

# REVUE DE STATISTIQUE APPLIQUÉE

HAYET MERABET

## **Méthodologie des essais préliminaires. Calcul de prévision de satisfaction dans le cas gaussien**

*Revue de statistique appliquée*, tome 52, n° 3 (2004), p. 93-106

[http://www.numdam.org/item?id=RSA\\_2004\\_\\_52\\_3\\_93\\_0](http://www.numdam.org/item?id=RSA_2004__52_3_93_0)

© Société française de statistique, 2004, tous droits réservés.

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

NUMDAM

Article numérisé dans le cadre du programme  
Numérisation de documents anciens mathématiques

<http://www.numdam.org/>

# MÉTHODOLOGIE DES ESSAIS PRÉLIMINAIRES. CALCUL DE PRÉVISION DE SATISFACTION DANS LE CAS GAUSSIEN

Hayet MERABET

*Département de Mathématiques, Faculté des Sciences, Université Mentouri,  
25000 Constantine, Algérie, merabet1@caramail.com*

## RÉSUMÉ

Le présent article propose les notions d'indice et de prévision de satisfaction relatifs à un test d'hypothèse pour certains protocoles d'essais cliniques en deux étapes dans le modèle Gaussien.

*Mots-clés : Essais cliniques, Indices de satisfaction, Méthodes prédictives.*

## ABSTRACT

The notions of satisfaction index and of prevision index adapted to some 2 steps protocols for clinical trials are presented in the gaussian model.

*Keywords : Clinical trials, Satisfaction index, Predictive methods.*

## 1. Problématique : Indice de prévision de satisfaction

### 1.1. Le modèle

En [5], nous avons introduit, par référence à des contextes expérimentaux, la notion d'*indice de satisfaction* relatif à un test d'hypothèse, dans un cadre où un tel test peut être construit à l'aide d'une fonction de test raisonnable.

Précisons ce cadre.

Soit un modèle  $(P_\theta)_{\theta \in \Theta}$  et soit à tester l'hypothèse nulle  $\theta \in \Theta_0$  contre l'hypothèse alternative  $\theta \in \Theta_1$ , définie à l'aide d'une application  $\psi$  de  $\Theta$  dans  $\mathbb{R}$  : on suppose qu'il existe un point  $t_0$  tel que

$$\theta \in \Theta_0 \iff \psi(\theta) \leq t_0.$$

Supposons par ailleurs qu'on dispose d'une application réelle  $\xi$ , définie sur l'ensemble des observations, telle que la loi de  $\xi$  pour la valeur  $\theta$  du paramètre ne dépende que de  $\psi(\theta)$ , et soit *stochastiquement croissante* avec  $\psi(\theta)$  : plus  $\psi(\theta)$  est élevé, plus  $\xi$  a

tendance à prendre de grandes valeurs, c'est-à-dire que

$$\psi(\theta_1) < \psi(\theta_2) \implies \forall t \ P_{\theta_1}[\xi \leq t] \geq P_{\theta_2}[\xi \leq t]$$

Un test, au seuil  $\alpha$ , de  $\Theta_0$  contre  $\Theta_1 (= \{\theta, \psi(\theta) > t_0\})$  est alors naturellement défini en rejetant l'hypothèse si le résultat expérimental, soit  $\underline{y}$ , vérifie  $\xi(\underline{y}) > g(\alpha)$ , où  $g(\alpha)$  est le  $(1 - \alpha)$  fractile de la loi de  $\xi$  quand  $\psi(\theta) = t_0$ .

En effet, la région critique de ce test,  $C$ , est alors l'ensemble des observations  $\underline{y}$  qui vérifient  $\xi(\underline{y}) > g(\alpha)$ , et on a bien, en raison de la croissance stochastique des lois de  $\xi$ , et pour tout  $\theta_0$  vérifiant  $\psi(\theta_0) = t_0$ ,

$$\forall \theta \in \Theta_0 \ P_{\theta}(C) \leq P_{\theta_0}(C) \leq 1 - \alpha.$$

De plus, si  $\theta_1 \in \Theta_1$  et  $\theta_2 \in \Theta_1$  avec  $\psi(\theta_1) < \psi(\theta_2)$ , il vient  $P_{\theta_1}(C) \leq P_{\theta_2}(C)$ , c'est-à-dire que la fonction puissance du test croît avec  $\psi(\theta)$ .

### 1.2. Indice de satisfaction

Cette notion trouve son origine dans des situations où le statisticien qui procède à un test «souhaite» détecter un effet significatif, c'est-à-dire rejeter l'hypothèse. Corrélativement, ce statisticien est d'autant plus satisfait que, en fonction du résultat expérimental, cet effet paraît plus significatif. C'est ce que les utilisateurs mettent souvent en évidence en donnant, à l'issue de la procédure de test, non seulement une conclusion «en tout ou rien» (significatif ou non significatif) mais aussi la plus petite valeur du seuil  $p$  pour laquelle le résultat  $\underline{y}$  obtenu serait considéré comme significatif.  $p$  est classiquement appelé la  $p$ -valeur et est dans ce cas

$$p(\underline{y}) = P_{\theta_0}[\xi > \xi(\underline{y})].$$

Un indice de satisfaction, pour le test de seuil  $\alpha$  considéré, se définit donc naturellement comme une application de l'ensemble des résultats dans  $\mathfrak{R}_+$  qui :

- prend la valeur 0 si on ne rejette pas l'hypothèse (c'est-à-dire si  $\xi(\underline{y}) \leq g(\alpha)$ ),
- est, si  $\xi(\underline{y}) > g(\alpha)$ , une fonction décroissante de  $P_{\theta_0}[\xi > \xi(\underline{y})]$ , qu'on notera  $L(P_{\theta_0}(\xi > \xi(\underline{y})))$  autrement dit  $L(1 - F_{\theta_0}(\xi(\underline{y})))$  où  $F_{\theta_0}$  est la fonction de répartition de  $\xi$  à la frontière, c'est-à-dire pour tout  $\theta_0$  tel que  $\psi(\theta_0) = t_0$ .

