

# Intervalles de confiance et de crédibilité pour le rapport de taux d'évènements rares

Bruno Lecoutre, Gérard Derzko

► **To cite this version:**

Bruno Lecoutre, Gérard Derzko. Intervalles de confiance et de crédibilité pour le rapport de taux d'évènements rares. 41èmes Journées de Statistique, SFdS, Bordeaux, 2009, Bordeaux, France, France. 2009. <inria-00386595>

**HAL Id: inria-00386595**

**<https://hal.inria.fr/inria-00386595>**

Submitted on 22 May 2009

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# INTERVALLES DE CONFIANCE ET DE CRÉDIBILITÉ POUR LE RAPPORT DE TAUX D'ÉVÈNEMENTS RARES

Bruno Lecoutre<sup>1</sup> & Gérard Derzko<sup>2</sup>

<sup>1</sup> *ERIS, UPRESA 6085, Laboratoire de Mathématiques Raphaël Salem  
C.N.R.S. et Université de Rouen  
Avenue de l'Université, BP 12, 76801 Saint-Etienne-du-Rouvray*

<sup>2</sup> *Sanofi-aventis Recherche & Développement  
371, rue du Professeur Joseph Blayac  
34184 Montpellier*

**Résumé :** *Nous comparons une sélection de méthodes fréquentistes et l'approche bayésienne de Jeffreys pour l'estimation de limites de confiance d'un rapport de paramètres de Poisson sur la base du critère fréquentiste du taux de couverture. Nous concluons que, pour des valeurs de paramètres de Poisson très petites, certaines méthodes fréquentistes deviennent peu satisfaisantes, mais l'approche bayésienne garde d'excellentes performances, similaires à celles de la meilleure des méthodes fréquentistes.*

**Mots-clef :** Modèle de Poisson, intervalle de confiance, intervalle de crédibilité, taux de couverture, inférence bayésienne.

**Abstract :** *A selection of frequentist methods and the Bayesian approach are compared for the estimation of confidence limits of the rates ratio of two Poisson parameters. Performance is evaluated through the frequentist criterion of coverage probabilities. The conclusion is that, when the Poisson parameters become very small, some of the selected frequentist methods perform very poorly, while the Bayesian approach keeps performing as well as the best of the frequentist method.*

**Keywords :** Poisson Model, confidence interval, credible interval, coverage probability, Bayesian inference.

Des effets indésirables rares mais sévères peuvent parfois être enregistrés dans des études cliniques comparatives de phase 3. Il convient alors d'estimer l'incertitude sur le rapport des taux de tels effets obtenus avec le traitement actif et un *placebo*, d'une façon éventuellement stratifiée. L'intérêt se porte principalement sur l'estimation d'une limite supérieure pour ce rapport. Dans le cadre d'un modèle Poissonien pour les évènements, un grand nombre de procédures fréquentistes (Price et Bonnett, 2000) fournissent des intervalles de confiance uni- ou bi-latéraux généralement associés à des tests et obtenus au moyen de différents types d'approximations. L'approche bayésienne fournit par ailleurs des intervalles de crédibilité. Nous comparons ici les performances fréquentistes de certaines de ces procédures.

Formellement, nous considérerons deux variables aléatoires  $X_t$  ( $t = 1, 2$ ) ayant les distributions d'échantillonnage indépendantes

$$X_t | \lambda_1, \lambda_2 \sim \text{Poi}(n_t \lambda_t)$$

$$Pr(X_t = x_t | \lambda_1, \lambda_2) = \frac{(n_t \lambda_t)^{x_t}}{(x_t - 1)!} e^{-x_t \lambda_t} \quad t = 1, 2$$

Nous nous intéresserons aux intervalles de confiance pour le rapport  $\tau = \lambda_1/\lambda_2$  et nous examinerons les propriétés de couverture fréquentiste de ces intervalles. Soit

$$[\ell_{inf}(X_1, X_2), \ell_{sup}(X_1, X_2)]$$

une procédure d'intervalle de confiance  $100(1 - \alpha)\%$ . Pour des valeurs fixées  $\lambda_1$  et  $\lambda_2$  des paramètres, un intervalle particulier  $[\ell_{inf}(x_1, x_2), \ell_{sup}(x_1, x_2)]$ , associé aux données  $(x_1, x_2)$  ne contient pas  $\tau$  si  $\ell_{inf}(x_1, x_2) > \tau$  (erreur inférieure) ou  $\ell_{sup}(x_1, x_2) < \tau$  (erreur supérieure). Les probabilités d'échantillonnage d'une erreur inférieure et d'une erreur supérieure sont respectivement :

