MATHÉMATIQUES ET SCIENCES HUMAINES

H. ROUANET
D. LÉPINE

Problèmes de méthodologie statistique. II. - Étude d'un conflit robustesse-efficacité dans le problème de la comparaison de deux moyennes (groupes indépendants)

Mathématiques et sciences humaines, tome 47 (1974), p. 61-71
<http://www.numdam.org/item?id=MSH_1974__47__61_0>

Math. Sci. hum. (12^e année, n° 47, 1974, p. 61-71)

PROBLÈMES DE MÉTHODOLOGIE STATISTIQUE

II. — ÉTUDE D'UN CONFLIT ROBUSTESSE-EFFICACITÉ DANS LE PROBLÈME DE LA COMPARAISON DE DEUX MOYENNES (GROUPES INDÉPENDANTS)

par
H. ROUANET ET D. LÉPINE

RÉSUMÉ

Le conflit robustesse-efficacité se trouve posé notamment chaque fois que l'on a à choisir entre deux méthodes d'inférence statistique dont l'une est privilégiée sous un modèle plus spécifique et l'autre sous un modèle plus général. Ce conflit est étudié dans le cas de la comparaison de deux moyennes (groupes indépendants), à propos du choix entre le modèle (spécifique) postulant l'égalité des variances intra-groupes et le modèle (général) à variances quelconques. On montre que le choix n'est crucial (risque de faussages ou de pertes d'efficacité importants) que dans le cas des petits échantillons, et que le cas des échantillons de taille égale est privilégié. On fait apparaître à ce propos les rôles différents que jouent, dans le problème de la comparaison de moyennes, les deux conditions classiques d'égalité des variances et de normalité des distributions.

SUMMARY

The conflict robustness-efficiency poses itself each time one has to choose between two methods of statistical inference where one method is privileged under a more specific model and the other method under a more general model. This conflict is studied in the case of the comparison of two means (independent groups) in connection with the choice between the model (specific) postulating the equality of the intragroup variances and the model (general) with any variances. One shows that the choice is only then crucial (risk of falsification or important loss of efficiency) in the case of small samples, and the case of samples where equal sizes are privileged. In this connection the different roles which the two classical conditions of equal variances and normal distribution play in the problem of mean comparison are brought to light.

1. INTRODUCTION

Lorsqu'on se pose un problème d'inférence statistique (interrogation sur un paramètre) il arrive que plusieurs méthodes soient envisageables, chaque méthode correspondant à un modèle probabiliste possible de la situation étudiée et étant privilégiée sous ce modèle. Choisir une méthode statistique revient ainsi à choisir un modèle de la situation. Plus précisément, un cas fréquent est celui où on a le choix entre deux modèles dont l'un est plus *spécifique*, ou *moins général*, que l'autre. Choisir le modèle spécifique, c'est courir le risque que seul le modèle général soit vrai ; or la méthode statistique fondée sur le modèle spécifique peut être *faussée* lorsque ce modèle spécifique n'est pas vrai, le modèle général étant seul vrai ; en termes de test, elle peut conduire à trop, ou trop peu de résultats significatifs ; en termes d'intervalles de confiance, elle peut conduire à des intervalles trop étroits ou trop larges. Autrement dit, la méthode statistique fondée sur le modèle spécifique peut être plus ou moins *robuste* lorsqu'on se place dans le modèle général en dehors du modèle spécifique. Inversement, choisir le modèle général, c'est courir le risque que le modèle spécifique soit vrai ; or la méthode fondée sur le modèle général peut être moins *efficace* que la méthode fondée sur le modèle spécifique lorsque ce dernier est vrai ; en termes de test, on dira qu'une méthode est plus efficace (ou plus puissante) si la probabilité d'observer un résultat significatif à un seuil donné pour une certaine contre-hypothèse est plus élevée ; en termes d'intervalles de confiance, une méthode est plus efficace si pour un seuil donné elle conduit à des intervalles plus étroits (les notions relativement peu techniques de robustesse et d'efficacité nous suffiront pour les développements de cet article).

On voit donc que pour choisir entre deux méthodes associées respectivement à un modèle spécifique et à un modèle général, on va être amené à examiner le *conflit robustesse-efficacité*, sous la forme suivante :

— étude de la robustesse, sous le modèle général, de la méthode fondée sur le modèle spécifique ;
— étude de l'efficacité, sous le modèle spécifique, de la méthode fondée sur le modèle général.

Par exemple, supposons qu'on ait à choisir entre deux méthodes statistiques, l'une fondée sur le modèle normal (modèle spécifique), l'autre non paramétrique (donc fondée sur un modèle plus général). On pourra préférer la méthode fondée sur le modèle normal, même si le modèle normal est peu réaliste, si on estime que cette méthode est suffisamment robuste sous le modèle général ; on pourra au contraire préférer la méthode non paramétrique, même si le modèle normal est réaliste, si cette méthode est suffisamment efficace sous le modèle normal.

Dans certains cas, on peut procéder à une étude systématique du conflit robustesse-efficacité. C'est ce que nous nous proposons de faire dans cette note à propos du problème de la comparaison de deux moyennes (groupes indépendants). Outre l'intérêt intrinsèque du problème étudié, nous tirerons de notre étude des implications méthodologiques d'une portée assez générale sur les rôles différents que peuvent jouer les modèles en inférence statistique.

2. Position du problème

Soit deux groupes d'observations numériques de tailles n', n'', de moyennes M', M'', de variances corrigées S'^2, S''^2 (nous appelons variance-corrigée la somme des carrés des écarts à la moyenne divisée par l'effectif du groupe moins un). On suppose que ces groupes constituent des échantillons indépendants extraits de distributions parentes de moyennes μ', μ'' et de variances σ'^2, σ''^2. L'inférence portera sur la différence $\mu' - \mu''$ (problème de comparaison de moyennes). En termes de test : on posera l'hypothèse nulle $\mu' = \mu''$; en termes d'intervalle de confiance, on cherchera un intervalle pour la différence $\mu' - \mu''$.

Dans ces conditions, nous envisageons les deux modèles suivants :

— (Modèle général) : σ'^2 et σ''^2 quelconques ; on appellera ce modèle *modèle à variances quelconques* ;
— (Modèle spécifique) : $\sigma'^2 = \sigma''^2$ (on désignera alors par σ^2 la valeur commune de σ'^2 et σ''^2) ; on appellera ce modèle *modèle équivarié.*

(On pourrait envisager d'autres modèles spécifiques, stipulant que les variances sont inégales dans un sens donné, par exemple des modèles du type $\sigma'^2 > \sigma''^2$, ou $\sigma'^2 = k\ \sigma''^2$, k étant une constante différente de 1 ; dans cette note nous nous limiterons à la considération du modèle équivarié.)

