

# CAHIERS DU BURO

R. A. S. WHITFIELD

## **Une stratégie optimale pour le contrôle douanier**

*Cahiers du Bureau universitaire de recherche opérationnelle.*  
*Série Recherche*, tome 22 (1975), p. 3-69

[http://www.numdam.org/item?id=BURO\\_1975\\_\\_22\\_\\_3\\_0](http://www.numdam.org/item?id=BURO_1975__22__3_0)

© Institut Henri Poincaré — Institut de statistique de l'université de Paris, 1975, tous droits réservés.

L'accès aux archives de la revue « Cahiers du Bureau universitaire de recherche opérationnelle. Série Recherche » 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/>

# CHAPITRE 1

## INTRODUCTION

### 1.1. – ORIGINE DU PROBLEME

Le sujet de cette thèse est tiré du domaine de la douane d'un pays européen. On s'intéresse au contrôle des importations. Le problème est donc réel, et nous allons l'aborder dans l'esprit de la recherche opérationnelle, nous intéressant aux aspects théoriques et pratiques.

### 1.2. – PORT SCHEMATIQUE

Nous considérons les étapes dans le passage des biens par un port schématique. C'est le contrôle douanier dans le port qui nous concerne.

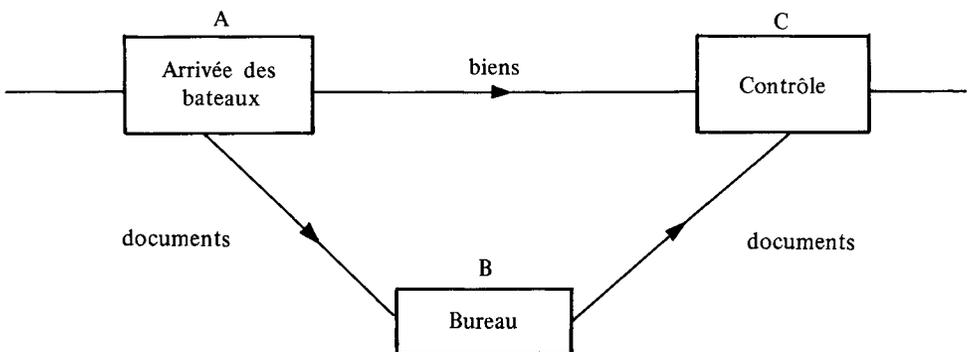


Figure (1.2.1.)

Le contrôle se fait en deux parties :

- 1/ Vérification des documents au bureau *B*.
- 2/ Inspection partielle des biens, comparaison avec les documents (point C).

### 1.3. – PROCEDURE D'IMPORTATION

Pour effectuer des importations à l'échelle commerciale les biens sont groupés. La douane emploie le mot (anglais) "entry" (pluriel : "entries") pour désigner ces groupes. Il est commode d'adopter ce mot dans son sens large, indiquant :

- 1/ Le bien appartenant à un groupe
- 2/ Les documents sur lesquels ces biens sont facturés et déclarés
- 3/ (1) et (2) ensemble.

Il n'y a pas de règle stricte qui gouverne la nature des biens appartenant à une seule entry. Il peut y avoir plusieurs articles, en quantités différentes, portant des droits de douane divers, compris dans une seule entry. Néanmoins une homogénéité approximative et certaines limitations quantitatives sont respectées. En premier lieu ce sont les importateurs qui effectuent le groupement des biens par entries.

Le passage des biens par le port est effectué par des spécialistes d'importation, dont les clients vont recevoir les biens par la suite. C'est le spécialiste qui met au point les documents, les présente à la douane, et qui est responsable de l'entry à ce stade. Une entry sera toujours importée par un seul spécialiste. Par contre, les articles faisant partie d'une seule entry peuvent être destinés à plusieurs clients différents du spécialiste.

Dans le contexte présent, nous entendons par le mot "importateur" l'ensemble des parties intéressées par l'importation d'une entry, et plus ou moins opposées à la douane. Nous préciserons davantage par la suite.

### 1.4. – PROCEDURE DE CONTROLE

Il convient d'abord de distinguer entre la douane entière, et son représentant chargé de contrôler les biens. "La douane" pourra être soit la direction centrale qui dirige la stratégie globale douanière, soit le "joueur" à qui reviennent les droits d'importation. Nous entendons par "le douanier" la

personne qui contrôle les importations et qui prend les décisions journalières nécessaires pour ce travail. Nous maintiendrons cette distinction dans cette étude.

La stratégie de contrôle utilisée par la douane consiste à donner au douanier certaines règles de procédure. Le douanier n'est pas censé contrôler tous les biens ; il en vérifie une certaine proportion. Comme les biens sont déjà groupés (par entries) il est naturel de définir cette proportion par rapport aux entries. La stratégie globale de la douane fixe le taux de contrôle pour tous les douaniers : chacun d'entre eux doit contrôler à long terme, disons, 1 entry sur  $n$ , soit une proportion  $\lambda$ , ( $0 \leq \lambda \leq 1$ ), d'entries.

A court terme, par contre, le douanier est libre d'agir comme il veut. L'échantillonnage des entries (tout en respectant le niveau  $\lambda$ ) ne se fait pas d'une manière purement aléatoire. Le douanier choisit parmi les entries celles qu'il soupçonne le plus d'être erronées.

Ce dernier point est important, car c'est un aspect fondamental de la méthode employée. Le douanier peut parfois avoir de bonnes raisons pour soupçonner une entry. Dans des cas extrêmes il peut même être sûr de la présence d'une erreur. Le choix des entries individuelles est donc fait selon des critères subjectifs connus uniquement du douanier et fondés sur son expérience du travail. La nature de ces "indications" est très diverse, et l'on n'est pas en mesure de les analyser. Pour donner une idée de cette diversité, on pourrait concevoir deux exemples hypothétiques :

1/ Les documents déclarent 90 caisses de quelque article, mais elles semblent être rangées  $4 \times 3 \times 8$ .

2/ Importation d'un exportateur dont la réputation en ce qui concerne les erreurs est mauvaise.

Le premier exemple représente une erreur presque certaine, le deuxième représente un risque statistiquement plus fort que la moyenne.

A titre d'évidence pour soutenir ces arguments, remarquons quelques aspects et statistiques de la procédure de contrôle actuellement en vigueur :

- C'est essentiellement la procédure décrite dans (1.4)
- La valeur de  $\lambda$  employée est de  $1/3$ .
- La fréquence des entries erronées est de l'ordre 1 %.
- La douane perd légèrement dans l'affaire (argent récupéré contre coût de contrôle).

Le réflexe est sans doute de se demander si un taux d'erreurs si faible justifie un  $\lambda$  aussi élevé. Nous espérons qu'en précisant bien les hypothèses que nous aurons à faire nous pourrions peser la valeur des conclusions du modèle mathématique en fonction du degré de "réalité" de ces hypothèses.

Le but de l'étude est donc d'arriver à une formalisation suffisamment convaincante pour que la douane puisse agir, dans une situation incertaine, avec plus de confiance que ce ne serait le cas si elle ne tenait compte que de l'avis d'un (ou plusieurs) douaniers expérimentés.

### 1.5. – ENONCE DU PROBLEME

Le problème que se pose la douane est le suivant :

- 1/ Est-il souhaitable d'imposer une stratégie générale aux douaniers ?
- 2/ Si oui, quelle est la valeur optimale de  $\lambda$  ? (défini dans (1.4)).

### 1.6. – COMMENTAIRE SUR LES METHODES ET HYPOTHESES EMPLOYEES

Avant de développer la description et la formalisation du problème il sera utile de préciser les styles qui seront adoptés. En recherchant une solution du problème de (1.5) –  $\lambda$  optimal – nous tacherons d'assimiler la formalisation à la réalité. Dans la deuxième partie de la thèse (chapitres 4 et 5) nous adopterons la pratique plus habituelle : choisir les hypothèses mathématiquement convenables ou maniables. Ainsi une situation qui se présente dans l'étude "appliquée" fournit l'idée de base d'un problème d'estimation statistique qui sera développé indépendamment.

## CHAPITRE 2

# FORMALISATION DU MODÈLE

### 2.1. – UN JEU ENTRE LA DOUANE ET L'IMPORTATEUR ?

Récapitulons d'abord les aspects évidents de la situation :

1/ Il y a un conflit d'intérêts.

2/ Les rôles des deux parties ne sont pas symétriques.

Les conditions semblent être réunies pour représenter le processus d'importation et de contrôle comme un jeu entre la douane et l'importateur.

En ce qui concerne les utilités, il y a tout d'abord une partie monétaire (les droits d'importations, etc.). Mais il y a aussi une partie non-mesurable : les retards de contrôle, les embouteillages de biens stockés, le degré d'entente entre douaniers et importateurs, etc. Quand il s'agit d'argent il est clair qu'il y a un conflit d'intérêts entre les joueurs. Par contre, si nous pensons uniquement aux aspects non-mesurables, il n'est pas évident que les intérêts des deux parties s'opposent. Nous verrons par la suite que, selon les douaniers eux-mêmes, le comportement typique d'un importateur n'est pas calculé pour rendre optimal son gain (monétaire) d'une manière très agressive. En tous cas les aspects non-mesurables d'une situation si complexe sont presque impossibles à analyser.

Nous nous proposons de traiter désormais uniquement des gains monétaires des deux parties, et de laisser à l'appréciation des experts (la douane) les conséquences non-mesurables des stratégies que nous pourrions être amenés à considérer. Cette optique risque de nous écartier de la théorie classique des jeux. L'intention est simplement de formaliser la situation réelle telle qu'elle existe.

## 2.2. – LA MATRICE DES UTILITES MONETAIRES

		L'importateur	
		pas d'erreur	erreur
La douane	pas de contrôle	0, 0	- m, m
	contrôle	- c, 0	- c, 0

Table (2.2.1)

Les valeurs dans chaque case à gauche sont celles de la douane, à droite celles de l'importateur. Une valeur négative représente une perte :

$c$  est le coût de contrôle (examen) d'une entry

$m$  est la valeur de l'erreur : la différence entre la valeur déclarée et la valeur correcte de ce qui est dû à la douane.

Nous choisissons comme origine le cas où une déclaration correcte a été faite et où le douanier ne perd pas son temps (coût de contrôle) à la vérifier.

La première ligne de la table ne requiert pas d'explication. C'est dans la deuxième ligne, par contre, que repose le caractère particulier de la situation. Si une erreur est découverte par le douanier l'importateur n'est pas directement pénalisé. Normalement le douanier fera le calcul nécessaire pour que l'entry soit réglementaire, et les documents seront renvoyés à l'importateur qui devra recommencer la procédure. Comme celle-ci suit un cycle journalier ou demi-journalier il est évident que l'importateur subit un certain retard dans ce cas. Nous ne considérons pas ce retard comme appartenant aux utilités monétaires.

Une deuxième possibilité existe, si la douane estime que l'erreur correspond à une négligence extrême, voire suspecte : elle peut frapper l'importateur d'une amende.

En cas de fraude flagrante ou éventuellement de contrebande la douane peut enfin engager des poursuites contre l'importateur.

Il est clair que la matrice des utilités (table (2.2.1)) ne traduit pas ces deux dernières éventualités. Dans la pratique la douane ne fait que si rarement appel à ces mesures que, vu la précision avec laquelle nous connaissons la situation générale, il est raisonnable de les négliger. Des statistiques pour soutenir cette prise de position se trouvent dans l'appendice 1 .

### 2.3. – LA FRAUDE ET LES ERREURS

L'argument avancé dans (2.2) revient à dire que les entries non-réglementaires correspondent à des erreurs et non à des fraudes. Cette interprétation est quelque peu étonnante, mais elle est défendue par les douaniers et confirmée par les statistiques de l'appendice 1.

La question immédiate que pose l'enquêteur est : "Existe-t-il des erreurs qui favorisent la douane ?" Plus généralement, quelle est la répartition de leurs valeurs ? En réponse à la première question on apprend que des erreurs "négatives" existent, mais qu'elles sont rares et peu importantes. Nous disposons de statistiques empiriques sur la répartition des valeurs des erreurs (appendice 2).

Pour concilier son point de vue avec les faits, apparemment contradictoires, le douanier rappelle que :

1/ Les diverses catégories de biens sont aujourd'hui si nombreuses et si complexes qu'il peut réellement exister des doutes quand il s'agit d'y placer une entry particulière.

2/ C'est le spécialiste d'importation qui fait ces calculs, alors que c'est son client qui profitera d'un moindre droit d'importation éventuel. Les deux parties étant financièrement indépendantes (obligatoirement), si le spécialiste commet une fraude préméditée ce n'est pas lui-même qui en récoltera le bénéfice éventuel. Inversement, si jamais le spécialiste propose un paiement trop élevé, c'est encore son client qui y perd. Evidemment, dans ce dernier cas, il met en cause son contrat, car le client ne cherche surtout pas un spécialiste qui lui fasse payer trop cher. En fin de compte les deux parties préféreraient probablement qu'il n'y ait jamais d'erreurs à leur compte.

3/ Cela dit, tous les intéressés faisant partie de "l'importateur" seraient d'accord pour que, au cas où il existe un doute réel entre deux ou plusieurs déclarations possibles, on choisisse la plus favorable à l'importateur. Si jamais ce n'est pas la bonne, la douane l'en informera sans risque de pénalisation.

Enfin, si le douanier estime parfois que l'importateur se place quelque part dans la région entre l'erreur et la fraude, il lui accorde presque toujours le bénéfice du doute.

Retournons à la matrice des utilités (Table (2.2.1)). Nous ne pouvons supposer que l'importateur choisisse sa stratégie afin de rendre maximal son gain monétaire, car il serait toujours amené à commettre des erreurs,

ce qui n'est pas le cas. Par contre il est certain que le taux des erreurs varie "inversement" avec le niveau de contrôle ; c'est-à-dire : moins strict le contrôle, plus fréquentes les erreurs. Nous allons supposer que l'importateur s'adapte ou réagit à la stratégie choisie par la douane. Nous ferons des hypothèses sur la manière dont cela s'effectue par la suite. Du point de vue de la douane, son adversaire dans ce "jeu" ressemble plutôt à la nature qu'à un stratège. Cette nature est pourtant d'une tendance hostile.

Il est intéressant à ce stade de remarquer que nous n'avons pas besoin de nous demander si les utilités monétaires sont linéaires puisque nous essayons simplement de prévoir les conséquences financières des stratégies. Les valeurs monétaires concernées sont alors simplement linéaires suivant les règles additives de la comptabilité.

#### 2.4. – LE POINT DE VUE DE LA DOUANE

Nous avons déjà indiqué (1.4) que le contrôle douanier comprend une "stratégie" dictée par la douane et une "tactique" à la discrétion du douanier. La stratégie veut que chaque douanier contrôle, à la longue, une proportion  $\lambda$  des entries. La tactique permet au douanier de choisir comme il veut les entries qui constituent cette proportion  $\lambda$  du total. Le problème que nous avons à résoudre concerne uniquement la stratégie ; nous supposons que la tactique reste toujours la même, quelle que soit la stratégie adoptée. Mais pour construire un modèle du processus qui permette de rechercher une stratégie optimale, il faudra d'abord formaliser la tactique douanière.

Le principe de base de cette tactique est d'exploiter l'expérience du douanier qui arrive, dans une certaine mesure, à "flairer" les erreurs (ainsi que l'absence d'erreurs). Les douaniers prétendent (sans pouvoir le démontrer) que la plupart des erreurs sont rattrapées dans la proportion  $\lambda$  contrôlée (c'est-à-dire qu'il n'y en a pas beaucoup dans le complément,  $1-\lambda$ ). Il faut supposer d'autre part qu'un douanier expérimenté se fait une idée non seulement de la "probabilité d'existence" d'une erreur, mais aussi de la valeur de cette erreur (éventuelle). Cette analyse est soutenue par les douaniers. Nous nous proposons de fonder une formalisation de la tactique douanière là-dessus.

Nous suggérons que le douanier fait une estimation "subjective" ou "personnelle" des valeurs inconnues. Malheureusement il ne semble pas que ces estimations dépendent d'une manière simple des attributs des entries.

On ne peut pas, par exemple, simplement placer chaque entry dans une table de contingence, et trouver la proportion des erreurs pour chaque catégorie. (Voir l'exemple hypothétique dans (1.4)).

## 2.5. – HYPOTHESES DE BASE DU MODELE TACTIQUE

Nous supposons que pour chaque entry le douanier arrive à la décision de l'examiner ou non en faisant deux "estimations subjectives" :

1/ de la probabilité d'erreur.

2/ de la quantité d'argent en jeu (la valeur estimée de l'erreur).

Ensuite il se fera un produit de ces deux quantités qui, suivant qu'il est supérieur ou inférieur à une certaine valeur critique, entraînera le contrôle ou non. Le niveau critique sera choisi de manière à ce qu'une proportion,  $\lambda$ , des entries soit examinée à long terme.

Il convient, dans la formalisation de ces idées, de traiter les deux processus d'estimation subjective d'abord indépendamment, et ensuite de les combiner.

## 2.6. – MODELE TACTIQUE DE L'ESTIMATION SUBJECTIVE DE LA PROBABILITE D'ERREUR

Supposons que le douanier décide d'examiner en fonction seulement de son estimation subjective de la probabilité d'erreur, sans tenir compte de la valeur éventuelle de l'erreur. Or, on pourrait supposer qu'à chaque entry le douanier associe une unique valeur dans l'intervalle, disons,  $[0, 1]$ . Mais une telle théorie est difficile à manier et à interpréter. Il est peut-être plus réaliste de supposer que le douanier place chaque entry dans une catégorie déterminée parmi  $n$  (par exemple, de risque fort, moyen, ou bas). Il est douteux qu'une personne puisse faire les très fines différences implicites dans l'hypothèse d'une échelle continue de probabilité subjective. D'autre part le système des catégories offre la possibilité d'une vérification expérimentale, comme nous le verrons par la suite. Dans un tel modèle tactique on suppose que le douanier contrôle tout d'abord les entries de la catégorie la plus soupçonnée, puis la deuxième catégorie, et ainsi de suite jusqu'à atteindre la proportion  $\lambda$ .

