

ANNALES DE L'I. H. P., SECTION B

C. HEUCHENNE

L'analyse de covariance mise en question

Annales de l'I. H. P., section B, tome 15, n° 4 (1979), p. 315-332

http://www.numdam.org/item?id=AIHPB_1979__15_4_315_0

© Gauthier-Villars, 1979, tous droits réservés.

L'accès aux archives de la revue « *Annales de l'I. H. P., section B* » (<http://www.elsevier.com/locate/anihpb>) 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/>

L'analyse de covariance mise en question

par

C. HEUCHENNE (*)

SUMMARY. — Alternatives to usual analysis of covariance are proposed. Powers are compared.

Le propos de cet article est d'étudier des alternatives à l'analyse de covariance dont je rappelle le contexte.

Une variable y et une variable x sont observées dans q groupes repérés par l'indice i . Avec des échantillons de taille n_i , le modèle I le plus général pose que $y_{ij} = \alpha_i + \beta_i x_{ij} + e_{ij}$ pour $i = 1, 2, \dots, q$ et $j = 1, 2, \dots, n_i$. En outre, les e_{ij} sont des variables normales centrées indépendantes de variance commune σ^2 .

La théorie générale des modèles linéaires [2] montre sans peine que β_i est estimé par $\hat{\beta}_i = \frac{E_{xyi}}{E_{xxi}}$ et α_i par

$$\hat{\alpha}_i = \bar{y}_i - \hat{\beta}_i \bar{x}_i = \frac{T_{yi} - \hat{\beta}_i T_{xi}}{n_i}$$

où

$$\begin{aligned} T_{yi} &= \sum_{j=1}^{n_i} y_{ij} \quad , \quad T_{xi} = \sum_{j=1}^{n_i} x_{ij} \quad , \quad E_{xyi} = \sum_{j=1}^{n_i} (x_{ij} - \bar{x}_i)(y_{ij} - \bar{y}_i) \\ &= \left(\sum_{j=1}^{n_i} x_{ij} y_{ij} \right) - \frac{T_{xi} T_{yi}}{n_i} \quad \text{et} \quad E_{xxi} = \sum_{j=1}^{n_i} (x_{ij} - \bar{x}_i)^2 = \left(\sum_{j=1}^{n_i} x_{ij}^2 \right) - \frac{T_{xi}^2}{n_i} . \end{aligned}$$

(*) Adresse postale : Université de Liège au Sart-Tilman, par 4000 Liège 1, Belgique.

Le modèle I ne fait qu'exprimer la régression de y sur x dans chaque groupe. La somme des carrés correspondante est $\sum_{i=1}^q \left(\frac{E_{xyi}^2}{E_{xxi}} + \frac{T_{yi}^2}{n_i} \right)$ et la somme des carrés résiduelle est

$$S_r = \left(\sum_{i=1}^q \sum_{j=1}^{n_i} y_{ij}^2 \right) - \sum_{i=1}^q \left(\frac{E_{xyi}^2}{E_{xxi}} + \frac{T_{yi}^2}{n_i} \right).$$

Dans l'hypothèse d'égalité de tous les coefficients de régression β_i à un coefficient commun β , le modèle I devient le modèle II : $y_{ij} = \alpha_i + \beta x_{ij} + e_{ij}$. Les estimations y sont

$$\hat{\beta} = \frac{E_{xy}}{E_{xx}} \quad \text{et} \quad \alpha_i^* = \bar{y}_i - \hat{\beta} \bar{x}_i$$

où

$$E_{xy} = \sum_{i=1}^q E_{xyi} = \left(\sum_{i=1}^q \sum_{j=1}^{n_i} x_{ij} y_{ij} \right) - \sum_{i=1}^q \frac{T_{xi} T_{yi}}{n_i}$$

et

$$E_{xx} = \sum_{i=1}^q E_{xxi} = \left(\sum_{i,j} x_{ij}^2 \right) - \sum_i \frac{T_{xi}^2}{n_i}.$$

La réduction de la somme des carrés par rapport au modèle I est

$$S_p = \left(\sum_i \frac{E_{xyi}^2}{E_{xxi}} \right) - \frac{E_{xy}^2}{E_{xx}}.$$

L'épreuve du parallélisme des droites de régression est fondée sur $F_p = \frac{S_p/q - 1}{S_r/n - 2q}$ qui est une variable de Fisher-Snedecor dans le modèle I, centrée dans le modèle II.

Enfin, dans le modèle II, on peut éprouver l'égalité des ordonnées à l'origine α_i à une constante α , c'est-à-dire le modèle III : $y_{ij} = \alpha + \beta x_{ij} + e_{ij}$.

Les estimations y sont $\beta^* = \frac{S_{xy}}{S_{xx}}$ et $\alpha^* = \bar{y} - \beta^* \bar{x}$ où

$$S_{xy} = \sum_{i,j} (x_{ij} - \bar{x})(y_{ij} - \bar{y}) = \left(\sum_{i,j} x_{ij} y_{ij} \right) - \frac{T_x T_y}{n}$$

et

$$S_{xx} = \sum_{i,j} (x_{ij} - \bar{x})^2 = \left(\sum_{i,j} x_{ij}^2 \right) - \frac{T_x^2}{n}.$$

La somme des carrés du modèle III est $\frac{S_{xy}^2}{S_{xx}} + \frac{T_y^2}{n}$ et sa réduction par rapport au modèle II est $S_c = \left(\sum_i \frac{T_{yi}^2}{n_i} \right) - \frac{T_y^2}{n} + \frac{E_{xy}^2}{E_{xx}} - \frac{S_{xy}^2}{S_{xx}}$. L'épreuve de l'égalité des α_i est fondée sur $F_c = \frac{S_c/q - 1}{(S_p + S_r)/n - q - 1}$ qui est une variable de Fisher-Snedecor dans le modèle II, centrée dans le modèle III.

On peut encore envisager $S_t = S_p + S_c = \sum_i \left(\frac{E_{xyi}^2}{E_{xxi}} + \frac{T_{yi}^2}{n_i} \right) - \frac{S_{xy}^2}{S_{xx}} - \frac{T_y^2}{n}$, réduction de la somme des carrés du modèle I au modèle III. $F_t = \frac{S_t/2q - 2}{S_r/n - 2q}$, variable de Fisher-Snedecor dans le modèle I, centrée dans le modèle III, permet l'épreuve globale de l'identité des droites de régression (égalité des α_i et des β_i).

Une objection sérieuse à la méthode classique utilisant F_c est que, pour être exécutée correctement, elle postule l'égalité des coefficients de régression β_i . Or il est clair que cette hypothèse n'est strictement réalisée en pratique que rarement; et ce n'est évidemment pas l'absence de résultat significatif à l'épreuve de parallélisme qui prouve quoi que ce soit.

D'autre part, cette condition est une nécessité technique, non une contrainte d'interprétation comme beaucoup de praticiens semblent le croire. En effet, le but d'un chercheur est de comparer le critère y entre les groupes quand il a été ramené à un même niveau de la covariable x ; peu lui importe que cet ajustement se fasse différemment ou non dans chaque groupe.

Évidemment, dans le cas de non-parallélisme, se pose le problème du choix du niveau de la covariable où comparer le critère, problème inexistant si les régressions sont parallèles. La nature de la question peut imposer ce choix. Imaginons la comparaison d'une performance y selon le milieu socio-culturel en neutralisant l'influence de l'âge x . Il y a fort à parier que la vitesse de développement de y avec l'âge est différente suivant la catégorie socio-culturelle envisagée : voilà un exemple typique où le parallélisme des droites de régression n'est pas tenable. La comparaison des niveaux intrinsèques de performance se fera tout naturellement près de l'origine du temps, à une époque où seul jouerait le matériel génétique.

