

**Record Number:**  
**Author, Monographic:** Paradis, M.//Bobée, B.  
**Author Role:**  
**Title, Monographic:** La distribution gamma généralisée et son application en hydrologie  
**Translated Title:**  
**Reprint Status:**  
**Edition:**  
**Author, Subsidiary:**  
**Author Role:**  
**Place of Publication:** Québec  
**Publisher Name:** INRS-Eau  
**Date of Publication:** 1983  
**Original Publication Date:**  
**Volume Identification:**  
**Extent of Work:** vi, 52  
**Packaging Method:** pages  
**Series Editor:**  
**Series Editor Role:**  
**Series Title:** INRS-Eau, Rapport de recherche  
**Series Volume ID:** 156  
**Location/URL:**  
**ISBN:** 2-89146-154-1  
**Notes:** Rapport annuel 1983-1984  
**Abstract:** 10.00\$  
**Call Number:** R000156  
**Keywords:** rapport/ ok/ dl

INRS-Eau  
Université du Québec  
C.P. 7500, Sainte-Foy  
Québec G1V 4C7

Rapport scientifique No 156  
1983

par

Marc Paradis et Bernard Bobée

La distribution gamma généralisée  
et son application en hydrologie

La distribution gamma généralisée  
et son application en hydrologie

Rapport scientifique No 156  
1983

par

Marc Paradis et Bernard Bobée

INRS-Eau  
Université du Québec  
C.P. 7500, Sainte-Foy  
Québec G1V 4C7

## Table des matières

	Page
Table des matières .....	i
Liste des tableaux .....	iii
Liste des figures .....	iv
Introduction .....	v
Chapitre 1: Caractéristiques générales de la fonction de densité de la loi gamma généralisée .....	1
1.1 Généralités .....	1
1.2 Propriétés mathématiques .....	3
Chapitre 2: Propriétés statistiques de la distribution gamma généralisée .....	17
2.1 Moments .....	17
2.2 Coefficients: $C_V$ , $C_S$ , $C_K$ .....	18
Chapitre 3: Estimation des paramètres par la méthode du maximum de vraisemblance de l'échantillon .....	25
3.1 Estimation des paramètres .....	25
Chapitre 4: Estimation des paramètres par la méthode des moments .....	30
4.1 Propriétés d'estimation de la méthode des moments ...	30
4.2 Estimation des paramètres .....	30

	<u>Page</u>
Chapitre 5: Calcul de la variance de l'évènement $X_T$ .....	35
5.1 Introduction .....	35
5.2 Calcul des variances et covariances des paramètres estimés par la méthode du maximum de vraisemblance ..	38
5.3 Calcul des variances et covariances des paramètres estimés par la méthode des moments .....	38
Conclusion .....	44
Références bibliographiques .....	46
Annexe 1: Calcul de la fonction de distribution de la loi G.G. à 3 paramètres .....	49
Annexe 2: Calcul de la variable standardisée gamma simple en utilisant un développement en série pour la fonction gamma incomplète .....	51

## Liste des tableaux

	Page
1. Détails relatifs aux fonctions de la figure 1 (Moments, coefficients, positions du mode et des points d'inflexion, ordonnée à l'origine) .....	9
2. Détails relatifs aux fonctions de la figure 2 .....	11
3. Détails relatifs aux fonctions de la figure 3 .....	13
4. Détails relatifs aux fonctions de la figure 4 .....	15
5. Exemple d'amplification d'erreur lors du calcul de la variable $x$ en fonction de $S$ , pour $\lambda = 1$ et pour 9 probabilités au non dépassement .....	39

Liste des figures

	Page
1. Quelques fonctions de densité de probabilité gamma généralisée pour des asymétries variant de -0,5 à 1,0 .	8
2. Influence du paramètre $\lambda$ sur la forme de la loi gamma généralisée .....	10
3. Influence du paramètre S sur la forme de la loi gamma généralisée .....	12
4. Influence du paramètre d'échelle a pour la loi gamma généralisée .....	14
5. Graphe du coefficient de variation $C_V$ en fonction du paramètre S pour quelques valeurs de $\lambda$ .....	19
6. Graphe du coefficient d'asymétrie $C_S$ en fonction du paramètre S pour quelques valeurs de $\lambda$ .....	20
7. Graphe du coefficient d'aplatissement $C_K$ en fonction de S pour quelques valeurs de $\lambda$ .....	22
8. La fonction de vraisemblance $U(S)$ pour $-10 \leq S \leq 10$ ...	29

## Introduction

Diverses fonctions de densité de probabilité (f.d.p.) ont été proposées afin d'effectuer des analyses statistiques sur des débits en rivière, dans le but de prévoir les crues extrêmes (ou les débits résiduels les plus faibles) dans le cadre des aménagements hydrauliques en général.

Dans ce rapport, on présente la f.d.p. gamma généralisée et on décrit deux méthodes classiques d'estimation des paramètres : La méthode des moments et la méthode du maximum de vraisemblance. Dans chacun des cas la variance de l'évènement de période de retour  $T$  est calculée; on peut ainsi comparer les deux méthodes d'estimation des paramètres et en déduire les intervalles de confiance, correspondant à un niveau donné de probabilité.

Les deux méthodes d'estimation des paramètres font appel à des processus itératifs qui sont développés en détail et qui assurent la convergence même lors de l'analyse d'échantillons de tailles réduites. Le calcul des variances et covariance des paramètres, ainsi que des intervalles de confiance de l'évènement de période de retour  $T$  est comparé pour les deux méthodes d'estimation. Les aspects numériques font que le traitement informatique de la f.d.p. gamma généralisée nécessite un travail important, la plupart des équations théoriques apparaissant dans ce rapport devant d'ailleurs être modifiées dans le but de les exprimer en termes de sommes de logarithmes pour pouvoir manipuler les nombreuses fonctions gamma.

La loi gamma généralisée semble être un outil très puissant pour l'étude statistique d'échantillons de tailles et d'asymétries les plus diverses. La distorsion plus ou moins grande de la forme qui est apportée par le paramètre  $S$  donne à cette loi une grande versatilité pour s'adapter à la réalité dont on doit rendre compte.

Ce rapport contient quelques résultats préliminaires provenant de génération d'échantillons aléatoires, qui tendent à démontrer une efficacité légèrement supérieure de la méthode d'estimation des paramètres selon le maximum de vraisemblance.

L'essentiel de ce travail permet donc de répondre aux questions importantes et pratiques pour l'étude des débits de crue.

- Quelle est la probabilité d'obtenir un débit plus grand (ou plus faible) qu'un débit  $Q$  fixé?
- Dans quel intervalle de confiance doit-on s'attendre de trouver le débit  $Q_T$  correspondant à une période de retour  $T$ ?

Ce travail constitue une partie d'une étude plus générale portant sur l'examen des modèles statistiques pour l'estimation des débits de crue; les fonds utilisés dans le cadre de ce rapport proviennent de la F.C.A.C. que nous tenons à remercier, ainsi que Hélène Boucher, mathématicienne, pour ses remarques constructives, également André Parent et Anne Provencher qui ont respectivement exécuté les tâches reliées au graphisme et à la dactylographie.

CHAPITRE 1

CARACTÉRISTIQUES GÉNÉRALES DE LA FONCTION DENSITÉ DE PROBABILITÉ DE LA LOI  
GAMMA GÉNÉRALISÉE

1.1 Généralités

La forme générale de la fonction densité de probabilité de la loi gamma généralisée à quatre paramètres peut s'écrire:

$$f(x;a,\lambda,s,m) = \frac{|s|}{\Gamma(\lambda) |a|} e^{-[(x-m)/a]^s} \left(\frac{x-m}{a}\right)^{s\lambda-1} \quad (1.1)$$

où a est un paramètre d'échelle  
 $\lambda$  est un paramètre de forme  
 S est un paramètre de distorsion de la forme  
 m est un paramètre de position

$\Gamma(\lambda)$  est la fonction gamma telle que:

$$\Gamma(\lambda) = \int_0^{\infty} e^{-x} x^{\lambda-1} dx$$

Les paramètres a, s et m peuvent théoriquement être positifs ou négatifs, alors que  $\lambda$  est toujours positif.

L'intervalle de définition de la variable x est tel que:

$$(x-m)/a \geq 0 \quad \text{c'est-à-dire que: } \begin{array}{l} \text{-si } a > 0 \text{ alors } x \geq m \\ \text{-si } a < 0 \text{ alors } x \leq m \end{array}$$

en effectuant dans (1.1) le changement de variable

$W = [(x-m)/a]^s$  on obtient la fonction densité g(W) de la variable réduite W:

$$g(W) = \frac{e^{-W} W^{\lambda-1}}{\Gamma(\lambda)} \quad W \geq 0, \lambda > 0 \quad (1.2)$$

qui représente la loi gamma à 1 paramètre (gamma simple).

Harter (1966) a étudié les propriétés de cette distribution dans le cas où  $S$ ,  $a$  et  $m$  sont positifs et  $x \geq m \geq 0$ . La fonction densité est alors:

$$f(x; a, \lambda, s, m) = \frac{s}{a} \frac{e^{-[(x-m)/a]^s}}{\Gamma(\lambda)} \left(\frac{x-m}{a}\right)^{s\lambda-1} \quad (1.3)$$

Lorsque  $S=1$ ,  $a$  pouvant être positif ou négatif, on retrouve à partir de (1.1) le cas particulier de la distribution Pearson type III (Bobée, 1975) avec asymétrie positive (si  $a > 0$ ) ou négative (si  $a < 0$ ).

Tadikamalla (1979) a développé une méthode de génération de variables aléatoires pour la loi gamma généralisée à quatre paramètres, cette méthode a d'ailleurs été utilisée dans le cadre de cet ouvrage pour les tests d'estimation des paramètres et des événements de période de retour  $T$ . Cet algorithme de génération semble être à la fois simple et rapide : C'est à partir de deux nombres aléatoire de distribution uniforme dans l'intervalle (0,1) qu'une variable gamma généralisée est créée, avec un taux de rejection de l'ordre de 20% cependant.

Stacy et Mihram (1965) ont considéré la loi gamma généralisée à 3 paramètres qui constitue un cas particulier de (1.1) lorsque le paramètre d'origine est égal à 0. Ces auteurs considèrent le cas où  $a$  est positif et  $S$  peut avoir un signe quelconque mais est non-nul, ils obtiennent ainsi la forme suivante de fonction densité:

$$f(x; a, \lambda, S) = \frac{|s| x^{S\lambda-1} e^{-(x/a)^S}}{a^{S\lambda} \Gamma(\lambda)} \quad (1.4)$$

avec  $x \geq 0$

Lorsque  $S$  est positif, la fonction (1.4) comprend plusieurs cas particuliers (exponentielle, gamma, Weibull, Rayleigh, chi-deux, ---) qui sont décrits par Stacy et Mihram (1965).

C'est la généralité de cette distribution qui en fait un outil puissant et souple pour l'ajustement d'un échantillon représentant une variable hydrologique.

Le calcul de la fonction de distribution de la loi gamma généralisée à 3 paramètres est décrit dans l'annexe 1, cette fonction nécessite l'évaluation d'une fonction gamma incomplète; il existe quelques méthodes indirectes à cet effet (Bobée, B. et al., 1983) dont celle qui est décrite dans l'annexe 2 et qui fait appel à un développement en série.

## 1.2 Propriétés mathématiques de la f.d.p. gamma généralisée

Dans tout ce qui suit, on considérera le cas de la loi gamma généralisée à 3 paramètres ( $m=0$ ) car les échantillons qu'on analyse en hydrologie statistique sont constitués de débits positifs.

En posant  $f(x;a,\lambda,S,0) = f(x;a,\lambda,S)$  on obtient:

$$f(x;a,\lambda,S) = \frac{|s| e^{-[x/a]^S} x^{S\lambda-1}}{a^{S\lambda} \Gamma(\lambda)} \quad (1.5)$$

avec  $x \geq 0$ ,  $a$  et  $\lambda$  sont positifs alors que  $S$  peut être positif ou négatif mais il ne doit pas être nul.

-----

L'examen de la fonction densité montre que si:

- $S\lambda > 1$  alors  $f(0;a,\lambda,S) = 0$
- $S\lambda = 1$  alors  $f(0;a,\lambda,S) = \frac{|S|}{a \Gamma(\lambda)}$
- $0 < S\lambda < 1$  alors  $f(0;a,\lambda,S) = +\infty$
- $S < 0$  alors  $f(0;a,\lambda,S) = 0$

Le dernier cas où  $S < 0$  n'est pas évident à l'examen, il convient d'exprimer l'équation 1.5 sous forme logarithmique, on a:

$$\ln f(x;a,\lambda,S) = \ln |S| - (x/a)^S + (S\lambda - 1) \ln x - S\lambda \ln a - \ln \Gamma(\lambda) \quad (1.6)$$

on peut alors montrer que le terme  $(x/a)^S$  augmente plus rapidement que le terme  $(S\lambda - 1) \log x$  lorsque  $x \rightarrow 0$

donc  $\ln f(x;a,\lambda,S) = -\infty$  et  $f(0;a,\lambda,S) = 0$  pour  $S < 0$   
 $x \rightarrow 0$

-----

### Mode

Il s'agit de déterminer  $x_m$  tel que  $f(x_m;a,\lambda,S)$  corresponde au maximum de la fonction densité. La première dérivée de l'équation 1.5 par rapport à  $x$  est:

$$\frac{d}{dx} f(x;a,\lambda,S) = \frac{|S|}{a^{S\lambda} \Gamma(\lambda)} e^{-(x/a)^S} x^{S\lambda-2} [S\lambda-1-S(x/a)^S] \quad (1.7)$$

Il y a un mode lorsque cette expression devient égale à 0, à ce moment

$$S\lambda - 1 = S (x_m/a)^S, \text{ donc } x_m = a \frac{(S\lambda-1)^{1/S}}{S} = a (\lambda-1/S)^{1/S} \quad (1.8)$$

lorsque  $0 < S\lambda \leq 1$  il n'y a pas de mode.

