



Le quotient de deux variances corrélées, sa distribution et son test

Louis Laurencelle^a,

^aUniversité du Québec à Trois-Rivières

Abstract ■ The joint sampling distribution of two correlated variances, i.e. variances stemming from a bivariate normal distribution or from two normal ρ -correlated distributions, is hardly known and used, by contrast with the distribution of F , the quotient of two independent, zero-correlated variances. The distribution of F_ρ , the quotient of two correlated variances, established by Bose (1935) and Finney (1938), is given along with its main characteristics, to which is added a handy F_ρ to F transformation. Finally, data based on Monte Carlo simulations document and compare the accuracy and power of two approximate tests of the difference between two correlated sample variances.

Keywords ■ Correlated variances, t-test on correlated variances.

Acting Editor ■ Denis Cousineau (Université d'Ottawa)

louis.laurencelle@gmail.com

LL: 0000-0003-3448-2872

10.20982/tqmp.12.2.r131

Introduction

Dans une population normale bivariable à cinq paramètres, $N(\mu_X, \mu_Y, \sigma_X^2, \sigma_Y^2, \rho)$, l'échantillon $\{(X_1, Y_1), (X_2, Y_2), \dots, (X_n, Y_n)\}$ donne lieu aux estimateurs correspondants $\bar{X}, \bar{Y}, s_X^2, s_Y^2$ et r . Soit le quotient :

$$q = \frac{s_X^2/\sigma_X^2}{s_Y^2/\sigma_Y^2} \tag{1a}$$

ou :

$$q = \frac{s_X^2}{s_Y^2} \text{ si } \sigma_X^2 = \sigma_Y^2. \tag{1b}$$

La fonction de densité de ce quotient est :

$$h_{F_\rho}(q) = \frac{\Gamma(n-1)}{[\Gamma(\frac{n-1}{2})]^2} \times \frac{(1-\rho^2)^{\frac{n-1}{2}} q^{\frac{n-3}{2}}}{(1+q)^{n-1}} \times \left\{ 1 - \frac{4\rho^2 q}{(1+q)^2} \right\}^{-n/2}, \tag{2}$$

définissant la distribution échantillonnale $F_\rho(q)$ pour deux variances corrélées (Bose 1935; Finney 1938). Pour $\rho = 0$, cette fonction $h(q)$ se ramène à la fonction de densité de

la loi F , dite de Fisher-Snedecor (Kendall et Stuart 1977; Laurencelle et Dupuis 2000).

La fonction de répartition de la loi F_ρ n'est pas connue. Cependant, Finney (1938) propose de transformer la variable q en une variable de distribution Bêta, qu'on transforme à son tour en une variable F régulière, c'est-à-dire de loi $F_{\rho=0}$, par :

$$F = \frac{\sqrt{(1+q)^2 - 4\rho^2 q} + q - 1}{\sqrt{(1+q)^2 - 4\rho^2 q} - q + 1} \sim F(n-1, n-1); \tag{3}$$

les tables d'intégrale ou de valeurs critiques de cette distribution sont largement disponibles. Le tableau 1 fournit quelques centiles illustratifs de la loi F_ρ , obtenus par intégration numérique directe de (2).

Caractéristiques de la distribution F_ρ

Nous avons amassé quelques informations permettant d'indiquer les principales caractéristiques de la distribution F_ρ . D'abord, les variances s_X^2 et s_Y^2 étant calculées dans un échantillon tiré d'une loi bivariable $N(\mu_X, \mu_Y, \sigma_X^2, \sigma_Y^2, \rho)$, nous avons (voir Appendice) :

$$\rho(s_X^2, s_Y^2) = \rho^2. \tag{4}$$

1. La formule d'espérance $E(q)$ a d'abord été conjecturée par Denis Allaire, Ph.D., Université de Sherbrooke.



Quant au quotient de variances $q = s_X^2/s_Y^2$, nous trouvons, par induction numérique :¹

$$E(q) = \mu(q) = \frac{n-1-2\rho^2}{n-3} \quad (5)$$

et :

$$\begin{aligned} \text{var}(q) &= \sigma^2(q) \\ &= \frac{[4n^2 - 12n + 8 - 4\rho^2(5n - 13)](1 - \rho^2)}{(n-3)^2(n-5)}. \end{aligned} \quad (6)$$

Une expression approximative pour la variance est :

$$\sigma^2(q) \approx \frac{4(n-1-5\rho^2)(1-\rho^2)}{(n-4,5)^2}. \quad (7)$$

La médiane, obligatoirement, est :

$$Md(q) = 1. \quad (8)$$

Quant au mode, nous proposons l'approximation :

$$Mo(q) \approx \frac{n-3+2\rho^2}{n+1-2\rho^2}. \quad (9)$$

Toutes ces caractéristiques conviennent aussi à la variable de loi F avec $\nu_1 = \nu_2 = n - 1$ (et pour laquelle $\rho = 0$), le mode (9) devenant alors exact.