## 2.7. – NOTATIONS

Définissons une variable binaire,  $z$ , pour chaque entry :

$z = 0$  si l'entry est réglementaire

$z = 1$  si l'entry est erronée .

Nous allons utiliser un indice inférieur pour indiquer la valeur de  $z$  :

$f_0 = P[z = 0]$  = la probabilité pour que l'entry soit réglementaire.

$f_1 = P[z = 1]$  = la probabilité pour que l'entry soit erronée.

L'estimation subjective d'erreur consiste à placer chaque entry dans une des  $n$  catégories. Pour indiquer la catégorie nous employons l'indice supérieur,  $j$  : ( $j = 1, 2, \dots, n$ ).

$f_i^j = P[z = i \text{ pour une entry dans la catégorie } j]$

$\lambda^j =$  probabilité de contrôle dans la catégorie  $j$ .

$f^j =$  fréquence marginale des entries dans la catégorie  $j$ .

$V^j =$  la valeur à la douane du sous-jeu formé par les entries de la catégorie  $j$ .

Certaines relations sont évidentes :

$$f^j = f_0^j + f_1^j$$

$$f_i = \sum_{j=1}^n f_i^j$$

$$\lambda = \sum_{j=1}^n \lambda^j f^j = \text{la proportion moyenne d'entries contrôlées.}$$

## 2.8. – FORMALISATION DE LA TACTIQUE

$$\text{Maximiser } V = \sum_{j=1}^n V^j f^j \quad (2.8.1)$$

$$\text{avec } \lambda = \sum_{j=1}^n \lambda^j f^j \quad (2.8.2)$$

$$0 \leq \lambda \leq 1 \quad (j = 1, 2, \dots, n) \quad (2.8.3)$$

Adoptant la convention que la catégorie  $n$  est celle qui contient la plus haute proportion d'erreurs, nous supposons que le douanier contrôle progressivement les catégories  $n, n - 1, \dots$  jusqu'à ce que (2.8.2) soit satisfaite.

Considérons le sous-jeu comprenant uniquement les entries de la catégorie  $j$ . Tout d'abord il faut noter les fréquences conditionnelles :

$$P[z = i \mid \text{catégorie} = j] = f_i^j / f^j \quad (2.8.4)$$

et

$$V^j = -c \frac{f_0^j \lambda^j}{f^j} - c \frac{f_1^j \lambda^j}{f^j} - m \frac{f_1^j}{f^j} (1 - \lambda^j) \quad (2.8.5)$$

La tactique douanière consiste à poser :

$$\left\{ \begin{array}{l} \lambda^n = \lambda^{n-1} = \dots = \lambda^{r+1} = 1 \\ \lambda^r = \alpha \quad (0 < \alpha < 1) \\ \lambda^{r-1} = \dots = \lambda^1 = 0. \end{array} \right. \quad (2.8.6)$$

## 2.9 - VARIATION DE $\lambda_1$ avec $\lambda$

$$\begin{aligned} \lambda_1 &= P[\text{Contrôle} \mid z = 1] \\ &= \sum_{j=1}^n \frac{\lambda^j f_1^j}{f_1} \end{aligned}$$

Or

$$\begin{aligned} \lambda^j &= 1 \text{ pour } r + 1 \leq j \leq n \\ &= \alpha \text{ pour } j = r \\ &= 0 \text{ pour } 1 \leq j \leq r - 1 \end{aligned}$$

Ainsi

$$\lambda_1 = \sum_{j=r+1}^n f_1^j / f_1 + \alpha f_1^r / f_1 \quad (2.9.1)$$

et

$$\lambda = \sum_{j=1}^n f^j \lambda^j = \sum_{j=r+1}^n f^j + \alpha f^r \quad (2.9.2)$$

Donc

$$\frac{d\lambda_1}{d\lambda} = \frac{d\lambda_1}{d\alpha} \frac{d\alpha}{d\lambda} = \frac{f_1^r}{f_1 f^r} .$$

La courbe de  $\lambda_1$  contre  $\lambda$  est constituée d'une série de droites, de pentes  $f_1^j/f_1 f^j$  ( $j = 1, 2 \dots n$ ).

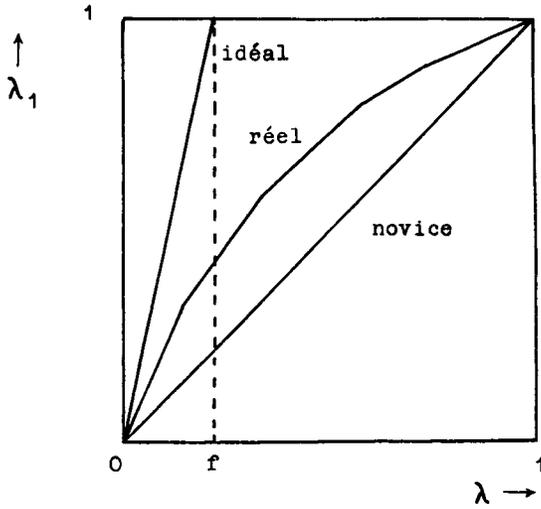


fig. (2.9.1)

## 2.10. – DOUANIERS IDEAL ET NOVICE

Nous pouvons imaginer le douanier "idéal", c'est celui qui sait à l'avance si l'entry est erronée, ainsi que le douanier "novice", qui contrôle une proportion des entries choisies au hasard.

Pour le douanier idéal on suppose :

$$f_0^j = 0 \text{ pour } 2 \leq j \leq n, \quad f_0^1 = 1$$

$$f_1^j = 0 \text{ pour } 1 \leq j \leq n-1, \quad f_1^n = 1$$

et pour le douanier novice on suppose  $f_1^j/f^j = f_1$  pour tout  $j$ .

Le domaine dans lequel doit se trouver la courbe du douanier réel est donc borné par ces deux courbes limites. Il sera commode de caractériser le douanier par la répartition des entrées dans les différentes catégories :

$$P[\text{Catégorie} = j \mid z = 0] = f_0^j/f_0 = p^j \quad (2.10.1)$$

$$P[\text{Catégorie} = j \mid z = 1] = f_1^j/f_1 = q^j \quad (2.10.2)$$

Les  $p^j$ ,  $q^j$  ( $j = 1, 2, \dots, n$ ) serviront à mesurer le degré de perfection du douanier en tant qu'estimateur subjectif.

## 2.11. – SOLUTION OPTIMALE

Substituant (2.9.1) et (2.9.2) dans (2.8.1) on obtient la valeur du jeu à la douane :

$$V = \sum_{j=r+1}^n -f^j c - (1 - \alpha) m f_1^r - \alpha f^r c - \sum_{j=1}^{r-1} m f_1^j \quad (2.11.1)$$

Si  $f_0^j$ ,  $f_1^j$ ,  $c$ ,  $m$  sont constants, alors

$$\frac{dV}{d\lambda} = \frac{dV}{d\alpha} \frac{d\alpha}{d\lambda} = m f_1^r / f^r - c \quad (2.11.2)$$

La solution optimale pour la douane correspondrait à l'un des sommets de la figure (2.9.1).

## 2.12. – MODELE TACTIQUE : ESTIMATION SUBJECTIVE DE LA VALEUR D'UNE ERREUR

Considérons maintenant le deuxième stade de l'estimation subjective d'erreurs. Le douanier ne s'occupera ici que de la valeur d'une erreur éventuelle. Nous postulons un modèle pour la génération d'une erreur suivant la réalisation de deux processus aléatoires. Le premier processus fournit une valeur à partir d'une densité qui correspond à la répartition des erreurs (app. 2). Le second est une variable binaire,  $z$ , avec paramètre,  $f_1$ . Si  $z = 0$

il n'y a pas d'erreur ; si  $z = 1$  l'erreur existe et prend la valeur du premier processus. Ainsi à chaque entry le douanier associe une valeur, disons  $\tilde{m}$ , qui est son estimation subjective de la valeur probable d'une erreur éventuelle. Le douanier très expérimenté sera capable de prévoir cette valeur, là où une erreur existe, avec une bonne précision. Le modèle tactique veut que soit contrôlée toute entry pour laquelle  $\tilde{m}$  dépasse une certaine valeur critique,  $S$ . La contrainte stratégique est respectée tant que :

$$\int_S^{\infty} f(\tilde{m}) d\tilde{m} = \lambda \quad (2.12.1)$$

### 2.13. – HYPOTHESES DU MODELE

Afin de réaliser une simulation du modèle les hypothèses suivantes sont faites :

1/ A chaque entry correspond une valeur hypothétique d'erreur,  $m$ . La densité,  $f(m)$ , est connue.

2/ Pour chaque entry le douanier fait une estimation subjective,  $\tilde{m}$ .

3/ La répartition de  $\tilde{m}$  dépend de  $m$  et de la précision du douanier. L'équation qui relie  $\tilde{m}$  à  $m$  est :

$$m = \tilde{m} + \epsilon$$

où  $\epsilon$  est une variable aléatoire qu'il est commode de supposer uniformément répartie dans l'intervalle  $[-d, +d]$ .

### 2.14. – NOTATION

$f(m)$ ,  $f(\tilde{m})$  sont les densités de  $m$ ,  $\tilde{m}$ .

$f(m, \tilde{m})$  est la densité du couple  $(m, \tilde{m})$ .

$V_m^{\tilde{m}}$  est la valeur du jeu à la douane d'une entry caractérisée par  $m$  et  $\tilde{m}$ .

$\lambda^{\tilde{m}}$  est la probabilité de contrôle (fonction de  $\tilde{m}$ ).

Donc

$$\begin{aligned} V_m^{\tilde{m}} &= -c(1 - f_1)\lambda^{\tilde{m}} - cf_1\lambda^{\tilde{m}} - mf_1(1 - \lambda^{\tilde{m}}) \quad (2.14.1) \\ &= -c\lambda^{\tilde{m}} + mf_1\lambda^{\tilde{m}} - mf_1. \end{aligned}$$

Or,  $V$  est tout simplement l'espérance de  $V_m^{\tilde{m}}$  pour tout  $m, \tilde{m}$  :

$$V = \iint V_m^{\tilde{m}} f(m, \tilde{m}) dm d\tilde{m} \quad (2.14.2)$$

$$= -c\lambda - mf_1 + \iint f_1 m \lambda^{\tilde{m}} f(m, \tilde{m}) dm d\tilde{m} \quad (2.14.3)$$

La décision de contrôler sera prise de manière à rendre optimale l'espérance de  $\tilde{m}$ , ou (2.14.3) avec  $\tilde{m}$  à la place de  $m$ .

## 2.15. – DOUANIERS IDEAL ET NOVICE

Le paramètre qui traduit l'expérience du douanier est évidemment  $d$  (2.13). Le douanier idéal pose toujours  $\tilde{m} = m$  ( $d = 0$ ). C'est-à-dire que si une erreur existe il en sait la valeur avant d'examiner. Le douanier novice établit  $\tilde{m}$  selon la règle :

$$\tilde{m} = m + \epsilon$$

où  $\epsilon$  est uniformément réparti dans l'intervalle  $[-d, +d]$ , et où  $d \rightarrow \infty$ . Cet aspect du modèle n'est pas très convaincant, surtout que des valeurs négatives de  $\tilde{m}$  deviennent possibles. Dans la simulation numérique il sera commode d'ignorer les valeurs négatives. Le modèle peut néanmoins être plausible pour des petites valeurs de  $d$ .

## 2.16. – UNE STRATEGIE OPTIMALE

Il est clair que cette formalisation permet de rechercher une stratégie optimale. Pour simplifier les choses, considérons le douanier idéal. Posant  $\tilde{m} = m$  dans (2.14.3) on trouve :

$$V = -c\lambda - \bar{m}f_1 + \int_S^{\infty} f_1 m f(m) dm \quad (2.16.1)$$

où  $S$  est défini par

$$\lambda = \int_S^{\infty} f(m) dm \quad (2.16.2)$$

$$\frac{d\lambda}{dS} = -f(m)_{m=S} \quad (2.16.3)$$

Dérivant (2.16.1) pour  $c$  et  $f_1$  supposés constants :

$$\frac{dV}{d\lambda} = -c - f_1 S f(m)_{m=S} \cdot 1/f(m)_{m=S} = -c + f_1 S \quad (2.16.4)$$

Le point d'équilibre, s'il existe, est donné par la condition :

$$S_{\text{opt}} = c/f_1$$

ce qui traduit le fait que l'espérance de gain pour les valeurs limites est égale au coût de contrôle. Dans le cas du douanier réel l'algèbre n'est pas aussi simple, car la valeur critique,  $S$ , définie par :

$$\int_S^\infty f(\tilde{m}) d\tilde{m} = \lambda \quad (2.16.5)$$

n'est pas la même que  $S'$  où :

$$\int_{S'}^\infty f(m) dm = \lambda. \quad (2.16.6)$$

On voit cependant que dans la structure de l'estimation subjective que nous avons postulée, un raisonnement essentiellement pareil permet d'établir la possibilité d'une solution optimale stratégique.

## 2.17. – MODELE COMPOSE

Nous proposons maintenant un modèle "composé", c'est le produit des deux stades précédents. Supposons que pour chaque entry le douanier :

1/ la place dans une catégorie,  $j$ , de risque d'erreur.

2/ fait une estimation de la valeur d'une erreur éventuelle.

Nous considérons ces deux processus d'estimation comme statistiquement indépendants. La notation correspond à une synthèse des deux modèles précédents :

$\lambda^{\tilde{m}j} = P[\text{contrôler une entry dont les estimations subjectives sont représentées par } \tilde{m} \text{ et } j]$

$V_m^{\tilde{m}j} = \text{Valeur à la douane du sous-jeu que constitue une entry individuelle, caractérisée par } m, \tilde{m}, \text{ et } j.$

La valeur,  $V_m^{\tilde{m}^j}$ , se compose comme dans les modèles précédents :

$$\begin{aligned} V_m^{\tilde{m}^j} &= -c \lambda^{\tilde{m}^j} f_0^j / f^j - c \lambda^{\tilde{m}^j} f_1^j / f^j - m(1 - \lambda^{\tilde{m}^j}) f_1^j / f^j \\ &= -c \lambda^{\tilde{m}^j} - m f_1^j / f^j + m \lambda^{\tilde{m}^j} f_1^j / f^j. \end{aligned} \quad (2.17.1)$$

La valeur du jeu à long terme,  $V$ , s'obtient en intégrant  $V_m^{\tilde{m}^j}$  par rapport à  $m$  et  $\tilde{m}$ , et en faisant la somme pour tout  $j$  :

$$V = \sum_{j=1}^n \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} [-c \lambda^{\tilde{m}^j} f^j - m f_1^j + m \lambda^{\tilde{m}^j} f_1^j] f(m, \tilde{m}) dm d\tilde{m}$$

et puisque les deux processus sont supposés indépendants

$$V = -c \lambda - \bar{m} f_1 + \sum_{j=1}^n \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} m \lambda^{\tilde{m}^j} f_1^j f(m, \tilde{m}) dm d\tilde{m} \quad (2.17.2)$$

où  $\bar{m}$  est la valeur moyenne des erreurs,  $m$ .

## 2.18. – SOLUTION TACTIQUE

Comme dans le modèle précédent, nous supposons que le contrôle s'effectue quand  $\tilde{m}$  est supérieur à une valeur critique,  $S$ . Il faut cependant tenir compte de la catégorie,  $j$ . Nous postulons l'existence d'un seuil,  $S^j$ , pour chaque catégorie, de manière à ce que l'entry soit contrôlée si  $\tilde{m} \geq S^j$  :

$$\tilde{m} < S^j \rightarrow \lambda^{\tilde{m}^j} = 0$$

$$\tilde{m} \geq S^j \rightarrow \lambda^{\tilde{m}^j} = 1.$$

La contrainte stratégique s'écrit :

$$\begin{aligned} \lambda &= \sum_{j=1}^n \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \lambda^{\tilde{m}^j} f^j f(m, \tilde{m}) dm d\tilde{m} \\ &= \sum_{j=1}^n \int_S^{\infty} f(\tilde{m}) d\tilde{m}. \end{aligned} \quad (2.18.1)$$

Or, le douanier contrôlera les entries afin de maximiser son espérance de gain. C'est-à-dire qu'il tachera de maximiser le dernier terme de (2.17.2), avec

$\tilde{m}$  à la place de  $m$ . Maximiser :

$$\begin{aligned} V' &= \sum_{j=1}^n \int_{-\infty}^{\infty} \tilde{m} \lambda^{\tilde{m}^j} f_1^j f(\tilde{m}) d\tilde{m} \\ &= \sum_{j=1}^n \int_{S^j}^{\infty} \tilde{m} f_1^j f(\tilde{m}) d\tilde{m} \end{aligned} \quad (2.18.2)$$

d'après (2.18.1). On applique la méthode des multiplicateurs de Lagrange :

$$dV' = \sum_{j=1}^n -S^j f_1^j f(\tilde{m} = S^j) dS^j$$

où  $f(\tilde{m} = S^j)$  désigne la densité de  $\tilde{m}$  au point où  $\tilde{m} = S^j$ .

$$d\lambda = \sum_{j=1}^n -f^j f(m = S^j) dS^j$$

Désignant par  $M$  le multiplicateur :

$$\sum_{j=1}^n S^j f_1^j f(\tilde{m} = S^j) - M \sum_{j=1}^n f^j f(\tilde{m} = S^j) = 0$$

ce qui donne

$$\begin{aligned} \sum_{j=1}^n (S^j f_1^j - M f^j) &= 0 \\ S^j f_1^j &= M = \text{const.} \end{aligned} \quad (2.18.3)$$

Cette relation servira à définir la solution tactique du modèle de contrôle que nous emploierons dans la simulation.

