GEORGES KAMENAROV

Analyse de l'estimation bornée à l'aide de simulations

Statistique et analyse des données, tome 6, nº 1 (1981), p. 1-27

<http://www.numdam.org/item?id=SAD_1981__6_1_1_0>

Statistique et Analyse des données
1981 - 1 - pp. 1-27

ANALYSE DE L'ESTIMATION BORNEE A L'AIDE DE SIMULATIONS

Georges KAMENAROV

Institut de Technologie Chimique
Département de Mathématiques
1156 SOFIA - BULGARIE

RESUME : *On compare, par simulation, les propriétés de plusieurs estimateurs bornés du paramètre* β *d'un modèle linéaire* $Y = X\beta+\varepsilon$.

ABSTRACT : *Properties of several ridge estimators of the parameter* β *in a linear model* $Y = X\beta+\varepsilon$ *are compared by simulations methods.*

Mots-clés : *Modèle linéaire, estimation bornée (ridge regression), simulation.*

INTRODUCTION

L'analyse de régression est un ensemble de méthodes statistiques très générales et couramment appliquées par les analystes à une grande variété de problèmes dans le domaine des recherches scientifiques. C'est un modèle visant à décrire l'évolution de phénomènes matériels, sociaux, biologiques. C'est en servant d'outil pour les recherches, que les méthodes de régression sont à leur tour devenues un objet d'études.

Il est bien connu, que l'application de la méthode des moindres carrés pour l'estimation des coefficients dans le modèle linéaire donne parfois des résultats aberrants. Les causes de ces effets troublants sont analysées dans les articles [6] et [10]. Des techniques récentes dont le but est de protéger les qualités de la régression sont parues dans la littérature mondiale [4], [5], [11], [12]. La plus connue de ces méthodes et celle dont les propriétés ont été le plus souvent l'objet des publications est la régression ridge ou bornée.

Aussi bien développée qu'elle soit, la théorie de l'estimation bornée n'a pas formulé une règle explicite, exprimée en termes de probabilités, qui justifie dans un cas donné l'application d'un estimateur borné à la place de l'estimateur des moindres carrés. Ce fait rend incrédule en ce qui concerne l'application des estimateurs du type borné. N'ayant pas donné une réponse fondée sur des résultats théoriques, les statisticiens ont procédé à des simulations pour trouver au moins une solution fondée sur des résultats empiriques. Telle est la méthode employée par Hoerl, Kennard et Baldwin dans [3]. Nous ne discuterons pas des avantages et des inconvénients de cette méthode, d'ailleurs bien connue. Maints résultats théoriques ont été suggérés par des problèmes numériques. Aussi connaît-on pas mal de cas où les mathématiciens ont prouvé ou rejeté des propositions à l'aide d'un exemple numérique.

1. SYMBOLES ET REMARQUES DIVERSES

Désignons par X une matrice de type $n \times p$, de rang p. Ses éléments seront notés par x_{ij} et soumis aux conditions :

$$(1.1) \qquad \sum_{i=1}^{n} x_{ij} = 0, \quad \sum_{i=1}^{n} x_{ij}^2 = 1, \quad j = 1,2,\ldots,p.$$

Nous notons X' la transposée de X. Soit β un vecteur constant, inconnu, appartenant à l'espace R^p et soit σ^2 une constante inconnue. Notons par ε un vecteur aléatoire à valeurs dans R^n et de loi de Laplace-Gauss centrée dont la matrice des variances-covariances est $D\varepsilon = \sigma^2 I_n$, I_n désignant la matrice unité $n \times n$. Le modèle linéaire s'écrit :

$$Y = X\beta+\varepsilon$$

Le vecteur

$$\hat{\beta} = (X'X)^{-1}X'Y \;=\; \beta + \underbrace{(X'X)^{-1}X'}_{Q\Lambda^{-1}Q'Q'Q\Gamma}\varepsilon$$

est l'estimateur par la méthode des moindres carrés.

On sait qu'il existe trois matrices Q, Λ, G telles que

$$X = G\,\Lambda^{1/2}\,Q', \qquad X'X = Q\,\Lambda\,Q'$$

où Q est une matrice orthogonale $p \times p$, $\Lambda = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_p)$ avec $\lambda_1 > \lambda_2 > \ldots > \lambda_p > 0$ et G est du type $n \times p$ avec $G'G = I_p$. De (1.1), il résulte que

$$\mathrm{Tr}(X'X) = \sum_{j=1}^{p} \lambda_j = p, \qquad \mathrm{Tr}(X'X)^{-1} = \sum_{j=1}^{p} \lambda_j^{-1} \geqq p$$

Posons :

$$\gamma = Q'\beta\ , \quad \hat{\gamma} = Q'\hat{\beta}, \quad \eta = G'\varepsilon \qquad \Rightarrow \beta'\beta = \gamma'\gamma$$

Le vecteur η suit alors une loi centrée de Laplace-Gauss, de variance $D\eta = \sigma^2 I_p$. On obtient :

$$\hat{\beta} = \beta + Q\,\Lambda^{-1/2}\eta \ , \qquad \hat{\gamma} = \gamma + \Lambda^{-1/2}\eta$$

Les variances $D\hat{\gamma}_j = \sigma^2\lambda_j^{-1}$, $j = 1,2,\ldots,p$, des coordonnées de $\hat{\gamma}$ montrent, que les fluctuations de celles-ci peuvent devenir très grandes dans le cas $\lambda_j < 1$ tandis que dans le cas $\lambda_j > 1$ elles s'amoindrissent. Compte tenu des relations ci-dessus, l'instabilité de $\hat{\beta}$ est le résultat des fluctuations de certaines seulement des coordonnées de $\hat{\gamma}$. Cette instabilité s'exprime par des valeurs absolues élevées et les signes fortuits des coordonnées de $\hat{\beta}$.

Notons par b un estimateur quelconque de β . Un critère facile à saisir est donné par l'égalité

(1.2.) $\quad M(b) = E(b-\beta)'(b-\beta) = \mathrm{Tr}(Db) + (Eb-\beta)'(Eb-\beta)$

Certains auteurs estiment qu'il est parfois préférable de se soustraire aux règles traditionnelles qui refusent tout compromis avec un estimateur biaisé. Ces auteurs proposent dans le cas où les grandeurs $M(\hat{\beta}) \equiv \mathrm{Tr}(D\hat{\beta})$ ont de grandes valeurs, de recourir à des méthodes plus souples en adoptant un estimateur biaisé b pour lequel les valeurs de M(b) et de Tr(Db) sont plus faibles.

Posons $\bar{K} = \text{diag}\ (K_1, K_2, \ldots, K_p)$, $K_j \geq 0$, $j = 1,2,\ldots,p$.
Le vecteur

$$(1.3) \qquad b_g^* = Q(\Lambda+\bar{K})^{-1}\ \Lambda\ Q'\hat{\beta}$$

appelé estimateur ridge ou borné a été introduit par Hoerl et Kennard [1].
De (1.2) on déduit :

$$(1.4) \qquad M(b_g^*) = \sigma^2 \text{Tr}[(\Lambda+\bar{K})^{-2}\Lambda]+\gamma'(\Lambda+\bar{K})^{-2}\ \bar{K}^2\gamma$$

Nous considèrerons trois versions du vecteur (1.3). Le vecteur b_g^* peut être défini aussi par la relation

$$(1.5) \qquad b_g^* = (X'X+Q\bar{K}Q')^{-1}X'Y$$

équivalente à (1.3). Pour $M(b_g^*)$ on obtient encore

σ^2

$$(1.6) \qquad M(b_g^*) = \sigma^2 \sum_{j=1}^{p} \frac{\lambda_j}{(\lambda_j+K_j)^2} + \sum_{j=1}^{p} \left(\frac{K_j}{\lambda_j+K_j}\right)^2 \gamma_j^2$$

Hoerl et Kennard ont prouvé, que si le point $(K_1,K_2,\ldots,K_p)$ appartient à un sous-ensemble déterminé (mais inconnu) de l'espace R^p, la fonction $M(b_g^*)$ prend des valeurs inférieures à $\text{Tr}(D\hat{\beta}) = \sigma^2\ \text{Tr}(X'X)^{-1}$. Le point dont les coordonnées sont $K_j = \sigma^2/\ \gamma_j^2$, $j = 1,2,\ldots,p$, minimise $M(b_g^*)$.

