ÉCOLE NATIONALE SUPÉRIEURE DES MINES DE PARIS

Fontainebleau/CGMM N-717

DEUX AUTRES FAMILLES,
NON MOINS DISCRETES, MAIS
PLUS NOMBREUSES

G. MATHERON Novembre 1981

CENTRE DE GÉOSTATISTIQUE : 35, RUE SAINT-HONORÉ — 77305 FONTAINEBLEAU (FRANCE)

ÉCOLE NATIONALE SUPÉRIEURE DES MINES DE PARIS

DEUX AUTRES FAMILLES, NON MOINS DISCRETES, MAIS PLUS NOMBREUSES

Par

G. MATHERON

Novembre 1981

CENTRE DE GÉOSTATISTIQUE : 35, RUE SAINT-HONORÉ — 77305 FONTAINEBLEAU (FRANCE)

TABLE DES MATIERES

REMERCIEMENTS

Je dédie ce travail à la Haute Provence,
sans le charme et la solitude de laquelle
il n'aurait peut-être pas pu venir au jour.

DEUX AUTRES FAMILLES, NON MOINS DISCRETES, MAIS PLUS NOMBREUSES

INTRODUCTION

Le présent travail constitue la suite d'une étude antérieure, intitulée " Quatre familles discrètes" (en abrégé dans ce qui suit : 4-F). Or cette étude 4-F laisse un sentiment d'insatisfaction : la comparaison avec une population expérimentale particulièrement nombreuse (12.000 échantillons !) montre, en effet, que le modèle utilisé (la loi de Schiel à 3 paramètres) est très proche de la réalité, mais, néanmoins, le test χ^2 classique conduit au rejet : on peut certes épiloguer sur la validité de ce test, car les échantillons, comme nous le verrons, ne peuvent en aucune manière être considérés comme indépendants, mais on serait quand même plus satisfait si le test classique était positif.

Un autre motif d'insatisfaction, non moins sérieux, provient de ce que la loi de Schiel ne contient la loi classique de Sichel ni comme cas particulier ni comme cas limite. La technique de recherche utilisée dans 4-F, en effet, qui consistait à faire l'inventaire des lois indéfiniment divisibles admettant une hypergéométrique comme fonction génératrice, nous a conduit à quatre familles disconnectées. Ces quatre familles dépendent chacune de trois paramètres, et deux seulement correspondent à des mélanges poissoniens, à savoir la loi de Sichel généralisée et la loi de Schiel à 3 paramètres. Cette propriété de mélange poissonien me semble très intéressante, du fait qu'elle permet d'introduire aussi des lois continues possédant de bonnes propriétés, et dans la présente étude je m'intéresserai uniquement aux mélanges poissoniens.

Dans un premier temps, nous allons donc chercher une famille de lois discrètes un peu plus générale que les lois de Schiel et de Sichel à trois paramètres, mais contenant ces dernières comme cas particuliers, et rétablissant donc entre celles-ci la connexion que 4-F avait brisée. Cette recherche nous conduira à une famille de lois discrètes dépendant de 4 paramètres (au lieu de trois) que nous appellerons lois de Schiel à 4 paramètres, ou simplement loi de Schiel s'il n'y a pas d'ambiguité.

Avec cette loi à 4 paramètres, les 12.000 données de 4-F nous donneront un χ^2 satisfaisant. Mais d'autres difficultés se présenteront. En premier lieu, avec la loi de Schiel, le problème du changement de support ne se présente pas de manière très simple. Ce ne serait pas réellement une objection suffisante. Mais les données expérimentales relatives au changement de support montreront que l'ajustement à la loi de Schiel, satisfaisant pour les petits supports, se détériore progressivement au fur et à mesure que l'on augmente la taille des échantillons, et devient franchement détestable pour les plus grands supports.

Il nous faudra donc prendre un nouveau départ. Après avoir essayé diverses possibilités (et il en reste d'autres encore que je n'ai pas testées), j'ai finalement trouvé un modèle à 4 paramètres auquel je propose de donner le nom de <u>loi de SICHEL-PASCAL</u> (ou, si l'on est pressé, SIPA). Cette fois, les résultats sont satisfaisants pour toutes les tailles de support étudiées, et, de plus, on peut envisager des modèles (relativement!) simples de changement de support, ainsi que des lois multivariables. Ce nouveau modèle contient comme cas particulier la loi de SICHEL à 3 paramètres, mais non la loi de SCHIEL : il n'est toutefois pas très éloigné de cette dernière. En fait, il apparait que toute variable obéissant à la loi de Schiel à 4 paramètres est la somme d'une variable de SICHEL-PASCAL et d'une binomiale négative. Il est remarquable qu'une différence aussi modeste soit responsable du succès ou de l'échec expérimental.

Notons qu'il serait possible encore d'introduire une famille à 5 paramètres, dont SCHIEL et SICHEL-PASCAL seraient des cas particuliers, et qui ne serait pas beaucoup plus difficile à traiter en pratique : je ne l'ai pas fait, puisque SICHEL-PASCAL m'a donné satisfaction, mais c'est une possibilité qu'il est bon de tenir en réserve.

Mettons-nous maintenant au travail et, comme nous le verrons, même en se limitant à ces deux familles à 4 paramètres, il y a pas mal à faire.

I - RECHERCHE D'UNE LOI DE SCHIEL A 4 PARAMETRES.

Interpoles de Schiel à Sichel.

Pour commencer, rappelons l'expression des fonctions génératrices des lois de Schiel et de Sichel à 3 paramètres :

Loi de Schiel :

$$G(s) = \left(\frac{1 - \alpha + \sqrt{(1-\alpha)^2 - \beta^2}}{1 - \alpha s + \sqrt{(1-\alpha s)^2 - \beta^2}} \right)^{\nu} \tag{I-1}$$

Loi de Sichel :

$$G(s) = \left(\frac{a + \sqrt{1-\theta}}{a + \sqrt{1-\theta s}} \right)^{2\nu} \tag{I-2}$$

Entre ces deux expressions, comment interpoler ? Vendons tout de suite la mèche : dans l'expression (I-2) de la loi de Sichel, c'est à dessein que j'ai mis en exposant 2ν au lieu de ν comme dans 4-F. De fait, considérons l'identité élémentaire :

$$\left(\sqrt{1+\beta - \alpha s} + \sqrt{1-\beta - \alpha s} \right)^2 = 2\left[1 - \alpha s + \sqrt{(1-\alpha s)^2 - \beta^2}\right]$$

qui nous permet de mettre la loi de Schiel sous la forme :

$$G(s) = \left(\frac{\sqrt{1+\beta-\alpha} + \sqrt{1-\beta-\alpha}}{\sqrt{1+\beta-\alpha s} + \sqrt{1-\beta-\alpha s}} \right)^{2\nu} \tag{I-1'}$$

Il est clair dès lors que (I-1) et (I-2) sont deux cas particuliers de la formule générale suivante :

$$G(s) = \left(\frac{a\sqrt{1-\theta'} + \sqrt{1-\theta}}{a\sqrt{1-\theta' s} + \sqrt{1-\theta s}} \right)^{2\nu} \tag{I-3}$$

Nous allons montrer que (I-3) est effectivement la fonction génératrice d'une loi que nous appellerons loi de SCHIEL (à quatre paramètres).

Cela n'est d'ailleurs pas très difficile à voir. De fait, (I-3) se met sous la forme :

$$G(s) = \left(\frac{1-\theta'}{1-\theta' s}\right)^{\nu} \left(\frac{a + \sqrt{\frac{1-\theta}{1-\theta'}}}{a + \sqrt{\frac{1-\theta s}{1-\theta' s}}}\right)^{2\nu}$$

Nous supposons $\theta' \leq \theta$ (ce qui n'est pas une limitation : si $\theta' \geq \theta$, il suffit de changer a en $1/a$ et d'échanger θ et θ' pour se ramener au cas précédent), et nous poserons pour plus de clarté :

$$\theta' = p \qquad 1 - \theta' = q$$

avec évidemment les conditions :

$$0 \leq p \leq \theta \leq 1$$

Posant :

$$\theta = p + q\tau \quad ; \quad \tau = \frac{\theta - p}{q}$$

nous voyons que la fonction $G(s)$ s'écrit encore :

$$\text{(I-3')} \qquad G(s) = \left(\frac{q}{1-ps}\right)^{\nu} \left(\frac{a + \sqrt{1-\tau}}{a + \sqrt{1-\tau \frac{qs}{1-ps}}}\right)^{2\nu}$$

Le premier facteur est la fonction génératrice d'une loi binomiale négative. Le second s'obtient en remplaçant s par $qs/(1-ps)$ dans la fonction :

$$\gamma(s) = \left(\frac{a + \sqrt{1-\tau}}{a + \sqrt{1-\tau s}}\right)^{2\nu}$$

Or, $\gamma(s)$ est une fonction génératrice (celle d'une loi de SICHEL) pourvu que

$$a + 1-\tau > 0 \quad ; \quad 0 \leq \tau \leq 1 \quad ; \quad \nu > 0$$

et $qs/(1-ps)$ est également une fonction génératrice (celle d'une loi de PASCAL). Le second facteur de la formule (I-3) est donc encore, sous les mêmes conditions, lui-même une fonction génératrice - à savoir celle de la loi de SICHEL - PASCAL que nous étudierons plus loin.

Il en résulte que (I-3) est bien une loi de probabilité, et chemin faisant nous avons même établi le résultat annoncé dans l'introduction : toute variable obéissant à la loi de SCHIEL à 4 paramètres (I-3) ou (I-3') est la somme d'une variable de SICHEL-PASCAL (second facteur dans (I-3')) et d'une binomiale négative (premier facteur) indépendantes.

Ce résultat essentiel étant établi, le lecteur peut, s'il le désire, sauter la suite de ce paragraphe et passer directement au Chapitre II où sont établies les propriétés utiles de cette loi. Dans l'immédiat, nous allons montrer que notre nouvelle loi est encore un mélange poissonien, et donner quelques indications sur la loi continue correspondante.

a/ Loi continue associée à la loi de Schiel à 3 paramètres.

Aux notations près, la loi (I-1') c'est à dire la loi de Schiel à 3 paramètres s'obtient en changeant λ en 1-s dans la fonction :

$$(1-4) \qquad \Phi(\lambda) = \left(\frac{\sqrt{a} + \sqrt{b}}{\sqrt{a+\lambda} + \sqrt{b+\lambda}} \right)^{2\nu}$$

où a, b et ν sont des constantes positives. Nous allons montrer que $\Phi(\lambda)$ est la transformée de Laplace d'une loi continue f(x), d'où résulte en particulier que G(s) est un mélange poissonien, et nous donnerons deux expressions différentes de la densité f(x).

Première forme pour la densité f(x).

Supposons (par exemple) a ≥ b, et posons :

$$c = a-b \geq 0 \quad ; \quad \tau = \frac{c}{a} = \frac{a-b}{a}$$

Alors, la fonction $\Phi(\lambda)$ définie en (I-4) s'écrit encore :

$$\Phi(\lambda) = \left(\frac{a}{a+\lambda}\right)^{\nu} \left(\frac{1+\sqrt{1-\tau}}{1+\sqrt{1-\tau \frac{a}{a+\lambda}}}\right)^{2\nu}$$

On reconnait dans le deuxième facteur la fonction $g_{2\nu}$ de la note 4-F, formule (1-9) : utilisant la relation (I-12') (de 4-F) on obtient donc le développement suivant :

$$\Phi(\lambda) = 2\nu \left(\frac{1+\sqrt{1-\tau}}{2}\right)^{2\nu} \sum_{n\geq 0} \frac{\Gamma(2\nu+2n)}{n!\,\Gamma(2\nu+1+n)} \left(\frac{\tau}{4}\right)^{n} \left(\frac{a}{a+\lambda}\right)^{n+\nu}$$

qui montre que $\Phi(\lambda)$ est effectivement la transformée de Laplace d'un mélange de lois gamma.

Cherchons la densité $f(x)$. Chaque terme $(a/a+\lambda)^{n+\nu}$ est la transformée de la loi gamma

$$\frac{a^{n+\nu}\, x^{n+\nu-1}}{\Gamma(n+\nu)}\, e^{-ax}$$

Compte tenu de la formule classique de duplication :

$$\Gamma(2\nu+2n) = 2^{2\nu+2n-1}\, \frac{\Gamma(\nu+n)\,\Gamma(\nu+n+\frac{1}{2})}{\sqrt{\pi}}$$

nous voyons que cette densité s'écrit :

$$f(x) = \frac{\nu}{\sqrt{\pi}}\,(1+\sqrt{1-\tau})^{2\nu} \sum_{n\geq 0} \frac{\Gamma(\nu+\frac{1}{2}+n)}{\Gamma(2\nu+1+n)}\, \frac{\tau^{n}}{n!}\, a^{n+\nu}\, x^{n+\nu-1}\, e^{-ax}$$

Il est agréable d'introduire ici la notation classique des hypergéométriques confluentes :

$$F(\alpha,\gamma;x) = 1 + \sum_{n=1}^{\infty} \frac{\alpha(\alpha+1)\ldots(\alpha+n-1)}{\gamma(\gamma+1)\ldots(\gamma+n-1)}\, \frac{x^{n}}{n!}$$

$$= \frac{\Gamma(\gamma)}{\Gamma(\alpha)} \sum_{n\geq 0} \frac{\Gamma(\alpha+n)}{\Gamma(\gamma+n)}\, \frac{x^{n}}{n!}$$

Avec cette écriture, notre densité f(x) se met sous la forme :

$$(\text{I-5}) \qquad f(x) = \left(\frac{1+\sqrt{1-\tau}}{2}\right)^{2\nu} \nu \frac{a^{\nu} x^{\nu-1} e^{-ax}}{\Gamma(1+\nu)} F(\nu+\frac{1}{2}, 2\nu+1 ; cx)$$

(on rappelle que $c = a - b$ et $\tau = c/a$, soit $c = a\tau$).

D'autre part, on sait que (pour $\gamma > \alpha > 0$), l'hypergéométrique confluente admet la représentation intégrale :

$$F(\alpha,\gamma;x) = \frac{\Gamma(\gamma)}{\Gamma(\alpha)\ \Gamma(\gamma-a)} \int_0^1 e^{tx}\ t^{\alpha-1}(1-t)^{\gamma-\alpha-1}\ dt$$

Appliquons cette représentation avec $\alpha = \nu + \frac{1}{2}$, $\gamma = 2\alpha = 2\nu + 1$: la formule (I-5) devient alors (moyennant de légères transformations) :

$$(\text{I-5'}) \qquad f(x) = \frac{(\sqrt{a}+\sqrt{b})^{2\nu\ 2}}{\sqrt{\pi}\ \Gamma(\nu+\frac{1}{2})} x^{\nu-1} e^{-bx} \int_0^1 e^{-ctx}\ t^{\nu-\frac{1}{2}}(1-t)^{\nu-\frac{1}{2}}\ dt$$

Dans le cas $b = 0$, des simplifications se produisent et suggèrent une interprétation probabiliste intéressante : pour $b = 0$, la densité $f_0(x)$ s'écrit en effet :

$$f_0(x) = \frac{a^{\nu} x^{\nu-1}}{\Gamma(\nu)} \int_0^1 t^{\nu}\ e^{-atx} \frac{\Gamma(\nu)}{\Gamma(\frac{1}{2})\ \Gamma(\nu+\frac{1}{2})} t^{-1/2}(1-t)^{\nu-1/2}\ dt$$

Sous cette forme, on voit que la variable X_0 admettant cette densité $f_0(x)$ est équivalente en loi au rapport d'une variable gamma (ν,a) et d'une variable bêta $(\frac{1}{2}, \nu+\frac{1}{2})$: avec des notations commodes, nous désignerons par :

$$\begin{cases} S_{\alpha} & \text{une V.A. de loi gamma réduite } (\alpha,1) \\ \Pi_{\alpha,\beta} & \text{une V.A. de loi beta } (\alpha,\beta) \end{cases}$$

Alors, notre équivalence en loi peut s'écrire :

(I-5")
$$\boxed{X_o \equiv \frac{1}{a} \frac{S_\nu}{\Pi_{1/2,\nu+1/2}}}$$

Dans le cas général où b n'est pas nulle, la densité f se déduit de la densité f_o ci-dessus à l'aide d'une simple pondération par l'exponentielle e^{-bx}. Il suffit donc d'examiner le cas b = 0. Comme a est un simple paramètre d'échelle, nous pouvons aussi prendre a = 1 (ce qui revient à;remplacer X_o par a X_o). Enfin, l'équivalence en loi bien classique :

$$\Pi_{\alpha,\beta} \equiv \frac{S_\alpha}{S_\alpha + S_\beta}$$

où S_α et S_β sont deux variables gamma indépendantes, nous donne avec $\alpha = 1/2$ et $\beta = \nu + 1/2$:

$$\frac{1}{\Pi_{1/2,\nu+1/2}} \equiv 1 + \frac{S_{\nu+\frac{1}{2}}}{S_{\frac{1}{2}}}$$

Ainsi, le prototype de la loi continue engendrant par mélange poissonien une loi de Schiel à 3 paramètres correspond à une variable de la forme :

(I-6)
$$\boxed{X = S_\nu\left(1 + \frac{S_{\nu+1/2}}{S_{\frac{1}{2}}}\right)}$$

C'est cette formule que nous allons généraliser. Rappelons auparavant la seconde forme de densité f(x), forme déjà mentionnée d'ailleurs dans 4-F.