Il serait donc intéressant de disposer d'une épreuve qui compare les y ramenés à une même valeur a de x , sans postuler l'égalité des β_i . C'est l'épreuve du modèle IV : $y_{ij} = \gamma + \beta_i(x_{ij} - a) + e_{ij}$ à l'intérieur du modèle I' : $y_{ij} = \gamma_i + \beta_i(x_{ij} - a) + e_{ij}$.

La procédure est basée sur une statistique F_m et permet de répondre à la

question : les droites de régression ont-elles une même ordonnée γ pour $x = a$?

En posant $\gamma_i = \alpha_i + \beta_i a$, le modèle I' est identique au modèle I. Si les β_i sont égaux, les méthodes F_c et F_m éprouvent la même hypothèse. La méthode F_c sera naturellement plus puissante puisqu'elle se fonde sur une information supplémentaire dont ne peut tenir compte la méthode F_m .

Par contre, si les β_i sont inégaux, l'épreuve F_m garde un sens contrairement à l'épreuve F_c . Si la perte de puissance de F_m en cas de parallélisme est minimale et compensée par un gain appréciable sur F_c pour des β_i modérément différents, la méthode F_m prend de l'intérêt puisqu'elle évite une hypothèse souvent gratuite. C'est ce que je vais tenter de montrer.

Le modèle I' s'écrit sous forme matricielle

$$\vec{y} = X \begin{bmatrix} \vec{\gamma} \\ \vec{\beta} \end{bmatrix} + \vec{e} = A\vec{\gamma} + B\vec{\beta} + \vec{e}$$

où

$$X = \begin{bmatrix} 1 & 0 & \dots & 0 & x_{11} & -a & 0 & \dots & \dots & 0 \\ \dots & \dots \\ 1 & 0 & \dots & 0 & x_{1n_1} & -a & 0 & \dots & \dots & 0 \\ 0 & 1 & \dots & 0 & 0 & x_{21} & -a & \dots & \dots & 0 \\ \dots & \dots \\ 0 & 1 & \dots & 0 & 0 & x_{2n_2} & -a & \dots & \dots & 0 \\ \dots & \dots \\ 0 & 0 & \dots & 1 & 0 & 0 & 0 & \dots & x_{q1} & -a \\ \dots & \dots \\ 0 & 0 & \dots & 1 & 0 & 0 & 0 & \dots & x_{qn_q} & -a \end{bmatrix} = [A \mid B].$$

$\underbrace{\hspace{10em}}_q$
 $\underbrace{\hspace{10em}}_q$

On a $X'X = \begin{bmatrix} A'A & A'B \\ B'A & B'B \end{bmatrix} = \begin{bmatrix} N & T \\ T & S \end{bmatrix}$ où

$N = \text{diag}(n_i)$, $T = \text{diag}(t_i) = \text{diag}(T_{xi} - n_i a)$

et

$$S = \text{diag}(s_i) = \text{diag} \left(\sum_j (x_{ij} - a)^2 \right).$$

Comme

$$s_i = \sum_j (x_{ij} - \bar{x}_i)^2 + n_i(\bar{x}_i - a)^2 = E_{xxi} + \frac{t_i^2}{n_i},$$

N et S sont inversibles pour peu que les n_i et E_{xxi} soient positifs pour tout $i = 1, 2, \dots, q$. Par la formule de Fröbenius-Schur, on a

$$(X'X)^{-1} = \left[\begin{array}{c|c} N^{-1} + N^{-1}TE^{-1}TN^{-1} & -N^{-1}TE^{-1} \\ \hline -E^{-1}TN^{-1} & E^{-1} \end{array} \right]$$

où

$$E = S - TN^{-1}T = \text{diag}(E_{xxi}).$$

Les estimations des paramètres sont données par $(X'X)^{-1}X'y$. Or

$$A'y = (T_{yi}) = \vec{T}_y$$

et

$$B'y = \left(\sum_j (x_{ij} - a)y_{ij} \right) = \left(E_{xyi} + \frac{t_i T_{yi}}{n_i} \right) = \vec{E}_{xy} + TN^{-1}\vec{T}_y.$$

Le vecteur des β_i est estimé par

$$\hat{\beta} = -E^{-1}TN^{-1}A'y + E^{-1}B'y = E^{-1}\vec{E}_{xy},$$

donc $\hat{\beta}_i = \frac{E_{xyi}}{E_{xxi}}$ comme dans le modèle I.

Le vecteur des γ_i est estimé par

$$\begin{aligned} \hat{\gamma} &= N^{-1}A'y + N^{-1}TE^{-1}TN^{-1}A'y - N^{-1}TE^{-1}B'y \\ &= N^{-1}\vec{T}_y - N^{-1}T\hat{\beta}, \end{aligned}$$

soit explicitement $\hat{\gamma}_i = \bar{y}_i - \hat{\beta}_i(\bar{x}_i - a)$.

Par rapport à l'estimation des α_i du modèle I, on voit que $\hat{\gamma} = \hat{\alpha} + a\hat{\beta}$.

La somme des carrés du modèle I' est $y'y(X'X)^{-1}X'y$, soit

$$\begin{aligned} [y'A \mid y'B] \begin{bmatrix} \hat{\gamma} \\ \hat{\beta} \end{bmatrix} &= \vec{T}'_y(N^{-1}\vec{T}_y - N^{-1}T\hat{\beta}) + (\vec{E}_{xy} + TN^{-1}\vec{T}_y)\hat{\beta} \\ &= \vec{T}'_yN^{-1}\vec{T}_y + \vec{E}'_{xy}\hat{\beta} = \sum_i \left(\frac{T_{yi}^2}{n_i} + \frac{E_{xyi}^2}{E_{xxi}} \right). \end{aligned}$$

La somme des carrés et le résidu du modèle I' sont donc ceux du modèle I. Ils ne dépendent pas de a .

Dans le but de comparer les modèles I' et IV, notons que

$$(X'X)^{-1} = \left[\begin{array}{c|c} M^{-1} & -M^{-1}TS^{-1} \\ \hline -S^{-1}TM^{-1} & S^{-1} + S^{-1}TM^{-1}TS^{-1} \end{array} \right]$$

où $M = N - TS^{-1}T = \text{diag}(m_i)$ avec $m_i = n_i - \frac{t_i^2}{s_i} = \frac{n_i E_{xxi}}{s_i}$.

On a
$$\widehat{\gamma} = \mathbf{M}^{-1}\mathbf{A}'\vec{y} - \mathbf{M}^{-1}\mathbf{TS}^{-1}\mathbf{B}'\vec{y}$$

et

$$\widehat{\beta} = -\mathbf{S}^{-1}\mathbf{TM}^{-1}\mathbf{A}'\vec{y} + \mathbf{S}^{-1}\mathbf{B}'\vec{y} + \mathbf{S}^{-1}\mathbf{TM}^{-1}\mathbf{TS}^{-1}\mathbf{B}'\vec{y} = \mathbf{S}^{-1}\mathbf{B}'\vec{y} - \mathbf{S}^{-1}\widehat{\mathbf{T}}\widehat{\gamma}.$$

La somme des carrés du modèle I' est aussi

$$\begin{aligned} [\vec{y}'\mathbf{A} \mid \vec{y}'\mathbf{B}] \begin{bmatrix} \widehat{\gamma} \\ \widehat{\beta} \end{bmatrix} &= \vec{y}'\mathbf{A}\widehat{\gamma} + \vec{y}'\mathbf{B}(\mathbf{S}^{-1}\mathbf{B}'\vec{y} - \mathbf{S}^{-1}\widehat{\mathbf{T}}\widehat{\gamma}) \\ &= (\vec{y}'\mathbf{A}\mathbf{M}^{-1}\mathbf{M} - \vec{y}'\mathbf{B}\mathbf{S}^{-1}\mathbf{TM}^{-1}\mathbf{M})\widehat{\gamma} + \vec{y}'\mathbf{B}\mathbf{S}^{-1}\mathbf{B}'\vec{y} = \widehat{\gamma}'\widehat{\mathbf{M}}\widehat{\gamma} + \vec{y}'\mathbf{B}\mathbf{S}^{-1}\mathbf{B}'\vec{y}. \end{aligned}$$

