REVUE DE STATISTIQUE APPLIQUÉE

G. HANSEL
D. GROUCHKO

Prévision séquentielle par la méthode de Bayes

Revue de statistique appliquée, tome 13, n^{o} 3 (1965), p. 67-81
<http://www.numdam.org/item?id=RSA_1965__13_3_67_0>

PRÉVISION SÉQUENTIELLE PAR LA MÉTHODE DE BAYES

G. HANSEL et D. GROUCHKO

INTRODUCTION

Une variable aléatoire X est caractérisée par une loi de probabilité f(X, u) dont la nature seule est connue, et non la valeur du paramètre u.

Des réalisations x_1, x_2, x_3, ..., x_k de la variable X étant observées, deux démarches peuvent conduire à une prévision de ses réalisations futures :

Première solution (Méthode classique).

Nous déterminons une estimation ponctuelle $\hat{u}$ du paramètre (par exemple, par la méthode du maximum de vraisemblance) et considérons que la loi $f(x, \hat{u})$, ainsi obtenue, est la loi exacte de la variable X, qui définit les probabilités des réalisations futures de X. La prévision, ainsi effectuée, néglige l'incertitude subsistant sur la valeur exacte de u ; par exemple, que le nombre d'observations soit grand ou petit, à une même estimation $\hat{u}$ correspond la même loi de la prévision $f(x, \hat{u})$; pourtant, une plus grande incertitude sur le paramètre u devrait conduire à une loi de la prévision plus dispersée.

Deuxième solution.

Une méthode fondée sur le théorème de Bayes sera décrite dans cet article, elle permet d'intégrer dans la prévision, l'incertitude sur l'estimation des paramètres.

Une propriété caractéristique de cette méthode est de conduire à une loi $z(x/x_1, x_2, \dots x_k)$ de la prévision dont la nature est, en général, différente de celle de la loi f (x, u).

Cependant, si le nombre K des observations augmente indéfiniment, l'incertitude sur u devient négligeable et $z(x/x_1, x_2, \dots x_k)$ tend vers f(x, u).

Le tableau ci-dessous, donne la nature des lois de prévision correspondant aux distributions les plus classiques :

Loi du phénomène $f(x, u)$	Loi de la prévision $Z(x/x_1, x_2 \dots x_k)$
1. Loi normale $\begin{cases}\text{moyenne inconnue} \\ \text{écart-type connu}\end{cases}$	Loi normale
2. Loi normale $\begin{cases}\text{écart-type inconnu,} \\ \text{moyenne connue ou} \\ \text{inconnue}\end{cases}$	Loi de Student
3. Loi de Poisson de moyenne inconnue	Loi binomiale négative (aussi appelée loi de Pascal)
4. Loi binomiale $\begin{cases}p \text{ inconnu} \\ n \text{ connu}\end{cases}$	Loi hyper-géométrique négative (aussi appelée loi Beta-binomiale)
5. Loi Gamma (paramètre d'échelle inconnu, paramètre de forme connu)	Loi Beta-inverse
6. Loi Binomiale négative p inconnu n connu	Loi Beta-Pascal
7. Loi uniforme $(0, u)$ u inconnu	Loi Rectangulaire-Hyperbolique

Nous ne donnerons pas dans cet article, une application de la méthode de prévision mise au point ; un exemple pratique a déjà été publié, en effet, dans la Revue de Statistique Appliquée. Référence [3].

Nous noterons seulement, que l'application de cette méthode à la prévision séquentielle est simple, lorsque la loi de probabilité $f(x,u)$ possède un résumé exhaustif.

Par ailleurs, le raisonnement effectué pour passer de la loi $f(x,u)$ à la loi $Z(x/x_1, x_2, \dots x_k)$ de la prévision, fait apparaître, comme intermédiaire de calcul, la distribution conjuguée naturelle $a(u/x_1, x_2, \dots, x_k)$ de $f(x, u)$. Cette distribution, qui sera définie de façon précise dans le corps de l'article, peut être utilisée pour la détermination des intervalles de confiance des paramètres ; les résultats obtenus sont le plus souvent identiques à ceux auxquels conduit la méthode de Neymann.

Les distributions $a(u/x_1, x_2, \dots, x_k)$ conjuguées des lois classiques $f(x,u)$ énumérées dans le tableau ci-dessus sont respectivement :

1. Loi Normale
2. Loi Gamma 2 - inverse
3. Loi Gamma
4. Loi Beta
5. Loi Gamma

6. Loi Beta

7. Loi hyperbolique

Le lecteur désireux d'approfondir l'appareillage analytique de la méthode de Bayes pourra se reporter aux références [1] et [2].

I - LE RAISONNEMENT BAYESIEN APPLIQUE A LA PREVISION SEQUENTIELLE

I-1 - Définitions.

