

Record Number: 12100
Author, Monographic: Ballay, S.//Perreault, L.//Bobée, B.
Author Role:
Title, Monographic: Loi des fuites
Translated Title:
Reprint Status:
Edition:
Author, Subsidiary:
Author Role:
Place of Publication: Québec
Publisher Name: INRS-Eau
Date of Publication: 1996
Original Publication Date: Septembre 1996
Volume Identification:
Extent of Work: ix, 97
Packaging Method: pages incluant 5 annexes
Series Editor:
Series Editor Role:
Series Title: INRS-Eau, rapport de recherche
Series Volume ID: 479
Location/URL:
ISBN: 2-89146-424-9
Notes: Rapport annuel 1996-1997
Abstract: Chaire en hydrologie statistique
Call Number: R000479
Keywords: rapport/ ok/ dl

LOI DES FUITES

LOI DES FUTES

par

**Sophie Ballay
Luc Perreault
Bernard Bobée**

**Chaire en Hydrologie statistique
Institut national de la recherche scientifique, INRS-Eau
2800, rue Einstein, Case postale 7500, SAINTE-FOY (Québec), G1V 4C7**

Rapport de recherche No R-479

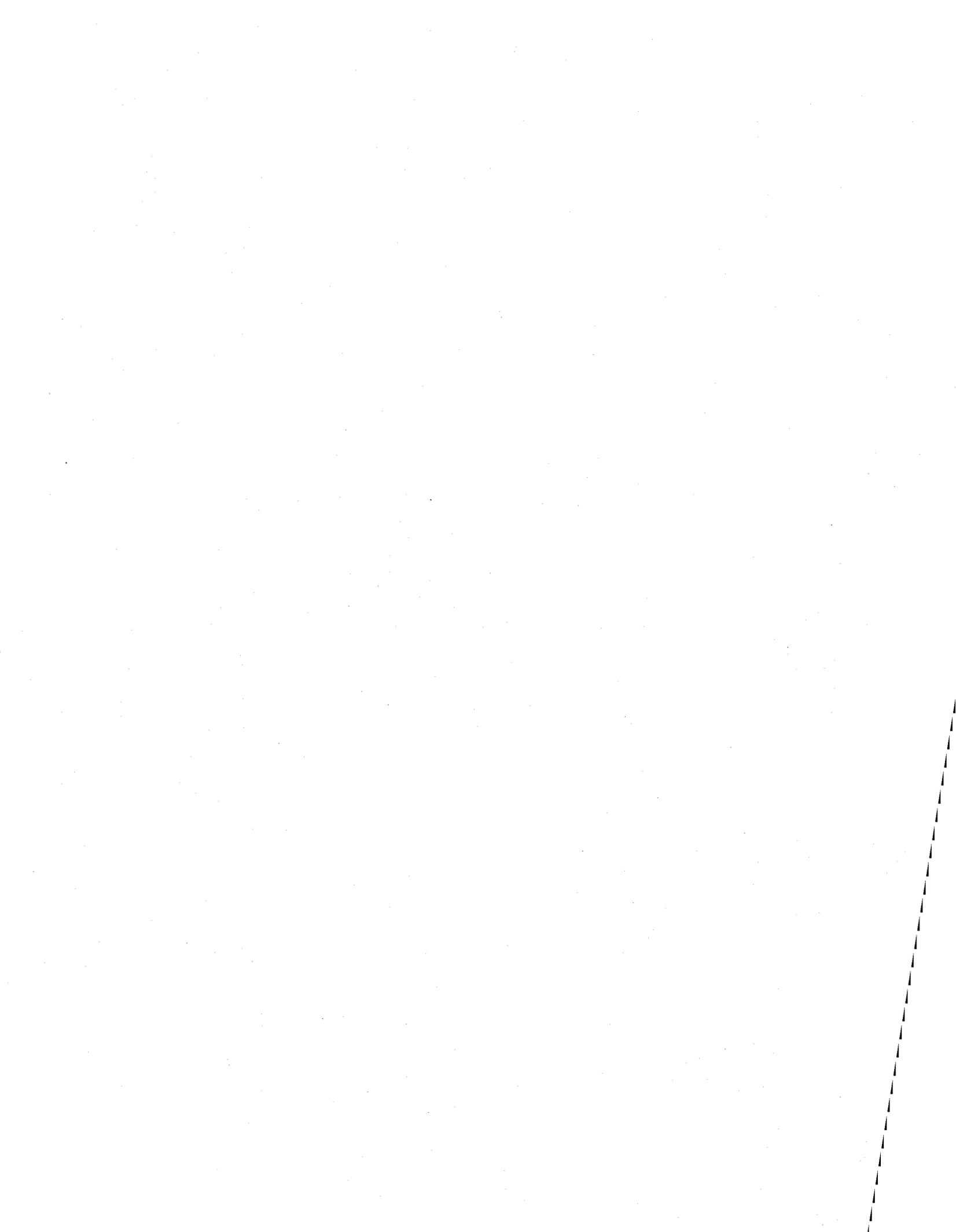
Septembre 1996

TABLE DES MATIERES

LISTE DES TABLEAUX.....	vii
LISTE DES FIGURES.....	ix
CHAPITRE 1 : Introduction	1
1.1 Problématique	1
1.2 Historique.....	2
1.3 Contenu du rapport	3
CHAPITRE 2 : Propriétés mathématiques et statistiques.....	5
2.1 Définition de la loi des fuites	5
2.2 Propriétés mathématiques	6
2.2.1 Fonction de densité.....	6
2.2.2 Fonction de répartition	8
2.3 Propriétés statistiques	9
2.3.1 Fonction caractéristique.....	9
2.3.2 Cumulants, moments et coefficients	10
2.4 Etude asymptotique lorsque λ tend vers l'infini	11
2.5 Comparaison avec la loi Gamma	12
CHAPITRE 3 : Estimation des paramètres	15
3.1 Méthode du maximum de vraisemblance.....	15
3.1.1 Fonction de vraisemblance	16
3.1.2 Système d'équations.....	16
3.1.3 Méthodes numériques	18
3.1.4 Variances et covariance asymptotiques	18
3.2 Méthode des moments.....	21
3.2.1 Système d'équations et solution.....	21
3.2.2 Variances et covariance asymptotiques	22

3.3 Autres méthodes	24
3.3.1 Méthodes faisant intervenir n_0/n	24
3.3.2 Méthode utilisant la moyenne des racines carrées des observations	25
CHAPITRE 4 : Estimation du quantile x_T.....	27
4.1 Estimation de x_T	27
4.2 Variance asymptotique et intervalle de confiance.....	28
4.2.1 Variance asymptotique.....	28
4.2.2 Intervalle de confiance	29
CHAPITRE 5 : Algorithmie.....	31
5.1 Moments et coefficients.....	31
5.2 Estimation des paramètres.....	32
5.2.1 Méthode du maximum de vraisemblance	32
5.2.2 Méthode des moments	33
5.2.3 Méthodes faisant intervenir n_0/n	33
5.3 Variances et covariance asymptotiques des paramètres	34
5.3.1 Méthode du maximum de vraisemblance	34
5.3.2 Méthode des moments	36
5.3.3 Méthodes faisant intervenir n_0/n	36
5.4 Détermination du quantile x_T	37
5.4.1 Calcul de la fonction de répartition $F(u)$	37
5.4.2 Calcul de x_T	38
5.5 Variance de \hat{x}_T	38
CHAPITRE 6 : Application	41
6.1 Exemple I.....	41
6.1.1 Estimation des paramètres	41
6.1.2 Variances et covariance des paramètres.....	43
6.1.3 Estimation des quantiles.....	44
6.2 Exemple II.....	46
6.2.1 Estimation des paramètres	46
6.2.2 Variances et covariance des paramètres.....	48
6.2.3 Estimation des quantiles.....	48

6.3 Exemple III.....	50
6.3.1 Estimation des paramètres	51
6.3.2 Variances et covariance des paramètres	51
6.3.3 Estimation des quantiles	52
CHAPITRE 7 : Générateur d'échantillons pour la loi des fuites	57
7.1 Méthodes de génération	57
7.1.1 Méthode d'inversion	57
7.1.2 Méthode d'acceptation-rejet	57
7.2 Génération de la loi des fuites par la méthode d'acceptation-rejet	58
7.2.1 $\lambda < 3 - \sqrt{3}$	58
7.2.2 $3 - \sqrt{3} \leq \lambda < 2$	59
7.2.3 $2 \leq \lambda < 3 + \sqrt{3}$	60
7.2.4 $\lambda \geq 3 + \sqrt{3}$	61
7.2.5 Conclusion	62
CHAPITRE 8 : Conclusion et recommandations	63
REFERENCES BIBLIOGRAPHIQUES	65
ANNEXE A : Principe de la méthode de Newton-Raphson.....	67
ANNEXE B : Calcul des coefficients de la matrice de Fisher	69
ANNEXE C : Définition de l'efficacité d'un estimateur.....	81
ANNEXE D : Calcul des dérivées partielles du quantile de période de retour T par rapport à chacun des paramètres de la loi.....	83
ANNEXE E : Programmation dans Matlab et Mathematica	87



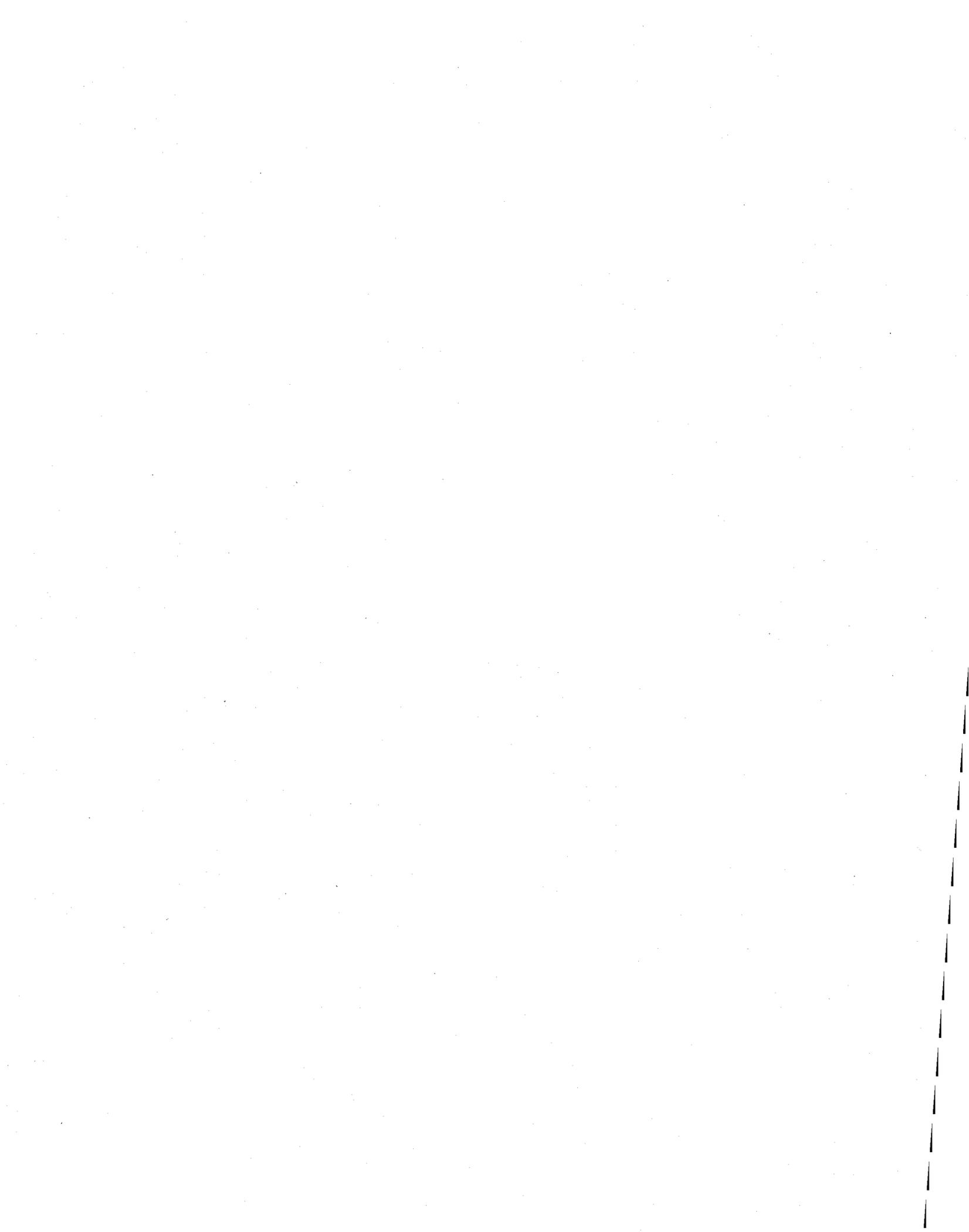
LISTE DES TABLEAUX

Tableau 2.1 : Tableau comparatif des coefficients de la loi des fuites et de la loi Gamma.....	13
Tableau 6.1 : Précipitations (mm) relevées à Massiac ($n=86, n_o=1$)	41
Tableau 6.2 : Estimation des paramètres λ et β	42
Tableau 6.3 : Résultats du test du khi-deux	43
Tableau 6.4 : Coefficients de la matrice d'information de Fisher.....	44
Tableau 6.5 : Variances et covariance des paramètres.....	44
Tableau 6.6 : Estimation des quantiles de période de retour T	45
Tableau 6.7 : Variances des quantiles de période de retour T	45
Tableau 6.8 : Intervalles de confiance associés aux quantiles de période de retour T	46
Tableau 6.9 : Précipitations (mm) relevées à Chateauneuf de Randon ($n=86, n_o=15$).....	46
Tableau 6.10 : Estimation des paramètres λ et β	47
Tableau 6.11 : Résultats du test du khi-deux	47
Tableau 6.12 : Coefficients de la matrice d'information de Fisher.....	48
Tableau 6.13 : Variances et covariance des paramètres.....	48
Tableau 6.14 : Estimation des quantiles de période de retour T pour les paramètres estimés par la méthode du maximum de vraisemblance.....	49
Tableau 6.15 : Estimation des quantiles de période de retour T pour les paramètres estimés par la méthode utilisant n_o et $n(1)$	49
Tableau 6.16 : Variances des quantiles de période de retour T	50
Tableau 6.17 : Intervalles de confiance associés aux quantiles de période de retour T	50
Tableau 6.18 : Estimation des paramètres λ et β	51
Tableau 6.19 : Coefficients de la matrice d'information de Fisher.....	51
Tableau 6.20 : Variances et covariance des paramètres.....	52
Tableau 6.21 : Estimation des quantiles de période de retour T pour les paramètres estimés par la méthode du maximum de vraisemblance.....	52
Tableau 6.22 : Estimation des quantiles de période de retour T pour les paramètres estimés par la méthode des moments.....	53

Tableau 6.23 : Estimation des quantiles de période de retour T pour les paramètres estimés par la méthode utilisant n_o et n (1)	53
Tableau 6.24 : Estimation des quantiles de période de retour T pour les paramètres estimés par la méthode utilisant n_o et n (2)	54
Tableau 6.25 : Variances des quantiles de période de retour T	54
Tableau 6.26 : Intervalles de confiance associés aux quantiles de période de retour T	55

LISTE DES FIGURES

Figure 2.1 : Fonction de densité de probabilité	8
Figure 2.1 : Fonction de répartition	9
Figure 7.1 : Fonction de densité et fonction majorante pour $\lambda = 1$	59
Figure 7.2 : Fonction de densité et fonction majorante pour $\lambda = 1,5$	60
Figure 7.3 : Fonction de densité et fonction majorante pour $\lambda = 3$	61
Figure 7.4 : Fonction de densité et fonction majorante pour $\lambda = 8$	62



1 INTRODUCTION

1.1 Problématique

Les activités des grandes compagnies hydroélectriques nécessitent de nombreuses études concernant les débits extrêmes de crue. La planification et le dimensionnement d'ouvrages hydrauliques tels que barrages, évacuateurs et dérivations provisoires reposent en effet sur une estimation adéquate des événements extrêmes de crue. Si une surestimation des crues et donc un sur-dimensionnement des ouvrages peut entraîner des coûts de construction supplémentaires inutiles, une sous-estimation des crues présente le risque d'une défaillance d'ouvrages conduisant à des inondations qui se traduisent par des dégâts matériels et parfois des pertes en vies humaines.

Un des outils privilégiés par les hydrologues pour estimer les débits extrêmes de crue est l'analyse de fréquence des crues (flood frequency analysis). Cette approche a pour objectif l'utilisation des mesures d'événements hydrologiques extrêmes passés pour estimer les probabilités futures d'occurrence. On tente en particulier d'estimer l'événement extrême x_T de période de retour T . En ajustant une loi de probabilité $F(x, \underline{\theta})$, où $\underline{\theta}$ désigne le vecteur de paramètres associé à la distribution F à un échantillon de mesures d'événements extrêmes observés historiquement, on peut obtenir une estimation \hat{x}_T de la valeur théorique du quantile inconnu x_T . Si l'échantillon satisfait certaines hypothèses statistiques de base, les paramètres $\underline{\theta}$ peuvent être estimés en ajustant la distribution $F(x, \underline{\theta})$ aux observations. On montre alors que le quantile estimé de période de retour T est donné par :

$$\hat{x}_T = F^{-1}(1 - 1/T; \hat{\underline{\theta}}) \quad (1.1)$$

où $F(x, \hat{\underline{\theta}})$ représente la fonction de répartition calculée avec les paramètres estimés $\hat{\underline{\theta}}$.

Le choix d'une loi statistique comme approximation de la distribution théorique inconnue sera justifié par la qualité de son adéquation aux données observées. Plusieurs lois de probabilité ont été développées en hydrologie afin de modéliser les événements extrêmes. Bobée et Ashkar ont présenté en 1991 une synthèse des principales méthodes d'estimation des paramètres des lois de la famille Gamma étudiées à l'INRS-Eau et ailleurs dans le monde. Un logiciel d'ajustement automatique de ces distributions incluant les méthodes d'estimation de leurs paramètres accompagnait cette synthèse. Ce logiciel est paru sous le nom de HFA, et AJUSTE-I pour sa version française. Une nouvelle version, AJUSTE-II a ensuite été développée dans le cadre d'un projet de partenariat financé par Hydro-Québec et

le Conseil de Recherche en Sciences Naturelles et en Génie (CRSNG). Les améliorations apportées concernent la structure et la convivialité du logiciel mais aussi les méthodes statistiques employées. Alors que HFA était écrit en Fortran, AJUSTE-II utilise le langage C++ orienté objet et la librairie XVT qui permet de travailler dans les environnements graphiques *Windows* (DOS) et *X-Windows* (UNIX). Parmi les nombreuses méthodes d'estimation des paramètres que proposait HFA pour les lois de la famille Gamma, AJUSTE-II ne retient que les plus efficaces. Des tests statistiques permettent de valider les hypothèses de base et l'adéquation des modèles, et quelques lois parmi les plus fréquemment utilisées en hydrologie ainsi que les trois types de lois de Halphen ont été ajoutées.

Mais les lois classique sont peu aptes à représenter des séries comportant des zéros, car elles obligent à un tronquage de la loi, entraînant des calculs supplémentaires qui peuvent être complexes. Or, l'hydrologue est souvent amené à traiter des échantillons qui présentent un grand nombre de valeurs nulles: hauteurs d'eau tombée par précipitation pendant un certain intervalle de temps, débits moyens journaliers ou même mensuels dans des régions arides ou semi-arides comme le Sahel par exemple. La loi des fuites qui est une loi mixte, c'est-à-dire qu'elle attache une probabilité non nulle à la valeur zéro tandis que la densité de probabilité est continue pour les valeurs strictement positives de la variable, s'avère donc particulièrement intéressante pour la représentation de ce type de série.

1.2 Historique

La loi des fuites a d'abord été exposée par Halphen (1955) puis Morlat a montré qu'elle permettait la représentation des débits de fuite le long d'une conduite de gaz, d'où son nom. Bernier (1968), qui travaillait au département Hydrologie de l'EDF comme Halphen et Morlat, a ensuite pensé à utiliser les lois I_γ dont la loi des fuites est la loi particulière I_0 pour l'étude des précipitations en un lieu donné. Il a suivi la thèse de Babuziaux (1969) "Etude statistique de la loi des fuites" qui porte essentiellement sur les modes d'estimation des paramètres. S'inspirant de ce document et à la suite d'une étude de Brunet-Moret (1978), Ribstein (1983), hydrologue à l'ORSTOM Montpellier, s'est intéressé aux problèmes de troncature des observations. Le logiciel de calculs statistiques et d'analyse fréquentielle adapté à l'évaluation du risque en hydrologie (Safarhy), mis en place par l'ORSTOM Montpellier et qui a pour objet d'aider l'hydrologue à ajuster une loi de probabilité théorique à la distribution de fréquence d'un échantillon réel, propose notamment la loi des fuites.

1.3 Contenu du rapport

Nous nous proposons d'étudier dans ce rapport la loi des fuites, les principales méthodes d'estimation de ses paramètres et des quantiles. Les développements théoriques qui y sont présentés sont pour la plupart une synthèse des principaux résultats que l'on retrouve dans la littérature consacrée à cette loi de probabilité. Ce travail a été effectué dans le but d'ajouter la loi des fuites au logiciel AJUSTE-II.

Le Chapitre 2 présente les principales propriétés mathématiques et statistiques de la loi. Dans le Chapitre 3 sont développées différentes méthodes d'estimation des paramètres, principalement la méthode du maximum de vraisemblance et la méthode des moments. Le Chapitre 4 concerne l'estimation des quantiles x_T de période de retour T . Au Chapitre 5, des algorithmes sont proposés en vue d'inclure la loi des fuites dans le logiciel AJUSTE-II. Le Chapitre 6 présente trois applications de l'ajustement de la loi des fuites à des ensembles de données. Enfin, nous abordons au chapitre 7 le problème de la génération d'échantillons.



2 PROPRIETES MATHEMATIQUES ET STATISTIQUES

Ce chapitre décrit les propriétés mathématiques et statistiques de la loi des fuites et donne en particulier les principaux moments centrés de cette distribution.

2.1 Définition de la loi des fuites

Nous utilisons ici la définition donnée par Morlat et reprise par H.A. Einstein (voir Babuziaux, 1969), basée sur la théorie du renouvellement, qui est l'étude d'événements chronologiques de longueur aléatoire et séparés par des durées aléatoires.

Soit Y une variable aléatoire pouvant prendre des valeurs positives ou nulles et qui suit une loi exponentielle simple, de paramètre β , de densité de probabilité:

$$f(y) = \frac{\exp(-y/\beta)}{\beta} \quad (2.1)$$

(Y peut par exemple représenter la hauteur d'eau tombée au cours d'une pluie).

Soit N une variable aléatoire prenant des valeurs entières et qui suit une loi de Poisson de paramètre λ . La probabilité de la valeur n est alors:

$$P(N = n) = \frac{\lambda^n}{n!} \exp(-\lambda) \quad (2.2)$$

où N peut par exemple représenter le nombre d'averses.

Nous allons montrer que la variable aléatoire X définie par $X=Y_1+Y_2+\dots+Y_N$, somme d'un nombre aléatoire N de variables aléatoires indépendantes Y , est distribuée selon une loi des fuites de paramètres λ et β .

On peut d'abord montrer (voir Bobée et Ashkar (1991) p.41), pour $N = n$ fixé, que la loi conditionnelle de X est une loi Gamma s'exprimant de la manière suivante:

$$g(x/n) = \frac{1}{\Gamma(n)} \exp\left(-\frac{x}{\beta}\right) \left(\frac{x}{\beta}\right)^{n-1} \frac{1}{\beta} \quad (2.3)$$

Or, pour obtenir la loi marginale $g(x)$ de X (loi inconditionnelle), il suffit de sommer sur toutes les valeurs possibles de n l'expression $g(x/n)P(N=n)$:

$$g(x) = \sum_{n=1}^{\infty} g(x/n)P(N=n) = \frac{\lambda}{\beta} \exp\left(-\lambda - \frac{x}{\beta}\right) \sum_{n=0}^{\infty} \frac{1}{n!(n+1)!} \left(\frac{\lambda x}{\beta}\right)^n \quad (2.4)$$

Après quelques développements algébriques, l'expression (2.4) peut s'écrire comme suit :

$$g(x) = \frac{\lambda}{\beta} \exp\left(-\lambda - \frac{x}{\beta}\right) \frac{I_1\left(2\sqrt{\frac{\lambda x}{\beta}}\right)}{\sqrt{\frac{\lambda x}{\beta}}} \quad x > 0 \quad (2.5)$$

où I_1 représente la fonction de Bessel d'ordre 1 définie par :

$$I_1(t) = \frac{1}{\pi} \int_0^{\pi} \cos\theta \exp(t \cos\theta) d\theta = \sum_{n=0}^{\infty} \frac{1}{n!(n+1)!} \left(\frac{t}{2}\right)^{2n+1} \quad (2.6)$$

On note en particulier que si $n=0$ alors $X=0$ et $P(X=0)=e^{-\lambda}$.

La loi des fuites peut aussi être introduite à partir de la solution d'une équation différentielle portant sur un processus de Markov (voir Babuziaux (1969)).