Le modèle IV s'écrit

$$\vec{y} = \gamma\mathbf{A}\vec{\mathbf{1}} + \mathbf{B}\vec{\beta} + \vec{e}$$

où $\vec{\mathbf{1}} = (1, 1, \dots, 1)$ de dimension q . On a maintenant

$$\begin{aligned} \mathbf{X} &= [\mathbf{A}\vec{\mathbf{1}} \mid \mathbf{B}] \quad , \quad \mathbf{X}'\vec{y} = \begin{bmatrix} \vec{\mathbf{1}}'\mathbf{A}\vec{y} \\ \vdots \\ \mathbf{B}'\vec{y} \end{bmatrix}, \\ \mathbf{X}'\mathbf{X} &= \begin{bmatrix} \vec{\mathbf{1}}'\mathbf{A}'\mathbf{A}\vec{\mathbf{1}} & \vec{\mathbf{1}}'\mathbf{A}'\mathbf{B} \\ \mathbf{B}'\mathbf{A}\vec{\mathbf{1}} & \mathbf{B}'\mathbf{B} \end{bmatrix} = \begin{bmatrix} \vec{\mathbf{1}}'\mathbf{N}\vec{\mathbf{1}} & \vec{\mathbf{1}}'\mathbf{T} \\ \mathbf{T}\vec{\mathbf{1}} & \mathbf{S} \end{bmatrix}, \\ (\mathbf{X}'\mathbf{X})^{-1} &= \begin{bmatrix} \frac{1}{m} & -\frac{\vec{\mathbf{1}}'\mathbf{TS}^{-1}}{m} \\ -\frac{\mathbf{S}^{-1}\mathbf{T}\vec{\mathbf{1}}}{m} & \mathbf{S}^{-1} + \frac{\mathbf{S}^{-1}\mathbf{T}\vec{\mathbf{1}}\vec{\mathbf{1}}'\mathbf{TS}^{-1}}{m} \end{bmatrix} \end{aligned}$$

où

$$m = \vec{\mathbf{1}}'\mathbf{N}\vec{\mathbf{1}} - \vec{\mathbf{1}}'\mathbf{TS}^{-1}\mathbf{T}\vec{\mathbf{1}} = \vec{\mathbf{1}}'\mathbf{M}\vec{\mathbf{1}} = \sum_i m_i.$$

Le paramètre γ est estimé par

$$\widehat{\gamma} = \frac{\vec{\mathbf{1}}'\mathbf{A}'\vec{y}}{m} - \frac{\vec{\mathbf{1}}'\mathbf{TS}^{-1}\mathbf{B}'\vec{y}}{m} = \frac{\vec{\mathbf{1}}'\mathbf{M}\widehat{\gamma}}{m}.$$

Le vecteur des β_i est estimé par

$$\vec{\beta}^* = \frac{-\mathbf{S}^{-1}\mathbf{T}\vec{\mathbf{1}}\vec{\mathbf{1}}'\mathbf{A}'\vec{y}}{m} + \mathbf{S}^{-1}\mathbf{B}'\vec{y} + \frac{\mathbf{S}^{-1}\mathbf{T}\vec{\mathbf{1}}\vec{\mathbf{1}}'\mathbf{TS}^{-1}\mathbf{B}'\vec{y}}{m} = \mathbf{S}^{-1}\mathbf{B}'\vec{y} - \widehat{\gamma}\mathbf{S}^{-1}\mathbf{T}\vec{\mathbf{1}}.$$

La relation entre les estimateurs des β_i dans le modèle I (ou I') et le modèle IV est donc

$$\vec{\beta}^* - \widehat{\beta} = \mathbf{S}^{-1}\mathbf{T}(\widehat{\gamma} - \gamma\vec{\mathbf{1}}) = \mathbf{S}^{-1}\mathbf{T}\left(\mathbf{I} - \frac{\vec{\mathbf{1}}\vec{\mathbf{1}}'\mathbf{M}}{m}\right)\widehat{\gamma}.$$

La somme des carrés du modèle IV est

$$[\vec{y}'\vec{A}\vec{I} \mid \vec{y}'\vec{B}] \begin{bmatrix} \hat{\gamma} \\ \hat{\beta}^* \end{bmatrix} = \hat{\gamma}\vec{y}'\vec{A}\vec{I} + \vec{y}'\vec{B}(\mathbf{S}^{-1}\vec{B}'\vec{y} - \hat{\gamma}\mathbf{S}^{-1}\vec{T}\vec{I})$$

$$= \hat{\gamma}(\vec{y}'\vec{A}\mathbf{M}^{-1}\vec{M}\vec{I} - \vec{y}'\vec{B}\mathbf{S}^{-1}\vec{T}\mathbf{M}^{-1}\vec{M}\vec{I}) + \vec{y}'\vec{B}\mathbf{S}^{-1}\vec{B}'\vec{y} = m\hat{\gamma}^2 + \vec{y}'\vec{B}\mathbf{S}^{-1}\vec{B}'\vec{y}$$

Finalement la réduction du modèle I' au modèle IV est $S_m = \hat{\gamma}'\vec{M}\hat{\gamma} - m\hat{\gamma}^2$.

On peut encore l'écrire

$$\hat{\gamma}'\left(\mathbf{M} - \frac{\vec{M}\vec{I}\vec{I}'\vec{M}}{m}\right)\hat{\gamma} = \left(\sum_i m_i \hat{\gamma}_i^2\right) - \frac{1}{m} \left(\sum_i m_i \hat{\gamma}_i\right)^2 = \frac{1}{m} \sum_{i < k} m_i m_k (\hat{\gamma}_i - \hat{\gamma}_k)^2.$$

Comme il se doit, S_m ne s'annule que lorsque tous les $\hat{\gamma}_i$ sont égaux entre eux. L'hypothèse $\gamma_i = \gamma$ pour tout i s'éprouve par $F_m = \frac{S_m/q - 1}{S_r/n - 2q}$ qui est une variable de Fisher-Snedecor dans le modèle I, centrée dans le modèle IV.

Si a n'est pas connu *a priori*, mais dépend des x_{ij} , une petite difficulté algorithmique se présente : il faut garder en mémoire $\bar{x}_i, \bar{y}_i, n_i, E_{xxi}$ et E_{xyi} (ou $\hat{\beta}_i$) jusqu'à la fin de l'introduction des données. C'est seulement alors qu'on pourra calculer successivement

$$\bar{x}_i - a, \hat{\gamma}_i = \bar{y}_i - \hat{\beta}_i(\bar{x}_i - a), \frac{1}{m_i} = \frac{1}{n_i} + \frac{(\bar{x}_i - a)^2}{E_{xxi}}$$

(car $\mathbf{M}^{-1} = \mathbf{N}^{-1} + \mathbf{N}^{-1}\vec{T}\mathbf{E}^{-1}\vec{T}\mathbf{N}^{-1}$), enfin S_m .

Comme je l'ai indiqué précédemment, le choix de a peut être imposé par l'application. Autrement il semble naturel de prendre $a = \bar{x}$. Cependant, pour $q = 2$, la valeur $a = \frac{E_{xx2}\bar{x}_1 + E_{xx1}\bar{x}_2}{E_{xx}}$ est intéressante. On vérifie qu'avec elle

$$\hat{\gamma}_1 - \hat{\gamma}_2 = \bar{y}_1 - \bar{y}_2 - \frac{E_{xy}}{E_{xx}}(\bar{x}_1 - \bar{x}_2) \quad \text{et} \quad \frac{m_1 m_2}{m} = \frac{n_1 n_2 E_{xx}}{n S_{xx}},$$

d'où résulte

$$S_m = \frac{m_1 m_2}{m} (\hat{\gamma}_1 - \hat{\gamma}_2)^2 = \frac{n_1 n_2 (\bar{y}_1 - \bar{y}_2)^2}{n} + \frac{E_{xy}^2}{E_{xx}} - \frac{S_{xy}^2}{S_{xx}} = S_c$$

et

$$F_m = \frac{S_c(n - 4)}{S_r} = \frac{F_c(F_p + n - 4)}{n - 3}.$$

Puisque $F_m > F_c$ quand $F_p > 1$, on entrevoit déjà par là que, lorsque le parallélisme est suspect, l'épreuve F_m peut être plus puissante que l'épreuve F_c .