L'estimateur borné

$$(1.7) \qquad b^* = (X'X+K\ I_p)^{-1}\ X'Y$$

s'obtient à partir du vecteur (1.3) en posant $\bar{K} = K\ I_p$, $K \geq 0$.
De (1.4) on déduit

$$(1.8) \qquad M(b^*) = \sigma^2 \sum_{j=1}^{p} \frac{\lambda_j}{(\lambda_j+K)^2} + K^2 \sum_{j=1}^{p} \frac{\gamma_j^2}{(\lambda_j+K\ )^2}$$

En général les minimums des fonctions $M(b^*)$ et $M(b_g^*)$ diffèrent et le second est inférieur au premier.

L'estimateur raccourci

(1.9) $\quad b_s^* = d\hat{\beta}, \quad 0 < d \leq 1$

provient de (1.3) en déterminant $\bar{K}$ par la résolution de l'équation matricielle $(\Lambda+\bar{K})^{-1}\,\Lambda = d\, I_p$, c'est-à-dire en posant $K_j = \lambda_j \frac{1-d}{d}$, $j = 1,2,\ldots,p$. De (1.4) nous obtenons

$d = \frac{\lambda_i}{\lambda_i + k_i} = \frac{1}{1+\frac{k_i}{\lambda_i}}$

(1.10) $\quad M(b_s^*) = d^2\sigma^2 \mathrm{Tr}(X'X)^{-1} + \beta'\beta \;\; (1-d)^2$

Le minimum de $M(b_s^*)$ s'obtient pour $d_o = \dfrac{\beta'\beta}{\beta'\beta+\sigma^2\mathrm{Tr}(X'X)^{-1}}$. Si l'hypothèse

$$H : \gamma_1^2\lambda_1 = \gamma_2^2\lambda_2 = \ldots = \gamma_p^2\lambda_p$$

$\frac{\sigma^2}{\gamma_i^2} \cdot \lambda_j \frac{1-d_0}{d_0} \Rightarrow \lambda_j \gamma_j^2$ cte

est vraie les minimums de $M(b_g^*)$ et de $M(b_s^*)$ coïncident.
Une procédure pour tester cette hypothèse est donnée par Obenchain dans [7].
On a démontré que si H n'est pas vraie, on a $M(b_g^*) < M(b_s^*)$.
Posons

(1.11) $\quad \Delta = \beta'\beta - \sigma^2\mathrm{Tr}(X'X)^{-1}$

Proposition 1

Si $\Delta > 0$, *il existe un nombre positif* K^* *tel que si* $0 < K < K^*$ *, on a* $(M(b^*) < \sigma^2\mathrm{Tr}(X'X)^{-1}$ *et si* $K > K^*$ *on a* $M(b^*) > \sigma^2\mathrm{Tr}(X'X)^{-1}$. *On a aussi* $M(b^*|K = K^*) = \sigma^2\mathrm{Tr}(X'X)^{-1}$.

Proposition 2

Si $\Delta \leq 0$, *on a* $M(b^*) < \sigma^2\, \mathrm{Tr}(X'X)^{-1}$ *pour tout* $K > 0$.

Ces propositions se démontrent facilement en tenant compte de la relation $\lim_{K\to\infty} M(b^*) = \gamma'\gamma = \beta'\beta$ et de la continuité de $M(b^*)$ pour $K \geq 0$. Les figures ci-après nous en donnent une illustration.

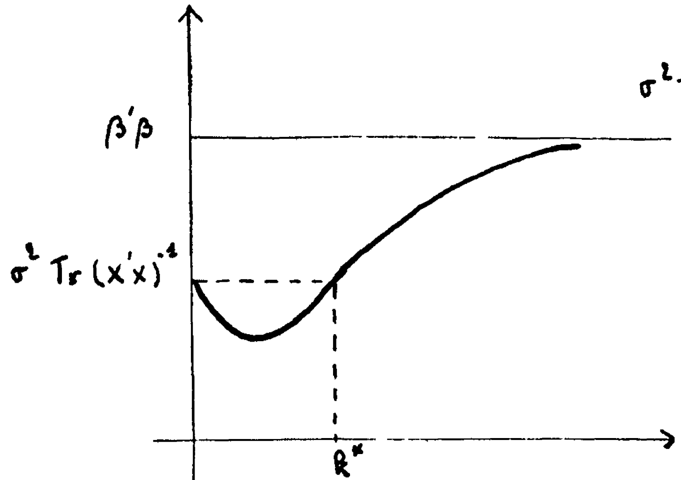

fig.1 $\Delta > 0$

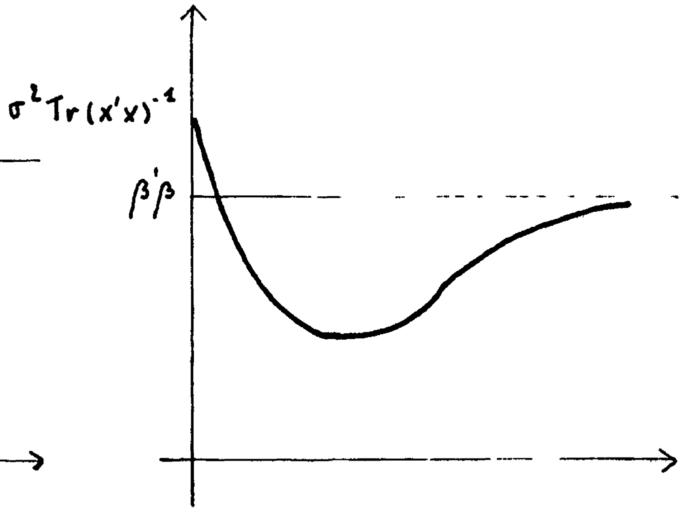

fig.2 $\Delta < 0$

Des propositions équivalentes à 1 et 2 ont été traitées par Marquardt [14].

Proposition 3

Soit $\Delta > 0$. Pour que K^* converge vers ∞ il faut et il suffit que Δ tende vers 0 .

Démonstration : La proposition 1 et la définition de Δ (1.11) donnent $M(b^*|K = K^*) = \beta'\beta - \Delta$. Suivant (1.8), cette dernière relation peut s'écrire sous la forme

$$(1.12) \qquad \sigma^2 \sum_{j=1}^{p} \frac{\lambda_j}{(\lambda_j+K^*)^2} + \sum_{j=1}^{p} \frac{\gamma_j^2}{\left(\frac{\lambda_j}{K^*}+1\right)^2} = \beta'\beta-\Delta$$

En posant $K^* \to \infty$, la relation (1.12) se ramène à l'égalité $\sum_{j=1}^{p} \gamma_j^2 = \beta'\beta-\Delta$, ce qui prouve la nécessité, pour que $\Sigma \gamma_j^2 = \delta'\delta - \beta'\beta$

Par définition K^* est positif. Il est unique. Pour démontrer la suffisance faisons tendre Δ vers 0 dans (1.12). Le côté gauche de l'égalité converge vers $\beta'\beta$. Comme toutes les valeurs propres $\lambda_1, \lambda_2, \ldots \lambda_p$ sont différentes de 0, les expressions $\sum_{j=1}^{p} \frac{\lambda_j}{(\lambda_j+K^*)^2}$ et $\frac{\lambda_j}{K^*}$ ne convergent simultanément vers 0 que dans le cas où $K^* \to \infty$.