## 2.19. – REACTION DE L'IMPORTATEUR

Le modèle tactique établi, on peut reprendre le thème des aspects "stratégiques" déjà entamé dans (2.4). Le problème à résoudre concerne le niveau optimal de  $\lambda$ . Il est évident qu'un facteur important dans toute prédiction sera la réaction de l'importateur face à un changement de la stratégie douanière. C'est-à-dire que l'on voudrait savoir la relation entre  $f_1$  et

$\lambda$ , ce qui, par la nature des circonstances, est inconnu. Néanmoins la faible fréquence de la fraude et la motivation que nous avons prêtée aux importateurs suggèrent une règle plausible pour gouverner ce comportement.

D'après l'analyse descriptive de (2.2) et (2.3) l'importateur se comporte comme un joueur qui, tout en préférant ne pas payer les droits d'importation, y est obligé à cause du risque d'être imposé par la douane. Un taux d'erreurs faible est cependant toléré sans que l'importateur soit soupçonné de fraude. Supposons donc que pour un  $\lambda$  donné (stratégie de la douane) un équilibre s'établit de sorte que la fréquence des erreurs reste tout juste au niveau admis par la douane. Ce raisonnement nous permet de construire un "indice" d'erreurs commises. Ce qui compte, évidemment, ce sont les erreurs qui sont exposés par la douane. C'est-à-dire :

$$\lambda_1 f_1 = C_1 = \text{const.} \quad (2.19.1)$$

ce qui traduit le fait que, de toutes les entries, la proportion qui est examinée et trouvée fautive reste constante.

La règle (2.19.1) traite toutes les entries à poids égal, ce que nous n'avons pas supposé être le cas en formulant la tactique douanière. Une autre manière évidente de pondérer les erreurs révélées consiste tout simplement à les multiplier par leur valeur :

$$V' = \int_0^{\infty} m \lambda_{m1} f(m) dm = C_2 = \text{const.} \quad (2.19.2)$$

Cette équation correspond à un équilibre où la quantité d'argent récupéré par la douane reste constante. Les valeurs actuelles de ces deux quantités nous sont fournies par les statistiques de l'appendice 1. Nous y reviendrons par la suite.



## CHAPITRE 3

# ÉVALUATION NUMÉRIQUE DU MODÈLE

### 3.1. – OBJECTIF DE L'ETUDE

Les diverses hypothèses faites dans le chapitre précédent nous ont amenés au point où il est possible de simuler numériquement le contrôle douanier. Comme certains des paramètres sont inconnus ou mal définis, il faudra utiliser une gamme large de valeurs. Un tel paramètre est l'expérience du douanier. Nous avons postulé un douanier "novice" et un douanier "idéal" et des cas intermédiaires peuvent se créer par des combinaisons linéaires de ces deux extrêmes.

### 3.2. – SIMULATION PAR ORDINATEUR

Un programme a été écrit en Fortran (sur un ordinateur Atlas) fournissant une série de graphiques aussi bien que des valeurs numériques. Les paramètres d'entrée sont les suivants :

1/  $p^j, q^j$  qui donnent une mesure du succès avec laquelle le douanier arrive à estimer la probabilité d'une erreur, (2.10).

2/  $d$ , qui représente la précision de l'estimation subjective de  $m$ , (2.13).

3/  $f_1$ , la fréquence des erreurs.

Le coût moyen d'un contrôle,  $c$ , est évalué à 0,45 (voir appendice 2).

Les valeurs de sortie (calculées par le programme) sont :

$\lambda$  : niveau stratégique de contrôle.

$\lambda_1$  : probabilité pour qu'une entry erronée soit contrôlée.

$V$  : valeur du jeu à la douane.

$V'$  : argent récupéré par la douane (2.19.2).

Les valeurs d'entrées fixées, le programme pose  $\lambda = 1$ , puis le diminue par petites étapes jusqu'à  $\lambda = 0$ , respectant la règle (2.18.3). Deux séries de graphiques ont été produites, illustrant  $\lambda_1$  contre  $\lambda$ , et  $V$  contre  $\lambda$ . Pour chaque ensemble de valeurs d'entrée le maximum de  $V$  et les valeurs de  $\lambda_{opt}$ ,  $V'$ , et  $\lambda_1$  qui y correspondent ont été déterminées.

Le nombre de catégories,  $n$ , associées à l'estimation subjective (2.6), a été fixé à trois (conforme à l'expérience pratique pour les mesurer décrite dans (3.4)). Six niveaux différents entre novice et idéal ont été définis comme suit :

$$\begin{array}{ll} \text{Novice (N)} : p^1 = q^1 = 1/3 & \text{Idéal (I)} : p^1 = 1 \quad q^1 = 0 \\ p^2 = q^2 = 1/3 & p^2 = 0 \quad q^2 = 0 \\ p^3 = q^3 = 1/3 & p^3 = 0 \quad q^3 = 1 \end{array}$$

Les valeurs intermédiaires s'obtenant par combinaison linéaire :

$$(1 - \alpha)N + \alpha I$$

avec  $\alpha$  égal à : 1,0 (Idéal)  
0,9  
0,75  
0,5  
0,25  
0,0 (Novice)

La fréquence des erreurs a pris 5 valeurs :

$f_1 = 0,005 ; 0,01 ; 0,02 ; 0,05 ; 0,1$ . (La valeur observée (appendice 1) est de l'ordre  $f_1 = 0,01$ ).

Le paramètre,  $d$ , a pris les quatre valeurs :

$$\begin{array}{ll} d = 0 & \text{(Idéal)} \\ d = 10 & \\ d = 100 & \\ d = 25\,000 & \text{(Novice)} \end{array}$$

L'estimateur,  $\tilde{m}$ , suit le traitement de (2.12), la règle exacte étant :  $\tilde{m} = m + \epsilon$ , où  $\epsilon$  est réparti uniformément

$$\begin{array}{ll} \text{dans } [-m, d] & \text{si } d > m \\ \text{dans } [-d, d] & \text{si } d \leq m. \end{array}$$

Une sélection de graphiques est présentée dans l'appendice 4. Les courbes de  $\lambda_1$  contre  $\lambda$  illustrent les différentes qualités du douanier. La

figure 4, avec  $d = 25\ 000$  correspond au modèle tactique de (2.6). A mesure que  $d$  diminue, les droites (fig. 4) se transforment en une courbe (figs. 3, 2, 1). Les figures 7 et 8 correspondent aux douaniers novice et idéal (2.10). Les figures 9 à 16 sont les traces de  $V$  contre  $\lambda$  ayant les mêmes valeurs de paramètres que les figures 1 à 8.

### 3.3. – INTERPRETATION DES RESULTATS NUMERIQUES

Les résultats de l'évaluation numérique du modèle sont présentés dans l'appendice 4. Les tables dans cet appendice donnent, pour  $f_1$ ,  $[p^j, q^j]$  et  $d$  spécifiés, les valeurs optimales de  $\lambda$ ,  $V$ ,  $\lambda_1$  et  $V'$ . En examinant les tables nous pouvons dégager les traits suivants :

- (i) Quelles que soient les valeurs de  $[p^j, q^j]$  et de  $d$ ,  $\lambda_{\text{opt}}$ ,  $\lambda_{1\text{opt}}$ ,  $V'_{\text{opt}}$  sont fonctions croissantes de  $f_1$  ;  $V_{\text{opt}}$  est une fonction décroissante de  $f_1$ .
- (ii) Suivant la notation de (3.2) où  $[p^j, q^j] = (1 - \alpha)N + \alpha I$ ,  $V$  est une fonction croissante de  $\alpha$  ( $d$  et  $f_1$  fixés).
- (iii) Quand  $f_1$  et  $[p^j, q^j]$  sont fixés,  $V$  est une fonction décroissante de  $d$ .

Toutes ces conclusions sont intuitivement correctes. Une autre tendance générale est apparente :

- (iv) Quand  $d$  est petit ( $\sim 10$ ) une augmentation de  $d$  affecte à peine le niveau optimal de  $\lambda$ , mais fait baisser  $V_{\text{opt}}$ .

Les deux facteurs les plus importants sont en fait la fréquence des erreurs,  $f_1$ , et la faculté chez les douaniers de "flairer" les erreurs, c'est-à-dire le paramètre  $[p^j, q^j]$ . Formellement donc,  $V$  serait maximal pour la douane si chaque douanier individuellement adoptait une stratégie optimale en fonction de ses paramètres personnels,  $[p^j, q^j]$ . Ce serait une solution fort théorique qui nécessiterait que chaque douanier prenne conscience, objectivement, de sa propre habileté. D'autre part la douane n'aurait aucun contrôle sur le nombre total d'entries examinées.

En ce qui concerne l'influence de  $f_1$ , il est fort possible que la fréquence des erreurs varie beaucoup entre les différents ports. Il serait donc logique d'imposer une stratégie locale,  $\lambda$ , en fonction du niveau d'erreurs dans chaque port ou autre lieu d'importation.

### 3.4. – UNE EXPERIENCE POUR MESURER LE “DOUANIER REEL”

Un facteur très important parmi ceux qui sont inconnus est “l’expérience” du douanier. Afin d’obtenir des valeurs approximatives des paramètres  $[p^j, q^j]$  et  $d$ , il fut décidé d’effectuer une expérience pratique destinée à les mesurer. Cette expérience devait permettre de déterminer si vraiment le douanier possède ce pouvoir de discernement auquel il prétend, et quelle est sa précision en tant qu’estimateur subjectif.

L’expérience proposée exigeait que le douanier, avant de procéder à un contrôle par sa propre méthode, et selon ses propres critères, remplisse un questionnaire pour chaque entry :

1/ Donner une indication de ce que vous considérez être la probabilité pour que cette entry soit erronée.

2/ Estimer la valeur d’une erreur éventuelle.

Puis : effectuer le contrôle ou non, comme d’habitude, et en noter les résultats.

Une telle expérience ne révélerait rien sur les entries non-contrôlées. Tant qu’il n’y a pas de contrôle à 100 % on ne sait pas combien d’erreurs échappent au contrôle partiel. Compte tenu des difficultés d’ordre psychologique (et aussi du coût) de ce genre d’exercice, un schéma simplifié fut retenu. Les entries seraient classées selon la probabilité estimée d’erreur dans trois catégories disjointes :

1/ Faible, moyenne, ou forte chance d’erreur.

D’autre part il a été décidé d’assimiler la forme de la deuxième question à celle de la première. Il fallait placer chaque entry selon la valeur estimée d’une erreur éventuelle :

2/ Moins de £ 5, entre £ 5 et £ 50, plus de £ 50.

Le choix de 5 et 50 comme limites d’intervalles a été fait par les douaniers.

Les résultats de cette expérience sont résumés dans l’appendice 3, table 1. On remarque qu’en 500 entries n’ont été trouvées que 3 erreurs. Ce taux d’erreurs est d’ailleurs assez représentatif de la situation générale. Par conséquent il faudrait un échantillon d’effectif prodigieux pour assurer un nombre d’erreurs qui soient statistiquement valable. La direction de la douane n’était pas disposée à autoriser une expérience aussi vaste, et un deuxième plan d’expérience a enfin été proposé. Au lieu de poser toutes ces questions au douanier au moment du contrôle des biens, on travaillerait sur les documents d’entries déjà passées par la douane, qui sont renvoyés au siège central de celle-ci. L’expérience serait effectuée par un employé expérimenté de la douane avec une connaissance intime de la procédure d’importation. L’échantillon comprendrait

uniquement des entries erronées et l'employé les trierait dans les catégories décrites ci-dessus. Il tacherait de prendre sa décision en fonction de l'information disponible par les documents et d'agir comme s'il effectuait un vrai contrôle lors de l'importation.

Il est évident qu'une telle expérience n'est pas scientifiquement très rigoureuse. Néanmoins, dans une situation d'information imparfaite il faut exploiter tous les moyens dont on dispose. Un échantillon supplémentaire de 71 entries erronées a été analysée de cette manière. Le résumé des résultats se trouve dans l'appendice 3, table 2.

Ces expériences nous fournissent une base, aussi peu solide soit-elle, par laquelle nous pouvons mesurer le degré d'expérience du douanier. En ce qui concerne la première question, les résultats des deux expériences se résument ainsi :

Catégorie ( $j$ )	Entries correctes	Entries erronées
1	209	2
2	275	22
3	13	50
Total	497	74

Table (3.4.1)

Nous supposons dans ce résumé que les entries du premier échantillon qui ne furent pas examinées lors de l'importation (à savoir quelques  $500 \times 2/3$  entries) sont toutes réglementaires. Le taux d'erreurs est en tous cas si faible que nous ne perdons pas ainsi beaucoup de précision. Ces chiffres nous permettent d'estimer les paramètres  $[p^j \ q^j]$  définis dans (2.10) :

$p^j = f_0^j/f_0$	$j$	$p$	$q$
	1	0,4205	0,0270
	2	0,5534	0,2972
$q^j = f_1^j/f_1$	3	0,0261	0,6758

Table (3.4.2)

L'analyse de la deuxième question – la précision d'estimation des valeurs d'erreurs – est plus délicate. Il s'agit de comparer la valeur observée de l'erreur avec l'intervalle dans lequel l'estimation se trouve. Le traitement de cette question, dans des hypothèses un peu différentes de celles qui ont

été adoptées pour la simulation par ordinateur, forme le sujet du chapitre 4. Dans le contexte présent il suffit de noter qu'à partir de cette analyse-là une valeur pour le paramètre  $d$ , de l'ordre de 10 semble correspondre à la précision des prévisions.

### 3.5. – UNE PREVISION DE L'EQUILIBRE OPTIMAL

Munis de ces statistiques – aussi inexactes soient-elles – sur le “douanier réel”, nous pouvons prévoir les points d'équilibre selon les deux critères de (2.19.1) et (2.19.2). Les statistiques de l'appendice 1 nous fournissent directement des valeurs pour les constantes de ces deux équations. Le nombre d'entries examinées et erronées s'élève à 1 500 dans une population totale de  $3 \times 82\,000$  ( $\lambda = 1/3$ )

$$C_1 = 1\,500 / (3 \times 82\,000) = 0,0061 \quad (3.5.1)$$

de même :

$$C_2 = 27\,000 / (3 \times 82\,000) = 0,11. \quad (3.5.2)$$

Connaissant  $C_1$  on peut très simplement résoudre l'équation :

$$\lambda_{1\text{ opt}} f_1 = 0,0061 \quad (3.5.3)$$

à partir de l'appendice 4 en construisant la courbe de  $\lambda_{1\text{ opt}} f_1$  contre  $\lambda_{\text{opt}}$ . De même la courbe de  $V'_{\text{opt}}$  contre  $\lambda_{\text{opt}}$  fournit la valeur de  $\lambda$  qui correspond à  $V'_{\text{opt}} = 0,11$ . La table (3.5.1) contient des valeurs (extraites de l'appendice 4) pour le “douanier réel”, c'est-à-dire avec les  $[p^j, q^j]$  de la table (3.4.2) et avec  $d = 10$ .

$f_1$	$\lambda_{1\text{ opt}}$	$f_1 \lambda_{1\text{ opt}}$	$V'_{\text{opt}}$	$\lambda_{\text{opt}}$
0,005	0,53	0,00265	0,0604	0,026
0,01	0,62	0,0062	0,136	0,053
0,02	0,67	0,0134	0,305	0,111

Table (3.5.1)

Les points d'équilibre prévus par les deux règles sont respectivement :  $\lambda_{\text{opt}} = 0,052$  et  $\lambda_{\text{opt}} = 0,044$ . Comme les estimations des paramètres  $d$  et  $[p^j, q^j]$  ont été obtenues d'une manière assez approximative on peut encore avoir recours à la technique de combinaisons linéaires.

Désignant le douanier réel de (3.4) par  $D$ , et les douaniers novice et idéal par  $N$  et  $I$  comme dans (3.2), on peut former un "intervalle" autour de  $D$  comme suit :

$$[(N + D)/2, D, (I + D)/2]$$

La procédure équivalente pour le paramètre  $d$  peut se réaliser en prenant la moitié et le double de la valeur centrale (= 10). De cette manière nous avons construit un domaine borné par les douaniers "deux fois moins expérimenté" et "deux fois plus expérimenté" que le douanier estimé en (3.4).

La table (3.5.2) résume les solutions des deux équations dans les divers cas. Dans chaque case la valeur écrite en haut correspond à la règle (2.19.1), et celle d'en bas correspond à (2.19.2).

Valeurs de  $\lambda_{opt}$

	$d = 5$	$d = 10$	$d = 20$
$0,5 I + 0,5 D$	0,026 0,023	0,027 0,025	0,027 0,026
$D$ (douanier estimé)	0,053 0,039	0,052 0,044	0,052 0,046
$0,5 D + 0,5 N$	0,20 0,098	0,18 0,098	0,18 0,08

Table (3.5.2)

### 3.6. – CONCLUSIONS POUR LA STRATEGIE DOUANIERE

L'un des avantages de cette méthode de solution est que l'on n'a pas besoin de savoir la valeur exacte de  $f_1$ , qui est "éliminé" du calcul. Implicitement nous avons supposé que les paramètres descriptifs douaniers ( $[p^j, q^j]$ ,  $d$  etc.) s'appliquent aussi bien dans les conditions dans lesquelles ont été recueillies les statistiques de l'appendice 1 que dans la situation qui se présenterait si la douane adoptait une nouvelle stratégie.

La table (3.5.2) indique qu'afin de s'assurer le meilleur gain monétaire, la douane devrait adopter une stratégie globale qui dépende surtout des  $[p^j, q^j]$ , mais aussi du comportement de l'importateur.