Pour $q > 2$, il n'existe pas de valeur de a (indépendante des y_{ij}) qui iden-

tifie S_m à S_c . La valeur $a = \left(\sum E_{xxi}^{-1} \right)^{-1} \sum E_{xxi}^{-1} \bar{x}_i$, qui minimise la trace de M^{-1} et qui est l'extension de la valeur ci-dessus à $q > 2$, ne semble pas avoir de propriété spéciale, mais la question reste ouverte.

Une autre procédure a été suggérée ([1], p. 275). Alors que dans le modèle II, l'estimation du coefficient de régression commun est $\hat{\beta} = \frac{E_{xy}}{E_{xx}}$, dans le modèle III, β est estimé par $\beta^* = \frac{S_{xy}}{S_{xx}}$. Or la meilleure estimation est $\hat{\beta}$ si les α_i sont inégaux. Ceci a même conduit à affirmer ([4], p. 764) que l'analyse de covariance exigeait $\hat{\beta} = \beta^*$! Mais le paradoxe n'est qu'apparent puisque pour éprouver l'hypothèse $y_{ij} = \alpha + \beta x_{ij} + e_{ij}$, il faut bien estimer β dans ce modèle.

Il n'empêche qu'on peut tenter naïvement de corriger ce défaut en faisant jouer à $\hat{\beta}$ dans S_c le rôle normalement tenu par β^* . On est alors conduit à une somme de carrés notée T_{yyA} en [1]. Comme il est dit à cet endroit, la méthode est incorrecte, au mieux approximative, puisqu'elle fait jouer à l'estimation d'échantillon $\hat{\beta}$ le rôle du paramètre β dans le modèle II.

Cette procédure est cependant alléchante parce que plus puissante : T_{yyA} est systématiquement supérieur à S_c , noté T_{yyR} en [1]. Elle reste intéressante quand on peut connaître une bonne estimation *a priori* b du coefficient de régression commun β , indépendante de l'analyse de covariance proprement dite.

Posons $z_{ij} = y_{ij} - bx_{ij}$. L'épreuve du modèle

$$y_{ij} = \alpha + bx_{ij} + e_{ij} \quad \text{ou} \quad z_{ij} = \alpha + e_{ij} \quad \text{ou} \quad \vec{z} = \alpha \vec{A} \vec{1} + \vec{e}$$

dans le modèle

$$y_{ij} = \alpha_i + bx_{ij} + e_{ij} \quad \text{ou} \quad z_{ij} = \alpha_i + e_{ij} \quad \text{ou} \quad \vec{z} = \vec{A} \vec{\alpha} + \vec{e}$$

n'est autre qu'une analyse de variance simple des z_{ij} . L'épreuve passe par

$$F_b = \frac{S_b/q - 1}{S_e/n - q}$$

où

$$\begin{aligned} S_b &= \vec{z}' \left(\vec{A} \vec{N}^{-1} \vec{A}' - \frac{\vec{A} \vec{1} \vec{1}' \vec{A}'}{n} \right) \vec{z} \\ &= \left(\sum_i \frac{T_{yi}^2}{n_i} - \frac{T_y^2}{n} \right) - 2b \left(\sum_i \frac{T_{xi} T_{yi}}{n_i} - \frac{T_x T_y}{n} \right) + b^2 \left(\sum_i \frac{T_{xi}^2}{n_i} - \frac{T_x^2}{n} \right) \end{aligned}$$

et

$$S_e = \vec{z}'(\mathbf{I} - \mathbf{A}\mathbf{N}^{-1}\mathbf{A}')\vec{z} \\ = \left(\sum_{i,j} y_{ij}^2 - \sum_i \frac{\mathbf{T}_{yi}^2}{n_i} \right) - 2b \left(\sum_{i,j} x_{ij}y_{ij} - \sum_i \frac{\mathbf{T}_{xi}\mathbf{T}_{yi}}{n_i} \right) + b^2 \left(\sum_{i,j} x_{ij}^2 - \sum_i \frac{\mathbf{T}_{xi}^2}{n_i} \right).$$

Pour comparer les puissances des différentes épreuves, nous calculerons les valeurs attendues de leurs statistiques en termes des paramètres α_i et β_i du modèle I, le plus général.

Le modèle I peut être considéré lui-même comme le cas particulier du modèle I' obtenu pour $a = 0$. J'affecterai d'un indice 0 tous les symboles dépendant de a quand ils sont pris avec $a = 0$; ainsi $\mathbf{T}_0 = \mathbf{A}'\mathbf{B}_0 = \text{diag}(\mathbf{T}_{xi})$, $\mathbf{S}_0 = \mathbf{B}'_0\mathbf{B}_0 = \text{diag} \left(\sum_j x_{ij}^2 \right)$, etc.

Par ([2], théor. 6.1), dans le modèle I, $(\hat{\alpha}, \hat{\beta})$ est un vecteur multinormal de moyenne $(\vec{\alpha}, \vec{\beta})$ et de matrice des covariances

$$\sigma^2(\mathbf{X}'_0, \mathbf{X}_0)^{-1} = \sigma^2 \left[\begin{array}{c|c} \mathbf{M}_0^{-1} & -\mathbf{N}^{-1}\mathbf{T}_0\mathbf{E}^{-1} \\ \hline -\mathbf{E}^{-1}\mathbf{T}_0\mathbf{N}^{-1} & \mathbf{E}^{-1} \end{array} \right]$$

En répliquant la procédure générale développée lors du passage du modèle I' au modèle IV, on verrait que la réduction de la somme des carrés du modèle I au modèle II est

$$S_p = \hat{\beta}' \left(\mathbf{E} - \frac{\mathbf{E}\vec{1}\vec{1}'\mathbf{E}}{\vec{1}'\mathbf{E}\vec{1}} \right) \hat{\beta}.$$

Par ([2], théor. 4.20), S_p/σ^2 est une variable chi-carré à $(q - 1)$ libertés dont la non-centralité est

$$p(\vec{\beta}) = \frac{\vec{\beta}'}{\sigma} \left(\mathbf{E} - \frac{\mathbf{E}\vec{1}\vec{1}'\mathbf{E}}{\mathbf{E}_{xx}} \right) \frac{\vec{\beta}}{\sigma}.$$

Cette non-centralité s'annule si et seulement si $\vec{\beta}$ est colinéaire à $\vec{1}$.

On notera que la non-centralité considérée ici est celle de [3] ou [4], c'est-à-dire le double de celle de [2].

