2}} \end{cases}$$

En ce qui concerne les p_n, pour $n = 0$ et $n = 1$:

$$\begin{cases} p_o = e^{-\alpha(1-\sqrt{1-\theta})} \\ p_1 = \dfrac{\alpha\theta}{2}\, p_o \end{cases}$$

Les autres p_n se calculent facilement par récurrence. De fait, G(s) vérifie l'équation différentielle :

$$(1\text{-}5) \qquad (1-\theta s)\, G'' - \frac{\theta}{2}\, G' - \frac{\alpha^2\,\theta^2}{4}\, G = 0$$

Il suffit de remplacer, dans (1-5), G(s) par son développement $\Sigma\, p_n\, s^n$ et d'annuler le coefficient du terme en s^n pour obtenir :

$$(1\text{-}6) \qquad (n+2)(n+1)\, p_{n+2} = \theta\,(n+1)(n+\tfrac{1}{2})\, p_{n+1} + \frac{\alpha^2\,\theta^2}{4}\, p_n$$

Ce sont ces deux propriétés : (1-4), qui exprime la divisibilité indéfinie, et (1-5), qui exprime l'existence d'une relation de récurrence entre les p_n , que nous allons chercher à généraliser. Avant de passer à cette étude, notons encore quelques points condernant la loi (1-4) de Sichel.

Décomposition en Paquets Poissoniens.

De manière générale, on sait que la condition nécessaire et suffisante pour qu'une variable discrète N soit indéfiniment divisible est que sa fonction génératrice $G(s) = E(s^N)$ soit de la forme :

$$G(s) = e^{-a(1-\gamma(s))} \tag{1-7}$$

où $\gamma(s)$ est elle-même une fonction génératrice. Ecrivant (1-7) sous la forme développée :

$$G(s) = \sum \frac{a^k}{k!} \; e^{-a} \; \gamma^k(s)$$

on voit que $G(s)$ se construit à partir de $\gamma(s)$ par poissonisation. Concrètement: introduisons une variable K admettant la loi de Poisson de paramètre a, et des variables indépendantes L_i admettant la même loi $\gamma(s)$. Alors la variable :

$$N = \sum_{i=1}^{K} L_i$$

(étant entendu que $N = 0$ pour $K = 0$) admet la loi G : la variable aléatoire N représentant un certain nombre d'objets, on peut considérer que cet ensemble d'objets est la réunion d'un nombre K (aléatoire poissonien) de sous-paquets, chaque sous-paquet contenant un nombre aléatoire L_i (de loi $\gamma(s)$) d'objets.

Notons aussi que, dans l'écriture

$$\gamma(s) = \varpi_0 + \sum_{n \geq 1} \varpi_n \; s^n$$

de la loi $\gamma(s)$, on peut toujours supposer $\varpi_0 = 0$: si ϖ_0 n'est pas nul, on aura toujours

$$a\;(1-\gamma(s)) = a_0\;(1-\gamma_0(s))$$

avec

$$\gamma_0(s) = \sum_{n \geq 1} \frac{\varpi_n}{1-\varpi_0} \; s^n \quad ; \quad a_0 = (1-\varpi_0)a$$

(concrètement, $\varpi_0 \neq 0$ correspond au cas où certains de nos sous-paquets d'objets peuvent être vides : dans certains modèles, cette possibilité est intéressante, mais il est en général plus simple de ne considérer que les paquets non vides).

Avec $\varpi_0 = 0$, la loi (1-4) est de la forme

$$G(s) = e^{-\nu(1-\gamma(s))}$$

avec

$$\nu = \alpha\,(1-\sqrt{1-\theta}) \quad ; \quad \gamma(s) = \frac{1-\sqrt{1-\theta s}}{1-\sqrt{1-\theta}}$$

Ainsi, la dédomposition de la loi de Sichel en paquets poissoniens est de la forme :

$$(1\text{-}8) \qquad G(s) = \sum \frac{\nu^n}{n!}\, e^{-\nu} \left(\frac{1-\sqrt{1-\theta s}}{1-\sqrt{1-\theta}} \right)^n$$

La loi $\gamma(s) = \dfrac{1-\sqrt{1-\theta s}}{1-\sqrt{1-\theta}}$ est celle du nombre d'objets constituant un paquet, et le nombre de ces paquets est lui-même un Poisson de moyenne ν. L'intérêt se concentre sur cette loi $\gamma(s)$ très particulière (elle ne dépend que d'un seul paramètre). Notons que l'on peut écrire :

$$\gamma(s) = s\,\frac{1+\sqrt{1-\theta}}{1+\sqrt{1-\theta s}}$$

Chaque paquet contient ainsi un objet plus un nombre aléatoire (éventuellement nul) admettant la fonction génératrice

$$g(s) = \frac{1+\sqrt{1-\theta}}{1+\sqrt{1-\theta s}}$$

De même, lorsqu'il y a n paquets d'objets, la loi du nombre total de ces objets est

$$\gamma^n(s) = s^n\, g^n(s)$$

Ce nombre total est donc n plus un nombre aléatoire (éventuellement nul) de loi $g^n(s)$.

Or, il se trouve que la loi g est <u>indéfiniment divisible</u> (nous la retrouverons ultérieurement). Pour tout a > 0 (entier ou non), la fonction

$$(1\text{-}9) \qquad g^a(s) = \left(\frac{1+\sqrt{1-\theta}}{1+\sqrt{1-\theta s}}\right)^a$$

est une fonction génératrice. De fait, g^a vérifie l'équation différentielle :

$$s(1-\theta s)\ F'' + (a+1 - \theta(a+\tfrac{3}{2})\ s)\ F' - \theta\ \frac{a(a+1)}{4}\ F = 0$$

Sous forme réduite (c'est-à-dire avec le changement de variable $x = \theta s$), cette équation devient :

$$s(1-s)\ F'' + (a+1 - (a+\tfrac{3}{2})\ s)\ F' - \frac{a(a+1)}{4}\ F = 0$$

C'est un cas particulier de l'équation des hypergéométriques réduites :

$$(1\text{-}10) \qquad s(1-s)\ F'' + (\gamma - (1+\alpha+\beta)s)\ F' - \alpha\ \beta\ F = 0$$

à savoir le cas $\gamma = a+1$, $\alpha = \frac{a}{2}$, $\beta = \frac{a+1}{2}$. Or on sait que l'unique solution régulière en $x = 0$ de l'équation (1-10) est, à un facteur près, la fonction hypergéométrique :

$$(1\text{-}11) \qquad F(\alpha,\beta,\gamma;x) = 1 + \sum_{n=1}^{\infty} \frac{\alpha(\alpha+1)..(\alpha+n-1)\ \beta(\beta+1)..(\beta+n-1)}{\gamma(\gamma+1)...(\gamma+n-1)} \frac{x^n}{n!}$$

Comme $g^a(s)$ est régulière à l'origine, nous concluons :

$$(1\text{-}12) \qquad g^a(s) = \left(\frac{1+\sqrt{1-\theta}}{1+\sqrt{1-\theta s}}\right)^a = \left(\frac{1+\sqrt{1-\theta}}{2}\right)^a F(\tfrac{a}{2}\ ,\ \tfrac{a+1}{2}\ ,\ a+1\ ;\theta s)$$

Explicitement, donc, les probabilités $\varpi_n(a)$ associées à la loi

$$g^a(s) = \sum \varpi_n(a)\ s^n$$

sont données par

$$\varpi_n(a) = \left(\frac{1+\sqrt{1-\theta}}{2}\right)^a \frac{\frac{a}{2}\ (\frac{a}{2}+1)..(\frac{a}{2}+n-1)(\frac{a+1}{2})..(\frac{a+1}{2}+n-1)}{(a+1)(a+2)..(a+n)} \frac{\theta_n}{n!}$$

Il est plus commode d'exprimer ceci à l'aide de la fonction eulerienne $\Gamma(x)$: on trouve

$$\varpi_n(a) = \left(\frac{1+\sqrt{1-\theta}}{2}\right)^a \frac{\Gamma(a+1)}{\Gamma(\frac{a}{2})\ \Gamma(\frac{a+1}{2})} \frac{\Gamma(\frac{a}{2}+n)\ \Gamma(\frac{a+1}{2})+n)}{\Gamma(a+1+n)} \frac{\theta^n}{n!}$$

On peut encore simplifier cette expression à l'aide de la formule classique de duplication de l'eulerienne Γ :

$$\Gamma(2b) = \frac{\Gamma(b)\ \Gamma(b+1/2)}{\sqrt{\pi}}\ 2^{2b-1}$$

Avec $b = a/2$ et $b = a/2 + n$, cette formule donne :

$$\Gamma(\tfrac{a}{2})\ \Gamma(\tfrac{a+1}{2}) = 2^{1-a}\sqrt{\pi}\,\Gamma(a)$$

$$\Gamma(\tfrac{a}{2}+n)\ \Gamma(\tfrac{a+1}{2}+n) = 2^{1-a-2n}\sqrt{\pi}\ \Gamma(a+2n)$$

et il vient

(1-12') $$\varpi_n(a) = \left(\frac{1+\sqrt{1-\theta}}{2}\right)^a a\,\frac{\Gamma(a+2n)}{n!\ \Gamma(a+1+n)} \left(\frac{\theta}{4}\right)^n$$

Revenons maintenant à la loi $\gamma = s\ g(s)$ du nombre d'objets dans un paquet. Du fait de la présence du facteur s^a dans $\gamma^a = s^a\ g^a(s)$, il n'est pas vrai que $\gamma(s)$ soit indéfiniment divisible. Mais, pour n entier, $\gamma^n = s^n\ g^n$ représente la loi du nombre k d'objets contenus dans la réunion de n paquets. La probabilité $p_k(n)$ correspondante est

$$p_k(n) = \left(\frac{1+\sqrt{1-\theta}}{2}\right)^n n\,\frac{\Gamma(2k-n)}{n!\ k!} \left(\frac{\theta}{4}\right)^{k-n}$$

(pour $k \geq n$, et $p_k(n) = 0$ pour $k < n$: pour $n = 0$, $p_k(n) = 0$ sauf pour $k = 0$ et $p_o(o) = 1$).

