Sciences & Technologie – N°17, Juin (2002), pp. 15-18.

INDICES STATISTIQUES DE SATISFACTION DANS LES ESSAIS CLINIQUES

Reçu le 10/03/2001 – Accepté le 16/02/2002

Résumé

Le présent article propose des procédures bayesiennes prédictives pour certains protocoles d'essais cliniques en deux étapes. Les notions d'indice de satisfaction et de prévision relatifs à un test d'hypothèse sont présentées et illustrées par des exemples où la loi a priori est une loi conjuguée.

***Mots clés**: Essais cliniques, Indices de satisfaction, Méthodes prédictives, Loi a priori conjuguée.*

Abstract

We propose Bayesian procedures adapted to some two steps protocols for clinical trials. The notions of satisfaction index and of prevision index are presented and illustrated through examples with a conjugate prior.

***Key words**: Clinical trials, Satisfaction index, Predictive methods, conjugate prior.*

H. MERABET
Département de Mathématiques
Faculté des Sciences
Université Mentouri
Constantine, Algérie

ملخص

في هذا البحث نقترح طرق "بايسية" توقعية (تنبئية) لبعض النماذج في التجارب الطبية التطبيقية على مرحلتين. مفهوم إشارة الرضا والتنبأ الإحصائي معرفة و مطبقة على أمثلة مناسبة.

***الكلمات المفتاحية**: التجارب الطبية التطبيقية، إشارة الرضاء، التنبأ الإحصائي.*

La méthodologie appliquée au contexte des essais cliniques est caractérisée par plusieurs contraintes et insatisfactions et fait l'objet d'un développement continu et approfondi. Une des raisons de cette situation tient vraisemblablement au fait que les autorités de la santé publique ont la responsabilité de l'autorisation de mise sur le marché des médicaments (A.M.M.) et jouent un rôle primordial dans l'élaboration d'une méthodologie rigoureuse des essais cliniques auprès de tous les acteurs dans ce domaine (Industries, Institut publics de recherche, Hôpitaux, Revues scientifiques). Les essais cliniques ont pour but d'évaluer l'efficacité et la tolérance d'un nouveau traitement. Ils sont caractérisés par des actions complexes qui peuvent pas être modélisées aisément et ne sont pas uniquement fonction de considérations statistiques. De plus, les statisticiens travaillant pour certains secteurs d'application, tels que les essais cliniques, se trouvent de plus en plus souvent face à des interlocuteurs qui trouvent trop sommaires les formulations tranchées qui leur ont été enseignées et qui leur sont traditionnellement fournies. Par exemple, l'usage de la théorie classique des tests ou des intervalles de confiance est souvent ressenti par le praticien comme arbitraire et mal adapté à des préoccupations de pilotage expérimental. On peut consulter à cet égard [1-3]. C'est ce contexte qui a conduit à introduire des outils statistiques dits prédictifs. On propose à cet effet au praticien l'emploi d'indices qui mesurent son degré de satisfaction face à tel ou tel résultat ou qui traduisent la prédiction qu'il effectue sur tel ou tel événement futur.

1. MODELE EXPERIMENTAL

On rappelle (voir [1]) que le contexte expérimental consiste en deux expérimentations successives, de résultats $\omega' \in \Omega'$ et $\omega'' \in \Omega''$, qui sont en général menées indépendamment. Leurs lois dépendent, dans le cadre d'un modèle bien établi, d'un paramètre $\theta \in \Theta$; seul ω'' sert à fonder la conclusion officielle de l'étude et détermine la satisfaction de l'usager qu'on notera

$\phi(\omega'')$ (et sur le choix de la quelle nous reviendrons en 2). Mais il est utile, sur la base du résultat ω' de la première phase, de prévoir ce que sera la satisfaction à l'issue de la seconde phase. Dans notre étude, comme dans [2], cette prévision est effectuée dans un contexte bayesien, c'est-à-dire fondée sur le choix d'une probabilité a priori sur Θ.
On note:

$P_{\Theta\times\Omega\times\Omega}$:	probabilité sur $\Theta\times\Omega'\times\Omega''$,
P_Θ:	probabilité a priori sur Θ,
$P_\Theta^{\omega'}$:	probabilité a posteriori sur θ, fondée sur le résultat de la première phase,
$P_{\Omega''}^{\theta}$:	loi d'échantillonnage de la deuxième phase.