L'indice de satisfaction ainsi défini sera noté  $S_{L,\alpha}$ . On a donc  $S_{L,\alpha}(\underline{y}) = 0$  si  $\xi(\underline{y}) \leq g(\alpha)$ , c'est-à-dire si  $p(\underline{y}) \geq 1 - \alpha$  et  $S_{L,\alpha}(\underline{y}) = L(1 - F_{\theta_0}(\xi(\underline{y}))) = L(p(\underline{y}))$  sinon.

Un indice particulièrement fruste est la fonction indicatrice de la région critique mais il ne rend pas compte de l'intérêt pour la  $p$ -valeur évoqué ci-dessus.

Nous nous attacherons ici à l'étude de deux familles d'indices bornés ou non bornés :

- a) Indices de type 1 définis par :

$$L(p) = (1 - p)^l \text{ où } l \geq 0$$

(dans le cas où  $l = 1$ ,  $1 - S_\alpha(y)$  est donc la  $p$ -valeur et dans le cas où  $l = 0$ , on retrouve la fonction indicatrice de la région critique).

On notera ces indices  $S_{1,l,\alpha}$ ; ils sont bornés par 1.

b) Indices de type 2 définis par :

$$L(p) = \left(\frac{1}{p}\right)^l \text{ où } l > 0$$

On notera ces indices  $S_{2,l,\alpha}$ ; ils ne sont pas bornés.

### 1.3. La prévision de satisfaction

Cette notion s'introduit quand, comme c'est souvent le cas pour des essais cliniques, (voir par exemple [3]), on doit mener une expérimentation en deux temps :

- un premier résultat,  $\underline{x}$ , sert à déterminer si on poursuit ou non l'expérimentation,
- si on poursuit effectivement l'expérimentation, le second résultat,  $\underline{y}$ , sert seul à fonder un test au seuil  $\alpha$ , préalablement choisi (voire imposé dans les protocoles cliniques).

La poursuite n'intervient donc que si on considère, sur le vu du premier résultat, qu'il y a «suffisamment de chances» que le second résultat apparaisse satisfaisant.

Suivant [2], nous adoptons ici, pour la décision de poursuite, un comportement de type bayésien. Le choix d'une probabilité *a priori* sur  $\Theta$  permet alors de donner un sens à la loi du second résultat,  $\underline{y}$ , conditionnellement au premier,  $\underline{x}$ . Cette loi ne fait plus intervenir le paramètre  $\theta$ , relativement auquel on a intégré; c'est ce qu'on appelle classiquement, en statistique bayésienne, la loi prédictive de  $\underline{y}$  conditionnellement à  $\underline{x}$ .

La prévision de satisfaction est alors définie comme l'espérance mathématique, pour cette loi prédictive conditionnelle, de l'indice de satisfaction relatif au test que l'on pourrait opérer en seconde phase. On la notera  $I_\alpha(\underline{x})$ .

C'est au praticien qu'il appartient de décider en-dessous de quelle valeur de la prévision de satisfaction il renonce à la poursuite de l'expérience.

Dans Grouin [2,4] est en fait utilisée à cette fin la probabilité prédictive conditionnelle du rejet de l'hypothèse nulle du test de la seconde étape, autrement dit la probabilité prédictive conditionnelle de la conclusion «le résultat  $\underline{y}$  est significatif». Il s'agit donc, en notre sens, de l'indice de prévision associé à l'indice de satisfaction rudimentaire qu'est  $S_{1,0,\alpha}$  la fonction indicatrice de la région critique.

Si on note  $P_{\underline{x}}$  la loi prédictive conditionnelle du second résultat sachant que le premier vaut  $\underline{x}$ , il vient

$$\begin{aligned} I_\alpha(\underline{x}) &= \int S_{L,\alpha}(\underline{y}) P_{\underline{x}}(d\underline{y}) \\ &= \int_{\{\underline{y}; \xi(\underline{y}) > g(\alpha)\}} L(1 - F_{\theta_0}[\xi(\underline{y})]) P_{\underline{x}}(d\underline{y}) \end{aligned}$$

Nous nous intéressons dans la suite à calculer, dans un modèle gaussien, les prévisions de satisfaction  $I(\underline{x})$  associées aux indices  $S_{1,l,\alpha}$  et  $S_{2,l,\alpha}$  définis ci-dessus.

## 2. La prévision de satisfaction dans le modèle gaussien

### 2.1. Introduction du modèle

On effectue des observations indépendantes et de même loi normale  $N(\theta, \sigma^2)$ . Dans toute la suite,  $\Phi$  (resp  $\varphi$ ) désigne la fonction de répartition (resp la densité) de la loi  $N(0, 1)$ .

Le premier résultat,  $\underline{x}$ , est une suite  $(x_1, \dots, x_k)$  de  $k$  observations et le second résultat,  $\underline{y}$ , est une suite  $(y_1, \dots, y_n)$ .

