REVUE DE STATISTIQUE APPLIQUÉE

R. SNEYERS

Sur l'estimation du nombre équivalent de répétitions

Revue de statistique appliquée, tome 19, nº 2 (1971), p. 35-47
<http://www.numdam.org/item?id=RSA_1971__19_2_35_0>

SUR L'ESTIMATION DU NOMBRE ÉQUIVALENT DE RÉPÉTITIONS

R. SNEYERS

Institut Royal Météorologique de Belgique

1 - INTRODUCTION

On sait que pour caractériser l'interdépendance des termes d'une série chronologique, Bartels (cf. par ex. [1]) a introduit la notion du nombre équivalent de répétitions défini par la relation :

$$\omega = n \text{ var } \bar{y}_n / \text{var } y \, , \tag{1}$$

où var $\bar{y}_n$ désigne la variance de la moyenne de n termes consécutifs de la série et où var y est la variance des termes eux-mêmes[1].

On se souvient aussi (cf. [2] p. 43) qu'en faisant appel aux cœfficients d'autocorrélation ρ_i d'ordre $i = 1, 2, \ldots, n - 1$, la relation (1) peut s'écrire :

$$\omega = 1 + 2 \, [(n - 1) \; \rho_1 + (n - 2) \; \rho_2 + \ldots + \rho_{n-1}]/n \, , \tag{2}$$

d'où il ressort que pour une série aléatoire indépendante, c'est-à-dire pour laquelle $\rho_i = 0$ quel que soit i, on a pour toute valeur de n :

$$\omega = 1 \, , \tag{3}$$

qu'à l'opposé, si y est une constante, c'est-à-dire si $\rho_i = 1$, quel que soit i, on a :

$$\omega = n \, , \tag{4}$$

et que, d'une manière générale, les relations $\rho_i \geqslant 0$ conduisent à la relation :

$$1 \leqslant \omega \leqslant n \tag{5}$$

----

(1) On voit sans peine que ω est le nombre par lequel il faut diviser l'effectif n pour obtenir le nombre d'observations indépendantes conduisant à une estimation de la moyenne de même précision.

En particulier, si la série chronologique y_j, $j = \ldots, -1, -1+1, \ldots, 0, 1, 2, \ldots, 1, \ldots$ forme une série de Markov de paramètre $\rho > 0$, c'est-à-dire qu'on a, quelque soit j (cf. [3] p. 418) :

$$y_j = \rho y_{j-1} + \varepsilon_j \, , \; 0 < \rho \leqslant 1 \tag{6}$$

où ε est une variable aléatoire indépendante, on a aussi quel que soit i, les égalités :

$$\rho_i = \rho^i \tag{7}$$

desquelles on déduit avec (2) et la condition $\rho > 0$, que $\frac{d\omega}{d\rho} > 0$ et qu'ainsi ω croît uniformément de 1 à n, lorsque ρ varie de 0 à 1. On vérifie en outre que pour $\rho = 0$ et $\rho = 1$, il vient respectivement :

$$\frac{d\omega}{d\rho} = 2(n-1)/n \quad \text{et} \quad \frac{d\omega}{d\rho} = (n-1)(n+1)/3 \tag{8}$$

Enfin pour le calcul de ω, il peut être utile de rappeler que, grâce à (7), la relation (2) peut se mettre sous la forme (cf. [4] p. 234) :

$$\omega = \frac{1+\rho}{1-\rho} - \frac{2\rho(1-\rho^n)}{n(1-\rho)^2} \tag{9}$$

ce qui permet d'écrire aussi :

$$\frac{d\omega}{d\rho} = 2(1-\rho)^{-2}\left[1 + \rho^n - \frac{1+\rho}{n} \cdot \frac{1-\rho^n}{1-\rho}\right] \tag{10}$$

L'étude qui suit a été plus spécialement entreprise en vue de son application aux séries persistantes ($\rho > 0$). On notera cependant que les relations qui ont été développées restent vraies pour $\rho < 0$.