Ce qui intéresse l'utilisateur potentiel de l'indicateur de satisfaction $\phi(\omega'')$, c'est, à l'issue de la première phase, la prévision de sa valeur moyenne connaissant ω'.

On définit donc l'indicateur de prédiction comme suit:

$$\pi(\omega')=\int_{\Omega''}\phi''\,P_{\Omega''}^{\omega'}(d\omega''), \tag{1}$$

où $\phi(\omega'')$ est l'indicateur de satisfaction, ou encore:

$$\pi(\omega')=\int_\Theta\left(\int_{\Omega''}\phi(\omega'')P_{\Omega''}^{\theta}\right)P_\Theta^{\omega'}(d\theta). \tag{2}$$

Considérons le cas où l'on dispose de densités relativement à des mesures μ, ν', ν'' sur Θ, Ω' et Ω'', celle de l'a priori P_Θ étant notée g, et celles des probabilités d'échantillonnage $P_{\Omega'}^{\theta}$, et $P_{\Omega''}^{\theta}$ étant notées respectivement $f'(./\theta)$ et $f''(./\theta)$.

Les expressions (1) et (2) deviennent alors:

$$\pi(\omega')=\frac{\int_{\Omega''}\phi(\omega'')\left[\int_\Theta f'(\omega'|\theta)f''(\omega''|\theta)g(\theta)\mu(d\theta)\right]\nu''(d\omega'')}{\int_\Theta f'(\omega'|\theta)g(\theta)\mu(d\theta)} \tag{3}$$

et

$$\pi(\omega')=\frac{\int_\Theta\left[\int_{\Omega''}\phi(\omega'')f''(\omega''|\theta)\nu''(d\omega'')\right]f'(\omega'|\theta)g(\theta)\mu(d\theta)}{\int_\Theta f'(\omega'|\theta)g(\theta)\mu(d\theta)}. \tag{4}$$

2. INDICE DE SATISFACTION

2.1. Cas de tests

Comme dans [5], nous nous plaçons dans le cadre où le statisticien "souhaite" observer un résultat significatif, c'est-à-dire rejeter une hypothèse nulle Θ_0. Sa "satisfaction" sera donc plus grande en cas de rejet, et même en général d'autant plus grande que l'observation qui conduit à ce rejet est plus significative.

2.1.1. Indice fruste

Etant fixé α, soit un test de niveau α défini par la région critique $\Omega_1''^{(\alpha)}$. Un premier indice de satisfaction (celui étudié par J. M. Grouin en [3]) est défini par:

$$\phi(\omega'')=1_{\Omega_1''^{(\alpha)}}(\omega''); \tag{5}$$

à ω' fixé, la prévision est alors

$$\pi(\omega')=\int_\Theta P_{\Omega''}^{\theta}\left(\Omega_1''^{(\alpha)}\right)P_\Theta^{\omega'}(d\theta), \tag{6}$$

où $P_{\Omega''}^{\theta}\left(\Omega_1''^{(\alpha)}\right)$ est la valeur en θ de la puissance du test.

Remarquons que l'a priori P_Θ peut aussi être utilisé pour fonder le test, et alors

$$\Omega_1''^{(\alpha)}=\left\{\omega''\in\Omega'';\,P_\Theta^{\omega''}(\Theta_1)\geq 1-\alpha\right\}, \tag{7}$$

mais ce n'est pas nécessairement le cas.

2.1.2. Indice amélioré

Le défaut de l'indice fruste ci-dessus est qu'il exprime une satisfaction en "tout ou rien". Il est plus intéressant de prendre en compte jusqu'à quel niveau le résultat apparaîtra toujours significatif. On utilise donc un nouvel indice de satisfaction défini par:

$$\begin{cases}\phi(\omega'')=0 & si\ \ \omega''\in\Omega_0''^{(\alpha)}\\ \quad\;\; =1-inf\left\{\beta;\,\omega\in\Omega_1''^{(\beta)}\right\} & si\ \ \omega''\in\Omega_1''^{(\alpha)}\end{cases}$$

et évidemment la prédiction est donnée par:

$$\begin{aligned}\pi(\omega')&=\int_{\Omega_1''^{(\alpha)}}\phi(\omega'')P_{\Omega''}^{\omega'}(d\omega'')\\ &=\int_\Theta\left(\int_{\Omega_1''^{(\alpha)}}\phi(\omega'')P_{\Omega''}^{\theta}(d\omega'')\right)P_\Theta^{\omega'}(d\theta').\end{aligned}$$

On remarque que $\int_{\Omega_1''^{(\alpha)}}\phi(\omega'')P_{\Omega''}^{\theta}(d\omega'')$ généralise la puissance du test dans la logique de l'indice de satisfaction proposé.