Le test de différence entre deux variances corrélées

Un chercheur mesure un groupe de personnes deux fois, obtenant d'une part la série $\{X_1, X_2, \dots, X_n\}$ et d'autre part la série $\{Y_1, Y_2, \dots, Y_n\}$. Les variances calculées sont σ_X^2 et σ_Y^2 , et la corrélation entre les séries, $r_{X,Y}$. Pour disposer de l'hypothèse nulle de non-différence entre les variances, soit :

$$H_0 : \sigma_X^2 = \sigma_Y^2 \quad (10a)$$

ou, équivalamment :

$$H_0 : \mu(s_X^2) = \mu(s_Y^2) = \sigma^2, \quad (10b)$$

le test correct consisterait à calculer $q = s_X^2/s_Y^2$, puis à comparer ce résultat à un centile extrême de F_ρ ou, après la transformation (3), à un centile de F . Cependant, pour appliquer ce procédé, il faut connaître ρ , la valeur paramétrique de corrélation dont dépendent la distribution de F_ρ , voire la transformation (3). Une option pratique consiste alors à substituer $r_{X,Y}$ à ρ , selon $\hat{\rho} = r_{X,Y}$ (voir cependant Laurencelle 2000a) et à considérer qu'on a, en

approximation, une variable de distribution F_r . L'autre option, la seule qu'on trouve dans la littérature (Walker et Lev 1953; Guilford 1965; Hinkle, Wiersma et Jurs 1988; Laurencelle et Dupuis 2000) utilise la formule suivante :

$$\frac{s_X^2 - s_Y^2}{2s_X s_Y \sqrt{\frac{1-r_{X,Y}^2}{n-2}}} \sim t_{n-2}; \quad (11)$$

il s'agit donc d'un test t , approximativement distribué comme un t de Student avec $\nu = n - 1$. Comparativement au test correct sur q (de distribution F_ρ) basé sur ρ , qu'est-ce que valent le test sur \hat{q} (de distribution approximative F_r) basé sur $r_{X,Y}$ et le test t (de distribution approximative t_{n-1}), lui aussi basé sur $r_{X,Y}$?

Un exemple. Supposons que, d'une série bivariée de $n = 11$ observations, on obtienne $s_X^2 = 48$, $s_Y^2 = 15$ et $r_{X,Y} = 0,612$. La valeur de ρ n'est pas connue, comme c'est le cas habituel; toutefois, dans une expérience où les mêmes personnes ou objets sont évalués deux fois, il est raisonnable de supposer que $\rho > 0$. Pour fins d'illustration uniquement, nous supposons ici que $\rho = 0,500$.² Nous effectuons un test bilatéral au seuil de signification de 5 %.

Le test correct implique $q = s_X^2/s_Y^2 = 48/15 = 3,2$ dans la distribution $F_{0,500}$ avec $\nu = n - 1 = 10$. Au tableau 1, vis-à-vis de $\nu = 10$, $P = 0,975$ et $\rho = 0,5$, nous trouvons $q = 3,174$; la valeur observée de 3,2 serait donc déclarée significative, les deux variances parentes (σ_X^2, σ_Y^2) différant l'une de l'autre. À défaut d'une table de valeurs critiques ou d'une évaluation de l'intégrale de la fonction (2), la transformation (3), inspirée de Finney (1938), nous fournit $F = 3,750$, avec $\nu_1 = \nu_2 = n - 1 = 10$. Or, dans une table du F (p. ex. Laurencelle et Dupuis 2000), on trouve $F_{10,10[0,975]} = 3,717$, la valeur observée de 3,750 apparaissant évidemment significative.

Le test approximatif $\hat{q} = 3,2$ utilise cette fois l'estimation $\hat{\rho} = r_{X,Y} = 0,612$. Avec $\nu = n - 1 = 10$ et $P = 0,975$, il est apparent que la valeur observée excède la valeur critique³; la conversion en $F = 4,179$ par la transformation (3) confirme cette décision.