-----

### Points d'inflexion

Un point d'inflexion apparaît lorsque la dérivée seconde devient nulle.

$$\frac{d^2}{dx^2} f(x;a,\lambda,S) = \frac{|S| e^{-(x/a)^S} x^{S\lambda-3}}{a^{S\lambda} \Gamma(\lambda)} \{2+S^2\lambda^2-3S\lambda-S^2(x/a)^S-2\lambda S^2(x/a)^S+3S(x/a)^S+S^2(x/a)^{2S}\} = 0 \quad (1.9)$$

cette condition peut encore s'exprimer de la façon suivante:

$$S^2(x/a)^{2S} + [3S-2\lambda S^2-S^2] (x/a)^S + [2-3S\lambda+S^2\lambda^2] = 0 \quad (1.10)$$

les racines de cette équation, lorsqu'elles existent, sont données par l'expression suivante:

$$x = a \left\{ \frac{(S+2\lambda S-3) \pm (S^2(1+4\lambda) - 6S+1)^{\frac{1}{2}}}{2S} \right\}^{1/S} \quad (1.11)$$

Ces racines sont réelles si le discriminant  $\Delta$  est positif ou nul  $S^2(1+4\lambda) - 6S + 1 \geq 0$ , on doit également pouvoir élever l'expression entre accolades à la puissance  $1/S$ .

Si  $\Delta < 0$  il n'y a pas de point d'inflexion. Si  $\Delta = 0$  il y a un seul point d'inflexion à condition que  $\lambda \geq \frac{3-S}{2S}$ , en cas d'égalité le point se trouve à l'origine.

Si  $\Delta > 0$  il y a deux points seulement si  $\lambda S \geq 2$  ou  $S < 0$ , sinon il n'y en a qu'un seul.

Ces conditions peuvent se résumer ainsi, en posant  $G = \frac{6S-1-S^2}{4S^2}$

- 1°  $\lambda < G$  aucun point réel d'inflexion  
 2°  $\lambda = G$  1 point si  $\lambda \geq \frac{3-S}{2S}$ , sinon aucun point.

Ces deux premier cas sont possibles seulement si  $3 - \sqrt{32}/2 < S < 3 + \sqrt{32}/2$ .

- 3°  $\lambda > G$  2 points si  $\lambda S \geq 2$  ou  $S < 0$  (en cas d'égalité un point se trouve à l'origine). Sinon il n'y a qu'un seul point réel d'inflexion.

La valeur maximale de  $G$  est 2, donc si  $\lambda > 2$  et ( $S \geq \frac{2}{\lambda}$  ou  $S < 0$ ) il y a deux points d'inflexion.

De plus si ( $S > 3 + \sqrt{32}/2$  ou  $S < 3 - \sqrt{32}/2$ ), alors  $\lambda$  est toujours plus grand que  $G$  et il suffit que ( $\lambda \geq \frac{2}{S}$  ou  $S < 0$ ) pour qu'il y ait la présence de deux points d'inflexion.

### Analyse graphique

La figure 1 montre l'allure de quelques fonctions de densité de la loi gamma généralisée pour des asymétries variant de -0,5 à 1,0 la courbe 5 est à asymétrie positive, les détails relatifs à ces courbes se trouvent sur le tableau 1. On montrera au chapitre 2 la théorie se rapportant aux moments et aux coefficients  $C_V$ ,  $C_S$  et  $C_K$ . Nos expériences nous ont démontré que la loi gamma généralisée possédait son meilleur potentiel d'utilisation pour  $C_S > 0$ , il est possible cependant d'obtenir des fonctions G.G. avec  $C_S$  de l'ordre de -2 mais alors le choix des coefficients de variation est très limité. On reviendra sur ces aspects plus loin.

La figure 2 montre 4 courbes qui ne diffèrent que par le paramètre  $\lambda$ ; à mesure que  $\lambda$  augmente la moyenne augmente et les courbes sont décalées vers la droite tout en devenant plus symétriques, la valeur de  $f(x)$  au mode diminue également. L'influence du paramètre  $\lambda$  est encore plus visible sur le tableau 2 qui contient les détails relatifs aux courbes de la figure 2.

Dans le même ordre d'idées, la figure 3 montre l'influence du paramètre  $S$  sur la forme de quelques fonctions G.G. La courbe 1 a dégénéré en exponentielle car  $\lambda S = 1$ , l'asymétrie passe de 2 à  $-0.087$  alors que  $S$  varie de 1 à 4. Les détails relatifs à ces courbes se trouvent sur le tableau 3, on se rend compte de l'absence de quelques points d'inflexion. Le paramètre  $S$  agit de façon marquée sur la forme de la loi gamma généralisée, on peut le considérer comme un paramètre de distorsion de la forme; ainsi l'action de l'exposant  $S$  peut transformer une loi gamma simple avec  $C_S = 4$  en une loi gamma généralisée d'asymétrie égale à  $-2$ .

On peut également considérer le paramètre  $S$  comme un paramètre d'échelle au sens logarithmique. Les figures 2 et 3 semblent démontrer que  $S$  est le paramètre majeur et que son introduction augmente de beaucoup les possibilités et le potentiel des f.d.p. de la famille gamma, au détriment toutefois de la simplicité du traitement mathématique.

L'influence du paramètre d'échelle  $a$  est montrée sur la figure 4, comme on peut le voir sur le tableau 4 toutes ces courbes possèdent les mêmes  $C_V$ ,  $C_S$  et  $C_K$ . Le rôle du paramètre  $a$  se fait sentir essentiellement au niveau de la moyenne et de la variance; la moyenne augmentant et  $f(x_{\text{MODE}})$  diminuant pour une augmentation de  $a$ .

---

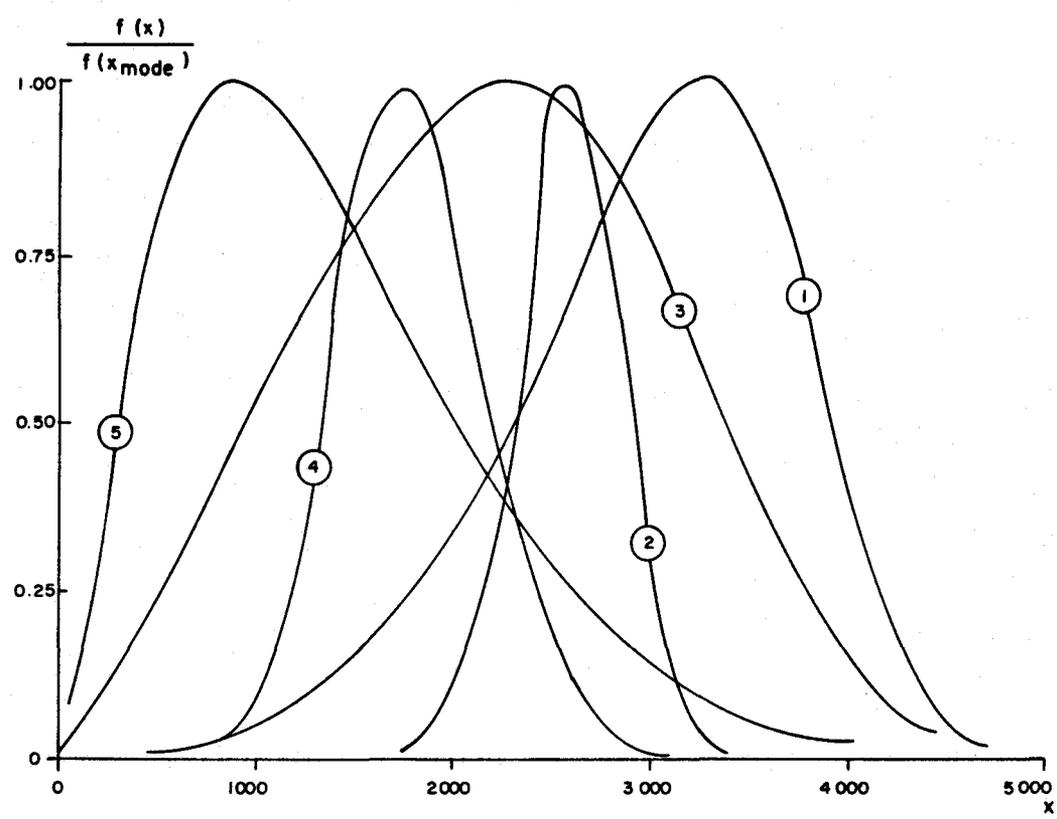


Figure 1: Quelques fonctions de densité G.G. pour des asymétries variant de -0,5 à 1,0.

(voir le Tableau 1 pour les détails relatifs à ces fonctions)

Tableau 1: Détails relatifs aux fonctions de la figure 1

	1	2	3	4	5
a	3707,0	2032,0	3141,0	750,3	811,8
$\lambda$	0,5	4,0	0,5	6,0	2,0
S	8,0	5,2	4,2	2,0	1,2
$\log \mu'_1$	3,477	3,415	3,342	3,255	3,146
$\log \mu'_2$	6,977	6,834	6,751	6,528	6,422
$\log \mu'_3$	10,495	10,257	10,203	9,818	9,794
$\log \mu'_4$	14,027	13,685	13,686	13,124	13,241
$\log \mu_2$	5,698	4,830	5,905	5,138	5,835
$\log  \mu_3 $	8,262	6,548	7,390	7,040	8,782
$\log \mu_4$	11,873	10,145	12,205	10,755	12,332
$C_V$	0,235	0,100	0,407	0,206	0,590
$C_S$	-0,519	-0,200	0,034	0,214	1,069
$C_K$	2,999	3,048	2,477	3,005	4,594
f(o)	0	0	0	0	0
MODE					
x	3 279	2 627	2 283	1 759	923
f(x)*	0,58	1,54	0,41	1,07	0,55
I					
x	3 920	2 886	3 331	2 136	1 644
N					
F					
f(x)*	0,30	0,90	0,22	0,67	0,39
L					
E					
x	2 607	2 368	883	1 381	182
X					
I					
f(x)*	0,40	0,96	0,19	0,62	0,15
O					
N					

\*  $f(x)^* = f(x) \times 1000$

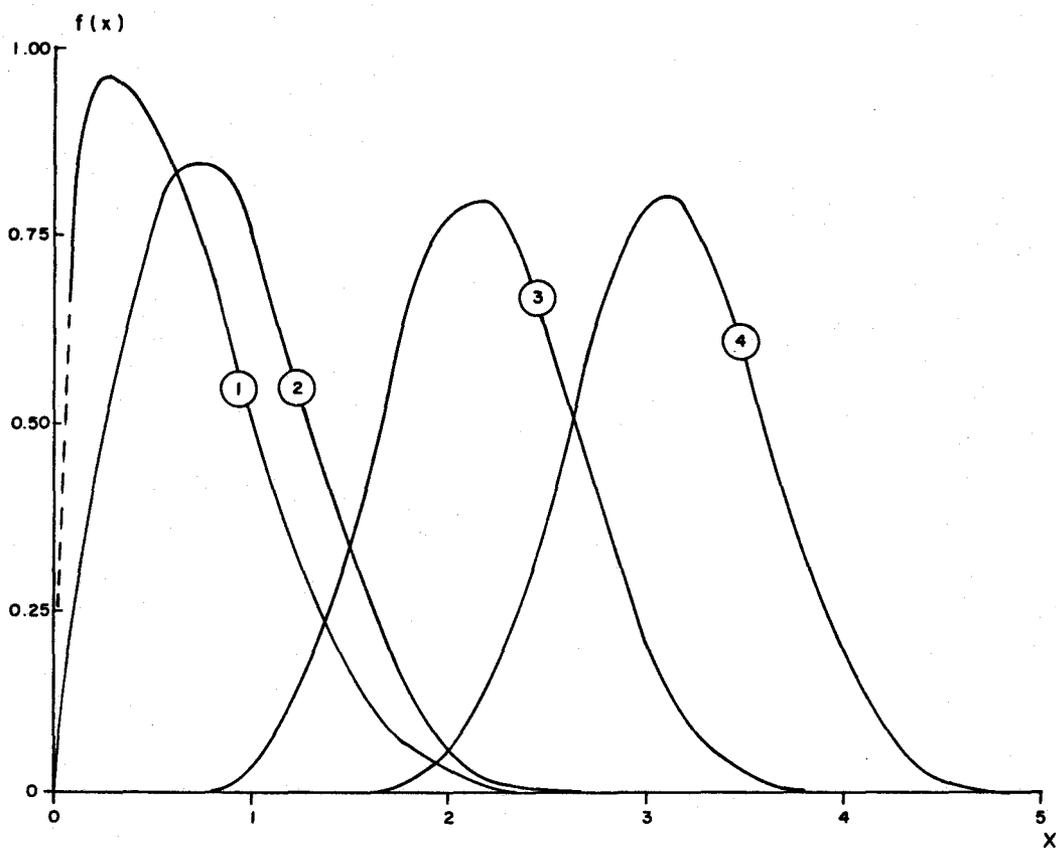


Figure 2: Influence du paramètre  $\lambda$  sur la forme de la loi G.G.  
(les détails relatifs à ces courbes se trouvent sur le Tableau 2)

