

Record Number: 1190
Author, Monographic: Bobée, B.//Boucher, P.
Author Role:
Title, Monographic: Propriétés mathématiques et statistiques de la distribution Wakeby
Translated Title:
Reprint Status:
Edition:
Author, Subsidiary:
Author Role:
Place of Publication: Québec
Publisher Name: INRS-Eau
Date of Publication: 1979
Original Publication Date:
Volume Identification:
Extent of Work: iii, 36
Packaging Method: pages et un annexe
Series Editor:
Series Editor Role:
Series Title: INRS-Eau, Rapport de recherche
Series Volume ID: 118
Location/URL:
ISBN: 2-89146-115-0
Notes: Rapport annuel 1979-1980
Abstract: 15.00\$
Call Number: R000118
Keywords: rapport/ ok/ dl

PROPRIETES MATHÉMATIQUES
ET STATISTIQUES DE LA DISTRIBUTION
WAKEBY

RAPPORT SCIENTIFIQUE No 118

par

B. Bobée et P. Boucher

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

Décembre 1979

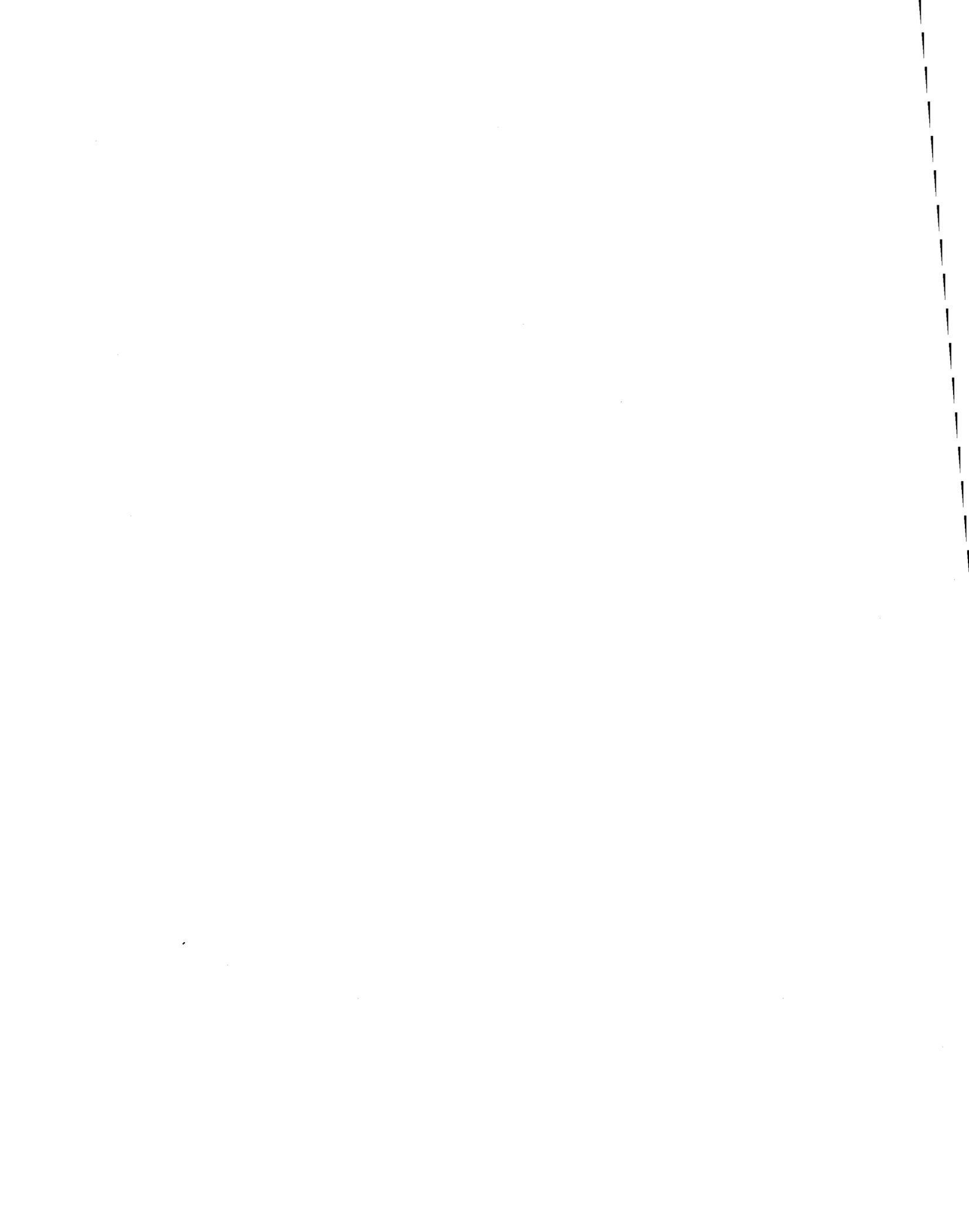


TABLE DES MATIERES

	PAGE
AVANT PROPOS.....	1
1. Propriétés mathématiques de la distribution Wakeby.....	2
1.1 Définition des fonctions de densité et de dis- tribution.....	2
1.2 Conditions d'existence de la fonction densité.....	3
1.3 Formes de la fonction densité.....	6
1.4 Formes de la fonction de distribution.....	6
2. Propriétés statistiques de la distribution Wakeby.....	10
2.1 Détermination des moments non centrés $\mu^i_{\gamma}(x)$	10
2.2 Détermination des moments centrés $\mu_{\gamma}(x)$	15
2.3 Calcul des statistiques d'ordre.....	18
2.4 Moments de probabilité pondérés (Probability weighted moments).....	26
REFERENCES.....	35

ANNEXE A

LISTE DES TABLEAUX

PAGE

Table 1. Conditions d'existence de la fonction densité (f)..... 5

LISTE DES FIGURES

	DATE
Figure 1. Tracé de la fonction $U = -(1-F)^b$ pour différentes valeurs positives de b.....	8
Figure 2. Tracé de la fonction $V = (1-F)^{-d}$ pour différentes valeurs positives de d.....	9
Figure 3. Fonctions de distribution des statistiques d'ordre de la distribution WAKEBY ($N = 39; k = 1, 10, 20, 30, 39$)	21
Figure 4. Exemple de calcul d'intervalles de confiance.....	25

AVANT-PROPOS

La distribution WAKEBY a récemment été proposée pour l'analyse des débits de crue (HOUGHTON, 1977). Cette distribution générale comportant cinq paramètres permet en particulier de tenir compte de l'effet de séparation mis en évidence par MATALAS et al. (1975); cet effet est relatif à l'écart entre les caractéristiques d'asymétrie d'échantillons tirés de lois usuelles et les caractéristiques d'asymétrie des débits de crue observés. HOUGHTON (1977, 1978a, 1978b) et LANDWEHR et al. (1978) ont déjà étudié quelques propriétés de la distribution WAKEBY; il s'agit dans ce rapport de regrouper ces résultats et de démontrer des résultats originaux afin d'en faire une présentation détaillée d'ensemble pour faciliter les études futures liées à cette distribution.

1. PROPRIETES MATHEMATQUES DE LA DISTRIBUTION WAKEBY

1.1 Définition des fonctions de densité et de distribution

La forme générale de la loi WAKEBY est donnée par:

$$x = -a (1-F)^b + c (1-F)^{-d} + e \quad (1)$$

a, b, c, d et e sont les paramètres de la distribution;

x est la variate;

F est la probabilité au non-dépassement ou encore la fonction de distribution telle que $F(x) = \Pr [X \leq x]$; on a: $0 \leq F \leq 1$.

Il est possible par différentiation d'en déduire la fonction densité $f(x)$, on a:

$$f(x) = \left[ab (1-F)^{b-1} + cd (1-F)^{-d-1} \right]^{-1} \quad (2)$$

on peut noter que la fonction densité:

- ne dépend pas directement du paramètre e;
- ne s'exprime pas directement en fonction de x mais plutôt en fonction de F.

En pratique $f(x)$ doit par définition être toujours positive ou nulle ce qui implique certaines contraintes sur les paramètres.

1.2 Conditions d'existence de la fonction densité

Par définition, f donnée par la relation (2) doit toujours être positive ou nulle ce qui implique que:

$$ab(1-F)^{b-1} + cd(1-F)^{-d-1} > 0 \quad (3)$$

avec $0 \leq F \leq 1$.

On peut envisager plusieurs cas:

1. Si $ab \geq 0$

a) Si $cd \geq 0$

la relation (3) est toujours satisfaite puisque $0 \leq F \leq 1$

b) Si $cd < 0$

la relation (3) est vérifiée si et seulement si: $-\frac{cd}{ab} < (1-F)^{b+d}$

• si $(b+d) > 0$, la fonction $(1-F)^{b+d}$ ayant un minimum de 0 l'inégalité précédente, donc la relation (3) ne peut être vérifiée puisque $(-cd/ab) > 0$;

• Si $(b+d) \leq 0$, la fonction $(1-F)^{b+d}$ a un minimum de 1 et la

relation (3) est vérifiée si $(-cd/ab) < 1$ c'est-à-dire si $(ab + cd) > 0$.

2. Si $ab < 0$

a) si $cd > 0$

la relation (3) peut alors s'écrire $(1-F)^{b+d} < -\frac{cd}{ab}$

• si $b + d \geq 0$ la fonction $(1-F)^{b+d}$ a un maximum de 1 et l'on doit donc avoir $1 < -(cd/ab)$ ou encore $(ab + cd) > 0$ pour que la relation (3) soit vérifiée;

• si $(b + d) < 0$, la fonction $(1-F)^{b+d}$ peut devenir infinie et l'inégalité (3) n'est pas vérifiée;

b) si $cd \leq 0$

la relation (3) ne peut être vérifiée puisque $(1-F)^{b+d}$ est toujours positive ou nulle.

On peut finalement dresser un tableau récapitulatif (table 1) de ces résultats qui recourent ceux obtenus par LANDWEHR et al. (1978). Les résultats de la table 1 peuvent être détaillés suivant les signes particuliers des paramètres a , b , c et d ; par exemple lorsque $ab > 0$ et $cd < 0$, la fonction densité n'existe pas si b et d sont positifs et

Table 1. Conditions d'existence de la fonction densité (f).

Signe de ab	Signe de cd	CONCLUSIONS SUR f
+	+	f est toujours définie
+	-	f existe si $(b+d) \leq 0$ et si $(ab + cd) > 0$ f n'existe pas si $(b+d) > 0$
-	+	f existe si $(b + d) \geq 0$ et si $(ab + cd) > 0$ f n'existe pas si $(b+d) < 0$
-	-	f n'est jamais définie

lorsque $ab < 0$ et $cd > 0$ la fonction densité n'existe pas si b et d sont négatifs, alors que dans tous les autres cas, l'existence de f est soumise aux valeurs respectives des paramètres.

1.3 Formes de la fonction densité

L'étude complète des différentes formes de la fonction densité est très difficile à effectuer en raison des nombreux cas possibles.

LANDWEHR et al. (1978) ont fait l'étude de certaines valeurs particulières des paramètres.

L'intervalle de définition de la variate x peut être obtenu à partir de la relation (1):

$$\text{pour } F = 0 \quad \text{on a } x_{\min} = -a + c + e = m$$

$$F = 1 \quad \text{on a } x_{\max} = e \text{ ou } +\infty \text{ (suivant les valeurs des paramètres).}$$

1.4 Formes de la fonction de distribution

L'étude complète de la fonction de distribution est complexe en raison des nombreuses valeurs que peuvent prendre les paramètres; cependant pour avoir une idée de la forme générale de cette fonction, on a étudié:

$$U(b) = -(1-F)^b$$

$$V(d) = (1-F)^{-d}$$

de sorte que l'on a $x = aU + cV + e$

1- Cas $b > 0$ et $d > 0$

La figure 1 indique le tracé de $U = -(1-F)^b$ pour $b = .25, .50, .75, 1.0, 2.0$ et 3.0 .

La figure 2 indique les tracés de $V = (1-F)^{-d}$ pour $d = .25, .50, .75$ et 1.0 .

Les fonctions aU et cV se déduisent de U et V par une affinité.

- . Si a et c sont positifs, la fonction densité est toujours définie et l'on obtient la fonction utilisée par HOUGHTON (1977);
- . Si a et c sont négatifs, la fonction f n'est jamais définie comme l'indique la table 1;
- . Si $a > 0$ et $c < 0$, d'après la table 1, la fonction densité n'est pas définie car on a $b + d > 0$;
- . Si $a < 0$ et $c > 0$, d'après la table 1, la fonction densité n'existe que si $ab + cd > 0$.

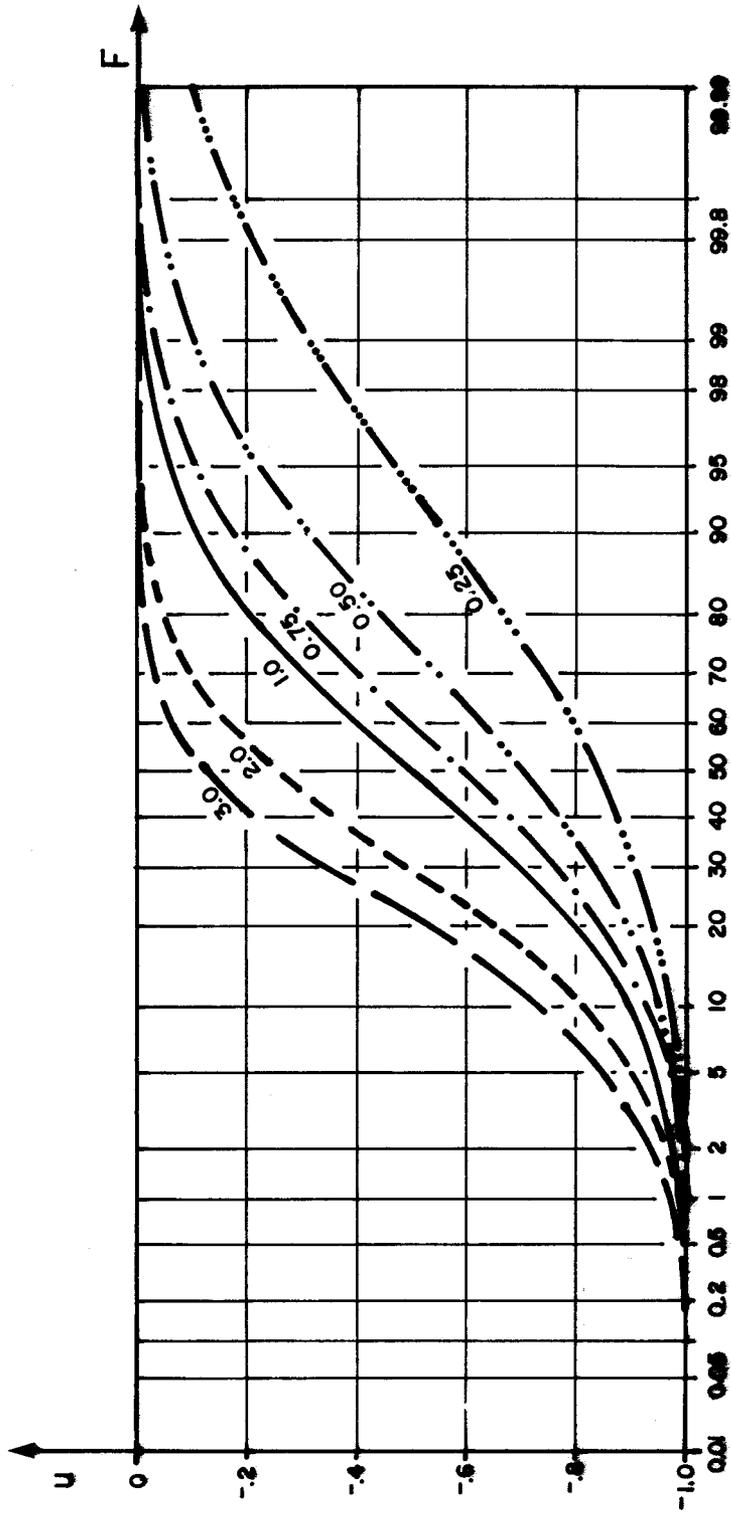


Figure 1 . Tracé de la fonction $u = -(1-F)^b$ pour différentes valeurs positives de b .