Il apparaît dans la table (3.5.2) que la valeur de  $\lambda$  en équilibre augmente quand le douanier est moins expérimenté. On peut se demander si cette tendance continue pour les douaniers de moins en moins expérimentés. Des valeurs supplémentaires se trouvent dans la table suivante :

Valeurs de $\lambda_{opt}$				
	$d = 5$	$d = 10$	$d = 20$	
$0,25 D + 0,75 N$	0,25	0,24	0,22	
	0,12	0,11	0,11	
	$d = 0$	$d = 10$	$d = 100$	$d = 25\ 000$
$N$ (novice)	0,27	0,26	0,22	0,12
	0,13	0,11	0,09	0,12

Table (3.6.1)

Notre modèle nous offre donc un conseil immédiat : dans *tous* les cas que nous avons envisagés, si une unique stratégie est imposée la valeur optimale de  $\lambda$  est inférieure à celle actuellement en vigueur (1/3). Les stratégies optimales qui correspondent aux "meilleures" estimations des paramètres douaniers ( $[p^j, q^j] = D, d = 10$ ) sont, suivant les deux règles (2.19.1 et 2) :

$$\lambda_{opt} = 5,2 \% \quad \text{et} \quad \lambda_{opt} = 44 \%$$

Les facteurs qui déterminent le choix de la meilleure stratégie sont, en ordre d'importance :

- 1/ La valeur de  $f_1$  et l'expérience du douanier :  $[p^j, q^j]$
- 2/ Le comportement de l'importateur
- 3/ L'expérience du douanier : valeur de  $d$ .

Nous n'avons pas étudié l'effet d'une augmentation du coût de contrôle,  $c$ , mais il est évident que  $\lambda_{opt}$  doit varier en sens inverse de  $c$  :

$$c \uparrow \rightarrow \lambda_{opt} \downarrow$$

Cependant il est peu probable que  $c$  augmenterait sans un accroissement correspondant des prix, des droits d'importation etc. Une étude détaillée de  $c$  exigerait plus de données que nous n'en possédons au stade actuel.

Résumant ces conclusions ainsi que celles de (3.3) nous pouvons répondre aux questions posées dans (1.5) :

1/ Le niveau de contrôle  $\lambda = 1/3$  est trop élevé. La douane pourrait gagner en adoptant les stratégies indiquées dans la table (3.5.2) dont la valeur centrale est de l'ordre  $\lambda = 0,05$ .

2/ Si la fréquence des erreurs varie selon le port (ou selon un autre classement convenable), une stratégie liée à  $f_1$  serait souhaitable.

3/ La faculté chez les douaniers d'estimer les  $[p^j, q^j]$  est importante.

### 3.7. – VALEUR DU JEU A LA DOUANE PAR RAPPORT A LA VALEUR ACTUELLE

Ayant trouvé toutes ces stratégies optimales on est enfin amené à se demander si les valeurs du jeu à la douane ( $V_{\text{opt}}$ ) sont sensiblement supérieures à celles qui correspondent à  $\lambda = 1/3$ . Celles-ci s'obtiennent facilement à partir de la simulation. Il suffit de noter les valeurs de  $V$  quand  $\lambda = 1/3$  pour différents  $f_1$  ( $[p^j, q^j]$  et  $d$  fixés). On détermine ainsi la valeur de  $f_1$  pour laquelle  $\lambda_1 f_1 = 0,0061$  et le  $V$  qui y correspond. Les valeurs de  $V_{\text{opt}}$  de la table (3.5.2) s'obtiennent de cette manière par interpolation à partir des données de l'appendice 4. Les résultats de ces calculs, pour les deux règles (2.19.1) et (2.19.2) sont résumés dans la table suivante :

	0,5 D + 0,5 I	D	0,5 D + 0,5 N
$V(\lambda = 1/3)$	- 0,151	- 0,155	- 0,185
$V_{\text{opt}}(\lambda_{\text{opt}})$ par (2.19.1)	- 0,032 (0,027)	- 0,067 (0,052)	- 0,18 (0,18)
$V_{\text{opt}}(\lambda_{\text{opt}})$ par (2.19.2)	- 0,03 (0,025)	- 0,057 (0,044)	- 0,12 (0,098)

$d = 10$  dans tous les cas

Table (3.7.1)

On voit que le gain relatif est considérable dans tous les cas sauf un (- 0,18 remplace - 0,185). Ce cas exceptionnel est l'un de ceux où le maximum est extrêmement aplati.  $V$  est donc presque invariable entre  $\lambda = 0,18$  et  $\lambda = 0,33$ .

La table indique que, si le modèle de cette étude est suffisamment réaliste, il y a de bonnes chances pour que la douane puisse réaliser d'importantes économies en adoptant une stratégie optimale.



## CHAPITRE 4

# MODÈLE THÉORIQUE SERVANT A ANALYSER LES RÉSULTATS DU QUESTIONNAIRE SUR L'ESTIMATION SUBJECTIVE DES VALEURS

### 4.1. – DISCUSSION DU PROBLEME

Dans ce chapitre nous traitons un problème un peu éloigné du thème central de l'étude. Le modèle mathématique a néanmoins été suggéré par le problème de (3.4), et il convient d'introduire le sujet dans ce contexte. D'ailleurs la méthode analytique sert à fournir une estimation du paramètre  $d$  utilisé dans le chapitre précédent.

Commençons par bien définir les termes que nous employons, car nous utilisons le mot "erreur" en un sens qui peut prêter à confusion. Nous désignons par "erreur" une disparité entre ce que déclare l'importateur (sur les documents) et la déclaration réglementaire correspondant aux biens de l'entry. Si une erreur a été décelée lors de l'importation nous possédons sa vraie valeur. Si les documents de l'entry concernée ont fait partie de notre (deuxième) échantillon expérimental nous possédons également l'estimation subjective de la valeur de l'erreur, à savoir :

- 1/ Moins de 5 livres.
- 2/ Entre 5 et 50 livres.
- 3/ Plus de 50 livres.

La situation est représentée dans la figure (4.1.1) page suivante.

Nous considérons quatre entries erronées. La vraie valeur de l'erreur est indiquée dans chaque cas par le point  $m_i (i = 1, 2, 3, 4)$ . L'estimation subjective de  $m_i$  consiste à spécifier l'intervalle dans lequel  $m_i$  est censé se trouver : moins de 5, entre 5 et 50, ou plus de 50. L'intervalle choisi est hachuré sur la figure. Désignant les trois intervalles respectivement par  $\omega_1$ ,  $\omega_2$  et  $\omega_3$ , on voit que pour l'entry 1, ( $m_i \simeq 15$ ),  $\omega_2$  est hachuré. C'est donc la meilleure estimation possible. Pour  $m_4 (\simeq 0)$ ,  $\omega_3$  est hachuré. C'est la plus mauvaise estimation possible.

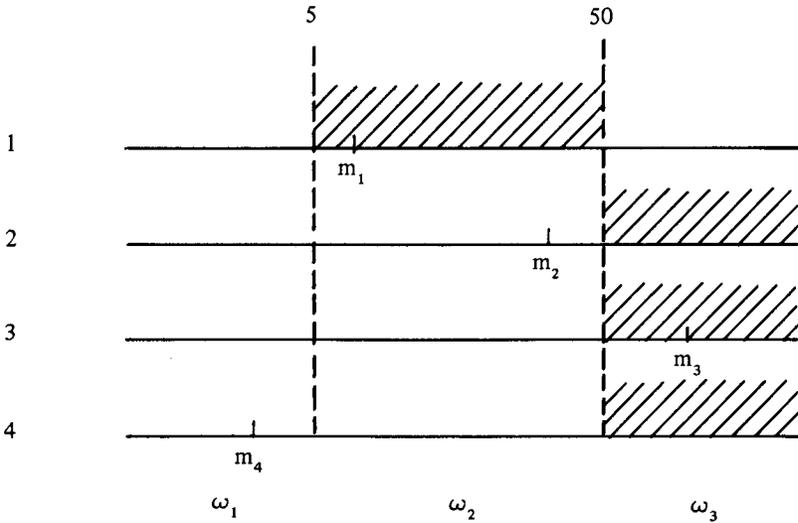


fig. (4.1.1)

Les  $m_i$ , qui sont les vraies valeurs d'erreurs, peuvent tomber théoriquement n'importe où. Leur fonction de répartition a été construite empiriquement à force d'avoir recueilli un grand nombre d'erreurs pendant une période étendue. (Voir appendice 2). Dans le cas présent nous considérons les  $m_i$  comme des valeurs fixes et connues. Leur répartition ne nous intéresse pas pour l'instant.

Dans cette étude nous voulons mesurer la précision avec laquelle le douanier peut deviner les valeurs d'erreur lorsqu'elles existent. Nous avons postulé un estimateur subjectif,  $\tilde{m}$ , fonction de  $m$ , la vraie valeur :

$$\tilde{m} = m + \epsilon .$$

Dans la simulation  $\epsilon$  a obéi à une répartition uniforme dans l'intervalle  $[-d, d]$ . En ce qui concerne l'analyse de l'expérience, cependant, une situation intéressante se produit si l'on suppose  $\epsilon$  être réparti normalement avec moyenne  $b$  et écart-type  $\sigma$ . Cette hypothèse est absolument arbitraire, et choisie pour des considérations purement mathématiques.

D'abord supposons que le douanier fait une estimation ponctuelle,  $\tilde{m}_i$ , de  $m_i$  (entry  $i$ ). Ensuite nous imposons la censure sur  $\tilde{m}_i$  : notre information est réduite à ne savoir que dans lequel de trois intervalles  $\omega_1$ ,  $\omega_2$  ou  $\omega_3$  se trouve  $\tilde{m}_i$ .

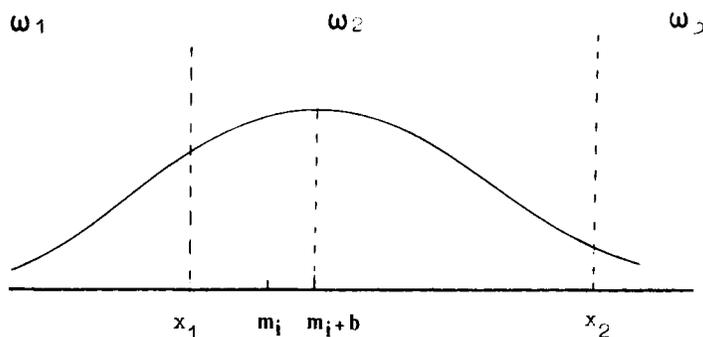


fig. (4.1.2)

$\tilde{m}_i$  est une variable normale  $(m_i + b, \sigma)$ .

Ecrivons

$$G(u) = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right) dx$$

Alors

$$P[\tilde{m}_i \in \omega_1] = P[\tilde{m}_i < x_1] = G\left[\frac{x_1 - m_i - b}{\sigma}\right]$$

$$P[\tilde{m}_i \in \omega_2] = G\left[\frac{x_2 - m_i - b}{\sigma}\right] - G\left[\frac{x_1 - m_i - b}{\sigma}\right]$$

et

$$P[\tilde{m}_i \in \omega_3] = 1 - G\left[\frac{x_2 - m_i - b}{\sigma}\right].$$

Dans le contexte de l'estimation subjective d'erreurs,  $b$  correspond à un biais systématique, et  $\sigma$  fournit une mesure générale de la précision. Dans ce qui suit ce problème sera analysé statistiquement sans faire allusion au problème douanier. Le nombre d'intervalles ne sera pas restreint à trois, et les estimateurs seront construits à partir d'un échantillon de taille  $n$ , ( $i = 1, 2, \dots, n$ ). Nous précisons les conditions dans lesquelles existent des solutions aux équations de maximum de vraisemblance, et il sera démontré que les estimateurs ainsi obtenus sont convergents et asymptotiquement efficaces. La matrice des covariances des estimateurs sera également obtenue.

## 4.2. — LA DISTRIBUTION NORMALE CENSUREE

Avant de procéder à l'analyse du cas présent, considérons un cas plus simple : l'estimation à partir d'observations censurées (groupées) tirées d'une population normale.

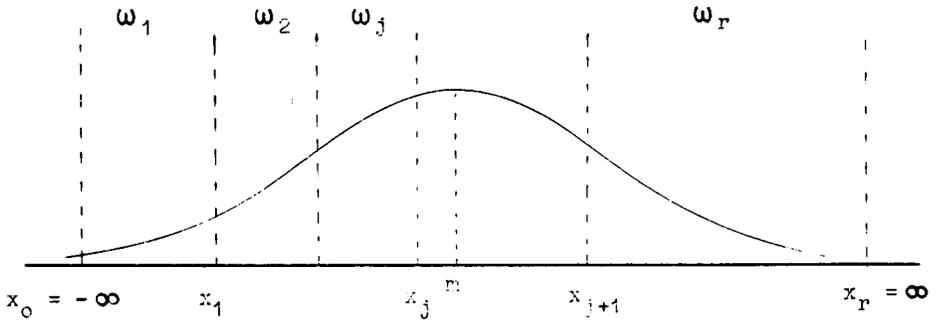


fig. (4.2.1)

La figure montre une courbe de densité normale ( $m, \sigma$ ). L'espace des observations est divisé en un nombre  $r$  d'intervalles disjoints.  $\omega_j$  est l'intervalle  $[x_{j-1}, x_j]$ . Si un échantillon de taille  $n$  est prélevé, l'information en résultant s'exprime par les nombres  $n_j (j = 1, 2, \dots, r)$  d'observations qui

tombent dans chaque intervalle,  $\omega_j \cdot \sum_{j=1}^r n_j = n$ .

Dans la figure (4.2.1),  $m$  est censé être fixe (connu ou inconnu). Dans le cas que nous voulons étudier,  $m$  prendra  $n$  valeurs différentes dans un échantillon de taille  $n$ ;  $m$  sera réparti selon une distribution quelconque. Le cas de la figure (4.2.1) est donc un cas particulier où, dans la distribution de  $m$ , tout le poids est concentré en un seul point.

Cette situation a fait l'objet de nombreuses études, et sans entrer dans les détails, il convient d'en noter trois résultats :

1/ Si  $r = 2$  il est impossible d'estimer  $m$  et  $\sigma$  à la fois, car l'on ne possède que la proportion  $\frac{n_1}{n_1 + n_2}$ .

2/ Si  $r > 3$  il n'y a pas de méthode immédiate pour résoudre les diverses équations de vraisemblance. En général on emploie une technique d'approximations successives (voir référence 1).

3/ Il a été démontré (réf. 2) que si l'un des paramètres,  $m$  et  $\sigma$ , est connu, alors la racine de l'équation de vraisemblance est, sous certaines conditions, unique et l'estimateur ainsi obtenu est asymptotiquement convergent et efficace.

### 4.3. - NOTATIONS

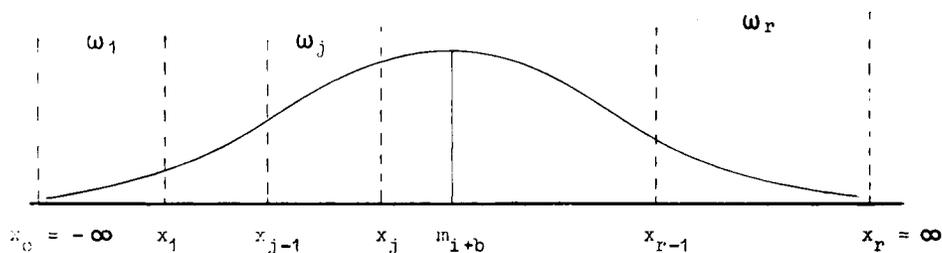


fig. (4.3.1)

La variable  $y_{ij}$  exprime le résultat du  $i$ -ième prélèvement :

$$y_{ij} = 0 \quad \text{si} \quad m_i \notin \omega_j$$

$$y_{ij} = 1 \quad \text{si} \quad m_i \in \omega_j$$

Ainsi

$$\sum_{j=1}^r y_{ij} = 1 \quad \text{et} \quad \sum_{i=1}^n \sum_{j=1}^r y_{ij} = n$$

Ecrivons aussi :

$$g(X) = \exp(-X^2/2)/\sqrt{2\pi} \quad \text{et} \quad G(X) = \int_{-\infty}^X g(t) dt$$

$$t_{ij} = (x_j - m_i - b)/\sigma$$

$$\begin{aligned} P_{ij} &= P[y_{ij} = 1] = G[x_j - m_i - b] - G[x_{j-1} - m_i - b] \\ &= G[t_{ij}] - G[t_{ij-1}] \end{aligned}$$

On doit aussi noter que :

$$\frac{\partial}{\partial b} G[t_{ij}] = -g(t_{ij})/\sigma \quad (4.3.1)$$

$$\frac{\partial}{\partial \sigma} G[t_{ij}] = -g(t_{ij}) t_{ij}/\sigma \quad (4.3.2)$$

$$\frac{\partial P_{ij}}{\partial b} = -\frac{1}{\sigma} [g(t_{ij}) - g(t_{ij-1})] \quad (4.3.3)$$

$$\frac{\partial^2 P_{ij}}{\partial b^2} = -\frac{1}{\sigma^2} [t_{ij} g(t_{ij}) - t_{ij-1} g(t_{ij-1})] \quad (4.3.4)$$

$$\frac{\partial P_{ij}}{\partial \sigma} = -\frac{1}{\sigma} [t_{ij} g(t_{ij}) - t_{ij-1} g(t_{ij-1})] \quad (4.3.5)$$

$$\frac{\partial^2 P_{ij}}{\partial \sigma^2} = -\frac{1}{\sigma^2} [g(t_{ij}) (t_{ij}^3 - 2t_{ij}) - g(t_{ij-1}) (t_{ij-1}^3 - 2t_{ij-1})] \quad (4.3.6)$$