On a montré dans [8], que l'inégalité $M(b_s^*) < \sigma^2 Tr(X'X)^{-1}$ est vraie si et seulement si $d^* < d < 1$ où $d^* = \dfrac{\Delta}{\beta'\beta+\sigma^2 Tr(X'X)^{-1}}$.

$\Sigma \gamma_j^2 = \Sigma \beta_j^2 = \beta'\beta$

Dans le cas $d^* > 0$ on obtient directement le résultat $M(b_s^* | d = d^*) = \sigma^2 Tr(X'X)^{-1}$. Mais le signe de d^* est le même que celui de Δ. On peut immédiatement évoquer des relations entre $M(b_s^*)$ et Δ analogues aux relations entre $M(b^*)$ et Δ, traitées par les propositions 1, 2 et 3.

Comme selon les formules (1.5), (1.7), (1.9) les estimateurs b_g^*, b^*, b_s^* se déduisent de la formule (1.3), il en découle que les propriétés de b_g^* sont à l'origine de certaines propriétés importantes de b^* et de b_s^*. Il convient de compléter les considérations citées plus haut dans le cas de b^* et de b_s^*, par une analyse qui traite l'effet du signe de Δ dans le cas de l'estimateur b_g^*, générateur de b^* et de b_s^*.

Désignons par le symbole $\tilde{K}$ le sous-ensemble de R^p donné par l'égalité

$$\tilde{K} = \{(K_1, K_2, \ldots, K_p) \ / \ K_j \geq 0, \ j = 1,2,\ldots,p \}$$

D'après (1.6) on voit que la fonction $M(b_g^*)$ de la variable $(K_1, K_2, \ldots, K_p)$ est définie et continue dans $\tilde{K}$.

Proposition 4

Si $\Delta>0$, *il existe une partition* $\{\tilde{K}_1, \tilde{K}_2, \tilde{K}_3\}$ *de l'ensemble* $\tilde{K}$, *qui a les propriété suivantes*

$$\begin{aligned}
&(K_1, K_2, \ldots, K_p) \in \tilde{K}_1 \iff M(b_g^*) < \sigma^2 Tr(X'X)^{-1} \\
(1.13) \qquad &(K_1, K_2, \ldots, K_p) \in \tilde{K}_2 \iff M(b_g^*) = \sigma^2 Tr(X'X)^{-1} \\
&(K_1, K_2, \ldots, K_p) \in \tilde{K}_3 \iff M(b_g^*) > \sigma^2 Tr(X'X)^{-1}
\end{aligned}$$

Démonstration : Soit $\Delta > 0$. On a

$$(1.14) \qquad \min_{K_1,K_2,\ldots,K_p} M(b_g^*) \leq M(b_g^*) < \beta'\beta$$

Si nous admettons que dans $\tilde{K}$ peut avoir lieu la relation $M(b_g^*) \geq \beta'\beta$ il suit en vertu de la continuité de $M(b_g^*)$ et de $\lim\limits_{K_1\to\infty,K_2\to\infty,\ldots,K_p\to\infty} M(b_g^*) = \beta'\beta$, que la fonction $M(b_g^*)$ a plus d'un point extrême, ce qui est impossible. Il existe alors trois sous-ensembles $\tilde{K}_1$, $\tilde{K}_2$, $\tilde{K}_3$ de $\tilde{K}$, disjoints, pour lesquels on a $\tilde{K}_1+\tilde{K}_2+\tilde{K}_3 = \tilde{K}$ et tels que les relations (1.13) soient vraies.

Proposition 5

Si $\Delta \leq 0$, *on a* $M(b_g^*) < \sigma^2 Tr(X'X)^{-1}$ *pour tout point de* $\tilde{K}$ *sauf pour l'origine* (0, 0, ..., 0).

Démonstration : Dans le cas où $\Delta \leq 0$ on obtient immédiatement

$$\min_{K_1, K_2, \ldots, K_p} M(b_g^*) < \beta'\beta \leq \sigma^2 Tr(X'X)^{-1}$$

La supposition que dans ce cas peut avoir lieu la relation $M(b_g^*) \geq \sigma^2 Tr(X'X')^{-1}$ pour des points de $\tilde{K}$ différents de l'origine, nous amène aux possibilités suivantes :

i) la fonction $M(b_g^*)$ a plus d'un point extrême dans $\tilde{K}$.

ii)La fonction $M(b_g^*)$ est constante dans un voisinage de l'origine où elle prend la valeur de $\sigma^2 Tr(X'X)^{-1}$.

L'absurdité de ces deux suppositions évant évidente, la proposition est prouvée.

Les considérations ci-dessus engendrent une classification essentielle des problèmes de régression traités par l'estimation bornée. Les problèmes qui se caractérisent par la relation $\Delta \leq 0$ peut être traités par l'estimation bornée sans aucun risque de faire augmenter la valeur de la fonction correspondante $M(.)$ au dessus de la valeur de $\sigma^2 Tr(X'X)^{-1}$. Pour les problèmes caractérisés par $\Delta > 0$,la probabilité de réaliser un choix des paramètres K, K_j, d, tel qu'on obtient $M(b_g^*) < \sigma^2 Tr(X'X)^{-1}$, tend vers 1 quand Δ tend vers 0.

Ceci montre l'importance que peuvent avoir des informations a priori dans le cas où on est enclin à utiliser un estimateur borné.

Dans la pratique la valeur de Δ et son signe sont inconnus. Les valeurs de K^* et de d^*, ainsi que les ensembles $\tilde{K}_1$, $\tilde{K}_2$, $\tilde{K}_3$, le sont aussi.

Il est impossible de choisir les valeurs des paramètres K, K_j, d, telles que l'on ait $M(.) < \sigma^2 Tr(X'X)^{-1}$ avec une probabilité égale à 1.

Dans certains cas, dans la pratique, on est en possession d'informations qui peuvent rendre plus efficaces les méthodes d'évaluation de K, K_j, d. Par exemple si on sait d'avance que les coordonnées de β sont soumises aux conditions $\beta_j \geq 0$, $j = 1,2,\ldots,p$, $\sum_{j=1}^{p} \beta_j = 1$, on déduit que $\beta'\beta \leq 1$. Alors si $\sigma^2\, Tr(X'X)^{-1} \geq 1$ est vrai, on a $\Delta \leq 0$. De même si $p\sigma^2 \geq \beta'\beta$ est vrai, on a aussi $\Delta \leq 0$, quelle que soit la valeur de $Tr(X'X)^{-1}$.

Nous nous rapporterons dans la suite aux relations suivantes :

i) si $\Delta \leq 0$, on a

$$(1.15) \qquad p\sigma^2/\beta'\beta \geq p/Tr(X'X)^{-1}, \qquad 0 < d_o \leq 1/2$$

ii) si $\Delta > 0$, on a

$$(1.16) \qquad p\sigma^2/\beta'\beta < p/Tr(X'X)^{-1}, \qquad 1/2 < d_o < 1$$

Posons

$$\hat{Y}_i = \sum_{j=1}^{p} x_{ij}\beta_j \qquad i = 1,2,\ldots,n$$

et

$$(1.17) \qquad \hat{\sigma}^2 = \frac{1}{n-p} \sum_{i=1}^{n} (Y_i - \hat{Y}_i)^2$$

Les statistiques

$$(1.18) \qquad T_1 = \hat{\beta}'\hat{\beta} - \hat{\sigma}^2\, Tr(X'X)^{-1} \; ; \quad T_2 = \hat{\beta}'\hat{\beta} - 2\,\hat{\sigma}^2 Tr(X'X)^{-1}$$

sont des estimateurs sans biais de $E(\hat{\beta}'\hat{\beta})$ et de Δ .