Si maintenant nous remontons à la loi $G(s)$ écrite en (1-8), nous voyons que la probabilité p_k correspondante est

$$p_k = \sum_{n=o}^{k} \frac{\nu^n}{n!}\ e^{-\nu}\ p_k(n)$$

et, compte tenu de l'expression trouvée ci-dessus, on retrouve l'expression de p_k au moyen de fonctions de Bessel telle que Sichel l'a établie. Mais il n'y a pas lieu d'insister sur ces calculs. Notre but essentiel était de montrer la divisibilité indéfinie de la loi $g^a(s)$ définie en (1-9), loi qui jouera un certain rôle dans la suite, et, notamment, d'obtenir les expressions explicites (1-12) et (1-12'). D'après ces expressions, on peut voir que lorsque k est grand

$$\varpi_k(a) \sim \frac{\theta^k}{k^{3/2}}$$

Ce comportement asymptotique est le même que pour la loi de Sichel.

Une Première Généralisation.

A partir de la loi de Sichel (1-4), nous pouvons dès maintenant former une famille à 3 paramètres vérifiant les conditions que nous nous sommes imposées (divisibilité indéfinie et relation de récurrence). Il suffit pour cela de randomiser le paramètre α en le remplaçant par une variable aléatoire X admettant une loi gamma de paramètres b et ν. Nous obtenons ainsi une nouvelle fonction génératrice G dont l'expression est :

$$G(s) = \frac{b^\nu}{\Gamma(\nu)} \int_0^\infty e^{-x(\sqrt{1-\theta s}) + x\sqrt{1-\theta}}\, x^{\nu-1}\, e^{-bx} dx$$

c'est-à-dire :

$$G(s) = \left(\frac{b}{b-\sqrt{1-\theta} + \sqrt{1-\theta s}} \right)^\nu$$

Il sera plus agréable de remplacer b par $a = b - \sqrt{1-\theta}$.

Ainsi, d'après le procédé même de construction, nous savons que la fonction

(1-13) $$\boxed{G(s) = \left(\frac{a + \sqrt{1-\theta}}{a + \sqrt{1-\theta s}} \right)^\nu}$$

est une fonction génératrice, pourvu seulement que :

$$\nu \geq 0 \quad , \quad a > \sqrt{1-\theta} \tag{1-13'}$$

On notera, en particulier, que le paramètre a peut prendre des valeurs négatives, puisqu'il peut varier de $-\sqrt{1-\theta}$ à $+\infty$.

Pour $a = 0$, en particulier, on obtient la loi binomiale négative. Pour $a = 1$, nous retrouvons la forme hypergéométrique réduite (1-12). Enfin, si a et ν tendent vers l'infini avec $\nu/a = \alpha = C^{ste}$, on obtient comme cas limite la loi de Sichel (1-4) dont nous sommes partis.

En ce qui concerne les caractéristiques de cette loi, et, notamment, l'équation différentielle et les relations de récurrence, nous les établirons plus loin dans un cadre général, après avoir procédé à l'inventaire exhaustif des possibilités. Notons seulement la décomposition en paquets de cette nouvelle famille de lois.

Décomposition en Paquets.

Comme toute loi indéfiniment divisible, la loi (1-13) admet une décomposition en paquets poissoniens, avec

$$\gamma(s) = 1 + \ell_n \left(\frac{a + \sqrt{1-\theta}}{a + \sqrt{1-\theta s}} \right)$$

Il n'est pas difficile de trouver les probabilités $\varpi_n(a)$ associées (la dérivée $\gamma'(s)$ s'exprime facilement comme hypergéométrique). Mais il est plus intéressant de mentionner ici une autre décomposition, en paquets non poissoniens, dont le nombre obéit à une loi binomiale négative.

Revenons pour cela à l'expression de la loi (1-13) comme randomisée de la loi de Sichel, mais en remplaçant cette fois la fonction génératrice de cette dernière loi par son développement (1-8) en paquets poissoniens. Cela donne (avec $\nu = x(1-\sqrt{1-\theta})$)

$$G(s) = \frac{b^\nu}{\Gamma(\nu)} \sum_{n=o}^{\infty} \int_o^{\infty} \frac{x^n}{n!} (1-\sqrt{1-\theta s})^n \; e^{-x(1-\sqrt{1-\theta})} \; x^{\nu-1} \; e^{-bx} dx$$

$$= \frac{b^\nu}{\Gamma(\nu)} \sum_{n=o}^{\infty} \frac{\Gamma(n+\nu)}{(b-\sqrt{1-\theta}+1)^{n+\nu}} \frac{(1-\sqrt{1-\theta s})^n}{n!}$$

Soit, avec la notation $a = b - \sqrt{1-\theta}$ qui est celle de (1-13)

$$G(s) = \left(\frac{a+\sqrt{1-\theta}}{a+1}\right)^{\nu} \frac{1}{\Gamma(\nu)} \sum_{n=o}^{\infty} \frac{\Gamma(n+\nu)}{n!} \left(\frac{1-\sqrt{1-\theta}}{a+1}\right)^{n} \left(\frac{1-\sqrt{1-\theta s}}{1-\sqrt{1-\theta}}\right)^{n}$$

D'une manière plus synthétique, en posant

$$p = \frac{1-\sqrt{1-\theta}}{a+1} \quad , \quad q = 1 - p = \frac{a+\sqrt{1-\theta}}{a+1}$$

et en désignant par ϖ_n les probabilités

$$\varpi_n = q^\nu \frac{\Gamma(n+\nu)}{n!} p^n$$

qui sont celles de la binomiale négative $(q/(1-ps))^\nu$, ce résultat s'écrit :

$$\left(\frac{a+\sqrt{1-\theta}}{a+\sqrt{1-\theta s}}\right)^{\nu} = \sum_n \varpi_n \left(\frac{1-\sqrt{1-\theta s}}{1-\sqrt{1-\theta}}\right)^{n}$$

et met en évidence la décomposition annoncée : <u>la loi du nombre d'objets par paquets est la même que pour la loi de Sichel</u>, mais, au lieu d'une loi de Poisson, nous trouvons une <u>binomiale négative</u> comme loi du nombre de paquets.

Une Simulation.

Avant de passer à l'inventaire général, donnons tout d'abord un procédé permettant de simuler les lois de cette première famille, y compris la loi de Sichel, et peut-être de mieux comprendre ainsi leurs propriétés.

Considérons deux processus de Poisson $N_o(t)$ et $N_1(t)$ admettant les densités (différentes) a_0 et a_1. A l'instant initial $t = 0$, prenons $N_0(o) = 0$ et $N_1(o) = 1$. Nous supposons $a_0 > a_1$, de manière à ce que le processus N_0 rattrape le processus N_1 avec une probabilité unité. Nous désignerons par

$$N = N_0(T) = N_1(T)$$

cette valeur commune des deux processus à l'instant aléatoire T où le premier processus rattrape le second. Je dis que cette variable N admet une loi du type

$$\gamma(s) = \frac{1 - \sqrt{1-\theta s}}{1 - \sqrt{1-\theta}}$$

c'est-à-dire la même loi que le nombre d'objets par paquets dans le modèle ci-dessus.

De fait, si nous réunissons les sauts des deux processus N_0 et N_1, nous obtenons un nouveau processus de Poisson, admettant comme densité la somme $a_0 + a_1$, et, indépendamment des autres, chacun des sauts de ce nouveau processus appartient à N_0 avec la probabilité

$$p = \frac{a_0}{a_0 + a_1}$$

où à N_1 avec la probabilité complémentaire $q = 1 - p$.

Raisonnons sur le premier saut du processus somme. S'il appartient à N_0, ce qui a lieu avec la;probabilité p, N_0 a rattrapé N_1 et donc $N = 1$. Si, au contraire, ce premier saut appartient à N_1 (probabilité q), il y aura maintenant deux échelons à rattraper au lieu d'un seul. Ainsi, dans cette hypothèse, N sera N' + N" où N' et N" sont deux variables indépendantes admettant la même loi G(s) que N. Au total, il vient ainsi :

$$G(s) = p\,s + q\,(G(s))^2$$

soit

$$G = \frac{1 \pm \sqrt{1 - 4\,pqs}}{2q}$$

Comme G(s) est une fonction croissante, c'est le signe - qui convient, soit

$$G(s) = \frac{1 - \sqrt{1 - 4\,pqs}}{2q}$$

Posons $\theta = 4\,pq$, d'où

$$\sqrt{1 - \theta} = \sqrt{(p+q)^2 - 4\,pq} = |p-q|$$

Comme on a supposé $p > q$, il vient en $s = 1$

$$G(1) = \frac{1 - p + q}{2q} = 1$$

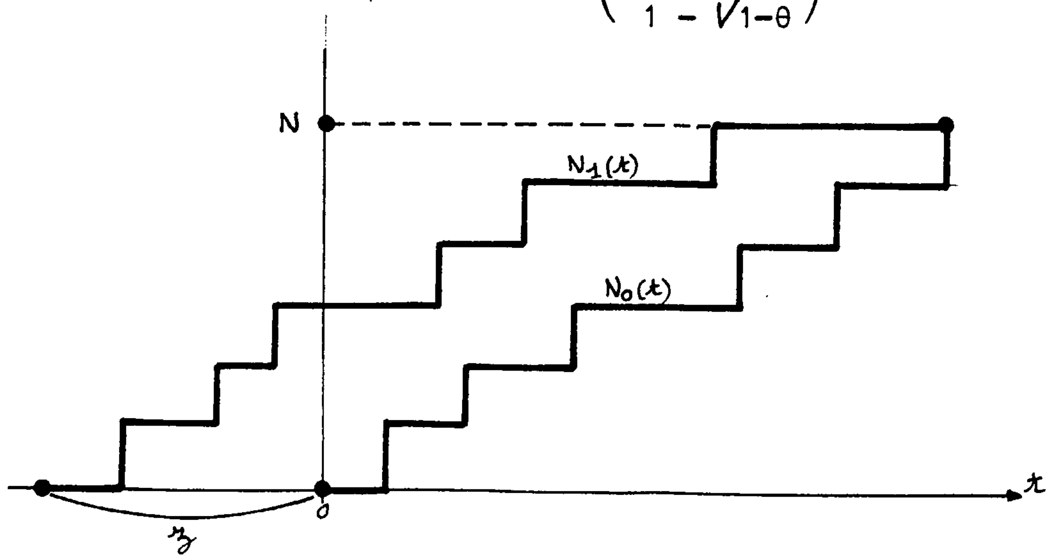
il s'agit donc bien d'une loi (non défective, i.e. le premier processus rattrape presque sûrement le second) et :

$$G(s) = \frac{1 - \sqrt{1-\theta s}}{1 - \sqrt{1-\theta}}$$

comme annoncé.