$$Q_{inf} = \sum_{x_1=0}^{\infty} \sum_{x_2=0}^{\infty} Pr(X_1 = x_1 | \lambda_1) Pr(X_2 = x_2 | \lambda_2) \mathbb{1}_{[\tau, \infty]}(\ell_{inf}(x_1, x_2))$$

$$Q_{sup} = \sum_{x_1=0}^{\infty} \sum_{x_2=0}^{\infty} Pr(X_1 = x_1 | \lambda_1) Pr(X_2 = x_2 | \lambda_2) \mathbb{1}_{[0, \tau]}(\ell_{sup}(x_1, x_2))$$

le taux de couverture bilatérale étant égal à  $1 - Q_{inf} - Q_{sup}$ . L'intervalle est dit *exact* si cette quantité est supérieure ou égale à  $1 - \alpha$  pour toute les valeurs des paramètres. Mais une procédure exacte (comme l'intervalle de Clopper-Pearson pour une proportion Binomiale) est souvent très conservatrice. Dans la pratique expérimentale, on préférera donc généralement une procédure approchée pour obtenir un taux le plus proche possible de  $1 - \alpha$  ("*approximate is better than exact*" : Agresti et Coul, 1998). Par ailleurs, on souhaitera également le plus souvent un intervalle ayant de bonne propriétés de couverture, non seulement bilatérale, mais aussi unilatérale ( $Q_{inf}$  et  $Q_{sup}$  proches de  $\alpha/2$ ).

En pratique, le calcul des taux d'erreurs sera rendu possible (il n'y a pas besoin de recourir à des simulations) par le fait que les sommes infinies peuvent être approximées par des sommes finies sur des ensembles  $\mathbf{X}_1$  et  $\mathbf{X}_2$ . On peut facilement contrôler l'erreur de troncature, puisque celle-ci est inférieure à

$$\sum_{x_1 \notin \mathbf{X}_1} Pr(X_1 = x_1 | \lambda_1) + \sum_{x_2 \notin \mathbf{X}_2} Pr(X_2 = x_2 | \lambda_2)$$

Dans le cas d'événements rares, même pour des  $n_t$  très élevés, les ensembles  $\mathbf{X}_t$  contiendront un nombre limité d'éléments.

## APPROCHE CONDITIONNELLE

C'est l'approche la plus fréquemment utilisée. Elle est basée sur la propriété suivante. Si  $S = X_1 + X_2$ , alors la distribution d'échantillonnage de  $X_1$  conditionnellement à la valeur observée  $s = x_1 + x_2$  est une distribution Binomiale qui ne dépend des paramètres  $\lambda_1$  et  $\lambda_2$  que par l'intermédiaire de  $\pi = n_1 \lambda_1 / (n_1 \lambda_1 + n_2 \lambda_2) = \tau / ((n_2/n_1) + \tau)$

$$X_1 | S = s, \lambda_1, \lambda_2 \sim X_1 | S = s, \tau \sim \text{Bin}\left(\frac{\tau}{\frac{n_2}{n_1} + \tau}, s\right)$$

$$Pr(X_1 = x_1 | S = s, \tau) = \frac{s!}{x_1! (s-x_1)!} \left(\frac{\tau}{\frac{n_2}{n_1} + \tau}\right)^{x_1} \left(\frac{\frac{n_2}{n_1}}{\frac{n_2}{n_1} + \tau}\right)^{s-x_1}$$

On peut alors appliquer au paramètre  $\pi$  toutes les procédures qui ont été proposées pour la construction d'un intervalle de confiance  $100(1 - \alpha)\%$  pour une proportion Binomiale, et en déduire un intervalle pour  $\tau$ . On trouvera une revue récente de celles-ci

dans Pires et Amado (2008). Si l'on s'en tient aux critères de choix d'une procédure approchée plutôt qu'exacte, et ayant de bonnes propriétés de couverture unilatérale, des solutions fréquentistes performantes pour une proportion Binomiale sont la procédure du "second-ordre corrigé" (Cai, 2005) et la procédure basée sur le test du "mid-p".