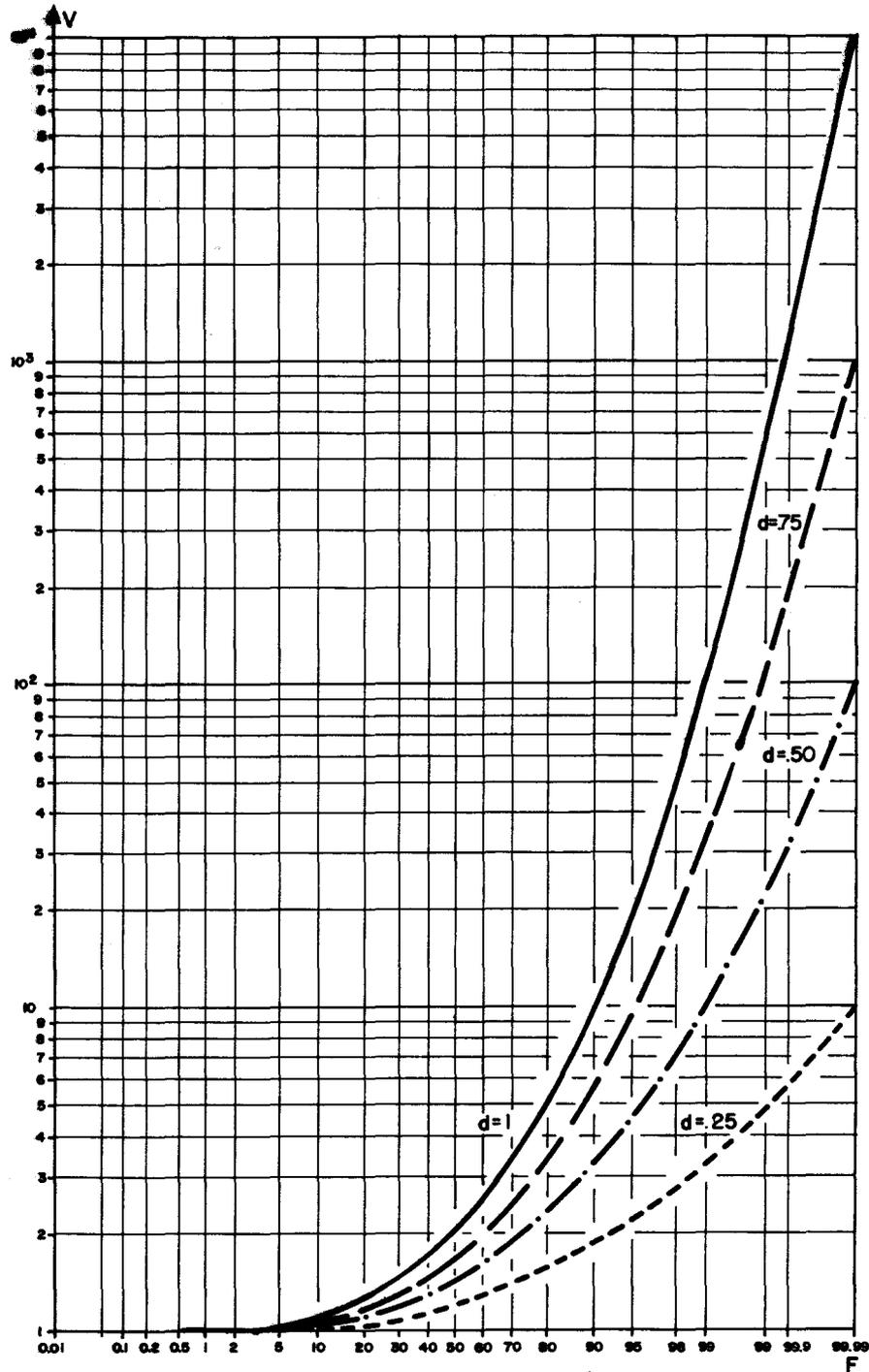


Figure 2. Tracé de la fonction $V = (1-F)^{-d}$ pour différentes valeurs positives de d .

2- Autres cas

Si $b < 0$ on a $U(b) = -V(-b)$

-b étant positif, $U(b)$ peut se déduire de la figure 2;

Si $d < 0$ on a $V(d) = -U(-d)$

-d étant positif, $V(d)$ peut se déduire de la figure 1.

Il en résulte que tous les cas suivant les combinaisons de signes de b et d , peuvent se déduire du cas $b > 0$ et $d > 0$; les conditions d'existence de la fonction densité étant soumises aux contraintes de la table 1 pour les paramètres a et c .

2. Propriétés statistiques de la distribution WAKEBY

2.1 Détermination des moments non centrés $\mu'_r(x)$

On a $x = -a(1-F)^b + c(1-F)^{-d} + e$. Le moment non centré d'ordre r est donné par:

$$\mu'_r(x) = \int_D x^r f(x) dx = \int_0^1 [x(F)]^r dF$$

d'où
$$\mu'_r(x) = \int_0^1 \left[-a(1-F)^b + c(1-F)^{-d} + e \right]^r dF$$

si l'on pose $u = 1-F$ $du = -dF$ et l'on a:

$$\mu'_r(x) = \int_0^1 (-au^b + cu^{-d} + e)^r du$$

$$\mu_r^i(x) = \int_0^1 \sum_{i=0}^r c_r^i e^i \left(\sum_{j=0}^{r-i} (-au^b)^j (cu^{-d})^{r-i-j} du \right)$$

ce qui s'écrit encore:

$$\mu_r^i(x) = \sum_{i=0}^r c_r^i e^i \sum_{j=0}^{r-i} (-a)^j c^{r-i-j} \Lambda_{ij}$$

avec:

$$\Lambda_{ij} = \int_0^1 u^{bj-d(r-i-j)} du$$

Lorsque Λ_{ij} est définie on a:

$$\Lambda_{ij} = \frac{1}{(b+d)j - dr + di + 1}$$

et

$$\mu_r^i(x) = \sum_{i=0}^r c_r^i e^i \sum_{j=0}^{r-i} c_{r-i}^j \frac{(-a)^j c^{r-i-j}}{(b+d)j + di - dr + 1} \quad (4)$$

Existence des intégrales Λ_{ij}

Λ_{ij} est de la forme $\int_0^1 u^\alpha du$, or cette intégrale converge si et seulement si $\alpha > -1$; pour que Λ_{ij} converge, il faut et il suffit que $[bj - d(r-i-j)] > -1$, i variant de 0 à r et j variant de 0 à $(r-i)$ pour le calcul du moment d'ordre 1.

. Si $r = 1$, l'exposant $\alpha = bj - d(r-i-j)$ prend les valeurs

$j \backslash i$	0	1
0	-d	b
1	0	*

Donc la moyenne (moment d'ordre 1) existe si et seulement si:

$$d < 1 \quad \text{et} \quad b > -1$$

. si $r = 2$, l'exposant $\alpha = bj - d(r-i-j)$ est donné par la table

$j \backslash i$	0	1	2
0	-2d	b-d	2b
1	-d	-b	*
2	0	*	*

Pour $i = 1, 2$ on retrouve la table du cas précédent $r = 1$

Si $d < \frac{1}{2}$ et $b > -\frac{1}{2}$ les exposants correspondant à $i = 0$ sont tous supérieurs à -1 et il y a convergence; quant aux autres valeurs de la table ($i = 1, 2$) il y a également convergence car:

$$d < \frac{1}{2} \iff d < 1 \iff -d > -1$$

$$b > -\frac{1}{2} \iff b > -1$$

Donc le moment non centré d'ordre 2 et par conséquent la variance ne sont définis que si et seulement si:

$$d < \frac{1}{2} \quad \text{et} \quad b > -\frac{1}{2}$$

De manière plus générale on obtient pour r la table:

$i \backslash j$	0	1	-----	j	-----	$(r-1)$	r
0	$-rd$	$b-d(r-1)$		$\frac{bj-d(r-j)}{d}$		$\frac{b(r-1)-d}{d}$	br
1	$-(r-1)d$	$\frac{b-d(r-2)}{d}$		$\frac{bj-d(r-j-1)}{d}$		$b(r-1)$	*
\vdots							
i	$-(r-i)d$	$\frac{b-d(r-i-1)}{d}$		$\frac{bj-d(r-i-j)}{d}$		*	*
\vdots							
$(r-1)$	$-d$	b		*		*	*
r	0	*		*		*	*

La partie encadrée de la table correspond au cas (r-1)

si $b > -\frac{1}{r}$ et $d < \frac{1}{r}$ alors quel que soit $j = 0, \dots, r$
 on a $bj - d(r-j) > -\frac{j}{r} + \frac{r-j}{r} = -1$ et il y a alors convergence des intégrales pour tous les exposants correspondant à la première ligne ($i = 0$).

Il y aura convergence pour les exposants du tableau correspondant à (r-1) si et seulement si:

$$b > -\frac{1}{r-1} \quad \text{et} \quad d < \frac{1}{r-1}$$

mais

$$\left. \begin{array}{l} b > -\frac{1}{r} \\ d < \frac{1}{r} \end{array} \right\} \longrightarrow \left\{ \begin{array}{l} b > -\frac{1}{r-1} \\ d < \frac{1}{r-1} \end{array} \right.$$

Donc de proche en proche on peut montrer que pour r il y a convergence des intégrales Λ_{ij} , donc que les moments d'ordre inférieurs ou égaux à r existent si et seulement si:

$$b > -\frac{1}{r} \quad \text{et} \quad d < \frac{1}{r}$$

En pratique:

- le moment non centré d'ordre 3 (et les moments d'ordre inférieur) donc le coefficient d'asymétrie existent si et seulement si:

$$b > -\frac{1}{3} \quad \text{et} \quad d < \frac{1}{3}$$

- le moment non centré d'ordre 4 (et les moments d'ordre inférieur)
donc le coefficient d'aplatissement existent si et seulement si:

$$b > -\frac{1}{4} \quad \text{et} \quad d < \frac{1}{4}$$

L'application de la relation (4) pour $r = 1$ donne la moyenne μ_1'

$$\mu_1' = \frac{c}{1-d} - \frac{a}{b+1} + e$$

qui existe si et seulement si: $d < 1$ et $b > -1$.

On trouve pour le moment d'ordre 2, μ_2' :

$$\mu_2' = \left(\frac{c^2}{1-2d} - \frac{2ac}{b-d+1} + \frac{a^2}{2b+1} \right) + 2e \left(\frac{c}{1-d} - \frac{a}{b+1} \right) + e^2$$

on en déduit la variance $\mu_2 = \mu_2' - \mu_1'^2$

$$\mu_2 = \left(\frac{c^2}{1-2d} - \frac{2ac}{b-d+1} + \frac{a^2}{2b+1} \right) - \left(\frac{c}{1-d} - \frac{a}{b+1} \right)^2$$

μ_2 et μ_2' existent si et seulement si $d < \frac{1}{2}$ et $b > -\frac{1}{2}$

2.2 Détermination des moments centrés $\mu_r(x)$

On peut théoriquement déduire le moment centré par rapport à la moyenne

μ_r en fonction des moments non centrés μ_1^j -- μ_r^j par:

$$\mu_r = \sum_{j=0}^r C_r^j \mu_j^j (-\mu_1^j)^{r-j} \quad (5)$$

Dans le cas de la variate x , compte tenu de l'expression complexe (4) des moments $\mu_r^j(x)$ la formule résultante est inextricable, c'est pourquoi nous utiliserons une autre approche.

La formule (5) montre cependant que pour que le moment centré d'ordre r existe, il faut et il suffit que les moments non centrés jusqu'à l'ordre r existent, donc que $b > -\frac{1}{r}$ et $d < \frac{1}{r}$.

Calcul des moments centrés de la variate $y = x - e$

On considère la variate $y = x - e$ on a $\bar{y} = \bar{x} - e$

on a
$$\mu_r(x) = \int_0^1 (x - \bar{x})^r dF_x$$

$$\mu_r(y) = \int_0^1 (y - \bar{y})^r dF_y$$

mais puisque les fonctions densité de $x, f(x)$ et de $y, g(y)$ sont identiques

on a:

$$f(x)dx = g(y)dy \quad \text{d'où} \quad dF_x = dF_y$$

de plus $x - \bar{x} = y - \bar{y}$

donc $\underline{\mu_r(x) = \mu_r(y)}$

Il est donc équivalent de calculer les moments centrés de x ou de y .
Mais les moments $\mu_r(y)$ peuvent s'exprimer en fonction des $\mu'_k(y)$ par la relation (5). On a donc:

$$\mu_r(x) = \mu_r(y) = \sum_{k=0}^r c_r^k \mu'_k(y) \left[-\mu'_1(y) \right]^{r-k} \quad (6)$$

Les moments $\mu'_k(y)$ peuvent se déduire de la relation (4) pour $e = 0$, car alors $y = x$, on a:

$$\mu'_k(y) = \sum_{j=0}^k c_k^j \frac{(-a)^j c^{k-j}}{(b+d)^j - dk+1} \quad (7)$$

En remplaçant $\mu'_k(y)$ donné par (7) dans la relation (6), on obtient $\mu_r(x)$

$$\mu_r(x) = \sum_{k=0}^r c_r^k \left[-\mu'_1(y) \right]^{r-k} \cdot \sum_{j=0}^k c_k^j \frac{(-a)^j c^{k-j}}{(b+d)^j - dk+1} \quad (8)$$

avec

$$\mu_1(y) = \left(\frac{c}{1-d} - \frac{a}{b+1} \right)$$

On peut en déduire:

- le coefficient d'asymétrie $\gamma(x) = \frac{\mu_3(x)}{[\mu_2(x)]^{3/2}}$
- le coefficient d'aplatissement $k(x) = \frac{\mu_4(x)}{\mu_2^2(x)}$

2.3. Calcul des statistiques d'ordre

Les éléments d'un échantillon de taille N peuvent être classés, on obtient ainsi des réalisations des statistiques d'ordre. Si les fonctions densité et de distribution de la loi WAKEBY sont f et F, la fonction densité de l'événement x_k d'un échantillon de taille N (statistique d'ordre k) est:

$$h(x_k) = \frac{N!}{(k-1)!(N-k)!} \left[F(x_k) \right]^{k-1} \left[1-F(x_k) \right]^{N-k} f(x_k)$$

avec:
$$x_k = -a(1-F)^b + c(1-F)^{-d} + e$$

$$f(x_k) = \left[ab(1-F)^{b-1} + cd(1-F)^{-d-1} \right]^{-1}$$

La distribution cumulée de la statistique d'ordre k est $H(Z_0, k)$ telle que:

$$P_r [x_k \leq z_0] = H(z_0, k) = \int_m^{z_0} h(x_k) dx_k$$

m est la borne inférieure de l'intervalle de définition de la variate qui suit une distribution WAKEBY (cf. 1.3). On peut écrire:

$$H(z_0, k) = \frac{N!}{(k-1)! (N-k)!} \int_0^{F_0} F^{k-1} (1-F)^{N-k} dF$$

avec $F(m) = 0$ et $F(z_0) = F_0$

La relation précédente peut s'écrire:

$$H(z_0, k) = \frac{N!}{(k-1)! (N-k)!} \int_0^{F_0} F^{k-1} \cdot \sum_{j=0}^{N-k} C_{N-k}^j (-F)^j dF$$

ou encore

$$H(z_0, k) = \frac{N!}{(k-1)! (N-k)!} \sum_{j=0}^{N-k} (-1)^j C_{N-k}^j \frac{F_0^{k+j}}{k+j} \quad (9)$$

avec
$$z_0 = -a (1-F_0)^b + c (1-F_0)^{-d} + e$$

quand
$$F_0 = 0 \quad z_0 = -a + c + e = m \quad \text{et} \quad H(m, k) = 0$$

$$F_0 = 1 \quad H(z_0, k) = 1$$

En pratique le calcul direct de $H(z_0, k)$ peut poser certains problèmes de précision en raison des nombres élevés qui interviennent. La figure 3 indique la forme de quelques statistiques d'ordre pour $N = 39$.