<u>Deuxième forme pour la densité f(x).</u>

La fonction $\Phi(\lambda)$ de la formule (I-4) peut encore s'écrire :

$$\Phi(\lambda) = \left(\frac{a+b + 2\sqrt{ab}}{a+b + 2\lambda + 2\sqrt{(a+\lambda)(b+\lambda)}}\right)^\nu$$

Il sera commode de poser

$$r = \frac{a+b}{2} \quad ; \quad t = \frac{a-b}{2} = \frac{c}{2} \quad ; \quad \gamma = \frac{t^2}{r^2}$$

Alors :

$$\Phi(\lambda) = \left(\frac{r + \sqrt{r^2 - t^2}}{r+\lambda + \sqrt{(r+\lambda)^2 - t^2}} \right)^\nu$$

$$= \left(\frac{r}{r+\lambda}\right)^\nu \left(\frac{1 + \sqrt{1-\gamma}}{1 + \sqrt{1-\gamma \frac{r^2}{(r+\lambda)^2}}} \right)^\nu$$

En utilisant à nouveau la formule (I-12') de 4-F (mais cette fois en l'exposant ν au lieu de 2ν) nous obtenons ensuite le développement suivant :

$$\Phi(\lambda) = \nu \left(\frac{1 + \sqrt{1-\gamma}}{2}\right)^\nu \sum_{n \geq 0} \frac{\Gamma(\nu+2n)}{n!\ \Gamma(\nu+1+n)} \left(\frac{\gamma}{4}\right)^n \left(\frac{r}{r+\lambda}\right)^{2n+\nu}$$

qui montre que la densité $f(x)$ se représente d'une seconde manière comme mélange de lois gamma :

$$f(x) = \nu \left(\frac{1 + \sqrt{1-\gamma}}{2}\right)^\nu r^\nu x^{\nu-1} e^{-rx} \sum \frac{1}{n!\ \Gamma(n+\nu+1)} \left(\frac{t^2 x^2}{4}\right)^n$$

Il est alors agréable d'introduire <u>la fonction de Bessel modifiée de 1ère espèce</u> :

$$I_\nu(y) = \left(\frac{y}{2}\right)^\nu \sum \left(\frac{y^2}{4}\right)^n \frac{1}{n!\ \Gamma(n+\nu+1)}$$

En revenant aux notations initiales (a et b au lieu de r, t et γ) on trouve ainsi pour la densité :

$$\text{(I-7)} \qquad f(x) = \nu \left(\frac{a+b}{2} + \sqrt{ab}\right)^\nu \frac{e^{-\frac{a+b}{2}x}}{x} I_\nu\left(\frac{a-b}{2}x\right)$$

Sous cette forme, le comportement asymptotique de $f(x)$ pour $x \to \infty$ est bien mis en évidence : on sait, en effet, que

$I_\nu(y)$ se comporte à l'infini comme $e^y/\sqrt{y}$. Ainsi, pour $x \to \infty$, $f(x)$ aura un comportement en :

$$\frac{e^{-bx}}{x^{3/2}}$$

(et en particulier en $1/x^{3/2}$ dans le cas prototypique $b = 0$). Ce comportement est le même que celui de la loi

$$\frac{A\, e^{-bx-\frac{a}{x}}}{x^{3/2}}$$

utilisée par Sichel pour former ses mélanges poissoniens.

En égalant les deux expressions (I-5) et (I-7) de la densité $f(x)$ on obtient une identité entre fonction de Bessel I_ν et hypergéométrique confluente $F(\nu+\frac{1}{2}, 2\nu+1)$. Cette identité, nullement triviale, est connue dans la littérature sous le nom d'identité de Kummer.

Enfin, la représentation intégrale (I-5'), et l'équivalence en loi (I-6) qui s'en déduit peuvent également être obtenues à partir de l'expression (I-7) : il suffit pour cela d'utiliser la représentation intégrale classique

$$I_\nu(x) = \frac{1}{\sqrt{\pi}\,\Gamma(\nu+\frac{1}{2})}\left(\frac{x}{2}\right)^\nu \int_{-1}^{1} e^{tx}\,(1-t^2)^{\nu-\frac{1}{2}}\,dt$$

et de faire les changements de variables voulus.

b/ Loi continue associée à la loi de Sichel à 3 paramètres.

La loi continue associée à la loi de Sichel à 2 paramètres admet la transformée de Laplace $\exp(a\sqrt{b} - a\sqrt{b+\lambda})$, dont le prototype est la loi stable :

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

On en déduit le prototype de la loi de Sichel à 3 paramètres en substituant à la constante a une variable gamma que nous prendrons, comme ci-dessus, d'indice 2ν (plutot que ν), soit $S_{2\nu}$. D'autre part, on sait que la loi stable ci-dessus est celle de l'inverse $1/S_{1/2}$ d'une gamma de paramètre 1/2. Plus précisément :

$$E(e^{-\lambda/S_{1/2}}) = e^{-2\sqrt{\lambda}}$$

Introduisant deux gamma indépendantes $S_{2\nu}$ et $S_{1/2}$, on trouve donc

$$E\left(e^{-\frac{\lambda}{4}\frac{S_{2\nu}^2}{S_{1/2}}}\right) = E\left(e^{-S_{2\nu}\sqrt{\lambda}}\right) = \left(\frac{1}{1+\sqrt{\lambda}}\right)^{2\nu}$$

Il s'agit bien du prototype de la loi de Sichel à 3 paramètres (la formule générale s'obtient en effectuant une transformation $\lambda \rightarrow a\lambda + b$: ce qui revient à un changement d'échelle sur la V.A. et une pondération de sa densité par une exponentielle e^{-bx}).

Ainsi, pour la loi de Sichel à 3 paramètres, la variable continue prototypique est du type :

(I-8) $$\boxed{X = \frac{S_{2\nu}^2}{4\,S_{1/2}}}$$

Pour comparer ceci avec la variable (I-6) engendrant la loi de Schiel à 3 paramètres, nous allons utiliser une interprétation probabiliste amusante de la formule classique de duplication de la fonction eulérienne. En effet, écrivant cette formule sous les deux formes :

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

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

nous obtenons

$$\frac{\Gamma(2\nu + 2\lambda)}{\Gamma(2\nu)} = 2^{2\lambda} \; \frac{\Gamma(\nu + \lambda)}{\Gamma(\nu)} \; \frac{\Gamma(\nu + 1/2 + \lambda)}{\Gamma(\nu + 1/2)}$$

Mais, comme on sait :

$$\frac{\Gamma(\alpha + \lambda)}{\Gamma(\alpha)} = E((S_\alpha)^\lambda)$$

La formule ci-dessus exprime donc une équivalence en loi entre variables gamma, à savoir :

(I-9) $$S_{2\nu}^2 \equiv 4 \; S_\nu \; S_{\nu+1/2}$$

Compte tenu de cette très curieuse relation, nous voyons que la variable (I-8), prototype de la loi de Sichel à 3 paramètres, peut aussi bien s'écrire :

(I-10) $$\boxed{X = \frac{S_\nu \; S_\nu + 1/2}{S_{1/2}}}$$

c/ <u>Loi continue de Schiel à 4 paramètres</u>.

Comparant les variables prototypiques (I-6) et (I-10) de nos deux lois à 3 paramètres, nous sommes fort tentés d'introduire la variable :

(I-11) $$\boxed{Z = S_\nu \left(\frac{S_\nu + 1/2}{S_{1/2}} + \beta \right)}$$

à partir de laquelle on peut constituer une famille à quatre paramètres, contenant comme cas particuliers nos deux lois à trois paramètres : pour $\beta = 0$, en effet, (I-11) se réduit à (I-10) et engendre la loi de Sichel, pour $\beta = 1$ au contraire (I-11) s'identifie à (I-6) et engendre la loi de Schiel à 3 paramètres. De fait, nous allons voir que la variable (I-11) engendre effectivement la loi de Schiel à 4 paramètres, telle que nous l'avons définie.

Pour cela, partons de l'équivalence en loi (I-9) établie ci-dessus

$$(\text{I-9}) \qquad S_{2\nu}^2 \equiv 4\, S_\nu\, S_{S+\,1/2}$$

S_ν et $S_{\nu+\,1/2}$ étant deux gamma indépendantes, posons :

$$X = 2\sqrt{S_\nu\, S_{\nu+\,1/2}} \quad ; \quad Y = S_\nu$$

de sorte que X, d'après (I-9), est une gamma de paramètre 2ν. Cherchons alors <u>la loi conditionnelle</u> de $Y = S_\nu$ lorsque $X = x$ est fixée : en désignant par f_α la densité de la variable gamma S_α, on trouve pour cette loi conditionnelle la densité

$$f(y/x) = \frac{\frac{x}{2y}\, f_{\nu+1/2}\left(\frac{x^2}{4y}\right) f_\nu(y)}{f_{2\nu}(x)}$$

Sous forme explicite, on constate que le paramètre ν s'élimine, et l'on trouve :

$$(\text{I-12}) \qquad f(y/x) = \frac{1}{2\sqrt{\pi}}\, \frac{x}{y^{3/2}}\, e^{x-y-\frac{x^2}{4y}}$$

Chose très curieuse, donc, cette loi conditionnelle de $Y = S_\nu$ à $X = 2\sqrt{S_\nu\, S_{\nu+1/2}}$ fixé coïncide avec la loi continue qui engendre, par mélange poissonien, la loi de Sichel sensu stricto (à 2 paramètres). En désignant par

$$\Phi_x(\lambda) = \int_0^\infty e^{-\lambda y}\, f(y/x)\, dy$$

sa transformée de Laplace, on trouve en particulier :

$$(\text{I-13}) \qquad \Phi_x(\lambda) \equiv E\left(e^{-\lambda S_\nu} / S_\nu\, S_\nu + \frac{1}{2} = \frac{x^2}{4}\right) = e^{x(1-\sqrt{1+\lambda})}$$

A partir de cette relation, il n'est plus très difficile de trouver la transformée de Laplace de la variable Z définie en

(I-11). Posons, pour abréger, les notations

$$\begin{cases} X = 2\sqrt{S_\nu \; S_{\nu+1/2}} \\ Y = S_\nu \\ Z = \dfrac{X^2}{4\, S_{1/2}} + \beta Y \end{cases}$$

Nous devons calculer la transformée de Laplace de Z, soit :

$$\Phi(\lambda) = E(e^{-\lambda Z}) = E\left(e^{-\lambda \frac{X^2}{4\, S_{1/2}} - \lambda \beta Y}\right)$$

Calculons d'abord cette espérance à $X = x$ et $S_{1/2} = s$ fixés. D'après (I-13), on trouve :

$$E(e^{-\lambda Z}/X = x, \; S_{1/2} = s) = e^{-\frac{\lambda x^2}{4s} + x(1 - \sqrt{1+\lambda\beta})}$$

Déconditionnalisons maintenant d'abord en $S_{1/2}$: il vient

$$E(e^{-\lambda Z}/X = x) = e^{x(1 - \sqrt{1+\lambda\beta} - \sqrt{\lambda})}$$

Déconditionnalisons ensuite en X, qui est une gamma $(2\nu, 1)$. Cela donne

$$\text{(I-14)} \qquad \Phi(\lambda) = \left(\frac{1}{\sqrt{\lambda} + \sqrt{1+\lambda\beta}}\right)^{2\nu}$$

Il s'agit donc bien de la forme prototypique de la loi continue engendrant par mélange poissonien notre loi de Schiel à quatre paramètres.

Il convient, maintenant, de chercher l'expression de la densité correspondante. Elle n'est malheureusement pas aussi simple que dans les cas précédents. On peut en donner une représentation intégrale ou un développement en série double.

Représentation intégrale de la densité $f(z)$.

Donnons cette expression seulement dans le cas de la variable prototypique (I-11) (le cas général s'en déduira en changeant Z en aZ et en pondérant la densité par e^{-bz}). Pour cela, remarquons que la variable beta $(\frac{1}{2}, \nu + \frac{1}{2})$ vérifie l'équivalence en loi déjà rencontrée :

$$\frac{1}{\Pi_{\frac{1}{2}, \nu+\frac{1}{2}}} \equiv 1 + \frac{S_{\nu+1/2}}{S_{1/2}}$$

de sorte que la variable (I-11) peut être remplacée par :

$$Z = S_\nu \left(\frac{1}{\Pi_{1/2, \nu+\frac{1}{2}}} + \beta - 1 \right)$$

Nous poserons, pour abréger

$$\alpha = \beta - 1 \qquad (\alpha \geq -1)$$

et il vient alors pour la densité :

$$\text{(I-15)} \qquad f(z) = \frac{\nu\, z^{\nu-1}}{\sqrt{\pi}\, \Gamma(\nu+\frac{1}{2})} \int_0^1 e^{-\frac{zu}{1+\alpha u}} \frac{u^{\nu-1/2}(1-u)^{\nu-1/2}}{(1+\alpha u)^\nu} du$$

<u>Développement en série double de la densité $f(z)$</u>

Sous sa forme générale, la version continue de la loi de Schiel à 4 paramètres admet une transformée de Laplace de la forme :

$$\Phi(\lambda) = \left(\frac{c}{c+\lambda}\right)^\nu \left(\frac{a + \sqrt{1-\tau}}{a + \sqrt{1-\tau \frac{c}{c+\lambda}}} \right)^{2\nu}$$

Désignons par $\varpi_n(a)$ les coefficients du développement

$$\left(\frac{a + \sqrt{1-\tau}}{a + \sqrt{1-\tau s}} \right)^{2\nu} = \sum_{n \geq 0} \varpi_n(a)\, s^n$$

On aura donc pour la densité :

$$\text{(I-16)} \qquad f(z) = \sum_{n \geq 0} \varpi_n(a) \frac{c^{n+\nu}\, z^{n+\nu-1}\, e^{-cz}}{\Gamma(n+\nu)}$$

Or, on a vu dans 4-F :

$$\left(\frac{a+\sqrt{1-\tau}}{a+\sqrt{1-\tau s}}\right)^{2\nu} = \sum_n q^\nu \frac{\Gamma(n+2\nu)}{n!} p^n \left(\frac{1-\sqrt{1-\tau s}}{1-\sqrt{1-\tau}}\right)^n$$

avec

$$p = \frac{1-\sqrt{1-\tau}}{a+1} \quad ; \quad q = \frac{a+\sqrt{1-\tau}}{a+1}$$

et, d'autre part

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

$$= \left(\frac{1+\sqrt{1-\tau}}{2}\right)^n n \sum_{k\geq o} \frac{\Gamma(n+2k)}{k!\ \Gamma(n+1+k)} \left(\frac{\tau}{4}\right)^k sk$$

On en déduit :

$$\varpi_n(a) = \left(\frac{a+\sqrt{1-\tau}}{a+1}\right)^\nu \frac{1}{n!} \left(\frac{\tau}{4}\right)^n \sum_{k=o}^{n} \frac{\Gamma(2\nu+n-k)}{(n-k)!} \frac{(n+k-1)!}{k!} \left(\frac{2}{a+1}\right)^{r-k}$$

En reportant dans (I-16), on obtient un développement de f(z) en série double.

On peut obtenir un autre développement (particulièrement intéressant lorsque a est voisin de 1) en utilisant le dévelop-pement en série de Taylor (par rapport à la variable a) de l'ex-pression

$$\left(\frac{1}{a+\sqrt{1-\tau s}}\right)^{2\nu} = \sum_{p\geq o} (-1)^p \frac{\Gamma(2\nu+p)}{\Gamma(2\nu)\ p!} (a-1)^p \left(\frac{1}{1+\sqrt{1-\tau s}}\right)^{2\nu+p}$$

Compte tenu de la formule (1-12') de 4-F, ceci donne $\varpi_n(a)$ sous la forme

$$\varpi_n(a) = \frac{(a+\sqrt{1-\tau})^{2\nu}}{2^{2n+2\nu}} \frac{\tau^n}{n!\ \Gamma(2\nu)} \sum_{p\geq o} (-1)^p \left(\frac{a-1}{2}\right)^p \frac{\Gamma(2\nu+p+1)\ \Gamma(2\nu+p+2n)}{\Gamma(2\nu+p+n)\ p!}$$

En substituant dans (I-16), on obtient un développement de f(z) en série double.

Ces résultats ne sont pas particulièrement simples, mais les calculs numériques semblent réalisables (je ne les ai toutefois pas essayés : en particulier, personne n'étant à l'abri d'une erreur de calcul, je conseille fortement à un utilisateur éventuel de vérifier soigneusement les calculs ci-dessus avant de programmer ces formules).

II - PROPRIETES DE LA LOI DE SCHIEL.

Nous en venons maintenant à une tâche ingrate, mais utile : calculer la moyenne, la variance, les deux premières probabilités p_o et p_1 de la loi de Schiel à 4 paramètres, puis établir l'équation différentielle qui vérifie la fonction génératrice, et en déduire les relations de récurrence qui permettent le calcul des probabilités p_n successives : nous verrons qu'il s'agit ici d'une relation de récurrence à quatre termes, et non plus à trois comme dans 4-F : il fallait bien s'y attendre, puisque les lois identifiées dans 4-F sont les seules lois indéfiniment divisibles admettant des relations de récurrence à trois termes.

Nous partons de la fonction génératrice :

$$G(s) = \left(\frac{b\sqrt{1-\theta'} + \sqrt{1-\theta}}{b\sqrt{1-\theta' s} + \sqrt{1-\theta s}}\right)^{2\nu} \tag{II-1}$$

Par dérivations successives, on trouve

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

$$\frac{G''}{G} - \frac{G'^2}{G^2} = \frac{\nu}{2}\left(\frac{b\,\theta'^2}{(1-\theta' s)^{3/2}} + \frac{\theta^2}{(1-\theta s)^{3/2}}\right)\frac{1}{b\sqrt{1-\theta' s} + \sqrt{1-\theta s}}$$

$$+ \frac{\nu}{2}\left(\frac{b\,\theta'}{\sqrt{1-\theta' s}} + \frac{\theta}{\sqrt{1-\theta s}}\right)^2 \frac{1}{(b\sqrt{1-\theta' s} + \sqrt{1-\theta s})^2}$$

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

$$\left\{\begin{aligned} m &= \frac{\nu}{b\sqrt{1-\theta'}+\sqrt{1-\theta}}\left(\frac{b\,\theta'}{\sqrt{1-\theta'}}+\frac{\theta}{\sqrt{1-\theta}}\right) \\ \sigma^2 &= m+\frac{\nu}{2}\frac{m^2}{\nu^2}+\frac{\nu}{2}\left(\frac{b\,\theta'^2}{(1-\theta')^{3/2}}+\frac{\theta^2}{(1-\theta)^{3/2}}\right)\frac{1}{b\sqrt{1-\theta'}+\sqrt{1-\theta}} \end{aligned}\right.$$

et les deux premières probabilités

$$\left\{\begin{aligned} p_o &= \left(\frac{b\sqrt{1-\theta'}+\sqrt{1-\theta}}{b+1}\right)^{2\nu} \\ p_1 &= \nu\,\frac{b\,\theta'+\theta}{b+1}\,p_o \end{aligned}\right.$$

<u>L'Equation différentielle</u>.

