

# REVUE DE STATISTIQUE APPLIQUÉE

B. COLIN

N. MELLAL

## **Indice d'efficacité relative fondé sur la statistique d'ordre**

*Revue de statistique appliquée*, tome 54, n° 1 (2006), p. 61-83

[http://www.numdam.org/item?id=RSA\\_2006\\_\\_54\\_1\\_61\\_0](http://www.numdam.org/item?id=RSA_2006__54_1_61_0)

© Société française de statistique, 2006, 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/>

## INDICE D'EFFICACITÉ RELATIVE FONDÉ SUR LA STATISTIQUE D'ORDRE

B. COLIN<sup>(1)</sup>, N. MELLAL<sup>(2)</sup>

<sup>(1)</sup> *Département de Mathématiques, Université de Sherbrooke, Sherbrooke J1K-2R1, (Québec) Canada*

<sup>(2)</sup> *Département de Mathématiques et de Statistique, Université d'Ottawa, Ottawa K1N-6N5, (Ontario) Canada*

### RÉSUMÉ

Soit  $(X, Y)$  un vecteur aléatoire de  $\mathbb{R}^2$  et soient respectivement  $\xi_p, \xi_q$  et  $\eta_r, \eta_s$  les quantiles d'ordres  $p$  et  $q$  ( $p > q$ ) et d'ordres  $r$  et  $s$  ( $r > s$ ) des composantes  $X$  et  $Y$ . On s'intéresse alors, en tant qu'indice relatif de dispersion, au rapport  $\frac{(\xi_p - \xi_q)^2}{(\eta_r - \eta_s)^2}$  et l'on montre que le rapport empirique correspondant est, pour ce dernier, un estimateur asymptotiquement normal, sans biais et convergent et qu'en outre il peut dans certains cas représenter, pour le rapport des variances  $\frac{\sigma_X^2}{\sigma_Y^2}$ , un estimateur concurrent du rapport des variances empiriques. Quelques simulations illustrent les résultats de normalité asymptotique et permettent de comparer les deux estimateurs.

**Mots-clés :** *Loi bidimensionnelle, étendues interquantiles, copule, normalité asymptotique, rapport des variances*

### ABSTRACT

Let  $(X, Y)$  be a bivariate random vector and let  $\xi_p, \xi_q$  ( $p > q$ ) and  $\eta_r, \eta_s$  ( $r > s$ ) be respectively the quantiles of order  $p$  and  $q$  and the quantiles of order  $r$  and  $s$  of the components  $X$  and  $Y$ . We consider, as a relative measure of dispersion, the ratio  $\frac{(\xi_p - \xi_q)^2}{(\eta_r - \eta_s)^2}$  for which we show that its empirical counterpart is an asymptotically normal, unbiased and efficient estimator and this one, moreover, may be consider under certain conditions, as a competitor of the usual estimator of the variances ratio  $\frac{\sigma_X^2}{\sigma_Y^2}$ . Some simulations illustrate the asymptotic normality results and provide us with comparison of the two estimators.

**Keywords :** *Bivariate random vector, interquantile range, copula, asymptotic normality, variances ratio*

## 1. Introduction

Les notions de risque, de rendement, de volatilité que l'on rencontre, entre autres, dans les domaines de l'économie, de la finance et des assurances, sont associées en général aux variances des variables aléatoires présentes dans un modèle donné. Parmi ces modèles, ceux qui font appel à des lois stables ou encore à des lois dites « asymptotiquement parétiennes », posent problème puisque pour ces dernières, les variances des variables étant pour certaines valeurs des paramètres infinies (voir par exemple R.E. Feldman, M.S. Taquq [7]), il n'est pas possible de considérer les notions précédentes telles qu'habituellement définies. Notant cependant que les étendues des intervalles interquartiles et plus généralement des intervalles interquantiles constituent des indices de dispersion, on peut alors, dans le but d'exprimer néanmoins une notion de risque relatif en l'absence de variances des variables aléatoires, considérer le rapport des carrés des étendues interquartiles puisque ce dernier s'exprime avec les mêmes unités que le rapport des variances. De plus, dans le cas où les variances sont finies, les estimations du risque relatif ou de l'efficacité relative, telle que définie dans le cadre de procédures statistiques, reposent sur des estimateurs de variances que l'on sait particulièrement sensibles aux valeurs aberrantes. Dans ce contexte, le rapport des carrés des étendues de deux intervalles interquantiles empiriques peut s'avérer également une alternative intéressante au rapport des variances empiriques en tant qu'estimateur du risque relatif ou de l'efficacité relative (tant asymptotique qu'à distance finie). Désignant respectivement par  $\hat{\xi}_p, \hat{\xi}_q$  et par  $\hat{\eta}_r, \hat{\eta}_s$  les quantiles empiriques d'ordres  $p, q$  et  $r, s$  de deux variables aléatoires  $X$  et  $Y$  quelconques, nous nous proposons dans les paragraphes qui suivent d'étudier, dans le cadre des lois bidimensionnelles absolument continues par rapport à la mesure de *Lebesgue* de  $\mathbb{R}^2$ , le rapport  $\frac{(\hat{\xi}_p - \hat{\xi}_q)^2}{(\hat{\eta}_r - \hat{\eta}_s)^2}$  (espérance, variance, loi asymptotique) et en particulier dans le cas où  $p = r$  et où  $q = s$  qui, dans un certain nombre de situations, correspond au rapport  $\frac{Var(X)}{Var(Y)}$ . Enfin, quelques résultats découlant de cette étude sont illustrés à l'aide de simulations.

## 2. Généralités

Considérons un vecteur aléatoire de  $\mathbb{R}^2$ , de composantes  $X$  et  $Y$ , auquel est associé une structure statistique paramétrique dominée de la forme :

$$(\mathbb{R}^2, \mathcal{B}_{\mathbb{R}^2}, \{\mathbb{P}_\theta\} ; \theta \in \Theta, \mu)$$

où  $\mu$  désigne la mesure de *Lebesgue* de  $\mathbb{R}^2$  telle que  $\mathbb{P}_\theta \ll \mu$  pour tout  $\theta \in \Theta$  et où  $\Theta$  désigne l'espace paramètre de la famille  $\{\mathbb{P}_\theta\}$  de mesures de probabilité. Soit :

$${}^t(X_1, Y_1), {}^t(X_2, Y_2), \dots, {}^t(X_n, Y_n)$$

un échantillon de taille  $n$  issu du modèle empirique  $(\mathbb{R}^2, \mathcal{B}_{\mathbb{R}^2}, \{\mathbb{P}_\theta\}, \mu) \otimes^n$  et soient respectivement  $\widehat{\xi}_p, \widehat{\xi}_q$  et  $\widehat{\eta}_r, \widehat{\eta}_s$  les quantiles empiriques d'ordres  $p$  et  $q$  ( $p > q$ ) et d'ordres  $r$  et  $s$  ( $r > s$ ) déduits des observations

$$(X_1, X_2, \dots, X_n) \text{ et } (Y_1, Y_2, \dots, Y_n)$$

des variables aléatoires  $X$  et  $Y$ . Dans le but d'introduire une notion «d'efficacité relative» pouvant se substituer au quotient de deux variances dans le cas où ces dernières n'existent pas (ou en cas de variances complexes à calculer, voire dans le cas où l'on ne dispose que d'approximations numériques de celles-ci), nous nous intéresserons dans ce qui suit au rapport des carrés des étendues interquantiles empiriques  $\widehat{R}^2_{(p,q),(r,s)}$  défini par :

$$\widehat{R}^2_{(p,q),(r,s)} = \frac{(\widehat{\xi}_p - \widehat{\xi}_q)^2}{(\widehat{\eta}_r - \widehat{\eta}_s)^2} \text{ où } p > q ; r > s$$

Plus précisément, nous déterminerons les approximations d'ordres 1 et 2 des quantités :

$$\mathbb{E} \left[ \widehat{R}^2_{(p,q),(r,s)} \right] \text{ et } Var \left[ \widehat{R}^2_{(p,q),(r,s)} \right]$$

et nous montrerons que  $\widehat{R}^2_{(p,q),(r,s)}$  est un estimateur asymptotiquement normal, sans biais et convergent du rapport des carrés des étendues interquantiles théoriques :

$$R^2_{(p,q),(r,s)} = \frac{(\xi_p - \xi_q)^2}{(\eta_r - \eta_s)^2}$$

On trouvera par ailleurs en annexe, une expression générale de la loi à distance finie de  $\widehat{R}^2_{(p,q),(r,s)}$ .