Pour des raisons d'exhaustivité évidentes, on va fonder tous les calculs sur  $x = \frac{1}{k} \sum_{i=1}^k x_i$  et  $y = \frac{1}{n} \sum_{j=1}^n y_j$ , de lois respectives  $N(\theta, \sigma_1^2)$  et  $N(\theta, \sigma_2^2)$ , où

$$\sigma_1^2 = \frac{\sigma^2}{k} \text{ et } \sigma_2^2 = \frac{\sigma^2}{n}$$

On suppose ici  $\sigma^2$  connu.  $\theta$  est inconnu et on choisit pour loi *a priori* pour  $\theta$  la conjuguée naturelle (voir [7]), c'est-à-dire ici la loi normale  $\mu = N(\delta, \tau^2)$ . On désire tester une hypothèse nulle de type  $\theta \leq \theta_0$ .

La loi du résultat  $y$  est évidemment stochastiquement croissante en  $\theta$ .

### 2.2. Le test et les indices

On utilise ici un test usuel portant sur  $y$ , dont la région critique est  $]q_0, +\infty[$ , où  $q_0 = \theta_0 + \sigma_2 u_\alpha^+$ ,  $u_\alpha^+$  désignant le  $\alpha$  quantile supérieur de  $N(0, 1)$  :  $\Phi(u_\alpha^+) = 1 - \alpha$ . La prévision de satisfaction est donnée par

$$I_{L,\alpha}(x) = \int_{q_0}^{+\infty} L \left( 1 - \Phi \left( \frac{y - \theta_0}{\sigma_2} \right) \right) f_x(y) dy \quad (1)$$

où  $f_x(y)$  est la densité de la loi prédictive conditionnelle de  $y$  sachant  $x$ .

### 2.3. Calcul de la loi prédictive conditionnelle

Les notations  $X$  et  $Y$  désignent les variables aléatoires dont les observations respectives sont  $x$  et  $y$ .

On utilisera les symboles suivants :

–  $E_\theta$  désigne l'espérance mathématique dans le modèle d'échantillonnage, à  $\theta$  fixé,

–  $E_\mu$  désigne l'espérance, relativement à l'*a priori*, et s'applique donc aux fonctions de  $\theta$ ,

–  $E$  désigne l'espérance relative à la loi prédictive.

On a les mêmes conventions d'indices pour la variance ( $V$ ) et la covariance ( $C$ ).

Rappelons les caractéristiques de la loi prédictive :

$$\begin{aligned} E(X) &= E_\mu(E_\theta X) = \delta = E(Y), \\ V(X) &= E_\mu(V_\theta(X)) + V_\mu(E_\theta(X)) = \sigma_1^2 + \tau^2, \\ V(Y) &= \sigma_2^2 + \tau^2, \\ C(X, Y) &= E_\mu(E_\theta(X) E_\theta(Y)) - \delta^2 = \delta^2 + \tau^2 - \delta^2 = \tau^2. \end{aligned}$$

La loi du triplet  $(\theta, X, Y)$  est une loi normale tridimensionnelle. Il en résulte que la loi prédictive, c'est-à-dire la loi marginale du couple  $(X, Y)$ , est une loi normale bidimensionnelle. On déduit alors la loi conditionnelle de  $Y$  sachant  $x$ , à l'aide des formules générales suivantes sur la loi normale bidimensionnelle (voir [1]), pour lesquelles on adopte les symboles  $E_x$  et  $V_x$  pour désigner l'espérance et la variance de la loi prédictive sachant  $X = x$  :

$$\begin{aligned} E_x(Y) &= E(Y) + \frac{C(X, Y)}{V(X)}x - \frac{C(X, Y)}{V(X)}E(X) \\ V_x(Y) &= V(Y) \left( 1 - \frac{C^2(X, Y)}{V(X)V(Y)} \right) \end{aligned}$$

On obtient donc, dans le modèle gaussien considéré ci-dessus :

$$\begin{aligned} E_x(Y) &= \frac{\tau^2 x + \sigma_1^2 \delta}{\sigma_1^2 + \tau^2} \\ V_x(Y) &= \sigma_2^2 + \frac{\tau^2 \sigma_1^2}{\sigma_1^2 + \tau^2} \end{aligned}$$

Dans (1),  $f_x$  est la densité de la loi normale avec les caractéristiques ci-dessus, mais, comme on ne dispose pas de forme explicite de la fonction  $\Phi$  à l'aide de fonctions élémentaires, on ne peut pas donner une forme analytique pour  $I_{L,\alpha}(x)$ . On va donc indiquer comment procéder à son évaluation numérique.

### 3. Calcul numérique de la prévision

#### 3.1. Expression formelle

On rappelle que, étant donnée  $\sigma_1^2, \sigma_2^2, \delta, \alpha$  et  $x$ , on veut calculer

$$I_{L,\alpha}(x) = \frac{1}{s} \int_{q_0}^{+\infty} L \left( 1 - \Phi \left( \frac{y - \theta_0}{\sigma_2} \right) \right) \varphi \left( \frac{y - m}{s} \right) dy \quad (2)$$

où

$$\begin{aligned} m &= \frac{\tau^2 x + \sigma_1^2 \delta}{\sigma_1^2 + \tau^2} \\ s^2 &= \sigma_2^2 + \frac{\tau^2 \sigma_1^2}{\sigma_1^2 + \tau^2} \\ q_0 &= \theta_0 + \sigma_2 u_\alpha^+ \end{aligned}$$