Cette dernière est construite à partir du test de Clopper-Pearson. Pour ce dernier, l'hypothèse nulle  $\tau = \tau_0$  est testée contre l'hypothèse alternative  $\tau > \tau_0$  (par exemple) en déclarant le résultat significatif au seuil  $\alpha$  si

$$\bar{p}_{exc}^{\tau_0} = \sum_{j=x_1+1}^s P(X_1 = j \mid S = s, \tau_0) \leq \alpha$$

Dans cette solution la sommation porte sur toutes les valeurs de  $X_1$  qui sont strictement plus grandes que la valeur observée  $x_1$  (d'où l'indice *exc* pour test *excluant*). Comme de test peut être très conservateur, une autre solution est d'inclure  $x_1$  dans la sommation. Soit

$$\bar{p}_{inc}^{\tau_0} = \sum_{j=x_1}^s P(X_1 = j \mid S = s, \tau_0) \leq \alpha$$

(*inc* pour test *incluant*) la probabilité correspondante. Ce test est libéral. Une solution typique pour résoudre le problème consiste à considérer la moyenne des probabilités excluanes et incluanes ("mid-p-value", voir par exemple Berry et Armitage, 1995)

$$\bar{p}_{mid}^{\tau_0} = \frac{1}{2} (\bar{p}_{exc}^{\tau_0} + \bar{p}_{inc}^{\tau_0})$$

Pour chaque test, la limite inférieure de l'intervalle de confiance  $100(1-\alpha)\%$  pour  $\tau$  est obtenue en résolvant  $\underline{p}^{\tau_0} = \alpha/2$ . De même, on définit le test contre l'hypothèse alternative  $\tau < \tau_0$  (d'où les seuils  $\underline{p}_{exc}^{\tau_0}$ ,  $\underline{p}_{inc}^{\tau_0}$  et  $\underline{p}_{mid}^{\tau_0}$ ), et on en déduit la limite supérieure.

Mais, appliquées au rapport  $\tau$ , les procédures pour une proportion Binomiale sont des procédures conditionnelles. Même si la distribution utilisée pour les construire ne dépend que de ce paramètre, la probabilité d'échantillonnage qui donne les taux de couverture fréquentiste dépend à la fois de  $\lambda_1$  et de  $\lambda_2$ , donc à la fois de  $\tau$  et d'un paramètre *parasite*. En conséquence ces procédures peuvent être beaucoup moins satisfaisantes que pour une proportion Binomiale, en particulier dans le cas d'événements rares.

## AUTRES APPROCHES

Nous retiendrons ici, d'une part l'approche de Sahai et Kurshid (1993), et d'autre part l'approche bayésienne.

*L'intervalle de Sahai et Khurshid*

Les limites inférieure et supérieure sont obtenues à partir des percentiles de la distribution Normale standard. On a respectivement (avec  $P(Z > z_{\alpha/2}) = \alpha/2$ ) :

$$\frac{n_2}{n_1} \left( \frac{\sqrt{(x_1+0.5)(x_2+0.5)} - 0.5z_{\alpha/2} \sqrt{x_1+x_2+1-0.25z_{\alpha/2}^2}}{x_2+0.5-0.25z_{\alpha/2}^2} \right)^2$$

$$\frac{n_2}{n_1} \left( \frac{\sqrt{(x_1+0.5)(x_2+0.5)} + 0.5z_{\alpha/2} \sqrt{x_1+x_2+1-0.25z_{\alpha/2}^2}}{x_2+0.5-0.25z_{\alpha/2}^2} \right)^2$$

*Procédures bayésiennes*

Une solution simple et usuelle postule deux distributions *a priori* indépendantes Gamma (conjuguées) pour  $\lambda_1$  et  $\lambda_2$ , de paramètres  $x_t^0$  et d'échelle  $1/n_t^0$

$$\lambda_1 \sim \frac{1}{n_1^0} \text{Gamma}(x_1^0) \text{ et } \lambda_2 \sim \frac{1}{n_2^0} \text{Gamma}(x_2^0)$$

Les distributions *a posteriori* sont encore deux distributions Gamma indépendantes

$$\lambda_1 \mid \text{données} \sim \frac{1}{n_1+n_1^0} \text{Gamma}(x_1+x_1^0) \text{ et } \lambda_2 \mid \text{données} \sim \frac{1}{n_2+n_2^0} \text{Gamma}(x_2+x_2^0)$$

On en déduit les distributions marginales des paramètres  $\pi$  et  $\tau$ . Dans le cas particulier  $n_1^0 = n_2^0 = 0$ , la première est simplement la distribution Bêta

$$\pi \mid \text{données} \sim \text{Bêta}(x_1+x_1^0, x_2+x_2^0)$$

tandis que la distribution de  $\tau$  est une distribution Bêta type II, d'échelle  $n_2/n_1$

$$\tau \mid \text{données} \sim \frac{n_2}{n_1} \text{Bêta}_{II}(x_1+x_1^0, x_2+x_2^0)$$