### 3. Moments d'ordres 1 et 2 du rapport $\widehat{R}^2_{(p,q),(r,s)}$

De façon générale il n'est pas possible de donner, pour tout  $n$ , une expression simple et explicite de l'espérance et de la variance de  $\widehat{R}^2_{(p,q),(r,s)}$  en raison de la complexité de sa loi (voir annexe). On peut cependant, pour des valeurs de  $n$  données, en obtenir des approximations en utilisant une démarche, en tout point comparable à celle utilisée dans [2], et dont le principe consiste à représenter la loi  $F_{X,Y}(x, y)$  du couple  $(X, Y)$  sous la forme d'une copule et à effectuer, par ailleurs, le développement limité jusqu'à l'ordre 2 de la fonction  $g(\xi_p, \xi_q, \eta_r, \eta_s)$  définie par :

$$g(\xi_p, \xi_q, \eta_r, \eta_s) = \frac{(\xi_p - \xi_q)^2}{(\eta_r - \eta_s)^2}$$

### 3.1. Moment d'ordre 1

Considérant le changement de variables usuel  $Z_1 = F_X(X)$  et  $Z_2 = F_Y(Y)$ , il découle du présent cadre théorique que les variables  $Z_1$  et  $Z_2$  suivent des lois uniformes  $\mathcal{U}_{[0,1]}$  sur l'intervalle  $[0, 1]$  et que la fonction de répartition

$$F_{Z_1, Z_2}(z_1, z_2) = F_{X, Y}(F_X^{-1}(z_1), F_Y^{-1}(z_2))$$

du couple  $(Z_1, Z_2)$ , constitue une copule (voir Sklar [14], Fréchet [8], Kimeldorf et Sampson [9]). Adoptant dans ce qui suit la notation  $X_{r:n}$  afin de désigner la  $r^e$  statistique d'ordre déduite d'un échantillon de taille  $n$  d'une variable aléatoire quelconque  $X$  on a :

$$\{X_{r:n}\}_{1 \leq r \leq n} = \{F_X^{-1}(Z_{1r:n})\}_{1 \leq r \leq n} ; \{Y_{s:n}\}_{1 \leq s \leq n} = \{F_Y^{-1}(Z_{2s:n})\}_{1 \leq s \leq n}$$

Notons alors par  $Z_n = {}^t(Z_{1u:n}, Z_{1t:n}, Z_{2w:n}, Z_{2v:n})$  le vecteur des statistiques d'ordre associées aux quantiles empiriques  $\hat{\xi}_p, \hat{\xi}_q, \hat{\eta}_r$  et  $\hat{\eta}_s$  des variables  $X$  et  $Y$  (on rappelle que  $\hat{\xi}_p = X_{r:n}$  avec  $r = np$  si  $np$  est un nombre entier et  $r = [np + 1]$  sinon, les autres quantiles empiriques se définissant de façon analogue). Désignons respectivement par  $\mu_n = {}^t(\mu_{u:n}, \mu_{t:n}, \nu_{w:n}, \nu_{v:n})$  et par  $\Sigma_n$  le vecteur moyen et la matrice de variance-covariance du vecteur  $Z_n$  et soit enfin :

$$g(z_n) = g(z_{1u:n}, z_{1t:n}, z_{2w:n}, z_{2v:n}) = \frac{(F_X^{-1}(z_{1u:n}) - F_X^{-1}(z_{1t:n}))^2}{(F_Y^{-1}(z_{2w:n}) - F_Y^{-1}(z_{2v:n}))^2}$$

dont le gradient et la matrice hessienne évalués en  $z_n = \mu_n$  seront notés respectivement par  $Gradg|_{\mu_n}$  et par  $D^{(2)}g|_{\mu_n}$ . Suivant la démarche présentée dans [2], on montre alors que les approximations d'ordres 1 et 2 de  $\mathbb{E}(g(Z_n)) = \mathbb{E}(\widehat{R}_{(p,q),(r,s)}^2)$  sont données par :

$$\frac{(F_X^{-1}(\mu_{u:n}) - F_X^{-1}(\mu_{t:n}))^2}{(F_Y^{-1}(\nu_{w:n}) - F_Y^{-1}(\nu_{v:n}))^2}$$

et par :

$$\frac{(F_X^{-1}(\mu_{u:n}) - F_X^{-1}(\mu_{t:n}))^2}{(F_Y^{-1}(\nu_{w:n}) - F_Y^{-1}(\nu_{v:n}))^2} + \frac{1}{2!} \mathbb{E} \left[ {}^t(Z_n - \mu_n) D^{(2)}g|_{\mu_n} (Z_n - \mu_n) \right]$$

où  $\mu_n = \frac{1}{n+1} {}^t(u, t, w, v)$

*Remarque 1.* – Si l'on suppose que  $n, u, t, w$  et  $v$  tendent vers l'infini de sorte que  $n^{-1}u \rightarrow p, n^{-1}t \rightarrow q, n^{-1}w \rightarrow r$  et que  $n^{-1}v \rightarrow s$ , on sait dans ce cas que, sous des conditions très générales, (voir Serfling [13], David [4], Lecoutre et Tassi

[10]), les quantiles empiriques convergent, respectivement, presque sûrement vers  $p, q, r$  et  $s$  et comme par ailleurs  $\mu_n$  converge vers  $\mu = {}^t(p, q, r, s)$ , nous aurons, en vertu du fait que les fonctions  $F_X^{-1}$  et  $F_Y^{-1}$  sont continues et que  $(Z_n - \mu_n) \xrightarrow{p.s} 0$  :

$$\lim_{n \rightarrow \infty} \mathbb{E} \left[ \frac{(\widehat{\xi}_p - \widehat{\xi}_q)^2}{(\widehat{\eta}_r - \widehat{\eta}_s)^2} \right] = \frac{(\xi_p - \xi_q)^2}{(\eta_r - \eta_s)^2}$$

### 3.2. Moment d'ordre 2

Toujours suivant la démarche utilisée dans [2], on montre sans peine que les approximations d'ordres 1 et 2 de  $Var [g(Z_n)] = Var [\widehat{R}^2_{(p,q),(r,s)}]$  sont données par :

$${}^t Gradg|_{\mu_n} \Sigma_n Gradg|_{\mu_n}$$

et par :

$${}^t Gradg|_{\mu_n} \Sigma_n Gradg|_{\mu_n} - \frac{1}{4} \left[ \mathbb{E} \left( {}^t (Z_n - \mu_n) D^{(2)}g|_{\mu_n} (Z_n - \mu_n) \right) \right]^2$$

*Remarque 2.* – À l'aide des dérivées partielles de la fonction  $g$  il est possible d'obtenir des approximations numériques de la moyenne et de la variance de  $\widehat{R}^2_{(p,q),(r,s)}$  en tenant compte du fait que :

$$\begin{aligned} Var(Z_{1_{u:n}}) &= \frac{u(n-u+1)}{(n+1)^2(n+2)} & ; & \quad Var(Z_{1_{t:n}}) = \frac{t(n-t+1)}{(n+1)^2(n+2)} \\ Var(Z_{2_{w:n}}) &= \frac{w(n-w+1)}{(n+1)^2(n+2)} & ; & \quad Var(Z_{2_{v:n}}) = \frac{v(n-v+1)}{(n+1)^2(n+2)} \end{aligned}$$

et que :

$$Cov(Z_{1_{u:n}}, Z_{1_{t:n}}) = \frac{u(n-t+1)}{(n+1)^2(n+2)} ; Cov(Z_{2_{w:n}}, Z_{2_{v:n}}) = \frac{w(n-v+1)}{(n+1)^2(n+2)}$$

Quant à la covariance de  $Z_{1_{r:n}}$  et de  $Z_{2_{s:n}}$  (pour  $r = u, t$  et pour  $s = w, v$ ) il est nécessaire de connaître la loi conjointe du couple  $(Z_{1_{r:n}}, Z_{2_{s:n}})$  pour la déterminer. Cependant, si les variables sont non corrélées ou si le couple  $(Z_{1_{r:n}}, Z_{2_{s:n}})$  est « asymptotiquement indépendant » (voir Srivastava [15]) les calculs précédents s'trouvent simplifiés.