Nous étudierons le conflit robustesse-efficacité associé à ces deux modèles ; plus précisément, nous nous placerons dans les deux cas suivants :

Grands échantillons : nous n'aurons pas alors besoin de préciser davantage les modèles en présence (nous ne supposerons pas, en particulier, la condition de normalité) ;

Petits échantillons, auquel cas nous adjoindrons la *condition de normalité*, c'est-à-dire nous supposerons que les distributions parentes des deux échantillons sont normales.

(La condition de normalité parente permet naturellement des inférences pour des échantillons de taille quelconque ; mais elle est en fait inutile dans le cas de grands échantillons.)

3. Variables aléatoires privilégiées : premières propriétés

En rapportant la variable aléatoire (v.a.) $M' - M'' - (\mu' - \mu'')$ à l'estimation de son écart-type $\frac{\sigma'^2}{n'} + \frac{\sigma''^2}{n''}$, on construira pour chacun des deux modèles une *v.a. privilégiée*, d'où on dérivera les méthodes d'inférence privilégiées sous chaque modèle.

Pour le modèle équivarié, on prendra, pour estimer *Var* $(M' - M'')$ l'estimateur classique W^2 défini par $W^2 = S^2 \left(\frac{1}{n'} + \frac{1}{n''}\right)$, S^2, estimateur de la variance commune σ^2, étant défini par $S^2 = \frac{(n'-1)\,S'^2 + (n''-1)\,S''^2}{n' + n'' - 2}$. On en déduit la v.a. privilégiée $\frac{M' - M'' - (\mu' - \mu'')}{W}$. Les propriétés suivantes sont classiques : sous le modèle équivarié :

— $E\,(W^2) = \sigma^2 \left(\frac{1}{n'} + \frac{1}{n''}\right)$;

— la distribution asymptotique de $\frac{M' - M'' - (\mu' - \mu'')}{W}$ est normale réduite (*cf.* Apprendice 1) ;

— sous la condition de normalité, W^2 est distribué comme $\sigma^2 \left(\frac{1}{n'} + \frac{1}{n''}\right) \frac{\chi^2\,(n' + n'' - 2)}{n' + n'' - 2}$ (*cf.* Appendice 2) d'où on déduit que $\frac{M' - M'' - (\mu' - \mu'')}{W}$ est distribuée comme $t_{[n'+n''-2]}$. (t de Student à $n' + n'' - 2$ d.l.)

On en tire les solutions classiques aux problèmes de test et d'intervalle de confiance sous le modèle équivarié. On prend pour statistique de test $T = \frac{M' - M''}{W}$ (T est obtenue en remplaçant $\mu' - \mu''$ par 0 dans la v.a. privilégiée) ; sous H_0, T est asymptotiquement une v.a. normale réduite, et, sous la condition de normalité, t est distribuée comme $t_{[n'+n''-2]}$, d'où le test (« t de Student » classique). De même, les limites de confiance au seuil α sont, pour de grands échantillons : $M' - M'' \pm z_\alpha\, W$ (où z_α est la valeur critique au seuil α de la distribution normale réduite), et, sous la condition de normalité : $M' - M'' \pm t_{[n'+n''-2]\alpha}\, W$ (où $t_{[n'+n''-2]\alpha}$ est la valeur critique au seuil α de la v.a. t de Student à $n' + n'' - 2$ d.l.).

Pour le modèle à variances quelconques, on prendra, pour estimer *Var* $(M' - M'')$, l'estimateur W^{*2} défini par $W^{*2} = \frac{S'^2}{n'} + \frac{S''^2}{n''}$ (obtenu en remplaçant dans *Var* $(M' - M'')$, σ'^2 par son estimateur S'^2 et σ''^2 par son estimateur S''^2). On en déduit la v.a. privilégiée $\frac{M' - M'' - (M' - M'')}{W^*}$. On peut alors démontrer les propriétés suivantes :

— $E\,(W^{*2}) = \frac{\sigma'^2}{n'} + \frac{\sigma''^2}{n''}$;

— la distribution asymptotique de $\frac{M' - M'' - (\mu' - \mu'')}{W^*}$ est normale réduite (*cf.* Appendice 1).

— sous la condition de normalité, W^{*2} est approximativement distribué comme $\left(\frac{\sigma'^2}{n'} + \frac{\sigma''^2}{n''}\right) \frac{\chi^2 \nu}{\nu}$, où $\nu = \dfrac{\left(\frac{\sigma'^2}{n'} + \frac{\sigma''^2}{n''}\right)^2}{\dfrac{\left(\frac{\sigma'^2}{n'}\right)^2}{n'-1} + \dfrac{\left(\frac{\sigma''^2}{n''}\right)^2}{n''-1}}$ (*cf.* Appendice 2), d'où on déduit que $\frac{M' - M'' - (\mu' - \mu'')}{W^*}$ est approximativement distribué comme $t_{[\nu]}$.

On en tire les solutions suivantes aux problèmes de test et d'intervalle de confiance sous le modèle à variances quelconques. On prend pour statistique de test $T^* = \frac{M' - M''}{W^*}$. Sous H_0 cette statistique est asymptotiquement normale réduite, et, sous la condition de normalité, elle est approximativement distribuée comme $t_{[\nu]}$; ν n'étant pas connu on l'estimera par $\bar{\nu}$ défini à partir de ν en remplaçant les variances parentes par leurs estimations, c'est-à-dire $\bar{\nu} = \frac{\left(\frac{S'^2}{n'} + \frac{S''^2}{n''}\right)^2}{\frac{\left(\frac{S'^2}{n'}\right)^2}{n'-1} + \frac{\left(\frac{S''^2}{n''}\right)^2}{n''-1}}$; T^* est alors approximativement distribué comme $t_{[\bar{\nu}]}$, d'où le test.

De même, les limites de confiance au seuil α sont, pour de grands échantillons, $M' - M'' \pm z_\alpha W^*$, et, sous la condition de normalité : $M' - M'' \pm t_{[\bar{\nu}]\alpha} W^*$.

