

Record Number: 1370
Author, Monographic: Boucher, P.//Boisvert, J.//Villeneuve, J. P.//Bobée, B.
Author Role:
Title, Monographic: Comparaison de méthodes de simulation de la loi Pearson 3
Translated Title:
Reprint Status:
Edition:
Author, Subsidiary:
Author Role:
Place of Publication: Québec
Publisher Name: INRS-Eau
Date of Publication: 1981
Original Publication Date: Février 1981
Volume Identification:
Extent of Work: 29
Packaging Method: pages
Series Editor:
Series Editor Role:
Series Title: INRS-Eau, Rapport de recherche
Series Volume ID: 136
Location/URL:
ISBN: 2-89146-134-7
Notes: Rapport annuel 1980-1981
Abstract: 10.00\$
Call Number: R000136
Keywords: rapport/ ok/ dl

COMPARAISON DE METHODES
DE SIMULATION DE LA LOI PEARSON 3

par

P. Boucher, J. Boisvert,
J.-P. Villeneuve et B. Bobée

RAPPORT SCIENTIFIQUE

No 136

Université du Québec
Institut national
de la recherche scientifique
INRS-Eau
C.P. 7500
Ste-Foy, Québec
G1V 4C7

Février 1981

COMPARAISON DE METHODES
DE SIMULATION DE LA LOI PEARSON 3

par

P. Boucher, J. Boisvert,
J.-P. Villeneuve et B. Bobée

RAPPORT SCIENTIFIQUE

No 136

Université du Québec
Institut national
de la recherche scientifique
INRS-Eau
C.P. 7500
Ste-Foy, Québec
G1V 4C7

Février 1981

INTRODUCTION

L'utilisation de la loi Pearson 3 et de ses formes dérivées (Gamma, log Pearson type 3) est très répandue en hydrologie, notamment pour représenter les phénomènes de crue et d'étiage, suscitant, de ce fait, de nombreuses études sur ces lois.

Le but de ce rapport est de fournir une méthode de génération de variate Pearson type 3, afin de permettre d'étudier par simulation les propriétés d'échantillonnage de ces lois. Quatre techniques sont comparées entre elles; il s'agit de:

- la méthode de Johnk;
- la transformation de Wilson-Hilferty;
- l'interpolation de la table de Harter;
- l'approximation polynomiale de la table de Harter.

Parmi les algorithmes de génération d'une variate Pearson type 3, l'algorithme de Johnk est celui qui conduit, en général, aux meilleurs résultats (Bobée et Boucher, 1979); c'est une technique relativement fiable mais qui demeure approximative. La transformation de Wilson-Hilferty simule directement une variate Pearson type 3 à partir d'une

variante normale et cette transformation est aussi une approximation. Quant aux deux dernières méthodes, ce sont des approximations numériques faites à partir de la table des variates Pearson 3 standardisée développée par Harter (Harter, 1969).

1. Méthodes de simulation

1.1 Méthode par interpolation de la table de Harter

La base de cette méthode consiste à interpoler une variate Pearson 3 standardisée K , de coefficient d'asymétrie C_S et de probabilité au non-dépassement P , à partir d'un polynôme du second degré de la forme:

$$Y = Ax^2 + Bx + C$$

où A, B, C sont déterminés à partir de la table de Harter.

Connaissant l'équation de la courbe passant par trois points de la table, soit (K_1^T, C_{S_1}) , (K_2^T, C_{S_2}) , (K_3^T, C_{S_3}) pour une même probabilité P_i au non-dépassement définie dans la table, il est possible d'interpoler un point (K_i, C_S) pour cette probabilité P_i et pour une valeur donnée C_S du coefficient d'asymétrie de la variate K recherchée. En répétant ce processus pour trois probabilités différentes mais à une même valeur de C_S nous interpolons trois points de coordonnées: $(K_1, P_1)_{C_S}$, $(K_2, P_2)_{C_S}$, $(K_3, P_3)_{C_S}$. Il suffit de déterminer l'équation de la courbe qui passe par ces trois points pour obtenir, par interpolation, le point cherché, soit $(K, P)_{C_S}$.

En pratique, le cheminement suivi peut se résumer en trois étapes.

- 1) Choix de C_{S_1} , C_{S_2} , C_{S_3} , P_1 , P_2 , P_3 soit des valeurs présentes dans la table et les plus près de C_S et de P . Ces six points fournissent 9 variates K_j^T :

$$K_j^T \equiv K^T (P_i, C_{S_i})$$

- où i varie de 1 à 3
 j varie de 1 à 9
 K^T variate Pearson 3 standardisée fournie par la table de Harter.

