

Record Number: 550
Author, Monographic: Bobée, B.
Author Role:
Title, Monographic: Étude des propriétés mathématiques et statistiques des lois Pearson type III et log-Pearson type III
Translated Title:
Reprint Status:
Edition:
Author, Subsidiary:
Author Role:
Place of Publication: Québec
Publisher Name: INRS-Eau
Date of Publication: 1975
Original Publication Date:
Volume Identification:
Extent of Work: 167
Packaging Method: pages
Series Editor:
Series Editor Role:
Series Title: INRS-Eau, Rapport de recherche
Series Volume ID: 55
Location/URL:
ISBN: 2-89146-021-9
Notes: Rapport annuel 1975-1976
Abstract: 25.00\$
Call Number: R000055
Keywords: rapport/ ok/ dl

Etude des propriétés mathématiques et
statistiques des lois Pearson type III
et log-Pearson type III

INRS-Eau
UNIVERSITE DU QUEBEC
C.P. 7500, Sainte-Foy
Québec G1V 4C7

RAPPORT SCIENTIFIQUE No 55
1975

Rapport rédigé pour
INRS-Eau

par
B. Bobée

ISBN 2-89146-021-9

DEPOT LEGAL 1975

Tous droits de traduction, de reproduction et d'adaptation réservés

c 1975 - Institut national de la recherche scientifique

TABLE DES MATIERES

	<u>Page</u>
Chapître 1: Origine et différentes formes de la loi Pearson type 3	
1.1 Origine de la loi Pearson type 3	1
1.2 Formes simplifiées de la loi Pearson type 3	4
1.3 Etude mathématique de la distribution Pearson type 3	6
Chapître 2: Distribution Log. Pearson type 3	
2.1 Fonction densité	9
2.2 Mode de la distribution	10
2.3 Etude des points d'inflexion	12
2.4 Signe de g' - Valeurs aux limites de g et g'	14
2.5 Forme des distributions	16
Chapître 3: Préléminaires statistiques	
3.1 Moments et coefficients	19
3.2 Fonction caractéristique, cumulants	21
3.3 Distribution des moments	23
3.4 Distribution d'une fonction de variables aléatoires	26
3.5 Statistiques d'ordre	28
Chapître 4: Propriétés de la loi Pearson III et de ses formes simplifiées	
4.1 Introduction	30
4.2 Fonction caractéristiques, cumulants et moments	30
4.3 Forme standardisée de la distribution Pearson III	33
4.4 Coefficients caractéristiques de la distribution Pearson III	38

TABLE DES MATIERES (suite)

	<u>Page</u>	
4.5	Distribution de la moyenne d'un échantillon tiré d'une loi Pearson III	40
4.6	Erreur type des moments et coefficients d'un échantillon	42
4.7	Erreur type d'un évènement de période de retour donnée	51
4.8	Génération d'une variable suivant une distribution Pearson III	57
4.9	Statistiques d'ordre de la distribution Pearson III	63
4.10	Détermination des intervalles de confiance de la distribution Pearson III	66
Chapître 5:	Propriétés statistiques de la loi Log. Pearson type III	
5.1	Moments de la distribution Log. Pearson III	87
5.2	Détermination des principaux coefficients de la loi Log. Pearson III	90
5.3	Comportement de $\left(C_v \right)_L$ et $\left(C_s \right)_L$	94
5.4	Lien entre les distributions Pearson III et Log. Pearson III	97
Chapître 6:	Estimation des paramètres d'une distribution	
6.1	Introduction	100
6.2	Estimateurs corrects et absolument corrects	100
6.3	Estimateurs efficaces	101
6.4	Statistiques exhaustives	104
6.5	Cas où l'intervalle de définition dépend d'un paramètre	106

TABLE DES MATIERES (suite)

	<u>Page</u>	
6.6	Estimation de plusieurs paramètres	107
6.7	Méthode du Maximum de vraisemblance	108
6.8	Méthode des moments	112
6.9	Estimation du paramètre d'échelle et de position d'une distribution	113
6.10	Conclusion	116
Chapître 7:	Estimation des paramètres de la distribution Pearson III	
7.1	Distribution Pearson III à 1 paramètre	117
7.2	Estimation des paramètres de la loi Gamma	120
7.3	Estimation des paramètres de la distribution Pearson III	127
Chapître 8:	Estimation des paramètres de la loi Log. Pearson III	
8.1	Estimation des paramètres de la loi Log. Gamma	141
8.2	Estimation des paramètres de la loi Log. Pearson III	153

LISTE DES TABLES

		<u>Page</u>
Table 1	Etude du Mode de la distribution Log. Pearson III	12
Table 2	Caractéristiques de la loi Pearson III et de ses formes simplifiées	31
Table 3	Cumulants de la loi Pearson III et de ses formes simplifiées	34
Table 4	Moments de la distribution Pearson III et de ses formes simplifiées	35
Table 5	Erreurs types pour les différentes formes de la loi Pearson III et pour la loi Normale	48
Table 6	Influence de la variabilité de K sur les intervalles de confiance lorsque $\bar{C}_s = 2 \bar{C}_v$	58
Table 7	Influence de la variabilité de K sur les intervalles de confiance lorsque $\bar{C}_s \neq 2 \bar{C}_v$	59
Table 8	Statistiques d'ordre de la loi Pearson III standardisée pour $C_s = 0$	67
Table 9	Statistiques d'ordre de la loi Pearson III standardisée pour $C_s = 1$	70
Table 10	Statistiques d'ordre de la forme dérivée de la loi Pearson III pour $C_s = -1$	73
Table 11	Intervalles de confiance de la loi Pearson III standardisée ($C_s = 0$)	84
Table 12	Caractéristiques générales des distributions Pearson III et Log. Pearson III	99
Table 13	Efficacité de la méthode des moments: loi Pearson III à 1 paramètre	120
Table 14	Efficacité de la méthode des moments pour la loi Gamma	126
Table 15	Efficacité relative des paramètres α, λ, m de la loi Pearson III	139

LISTE DES FIGURES

		<u>Page</u>
Fig. 1	Fonction densité de la distribution Pearson 3	7
Fig. 2	Relation entre les formes A et B de la loi Pearson 3	7
Fig. 3	Signe de δ	12
Fig. 4	Détermination du nombre de points d'inflexion	15
Fig. 5	Différentes formes de la fonction densité de la loi Log. Pearson III	17
Fig. 6	Erreur-type, loi Normale	49
Fig. 7	Erreur-type, loi Gamma	50
Fig. 8	Statistiques d'ordre de la loi Pearson III standardisée ($C_s = 0$, $N = 101$)	69
Fig. 9	Statistiques d'ordre de la loi Pearson III standardisée ($C_s = 1$, $N = 101$)	72
Fig. 10	Statistiques d'ordre de la loi Pearson III standardisée ($C_s = -1$, $N = 101$)	75
Fig. 11	Relation entre les limites des intervalles de confiance	79
Fig. 12	Intervalles de confiance à 95% de la loi Pearson III pour $C_s = 0$	81
Fig. 13	Intervalles de confiance à 90% de la loi Pearson III pour $C_s = 0$	82
Fig. 14	Intervalles de confiance à 80% de la loi Pearson III pour $C_s = 0$	83
Fig. 15	Relation entre $\left(C_s\right)_L$ et $\left(C_v\right)_L$ pour λ fixé	92
Fig. 16	Relation entre $\left(C_s\right)_L$ et $\left(C_v\right)_L$ pour β fixé	93
Fig. 17	Efficacité de la méthode des moments, loi Log. Gamma (paramètres d'échelle α)	154
Fig. 18	Efficacité de la méthode des moments, loi Log. Gamma (paramètre de forme λ)	155

LISTE DES TABLES (suite)

		<u>Page</u>
Table 16	Efficacité des paramètres de la loi Pearson III, λ étant fixé	140
Table 17	Relation entre α et A (loi Log. Gamma)	146
Table 18	Relation entre β et B (loi Log. Pearson III)	161
Table 19	Relation entre α et B (logarithmes décimaux)	164

AVANT PROPOS

Les fonctions de distributions statistiques jouent un rôle important pour représenter des séries de données expérimentales indépendantes. La famille des lois Pearson en raison des formes variées qu'elle peut prendre et de sa généralité, a été employée dans de nombreux domaines. De nombreuses applications en hydrologie et météorologie mais aussi dans le domaine médical et électronique, ont montré que les diverses formes de la loi Pearson III sont souvent satisfaisantes et ont conduit à l'emploi fréquent de ce type de distribution.

En 1968, le Conseil des Ressources en Eau des Etats-Unis a recommandé l'usage systématique de la distribution Log.Pearson III pour représenter les séries de débits de crue annuels et a insisté sur la nécessité de poursuivre les études concernant cette loi, qui a pour cas limite la distribution Log.Normale. D'autre part, dans de nombreux pays, en particulier en Europe de l'Est, la distribution Pearson III est d'un usage courant en hydrologie.

Au cours des dernières années on a donc assisté à une intensification des études concernant les lois Pearson III et Log. Pearson III. C'est dans ce courant d'idées que s'inscrit cet ouvrage dont les objectifs principaux sont, de faire d'une part la synthèse des travaux connus tout en présentant la loi Pearson III sous sa forme générale; de montrer d'autre part les développements récents qui concernent la loi Log Pearson III.

On a particulièrement insisté dans ce travail sur les aspects qui ont une importance pratique (intervalles de confiance, méthodes d'ajustement) tout en rappelant les principales bases statistiques nécessaires à une bonne compréhension. Outre les références citées au cours des différents chapîtres, le lecteur pourra trouver à la fin de l'ouvrage une liste de publications, essen-

tiellement dans le domaine hydrologique, qui montrent certaines utilisations pratiques des lois Pearson III, Log. Pearson III et de leurs formes simplifiées.

Ce livre est le résultat d'un travail qui s'est poursuivi durant plusieurs années dans le cadre de l'Institut National de la Recherche Scientifique, avec l'aide partielle de subventions du Conseil National des Recherches du Canada. Je tiens à remercier messieurs Michel Slivitzky et Jean-Pierre Villeneuve, l'intérêt qu'ils ont montré pour ces travaux m'a profondément stimulé. Mes remerciements vont également à monsieur Yves Brunet-Moret de l'ORSTOM pour les fructueuses discussions que nous avons eues et à monsieur Nicholas Matalas pour les commentaires qu'il m'a exprimés à l'occasion de publications dans la revue Water Resource Research. Je remercie particulièrement messieurs Roland Robitaille et Guy Morin pour leur collaboration dans l'établissement des tables, messieurs Magella Cantin et André Parent pour leur support technique et madame Solanges Marcoux pour son travail de déchiffrement et de dactylographie.

1. ORIGINE ET DIFFERENTES FORMES DE LA LOI PEARSON TYPE III

1.1 Origine de la loi Pearson type III

On considère une fonction densité de probabilité $y = f(x)$ continue de la variable x , satisfaisant l'équation différentielle:

$$\frac{1}{f} \frac{df}{dx} = \frac{a - x}{b_0 + b_1 x + b_2 x^2} \quad (1.1)$$

où a, b_0, b_1, b_2 sont des paramètres

La dérivée $\frac{df}{dx}$ s'annule lorsque $f(x)$ s'annule ou lorsque $x = a$,

ce qui montre que la distribution représentée par $f(x)$ peut avoir un mode pour $x = a$ et que la courbe représentative peut être tangente à l'axe des x à l'une ou aux deux extrémités de l'intervalle de variation de x . Une fonction densité de probabilité possédant ces propriétés générales représente souvent adéquatement des séries de données; la famille des fonctions densité déduites de l'intégration de (1.1), qui constitue l'ensemble des distributions Pearson joue un rôle très important en statistique; cet ensemble de distribution qui a été étudié en détail et caractérisé par CRAIG (1935), ELDERTON (1938), HOADLEY (1968) comprend en particulier :

- la loi Bêta de 1ère espèce (Pearson type I) obtenue lorsque les racines du trinôme $(b_0 + b_1 x + b_2 x^2)$ sont réelles et de signe opposé

- la loi Bêta de 2ième espèce (Pearson type VI) obtenue lorsque les racines du trinôme $(b_0 + b_1 x + b_2 x^2)$ sont réelles et de même signe

- la loi Pearson type III obtenue lorsque $b_2 = 0$, qui comprend comme cas limite la loi normale si de plus $b_1 = 0$.

A partir de l'équation différentielle générale dans laquelle $b_2 = 0$, il est possible de déduire par intégration la forme générale de la distribution Pearson type III, l'équation (1.1) s'écrit:

$$\frac{1}{f} \frac{df}{dx} = -\alpha + \frac{(\lambda-1)}{(x-m)} \quad (1.2)$$

avec:

$$\alpha = \frac{1}{b_1}$$

$$m = -\frac{b_0}{b_1}$$

$$(\lambda-1) = \frac{1}{b_1} \left(a + \frac{b_0}{b_1} \right)$$

L'intégration de (1.2) conduit à:

$$\text{Log } |f| = -\alpha x + (\lambda-1) \text{Log } |x-m| + \text{Constante}$$

f est toujours positive puisque c'est une fonction de densité d'où

$$f = K e^{-\alpha x} |x-m|^{\lambda-1}$$

K étant une constante telle que si Δ est l'intervalle de définition de x :

$$\int_{\Delta} f(x) dx = 1$$

Deux cas sont à envisager:

a) $x - m \geq 0$ ou $x \geq m$, on a alors:

$$\int_m^{+\infty} K e^{-\alpha x} (x-m)^{\lambda-1} dx = 1$$

en posant $v = |\alpha| (x-m)$ la relation précédente devient :

$$K \frac{e^{-\alpha m}}{|\alpha|^\lambda} \int_0^{+\infty} e^{-\frac{v}{|\alpha|}} v^{\lambda-1} dv = 1$$

l'intégrale converge si et seulement si $\alpha > 0$ et $\lambda > 0$ et l'on a alors :

$$K = \frac{\alpha^\lambda}{e^{-\alpha m} \Gamma(\lambda)} \quad \text{avec} \quad \Gamma(\lambda) = \int_0^{+\infty} e^{-v} v^{\lambda-1} dv \quad (\text{fonction Gamma})$$

La fonction densité s'écrit :

$$f(x) = \frac{\alpha}{\Gamma(\lambda)} e^{-\alpha(x-m)} [\alpha(x-m)]^{\lambda-1} \quad (1.3)$$

FORME A

avec $m \leq x$; $\lambda > 0$; $\alpha > 0$

b) $x-m \leq 0$ ou $x \leq m$, K est alors déterminé par :

$$\int_{-\infty}^m K e^{-\alpha x} (m-x)^{\lambda-1} dx = 1$$

si $v = |\alpha| (m-x)$ la relation précédente devient :

$$K \frac{e^{-\alpha m}}{|\alpha|^\lambda} \int_0^{+\infty} e^{\frac{v}{|\alpha|}} v^{\lambda-1} dv = 1$$

L'intégrale converge si et seulement si $\alpha < 0$ et $\lambda > 0$ et l'on a :

$$f(x) = \frac{-\alpha}{\Gamma(\lambda)} e^{-\alpha(x-m)} [\alpha(x-m)]^{\lambda-1} \quad (1.4)$$

Forme B

avec $x \leq m$; $\alpha < 0$; $\lambda > 0$

La distribution Pearson type III possède donc 2 formes suivant le signe de α , l'une (Forme A) pour $\alpha > 0$ correspond comme nous le verrons à une asymétrie positive, tandis que pour $\alpha < 0$ (Forme B) la distribution a une asymétrie négative.

Puisque l'on a toujours $\alpha(x-m) \geq 0$ la synthèse des deux cas est donnée par:

$$f(x) = \frac{|\alpha|}{\Gamma(\lambda)} e^{-\alpha(x-m)} [\alpha(x-m)]^{\lambda-1} \quad (1.5)$$

L'intervalle de définition de x est tel que $\alpha(x-m) \geq 0$, donc:

$$\text{si } \alpha > 0 \quad m \leq x < +\infty \quad (\text{Forme A})$$

$$\text{si } \alpha < 0 \quad -\infty < x \leq m \quad (\text{Forme B})$$

1.2 Formes simplifiées de la loi Pearson type III

La distribution Pearson type III définie par (1.5) dépend de 3 paramètres:

m paramètre de position, est la borne inférieure ou supérieure de l'intervalle de définition de x suivant que α est positif ou négatif

α est un paramètre d'échelle

λ est un paramètre de forme

Suivant les valeurs de ces paramètres, la distribution Pearson type III peut prendre des formes simplifiées.

Si $m = 0$, on obtient la distribution gamma qui bien que définie, quel que soit α est surtout connue et utilisée pour $\alpha > 0$:

$$f(x) = \frac{|\alpha|}{\Gamma(\lambda)} e^{-\alpha x} (\alpha x)^{\lambda-1} \quad \underline{\text{Loi gamma}} \quad (1.6)$$

avec $\lambda > 0$

$$0 \leq x < +\infty \quad \text{si } \alpha > 0$$

$$-\infty < x \leq 0 \quad \text{si } \alpha < 0$$

La loi gamma comprend comme cas particulier la distribution chi-deux (χ^2), en effet χ^2 avec ν degrés de liberté est une loi gamma pour laquelle $\alpha = \frac{1}{2}$ $\lambda = \frac{\nu}{2}$.

La loi gamma est reliée à d'autres distributions en effet:

- la loi binomiale se comporte asymptotiquement comme une loi gamma.
- la distribution d'Erlang (qui intervient dans la théorie des queues) est un cas particulier de la loi gamma pour lequel λ prend des valeurs entières (GREENWOOD et DURAND; 1960) lorsque $\lambda = 1$ on obtient la distribution exponentielle négative.
- la distribution binomiale peut être obtenue à partir d'une loi de Poisson dont le paramètre suit une loi gamma (CRAMER 1971)
- la distribution Weibull-Goodrich est un cas particulier de la loi gamma (KRITSKY et MENKEL, 1969).

Si $m = 0$ et $\alpha = 1$, on obtient la loi Pearson type III à un paramètre:

$$g(u) = \frac{e^{-u}}{\Gamma(\lambda)} u^{\lambda-1} \quad \begin{array}{l} 0 \leq u < +\infty \\ 0 < \lambda \end{array} \quad (1.7)$$

Cette forme simplifiée peut se déduire de la forme (1.5) par le changement de variable $u = \alpha(x-m)$, on a en effet:

$$g(u) = f \left[x(u) \right] \left| \frac{dx}{du} \right|$$

La distribution Pearson type III est donc très générale, ce qui explique le rôle important qu'elle joue en statistique.

1.3 Etude mathématique de la distribution Pearson type III

Cette étude a été effectuée par plusieurs auteurs en particulier par BRUNET-MORET (1969), dans le cas $\alpha > 0$:

$0 < \lambda < 1$: pas de mode réel ni de point d'inflexion.

$$f(m) = +\infty ; f(+\infty) = 0 ; f'(m) = -\infty ; f'(+\infty) = 0,$$

on a une courbe en forme de J renversé.

$\lambda = 1$ loi exponentielle, ni mode réel ni point d'inflexion

$$f(m) = 1 ; f(+\infty) = 0 ; f'(m) = -1 ; f'(+\infty) = 0$$

$1 < \lambda < 2$ Mode pour $x = m + \frac{1}{\alpha} (\lambda - 1)$, un point d'inflexion pour

$$x = m + \frac{1}{\alpha} (\lambda - 1 + \sqrt{\lambda - 1})$$

$$f(m) = 0 ; f(+\infty) = 0 ; f'(m) = +\infty ; f'(+\infty) = 0$$

$\lambda = 2$ Mode pour $x = m + \frac{1}{\alpha}$, point d'inflexion pour $x = m + \frac{2}{\alpha}$.

$$f(m) = 0 ; f(+\infty) = 0 ; f'(m) = +1 ; f'(+\infty) = 0$$

$\lambda > 2$ Mode pour $x = m + \frac{1}{\alpha} (\lambda - 1)$, 2 points d'inflexion pour

$$x = m + \frac{1}{\alpha} \left[\lambda - 1 \pm \sqrt{\lambda - 1} \right]$$

$$f(m) = 0 ; f(+\infty) = 0 ; f'(m) = 0 ; f'(+\infty) = 0$$

La figure 1 donne les différentes formes de la fonction densité suivant les valeurs de λ .

Si $\alpha < 0$, les courbes densités de probabilité de la forme B de la

loi Pearson III sont déduites de celle de la forme A par une symétrie par rapport à la droite $x = 0$; en effet:

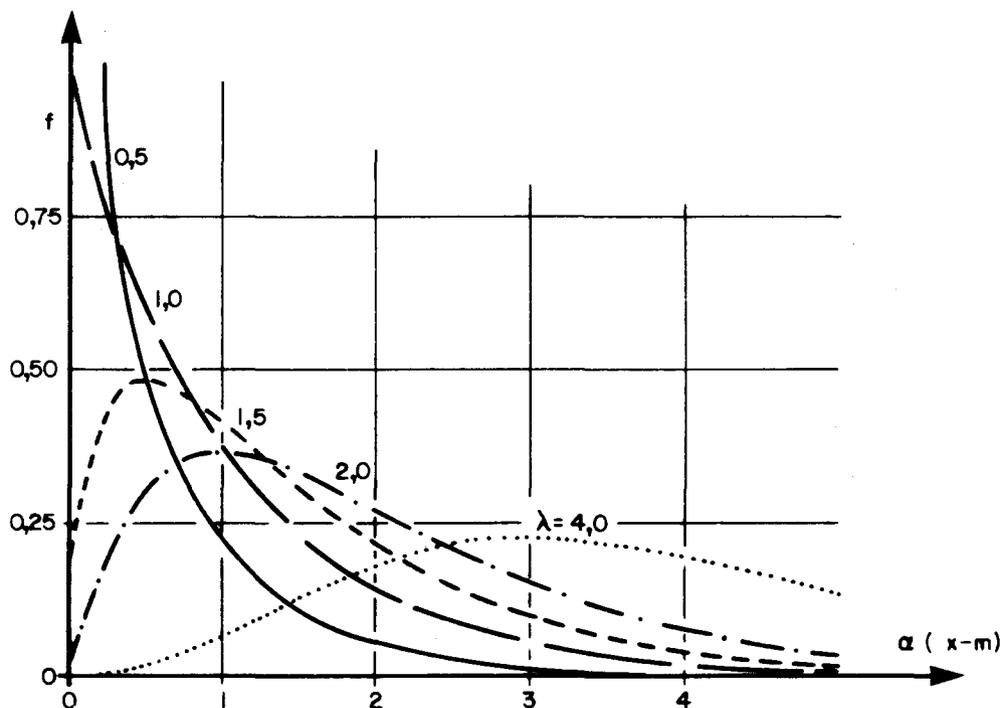


FIG 1 . Densité de la distribution Pearson III

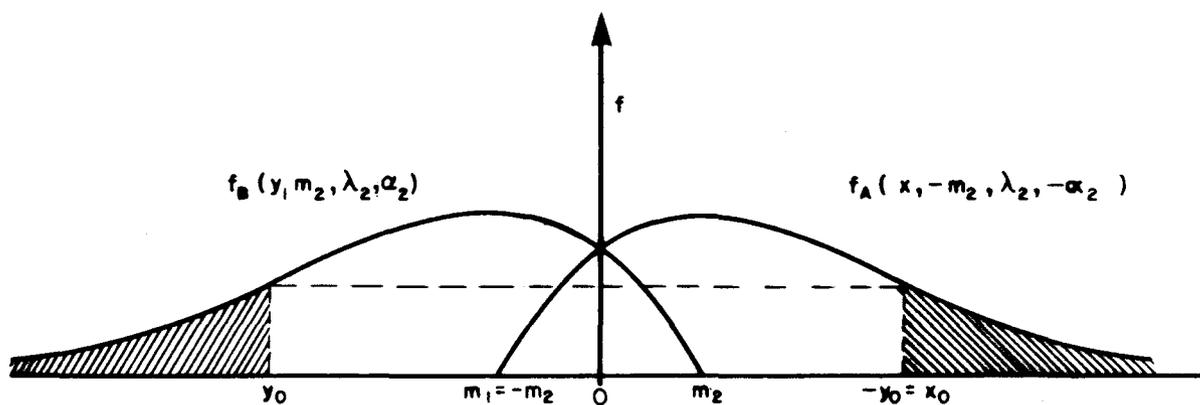


FIG 2 . Relation entre les densités des formes A et B de la loi Pearson III

soient x et y deux variables suivant la loi Pearson III respectivement de forme A ($\alpha > 0$) de densité $f_A(x, \alpha_1, \lambda_1, m_1)$, et, de forme B ($\alpha < 0$) de densité $f_B(y, \alpha_2, \lambda_2, m_2)$. En considérant les relations (1.3) et (1.4) on a :

$$f_B(y, m_2, \lambda_2, \alpha_2) = f_A(-y, -m_2, \lambda_2, -\alpha_2) \quad (1.8)$$

$$\alpha_2 < 0 ; \lambda_2 > 0 ; -\infty < y < m_2$$

La fonction f_B se déduit donc de f_A par symétrie (fig. 2) avec $m_1 = -m_2$;

$$\lambda_2 = \lambda_1 ; \alpha_2 = -\alpha_1.$$

Il en résulte que la probabilité au non dépassement, $P[y \leq y_0]$

de la loi Pearson III de forme B est égale à la probabilité au dépassement de la loi de forme A, en effet :

$$P[y \leq y_0] = \int_{-\infty}^{y_0} f_B(y) dy = \int_{-y_0}^{+\infty} f_A(x) dx = P[x \geq -y_0] \quad (1.9)$$

Cette propriété permet donc de déterminer les percentiles de la forme B de la distribution Pearson III connaissant ceux de la forme A.

2. DISTRIBUTION LOG-PEARSON TYPE III

La distribution Log-Pearson III ne joue un rôle en hydrologie statistique que depuis peu de temps, puisque ce n'est qu'en 1968 que le comité d'Hydrologie du Water Resources Council des Etats-Unis (W.R.C.) a recommandé son emploi systématique pour la représentation des données de crues annuelles (BENSON, 1968).

De la même manière que la loi Pearson III comprend comme cas limite la loi Normale, la loi Log-Normale est un cas particulier de la distribution Log-Pearson III.

Si du point de vue théorique certaines propriétés de la distribution Log-Pearson III peuvent se déduire de la loi Pearson III, il est cependant important d'étudier la distribution Log. Pearson III pour mettre en évidence ses particularités. Dans ce chapitre nous étudions en détail les propriétés mathématiques de la fonction densité de la loi Log-Pearson III.

2.1 Fonction densité de la loi Log-Pearson III

Si $y = \log_a x$ suit une loi Pearson III, x suit une loi Log-Pearson III, la fonction densité de y est d'après la relation (1.5) :

$$f(y) = \frac{|\alpha|}{\Gamma(\lambda)} e^{-\alpha(y-m)} [\alpha(y-m)]^{\lambda-1}$$

on peut en déduire $g(x)$ fonction densité de x , puisque

$$g(x) = f \left[y(x) \right] \left| \frac{dy}{dx} \right|$$

$$y = \log_a x = k \ln x \text{ ou } x = a^y = e^{y/k}$$

$$\text{avec } k = \log_a e = \frac{1}{\ln a} \quad (\ln \text{ signifie logarithme népérien})$$

On suppose $k > 0$ c'est-à-dire $a > 1$

on a :

$$g(x) = \frac{|\alpha|}{\Gamma(\lambda)} e^{-\alpha(\log_a x - m)} \left[\alpha(\log_a x - m) \right]^{\lambda-1} \frac{k}{x}$$

avec $0 < \lambda$ et $-\infty < m < +\infty$

L'intervalle de variation est D tel que

$$\text{si } \alpha > 0 \quad a^m = e^{m/k} \leq x < +\infty$$

$$\text{Si } \alpha < 0 \quad 0 \leq x \leq a^m = e^{m/k}$$

La fonction $g(x)$ peut encore s'écrire :

$$g(x) = \frac{|\alpha| k}{\Gamma(\lambda)} e^{\alpha m} \frac{1}{x^{1+\alpha k}} \left[\alpha(k \ln x - m) \right]^{\lambda-1} \quad (2.1)$$

Le calcul des dérivées première et seconde de $g(x)$ donne :

$$g'(x) = \frac{|\alpha| k \alpha}{\Gamma(\lambda)} \frac{e^{\alpha m}}{x^{2+\alpha k}} \left[\alpha(k \ln x - m) \right]^{\lambda-2} * A \quad (2.2)$$

avec $A = (\lambda-1) k - (1 + \alpha k) (k \ln x - m)$

$$g''(x) = \frac{|\alpha| k}{\Gamma(\lambda)} \frac{e^{\alpha m}}{x^{3+\alpha k}} \left[\alpha(k \ln x - m) \right]^{\lambda-3} * T \quad (2.3)$$

avec

$$T = U^2 (1+\alpha k) (2+\alpha k) - U \alpha k (\lambda-1) (3+2\alpha k) + \alpha^2 k^2 (\lambda-1) (\lambda-2) \quad (2.4)$$

$$U = \alpha (k \ln x - m)$$

2.2 Mode de la distribution

Le mode de la distribution est obtenu lorsque A s'annule.

La valeur modale x_M est telle que :

$$(\lambda-1) k - (1 + \alpha k) (k \ln x_M - m) = 0$$

puisque l'on a toujours quelque soit le signe de α

$\alpha (k \ln x - m) \geq 0$, le mode n'existe que si:

$$\frac{(\lambda-1) \alpha k}{1 + \alpha k} > 0$$

La valeur modale lorsqu'elle existe est atteinte pour:

$$x_M = e^{\frac{m}{k}} \cdot e^{\frac{(\lambda-1)}{1 + \alpha k}} \quad (2.5)$$

on a effectivement un mode (maximum) si $g''(x_M) < 0$, autrement il s'agit d'un minimum, la table 1 déduite de l'étude du signe de $g''(x_M)$ indique les différents cas possibles en fonction de α et λ (k et λ sont positifs).

La fonction densité de probabilité de x_M est

$$g(x_M) = e^{-\frac{m}{k}} \frac{|\alpha| k}{\Gamma(\lambda)} \cdot \frac{1}{e^{\lambda-1}} \left[\frac{(\lambda-1) \alpha k}{1 + \alpha k} \right]^{\lambda-1} \quad (2.6)$$

quand $\lambda=1$ l'extremum est atteint à une borne.

Cas particuliers

Si $m = 0$, on est dans le cas de la loi Log-gamma, c'est-à-dire que $y = \log_a x$ suit une loi gamma. Dans ce cas, le mode $x_{M,0}$

et la fonction densité $g(x_{M,0})$ sont données par:

$$x_{M,0} = e^{\frac{\lambda-1}{1+\alpha k}} \quad (2.7)$$

$$g(x_{M,0}) = \frac{|\alpha| \cdot k}{\Gamma(\lambda)} \cdot \frac{1}{e^{\lambda-1}} \left[\frac{(\lambda-1) \alpha k}{1+\alpha k} \right]^{\lambda-1} \quad (2.8)$$

Quel que soit m , on a $x_M \cdot g(x_M) = x_{M,0} g(x_{M,0}) = \text{constante}$.

	$\alpha < -\frac{1}{k}$	$-\frac{1}{k} < \alpha < 0$	$\alpha > 0$
$0 < \lambda < 1$	Pas d'extremum	Minimum	Pas d'extremum
$\lambda > 1$	Extremum: Mode	Pas d'extremum	Extremum: Mode

TABLE I : Etude du mode de la distribution Log Pearson III

Donc le point de coordonnées $[x_M, g(x_M)]$ se déplace sur une hyperbole quand m varie, pour λ et α fixés.

Suivant la base a considérée dans $y = \log_a x$, le paramètre k prend différentes valeurs, en particulier:

si $y = \ln x$ (base e) on a: $k = 1$

si $y = \log_{10} x$ (base 10) on a: $k \approx 0.434$

2.3 Etude des points d'inflexions

Les points d'inflexion sont obtenus lorsque g'' s'annule, donc d'après l'expression (2.3) si le trinôme T s'annule. La recherche des points d'inflexion nous conduit à déterminer les racines positives de T puisque l'on a toujours :

$$U = \alpha(k \ln x - m) > 0.$$

(le cas $U = 0$ n'est pas considéré) il y aura autant de points d'inflexion que de racines réelles positives de T .

Le discriminant du trinôme est:

$$\delta = \beta^2 (\lambda - 1) \left[4\beta^2 + 12\beta + 7 + \lambda \right] \quad (2.9)$$

avec $\beta = \alpha k$

Le signe de δ est donné par la figure 3.

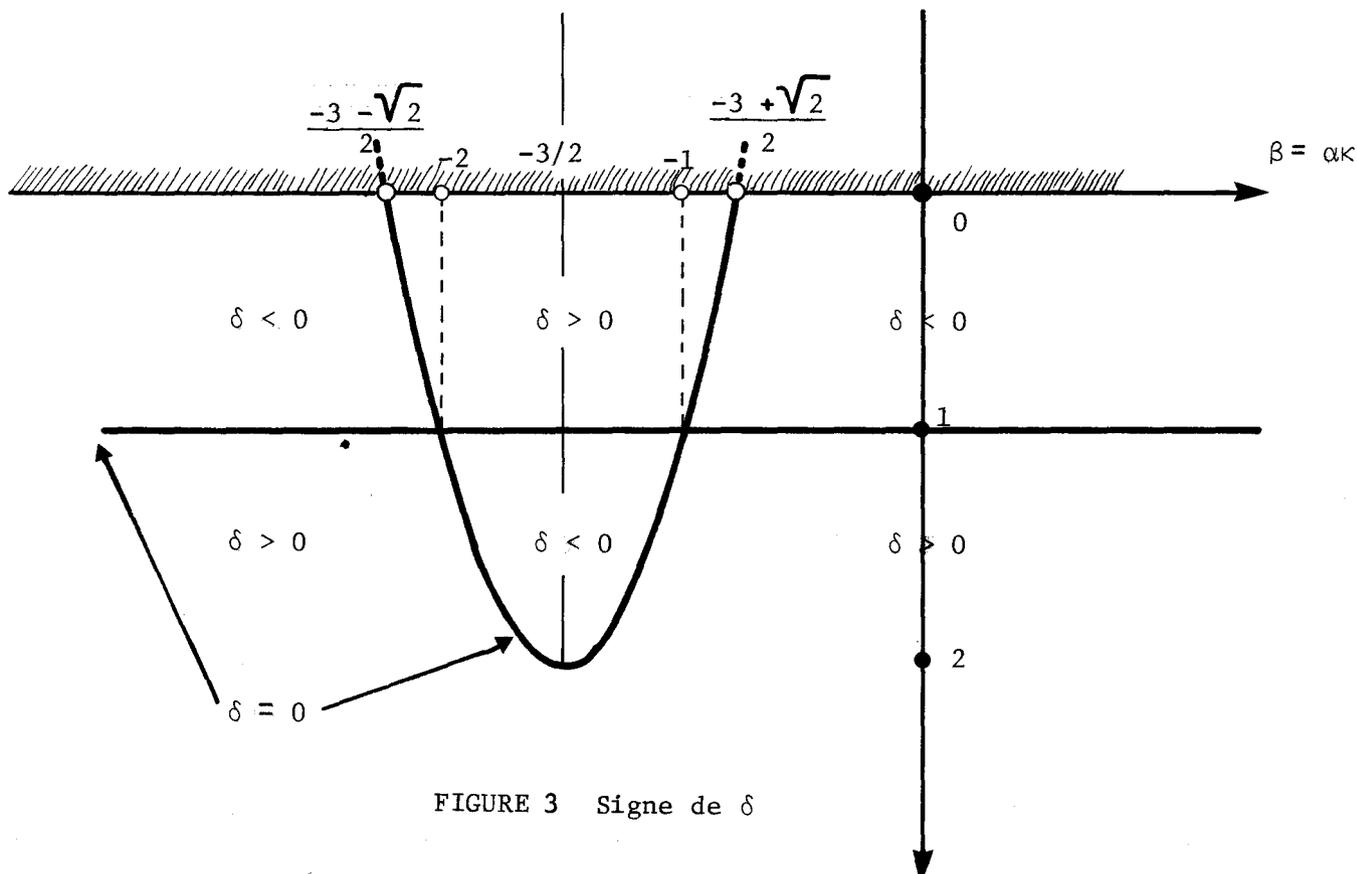


FIGURE 3 Signe de δ

- Si $\delta < 0$, il n'y a pas de point d'inflexion et g'' est du signe de T donc de $(1 + \beta)(2 + \beta)$
- Si $\delta > 0$, il y a:
 - 2 points d'inflexions si les 2 racines sont positives
 - 1 point d'inflexion si une seule racine est positive
 - aucun point d'inflexion si les 2 racines sont négatives

g'' est du signe de $(1 + \beta) (2 + \beta)$ à l'extérieur des racines

g'' est du signe $-(1 + \beta) (2 + \beta)$ entre les racines.

- Si $\delta = 0$, lorsque la racine double est négative, il ne peut y avoir inflexion, et lorsque la racine double est positive, il n'y a pas inflexion en réalité car g'' s'annule également pour la valeur de la racine double (branche AB de la parabole fig. 4)

Lorsqu'il y a possibilité de point d'inflexion ($\delta > 0$) l'étude de la somme et du produit des racines du trinôme T permet la détermination du nombre effectif des points d'inflexion (fig.4).

Lorsqu'il y a 1 ou 2 points d'inflexion, la somme et le produit des racines de T, ainsi que l'étude de g' et g'' permettent de situer ces points par rapport au mode de la distribution s'il existe.

2.4 Signe de g' - Valeurs aux limites de g et g'

Pour pouvoir donner la forme générale des courbes, nous devons connaître le signe de g' ainsi que les limites de g et g' pour les bornes de l'intervalle de variation.

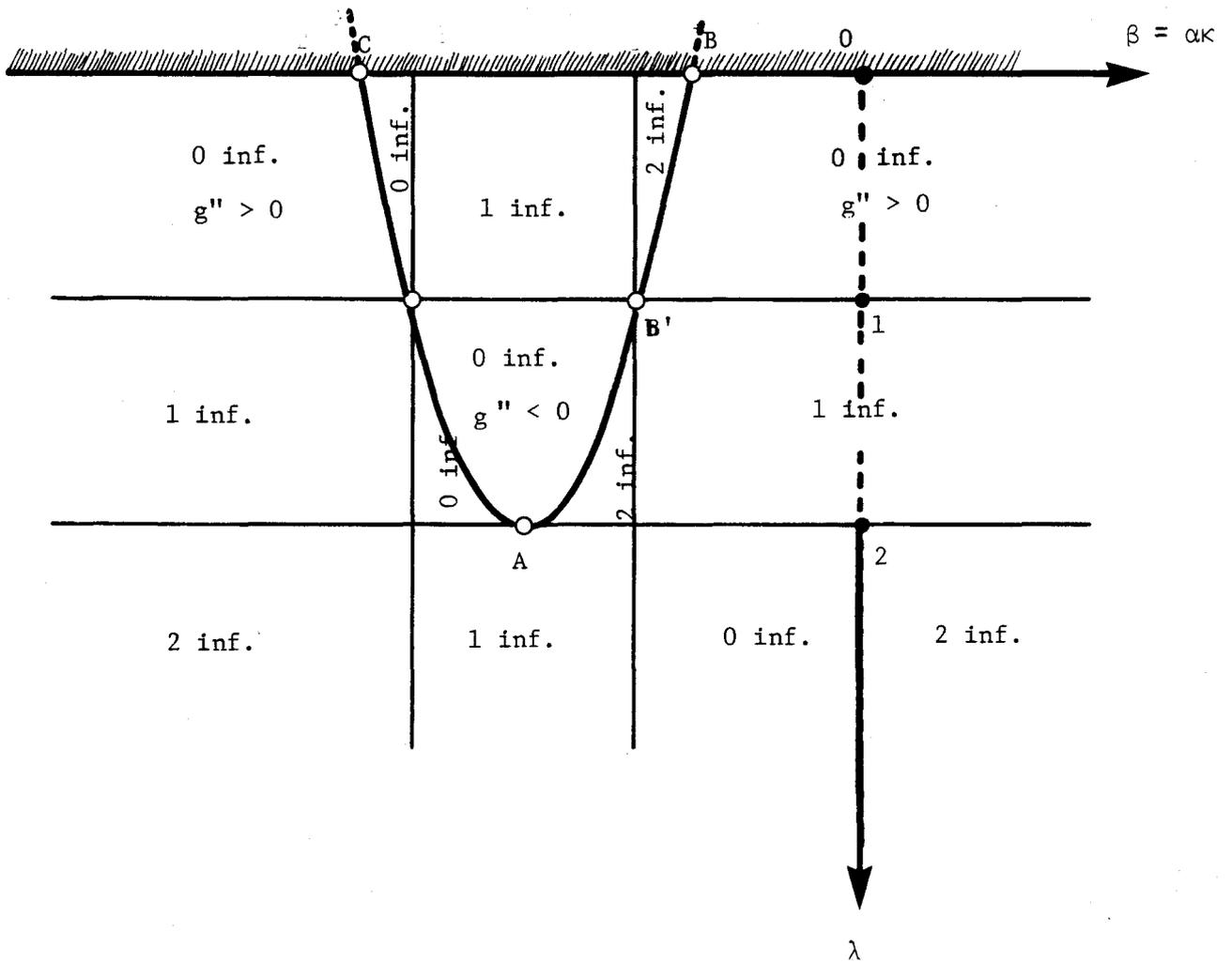
Signe de g'

Lorsque la fonction densité possède un extremum, on a, en tenant compte du fait que x et $U = \alpha(k \ln x - m)$ sont positifs:

$$\text{signe } g' = \text{signe} \left[\alpha(1+\alpha k) \ln \frac{x_M}{x} \right] = \text{signe} \left[\beta(1+\beta) \ln \frac{x_M}{x} \right]$$

Si g' ne s'annule pas, et garde un signe constant dans l'intervalle de variation, la fonction g n'a pas d'extremum et:

$$\text{signe } g' = \text{signe} \left[\alpha(\lambda-1) \right] = \text{signe} \left[\beta(\lambda-1) \right]$$



quand $\delta > 0$ on indique le nombre de points d'inflexion
quand $\delta < 0$ on indique le signe de g''

FIGURE 4 Détermination du nombre de points d'inflexion

Valeurs aux limites de g et g'

Les valeurs de g et g' pour les bornes de l'intervalle de variation sont telles que:

Si $\alpha > 0$; $e^{m/k} \leq x < +\infty$

$$g \left(e^{m/k} \right) = \lim_{U \rightarrow 0} U^{\lambda-1}$$

$$g(+\infty) = \lim_{x \rightarrow +\infty} x^{-(1+\beta)}$$

$$\left| g' \left(e^{m/k} \right) \right| = \lim_{U \rightarrow 0} U^{\lambda-2}$$

$$\left| g'(+\infty) \right| = \lim_{x \rightarrow +\infty} x^{-(2+\beta)}$$

Si $\alpha < 0$; $0 \leq x < e^{m/k}$

$$g(0) = \lim_{x \rightarrow 0} x^{-(1+\beta)}$$

$$g \left(e^{m/k} \right) = \lim_{U \rightarrow 0} U^{\lambda-1}$$

$$\left| g'(0) \right| = \lim_{x \rightarrow 0} x^{-(2+\beta)}$$

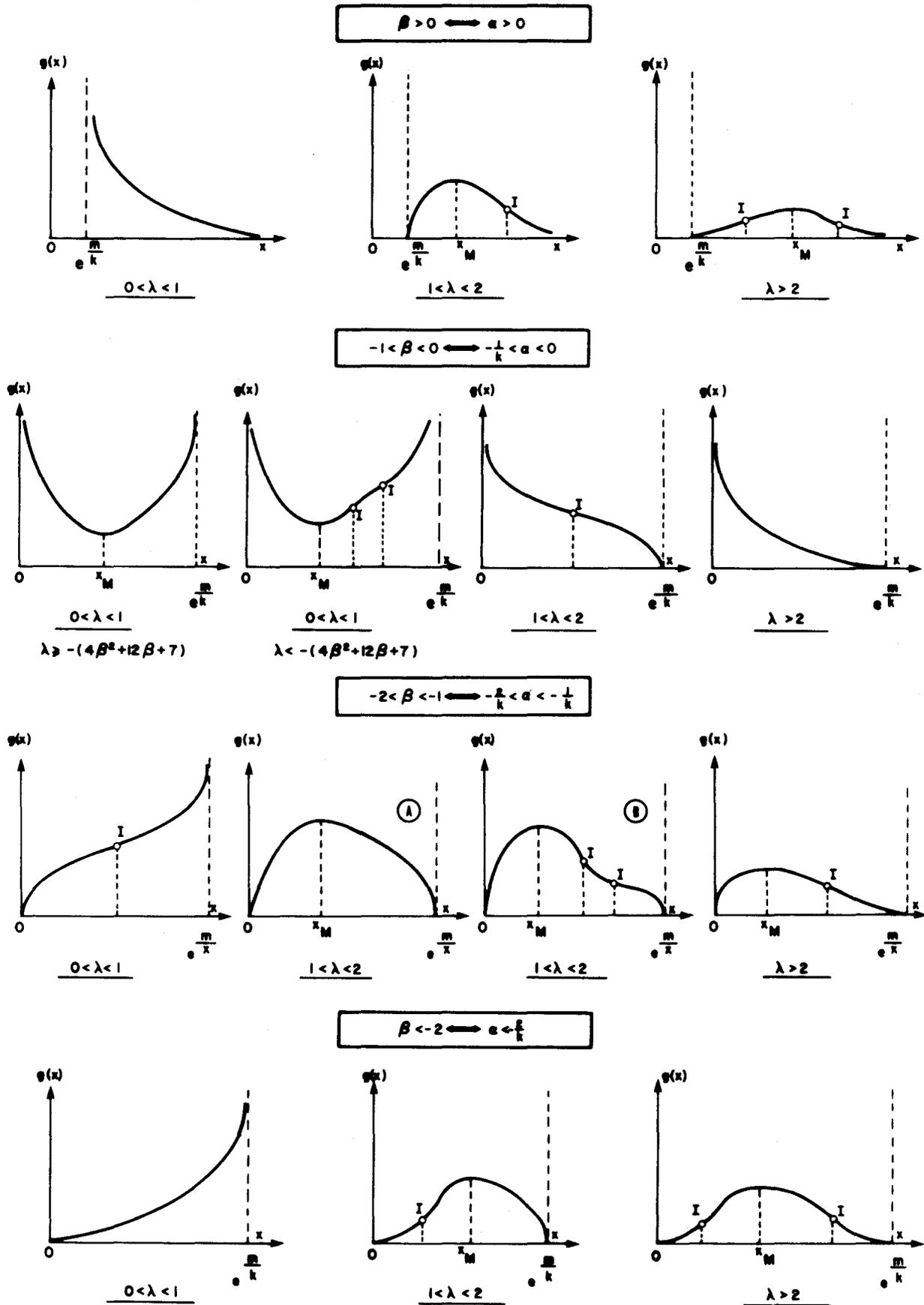
$$\left| g' \left(e^{m/k} \right) \right| = \lim_{U \rightarrow 0} U^{\lambda-2}$$

2.5 Forme des distributions

La figure 5 donne la forme des distributions pour les différentes valeurs de λ et de $\beta = \alpha k$

Ces formes sont définies en fonction des paramètres α et λ de la loi Pearson III et en fonction de k donc de la base choisie dans la transformation logarithmique, en particulier:

FIGURE 5. Différentes formes de la fonction densité de la loi log-Pearson III



Si $a = e$, on a: $y = \ln x$ avec $k = 1$

Si $a = 10$ $y = \log_{10} x$, $k = \log_{10} e \approx .434$

Les cas spéciaux tels que $\lambda = 1$, $\beta = -1$, $\beta = -2$, ne présentent pas beaucoup d'intérêt et s'étudient facilement en considérant les relations (2.1), (2.2), et (2.3) donnant la fonction et ses dérivées.