Soit une variable aléatoire X dont la distribution de probabilité dépend d'un ensemble de paramètres $u = \{u_1\ u_2 \dots\ u_r\}$. Les valeurs exactes des paramètres n'étant pas connues avec certitude le raisonnement bayésien attribue une distribution de probabilité aux diverses valeurs possibles de ces paramètres qui sont alors considérés comme des variables aléatoires $U = \{U_1\ U_2\ \dots\ U_r\}$.

I-2 - Le raisonnement séquentiel.

Nous allons étudier la séquence des lois de probabilités de X et de U lorsque des réalisations successives $x_1,\ x_2,\ \dots\ x_k$ de la variable X sont observées.

a) Avant toute observation, il faut se donner :

- La loi de probabilité conditionnelle de X, u étant connu : $f(x/u)$. (Cette loi est supposée être la loi exacte du phénomène. Elle devrait être confirmée si un grand nombre d'observations était effectué).

- La loi de probabilité à priori de U : $a_o(u)$. Nous pouvons alors déduire la distribution marginale (à priori) de X :

$$z_o(x) = \int_{D(u)} a_o(u)\ f(x/u)\ du \tag{1}$$

(D(u) étant le domaine de variation de u et cette intégrale étant multiple lorsque le nombre r de dimensions de u est supérieur à 1).

b) Une réalisation particulière x_1 de la variable aléatoire X ayant été observée, les lois de probabilité doivent être adaptées pour tenir compte de cette information.

L'application de l'axiome des probabilités composées conduit aux formules suivantes :

$$a_1\ (u/x_1) = \frac{a_o(u)\ f(x_1/u)}{z_o(x_1)}$$

soit

$$a_1(u/x_1) = \frac{a_o(u)\ f(x_1/u)}{\int_{D(u)} a_o(u)\ f(x_1/u)\ du} \tag{2}$$

(Formule de Bayes).

$a_1(u/x_1)$ est la densité de probabilité de U après observation de x_1. Elle est souvent appelée densité de probabilité "à postériori" de u.

$$z_1(x/x_1) = \int_{D(u)} a_1(u/x_1)\ f(x/u)\ du$$

Soit
$$z_1(x/x_1) = \frac{\int_{D(u)} a_o(u)\ f(x_1/u)\ f(x/u)\ du}{\int_{D(u)} a_o(u)\ f(x_1/u)\ du} \tag{3}$$

$z_1(x/x_1)$ est la densité de probabilité de X après observation de x_1.

c) Après k réalisations de X supposées indépendantes, la répétition du raisonnement conduit aux expressions générales :

$$a_k(u/x_1,\ x_2 \ldots,\ x_k) = \frac{a_o(u)\ f(x_1/u)\ f(x_2/u) \ldots f(x_k/u)}{\int_{D(u)} a_o(u)\ f(x_1/u)\ f(x_2/u) \ldots\ f(x_k/u)\ du} \tag{4}$$

$$z_k(x/x_1\ x_2 \ldots,\ x_k) = \frac{\int_{D(u)} a_o(u)\ f(x_1/u)\ f(x_2/u) \ldots f(x_k/u)\ f(x/u)\ du}{\int_{D(u)} a_o(u)\ f(x_1/u)\ f(x_2/u) \ldots f(x_k/u)\ du} \tag{5}$$

La prévision du phénomène X doit être effectuée au moyen de cette distribution de probabilité z_k. Notons que dans l'expression des densités de probabilité a_k et z_k les dénominateurs ne font pas intervenir les variables aléatoires X et U. Ils jouent le rôle de constantes de normalisation rendant les intégrales des fonctions a_k et z_k égales à 1 et ne dépendant que des observations $x_1\ x_2 \ldots\ x_k$.

I-3 - Lois conjugués naturelles - Résumés exhaustifs.

L'exploitation pratique des expressions (4) et (5) données ci-dessus ne peut être envisagée que si elles conservent une forme stable lorsque le nombre k des observations varie. Cette condition est réalisée si la loi de probabilité f(x/u) possède un résumé exhaustif (cf. Annexe A) ce qui est vrai de la plupart des lois courantes.

Dans ce cas la vraisemblance de l'échantillon peut être mise sous la forme :

$$f(x_1/u) \times f(x_2/u) \ldots f(x_k/u) = r(x_1\ x_2 \ldots\ x_k)\ b(t_k,\ u) \tag{6}$$

(où t_k est le résumé exhaustif de l'échantillon). Les formules (4) et (5) deviennent alors :

$$a_k(u/x_1\ x_2 \ldots\ x_k) = \frac{a_o(u)\ b(t_k,\ u)}{\int_{D(u)} a_o(u)\ b(t_k,\ u)\ du} \tag{7}$$

$$z_k(x/x_1\ x_2 \ldots\ x_k) = \frac{\int_{D(u)} a_o(u)\ b(t_k,\ u)\ f(x/u)\ du}{\int_{D(u)} a_o(u)\ b(t_k,\ u)\ du} \tag{8}$$

Dans ces nouvelles expressions de a_k et z_k, les fonctions a_o, b et f ont la même structure quel que soit k. Il en est donc de même de a_k et z_k qui constituent alors des familles de distributions.