- 2) Trois variates K_1 , K_2 , K_3 sont calculées aux probabilités P_1 , P_2 , P_3 et au coefficient d'asymétrie C_S .

$$K_1 = K (P_1, C_S) = A_1 C_S^2 + B_1 C_S + C_1$$

$$K_2 = K (P_2, C_S) = A_2 C_S^2 + B_2 C_S + C_2$$

$$K_3 = K (P_3, C_S) = A_3 C_S^3 + B_3 C_S + C_3$$

où les coefficients A_i , B_i et C_i sont les solutions des équations suivantes:

$$Y_{1, i} = K^T (P_i, C_{S_1}) = A_i C_{S_1}^2 + B_i C_{S_1} + C_i$$

$$Y_{2, i} = K^T (P_i, C_{S_2}) = A_i C_{S_2}^2 + B_i C_{S_2} + C_i$$

$$Y_{3,i} = K^T (P_i, C_{S_3}) = A_i C_{S_3}^2 + B_i C_{S_3} + C_i$$

$$B_i = \frac{((Y_{3,i} - Y_{2,i}) \cdot (C_{S_2}^2 - C_{S_1}^2)) - ((Y_{2,i} - Y_{1,i}) \cdot (C_{S_3}^2 - C_{S_2}^2))}{(C_{S_3} - C_{S_2}) \cdot (C_{S_2} - C_{S_1}) \cdot (C_{S_1} - C_{S_3})}$$

$$A_i = \frac{(Y_{2,i} - Y_{1,i}) - B_i \cdot (C_{S_2} - C_{S_1})}{(C_{S_2}^2 - C_{S_1}^2)}$$

$$C_i = Y_{2,i} - (A_i \cdot C_{S_2}^2) - (B_i \cdot C_{S_2})$$

3) La variate K recherchée pour C_S et P s'exprime maintenant ainsi:

$$K = K(P, C_S) = A_4 P^2 + B_4 P + C_4$$

où A_4 , B_4 et C_4 sont les solutions des équations:

$$K_1 = A_4 P_1^2 + B_4 P_1 + C_4$$

$$K_2 = A_4 P_2^2 + B_4 P_2 + C_4$$

$$K_3 = A_4 P_3^2 + B_4 P_3 + C_4$$

1.2 Approximation polynomiale de la table de Harter

A une probabilité P , la variate Pearson type 3 standardisée K , s'exprime par l'équation suivante:

$$K = \sum_{i=0}^8 A_i C_S^i$$

où C_S est le coefficient d'asymétrie.

Les coefficients A_i sont des expressions polynomiales fonction de la probabilité P :

$$A_i = \sum_{j=0}^n L_{i,j} P^j$$

où $n = 8$ si $.0001 \leq P < .04$ ou $P \geq .95$
 $n = 7$ si $.04 \leq P < .95$

Les coefficients $L_{i,j}$ ont été déterminés par régression polynomiale. La procédure suivie peut se résumer ainsi:

- a) une matrice de coefficients $B_{i,k}$ est formée au moyen de polynomes orthogonaux pour chacune des 27 probabilités de la table de Harter. Cette matrice compte 8 colonnes (i) et 27 lignes (k).
- b) les coefficients $L_{i,j}$ sont obtenus en résolvant huit fois (pour chaque i), le système d'équation suivant:

$$B_{i,1} = f(p_1) = L_{i,0} p_1^0 + \dots + L_{i,n} p_1^n$$

$$B_{i,k} = f(p_k) = L_{i,0} p_k^0 + \dots + L_{i,n} p_k^n$$

où p_1, \dots, p_k sont des probabilités définies dans la table de Harter
i varie de 0 à 8
n et *k* sont fonction de la probabilité (voir paragraphe c ci-dessous).

c) les meilleurs résultats ont été obtenus en regroupant les probabilités en 4 classes. Ceci signifie qu'il a fallu, en fait, résoudre 32 fois le système d'équation ci-haut, soit 8 fois pour chacune des 4 classes.