Si maintenant le second processus est en $N_1(o) = n$ au temps 0, la cote à laquelle aura lieu le rattrapage aura la loi

$$G(s)^n = \left(\frac{1 - \sqrt{1-\theta s}}{1 - \sqrt{1-\theta}}\right)^n$$



Supposons maintenant, pour corser l'affaire, que le processus $N_1(t)$ soit parti de 0 à l'instant $t = -z$ antérieur au départ de $N_0(t)$. A l'instant $t = 0$, $N_1(t)$ est alors une variable poissonienne de moyenne $a_1 z$. Posant $a_1 z = \nu$, nous trouvons cette fois pour N (altitude du rattrapage) la loi poissonisée :

$$G(s) = \sum_{n=o}^{\infty} \frac{\nu^n}{n!} e^{-\nu} \left(\frac{1 - \sqrt{1-\theta s}}{1 - \sqrt{1-\theta}} \right)^n = e^{\alpha(\sqrt{1-\theta} - \sqrt{1-\theta s})}$$

C'est la <u>loi de Sichel</u> avec $\alpha = \nu/(1 - \sqrt{1-\theta})$.

Enfin, pour terminer, nous pouvons encore supposer que z (l'avance de départ du processus N_1 sur celui de N_0) est une variable aléatoire avec une loi gamma : d'après le procédé même qui nous a servi à construire la loi (1-13), nous voyons que cette fois la loi de N est justement du type général (1-13).

<u>NB</u> On peut aussi s'intéresser à la loi simultanée de N et à l'instant T où le rattrapage a lieu : c'est là un exercice de probabilité fort intéressant que j'abandonne à la sagacité de mon lecteur.

2 - L'INVENTAIRE.

Abordons maintenant une tâche plus ingrate. Nous cherchons à déterminer la classe des fonctions ϕ possédant les propriétés suivantes :

~ Pour tout $\nu > 0$, la fonction

$$G(s) = e^{\nu\phi(s)} \tag{2-1}$$

est une fonction génératrice (ou bien, si nous nous intéressons au cas continu : $\Phi_\nu(\lambda) = \exp(\nu\phi(\lambda))$ est transformée de Laplace d'une loi de probabilité).

$\sim G_\nu(s)$ (ou $\Phi_\nu(\lambda)$) vérifie (pour tout $\nu > 0$) une équation différentielle du second ordre, du type :

$$(2\text{-}2) \qquad A_\nu(s)\ G_\nu'' + B_\nu(s)\ G_\nu' + C_\nu\ G_\nu = 0$$

où A_ν est un polynome du second degré en s, B_ν un polynome du premier degré et C_ν une constante.

(Cette seconde condition nous garantit l'existence de relations de récurrence à trois termes). Cette équation (2-2) est celle de la classe générale des fonctions hypergéométriques. Notre tâche consiste donc à trouver les fonctions hypergéométriques qui sont les fonctions génératrices de lois indéfiniment divisibles.

Portant l'expression (2-1) de G dans l'équation différentielle (2-2), nous voyons que la fonction ψ (qui ne dépend pas de ν) doit vérifier <u>pour tout $\nu > 0$</u> l'équation différentielle non linéaire :

$$(2\text{-}3) \qquad A_\nu\ \psi'' + \nu\ A_\nu\ \psi'^2 + B_\nu\ \psi' + \frac{C_\nu}{\nu} = 0$$

Dans ce paragraphe, nous examinons les conditions que doivent vérifier les polynomes A_ν, B_ν et la constante C_ν pour que ceci soit possible, et nous déterminerons les fonctions ψ (ou plutôt leurs dérivées ψ'). Dans les paragraphes suivants, nous sélectionnerons, parmi ces solutions ψ possibles, celles qui correspondent à des lois de probabilité, et nous étudierons ensuite les propriétés de ces lois.

Voici un premier théorème :

Pour que la condition (2-3) soit satisfaite, il est <u>nécessaire</u> que l'on puisse trouver un polynome A(s) du second degré en s, deux polynomes B(s) et $B_o(s)$ du premier degré, et deux constantes C et C_o telles que l'on ait :

$$(2\text{-}4) \qquad \begin{cases} A\ \psi'' + B_o\ \psi' + C_o = 0 \\ A\ \psi'^2 + 2\ B\ \psi' + C = 0 \end{cases}$$

Le point essentiel est que c'est le même polynome A(s) qui figure dans les deux équations (2-4).

Ce théorème est presque intuitif, mais la démonstration rigoureuse est assez ingrate. Je me contente de l'esquisser. Partons de (2-3) supposée vérifiée. En écrivant ces relations pour deux valeurs différentes λ_0 et λ, nous pouvons éliminer le terme en x^2 qui figure dans les coefficients A_{λ_0} et A_λ de ψ'' et former une nouvelle relation duntype

$$\dot{A}_\lambda \psi'' + P_\lambda \psi'^2 + \dot{B}_\lambda \psi' + \dot{C}_\lambda = 0$$

où $\dot{A}_\lambda$ est de degré 1 en s, P_λ de degré 2, $\dot{B}_\lambda$ de degré 1 et $\dot{C}_\lambda$ constante en s. Réitérant cette opération, nous aboutissons à une équation au moins de la forme

(a) $$A \psi'^2 + 2 B \psi' + C = 0$$

En éliminant de la même façon le terme en ψ'^2, on aboutit à

(b) $$A_0 \psi'' + B_0 \psi' + C_0 = 0$$

avec A et A_0 de degré ≤ 2, B et B_0 de degré ≤ 1, C et C_0 constantes. Tout se ramène à montrer que A et A_0 sont égaux (à un facteur près).

Aux deux équations (a) et (b) ci-dessus, adjoignons l'équation (2-3). Si cette équation (2-3) n'est pas une conséquence des deux premières, le problème est largement surdéterminé. De fait, on voit facilement que l'ensemble des λ pour lesquels (2-3) ne résulte pas de (a) et (b) est ouvert : si donc il n'est pas vide, il existera une infinité de valeurs de λ pour lesquelles (2-3) sera indépendante de (a) et (b). On pourra donc former un second système (a'), (b') analogue au premier, mais avec des polynomes différents A', B', C', A_0' , B_0' , C_0' . Mais (a) est une équation du second degré en ψ', donc détermine deux solutions possibles ψ_1' et ψ_2' dont l'une au moins doit être solution de (a') ainsi que de (b) et (B'). Sauf cas de dégénérescence extrême (que nous retrouverons de toute façon dans la discussion ultérieure), on aboutit à la conclusion que les polynomes A, A' etc.. sont proportionnels, et donc à une contradiction.

On conclut donc que (2-3) est une conséquence de (a) et (b) : il existe ainsi deux fonctions $\alpha(\lambda)$ et $\beta(\lambda)$ telles que

$$A_\lambda = \alpha(\lambda)\ A_o$$

$$\lambda\ A_\lambda = \beta(\lambda)\ A$$

$$B_\lambda = \alpha(\lambda)\ B_o + \beta(\lambda)\ B$$

$$\frac{C_\lambda}{\lambda} = \alpha(\lambda)\ C_o + \beta(\lambda)\ C$$

Mais les deux premières relations indiquent justement que A et A_o sont proportionnels : on peut donc se ramener à :

$$A = A_o$$

Ayant maintenant les conditions <u>nécessaires</u> (2-4), examinons sous quelles conditions supplémentaires il existe une solution.

Comme la seconde relation (2-4) est une simple équation de second degré par ψ', on a :

$$\psi' = -\frac{B}{A} + \varepsilon\frac{\sqrt{\Delta}}{A}$$

$$(\varepsilon = \pm 1,\ \Delta = B^2 - AC)$$

et, en dérivant :

$$\psi'' = -\frac{AB' - BA'}{A^2} - \varepsilon\frac{A'}{A^2}\sqrt{\Delta} + \frac{\varepsilon}{2A}\ \frac{\Delta'}{\Delta}\sqrt{\Delta}$$

En remplaçant $\varepsilon\frac{\sqrt{\Delta}}{A}$ par $\psi' + \frac{B}{A}$, il vient ensuite :

$$\psi'' = -\frac{B'}{A} + \frac{1}{2}\frac{B}{A}\frac{\Delta'}{\Delta} + \psi' \left(\frac{1}{2}\frac{\Delta'}{\Delta} - \frac{A'}{A}\right)$$

ce que nous écrivons sous la forme :

$$\text{(c)} \qquad A\,\psi'' + \left(A' - \frac{1}{2}\frac{A\Delta'}{\Delta}\right)\psi' + B' - \frac{1}{2}\frac{B\Delta'}{\Delta} = 0$$

Cette équation (c) doit être identique à la première équation (2-4). Or ceci n'est possible que si le discriminant Δ divise $B\Delta'$ et $A\Delta'$.

(Nous laissons de côté le cas trivial où le discriminant Δ serait nul : on a alors $B^2 = AC$ et $\psi' = -B/A = -C/B$: ψ' est alors l'inverse d'un polynome de degré 1. Ce cas trivial correspond pour $G(s)$ à une loi binomiale négative : nous le retrouverons d'ailleurs à plusieurs reprises).

Examinons les différents cas possibles.

Les Cas Possibles.

Cas 1 : Degré $A = 2$, Degré $\Delta = 2$, $B \neq 0$, $C \neq 0$.

Comme Δ divise $B\Delta'$, qui est de degré 2, on a $\Delta \sim B\Delta'$. En particulier, cela implique que Δ' divise Δ, donc que Δ est un carré parfait. Mais, puisque $\Delta \sim B\,\Delta'$, cela entraine $\Delta \sim B^2$ et $\Delta' \sim B$. Or $\Delta = B^2 - AC$. Donc $\Delta \sim B^2$ et $C \neq 0$ entrainent $A \sim B^2$. On retombe sur la même conclusion que pour $\Delta = 0$:

$$\psi' = \frac{\alpha}{B}$$

où α est une constante. Les solutions correspondantes existent ($G_\nu(s)$ binomiale négative).

Cas 2 : $C = 0$.