#### 4. Loi asymptotique du rapport des carrés des étendues interquantiles empiriques

À l'aide du théorème de représentation des quantiles de Bahadur [1], on montre (Colin et Mellal [3]) que  $\frac{\widehat{\xi}_p - \widehat{\xi}_q}{\widehat{\eta}_r - \widehat{\eta}_s}$  est un estimateur asymptotiquement normal sans biais et convergent du rapport des étendues interquantiles théoriques  $\frac{\xi_p - \xi_q}{\eta_r - \eta_s}$ . Autrement dit :

$$\sqrt{n} \left( \frac{\widehat{\xi}_p - \widehat{\xi}_q}{\widehat{\eta}_r - \widehat{\eta}_s} - \frac{\xi_p - \xi_q}{\eta_r - \eta_s} \right) \xrightarrow{L} N(0, \sigma^2)$$

où l'expression de  $\sigma^2$  est donnée en annexe.

Ainsi, posant  $g \left( \frac{\widehat{\xi}_p - \widehat{\xi}_q}{\widehat{\eta}_r - \widehat{\eta}_s} \right) = \frac{(\widehat{\xi}_p - \widehat{\xi}_q)^2}{(\widehat{\eta}_r - \widehat{\eta}_s)^2}$ , il vient (voir Serfling [13]) :

$$\sqrt{n} \left( g \left( \frac{\widehat{\xi}_p - \widehat{\xi}_q}{\widehat{\eta}_r - \widehat{\eta}_s} \right) - g \left( \frac{\xi_p - \xi_q}{\eta_r - \eta_s} \right) \right) \xrightarrow{L} N \left( 0, \left( g' \left( \frac{\xi_p - \xi_q}{\eta_r - \eta_s} \right) \right)^2 \sigma^2 \right)$$

c'est-à-dire :

$$\sqrt{n} \left( \frac{(\widehat{\xi}_p - \widehat{\xi}_q)^2}{(\widehat{\eta}_r - \widehat{\eta}_s)^2} - \frac{(\xi_p - \xi_q)^2}{(\eta_r - \eta_s)^2} \right) \xrightarrow{L} N \left( 0, 4 \frac{(\xi_p - \xi_q)^2}{(\eta_r - \eta_s)^2} \sigma^2 \right)$$

ce qui permet de conclure que  $\widehat{R}_{(p,q),(r,s)}^2$  est un estimateur asymptotiquement normal, sans biais et convergent de  $R_{(p,q),(r,s)}^2$ .

Considérons alors le cas  $p = r$  et  $q = s$  et supposons de plus que les lois marginales de  $X$  et de  $Y$  soient, en utilisant des notations évidentes, de la forme :

$$F_X(x, \mu_X, \sigma_X) = F_0 \left( \frac{x - \mu_X}{\sigma_X} \right) \quad \text{et} \quad F_Y(y, \mu_Y, \sigma_Y) = F_0 \left( \frac{y - \mu_Y}{\sigma_Y} \right)$$

Dans ce cas, il vient :

$$\begin{aligned} \mathbb{P}(X \leq \xi_p) &= p = \mathbb{P} \left( \frac{X - \mu_X}{\sigma_X} \leq \frac{\xi_p - \mu_X}{\sigma_X} \right) = F_0 \left( \frac{\xi_p - \mu_X}{\sigma_X} \right) \\ \mathbb{P}(X \leq \xi_q) &= q = \mathbb{P} \left( \frac{X - \mu_X}{\sigma_X} \leq \frac{\xi_q - \mu_X}{\sigma_X} \right) = F_0 \left( \frac{\xi_q - \mu_X}{\sigma_X} \right) \end{aligned}$$

$$\mathbb{P}(Y \leq \eta_p) = p = \mathbb{P}\left(\frac{Y - \mu_Y}{\sigma_Y} \leq \frac{\eta_p - \mu_Y}{\sigma_Y}\right) = F_0\left(\frac{\eta_p - \mu_Y}{\sigma_Y}\right)$$

$$\mathbb{P}(Y \leq \eta_q) = q = \mathbb{P}\left(\frac{Y - \mu_Y}{\sigma_Y} \leq \frac{\eta_q - \mu_Y}{\sigma_Y}\right) = F_0\left(\frac{\eta_q - \mu_Y}{\sigma_Y}\right)$$

On peut donc déduire des relations précédentes que :

$$\frac{\xi_p - \mu_X}{\sigma_X} = \frac{\eta_p - \mu_Y}{\sigma_Y}$$

et que :

$$\frac{\xi_q - \mu_X}{\sigma_X} = \frac{\eta_q - \mu_Y}{\sigma_Y}$$

Par conséquent il vient :

$$\frac{(\xi_p - \xi_q)^2}{(\eta_p - \eta_q)^2} = \frac{\sigma_X^2}{\sigma_Y^2}$$

et ce quels que soient  $p$  et  $q$  ( $p \neq q$ ) appartenant à l'intervalle  $]0, 1[$ . D'où :

$$\sqrt{n} \left( \frac{(\widehat{\xi}_p - \widehat{\xi}_q)^2}{(\widehat{\eta}_p - \widehat{\eta}_q)^2} - \frac{\sigma_X^2}{\sigma_Y^2} \right) \xrightarrow{L} N\left(0, 4 \frac{\sigma_X^2}{\sigma_Y^2} \sigma^2\right)$$

ce qui montre que

$$\widehat{R}^2_{(p,q),(p,q)} = \frac{(\widehat{\xi}_p - \widehat{\xi}_q)^2}{(\widehat{\eta}_p - \widehat{\eta}_q)^2}$$

est un estimateur asymptotiquement normal, sans biais et convergent de  $\frac{\sigma_X^2}{\sigma_Y^2}$ .