Par changement de variable, on obtient

$$I_{L,\alpha}(x) = \int_{-bx' + d\theta'_0 + tu_\alpha^+}^{\infty} L \left( 1 - \Phi \left( \frac{z - (-bx' + d\theta'_0)}{t} \right) \right) \varphi(z) dz$$

où

$$\begin{aligned} x' &= \frac{x - \delta}{\tau}, \quad \theta'_0 = \frac{\theta_0 - \delta}{\tau}, \quad \sigma'_1 = \frac{\sigma_1}{\tau}, \quad \sigma'_2 = \frac{\sigma_2}{\tau}, \\ b &= \left[ (1 + \sigma_1'^2)(\sigma_1'^2 + \sigma_2'^2 + \sigma_1'^2 \sigma_2'^2) \right]^{-\frac{1}{2}}, \\ d &= (1 + \sigma_1'^2)b, \\ t &= \sigma_2' d. \end{aligned}$$

On constate que  $I_{L,\alpha}(x)$  ne dépend que de deux nombres réels que nous appellerons *paramètres essentiels* : un paramètre d'échelle  $t$  dans l'expression duquel n'interviennent que les rapports de variances

$$\frac{\sigma_1^2}{\tau^2} \quad \text{et} \quad \frac{\sigma_2^2}{\tau^2}.$$

et un paramètre de position

$$a = -bx' + d\theta'_0.$$

A seuil  $\alpha$  et à paramètres d'échelle fixés, une modification de  $\theta_0$  et  $\delta$  a un effet de translation sur  $I_{L,\alpha}$  : si  $\theta'_0$  augmente de  $\Delta\theta'_0$ , la courbe représentative de  $I_{L,\alpha}$  subit une translation horizontale d'amplitude  $\frac{d}{b}\Delta\theta'_0$  où  $\frac{d}{b} = 1 + \sigma_1'^2$ .

### 3.2. La prévision de satisfaction est-elle finie ?

#### a) Indices de type 1

La prévision de satisfaction ainsi définie est finie dans le cas des indices de type 1,  $S_{1,l,\alpha}$ , quel que soit  $l > 0$ , car la satisfaction est dans ce cas bornée par 1; il en est donc de même de la prévision associée,  $I_{1,l,\alpha}$ , où

$$I_{1,l,\alpha} = \int_{a+tu_\alpha^+}^{+\infty} \left[ \Phi \left( \frac{z - a}{t} \right) \right]^l \varphi(z) dz.$$

## b) Indices de type 2

Là se pose le problème de la convergence de l'intégrale

$$I_{2,l,\alpha} = \int_{a+tu_{\alpha}^+}^{\infty} \left[ 1 - \Phi \left( \frac{z-a}{t} \right) \right]^{-l} \varphi(z) dz.$$

On sait que la fonction de répartition de la loi normale centrée réduite vérifie, au voisinage de  $+\infty$

$$1 - \Phi(u) \sim \frac{1}{\sqrt{2\pi}} \frac{1}{u} e^{-\frac{u^2}{2}}.$$

Donc, dans  $I_{2,l,\alpha}$ , l'intégrande  $h$  vérifie :

$$h(z) = \left[ 1 - \Phi \left( \frac{z-a}{t} \right) \right]^{-l} \varphi(z) \sim (2\pi)^{\frac{l-1}{2}} \frac{(z-a)^l}{t^l} \exp \left[ -\frac{1}{2} \left( z^2 - l \left( \frac{z-a}{t} \right)^2 \right) \right].$$

On distingue alors deux cas

a)  $l \neq t^2$ . Alors en posant

$$u = z + \frac{al}{t^2 - l},$$

il vient

$$h(z) \sim \frac{(2\pi)^{\frac{l-1}{2}}}{t^l} \exp \left( +\frac{1}{2} \frac{la^2}{t^2 - l} \right) u^l \exp - \frac{1}{2} \left( \frac{t^2 - l}{t^2} u^2 \right)$$

et l'intégrale est divergente si  $l > t^2$  et convergente si  $l < t^2$ .

b)  $l = t^2$ . Alors

$$h(z) \sim \frac{(2\pi)^{\frac{l^2-1}{2}}}{t^2} e^{\frac{a^2}{2}} z^l e^{-az}$$

et l'intégrale est divergente si  $a \leq 0$  et convergente si  $a > 0$ .

Etudions quel est le poids de ces contraintes, faisant intervenir les paramètres essentiels  $a$  et  $t$ , pour assurer l'existence d'une prévision de satisfaction finie.

On remarque que

$$t = \sigma'_2 d = \frac{\sigma'_2 \sqrt{1 + \sigma_1'^2}}{\sqrt{\sigma_1'^2 + \sigma_2'^2 + \sigma_1'^2 \sigma_2'^2}} < 1.$$



La condition  $l < t^2$  impose donc que la fonction  $L$ , si elle est de type puissance, converge lentement vers  $\infty$  quand  $p$  tend vers 0; en particulier le cas  $l = 1$ , c'est-à-dire

$$S_{2,1,\alpha} = \frac{1}{1 - F_{\theta_0}(\xi(\underline{y}))}$$

conduit ici à une prévision de satisfaction infinie quel que soit le résultat expérimental, donc inutilisable.

Dans le cas où  $l = t^2$  (condition qui ne fait intervenir que les variances connues, mais ne dépend pas des observations), la condition  $a > 0$ , signifie que

$$x' < \frac{d}{b}\theta'_0 = (1 + \sigma_1'^2)\theta'_0;$$

la prévision de satisfaction est donc finie pour  $x'$  strictement inférieur à un certain seuil, et infinie au delà.