Tableau 2: Détails relatifs aux fonctions de la figure 2

	1	2	3	4
a	1,0	1,0	1,0	1,0
$\lambda$	0,6	1,0	5,0	10,0
S	2,0	2,0	2,0	2,0
$\log \mu'_1$	-0,194	-0,052	-0,338	0,494
$\log \mu'_2$	-0,021	0,000	0,698	1,000
$\log \mu'_3$	-0,153	0,123	1,079	1,515
$\log \mu'_4$	-0,017	0,301	1,477	2,041
$\log \mu_2$	-0,716	-0,668	-0,613	-0,607
$\log  \mu_3 $	-1,129	-1,202	-1,544	-1,699
$\log \mu_4$	-0,873	-0,825	-0,748	-0,737
$C_V$	0,685	0,522	0,226	0,159
$C_S$	0,883	0,631	0,237	0,163
$C_K$	3,634	3,245	3,008	3,002
f(o)	0	0	0	0
MODE x	0,316	0,707	2,121	3,082
f(x)	0,965	0,857	0,805	0,801
I x	0,867	1,224	2,624	3,583
N f(x)	0,614	0,546	0,502	0,496
F x	-	0,000	1,616	2,580
L f(x)	-	0,000	0,460	0,469
E x	-	0,000	0,460	0,469
X f(x)	-	0,000	0,460	0,469
I x	-	0,000	0,460	0,469
O f(x)	-	0,000	0,460	0,469
N x	-	0,000	0,460	0,469

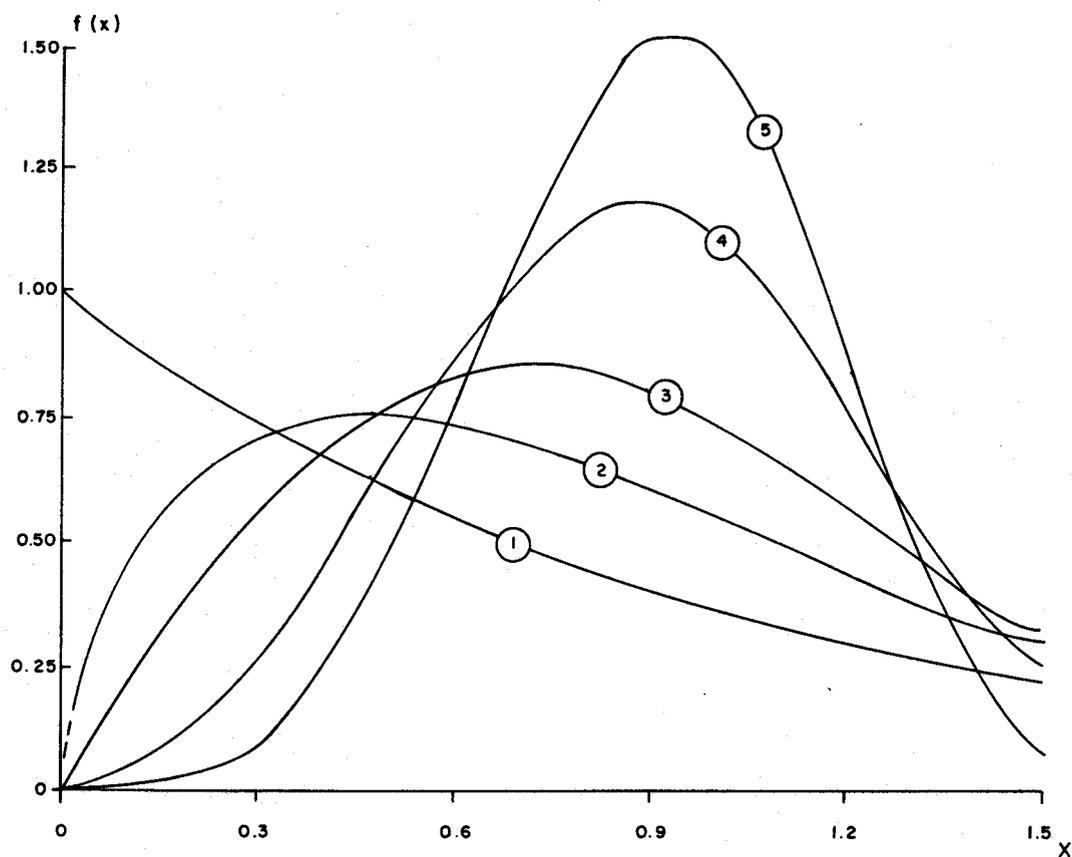


Figure 3: Influence du paramètre  $S$  sur la forme de la loi G.G.  
(les détails relatifs à ces courbes sont montrés sur le Tableau 3)

Tableau 3: Détails relatifs aux fonctions de la figure 3

	1	2	3	4	5
a	1,0	1,0	1,0	1,0	1,0
$\lambda$	1,0	1,0	1,0	1,0	1,0
S	1,0	1,5	2,0	3,0	4,0
$\log \mu'_1$	0,000	-0,044	-0,052	-0,049	-0,042
$\log \mu'_2$	0,301	0,075	0,000	-0,044	-0,052
$\log \mu'_3$	0,778	0,301	0,123	0,000	-0,036
$\log \mu'_4$	1,380	0,603	0,301	0,075	0,000
$\log \mu_2$	0,000	-0,425	-0,668	-0,977	-1,189
$\log  \mu_3 $	0,301	-0,607	-1,202	-2,240	-2,843
$\log \mu_4$	0,954	-0,207	-0,825	-1,518	-1,939
$C_V$	1,000	0,678	0,522	0,363	0,280
$C_S$	2,000	1,071	0,631	0,168	-0,087
$C_K$	9,000	4,390	3,245	2,729	2,747
f(o)	1,000	0,000	0,000	0,000	0,000
MODE x	0,000	0,480	0,707	0,873	0,930
f(x)	-	0,745	0,857	1,175	1,522
I x	0,000	1,066	1,224	1,234	1,999
N f(x)	0,000	0,515	0,546	0,696	0,871
F x	0,000	-	0,000	0,490	0,652
L x	0,000	-	0,000	0,490	0,652
E x	0,000	-	0,000	0,490	0,652
X f(x)	0,000	-	0,000	1,641	0,927
I					
O					
N					

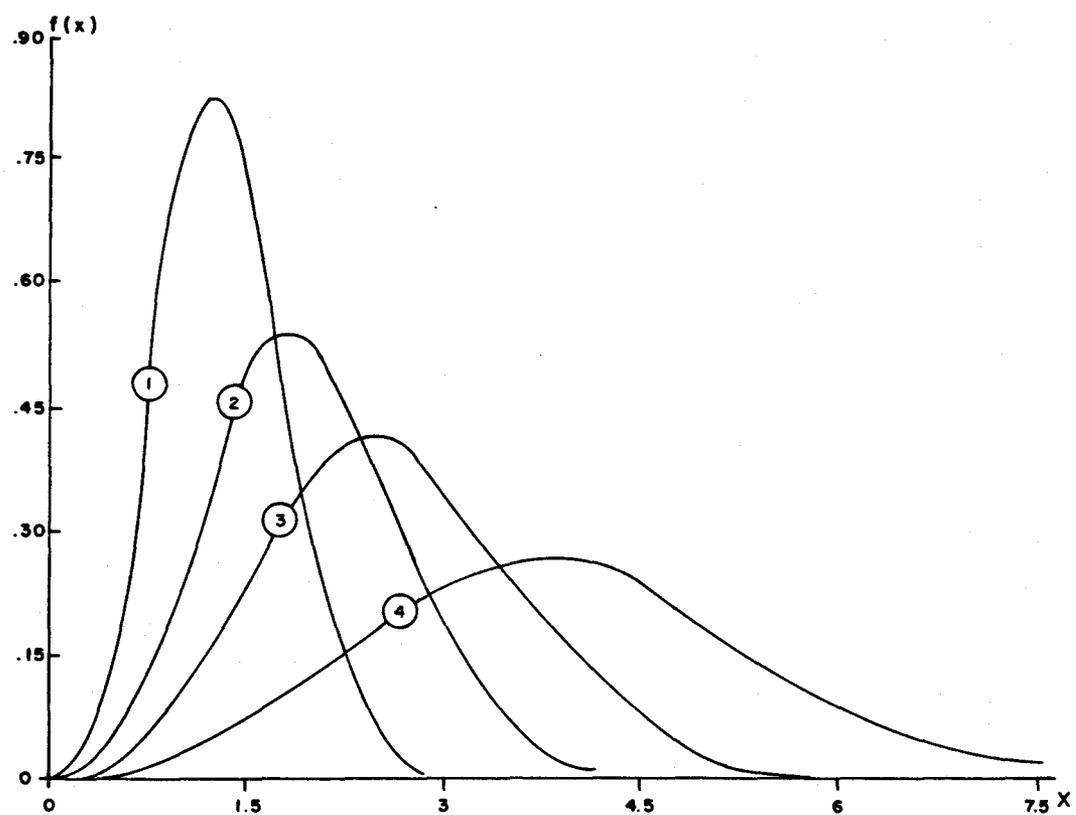


Figure 4: Influence du paramètre d'échelle  $\alpha$  pour la loi G.G.  
(les détails relatifs à ces courbes se trouvent au Tableau 4)

Tableau 4: Détails relatifs aux fonctions de la figure 4

	1	2	3	4
a	1,0	1,5	2,0	3,0
$\lambda$	2,0	2,0	2,0	2,0
S	2,0	2,0	2,0	2,0
$\log \mu'_1$	0,123	2,299	0,424	0,600
$\log \mu'_2$	0,301	0,653	0,903	1,255
$\log \mu'_3$	0,521	1,049	1,424	1,952
$\log \mu'_4$	0,778	1,482	1,982	2,686
$\log \mu_2$	-0,632	-0,280	-0,030	0,321
$\log  \mu_3 $	-1,341	-0,812	-0,438	0,090
$\log \mu_4$	-0,780	-0,075	0,423	1,128
$C_V$	0,363	0,363	0,363	0,363
$C_S$	0,405	0,405	0,405	0,405
$C_K$	3,059	3,059	3,059	3,059
f(o)	0,000	0,000	0,000	0,000
MODE x	1,224	1,837	2,449	3,674
f(x)	0,819	0,546	0,409	0,273
I x	1,732	2,598	3,464	5,196
N f(x)	0,517	0,344	0,258	0,172
F x	0,707	1,060	1,414	2,121
L f(x)	0,428	0,285	0,214	0,142
E x				
X f(x)				
I x				
O f(x)				
N x				

### Étude de quelques transformations usuelles

Il est intéressant d'examiner le comportement de la loi gamma généralisée lorsqu'on effectue un changement quelconque de variable  $y=h(x)$ , l'équation 1.5 s'exprime alors de la façon suivante:

$$g(y) = \frac{|S|}{a^{S\lambda}} \frac{(h^{-1}(y))^{S\lambda-1} e^{-(h^{-1}(y)/a)^S}}{\Gamma(\lambda)} \left| \frac{d(h^{-1}(y))}{dy} \right| \quad (1.12)$$

pour que  $g(y)$  demeure une f.d.p. gamma généralisée il faut et il suffit que  $g(y) = f(y; a^1, \lambda^1, S^1)$

si on pose  $y = x + c$  on retrouve une loi à quatre paramètres telle que définie au début:

$$f(x; a, \lambda, S) \rightarrow f(y; a, \lambda, S, c)$$

pour  $y = Kx$  on aura:  $f(x; a, \lambda, S) \rightarrow f(y; Ka, \lambda, S)$

On peut également démontrer que si  $x$  est tiré d'une population gamma généralisée  $f(x; a, \lambda, S)$ , alors  $x^P$  appartiendra à une population  $f(x^P; a^{1/P}, \lambda, PS)$ .