Dans le cas $C = 0$, le raisonnement précédent est en défaut. De façon générale, pour $C = 0$, la seconde équation (2-4) donne soit $\psi' = 0$, soit

$$\psi' = \frac{2B}{A}$$

c'est-à-dire ψ' égal au rapport d'un polynome de degré ≤ 1 à un polynome de degré ≤ 2. Cette solution convient, et nous verrons qu'il lui correspond effectivement des lois de probabilités : $\Delta = 2B$ ne divise pourtant pas $A\Delta'$, mais un regard sur l'équation (c) montre que le terme $\frac{A\Delta'}{\Delta}\psi' = -4B'$ est constant, de sorte que la condition : Δ divise $A\Delta'$ devient superflue dans ce cas très particulier. En sens inverse, cette anomalie

$$\frac{A\Delta'}{\Delta}\psi' = \alpha$$

où α est une constante, entraine

$$\alpha\frac{\Delta}{\Delta'} = -B + \varepsilon\sqrt{\Delta}$$

ce qui n'est possible que pour $\Delta \sim B^2$: si $C \neq 0$ on retombe sur le cas 1/, si $C = 0$ on retombe sur le cas 2/ , de sorte qu'il n'y a pas d'autre possibilité pour que cette anomalie se produise.

Cas 3 : Degré $A = 2$, Degré $\Delta = 2$, $B = 0$.

Comme B est nul, on doit avoir $C < 0$, soit $C = \beta^2$ pour une constante β et

$$\psi' = \frac{\beta}{\sqrt{A}}$$

Comme $\Delta = \beta^2 A$ divise $A\Delta'$, c'est là une solution possible, et nous verrons qu'il lui correspond effectivement deux familles de lois.

Cas 4 : Degré $A = 2$, Degré $\Delta = 1$, $B \neq 0$.

Ici Δ' est une constante. Comme Δ doit diviser $B\Delta'$, on doit donc avoir $B = \beta\Delta$ avec une constante β. A doit nécessairement contenir Δ en facteur, puisque Δ divise aussi $A\Delta'$. Donc A est de la forme $A = \lambda\Delta$, où λ est un polynome de degré 1. Ecrivant $\Delta = B^2 - AC$, il vient alors $\beta^2\Delta^2 - C\lambda\Delta = \Delta$, donc

$$\beta^2\Delta - C\lambda = 1$$

Posant

$$\mu = 1 + C\lambda$$

(polynome de degré 1 qui peut être d'ailleurs quelconque), il vient

$$\Delta = \frac{\mu}{\beta^2} \qquad \lambda = \frac{\mu-1}{C}$$

$$B = \frac{\mu}{\beta} \qquad A = \frac{\mu(\mu-1)}{C\beta^2}$$

On trouve alors :

$$\psi' = \frac{C\beta}{\mu-1}\left(1 + \frac{\varepsilon}{\sqrt{\mu}}\right) = \frac{C\beta}{\sqrt{\mu}}\ \frac{1}{\sqrt{\mu} - \varepsilon}$$

Cette solution convient effectivement, et nous verrons qu'il lui correspond effectivement deux nouvelles familles de lois à trois paramètres.

Les cas restants peuvent être examinés plus rapidement.

Cas 5 : Deg $A = 2$, Deg $\Delta = 1$, $B = 0$.

Ceci est impossible, puisque, pour $B = 0$, $\Delta = - AC$ est de degré 2 ou 0 selon que C est $\neq 0$ ou $= 0$.

Cas 6 : Deg $A = 2$, Deg $\Delta = 0$ et $B \neq 0$.

Δ étant constant, il existe une constante $\alpha \geq 0$ telle que

$$B^2 = \alpha + AC \text{ , soit } C \neq 0 \text{ et}$$

$$A = \frac{B^2-\alpha}{C} \text{ et finalement :}$$

$$\psi' = - \frac{CB}{B^2-\alpha} + \varepsilon\frac{C\sqrt{\alpha}}{B^2-\alpha} = - \frac{C}{B + \varepsilon\sqrt{\alpha}}$$

A nouveau, nous rencontrons le cas où ψ' est l'inverse d'un polynome du 1er degré, cas auquel correspondent des <u>lois binomiales négatives</u>.

Cas 7 : Deg $A = 1$, Deg $\Delta = 2$, $B \neq 0$.

(la même configuration avec $B = 0$ est impossible). Comme Δ divise $B\Delta'$ on a $\Delta \sim B\Delta'$ et, Δ divisant $A\Delta'$ avec ici A de degré 1, on a également $\Delta \sim A\Delta'$: il en résulte que A et B sont proportionnels $A \sim B$, et $\Delta \sim A^2 \sim B^2$. On a donc

$$\psi' = C^{ste}$$

Cette solution convient effectivement : il lui correspond les lois de Poisson.

Cas 8 : Deg $A = 1$, Deg $\Delta = 1$.

Comme $\Delta = B^2 - AC$, ceci n'est possible que si $B = \beta = C^{ste}$. Mais Δ doit diviser $B\Delta' = \beta\Delta'$: comme Δ' est une constante non nulle, il en résulte $\beta = 0$ c'est-à-dire $B = 0$. Alors $\Delta = - AC$ divise $A\Delta'$: la solution convient. On trouve

$$\psi' = \frac{\varepsilon\sqrt{-C}}{\sqrt{A}}$$

et il est favile de voir qu'il lui correspond la loi de Sichel.

Cas 9 : Deg $A = 1$, Deg $\Delta = 0$.

Comme $\Delta = B^2 - AC$, ceci entraine $B = C = 0$, et il reste $\psi' = 0$.

Cas 10 : Deg $A = 0$, Deg $\Delta = 2$.

Comme Δ doit diviser $A\Delta'$, on doit avoir $A = 0$, et il reste $\psi' = - C/2B$: à nouveau la binomiale négative.

Cas 11 : Deg $A = 0$, Deg $\Delta = 1$: Impossible.

Cas 12 : Deg $A = 0$, Deg $\Delta = 0$.

A et Δ étant constantes, B est également constante. On retrouve $\psi' = C^{ste}$, c'est-à-dire la loi de Poisson.

En résumé, trois cas seulement vont mériter un examen plus détaillé : le cas "rationnel" 2/, correspondant à C = 0 et ψ' = - 2B/A, ainsi que les deux cas 3/ et 4/ qui sont de loin les plus intéressants.

3 - LE CAS RATIONNEL.

Examinons en premier lieu le cas "rationnel", c'est-à-dire le cas où C = 0 et

$$\psi'(s) = \frac{P_1(s)}{P_2(s)}$$

$P_1(s)$, $P_2(s)$ = polynomes de degrés ≤ 1 et ≤ 2 respectivement.

3-a Cas où Deg. $P_2 = 2$ effectivement.

La discussion se subdivise en trois cas, selon que les racines de P_2 sont réelles et distinctes, réelles et égales, ou complexes conjuguées.

Un examen ennuyeux que je ne voudrais pas infliger aux lecteurs montre que le cas des racines complexes conjuguées ne saurait conduire à aucune loi de probabilité.

Supposons qu'il existe deux racines réelles s_o et s_1 distinctes. Comme $\psi'(s)$ doit admettre un développement en série entière convergeant pour $s < 1$, on a nécessairement $|s_1|$ et $|s_o| \geq 1$ (nous verrons dans un instant que l'égalité stricte est impossible).

Supposons, par exemple $|s_1| > |s_o|$, et posons

$$1/|s_o| = p \qquad 1/|s_1| = \alpha \qquad \text{avec } \alpha < p$$

Alors $\psi'(s)$ se met sous la forme :

$$\psi'(s) = \frac{A_o}{1-\varepsilon_1\, ps} + \frac{A_1}{1-\varepsilon_2\, \alpha s}$$

($\varepsilon_1 = \pm 1$, $\varepsilon_2 = \pm 1$ selon le signe de s_o et s_1 , et :

$$\alpha < p < 1$$

Si les deux racines s_o et s_1 sont ≥ 0, on a $\varepsilon_1 = \varepsilon_2 = +1$, d'où

$$\psi'(s) = \frac{A_o}{1-ps} + \frac{A_1}{1-\alpha s} = \sum_{n \geq o} (A_o\, p^n + A_1\, \alpha^n)\, s^n$$

Le coefficient de s^n sera ≥ 0 si et seulement si

$$A_o \geq 0 \quad , \quad A_o + A_1 \geq 0$$

En intégrant, il viendra :

$$\psi(s) = \frac{A_o}{p} \log\left(\frac{1-p}{1-ps}\right) + \frac{A_1}{\alpha} \log\left(\frac{1-\alpha}{1-\alpha s}\right)$$

(On voit que le cas $p = 1$ est impossible). Pour alléger les notations, prenons $A_o = p$, $A_1/\alpha = \beta$:

$$\psi(s) = \log\left(\frac{1-p}{1-ps}\right) + \beta \log\left(\frac{1-\alpha}{1-\alpha s}\right)$$

$$p + \alpha\beta \geq 0$$

On obtient la famille <u>à quatre paramètres</u> :

$$(3\text{-}1) \qquad \boxed{G_\nu(s) = \left(\frac{1-p}{1-ps}\right)^{\nu} \left(\frac{1-\alpha}{1-\alpha s}\right)^{\nu\beta}}$$

avec $\nu \geq 0$, $0 < \alpha < p < 1$ $\quad p + \alpha\beta \geq 0$.

Pour $\beta > 0$, la variable correspondante est somme de deux binomiales négatives indépendantes. Pour β négatif (mais $\geq -\frac{p}{\alpha}$), on n'a plus d'interprètation aussi simple.

Examinons le cas où l'une au moins des racines s_o et s_1 est négative. Le développement :

$$\psi'(s) = \sum_{n \geq o} (A_o(\varepsilon_1)^n\, p^n + A_1\, (\varepsilon_2)^n\, \alpha^n)\, s^n$$

montre que l'on doit nécessairement avoir $A_o > 0$ et $\varepsilon_1 = +1$, (le premier terme, en effet, subsiste seul asymptotiquement pour $n \to \infty$). On a donc :

$$\psi'(s) = \frac{A_0}{1-ps} + \frac{A_1}{1+\alpha s} = \Sigma \left(A_0\, p^n + (-1)^n A_1\, \alpha^n\right) s^n$$

En faisant $s = 0$, on trouve la condition

$$A_0 + A_1 \geq 0$$

Avec $n = 1$, il vient

$$A_0' \, p - A_1 \, \alpha \geq 0$$

Ces conditions (jointes à $0 < \alpha < p < 1$) suffisent, inversement, pour garantir la positivité du coefficient de s^n.