La matrice des covariances de $(\vec{\alpha}^*, \hat{\beta})$ dans le modèle II est

$$\sigma^2 \left[\begin{array}{c|c} \mathbf{D}^{-1} & -\frac{\mathbf{N}^{-1}\mathbf{T}_0\vec{1}}{\mathbf{E}_{xx}} \\ \hline -\frac{\vec{1}'\mathbf{T}_0\mathbf{N}^{-1}}{\mathbf{E}_{xx}} & \frac{1}{\mathbf{E}_{xx}} \end{array} \right] \quad \text{où} \quad \mathbf{D} = \mathbf{N} - \frac{\mathbf{T}_0\vec{1}\vec{1}'\mathbf{T}_0}{\vec{1}'\mathbf{S}_0\vec{1}}.$$

Le schéma général de réduction montre que, pour le passage du modèle II au modèle III,

$$S_c = \vec{\alpha}^{*'} \left(\mathbf{D} - \frac{\vec{\mathbf{D}} \vec{\mathbf{I}}' \vec{\mathbf{D}}}{\vec{\mathbf{I}}' \vec{\mathbf{D}} \vec{\mathbf{I}}} \right) \vec{\alpha}^*$$

qui est encore une forme quadratique en $(\hat{\alpha}, \hat{\beta})$ puisque

$$\alpha_i^* = \bar{y}_i - \hat{\beta} \bar{x}_i = \hat{\alpha}_i + \bar{x}_i (\hat{\beta}_i - \hat{\beta}) = \hat{\alpha}_i + \frac{\mathbf{T}_{xi}}{n_i} \left(\hat{\beta}_i - \frac{\sum_k \mathbf{E}_{xxk} \hat{\beta}_k}{\mathbf{E}_{xx}} \right)$$

soit

$$\vec{\alpha}^* = \hat{\alpha} + \mathbf{N}^{-1} \mathbf{T}_0 \left(\mathbf{I} - \frac{\vec{\mathbf{I}} \vec{\mathbf{I}}' \mathbf{E}}{\mathbf{E}_{xx}} \right) \hat{\beta}.$$

De cette dernière égalité résulte la multinormalité de $\vec{\alpha}^*$ ([2], théor. 3.22). On vérifie sans peine que la matrice

$$\left(\mathbf{D} - \frac{\vec{\mathbf{D}} \vec{\mathbf{I}}' \vec{\mathbf{D}}}{\vec{\mathbf{I}}' \vec{\mathbf{D}} \vec{\mathbf{I}}} \right) \mathbf{D}^{-1}$$

est idempotente. Par ([2], théor. 4.9), S_c/σ^2 est une variable chi-carré à $(q - 1)$ libertés dont la non-centralité est

$$\frac{\mathcal{E}(\vec{\alpha}^{*'})}{\sigma} \left(\mathbf{D} - \frac{\vec{\mathbf{D}} \vec{\mathbf{I}}' \vec{\mathbf{D}}}{\vec{\mathbf{I}}' \vec{\mathbf{D}} \vec{\mathbf{I}}} \right) \frac{\mathcal{E}(\vec{\alpha}^*)}{\sigma}.$$

Cette non-centralité s'annule si et seulement si

$$\hat{\alpha} + \mathbf{N}^{-1} \mathbf{T}_0 \left(\mathbf{I} - \frac{\vec{\mathbf{I}} \vec{\mathbf{I}}' \mathbf{E}}{\mathbf{E}_{xx}} \right) \hat{\beta}$$

est colinéaire à $\vec{\mathbf{I}}$.

Dans le modèle I', $\hat{\gamma}$ est un vecteur multinormal de moyenne $\vec{\gamma}$ et de matrice de covariances $\sigma^2 \mathbf{M}^{-1}$. Par ([2], théor. 4.20), puisque

$$S_m = \hat{\gamma}' \left(\mathbf{M} - \frac{\vec{\mathbf{M}} \vec{\mathbf{I}}' \vec{\mathbf{M}}}{\vec{\mathbf{I}}' \vec{\mathbf{M}} \vec{\mathbf{I}}} \right) \hat{\gamma},$$

S_m/σ^2 est une variable chi-carré à $(q - 1)$ libertés dont la non-centralité est

$$\frac{\vec{\gamma}'}{\sigma} \left(\mathbf{M} - \frac{\vec{\mathbf{M}} \vec{\mathbf{I}}' \vec{\mathbf{M}}}{\vec{\mathbf{I}}' \vec{\mathbf{M}} \vec{\mathbf{I}}} \right) \frac{\vec{\gamma}}{\sigma} = \frac{(\vec{\alpha} + a\vec{\beta})'}{\sigma} \left(\mathbf{M} - \frac{\vec{\mathbf{M}} \vec{\mathbf{I}}' \vec{\mathbf{M}}}{m} \right) \frac{(\vec{\alpha} + a\vec{\beta})}{\sigma} = m(\vec{\alpha}, \vec{\beta}, a).$$

Dans le modèle I, \vec{y} est un vecteur multinormal de moyenne $\vec{\mu} = A\vec{\alpha} + B_0\vec{\beta}$ et de matrice des covariances $\sigma^2 I$. On a

$$\frac{\vec{y}'\vec{y}}{\sigma^2} = \frac{1}{\sigma^2} \left(S_r + S_p + S_c + \frac{T_y^2}{n} + \frac{S_{xy}^2}{S_{xx}} \right).$$

Chaque composante est une forme quadratique en \vec{y} et il est bien connu que leurs rangs respectifs sont $n - 2q$, $q - 1$, $q - 1$, 1 et 1 de somme n . Par le théorème de Cochran ([2], théor. 4.14), ce sont des chi-carré indépendants dont les non-centralités s'ajoutent. La non-centralité totale est

$$\frac{\vec{\mu}'\vec{\mu}}{\sigma^2} = \frac{1}{\sigma^2} (\vec{\alpha}'N\vec{\alpha} + 2\vec{\alpha}'T_0\vec{\beta} + \vec{\beta}'S_0\vec{\beta})$$

qui est aussi celle de $\frac{1}{\sigma^2} \left(S_r + \frac{T_y^2}{n} + \frac{S_{xy}^2}{S_{xx}} \right)$ puisque S_r/σ^2 est centré.

$$\frac{T_y^2}{n} = \frac{(\vec{1}'A'\vec{y})^2}{n} = \vec{y}' \left(\frac{A\vec{1}\vec{1}'A'}{n} \right) \vec{y}.$$

Comme $\frac{A\vec{1}\vec{1}'A'}{n}$ est idempotente, par ([2], théor. 4.7), la non-centralité

de $\frac{T_y^2}{n\sigma^2}$ est

$$\begin{aligned} \frac{\vec{\mu}' \left(\frac{A\vec{1}\vec{1}'A'}{n} \right) \vec{\mu}}{\sigma} &= \frac{(\vec{1}'A'\vec{\mu})^2}{n\sigma^2} = \frac{(\vec{1}'N\vec{\alpha} + \vec{1}'T_0\vec{\beta})^2}{n\sigma^2} \\ S_{xy} &= \left(\sum_{i,j} x_{ij}y_{ij} \right) - \frac{T_x T_y}{n} = \vec{1}'B_0'\vec{y} - \frac{T_x}{n} (\vec{1}'A'\vec{y}), \end{aligned}$$

donc

$$\frac{S_{xy}^2}{S_{xx}} = \frac{1}{S_{xx}} \left(\vec{y}' \left(B_0\vec{1} - \frac{T_x}{n} A\vec{1} \right) \left(\vec{1}'B_0' - \frac{T_x}{n} \vec{1}'A' \right) \vec{y} \right).$$

La matrice de cette forme est idempotente car

$$\left(\vec{1}'B_0' - \frac{T_x}{n} \vec{1}'A' \right) \left(B_0\vec{1} - \frac{T_x}{n} A\vec{1} \right) = \vec{1}'S_0\vec{1} - 2 \frac{T_x}{n} \vec{1}'T_0\vec{1} + \frac{T_x^2}{n^2} \vec{1}'N\vec{1} = S_{xx}.$$

Par ([2], théor. 4.7), la non-centralité de $\frac{S_{xy}^2}{S_{xx}\sigma^2}$ est

$$\begin{aligned} \frac{\vec{\mu}' \left(B_0\vec{1} - \frac{T_x}{n} A\vec{1} \right) \left(\vec{1}'B_0' - \frac{T_x}{n} \vec{1}'A' \right) \vec{\mu}}{S_{xx}\sigma^2} &= \left(\vec{1}'B_0'\vec{\mu} - \frac{T_x}{n} \vec{1}'A'\vec{\mu} \right)^2 \frac{1}{S_{xx}\sigma^2} \\ &= \frac{1}{S_{xx}\sigma^2} \left(\vec{1}'T_0\vec{\alpha} + \vec{1}'S_0\vec{\beta} - \frac{T_x}{n} (\vec{1}'N\vec{\alpha} + \vec{1}'T_0\vec{\beta}) \right)^2 \end{aligned}$$