Le cas de la loi Log-gamma ($m=0$) se déduit de la figure 5, l'intervalle de variation devient $(0,1)$ pour $\alpha < 0$ et $(1, +\infty)$ pour $\alpha > 0$.

A partir des résultats concernant les dérivées première et seconde, l'existence d'un mode et de points d'inflexions, les limites de la fonction et de sa dérivée aux bornes de l'intervalle de variation, on peut mettre en évidence les différentes formes de la fonction densité de la loi Log-Pearson III, (Fig. 5).

Dans le cas $\beta < 0$ on est amené à considérer plusieurs cas comme le montre la figure 5.

Plus particulièrement lorsque $-2 < \beta < -1$ et $1 < \lambda < 2$ on doit envisager les deux cas de figure, A et B tels que:

$$\begin{aligned} \textcircled{A} & \left\{ \begin{array}{ll} \text{si} & -2 < \beta < -1 \quad \text{et} \quad \lambda \leq - (4\beta^2 + 12\beta + 7) \\ \text{si} & -2 < \beta < -\frac{3}{2} \quad \text{et} \quad \lambda > - (4\beta^2 + 12\beta + 7) \end{array} \right. \\ \textcircled{B} & \text{ si} \quad -\frac{3}{2} < \beta < -1 \quad \text{et} \quad \lambda > - (4\beta^2 + 12\beta + 7) \end{aligned}$$

Cette étude montre que la fonction densité de la distribution Log-Pearson III prend des formes très variées, ce qui met en évidence la souplesse de cette loi.

3. PRELIMINAIRES STATISTIQUES

Ce chapitre a pour but de rappeler de manière succincte des résultats importants de la théorie statistique, que l'on appliquera dans les chapitres suivants aux distributions Pearson III et Log-Pearson III.

3.1 Moments et coefficients

On considère une distribution continue, définie par sa fonction densité de probabilité $f(x)$, l'intervalle de définition de la variable aléatoire est D .

Le moment non centré d'ordre r de la distribution, défini pour r entier positif est:

$$\mu'_r = \int_D x^r f(x) dx \quad (3.1)$$

Le moment non centré d'ordre 1, μ'_1 est la moyenne de la distribution, et l'on peut définir les moments d'ordre r (r entier positif) centrés par rapport à la moyenne par:

$$\mu_r = \int_D \left(x - \mu'_1 \right)^r f(x) dx \quad (3.2)$$

Les moments centrés μ_r sont liés aux moments non centrés μ'_r par la relation:

$$\mu_r = \sum_{j=0}^{j=r} C_r^j \mu'_{r-j} \left(-\mu'_1 \right)^j \quad (3.3)$$

Le moment centré μ_1 est toujours nul, le moment centré μ_2 appelé variance, caractérise la dispersion de la distribution autour de la moyenne, (l'écart type est la racine carrée de la variance). Il est possible de définir à partir des moments de la distribution

certaines coefficients caractéristiques tels que:

$$\text{coefficient de variation} \quad C_v = \frac{\mu_2^{\frac{1}{2}}}{\mu_1} \quad (3.4)$$

$$\text{coefficient d'asymétrie} \quad C_s = \frac{\mu_3}{\mu_2^{3/2}} \quad (3.5)$$

$$\text{coefficient de finesse (ou d'aplatissement)} \quad C_k = \frac{\mu_4}{\mu_2^2} \quad (3.6)$$

Le signe de C_s indique si la distribution est à asymétrie positive ($C_s > 0$) ou négative ($C_s < 0$); lorsque $C_s = 0$, la distribution est symétrique par rapport à la moyenne.

Dans le cas de la Loi Normale, on a $C_s = 0$ et $C_k = 3$.

Lorsque l'on considère un échantillon de taille N constitué d'évènements indépendants tirés d'une population, il est également possible de définir les moments de l'échantillon, qui sont des variables aléatoires dont une réalisation donnée est une estimation du moment correspondant de la population.

Si x_i ($i = 1, \dots, N$) sont les éléments de l'échantillon, on définit pour un ordre r entier positif, les moments non centré m'_r et centré par rapport à la moyenne m_r par:

$$m'_r = \frac{1}{N} \sum_{i=1}^N x_i^r \quad (3.7)$$

$$m_r = \frac{1}{N} \sum_{i=1}^N \left(x_i - m'_1 \right)^r \quad (3.8)$$

De la même manière que précédemment, il est possible de définir certains coefficients comme fonction des moments de l'échantillon.

3.2 Fonction caractéristique, cumulants

La fonction caractéristique d'une distribution est définie par:

$$\psi(t) = \int_D e^{itx} f(x) dx \quad (3.9)$$

t est un nombre réel, i est imaginaire ($i^2 = -1$)

Cette fonction est d'une grande importance car elle permet de déterminer les moments et les cumulants d'une distribution. Si $\psi(t)$ peut se développer suivant les puissances de t, le moment non centré d'ordre r est le coefficient de $\frac{(it)^r}{r!}$ du développement et l'on a:

$$\psi(t) = \sum_{r=0} \frac{(it)^r}{r!} \mu'_r \quad (3.10)$$

Connaissant les moments non centrés, il est possible d'en déduire les moments centrés par rapport à la moyenne à l'aide de la relation (3.3).

Pour caractériser une distribution il est possible de déterminer une série de constantes appelées cumulants qui possèdent d'importantes propriétés. Le cumulant d'ordre r, κ_r est défini comme le coefficient de $\frac{(it)^r}{r!}$ dans le développement de $\ln \psi(t)$ si il existe, on a alors:

$$\ln \psi(t) = \sum_{r=1} \kappa_r \frac{(it)^r}{r!} \quad (3.11)$$

A partir des relations (3.10) et (3.11) il est possible d'établir des relations entre les moments non centrés et les cumulants, et en tenant compte de (3.3), entre les moments centrés par rapport à la moyenne, et les cumulants; on a en particulier:

$$\left. \begin{aligned}
\mu_1' &= \kappa_1 \\
\mu_2 &= \kappa_2 \\
\mu_3 &= \kappa_3 \\
\mu_4 &= \kappa_4 + 3\kappa_2^2 \\
\mu_5 &= \kappa_5 + 10 \kappa_3 \kappa_2 \\
\mu_6 &= \kappa_6 + 15 \kappa_4 \kappa_2 + 10 \kappa_3^2 + 15 \kappa_2^3
\end{aligned} \right\} (3.12)$$

La fonction caractéristique permet donc de déduire rapidement les moments et les cumulants de la distribution. Cette fonction possède de plus les propriétés suivantes:

- Si $z = \sum_{j=1}^N x_j$ est la somme de N variables indépendantes

x_j de fonctions caractéristiques respectives $\psi_j(t)$, la fonction caractéristique de z est le produit des $\psi_j(t)$.

Cette propriété résulte directement de la structure de la fonction caractéristique et l'on a:

$$\psi_z(t) = \psi_1(t) * \dots * \psi_j(t) * \dots * \psi_N(t) \quad (3.13)$$

- Si l'on considère la transformation linéaire $y = \ell x + p$, on peut montrer que:

$$\psi_y(t) = e^{ipt} \psi_x(\ell t) \quad (3.14)$$

et, il est possible d'en déduire les relations entre les cumulants:

$$\kappa_{r,y} = \ell^r \kappa_{r,x} \quad r = 2, 3, \dots \quad (3.15)$$

$$\kappa_{1,y} = \ell \kappa_{1,x} + p$$

$\kappa_{r,y}$ et $\kappa_{r,x}$ sont respectivement les cumulants d'ordre r des variables y et x .

Un cas particulier important est celui où l'on considère la distribution standardisée, la variable standardisée t est telle que:

$$t = \frac{x - \kappa_{1,x}}{\sqrt{\kappa_{2,x}}}$$

on a alors:

$$\kappa_{r,t} = \frac{\kappa_{r,x}}{(\kappa_{2,x})^{r/2}} \quad (\text{en particulier } \kappa_{2,t} = 1)$$

$$\kappa_{1,t} = 0$$

3.3 Distribution des moments

Les moments d'un échantillon de taille N , tiré d'une population donnée, sont des variables aléatoires qui suivent des distributions dont certaines caractéristiques peuvent être déterminées (KENDALL 1969):

a) On peut montrer que le moment non centré de l'échantillon, m'_r est distribué avec:

$$\text{une moyenne} \quad E(m'_r) = \mu'_r \quad (3.16)$$

$$\text{une variance} \quad \text{Var}(m'_r) = \frac{1}{N} \left[\mu'_{2r} - \mu'^2_r \right] \quad (3.17)$$

La distribution de m'_r est asymptotiquement normale quand $N \rightarrow \infty$.

La covariance de 2 moments m'_r et m'_q est telle que :

$$\text{Cov}(m'_r, m'_q) = \frac{1}{N} \left[\mu'_{q+r} - \mu'_q \mu'_r \right] \quad (3.18)$$

- b) Si l'on considère les moments centrés par rapport à la moyenne $\left(m'_r \right)$, les résultats sont moins simples car la moyenne de l'échantillon est soumise à des erreurs d'échantillonnage, on peut cependant montrer que m_r est distribué avec :

$$\text{une moyenne } E(m_r) \approx \mu_r \quad (3.19)$$

$$\text{une variance } \text{Var}(m_r) \approx \frac{1}{N} \left[\mu_{2r} - \mu_r^2 + r^2 \mu_2 \mu_{r-1}^2 - 2r \mu_{r-1} \mu_{r+1} \right] \quad (3.20)$$

La distribution de m_r est asymptotiquement normale quand $N \rightarrow \infty$.

D'autre part la covariance de deux moments centrés m_r et m_q est :

$$\text{Cov}(m_r, m_q) \approx \frac{1}{N} \left[\mu_{r+q} - \mu_r \mu_q + r q \mu_2 \mu_{r-1} \mu_{q-1} - r \mu_{r-1} \mu_{q+1} - q \mu_{r+1} \mu_{q-1} \right] \quad (3.21)$$

Dans le cas des moments centrés, ces formules sont des approximations à l'ordre $1/\sqrt{N}$, c'est-à-dire que les termes en $\frac{1}{N}$, $\frac{1}{N^{3/2}}$, ... etc ... du développement de m_r suivant les puissances de r sont négligés.

Il est également possible d'établir au même ordre d'approximation des relations donnant la covariance entre un moment centré m_r et la moyenne m'_1 :

$$\text{Cov}(m'_1, m_r) \approx \frac{1}{N} (\mu_{r+1} - r \mu_2 \mu_{r-1}) \quad (3.22)$$

La relation 3.19 n'est pas exacte, et, l'on dit alors que m_r est une estimation biaisée du moment μ_r de la population. (Une valeur non biaisée \hat{m}_r est telle que $E(\hat{m}_r) = \mu_r$). On peut montrer en particulier que des estimateurs non biaisés de μ_2, μ_3 sont :

$$\hat{m}_2 = \frac{N}{N-1} m_2 \quad (3.23)$$

$$\hat{m}_3 = \frac{N^2}{(N-1)(N-2)} m_3 \quad (3.24)$$

m_2, m_3 sont déterminés par la relation (3.8)

On tient habituellement compte du biais introduit par la faible taille d'un échantillon dans le calcul de C_s en considérant le coefficient corrigé \hat{C}_s tel que:

$$\hat{C}_s = \frac{\hat{m}_3}{(\hat{m}_2)^{3/2}} = \frac{\sqrt{N(N-1)}}{N-2} \bar{C}_s \quad (3.25)$$

avec $\bar{C}_s = \frac{m_3}{m_2^{3/2}}$

Cependant, WALLIS et al (1974) ont étudié pour plusieurs distributions, et en particulier pour la loi Pearson III, la distribution de \bar{C}_s . Pour une taille donnée N ces auteurs génèrent 100,000 échantillons tirés d'une loi Pearson III d'asymétrie théorique C_s fixée.

Les résultats obtenus prouvent, comme KIRBY (1974) l'a démontré que \bar{C}_s (donc \hat{C}_s) est borné par une fonction de N quelle que soit la loi considérée. On a en effet:

$$\left| \bar{C}_s \right| \leq \frac{N-2}{\sqrt{N-1}} \quad \text{ou} \quad \left| \hat{C}_s \right| \leq \sqrt{N} \quad (3.26)$$

A partir des données de WALLIS et al (1974) obtenues pour différentes asymétries théoriques C_s de la loi Pearson III, BOBEE et ROBITAILLE (1975) ont établi une relation donnant C_s à partir de $E(\bar{C}_s)$ moyenne pour un grand nombre d'échantillons de \bar{C}_s .

On a pour $20 \leq N \leq 90$ et $.2 \leq C_s \leq 5$

$$C_s = E(\bar{C}_s) \left[\left(1 + \frac{6.51}{N} + \frac{20.20}{N^2} \right) + \left(\frac{1.48}{N} + \frac{6.77}{N^2} \right) \left(E(\bar{C}_s) \right)^2 \right] \quad (3.27)$$

En pratique lorsque l'on ne dispose que d'un seul échantillon on remplace $E(\bar{C}_s)$ par son estimation \bar{C}_s . La correction donnée par (3.27), valable quel que soit le signe de l'asymétrie de la distribution Pearson III, permet une meilleure estimation que (3.25) du coefficient d'asymétrie.

3.4 Distribution d'une fonction de variables aléatoires

Soit: $g(y_1, \dots, y_k)$ une fonction des variables aléatoires

y_i ($i = 1, \dots, k$) telles que:

$$E(y_i) \approx \theta_i$$

Si l'on désigne par:

θ le point de coordonnées $(\theta_1, \dots, \theta_1, \dots, \theta_k)$

y le point de coordonnées $(y_1, \dots, y_1, \dots, y_k)$

$g(y)$ est distribué avec:

$$\text{une moyenne: } E[g(y)] \approx g(\theta) \quad (3.28)$$

une variance :

$$\begin{aligned} \text{var} [g(y)] &\approx \sum_{i=1}^k \left[\left[\frac{\partial g(y)}{\partial y_i} \right]_{y=\theta} \right]^2 * \text{var } y_i \\ + \sum_{\substack{i=1 \\ i \neq j}}^k \sum_{j=1}^k \left[\frac{\partial g(y)}{\partial y_i} \right]_{y=\theta} \left[\frac{\partial g(y)}{\partial y_j} \right]_{y=\theta} * \text{cov}(y_i, y_j) \end{aligned} \quad (3.29)$$

La distribution de $g(y)$ est asymptotiquement normale quand $N \rightarrow \infty$.

De même la covariance de 2 fonctions de variables aléatoires $g(y)$ et $h(y)$ est:

$$\begin{aligned} \text{Cov} [g(y), h(y)] &\approx \sum_{i=1}^k \left[\frac{\partial g(y)}{\partial y_i} \right]_{y=\theta} \left[\frac{\partial h(y)}{\partial y_i} \right]_{y=\theta} \text{var } x_i \\ + \sum_{\substack{i=1 \\ j \neq i}}^k \sum_{j=1}^k \left[\frac{\partial g(y)}{\partial y_i} \right]_{y=\theta} \left[\frac{\partial h(y)}{\partial y_j} \right]_{y=\theta} * \text{cov}(y_i, y_j) \end{aligned} \quad (3.30)$$

Ces relations sont des approximations du même ordre que celles relatives à la moyenne, et à la variance des variables y_1, \dots, y_k

A partir de ces résultats il est possible de déduire la moyenne et la variance de toute fonction de variables aléatoires et en particulier de toute fonction de moments comme par exemple les cumulants, et, les coefficients de variation et d'asymétrie.

Les différentes formules, concernant les moments centrés et les fonctions de variables aléatoires, sont des approximations qui sont d'autant plus satisfaisantes que la taille de l'échantillon considéré est élevée.

Le caractère asymptotiquement normal des distributions est également fonction de la taille N .

Des simulations par la méthode de Monte Carlo peuvent dans le cas d'une distribution donnée préciser ces différents points.

3.5 Statistiques d'ordre

Considérons un échantillon de N observations x_1, x_2, \dots, x_N , par ordre chronologique, il est possible de classer ces valeurs par ordre croissant et l'on obtient la série ordonnée:

$$y_1 < y_2 \dots < y_N$$

y_1 est la plus petite valeur

y_N est la plus grande valeur

On suppose que l'échantillon considéré est tiré d'une population dont:

- la densité de probabilité est $f(x)$;
- la fonction de distribution cumulée est $F(x)$.

On peut montrer (KENDALL 1969) que la fonction densité de probabilité de l'événement ordonné Y_k est:

$$h(Y_k) = \frac{N!}{(k-1)! (N-k)!} [F(Y_k)]^{k-1} \cdot [1 - F(Y_k)]^{N-k} f(Y_k) \quad (3.31)$$

C'est la densité que l'on obtiendrait si l'on avait une infinité de réalisations y_k de la variable aléatoire Y_k .

La fonction de distribution de la variate Y_k est telle que:

$$H(z, k) = \text{Prob} [Y_k \leq z] \quad (\text{Probabilité au non-dépassement})$$

ou encore

$$dH(z, k) = h(Y_k) dY_k = \frac{[F(Y_k)]^{k-1} [1-F(Y_k)]^{N-k} dF(Y_k)}{B(k, N-k+1)} \quad (3.32)$$

ce qui montre que la distribution de $F(Y_k)$ associé au rang k est une loi Bêta de 1ère espèce (B est la fonction Bêta), on peut en déduire que l'espérance mathématique de $F(Y_k)$ est:

$$E [F(Y_k)] = \frac{k}{N+1} \quad (3.33)$$

Cette valeur est souvent utilisée comme probabilité empirique (plotting position) pour la représentation de l'évènement d'ordre k d'un échantillon de taille N . Cependant BRUNET-MORET (1973 b) a montré que :

- pour un échantillon de taille N tiré d'une population parfaitement connue, il est préférable d'associer à l'évènement d'ordre k , la probabilité empirique au non dépassement :

$$P_k = \frac{k - .3}{N + .4} \quad (\text{formule de Chegodayev}) \quad (3.34)$$

Cette relation est une approximation de la médiane de $F(Y_k)$

- pour un échantillon de taille N , lorsque la population mère n'est pas connue (par exemple lorsque l'estimation des paramètres de la population est effectuée à partir de l'échantillon), la probabilité empirique au non dépassement que l'on peut associer à l'évènement d'ordre k est :

$$P_k = \frac{k - .5}{N} \quad (\text{formule de Hazen}) \quad (3.35)$$

On associe à la probabilité au dépassement $F(x)$ de la valeur x , la période de retour T telle que :

$$\frac{1}{T} = 1 - F(x) = \Pr [X \geq x] \quad (3.36)$$

T est le temps moyen séparant deux évènements supérieurs à x , ce qui ne signifie pas, que l'on n'observe pas en réalité durant un temps T plusieurs évènements supérieurs à x , on montre en effet (REMINIERAS 1967) que la probabilité qu'un évènement x_T soit égalé ou dépassé pendant le temps r est :

$$P = \Pr [X \geq x_T] = 1 - \left(1 - \frac{1}{T}\right)^r \quad (3.37)$$

en particulier pour T grand on montre qu'il y a 63% de chances qu'un évènement supérieur ou égal à x_T se produise pendant la durée T .

D'autre part BENSON (1967) et APPEL (1968) ont montré que la probabilité moyenne de tous les évènements supérieurs ou égaux à x_T est $\frac{1}{2T}$.

4. PROPRIETES DE LA LOI PEARSON III ET DE SES FORMES SIMPLIFIEES

4.1 Introduction

Les différentes formes de la distribution Pearson III suivant le nombre de paramètres et le signe du paramètre d'échelle α ont été mises en évidence dans le chapitre I, les caractéristiques de ces distributions sont récapitulées dans la table 2. Le but de ce chapitre est de développer les propriétés statistiques importantes de la distribution Pearson III et de ses formes simplifiées. Nous insistons plus particulièrement sur les parties originales telles que la détermination de l'erreur-type d'un événement de période de retour donnée, la construction de tables et graphiques d'intervalles de confiance qui sont d'un grand intérêt pratique pour l'utilisation de la loi Pearson III et plus particulièrement pour son application en hydrologie.

4.2 Fonction caractéristique, cumulants et moments

La fonction caractéristique de la distribution Pearson III à 3 paramètres est obtenue à partir de (1.6) et (3.9) et l'on a :

$$\psi(t) = \int_{\Delta} e^{itx} \frac{|\alpha|}{\Gamma(\lambda)} \left[\alpha(x-m) \right]^{\lambda-1} e^{-\alpha(x-m)} dx$$

$$\text{avec: } \Delta \begin{cases} (m, \infty) & \text{si } \alpha > 0 \\ (-\infty, m) & \text{si } \alpha < 0 \end{cases}$$

Distribution	Fonction densité	Domaine de variation	Signe des paramètres	Caractéristique
Pearson III (forme A)	$f(x) = \frac{\alpha}{\Gamma(\lambda)} e^{-\alpha(x-m)} [\alpha(x-m)]^{\lambda-1}$	$m \leq x$	$\alpha > 0$ $\lambda > 0$ m quelconque	Asymétrie positive
Pearson III (forme B)	$f(x) = \frac{-\alpha}{\Gamma(\lambda)} e^{-\alpha(x-m)} [\alpha(x-m)]^{\lambda-1}$	$x \leq m$	$\alpha < 0$ $\lambda > 0$ m quelconque	Asymétrie négative
Gamma (forme A)	$g(y) = \frac{\alpha}{\Gamma(\lambda)} e^{-\alpha y} (\alpha y)^{\lambda-1}$	$0 \leq y$	$\alpha > 0$ $\lambda > 0$ $m = 0$	Asymétrie positive
Gamma (forme B)	$g(y) = \frac{-\alpha}{\Gamma(\lambda)} e^{-\alpha y} (\alpha y)^{\lambda-1}$	$y \leq 0$	$\alpha < 0$ $\lambda > 0$ $m = 0$	Asymétrie négative Variate négative
Pearson III à 1 paramètre	$h(u) = \frac{e^{-u}}{\Gamma(\lambda)} u^{\lambda-1}$	$u > 0$	$\lambda > 0$ $\alpha = 1$ $m = 0$	Asymétrie positive

Table 2 : Caractéristiques de la loi Pearson III et de ses formes simplifiées

en effectuant le changement de variable $v = (\alpha - it)(x-m)$
la relation précédente s'écrit:

$$\psi(t) = \frac{e^{itm}}{\Gamma(\lambda)} \cdot \frac{1}{\left(1 - \frac{it}{\alpha}\right)^\lambda} \int_0^{+\infty} e^{-v} v^{\lambda-1} dv$$

donc

$$\psi(t) = \frac{e^{itm}}{\left(1 - \frac{it}{\alpha}\right)^\lambda} \quad (4.1)$$

Cette relation est valable pour la forme la plus générale de la distribution Pearson III à 3 paramètres quel que soit le signe de α .

Le développement de $\text{Ln } \psi(t)$ permet de déterminer les cumulants de la distribution on a:

$$\text{Ln } \psi(t) = itm - \lambda \text{Ln} \left(1 - \frac{it}{\alpha}\right)$$

ou encore

$$\text{Ln } \psi(t) = \left(m + \frac{\lambda}{\alpha}\right) it + \sum_{r=2} \lambda \left(\frac{it}{\alpha}\right)^r \cdot \frac{1}{r}$$

Par comparaison avec la relation (3.11) on en déduit les cumulants de la distribution Pearson III de variable x

$$\kappa_{r,x} = \lambda \frac{(r-1)!}{\alpha^r} \quad \text{pour } r \geq 2 \quad (4.2)$$

$$\kappa_{1,x} = m + \frac{\lambda}{\alpha}$$

On peut passer de la distribution Pearson III à la loi Gamma par la transformation $y = x-m$ et la relation (3.15) permet de déterminer les cumulants de la distribution Gamma:

$$\begin{aligned} \kappa_{r,y} &= \kappa_{r,x} = \lambda \frac{(r-1)!}{\alpha^r} \\ \kappa_{1,y} &= \kappa_{1,x} - m = \frac{\lambda}{\alpha} \end{aligned} \quad (4.3)$$

De la même manière les cumulants de la distribution à 1 paramètre obtenue par la transformation $u = \alpha y$ sont donnés par:

$$\kappa_{r,u} = \alpha^r \kappa_{r,y} = \lambda(r-1)! \quad (4.4)$$

La table (3) résume les différentes valeurs obtenues pour les cumulants jusqu'à l'ordre 6 (les notations utilisées sont identiques à celles de la table 2.)

La moyenne et les moments centrés par rapport à la moyenne sont calculables à partir des cumulants (relation 3.12), et les valeurs obtenues figurent dans la table 4. Le calcul des moments non centrés peut s'obtenir à partir du développement de $\psi(t)$ mais présente peu d'intérêt pratique.

4.3 Forme standardisée de la distribution Pearson III

Si l'on considère la forme la plus générale de la distribution Pearson III dont la fonction densité $f(x)$ est donnée par (1.5), la forme standardisée est obtenue par la transformation

$$K = \frac{x - \kappa_{1,x}}{\sqrt{\kappa_{2,x}}} = \frac{|\alpha|}{\sqrt{\lambda}} \left[x - \left(m + \frac{\lambda}{\alpha} \right) \right] \quad (4.5)$$

la fonction densité de probabilité $p(K)$ de la variable standardisée est telle que:

$$p(K) = f \left[x(K) \right] \left| \frac{dx}{dK} \right|$$

en exprimant p en fonction des paramètres α, λ, m il vient:

$$p(K) = \frac{e^{-\lambda(\sqrt{\lambda})^\lambda}}{\Gamma(\lambda)} \left[e^{-K \frac{\alpha}{|\alpha|} \sqrt{\lambda}} \left(\sqrt{\lambda} + K \frac{\alpha}{|\alpha|} \right)^{\lambda-1} \right] \quad (4.6)$$

Cumulants Distribution	κ_1	κ_2	κ_3	κ_4	κ_5	κ_6	κ_r
	Pearson III 3 paramètres	$m + \frac{\lambda}{\alpha}$	$\frac{\lambda}{\alpha^2}$	$\frac{2\lambda}{\alpha^3}$	$\frac{6\lambda}{\alpha^4}$	$\frac{24\lambda}{\alpha^5}$	$\frac{120\lambda}{\alpha^6}$
Gamma	$\frac{\lambda}{\alpha}$	$\frac{\lambda}{\alpha^2}$	$\frac{2\lambda}{\alpha^3}$	$\frac{6\lambda}{\alpha^4}$	$\frac{24\lambda}{\alpha^5}$	$\frac{120\lambda}{\alpha^6}$	$\frac{\lambda(r-1)!}{\alpha^r}$
Pearson III 1 paramètre	λ	λ	2λ	6λ	24λ	120λ	$\lambda(r-1)!$
Distribution standardisée $\alpha > 0$ (asymétrie posi- tive)	0	1	$\frac{2}{\sqrt{\lambda}}$	$\frac{6}{\lambda}$	$\frac{24}{\lambda^{3/2}}$	$\frac{120}{\lambda^2}$	$\frac{(r-1)!}{\lambda^{r/2-1}}$
Distribution standardisée $\alpha < 0$ (asymétrie néga- tive)	0	1	$-\frac{2}{\sqrt{\lambda}}$	$\frac{6}{\lambda}$	$-\frac{24}{\lambda^{3/2}}$	$\frac{120}{\lambda^2}$	$(-1)^r \frac{(r-1)!}{\lambda^{r/2-1}}$

Table 3: Cumulants de la loi Pearson III et de ses formes simplifiées

Moments Distribution	μ'_1 (moyenne)	μ_2 variance	μ_3	μ_4	μ_5	μ_6
Pearson III 3 paramètres	$m + \frac{\lambda}{\alpha}$	$\frac{\lambda}{\alpha^2}$	$\frac{2\lambda}{\alpha^3}$	$\frac{3\lambda}{\alpha^4} (\lambda+2)$	$\frac{4\lambda}{\alpha^5} (5\lambda + 6)$	$\frac{5\lambda}{\alpha^6} (3\lambda^2+26\lambda+24)$
Gamma	$\frac{\lambda}{\alpha}$	$\frac{\lambda}{\alpha^2}$	$\frac{2\lambda}{\alpha^3}$	$\frac{3\lambda}{\alpha^4} (\lambda+2)$	$\frac{4\lambda}{\alpha^5} (5\lambda + 6)$	$\frac{5\lambda}{\alpha^6} (3\lambda^2+26\lambda+24)$
Pearson III 1 paramètre	λ	λ	2λ	$3\lambda (\lambda+2)$	$4\lambda (5\lambda + 6)$	$5\lambda (3\lambda^2+26\lambda+24)$
Distribution standardisée $\alpha > 0$	0	1	$\frac{2}{\sqrt{\lambda}}$	$3 + \frac{6}{\lambda}$	$\frac{4}{\sqrt{\lambda}} (5 + \frac{6}{\lambda})$	$\frac{5}{\lambda^2} (24+26\lambda+3\lambda^2)$
Distribution standardisée $\alpha < 0$	0	1	$-\frac{2}{\sqrt{\lambda}}$	$3 + \frac{6}{\lambda}$	$-\frac{4}{\sqrt{\lambda}} (5 + \frac{6}{\lambda})$	$\frac{5}{\lambda^2} (24+26\lambda+3\lambda^2)$

Table 4: Moments de la distribution Pearson III et de ses formes simplifiées

Cette fonction ne dépend en fait que de λ puisque l'on a $\frac{\alpha}{|\alpha|} = \pm 1$ suivant le signe de α .

Si α est positif, l'intervalle de variation est:

$$-\sqrt{\lambda} < K < +\infty$$

Si α est négatif, l'intervalle de variation est:

$$-\infty < K < +\sqrt{\lambda}$$

La fonction caractéristique $\psi_K(t)$ de la distribution standardisée peut être obtenue à partir de celle de la variable x en appliquant la relation (3.14) à la transformation linéaire (4.5) il vient:

$$\psi_K(t) = \frac{e^{-i \frac{|\alpha|}{\alpha} \sqrt{\lambda} t}}{\left(1 - i \frac{|\alpha|}{\alpha} \frac{t}{\sqrt{\lambda}}\right)^\lambda} \quad (4.7)$$

Le développement de $\ln \psi_K(t)$ permet ensuite la détermination des cumulants de la distribution standardisée

Si $\alpha > 0$

$$\ln \psi_K(t) = it \left(-\sqrt{\lambda} + \sqrt{\lambda}\right) + \frac{1}{2} (it)^2 + \dots + \frac{\lambda(it)^r}{r(\sqrt{\lambda})^r} + \dots$$

d'où:

$$\kappa_1 = 0$$

$$\kappa_r = \lambda^{1-r/2} (r-1)! \quad (4.8)$$

Si $\alpha < 0$

$$\ln \psi_K(t) = it \left(\sqrt{\lambda} - \sqrt{\lambda}\right) + \frac{(it)^2}{2} + \dots + (-1)^r \frac{\lambda(it)^r}{r(\sqrt{\lambda})^r} + \dots$$

d'où :

$$\kappa_1 = 0$$

(4.9)

$$\kappa_r = (-1)^r \lambda^{1-r/2} (r-1)!$$

A partir des cumulants il est possible de déduire la moyenne et les moments centrés par rapport à la moyenne à l'aide de (3.12). On vérifie en particulier que la moyenne est nulle, les moments centrés et non centrés sont donc identiques, la variance est égale à l'unité, tandis que les autres moments et cumulants ne dépendent que du paramètre λ , leurs valeurs jusqu'à l'ordre 6 figurent dans les tables 3 et 4.

Les tables donnant la valeur de la variable Pearson III standardisée en fonction de l'asymétrie, donc en fonction de λ , pour différentes probabilités au non-dépassement ont été calculées par HARTER (1969).

Comportement asymptotique de la distribution Pearson III

L'examen des moments et cumulants de la distribution Pearson III standardisée montre que lorsque $\lambda \rightarrow \infty$ les valeurs limites obtenues correspondent aux moments et cumulants de la loi Normale standardisée. Ce résultat peut être étendu en montrant que la fonction densité de la distribution Pearson III standardisée a pour limite la loi normale standardisée lorsque $\lambda \rightarrow \infty$.

D'après (4.6) on a :

$$\ln p(K) = -\lambda + \frac{\lambda}{2} \ln \lambda - \ln \Gamma(\lambda) - K \frac{\alpha}{|\alpha|} \sqrt{\lambda} + (\lambda-1) \ln \left(\sqrt{\lambda+K} \frac{\alpha}{|\alpha|} \right)$$

D'autre part il est possible de remplacer $\text{Ln } \Gamma(\lambda)$ par l'approximation de STIRLING:

$$\text{Ln } \Gamma(\lambda) = (\lambda - \frac{1}{2}) \text{Ln } \lambda - \lambda + \frac{1}{2} \text{Ln } 2\pi + \frac{1}{12\lambda} + \varepsilon \left(\frac{1}{\lambda^2} \right) \quad (4.10)$$

où $\varepsilon \left(\frac{1}{\lambda^2} \right)$ est un terme qui tend vers 0 quand λ devient grand.

On peut montrer que:

$$\text{Ln } p(K) = -\frac{1}{2} \text{Ln } 2\pi - \frac{1}{2} K^2 \left(\frac{\alpha}{|\alpha|} \right)^2 + \varepsilon_1$$

où ε_1 est une fonction de λ qui tend vers 0 quand λ est grand, on a alors:

$$\lim_{\lambda \rightarrow \infty} p(K) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}K^2} \quad (4.11)$$

Donc lorsque λ devient grand, la distribution Pearson III standardisée est asymptotiquement normale quel que soit le signe de α , ce qui signifie en pratique que:

- La distribution Pearson III de faible asymétrie (positive ou négative) est voisine de la loi Normale.
- La distribution Normale peut être considérée comme un cas particulier de la loi Pearson III, correspondant à λ infini (ou à asymétrie nulle).

4.4 Coefficients caractéristiques de la distribution Pearson III

Le coefficient de variation $\left(C_v \right)$ défini par la relation (3.4) dépend de la forme de la distribution Pearson III que l'on considère, dans le cas de la loi à trois paramètres on a:

$$C_v = \frac{\alpha}{|\alpha|} \frac{\sqrt{\lambda}}{\lambda + m\alpha} \quad (4.12)$$

ou encore :

$$\text{si } \alpha > 0 \quad C_v = \frac{\sqrt{\lambda}}{\lambda + m\alpha}$$

$$\text{si } \alpha < 0 \quad C_v = -\frac{\sqrt{\lambda}}{\lambda + m\alpha}$$

Dans le cas de la loi Gamma la valeur de C_v est obtenue à partir de la relation (4.12) dans laquelle $m = 0$ d'où :

$$C_v = \frac{\alpha}{|\alpha|} \cdot \frac{1}{\sqrt{\lambda}} \quad (4.13)$$

Le coefficient d'asymétrie (C_s) défini par (3.5) est tel que :

$$C_s = \frac{\alpha}{|\alpha|} \cdot \frac{2}{\sqrt{\lambda}} \quad (4.14)$$

Le coefficient de finesse (C_k) est indépendant du signe de α et l'on a :

$$C_k = 3 + \frac{6}{\lambda} \quad (4.15)$$

Par analogie, il est possible de définir des coefficients C_1 et C_2 faisant intervenir les moments μ_5 et μ_6 dont l'expression pour la loi Pearson III est :

$$C_1 = \frac{\mu_5}{\mu_2^{5/2}} = \frac{\alpha}{|\alpha|} \frac{4}{\sqrt{\lambda}} \left(5 + \frac{6}{\lambda} \right) \quad (4.16)$$

$$C_2 = \frac{\mu_6}{\mu_2^3} = 5 \left(3 + \frac{26}{\lambda} + \frac{24}{\lambda^2} \right) \quad (4.17)$$

Ces différents coefficients (sauf C_v) sont indépendants du nombre de paramètres considérés, et prennent donc au signe près les mêmes valeurs pour les différentes formes de la distribution Pearson III. Le signe de C_s et C_1 est celui de α , tandis que C_k et C_2 sont indépendants du signe de α .

Ces coefficients peuvent se déduire directement des moments de la forme standardisée pour laquelle on a $\mu_2 = 1$, donc $C_s = \mu_3$ etc...

Dans le cas général de la distribution Pearson III à 3 paramètres, il est possible de montrer que C_k, C_1, C_2 s'expriment en fonction de C_s par:

$$C_k = 3 \left(1 + \frac{1}{2} C_s^2 \right) \quad (4.18)$$

$$C_1 = C_s \left(10 + 3C_s^2 \right) \quad (4.19)$$

$$C_2 = 5 \left(3 + \frac{13}{2} C_s^2 + \frac{3}{2} C_s^4 \right) \quad (4.20)$$

Dans le cas particulier de la distribution Gamma et de la distribution à un seul paramètre on a de plus:

$$C_s = 2 C_v \quad (4.21)$$

et C_s, C_k, C_1, C_2 peuvent s'exprimer en fonction de C_v .

4.5 Distribution de la moyenne d'un échantillon tiré d'une loi Pearson III

Soient $x_1, x_2, \dots, x_j, \dots, x_N$, les N éléments indépendants d'un échantillon tiré d'une distribution Pearson III. La fonction caractéristique de x_j est d'après (4.1):

$$\psi_j(t) = \frac{e^{itm}}{\left(1 - \frac{it}{\alpha} \right)^\lambda}$$

La fonction caractéristique de $Z = \sum_{j=1}^N x_j$ est donnée par (3.13)

$$\psi_Z(t) = \psi_1(t) * \dots * \psi_j(t) * \dots * \psi_N(t)$$

A l'aide de la relation (3.14) il est possible d'en déduire la fonction caractéristique de la moyenne m'_1 de l'échantillon

$$\psi_{m'_1}(t) = \psi_Z\left(\frac{t}{N}\right) = \frac{e^{itm}}{\left(1 - \frac{it}{N\alpha}\right)^{\lambda N}} \quad (4.22)$$

La comparaison des relations (4.1) et (4.22) montre que la moyenne d'un échantillon de taille N , tiré d'une loi Pearson III de paramètres m, α, λ est distribuée suivant une loi Pearson III de paramètres $m, N\alpha; N\lambda$, la fonction densité de m'_1 est donc:

$$f(m'_1) = \frac{|N\alpha|}{\Gamma(N\lambda)} e^{-N\alpha(x-m)} \left[N\alpha(x-m) \right]^{N\lambda-1} \quad (4.23)$$

La distribution de m'_1 est caractérisée par:

$$\text{une moyenne } E(m'_1) = m + \frac{\lambda}{\alpha} = \mu'_1 \quad (4.24)$$

$$\text{une variance } \text{var}(m'_1) = \frac{\lambda}{N\alpha^2} = \frac{1}{N} \mu_2 \quad (4.25)$$

$$\text{un coefficient d'asymétrie } C_s(m'_1) = \frac{\alpha}{|\alpha|} \frac{2}{\sqrt{N\lambda}} = \frac{1}{\sqrt{N}} C_s \quad (4.26)$$

où μ'_1, μ_2, C_s sont relatifs à la population-mère.

Les relations (4.24) et (4.25) correspondent à (3.16) et (3.17) avec la propriété supplémentaire que m'_1 est distribuée exactement suivant une loi Pearson III, dont le coefficient d'asymétrie est de même signe que celui de la population-mère d'après la relation (4.26) Si $N\lambda$ est grand, m'_1 suit une distribution approximativement Normale, mais pour les échantillons de faible taille ou lorsque la population-mère a une forte asymétrie, la distribution de m'_1 ne peut être assimilée à une loi Normale.

4.6 Erreur-type des moments et coefficients d'un échantillon

On considère un échantillon de taille N constitué d'observations indépendantes (autocorrélation nulle) tirées d'une population donnée; les moments et coefficients de cet échantillon sont des estimations des moments et coefficients correspondants de la population. Si l'on considère un grand nombre d'échantillons de taille N , un moment m'_r par exemple sera distribué autour de la valeur μ'_r de la population avec une variance (ou un écart-type) que l'on peut évaluer. La loi suivie par la variate m'_r est en général difficile sinon impossible à déterminer (dans le cas de m'_1 nous avons vu que cette détermination est possible), on peut alors admettre en première approximation que m'_r suit une distribution normale bien que cela ne soit rigoureux que pour N très grand. L'écart-type de cette distribution est appelé aussi l'erreur-type (E.T.)

Le calcul de l'erreur-type d'un moment ou d'un coefficient d'un échantillon de taille donnée a une grande importance pratique car il permet:

- la détermination de l'intervalle de confiance de la valeur correspondante de la population (sous l'hypothèse de normalité)
- l'évaluation de l'erreur d'échantillonnage associée à ce moment ou à ce coefficient;
- la détermination de la taille de l'échantillon pour atteindre une précision donnée sur ce moment ou ce coefficient.

HARDISON (1969) a effectué le calcul des erreurs-types de moments et de coefficients dans le cas d'un échantillon tiré d'une population Normale. Le cas de la loi Gamma a également été considéré (BOBEE et VILLENEUVE, 1973). En hydrologie cependant, la loi Pearson III rend en général bien compte de séries observées d'asymétrie non nulle, ce qui justifie l'étude des erreurs-types des principaux moments et coefficients d'un échantillon tiré de cette distribution.

Dans ce qui suit on note $\bar{C}_v, \bar{C}_s, \bar{C}_k, \bar{C}_1, \bar{C}_2$ les valeurs de l'échantillon pour les distinguer des valeurs correspondantes de la population C_v, C_s, C_k, C_1, C_2 .

a. Erreur-type sur la variance, (E.T.)_{m₂}

D'après (3.20) on a:

$$\text{var } m_2 \approx \frac{1}{N} \left(\mu_4 - \mu_2^2 \right) = \frac{\mu_2^2}{N} \left(C_k - 1 \right)$$

En tenant compte de la relation (4.18) l'erreur type de la variance de l'échantillon m_2 est dans le cas de la forme générale de la loi Pearson III.

$$(\text{E.T.})_{m_2} \approx \mu_2 \sqrt{\frac{2}{N} \left(1 + \frac{3}{4} C_s^2 \right)}$$

en pratique les valeurs de la population μ_2 et C_s sont inconnues et peuvent être estimées par les valeurs correspondantes de l'échantillon; on a alors

$$(\text{E.T.})_{m_2} \approx m_2 \sqrt{\frac{2}{N} \left(1 + \frac{3}{4} \bar{C}_s^2 \right)} \quad (4.27)$$

Dans le cas de la loi Gamma ou de la loi Pearson à 1 paramètre en remplaçant C_s en fonction de C_v d'après (4.21) on a:

$$(\text{E.T.})_{m_2} \approx m_2 \sqrt{\frac{2}{N} \left(1 + 3 \bar{C}_v^2 \right)} \quad (4.28)$$

b. Erreur-type sur l'écart-type

L'écart type de l'échantillon est $s = \sqrt{m_2}$ qui correspond à l'écart type σ de la population. En appliquant la relation (3.29) à la fonction $g(m_2) = \sqrt{m_2}$ on a:

$$\text{var } s = \left[\left(\frac{\partial \sqrt{m_2}}{\partial m_2} \right)_{m_2 = \mu_2} \right]^2 * \text{var } m_2$$

En remplaçant $\text{var } m_2$ par son expression déduite de (3.20)

il vient:

$$\text{var } s = \frac{1}{4N} \left(\frac{\mu_4 - \mu_2^2}{\mu_2} \right) = \frac{\mu_2}{4N} (C_k - 1)$$

Dans le cas de la loi Pearson III à 3 paramètres on peut remplacer C_k en fonction de C_s d'où :

$$(\text{E.T.})_s = \frac{\sigma}{\sqrt{2N}} \sqrt{1 + \frac{3}{4} C_s^2} \quad (4.29)$$

Dans le cas de la loi Gamma on a de plus $C_s = 2C_v$ et la relation précédente devient

$$(\text{E.T.})_s = \frac{\sigma}{\sqrt{2N}} \sqrt{1 + 3 C_v^2} \quad (4.30)$$

En pratique les valeurs de la population σ, C_s, C_v sont remplacées par leurs estimations s, \bar{C}_s, \bar{C}_v .

c. Erreur-type sur le coefficient de variation

Le coefficient de variation de l'échantillon est défini par

$$\bar{C}_v = \frac{\sqrt{m_2}}{m'_1}, \text{ l'application de (3.29) conduit à:}$$

$$\begin{aligned} \text{var } \bar{C}_v &= \left[\begin{array}{c} \left[\frac{1}{2\sqrt{m_2} m'_1} \right] \\ m_2 = \mu_2 \\ m'_1 = \mu'_1 \end{array} \right]^2 * \text{var } m_2 + \left[\begin{array}{c} \left[-\frac{\sqrt{m_2}}{m_1'^2} \right] \\ m_2 = \mu_2 \\ m'_1 = \mu'_1 \end{array} \right]^2 * \text{var } m'_1 \\ &\quad - 2 \left[\begin{array}{c} \left[\frac{1}{2\sqrt{m_2} m'_1} \right] \\ m_2 = \mu_2 \\ m'_1 = \mu'_1 \end{array} \right] * \left[\begin{array}{c} \left[\frac{\sqrt{m_2}}{m_1'^2} \right] \\ m_2 = \mu_2 \\ m'_1 = \mu'_1 \end{array} \right] \text{Cov } (m'_1, m_2) \end{aligned}$$

Ce qui s'écrit:

$$\text{var} (\bar{C}_v) = \frac{1}{(2\mu_1' \sqrt{\mu_2})^2} \text{var } m_2 + \frac{\mu_2}{\mu_1'^4} \text{var } m_1' - \frac{1}{\mu_1'^3} \text{cov} (m_1', m_2)$$

avec

$$\text{var } m_2 \approx \frac{1}{N} (\mu_4 - \mu_2^2) \quad (\text{d'après 3.20})$$

$$\text{var } m_1' \approx \frac{\mu_2}{N} \quad (\text{d'après 3.17})$$

$$\text{cov} (m_2, m_1') \approx \frac{\mu_3}{N} \quad (\text{d'après 3.22})$$

on obtient alors:

$$\text{var} (\bar{C}_v) = \frac{C_v^2}{N} \left[\frac{\mu_4 - \mu_2^2}{4\mu_2^2} + \frac{\mu_2}{\mu_1'^2} - \frac{\mu_3}{\mu_1' \mu_2} \right] \quad (4.31)$$

Dans le cas de la loi Pearson III, la relation (4.31) devient:

$$(\text{E.T.}) \bar{C}_v = \frac{C_v}{\sqrt{2N}} \sqrt{\left(1 + 2C_v^2\right) + C_s^2 \left(\frac{3}{4} - \frac{2C_v}{C_s}\right)} \quad (4.32)$$

Dans le cas de la distribution Gamma, puisque $C_s = 2C_v$ la relation précédente peut être simplifiée:

$$(\text{E.T.}) \bar{C}_v = \frac{C_v}{\sqrt{2N}} \sqrt{1 + C_v^2} \quad (4.33)$$

d. Erreur type sur le coefficient d'asymétrie

Le coefficient d'asymétrie de l'échantillon est $\bar{C}_s = \frac{m_3}{m_2^{3/2}}$,

en utilisant la relation (3.29) on peut montrer que l'on a:

$$\text{var } \bar{C}_s \approx \frac{1}{\mu_2^3} \text{var } m_3 + \frac{9}{4} \frac{\mu_3^2}{\mu_2^5} \text{var } m_2 - \frac{3\mu_3}{\mu_2^4} \text{cov} (m_2, m_3)$$

or:

$$\text{var } (m_3) = \frac{1}{N} \left[\mu_6 - \mu_3^2 + 9\mu_2^3 - 6\mu_2\mu_4 \right] \quad (\text{d'après 3.20})$$

$$\text{var } (m_2) \approx \frac{1}{N} \left[\mu_4 - \mu_2^2 \right]$$

$$\text{cov } (m_2, m_3) \approx \frac{1}{N} \left[\mu_5 - 4\mu_2\mu_3 \right] \quad (\text{d'après 3.21})$$

en remplaçant il vient:

$$\text{var } (\bar{C}_s) = \frac{1}{N\mu_2^3} \left[\left(\mu_6 - \mu_3^2 + 9\mu_2^3 - 6\mu_2\mu_4 \right) + \frac{9}{4} \frac{\mu_3^2}{\mu_2} \left(\mu_4 - \mu_2^2 \right) - \frac{3\mu_3}{\mu_2} \left(\mu_5 - 4\mu_2\mu_3 \right) \right]$$

En remplaçant dans la relation précédente les moments par les différents coefficients on a:

$$\text{var } (\bar{C}_s) = \frac{1}{N} \left[\left(C_2 - C_s^2 + 9 - 6C_k \right) + \frac{9}{4} C_s^2 (C_k - 1) - 3(C_s C_1 - 4C_s^2) \right]$$

et puisque dans le cas de la loi Pearson III les coefficients C_k , C_1 , C_2 peuvent s'exprimer en fonction de C_s par les relations (4.18), (4.19) et (4.20) on a:

$$(\text{E.T.})_{\bar{C}_s} = \sqrt{\frac{6}{N} \left(1 + \frac{3}{2} C_s^2 + \frac{5}{16} C_s^4 \right)} \quad (4.34)$$

Dans le cas de la loi Gamma on peut exprimer l'erreur type en fonction de C_v :

$$(\text{E.T.})_{\bar{C}_s} = \sqrt{\frac{6}{N} \left(1 + 6 C_v^2 + 5 C_v^4 \right)} \quad (4.35)$$

Les différentes formules donnant les écarts-types de moments centrés ou de coefficients d'un échantillon ont été établies à partir d'approximations dont la validité dépend de la taille de l'échantillon. De manière générale pour une taille d'échantillon donnée, les formules relatives à des moments d'ordre peu élevé sont plus précises.