### 3.3. Méthode de Monte-Carlo

Pour procéder au calcul de  $I_{L,\alpha}(x)$  par une méthode de Monte-Carlo, on le réécrit sous la forme

$$I_{L,\alpha}(x) = [1 - \Phi(a + tu_\alpha^+)] \int_{\mathfrak{R}} L \left( 1 - \Phi \left( \frac{z - a}{t} \right) \right) \frac{\varphi(z)}{1 - \Phi(a + tu_\alpha^+)} 1_{[a + tu_\alpha^+, \infty[}(z) dz,$$

où  $\frac{\varphi}{1 - \Phi(a + tu_\alpha^+)} 1_{[a + tu_\alpha^+, \infty[}$  est la densité de la probabilité  $Q$ , déduite de la loi normale centrée réduite par conditionnement par l'événement  $[a + tu_\alpha^+, \infty[$ .

La méthode de Monte-Carlo consiste alors à approcher  $I_{L,\alpha}(x)$  par

$$[1 - \Phi(a + tu_\alpha^+)] \left[ \frac{1}{N} \sum_{i=1}^N L \left( 1 - \Phi \left( \frac{Z_i - a}{t} \right) \right) \right],$$

où les  $Z_i$  sont  $N$  réalisations de la probabilité  $Q$ . Le tirage des  $Z_i$  se déroule de la manière suivante :

$U_i$  est tiré selon la loi uniforme sur  $[0, 1]$ ,

$V_i = \Phi(a + tu_\alpha^+) + (1 - \Phi(a + tu_\alpha^+))U_i$ , c'est-à-dire que  $V_i$  suit la loi uniforme sur  $[\Phi(a + tu_\alpha^+), 1]$ ,

$Z_i = \Phi^{-1}(V_i)$ , c'est-à-dire que  $Z_i$  suit la loi  $Q$ .

### 3.4. Présentation des résultats

On trouvera en annexe des courbes représentatives de  $I_{L,\alpha}$  en fonction de l'observation  $x = \frac{1}{k} \sum_{i=1}^k x_i$ .

Dans chaque graphique, on a pris  $\delta = 0$  et  $\tau = 1$ , ce qui ne diminue en rien la généralité puisque les paramètres essentiels ne dépendent de  $x$  et  $\theta_0$  que via

$$x' = \frac{x - \delta}{\tau} \quad \text{et} \quad \theta'_0 = \frac{\theta_0 - \delta}{\tau}.$$

D'un graphique à l'autre varient donc le choix de  $L$  et les variances  $\sigma_1^2$  et  $\sigma_2^2$ . On a adopté pour  $\alpha$  la valeur 0,05.

Pour  $\sigma_1^2$  et  $\sigma_2^2$ , on a considéré des situations où, dans le premier comme dans le second échantillon, les observations sont de même variance unitaire  $\sigma^2$ , mais où les effectifs peuvent varier. On a pris :

– d'une part  $\sigma^2 = 1$  et  $\sigma^2 = 4$ , (autrement dit le rapport des écarts-types des observations à l'écart-type de l'a priori est 1 ou 2),

– d'autre part  $k = 10$  et  $n = 10$  ou 20 (autrement dit le rapport des effectifs de la deuxième phase éventuelle à la première phase exploratoire est 1 ou 2).

Le tableau ci-dessous donne, dans chacun de ces cas, les valeurs de  $\sigma_1^2$  et  $\sigma_2^2$ , ainsi que les paramètres essentiels  $t$  et  $a$ , ce dernier en fonction de  $x$  et  $\theta_0$ .

$n$	$\sigma^2 = 1$
10	cas A ( $\sigma_1^2 = 0, 1; \sigma_2^2 = 0, 1; t = 0, 724; a = -2, 0806x + 2, 2887\theta_0$ )
20	cas B ( $\sigma_1^2 = 0, 1; \sigma_2^2 = 0, 05; t = 0, 596; a = -2, 4218x + 2, 664\theta_0$ )
$n$	$\sigma^2 = 4$
10	cas C ( $\sigma_1^2 = 0, 4; \sigma_2^2 = 0, 4; t = 0, 764; a = -0, 8626x + 1, 2076\theta_0$ )
20	cas D ( $\sigma_1^2 = 0, 4; \sigma_2^2 = 0, 2; t = 0, 642; a = -1, 0249x + 1, 4349\theta_0$ )

Les tracés sont tous effectués pour  $\theta_0 = 0$ ; les tracés relatifs à une valeur  $\theta_0 \neq 0$  s'en déduiraient par une translation d'amplitude  $(1 + \sigma_1^2)\theta_0$  (cf. fin du § 3.1). On a pris d'autre part  $\alpha = 0,05$ . Les graphiques 1 à 4 représentent simultanément, pour respectivement chacun des cas A, B, C et D, les courbes de prévision de satisfaction relatives aux indices  $S_{1,1,\alpha}$  et  $S_{1,0,\alpha}$ ; on rappelle que cette dernière n'est autre que la probabilité prédictive de la région critique, déjà utilisée en [2].