En particulier, avec $\rho = -1$, de (9), on tire, $\omega = 0$ ou $1/n$ selon que n est pair ou impair. De plus, comme on a aussi :

$$\frac{d\omega}{d\rho} = 2(1-\rho)^{-2}(1+\rho)\left[(n-1) - (n+1)\rho + (n-1)\rho^2 - \ldots\right]/n \tag{11}$$

on voit immédiatement que si $-1 < \rho < 0$, on a aussi $\frac{d\omega}{d\rho} > 0$, avec $\frac{d\omega}{d\rho} = 0$ pour $\rho = -1$.

Il s'ensuit que ω croît uniformément avec ρ sur tout l'intervalle de -1 à $+1$.

2 - L'ESTIMATION OPTIMALE

Dans ce qui suit nous supposons que la série chronologique considérée est une série de Markov normale et que, pour l'estimation de ω, on dispose d'un échantillon formé par k séries indépendantes de n valeurs consécutives :

$$y_{1j}, y_{2j}, \ldots, y_{kj}, \; j = 1, 2, \ldots, n. \tag{1}$$

Comme il est établi que dans ce cas une relation biunivoque et réciproque lie ω à ρ, on en déduit que l'estimateur par la méthode de la vraisemblance maximale de ω sera obtenu en introduisant dans 1. (2) l'estimateur correspondant de ρ, en tenant compte de 1. (7).

Si l'on pose alors $m = E(y)$ et $\sigma^2 = \text{var } \varepsilon$, la fonction de vraisemblance peut s'écrire :

$$\log L = -\sum_{ij} \left[\left(u_{ij} - \rho u_{i(j-1)}\right)^2 / (2\sigma^2) - \frac{1}{2} \log 2\pi - \log \sigma \right] \tag{2}$$

$$(i = 1, 2, \ldots, k, j = 2, 3, \ldots, n)$$

avec :

$$u_{ij} = y_{ij} - m \tag{3}$$

On en déduit, par élimination de m et de ρ entre les équations :

$$\frac{\partial \log L}{\partial m} = 0 \quad \text{et} \quad \frac{\partial \log L}{\partial \rho} = 0 \tag{4}$$

que les estimations cherchées de m et de ρ sont les solutions d'équations du second degré dont les cœfficients s'expriment en fonction des termes y_{ij}.

De plus, le calcul de $\frac{\partial^2 \log L}{\partial m^2}$ et de $\frac{\partial^2 \log L}{\partial \rho^2}$ montre que pour ces estimateurs, on a :

$$\text{var } \hat{m} = \omega \text{ var } y/(n-1)k \cong \text{var } \bar{y}_n/k \tag{5}$$

et

$$\text{var } \hat{\rho} = (1 - \rho^2)/(n-1)k \tag{6}$$

La complexité du calcul de ces estimations peut toutefois être évitée en adoptant comme estimateur de m la solution de l'équation :

$$\sum_{ij} u_{ij} = 0 \tag{7}$$

et pour ρ, un estimateur de la forme :

$$r_1 = \sum_{ij} u_{ij}\, u_{i(j-1)} / \sum_{ij} u_{ij}^2 \tag{8}$$

En particulier, en associant à l'estimateur de m défini par (7) celui de ρ donné en (8) sous la forme circulaire, c'est-à-dire en faisant $j = 2, 3, \ldots, (n+1)$, avec :

$$u_{i(n+1)} \equiv u_{i1} \tag{9}$$

on obtient pour m et ρ des estimateurs dont les variances se confondent pratiquement avec celles données en (5) et (6).

On notera toutefois que, comme dans les cas analogues (cf. [3] p. 433), l'estimateur r_1 de ρ possède un biais. La détermination de ce biais ainsi que le calcul de var r_1 seront examinés plus loin.

Finalement, en adoptant pour ρ l'estimateur r_1 corrigé pour le biais, soit r_1', on en déduit, grâce aux relations 1. (7) et 1. (2), un estimateur $\hat{\omega}$ de ω de variance :

$$\text{var } \hat{\omega} = \left(\frac{d\omega}{d\rho}\right)^2 \text{ var } r_1' \tag{10}$$

3 - UNE ESTIMATION SIMPLIFIEE DE ω

L'estimateur de ω qui sera considéré ici est celui qui découle directement de sa définition 1.(1), soit :

$$\omega^* = \sum_i \left(\sum_j \hat{u}_{ij}\right)^2 / \Sigma\ \hat{u}_{ij}^2 \text{ , avec } \hat{u}_{ij} = y_{ij} - \hat{m} \text{ ,} \tag{1}$$

pour lequel nous déterminerons successivement le biais et la variance.