Une nouvelle simplification peut être obtenue en supposant $a_o(u)$ uniforme dans le domaine D(u) de variation des paramètres (Postulat de Bayes). Les expressions (7) et (8) deviennent alors :

$$a_k(u/x_1\ x_2\ \dots\ x_k) = \frac{b(t_k,\ u)}{\int_{D(u)} b(t_k,\ u)\ du} \tag{9}$$

$$z_k(x/x_1\ x_2\ \dots\ x_k) = \frac{\int_{D(u)} b(t_k,\ u)\ f(x/u)\ du}{\int_{D(u)} b(t_k,\ u)\ du} \tag{10}$$

La famille des lois a_k est dite "conjuguée naturelle" de la distribution f(x/u).

La distribution à posteriori $a_k\ (u/x_1\ x_2\ \dots\ x_k)$ peut être utilisée pour calculer des intervalles de confiance des paramètres $u_1\ u_2\ \dots\ u_r$. Les intervalles de confiance ainsi déterminés sont le plus souvent identiques à ceux auxquels conduit la méthode de Neyman. La considération des lois naturelles conjuguées éclaire la signification générale de notre calcul ; il consiste à effectuer la prévision du phénomène sans recourir à une estimation ponctuelle des paramètres mais en tenant compte de l'incertitude attachée à ces paramètres, incertitude caractérisée habituellement par les intervalles de confiance. Cependant pour le calcul des z_k il n'est pas nécessaire de déterminer explicitement l'expression de a_k qui n'est qu'un intermédiaire de calcul. L'étude des lois conjuguées sera donc dans son ensemble repoussée en annexe.

II - APPLICATION A UN PROCESSUS GAUSSIEN

Nous supposons que la variable aléatoire X est distribuée suivant une loi normale

$$f(x/m,\ \sigma) = \frac{1}{\sqrt{2\pi}\,\sigma}\ e^{-\frac{(x-m)^2}{2\sigma^2}} \tag{11}$$

Nous distinguerons 3 cas :

A - Paramètre σ connu, m inconnu et distribué uniformément entre $-M$ et $+M$. Nous ferons tendre ensuite M vers l'infini.

B - Paramètre m connu, σ inconnu, Log σ distribué uniformément entre $-N$ et $+N$ $(N \longrightarrow \infty)$.

C - m et σ inconnus.
m distribué uniformément entre $-M$ et $+M$ $(M \longrightarrow \infty)$.
Log σ distribué uniformément entre $-N$ et $+N$ $(N \longrightarrow \infty)$.

Les raisons pour lesquelles nous avons choisi de distribuer uniformément Log σ et non σ sont les suivantes :

a) Les distributions à posteriori des paramètres conduisent alors à des estimations par intervalles identiques à celles obtenues par la méthode de Neyman.

b) Le nombre de réalisations minimum de X nécessaire pour que la distribution à posteriori de σ ne soit pas nulle en tout point est conforme à l'interprétation de σ comme dispersion : il suffit d'une expérience dans le cas où la moyenne m est connue et il faut deux expériences au minimum si m est σ sont inconnus. Au contraire, si l'on distribue uniformément le paramètre σ lui-même les nombres de réalisations minimum de X dans les cas précédents sont respectivement 2 et 3.

D'une manière générale si l'on distribue uniformément σ^n entre 0 et l'infini il est nécessaire d'effectuer respectivement n + 1 et n + 2 expériences pour que la distribution à postériori ne soit pas nulle en tout point.

Cas A - σ connu, m inconnu uniforme entre $-M$ et M ($M \longrightarrow \infty$).
Avant toute observation :

$$a_o(m) = \frac{1}{2M} \tag{12}$$

$$f(x/m) = \frac{1}{\sqrt{2\pi}\,\sigma}\, e^{-\frac{(x-m)^2}{2\sigma^2}} \tag{13}$$

Supposons que k réalisations $x_1\ x_2 \ldots\ x_k$ de X aient été observées. Posons :

$$\frac{1}{k}\sum_{i=1}^{k} x_i = \bar{x}$$

$$\frac{1}{k}\sum_{i=1}^{k} (x_i - \bar{x})^2 = s^2$$

La vraisemblance de l'échantillon $\{x_1\ x_2 \ldots\ x_k\}$ est alors :

$$\frac{1}{(\sqrt{2\pi}\,\sigma)^k}\, e^{-\frac{\sum_{i=1}^{k}(x_i-m)^2}{2\sigma^2}} = \frac{e^{-\frac{ks^2}{2\sigma^2}}}{(\sqrt{2\pi}\,\sigma)^k}\, e^{-\frac{k(\bar{x}-m)^2}{2\sigma^2}}$$