#### 4.4. - L'EQUATION DE VRAISEMBLANCE

La vraisemblance de l'échantillon de taille  $n$  s'écrit :

$$L = \prod_{i=1}^n \prod_{j=1}^r [P_{ij}]^{y_{ij}}$$

$$\log L = \sum_{i=1}^n \sum_{j=1}^r y_{ij} \log P_{ij} \quad (4.4.1)$$

$$\frac{\partial \log L}{\partial \phi} = \sum_{i=1}^n \sum_{j=1}^r \frac{y_{ij}}{P_{ij}} \frac{\partial P_{ij}}{\partial \phi} \quad (\phi = b \text{ ou } \sigma) \quad (4.4.2)$$

$$\frac{\partial^2 \log L}{\partial \phi^2} = \sum_{i=1}^n \sum_{j=1}^r y_{ij} \left\{ \frac{-1}{P_{ij}^2} \left( \frac{\partial P_{ij}}{\partial \phi} \right)^2 + \frac{1}{P_{ij}} \frac{\partial^2 P_{ij}}{\partial \phi^2} \right\} \quad (4.4.3)$$

Substituant dans (4.4.2)

$$\frac{\partial \log L}{\partial b} = \sum_{i=1}^n \sum_{j=1}^r \frac{-y_{ij} [g(t_{ij}) - g(t_{ij-1})]}{\sigma [G(t_{ij}) - G(t_{ij-1})]} \quad (4.4.4)$$

$$\frac{\partial \log L}{\partial \sigma} = \sum_{i=1}^n \sum_{j=1}^r \frac{-y_{ij} [t_{ij} g(t_{ij}) - t_{ij-1} g(t_{ij-1})]}{\sigma [G(t_{ij}) - G(t_{ij-1})]} \quad (4.4.5)$$

et substituant dans (4.4.3) :

$$\frac{\partial^2 \log L}{\partial b^2} = \sum_{i=1}^n \sum_{j=1}^r \frac{-y_{ij}}{\sigma^2} \left\{ \left[ \frac{g(t_{ij}) - g(t_{ij-1})}{G(t_{ij}) - G(t_{ij-1})} \right]^2 + \left[ \frac{t_{ij} g(t_{ij}) - t_{ij-1} g(t_{ij-1})}{G(t_{ij}) - G(t_{ij-1})} \right] \right\} \quad (4.4.6)$$

$$\frac{\partial^2 \log L}{\partial \sigma^2} = \sum_{i=1}^n \sum_{j=1}^r \frac{y_{ij}}{\sigma^2} \left\{ \frac{-[t_{ij} g(t_{ij}) - t_{ij-1} g(t_{ij-1})]^2}{[G(t_{ij}) - G(t_{ij-1})]^2} + \frac{[g(t_{ij})(2t_{ij} - t_{ij}^3) - g(t_{ij-1})(2t_{ij-1} - t_{ij-1}^3)]}{[G(t_{ij}) - G(t_{ij-1})]} \right\} \quad (4.4.7)$$

Pour trouver la matrice des covariances il nous faut l'espérance des  $\frac{\partial^2 \log L}{\partial \phi^2}$  (4.4.3). Mais l'espérance de  $y_{ij}$  n'est autre chose que  $P_{ij}$ , donc :

$$E \left\{ \sum_{i=1}^n \sum_{j=1}^r \frac{y_{ij}}{P_{ij}} \frac{\partial^2 P_{ij}}{\partial \phi^2} \right\} = E \left\{ \frac{\partial^2}{\partial \phi^2} \sum_{i=1}^n \sum_{j=1}^r P_{ij} \right\} = 0$$

ce qui donne :

$$- E \left[ \frac{\partial^2 \log L}{\partial b^2} \right] = \sum_{j=1}^r \sum_{i=1}^n \frac{[g(t_{ij}) - g(t_{ij-1})]^2}{\sigma^2 [G(t_{ij}) - G(t_{ij-1})]} \quad (4.4.8)$$

$$- E \left[ \frac{\partial^2 \log L}{\partial b \partial \sigma} \right] = \sum_{i=1}^n \sum_{j=1}^r \frac{[g(t_{ij}) - g(t_{ij-1})] [t_{ij} g(t_{ij}) - t_{ij-1} g(t_{ij-1})]}{\sigma^2 [G(t_{ij}) - G(t_{ij-1})]} \quad (4.4.9)$$

$$- E \left[ \frac{\partial^2 \log L}{\partial \sigma^2} \right] = \sum_{i=1}^n \sum_{j=1}^r \frac{[t_{ij} g(t_{ij}) - t_{ij-1} g(t_{ij-1})]^2}{\sigma^2 [G(t_{ij}) - G(t_{ij-1})]} \quad (4.4.10)$$

4.5. – ESTIMATION DE  $b$  QUAND  $\sigma$  EST CONNU

Nous allons préciser les conditions nécessaires et suffisantes pour l'existence et l'unicité d'une racine de (4.4.4). Nous suivons les lignes indiquées par Kulldorf [2] dans l'analyse du cas classique de la distribution normale censurée. Nous considérons le comportement de (4.4.4) quand  $b$  tend vers les deux limites dans l'espace du paramètre :  $-\infty < b < \infty$ . Il faut tenir compte de trois cas :

$$(i) y_{i1} = 1$$

$$(ii) y_{ij} = 1, 1 < j < r$$

$$(iii) y_{ir} = 1$$

Par le théorème de la moyenne on trouve que :

$$\frac{g(t_{ij}) - g(t_{ij-1})}{G(t_{ij}) - G(t_{ij-1})} = -T_{ij}, t_{ij-1} < T_{ij} < t_{ij}$$

Ecrivant

$$Q_{ij} = \frac{y_{ij} [g(t_{ij}) - g(t_{ij-1})]}{\sigma [G(t_{ij}) - G(t_{ij-1})]}$$

on trouve, quand  $b \rightarrow -\infty$  :

$$\text{cas (i)} \quad \lim_{b \rightarrow -\infty} Q_{i1} = \frac{[g(\infty) - g(-\infty)]}{[G(\infty) - G(-\infty)]} = 0$$

$$\text{cas (ii)} \quad \lim_{b \rightarrow -\infty} Q_{ij} = -\lim_{b \rightarrow -\infty} T_{ij} = -\infty$$

$$\text{cas (iii)} \quad \lim_{b \rightarrow -\infty} Q_{ir} = -\infty$$

et quand  $b \rightarrow \infty$  :

$$\text{cas (i)} \quad \lim_{b \rightarrow \infty} Q_{i1} = -\lim_{b \rightarrow \infty} T_{i1} = \infty$$

$$\text{cas (ii)} \quad \lim_{b \rightarrow \infty} Q_{ij} = \infty$$

$$\text{cas (iii)} \quad \lim_{b \rightarrow \infty} Q_{ir} = \frac{[g(\infty) - g(-\infty)]}{[G(\infty) - G(-\infty)]} = 0$$

A condition que toutes les observations ne tombent pas dans un des intervalles extrêmes, c'est-à-dire à condition que :

$$\sum_{i=1}^n y_{i1} < n \quad \text{et} \quad \sum_{i=1}^n y_{ir} < n$$

il existe au moins une racine de (4.4.4), fonction continue de  $b$ .

Cette racine sera unique si  $\frac{\partial^2 \log L}{\partial b^2} < 0$  pour tout  $b$ , ce qui sera vrai si la quantité entre accolades dans (4.4.6) est positive. Or, la fonction,  $tg(t)$ , s'esquisse comme suit :

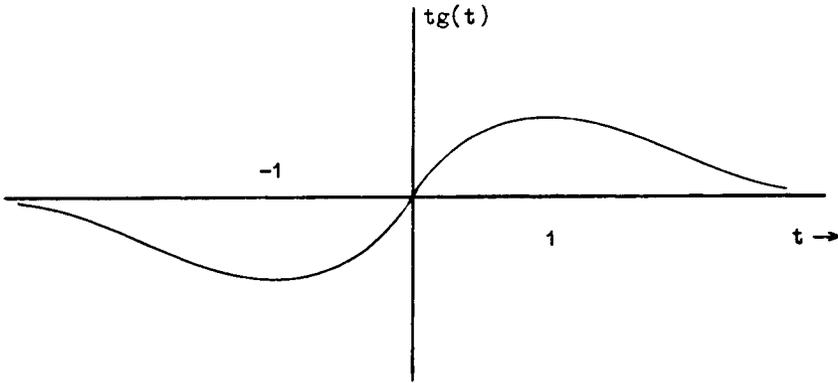


fig. (4.5.1)

$t_{ij} g(t_{ij}) - t_{ij-1} g(t_{ij-1})$  ne peut être négatif que si  $t_{ij} > 1$  et  $t_{ij-1} > 0$  (ou dans le cas symétrique quand  $t_{ij-1} < -1$  et  $t_{ij} < 0$ )

Nous verrons que même dans ces cas (il suffit d'en traiter le premier)  $\frac{\partial^2 \log L}{\partial b^2} < 0$ .

Ecrivons  $\phi(t) = \frac{g(t)}{t} + G(t)$ , fonction décroissante pour  $t > 0$ ,

car  $\frac{d\phi}{dt} = -\frac{g(t)}{t^2}$ , ce qui donne :

$$G(t_{ij}) - G(t_{ij-1}) = -\frac{g(t_{ij})}{t_{ij}} + \frac{g(t_{ij-1})}{t_{ij-1}} + \phi(t_{ij}) - \phi(t_{ij-1})$$

$$G(t_{ij}) - G(t_{ij-1}) > -\frac{g(t_{ij})}{t_{ij}} + \frac{g(t_{ij-1})}{t_{ij-1}} \quad (4.5.1)$$

Et la quantité entre accolades dans (4.4.6) est proportionnelle à :

$$\begin{aligned}
& [t_{ij} g(t_{ij}) - t_{ij-1} g(t_{ij-1})] [G(t_{ij}) - G(t_{ij-1})] + [g(t_{ij}) - g(t_{ij-1})]^2 \\
> & [-g(t_{ij})^2 - g(t_{ij-1})^2] + \left[ \frac{t_{ij-1}}{t_{ij}} + \frac{t_{ij}}{t_{ij-1}} \right]^2 g(t_{ij}) g(t_{ij-1}) \\
& \qquad \qquad \qquad + [g(t_{ij}) - g(t_{ij-1})]^2 \\
> & g(t_{ij}) g(t_{ij-1}) \left[ \frac{-2t_{ij} t_{ij-1} + t_{ij-1}^2 + t_{ij}^2}{t_{ij} t_{ij-1}} \right] > 0
\end{aligned}$$

Donc si une racine de (4.4.4) existe elle est unique.

#### 4.6. – ESTIMATION DE $\sigma$ QUAND $b$ EST CONNU

Nous considérons le comportement de (4.4.5) quand  $\sigma \rightarrow 0$  et puis quand  $\sigma \rightarrow \infty$ , pour établir l'existence d'une racine.

Puisque  $y_{ij} = 1$  pour une seule valeur de  $j$  ( $j = 1, 2, \dots, r$ ) nous pouvons traiter les deux cas :

$$y_{ij} = 1 \quad \text{quand} \quad 1 < j < r$$

et

$$y_{ij} = 1 \quad \text{quand} \quad j = 1 \quad \text{ou} \quad j = r.$$

Dans chaque cas il faut différencier encore selon que  $(m_i + b)$  tombe dans un intervalle central ou extrême. Tous les cas possibles se ramènent à un des cinq suivants :  $t_{ij} > t_{ij-1}$  dans tous les cas :

- |                                 |   |                         |
|---------------------------------|---|-------------------------|
| (i) $m_i + b < x_{j-1}$         | ⟶ | $t_{ij-1} > 0$          |
| (ii) $m_i + b = x_{j-1}$        | ⟶ | $t_{ij-1} = 0$          |
| (iii) $x_{j-1} < m_i + b < x_j$ | ⟶ | $t_{ij-1} < 0 < t_{ij}$ |
| (iv) $m_i + b = x_{r-1}$        | ⟶ | $t_{r-1} = 0$           |
| (v) $m_i + b > x_{r-1}$         | ⟶ | $t_{r-1} < 0$           |

Il est utile encore une fois de se servir du théorème de la moyenne :

$$\frac{t_{ij} g(t_{ij}) - t_{ij-1} g(t_{ij-1})}{G(t_{ij}) - G(t_{ij-1})} = 1 - T_{ij}^2 \quad \text{où} \quad t_{ij-1} < T_{ij} < t_{ij}$$

Posons :

$$R_{ij} = \frac{y_{ij} [t_{ij} g(t_{ij}) - t_{ij-1} g(t_{ij-1})]}{\sigma [G(t_{ij}) - G(t_{ij-1})]}$$

$\sigma \rightarrow 0$

cas (i)  $y_{ij} = 1 : \lim_{\sigma \rightarrow 0} R_{ij} = -\infty$

$y_{ir} = 1 : \lim_{\sigma \rightarrow 0} R_{ir} = -\infty$

cas (ii)  $y_{ij} = 1 : \lim_{\sigma \rightarrow 0} R_{ij} = \lim_{\sigma \rightarrow 0} \frac{t_{ij} g(t_{ij})}{\sigma \left[ G(t_{ij}) - \frac{1}{2} \right]} = 0$

$y_{ir} = 1 : \lim_{\sigma \rightarrow 0} R_{ir} = -\infty$

cas (iii)  $y_{ij} = 1 : \lim_{\sigma \rightarrow 0} R_{ij} = 0$

$y_{ir} = 1 : \lim_{\sigma \rightarrow 0} R_{ir} = -\infty$

cas (iv)  $y_{ij} = 1 : \lim_{\sigma \rightarrow 0} R_{ij} = -\infty$

$y_{ir} = 1 : \lim_{\sigma \rightarrow 0} R_{ir} = 0$

cas (v)  $y_{ij} = 1 : \lim_{\sigma \rightarrow 0} R_{ij} = -\infty$

$y_{ir} = 1 : \lim_{\sigma \rightarrow 0} R_{ir} = 0$

On en conclut qu'à condition qu'il y ait au moins une observation qui ne tombe pas dans l'intervalle qui recouvre la moyenne, c'est-à-dire que la condition

$$(y_{ij} = 0) \cap (x_{j-1} < m_i + b < x_j)$$

soit au moins une fois remplie, alors :

$$\lim_{\sigma \rightarrow 0} \frac{\partial \log L}{\partial \sigma} = - \sum_{j=1}^n \sum_{i=1}^r \lim_{\sigma \rightarrow 0} R_{ij} = +\infty \quad (4.6.1)$$

$\sigma \rightarrow \infty$

$$\text{cas (i)} \quad y_{ij} = 1 : \lim_{\sigma \rightarrow \infty} R_{ij} = \lim_{\sigma \rightarrow \infty} \frac{(1 - T_{ij}^2)}{\sigma} = + 0$$

$$y_{ir} = 1 : \lim_{\sigma \rightarrow \infty} R_{ir} = \lim_{\sigma \rightarrow \infty} \frac{[0 - t_{ir-1} g(t_{ir-1})]}{\sigma [1 - G(t_{ir-1})]} = 0$$

$$\text{cas (iii)} \quad y_{ij} = 1 : \lim_{\sigma \rightarrow \infty} R_{ij} = + 0$$

$$y_{ir} = 1 : \lim_{\sigma \rightarrow \infty} R_{ir} = \frac{(0 - 0)}{(1 - 0,5) \sigma} = 0$$

cas (iii), (iv) et (v) :

$$y_{ij} = 1 : \lim_{\sigma \rightarrow \infty} R_{ij} = \lim_{\sigma \rightarrow \infty} \frac{(1 - T_{ij}^2)}{\sigma} = + 0$$

$$y_{ir} = 1 : \lim_{\sigma \rightarrow \infty} R_{ir} = 0$$

on voit que si toutes les observations ne tombent pas dans les deux intervalles extrêmes, c'est-à-dire si

$$\sum_{j=1}^n (y_{i1} + y_{ir}) < n$$

alors

$$\lim_{\sigma \rightarrow \infty} \frac{\partial \log L}{\partial \sigma} = \sum_{i=1}^n \sum_{j=1}^r -\lim_{\sigma \rightarrow \infty} P_{ij} = - 0 \quad (4.6.2)$$

Donc une racine de (4.4.5) existe à condition que :

1/ Au moins une observation ne tombe pas dans l'intervalle qui recouvre la moyenne.

2/ Toutes les observations ne tombent pas dans les deux intervalles extrêmes.

Pour établir l'unicité de la racine il suffira de démontrer que  $\left(\frac{\partial \log L}{\partial \sigma}\right)_{\sigma=\hat{\sigma}}$  est une fonction décroissante de  $\sigma$ .

Si l'on substitue  $\frac{\partial \log L}{\partial \sigma} = 0$  de (4.4.5) dans (4.4.7) on trouve :

$$\frac{\partial^2 \log L}{\partial \sigma^2} = \sum_{i=1}^n \sum_{j=1}^r \frac{-y_{ij}}{\sigma^2} \left\{ \frac{[t_{ij} g(t_{ij}) - t_{ij-1} g(t_{ij-1})]^2}{[G(t_{ij}) - G(t_{ij-1})]^2} + \frac{t_{ij}^3 g(t_{ij}) - t_{ij-1}^3 g(t_{ij-1})}{G(t_{ij}) - G(t_{ij-1})} \right\} \quad (4.6.3)$$

Il s'agit de démontrer que la quantité entre accolades est toujours positive. La fonction  $t^3 g(t)$  s'esquisse comme suit :

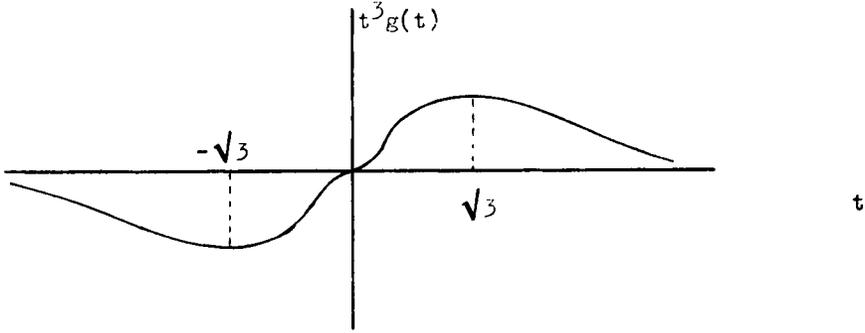


fig. (4.6.1)

La pente est égale à zéro quand  $t = 0, \pm\sqrt{3}, \pm\infty$ . Il s'ensuit que  $t_{ij}^3 g(t_{ij}) - t_{ij-1}^3 g(t_{ij-1})$  ne peut être négatif que si  $t_{ij} > \sqrt{3}$  et  $t_{ij-1} > 0$  (ou dans le cas symétrique quand  $t_{ij-1} < -\sqrt{3}$  et  $t_{ij} < 0$ ). Il suffit donc de considérer le premier de ces deux cas symétriques.