A ce sujet, on notera que si l'on pose les relations :

$$\omega^* = a/b \text{ , } E(a) = \alpha \text{ , } E(b) = \beta \text{ , } a = \alpha + da \text{ , } b^2 = \beta + db \text{ ,} \tag{2}$$

on trouve, en développant $\left(1 + \frac{db}{\beta}\right)^{-1}$ en série et en se limitant aux termes ne dépassant pas le second ordre :

$$\omega^* = \frac{\alpha}{\beta}\left(1 - \frac{d\,b}{\beta} + \frac{d\,b^2}{\beta^2} + \frac{d\,a}{\alpha} - \frac{da\,db}{\alpha\,\beta}\right) \tag{3}$$

d'où on tire :

$$E(\omega^*) = \frac{\alpha}{\beta}\left(1 + \frac{\text{var } b}{\beta^2} - \frac{\text{cov}(a\text{ , }b)}{\alpha\,\beta}\right) \tag{4}$$

ainsi que

$$\text{var } \omega^* = \frac{\alpha^2}{\beta^2} \cdot \left(\frac{\text{var } a}{\alpha^2} + \frac{\text{var } b}{\beta^2} - \frac{2\text{ cov}(a\text{ , }b)}{\alpha\ \beta}\right) \tag{5}$$

relations qui montrent qu'à l'approximation adoptée, la détermination du biais et de la variance de l'estimateur ω^* se ramènent au calcul de α, β, var b et cov(a , b).

3.1 - Calcul de α, β, var a, var b et cov(a , b)

Si l'on écrit, pour simplifier :

$$\sum_j u = \sum_j u_{ij} \quad \text{et} \quad \sum_j \sum_j u = \sum_i \sum_j u_{ij} \tag{1}$$

il vient :

$$a = \sum_i \left(\sum_j u\right)^2 - \left(\sum_i \sum_j u\right)^2 / k \quad \text{et} \quad b = \sum_i \sum_j u^2 - \left(\sum_i \sum_j u\right)^2 / (kn) \tag{2}$$

Si l'on pose alors :

$$E(u^2) = \text{var } y = 1 \tag{3}$$

avec 1.(1), on trouve :

$$E\left[\left(\sum_j u\right)^2\right] = \text{var} \sum_j u = \text{var } n\, \bar{y}_n = n^2 \text{ var } \bar{y}_n = n\omega \tag{4}$$

De plus, en vertu de l'indépendance des k séries y_{ij}, $i = 1, 2, \ldots, k$, on a aussi :

$$E\left[\left(\sum_i \sum_j u\right)^2\right] = E\left[\sum_i \left(\sum_j u\right)^2\right] = kn\omega \tag{5}$$

Avec (2), on en déduit pour α et pour β :

$$\alpha = E(a) = (k-1)\, n\omega \quad \text{et} \quad \beta = E(b) = kn - \omega \tag{6}$$

Pour le calcul de var a, var b et cov(a , b) on se souviendra tout d'abord que la fonction génératrice des moments de la distribution conjointe de variables normales x_1 et x_2 de corrélation ρ et d'écarts-types σ_1 et σ_2 permet de déduire la relation :

$$\text{cov}(x_1^2, x_2^2) = 2(\rho\, \sigma_1\, \sigma_2)^2 = 2 \text{ cov}^2(x_1, x_2) \tag{7}$$

d'où il vient également :

$$\text{var } x_1^2 = 2 \text{ var}^2 x_1 \tag{8}$$

Il en découle immédiatement pour chacune des séries:

$$\text{var}\left(\sum_j u\right)^2 = 2 \text{ var}^2 \sum_j u = 2\, n^2\, \omega^2 , \tag{9}$$

$$\text{var} \sum_j u^2 = \sum_j \text{var } u^2 + 2 \sum_j \sum_{j'} \text{cov}(u_j^2, u_{j'}^2) = 2n\omega(\rho^2) \tag{10}$$

où $\omega(\rho^2)$ désigne la valeur de ω pour le cœfficient de corrélation ρ^2, ainsi que :

$$\text{cov}\left[\sum_j u^2, \left(\sum_j u\right)^2\right] = \sum_j \text{cov}\left[u^2, \left(\sum_j u\right)^2\right] = 2 \sum_j \text{cov}^2\left(u, \sum_j u\right) \tag{11}$$