Dans le cas du coefficient d'asymétrie les formules 4.34 et 4.35 ont été établies en faisant intervenir des moments centrés jusqu'à l'ordre 6, et leur validité pour des petits échantillons est douteuse, MATALAS et BENSON (1968) conseillent de considérer l'erreur-type du coefficient d'asymétrie lorsque $N > 100$. Pour les échantillons de faible taille on peut cependant tester si \bar{C}_s est significativement différent de 0, en effet lorsque $C_s = 0$ la distribution Pearson III tend vers la loi Normale et FISHER (1931) a montré que pour des échantillons tirés de la loi Normale on a :

$$(E.T.)_{\bar{C}_s} = \sqrt{\frac{6N(N-1)}{(N-2)(N+1)(N+3)}} \quad (4.36)$$

Lorsque N est grand $(E.T.)_{\bar{C}_s} \rightarrow \frac{6}{N}$, qui est la limite obtenue dans (4.34) lorsque $C_s \rightarrow 0$.

De manière générale les formules établies précédemment pour la distribution Pearson III sont aussi valables pour la loi Normale, il suffit de faire $C_s = 0$ dans ces formules; la table 5 récapitule les résultats obtenus pour la loi Pearson III, la loi Gamma et la loi Normale. Les figures 6 et 7 donnent les erreurs-types relatives pour la moyenne, la variance, l'écart-type et le coefficient de variation respectivement dans le cas de la loi Normale et de la loi Gamma.

Les formules établies précédemment permettent donc de déterminer l'erreur-type de certaines caractéristiques d'un échantillon tiré d'une population qui suit la distribution Pearson III où l'une de ses formes simplifiées. Ces formules cependant ne sont que des approximations, d'autant plus valables que la taille de l'échantillon est élevée. Dans le cas d'échantillons montrant une asymétrie et pouvant provenir d'une distribution Pearson III, il est donc important d'utiliser les relations précédentes et non comme cela est souvent fait des formules d'erreur-type relatives à la loi Normale qui conduisent à une sous-estimation.

Distribution / Erreur-type	Loi Pearson III	Gamma Pearson III à 1 paramètre	Normale
Moyenne (E.T.) _{m₁}	$\frac{\sigma}{\sqrt{N}}$	$\frac{\sigma}{\sqrt{N}}$	$\frac{\sigma}{\sqrt{N}}$
Variance (E.T.) _{m₂}	$\frac{\mu_2}{\sqrt{N}} \sqrt{2 \left(1 + \frac{3}{4} C_s^2 \right)}$	$\frac{\mu_2}{\sqrt{N}} \sqrt{2 \left(1 + 3 C_v^2 \right)}$	$\mu_2 \sqrt{\frac{2}{N}}$
Ecart-type (E.T.) _s	$\frac{\sigma}{\sqrt{2N}} \sqrt{1 + \frac{3}{4} C_s^2}$	$\frac{\sigma}{\sqrt{2N}} \sqrt{1 + 3 C_v^2}$	$\frac{\sigma}{\sqrt{2N}}$
Coefficient de variation (E.T.) _{C_v}	$\frac{C_v}{\sqrt{2N}} \sqrt{\left(1 + 2C_v^2 \right) + C_s^2 \left(\frac{3}{4} - 2\frac{C_v}{C_s} \right)}$	$\frac{C_v}{\sqrt{2N}} \sqrt{1 + C_v^2}$	$\frac{C_v}{\sqrt{2N}} \sqrt{1 + 2 C_v^2}$
Coefficient d'asymétrie (E.T.) _{C_s}	$\sqrt{\frac{6}{N} \left(1 + \frac{3}{2} C_s^2 + \frac{5}{16} C_s^4 \right)}$	$\sqrt{\frac{6}{N} \left(1 + 6C_v^2 + 5 C_v^4 \right)}$	$\sqrt{\frac{6}{N}}$

Table 5: Erreurs-types pour les différentes formes de la loi Pearson III et la loi Normale

$e\sqrt{N} \times 10^2$

LOI NORMALE

$$e = \frac{(E.T) \text{ Variate}}{\text{Variate}}$$

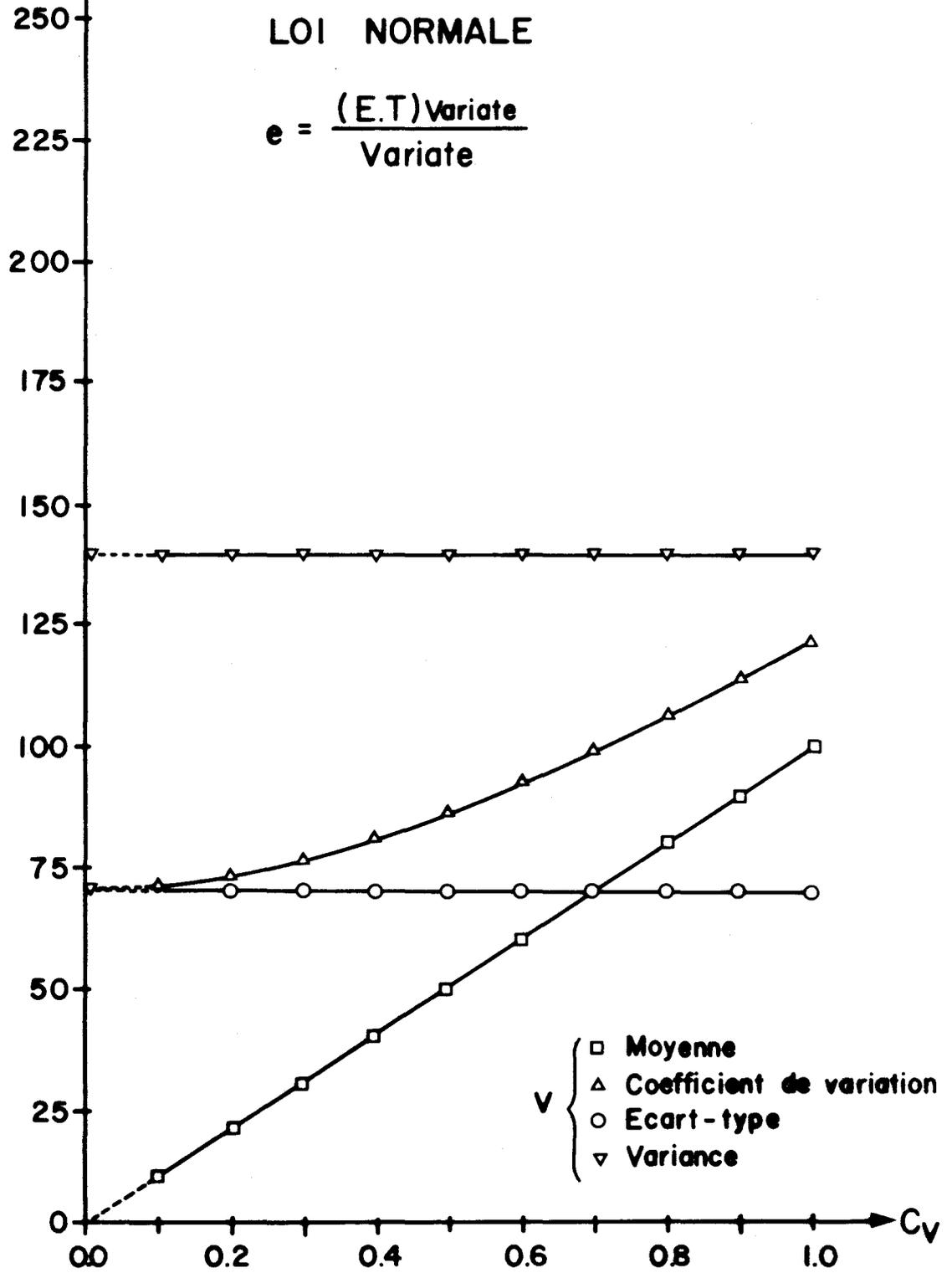


Fig. 6

Erreur-type

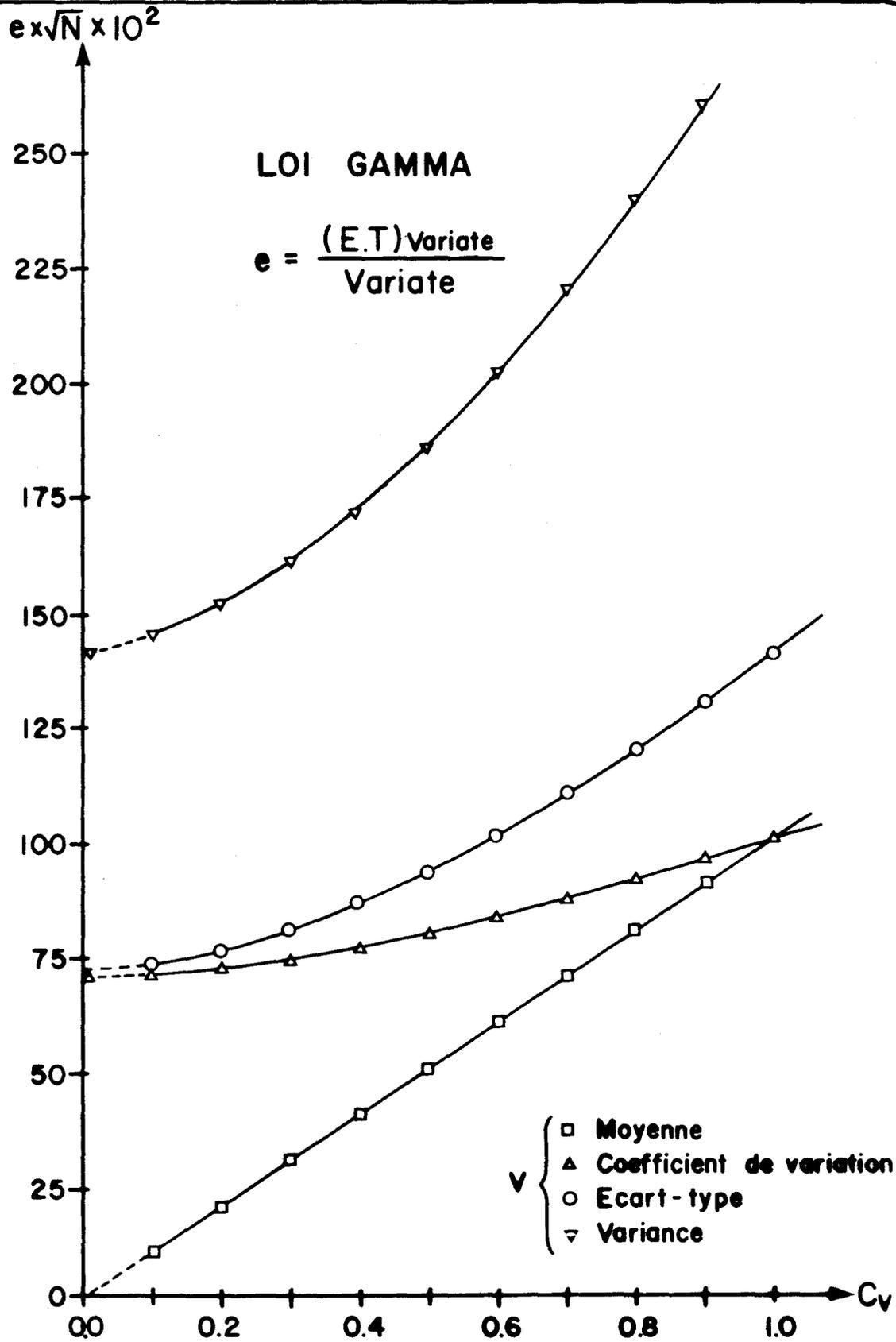


Fig. 7 Erreur-type

4.7 Erreur-type d'un évènement de période de retour donnée

L'un des problèmes les plus importants de l'hydrologie concerne l'estimation du débit de crue correspondant à une période de retour donnée T , c'est-à-dire à une probabilité au dépassement $p = \frac{1}{T}$. Ce problème se pose notamment avec acuité dans les projets de contrôle de crue, d'aménagements hydroélectriques où le dimensionnement des ouvrages, le calcul des risques encourus sont conditionnés par la grandeur de la crue de période de retour T ($T = 1000$ ans par exemple) et par l'erreur commise sur son estimation.

Cette estimation est soumise à deux types d'erreur:

- erreur due à l'inadéquation de la distribution statistique choisie pour représenter la population; cette erreur est difficile à évaluer car en pratique plusieurs distributions sont susceptibles de représenter un phénomène hydrologique donné et aucun test statistique ne permet de déterminer la meilleure loi.
- erreur d'échantillonnage dans la détermination des paramètres de la distribution choisie, car l'on dispose seulement d'un échantillon de taille N , ce qui se traduit par une erreur-type $(E.T.)_{Y_T}$ sur l'évènement estimé Y_T de période de retour T .

NASH et AMOROCHO (1966) ont étudié l'erreur-type d'un évènement extrême de période de retour T lorsque l'échantillon est tiré d'une loi Normale ou d'une loi FISHER-TIPPET (double exponentielle). Ici nous considérons le cas de la distribution Pearson III qui permet souvent de rendre compte adéquatement des évènements hydrologiques extrêmes. On considère donc un échantillon de taille N dont les éléments sont des évènements extrêmes (par exemple la crue maximum annuelle) tirés d'une distribution Pearson type III, l'évènement de période de retour T de la population est Y_T tel que :

$$Y_T = \mu_1' + \chi \sqrt{\mu_2} \quad (4.37)$$

μ_1' et μ_2 étant respectivement la moyenne et la variance de la population, χ est la variable standardisée qui est fonction de la période de retour T et du coefficient d'asymétrie C_s de la population.

La valeur estimée Y_T est donnée par:

$$Y_T = m_1' + K\sqrt{m_2} \quad (4.38)$$

ou

$$Y_T = m_1' \left(1 + K \frac{\bar{C}_v}{v} \right) \quad (4.39)$$

K , variable standardisée, dépend de la période de retour T et de la valeur estimée à partir de l'échantillon du coefficient d'asymétrie \bar{C}_s .

L'estimation Y_T est une variable aléatoire distribuée approximativement selon une loi Normale de moyenne y_T et d'écart type $(E.T.)_{Y_T}$; cette tendance vers la normalité qui dépend fortement de la taille de l'échantillon est due au fait que Y_T est fonction des moments estimés, donc de sommes de variates soumises à l'effet du théorème Central Limite.

La détermination de σ_{Y_T} , dans le cas où l'on suppose connu à priori le coefficient d'asymétrie C_s de la population, ce qui revient dans la relation (4.38) à négliger la variabilité de K , a été effectuée par HARDISON (1969) qui obtient:

$$\left[(E.T.)_{Y_T} \right]^2 = \frac{\mu_2}{N} \left[1 + \frac{b^2}{2} \chi^2 + \sqrt{2} \rho b \chi \right] \quad (4.40)$$

avec: $b^2 = \left(1 + \frac{3}{4} C_s^2 \right)$

Dans la relation précédente ρ est le coefficient de corrélation entre la moyenne (m_1') et l'écart type (s) de l'échantillon, il est cependant possible de montrer BOBEE (1973) que dans le cas de la distribution Pearson III on a :

$$\rho = \frac{C_s}{\sqrt{2 \left(1 + \frac{3}{4} C_s^2 \right)}}$$

La relation (4.40) peut alors se mettre sous la forme

$$\left[(E.T.)_{Y_T} \right]^2 = \frac{\mu_2}{N} \left[1 + \frac{\chi^2}{2} \left(1 + \frac{3}{4} C_s^2 \right) + \chi C_s \right] \quad (4.41)$$

Ces différentes formules cependant sont peu réalistes puisqu'elles sont établies en supposant C_s connu, alors qu'en pratique C_s est inconnu ; on ne connaît que \bar{C}_s , coefficient d'asymétrie de l'échantillon qui est soumis à de fortes erreurs.

Dans ce qui suit nous déterminons $(E.T.)_{Y_T}$ en tenant compte de la variabilité de K due à l'erreur d'échantillonnage commise sur C_s .

En ce qui concerne les différentes formes de la loi Pearson III on doit envisager deux cas suivant que l'on admet ou non la relation

$$\bar{C}_s = 2 \bar{C}_v$$

a) Calcul de $(E.T.)_{Y_T}$, quand $\bar{C}_s = 2 \bar{C}_v$

Dans le cas de la loi Pearson III à deux paramètres (Gamma) ou à un seul paramètre on a la relation (4.21) : $C_s = 2 C_v$ entre les coefficients de la population. Cette relation sera encore vraie pour les valeurs estimées si l'on utilise la méthode du maximum de vraisemblance (cf. Chap. 7) pour estimer

les paramètres de la distribution (la méthode des moments conduit en général à $\bar{C}_s \neq 2 \bar{C}_v$).

Lorsque l'on a $\bar{C}_s = 2 \bar{C}_v$, Y_T donné par la relation (4.39) est alors fonction de m'_1 et \bar{C}_v , on a donc d'après (3.29):

$$\left[(E.T.)_{Y_T} \right]^2 = \left(\frac{\partial Y_T}{\partial m'_1} \right)_0^2 \text{ var } m'_1 + \left(\frac{\partial Y_T}{\partial C_v} \right)_0^2 \text{ var } C_v + 2 \left(\frac{\partial Y_T}{\partial m'_1} \right)_0 \left(\frac{\partial Y_T}{\partial C_v} \right)_0 \text{ Cov } (m'_1, C_v)$$

Les dérivées partielles sont calculées pour $m'_1 = \mu'_1$; $m_2 = \mu_2$; $\bar{C}_v = C_v$; $\bar{C}_s = C_s$; ce qui est symbolisé par l'indice 0; on peut alors montrer (BOBEE 1973 a) que l'on obtient:

$$\left[(E.T.)_{Y_T} \right]^2 = \frac{\mu_2}{N} \left\{ \left(1 + \chi C_v \right)^2 + \frac{1}{2} \left[\chi + 2 C_v \left(\frac{\partial K}{\partial C_s} \right)_0 \right]^2 \left(1 + C_v^2 \right) \right\} \quad (4.42)$$

Si l'on néglige la variabilité de K, ce qui revient à faire

$$\left(\frac{\partial K}{\partial C_s} \right)_0 = 0 \quad \text{dans (4.42), on obtient une relation simplifiée}$$

qui se ramène à (4.41) en tenant compte de $C_s = 2 C_v$.

La formule (4.42) fait intervenir les valeurs de la population pour les moments et les coefficients, en pratique on leur substitue les estimations obtenues à partir de l'échantillon,

$$\left(\frac{\partial K}{\partial C_s} \right)_0 \quad \text{peut d'autre part être calculé à partir des tables}$$

de HARTER (1969).

b) Calcul de $(E.T.)_{Y_T}$ quand $\bar{C}_s \neq 2 \bar{C}_v$

On se trouve dans ce cas lorsque l'on considère la distribution Pearson III à 3 paramètres, ou lorsque pour les formes à 2 ou 1 paramètres on utilise la méthode des moments.

L'évènement Y_T est alors fonction de m'_1 , \bar{C}_v , et \bar{C}_s et l'on a d'après (3.29):

$$\begin{aligned} \left[(E.T.)_{Y_T} \right]^2 &= \left(\frac{\partial Y_T}{\partial m'_1} \right)_o^2 \text{ var } m'_1 + \left(\frac{\partial Y_T}{\partial \bar{C}_v} \right)_o^2 \text{ var } C_v + \left(\frac{\partial Y_T}{\partial \bar{C}_s} \right)_o^2 \text{ var } C_s \\ &+ 2 \left(\frac{\partial Y_T}{\partial m'_1} \right)_o \left(\frac{\partial Y_T}{\partial \bar{C}_v} \right)_o \text{ Cov } (m'_1, C_v) + 2 \left(\frac{\partial Y_T}{\partial m'_1} \right)_o \left(\frac{\partial Y_T}{\partial \bar{C}_s} \right)_o \text{ Cov } (m'_1, C_s) \\ &+ 2 \left(\frac{\partial Y_T}{\partial \bar{C}_v} \right)_o \left(\frac{\partial Y_T}{\partial \bar{C}_s} \right)_o \text{ Cov } (C_v, C_s). \end{aligned}$$

Les dérivées partielles sont calculées au point $m'_1 = \mu_1$, $m'_2 = \mu_2$ etc...., ce qui est symbolisé par l'indice o on peut alors montrer (BOBEE 1973'a) que l'on a:

$$\begin{aligned} \left[(E.T.)_{Y_T} \right]^2 &= \frac{\mu_2}{N} \left\{ 1 + \frac{\chi^2}{2} \left(1 + \frac{3}{4} C_s^2 \right) + \chi C_s + 6 \left(1 + \frac{C_s^2}{4} \right) \left(\frac{\partial K}{\partial C_s} \right)_o \right. \\ &\cdot \left. \left[\left(\frac{\partial K}{\partial C_s} \right)_o \left(1 + 5 \frac{C_s^2}{4} \right) + \frac{\chi}{2} C_s \right] \right\} \quad (4.43) \end{aligned}$$

Pour l'utilisation de cette formule on remplace les moments et coefficients de la population par leurs estimations à partir

de l'échantillon, $\frac{\partial K}{\partial C_s}$ est déduit des tables de HARTER (1969). Lors-

que l'on néglige la variabilité de K , c'est-à-dire si $\frac{\partial K}{\partial C_s} = 0$, on obtient

la relation simplifiée (4.41).

Lors de l'estimation d'un évènement de période de retour T , on détermine à partir de l'échantillon de taille N , la valeur Y_T ainsi que l'erreur-type $(E.T.)_{Y_T}$, pour obtenir en particulier les limites de l'intervalle de confiance de y_T correspondant à un niveau donné. Il est donc important de mettre en évidence l'erreur commise dans la détermination des intervalles de confiance lorsque l'on utilise les formules simplifiées au lieu des relations qui tiennent compte de la variabilité de K .

Si on suppose que Y_T est distribué normalement avec une moyenne y_T et un écart-type $(E.T.)_{Y_T}$, ce qui est une approximation d'autant plus valable que N est grand, l'intervalle de confiance de y_T correspondant à un niveau $(1 - \alpha) \%$ est tel que :

$$Y_T - u_{\alpha/2} * (E.T.)_{Y_T} < y_T < Y_T + u_{\alpha/2} * (E.T.)_{Y_T}$$

La taille de l'intervalle de confiance pour une période de retour donnée T est donc $2 u_{\alpha/2} * (E.T.)_{Y_T}$, $u_{\alpha/2}$ étant la variable normale standardisée correspondant à la probabilité au dépassement $\alpha/2$.

Pour α et N fixés considérons les rapports :

$$k_1 = \frac{\left[(E.T.)_{Y_T} \right]_2}{\left[(E.T.)_{Y_T} \right]_1} \quad k_2 = \frac{\left[(E.T.)_{Y_T} \right]_3}{\left[(E.T.)_{Y_T} \right]_1}$$

Les indices 1, 2, 3, sont relatifs respectivement aux relations donnant les erreurs-types, (4.41); (4.42); (4.43); k_1 et k_2 représentent les rapports entre les tailles des intervalles de confiance obtenues par la formule complète et la formule simplifiée respectivement dans les cas $\bar{C}_s = 2 \bar{C}_v$ et $\bar{C}_s \neq 2 \bar{C}_v$. Les tables 6 et 7 donnent les valeurs de k_1 et k_2 pour une probabilité au dépassement $P = \frac{1}{T}$ et une asymétrie C_s données et permettent de mesurer l'influence de la variabilité de K sur les intervalles de confiance correspondant à un niveau donné lors de l'estimation d'un évènement de période de retour donnée T , à partir d'un échantillon de taille N .

Les fonctions $K(C_s)$ pour P (ou T) donné ont été ajustées, à partir des tables de HARTER (1969), par des polynomes pour déterminer $\left(\frac{\partial K}{\partial C_s} \right)$.

Les résultats obtenus, représentent en fait la limite des rapports k_1 et k_2 lorsque N devient très grand. Ils montrent cependant que pour les périodes de retour élevées et surtout dans le cas $\bar{C}_s \neq 2 \bar{C}_v$ l'emploi des formules simplifiées conduit à une sous-estimation importante des limites d'intervalle de confiance.

4.8 Génération d'une variable suivant une distribution Pearson III

Il est souvent utile de construire une série tirée d'une population distribuée suivant une loi Pearson III dont les paramètres α , λ , m sont connus ou encore dont la moyenne la variance, le coefficient d'asymétrie sont fixés. Ce problème existe en particulier si l'on veut:

- comparer deux méthodes d'estimation des paramètres par simulation;
- étudier par simulation le comportement asymptotiquement normal de la distribution des moments d'un échantillon;
- simuler une longue série conservant des caractéristiques données.

P Cs									
	.001	.002	.005	.010	.020	.050	.100	.200	.500
0.0	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
.1	1.036	1.032	1.026	1.021	1.016	1.009	1.003	.998	1.000
.2	1.066	1.059	1.048	1.039	1.030	1.016	1.005	.997	1.001
.3	1.091	1.081	1.067	1.054	1.041	1.022	1.007	.995	1.002
.4	1.113	1.100	1.082	1.067	1.050	1.026	1.008	.992	1.003
.5	1.131	1.116	1.095	1.077	1.058	1.030	1.008	.990	1.006
.6	1.145	1.129	1.105	1.086	1.064	1.032	1.007	.988	1.008
.7	1.157	1.140	1.114	1.093	1.069	1.034	1.007	.985	1.012
.8	1.167	1.149	1.121	1.098	1.073	1.035	1.005	.982	1.016
.9	1.176	1.156	1.127	1.103	1.076	1.036	1.004	.980	1.022
1.0	1.182	1.162	1.132	1.107	1.078	1.036	1.002	.977	1.028
1.1	1.188	1.167	1.136	1.110	1.080	1.035	1.000	.974	1.035
1.2	1.192	1.171	1.140	1.112	1.081	1.035	.998	.971	1.044
1.3	1.196	1.175	1.142	1.114	1.082	1.034	.995	.968	1.054
1.4	1.199	1.177	1.144	1.115	1.083	1.033	.992	.965	1.066
1.5	1.202	1.180	1.146	1.117	1.083	1.031	.989	.961	1.079
1.6	1.204	1.181	1.147	1.118	1.083	1.029	.986	.958	1.093
1.7	1.205	1.183	1.148	1.118	1.082	1.027	.982	.955	1.109
1.8	1.207	1.184	1.149	1.118	1.081	1.025	.978	.953	1.127
1.9	1.208	1.185	1.150	1.118	1.080	1.022	.975	.950	1.145
2.0	1.209	1.186	1.150	1.117	1.079	1.019	.970	.948	1.165
2.1	1.209	1.186	1.150	1.117	1.078	1.016	.966	.946	1.184
2.2	1.210	1.186	1.150	1.116	1.076	1.013	.962	.944	1.204
2.3	1.210	1.186	1.149	1.115	1.075	1.010	.957	.943	1.223
2.4	1.210	1.186	1.149	1.114	1.073	1.006	.953	.942	1.240
2.5	1.210	1.186	1.148	1.113	1.071	1.003	.948	.943	1.255
2.6	1.210	1.186	1.147	1.112	1.069	.999	.943	.944	1.267
2.7	1.210	1.185	1.147	1.111	1.067	.995	.939	.946	1.275
2.8	1.209	1.185	1.146	1.110	1.064	.991	.934	.949	1.277
2.9	1.209	1.184	1.145	1.109	1.061	.986	.929	.953	1.274
3.0	1.209	1.184	1.143	1.107	1.059	.982	.925	.959	1.265
3.1	1.208	1.183	1.142	1.105	1.056	.978	.920	.965	1.249
3.2	1.208	1.182	1.141	1.103	1.053	.973	.916	.975	1.226
3.3	1.207	1.182	1.140	1.101	1.050	.969	.911	.985	1.196
3.4	1.207	1.181	1.138	1.100	1.047	.964	.908	.997	1.159
3.5	1.207	1.180	1.137	1.098	1.044	.960	.904	1.011	1.115
3.6	1.206	1.179	1.136	1.096	1.041	.955	.901	1.027	1.065
3.7	1.205	1.179	1.135	1.094	1.038	.951	.897	1.044	1.008
3.8	1.205	1.178	1.133	1.092	1.035	.946	.895	1.063	.945
3.9	1.205	1.178	1.132	1.090	1.032	.942	.893	1.084	.877
4.0	1.205	1.177	1.131	1.088	1.030	.938	.891	1.105	.803

TABLE 6 : INFLUENCE DE LA VARIABILITE DE K SUR LES INTERVALLES DE CONFIANCE LORSQUE $\bar{C}_s = 2 \bar{C}_v$

C_s \ P	.001	.002	.005	.010	.020	.050	.100	.200	.500
0.0	1.774	1.657	1.497	1.372	1.249	1.099	1.019	1.005	1.073
.1	1.770	1.659	1.505	1.384	1.262	1.110	1.024	1.002	1.076
.2	1.776	1.668	1.517	1.398	1.275	1.120	1.028	.998	1.083
.3	1.786	1.680	1.531	1.412	1.288	1.128	1.029	.994	1.093
.4	1.798	1.694	1.545	1.425	1.300	1.134	1.029	.990	1.106
.5	1.812	1.708	1.559	1.438	1.310	1.138	1.028	.986	1.122
.6	1.825	1.721	1.572	1.449	1.319	1.141	1.026	.983	1.143
.7	1.837	1.734	1.584	1.459	1.326	1.143	1.022	.980	1.168
.8	1.849	1.745	1.594	1.468	1.332	1.143	1.018	.978	1.198
.9	1.859	1.756	1.603	1.475	1.336	1.141	1.012	.977	1.232
1.0	1.868	1.765	1.611	1.481	1.339	1.139	1.006	.978	1.271
1.1	1.876	1.772	1.618	1.486	1.341	1.135	1.000	.980	1.315
1.2	1.883	1.779	1.624	1.490	1.342	1.131	.993	.984	1.365
1.3	1.889	1.785	1.628	1.493	1.342	1.125	.986	.991	1.419
1.4	1.894	1.790	1.632	1.495	1.340	1.119	.978	1.000	1.478
1.5	1.898	1.794	1.635	1.495	1.338	1.112	.971	1.013	1.541
1.6	1.901	1.797	1.636	1.495	1.335	1.104	.964	1.030	1.607
1.7	1.904	1.799	1.638	1.495	1.331	1.095	.957	1.050	1.677
1.8	1.906	1.800	1.638	1.493	1.327	1.086	.951	1.076	1.748
1.9	1.908	1.802	1.638	1.491	1.321	1.076	.945	1.106	1.820
2.0	1.909	1.802	1.637	1.487	1.315	1.065	.941	1.142	1.890
2.1	1.909	1.802	1.635	1.484	1.308	1.054	.938	1.183	1.957
2.2	1.909	1.802	1.633	1.479	1.300	1.043	.937	1.230	2.020
2.3	1.908	1.800	1.630	1.474	1.292	1.031	.937	1.284	2.074
2.4	1.907	1.799	1.626	1.468	1.283	1.019	.940	1.343	2.119
2.5	1.906	1.797	1.622	1.462	1.274	1.008	.946	1.409	2.151
2.6	1.904	1.794	1.618	1.455	1.264	.996	.955	1.482	2.168
2.7	1.902	1.792	1.613	1.448	1.253	.985	.966	1.561	2.167
2.8	1.900	1.788	1.608	1.440	1.242	.974	.981	1.646	2.147
2.9	1.898	1.785	1.603	1.432	1.231	.963	1.000	1.737	2.105
3.0	1.895	1.782	1.597	1.424	1.219	.953	1.023	1.834	2.041
3.1	1.892	1.778	1.591	1.416	1.208	.944	1.049	1.936	1.954
3.2	1.890	1.774	1.585	1.407	1.196	.936	1.078	2.044	1.844
3.3	1.887	1.771	1.579	1.398	1.183	.929	1.112	2.157	1.712
3.4	1.885	1.767	1.573	1.389	1.171	.924	1.149	2.274	1.560
3.5	1.882	1.763	1.566	1.379	1.159	.920	1.190	2.395	1.391
3.6	1.880	1.760	1.560	1.370	1.147	.917	1.234	2.520	1.210
3.7	1.878	1.756	1.554	1.361	1.135	.916	1.280	2.646	1.024
3.8	1.877	1.753	1.548	1.353	1.124	.917	1.329	2.774	.850
3.9	1.876	1.751	1.542	1.344	1.112	.919	1.380	2.902	.721
4.0	1.875	1.748	1.537	1.336	1.102	.923	1.432	3.028	.689

TABLE 7 : INFLUENCE DE LA VARIABILITE DE K SUR LES INTERVALLES DE
CONFIANCE LORSQUE $\bar{C}_s \neq 2 \bar{C}_v$

Les méthodes existantes de génération d'une série de variables Pearson III utilisent les propriétés de la distribution chi-deux (χ^2), qui est une forme particulière de la distribution Pearson III (paragraphe 1.2).

Soit Z une variable suivant une distribution χ^2 à ν degrés de liberté, Z est également une loi Gamma de paramètre

$$\alpha_0 = \frac{1}{2}; \quad \lambda_0 = \frac{\nu}{2};$$

donc:

$$\mu_1'(Z) = \nu; \quad \mu_2'(Z) = 2\nu; \quad c_s'(Z) = \sqrt{\frac{8}{\nu}}$$

La variable standardisée Pearson III d'asymétrie $\sqrt{\frac{8}{\nu}}$ est alors:

$$K = \frac{Z - \mu_1'(Z)}{\sqrt{\mu_2'(Z)}} = \frac{Z - \nu}{\sqrt{2\nu}}$$

Une variable X tirée d'une distribution Pearson III à asymétrie positive définie par $\mu_1'(X)$, $\mu_2(X)$, $c_s(X)$ est telle que

$$X = \mu_1'(X) + \sqrt{\mu_2(X)} \left[\frac{Z - \nu}{\sqrt{2\nu}} \right] \quad (4.44)$$

avec: $c_s(X) = \sqrt{\frac{8}{\nu}}$

Si la distribution Pearson III est définie en fonction des paramètres

α , λ , m on a:

$$X = \left(m + \frac{\lambda}{\alpha} \right) + \frac{\sqrt{\lambda}}{\alpha} \left[\frac{Z - \nu}{\sqrt{2\nu}} \right] \quad (4.45)$$

avec: $\nu = 2\lambda$

Puisque ν doit être un entier supérieur ou égal à 1, les valeurs possibles de λ sont des multiples de 1/2 et l'on a:

$$c_s(X) \leq 2\sqrt{2} \quad \text{ou} \quad \lambda \geq \frac{1}{2}$$

Pour la génération d'une série tirée d'une distribution Pearson III à asymétrie négative on utilise la relation qui existe entre les probabilités des formes A et B de la distribution (Ch. 1)

Pour obtenir une série tirée d'une distribution Pearson III de paramètres α, λ, m il suffit de construire une série suivant une distribution χ^2 à $\nu = 2\lambda$ degrés de liberté, pour cela différentes méthodes sont possibles:

- a) Si l'on considère ν variables normales standardisées u_i , la variable Z telle que:

$$Z = \sum_{i=1}^{\nu} u_i^2$$

est distribuée suivant une distribution χ^2 à ν degrés de liberté et l'on peut déterminer la variable Pearson III, X par (4.45)

Il est possible à partir du même principe de construire une série tirée de la loi Pearson III et dont le coefficient d'autocorrélation d'ordre 1 n'est pas nul (YEVJEVICH 1966)

- b) WILSON et HILFERTY (1931) ont montré que $\left(\frac{\chi^2}{\nu}\right)^{1/3}$ suit approximativement une loi normale de moyenne $(1 - 2/9 \nu)$ et de variance $(2/9 \nu)$; il en résulte que si u est une variable normale standardisée, K tel que:

$$K = \frac{2}{C_s} \left\{ \left[1 - \left(\frac{C_s}{6}\right)^2 + \left(\frac{C_s}{6}\right) u \right]^3 - 1 \right\} \quad (4.46)$$

suit une distribution Pearson III standardisée d'asymétrie C_s .

KIRBY (1972) a montré que pour les fortes asymétries ($C_s > 3$) une transformation modifiée est préférable.

- c) FISHER et CORNISH (1960) ont montré que l'on peut, pour une distribution donnée, exprimer la valeur de la variable aléatoire

correspondant à une probabilité au dépassement donnée en fonction de la variable normale de même probabilité. Dans le cas de la distribution χ^2 à ν degrés de liberté, la variable Z de même probabilité que la variable normale standardisée u est telle que:

$$\begin{aligned}
 Z = & \nu + \sqrt{2\nu} u + \frac{2}{3} (u^2 - 1) + \frac{1}{\sqrt{\nu}} \frac{(u^3 - 7u)}{9\sqrt{2}} \\
 & - \frac{1}{\nu} \frac{(6u^4 + 14u^2 - 32)}{405} + \frac{1}{\nu\sqrt{\nu}} \frac{(9u^5 + 256u^3 - 433u)}{4860\sqrt{2}} \\
 & - \frac{1}{\nu^2} \left(\frac{12u^6 - 243u^4 - 923u^2 + 1472}{25515} \right) \\
 & - \frac{1}{\nu^2\sqrt{\nu}} \left(\frac{3753u^7 + 4353u^5 - 289517u^3 - 289717u}{9185400\sqrt{2}} \right) \quad (4.47)
 \end{aligned}$$

BRUNET-MORET (1971) a montré la convergence de ce développement. Les tables que l'on peut établir à partir de (4.47) pour la distribution Pearson III standardisée concordent avec celles de HARTER (1969).

d) JÖHNK (1964) a mis au point un algorithme permettant la génération d'une variate Pearson III à partir d'une distribution uniforme sur l'intervalle (0,1). BERMAN (1971) décrit également ce processus de génération. Considérons la variate y qui suit une distribution Pearson III à 1 paramètre λ , si $[\lambda]$ et s sont respectivement les parties entière et décimale de λ on a:

$$y_i = - \sum_{k=1}^{[\lambda]} \log u_{i,k} - B_i \text{Log } t_i \quad (4.48)$$

avec:

$$B_i = \frac{v_i^{1/s}}{v_i^{1/s} + w_i^{1/(1-s)}} \quad (4.49)$$

u_{ik} , t_i , v_i et w_i sont des variates tirées d'une loi uniforme définie sur $[0,1]$

B_i défini par la relation (4.49) suit une distribution Bêta de paramètres s et $(1-s)$.

Au cours de cette génération on doit vérifier que:

$$\left(v_i \frac{1}{s} + w_i \frac{1}{(1-s)} \right) \leq 1$$

Pour générer une variate x_i qui suit une loi Pearson III de paramètres (α, λ, m) on utilise la relation:

$$x_i = m + \frac{1}{\alpha} y_i$$

Cette génération est valable pour la distribution Pearson III à asymétrie positive ($\alpha > 0$) mais aussi à asymétrie négative ($\alpha < 0$)

4.9 Statistiques d'ordre de la distribution Pearson III

On a vu (paragraphe 3.5) que la valeur d'ordre k des événements classés d'un échantillon de taille N , tiré d'une population donnée est une réalisation de la statistique d'ordre k . La fonction densité de cette variate peut s'exprimer en fonction de k , N , et de la fonction densité de la population considérée, par la relation (3.31).

La connaissance de la distribution des statistiques d'ordre, donc de la valeur correspondant à une probabilité au dépassement donnée, est utile notamment en ce qui concerne les valeurs extrêmes et la valeur médiane.

GUPTA (1960) a déterminé la distribution des statistiques d'ordre, dans le cas de la distribution Pearson III à un seul paramètre prenant des valeurs entières, et, a considéré différentes applications relatives aux tests de durée de vie. Ici nous considérons le cas de la distribution Pearson III standardisée en mettant en évidence la relation entre la distribution de la forme à asymétrie positive et celle de la forme à asymétrie négative (BOBEE et MORIN 1973 a).

On considère deux échantillons de taille N , l'un tiré de la loi Pearson III standardisée à asymétrie positive dont la fonction de densité est f et la fonction de distribution F , l'autre tiré de la loi Pearson III à asymétrie négative définie par f_a et F_a , les deux coefficients d'asymétrie étant opposés

$$\left[\left(C_s \right)_a = - C_s \right]$$

On désigne par X_k l'évènement d'ordre k relatif à la loi f et par Y_k celui relatif à la loi f_a .

Soient $h(X_k)$ et $h_a(Y_k)$ les densités de probabilité de X_k et Y_k qui sont définies par la relation (3.31).

$H(z, k)$ et $H_a(z, k)$ les fonctions de distribution cumulée des X_k et Y_k .

On a d'après la relation (3.32)

$$H(z, k) = \Pr \left[X_k \leq z \right] = \int_{-\sqrt{\lambda}}^z h(X_k) dX_k$$

$$H_a(z, k) = \Pr \left[Y_k \leq z \right] = \int_{-\infty}^z h_a(Y_k) dY_k$$

Il est possible à partir de la relation (1.8) entre f et f_a de montrer que H et H_a vérifient:

$$H_a(z, k) = 1 - H(-z, N-k + 1) \quad (4.50)$$

Cette relation permet de déterminer les distributions cumulées des évènements ordonnés Y_k ($k = 1, N$) de la loi Pearson III à asymétrie négative, lorsqu'on connaît les distributions cumulées des évènements ordonnés X_k ($k = 1, N$) de la loi Pearson III classique à asymétrie positive. (Le N est le même pour les 2 lois, tandis que les C_s sont opposés).

Graphiquement la relation (4.50) montre que la statistique d'ordre k de la forme à asymétrie négative se déduit de la statistique d'ordre $(N - k + 1)$ de la forme à asymétrie positive par une symétrie par rapport au point o de coordonnées $Z = 0, P = .5$ (P étant la probabilité au dépassement).

Dans le cas particulier de la distribution normale, on a $C_s = \left(C_s \right)_a = 0$ et la relation (4.50) devient:

$$H(Z, k) = 1 - H(-Z, N - k - 1) \quad (4.51)$$

et la statistique d'ordre $(N - k + 1)$ est déduite de la statistique d'ordre k par une symétrie par rapport au point o .

Résultats

Les calculs ont été effectués par intégration numérique pour $C_s > 0$ et permettent k, N, C_s étant fixés, de déterminer Z pour une probabilité au non-dépassement H , (la détermination de la valeur de la variable Pearson III standardisée correspondant à une probabilité donnée est faite à l'aide du développement de CORNISH-FISHER (4.8.c)).

Les valeurs relatives à un coefficient d'asymétrie négatif sont déduits par symétrie (relation 4.50).

L'ensemble des résultats est présenté sous forme de tables (BOBEE et MORIN 1972a), qui pour chaque couple (C_s, N) envisage les valeurs de k telles que:

$$k = i \left(\frac{N - 1}{10} \right) + 1 \quad i = 0, 1, \dots, 10$$

C_s varie de 0 à 1.8 avec un pas de .1

N prend les valeurs 11, 21, 41, 61, 81, 101.

Ici nous présentons les tables et figures correspondant à :

$C_s = 0$ (table 8; figure 8)

$C_s = +1$ (table 9; figure 9)

$C_s = -1$ (table 10; figure 10)

4.10 Détermination des intervalles de confiance de la distribution Pearson III

Lorsqu'à partir d'un échantillon de taille N on détermine les paramètres de la distribution dont l'échantillon est tiré, il est alors possible d'estimer la valeur d'un évènement de probabilité donnée. Cette estimation est soumise à une certaine erreur d'échantillonnage et le problème qui se pose est de déterminer l'intervalle de confiance de l'évènement de probabilité donné.