---

## CHAPITRE 2

### PROPRIÉTÉS STATISTIQUES DE LA DISTRIBUTION GAMMA GÉNÉRALISÉE

Les expressions mathématiques pour les moments centrés et non centrés, ainsi que pour les coefficients adimensionnels de la loi gamma généralisée font intervenir des termes du type:  $\Gamma(\lambda + r/s)$  avec  $r = 0, 1, 2, 3, \dots$

Il peut arriver que  $S$  soit près de 0 et  $\lambda$  très grand, d'ailleurs lorsque  $C_S \rightarrow C_V^3 + 3 C_V$   $S \rightarrow 0$  et  $\lambda \rightarrow \infty$ , la loi gamma généralisée tend alors vers une loi log-normale. Il convient mieux de toujours utiliser les logarithmes de ces fonctions gammas.

#### 2.1 Moments

On définit les moments non-centrés de la façon suivante:

$$\mu'_r = \frac{a^r \Gamma(\lambda + r/s)}{\Gamma(\lambda)} \quad (2.1)$$

Les moments sont définis seulement si  $S$  est positif ou si  $S < -r/\lambda$  lorsque  $S$  est négatif

$$\text{la moyenne est: } \mu'_1 = \frac{a \Gamma(\lambda + 1/S)}{\Gamma(\lambda)} \quad (2.2)$$

Les moments centrés  $\mu_r$  sont reliés aux moments non-centrés  $\mu'_r$  par la relation suivante:

$$\mu_r = \sum_{j=0}^r C_r^j \mu'_{r-j} (-\mu'_1)^j \quad (2.3)$$

$C_r^j$  étant les coefficients correspondant au développement d'un binôme de degré  $r$ .

La variance en particulier est donnée par:

$$\mu_2 = \frac{a^2}{\Gamma^2(\lambda)} [\Gamma(\lambda) \Gamma(\lambda+2/S) - \Gamma^2(\lambda+1/S)] \quad (2.4)$$

On remarque que le paramètre d'échelle  $a$ , qui possède la même dimension que  $x$ , intervient toujours seul sous la forme  $a^r$ , alors que  $\lambda$  et  $S$  sont étroitement reliés ensemble, il est d'ailleurs normal de s'attendre à une disparition par simplification de celui-ci lors du calcul des coefficients adimensionnels caractérisant la distribution.

## 2.2 Coefficients

Le coefficient de variation est défini de la façon suivante:

$$C_V = \frac{\sqrt{\mu_2}}{\mu'} = \frac{\sqrt{\Gamma(\lambda) \Gamma(\lambda+2/S) - \Gamma^2(\lambda+1/S)}}{\Gamma(\lambda+1/S)} \quad (2.5)$$

La figure 5 représente  $C_V$  pour  $S$  variant de  $-4$  à  $6$  et pour  $\lambda$  étant égal à  $0.1$   $0.5$   $1$   $2$   $5$   $10$   $50$ . On remarque que la droite  $S=0$  constitue une asymptote. Lorsque  $S \rightarrow 0^+$   $C_V$  devient de plus en plus grand, cependant lorsque  $S \rightarrow 0^-$ ,  $\Gamma(\lambda+2/S)$  et  $C_V$  deviennent éventuellement indéfinis.

-----

Le coefficient d'asymétrie possède une expression plus complexe:

$$C_S = \frac{\mu_3}{(\mu_2)^{3/2}} = \frac{\Gamma^2(\lambda) \Gamma(\lambda+3/S) - 3\Gamma(\lambda) \Gamma(\lambda+2/S) \Gamma(\lambda+1/S) + 2\Gamma^3(\lambda+1/S)}{[\Gamma(\lambda) \Gamma(\lambda+2/S) - \Gamma^2(\lambda+1/S)]^{3/2}} \quad (2.6)$$

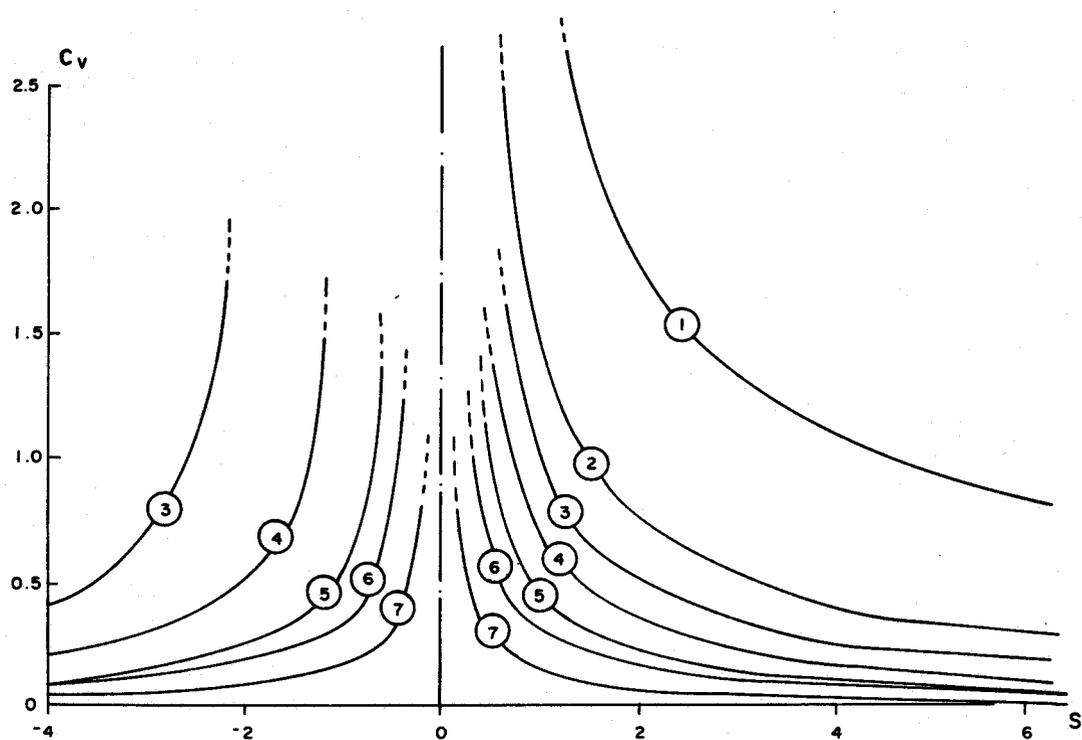


Figure 5: Graphe du coefficient de variation en fonction du paramètre  $S$ , pour quelques valeurs de  $\lambda$ .

1:  $\lambda=0,1$     2:  $\lambda=0,5$     3:  $\lambda=1,0$     4:  $\lambda=2,0$     5:  $\lambda=5$     6:  $\lambda=10$     7:  $\lambda=50$

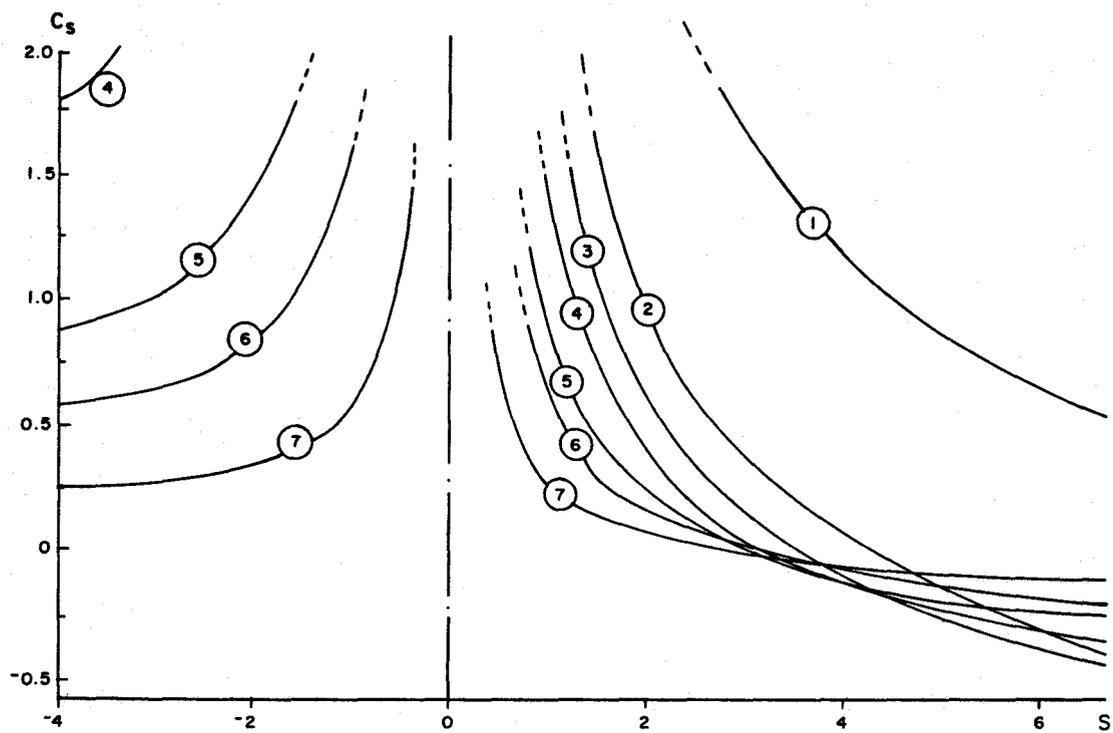


Figure 6: Graphe du coefficient d'asymétrie  $C_s$  en fonction du paramètre  $S$ , pour quelques valeurs de  $\lambda$ .

1:  $\lambda=0,1$  2:  $\lambda=0,5$  3:  $\lambda=1,0$  4:  $\lambda=2,0$  5:  $\lambda=5,0$  6:  $\lambda=10,0$  7:  $\lambda=50,0$

Cette relation a été représentée graphiquement sur la figure 6 pour les mêmes conditions que précédemment; les mêmes remarques s'appliquent pour le comportement de  $C_S$  près de  $S = 0$ .

Ce n'est que pour les grandes valeurs de  $S$  que  $C_S$  peut être négatif, encore là il est nécessaire que  $\lambda$  ne soit pas trop grand. Lorsque  $S < 0$ ,  $C_S$  est toujours positif. La loi gamma généralisée peut fournir des distributions possédant des asymétries ayant, semble-t-il, une limite inférieure de  $-2$ . Ceci est bien suffisant en ce qui concerne les échantillons rencontrés habituellement en hydrologie, mais alors  $S$  est grand et la gamme des coefficients de variation est peu étendue ( $C_V$  est très faible comme le montre la figure 5).

Le coefficient d'aplatissement défini par  $C_K = \frac{\mu_4}{(\mu_2)^2}$  est donné par:

$$C_K = \frac{\Gamma^3(\lambda)\Gamma(\lambda+4/S) - 4\Gamma^2(\lambda)\Gamma(\lambda+3/S)\Gamma(\lambda+1/S) + 6\Gamma(\lambda)\Gamma(\lambda+2/S)\Gamma^2(\lambda+1/S) - 3\Gamma^4(\lambda+1/S)}{[\Gamma(\lambda)\Gamma(\lambda+2/S) - \Gamma^2(\lambda+1/S)]^2}$$

La figure 7 représente  $C_K$  pour les mêmes conditions que les deux figures précédentes; ici également on remarque le comportement qui est propre à tous les coefficients adimensionnels de la loi G.G. autour de la droite  $S=0$ .

Pour  $\lambda = 50$ , la courbe plonge rapidement lorsque  $S$  augmente à partir de  $S = 0$ , elle atteint rapidement un plateau et reste horizontale à  $C_K = 3$ .

