Un nouveau regard sur le "maximum de vraisemblance"

par Jörg W. Müller

Bureau International des Poids et Mesures, F-92310 Sèvres

Résumé - On montre que le maximum de vraisemblance L_{max} pour n mesures doit se situer aux environs d'une valeur numérique Λ^n qui peut être calculée pour une loi donnée. Dans le cas d'une distribution normale, des expressions explicites sont indiquées pour l'espérance et l'écart-type de Λ.

Dans cette méthode d'estimation très utilisée, on obtient la "meilleure" valeur d'un paramètre θ, qui détermine une loi de distribution f supposée autrement connue, en utilisant les observations $\vec{x} = (x_1, x_2, \ldots, x_n)$ sous la forme

$$L(\vec{x}\,|\,\theta) = \prod_{j=1}^{n} f(x_j\,|\,\theta).$$

La valeur maximale de L, considérée comme fonction de θ, est déterminée à l'aide de la condition

$$\frac{\partial L}{\partial \theta} = 0.$$

Admettons, par exemple, que la variable mesurée suit une loi de Poisson, donc

$$f_\mu(x_j) = \frac{\mu^{x_j}}{x_j!}\, e^{-\mu}, \qquad \text{avec } x_j = 0, 1, 2, \ldots$$

$$\text{et } \mu \equiv \theta > 0.$$

Cela nous amène à

$$L = C\, e^{-n\mu}\, \mu^{\sum x_j},$$

où C est une constante qui ne dépend pas de μ. L atteint son maximum si μ prend la valeur

$$\tilde{\mu} = \frac{1}{n} \sum x_j = \bar{x}.$$

Il s'ensuit, par exemple, que pour les deux séries hypothétiques à cinq observations

$$\vec{x}_A = (1, 0, 3, 0, 2) \qquad \text{et} \qquad \vec{x}_B = (0, 0, 0, 6, 0),$$

on serait amené sans distinction à l'estimation $\tilde{\mu}_A = \tilde{\mu}_B = 6/5 = 1{,}2$. Ce résultat ne tient visiblement pas compte de notre "sentiment" que la série A de résultats semble pourtant plus probable que la série B.

Dans une tentative de remédier à cette situation jugée peu satisfaisante, nous proposons d'examiner de plus près la constante C, jusuqu'ici volontairement négligée.

Pour une distribution de Poisson on a

$$L_{max} \equiv \Lambda^n = \frac{(\bar{x}/e)^{n\bar{x}}}{\prod_{j=1}^{n} x_j!} .$$

Comme on le voit, cette valeur ne dépend pas seulement de $\bar{x}$, mais aussi de la répartition des valeurs x_j mesurées. Or, est-elle utilisable? Cela est le cas seulement si l'on peut connaître sa distribution théorique, à laquelle la valeur empirique L_{max} (ou Λ) pourra alors être comparée.

On peut montrer que les deux premiers moments de $\lambda \equiv \ln \Lambda$ sont donnés de façon générale par

$$m_1(\lambda) = \sum_{j=0}^{\infty} f_\mu(j) \ln f_\mu(j) \qquad \text{et}$$

$$m_2(\lambda) = \sum_{j=0}^{\infty} f_\mu(j) \left[\ln f_\mu(j)\right]^2 ,$$

d'où l'on déduit espérance et variance de Λ. Malheureusement, pour une loi de Poisson ces sommations ne s'effectuent que numériquement. Dans notre cas, avec $\mu = 1{,}2$ et $n = 5$, on obtient ainsi $\Lambda = 0{,}27 \pm 0{,}07$. Ce résultat est à comparer avec les valeurs empiriques des deux séries, c'est-à-dire $\Lambda_A = 0{,}23$ et $\Lambda_B = 0{,}10$.

Il se révèle que la série A donne une valeur de Λ qui est en bon accord avec l'espérance théorique, mais que ce n'est point le cas pour la série B qui nous a déjà paru suspecte. Or, ce doute peut maintenant être justifié et exprimé de manière quantitative.

La distribution normale est d'une importance particulière, car c'est vers elle que tendra la vraisemblance $L(\theta)$ pour un nombre n d'observations suffisamment élevé. Ceci est vrai pour presque n'importe quelle fonction $f(x|\theta)$. Il est donc très heureux que l'évaluation de Λ puisse se faire de façon rigoureuse pour cette loi. Ainsi, pour

$$f(x|\mu,s) = \frac{1}{s\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{x-\mu}{s}\right)^2\right],$$

où s est supposé connu, on peut montrer que

$$E(\Lambda) = \frac{1}{s\sqrt{2\pi}} \exp\left[-\frac{n-1}{2n}\right] \cong \frac{0{,}242\ 0}{s} \exp\left(\frac{1}{2n}\right)$$

et

$$\sigma(\Lambda) = \frac{1}{n}\sqrt{\frac{n-1}{2}}\, E(\Lambda) \cong \frac{0{,}171\ 1}{s}\,\frac{\sqrt{n-1}}{n} \exp\left(\frac{1}{2n}\right).$$

En particulier, pour une seule mesure ($n = 1$) on obtient donc $\Lambda = (2\pi\, s^2)^{-1/2}$ (sans incertitude), comme il fallait s'y attendre.

Le résultat général est très utile, car il peut nous procurer des valeurs approximatives pour des cas où l'évaluation exacte serait trop pénible ou même impossible. Comme illustration, prenons l'exemple précédent d'une loi de Poisson avec $\mu = 1{,}2$ et $n = 5$. Dans ce cas, puisque $s = \sqrt{\mu}$, l'approximation normale aurait donné

$$E(\Lambda) \cong 0{,}24 \qquad \text{et} \qquad \sigma(\Lambda) \cong 0{,}07.$$

La comparaison avec le résultat exact $\Lambda = 0{,}27 \pm 0{,}07$ indiqué plus haut montre que, même pour cette valeur de μ peu élevée, l'approximation est déjà utilisable bien que les lois limites ne le garantissent que pour $\mu \to \infty$.

Il va sans dire que ces considérations n'ont pas d'influence sur la valeur numérique d'un paramètre θ estimé par la méthode du maximum de vraisemblance, mais elles nous permettent de savoir si cette démarche est bien utile ou si, le cas échéant, il ne vaudrait pas mieux y renoncer quand la valeur expérimentale de Λ nous indique que la cohérence interne des résultats est insuffisante.

(Novembre 1981)