Si l'on note alors, pour la dernière relation, qu'avec (4) l'on a, par définition :

$$n\omega = \text{var} \sum_j u = \sum_j \text{cov}\left(u, \sum_j u\right) \tag{12}$$

en posant :

$$\text{cov}\left[\sum_j u^2, \left(\sum_j u\right)^2\right] = 2\, n\omega'^2 \tag{13}$$

on pourra écrire, d'une manière approchée :

$$\omega'^2 = \omega^2 \tag{14}$$

En réalité, comme pour $\rho = 0$ et $\rho = 1$, on a respectivement :

$$\text{cov}\left(u, \sum_j u\right) = 1 \quad \text{et} \quad \text{cov}\left(u, \sum_j u\right) = n \tag{15}$$

on a, de même que pour ω^2 :

$$\omega'^2 = 1 \quad \text{ou} \quad \omega'^2 = n^2 \tag{16}$$

Pour le reste, le calcul montre qu'on a :

$$\omega'^2 = (1-\rho)^{-2}\left[(1+\rho)^2 + 2\rho^{n+1} + \frac{2\rho^2}{n} \cdot \frac{1-\rho^{2n}}{1-\rho^2} - \frac{4\rho(1+\rho)}{n} \cdot \frac{1-\rho^n}{1-\rho}\right] \tag{17}$$

ce qui donne :

$$\omega'^2 - \omega^2 = (1-\rho)^{-2}\left[\frac{2\rho^2}{n} \cdot \frac{1-\rho^{2n}}{1-\rho^2} - \frac{4\rho^2(1-\rho^n)^2}{n^2(1-\rho)^2} + 2\rho^{n+1}\right] \tag{18}$$

Finalement, si l'on néglige la différence précédente, on obtient en tenant compte de l'indépendance des k séries :

$$\text{var } a = (k-1) \text{ var}\left(\sum_j u\right)^2 = 2(k-1)\, n^2\, \omega^2$$

$$\text{var } b = k \text{ var}\left(\sum_j u^2\right) + \frac{1}{n^2} \text{ var}\left(\sum_j u\right)^2 - \frac{2}{n} \text{ cov}\left[\sum_j u^2, \left(\sum_j u\right)^2\right]$$

$$= 2kn\, \omega(\rho^2) - 2\omega^2$$

$$\text{cov}(a, b) = (k-1) \text{ cov}\left[\left(\sum_j u\right)^2, \sum_j u^2\right] = 2(k-1)\, n\omega^2 \tag{19}$$

3.2 - Moyenne et variance de l'estimateur $\omega^* = a/b$

Des valeurs trouvées pour var a, var b et cov(a , b) en 3.1.(19), on déduit, grâce à 3. (4) et 3. (5) :

$$E(\omega^*) = E(a/b) = \frac{(k-1)\, n\omega}{kn - \omega}\left[1 + 2kn \frac{\omega(\rho^2) - \omega}{(kn-\omega)^2}\right] \tag{1}$$

et

$$\text{var } \omega^* = \text{var}(a/b) = 2\, \frac{(k-1)^2\, n^2\, \omega^2}{(kn-\omega)^2}\left[\frac{1}{k-1} + \frac{kn\,[\omega(\rho^2) - 2\omega] + \omega^2}{(kn-\omega)^2}\right] \tag{2}$$

On en conclut que si on corrige les quantités a et b de leur biais propre, et si on pose :

$$\delta = 2kn \frac{\omega(\rho^2) - \omega}{(kn-\omega)^2}, \tag{3}$$

le biais résiduel de ω^* est égal à la quantité $\delta\omega$. Cette quantité est nulle pour $\rho = 0$ et $\rho = 1$, tandis que pour $0 < \rho < 1$, elle est négative puisqu'alors

$\omega(\rho^2) < \omega$. De plus, cette quantité devient négligeable dès que k et n sont suffisamment grands.