## 5. Cas particulier : test et estimation du rapport de deux variances

### 5.1. Construction du test

Supposons que les variables aléatoires  $X$  et  $Y$  soient indépendantes et que les lois de  $X$  et  $Y$  soient de nouveau de la forme :  $F_X(x, \mu_X, \sigma_X) = F_0\left(\frac{x - \mu_X}{\sigma_X}\right)$  et  $F_Y(y, \mu_Y, \sigma_Y) = F_0\left(\frac{y - \mu_Y}{\sigma_Y}\right)$  et posons, afin d'obtenir une expression relativement simple de la variance  $\sigma^2$ ,  $q = 1 - p$  (avec  $p > 0,5$ ). La loi exacte du rapport  $\widehat{R}^2_{(p,1-p),(p,1-p)}$  étant en général beaucoup trop complexe pour être utilisée



dans la pratique, nous nous situerons dans le cadre où l'approximation normale de cette dernière est valide, c'est-à-dire pour des valeurs de  $n$  suffisamment grandes ce qui, bien entendu, limite la portée pratique du test proposé ci-dessous mais qui suffit néanmoins dans bon nombre d'applications. Il est alors aisé dans ces conditions de vérifier que  $\widehat{R}^2_{(p,1-p),(p,1-p)}$  suit approximativement, en supposant, par souci de simplification, que les lois de  $X$  et  $Y$  sont symétriques, une loi normale de moyenne  $\frac{\sigma_X^2}{\sigma_Y^2}$  et de variance  $\frac{4}{n} \frac{\sigma_X^2}{\sigma_Y^2} \sigma^2$  où  $\sigma^2$  est donné par (voir annexe) :

$$\frac{2(1-p)(2p-1)}{(\eta_p - \eta_{1-p})^2} \left[ \frac{1}{f_X^2(\xi_p)} + \frac{\sigma_X^2}{\sigma_Y^2} \frac{1}{f_Y^2(\eta_p)} \right]$$

Considérons alors l'hypothèse nulle  $\mathcal{H}_0$  :

$$\sigma_X^2 = k^2 \sigma_Y^2$$

contre l'hypothèse alternative  $\mathcal{H}_1$  :

$$\sigma_X^2 \neq k^2 \sigma_Y^2$$

Si l'hypothèse nulle est vraie, alors la statistique  $\widehat{R}^2_{(p,1-p),(p,1-p)}$  suit approximativement une loi normale de moyenne  $k^2$  et de variance  $\tau^2$  donnée par :

$$\tau^2 = \frac{8k^2(1-p)(2p-1)}{n(\eta_p - \eta_{1-p})^2} \left[ \frac{1}{f_X^2(\xi_p)} + k^2 \frac{1}{f_Y^2(\eta_p)} \right]$$

Désignons par  $f_0$  la densité associée à  $F_0$  et par  $\xi_p^0, \xi_{1-p}^0$  les quantiles d'ordres  $p$  et  $1-p$  déduits de  $F_0$ . Il vient :

$$\begin{aligned} f_X^2(\xi_p) &= \frac{1}{\sigma_X^2} f_0^2(\xi_p^0) \\ &= \frac{1}{k^2 \sigma_Y^2} f_0^2(\xi_p^0) \\ f_Y^2(\eta_p) &= \frac{1}{\sigma_Y^2} f_0^2(\xi_p^0) \end{aligned}$$

D'où :

$$\tau^2 = \frac{16k^4(1-p)(2p-1)}{n(\eta_p - \eta_{1-p})^2} \left[ \frac{\sigma_Y^2}{f_0^2(\xi_p^0)} \right]$$

soit, en notant que  $\left(\frac{\eta_p}{\sigma_Y} - \frac{\eta_{1-p}}{\sigma_Y}\right)^2 = (\xi_p^0 - \xi_{1-p}^0)^2$  :

$$\tau^2 = \frac{16k^4 (1-p)(2p-1)}{n (\xi_p^0 - \xi_{1-p}^0)^2 f_0^2 (\xi_p^0)}$$

Étant donné un risque de première espèce  $\alpha$ , le test bilatéral construit sur  $\widehat{R}^2_{(p,1-p),(p,1-p)}$  admettra une région d'acceptation de la forme :

$$k^2 - z_{\frac{\alpha}{2}} \tau \leq \widehat{R}^2_{(p,1-p),(p,1-p)} \leq k^2 + z_{\frac{\alpha}{2}} \tau$$

où  $z_{\frac{\alpha}{2}}$  désigne le quantile d'ordre  $1 - \frac{\alpha}{2}$  de la loi normale centrée réduite. La précédente région d'acceptation peut également s'écrire en posant :

$$\frac{4\sqrt{(1-p)(2p-1)/n}}{(\xi_p^0 - \xi_{1-p}^0) f_0 (\xi_p^0)} = \gamma(p, n, f_0)$$

sous la forme :

$$k^2 \left(1 - z_{\frac{\alpha}{2}} \gamma(p, n, f_0)\right) \leq \widehat{R}^2_{(p,1-p),(p,1-p)} \leq k^2 \left(1 + z_{\frac{\alpha}{2}} \gamma(p, n, f_0)\right)$$

ou encore sous la forme :

$$\left| \frac{\widehat{R}^2_{(p,1-p),(p,1-p)}}{k^2} - 1 \right| \leq z_{\frac{\alpha}{2}} \gamma(p, n, f_0)$$

Dans le cas où, par exemple,  $f_0$  correspond à la loi normale centrée réduite, on a pour  $p = 0,75$ , pour  $\alpha = 0,05$ , et pour  $n = 100$  :

$$z_{\frac{\alpha}{2}} \gamma(p, n, f_0) \simeq 0,66$$

ce qui, pour une valeur donnée de  $k^2$  donne comme région d'acceptation de l'hypothèse  $\mathcal{H}_0 : \frac{\sigma_X^2}{\sigma_Y^2} = k^2$  :

$$0,34k^2 \leq \widehat{R}^2_{(p,1-p),(p,1-p)} \leq 1,66k^2$$

En particulier, on acceptera l'hypothèse d'égalité des variances ( $k^2 = 1$ ) si :

$$0,34 \leq \frac{(\widehat{\xi}_p - \widehat{\xi}_q)^2}{(\widehat{\eta}_p - \widehat{\eta}_q)^2} \leq 1,66$$

Le test usuel d'égalité des variances dans le cas de la normalité, donne pour  $n = 100$  et pour  $\alpha = 0,05$ , une région d'acceptation bilatérale de la forme :

$$0,60 \leq \frac{s_x^2}{s_y^2} \leq 1,63$$

Ainsi, malgré le fait que le test construit sur la statistique  $\widehat{R}^2_{(p,1-p),(p,1-p)}$  (qui n'est fonction d'ailleurs que de quatre valeurs numériques seulement) découle d'une approximation normale, les deux intervalles obtenus sont tout à fait comparables.

Si les lois de  $X$  et  $Y$  ne sont pas symétriques, un calcul élémentaire montre que l'on obtient également un intervalle d'acceptation de la forme :

$$k^2 \left( 1 - z_{\frac{\alpha}{2}} \gamma' (p, n, f_0) \right) \leq \widehat{R}^2_{(p,1-p),(p,1-p)} \leq k^2 \left( 1 + z_{\frac{\alpha}{2}} \gamma' (p, n, f_0) \right)$$

mais où la fonction  $\gamma' (p, n, f_0)$  a une expression plus complexe que celle de la fonction  $\gamma (p, n, f_0)$ .

*Remarque 3.* – Il est clair que le test construit ci-dessus permet, dans le cas général et moyennant quelques modifications élémentaires, de tester l'hypothèse nulle :

$$\mathcal{H}_0 : \frac{(\xi_p - \xi_q)^2}{(\eta_p - \eta_q)^2} = k^2$$

contre l'hypothèse alternative :

$$\mathcal{H}_1 : \frac{(\xi_p - \xi_q)^2}{(\eta_p - \eta_q)^2} \neq k^2$$

On peut alors vérifier sans peine dans ce cas que l'intervalle d'acceptation est donné, pour un risque  $\alpha$  de première espèce, par :

$$k^2 - z_{\frac{\alpha}{2}} \tau' \leq \widehat{R}^2_{(p,q),(p,q)} \leq k^2 + z_{\frac{\alpha}{2}} \tau'$$

où  $\tau'$  est une fonction de  $n$ ,  $k$ ,  $p$  et  $q$ .