2. ALGORITHMES POUR EVALUER LES PARAMETRES K, K_j, d.

Le but de notre étude est de tirer des conclusions sur la conduite des estimateurs (1.5), (1.7), (1.9) en comparaison avec l'estimateur des moindres carrés $\hat{\beta}$ et le vecteur inconnu β et $\sigma^2 Tr(X'X)^{-1}$, tout en les comparant entre eux. Or les estimateurs du type borné dépendent à leur tour des paramètres K, K_j, d. Les méthodes connues, propres à évaluer ces paramètres dans la pratique, sont trés variées [7]. Ces méthodes ont en général des fondements différents, ce qui est la cause des effets suivants :

i) les valeurs des paramètres varient pour un problème donné avec la méthode ;

ii) il est difficile de comparer entre elles ces méthodes. Ces raisonnements nous ont amené à la conclusion de soumettre à l'épreuve plusieurs méthodes.

Premier algorithme

Il sert à évaluer le paramètre K de (1.7) et repose sur le test de l'hypothèse nulle H_o : $\Delta > 0$ contre l'hypothèse alternative H_1 : $\Delta \leq 0$. Pour limiter les valeurs acceptables de K dans le cas où l'hypothèse H_0 est acceptée, nous nous réfèrerons à un résultat publié dans [9] : si $0 < K < \frac{2\sigma^2}{\beta'\beta}$ on a $M(b^*) < \sigma^2 Tr(X'X)^{-1}$. La valeur de $\frac{2\sigma^2}{\beta'\beta}$ étant inconnue, nous l'estimerons par $\frac{2\sigma^2}{T_1}$, dans le cas $T_1 > 0$. Si l'hypothèse H_o est rejetée, c'est-à-dire si on accepte H_1, on doit choisir la valeur de K dans l'intervalle $(0,\infty)$. Or dans le cas $\Delta \leq 0$ la fonction (1.8) a une faible variation dans un intervalle assez grand, dans lequel elle prend des valeurs voisines de son minimum. Nous proposons dans le cas $\Delta \leq 0$ de restreindre le domaine des valeurs acceptables de K en choisissant pour la limite supérieure de K un nombre arbitraire K_o^*, tel qu'on ait $K_o^* > p/Tr(X'X)^{-1}$. Munissons les intervalles $J_1 = (0\ \frac{2\sigma^2}{T_1})$ et $J_2 = (p/Tr(X'X)^{-1}, K_o^*)$ de distributions uniformes et soit q positif. Nous formulons la règle suivante :

I/ Si $T_1 > 0$, $T_2 > -q$, on choisit $K \in J_1$;

II/ Si $T_2 \leq -q$, on choisit $K \in J_2$.

La valeur de q sert à déterminer la région critique du test. Elle n'est pas simplement positive, il faut aussi que q vérifie l'inégalité $q \le \sigma^2 \mathrm{Tr}(X'X)^{-1}$ afin de ne pas appliquer la partie I/ de la règle quand $T_1 \le 0$. Les distributions de T_1 et de T_2 restant inconnues les risques de première et deuxième espèce α et β restent indéterminés. Pourtant on peut se rendre compte de la perte consécutive à une décision fausse. Si on accepte H_0 alors que H_1 est vraie la perte est nulle, car dans le cas $\Delta \le 0$, on a $M(b^*) < \sigma^2 \mathrm{Tr}(X'X)^{-1}$ pour tout $K > 0$. Au contraire, si on accepte H_1 alors que H_0 est vraie, la perte a une grande probabilité d'être importante si on a $M(b^*) \ge \sigma^2 \mathrm{Tr}(X'X)^{-1}$. Il en résulte qu'il faut se garder surtout de rejeter H_0 si elle est vraie. Ceci implique, que la valeur du risque de première espèce α soit minimale, ce qui a lieu si q est maximal. Si on choisit $q = \sigma^2 \mathrm{Tr}(X'X)^{-1}$, $T_2 > -q$ entraîne $T_1 > 0$.

Deuxième algorithme

Il sert de nouveau à évaluer K et il a été publié dans [3]. Cette règle implique

$$K = p\, \hat{\sigma}^2 / \hat{\beta}'\hat{\beta}$$

L'efficacité de cette évaluation fait l'objet de considérations dans [3].

Troisième algorithme

Nous proposons cet algorithme pour l'évaluation des éléments diagonaux de la matrice $\bar{K}$, (1.5). Comme nous l'avons déjà observé, dans le cas $\lambda_j > 1$ il y a une grande probabilité pour que la valeur absolue $|\hat{\gamma}_j - \gamma_j|$ ne soit pas grande puisque $D\hat{\gamma}_j < \sigma^2$. On peut alors estimer $\frac{\sigma^2}{\gamma_j^2}$ par $\frac{\hat{\sigma}^2}{\hat{\gamma}_j^2}$. Formulons la règle suivante :

$$\left.\begin{array}{ll} \text{I/} \quad \text{Si } \lambda_j > 1 \text{ on pose } K_j = \dfrac{\hat{\sigma}^2}{\hat{\gamma}_j^2} \;; & \\ & j = 1,2,\ldots,p. \\ \text{II/} \quad \text{Si } \lambda_j \le 1 \text{ on pose } K = p\, \hat{\sigma}^2/\hat{\beta}'\hat{\beta} & \end{array}\right)$$

Quatrième algorithme

Il sert à choisir le paramètre d de (1.9) et repose sur le même test que le premier algorithme. Supposons, que $T_1 > 0$, il résulte que

$$0 < \frac{T_1}{\hat{\beta}'\hat{\beta}} < 1.$$

Comme

$$ET_1 = \beta'\beta, \quad E(\hat{\beta}'\hat{\beta}) = \beta'\beta + \sigma^2 Tr(X'X)^{-1}$$

et comme

$$d^* = \frac{\Delta}{E(\hat{\beta}'\hat{\beta})} < \frac{ET_1}{E(\hat{\beta}'\hat{\beta})}$$

nous proposons dans le cas où on accepte H_o, d'estimer la limite inférieure des valeurs acceptables de d par $\frac{T_1}{\hat{\beta}'\hat{\beta}}$.

D'autre part, on a

$$\frac{\hat{\beta}'\hat{\beta}}{\hat{\beta}'\hat{\beta}+\hat{\sigma}^2 Tr(X'X)^{-1}} > \frac{T_1}{\hat{\beta}'\hat{\beta}}$$

Posons

$$J_1 = \left(\frac{T_1}{\hat{\beta}'\hat{\beta}}, \frac{\hat{\beta}'\hat{\beta}}{\hat{\beta}'\hat{\beta}+\hat{\sigma}^2 Tr(X'X)^{-1}}\right).$$

Si $T_1 \leq 0$ et si $K_o^{**} > 1$ est un nombre arbitraire on a

$$1/2 > \frac{\hat{\beta}'\hat{\beta}}{\hat{\beta}'\hat{\beta}+\hat{\sigma}^2 Tr(X'X)^{-1}} > \frac{\hat{\beta}'\hat{\beta}}{\hat{\beta}'\hat{\beta}+K_o^{**}\sigma^2 Tr(X'X)^{-1}}$$

Posons

$$J_2 = \left(\frac{\hat{\beta}'\hat{\beta}}{\hat{\beta}'\hat{\beta}+K_o^{**}\hat{\sigma}^2 Tr(X'X)^{-1}}, \frac{\hat{\beta}'\hat{\beta}}{\hat{\beta}'\hat{\beta}+\hat{\sigma}^2 T(X'X)^{-1}}\right)$$

et munissons les intervalles J_1 et J_2 de distributions uniformes. Nous formulons la règle suivante :

I/ si $T_1 > 0$, $T_2 > -q$, on choisit $d \in J_1$;

II/ si $T_2 \leq -q$, on choisit $d \in J_2$

Remarque - Les intervalles J_1 et J_2 dans les algorithmes 1 et 4 ont été choisis en raison des relations (1.15) et (1.16). Dans les considérations qui suivent on a :

$$q = \hat{\sigma}^2 \mathrm{Tr}(X'X)^{-1},$$

$$K_o^* = \frac{2\hat{\sigma}^2 \mathrm{Tr}(X'X')^{-1}}{\hat{\beta}'\hat{\beta}} \quad \text{et} \quad K_o^{**} = 2.$$

Remarque - L'introduction des algorithmes 1 et 4 a un caractère plutôt expérimental. Ces algorithmes, quoique imparfaits reposent sur la différence des cas $\Delta > 0$ et $\Delta \leq 0$ qui représente l'objet fondamental de cette étude. Leurs formes nous ont été dictées plutôt par le caractère de la méthode : les simulations. Il est possible de mettre sur pied des algorithmes de ce genre et de les perfectionner.