(Le test et l'intervalle de confiance approchés précédents sous la condition de normalité constituent « la solution de Welch » au problème de la comparaison de deux moyennes dans le cas de variances quelconques, sous la condition de normalité-problème dit de « Behrens-Fisher », auquel ont été proposés des solutions diverses ; *cf.* Welch, 1939.)

N.B. — La valeur de $\bar{\nu}$ sera en général non entière, on devra donc interpoler dans les tables usuelles du t (il sera préférable alors d'interpoler selon les *inverses* des d.l.).

De la définition des v.a. privilégiées on tire les propriétés suivantes, faciles à vérifier :

— Si $S'^2 = S''^2$, ou si $n' = n''$ (cas « équilibré »), alors $W^{*2} = W^2$ et $T^* = T$; par contre si $S'^2 \neq S''^2$ et $n' \neq n''$: $W^{*2} \neq W^2$ et $T^{*2} \neq T^2$; cette propriété étant vraie même lorsque les effectifs sont élevés : on note donc que les solutions associées aux deux modèles ne tendent nullement à coincider lorsque les effectifs sont élevés ;

— lorsque $n' \neq n''$: $W^{*2} > W^2$ (et $|T^*| < |T|$) si et seulement si le plus grand échantillon a la plus petite variance ; $W^{*2} < W^2$ (et $|T^*| > |T|$) si et seulement si le plus grand échantillon a la plus grande variance (il s'agit naturellement ici des variances observées) ;

$$min\ (n'-1,\ n''-1) \leqslant \nu \text{ ou } \bar{\nu} \leqslant n' + n'' - 2$$

$\nu = n' + n'' - 2$ si et seulement si $\frac{\sigma'^2}{n'(n'-1)} = \frac{\sigma''^2}{n''(n''-1)}$;

$\bar{\nu} = n' + n'' - 2$ si et seulement si $\frac{S'^2}{n'(n'-1)} = \frac{S''^2}{n''(n''-1)}$; $\nu = n' - 1$ si et seulement si $\sigma''^2 = 0$ (si $\sigma''^2 = 0$, on a $M'' = \mu''$, $S''^2 = 0$ et $T^* = \frac{M' - \mu'}{\frac{S'}{\sqrt{n}}}$: on est ramené au test de la comparaison d'une moyenne à la valeur donnée μ''. Lorsque $n' = n''$, $\nu = n' + n'' - 2$ si $\sigma'^2 = \sigma''^2$, $\bar{\nu} = n' + n'' - 2$ si $S'^2 = S''^2$. Plus généralement, lorsque $n' = n''$, $\nu = \frac{n' + n'' - 2}{2} \frac{(1+\lambda)^2}{1+\lambda^2}$ en posant $\lambda = \frac{\sigma'^2}{\sigma''^2}$, on retrouve que $n' - 1 = n'' - 1 \leqslant \nu \leqslant n' + n'' - 2$, avec $\nu = n' - 1 = n'' - 1$ pour $\lambda = 0$ ou $\lambda = \infty$.

Ces propriétés entraînent les conséquences suivantes :

— Si $n' \neq n''$ et si le plus *grand échantillon* a la plus *petite variance* : l'utilisation de T^* conduit à un résultat *moins significatif* que l'utilisation de T : en effet $|T^*| < |T|$ et pour tout α, $t_{\bar{\nu}[\alpha]} \geqslant t_{[n'+n''-2]\alpha}$ (puisque $\bar{\nu} \leqslant n' + n'' - 2$).

— Si $n' \neq n''$ et si le plus *grand échantillon* a la plus *grande variance* : la méthode du T^* conduira *généralement* à un résultat *plus significatif* que celle du T car $|T^*| > |T|$ (cependant $t_{[\bar{\nu}]\alpha} \geqslant t_{[n'+n''-2]\alpha}$ ce qui conduit à un effet en sens contraire) ; mais si les effectifs sont élevés on compare $|T^*|$ et $|T|$ au même z_α ; le seul effet qui demeure est $|T^*| > |T|$, donc T^* conduit toujours à un résultat plus significatif que T.

— Si $S'^2 = S''^2$ ou $n' = n''$, l'utilisation de T^* conduit à un résultat *moins significatif* que celle du T : en effet on a bien $T^* = T$ mais comme $\bar{\nu} \leqslant n' + n'' - 2$, $t_{[\bar{\nu}]\alpha} \geqslant t_{[n'+n''-2]\alpha}$, cependant, si les effectifs sont élevés $n' + n'' - 2$ et $\bar{\nu}$ sont élevés, et la différence s'estompe puisque les deux méthodes tendent à donner le même résultat ($t_{[\bar{\nu}]\alpha}$ et $t_{[n'+n''-2]\alpha}$ étant remplacés par z_α). Dans le cas où $n' = n''$, on a toujours $\nu \geqslant n' - 1 = n'' - 1 = \dfrac{n' + n'' - 2}{2}$: la différence entre $t_{[\bar{\nu}]\alpha}$ et $t_{[n'+n''-2]\alpha}$ sera faible sauf pour de très petits effectifs.

4. Etude de la robustesse, sous le modèle a variances quelconques, de la méthode fondée sur le modèle équivarié

Pour étudier la robustesse, sous le modèle à variances quelconques, de la méthode fondée sur le modèle équivarié, nous raisonnerons en termes de test. Lorsqu'on utilise la méthode du T, fondée sur la v.a. $\dfrac{M' - M'' - (\mu' - \mu'')}{W}$, alors que le modèle équivarié n'est pas vrai, on a un faussage que nous allons évaluer. Nous calculerons tout d'abord $E\,(W^2)$, ce qui nous permettra d'introduire la notion de « facteur de faussage ». On a : $E\,(W^2) = (\dfrac{1}{n'} + \dfrac{1}{n''})\,\dfrac{(n'-1)\,\sigma'^2 + (n''-1)\,\sigma''^2}{n' + n'' - 2}$ qu'on peut écrire :

$$E\,(W^2) = \frac{1}{\mathcal{F}^2}\,(\frac{\sigma'^2}{n'} + \frac{\sigma''^2}{n''})$$

en posant $\mathcal{F}^2 = \dfrac{(\dfrac{\sigma'^2}{n'} + \dfrac{\sigma''^2}{n''})\,(n' + n'' - 2)}{[(n'-1)\,\sigma'^2 + (n''-1)\,\sigma''^2]\,(\dfrac{1}{n'} + \dfrac{1}{n''})}$

(on remarquera que $\mathcal{F}^2$ ne dépend de σ'^2 et σ''^2 qu'à travers leur rapport $\dfrac{\sigma'^2}{\sigma''^2}$)

On a donc $E\,(W^2) = \dfrac{1}{\mathcal{F}^2}$ ($E\,(W^2)$ sous le modèle équivarié) ; si $\sigma'^2 = \sigma''^2$, $\mathcal{F}^2 = 1$: nous appellerons donc $\mathcal{F}^2$ (ou sa racine carrée) *facteur de faussage*. On vérifie les propriétés :

— Lorsque $n' = n''$, on a $\mathcal{F}^2 = 1$, quels que soient σ'^2 et σ''^2.