2.2 Propriétés mathématiques

2.2.1 Fonction de densité

Nous venons de voir que la loi des fuites s'écrit :

$$g(x) = \begin{cases} e^{-\lambda}, & x = 0 \\ \frac{\lambda}{\beta} \exp\left(-\lambda - \frac{x}{\beta}\right) \frac{I_1\left(2\sqrt{\frac{\lambda x}{\beta}}\right)}{\sqrt{\frac{\lambda x}{\beta}}}, & x > 0 \end{cases} \quad (2.7)$$

Si on pose $U=X/\beta$, la densité de probabilité de la loi des fuites peut s'exprimer au moyen du seul paramètre λ par :

Si on pose $U=X/\beta$, la densité de probabilité de la loi des fuites peut s'exprimer au moyen du seul paramètre λ par :

$$f(u) = \begin{cases} e^{-\lambda}, & u = 0 \\ \lambda \exp(-\lambda - u) \frac{I_1(2\sqrt{\lambda u})}{\sqrt{\lambda u}}, & u > 0 \end{cases} \quad (2.8)$$

Cette distribution est mixte puisqu'il y a discontinuité en zéro (sauf dans le cas où $\lambda=1$). En effet, l'ordonnée à l'origine de la fonction f est $\lambda e^{-\lambda}$ alors que $P(X=0)=e^{-\lambda}$.

De l'étude des dérivées première et seconde de f , on déduit les formes que peut admettre la fonction de densité selon les valeurs de λ . Les dérivées première et seconde s'expriment respectivement comme suit :

$$f'(u) = \lambda \exp(-\lambda - u) \left[\left(\frac{\lambda}{2} - 1 \right) + \sum_{i=1}^{\infty} \frac{(\lambda u)^i}{i!(i+1)!} \left(\frac{\lambda}{i+2} - 1 \right) \right] \quad (2.9)$$

$$f''(u) = \lambda \exp(-\lambda - u) \left[\left(1 - \lambda + \frac{\lambda^2}{6} \right) + \sum_{i=1}^{\infty} \frac{(\lambda u)^i}{i!(i+1)!} \left(1 - \frac{2\lambda}{i+2} + \frac{\lambda^2}{(i+2)(i+3)} \right) \right] \quad (2.10)$$

Ainsi, d'après les expressions (2.9) et (2.10), on a :

- $\lambda < (3-\sqrt{3})$: forme en J renversé, pas d'inflexion, mode en $u=0$
- $(3-\sqrt{3}) \leq \lambda < 2$: forme en J renversé, un point d'inflexion, mode en $u=0$
- $2 \leq \lambda < (3+\sqrt{3})$: un point d'inflexion, mode en $u > 0$
- $\lambda \geq (3+\sqrt{3})$: deux points d'inflexion, mode en $u > 0$

La figure 2.1 illustre la variété des formes que peut prendre la densité f de la loi des fuites pour différentes valeurs du paramètre λ . La courbe prend une allure en cloche, d'autant plus prononcée que λ est grand. On montrera dans le paragraphe 2.4, au moyen des fonctions caractéristiques, que pour $\lambda \rightarrow \infty$, la loi des fuites tend vers la loi normale.

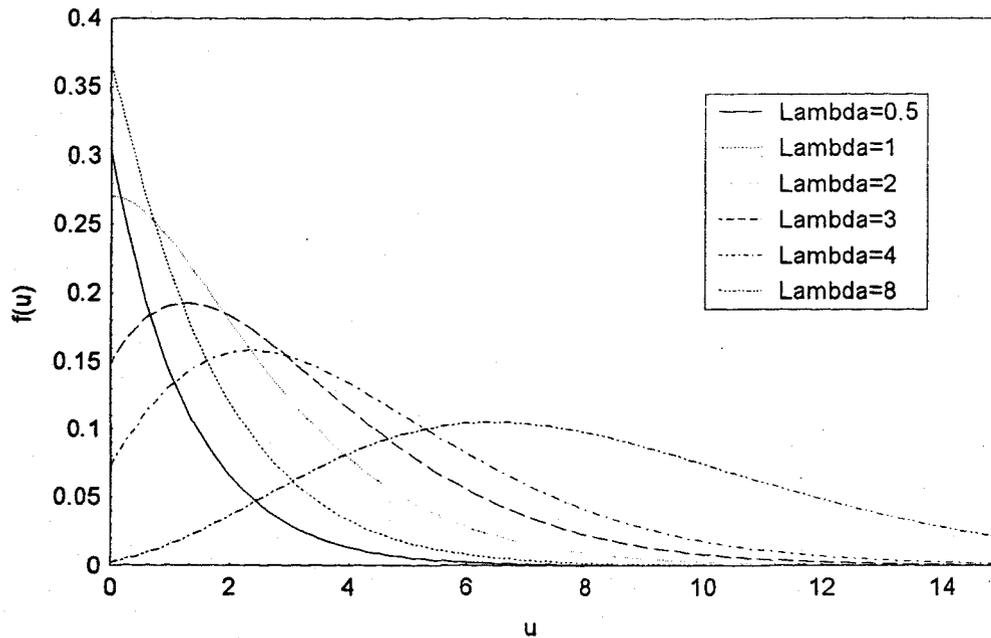


Figure 2.1. Fonction de densité de probabilité.

2.2.2 Fonction de répartition

Par définition, la fonction de répartition s'obtient par intégration de la fonction de densité:

$$\begin{aligned}
 F(u) &= P(X=0) + \int_0^u f(u) du \\
 &= e^{-\lambda} + \lambda e^{-\lambda} \int_0^u e^{-u} \frac{I_1(2\sqrt{\lambda u})}{\sqrt{\lambda u}} du \\
 &= e^{-\lambda} + \lambda e^{-\lambda} \int_0^u e^{-u} \left(\sum_{i=0}^{\infty} \frac{(\lambda u)^i}{i!(i+1)!} \right) du
 \end{aligned} \tag{2.11}$$

En utilisant le résultat suivant :

$$\int_0^u \frac{e^{-u} u^i}{i!} du = -\frac{e^{-u} u^i}{i!} + \int_0^u \frac{e^{-u} u^{i-1}}{(i-1)!} du = -e^{-u} \sum_{j=0}^i \frac{u^j}{j!} + 1 \tag{2.12}$$

on obtient finalement que :

$$F(u) = 1 - e^{-\lambda} e^{-u} \sum_{i=0}^{\infty} \frac{\lambda^{i+1}}{(i+1)!} \left(\sum_{j=0}^i \frac{u^j}{j!} \right) \quad (2.13)$$

La fonction de répartition F dépend d'une somme infinie et ne peut pas s'exprimer de façon explicite. Son calcul nécessite donc une approximation numérique (voir section 5.4.1). La figure 2.2 donne l'allure de la fonction F pour quelques valeurs de λ .

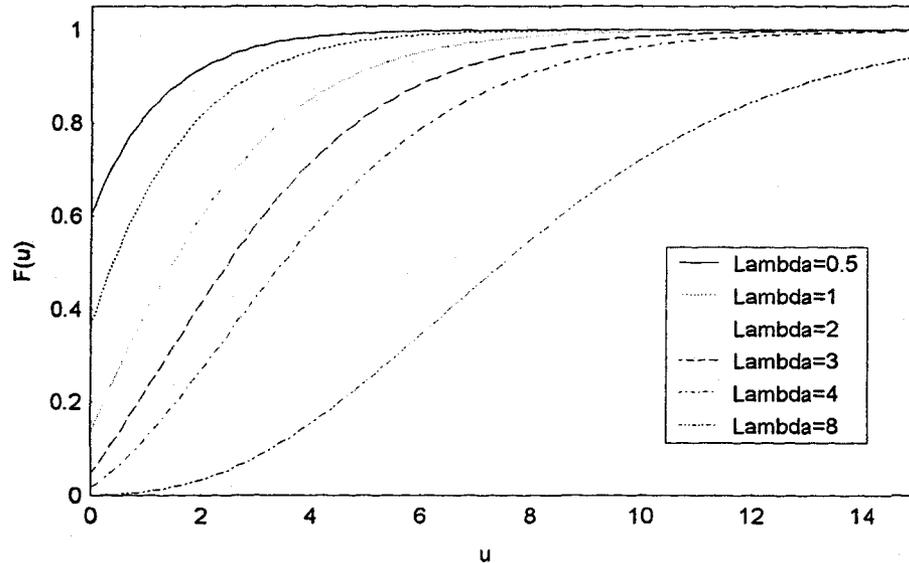


Figure 2.2. Fonction de répartition.

2.3 Propriétés statistiques

2.3.1 Fonction caractéristique

Pour la loi des fuites, la fonction caractéristique s'évalue par l'intégrale suivante :

$$\begin{aligned} \Phi(t) &= e^{it0} P(X=0) + \int_0^{\infty} e^{itx} g(x) dx \\ &= e^{it0} P(X=0) + \int_0^{\infty} e^{it\beta u} f(u) du \\ &= e^{-\lambda} + \lambda e^{-\lambda} \int_0^{\infty} e^{(\beta it - 1)u} \left(\sum_{k=0}^{\infty} \frac{(\lambda u)^k}{k!(k+1)!} \right) du \end{aligned} \quad (2.14)$$

Pour calculer cette série terme à terme, on pose $z = (1-\beta it) u$ et on obtient :

$$\int_0^{\infty} e^{(\beta it - 1)u} u^k du = \int_0^{\infty} e^{-z} \frac{z^k}{(1-it)^{k+1}} dz = \frac{1}{(1-it)^{k+1}} (1 - e^{-z} \sum_{j=0}^{\infty} \frac{z^j}{j!}) \quad (2.15)$$

en utilisant l'équation (2.12).

En développant les calculs, on a finalement que :

$$\begin{aligned} \Phi(t) &= e^{-\lambda} + e^{-\lambda} \sum_{i=1}^{\infty} \frac{1}{i!} \left(\frac{\lambda}{1-\beta it} \right)^i \\ &= e^{-\lambda} e^{\frac{\lambda}{1-\beta it}} \\ &= \exp\left(\frac{it\beta\lambda}{1-it\beta}\right) \end{aligned} \quad (2.16)$$

2.3.2 Cumulants, moments et coefficients

- Les cumulants K_n correspondent aux coefficients du développement en série entière du logarithme de la fonction caractéristique :

$$\ln \Phi(t) = \frac{it\beta\lambda}{1-it\beta} = it\beta\lambda(1+it\beta+\dots+(it\beta)^n+\dots) = \sum_{n=1}^{\infty} \frac{(it)^n}{n!} (n! \lambda \beta^n) \quad (2.17)$$

On en déduit que :

$$\begin{aligned} K_1 &= \beta\lambda \\ K_2 &= 2\beta^2\lambda \\ K_3 &= 6\beta^3\lambda \\ K_4 &= 24\beta^4\lambda \end{aligned} \quad (2.18)$$

- Les moments centrés μ_n s'expriment facilement à partir des cumulants. En effet, on a que:

$$\begin{aligned} E\{X\} &= \mu = K_1 = \beta\lambda \\ \text{Var}\{X\} &= \mu_2 = K_2 = 2\beta^2\lambda \\ \mu_3 &= K_3 = 6\beta^3\lambda \\ \mu_4 &= K_4 + 3\mu_2^2 = 12\beta^4\lambda(2+\lambda) \end{aligned} \quad (2.19)$$

- Le coefficient d'asymétrie s'exprime alors de la façon suivante :

$$C_s = \frac{\mu_3}{\mu_2^{3/2}} = \frac{3}{\sqrt{2\lambda}} \quad (2.20)$$

La distribution présente donc une asymétrie positive, qui diminue lorsque λ augmente.

Le coefficient d'aplatissement est :

$$C_k = \frac{\mu_4}{\mu_2^2} = 3 + \frac{6}{\lambda} \quad (2.21)$$

La fonction de densité de la loi des fuites possède des extrémités plus lourdes que la loi normale. Toutefois, le poids des extrémités diminue à mesure que λ augmente et à la limite ($\lambda \rightarrow \infty$), il correspond au poids des extrémités de la loi normale.

Le coefficient de variation est :

$$C_v = \frac{\sqrt{\mu_2}}{\mu_1} = \sqrt{\frac{2}{\lambda}} \quad (2.22)$$

2.4 Etude asymptotique lorsque λ tend vers l'infini

Effectuons le changement de variable suivant :

$$W = \frac{X - E\{X\}}{\sqrt{\text{Var}\{X\}}} = \frac{X - \lambda\beta}{\sqrt{2\beta^2\lambda}} \quad (2.23)$$

et calculons la fonction caractéristique de W .

$$\begin{aligned} \Phi_w(t) &= E\{e^{itw}\} \\ &= E\left\{\exp\left(\frac{itx}{\sqrt{2\beta^2\lambda}} - \frac{i\lambda\beta}{\sqrt{2\beta^2\lambda}}\right)\right\} \\ &= \Phi_x\left(\frac{t}{\sqrt{2\beta^2\lambda}}\right) \exp\left(-\frac{i\lambda\beta}{\sqrt{2\beta^2\lambda}}\right) \end{aligned} \quad (2.24)$$

L'expression de Φ_x étant donnée par (2.16), on trouve finalement que :

$$\Phi_w(t) = \exp\left(\frac{-t^2}{2 - it\sqrt{\frac{2}{\lambda}}}\right) \quad (2.25)$$

Lorsque λ tend vers l'infini, la fonction caractéristique de W tend vers $\exp(-t^2/2)$ qui correspond à la fonction caractéristique de la loi normale centrée réduite; la loi limite de la loi des fuites pour λ tendant vers l'infini est donc une loi normale. Cela est conforme à ce qui a été observé à la section 2.3.2, à savoir que lorsque λ tend vers l'infini, le coefficient d'aplatissement de la loi des fuites tend vers 3.

En pratique, on peut confondre la loi des fuites avec la loi normale pour $\lambda \geq 15$.

2.5 Comparaison avec la loi de Pearson

Soit Y une variable aléatoire distribuée selon une loi exponentielle. Nous avons vu au paragraphe 2.1 que la somme d'un nombre fixé n de variables aléatoires Y suit une loi Gamma tandis que la somme d'un nombre aléatoire N de variables aléatoires Y suit une loi des fuites. Il est donc intéressant de comparer les allures de ces deux distributions.

La densité d'une loi Gamma, de paramètre de forme γ et de paramètre d'échelle α s'exprime de la façon suivante :

$$f(z) = \frac{1}{\alpha\Gamma(\gamma)} \left(\frac{z}{\alpha}\right)^{\gamma-1} \exp\left(-\frac{z}{\alpha}\right) \quad (2.26)$$

Les premiers moments centrés s'écrivent :

$$\begin{aligned} E\{X\} &= \alpha\gamma \\ \text{Var}\{X\} &= \alpha^2\gamma \\ \mu_3 &= 2\alpha^3\gamma \\ \mu_4 &= 3\alpha^4\gamma(2 + \gamma) \end{aligned} \quad (2.27)$$

Ils ont une forme analogue à celle des moments de la loi des fuites, exprimés par les équations (2.19).

A l'aide de ces moments, nous pouvons calculer les principaux coefficients :

$$C_s = \frac{2}{\sqrt{\gamma}} \quad (2.28)$$

$$C_k = 3 + \frac{6}{\gamma} \quad (2.29)$$

$$C_v = \frac{1}{\sqrt{\gamma}} \quad (2.30)$$

Le tableau 2.1 compare l'allure des densités de la loi des fuites et de la loi Gamma à coefficient de variation égal (c'est-à-dire pour $\gamma = \lambda / 2$), en donnant pour chacune d'elles l'expression des coefficients d'asymétrie et d'aplatissement.

Tableau 2.1. Tableau comparatif des coefficients de la loi des fuites et de la loi Gamma

	LOI DE PEARSON	LOI DES FUTES
C_S	$\frac{2\sqrt{2}}{\sqrt{\lambda}}$	$\frac{3}{\sqrt{2\lambda}}$
C_K	$3 + \frac{12}{\lambda}$	$3 + \frac{6}{\lambda}$

La loi des fuites est donc plus symétrique et possède des extrémités plus légères que celles de la loi Gamma pour un même coefficient de variation. Toutes les deux ont pour limite la loi normale lorsque le paramètre de forme λ tend vers l'infini.

3 ESTIMATION DES PARAMETRES

Les méthodes d'estimation des paramètres les plus utilisées sont la méthode du maximum de vraisemblance et la méthode des moments, la première à cause de ses propriétés asymptotiques intéressantes, la seconde pour sa simplicité.

En ce qui concerne la loi des fuites, d'autres méthodes ont été proposées, faisant intervenir le nombre n_0 de valeurs nulles dans l'échantillon, ainsi qu'une méthode utilisant la moyenne des racines carrées des n observations.

Nous présentons dans ce chapitre les développements et les résultats menant aux estimateurs des paramètres de la loi des fuites pour la méthode du maximum de vraisemblance et la méthode des moments ainsi que les calculs des variances et covariances asymptotiques des estimateurs obtenus. Nous exposons ensuite brièvement les autres méthodes d'estimation en donnant les principaux résultats.

3.1 Méthode du maximum de vraisemblance

La méthode du maximum de vraisemblance consiste à déterminer les paramètres de la loi de façon à rendre l'échantillon le plus vraisemblable possible.

Soient X_1, X_2, \dots, X_n n variables aléatoires indépendantes que l'on suppose identiquement distribuées selon une loi donnée de densité $f(x, \underline{\theta})$. Soient x_1, x_2, \dots, x_n les réalisations de ces variables aléatoires qui forment l'échantillon observé. La probabilité que les événements $\{X_1 = x_1\}, \{X_2 = x_2\}, \dots, \{X_n = x_n\}$ se réalisent simultanément est donnée par la fonction de densité jointe, qui est définie comme le produit des densités de chaque variable aléatoire X_1, X_2, \dots, X_n et s'exprime donc de la façon suivante :

$$f(x_1, x_2, \dots, x_n) = \prod_{i=1}^n f(x_i, \underline{\theta}) \quad (3.1)$$

Cette probabilité peut être considérée comme une fonction des paramètres $\underline{\theta} = (\theta_1, \theta_2, \dots, \theta_p)$. On l'appelle alors fonction de vraisemblance et elle est notée $L(\theta_1, \theta_2, \dots, \theta_p)$. Supposons que pour l'échantillon observé x_1, x_2, \dots, x_n on peut déterminer

p fonctions de ces observations $\hat{\theta}_1(x_1, x_2, \dots, x_n), \dots, \hat{\theta}_p(x_1, x_2, \dots, x_n)$, telles que lorsque chacun des paramètres $\theta_1, \theta_2, \dots, \theta_p$ est remplacé par sa fonction correspondante, la fonction L soit maximum. Alors les statistiques $\hat{\theta}_1(x_1, x_2, \dots, x_n), \dots, \hat{\theta}_p(x_1, x_2, \dots, x_n)$ sont les estimateurs du maximum de vraisemblance de $\theta_1, \theta_2, \dots, \theta_p$ respectivement.

En pratique, pour des raisons de simplicité, on maximise plutôt le logarithme de la fonction L . Les estimateurs du maximum de vraisemblance sont alors obtenus par différenciation puisque la fonction de vraisemblance logarithmique $\ln L$ atteint son maximum lorsque toutes les dérivées partielles par rapport à chacun des paramètres sont nulles. Ainsi, il suffit de résoudre le système à p équations et p inconnues suivant :

$$\frac{\partial \ln L(\theta_1, \dots, \theta_p)}{\partial \theta_j} = 0, \quad j = 1, \dots, p \quad (3.2)$$

Remarquons que la fonction de vraisemblance L et la fonction de vraisemblance logarithmique $\ln L$ atteignent bien leur maximum aux mêmes valeurs de $\theta_1, \theta_2, \dots, \theta_p$ puisque :

$$\frac{\partial \ln L}{\partial \theta_j} = \frac{1}{L} \frac{\partial L}{\partial \theta_j} \quad (3.3)$$

3.1.1 Fonction de vraisemblance

La fonction de vraisemblance L est égale au produit des densités de probabilité :

$$L = [F(0)]^{n_0} \prod_{i=1}^{n-n_0} f(x_i) = e^{-n\lambda} \left(\frac{\lambda}{\beta}\right)^{\frac{n-n_0}{2}} \prod_{i=1}^{n-n_0} \exp\left(\frac{-x_i}{\beta}\right) \frac{I_1\left(2\sqrt{\frac{\lambda x_i}{\beta}}\right)}{\sqrt{x_i}} \quad (3.4)$$

où n_0 représente le nombre de valeurs $x_i = 0$ et n la taille de l'échantillon. La fonction de vraisemblance logarithmique est obtenue en appliquant la transformation \log à la fonction L :

$$\ln L = -\lambda n + \frac{(n-n_0)}{2} \ln\left(\frac{\lambda}{\beta}\right) - \sum_{i=1}^{n-n_0} \frac{x_i}{\beta} - \frac{1}{2} \sum_{i=1}^{n-n_0} \ln x_i + \sum_{i=1}^{n-n_0} \ln I_1\left(2\sqrt{\frac{\lambda x_i}{\beta}}\right) \quad (3.5)$$

3.1.2 Système d'équations

Pour maximiser la fonction de vraisemblance, il faut annuler les dérivées partielles de $\ln L$ par rapport à chacun des paramètres.

En remarquant que

$$\bar{x} = \frac{1}{n} \sum_{i=0}^{n-n_0} x_i, \quad (3.6)$$

on obtient le système d'équations suivant:

$$\frac{\partial \ln L}{\partial \lambda} = -n + \frac{n-n_0}{2\lambda} + \sum_{i=1}^{n-n_0} \frac{\sqrt{x_i}}{\lambda\beta} \frac{I_1'(2\sqrt{\frac{\lambda x_i}{\beta}})}{I_1(2\sqrt{\frac{\lambda x_i}{\beta}})} = 0 \quad (3.7)$$

$$\frac{\partial \ln L}{\partial \beta} = -\frac{n-n_0}{2\beta} + \frac{n\bar{x}}{\beta^2} - \frac{1}{\beta} \sum_{i=1}^{n-n_0} \frac{\sqrt{\lambda x_i}}{\beta} \frac{I_1'(2\sqrt{\frac{\lambda x_i}{\beta}})}{I_1(2\sqrt{\frac{\lambda x_i}{\beta}})} = 0 \quad (3.8)$$

où I_1' désigne la dérivée de la fonction de Bessel $I_1(z)$ par rapport à son argument z .

En combinant ces deux équations, on peut écrire que :

$$\bar{x} = \lambda\beta \quad (3.9)$$

D'autre part, les fonctions de Bessel ont la propriété suivante (cf. Abramowitz et Stegun (1970), p.376):

$$I_1'(z) = I_0(z) - \frac{1}{z} I_1(z) \quad (3.10)$$

où I_0 désigne la fonction de Bessel d'ordre 0 définie par :

$$I_0(z) = \frac{1}{\pi} \int_0^\pi \exp(z \cos \theta) d\theta = \sum_{i=0}^{\infty} \frac{(\frac{z}{2})^{2i}}{(i!)^2} \quad (3.11)$$

En utilisant cette propriété dans l'équation (3.7), on obtient alors :

$$\sum_{i=1}^{n-n_0} \frac{I_0(2\lambda\sqrt{\frac{x_i}{\bar{x}}})}{I_1(2\lambda\sqrt{\frac{x_i}{\bar{x}}})} \sqrt{x_i} - n\sqrt{\bar{x}} = 0 \quad (3.12)$$

L'équation (3.12) est de la forme $h(\lambda) = 0$ et on peut montrer que cette fonction h est décroissante en λ pour $\lambda > 0$ et qu'elle tend vers $+\infty$ en 0 et vers $K < 0$ en $+\infty$. Il existe donc une et une seule valeur $\hat{\lambda}$ du paramètre λ , positive, qui vérifie (3.12). On estimera ensuite le paramètre β par :

$$\hat{\beta} = \frac{\bar{x}}{\hat{\lambda}} \quad (3.13)$$

La détermination de $\hat{\lambda}$ requiert l'emploi d'une méthode numérique puisque l'équation (3.12) ne peut se résoudre explicitement.

3.1.3 Méthodes numériques

Pour résoudre l'équation (3.12), on peut rechercher la racine $\hat{\lambda}$ par dichotomie dans l'intervalle $[10^{-4} ; 5\hat{\lambda}_m]$ où $\hat{\lambda}_m$ désigne l'estimation du paramètre λ donnée par la méthode des moments. On peut aussi utiliser la méthode de Newton-Raphson, dont le principe général est décrit dans l'annexe A, en prenant $\hat{\lambda}_m$ comme valeur initiale; pour calculer la fonction $h'(\lambda)$, on se sert de la propriété des fonctions de Bessel (3.10) et on obtient l'expression suivante:

$$h'(\lambda) = 2n\sqrt{\bar{x}} - \frac{2}{\sqrt{\bar{x}}} \sum_{i=1}^{n-n_0} x_i \frac{I_0^2(2\lambda\sqrt{\frac{x_i}{\bar{x}}})}{I_1^2(2\lambda\sqrt{\frac{x_i}{\bar{x}}})} + \frac{1}{\lambda} \sum_{i=1}^{n-n_0} \sqrt{x_i} \frac{I_0(2\lambda\sqrt{\frac{x_i}{\bar{x}}})}{I_1(2\lambda\sqrt{\frac{x_i}{\bar{x}}})} \quad (3.14)$$

3.1.4 Variances et covariance asymptotiques

Les propriétés des estimateurs du maximum de vraisemblance, lorsque la taille d'échantillon est grande, sont bien connues. Ils sont convergents, asymptotiquement non-biaisés et asymptotiquement efficaces. De plus, comme la fonction de vraisemblance $L(\lambda, \beta)$ admet un seul maximum, les variables aléatoires $\sqrt{n}(\hat{\lambda} - \lambda)$ et $\sqrt{n}(\hat{\beta} - \beta)$ sont distribuées selon une loi normale bidimensionnelle de moyenne nulle et de matrice des variances et covariance Σ dont les éléments correspondent à ceux de l'inverse de la matrice d'information de Fisher I_f , donnée par :

$$I_f = \begin{bmatrix} -E\left(\frac{\partial^2 \ln L}{\partial \lambda^2}\right) & -E\left(\frac{\partial^2 \ln L}{\partial \lambda \partial \beta}\right) \\ -E\left(\frac{\partial^2 \ln L}{\partial \lambda \partial \beta}\right) & -E\left(\frac{\partial^2 \ln L}{\partial \beta^2}\right) \end{bmatrix} \quad (3.15)$$

Ainsi,

$$\Sigma = \frac{1}{n} I_f^{-1} = \begin{bmatrix} \text{Var}\{\hat{\lambda}\} & \text{Cov}\{\hat{\lambda}, \hat{\beta}\} \\ \text{Cov}\{\hat{\lambda}, \hat{\beta}\} & \text{Var}\{\hat{\beta}\} \end{bmatrix} \quad (3.16)$$

Les éléments de la matrice de Fisher (équation (3.15)) pour la loi des fuites peuvent être exprimés sous deux formes différentes.