Enfin, avec 2.(10), 2.(6) et 1.(8), pour $\rho = 0$, on obtient :

$$\text{var } \omega^*/\text{var } \hat{\omega} = \frac{(kn)^2}{(kn-1)^2} \frac{(k-1)\, n^2}{2(kn-1)} \simeq \frac{n}{2} \tag{4}$$

Ce qui montre qu'au voisinage de $\rho = 0$, l'estimateur ω^* est moins efficace que l'estimateur $\hat{\omega}$ dès que $n > 2$.

Pour $\rho = 1$, on trouve, par contre :

$$\text{var } \omega^* = 0 \tag{5}$$

Le comportement du rapport var ω^*/var $\hat{\omega}$ au voisinage de $\rho = 1$ ne sera toutefois examiné qu'après le calcul de la moyenne et de la variance de r_1.

4 - CALCUL DE LA MOYENNE ET DE LA VARIANCE DE R_1

Si l'on pose :

$$a' = \sum_i \sum_j u_{ij}\, u_{i(j-1)} - \Big(\sum_i \sum_j u\Big)^2 / kn \,, \tag{1}$$

il vient, avec 3.1.(2) :

$$r_1 = a'/b \tag{2}$$

De plus, en écrivant :

$$\alpha' = E(a') \tag{3}$$

on aura, comme en 3.(4) et 3.(5) :

$$E(r_1) = \frac{\alpha'}{\beta}\left(1 + \frac{\text{var } b}{\beta^2} - \frac{\text{cov}(a', b)}{\alpha' \beta}\right) \tag{4}$$

et :

$$\text{var } r_1 = \frac{\alpha'^2}{\beta^2}\left(\frac{\text{var } a'}{\alpha'^2} + \frac{\text{var } b}{\beta^2} - \frac{2\,\text{cov}(a', b)}{\alpha' \beta}\right) \tag{5}$$

d'où il ressort que α', cov(a' , b) et var a' sont les seules quantités qui restent à déterminer.

4.1 - Calcul de α', var a' et cov(a' , b)

Pour α', on trouve immédiatement :

$$\alpha' = E(a') = k(n-1)\,\rho + k\rho^{n-1} - \omega \tag{1}$$

Pour var a' et cov(a' , b), il convient tout d'abord de calculer les expressions :

$$\text{var}\left[\sum_i \sum_j u_j\ u_{j-1}\right] , \text{cov}\left[\sum_i \sum_j u_j\ u_{j-1} , \left(\sum_i \sum_j u\right)^2\right]$$

et

$$\text{cov}\left[\sum_i \sum_j u_j\ u_{j-1} , \sum_i \sum_j u^2\right]$$

A cet effet, on se souviendra que la fonction génératrice des moments de la distribution conjointe des variables normales x_1, x_2, x_3 et x_4 permet de tirer la relation :

$$\text{cov}(x_1\ x_2 , x_3\ x_4) = \text{cov}(x_2 , x_3)\ \text{cov}(x_1 , x_4) + \text{cov}(x_1 , x_3)\ \text{cov}(x_2 , x_4) \qquad (2)$$

donc aussi, la relation :

$$\text{cov}(x_1\ x_2 , x_3^2) = 2\ \text{cov}(x_1 , x_3)\ \text{cov}(x_2 , x_3) \qquad (3)$$

Si l'on note alors qu'on a, en vertu de l'indépendance des k séries y_{ij}, $i = 1 , 2 , \dots , k$:

$$\text{var} \sum_i \sum_j u_j\ u_{j-1} = k \left[\sum_j \text{var}\ u_j\ u_{j-1} + 2 \sum_{j'>j} \text{cov}(u_j\ u_{j-1} , u_{j'}\ u_{j'-1})\right] \qquad (4)$$

$$\text{cov}\left[\sum_i \sum_j u_j\ u_{j-1} , \left(\sum_i \sum_j u\right)^2\right] = k\ \text{cov}\left[\sum_j u_j\ u_{j-1} , \left(\sum_j u\right)^2\right] \qquad (5)$$

et

$$\text{cov}\left[\sum_i \sum_j u_j\ u_{j-1} , \sum_i \sum_j u^2\right] = k\ \text{cov}\left[\sum_j u_j\ u_{j-1} , \sum_j u^2\right] \qquad (6)$$

en posant, comme ci-dessus la relation :

$$\text{cov}\left[\sum_j u_j\ u_{j-1} , \left(\sum_j u\right)^2\right] = 2\ n\omega''^2 \qquad (7)$$