Pour établir <u>l'équation différentielle</u> qui vérifie G(s), nous partirons plutot de la forme

$$G(s) = \left(\frac{b^2(1-\theta')+1-\theta+2b\sqrt{(1-\theta)(1-\theta')}}{b^2(1-\theta's)+(1-\theta s)+2b\sqrt{(1-\theta s)(1-\theta' s)}}\right)^{\nu}$$

Posons pour abréger

$$P(s) = 1-\theta s \quad ; \quad \Pi(s) = b^2(1-\theta' s)$$

d'où

$$G = \left(\frac{C^{ste}}{P+\Pi+2\sqrt{P\Pi}}\right)^{\nu}$$

Ceci est de la forme

$$G = e^{\nu\psi}$$

avec

$$\psi = -\log\left(P+\Pi+2\sqrt{P\Pi}\right)+C^{ste}$$

En dérivant ψ il vient :

$$-\psi' = \left(P' + \Pi' + \frac{P\Pi' + \Pi P'}{\sqrt{P\Pi}}\right) \frac{1}{P + \Pi + 2\sqrt{P\Pi}}$$

$$= \frac{P\Pi' + \Pi P' + (P' + \Pi')\sqrt{P\Pi}}{\sqrt{P\Pi}} \times \frac{P + \Pi - 2\sqrt{P\Pi}}{(P-\Pi)^2}$$

Cette expression se simplifie notablement, et on trouve

(II-2) $$-\psi' = \frac{P' - \Pi'}{P-\Pi} + \frac{\gamma}{(P-\Pi)\sqrt{P\Pi}}$$

où γ est la constante définie par :

$$\gamma = P\,\Pi' - \Pi\,P' = b^2(\theta-\theta')$$

On en déduit :

$$\psi'^2 + 2\,\frac{(P' - \Pi')}{P-\Pi}\,\psi' + \frac{(P' - \Pi')^2\,P\,\Pi - \gamma^2}{(P-\Pi)^2\,P\,\Pi} = 0$$

Une nouvelle simplification dans le terme indépendant de ψ' conduit finalement à l'équation :

(II-3) $$P\,\Pi\,(P-\Pi)\,\psi'^2 + 2(P' - \Pi')\,\psi' + (\Pi\,P'^2 - P\,\Pi'^2) = 0$$

En dérivant (Π-2), on trouve ensuite :

$$\psi'' = -\frac{P' - \Pi'}{P-\Pi}\,\psi' + \frac{P\,\Pi' + \Pi\,P'}{2\,P\,\Pi}\;\frac{\gamma}{(P-\Pi)\sqrt{P\Pi}}$$

$$= -\frac{P' - \Pi'}{P-\Pi}\,\psi' - \frac{P\,\Pi' + \Pi\,P'}{2\,\Pi\,P}\left(\psi' + \frac{P' - \Pi'}{P-\Pi}\right)$$

ce qui conduit à l'équation différentielle

(II-4) $$P\,\Pi(P-\Pi)\,\psi'' + \left[P\,\Pi(P' - \Pi') + (P-\Pi)\left(\frac{P\,\Pi' + \Pi P'}{2}\right)\right]\psi' + (P' - \Pi')\,\frac{P\,\Pi' + \Pi P'}{2} = 0$$

Ainsi, en posant

$$\begin{cases} A(s) = P\,\Pi\,(P-\Pi) = b^2(1-\theta s)(1-\theta' s)\,[1 - b^2 + \mu s] \\ B(s) = P\,\Pi\,(P'-\Pi') = b^2\,\mu(1-\theta s)(1-\theta' s) \\ C = \Pi\,P'^2 - P\,\Pi'^2 = b^2(\theta^2 - b^2\,\theta'^2) + \theta\theta'\,\mu s \\ B_o = B + (P-\Pi)\,\dfrac{P\,\Pi' + \Pi P'}{2} = B + b^2(\theta\theta' s - \dfrac{\theta+\theta'}{2})(1 - b^2 + \mu s) \\ C_o = (P'-\Pi')\,\dfrac{P\,\Pi' + \Pi P'}{2} = b^2\mu(\theta\theta' s - \dfrac{\theta+\theta'}{2}) \end{cases}$$

avec, pour abréger :

$$\mu = b^2\,\theta' - \theta$$

on voit que $\psi = \frac{1}{\nu}\log G$ vérifie les deux équations :

$$A\,\psi'^2 + 2\,B\,\psi' + C = 0$$

$$A\,\psi'' + B_o\,\psi' + C_o = 0$$

Compte tenu de :

$$G' = \nu\,\psi'\,G \quad ; \quad G'' = (\nu\,\psi'' + \nu^2\,\psi'^2)\,G$$

il en résulte que G lui-même obéit à l'équation différentielle du second ordre :

$$A\,G'' + (B_o + 2\nu\,B)G' + (\nu\,C_o + \nu^2\,C)G = 0 \tag{II-5}$$

La Relation de récurrence

Les coefficients de G", G' et G étant, respectivement, de degré 3, 2 et 1, les p_n vont donc vérifier une relation de récurrence à quatre termes, relation que l'on obtient en annulant le coefficient du terme en s^n dans l'expression explicite de l'équation (II-5). Ce petit travail de patience conduit au résultat suivant :

$$(II-6)\quad \begin{cases} (1-b^2)(n+2)(n+1)\ p_{n+2} \\ +\ [\mu(n+1)(n+2\nu+1) - (1-b^2)(\theta+\theta')(n+1)(n+\frac{1}{2})]\ p_{n+1} \\ +\ [\nu^2(\theta^2-b^2\ \theta'^2) + (1-b^2)\ \theta\theta' n^2 - \mu(\theta+\theta')(\frac{\nu}{2} + (2\nu+\frac{3}{2})n + n(n-1)]p_n \\ +\ \mu\theta\theta'[\nu(\nu+1) + 2(\nu+1)(n-1) + (n-1)(n-2)]\ p_{n-1} = 0 \end{cases}$$

Sauf erreur de transcription, cette formule doit être correcte : je l'ai abondamment testée numériquement.

Lorsque le terme

$$\mu = b^2\ \theta'-\theta$$

s'annule, on retombe sur la loi de Schiel à 3 paramètres. Le terme en p_{n-1} de la relation de récurrence (II-6) disparait, et il reste une relation à 3 termes équivalente (aux notations près) à celle qui a été établie dans 4-F.

Lorsque $\theta' = 0$, la même circonstance se produit : on retombe sur la loi de Sichel à 3 paramètres de la note 4-F, avec la seule différence que le coefficient ν de 4-F est ici remplacé par 2ν.

Une erreur de transcription s'est glissée dans la formule (5-3) de la note 4-F : la relation de récurrence correcte pour la loi de Sichel (avec ν et non 2ν) est :

$$\begin{array}{l} (5\text{-}3) \\ \text{de } 4\text{-F} \end{array} \quad \begin{cases} (a^2-1)(n+2)(n+1)\ p_{n+2} = \\ -\ \theta(n+1)[\nu + \frac{3}{2} - \frac{a^2}{2} + (2-a^2)n]\ p_{n+1} \\ +\ \theta^2[n(n-1) + n(\nu+\frac{3}{2}) + \frac{\nu(\nu+1)}{4}]\ p_n \end{cases}$$

C'est cette formule qui doit remplacer la relation (5-3) de 4-F. Compte tenu de cette rectification, on retrouve bien (5-3) de 4-F en annulant θ' dans (II-6), et en remplaçant ν par $\nu/2$.

Lorsque $b = 1$, le terme en p_{n+2} s'évanouit dans (II-6), et il reste une relation à 3 termes seulement. Cette simplification, déjà rencontrée dans F-4, a comme contrepartie une très grande instabilité numérique lorsque b est voisin de 1. Dès que $|b^2 - 1|$ est inférieur à 0,6 ou 0,7, le calcul des p_n devient numériquement impossible. Cette circonstance m'a beaucoup gêné dans les ajustements.

N.B. : La relation (II-6) est valable à partir de n = 0 (avec $p_{-1} = 0$), de sorte qu'il suffit de calculer p_o et p_1 pour amorcer la récurrence.

III - EXEMPLES D'AJUSTEMENTS A LA LOI DE SCHIEL

A tout seigneur tout honneur : pour le premier ajustement, j'ai repris les 12 017 données que j'avais ajustées dans F-4 à la loi de Schiel à 3 paramètres. La technique d'ajustement utilisée est assez rudimentaire : elle consiste à imposer pour p_o, p_1 et m les valeurs expérimentales (la variance n'est pas assez fiable pour être retenue), ce qui donne trois relations pour quatre paramètres : l'un des paramètres (en pratique θ) étant choisi arbitrairement, les trois autres se trouvent fixés. Il suffit alors de faire varier le paramètre libre (θ), et de retenir la valeur jugée la plus satisfaisante (par exemple celle qui minimise un χ^2). On peut d'ailleurs ensuite faire varier les paramètres fixés (p_o, p_1 et m) pour améliorer le résultat : je l'ai fait dans d'autres cas, mais non dans celui-ci.

Le résultat de cet ajustement, reproduit ci-après, est assez spectaculaire : le χ^2, qui était de 56,88 dans le cas de la loi à 3 paramètres, tombe à 40,80 , ce qui, avec 34 degrés de liberté, correspond à une probabilité $P(\chi^2 \geq 40{,}80) = 20/100$ (au lieu de 1/100) et permet donc, classiquement, d'accepter le modèle.

En réalité, l'ajustement, qui est extraordinairement bon jusqu'à n = 60, et très bon jusqu'à n = 100, représente mal l'extrême queue de la distribution : la classe des 100-200 apporte, en effet, à elle seule une contribution énorme (11,77!) au χ^2. Il est intéressant de comparer les > 100 et les > 200 calculées et observées :

Ajustement à la loi de SCHIEL à 4 paramètres

Toutes plages, Support 5 m, 12017 données

classe	Fréquence observée	Fréquence calculée	χ^2
0	7603	7603	0,00
1	1376	1376	0,00
2	722	712	0,13
3	440	452	0,29
4	294	314	1,23
5	233	230	0,04
6	198	175	2,97
7	137	137	0,00
8	109	110	0,01
9	90	90	0,00
10	76	74	0,04
11	71	62	1,15
12	42	53	2,36
13	53	45,8	1,13
14	41	39,8	0,04
15	37	34,9	0,13
16	37	30,8	1,23
17	29	27,5	0,09
18	30	24,6	1,18
19	22	22,2	0,00
20	16	20,1	0,84
21	23	18,3	1,19
22	18	16,8	0,09
23	22	15,4	2,81
24	17	14,2	0,54
25	13	13,2	0,00
26	9	12,2	0,85
27	10	11,4	0,17
28	16	10,6	2,70
29	9	10,0	0,09
30	9	9,3	0,01
31-40	66	69,4	0,17
41-50	40	43,2	0,23
51-60	30	29,4	0,01
61-70	13	21,1	3,13
71-80	17	15,8	0,09
81-100	15	21,8	2,11
101-200	17	38,2	11,77
≥ 201	17	12,1	1,97

$\chi^2 = 40,80$

N = 12017 données

Ajustement sur p_0, b_1 et m

$\theta = 0,9932$
$\theta' = 0,7733$
$\beta = 1,93665$
$\alpha = 0,21338$

Moyenne obs. 3,2264
" calc. 3,2264

Variance obs. 426,57
" calc. 215,59

Degrés de liberté : 34

$\boxed{P(\chi^2 \geq 40,80) = 0,20}$

Les > 100

	Calc.	Obs.
Fréquence en %	0,42%	0,29%
Fréquence en nombre	50,3	34
N. de pierres	8 738	9 751
% des pierres	22,54%	25,15%
Moyenne	173,7	286,8

Les > 200

	Calc.	Obs.
Fréquence en %	0,10%	0,14%
Fréquence en nombre	12,1	17
N. de pierres	3 520	7 175
% des pierres	9,08%	18,51%
Moyenne	290,7	422,1

Il est certain que ces classes extrêmes sont assez mal représentées par le modèle. A dire vrai, la population étudiée ici est hétérogène : elle résulte, en effet, du mélange de 6 populations qui présentent, nous le verrons, des comportements asymptotiques (i.e. des paramètres θ) assez différents : il ne faut peut-être pas s'étonner qu'il ne soit pas possible de bien représenter la queue de distribution à l'aide d'un modèle simple.

Il reste que, dans l'ensemble, cet ajustement constitue un succès spectaculaire pour le modèle de Schiel à 4 paramètres. Mais il faut aller plus loin, et regarder comment se comporte notre population sous l'effet d'un changement de support.

Changement de support

Les données précédentes proviennent de tranchées découpées en tronçons de 5m (support 5m). Ayant regroupé ces tronçons (là où c'était possible) par 2, par 3, par 5 et par 10, on dispose des distributions relatives aux supports 10, 15, 25 et 50m.

On constate que les ajustements se détériorent lorsque le support augmente, et deviennent inacceptables à partir de 25m. Le test χ^2 a été effectué à partir d'un découpage en 22 classes (les mêmes pour tous les supports), différent du découpage précédent. Il lui correspond 22 - 5 = 17 degrés de liberté. Voici les résultats des 5 tests χ^2 :

χ^2 toutes plages (Schiel)		
	χ^2	P_{17} ($\chi^2 \geq$.)
Support 5 m	22,42	17/100
" 10 m	22,16	17,9/100
" 15 m	25,29	8,8/100
" 25 m	56,94	$3,34 \times 10^{-6}$
" 50 m	68,90	$3,32 \times 10^{-8}$

L'échec est donc total à partir de 25 m. A dire vrai, du fait que les échantillons ne sont nullement indépendants, on ne peut pas prendre à la lettre les probabilités indiquées pour les χ^2, mais les ordres de grandeurs sont décisifs : en particulier le saut très brutal du χ^2 entre 15 et 25 m (alors que l'effectif de la population diminue) est certainement très significatif. A partir de 25 m, le modèle est inacceptable.

Cet échec n'est pas imputable à l'hétérogénéité de la population. Voici, à titre indicatif, les résultats obtenus pour les deux premières sous-populations homogènes (plages A et B):

.../

χ^2 Plage A (Schiel)				
	N	χ^2	Deg. de Lib.	$P(\chi^2 \geq .)$
Support 5 m	2 357	27,61	14	1,6/100
" 10 m	1 145	28,42	14	1,3/100
" 15 m	747	12,20	13	51,1/100
" 25 m	422	15,69	13	26,6/100
" 50 m	178	27,29	9	0,13/100

χ^2 Plage B (Schiel)				
	N	χ^2	Deg. de Lib.	$P(\chi^2 \geq .)$
Support 5 m	3 026	22,71	21	36,0/100
" 10 m	1 474	17,85	21	65,8/100
" 15 m	964	19,23	21	57,0/100
" 25 m	540	19,88	14	13,4/100
" 50 m	227	22,51	11	2,1/100

Dans 6 cas sur 10, les résultats sont acceptables ou même très bons, dans 4 cas sur 10 ils sont mauvais ou même très mauvais (A-50). Dans les deux cas, le support 50 donne de mauvais résultats.

Evolution des paramètres

Avant de passer à la loi de SICHEL-PASCAL, qui, nous le verrons, donne des résultats franchement meilleurs, il est intéressant de jeter un coup d'oeil sur l'évolution des paramètres en fonction du support (à titre de comparaison, mentionnons que, si les échantillons de 5 m pouvaient être considérés comme indépendants, les paramètres θ, θ' et b devraient rester constants, tandis que ν varierait proportionnellement au support : il n'en est pas du tout ainsi, les résultats ci-dessous montrent que les échantillons ne peuvent absolument pas être regardés comme indépendants).

Evolution des paramètres - Toutes plages				
	θ	θ'	b	ν
Support 5 m	0,99230	0,7796	1,920	0,212
" 10 m	0,99470	0,8560	1,871	0,281
" 15 m	0,99384	0,8519	1,286	0,342
" 25 m	0,99288	0,9402	1,136	0,373
" 50 m	0,99147	0,9644	0,419	0,494
	Plage A			
Support 5 m	0,98923	0,6550	0,466	0,106
" 10 m	0,98612	0,8113	0,239	0,132
" 15 m	0,98374	0,8904	0,122	0,160
" 25 m	0,98729	0,9472	0,013	0,176
" 50 m	0,98591	0,9524	0,039	0,371
	Plage B			
Support 5 m	0,99385	0,6342	1,162	0,192
" 10 m	0,99556	0,7572	1	0,239
" 15 m	0,99564	0,9006	1	0,229
" 25 m	0,99760	0,9520	1,721	0,300
" 50 m	0,99840	0,9730	1,554	0,436

Quelques traits se dégagent nettement : θ' et ν augmentent systématiquement avec le support (mais ν croît beaucoup moins vite que celui-ci : il varie de 1 à 2 ou 3 et non de 1 à 10 lorsque le support passe de 5 m à 50 m : d'où le rejet de l'hypothèse d'indépendance des échantillons). Dans l'ensemble b diminue lorsque le support augmente (sauf pour B-25 et B-50). L'évolution de θ est plus plus indécise, sauf dans le cas B où l'augmentation est assez régulière.

(N.B. : les nécessités du regroupement ont fait que les données retenues pour l'ensemble des plages ne sont pas exactement les mêmes que dans le premier ajustement, d'où les valeurs légèrement différentes des paramètres).

IV - LA LOI DE SICHEL-PASCAL

Venons en maintenant à la loi de Sichel-Pascal à 4 paramètres, loi que nous avons déjà rencontrée au premier paragraphe : sa fonction génératrice se déduit de (I-3') par simple suppression du premier facteur (binomial négatif). Nous l'écrirons sous la forme suivante (avec ν au lieu de 2ν en exposant) :

$$\text{(IV-1)} \qquad G(s) = \left(\frac{a + \sqrt{1-\tau}}{a + \sqrt{1-\tau \frac{qs}{1-ps}}} \right)^{\nu}$$

ou encore, avec

$$\boxed{\theta = p + q\,\tau}$$

$$\text{(IV-1')} \qquad G(s) = \left(\frac{a + \sqrt{1-\tau}}{a + \sqrt{\frac{1-\theta s}{1-ps}}} \right)^{\nu}$$

Ce paramètre θ a la même signification que dans les cas précédents : pour n tendant vers l'infini, le comportement asymptotique de la probabilité p_n est en $\theta^n/n^{3/2}$. On a évidemment

$$\nu > 0 \qquad 0 \leq \tau \leq 1 \qquad p \leq 1 \qquad a + \sqrt{1-\tau} > 0$$

Le paramètre p peut prendre des valeurs négatives (pas trop fortes en valeur absolue), bien que l'interprétation probabiliste (loi de Pascal) ne subsiste plus dans ce cas.