Relation de symétrie entre statistiques d'ordre

On a d'après ce qui précède:

$$H [z(F_0), k] = \int_0^{F_0} \frac{N!}{(k-1)! (N-k)!} F^{k-1} (1-F)^{N-k} dF$$

$$H [z(F_0), N-k+1] = \int_0^{F_0} \frac{N!}{(N-k)! (k-1)!} F^{N-k} (1-F)^{k-1} dF$$

Si on pose $u = 1-F \quad du = -dF$

$$H [z(F_0), N-k+1] = \int_1^{1-F_0} \frac{N!}{(N-k)! (k-1)!} (1-u)^{N-k} u^{k-1} du$$

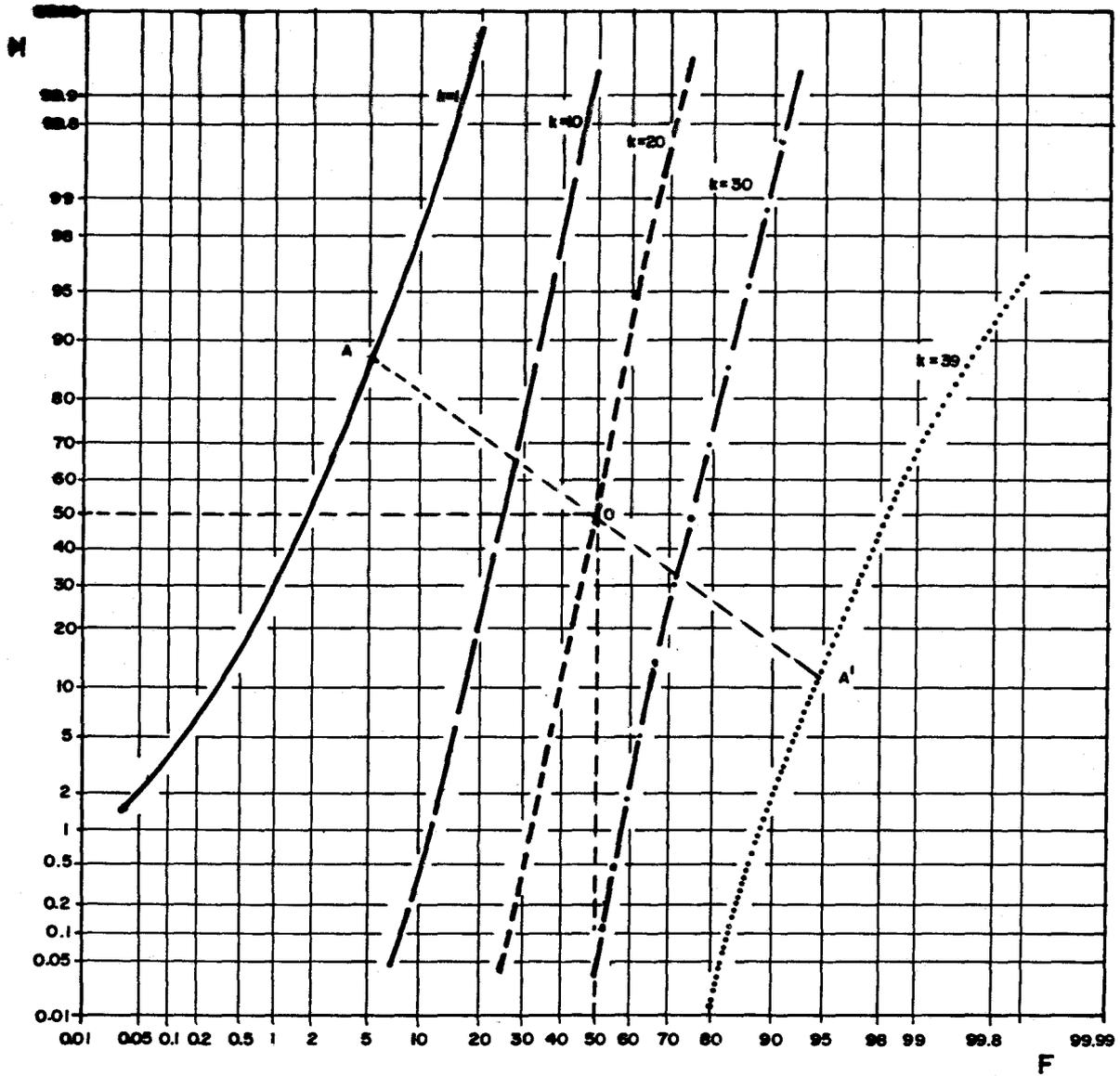


Figure 3 . Fonctions de distribution des statistiques d'ordre de la distribution Wakeby
($N=39$; $k=1, 10, 20, 30, 39$).

$$H \left[z(1-F_0), N-k+1 \right] = \int_{F_0}^1 \frac{N!}{(N-k)!(k-1)!} (1-u)^{N-k} u^{k-1} du$$

$$H \left[z(1-F_0), N-k+1 \right] = 1 - \int_0^{F_0} \frac{N!}{(N-k)!(k-1)!} (1-u)^{N-k} u^{k-1} du$$

Donc finalement:

$$H \left[z(1-F_0), N-k+1 \right] = 1 - H \left[z(F_0), k \right]$$

ou encore:

$$\boxed{H \left[z(F_0), N-k+1 \right] = 1 - H \left[z(1-F_0), k \right]} \quad (10)$$

En pratique on peut donc utiliser la statistique d'ordre k pour en déduire celle d'ordre $N-k+1$.

Par exemple si $N = 40$, $k = 1$ on a:

$$H \left[z(.05), 1 \right] = .871$$

On en déduit:

$$H \left[z(.95), 40 \right] = 1 - 0.871 = 0.129$$

Le point A situé sur la statistique d'ordre 1 ayant pour composantes $F_0 = .05$ et $H = 0.871$ et le point A' situé sur la statistique d'ordre 40 et ayant pour composantes $F_0 = .95$ et $H = .129$ sont symétriques dans le diagramme de la figure 3 par rapport au point 0 de composantes $F_0 = .5$ et $H = .5$; de manière plus générale les statistiques d'ordre k et $(N-k+1)$ sont symétriques par rapport au point 0.

Application à la détermination d'intervalles de confiance de la distribution WAKEBY

Nous utiliserons ici la méthode décrite par GLADWELL et CHENG-NAN LIN (1969) et appliquée à la loi Pearson type 3 par BOBEE et MORIN (1973). Cette méthode est ici particulièrement pratique en raison de la forme explicite (formule 9) des statistiques d'ordre de la distribution WAKEBY. Elle a d'autant plus d'intérêt que les méthodes usuelles, basées sur la variance asymptotique d'un événement de période de retour donnée, ne sont pas applicables pour la distribution WAKEBY car les méthodes des moments et celle du maximum de vraisemblance ne peuvent être utilisées dans le cas de cette distribution.

A chaque événement ordonné x_k on peut attribuer une probabilité empirique P_k ; on connaît par ailleurs la distribution cumulée $H(z,k)$ de la statistique d'ordre k associée à l'événement x_k . Les limites de l'intervalle de confiance au niveau $(1-\alpha)$ auront pour composantes:

limite inférieure: $A_{1,k}: (P_k, z_{1,k})$

limite supérieure: $A_{2,k}: (P_k, z_{2,k})$

$z_{1,k}$ et $z_{2,k}$ sont tels que:

$$H(z_{1,k}, k) = \alpha/2$$

$$H(z_{2,k}, k) = 1 - \alpha/2$$

En pratique, à partir de la relation (9) on détermine par itération les valeurs $F_{0,1}$ et $F_{0,2}$ telles que $H(z_{1,k}, k) = \alpha/2$ et $H(z_{2,k}, k) = 1 - \alpha/2$ et l'on en déduit:

$$z_{i,k} = -a (1 - F_{0,i})^b + c (1 - F_{0,i})^{-d} + e \quad i = 1, 2$$

Les valeurs a , b , c , d et e sont préalablement obtenues par ajustement de la distribution WAKEBY à l'échantillon des valeurs observées. La figure 4 donne un exemple de détermination des intervalles de confiance à 99% et à 90% de la distribution WAKEBY correspondant à: $a = 0.5$; $b = 2.0$; $c = 1.5$; $d = 0.2$ et $e = -1.0$.

En appendice figure un programme de calcul permettant la détermination automatique des intervalles de confiance de la distribution WAKEBY.

Le choix de la probabilité empirique P_k peut avoir une certaine

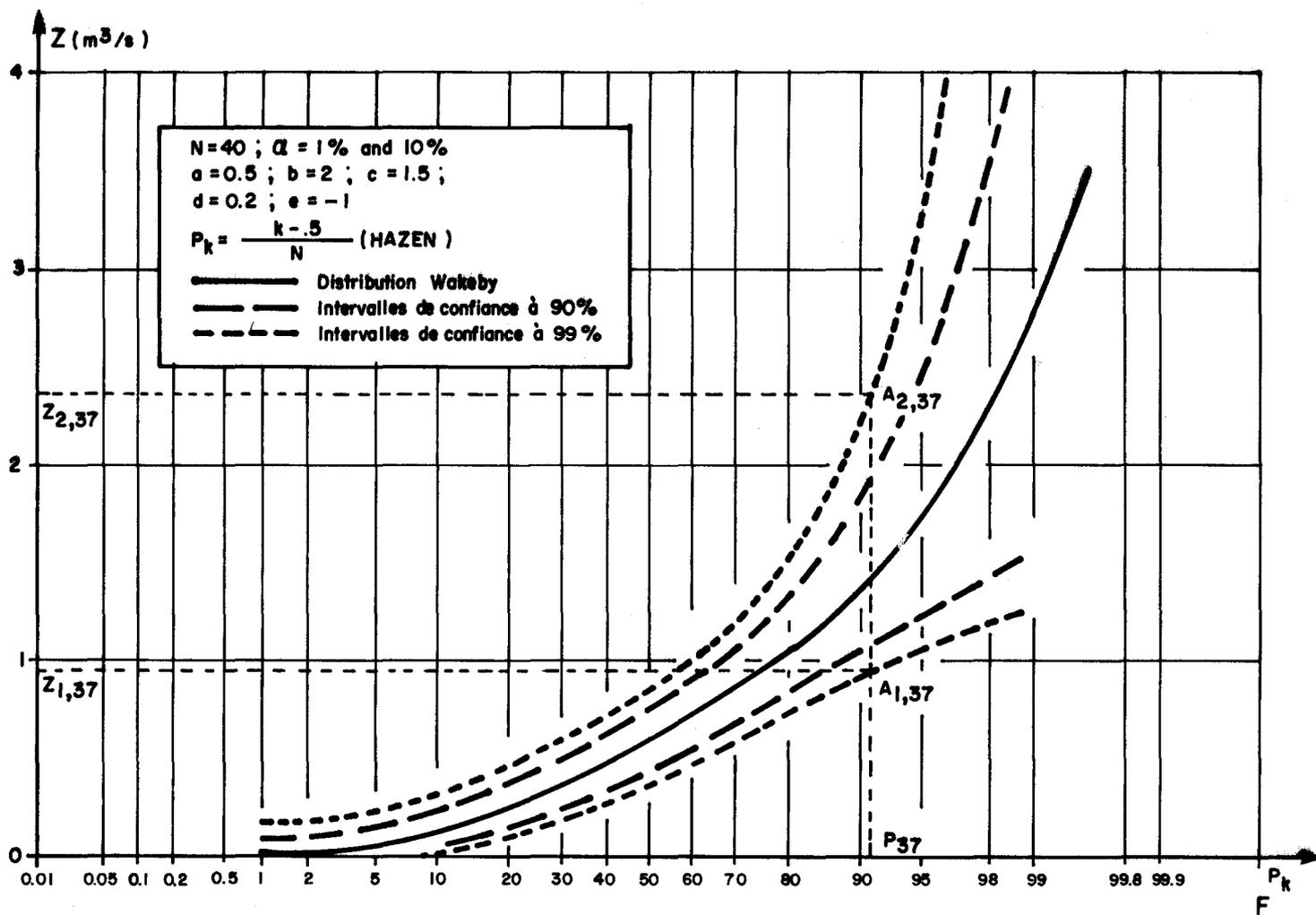


Figure 4 . Intervalles de confiance à 99% et 90% de la distribution Wakeby .

influence dans l'établissement des intervalles de confiance; en pratique on peut utiliser:

$$P_k = \frac{k - .5}{N} \quad (\text{HAZEN})$$

$$P_k = \frac{k}{N + 1} \quad (\text{WEIBULL})$$

$$P_k = \frac{k - .3}{N + .4} \quad (\text{CHEGODAYEV})$$

Pour une période de retour T donnée, la formule de WEIBULL conduit à la valeur la plus élevée de la limite supérieure de l'intervalle de confiance, la formule de HAZEN à la valeur la plus faible et l'on obtient pour la formule de CHEGODAYEV une valeur intermédiaire; pour les périodes de retour élevées, les écarts peuvent être assez grands. Il n'existe pas actuellement un critère conduisant à la sélection de la meilleure probabilité empirique, une étude par simulation de MONTE-CARLO pourrait fournir des indications à ce sujet.

2.4 Moments de probabilité pondérés (Probability weighted moments)

GREENWOOD et al., (1978) définissent le moment de probabilité pondéré (MPP) par:

$$M_{\ell, j, k} = \int_0^1 [x(F)]^\ell F^j (1-F)^k dF$$

ℓ, j, k sont des entiers non négatifs.