qui donne également d'une manière approchée :

$$\omega''^2 = \omega^2 \qquad (8)$$

pour $0 < \rho < 1$, tandis que pour $\rho = 0$ et $\rho = 1$, l'égalité est réalisée, l'application aux développements (4), (5) et (6) des égalités (2) et (3) ainsi que de leurs analogues en 3.1.(7) et 3.1.(8), permet de trouver, après réduction des termes semblables :

$$\text{var a}' = kn + k(n-1)\ \rho^2 + k\rho^{2(n-1)}$$
$$+ 2\ k(n-1)\ [\rho^n + \rho^{n-2} + \omega_{n-1}(\rho^2) - 1] - 2\omega^2 \qquad (9)$$

où $\omega_{n-1}(\rho^2)$ est la valeur de ω dans laquelle $(n-1)$ remplace n et ρ^2 remplace ρ, ainsi que :

$$\text{cov}(a' , b) = 2\ kn\ [\omega(\rho^2) - 1 + \rho^n]\ \rho^{-1} - 2\ \omega^2 \qquad (10)$$

4.2 - Moyenne et variance de l'estimateur $r_1 = a'/b$

Pour une valeur quelconque de ρ, cette moyenne et cette variance s'obtiennent en substituant dans 4. (4) et 4. (5) les valeurs trouvées en 4.1 pour α', var a' et cov(a' , b) ainsi que celles obtenues dans 3.1 pour β et var b.

En particulier, si on se limite aux termes en 1/kn du développement de $E(r_1)$, on trouve pour $0 < \rho < 1$:

$$E(r_1) = \left(1 - \frac{1}{n}\right) \rho - \frac{1 + 3\rho}{kn} \tag{1}$$

formule qui généralise celle donnée par Kendall [5] dans le cas où k = 1, tandis que pour var r_1, on obtient :

$$\text{var } r_1 = \frac{1 - \rho^2}{kn} \tag{2}$$

valeur qui, à l'approximation considérée, se confond comme annoncé avec 2. (6).

De plus, pour $\rho = 0$, on a :

$$\alpha' = -1 \text{ , } \beta = kn - 1 \text{ , var } a' = kn - 2 \text{ , cov}(a' , b) = -2 \text{ ,}$$

$$\text{var } b = 2(kn - 1) \tag{3}$$

ce qui, avec 4. (4) et 4. (5) donne successivement :

$$E(r_1) = -\frac{1}{kn - 1}$$

et

$$\text{var } r_1 = \frac{kn(kn - 3)}{(kn - 1)^3} \tag{4}$$

Enfin, pour $\rho = 1$, on a :

$$\alpha' = \beta = (k - 1)\, n \tag{5}$$

et

$$\text{var } a' = \text{cov}(a' , b) = \text{var } b = 2(k - 1)\, n^2 \tag{6}$$

ce qui pour k > 1 entraîne :

$$E(r_1) = 1 \quad \text{et} \quad \text{var } r_1 = 0 \tag{7}$$

Il s'ensuit que le rapport var ω^*/var $\hat{\omega}$ est indéterminé pour $\rho = 1$.

Par ailleurs, on constate aussi que pour k = 1, les quantités $E(r_1)$ et var r_1 sont également indéterminées.

4.3 - Valeurs limites pour $\rho = 1$ de $E(r_1)$ et de var r_1 lorsque $k = 1$ et du rapport var ω^*/var $\hat{\omega}$ pour $k > 1$

Comme il s'agit dans chaque cas de rapports dont les termes s'annulent pour $\rho = 1$, l'indétermination sera levée en calculant pour $\rho = 1$ la dérivée de chacun des termes du rapport, en répétant l'opération de dérivation autant de fois qu'il est nécessaire.