Le test t , en application de la formule (11), fournit $t = 2,333$; le centile 97,5 du t_9 égale 2,262. La valeur calculée est donc, encore une fois, significative selon ce critère approximatif.

Comparaisons Monte Carlo. Trois courtes études Monte Carlo viennent éclairer la question. Dans une première étude, basée sur 10000 itérations, nous avons simulé une population normale bivariée respectant l'égalité (9a), i.e. $\sigma_X^2 = \sigma_Y^2$, ce pour $\rho = 0, \frac{1}{2}, \sqrt{\frac{1}{2}}$ et $\sqrt{\frac{3}{4}}$, et en formant des échantillons de $n = 5, 10, 25, 50$ ou 100 données paires.

2. Cette valeur putative de ρ correspond en fait à l'estimateur modal $\hat{\rho}(n, r)$ pour $n = 11$ et $r = 0,612$, tel que $r = 0,612$ est le mode de la distribution échantillonnale de r pour $\rho = 0,500$ à $n = 11$ (voir Laurencelle, 2000a).

3. L'évaluation numérique de l'intégrale de (2) pour $n = 11, \rho = 0,612$ et $q = 3,2$ indique $P = 0,9832$.



Tableau 1 ■ Quelques centiles de la distribution F_ρ selon différentes valeurs de $\nu = n - 1$, en fonction de ρ

	ρ	$P = 0,75$	$P = 0,95$	$P = 0,975$	$P = 0,99$	$P = 0,995$
$\nu = 5(n = 6)$	0	1,895	5,050	7,146	10,967	14,941
	0,5	1,743	4,198	5,792	8,679	11,671
	0,9	1,326	2,153	2,623	3,429	4,235
	0,99	1,096	1,289	1,381	1,524	1,654
$\nu = 10(n = 11)$	0	1,551	2,978	3,717	4,849	5,847
	0,5	1,464	2,601	3,174	4,044	4,805
	0,9	1,212	1,640	1,831	2,106	2,335
	0,99	1,064	1,175	1,220	1,279	1,326
$\nu = 15(n = 16)$	0	1,426	2,403	2,862	3,522	4,070
	0,5	1,361	2,149	2,510	3,024	3,446
	0,9	1,168	1,480	1,608	1,782	1,919
	0,99	1,052	1,136	1,168	1,208	1,239

Dans chaque itération, les variances s_X^2 et s_Y^2 étaient calculées de même que la corrélation $r_{X,Y}$, le quotient de variances q , l'intégrale exacte selon F_ρ , l'intégrale approchée selon F_r et le test t selon (11), avec son intégrale de probabilité. Les tests selon F_r et t , tous deux basés sur la valeur échantillonnale $r_{X,Y}$, ayant un comportement quasi identique, nous nous contenterons de comparer le test F_ρ au test t approximatif.

Le tableau 2 résume l'essentiel des résultats. L'approximation t dérape d'autant plus que ρ s'écarte de 0 et elle s'améliore quand n croît. L'erreur, qu'il s'agisse de l'erreur absolue maximale ou moyenne, tend donc vers zéro pour n croissant mais cette progression est d'autant plus lente que ρ est proche de 1.

Dans la seconde étude Monte Carlo, utilisant 20000 itérations de même type, nous avons noté les cas dans lesquels la paire (s_X^2, s_Y^2) atteignait la significativité statistique, au seuil de 5% unilatéral, selon l'un ou l'autre des trois critères : quotient q évalué selon F_ρ , quotient q évalué selon F_r , test t (basé sur $r_{X,Y}$). Globalement, les comportements des deux tests approximatifs, selon F_r et t , se sont encore montrés presque identiques : la comparaison entre le test correct (selon F_ρ) et l'approximation t suffit. Le tableau 3 illustre les résultats de cette étude.

On constate, encore une fois, que la performance du test t est affectée par l'augmentation de ρ , les conclusions du t perdant alors de la concordance avec celles du test F_ρ . Quand $\rho = 0$, l'approximation t retrouve au moins 90% des valeurs extrêmes du F_ρ ("Sensibilité") dès $n = 10$, et plus de 90 % de ses déclarations significatives sont confirmées par le F_ρ ("Spécificité"). Pour $\rho = \sqrt{\frac{3}{4}}$, cette qualité d'approximation n'est pas encore atteinte quand $n = 100$.