Pour établir les propriétés de cette loi, le plus simple est de partir de la loi de Sichel à 3 paramètres :

$$\text{(IV-2)} \qquad \gamma(s) = \left(\frac{a + \sqrt{1-\tau}}{a + \sqrt{1-\tau\, s}} \right)^{\nu}$$

et de faire la substitution

$$s \to \frac{qs}{1-ps}$$

Interprétation probabiliste

L'interprétation probabiliste de cette transformation est très simple : si N_S désigne une variable obéissant à la loi de Sichel (IV-2), la variable N de loi (IV-1) peut s'écrire :

$$N = \sum_{i=1}^{N_S} N_i$$

(et $N = 0$ si $N_S = 0$), les N_i désignant des pascaliennes (loi $qs/(1-ps)$) indépendantes entre elles et de N_S : nous avons donc affaire à la somme d'un nombre N_S aléatoire (sichélien) de variables pascaliennes.

Nous avons déjà rencontré dans 4-F la décomposition de la loi de Sichel en paquets, la loi du nombre d'objets par paquets étant l'hypergéométrique :

$$\left(\frac{1 - \sqrt{1-\tau\, s}}{1 - \sqrt{1-\tau}}\right)$$

Cette décomposition est de la forme :

$$\begin{cases} \gamma(s) = \sum \varpi_n \left(\dfrac{1 - \sqrt{1-\tau\, s}}{1 - \sqrt{1-\tau}}\right)^n \\ \varpi_n = \left(\dfrac{a + \sqrt{1-\tau}}{a+1}\right)^\nu \dfrac{\Gamma(n+\nu)}{n!} \left(\dfrac{1 - \sqrt{1-\tau}}{a+1}\right)^n \end{cases}$$

Il s'agissait donc de la somme d'un nombre aléatoire (binomial négatif) de paquets contenants eux mêmes un nombre aléatoire d'objet obéissant à la loi hypergéométrique ci-dessus. Pour la clarté de l'exposé, nous dirons que ces paquets sont des paquets sicheliens.

Avec la loi de Sichel-Pascal, cette interprétation subsiste : nous avons encore affaire aux mêmes paquets sicheliens, avec la même loi binomiale négative que ci-dessus pour le nombre de ces paquets. Simplement, au lieu d'être constitués d'individus, les paquets sicheliens sont maintenant constitués de sous-paquets pascaliens : chaque sous-paquet contient un nombre d'individus obéissant à la loi de Pascal, le nombre de ces sous-paquets dans un paquet sichelien restant soumis à la même loi hypergéométrique que ci-dessus.

Cette hiérarchie en paquets sicheliens et sous-paquets pascaliens présente un grand intérêt pour la recherche de modèle de changement de support, et permet en particulier de construire assez facilement des lois multivariables. C'est là l'intérêt principal de ce modèle du point de vue de la géostatistique. Dans certains cas, cette décomposition en paquets et sous-paquets ne sera rien de plus qu'un artifice de calcul. Dans le cas qui nous sert de test, l'existence effective des paquets est une réalité géologique bien établie. Les sous-paquets ne sont pas mentionnés, peut-être simplement parce qu'ils ne sont pas observables (on s'attend intuitivement à ce que le rayon de dispersion des individus d'un sous-paquet soit nettement plus petit que celui d'un paquet). On leur attribuerait volontiers comme origine la fragmentation des pierres initiales, soit au cours de leur histoire géologique, soit au cours des processus d'échantillonnage et de traitement. Naturellement, tout ceci est hypothétique. Même en ce qui concerne les paquets, dont l'existence est assurée, nous ne disposons d'aucune information quantitative, et nous n'avons donc aucune preuve expérimentale directe pour affirmer que le nombre de pierres qu'ils contiennent obéit (ou non) à la loi ci-dessus. C'est uniquement sur la loi résultante (IV-1) que porteront les contrôles expérimentaux.

Calcul des paramètres

Venons en maintenant à la tâche ingrate, qui consiste à calculer m, σ^2, p_0, p_1 ainsi que les relations de récurrence (qui porteront cette fois sur 5 termes simultanément). Pour faciliter (un peu) ces calculs laborieux, nous partirons de :

$$G(s) = \gamma\left(\frac{qs}{1-ps}\right)$$

où γ est la loi de Sichel (IV-2). En dérivant, il vient :

$$\begin{cases} G'(s) = \dfrac{q}{(1-ps)^2} \; \gamma'\left(\dfrac{qs}{1-ps}\right) \\[2ex] G''(s) = \dfrac{2\,pq}{(1-ps)^3} \; \gamma'\left(\dfrac{qs}{1-ps}\right) + \dfrac{q^2}{(1-ps)^4} \; \gamma''\left(\dfrac{qs}{1-ps}\right) \end{cases}$$

Les expressions des γ' et γ'' ont été données dans 4-F, et il n'y a pas lieu de les réécrire ici.

Avec s = 1, nous obtenons la <u>moyenne</u> et la <u>variance</u>

$$\text{(IV-3)} \quad \begin{cases} m = \dfrac{\nu\tau}{2q} \; \dfrac{1}{\sqrt{1-\tau}} \; \dfrac{1}{a + \sqrt{1-\tau}} \\[2ex] \sigma^2 = \dfrac{m}{q} \left(1+p + \dfrac{\tau}{2(1-\tau)} + \dfrac{\tau}{2(1-\tau + a\sqrt{1-\tau})}\right) \end{cases}$$

De même, pour <u>les deux premières probabilités</u>, on trouve en faisant s = 0

$$\text{(IV-3')} \quad \begin{cases} p_0 = \left(\dfrac{a + \sqrt{1-\tau}}{a+1}\right)^{\nu} \\[2ex] p_1 = \dfrac{\nu\tau q}{2(a+1)} \; p_0 \end{cases}$$

<u>L'équation diférentielle</u>

Passons à <u>l'équation différentielle</u>. On a vu dans 4-F que $\gamma(s)$ vérifie l'équation :

$$(a^2-1+\tau s)\;\gamma''(s) + \tau\left[\nu + \frac{3}{2} - \frac{a^2}{2} - \left(\nu + \frac{3}{2}\right)\tau s\right] \gamma' - \tau^2 \, \frac{\nu(\nu+1)}{4} \, \gamma = 0$$

Nous allons substituer partout dans cette équation $qs/(1-ps)$ à s et exprimer les dérivées de γ en fonction de celles de G :

$$\left\{\begin{array}{l}\gamma\left(\frac{qs}{1-ps}\right) = G(s)\\[2ex] \gamma'\left(\frac{qs}{1-ps}\right) = \frac{(1-ps)^2}{q}\, G'(s)\\[2ex] \gamma''\left(\frac{qs}{1-ps}\right) = -\frac{2p}{q^2}\,(1-ps)^3\, G'(s) + \frac{(1-ps)^4}{q^2}\, G''(s)\end{array}\right.$$

Le résultat de cette substitution conduit à l'équation différentielle suivante :

$$\text{(IV-4)}\quad \left\{\begin{array}{l}P_4(s)\, G'' + P_3(s)\, G' + P_o\, G = 0\\ \qquad\qquad \text{avec (en posant } \theta = p + \tau q) :\\[1ex] P_4(s) = (1-ps)^2\,(1-\theta s)\,[a^2(1-ps) - (1-\theta s)]\\[1ex] P_3(s) = -2p\,(1-ps)(1-\theta s)\,[a^2(1-ps) - (1-\theta s)]\\ \qquad\qquad + \tau\, q(1-ps)\,[(\nu + \frac{3}{2})(1-\theta s) - \frac{a^2}{2}(1-ps)]\\[1ex] P_o = -\tau^2 q^2\,\frac{\nu(\nu+1)}{4}\end{array}\right.$$

La relation de récurrence

Pour passer à la relation de récurrence (qui est d'ici d'ordre 5) il est commode de poser

$$\left\{\begin{array}{l}\alpha = a^2 - 1\\ \theta = p + \tau q\\ A = (a^2-1)\, p - \tau q\\ C = \nu + \frac{3}{2} - \frac{a^2}{2}\\ D = (\nu + \frac{3}{2})(p + \tau q) - \frac{a^2 p}{2}\end{array}\right.$$

On trouve alors, avec ces notations :

$$
\text{(IV-5)} \quad \left\{ \begin{aligned}
\alpha\, n(n-1)\, p_n = & \left[T_1(n-2)(n-1) + T_2(n-1)\right] p_{n-1} \\
& + \left[T_3(n-2)(n-3) + T_4(n-2) + T_5\right] p_{n-2} \\
& + \left[T_6(n-3)(n-4) + T_7(n-3)\right] p_{n-3} \\
& + T_8(n-3)(n-4)\, p_{n-4}
\end{aligned} \right.
$$

avec :

$$
\text{(IV-5')} \quad \left\{ \begin{aligned}
T_1 &= A + \alpha\theta + 2p\alpha \\
T_2 &= 2p\alpha - \tau q C \\
T_3 &= -(\alpha p^2 + 2p\alpha\theta + 2pA + A\theta) \\
T_4 &= \tau q p C + \tau q p D - 2pA - 2p\alpha\theta - 2p^2\alpha \\
T_5 &= \tau^2 q^2 \frac{\nu(\nu+1)}{4} \\
T_6 &= p^2\alpha\theta + p^2 A + 2pA\theta \\
T_7 &= 2p^2\alpha\theta + 2p^2 A + 2pA\theta - \tau p q D \\
T_8 &= -p^2 A\theta
\end{aligned} \right.
$$

<u>N.B.</u> : Ces relations sont utilisables dès que $n \geq 2$ (étant entendu que $p_{-1} = p_{-2} = 0$), de sorte qu'il suffit, pour amorcer la récurrence, de calculer p_o et p_1.

Sauf erreur de transcription, ces formules sont exactes : je les ai testées dans quelques centaines de cas de figure. Elles marchent numériquement très bien, sauf toujours dans le cas où a est voisin de 1 (α voisin de 0 dans la relation (IV-5).

Pour $|a^2 - 1| \leq 0{,}7$ ou $0{,}8$, le calcul numérique n'est pas possible. Lorsque les ajustements indiquaient que la valeur de a allait tomber trop près de 1, je me suis résigné à prendre exactement a = 1 (c'est-à-dire à travailler avec une famille à 3 paramètres seulement). Comme on le verra, ce cas se présente malheureusement assez souvent : les ajustements correspondants, déjà bons comme on le verra, auraient été encore meilleurs, et surtout l'évolution des paramètres en fonction de la taille du support aurait été plus significative.

Pour a = 1 (exactement) le terme en p_n disparait dans (IV-5), et il reste une relation de récurrence à 4 termes seulement, dont le calcul ne pose aucun problème numérique. Dans le cas très difficile où a est voisin de 1, je pense qu'il conviendrait d'utiliser un développement limité par rapport à (a - 1). Partant de :

$$\left(\frac{1}{a+\sqrt{1-x}}\right)^{\nu} = \sum_{k\geq 0} (-1)^k \frac{\Gamma(\nu+k)}{\Gamma(\nu)k!} (a-1)^k \left(\frac{1}{1+\sqrt{1-x}}\right)^{\nu+k}$$

on voit, en effet, que les probabilités $p_n(a,\nu)$ correspondant aux valeurs a (voisin de 1) et ν des paramètres se mettent sous la forme

$$\text{(IV-6)} \qquad p_n(a,\nu) = \sum_{k\geq 0} (-1)^k \frac{\Gamma(\nu+k)}{k!\Gamma(\nu)} \left(\frac{a-1}{a+\sqrt{1-\tau}}\right)^k p_n(1,\nu+k)$$

Lorsque a est voisin de 1, cette série doit converger assez vite : τ en effet est, lui aussi, voisin de 1 dans les cas que j'ai étudiés, de sorte que le terme $(a-1)/(a+\sqrt{1-\tau})$ est à peu près égal à $(a-1)/2$, ce qui assure une convergence rapide dès que $|a-1| \leq 1$. Je n'ai malheureusement pas pu essayer cette méthode, faute de moyens de calcul au moment où j'ai effectué ce travail : ceci est certainement à reprendre.

Technique d'ajustement utilisée (paramétrage en ν)

J'ai utilisé la technique d'ajustement, assez sommaire, que voici : il est facile d'exprimer les trois paramètres τ, p et a en fonction de ν, de m, et des 2 premières probabilités p_0 et p_1. Dans un premier temps, on donne à m, p_0 et p_1 les valeurs déduites de la distribution empirique, et, faisant varier ν , on cherche le minimum du χ^2. Dans un deuxième temps, on peut faire varier p_o, p_1 et m et recommencer l'opération, de façon à améliorer le résultat.

Voici les relations donnant τ, p et a en fonction de ν, p_o, p_1 et m. Posons, pour simplifier les écritures :

$$x = \sqrt{1-\tau} \ ; \qquad A = \frac{p_1}{p_o m} \ ; \qquad \beta = \left(\frac{1}{p_o}\right)^{1/\nu}$$

Compte tenu des relations (IV-3) et (IV-3'), nous trouvons :

$$\begin{cases} \beta = \dfrac{a+1}{a+x} \ ; \qquad \beta A = q^2 x \\ m = \dfrac{\nu(\beta-1)}{2q} \ \dfrac{(1+x)}{x} \end{cases}$$

On en tire :

$$\begin{cases} a = \dfrac{1-\beta x}{\beta - 1} \\ q = \dfrac{\nu(\beta-1)}{2m} \ \dfrac{1+x}{x} \end{cases}$$

Substituant cette valeur de q dans $\beta A = q^2 x$, on trouve que x vérifie une équation du second degré

$$(1+x)^2 = \frac{4m^2 \beta A}{\nu^2(\beta-1)^2} \ x$$

Posons, pour abréger :

$$C = \frac{2m^2 \beta A}{\nu^2(\beta-1)^2} = \frac{2m\beta}{\nu^2(\beta-1)^2} \ \frac{p_1}{p_0}$$

Il vient donc :

(IV-7) $$x^2 + 2(1-C)\, x + 1 = 0$$

et (x étant ≤ 1)

(IV-7') $$x = C-1 - \sqrt{C(C-2)}$$

ce qui suppose $C \leq 2$. Connaissant $x = \sqrt{1-\tau}$, les autres paramètres s'en déduisent sans peine.

La condition $C \leq 2$, c'est-à-dire

$$\frac{m\,\beta}{\nu^2(\beta-1)^2}\;\frac{p_1}{p_o} \leq 1$$

est une limitation absolue (p_1, p_o et m étant donnés, toutes les valeurs ne sont pas possibles pour ν).

Une autre limitation est la suivante. D'après (IV-7), C et x varient en sens inverse (pour x variant de 0 à 1). Ainsi (m, p_1 et p_o étant donnés, et ν variant, ainsi que β qui est une fonction de ν) la valeur maximale possible pour τ (minimum de x) correspond au maximum de C : il est facile de voir que ce maximum de C correspond à $\nu = \infty$. Si l'on pose

$$\alpha = \log 1/p_o$$

on a, en effet, $\beta = e^{\alpha/\nu}$ et

$$\frac{\beta}{\nu^2(\beta-1)^2} = \frac{e^{\alpha/\nu}}{\nu^2(e^{\alpha/\nu}-1)^2}$$

Pour $\nu \to \infty$, cette expression tend vers sa valeur maximale, qui est

$$\frac{1}{\alpha^2} = \frac{1}{(\log p_o)^2}$$

Ainsi, en prenant

$$C_o = \frac{2m}{(\log p_o)^2}\;\frac{p_1}{p_o}$$

$$x_o = C_o - 1 - \sqrt{C_o(C_o - 2)}$$

on trouve :

$$\tau \leq \tau_o = \sqrt{1 - x_o}$$

Autrement dit : pour des valeurs de m, p_o et p_1 données, tout n'est pas possible, on ne peut pas choisir τ aussi voisin de 1 que l'on veut (cette circonstance explique que certains de nos ajustements ne donnent pas une représentation entièrement satisfaisante des queues de distribution).

<u>REMARQUE</u> : Nous avons utilisé le paramètre $A = p_1/m\ p_o$. Il est intéressant de remarquer qu'il est nécessairement ≤ 1. Plus généralement, <u>pour toute loi discrète indéfiniment divisible</u>, on a

$$\frac{p_1}{p_o} \leq m$$

<u>l'égalité n'étant réalisée que dans le cas d'une loi de Poisson</u>.

En effet, si G est la fonction génératrice d'une loi indéfiniment divisible, log G et G'/G sont des fonctions croissantes de s. On a donc

$$\frac{G'(0)}{G(0)} \leq \frac{G'(1)}{G(1)}$$

c'est-à-dire $p_1/p_o \leq m$. L'égalité n'est possible que si G'/G est constante, donc log G linéaire en 1, ce qui correspond à la loi de Poisson.

<u>Comportement asymptotique</u>

Pour n grand, p_n est de l'ordre de $\theta^n/n^{3/2}$ $(\theta = p + \tau q)$, sauf dans le cas d'exception $a = 0$ où la loi de Sichel-Pascal se réduit à une binomiale négative d'indice $\nu/2$. Pour a voisin de 0, le comportement limite en $\theta^n/n^{3/2}$ ne se réalise qu'aux très grandes valeurs de n. Dès que a est nettement différent de 0, ce comportement est atteint assez rapidement : en pratique, dès n = 40 pourvu que a soit supérieur à 0,4 ou même 0,3.

On a donc, asymptotiquement

$$p_n \simeq C \frac{\theta^n}{n^{3/2}}$$

Le paramètre C dépend de τ, p, a et ν de manière assez complexe, mais ne dépend pas de n : par conséquent, l'espérance conditionnelle de la variable N pour N $\geq$ n donné, soit

$$m(N \geq n) = \frac{\sum\limits_{k \geq n} k p_k}{\sum\limits_{k \geq n} p_k}$$

ne dépend (pratiquement) plus que de θ et de n seulement dès que n est de l'ordre de 40 :

$$m(N \geq n) = \frac{\sum\limits_{k \geq n} \theta^K / \sqrt{K}}{\sum\limits_{k \geq n} \theta^K / K^{3/2}}$$

Choisissant, par exemple, la valeur n = 40, il apparait que m(40) est une fonction (évidemment croissante) du seul paramètre θ : d'où une possibilité d'évaluer très vite l'ordre de grandeur de ce paramètre à partir d'une distribution empirique.