La première est celle donnée par Babusiaux (1969). Les éléments de I_f s'expriment dans ce cas comme suit :

$$I_f(1,1) = n \left(-1 + \frac{1}{\lambda} - \frac{1}{4\lambda^2} + \frac{e^{-\lambda}}{4\lambda^2} + \frac{e^{-\lambda}}{4\lambda^2} J(\lambda) \right) \quad (3.17)$$

$$I_f(1,2) = I_f(2,1) = \frac{n}{\beta} \left(\lambda + \frac{1}{4\lambda} - \frac{e^{-\lambda}}{4\lambda} - \frac{e^{-\lambda}}{4} J(\lambda) \right) \quad (3.18)$$

$$I_f(2,2) = \frac{n}{\beta^2} \left(-\lambda^2 + \lambda - \frac{1}{4} + \frac{e^{-\lambda}}{4} + \frac{\lambda e^{-\lambda}}{4} J(\lambda) \right) \quad (3.19)$$

où $J(\lambda)$ est donnée par :

$$J(\lambda) = \int_0^{\infty} e^{-u} \frac{\left[\sum_{i=0}^{\infty} \frac{(2i+1)(\lambda u)^i}{i!(i+1)!} \right]^2}{\sum_{i=0}^{\infty} \frac{(\lambda u)^i}{i!(i+1)!}} du \quad (3.20)$$

Puisque l'expression de la fonction $J(\lambda)$ qui intervient est complexe, le calcul par voie numérique risque d'être peu précis. Les détails sont développés dans l'annexe B-1.

La seconde forme possible, dont les calculs sont effectués dans l'annexe B-2, fait intervenir les fonctions de Bessel; si celles-ci sont correctement programmées¹, on la privilégiera par rapport à la précédente. Dans ce cas, les éléments de la matrice d'information de Fisher I_f s'expriment de la manière suivante :

¹ Les logiciels mathématiques (Matlab, Mathematica, etc.) ainsi que plusieurs bibliothèques de routines mathématiques (ISML par exemple) offrent des programmes permettant d'évaluer les fonctions de Bessel.

$$I_f(1,1) = n \left(-1 + \frac{e^{-\lambda}}{\sqrt{\lambda}} \int_0^{\infty} e^{-u} \sqrt{u} \frac{I_0^2(2\sqrt{\lambda u})}{I_1(2\sqrt{\lambda u})} du \right) \quad (3.21)$$

$$I_f(1,2) = I_f(2,1) = \frac{n}{\beta} (\lambda + 1 - e^{-\lambda} \sqrt{\lambda} \int_0^{\infty} e^{-u} \sqrt{u} \frac{I_0^2(2\sqrt{\lambda u})}{I_1(2\sqrt{\lambda u})} du) \quad (3.22)$$

$$I_f(2,2) = \frac{\lambda n}{\beta^2} (-\lambda + e^{-\lambda} \sqrt{\lambda} \int_0^{\infty} e^{-u} \sqrt{u} \frac{I_0^2(2\sqrt{\lambda u})}{I_1(2\sqrt{\lambda u})} du) \quad (3.23)$$

On peut aussi utiliser la matrice de Fisher observée, définie par :

$$I_{f(ob)} = \begin{pmatrix} \frac{\partial^2 \ln L}{\partial \lambda^2} & \frac{\partial^2 \ln L}{\partial \lambda \partial \beta} \\ \frac{\partial^2 \ln L}{\partial \lambda \partial \beta} & \frac{\partial^2 \ln L}{\partial \beta^2} \end{pmatrix} \quad (3.24)$$

et qui s'écrit ici :

$$I_{f(ob)}(1,1) = \frac{1}{\lambda \beta} \left[\sum_{i=1}^{n-n_0} x_i \frac{I_0^2(2\sqrt{\frac{\lambda x_i}{\beta}})}{I_1^2(2\sqrt{\frac{\lambda x_i}{\beta}})} - \sum_{i=1}^{n-n_0} x_i \right] \quad (3.25)$$

$$I_{f(ob)}(1,2) = I_{f(ob)}(2,1) = \frac{1}{\beta} \left[\frac{1}{\beta} \sum_{i=1}^{n-n_0} x_i - \frac{1}{\beta} \sum_{i=1}^{n-n_0} x_i \frac{I_0^2(2\sqrt{\frac{\lambda x_i}{\beta}})}{I_1^2(2\sqrt{\frac{\lambda x_i}{\beta}})} + \frac{1}{\sqrt{\lambda} \beta} \sum_{i=1}^{n-n_0} \sqrt{x_i} \frac{I_0(2\sqrt{\frac{\lambda x_i}{\beta}})}{I_1(2\sqrt{\frac{\lambda x_i}{\beta}})} \right] \quad (3.26)$$

$$I_{f(ob)}(2,2) = \frac{1}{\beta^2} \left[\frac{2-\lambda}{\beta} \sum_{i=1}^{n-n_0} x_i + \frac{\lambda}{\beta} \sum_{i=1}^{n-n_0} x_i \frac{I_0^2(2\sqrt{\frac{\lambda x_i}{\beta}})}{I_1^2(2\sqrt{\frac{\lambda x_i}{\beta}})} - 2\sqrt{\frac{\lambda}{\beta}} \sum_{i=1}^{n-n_0} \sqrt{x_i} \frac{I_0(2\sqrt{\frac{\lambda x_i}{\beta}})}{I_1(2\sqrt{\frac{\lambda x_i}{\beta}})} \right] \quad (3.27)$$

Ces coefficients sont plus simples à obtenir puisqu'on évite ainsi le calcul de l'espérance mathématique. Cette approche est toutefois plus approximative, c'est pourquoi on l'évitera.

3.2 Méthode des moments

La méthode des moments est généralement utilisée parce qu'elle est très simple à appliquer. On s'en sert aussi pour initialiser les paramètres dans le cadre du processus itératif de la méthode du maximum de vraisemblance.

La méthode des moments consiste à résoudre le système d'équations formé en égalant les moments théoriques $\mu_r(X)$ à ceux de l'échantillon m'_r . Pour obtenir une solution unique, on doit disposer d'autant d'équations indépendantes que de paramètres à estimer. En pratique, on utilise les moments d'ordre r petit. Ainsi, si on considère n variables aléatoires X_1, X_2, \dots, X_n indépendantes et identiquement distribuées selon une loi $F(x, \underline{\theta})$ donnée dépendant de p paramètres $\underline{\theta} = (\theta_1, \theta_2, \dots, \theta_p)$, et x_1, x_2, \dots, x_n les réalisations correspondantes, les solutions $\hat{\theta}_1, \hat{\theta}_2, \dots, \hat{\theta}_p$ du système à p équations et p inconnues suivant:

$$\mu_j(X) = m_j, \quad j = 1, 2, \dots, p \quad (3.28)$$

sont des estimateurs des paramètres $\theta_1, \theta_2, \dots, \theta_p$ obtenus par la méthode des moments.

3.2.1 Système d'équations et solution

Puisque la loi des fuites est une loi à deux paramètres, la méthode des moments consiste ici à égaliser les deux premiers moments (moyenne et variance) de l'échantillon aux moments théoriques correspondants :

$$\bar{x} = \mu(X) \quad (3.29)$$

$$s^2 = \mu_2(X) \quad (3.30)$$

avec :

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i \quad (3.31)$$

$$s^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2 \quad (3.32)$$

A partir des expressions des moments données par les équations (2.19), on déduit le système à 2 équations et 2 inconnues suivant :

$$\bar{x} = \lambda\beta \quad (3.33)$$

$$s^2 = 2\lambda\beta^2 \quad (3.34)$$

ce qui nous donne comme estimateurs respectifs de λ et β :

$$\hat{\lambda} = \frac{2\bar{x}^2}{s^2} \quad (3.35)$$

$$\hat{\beta} = \frac{s^2}{2\bar{x}} \quad (3.36)$$

3.2.2 Variances et covariance asymptotiques

Les variances et covariances asymptotiques des estimateurs déduits de la méthode des moments sont reliées à celles des moments de l'échantillon (voir Bobée et Ashkar (1991), Appendix D). Plus précisément, on a :

$$Var(\hat{\lambda}) = \left(\frac{\partial\hat{\lambda}}{\partial s^2}\right)^2 Var(s^2) + \left(\frac{\partial\hat{\lambda}}{\partial\bar{x}}\right)^2 Var(\bar{x}) + 2\left(\frac{\partial\hat{\lambda}}{\partial s^2}\right)\left(\frac{\partial\hat{\lambda}}{\partial\bar{x}}\right)Cov(s^2, \bar{x}) \quad (3.37)$$

$$Var(\hat{\beta}) = \left(\frac{\partial\hat{\beta}}{\partial s^2}\right)^2 Var(s^2) + \left(\frac{\partial\hat{\beta}}{\partial\bar{x}}\right)^2 Var(\bar{x}) + 2\left(\frac{\partial\hat{\beta}}{\partial s^2}\right)\left(\frac{\partial\hat{\beta}}{\partial\bar{x}}\right)Cov(s^2, \bar{x}) \quad (3.38)$$

$$Cov(\hat{\lambda}, \hat{\beta}) = \left(\frac{\partial\hat{\lambda}}{\partial s^2}\right)\left(\frac{\partial\hat{\beta}}{\partial s^2}\right)Var(s^2) + \left(\frac{\partial\hat{\lambda}}{\partial\bar{x}}\right)\left(\frac{\partial\hat{\beta}}{\partial\bar{x}}\right)Var(\bar{x}) + \left(\frac{\partial\hat{\lambda}}{\partial s^2}\frac{\partial\hat{\beta}}{\partial\bar{x}} + \frac{\partial\hat{\lambda}}{\partial\bar{x}}\frac{\partial\hat{\beta}}{\partial s^2}\right)Cov(s^2, \bar{x}) \quad (3.39)$$

- Calculons d'abord les variances et covariance asymptotiques des moments de l'échantillon.

Lorsque n est assez grand (voir Kendall et Stuart (1987), Chapitre 10), elles sont données par :

$$Var(s^2) = \frac{\mu_4 - \mu_2^2}{n} \quad (3.40)$$

$$Var(\bar{x}) = \frac{\mu_2}{n} \quad (3.41)$$

$$\text{Cov}(s^2, \bar{x}) = \frac{\mu_3}{n} \quad (3.42)$$

Les moments centrés μ_2 , μ_3 et μ_4 ont été calculés au chapitre 2, paragraphe 2.3.2, nous rappelons ici leur expression :

$$\begin{aligned} \mu_2 &= 2\beta^2\lambda \\ \mu_3 &= 6\beta^3\lambda \\ \mu_4 &= 12\beta^4\lambda(2+\lambda) \end{aligned} \quad (3.43)$$

Nous obtenons donc :

$$\text{Var}(s^2) = \frac{8\lambda\beta^4(3+\lambda)}{n} \quad (3.44)$$

$$\text{Var}(\bar{x}) = \frac{2\lambda\beta^2}{n} \quad (3.45)$$

$$\text{Cov}(s^2, \bar{x}) = \frac{6\lambda\beta^3}{n} \quad (3.46)$$

- Calculons $\text{Var}(\hat{\lambda})$, $\text{Var}(\hat{\beta})$, $\text{Cov}(\hat{\lambda}, \hat{\beta})$:

En reportant les expressions des variances et covariance asymptotiques des moments de l'échantillon dans les équations (3.37), (3.38) et (3.39), nous obtenons :

$$\text{Var}(\hat{\lambda}) = \frac{2\lambda(\lambda+1)}{n} \quad (3.47)$$

$$\text{Var}(\hat{\beta}) = \frac{2\beta^2(\lambda+1)}{n\lambda} \quad (3.48)$$

$$\text{Cov}(\hat{\lambda}, \hat{\beta}) = -\frac{\beta(2\lambda+1)}{n} \quad (3.49)$$

En prenant comme définition de l'efficacité la définition classique que nous rappelons dans l'annexe C, Babuziaux (1969) montre que la méthode des moments est peu efficace pour les faibles valeurs de λ , c'est-à-dire lorsque les valeurs observées sont proches de zéro. Il s'est

donc intéressé plus particulièrement au cas où les observations sont de faible valeur et a proposé d'utiliser dans les estimations le nombre n_o de valeurs nulles.

3.3 Autres méthodes

3.3.1 Méthodes faisant intervenir n_o/n

Puisque le nombre de valeurs nulles suit une loi binomiale de probabilité $p = e^{-\lambda}$ (voir Annexe B-1, équation (B.20)), il semble naturel d'estimer le paramètre λ par

$$\hat{\lambda} = Ln\left(\frac{n_o}{n}\right) \quad (3.50)$$

Or, comme \bar{x} est l'estimateur du maximum de vraisemblance de $\lambda\beta$ (voir équation (3.9)), on peut, selon (3.50), estimer β par :

$$\hat{\beta} = \frac{\bar{x}}{\lambda} = \frac{\bar{x}}{Ln\left(\frac{n_o}{n}\right)} \quad (3.51)$$

Les variances et covariance asymptotiques de ces estimateurs ont été déterminées par Babuziaux (1969) :

$$Var(\hat{\lambda}) = \frac{e^{\lambda} - 1}{n} \quad (3.52)$$

$$Var(\hat{\beta}) = \frac{\beta^2 (e^{\lambda} - 1)}{n\lambda^2} \quad (3.53)$$

$$Cov(\hat{\lambda}, \hat{\beta}) = \frac{\beta}{n} \left(1 - \frac{1}{\lambda} (e^{\lambda} - 1)\right) \quad (3.54)$$

Ce dernier a constaté que pour pouvoir utiliser cette approche, il faut un nombre non négligeable de valeurs nulles dans l'échantillon car, si ce mode d'estimation est bon pour les valeurs faibles de λ , son efficacité décroît rapidement lorsque λ augmente. Il a donc été amené à chercher de nouveaux estimateurs, combinaisons linéaires des estimateurs

précédents et des estimateurs issus de la méthode des moments. Les coefficients ont été déterminés de façon à obtenir une efficacité maximum, et le calcul donne :

$$\hat{\lambda} = (1 - (2 - \sqrt{2})\sqrt{\frac{n_0}{n}})\left(\frac{2\bar{x}^2}{s^2} - \text{Ln}\left(\frac{n_0}{n}\right)\right) + \text{Ln}\left(\frac{n_0}{n}\right) \quad (3.55)$$

$$\hat{\beta} = (1 - (2 - \sqrt{2})\sqrt{\frac{n_0}{n}})\left(\frac{s^2}{2\bar{x}} - \frac{\bar{x}}{\text{Ln}\left(\frac{n_0}{n}\right)}\right) + \frac{\bar{x}}{\text{Ln}\left(\frac{n_0}{n}\right)} \quad (3.56)$$

Pour les variances et covariance asymptotiques de ces estimateurs, nous reprenons ici encore une fois les résultats de Babuziaux (1969). Il obtient :

$$\text{Var}(\hat{\lambda}) = \frac{1}{n} \left[(p-1)^2 (e^\lambda - 1) + p^2 \lambda^2 + 2p\lambda + p\lambda^2 \right] \quad (3.57)$$

$$\text{Var}(\hat{\beta}) = \frac{\beta^2}{n\lambda} \left[2p + p^2 \lambda + p\lambda + \frac{1}{\lambda} (p-1)^2 (e^\lambda - 1) \right] \quad (3.58)$$

$$\text{Cov}(\hat{\lambda}, \hat{\beta}) = \frac{\beta}{n} \left[1 - 2p - p^2 \lambda - p\lambda - \frac{1}{\lambda} (p-1)^2 (e^\lambda - 1) \right] \quad (3.59)$$

où

$$p = 1 - (2 - \sqrt{2})\sqrt{\frac{n_0}{n}} \quad (3.60)$$

3.3.2 Méthode utilisant la moyenne des racines carrées des observations

Nous avons vu au chapitre 2, paragraphe 2.4, que lorsque λ tend vers l'infini, la loi de x (loi des fuites) a pour limite une loi normale. On peut montrer que la loi de \sqrt{x} tend également vers une distribution normale et que cette convergence en loi est plus rapide. Babuziaux a donc proposé d'utiliser la méthode des moments appliquée à la variable transformée \sqrt{x} , son efficacité devant augmenter plus rapidement avec λ que si on l'applique à la variable d'origine x . Il constate finalement que si cette approche se révèle tout à fait valable quelle que soit la valeur de λ , les calculs à effectuer sont extrêmement complexes.



4 ESTIMATION DU QUANTILE x_T

La notion de quantile a déjà été abordée dans l'introduction de ce rapport, au chapitre 1. Par définition le quantile d'ordre p , x_p , est la valeur dont la probabilité au non-dépassement est p . Plus précisément, x_p est tel que

$$\text{Prob} \{ X \leq x_p \} = p \quad (4.1)$$

En hydrologie, on définira l'événement extrême x_T de période de retour T comme étant la valeur dont la probabilité au dépassement est $1/T$, où T est un intervalle de temps donné. Ainsi, x_T est tel que

$$\text{Prob} \{ X > x_T \} = 1/T \quad (4.2)$$

Nous présentons dans ce chapitre l'estimation des quantiles de la loi des fuites, le calcul de leur variance asymptotique ainsi que la détermination des intervalles de confiance associés.

4.1 Estimation de x_T

Nous avons vu au chapitre 1 qu'un estimateur du quantile théorique de période de retour T est donné par :

$$\hat{x}_T = F^{-1}(1 - 1/T; \hat{\theta}) \quad (1.1)$$

où $F(x, \hat{\theta})$ représente la fonction de répartition calculée avec les paramètres estimés $\hat{\theta}$.

Pour la loi des fuites on aura ainsi :

$$\hat{x}_T = F^{-1}(1 - 1/T; \hat{\lambda}, \hat{\beta}) \quad (4.3)$$

Cet estimateur sera celui du maximum de vraisemblance ou des moments selon la méthode utilisée pour déterminer $\hat{\lambda}$ et $\hat{\beta}$ (sections 3.1 et 3.2).

Comme on ne peut pas calculer F^{-1} , fonction inverse de la fonction de répartition, de façon explicite, il n'existe pas de méthode directe pour déterminer \hat{x}_T . On propose ici deux méthodes numériques, la méthode de Newton-Raphson (voir annexe-A) et une méthode par dichotomie.

Pour la méthode de Newton-Raphson, on se souviendra que la dérivée de la fonction de répartition F est égale à la densité. La fonction de répartition étant monotone (cf. figure (2.2)), l'équation à résoudre possède une solution unique. Le choix de la valeur initiale de la racine n'influe donc pas sur la solution comme c'est parfois le cas avec la méthode de Newton-Raphson.

La croissance monotone de la fonction de répartition F nous permet d'envisager une autre méthode d'approximation par itérations successives, basée sur une approche par dichotomie. On se donne une valeur initiale s_0 du quantile x_T et on calcule $\text{Prob}(X < s_0) = F(s_0)$. Si $F(s_0) > 1-1/T$ [respectivement $F(s_0) < 1-1/T$] alors on pose $s_1 = s_0 / 2$ [respectivement $s_1 = 2s_0$] -puis on calcule $F(s_1)$ et ainsi de suite. Lorsqu'on a $F(s_i) > 1-1/T$ et $F(s_{i+1}) < 1-1/T$ [respectivement $F(s_i) < 1-1/T$ et $F(s_{i+1}) > 1-1/T$], on prend alors $s_{i+2} = (s_i + s_{i+1})/2$ et on procède ensuite par dichotomie.

Pour les deux méthodes on pourra choisir comme valeur initiale la moyenne des observations \bar{x} qui est égale à $\hat{\lambda}\hat{\beta}$.

4.2 Variance asymptotique et intervalle de confiance

4.2.1 Variance asymptotique

La variance asymptotique est une mesure de la précision de l'estimateur.

Selon la théorie des grands nombres (théorème de la limite centrale, Bickel et Docksum, 1977), l'estimateur \hat{x}_T est distribué asymptotiquement selon une loi normale de moyenne x_T et de variance donnée par :

$$\text{Var}\{\hat{x}_T\} = \left(\frac{\partial x_T}{\partial \lambda}\right)^2 \text{Var}\{\hat{\lambda}\} + \left(\frac{\partial x_T}{\partial \beta}\right)^2 \text{Var}\{\hat{\beta}\} + 2\left(\frac{\partial x_T}{\partial \lambda}\right)\left(\frac{\partial x_T}{\partial \beta}\right)\text{Cov}\{\hat{\lambda}, \hat{\beta}\} \quad (4.4)$$

Les expressions de $\text{Var}\{\hat{\lambda}\}$, $\text{Var}\{\hat{\beta}\}$ et $\text{Cov}\{\hat{\lambda}, \hat{\beta}\}$ ont été établies au Chapitre 3 pour chaque méthode d'estimation des paramètres $\hat{\lambda}$ et $\hat{\beta}$.

Pour calculer les dérivées partielles du quantile par rapport à chacun des deux paramètres, on écrit la définition de x_T :

$$F(x_T) = 1 - \frac{1}{T} \quad (4.5)$$

sous la forme

$$e^{-\lambda} + \int_0^{x_T} g(x) dx = 1 - \frac{1}{T} \quad (4.6)$$

et on dérive cette dernière équation par rapport à chacun des paramètres en utilisant la formule de Leibniz. Les calculs sont développés dans l'annexe D et conduisent aux résultats suivants :

$$\frac{\partial x_T}{\partial \lambda} = \frac{\sqrt{\frac{\hat{\beta} \hat{x}_T}{\hat{\lambda}}} I_0\left(2\sqrt{\frac{\hat{\lambda} \hat{x}_T}{\hat{\beta}}}\right)}{I_1\left(2\sqrt{\frac{\hat{\lambda} \hat{x}_T}{\hat{\beta}}}\right)} \quad (4.7)$$

$$\frac{\partial x_T}{\partial \beta} = \frac{\hat{x}_T}{\hat{\beta}} \quad (4.8)$$

La variance asymptotique est particulièrement intéressante pour comparer deux estimations issues de méthodes différentes, mais elle demeure en pratique difficile à interpréter. Une approche permettant d'avoir une meilleure idée de l'exactitude de l'estimation est la construction d'un intervalle de confiance.

4.2.2 Intervalle de confiance

Un intervalle de confiance de niveau $100(1-\alpha)$ % pour la quantité x_T inconnue est un intervalle $[L_1, L_2]$ tel que :

$$Prob\{L_1 \leq x_T \leq L_2\} = 1-\alpha \quad (4.9)$$

où L_1 et L_2 sont des statistiques indépendantes de x_T . Cette approche nous permet donc d'obtenir un intervalle contenant la valeur théorique inconnue avec une probabilité de $(1-\alpha)$.

Nous avons vu au paragraphe précédent que l'estimateur \hat{x}_T est distribué asymptotiquement selon une loi normale de moyenne x_T et de variance $Var\{\hat{x}_T\}$. La statistique

$$S = \frac{\hat{x}_T - x_T}{\sqrt{\text{Var}\{\hat{x}_T\}}} \quad (4.10)$$

est alors distribuée asymptotiquement selon une loi normale centrée-réduite. On en déduit donc que

$$\text{Prob}\left\{z_{\frac{\alpha}{2}} \leq \frac{\hat{x}_T - x_T}{\sqrt{\text{Var}\{\hat{x}_T\}}} \leq z_{1-\frac{\alpha}{2}}\right\} = 1 - \alpha \quad (4.11)$$

où $z_{\alpha/2}$ et $z_{1-\alpha/2}$ sont respectivement les quantiles de probabilité au non-dépassement $\alpha/2$ et $1-\alpha/2$ de la loi normale centrée-réduite. En isolant x_T dans l'expression (4.11) et en remarquant qu'en vertu de la symétrie de la loi normale par rapport à zéro, $z_{\alpha/2} = -z_{1-\alpha/2}$, on obtient

$$\text{Prob}\left\{\hat{x}_T - z_{1-\frac{\alpha}{2}}\sqrt{\text{Var}\{\hat{x}_T\}} \leq x_T \leq \hat{x}_T + z_{1-\frac{\alpha}{2}}\sqrt{\text{Var}\{\hat{x}_T\}}\right\} = 1 - \alpha \quad (4.12)$$

L'intervalle de confiance de niveau $100(1-\alpha)\%$ pour le quantile théorique x_T de période de retour T est alors :

$$\left\{\hat{x}_T - z_{1-\frac{\alpha}{2}}\sqrt{\text{Var}\{\hat{x}_T\}} ; \hat{x}_T + z_{1-\frac{\alpha}{2}}\sqrt{\text{Var}\{\hat{x}_T\}}\right\} \quad (4.13)$$

Cet intervalle est dit intervalle de confiance asymptotique puisqu'il est construit à l'aide d'une approximation de la loi exacte de \hat{x}_T (loi asymptotique).