Considérant la formule (6) nous voyons que $\bar{x}$ est un résumé exhaustif de l'échantillon : le paramètre m dont nous cherchons la loi étant distribué uniformément, à priori, la formule (9) nous donne (après avoir fait tendre M vers l'infini) :

$$a_k(m/x_1\ x_2 \ldots\ x_k) = \frac{e^{-\frac{k(\bar{x}-m)^2}{2\sigma^2}}}{\int_{-\infty}^{+\infty} e^{-\frac{k(\bar{x}-m)^2}{2\sigma^2}}\, dm}$$

soit

$$a_k(m/x_1\ x_2 \ldots\ x_k) = \frac{1}{\sqrt{2\pi}\,\frac{\sigma}{\sqrt{k}}}\, e^{-\frac{k(\bar{x}-m)^2}{2\sigma^2}} \tag{14}$$

La distribution de m à posteriori est donc une loi normale de moyenne $\bar{x}$ et d'écart-type $\frac{\sigma}{\sqrt{k}}$. Elle conduit à une estimation des intervalles de confiance pour m identique à ceux obtenus par la méthode de Neyman.

La formule (10) donne la loi à utiliser pour la prévision :

$$z_k(x/x_1\, x_2 \ldots x_k) = \frac{\dfrac{1}{\sqrt{2\pi}\,\sigma}\displaystyle\int_{-\infty}^{+\infty} e^{-\left[\frac{k(\bar{x}-m)^2+(x-m)^2}{2\sigma^2}\right]}\, dm}{\displaystyle\int_{-\infty}^{+\infty} e^{\frac{k(\bar{x}-m)^2}{2\sigma^2}}\, dm}$$

Soit, en effectuant les intégrations :

$$z_k(x/x_1\, x_2 \ldots x_k) = \frac{1}{\sqrt{2\pi}\,\sigma\sqrt{\dfrac{k+1}{k}}}\, e^{-\frac{k}{2\sigma^2(k+1)}(x-\bar{x})^2} \tag{15}$$

La loi z_k est normale de moyenne $\bar{x}$ et d'écart-type $\sigma\sqrt{\dfrac{k+1}{k}}$.

Une estimation ponctuelle de m par $\bar{x}$ aurait conduit à prendre une loi normale de moyenne $\bar{x}$ et d'écart-type σ pour effectuer les prévisions.

Cas B - m connu, σ inconnu, Log σ distribué uniformément entre $-N$ et $+N$ ($N \longrightarrow \infty$).

Avant toute observation :

$$a_o(\sigma)\, d\sigma = \frac{1}{2N}\, d(\text{Log}\ \sigma) = \frac{1}{2N}\,\frac{d\sigma}{\sigma}$$

Soit $a_o(\sigma) = \dfrac{1}{2N\sigma}$. (16)

$$f(x/\sigma) = \frac{1}{\sqrt{2\pi}\,\sigma}\, e^{-\frac{(x-m)^2}{2\sigma^2}} \tag{17}$$

La vraisemblance d'un échantillon de k observations $x_1\, x_2 \ldots x_k$ s'écrit :

$$\frac{1}{(\sqrt{2\pi})^k}\, e^{-\frac{kA^2}{2\sigma^2}} \times \frac{1}{\sigma^k}$$

en posant $\qquad \dfrac{1}{k}\displaystyle\sum_{i=1}^{k}(x_i - m)^2 = A^2$

D'après l'expression (6) A constitue un résumé exhaustif des observations. L'expression (7) devient alors, en tenant compte de (16) :

$$a_k(\sigma/x_1\, x_2 \ldots x_k) = \frac{\dfrac{1}{(\sqrt{2\pi})^k}\, e^{-\frac{kA^2}{2\sigma^2}} \times \dfrac{1}{\sigma^{k+1}}}{\displaystyle\int_0^{\infty} \frac{1}{(\sqrt{2\pi})^k}\, e^{\frac{kA^2}{2\sigma^2}}\, \frac{1}{\sigma^k}\,\frac{d\sigma}{\sigma}}$$

Soit après calcul de dénominateur

$$a_k(\sigma/x_1\, x_2 \ldots x_k) = \frac{A^k\, k^{\frac{k}{2}}}{2^{\frac{k}{2}-1}\,\Gamma\left(\frac{k}{2}\right)}\,\frac{e^{-\frac{kA^2}{2\sigma^2}}}{\sigma^{k+1}} \tag{18}$$

la quantité $\dfrac{kA^2}{\sigma^2} = \dfrac{\Sigma\,(x_i - m)^2}{\sigma^2}$ (où la variable aléatoire est σ) suit une loi de χ^2 à k degrés de liberté. Les intervalles de confiance obtenus pour σ sont identiques à ceux de la méthode de Neyman. La loi de la prévision est alors (formule 8).