Pour illustrer l'idée de ces algorithmes, considérons à titre d'exemple le cas suivant : $n = 12$ et $p = 1$. Dans ce cas le vecteur $\hat{\gamma}$ a une seule coordonnée que nous désignons aussi par $\hat{\gamma}$. Comme $\mathrm{Tr}(X'X)^{-1} = 1$, on a $E\hat{\gamma} = \gamma$ et $D\hat{\gamma} = \sigma^2$. La valeur de σ^2 peut être estimée par $\hat{\sigma}^2$ qui a une distribution χ^2 à $f = 10$ degrés de liberté. La variable aléatoire $\frac{\hat{\gamma}}{\hat{\sigma}}$ a alors une distribution de Student non-centrée à $f = 10$ degrés de liberté et de paramètre δ . Pour plus de détails sur cette distribution on se reportera à [16].

Selon les algorithmes 1 et 4 le choix des valeurs des paramètres K et d repose sur un test dont la région critique se détermine par la relation $\hat{\gamma}^2-\hat{\sigma}^2 \leq 0$, si nous choisissons pour la valeur de q celle de $\hat{\sigma}^2$. Le risque de premiere espèce α est donné par les relations

$$\alpha = P([-1 \leq \frac{\hat{\gamma}}{\hat{\sigma}} < 1 / \left|\frac{\gamma}{\sigma}\right| > 1]) =$$

$$= P([-1 \leq \frac{\hat{\gamma}}{\hat{\sigma}} < 1 / \frac{\gamma}{\sigma} > 1]) + P([-1 \leq \frac{\hat{\gamma}}{\hat{\sigma}} < 1 / \frac{\gamma}{\sigma} < -1])$$

Posant $\delta = \left|\frac{\gamma}{\sigma}\right|$, pour une valeur fixe de ce paramètre et ayant en vue la relation $P[T < t|\delta] = 1-P[T < -t|-\delta]$, [16] on obtient :

$$\alpha = 2\{P([\frac{\hat{\gamma}}{\hat{\sigma}} < 1/\delta]) - P([\frac{\hat{\gamma}}{\hat{\sigma}} < -1/\delta])\}$$

Pour illustrer le risque α, nous avons calculé ses valeurs en faisant varier δ . Ces calculs ont été effectués à l'aide de la formule donnée à cette fin, pour f pair, dans le chapître 5 de [16]:

$$P[T_f \leq t] = P(-\delta)+\sqrt{2\pi}\ [M_o + \frac{1}{1!2} M_2 + \frac{1}{2!2^2} M_4 +$$

$$\ldots + \frac{1}{(\frac{f-2}{2})!\ 2^{\frac{f-2}{2}}} M_{f-2}\ ]$$

Les résultats obtenus sont présentés dans la table suivante :

δ	1,0	1,1	1,2	1,3	1,4	1,5	1,6	1,7
α	0.280	0.272	0.367	0,259	0,253	0,248	0,246	0,241

3. SIMULATION

Pour obtenir de l'information sur les avantages de l'application d'un estimateur borné à la place de l'estimateur des moindres carrés, nous nous proposons d'effectuer un contrôle à l'aide d'un nombre limité d'exemples numériques dont nous connaissons la structure probabiliste. Cette structure consiste en un triplet (X,β,σ^2) et une loi de distribution de Laplace-Gauss. Ces conditions suffisent pour définir le modèle linéaire dans les termes que nous avons détaillés dans la partie 1.
Les indéterminations liées à l'emploi d'un estimateur borné reposent sur les faits suivants :

I) Le vecteur β et la constante σ^2 sont inconnus et restent inconnus après l'expérience ;

II) Les grandeurs σ^2/γ_j^2, $j = 1,2,\ldots,p$, ne peuvent pas se calculer dans la pratique ;

III) La fonction M(.) est inconnue dans la pratique.

Ces considérations peuvent se ramener à la première dont les deux autres sont des conséquences immédiates .

Nous nous proposons d'effectuer un contrôle aux niveaux suivants :

i) distance moyenne d'un estimateur borné à β ;

ii) fréquence avec laquelle un estimateur borné est à une distance du vecteur β inférieure à celle de l'estimateur des moindres carrés ;

iii) forme de l'estimateur borné dans le sens des relations (1.5), (1.7), (1.9) et algorithme pour l'évaluation des paramètres K, $\dot{K}_j$, d.

Notre étude repose sur une seule matrixe X. Celle-ci est du type 13 × 4 et a été publiée dans le livre de A. Hald [13] en 1952. Avant d'être employée ici, elle a subit les transformations nécessaires qui la soumettent aux exigences (1.1). Pour mettre au courant les lecteurs pour lesquels l'ouvrage de Hald, quoique cité à plusieurs reprises, est peu accessible, nous citons sans transformation cette matrice que voici :

$$X = \begin{vmatrix} 7 & 26 & 6 & 60 \\ 1 & 29 & 15 & 52 \\ 11 & 56 & 8 & 20 \\ 11 & 31 & 8 & 47 \\ 7 & 52 & 6 & 33 \\ 11 & 55 & 9 & 22 \\ 3 & 71 & 17 & 6 \\ 1 & 31 & 22 & 44 \\ 2 & 54 & 18 & 22 \\ 21 & 47 & 4 & 26 \\ 1 & 40 & 23 & 34 \\ 11 & 66 & 9 & 12 \\ 10 & 68 & 8 & 12 \end{vmatrix}$$

La matrice transformée avec laquelle nous opérons dans la suite pour des raisons de simplicité est de nouveau désignée par X et ses éléments par x_{ij}.

Les valeurs propres de la matrice X'X (X transformée) sont :

$\lambda_1 = 2{,}235698$; $\lambda_2 = 1{,}576062$; $\lambda_3 = 0{,}1866056$; $\lambda_4 = 0{,}0016239$

On a obtenu $\det(X'X) = 0{,}0010678$ et $Tr(X'X)^{-1} = 622{,}3208$. Nous effectuons l'expérience en variant le vecteur β et la constante σ^2 de manière à obtenir 7 valeurs différentes pour la grandeur $\beta'\beta$ et 7 valeurs pour σ^2. Ainsi on a obtenu un plan d'expérience du type 7^2. L'expérience a pour but d'étudier les effets des "facteurs" $\beta'\beta$ et σ^2, ainsi que l'effet du "facteur" Δ, ce dernier étant un point de vue nouveau dans les études de ce genre.

Le plan est présenté par la table 1^A. La première ligne de celle-ci contient les valeurs de $\beta'\beta$, la première colonne les valeurs de σ^2, tous les autres éléments de la table représentent les valeurs de Δ correspondant aux différents couples $(\beta'\beta,\sigma^2)$.

σ^2 \ $\beta'\beta$	0,3	207,45	415,20	622,90	830,10	1037,55	1245,00
0,0004820	0	207,45	414,90	621,90	829,80	1037,25	1244,70
0,3338793	- 207,45	0	207,45	414,90	621,90	829,80	1037,25
0,6672765	- 414,90	- 207,45	0	207,45	414,90	621,90	829,80
1,0006737	- 621,90	- 414,90	-207,45	0	207,45	414,90	621,90
1,3340708	- 829,80	- 621,90	-414,90	-207,45	0	207,45	414,90
1,6674680	-1037,25	- 829,80	-621,90	-414,90	-207,45	0	207,45
2,0008650	-1244,70	-1037,25	-829,80	-621,90	-414,90	-207,45	0

Table 1^A

Pour chaque couple $(\beta'\beta,\sigma^2)$, c'est-à-dire pour chaque triplet (X,β,σ^2) on a effectué 100 générations du vecteur Y, tel que $EY = X\beta$ et $DY = \sigma^2 I_n$. Nous avons traité les 100 problèmes qui correspondent à chaque triplet (X,β,σ^2) par la méthode des moindres carrés, par les estimateurs du type borné correspondant aux algorithmes cités, par l'estimateur borné (1.5) avec $K_j = \sigma^2/\gamma_j^2$, $j = 1,2,\ldots,p$, et par l'estimateur borné (1.7) avec $K = p\,\sigma^2/\,\beta'\beta$.