On a utilisé comme nombre de répétitions pour la méthode de Monte-Carlo  $N = 50$  et tracé les courbes avec un pas de 0,05 pour  $x$ . On remarque que ces courbes, très voisines pour les valeurs de  $x$  négatives ou faiblement positives se distinguent pour  $x$  assez grand, c'est-à-dire assez nettement supérieur à  $\theta_0$ ; ceci traduit bien le fait que la prise en compte, dans l'indice de satisfaction, de la  $p$ -valeur plutôt que du seul rejet de l'hypothèse, est d'autant plus informatif que  $x$  est plus grand.

Les graphiques 5 à 8 reprennent respectivement les cas A à D, mais cette fois avec une fonction  $L$  de type 2,  $L(p) = \left(\frac{1}{p}\right)^l$ .

On a vu que la convergence de l'intégrale qui définit la prévision de satisfaction nécessite que  $l$  soit assez faible et en tout cas plus petit que les valeurs de  $t^2$  apparaissant dans ces exemples. Pour que les valeurs prises, pour  $x$  assez proche

de 0 (ici dans l'intervalle  $[-1, 1.5]$  dans les cas A et B et dans l'intervalle  $[-3, 4]$  dans les cas C et D) soient utilisables, nous avons été amenés à prendre  $l = \frac{1}{60}$  et avons effectué les simulations de Monte-Carlo avec  $N = 300$ . Le programme de simulation (écrit en MATLAB) est disponible en s'adressant à l'auteur de cet article.

#### 4. Conclusion

L'étude théorique et l'étude numérique comparative que nous avons menées nous amènent à conclure que les indices de satisfaction de type 1, avec le choix de la valeur 1 pour le paramètre  $l$  (autrement dit la fonction  $L(p) = 1 - p$ ) ont le double avantage d'être plus informatifs que l'indice de Grouin (type 1,  $l = 0$ ) et plus faciles de maniement et plus stables dans leurs valeurs que les indices de type 2. C'est cet indice  $I_{1,1,\alpha}$  que nous utilisons également dans [6] pour les calculs dans le modèle de lois binomiales, Gamma et de Poisson.

#### 5. Références

- [1] FOURGEAUD C. and FUCHS A. (1967), *Statistique*, Dunod, Paris.
- [2] GROUIN J.M. (1994), Procédures bayésiennes prédictives pour les essais expérimentaux, *Thèse de Doctorat* de l'Université René Descartes, Paris.
- [3] HOLST E., THYREGOD P. and WILRICH Peter-Th. (2001), On conformity testing and the use of two stage procedures, *International Statistical Review*, **69**, 3, 419-432.
- [4] LECOUTRE B., DERZKO G. and GROUIN J.M. (1995), Bayesian predictive approach for inference about proportions, *Statistics in Medicine*, **14**, 1057-1063.
- [5] MERABET H. and RAOULT J.P. (1995), Les indices de satisfaction, outils classiques et bayésiens pour la prédiction statistique, *ASU, XXVIIèmes Journées de Statistique*, Jouy-en-Josas.
- [6] MERABET H. and RAOULT J.P. (2001), Indices statistiques de satisfaction dans les essais cliniques, *Prépublications de l'Équipe d'Analyse et de Mathématiques Appliquées*, N° 5/2001.
- [7] ROBERT C. (1992), L'analyse statistique bayésienne, *Economica*, Paris.

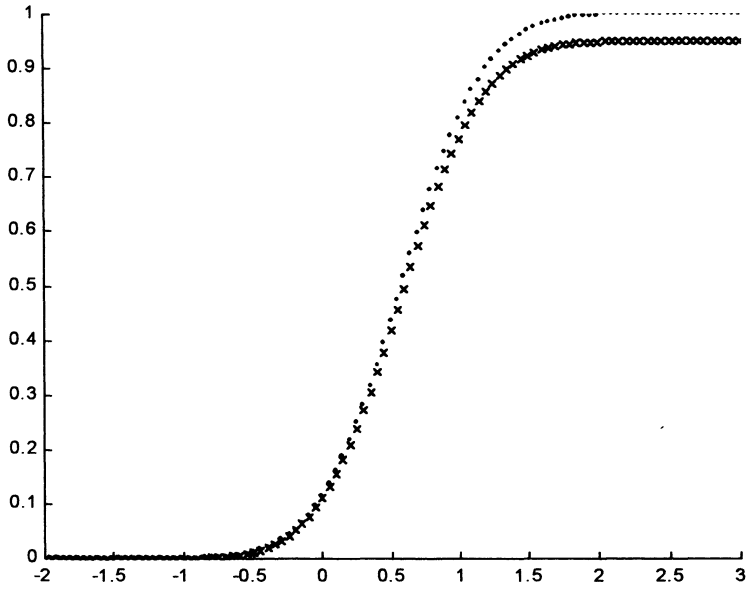


FIGURE 1

*Prévision de satisfaction (cas A);* *Prévision relative à  $S_{1,1,0.05}$  : .....;*  
*Prévision relative à  $S_{1,0,0.05}$  : × × × ×*