La fonction gamma  $\Gamma(x)$  possède un minimum pour  $x = 1.46163 21449 68362$  (Abramowitz, M. et I.A. Stegun) de part et d'autre de ce point il existe deux branches dont les pentes deviennent de plus en plus fortes et de signes opposées; les quantités  $\Gamma(\lambda)$ ,  $\Gamma(\lambda+1/S)$ ,  $\Gamma(\lambda+2/S)$ , .....  $\Gamma(\lambda+n/S)$  peuvent se situer sur ces deux branches et lorsqu'on fait varier  $S$  elles se déplacent mais certaines de ces quantités peuvent augmenter alors que d'autres diminuent. C'est pour ces raisons que les courbes des coefficients sont assez complexes, d'autre part il existe plus de combinaisons possibles entre les

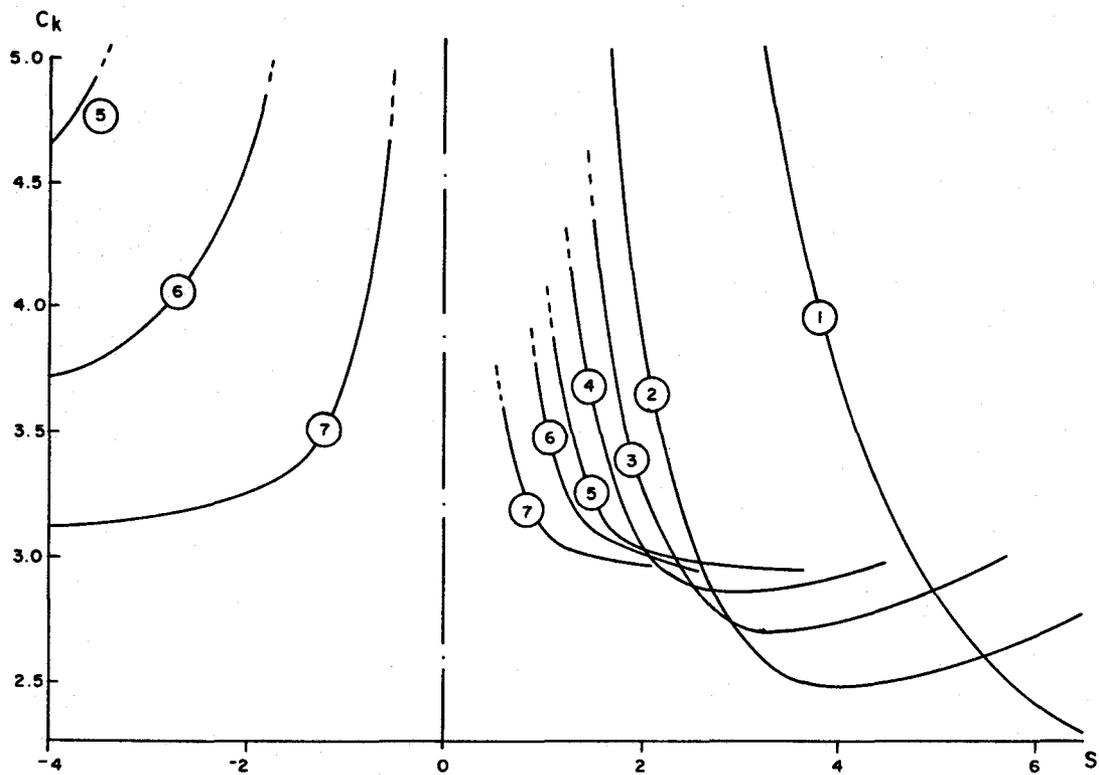


Figure 7: Graphe du coefficient d'aplatissement  $C_k$  en fonction du paramètre  $S$  pour quelques valeurs de  $\lambda$

1:  $\lambda=0,1$  2:  $\lambda=0,5$  3:  $\lambda=1,0$  4:  $\lambda=2,0$  5:  $\lambda=5,0$  6:  $\lambda=10,0$  7:  $\lambda=50,0$

coefficients lors de l'ajustement des paramètres à partir d'un échantillon quelconque, d'où le potentiel accru de la généralisation de la fonction gamma.

---

En pratique il est préférable de calculer les coefficients à partir des moments non-centrés en posant  $a = 1$  puisque ce paramètre se simplifie toujours, il convient également d'utiliser les logarithmes népériens des moments et de prendre préalablement les valeurs absolues lorsqu'il y a lieu, en conservant le signe adéquat en mémoire.

Cette procédure est d'autant plus importante que  $C_S$  est près de sa valeur critique  $C_S = C_V^3 + 3 C_V$ , comme mentionné au début de ce chapitre, car à ce moment  $\lambda$  peut-être très grand.

Ainsi on a obtenu de bons résultats en utilisant pour  $C_5$  et  $C_6$ :

$$|C_5| = e^{\ln|\mu_5| - 2.5(\ln \mu_2)} \quad (2.8)$$

$$C_6 = e^{\ln \mu_6 - 3(\ln \mu_2)} \quad (2.9)$$

$$\text{avec } \ln|\mu_5| = \ln \left| (1 - 5e^{G_4+G_1-G_5-G_0} + 10e^{G_3+2G_1-G_5-2G_0} - 10e^{G_2+3G_1-G_5-3G_0} + 4e^{5G_1-G_5-4G_0}) \right| + G_5 - G_0 \quad (2.10)$$

$$\ln \mu_6 = \ln (1 - 6e^{G_5+G_1-G_6-G_0} + 15e^{G_4+2G_1-G_6-2G_0} - 20e^{G_3+3G_1-G_6-3G_0} + 15e^{G_2+4G_1-G_6-4G_0} - 5e^{6G_1-G_6-5G_0}) + G_6 - G_0 \quad (2.11)$$

où  $Gr = \ln \Gamma(\lambda+r/S)$

Ces équations doivent être utilisées aux fins du traitement numérique.

Les tables de mathématiques telles celles de Abramowitz et Stegun donnent les expressions à utiliser pour le calcul de  $\ln \Gamma(x)$ , la fonction digamma  $\psi(x)$ , la fonction trigamma  $\psi'(x)$  et la fonction gamma incomplète  $\Gamma(\lambda, x)$ , qui seront employées ultérieurement.

Hoshi, K. et I. Yamaoka (1980) ont présenté des tables et des graphiques intéressants sur les relations entre  $C_V$ ,  $C_S$  et  $\lambda$ ,  $S$ . On peut se rendre compte de la situation de la zone discontinue où  $\lambda \rightarrow \infty$  et  $S \rightarrow 0$ .

---

### CHAPITRE 3

#### ESTIMATION DES PARAMÈTRES PAR LA MÉTHODE DU MAXIMUM DE VRAISEMBLANCE

Cette méthode classique d'ajustement conduit généralement à des intervalles de confiance qui sont plus restreints que les autres méthodes. En effet les variances des paramètres, qui tendent asymptotiquement vers 0 à mesure que les tailles des échantillons augmentent, sont plus petites que celles fournies par la méthode des moments par exemple. Cette efficacité accrue semble également être complétée par une meilleure estimation des événements  $x$  correspondant à une période de retour  $T$  selon les résultats obtenus lors des tests préliminaires de génération et d'estimation. Des résultats plus substantiels feront l'objet d'un autre rapport où on prévoit analyser plusieurs milliers d'échantillons synthétiques de petites tailles selon les lois de la famille gamma.

---

#### 3.1 Estimation des paramètres

La solution optimale recherchée est obtenue (Parr, V.B. et J.T. Webster, 1965) en maximisant la fonction de vraisemblance définie par:

$$U = \sum_{i=1}^N \ln f(x_i; a, \lambda, S)$$

$$U = N \ln |S| + N \lambda \ln \alpha - N \ln \Gamma(\lambda) - \alpha \sum_{i=1}^N x_i^S + (S\lambda - 1) \sum_{i=1}^N \ln x_i \quad (3.1)$$

$$\text{On a préalablement posé } \alpha = a^{-S} \quad (3.2)$$

car ce changement apporte une simplification par la disparition d'un paramètre  $S$ .  $N$  est la taille de l'échantillon.

En égalant à 0 les dérivées partielles de la fonction de vraisemblance par rapport à chacun des paramètres on obtient un système non linéaire de trois équations à trois inconnues.

$$\frac{\partial U}{\partial \alpha} \rightarrow \frac{\lambda}{\alpha} - \frac{1}{N} \sum_{i=1}^N x_i^S = 0 \quad (3.3)$$

$$\frac{\partial U}{\partial S} \rightarrow \frac{1}{S} - \alpha \frac{1}{N} \sum_{i=1}^N (x_i^S \ln x_i) + \lambda \frac{1}{N} \sum_{i=1}^N \ln(x_i) = 0 \quad (3.4)$$

$$\frac{\partial U}{\partial \lambda} \rightarrow \ln \alpha - \frac{\partial \ln \Gamma(\lambda)}{\partial \lambda} + \frac{S}{N} \sum_{i=1}^N \ln x_i = 0 \quad (3.5)$$

Les équations 3.3 et 3.4 sont linéaires vis-à-vis des paramètres  $\alpha$  et  $\lambda$ ; on peut ainsi les exprimer en fonction de  $S$  et des observations  $x_i$  de l'échantillon, de la façon suivante:

$$\alpha = \left( S \left[ \frac{\sum_{i=1}^N (x_i^S \ln x_i)}{N} - \frac{\sum_{i=1}^N x_i^S \sum_{i=1}^N \ln x_i}{N^2} \right] \right)^{-1} \quad (3.6)$$

$$\lambda = \frac{\alpha \sum_{i=1}^N x_i^S}{N} \quad (3.7)$$

L'approche initiale serait de substituer 3.6 et 3.7 dans 3.5, en tentant de résoudre pour  $S$  selon la méthode itérative de Newton-Raphson, ce qui implique une nouvelle dérivation par rapport à  $S$ , de l'équation ainsi transformée.

Les premiers tests ne nous ont guère poussé à persévérer dans cette voie. Des problèmes numériques importants (Lawless, J.F., 1980; Hager et al., 1971), sans doute introduits par des termes très près de 0, apparaissent sous la forme d'un comportement erratique dans l'évaluation de 3.5. La

méthode des sécantes est également affligée des mêmes tares. De plus, Hager, W.H. et L.J. Bain (1970) avaient déjà observé le même genre de problèmes et conseillé une méthode de tâtonnement (a simple search procedure).

Il est en effet beaucoup plus simple de faire varier  $S$  et de calculer à chaque fois  $\alpha$ ,  $\lambda$  et  $U$ . Ainsi on peut trouver des valeurs  $S$  pour lesquelles  $U$  devient de plus en plus grand. La procédure qui a été programmée se présente sous la forme de l'algorithme simplifié suivant:

- 1° On pose une valeur de départ pour  $S$  et pour l'incrément  $\Delta S$ .
- 2° On calcule  $U(S)$ ,  $U(S+\Delta S)$ ,  $U(S-\Delta S)$ .
- 3° On choisit le nouveau  $S$  qui augmente la fonction de vraisemblance  $U$ , lorsqu'il n'y a pas d'amélioration ou si  $S \pm \Delta S = 0$ , on divise l'intervalle par deux.

Éventuellement, après quelques oscillations autour de la valeur optimale de  $S$ , l'incrément  $\Delta S$  sera devenu suffisamment petit pour qu'on puisse arrêter la procédure. Bien sur, un programme beaucoup plus performant peut être imaginé qui saute automatiquement des étapes inutiles lorsque les renseignements désirés sont obtenus, en minimisant ainsi le nombre de fois où la fonction de vraisemblance doit être calculée.

Un avantage important de cette méthode est qu'elle ne peut pas diverger comme le ferait facilement un algorithme itératif Newton-Raphson, de plus on peut fixer une valeur de départ quelconque pour  $S$ .

Lors de l'analyse de petits échantillons, il peut arriver que  $|S|$  devienne très grand; à ce moment on peut considérer que l'échantillon est intraitable selon un modèle gamma généralisée, quoique les résultats pour la fonction de vraisemblance et pour le calcul des  $x_T$  puissent être excellents dans certains cas où l'estimation optimale semble se situer asymptotiquement pour  $|S|$  élevé et  $\lambda$  petit. Ceci est surtout susceptible de se produire lorsque le coefficient d'asymétrie de l'échantillon est assez négatif et que le coefficient de variation est assez grand ( $C_S < -1$  et  $C_V > .5$ ). On suggère d'imposer une limite telle que  $|S| < 30$  pour éviter les risques d'over-

flow, de plus lorsque  $|S|$  est grand,  $\lambda$  est tellement faible que les méthodes standards d'évaluation de l'événement  $x$  correspondant à une période de retour  $T$  ne sont plus applicables. On reviendra sur cet aspect au chapitre 5 lors de la présentation de la théorie sur les variances et covariances des paramètres. Parfois l'estimation des paramètres selon cette méthode nous fournira un estimateur  $S < 0$  et pour lequel certains moments ne sont pas définis; théoriquement la population obtenue par l'ajustement selon le maximum de vraisemblance peut donc posséder ni coefficient de variation, ni coefficient d'asymétrie; ce phénomène est très rare et n'entrave pas les expériences de génération et d'estimation sur des échantillons de petites tailles.

Les populations estimées par la méthode des moments possèdent toujours des coefficients de variation et des coefficients d'asymétrie qui sont bien définis, ceci est dû à l'essence même de cette méthode qui est décrite au chapitre suivant.

La figure 8 montre l'allure de la fonction de vraisemblance  $U(S)$  pour  $S$  variant de  $-10$  à  $10$ ; il s'agit ici d'un échantillon de taille 23 dont la solution optimale correspond à  $S = 2,4801$ . Dans la majorité des échantillons qui ont été analysés par cette méthode du maximum de vraisemblance, la fonction  $U(S)$  était bien définie (sauf pour  $S = 0$ ) et ne semblait posséder qu'un maximum distinct.

L'ajustement d'un échantillon typique nécessite environ 20 fois le calcul de la fonction de vraisemblance  $U(S)$ , cependant pour des échantillons de petites tailles ceci ne représente guère plus de calculs que par la méthode des moments qui nécessite moins d'itérations mais qui est plus complexe.