Utilisant (4.5.1) on trouve :

$$\begin{aligned} & [t_{ij} g(t_{ij}) - t_{ij-1} g(t_{ij-1})]^2 + [t_{ij}^3 g(t_{ij}) - t_{ij-1}^3 g(t_{ij-1})] \times [G(t_{ij}) - G(t_{ij-1})] \\ & > [t_{ij} g(t_{ij}) - t_{ij-1} g(t_{ij-1})]^2 + [t_{ij}^3 g(t_{ij}) - t_{ij-1}^3 g(t_{ij-1})] \\ & \quad \times \left[ \frac{-g(t_{ij})}{t_{ij}} + \frac{g(t_{ij-1})}{t_{ij-1}} \right] \\ & > -2 t_{ij} t_{ij-1} g(t_{ij}) g(t_{ij-1}) + \left[ \frac{t_{ij}^3}{t_{ij-1}} + \frac{t_{ij-1}^3}{t_{ij}} \right] g(t_{ij}) g(t_{ij-1}) \\ & > \frac{[t_{ij}^2 - t_{ij-1}^2]^2}{t_{ij} t_{ij-1}} g(t_{ij}) g(t_{ij-1}) > 0 \end{aligned}$$

$\frac{\partial^2 \log L}{\partial \sigma^2}$  est donc une fonction décroissante pour  $\sigma$  qui satisfait à  $\frac{\partial \log L}{\partial \sigma} = 0$ .

Comme  $\frac{\partial \log L}{\partial \sigma}$  est une fonction continue de  $\sigma$  ( $0 < \sigma < \infty$ ) il ne peut exister qu'une racine de (4.4.5)

#### 4.7. – PROPRIETES ASYMPTOTIQUES DES ESTIMATEURS

Il faut remarquer que les estimateurs fournis par les solutions des équations de maximum de vraisemblance n'existent que si les conditions énoncées ci-dessus sont remplies. Dans le cas (4.2) de la répartition normale censurée (ou groupée) la probabilité pour qu'elles ne soient pas remplies tend vers zéro quand  $n \rightarrow \infty$ . Dans le cas plus général nous n'avons pas tenu compte de la répartition des  $m_i$ . Il est évident que si cette répartition correspond à presque n'importe quelle densité habituellement rencontrée la même conclusion peut en être tirée. Cependant, pour éviter cette difficulté théorique nous pouvons choisir de considérer les propriétés de la racine éventuelle — les conditions des sections (4.5) et (4.6) supposées satisfaites.

Nous rappelant la répartition normale censurée (4.2), Kulldorf [8] a donné un ensemble de conditions suffisantes pour que les estimateurs soient asymptotiquement convergents et efficaces. Nous nous contenterons ici d'énumérer ces conditions, adaptées au cas plus général de (4.3), et d'indiquer que la plupart d'entre elles sont immédiatement satisfaites.

#### CONDITIONS SUFFISANTES (DANS L'ESTIMATION DE $\sigma$ AVEC $b$ CONNU)

Les  $P_{ij}$  sont définis dans (4.3)

- (i)  $\frac{\partial \log P_{ij}}{\partial \sigma}$  existe pour tout  $i, j$  et  $0 < \sigma < \infty$
- (ii)  $\frac{\partial^2 \log P_{ij}}{\partial \sigma^2}$  existe pour tout  $i, j$  et  $0 < \sigma < \infty$
- (iii)  $\frac{\partial^3 \log P_{ij}}{\partial \sigma^3}$  existe pour tout  $i, j$  et  $0 < \sigma < \infty$

$$(iv) \sum_{j=1}^r \frac{\partial P_{ij}}{\partial \sigma} = 0 \quad \text{pour tout } i \text{ et } 0 < \sigma < \infty$$

$$(v) \sum_{j=1}^r \frac{\partial^2 P_{ij}}{\partial \sigma^2} = 0 \quad \text{pour tout } i \text{ et } 0 < \sigma < \infty$$

$$(vi) -\infty < \sum_{i=1}^r P_{ij} \frac{\partial^2 \log P_{ij}}{\partial \sigma^2} < 0 \quad \text{pour tout } i \text{ et } 0 < \sigma < \infty$$

(vii) Il existe une fonction,  $a(\sigma)$ , positive et deux fois dérivable pour  $0 < \sigma < \infty$ , et une suite de nombres  $H_{ij}$  tels que, pour  $0 < \sigma < \infty$  :

$$\left| \frac{\partial^2}{\partial \sigma^2} \left( a(\sigma) \frac{\partial \log P_{ij}}{\partial \sigma} \right) \right| < H_{ij} \quad \text{pour tout } i, j.$$

Quand on précise "pour tout  $i$ " on entend "quelle que soit la valeur (finie) de  $m_i$ ", car  $m_i$  entre dans la densité discrète  $P_{ij}$ . Dans le cas simple ou  $m$  est fixe on supprime l'indice  $i$ , mais les arguments sont valables pour tout  $m$  (fini). On voit tout de suite que les cinq premières conditions sont satisfaites. (vi) se démontre facilement, et  $a(\sigma) = \sigma^3$  convient pour (vii).

Les conditions équivalentes pour l'estimation de  $m$  dans le cas censuré de (4.2) ont été données par Kulldorf [2]. De la même manière on en arrive à les adapter au cas général.



## CHAPITRE 5

# CONSIDÉRATIONS DE PLAN D'EXPÉRIENCE

### 5.1. – EVALUATION NUMERIQUE DES EQUATIONS DE VRAISEMBLANCE

Dans la simulation du chapitre 3 nous avons utilisé une valeur de l'ordre 10 pour le paramètre  $d$ . Une estimation du paramètre  $\sigma$  peut s'obtenir à partir de l'équation (4.4.5), avec  $b = 0$ , et en utilisant les données de l'appendice 3, table 3. Une inspection de ces données montre que les conditions pour l'existence d'une racine sont satisfaites. La solution ainsi obtenue (par itération) est :

$$\hat{\sigma} = 6,6$$

Les équations (4.4.8–10) fournissent la matrice des covariances :

$$V = \begin{bmatrix} 1,3740 & 0,1836 \\ 0,1836 & 2,6230 \end{bmatrix} \quad \begin{array}{l} b = 0 \\ \sigma = 6,6 \end{array}$$

Il est évident que les estimateurs ne convergent pas très rapidement. Mais cette convergence dépend, entre autre, des positions des points de troncature, et l'on peut se demander s'il est possible de les placer d'une manière optimale dans un schéma tel que celui de (4.3).

### 5.2. – POINTS DE TRONCATURE OPTIMAUX

Le critère d'optimalisation est la variance asymptotique de l'estimateur (MV). Il y aura donc certainement deux solutions, correspondant à  $\sigma$  et à  $b$ .

L'étude de ce problème dans le cas simple (4.2) est assez bien documentée, en particulier lorsqu'il n'y a que deux points de troncature, [9]. Plusieurs tables ont été publiées, destinées à aider à la solution de certains cas particuliers, [1, 9].

Il est intéressant de se demander si l'on peut poursuivre une analyse comparable de la situation générale de (4.3). Evidemment le choix des points de troncature dépendra de la répartition des  $m_i$ .

Considérons d'abord le cas où l'on veut estimer  $\sigma$  quand  $b$  est connu. On sait que la variance asymptotique de  $\hat{\sigma}$  s'écrit :

$$V[\hat{\sigma}] = \frac{-1}{E \left[ \frac{\partial^2 \log L}{\partial \sigma^2} \right]}$$

Reprenant (4.4.10), rendre minimal  $V[\sigma]$ , c'est rendre maximal :

$$\frac{1}{V[\hat{\sigma}]} = \sum_{i=1}^n \sum_{j=1}^r \left[ \frac{[t_{ij} g(t_{ij}) - t_{ij-1} g(t_{ij-1})]^2}{\sigma^2 [G(t_{ij}) - G(t_{ij-1})]} \right] \quad (5.2.1)$$

L'expression équivalente dans l'estimation de  $b$ , quand  $\sigma$  est connu, est :

$$\frac{1}{V[\hat{b}]} = \sum_{i=1}^n \sum_{j=1}^r \frac{[g(t_{ij}) - g(t_{ij-1})]^2}{\sigma^2 [G(t_{ij}) - G(t_{ij-1})]} \quad (5.2.2)$$

Rappelons-nous que les  $t_{ij}$  sont fonctions des  $m_i$  qui sont considérés comme des valeurs fixes et connues. Mais du point de vue expérimental on ne sait pas à l'avance les valeurs exactes des  $m_i$  ; on ne sait que leur répartition. Nous supposons que cette répartition est continue. Plus exactement donc, la fonction à rendre maximale s'écrit :

$$\frac{1}{V[\hat{\sigma}]} = E \left\{ \sum_{i=1}^n \sum_{j=1}^r \frac{[t_{ij} g(t_{ij}) - t_{ij-1} g(t_{ij-1})]^2}{\sigma^2 [G(t_{ij}) - G(t_{ij-1})]} \right\}$$

qui s'applique à une expérience à  $n$  observations et où les  $m_i$  ( $i = 1, 2, \dots, n$ ) sont répartis selon la densité  $f(m)$ . La fonction à rendre optimale s'écrit, indépendamment de  $n$  :

$$I[\hat{\sigma}] = \int_{-\infty}^{\infty} \sum_{j=1}^r \frac{[t_j g(t_j) - t_{j-1} g(t_{j-1})]^2}{\sigma^2 [G(t_j) - G(t_{j-1})]} f(m) dm \quad (5.2.3)$$

où 
$$t_j = \frac{x_j - m - b}{\sigma} \quad \text{et :}$$

$$I[\hat{b}] = \int_{-\infty}^{\infty} \sum_{j=1}^r \frac{[g(t_j) - g(t_{j-1})]^2}{\sigma^2 [G(t_j) - G(t_{j-1})]} f(m) dm \quad (5.2.4)$$

Considérons maintenant le cas simple de (4.2). Les valeurs de  $m$  ne sont plus réparties, et par conséquent  $f(m)$  disparaît. On obtient alors :

$$\frac{1}{V[\hat{\sigma}]} = \frac{n}{\sigma^2} \sum_{j=1}^r \frac{[t_j g(t_j) - t_{j-1} g(t_{j-1})]^2}{[G(t_j) - G(t_{j-1})]} \quad (5.2.5)$$

où  $t_j = \frac{x_j - m}{\sigma}$  et l'échantillon est de taille  $n$ .

Si l'on compare  $V[\hat{\sigma}]$  avec la variance de  $\hat{\sigma}$  qui résulte d'un échantillon non-censuré on obtient une expression que nous pouvons appeler, "l'efficacité absolue de  $\hat{\sigma}$ , dans le cas simple, pour un échantillon de taille  $n$ ". Dans une notation abrégée :

$$\begin{aligned} Ef. \sigma. S. (n) &= \frac{\text{Variance de } \hat{\sigma} \text{ (non-censurée)}}{V[\hat{\sigma}]} \\ &= \frac{\sigma^2}{2n} \cdot \frac{n}{\sigma^2} \sum_{j=1}^r \left\{ \frac{[t_j g(t_j) - t_{j-1} g(t_{j-1})]^2}{[G(t_j) - G(t_{j-1})]} \right\} \\ &= \frac{1}{2} \sum_{j=1}^r \frac{[t_j g(t_j) - t_{j-1} g(t_{j-1})]^2}{[G(t_j) - G(t_{j-1})]} \end{aligned}$$

Supposons maintenant que les  $n$  observations dans le cas général consistent en  $n_1$  observations avec  $m = m_1$ ,  $n_2$  avec  $m = m_2$ , ...  $n_s$  avec  $m = m_s$  :

$$\sum_{i=1}^s n_i = n$$

on obtient alors :

$$\frac{1}{V[\hat{\sigma}]} = \sum_{i=1}^s \frac{n_i}{\sigma^2} \sum_{j=1}^r \left\{ \frac{[t_{ij} g(t_{ij}) - t_{ij-1} g(t_{ij-1})]^2}{[G(t_{ij}) - G(t_{ij-1})]} \right\}$$

Enfin appelons  $Eff. \hat{\sigma}. G. (n)$  l'efficacité absolue de  $\sigma$ , dans le cas général, pour un échantillon de taille  $n$ .

$$\begin{aligned}
 Ef. \hat{\sigma} . G . (n) &= \frac{\sigma^2}{2n} \cdot \frac{1}{V[\hat{\sigma}]} \\
 &= \frac{1}{n} \sum_{i=1}^s \frac{n_i}{2} \sum_{j=1}^r \left\{ \frac{[t_{ij} g(t_{ij}) - t_{ij-1} g(t_{ij-1})]^2}{[G(t_{ij}) - G(t_{ij-1})]} \right\} \\
 &= \frac{1}{n} \sum_{i=1}^s \{n_i [Ef. \hat{\sigma} . S . (n_i)]\}
 \end{aligned}$$

C'est-à-dire que l'efficacité absolue dans le cas général consiste en une combinaison linéaire des efficacités des cas simples constituants. Stevens [9] donne des tables d'efficacité de  $\hat{m}$  et  $\hat{\sigma}$  pour le cas où  $r = 3$  (deux points de troncature). En général, cependant, ce genre de travail devra se faire sur ordinateur.

## APPENDICE 1

### QUELQUES STATISTIQUES SUR LA FRÉQUENCE DE LA FRAUDE ET DES ERREURS

Si la douane relève une entry qui est incorrecte elle peut avoir recours à l'une de trois procédures :

1/ L'entry est inacceptable et doit être à nouveau soumise à la douane. (Normalement les erreurs dans la déclaration auront été mises en évidence et corrigées par la douane).

2/ Jusqu'à une certaine limite la douane peut frapper l'importateur d'une amende.

3/ La douane peut mettre en marche des poursuites juridiques menant à des amendes et éventuellement à des peines infamantes en cas de fraude ou contrebande flagrante.

Un recensement de statistiques douanières dans tous les ports du pays pendant un mois a fourni les chiffres suivants :

Nombre d'entries contrôlées :	82 000
dont non-réglementaires :	1 500
somme ainsi récupérée :	£ 27 000
Valeur totale des amendes imposées par la procédure 2 :	£ 500
Aucune poursuite selon la procédure 3.	

## APPENDICE 2

### QUELQUES STATISTIQUES SUR LA RÉPARTITION DES ERREURS

Un recensement des erreurs révélées a fourni les chiffres suivants :

valeur (livres)	0-1	1-2	2-3	3-4	4-5	5-10	10-25	25-50	> 50
fréquence	158	139	91	63	46	141	160	87	90

Total 975.

A partir de ces chiffres a été construite une fonction de densité continue  $f(m)$ , telle que :

- pour  $m < 0$ ,  $f(m) = 0$  ;
- pour  $0 \leq m \leq 250$ ,  $f(m) > 0$  ;
- pour  $250 < m$ ,  $f(m) = 0$ .

Le choix de 250 comme limite supérieure est arbitraire. Dans l'intervalle  $0 \leq m \leq 250$ , la manière d'obtenir la densité consiste à pondérer les fréquences inversement à la longueur des intervalles de l'histogramme, puis à relier les neuf points ainsi générés par des droites (en extrapolant jusqu'à  $m = 0$ ), et enfin à normaliser de sorte que l'aire totale sous la courbe soit égale à l'unité.

La densité ainsi construite est définie par les coordonnées suivantes (reliés par des droites) :

$m$	0	1,5	2,5	3,5	4,5	7,5	17,5	37,5	100,0	250,0
$f(m)$	0,1556	0,1291	0,0845	0,0585	0,0427	0,0262	0,0099	0,00323	0,00038	0

La moyenne de cette densité est égale à 18,01.

Le coût moyen à la douane du contrôle d'une entry ( $. = c$ ) a été calculé à partir des salaires, etc., et s'élève à :

$$c = 0,45 \text{ (livre)}$$

## APPENDICE 3

### RÉSULTATS EXPÉRIMENTAUX

On se réfère au questionnaire de (3.4). L'échantillon d'entries est trié selon deux critères :

1/ La probabilité d'erreur.  $j = 1, 2$  ou  $3$ . La catégorie  $3$  est celle qui est la plus soupçonnée d'erreur.

2/ La valeur d'une erreur éventuelle. Les trois intervalles disjoints sont :  $\leq 5$ , entre  $5$  et  $50$ ,  $> 50$  (valeurs en livres sterling).

"c" indique que l'entry a été contrôlée ; "c̄" en indique le contraire.

Table I  
Valeur éventuelle

$j$	$\leq 5$		$]5, 50]$		$> 50$		Total	
	c	c̄	c	c̄	c	c̄	c	c̄
3	3	0	5	0	6 <sup>°</sup>	0	14	0
2	32	88	64 <sup>+</sup>	59	28	6	124	153
1	22	176	1	5	2	3	25	184
	57	264	70	64	36	9	163	337

(+) = 1 erreur retrouvée

(°) = 2 erreurs retrouvées.