$$z_k(x/x_1\ x_2\ \dots\ x_k) = \frac{\displaystyle\int_0^{\infty} \frac{1}{(\sqrt{2\pi})^{k+1}}\ \frac{e^{-\frac{kA^2+(x-m)^2}{2\sigma^2}}}{\sigma^{k+1}}\ \frac{d\sigma}{\sigma}}{\displaystyle\int_0^{\infty} \frac{1}{(\sqrt{2\pi})^{k}}\ \frac{e^{-\frac{kA^2}{2\sigma^2}}}{\sigma^{k}}\ \frac{d\sigma}{\sigma}}$$

Soit, après calculs

$$z_k(x/x_1\ x_2\ \dots\ x_k) = \frac{1}{\sqrt{\pi}}\ \frac{\Gamma\left(\frac{k+1}{2}\right)}{\Gamma\left(\frac{k}{2}\right)}\ \frac{k^{\frac{k}{2}} A^k}{[kA^2 + (x-m)^2]^{\frac{k+1}{2}}} \tag{19}$$

Cette expression montre que *la quantité* $\frac{(x-m)}{A}$ *suit une loi de student à k degrés de liberté.*

Une estimation ponctuelle de la variance σ^2 aurait conduit à admettre que x suit une loi normale de moyenne m et d'écart-type A, c'est-à-dire, que $\frac{(x-m)}{A}$ *suit une loi normale centrée réduite.*

Cas C - m inconnu, uniforme entre $-M$ et $+M$ ($M \longrightarrow \infty$).
σ inconnu, Log σ uniforme entre $-N$ et $+N$ ($N \longrightarrow \infty$).

La vraisemblance d'un échantillon $\{x_1\ x_2\ \dots\ x_k\}$ s'écrit

$$\frac{1}{(\sqrt{2\pi})^k\ \sigma^k}\ e^{-\frac{\sum_{i=1}^{k}(x_i-m)^2}{2\sigma^2}} = \frac{1}{(\sqrt{2\pi})^k}\ \frac{e^{-\frac{k[s^2+(\bar{x}-m)^2]}{2\sigma^2}}}{\sigma^k} \tag{20}$$

Les quantités $\bar{x}$ et s constituent un résumé de l'échantillon [cf. (6)].

La loi à posteriori de l'ensemble des variables aléatoires (m, σ) sera d'après (7).

$$a_k(m, \sigma/x_1\ \dots\ x_k) = \frac{e^{-\frac{k[s^2+(\bar{x}-m)^2]}{2\sigma^2}}\ \frac{1}{\sigma^{k+1}}}{\displaystyle\int_{\sigma=0}^{\infty} \int_{m=-\infty}^{+\infty} \frac{e^{-\frac{k[s^2+(x-m)^2]}{2\sigma^2}}}{\sigma^k}\ \frac{d\sigma}{\sigma}\ dm}$$

Soit après calcul du dénominateur

$$a_k(m, \sigma/x_1\ x_2\ \dots\ x_k) = \frac{\sqrt{\frac{2k}{\pi}}\left(\frac{ks^2}{2}\right)^{\frac{k-1}{2}}}{\Gamma\left(\frac{k-1}{2}\right)}\ \frac{e^{-\frac{k[s^2+(\bar{x}-m)^2]}{2\sigma^2}}}{\sigma^{k+1}} \tag{21}$$

Cette loi n'est définie que si $k \geqslant 2$.

Les lois marginales à posteriori de m et σ sont obtenues en intégrant l'expression précédente par rapport respectivement à σ et m, ce qui donne :

$$a_k(m/x_1\ x_2 \ldots x_k) = \frac{1}{\sqrt{\pi}} \frac{\Gamma\left(\frac{k}{2}\right)}{\Gamma\left(\frac{k-1}{2}\right)} \frac{s^{k-1}}{[s^2 + (m - \bar{x})^2]^{\frac{k}{2}}} \tag{22}$$

la quantité $\sqrt{k-1}\ \frac{(m - \bar{x})}{s}$ suit une loi de Student à k − 1 degrés de liberté.

$$a_k(\sigma/x_1\ x_2 \ldots x_k) = \frac{1}{2^{\frac{k-3}{2}}\ \Gamma\left(\frac{k-1}{2}\right)}\ e^{-\frac{ks^2}{2\sigma^2}} \left(\frac{ks^2}{\sigma^2}\right)^{\frac{k-1}{2}} \frac{1}{\sigma} \tag{23}$$

la quantité $\frac{ks^2}{\sigma^2}$ suit une loi de χ^2 à k − 1 degrés de liberté. Si en utilisant ces lois à postériori on construit des intervalles de confiance pour m et σ on obtient des résultats identiques à ceux de la méthode de Neyman.