Les deux dernières estimations ne peuvent pas se réaliser dans la pratique. Nous en avons besoin pour nous faire une idée de l'estimateur borné "modèle"

Introduisons les symbôles suivants :

a) Nous notons b_{01} l'estimateur (1.5) avec $K_j = \sigma^2/\gamma_j^2$, $j = 1,2,\ldots,p$ et b_{02} l'estimateur (1.7) avec $K = p\,\sigma^2/\beta'\beta$;

b) Nous notons K_u et K_v les valeurs de K obtenues à l'aide du premier et du deuxième algorithmes, de même nous notons b_u et b_v les estimateurs du type (1.7) correspondant ;

c) Nous notons K_w^j , j = 1,2,...,p, les valeurs des éléments diagonaux de $\bar{K}$ obtenues au moyen du troisième algorithme et b_w l'estimateur (1.5) correspondant ;

d) Nous notons d la valeur du paramètre de l'estimateur (1.9) et celui-ci (estimateur raccourci) par b_s.

Pour un estimateur quelconque b nous introduisons les grandeurs suivantes :

e) $\bar{\delta}(b)$ est la moyenne arithmétique des 100 valeurs de la distance. $\delta(b) = \sqrt{(\beta-b)'(\beta-b)}$, valeurs qui correspondent à un triplet (X,β,σ^2) ;

f) $a(b) = |\beta'\beta-b'b|$

g) Fr(r) est la fréquence absolue avec laquelle la relation "r" a eu lieu lors des 100 expériences qui se rapportent à un triplet (X,β,σ^2).

Les résultats des simulations sont donnés dans les tables qui suivent. Ces tables sont numérotées de 1 à 24. Chaque table est suivie par une explication de la nature des données qu'elle contient. Chaque table a 7 lignes et 7 colonnes. Celles-ci se rapportent aux valeurs de $\beta'\beta$ et de σ^2 auxquelles correspondent les lignes et les colonnes de la table 1^A. Pour cette raison les lignes et les colonnes des différentes tables dans la suite seront marquées : 1, 2, 3, 4, 5, 6, 7.

Nous évoquerons aussi l'article [15] dont l'objet est similaire à celui que nous traitons. L'intérêt suscité par ce travail a pour témoignage irréfutable les commentaires publiés à son sujet. Il est inutile ici d'y revenir. Nous ferons remarquer seulement qu'il prouve une fois de plus l'efficience de la méthode des simulations dans le cas de l'analyse des procédures de regression alternatives à celle des moindres carrés. Les méthodes que les auteurs de [15] mettent en pratique, diffèrent des notres tant qu'au point de vue du plan de l'expérience que par le but qu'ils se proposent. Aussi les estimateurs qu'ils ont soumis à l'épreuve sont en général de nature différente de ceux qui sont présentés ici à l'exclusion de l'estimateur (1.7).

Néanmoins, les conclusions des auteurs de [15] sur la dépendance des estimateurs étudiés et en particulier de l'estimateur (1.7), du caractère des coordonnées de β et de la valeur de σ^2 sont en accord avec les considérations que nous citons à ce sujet dans la partie suivante de notre travail.

4. CONCLUSIONS

Le souci de préserver les qualités du modèle linéaire et d'augmenter s'il est possible sa stabilité et sa précision ont conduit les praticiens à tenir compte dans la pratique, de toute l'information disponible. Cette information peut suggérer le choix d'un critère éventuellement différent de celui des moindres carrés. Ici nous nous bornons à étudier le cas où toute l'information se réduit à la matrice X et au vecteur Y. Nous ne poserons pas des contraintes au niveau des coefficients et qui sont données par un système d'équations et d'inéquations linéaires. Les données étant supposées hors de soupçon et le modèle étant posé linéaire, la décision d'appliquer un estimateur borné est fondée sur l'étude du spectre de X'X. Or une petite valeur de det(X'X) est déjà un indice suffisant, que l'aberration des résultats obtenus par la méthode des moindres carrés est très probable. Nous espérons que nos résultats n'entreront pas en conflit avec les besoins des praticiens qui disposent d'informations supplémentaires.

La table 1 montre que la distance $\delta(\hat{\beta})$ dépend essentiellement de σ^2 et non de $\beta'\beta$, ce qui est dû à l'absence de biais. En comparant les tables 2 - 7 la table 1, on peut juger, en fonction de la distance euclidienne, des effets des différents estimateurs du type borné. Les valeurs des différents $\bar{\delta}(.)$, montrent, que les estimateurs du type borné sont préférables à l'estimateur des moindres carrés pour les valeurs de $\beta'\beta$ qui sont petites, indépendamment de la valeur de σ^2. En général, les deux premières colonnes des tables 2-7 contiennent des valeurs petites par rapport aux valeurs des colonnes correspondantes de la table 1. Les colonnes qui se rapportent à des valeurs élevées de $\beta'\beta$ montrent que l'estimateur des moindres carrés est équivalent à tout estimateur borné. On peut aussi marquer la tendance exprimée par la première ligne de toutes les tables 1-7, qui est en faveur de l'équivalence de tous les estimateurs dans le cas où σ^2 est petit.

Les tables 8-13 marquent l'effet du facteur Δ sur l'opportunité de l'emploi des estimateurs du type borné, du point de vue fréquences. En général, les fréquences appartenant à la diagonale d'une table quelconque, sont presque égales. Les fréquences qui sont au-dessous de la diagonale sont supérieure à celles qui sont au-dessus. Les premières se rapportent aux valeurs négatives de Δ et les secondes à ses valeurs positives (table 1^A). Ce fait prouve que dans le cas $\Delta \leq 0$ les estimateurs du type borné sont fréquemment plus proches de β (dans le sens euclidien), que ne l'est l'estimateur des moindres carrés. Les fréquences au-dessus de la diagonale montrent que pour $\Delta > 0$, l'intérêt de l'utilisation d'un estimateur borné à la place de l'estimateur des moindres carrés est moins fréquent.

La table 14 montre que la valeur et le signe de Δ sont étroitement liés aux fréquences de l'inégalité $\hat{\beta}'\hat{\beta} > \beta'\beta$. Au-dessous de la diagonale les fréquences sont essentiellement plus grands que les fréquences au-dessus. La "dynamique" au-dessous de la diagonale est aussi plus grande que celle qui est au-dessus. Les tables 15-20 ont une forte ressemblance avec la table 14. Elles prouvent, du point de vue des grandeurs a(.), que les estimateurs du type borné ont des effets avantageux plus fréquents dans le cas $\Delta \leq 0$.

L'estimateur borné b_{01} est l'estimateur "modèle" du point de vue critère (1.2). C'est à lui (et à tous les autres estimateurs du type borné), qu'on peut appliquer la phrase de M. Gary Mc Donald citée dans la discussion qui accompagne l'article [7] et que nous citerons aussi : "la démonstration du fait qu'un estimateur borné, ayant une forme quelconque, est à une distance de β (critère MSE) moindre que celle de l'estimateur des moindres carrés, N'EST PAS un argument convainquant pour qu'on renonce à la technique de la méthode des moindres carrés". Et en effet, la table 2 nous montre, que l'estimateur b_{01} est meilleur que l'estimateur des moindres carrés dans le cas où $\beta'\beta$ est petit et que pour des valeurs de $\beta'\beta$ élevées il lui est équivalent. L'estimateur b_{02} a presque les mêmes caractéristiques que b_{01}.