Cette valeur m(40) peut être tabulée en fonction de θ, en utilisant par exemple la loi de référence

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

(soit a = ν = 1, p = 0, $\tau = \theta$). On obtient la table suivante, suffisante pour apprécier des ordres de grandeurs :

TABLE 1

θ	$m(\geq 40)$	θ	$m(\geq 40)$
0,80	43,46	0,991	84,16
0,85	44,71	0,992	87,62
0,90	47,02	0,993	91,81
0,92	48,60	0,994	97,03
0,94	51,05	0,995	103,78
0,95	52,85	0,996	112,96
0,96	55,37	0,997	126,48
0,97	59,17	0,998	149,27
0,975	61,95	0,9985	168,45
0,980	65,77	0,9990	200,97
0,9825	68,31	0,9995	274,35
0,9850	71,48	0,9999	584,78
0,9875	75,60	0,99999	1 800,32
0,9900	81,25	0,999999	5 644,79

(pour les valeurs $\theta < 0,8$, on a à peu près $m(40) = 40 + \frac{\theta}{1-\theta}$).

Dans les ajustements qui suivent, on trouvera la valeur exacte de m(40) calculée pour 30 cas de figures différents. On pourra constater que la table ci-dessus donne un bon ordre de grandeur (les deux seuls cas d'exception, A-50 et B-50 correspondent à de très petites valeurs de a).

V - EXEMPLES D'AJUSTEMENTS A LA LOI DE SICHEL-PASCAL

Trente ajustements ont été effectués sur les mêmes données que précédemment, à savoir pour les 5 supports différents (5, 10, 15, 25 et 50), la distribution générale (toutes plages confondues, en abrégé TP) ainsi que les distributions relatives aux plages homogènes A, B, D, E et F (la plage C a été laissée de côté à cause de son effectif trop faible).

29 sur 30

Le résultat est assez spectaculaire. Le test χ^2, en effet, conduit à l'acceptation du modèle dans 29 cas sur 30 (si l'on prend comme critère $P(\chi^2 \geq .) \geq 5/100$), et peut être considéré comme très bon (si l'on prend comme critère $P(\chi^2 \geq .) > 50/100$) dans 19 cas sur 30.

Le seul cas qui concluerait au rejet (d'ailleurs catégorique) est celui de la population totale sur un support de 5 m (TP5). Il est intéressant d'ailleurs de noter, à propos de cette population totale, que l'**évolution** en fonction du support est exactement l'inverse de celle que nous avons observée avec la loi de Schiel : ici, en effet, l'ajustement s'améliore lorsque le support augmente. Absolument détestable pour le support 5 m, il est bon pour les supports 10 et 15 et très bon pour les supports 25 et 50.

A dire vrai, cette population "toutes plages" étant hétérogène, on ne peut pas attribuer grande signification à l'échec de l'ajustement TP-5. Si donc on élimine la population TP, on constate que le modèle doit être considéré comme acceptable dans 25 cas sur 25, assez bon ($P(\chi^2 \geq .) \geq 10\%$) dans 23 cas sur 25, bon ($P(\chi^2) \geq 25\%$) dans 20 cas sur 25, très bon ($P(\chi^2) \geq 50\%$) dans 17 cas sur 25 et exceptionnel ($P(\chi^2) > 90\%$) dans 5 cas sur 25.

Si nos ajustements pouvaient être considérés comme indépendants (ce qui n'est d'ailleurs pas le cas, puisque, pour chaque plage, il s'agit des mêmes échantillons regroupés sur des supports différents), nous devrions nous attendre à ce que les variables $P(\chi^2 \geq .)$ obéissent à une loi de distribution uniforme. Il est amusant de constater qu'il en est bien à peu près ainsi, ou du moins que le test χ^2 ne conclut pas à un rejet :

.../

Test χ^2 sur les variables $P(\chi^2 \geq .)$ - loi uniforme			
Classe	Fréquence observée	Fréquence calculée	χ^2
0 à 0,2	6	6	0,0
0,2 à 0,4	3	6	1,5
0,4 à 0,6	3	6	1,5
0,6 à 0,8	9	6	1,5
0,8 à 1,0	9	6	1,5
			$\chi^2 = 6,0$
Avec 4 degrés de liberté : $P(\chi^2 \geq 6) = 20\%$			

On trouvera en annexe le détail de ces 30 ajustements.

.../

	Tests χ^2 (SIPA)			
	N	χ^2	Degrés de liberté	$P(\chi^2 \geq .)$ en %
TP-5	12 021	63,80	36	0,29 %
TP-10	5 857	40,84	37	30 5 %
TP-15	3 807	43,52	37	21,4 %
TP-25	2 160	31,33	37	73,2 %
TP-50	922	32,06	33	51 4 %
A-5	2 357	21,31	14	9,4 %
A-10	1 145	15,41	14	35,0 %
A-15	747	9,40	14	80,5 %
A-25	422	5,56	14	97,6 %
A-50	178	12,73	9	17,5 %
B-5	3 028	22,76	22	41,5 %
B-10	1 474	18,26	21	63,2 %
B-15	964	15,48	21	79,8 %
B-25	540	13,74	18	74,6 %
B-50	227	18,66	12	9,7 %
D-5	2 790	20,10	14	12,7 %
D-10	1 370	12,70	16	69,5 %
D-15	897	19,09	14	16,2 %
D-25	518	7,69	14	90,5 %
D-50	235	6,06	13	94,4 %
E-5	1 934	14,00	20	83,0 %
E-10	944	13,37	20	86,1 %
E-15	612	13,12	19	83,2 %
E-25	347	16,67	19	61,2 %
E-50	146	8,50	11	66,8 %
F-5	14406	18,68	21	60,6 %
F-10	677	12,12	21	93,6 %
F-15	429	10,72	19	93,3 %
F-25	242	17,64	18	48,0 %
F-50	97	5,31	7	62,2 %

Moyenne des ≥ 40

On a donc un peu envie de chanter victoire. Pour atténuer cet excès d'optimisme, observons cependant que le comportement de l'extrême queue de distribution n'est pas toujours parfait. Un bon critère pour en juger nous est fourni par la comparaison des valeurs calculées et expérimentales de la moyenne des échantillons cpntenant au moins 40 pierres. On trouvera ces valeurs sur le tableau suivant. Sur le même tableau, on a également indiqué la valeur du paramètre θ ajusté, pour permettre la comparaison avec la Table 1 : on voit que les valeurs de θ et de m (≥ 40) calculées sont, effectivement, compatibles avec cette table (aux deux exceptions près A-50 et B-50 pour lesquelles a est très petit : a = 0,1 pour A-50 et a = 0,26 pour B-50).

Pour la population totale (TP) les valeurs de m (≥ 40) calculées sont systématiquement trop faibles (sauf pour TP-50). Si nous laissons de côté cette population hétérogène, l'accord apparait, dans l'ensemble, comme assez bon, sauf peut-être pour la population A où les valeurs calculées sont trop faibles. Pour les autres populations, l'accord est meilleur. Pour les populations B et E, il est très bon et les fluctuations ne sont pas systématiques. Par contre, on trouve des valeurs calculées systématiquement trop faibles pour TP, A et D, et systématiquement trop fortes pour D et F.

On note aussi que l'accord entre les m(40) calculées et observées est d'autant meilleur, dans l'ensemble, que le χ^2 est plus petit, ce qui est d'ailleurs assez naturel. Au total, les résultats de ce deuxième test sont assez favorables, sans cependant que ce succès très réel présente un aspect aussi massif que pour le test χ^2.

.../

Moyenne des Echantillons ≥ 40			
	θ	m (≥40) calculée	m (≥40) expérimentale
TP-5	0,98989	80,59	107,60
TP-10	0,99289	90,99	108,28
TP-15	0,99360	94,66	110,21
TP-25	0,99486	103,32	119,75
TP-50	0,99760	142,37	143,19
A-5	0,99428	98,15	117,00
A-10	0,99593	111,46	133,44
A-15	0,99678	123,22	133,28
A-25	0,99659	121,39	144,00
A-50	0,99562	129,53	165,24
B-5	0,99214	88,03	96,40
B-10	0,99495	103,46	104,24
B-15	0,99603	113,65	110,87
B-25	0,99763	140,68	126,72
B-50	0,99751	152,71	146,13
D-5	0,97068	59,36	82,69
D-10	0,97948	65,10	85,38
D-15	0,98260	68,24	79,31
D-25	0,98845	77,33	82,29
D-50	0,99369	95,49	96,72
E-5	0,97209	60,21	63,52
E-10	0,98269	68,46	62,65
E-15	0,98711	75,09	68,72
E-25	0,99137	86,04	75,90
E-30	0,99496	107,58	104,84
F-5	0,97778	63,99	59,43
F-10	0,98715	75,06	68,03
F-15	0,98929	80,09	76,60
F-25	0,99278	98,63	90,73
F-50	0,99740	145,02	130,84

Evolution des Paramètres.

Voyons maintenant l'évolution des paramètres en fonction de la taille du support. Celle du paramètre θ (qui fixe le comportement asymptotique des p_n) ressort assez clairement du tableau précédent : il y a une augmentation systématique de θ avec la taille du support. Il ne s'agit nullement d'un artefact de nos ajustements, puisque les m (≥ 40) expérimentaux présentent la même tendance systématique à l'augmentation. Cet effet est, du reste, bien celui que l'on attendait intuitivement (dans les ajustements à la loi de Schiel, l'évolution de θ restait beaucoup plus indécise. Il y a là, sans doute, un argument supplémentaire assez fort en faveur du modèle de Sichel-Pascal).

L'évolution des autres paramètres (τ, p, a et ν) est présentée sur le tableau ci-après. Un seul trait ressort de manière vraiment nette : l'augmentation de τ avec la taille du support (à rapprocher de l'augmentation de θ). La seule population à présenter une évolution cohérente pour les autres paramètres (augmentation de p et de ν, et diminution de a) est la population hétérogène "toutes plages", peut-être simplement d'ailleurs parce que son effectif plus nombreux a permis des ajustements plus fiables. pour les autres population, cette évolution reste indécise, parfois même un peu incohérente (par exemple, pour les populations A, E et F, l'apparition de valeurs négatives pour p aux grands supports, succédant brusquement à une tendance antérieure à la croissance, surprend quelque peu). On notera que beaucoup d'ajustements ont été faussés par la contrainte qui nous a imposé de prendre a = 1 (alors que a = 1,3 ou a = 0,5 aurait peut-être été bien meilleur), pour éviter l'instabilité numérique signalée plus haut : cette contrainte a certainement perturbé l'ajustement des autres paramètres, et masqué leur évolution véritable.

L'estimation des paramètres p, a, et ν est probablement très incertaine. Dans certains cas, la qualité de l'ajustement est particulièrement peu sensible à la variation de certains d'entre eux. Par exemple, pour la population D, on avait l'impression que le paramètre ν pouvait être pris à peu près

Evolution des Paramètres				
	τ	p	a	ν
TP-5	0,98685	0,2312	4,0602	2,3986
TP-10	0,99011	0,2811	2,8191	2,7105
TP-15	0,99075	0,3080	2,1503	2,8248
TP-25	0,99154	0,3928	1,6629	3,2313
TP-50	0,99534	0,4851	1,6350	4,9000
A-5	0,99138	0,3356	2,1574	0,8662
A-10	0,99273	0,4394	2,4750	1,5067
A-15	0,99497	0,3609	1	0,9665
A-25	0,99255	0,5420	1	1,4237
A-50	0,99570	-0,0194	0,1008	0,8726
B-5	0,99165	0,0583	1	0,6452
B-10	0,99406	0,1505	1	0,9928
B-15	0,99440	0,2910	1	1,2359
B-25	0,99636	0,3495	1	1,7668
B-50	0,99689	0,1998	0,2561	1,3502
D-5	0,96516	0,1585	20,4205	10,0076
D-10	0,97293	0,2418	15,2522	11,9983
D-15	0,97809	0,2059	14,9848	17,2110
D-25	0,98329	0,3089	13,5186	20,0000
D-50	0,99108	0,2928	8,6224	20,0000
E-5	0,96838	0,1174	1,4606	1,7000
E-10	0,97924	0,1665	4,1067	6,5613
E-15	0,98558	0,1062	2,3291	5,4702
E-25	0,98895	0,2188	4,2012	13,0977
E-50	0,99527	-0,0666	3,8940	24,9433
F-5	0,97163	0,2164	1	1,3055
F-10	0,98037	0,3453	1,9586	3,3547
F-15	0,98544	0,2646	1,4753	3,8096
F-25	0,99419	-0,0579	1	4,4877
F-50	0,99814	-0,2921	2,1719	14,7563

quelconque, pourvu qu'il ne soit pas petit : a variait alors, en gros, proportionnellement à ν, τ et p ne bougeant presque pas (c'est ce qui explique les valeurs quasiment symboliques $\nu = 20$ pour D-25 et D-50 : on aurait pu prendre $\nu = 15$ ou 25 sans modifier la qualité de l'ajustement, d'ailleurs très bon).

La conclusion de ce tableau un peu décevant est sans doute qu'il faudra mettre au point des techniques plus fiables d'estimation des paramètres, avant de pouvoir étudier sérieusement leur évolution avec le support. On trouvera dans le prochain paragraphe quelques suggestions pour des modèles possibles de changement de support : mais il sera difficile de les confronter avec les valeurs incertaines dont nous disposons.

VI - LA LOI CONTINUE DE SICHEL-PASCAL.

Propriétés d'Invariance.

Toutes les lois que nous avons rencontrées possèdent l'intéressante propriété d'être invariantes par décimation. Par décimation, j'entends l'opération qui consiste à tuer avec une probabilité ϖ donnée chaque individu d'une population : si la population a un effectif N = n donné, la nouvelle population obéit à la loi binomiale $(\chi + \varpi s)^n$ $(\chi = 1-\varpi)$. Si la population N est elle-même aléatoire, avec une fonction génératrice G(s), la population décimée admet la loi $G(\chi + \varpi s)$. Il s'agit donc de la transformation :

$$s \to \chi + \varpi s \qquad \text{ou} \qquad 1-s \to \varpi(1-s)$$

Cette propriété est d'ailleurs fort intéressante dans les applications analogues à celles que nous étudions ici : on n'est jamais certain, en effet, que les procédés de traitement utilisés récupèrent toutes les pierres réellement contenues dans un échantillon. Mais, si l'on admet que les causes de pertes agissent indépendamment sur chacune des pierres de l'échantillon, chaque pierre ayant la même probabilité ϖ d'être perdue, il apparaît que la population observée se déduit de la population réelle par décimation : il;en résulte que ces pertes affecteront les paramètres de nos lois mais non leur forme, puisque celle-ci est invariante par décimation.

La loi de Sichel-Pascal possède une propriété d'invariance beaucoup plus forte. La fonction génératrice :

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

est invariante quant à sa forme, non seulemant par la décimation $s \to \chi + \varpi s$, mais également par la transformation de Pascal

$$s \to \frac{\chi\, s}{1 - \varpi s} \qquad \text{ou} \qquad s \to \frac{\chi}{1 - \varpi s}$$

(selon que l'on considère la loi de Pascal sans zero ou avec zero), transformation consistant à remplacer chaque individu par un paquet pascalien. Combinant décimation et transformation de Pascal, nous obtenons la transformation suivante :

$$s \to \chi + \varpi \frac{\chi' s}{1 - \varpi' s}$$

que nous appellerons pascalisation. Comme le produit de deux pascalisations successives est encore une pascalisation, ces transformations forment un demi-groupe (plus vaste que le demi-groupe des décimations).

Il est clair, d'après l'expression même de sa fonction génératrice, que la loi de Sichel-Pascal est invariante par le demi-groupe des pascalisations. Cette propriété est intéressante dans les applications que nous avons en vue, car la pascalisation représente assez bien non seulement les pertes (ce que faisait déjà la décimation) mais également les fragmentations auxquelles nos pierres ont pu être soumises par suite de processus naturels et/ou de nos procédés de traitement : ces pertes et ces fragmentations laisseront invariante la forme de la loi de Sichel-Pascal (tout en affectant, évidemment, les paramètres que nous lui attribuons). C'est peut-être cette robustesse vis-à-vis de la pascalisation qui explique le succès de cette loi dans les cas que nous avons étudiés.

<u>Version continue de la Pascalisation</u>.

Voici, maintenant, une autre manière d'introduire les pascalisations. Partons d'une loi quelconque (discrète ou non, mais concentrée sur $x \geq 0$) caractérisée par sa transformée de Laplace $\Phi(\lambda)$. Changeant λ en $a(1-s)$, nous obtenons <u>le mélange poissonien</u> correspondant, qui admet la fonction génératrice :

$$G(s) = \Phi(a(1-s)) \tag{VI-2}$$

Mais, inversement, à partir d'une loi discrète, définie par sa fonction génératrice :

$$G(s) = \Sigma\, p_n\, s^n$$

nous pouvons former <u>le mélange gamma</u>

$$\Phi(\lambda) = \Sigma\, p_n \left(\frac{c}{c+\lambda}\right)^n = G\left(\frac{c}{c+\lambda}\right)$$

(noter que cette loi admet <u>un atome</u> p_0 en $x = 0$, et admet pour le reste une densité continue). Il est facile de voir qu'en itérant ces opérations successives de mélanges poissoniens et de mélange gamma, on constitue un demi-groupe de transformation qui généralise la pascalisation.

Partant d'une loi discrète, et effectuant d'abord un mélange gamma, puis un mélange poissonien, on obtient la transformation de Pascal particulière $s \to q/(1-ps)$. En réitérant l'opération, on obtient la transformation générale que nous avons appelée pascalisation.

Partant, maintenant, d'une loi quelconque (discrète ou non), effectuons d'abord un mélange poissonien, puis un mélange gamma :

nous obtenons la transformation

$$\lambda \to a \frac{\lambda}{c+\lambda}$$

qui forme évidemment un demi-groupe, et constitue un analogue continu de la pascalisation.

La Loi continue de Sichel-Pascal.

Naturellement, la loi continue de SICHEL-PASCAL, d'où la loi discrète se déduit par mélange poissonien, restera elle-même invariante par la transformation ci-dessus : il en résulte une propriété fort intéressante de ces lois : une même loi de Sichel-Pascal continue peut, indifféremment, être considérée comme la mère d'une loi discrète, qui s'en déduit par mélange poissonien, ou comme la fille d'une autre loi discrète, dont elle se déduit par mélange gamma. Ainsi, notre famille de lois possède un arbre généalogique qui se prolonge à l'infini dans les deux sens.

Du point de vue pratique, on peut utiliser cette propriété pour calculer commodément la loi continue dont une loi discrète donnée est la fille par mélange poissonien : on calculera cette loi continue en faisant un mélange gamma à partir d'une autre loi discrète; de Sichel-Pascal convenablement choisie.