$$- \quad .0001 \leq p_k < .04$$

$$k = 7, n = 8$$

$$- \quad .04 \leq p_k < .5$$

$$k = 6, n = 7$$

$$- \quad .5 \leq p_k < .95$$

$$k = 6, n = 7$$

$$- \quad .95 \leq p_k$$

$$k = 7, n = 8$$

L'expression finale de la variate *K* à la probabilité *P* et pour le coefficient d'asymétrie C_S devient:

$$K = \sum_{i=0}^8 \left(\sum_{j=0}^n L_{i,j} p^j \right) C_S^i$$

1.3 Transformation de Wilson-Hilferty

La transformation de Wilson-Hilferty peut s'écrire:

$$K = \frac{2}{C_S} \left\{ \left[1 - \left(\frac{C_S}{6} \right)^2 + \left(\frac{C_S}{6} \right) t \right]^3 - 1 \right\}$$

où K est la variate Pearson 3 standardisée, C_S l'asymétrie et t la variate normale standardisée de même probabilité que K . La borne inférieure théorique de la loi Pearson type 3 est $-2/C_S$. Cette borne n'est pas toujours respectée par la transformation de Wilson-Hilferty; on utilise alors la transformation de Wilson-Hilferty tronquée définie par:

$$K^T = \max \left(-2/C_S, K \right)$$

En pratique, on génère une probabilité P au non-dépassement (variate uniforme), on en déduit t et K correspondant à une asymétrie donnée.

1.4 Méthode de Johnk

Soient U et U_i des variates uniformes $(0, 1)$.

$$Z = -\ln \prod_{i=1}^{[\lambda]} U_i \quad \text{où } [\lambda] \text{ est la partie entière de } \lambda$$

$$Y = -\ln U$$

$$W = \beta [(\lambda - [\lambda], 1 - \lambda + [\lambda])] \quad (\text{variate Bêta})$$

Alors $J = \frac{Z + WY}{\alpha}$ est une variate Gamma de paramètres α, λ .

La variate $B \sim \beta(a, b)$ est obtenue par $B = \frac{X}{X + Y}$

où $X = U_1^{1/a}$

$$Y = U_2^{1/b}$$

avec $X + Y < 1$

La variate Pearson type 3 de paramètres α, λ et m est obtenue par

$$J_{p3} = J + m$$

m étant le paramètre d'origine de la loi Pearson 3.

En pratique, on génère des variates uniformes et après transformation on déduit J_{p3} .

2. Comparaisons des méthodes de simulation

Pour comparer les méthodes de simulation, on génère, pour certaines valeurs de α et λ un nombre important d'échantillons de taille N . Le paramètre d'origine m de la loi Pearson 3 est fixé à 0; on a

donc des échantillons provenant d'une loi Gamma de paramètres α et λ . On effectue l'ajustement de la loi Gamma par le maximum de vraisemblance et on peut, à ce moment, comparer entre elles les valeurs des paramètres estimées ainsi qu'avec les valeurs initiales (α_0, λ_0) .

Les premières simulations nous amènent à éliminer au départ la méthode d'approximation polynomiale, cette méthode demandant beaucoup trop de temps d'exécution (environ 7 fois plus que la méthode d'interpolation) tout en n'apportant aucun gain quant à la qualité de la simulation. Ce qui suit s'applique seulement aux 3 autres méthodes de simulation.

2.1 Plan de simulation

Les paramètres α et λ sont déterminés en fixant l'asymétrie C_S et la variance. Huit valeurs de C_S sont retenues variant de .25 à 4, la variance est fixée à 1. Le tableau 2.1 donne les valeurs des paramètres calculés à partir de C_S et de la variance ($\frac{\lambda}{\alpha^2} = 1$).

C_S	.25	.5	$1/\sqrt{2}$	1	$\sqrt{2}$	2	3	4
λ	64	16	8	4	2	1	4/9	1/4
α	8	4	$\sqrt{8}$	2	$\sqrt{2}$	1	2/3	1/2

TABEAU 2.1 : Valeurs de C_S , λ et α .

Par chaque méthode de simulation, Wilson-Hilferty tronquée, Johnk, interpolation de la table, on génère 100,000 variates Gamma pour chacune des 8 valeurs de C_S .

Ces 100,000 variates sont divisées en $p = 1000$ échantillons de taille 100. Sur chacun des échantillons de taille 100, on considère d'autres échantillons de taille 25, 50 et 75 formés à partir de 25, 50 ou 75 premiers éléments de l'échantillon de taille 100.

On dispose donc de $p = 1,000$ échantillons de tailles, 25, 50, 75 et 100 pour 8 valeurs de C_S correspondant à 8 couples (α_0, λ_0) . Sur chacun des échantillons, on ajuste la loi Gamma par le maximum de vraisemblance si le coefficient d'asymétrie de l'échantillon est positif. Le nombre d'échantillons pour lesquels le coefficient d'asymétrie est positif est noté p' .