BRUNET-MORET (1973 b) a proposé une méthode générale valable en particulier pour la loi Pearson III qui consiste pour un évènement de période de retour T (ou de probabilité au dépassement $p = \frac{1}{T}$) à :

- déterminer les paramètres de la distribution à partir de l'échantillon de taille N par la méthode la plus adéquate;
- simuler r échantillons de taille N à partir de la distribution dont les paramètres sont connus (r doit être assez grand)
- recalculer pour chaque échantillon les valeurs des para-

STATISTIQUES D'ORDRE POUR LA LOI PEARSON III STANDARDISEE

ASYMETRIE = 0.000

NOMBRE D'OBSERVATIONS N = 11

PROB. /K	1	2	3	4	5	6	7	8	9	10	11
.010	-3.117	-2.196	-1.719	-1.381	-1.106	-.863	-.636	-.413	-.182	.075	.406
.020	-2.905	-2.049	-1.596	-1.269	-1.000	-.761	-.536	-.313	-.080	.183	.526
.050	-2.601	-1.835	-1.413	-1.103	-.843	-.609	-.385	-.162	.075	.346	.711
.100	-2.345	-1.651	-1.255	-.957	-.704	-.474	-.252	-.027	.214	.495	.882
.200	-2.052	-1.436	-1.067	-.783	-.538	-.311	-.089	.137	.385	.679	1.098
.300	-1.854	-1.285	-.935	-.659	-.418	-.193	.028	.257	.510	.814	1.261
.400	-1.692	-1.160	-.823	-.554	-.317	-.093	.128	.360	.618	.933	1.406
.500	-1.546	-1.045	-.720	-.457	-.222	-.000	.222	.457	.720	1.045	1.546
.600	-1.406	-.933	-.618	-.360	-.128	.093	.317	.554	.823	1.160	1.692
.700	-1.261	-.814	-.510	-.257	-.028	.193	.418	.659	.935	1.285	1.854
.800	-1.098	-.679	-.385	-.137	.089	.311	.538	.783	1.067	1.436	2.053
.900	-.882	-.495	-.214	.027	.252	.474	.704	.957	1.255	1.651	2.345
.950	-.711	-.346	-.075	.162	.385	.609	.843	1.103	1.413	1.835	2.601
.980	-.526	-.183	.080	.313	.536	.761	1.000	1.269	1.595	2.049	2.906
.990	-.406	-.075	.182	.413	.636	.863	1.106	1.381	1.719	2.196	3.118

ASYMETRIE = 0.000

NOMBRE D'OBSERVATIONS N = 21

PROB. /K	1	3	5	7	9	11	13	15	17	19	21
.010	-3.303	-2.023	-1.512	-1.164	-.881	-.630	-.392	-.154	.101	.401	.852
.020	-3.102	-1.912	-1.422	-1.082	-.805	-.556	-.319	-.080	.177	.484	.954
.050	-2.815	-1.750	-1.288	-.961	-.690	-.445	-.209	.031	.293	.611	1.113
.100	-2.576	-1.611	-1.172	-.855	-.589	-.346	-.111	.131	.398	.725	1.260
.200	-2.306	-1.447	-1.034	-.728	-.467	-.227	.008	.253	.526	.867	1.448
.300	-2.124	-1.333	-.935	-.637	-.380	-.141	.094	.341	.619	.972	1.592
.400	-1.977	-1.237	-.852	-.560	-.306	-.068	.168	.417	.700	1.062	1.721
.500	-1.846	-1.149	-.776	-.488	-.237	-.000	.237	.488	.776	1.149	1.846
.600	-1.721	-1.062	-.700	-.417	-.168	.068	.306	.560	.852	1.237	1.977
.700	-1.592	-.972	-.619	-.341	-.094	.141	.380	.637	.935	1.333	2.124
.800	-1.448	-.867	-.526	-.253	-.008	.227	.467	.728	1.034	1.447	2.306
.900	-1.260	-.725	-.398	-.131	.111	.346	.589	.855	1.172	1.611	2.576
.950	-1.112	-.611	-.293	-.031	.209	.445	.690	.961	1.288	1.750	2.815
.980	-.954	-.484	-.177	.080	.319	.556	.804	1.082	1.422	1.912	3.103
.990	-.852	-.401	-.101	.154	.392	.630	.881	1.164	1.512	2.023	3.304

ASYMETRIE = 0.000

NOMBRE D'OBSERVATIONS N = 41

PROB. /K	1	5	9	13	17	21	25	29	33	37	41
.010	-3.486	-1.849	-1.337	-.990	-.707	-.453	-.209	.039	.312	.650	1.247
.020	-3.295	-1.769	-1.272	-.932	-.652	-.400	-.156	.093	.369	.712	1.335
.050	-3.023	-1.652	-1.177	-.846	-.570	-.320	-.077	.174	.454	.808	1.473
.100	-2.799	-1.551	-1.094	-.770	-.498	-.249	-.006	.247	.531	.894	1.602
.200	-2.548	-1.431	-.994	-.678	-.411	-.164	.080	.335	.625	1.001	1.769
.300	-2.380	-1.347	-.923	-.613	-.348	-.102	.142	.399	.693	1.079	1.897
.400	-2.245	-1.276	-.863	-.557	-.294	-.049	.195	.454	.752	1.147	2.013
.500	-2.126	-1.211	-.807	-.505	-.245	-.000	.245	.505	.807	1.211	2.126
.600	-2.013	-1.147	-.752	-.454	-.195	.049	.294	.557	.863	1.276	2.246
.700	-1.897	-1.079	-.693	-.399	-.142	.102	.348	.613	.923	1.347	2.380
.800	-1.769	-1.001	-.625	-.335	-.080	.164	.411	.678	.994	1.431	2.548
.900	-1.602	-.894	-.531	-.247	.006	.249	.498	.770	1.094	1.551	2.799
.950	-1.472	-.808	-.454	-.174	.077	.320	.570	.846	1.177	1.652	3.024
.980	-1.334	-.712	-.369	-.093	.156	.400	.652	.932	1.272	1.769	3.295
.990	-1.246	-.650	-.312	-.039	.209	.453	.707	.990	1.337	1.849	3.487

TABLE 8 STATISTIQUES D'ORDRE DE LA LOI PEARSON III STANDARDISEE $C_s = 0$

STATISTIQUES D'ORDRE POUR LA LOI PEARSON III STANDARDISEE

ASYMETRIE = 0.000

NOMBRE D'OBSERVATIONS N = 61

PROB. /K	1	7	13	19	25	31	37	43	49	55	61
.010	-3.590	-1.758	-1.252	-.909	-.627	-.372	-.126	.127	.408	.763	1.456
.020	-3.404	-1.693	-1.200	-.861	-.582	-.328	-.082	.171	.455	.816	1.537
.050	-3.142	-1.598	-1.122	-.791	-.515	-.263	-.017	.239	.526	.896	1.666
.100	-2.925	-1.515	-1.054	-.728	-.455	-.205	.041	.298	.590	.969	1.786
.200	-2.683	-1.416	-.972	-.654	-.384	-.134	.112	.371	.668	1.058	1.943
.300	-2.523	-1.346	-.914	-.600	-.332	-.084	.163	.424	.724	1.123	2.064
.400	-2.394	-1.288	-.864	-.554	-.288	-.040	.206	.469	.773	1.180	2.173
.500	-2.281	-1.233	-.818	-.512	-.247	-.000	.247	.512	.818	1.233	2.281
.600	-2.173	-1.180	-.773	-.469	-.206	.040	.288	.554	.864	1.288	2.394
.700	-2.064	-1.123	-.724	-.424	-.163	.084	.332	.600	.914	1.346	2.523
.800	-1.943	-1.058	-.668	-.371	-.112	.134	.384	.654	.972	1.416	2.683
.900	-1.786	-.969	-.590	-.298	-.041	.205	.455	.728	1.054	1.515	2.925
.950	-1.665	-.896	-.526	-.239	.017	.263	.515	.791	1.122	1.598	3.142
.980	-1.537	-.816	-.455	-.171	.082	.328	.582	.861	1.200	1.693	3.405
.990	-1.455	-.763	-.408	-.127	.126	.372	.627	.909	1.252	1.758	3.592

ASYMETRIE = 0.000

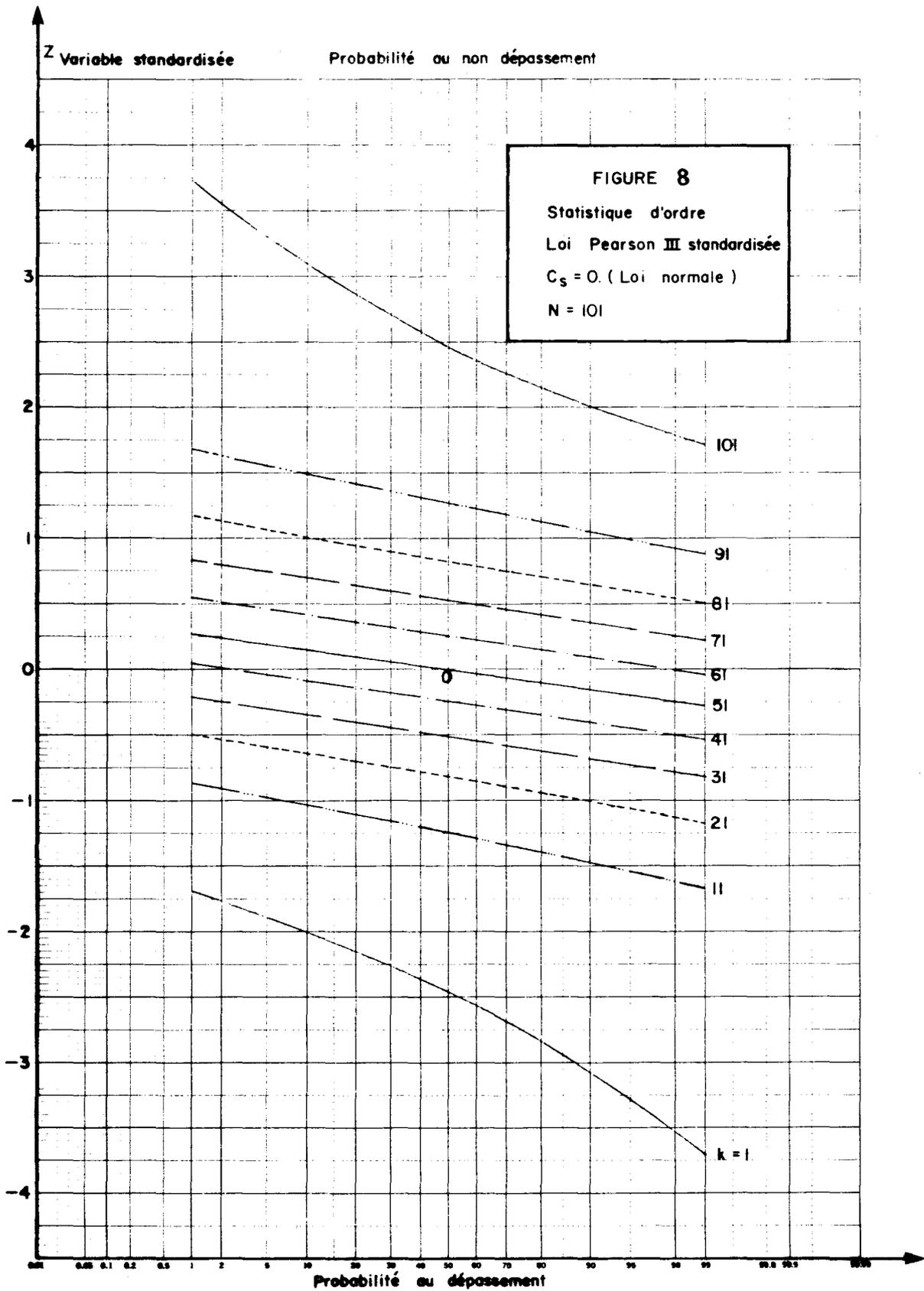
NOMBRE D'OBSERVATIONS N = 81

PROB. /K	1	9	17	25	33	41	49	57	65	73	81
.010	-3.663	-1.701	-1.200	-.859	-.578	-.323	-.076	.180	.466	.832	1.596
.020	-3.481	-1.645	-1.155	-.818	-.539	-.285	-.038	.219	.507	.879	1.673
.050	-3.223	-1.562	-1.088	-.757	-.481	-.228	.019	.277	.569	.949	1.796
.100	-3.012	-1.490	-1.029	-.703	-.429	-.178	.069	.329	.625	1.013	1.911
.200	-2.776	-1.404	-.958	-.638	-.367	-.117	.131	.392	.693	1.091	2.061
.300	-2.621	-1.344	-.907	-.591	-.322	-.073	.175	.438	.742	1.149	2.177
.400	-2.496	-1.292	-.864	-.552	-.284	-.035	.213	.478	.784	1.198	2.282
.500	-2.386	-1.245	-.824	-.515	-.249	-.000	.249	.515	.824	1.245	2.386
.600	-2.282	-1.198	-.784	-.478	-.213	.035	.284	.552	.864	1.292	2.496
.700	-2.177	-1.149	-.742	-.438	-.175	.073	.322	.591	.907	1.344	2.621
.800	-2.061	-1.091	-.693	-.392	-.131	.117	.367	.638	.958	1.404	2.777
.900	-1.911	-1.013	-.625	-.329	-.069	.178	.429	.703	1.029	1.490	3.012
.950	-1.795	-.949	-.569	-.277	-.019	.228	.481	.757	1.088	1.562	3.224
.980	-1.673	-.879	-.507	-.219	.038	.285	.539	.818	1.155	1.645	3.482
.990	-1.595	-.832	-.466	-.180	.076	.323	.578	.859	1.200	1.701	3.665

ASYMETRIE = 0.000

NOMBRE D'OBSERVATIONS N = 101

PROB. /K	1	11	21	31	41	51	61	71	81	91	101
.010	-3.719	-1.660	-1.164	-.825	-.544	-.289	-.041	.216	.505	.879	1.700
.020	-3.539	-1.610	-1.124	-.788	-.509	-.256	-.008	.251	.543	.921	1.775
.050	-3.286	-1.536	-1.064	-.734	-.458	-.205	.043	.303	.598	.985	1.893
.100	-3.078	-1.472	-1.011	-.685	-.411	-.159	.089	.350	.648	1.043	2.004
.200	-2.847	-1.395	-.947	-.627	-.356	-.105	.144	.407	.709	1.114	2.150
.300	-2.695	-1.341	-.902	-.585	-.316	-.065	.184	.448	.754	1.165	2.262
.400	-2.573	-1.295	-.863	-.550	-.281	-.031	.218	.483	.792	1.210	2.365
.500	-2.466	-1.252	-.827	-.516	-.250	-.000	.250	.516	.827	1.252	2.466
.600	-2.365	-1.210	-.792	-.483	-.218	.031	.281	.550	.863	1.295	2.573
.700	-2.262	-1.165	-.754	-.448	-.184	.065	.316	.585	.902	1.341	2.695
.800	-2.149	-1.114	-.709	-.407	-.144	.105	.356	.627	.947	1.395	2.847
.900	-2.004	-1.043	-.648	-.350	-.089	.159	.411	.685	1.011	1.472	3.078
.950	-1.892	-.985	-.598	-.303	-.043	.205	.458	.734	1.064	1.536	3.287
.980	-1.774	-.921	-.543	-.251	.008	.256	.509	.788	1.124	1.610	3.541
.990	-1.699	-.879	-.505	-.216	.041	.289	.544	.825	1.164	1.660	3.722



STATISTIQUES D'ORDRE POUR LA LOI PEARSON III STANDARDISEE

ASYMETRIE = 1.000

NOMBRE D'OBSERVATIONS N = 11

PROB. /K	1	2	3	4	5	6	7	8	9	10	11
.010	-1.791	-1.544	-1.352	-1.183	-1.024	-.867	-.704	-.530	-.332	-.091	.251
.020	-1.747	-1.490	-1.293	-1.121	-.958	-.796	-.628	-.446	-.239	.016	.383
.050	-1.671	-1.403	-1.200	-1.022	-.853	-.684	-.507	-.314	-.092	.186	.599
.100	-1.594	-1.320	-1.112	-.929	-.755	-.578	-.393	-.190	.047	.348	.808
.200	-1.491	-1.212	-1.000	-.811	-.629	-.444	-.248	-.030	.227	.560	1.088
.300	-1.411	-1.130	-.915	-.721	-.534	-.342	-.137	.092	.365	.724	1.311
.400	-1.339	-1.057	-.839	-.642	-.449	-.252	-.039	.200	.488	.872	1.518
.500	-1.269	-.986	-.766	-.565	-.368	-.164	.056	.306	.609	1.018	1.726
.600	-1.196	-.913	-.690	-.485	-.283	-.074	.155	.415	.735	1.172	1.949
.700	-1.116	-.833	-.607	-.398	-.190	.027	.264	.537	.875	1.345	2.208
.800	-1.019	-.736	-.506	-.292	-.077	.148	.397	.686	1.048	1.561	2.539
.900	-.879	-.595	-.360	-.138	.086	.325	.591	.904	1.303	1.885	3.055
.950	-.759	-.474	-.234	-.005	.228	.478	.759	1.095	1.529	2.177	3.537
.980	-.619	-.333	-.087	.150	.395	.659	.960	1.323	1.800	2.533	4.144
.990	-.523	-.235	.015	.258	.510	.785	1.099	1.482	1.992	2.788	4.590

ASYMETRIE = 1.000

NOMBRE D'OBSERVATIONS N = 21

PROB. /K	1	3	5	7	9	11	13	15	17	19	21
.010	-1.825	-1.480	-1.252	-1.059	-.879	-.700	-.512	-.307	-.067	.245	.771
.020	-1.788	-1.435	-1.205	-1.009	-.826	-.643	-.451	-.239	.010	.337	.900
.050	-1.726	-1.366	-1.132	-.932	-.744	-.555	-.356	-.134	.130	.480	1.108
.100	-1.664	-1.300	-1.064	-.861	-.669	-.474	-.268	-.036	.241	.616	1.310
.200	-1.582	-1.218	-.979	-.772	-.573	-.372	-.156	.087	.383	.789	1.580
.300	-1.518	-1.156	-.915	-.705	-.502	-.295	-.073	.180	.490	.922	1.795
.400	-1.462	-1.102	-.859	-.646	-.440	-.228	.000	.262	.585	1.041	1.995
.500	-1.407	-1.050	-.806	-.590	-.380	-.164	.071	.341	.677	1.157	2.195
.600	-1.352	-.997	-.751	-.532	-.319	-.098	.143	.422	.771	1.278	2.412
.700	-1.291	-.939	-.691	-.469	-.252	-.026	.222	.511	.876	1.412	2.663
.800	-1.218	-.869	-.619	-.394	-.172	.061	.318	.619	1.003	1.578	2.985
.900	-1.115	-.770	-.517	-.286	-.056	.186	.455	.775	1.188	1.824	3.488
.950	-1.027	-.685	-.429	-.194	.042	.293	.574	.909	1.350	2.041	3.960
.980	-.927	-.587	-.328	-.087	.157	.417	.712	1.067	1.541	2.303	4.557
.990	-.858	-.520	-.258	-.013	.235	.503	.807	1.177	1.674	2.488	4.995

ASYMETRIE = 1.000

NOMBRE D'OBSERVATIONS N = 41

PROB. /K	1	5	9	13	17	21	25	29	33	37	41
.010	-1.853	-1.409	-1.159	-.951	-.757	-.562	-.356	-.126	.150	.526	1.292
.020	-1.823	-1.374	-1.123	-.913	-.716	-.519	-.309	-.074	.210	.600	1.415
.050	-1.772	-1.320	-1.067	-.855	-.654	-.452	-.236	.007	.303	.716	1.616
.100	-1.722	-1.271	-1.017	-.802	-.598	-.391	-.170	.081	.389	.823	1.810
.200	-1.656	-1.210	-.954	-.735	-.527	-.315	-.087	.174	.497	.960	2.071
.300	-1.606	-1.164	-.907	-.687	-.475	-.259	-.026	.243	.577	1.063	2.279
.400	-1.561	-1.124	-.866	-.644	-.430	-.210	.028	.303	.648	1.154	2.471
.500	-1.519	-1.087	-.828	-.604	-.387	-.164	.079	.360	.715	1.242	2.666
.600	-1.476	-1.049	-.789	-.562	-.343	-.117	.131	.419	.784	1.332	2.876
.700	-1.429	-1.007	-.746	-.518	-.296	-.065	.188	.483	.860	1.433	3.120
.800	-1.374	-.957	-.695	-.464	-.239	-.004	.255	.560	.952	1.555	3.434
.900	-1.296	-.887	-.623	-.389	-.158	.084	.352	.669	1.083	1.733	3.927
.950	-1.230	-.828	-.563	-.325	-.090	.158	.434	.763	1.196	1.887	4.389
.980	-1.156	-.760	-.493	-.251	-.011	.244	.529	.872	1.327	2.071	4.976
.990	-1.106	-.714	-.445	-.201	.043	.302	.594	.946	1.418	2.200	5.409

TABLE 9 - STATISTIQUES D'ORDRE POUR LA LOI PEARSON III STANDARDISEE $C_s = + 1$

STATISTIQUES D'ORDRE POUR LA LOI PEARSON III STANDARDISEE

ASYMETRIE = 1.000

NOMBRE D'OBSERVATIONS N = 61

PROB. /K	1	7	13	19	25	31	37	43	49	55	61
.010	-1.868	-1.370	-1.111	-.898	-.697	-.496	-.282	-.040	.253	.662	1.591
.020	-1.841	-1.340	-1.081	-.866	-.663	-.459	-.242	.004	.304	.726	1.713
.050	-1.796	-1.294	-1.034	-.817	-.611	-.403	-.180	.073	.384	.826	1.908
.100	-1.751	-1.253	-.992	-.772	-.564	-.352	-.125	.135	.456	.919	2.099
.200	-1.693	-1.201	-.939	-.717	-.505	-.289	-.056	.212	.547	1.035	2.354
.300	-1.649	-1.164	-.901	-.677	-.462	-.243	-.005	.270	.614	1.122	2.558
.400	-1.610	-1.131	-.867	-.642	-.425	-.202	.040	.320	.673	1.199	2.748
.500	-1.573	-1.100	-.836	-.608	-.389	-.164	.082	.367	.729	1.273	2.939
.600	-1.536	-1.068	-.803	-.575	-.354	-.125	.124	.415	.786	1.349	3.146
.700	-1.495	-1.034	-.769	-.538	-.315	-.083	.171	.468	.848	1.432	3.387
.800	-1.448	-.994	-.727	-.495	-.268	-.033	.226	.530	.923	1.533	3.696
.900	-1.381	-.937	-.669	-.433	-.203	.038	.305	.620	1.030	1.678	4.183
.950	-1.326	-.889	-.620	-.382	-.148	.098	.371	.695	1.121	1.804	4.641
.980	-1.263	-.834	-.564	-.322	-.084	.167	.447	.782	1.227	1.951	5.223
.990	-1.220	-.797	-.525	-.282	-.041	.213	.499	.842	1.299	2.054	5.653

ASYMETRIE = 1.000

NOMBRE D'OBSERVATIONS N = 81

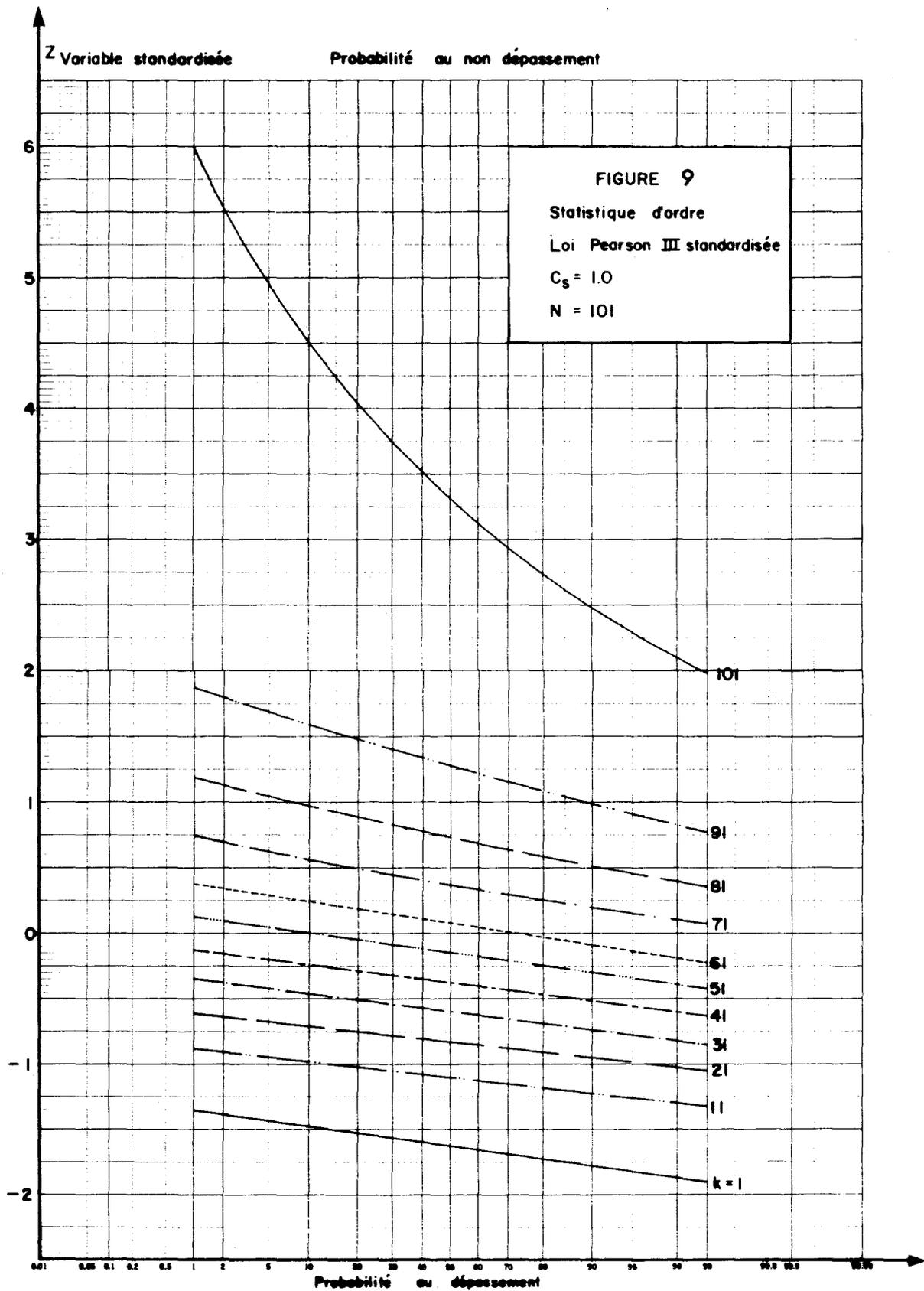
PROB. /K	1	9	17	25	33	41	49	57	65	73	81
.010	-1.878	-1.343	-1.081	-.864	-.660	-.455	-.235	.013	.316	.746	1.801
.020	-1.853	-1.317	-1.054	-.836	-.630	-.422	-.200	.052	.362	.804	1.921
.050	-1.811	-1.277	-1.013	-.793	-.584	-.373	-.146	.113	.432	.894	2.114
.100	-1.770	-1.240	-.976	-.754	-.543	-.328	-.097	.167	.497	.976	2.302
.200	-1.716	-1.195	-.930	-.706	-.492	-.273	-.036	.235	.577	1.080	2.554
.300	-1.676	-1.162	-.896	-.670	-.454	-.232	.008	.286	.636	1.157	2.755
.400	-1.641	-1.134	-.867	-.640	-.421	-.197	.047	.329	.687	1.224	2.942
.500	-1.607	-1.107	-.840	-.611	-.391	-.164	.083	.370	.736	1.289	3.131
.600	-1.573	-1.079	-.812	-.581	-.359	-.130	.120	.412	.786	1.355	3.336
.700	-1.537	-1.050	-.781	-.550	-.326	-.094	.161	.458	.840	1.428	3.574
.800	-1.494	-1.015	-.746	-.512	-.286	-.051	.208	.512	.905	1.516	3.882
.900	-1.434	-.965	-.695	-.459	-.229	.011	.276	.589	.997	1.641	4.365
.950	-1.385	-.924	-.653	-.415	-.182	.062	.333	.654	1.075	1.749	4.819
.980	-1.329	-.877	-.605	-.364	-.127	.121	.398	.729	1.165	1.876	5.398
.990	-1.291	-.845	-.572	-.330	-.091	.161	.443	.780	1.227	1.963	5.826

ASYMETRIE = 1.000

NOMBRE D'OBSERVATIONS N = 101

PROB. /K	1	11	21	31	41	51	61	71	81	91	101
.010	-1.885	-1.324	-1.060	-.841	-.634	-.426	-.204	.049	.360	.805	1.963
.020	-1.861	-1.300	-1.035	-.815	-.607	-.397	-.171	.085	.402	.858	2.080
.050	-1.822	-1.264	-.998	-.776	-.566	-.352	-.122	.140	.466	.940	2.271
.100	-1.783	-1.231	-.964	-.741	-.528	-.312	-.078	.190	.524	1.015	2.457
.200	-1.733	-1.190	-.923	-.697	-.482	-.262	-.023	.251	.597	1.109	2.707
.300	-1.696	-1.161	-.893	-.666	-.448	-.225	.017	.296	.650	1.179	2.906
.400	-1.663	-1.135	-.867	-.638	-.419	-.194	.051	.335	.696	1.241	3.092
.500	-1.632	-1.111	-.842	-.612	-.391	-.164	.084	.373	.740	1.299	3.280
.600	-1.600	-1.086	-.817	-.586	-.363	-.134	.117	.410	.785	1.359	3.483
.700	-1.567	-1.060	-.790	-.558	-.333	-.101	.153	.451	.833	1.424	3.719
.800	-1.527	-1.028	-.758	-.524	-.297	-.063	.196	.499	.891	1.502	4.025
.900	-1.472	-.984	-.713	-.477	-.247	-.008	.256	.568	.973	1.615	4.505
.950	-1.426	-.948	-.676	-.438	-.205	.038	.307	.626	1.043	1.711	4.957
.980	-1.375	-.905	-.632	-.392	-.156	.090	.365	.692	1.123	1.823	5.534
.990	-1.340	-.877	-.603	-.362	-.124	.126	.404	.737	1.178	1.900	5.960

TABLE 9 - (S U I T E)



STATISTIQUES D'ORDRE POUR LA LOI PEARSON III STANDARDISEE

ASYMETRIE = -1.000

NOMBRE D'OBSERVATIONS N = 11

PROB. /K	1	2	3	4	5	6	7	8	9	10	11
.010	-4.590	-2.788	-1.992	-1.482	-1.099	-.785	-.510	-.258	-.015	.235	.523
.020	-4.144	-2.533	-1.800	-1.323	-.960	-.659	-.395	-.150	.087	.333	.619
.050	-3.537	-2.177	-1.529	-1.095	-.759	-.478	-.228	.005	.234	.474	.759
.100	-3.055	-1.885	-1.303	-.904	-.591	-.325	-.086	.138	.360	.595	.879
.200	-2.539	-1.561	-1.048	-.686	-.397	-.148	.077	.292	.506	.736	1.019
.300	-2.208	-1.345	-.875	-.537	-.264	-.027	.190	.398	.607	.833	1.116
.400	-1.949	-1.172	-.735	-.415	-.155	.074	.283	.485	.690	.913	1.196
.500	-1.726	-1.018	-.609	-.306	-.056	.164	.368	.565	.766	.986	1.269
.600	-1.518	-.872	-.488	-.200	.039	.252	.449	.642	.839	1.057	1.339
.700	-1.311	-.724	-.365	-.092	.137	.342	.534	.721	.915	1.130	1.411
.800	-1.088	-.560	-.227	.030	.248	.444	.629	.811	1.000	1.212	1.491
.900	-.808	-.348	-.047	.190	.393	.578	.755	.929	1.112	1.320	1.594
.950	-.599	-.186	.092	.314	.507	.684	.853	1.022	1.200	1.403	1.671
.980	-.383	-.016	.239	.446	.628	.796	.958	1.121	1.293	1.490	1.747
.990	-.251	.091	.332	.530	.704	.867	1.024	1.183	1.352	1.544	1.791

ASYMETRIE = -1.000

NOMBRE D'OBSERVATIONS N = 21

PROB. /K	1	3	5	7	9	11	13	15	17	19	21
.010	-4.995	-2.488	-1.674	-1.177	-.807	-.503	-.235	.013	.258	.520	.858
.020	-4.557	-2.303	-1.541	-1.067	-.712	-.417	-.157	.087	.328	.587	.927
.050	-3.960	-2.041	-1.350	-.909	-.574	-.293	-.042	.194	.429	.685	1.027
.100	-3.488	-1.824	-1.188	-.775	-.455	-.186	.056	.286	.517	.770	1.115
.200	-2.985	-1.578	-1.003	-.619	-.318	-.061	.172	.394	.619	.869	1.218
.300	-2.663	-1.412	-.876	-.511	-.222	.026	.252	.469	.691	.939	1.291
.400	-2.412	-1.278	-.771	-.422	-.143	.098	.319	.532	.751	.997	1.352
.500	-2.195	-1.157	-.677	-.341	-.071	.164	.380	.590	.806	1.050	1.407
.600	-1.995	-1.041	-.585	-.262	-.000	.228	.440	.646	.859	1.102	1.462
.700	-1.795	-.922	-.490	-.180	.073	.295	.502	.705	.915	1.156	1.518
.800	-1.580	-.789	-.383	-.087	.156	.372	.573	.772	.979	1.218	1.582
.900	-1.310	-.616	-.241	.036	.268	.474	.669	.861	1.064	1.300	1.664
.950	-1.108	-.480	-.130	.134	.356	.555	.744	.932	1.132	1.366	1.726
.980	-.900	-.337	-.010	.239	.451	.643	.826	1.009	1.205	1.435	1.788
.990	-.771	-.245	.067	.307	.512	.700	.879	1.059	1.252	1.480	1.825

ASYMETRIE = -1.000

NOMBRE D'OBSERVATIONS N = 41

PROB. /K	1	5	9	13	17	21	25	29	33	37	41
.010	-5.409	-2.200	-1.418	-.946	-.594	-.302	-.043	.201	.445	.714	1.106
.020	-4.976	-2.071	-1.327	-.872	-.529	-.244	.011	.251	.493	.760	1.156
.050	-4.389	-1.887	-1.196	-.763	-.434	-.158	.090	.325	.563	.828	1.230
.100	-3.927	-1.733	-1.083	-.669	-.352	-.084	.158	.389	.623	.887	1.296
.200	-3.434	-1.555	-.952	-.560	-.255	.004	.239	.464	.695	.957	1.374
.300	-3.120	-1.433	-.860	-.483	-.188	.065	.296	.518	.746	1.007	1.429
.400	-2.876	-1.332	-.784	-.419	-.131	.117	.343	.562	.789	1.049	1.476
.500	-2.666	-1.242	-.715	-.360	-.079	.164	.387	.604	.828	1.087	1.519
.600	-2.471	-1.154	-.648	-.303	-.028	.210	.430	.644	.866	1.124	1.561
.700	-2.279	-1.063	-.577	-.243	.026	.259	.475	.687	.907	1.164	1.606
.800	-2.071	-.960	-.497	-.174	.087	.315	.527	.735	.954	1.210	1.656
.900	-1.810	-.823	-.389	-.081	.170	.391	.598	.802	1.017	1.271	1.722
.950	-1.616	-.716	-.303	-.007	.236	.452	.654	.855	1.067	1.320	1.772
.980	-1.415	-.600	-.210	.074	.309	.519	.716	.913	1.123	1.374	1.823
.990	-1.292	-.526	-.150	.126	.356	.562	.757	.951	1.159	1.409	1.853

TABLE 10 STATISTIQUES D'ORDRE POUR LA FORME DERIVEE DE LA LOI PEARSON III $C_s = -1$

STATISTIQUES D'ORDRE POUR LA LOI PEARSON III STANDARDISEE

ASYMETRIE = -1.000

NOMBRE D'OBSERVATIONS N = 61

PROB. /K	1	7	13	19	25	31	37	43	49	55	61
.010	-5.653	-2.054	-1.299	-.842	-.499	-.213	.041	.282	.525	.797	1.220
.020	-5.223	-1.951	-1.227	-.782	-.447	-.167	.084	.322	.564	.834	1.263
.050	-4.641	-1.804	-1.121	-.695	-.371	-.098	.148	.382	.620	.889	1.326
.100	-4.183	-1.678	-1.030	-.620	-.305	-.038	.203	.433	.669	.937	1.381
.200	-3.696	-1.533	-.923	-.530	-.226	.033	.268	.495	.727	.994	1.448
.300	-3.387	-1.432	-.848	-.468	-.171	.083	.315	.538	.769	1.034	1.495
.400	-3.146	-1.349	-.786	-.415	-.124	.125	.354	.575	.803	1.068	1.536
.500	-2.939	-1.273	-.729	-.367	-.082	.164	.389	.608	.836	1.100	1.573
.600	-2.748	-1.199	-.673	-.320	-.040	.202	.425	.642	.867	1.131	1.610
.700	-2.558	-1.122	-.614	-.270	.005	.243	.462	.677	.901	1.164	1.649
.800	-2.354	-1.035	-.547	-.212	.056	.289	.505	.717	.939	1.201	1.693
.900	-2.099	-.919	-.456	-.135	.125	.352	.564	.772	.992	1.253	1.751
.950	-1.908	-.826	-.384	-.073	.180	.403	.611	.817	1.034	1.294	1.796
.980	-1.713	-.726	-.304	-.004	.242	.459	.663	.866	1.081	1.340	1.841
.990	-1.591	-.662	-.253	.040	.282	.496	.697	.898	1.111	1.370	1.868

ASYMETRIE = -1.000

NOMBRE D'OBSERVATIONS N = 81

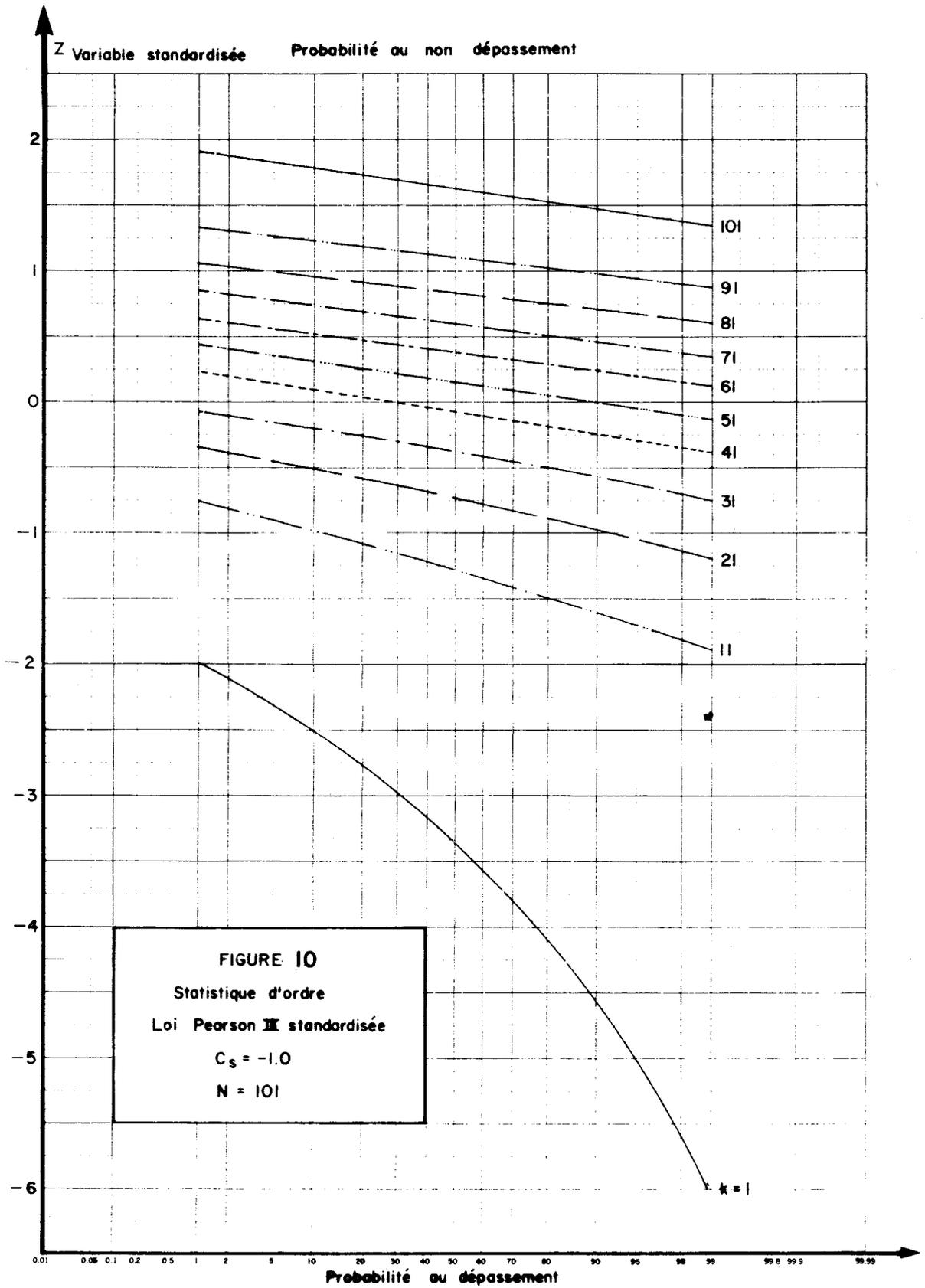
PROB. /K	1	9	17	25	33	41	49	57	65	73	81
.010	-5.826	-1.963	-1.227	-.780	-.443	-.161	.091	.330	.572	.845	1.291
.020	-5.398	-1.876	-1.165	-.729	-.398	-.121	.127	.364	.605	.877	1.329
.050	-4.819	-1.749	-1.075	-.654	-.333	-.062	.182	.415	.653	.924	1.385
.100	-4.365	-1.641	-.997	-.589	-.276	-.011	.229	.459	.695	.965	1.434
.200	-3.882	-1.516	-.905	-.512	-.208	.051	.286	.512	.746	1.015	1.494
.300	-3.574	-1.428	-.840	-.458	-.161	.094	.326	.550	.781	1.050	1.537
.400	-3.336	-1.355	-.786	-.412	-.120	.130	.359	.581	.812	1.079	1.573
.500	-3.131	-1.289	-.736	-.370	-.083	.164	.391	.611	.840	1.107	1.607
.600	-2.942	-1.224	-.687	-.329	-.047	.197	.421	.640	.867	1.134	1.641
.700	-2.755	-1.157	-.636	-.286	-.008	.232	.454	.670	.896	1.162	1.676
.800	-2.554	-1.080	-.577	-.235	.036	.273	.492	.706	.930	1.195	1.716
.900	-2.302	-.976	-.497	-.167	.097	.328	.543	.754	.976	1.240	1.770
.950	-2.114	-.894	-.432	-.113	.146	.373	.584	.793	1.013	1.277	1.811
.980	-1.921	-.804	-.362	-.052	.200	.422	.630	.836	1.054	1.317	1.853
.990	-1.801	-.746	-.316	-.013	.235	.455	.660	.864	1.081	1.343	1.878

ASYMETRIE = -1.000

NOMBRE D'OBSERVATIONS N = 101

PROB. /K	1	11	21	31	41	51	61	71	81	91	101
.010	-5.960	-1.900	-1.178	-.737	-.404	-.126	.124	.362	.603	.877	1.340
.020	-5.534	-1.823	-1.123	-.692	-.365	-.090	.156	.392	.632	.905	1.375
.050	-4.957	-1.711	-1.043	-.626	-.307	-.038	.205	.438	.676	.948	1.426
.100	-4.505	-1.615	-.973	-.568	-.256	.008	.247	.477	.713	.984	1.472
.200	-4.025	-1.502	-.891	-.499	-.196	.063	.297	.524	.758	1.028	1.527
.300	-3.719	-1.424	-.833	-.451	-.153	.101	.333	.558	.790	1.060	1.567
.400	-3.483	-1.359	-.785	-.410	-.117	.134	.363	.586	.817	1.086	1.600
.500	-3.280	-1.299	-.740	-.373	-.084	.164	.391	.612	.842	1.111	1.632
.600	-3.092	-1.241	-.696	-.335	-.051	.194	.419	.638	.867	1.135	1.663
.700	-2.906	-1.179	-.650	-.296	-.017	.225	.448	.666	.893	1.161	1.696
.800	-2.707	-1.109	-.597	-.251	.023	.262	.482	.697	.923	1.190	1.733
.900	-2.457	-1.015	-.524	-.190	.078	.312	.528	.741	.964	1.231	1.783
.950	-2.271	-.940	-.466	-.140	.122	.352	.566	.776	.998	1.264	1.822
.980	-2.080	-.858	-.402	-.085	.171	.397	.607	.815	1.035	1.300	1.861
.990	-1.963	-.805	-.360	-.049	.204	.426	.634	.841	1.060	1.324	1.885

TABLE 10 (S U I T E)



mètres de la distribution pour en déduire une réalisation de Y_T .

- Les limites de l'intervalle de confiance symétrique au niveau $(1 - p)$ sont données par les valeurs classées d'ordre $r \frac{p}{2}$ et $r(1 - \frac{p}{2})$, de la série simulée des r valeurs de Y_T .

Par exemple les limites inférieure et supérieure de l'intervalle de confiance à 90% ($p = .10$), en simulant $r = 200$ échantillons, sont la 10ème et la 190ème valeur classée.

Cette méthode doit être employée dans les cas exigeant une bonne précision car elle est assez longue, et pour les cas pratiques d'autres moyens peuvent être utilisés. Nous indiquons ici deux autres méthodes.

Il est possible d'utiliser les résultats du paragraphe (4.7) puisque Y_T est distribué asymptotiquement suivant une loi normale, d'erreur type $(E.T.)_{Y_T}$ que l'on sait calculer. Cette méthode doit conduire à une bonne approximation de l'intervalle de confiance si N est suffisamment grand.

GLADWELL et CHENG-NAN LIN (1969) ont déterminé les intervalles de confiance de la distribution normale standardisée en utilisant les statistiques d'ordre de cette loi; l'adaptation de cette technique au cas de la distribution Pearson III standardisée (BOBEE et MORIN 1973 b) est décrite ici.

Détermination des intervalles de confiance par les statistiques d'ordre.

On considère la distribution Pearson III standardisée, d'asymétrie positive $C_s = a > 0$. La fonction de distribution cumulée de l'évènement X_k d'ordre k dans un échantillon ordonné de taille N , $H(Z, k)$ peut être calculée pour la distribution Pearson III standardisée; les limites inférieure et supérieure de l'intervalle

de confiance au niveau $(1 - p)$ de l'évènement X_k sont Z_{1k} et Z_{2k} telles que :

$$H(Z_{1k}, k) = \frac{p}{2} \quad ; \quad H(Z_{2k}, k) = 1 - \frac{p}{2} \quad (4.52)$$

D'autre part la variable d'ordre k correspond à une probabilité empirique P_k , donc les points $A_{1k}(P_k, Z_{1k})$ et $A_{2k}(P_k, Z_{2k})$, lorsque k varie de 1 à N , permettent de tracer sur un papier de probabilité les limites inférieure et supérieure de l'intervalle de confiance au niveau $(1-p)$ de la loi Pearson III pour C_s et N connus. Les intervalles de confiance pour la forme à asymétrie négative (B) ($C_s = -a$), de fonction de distribution H_a , peuvent être déduits de ceux de la forme à asymétrie positive (A) ($C_s = +a$), en considérant la relation (4.50) (la taille de l'échantillon N étant fixée).