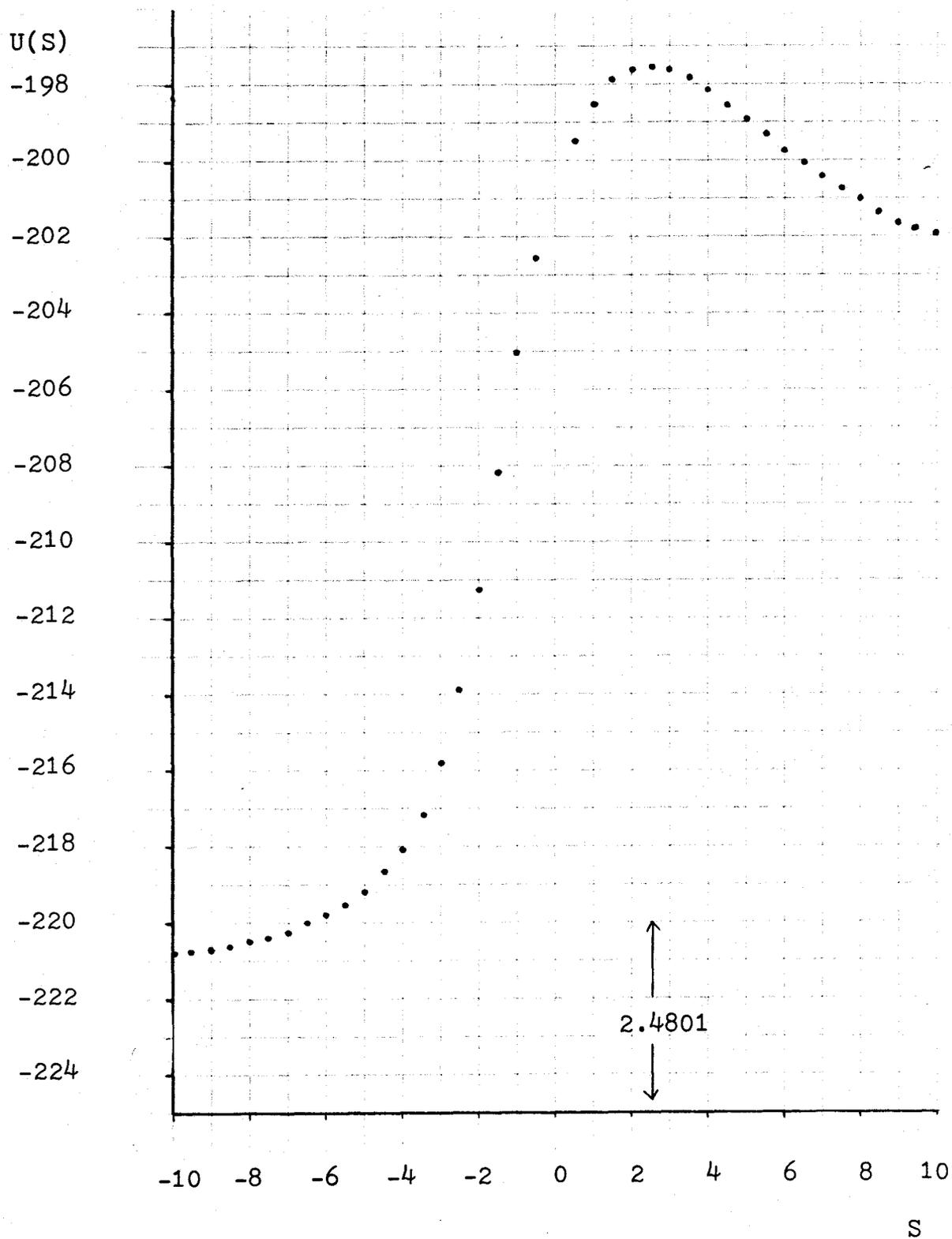


Figure 8 : La fonction de vraisemblance  $U(S)$   
pour  $-10 \leq S \leq 10$

## CHAPITRE 4

### ESTIMATION DES PARAMÈTRES PAR LA MÉTHODE DES MOMENTS

#### 4.1 Généralités

La méthode des moments appliquée à la loi gamma généralisée à trois paramètres consiste à égaliser les trois premiers moments non centrés de l'échantillon aux expressions théoriques des moments correspondants de la distribution exprimés en fonction des paramètres.

On dispose ainsi de 3 équations à 3 inconnues, dont la solution conduit à l'estimation des paramètres par la méthode des moments. Comme l'indique Bobée (1975) cette méthode s'appuie sur le fait que les moments non centrés de l'échantillon sont des estimateurs absolument corrects des moments correspondants de la distribution.

Les estimations obtenues par cette méthode sont correctes mais biaisées et généralement non efficaces.

#### 4.2 Estimation des paramètres

Si on postule l'égalité entre la moyenne de l'échantillon et le premier moment non centré de la population on obtient après transformation:

$$a = \frac{\bar{X} \Gamma(\lambda)}{\Gamma(\lambda + 1/S)} \quad (4.1)$$

Les estimateurs des paramètres  $\lambda$  et  $S$  nous sont préalablement fournis par la résolution du système non linéaire obtenu en égalant les coefficients de variation et d'asymétrie de l'échantillon, après corrections pour le biais,

aux expressions théoriques correspondantes de la loi gamma généralisée. On a vu précédemment qu'elles sont indépendantes du paramètre  $a$  (équations 2.5 et 2.6).

Hoshi, K. et I. Yamaoka (1980) ont développé une méthode itérative Newton-Raphson dans le but de résoudre ce système à deux équations et à deux inconnues, on reprend ici le développement essentiel au bon déroulement du processus en posant d'abord:

$$C_V = C_V(\lambda, S) + \Delta\lambda \frac{\partial C_V(\lambda, S)}{\partial \lambda} + \Delta S \frac{\partial C_V(\lambda, S)}{\partial S} \quad (4.2)$$

$$C_S = C_S(\lambda, S) + \Delta\lambda \frac{\partial C_S(\lambda, S)}{\partial \lambda} + \Delta S \frac{\partial C_S(\lambda, S)}{\partial S} \quad (4.3)$$

$C_V$  et  $C_S$  étant les coefficients corrigés de l'échantillon,  $C_V(\lambda, S)$  et  $C_S(\lambda, S)$  étant ceux de la loi gamma généralisée.

Si on suppose des valeurs de départ pour  $\lambda$  et  $S$  on peut résoudre 4.2 et 4.3 pour  $\Delta\lambda$  et  $\Delta S$ .

On obtient successivement de meilleurs estimateurs pour  $S$ :  $S_{i+1} = S_i + \Delta S$  et pour  $\lambda$  :  $\lambda_{i+1} = \lambda_i + \Delta\lambda$

Les itérations sont poursuivies jusqu'à ce que  $\Delta\lambda/\lambda$  et  $\Delta S/S$  soient devenus suffisamment petits, on utilise également:

$$\frac{\partial C_V(\lambda, S)}{\partial \lambda} = C_V(\lambda, S) \left[ \frac{1}{2} \frac{\partial A}{\partial \lambda} \frac{1}{A} - \frac{\partial B}{\partial \lambda} \frac{1}{B} \right] \quad (4.4)$$

$$\frac{\partial C_V(\lambda, S)}{\partial S} = C_V(\lambda, S) \left[ \frac{1}{2} \frac{\partial A}{\partial S} \frac{1}{A} - \frac{\partial B}{\partial S} \frac{1}{B} \right] \quad (4.5)$$

$$\frac{\partial C_S(\lambda, S)}{\partial \lambda} = C_S(\lambda, S) \left[ \frac{\partial C}{\partial \lambda} \frac{1}{C} - \frac{3}{2} \frac{\partial A}{\partial \lambda} \frac{1}{A} \right] \quad (4.6)$$

$$\frac{\partial C_S(\lambda, S)}{\partial S} = C_S(\lambda, S) \left[ \frac{\partial C}{\partial S} \frac{1}{C} - \frac{3}{2} \frac{\partial A}{\partial S} \frac{1}{A} \right] \quad (4.7)$$

$$\text{avec } A = \Gamma(\lambda + 2/S) \Gamma(\lambda) - \Gamma^2(\lambda+1/S) \quad (4.8)$$

$$B = \Gamma(\lambda+1/S) \quad (4.9)$$

$$C = \Gamma(\lambda+3/S) \Gamma^2(\lambda) - 3\Gamma(\lambda+2/S)\Gamma(\lambda+1/S)\Gamma(\lambda) + 2\Gamma^3(\lambda+1/S) \quad (4.10)$$

$$\frac{\partial A}{\partial \lambda} = [\psi(\lambda) + \psi(\lambda+2/S)] \Gamma(\lambda+2/S) \Gamma(\lambda) - 2\psi(\lambda+1/S) \Gamma^2(\lambda+1/S) \quad (4.11)$$

$$\frac{\partial A}{\partial S} = -\frac{2}{S^2} [\psi(\lambda+2/S) \Gamma(\lambda) \Gamma(\lambda+2/S) - \psi(\lambda+1/S) \Gamma^2(\lambda+1/S)] \quad (4.12)$$

$$\frac{1}{B} \frac{\partial B}{\partial \lambda} = \psi(\lambda+1/S) \quad (4.13)$$

$$\frac{1}{B} \frac{\partial B}{\partial S} = \frac{-\psi(\lambda+1/S)}{S^2} \quad (4.14)$$

$$\begin{aligned} \frac{\partial C}{\partial \lambda} = & \Gamma(\lambda+3/S) \Gamma^2(\lambda) [\psi(\lambda+3/S) + 2\psi(\lambda)] \\ & - 3\Gamma(\lambda+2/S) \Gamma(\lambda+1/S) \Gamma(\lambda) [\psi(\lambda+2/S) + \psi(\lambda+1/S) + \psi(\lambda)] \\ & + 6\Gamma^3(\lambda+1/S) \psi(\lambda+1/S) \end{aligned} \quad (4.15)$$

$$\begin{aligned} \frac{\partial C}{\partial S} = & -\frac{3}{S^2} [\Gamma(\lambda+3/S) \Gamma^2(\lambda) \psi(\lambda+3/S) + 2\Gamma^3(\lambda+1/S) \psi(\lambda+1/S) \\ & - \Gamma(\lambda+2/S) \Gamma(\lambda+1/S) \Gamma(\lambda) [2\psi(\lambda+2/S) + \psi(\lambda+1/S)]] \end{aligned} \quad (4.16)$$

et  $\psi(\lambda) = \frac{d}{d\lambda} \ln \Gamma(\lambda)$  est la fonction digamma qui est disponible sous la forme d'un développement en série asymptotique dans les tables de mathématiques avancées (Abramowitz, M. et I.A. Stegun).

-----

Les manipulations algébriques sont beaucoup plus lourdes que dans le cas de l'estimation par le maximum de vraisemblance. Il est nécessaire à chaque itération d'évaluer quatre fonctions gammas et quatre fonctions digammas, on peut cependant couper court à cette méthode en évaluant les quatre dérivées (4.4 à 4.7) d'une façon numérique plutôt que théorique, ainsi:

$$\frac{\partial C_V(\lambda, S)}{\partial \lambda} = \frac{C_V(\lambda + \epsilon\lambda, S) - C_V(\lambda - \epsilon\lambda, S)}{2\epsilon\lambda} \quad (4.17)$$

on a obtenu de bons résultats sur une calculatrice de poche en utilisant  $\epsilon = 0,001$ .

La méthode suggérée par Hoski K. et I. Yamaoka (1980) est une méthode qui nécessite environ 5 itérations pour en arriver à des résultats précis à 0,001%, cependant le processus itératif a de fortes tendances à diverger lorsque les valeurs de départ sont trop éloignées des valeurs optimales théoriques. Les auteurs précédents proposent des tables étendues dont on peut se servir en interpolant pour obtenir de bonnes valeurs de départ. On a cependant choisi, dans le présent travail, d'effectuer des régressions linéaires multiples à partir des données de  $\lambda$ ,  $S$ ,  $C_V$ ,  $C_S$  qui ont été générées par ordinateur, ainsi en connaissant  $C_V$  et  $C_S$  de l'échantillon on peut obtenir des valeurs initiales en employant le type d'équations suivantes:

$$\log \lambda = \sum_{i=0}^n a_i C_V^i \times \sum_{i=0}^n b_i C_S^i \quad (4.18)$$

De façon semblable, on trouve une valeur de départ, pour le paramètre  $S$ , le cinquième degré s'est avéré suffisant ( $n = 5$ ). On dispose finalement d'une méthode d'estimation de paramètres pour les moments de l'échantillon

qui fonctionne de façon automatique dans tous les cas et qui est apte à être utilisée dans les tests d'estimation des  $x_T$  à partir d'échantillons obtenus par génération de variables aléatoires.

---

## CHAPITRE 5

### CALCUL DE LA VARIANCE DE L'ÉVÉNEMENT $X_T$

#### 5.1 Introduction

Le calcul de  $\text{var}(X_T)$ , variance de l'événement de période de retour  $T$  est important en particulier pour la détermination de l'intervalle de confiance de la vraie valeur inconnue  $X_T$ ; en effet si l'on admet que  $X_T$  est approximativement distribué suivant une loi normale de moyenne  $X_T$  et de variance  $\text{var}(X_T)$ , la connaissance de  $\text{var}(X_T)$  permet d'estimer rapidement l'intervalle de confiance de  $X_T$ , on a pour l'intervalle de confiance au niveau  $(1-\alpha)$ :

$$X_T - U_{\alpha/2} \sqrt{\text{var } X_T} \leq x_T \leq x_T + U_{\alpha/2} \sqrt{\text{var } X_T}$$

$U_{\alpha/2}$  étant la variable normale centrée réduite de probabilité au dépassement  $\alpha/2$ .