Dans la troisième étude, nous avons estimé la puissance des trois tests sous les mêmes conditions, mais en variant cette fois le rapport des variances paramétriques,

soit en imposant aux variables X de chaque échantillon une variance $\sigma_X^2 = 1, 3/2, 2, 4, 6$ et 8, tout en gardant $\sigma_Y^2 = 1$. L'hypothèse nulle (10a) est rejetée unilatéralement. Les tableaux 4a à 4d présentent les résultats se rapportant au seuil de signification unilatéral $\alpha = 0,05$, les données d'autres seuils se montrant en tout point similaires. Chaque proportion de rejets est basée sur 20000 échantillons.

La puissance des trois tests augmente en fonction directe du quotient σ_X^2/σ_Y^2 , de la taille d'échantillon n et de la corrélation ρ (ou ρ^2 , dénotant la corrélation entre s_X^2 et s_Y^2). L'augmentation de puissance selon n et ρ est surtout attribuable à un resserrement de la distribution échantillonnale du quotient $q = s_X^2/s_Y^2$, comme le suggère la formule de variance $\sigma^2(q)$ (voir (7)).

Sous H_0 , soit $\sigma_X^2/\sigma_Y^2 = 1$, les tests F_ρ et t s'avèrent exacts, en rejetant à peu près 5 % des échantillons, alors que le test F_r présente un taux de rejets excessif : cet excès du test F_r , par ailleurs peu influencé par ρ , diminue avec n . On constate aussi que le test F_r reste généreux, voire trop généreux comparativement au test F_ρ , et qu'il serait donc plus 'puissant' que le test correct lorsque l'hypothèse est fausse.

Le test t , exact sous H_0 , montre une puissance qui approche de façon surprenante celle du test F_ρ . En fait, s'il apparaît quelque peu moins puissant que F_ρ , avec des quotients de puissance variant de 0,87 à 1,02 pour $n = 5$, ce déficit diminue pour n croissant, les quotients de puissance allant de 0,96 à 1,02 pour $n = 10$, et de 0,99 à 1,02 pour $n = 25$.

En conclusion, il semble donc raisonnable de formuler les recommandations suivantes quant aux différents tests de l'égalité de deux variances corrélées.



Tableau 2 ■ Différence absolue maximale ($|\varepsilon|_{max}$) et moyenne ($|\varepsilon|_{moy}$) entre l'intégrale $P(F_\rho)$ du test correct du quotient de variances q et l'intégrale $P(t)$ du test t approximatif (11) dans la distribution de Student, selon ρ et n (10000 échantillons Monte Carlo)

n	$ \varepsilon _{max}$				$ \varepsilon _{moy}$			
	$\rho = 0$	$\frac{1}{2}$	$\sqrt{\frac{1}{2}}$	$\sqrt{\frac{3}{4}}$	$\rho = 0$	$\frac{1}{2}$	$\sqrt{\frac{1}{2}}$	$\sqrt{\frac{3}{4}}$
5	,314	,363	,408	,380	,025	,038	,049	,059
10	,180	,242	,291	,291	,009	,023	,031	,037
25	,052	,111	,146	,152	,003	,013	,018	,023
50	,050	,088	,100	,093	,002	,009	,013	,016
100	,016	,049	,059	,071	,001	,006	,009	,011

Tableau 3 ■ Indices de spécificité et de sensibilité du test t approximatif (11) pour le test d'une différence entre deux variances corrélées au seuil de 5% (test unilatéral) comparativement au test exact basé sur la distribution F_ρ (20000 échantillons Monte Carlo)

N	Spécificité				Sensibilité			
	$\rho = 0$	$\frac{1}{2}$	$\sqrt{\frac{1}{2}}$	$\sqrt{\frac{3}{4}}$	$\rho = 0$	$\frac{1}{2}$	$\sqrt{\frac{1}{2}}$	$\sqrt{\frac{3}{4}}$
5	,742	,625	,541	,442	,755	,620	,529	,446
10	,910	,783	,691	,605	,902	,777	,689	,608
25	,963	,872	,797	,742	,959	,857	,794	,763
100	,996	,931	,901	,876	,990	,940	,894	,898

Note. (Spécificité) = $[F_\rho^+, t+]/[t+]$, où $[t+]$ indique le nombre d'échantillons dans lesquels le test t s'est révélé significatif, et $[F_\rho^+, t+]$ indique le nombre d'échantillons dans lesquels les tests q (selon F_ρ) et t se sont révélés conjointement significatifs. (Sensibilité) = $[F_\rho^+, t+]/[F_\rho^+]$.