La formule (8) nous donne alors la loi de la prévision pour X :

$$z_k(x/x_1\ x_2 \ldots x_k) = \frac{\frac{1}{\sqrt{2\pi}} \int_{\sigma=0}^{\infty} \int_{m=-\infty}^{+\infty} \frac{e^{-\frac{k[s^2+(\bar{x}-m)^2]+(x-m)^2}{2\sigma^2}}}{\sigma^{k+1}} \frac{d\sigma}{\sigma}\, dm}{\int_{\sigma=0}^{\infty} \int_{m=-\infty}^{+\infty} \frac{e^{-\frac{k[s^2+(\bar{x}-m)^2]}{2\sigma^2}}}{\sigma^k} \frac{d\sigma}{\sigma}\, dm}$$

Soit

$$z_k(x/x_1\ x_2 \ldots x_k) = \frac{1}{\sqrt{\pi}} \frac{\Gamma\left(\frac{k}{2}\right)}{\Gamma\left(\frac{k-1}{2}\right)} \times \frac{1}{\sqrt{k+1}} \frac{s^{k-1}}{\left[s^2 + \frac{(x-\bar{x})^2}{k+1}\right]^{\frac{k}{2}}}$$

Ceci montre que la quantité $\sqrt{k-1}\ \frac{x - \bar{x}}{\sqrt{k+1}\ s}$ *suit une loi de Student à k-1 degrés de liberté.*

Si m et σ^2 avaient été estimés ponctuellement par $\bar{x}$ et

$$\frac{1}{k-1} \sum (x_i - \bar{x})^2 = \frac{k}{k-1}\ s^2$$

la prévision sur x aurait été effectuée en considérant que la quantité

$$\sqrt{k-1}\,\frac{(x-\bar{x})}{\sqrt{k}\,s}$$

suit une loi normale centrée réduite.

III - APPLICATION A D'AUTRES PROCESSUS CLASSIQUES

Le modèle décrit au chapitre I peut être appliqué à tous les processus classiques. Nous ne reprendrons pas l'exposé des calculs qui a été donné dans le cas particulier de la loi normale (cf. chap. 2). Nous indiquerons seulement les lois de prévision correspondant aux processus ci-après :

- Loi de Poisson,
- Loi Binomiale,
- Loi Gamma,
- Loi Binomiale négative,
- Loi uniforme.

3-1 - Loi de Poisson.

La loi de probabilité d'une variable aléatoire X de Poisson est donnée par l'expression :

$$f(x/m) = e^{-m}\frac{m^x}{x!}$$

m inconnu, est supposé a priori distribué uniformément entre O et M ; puis nous faisons tendre M vers l'infini. Nous obtenons la loi de la prévision (après k observations $x_1\ x_2\ \ldots\ x_k$)

$$z_k(x/x_1\ x_2\ \ldots\ x_k) = C^x_{k\bar{x}+x}\frac{k^{k\bar{x}+1}}{(k+1)^{k\bar{x}+x+1}}$$

avec

$$\bar{x} = \frac{1}{k}\sum_{i=1}^{k} x_i$$

x suit donc une loi binomiale négative.

3-2 - Loi Binomiale.

La loi de probabilité d'une variable aléatoire X binomiale est donnée par l'expression :

$$f(x/n,\ p) = C^x_n p^x\,(1-p)^{n-x}$$

Le paramètre n est supposé connu.

p inconnu, est distribué à priori uniformément entre 0 et 1.

La loi de la prévision est :

$$z_k(x/x_1\ x_2 \ldots\ x_k) = \frac{kn+1}{(k+1)\,n+1}\ \frac{C_{nk}^{k\bar{x}}\ .\ C_n^x}{C_{(k+1)n}^{k\bar{x}+x}}$$

x suit une loi hypergéométrique négative.

Remarques :

Si n est inconnu la loi binomiale ne possède pas de résumé exhaustif. Le raisonnement bayesien peut cependant être effectué. Il conduit lorsque n et p sont supposés distribués uniformément, n entre 0 et l'infini par valeurs entières, p entre 0 et 1, à la même loi binomiale négative que dans le cas du Processus de Poisson.

3-3 - Loi gamma.

Le distribution Gamma est de la forme :

$$f(x/u, v) = \frac{e^{-vx}\ (vx)^{u-1}\ v}{\overline{u-1}\ !}$$

Le paramètre u qui caractérise la forme de la loi est supposé connu, alors que v, paramètre d'échelle inconnu, est distribué à priori uniformément entre O et V. Puis nous faisons tendre V vers l'infini et nous obtenons la loi de la prévision :

$$z_k(x/x_1\ x_2 \ldots\ x_k) = \frac{(k\bar{x})^{ku+1}}{B(ku+1,\ u)}\ \frac{x^{u-1}}{(k\bar{x}+x)^{(k+1)u+1}}$$

x suit donc une loi Beta-inverse.

$$\left[\ B(ku+1,\ u) = \frac{(ku)\ !\ \overline{u-1}\ !}{(k+1)\ u\ !}\ \right]$$

Remarques :

Si u est inconnu la loi Gamma ne possède pas de résumé exhaustif. L'application du raisonnement bayésien au cas où u et v sont à priori distribués uniformément entre 0 et l'infini conduit à une loi de probabilité complexe pour les grandes valeurs de k. La densité $z_k(x/x_1\ x_2 \ldots\ x_k)$ possède une particularité analytique intéressante si

$$x_1 = x_2 = \ldots = x_k\ :$$

elle est infinie pour $x = x_1 = \ldots = x_k$ et finie pour toute autre valeur de x.