Puisque la f.d.p. de la loi gamma généralisée à 3 paramètres dépend de  $a$ ,  $\lambda$  et  $S$ ,  $X_T$  est fonction de ces paramètres et l'on a de manière générale:

$$\begin{aligned} \text{Var}(X_T) = & \left(\frac{\partial X_T}{\partial a}\right)^2 \text{Var } a + 2 \left(\frac{\partial X_T}{\partial a}\right) \left(\frac{\partial X_T}{\partial \lambda}\right) \text{cov}(a, \lambda) + 2 \left(\frac{\partial X_T}{\partial a}\right) \left(\frac{\partial X_T}{\partial S}\right) \text{cov}(a, S) \\ & + \left(\frac{\partial X_T}{\partial \lambda}\right)^2 \text{Var } \lambda + 2 \left(\frac{\partial X_T}{\partial \lambda}\right) \left(\frac{\partial X_T}{\partial S}\right) \text{cov}(\lambda, S) + \left(\frac{\partial X_T}{\partial S}\right)^2 \text{Var } S \quad (5.1) \end{aligned}$$

Les dérivées partielles sont calculées pour les valeurs estimées de paramètres.

Pour calculer  $\text{var}(X_T)$  on doit donc estimer:

- les dérivées partielles de  $X_T$  par rapport aux paramètres;

- Les variances et covariances d'estimation des paramètres qui dépendent de la méthode d'estimation considérée, qui sont données par la matrice de dispersion (V).

On a vu en 1.1 qu'en considérant la transformation:

$$w = \frac{X^S}{a} \quad (\text{ici } m = 0)$$

la loi gamma généralisée est transformée en loi gamma à 1 paramètre (équation 1.2)

Puisque  $X_T = aW_T^{1/S}$  on a:

$$\frac{\partial X_T}{\partial a} = W_T^{1/S} \quad (5.2)$$

$$\frac{\partial X_T}{\partial S} = -\frac{aW_T^{1/S}}{S^2} \ln W_T \quad (5.3)$$

$$\frac{\partial X_T}{\partial \lambda} = -\frac{aW_T^{1/S-1}}{S} \frac{\partial W_T}{\partial \lambda} \quad (5.4)$$

$W_T = \Phi(\lambda, T)$  est une fonction de  $\lambda$  et  $T$  seulement.

$W_T$  et  $(dW_T/d\lambda)$  peuvent être calculés à l'aide de l'approximation de Wilson Hilferty :

$$W_T = \lambda \left( 1 - \frac{1}{9\lambda} + \frac{U_T^3}{3\sqrt{\lambda}} \right) \quad (5.5)$$

$$\frac{\partial W_T}{\partial \lambda} = \left( 1 - \frac{1}{3\lambda} + \frac{U_T^2}{3\sqrt{\lambda}} \right) \left( 1 + \frac{2}{9\lambda} - \frac{U_T}{6\sqrt{\lambda}} \right) \quad (5.6)$$

$U_T$  est la variate normale centrée réduite correspondant à la période de retour  $T$ , c'est-à-dire à la probabilité au dépassement  $P = \frac{1}{T}$ .

Hoshi et Yamaoka (1980) suggèrent l'utilisation de ces formules pour  $\lambda \geq 20$ .

-----

On peut également se ramener à la variate Pearson type III standardisée  $K$  pour laquelle on dispose des tables de Harter (1969).  $K$  est comme  $W$  une fonction de  $T$  et  $\lambda$ , ou encore de  $T$  et  $C_S$  (asymétrie de la loi Pearson type III) puisque  $C_S = 2/\sqrt{\lambda}$ .

Puisque  $W_T$  est une loi gamma de moyenne  $\lambda$  et de variance  $\lambda$  on a :

$$K_T = \frac{W_T - \lambda}{\sqrt{\lambda}}$$

d'où l'on peut déduire les relations :

$$W = K \sqrt{\lambda} + \lambda \quad (5.7)$$

$$\frac{dW}{d\lambda} = 1 + \frac{K}{2\sqrt{\lambda}} - \frac{1}{\lambda} \frac{dK}{dC_S} \quad (5.8)$$

ou de manière inverse

$$\frac{dK}{dC_S} = \lambda \left( 0.5 - \frac{dW}{d\lambda} \right) + \frac{W}{2} \quad (5.9)$$

Les dérivées données dans les équations 5.2 à 5.4 peuvent donc être calculée, en fonction de  $K_T$  et  $\frac{\partial K_T}{\partial C_S}$ .

Tableau 5: Exemple d'amplification d'erreur lors du calcul de la variable x en fonction de S, pour  $\lambda = 1$  et pour 9 probabilités au non dépassement

$\frac{x_\epsilon}{x}$ correspondant à une probabilité au non dépassement P											
S	$C_V$	$C_S$	P=.9999	P=.999	P=.99	P=.9	P=.5	P=.1	P=.01	P=.001	P=.0001
-9	0,1569	2,0280	0,7661	0,9259	0,9895	0,9990	0,9998	1,0000	1,0000	1,0000	1,0000
-6	0,2505	2,8056	0,6706	0,8909	0,9843	0,9984	0,9998	0,9999	1,0000	1,0000	1,0000
-5	0,3141	3,5351	0,6190	0,8706	0,9812	0,9981	0,9997	0,9999	1,0000	1,0000	1,0000
-4.5	0,3607	4,2388	0,5869	0,8572	0,9791	0,9979	0,9997	0,9999	1,0000	1,0000	1,0000
0.7	1,4624	3,4984	1,0002	1,0002	1,0003	1,0006	1,0021	1,0136	1,1451	2,6918	30,7400
0.9	1,1130	2,3450	1,0001	1,0002	1,0002	1,0005	1,0016	1,0106	1,1112	2,1601	14,3583
1.5	0,6790	1,0720	1,0001	1,0001	1,0001	1,0003	1,0010	1,0063	1,0653	1,5874	4,9461
3	0,3634	0,1681	1,0000	1,0000	1,0001	1,0001	1,0005	1,0032	1,0321	1,2599	2,2240
5	0,2291	-0,2541	1,0000	1,0000	1,0000	1,0001	1,0003	1,0019	1,0192	1,1487	1,6154
7	0,1680	-0,4632	1,0000	1,0000	1,0000	1,0001	1,0002	1,0014	1,0136	1,1041	1,4085

Note:  $\frac{x_\epsilon}{x} = \left(1 + \frac{\epsilon}{K + \sqrt{\lambda}}\right)^{1/S}$  avec  $x_\epsilon$  = valeur de l'événement x évaluée en supposant une erreur de + 0,001 sur k ( $\epsilon = 0,001$ )  
 et  $x$  = valeur exacte de x

K est la variable Pearson-III standardisée avec une précision de  $\pm 0,00001$  telle que prise dans les tables de Harter (1969).

$$\begin{bmatrix} \text{Var } a \\ \text{Cov } (a, \lambda) \\ \text{Cov } (a, S) \\ \text{Var } \lambda \\ \text{Cov } (\lambda, S) \\ \text{Var } S \end{bmatrix} = \begin{bmatrix} S_{11} & S_{12} & S_{13} & S_{14} & S_{15} & S_{16} \\ 0 & S_{22} & S_{23} & S_{24} & S_{25} & S_{26} \\ 0 & S_{32} & S_{33} & S_{34} & S_{35} & S_{36} \\ 0 & 0 & 0 & S_{44} & S_{45} & S_{46} \\ 0 & 0 & 0 & S_{54} & S_{55} & S_{56} \\ 0 & 0 & 0 & S_{64} & S_{65} & S_{66} \end{bmatrix}^{-1} \begin{bmatrix} \text{Var } \bar{X} \\ \text{Cov } (\bar{X}, C_V) \\ \text{Cov } (\bar{X}, C_S) \\ \text{Var } C_V \\ \text{Cov } (C_V, C_S) \\ \text{Var } C_S \end{bmatrix} \quad (5.11)$$

$$\text{avec } S_{11} = \left( \frac{\partial \bar{X}}{\partial a} \right)^2 \quad S_{12} = 2 \left( \frac{\partial \bar{X}}{\partial a} \right) \left( \frac{\partial \bar{X}}{\partial \lambda} \right) \quad S_{13} = 2 \left( \frac{\partial \bar{X}}{\partial a} \right) \left( \frac{\partial \bar{X}}{\partial S} \right)$$

$$S_{14} = \left( \frac{\partial \bar{X}}{\partial \lambda} \right)^2 \quad S_{15} = 2 \left( \frac{\partial \bar{X}}{\partial \lambda} \right) \left( \frac{\partial \bar{X}}{\partial S} \right) \quad S_{16} = \left( \frac{\partial \bar{X}}{\partial S} \right)^2$$

$$S_{22} = \left( \frac{\partial \bar{X}}{\partial a} \right) \left( \frac{\partial C_V}{\partial \lambda} \right) \quad S_{23} = \left( \frac{\partial \bar{X}}{\partial a} \right) \left( \frac{\partial C_V}{\partial S} \right)$$

$$S_{24} = \left( \frac{\partial \bar{X}}{\partial \lambda} \right) \left( \frac{\partial C_V}{\partial \lambda} \right) \quad S_{25} = \left( \frac{\partial \bar{X}}{\partial \lambda} \right) \left( \frac{\partial C_V}{\partial S} \right) + \left( \frac{\partial \bar{X}}{\partial S} \right) \left( \frac{\partial C_V}{\partial \lambda} \right)$$

$$S_{26} = \left( \frac{\partial \bar{X}}{\partial S} \right) \left( \frac{\partial C_V}{\partial S} \right)$$

$$S_{32} = \left(\frac{\partial \bar{X}}{\partial a}\right) \left(\frac{\partial C_S}{\partial \lambda}\right) \quad S_{33} = \left(\frac{\partial \bar{X}}{\partial a}\right) \left(\frac{\partial C_S}{\partial S}\right)$$

$$S_{34} = \left(\frac{\partial \bar{X}}{\partial \lambda}\right) \left(\frac{\partial C_S}{\partial \lambda}\right) \quad S_{35} = \left(\frac{\partial \bar{X}}{\partial \lambda}\right) \left(\frac{\partial C_S}{\partial S}\right) + \left(\frac{\partial \bar{X}}{\partial S}\right) \left(\frac{\partial C_S}{\partial \lambda}\right)$$

$$S_{36} = \left(\frac{\partial \bar{X}}{\partial S}\right) \left(\frac{\partial C_S}{\partial S}\right)$$

$$S_{44} = \left(\frac{\partial C_V}{\partial \lambda}\right)^2 \quad S_{45} = 2 \left(\frac{\partial C_V}{\partial \lambda}\right) \left(\frac{\partial C_V}{\partial S}\right)$$

$$S_{46} = \left(\frac{\partial C_V}{\partial S}\right)^2$$

$$S_{54} = \left(\frac{\partial C_V}{\partial \lambda}\right) \left(\frac{\partial C_S}{\partial \lambda}\right) \quad S_{55} = \left(\frac{\partial C_V}{\partial \lambda}\right) \left(\frac{\partial C_S}{\partial S}\right) + \left(\frac{\partial C_V}{\partial S}\right) \left(\frac{\partial C_S}{\partial \lambda}\right)$$

$$S_{56} = \left(\frac{\partial C_V}{\partial S}\right) \left(\frac{\partial C_S}{\partial S}\right)$$

$$S_{64} = \left(\frac{\partial C_S}{\partial \lambda}\right)^2 \quad S_{65} = 2 \left(\frac{\partial C_S}{\partial \lambda}\right) \left(\frac{\partial C_S}{\partial S}\right)$$

$$S_{66} = \left(\frac{\partial C_S}{\partial S}\right)^2$$

avec

$$\frac{\partial \bar{X}}{\partial a} = \frac{\Gamma(\lambda+1/S)}{\Gamma(\lambda)}$$

$$\frac{\partial \bar{X}}{\partial \lambda} = a \frac{\Gamma(\lambda+1/S)}{\Gamma(\lambda)} (\psi(\lambda+1/S) - \psi(\lambda))$$

$$\frac{\partial \bar{X}}{\partial S} = - \frac{a}{S^2} \frac{\Gamma(\lambda+1/S)}{\Gamma(\lambda)} \psi(\lambda+1/S)$$

et

$$\frac{\partial C_V}{\partial \lambda}, \frac{\partial C_V}{\partial S}, \frac{\partial C_S}{\partial \lambda}, \frac{\partial C_S}{\partial S}$$

sont données respectivement par les relations 4.4 à 4.7.