— Lorsque $n' \neq n''$: $\mathcal{F}^2 > 1$ si et seulement si le plus *grand échantillon* a la plus *petite variance parente* ; $\mathcal{F}^2 < 1$ si et seulement si le plus *grand échantillon* a la plus *grande variance parente.*

— Lorsque n' et n'' sont grands : $\mathcal{F}^2 \simeq \dfrac{\dfrac{\sigma'^2}{n'} + \dfrac{\sigma''^2}{n''}}{\dfrac{\sigma'^2}{n''} + \dfrac{\sigma''^2}{n'}}$ (qu'on peut écrire aussi $\dfrac{\lambda + l}{\lambda l + 1}$ en posant $\dfrac{\sigma'^2}{\sigma''^2} = \lambda$, $\dfrac{n'}{n''} = l$) (sur cette formule on voit immédiatement que $\mathcal{F}^2 = 1 \Rightarrow \lambda = 1$ ou $l = 1$).

Nous étudierons maintenant la *robustesse asymptotique* ; on démontre (*cf.* Appendice 1) que la v.a. $\dfrac{M' - M'' - (\mu' - \mu'')}{W}$ est asymptotiquement normale d'espérance 0 et de variance $\mathcal{F}^2$ (ce que nous écrirons $N\,(0, \mathcal{F}^2)$), d'où on tire les conditions suivantes :

— Si $n' = n''$: *pas de faussage* (résultat qui était à prévoir, puisque alors $W^* = W$).

— Si $n' \neq n''$ et si le plus grand échantillon correspond à la plus petite variance parente ($\mathcal{F}^2 > 1$), le test est *faussé positivement* (c'est-à-dire : le seuil réel est supérieur au seuil nominal on obtient trop de résultats significatifs) ; si le plus grand échantillon correspond à la plus grande variance parente ($\mathcal{F}^2 < 1$), le test est *faussé négativement* (le seuil réel est inférieur au seuil normal, on obtient trop peu de résultats significatifs).

L'amplitude du faussage dépend de la valeur de $\mathcal{G}^2$, laquelle est fonction des rapports $\lambda = \frac{\sigma'^2}{\sigma''^2}$ et $l = \frac{n'}{n''}$ (pour une valeur de $\mathcal{G}^2 \neq 1$ donnée, le faussage ne tend pas à s'atténuer lorsqu'on augmente la taille des échantillons).

Pour un contre-modèle donné, on peut facilement évaluer l'amplitude du faussage. Posons en effet $X = \frac{M' - M'' - (\mu' - \mu'')}{W}$. On peut écrire $X = \mathcal{G} Z$, où Z est une v.a. normale réduite. Soit α le seuil nominal ; il lui correspond une valeur z_α telle que $Prob\ (|Z| \geqslant z_\alpha) = \alpha$ (z_α est la valeur lue dans la table de la distribution normale réduite au seuil bilatéral α). L'événement critique au seuil α est $(|X| \geqslant z_\alpha)$; le seuil réel α' est donc

$$\alpha' = Prob\ (|X| \geqslant z_\alpha) = Prob\ (|Z| \geqslant \frac{z}{\mathcal{G}}).$$

A titre d'illustration, donnons quelques exemples numériques :

— Prenons $\frac{n'}{n''} = 5$ et $\frac{\sigma'^2}{\sigma''^2} = \frac{1}{5}$; le faussage est positif ; on trouve $\mathcal{G}^2 = 2{,}60$ d'où $\mathcal{G} \simeq 1{,}61$:

Pour $\alpha = 0{,}05$: $\frac{z_\alpha}{\mathcal{G}} = \frac{1{,}96}{1{,}61} = 1{,}217$. D'où $\alpha' = Prob\ (|Z| \geqslant \frac{z_\alpha}{\mathcal{G}}) = Prob\ (|Z| \geqslant 1{,}217)$ soit $\alpha' = 0{,}22$.

— Prenons $\frac{n'}{n''} = 5$ et $\frac{\sigma'^2}{\sigma''^2} = 5$; le faussage est négatif ; on trouve $\mathcal{G}^2 \simeq 0{,}385$ d'où $\mathcal{G} \simeq 0{,}62$.

Considérons un exemple moins extrême : pour $\frac{n'}{n''} = 2$, $\frac{\sigma'^2}{\sigma''^2} = \frac{1}{2}$, on trouve, pour $\alpha = 0{,}05$, $\alpha' = 0{,}08$; pour $\frac{n'}{n''} = 2$, $\frac{\sigma'^2}{\sigma''^2} = 2$, on trouve, toujours pour $\alpha = 0{,}05$; $\alpha' = 0{,}029$. Sans être aussi énormes, les faussages ne sont pas négligeables.

Nous étudierons enfin la *robustesse sous le modèle normal* : on démontre alors (*cf.* Appendice 2) que W^2 est approximativement distribué comme

$$\left(\frac{\sigma'^2}{n'} + \frac{\sigma''^2}{n''}\right) \frac{1}{\mathcal{G}^2} \frac{\chi^2[\nu]'}{\nu'} \text{ avec } \nu' = \frac{[(n'-1)\ \sigma'^2 + (n''-1)\ \sigma''^2]^2}{(n'-1)\ \sigma'^4 + (n''-1)\ \sigma''^4}.$$

(On vérifie que $\nu' \leqslant n' + n'' - 2$, avec l'égalité si et seulement si $\sigma'^2 = \sigma''^2$). Il en résulte que $\frac{M' - M'' - (\mu' - \mu'')}{W}$ est approximativement distribué comme $\mathcal{G}\ t_{[\nu']}$, d'où on tire les conclusions suivantes.