5 ALGORITHMIE

Ce chapitre a pour objet de fournir tous les éléments nécessaires à la programmation éventuelle de la loi des fuites dans le logiciel AJUSTE-II. Afin de permettre au lecteur de retrouver facilement l'origine des équations, celles-ci ont conservé dans ce chapitre le numéro qui leur était attribué dans la section où elles ont été calculées.

Nous rappelons donc ici les principales formules des chapitres précédents et discutons des méthodes numériques à employer au vu des résultats de quelques essais effectués à l'aide des logiciels Matlab et Mathematica.

On rappelle d'abord les deux expressions possibles pour la densité de la loi des fuites,

$$g(x) = \begin{cases} e^{-\lambda}, & x = 0 \\ \frac{\lambda}{\beta} \exp\left(-\lambda - \frac{x}{\beta}\right) \frac{I_1\left(2\sqrt{\frac{\lambda x}{\beta}}\right)}{\sqrt{\frac{\lambda x}{\beta}}}, & x > 0 \end{cases} \quad (2.7)$$

$$f(u) = \begin{cases} e^{-\lambda}, & u = 0 \\ \lambda \exp(-\lambda - u) \frac{I_1(2\sqrt{\lambda u})}{\sqrt{\lambda u}}, & u > 0 \end{cases} \quad (2.8)$$

où I_1 désigne la fonction de Bessel d'ordre 1.

5.1 Moments et coefficients

$$\begin{aligned} E\{X\} &= \mu = \beta\lambda \\ \text{Var}\{X\} &= \mu_2 = 2\beta^2\lambda \\ \mu_3 &= 6\beta^3\lambda \\ \mu_4 &= 12\beta^4\lambda(2 + \lambda) \end{aligned} \quad (2.19)$$

$$C_s = \frac{\mu_3}{\mu_2^{3/2}} = \frac{3}{\sqrt{2\lambda}} \quad (2.20)$$

$$C_k = \frac{\mu_4}{\mu_2^2} = 3 + \frac{6}{\lambda} \quad (2.21)$$

$$C_v = \frac{\sqrt{\mu_2}}{\mu_1} = \sqrt{\frac{2}{\lambda}} \quad (2.22)$$

5.2 Estimation des paramètres

5.2.1 Méthode du maximum de vraisemblance

$\hat{\lambda}$ est donné par la résolution de l'équation suivante :

$$h(\lambda) = \sum_{i=1}^{n-n_0} \frac{I_0(2\lambda\sqrt{\frac{x_i}{\bar{x}}})}{I_1(2\lambda\sqrt{\frac{x_i}{\bar{x}}})} \sqrt{x_i} - n\sqrt{\bar{x}} = 0 \quad (3.12)$$

La résolution peut se faire de deux façons, par la méthode de Newton-Raphson ou par dichotomie.

On estime ensuite le paramètre β par :

$$\hat{\beta} = \frac{\bar{x}}{\hat{\lambda}} \quad (3.13)$$

1. Méthode de Newton-Raphson (voir Annexe A)

On initialise $\hat{\lambda}$ à la valeur $\lambda_0 = \hat{\lambda}_m$ où $\hat{\lambda}_m$ désigne l'estimation du paramètre λ donnée par la méthode des moments.

On répète le processus suivant :

$$\lambda_i = \lambda_{i-1} - \frac{h(\lambda_{i-1})}{h'(\lambda_{i-1})} \quad (5.1)$$

avec $h'(\lambda)$ définie par

$$h'(\lambda) = 2n\sqrt{\bar{x}} - \frac{2}{\sqrt{\bar{x}}} \sum_{i=1}^{n-n_0} x_i \frac{I_0^2(2\lambda\sqrt{\frac{x_i}{\bar{x}}})}{I_1^2(2\lambda\sqrt{\frac{x_i}{\bar{x}}})} + \frac{1}{\lambda} \sum_{i=1}^{n-n_0} \sqrt{x_i} \frac{I_0(2\lambda\sqrt{\frac{x_i}{\bar{x}}})}{I_1(2\lambda\sqrt{\frac{x_i}{\bar{x}}})} \quad (3.14)$$

En programmant cette méthode dans le logiciel Matlab, nous avons constaté que la condition d'arrêt $|\lambda_N - \lambda_{N-1}| < 10^{-6}$ convient. Le programme est donné dans l'annexe E-1.

2. Méthode par dichotomie

On cherche la racine $\hat{\lambda}$ par dichotomie dans l'intervalle $[10^{-4} ; 5\hat{\lambda}_m]$ où $\hat{\lambda}_m$ désigne l'estimation du paramètre λ donnée par la méthode des moments. On arrêtera le processus dès que $|h(\lambda)| < 10^{-6}$

Le programme Matlab de cette approche est donné dans l'annexe E-1. Les résultats obtenus lors des essais sont identiques pour les deux méthodes.

5.2.2 Méthode des moments

Ici, la solution est directe (cf. section 3.2.1) et λ et β sont estimés respectivement par :

$$\hat{\lambda} = \frac{2\bar{x}^2}{s^2} \quad (3.35)$$

$$\hat{\beta} = \frac{s^2}{2\bar{x}} \quad (3.36)$$

5.2.3 Méthodes faisant intervenir n_0/n

1. Nous estimons λ et β par :

$$\hat{\lambda} = \text{Ln}\left(\frac{n_0}{n}\right) \quad (3.50)$$

$$\hat{\beta} = \frac{\bar{x}}{\text{Ln}\left(\frac{n_0}{n}\right)} \quad (3.51)$$

2. ou encore par :

$$\hat{\lambda} = (1 - (2 - \sqrt{2})\sqrt{\frac{n_0}{n}})\left(\frac{2\bar{x}^2}{s^2} - \text{Ln}\left(\frac{n_0}{n}\right)\right) + \text{Ln}\left(\frac{n_0}{n}\right) \quad (3.55)$$

$$\hat{\beta} = (1 - (2 - \sqrt{2})\sqrt{\frac{n_0}{n}})\left(\frac{s^2}{2\bar{x}} - \frac{\bar{x}}{\text{Ln}\left(\frac{n_0}{n}\right)}\right) + \frac{\bar{x}}{\text{Ln}\left(\frac{n_0}{n}\right)} \quad (3.56)$$

5.3 Variances et covariance asymptotiques des paramètres

5.3.1 Méthode du maximum de vraisemblance

On inverse la matrice de Fisher dont les éléments peuvent être exprimés de deux façons.

1. Expression à l'aide de la fonction $J(\lambda)$

$$I_f(1,1) = n \left(-1 + \frac{1}{\lambda} - \frac{1}{4\lambda^2} + \frac{e^{-\lambda}}{4\lambda^2} + \frac{e^{-\lambda}}{4\lambda^2} J(\lambda) \right) \quad (3.17)$$

$$I_f(1,2) = I_f(2,1) = \frac{n}{\beta} \left(\lambda + \frac{1}{4\lambda} - \frac{e^{-\lambda}}{4\lambda} - \frac{e^{-\lambda}}{4} J(\lambda) \right) \quad (3.18)$$

$$I_f(2,2) = \frac{n}{\beta^2} \left(-\lambda^2 + \lambda - \frac{1}{4} + \frac{e^{-\lambda}}{4} + \frac{\lambda e^{-\lambda}}{4} J(\lambda) \right) \quad (3.19)$$

où la fonction $J(\lambda)$ est définie par

$$J(\lambda) = \int_0^{\infty} j(\lambda, u) du = \int_0^{\infty} e^{-u} \frac{\left[\sum_{i=0}^{\infty} \frac{(2i+1)(\lambda u)^i}{i!(i+1)!} \right]^2}{\sum_{i=0}^{\infty} \frac{(\lambda u)^i}{i!(i+1)!}} du \quad (3.20)$$

Le calcul de la fonction $J(\lambda)$, dont les détails sont donnés dans l'annexe B-1, a été programmé dans Matlab et dans Mathematica (voir Annexe E-2).

- Dans Mathematica, nous pouvons exprimer la fonction $j(\lambda, u)$ au moyen d'une somme infinie, mais pour pouvoir ensuite l'intégrer à partir de $u=0$, il faut écrire les sommes sous la forme :

$$\sum_{i=0}^{\infty} \frac{(2i+1)(\lambda u)^i}{i!(i+1)!} = 1 + \sum_{i=1}^{\infty} \frac{(2i+1)(\lambda u)^i}{i!(i+1)!} \quad (5.2)$$

$$\sum_{i=0}^{\infty} \frac{(\lambda u)^i}{i!(i+1)!} = 1 + \sum_{i=1}^{\infty} \frac{(\lambda u)^i}{i!(i+1)!} \quad (5.3)$$

Comme il n'est pas possible d'intégrer jusqu'à l'infini, nous devons calculer

$$J(\lambda) \approx \int_0^M j(\lambda, u) du \quad (5.4)$$

On voudrait trouver une valeur de M qui satisfasse la condition suivante :

$$\int_M^{M+\varepsilon} j(\lambda, u) du \ll \int_0^M j(\lambda, u) du \quad (5.5)$$

Il est clair que M dépend de λ et on constate que lorsque λ est inférieur à 9,

$$\frac{j(\lambda, 40)}{\int_0^{40} j(\lambda, u) du} < 10^{-6} \quad (5.6)$$

La valeur $M = 40$ semble donc appropriée tant que $\lambda \leq 9$; pour des valeurs supérieures de λ , il faudra choisir $M > 40$.

- Dans Matlab, le calcul de la fonction $j(\lambda, u)$ n'est pas exact. Nous ne pouvons, en effet, programmer que des sommes finies. Nous avons choisi d'arrêter la sommation d'une série T lorsque le terme T_i devient suffisamment petit devant la somme des i premiers termes, c'est-à-dire lorsque :

$$\frac{T_i}{\sum_{p=0}^i T_p} < 10^{-6} \quad (5.7)$$

On intègre ensuite numériquement la fonction $j(\lambda, u)$: on initialise $u_0 = 0$ puis on répète :

$$u_s = u_{s-1} + pas \quad (5.8)$$

$$J_s(\lambda) = J_{s-1}(\lambda) + j(\lambda, u_{s-1})pas \quad (5.9)$$

en considérant qu'on a atteint la valeur $J(\lambda)$ dès que

$$\frac{j(\lambda, u_s)pas}{J_s(\lambda)} < 10^{-6} \quad (5.10)$$

Le choix du pas a une influence importante sur le résultat. En diminuant le pas, on augmente la précision mais le temps d'exécution devient vite très long.

2. Expression à l'aide des fonctions de Bessel

Les coefficients de la matrice de Fisher s'expriment à l'aide des fonctions de Bessel par les formules suivantes :

$$I_f(1,1) = n \left(-1 + \frac{e^{-\lambda}}{\sqrt{\lambda}} \int_0^{\infty} e^{-u} \sqrt{u} \frac{I_0^2(2\sqrt{\lambda u})}{I_1(2\sqrt{\lambda u})} du \right) \quad (3.21)$$

$$I_f(1,2) = I_f(2,1) = \frac{n}{\beta} (\lambda + 1 - e^{-\lambda} \sqrt{\lambda} \int_0^{\infty} e^{-u} \sqrt{u} \frac{I_0^2(2\sqrt{\lambda u})}{I_1(2\sqrt{\lambda u})} du) \quad (3.22)$$

$$I_f(2,2) = \frac{\lambda n}{\beta^2} (-\lambda + e^{-\lambda} \sqrt{\lambda} \int_0^{\infty} e^{-u} \sqrt{u} \frac{I_0^2(2\sqrt{\lambda u})}{I_1(2\sqrt{\lambda u})} du) \quad (3.23)$$

Les programmes de ce calcul dans Mathematica et Matlab sont donnés dans l'annexe E-2. Nous avons utilisé les mêmes méthodes d'intégration numérique que lors du calcul de $J(\lambda)$.

5.3.2 Méthode des moments

Les variances et covariance des paramètres s'expriment de la façon suivante :

$$Var(\hat{\lambda}) = \frac{2\lambda(\lambda + 1)}{n} \quad (3.47)$$

$$Var(\hat{\beta}) = \frac{2\beta^2(\lambda + 1)}{n\lambda} \quad (3.48)$$

$$Cov(\hat{\lambda}, \hat{\beta}) = -\frac{\beta(2\lambda + 1)}{n} \quad (3.49)$$

5.3.3 Méthodes faisant intervenir n_o/n

1. Les variances et covariance des paramètres sont données par :

$$Var(\hat{\lambda}) = \frac{e^{\lambda} - 1}{n} \quad (3.52)$$

$$\text{Var}(\hat{\beta}) = \frac{\beta^2 (e^\lambda - 1)}{n\lambda^2} \quad (3.53)$$

$$\text{Cov}(\hat{\lambda}, \hat{\beta}) = \frac{\beta}{n} \left(1 - \frac{1}{\lambda} (e^\lambda - 1)\right) \quad (3.54)$$

2. ou bien par :

$$\text{Var}(\hat{\lambda}) = \frac{1}{n} \left[(p-1)^2 (e^\lambda - 1) + p^2 \lambda^2 + 2p\lambda + p\lambda^2 \right] \quad (3.57)$$

$$\text{Var}(\hat{\beta}) = \frac{\beta^2}{n\lambda} \left[2p + p^2 \lambda + p\lambda + \frac{1}{\lambda} (p-1)^2 (e^\lambda - 1) \right] \quad (3.58)$$

$$\text{Cov}(\hat{\lambda}, \hat{\beta}) = \frac{\beta}{n} \left[1 - 2p - p^2 \lambda - p\lambda - \frac{1}{\lambda} (p-1)^2 (e^\lambda - 1) \right] \quad (3.59)$$

avec

$$p = 1 - (2 - \sqrt{2}) \sqrt{\frac{n_0}{n}} \quad (3.60)$$

5.4 Détermination du quantile x_T

5.4.1 Calcul de la fonction de répartition $F(u)$

$$F(u) = 1 - e^{-\lambda} e^{-u} \sum_{i=0}^{\infty} \frac{\lambda^{i+1}}{(i+1)!} \left(\sum_{j=0}^i \frac{u^j}{j!} \right) \quad (2.13)$$

Nous proposons dans l'annexe E-3 deux algorithmes pour calculer cette fonction dans Matlab, le premier est la programmation directe de (2.13), le second a été établi par Lubès *et al.* (1994). Ils fournissent des résultats identiques sauf pour des valeurs de λ très grandes, auquel cas on constate que $F(u) \rightarrow e^{-\lambda} = F(0)$ lorsque $u \rightarrow 0$ plus rapidement avec la méthode proposée par Lubès.

Dans Mathematica, la fonction de répartition $F(u)$ peut être définie directement comme l'intégrale de la densité, ou bien au moyen d'une somme infinie (équation (2.13)).

5.4.2 Calcul de x_T

On détermine \hat{x}_T en solutionnant l'équation suivante :

$$F(x_T) = 1 - \frac{1}{T} \quad (4.5)$$

La résolution a été programmée dans Mathématica et dans Matlab (voir Annexe E-3)

- Dans Mathématica, si on introduit une seule valeur initiale pour x_T , la méthode employée est celle de Newton, si on en spécifie deux, c'est une méthode par approches successives. Si λ est trop petit et T est grand, c'est-à-dire pour des valeurs de x_T proches de zéro, aucune des deux méthodes ne converge.
- Dans Matlab, nous avons programmé les deux méthodes expliquées dans la section 4.2 : la méthode de Newton-Raphson et une méthode basée sur le principe de la dichotomie. Là encore, pour des valeurs de λ trop faibles et lorsque T est grand, c'est-à-dire pour des valeurs de x_T proches de zéro, la méthode de Newton ne fonctionne pas car

$$x_{T_i} = x_{T_{i-1}} - \frac{F(x_{T_{i-1}}) - (1 - \frac{1}{T})}{f(x_{T_{i-1}})} \quad (5.12)$$

devient négatif, tandis que le résultat de la méthode par dichotomie dépend du nombre maximum d'itérations imposé, car la convergence n'est pas atteinte.

5.5 Variance de \hat{x}_T

La variance du quantile \hat{x}_T est donnée par :

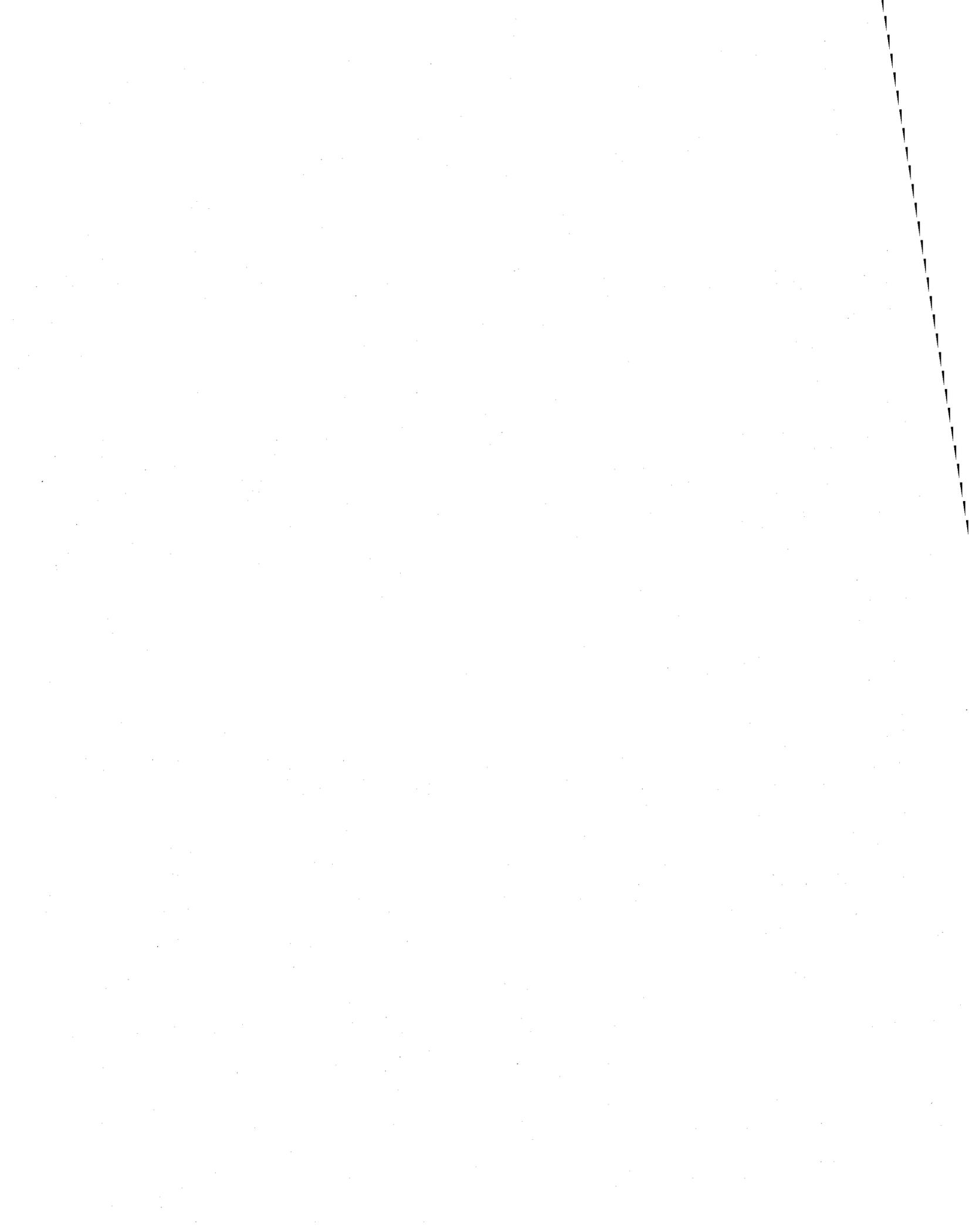
$$Var\{\hat{x}_T\} = \left(\frac{\partial x_T}{\partial \lambda}\right)^2 Var\{\hat{\lambda}\} + \left(\frac{\partial x_T}{\partial \beta}\right)^2 Var\{\hat{\beta}\} + 2\left(\frac{\partial x_T}{\partial \lambda}\right)\left(\frac{\partial x_T}{\partial \beta}\right) Cov\{\hat{\lambda}, \hat{\beta}\} \quad (4.4)$$

où

$$\frac{\partial x_T}{\partial \lambda} = \sqrt{\frac{\hat{\beta}\hat{x}_T}{\hat{\lambda}}} \frac{I_0\left(2\sqrt{\frac{\hat{\lambda}\hat{x}_T}{\hat{\beta}}}\right)}{I_1\left(2\sqrt{\frac{\hat{\lambda}\hat{x}_T}{\hat{\beta}}}\right)} \quad (4.7)$$

$$\frac{\partial x_T}{\partial \beta} = \frac{\hat{x}_T}{\hat{\beta}} \quad (4.8)$$

Les variances et covariance asymptotiques des estimateurs des paramètres ont été calculées au paragraphe 5.3.



6 APPLICATION

Pour illustrer les résultats établis dans les chapitres précédents, nous allons traiter trois exemples, en suivant pour chacun la même démarche. Après avoir estimé les paramètres λ et β par quatre méthodes différentes, nous utiliserons les résultats d'un test du khi-deux comme critère pour ne retenir que les estimateurs qui nous semblent adéquats. Nous calculerons ensuite les variances et covariances des paramètres, puis les quantiles et les intervalles de confiance qui leur sont associés. Au cours de chacune des trois études, nous essaierons de comparer les différentes méthodes numériques employées.

6.1 Exemple I

Nous disposons d'un échantillon de 86 observations, chacune de ces observations représentant la précipitation en millimètres d'eau relevée à Massiac (Cantal, France), sur une période de 20 jours. Les données sont présentées dans le tableau 6.1

Tableau 6.1. Précipitations (mm) relevées à Massiac ($n = 86$; $n_0 = 1$).

0	1.5	3.5	5.5	9.2	11.7	19.8	28.9
0.1	1.8	3.6	7	9.2	11.9	20.1	29.7
0.1	1.9	4.2	7.3	9.5	12	22.5	32
0.2	2	4.3	7.8	9.6	13.2	22.7	40.7
0.3	2.1	4.4	8	10	13.5	23.5	52.2
0.3	2.5	4.7	8.3	10.2	15	23.9	56.3
0.9	2.8	4.8	8.4	10.3	15.3	24.3	58.2
1	2.9	5.1	8.5	11.2	15.6	24.3	60.6
1.1	3.1	5.1	8.9	11.2	16.5	24.9	69.9
1.2	3.3	5.2	9	11.2	17.6	25.3	
1.4	3.3	5.2	9.1	11.5	17.8	26.9	

6.1.1 Estimation des paramètres

- Nous avons tout d'abord estimé les paramètres λ et β par quatre méthodes différentes. Les estimateurs obtenus sont présentés au tableau 6.2.

Tableau 6.2. Estimation des paramètres λ et β .

Méthode	$\hat{\lambda}$	$\hat{\beta}$
Méthode du maximum de vraisemblance (cf. section 3.1)	2.6440	4.9756
Méthode des moments (cf. section 3.2)	1.6520	7.9651
Méthode utilisant n_o et n (1) (cf. section 3.3.1)	4.4543	2.9540
Méthode utilisant n_o et n (2) (cf. section 3.3.1)	1.8290	7.6486

La méthode utilisant n_o et n (2) donne des estimations très similaires à celles de la méthode des moments. Ce n'est pas étonnant puisque le nombre de valeurs nulles de l'échantillon est très faible ($n_o = 1$).

- Appliquons le test du khi-deux (voir Perreault *et al.* (1994), p.60) pour vérifier que la loi des fuites avec les paramètres estimés est compatible avec les observations.

Le domaine de variation de notre variable X ayant été découpé en 10 classes, on note N_i le nombre d'observations de notre échantillon qui appartiennent à la classe C_i , et p_i la probabilité théorique pour la variable X d'être dans l'intervalle C_i . Soit χ^2 la statistique suivante :

$$\chi^2 = \sum_{i=1}^{10} \frac{(N_i - np_i)^2}{np_i} \quad (6.1)$$

On peut montrer que est χ^2 distribué asymptotiquement suivant une loi du χ^2 à 7 degrés de liberté. Le tableau 6.3 présente les résultats du test appliqué aux trois premières méthodes :

Tableau 6.3. Résultats du test du khi-deux.

CLASSES C_i	N_i	np_i		
		Maximum de vraisemblance	Méthode des moments	Méthode utilisant n_0 et $n(1)$
$X < 1$	7	9.4553	19.8631	2.8161
$1 \leq X < 2.5$	9	5.2799	4.9092	3.8
$2.5 \leq X < 4.5$	11	7.2741	6.2094	6.7073
$4.5 \leq X < 7$	7	9.0243	7.1909	10.0591
$7 \leq X < 9$	8	6.8853	5.2429	8.5660
$9 \leq X < 11$	9	6.4250	4.8016	8.4206
$11 \leq X < 13.5$	8	7.2253	5.3955	9.7256
$13.5 \leq X < 22.5$	9	18.6220	14.4417	23.6001
$22.5 \leq X < 29$	10	7.3927	6.5624	7.4449
$29 \leq X$	8	8.3861	11.4033	4.8593
Statistique χ^2		12.8258	26.6849	29.3321

La table du khi-deux nous donne une valeur limite pour χ^2 correspondant à la probabilité 5% de dépassement égale à 14.0671. Les résultats obtenus semblent indiquer que la méthode des moments et la méthode faisant intervenir n_0 et $n(1)$ sont moins adéquates. Cela confirme ce que nous avons vu au Chapitre 3 (section 3.2.2), à savoir que la méthode des moments est peu efficace pour les faibles valeurs de λ . D'autre part les méthodes utilisant n_0 et n ne sont intéressantes que lorsque l'échantillon possède un nombre important de valeurs nulles.