## 5.2. Intervalle de confiance

Les expressions obtenues ci-dessus permettent d'écrire immédiatement que :

$$\frac{\widehat{R}^2_{(p,1-p),(p,1-p)}}{\left( 1 + z_{\frac{\alpha}{2}} \gamma (p, n, f_0) \right)} \leq k^2 \leq \frac{\widehat{R}^2_{(p,1-p),(p,1-p)}}{\left( 1 - z_{\frac{\alpha}{2}} \gamma (p, n, f_0) \right)}$$

où que :

$$\frac{\widehat{R}^2_{(p,1-p),(p,1-p)}}{\left(1 + z_{\frac{\alpha}{2}} \gamma'(p, n, f_0)\right)} \leq k^2 \leq \frac{\widehat{R}^2_{(p,1-p),(p,1-p)}}{\left(1 - z_{\frac{\alpha}{2}} \gamma'(p, n, f_0)\right)}$$

ce qui donne un intervalle de confiance au seuil  $1 - \alpha$  pour le rapport  $\frac{\sigma_X^2}{\sigma_Y^2}$ .

La même remarque que ci-dessus est valable également en ce qui concerne l'intervalle de confiance, au seuil  $1 - \alpha$ , du rapport :

$$\frac{(\xi_p - \xi_q)^2}{(\eta_p - \eta_q)^2}$$

## 6. Simulations

### 6.1. Distribution empirique du rapport $\widehat{R}^2_{(p,1-p),(p,1-p)}$

Les simulations présentées ci-dessous, illustrent, dans le cas de trois types de lois bidimensionnelles, la normalité asymptotique du rapport  $\widehat{R}^2_{(p,1-p),(p,1-p)}$  avec  $p = 0,75$  (Figures 1, 2 et 3). Les simulations ont été réalisées à l'aide du logiciel S-Plus. On notera à ce sujet que la présence, dans certains graphiques, de valeurs négatives pour le rapport  $\widehat{R}^2_{(p,1-p),(p,1-p)}$  est due au logiciel utilisé qui, à des fins de présentation, complète les extrémités des distributions empiriques. Enfin, pour chaque type de loi et pour chaque taille  $n$  d'échantillon, on donne les valeurs empiriques  $m$  et  $s^2$  de la moyenne et de la variance du rapport  $\widehat{R}^2_{(p,1-p),(p,1-p)}$  ainsi, qu'entre parenthèses et en caractères gras, la valeur de  $n$  et les valeurs théoriques  $\mu = \frac{(\xi_p - \xi_q)^2}{(\eta_p - \eta_q)^2}$  et  $\sigma_n^2 = \frac{4}{n} \frac{(\xi_p - \xi_q)^2}{(\eta_p - \eta_q)^2} \sigma^2$  correspondantes.

#### 6.1.1. Loi normale

On considère 500 échantillons, de tailles respectives  $n = 50, 100$  et  $1000$ , déduits de la loi normale  $N(\nu, \Sigma)$  donnée par :

$$\nu = {}^t(0, 0) \text{ et } \Sigma = \begin{bmatrix} 0,5 & 0,7 \\ 0,7 & 2 \end{bmatrix}$$

On a donc ici  $\mu = \frac{\sigma_X^2}{\sigma_Y^2} = 0,25$

Les distributions empiriques du rapport  $\widehat{R}^2_{(0,75, 0,25),(0,75, 0,25)}$  sont illustrées aux figures 1, 2 et 3 qui suivent :

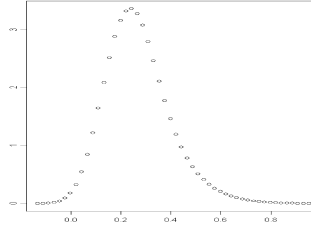


FIGURE 1

$$m = 0,2692; s^2 = 0,0121 \quad (n = 50, \mu = 0,25, \sigma_{50}^2 = 0,0102)$$

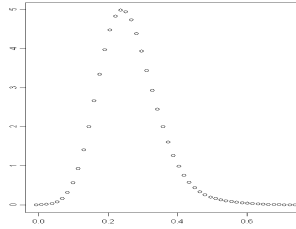


FIGURE 2

$$m = 0,2603; s^2 = 0,0052 \quad (n = 100, \mu = 0,25, \sigma_{100}^2 = 0,00512)$$

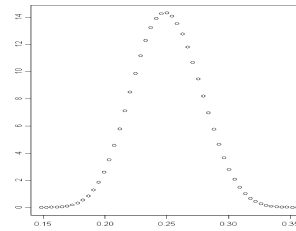


FIGURE 3

$$m = 0,2502; s^2 = 0,00054 \quad (n = 1000, \mu = 0,25, \sigma_{1000}^2 = 0,00051)$$

Puisque l'on se trouve dans le cas où  $R_{(p,q),(p,q)}^2 = \frac{\sigma_X^2}{\sigma_Y^2}$ , on présente à titre indicatif, à l'aide des figures 4, 5 et 6 ci-dessous, les distributions empiriques du rapport  $\frac{\widehat{\sigma_X^2}}{\widehat{\sigma_Y^2}}$ , où  $m$  et  $s^2$  désignent respectivement les valeurs empiriques de la moyenne et de la variance de celui-ci.

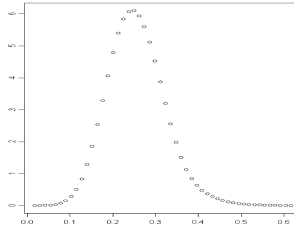


FIGURE 4

$$m = 0,2553; s^2 = 0,0031 \left( n = 50, \frac{\sigma_X^2}{\sigma_Y^2} = 0,25 \right)$$

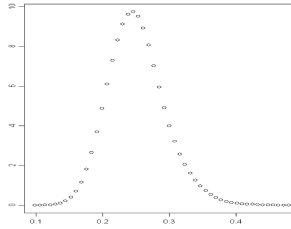


FIGURE 5

$$m = 0,2512; s^2 = 0,0013 \left( n = 100, \frac{\sigma_X^2}{\sigma_Y^2} = 0,25 \right)$$

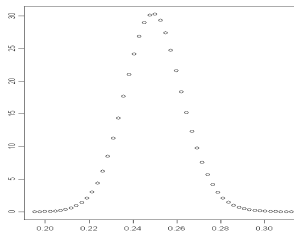


FIGURE 6

$$m = 0,2497; s^2 = 0,0001 \left( n = 1000, \frac{\sigma_X^2}{\sigma_Y^2} = 0,25 \right)$$

### 6.1.2. Loi exponentielle

On considère 500 échantillons, de tailles respectives  $n = 50$  et  $100$ , déduits de la loi exponentielle, ou encore loi de Gumbel,  $\mathcal{G}m(\alpha)$  de paramètre  $\alpha = 0,5$  dont la fonction de répartition  $F_{X,Y}(x,y)$  a pour expression :

$$F_{X,Y}(x,y) = 1 - e^{-x} - e^{-y} + e^{-x-y-0,5xy} \text{ où } x > 0 \text{ et } y > 0$$

On a donc ici  $\mu = 1$  puisque les variables  $X$  et  $Y$  ont même loi. Les distributions empiriques du rapport  $\widehat{R}^2_{(0,75, 0,25),(0,75, 0,25)}$  sont illustrées aux figures 7 et 8 ci-dessous :

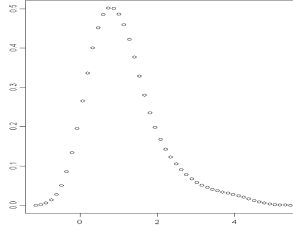


FIGURE 7

$$m = 1,1709; s^2 = 0,7206 \quad (n = 50, \mu = 1, \sigma_{50}^2 = 0,4959)$$

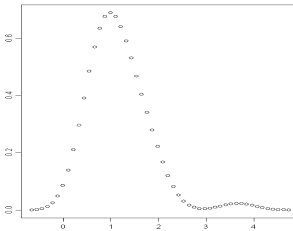


FIGURE 8

$$m = 1,1525; s^2 = 0,3169 \quad (n = 100, \mu = 1, \sigma_{100}^2 = 0,2479)$$

À titre de comparaison, les figures 9 et 10 suivantes illustrent la distribution empirique du rapport  $\frac{\widehat{\sigma}_X^2}{\widehat{\sigma}_Y^2}$ , où  $m$  et  $s^2$  désignent respectivement les valeurs empiriques de la moyenne et de la variance de celui-ci.