Une situation standard est celle où il existe une application Ψ ($\Theta\rightarrow R$) telle que

$$\Theta_0=\left\{\theta;\quad\Psi(\theta)\leq t^*\right\}$$

et il où existe aussi ξ ($\Omega''\rightarrow R$) et $g\ \left(]0,1[\rightarrow R\right)$ tels que

$$\Omega_0''^{(\alpha)}=\left\{\omega'');\,\xi(\omega'')\leq g(\alpha)\right\}.$$

On suppose, de plus, que la loi de ξ sous $P_{\Omega''}^{\theta}$ ne dépend que de Ψ(θ) (on la note $Q_{\Psi(\theta)}$) et que la famille des lois Q_t est stochastiquement croissante, dans le sens où ξ a de plus en plus tendance à prendre de grandes valeurs lorsque ψ(θ) devient de plus en plus grand.

Soit alors G_t la fonction de répartition de Q_t. Il est clair que:

$$\begin{cases}\phi(\omega'')=0 & si\ \ \omega''\notin\Omega_1''^{(\alpha)}\\ \quad\;\; =G_{t^*}(\omega'') & si\ \ \omega''\in\Omega_1''^{(\alpha)},\end{cases}$$

où G_{t^*} s'interprète comme la fonction de répartition "à la frontière" de ξ. La prévision est alors donnée par:

$$\pi(\omega')=\int_\Theta\left[\int_{g(\alpha)}^{\infty}G_{t^*}(dt)\,dG_{\Psi(\theta)}(x)\right]P_\Theta^{\omega'}(d\theta). \tag{8}$$

Comme pour l'indice fruste, on peut considérer le cas particulier où l'on utilise des tests de type bayesien, fondés sur le même a priori P_Θ que le calcul de la prévision de satisfaction; alors l'indice de satisfaction est:

$$\begin{cases}\tilde\phi(\omega'')=0 & si\ \ \omega''\notin\tilde\Omega_1''^{(\alpha)}\\ \quad\;\; =P_\Theta^{\omega''}(\Theta_1) & si\ \ \omega''\in\tilde\Omega_1''^{(\alpha)}\end{cases}$$

et on a bien:

$$\tilde\phi(\omega'')\geq 1-\alpha.$$

2.2. Intervalles de confiance au niveau α

Un indice de "mécontentement" est la longueur de l'intervalle de confiance. Si on fixe une longueur critique l_0 au-dessus de laquelle l'intervalle est "inadmissible", un indice de satisfaction, à valeur dans [0, 1] est:

$$\begin{cases} 0 & \text{si } l(I) \geq l_0 \\ 1 - \dfrac{l(I)}{l_0} & \text{si } l(I) < l_0 \end{cases}$$

2.3. Estimateurs

Soit à estimer un sous-paramètre $\xi = \Psi(\theta)$, dans un contexte décisionnel utilisant une fonction de coût c: $c(\xi, \xi')$ est le coût subi si on conclut ξ' alors que ξ est vrai.

Il semble difficile d'introduire un indice de satisfaction, lié à un estimateur $\hat{\xi}$, en dehors du contexte bayesien. On peut proposer l'espérance a posteriori du coût si on a observé ω'', c'est-à-dire

$$\phi(\omega'') = \int_\Theta c\left(\Psi(\theta), \hat{\xi}(\omega'')\right) P_\Theta^{\omega''}(d\theta) \qquad (9)$$

et alors la prévision est

$$\pi(\omega') = \int_{\Omega''} \left[\int_\Theta c\left(\Psi(\theta), \hat{\xi}(\omega'')\right) P_{\Omega''}^{\omega''}(d\theta)\right] P_{\Omega''}^{\omega'}(d\omega''). \qquad (10)$$

3. APPLICATIONS

On se propose de calculer explicitement ou numériquement selon le cas l'indice et la prévision de satisfaction dans plusieurs modèles exponentiels lorsque la loi a priori du paramètre inconnu θ est une loi a priori conjuguée (voir [4]), et dans le cas d'un test au seuil α où l'hypothèse nulle est du type $\theta \leq \theta_0$.