L'utilisation de ces moments est particulièrement intéressante dans le cas de la distribution WAKEBY pour laquelle il est impossible d'exprimer de manière explicite les paramètres en fonction des moments usuels.

Relation avec les statistiques d'ordre

Si $\mu_{\ell}^j(j+1; k+j+1)$ est le moment non centré de la statistique d'ordre $(j+1)$ tirée d'un échantillon de taille $(k+j+1)$, on peut montrer (GREENWOOD et al., 1978) que:

$$M_{\ell, j, k} = B(j+1, k+1) \cdot \mu_{\ell}^j(j+1, k+j+1) \quad (11)$$

$B(j+1, k+1)$ est la fonction Bêta telle que:

$$B(j+1, k+1) = \frac{j! \cdot k!}{(j+k+1)!}$$

Il est donc possible à partir de la relation (11) de déterminer les moments des statistiques d'ordre définis en (2.3).

Relation entre les MPP et les paramètres de la distribution WAKEBY

La forme générale de la distribution WAKEBY (1) peut s'écrire

$$x = m + a [1 - (1-F)^b] - c [1 - (1-F)^{-d}] \quad (12)$$

(avec $m + a - c = e$)

On peut montrer que l'on a:

$$M_{1, 0, k} = \frac{m + a - c}{k + 1} - \frac{a}{b + k + 1} + \frac{c}{k - d + 1} \quad (13)$$

Ce moment étant défini si $(b + k) > -1$ et $(k - d) > -1$

On pose:

$$M_{(k)} = M_{1, 0, k}$$

$$\{k\} = (k + 1) (k + 1 + b) (k + 1 - d) M_{(k)}$$

$$\text{Si } k_j = k + i$$

On peut montrer que de manière générale:

$$\begin{aligned} \{k_j\} = \{k + i\} &= m [(k_j + 1 + b) (k_j + 1 - d)] + ab [k_j + 1 - d] \\ &+ cd [k_j + 1 + b] \end{aligned} \quad (14)$$

Suivant les valeurs de m on peut montrer qu'il existe une relation entre les $\{k_j\}$:

$$\text{Si } m = 0 \quad \text{on a:} \quad - \{k_2\} + 2 \{k_1\} - \{k_0\} = 0 \quad (15)$$

$$\text{Si } m \neq 0 \quad \text{on a:} \quad \{k_3\} - 3 \{k_2\} + 3 \{k_1\} - \{k_0\} = 0 \quad (16)$$

Ces équations sont valables quel que soit k entier non négatif.

Il est possible d'exprimer de manière générale les paramètres de la distribution WAKEBY en fonction de MPP donc des $\{k_i\}$.

On définit préalablement les quantités $A(j, k)$, N_{4-j} , C_{4-j} telles que:

$$\text{Si } m \neq 0 \quad A(j, k) = (k_3 + 1)^j M_{(k_3)} - 3(k_2 + 1)^j M_{(k_2)} + 3(k_1 + 1)^j M_{(k_1)} - (k_0 + 1)^j M_{(k_0)}$$

$$\text{Si } m = 0 \quad A(j, k) = -(k_2 + 1)^j M_{(k_2)} + 2(k_1 + 1)^j M_{(k_1)} - (k_0 + 1)^j M_{(k_0)}$$

$$N_{4-j} = A(j, K) \quad \text{et} \quad C_{4-j} = A(j, I)$$

K et I sont des entiers non négatifs tels que $I \neq K$.

On obtient alors pour les paramètres:

$$b = \frac{(N_3 C_1 - N_1 C_3) \pm \sqrt{(N_1 C_3 - N_3 C_1)^2 - 4(N_1 C_2 - N_2 C_1)(N_2 C_3 - N_3 C_2)}}{2(N_2 C_3 - N_3 C_2)} \quad (17)$$

$$d = \frac{A(3,k) + b A(2,k)}{A(2,k) + b A(1,k)} \quad (18)$$

$$m = \frac{\{k_2\} + \{k_0\} - 2 \{k_1\}}{2} \quad (19)$$

$$a = \frac{(k_0 + 1 + b)(k_1 + 1 + b)}{b(b + d)} \left[-m + \frac{\{k_1\}}{k_1 + 1 + b} - \frac{\{k_0\}}{k_0 + 1 + b} \right] \quad (20)$$

$$c = \frac{(k_0 + 1 - d)(k_1 + 1 - d)}{d(b + d)} \left[m + \frac{\{k_0\}}{k_0 + 1 - d} - \frac{\{k_1\}}{k_1 + 1 - d} \right] \quad (21)$$

GREENWOOD et al., (1978) au lieu de la relation (19) donnent:

$$m = \frac{\{k_3\} - \{k_2\} - \{k_1\} + \{k_0\}}{4} \quad (22)$$

L'équation (17) donne 2 solutions possibles pour b mais GREENWOOD et al., (1978) montrent qu'en fait ces 2 possibilités conduisent à la même solution générale pour l'équation (12) qui définit la distribution WAKEBY.

En pratique, les relations précédentes permettent le calcul des

paramètres a , b , c , d et m lorsque les MPP sont connus; en effet:

- Si les $M_{(k_i)}$ sont connus à partir d'un échantillon, on peut en déduire $A(j,k)$, C_{4-j} , N_{4-j} donc b par la relation (17);
- quand b est connu, on peut calculer d par la relation (18);
- quand b et d sont calculés, on en déduit m (lorsque $m \neq 0$) par (19) ou (22);
- quand b , d et m ont été calculés, on peut en déduire a et c respectivement par les relations (20) et (21).

Les équations 17 à 22 sont valables quels que soient les entiers non négatifs K , I et k ; en pratique, pour faire intervenir des moments d'ordre le plus bas possible, on prend $K = 0$, $I = 1$, $k = 0$ et l'on obtient une solution particulière des équations (17) à (22), on a:

- Si $m \neq 0$

$$N_{4-j} = 4^j M_{(3)} - 3^{j+1} M_{(2)} + 3 * 2^j M_{(1)} - M_{(0)}$$

$$C_{4-j} = 5^j M_{(4)} - 3 * 4^j M_{(3)} + 3^{j+1} M_{(2)} - 2^j M_{(1)}$$

- Si $m = 0$

$$N_{4-j} = -(3)^j M_{(2)} + (2)^{j+1} M_{(1)} - M_{(0)}$$

$$C_{4-j} = -(4)^j M_{(3)} + 2 * 3^j M_{(2)} - 2^j M_{(1)}$$

On a alors:

$$b = \frac{(N_3 C_1 - N_1 C_3) + \sqrt{(N_1 C_3 - N_3 C_1)^2 - 4(N_1 C_2 - N_2 C_1)(N_2 C_3 - N_3 C_2)}}{2(N_2 C_3 - N_3 C_2)} \quad (23)$$

$$d = \frac{N_1 + b N_2}{N_2 + b N_3} \quad (24)$$

$$m = \frac{\{2\} + \{0\} - 2\{1\}}{2} \quad (25)$$

ou

$$m = \frac{\{3\} - \{2\} - \{1\} + \{0\}}{4} \quad (26)$$

$$a = \frac{(1+b)(2+b)}{b(b+d)} \left[-m + \frac{\{1\}}{2+b} - \frac{\{0\}}{1+b} \right] \quad (27)$$

$$c = \frac{(1-d)(2-d)}{d(b+d)} \left[m + \frac{\{0\}}{1-d} - \frac{\{1\}}{2-d} \right] \quad (28)$$

Estimation des MPP à partir d'un échantillon.

Pour pouvoir calculer les paramètres de la distribution WAKEBY on doit être capable d'estimer les MPP à partir d'un échantillon de taille N.

En utilisant la relation (11), LANDWEHR et WALLIS (1978) montrent que l'estimation $\hat{M}_{(k)}$ de $M_{(k)}$ est donnée par:

$$\hat{M}_{(k)} = \frac{1}{N} \sum_{i=1}^{N-k} x_i \frac{C_{N-i}^k}{C_{N-1}^k} \quad (29)$$

Cette relation peut encore s'écrire :

$$\hat{M}_{(k)} = \frac{\sum_{i=1}^{N-k} x_i \left[\prod_{j=0}^{k-1} (N-i-j) \right]}{\prod_{j=0}^k (N-j)} \quad (30)$$

\prod représentant l'opérateur produit.

On trouve ainsi pour:

$$k = 0 \quad \hat{M}_{(0)} = \frac{1}{N} \sum_{i=1}^N x_i \quad (\text{moyenne de l'échantillon})$$

$$k = 1 \quad \hat{M}_{(1)} = \frac{\sum_{i=1}^{N-1} x_i (N-i)}{N(N-1)}$$

$$k = 2 \quad \hat{M}_{(2)} = \frac{\sum_{i=1}^{N-2} x_i (N-i)(N-i-1)}{N(N-1)(N-2)}$$

etc...

REFERENCES

BOBEE, B. et G. MORIN. (1973).

Détermination des intervalles de confiance de la loi Pearson type 3 par les statistiques d'ordre. Journ. Hydr., 20, 137-153.

GLADWELL, J.S. and CHENG-NAN LIN. (1969).

Confidence limits determined using ordre statistics. Wat. Resour. Res. 5(5), 1120-1123.

GREENWOOD, J.A., J. M. LANDWEHR and J.R. WALLIS. (1978).

Probability weighted moments: definition and application. Research Report, IBM Research Division RC7108, Environmental Sciences.

HOUGHTON, J.C. (1977).

Robust estimation of the frequency of extreme events in a flood frequency context. Ph.D. dissertation, Harvard University, Cambridge, Mass.

HOUGHTON, J.C. (1978a).

Birth of a parent: the Wakeby distribution for modeling flood flows. Wat. Resour. Res. 14(6), 1105-1109.

HOUGHTON, J.C. (1978b).

The incomplete means estimation procedure applied to flood frequency analysis. Wat. Resour. Res. 14(6), 1111-1115.

LANDWEHR, J. M., N.C. MATALAS and J.R. WALLIS. (1978).

Some comparisons of flood statistics in real and log-space. Wat. Resour. Res. 14(5), 902-920.

LANDWEHR, J.M. and J.R. WALLIS. (1978).

Probability weighted moments compared with some traditional techniques in estimating Gumbel parameters and quantiles: Research Report, IBM Research Division RC 7281, Environmental Science.

MATALAS, N.C., J.R. SLACK and J.R. WALLIS. (1975).

Regional skew in search of a parent. Wat. Resour. Res. 11(6), 815-826.

ANNEXE A

Détermination automatique des intervalles de
confiance d'une distribution WAKEBY

TABLE DES MATIERES

	PAGE
A.1 Calculs théoriques	A1
A.1.1 Détermination du maximum de $A_{k, j}$	A1
A.1.1.1 Etude du minimum de $P_j = j! (N-k-j)!$	A2
A.1.1.2 Etude du minimum de $Q_{k, j} = (k-1)!j!(N-k-j)!$	A3
A.1.2 Détermination du maximum de $B_{k, j}$	A11
A.1.2.1 Cas où $F_0 = 1$	A11
A.1.2.2 Cas où $F_0 \neq 1$	A19
A.2 Calcul pratique de $H(Z, k)$	A19
A.2.1 Cas où $F_0 = 1$	A19
A.2.2 Cas où $F_0 \neq 1$	A23
A.2.3 Utilisation de la relation de symétrie	A24
A.3 Calcul des intervalles de confiance	A25
A.4 Programme de calcul ITCO	A25
A.4.1 Détermination de différentes valeurs de k	A27
A.4.2 Utilisation du programme	A27
A.4.2.1 Données d'entrée	A27
A.4.2.2 Sortie des résultats	A28
A.5 Listing du programme et exemple de traitement	A29

A.1 Calculs théoriques

Les calculs de $H(Z, k)$ lorsque N est élevé, font apparaître des problèmes de précision. On doit en effet toujours trouver $0 \leq H(Z, k) \leq 1$; or, en pratique, en raison des nombres importants intervenant dans le calcul de $H(Z, k)$, on peut obtenir, même en effectuant des calculs en double précision (28 chiffres significatifs), des valeurs aberrantes de $H(Z, k)$. Ceci nous a conduit à étudier plus en détail les fonctions qui interviennent dans le calcul de $H(Z, k)$.

La fonction $H(Z, k)$ peut s'écrire pour N donné:

$$H(Z, k) = \sum_{j=0}^{N-k} (-1)^j B_{k,j} \quad k = 1, \dots, N$$

avec:

$$B_{k,j} = \frac{F_0^{k+j}}{k+j} A_{k,j} \quad \text{et} \quad A_{k,j} = \frac{N!}{(k-1)! j! (N-k-j)!}$$

D'un point de vue pratique H est toujours inférieur à 1 mais $A_{k,j}$ peut prendre des valeurs très élevées; les limites de précision de l'ordinateur rendent importante la connaissance des valeurs maximales atteintes par $B_{k,j}$ et $A_{k,j}$ pour N donné.

A.1.1 Détermination du maximum de $A_{k,j}$

Lorsque N est fixé $A_{k,j}$ est maximum lorsque $\left[(k-1)! j! (N-k-j)! \right]$ est minimum.

Pour étudier le maximum de $A_{k,j}$ on doit étudier successivement les valeurs minimales de:

a) $P_j = j! (N-k-j)!$

b) $Q_{k,j} = (k-1)! j! (N-k-j)!$

A.1.1.1 Etude du minimum de $P_j = j! (N-k-j)!$

Lorsque N et k sont fixés, j varie de 0 à $(N-k)$, on doit donc trouver la valeur de j qui rend minimum le produit de deux factoriels $j!$ et $(N-k-j)!$ dont la somme est constante ($j + N-k-j = N-k$)

α) Si $(N-k)$ est pair : $N-k = 2\phi$ (ϕ entier)

on a $P_j = j! (2\phi-j)!$, pour déterminer la valeur de j correspondant au minimum on considère $\frac{P_j}{P_{j-1}} = p_j$, on a $p_j = \frac{j! (2\phi-j)!}{(j-1)! (2\phi-j+1)!} = \frac{j}{2\phi-j+1}$

$p_j < 1$ entraîne $j < \phi + \frac{1}{2}$

$p_j > 1$ entraîne $j > \phi + \frac{1}{2}$

$p_j = 1$ entraîne $j = \phi + \frac{1}{2}$

Les relations précédentes montrent que:

- il ne peut y avoir 2 minimums car on aurait alors $p=1$ ce qui est impossible puisque ϕ est entier;
- lorsque j varie de 0 à ϕ il y a décroissance $P_0 > P_1 > \dots > P_{\phi-1} > P_\phi$

- lorsque j varie de $\phi+1$ à 2ϕ il y a croissance $P_\phi < P_{\phi+1} < P_{\phi+2} \dots < P_{2\phi}$

Donc, lorsque $(N-k)$ est pair P_j passe par un minimum unique lorsque $j = \phi = \frac{N-k}{2}$
qui vaut $P_\phi = \phi! \phi!$

$\beta)$ Si $(N-k)$ est impair : $N-k = 2\phi + 1$ (ϕ entier)

On a alors: $P_j = j! (2\phi + 1 - j)!$

$$p_j = \frac{P_j}{P_{j-1}} = \frac{j}{2\phi - j + 2}$$

$p_j < 1$	entraîne	$j < \phi + 1$
$p_j > 1$	entraîne	$j > \phi + 1$
$p_j = 1$	entraîne	$j = \phi + 1$

On peut donc en conclure que:

$$P_0 > P_1 \dots > P_\phi = P_{\phi+1} < P_{\phi+2} \dots < P_{2\phi}$$

Lorsque $(N-k)$ est impair il y a donc 2 minimums de P_j pour $j = \phi = \frac{N-k-1}{2}$

et $j = \phi + 1 = \frac{N-k+1}{2}$ qui valent $P_\phi = P_{\phi+1} = \phi! (\phi+1)!$

A.1.1.2 Etude du minimum de $Q_{k,j} = (k-1)! j! (N-k-j)!$

On a $Q_{k,j} = (k-1)! P_j$ k varie de 1 à N

Pour N fixé, à chaque valeur de k on fait varier j de 0 à $(N-k)$ et l'on obtient $(N-k+1)$ valeurs de $Q_{k,j}$ mais le minimum de $Q_{k,j}$ est obtenu pour la valeur de j qui donne le minimum de P_j , puisqu'alors k est fixé. Pour N fixé, on obtient donc pour chaque valeur de k un minimum M_k , j étant donné par la valeur qui conduit au minimum de P_j ; si par exemple $N=10$, $k=2$ on aura le minimum M_2 pour $j = \frac{N-k}{2} = 4$ et l'on a:

$$M_2 = 1! 4! 4!$$

Pour N donné le maximum de $A_{k,j}$, donc le minimum absolu de $Q_{k,j}$, est aussi le minimum absolu des M_k lorsque k varie.