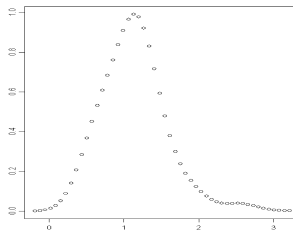


FIGURE 9

$$m = 1,1535; s^2 = 0,5670 \quad \left( n = 50, \frac{\sigma_X^2}{\sigma_Y^2} = 1 \right)$$

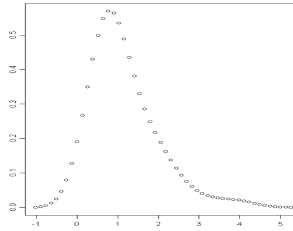


FIGURE 10

$$m = 1,1349; s^2 = 0,1612 \left( n = 100, \frac{\sigma_X^2}{\sigma_Y^2} = 1 \right)$$

6.1.3. Loi de Cauchy

On considère enfin 200 échantillons de tailles respectives 20, 30, 50, 100 et 1000, tirés de la loi de Cauchy bidimensionnelle dont la densité de probabilité est donnée par :

$$f(x, y) = \frac{1}{2\pi (1 + x^2 + y^2)^{3/2}}$$

On sait alors dans ce cas (voir par exemple Devroye [5]) que les lois marginales sont des lois de Cauchy et que les lois conditionnelles d'une des variables étant donnée l'autre sont des lois de Student à 2 degrés de liberté ce qui permet de réaliser aisément les simulations voulues. On notera que l'on a encore ici  $\mu = 1$  alors que le rapport des variances n'existe pas, puisque les lois marginales sont des lois de Cauchy.

Les distributions empiriques du rapport  $\widehat{R}^2_{(0,75, 0,25),(0,75, 0,25)}$  sont illustrées aux figures 11 à 15 ci-dessous :

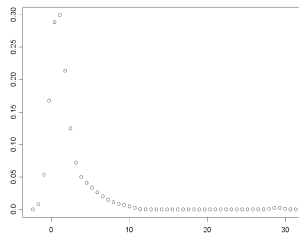


FIGURE 11

$$m = 1,8972; s^2 = 7,3409 \left( n = 20, \mu = 1, \sigma_{20}^2 = 0,6581 \right)$$



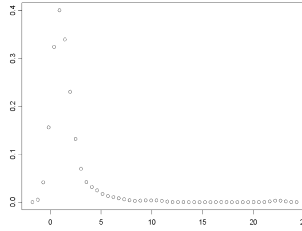


FIGURE 12

$$m = 1,6256; s^2 = 4,4121 \quad (n = 30, \mu = 1, \sigma_{30}^2 = 0,4387)$$

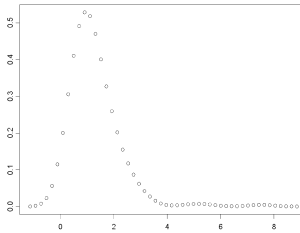


FIGURE 13

$$m = 1,2736; s^2 = 0,7772 \quad (n = 50, \mu = 1, \sigma_{50}^2 = 0,2632)$$

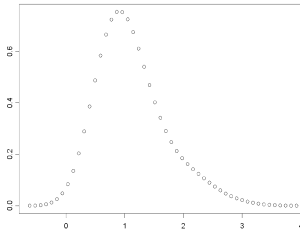


FIGURE 14

$$m = 1,0860; s^2 = 0,2683 \quad (n = 100, \mu = 1, \sigma_{100}^2 = 0,1316)$$

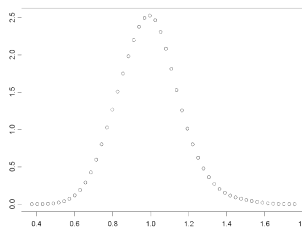


FIGURE 15

$$m = 0,9918; s^2 = 0,0203 \quad (n = 1000, \mu = 1, \sigma_{1000}^2 = 0,0132)$$

**6.2. Tests de normalité**

À titre indicatif, on donne, dans le cas des lois  $N(\nu, \Sigma)$  et  $\mathcal{G}m(0, 5)$  ci-dessus, les tailles des échantillons à partir desquelles on accepte, avec un risque de première espèce de 0,05 l'hypothèse  $\mathcal{H}_0$  suivant laquelle la statistique :

$$\widehat{R}^2_{(p,1-p),(p,1-p)} = \frac{(\widehat{\xi}_p - \widehat{\xi}_{1-p})^2}{(\widehat{\eta}_p - \widehat{\eta}_{1-p})^2}$$

suit, pour  $p = 0,75$ , une loi normale.

Le test de normalité utilisé dans le cas présent, est le test de *Shapiro-Wilk*. Les simulations, effectuées également à l'aide du logiciel S-Plus, comportent dans chacun des cas, 500 répétitions et les tailles des échantillons sont, à chaque pas, augmentées de 10 unités. Les résultats qui découlent de ces simulations sont présentés dans le tableau ci-dessous où  $W$  désigne la valeur empirique de la statistique du test de *Shapiro-Wilk* et où  $d_0$  et  $d_1$  correspondent respectivement aux décisions d'accepter ou de rejeter l'hypothèse  $\mathcal{H}_0$ .

TABLEAU 1

Loi $N(0, \Sigma)$			Loi $\mathcal{G}m(0, 5)$		
Taille $n$	$W$	Décision	Taille $n$	$W$	Décision
30	0,8199	$d_1$	30	0,7726	$d_1$
40	0,9133	"	40	0,8284	"
50	0,9304	"	50	0,8373	"
60	0,9378	"	60	0,8886	"
70	0,9387	"	70	0,9085	"
80	0,9430	"	80	0,9146	"
90	0,9568	"	90	0,9244	"
100	0,9610	"	100	0,9341	"
⋮	⋮	⋮	⋮	⋮	⋮
350	0,9994	$d_0$	700	0,9992	$d_0$
360	0,9996	"	710	0,9994	"
370	0,9998	"	720	0,9996	"
⋮	⋮	⋮	⋮	⋮	⋮

Ces résultats appellent les commentaires suivants :

- 1) De fait, dans le cadre de la présente simulation, l'hypothèse de normalité a été acceptée pour la première fois pour des tailles d'échantillons respectivement de 347, dans le cas de la loi normale  $N(0, \Sigma)$  et de 702, dans le cas de la loi

de Gumbel  $\mathcal{G}m(0, 5)$ . C'est uniquement pour une question de présentation que ces deux nombres ont été arrondis puisque le pas avait été fixé à 10 unités. Par ailleurs, pour de nombreuses tailles d'échantillons supérieures à ces deux valeurs, on a constaté empiriquement que la valeur de la statistique du test de *Shapiro-Wilk* était supérieure systématiquement à la valeur critique du test, pour un risque de première espèce constant et égal à 0,05, entraînant ainsi toujours la même décision à savoir  $d_0$ .

- 2) Il peut sembler, à première vue, que les tailles des échantillons à partir desquelles l'approximation normale est valide (et ceci en particulier dans le cas de la loi  $\mathcal{G}m(0, 5)$ ), soient quelque peu élevées. Or, pour une valeur de  $p$  donnée, la statistique  $\widehat{R}^2_{(p,1-p),(p,1-p)}$  prend en compte uniquement les quatre valeurs numériques  $\widehat{\xi}_p, \widehat{\xi}_{1-p}, \widehat{\eta}_p$  et  $\widehat{\eta}_{1-p}$  et non les  $2n$  valeurs numériques des observations (si ce n'est pour déterminer les quantiles empiriques). Dans ce contexte, les ordres de grandeur des tailles déterminées ci-dessus empiriquement semblent tout à fait raisonnables.
- 3) Lorsque les tailles des échantillons considérés dépassent, comme c'est souvent le cas, la centaine, il peut être plus judicieux de choisir une valeur de  $p$  égale à 0,9 voire à 0,95.