En fait, le plus simple est de prendre comme ancêtre une loi discrète de Sichel à 3 paramètres, admettant donc une fonction génératrice de la forme :

$$\text{(VI-3)} \qquad G(s) = \left(\frac{A + \sqrt{1-T}}{A + \sqrt{1-Ts}}\right)^{\nu}$$

Nous voulons que, par les substitutions successives :

$$s \to \frac{c}{c+\lambda} \quad ; \quad \lambda \to (1-s)$$

cette loi devienne

$$\text{(VI-4)} \qquad G(s) = \left(\frac{a + \sqrt{1-\tau}}{a + \sqrt{\frac{1-\theta s}{1-ps}}}\right)^{\nu} \qquad (\theta = p + q\tau)$$

Il est facile de voir qu'il suffit, pour cela, de prendre dans (VI-3)

$$\text{(VI-5)} \qquad A = \left(\sqrt{\frac{p}{\theta}}\right)a \quad ; \quad T = \frac{\theta - p}{q\ \theta} = \frac{\tau}{\theta} \quad ; \quad c = \frac{q}{p}$$

Avec ces valeurs des paramètres A, T et c, on voit donc que la loi continue dont la loi (VI-4) se déduit par poissonisation admet la transformée de Laplace :

$$\Phi(\lambda) = \left(\frac{A + \sqrt{1-T}}{A + \sqrt{1 - T\frac{c}{c+\lambda}}}\right)^{\nu}$$

Il n'est pas très difficile d'en déduire l'expression de la loi continue correspondante. Désignons, en effet, par ϖ_n les probabilités affectées à la loi discrète mère :

$$\left(\frac{A + \sqrt{1-T}}{A + \sqrt{1-Ts}}\right)^{\nu} = \sum_{n \geq 0} \varpi_n\ s^n$$

(les expressions explicites des ϖ_n ont été données dans 4-F : leur calcul est facile à partir de la relation de récurrence à 3 termes que vérifie l'ancêtre Sichelienne). Dès lors, la transformée de Laplace de la loi continue de Sichel-Pascal s'écrit :

$$\Phi(\lambda) = \varpi_o + \sum_{n=1}^{\infty} \varpi_n \left(\frac{c}{c+\lambda}\right)^n$$

Il n'est pas très difficile d'en déduire la loi de probabilité correspondante. Elle admet <u>un atome</u> :

$$\varpi_o = \left(\frac{A + \sqrt{1-T}}{A + 1}\right)^{\nu} = \left(\frac{a\sqrt{p} + \sqrt{\frac{p}{q}}\,(1-\theta)}{a\sqrt{p} + \sqrt{\theta}}\right)^{\nu}$$

en x = 0, et pour le reste la densité :

$$f(x) = \sum_{n=1}^{\infty} \varpi_n \frac{c^n\ x^{n-1}}{(n-1)!}\ e^{-cx}$$

Sous cette forme, le calcul numérique de la loi mère ne pose pas de problèmes particuliers. On peut aussi donner un développement en série double, analogue à celui dont la loi continue de Schiel a fait l'objet, mais cela semble nettement moins intéressant.

VII - SUGGESTIONS POUR UN MODELE DE CHANGEMENT DE SUPPORT.

Si nous désignons par N_V et N_v des variables discrètes (nombre de pierres, par exemple) attachées aux supports V et v, avec $v \leq V$, la Géostatistique nous enseigne qu'il existe nécessairement une loi à deux variables discrètes N_V et N_v , telle que l'on ait :

$$E(N_v/N_V) = \frac{v}{V} N_V$$

Ainsi, l'espérance conditionnelle de la variable attachée au petit support doit, pour cette loi à deux variables, être proportionnelle à la variable attachée au grand support. Or nous disposons de trois transformations simples vérifiant cette condition :

a/ Allant dans le sens $V \to v$ (du grand support vers le petit) on peut effectuer une <u>pascalisation</u> (transformation dépendant de deux paramètres).

b/ Toujours dans le sens $V \to v$, on peut effectuer une <u>décimation des paquets Sicheliens</u> (1 paramètre).

c/ Les transformations précédentes laissent invariant le paramètre ν. On peut leur adjoindre la transformation suivante, qui, à l'inverse, affecte le seul paramètre ν, laissant les trois autres paramètres invariants. Dans le sens $v \to V$ (supports croissants), cette transformation consiste simplement dans <u>l'addition de variables indépendantes</u> : de manière générale, à partir d'une loi indéfiniment divisible :

$$G_\nu = e^{\nu\phi}$$

où ϕ reste constant, tandis que ν varie, on a l'équivalence en loi :

$$X_{\nu+\nu'} \overset{\mathcal{L}}{\equiv} X_\nu + X_{\nu'}$$

où X_α désigne évidemment la variable de loi G_α. Il est remarquable que la transformation inverse (dans le sens $V \to v$ des supports décroissants) possède aussi la propriété requise quant aux espérances conditionnelles : il n'est pas très difficile, en effet, de montrer que si l'on pose

$$Z = X_\nu + X_{\nu'}$$

où X_ν et $X_{\nu'}$ sont indépendantes, l'espérance conditionnelle de X_ν à Z fixée est simplement :

$$E(X_\nu/Z) = \frac{\nu}{\nu+\nu'}\, Z$$

Cette transformation simple par addition de variables indéfiniment divisibles est, d'ailleurs, celle qui convient dans le cas particulier où les variables affectées à des supports disjoints peuvent être considérées comme indépendantes.

Au total, si nous combinons entre elles les trois opérations précédentes, nous obtenons un demi-groupe de transformations qui dépend de 4 paramètres : mais c'est exactement ce qui nous convient, puisque nous voulons transformer des lois qui dépendent justement de quatre paramètres.

Comme la transformation c/ qui agit sur ν et laisse les 3 autres paramètres invariants, est indépendante des transformations a/ et b/, qui laissent, au contraire, ν invariant, il nous suffit pour l'instant d'établir les formules relatives au demi-groupe à 3 paramètres engendré par les transformations a/ et b/.

Pour cela, nous allons changer les notations, et substituer aux paramètres τ (ou θ), p et a des expressions plus maniables. De fait, toutes nos transformations, considérées comme agissant sur l'argument s d'une fonction génératrice G(s), laissent nécessairement invariant le point s = 1. Nous allons donc adopter une écriture mettant en évidence l'expression 1-s.

Posant :

$$\text{(VII-1)} \qquad b = \frac{a}{\sqrt{1-\tau}} \quad ; \quad \alpha = \frac{\theta}{1-\theta} = \frac{p+\tau q}{q(1-\tau)} \quad ; \quad \beta = \frac{p}{q}$$

la fonction génératrice G(s) de la loi de Sichel-Pascal se met sous la forme que nous souhaitons :

$$\text{(VII-2)} \qquad G(s) = \left(\frac{1+b}{b + \sqrt{\dfrac{1+\alpha(1-s)}{1+\beta(1-s)}}} \right)^{\nu}$$

La pascalisation, telle que nous l'avons définie, peut s'écrire :

$$1 - s \to \frac{\gamma(1-s)}{1 + r(1-s)}$$

les paramètres r et γ devant simplement vérifier les contraintes :

$$0 \leq \gamma \leq 1 + r$$

Sous la forme usuelle (d'interpretation probabiliste plus aisée) la pascalisation s'écrit

$$s \to \chi_P + \varpi_P \frac{q_P\, s}{1 - p_P\, s}$$

(l'indice P rappelle qu'il s'agit de la pascalisation) On a évidemment $\chi_P = 1 - \varpi_P$, $q_P = 1 - p_P$ et, en principe, ces paramètres, qui représentent des probabilités, doivent rester compris entre 0 et 1). Il est facile de voir que p_P et ϖ_P se relient aux paramètres γ et r ci-dessus (qui sont d'emploi plus facile) par les formules :

$$p_P = \frac{r}{1+r} \qquad \varpi_P = \frac{\gamma}{1+r}$$

ou, en sens inverse

$$r = \frac{p_P}{q_P} \quad ; \quad \gamma = \frac{\varpi_P}{q_P}$$

La loi de composition de deux pascalisations successives est plus facile à écrire en notation (γ,r) : si T est la pascalisation de paramètres (γ,r), et T' la pascalisation (γ',r'), alors le produit T'T (d'abord T, ensuite T') est la pascalisation T" = T'T de paramètres :

$$\gamma'' = \gamma\,\gamma' \quad ; \quad r'' = r' + r\,\gamma'$$

(le demi-groupe n'est pas commutatif : $T'T \neq TT'$).

Effectuons, sur la loi de Sichel-Pascal écrite sous forme (VII-2) la pascalisation T de paramètres r et γ. Alors, le groupement

$$\frac{1+\alpha(1-s)}{1+\beta(1-s)}$$

se transforme en

$$\frac{1+(r+\alpha\gamma)(1-s)}{1+(r+\beta\gamma)(1-s)}$$

Ainsi, les paramètres α' et β' de la loi transformée par la Pascalisation $T(r,\gamma)$ se déduisent des paramètres α et β de la loi initiale selon les formules très simples :

$$\text{(VII-3)} \qquad \begin{aligned} \alpha' &= r + \alpha\gamma \\ \beta' &= r + \beta\gamma \end{aligned}$$

(les paramètres b et ν restant invariants).

Passons maintenant à <u>la décimation des paquets Sicheliens</u>. Pour cela, souvenons-nous que la loi de Sichel-Pascal peut s'obtenir en randomisant l'exposant z d'une fonction génératrice du type

$$e^{-z\left(\sqrt{\frac{1+\alpha(1-s)}{1+\beta(1-s)}}-1\right)}$$

Explicitement :

$$G(s) = \int_0^\infty e^{-z\left(\sqrt{\frac{1+\alpha(1-s)}{1+\beta(1-s)}}-1\right) - cz} \frac{c^\nu z^{\nu-1}}{\Gamma(\nu)} dz$$

avec simplement :

$$c = b + 1$$

Par conséquent, une décimation de probabilité ϖ_S agissant sur les paquets Sicheliens laisse α et β invariants et transforme c en $c' = \frac{c}{\varpi_S}$, soit :

$$\text{(VII-4)} \qquad b' + 1 = \frac{b+1}{\varpi_S}$$

Au total, après une transformation comportant à la fois une décimation ϖ_S des paquets Sicheliens et une pascalisation de paramètres :

$$p_P = \frac{r}{1+r} \qquad \varpi_P = \frac{\gamma}{1+r}$$

on transforme les paramètres (α,β,b) du support V en (α',β',b'), paramètres de v, et l'on trouve :

$$\varpi_S = \frac{b+1}{b'+1} \quad ; \quad \gamma = \frac{\alpha'-\beta'}{\alpha-\beta} \quad ; \quad r = \alpha' - \alpha\gamma = \beta' - \beta\gamma$$

ou, en notations p_P et ϖ_P d'interprétation plus facile :

$$\text{(VII-3)} \qquad \begin{cases} \varpi_S = \dfrac{b+1}{b'+1} \\ p_P = \dfrac{\alpha\beta' - \beta\alpha'}{\alpha-\beta + \alpha\beta' - \beta\alpha'} \\ \varpi_P = \dfrac{\alpha' - \beta'}{\alpha-\beta + \alpha\beta' - \beta\alpha'} \end{cases}$$

Connaissant les paramètres α, β et b (qui se déduisent des paramètres usuels θ, p et a par les formules (VII-1)) du grand support V et ceux du petit support v, soit α', β' et b', les formules (VII-3) permettent de reconstituer la transformation (par décimation de paquets Sicheliens, + une pascalisation) qui, jointe à une transformation de type c/ transformant ν en ν', fait passer de V à v : encore faut-il que ϖ_S, p_P et ϖ_P restent compris entre 0 et 1, et c'est là que le bât va nous blesser.

Résultats Expérimentaux.

Les tableaux ci-après donnent les résultats de l'application des formules (VII-3) à toutes les transitions par supports décroissants (50 → 25, 50 → 15.... jusqu'à 10 → 5), plage par plage, et aussi pour la population totale TP.

De prime abord, le résultat ne semble pas miraculeux : les probabilités ≥ 1 ou < 0 abondent. A dire vrai, l'évolution un peu incohérente des paramètres en fonction de la taille du support, telle que nous l'avons notée dans le chapitre V, ne permettait guère d'espérer mieux. Les résultats sont même, en réalité, moins mauvais qu'on aurait pu le craindre. Compte tenu de l'imprécision qui affecte l'estimation de nos paramètres, nous ne devons pas, en effet, prendre au tragique des valeurs faiblement négatives (comme celles de p_P pour la population totale TP), ou faiblement supérieures à 1. On peut les remplacer par des 0 et des 1 respectivement. De ce point de vue, les transitions TP (à l'exception de 50 → 25 qui donne pour ϖ_S une valeur trop fortement supérieure à 1) fournissent un tableau assez cohérent : ϖ_S et ϖ_P diminuent avec le support V d'origine et augmentent avec le support v d'arrivée, comme on devait s'y attendre. Quant à p_P , il reste partout peu significativement différent de 0 (ce qui voudrait dire que la pascalisation se réduit à une simple décimation).

Les valeurs aberrantes de ϖ_S , observées pour les populations E et F n'ont, en fait, pas grande signification. De fait, ϖ_S reflète l'évolution du paramètre b, c'est-à-dire, pour l'essentiel, celle du paramètre a des notations de départ. Or, justement, pour ces deux populations, nous avions noté l'extrême insensibilité des ajustements au choix du paramètre ν : on pouvait, sans pratiquement modifier ni les paramètres τ et p, ni la qualité de l'ajustement, choisir ν presque arbitrairement, à condition de maintenir constant le rapport ν/a : seul ce rapport possède une signification fiable, les valeurs retenues pour ν (et donc pour a) restent très largement arbitraires. C'est cette grande marge d'arbitraire qui explique, semble-t-il, l'incohérence des résultats relatifs à ϖ_S .

On reste davantage perplexe devant certaines valeurs fortement négatives de p_P, ou fortement supérieures à 1 de ϖ_P, comme par exemple pour la transition A25 → A15.

A ces exceptions près, toutefois, on note que dans l'ensemble ϖ_S et ϖ_P apparaissent comme des fonctions croissantes du support d'arrivée v, et décroissantes du support de départ V, ce qui est assez satisfaisant. De son côté, p_P (dans les cas de figure où il présente une évolution interprétable) semble obéir à une loi analogue, mais dans bien des cas reste peu significativement différent de 0.

Au total, ces résultats sont ambigus, mais non franchement négatifs. Il faudrait sans doute, en premier lieu, améliorer la technique d'ajustement pour obtenir des estimations plus fiables des paramètres : c'est alors seulement sans doute que l'on pourra porter un jugement sur ce modèle de changement de support.

Transitions TP			
	ϖ_S	p_P	ϖ_P
50 → 5	0,789	0,028	0,235
50 → 10	0,850	0,070	0,312
50 → 15	1,069	0,086	0,341
50 → 25	1,308	0,173	0,384
25 → 5	0,603	-0,087	0,565
25 → 10	0,650	-0,082	0,782
25 → 15	0,817	-0,080	0,866
15 → 5	0,739	-0,033	0,670
15 → 10	0,796	-0,010	0,909
10 → 5	0,928	-0,026	0,739

Transitions A			
	ϖ_S	p_P	ϖ_P
50 → 5	0,105	0,342	0,502
50 → 10	0,084	0,446	0,594
50 → 15	0,168	0,371	0,856
50 → 25	0,202	0,547	0,580
25 → 5	0,519	-0,250	0,744
25 → 10	0,419	-0,262	1,058
25 → 15	0,834	-2,264	3,471
15 → 5	0,623	0,159	0,471
15 → 10	0,502	0,253	0,589
10 → 5	1,239	-0,055	0,750

Transitions B			
	ϖ_S	p_P	ϖ_P
50 → 5	0,488	-0,017	0,320
50 → 10	0,400	0,051	0,466
50 → 15	0,389	0,203	0,498
50 → 25	0,318	0,216	0,824
25 → 5	1,471	-0,110	0,333
25 → 10	1,257	-0,080	0,506
25 → 15	1,223	0,083	0,546
15 → 5	1,202	-0,169	0,589
15 → 10	1,028	-0,170	0,920
10 → 5	1,170	-0,054	0,675

Transitions D			
	ϖ_S	p_P	ϖ_P
50 → 5	0,836	0,092	0,190
50 → 10	0,985	0,162	0,253
50 → 15	0,903	0,100	0,322
50 → 25	0,874	0,182	0,443
25 → 5	0,956	0,0153	0,381
25 → 10	1,127	0,065	0,520
25 → 15	1,033	-0,037	0,685
15 → 5	0,926	0,035	0,565
15 → 10	1,091	0,091	0,766
10 → 5	0,849	-0,034	0,718

Transitions E			
	ϖ_S	p_P	ϖ_P
50 → 5	6,256	0,126	0,154
50 → 10	1,954	0,179	0,235
50 → 15	2,826	0,125	0,339
50 → 25	1,407	0,240	0,441
25 → 5	4,447	0,046	0,289
25 → 10	1,389	0,058	0,465
25 → 15	2,009	-0,073	0,716
15 → 5	2,214	0,073	0,420
15 → 10	0,691	0,101	0,665
10 → 5	3,202	0,010	0,607

Transitions F			
	ϖ_S	p_P	ϖ_P
50 → 5	7,398	0,231	0,081
50 → 10	3,426	0,363	0,117
50 → 15	3,880	0,291	0,157
50 → 25	3,636	0,032	0,377
25 → 5	2,035	0,225	0,209
25 → 10	0,942	0,356	0,304
25 → 15	1,067	0,281	0,409
15 → 5	1,906	0,095	0,430
15 → 10	0,883	0,186	0,674
10 → 5	2,159	-0,027	0,588

A N N E X E

DETAIL DES 30 AJUSTEMENTS

A LA LOI DE SICHEL-PASCAL

Toutes Plages — Support 5 m
Ajustement à Sichel Pascal

classe	Frequence OBS	Frequence CALC	χ^2
0	7608	7579	0,11
1	1376	1362	0.13
2	722	747	0,85
3	440	456	0,54
4	294	304	0.30
5	233	217	1.22
6	198	163	7,45
7	137	128	0,67
8	109	103	0,33
9	90	85	0,26
10	76	72	0,23
11	71	62	1,44
12	42	53	2,43
13	53	47	0,83
14	41	41	0.00
15	37	37	0.00
16	37	33	0.46
17	29	30	0.03
18	30	27	0.30
19	22	25	0.30
20	16	22.6	1.96
21	23	20.8	0.23
22	18	19.2	0.08
23	22	17.8	1.00
24	17	16.5	0.01
25	13	15.4	0.36
26	9	14.3	1.98
27	10	13.4	0.86
28	16	12,6	0.94
29	9	11.8	0.66
30-34	44	49.5	0.61
35-39	25	37.8	4.31
40-44	23	29.6	1.48
45-49	19	23.8	0.95
50-59	33	35.4	0.16
60-69	13	24.5	5.60
70-79	18	18.0	0.00
80-99	15	23.7	3.20
100-149	13	26.5	6.87
150-249	10	13.2	0.80
≥ 250	10	3.3	13,84
		$\chi^2 =$	63,80

$N = 12021$

m exp. $= 3,2251$
m calc. $= 3,2150$

σ^2 exp. $= 426,40$
σ^2 calc. $= 166,32$

$r = 0,986845780$
$p = 0,231202068$
$a = 4,060255791$
$D = 2,398565681$

$\theta = 0,989887$

$\chi^2 = 63,80$

$41 - 5 = 36$ deg. de lib.