En effet, les limites inférieure et supérieure de l'intervalle de confiance au niveau $(1 - p)$ pour l'ordre $(N - k + 1)$ sont $Z'_{1, N - k + 1}$ et $Z'_{2, N - k + 1}$ telles que :

$$H_a(Z'_{1, N - k + 1} ; N - k + 1) = p/2 \quad (4.53)$$

$$H_a(Z'_{2, N - k + 1} ; N - k + 1) = 1 - p/2$$

On peut montrer à l'aide des relations (4.50); (4.52); (4.53) que :

$$Z'_{1, N - k + 1} = -Z_{2,k}$$

$$Z'_{2, N - k + 1} = -Z_{1,k}$$

Les limites inférieure et supérieure de l'intervalle de confiance correspondant à la probabilité empirique $P_{N - k + 1}$ sont :

$$B_{1, N - k + 1} \text{ de coordonnées } \left(P_{N - k + 1} ; Z'_{1, N - k + 1} \right)$$

$$B_{2, N - k + 1} \text{ de coordonnées } \left(P_{N - k + 1} ; Z'_{2, N - k + 1} \right)$$

Puisque de plus quelle que soit la forme de la probabilité empirique choisie on a $P_k + P_{N-k+1} = 1$, les points $B_{1, N-k+1}$ et $B_{2, N-k+1}$ sont respectivement les symétriques, par rapport à 0 ($P = .5$, $Z = 0$), des points $A_{2, k}$ et $A_{1, k}$ (figure 11).

Donc pour N, P , fixés la limite supérieure (inférieure) de la distribution Pearson III standardisée à asymétrie négative, $C_s = -a$, (forme B) est déduite de la limite inférieure (supérieure) de la distribution Pearson III standardisée à asymétrie positive, $(C_s) = -a$, (forme A).

Dans le cas particulier de la loi normale puisque $C_s = \left(C_s \right)_a = 0$, pour un niveau $(1-p)$ et une taille N donnés, la limite inférieure de l'intervalle de confiance peut être déduite de la limite supérieure par une symétrie par rapport à 0 ($P = .5$, $Z = 0$) comme le montrent les figures 12, 13, 14.

En utilisant les résultats précédents il est possible de déterminer les intervalles de confiance pour toute forme de la distribution Pearson III, puisque si l'on connaît la moyenne et la variance (après estimation des paramètres), la variable standardisée est calculable.

Des tables et graphiques ont été établis (BOBEE, MORIN 1972 b) qui permettent une détermination rapide des intervalles de confiance pour les cas rencontrés pratiquement, tout au moins en hydrologie.

Les calculs ont été effectués pour:

$$C_s = (0; 1.9) \text{ avec un pas de } 1$$

($C_s = 0$ correspond à la loi Normale)

$$N = 10, 20; 40; 60; 80; 100.$$

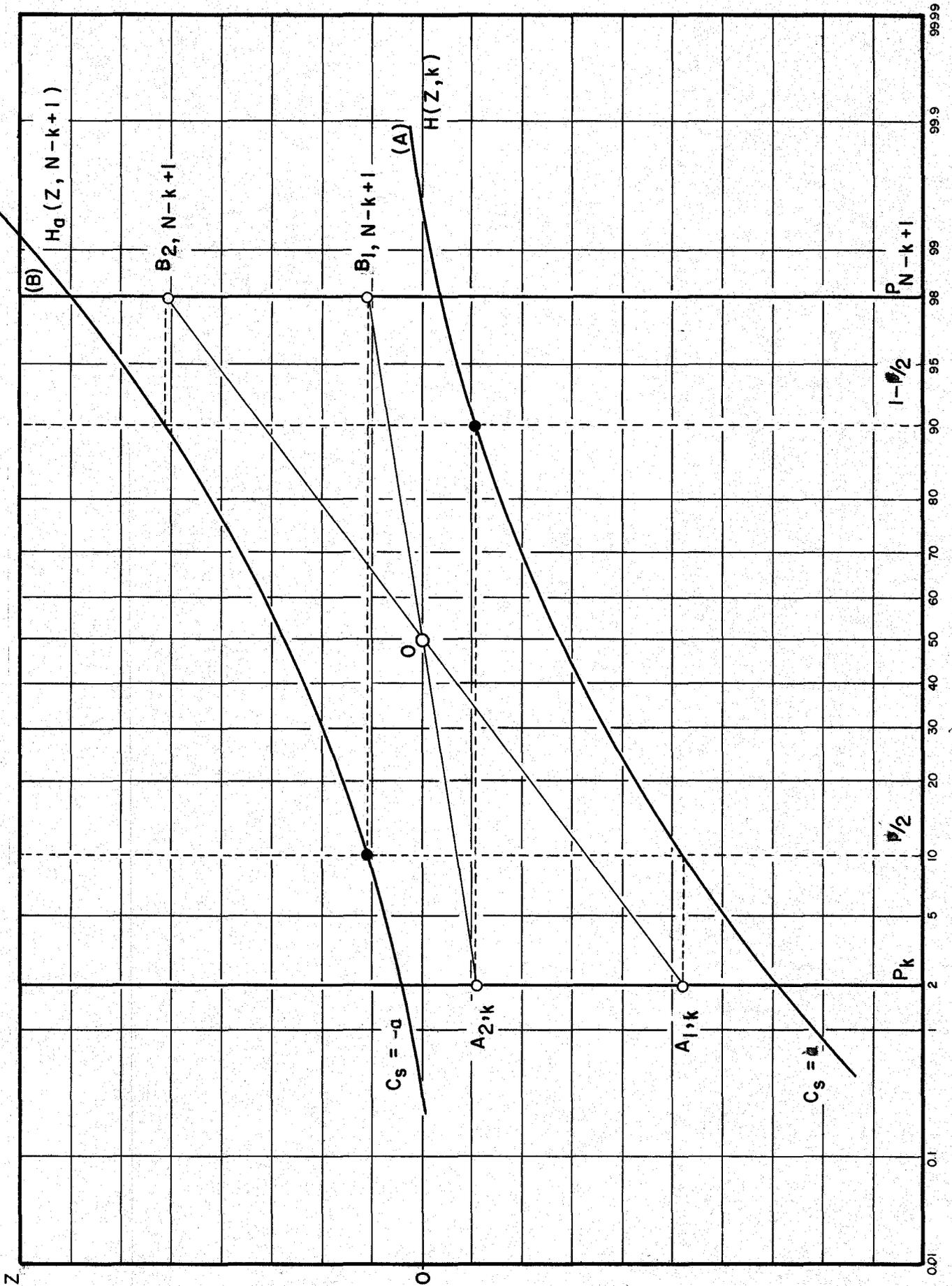


FIG - 11 . Relations entre les limites des intervalles de confiance.

Pour chaque couple (C_s, N) on considère les niveaux de confiance 95%, 90%, 80%.

Les tables correspondant aux valeurs négatives $C_s = 0, -.1, \dots, -1.9$, sont déduites par symétrie.

Les figures 12, 13, 14 et la table 11 donnent les intervalles de confiance déterminés par la méthode proposée pour $C_s = 0$ (Loi Normale) et les niveaux de confiance 95%, 90%, 80%.

La formule de probabilité empirique utilisée est celle de CHEGODAYEV (3.34), mais le programme de calcul permet d'utiliser toute autre formule pour P_k .

La méthode utilisée a donc permis l'établissement de tables et graphiques facilement utilisables, et, qui permettent de donner rapidement une bonne estimation de l'intervalle de confiance à un niveau fixé, pour un évènement de probabilité au dépassement donné. Lors de l'intégration numérique, certaines difficultés prévisibles ont été rencontrées pour les probabilités extrêmes surtout lorsque N est petit et que l'asymétrie devient élevée. En effet quelle que soit la méthode choisie, si N est petit, la détermination des intervalles de confiance relatifs à une période de retour T élevée est très imprécise sinon impossible.

Cependant, ces tables et graphiques ont été établis pour répondre aux cas fréquemment rencontrés en pratique, lorsque l'une des formes de la distribution Pearson III convient pour représenter une série de données.

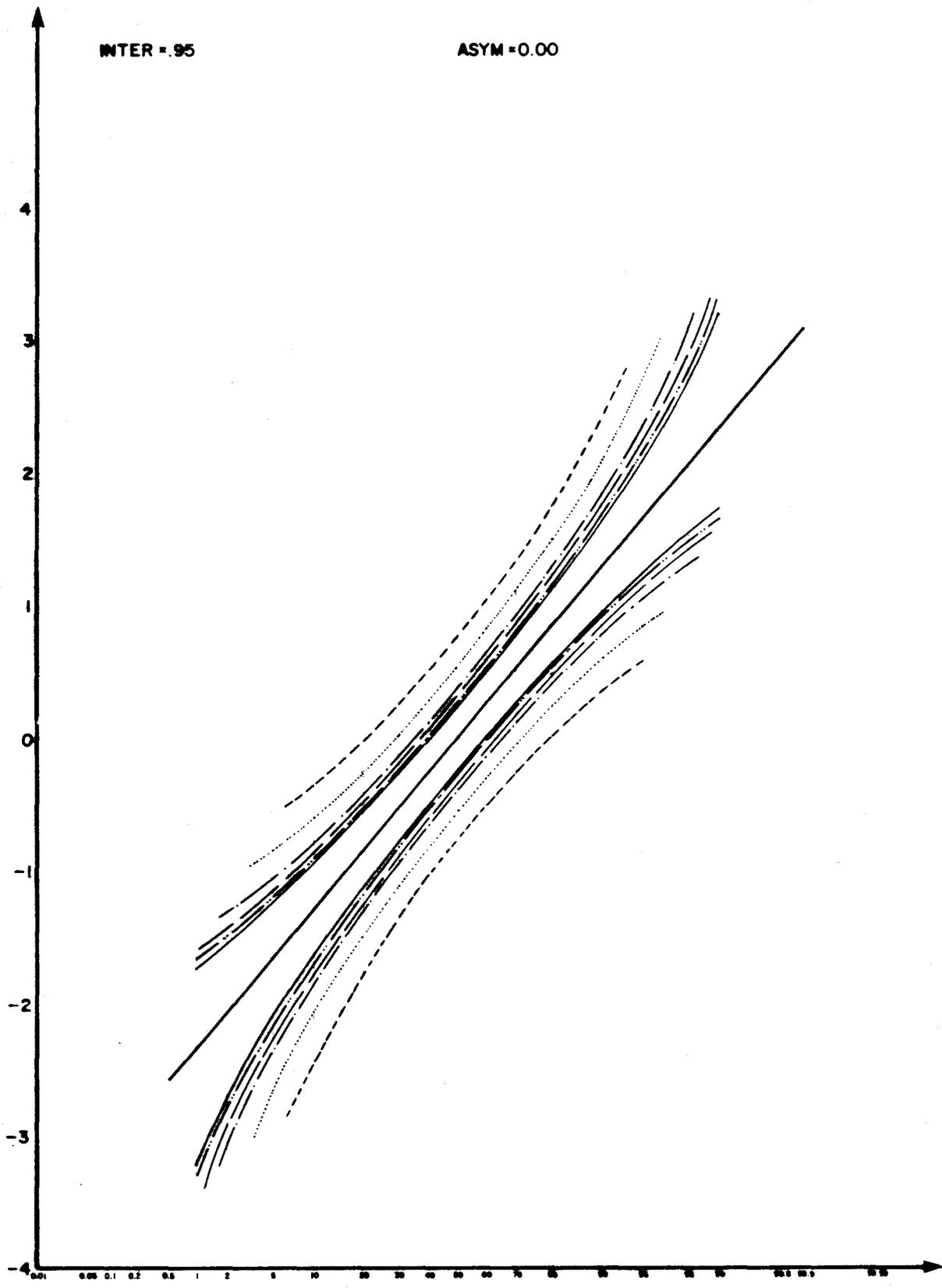


FIG 12 . Intervalle de confiance à 95 % ; loi Pearson III $C_3=0$

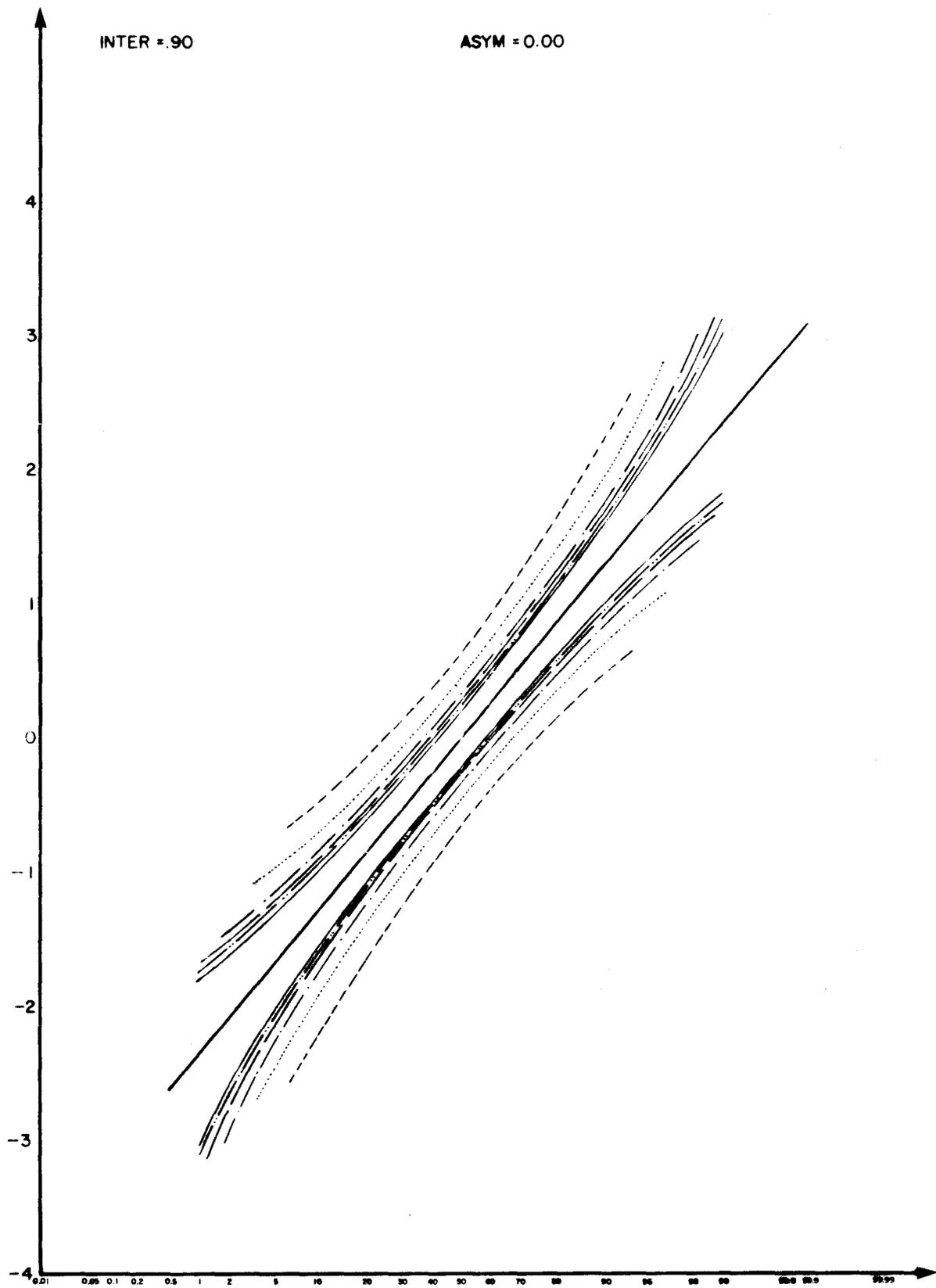


FIG 13 Intervalle de confiance à 90 % ; loi Pearson III $C_s = 0$

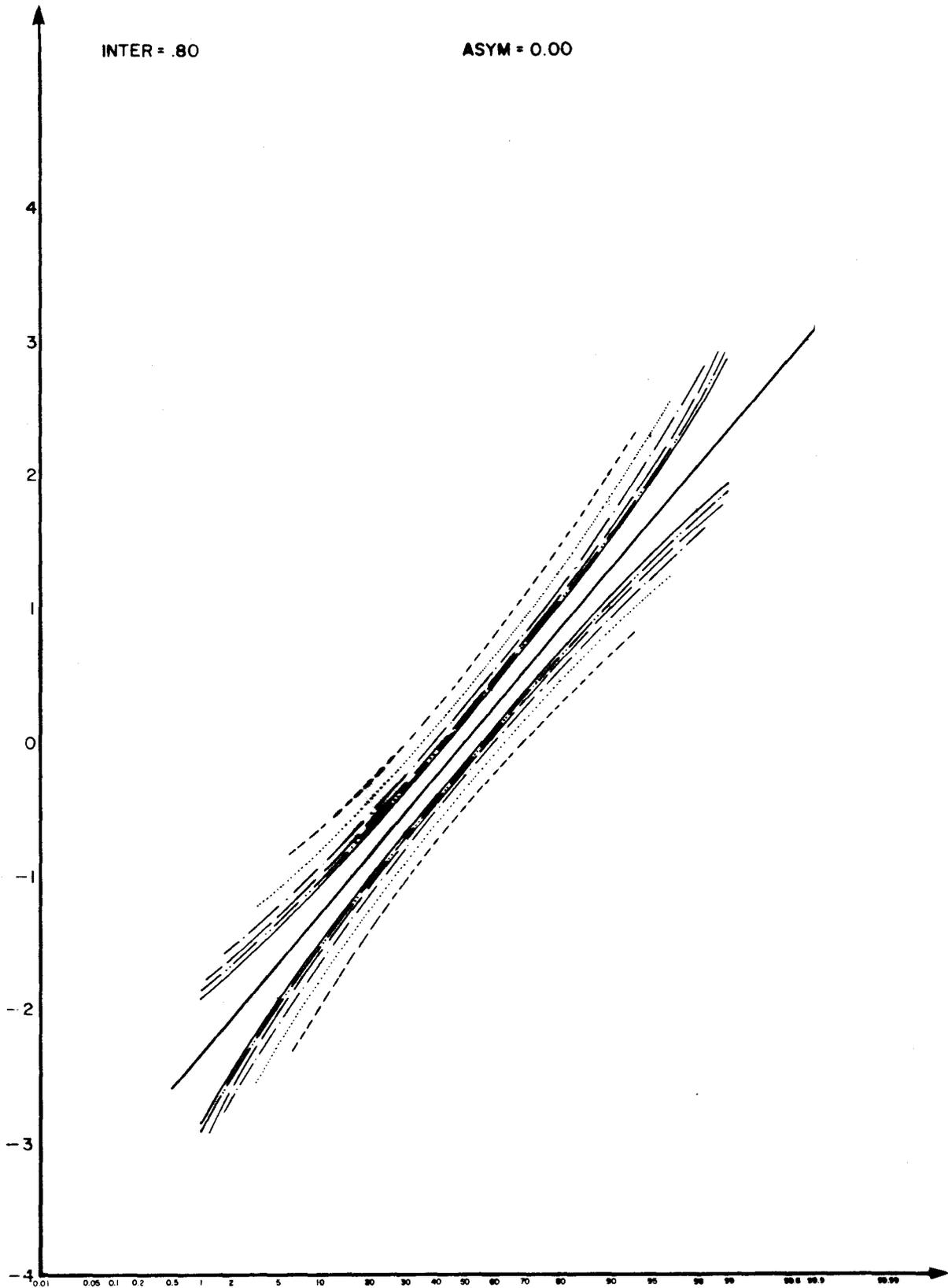


FIG 14 . Intervalle de confiance à 80 % ; loi Pearson III $C_s = 0$

TABLE 11: INTERVALLES DE CONFIANCE DE LA LOI PEARSON III STANDARDISEE

($C_s = 0$ Loi Normale)

Intervalles de confiance pour la loi Pearson III N = 10

Asymétrie = 0.00

Probabilité	Intervalle 95		Intervalle 90		Intervalle 80	
	.025	.975	.050	.950	.100	.900
.067	.500 2.804	.646 2.568	.821 2.309			
.100	.360 2.427	.499 2.225	.663 2.002			
.200	.027 1.761	.153 1.606	.302 1.431			
.300	-.248 1.351	-.126 1.214	.016 1.059			
.400	-.504 1.035	-.382 .907	-.242 .761			
.500	-.760 .760	-.637 .637	-.496 .496			
.600	-1.035 .504	-.907 .382	-.761 .242			
.700	-1.351 .248	-1.214 .126	-1.059 -.016			
.800	-1.161 -.027	-1.606 -.153	-1.431 -.302			
.900	-2.427 -.360	-2.225 -.499	-2.002 -.663			
.933	-2.804 -.500	-2.568 -.646	-2.309 -.821			

Intervalles de confiance pour la loi Pearson III N = 20

Asymétrie = 0.00

Probabilité	Intervalle 95		Intervalle 90		Intervalle 80	
	.025	.975	.050	.950	.100	.900
.034	.960 3.020	1.084 2.800	1.233 2.559			
.050	.855 2.688	.972 2.500	1.111 2.294			
.100	.604 2.091	.707 1.951	.830 1.796			
.200	.250 1.491	.342 1.383	.451 1.260			
.300	-.033 1.111	.055 1.015	.158 .985			
.400	-.290 .810	-.203 .719	-.102 .615			
.500	-.543 .543	-.456 .456	-.355 .355			
.600	-.810 .290	-.719 .203	-.615 .103			
.700	-1.111 .033	-1.015 -.055	-.905 -.158			
.800	-1.491 -.250	-1.383 -.342	-1.260 -.451			
.900	-2.091 -.604	-1.951 -.707	-1.796 -.830			
.950	-2.688 -.855	-2.500 -.972	-2.294 -1.111			
.966	-3.020 -.960	-2.800 -1.084	-2.559 -1.233			

TABLE 11: INTERVALLES DE CONFIANCE DE LA LOI PEARSON III STANDARDISEE (suite)

($C_s = 0$ Loi Normale)

Intervalles de confiance pour la loi Pearson III N = 40

Asymétrie = 0.00

Probabilité	Intervalle 95		Intervalle 90		Intervalle 80	
		.025	.975	.050	.950	.100
.017	1.352	3.224	1.460	3.016	1.590	2.791
.020	1.321	3.104	1.425	2.910	1.552	2.698
.050	1.062	2.373	1.150	2.245	1.254	2.104
.100	.788	1.846	.865	1.750	.954	1.643
.200	.415	1.298	.482	1.223	.561	1.137
.300	.125	.938	.189	.871	.262	.794
.400	-.134	.647	-.072	.584	-.001	.510
.500	-.386	.386	-.324	.324	-.252	.252
.600	-.647	.134	-.584	.072	-.510	.001
.700	-.938	-.125	-.871	-.189	-.794	-.262
.800	-1.298	-.415	-1.223	-.482	-1.137	-.561
.900	-1.846	-.788	-1.750	-.865	-1.643	-.954
.950	-2.378	-1.062	-2.245	-1.150	-2.104	-1.254
.980	-3.105	-1.321	-2.910	-1.426	-2.698	-1.552
.983	-3.324	-1.353	-3.016	-1.460	-2.791	-1.590

Intervalles de confiance pour la loi Pearson III N = 60

Asymétrie = 0.00

Probabilité	Intervalle 95		Intervalle 90		Intervalle 80	
		.025	.975	.050	.950	.100
.012	1.558	3.338	1.658	3.137	1.779	2.919
.020	1.436	2.906	1.526	2.750	1.634	2.581
.050	1.160	2.231	1.233	2.130	1.321	2.017
.100	.873	1.738	.937	1.662	1.011	1.575
.200	.491	1.213	.546	1.152	.611	1.082
.300	.197	.862	.249	.807	.309	.744
.400	-.064	.575	-.013	.523	.045	.463
.500	-.316	.316	-.265	.265	-.206	.206
.600	-.575	.064	-.523	.013	-.463	-.045
.700	-.862	-.197	-.807	-.249	-.744	-.309
.800	-1.213	-.491	-1.152	-.546	-1.082	-.611
.900	-1.738	-.873	-1.662	-.937	-1.575	-1.011
.950	-2.231	-1.160	-2.130	-1.233	-2.017	-1.321
.980	-2.906	-1.436	-2.750	-1.526	-2.581	-1.634
.988	-3.339	-1.558	-3.137	-1.658	-2.920	-1.779

TABLE 11: INTERVALLES DE CONFIANCE DE LA LOI PEARSON III STANDARDISEE (suite)

($C_s = 0$ Loi Normale)

Intervalles de confiance pour la loi Pearson III N = 80

Asymétrie = 0.00

Probabilité	Intervalle 95		Intervalle 90		Intervalle 80	
	.025	.975	.050	.950	.100	.900
.010	1.668	3.308	1.761	3.123	1.873	2.924
.020	1.508	2.784	1.588	2.652	1.684	2.508
.050	1.219	2.148	1.284	2.062	1.361	1.965
.100	.925	1.675	.981	1.609	1.046	1.535
.200	.536	1.162	.585	1.110	.641	1.050
.300	.240	.816	.285	.769	.337	.715
.400	-.022	.532	.022	.487	.073	.435
.500	-.274	.274	-.230	.230	-.179	.179
.600	-.532	.022	-.487	-.022	-.435	-.073
.700	-.816	-.240	-.769	-.285	-.715	-.337
.800	-1.162	-.536	-1.110	-.585	-1.050	-.641
.900	-1.675	-.925	-1.609	-.981	-1.535	-1.046
.950	-2.148	-1.219	-2.062	-1.284	-1.965	-1.361
.980	-2.784	-1.508	-2.652	-1.588	-2.508	-1.684
.990	-3.308	-1.668	-3.124	-1.761	-2.924	-1.873

Intervalles de confiance pour la loi Pearson III N = 100

Asymétrie = 0.00

Probabilité	Intervalle 95		Intervalle 90		Intervalle 80	
	.025	.975	.050	.950	.100	.900
.010	1.726	3.200	1.812	3.037	1.915	2.861
.020	1.559	2.700	1.632	2.585	1.720	2.458
.050	1.261	2.091	1.320	2.015	1.390	1.930
.100	.961	1.632	1.011	1.574	1.070	1.508
.200	.568	1.128	.611	1.081	.662	1.027
.300	.269	.785	.310	.743	.357	.694
.400	.006	.508	.046	.462	.092	.416
.500	-.245	.245	-.206	.206	-.160	.160
.600	-.508	-.006	-.462	-.046	-.416	-.092
.700	-.785	-.269	-.743	-.310	-.694	-.357
.800	-1.128	-.568	-1.081	-.611	-1.027	-.662
.900	-1.632	-.961	-1.574	-1.011	-1.508	-1.070
.950	-2.091	-1.261	-2.015	-1.320	-1.930	-1.390
.980	-2.701	-1.559	-2.585	-1.632	-2.458	-1.720
.990	-3.201	-1.726	-3.038	-1.812	-2.861	-1.915

5. PROPRIETES STATISTIQUES DE LA LOI LOG-PEARSON TYPE III

5.1 Moments de la distribution Log-Pearson III

On reprend dans ce chapitre les mêmes notations qu'au chapitre 2.

La fonction caractéristique présentant des difficultés d'intégration, la détermination des moments non centrés est effectuée directement à l'aide de la relation 3.1. Si $\left(\mu_r'\right)_L$ représente le moment non centré d'ordre r de la distribution Log-Pearson III, on a :

$$\left(\mu_r'\right)_L = \int_D x^r g(x) dx$$

où D est l'intervalle de variation défini en (2.1) et $g(x)$ est donné par la relation (2.1); en posant $\beta = \alpha k$ (avec $k = 1/\ln a$) on a :

$$\left(\mu_r'\right)_L = \int_D \frac{|\beta|}{\Gamma(\lambda)} e^{\frac{\beta m}{k}} x^{r-(1+\beta)} \left[\beta(\ln x - \frac{m}{k}) \right]^{\lambda-1} dx \quad (5.1)$$

a) Si $\left(1 - \frac{r}{\beta}\right) > 0$ (c'est-à-dire $\beta < 0$ ou $\beta > r > 0$)

effectuons le changement $v = \beta(\ln x - \frac{m}{k}) \left(1 - \frac{r}{\beta}\right)$

Le domaine de variation de v est D' , on a alors :

$$\left(\mu_r'\right)_L = \frac{|\beta|}{\beta} \frac{1}{\Gamma(\lambda)} \frac{e^{\frac{mr}{k}}}{\left(1 - \frac{r}{\beta}\right)^\lambda} \int_{D'} e^{-v} v^{\lambda-1} dv$$

$$\text{Soit } I = \int_{D'} e^{-v} v^{\lambda-1} dv$$

Si $\beta < 0$ on a $I = \int_{+\infty}^0 e^{-v} v^{\lambda-1} dv = -\Gamma(\lambda)$

donc:

$$\left(\mu'_r\right)_L = \frac{e^{mr/k}}{\left(1 - \frac{r}{\beta}\right)^\lambda} \quad (5.2)$$

Si $\beta > r > 0$ on a $I = \int_0^{+\infty} e^{-v} v^{\lambda-1} dv = \Gamma(\lambda)$

donc

$$\left(\mu'_r\right)_L = \frac{e^{mr/k}}{\left(1 - \frac{r}{\beta}\right)^\lambda} \quad (5.3)$$

b) Si $\left(1 - \frac{r}{\beta}\right) < 0$ (c'est-à-dire si $0 < \beta < r$) effectuons le changement $v = -\beta \left(\ln x - \frac{m}{k}\right) \left(1 - \frac{r}{\beta}\right)$

on a d'après (5.1):

$$\left(\mu'_r\right)_L = \frac{e^{rm/k}}{\Gamma(\lambda)} \frac{1}{\left(\frac{r}{\beta} - 1\right)^\lambda} \int_0^{+\infty} e^v v^{\lambda-1} dv$$

dans ce cas l'intégrale est divergente et $\left(\mu'_r\right)_L$ n'est pas défini.

c) Si $\beta = \alpha k$ est entier, on peut montrer que pour $r_0 = \beta$

$\left(\mu'_{r_0}\right)_L$ n'est pas défini.

Donc finalement :

- si $\beta = \alpha k < 0$ (c'est-à-dire $\alpha < 0$) les moments non centrés de la distribution log-Pearson III sont toujours définis;

- si $\beta = \alpha k > 0$ (c'est-à-dire $\alpha > 0$) les moments non centrés de la distribution Log-Pearson III n'existent que jusqu'à l'ordre $r < \beta = \alpha k$.

Dans le cas particulier de la loi Log-Gamma, ($m = 0$) les relations (5.2) et (5.3) deviennent:

$$\left(\mu'_r\right)_{LG} = \frac{1}{\left(1 - \frac{r}{\beta}\right)^\lambda} \quad (5.4)$$

où $\left(\mu'_r\right)_{LG}$ est le moment non centré d'ordre r de la loi Log-Gamma.

Il est possible de relier les moments $\left(\mu'_r\right)_L$ aux cumulants de la distribution Pearson III, la relation (5.2) ou (5.3) peut en effet s'écrire:

$$\text{Ln} \left(\mu'_r\right)_L = \frac{mr}{k} - \lambda \text{Ln} \left(1 - \frac{r}{\beta}\right)$$

Si $\frac{r}{|\beta|} < 1$ on peut remplacer $\text{Ln} \left(1 - \frac{r}{\beta}\right)$ par son développement, d'où:

$$\text{Ln} \left(\mu'_r\right)_L = \frac{mr}{k} + \lambda \sum_j \left(\frac{r}{\beta}\right)^j \cdot \frac{1}{j}$$

Si l'on fait intervenir les cumulants κ_j de la distribution Pearson III, on peut écrire d'après (4.2):

$$\text{Ln} \left(\mu'_r\right)_L = \sum_j \frac{1}{j!} \left(\frac{r}{k}\right)^j \kappa_j$$

ou encore:

$$\left(\mu'_r\right)_L = e^{\sum_{j=1} \frac{1}{j!} \left(\frac{r}{k}\right)^j \kappa_j} \quad (5.5)$$

La relation (5.5) présente surtout un intérêt quand $|\beta|$ (donc $|\alpha|$) est grand par rapport à r car alors les cumulants κ_j deviennent très petits quand r augmente.

En particulier si $|\beta| \gg r$ on a :

$$\left(\mu'_r\right)_L \approx e^{\frac{r}{k}} \kappa_1 = e^{\frac{r}{k}} \mu'_1 \quad (5.6)$$

où μ'_1 est la moyenne de la distribution Pearson III de paramètres λ, α, m .

Le moment centré d'ordre r , $\left(\mu_r\right)_L$ de la distribution Log-Pearson III est déterminé à partir de $\left(\mu'_r\right)_L$ à l'aide de la relation (3.3),

son expression devient vite compliquée lorsque r augmente, on a par exemple pour l'ordre 2 (variance):

$$\left(\mu_2\right)_L = e^{2m/k} \left[\frac{1}{\left(1 - \frac{2}{\beta}\right)^\lambda} - \frac{1}{\left(1 - \frac{1}{\beta}\right)^{2\lambda}} \right] \quad (5.7)$$

5.2 Détermination des principaux coefficients de la loi Log-Pearson III

Certains coefficients sont importants pour préciser les caractéristiques d'une distribution et présentent une grande utilité pour l'ajustement à une série observée.

En ce qui concerne la distribution Log-Pearson III nous considérons plus particulièrement les coefficients de variation $\left(C_v\right)_L$ et d'asymétrie $\left(C_s\right)_L$ définis par les relations générales (3.4) et (3.5),

dont l'expression est:

$$\left(C_v \right)_L = \sqrt{\left[\frac{\left(1 - 1/\beta \right)^2}{\left(1 - 2/\beta \right)} \right]^\lambda - 1} \quad (5.8)$$

$$\left(C_s \right)_L = \frac{\left[\frac{1}{\left(1 - \frac{3}{\beta} \right)^\lambda} - \frac{3}{\left(1 - \frac{2}{\beta} \right)^\lambda \left(1 - \frac{1}{\beta} \right)^\lambda} + \frac{2}{\left(1 - \frac{1}{\beta} \right)^3} \right]}{\left[\frac{1}{\left(1 - \frac{2}{\beta} \right)^\lambda} - \frac{1}{\left(1 - \frac{1}{\beta} \right)^{2\lambda}} \right]^{3/2}} \quad (5.9)$$

Ces coefficients sont fonctions de λ et β mais sont indépendants de m , il en résulte en particulier que les expressions (5.8) et (5.9) sont également valables pour la distribution Log - Gamma ($m = 0$). Les valeurs de $\left(C_v \right)_L$ et $\left(C_s \right)_L$ ont été calculées de manière systématique et tabulées en fonction de λ et β (BOBEE 1973b) pour étudier les relations entre ces 2 coefficients.

- Pour λ fixé, c'est-à-dire pour une valeur donnée du coefficient d'asymétrie C_s de la loi Pearson III (d'après 4.14), on obtient une courbe sur laquelle on se déplace lorsque α varie.

En faisant varier λ (ou C_s) on obtient un ensemble de courbes (fig. 15).

- Pour β fixé (donc pour $\alpha = \beta/k$) on obtient une courbe dont chaque point correspond à une valeur de λ . En faisant varier le paramètre β on obtient une famille de courbes (fig. 16)

Ces courbes permettent d'associer à tout couple $\left(\left(C_s \right)_L, \left(C_v \right)_L \right)$ un couple $\left(\beta, C_s \right)$ donc $\left(\alpha, \lambda \right)$; elles sont d'autre part valables quelle que soit la transformation logarithmique utilisée qui est caractérisée par $k = 1/\text{Ln } a$.

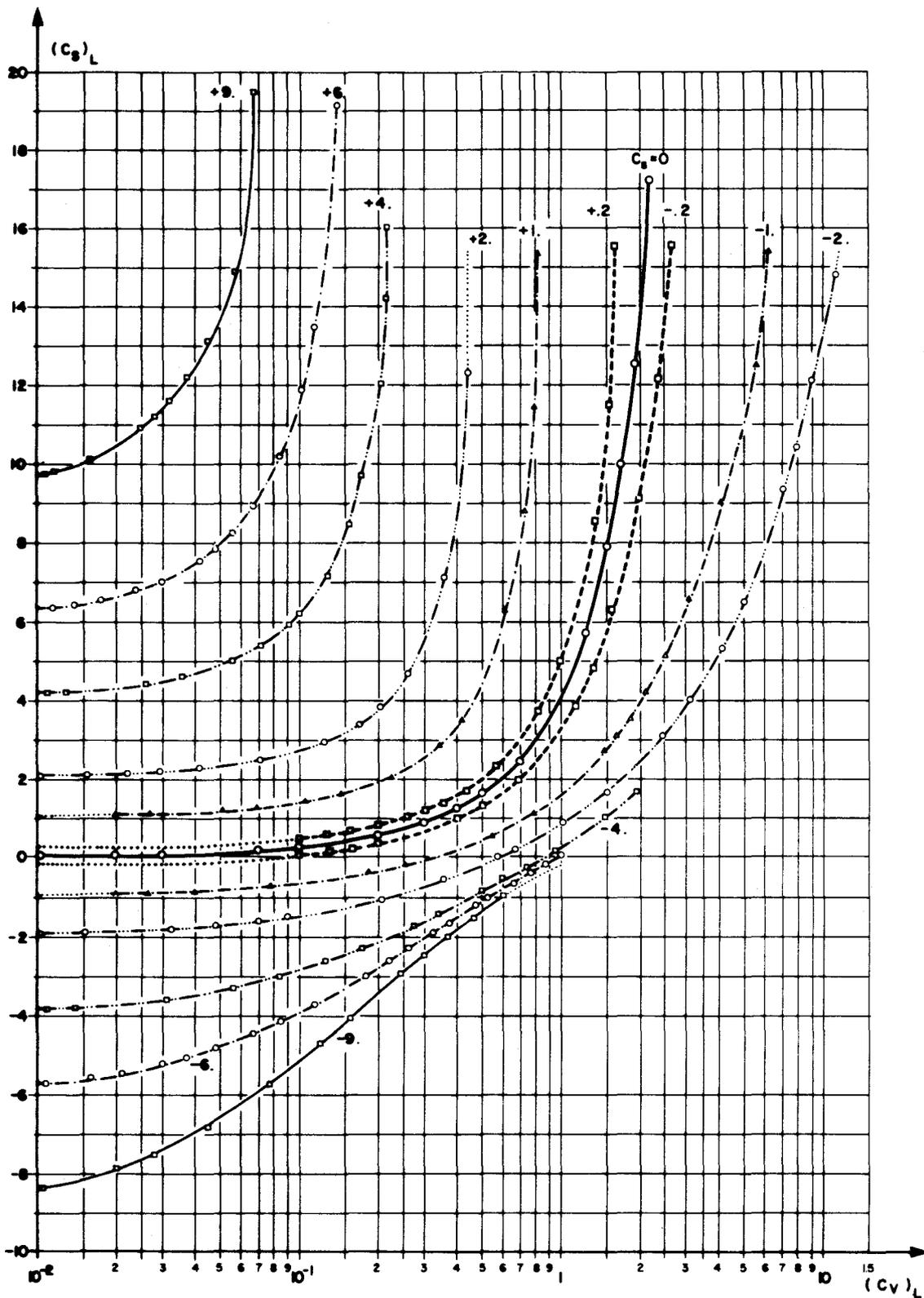


FIGURE 15: Relation entre $(C_S)_L$ et $(C_V)_L$ pour λ fixé.

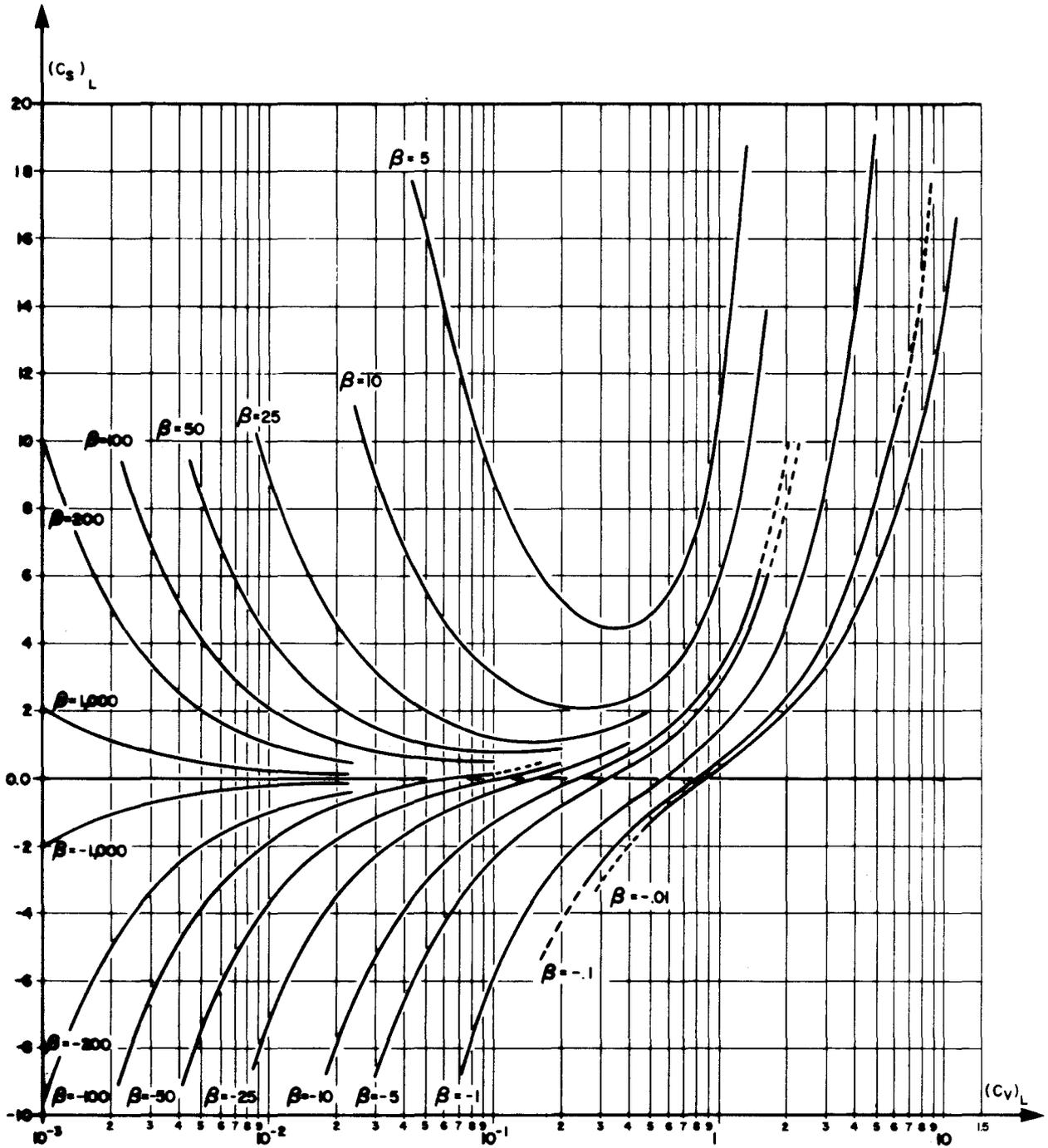


FIGURE 16 . Relation entre $(C_g)_L$ et $(C_v)_L$ pour β fixé.

Dans le cas particulier $C_s = 0$ (fig. 15), on obtient la distribution Log-Normale pour laquelle la relation entre le coefficient d'asymétrie et le coefficient de variation est (AITCHISON ET BROWN 1957):

$$\left(C_s\right)_L = 3\left(C_v\right)_L + \left[\left(C_v\right)_L\right]^3 \quad (5.10)$$

Sur un papier de probabilité Log-Normal, où les ordonnées sont en échelles logarithmiques et où les abscises représentent les probabilités au non dépassement, la courbe Log-Normale est représentée par une droite.

En ce qui concerne la distribution Log-Pearson III deux cas sont à considérer:

$$a) \quad \left(C_s\right)_L > 3\left(C_v\right)_L + \left[\left(C_v\right)_L\right]^3$$

Les courbes correspondantes sont telles que $C_s > 0$, donc α et β sont positifs et la représentation sur un papier Log-Normal donne une fonction de distribution convexe.

$$b) \quad \left(C_s\right)_L < 3\left(C_v\right)_L + \left[\left(C_v\right)_L\right]^3$$

Les courbes correspondantes sont telles que $C_s < 0$, donc α et β sont négatifs, la représentation sur un papier de probabilité Log-Normale donne une fonction de distribution concave. Cette fonction vers une asymptote horizontale pour les fortes probabilités au non-dépassement, puisque l'intervalle de variation de la variate Log Pearson III pour $\alpha < 0$ est borné supérieurement.

5.3 Comportement de $\left(C_v\right)_L$ et $\left(C_s\right)_L$

a) L'examen des figures (15) et (16) conduit à certaines conclusions concernant le comportement des coefficients $\left(C_v\right)_L$ et $\left(C_s\right)_L$:

- Si l'on considère la distribution Log-Pearson III correspondant à $\alpha > 0$, (donc associée à une distribution Pearson III à asymétrie positive) elle a toujours une asymétrie positive; par contre, si $\alpha < 0$, l'asymétrie de la distribution Log-Pearson III peut-être positive ou négative.

- Si λ est fixé (donc C_s fixé)

quand $|\beta|$ croit, $\left(C_v\right)_L$ et $\left(C_s\right)_L$ sont des fonctions décroissantes

et $\left(C_s\right)_L$ est une fonction croissante de $\left(C_v\right)_L$; d'autre part

on a toujours $\left(C_s\right)_L > C_s$ comme le montre la figure (15), donc

si l'on considère une population Log-Pearson III d'asymétrie $\left(C_s\right)_L$, la population correspondante des logarithmes a une asymétrie

$C_s < \left(C_s\right)_L$

- Si β est fixé

quand λ croit, $\left(C_v\right)_L$ est toujours croissante, mais $\left(C_s\right)_L$

est croissante si $\beta < 0$. Lorsque $\beta > 0$, et que λ croit $\left(C_s\right)_L$ est décroissante puis croissante.

Si $\beta < 0$, $\left(C_s\right)_L$ est une fonction croissante de $\left(C_v\right)_L$ comme

le montre la figure 16.

b) Il est également intéressant d'étudier le comportement asymptotique de $\left(C_v\right)_L$ et $\left(C_s\right)_L$ quand $|\beta|$ devient grand. Si l'on

remplace dans la relation (5.5) les cumulants par leur expression (4.3) on a:

$$\frac{\left(\mu'_r\right)_L}{e^{rm/k}} = e^{\lambda \left(\frac{r}{\beta} + \frac{r^2}{2\beta^2} + \frac{r^3}{3\beta^3} + \frac{r^4}{4\beta^4} + \dots \right)}$$

Si $|\beta|$ est élevé et si $\frac{\lambda r}{\beta}$ est petit, le coefficient de l'exponentielle est petit et on peut effectuer un développement limité, on obtient à l'ordre 4:

$$\frac{\left(\mu'_r\right)_L}{e^{rm/k}} = 1 + \frac{\lambda r}{\beta} + \frac{\lambda(\lambda+1)r^2}{2\beta^2} + \frac{\lambda r^3}{6\beta^3} (2+3\lambda+\lambda^2) + \frac{\lambda r^4}{4\beta^4} \left(1 + \frac{11}{6}\lambda + \lambda^2 + \frac{\lambda^3}{6}\right)$$

Pour obtenir le développement de $\left[\frac{\left(\mu'_r\right)_L}{e^{rm/k}}\right]^p$, p étant un nombre entier positif, on remplace dans le développement précédent λ par λp .