On est alors en mesure de calculer diverses statistiques sur les échantillons des valeurs des paramètres ajustés. Ces statistiques sont les moyennes $\bar{\alpha}$ et $\bar{\lambda}$, les écarts-types s_α et s_λ , les déviations relatives moyennes $\frac{\bar{\alpha} - \alpha_0}{\alpha_0}$ et $\frac{\bar{\lambda} - \lambda_0}{\lambda_0}$ ainsi que les quantités

$\frac{\bar{\alpha} - \alpha_0}{s_\alpha/\sqrt{p'}}$ et $\frac{\bar{\lambda} - \lambda_0}{s_\lambda/\sqrt{p'}}$ qui suivent une loi de Student, les statistiques

sont calculées pour chaque taille N et chaque couple (α_0, λ_0) . Les résultats apparaissent aux tableaux 2.2 à 2.5.

METHODE DE SIMULATION: WILSON-HILFERTY

PARAMETRES THEORIQUES: ALPHA= 8,000 LAMBDA= 64,000

TAILLE	ALPHA				LAMBDA				CS<0
	MOY.	E.T.	DIF.REL.	T	MOY.	E.T.	DIF.REL.	T	
25	8,9753	2,6876	.1219	9,3435	71,7876	21,5203	.1217	9,3143	337
50	8,3574	1,6867	.0447	5,7257	66,8098	13,4292	.0439	5,6532	270
75	8,2498	1,3904	.0312	5,0567	65,9490	11,1060	.0305	4,9388	208
100	8,1511	1,1737	.0189	3,7338	65,1759	9,3831	.0184	3,6344	159

PARAMETRES THEORIQUES: ALPHA= 4,000 LAMBDA= 16,000

TAILLE	ALPHA				LAMBDA				CS<0
	MOY.	E.T.	DIF.REL.	T	MOY.	E.T.	DIF.REL.	T	
25	4,4806	1,3764	.1202	9,9621	17,8583	5,3911	.1161	9,8347	186
50	4,2092	.8666	.0523	7,3096	16,8317	3,4166	.0520	7,3717	83
75	4,1412	.6763	.0353	6,4556	16,5632	2,6652	.0352	6,5334	44
100	4,1072	.5727	.0268	5,8512	16,4192	2,2427	.0262	5,8460	22

PARAMETRES THEORIQUES: ALPHA= 2,828 LAMBDA= 8,000

TAILLE	ALPHA				LAMBDA				CS<0
	MOY.	E.T.	DIF.REL.	T	MOY.	E.T.	DIF.REL.	T	
25	3,1663	.9722	.1195	10,4672	8,9143	2,6517	.1143	10,3838	93
50	3,0063	.6612	.0629	8,3960	8,4847	1,7790	.0606	8,5025	26
75	2,9288	.4987	.0355	6,3382	8,2736	1,3582	.0342	6,3436	8
100	2,9081	.4361	.0282	5,7771	8,2133	1,1843	.0267	5,6929	1

PARAMETRES THEORIQUES: ALPHA= 2,000 LAMBDA= 4,000

TAILLE	ALPHA				LAMBDA				CS<0
	MOY.	E.T.	DIF.REL.	T	MOY.	E.T.	DIF.REL.	T	
25	2,2737	.7289	.1369	11,6415	4,4803	1,3589	.1201	10,9566	39
50	2,1228	.4579	.0614	8,4753	4,2068	.8446	.0517	7,7345	2
75	2,0809	.3587	.0405	7,1262	4,1414	.6684	.0353	6,6814	2
100	2,0578	.3086	.0289	5,9165	4,1029	.5768	.0257	5,6375	1

TABLEAU 2.2 : Ajustements de la loi Gamma sur les échantillons générés par la méthode de Wilson-Hilferty tronquée.

METHODE DE SIMULATION; WILSON-HILFERTY

PARAMETRES THEORIQUES; ALPHA= 1,414 LAMBDA= 2,000

TAILLE	ALPHA				LAMBDA				CS<0
	MOY.	E.T.	DIF.REL.	T	MOY.	E.T.	DIF.REL.	T	
25	1,6052	.5238	.1351	11,4515	2,2364	.6711	.1182	11,0601	14
50	1,4895	.3342	.0532	7,1245	2,0832	.4256	.0416	6,1854	0
75	1,4575	.2655	.0306	5,1566	2,0458	.3367	.0229	4,3036	0
100	1,4320	.2257	.0126	2,4969	2,0169	.2879	.0085	1,8599	0

PARAMETRES THEORIQUES; ALPHA= 1,000 LAMBDA= 1,000

TAILLE	ALPHA				LAMBDA				CS<0
	MOY.	E.T.	DIF.REL.	T	MOY.	E.T.	DIF.REL.	T	
25	1,0640	.3872	.0640	5,2271	1,0285	.3231	.0285	2,7872	1
50	.9957	.2523	-.0043	-.5399	.9786	.2150	-.0214	-3,1467	0
75	.9728	.2054	-.0272	-4,1846	.9606	.1734	-.0394	-7,1897	0
100	.9631	.1738	-.0369	-6,7184	.9525	.1481	-.0475	-10,1340	0