$\boxed{P(\chi^2 \geq 63.8) = 0,29/100}$

T.P. – 10m – Ajustement SIPA

classe	Freq. OBS	Freq. CALC.	χ^2
0	2822	2826	0,00
1	711	714	0,01
2	458	451	0,11
3	300	304	0.06
4	223	217	0.15
5	164	163	0.01
6	124	126	0.04
7	107	101	0.33
8	94	83	1.43
9	70	70	0.00
10	55	59	0.31
11	46	51	0.53
12	57	45	3.35
13	33	39	1.07
14	43	35	1.75
15	40	32	2.29
16	30	28	0.09
17	15	26	4.52
18	27	24	0.51
19	25	22	0.55
20	19	19.8	0.03
21	16	18.3	0.29
22	23	17.0	2.15
23	18	15.8	0.32
24	14	14.7	0.03
25	13	13.7	0.04
26	11	12.8	0.27
27	10	12.1	0.35
28	8	11.3	0.98
29	13	10.7	0.50
30-34	47	45.3	0.06
35-39	32	35.2	0.29
40-44	32	28.1	0.55
45-49	22	22.9	0.03
50-59	35	34.9	0.00
60-69	18	25.2	2.06
70-79	17	18.9	0.19
80-89	11	14.6	0.87
90-99	9	11.4	0.53
100-149	21	32.0	3.80
150-249	10	19.6	4.70
≥ 250	14	7.5	5.67
			$\chi^2 = 40,84$

N = 5857

m exp. = 6,6198
m calc. = 6,4323
σ^2 exp. = 1310,15
σ^2 calc. = 474,77

$r = 0,990113069$
$p = 0,281144375$
$a = 2,819057182$
$\nu = 2,710454249$

$\theta = 0,99289$

$\chi^2 = 40,84$

42 − 5 = 37 Deg. de Lib.

$$\boxed{P(\chi^2 \geq 40,84) = 30,5\ \%}$$

T.P. - 15m. Ajustement SIPA

classe	Fréq. OBS.	Fréq. CALC.	χ^2
0	1458	1465	0.03
1	451	450	0.00
2	321	310	0.43
3	228	223	0.12
4	148	167	2.20
5	121	130	0.59
6	105	104	0.02
7	96	85	1.48
8	78	71	0.74
9	60	60	0.00
10	49	52	0.14
11	38	45	1.09
12	31	40	1.87
13	32	35	0.28
14	41	31	2.90
15	36	28	2.09
16	34	26	2.72
17	24	23	0.02
18	16	21	1.35
19	24	20	0.97
20	16	18	0.25
21	23	17	2.32
22	20	16	1.26
23	11	14	0.84
24	12	14	0.17
25	7	13	2.54
26	15	12	0.82
27	9	11	0.42
28	10	11	0.03
29	7	10	0.87
30-34	39	42	0.26
35-39	37	33	0.47
40-44	26	27	0.01
45-49	29	22	2.43
50-59	42	33	2.21
60-69	17	24	2.22
70-79	22	18	0.70
80-89	17	14	0.51
90-99	7	11	1.66
100-149	22	32	3.32
150-249	17	21	0.70
≥ 250	11	9	0.49

$\chi^2 = 43,52$

$N = 3807$

m exp. = 10,1843
m calc. = 9,3594
σ^2 exp. = 3123,58
σ^2 calc. = 772,97

$\nu = 0,990748632$
$p = 0,308050573$
$a = 2,150279076$
$D = 2,824809525$

$\theta = 0,9935985$

42 - 5 = 37 Deg de Lib

$\chi^2 = 43,52$

$\boxed{P(\chi^2 \geq 43,52) = 21,4\%}$

T.P. - 25 m - Ajustement 31,1,1

classe	Freq. OBS.	Freq. CALC.	χ^2
0	563	561	0.00
1	204	205	0.01
2	159	160	0.01
3	147	128	2.78
4	90	104	1.95
5	81	86	0.33
6	74	73	0.03
7	61	62	0.01
8	68	53	4.04
9	42	46	0.43
10	36	41	0.59
11	30	36	1.09
12	33	32	0.01
13	28	29	0.05
14	26	26	0.01
15	23	24	0.05
16	24	22	0.18
17	22	20	0.16
18	17	19	0.14
19	19	17	0.18
20	13	16	0.56
21	11	15	1.01
22	22	14	4.72
23	10	13	0.70
24	13	12	0.05
25	9	11	0.53
26	12	11	0.13
27	10	10	0.00
28	6	9	1.38
29	7	9	0.50
30-34	47	39	1.52
35-39	34	31	0.27
40-44	27	25	0.12
45-49	13	21	3.01
50-59	39	33	1.23
60-69	20	24	0.74
70-79	25	19	2.19
80-89	15	15	0.01
90-99	12	12	0.00
100-149	32	35	0.26
150-249	22	25	0.29
≥ 250	14	13	0.05
		$\chi^2 =$	31,33

$N = 2160$

$m_{exp} = 17{,}9501$

$m_{calc.} = 16{,}340$

$\sigma^2_{exp.} = 4894.11$

$\sigma^2_{calc.} = 1696{,}20$

$r = 0{,}99153\,5061$

$p = 0{,}392800136$

$a = 1{,}662941991$

$\nu = 3{,}231347667$

$\theta = 0{,}99486$

$42 - 5 = 37$ Deg. de Lib.

$\chi^2 = 31{,}33$

$$\boxed{P(\chi^2 \geq 31{,}33) = 73{,}2\,\%}$$

T-P. - 50 m. Ajustement SIPA

classe	Freq. OBS.	Freq. CALC.	χ^2
0	109	109	0.00
1	53	52	0.03
2	42	47	0.46
3	48	42	0.90
4	42	38	0.51
5	32	34	0.10
6	32	31	0.01
7	25	28	0.25
8	27	25	0.14
9	24	23	0.05
10	29	21	3.11
11	24	19	1.20
12	11	18	2.53
13	13	16	0.68
14	15	15	0.00
15	12	14	0.30
16	14	13	0.06
17	9	12	0.86
18-19	15	22	2.36
20-21	16	20	0.69
22-23	12	18	1.75
24-25	20	16	1.14
26-27	16	14	0.22
28-29	12	13	0.07
30-31	15	12	0.86
32-34	15	16	0.05
35-39	16	23	1.96
40-44	19	19	0.00
45-49	22	16	2,20
50-59	34	26	2.57
60-69	17	20	0.44
70-79	18	16	0.28
80-89	11	13	0.30
90-99	16	11	2.51
100-149	39	35	0.53
150-199	12	19	2.34
200-249	9	11	0.48
≥ 250	27	29	0.11

$\chi^2 = 32.06$

$N = 922$

m exp. = 42,0522
m calc. = 40,745

σ^2 exp. = 21216,55
σ^2 calc. = 8912,77

r = 0,99534 3122
p = 0,48508 7786
a = 1,63503 4622
D = 4,90

θ = 0,997602

38 - 5 = 33 Deg. de Lib.

$\chi^2 = 32,06$

$\boxed{P(\chi^2 \geq 32,06) = 51,4\,\%}$

A-5m. Adjustement SIPA

classe	Freq. OBS	Freq. CALC	χ^2
0	1765	1758	0.03
1	165	159	0.24
2	80	95	2.34
3	80	61	5.88
4	30	42	3.41
5	28	30	0.21
6	19	23	0.76
7	22	18	0.74
8	19	15	1.15
9-11	39	32	1.62
12-14	17	21	0.71
15-17	15	15	0.00
18-20	10	11	0.14
21-23	10	9	0.14
24-29	13	13	0.00
30-35	10	9	0.05
36-49	12	14	0.25
50-149	16	25	3.52
≥ 150	7	6	0.10
			21,31

$N = 2357$

m exp : 2,9499
m calc. : 3,0941
σ^2 exp : 311,86
σ^2 calc. : 285,22

τ : 0,99 138 4798
f : 0,33 559 3742
a : 2,15 743 4519
δ : 0,86 619 8221

θ : 0,99 4276

19-5 = 14 Deg. d. Lib.

$\chi^2 = 21,31$

$\boxed{P(\chi^2 \geq 21,31) = 9,4\%}$

A-10m. Ajustement SIPA

classe	Freq. OBS	Freq. CALC	χ^2
0	729	723	0.06
1	88	87	0.01
2	49	59	1.73
3	45	42	0.24
4	36	31	0.93
5	23	23	0.00
6	14	18	0.97
7	20	15	1.97
8	17	12	2.06
9-11	18	26	2.49
12-14	14	17	0.62
15-17	11	12	0.17
18-20	13	9	1.33
21-23	10	7	0.85
24-29	13	11	0.31
30-39	12	12	0.00
40-59	12	13	0.16
60-149	13	18	1.48
≥ 150	8	8	0.01
			15,41

$N = 1145$

$m\,exp. = 6{,}0725$

$m\,calc. = 6{,}1126$

$\sigma^2 exp. = 954{,}01$

$\sigma^2 calc. = 785{,}28$

$\tau = 0{,}99\ 273\ 2581$

$p = 0{,}43\ 944\ 8548$

$a = 2{,}47\ 502\ 9806$

$b = 1{,}50\ 666\ 8127$

$\theta = 0{,}995\ 926$

19-5=14 Deg. d Lib.

$\chi^2 = 15{,}41$

$$\boxed{P(\chi^2 \geq 15{,}41) = 35{,}0\%}$$

A-15m - Ajustement SIPA

classe	Frq. OBS	Frq. CALC	χ^2
0	411	408	0,02
1	67	63	0,29
2	37	42	0,70
3	29	30	0,05
4	22	23	0.02
5	14	18	0.73
6-7	27	26	0.07
8-9	23	18	1.38
10-11	11	13	0.45
12-13	9	11	0.22
14-15	11	9	0.72
16-19	14	13	0.07
19-24	14	12	0.51
25-29	12	8	1.47
30-39	12	12	0.01
40-54	11	11	0.01
55-149	14	22	2.64
$\geq$ 150	9	10	0.03
			9,40

$N = 747$

m exp. = 9,3079
m calc. = 9,8998
σ^2 exp = 1623,8
σ^2 calc. = 1653,0

$\tau = 0,99\ 496\ 5083$
$p = 0,36\ 090\ 2636$
$a = 1$
$\nu = 0,96\ 466\ 3540$

$\theta = 0,996\ 782$

18 - 4 = 14 Deg. de Lib.
$\chi^2 = 9,40$

$\boxed{P(\chi^2 \geq 9,40) : 80,5\%}$

A-25 m. Ajustement SIPA

classe	Freq OBS	Freq CALC	χ^2
0	180	177	0.05
1	30	29	0.06
2	23	23	0.00
3	18	18	0.01
4	13	15	0.29
5	11	13	0.21
6	9	11	0.27
7-8	17	17	0.00
9-10	15	13	0.27
11-12	12	10	0.24
13-15	12	12	0.00
16-20	13	15	0.17
21-25	15	10	2.13
26-31	10	9	0.10
32-38	9	8	0.20
39-54	8	11.5	1.09
55-149	19	21	0.28
≥ 150	8	9	0.18
			5.56

$N = 422$

m exp. = 16,476
m calc. = 16,442
σ^2 exp. = 3950,7
σ^2 calc. = 2636,4

$\tau = 0,99\ 254\ 7023$
$p = 0.54\ 195\ 2482$
$a = 1$
$\nu = 1,42\ 369\ 15\ 12$

$\theta = 0,996\ 586$

18-4 = 14 Deg de Lib.

$\chi^2 = 5.56$

$\boxed{P(\chi^2 \geq 5,56) = 97,6\ \%}$

A. Som. Ajustement SIPA

classe	Freq. OBS	Freq. CALC	χ^2
0	37	34	0.22
1	15	14	0.11
2-4	20	22	0.18
5-6	9	9	0.01
7-9	14	10	1.24
10-12	10	8	0.50
13-18	10	12	0.31
19-25	9	10	0.10
26-30	9	6	2.11
31-39	11	8	1.19
40-49	9	7	0.79
50-99	9	18	4.71
100-149	8	8	0.00
≥ 150	8	12	1.26
			12,73

$N = 178$

$m_{exp} = 39,062$

$m_{calc} = 39,057$

$\sigma^2_{exp} = 11\,303$

$\sigma^2_{calc.} = 6221,8$

$\tau = 0,99\,570\,0315$

$p = -0,01941\,9863$

$a = 0,10\,083\,0887$

$D = 0,87\,262\,9409$

$\theta = 0,995\,617$

$14 - 5 = 9$ Deg. de Lib.

$\chi^2 = 12,73$

$$\boxed{P(\chi^2 \geq 12,73) = 17,5\,\%}$$

B-5m · Ajustement SIP/1

classe	Freq. OBS	Freq. CALC.	χ^2
0	2056	2048	0.03
1	312	309	0.04
2	138	149	0.85
3	79	90	1.31
4	72	61	1.92
5	52	45	1.12
6	45	35	3.08
7	30	28	0.18
8	22	23	0.03
9	13	19	1.98
10	18	16	0.17
11	14	14	0.00
12	7	12	2.34
13	10	11	0.08
14	14	10	1.84
15	11	9	0.57
16-17	16	15	0.05
18-20	9	18	4.58
21-24	17	18	0.10
25-29	19	17	0.25
30-39	18	22	0.89
40-49	14	14	0.00
50-69	14	17	0.40
70-89	10	9	0.09
90-149	9	12	0.57
≥ 150	8	7	0.29
			22.76

$N = 3028$

m exp. = 3,4333
m calc. = 3,4069
σ^2 exp. = 315,89
σ^2 calc. = 236,70

τ = 0,99165 1615
r = 0,05834 6844
a = 1
ν = 0,64520 6926

θ = 0,9921387

26 - 4 = 22 Deg. d. Lib.

$\chi^2 = 22,76$

$\boxed{P(\chi^2 \geq 22,76) = 41,5\ \%}$

B-10 m. Ajust[illegible]

classe	Frq. OBS	Frq. CALC.	χ^2
0	801	797	0.02
1	168	167	0.00
2	101	96	0.31
3	52	62	1.65
4	42	44	0.10
5	30	33	0.31
6	38	26	5,42
7	26	21	1.10
8	21	18	0.65
9	10	15	1.63
10	13	13	0.00
11	11	11	0.00
12	13	10	0.99
13-15	14	24	3.99
16-18	15	18	0.41
19-21	13	14	0.03
22-24	10	11	0.11
25-29	13	14	0.14
30-34	12	11	0.10
35-39	9	9	0.02
40-49	16	13	0.84
50-69	16	16	0.01
70-99	11	12	0.14
100-149	8	9.5	0.25
≥ 150	11	10	0.04
			18.26

$N = 1474$

m exp. = 7,0529
m calc. = 6,9956

σ^2 exp = 830,13
σ^2 calc. = 747,52

r = 0,99405 7641
p = 0,15 045 0756
a = 1
ϑ = 0,99 281 0201

θ = 0,99495

25 - 4 = 21 Deg. de Li. l.

χ^2 = 18,26

$\boxed{P(\chi^2 \geq 18{,}26) = 63{,}2\ \%}$

B-15m. Ajustement SIPA

classe	Freq. OBS	Freq. CALC.	χ^2
0	451	447	0.03
1	98	97	0.00
2	70	65	0.42
3	36	46	2.14
4	38	34	0.40
5	24	27	0.29
6	28	21	1.99
7	19	18	0.10
8	12	15	0.56
9	10	13	0.59
10	14	11	0.78
11-12	16	18	0.29
13-14	11	15	0.89
15-16	15	12	0.77
17-19	12	15	0.42
20-22	13	11	0.20
23-29	14	20	1.59
30-34	11	10	0.09
35-39	8	8	0.00
40-49	14	12	0.36
50-59	13	8	2.39
60-79	12	11	0.04
80-109	9	10	0.07
110-179	7	10	1.08
≥ 180	9	9	0.00
			15,48

$N = 964$

m exp. = 10,784
m calc. = 10,773
σ^2 exp. = 1744,7
$\hat{\sigma}^2$ calc. = 1462,2

$\tau = 0,99\,439\,8618$
$p = 0,29\,097\,3353$
$a = 1$
$\nu = 1,23\,585\,95\,44$

$\theta = 0,996\,028$

25-4 = 21 Deg. de Lib.

$\chi^2 = 15,48$

$$\boxed{P(\chi^2 \geq 15,48) = 79,8\,\%}$$

B.25m. Ajustement SIPA

classe	Frq. OBS	Frq. CALC	χ^2
0	178	176	0.02
1	51	50	0.01
2	38	37	0.02
3	28	28	0.00
4	18	22	0.87
5	17	18	0.08
6	18	15	0.55
7	18	13	2.13
8	17	11	3.29
9-10	12	18	1.98
11-12	12	14	0.32
13-14	10	12	0.20
15-16	10	10	0.02
17-19	11	12	0.05
20-24	12	15	0.55
25-29	12	11	0.08
30-39	17	16	0.14
40-59	16	18	0.27
60-79	15	10	1.93
80-99	7	7	0.00
100-199	11	15	1.20
≥ 200	12	11	0.03
			13.74

$N = 540$

m exp. = 19,252
m calc. = 21,147
σ^2 exp. = 3384,4
σ^2 calc. = 4744,4

$\tau = 0{,}99\ 635\ 8313$
$p = 0{,}34\ 953\ 3639$
$a = 1$
$\nu = 1{,}76\ 680\ 1799$

$\theta = 0{,}99\ 763$

22-4 = 18 Deg. d Lib
$\chi^2 = 13{,}74$

$\boxed{P(\chi^2 \geq 13{,}74) = 74{,}6\ \%}$

B-50m. Ajustement SIPA

$N = 227$

classe	Frq. OBS	Frq. CALC	χ^2
0	35	35	0.00
1	15	15	0.00
2	10	11	0.19
3	12	9	0.76
4	8	7.9	0.00
5	8	6.8	0.20
6-8	18	16.2	0.19
9-11	22	12.2	7.88
12-19	15	22.3	2.40
20-29	13	17.6	1.19
30-39	9	11.9	0.72
40-49	8	8.8	0.07
50-59	10	6.8	1.53
60-89	11	13.6	0.48
90-119	13	8.2	2.84
120-199	13	11.5	0.20
≥200	13	13.0	0.00
			18.66

m exp. = 45,797
m calc. = 48,343
σ^2 exp. = 9947,9
σ^2 calc. = 11 482,6

τ = 0,99 688 8947
p = 0,19 979 2764
a = 0, 25 612 0969
γ = 1,35 015 7633

θ = 0,99751

17 - 5 = 12 Deg. d. Lib.