1. Lorsqu'on connaît ρ ou qu'on peut lui supposer une valeur approximative, surtout si celle-ci est élevée, mieux vaut utiliser cette valeur (même approchée) en appliquant le test correct F_ρ basé sur le quotient de variances q . Pour pallier l'absence de valeurs critiques de la loi F_ρ , l'utilisateur peut recourir à la transformation de Finney (éq. 3), qui ramène la loi F_ρ à une loi F standard.
2. Le test F_r n'est pas exact (i.e. ne respecte pas le seuil de signification α), sauf pour n élevé. Étant donné, d'une part, que son exploitation suppose la disponibilité d'une table étendue de valeurs critiques et, d'autre part, que le test t pour la même situation convient mieux, nous recommandons sa mise aux oubliettes.
3. Le test t , selon la formule (11), est exact (c.-à-d. respecte le seuil de signification α) et il manifeste une puissance qui rejoint rapidement celle du test F_ρ , tout en étant satisfaisante dès $n = 5$. Ces qualités du test t s'affirment malgré la concordance échantillonnale imparfaite entre les deux tests (voir tableau 2).

Voilà tout l'éclairage que nous pouvons fournir à ce jour sur le quotient de deux variances corrélées, sa distribution et ses tests.

Note de l'auteur

Une version antérieure de cet article a paru dans Laurencelle (2000b).

Références

Bose, S. S. (1935). On the distribution of the ratio of variances of two samples drawn from a given normal bivariate population. *Sankhy* ?, 2, 65.

Finney, D. J. (1938). The distribution of the ratio of estimates of the two variances in a sample from a normal bi-variate population. *Biometrika*, 30, 190–192.

Guilford, J. P. (1965). *Fundamental statistics in psychology and education*. McGraw-Hill : New York.

Hinkle, D. E., Wiersma, W., & Jurs, S. G. (1988). *Applied statistics for the behavioral sciences (2e édition)*. Houghton Mifflin : Boston.

Johnson, N. L., Kotz, S., & Balakrishnan, N. (1995). *Continuous univariate distributions (2 volumes) (2e édition)*. Wiley : New York.

Kendall, M. G. & Stuart, A. (1977). *The advanced theory of statistics. Volume, 1*.

Laurencelle, L. [L.]. (2000a). Distribution, moments et estimation de la corrélation dans une population normale bivariée. *Lettres Statistiques*, 11, 31–45.

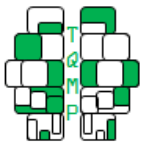


Tableau 4 ■ Puissance estimée des tests F_ρ , F_r et t au seuil $\alpha = 0,05$ en fonction du quotient de variances σ_X^2/σ_Y^2 et de $\rho_{X,Y}$, pour $n = 5$

$\sigma_X^2/\sigma_Y^2 =$		1	3/2	2	4	6	8
$\rho = 0$	F_ρ	,051	,095	,147	,330	,471	,582
	F_r	,083	,143	,207	,413	,558	,662
	t	,049	,092	,140	,306	,434	,539
$\rho = \frac{1}{2}$	F_ρ	,047	,102	,155	,388	,556	,674
	F_r	,076	,156	,235	,479	,634	,727
	t	,045	,100	,156	,361	,515	,618
$\rho = \sqrt{\frac{1}{2}}$	F_ρ	,053	,110	,181	,502	,699	,811
	F_r	,083	,179	,273	,588	,733	,824
	t	,050	,116	,184	,463	,629	,737
$\rho = \sqrt{\frac{3}{4}}$	F_ρ	,051	,139	,278	,743	,906	,957
	F_r	,080	,230	,395	,753	,876	,928
	t	,050	,154	,285	,643	,806	,876

Tableau 5 ■ Puissance estimée des tests F_ρ , F_r et t au seuil $\alpha = 0,05$ en fonction du quotient de variances σ_X^2/σ_Y^2 et de $\rho_{X,Y}$, pour $n = 5$