En posant $d = \vec{1}'\left(T_0 - \frac{T_x}{n}N\right)\vec{\alpha} + \vec{1}'\left(S_0 - \frac{T_x}{n}T_0\right)\vec{\beta}$, la non-centralité de $S_t = S_p + S_c$ est finalement

$$t(\vec{\alpha}, \vec{\beta}) = \frac{1}{\sigma^2} \left(\vec{\alpha}'N\vec{\alpha} + 2\vec{\alpha}'T_0\vec{\beta} + \vec{\beta}'S_0\vec{\beta} - \frac{d^2}{S_{xx}} - \frac{1}{n}(\vec{1}'N\vec{\alpha} + \vec{1}'T_0\vec{\beta})^2 \right).$$

Cette non-centralité s'annule quand le modèle I coïncide avec le modèle III, c'est-à-dire quand $\vec{\alpha}$ d'une part, $\vec{\beta}$ d'autre part, sont colinéaires à $\vec{1}$. La non-centralité de S_c est encore $t(\vec{\alpha}, \vec{\beta}) - p(\vec{\beta})$.

Il est facile de voir que le modèle III' : $y_{ij} = \gamma + \beta(x_{ij} - a) + e_{ij}$ a la même somme de carrés que le modèle III. Il en résulte que la réduction du modèle I' au modèle III' n'est autre que S_t , réduction du modèle I au modèle III, et ne dépend pas de a .

Appliquant encore le même schéma général, on verrait que la réduction du modèle IV au modèle III', correspondant à l'hypothèse $\beta_i = \beta$ pour tout i dans $y_{ij} = \gamma + \beta_i(x_{ij} - a) + e_{ij}$ est la forme quadratique

$$\vec{\beta}^{*'} \left(F - \frac{\vec{F}\vec{1}\vec{1}'\vec{F}}{\vec{1}'\vec{F}\vec{1}} \right) \vec{\beta}^* \quad \text{où} \quad F = S - \frac{\vec{T}\vec{1}\vec{1}'\vec{T}}{n},$$

la matrice des covariances de $(\hat{\gamma}, \hat{\beta}^*)$ dans le modèle IV étant

$$\left[\begin{array}{c|c} \frac{1}{m} & -\frac{\vec{1}'\vec{T}S^{-1}}{m} \\ \hline -\frac{S^{-1}\vec{T}\vec{1}}{m} & F^{-1} \end{array} \right]$$

Cette réduction est encore une forme quadratique en $(\hat{\alpha}, \hat{\beta})$ puisque

$$\vec{\beta}^* = \hat{\beta} + S^{-1}\vec{T}\left(\vec{I} - \frac{\vec{1}\vec{1}'\vec{M}}{m}\right)\hat{\gamma}$$

Comme elle vaut $S_t - S_m$, sa non-centralité est $t(\vec{\alpha}, \vec{\beta}) - m(\vec{\alpha}, \vec{\beta}, a)$ par ([2], théor. 4.22).

$\left(F - \frac{\vec{F}\vec{1}\vec{1}'\vec{F}}{\vec{1}'\vec{F}\vec{1}}\right)F^{-1}$ étant une matrice idempotente, par ([2], théor. 4.9),

$$t(\vec{\alpha}, \vec{\beta}) - m(\vec{\alpha}, \vec{\beta}, a) = \frac{\mathcal{E}(\hat{\beta}^*)'}{\sigma} \left(F - \frac{\vec{F}\vec{1}\vec{1}'\vec{F}}{\vec{1}'\vec{F}\vec{1}} \right) \frac{\mathcal{E}(\hat{\beta}^*)}{\sigma}$$

qui ne s'annule que lorsque $\vec{\beta} + S^{-1}T\left(I - \frac{\vec{1}\vec{1}'M}{m}\right)(\vec{\alpha} + a\vec{\beta})$ est colinéaire à $\vec{1}$.

Dans le modèle I, les z_{ij} sont des variables normales indépendantes de moyennes $\mathcal{E}(y_{ij}) - bx_{ij} = \alpha_i + (\beta_i - b)x_{ij}$ et de variance σ^2 .

$$\text{Donc } \mathcal{E}(\vec{z}) = A\vec{\alpha} + B_0(\vec{\beta} - b\vec{1}).$$

Par ([2], théor. 4.7), la non-centralité de $S_e/\sigma^2 = \frac{\vec{z}'}{\sigma}(I - AN^{-1}A')\frac{\vec{z}}{\sigma}$, de matrice visiblement idempotente, est

$$\begin{aligned} & \frac{1}{\sigma^2}(\vec{\alpha}'A' + (\vec{\beta} - b\vec{1})'B_0')(I - AN^{-1}A')(A\vec{\alpha} + B_0(\vec{\beta} - b\vec{1})) \\ &= \frac{1}{\sigma^2}(\vec{\alpha}'(N - NN^{-1}N)\vec{\alpha} + 2\vec{\alpha}'(T_0 - NN^{-1}T_0)(\vec{\beta} - b\vec{1}) \\ &+ (\vec{\beta} - b\vec{1})'(S_0 - T_0N^{-1}T_0)(\vec{\beta} - b\vec{1})) = \frac{1}{\sigma^2}(\vec{\beta} - b\vec{1})'E(\vec{\beta} - b\vec{1}). \end{aligned}$$

Sous cette forme, on voit que la non-centralité de S_e/σ^2 ne s'annule que pour $\vec{\beta} = b\vec{1}$.

En posant $e(\vec{\beta}, b) = \frac{E_{xx}}{\sigma^2} \left(\frac{\vec{1}'E\vec{\beta}}{E_{xx}} - b \right)^2$, la non-centralité de S_e/σ^2 s'écrit $p(\vec{\beta}) + e(\vec{\beta}, b)$.

Par ([2], théor. 4.7), la non-centralité de $\frac{S_b}{\sigma^2} = \frac{\vec{z}'}{\sigma} \left(AN^{-1}A' - \frac{A\vec{1}\vec{1}'A'}{n} \right) \frac{\vec{z}}{\sigma}$ de matrice visiblement idempotente est

$$\begin{aligned} & \frac{1}{\sigma^2} \left(\vec{\alpha}'A' + (\vec{\beta} - b\vec{1})'B_0' \right) \left(AN^{-1}A' - \frac{A\vec{1}\vec{1}'A'}{n} \right) (A\vec{\alpha} + B_0(\vec{\beta} - b\vec{1})) \\ &= \frac{1}{\sigma^2} (\vec{\alpha}' + (\vec{\beta} - b\vec{1})'T_0N^{-1}) \left(N - \frac{N\vec{1}\vec{1}'N}{n} \right) (\vec{\alpha} + N^{-1}T_0(\vec{\beta} - b\vec{1})). \end{aligned}$$

Sous cette forme, on voit que la non-centralité de S_b/σ^2 ne s'annule que lorsque $\vec{\alpha} + N^{-1}T_0(\vec{\beta} - b\vec{1})$ est colinéaire à $\vec{1}$.

En posant $s(\vec{\alpha}, \vec{\beta}, b) = \frac{S_{xx}}{\sigma^2} \left(\frac{d}{S_{xx}} - b \right)^2$, un calcul un peu long montre que la non-centralité de S_b/σ^2 est $t(\vec{\alpha}, \vec{\beta}) - p(\vec{\beta}) + s(\vec{\alpha}, \vec{\beta}, b) - e(\vec{\beta}, b)$.