Il y a une correspondance entre les seuils des tests conditionnels excluant (Clopper-Pearson) et incluant de l'hypothèse nulle  $\tau = \tau_0$  et les probabilités bayésiennes *a posteriori* que  $\tau < \tau_0$  associée à des choix particuliers de la distribution *a priori*. On a les égalités suivantes.

$$\bar{p}_{exc}^{\tau_0} = 1 - \underline{p}_{inc}^{\tau_0} = P(\tau < \tau_0 \mid \text{données})$$

$$\text{pour la distribution } a \text{ priori } (x_1^0 = 1, n_1^0 = 0; x_2^0 = 0, n_2^0 = 0)$$

$$\bar{p}_{inc}^{\tau_0} = 1 - \underline{p}_{exc}^{\tau_0} = P(\tau < \tau_0 \mid \text{données})$$

$$\text{pour la distribution } a \text{ priori } (x_1^0 = 0, n_1^0 = 0; x_2^0 = 1, n_2^0 = 0)$$

En effet, pour  $n_1^0 = n_2^0 = 0$ ,  $P(\tau < \tau_0 \mid \text{données})$  est la fonction Bêta incomplète  $I_{\pi_0}(x_1+x_1^0, x_2+x_2^0)$  ( $\pi_0 = \tau_0 / ((n_2/n_1) + \tau_0)$ ). Donc, si  $x_1^0$  et  $x_2^0$  sont des entiers, c'est la probabilité qu'une variable ayant la distribution Binomiale de paramètres  $x_1+x_2+x_1^0+x_2^0$  et  $\pi_0$  soit supérieure ou égale à  $x_1+x_1^0$  (Johnson, Kotz et Kemp, 1993, page 117). On en déduit les égalités précédentes.

En conséquence, pour une distribution *a priori* objective (Berger, 2004; Lecoutre, 2008), une solution raisonnable simple est la distribution intermédiaire

$$(x_1^0 = 1/2, n_1^0 = 0; x_2^0 = 1/2, n_2^0 = 0)$$

Une justification formelle est que cette distribution est l'*a priori* de Jeffreys.

## COMPARAISON DES PROCEDURES : EXEMPLE

La Figure 1 donne les probabilités fréquentistes d'erreur inférieure et supérieure, pour  $n_1 = n_2 = 1000$  et  $\tau$  variant de 0.10 à 4. Deux cas sont considérés.

Dans le premier cas le paramètre (parasite)  $\lambda_2$  vaut 0.04. Les quatre procédures considérées sont très performantes.

Dans le second cas, qui correspond à des proportions  $\lambda_1$  et  $\lambda_2$  plus extrêmes, le paramètre (parasite)  $\lambda_2$  vaut 0.004. La procédure du second-ordre corrigé n'est pas satisfaisante : elle donne un taux de couverture bilatérale sensiblement inférieur à  $1 - \alpha$ . La procédure du *mid-p* a encore un taux de couverture bilatérale proche de  $1 - \alpha$ , mais les taux d'erreur unilatéraux sont moins satisfaisants. Les procédures bayésienne et de Sahai et Kurshid restent au contraire très performantes.