3-4 - Loi binomiale négative.

La loi de probabilité binomiale négative est de la forme :

$$f(x/n,\ p) = C_{n+x}^n\ p^{n+1}\ (1-p)^x$$

Le paramètre n est supposé connu.

p inconnu, est distribué à priori uniformément entre 0 et 1. Nous obtenons la loi de la prévision :

$$z_k(x/x_1\ x_2 \dots\ x_k) = \frac{1 + k(n+1)}{1 + (k+1)(n+1)}\ C_{kn+k\bar{x}+1}^{k(n+1)+1}\ \frac{C_{n+x}^{n}}{C_{(k+1)(n+1)+k\bar{x}+x+1}^{(k+1)(n+1)+1}}$$

Cette loi n'a pas à notre connaissance de désignation classique.

Remarque :

Si n est inconnu la remarque faite dans le cas de la loi binomiale s'applique identiquement.

3-5 - Loi uniforme.

La loi de probabilité uniforme (0, u) est de la forme :

$$\begin{cases} f(x/u) = \dfrac{1}{u} & \text{si} \quad x \leqslant u \\ f(x/u) = 0 & \text{si} \quad x > u \end{cases}$$

Nous supposons à priori Log u distribué uniformément entre 0 et l'infini.

Nous obtenons la loi de la prévision :

$$z_k(x/x_1\ x_2 \dots\ x_k) = \frac{k}{k+1}\ \frac{[\mathrm{Max}\ (x_1,\ x_2 \dots,\ x_k)]^k}{[\mathrm{Max}\ (x_1,\ x_2 \dots,\ x_k, x)]^{k+1}}$$

Cette loi est uniforme pour $0 \leqslant x \leqslant \mathrm{Max}\ (x_1,\ x_2 \dots\ x_k)$.

Elle prend une forme hyperbolique pour $x > \mathrm{Max}\ (x_1,\ x_2 \dots,\ x_k)$ ($\mathrm{Max}\ (x_1,\ x_2 \dots\ x_k)$ désigne la plus grande des valeurs $x_1\ x_2 \dots\ x_k$).

ANNEXE A

RESUMES EXHAUSTIFS

1/ Définition Bayésienne des résumés exhaustifs :

Soit une application $t_k\ (x_1, x_2 \dots\ x_k)$ de l'ensemble des observations sur un espace R_p, le nombre de p de dimensions étant fixe alors que le nombre k d'observations est supposé variable.

Par définition, t est un résumé exhaustif pour la loi f(x/u), si pour toute distribution à priori $a_o(u)$ des paramètres $(u_1\ u_2 \dots\ u_r)$ la densité de probabilité à postériori $a_k(u/x_1\ x_2 \dots\ x_k)$ ne dépend que de t et de u ; c'est-à-dire si pour deux échantillons quelconques $(x_1,\ x_2 \dots,\ x_k)$ et $(x_1',\ x_2' \dots\ x_k')$ tels que $t_k(x_1 \dots\ x_k) = t_k(x_1'\ x_2' \dots\ x_k')$ on a :

$$a_k(u/x_1\ x_2 \dots\ x_k) = a_k(u/x_1'\ x_2' \dots\ x_k') \qquad (A_1)$$

2/ Théorème (2e définition des résumés exhaustifs).

La condition nécessaire et suffisante pour que t soit un résumé exhaustif pour une loi f(x/u) est que la vraisemblance d'un échantillon $(x_1,\ x_2 \dots,\ x_k)$ puisse s'écrire sous la forme

$$V(x_1\ x_2 \dots\ x_k/u) = b(t_k,\ u)\ .\ r(x_1\ x_2 \dots\ x_k) \qquad (A_2)$$

Démonstration :

a) Condition suffisante : compte tenu de l'équation (4) l'égalité (A_2) est équivalente à :

$$a_k(u/x_1\ x_2 \dots\ x_k) = \frac{r(x_1\ x_2 \dots\ x_k)}{\int_{D(u)} a_o(u)\ V(x_1\ x_2 \dots\ x_k/u)\,du}\ .\ a_o(u)\ b(t_k,\ u) \qquad (A_3)$$

L'intégrale de cette densité de probabilité pour les diverses valeurs possibles de u devant être égale à 1 on doit avoir :

$$\frac{r(x_1\ x_2 \dots\ x_k)}{\int_{D(u)} a_o(u)\ V(x_1\ x_2 \dots\ x_k/u)\ du} = \frac{1}{\int_{D(u)} a_o(u)\ b(t_k,\ u)\ du} \qquad (A_4)$$