Dans le cadre du modèle I, F_c et F_b ne sont généralement pas des variables de Fisher-Snedecor. En effet, la non-centralité $p(\vec{\beta})$ du dénominateur $(S_p + S_r)/\sigma^2$ de F_c ne s'annule que si les β_i sont égaux; la non-centralité $p(\vec{\beta}) + e(\vec{\beta}, b)$ du dénominateur S_e/σ^2 de F_b ne s'annule qu'en cas de parallélisme

et si b vaut la pente commune. En outre, en l'absence de parallélisme, la non-centralité de leur numérateur S_c ou S_b s'annule dans des conditions curieuses.

Le résultat suivant permet d'évaluer la valeur attendue de variables

telles F_c et F_b . La moyenne de $F = \frac{\chi^2(v_1, \lambda_1)/v_1}{\chi^2(v_2, \lambda_2)/v_2}$ est $\frac{1 + \frac{\lambda_1}{v_1}}{1 - \frac{\lambda_2}{v_2}}$ si $\lambda_2 = 0$

(moyenne d'un F non centré) ([4], p. 832), appartient à l'intervalle

$$\left[\frac{1 + \frac{\lambda_1}{v_1}}{1 + \frac{\lambda_2 - 2}{v_2}}, \frac{1 + \frac{\lambda_1}{v_1}}{1 - \frac{2}{v_2} + \frac{(v_2 - 2)\lambda_2}{v_2^2}} \right] \quad \text{si } \lambda_2 > 0,$$

donc vaut approximativement $\frac{1 + \frac{\lambda_1}{v_1}}{1 - \frac{2}{v_2} + \frac{(v_2 - 1)\lambda_2}{v_2^2}}$ si v_2 est grand.

Par intégration sur la densité de probabilité ([3], p. 252), la moyenne pour $v_2 > 2$ est

$$\frac{v_2 \exp\left(-\frac{\lambda_1 + \lambda_2}{2}\right)}{v_1} \sum_{r=0}^{\infty} \sum_{s=0}^{\infty} \frac{(\lambda_1/2)^r (\lambda_2/2)^s}{r! s! \mathbf{B}\left(\frac{v_1}{2} + r, \frac{v_2}{2} + s\right)} \int_0^{\infty} \frac{u^{\frac{v_1}{2} + r}}{(1+u)^{\frac{v_1 + v_2}{2} + r + s}} du$$

L'intégrale valant $\mathbf{B}\left(\frac{v_1}{2} + r + 1, \frac{v_2}{2} + s + 1\right)$, c'est encore

$$\begin{aligned} \frac{v_2 \exp\left(-\frac{\lambda_1 + \lambda_2}{2}\right)}{v_1} \sum_{r=0}^{\infty} \left(\frac{v_1}{2} + r\right) \frac{(\lambda_1/2)^r}{r!} \sum_{s=0}^{\infty} \frac{(\lambda_2/2)^s}{s! \left(\frac{v_2}{2} + s - 1\right)} \\ = \left(1 + \frac{\lambda_1}{v_1}\right) \frac{v_2}{2} \exp\left(-\frac{\lambda_2}{2}\right) \sum_{s=0}^{\infty} \frac{(\lambda_2/2)^s}{s! \left(\frac{v_2}{2} + s - 1\right)} \end{aligned}$$

Pour a quelconque,

$$\begin{aligned} \left(\frac{v_2}{2} + a \frac{\lambda_2}{2} - 1\right) \sum_{s=0}^{\infty} \frac{(\lambda_2/2)^s}{s! \left(\frac{v_2}{2} + s - 1\right)} \\ = \exp\left(\frac{\lambda_2}{2}\right) + \frac{\lambda_2}{2} \sum_{s=0}^{\infty} \frac{(\lambda_2/2)^s}{s!} \left(\frac{a}{\frac{v_2}{2} + s - 1} - \frac{1}{\frac{v_2}{2} + s}\right) \end{aligned}$$

En prenant $a = 1$, puis $a = \frac{v_2 - 2}{v_2}$, on en tire

$$\left(\frac{v_2}{2} + \frac{\lambda_2}{2} - 1\right) \sum_{s=0}^{\infty} \frac{(\lambda_2/2)^s}{s! \left(\frac{v_2}{2} + s - 1\right)} > \exp\left(\frac{\lambda_2}{2}\right) \\ > \left(\frac{v_2}{2} + \frac{(v_2 - 2)\lambda_2}{2v_2} - 1\right) \sum_{s=0}^{\infty} \frac{(\lambda_2/2)^s}{s! \left(\frac{v_2}{2} + s - 1\right)}$$

avec la double égalité si $\lambda_2 = 0$. Il en résulte

$$\frac{1 + \frac{\lambda_1}{v_1}}{1 - \frac{2}{v_2} + \frac{\lambda_2}{v_2}} < \mathcal{E}(F) < \frac{1 + \frac{\lambda_1}{v_1}}{1 - \frac{2}{v_2} + \frac{(v_2 - 2)\lambda_2}{v_2^2}}$$

avec la double égalité si $\lambda_2 = 0$.

Dans le modèle I, on a ainsi

$$\mathcal{E}(F_p) = \frac{1 + \frac{p(\vec{\beta})}{q-1}}{1 - \frac{n-2q}{2}}, \quad \mathcal{E}(F_m) = \frac{1 + \frac{m(\vec{\alpha}, \vec{\beta}, a)}{q-1}}{1 - \frac{n-2q}{2}}, \quad \mathcal{E}(F_t) = \frac{1 + \frac{t(\vec{\alpha}, \vec{\beta})}{2(q-1)}}{1 - \frac{n-2q}{2}}$$

$$\mathcal{E}(F_c) \sim \frac{1 + \frac{t(\vec{\alpha}, \vec{\beta}) - p(\vec{\beta})}{q-1}}{1 - \frac{2}{n-q-1} + \frac{(n-q-2)p(\vec{\beta})}{(n-q-1)^2}}$$

$$\mathcal{E}(F_b) \sim \frac{1 + \frac{t(\vec{\alpha}, \vec{\beta}) - p(\vec{\beta}) + s(\vec{\alpha}, \vec{\beta}, b) - e(\vec{\beta}, b)}{q-1}}{1 - \frac{2}{n-q} + \frac{(n-q-1)(p(\vec{\beta}) + e(\vec{\beta}, b))}{(n-q)^2}}$$

Dans la discussion qui suit, nous supposons n grand devant q ce qui a pour effet que les degrés de liberté $v_2 = n - 2q$, $n - q - 1$, $n - q$ des dénominateurs sont grands et peu différents entre eux.

La qualité de l'épreuve F_b dépend essentiellement du choix de b . Si ce

dernier peut être pris très voisin de $\frac{\vec{1}'E\vec{\beta}}{E_{xx}} = \frac{\sum_i E_{xxi}\beta_i}{E_{xx}}$, $e(\vec{\beta}, b) \sim 0$, les dénominateurs de $\mathcal{E}(F_c)$ et $\mathcal{E}(F_b)$ sont presque identiques et le numérateur de

$\mathcal{E}(F_b)$ est supérieur à celui de $\mathcal{E}(F_c)$. L'avantage de F_b sur F_c croît avec $s\left(\vec{\alpha}, \vec{\beta}, \frac{\vec{1}'E\vec{\beta}}{E_{xx}}\right)$, c'est-à-dire avec $\left|\frac{d}{S_{xx}} - \frac{\vec{1}'E\vec{\beta}}{E_{xx}}\right|$ qui est dans la population ce qu'est $\left|\frac{S_{xy}}{S_{xx}} - \frac{E_{xy}}{E_{xx}}\right| = |\beta^* - \hat{\beta}|$ dans l'échantillon. On a intérêt à employer F_b quand les deux estimations β^* et $\hat{\beta}$ sont fort différentes ce que l'intuition indiquait.