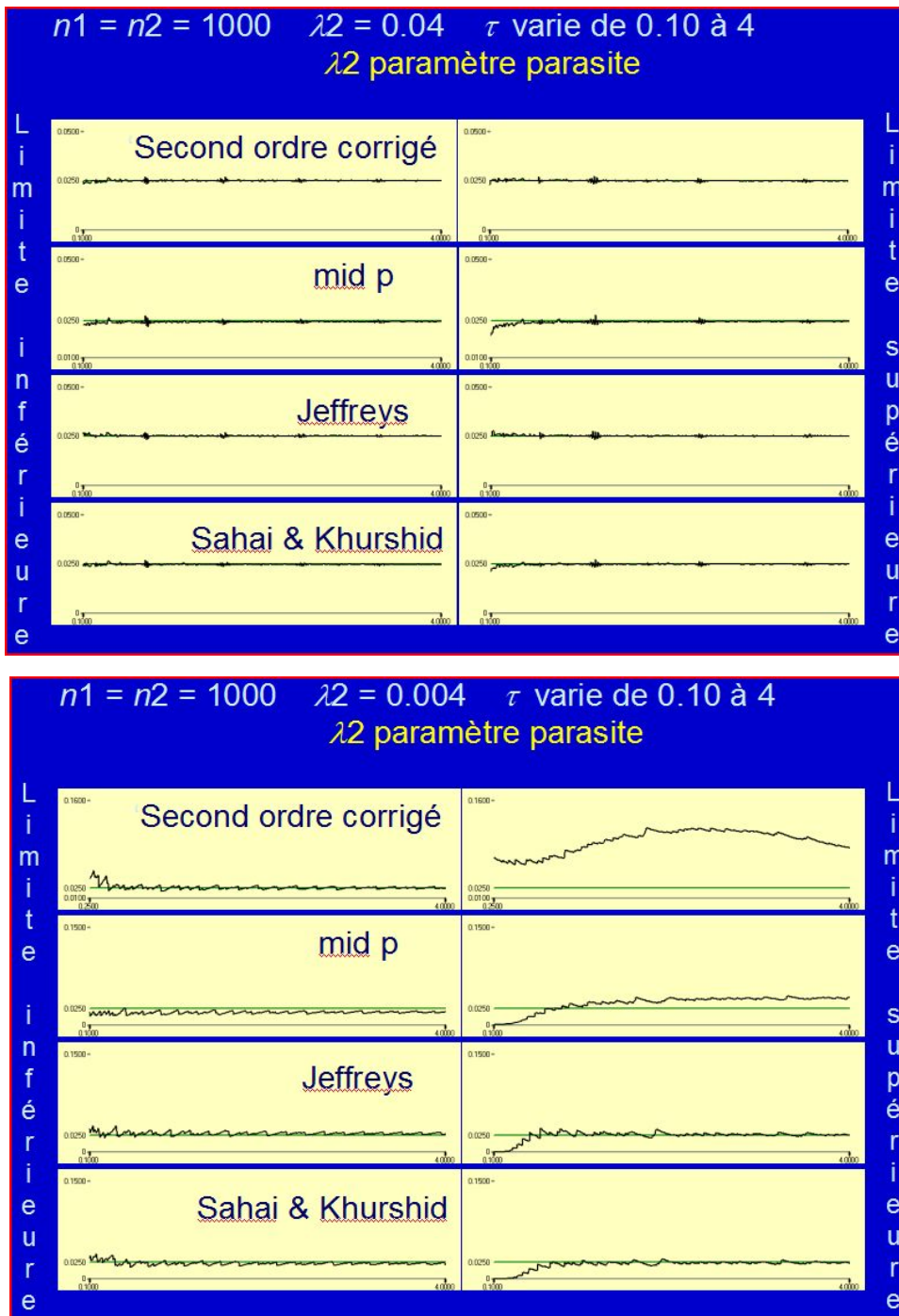


FIG. 1 – Probabilités fréquentistes d’erreur inférieure et supérieure, pour  $n_1 = n_2 = 1000$  et  $\tau$  variant de 0.10 à 4, quand  $\lambda_2 = 0.04$  (en haut) et  $\lambda_2 = 0.004$  (en bas).

## GENERALISATIONS

Alors que les procédures fréquentistes nécessitent une solution *ad hoc* pour chaque nouvelle situation, l'approche bayésienne fournit des procédures générales et s'applique de la même manière, par exemple à l'inférence sur la différence  $\lambda_1 - \lambda_2$  ou encore dans le cas de groupes stratifiés. Dans chaque situation, l'approche bayésienne objective conduit en outre à des procédures, qui sont très performantes d'un point de vue fréquentiste.

## Bibliographie

- [1] Agresti, A. et Coull, B.A. (1998) Approximate is better than "exact" for interval estimation of binomial proportions. *American Statistician*, 52, 119–126.
- [2] Berger, J. (2004) The case for objective Bayesian analysis. *Bayesian Analysis*, 1, 1–17.
- [3] Berry, G. et Armitage, P. (1995) Mid-P confidence intervals : A brief review. *The Statistician*, 44, 417–423.
- [4] Cai, T.T. (2005) One-sided confidence intervals in discrete distributions. *Journal of Statistical Planning and Inference*, 131, 63–88.
- [5] Johnson, N.L., Kotz, S. et Kemp, A.W. (1993) *Univariate Discrete Distributions* (2nd Edition). New York : John Wiley.
- [6] Lecoutre B. (2008) Bayesian methods for experimental data analysis. In C.R. Rao, J. Miller & D.C. Rao (Eds.), *Handbook of statistics : Epidemiology and Medical Statistics* (Vol 27). Amsterdam : Elsevier, 775–812.
- [7] Pires, A.M. et Amado, C. (2008) Interval estimators for a binomial proportion : comparison of twenty methods. *REVSTAT, Statistical Journal*, 6, 1–33
- [8] Price, M.P. et Bonett, D.G. (2000) Estimating the ratio of two Poisson rates. *Computational Statistics & Data Analysis*, 34, 345–356.
- [9] Sahai, H., Khurshid, A., (1993) Confidence intervals for the ratio of two Poisson means. *The Mathematical Scientist*, 18, 43–50.