A partir de la relation (3.3), il est possible d'obtenir le développement limité à l'ordre 4 des moments centrés de la loi Log Pearson III lorsque $|\beta|$ est élevé, et d'en déduire le développement de $\left(C_v\right)_L$ et $\left(C_s\right)_L$.

Comportement de $\left(C_v\right)_L$ quand $|\beta|$ est grand

Quand $|\beta|$ est élevé on obtient:

$$\left(C_v\right)_L \sim \frac{\sqrt{\lambda}}{|\beta|} \left(1 + \frac{1}{\beta}\right) \quad (5.11)$$

Cette relation montre que:

- Si $\beta > 0$, $\left(C_v\right)_L$ tend vers $\frac{\sqrt{\lambda}}{\beta}$ par valeurs supérieures

lorsque β croît et que $\left(\left(C_v\right)_L - \frac{\sqrt{\lambda}}{\beta}\right) \sim \frac{\sqrt{\lambda}}{\beta^2}$

- Si $\beta < 0$, $\left(c_v\right)_L$ tend vers $-\frac{\sqrt{\lambda}}{\beta}$ par valeurs inférieures lorsque $|\beta|$ croît, (donc lorsque β décroît) et que

$$\left[\left(c_v\right)_L - \left(-\frac{\sqrt{\lambda}}{\beta}\right)\right] \sim -\frac{\sqrt{\lambda}}{\beta^2}$$

Comportement de $\left(c_s\right)_L$ quand $|\beta|$ est grand

En procédant comme précédemment on peut montrer que lorsque $|\beta|$ est élevé et $\frac{\lambda}{|\beta|}$ petit

$$\left(c_s\right)_L \sim \frac{2}{\sqrt{\lambda}} \frac{|\beta|}{\beta} \left(1 + \frac{3}{2} \frac{\lambda+1}{\beta}\right) \quad (5.12)$$

Donc quand $\frac{\lambda}{|\beta|}$ est petit, $\left(c_s\right)_L$ tend vers $\frac{2}{\sqrt{\lambda}} \frac{|\beta|}{\beta}$,

c'est-à-dire vers le coefficient d'asymétrie de la loi Pearson III.

D'autre part $\left[\left(c_s\right)_L - \frac{2}{\sqrt{\lambda}} \frac{|\beta|}{\beta}\right] \sim \frac{3(\lambda+1)}{\sqrt{\lambda}} \frac{|\beta|}{\beta^2} = \frac{3(\lambda+1)}{\sqrt{\lambda}|\beta|} > 0$

Donc:

Si $\beta > 0$, $\left(c_s\right)_L$ tend par valeurs supérieures vers $c_s = \frac{2}{\sqrt{\lambda}}$

Si $\beta < 0$, $\left(c_s\right)_L$ tend par valeurs supérieures vers $c_s = -\frac{2}{\sqrt{\lambda}}$

5.4 Lien entre les distributions Pearson III et Log-Pearson III

Si la variate x est distribuée suivant la loi Log-Pearson III $y = \log_a x$ est distribuée suivant Pearson III. Les résultats établis pour la loi Pearson III dans le cas des erreurs-types (4.6, 4.7) et des intervalles de confiance (4.10), peuvent être utiles pour la variate y (logarithme de l'évènement), il est

ensuite possible d'en déduire les résultats concernant x . En effet, pour une transformation logarithmique de base a , on a la relation:

$$\frac{\Delta X}{X} \approx \frac{1}{2} \left[a^{\Delta Y} - a^{-\Delta Y} \right] \quad (5.13)$$

où:

ΔY est un accroissement en unités logarithmiques

$\frac{\Delta X}{X}$ l'accroissement relatif correspondant à ΔY

Donc à une erreur type (E.T.) = ΔY exprimée en unité logarithmique, il est possible d'associer l'erreur relative $\frac{\Delta X}{X}$.

L'étude de la distribution Log-Pearson III est cependant nécessaire pour mettre en évidence certaines de ses propriétés mathématiques et statistiques.

L'intervalle de définition de la variate x , le signe du coefficient d'asymétrie $\left(C_s \right)_L$, la relation existant entre $\left(C_s \right)_L$ et $\left(C_v \right)_L$ sont en effet autant d'éléments qui guident dans le choix d'une distribution susceptible de représenter un phénomène donné. La table 12 récapitule ces informations pour les distributions Pearson III, Log-Pearson III et leurs formes simplifiées.

En ce qui concerne la loi Log-Pearson III, les résultats ont été établis en considérant la transformation logarithmique sous sa forme générale dans une base a .

En pratique les transformations que l'on peut utiliser sont celles de

- base 10: (logarithmes décimaux) on a alors

$$k = 1/\text{Ln } 10 \approx .434$$

- base e : (logarithmes népériens) on a alors

$$k = 1$$

Loi	Paramètres	Intervalle de la variate	Asymétrie	Particularités	
Pearson III (A)	$\alpha > 0, m$ quelc.	$(m, +\infty)$	$C_s = \frac{2}{\sqrt{\lambda}} > 0$	$C_s = 2 C_v \left(1 + m \frac{\lambda}{\alpha}\right)$	
Pearson III (B)	$\alpha < 0, m$ quelc.	$(-\infty, m)$	$C_s = -\frac{2}{\sqrt{\lambda}} < 0$		
Gamma (A)	$\alpha > 0, m = 0$	$(0, +\infty)$	$C_s = \frac{2}{\sqrt{\lambda}} > 0$	$C_s = 2 C_v$	
Gamma (B)	$\alpha < 0, m = 0$	$(-\infty, 0)$	$C_s = -\frac{2}{\sqrt{\lambda}} < 0$		
Normale	$\lambda \rightarrow \infty$	$(-\infty, +\infty)$	$C_s = 0$		
Log-Pearson (A)	$\alpha > 0, m$ quelc.	$(e^{m/k}, \infty)$	$(C_s)_L > 0$	$(C_s)_L > 3(C_v)_L + [(C_v)_L]^3$	$\beta = \alpha k$ $k = 1/\text{Ln } a$ • <u>Log-décimal</u> (a = 10) $k \approx .434$ • <u>Log Népérien</u> (a = e) $k = 1$
Log-Pearson (B)	$\alpha < 0, m$ quelc.	$(0, e^{m/k})$	$(C_s)_L$ quelconque	$(C_s)_L < 3(C_v)_L + [(C_v)_L]^3$	
Log-Gamma (A)	$\alpha > 0, m = 0$	$(1, \infty)$	$(C_s)_L > 0$	$(C_s)_L > 3(C_v)_L + [(C_v)_L]^3$	
Log-Gamma (B)	$\alpha < 0, m = 0$	$(0, 1)$	$(C_s)_L$ quelconque	$(C_s)_L < 3(C_v)_L + [(C_v)_L]^3$	
Log-Normale	$\lambda \rightarrow \infty$	$(0, \infty)$	$(C_s)_L > 0$	$(C_s)_L = 3(C_v)_L + [(C_v)_L]^3$	

TABLE 12 : CARACTERISTIQUES GENERALES DE DISTRIBUTION PEARSON III ET LOG-PEARSON III

6. ESTIMATION DES PARAMETRES D'UNE DISTRIBUTION

6.1 Introduction

Lorsque l'on dispose d'une série d'observations que l'on suppose tirée d'une population mère donnée, on doit en pratique estimer à partir des valeurs de l'échantillon, les paramètres de la distribution et leur variance d'estimation.

Pour une fonction de distribution représentative d'un échantillon, il existe en général plusieurs méthodes de détermination des paramètres plus ou moins faciles à mettre en oeuvre et conduisant à des estimations plus ou moins bonnes.

Dans le but de comparer les différentes méthodes d'estimation il est nécessaire de caractériser les estimateurs auxquels chaque méthode conduit. Nous rappellerons donc les propriétés générales des estimateurs et les principales méthodes d'estimation des paramètres d'une distribution en utilisant les ouvrages de KENDALL (1967), CRAMER (1971), RAO (1952) et DUMAS de RAULY (1968).

6.2 Estimateurs corrects et absolument corrects

a. Estimateurs corrects

Si l'on considère un échantillon de taille N , x_1, \dots, x_N , on peut déterminer un estimateur $t_N = \phi(x_1, \dots, x_N)$, pour un échantillon donné on obtient une valeur de t_N qui est l'estimation cherchée, l'estimateur t_N est donc une variable aléatoire.

L'estimateur t_N du paramètre θ est correct (ou convergent) s'il converge en probabilité vers la vraie valeur θ c'est-à-dire si quels que soient ϵ et η arbitrairement petits, on peut trouver un nombre entier p tel que pour tout $N > p$:

$$\Pr \left[\left| t_N - \theta \right| < \epsilon \right] > 1 - \eta \quad (6.1)$$

En pratique si:

$$E \left(t_N \right) = \theta + f(N)$$
$$\text{var} \left(t_N \right) = g(N)$$

Il est possible de montrer d'après le théorème de Bienaymé-Tchebychev que si:

$$\lim_{N \rightarrow \infty} g(N) = 0 \quad \text{et} \quad \lim_{N \rightarrow \infty} f(N) = 0$$

alors t_N est un estimateur correct de θ .

b. Estimateurs absolument corrects

Un estimateur correct peut être biaisé, on a alors $E \left(t_N \right) \neq \theta$.

Un estimateur correct et non biaisé sera dit absolument correct.

L'obtention d'un estimateur absolument correct est souhaitable, surtout lorsque l'on considère des échantillons de petite taille, mais on peut être amené à préférer des estimateurs corrects et biaisés possédant d'autres propriétés.

6.3 Estimateurs efficaces

Il est possible de trouver plusieurs estimateurs absolument corrects d'un même paramètre, ces estimateurs de même moyenne ont, en général, des variances différentes. On choisit alors l'estimateur qui a la plus petite variance, car sa distribution est moins dispersée autour de la valeur vraie du paramètre. Il est cependant possible de montrer qu'il existe une valeur limite de la variance.

a. Variance minimum d'un estimateur

On considère la fonction densité de probabilité $f(x, \theta)$ où θ est le seul paramètre à déterminer, la fonction vraisemblance est définie par:

$$L = L \left(x_1, x_2, \dots, x_N; \theta \right) = f \left(x_1, \theta \right) \cdot f \left(x_2, \theta \right) \dots \dots \dots \cdot f \left(x_N, \theta \right)$$

L est aussi la densité de la distribution jointe des observations.

Soit $t = \phi(x_1, \dots, x_N)$ un estimateur non biaisé de la fonction $\psi(\theta)$ du paramètre θ , (on a $E(t) = \psi(\theta)$ puisque t est non biaisé).

Si de plus:

- . l'intervalle de définition de la variate x est indépendant de θ
- . les deux premières dérivées de L par rapport à θ existent
- . la dérivée de la fonction ψ par rapport à θ existe.

On peut alors montrer que $\text{var}(t)$ est toujours supérieur ou égal à une limite qui est la variance minimum possible

$$\text{var } t \geq \left(\frac{d\psi}{d\theta} \right)^2 / I \quad (6.2)$$

$$\text{avec } I = \text{var} \left(\frac{1}{L} \frac{\partial L}{\partial \theta} \right) = E \left(\frac{\partial \text{Log } L}{\partial \theta} \right)^2 = - E \left(\frac{\partial^2 \text{Log } L}{\partial \theta^2} \right) \quad (6.3)$$

I est la quantité d'information du paramètre θ .

La limite est indépendante de la méthode d'estimation choisie et tout estimateur non biaisé conduisant à cette limite est le meilleur estimateur non biaisé. On dit alors que t est un estimateur efficace.

- . Dans le cas particulier où t est un estimateur non biaisé de θ (on a alors $\psi(\theta) = \theta$), la relation (6.2) devient:

$$\text{var } t \geq \frac{1}{I} \quad (6.4)$$

- . Dans le cas où t est un estimateur biaisé de θ tel que:

$E(t) = \theta + b(\theta) = \psi(\theta)$, on peut montrer que:

$$E \left[(t - \theta)^2 \right] \geq \left[b(\theta) \right]^2 + \left[\frac{1 + db/d\theta}{I} \right]^2 \quad (6.5)$$

b. Forme de la fonction vraisemblance, pour un estimateur efficace

Il est possible de montrer que la condition nécessaire et suffisante pour que la variance de l'estimateur t atteigne la valeur minimum est que la fonction vraisemblance soit de la forme:

$$L = L_1 \exp (t \theta_1 + \theta_2) \quad (6.6)$$

où:

L_1 et t sont fonction des observations (x_1, \dots, x_N) seulement
 θ_1 et θ_2 sont fonction de θ seulement

avec:

$$E (t) = \psi(\theta) = - \frac{d\theta_2}{d\theta_1} \quad \text{var } t = - \frac{d^2\theta_2}{d\theta_1^2} \quad (6.7)$$

La condition (6.6) peut encore s'écrire:

$$\frac{\partial \text{Log} L}{\partial \theta} = A (\theta) \left[t - \psi(\theta) \right] \quad (6.8)$$

où:

A est une fonction de θ seulement $\left(A = \frac{d\theta_1}{d\theta} \right)$

et l'on a:

$$\text{var } t = \left| \frac{d\psi/d\theta}{A(\theta)} \right| \quad (6.9)$$

Si donc la fonction de vraisemblance vérifie (6.6) ou (6.8) on est certain de pouvoir trouver un estimateur efficace t de $\psi(\theta)$ et les relations précédentes permettent de déterminer $\psi(\theta)$ à partir de la forme de L . La variance minimum sera atteinte pour une seule fonction $\psi(\theta)$.

c. Efficacité d'une estimation

L'efficacité absolue d'un estimateur t est définie par:

$$e (\theta) = \frac{1}{I \text{ var}(t)} \quad (6.10)$$

Si t est un estimateur efficace de θ on a $e(\theta) = 1$

L'efficacité relative de l'estimateur t , par rapport à l'estimateur t' est définie par

$$e_{t/t'} = \frac{E[(t' - \theta)^2]}{E[(t - \theta)^2]}$$

Lorsque t et t' sont non biaisés on a:

$$e_{t/t'} = \frac{\text{var } t'}{\text{var } t} \quad (6.11)$$

Dans certains cas, on définit l'efficacité relative par le rapport des tailles d'échantillons conduisant à des variances égales. Souvent, lorsqu'il n'existe pas d'estimateur efficace pour une taille N donnée d'échantillon, la variance minimum est atteinte lorsque N devient infini; on dit alors que l'estimateur considéré est asymptotiquement efficace.

6.4 Statistique exhaustive

Nous nous plaçons encore ici dans le cas d'un seul paramètre à déterminer.

a. Définition

En général, il est possible de déterminer plusieurs estimateurs d'un paramètre θ résumant plus ou moins bien l'information contenue dans un échantillon.

Une statistique T est exhaustive si elle contient toute l'information de l'échantillon relative au paramètre θ . On montre que la condition nécessaire et suffisante pour que T soit exhaustive pour le paramètre θ est que la fonction vraisemblance soit de la forme:

$$L(x_1, \dots, x_N; \theta) = g(T; \theta) k(x_1, \dots, x_N) \quad (6.12)$$

où:

$g(T, \theta)$ dépend uniquement de T et θ

k est une fonction des observations et ne dépend pas de θ

b. Rapport avec la variance minimum

Si l'on dérive (6.12) par rapport à θ on a :

$$\frac{\partial \text{Log } L}{\partial \theta} = \frac{\partial g(T; \theta)}{\partial \theta} \quad (6.13)$$

On peut alors voir que si la relation (6.8), (condition nécessaire et suffisante pour avoir un estimateur efficace) est vraie alors (6.13) l'est aussi, la réciproque est fautive. Donc un estimateur efficace est exhaustif alors qu'un estimateur exhaustif peut ne pas être efficace.

On peut montrer que lorsque T est exhaustif pour θ , toutes les fonctions biunivoques de T sont exhaustives pour θ .

Parmi cette famille de fonctions il existe une seule fonction $t_0(T)$ estimateur non biaisé de $\psi(\theta)$ qui peut atteindre la variance minimum si elle existe.

On recherchera donc l'estimateur efficace lorsqu'il existe parmi les fonctions non biaisées de la statistique exhaustive T . En particulier si il y a une seule fonction non biaisée de T et si un estimateur efficace existe, ce sera cette fonction. On peut de plus montrer que :

- Une condition nécessaire pour qu'une distribution possède une statistique exhaustive T est qu'elle soit de la forme :

$$f(x, \theta) = \exp [A(\theta) B(x) + C(x) + D(\theta)] \quad (6.14)$$

où :

$C(x)$ et $B(x)$ sont fonction seulement des observations

$$\left(x_1 \dots \dots x_N \right)$$

$A(\theta)$ et $D(\theta)$ sont seulement fonction du paramètre à estimer.

- Si de plus l'intervalle de la variate x est indépendant de θ la condition (6.14) est suffisante.

- Il existe toujours une fonction $t_0(T)$ de la statistique exhaustive T de θ qui est le meilleur estimateur de $E(t_0) = \psi(\theta)$

6.5 Cas où l'intervalle de définition dépend d'un paramètre

Nous considérons une fonction densité $f(x, \theta)$ l'intervalle de définition de la variate x étant fonction du paramètre à estimer θ . Nous nous intéressons ici au cas où une seule des bornes dépend de θ , par exemple la borne inférieure on a alors:

$$a(\theta) \leq x \leq b.$$

Il est possible de montrer que:

- Si $f[a; \theta] = 0$ la relation donnant la variance minimum reste vraie :

$$\text{var } t \geq \frac{\left(\frac{d\psi}{d\theta}\right)^2}{E\left[\left(\frac{\partial \text{Log} L}{\partial \theta}\right)^2\right]}$$

- Si de plus $\left[\frac{\partial f(x; \theta)}{\partial \theta}\right]_{x=a} = 0$

$$\text{on a: } E\left[\left(\frac{\partial \text{Log} L}{\partial \theta}\right)^2\right] = -E\left(\frac{\partial^2 \text{Log} L}{\partial \theta^2}\right)$$

et les relations (6.2) et (6.3) restent vraies.

- Si $a(\theta)$ est une fonction monotone, on peut montrer que s'il existe une statistique exhaustive, ce ne peut être que x_1 , et x_1 est exhaustif si $f(x, \theta)$ est de la forme:

$$f(x, \theta) = g(x) / h(\theta) \tag{6.15}$$

Les résultats précédents se transposent aisément au cas où la borne supérieure seule est fonction de θ .

6.6 Estimation de plusieurs paramètres

Soit $f(x; \theta_1, \dots, \theta_k)$, une fonction densité dépendante de k paramètres que l'on estime conjointement.

Si la fonction vraisemblance L peut se mettre sous la forme:

$$L = g(T_1, \dots, T_r; \theta_1, \dots, \theta_k) * k(x_1, \dots, x_N) \quad (6.16)$$

on dit que T_1, \dots, T_r sont conjointement exhaustifs pour les paramètres $\theta_1, \dots, \theta_k$. Ceci n'implique pas que T_1 par exemple est individuellement exhaustif pour θ_1 . De même si T_1, T_2, \dots sont individuellement exhaustifs pour $\theta_1, \theta_2, \dots$, cela n'implique pas que T_1, T_2, \dots sont conjointement exhaustifs.

Il est possible de montrer (sous les hypothèses de régularité décrites en 6.3 a) que pour $r \leq k$ l'estimation de r fonctions $\psi_i(\theta_1, \dots, \theta_k)$ par des fonctions non biaisées t_i des statistiques T_1, \dots, T_r sont telles que:

$$\text{var } t_i \geq \sum_{j=1}^k \sum_{l=1}^k \frac{\partial \psi_i}{\partial \theta_j} \frac{\partial \psi_i}{\partial \theta_l} I_{jl}^{-1} \quad (6.17)$$

I_{jl}^{-1} étant le terme de la ligne j et de la colonne l de l'inverse de la matrice d'information I définie par:

$$I_{jl} = E \left[\frac{\partial \text{Log} L}{\partial \theta_j} \frac{\partial \text{Log} L}{\partial \theta_l} \right] \quad (6.18)$$

La relation (6.17) est la généralisation de (6.2).

On montre par ailleurs qu'une fonction densité dépendant de k paramètres (l'intervalle de définition de la variate étant indépendant des paramètres) possède une série de k statistiques conjointement exhaustives pour les k paramètres si elle peut se mettre sous la forme:

$$f(x; \theta_1, \dots, \theta_k) = \exp \left[D(\theta_1, \dots, \theta_k) + C(x) + \sum_{j=1}^k A_j(\theta_1, \dots, \theta_k) B_j(x) \right] \quad (6.19)$$

où: les fonctions B_j et C sont indépendantes des paramètres
les fonctions D et A_j ne dépendent que des k paramètres.

6.7 Méthode du Maximum de Vraisemblance

a. Principe de la méthode

Lorsque nous considérons la fonction vraisemblance $L(x_1, \dots, x_N; \theta)$, pour un échantillon donné, cette fonction ne dépend que de θ . La méthode du maximum de vraisemblance (M.V.) consiste à déterminer θ pour que L (ou ce qui est équivalent $\text{Log}L$) soit maximum étant donné l'échantillon (x_1, \dots, x_N) ce qui se traduit par:

$$\frac{\partial \text{Log}L}{\partial \theta} = 0 \quad \text{avec} \quad \frac{\partial^2 \text{Log}L}{\partial \theta^2} < 0 \quad (6.20)$$

Si l'on considère une distribution à k paramètres $\theta_1, \dots, \theta_k$ on a les conditions:

$$\frac{\partial \text{Log}L}{\partial \theta_j} = 0 \quad j = 1, \dots, k \quad (6.21)$$

avec la condition que la matrice de terme général:

$$a_{ij} = \left(\frac{\partial^2 \text{Log}L}{\partial \theta_i \partial \theta_j} \right) \text{ soit définie négative}$$

b. Propriétés de la méthode dans le cas d'un seul paramètre

- Les estimateurs par le M.V. sont corrects quoiqu'en général biaisés (ils ne sont donc pas absolument corrects).

- Si une statistique exhaustive T du paramètre θ existe
(relation 6.12), la méthode du M.V. conduit à une solution unique
 θ_* fonction de T .

D'après les résultats de (6.4b), il suffira de rechercher la fonction $t(\theta_*)$ estimateur non biaisé de $\psi(\theta)$. S'il existe un estimateur efficace de θ , c'est alors la fonction $t(\theta_*)$ dont la variance est minimum, et est donnée par la relation (6.2).

Donc s'il existe un estimateur efficace non biaisé il est donné par la méthode du maximum de vraisemblance.

- Dans le cas où une borne de l'intervalle de définition de la variate est fonction du paramètre à déterminer θ , si la relation (6.15) est vérifiée on sait qu'il est possible de trouver une statistique exhaustive, et les conclusions précédentes s'appliquent. On peut alors montrer que dans la relation (6.15), on doit déterminer θ pour minimiser $h(\theta) = \int g(x) dx$

- S'il n'existe pas de statistique exhaustive de θ on peut montrer que lorsque:

- les dérivées première et seconde de L existent

- $E \left(\frac{\partial \text{Log} L}{\partial \theta} \right) = 0$

- $R^2(\theta) = - E \left(\frac{\partial^2 \text{Log} L}{\partial \theta^2} \right) = E \left[\left(\frac{\partial \text{Log} L}{\partial \theta} \right)^2 \right]$ est définie et non nulle

pour tout θ

Alors la solution θ_* de l'équation du maximum de vraisemblance est asymptotiquement normale avec:

- une moyenne θ_0 , θ_0 étant la valeur vraie de θ
- une variance asymptotique $\frac{1}{R^2(\theta_0)}$

Puisque la variance asymptotique de θ_* tend vers la limite d'un estimateur efficace, on dit que dans ce cas la solution du M.V. est asymptotiquement efficace.

Dans le cas particulier où une borne de l'intervalle de définition de la variate dépend du paramètre θ et si il n'existe pas de statistique exhaustive (relation 6.15 non vérifiée), on peut montrer que la solution obtenue par le M.V. peut avoir une variance supérieure à celle de l'estimation obtenue par une autre méthode. Ceci est vrai en particulier si la fonction densité ne s'annule pas pour la borne dépendant du paramètre puisque (6.2) et (6.3) ne sont plus vérifiées. Le M.V. peut donc ne pas être optimum dans ce cas même pour les grands échantillons.

De manière générale il est donc souhaitable d'avoir une statistique exhaustive pour que la solution du Maximum de vraisemblance soit optimum. Si il n'existe pas une telle statistique exhaustive, la solution du Maximum de vraisemblance est douteuse pour les petits échantillons.

c. Cas de plusieurs paramètres

Les équations du M.V. pour plusieurs paramètres sont données par (6.21).

Si pour une distribution à k paramètres il existe k statistiques t_1, \dots, t_k conjointement exhaustives pour $\theta_1, \dots, \theta_k$, les estimateurs $\hat{\theta}_1, \dots, \hat{\theta}_k$ obtenus par le maximum de vraisemblance sont des fonctions de t_1, \dots, t_k .

Dans le cas où les dérivées premières et secondes de la fonction de vraisemblance existent, et si l'intervalle de définition est indépendant des paramètres, on montre que les équations du M.V. conduisent à une solution unique fonction des k statistiques exhaustives.

Dans le cas particulier où une borne de l'intervalle de définition est fonction d'un paramètre à déterminer pour lequel il existe une statistique exhaustive il est encore possible de déterminer la solution du M.V. comme dans le cas d'un seul paramètre.

Dans le cas où il n'existe pas k statistiques exhaustives pour les k paramètres, les estimateurs possèdent des propriétés optimales pour les grands échantillons seulement. On montre alors que les estimateurs conjoints du M.V. suivent asymptotiquement une distribution multinormale dont la matrice de dispersion V est l'inverse de la matrice d'information.

$$\left(V^{-1} \right)_{ij} = -E \left(\frac{\partial^2 \text{Log}L}{\partial \theta_i \partial \theta_j} \right) = E \left[\begin{array}{cc} \frac{\partial \text{Log}L}{\partial \theta_i} & \frac{\partial \text{Log}L}{\partial \theta_j} \end{array} \right] \quad (6.22)$$

Il est possible de montrer que le déterminant de la matrice de dispersion de n'importe quel ensemble de k estimateurs corrects est toujours supérieur ou égal à $|V|$ obtenu pour la solution du M.V. Donc pour de grands échantillons le déterminant de la matrice de dispersion (variance généralisée) est minimum pour la solution du M.V.

- Remarque:

Lorsque les estimateurs $\hat{\theta}_1, \dots, \hat{\theta}_k$ du M.V. ont des statistiques exhaustives, la relation précédente pour des échantillons de taille élevée devient:

$$\left(V^{-1} \right)_{ij} = - \left(\frac{\partial^2 \text{Log}L}{\partial \theta_i \partial \theta_j} \right)_{\theta = \hat{\theta}} \quad (6.23)$$

6.8 Méthode des moments

a. Principe de la méthode

La méthode des moments appliquée à une loi à k paramètres consiste à évaluer les k premiers moments non centrés de l'échantillon aux k premiers moments théoriques correspondants de la distribution exprimés en fonction des paramètres.

On dispose ainsi de k équations à k inconnues, dont la solution conduit à l'estimation des paramètres par la méthode des moments.

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.

En pratique, cette méthode est souvent utilisée en considérant les moments centrés, les cumulants ou des coefficients (variation, asymétrie) et dans ce cas on tient compte du biais éventuel de ces grandeurs.

b. Variance des paramètres

Le calcul de la variance des paramètres est important en particulier pour déterminer l'efficacité de la méthode des moments. Si l'on peut exprimer les paramètres en fonction des moments, il est alors possible de déterminer la variance des paramètres (cf. ch. 3.3 et 3.4).

6.9 Estimation du paramètre d'échelle et de position d'une distribution

On considère une fonction de distribution définie par:

$$dF(x) = f\left(\frac{x-m}{\beta}\right) d\left(\frac{x-m}{\beta}\right)$$

Cette distribution est caractérisée par deux paramètres, un paramètre de position m , un paramètre d'échelle β .

a. Méthode du maximum de vraisemblance

On considère la variable réduite $u = \frac{x-m}{\beta}$ et l'on pose

$$g(u) = \text{Log } f(u).$$

On démontre (KENDALL 1967) que si

- $g'(u)$ et $ug'(u)$, $g''(u)$ existent pour tout u
- $f(u)$ est nulle aux bornes

la matrice de dispersion des paramètres m et β dans cet ordre est:

$$V = -\frac{\beta^2}{N\Delta} \begin{bmatrix} E[u^2 g''(u) - 1] & E[-ug''(u)] \\ E[-u g''(u)] & E[g''(u)] \end{bmatrix} \quad (6.24)$$

avec

$$\Delta = E[g''(u)] * E[u^2 g''(u) - 1] - \left\{ E[-u g''(u)] \right\}^2$$

Si $g(u)$ est une fonction paire, c'est-à-dire si la distribution est symétrique autour de m (loi Normale par exemple) la matrice de dispersion est diagonale. On peut alors dire que les paramètres m et β sont asymptotiquement indépendants, puisque leur covariance asymptotique est nulle.

- Remarques

- La validité des résultats précédents est soumise aux hypothèses générales de régularité de la fonction vraisemblance (cf. 6.7b)
- Si la distribution possède un ou plusieurs autres paramètres que m et β , on doit les supposer connus pour appliquer les résultats précédents, car l'application du M.V. pour trois paramètres par exemple ne conduit pas pour m et β aux mêmes résultats que (6.24).

b. Méthode de moments

L'application de la méthode des moments (BRUNET-MORET 1973 a)

donne:

$$\beta = \frac{\sqrt{m'_2 - m_1'^2}}{\sqrt{\mu_2}} \quad (6.25)$$

$$m = m_1' - \frac{\mu_1'}{\sqrt{\mu_2}} \sqrt{m'_2 - m_1'^2} \quad (6.26)$$

m_1' et m_2' sont les moments non centrés de l'échantillon ($x_1 \dots x_N$)

μ_1' est la moyenne théorique de la distribution de u

μ_r est le moment centré par rapport à la moyenne de la distribution de u .

Si N est la taille de l'échantillon, la variance des paramètres est donnée par:

$$\text{var } \beta = \frac{\beta^2}{4N\mu_2} (\mu_4 - \mu_2^2) \quad (6.27)$$

$$\text{var } m = \frac{\beta^2}{N} \left[\mu_2 - \frac{\mu_1' \mu_3}{\mu_2} + \frac{\mu_1'^2}{4\mu_2} (\mu_4 - \mu_2^2) \right] \quad (6.28)$$

$$\text{Cov } (m, \beta) = \frac{\beta^2}{4N\mu_2} \left[2\mu_3 - \frac{\mu_1'}{\mu_2} (\mu_4 - \mu_2^2) \right] \quad (6.29)$$

Ces formules sont des approximations à l'ordre $1/N$

6.10 Conclusion

En pratique, lorsque l'on veut estimer les paramètres d'une loi, le choix de la méthode d'estimation est très important, puisqu'il conditionne la validité de l'ajustement.

La méthode des moments et la méthode du maximum de vraisemblance sont les plus fréquemment utilisées. La première est en général caractérisée par sa facilité d'application, la seconde par son efficacité mais aussi par sa difficulté d'application pratique.

Le choix de la méthode est souvent effectué suivant l'objectif poursuivi et l'usage que l'on compte faire de la distribution ajustée, cependant l'utilisation croissante des ordinateurs favorise la méthode du maximum de vraisemblance. Cette dernière permet en effet lorsqu'il existe des estimateurs efficaces de les déterminer. Cependant, lorsqu'il n'existe pas de statistiques exhaustives des paramètres, l'optimalité du maximum de vraisemblance n'est qu'asymptotique et il est dangereux d'étendre les propriétés de cette méthode à des échantillons de faible taille. D'autres méthodes d'estimations sont parfois employées (χ^2 minimum, quantiles, statistiques d'ordre), mais leur efficacité ne peut être examinée que dans des cas bien précis.

7. ESTIMATION DES PARAMETRES DE LA DISTRIBUTION PEARSON III

7.1 Distribution Pearson III à 1 paramètre

a. Propriétés générales d'estimation

La fonction densité de la loi Pearson III à 1 paramètre est (1.7):

$$f(x) = \frac{e^{-x}}{\Gamma(\lambda)} x^{\lambda-1} \quad (7.1)$$

avec $\lambda > 0$; $0 \leq x$

La fonction densité peut se mettre sous la forme (6.14):

$$f(x, \lambda) = \exp \left[(\lambda-1) \text{Log } x - x - \text{Log } \Gamma(\lambda) \right]$$

avec

$$A(\lambda) = \lambda-1 ; B(x) = \text{Log } x ; C(x) = -x ; D(\lambda) = -\text{Log } \Gamma(\lambda)$$

La distribution possède donc une statistique exhaustive de λ , il est donc possible de trouver un estimateur efficace t d'une fonction de λ qui atteint la variance minimum.

La fonction vraisemblance pour un échantillon de taille N est telle que:

$$\text{Log } L = - \sum_{i=1}^N x_i - N \text{Log } \Gamma(\lambda) + (\lambda-1) \sum_{i=1}^N \text{Log } x_i \quad (7.2)$$

$\frac{d \text{Log } L}{d\lambda}$ peut se mettre sous la forme (6.8) en effet:

$$\frac{d \text{Log } L}{d\lambda} = N \left[\frac{1}{N} \sum \text{Log } x_i - \frac{d \text{Log } \Gamma(\lambda)}{d\lambda} \right] \quad (7.3)$$

Donc l'estimateur $t = \frac{1}{N} \sum_{i=1}^N \text{Log } x_i$ de la fonction $E(t)$ telle que:

$$E(t) = \psi(\lambda) = + \frac{d \text{Log } \Gamma(\lambda)}{d\lambda} \quad \text{atteint la variance minimum:}$$

$$\text{var } t = \frac{1}{N} \frac{d^2 \text{Log } \Gamma(\lambda)}{d\lambda^2} \quad (7.4)$$

Donc le seul estimateur efficace atteignant la variance minimum est

$t = \frac{1}{N} \sum \text{Log } x_i$ (logarithme de la moyenne géométrique). Un estima-

teur t_0 tel que $E(t_0) = \lambda$ aura toujours une variance supérieure

à la variance limite, donc:

$$\text{var } t_0 > \frac{1}{-E \left[\frac{d^2 \text{Log } L}{d\lambda^2} \right]} = \frac{1}{N \frac{d^2 \text{Log } \Gamma(\lambda)}{d\lambda^2}}$$

b. Méthode du maximum de vraisemblance

L'équation donnant la solution λ_* du maximum de vraisemblance est obtenue en annulant (7.3) d'où:

$$\left(\frac{d \text{Log } \Gamma(\lambda)}{d\lambda} \right) = \frac{1}{N} \sum_{i=1}^N \text{Log } x_i \quad (7.5)$$

La solution λ_* du M.V. est unique, mais dans le cas de cet estimateur exhaustif la variance minimum n'est pas atteinte (puisque $\psi(\lambda) \neq \lambda$). Cependant pour les grands échantillons λ_* a une variance asymptotique telle que:

$$\text{var } (\lambda_*) = \frac{1}{-E \left[\frac{d^2 \text{Log } L}{d\lambda^2} \right]} = \frac{1}{N \frac{d^2 \text{Log } \Gamma(\lambda)}{d\lambda^2}} \quad (7.6)$$

λ_* a une efficacité asymptotique de 1.

La résolution de l'équation (7.5) du M.V. donnant λ_* peut être effectuée à l'aide des tables de la fonction Digamma (DAVIS 1933).

c. Méthode des moments

Les moments de la distribution Pearson III à 1 paramètre sont, (4.2):

moyenne: $\mu_1' = \lambda$

variance: $\mu_2 = \lambda$

La détermination du paramètre λ par la méthode des moments a pour solution:

$$\lambda_o = \frac{1}{N} \sum_{i=1}^N x_i = m_1' \quad (7.7)$$

Cet estimateur est absolument correct et a pour variance:

$$\text{var } \lambda_o = \text{var } m_1' = \frac{\lambda}{N} \quad (7.8)$$

D'après (6.11) l'efficacité de cette estimation est pour de grands échantillons:

$$e(\lambda_o) = \frac{1}{\lambda \cdot \frac{d^2 \text{Log } \Gamma(\lambda)}{d\lambda^2}} \quad (7.9)$$

$e(\lambda_o)$ est toujours inférieur à 1 et tend vers 0 quand $\lambda \rightarrow 0$.

Quand λ devient grand $e(\lambda_o)$ tend vers 1. La méthode des moments est donc d'autant plus efficace que la distribution Pearson III à 1 paramètre a une faible asymétrie, comme l'indique la table 13.

λ_0	64	16	8	4	2	1	.36	.25	.16	.04
C_s	.25	.50	$\frac{\sqrt{2}}{2}$	1	$\sqrt{2}$	2	3.33	4	5	10
$e(\lambda_0)$.99	.97	.94	.88	.78	.61	.32	.23	.15	.04

Table 13 : Efficacité de la méthode des moments, Pearson III à 1 paramètre.

7.2 Estimation des paramètres de la loi Gamma

a. Propriétés générales d'estimation

La forme générale de la fonction densité de la loi Gamma est:

$$f(x) = \frac{|\alpha|}{\Gamma(\lambda)} e^{-\alpha x} (\alpha x)^{\lambda-1} \quad (7.10)$$

avec : $\lambda > 0$ et $\alpha x \geq 0$ (α peut être positif ou négatif)

$f(x; \alpha, \lambda)$ peut se mettre sous la forme (6.19), en effet:

$$f(x; \alpha, \lambda) = \exp \left[\left(\lambda \text{Log } |\alpha| - \text{Log } \Gamma(\lambda) \right) - \alpha x + (\lambda-1) \text{Log } |x| \right]$$

il est donc théoriquement possible de trouver 2 statistiques conjointement exhaustives pour les paramètres α et λ .

. Si α est fixé

on a: $\text{Log } L = N \text{Log } |\alpha| - N \text{Log } \Gamma(\lambda) - \alpha \sum x_i + (\lambda-1) \sum \text{Log } (\alpha x_i)$

$$\frac{\partial \text{Log } L}{\partial \lambda} = N \left[- \frac{d \text{Log } \Gamma(\lambda)}{d \lambda} + \frac{1}{N} \sum \text{Log } (\alpha x_i) \right] \quad (7.11)$$

La relation (7.11) peut se mettre sous la forme (6.8) avec

$$t_1 = \frac{1}{N} \sum \text{Log } \alpha x_i ; \psi(\lambda) = \frac{d \text{Log } \Gamma(\lambda)}{d \lambda} ; A = N$$

L'estimateur $t_1 = \frac{1}{N} \sum \text{Log } \alpha x_i$ est efficace et exhaustif pour λ

si α est fixé et l'on a :

$$E(t_1) = \frac{d \text{Log } \Gamma(\lambda)}{d\lambda}$$

$$\text{var } t_1 = \frac{1}{N} \frac{d^2 \text{Log } \Gamma(\lambda)}{d\lambda^2}$$

Pour $\alpha = 1$, on retrouve comme cas particulier (7.1a)

. Si λ est fixé

on a : $\frac{\partial \text{Log } L}{\partial \alpha} = \frac{N}{\alpha} - \sum x_i + \frac{N}{\alpha} (\lambda - 1)$

$$\frac{\partial \text{Log } L}{\partial \alpha} = -N\lambda \left[\frac{\sum x_i}{N\lambda} - \frac{1}{\alpha} \right] \quad (7.12)$$

La relation (7.12) est de la forme (6.8) avec :

$$A(\alpha) = -N\lambda \quad ; \quad t_2 = \frac{1}{\lambda} \frac{\sum x_i}{N} \quad ; \quad \psi(\alpha) = \frac{1}{\alpha}$$

Donc l'estimateur $t_2 = \frac{m'_1}{\lambda}$ $\left(m'_1 = \frac{1}{N} \sum x_i \text{ est la moyenne de l'échantillon} \right)$

est efficace et exhaustif pour le paramètre α , lorsque λ est fixé et il est tel que :

$$E(t_2) = \frac{1}{\alpha}$$

$$\text{var } t_2 = \frac{1}{N\lambda\alpha^2}$$

Cet estimateur atteint donc la variance minimum.

b. Méthode du Maximum de vraisemblance

La détermination des paramètres α et λ par la méthode du M.V. est obtenue par la résolution des équations:

$$\left\{ \begin{array}{l} \frac{\partial \text{Log} L}{\partial \alpha} = N \lambda \left(\frac{1}{\alpha} - \frac{m'_1}{\lambda} \right) = 0 \\ \frac{\partial \text{Log} L}{\partial \lambda} = N \left[\frac{1}{N} \sum \text{Log}(\alpha x_i) - \frac{d \text{Log} \Gamma(\lambda)}{d\lambda} \right] = 0 \end{array} \right.$$

Puisque l'on sait qu'il existe des estimateurs conjointement exhaustifs de α et λ , la solution α_* , λ_* du M.V. est unique.

Le système des équations du M.V. peut se mettre sous la forme:

$$\left\{ \begin{array}{l} \alpha = \frac{\lambda}{m'_1} \\ \text{Log} \lambda - \frac{d \text{Log} \Gamma(\lambda)}{d\lambda} = A \end{array} \right. \quad (7.13)$$

où : m'_1 est la moyenne arithmétique de l'échantillon.

$$A = - \frac{1}{N} \sum_{i=1}^N \text{Log} \left(\frac{x_i}{m'_1} \right)$$

L'équation (7.14) est implicite en λ et peut être résolue par itération en utilisant les tables de DAVIS (1933) qui donnent la fonction digamma

$$\frac{d \text{Log} \Gamma(\lambda)}{d\lambda}$$

Cependant il existe des tables permettant de déterminer directement λ_* connaissant A, citons en particulier les tables de MUSUYAMA et KUROIWA (1951), et celles de GREENWOOD et DURAND (1960).

THOM (1958) montre que l'on peut directement déterminer une valeur approchée de λ par:

$$\lambda_* = \frac{1 + \sqrt{1 + \frac{4}{3} A}}{4 A} \quad (7.14)$$

Connaissant λ_* , α_* est déterminé par (7.13).

Les estimateurs λ_* et α_* obtenus par la méthode du M.V. sont fonction des statistiques conjointement exhaustives, mais la distribution conjointe de α_* et λ_* n'est pas connue, on peut cependant déterminer la matrice de dispersion de la distribution normale asymptotique de λ_* et α_* .

L'inverse de la matrice de dispersion est obtenue en utilisant la relation (6.23) on a:

$$V^{-1} = N \begin{bmatrix} \lambda/\alpha^2 & -1/\alpha \\ -1/\alpha & \frac{d^2 \text{Log } \Gamma(\lambda)}{d\lambda^2} \end{bmatrix}$$

En posant : $\psi' = \frac{d^2 \text{Log } \Gamma(\lambda)}{d\lambda^2}$ (fonction Trigamma) et $\eta = \psi' - \frac{1}{\lambda}$

$$\text{on obtient: } \text{var } \alpha_* = \frac{\alpha^2 \psi'}{N \lambda \eta} \quad (7.15)$$

$$\text{var } \lambda_* = \frac{1}{N \eta} \quad (7.16)$$

$$\text{Cov} \left(\alpha_*, \lambda_* \right) = \frac{\alpha}{N \lambda \eta} \quad (7.17)$$

Les relations précédentes sont des limites asymptotiques et ne sont rigoureusement valables que pour les grands échantillons.

Il est possible de déterminer le coefficient de corrélation entre

α_* et λ_* on a:

$$\rho_{\alpha_*, \lambda_*} = \frac{\alpha}{|\alpha|} \frac{1}{\sqrt{\lambda \psi'}} \quad (7.18)$$

Le coefficient de corrélation est du signe de α , quand λ devient grand il tend vers 1.

- Remarque

- Si α est connu, on détermine seulement λ_* par (7.14) l'estimateur λ_* est exhaustif, mais il n'est pas efficace sa variance asymptotique est alors:

$$\text{var } \lambda_* = \frac{1}{N\psi'}$$

- Si λ est fixé, on détermine seulement α_* par (7.13), cet estimateur est exhaustif mais il n'est pas efficace, il est seulement asymptotiquement efficace et l'on a:

$$\text{var } \alpha_* = \frac{\alpha^2}{N\lambda}$$

c. Méthode des moments

Si m'_1 et m'_2 sont respectivement la moyenne et la variance de l'échantillon, la méthode des moments conduit à:

$$\begin{cases} m'_1 = \frac{\lambda}{\alpha} \\ m'_2 = \frac{\lambda}{\alpha^2} \end{cases}$$

La première équation est identique à (7.13).

On déduit du système précédent, la solution α_o, λ_o

$$\alpha_o = \frac{m'_1}{m'_2} \tag{7.19}$$

$$\lambda_o = \frac{m'^2_1}{m'_2} \tag{7.20}$$

L'application de la relation (3.29) donne:

$$\text{var } \alpha_o = \frac{1}{\mu_2} \text{var } m'_1 + \frac{\mu_1'^2}{\mu_2} \text{var } m'_2 - 2 \frac{\mu_1'}{\mu_2} \text{Cov} \left(m'_1, m'_2 \right)$$

En utilisant (3.17), (3.20), (3.22) et en remplaçant les moments théoriques en fonction des paramètres de la distribution (Table 4), on obtient:

$$\text{var } \alpha_o = \frac{\alpha^2}{\lambda N} (2\lambda + 3) \quad (7.21)$$

De la même manière, on obtient:

$$\text{var } \lambda_o = \frac{2\lambda}{N} (\lambda + 1) \quad (7.22)$$

$$\text{Cov} \left(\alpha_o, \lambda_o \right) = \frac{2\alpha}{N} (1 + \lambda) \quad (7.23)$$

Le coefficient de corrélation entre α_o et λ_o est:

$$\rho_{\alpha_o, \lambda_o} = \frac{\alpha}{|\alpha|} \sqrt{\frac{1 + \lambda}{\lambda + 3/2}} \quad (7.24)$$

Le coefficient de corrélation est du signe de α , il augmente quand λ croît, donc quand l'asymétrie diminue.

d. Efficacité de la méthode des moments

La méthode du M.V. conduit à des estimations asymptotiquement efficaces, les variances des paramètres α_* et λ_* sont des valeurs limites quand l'échantillon devient grand. Dans la méthode des moments, les variances des paramètres sont également des approximations valables quand N est suffisamment élevé.