Nous ne retiendrons donc ici que la méthode du maximum de vraisemblance pour estimer λ et β .

6.1.2 Variances et covariance des paramètres

Nous avons calculé à l'aide de Mathematica les coefficients de la matrice de Fisher au moyen de deux expressions différentes, la première faisant intervenir la fonction $J(\lambda)$, la seconde les fonctions de Bessel, ainsi que les coefficients de la matrice de Fisher observée (voir section 3.1.4). Les algorithmes sont exposés au Chapitre 5, section 5.2. Les résultats ont été regroupés dans le tableau 6.4.

Tableau 6.4. Coefficients de la matrice d'information de Fisher.

Méthode de calcul	$I_f(1,1)$	$I_f(1,2)$	$I_f(2,2)$
A l'aide de $J(\lambda)$	20.5236	31.9356	130.704
A l'aide des fonctions de Bessel	20.5236	31.9356	130.704
Matrice observée	20.4201	32.1968	130.047

Le calcul au moyen de la fonction $J(\lambda)$ ou des fonctions de Bessel donnent les mêmes valeurs pour les coefficients de la matrice de Fisher. Le calcul de la matrice observée est une approche approximative qui donne ici d'assez bons résultats.

Les variances et covariances des paramètres sont obtenues en inversant la matrice de Fisher, le tableau 6.5 donne les résultats pour la matrice de Fisher évaluée à l'aide des fonctions de Bessel (tableau 6.4).

Tableau 6.5. Variances et covariance des paramètres.

Variances et covariance	$Var(\hat{\lambda})$	$Cov(\hat{\lambda}, \hat{\beta})$	$Var(\hat{\beta})$
	0.0013	-0.0019	0.0049

6.1.3 Estimation des quantiles

1. Estimation des quantiles

Les quantiles x_T de période de retour T ont été estimés par plusieurs méthodes. Mathematica inverse la fonction de répartition, celle-ci pouvant être définie de deux façons différentes, au moyen d'une somme infinie (équation (2.13)) ou par intégration numérique de la densité (équation (2.11)). Dans Matlab, nous pouvons utiliser la méthode de Newton-Raphson ou une méthode par dichotomie (la formule employée pour définir la fonction de répartition n'influe pas sur le résultat). Les algorithmes utilisés sont expliqués au Chapitre 5, paragraphe 5.4. Les résultats obtenus sont présentés au tableau 6.6.

Tableau 6.6. Estimation des quantiles de période de retour T .

Probabilité au dépassement $1/T$	x_T						
	0.9	0.75	0.5	0.25	0.1	0.01	0.001
Méthode de calcul							
MATHEMATICA							
F définie comme somme infinie	0.1505	0.8762	2.1235	3.8376	5.7771	9.9904	13.7689
F définie comme intégrale	0.1505	0.8763	2.1236	3.8376	5.7771	9.9906	13.7718
MATLAB							
Méthode par dichotomie	0.1505	0.8762	2.1235	3.8376	5.7771	9.9905	13.7700
Méthode de Newton-Raphson	0.1505	0.8762	2.1235	3.8376	5.7771	9.9905	13.7700

Les méthodes employées donnent des résultats très similaires. De petites différences (sur la troisième décimale) sont observées pour les grandes périodes de retour.

2. Variance des quantiles

Le tableau 6.7 présente les variances asymptotiques des quantiles (voir section 4.2.1), en prenant les estimations des quantiles données par le logiciel Matlab (tableau 6.6).

Tableau 6.7. Variances des quantiles de période de retour T .

Probabilité au dépassement $1/T$	0.9	0.75	0.5	0.25	0.1	0.01	0.001
$Var(\hat{x}_T)$	0.0198	0.0197	0.0195	0.0196	0.0201	0.0230	0.0282

3. Intervalles de confiance

La méthode pour évaluer les intervalles de confiance est donnée dans la section 4.2.2. On lit dans la table de la loi normale centrée-réduite, pour $\alpha = 5\%$,

$$z_{1-\frac{\alpha}{2}} = 1.96 \quad (6.2)$$

ce qui nous donne les résultats suivants, présentés dans le tableau 6.8 :

Tableau 6.8. Intervalles de confiance associés aux quantiles de période de retour T .

Probabilité au dépassement $1/T$	0.9	0.75	0.5	0.25	0.1	0.01	0.001
\hat{x}_T	0.1505	0.8762	2.1235	3.8376	5.7771	9.9905	13.7700
Intervalle de confiance (95%)	0 0.4263	0.6011 1.1513	1.8498 2.3972	3.5632 4.1120	5.4992 6.0550	9.6933 10.2877	13.4409 14.0991

6.2 Exemple II

Nous disposons d'un échantillon de 86 observations, chacune de ces observations représentant la précipitation en millimètres d'eau relevée à Chateauneuf de Randon (France), sur une période de 10 jours. Les données sont présentées dans le tableau 6.9.

Tableau 6.9. Précipitations (mm) relevées à Chateauneuf de Randon ($n = 86$; $n_o = 15$).

0	0	0.5	2.4	4	7.1	11.4	20.3
0	0	0.9	2.7	4.1	8	11.5	20.7
0	0	0.9	2.7	4.3	8	12.6	22.6
0	0	1	2.7	4.4	8.3	14.3	24.8
0	0.1	1.1	2.9	4.7	8.4	15.2	29.6
0	0.2	1.7	2.9	5	8.6	15.2	31.2
0	0.3	1.7	3	5	8.8	15.9	35.5
0	0.3	1.7	3.3	5.7	8.9	16.3	36.7
0	0.4	1.7	3.4	6.1	9.3	17.4	41.7
0	0.4	1.7	3.5	6.5	10	17.9	
0	0.5	1.7	3.7	6.6	11.1	19	

L'approche adoptée pour traiter cet échantillon est identique à celle de l'exemple I.

6.2.1 Estimation des paramètres

- Nous avons d'abord estimé les paramètres λ et β par quatre méthodes différentes. Les estimations obtenues sont présentées au tableau 6.10.

Tableau 6.10. Estimation des paramètres λ et β .

Méthode	$\hat{\lambda}$	$\hat{\beta}$
Méthode du maximum de vraisemblance (cf. section 3.1)	1.6106	4.5680
Méthode des moments (cf. section 3.2)	1.2618	5.8305
Méthode utilisant n_o et n (1) (cf. section 3.3.1)	1.7463	4.2129
Méthode utilisant n_o et n (2) (cf. section 3.3.1)	1.3803	5.4348

- Appliquons le test du khi-deux pour vérifier que la loi des fuites avec les paramètres estimés est compatible avec les observations.

Le domaine de variation de notre variable X a été découpé en 8 classes. Les résultats du test appliqué aux quatre méthodes sont regroupés dans le tableau 6.11.

Tableau 6.11. Résultats du test du khi-deux.

CLASSES C_i	N_i	np_i			
		Maximum de vraisemblance	Méthode des moments	Méthode utilisant n_o et n (1)	Méthode utilisant n_o et n (2)
$X < 0.2$	16	18.3869	25.3943	16.2398	22.7218
$0.2 \leq X < 1.1$	10	5.2903	4.5480	5.4695	4.7586
$1.1 \leq X < 2.8$	11	9.3049	7.8646	9.7220	8.2775
$2.8 \leq X < 4.4$	10	7.8614	6.5663	8.2791	6.9458
$4.4 \leq X < 8.1$	11	14.7972	12.3609	15.6308	13.1166
$8.1 \leq X < 11.5$	9	9.8271	8.3963	10.3149	8.8955
$11.5 \leq X < 19$	9	12.4340	11.3481	12.7442	11.8828
$19 \leq X$	10	8.0982	9.5176	7.5997	9.4013
Statistique χ^2		7.8323	13.7601	7.6796	11.0804

χ^2 est distribué asymptotiquement suivant une loi du khi-deux à 5 degrés de liberté. Les tables nous donnent une valeur limite pour χ^2 correspondant à la probabilité 5% de dépassement égale à 11.0705, Les résultats obtenus semblent indiquer que la méthode des moments ainsi que la méthode utilisant n_o et n (2) sont moins adéquates, nous les abandonnons donc pour la suite de l'étude.

6.2.2 Variances et covariance des paramètres

1. Matrice de Fisher

Les coefficients de la matrice d'information de Fisher, calculés par trois méthodes différentes, sont présentés au tableau 6.12.

Tableau 6.12. Coefficients de la matrice d'information de Fisher.

Méthode de calcul	$I_f(1,1)$	$I_f(1,2)$	$I_f(2,2)$
A l'aide de $J(\lambda)$	33.8928	6.87658	4.21338
A l'aide des fonctions de Bessel	33.8928	6.87658	4.21338
Matrice observée	34.6283	6.6169	4.3046

2. Variances et covariance

Le tableau 6.13 donne les variances et covariance des paramètres estimés par les deux méthodes retenues. Pour la méthode faisant intervenir n_0 et $n(1)$, nous avons utilisé les équations (3.52) à (3.54).

Tableau 6.13. Variances et covariance des paramètres.

Méthode	$Var(\hat{\lambda})$	$Cov(\hat{\lambda}, \hat{\beta})$	$Var(\hat{\beta})$
Méthode du maximum de vraisemblance	0.0005	-0.0008	0.0041
Méthode utilisant n_0 et $n(1)$	0.0550	-0.0838	0.3203

6.2.3 Estimation des quantiles

1. Estimation des quantiles

Les valeurs estimées des quantiles dépendent des méthodes d'estimation des paramètres. Les résultats obtenus sont présentés aux tableaux 6.14 et 6.15.

Tableau 6.14. Estimation des quantiles de période de retour T pour les paramètres estimés par la méthode du maximum de vraisemblance.

Probabilité au dépassement $1/T$	x_T					
	0.75	0.5	0.25	0.1	0.01	0.001
Méthode de calcul						
MATHEMATICA						
F définie comme somme infinie	0.1587	1.0670	2.4320	4.0573	7.7325	11.1195
F définie comme intégrale	0.1587	1.0670	2.4320	4.0573	7.7330	11.1241
MATLAB						
Méthode par dichotomie	0.1587	1.0670	2.4320	4.0573	7.7325	11.1201
Méthode de Newton-Raphson	0.1587	1.0670	2.4320	4.0573	7.7325	11.1201

Tableau 6.15. Estimation des quantiles de période de retour T pour les paramètres estimés par la méthode utilisant n_0 et $n(1)$.

Probabilité au dépassement $1/T$	x_T					
	0.75	0.5	0.25	0.1	0.01	0.001
Méthode de calcul						
MATHEMATICA						
F définie comme somme infinie	0.2528	1.2083	2.6239	4.2952	8.0494	11.4938
F définie comme intégrale	0.2528	1.2083	2.6239	4.2953	8.0499	11.4982
MATLAB						
Méthode par dichotomie	0.2528	1.2083	2.6239	4.2952	8.0494	11.4942
Méthode de Newton-Raphson	0.2528	1.2083	2.6239	4.2952	8.0494	11.4942

2. Variance des quantiles

Nous avons calculé les variances des quantiles pour les deux méthodes d'estimation des paramètres, les résultats sont présentés au tableau 6.16.

Tableau 6.16. Variances des quantiles de période de retour T .

Probabilité au dépassement $1/T$	$Var(\hat{x}_T)$					
	0.75	0.5	0.25	0.1	0.01	0.001
Méthode d'estimation des paramètres						
Méthode du maximum de vraisemblance	0.0160	0.0156	0.0154	0.0158	0.0192	0.0255
Méthode utilisant n_0 et $n(1)$	1.2660	1.2225	1.1905	1.2047	1.4537	1.9624

3. Intervalles de confiance

On lit dans la table de la loi normale centrée-réduite, pour $\alpha = 5\%$:

$$z_{1-\frac{\alpha}{2}} = 1.96 \quad (6.3)$$

ce qui nous donne les résultats suivants, présentés dans le tableau 6.17 :

Tableau 6.17. Intervalles de confiance associés aux quantiles de période de retour T .

Probabilité au dépassement $1/T$	0.75	0.5	0.25	0.1	0.01	0.001
Méthode d'estimation des paramètres						
METHODE DU MAXIMUM DE VRAISEMBLANCE						
\hat{x}_T	0.1587	1.0670	2.4320	4.0573	7.7325	11.1201
Intervalle de confiance (95%)	0	0.8222	2.1888	3.8109	7.4609	10.8071
	0.4066	1.3118	2.6752	4.3037	8.0041	11.4331
METHODE UTILISANT n_0 ET $n(1)$						
\hat{x}_T	0.2528	1.2083	2.6239	4.2952	8.0494	11.4942
Intervalle de confiance (95%)	0	0	0.4853	2.1439	5.6862	8.7485
	2.4581	3.3754	4.7625	6.4465	10.4126	14.2399

6.3 Exemple III

Nous avons généré au moyen de l'ordinateur 50 valeurs distribuées selon une loi normale auxquelles nous avons joint 10 zéros pour obtenir un échantillon de 60 valeurs ($n = 60$; $n_0 = 10$). La démarche pour cette étude est celle des deux exemples précédents.

6.2.1 Estimation des paramètres

Nous avons d'abord estimé les paramètres λ et β par quatre méthodes différentes. Les estimations obtenues sont présentées au tableau 6.18.

Tableau 6.18. Estimation des paramètres λ et β .

Méthode	$\hat{\lambda}$	$\hat{\beta}$
Méthode du maximum de vraisemblance (cf. section 3.1)	1.8084	0.7166
Méthode des moments (cf. section 3.2)	1.7448	0.7427
Méthode utilisant n_o et n (1) (cf. section 3.3.1)	1.7918	0.7233
Méthode utilisant n_o et n (2) (cf. section 3.3.1)	1.7561	0.7381

6.3.2 Variances et covariance des paramètres

1. Matrice de Fisher

Les coefficients de la matrice d'information de Fisher, calculés par trois méthodes différentes, sont présentés au tableau 6.19.

Tableau 6.19. Coefficients de la matrice d'information de Fisher.

Méthode de calcul	$I_f(1,1)$	$I_f(1,2)$	$I_f(2,2)$
A l'aide de $J(\lambda)$	20.5236	31.9356	130.704
A l'aide des fonctions de Bessel	20.5236	31.9356	130.704
Matrice observée	20.4201	32.1968	130.0471

2. Variances et covariance

Le tableau 6.20 donne les variances et covariance des paramètres estimés par les quatre méthodes différentes. Pour la méthode des moments, nous avons utilisé les équations (3.47) à (3.49) et pour la méthode faisant intervenir n_o et n (2), les équations (3.57) à (3.60).

Tableau 6.20. Variances et covariance des paramètres.

Méthode	$Var(\hat{\lambda})$	$Cov(\hat{\lambda}, \hat{\beta})$	$Var(\hat{\beta})$
Méthode du maximum de vraisemblance	0.0013	-0.0003	0.0002
Méthode des moments	0.1596	-0.0556	0.0289
Méthode utilisant n_0 et n (1)	0.0833	-0.0216	0.0136
Méthode utilisant n_0 et n (2)	0.1180	-0.0373	0.0208

6.3.3 Estimation des quantiles

1. Estimation des quantiles

Les valeurs estimées des quantiles dépendent des méthodes d'estimation des paramètres. Les tableaux 6.21 à 6.24 présentent les résultats obtenus pour chaque méthode.

Tableau 6.21. Estimation des quantiles de période de retour T pour les paramètres estimés par la méthode du maximum de vraisemblance.

Probabilité au dépassement $1/T$	x_T					
	0.75	0.5	0.25	0.1	0.01	0.001
Méthode de calcul						
MATHEMATICA						
F définie comme somme infinie	0.2956	1.2725	2.7107	4.4026	8.19120	11.6621
F définie comme intégrale	0.2956	1.2725	2.7108	4.4026	8.1924	11.6650
MATLAB						
Méthode par dichotomie	0.2956	1.2725	2.7107	4.4026	8.1921	11.6628
Méthode de Newton-Raphson	0.2956	1.2725	2.7107	4.4026	8.1921	11.6628

Tableau 6.22. Estimation des quantiles de période de retour T pour les paramètres estimés par la méthode des moments.

Probabilité au dépassement $1/T$	x_T					
	0.75	0.5	0.25	0.1	0.01	0.001
Méthode de calcul						
MATHEMATICA						
F définie comme somme infinie	0.2518	1.2067	2.6218	4.2926	8.0460	11.4897
F définie comme intégrale	0.2518	1.2067	2.6218	4.2927	8.0465	11.4944
MATLAB						
Méthode par dichotomie	0.2517	1.2067	2.6218	4.2926	8.0461	11.4911
Méthode de Newton-Raphson	0.2517	1.2067	2.6218	4.2926	8.0461	11.4911

Tableau 6.23. Estimation des quantiles de période de retour T pour les paramètres estimés par la méthode utilisant n_0 et $n(1)$.

Probabilité au dépassement $1/T$	x_T					
	0.75	0.5	0.25	0.1	0.01	0.001
Méthode de calcul						
MATHEMATICA						
F définie comme somme infinie	0.2842	1.2553	2.6876	4.3740	8.1540	11.6169
F définie comme intégrale	0.2842	1.2553	2.6876	4.3740	8.1545	11.6218
MATLAB						
Méthode par dichotomie	0.2842	1.2553	2.6876	4.3740	8.1541	11.6187
Méthode de Newton-Raphson	0.2842	1.2553	2.6876	4.3740	8.1541	11.6187

Tableau 6.24. Estimation des quantiles de période de retour T pour les paramètres estimés par la méthode utilisant n_0 et n (2).

Probabilité au dépassement $1/T$	x_T					
	0.75	0.5	0.25	0.1	0.01	0.001
Méthode de calcul						
MATHEMATICA						
F définie comme somme infinie	0.2596	1.2184	2.6376	4.3122	8.0721	11.5204
F définie comme intégrale	0.2596	1.2184	2.6377	4.3129	8.0726	11.5252
MATLAB						
Méthode par dichotomie	0.2595	1.2183	2.6376	4.3122	8.0720	11.5210
Méthode de Newton-Raphson	0.2595	1.2183	2.6376	4.3122	8.0720	11.5210

2. Variance des quantiles

Nous avons calculé les variances des quantiles pour les chaque méthode d'estimation des paramètres, les résultats sont présentés au tableau 6.25.

Tableau 6.25. Variances des quantiles de période de retour T .

Probabilité au dépassement $1/T$	$Var(\hat{x}_T)$					
	0.75	0.5	0.25	0.1	0.01	0.001
Méthode d'estimation des paramètres						
Méthode du maximum de vraisemblance	7.91e-4	9.88e-4	0.0021	0.005	0.0175	0.0371
Méthode des moments	0.1019	0.0746	0.1226	0.3401	1.5587	3.6483
Méthode utilisant n_0 et n (1)	0.0512	0.0579	0.1227	0.2926	1.0785	2.3227
Méthode utilisant n_0 et n (2)	0.0753	0.0659	0.1222	0.3129	1.2945	2.9238

3. Intervalles de confiance

On lit dans la table de la loi normale centrée-réduite, pour $\alpha = 5\%$:

$$z_{1-\frac{\alpha}{2}} = 1.96 \quad (6.4)$$

ce qui nous donne les résultats suivants, présentés dans le tableau 6.26 :

Tableau 6.26. Intervalles de confiance associés aux quantiles de période de retour T .

Probabilité au dépassement $1/T$	0.75	0.5	0.25	0.1	0.01	0.001
Méthode d'estimation des paramètres						
METHODE DU MAXIMUM DE VRAISEMBLANCE						
\hat{x}_T	0.2956	1.2725	2.7107	4.4026	8.1921	11.6628
Intervalle de confiance (95%)	0.2405	1.2109	2.6209	4.2640	7.9328	11.2853
	0.3507	1.3341	2.8005	4.5412	8.4514	12.0403
METHODE DES MOMENTS						
\hat{x}_T	0.2517	1.2067	2.6218	4.2926	8.0461	11.4911
Intervalle de confiance (95%)	0	0.6714	1.9355	3.1496	5.5991	7.7474
	0.8774	1.7420	3.3081	5.4356	10.4931	15.2348
METHODE UTILISANT n_0 ET $n(1)$						
\hat{x}_T	0.2842	1.2553	2.6876	4.3740	8.1541	11.6187
Intervalle de confiance (95%)	0	0.7837	2.0010	3.3138	6.1186	7.8750
	0.7277	1.7269	3.3742	5.4342	10.1896	15.3624
METHODE UTILISANT n_0 ET $n(2)$						
\hat{x}_T	0.2595	1.2183	2.6376	4.3122	8.0720	11.5210
Intervalle de confiance (95%)	0	0.7151	1.9524	3.2158	5.8420	8.1696
	0.7973	1.7215	3.3228	5.4086	10.3020	14.8724

La méthode utilisant n_0 et $n(1)$ est meilleure que la méthode utilisant n_0 et $n(2)$ car λ est faible. Les meilleurs résultats sont obtenus par la méthode du maximum de vraisemblance.

7 GÉNÉRATEUR D'ÉCHANTILLONS POUR LA LOI DES FUITES

En simulation, il est souvent utile de générer des variables aléatoires à partir de distributions dont la forme analytique est connue. Il existe pour cela plusieurs techniques, qui nécessitent généralement l'emploi d'un générateur pour la loi uniforme sur $(0,1)$. Dans ce qui suit, si U est une variable aléatoire suivant la loi uniforme sur $(0,1)$, on notera $U \sim U(0,1)$.

Nous n'aborderons ici que deux méthodes, la méthode d'inversion et la méthode d'acceptation-rejet, qui sont parmi les plus utilisées.

7.1 Méthodes de génération

7.1.1 Méthode d'inversion

Soit X une variable aléatoire, dont la fonction de répartition F est continue et strictement croissante, et admet une fonction inverse F^{-1} . L'algorithme qui permet de générer la variable aléatoire X comprend alors les étapes suivantes :

- 1- Générer $u \sim U(0,1)$,
- 2- Calculer $x = F^{-1}(u)$.

Cette méthode est particulièrement intéressante lorsqu'on dispose d'une formule explicite pour F^{-1} , ce qui n'est pas le cas pour la loi des fuites. On est donc obligé d'avoir recours aux méthodes numériques d'inversion, comme lors de l'estimation des quantiles (voir sections 4.1 et 5.4). Puisque l'utilisation de ces méthodes risque d'avoir une influence sur la précision des résultats, nous préférons envisager une autre méthode, dite d'acceptation-rejet.

7.1.2 Méthode d'acceptation-rejet

Soit f la densité de probabilité de la loi dont on veut tirer un échantillon aléatoire, et t une fonction qui majore f , c'est-à-dire telle que $t(x) \geq f(x)$ pour tout x . A priori, t n'est pas une densité de probabilité puisque :

$$c = \int_{-\infty}^{+\infty} t(x) dx \geq \int_{-\infty}^{+\infty} f(x) dx = 1 \quad (7.1)$$

Mais si t est choisie telle que l'intégrale c soit finie, alors la fonction $r(x) = t(x)/c$ est une densité de probabilité et l'algorithme de base de la méthode d'acceptation-rejet est le suivant:

- 1- Générer y de densité r ,
- 2- Générer $u \sim U(0,1)$, indépendante de y ,
- 3- Si $u \leq f(y)/t(y)$ alors accepter $x = y$, sinon rejeter y et reprendre à l'étape 1.

Le choix de la fonction t joue un rôle très important dans l'efficacité de la méthode. Puisque l'étape 1 consiste à générer une variable aléatoire y de densité $r(x) = t(x)/c$, il faut rendre cette opération facile et rapide, c'est pourquoi on cherchera en général une fonction t telle que y puisse être générée par la méthode d'inversion. D'autre part, on veut que la probabilité de rejet à l'étape 3 soit aussi petite que possible, puisqu'en cas de rejet, l'algorithme reprend à l'étape 1, ce qui augmente le temps de génération. On peut montrer (voir Law et Kelton (1982), p.273) que, pour chaque itération, la probabilité de rejet à l'étape 3 est égale à $(1-1/c)$. Il faut donc choisir la fonction t de façon à ce que c soit aussi proche que possible de 1, ce qui revient à minimiser la surface comprise entre les courbes des fonctions $f(x)$ et $t(x)$.

Nous allons proposer, dans ce qui suit, une expression de t pour la loi des fuites.

7.2 Génération de la loi des fuites par la méthode d'acceptation-rejet

En nous inspirant des travaux concernant la loi de Halphen de type A (Essadaoui *et al.* (1995)), nous avons pensé à une fonction majorante définie par morceaux, exponentielle croissante, constante, puis exponentielle décroissante.

L'étude de la fonction de densité $f(u)$ selon la valeur du paramètre λ (voir chapitre 2) nous amène à envisager quatre cas.

7.2.1 $\lambda < 3-\sqrt{3}$

Pour $\lambda < 3-\sqrt{3}$, $f(u)$ présente une forme en J renversé, sans inflexion et son mode est observé en $u = 0$. On cherche $t(u)$ sous forme exponentielle décroissante, et telle que sa courbe représentative tangente celle de f à l'origine, soit :

$$t(u) = \lambda \exp\left[-\lambda + u\left(\frac{\lambda}{2} - 1\right)\right] \quad (7.2)$$

Le calcul de c nous donne :

$$c = \int_{-\infty}^{+\infty} t(u) du = \frac{2\lambda e^{-\lambda}}{2-\lambda} \quad (7.3)$$

La figure (7.1) représente les fonctions $f(u)$ et $t(u)$ pour $\lambda=1$.