Dans ces conditions, en tenant compte de 1.(8), on trouve pour $k = 1$ et $\rho = 1$:

$$\frac{d\alpha'}{d\rho} = -(n-1)(n-5)/3 \quad \text{et} \quad \frac{d\beta}{d\rho} = -(n^2-1)/3 \tag{1}$$

d'où :

$$\lim_{\rho=1} \frac{\alpha'}{\beta} = (n-5)/(n+1) \tag{2}$$

De même, on trouve pour les mêmes conditions :

$$\frac{d^2 \text{ var } a'}{d\rho^2} = \frac{2}{9}(n-1)(n^3 - 5n^2 - 4n + 38) \tag{3}$$

$$\frac{d^2 \text{ var } b}{d\rho^2} = \frac{2}{9}(n^2-1)(n^2+2) \tag{4}$$

$$\frac{d^2 \text{ cov}(a', b)}{d\rho^2} = \frac{2}{9}(n-1)(n^3 - 2n^2 - 10n + 2) \tag{5}$$

d'où l'on tire respectivement avec 4.(4) et 4.(5) :

$$\lim_{\rho=1} E(r_1) = \frac{n-5}{n+1} - \frac{3n^2 - 12n + 12}{(n+1)(n^2-1)} \tag{6}$$

et

$$\lim_{\rho=1} \text{var } r_1 = \frac{18n^2 - 90n + 108}{(n+1)^3(n-1)} \tag{7}$$

En ce qui concerne le rapport var ω^*/var $\hat{\omega}$ on remarque tout d'abord que si l'on pose :

$$A = \beta^2 \text{ var } a + \alpha^2 \text{ var } b - 2\alpha\beta \text{ cov}(a, b) \tag{8}$$

et

$$A' = \beta^2 \text{ var } a' + \alpha'^2 \text{ var } b - 2\alpha'\beta \text{ cov}(a', b), \tag{9}$$

il vient immédiatement :

$$\text{var } \omega^*/\text{var } r_1 = A/A' \tag{10}$$

Si l'on note en outre que l'on a aussi pour $\rho = 1$:

$$\frac{d^2 A}{d\rho^2} = \frac{d^2 A}{d\omega^2} \left(\frac{d\omega}{d\rho}\right)^2 \tag{11}$$

avec 2. (10), on obtient :

$$\text{var } \omega^*/\text{var } \hat{\omega} = \frac{d^2 A}{d\omega^2} \Big/ \frac{d^2 A'}{d\rho^2} \tag{12}$$

Dans ces conditions, sachant que pour $\rho = 1$, en plus de 4.2.(5) et 4.2.(6) on a aussi :

$$\text{var } a = n \text{ cov}(a , b) = n^2 \text{ var } b \text{ et } 2\alpha = \text{var } b \tag{13}$$

en posant var b = v, après quelques calculs, on trouve, pour $\rho = 1$:

$$\frac{d^2 A}{d\omega^2} = 2 \text{ v}(kn)^2 + \frac{1}{2} k \text{ v}^2 \left[\frac{3 n^2}{n^2 - 1} - 2\right] \tag{14}$$

et

$$\frac{d^2 A'}{d\rho^2} = 2 \text{ v } [2 \text{ k}(n - 1)]^2 + 4 \text{ k } \beta^2 (n - 1)(n - 2) \tag{15}$$

ce qui avec (12) donne :

$$\text{var } \omega^*/\text{var } \hat{\omega} \cong 3/10 \tag{16}$$

Il en résulte qu'il existe une valeur de ρ à partir de laquelle l'estimateur ω^* devient plus efficace que l'estimateur $\hat{\omega}$.

On notera immédiatement que ce résultat n'est pas en contradiction avec les propriétés optimales des estimateurs par la méthode de la vraisemblance maximale puisque l'estimateur $\hat{\omega}$ est dérivé de l'estimateur r_1 qui n'est qu'une forme approchée de la solution par la vraisemblance maximale.

5 - UN EXEMPLE

Pour concrétiser les résultats trouvés, nous avons considéré le cas où n = 30 et k = 15, cas qui correspond, en particulier, à des séries d'observations journalières d'un mois de trente jours au cours de quinze années, en supposant que les séries d'observations sont indépendantes d'une année à l'autre.

Il se confirme ainsi (tableau I), grâce au calcul des éléments nécessaires au moyen des formules qui précèdent :

1/ que la substitution de ω'^2 par ω^2 est justifiée par la petitesse de la différence entre ces deux quantités,

2/ que, si l'on corrige les quantités a et b pour le biais, le biais résiduel reste petit en regard de l'écart-type de l'erreur d'estimation qui affecte ω^*, 3/ que l'efficacité relative de l'estimateur ω^* par rapport à l'estimateur $\hat{\omega}$ croit avec ρ et que, dans ce cas-ci, la valeur de ρ à partir de laquelle le premier est plus efficace que le second est proche de 0,95,

4/ que même pour k = 15, le biais de r_1 reste sensible pour les valeurs élevées de ρ (exception faite de la limite) puisque, ainsi qu'on peut le vérifier, ce biais dépasse l'écart-type de r_1 pour $\rho > 0,8$,

5/ que pour k = 1, l'estimateur r_1 présente un biais régulièrement croissant de 0 à 1.