On évalue cependant l'efficacité de la méthode des moments par rapport à la méthode du Maximum de vraisemblance en considérant le rapport des variances des estimations d'un même paramètre par ces deux méthodes.

$$\text{On a: } e \left(\alpha_o \right) = \frac{\text{var} \left(\alpha_* \right)}{\text{var} \left(\alpha_o \right)} = \frac{\psi'}{\eta (2\lambda + 3)} \quad (7.25)$$

ψ' et η sont définis en (7.2 b)

$$e(\lambda_0) = \frac{\text{var}(\lambda_*)}{\text{var}(\lambda_0)} = \frac{1}{\eta \cdot 2\lambda(\lambda+1)} \quad (7.26)$$

Quand λ devient grand, c'est-à-dire quand l'asymétrie de la distribution tend vers 0, l'efficacité relative de α et λ tend vers 1. Ce résultat est logique puisque la loi Gamma tend vers la loi Normale lorsque λ devient grand et dans ce cas la méthode des moments et du maximum de vraisemblance sont équivalentes.

La table 14 donne les valeurs des efficacités $e(\alpha_0)$ et $e(\lambda_0)$ pour quelques valeurs de λ

λ	64	16	8	4	2	1	.36	.25	.16	.04
c_s	.25	.50	$\frac{\sqrt{2}}{2}$	1	$\sqrt{2}$	2	3.33	4	5	10
$e(\alpha_0)$.98	.92	.86	.76	.64	.51	.39	.37	.36	.02
$e(\lambda_0)$.98	.92	.85	.74	.57	.39	.17	.12	.08	.02

Table 14: Efficacité de la méthode des moments pour la loi Gamma

Il est également possible de donner des valeurs approchées de $e(\alpha_0)$ et $e(\lambda_0)$ pour λ élevé, en effet la formule de Stirling donne pour λ élevé:

$$\text{Log } \Gamma(\lambda) \approx \left(\lambda - \frac{1}{2}\right) \text{Log } \lambda - \lambda + \frac{1}{2} \text{Log } 2\pi + \frac{1}{12\lambda} - \frac{1}{360\lambda^3} + \dots$$

d'où:

$$\psi' = \frac{d^2 \text{Log } \Gamma(\lambda)}{dx^2} \approx \frac{1}{\lambda} + \frac{1}{2\lambda^2} + \frac{1}{6\lambda^3} - \frac{1}{30\lambda^5} + \dots$$

$$\eta = \psi' - \frac{1}{\lambda} \approx \frac{1}{2\lambda^2} \left(1 + \frac{1}{3\lambda} - \frac{1}{15\lambda^3} + \dots \right)$$

On a alors pour λ grand:

$$e(\alpha_0) \approx 1 - \frac{4}{3\lambda} + \frac{19}{9\lambda^2} + \dots \quad (7.27)$$

$$e(\lambda_0) \approx 1 - \frac{4}{3\lambda} + \frac{13}{9\lambda^2} + \dots \quad (7.28)$$

Pour $\lambda > 4$ ces formules approchées donnent une erreur relative inférieure à 5% sur $e(\alpha_0)$ et $e(\lambda_0)$

Remarque:

. Si l'on considère la loi Gamma, avec α fixé à priori on est dans le cas de la loi Pearson III à un seul paramètre et l'on a:

$$e(\lambda) = \frac{1}{\lambda\psi'}$$

On retrouve la relation (7.9)

. Si l'on considère la loi Gamma avec λ fixé à priori, α étant le seul paramètre à déterminer, la solution de la méthode des moments est:

$$\alpha_0 = \frac{\lambda}{m_1} \quad \text{avec} \quad \text{var } \alpha_0 = \frac{\alpha^2}{\lambda N}$$

On retrouve le même résultat que dans la méthode du M.V. (cf 7.2 b) on a donc dans ce cas

$$e(\alpha_0) = 1$$

7.3 Estimation des paramètres de la distribution Pearson III

a. Propriétés générales d'estimation

La forme générale de la distribution Pearson III donnée par (1.5) est:

$$f(x; \alpha, \lambda, m) = \frac{|\alpha|}{\Gamma(\lambda)} e^{-\alpha(x-m)} [\alpha(x-m)]^{\lambda-1}$$

l'intervalle de définition est tel que $\alpha(x-m) \geq 0, \lambda > 0$

- . Si α , λ , m sont inconnus:

La fonction densité ne peut se mettre sous la forme (6.14), il n'existe donc pas de statistiques conjointement exhaustives pour les paramètres α , λ , m . On ne pourra donc trouver d'estimateurs efficaces, c'est-à-dire atteignant la variance minimum, pour ces paramètres.

- . Si α et λ sont inconnus, m étant connu:

On est ramené au cas de la loi Gamma avec le changement d'origine $y = (x-m)$, puisque m est connu. Il est alors possible (7.2 a) de trouver deux statistiques conjointement exhaustives pour α et λ . Tout ce qui a été vu pour la loi Gamma s'applique dans ce cas.

- . Si α et m sont inconnus, λ étant connu:

Il n'est pas possible de mettre en évidence deux statistiques conjointement exhaustives pour α et m .

- . Si λ et m sont inconnus, α étant connu:

Comme dans le cas précédent, il n'y a pas de statistiques conjointement exhaustives.

- . Si α est inconnu, λ et m étant connus:

La fonction densité peut se mettre sous la forme (6.14) et l'on a:

$$f(x;\alpha) = \exp \left[-\alpha(x-m) + (\lambda-1) \text{Log} |x-m| + \left(\lambda \text{Log} |\alpha| - \text{Log} \Gamma(\lambda) \right) \right]$$

On retrouve le cas de la loi Gamma appliqué à la variate $y = (x-m)$ lorsque λ est fixé (7.2 a). Il est encore possible de montrer que:

$$t = \frac{1}{\lambda} \frac{\sum (x_i - m)}{N} \text{ est un estimateur efficace de } \psi(\alpha) = \frac{1}{\alpha}$$

qui atteint la variance minimum $\frac{1}{N \lambda \alpha^2}$

- Si λ est inconnu, α et m étant fixés

On retrouve le cas de la loi Gamma avec α fixé. Il existe un estimateur efficace atteignant la variance minimum.

- Si m est inconnu, α et λ étant fixés

Dans ce cas on ne peut appliquer la relation (6.14) puisque l'intervalle de définition de la variate est fonction du paramètre m , que l'on doit estimer.

D'autre part, la fonction densité peut se mettre sous la forme (6.15) seulement et seulement si $\lambda = 1$, on a alors:

$$f(x; m) = \frac{|\alpha|}{\Gamma(\lambda)} e^{-\alpha x} * e^{\alpha m}$$

Donc si l'on doit déterminer seulement m (α et λ étant fixés) ce paramètre possède une statistique exhaustive lorsque $\lambda = 1$, cette statistique est:

- x_1 , plus petite valeur de l'échantillon lorsque $\alpha > 0$ et $\lambda = 1$
- x_N plus grande valeur de l'échantillon lorsque $\alpha < 0$ et $\lambda = 1$

De plus, la variance minimum n'est pas atteinte car pour $\lambda = 1$, la fonction densité ne s'annule pas à la borne $x = m$. Les relations (6.2) et (6.3) donnant la valeur de la variance minimum (que celle-ci soit ou non atteinte) restent vraies si la fonction et sa dérivée s'annulent pour $x = m$, ce qui est obtenu pour $\lambda > 2$.

b. Méthode du Maximum de vraisemblance

Cas où les 3 paramètres α , λ , m sont inconnus

Nous avons vu que dans ce cas, il n'existe pas de statistiques conjointement exhaustives des paramètres, les propriétés du M.V. sont donc asymptotiques et valables pour de grands échantillons.

La fonction vraisemblance L est telle que:

$$\text{Log } L = N \text{Log } |\alpha| - N \text{Log } \Gamma(\lambda) - \alpha \sum_{i=1}^N (x_i - m) + (\lambda - 1) \sum_{i=1}^N \text{Log} [\alpha(x_i - m)] \quad (7.29)$$

La solution m_* , α_* , λ_* du M.V. est donnée par:

$$\frac{\partial \text{Log } L}{\partial \alpha} = \frac{\lambda N}{\alpha} - \sum_{i=1}^N (x_i - m) = 0 \quad (7.30)$$

$$\frac{\partial \text{Log } L}{\partial \lambda} = -N \frac{d \text{Log } \Gamma(\lambda)}{d\lambda} + \sum_{i=1}^N \text{Log} [\alpha(x_i - m)] = 0 \quad (7.31)$$

$$\frac{\partial \text{Log } L}{\partial m} = \alpha N - (\lambda - 1) \sum_{i=1}^N \frac{1}{(x_i - m)} = 0 \quad (7.32)$$

Ces équations permettent plusieurs remarques:

. Puisque l'on a $\alpha(x-m) > 0$, quel que soit le signe de α , l'équation (7.32) conduit toujours à une solution $\lambda_* \geq 1$. Si l'échantillon provient d'une population telle que $0 < \lambda < 1$, la solution du M.V. est donc biaisée.

. Si $\lambda \leq 1$, la fonction densité ne s'annule pas pour la borne $x = m$, et d'après ce qui a été vu en (6.5), la relation donnant la variance minimum n'est plus vraie. La méthode du M.V. n'est pas optimale même asymptotiquement pour de grands échantillons, et une autre méthode que le M.V. peut conduire à des estimations de plus faible variance.

. La méthode du M.V. présente donc un intérêt pour les grands échantillons lorsque $\lambda > 1$. Si de plus $\lambda > 2$, la fonction densité et sa dérivée s'annulent pour $x = m$ et le calcul de la variance minimum est simplifié (6.5).

La **résolution** des équations du M.V. est difficile à obtenir puisque l'on doit résoudre un système de trois équations non linéaires à trois inconnues.

Il est possible d'utiliser les méthodes décrites pour la loi Gamma (7.2 b) pour m fixé en considérant les équations (7.30) et (7.31), et de calculer le premier membre de (7.32); on fait varier m jusqu'à ce que cette quantité s'annule.

Une autre méthode décrite par MATALAS et WALLIS (1973) consiste à déterminer α et λ à partir de (7.30) et (7.32) en supposant m connu, on a:

$$\lambda_1 = \frac{\sum_{i=1}^N \left(\frac{1}{x_i - m} \right)}{\sum_{i=1}^N \left(\frac{1}{x_i - m} \right) - \frac{N^2}{\sum_{i=1}^N (x_i - m)}} \quad (7.33)$$

$$\alpha_1 = \lambda_1 \frac{N}{\sum_{i=1}^N (x_i - m)} \quad (7.34)$$

On reporte α_1 et λ_1 dans R tel que

$$R = -N \frac{d \operatorname{Log} \Gamma (\lambda_1)}{d\lambda} + \sum_{i=1}^N \operatorname{Log} \left[\alpha_1 (x_i - m) \right] \quad (7.35)$$

et l'on fait varier m , jusqu'à ce que $|R| < \varepsilon$, ε étant une valeur petite fixée à priori en fonction de la précision désirée.

Il est possible de calculer les variances asymptotiques des estimations α_* , λ_* et m_* obtenues par la méthode du M.V., en utilisant la relation (6.22), valable pour $\lambda > 2$.

L'inverse de la matrice de dispersion est en prenant les paramètres suivant l'ordre α , λ , m

$$V^{-1} = N \begin{bmatrix} \frac{\lambda}{\alpha^2} & -\frac{1}{\alpha} & -1 \\ -\frac{1}{\alpha} & \psi' & \frac{\alpha}{\lambda-1} \\ -1 & \frac{\alpha}{\lambda-1} & \frac{\alpha^2}{\lambda-2} \end{bmatrix}$$

avec $\psi' = \psi'(\lambda) = \frac{d^2 \text{Log } \Gamma(\lambda)}{d\lambda^2}$ (fonction trigamma)

Le déterminant de V^{-1} est Δ_* tel que:

$$\Delta_* = \frac{1}{(\lambda-2)} \left[2\psi' - \frac{2}{\lambda-1} + \frac{1}{(\lambda-1)^2} \right] \quad (7.36)$$

Et l'on a en particulier pour les variances asymptotiques:

$$\text{var}(\alpha_*) = \frac{\alpha^2}{N\Delta_*} \left[\frac{\psi'}{\lambda-2} - \frac{1}{(\lambda-1)^2} \right] \quad (7.37)$$

$$\text{var}(\lambda_*) = \frac{2}{N(\lambda-2)\Delta_*} \quad (7.38)$$

$$\text{var}(m_*) = \frac{\lambda\psi' - 1}{N\alpha^2\Delta_*} \quad (7.39)$$

Ces résultats sont vrais pour de grands échantillons, lorsque $\lambda > 2$

Lorsque $1 < \lambda \leq 2$, il est encore possible théoriquement de déterminer la matrice de dispersion asymptotique, mais les calculs sont plus complexes, puisque l'on ne peut utiliser la dérivée seconde de la fonction vraisemblance.

Cas où α et m sont inconnus, λ étant fixé

Ce cas peut être important en pratique, car il est parfois possible de fixer à priori le coefficient d'asymétrie de la distribution, donc λ est connu.

La solution α_* et m_* est obtenue par la résolution des équations (7.30) et (7.32), on suppose λ fixé tel que $\lambda > 2$ pour pouvoir appliquer les résultats de (6.9 a)

$$\text{On pose } u = \frac{x-m}{\beta} \quad \text{avec} \quad \beta = \frac{1}{\alpha}$$

La fonction densité de la variate u est telle que:

$$f(u) = \frac{e^{-u} u^{\lambda-1}}{\Gamma(\lambda)} \quad u \geq 0 \quad ; \quad \lambda > 2$$

$$g(u) = \text{Log } f(u) = -u + (\lambda-1) \text{Log } u - \text{Log } \Gamma(\lambda)$$

$$g'(u) = -1 + \frac{\lambda-1}{u}$$

$$g''(u) = -\frac{\lambda-1}{u^2}$$

On peut en déduire:

$$E \left[g''(u) \right] = \int_0^{+\infty} g''(u) f(u) du$$

$$E \left[g''(u) \right] = -\frac{(\lambda-1)}{\Gamma(\lambda)} \int_0^{+\infty} e^{-u} u^{\lambda-3} du$$

$$E \left[g''(u) \right] = -(\lambda-1) \frac{\Gamma(\lambda-2)}{\Gamma(\lambda)} = -\frac{1}{(\lambda-2)}$$

De la même manière:

$$E \left[-u g''(u) \right] = +1$$

$$E \left[u^2 g''(u) -1 \right] = -\lambda$$

D'après (6.24) la matrice de dispersion des paramètres m et β pris dans cet ordre est:

$$V = -\frac{\beta^2}{N\Delta} \begin{bmatrix} -\lambda & 1 \\ 1 & -\frac{1}{\lambda-2} \end{bmatrix} \quad \text{avec} \quad \Delta = \frac{2}{\lambda-2}$$

On peut alors déduire:

$$\text{var } m_* = \frac{\lambda (\lambda-2) \beta^2}{2N}$$

$$\text{var } \beta_* = \frac{\beta^2}{2N}$$

Si l'on revient aux paramètres α et m on peut écrire:

$$\text{var } m_* = \frac{\lambda (\lambda-2)}{2N \alpha^2} \quad (7.40)$$

$$\text{var } \alpha_* = \frac{\alpha^2}{2N} \quad (7.41)$$

$$\text{Cov} \left(\alpha_*, m_* \right) = \frac{\lambda-2}{2N} \quad (7.42)$$

Ces relations donnent les variances et la covariance asymptotiques des paramètres déterminés par la méthode du maximum de vraisemblance qui est optimale pour de grands échantillons lorsque λ est plus grand que 2.

c. Méthode des moments

Cas où les 3 paramètres α , λ , m sont inconnus

Puisque l'on doit déterminer trois paramètres, nous considérons trois moments (ou fonctions de moments) indépendants.

Soient k_1, k_2, k_3 les estimations non biaisées des trois premiers cumulants, obtenues à partir de l'échantillon.

Les trois équations de la méthode des moments s'écrivent alors:

$$k_1 = m + \frac{\lambda}{\alpha} \quad (7.43)$$

$$k_2 = \frac{\lambda}{\alpha^2} \quad (7.44)$$

$$k_3 = \frac{2\lambda}{\alpha^3} \quad (7.45)$$

On peut également considérer comme troisième équation, le coefficient

d'asymétrie $C_s = \frac{k_3}{k_2^{3/2}}$, ce qui conduit à la même solution que le

système précédent.

Connaissant k_1 , k_2 , k_3 déduits de l'échantillon, on peut déterminer explicitement la solution α_0 , λ_0 , m_0 :

$$\lambda_0 = 4 \frac{k_2^3}{k_3^2} \quad (7.46)$$

$$\alpha_0 = 2 \frac{k_2}{k_3} \quad (7.47)$$

$$m_0 = k_1 - \frac{\lambda_0}{\alpha_0} = k_1 - 2 \frac{k_2^2}{k_3} \quad (7.48)$$

Le signe de α_0 est celui de k_3 (ou du coefficient d'asymétrie). Pour de grands échantillons k_1 , k_2 , k_3 sont respectivement la moyenne m_1' , la variance m_2 , le moment centré du 3^o ordre m_3 de l'échantillon, puisque la correction de biais n'a plus d'influence si N est grand. Il est donc possible de déterminer les variances asymptotiques des paramètres α_0 , λ_0 , m_0 obtenus par la méthode des moments en utilisant la relation (3.29).

On a :

$$\text{var } \alpha_o = 4 \text{ var } \left(\frac{m_2}{m_3} \right) = \frac{1}{\mu_3} \text{ var } m_2 + \frac{\mu_2^2}{\mu_3^4} \text{ var } m_3 - 2 \frac{\mu_2}{\mu_3} \text{ Cov}(m_2, m_3)$$

En utilisant (3.20) et (3.21) puis en remplaçant les moments théoriques μ_r en fonction des paramètres on obtient les variances asymptotiques :

$$\text{var } \alpha_o = \frac{\alpha^2}{2\lambda N} (24 + 25\lambda + 3\lambda^2) \quad (7.49)$$

$$\text{var } \lambda_o = \frac{6\lambda}{N} (\lambda + 1) (\lambda + 5) \quad (7.50)$$

$$\text{var } m_o = \frac{\lambda (\lambda + 1) (3\lambda + 10)}{2N\alpha^2} \quad (7.51)$$

Cas où α et m sont inconnus, λ étant fixé

Puisque λ est connu, les équations (7.43) et (7.44) suffisent pour déterminer α_o et m_o :

$$\alpha_o = (\text{signe } k_3) \sqrt{\frac{\lambda}{k_2}} \quad (7.52)$$

$$m_o = k_1 - \frac{\lambda}{\alpha_o} \quad (7.53)$$

Les variances asymptotiques sont alors :

$$\text{var } \alpha_o = \frac{\alpha^2 (\lambda + 3)}{2 \lambda N} \quad (7.54)$$

$$\text{var } m_o = \frac{\lambda (\lambda + 1)}{2 N \alpha^2} \quad (7.55)$$

d. Efficacité de la méthode des moments

La méthode du Maximum de vraisemblance permet l'estimation des paramètres de la loi Pearson III lorsque le paramètre de forme λ est plus grand que 1. La méthode est alors asymptotiquement efficace, et les variances limites d'estimation des paramètres pour des échantillons de taille élevée ont été calculés en (7.3 b) pour $\lambda > 2$.

La méthode des moments est applicable indépendamment de la valeur de λ , ici encore le calcul des variances effectué n'est valable que pour de grands échantillons.

Pour $\lambda > 2$ il est possible de comparer les variances asymptotiques obtenues par les deux méthodes.

Cas où les 3 paramètres α , λ , m sont inconnus

L'efficacité pour le paramètre λ est donnée par

$$e(\lambda) = \frac{\text{var } \lambda_*}{\text{var } \lambda_0}$$

λ_* et λ_0 étant respectivement les estimations par le M.V. et par la méthode des moments, données par les relations (7.38) et (7.50).

$$e(\lambda) = \frac{1}{3K \lambda (\lambda + 1) (\lambda + 5)} \quad (7.56)$$

avec

$$K = (\lambda - 2) \Delta_* = \left[2 \psi'(\lambda) - \frac{2}{\lambda-1} + \frac{1}{(\lambda-1)^2} \right] \quad (7.57)$$

Connaissant la fonction Trigamma $\psi'(\lambda)$ à partir des tables de DAVIS (1933), on peut calculer $e(\lambda)$ pour différentes valeurs de λ (table 15).

Pour λ suffisamment grand, on peut obtenir une valeur approchée, en effet la formule de STIRLING donne $\text{Log } \Gamma(\lambda)$

$$\text{Log } \Gamma(\lambda) \approx \left(\lambda - \frac{1}{2} \right) \text{Log } \lambda - \lambda + \frac{1}{2} \text{Log } 2\pi + \frac{1}{12\lambda} - \frac{1}{360\lambda^3}$$

D'où l'on déduit $\psi'(\lambda) = \frac{d^2 \text{Log } \Gamma(\lambda)}{d\lambda^2} :$

$$\psi'(\lambda) \approx \frac{1}{\lambda} + \frac{1}{2\lambda^2} + \frac{1}{6\lambda^3} - \frac{1}{30\lambda^5} \quad (7.58)$$

Puisque $\Gamma(\lambda) = (\lambda-1) \Gamma(\lambda-1)$

on a
$$\psi'(\lambda) = \psi'(\lambda-1) - \frac{1}{(\lambda-1)^2}$$

En remplaçant dans (7.57) $\psi'(\lambda)$ en fonction de $\psi'(\lambda-1)$ et en développant $\psi'(\lambda-1)$ d'après (7.58) on obtient:

$$K \approx \frac{1}{3(\lambda-1)^3} \left[1 - \frac{1}{5(\lambda-1)^2} \right]$$

$$e(\lambda) = \frac{\left[(\lambda-1)^3 + \frac{1}{5} (\lambda-1) \right]}{\lambda (\lambda+1) (\lambda+5)} \quad (7.59)$$

Pour $\lambda > 3$ cette formule approchée conduit à une erreur relative inférieure à 1%.

Quant λ devient très grand, l'efficacité $e(\lambda)$ tend vers 1

. L'efficacité relative de α est obtenue en utilisant (7.37) et (7.49)

$$e(\alpha) = \frac{2\lambda \left[\psi' - \frac{\lambda-2}{(\lambda-1)^2} \right]}{K (3\lambda^2 + 25\lambda + 24)} \quad (7.60)$$

K est défini par (7.57)

Une formule approchée donnant une erreur relative inférieure à 1% pour $\lambda > 3$ est:

$$e(\alpha) \approx \frac{\left(\lambda^2 - \frac{2}{3}\lambda + \frac{1}{5} \frac{\lambda}{\lambda-1} \right)}{\left(\lambda^2 + \frac{25}{3}\lambda + 8 \right)} \quad (7.61)$$

L'efficacité $e(\alpha)$ calculée pour quelques valeurs de λ est indiquée table (15).

L'efficacité relative de m est obtenue à l'aide des relations (7.39) et (7.51):

$$e(m) = \frac{2(\lambda \psi' - 1)}{K} * \frac{(\lambda-2)}{\lambda(\lambda+1)(3\lambda+10)} \quad (7.62)$$

Pour $\lambda > 3$, on peut calculer $e(m)$ avec une erreur relative inférieure à 1% par la formule approchée:

$$e(m) = \frac{(\lambda-2)}{(\lambda^2-1) \left(\lambda + \frac{10}{3} \right)} \left[\lambda^2 - \frac{11}{3}\lambda + \frac{73}{15} - \frac{12}{5\lambda} \right] \quad (7.63)$$

La relation (7.63) montre que $e(m)$ tend vers 1 quand λ devient grand.

La table 15 donne les efficacités relatives des différents paramètres pour N élevé.

λ	64	16	8	5	4	3
C_s	.25	.50	.71	.89	1	1.15
$e(\lambda)$.87	.59	.37	.22	.15	.09
$e(\alpha)$.88	.62	.42	.29	.24	.17
$e(m)$.87	.57	.33	.17	.10	.04

Table 15: Efficacité relative des paramètres α , λ , m de la loi Pearson III

Cas où m et α sont inconnus, λ étant fixé

Ici encore la comparaison n'est faite que pour $\lambda > 2$, en supposant les échantillons de taille élevée. On a en utilisant les relations (7.41) et (7.54) d'une part, (7.40) et (7.55) d'autre part:

$$e(\alpha) = \frac{\lambda}{\lambda+3} \quad (7.64)$$

$$e(m) = \frac{\lambda-2}{\lambda+1} \quad (7.65)$$

Ces efficacités tendent vers 1 quand λ est élevé. La table 16 indique les valeurs de quelques efficacités pour différentes valeurs de λ

λ	100	64	16	8	5	4	3
C_s	.20	.25	.50	.71	.89	1	1.15
$e(\alpha)$.97	.96	.84	.73	.63	.57	.50
$e(m)$.97	.95	.82	.67	.50	.40	.25

Table 16: Efficacité des paramètres de la loi Pearson III, λ étant fixé

Remarque:

Il est possible de faire des calculs similaires lorsque 1 ou 2 paramètres sont fixés et de comparer les méthodes des moments et du M.V. Certains de ces cas se ramènent à la loi Pearson III à 1 paramètre ou à la loi Gamma, et les autres cas présentent peu d'intérêt pratique tout au moins en hydrologie.

8. ESTIMATION DES PARAMETRES DE LA LOI LOG. PEARSON III

Nous envisagerons successivement le cas de la Loi Log.Gamma ($m=0$) et celui de la loi Log.Pearson III à 3 paramètres, dans le cas général d'une transformation logarithmique de base quelconque.

8.1 Estimation des paramètres de la Loi Log.Gamma

a. Propriétés générales d'estimation

La fonction densité de la Loi Log.Gamma est obtenue par la relation (2.1) avec $m=0$:

$$g(x) = \frac{|\alpha|k}{\Gamma(\lambda)} \frac{1}{x^{1+\alpha k}} (\alpha k \text{ Log } x)^{\lambda-1} \quad (8.1)$$

λ est positif et l'on a : $(\alpha \text{ Log } x) \geq 0$

La variate $y = \log_a x$ suit une loi Gamma, la transformation logarithmique est considérée dans la base a quelconque ($a > 1$).

On a : $\log_a x = k \text{ Log } x$ (Log = log. Népérien)

avec $k = \log_a e = \frac{1}{\text{Log } a}$ (k est > 0 puisque $a > 1$)

La fonction densité $g(x)$ peut se mettre sous la forme (6.19), en effet en tenant compte de l'inégalité $(\alpha \text{ Log } x) \geq 0$ on a :

$$g(x) = \exp \left[\left(\lambda \text{ Log } (k|\alpha|) - \text{Log } \Gamma(\lambda) \right) - (1 + \alpha k) \text{ Log } x + (\lambda-1) \text{ Log } \left| \text{Log } x \right| \right]$$

Il est donc théoriquement possible de trouver 2 statistiques conjointement exhaustives pour les paramètres α et λ , cette détermination est délicate en pratique lorsque α et λ varient.

b. Méthode du Maximum de vraisemblance

L'existence de statistiques conjointement exhaustives pour α et λ assure l'unicité de la solution obtenue par la méthode du maximum de vraisemblance. Cette solution (α_*, λ_*) est de plus

fonction du couple de statistiques conjointement exhaustives, mais α_* et λ_* ne sont pas conjointement exhaustifs et n'atteignent la variance minimale que de manière asymptotique.

La fonction vraisemblance L de la loi Log.Gamma pour l'échantillon (x_1, \dots, x_N) est telle que:

$$\text{Log } L = N\lambda \text{ Log } |\alpha|^k - N \text{ Log } \Gamma(\lambda) - (1+\alpha k) \sum_{i=1}^N \text{Log } x_i + (\lambda-1) \sum_{i=1}^N \text{Log} \left| \text{Log } x_i \right| \quad (8.2)$$

La solution α_* , λ_* est retenue par la résolution des équations:

$$\left\{ \begin{array}{l} \frac{\partial \text{Log } L}{\partial \alpha} = \frac{\lambda N}{\alpha} - k \sum_{i=1}^N \text{Log } x_i = 0 \\ \frac{\partial \text{Log } L}{\partial \lambda} = - \frac{d \text{Log } \Gamma(\lambda)}{d\lambda} + \text{Log } |\alpha|^k + \frac{1}{N} \sum_{i=1}^N \text{Log} \left| \text{Log } x_i \right| = 0 \end{array} \right.$$

En tenant compte de la relation entre y_i et x_i ($y_i = k \text{Log } x_i$),

le système précédent se met sous la forme:

$$\left\{ \begin{array}{l} \alpha = \frac{\lambda}{\frac{1}{N} \sum_{i=1}^N y_i} = \frac{\lambda}{\bar{y}} \end{array} \right. \quad (8.3)$$

$$\left\{ \begin{array}{l} \text{Log } \lambda - \frac{d \text{Log } \Gamma(\lambda)}{d\lambda} = - \frac{1}{N} \sum_{i=1}^N \text{Log} \frac{y_i}{\bar{y}} \end{array} \right. \quad (8.4)$$

avec $\bar{y} = \frac{1}{N} \sum_{i=1}^N y_i$

Ces équations sont celles du maximum de vraisemblance de la Loi Gamma appliquée à l'échantillon (y_1, \dots, y_N) (cf 7.2 b)

En pratique pour ajuster par le maximum de vraisemblance la loi Log.Gamma à un échantillon (x_1, \dots, x_N) , on ajuste la loi Gamma par le M.V. à l'échantillon (y_1, \dots, y_N) , où $y_i = \log_a x_i$. La solution (α_*, λ_*) est la même dans les 2 cas. Il s'ensuit que les propriétés asymptotiques des estimateurs α_* et λ_* déterminées dans le cas de la loi Gamma de variate y sont valables pour la loi Log.Gamma de variate x telle que $y = \log_a x$.

Ces estimateurs sont en particulier asymptotiquement efficaces ce qui constitue l'intérêt principal du M.V. pour de grands échantillons. Dans le cas de petits échantillons, comme ceux que l'on rencontre en hydrologie par exemple, la méthode du M.V. a un intérêt plus réduit puisqu'elle n'est plus optimale.

c. Méthode des moments

L'application de la méthode des moments à un échantillon (x_1, \dots, x_N) que l'on suppose tiré d'une loi Log.Gamma consiste à identifier les moments de l'échantillon (x_i) et ceux de la population Log.Gamma. On considère encore la transformation logarithmique dans une base a quelconque ($a > 1$).

Soient l_1 et l_2 les moments non centrés d'ordre 1 et 2 de l'échantillon (x_1, \dots, x_N) , en utilisant la relation (5.4) on obtient les équations permettant de déterminer les paramètres α et λ , on a :

$$\left\{ \begin{array}{l} \text{Log } l_1 = - \lambda \text{ Log } (1 - 1/\alpha k) \quad (8.5) \\ \text{Log } l_2 = - \lambda \text{ Log } (1 - 2/\alpha k) \quad (8.6) \end{array} \right.$$

k est défini par la base a choisie.

Connaissant l_1 et l_2 déduits de l'échantillon on peut alors déterminer α et λ à partir des équations. Lorsque α est positif, l'ajustement n'est possible que si $\alpha k > 2$ car autrement les moments ne sont pas définis, comme on l'a montré en 5.1 .

Variance des paramètres

Si la solution obtenue par la méthode des moments est (α_o, λ_o) , il est possible de déterminer la variance de ces paramètres pour de grands échantillons, on a en utilisant la relation générale (3.29):

$$\text{var } \alpha_o = \left(\frac{\partial \alpha}{\partial z_1} \right)^2 \text{var } z_1 + \left(\frac{\partial \alpha}{\partial z_2} \right)^2 \text{var } z_2 + 2 \left(\frac{\partial \alpha}{\partial z_1} \right) \left(\frac{\partial \alpha}{\partial z_2} \right) \text{Cov} \left(z_1, z_2 \right)$$

$$\text{var } \lambda_o = \left(\frac{\partial \lambda}{\partial z_1} \right)^2 \text{var } z_1 + \left(\frac{\partial \lambda}{\partial z_2} \right)^2 \text{var } z_2 + 2 \left(\frac{\partial \lambda}{\partial z_1} \right) \left(\frac{\partial \lambda}{\partial z_2} \right) \text{Cov} \left(z_1, z_2 \right)$$

La détermination des dérivées partielles de α et λ par rapport à z_1 et z_2 est obtenue à partir des équations (8.5) et (8.6), le calcul de la variance et de la covariance des moments z_1 et z_2 est effectuée par les relations (3.17) et 3.18).

En posant:

$$A_r = 1 - \frac{r}{\alpha k} \quad (r = 1, 2, 3, 4)$$

$$K = 2A_1 \text{Log } A_1 - A_2 \text{Log } A_2$$

$$M = A_1^\lambda \text{Log } A_2$$

$$N = A_2^\lambda \text{Log } A_1$$

$$P = 2A_1^{\lambda+1}$$

$$Q = A_2^{\lambda+1}$$

on obtient finalement:

$$\text{var } \alpha_o = \frac{\alpha^4 k^2}{N \lambda^2 K^2} \left(A_1 A_2 \right)^2 \left[\frac{M^2}{A_2^\lambda} + \frac{N^2}{A_1^\lambda} - 2 \frac{MN}{A_3^\lambda} - \left[\text{Log} \left(\frac{A_1}{A_2} \right) \right]^2 \right] \quad (8.7)$$

$$\text{var } \lambda_0 = \frac{1}{NK^2} \left[\frac{P^2}{A_2^\lambda} + \frac{Q^2}{A_2^\lambda} - 2 \frac{PQ}{A_3^\lambda} - \left(2 A_1 - A_2 \right)^2 \right] \quad (8.8)$$

Méthode pratique d'ajustement

La solution des équations (8.5) et (8.6) ne conduit pas à une forme explicite des paramètres α et λ , mais on peut remplacer ces équations par le système:

$$\left\{ \begin{array}{l} \frac{\text{Log } l_2}{\text{Log } l_1} = \frac{\text{Log } (1 - 2/\alpha k)}{\text{Log } (1 - 1/\alpha k)} \end{array} \right. \quad (8.9)$$

$$\left\{ \begin{array}{l} \lambda = - \frac{\text{Log } l_1}{\text{Log } (1 - 1/\alpha k)} \end{array} \right. \quad (8.10)$$

Nous considérons le cas de la transformation logarithme décimal ($a = 10$) on a alors $k = 1 / \text{Log } 10 \approx .434$.

A partir de l'échantillon on calcule $A = \frac{\text{Log } l_2}{\text{Log } l_1}$ et l'on déduit

α_0 à l'aide de la table 17, la relation (8.10) donne ensuite λ_0 .

Le cas $\alpha < 0$ pour lequel l'intervalle de définition de la variate est $(0,1)$ présentant moins d'intérêt en pratique, la table (17) a été établie pour $\alpha > 0$, de plus on doit avoir $\alpha > 4.60517$ pour que les 2 premiers moments de la Loi Log.Gamma existent.

Cette table peut en fait être utilisée quelle que soit la base a de la transformation logarithmique, en effet, si l'on connaît α_0 (pour la base 10) donné par la table 17, on détermine α_a pour la base a par:

$$\alpha_a = \alpha_0 \frac{\text{Log } a}{\text{Log } 10} \quad (8.11)$$

En particulier pour la transformation logarithme népérien, base e ,

$$\text{on a : } \alpha_e = \frac{\alpha_0}{\text{Log } 10} \approx .434 \alpha_0 \quad (8.12)$$

α	A	α	A	α	A
4.606	12.4417	4.648	6.85253	4.85	4.63763
4.607	11.3044	4.649	6.82121	4.86	4.59191
4.608	10.6787	4.65	6.79063	4.87	4.54825
4.609	10.2452	4.652	6.73155	4.88	4.50648
4.61	9.91356	4.654	6.67508	4.89	4.46647
4.611	9.645	4.656	6.62099	4.9	4.42809
4.612	9.41944	4.658	6.5691	4.91	4.39123
4.613	9.22508	4.66	6.51924	4.92	4.35579
4.614	9.05439	4.662	6.47128	4.93	4.32167
4.615	8.90228	4.664	6.42507	4.94	4.28879
4.616	8.76512	4.666	6.38049	4.95	4.25707
4.617	8.64028	4.668	6.33745	4.96	4.22645
4.618	8.52575	4.67	6.29585	4.97	4.19685
4.619	8.41997	4.672	6.25559	4.98	4.16823
4.62	8.32172	4.674	6.21659	4.99	4.14053
4.621	8.23002	4.676	6.17879	5	4.11369
4.622	8.14406	4.678	6.14211	5.01	4.08767
4.623	8.06317	4.68	6.1065	5.02	4.06244
4.624	7.98681	4.682	6.07189	5.03	4.03794
4.625	7.91449	4.684	6.03824	5.04	4.01415
4.626	7.84582	4.686	6.0055	5.05	3.99103
4.627	7.78047	4.688	5.97362	5.06	3.96855
4.628	7.71812	4.69	5.94256	5.07	3.94668
4.629	7.65852	4.692	5.91227	5.08	3.92538
4.63	7.60145	4.694	5.88274	5.09	3.90465
4.631	7.5467	4.696	5.85391	5.1	3.88444
4.632	7.4941	4.698	5.82577	5.11	3.86475
4.633	7.44348	4.7	5.79828	5.12	3.84554
4.634	7.39472	4.71	5.66966	5.13	3.8268
4.635	7.34767	4.72	5.55381	5.14	3.8085
4.636	7.30224	4.73	5.44857	5.15	3.79064
4.637	7.2583	4.74	5.35225	5.16	3.77319
4.638	7.21578	4.75	5.26357	5.17	3.75614
4.639	7.17458	4.76	5.18149	5.18	3.73948
4.64	7.13463	4.77	5.10515	5.19	3.72318
4.641	7.09586	4.78	5.03387	5.2	3.70724
4.642	7.0582	4.79	4.96706	5.21	3.69164
4.643	7.02159	4.8	4.90426	5.22	3.67636
4.644	6.98598	4.81	4.84504	5.23	3.66141
4.645	6.95131	4.82	4.78905	5.24	3.64677
4.646	6.91754	4.83	4.73598	5.25	3.63242
4.647	6.88463	4.84	4.68559	5.26	3.61837

TABLE 17 : Relation entre α et A

α	A	α	A	α	A
5.27	3.60459	5.69	3.19543	6.11	2.96262
5.28	3.59107	5.7	3.18845	6.12	2.95825
5.29	3.57782	5.71	3.18156	6.13	2.95392
5.3	3.56482	5.72	3.17476	6.14	2.94964
5.31	3.55207	5.73	3.16804	6.15	2.94539
5.32	3.53955	5.74	3.1614	6.16	2.94119
5.33	3.52726	5.75	3.15485	6.17	2.93703
5.34	3.5152	5.76	3.14838	6.18	2.9329
5.35	3.50335	5.77	3.14199	6.19	2.92882
5.36	3.49171	5.78	3.13557	6.2	2.92477
5.37	3.48028	5.79	3.12943	6.21	2.92076
5.38	3.46904	5.8	3.12327	6.22	2.91678
5.39	3.458	5.81	3.11717	6.23	2.91285
5.4	3.44714	5.82	3.11115	6.24	2.90895
5.41	3.43646	5.83	3.1052	6.25	2.90508
5.42	3.42596	5.84	3.09932	6.26	2.90125
5.43	3.41564	5.85	3.09351	6.27	2.89745
5.44	3.40548	5.86	3.08776	6.28	2.89369
5.45	3.39548	5.87	3.08208	6.29	2.88996
5.46	3.38565	5.88	3.07646	6.3	2.88627
5.47	3.37596	5.89	3.07091	6.31	2.8826
5.48	3.36643	5.9	3.06541	6.32	2.87897
5.49	3.35705	5.91	3.05998	6.33	2.87537
5.5	3.34781	5.92	3.05461	6.34	2.8718
5.51	3.33871	5.93	3.04929	6.35	2.86826
5.52	3.32974	5.94	3.04404	6.36	2.86475
5.53	3.32091	5.95	3.03884	6.37	2.86128
5.54	3.31221	5.96	3.03369	6.38	2.85783
5.55	3.30364	5.97	3.0286	6.39	2.85441
5.56	3.29519	5.98	3.02357	6.4	2.85102
5.57	3.28686	5.99	3.01859	6.41	2.84766
5.58	3.27865	6	3.01365	6.42	2.84432
5.59	3.27056	6.01	3.00877	6.43	2.84101
5.6	3.26257	6.02	3.00395	6.44	2.83774
5.61	3.2547	6.03	2.99917	6.45	2.83448
5.62	3.24694	6.04	2.99443	6.46	2.83126
5.63	3.23928	6.05	2.98975	6.47	2.82806
5.64	3.23173	6.06	2.98511	6.48	2.82488
5.65	3.22427	6.07	2.98052	6.49	2.82174
5.66	3.21692	6.08	2.97598	6.5	2.81861
5.67	3.20966	6.09	2.97148	6.51	2.81551
5.68	3.2025	6.1	2.96703	6.52	2.81244

TABLE 17 (suite)

α	A	α	A	α	A
6.53	2.80939	6.95	2.69992	7.74	2.55971
6.54	2.80637	6.96	2.69768	7.76	2.5569
6.55	2.80337	6.97	2.69547	7.78	2.55413
6.56	2.80039	6.98	2.69327	7.8	2.55138
6.57	2.79743	6.99	2.69108	7.82	2.54866
6.58	2.7945	7	2.6889	7.84	2.54597
6.59	2.79159	7.02	2.6846	7.86	2.54331
6.6	2.78871	7.04	2.68035	7.88	2.54067
6.61	2.78584	7.06	2.67616	7.9	2.53806
6.62	2.783	7.08	2.67202	7.92	2.53547
6.63	2.78018	7.1	2.66793	7.94	2.53291
6.64	2.77738	7.12	2.66389	7.96	2.53038
6.65	2.7746	7.14	2.6599	7.98	2.52787
6.66	2.77184	7.16	2.65597	8	2.52538
6.67	2.7691	7.18	2.65208	8.02	2.52292
6.68	2.76638	7.2	2.64824	8.04	2.52048
6.69	2.76369	7.22	2.64444	8.06	2.51807
6.7	2.76101	7.24	2.64069	8.08	2.51567
6.71	2.75835	7.26	2.63699	8.1	2.5133
6.72	2.75571	7.28	2.63333	8.12	2.51096
6.73	2.75309	7.3	2.62971	8.14	2.50863
6.74	2.75049	7.32	2.62614	8.16	2.50633
6.75	2.74791	7.34	2.62261	8.18	2.50404
6.76	2.74535	7.36	2.61912	8.2	2.50178
6.77	2.7428	7.38	2.61567	8.22	2.49954
6.78	2.74028	7.4	2.61225	8.24	2.49732
6.79	2.73777	7.42	2.60888	8.26	2.49512
6.8	2.73528	7.44	2.60555	8.28	2.49294
6.81	2.7328	7.46	2.60225	8.3	2.49078
6.82	2.73035	7.48	2.59899	8.32	2.48864
6.83	2.72791	7.5	2.59577	8.34	2.48651
6.84	2.72549	7.52	2.59258	8.36	2.48441
6.85	2.72308	7.54	2.58943	8.38	2.48232
6.86	2.72069	7.56	2.58631	8.4	2.48026
6.87	2.71832	7.58	2.58323	8.42	2.47821
6.88	2.71597	7.6	2.58018	8.44	2.47618
6.89	2.71363	7.62	2.57716	8.46	2.47416
6.9	2.7113	7.64	2.57417	8.48	2.47217
6.91	2.70899	7.66	2.57122	8.5	2.47019
6.92	2.7067	7.68	2.56829	8.52	2.46823
6.93	2.70443	7.7	2.5654	8.54	2.46628
6.94	2.70216	7.72	2.56254	8.56	2.46435

TABLE 17 (suite)

α	A	α	A	α	A
8.58	2.46244	9.44	2.39305	11.4	2.29343
8.6	2.46054	9.46	2.39169	11.5	2.2897
8.62	2.45866	9.48	2.39034	11.6	2.28606
8.64	2.45679	9.5	2.38899	11.7	2.28251
8.66	2.45494	9.52	2.38766	11.8	2.27905
8.68	2.45311	9.54	2.38633	11.9	2.27568
8.7	2.45129	9.56	2.38502	12	2.27239
8.72	2.44949	9.58	2.38371	12.1	2.26917
8.74	2.4477	9.6	2.38241	12.2	2.26603
8.76	2.44592	9.62	2.38112	12.3	2.26296
8.78	2.44416	9.64	2.37984	12.4	2.25997
8.8	2.44241	9.66	2.37857	12.5	2.25704
8.82	2.44068	9.68	2.37731	12.6	2.25418
8.84	2.43896	9.7	2.37605	12.7	2.25138
8.86	2.43725	9.72	2.37481	12.8	2.24864
8.88	2.43556	9.74	2.37357	12.9	2.24596
8.9	2.43388	9.76	2.37234	13	2.24334
8.92	2.43222	9.78	2.37112	13.1	2.24078
8.94	2.43056	9.8	2.36991	13.2	2.23827
8.96	2.42892	9.82	2.3687	13.3	2.23581
8.98	2.4273	9.84	2.36751	13.4	2.2334
9	2.42568	9.86	2.36632	13.5	2.23104
9.02	2.42408	9.88	2.36514	13.6	2.22873
9.04	2.42249	9.9	2.36396	13.7	2.22646
9.06	2.42091	9.92	2.3628	13.8	2.22424
9.08	2.41934	9.94	2.36164	13.9	2.22206
9.1	2.41779	9.96	2.36049	14	2.21993
9.12	2.41625	9.98	2.35934	14.1	2.21783
9.14	2.41472	10	2.35821	14.2	2.21578
9.16	2.41319	10.1	2.35708	14.3	2.21376
9.2	2.41169	10.2	2.35596	14.4	2.21179
9.22	2.41019	10.3	2.35485	14.5	2.20984
9.24	2.4087	10.4	2.35375	14.6	2.20794
9.26	2.40723	10.5	2.35266	14.7	2.20606
9.28	2.40576	10.6	2.35158	14.8	2.20423
9.3	2.40431	10.7	2.35051	14.9	2.20242
9.32	2.40287	10.8	2.34945	15	2.20065
9.34	2.40144	10.9	2.3484	15.1	2.1989
9.36	2.40001	11	2.34736	15.2	2.19719
9.38	2.3986	11.1	2.34632	15.3	2.19551
9.4	2.3972	11.2	2.34529	15.4	2.19385
9.42	2.39581	11.3	2.34427	15.5	2.19223
	2.39442				

TABLE 17 (suite)

α	A	α	A	α	A
15.6	2.19063	19.8	2.1413	24	2.11229
15.7	2.18905	19.9	2.14044	24.1	2.11175
15.8	2.18751	20	2.13958	24.2	2.11121
15.9	2.18598	20.1	2.13874	24.3	2.11067
16	2.18449	20.2	2.13791	24.4	2.11014
16.1	2.18301	20.3	2.13708	24.5	2.10961
16.2	2.18156	20.4	2.13627	24.6	2.10909
16.3	2.18014	20.5	2.13547	24.7	2.10858
16.4	2.17873	20.6	2.13467	24.8	2.10807
16.5	2.17735	20.7	2.13389	24.9	2.10756
16.6	2.17599	20.8	2.13311	25	2.10706
16.7	2.17465	20.9	2.13234	25.1	2.10657
16.8	2.17333	21	2.13159	25.2	2.10607
16.9	2.17203	21.1	2.13084	25.3	2.10559
17	2.17074	21.2	2.1301	25.4	2.1051
17.1	2.16948	21.3	2.12936	25.5	2.10463
17.2	2.16824	21.4	2.12864	25.6	2.10415
17.3	2.16701	21.5	2.12792	25.7	2.10368
17.4	2.16581	21.6	2.12721	25.8	2.10322
17.5	2.16462	21.7	2.12651	25.9	2.10275
17.6	2.16344	21.8	2.12582	26	2.1023
17.7	2.16229	21.9	2.12513	26.1	2.10184
17.8	2.16115	22	2.12446	26.2	2.1014
17.9	2.16002	22.1	2.12379	26.3	2.10095
18	2.15891	22.2	2.12312	26.4	2.10051
18.1	2.15782	22.3	2.12247	26.5	2.10007
18.2	2.15674	22.4	2.12182	26.6	2.09964
18.3	2.15568	22.5	2.12117	26.7	2.09921
18.4	2.15463	22.6	2.12054	26.8	2.09878
18.5	2.1536	22.7	2.11991	26.9	2.09836
18.6	2.15257	22.8	2.11929	27	2.09794
18.7	2.15157	22.9	2.11867	27.1	2.09752
18.8	2.15057	23	2.11806	27.2	2.09711
18.9	2.14959	23.1	2.11746	27.3	2.0967
19	2.14862	23.2	2.11686	27.4	2.0963
19.1	2.14767	23.3	2.11627	27.5	2.0959
19.2	2.14672	23.4	2.11568	27.6	2.0955
19.3	2.14579	23.5	2.1151	27.7	2.0951
19.4	2.14487	23.6	2.11453	27.8	2.09471
19.5	2.14396	23.7	2.11396	27.9	2.09432
19.6	2.14306	23.8	2.1134	28	2.09394
19.7	2.14218	23.9	2.11284	28.1	2.09356

TABLE 17 (suite)

α	A	α	A	α	A
28.2	2.09318	32.4	2.07963	95	2.02515
28.3	2.0928	32.5	2.07935	100	2.02385
28.4	2.09243	32.6	2.07908	125	2.01895
28.5	2.09206	32.7	2.07881	150	2.01571
28.6	2.09169	32.8	2.07854	175	2.01342
28.7	2.09133	32.9	2.07827	200	2.01172
28.8	2.09097	33	2.07801	250	2.00934
28.9	2.09061	33.1	2.07774	300	2.00776
29	2.09025	33.2	2.07748	400	2.00581
29.1	2.0899	33.3	2.07722	500	2.00464
29.2	2.08955	33.4	2.07696	600	2.00386
29.3	2.0892	33.5	2.07671	700	2.00331
29.4	2.08886	33.6	2.07645	800	2.00289
29.5	2.08851	33.7	2.0762	900	2.00257
29.6	2.08818	33.8	2.07595	1000	2.00231
29.7	2.08784	33.9	2.0757		
29.8	2.0875	34	2.07545		
29.9	2.08717	34.1	2.0752		
30	2.08684	34.2	2.07496		
30.1	2.08652	34.3	2.07471		
30.2	2.08619	34.4	2.07447		
30.3	2.08587	34.5	2.07423		
30.4	2.08555	34.6	2.07399		
30.5	2.08523	34.7	2.07376		
30.6	2.08492	34.8	2.07352		
30.7	2.08461	34.9	2.07329		
30.8	2.0843	35	2.07305		
30.9	2.08399	36	2.07081		
31	2.08368	37	2.06869		
31.1	2.08338	38	2.0667		
31.2	2.08308	39	2.06482		
31.3	2.08278	40	2.06305		
31.4	2.08248	45	2.05545		
31.5	2.08219	50	2.04949		
31.6	2.0819	55	2.04468		
31.7	2.08161	60	2.04073		
31.8	2.08132	65	2.03742		
31.9	2.08103	70	2.03461		
32	2.08075	75	2.03219		
32.1	2.08046	80	2.03009		
32.2	2.08018	85	2.02824		
32.3	2.0799	90	2.02661		

TABLE 17 (suite)

Remarque:

Lorsque la variate x suit une loi Log.Gamma, la variate $y = \log_a x$ suit une loi Gamma. Pour un échantillon de très grande taille, l'application de la méthode des moments à l'échantillon valeurs transformées y_i (cf 7.2 c) doit conduire aux valeurs α_0, λ_0 obtenues par l'application de la méthode des moments à l'échantillon des valeurs originales x_i . Pour un échantillon de taille réduite comme ceux que l'on rencontre en hydrologie ce n'est plus vrai. Il est préférable d'appliquer la méthode des moments à l'échantillon (x_i) car on donne ainsi le même poids à chaque observation, alors que la méthode des moments appliquée aux variables transformées y_i donne le même poids aux logarithmes des valeurs observées.

d. Efficacité de la méthode des moments

Pour mesurer l'efficacité de la méthode des moments on compare les variances asymptotiques des paramètres obtenues par la méthode des moments et par le M.V. Cette efficacité n'a un sens que pour de grands échantillons, puisque l'on sait que la méthode du M.V. est asymptotiquement optimale.