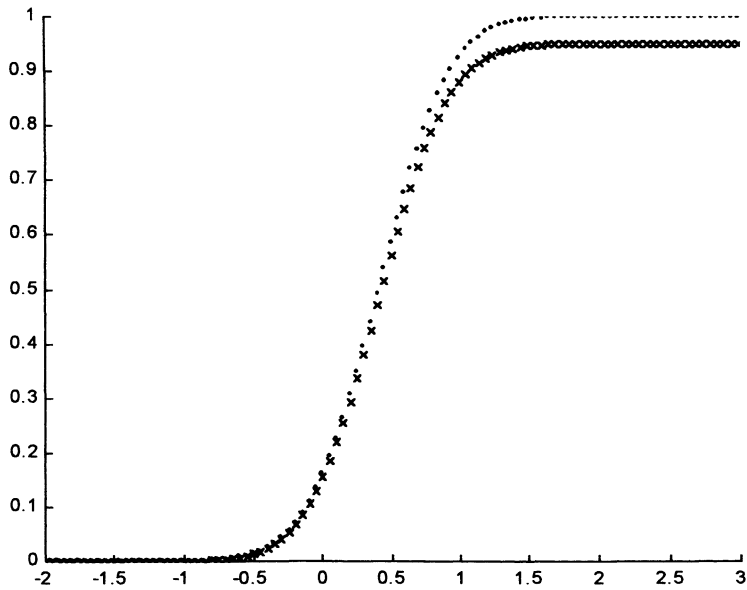


FIGURE 2

*Prévision de satisfaction (cas B);* *Prévision relative à  $S_{1,1,0.05}$  : .....;*  
*Prévision relative à  $S_{1,0,0.05}$  : × × × ×*

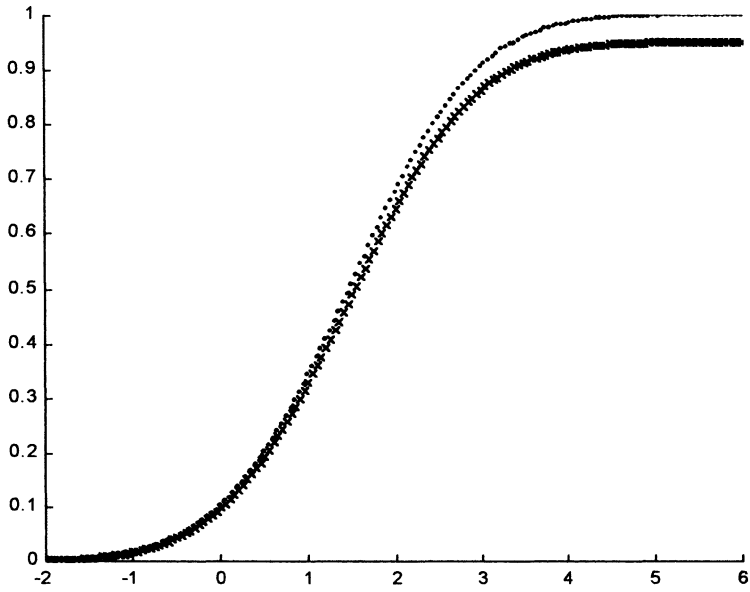


FIGURE 3

*Prévision de satisfaction (cas C); Prévision relative à  $S_{1,1,0.05}$  : .....;  
Prévision relative à  $S_{1,0,0.05}$  :  $\times \times \times \times$*

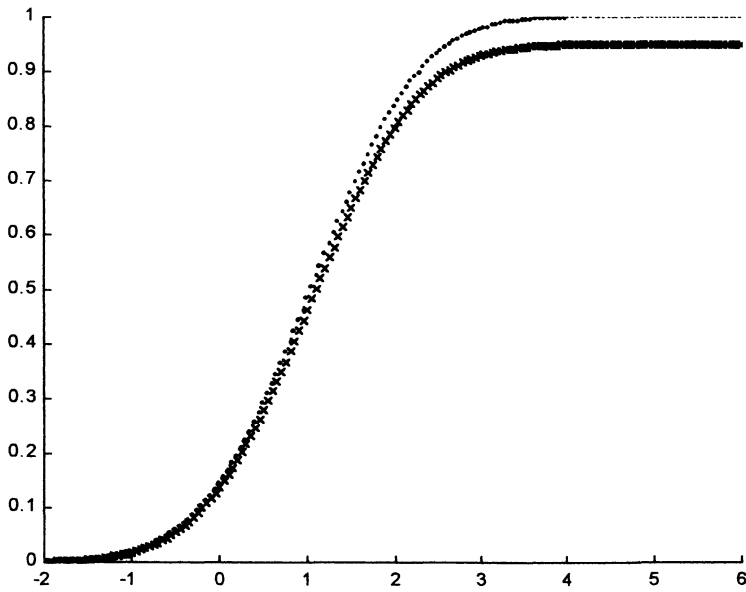


FIGURE 4

*Prévision de satisfaction (cas D); Prévision relative à  $S_{1,1,0.05}$  : .....;  
Prévision relative à  $S_{1,0,0.05}$  :  $\times \times \times \times$*

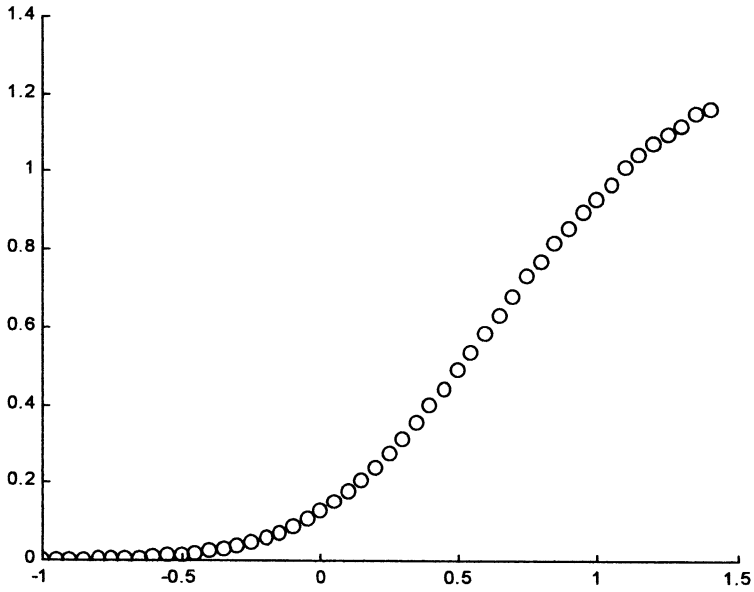


FIGURE 5

Prévision de satisfaction relative à  $S_{2, \frac{1}{60}, 0.05}$  (cas A)