Table 2  
Répartition de 71 entries (erronnées) supplémentaires

$j$	$\leq 5$	$]5, 50]$	$> 50$
3	21	25	4
2	11	10	1
1	1	1	0

Table 3 – Classement des entries erronnées par valeur de  $m$

Les soixante-quatorze entries erronnées des deux tables précédentes permettent de comparer la vraie valeur,  $m_i$ , de l'erreur et l'intervalle,  $\omega_{ij}$ , dans lequel le douanier l'a placée. Les  $n_i$  sont des fréquences, car certaines valeurs des  $m_i$  (à  $\pounds 1$  près) se répètent.

$n_i$	$m_i$	$\omega_{i1}$	$\omega_{i2}$	$\omega_{i3}$
10	1	1		
2	1		1	
5	2	1		
1	2		1	
6	3	1		
2	3		1	
6	4	1		
1	5	1		
1	5		1	
4	6		1	
1	7		1	
2	8		1	
3	9		1	
1	11	1		
2	11		1	
1	12		1	
3	13		1	
1	14		1	

$n_i$	$m_i$	$\omega_{i1}$	$\omega_{i2}$	$\omega_{i3}$
2	15		1	
1	16	1		
3	16		1	
1	17	1		
1	17		1	
2	18		1	
2	19		1	
1	21		1	
1	27		1	
1	28		1	
1	30		1	
1	31		1	
1	59			1
1	76			1
1	250			1
1	281			1
1	340			1

Total :

52

22

## APPENDICE 4

Dans cet appendice sont présentées les principales valeurs numériques fournies par le programme qui est décrit dans le chapitre 3. Les quantités  $\lambda$ ,  $\lambda_1$ ,  $V$  et  $V'$  sont définies dans (3.2). Les six tables correspondent aux combinaisons linéaires de novice ( $N$ ) et idéal ( $I$ ) expliquées dans (3.2). A côté de chaque table se trouvent les coefficients de  $N$  et  $I$ , ainsi que la matrice des  $[p^j, q^j]$  qui en résulte.

A titre d'exemple, la combinaison :

$$0,5 N + 0,5 I = \begin{bmatrix} 0,6667 & 0,1667 \\ 0,1667 & 0,1667 \\ 0,1667 & 0,6667 \end{bmatrix}$$

est expliquée plus en détail dans (3.2).

Après les tables se trouvent les graphiques qui illustrent le comportement général du modèle dans les différents cas. Toutes les lettres imprimées par cet appareil sont des majuscules. Les valeurs des  $[p^j, q^j]$  sont imprimées sans confusion, mais le paramètre,  $d$ , s'est écrit comme  $D$ . Les graphiques (figures 1 à 16) sont commentées dans (3.2).

Table 1

$$[p^j, q^j] = 1 I$$

=

$$\begin{bmatrix} 0,9999 & 0,0001 \\ 0,0 & 0,0 \\ 0,0001 & 0,9999 \end{bmatrix}$$

$d$	$f_1 \rightarrow$	0,005	0,01	0,02	0,05	0,1
0	$\lambda_{\text{opt}}$	0,0049	0,0091	0,019	0,046	0,093
	$V_{\text{opt}}$	-0,0022	-0,0044	-0,0087	-0,022	-0,043
	$\lambda_{1\text{opt}}$	0,96	0,898	0,936	0,923	0,929
	$V'_{\text{opt}}$	0,0901	0,18	0,360	0,9	1,8
10	$V_{\text{opt}}$	0,051	0,01	0,02	0,05	0,1
	$V_{\text{opt}}$	-0,0023	-0,0045	-0,009	-0,022	-0,045
	$\lambda_{1\text{opt}}$	1,0	1,0	1,0	1,0	1,0
	$V'_{\text{opt}}$	0,0901	0,18	0,36	0,901	1,8
100	$\lambda_{\text{opt}}$	0,0051	0,01	0,02	0,05	0,1
	$V_{\text{opt}}$	-0,0023	-0,0045	-0,009	-0,022	-0,045
	$\lambda_{1\text{opt}}$	1,0	1,0	1,0	1,0	1,0
	$V'_{\text{opt}}$	0,0901	0,18	0,36	0,901	1,8
25 000	$\lambda_{\text{opt}}$	0,0051	0,01	0,02	0,05	0,1
	$V_{\text{opt}}$	-0,0023	-0,0045	-0,009	-0,022	-0,045
	$\lambda_{1\text{opt}}$	1,0	1,0	1,0	1,0	1,0
	$V'_{\text{opt}}$	0,0901	0,1802	-0,36	0,901	1,8

Table 2

$$[p^j, q^j] = 0,15 N + 0,75 I$$

=

$$\begin{bmatrix} 0,8333 & 0,0833 \\ 0,0833 & 0,0833 \\ 0,0833 & 0,8333 \end{bmatrix}$$

$d$	$f_1 \rightarrow$	0,005	0,01	0,02	0,05	0,1
0	$\lambda_{\text{opt}}$	0,038	0,058	0,086	0,162	0,281
	$V_{\text{opt}}$	-0,0376	-0,0552	-0,0819	-0,1388	-0,1994
	$\lambda_{1\text{opt}}$	0,338	0,456	0,575	0,713	0,783
	$V'_{\text{opt}}$	0,0695	0,151	0,317	0,836	1,73
10	$\lambda_{\text{opt}}$	0,033	0,045	0,108	0,171	0,325
	$V_{\text{opt}}$	-0,0393	-0,0621	-0,0909	-0,1514	-0,2183
	$\lambda_{1\text{opt}}$	0,306	0,3949	0,75	0,809	0,878
	$V'_{\text{opt}}$	0,066	0,1386	0,318	0,826	1,73
100	$\lambda_{\text{opt}}$	0,079	0,089	0,157	0,217	0,291
	$V_{\text{opt}}$	-0,055	-0,0746	-0,1116	-0,1719	-0,2592
	$\lambda_{1\text{opt}}$	0,732	0,756	0,868	0,893	0,898
	$V'_{\text{opt}}$	0,071	0,146	0,319	0,827	1,67
25 000	$\lambda_{\text{opt}}$	0,078	0,081	0,172	0,193	0,232
	$V_{\text{opt}}$	-0,0579	-0,0827	-0,1140	-0,1814	-0,3010
	$\lambda_{1\text{opt}}$	0,745	0,742	0,897	0,894	0,89
	$V'_{\text{opt}}$	0,067	0,134	0,324	0,807	1,61

Table 3

$$[p^j, q^j] = 0,5 N + 0,5 I$$

=

$$\begin{bmatrix} 0,6667 & 0,1667 \\ 0,1667 & 0,1667 \\ 0,1667 & 0,6667 \end{bmatrix}$$

$d$	$f_1 \rightarrow$	0,005	0,01	0,02	0,05	0,1
0	$\lambda_{opt}$	0,043	0,085	0,15	0,299	0,435
	$V_{opt}$	-0,061	-0,0947	-0,14	-0,214	-0,275
	$\lambda'_{1opt}$	0,154	0,258	0,378	0,546	0,667
	$V'_{opt}$	0,049	0,124	0,288	0,821	1,72
10	$\lambda_{opt}$	0,043	0,078	0,126	0,313	0,499
	$V_{opt}$	-0,061	-0,097	-0,152	-0,237	-0,311
	$\lambda'_{1opt}$	0,155	0,252	0,369	0,653	0,787
	$V'_{opt}$	0,049	-0,118	0,265	0,805	1,72
100	$\lambda_{opt}$	0,017	0,139	0,155	0,364	1,0
	$V_{opt}$	-0,0757	-0,133	-0,204	-0,311	-0,45
	$\lambda'_{1opt}$	0,068	0,508	0,531	0,763	1,0
	$V'_{opt}$	0,022	0,109	0,227	0,754	1,80
25 000	$\lambda_{opt}$	0,0	0,130	0,288	0,305	1,0
	$V_{opt}$	-0,0901	-0,149	-0,221	-0,369	-0,45
	$\lambda'_{1opt}$	0,0	0,499	0,745	0,741	1,0
	$V'_{opt}$	0,0	0,090	0,269	0,669	1,8

Table 4

$$[p^j, q^j] = 0,75 N + 0,25 I$$

=

$$\begin{bmatrix} 0,5 & 0,25 \\ 0,25 & 0,25 \\ 0,25 & 0,5 \end{bmatrix}$$

$d$	$f_1 \rightarrow$	0,005	0,01	0,02	0,05	0,1
0	$\lambda_{opt}$	0,043	0,105	0,211	0,378	0,529
	$V_{opt}$	-0,0757	-0,1215	-0,177	-0,257	-0,315
	$\lambda'_{1opt}$	0,071	0,153	0,271	0,443	0,586
	$V'_{opt}$	0,033	0,106	0,278	0,814	1,72
10	$\lambda_{opt}$	0,043	0,10	0,195	0,377	0,57
	$V_{opt}$	-0,0759	-0,122	-0,182	-0,286	-0,371
	$\lambda'_{1opt}$	0,072	0,146	0,259	0,483	0,669
	$V'_{opt}$	0,033	0,103	0,266	0,785	1,69
100	$\lambda_{opt}$	0,017	0,028	0,170	0,796	1,0
	$V_{opt}$	-0,083	-0,156	-0,272	-0,448	-0,45
	$\lambda'_{1opt}$	0,032	0,056	0,302	0,847	1,0
	$V'_{opt}$	0,0145	0,0365	0,165	0,811	1,8
25 000	$\lambda_{opt}$	0,0	0	0,308	1,0	1,0
	$V_{opt}$	-0,09	-0,18	-0,3227	-0,45	-0,45
	$\lambda'_{1opt}$	0	0	0,489	1,0	1,0
	$V'_{opt}$	0	0	0,176	0,9	1,8

Table 5

$$[p^j, p^j] = 0,9 N + 0,1 I$$

=

$$\begin{bmatrix} 0,4 & 0,3 \\ 0,3 & 0,3 \\ 0,3 & 0,4 \end{bmatrix}$$

$d$	$f_1 \rightarrow$	0,005	0,01	0,02	0,05	0,1
0	$\lambda_{opt}$	0,038	0,114	0,223	0,409	0,552
	$V_{opt}$	-0,0805	-0,130	-0,188	-0,269	-0,326
	$\lambda_{1, opt}$	0,043	0,122	0,233	0,420	0,561
	$V'_{opt}$	0,027	0,102	0,272	0,816	1,72
10	$\lambda_{opt}$	0,037	0,113	0,214	0,352	0,534
	$V_{opt}$	-0,0806	-0,131	-0,191	-0,292	-0,394
	$\lambda_{1, opt}$	0,042	0,121	0,225	0,368	0,555
	$V'_{opt}$	0,026	0,10	0,266	0,768	1,65
100	$\lambda_{opt}$	0,018	0,041	0,141	1,0	1,0
	$V_{opt}$	-0,0853	-0,157	-0,284	-0,45	-0,45
	$\lambda_{1, opt}$	0,0217	0,049	0,17	1,0	1,0
	$V'_{opt}$	0,0129	0,041	0,14	0,901	1,8
25 000	$\lambda_{opt}$	0	0	0,076	1,0	1,0
	$V_{opt}$	-0,09	-0,18	-0,358	-0,45	-0,45
	$\lambda_{1, opt}$	0	0	0,100	1,0	1,0
	$V'_{opt}$	0	0	0,037	0,901	1,8

Table 6

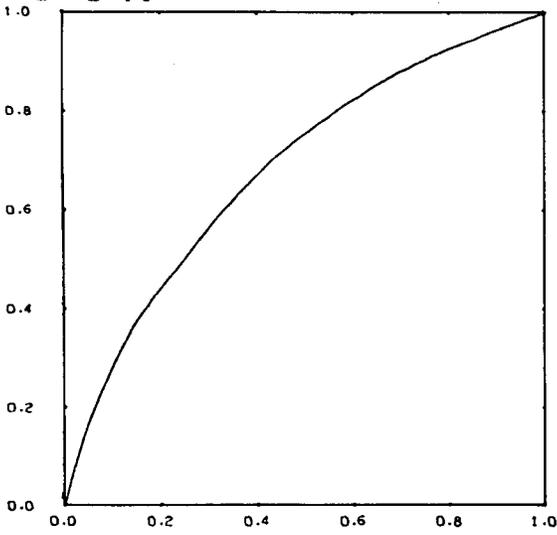
$$[p^j, q^j] = 1 N$$

=

$$\begin{bmatrix} 0,3333 & 0,3333 \\ 0,3333 & 0,3333 \\ 0,3333 & 0,3333 \end{bmatrix}$$

$d$	$f_1 \rightarrow$	0,005	0,01	0,02	0,05	0,1
0	$\lambda_{opt}$	0,034	0,119	0,231	0,416	0,559
	$V_{opt}$	-0,081	-0,132	-0,191	-0,272	-0,328
	$\lambda_{1, opt}$	0,034	0,119	0,231	0,416	0,559
	$V'_{opt}$	0,024	0,102	0,274	0,816	1,725
10	$\lambda_{opt}$	0,034	0,118	0,220	0,369	0,468
	$V_{opt}$	-0,081	-0,132	-0,193	-0,291	-0,398
	$\lambda_{1, opt}$	0,034	0,118	0,220	0,369	0,468
	$V'_{opt}$	0,024	0,101	0,266	0,776	1,615
100	$\lambda_{opt}$	0,021	0,058	0,10	1,0	1,0
	$V_{opt}$	-0,085	-0,154	-0,273	-0,45	-0,45
	$\lambda_{1, opt}$	0,021	0,058	0,10	1,0	1,0
	$V'_{opt}$	0,014	0,052	0,132	0,901	1,802
25 000	$\lambda_{opt}$	0	0	0	1,0	1,0
	$V_{opt}$	-0,09	-0,18	-0,36	-0,45	-0,45
	$\lambda_{1, opt}$	0	0	0	1,0	1,0
	$V'_{opt}$	0	0	0	0,901	1,802

LAMDA.1

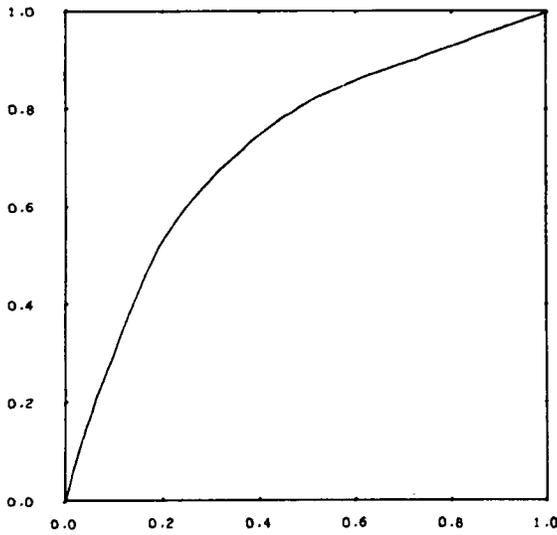


F1 = 0.020  
D = 0.0

P1=0.6667 Q1=0.1667  
P2=0.1667 Q2=0.1667  
P3=0.1667 Q3=0.6667

LAMDA FIG. [1]

LAMDA.1

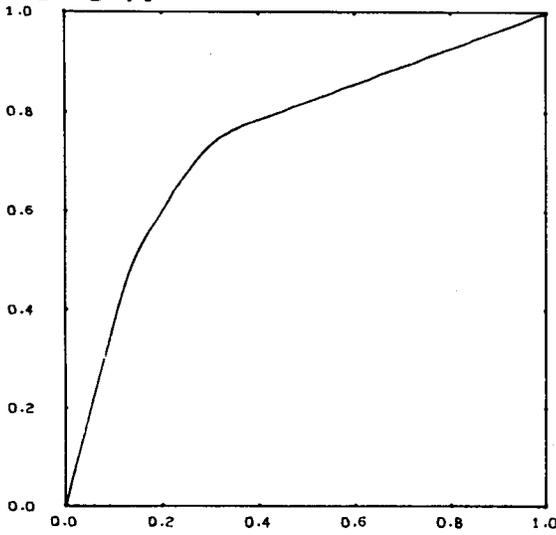


F1 = 0.020  
D = 10.0

P1=0.6667 Q1=0.1667  
P2=0.1667 Q2=0.1667  
P3=0.1667 Q3=0.6667

LAMDA FIG. [2]

FRAC<sup>2</sup> LAMDA.1

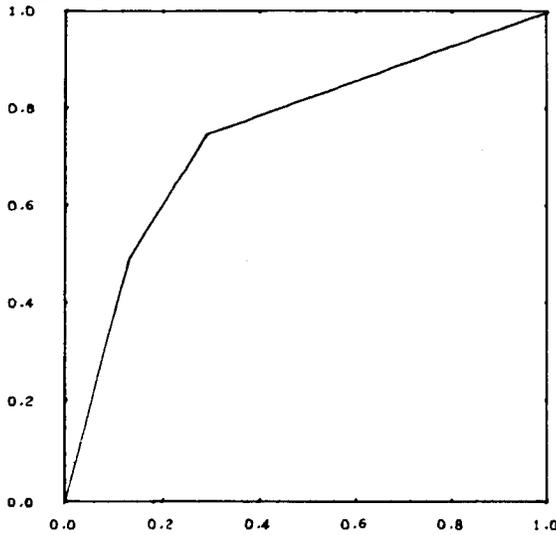


F1 = 0.020  
D = 100.0

P1=0.6667 Q1=0.1667  
P2=0.1667 Q2=0.1667  
P3=0.1667 Q3=0.6667

LAMDA FIG. [3]