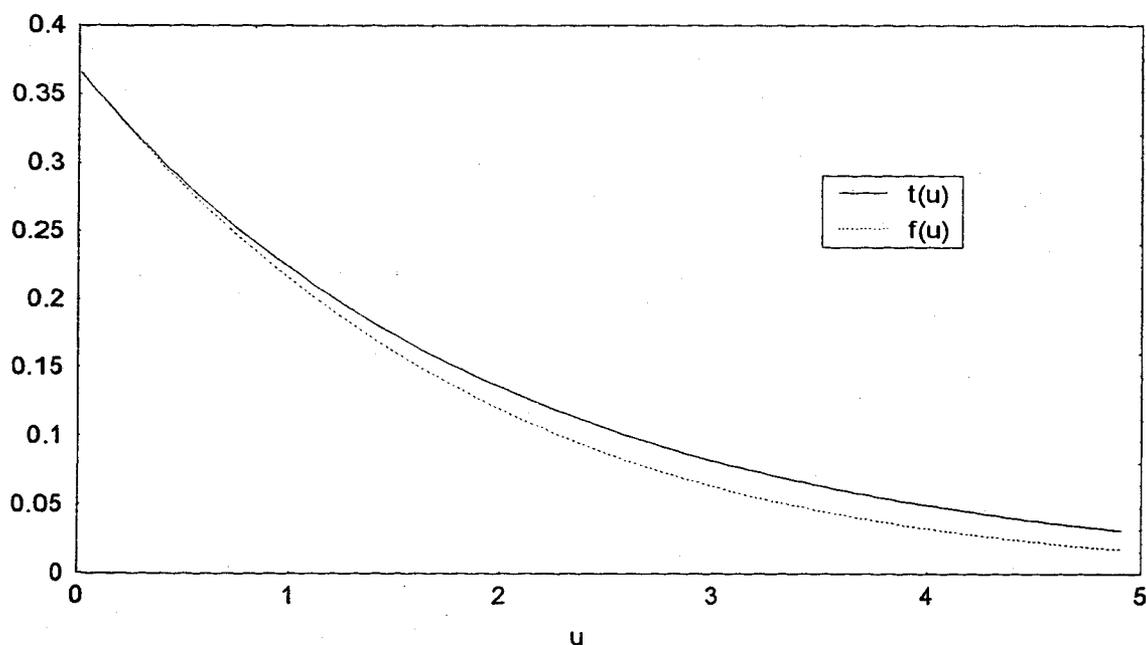


Figure 7.1. Fonction de densité et fonction majorante pour $\lambda = 1$.

7.2.2 $3-\sqrt{3} \leq \lambda < 2$

Pour $\lambda < 3-\sqrt{3}$, $f(u)$ présente une forme en J renversé, avec un point d'inflexion observé en $u = u_0$ et son mode est atteint en $u = 0$. On cherche $t(u)$ constante sur $[0, u_0[$ et exponentielle décroissante sur $[u_0, \infty[$, et telle que sa courbe représentative tangente celle de f en $u = u_0$, soit :

$$t(u) = \begin{cases} \lambda e^{-\lambda}, & u < u_0 \\ f(u_0) \exp\left[\frac{f'(u_0)}{f(u_0)}(u - u_0)\right], & u \geq u_0 \end{cases} \quad (7.4)$$

Le calcul de c nous donne :

$$c = u_0 \lambda e^{-\lambda} - \frac{f(u_0)^2}{f'(u_0)} \quad (7.5)$$

La figure (7.2) représente les fonctions $f(u)$ et $t(u)$ pour $\lambda=1,5$.

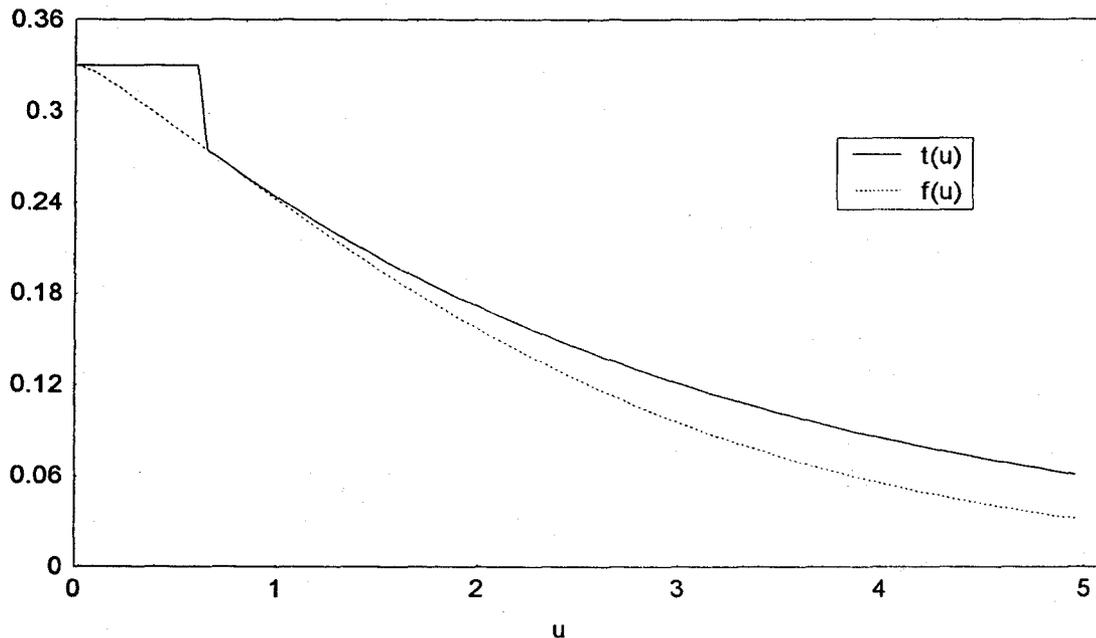


Figure 7.2. Fonction de densité et fonction majorante pour $\lambda = 1,5$.

7.2.3 $2 \leq \lambda < 3 + \sqrt{3}$

Pour $2 \leq \lambda < 3 + \sqrt{3}$, $f(u)$ présente une forme en cloche, avec un point d'inflexion observé en $u = u_1$ et son mode est atteint en $u = u_0$ ($u_0 < u_1$). On cherche $t(u)$ constante sur $[0, u_1[$ et exponentielle décroissante sur $[u_1, \infty[$, et telle que sa courbe représentative tangente celle de f en $u = u_1$, soit :

$$t(u) = \begin{cases} f(u_0), & u < u_1 \\ f(u_1) \exp\left[\frac{f'(u_1)}{f(u_1)}(u - u_1)\right], & u \geq u_1 \end{cases} \quad (7.6)$$

Le calcul de c nous donne :

$$c = u_0 f(u_0) - \frac{f(u_1)^2}{f'(u_1)} \quad (7.7)$$

La figure (7.3) représente les fonctions $f(u)$ et $t(u)$ pour $\lambda=3$.

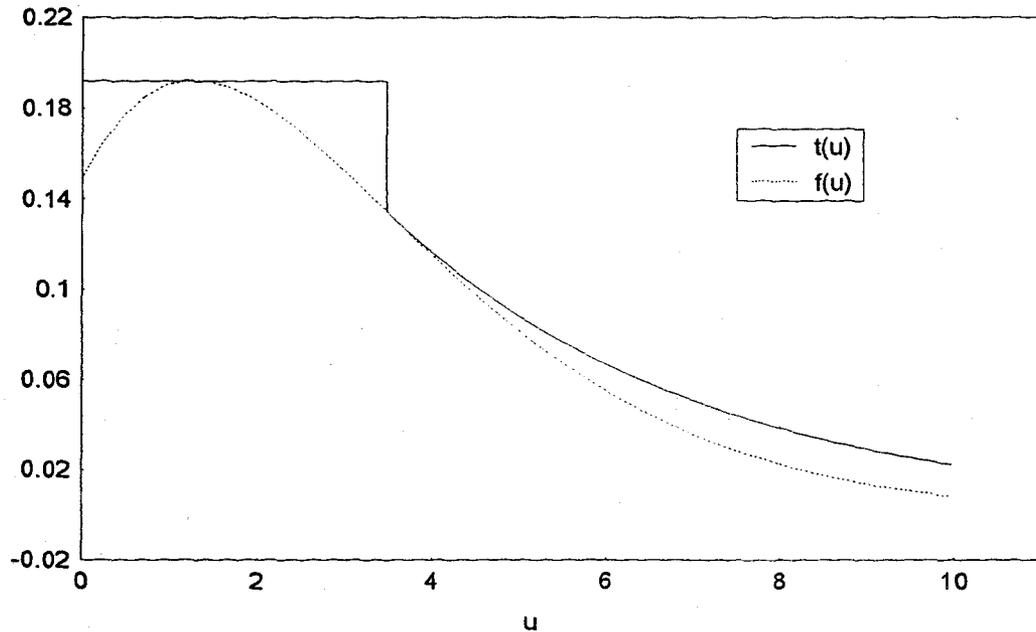


Figure 7.3. Fonction de densité et fonction majorante pour $\lambda = 3$.

7.2.4 $\lambda \geq 2$

Pour $\lambda \geq 2$, $f(u)$ présente une forme en cloche, avec deux points d'inflexion observés en $u = u_0$ et $u = u_2$, et son mode est atteint en $u = u_1$ ($u_0 < u_1 < u_2$). On cherche $t(u)$ exponentielle croissante sur $[0, u_0]$, constante sur $]u_0, u_2[$ et exponentielle décroissante sur $[u_2, \infty[$, et telle que sa courbe représentative tangente celle de f en $u = u_0$ et $u = u_2$, soit :

$$t(u) = \begin{cases} f(u_0) \exp\left[-\frac{f'(u_0)}{f(u_0)}(u_0 - u)\right], & u \leq u_0 \\ f(u_1), & u_0 < u < u_2 \\ f(u_2) \exp\left[\frac{f'(u_2)}{f(u_2)}(u - u_2)\right], & u \geq u_2 \end{cases} \quad (7.8)$$

Le calcul de c nous donne :

$$c = \frac{f(u_0)^2}{f'(u_0)} \left[1 - \exp\left(-u_0 \frac{f'(u_0)}{f(u_0)}\right) \right] + (u_2 - u_0) f(u_1) - \frac{f(u_2)^2}{f'(u_2)} \quad (7.9)$$

La figure (7.4) représente les fonctions $f(u)$ et $t(u)$ pour $\lambda=8$.

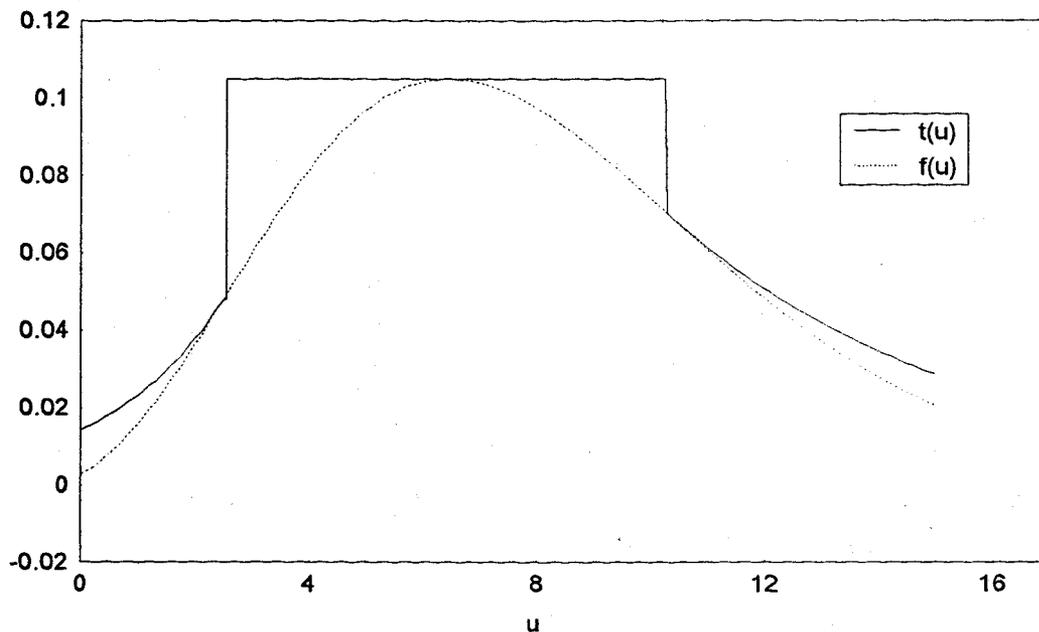


Figure 7.4. Fonction de densité et fonction majorante pour $\lambda = 8$.

REMARQUE : Pour réaliser les figures de ce chapitre, les maxima et les points d'inflexion des densités ont été calculés à l'aide du logiciel Mathematica, en déterminant les zéros respectifs des dérivées première et seconde des fonctions de densité.

7.2.5 Conclusion

Nous avons proposé une fonction majorante t pour la densité de la loi des fuites, que nous n'avons cependant pas testée. Il faudrait donc effectuer des essais afin de voir si l'expression donnée dans ce chapitre convient et éventuellement la modifier. On pourra pour cela se référer aux travaux concernant la loi de Halphen (Essadaoui *et al.* (1995)).

8 CONCLUSION ET RECOMMANDATIONS

Ce rapport effectue un rappel des propriétés mathématiques et statistiques de la loi des fuites. Nous avons présenté quatre méthodes d'estimation des paramètres, parmi lesquelles figurent deux méthodes classiques : la méthode du maximum de vraisemblance et la méthode des moments. Pour chacune des méthodes d'estimation, l'étude théorique des variances et des covariances asymptotiques des estimateurs a été effectuée. De plus, nous avons déterminé les estimateurs des quantiles x_T et les intervalles de confiance associés. A côté de ces développements théoriques ont été proposés une série d'algorithmes en vue d'introduire la loi des fuites dans le logiciel AJUSTE-II ainsi que trois applications de l'ajustement de cette loi à des ensembles de données. Enfin, nous avons suggéré la méthode d'acceptation-rejet pour la génération d'échantillons.

En tentant de faire le point sur la loi des fuites, ce travail devrait permettre de faciliter l'utilisation en pratique de cette loi de probabilité qui possède d'intéressantes propriétés. La présente étude est cependant loin d'être complète, et certains points mériteraient d'être approfondis. Il serait notamment intéressant de poursuivre l'étude de la relation entre la loi des fuites et la loi Gamma, en comparant par exemple leurs quantiles respectifs. D'autre part, afin de déterminer la méthode d'estimation des paramètres la mieux adaptée selon la proportion de valeurs nulles contenues dans l'échantillon, il est recommandé d'effectuer une simulation de Monte-Carlo et de comparer les biais et les variances des estimateurs des différentes méthodes, pour des valeurs de n et de n_0/n variables.

9 REFERENCES BIBLIOGRAPHIQUES

- Abramowitz M. et Stegun I.A. (1970) *Handbook of Mathematical Functions*. Dover Publications, Inc., New York. 9, 355-429
- Babuziaux C. (1969) 'Etude statistique de la loi des fuites'. Service des études et recherches nucléaires, thermiques et hydrauliques, Département laboratoire national d'hydraulique, Electricité de France.
- Bernier J. (1968) 'Sur les propriétés des lois $I\gamma$ '. Service des études et recherches nucléaires, thermiques et hydrauliques, Département laboratoire national d'hydraulique, Electricité de France.
- Bickel P.J. et Docksum K.A. (1977) *Mathematical Statistics*. Holden-Day, Inc., Oakland, California.
- Bobée B. et Ashkar F. (1991) *The Gamma Family and Derived Distributions Applied in Hydrology*. Water Resources Publications, Littleton, Colorado.
- Brunet-Moret Y. (1978) 'La loi des fuites'. Note provisoire, ORSTOM.
- Burden R.L. et Faires J.D. (1988) *Numerical Analysis, fourth Edition*. PWS-KENT Publishing Company, Boston. 2, 47-49.
- Einstein H.A. *The application of radioactive tracers in the study of sands movements in rivers*. Institute for development of water resources, Ioroslav Cerni, Beograd.
- Essadaoui M., Perreault L. et Bobée B. (1995) 'Générateur d'échantillons pour la loi de Halphen de type A'. INRS-Eau.
- Kendall M.G. et Stuart A. (1987) *Advanced Theory of Statistics. Volume 1 : Distribution Theory*. Oxford University Press, New York.
- Law A.M. et Kelton W.D. (1982) *Simulation modeling and analysis*. New York, McGraw-Hill Book Company.

Lubès H., Masson J.M., Raous P. et Tapiou M. (1994) 'Safarhy, Logiciel de calculs statistiques et d'analyse fréquentielle adapté à l'évaluation du risque en hydrologie, Manuel de référence', Orstom et Université Montpellier II.

Morlat 'La distribution des débits de fuite sur une conduite de gaz', EDF, Etudes économiques générales n° 1021.

Perreault L., Bobée B. et Legendre P. (1994) 'Rapport général du logiciel AJUSTE-II: Théorie et application'. Rapport de recherche R-421, INRS-Eau.

Ribstein P. 'Loi des fuites'. (1983) Cahier ORSTOM, série Hydrologie, Vol.XX,n°2. 2, 117-145.

ANNEXE A - PRINCIPE DE LA METHODE DE NEWTON-RAPHSON

La méthode de Newton-Raphson est une des méthodes numériques les plus puissantes et les plus employées pour calculer les racines d'une équation du type $f(x) = 0$.

Supposons que la fonction f est de classe C^2 , c'est-à-dire continûment dérivable sur un intervalle $[a; b]$. Soit p une racine de l'équation $f(x) = 0$ et \bar{x} une approximation de p telle que $f'(\bar{x})$ soit non nul et $|\bar{x} - p|$ " suffisamment petit ". On peut écrire le développement de Taylor de f au voisinage de \bar{x} :

$$f(x) = f(\bar{x}) + (x - \bar{x})f'(\bar{x}) + \frac{(x - \bar{x})^2}{2} f''(\xi(x)) \quad (\text{A.1})$$

avec $\xi(x) \in [x, \bar{x}]$ (ou $[\bar{x}, x]$).

En particulier pour $x = p$, on aura

$$0 = f(\bar{x}) + (p - \bar{x})f'(\bar{x}) + \frac{(p - \bar{x})^2}{2} f''(\xi(p)) \quad (\text{A.2})$$

Comme $|\bar{x} - p|$ est petit, on peut négliger le terme en $(p - \bar{x})^2$, pour écrire

$$0 \approx f(\bar{x}) + (p - \bar{x})f'(\bar{x}) \quad (\text{A.3})$$

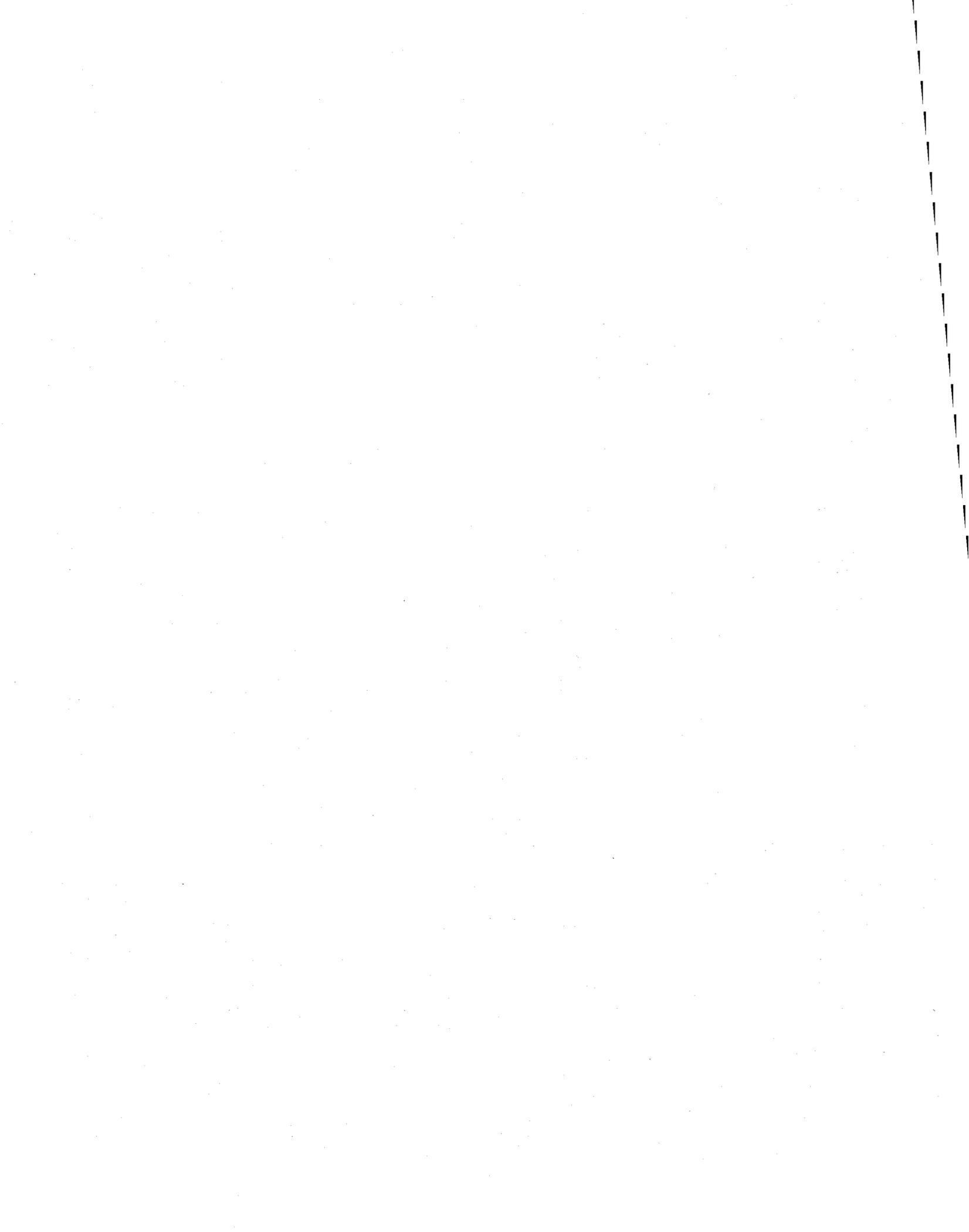
que l'on peut exprimer sous la forme suivante :

$$p = \bar{x} - \frac{f(\bar{x})}{f'(\bar{x})} \quad (\text{A.4})$$

C'est cette dernière équation qu'utilise la méthode de Newton-Raphson, en générant à partir d'une valeur initiale p_0 la suite $\{p_n\}$ définie comme suit :

$$p_n = p_{n-1} - \frac{f(p_{n-1})}{f'(p_{n-1})} \quad (\text{A.5})$$

On estime que $p_N = p$ lorsque $|p_N - p_{N-1}| < \varepsilon$, où ε est appelée tolérance et est fixée par l'utilisateur.