Rappelons toutefois que F_b , tout autant que F_c , n'est correcte qu'en cas de parallélisme. Si $\vec{\beta} = \beta\vec{1}$, on a

$$\frac{d}{S_{xx}} = \frac{1}{S_{xx}} \left(\vec{1}' \left(T_0 - \frac{T_x}{n} N \right) \vec{\alpha} + \beta \vec{1}' \left(S_0 - \frac{T_x}{n} T_0 \right) \vec{1} \right) = \beta + \frac{\vec{1}' \left(T_0 - \frac{T_x}{n} N \right) \vec{\alpha}}{S_{xx}}.$$

Dans ce cas, pour $b = \frac{\vec{1}'E\vec{\beta}}{E_{xx}} = \beta$, le gain de F_b sur F_c se marque par la magnitude de

$$\vec{1}' \left(T_0 - \frac{T_x}{n} N \right) \vec{\alpha} = \sum_i (T_{xi} - n_i \bar{x}) \alpha.$$

$\mathcal{E}(F_m)$ est $>$ ou $<$ $\mathcal{E}(F_t)$ selon que $2m(\vec{\alpha}, \vec{\beta}, a)$ est $>$ ou $<$ $t(\vec{\alpha}, \vec{\beta})$. Mais le degré de libertés du numérateur de F_t doublant, les valeurs critiques de F_t sont, *grosso modo*, doubles de celles de F_m . F_t sera généralement une épreuve moins puissante que F_m , ce qui est logique puisqu'elle s'attaque à deux hypothèses en même temps : $\alpha_i = \alpha$ et $\beta_i = \beta$ pour tout i .

En ce qui concerne la comparaison de $\mathcal{E}(F_m)$ et $\mathcal{E}(F_c)$, on a

$$1 - \frac{2}{n-2q} > \frac{1}{1 - \frac{2}{n-q-1} + \frac{(n-q-2)p(\vec{\beta})}{(n-q-1)^2}}$$

Il importe donc de comparer $m(\vec{\alpha}, \vec{\beta}, a)$ à $t(\vec{\alpha}, \vec{\beta}) - p(\vec{\beta})$. Si $\vec{\beta} = \beta\vec{1}$, la non-centralité de S_c/σ^2 est supérieure à celle de S_m/σ^2 puisque alors $t(\vec{\alpha}, \beta\vec{1}) - p(\beta\vec{1}) = t(\vec{\alpha}, \beta\vec{1}) \geq m(\vec{\alpha}, \beta\vec{1}, a)$. Cette dernière inégalité compensera vraisemblablement l'inégalité ci-dessus et F_c gagnera sur F_m .

Rappelons ici que $p(\vec{\beta}) \leq t(\vec{\alpha}, \vec{\beta})$ (resp. $m(\vec{\alpha}, \vec{\beta}, a) \leq t(\vec{\alpha}, \vec{\beta})$) puisque le modèle II (resp. IV) est intermédiaire entre les modèles I (resp. I') et III (resp. III').

Si les droites de régression ne sont pas parallèles, il n'y a pas de relation systématique entre $m(\vec{\alpha}, \vec{\beta}, a)$ et $t(\vec{\alpha}, \vec{\beta}) - p(\vec{\beta})$. On peut cependant s'attendre à ce que ces deux quantités soient du même ordre de grandeur. En effet, si

les moyennes \bar{x}_i sont égales entre elles, on peut voir que les non-centralités de S_m/σ^2 avec $a = \bar{x}$, S_c/σ^2 et S_b/σ^2 avec b quelconque valent toutes

$$\frac{1}{n\sigma^2} \sum_{i < k} n_i n_k ((\alpha_i - \alpha_k) + \bar{x}(\beta_i - \beta_k))^2.$$

D'ailleurs $m(\vec{\alpha}, \vec{\beta}, a)$ peut dépasser $t(\vec{\alpha}, \vec{\beta}) - p(\vec{\beta})$ comme l'indiquent les deux cas particuliers suivants. Si

$$\vec{\alpha} + N^{-1}T_0 \left(I - \frac{\vec{1}\vec{1}'E}{E_{xx}} \right) \vec{\beta}$$

est colinéaire à $\vec{1}$, la non-centralité de S_c/σ^2 est nulle alors que celle de S_m/σ^2 est positive pour presque tout a . Si

$$\vec{\beta} + S^{-1}T \left(I - \frac{\vec{1}\vec{1}'M}{m} \right) (\vec{\alpha} + a\vec{\beta})$$

est colinéaire à $\vec{1}$, la non-centralité de S_c/σ^2 est inférieure à $m(\vec{\alpha}, \vec{\beta}, a) = t(\vec{\alpha}, \vec{\beta})$.

Enfin il faut tenir compte de ce que, si $\vec{\beta} \neq \vec{\beta}\vec{1}$, la différence entre

$$\frac{1}{1 - \frac{2}{n - 2q}} \quad \text{et} \quad \frac{1}{1 - \frac{2}{n - q - 1} + \frac{(n - q - 2)p(\vec{\beta})}{(n - q - 1)^2}}$$

s'approfondit en faveur de F_m . Il est donc probable qu'en cas de non-parallélisme accusé l'épreuve par F_m soit plus puissante que l'épreuve par F_c .

Pour montrer la supériorité de la méthode F_m , j'illustre la comparaison des performances selon la catégorie socio-culturelle par deux exemples non forcés. De trois populations où la performance a une variance unité, on extrait un échantillon de 30 enfants de moyenne d'âge 5 et d'écart type 2. On sait que le taux linéaire de croissance de la performance avec l'âge est 0,2 environ; en fait, les taux réels inconnus sont 0,1, 0,2 et 0,3. Enfin on décide de faire la comparaison à un an.

Les données sont donc $n_i = 30$, $T_{xi} = 150$, $E_{xxi} = 120$, $q = 3$, $\sigma^2 = 1$, $\vec{\beta} = (0,1, 0,2, 0,3)$, $b = 0,2$ et $a = 1$. $\mathcal{E}(F_p) = 2,2537$ à 2 et 84 libertés n'est pas significatif à 0,10 et permet d'envisager les épreuves F_c et F_b .

Si $\vec{\alpha} = (2,3, 1,6, 0,9)$, $\mathcal{E}(F_c) = 2,1905$ et $\mathcal{E}(F_b) = 2,1906$ ne sont pas significatifs à 0,10; $\mathcal{E}(F_t) = 2,2537$ n'est pas significatif à 0,05; $\mathcal{E}(F_m) = 3,2371$ est significatif à 0,05.

Si $\vec{\alpha} = (1,6, 1,6, 1,6)$, $\mathcal{E}(F_c) = 8,4633$, $\mathcal{E}(F_b) = 8,4637$ et $\mathcal{E}(F_t) = 5,4805$ sont tous trois significatifs à 0,001; $\mathcal{E}(F_m) = 1,0859$ n'est pas significatif à 0,25.

Ici la méthode F_m gagne sur tous les fronts.

En conclusion, on pourrait proposer la stratégie suivante.

1) Des raisons théoriques sérieuses permettent de postuler l'égalité des coefficients de régression.

a) Si dans un premier échantillon de taille importante, on constate une forte disparité entre $\hat{\beta}$ et β^* , et s'il est possible d'extraire indépendamment un second échantillon, on effectue l'épreuve F_b dans ce dernier avec b égal à l'estimation $\hat{\beta}$ du premier.

b) Sinon on éprouve l'hypothèse par F_c .

2) On ne peut supposer *a priori* le parallélisme.

a) On emploie F_m si c'est l'égalité des ordonnées en une certaine abscisse a qui est en question.

b) Si c'est l'identité des droites de régression qu'il faut éprouver, on emploie F_t .

BIBLIOGRAPHIE

- [1] W. G. COCHRAN, *Analysis of covariance: its nature and uses*. *Biometrics*, vol. 13, n° 3, 1957.
 [2] F. A. GRAYBILL, *An introduction to linear statistical models*. McGraw-Hill, 1961.
 [3] M. G. KENDALL, A. STUART, *The advanced theory of statistics*, vol. 2. Griffin, 1967.
 [4] B. J. WINER, *Statistical principles in experimental design*. McGraw-Hill, 1971.

(Manuscrit reçu le 8 mai 1978)