## 7. Conclusions

Afin de généraliser la notion d'efficacité relative ou de risque relatif, tant en ce qui concerne les procédures statistiques, que les applications que l'on rencontre, entre autres, dans les domaines de la finance, de l'économie ou des assurances, nous proposons de considérer le rapport

$$\widehat{R}^2_{(p,q),(r,s)} = \frac{(\widehat{\xi}_p - \widehat{\xi}_q)^2}{(\widehat{\eta}_r - \widehat{\eta}_s)^2} \quad \text{où } p > q; r > s$$

des carrés des étendues interquantiles empiriques des composantes d'une loi bidimensionnelle. Nous avons montré que ce rapport est asymptotiquement normal et qu'il constitue un estimateur asymptotiquement sans biais et convergent du rapport théorique

$$R^2_{(p,q),(r,s)} = \frac{(\xi_p - \xi_q)^2}{(\eta_r - \eta_s)^2}$$

ce que confirment les simulations faites dans le cas de trois lois différentes. En particulier, dans le cas de la loi de Cauchy bidimensionnelle, il semble que cet estimateur se comporte bien dès que la taille  $n$  de l'échantillon est suffisamment grande ( $n \geq 100$ ). De surcroît, dans le cas où les lois admettent des variances finies, le rapport

$\widehat{R}^2_{(p,q),(r,s)}$  se compare fort bien au rapport usuel  $\frac{\sigma_X^2}{\sigma_Y^2}$  même si, comparativement à

celui-ci, il demeure moins précis, inconvénient d'ailleurs qui s'atténue lorsque la taille de l'échantillon devient de plus en plus grande. Par ailleurs le rapport  $\widehat{R}^2_{(p,q),(r,s)}$  est particulièrement facile à calculer, puisque il nécessite seulement, afin de déterminer les quantiles empiriques, d'ordonner les observations pour chacune des composantes de l'échantillon, ce que l'on peut très facilement réaliser à l'aide d'un programme élémentaire.

De plus, s'exprimant sous la forme d'un rapport d'indices de dispersion de deux variables aléatoires, analogue en ce sens au rapport des variances, l'expression  $\widehat{R}^2_{(p,q),(r,s)}$  s'apparente à un indice d'efficacité ou de risque relatif et s'avère particulièrement utile dans le cas où, si l'une au moins des variables aléatoires considérées n'admet pas de variance, l'indice d'efficacité relative, tel que défini habituellement, n'existe pas. On dispose donc ainsi d'un outil permettant de prolonger cette notion lorsque le rapport usuel des variances ne peut être utilisé.

Enfin, il est clair que le rapport  $\widehat{R}^2_{(p,q),(r,s)}$  est beaucoup moins sensible à la présence de valeurs aberrantes dans les observations que ne l'est le rapport des variances  $\frac{\sigma_X^2}{\sigma_Y^2}$  et nous nous proposons, à ce sujet, d'aborder prochainement tout ce qui concerne l'étude de la robustesse de celui-ci.

**Annexe**

**1) Loi du rapport**  $\frac{(X_{(u)} - X_{(t)})^2}{(Y_{(w)} - Y_{(v)})^2}$

Nous nous contenterons de donner ici l'expression générale de la densité de probabilité de la variable aléatoire  $\frac{(X_{(u)} - X_{(t)})^2}{(Y_{(w)} - Y_{(v)})^2}$  (cette dernière étant, dans le cadre retenu, définie presque sûrement). Considérons le changement de variables suivant :

$$S_1 = X_{(t)}, S_2 = X_{(u)} - X_{(t)}, S_3 = Y_{(v)}, S_4 = Y_{(w)} - Y_{(v)}$$

La valeur absolue du Jacobien de la transformation étant égale à 1, il s'ensuit que la densité de probabilité  $f_{S_1, S_2, S_3, S_4}(s_1, s_2, s_3, s_4)$  du vecteur aléatoire

$${}^t(S_1, S_2, S_3, S_4)$$

a pour expression :

$$f_{S_1, S_2, S_3, S_4}(s_1, s_2, s_3, s_4) = f_{(t,u),(v,w)}(s_1, s_1 + s_2, s_3, s_3 + s_4) \times \mathbb{I}_{\{\mathbb{R} \times \mathbb{R}^+ \times \mathbb{R} \times \mathbb{R}^+\}}(s_1, s_2, s_3, s_4)$$

où l'expression  $f_{(t,u),(v,w)}$ , telle qu'introduite dans [2], désigne la densité de probabilité du vecteur aléatoire  $(\xi_t, \xi_u, \eta_v, \eta_w)$ . En conséquence, la densité de probabilité  $f_{S_2, S_4}(s_2, s_4)$  du couple

$$(X_{(u)} - X_{(t)}, Y_{(w)} - Y_{(v)})$$

est donnée par :

$$f_{S_2, S_4}(s_2, s_4) = \int_{\mathbb{R}} \int_{\mathbb{R}} f_{(t,u),(v,w)}(s_1, s_1 + s_2, s_3, s_3 + s_4) ds_1 ds_3 \\ \times \mathbb{I}_{\{\mathbb{R}^+ \times \mathbb{R}^+\}}(s_2, s_4)$$

On en déduit alors aisément la densité de probabilité  $f_R(r)$  du rapport  $R = \frac{(X_{(u)} - X_{(t)})}{(Y_{(w)} - Y_{(v)})}$  en effectuant le changement de variables :

$$S = X_{(u)} - X_{(t)}, R = \frac{(X_{(u)} - X_{(t)})}{(Y_{(w)} - Y_{(v)})}$$

On obtient alors, en notant que la valeur absolue du Jacobien est donnée par  $\frac{S}{R^2}$  :

$$f_R(r) = \left[ \int_{\mathbb{R}^+} f_{S_2, S_4}\left(s, \frac{s}{r}\right) \frac{s}{r^2} ds \right] \mathbb{I}_{\{\mathbb{R}^+ \setminus \{0\}\}}(r)$$

ce qui entraîne que la loi  $f_Z(z)$  de  $Z = R^2$  est donnée par :

$$f_Z(z) = \left[ \int_{\mathbb{R}^+} f_{S_2, S_4}\left(s, \frac{s}{\sqrt{z}}\right) \frac{s}{z} ds \right] \frac{1}{2\sqrt{z}} \mathbb{I}_{\{\mathbb{R}^+ \setminus \{0\}\}}(z)$$

## 2) Loi du rapport $\widehat{R}_{(p,q),(r,s)}^2$

Soit  $^t(X_i, Y_i)$   $i = 1, 2, \dots, n$  un échantillon de taille  $n$  issu du modèle empirique  $(\mathbb{R}^2, \mathcal{B}_{\mathbb{R}^2}, \{\mathbb{P}_\theta\}, \mu)^{\otimes n}$ . Les quantiles empiriques  $\widehat{\xi}_p, \widehat{\xi}_q, \widehat{\eta}_r$  et  $\widehat{\eta}_s$  étant donnés respectivement par  $X_{u:n}, X_{t:n}, Y_{w:n}$  et  $Y_{v:n}$ , il s'ensuit que la loi de  $(\widehat{\xi}_p, \widehat{\xi}_q, \widehat{\eta}_r, \widehat{\eta}_s)$  se déduit aisément des résultats précédents et qu'il en est de même pour la loi du

$$\text{rapport } \widehat{R}_{(p,q),(r,s)}^2 = \frac{(\widehat{\xi}_p - \widehat{\xi}_q)^2}{(\widehat{\eta}_r - \widehat{\eta}_s)^2}.$$