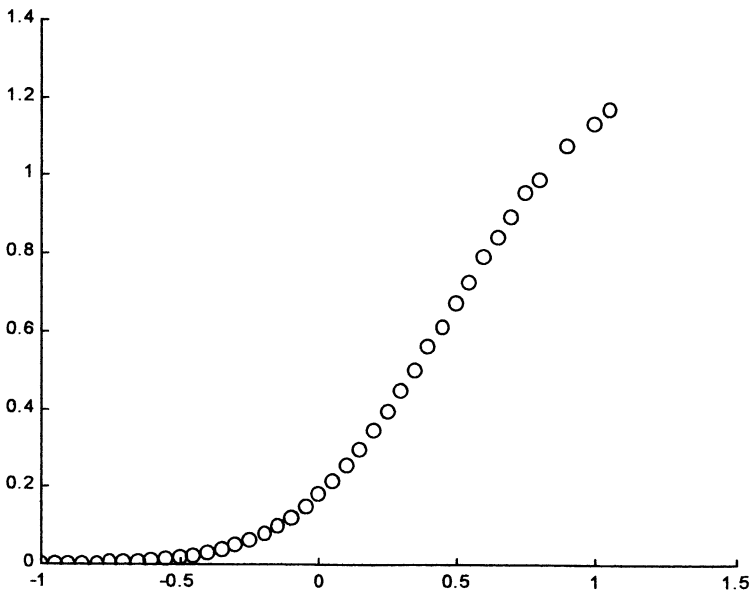


FIGURE 6

Prévision de satisfaction relative à  $S_{2, \frac{1}{60}, 0.05}$  (cas B)

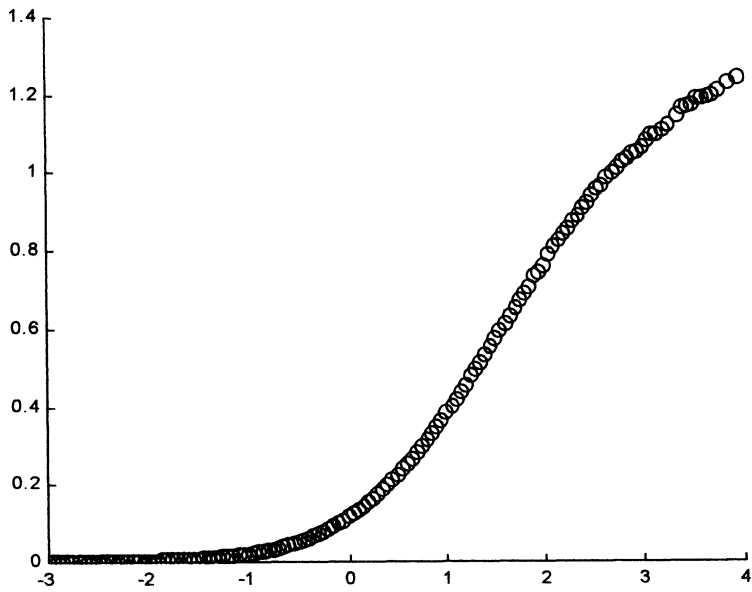


FIGURE 7

*Prévision de satisfaction relative à  $S_{2, \frac{1}{60}, 0.05}$  (cas C)*

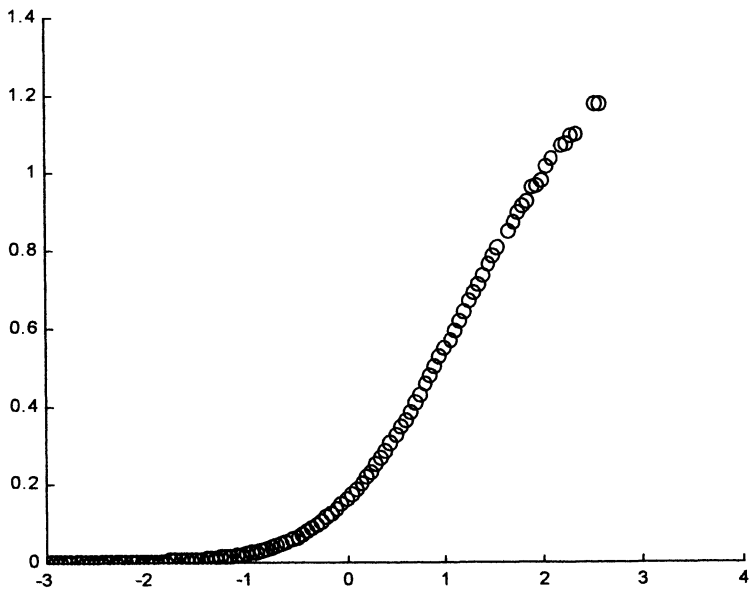


FIGURE 8

*Prévision de satisfaction relative à  $S_{2, \frac{1}{60}, 0.05}$  (cas D)*