— Si $n' = n''$: $\mathcal{G} = 1$, mais on a toujours $\nu' \leqslant n' + n'' - 2$: la dispersion de la v.a. $t_{[\nu']}$ étant plus grande que celle de $t_{[n'+n''-2]}$, on a un *faussage positif* (faible sauf pour de très petits échantillons, et d'autant moins marqué que les effectifs sont plus élevés) ;

— Si $n' \neq n''$ et si le plus grand échantillon correspond à la plus petite variance parente ($\mathcal{G}^2 > 1$), le test est *faussé positivement* ; en effet, la dispersion de la v.a. $\mathcal{G}\ t_{[\nu']}$ est plus grande que celle de $t_{[n'+n''-2]}$;

— Si $n' \neq n''$ et si le plus grand échantillon correspond à la plus grande variance parente ($\mathcal{G}^2 < 1$), le test est faussé de deux façons : comme $\mathcal{G}^2 < 1$ la distribution de $\mathcal{G}\ t_{[\nu']}$ tend à être moins dispersée que celle de $t_{[n'+n''-2]}$ d'où source de faussage négatif ; mais d'autre part $\nu' \leqslant n' + n'' - 2$ d'où source de faussage positif ; cependant, en général le premier effet sera de beaucoup le plus important, et le faussage global sera *négatif*.

Pour étudier l'amplitude du faussage, on raisonnera comme dans le cas asymptotique. Soit $T_{\nu'}$ une v.a. ayant la distribution du t à ν' d.l. Appelons $t_{[\nu']\alpha}$ la valeur lue dans la table au seuil bilatéral α : $Prob\ (|T|_{[\nu']\alpha} \geqslant t_{[\nu']\alpha}) = \alpha$. Soit α le seuil nominal, le seuil réel est $\alpha' = Prob\ (|X| \geqslant t_{[n'+n''-2]\alpha}) = Prob\ (T_{[\nu']} \geqslant \frac{t_{[n'+n''-2]\alpha}}{\mathcal{G}})$

Considérons quelques exemples numériques :

Prenons $n' = 20$, $n'' = 4$ et $\frac{\sigma'^2}{\sigma''^2} = \frac{1}{5}$ d'où (faussage positif) : on trouve $\mathcal{G}^2 = 2{,}804$ d'où $\mathcal{G} \simeq 1{,}67$, et d'autre part $\nu' \simeq 12{,}3$.

Pour $\alpha = 0{,}05$, $t_{[n'+n''-2]\alpha} = 2{,}074$. D'où $\frac{t_{[n'+n''-2]\alpha}}{\mathcal{G}} = \frac{2{,}074}{1{,}67} = 1{,}24$.

En consultant la table du t pour 12 et 13 d.l. on trouve que $\alpha' = Prob\ (|T_{[\alpha]}| \geqslant 1{,}24)$ est de l'ordre de 0,30.

On remarque que pour un même rapport $l = \frac{n'}{n''}$ le test est plus faussé que dans le cas asymptotique.

Prenons maintenant $n' = 20$, $n'' = 4$ et $\frac{\sigma'^2}{\sigma''^2} = 5$ (d'où faussage négatif) : on trouve $\mathcal{G}^2 = 0{,}374$ d'où $\mathcal{G} \simeq 0{,}61$ et d'autre part $\nu' = 20{,}1$

pour $\alpha = 0{,}05$, $\frac{t_{[n'+n''-2]\alpha}}{\mathcal{G}} = \frac{2{,}074}{0{,}61} = 3{,}40$.

En consultant la table du t pour 20 d.l., on trouve $\alpha' = Prob\ (|T_{[\nu']}| \geqslant 3{,}40)$ de l'ordre de 0,002. On constate que la perte en d.l. (source de faussage positif) est loin de compenser le faussage négatif dû au facteur de faussage $\mathcal{G}$.

Prenons enfin $n' = 4$, $n'' = 4$ et $\frac{\sigma'^2}{\sigma''^2} = 5$: $\mathcal{G}^2 = 1$ et $\nu' \simeq 4{,}1$. Pour $\alpha = 0{,}05$, $t_{[n'+n''-2]\alpha} = 2{,}45$. En consultant la table du t à 4 d.l. on trouve $\alpha' = Prob\ (T_{[\nu']}) \geqslant 2{,}45) \simeq 0{,}07$: le faussage positif est assez notable (les effectifs sont ici vraiment très faibles).

5. Etude de l'efficacité, sous le modèle équivarié, de la méthode fondée sur le modèle a variances quelconques

Pour étudier l'efficacité, sous le modèle équivarié de la méthode fondée sur le modèle à variances quelconques, nous raisonnerons en termes d'intervalles de confiance (on transposera facilement les raisonnements en termes de test).

Nous envisagerons d'abord *l'efficacité asymptotique*. La méthode fondée sur le modèle à variances quelconques conduit à prendre $M' - M'' \pm z_\alpha W^*$ alors que sous le modèle équivarié on devrait prendre $M' - M'' \pm z_\alpha W$. Or sous le modèle équivarié W^2 et W^{*2} estiment la même variance $\sigma^2\ (\frac{1}{n'} = \frac{1}{n''})$, donc W^* ne tend pas en moyenne à être supérieur à W, et l'intervalle de confiance $M' - M'' \pm z_\alpha W^*$ ne tend pas en moyenne à être plus large que $M' - M'' \pm z_\alpha W$: lorsque les effectifs sont élevés, on n'a donc *pas de perte d'efficacité* — lorsqu'on utilise la méthode fondée sous le modèle à variances quelconques alors que le modèle équivarié est vrai [1].

Nous envisagerons ensuite l'*efficacité sous le modèle normal* : la méthode fondée sur le modèle normal à variances quelconques conduit à prendre $M' - M'' \pm t_{[\nu]\alpha}\ W^*$ alors que sous le modèle normal équivarié on devrait prendre $M' - M'' \pm t_{[n'+n''-2]\alpha}\ W$. Or sous le modèle équivarié W^{*2} et W^2 estiment toujours la même variance, et W^* ne tend donc pas à être supérieur à W, mais comme $\bar{\nu} < n' + n'' - 2$, $t_{[\bar{\nu}]\alpha} \geqslant t_{[n'+n''-2]\alpha}$, donc l'intervalle $M' - M'' \pm t_{[\bar{\nu}]\alpha}\ W$ tend à être plus large que $M' - M'' \pm t_{[n'+n''-2]\alpha}\ W$, on a donc une *perte d'efficacité* lorsqu'on utilise la méthode fondée sur le modèle à variances quelconques alors que le modèle équivarié est vrai.