En intégrant, il vient :

$$\psi(s) = \frac{A_0}{p} \log \frac{1-p}{1-ps} + \frac{A_1}{\alpha} \log \frac{1+\alpha s}{1+\alpha}$$

Prenons ici encore $A_0 = p$ et $A_1/\alpha = \beta$. Sous les conditions

$$\begin{cases} 0 < \alpha < p < 1 \\ p + \alpha\beta \geq 0 \\ p^2 - \alpha^2\beta \geq 0 \end{cases}$$

et, bien sûr, $\nu \geq 0$, on voit que la fonction

$$\boxed{G_\nu(s) = \left(\frac{1-p}{1-ps}\right)^\nu \left(\frac{1+\alpha s}{1+\alpha}\right)^{\beta\nu}} \qquad (3\text{-}2)$$

est effectivement la génératrice d'une loi I.D. à <u>quatre paramètres</u>. Le cas d'une <u>racine double</u>, enfin, est immédiat. $\psi'(s)$ est obligatoirement de la forme

$$\psi'(s) = \frac{a}{1-ps} + \frac{b}{(1-ps)^2}$$

avec $0 < p < 1$, $b \geq 0$, $a + b \geq 0$.

En intégrant, on trouve à un facteur près (égal à $1/p$)

$$\psi(s) = a \log \frac{1-p}{1-ps} + \frac{b}{1-ps} - \frac{b}{q}$$

D'où la famille de lois à 3 paramètres.:

$$\boxed{G_\nu(s) = \left(\frac{q}{1-ps}\right)^{a\nu} e^{\nu \frac{b}{q} \left(\frac{q}{1-ps} - 1\right)}} \qquad (3\text{-}3)$$

avec les conditions $\nu > 0$ et

$$0 < p < 1 \qquad p + q = 1$$

$$b \geq 0 \qquad a + b = 1$$

Pour $a > 0$, il s'agit de la somme d'une binomiale négative et d'une binomiale négative poissonisée. Pour $a < 0$, il n'y a plus d'interprétation aussi simple.

3-b Cas où P_2 est de degré 1.

Dans ce cas, ψ' est de la forme

$$\psi' = a + \frac{b}{1-ps}$$

avec $|p| \leq 1$, puisque $\psi'(s)$ doit se développer en série convergente pour $s < 1$. En développant ψ', on trouve les conditions suivantes :

$$b \geq 0 \qquad p \geq 0 \qquad a + b \geq 0$$

nécessaires et suffisantes pour que le coefficient de s^n soit ≥ 0. En intégrant, il vient

$$\psi(s) = a(s-1) + \frac{b}{p} \log \frac{1-p}{1-ps}$$

En prenant $b > 0$, on peut supposer $b = p$. Posant $q = 1-p$, on voit que sous les conditions

$$\begin{cases} 0 \leq p < 1 \\ a + p \geq 0 \end{cases}$$

la famille à 3 paramètres :

(3-4) $$\boxed{G_\nu(s) = \left(\frac{q}{1-ps}\right)^\nu e^{-a\nu(1-s)}}$$

convient. Pour $a > 0$, la variable correspondante est la somme d'une binomiale négative et d'une poissonienne indépendantes. Pour $a < 0$ (mais $aq + p^2 \geq 0$), on n'a plus d'interprétation aussi simple.

3-b Cas où P_2 est une constante.

Dans ce cas :

$$\psi'(s) = b + 2\ as$$

avec nécessairement

$$a \geq 0 \quad b \geq 0$$

En intégrant :

$$\psi(s) = b(s-1) + a(s^2-1)$$

En posant $a + b = 1$, $a = p$, $b = q$, on obtient la famille à deux paramètres :

(3-5) $$\boxed{G_\nu(s) = e^{-\nu\ p(1-s)\ -\ \nu q(1-s^2)}}$$

avec $$0 \leq p \leq 1 \quad , \quad p + q = 1$$

Il s'agit de la somme d'une poissonienne et d'un doublet poissonien.

Mélanges Poissoniens.

Parmi ces lois, lesquelles correspondent à des mélanges poissoniens ? Pour le voir, il faut examiner si

$$\Phi_\nu(\lambda) = G_\nu(1-\lambda)$$

est une transformée de Laplace. D'après les théorèmes généraux, il suffit même de s'assurer que

$$\psi'(1-\lambda)$$

est complètement monotone (i.e. transformée d'une mesure positive non nécessairement bornée).

D'entrée de jeu, le cas des deux racines de signes opposées, c'est-à-dire le cas (3-2), est exclu ($\psi'(1-\lambda)$ devient infini pour une valeur > 0 de λ). Donc : <u>la loi (3-2) n'est pas un mélange poissonien</u>.

En ce qui concerne la loi (3-1), il est immédiat qu'il s'agit <u>effectivement d'un mélange poissonien</u> lorsque β est ≥ 0 : $\Phi_\nu(\lambda)$ représente, en effet, la somme de deux variables gamma. Pour $\beta < 0$ un examen attentif montre que $\Phi_\nu(\lambda)$ est encore une transformée de Laplace, mais sous la condition

$$1 + \beta \geq 0$$

un peu plus sévère que la condition $p + \alpha\beta \geq 0$ requise pour G_ν.

<u>La loi (3-3) est toujours un mélange poissonien pour $a \geq 0$ (mais non pour $a < 0$)</u>.

En ce qui concerne (3-4), la réponse est également positive si $a \geq 0$: Φ_ν représente une translatée (par une translation $a \geq 0$) d'une loi gamma. Pour $a < 0$, la variable continue prend des valeurs négatives, et on ne peut plus parler de mélange poissonien.

Enfin (3-5) n'est un mélange poissonien que dans le cas trivial $q = 0$, où G_ν se réduit à une loi de Poisson : Φ_ν représente alors un simple Dirac.

<u>Propriétés de ces Lois</u>.

En dérivant la fonction génératrice (3-1), on trouve

$$\left\{\begin{array}{l} G' = \nu\left(\dfrac{p}{1-ps} + \dfrac{\alpha\beta}{1-\alpha s}\right) G \\[2ex] G'' = \nu^2\left(\dfrac{p}{1-ps} + \dfrac{\alpha\beta}{1-\alpha s}\right)^2 G + \nu\left(\dfrac{p^2}{(1-ps)^2} + \dfrac{\nu\beta\ \alpha^2}{(1-\alpha s)^2}\right) G \end{array}\right.$$

D'où la <u>moyenne</u> et la <u>variance</u>

$$\left\{\begin{array}{l} m = \dfrac{\nu\ p}{1-p} + \dfrac{\nu\alpha\beta}{1-\alpha} \\[2ex] \sigma^2 = m + \nu\left(\left(\dfrac{p}{1-p}\right)^2 + \beta\left(\dfrac{\alpha}{1-\alpha}\right)^2\right) \end{array}\right.$$

ainsi que les <u>deux premières probabilités</u>

$$\left\{\begin{array}{l} p_o = (1-p)^{\nu}\ (1-\alpha)^{\nu\beta} \\[1ex] p_1 = \nu\ (p+\alpha\beta)\ p_o \end{array}\right.$$

Pour établir l'équation différentielle, nous noterons d'une manière générale que, pour

$$\psi' = \frac{B}{A}$$

on trouve pour la fonction $G_\nu = \exp(\nu\psi)$:

$$A\ G'' + (A'-\nu B)\ \psi'' - \nu B' = 0$$

Ici

$$A = (1-ps)(1-\alpha s)$$

$$B = p(1-\alpha s) + \alpha\beta(1-ps) = p + \alpha\beta - \alpha\ p(1+\beta)s$$

Donc

$$A' = 2\ \alpha ps - (\alpha+p)$$

$$B' = -\ \alpha\ p(1+\beta)$$

et <u>l'équation différentielle</u> de (3-1) s'écrit

$$(3\text{-}1')\qquad (1-ps)(1-\alpha s)\ G'' + [2\ \alpha ps - (\alpha+p) - \nu(p+\alpha\beta) + \nu\alpha p(1+\beta)s]\ G' + \nu a p(1+\beta)\ G$$

et on en déduit la <u>relation de récurrence</u>

$$(3\text{-}1') \qquad (n+1)(n+2)\, p_{n+2} = (n+1)\,[(n+1)(\alpha+p) + \nu(p+\alpha\beta)]\, p_{n+1}$$
$$- \alpha p\,[2n + \nu(n+1)(1+\beta)]\, p_n$$

En ce qui concerne la loi (3-2), les formules sont exactement les mêmes, à condition de remplacer α par $-\alpha$ et β par $-\beta$. Il est donc inutile de les écrire explicitement.

Passons maintenant au cas de la <u>racine double</u>. En dérivant (3-3), il vient :

$$G' = \left(\frac{a\nu p}{1-ps} + \frac{\nu bp}{(1-ps)^2}\right) G$$

$$G'' = \left(\frac{a\nu p}{1-ps} + \frac{\nu bp}{(1-ps)^2}\right)^2 G + \left(\frac{a\nu p^2}{(1-ps)^2} + \frac{2\,\nu bp^3}{(1-ps)^3}\right) G$$

D'où la <u>moyenne et la variance</u>

$$\begin{cases} m = \nu\, p\left(\dfrac{a}{q} + \dfrac{b}{q^2}\right) \\ \sigma^2 = m + \nu p^2\left(\dfrac{a}{q^2} + \dfrac{2b}{q^3}\right) \end{cases}$$

etles deux <u>premières probabilités</u> (avec a + b = 1)

$$\begin{cases} p_o = q^{a\nu}\, e^{-\frac{\nu bp}{q}} \\ p_1 = a\,\nu\, p\, p_o \end{cases}$$

Avec $A = (1-ps)^2$ et $B = p(1-aps)$, <u>l'équation différentielle</u> s'écrit :

$$(3\text{-}3') \qquad (1-ps)^2\, G'' + [2\,p^2 s - 2p - \nu p + \nu a\, p^2 s]\, G' + \nu a\, p^2\, G = 0$$

et on en déduit les relations de récurrence

$$(3\text{-}3'') \qquad (n+2)(n+1)p_{n+2} = 2\,p(n+1)(n+1 + \frac{\nu}{2})p_{n+1} - p^2[n(n+1+a\nu) + a\nu]p_n$$

Examinons maintenant le cas de <u>la loi (3-4)</u>. En dérivant on trouve

$$\begin{cases} G' = \nu \left(a + \frac{p}{1-ps}\right) G \\ G'' = \nu^2 \left(a + \frac{p}{1-ps}\right)^2 G + \frac{\nu\, p^2}{(1-ps)^2} \; G \end{cases}$$