Bobée, B. (1973) a développé les expressions pour les variances et covariances de  $\bar{X}$ ,  $C_V$  et  $C_S$  d'un échantillon de taille  $N$  en fonction des caractéristiques théoriques de la population considérée:

$$\text{Var}(\bar{X}) = \mu_2 / N$$

$$\text{Cov}(\bar{X}, C_V) = \sqrt{\mu_2} C_2 (C_3/2 - C_2) / N$$

$$\text{Cov}(\bar{X}, C_S) = \sqrt{\mu_2} (C_4 - 3(1 + C_3^2 / 2)) / N$$

$$\text{Var}(C_V) = C_2^2 ((C_4 - 1) / 4 + C_2 (C_2 - C_3)) / N$$

$$\text{Cov}(C_V, C_S) = C_2^2 (3 - C_4 - C_3 (3C_4 + 5) / (4C_2) + 3C_3^2 / 2 + C_5 / (2C_2)) / N$$

$$\text{Var}(C_S) = (9 + C_6 - 3C_3C_5 + (9 C_3^2 / 4 - 6) + 35C_3^2 / 4) / N$$

et avec  $C_2 = C_V$  (2.5)

$$C_3 = C_S \quad (2.6)$$

$$C_4 = C_K \quad (2.7)$$

$$C_5 \quad (2.8)$$

$$C_6 \quad (2.9)$$

$$\mu_2 \quad (2.4)$$

Ces termes sont évalués à partir des paramètres fournis par la méthode d'estimation.

---

## Conclusion

Ce rapport met en évidence quelques propriétés mathématiques et statistiques de la fonction gamma généralisée telle que présentée par Stacy (1962). Deux méthodes d'estimation sont exposées : La méthode des moments permet d'identifier la moyenne, le coefficient de variation et le coefficient d'asymétrie de l'échantillon aux expressions théoriques de la loi gamma généralisée, donnant un système non linéaire de trois équations à trois inconnues. Le traitement mathématique est celui proposé par Hoski, K. et I. Yamaoka (1980) qui donne la solution du système selon une méthode itérative Newton-Raphson, on a développé une sous-routine dans le cadre d'un programme de génération d'échantillons aléatoires qui fournit les valeurs de départ pour le processus itératif devant amener l'estimation des paramètres de façon automatique pour chacun de ces échantillons. La méthode d'estimation des paramètres par le maximum de vraisemblance utilise également un processus itératif qui assure la convergence dans la presque totalité des cas, même pour les échantillons de tailles réduites qui sont souvent rencontrés dans le domaine des débits de crue. Les outils mathématiques qui ont ainsi été développés devraient permettre l'utilisation à une plus grande échelle de cette prometteuse fonction de distribution, qui est d'ailleurs d'un usage courant dans les pays de l'est.

L'étude théorique des variances et des covariances des paramètres ainsi que des variances des événements  $x_T$  correspondant à une période de retour  $T$ , ceci pour un grand nombre de populations G.G. typiques de l'hydrologie, démontre que la méthode d'estimation des variances et des covariances par le maximum de vraisemblance est toujours plus efficace que l'autre méthode qui donne d'ailleurs des intervalles de confiance plus grands pour les événements  $x_T$ . Même lors de l'estimation et de l'analyse d'échantillons de petites tailles, où on obtient deux populations estimées différentes, la méthode des moments conduit toujours à des variances de  $x_T$  plus grandes ou égales à celles obtenues par la méthode du maximum de vraisemblance.

Les résultats préliminaires d'estimation des paramètres par les deux méthodes, à partir d'échantillons obtenus par génération semble démontrer qu'elles sont aussi bonnes l'une que l'autre pour l'estimation de la valeur moyenne de  $x_T$  qui est alors connue théoriquement au départ, ces aspects fe-

ront l'objet de travaux plus étendus dont les résultats seront présentés dans un autre rapport.

Il existe un phénomène d'amplification d'erreur lors du calcul de l'événement  $x_T$  qui peut causer des problèmes surtout lors de génération d'échantillons aléatoires; en effet une petite erreur dans l'évaluation de la variable gamma simple peut se traduire par une erreur appréciable sur la variable gamma généralisée. L'interpolation au troisième degré dans les tables de Harter (1969) semble être la meilleure méthode d'évaluation de la variable gamma simple, d'autant plus qu'on peut dériver ce polynôme d'interpolation pour obtenir un élément essentiel au calcul des variances et covariances. Ces mêmes tables ont fait l'objet de régressions polynomiales (Bobée et al., 1983B) qui sont utilisables avec précautions cependant; lorsque le paramètre  $S$  est négatif l'imprécision de calcul est reporté sur les débits de crue, alors que si  $S > 0$  ce sont les petits débits qui risquent d'être mal estimés. Le tableau 5 contient les éléments utiles pour effectuer un calcul d'erreur.

On peut dire finalement que la présence supplémentaire d'un paramètre de puissance  $S$  permet à la loi gamma généralisée de s'adapter à la très grande majorité des échantillons de débits de crue, le plan  $(C_V, C_S)$  est pratiquement couvert dans son entier aux fins de l'hydrologie par la méthode d'estimation par les moments, celle du maximum de vraisemblance semble être encore plus efficace.

La théorie qui a été exposée dans cet ouvrage a fait l'objet de tests numériques qui ont permis de vérifier les aspects pratiques de calcul pour les deux méthodes d'estimation des paramètres. Les programmes relatifs à la loi gamma généralisée et aux sous-routines utilitaires ( $\ln \Gamma(x)$ ,  $\psi(x)$ ,  $\psi'(x)$ ) sont disponibles (Bobée et al., 1983).

Références bibliographiques

- ABRAMOWITZ, W. et I.A. STEGUN. (1972).  
Handbook of mathematical functions. DOVER edition.
- BOBÉE, B. (1973).  
Sample error for T-year events computed by fitting a Pearson type III distribution. Wat. Res. Res., 9(5): 1264-1270.
- BOBÉE, B. (1975).  
Étude des propriétés mathématiques et statistiques des lois Pearson type III et log-Pearson type III. Rapport scientifique No 55, INRS-Eau, Québec.
- BOBÉE, B. et al. (1983).  
Ajustement des distributions Pearson type III, Gamma, gamma généralisée, Log-Pearson type III et Log-gamma. Rapport scientifique No 105, INRS-Eau, Québec.
- BOBÉE, B. et al. (1983B).  
Ajustement Polynomial de la variable Pearson type III standardisée (en préparation), INRS-Eau, Québec.
- HAGER, H.W. et L.J. BAIN. (1970).  
Inferential procedures for the generalized gamma distribution. J. Amer. Statist. Assoc., Vol. 65, (332): 1601-1609.
- HAGER, H.W., L.J. BAIN, C.E. ANTIE. (1971).  
Reliability estimation for the generalized gamma distribution and robustness of the Weibull Model. Technometrics Vol. 13, (3): 547-557.
- HARTER, H.L. (1967).  
Maximum-likelihood estimation of the parameters of a four-parameter generalized gamma population from complete and censored samples. Technometrics, Vol. 9, (1): 159-165.

HARTER, H.L. (1969).

A new table of percentage points of the Pearson type III distribution. *Technometrics*, 2(1): 177-187.

HARTER, H.L. (1966).

Asymptotic variances and covariances of maximum-likelihood estimators, from censored samples, of the parameters of a four-parameter generalized gamma population. Office of aerospace research, U.S. Air Force.

HOSHI, K. et I. YAMAOKA. (1980).

The generalized gamma probability distribution and its application in hydrology. *Hydraulics papers*, No 7. The research laboratory of civil and environmental engineering, Hokkaido University, Japan, 256 p.

LAWLESS, J.F. (1980).

Inference in the generalized gamma and log-gamma distributions. *Technometrics*, Vol. 22, (3): 409-419.

PARR, V.B. et J.T. WEBSTER. (1965).

A method for discriminating between failure density functions used in reliability predictions. *Technometrics* Vol. 7, (1): 1-10.

STACY, E.W. (1962).

A generalization of the gamma distribution. *Annals of mathematical statistics* 33(3): 1187-1192.

STACY, E.W. et G.A. MIHRAM. (1965).

Parameter estimation for a generalized gamma distribution. *Technometrics*, Vol. 7, (3): 349-358.

TADIKAMALLA, P.R. (1979).

Random sampling from the generalized gamma distribution. *Computing* 23, p. 199-203.

THOM, H.C.S. (1968).

Direct and inverse tables of the gamma distribution. Environmental data service EDS-2, U.S. Department of Commerce, 30 p.

## Annexe 1

Calcul de la fonction de distribution de la loi gamma généralisée à 3 paramètres

Cette fonction de distribution est obtenue en intégrant la fonction de densité de probabilité (équation 1.5).

$$F(x) = \int_0^x f(x;a,\lambda,S) dx \quad (A.1)$$

$$F(x) = \frac{|s|}{a^{S\lambda} \Gamma(\lambda)} \int_0^x e^{-\left(\frac{x}{a}\right)^S} x^{S\lambda-1} dx \quad (A.2)$$

On effectue le changement de variable  $W = \left(\frac{x}{a}\right)^S$ ,  $W$  suit ainsi une loi gamma simple, de plus:  $X = a W^{1/S}$   
 $dx = \frac{a}{S} W^{1/S-1} dW$

Il convient d'examiner séparément les deux cas où  $S > 0$  et  $S < 0$ .

-----

Cas  $S > 0$                       lorsque  $x = 0$  ,  $w = 0$   
   lorsque  $x = x$  ,  $w = w$

l'équation A.2 devient donc:

$$F(x) = \frac{|s|}{a^{S\lambda} \Gamma(\lambda)} \int_0^w e^{-W} (a W^{1/S})^{S\lambda-1} \frac{a}{S} W^{1/S-1} dW \quad (A.3)$$

après simplification, il reste:

## Annexe 2

Calcul de la fonction de distribution d'une loi gamma simple à l'aide d'un développement en série pour  $\Gamma(W, \lambda)$

On a vu précédemment que si  $x$  obéit à une loi gamma généralisée  $f(x; a, \lambda, S)$ , alors  $W = \left(\frac{x}{a}\right)^S$  est distribué selon une loi gamma simple. Il existe un développement en série qui converge ultimement vers la valeur théorique de la fonction gamma incomplète qu'on doit évaluer, en utilisant ce développement, qui est montré dans les tables (Abramowitz, M. et I.A. Stegun, on obtient:

$$F(x) = \frac{\Gamma(W, \lambda)}{\Gamma(\lambda)} = \frac{W^\lambda}{\Gamma(\lambda+1)e^{-W}} \left[ 1 + \frac{W}{\lambda+1} + \frac{W^2}{(\lambda+1)(\lambda+2)} + \dots + \frac{W^N}{(\lambda+1)(\lambda+2)\dots(\lambda+N)} + \dots \right]$$

(A.7)

On a utilisé cette expression comme base pour les comparaisons entre diverses méthodes plus approximatives d'évaluation de  $W$ , habituellement on fixe  $F(x)$  et on trouve  $W$  qui y correspond par la méthode itérative de Newton-Raphson, pour cela on peut dériver facilement l'expression (A.7) par rapport à  $W$ ; on obtient ainsi une autre série convergente dont les termes sont dérivés de la série initiale. Thom (1968) a d'ailleurs employé cette méthode pour le calcul de tables relatives à la loi gamma simple.

Cette façon de procéder possède le meilleur potentiel de précision pour le calcul de la variable  $W$  correspondant à un niveau de probabilité au non dépassement  $F(x)$ , le calcul de  $\frac{\partial W}{\partial \lambda}$  peut ainsi s'effectuer de la façon suivante:

$$F(x) = \frac{1}{\Gamma(\lambda)} \int_0^x e^{-W} W^{\lambda-1} dW \quad (\text{A.4})$$

l'intégrale est la définition même d'une fonction gamma incomplète, donc:

$$F(x) = \frac{\Gamma(W, \lambda)}{\Gamma(\lambda)} \quad \text{pour } S > 0 \quad (\text{A.5})$$

-----

#### Cas $S < 0$

On peut démontrer de façon semblable que la fonction de distribution s'exprime ainsi:

$$F(x) = 1 - \frac{\Gamma(W, \lambda)}{\Gamma(\lambda)} \quad \text{pour } S < 0 \quad (\text{A.6})$$

Sur le plan pratique, on peut conserver exactement la même procédure qui a été établie pour le cas  $S > 0$  mais en considérant que les probabilités au dépassement deviennent les probabilités au non dépassement.

$$\frac{\partial W}{\partial \lambda} = \frac{W(\lambda + \epsilon \lambda, F(x)) - W(\lambda - \epsilon \lambda, F(x))}{2\epsilon \lambda} \quad (\text{A.8})$$

où  $W(\lambda, F(x))$  est la variable  $W$  obtenue pour un niveau de probabilité au non dépassement  $F(x)$  et pour le paramètre  $\lambda$  considéré.

$\epsilon$  est posé arbitrairement très petit, on suggère  $\epsilon = 0,0001$ .

Le nombre de termes à utiliser pour le développement en série varie selon  $\lambda$  et  $F(x)$ , il peut être assez grand ce qui est alors un désavantage pour le temps de calcul, surtout si on désire obtenir également  $\frac{\partial W}{\partial \lambda}$ . Cependant la méthode itérative nécessite qu'environ 5 itérations et on n'a guère constaté de problèmes pour les valeurs de départ de la variable  $W$ .

Il nous a été possible grâce à ce développement et avec l'aide d'une calculatrice travaillant à 13 chiffres significatifs d'obtenir des valeurs de la variable standardisée Pearson type III, Harter (1969), avec une précision de  $10^{-6}\%$ .