PARAMETRES THEORIQUES; ALPHA= .667 LAMBDA= .444

TAILLE	ALPHA				LAMBDA				CS<0
	MOY.	E.T.	DIF.REL.	T	MOY.	E.T.	DIF.REL.	T	
25	.7466	.3093	.1199	8,1736	.4490	.1098	.0102	1,3101	0
50	.6782	.1815	.0174	2,0157	.4279	.0714	-.0372	-7,3301	0
75	.6564	.1442	-.0153	-2,2423	.4190	.0569	-.0572	-14,1190	0
100	.6526	.1251	-.0211	-3,5568	.4171	.0486	-.0616	-17,7958	0

PARAMETRES THEORIQUES; ALPHA= .500 LAMBDA= .250

TAILLE	ALPHA				LAMBDA				CS<0
	MOY.	E.T.	DIF.REL.	T	MOY.	E.T.	DIF.REL.	T	
25	.7451	.3799	.4901	20,3998	.2973	.0576	.1893	25,9940	0
50	.6595	.2271	.3190	22,2131	.2877	.0386	.1508	30,8979	0
75	.6313	.1721	.2626	24,1250	.2849	.0305	.1394	36,1122	0
100	.6171	.1453	.2342	25,4837	.2825	.0254	.1301	40,4150	0

TABLEAU 2.2 (suite) : Ajustements de la loi Gamma sur les échantillons générés par la méthode de Wilson-Hilferty tronquée.

METHODE DE SIMULATION: JOHNK

PARAMETRES THEORIQUES: ALPHA= 8,000 LAMBDA= 64,000

TAILLE	ALPHA				LAMBDA				
	MOY.	E.T.	DIF.REL.	T	MOY.	E.T.	DIF.REL.	T	CS<0
25	8,8931	2,6675	.1116	8,5488	71,1084	21,1987	.1111	8,5622	348
50	8,3396	1,6838	.0425	5,4462	66,7029	13,3651	.0422	5,4603	271
75	8,2541	1,4100	.0318	5,1201	66,0073	11,2248	.0314	5,0801	193
100	8,1396	1,1663	.0174	3,4475	65,1102	9,2856	.0173	3,4445	170

PARAMETRES THEORIQUES: ALPHA= 4,000 LAMBDA= 16,000

TAILLE	ALPHA				LAMBDA				
	MOY.	E.T.	DIF.REL.	T	MOY.	E.T.	DIF.REL.	T	CS<0
25	4,4175	1,3757	.1044	8,5992	17,6479	5,4429	.1030	8,5792	197
50	4,2087	.9225	.0522	6,8183	16,8310	3,6479	.0519	6,8648	92
75	4,1318	.7173	.0330	5,6437	16,5349	2,8273	.0334	5,8094	57
100	4,1146	.6185	.0287	5,7843	16,4612	2,4317	.0288	5,9191	26

PARAMETRES THEORIQUES: ALPHA= 2,828 LAMBDA= 8,000

TAILLE	ALPHA				LAMBDA				
	MOY.	E.T.	DIF.REL.	T	MOY.	E.T.	DIF.REL.	T	CS<0
25	3,1686	.9851	.1203	10,3710	8,8781	2,6551	.1098	9,9331	98
50	3,0077	.6388	.0634	8,7349	8,4673	1,7459	.0584	8,3313	31
75	2,9629	.5292	.0476	8,0180	8,3542	1,4478	.0443	7,7167	5
100	2,9318	.4455	.0365	7,3375	8,2736	1,2306	.0342	7,0306	0

PARAMETRES THEORIQUES: ALPHA= 2,000 LAMBDA= 4,000

TAILLE	ALPHA				LAMBDA				
	MOY.	E.T.	DIF.REL.	T	MOY.	E.T.	DIF.REL.	T	CS<0
25	2,2516	.7354	.1258	10,5632	4,4416	1,3473	.1104	10,1184	47
50	2,1259	.4766	.0629	8,3350	4,2162	.8967	.0541	7,6102	4
75	2,0810	.3747	.0405	6,8381	4,1379	.6963	.0345	6,2608	0
100	2,0620	.3186	.0310	6,1532	4,1042	.5935	.0260	5,5520	0

TABLEAU 2.3 : Ajustements de la loi Gamma sur les échantillons générés par la méthode de Johnk.