On voit donc, en résumé, que pour les valeurs faibles ou moyennes de ρ, l'estimateur ω^* est un estimateur assez peu intéressant, en raison de son efficacité médiocre.

Par contre, pour les valeurs élevées de ρ, il nous semble qu'il peut rendre d'utiles services. Son efficacité croissante et la simplicité de son calcul peuvent même être déterminants pour son emploi. principalement lorsqu'il s'agit de grands échantillons (k et n élevés) pour lesquels le traitement optimal n'est pas requis.

TABLEAU I

Valeurs, en fonction du coefficient ρ, du nombre équivalent de répétitions ω, de l'écart relatif entre ω'^2 et ω^2, du biais résiduel relatif $|\delta|$ de ω^*, des écarts-types des estimateurs ω^* et $\hat{\omega}$ et du rapport de ces écarts-types (n = 30 et k = 15). Valeurs de $E(r_1)$ pour k = 15 et k = 1, ainsi que de l'écart-type de r_1 pour k = 15

ρ	ω	$(\omega'^2 - \omega^2)/\omega^2$	$\|\delta\|$	$\sigma(\omega^*)$	$\sigma(\hat{\omega})$	$\sigma(\omega^*)/\sigma(\hat{\omega})$	$E(r_1)$	$\sigma(r_1)$	$E(r_1)$ (k = 1)
0,0	1	0	0	0,348	0,091	3,82	− 0,002	0,047	− 0,034
0,1	1,214	0,0005	0,0009	0,420	0,111	3,78	0,094	0,094	0,052
0,2	1,479	0,0018	0,0018	0,508	0,138	3,69	0,190	0,046	0,138
0,3	1,816	0,0036	0,0028	0,618	0,173	3,57	0,286	0,045	0,225
0,4	2,259	0,0058	0,0040	0,760	0,224	3,40	0,382	0,044	0,310
0,5	2,867	0,0086	0,0055	0,950	0,300	3,17	0,478	0,042	0,395
0,6	3,750	0,0122	0,0076	1,215	0,421	2,88	0,574	0,039	0,478
0,7	5,148	0,0167	0,0107	1,605	0,635	2,53	0,670	0,035	0,559
0,8	7,668	0,0211	0,0158	2,213	1,099	2,01	0,766	0,031	0,635
0,9	13,254	0,0170	0,0246	3,057	2,166	1,41	0,864	0,025	0,699
0,95	19,104	0,0075	0,0273	3,054	3,080	0,99	0,918	0,020	0,720
0,96	20,754	0,0054	0,0263	2,864	3,177	0,90	0,931	0,018	0,723
0,97	22,628	0,0033	0,0239	2,538	3,128	0,81	0,945	0,015	0,724
0,98	24,762	0,0017	0,0196	2,016	2,781	0,73	0,961	0,012	0,725
0,99	27,202	0,0004	0,0122	1,211	1,891	0,64	0,979	0,007	0,724
1,00	30,000	0	0	0	0	0,56	1,000	0	0,722

BIBLIOGRAPHIE

[1] BARTELS, J. - Gesetz und Zufall in der Geophysik. Naturwiss. 31, p. 421-435, 1943.

[2] VAN DER BIJL, W. - Toepassing van statistische methoden in de klimatotologie. Kon. Ned. Met. Inst., Med. en Verhand. 58, p. 107, 1952.

[3] KENDALL, M.G. and STUART, A. - The Advanced Theory of Statistics, Vol. 3, London, 1966.

[4] SNEYERS, R. - Sur la représentation des séries météorologiques au moyen de processus aléatoires stationnaires persistants, Arch. Met. Geoph. Biokl. B, 10, p. 232-242, 1960.

[5] KENDALL, M.G. - Note on bias in the estimation of autocorrelation, Biometrika, 41, p. 403, 1954.