Les tables concernant l'estimateur b_w montrent que pour des valeurs petites de $\beta'\beta$ il est meilleur que tous les autres estimateurs du type borné qui peuvent être utilisés dans la pratique. Pour des valeurs élevées de $\beta'\beta$ l'estimateur b_v est le meilleur de tous les estimateurs du type borné présentés ici, utilisables dans la pratique. Les tables concernant les estimateurs b_u et b_s ont trop de ressemblances, sans doute dues aux algorithmes pour le choix des paramètres K et d. Ces algorithmes se caractérisent par des fréquences élevées dans les tables 20 et 24.

L'analyse des résultats de notre expérience nous permet d'émettre des conclusions qui, tout en ayant le caractère d'hypothèses, fournissent la possibilité de juger de l'utilisation des estimateurs du type borné. Ces résultats suggèrent que le facteur Δ joue un rôle décisif. La condition, que le spectre de X'X contient des valeurs proches de 0 n'est pas une condition suffisante pour l'utilisation d'un estimateur borné. C'est surtout le cas $\Delta \leq 0$ qu'il faut considérer comme propice pour le traitement du problème à l'aide de l'estimation bornée.

Nous estimons qu'une analyse préalable des données est d'une grande utilité dans le cas où l'estimateur des moindres carrés a une grande probabilité d'être aberrant. Primo c'est un avertissement pour le praticien, secundo c'est une analyse minutieuse qui doit être à la base de toute décision du choix de critère.

En définitive, nous ferons remarquer qu'un estimateur borné est un mécanisme plus souple que l'estimateur des moindres carrés, ce dernier étant un cas particulier de l'estimateur (1.3) pour $\bar{K} = 0$. Notre conviction est que ce mécanisme doit être manié avec plus de circonspection et de mesure pour donner les résultats qu'on attend de lui.

	1	2	3	4	5	6	7
1	0,4272	0,4271	0,4272	0,4272	0,4276	0,4277	0,4269
2	11,24440	11,2439	11,2442	11,2441	11,2443	11,2441	11,2440
3	15,8958	15,8956	15,8959	15,8958	11,8961	15,8963	15,8958
4	19,4660	19,4658	19,4661	19,4659	19,4662	19,4663	19,4659
5	22,4759	22,4760	22,4762	22,4760	22,4762	22,4763	22,4758
6	25,1280	25,1279	25,1281	25,1282	25,1283	25,1283	25,1280
7	27,5271	27,5271	27,5273	27,5272	27,5274	27,5277	27,5277

T.1

$\bar{\delta}(\hat{\beta})$

	1	2	3	4	5	6	7
1	0,3332	0,7018	0,5020	2,3579	0,4330	0,4593	0,4272
2	0,5178	7,3015	13,2721	22,0389	26,2444	27,6422	19,7659
3	0,5277	7,8243	14,1571	22,2505	29,1474	29,8471	24,8887
4	0,5315	8,2312	14,8033	22,3852	30,2681	30,7182	27,2101
5	0,5337	8,5649	15,3467	22,4741	30,8637	31,2140	28,5299
6	0,5353	8,8436	15,8186	22,5315	31,2333	31,5511	29,3755
7	0,5364	9,0809	16,2339	22,5825	31,4855	31,8052	29,9591

T.2

$\bar{\delta}(b_{01})$

	1	2	3	4	5	6	7
1	0,3971	0,4209	0,4231	0,4239	0,4241	0,4245	0,4248
2	0,5517	9,8350	9,8756	11,0379	12,8942	12,1714	12,0882
3	0,5493	11,0770	11,9873	14,1099	17,2531	16,6977	16,9377
4	0,5481	11,6064	12,9632	15,6412	19,5426	19,2573	19,8323
5	0,5474	11,9183	13,5419	16,5650	20,9518	20,8912	21,7342
6	0,5469	12,1330	13,9360	17,1904	21,9112	22,0274	23,0814
7	0,5466	12,2950	14,2287	17,6478	22,6110	22,8662	24,0872

T.3

$\bar{\delta}\ (b_{02})$

	1	2	3	4	5	6	7
1	0,4357	0,1897	0,4271	0,4271	0,4268	0,4270	0,4264
2	5,5270	11,1658	11,5531	12,1154	13,0357	12,4136	12,2633
3	7,7366	13,8959	14,9709	16,3137	18,0873	17,7096	17,8347
4	9,4643	15,6661	17,1079	18,7377	20,9029	21,0462	21,3554
5	10,9172	17,0759	18,5366	20,7147	22,9738	23,1769	23,8510
6	12,1961	18,2623	19,8593	22,1233	24,6780	24,9606	25,6366
7	13,3532	19,3392	21,1325	23,2489	26,1614	26,5972	27,3462

7.4

$\bar{\delta}(b_n)$

	1	2	3	4	5	6	7
1	0,3812	0,4220	0,4225	0,4631	0,4239	0,4242	0,4236
2	4,1613	9,6426	10,3598	11,6678	13,6353	13,3324	13,5082
3	5,8464	11,3601	12,4773	14,2910	16,9373	16,9763	17,5293
4	7,1448	12,5624	13,8277	15,8495	18,7766	19,0208	19,7931
5	8,2414	13,5534	14,8898	17,0143	20,0848	20,4570	21,3660
6	9,2086	14,4202	15,7944	17,9782	21,1284	21,5861	22,5861
7	10,8400	15,2030	16,6061	18,8196	22,0176	22,5340	23,5980

T.5

$\bar{\delta}(b_v)$

	1	2	3	4	5	6	7
1	0,3210	0,7447	0,5274	2,3644	0,4315	0,4661	0,4198
2	0,7538	7,5532	13,5531	21,9854	24,3552	26,4648	18,1438
3	0,9959	8,1244	14,7083	21,9829	27,1634	28,7203	22,5442
4	1,1894	8,5569	15,5464	21,9886	28,3328	29,6411	24,6229
5	1,3555	8,9032	16,2188	21,9973	29,0045	30,1815	25,8481
6	1,5035	9,1880	16,7701	22,0061	29,4468	30,5586	26,6505
7	1,6384	9,4282	17,2297	22,0145	29,7581	30,8466	27,2163

T.6

$\bar{\delta}(b_w)$

	1	2	3	4	5	6	7
1	0,4199	0,4249	0,4258	0,4259	0,4251	0,4265	0,4257
2	5,5286	11,0304	11,9639	12,2119	12,4177	12,2446	12,0938
3	7,7701	13,7695	15,8196	16,4912	16,8728	17,1258	17,3072
4	9,4805	15,6818	18,1534	19,4659	19,4665	20,4167	20,6771
5	10,9240	17,0788	19,7182	21,3112	22,0805	22,8277	23,2217
6	12,1967	18,3648	21,2412	22,6063	23,8062	24,6917	25,1280
7	13,3485	19,5682	22,5863	24,0151	25,1695	26,4379	27,1383

T.7

$\bar{\delta}\ (b_s)$

	1	2	3	4	5	6	7
1	52	27	38	1	48	44	53
2	100	58	38	10	6	4	19
3	100	65	47	30	13	10	23
4	100	69	54	39	22	22	27
5	100	76	57	44	30	30	33
6	100	79	58	47	36	35	33
7	100	79	59	51	38	39	41

T.8 $Fr[\delta(\hat{\beta}) > \delta(b_{01})]$

	1	2	3	4	5	6	7
1	45	53	55	55	55	55	55
2	100	47	43	38	34	33	32
3	100	53	49	44	36	38	35
4	100	59	56	48	41	42	40
5	100	63	59	50	45	44	42
6	100	67	61	54	47	47	45
7	100	68	63	58	49	48	47