METHODE DE SIMULATION: JOHNK

PARAMETRES THEORIQUES: ALPHA= 1,414 LAMBDA= 2,000

TAILLE	ALPHA				LAMBDA				CS<0
	MOY.	E.T.	DIF.REL.	T	MOY.	E.T.	DIF.REL.	T	
25	1,6298	.5817	.1525	11,6438	2,2578	.6761	.1289	11,9788	13
50	1,5199	.3524	.0747	9,4828	2,1255	.4246	.0628	9,3493	0
75	1,4811	.2665	.0473	7,9333	2,0750	.3259	.0375	7,2790	0
100	1,4610	.2200	.0331	6,7229	2,0525	.2699	.0262	6,1499	0

PARAMETRES THEORIQUES: ALPHA= 1,000 LAMBDA= 1,000

TAILLE	ALPHA				LAMBDA				CS<0
	MOY.	E.T.	DIF.REL.	T	MOY.	E.T.	DIF.REL.	T	
25	1,1556	.3982	.1556	12,3580	1,1081	.3169	.1081	10,7912	0
50	1,0835	.2566	.0835	10,2883	1,0552	.1967	.0552	8,8798	0
75	1,0550	.2127	.0550	8,1710	1,0358	.1598	.0358	7,0847	0
100	1,0403	.1734	.0403	7,3535	1,0254	.1348	.0254	5,9575	0

PARAMETRES THEORIQUES: ALPHA= .667 LAMBDA= .444

TAILLE	ALPHA				LAMBDA				CS<0
	MOY.	E.T.	DIF.REL.	T	MOY.	E.T.	DIF.REL.	T	
25	.7987	.3418	.1981	12,2202	.4835	.1202	.0879	10,2713	0
50	.7324	.2115	.0987	9,8330	.4619	.0752	.0393	7,3481	0
75	.7076	.1670	.0614	7,7525	.4549	.0621	.0235	5,3188	0
100	.6964	.1378	.0445	6,8133	.4514	.0524	.0155	4,1701	0

PARAMETRES THEORIQUES: ALPHA= .500 LAMBDA= .250

TAILLE	ALPHA				LAMBDA				CS<0
	MOY.	E.T.	DIF.REL.	T	MOY.	E.T.	DIF.REL.	T	
25	.6445	.3450	.2890	13,2474	.2722	.0645	.0890	10,8990	0
50	.5644	.1964	.1289	10,3764	.2609	.0420	.0436	8,2038	0
75	.5434	.1493	.0868	9,1937	.2570	.0333	.0281	6,6713	0
100	.5338	.1274	.0676	8,3914	.2556	.0287	.0225	6,1778	0

TABLEAU 2.3 (suite) : Ajustements de la loi Gamma sur les échantillons générés par la méthode de Johnk.

METHODE DE SIMULATION: INTERPOLATION DE LA TABLE

PARAMETRES THEORIQUES: ALPHA= 8,000 LAMBDA= 64,000

TAILLE	ALPHA				LAMBDA				
	MOY.	E.T.	DIF.REL.	T	MOY.	E.T.	DIF.REL.	T	CS<0
25	8,8500	2,5944	.1063	8,5313	70,7566	20,7814	.1056	8,4659	322
50	8,3498	1,6772	.0437	5,7730	66,7838	13,4612	.0435	5,7236	234
75	8,2062	1,3257	.0258	4,3905	65,6670	10,5635	.0260	4,4551	203
100	8,1449	1,1376	.0181	3,6786	65,1804	9,1014	.0184	3,7456	166

PARAMETRES THEORIQUES: ALPHA= 4,000 LAMBDA= 16,000

TAILLE	ALPHA				LAMBDA				
	MOY.	E.T.	DIF.REL.	T	MOY.	E.T.	DIF.REL.	T	CS<0
25	4,5100	1,4534	.1275	10,0607	18,0339	5,7066	.1271	10,2185	178
50	4,2107	.8433	.0527	7,5109	16,8491	3,3440	.0531	7,6346	96
75	4,1046	.6314	.0261	5,1261	16,4151	2,4680	.0259	5,2054	42
100	4,0847	.5592	.0212	4,7582	16,3371	2,2040	.0211	4,8059	13

PARAMETRES THEORIQUES: ALPHA= 2,828 LAMBDA= 8,000

TAILLE	ALPHA				LAMBDA				
	MOY.	E.T.	DIF.REL.	T	MOY.	E.T.	DIF.REL.	T	CS<0
25	3,1415	.9230	.1107	10,1940	8,8180	2,5070	.1022	9,8051	97
50	3,0009	.6320	.0610	8,4926	8,4700	1,7245	.0587	8,4790	32
75	2,9325	.5048	.0368	6,4866	8,2863	1,3838	.0358	6,5073	11
100	2,9045	.4248	.0269	5,6465	8,2080	1,1682	.0260	5,6156	5