D'où la <u>moyenne</u> et la <u>variance</u> de cette loi

$$\begin{cases} m = \nu\left(a + \frac{p}{q}\right) \\ \sigma^2 = m + \nu \frac{p^2}{q^2} \end{cases}$$

et les <u>deux premières probabilités</u>

$$\begin{cases} p_o = q^\nu \, e^{-a\nu} \\ p_1 = \nu(a+p) \; p_o \end{cases}$$

Pour <u>l'équation différentielle</u>, on trouve

$$(3\text{-}4') \qquad (1-ps)\, G'' - [\nu(a+p - aps) + p]\, G' + \nu ap\, G = 0$$

et on en tire la relation de récurrence

$$(3\text{-}4'') \quad (n+1)(n+2)p_{n+2} = (n+1)[p(n+1) + \nu(a+p)]p_{n+1} - ap\nu(n+1)p_n$$

Enfin, dans le cas de <u>la loi (3-5)</u>, on trouve :

$$\begin{cases} m = \nu(1+q) \\ \sigma^2 = m + 2\,\nu q \end{cases}$$

$$\begin{cases} p_o = e^{-\nu} \\ p_1 = \nu \; p \; e^{-\nu} \end{cases}$$

<u>L'équation différentielle</u> est :

$$G'' - \nu(p + 2\, qs)\, G' - 2\, \nu q\, G = 0$$

et la relation de récurrence prend la forme très simple

$$(n+2)p_{n+2} = \nu\ p\ p_{n+1} + 2\ \nu q\ p_n \qquad (3\text{-}4''')$$

4 - LES LOIS DE SCHIEL ET DE LEISCH.

Venons-en maintenant au premier groupe de lois à longue queue, je veux dire le cas 3/ de l'inventaire du paragraphe 2, i.e. le cas B = 0 et degré A = 2. Dans ce cas, ψ' est de la forme

$$\psi' = \frac{C}{\sqrt{P_2}}$$

où C est une constante et P_2 un polynome de second degré en s. Comme d'habitude, il convient de distinguer : 2 racines réelles, une racine double, deux racines complexes conjuguées. Pour abréger la discussion, nous pouvons heureusement éliminer d'entrée de jeu le cas des deux racines complexes conjuguées : en effet, si $\psi'(s)$ admet un développement en série convergeant pour $|s| < 1$, tel que le coefficient de s^n soit ≥ 0, ψ'^2 possède la même propriété. Or nous avons vu au paragraphe précédent que $1/P_2(s)$ ne pouvait pas vérifier cette propriété dans le cas des racines complexes conjuguées. Nous pouvons aussi régler son sort au cas de la racine double : si P_2 est un carré parfait, ψ' est du type $1/1-ps$, déjà maintes fois rencontré, qui conduit pour G_ν à la loi binomiale négative.

Reste le cas où $P_2(s)$ possède deux racines distinctes. Comme $P_2(s)$ doit être > 0 sur le domaine $0 \leq s < 1$, il convient de distinguer deux cas selon le signe du coefficient de s^2. Ou, si l'on préfère, en désignant par

$$\mu(s) = as + b$$

un polynome du premier degré en s, il existe deux cas possibles : ou bien

$$\psi'(s) = \frac{1}{\sqrt{\mu^2(s)-1}} \quad \text{avec} \quad \mu^2(s) > 1 \quad \text{pour} \quad 0 \leq s < 1$$

ou bien

$$\psi'(s) = \frac{1}{\sqrt{1-\mu^2(s)}} \quad \text{avec} \quad \mu^2(s) < 1 \quad \text{pour} \quad 0 \leq s < 1$$

Le premier cas va nous conduire à des lois du type Arg Sh , ou lois de Schiel, le second à des lois de type Arc sin ou lois de Leisch.

4-a La Loi de SCHIEL, ou loi en Arg Sh.

Partons de l'expression

$$\psi'(s) = \frac{1}{\sqrt{(b-as)^2-1}}$$

où nous pouvons toujours supposer $b \geq 0$. On doit avoir $(b-as)^2 > 1$ pour $0 \leq s < 1$. Donc, en fait, ayant choisi $b \geq 0$, il faut :

$$b > 1 \quad \text{et} \quad B-a \geq 1$$

Comme, de plus, ψ' doit être une fonction croissante de s, il faut également avoir

$$a > 0$$

Il est commode de changer les notations, en posant

$$\beta = \frac{1}{b} \quad , \quad \frac{a}{b} = \alpha$$

Les conditions précédentes s'écrivent

$$\alpha > 0 \quad , \quad \beta \geq 0 \quad , \text{et } \alpha + \beta \leq 1$$

et (à un facteur positif près) il vient :

$$\psi'(s) = \frac{1}{\sqrt{(1-\alpha s)^2-\beta^2}}$$

Mis sous la forme :

$$\psi'(s) = \frac{1}{(1-\alpha s)} \frac{1}{\sqrt{1-\frac{\beta^2}{(1-\alpha s)^2}}}$$

$\psi'(s)$ se développe alors en série du type

$$\psi'(s) = \sum \frac{a_n}{(1-\alpha s)^{2n+1}}$$

avec des coefficients > 0 : comme $1/(1-\alpha s)$ représente une loi binomiale négative, on est assuré que le développement de $\psi'(s)$ en série entière en s est à coefficients positifs.

En intégrant $\psi'(s)$, on trouve (au facteur $1/\alpha$ près)

$$\psi(s) = \log \frac{1-\alpha+\sqrt{(1-\alpha)^2-\beta^2}}{1-\alpha s+\sqrt{(1-\alpha s)^2-\beta^2}} + 1$$

qui est une loi, puisque $\psi(1) = 1$, pourvu seulement que $\alpha < 1$ strictement.

Ainsi, sous les conditions :

$$0 < \alpha < 1 \ , \ \beta \geq 0 \ , \ \alpha + \beta \leq 1$$

l'expression suivante constitue la fonction génératrice d'une famille à 3 paramètres de lois indéfiniment divisibles :

$$\boxed{G_\nu(s) = \left(\frac{1-\alpha+\sqrt{(1-\alpha)^2-\beta^2}}{1-\alpha s+\sqrt{(1-\alpha s)^2-\beta^2}}\right)^\nu} \qquad (4-1)$$

C'est la loi de SCHIEL, ou loi en Arg Sh.

On peut en obtenir une interprétation plus parlante en posant

$$\theta = \left(\frac{\beta}{1-\alpha}\right)^2 \qquad (0 \leq \theta \leq 1)$$

et en écrivant :

$$G_\nu(s) = \left(\frac{1-\alpha}{1-\alpha s}\right)^\nu \left(\frac{1+\sqrt{1-\theta}}{1+\sqrt{1-\theta\,(\frac{1-\alpha}{1-\alpha s})^2}}\right)^\nu$$

Nous souvenant de l'expression :

$$g_\nu(s) = \left(\frac{1+\sqrt{1-\theta}}{1+\sqrt{1-\theta s}}\right)^\nu$$

rencontrée au premier paragraphe (cette fonction hypergéométrique réduite est la fonction génératrice d'une loi I.D.), il vient :

$$G_\nu(s) = \left(\frac{1-\alpha}{1-\alpha s}\right)^\nu g_\nu\left(\frac{1-\alpha}{1-\alpha s}\right)$$

Ainsi, G_ν représente la somme d'une binomiale négative $((1-\alpha)/(1-\alpha s))^\nu$ et d'une binomiale négative randomisée par la loi g_ν. En nous reportant à (1-12'), nous obtenons l'expression explicite de cette randomisation :

$$(4\text{-}1') \qquad G_\nu(s) = \nu\left(\frac{1+\sqrt{1-\theta}}{2}\right)^\nu \sum_{n=0}^{\infty} \frac{\Gamma(\nu+2n)}{n!\,\Gamma(\nu+n+1)} \left(\frac{\theta}{4}\right)^n \left(\frac{1-\alpha}{1-\alpha s}\right)^{n+\nu}$$

Pour obtenir la moyenne et la variance, calculons les deux premières dérivées de (4-1) :

$$G' = \frac{\alpha\nu}{\sqrt{(1-\alpha s)^2-\beta^2}}\;G$$

$$G'' = \frac{\alpha^2\nu^2}{(1-\alpha s)^2-\beta^2}\;G + \frac{\alpha^2\nu(1-\alpha s)}{[(1-\alpha s)^2-\beta^2]^{3/2}}\;G$$

Faisant s = 1, nous obtenons <u>moyenne</u> et <u>variance</u> :

$$\begin{cases} m = \dfrac{\alpha\,\nu}{\sqrt{(1-a)^2-\beta^2}} \\[2ex] \sigma^2 = m + \dfrac{\alpha^2\,(1-\alpha)\nu}{[(1-\alpha)^2-\beta^2]^{3/2}} \end{cases}$$

Faisant, au contraire, s = 0 dans G et G', nous obtenons les expressions des <u>deux premières probabilités</u> p_o et p_1

$$\begin{cases} p_o = \left(\dfrac{1-\alpha + \sqrt{(1-\alpha)^2 - \beta^2}}{1 + \sqrt{1 - \beta^2}} \right)^\nu \\ p_1 = \dfrac{\alpha \nu}{\sqrt{1-\beta^2}} \quad p_o \end{cases}$$

Pour obtenir les p_n suivantes, on devra utiliser les relations de récurrence (dont l'existence est assurée par construction). De fait, d'après les expressions mêmes de G' et G" calculées ci-dessus, il apparaît que G est solution de <u>l'équation différentielle</u>

$$(4\text{-}2) \qquad [(1-\alpha s)^2 - \beta^2]\, G'' - \alpha(1-\alpha s)\, G' - \alpha^2 \nu^2\, G = 0$$

De cette équation différentielle résultent les <u>relations de récurrence</u> à trois termes :

$$(4\text{-}3) \qquad (1-\beta^2)(n+2)(n+1)p_{n+2} = \alpha(n+1)(2\,n+1)p_{n+1} + \alpha^2(\nu^2-n^2)p_n$$

qui, compte tenu des expressions de p_o et de p_1, permettent le calcul numérique de proche en proche de tous les p_n .

<u>Exemples d'Ajustements</u>.