LAMDA.1

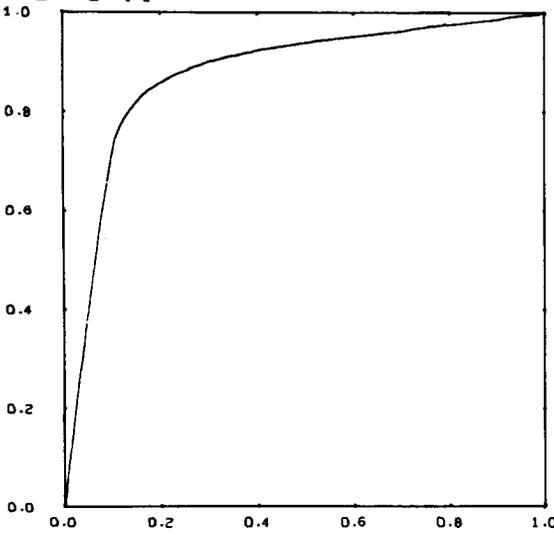


F1 = 0.020  
D = 25000.0

P1=0.6667 Q1=0.1667  
P2=0.1667 Q2=0.1667  
P3=0.1667 Q3=0.6667

LAMDA FIG. [4]

LAMDA.1

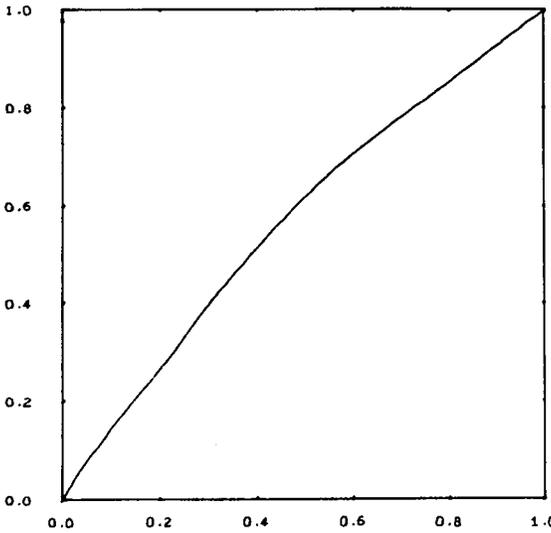


F1 = 0.020  
D = 10.0

P1 = 0.8333 Q1 = 0.0833  
P2 = 0.0833 Q2 = 0.0833  
P3 = 0.0833 Q3 = 0.8333

LAMDA FIG. [5]

LAMDA.1

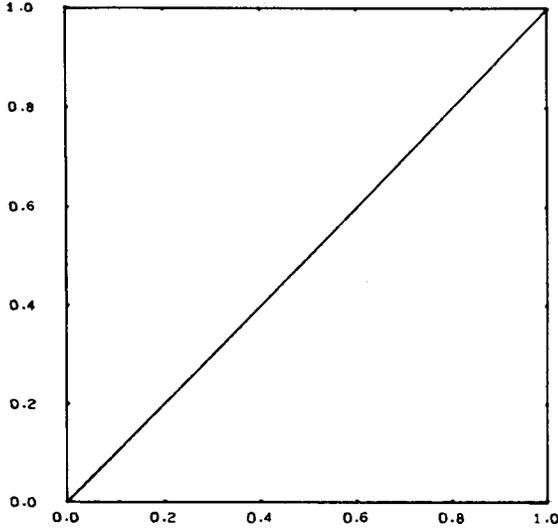


F1 = 0.020  
D = 10.0

P1 = 0.5000 Q1 = 0.2500  
P2 = 0.2500 Q2 = 0.2500  
P3 = 0.2500 Q3 = 0.5000

LAMDA FIG. [6]

FRAME 1 LAMDA.1

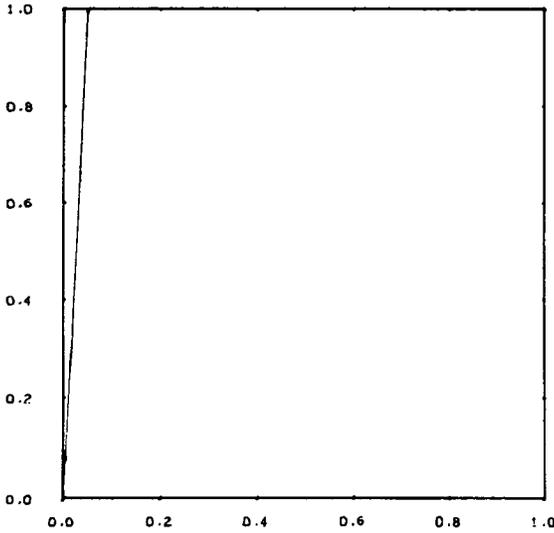


F1 = 0.050  
D = 100.0

P1 = 0.33333 Q1 = 0.33333  
P2 = 0.33333 Q2 = 0.33333  
P3 = 0.33333 Q3 = 0.33333

LAMDA FIG. [7]

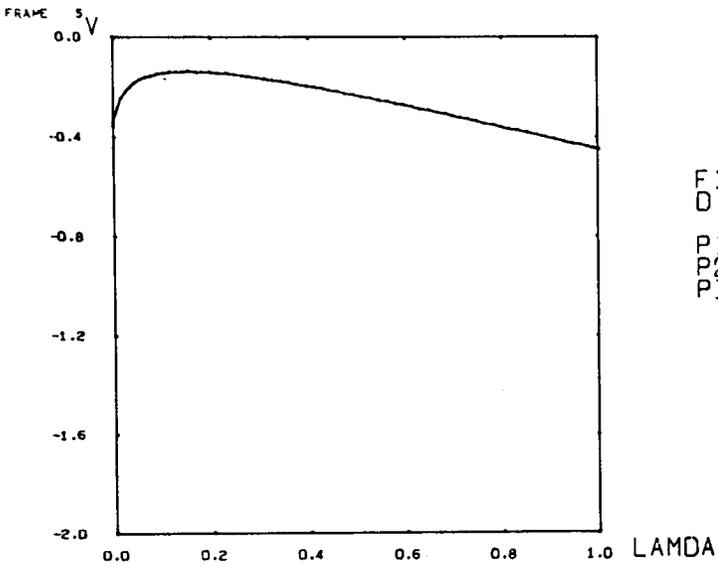
LAMDA.1



F1 = 0.050  
D = 0.0

P1 = 0.9999 Q1 = 0.0001  
P2 = 0.0000 Q2 = 0.0000  
P3 = 0.0001 Q3 = 0.9999

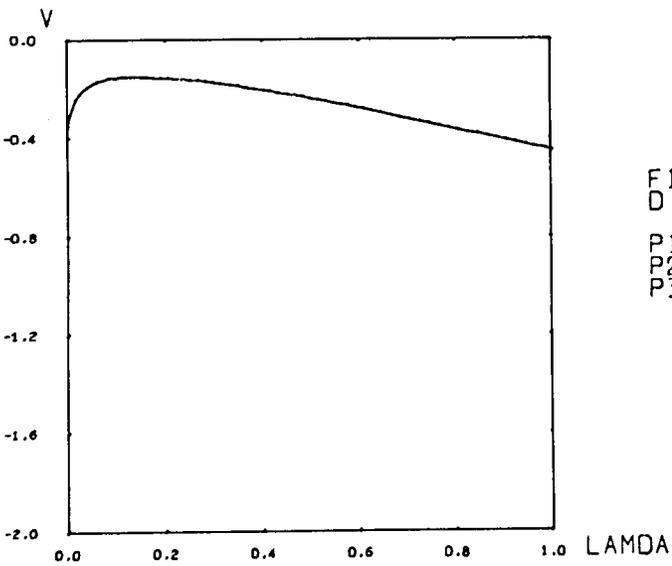
LAMDA FIG. [8]



F1 = 0.020  
D = 0.0

P1=0.6667 Q1=0.1667  
P2=0.1667 Q2=0.1667  
P3=0.1667 Q3=0.6667

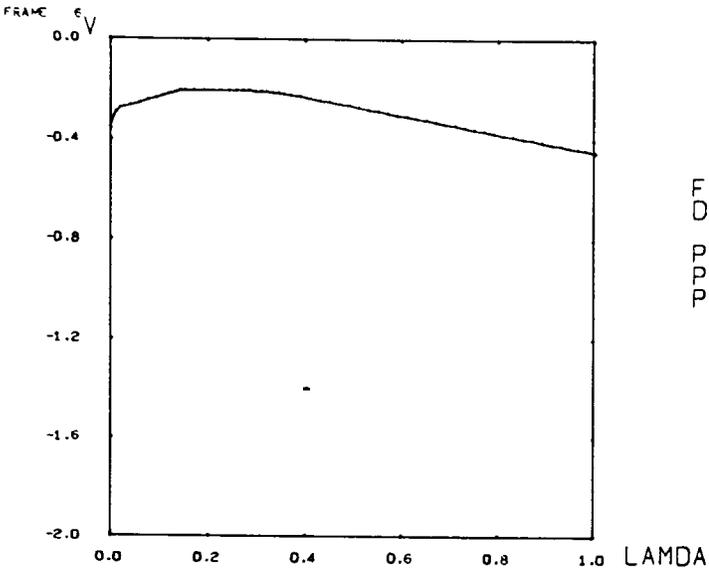
FIG. [9]



F1 = 0.020  
D = 10.0

P1=0.6667 Q1=0.1667  
P2=0.1667 Q2=0.1667  
P3=0.1667 Q3=0.6667

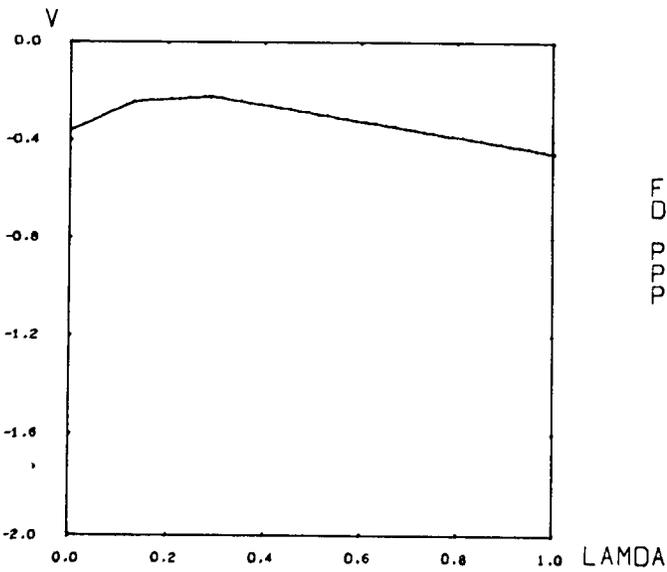
FIG. [10]



F1 = 0.020  
D = 100.0

P1=0.6667 Q1=0.1667  
P2=0.1667 Q2=0.1667  
P3=0.1667 Q3=0.6667

FIG. [11]



F1 = 0.020  
D = 25000.0

P1=0.6667 Q1=0.1667  
P2=0.1667 Q2=0.1667  
P3=0.1667 Q3=0.6667

FIG. [12]

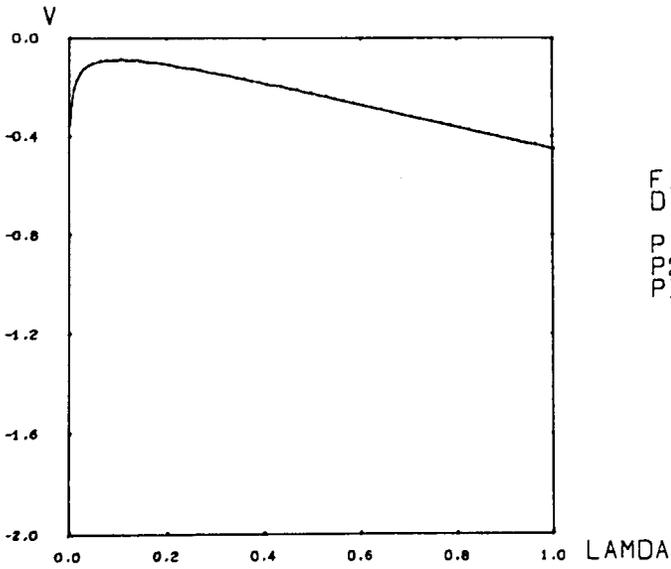


FIG. [13]

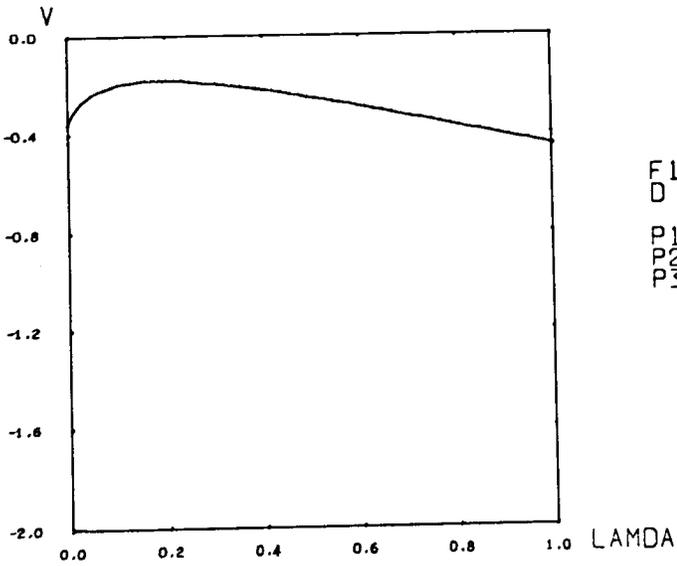
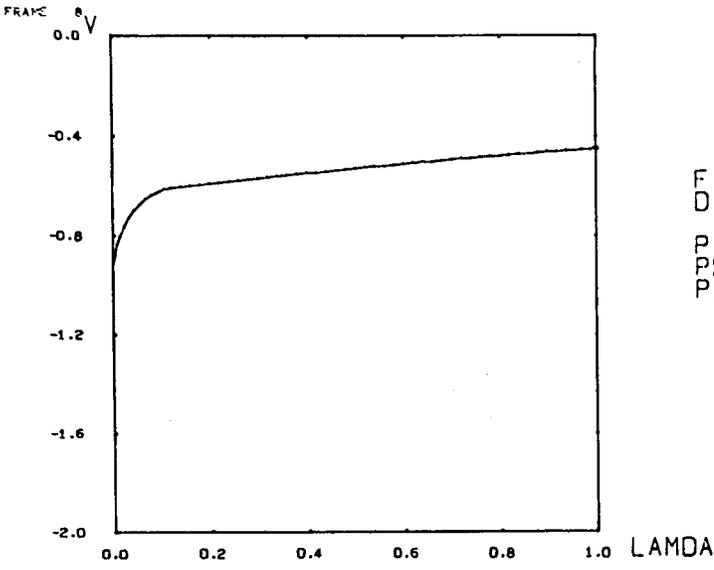


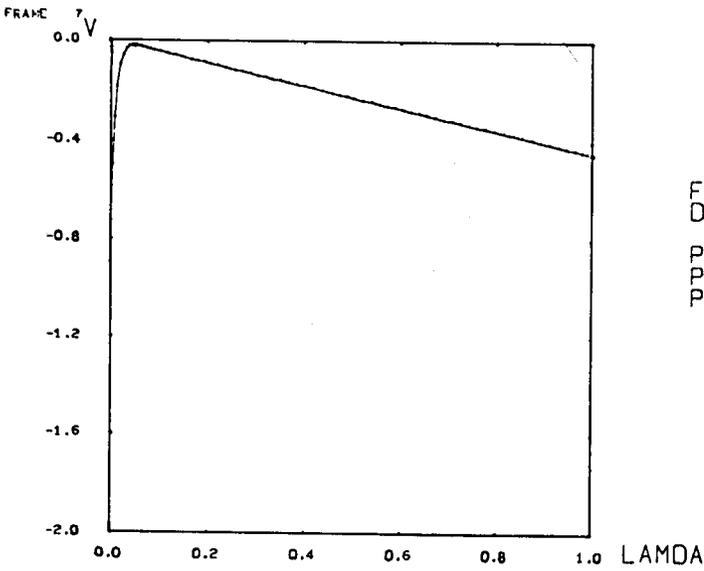
FIG. [14]



F1 = 0.050  
D = 100.0

P1=0.3333 Q1=0.3333  
P2=0.3333 Q2=0.3333  
P3=0.3333 Q3=0.3333

FIG. [15]



F1 = 0.050  
D = 0.0

P1=0.9999 Q1=0.0001  
P2=0.0000 Q2=0.0000  
P3=0.0001 Q3=0.9999

FIG. [16]

## BIBLIOGRAPHIE

- [1] GJEDDEBAEK, N.F. – “Contributions to the study of grouped observations, part I” SKANDINAVISK AKTUARIETIDSKRIFT vol. 32 (1949) pp. 135-59.
- [2] KULLDORF, G. – “Maximum likelihood estimation of the mean of a normal random variable when the sample is grouped”. SKAND. AKT. Vol. 41 (1958) pp. 1-36.
- [3] GJEDDEBAEK, N.F. – “Contributions to the study of grouped observations, part II” SKAND. AKT. Vol. 39 (1956) pp. 154-9.
- [4] GJEDDEBAEK, N.F. – Part III SKAND. AKT., Vol. 40 (1957) pp. 20-5
- [5] GJEDDEBAEK, N.F. – Part IV BIOMETRICS, 15, (1959) pp. 433-9.
- [6] GJEDDEBAEK, N.F. – Part V SKAND. AKT., Vol. 42 (1959) pp. 194-207.
- [7] GJEDDEBAEK, N.F. – Part VI SKAND. AKT., Vol. 44 (1961) pp. 55-73.
- [8] KULLDORF, G. – “On the conditions for consistency and asymptotic efficiency of maximum likelihood estimates”. SKAND. AKT. Vol.40 (1957) pp. 129-44.
- [9] STEVENS, W.L. – “Control by gauging”. JOURNAL OF THE ROYAL STATISTICAL SOCIETY, SERIES B Vol. 10 (1948) pp. 54-108.