Lorsque k est tel que $(N-k)$ est impair, on a vu précédemment qu'il y avait deux minimums de P_j , mais ils conduisent à la même valeur de M_k en effet: P_j est minimum pour $j_1 = \frac{N-k-1}{2}$ et $j_2 = \frac{N-k+1}{2}$

$$\text{pour } j_1 = \frac{N-k+1}{2}; M_k = (k-1)! \frac{(N-k+1)!}{2} \frac{(N-k-1)!}{2}$$

$$\text{pour } j_2 = \frac{N-k-1}{2}; M_k = (k-1)! \frac{(N-k-1)!}{2} \frac{(N-k+1)!}{2}$$

Il suffit donc de considérer une seule des 2 valeurs de j pour connaître le minimum.

Si l'on revient au but initial de l'étude qui est la recherche des valeurs maximales de $B_{k,j}$, on voit que $B_{k,j}$ fait intervenir $(k+j)$. Si donc pour N donné le minimum M_k , donc le maximum de $A_{k,j}$, est atteint pour 2 valeurs j_1 et j_2 telles que $j_1 < j_2$ (lorsque $N-k$ est impair) il suffit de considérer la plus petite valeur j_1 en effet:

$$A_{k,j_1} = A_{k,j_2} \quad \text{entraîne} \quad \frac{A_{k_1,j_1}}{k+j_1} > \frac{A_{k_1,j_2}}{k+j_2}$$

avec $j_1 < j_2$

Donc en pratique, lorsque $(N-k)$ est impair et qu'il y a 2 valeurs de j conduisant au minimum M_k , on considère seulement la valeur $j = \frac{N-k-1}{2}$

Pour déterminer le minimum absolu de $A_{k,j}$ on considère donc la suite des M_k ; d'après ce qui précède :

lorsque $(N-k)$ est pair on a $M_k = (k-1)! \left(\frac{N-k}{2}\right)! \left(\frac{N-k}{2}\right)!$

lorsque $(N-k)$ est impair on a $M_k = (k-1)! \left(\frac{N-k-1}{2}\right)! \left(\frac{N-k+1}{2}\right)!$

Pour déterminer le minimum absolu de $Q_{k,j}$ (donc des M_k) on considère 2 cas suivant la parité de N et l'on examinera pour quelles valeurs de j et de k le minimum ou les minimums sont atteints.

α) N est impair, $N = 2\mu + 1$

k varie de 1 à $(2\mu + 1)$

$$k = 1 \quad N-k = 2\mu \text{ est pair} \quad M_1 = 0! \mu! \mu!$$

$$k = 2 \quad N-k = (2\mu - 1) \text{ est impair} \quad M_2 = 1!(\mu - 1)! \mu!$$

$k = 2p-1$	$N-k = (2\mu-2p+2)$ est pair	$M_{2p-1} = (2p-2)! (\mu-p+1)! (\mu-p+1)!$
$k = 2p$	$N-k = (2\mu-2p+1)$ est impair	$M_{2p} = (2p-1)! (\mu-p)! (\mu-p+1)!$
$k = 2p+1$	$N-k = (2\mu-2p)$ est pair	$M_{2p+1} = (2p)! (\mu-p)! (\mu-p)!$
$k = 2p+2$	$N-k = (2\mu-2p-1)$ est impair	$M_{2p+2} = (2p+1)! (\mu-p-1)! (\mu-p)!$

Différents cas doivent être considérés

a. 1 minimum unique pour $k = 2p$, on a alors:

$$M_1 > M_2 \text{ -----} > M_{2p-1} > M_{2p} < M_{2p+1} \text{ ---} < M_N$$

on doit donc déterminer si il existe p tel que:

$$\frac{M_{2p}}{M_{2p-1}} < 1 \quad \text{et} \quad \frac{M_{2p+1}}{M_{2p}} > 1$$

b. 1 minimum unique pour $k = 2p+1$, on a alors:

$$M_1 > M_2 \text{ -----} > M_{2p} > M_{2p+1} < M_{2p+2} < \text{---} < M_N$$

on doit déterminer s'il existe p tel que:

$$\frac{M_{2p+1}}{M_{2p}} < 1 \quad \text{et} \quad \frac{M_{2p+2}}{M_{2p+1}} > 1$$

c. 1 minimum double pour $k = 2p$ et $k = 2p-1$, on a alors

$$M_1 > M_2 \text{ -----} > M_{2p-1} = M_{2p} < M_{2p+1} \text{ ---} < M_N$$

on doit déterminer si il existe p tel que:

$$\frac{M_{2p}}{M_{2p-1}} = 1$$

d. 1 minimum double pour $k = 2p$ et $k = 2p+1$, on a alors

$$M_1 > M_2 \text{ -----} > M_{2p-1} > M_{2p} = M_{2p+1} < M_{2p+2} \text{ ---} < M_N$$

on doit déterminer s'il existe p tel que:

$$\frac{M_{2p}}{M_{2p+1}} = 1$$

e. 1 minimum triple pour $k = 2p-1$, $k = 2p$, $k = 2p+1$, on a alors:

$$M_1 > M_2 \text{ -----} > M_{2p-1} = M_{2p} = M_{2p+1} < M_{2p+2} \text{ ---} < M_N$$

on doit déterminer s'il existe p tel que:

$$\frac{M_{2p}}{M_{2p-1}} = 1 \quad \text{et} \quad \frac{M_{2p+1}}{M_{2p}} = 1$$

f. l minimum triple pour $k = 2p, k = 2p+1, k = 2p+2$, on a alors:

$$M_1 > M_2 \text{ -----} > M_{2p-1} > M_{2p} = M_{2p+1} = M_{2p+2} < \text{---} < M_N$$

on doit déterminer si il existe p tel que:

$$\frac{M_{2p+1}}{M_{2p}} = 1 \quad \text{et} \quad \frac{M_{2p+2}}{M_{2p+1}} = 1$$

Considérons successivement ces différents cas en tenant compte que $N = 2\mu + 1$;

on a:

$$\frac{M_{2p}}{M_{2p-1}} = \frac{2_{p-1}}{\mu-p+1} ; \quad \frac{M_{2p+1}}{M_{2p}} = \frac{2p}{\mu-p+1} ; \quad \frac{M_{2p+2}}{M_{2p+1}} = \frac{2p+1}{\mu-p}$$

a. les conditions à vérifier entraînent $\frac{N+1}{6} < p < \frac{N+3}{6}$

N est impair et p doit être un entier; en posant $N = 6\phi - i$

avec $\phi = 1, 2 \text{ ---}$ et $i = 1, 3, 5$ (puisque N est impair), on montre que l'inégalité précédente n'est jamais vérifiée. Il ne peut donc y avoir de minimum pour k pair lorsque N est impair.

b. On doit vérifier que $\frac{N-3}{6} < p < \frac{N+1}{6}$ ce qui n'est possible que si N est de la forme $N = 6\phi - 5$ et le minimum est atteint pour $p = \phi - 1$, il vaut $M_{2\phi-1}$

- c. On doit vérifier que $p = \frac{N+3}{6}$, ce qui est possible si $N = 6\phi - 3$, on a alors $p = \phi$ et $M_{2\phi-1} = M_{2\phi}$
- d. On doit vérifier que $p = \frac{N+1}{6}$ ce qui est possible si $N = 6\phi - 1$, on a alors $p = \phi$ et $M_{2\phi} = M_{2\phi+1}$
- e. On doit avoir simultanément $p = \frac{N+3}{6}$ et $p = \frac{N+1}{6}$ ce qui est impossible
- f. On doit avoir simultanément $p = \frac{N+1}{6}$ et $p = \frac{N-3}{6}$ ce qui est impossible

En résumé, lorsque N est impair

- Si $N = 6\phi - 1$ et $N = 6\phi - 3$ on a un minimum double
- Si $N = 6\phi - 5$ on a un minimum unique
- Il ne peut y avoir de minimum triple

β) N est pair; on pose $N = 2\mu$

k varie de 1 à 2μ . Ici encore, on peut considérer les 6 cas précédents (de a à f). Mais ici puisque $N = 2\mu$, on a:

$$\frac{M_{2p}}{M_{2p-1}} = \frac{2p+1}{\mu-p+1} \quad ; \quad \frac{M_{2p+1}}{M_{2p}} = \frac{2p}{\mu-p} \quad ; \quad \frac{M_{2p+2}}{M_{2p+1}} = \frac{2p+1}{\mu-p}$$

on considère N de la forme $N = 6\phi - i$ avec $i = 0, 2, 4$ puisque N est pair

- a. on aura un minimum unique M_{2p} si $\frac{N}{6} < p < \frac{N+4}{6}$.

on peut montrer que l'on peut trouver p entier satisfaisant ces conditions si $N = 6\phi - 2$. On a alors $p = \phi$ et le minimum est $M_{2\phi}$.

- b. on aura un minimum M_{2p+1} si $\frac{N-2}{6} < p < \frac{N}{6}$

on peut montrer qu'il est impossible de réaliser ces conditions pour N pair ($i = 0, 2, 4$).

- c. on aura un minimum double si $p = \frac{N+4}{6}$; ceci est vérifié pour $N = 6\phi-4$ et le minimum vaut $M_{2\phi} = M_{2\phi-1}$
- d. on aura un minimum double si $p = \frac{N}{6}$; ceci est vérifié pour $N = 6\phi$; on a alors $p = \phi$ et le minimum vaut $M_{2\phi} = M_{2\phi+1}$
- e. on doit aussi simultanément $p = \frac{N+4}{6}$ et $p = \frac{N}{6}$ ce qui est impossible
- f. on doit aussi simultanément $p = \frac{N}{6}$ et $p = \frac{N-2}{6}$ ce qui est impossible

En résumé, lorsque N est pair

- Si $N = 6\phi$ et $N = 6\phi-4$, on a un minimum double
- Si $N = 6\phi-2$, on a un minimum unique
- Il ne peut y avoir de minimum triple

L'ensemble des résultats sont résumés dans la table 1 où l'on donne également les valeurs maximales pour $A_{k,j}$; car pour N fixé lorsque $Q_{k,j}$ est minimum $A_{k,j}$ est maximum.

N	Nombre de minimum	p	k	j	Minimum de $Q_{k,j}$	Maximum de $A_{k,j}$
6ϕ	2	ϕ	2ϕ $2\phi+1$	2ϕ $2\phi-1$ 2ϕ	$(2\phi)!(2\phi)!(2\phi-1)!$	$(6\phi)! / [(2\phi)!(2\phi)!(2\phi-1)!]$
$6\phi-1$	2	ϕ	2ϕ $2\phi+1$	$2\phi-1$ 2ϕ $2\phi-1$	$(2\phi)!(2\phi-1)!(2\phi-1)!$	$(6\phi-1)! / [(2\phi)!(2\phi-1)!(2\phi-1)!]$
$6\phi-1$	1	ϕ	2ϕ	$2\phi-1$	$[(2\phi-1)!]^3$	$(6\phi-2)! / [(2\phi-1)!]^3$
$6\phi-3$	2	ϕ	$2\phi-1$ 2ϕ	$2\phi-1$ $2\phi-2$ $2\phi-1$	$(2\phi-2)!(2\phi-1)!(2\phi-1)!$	$(6\phi-3)! / (2\phi-2)!(2\phi-1)!(2\phi-1)!$
$6\phi-4$	2	ϕ	$2\phi-1$ 2ϕ	$2\phi-2$ $2\phi-1$ $2\phi-2$	$(2\phi-1)!(2\phi-2)!(2\phi-2)!$	$(6\phi-4)! / (2\phi-1)!(2\phi-2)!(2\phi-2)!$
$6\phi-5$	1	$\phi-1$	$2\phi-1$	$2\phi-2$	$[(2\phi-2)!]^3$	$(6\phi-5)! / [(2\phi-2)!]^3$

Table 1: Maximum de $A_{k,j}$ suivant les valeurs de N

A.1.2 Détermination du maximum de $B_{k,j}$

On peut déduire $B_{k,j}$ de $A_{k,j}$, mais si on a pu mettre en évidence pour $A_{k,j}$ l'existence d'un maximum pour des valeurs particulières de k_0 et j_0 pour N donné et il est plus difficile d'en déduire les valeurs de k et j conduisant au maximum de $B_{k,j}$. On peut cependant dire, étant donné la forme de $B_{k,j}$ (et en tenant compte que $F_0 < 1$) que $B_{k,j}$ devient maximum pour $j \leq j_0$.