$\chi^2 = 18,66$

$\boxed{P(\chi^2 \geq 18,66) = 9,7\ \%}$

D. Sm. Ajustement SIPA

classe	Freq. OBS	Freq CALC	χ^2
0	1896	1894	0.00
1	356	359	0.03
2	192	167	3.61
3	82	92	1.03
4	48	57	1.45
5	46	39	1.27
6	32	28	0.47
7	13	22	3.43
8	18	17	0.06
9	11	14	0.53
10	14	11	0.62
11-12	11	18	2.46
13-14	13	13	0.00
15-16	12	10	0.47
17-19	11	11	0.00
20-23	11	10	0.09
24-27	8	7	0.18
28-49	8	15	3.56
≥ 50	8	6	0.82
			20.10

$N = 2790$

m exp. = 1,5118
m calc. = 1,4920

σ^2 exp. = 51,360
σ^2 calc. = 26,834

$\tau = 0,96515\,8488$
$p = 0,15848\,3510$
$a = 20,420\,54747$
$\nu = 10,007\,64991$

$\theta = 0,97068$

19-5 = 14 Deg. d. Lib.

$\chi^2 = 20,10$

$\boxed{P(\chi^2 \geq 20,10) = 12,7\,\%}$

D-10 · Ajustement SIPA

classe	Frq. OBS	Frq. CALC	χ^2
0	731	727	0,02
1	199	198	0.00
2	112	114	0.02
3	76	71	0.40
4	45	47	0.11
5	38	34	0.57
6	18	25	2.01
7	23	19	0.64
8	12	16	0.81
9	9	13	1.09
10	12	11	0.18
11	10	9	0.12
12	9	8	0.22
13	8	7	0.26
14	10	6	2.97
15-16	9	10	0.05
17-20	9	14	1.81
21-24	9	10	0.04
25-30	10	10	0.02
31-49	10	14	1.18
≥ 50	11	10	0.17
			12,70

$N = 1370$

m exp. = 3,0788
m calc. = 3,0350
σ^2 exp. = 128,54
σ^2 calc. = 77,68

$r = 0,97\,293\,2320$
$p = 0,24\,178\,3322$
$a = 15,252\,222\,141$
$\nu = 11,998\,263$

$\theta = 9,97948$

21-5 = 16 Deg de Lib

$\chi^2 = 12,70$

$\boxed{P(\chi^2 \geq 12,70) = 69,5\,\%}$

D-15m. Ajustement SIPA

dosse	Freq. OBS	Freq. CALC.	χ^2
0	370	349	1.21
1	141	146	0.18
2	101	91	1.13
3	66	60	0.64
4	25	42	6.66
5	30	30	0.01
6	22	23	0.06
7	20	18	0.16
8	8	15	3.11
9	9	12	0.84
10	11	10	0.03
11-12	14	16	0.33
13-14	18	12	2.56
15-17	12	14	0.24
18-21	11	13	0.28
22-29	12	16	0.89
30-39	8	11	0.67
40-59	10	10	0.00
≥ 60	9	8	0.05
			19,09

$N = 897$

$m_{exp.} = 4,7023$

$m_{calc.} = 4,4569$

$\sigma^2_{exp.} = 257,38$

$\sigma^2_{calc.} = 133,27$

$\tau = 0,97809\,1575$

$p = 0,20585\,2035$

$a = 14,984\,82\,641$

$\nu = 17,211$

$\theta = 0,98260$

$19 - 5 = 14$ Deg d Li B.

$\chi^2 = 19.09$

$\boxed{P(\chi^2 \geq 19.09) = 16,2\,\%}$

D-25- Ajustement SIPA

classe	Freq. OBS	Freq. CALC	χ^2
0	151	150	0.00
1	70	70	0.00
2	47	51	0.31
3	42	38	0.49
4	25	29	0.44
5	25	28	0.37
6	21	18	0.66
7	17	14	0.52
8	12	12	0.00
9-10	16	18	0.31
11-12	12	14	0.21
13-14	13	11	0.54
15-16	8	8	0.03
17-19	10	10	0.00
20-25	8	14	2.34
26-32	9	10	0.12
33-49	10	13	0.74
50-69	9	7	0.53
≥ 70	10	9	0.08
			7,69

$N = 518$

m exp $= 8,1429$

m calc. $= 8,0642$

σ^2 exp. $= 552,50$

$\hat{\sigma}^2$ calc. $= 361,82$

$r = 0,98\,328\,93\,85$

$p = 0,30\,887\,47\,01$

$a = 13,518\,627\,10$

$\nu = 20$

$\theta = 0,98845$

$19 - 5 = 14$ Deg d Lib.

$\chi^2 = 7,69$

$$\boxed{P(\chi^2 \geq 7,69) = 90,5\ \%}$$

[illegible]

$N = 235$

classe	Freq. OBS	Freq. CALC	χ^2
0	33	33	0.00
1	19	24	0.94
2	24	20	0.73
3	19	17	0.28
4	14	14	0.00
5	10	12	0.23
6	12	10	0.46
7	9	8	0.04
8	9	7	0.45
9-11	13	17	0.76
12-13	9	8	0.09
14-15	7	7	0.07
16-19	9	10	0.12
20-24	8	9	0.10
25-29	9	6	1.00
30-49	11	14	0.79
50-89	11	11	0.01
≥ 90	9	9	0.01
			6,06

m exp. = 17,949
m calc. = 17,025

σ^2 exp = 2194,9
σ^2 calc. = 1383,52

τ = 0,99 108 3670
p = 0,29 276 3165
a = 8,62 236 1429
ν = 20

θ = 0,99369

18-5 = 13 Deg. de Lib.

χ^2 = 6,06

$$\boxed{P(\chi^2 \geq 6,06) = 94,4\%}$$

E-5m. Ajustement SIPA

classe	Freq. OBS	Freq. CALC.	χ^2
0	969	969	0.00
1	283	286	0.03
2	179	162	1.84
3	98	103	0.26
4	62	71	1.26
5	49	52	0.24
6	46	40	0.82
7	30	32	0.11
8	20	26	1.30
9	19	21	0.25
10	20	18	0.25
11	20	15	1.51
12	12	13	0.09
13	15	11	1.18
14	8	10	0.37
15	9	9	0.01
16	7	8	0.07
17-18	11	13	0.32
19-20	14	11	1.14
21-22	12	9	1.31
23-26	9	13	1.32
27-31	10	11	0.16
32-39	12	11	0.04
40-54	11	10	0.04
≥ 55	8	9	0.07
			14,00

$N = 1934$

m exp. $= 3.2223$
m calc. $= 3.2010$

σ^2 exp. $= 69,654$
σ^2 calc. $= 65,616$

$\tau = 0,96\ 838\ 2002$
$p = 0,11\ 735\ 6426$
$a = 1,46\ 062\ 8771$
$\nu = 1,7$

$\theta = 0,97209$

$25 - 5 = 20$ Deg. de Lib.

$\chi^2 = 14,00$

$$\boxed{P(\chi^2 \geq 14) = 83,0\%}$$

E-10m. Ajustement SIPA

classe	Frq. OBS	Frq. CALC	χ^2
0	281	283	0.02
1	140	149	0.49
2	115	100	2.28
3	66	70	0.24
4	50	51	0.03
5	40	39	0.03
6	24	30	1.38
7	19	24	1.25
8	21	20	0.04
9	19	17	0.27
10	10	14	1.30
11	12	12	0.01
12	11	11	0.01
13	9	9	0.02
14	10	8	0.35
15	11	7	1.78
16-17	11	13	0.19
18-19	10	10	0.01
20-22	12	12	0.01
23-25	13	10	1.23
26-29	9	10	0.07
30-34	12	9	0.92
35-49	11	16	1.45
50-69	9	9	0.01
≥70	9	9	0.00
			13,37

$N = 944$

m exp. $= 6,6017$
m calc. $= 6,2929$

σ^2 exp. $= 188,58$
σ^2 calc. $= 192,90$

$\tau = 0,97\ 923\ 7905$
$p = 0,16\ 653\ 4481$
$a = 4,10\ 674\ 8353$
$D = 6,56\ 134\ 7340$

$\theta = 0,98269$

25-5 = 20 Deg. d. Lib.

$\chi^2 = 13,37$

$\boxed{P(\chi^2 \geq 13,37) = 86,1\%}$

E-15 m. Ajustement SIPA

classe	Frq. OBS	Frq. CALC	χ^2
0	115	114	0.01
1	82	83	0.00
2	62	62	0.00
3	54	48	0.87
4	35	37	0.12
5	27	30	0.23
6	25	24	0.03
7	21	20	0.04
8	23	17	2.16
9	16	14	0.16
10	8	13	1.65
11-12	14	21	2.15
13-14	14	16	0.32
15-16	17	13	1.10
17-18	9	11	0.33
19-20	9	9	0.00
21-24	12	14	0.42
25-29	11	13	0.37
30-34	9	10	0.05
35-39	9	7	0.36
40-49	14	10	1.31
50-59	8	7	0.24
60-79	10	8	0.57
≥ 80	8	11	0,62
			13,12

$N = 612$

m exp. = 10,183
m calc. = 10,256
σ^2 exp. = 391,82
σ^2 calc. = 424,13

$r = 0,98\ 558\ 2257$
$p = 0,10\ 623\ 1838$
$a = 2,32\ 906\ 8395$
$D = 5,47\ 019\ 6660$

$\theta = 0,987114$

24 - 5 = 19 Deg. de Lib.

$\chi^2 = 13,12$

$\boxed{P(\chi^2 \geq 13,12) = 83,2\%}$

E-25m - Ajustement SIPA

classe	Frq. OBS	Frq. CALC	χ^2
0	31	29	0.10
1	23	28	1.05
2	27	27	0.01
3	35	24	5.34
4	15	21	1.60
5	15	18	0.51
6	18	16	0.34
7	12	14	0.21
8	17	12	2.08
9	8	11	0.64
10-11	13	18	1.30
12-13	16	14	0.18
14-15	11	12	0.06
16-17	7	10	0.86
18-19	9	8	0.04
20-21	9	7	0.43
22-25	10	12	0.28
26-31	13	13	0.00
32-37	10	10	0.02
38-44	9	8	0.07
45-54	12	8	1.54
55-69	8	8	0.01
70-99	9	9	0.00
≥ 100	10	10	0.01
			16,67

$N = 347$

m exp = 17,960
m calc. = 18,317

σ^2 exp. = 799,03
σ^2 calc. = 1103,79

r = 0,98 895 4125
p = 0,21 876 8736
a = 4,20 121 4963
v = 13,09 769 9182

θ = 0,9913706

24-5 = 19 Deg de Lib.

$\chi^2 = 16,67$

$$\boxed{P(\chi^2 \geq 16,67) = 61,2\,\%}$$

E-50m. Ajustement SIPA

classe	Frq OBS	Frq CALC	χ^2
0-3	12	10	0,30
4-5	8	10	0,26
6-7	9	10	0,05
8-9	10	9	0,10
10-11	12	8	1,84
12-16	10	17	2,63
17-19	7	8	0,10
20-22	7	7	0,01
23-25	7	6	0,27
26-30	7	8	0,12
31-35	8	6	0,39
36-49	10	13	0,59
50-64	12	9	1,27
65-99	10	11	0,13
100-149	9	7	0,43
≥ 150	8	8	0,00
			8,50

$N = 146$

m exp. = 42,685
m calc. = 42,719
σ^2 exp. = 3842,4
σ^2 calc. = 4328,2

$r = 0,99\ 527\ 4349$
$p = -0,06\ 662\ 5486$
$a = 24,9\ 432\ 8999$
$a = 3,89\ 400\ 8260$

$\theta = 0,994\ 959$

16-5 = 11 Deg. de Lib.

$\chi^2 = 8,50$

$\boxed{P(\chi^2 \geq 8,5) = 66,8\%}$

F-5 m. Ajustement SIPA

classe	Frq. OBS	Frq. CALC.	χ^2
0	697	697	0,00
1	179	173	0,19
2	97	108	1,21
3	75	74	0,02
4	57	53	0,23
5	38	41	0,17
6	36	32	0,52
7	31	26	1,06
8	20	21	0,07
9	23	18	1,49
10	8	15	3,40
11	12	13	0,09
12	16	11	1,89
13	11	10	0,11
14	8	9	0,07
15	6	8	0,43
16	9	7	0,57
17-18	10	12	0,32
19-20	7	10	0,81
21-22	11	8	0,97
23-29	12	20	3,23
30-36	10	12	0,36
37-44	10	9	0,20
45-59	12	9	1,10
≥ 60	11	10	0,16
			18,68

$N = 1406$

m exp. = 4,1131
m calc. = 4,1130

σ^2 exp. = 106,73
σ^2 calc. = 109,24

$\tau = 0{,}97\ 163\ 2850$
$p = 0{,}21\ 641\ 2795$
$a = 1$
$\nu = 1{,}30554$

$\theta = 0{,}977\ 772$

25 - 4 = 21 Deg. de Lib.

$\chi^2 = 18{,}68$

$\boxed{P(\chi^2 \geq 18{,}68) = 60{,}6\%}$

F-10m Ajustement SIPA

classe	Freq. OBS	Freq. CALC.	χ^2
0	213	214	0.00
1	77	78	0.01
2	58	58	0.00
3	45	44	0,02
4	41	34	1,22
5	22	28	1,15
6	21	23	0,11
7	14	19	1,21
8	16	16	0,00
9	17	14	0,88
10	11	12	0,04
11	7	10	1,02
12	12	9	0.99
13	6	8	0.50
14	10	7	1.13
15	8	6	0.38
16-17	14	11	0.77
18-20	11	13	0.39
21-24	12	13	0.14
25-28	8	10	0,41
29-33	8	9	0,22
34-39	8	8	0,01
40-49	11	9	0,25
50-59	7	6	0,12
60-79	10	7	1,09
≥ 80	9	10	0,04
			12,12

$N = 677$

m exp = 8,5421
m calc = 8,5411

σ^2 exp. = 306,71
σ^2 calc. = 365,00

τ : 0,98 037 2705
p : 0,34 526 4313
a = 1,95 863 4845
D : 3,35 472 3372

θ : 0,987 146

26-5 = 21 Deg. de Lib.

χ^2 = 12,12

$\boxed{P(\chi^2 \geq 12,12) = 93,6\,\%}$

F-15m. Ajustement SIPA

classe	Frq. OBS	Frq. CALC	χ^2
0	81	81	0.00
1	44	45	0,02
2	36	36	0,00
3	35	29	1,24
4	22	24	0,14
5	18	20	0,17
6	17	17	0,00
7	9	14	1,99
8	14	12	0,21
9	13	11	0,43
10	9	10	0,03
11	7	8	0,26
12	8	8	0,02
13-14	11	13	0,30
15-16	15	11	1,73
17-20	14	17	0,41
21-24	13	12	0,04
25-29	8	11	1,05
30-33	7	7	0,00
34-39	8	8	0,00
40-49	9	10	0,03
50-69	12	11	0,11
70-89	9	6	1,86
≥ 90	7	10	0,68
			10,72

$N = 429$

m exp. = 13,480
m calc. = 13,254

σ^2 exp. = 589,66
σ^2 calc. = 678,81

τ = 0,98 543 8511
p = 0,26 464 7216
a = 1,47 529 9736
ν = 3,80 962 37 67

$\theta = 0,989292$

24 − 5 = 19 Deg. de Lib.

$\chi^2 = 10,72$

$$\boxed{P(\chi^2 \geq 10,72) = 93,3\ \%}$$

F-25 - Ajustement SIPA

classe	Freq. OBS	Freq CALC	χ^2
0	15	15	0,00
1	20	18	0,30
2	17	16	0,02
3	15	14	0,02
4	16	13	0,89
5	10	11	0,10
6	8	10	0,32
7	5	9	1,57
8	9	8	0,21
9	7	7	0,00
10-11	10	12	0,32
12-14	9	14	1,97
15-16	9	8	0,19
17-18	12	7	4,16
19-21	9	9	0,03
22-25	10	9	0,07
26-32	8	12	1,36
33-39	12	9	1,22
40-64	13	18	1,36
65-89	10	9	0,17
90-114	9	5	3,14
≥ 115	9	11	0,25
			17,64

$N = 242$

m exp. = 23,897
m calc. = 25,698
σ^2 exp. = 1392
σ^2 calc. = 2247

τ = 0,99 418 6798
p = −0,05 791 8125
a = 1
D = 4,4877

θ = 0,993 850

22 − 4 = 18 Deg. d Lib.

χ^2 = 17,64

$\boxed{P(\chi^2 \geq 17{,}64) = 48{,}0\,\%}$

F.50m. Ajustement SIPA

classe	Frq. OBS	Frq. CALC	χ^2
0-3	6	7	0,22
4-5	8	6	0,40
6-8	8	9	0,11
9-11	8	8	0,02
12-15	7	8	0,16
16-20	8	8	0,00
21-27	7	8	0,14
28-44	8	12	1,39
45-69	12	9	0,69
70-119	8	9	0,09
120-199	9	6	1,74
$\geq$ 200	8	6	0,35
			5,31

$N = 97$

m exp. = 59,619
m calc. = 59,619

σ^2 exp. = 6361,8
σ^2 calc. = 11703,6

τ = 0,99 813 7297
p = −0,29 211 15 84
a = 2,17 188 1055
ν = 14,7 563 1235

θ = 0,997 396

12 − 5 = 7 Deg. de Lib.

χ^2 = 5,31

$\boxed{P(\chi^2 \geq 5{,}31) = 62{,}2\,\%}$