Exemple numérique $n' = 4$, $n'' = 20$: on a $n' + n'' = 22$; d'autre part, si $\sigma'^2 = \sigma''^2$,

$$\nu = \frac{(\frac{1}{n'} + \frac{1}{n''})^2}{\frac{1}{n'^2\ (n' - 1)} + \frac{1}{n''^2\ (n'' - 1)}} \simeq 4{,}5.$$

Donc $\bar{\nu}$ sera de l'ordre de 4,5. Au seuil $\alpha = 0{,}05$ on a $t_{[n'+n''-2]\alpha} = 2{,}07$ alors que $t_{[\nu]\alpha} = 2{,}67$.

1. Cependant sous le modèle équivarié on démontre que $Var\ (W^2) < Var\ (W^{*2})$ en prenant $M' - M'' \pm z_\alpha$ W, on aura donc des intervalles dont la largeur sera moins dispersée.

Dans le cas où $n' = n''$, lorsque $\sigma'^2 = \sigma''^2$, $\nu = n' + n'' - 2$; on a alors $S'^2 \simeq S''^2$, donc $\bar{\nu}$ de l'ordre de $n' + n'' - 2$: la perte d'efficacité sera généralement minime.

6. Conclusions de l'étude du conflit robustesse-efficacité

Pour résumer l'étude du conflit robustesse-efficacité, nous avons trouvé : d'une part que le faussage, sous le modèle à variances quelconques, de la méthode fondée sur le modèle équivarié, ne disparaît pas pour de grands échantillons, sauf dans le cas équilibré, pour lequel, par ailleurs, le faussage est faible pour les petits échantillons — et d'autre part, que la perte d'efficacité, sous le modèle équivarié, des méthodes fondées sur le modèle à variances quelconques devient nulle pour de grands échantillons, cette perte étant faible pour les petits échantillons dans le cas équilibré.

Ces résultats nous permettent de formuler plusieurs conclusions.

Une conclusion concerne le *caractère privilégié du cas équilibré* du double point de vue de la robustesse et de l'efficacité : dans ce cas, le choix d'une méthode de comparaison de moyenne n'est pas crucial. En conséquence, lorsque l'expérimentateur est libre de choisir les effectifs des groupes, et qu'il n'est pas assuré de l'égalité des variances, il a intérêt à prendre des effectifs égaux [1].

Une autre conclusion concerne le cas des *grands échantillons* : les méthodes fondées sur le modèle à variances quelconques dominent toujours celles fondées sur le modèle équivarié, puisque si le modèle équivarié est vrai, on ne gagne rien en efficacité à utiliser ces dernières méthodes et que s'il est faux on risque un faussage : pour de grands échantillons, on choisira donc toujours les méthodes fondées sur le modèle à variances quelconques.

Finalement, le *conflit robustesse-efficacité* ne se pose que si les effectifs sont *faibles* : si on peut admettre que les variances parentes sont égales (ou tout au moins ne diffèrent pas « trop », l'interprétation de « trop » dépendant du rapport des effectifs) on aura intérêt à utiliser les méthodes fondées sur le modèle équivarié puisque ces méthodes fournissent un gain d'efficacité.

Si on ne peut admettre que les variances parentes sont égales, on aura intérêt à utiliser les méthodes fondées sur le modèle à variances quelconques puisque en utilisant le modèle équivarié on risque un faussage.

Les résultats précédents nous permettent également de porter un jugement sur le *statut méthodologique du modèle équivarié* : On voit que pour de petits échantillons ce modèle peut être *utile* (s'il est vrai, il permet un gain d'efficacité) alors que pour de grands échantillons il ne peut être que nuisible (s'il est vrai, on ne gagne rien en efficacité et s'il est faux on a un faussage).

Ce statut est à comparer avec celui du *modèle normal* : pour de petits échantillons le modèle normal est *utile* (s'il est vrai, il permet des conclusions précises) mais pour de grands échantillons il est non pas nuisible mais simplement *inutile*, puisque les méthodes asymptotiques sous la condition de normalité sont les mêmes que sous le modèle général (on ne gagne rien à poser le modèle normal, mais on ne risquerait pas de faussage en le posant à tort) ; on peut donc dire que la condition de normalité est ici « inoffensive ».

(N.B. : Cette propriété du modèle normal dans le problème de comparaison de moyennes ne se généralise pas à tout problème faisant intervenir le modèle normal. *Cf.* § 7.2.)

7. Compléments sur le problème de la comparaison de deux moyennes

1) Nous avons introduit, dans le cas des petits échantillons, la condition de normalité. On peut montrer que les méthodes fondées sur cette condition sont robustes vis-à-vis de modèles beaucoup plus généraux que le modèle normal (distributions parentes pas trop dissymétriques et pas trop étalées aux extrémités, disons, de variance finie). Cette robustesse est d'autant plus grande que les effectifs sont plus élevés, ce qui se conçoit aisément puisque asymptotiquement la condition de normalité est inutile. Ce n'est donc que pour de très petits échantillons qu'il sera nécessaire, avant d'appliquer l'une des méthodes de comparaison de moyennes présentées dans cette note, de supposer que la condition de normalité est remplie. Il importe de préciser que dans ce cas cette supposition devra généralement être faite à partir de considérations extérieures à l'information expérimentale, c'est-à-dire portant sur le processus ayant conduit aux observations, plutôt qu'à partir de l'examen de l'échantillon. En effet, si les effectifs sont faibles, on peut difficilement se fonder sur les distributions observées pour en tirer des informations valables, sur les formes des distributions parentes (le test de la normalité, par exemple, sera généralement trop peu puissant pour qu'on puisse conclure valablement à partir d'un résultat non-significatif à la validité de la normalité.) Par contre, si les effectifs ne sont plus très faibles, on pourra bien déceler à partir des échantillons des écarts à la normalité, mais cette information sera inutile pour le problème de la comparaison des moyennes, puisque alors la condition de normalité est peu importante ou même inutile.

1. Rappelons que nous avons laissé ici de côté le cas où on disposerait de modèles spécifiques indiquant que les variances sont dans un sens donné : dans le cas où l'expérimentateur saurait a priori que $\sigma'^2 > \sigma''^2$, il pourrait envisager de prendre $n' > n''$.