Pour N et k donnés, soit j_1 , la valeur conduisant au maximum de $A_{k,j}$, on a (cf.

$$A.2.1) \quad j_1 = \frac{N-k}{2} \quad \text{si } (N-k) \text{ pair}$$

$$j_1 = \frac{N-k-1}{2} \quad \text{si } (N-k) \text{ impair (étant donné la forme de } B_{k,j} \text{ la solution } j_1 = \frac{N-k+1}{2} \text{ ne présente pas d'intérêt)}$$

a. Si $(N-k)$ est pair, on peut montrer que:

$$\frac{B_{k,j_1}}{B_{k,j_1-1}} = \frac{N^2 - (-2)^2}{N^2 - k^2} F_0 \quad (\phi)$$

b. Si $(N-k)$ est impair, on peut montrer que:

$$\frac{B_{k,j_1}}{B_{k,j_1-1}} = \frac{N^2 - (k-3)^2}{(N-1)^2 - k^2} F_0 \quad (\beta)$$

A.1.2.1 Cas où $F_0 = 1$

a. Si $(N-k)$ est pair, on a:

$$\frac{B_{k,j_1}}{B_{k,j_1-1}} = \frac{N^2 - (k-2)^2}{N^2 - k^2}$$

Ce rapport est toujours supérieur à 1 puisque $(k-2) < k$

b. Si $(N-k)$ est impair, on a:

$$\frac{B_{k,j_1}}{B_{k,j_1-1}} = \frac{N^2 - (k-3)^2}{(N-1)^2 - k^2}$$

Si $N = 1$: $(N-k)$ est pair puisque $k = 1$

Si $N = 2$: $(K-k)$ est impair pour $k = 1$, on a alors $j_1 = 0$

Si $N \geq 3$, on a toujours $\left(\frac{B_{k,j_1}}{B_{k,j_1-1}} \right) > 1$

Donc, pour N et k donnés, lorsque $F = 1$, la valeur j_1 conduisant au maximum local de $A_{k,j}$ conduit au maximum local de $B_{k,j}$

On doit cependant encore examiner si le maximum général de $B_{k,j}$ est obtenu pour la valeur k_0 qui conduit au maximum général de $A_{k,j}$,

soient: j_1 la valeur donnant le maximum de $B_{k,j}$ pour $k = k_1$

j'_1 la valeur donnant le maximum de $B_{k,j}$ pour $k = k_1 - 1$

j''_1 la valeur donnant le maximum de $B_{k,j}$ pour $k = k_1 + 1$

k_1 étant une valeur quelconque

a. Si $(N-k_1)$ est pair

on a $j_1 = \frac{N-k_1}{2}$; $j'_1 = \frac{N-k_1}{2}$; $j''_1 = \frac{N-k_1-2}{2}$

et:

$$\frac{B_{k_1, j_1}}{B_{k_1-1, j'_1}} = \frac{N^2 - (k_1-2)^2}{2(k_1-1)(N+k_1)} \quad (1)$$

$$\frac{B_{k_1+1, j''_1}}{B_{k_1, j_1}} = \frac{N-k_1}{2k_1} \quad (2)$$

b. Si $(N-k_1)$ est impair

on a $j_1 = \frac{N-k_1-1}{2}$; $j'_1 = \frac{N-k_1-1}{2}$; $j''_1 = \frac{N-k_1+1}{2}$

et

$$\frac{B_{k_1, j_1}}{B_{k_1-1, j'_1}} = \frac{N-k_1+1}{2(k_1-1)} \quad (3)$$

$$\frac{B_{k_1+1, j_1''}}{B_{k_1, j_1}} = \frac{N^2 - (k_1-1)^2}{2k_1(N+k_1+1)} \quad (4)$$

En remplaçant k_1 par (k_1+1) dans (1) et (3), on obtient respectivement (4) et (2).

Examinons suivant la parité de N , en utilisant les résultats de la table 1, si le maximum de $B_{k,j}$ est obtenu pour k_0 qui conduit au maximum de $A_{k,j}$. Lorsque plusieurs valeurs de k_0 sont possibles, on les examine successivement et l'on choisit toujours j_0 de sorte que (k_0+j_0) soit le plus petit car on cherche le maximum de $B_{k,j}$.

1. Si $N = 6\phi$ avec $k_1 = k_0$ $\frac{B_{k_0}}{B_{k_0-1}} = \frac{(4\phi-1)(2\phi+1)}{(4\phi-2)(2\phi)} > 1$

• $k_0 = 2\phi$

d'après (1) avec $k_1 = k_0$ $\frac{B_{k_0}}{B_{k_0-1}} = \frac{(4\phi-1)(2\phi+1)}{(4\phi-2) 2\phi} > 1$

d'après (2) avec $k_1 = k_0$ $\frac{B_{k_0+1}}{B_{k_0}} = \frac{4\phi}{4\phi} = 1$

d'après (4) avec $k_1 = k_0+1$ $\frac{B_{k_0+2}}{B_{k_0+1}} = \frac{(2\phi)(4\phi)}{(2\phi+1)(4\phi+1)} < 1$

donc, $B_{k_0-1} < B_{k_0} = B_{k_0+1} > B_{k_0+2}$

ou encore $B_{2\phi-1} < B_{2\phi} = B_{2\phi+1} > B_{2\phi+2}$

• $k_0 = 2\phi+1$

d'après (3) avec $k_1 = k_0$ $\frac{B_{k_1}}{B_{k_0-1}} = \frac{4\phi}{4\phi} = 1$

d'après (1) avec $k_1 = k_0 - 1$
$$\frac{B_{k_0-1}}{B_{k_0-2}} = \frac{(4\phi-1)(2\phi+1)}{(4\phi-2)(2\phi)} > 1$$

d'après (4) avec $k_1 = k_0$
$$\frac{B_{k_0+1}}{B_{k_0}} = \frac{(2\phi)(4\phi)}{(2\phi+1)(4\phi+1)} < 1$$

Donc, $B_{k_0-2} < B_{k_0-1} = B_{k_0} > B_{k_0+1}$

ou encore $B_{2\phi-1} < B_{2\phi} = B_{2\phi+1} > B_{2\phi+2}$

Donc pour $N = 6\phi$, le maximum de $B_{k,j}$ est bien atteint pour la même valeur que le maximum de $A_{k,j}$ et l'on peut montrer que l'on a dans tous les cas:

$$B_{2\phi-1} < B_{2\phi} = B_{2\phi+1} > B_{2\phi+2}$$

Le maximum vaut:

$$B_M = B_{2\phi} = B_{2\phi+1} = \frac{6\phi!}{2\phi! 2\phi! (2\phi-1)! 4\phi}$$

2. Si $N = 6\phi - 1$

• $k_0 = 2\phi$

d'après (3) avec $k_1 = k_0$
$$\frac{B_{k_0}}{B_{k_0-1}} = \frac{4\phi}{4\phi-2} > 1$$

d'après (4) avec $k_1 = k_0$
$$\frac{B_{k_0+1}}{B_{k_0}} = \frac{8\phi-2}{8\phi+1} < 1$$

Donc $B_{2\phi-1} < B_{2\phi} > B_{2\phi+1}$

• $k_0 = 2\phi + 1$

d'après (1) pour $k_1 = k_0$
$$\frac{B_{k_0}}{B_{k_0-1}} = \frac{8\phi-2}{8\phi} < 1$$

d'après (3) pour $k_1 = k_0 - 1$
$$\frac{B_{k_0-1}}{B_{k_0-2}} = \frac{4\phi}{4\phi-2} > 1$$

Donc,

$$B_{2\phi-1} < B_{2\phi} > B_{2\phi+1}$$

Le maximum est $B_{2\phi}$, il vaut:

$$B_M = B_{2\phi} = \frac{(6\phi-1)!}{2\phi! (2\phi-1)! (2\phi-1)!} \cdot \frac{1}{4\phi-1}$$

3. Si $N = 6\phi-2$

• $k_0 = 2\phi$

d'après (1) pour $k_1 = k_0$

$$\frac{B_{k_0}}{B_{k_0-1}} = \frac{4\phi}{4\phi-1} > 1$$

d'après (2) pour $k_1 = k_0$

$$\frac{B_{k_0+1}}{B_{k_0}} = \frac{4\phi-2}{4\phi} < 1$$

Donc,

$$B_{2\phi-1} < B_{2\phi} > B_{2\phi+1}$$

Le maximum est $B_{2\phi}$, il vaut

$$B_M = B_{2\phi} = \frac{(6\phi-2)!}{[(2\phi-1)!]^3 (4\phi-1)}$$

4. Si $N = 6\phi-3$

• $k_0 = 2\phi-1$

d'après (1) on a pour $k_1 = k_0$

$$\frac{B_{k_0}}{B_{k_0-1}} = \frac{2\phi (4\phi-3)}{(2\phi-1)(4\phi-4)} > 1$$

d'après (2) on a pour $k_1 = k_0$

$$\frac{B_{k_0+1}}{B_{k_0}} = \frac{2\phi-1}{2\phi-1} = 1$$

d'après (4) on a pour $k_1 = k_0+1$

$$\frac{B_{k_0+2}}{B_{k_0+1}} = \frac{(4\phi-2)(2\phi-1)}{(4\phi-1)(2\phi)} < 1$$

$$B_{2\phi-2} < B_{2\phi-1} = B_{2\phi} > B_{2\phi+1}$$

• $k_0 = 2\phi$

d'après (1) on a pour $k_1 = k_0 + 1$ $\frac{B_{k_0-1}}{B_{k_0-2}} = \frac{2\phi(4\phi-3)}{(2\phi-1)(4\phi-4)} > 1$

d'après (2) on a pour $k_1 = k_0 - 1$ $\frac{B_{k_0}}{B_{k_0-1}} = \frac{2\phi-1}{2\phi-1} = 1$

d'après (4) on a pour $k_1 = k_0$ $\frac{B_{k_0+1}}{B_{k_0}} = \frac{(4\phi-2)(2\phi-1)}{(4\phi-1)(2\phi)} < 1$

Finalement $B_{2\phi-2} < B_{2\phi-1} = B_{2\phi} > B_{2\phi+1}$

Le maximum vaut:

$$B_M = B_{2\phi-1} = B_{2\phi} = \frac{(6\phi-3)!}{(2\phi-2)!(2\phi-1)!(2\phi-1)!} \cdot \frac{1}{4\phi-2}$$

5. $N = 6\phi-4$

• $k_0 = 2\phi-1$

d'après (3) on a pour $k_1 = k_0$ $\frac{B_{k_0}}{B_{k_0-1}} = \frac{2\phi-1}{2\phi-2} > 1$

d'après (4) on a pour $k_1 = k_0$ $\frac{B_{k_0+1}}{B_{k_0}} = \frac{4\phi-3}{4\phi-2} < 1$

Donc,

$$B_{2\phi-2} < B_{2\phi-1} > B_{2\phi}$$

• $k_0 = 2\phi$

d'après (3) on a pour $k_1 = k_0 - 1$ $\frac{B_{k_0-1}}{B_{k_0-2}} = \frac{2\phi-1}{2\phi-2} > 1$

d'après (4) on a pour $k_1 = k_0 - 1$
$$\frac{B_{k_0}}{B_{k_0-1}} = \frac{4\phi-3}{4\phi-2} < 1$$

$$B_{2\phi-2} < B_{2\phi-1} > B_{2\phi}$$

Le maximum vaut:

$$B_M = B_{2\phi-1} = \frac{(6\phi-4)!}{(2\phi-1)! (2\phi-2)! (2\phi-2)!} \cdot \frac{1}{4\phi-3}$$

6. $N = 6\phi-5$

• $k_0 = 2\phi-1$

d'après (1) on a pour $k_1 = k_0$
$$\frac{B_{k_0}}{B_{k_0-1}} = \frac{4\phi-2}{4\phi-3} > 1$$

d'après (2) on a pour $k_1 = k_0$
$$\frac{B_{k_0+1}}{B_{k_0}} = \frac{2\phi-2}{2\phi-1} < 1$$

Donc,

$$B_{2\phi-2} < B_{2\phi-1} > B_{2\phi}$$

Le maximum est atteint pour $k = 2\phi-1$ et il vaut:

$$B_M = B_{2\phi-1} = \frac{(6\phi-5)!}{[(2\phi-2)!]^3} \cdot \frac{1}{4\phi-3}$$

N	k	j	Maximum de $B_{k,j}$ pour $F_0 = 1$
6ϕ	$\frac{2\phi}{2\phi-1}$	$\frac{2\phi}{2\phi-1}$	$\frac{6\phi!}{2\phi! 2\phi! (2\phi-1)! 4\phi}$
$6\phi-1$	2ϕ	$2\phi-1$	$\frac{(6\phi-1)!}{2\phi! (2\phi-1)! (2\phi-1)! (4\phi-1)}$
$6\phi-2$	2ϕ	$2\phi-1$	$\frac{(6\phi-2)!}{[(2\phi-1)!]^3 (4\phi-1)}$
$6\phi-3$	$\frac{2\phi-1}{2\phi}$	$\frac{2\phi-1}{2\phi-2}$	$\frac{(6\phi-3)!}{(2\phi-2)! (2\phi-1)! (2\phi-1)!} \cdot \frac{1}{(4\phi-2)}$
$6\phi-4$	$2\phi-1$	$2\phi-2$	$\frac{(6\phi-4)!}{(2\phi-1)! (2\phi-2)! (2\phi-2)!} \cdot \frac{1}{(4\phi-3)}$
$6\phi-5$	$2\phi-1$	$2\phi-2$	$\frac{(6\phi-5)!}{[(2\phi-2)!]^3 (4\phi-3)}$

TABLE 2: Maximum de $B_{k,j}$ lorsque $F_0 = 1$

Le maximum absolu de $B_{k,j}$ pour N donné est atteint pour $F_0=1$; en effet lorsque $F_0 \neq 1$ le terme F_0^{k+j} a tendance à diminuer $B_{k,j}$ puisque $F_0 < 1$ et k et j sont entiers.

A.1.2.2 Cas où $F_0 \neq 1$

Lorsque $F_0 \neq 1$, en raison de l'influence du terme F_0^{k+j} et puisque $0 \leq F_0 \leq 1$ et que $(k+j)$ est entier, le maximum B'_M de $B_{k,j}$ pour N donné est inférieur au maximum B_M que l'on a obtenu pour $F_0 = 1$.

Pour F_0 donné différent de 1 il est impossible pour une valeur de N fixée de déterminer théoriquement quelles valeurs de k et de j conduisent au maximum B'_M . Pratiquement on peut cependant effectuer les calculs des $B_{k,j}$ pour déterminer B'_M (cf. A.2.2).

A.2 Calcul pratique de $H(Z,k)$

A.2.1 Cas où $F_0 = 1$

L'équation (9) de la distribution cumulée de la statistique d'ordre k peut s'écrire:

$$H(Z,k) = \sum_{j=0}^{N-k} (-1)^j B_{k,j}$$

avec

$$B_{k,j} = \frac{A_{k,j}}{(k+j)} \cdot F_0^{k+j}$$

$$\text{et } A_{k,j} = \frac{N!}{(k-1)! j! (N-k-j)!}$$

Pour N fixé on peut montrer (A.1.2) qu'il y a une valeur de k correspondant à B_M , la valeur maximum de $B_{k,j}$, lorsque $F_0 = 1$.

Avec $F_0 = 1$ théoriquement on doit trouver $H(Z,k) = 1$. En pratique en raison des valeurs élevées que peut prendre $B_{k,j}$, il peut en résulter une certaine perte de précision sur $H(Z,k)$; car on effectue la somme de nombres pouvant être très élevés et ayant un signe alterné, le résultat de la somme étant petit et égal à 1 pour $F_0 = 1$.

Deux éléments interviennent dans cette perte possible de précision:

- la grandeur de $B_{k,j}$
- le nombre $(N-k+1)$ de $B_{k,j}$ intervenant dans la sommation qui conduit au calcul de $H(Z,k)$.

Exemple $N = 60$

On peut montrer (A.1.2.1) que B_M est atteint à $k = 20$, $j = 20$ ainsi que à $k = 21$, $j = 19$ on a:

$$B_M = \frac{60!}{(40) (20!) (20!) (19!)} = .288 \times 10^{27}$$

$$B_M = \underbrace{288}_{27 \text{ chiffres}} + 1 \text{ chiffre}$$

En travaillant en double précision sur CDC on obtient 28 chiffres significatifs, le 28 ième étant arrondi, on a donc sur B_M une erreur maximale de 5.10^{-2} . Le calcul de $H(Z,20)$ sera l'addition de 41 termes de signes alternés. Lorsque j varie, les $B_{20,j}$ sont inférieurs à B_M mais peuvent être du même ordre de grandeur surtout pour j voisin de 20. On ne peut connaître l'erreur maximum sur $H(Z,20)$ si on ne calcule pas chaque $B_{k,j}$, mais on peut dire que dans ce cas puisque le résultat doit être $H(Z,20) = 1$, on a une erreur pouvant atteindre le premier chiffre après le point.