$\sigma_X^2/\sigma_Y^2 =$		1	3/2	2	4	6	8
$\rho = 0$	F_ρ	,050	,141	,255	,628	,822	,907
	F_r	,061	,164	,288	,662	,843	,920
	t	,050	,139	,253	,621	,817	,903
$\rho = \frac{1}{2}$	F_ρ	,049	,155	,299	,736	,902	,958
	F_r	,061	,184	,338	,759	,905	,958
	t	,049	,157	,299	,723	,886	,947
$\rho = \sqrt{\frac{1}{2}}$	F_ρ	,051	,189	,382	,863	,965	,990
	F_r	,061	,227	,429	,859	,958	,986
	t	,050	,197	,388	,833	,950	,982
$\rho = \sqrt{\frac{3}{4}}$	F_ρ	,051	,283	,611	,982	,998	1
	F_r	,063	,332	,628	,966	,993	,999
	t	,052	,295	,587	,959	,992	,998

Laurencelle, L. [L.]. (2000b). Le quotient de deux variances corrélées : distribution, moments et approximations. *Lettres Statistiques*, 11, 47–58.

Laurencelle, L. [L.] & Dupuis, F. A. (2000). *Tables statistiques expliquées et appliquées (2e édition)*. Le Griffon d'argile : Sainte-Foy.

Walker, H. M. & Lev, J. (1953). *Statistical inference*. Holt, Rinehart : New York.

Appendice : Démontrer que $\rho(s_X^2, s_Y^2) = \rho^2$

(gracieusement communiquée par Denis Allaire, Ph.D., université de Sherbrooke)

Soit A, B et C, des variables aléatoires (v.a.) indépendantes de loi normale standard, $N(0, 1)$, d'où il suit notamment que, par exemple :

$$E(s_{A,B}) = \sigma_{A,B} = 0 \tag{d1a}$$

et

$$E(s_A^2 \cdot s_B^2) = \sigma_A^2 \cdot \sigma_B^2 = 1. \tag{d1b}$$



Tableau 6 ■ Puissance estimée des tests F_ρ , F_r et t au seuil $\alpha = 0,05$ en fonction du quotient de variances σ_X^2/σ_Y^2 et de $\rho_{X,Y}$, pour $n = 25$

$\sigma_X^2/\sigma_Y^2 =$		1	3/2	2	4	6	8
$\rho = 0$	F_ρ	,051	,247	,509	,953	,996	,999
	F_r	,055	,258	,523	,955	,996	,999
	t	,051	,248	,508	,952	,995	,999
$\rho = \frac{1}{2}$	F_ρ	,051	,298	,600	,985	,999	1
	F_r	,055	,312	,611	,985	,999	1
	t	,052	,299	,596	,983	,999	1
$\rho = \sqrt{\frac{1}{2}}$	F_ρ	,050	,386	,759	,998	1	1
	F_r	,054	,402	,764	,997	1	1
	t	,050	,387	,751	,997	1	1
$\rho = \sqrt{\frac{3}{4}}$	F_ρ	,049	,614	,949	1	1	1
	F_r	,054	,624	,941	1	1	1
	t	,050	,610	,937	1	1	1

Tableau 7 ■ Puissance estimée des tests F_ρ , F_r et t au seuil $\alpha = 0,05$ en fonction du quotient de variances σ_X^2/σ_Y^2 et de $\rho_{X,Y}$, pour $n = 100$

$\sigma_X^2/\sigma_Y^2 =$		1	3/2	2	4	6	8
$\rho = 0$	F_ρ	,051	,643	,962	1	1	1
	F_r	,052	,646	,962	1	1	1
	t	,052	,642	,961	1	1	1
$\rho = \frac{1}{2}$	F_ρ	,051	,747	,989	1	1	1
	F_r	,052	,750	,989	1	1	1
	t	,051	,747	,989	1	1	1
$\rho = \sqrt{\frac{1}{2}}$	F_ρ	,050	,883	,999	1	1	1
	F_r	,050	,883	,999	1	1	1
	t	,049	,881	,999	1	1	1
$\rho = \sqrt{\frac{3}{4}}$	F_ρ	,048	,991	1	1	1	1
	F_r	,049	,990	1	1	1	1
	t	,048	,989	1	1	1	1