2) Dans le cas de petits échantillons, pour choisir entre le modèle équivarié et le modèle à variances quelconques, on s'appuiera également surtout sur des considérations extérieures à l'information expérimentale. La prise en compte des informations expérimentales, c'est-à-dire des valeurs observées des variances corrigées S'^2 et S''^2, pose des problèmes délicats. On pourrait en effet être tenté de procéder à une comparaison statistique préalable des variances au moyen du test F classique de comparaison de variances. Or, ce test, valide sous la condition de normalité, est peu robuste vis-à-vis de la non-normalité : si la condition de normalité n'est pas strictement remplie, ce test risque donc de conduire à des résultats gravement faussés dans un sens ou dans l'autre, ainsi que nous l'avons indiqué dans un précédent article (Rouanet et Lépine, 1973).

On *évitera* donc de procéder à ce test sauf si l'on est absolument assuré de la condition de normalité, ou on procédera à un test ne reposant pas sur la condition de normalité [1]. Même dans ce dernier cas, on devra conclure avec précautions ; ce qui importe en effet n'est pas de déceler un écart quelconque à l'homogénéité des variances, mais d'apprécier si la non-homogénéité est suffisamment marquée pour fausser gravement les méthodes fondées sur le modèle équivarié. La grandeur de l'inégalité à déceler dépendra elle-même, conformément à ce que nous avons vu, des tailles relatives des échantillons. A partir de telles considérations, on sera ainsi conduit à être d'autant plus exigeant (c'est-à-dire à travailler à un seuil d'autant plus bas) avant de conclure à l'hétérogénéité qu'on est plus proche du cas équilibré.

En l'absence de considérations extérieures à l'information expérimentale ou tirées de celles-ci, conduisant à admettre de façon assurée que les variances parentes sont égales ou tout au moins ne diffèrent pas trop, il sera plus prudent de conclure à partir des méthodes fondées sur le modèle à variances quelconques, même si à titre indicatif on peut juger utile de calculer les valeurs de statistiques fondées sur le modèle équivarié.

La « correction » ainsi effectuée par rapport au modèle équivarié pourra être importante si les effectifs sont très différents ainsi que les variances corrigées. Ainsi, si au plus petit échantillon correspond la plus grande variance parente, on sait que le test T est faussé positivement ; en pareil cas, le plus petit échantillon aura en général la plus grande variance observée, donc on aura $|T^*| < |T|$ et le résultat fourni par la méthode du T^* sera moins significatif, ce qui revient bien à corriger le faussage positif du T. Cet exemple montre en passant combien il serait incorrect de conclure systématiquement à partir de celle des deux statistiques conduisant au résultat le plus significatif (ou aussi bien, en l'occurrence, à partir de celle conduisant au résultat le moins significatif).

3) Dans certains cas, on peut considérer que le phénomène étudié est tel que lorsque la moyenne varie la variance varie d'une manière qui peut être précisée par une relation fonctionnelle. On montre que dans ce cas il existe une transformation de la variable telle que la variable transformée ait une variance à peu près indépendante de la moyenne (on dit que la transformation « stabilise la variance »). Par exemple, à la relation $\sigma = k\mu$ correspond la transformation logarithmique. En pareils cas certains auteurs recommandent de procéder à la transformation appropriée de la variable afin de se ramener au cas équivarié. En fait, il n'est pas conseillé d'utiliser systématiquement cette procédure. En effet, une fois effectuée la comparaison des moyennes sur le protocole transformé, il convient de revenir à la variable primitive pour interpréter le résultat : or cela pose souvent des problèmes délicats. L'existence d'une méthode appropriée dans le cas non-équivarié permettra souvent d'éviter le recours à une transformation de la variable.

4) Les principales conclusions de cet article se généralisent à la comparaison des moyennes dans le cas de *plusieurs* groupes indépendants au moyen du test F classique d'analyse de la variance ; les résultats essentiels sont les suivants. Dans le cas non-équilibré, le faussage est positif ou négatif selon que la moyenne simple des variances est supérieure ou inférieure à la moyenne pondérée par les effectifs des groupes, donc si les effectifs élevés correspondent à des variances petites on a un faussage positif, et si les effectifs élevés correspondent à des variances élevées on a un faussage négatif. Dans le cas équilibré, on a un léger faussage positif qui subsiste asymptotiquement (à cet égard, le cas de deux groupes est donc exceptionnellement privilégié). Une statistique de rechange à la statistique F, dans le cas où les variances sont inégales, est $F^* = \dfrac{1}{G-1} \sum\limits_{G} \dfrac{n_g (M_g - M)^2}{S_g^2}$ où G désigne le nombre de groupes, n_g l'effectif du groupe g, M_g sa moyenne, S_g^2 sa variance-corrigée et où $M = \dfrac{\sum\limits_{G} n_g M_g / S_g^2}{\sum\limits_{G} n_g / S_g^2}$; sous H_0 (les G moyenne parentes égales), F^* est asymptotiquement distribué comme $\chi^2_{[G-1]}$ (F^* est la généralisation de la statistique T^* considérée dans le cas de deux groupes ; on vérifie que pour deux groupes $F^* = T^{*2}$).

1. Dans l'article précédemment cité nous avons indiqué un test d'égalité des variances moins puissant que le test F mais ne reposant pratiquement pas sur la condition de normalité.

BIBLIOGRAPHIE

[1] ROUANET H. et LÉPINE D., « Problèmes de méthodologie statistique. I - Introduction à l'étude de la robustesse des méthodes usuelles d'inférence statistique », *Math. Sci. hum.*, n° 46, 1973.

[2] WELSH B.-L., « The significance of the difference between two means when the population variances are unequal », *Biometrika*, 1937, 29, pp. 350-362.

APPENDICES

Appendice 1 : ***Distributions asymptotiques (sans condition de normalité).*** Formellement, on peut définir le cas asymptotique de la manière suivante : on pose $n' = k'n$, $n'' = k''n$; les facteurs k' et k'' restant fixes, on fait tendre n vers l'infini. On peut alors montrer, à partir de théorémes de convergence classiques [1], que la v.a. $\dfrac{M' - M'' - (\mu' - \mu'')}{W*} = \dfrac{M' - M'' - (\mu' - \mu'')}{\sqrt{\dfrac{S'^2}{n'} + \dfrac{S''^2}{n''}}}$ est asymptotiquement $\mathcal{N}$ (0,1).