Pour $k = 60$, la sommation ne contient qu'un terme calculé à partir de nombres très élevés

$$B_{60,0} = \frac{60!}{59! \cdot 60}$$

Même en calculant séparément les factorielles on arrive à $H(Z,60) = 1$ avec 28 chiffres de précision.

Pour $k = 1$, la sommation contient 60 termes de grandeur variable et de signes alternés, on peut montrer (A.2.1.2) que le maximum de ces termes est atteint pour $j = 29$ et on a alors:

$$B_{1,29} = \frac{60!}{30 \cdot 0! \cdot 29! \cdot 30!} = .118 \times 10^{18}$$

Quand j varie les autres $B_{1,j}$ sont inférieurs, on est donc assuré d'avoir pour chaque $B_{1,j}$ au moins 10 chiffres significatifs après le point donc le résultat $H(Z,1)$ sera précis avec environ 10 chiffres significatifs après le point; en fait on trouve $H(Z,1) = 0.999999999$.

Le problème de précision se pose donc lorsque k est proche de la valeur conduisant au maximum B_M ($k = 20$ pour $N = 60$) car les $B_{k,j}$ deviennent élevés et la sommation comprend un nombre relativement important de termes (de l'ordre de 40).

A titre d'exemple en effectuant les calculs en double précision (28 chiffres significatifs) et en se limitant à 3 chiffres significatifs après le point pour $H(Z,k)$ on a pour $F_0 = 1$:

$k = 1$ à 12	$H(Z,k) = 1.000$
$k = 13$	$H(Z,13) = .998$
$k = 14$	$H(Z,14) = 1.001$
$k = 15$	$H(Z,15) = 1.002$
$k = 16$	$H(Z,16) = .994$
$k = 17$	$H(Z,17) = .989$
$k = 18$	$H(Z,18) = 1.008$
$k = 19$	$H(Z,19) = 1.011$
$k = 20$	$H(Z,20) = 1.018$
$k = 21$	$H(Z,21) = .974$
$k = 22$	$H(Z,22) = 1.033$
$k = 23$	$H(Z,23) = .999$
$k = 24$	$H(Z,24) = 1.003$
$k = 25$	$H(Z,25) = .996$
$k = 26$	$H(Z,26) = .992$
$k = 27$	$H(Z,27) = .997$
$k = 28$	$H(Z,28) = 1.001$
$k = 29$	$H(Z,29) = 1.000$
$k = 30$	$H(Z,30) = .999$
$k = 31$ à 60	$H(Z,k) = 1.000$

Théoriquement on devrait trouver rigoureusement $H(Z,k) = 1.000$ quelque soit k .

Exemple N = 100

Pour $N = 100$, B_M est atteint à $k = 34$, $j = 33$, on a :

$$B_M = .213 \times 10^{46}$$

Avec 28 chiffres significatifs l'erreur absolue sur un $B_{k,j}$ peut être de l'ordre de 10^{18} ; on ne peut donc s'attendre à trouver des bonnes valeurs de $H(Z,k)$ avec une telle erreur sur plusieurs termes de la sommation.

Choix du N optimum

On a calculé que pour $N = 54$ on a $B_M = .440 \times 10^{24}$; donc on peut obtenir $H(Z,k)$ avec une précision minimum de 3 chiffres significatifs après le point pour $N \leq 54$ et $F_0 = 1$.

A.2.2 Cas où $F_0 \neq 1$

F_0 étant compris entre 0 et 1 et puisqu'il intervient à la puissance $(k + j)$ entière et positive, il est clair que B'_M , le maximum de $B_{k,j}$ pour $F_0 \neq 1$, sera inférieur à B_M (obtenu pour $F_0 = 1$). De plus, en pratique on constate que pour N fixé B'_M est atteint à une valeur de k inférieure ou égale à celle correspondant à B_M ; ceci est dû à l'effet pondérateur de F_0^{k+j} ; B'_M est d'autant plus proche de B_M que F_0 est près de 1.

Par exemple on a avec $F_0 = .5$:

N = 40	$B_M = .125 \times 10^{18}$	$k = 14, j = 13$
	$B'_M = .123 \times 10^{11}$	$k = 10, j = 9$
N = 80	$B_M = .763 \times 10^{36}$	$k = 27, j = 26$
	$B'_M = .674 \times 10^{22}$	$k = 20, j = 19$
N = 100	$B_M = .213 \times 10^{46}$	$k = 34, j = 33$
	$B'_M = .567 \times 10^{28}$	$k = 25, j = 24$

A.2.3 Utilisation de la relation de symétrie

On a montré (équation 10 dans 2.3) que:

$$H[Z(F_0), N-k+1] = 1 - H[Z(1-F_0), k]$$

Cette relation permet de déduire $H(Z,k)$ pour $F_0 > .5$ à partir de $H(Z,k)$ obtenu pour $F_0 < .5$. B'_M étant plus petit lorsque F_0 diminue (A.1.2) on peut donc calculer $H(Z,k)$ pour un N plus grand tout en conservant une certaine précision. Par exemple pour garantir une précision de 3 chiffres significatifs sans utiliser la relation de symétrie on est limité à $N \leq 54$ tandis qu'en utilisant cette relation on peut se rendre à $N = 85$ puisque pour $N = 85$ on a $B'_M = .203 \times 10^{24}$ pour $F_0 = .5$ et que $H(Z,k)$ est obtenu en effectuant la somme de termes inférieurs à B'_M . De plus lorsque F_0 décroît B'_M diminue pour un N donné.

A.3 Calcul des intervalles de confiance

La méthode utilisée dans le programme est décrite en 2.3. Cette méthode nécessite la recherche de $F_{0,1}$ et $F_{0,2}$ tels que:

$$H(Z_{1,k},k) = \alpha/2 \text{ et } H(Z_{2,k},k) = 1 - \alpha/2$$

$$\text{avec } Z_{i,k} = -a (1 - F_{0,i})^b + c (1 - F_{0,i})^{-d} + e \text{ pour } i = 1, 2$$

On peut affirmer que la solution $F_{0,i}$ est comprise dans l'intervalle $(F_1 = .00001, .99999 = F_3)$; soit F_2 le centre de cet intervalle, si $H(Z(F_2),k) = \beta$ ($\beta = \alpha/2$ ou $1 - \alpha/2$) avec une précision de 10^{-5} alors F_2 est la solution cherchée sinon $F_{0,i}$ est dans l'intervalle (F_1, F_2) ou (F_2, F_3) selon que β est plus petit ou plus grand que $H(Z(F_2),k)$ (resp.). On divise alors (F_1, F_2) ou (F_2, F_3) et on compare avec β la valeur de $H(Z,k)$ calculée au centre de cet intervalle; on arrête si les valeurs sont égales sinon on divise en deux l'intervalle dans lequel se trouve $F_{0,i}$, on continue ainsi jusqu'à ce que l'on obtienne F_2 tel que $|H(Z(F_2),k) - \beta| < 10^{-5}$.

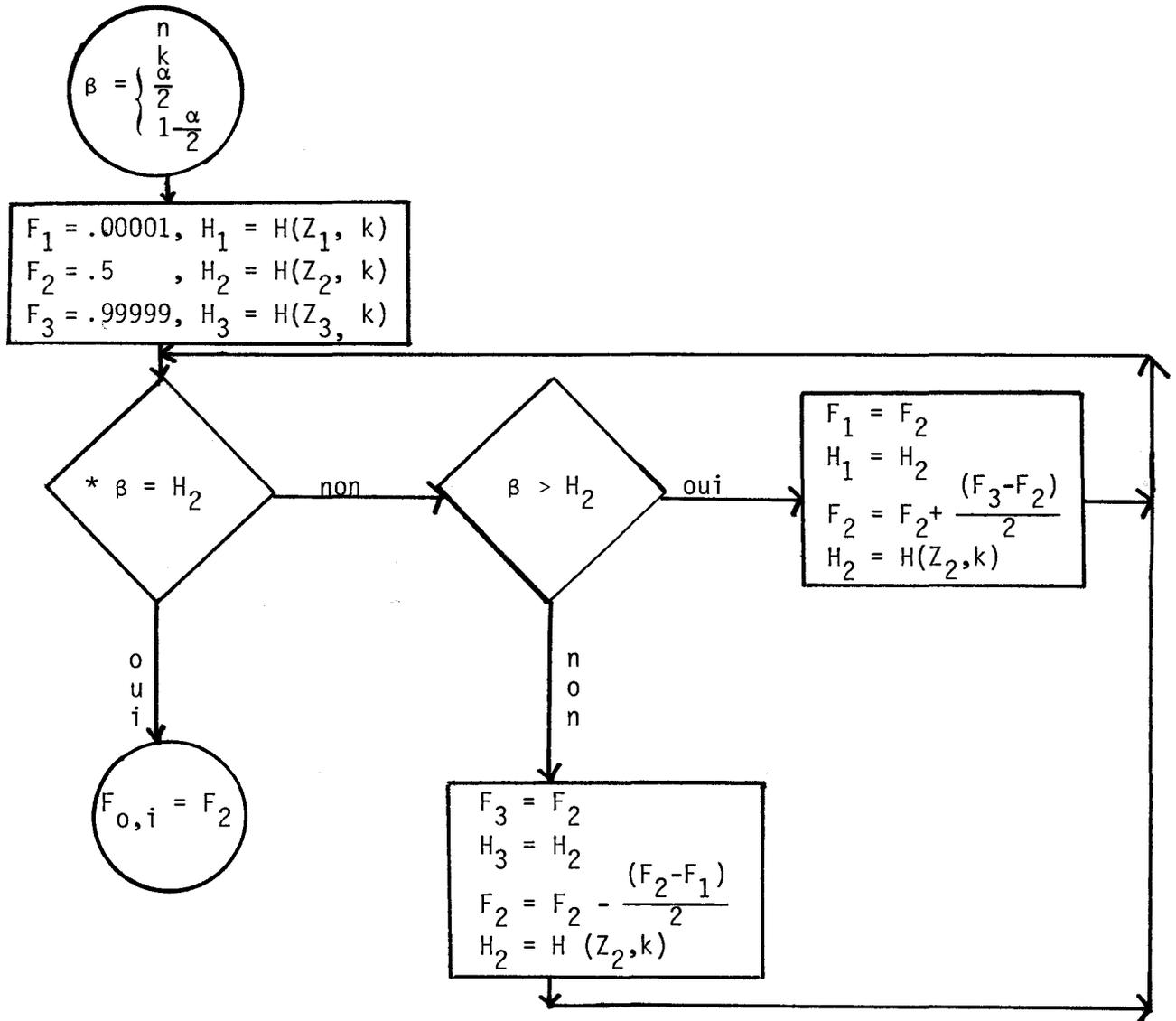
On peut illustrer ce procédé itératif par le diagramme de la page suivante.

A.4 Programme de calcul ITCO

Le programme ITCO calcule de façon automatique les intervalles de confiance de la loi de Wakeby. Le programme est composé:

- d'un programme principal qui détermine différentes valeurs de k
- de la sous-routine CONFI qui effectue la recherche de $F_{0,i}$ et calcule $Z(i,k)$

On peut illustrer le procédé itératif par le diagramme suivant:



* En pratique, on considère que $\beta = H_2$ si $|\beta - H_2| < .00001$

- de la fonction HZK qui effectue le calcul de $H(Z,k)$

A.4.1 Détermination de différentes valeurs de k

Pour tracer les intervalles de confiance ou la courbe théorique, on cherche à obtenir de 20 à 30 valeurs de Z correspondant à autant de valeurs de k.

Le choix des valeurs de k se fait de la manière suivante:

soient: N la taille de l'échantillon

$$N1 = \left[\frac{N}{6} + .5 \right],$$

$$N2 = \left[\frac{5N}{6} + .5 \right] + 1 \quad [] \text{ signifie partie entière}$$

$$N3 = \left[\frac{N}{15} \right] + 1$$

Les N1 premières valeurs de k ($k = 1, \dots, N1$) sont conservées, ensuite on conserve les valeurs comprises entre N1 et N2 avec un pas de N3 ($N1 + N3, N1 + 2N3, \dots$) et toutes les valeurs de k de N2 à N. Pour ces différentes valeurs de k on calcule les limites $Z_{1,k}$ et $Z_{2,k}$ de l'intervalle de confiance à un niveau donné $(1-\alpha)\%$.

A.4.2 Utilisation du programme

A.4.2.1 Données d'entrée

La lecture des différents paramètres se fait sur une seule carte.

N: taille de l'échantillon

IPE: code de la probabilité empirique choisie

= 0 Hazen

= 1 Weibull

= 2 Chegodayev

NIC: nombre d'intervalles de confiance que l'on désire calculer

(NIC \leq 3)

IC(I): niveau de confiance (en %) des intervalles désirés

(I = 1, 2, 3)

A,B,C,D,E: paramètres de la distribution Wakeby.

Le format de lecture est (6I5,5F10.2)

On répète cette carte si on désire traiter un autre échantillon, on termine par une carte blanche pour indiquer la fin des cas traités.

A.4.2.2 Sortie des résultats

- 1) Titre
- 2) Probabilité empirique choisie
- 3) Définition de Z, k, F, Z(1,k) et Z(2,k)
- 4) Paramètres de la distribution
- 5) Taille de l'échantillon (N)
- 6) Pour chacun des niveaux de confiance choisis:
 - niveau de confiance
 - pour chaque k: on écrit k, probabilité empirique, Z, Z(1,k), Z(2,k)

A.5 Listing du programme et exemple de traitement

On suppose un échantillon de taille $N = 40$, les paramètres de la distribution Wakeby valent:

$$A = 0.5 \quad B = 2.0 \quad C = 1.5 \quad D = 0.2 \quad E = -1.0$$

La probabilité empirique choisie est celle de Hazen. On veut obtenir l'intervalle de confiance aux niveaux 99%, 95% et 90%.

La carte de donnée est:

```

_40_ _0_ _3_ _99_ _95_ _90_ _5_ _2_ _15_ _2_
-----
-----1-----

```

Les pages suivantes donnent le listing du programme et son application à l'exemple considéré.