On trouvera ci-joint deux exemples d'ajustements à la loi de Schiel. Le premier concerne des données de Sichel : l'ajustement à la loi de Sichel (fait par Sichel lui-même) est déjà excellent, mais la loi de Schiel donne un résultat encore meilleur : toute la différence provient des deux premières classes. Avec la loi de Sichel à deux paramètres, on n'arrive pas à reproduire de manière satisfaisante le rapport p_o/p_1 = 41/19 ≃ 2, alors que cela se fait sans difficulté avec la loi de Schiel à 3 paramètres.

Le second ajustement concerne une population d'effectif très élevé (N = 12 017). Le χ^2 obtenu (56,88) conduirait en principe à un rejet ($P(\chi^2 \geq 56,88) = 1/100$), bien qu'en fait il s'agisse d'un résultat déjà très bon pour un effectif aussi nombreux. En réalité l'ajustement est extrêmement bon jusqu'à 30, et bon jusqu'à 100. Seules les deux dernières classes (100-200 et > 200) refusent de s'ajuster. Il suffit d'ailleurs de les regrouper pour obtenir un très bon χ^2 ($P(\chi^2 \geq 36,59) = 35/100$). Seule donc l'extrême queue de la distribution n'est pas rendue de manière satisfaisante (34 échantillons sur 12 017, mais ces 34 échantillons, soit 0,3% de l'effectif, contiennent 25% du nombre total des pierres !)

Loi de Sichel

Classes	Fréquence observée	Fréquence calculée	χ^2
0	41	36,0	0,694
1	19	26,8	2,270
2	16	16,2	0,002
3	8	10,1	0,437
4	6	6,6	0,054
5-6	11	8,0	1,125
7-8-9	6	6,0	0,000
≥ 10	11	8,3	0,878
	118		$\chi^2 = 5,46$

Avec 5 degrés de liberté :

$P(\chi^2 \geq 5,46) = 36/100$

Loi de Schiel

0	41	42,8	0,08
1	19	19,8	0,03
2	16	12,9	0,72
3	8	9,2	0,17
4	6	6,9	0,12
5-6	11	9,4	0,26
7-8-9	6	7,9	0,44
≥ 10	11	9,0	0,46
	118		$\chi^2 = 2,30$

Avec 4 degrés de liberté :

$P(\chi^2 \geq 2,30) = 68/100$

Ajustement à la Loi de SCHIEL

Classe	Fréquence observée	Fréquence calculée	χ^2
0	7603	7607	0,00
1	1376	1377	0,00
2	722	699	0,77
3	440	436	0,03
4	294	301	0,14
5	233	220	0,79
6	198	168	5,46
7	137	132	0,17
8	109	107	0,04
9	90	88	0,03
10	76	74	0,05
11	71	63	0,96
12	42	54,6	2,90
13	53	47,7	0,60
14	41	42,0	0,02
15	37	37,3	0,00
16	37	33,4	0,39
17	29	30,1	0,04
18	30	27,2	0,28
19	22	24,8	0,31
20	16	22,7	1,95
21	23	20,8	0,23
22	18	19,2	0,07
23	22	17,7	1,04
24	17	16,4	0,02
25	13	15,3	0,34
26	9	14,2	1,93
27	10	13,3	0,82
28	16	12,4	1,01
29	9	11,7	0,62
30	9	11,0	0,36
31-40	66	81,8	3,06
41-50	40	50,5	2,19
51-60	30	33,8	0,42
61-70	13	23,8	4,88
71-80	17	17,4	0,01
81-100	15	23,0	2,79
101-200	17	35,5	9,65
≥ 201	17	7,4	12,49
			$\chi^2 = 56,88$

N = 12 017 données

Ajustement sur p_0, p_1 et m.

$\alpha = 0,805$

$\beta = 0,187052325$

$\nu = 0,220853248$

	Moyenne	Variance
Obs.	3,2264	426,57
Calc.	3,2264	758,46

Avec 35 degrés de liberté

$P(\chi^2 \geq 56,88) = 1/100$

<u>NB</u>. En regroupant les deux dernières classes, le χ^2 tombe à

$$\chi^2 = 36,59$$

Avec 34 degrés de liberté

$P(\chi^2 \geq 36,59) = 35/100$

4-b La Loi de LEISCH, ou loi en Arc Sinus.

Nous considérons maintenant le cas où ψ' est de la forme

$$\psi'(s) = \frac{1}{\sqrt{1-(as+b)^2}}$$

où nous pouvons toujours choisir de prendre $b \geq 0$. Comme on doit avoir $(as+b)^2 < 1$ pour $0 \leq s < 1$, il faut $b < 1$ et $b+a \leq 1$. D'ailleurs, $\psi'(s)$ devant être croissante, a est nécessairement positif. Ainsi :

$$a > 0 \ , \quad b \geq 0 \quad \text{et} \quad a+b \leq 1$$

sont des conditions nécessaires. Dans le cas $a = 1$, $b = 0$, on obtient $1/\sqrt{1-s^2}$, dont le développement en puissances de s admet bien des coefficients ≥ 0. Mais c'est là une propriété qui subsiste lorsque l'on change s en $as + b$ avec a et $b \geq 0$ et $a + b \leq 1$. Donc $\psi'(s)$, sous les conditions ci-dessus, admet bien un développement à coefficient positif.

En intégrant, on trouve (au facteur $1/a$ près) :

$$\psi(s) = \text{Arc}\sin(as+b) - \text{Arc}\sin(a+b)$$

$\psi(1)$ étant fini, cette expression convient. Sous les conditions

$$a > 0 \ , \quad b \geq 0 \ , \quad a+b \leq 1$$

on obtient une famille de lois à 3 paramètres, ou loi de Leisch, dont la fonction génératrice est :

$$\boxed{G_\nu(s) = e^{\nu \, \text{Arc}\sin(as+b) \, - \, \nu \, \text{Arc}\sin(a+b)}} \tag{4-4}$$

Il n'est pas facile de donner une interprétation simple de ces lois. Contentons-nous ici des propriétés les plus utiles.

En dérivant (4-4), il vient

$$\begin{cases} G'(s) = \dfrac{a\nu}{\sqrt{1-(as+b)^2}} \; G(s) \\ \\ G''(s) = \dfrac{a^2\,\nu^2}{1-(as+b)^2} \; G + \dfrac{a^2\,\nu(as+b)}{[1-(as+b)^2]^{3/2}} \; G \end{cases}$$

D'où la moyenne et la variance :

$$\begin{cases} m = \dfrac{a\nu}{\sqrt{1-(a+b)^2}} \\ \\ \sigma^2 = m + \dfrac{a^2\,(a+b)\nu}{[1-(a+b)^2]^{3/2}} \end{cases}$$

et les deux premières probabilités

$$\begin{cases} p_o = e^{-\nu(\text{Arc sin}(a+b) - \text{Arc sin } b)} \\ \\ p_1 = \dfrac{\alpha\,\nu}{\sqrt{1-b^2}} \; p_o \end{cases}$$

Des expressions de G' et G", d'autre part, résulte l'équation différentielle

$$(4\text{-}5) \qquad [1 - (as+b)^2]\; G'' - a\,(as+b)\; G' - a^2\,\nu^2\; G = 0$$

d'où l'on déduit les relations de récurrence à 3 termes :

$$(4\text{-}6) \qquad (1-b^2)(n+2)(n+1)p_{n+2} = a\; b(n+1)(2n+1)p_{n+1} + a^2(n^2 + \nu^2)p_n$$

dont l'expression supprime le moindre doute quant à la positivité des p_n .

Mélanges Poissoniens.

La loi de Leisch ne peut pas être un mélange poissonien, du fait que $\psi'(1-\lambda)$ devient infini pour une valeur > 1 de λ. C'est ce qui fait le caractère un peu énigmatique de cette loi.

Au contraire, la loi de Schiel est un mélange poissonien. Cela est visible par lecture directe du développement (4-1') : changeant s en $1-\lambda$, nous trouvons pour $\Phi_\nu(\lambda)$ un mélange de lois gamma (avec des poids qui sont les probabilités ϖ_n de la loi $g_\nu(s)$). La loi continue correspondante admet une densité de la forme :

$$f(x) = A \frac{e^{-Cx}\, I_\nu(Cx\sqrt{\theta})}{x}$$

(avec $C = \frac{1-\alpha}{\alpha}$, I_ν : fonction de Bessel modifiée).

5 - LES LOIS DE SICHEL et de LECHIS.

Il nous reste à examiner le cas 4 de l'inventaire du paragraphe 2, c'est-à-dire le cas où le discriminant Δ est de degré 1 seulement. On a u qu'il lui correspond une fonction ψ' de la forme

$$\psi'(s) = \frac{C}{\mu-1}\left(1 + \frac{\varepsilon}{\sqrt{\mu}}\right) = \frac{C}{\sqrt{\mu}}\,\frac{1}{\sqrt{\mu}-\varepsilon}$$

où $\varepsilon = \pm 1$, et $\mu(s) = b - as$ est un polynome du premier degré.

Examinons d'abord le cas le plus simple, qui est $\varepsilon = -1$, soit

$$\psi'(s) = \frac{C}{\sqrt{\mu}(1+\sqrt{\mu})}$$

On doit avoir $\mu(s) > 0$ pour $0 \leq s < 1$, donc

$$b > 0 \qquad b-a \geq 0$$

D'ailleurs, ψ' devant être positive et croissante en s, a et C sont nécessairement > 0. Soit, à un facteur près

$$\psi'(s) = \frac{2/a}{\sqrt{b-as}\,(1+\sqrt{b-as})}$$

Ceci s'intègre sous la forme

$$\psi(s) = \log\left(\frac{1+\sqrt{b-a}}{1+\sqrt{b-as}}\right)$$

d'où

$$G_\nu(s) = e^{\nu\psi} = \left(\frac{1 + \sqrt{b-a}}{1 + \sqrt{b-as}}\right)^\nu$$

Compte tenu des conditions

$$a > 0 \ , \ b - a \geq 0$$

ceci s'identifie à l'expression (1-13) écrite avec un coefficient a (celui de la relation (1-13)) positif : il s'agit bien d'une famille de lois, ou plutôt d'une partie seulement de la famille (1-13).