On effectue des observations indépendantes et de même loi. Le premier résultat est une suite $X_1^{'},\ldots, X_k^{'}$ de k observations et le second est aussi une suite $X_1^{''},\ldots X_n^{''}$ de n observations. Tous les calculs sont fondés, pour des raisons d'exhaustivité, sur $\omega' = \sum_{i=1}^k X_i^{'}$ et $\omega'' = \sum_{i=1}^n X_i^{''}$.

Dans tout ce qui suit, on a les notations simplifiées suivantes:

g: densité de loi a priori de θ,
l: densité de la loi du couple (ω', ω''),
v: densité de la loi prédictive conditionnelle de ω'' sachant ω'.
F: fonction de répartition de la loi de ω'' pour la valeur θ_0 du paramètre.

3.1. Cas du modèle Gaussien

Ce modèle a été étudié en détail et fait l'objet d'une autre publication ([5]), où divers indices de satisfaction sont envisagés.

Des exemples numériques y sont présentés.

3.2. Cas de la loi Gamma

On suppose que toutes les v.a. indépendantes $X_i^{'}(1 \leq i \leq k)$ et $X_j^{''}(1 \leq j \leq n)$ suivent une loi Gamma $G(p, \theta)$ où p est connu et θ inconnu. Alors ω' est de loi Gamma $G(kp,\theta)$ et ω'' est de loi Gamma $G(np, \theta)$.

On pose $kp = K$ et $np = N$. Si θ suit la loi a priori Gamma $G(a, b)$, alors (voir [4]) la loi a posteriori de θ sachant ω' est une loi Gamma $G(a+K, b+\omega')$. L'indice de satisfaction est

$$\begin{cases} \phi(\omega'') = 0 & si\ \ \omega'' \leq q_0 \\ \quad = F(\omega'') = \displaystyle\int_0^{\omega''} \frac{\theta_0^N}{\Gamma(N)} e^{-\theta_0 t} t^{N-1} dt & sinon, \end{cases}$$

où q_0 est défini par: $F(q_0) = 1\alpha$.

Alors

$$l(\omega', \omega'') = \frac{\Gamma(K+N+a)}{(\omega'+\omega''+b)^{K+N+a}} \cdot \frac{\omega'^{K-1}.\omega''^{N-1}.b^a}{\Gamma(K).\Gamma(N).\Gamma(a)} \qquad (11)$$

et

$$v(\omega''|\omega') = \frac{\omega''^{N-1}}{(\omega'+\omega''+b)^{K+N+a}} \cdot \frac{1}{\displaystyle\int_0^\infty \frac{\omega''^{N-1}}{(\omega'+\omega''+b)^{K+N+a}} d\omega''}. \qquad (12)$$

Finalement la prédiction de satisfaction est donnée par:

$$\pi(\omega') = \int_{q_0}^{+\infty} \left[\int_0^{\omega''} \frac{\theta_0^K}{\Gamma(N)} e^{-\theta_0 u} u^{N-1} du\right] v(\omega''|\omega') d\omega'' \qquad (13)$$

qui peut être évaluée numériquement.

3.3. Cas de la loi binomiale

On suppose que toutes les v.a. indépendantes $X_1^{'},\ldots, X_k^{'}$, $X_1^{''},\ldots X_n^{''}$ suivent la loi binomiale $B(s, \theta)$, où s est connu et θ inconnu. Alors $\omega' = \sum_{i=1}^k X_i^{'}$ est de loi binomiale $B(ks, \theta)$ et $\omega'' = \sum_{i=1}^n X_i^{''}$ est de loi binomiale $B(ns, \theta)$.

On pose $ks = K$ et $ns = N$.

On suppose que θ a pour loi a priori une loi Beta de paramètres a et b.

Par conséquent (voir [4]), θ, sachant ω', est de loi Beta de paramètres $a + \omega'$ et $K + b - \omega'$.