Soit les variables construites X, Y :

$$X = \sqrt{\rho} A + \sqrt{1 - \rho} B; \tag{d2a}$$

$$Y = \sqrt{\rho} A + \sqrt{1 - \rho} C; \tag{d2b}$$

on voit que X, Y sont des v.a. bivariées de loi $N(0, 0, 1, 1, \rho)$. Il suit de (d2) que :

$$s_X^2 = \rho s_A^2 + (1 - \rho) s_B^2 + 2\sqrt{\rho(1 - \rho)} s_{A,B}; \tag{d3a}$$

$$s_Y^2 = \rho s_B^2 + (1 - \rho) s_A^2 + 2\sqrt{\rho(1 - \rho)} s_{A,C}. \tag{d3b}$$

Or, nous cherchons à évaluer :

$$\begin{aligned} \rho(s_X^2, s_Y^2) &= cov(s_X^2, s_Y^2) / \sqrt{var(s_X^2)var(s_Y^2)} \\ &= \{E(s_X^2, s_Y^2) - E(s_X^2)E(s_Y^2)\} / \sqrt{var(s_X^2)var(s_Y^2)}. \end{aligned} \tag{d4}$$



Puisque X et Y sont chacune de loi $N(0, 1)$, nous savons que :

$$E(s_X^2) = E(s_Y^2) = \sigma^2 = 1 \tag{d5a}$$

$$var(s_X^2) = var(s_Y^2) = 2\sigma^4/(n - 1) = 2/(n - 1) \tag{d5b}$$

d'où :

$$E(s_X^2)^2 = E(s_Y^2)^2 = \sigma^4[1 + 2/(n - 1)] = 1 + 2/(n - 1). \tag{d5c}$$

Quant au terme clé $E(s_X^2 \cdot s_Y^2)$ apparaissant au numérateur de (d4), on l'obtient en multipliant les expressions en (d3) et en en prenant l'espérance. Ce produit contient 9 termes, dont voici la teneur :

$E\{s_X^2 s_Y^2\} = E\{\rho^2(s_A^2)^2\} = \rho^2[1 + 2/(n - 1)],$	$= \rho^2 + 2\rho^2/(n - 1)$ par (d5c)	(d6a)
$+ E\{\rho(1 - \rho)s_A^2 s_B^2\} = \rho(1 - \rho)\sigma_A^2 \sigma_B^2$	$= \rho(1 - \rho)$ par (d1b)	(d6b)
$+ E\{2\rho\sqrt{\rho(1 - \rho)}s_A^2 s_{A,B}\}$	$= 0$ par (d1a)	(d6c)
$+ E\{\rho(1 - \rho)s_A^2 s_C^2\} = \rho(1 - \rho)\sigma_A^2 \sigma_C^2$	$= \rho(1 - \rho)$, par (d1b)	(d6d)
$+ E\{(1 - \rho)^2 s_B^2 s_C^2\} = (1 - \rho)^2 \sigma_B^2 \sigma_C^2$	$= (1 - \rho)^2$, par (d1b)	(d6e)
$+ E\{2(1 - \rho)\sqrt{\rho(1 - \rho)}s_A^2 s_{A,B}\}$	$= 0$, par (d1a)	(d6f)
$+ E\{2\rho\sqrt{\rho(1 - \rho)}s_A^2 s_{A,C}\}$	$= 0$, par (d1a)	(d6g)
$+ E\{2(1 - \rho)\sqrt{\rho(1 - \rho)}s_A^2 s_{A,C}\}$	$= 0$, par (d1a)	(d6h)
$+ E\{4\rho(1 - \rho)s_{A,B}s_{A,C}\} = 0,$	par (d1a)	(d6i)
$= \rho^2 + 2\rho(1 - \rho) + (1 - \rho)^2 + 2\rho^2/(n - 1) = 1 + 2\rho^2/(n - 1).$		(d6j)

Incorporant (d6j) dans (d4) et utilisant (d5b), nous obtenons enfin :

$$\begin{aligned} \rho(s_X^2, s_Y^2) &= [1 + 2\rho^2/(n - 1) - 1 \times 1] / \sqrt{\frac{2}{n - 1} \times \frac{2}{n - 1}} \\ &= \rho^2 \end{aligned} \tag{d7}$$

QED

Citation

Laurencelle, L. [Louis]. (2016). Le quotient de deux variances corrélées, sa distribution et son test. *The Quantitative Methods for Psychology*, 12(2), 131–137. doi :[10.20982/tqmp.12.2.r131](https://doi.org/10.20982/tqmp.12.2.r131)

Copyright © 2016, Laurencelle . This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Received: 26/10/2015 ~ Accepted: 01/01/2016