### 3) Expression de la variance de la loi asymptotique du rapport des étendues interquantiles empiriques

On montre dans [3] que :

$$\frac{\widehat{\xi}_p - \widehat{\xi}_q}{\widehat{\eta}_r - \widehat{\eta}_s} \rightsquigarrow AN \left( h(p, q, r, s), \frac{1}{n} {}^t Gradh|_{\mu} \Sigma Gradh|_{\mu} \right)$$

où le symbole  $AN$  désigne l'expression « asymptotiquement normal » et où  $h(p, q, r, s)$ ,  $\mu$ ,  $\Sigma$  et  $Gradh|_{\mu}$  sont donnés respectivement par :

$$h(p, q, r, s) = \frac{F_X^{-1}(p) - F_X^{-1}(q)}{F_Y^{-1}(r) - F_Y^{-1}(s)} = \frac{\xi_p - \xi_q}{\eta_r - \eta_s} ; t\mu = {}^t(p, q, r, s)$$

$$\Sigma = \begin{bmatrix} p(1-p) & q(1-p) & F_{Z_1, Z_2}(p, r) - pr & F_{Z_1, Z_2}(p, s) - ps \\ q(1-p) & q(1-q) & F_{Z_1, Z_2}(q, r) - qr & F_{Z_1, Z_2}(q, s) - qs \\ F_{Z_1, Z_2}(p, r) - pr & F_{Z_1, Z_2}(q, r) - qr & r(1-r) & s(1-r) \\ F_{Z_1, Z_2}(p, s) - ps & F_{Z_1, Z_2}(q, s) - qs & s(1-r) & s(1-s) \end{bmatrix} ;$$

$${}^t Gradh|_{\mu} = {}^t \left( \frac{1}{f_X(\xi_p)(\eta_r - \eta_s)}, \frac{-1}{f_X(\xi_q)(\eta_r - \eta_s)}, \frac{-(\xi_p - \xi_q)}{f_Y(\eta_r)(\eta_r - \eta_s)^2}, \frac{(\xi_p - \xi_q)}{f_Y(\eta_s)(\eta_r - \eta_s)^2} \right)$$

avec  $F_{Z_1, Z_2}(z_1, z_2) = F_{X, Y}(F_X^{-1}(z_1), F_Y^{-1}(z_2))$

On en déduit alors que :

$$\sqrt{n} \left( \frac{\widehat{\xi}_p - \widehat{\xi}_q}{\widehat{\eta}_r - \widehat{\eta}_s} - \frac{\xi_p - \xi_q}{\eta_r - \eta_s} \right) \xrightarrow{L} N(0, \sigma^2)$$

où  $\sigma^2 = {}^t Gradh|_{\mu} \Sigma Gradh|_{\mu}$  a pour expression :

$$\begin{aligned} & \frac{1}{(\eta_r - \eta_s)^2} \left( \frac{p(1-p)}{f_X^2(\xi_p)} + \frac{q(1-q)}{f_X^2(\xi_q)} - \frac{2q(1-p)}{f_X(\xi_p) f_X(\xi_q)} \right) \\ & + \frac{2(\xi_p - \xi_q)}{(\eta_r - \eta_s)^3} \left( \frac{(F_{X, Y}(\xi_p, \eta_s) - ps)}{f_X(\xi_p) f_Y(\eta_s)} + \frac{(F_{X, Y}(\xi_q, \eta_r) - qr)}{f_X(\xi_q) f_Y(\eta_r)} \right. \\ & \left. - \frac{(F_{X, Y}(\xi_p, \eta_r) - pr)}{f_X(\xi_p) f_Y(\eta_r)} - \frac{(F_{X, Y}(\xi_q, \eta_s) - qs)}{f_X(\xi_q) f_Y(\eta_s)} \right) \\ & + \frac{(\xi_p - \xi_q)^2}{(\eta_r - \eta_s)^4} \left( \frac{r(1-r)}{f_Y^2(\eta_r)} + \frac{s(1-s)}{f_Y^2(\eta_s)} - \frac{2s(1-r)}{f_Y(\eta_r) f_Y(\eta_s)} \right) \end{aligned}$$

Dans le cas où les variables aléatoires  $X$  et  $Y$  sont indépendantes, tous les termes faisant intervenir la fonction de répartition conjointe s'annulent et l'expression de  $\sigma^2$  devient :

$$\frac{1}{(\eta_r - \eta_s)^2} \left( \frac{p(1-p)}{f_X^2(\xi_p)} + \frac{q(1-q)}{f_X^2(\xi_q)} - \frac{2q(1-p)}{f_X(\xi_p) f_X(\xi_q)} \right) + \frac{(\xi_p - \xi_q)^2}{(\eta_r - \eta_s)^4} \left( \frac{r(1-r)}{f_Y^2(\eta_r)} + \frac{s(1-s)}{f_Y^2(\eta_s)} - \frac{2s(1-r)}{f_Y(\eta_r) f_Y(\eta_s)} \right)$$

Si de plus l'on pose  $r = p$  et  $s = q = 1 - p$  (avec  $p > 0,5$ ) et si de surcroît les lois de  $X$  et de  $Y$  sont symétriques, il est aisé de vérifier que l'expression précédente se réduit à :

$$\frac{2(1-p)(2p-1)}{(\eta_p - \eta_{1-p})^2} \left[ \frac{1}{f_X^2(\xi_p)} + R_{(p,1-p),(p,1-p)}^2 \frac{1}{f_Y^2(\eta_p)} \right]$$

### Références

- [1] BAHADUR R.R. (1966), *A note on quantiles in large samples*, Ann. Math. Statist., **37**, 577-580.
- [2] COLIN B. et MELLAL N. (2002), *Rapport des quantiles empiriques des composantes d'une loi bidimensionnelle*, Revue de Statistique Appliquée, **L(4)**, 55-80.
- [3] COLIN B. et MELLAL N. (2003), *Rapport des étendues interquantiles empiriques des composantes d'une loi bidimensionnelle*, Pub.Inst.Stat.Univ.Paris XXXXVII, fasc.1-2, 49-73.
- [4] DAVID H.A. (1981), *Order statistics*, John Wiley & Sons.
- [5] DEVROYE L. (1986), *Non-Uniform Random Variate Generation*, Springer-Verlag.
- [6] EICKER F. (1966), *On the asymptotic representation of sample quantiles*, (Abstract) Ann. Math. Statist., **37**, 1424.
- [7] FELDMAN R.S. and TAQQU M.S. (2002), *A Practical Guide to Heavy Tails. Statistical Techniques and Applications*, Birkhäuser.
- [8] FRÉCHET M. (1951), *Sur les tableaux de corrélation dont les marges sont données*, Ann. Univ. Lyon, Sér.3, **14**, 53-77.
- [9] KIMELDORF G. and SAMPSON A.R. (1975), *Uniform representations of bivariate distributions*, Com. Statist. **4**, 617-627.
- [10] LECOUTRE J.P. et TASSI P. (1987), *Statistique non paramétrique et Robustesse*, Economica.

- [11] MARITZ J.S. (1991), *Estimating the covariance matrix of bivariate medians*, Statist. Prob. Lett. **12**, n° 4 (1991) 305-309.
- [12] MELLAL N. (Mars 2000), *Rapport des quantiles des composantes d'une loi bidimensionnelle*, Thèse de doctorat, Université de Sherbrooke, Canada.
- [13] SERFLING R.J. (1980), *Approximation theorems of mathematical statistics*, John Wiley & Sons.
- [14] SKLAR A. (1959), *Fonctions de répartition à n dimensions et leurs marges*, Publ. Inst. Statist. Univ. Paris, **8**, 229-32.
- [15] SRIVASTAVA M.S. (1967), *Asymptotic independence of certain statistics connected with the extreme order statistics in a bivariate population*, Sankya : The Indian Journal of Statistics : **Series A**, vol.**29**, **2**, 175-182.