PROGRAM ITCO

73/171 OPT=0 TRACE

FTN 4,8+506

81/01/16

```

1      PROGRAM ITCO(INPUT,OUTPUT,TAPE1)
      C
      C      PROGRAMME DE CALCUL AUTOMATIQUE DES INTERVALLES DE
      C      CONFIANCE DE LA DISTRIBUTION DE WAKEBY, LE CALCUL SE
5      C      FAIT AU MOYEN DES STATISTIQUES D ORDRE
      C
      C      VARIABLES LUES
      C
      C      - N TAILLE DE L ECHANTILLON, N ≤ 85
10     C
      C      - IPE PROBABILITE EMPIRIQUE
      C          =0 HAZEN
      C          =1 WETBULL
      C          =2 CHEGODAYEV
15     C
      C      - NIC NOMBRE D INTERVALLES DE CONFIANCE
      C          QUE L ON VEUT CALCULER, NIC ≤ 3
      C
      C      - IC(I) NIVEAU DE CONFIANCE (EN %) DES
20     C          INTERVALLES DE CONFIANCE
      C
      C      - A,B,C,D,E PARAMETRES DE LA DISTRIBUTION
      C
25     C
      C      DIMENSION KK(100),PE(4,100),IC(3)
      C      1 READ 900,N,IPE,NIC,IC(1),IC(2),IC(3),A,B,C,D,E
      C      IF(N,EQ,0) STOP
      C
      C      DETERMINATION DES K (ORDRE DE LA STATISTIQUE)
30     C
      C      N1=INT(N/6,+5)
      C      N2=INT(5.*N/6,+5)+1
      C      N3=INT(N/15.)+1
      C      IM=1
35     C      IN=0
      C      DO 4 IK=1,100
      C      IF(IK.GT.N1) GO TO 2
      C      KK(IK)=IK
      C      GO TO 4
40     C      2 IT=N1+IM*N3
      C      IF(IT.GE.N2) GO TO 3
      C      KK(IK)=IT
      C      IM=IM+1
      C      GO TO 4
45     C      3 IT=N2+IN
      C      IF(IT.GT.N) GO TO 5
      C      KK(IK)=IT
      C      IN=IN+1
50     C      4 CONTINUE
      C      5 IK=IK-1
      C      PRINT 901
      C
      C      PROBABILITE EMPIRIQUE
55     C
      C      IF(IPE=1) 6,8,10
      C      6 PRINT 902
      C      DO 7 KI=1,IK

```

PROGRAM ITCU

73/171 OPT=0 TRACE

FTN 4,8+508

81/01/10

```

        PE(1,KI)=(KK(KI)-0,5)/N
60      7 CONTINUE
        GO TO 12
        8 PRINT 903
        DO 9 KI=1,IK
          PE(1,KI)=KK(KI)/(N+1,.)
65      9 CONTINUE
        GO TO 12
        10 PRINT 904
        DO 11 KI=1,IK
          PE(1,KI)=(KK(KI)-0,3)/(N+0,4)
70      11 CONTINUE
        C
        12 PRINT 905
        PRINT 906
        PRINT 907
        PRINT 908
75      PRINT 909
        PRINT 910
        PRINT 911,A,B,C,D,E
        PRINT 912,N
        DO 14 ICI=1,NIC
80      ALPHA=(100,-IC(ICI))/100,
        CALL CONF1(N,KK,PE,IK,ALPHA,A,B,C,D,E)
        PRINT 913,IC(ICI)
        PRINT 914
        DO 13 J=1,IK
85      PRINT 915,KK(J),PE(1,J),PE(2,J),PE(3,J),PE(4,J)
        WRITE(1,916) KK(J),PE(1,J),PE(2,J),PE(3,J),PE(4,J)
        13 CONTINUE
        14 CONTINUE
        GOTO 1
90      C
900     FORMAT(6I5,5F10,2)
901     FORMAT(1H1,/,30X,*INTERVALLE DE CONFIANCE DE*,/,32X,*LA DISTRIBUTI
        .ON WAKEBY*,//)
95     902 FORMAT(5X,*PROBABILITE EMPIRIQUE CHOISIE PK=(K-0,5)/N (HAZEN)*/)
        903 FORMAT(5X,*PROBABILITE EMPIRIQUE CHOISIE PK=K/(N+1) (WEIBULL)*/)
        904 FORMAT(5X,*PROBABILITE EMPIRIQUE CHOISIE PK=(K-0,3)/(N+0,4) (CHE
        .GUDAYEV)*,/)
        905 FORMAT(5X,*Z: COURBE THEORIQUE POUR LES PARAMETRES DONNES*,/)
        906 FORMAT(5X,31HZ=-A*(1-F)**B + C*(1-F)**D + E,/)
100    907 FORMAT(5X,*K: ORDRE DE LA STATISTIQUE (K=1,.,.,N)*,/)
        908 FORMAT(5X,*F: FONCTION DE DISTRIBUTION DE LA LOI WAKEBY*,/)
        909 FORMAT(5X,*Z(1,K): LIMITE INFERIEURE DE L INTERVALLE DE CONFIANCE P
        .OUR L ORDRE K*,/)
105    910 FORMAT(5X,*Z(2,K): LIMITE SUPERIEURE DE L INTERVALLE DE CONFIANCE P
        .OUR L ORDRE K*,/)
        911 FORMAT(5X,*PARAMETRES DE LA DISTRIBUTION: *,*A=*,F7,3,4X,*B=*,F7,3
        .,4X,*C=*,F7,3,4X,*D=*,F7,3,4X,*E=*,F7,3,/)
        912 FORMAT(5X,*TAILLE DE L ECHANTILLON N=*,I4,/)
110    913 FORMAT(1H1,/,30X,*NIVEAU DE CONFIANCE:*,I4,*X*,//)
        914 FORMAT(15X,*K*,8X,*PROB. EMP. *,7X,*Z*,12X,*Z(1,K)*,7X,*Z(2,K)*,/)
        915 FORMAT(13X,I3,7X,F7,3,7X,F7,3,9X,F7,3,6X,F7,3,/)
        916 FORMAT(1X,I4,4F12,4)
        END

```


SUBROUTINE CONF1

PE(3,J)=SNGL(Z(1))

PE(4,J)=SNGL(Z(2))

5 CONTINUE

RETURN

END

60

FUNCTION HZK

73/171 DPT=0 TRACE

FTN 4,8+508

81/01/16. 23.00,0

```

1      DOUBLE FUNCTION HZK(F1,N,KA)
      C
      C      FONCTION QUI CALCULE H(Z(F),K)
      C
5      C      VARIABLES
      C          = F1 F ( PROB. NON-DEPAS. )
      C          = N TAILLE DE L ECHANTILLON
      C          = KA K,ORDRE DE LA STATISTIQUE
10     C
      C      DOUBLE H,F,F1,X,Y,Y1,Q1
      C      DOUBLE FAC
      C      Y=1,D0
15     C      Y1=-1,D0
      C      F=F1
      C      K=KA
      C
      C      UTILISATION DE LA SYMETRIE SI F > ,5
20     C
      C          H(Z(F),K)=1-H(Z(1-F),N-K+1)
      C
      C      IF(F.GT.,.5D0) K=N-KA+1
      C      IF(F.GT.,.5D0) F=1,D0-F
25     C      NK=N-K
      C      NK1=NK+1
      C      H=0,D0
      C      IF(K.NE.N) GO TO 1
      C      H=F**N
30     C      GO TO 3
      C
      C      CALCUL DE H
      C
35     C      1 Q1=(F**K)*(FAC(N)/FAC(K-1))
      C      DO 2 JJ=1,NK1
      C      J=JJ-1
      C      X=(F**J)*(Q1/(FAC(J)*FAC(N-K-J)))/(K+J)
      C      H=H+X*(Y1**J)
      C      2 CONTINUE
40     C      3 IF(F1.GT.,.5D0) H=1,D0-H
      C      HZK=H
      C      RETURN
      C      END

```

SYMBOLIC REFERENCE MAP (R=1)

ENTRY POINTS
6 HZK

VARIABLES	SN	TYPE	RELOCATION			F.P.
246 F		DOUBLE		0	F1	DOUBLE
244 H		DOUBLE		242	HZK	DOUBLE
264 J		INTEGER		263	JJ	INTEGER

```
FUNCTION    FAC
           DOUBLE FUNCTION FAC(N)
           C
           C    FONCTION CALCULANT FACTORIELLE DE N
           C
5          DOUBLE A
           A=1.D0
           IF(N.EQ.0) GO TO 2
           DO 1 J=1,N
           A=A*J
10         1 CONTINUE
           2 FAC=A
           RETURN
           END
```

INTERVALLE DE CONFIANCE DE
LA DISTRIBUTION DE WAKEBYPROBABILITE EMPIRIQUE CHOISIE $P_k = (k-0.5)/N$ (HAZEN)

Z: COURBE THEORIQUE POUR LES PARAMETRES DONNES

 $Z = -A*(1-F)**B + C*(1-F)**D + E$

K: ORDRE DE LA STATISTIQUE (K=1,...,N)

F: FONCTION DISTRIBUTION DE LA LOI DE WAKEBY

Z(1,K): LIMITE INFERIEURE DE L'INTERVALLE DE CONFIANCE POUR L'ORDRE K

Z(2,K): LIMITE SUPERIEURE DE L'INTERVALLE DE CONFIANCE POUR L'ORDRE K

PARAMETRES DE LA DISTRIBUTION: A= .50 B= 2.00 C= 1.50 D= .20 E= -1.00

TAILLE DE L'ECHANTILLON N= 40

NIVEAU DE CONFIANCE: 90%

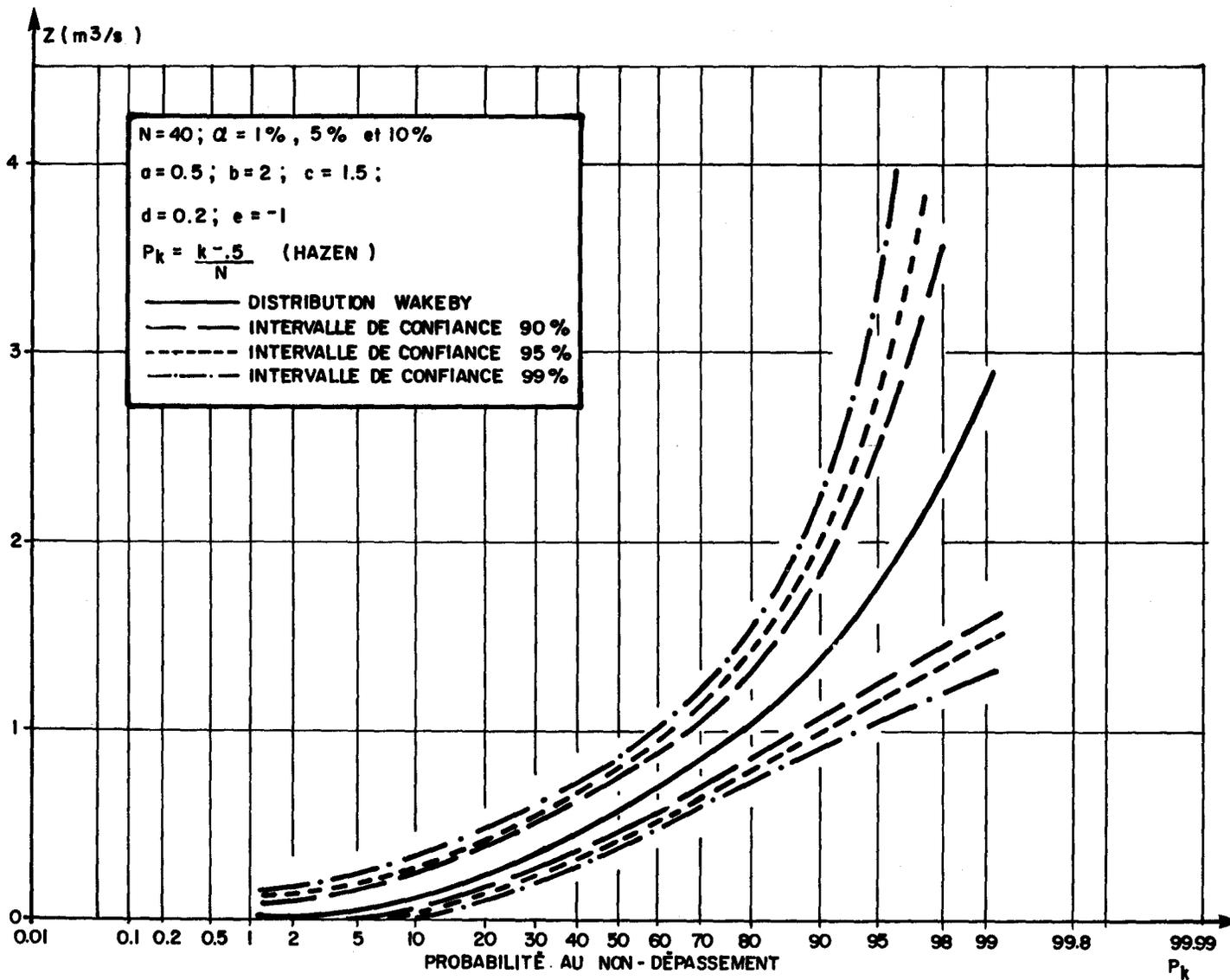
K	PROB. EMP.	Z	Z(1,K)	Z(2,K)
1	.012	.016	.002	.092
2	.037	.048	.012	.143
3	.063	.080	.027	.187
4	.087	.111	.045	.228
5	.112	.142	.065	.266
6	.137	.173	.086	.302
7	.162	.203	.108	.337
10	.237	.293	.179	.435
13	.313	.380	.253	.528
16	.387	.467	.330	.619
19	.462	.554	.410	.710
22	.537	.643	.492	.807
25	.612	.738	.578	.914
28	.688	.844	.671	1.040
31	.762	.971	.776	1.205
34	.837	1.144	.905	1.452
35	.862	1.221	.957	1.570
36	.887	1.316	1.018	1.723
37	.912	1.438	1.091	1.934
38	.938	1.610	1.184	2.256
39	.962	1.892	1.313	2.852
40	.987	2.603	1.535	4.682

NIVEAU DE CONFIANCE: 95%

K	PROB. EMP.	Z	Z(1,K)	Z(2,K)
1	.012	.016	.001	.112
2	.037	.048	.008	.166
3	.063	.080	.020	.212
4	.087	.111	.036	.253
5	.112	.142	.054	.292
6	.137	.173	.073	.329
7	.162	.203	.094	.364
10	.237	.293	.160	.464
13	.313	.380	.231	.557
16	.387	.467	.306	.649
19	.462	.554	.383	.742
22	.537	.643	.464	.840
25	.612	.738	.549	.950
28	.688	.844	.640	1.083
31	.762	.971	.742	1.259
34	.837	1.144	.866	1.526
35	.862	1.221	.916	1.658
36	.887	1.316	.973	1.829
37	.912	1.438	1.041	2.068
38	.938	1.610	1.126	2.441
39	.962	1.892	1.242	3.158
40	.987	2.603	1.434	5.544

NIVEAU DE CONFIANCE: 99%

K	PROB. EMP.	Z	Z(1,K)	Z(2,K)
1	.012	.016	.000	.157
2	.037	.048	.003	.214
3	.063	.080	.011	.262
4	.087	.111	.022	.306
5	.112	.142	.036	.346
6	.137	.173	.052	.383
7	.162	.203	.069	.419
10	.237	.293	.127	.520
13	.313	.360	.191	.615
16	.387	.467	.261	.708
19	.462	.554	.334	.804
22	.537	.643	.412	.906
25	.612	.738	.493	1.026
28	.688	.844	.581	1.173
31	.762	.971	.679	1.373
34	.837	1.144	.794	1.691
35	.862	1.221	.840	1.852
36	.887	1.316	.891	2.066
37	.912	1.438	.951	2.376
38	.938	1.610	1.024	2.880
39	.962	1.892	1.119	3.926
40	.987	2.603	1.269	8.046



Détermination des intervalles de confiance d'une distribution Wakeby à l'aide des statistiques d'ordre.