Quant à la v.a. $\dfrac{M' - M'' - (\mu' - \mu'')}{W}$, on peut l'écrire $\dfrac{M' - M'' - (\mu' - \mu'')}{\sqrt{\dfrac{S'^2}{n'} + \dfrac{S''^2}{n''}}} \sqrt{\dfrac{\dfrac{S'^2}{n'} + \dfrac{S''^2}{n''}}{\dfrac{S'^2}{n''} + \dfrac{S''^2}{n'}}}$

qui est le produit d'une v.a. asymptotiquement $\mathcal{N}$ (0,1) par une v.a. qui converge en probabilité vers $\mathcal{G} = \dfrac{\dfrac{\sigma'^2}{k'} + \dfrac{\sigma''^2}{k''}}{\dfrac{\sigma'^2}{k''} + \dfrac{\sigma''^2}{k'}}$. Il en résulte que $\dfrac{M' - M'' - (\mu' - \mu'')}{W}$ est asymptotiquement $\mathcal{N}$ $(0, \mathcal{G}^2)$. Sous le modèle équivarié, $\mathcal{G} = 1$ et $\dfrac{M' - M'' - (\mu' - \mu'')}{W}$ est asymptotiquement $\mathcal{N}$ (0,1).

Appendice 2 : ***Distributions approchées sous la condition de normalité.***
Soit deux v.a. $\chi^2_{[\nu']}$ et $\chi^2_{[\nu'']}$. On sait qu'une combinaison linéaire $a'\,\chi^2_{[\nu']} + a''\,\chi^2_{[\nu'']}$ n'est distribuée exactement comme une v.a. $a\,\chi^2_{[\nu]}$ que si $a' = a''$ (auquel cas $a = a' = a''$ et $\nu = \nu' + \nu''$: théorème classique de Cochran). Cependant on montre qu'on peut approcher la distribution de $a'\,\chi^2_{[\nu']} + a''\,\chi^2_{[\nu'']}$ par celle d'une v.a. $a\,\chi^2_{[\nu]}$, ν et a étant choisis de telle sorte que la v.a. $a\,\chi^2_{[\nu]}$ ait même espérance et même variance que la v.a. $a'\,\chi^2_{[\nu']} + a''\,\chi^2_{[\nu'']}$, c'est-à-dire vérifiant :

$$\begin{cases} a\nu = a'\nu' + a''\nu'' \\ a^2\nu = a'^2\nu' + a''^2\nu'' \end{cases}$$

D'où on tire ν et a :

$$\begin{cases} \nu = \dfrac{(a'\nu' + a''\nu'')^2}{a'^2\nu' + a''^2\nu''} \\[2ex] a = \dfrac{a'^2\nu' + a''^2\nu''}{a'\nu' + a''\nu''} = \dfrac{1}{\nu}\,(a'\nu' + a''\nu'') \end{cases}$$

Lorsque $a' = a''$ on retrouve $a = a' = a''$, $\nu = \nu' + \nu''$.

1. *Cf.* H. ROUANET : « Un théorème de convergence probabiliste et ses applications à des méthodes statistiques courantes », *Math. et Sci. hum.*, n° 42, 1972.

Applications

1) *Distribution approchée de W^{*2} sous le modèle normal à variances quelconques*

On a $W^{*2} = \frac{S'^2}{n'} + \frac{S''^2}{n''}$. Comme $S'^2 = \sigma'^2 \frac{\chi^2[n'-1]}{n'-1}$, on peut écrire :

$$W^{*2} = \frac{\sigma'^2}{n'(n'-1)} \chi^2{}_{[n'-1]} + \frac{\sigma''^2}{n''(n''-1)} \chi^2{}_{[n''-1]} = a' \chi^2{}_{[n'-1]} + a'' \chi^2{}_{[n''-1]}$$

en posant $a' = \frac{\sigma'^2}{n'(n'-1)}$, $a'' = \frac{\sigma''^2}{n''(n''-1)}$. On en tire l'approximation définie par

$$\nu = \frac{(\frac{\sigma'^2}{n'} + \frac{\sigma''^2}{n''})^2}{\frac{(\sigma'^2/n')^2}{n'-1} + \frac{(\sigma''^2/n'')^2}{n''-1}}, \quad a = \frac{\frac{\sigma'^2}{n'} + \frac{\sigma''^2}{n''}}{\nu}, \text{ c'est-à-dire :}$$

W^{*2} est approximativement distribuée comme $(\frac{\sigma'^2}{n'} + \frac{\sigma''^2}{n''}) \frac{\chi^2{}_{[\nu]}}{\nu}$.

2) *Distribution approchée de W^2 sous le modèle normal à variances quelconques*

Par définition $W^2 = (\frac{1}{n'} + \frac{1}{n''}) \frac{1}{n'+n''-2} [(n'-1) S'^2 + (n''-1) S''^2]$ qu'on peut écrire

$$W^2 = (\frac{1}{n'} + \frac{1}{n''}) \frac{1}{n'+n''-2} (\sigma'^2 \chi^2{}_{[n'-1]} + \sigma''^2 \chi^2{}_{[n''-1]}) = a' \chi^2{}_{[n'-1]} + a'' \chi^2{}_{[n''-1]}$$

en posant $a' = (\frac{1}{n'} + \frac{1}{n''}) \frac{1}{n'+n''-2} \sigma'^2$, $a'' = (\frac{1}{n'} + \frac{1}{n''}) \frac{1}{n'+n''-2} \sigma''^2$.

On en tire l'approximation définie par $\nu = \frac{[(n'-1)\sigma'^2 + (n''-1)\sigma''^2]^2}{(n'-1)\sigma'^4 + (n''-1)\sigma''^4}$,

$$a = \frac{1}{\nu} (\frac{1}{n'} + \frac{1}{n''}) \frac{(n'-1)\sigma'^2 + (n''-1)\sigma''^2}{n'+n''-2} = \frac{1}{\nu} (\frac{\sigma'^2}{n'} + \frac{\sigma''^2}{n''}) \frac{1}{\mathcal{F}2},$$

c'est-à-dire : W^2 est approximativement distribuée comme $(\frac{\sigma'^2}{n'} + \frac{\sigma''^2}{n''}) \frac{1}{\mathcal{F}2} \frac{\chi^2[\nu]}{\nu}$.

Sous le modèle équivarié : $a' = a'' = \sigma^2 (\frac{1}{n'} + \frac{1}{n''}) \frac{1}{n'+n''-2}$, $\nu = n' + n'' - 2$, résultats conformes au résultat classique selon lequel W^2 est *exactement* distribuée comme

$$\sigma^2 (\frac{1}{n'} + \frac{1}{n''}) \frac{\chi^2[n'+n''-2]}{n'+n''-2}.$$