PARAMETRES THEORIQUES: ALPHA= 2,000 LAMBDA= 4,000

TAILLE	ALPHA				LAMBDA				
	MOY.	E.T.	DIF.REL.	T	MOY.	E.T.	DIF.REL.	T	CS<0
25	2,2384	.6732	.1192	10,9162	4,4323	1,2988	.1081	10,2582	50
50	2,1287	.4604	.0643	8,8035	4,2222	.8800	.0556	7,9534	8
75	2,0903	.3605	.0452	7,9222	4,1510	.6852	.0377	6,9677	0
100	2,0679	.3068	.0339	6,9970	4,1162	.5763	.0291	6,3770	0

TABLEAU 2.4 : Ajustements de la loi Gamma sur les échantillons générés par interpolation de la table de Harter.

METHODE DE SIMULATION: INTERPOLATION DE LA TABLE

PARAMETRES THEORIQUES: ALPHA= 1.414 LAMBDA= 2.000

TAILLE	ALPHA				LAMBDA				CS<0
	MOY.	E.T.	DIF.REL.	T	MOY.	E.T.	DIF.REL.	T	
25	1.6125	.5265	.1402	11.8551	2.2250	.6543	.1125	10.8230	9
50	1.5047	.3376	.0640	8.4693	2.0956	.4123	.0478	7.3320	1
75	1.4744	.2657	.0426	7.1684	2.0632	.3249	.0316	6.1539	0
100	1.4562	.2274	.0297	5.8449	2.0431	.2787	.0216	4.8935	0

PARAMETRES THEORIQUES: ALPHA= 1.000 LAMBDA= 1.000

TAILLE	ALPHA				LAMBDA				CS<0
	MOY.	E.T.	DIF.REL.	T	MOY.	E.T.	DIF.REL.	T	
25	1.1611	.4129	.1611	12.3319	1.1134	.3043	.1134	11.7766	1
50	1.0612	.2532	.0612	7.6425	1.0451	.1957	.0451	7.2875	0
75	1.0360	.2000	.0360	5.6986	1.0273	.1513	.0273	5.7120	0
100	1.0204	.1699	.0204	3.7915	1.0184	.1312	.0184	4.4236	0

PARAMETRES THEORIQUES: ALPHA= .667 LAMBDA= .444

TAILLE	ALPHA				LAMBDA				CS<0
	MOY.	E.T.	DIF.REL.	T	MOY.	E.T.	DIF.REL.	T	
25	.8029	.3272	.2043	13.1660	.4923	.1242	.1076	12.1777	0
50	.7308	.1976	.0963	10.2700	.4686	.0749	.0544	10.2185	0
75	.7033	.1516	.0549	7.6358	.4599	.0587	.0348	8.3390	0
100	.6978	.1310	.0467	7.5149	.4569	.0514	.0281	7.6688	0

PARAMETRES THEORIQUES: ALPHA= .500 LAMBDA= .250

TAILLE	ALPHA				LAMBDA				CS<0
	MOY.	E.T.	DIF.REL.	T	MOY.	E.T.	DIF.REL.	T	
25	.6012	.3386	.2023	9.4484	.2563	.0726	.0253	2.7546	0
50	.5148	.1865	.0295	2.5027	.2421	.0473	-.0316	-5.2754	0
75	.4956	.1429	-.0088	-.9777	.2376	.0373	-.0495	-10.5070	0
100	.4882	.1228	-.0237	-3.0483	.2360	.0320	-.0561	-13.8742	0

TABLEAU 2.4 (suite) : Ajustements de la loi Gamma sur les échantillons générés par interpolation de la table de Harter.

METHODE DE SIMULATION: APPROXIMATION POLYNOMIALE

PARAMETRES THEORIQUES: ALPHA= 8,000 LAMBDA= 64,000

TAILLE	ALPHA				LAMBDA				CS<0
	MOY,	E,T,	DIF,REL,	T	MOY,	E,T,	DIF,REL,	T	
25	8,9238	2,6156	,1155	9,2033	71,3457	20,9502	,1148	9,1365	321
50	8,4171	1,6847	,0521	6,8526	67,3182	13,5270	,0518	6,7892	234
75	8,2809	1,3404	,0351	5,9157	66,2618	10,6892	,0353	5,9736	203
100	8,2159	1,1453	,0270	5,4517	65,7458	9,1617	,0273	5,5096	164

TABLEAU 2.5 : Ajustements de la loi Gamma sur les échantillons générés par approximation polynomiale de la table de Harter.