ANNEXE B - CALCUL DES COEFFICIENTS DE LA MATRICE DE FISHER

B.1. Expression à l'aide de la fonction $J(\lambda)$

On a vu au paragraphe 3.1.1 (équation (3.5)) que la fonction $\ln L$ était donnée par

$$\ln L = -\lambda n + \frac{(n-n_0)}{2} \ln\left(\frac{\lambda}{\beta}\right) - \sum_{i=1}^{n-n_0} \frac{x_i}{\beta} - \frac{1}{2} \sum_{i=1}^{n-n_0} \ln x_i + \sum_{i=1}^{n-n_0} \ln I_1\left(2\sqrt{\frac{\lambda x_i}{\beta}}\right)$$

- Calculons $E\left\{\frac{\partial^2 \ln L}{\partial \lambda^2}\right\}$:

En dérivant deux fois $\ln L$ par rapport à λ , nous obtenons :

$$\frac{\partial^2 \ln L}{\partial \lambda^2} = -\frac{n-n_0}{2\lambda^2} + \sum_{i=1}^{n-n_0} \left[\frac{\frac{\partial^2 I_1}{\partial \lambda^2}\left(2\sqrt{\frac{\lambda x_i}{\beta}}\right)}{I_1\left(2\sqrt{\frac{\lambda x_i}{\beta}}\right)} - \left(\frac{\frac{\partial I_1}{\partial \lambda}\left(2\sqrt{\frac{\lambda x_i}{\beta}}\right)}{I_1\left(2\sqrt{\frac{\lambda x_i}{\beta}}\right)} \right)^2 \right] \quad (\text{B.1})$$

Commençons par fixer n_0 . Nous travaillons avec deux variables aléatoires, X et Z . X peut prendre des valeurs positives ou nulles tandis que Z , qui est la restriction de X à $]0; \infty[$ est strictement positive. Si $\{x_1, \dots, x_n\}$ est un échantillon de valeurs prises par X alors on définit $\{z_1, \dots, z_{n-n_0}\}$ à partir des x_i par $z_i = x_i$ lorsque $x_i \neq 0$. n_0 est donc le nombre de valeurs nulles de l'échantillon $\{x_1, \dots, x_n\}$. Soient g et k les densités de probabilité respectives de X et Z , g vérifie alors :

$$P(X=0) + \int_0^{\infty} g(x) dx = e^{-\lambda} + \int_0^{\infty} g(x) dx = 1 \quad (\text{B.2})$$

On cherche k sous la forme Kg où K est une constante. Puisque k doit satisfaire :

$$\int_0^{\infty} k(z) dz = K \int_0^{\infty} g(x) dx = 1 \quad (\text{B.3})$$

on en déduit $K(1-e^{-\lambda}) = 1$ et donc que :

$$k(z) = \frac{1}{1-e^{-\lambda}} g(x) \quad (\text{B.4})$$

L'espérance à n_0 fixé sera donnée par l'équation suivante:

$$E_{n_0} \left\{ \frac{\partial^2 \ln L}{\partial \lambda^2} \right\} = -\frac{n-n_0}{2\lambda^2} + (n-n_0) \left[\int_0^\infty \frac{\frac{\partial^2 I_1(2\sqrt{\frac{\lambda x}{\beta}})}{\partial \lambda^2}}{I_1(2\sqrt{\frac{\lambda x}{\beta}})} \frac{g(x)}{1-e^{-\lambda}} dx - \int_0^\infty \left(\frac{\frac{\partial I_1(2\sqrt{\frac{\lambda x}{\beta}})}{\partial \lambda}}{I_1(2\sqrt{\frac{\lambda x}{\beta}})} \right)^2 \frac{g(x)}{1-e^{-\lambda}} dx \right] \quad (\text{B.5})$$

En posant $\alpha = \lambda\beta$, et d'après la définition des fonctions de Bessel donnée par (2.6),

$$I_1(2\sqrt{\frac{\lambda x}{\beta}}) = I_1(2\sqrt{\alpha x}) = \sqrt{\alpha x} \sum_{i=1}^{\infty} \frac{(\alpha x)^i}{i!(i+1)!} \quad (\text{B.6})$$

Les dérivées premières et secondes de I_1 par rapport à α , en posant $u = \alpha x/\lambda$ sont respectivement :

$$\frac{\partial I_1(2\sqrt{\alpha x})}{\partial \alpha} = \frac{\beta}{2} \sqrt{\frac{u}{\lambda}} \sum_{i=0}^{\infty} \frac{(2i+1)(\lambda u)^i}{i!(i+1)!} \quad (\text{B.7})$$

et

$$\frac{\partial^2 I_1(2\sqrt{\alpha x})}{\partial \alpha^2} = \frac{\beta^2}{4\lambda} \sqrt{\frac{u}{\lambda}} \sum_{i=0}^{\infty} \frac{(4i^2-1)(\lambda u)^i}{i!(i+1)!} \quad (\text{B.8})$$

Il faut calculer les intégrales suivantes, qui apparaissent dans (B.5) :

$$A = \int_0^\infty \frac{\frac{\partial^2 I_1(2\sqrt{\alpha x})}{\partial \alpha^2}}{I_1(2\sqrt{\alpha x})} \frac{g(x)}{1-e^{-\lambda}} dx \quad (\text{B.9})$$

$$B = \int_0^\infty \left(\frac{\frac{\partial I_1(2\sqrt{\alpha x})}{\partial \alpha}}{I_1(2\sqrt{\alpha x})} \right)^2 \frac{g(x)}{1-e^{-\lambda}} dx \quad (\text{B.10})$$

En se rappelant que $g(x)dx = f(u)du$, avec f définie par (2.8), et en utilisant le résultat suivant:

$$\int_0^{\infty} u^i e^{-u} du = i! \quad (\text{B.11})$$

on obtient :

$$A = \frac{\beta^2}{1-e^{-\lambda}} \left(1 - \frac{1}{\lambda} + \frac{3}{4\lambda^2} - \frac{3e^{-\lambda}}{4\lambda^2}\right) \quad (\text{B.12})$$

et

$$B = \frac{\beta^2 e^{-\lambda}}{4\lambda(1-e^{-\lambda})} J(\lambda) \quad (\text{B.13})$$

En remarquant que:

$$\frac{\partial I_1}{\partial \lambda} \left(2\sqrt{\frac{\lambda x}{\beta}}\right) = \frac{\partial I_1}{\partial a} (2\sqrt{ax}) \left(\frac{\partial a}{\partial \lambda}\right) = \frac{\partial I_1}{\partial a} (2\sqrt{ax}) \frac{1}{\beta} \quad (\text{B.14})$$

et

$$\frac{\partial^2 I_1}{\partial \lambda^2} \left(2\sqrt{\frac{\lambda x}{\beta}}\right) = \frac{\partial^2 I_1}{\partial a^2} (2\sqrt{ax}) \left(\frac{\partial a}{\partial \lambda}\right)^2 = \frac{\partial^2 I_1}{\partial a^2} (2\sqrt{ax}) \frac{1}{\beta^2} \quad (\text{B.15})$$

l'équation (B.5) devient:

$$E_{n_0} \left\{ \frac{\partial^2 \ln L}{\partial \lambda^2} \right\} = -\frac{n-n_0}{2\lambda^2} + (n-n_0) \frac{1}{\beta^2} [A-B] \quad (\text{B.16})$$

soit en remplaçant A et B par leurs expressions (B.12) et (B.13),

$$E_{n_0} \left\{ \frac{\partial^2 \ln L}{\partial \lambda^2} \right\} = \frac{n-n_0}{1-e^{-\lambda}} \left(1 - \frac{1}{\lambda} + \frac{1}{4\lambda^2} - \frac{e^{-\lambda}}{4\lambda^2} - \frac{e^{-\lambda}}{4\lambda} J(\lambda)\right) \quad (\text{B.17})$$

Pour calculer l'espérance lorsque n_0 varie à partir de l'espérance à n_0 fixé (équation (B.17)), on applique la formule suivante:

$$E\{Y\} = E_x\{E_y\{Y/X\}\} \quad (\text{B.18})$$

ce qui se traduit ici par:

$$E\left\{\frac{\partial^2 \ln L}{\partial \lambda^2}\right\} = E\left\{E_{n_0}\left\{\frac{\partial^2 \ln L}{\partial \lambda^2}\right\}\right\} = \frac{n - E\{n_0\}}{1 - e^{-\lambda}} \left(1 - \frac{1}{\lambda} + \frac{1}{4\lambda^2} - \frac{e^{-\lambda}}{4\lambda^2} - \frac{e^{-\lambda}}{4\lambda} J(\lambda)\right) \quad (\text{B.19})$$

Or, la probabilité d'avoir n_0 valeurs nulles dans un échantillon de n observations est:

$$P(n_0) = C_n^{n_0} P(0)^{n_0} (1 - P(0))^{n - n_0} \quad (\text{B.20})$$

ce qui correspond à une loi binomiale de moyenne $E\{n_0\} = nP(0) = ne^{-\lambda}$. On peut donc réécrire (B.19) sous la forme:

$$E\left\{\frac{\partial^2 \ln L}{\partial \lambda^2}\right\} = n \left(1 - \frac{1}{\lambda} + \frac{1}{4\lambda^2} - \frac{e^{-\lambda}}{4\lambda^2} - \frac{e^{-\lambda}}{4\lambda} J(\lambda)\right) \quad (\text{B.21})$$

• Calculons maintenant $E\left\{\frac{\partial^2 \ln L}{\partial \beta^2}\right\}$:

En dérivant deux fois $\ln L$ par rapport à β , nous obtenons :

$$\frac{\partial^2 \ln L}{\partial \beta^2} = \frac{n - n_0}{2\beta^2} - \frac{2}{\beta^3} \sum_{i=1}^{n - n_0} x_i + \sum_{i=1}^{n - n_0} \left[\frac{\frac{\partial^2 I_1}{\partial \beta^2} \left(2\sqrt{\frac{\lambda x_i}{\beta}}\right)}{I_1 \left(2\sqrt{\frac{\lambda x_i}{\beta}}\right)} - \left(\frac{\frac{\partial I_1}{\partial \beta} \left(2\sqrt{\frac{\lambda x_i}{\beta}}\right)}{I_1 \left(2\sqrt{\frac{\lambda x_i}{\beta}}\right)} \right)^2 \right] \quad (\text{B.22})$$

Comme précédemment, nous allons d'abord calculer l'espérance à n_0 fixé, donnée par l'équation suivante:

$$E_{n_0} \left\{ \frac{\partial^2 \ln L}{\partial \beta^2} \right\} = \frac{n - n_0}{2\beta^2} - \frac{2}{\beta^3} (n - n_0) E_{n_0} \{x\} + (n - n_0) \left[\int_0^\infty \frac{\frac{\partial^2 I_1}{\partial \beta^2} \left(2\sqrt{\frac{\lambda x}{\beta}}\right)}{I_1 \left(2\sqrt{\frac{\lambda x}{\beta}}\right)} \frac{g(x)}{1 - e^{-\lambda}} dx - \int_0^\infty \left(\frac{\frac{\partial I_1}{\partial \beta} \left(2\sqrt{\frac{\lambda x}{\beta}}\right)}{I_1 \left(2\sqrt{\frac{\lambda x}{\beta}}\right)} \right)^2 \frac{g(x)}{1 - e^{-\lambda}} dx \right] \quad (\text{B.23})$$

On remarque que :

$$\frac{\partial I_1}{\partial \beta} \left(2\sqrt{\frac{\lambda x}{\beta}}\right) = \frac{\partial I_1}{\partial a} (2\sqrt{ax}) \left(\frac{\partial a}{\partial \beta}\right) = \frac{\partial I_1}{\partial a} (2\sqrt{ax}) \left(\frac{-\lambda}{\beta^2}\right) \quad (\text{B.24})$$

et

$$\begin{aligned} \frac{\partial^2 I_1}{\partial \beta^2} (2\sqrt{\frac{\lambda x}{\beta}}) &= \frac{\partial^2 I_1(2\sqrt{ax})}{\partial a^2} \left(\frac{\partial a}{\partial \beta}\right)^2 + \frac{\partial I_1(2\sqrt{ax})}{\partial a} \left(\frac{\partial^2 a}{\partial \beta^2}\right) \\ &= \frac{\partial^2 I_1(2\sqrt{ax})}{\partial a^2} \frac{\lambda^2}{\beta^4} + \frac{\partial I_1(2\sqrt{ax})}{\partial a} \left(\frac{2\lambda}{\beta^3}\right) \end{aligned} \quad (\text{B.25})$$

Calculons l'intégrale suivante, qui apparait dans (B.25) :

$$C = \int_0^{\infty} \frac{\frac{\partial I_1}{\partial a}(2\sqrt{ax})}{I_1(2\sqrt{ax})} \frac{g(x)}{1-e^{-\lambda}} dx \quad (\text{B.26})$$

En se servant de (B.7), on trouve:

$$C = \frac{\beta}{1-e^{-\lambda}} \left(1 - \frac{1}{2\lambda} + \frac{e^{-\lambda}}{2\lambda}\right) \quad (\text{B.27})$$

En remarquant que d'après (B.4), $E_{n_0}\{x\} = E\{z\} = E\{x\}/(1-e^{-\lambda}) = \lambda\beta/(1-e^{-\lambda})$, et en introduisant les intégrales A , B et C définies respectivement par (B.9), (B.10) et (B.26), on peut écrire (B.23) de la façon suivante:

$$E_{n_0} \left\{ \frac{\partial^2 \ln L}{\partial \beta^2} \right\} = \frac{n-n_0}{2\beta^2} - \frac{2}{\beta^2} (n-n_0) \frac{\lambda}{1-e^{-\lambda}} + (n-n_0) \left(\frac{\lambda^2}{\beta^4} A - \frac{\lambda^2}{\beta^4} B + \frac{2\lambda}{\beta^3} \right) \quad (\text{B.28})$$

En utilisant ensuite pour A , B et C les expressions (B.12), (B.13) et (B.27), on a que

$$E_{n_0} \left\{ \frac{\partial^2 \ln L}{\partial \beta^2} \right\} = \frac{n-n_0}{\beta^2} \left[\frac{1}{2} + \frac{1}{1-e^{-\lambda}} \left(\lambda^2 - \lambda - \frac{1}{4} + \frac{e^{-\lambda}}{4} - \frac{\lambda}{4} e^{-\lambda} J(\lambda) \right) \right] \quad (\text{B.29})$$

A l'aide de (B.18) et (B.20), on obtient finalement l'espérance suivante :

$$E \left\{ \frac{\partial^2 \ln L}{\partial \beta^2} \right\} = \frac{n}{\beta^2} \left(\lambda^2 - \lambda + \frac{1}{4} - \frac{e^{-\lambda}}{4} - \frac{\lambda}{4} e^{-\lambda} J(\lambda) \right) \quad (\text{B.30})$$

- Calculons à présent $E \left\{ \frac{\partial^2 \ln L}{\partial \lambda \partial \beta} \right\}$:

En dérivant $\ln L$ par rapport à λ puis par rapport à β , nous obtenons :

$$\frac{\partial^2 \ln L}{\partial \lambda \partial \beta} = \sum_{i=1}^{n-n_0} \left[\frac{\frac{\partial^2 I_1}{\partial \lambda \partial \beta} \left(2\sqrt{\frac{\lambda x_i}{\beta}}\right)}{I_1 \left(2\sqrt{\frac{\lambda x_i}{\beta}}\right)} - \frac{\frac{\partial I_1}{\partial \lambda} \left(2\sqrt{\frac{\lambda x_i}{\beta}}\right) \frac{\partial I_1}{\partial \beta} \left(2\sqrt{\frac{\lambda x_i}{\beta}}\right)}{\left(I_1 \left(2\sqrt{\frac{\lambda x_i}{\beta}}\right)\right)^2} \right] \quad (\text{B.31})$$

Comme précédemment, nous allons d'abord calculer l'espérance à n_0 fixé de l'expression (B.31) pour obtenir :

$$E_{n_0} \left\{ \frac{\partial^2 \ln L}{\partial \lambda \partial \beta} \right\} = (n - n_0) \times \left[\int_0^{\infty} \frac{\frac{\partial^2 I_1}{\partial \lambda \partial \beta} \left(2\sqrt{\frac{\lambda x}{\beta}}\right)}{I_1 \left(2\sqrt{\frac{\lambda x}{\beta}}\right)} \frac{g(x)}{1 - e^{-\lambda}} dx - \int_0^{\infty} \frac{\frac{\partial I_1}{\partial \lambda} \left(2\sqrt{\frac{\lambda x}{\beta}}\right) \frac{\partial I_1}{\partial \beta} \left(2\sqrt{\frac{\lambda x}{\beta}}\right)}{\left(I_1 \left(2\sqrt{\frac{\lambda x}{\beta}}\right)\right)^2} \frac{g(x)}{1 - e^{-\lambda}} dx \right] \quad (\text{B.32})$$

Les équations (B.14), (B.24) et l'équation suivante :

$$\begin{aligned} \frac{\partial^2 I_1}{\partial \lambda \partial \beta} \left(2\sqrt{\frac{\lambda x}{\beta}}\right) &= \frac{\partial^2 I_1(2\sqrt{ax})}{\partial a^2} \left(\frac{\partial a}{\partial \lambda}\right) \left(\frac{\partial a}{\partial \beta}\right) + \frac{\partial I_1(2\sqrt{ax})}{\partial a} \left(\frac{\partial^2 a}{\partial \lambda \partial \beta}\right) \\ &= \frac{\partial^2 I_1(2\sqrt{ax})}{\partial a^2} \left(-\frac{\lambda}{\beta^3}\right) + \frac{\partial I_1(2\sqrt{ax})}{\partial a} \left(-\frac{1}{\beta^2}\right) \end{aligned} \quad (\text{B.33})$$

nous permettent d'écrire (B.32) au moyen des intégrales A , B et C définies respectivement par (B.9), (B.10) et (B.26) :

$$E_{n_0} \left\{ \frac{\partial^2 \ln L}{\partial \lambda \partial \beta} \right\} = (n - n_0) \left(-\frac{\lambda}{\beta^3} A + \frac{\lambda}{\beta^3} B - \frac{1}{\beta^2} C \right) \quad (\text{B.34})$$

Cette équation devient, en utilisant pour A , B et C les expressions (B.12), (B.13) et (B.27),

$$E_{n_0} \left\{ \frac{\partial^2 \ln L}{\partial \lambda \partial \beta} \right\} = \frac{n - n_0}{\beta(1 - e^{-\lambda})} \left(\lambda - \frac{1}{4\lambda} + \frac{e^{-\lambda}}{4\lambda} + \frac{e^{-\lambda}}{4} J(\lambda) \right) \quad (\text{B.35})$$

A l'aide de (B.18) et (B.20), on obtient finalement l'espérance suivante :

$$E \left\{ \frac{\partial^2 \ln L}{\partial \lambda \partial \beta} \right\} = \frac{n}{\beta} \left(-\lambda - \frac{1}{4\lambda} + \frac{e^{-\lambda}}{4\lambda} + \frac{e^{-\lambda}}{4} J(\lambda) \right) \quad (\text{B.36})$$

B.2. Expression à l'aide des fonctions de Bessel

- Calculons $E\left\{\frac{\partial^2 \ln L}{\partial \lambda^2}\right\}$:

Reprenons l'équation (B.5) de l'annexe B-1,

$$E_{n_0}\left\{\frac{\partial^2 \ln L}{\partial \lambda^2}\right\} = -\frac{n-n_0}{2\lambda^2} + (n-n_0) \left[\int_0^\infty \frac{\frac{\partial^2 I_1(2\sqrt{\lambda x})}{\partial \lambda^2}}{I_1(2\sqrt{\lambda x})} \frac{g(x)}{1-e^{-\lambda}} dx - \int_0^\infty \left(\frac{\frac{\partial I_1(2\sqrt{\lambda x})}{\partial \lambda}}{I_1(2\sqrt{\lambda x})} \right)^2 \frac{g(x)}{1-e^{-\lambda}} dx \right]$$

En utilisant la propriété (3.10) des fonctions de Bessel, et en posant $u = x/\beta$ on calcule les dérivées première et seconde de I_1 par rapport à λ . Elles s'expriment respectivement par :

$$\frac{\partial I_1(2\sqrt{\lambda u})}{\partial \lambda} = \sqrt{\frac{u}{\lambda}} I_0(2\sqrt{\lambda u}) - \frac{1}{2\lambda} I_1(2\sqrt{\lambda u}) \quad (\text{B.37})$$

et

$$\frac{\partial^2 I_1(2\sqrt{\lambda u})}{\partial \lambda^2} = -\frac{1}{\lambda} \sqrt{\frac{u}{\lambda}} I_0(2\sqrt{\lambda u}) + \left(\frac{u}{\lambda} + \frac{3}{4\lambda^2}\right) I_1(2\sqrt{\lambda u}) \quad (\text{B.38})$$

Comme $g(x)dx = f(u)du$, on obtient que :

$$E_{n_0}\left\{\frac{\partial^2 \ln L}{\partial \lambda^2}\right\} = \frac{n-n_0}{\lambda} \left[-\frac{1}{2\lambda} + \int_0^\infty \left(u + \frac{1}{2\lambda} - u \frac{I_0^2(2\sqrt{\lambda u})}{I_1^2(2\sqrt{\lambda u})}\right) \frac{f(u)}{1-e^{-\lambda}} du \right] \quad (\text{B.39})$$

Le calcul de l'espérance lorsque n_0 varie à partir de l'espérance à n_0 fixé se fait comme au 1. à l'aide des résultats (B.18) et (B.20) et nous donne :

$$E\left\{\frac{\partial^2 \ln L}{\partial \lambda^2}\right\} = \frac{n}{\lambda} \left[\frac{e^{-\lambda} - 1}{2\lambda} + \int_0^\infty \left(u + \frac{1}{2\lambda} - u \frac{I_0^2(2\sqrt{\lambda u})}{I_1^2(2\sqrt{\lambda u})}\right) f(u) du \right] \quad (\text{B.40})$$

Pour calculer l'intégrale suivante, qui apparait dans (B.40) :

$$\int_0^\infty u f(u) du = \frac{1}{\beta} \int_0^\infty x g(x) dx = \frac{1}{\beta} E\{X\} \quad (\text{B.41})$$

on se sert de l'estimateur \bar{x} de $E\{X\}$. Puisqu'on a choisi β (cf. section 3.1.2, équation (3.9)) tel que :

$$\bar{x} = \lambda\beta \quad (\text{B.42})$$

on obtient que :

$$\int_0^{\infty} u f(u) du = \lambda \quad (\text{B.43})$$

D' autre part on sait d'après (B.2) que :

$$\int_0^{\infty} f(u) du = 1 - e^{-\lambda} \quad (\text{B.44})$$

En exprimant la densité $f(u)$ à l'aide des fonctions de Bessel, et en utilisant les résultats (B.42) et (B.43), l'équation (B.40) se met alors sous la forme

$$E\left\{\frac{\partial^2 \ln L}{\partial \lambda^2}\right\} = n \left(1 - \frac{e^{-\lambda}}{\sqrt{\lambda}} \int_0^{\infty} \sqrt{u} e^{-u} \frac{I_0^2(2\sqrt{\lambda u})}{I_1(2\sqrt{\lambda u})} du\right) \quad (\text{B.45})$$

• Calculons maintenant $E\left\{\frac{\partial^2 \ln L}{\partial \beta^2}\right\}$:

Reprenons l'équation (B.23) de l'annexe B-1,

$$E_{n_0}\left\{\frac{\partial^2 \ln L}{\partial \beta^2}\right\} = \frac{n - n_0}{2\beta^2} - \frac{2}{\beta^3} (n - n_0) E_{n_0}\{x\} + (n - n_0) \left[\int_0^{\infty} \frac{\frac{\partial^2 I_1(2\sqrt{\frac{\lambda x}{\beta}})}{\partial \beta^2}}{I_1(2\sqrt{\frac{\lambda x}{\beta}})} \frac{g(x)}{1 - e^{-\lambda}} dx - \int_0^{\infty} \left(\frac{\frac{\partial I_1(2\sqrt{\frac{\lambda x}{\beta}})}{\partial \beta}}{I_1(2\sqrt{\frac{\lambda x}{\beta}})} \right)^2 \frac{g(x)}{1 - e^{-\lambda}} dx \right]$$

En utilisant la propriété (3.10) des fonctions de Bessel, on calcule les dérivées première et seconde de I_1 par rapport à β . Elles s'expriment respectivement de la manière suivante :

$$\frac{\partial I_1(2\sqrt{\frac{\lambda x}{\beta}})}{\partial \beta} = -\frac{1}{\beta} \sqrt{\frac{\lambda x}{\beta}} I_0(2\sqrt{\frac{\lambda x}{\beta}}) + \frac{1}{2\beta} I_1(2\sqrt{\frac{\lambda x}{\beta}}) \quad (\text{B.46})$$

et

$$\frac{\partial^2 I_1(2\sqrt{\frac{\lambda x}{\beta}})}{\partial \beta^2} = \frac{1}{\beta^2} \sqrt{\frac{\lambda x}{\beta}} I_0(2\sqrt{\frac{\lambda x}{\beta}}) + \frac{1}{\beta^2} \left(\frac{\lambda x}{\beta} - \frac{1}{4}\right) I_1(2\sqrt{\frac{\lambda x}{\beta}}) \quad (\text{B.47})$$

En introduisant les expressions (B.46) et (B.47) dans l'équation (B.23) et en remarquant que d'après (B.4), $E_{n_0}\{x\} = E\{z\} = E\{x\}/(1-e^{-\lambda}) = \lambda\beta/(1-e^{-\lambda})$, on obtient l'équation suivante :

$$E_{n_0} \left\{ \frac{\partial^2 \ln L}{\partial \beta^2} \right\} = \frac{n - n_0}{2\beta^2} - \frac{2n\lambda}{\beta^2} + \frac{n - n_0}{\beta^2} \int_0^{\infty} \left[\frac{\lambda x}{\beta} - \frac{1}{2} + 2\sqrt{\frac{\lambda x}{\beta}} \frac{I_0(2\sqrt{\frac{\lambda x}{\beta}})}{I_1(2\sqrt{\frac{\lambda x}{\beta}})} - \frac{\lambda x}{\beta} \frac{I_0^2(2\sqrt{\frac{\lambda x}{\beta}})}{I_1^2(2\sqrt{\frac{\lambda x}{\beta}})} \right] \frac{g(x)}{1 - e^{-\lambda}} dx \quad (\text{B.48})$$

Le calcul de l'espérance lorsque n_0 varie à partir de l'espérance à n_0 fixé se fait comme au 1. à l'aide des équations (B.18) et (B.20). En utilisant les résultats suivants déjà vus :

$$\int_0^{\infty} x g(x) dx = E\{X\} \approx \bar{x} = \lambda\beta \quad (\text{B.49})$$

et

$$\int_0^{\infty} g(x) dx = 1 - e^{-\lambda} \quad (\text{B.50})$$

nous pouvons finalement établir l'équation suivante :

$$E \left\{ \frac{\partial^2 \ln L}{\partial \beta^2} \right\} = \frac{n(1 - e^{-\lambda})}{2\beta^2} - \frac{2n\lambda}{\beta^2} + \frac{n}{\beta^2} \left[\lambda^2 - \frac{1 - e^{-\lambda}}{2} + 2\frac{\lambda}{\beta} e^{-\lambda} \int_0^{\infty} e^{-x/\beta} I_0(2\sqrt{\frac{\lambda x}{\beta}}) dx \right] - \frac{\lambda n}{\beta^3} e^{-\lambda} \int_0^{\infty} \sqrt{\frac{\lambda x}{\beta}} e^{-x/\beta} \frac{I_0^2(2\sqrt{\frac{\lambda x}{\beta}})}{I_1^2(2\sqrt{\frac{\lambda x}{\beta}})} dx \quad (\text{B.51})$$