Donc, en posant

$$\beta(x, y) = \frac{\Gamma(x).\Gamma(y)}{\Gamma(x+y)},$$

il vient, pour tout (ω', ω'') $\in \{0, 1,\ldots, K\} \times \{0, 1, \ldots, N\}$,

$$l(\omega', \omega'') = \frac{C_K^{\omega'}.C_N^{\omega''}}{\beta(a,b)}.\beta(a+\omega'+\omega'', b+K+N-\omega'-\omega'') \quad (14)$$

et

$$v(\omega''|\omega') = \frac{C_N^{\omega''}.\beta(a+\omega'+\omega'', b+K+N-\omega'-\omega'')}{\sum_{\omega''=0}^N C_N^{\omega''} \beta(a+\omega'+\omega'', b+K+N-\omega'-\omega'')}. \quad (15)$$

L'indice de satisfaction est donnée par:

$$\begin{cases} \phi(\omega'') = 0 & si\ \omega'' < q_0 \\ \quad = \sum_{t=0}^{\omega''-1} C_N^t \theta_0^t (1-\theta_0)^{N-t} & si\ \omega'' \geq q_0 \end{cases},$$

où

$$q_0 = inf\left\{ t;\ \sum_{t=q_0}^{N} C_N^t \theta_0^t (1-\theta_0)^{N-t} \leq \alpha \right\}.$$

Finalement, la prévision de satisfaction est

$$\pi(\omega') = \sum_{\omega''=q}^{N} \sum_{t=0}^{\omega''-1} C_N^t . \theta_0^t (1-\theta_0)^{N-t} \times$$
$$\times \frac{C_N^{\omega''} . \beta(a+\omega'+\omega'', b+K+N-\omega'-\omega'')}{\sum_{\omega''=0}^{N} C_N^{\omega''} \beta(a+\omega'+\omega'', b+K+N-\omega'-\omega'')} \quad (16)$$

3.4. Cas d'un échantillonnage de Poisson

On suppose maintenant que les v.a. indépendantes $X'_i (1 \leq i \leq k)$ et $X''_j (1 \leq j \leq n)$ suivent toutes la loi de Poisson $P(\theta)$, où θ est inconnu. Alors, $\omega' = \sum_{i=1}^{k} X'_i$ est de loi de Poisson $P(k\theta)$ et $\omega'' = \sum_{i=1}^{n} X''_i$ est de loi de Poisson $P(n\theta)$.

Si θ suit a priori la loi Gamma de paramètres a et b, alors, a posteriori, θ, sachant ω', suit la loi Gamma de paramètres $a+\omega'$ et $b+k$. L'indice de satisfaction s'écrit alors:

$$\begin{cases} \phi''^{(\alpha)}(\omega'') = 0 & si\ \omega'' < q_0 \\ \quad = \sum_{s=0}^{\omega''-1} e^{-n\theta_0} \frac{(n\theta_0)^s}{s!} & si\ \omega'' \geq q_0 \end{cases}$$

où

$$q_0 = inf\left\{ s;\ \sum_{0}^{q_0-1} e^{-n\theta_0} \frac{(n\theta_0)^s}{s!} \geq 1-\alpha \right\},$$

et la prévision de satisfaction est donnée par:

$$\pi(\omega') = \sum_{\omega''=q}^{\infty} \sum_{s=0}^{\omega''-1} e^{-N\theta_0} \frac{(N\theta_0)^s}{s!} . \frac{N^{\omega''}}{\omega''!} \times$$
$$\times \frac{\Gamma(a+\omega'+\omega'')}{(b+K+N)^{a+\omega'+\omega''}} . \frac{1}{\sum_{\omega''=0}^{\infty} \frac{N\omega''}{\omega''!} \frac{\Gamma(a+\omega'+\omega'')}{(b+K+N)^{a+\omega'+\omega''}}} \quad (17)$$

REFERENCES

[1]- Merabet H., Raoult J.P., "Les indices de satisfaction, outils classiques et bayesiens pour la prédiction statistique", *ASU*. XXVIIèmes Journées de Statistique, Jouyen, Josas (1995).

[2]- Grieve A.P., "Predictive probability in clinical trials", *Biometrics*, 41, (1992), pp. 979-990.

[3]- Grouin J.M., "Procédures bayesiennes prédictives pour les essais expérimentaux", thèse de Doctorat de l'Université René Descartes, Paris, (1994).

[4]- Robert C., "L'analyse bayesienne", Economica, Paris, (1992).

[5]- Merabet H., Raoult J.P., "Méthodologie des essais préliminaires. Calcul de prévision de satisfaction dans le cas Gaussien", Prépublication du Laboratoire d'Analyse et de Mathématiques Appliquées, Université de Marne la Vallée, (2001). □