Si l'on pose $\beta = \alpha k$ on a:

$$\text{var } \beta = k^2 \text{ var } \alpha$$

La méthode du M.V. donne d'après (7.15)

$$\text{var } \beta_* = \frac{\beta^2 \psi'}{N \lambda \eta}$$

La méthode des moments donne d'après (8.7)

$$\text{var } \beta_o = \frac{\beta^4}{N \lambda^2 K^2} \left(A_1 A_2 \right)^2 \left[\frac{M^2}{A_2 \lambda} + \frac{N^2}{A_4 \lambda} - 2 \frac{MN}{A_3 \lambda} - \left[\text{Log} \left(\frac{A_1}{A_2} \right) \right]^2 \right] \quad (8.13)$$

On peut alors calculer:

$$e(\beta) = \frac{\text{var } \beta_*}{\text{var } \beta_o} \quad \text{on a } e(\beta) = e(\alpha)$$

Il est possible de la même manière de calculer $e(\lambda)$ à partir des relations (7.16) et (8.8).

Les efficacités des paramètres β et λ dépendent de λ et de β alors que dans le cas de la loi Gamma elles sont fonction de λ seulement. Le calcul de $e(\lambda)$ et $e(\beta)$ pour différentes valeurs de λ et β permet de tracer les figures 17 et 18.

Les courbes de ces 2 figures ont sensiblement la même forme pour les paramètres α et λ , il est à remarquer que pour λ fixé, l'efficacité croît notablement avec β (donc avec α).

Quand $\beta > 50$, l'efficacité ne croît plus très rapidement pour λ fixé, et elle est d'autant plus élevée que λ est grand.

Dans le cas d'un échantillon de taille suffisante, les figures 17 et 18 permettent rapidement de voir si un ajustement obtenu par la méthode des moments conduit à de bonnes estimations.

8.2 Estimation des paramètres de la loi Log.Pearson III

a. Propriétés générales d'estimation

La fonction densité de la loi Log.Pearson III à 3 paramètres donnée par (2.1) ne peut se mettre sous la forme de la relation (5.19). Il n'est donc pas possible de trouver des statistiques conjointement exhaustives des paramètres α, λ, m .

Si m est connu, on est ramené au cas de la loi Log.Gamma par le changement de variable $y = \left(\log_a x - m \right)$. L'ensemble des résultats du paragraphe précédent peuvent alors être appliqués à la variate y . Nous examinons dans ce qui suit, lorsque m est inconnu, les principales méthodes d'estimation des paramètres α, λ, m de la loi Log. Pearson III.

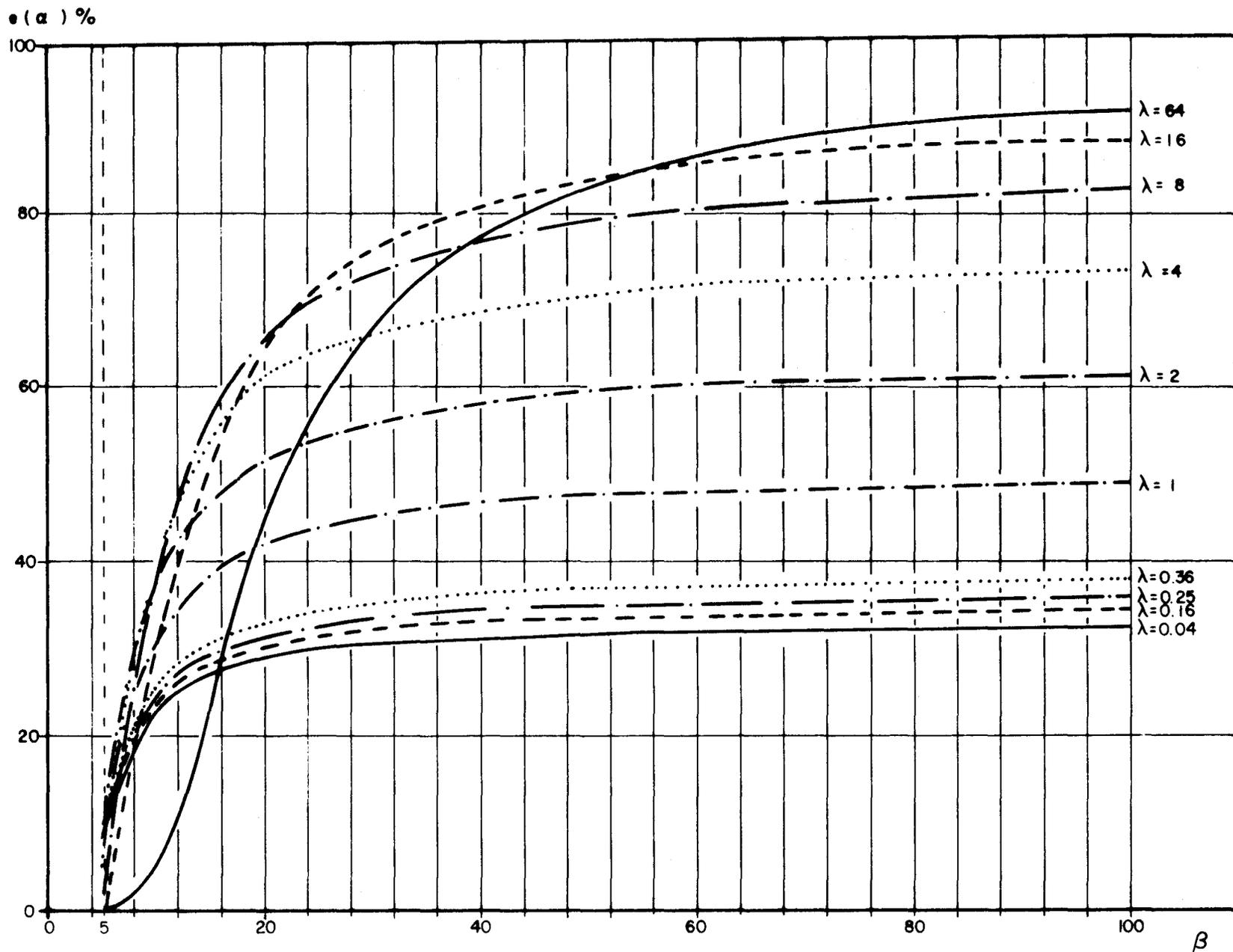


Fig. 17. Efficacité de la méthode des moments, loi log gamma. (Paramètre d'échelle α)

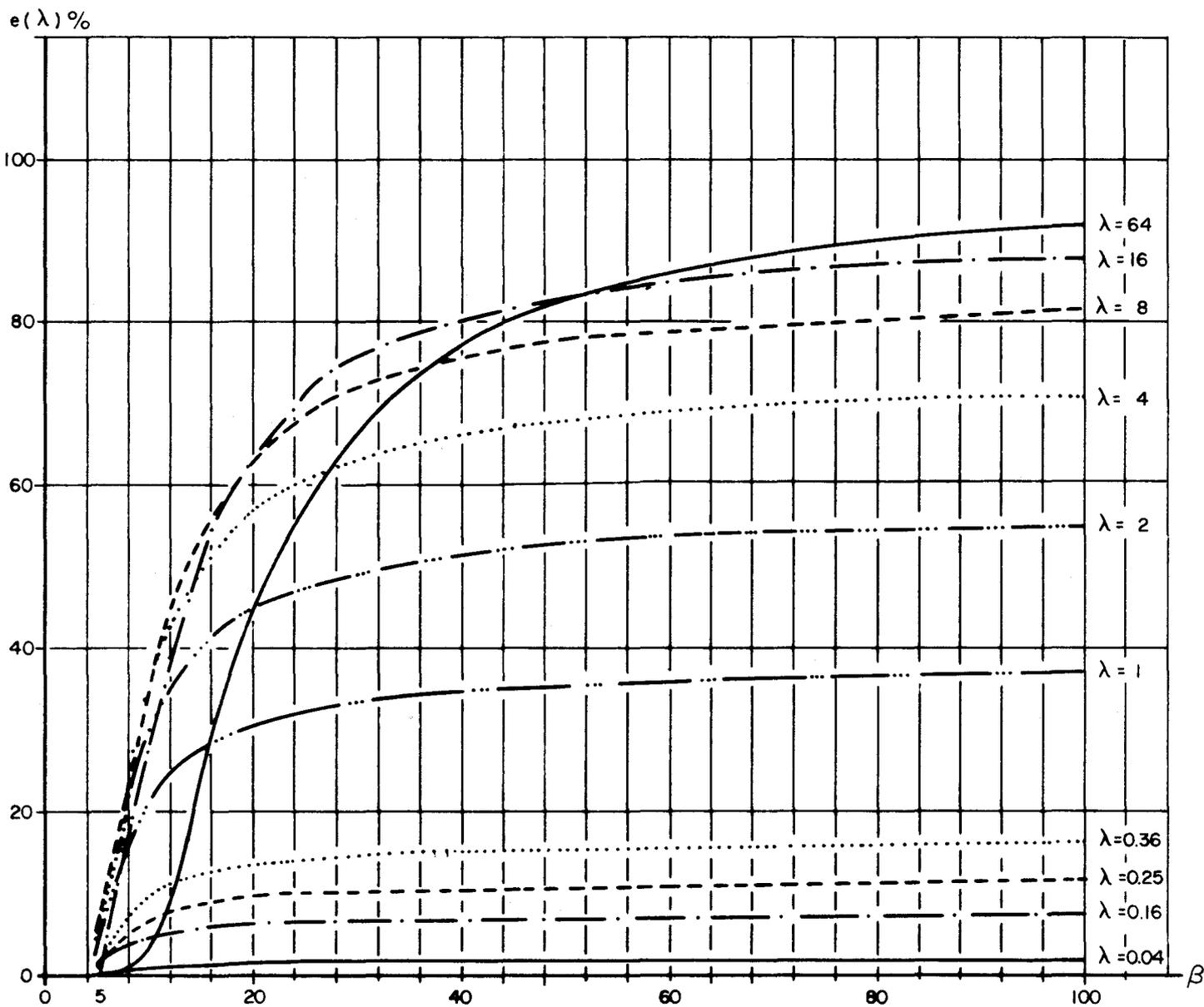


Fig. 18 . Efficacité de la méthode des moments, loi log-gamma .(Paramètre de forme λ)

b. Méthode du maximum de vraisemblance

Si l'on considère la fonction densité donnée par (2.1) les équations du M.V. sont:

$$\frac{\partial \text{Log } L}{\partial \alpha} = \frac{\lambda N}{\alpha} - \sum_{i=1}^N (k \text{ Log } x_i - m) = 0 \quad (8.14)$$

$$\frac{\partial \text{Log } L}{\partial \lambda} = -N \frac{d \text{Log } \Gamma(\lambda)}{d\lambda} + \sum_{i=1}^N \text{Log} \left[\alpha (k \text{ Log } x_i - m) \right] = 0 \quad (8.15)$$

$$\frac{\partial \text{Log } L}{\partial m} = N \alpha - (\lambda - 1) \sum_{i=1}^N \frac{1}{(k \text{ Log } x_i - m)} = 0 \quad (8.16)$$

Ces équations sont les mêmes que celles obtenues pour la loi Pearson III (cf 7.2 c) x étant remplacé par $(k \text{ Log } x)$. Il suffit donc de transformer l'échantillon des valeurs observées (x_1, \dots, x_N) pour obtenir l'échantillon (y_1, \dots, y_N) tel que $y_i = \log_a x_i = k \text{ Log } x_i$.

L'application de la méthode du M.V. à l'échantillon (y_1, \dots, y_N) que l'on suppose tiré d'une distribution Pearson III conduit à la solution $(\alpha_*, \lambda_*, m_*)$.

Les restrictions sur la méthode du M.V. décrites dans le cas de la loi Pearson III existent de la même manière pour la loi Log. Pearson III:

- Le M.V. conduit à des estimateurs asymptotiquement efficaces pour $\lambda > 1$
- Pour de petits échantillons la méthode du M.V. n'est pas optimale.
- Pour $\lambda < 1$, on ne peut définir de variance minimum même comme limite asymptotique.

Les variances asymptotiques des paramètres, pour $\lambda > 2$ sont données par les relations (7.37, 7.38 et 7.39).

c. Méthode des moments (W.R.C.)

Le Conseil des Ressources en Eau des Etats-Unis (Water Resources Council) a suggéré l'emploi généralisé de la distribution Log.Pearson III pour l'étude des crues annuelles. BENSON (1968) indique les raisons qui ont conduit à ce choix et décrit une méthode permettant la détermination des paramètres. La technique proposée consiste à appliquer la méthode des moments sur la série des logarithmes des valeurs observées, en supposant que cette série transformée est tirée d'une distribution Pearson III. Nous en décrivons les étapes:

- . Soient (x_1, \dots, x_N) l'échantillon des valeurs observées que l'on suppose tiré d'une distribution Log.Pearson III
- . on en déduit l'échantillon des valeurs transformées (y_1, \dots, y_N) tel que $y_i = \log x_i$
(Nous considérons ici la transformation logarithme décimal, mais la méthode est générale quelle que soit la base des logarithmes utilisée.)
- . on peut calculer les moments de l'échantillon des y_i
 - . la moyenne M est donnée par (3.7)
 - . la variance S^2 non biaisée est donnée par (3.23)
- . Une estimation du coefficient d'asymétrie est obtenue par la relation (3.25), qui donne \hat{C}_s
- . La méthode des moments permet alors de déterminer les paramètres α, λ, m , en considérant que la série des y_i suit une distribution Pearson III.
- . Cette méthode d'ajustement est systématiquement utilisée par les agences fédérales américaines pour l'étude des débits de crue. L'évaluation d'un débit de crue Q_T de période de retour T , c'est-à-dire ayant une probabilité au dépassement $1/T$, est effectuée par la relation:

$$Q_T = 10^{(M + K S)} \quad (8.17)$$

K est la valeur de la variate Pearson III standardisée correspondante à une probabilité au dépassement $\frac{1}{T}$, et à une asymétrie \hat{C}_s , qui est donnée par la table de HARTER (1969).

Plusieurs critiques peuvent être adressées à cette technique d'ajustement:

- . Le coefficient d'asymétrie de l'échantillon est déterminé par la relation (3.25), qui conduit en général à une estimation assez fortement biaisée, (WALLIS et al 1974).
- . La détermination des paramètres est effectuée en donnant le même poids aux logarithmes des valeurs observées ce qui a pour effet de diminuer l'influence des valeurs extrêmes, et de conduire à un biais dans l'estimation des paramètres. En effet, lorsque la variate x suit une distribution Log.Pearson III de paramètres α_0, λ_0, m_0 , la population des logarithmes de x suit une distribution Pearson III de paramètres α_0, λ_0, m_0 , mais quand l'on considère un échantillon de taille N l'application de la méthode des moments sur la série observée d'une part, et sur la série des logarithmes des valeurs observées d'autre part, conduisent à des estimations différentes des paramètres. Ceci revient à dire que l'on doit choisir entre conserver les moments de la série observée, ou conserver les moments de la série des logarithmes des valeurs observées (ce que fait la méthode W.R.C. qui a été décrite).

Cette dernière critique en particulier, nous a conduit à proposer une autre technique qui conserve les moments de la série observée et donne le même poids à chaque observation de cette série.

d. Méthode des moments appliquée à la série observée

Les moments non centrés de la distribution Log.Pearson III sont donnés par la relation (5.2), lorsque l'on considère la transformation logarithmique dans une base quelconque a .

Pour appliquer la méthode des moments on estime les 3 premiers moments non centrés de la population par ceux de l'échantillon l_1, l_2, l_3 ; on obtient ainsi les 3 relations permettant de déterminer les paramètres α, λ et m (BOBEE 1975).

$$l_1 = \frac{e^{m/k}}{\left(1 - \frac{1}{\beta}\right)^\lambda} \quad (8.18)$$

$$l_2 = \frac{e^{2m/k}}{\left(1 - \frac{2}{\beta}\right)^\lambda} \quad (8.19)$$

$$l_3 = \frac{e^{3m/k}}{\left(1 - \frac{3}{\beta}\right)^\lambda} \quad (8.20)$$

. k est déterminé par la base a de la transformation logarithmique choisie ($k = 1/\text{Log } a$)

. Connaissant β et k on en déduit $\alpha = \beta/k$

Le système précédent peut encore s'écrire :

$$\frac{\text{Log} \left[\left(1 - \frac{1}{\beta}\right)^3 / \left(1 - \frac{3}{\beta}\right) \right]}{\text{Log} \left[\left(1 - \frac{1}{\beta}\right)^2 / \left(1 - \frac{2}{\beta}\right) \right]} = \frac{\text{Log } l_3 - 3 \text{Log } l_1}{\text{Log } l_2 - 2 \text{Log } l_1} \quad (8.21)$$

$$\lambda = \frac{\text{Log } l_2 - 2 \text{Log } l_1}{\text{Log} \left[\left(1 - \frac{1}{\beta}\right)^2 / \left(1 - \frac{2}{\beta}\right) \right]} \quad (8.22)$$

$$m = k \left[\text{Log } l_1 + \lambda \text{Log} \left(1 - \frac{1}{\beta}\right) \right] \quad (8.23)$$

Lorsque l_1, l_2, l_3 sont calculés à partir de l'échantillon, la relation (8.21) permet de déterminer β_0 , on calcule ensuite λ_0 par (8.22) et m_0 par (8.23).

On obtient ainsi la solution $\alpha_0 = \frac{\beta_0}{k}$, λ_0 , m_0 de l'ajustement de la loi Log.Pearson par la méthode des moments.

La méthode proposée détermine la solution α_0, λ_0, m_0 à partir des moments de la série observée (et non comme la méthode précédente à partir des moments de la série des logarithmes des valeurs observées).

Cette méthode est valable quel que soit le signe du paramètre α , on doit cependant remarquer que pour $0 < \beta \leq 3$, il est impossible de faire l'ajustement décrit puisque les moments non centrés jusqu'à l'ordre 3 ne sont pas tous définis.

L'équation (8.21) est implicite en β et peut être résolue par approximations successives, il est cependant plus pratique d'établir des tables donnant β lorsque l'on connaît B tel que:

$$B = \frac{\text{Log } l_3 - 3 \text{ Log } l_1}{\text{Log } l_2 - 2 \text{ Log } l_1} \quad (8.24)$$

La table (18) permet de déterminer β_0 connaissant B à partir de l'échantillon.

- . Dans le cas particulier où c'est le logarithme népérien de la variate qui suit une distribution Pearson III on a $k = 1$, et la valeur α est directement donnée par la table 18.
- . Dans le cas où l'on considère la transformation logarithme décimal (base 10), on a $k = 1/\text{Log } 10 \approx .434$, et l'on détermine α_0 par $\alpha_0 = \beta_0 \text{ Log } 10$

3.001	23.6326	5.5	3.68151	31	3.0720	72	3.02878
3.002	21.2479	6.0	3.58097	32	3.06782	74	3.02797
3.003	19.8610	6.5	3.50651	33	3.06559	76	3.02721
3.004	18.8823	7.0	3.44908	34	3.06351	78	3.02649
3.005	18.1268	7.5	3.40343	35	3.06155	80	3.02581
3.006	17.5125	8.0	3.36624	36	3.05971	82	3.02516
3.007	16.9954	8.5	3.33536	37	3.05798	84	3.02454
3.008	16.5495	9.0	3.30930	38	3.05635	86	3.02395
3.009	16.1578	9.5	3.28701	39	3.05480	88	3.02339
3.010	15.8088	10	3.26773	40	3.05334	90	3.02286
3.02	13.5550	11	3.23603	41	3.05196	92	3.02235
3.03	12.2811	12	3.21106	42	3.05064	94	3.02186
3.04	11.4038	13	3.19087	43	3.04939	96	3.02139
3.05	10.7415	14	3.17422	44	3.04820	98	3.02094
3.06	10.2137	15	3.16024	45	3.04707	100	3.02051
3.07	9.77777	16	3.14834	46	3.04598	125	3.01633
3.08	9.40848	17	3.13809	47	3.04495	150	3.01356
3.09	9.08955	18	3.12916	48	3.04396	175	3.01159
3.1	8.80997	19	3.12132	49	3.04302	200	3.01013
3.2	7.12276	20	3.11437	50	3.04211	250	3.00808
3.3	6.27521	21	3.10818	52	3.04041	300	3.00672
3.4	5.74308	22	3.10263	54	3.03884	350	3.00576
3.5	5.37124	23	3.09762	56	3.03739	400	3.00503
3.6	5.09411	24	3.09307	58	3.03604	450	3.00447
3.7	4.87836	25	3.08893	60	3.03479	500	3.00402
3.8	4.70498	26	3.08514	62	3.03362	600	3.00335
3.9	4.56225	27	3.08167	64	3.03252	700	3.00287
4.0	4.44247	28	3.07846	66	3.03150	800	3.00251
4.5	4.04742	29	3.07550	68	3.03054	900	3.00223
5.0	3.82500	30	3.07275	70	3.02963	1000	3.00201

TABLE 18: Relation entre β et B

-0.001	2.04623	-2.5	2.59470	-31	2.94028	-72	2.97315
-0.002	2.05195	-3.0	2.63252	-32	2.94202	-74	2.97386
-0.003	2.05599	-3.5	2.66379	-33	2.94365	-76	2.97452
-0.004	2.05923	-4	2.69009	-34	2.94519	-78	2.97515
-0.005	2.06201	-5	2.73194	-35	2.94666	-80	2.97576
-0.006	2.06446	-6	2.76377	-36	2.94804	-82	2.97633
-0.007	2.06669	-7	2.78881	-37	2.94936	-84	2.97688
-0.008	2.06873	-8	2.80904	-38	2.95061	-86	2.97740
-0.009	2.07063	-9	2.82572	-39	2.95180	-88	2.97790
-0.01	2.07242	-10	2.83972	-40	2.95293	-90	2.97838
-0.02	2.08654	-11	2.85163	-41	2.95402	-92	2.97884
-0.03	2.09727	-12	2.86189	-42	2.95505	-94	2.97927
-0.04	2.10634	-13	2.87082	-43	2.95604	-96	2.97969
-0.05	2.11440	-14	2.87867	-44	2.95698	-98	2.98010
-0.06	2.12175	-15	2.88562	-45	2.95789	-100	2.98049
-0.07	2.12857	-16	2.89181	-46	2.95876	-125	2.98431
-0.08	2.13498	-17	2.89736	-47	2.95959	-150	2.98689
-0.09	2.14105	-18	2.90238	-48	2.96039	-175	2.98873
-0.1	2.14684	-19	2.90692	-49	2.96116	-200	2.99012
-0.2	2.19521	-20	2.91107	-50	2.96190	-250	2.99208
-0.3	2.23391	-21	2.91485	-52	2.96330	-300	2.99339
-0.4	2.26716	-22	2.91833	-54	2.96460	-350	2.99433
-0.5	2.29663	-23	2.92154	-56	2.96581	-400	2.99503
-0.6	2.32320	-24	2.92450	-58	2.96694	-450	2.99558
-0.7	2.34742	-25	2.92725	-60	2.96800	-500	2.99602
-0.8	2.36969	-26	2.92980	-62	2.96899	-600	2.99668
-0.9	2.39028	-27	2.93218	-64	2.96992	-700	2.99715
-1	2.40942	-28	2.93441	-66	2.97080	-800	2.99751
-1.5	2.48842	-29	2.93649	-68	2.97163	-900	2.99778
-2	2.54794	-30	2.93845	-70	2.97241	-1000	2.99800

TABLE 18 (suite) : Relation entre β et B

La table 19 a été spécialement construite pour ce cas fréquent en pratique, elle permet de déduire directement α_0 connaissant B.

La relation (8.23) s'écrit pour $a \approx 10$:

$$m = \log \left[Z_1 \left(1 - \frac{1}{\beta} \right)^\lambda \right] \quad (8.25)$$

Connaissant α_0 , λ_0 et m_0 il est possible de déterminer les moments de la distribution Pearson III. Ces moments ne sont pas égaux à ceux de l'échantillon des logarithmes des valeurs observées.

On peut en particulier calculer la moyenne, la variance et le coefficient d'asymétrie et en déduire la valeur d'un évènement X_T de période de retour T. Dans le cas d'une transformation log décimal par exemple on a :

$$\log X_T = \left(m_0 + \frac{\lambda_0}{\alpha_0} \right) + K \frac{\sqrt{\lambda}}{|\alpha_0|}$$

ou $X_T = 10^{\left(m_0 + \frac{\lambda_0}{\alpha_0} + K \frac{\sqrt{\lambda}}{|\alpha_0|} \right)} \quad (8.26)$

K est la variable standardisée Pearson III fonction de T et du coefficient d'asymétrie

$$C_s = \frac{\alpha_0}{|\alpha_0|} \frac{2}{\sqrt{\lambda_0}}$$

Remarque:

Lorsque $|\beta|$ est assez élevé on peut déterminer ce paramètre directement par la formule approchée

$$\beta_0 = \frac{1 + \sqrt{5B - 14}}{B - 3} \quad (8.27)$$

Pour $|\beta| > 30$, la détermination est effectuée avec une erreur relative inférieure à 1%.

6.910	23.7204	7.13	8.90335	28	3.20754	68	3.07401
6.912	21.5275	7.14	8.78742	29	3.19857	70	3.07171
6.914	20.2055	7.15	8.67731	30	3.19035	72	3.06954
6.916	19.2585	7.2	8.19863	31	3.18278	74	3.06750
6.918	18.5215	7.3	7.48802	32	3.17579	76	3.06558
6.920	17.9189	7.4	6.97594	33	3.16932	78	3.06376
6.922	17.4097	7.5	6.58378	34	3.16330	80	3.06204
6.924	16.9693	7.6	6.27107	35	3.15770	82	3.06041
6.926	16.5815	7.7	6.01435	36	3.15248	84	3.05887
6.928	16.2353	7.8	5.79891	37	3.14758	86	3.05740
6.93	15.9230	8	5.45573	38	3.14300	88	3.05600
6.94	14.7039	9	4.55107	39	3.13868	90	3.05467
6.95	13.8309	10	4.14523	40	3.13463	92	3.05341
6.96	13.1544	11	3.91092	41	3.13080	94	3.05220
6.97	12.6045	12	3.75741	42	3.12718	96	3.05104
6.98	12.1428	13	3.64872	43	3.12376	98	3.04993
6.99	11.7461	14	3.56759	44	3.12052	100	3.04887
7.00	11.3992	15	3.50465	45	3.11744	125	3.03862
7.01	11.0916	16	3.45436	46	3.11452	150	3.03193
7.02	10.8159	17	3.41325	47	3.11174	175	3.02721
7.03	10.5665	18	3.37900	48	3.10909	200	3.02371
7.04	10.3391	19	3.35002	50	3.10415	250	3.01886
7.05	10.1304	20	3.32517	52	3.09964	300	3.01565
7.06	9.9379	21	3.30363	54	3.09551	400	3.01168
7.07	9.7594	22	3.28477	56	3.09170	500	3.00932
7.08	9.59317	23	3.26813	58	3.08819	600	3.00775
7.09	9.43780	24	3.25333	60	3.08493	700	3.00663
7.10	9.29207	25	3.24008	62	3.08191	800	3.00580
7.11	9.15498	26	3.22816	64	3.07910	900	3.00515
7.12	9.02565	27	3.21736	66	3.07647	1000	3.00463

TABLE 19 : Relation entre α et B (logarithme décimaux).

α	B	α	B	α	B	α	B
-0.001	2.04079	-2.5	2.42480	-29	2.86735	-70	2.93920
-0.002	2.04521	-3.0	2.46013	-30	2.87106	-72	2.94076
-0.003	2.04826	-3.5	2.49112	-31	2.87458	-74	2.94225
-0.004	2.05068	-4.0	2.51858	-32	2.87791	-76	2.94366
-0.005	2.05273	-4.5	2.54311	-33	2.88106	-78	2.94501
-0.006	2.05452	-5.0	2.56519	-34	2.88406	-80	2.94629
-0.007	2.05613	-5.5	2.58517	-35	2.88691	-82	2.94751
-0.008	2.05760	-6.0	2.60335	-36	2.88962	-84	2.94868
-0.009	2.05896	-7.0	2.63524	-37	2.89221	-86	2.94980
-0.01	2.06023	-8.0	2.66231	-38	2.89468	-88	2.95087
-0.02	2.07005	-9	2.68560	-39	2.89703	-90	2.95190
-0.03	2.07728	-10	2.70585	-40	2.89929	-92	2.95288
-0.04	2.08329	-11	2.72363	-41	2.90144	-94	2.95383
-0.05	2.08855	-12	2.73937	-42	2.90351	-96	2.95474
-0.06	2.09329	-13	2.75340	-43	2.90549	-98	2.95561
-0.07	2.09765	-14	2.76599	-44	2.90739	-100	2.95645
-0.08	2.10173	-15	2.77735	-45	2.90922	-125	2.96478
-0.09	2.10556	-16	2.78766	-46	2.91098	-150	2.97043
-0.1	2.10920	-17	2.79705	-47	2.91267	-175	2.97452
-0.2	2.13918	-18	2.80564	-48	2.91429	-200	2.97762
-0.3	2.16304	-19	2.81353	-50	2.91737	-250	2.98199
-0.4	2.18374	-20	2.82081	-52	2.92023	-300	2.98494
-0.5	2.20235	-21	2.82753	-54	2.92290	-350	2.98706
-0.6	2.21944	-22	2.83377	-56	2.92540	-400	2.98865
-0.7	2.23533	-23	2.83958	-58	2.92774	-500	2.99089
-0.8	2.25023	-24	2.84499	-60	2.92994	-600	2.99240
-0.9	2.26430	-25	2.85004	-62	2.93201	-700	2.99347
-1	2.27764	-26	2.85478	-64	2.93396	-800	2.99428
-1.5	2.33592	-27	2.85922	-66	2.93581	-900	2.99492
-2	2.38398	-28	2.86341	-68	2.93755	-1000	2.99542

TABLE 19 (suite): Relation entre α et B (logarithmes décimaux)

e. Comparaison des différentes méthodes

Pour cette comparaison il est nécessaire de considérer la taille de l'échantillon à partir duquel on effectue l'ajustement. Pour de très grands échantillons on sait en effet que la méthode du Maximum de vraisemblance devient optimale et est donc préférable à la méthode des moments qu'elle soit appliquée directement aux valeurs observées (8.2 d) ou aux logarithmes des valeurs observées (8.2 c).

Il est d'ailleurs possible d'utiliser les résultats de (7.3 d) pour comparer la méthode du Maximum de vraisemblance et la méthode des moments appliquées aux logarithmes des valeurs observées. Dans ces deux méthodes on suppose en effet que l'échantillon du logarithme des valeurs observées est tiré d'une distribution Pearson III.

En ce qui concerne la méthode des moments appliquée directement à la série des valeurs observées (8.2 d) le calcul des variances d'échantillonnage des paramètres α , λ , m est possible mais très laborieux. Ce calcul n'est valable que pour des échantillons de taille élevée et présente un intérêt réduit puisque dans ce cas l'utilisation de la méthode du Maximum de vraisemblance est préférable.

En pratique les échantillons dont on dispose sont de taille réduite et quelle que soit la méthode d'ajustement utilisée il n'est pas possible de déterminer précisément les variances d'échantillonnage des paramètres. En particulier dans le cas du Maximum de vraisemblance les variances d'échantillonnage calculées en 7.3 b sont des limites asymptotiques, et l'optimalité de la méthode n'est plus vraie pour de petits échantillons.

En ce qui concerne l'application de la méthode des moments à des échantillons que l'on suppose tirés d'une distribution Log. Pearson III, il est préférable, si l'on veut faire jouer un rôle

identique à chaque individu de l'échantillon, de donner le même poids aux valeurs observées plutôt qu'aux logarithmes des valeurs observées. C'est pourquoi la méthode des moments (8.2 d) appliquée directement à la série observée devrait conduire à un meilleur ajustement que la méthode des moments (8.2 c) appliquée à la série des logarithmes des valeurs observées. Dans le premier cas on conserve les moments de la série observée alors que dans le second on conserve les moments de la série transformée obtenue en prenant le logarithme de chaque valeur observée.

Pour des échantillons de taille réduite, comme ceux que l'on rencontre en hydrologie pour l'étude des débits de crue annuels par exemple, on observe un grand écart entre les résultats obtenus par ces deux méthodes.

De manière générale, la comparaison des différentes méthodes pour l'ajustement de la loi Log. Pearson III à des petits échantillons devrait être effectuée par simulation. Une telle étude reste à faire et devrait s'appuyer sur un algorithme de génération d'une variate Log. Pearson III de très bonne qualité et bien testée. A partir d'un tel algorithme il serait possible de générer un très grand nombre d'échantillons de taille donnée, tirés d'une distribution Log. Pearson III de paramètres fixés et ensuite d'appliquer à chaque échantillon les différentes méthodes d'ajustement.

REFERENCES BIBLIOGRAPHIQUES

- AITCHISON J. and J.A.C. BROWN (1957). The Log. Normal distribution. Cambridge University Press.
- APPEL C.A. (1968). A note on the average probability of extreme events. Water Resources Research 4(6) p. 1359.
- BENSON M.A. (1967). Average probability of extreme events. Water Resources Research 3(1) p. 225.
- BENSON M.A. (1968) Uniform flood frequency estimating methods for federal agencies. Water Resources Research 4(5) pp. 891-908.
- BERMAN M.B. (1971) Generating Gamma distributed variates for computer simulation models. Rep. R.641 PR Rand, Santa Monica Calif.
- BOBEE B. et G. MORIN (1972 a). Extension de la loi Pearson III: cas de l'asymétrie négative. Rapport INRS-Eau No. 5 Université du Québec, Québec.
- BOBEE B. et G. MORIN (1972 b) Détermination des intervalles de confiance de la loi Pearson III par les statistiques d'ordre. Rapport INRS-Eau No. 6 Université du Québec, Québec.
- BOBEE B. (1973 a) Sample error of T-Year events computed by fitting a Pearson type 3 distribution. Water Resources Research 9(5) pp. 1264 - 1270.
- BOBEE B. (1973 b) Etude de la loi Log. Pearson III et ses applications. Rapport INRS-Eau No. 22 Université du Québec, Québec.
- BOBEE B. et G. MORIN (1973 a). Statistique d'ordre de la loi Pearson III et de sa forme dérivée à asymétrie négative. Revue de Statistique Appliquée Vol. XXI No. 4 pp. 69-80.
- BOBEE B. et G. MORIN (1973 b). Détermination des intervalles de confiance de la loi Pearson III par les statistiques d'ordre. Journal of Hydrology 20, pp. 137-153

- BOBEE B. et J.P. VILLENEUVE (1973). Précision des variables hydrologiques. Bulletin des Sciences Hydrologiques (IAHS) Vol. 18 No. 1 pp. 7-13.
- BOBEE B. (1975). The Log. Pearson type 3 distribution and its application in hydrology. Water Resources Research (accepté pour publication).
- BOBEE B. and R. ROBITAILLE (1975). Correction of bias in the estimation of the coefficient of skewness. Water Resources Research (accepté pour publication).
- BRUNET-MORET Y. Etude de quelques lois statistiques utilisées en hydrologie. Cahier ORSTOM, série hydrologie, Vol. VI, No. 3.
- BRUNET-MORET Y. (1971). Fonctions de distribution en développement limité. Cahier ORSTOM, série hydrologie, Vol. VIII, No. 3.
- BRUNET-MORET Y. (1973 a). Estimation de paramètres. Cahier ORSTOM, série hydrologie, Vol. X, No. 2 pp. 111-132.
- BRUNET-MORET Y. (1973 b). Statistique de rangs. Cahier ORSTOM, série hydrologie, Vol. X, No. 2.
- CRAIG C.C. (1935). A new exposition and chart for the Pearson system of frequency curves. The Annals of Mathematical Statistics, Vol. VI.
- CRAMER H. (1971). Mathematical methods of statistics. Princeton University Press, 12^e Edition.
- DAVIS H.T. (1933). Tables of higher mathematical functions Vol. 1 & 2 Princepia Press, Bloomington.
- DUMAS de RAULY D. (1968). L'estimation statistique. Gauthier-Villars, Paris.
- ELDERTON W.P. (1938). Frequency curves and correlation. 3^e Edition. Cambridge University Press.
- FISHER R.A. (1931). The moments of the distribution for normal samples of measures of departure from normality. Proc. Roy. Soc. London (A) 130, pp. 16-28.

- FISHER R.A. and E.A. CORNISH (1960). The percentile points of distribution having known cumulants. *Technometrics* 2, p. 209.
- GLADWELL J.S. and Lin CHENG-NAN (1969). Confidence limits determined using order statistics. *Water Resources Research* 5(5) pp. 1120-1123.
- GREENWOOD J.A. and D. DURAND (1960). Aids for fitting the Gamma distribution by Maximum Likelihood. *Technometrics* Vol. 2, No. 1.
- GUPTA S.S. (1960). Order statistics from the Gamma distribution. *Technometrics* Vol. 2, pp. 243-262.
- HARDISON C.H. (1969). Accuracy of streamflow characteristics. U.S. Geol. Surv. Prof. Pap. 650D, D210 - D 214.
- HARTER H.L. (1969). A new table of percentage points of the Pearson type 3 distribution. *Technometrics* 2(1) pp. 177-187.
- HOADLEY A.B. (1968). Use of the Pearson densities for approximating a skew density whose left terminal and first three moments are known. *Biometrika* 55, 3 p. 559.
- JÖHNK M.D. (1964). Erzeugung von betaverteilten und gamma verteilten Zufallszahlen. *Metrika* 8(1), pp. 5-15.
- KENDALL M.G. and A. STUART (1967). *The Advanced Theory of Statistics* T. 2 2nd Edition. Hafner Publishing Company, N.Y.
- KENDALL M.G. and A. STUART (1969). *The Advanced Theory of Statistics* T. 1 3rd Edition. Hafner Publishing Company, N.Y.
- KIRBY W. (1972). Computer-oriented Wilson-Hilferty transformation that preserves the first three moments and the lower bound of the Pearson type 3 distribution. *Water Resources Research* 8(5). pp. 1251-1254.
- KIRBY W. (1974). Algebraic Boundedness of Sample Statistics. *Water Resources Research* 10(2) pp. 220-222.

- KRITSKILY S.N. and M.F. MENKEL (1969). Fundamental principles of the method for computing maximum runoff. Soviet Hydrology, selected papers No. 1.
- MASUYAMA M. and Y. KUROIWA (1951). Table for the likelihood solution of Gamma distribution and its medical applications. Univ. of Japan. Sci. and Eng. Rev. of Stat. Appl. Res. 1(1).
- MATALAS N.C. and M.A. BENSON (1968). Note on the standard error of the coefficient of skewness. Water Resources Research 4(1). pp. 204-205.
- MATALAS N.C. and WALLIS J.R. (1973). Eureka! It fits a Pearson type 3 distribution. Water Resources Research 9(2). pp. 281-289.
- NASH J.E. and J. AMOROCHO (1966). The accuracy of the prediction of floods of high return period. Water Resources Research 2(2) pp. 191-198.
- RAO C.R. (1952). Advanced statistical methods in biometric research. J. Wiley N.Y.
- REMINIERAS G. (1967). Statistical methods of flood frequency analysis. Water Resources series No. 30 W.M.O.
- THOM H.C.S. (1958). A Note on the Gamma Distribution. Monthly weather review 86(4) pp. 117-122.
- WALLIS J.R. et al (1974). Just a moment! Water Resources Research 10(2) pp. 211-222.
- WILSON E.B. and M.M. HILFERTY (1931). The distribution of chi-square Proc. Nat. Acad. Sci. U.S.A. 17, pp. 684-688.
- YEVDJEVICH V.M. (1966). Stochastic problems in the design of reservoirs. Water Research, John Hopkins press for Resources for the Future Baltimore Md. pp. 375-411. Kneese and Smith Eds.

BIBLIOGRAPHIE DES APPLICATIONS

- BLOKHINOV Y.G. (1969). Distribution of estimates of quantiles computed by different methods (as applied to streamflow computations) Soviet Hydrology, Selected Papers No. 1.
- BOIS P. (1968). Contribution à l'étude de la variabilité interannuelle des débits de rivières. Thèse de doctorat. Faculté des Sciences, Université de Grenoble (FRANCE).
- COBB E.D. (1968). Skew parameter for Log. Pearson type III frequency distribution. Water Resources Div. Bull. Avril 1968. p. 37.
- CRIPPEN R.G. (1968). Composite Log. type III frequency magnitude curve of annual flood. Wat. Res. Div. Bull. Avril 1968. pp. 32-36.
- CSOMA J. (1969). An estimate of discharge probabilities. Publication No. 84, AIHS Les crues et leur évaluation.
- DAS S.C. (1955). The fitting of truncated type III curves to daily rainfall data. Australian Journal of Physics. Vol. 8 pp. 298-304.
- DAVIS D. et al. (1972). Uncertainty in the return period of maximum events: A Bayesian Approach. Int. Symp. on uncertainties in hydrologic and water resources systems. TUCSON.
- GUPTA S.S. and P.A. GROLL (1961). Gamma distribution in acceptance sampling based on life tests. Jour. Amer. Stat. Assoc. Vol. 56, pp. 942-970.
- JENNINGS M.E. and M.A. BENSON (1969). Frequency curves for annual floods series with some zero events or incomplete data. Water Resources Research 5(1).
- HERBST P.H. et al. (1973). Flood estimation by determination of regional parameters from limited data. Symposium on the design of water resources projects with inadequate data. Madrid. Tome 1 pp. 197-203.
- MARKOVIC R.D. (1965). Probability functions of best fit to distributions of annual precipitation and runoff. Hydrology Papers No. 8, Colorado State University.

BIBLIOGRAPHIE DES APPLICATIONS (suite)

- MATALAS N.C. (1963). Probability distribution of low flows. U.S. Geological Survey. Prof. Pap. 434 A.
- MOSS M.E. and M.R. KARLINGER (1974). Surface water network design by regression analysis simulation. Water Resources Research 10(3) pp. 427-433.
- QUELENNEC R.E. (1973). Contribution à l'étude probabiliste des phénomènes pluvieux. La Houille Blanche No. 1 1973 pp. 21-33.
- REICH B.M. (1973). Floods series compared to rainfall extremes. Water Resources Research 6(6) pp. 1655-1667.