T.9 $Tr[\delta(\hat{\beta}) > \delta(b_{02})]$

	1	2	3	4	5	6	7
1	55	55	53	53	47	44	47
2	100	57	57	53	49	51	50
3	100	62	60	57	52	53	50
4	100	65	63	59	55	56	54
5	100	67	65	60	55	56	56
6	100	69	65	62	57	57	57
7	100	70	68	64	58	58	57

T.10 $Fr[\delta(\hat{\beta}) > \delta(b_u)]$

	1	2	3	4	5	6	7
1	52	52	52	54	53	55	55
2	100	54	50	45	39	39	38
3	100	61	57	52	42	43	40
4	100	64	62	56	48	48	46
5	100	67	64	58	52	51	49
6	100	68	65	62	54	54	52
7	100	69	67	63	57	56	54

T.11 $Fr[\delta(\hat{\beta}) > \delta(b_v)]$

	1	2	3	4	5	6	7
1	52	24	37	3	48	42	52
2	100	57	37	12	9	8	24
3	100	65	44	31	18	12	27
4	100	70	53	40	28	23	30
5	100	74	56	47	34	32	36
6	100	77	59	48	39	37	40
7	100	78	59	51	41	40	41

T.12 $Fr[\delta(\hat{\beta}) > \delta(b_w)]$

	1	2	3	4	5	6	7
1	55	54	54	54	56	56	56
2	100	54	61	48	48	48	49
3	100	60	53	52	49	47	48
4	100	65	57	55	53	52	49
5	100	70	61	56	55	53	52
6	100	71	63	58	56	54	53
7	100	72	64	61	59	55	54

T.13 $Fr[\delta(\hat{\beta}) > \delta(b_s)]$

	1	2	3	4	5	6	7
1	60	56	55	56	56	56	56
2	100	61	59	56	56	56	56
3	100	67	62	60	56	56	56
4	100	71	67	63	58	57	57
5	100	77	70	65	59	59	58
6	100	77	76	67	61	61	69
7	100	78	76	70	62	61	61

T.14 $Fr\ [\hat{\beta}'\hat{\beta} > \beta'\beta\ ]$

	1	2	3	4	5	6	7
1	50	30	37	19	49	44	53
2	100	53	35	34	29	28	28
3	100	53	39	32	38	34	31
4	100	58	45	36	42	36	35
5	100	65	51	40	40	37	39
6	100	68	54	42	42	40	38
7	100	72	57	45	46	42	40

T.15 $Fr[a(\hat{\beta}) > a(b_{01})]$

	1	2	3	4	5	6	7
1	40	53	54	54	54	55	56
2	100	41	39	35	32	34	35
3	100	52	46	39	35	35	35
4	100	57	49	45	36	36	36
5	100	62	57	48	40	37	37
6	100	70	63	51	43	41	38
7	100	71	65	53	45	43	42

T.16 $Fr\ [a(\hat{\beta}) > a(b_{02})]$

	1	2	3	4	5	6	7
1	45	54	51	52	50	46	48
2	100	46	48	46	46	48	49
3	100	53	50	47	46	46	46
4	100	58	53	48	45	46	46
5	100	65	59	51	44	46	46
6	100	70	63	54	47	46	46
7	100	71	64	57	48	48	46

T.17 $Fr[a(\hat{\beta}) > a(b_u)]$

	1	2	3	4	5	6	7
1	42	52	53	54	52	54	54
2	100	44	41	39	38	39	39
3	100	53	48	43	38	38	38
4	100	59	52	47	40	39	39
5	100	65	60	49	42	42	41
6	100	70	63	52	44	44	42
7	100	71	66	56	48	46	44

T.18 Fr $[a(\hat{\beta}) > a(b_v)]$

	1	2	3	4	5	6	7
1	52	25	36	4	50	42	52
2	100	57	34	25	25	24	25
3	100	54	40	34	31	32	30
4	100	60	44	37	36	34	34
5	100	63	50	42	37	36	36
6	100	68	55	43	41	40	38
7	100	72	56	47	44	41	45

T.19 $[a(\hat{\beta}) > a(b_w)]$

	1	2	3	4	5	6	7
1	44	54	54	54	56	56	55
2	100	44	43	42	45	45	47
3	100	53	44	43	42	42	43
4	100	57	47	45	42	42	41
5	100	58	51	46	44	43	42
6	100	63	53	49	46	44	43
7	100	69	57	51	48	45	44

T.20 Fr $[a(\hat{\beta}) > a(b_s)]$

	1	2	3	4	5	6	7
1	100	52	100	100	98	98	98
2		100	99	91	75	89	91
3			100	100	66	78	83
4				100	94	85	75
5					100	99	93
6						100	100
7							100

T.21 Fr$[M(b_u) < \sigma^2 Tr(X'X)^{-1}]$

	1	2	3	4	5	6	7
1	100	27	45	38	17	25	25
2	100	100	100	85	45	48	44
3			100	100	68	75	65
4				100	93	92	84
5					100	100	98
6						100	100
7							100

T.22 Fr$[M(b_v) < \sigma^2 Tr(X'X)^{-1}]$

	1	2	3	4	5	6	7
1	100	27	45	32	19	25	25
2		100	100	85	45	48	44
3			100	100	68	75	65
4				100	93	92	84
5					100	100	98
6						100	100
7							100

T.23 Fr$[M(b_w) < \sigma^2 Tr(X'X)^{-1}]$

	1	2	3	4	5	6	7
1	100	95	95	95	95	96	95
2		100	91	88	81	87	87
3			100	100	83	86	84
4				100	94	91	86
5					100	99	95
6						100	99
7						100	100

T.24 $\quad Fr[M(b_s) < \sigma^2 Tr(X'X)^{-1}]$

BIBLIOGRAPHIE

1. Hoerl A., R. Kennard - Ridge regression : biased estimation for nonorthogonal problems, Technometrics, 12, 1970, 55-67
2. Hoerl A., R. Kennard - Ridge regression : application to nonorthogonal problems, Technometrics, 12, 1970, 69-82
3. Hoerl A., R. Kennard, K. Baldwin - Ridge regression : some simulations Communication in Statistics, 4, (2), 1975, 105-123.
4. Gunst R., R. Mason - Biased estimation in regression : an evaluation using mean squared error, JASA, 72, n° 359, 1977, 616-627
5. Hocking R., F. Speed, M. Lynn - A class of biased estimators in linear regression, Technometrics, 18, n° 4, 1976, 425-437
6. Mason R., R. Gunst, J. Westeber - Regression analysis and problems of multicollinearity, Communication in Statistics, 4, (3), 1975, 277-293
7. Obenchain R. - Ridge analysis following a preliminary test of shrunken hypothesis, Technometrics, 17, 1975, 431-445.
8. Mayer L, T. Wilke - On biased estimation in linear models, Technometrics, 15, 1973, 497-508
9. Theobald C.M. - Generalization of mean square error applied to ridge regression, Journal of the Royal Stat. Soc., (serie B), 36, 1974, 103-106
10. Silvey S.D. - Multicollinearity and imprecise estimation, J.R. Stat. Soc., (serie B), 31, 1969, 539-552.
11. Brenot J., P. Cases, N. Lacourly - Pratique de la regression : qualite et protection, Cahiers du B.U.R.O., 1975, n° 33 Paris.
12. Cases P. - Protection de la regression par utilisation de contraintes linéaires et nonlinéaires, Revue de Statistique Appliquée, 1975, vol. XXIII, n° 3, 37-57.
13. Hald A. - Statistical theory with engineering applications, New-York-London, 1952.
14. Marquardt D. - Generalized inverse, ridge regression, biases linear estimation, Technometrics, 12, n° 3, 1970, 591-612.
15. Dempster A.P., M. Schatzoff, N. Wermouth - A simulation study of alternatives to ordinary least squares, JASA, Vol.72, n° 357, 1977, 77-107.
16. Owen D.B. - Handbook of Statistical Tables, Addison-Wesley Publishing Company, INC, Palo Alto, London.