Venons-en au cas plus délicat où $\varepsilon = -1$. Il existe deux possibilités

$$\psi'(s) = \frac{1}{\sqrt{\mu}\ (\sqrt{\mu}-1)} \qquad \text{avec } \mu(s) > 1 \text{ pour } 0 \leq s < 1$$

ou bien

$$\psi'(s) = \frac{1}{\sqrt{\mu}\ (1-\sqrt{\mu})} \qquad \text{avec } \mu(s) < 1 \text{ pour } 0 \leq s < 1$$

Considérons d'abord le premier cas, soit :

$$\psi'(s) = \frac{1}{\sqrt{b-as}} \ \frac{1}{(\sqrt{b-as} - 1)}$$

avec $a > 0$ (puisque ψ' doit être croissante) et $b-a > 1$. Ceci s'intègre dans les mêmes conditions que précédemment, et donne à un facteur près

$$\psi(s) = \log\left(\frac{\sqrt{b-a} - 1}{\sqrt{b-as} - 1}\right)$$

D'où

$$G_\nu(s) = \left(\frac{\sqrt{b-a} - 1}{\sqrt{b-as} - 1}\right)^\nu$$

Compte tenu des conditions

$$a > 0 \ , \quad b-a > 1$$

on reconnait l'expression (1-13), écrite cette fois avec son coefficient (celui de (1-13)) $a < 0$: il s'agit bien d'une fonction génératrice, et, plus précisément, le présent cas et le précédent peuvent être regroupés dans une famille unique, celle qui est définie en (1-13) :

5-a La Loi de Sichel (généralisée).

Réécrivons l'expression de cette loi :

(5-1)
$$\boxed{G_\nu(s) = \left(\frac{a + \sqrt{1-\theta}}{a + \sqrt{1-\theta s}}\right)^\nu}$$

avec $\nu > 0$, $0 < \theta \leq 1$ et $a > -\sqrt{1-\theta}$.

Nous conviendrons (ce qui semble justifié) de donner à cette loi le nom de loi de Sichel généralisée : il s'agit bien, en effet, d'une généralisation élémentaire de la loi de Sichel sensu stricto, qui est la loi (1-4).

Examinons rapidement les propriétés élémentaires de cette loi. En dérivant deux fois la relation (5-1), nous obtenons :

$$\left\{\begin{aligned} G'(s) &= \frac{\nu\ \theta}{2\sqrt{1-\theta s}}\ \frac{1}{a+\sqrt{1-\theta s}}\ G \\ G''(s) &= \frac{\nu^2\ \theta^2}{4(1-\theta s)\ [a+\sqrt{1-\theta s}]^2}\ G + \frac{\nu\ \theta^2}{4(1-\theta s)^{3/2}}\ \frac{G}{a+\sqrt{1-\theta s}} \\ &\quad + \frac{\nu\ \theta^2}{4(1-\theta s)}\ \frac{G}{(a+\sqrt{1-\theta s})^2} \end{aligned}\right.$$

On en déduit la <u>moyenne</u> et la <u>variance</u> :

$$\left\{\begin{aligned} m &= \frac{\nu\ \theta}{2\sqrt{1-\theta}\ (a+\sqrt{1-\theta})} \\ \sigma^2 &= m\left(1 + \frac{\theta}{2(1-\theta)} + \frac{\theta}{2[1-\theta + a\sqrt{1-\theta}]}\right) \end{aligned}\right.$$

Pour <u>les deux premières probabilités</u>, il vient :

$$\begin{cases} p_0 = \left(\dfrac{a + \sqrt{1-\theta}}{a+1}\right)^{\nu} \\[2ex] p_1 = \dfrac{\nu\,\theta}{2(1+a)}\, p_0 \end{cases}$$

D'autre part, en éliminant $\sqrt{1-\theta s}$ entre les deux équations ci-dessus donnant G' et G", on obtient <u>l'équation différentielle</u> :

$$(5\text{-}2) \qquad (a^2-1+\theta s)(1-\theta s)\, G'' + \theta\left[\nu + \frac{3}{2} - \frac{a^2}{2} - \left(\nu + \frac{3}{2}\right)\theta s\right] G' - \theta^2\, \frac{\nu(\nu+1)}{4}\, G = 0$$

d'où l'on déduit les <u>relations de récurrence</u> :

$$(5\text{-}3) \quad \begin{cases} (a^2-1)(n+2)(n+1)p_{n+2} = -\,\theta(n+1)\left[\nu + \frac{3}{2} - \frac{a}{2} + (2-a^2)n\right]p_{n+1} \\[1ex] \qquad\qquad + \,\theta^2\left[n(n+1) + n\left(\nu + \frac{3}{2}\right) + \frac{\nu(\nu+1)}{4}\right]p_n \end{cases}$$

On notera la dégénérescence qui se produit pour $a = 1$: le terme en p_{n+2} disparait, et il reste une relation de récurrence à deux termes seulement : cette circonstance provient de ce que pour $a = 1$ la fonction G_ν coïncide avec la loi g_ν déjà rencontrée à plusieurs reprises, et qui est une hypergéométrique <u>réduite</u> (voir relation (1-9) et suivantes).

Cette simplification considérable qui se produit dans le cas $a = 1$ constitue aussi, si l'on n'y prend pas garde, un piège assez redoutable pour le calcul numérique. En pratique, pour a voisin de 1, le calcul numérique effectif des p_n au moyen de la relation (5-3) risque de devenir très difficile : (pour a^2 nettement différent de 1, il n'y a pas de difficulté sur le plan numérique).

5-b La Loi de LECHIS.

Il reste à examiner le dernier cas de figure, celui où ψ' est du type

$$\psi'(s) = \frac{1}{\sqrt{\mu}\,(1-\sqrt{\mu})} \quad \text{avec} \quad \mu(s) = b + as$$

On doit avoir $\mu < 1$ pour $a \leq s < 1$, donc

$$a < b < 1 \quad \text{et} \quad b+a \leq 1$$

Le signe de a, par contre, n'est pas évident au départ. Mettant toutefois ψ' sous la forme

$$\psi'(s) = \frac{1+\sqrt{\mu}}{\sqrt{\mu}\,(1-\mu)}$$

nous pouvons écrire :

$$\psi'(s) = \frac{1}{1-\mu} + \frac{1}{\sqrt{1-\mu}}\;\frac{1}{\sqrt{\mu\,(1-\mu)}}$$

Le terme

$$\frac{1}{1-\mu} = \frac{1}{1-b-as}$$

possède les propriétés voulues pourvu que :

$$0 < a \leq 1-b$$

Sous ces mêmes conditions, c'est-à-dire :

$$a > 0 \quad , \quad b > 0 \quad , \quad a+b \leq 1$$

nous pouvons écrire :

$$\mu\,(1-\mu) = \frac{1}{4}\,(1-(1-2\mu)^2) = \frac{1}{4}\,[1-(2as+(2b-1)]^2$$

Sous cette forme nous constatons, en nous reportant à la discussion du début du paragraphe 5-a que le terme $1/\sqrt{\mu(1-\mu)}$ possèdera lui aussi un développement à coefficients positifs dès que les conditions

$$2a > 0 \ , \quad 2b-1 > 0 \quad , \quad 2a + (2\,b-1) < 1$$

seront remplies. Compte tenu des conditions déjà obtenues, nous voyons que $\psi'(s)$ possèdera les propriétés requises au moins dans le cas où l'on a :

$$a > 0 \quad b > \frac{1}{2} \quad a + b < 1$$

Sous ces conditions (qu'il est peut-être possible d'améliorer) on obtient effectivement une quatrième famille de lois à 3 paramètres, que nous appellerons lois de LECHIS (je compte sur la lassitude de mon lecteur pour que celui-ci ne s'aperçoive pas que j'ai oublié de traiter le cas $a < 0$).

En intégrant, on trouve, à un facteur près)

$$\psi(s) = \log \frac{1-\sqrt{b+a}}{1-\sqrt{b+as}}$$

et

$$\boxed{G_\nu(s) = \left(\frac{1-\sqrt{a+b}}{1-\sqrt{b+as}}\right)^\nu} \tag{5-4}$$

avec

$$a > 0 \quad b \geq \frac{1}{2} \quad a + b < 1$$

En dérivant, il vient

$$\begin{cases} G'(s) = \dfrac{a\nu}{2\sqrt{b+as}} \; \dfrac{G}{1-\sqrt{b+as}} \\[2ex] G''(s) = \dfrac{a^2\,\nu^2}{4(b+as)} \; \dfrac{1}{(1-\sqrt{b+as})^2}\, G + \dfrac{a^2\,\nu}{4(b+as)} \; \dfrac{G}{(1-\sqrt{b+as})^2} \\[2ex] \qquad\qquad - \dfrac{a^2\,\nu}{4(b+as)^{3/2}} \; \dfrac{G}{1-\sqrt{b+as}} \end{cases}$$

D'où la moyenne et la variance de la loi de Lechis :

$$\begin{cases} m = \dfrac{a\,\nu}{2\sqrt{a+b}} \, \dfrac{1}{1-\sqrt{a+b}} \\ \sigma^2 = m \left[1 - \dfrac{a}{2(a+b)} + \dfrac{a}{2\,(\sqrt{b+a} - a-b)} \right] \end{cases}$$

et les deux premières probabilités :

$$\begin{cases} p_o = \left(\dfrac{1-\sqrt{a+b}}{1-\sqrt{b}} \right)^{\nu} \\ p_1 = \dfrac{a\,\nu}{2\sqrt{b}(1-\sqrt{b})} \, p_o \end{cases}$$

Comme équation différentielle, on trouve :

$$(5\text{-}5) \quad (b+as)(1-b-as)\,G'' - a\left[(\nu + \tfrac{3}{2})(b+as) - \tfrac{1}{2}\right] G' - a^2 \frac{\nu(\nu+1)}{4} G = 0$$

et on en déduit les relations de récurrence :

$$(5\text{-}6) \quad b(1-b)(n+1)(n+2)p_n = a(n+1)\left[n(2b-1) + (\nu + \tfrac{3}{2})b - \tfrac{1}{2}\right] p_{n+1}$$

$$+ a^2\left[n(n-1) + n(\nu + \tfrac{3}{2}) + \frac{\nu(\nu+1)}{4}\right] p_n$$

Mélanges Poissoniens

La loi de Lechis ne peut pas être un mélange poissonien : on le voir immédiatement sur (5-4) en remarquant que $b+a(1-\lambda)$ devient négatif pour λ assez grand, ce qui n'est guère permis à une expression figurant sous un radical.

Au contraire, la loi de Sichel est effectivement un mélange poissonien, comme cela résulte du mode même de construction de cette loi par randomisation d'un premier mélange poissonien.