L'équation A_3 devient alors :

$$a_k(u/x_1\ x_2 \dots\ x_k) = \frac{a_o(u)\ b(t_k,\ u)}{\int_{D(u)} a_o(u)\ b(t_k,\ u)\,du} \qquad (A_5)$$

donc $a_k(u/x_1\ x_2 \dots\ x_k)$ ne dépend que de t_k.

b) Condition nécessaire : supposons que $a_k(u/x_1 \dots\ x_k)$ ne dépende que de t_k et de u ; nous pouvons poser :

$$a_k(u/x_1\ x_2 \dots\ x_k) = B(t_k,\ u) \qquad (A_6)$$

d'après l'équation (4) il vient :

$$B(t_k,\ u) = \frac{a_o(u)\ V(x_1\ x_2 \dots\ x_k/u)}{\int_{D(u)} a_o(u)\ V(x_1\ x_2 \dots\ x_k/u)\,du} \qquad (A_7)$$

Le dénominateur du second membre est une fonction de $x_1, x_2 \dots, x_k$. Nous posons :

$$\int_{D(u)} a_o(u)\ V(x_1\ x_2 \dots\ x_k/u)\ du = r(x_1\ x_2 \dots\ x_k) \qquad (A_8)$$

Il en résulte, en posant encore $b(t_k,\ u) = \dfrac{B(t_k,\ u)}{a_o(u)}$:

$$V(x_1\ x_2 \dots\ x_k/u) = r(x_1\ x_2 \dots\ x_k)\ .\ b(t_k,\ u) \qquad (A_9)$$

3/ Théorème de Darmois, Pitman - Koopman.

Une condition nécessaire et suffisante pour qu'une loi de probabilité $f(x/u)$ à r paramètre $u = \{u_1\ u_2, \dots\ u_r\}$ possède un résumé exhaustif est qu'elle puisse être mise sous la forme

$$f(x/u) = A(u)\ C(x)\ e^{\sum_{i=1}^{r} D_i(u)C_i(x)}$$

La plupart des lois de probabilités classiques peuvent être mises sous cette forme et possèdent par conséquent des résumés exhaustifs.

ANNEXE B - LOIS CONJUGUEES

Les lois conjuguées des processus classiques étudiés au chapitre III sont :

Processus	Loi conjuguée
- Loi de Poisson	- Loi Gamma
$f(x/m) = e^{-m} \dfrac{m^x}{x!}$	$a_k(m/x_1 \dots x_k) = \dfrac{e^{-km} (km)^{k\bar{x}} k}{(k\bar{x})!}$
m uniforme $(0\ \infty)$	
- Loi Binomiale	- Loi Beta
$f(x/n, p) = C_n^x p^x (1-p)^{n-x}$	$a_k(m/x_1 \dots x_k) =$
p uniforme (0, 1)	$\dfrac{kn+1\,!}{(k\bar{x})!\, k(n-\bar{x})!} p^{k\bar{x}} (1-p)^{k(n-\bar{x})}$
- Loi Gamma	- Loi Gamma
$f(x/u, v) = \dfrac{e^{-vx} (vx)^{u-1} v}{u-1\,!}$	$a_k(v/x_1 \dots x_k) = \dfrac{e^{-k\bar{x}v} (k\bar{x}v)^{ku} k\bar{x}}{(ku)!}$
v uniforme $(0, \infty)$	
- Loi Binomiale négative	- Loi Beta
$f(x/n, p) = C_{n+x}^n p^{n+1} (1-p)^x$	$a_k(t/x_1 \dots x_k) =$
p uniforme (0, 1)	$\dfrac{[(k(n+\bar{x}+1)+1]\,!}{[k(n+1)]\,!\,(k\bar{x})\,!} p^{k(n+1)} (1-p)^{k\bar{x}}$
- Loi uniforme (0, u)	- Loi Hyperbolique
$f(x/u) = \dfrac{1}{u}$ Pour $x \leqslant u$	$a_k(u/x_1 \dots x_k) =$
$f(x/u) = 0$ pour $x > u$	$\dfrac{k(\text{Max } x_1, x_2, \dots x_k)}{u^{k+1}}$
Log u uniforme $(0, \infty)$	si $u \geqslant \text{Max}(x_1\ x_2 \dots x_k)$
	$a_k(u/x_1 \dots x_k) = 0$
	si $u < \text{Max}(x_1, x_2 \dots x_k)$

BIBLIOGRAPHIE

[1] JEFFREYS - Theory of probability. Clarendon Press. Oxford.

[2] SCHLAIFER and RAIFFA - Applied Statistical decision Theory. Harvard.

[3] D. GROUCHKO - Application de la formule de Bayes au calcul d'un volant de rechanges pour des pièces à faible consommation. Revue de statistique appliquée, 1964, Vol. 12, n° 4.

6