Pour calculer l'intégrale D , qui apparait dans (B.51) et qui est définie par :

$$D = \int_0^{\infty} e^{-x/\beta} I_0\left(2\sqrt{\frac{\lambda x}{\beta}}\right) dx \quad (\text{B.52})$$

on intègre par parties l'équation (B.52), puis on utilise la propriété des fonctions de Bessel suivante :

$$I_0' = I_1 \quad (\text{B.53})$$

On se sert ensuite du résultat donné par l'équation (B.50) pour obtenir enfin :

$$D = \beta e^\lambda \quad (\text{B.54})$$

En injectant ce résultat dans (B.51) et en posant $u = x/\beta$, il vient que :

$$E\left\{\frac{\partial^2 \ln L}{\partial \beta^2}\right\} = \frac{\lambda n}{\beta^2} (\lambda + e^{-\lambda} \sqrt{\lambda} \int_0^{\infty} \sqrt{u} e^{-u} \frac{I_0^2(2\sqrt{\lambda u})}{I_1^2(2\sqrt{\lambda u})} du) \quad (\text{B.55})$$

• Calculons maintenant $E\left\{\frac{\partial^2 \ln L}{\partial \lambda \partial \beta}\right\}$:

Reprenons l'équation (B.32) de l'annexe B-1,

$$E_{n_0}\left\{\frac{\partial^2 \ln L}{\partial \lambda \partial \beta}\right\} = (n - n_0) \left[\int_0^{\infty} \frac{\frac{\partial^2 I_1}{\partial \lambda \partial \beta}\left(2\sqrt{\frac{\lambda x}{\beta}}\right)}{I_1\left(2\sqrt{\frac{\lambda x}{\beta}}\right)} \frac{g(x)}{1 - e^{-\lambda}} dx - \int_0^{\infty} \frac{\frac{\partial I_1}{\partial \lambda}\left(2\sqrt{\frac{\lambda x}{\beta}}\right) \frac{\partial I_1}{\partial \beta}\left(2\sqrt{\frac{\lambda x}{\beta}}\right)}{\left(I_1\left(2\sqrt{\frac{\lambda x}{\beta}}\right)\right)^2} \frac{g(x)}{1 - e^{-\lambda}} dx \right]$$

En utilisant la propriété (3.10) des fonctions de Bessel et en posant $u = x/\beta$, on calcule la dérivée suivante :

$$\frac{\partial^2 I_1}{\partial \lambda \partial \beta}\left(2\sqrt{\frac{\lambda x}{\beta}}\right) = \frac{\partial^2 I_1}{\partial \lambda \partial \beta}(2\sqrt{\lambda u}) = \frac{1}{\beta} \left(-u - \frac{1}{4\lambda}\right) I_1(2\sqrt{\lambda u}) \quad (\text{B.56})$$

Multiplions membre à membre les équations (B.38) et (B.46) pour établir l'égalité qui suit, exprimée en fonction de u :

$$\begin{aligned} \frac{\partial I_1}{\partial \lambda}(2\sqrt{\lambda u}) \frac{\partial I_1}{\partial \beta}(2\sqrt{\lambda u}) &= -\frac{1}{\beta} \left[u I_0^2(2\sqrt{\lambda u}) \right. \\ &\quad \left. + \frac{1}{\beta} \left[\sqrt{\frac{u}{\lambda}} I_0(2\sqrt{\lambda u}) I_1(2\sqrt{\lambda u}) - \frac{1}{4\lambda} I_1^2(2\sqrt{\lambda u}) \right] \right] \quad (\text{B.57}) \end{aligned}$$

En injectant ces résultats dans (B.33) et en se servant de (B.43) nous trouvons :

$$E_{n_0} \left\{ \frac{\partial^2 \ln L}{\partial \lambda \partial \beta} \right\} = \frac{1}{\beta} \frac{n - n_0}{1 - e^{-\lambda}} \left[-\lambda + e^{-\lambda} \sqrt{\lambda} \int_0^{\infty} e^{-u} \sqrt{u} \frac{I_0^2(2\sqrt{\lambda u})}{I_1(2\sqrt{\lambda u})} du - e^{-\lambda} \int_0^{\infty} e^{-u} I_0(2\sqrt{\lambda u}) du \right] \quad (\text{B.58})$$

En remarquant l'égalité suivante :

$$\int_0^{\infty} e^{-u} I_0(2\sqrt{\lambda u}) du = \frac{D}{\beta} \quad (\text{B.59})$$

où D est l'intégrale définie par (B.52), on a d'après (B.54),

$$\int_0^{\infty} e^{-u} I_0(2\sqrt{\lambda u}) du = e^{\lambda} \quad (\text{B.60})$$

Le calcul de l'espérance lorsque n_0 varie à partir de l'espérance à n_0 fixé se fait ensuite comme au 1. à l'aide des résultats (B.18) et (B.20) et nous donne :

$$E \left\{ \frac{\partial^2 \ln L}{\partial \lambda \partial \beta} \right\} = \frac{n}{\beta} (-\lambda - 1 + e^{-\lambda} \sqrt{\lambda} \int_0^{\infty} e^{-u} \sqrt{u} \frac{I_0^2(2\sqrt{\lambda u})}{I_1(2\sqrt{\lambda u})} du) \quad (\text{B.61})$$



ANNEXE C - DEFINITION DE L'EFFICACITE D'UN ESTIMATEUR

Soit Σ la matrice des variances et covariance des paramètres λ et β ,

$$\Sigma = \begin{bmatrix} \text{Var}\{\hat{\lambda}\} & \text{Cov}\{\hat{\lambda}, \hat{\beta}\} \\ \text{Cov}\{\hat{\lambda}, \hat{\beta}\} & \text{Var}\{\hat{\beta}\} \end{bmatrix} \quad (\text{C.1})$$

Soit I la matrice des variances et covariance de $\frac{\partial \ln L}{\partial \lambda}$ et $\frac{\partial \ln L}{\partial \beta}$,

$$I = \begin{bmatrix} \text{Var}\left(\frac{\partial \ln L}{\partial \lambda}\right) & \text{Cov}\left(\frac{\partial \ln L}{\partial \lambda}, \frac{\partial \ln L}{\partial \beta}\right) \\ \text{Cov}\left(\frac{\partial \ln L}{\partial \lambda}, \frac{\partial \ln L}{\partial \beta}\right) & \text{Var}\left(\frac{\partial \ln L}{\partial \beta}\right) \end{bmatrix} \quad (\text{C.2})$$

L'efficacité E d'un estimateur est alors définie par :

$$E = \frac{1}{|I| |V|} \quad (\text{C.3})$$

D'après cette définition, l'efficacité des estimateurs du maximum de vraisemblance est égale à 1. Pour les autres méthodes, on cherchera à déterminer les estimateurs pour lesquels la courbe E est la plus proche possible de 1.

ANNEXE D - CALCUL DES DERIVEES PARTIELLES DU QUANTILE DE PERIODE DE RETOUR T PAR RAPPORT A CHACUN DES PARAMETRES DE LA LOI

Nous avons vu au Chapitre 4, pour la loi des fuites, que le quantile x_T de période de retour T est défini par :

$$e^{-\lambda} + \int_0^{x_T} g(x) dx = 1 - \frac{1}{T} \quad (\text{D.1})$$

• Calculons $\frac{\partial x_T}{\partial \lambda}$:

En posant $u = x / \beta$, l'équation (D.1) devient

$$e^{-\lambda} + \int_0^{u_T} f(u) du = 1 - \frac{1}{T} \quad (\text{D.2})$$

Nous dérivons l'équation (D.2) par rapport à λ pour obtenir :

$$-e^{-\lambda} + \frac{\partial}{\partial \lambda} \left(\int_0^{u_T} f(u) du \right) = 0 \quad (\text{D.3})$$

Le théorème de Leibniz sur la différentiation des intégrales nous permet d'écrire :

$$\frac{\partial}{\partial \lambda} \left(\int_0^{u_T} f(u) du \right) = \int_0^{u_T} \frac{\partial}{\partial \lambda} f(u) du + f(u_T) \frac{\partial u_T}{\partial \lambda} \quad (\text{D.4})$$

Nous en déduisons que :

$$\frac{\partial u_T}{\partial \lambda} = \frac{1}{f(u_T)} \left(e^{-\lambda} - \int_0^{u_T} \frac{\partial}{\partial \lambda} f(u) du \right) \quad (\text{D.5})$$

La densité f est donnée (voir Chapitre 2, équation (2.8)) par :

$$f(u) = \lambda \exp(-\lambda - u) \frac{I_1(2\sqrt{\lambda u})}{\sqrt{\lambda u}} \quad (\text{D.6})$$

En dérivant cette expression par rapport à λ et en utilisant la propriété des dérivées des fonctions de Bessel (3.10), on obtient :

$$\frac{\partial}{\partial \lambda} f(u) = \exp(-\lambda - u) I_0(2\sqrt{\lambda u}) - f(u) \quad (\text{D.7})$$

Nous pouvons donc écrire, en faisant appel à l'équation (D.2), que :

$$\int_0^{u_T} \frac{\partial}{\partial \lambda} f(u) du = e^{-\lambda} \int_0^{u_T} e^{-u} I_0(2\sqrt{\lambda u}) du - \left(1 - \frac{1}{T}\right) + e^{-\lambda} \quad (\text{D.8})$$

Par intégration par parties et en se souvenant que $I_0' = I_1$ et $I_0(0) = 1$, nous calculons l'intégrale suivante :

$$\int_0^{u_T} e^{-u} I_0(2\sqrt{\lambda u}) du = -e^{-u_T} I_0(2\sqrt{\lambda u_T}) + e^{-\lambda} \left(1 - \frac{1}{T}\right) \quad (\text{D.9})$$

que nous reportons dans l'équation (D.8) pour trouver

$$\int_0^{u_T} \frac{\partial}{\partial \lambda} f(u) du = e^{-\lambda} (-e^{-u_T} I_0(2\sqrt{\lambda u_T}) + 1) \quad (\text{D.10})$$

En injectant (D.10) dans (D.5) et en exprimant $f(u_T)$ au moyen de (D.6), nous obtenons finalement :

$$\frac{\partial u_T}{\partial \lambda} = \sqrt{\frac{u_T}{\lambda}} \frac{I_0(2\sqrt{\lambda u_T})}{I_1(2\sqrt{\lambda u_T})} \quad (\text{D.11})$$

qui s'exprime en fonction de x de la façon suivante :

$$\frac{\partial x_T}{\partial \lambda} = \sqrt{\frac{\beta x_T}{\lambda}} \frac{I_0\left(2\sqrt{\frac{\lambda x_T}{\beta}}\right)}{I_1\left(2\sqrt{\frac{\lambda x_T}{\beta}}\right)} \quad (\text{D.12})$$

- Calculons $\frac{\partial x_T}{\partial \beta}$:

La méthode est identique à celle employée dans le calcul de $\partial x_T / \partial \lambda$: en dérivant (D.1) par rapport à β , puis en utilisant le théorème de Leibniz sur la différentiation des intégrales, nous pouvons écrire :

$$\frac{\partial x_T}{\partial \beta} = \frac{-1}{g(x_T)} \int_0^{x_T} \frac{\partial}{\partial \beta} g(x) dx \quad (\text{D.13})$$

La densité g est donnée (voir Chapitre 2, équation (2.7)) par :

$$g(x) = \frac{\lambda}{\beta} \exp\left(-\lambda - \frac{x}{\beta}\right) \frac{I_1\left(2\sqrt{\frac{\lambda x}{\beta}}\right)}{\sqrt{\frac{\lambda x}{\beta}}} \quad (\text{D.14})$$

En dérivant cette expression par rapport à β et en utilisant la propriété des dérivées des fonctions de Bessel (3.10), on obtient :

$$\frac{\partial}{\partial \beta} g(x) = \frac{1}{\beta^2} \exp\left(-\lambda - \frac{x}{\beta}\right) \left[\sqrt{\frac{\lambda x}{\beta}} I_1\left(2\sqrt{\frac{\lambda x}{\beta}}\right) - \lambda I_0\left(2\sqrt{\frac{\lambda x}{\beta}}\right) \right] \quad (\text{D.15})$$

Nous pouvons donc écrire, en faisant appel à l'équation (D.1)

$$\int_0^{x_T} \frac{\partial}{\partial \beta} g(x) dx = \frac{e^{-\lambda}}{\beta} \left[\sqrt{\lambda} \int_0^{x_T} \frac{1}{\beta} e^{-\frac{x}{\beta}} \sqrt{\frac{x}{\beta}} I_1\left(2\sqrt{\frac{\lambda x}{\beta}}\right) dx - \lambda \int_0^{x_T} \frac{1}{\beta} e^{-\frac{x}{\beta}} I_0\left(2\sqrt{\frac{\lambda x}{\beta}}\right) dx \right] \quad (\text{D.16})$$

Effectuons le changement de variable $u = x / \beta$ pour obtenir

$$\int_0^{x_T} \frac{1}{\beta} e^{-\frac{x}{\beta}} \sqrt{\frac{x}{\beta}} I_1\left(2\sqrt{\frac{\lambda x}{\beta}}\right) dx = \int_0^{u_T} e^{-u} \sqrt{u} I_1(2\sqrt{\lambda u}) du \quad (\text{D.17})$$

En intégrant par parties (D.17), et en utilisant la propriété des dérivées des fonctions de Bessel (3.10), nous retrouvons les intégrales des équations (D.2) et (D.9), pour aboutir finalement au résultat suivant :

$$\int_0^{x_T} \frac{1}{\beta} e^{-\frac{x}{\beta}} \sqrt{\frac{x}{\beta}} I_1\left(2\sqrt{\frac{\lambda x}{\beta}}\right) dx = -e^{-u_T} \sqrt{u_T} I_1(2\sqrt{\lambda u_T}) - \sqrt{\lambda} e^{-u_T} I_0(2\sqrt{\lambda u_T}) + \sqrt{\lambda} e^{-\lambda} \left(1 - \frac{1}{T}\right) \quad (\text{D.18})$$

En remarquant que

$$\int_0^{x_T} \frac{1}{\beta} e^{-\frac{x}{\beta}} I_0\left(2\sqrt{\frac{\lambda x}{\beta}}\right) dx = \int_0^{u_T} e^{-u} I_0(2\sqrt{\lambda u}) du \quad (\text{D.19})$$

nous retombons sur une intégrale connue dont le résultat est donné par (D.9), et nous pouvons revenir à l'équation (D.16) qui s'écrit maintenant :

$$\int_0^{x_T} \frac{\partial}{\partial \beta} g(x) dx = -\frac{\sqrt{\lambda} e^{-\lambda}}{\beta} e^{-\frac{x_T}{\beta}} \sqrt{\frac{x_T}{\beta}} I_1\left(2\sqrt{\frac{\lambda x_T}{\beta}}\right) \quad (\text{D.20})$$

En injectant ce dernier résultat dans (D.13) et en exprimant $g(x_T)$ au moyen de (D.14), on déduit finalement que :

$$\frac{\partial x_T}{\partial \beta} = \frac{\hat{x}_T}{\beta} \quad (\text{D.21})$$

ANNEXE E - PROGRAMMATION DANS MATLAB ET MATHEMATICA

E.1. Estimation des paramètres par la méthode du maximum de vraisemblance

On résoud l'équation (3.12), qui est de la forme $h(\lambda) = 0$ (voir sections 3.1.2 et 3.1.3) pour calculer $\hat{\lambda}$. On déduit ensuite $\hat{\beta}$ grâce à l'équation (3.13).

Programmation dans Matlab

E.1.1 Méthode de Newton-Raphson

fonction [L,B] = maxvrai (x , n , no , Lmom)

----- x est un vecteur de taille n-no, de valeurs strictement positives -----

Initialisation des variables

```
v=Lmom;
moy=mean(x)*(n-no)/n;
yy=sqrt(x).*BESSELI(0,2.*v.*sqrt(x./moy))./BESSELI(1,2.*v.*sqrt(x./moy));
y=sum(yy);
zz=yy.^2;
z=sum(zz);
gprim=2/sqrt(moy)*(sum(x)-z)+1/v*y;
p=sqrt(x).*BESSELI(0,2.*v.*sqrt(x./moy))./BESSELI(1,2.*v.*sqrt(x./moy));
g=sum(p)-n*sqrt(moy);
u=v-g/gprim;
```

```
while abs(u-v) >= 1e-6
```

```
    v=u;
```

```
    Calcul de la dérivée de h, notée gprim (équation (3.14)):
```

```
        yy=sqrt(x).*BESSELI(0,2.*v.*sqrt(x./moy))./BESSELI(1,2.*v.*sqrt(x./moy));
```

```
        y=sum(yy);
```

```
        zz=yy.^2;
```

```
        z=sum(zz);
```

```
        gprim=2/sqrt(moy)*(sum(x)-z)+1/v*y;
```

```
    Calcul de la fonction h, notée g (équation (3.12)):
```

```
        p=sqrt(x).*BESSELI(0,2.*v.*sqrt(x./moy))./BESSELI(1,2.*v.*sqrt(x./moy));
```

```
        g=sum(p)-n*sqrt(moy);
```

```
        u=v-g/gprim;
```

```
end
```

```
L=u
```

```
B=moy/L
```

E.1.2 Méthode par dichotomie

fonction [L,B] = maxvrai (x , n , no , Lmom)

----- x est un vecteur de taille n-no, de valeurs strictement positives -----

Initialisation des variables

a=1.e-4;

b=5*Lmom;

u=(a+b)/2;

moy=mean(x)*(n-no)/n;

y=sqrt(x).*BESSELI(0,2.*u.*sqrt(x./moy))./BESSELI(1,2.*u.*sqrt(x./moy));

f=sum(y)-n*sqrt(moy);

while abs(f) > 0.000001

 if f < 0

 b=u;

 u=(a+b)/2;

Calcul de la fonction h, notée f (équation (3.12)):

 y=sqrt(x).*BESSELI(0,2.*u.*sqrt(x./moy))./BESSELI(1,2.*u.*sqrt(x./moy));

 f=sum(y)-n*sqrt(moy);

 else

 a=u;

 u=(a+b)/2;

Calcul de la fonction h, notée f (équation (3.12)):

 y=sqrt(x).*BESSELI(0,2.*u.*sqrt(x./moy))./BESSELI(1,2.*u.*sqrt(x./moy));

 f=sum(y)-n*sqrt(moy);

 end

end

L=u

B=moy/L

E.2. Calcul des coefficients de la matrice de Fisher

E.2.1 Calcul de la fonction $J(\lambda)$

La fonction $J(\lambda)$ est définie au paragraphe 3.1.4 par l'équation (3.20).

Programmation dans Matlab

L'équation (3.20) peut s'écrire sous la forme suivante :

$$J(\lambda) = \int_0^{\infty} N(\lambda, u) du = \int_0^{\infty} e^{-\lambda} \frac{KK(\lambda, u)}{P(\lambda, u)} du \quad (\text{E.1})$$

Calcul de la fonction $KK(\lambda, u)$:

```
function [ KK ] = numerat ( l , u )
V=1;
I=1;
K=1;
V=V*I*u*(2*I+1)/(I*(I+1)*(2*I-1));
K=K+V;
while V/K >= 1e-6
    I=I+1;
    V=V*I*u*(2*I+1)/(I*(I+1)*(2*I-1));
    K=K+V;
end
KK=K^2;
```

Calcul de la fonction $P(\lambda, u)$:

```
function [ P ] = denomin ( l , u )
V=1;
I=1;
P=1;
V=V*I*u/I/(I+1);
P=P+V;
while V/P >= 1e-6
    I=I+1;
    V=V*I*u/I/(I+1);
    P=P+V;
end
```

Calcul de la fonction $N(\lambda, u)$:

```
function [ N ] = a_integ ( l , u )
N=exp(-u)*numerat ( l , u )/denomin ( l , u );
```

Calcul de la fonction $J(\lambda)$:

```
function [ J ] = fonct_J ( l , pas )
J=0;
u=0;
N=a_integ ( l , u );
while N*pas/J >= 0.00001
    J=J+N*pas;
    u=u+pas;
    N=a_integ ( l , u );
end
```

Programmation dans Mathématique

```
Numerat[u_l_i_] := ((2i+1)*(l*u)^i) / (i*(i+1)!)
```

```
Denom[u_,l_,i_]:=((l*u)^i)/(i*(i+1)!)
Rapp[u_,l_]:=((1+NSum[Numera[u,l,i],{i,1,Infinity}])^2)*Exp[-u]
/(1+NSum[Denom[u,l,i],{i,1,Infinity}])
```

```
FonctJ[l_]:=NIntegrate[Rapp[u,l],{u,0,40}]
```

E.2.2 Expression à l'aide des fonctions de Bessel

Les coefficients de la matrice de Fisher sont exprimés à partir des fonctions de Bessel par les équations (3.21), (3.22) et (3.23).

Programmation dans Matlab

```
function [ N ] = a_integ2 ( l , u )
N=exp(-u)*sqrt(u)*(BESSELI(0,2*sqrt(l*u))^2/BESSELI(1,2*sqrt(l*u)));
```

```
function [ If11 , If12 , If22 ] = fisher_b ( l , b , n , pas )
J=0;
u=0.001;
N=a_integ2 ( l , u );
while N*pas/J >= 0.00001
    J=J+N*pas;
    u=u+pas;
    N=a_integ2 ( l , u );
end
```

```
If11=n*(-1+exp(-l)/sqrt(l)*J)
If12=n/b*(1+1-exp(-l)*sqrt(l)*J)
If22=l*n/b^2*(-l+exp(-l)*sqrt(l)*J)
```

Programmation dans Mathématique

```
Fon[u_,l_]:=Sqrt[u]*Exp[-u]*(BesselI[0,2*Sqrt[l*u]])^2 /
(BesselI[1,2*Sqrt[l*u]])
```

```
FonctA[l_]:=NIntegrate[Fon[u,l],{u,0,60}]
```

```
Iff11[l_,n_]:=n*(-1+(Exp[-l]/Sqrt[l])*FonctA[l])
Iff12[l_,n_,b_]:=n/b*(1+l-(Exp[-l]*Sqrt[l])*FonctA[l])
Iff22[l_,n_,b_]:=l*n/(b^2)*(-l+(Exp[-l]*Sqrt[l])*FonctA[l])
```

E.3. Détermination du quantile x_T

E.3.1. Calcul de la fonction de répartition $F(u)$

On se référera à la section 5.4.1.

Programmation dans Matlab

- PREMIERE METHODE (programmation directe de l'équation (2.13))

```

function [ F ] = Reparti1 ( u , l )
    j=2;
    t=l;
    k=1;
    v=1;
    w=1;
    v=v.*u./k;
    w=w+v;
    k=k+1;
    z=w.*l^j./prod(1:j);
    while z./t > 1e-6
        t=t+z;
        j=j+1;
        v=v.*u./k;
        w=w+v;
        k=k+1;
        z=w.*l^j./prod(1:j);
    end
    F=1-exp(-l-u).*t;

```

- SECONDE METHODE (algorithme proposé par Lubès *et al.*(1994))

```

function [ F ] = Reparti2 ( u , l )
    G=exp(l/2);
    A=G-1/G;
    B=exp(-u/2);
    C=1;
    D=A*B;
    E=1;
    B=B*u/E;
    C=C/l/E;
    A=A-C/G;
    D=D+A*B;
    while A*B/D >= 1e-6
        E=E+1;

```

```

B=B*u/E;
C=C*l/E;
A=A-C/G;
D=D+A*B;
end
F=1-D*exp(-u/2)/G;

```

Programmation dans Mathematica

- PREMIERE METHODE (programmation directe de l'équation (2.13))

```

Densite[u_,l_]:=Exp[-l-u]*Sqrt[l/u]*BesselI[1,2*Sqrt[l*u]]
Repartition[u_,l_]:=Exp[-l]+NIntegrate[Densite[t,l],{t,0.00001,u}]

```

- SECONDE METHODE (algorithme proposé par Lubès *et al.* (1994))

```

f[u_,j_]:=u^j / j!
S1[u_,l_,i_]:=NSum[f[u,j],{j,0,i}] * l^(i+1)/(i+1)!
Repartition[u_,l_]:=1-Exp[-l-u]*NSum[S1[u,l,i],{i,0,Infinity}]

```

E.3.2. Calcul de x_T

Voir section 5.4.2.

Programmation dans Mathematica

Exemple : Calcul du quantile de période de retour $T=10$ ans pour $\lambda=1.2618$.

```

FindRoot[Repartition[u,1.2618]==0.9,{u,{2,5}}]
FindRoot[Repartition[u,1.2618]==0.9,{u,2}]

```

Programmation dans Matlab

- METHODE DE NEWTON-RAPHSON

Calcul de la densité, notée D :

```

function [ D ] = Densite ( v , l )
D=exp(-l-v)*sqrt(l/v)*BESSELI(1,2*sqrt(l*v));

```

```

function [ X ] = Quantil2 ( p , l , xo )

```

```

v=xo

```

```

F= repartit(v,l)-p

```

```

D= densite(v,l)

```

```

u=v-F/D

```

```

while abs(u-v) >= 1e-6

```

```
v=u
F= repartit(v,l)-p
D= densite(v,l)
u=v-F/D
end
X=u
```

- METHODE PAR DICHOTOMIE

```
function [ X ] = Quantil1 ( p , l , beta )
b=0;
c=0;
i=0;
k=4;
u=l;
F= repartit(u,l);
while abs(F-p)./p >= 1e-6
    if F >= p
        b=u;
        u=(u+c)./2;
        k=1;
        F= repartit(u,l);
        i=i+1;
        if i > 1000, break, end
    else
        c=u;
        u=(k.*u+b)./2;
        F= repartit(u,l);
        i=i+1;
        if i > 1000, break, end
    end
end
X=u;
```