2.2 Présentation des résultats

Un premier aperçu aux tableaux 2.2, 2.3 et 2.4 révèle une assez forte similitude de comportement entre les trois méthodes de simulations.

D'abord, pour les trois méthodes de simulations, les paramètres α et λ sont pratiquement toujours surestimés. Les déviations relatives moyennes sont sensiblement les mêmes, pour une taille donnée, indépendamment de la valeur de λ_0 sauf pour les valeurs de λ_0 inférieures à 1 où la transformation de Wilson-Hilferty tronquée performe moins bien que les autres méthodes. Ces mêmes déviations diminuent lorsque la taille augmente passant d'environ 10% pour $N = 25$ à 2% pour $N = 100$. On remarque également que les déviations relatives moyennes sont légèrement supérieures pour α .

Les statistiques de Student, calculées pour α et λ sont nettement plus élevées que la valeur critique 2.6 correspondant à un niveau de signification 1% pour un test sur les moyennes des paramètres estimés. Cette remarque vaut pour les trois méthodes de simulation et quelque soit λ_0 ; cependant, les valeurs calculées se rapprochent de la valeur critique lorsque N augmente. On peut donc en conclure que la méthode du maximum de vraisemblance qui, théoriquement, est optimale, dans le cas de la loi Gamma, conduit généralement à une surestimation des paramètres α et λ qui est très significative et ce, indépendamment de la méthode de simulation utilisée.

Le nombre d'échantillons avec un coefficient d'asymétrie négatif est à peu près le même pour les trois méthodes de simulations et diminue considérablement lorsque λ_0 diminue, c'est-à-dire lorsque C_{S_0} augmente.

2.3 Temps de simulation

C'est le temps d'occupation de l'ordinateur qui détermine le coût d'utilisation d'une méthode de simulation et pour des plans d'échantillonnage assez élaborés, le coût d'utilisation devient souvent une contrainte.

Pour trois des quatre méthodes de simulations étudiées, le temps de simulation est indépendant de la valeur des paramètres initiaux. Pour la méthode de Johnk, le temps de simulation augmente avec la valeur de λ .

Les temps de simulations de 800,000 variates (100,000 pour chacune des 8 valeurs de λ_0) sont: transformation de Wilson-Hilferty tronquée, 1,044 secondes, technique de Johnk, 1,700 secondes et méthode d'interpolation de la table de Harter, 757 secondes. La méthode d'approximation polynomiale met 666 secondes à générer les 100,000 variates pour $\lambda_0 = 64$.

2.4 Conclusion

L'examen des tableaux 2.2 à 2.4 montre que les méthodes de simulations étudiées performant également quant à la qualité de la simulation qui est très acceptable pour la gamme d'asymétrie étudiée.

La méthode d'interpolation de la table de Harter est à retenir pour les gros plans de simulation entraînant des coûts d'exploitation élevés. Par contre, lors de simulations où les coûts ne sont pas importants, la transformation de Wilson-Hilferty tronquée est préférable pour sa simplicité d'utilisation.

La méthode de Johnk conduit également à de bons échantillons mais son temps de génération est un handicap surtout aux valeurs élevées de λ .

REFERENCES BIBLIOGRAPHIQUES

ABRAMOWITZ, M. et I.A. STEGUN (1964).

Handbook of mathematical functions. Dover Publications Inc., New York, 9^{ème} édition, 1045 pages.

BOBEE, B. et P. BOUCHER (1979).

Comparaison des algorithmes de génération de la distribution Gamma et de ses formes dérivées. INRS-Eau, rapport scientifique No 112, 78 pages.

HARTER, H. LEON (1969).

A new table of percentage points of the Pearson type III distribution. Technometrics, vol. 11, No 1, février 1969.

APPENDICEMode d'utilisation des simulateurs -Genere 1 et Genere 2

Le simulateur Genere 1 calcule une variate de Pearson 3 par la méthode de la table de Harter (section 1.1). L'évaluation de la variate de Pearson type 3 par approximation polynomiale (section 1.2) se fait à l'aide du simulateur Genere 2.

L'utilisateur doit fournir certaines informations, soit:

- 1) un programme principal ou une sous-routine qui introduit les cinq paramètres suivants:

A: α
B: λ
C: m (paramètre d'origine de la loi Pearson 3)
N: Nombre de probabilités à simuler
INIT: Valeur initiale nécessaire à la simulation des probabilités

- 2) un fichier contenant des données de base et lu par les simulateurs sur l'unité 10:

Genere 1: fichier PROB contenant la table de Harter.

Genere 2: fichier PBCOEF contenant les valeurs des coefficients $L_{i,j}$ déjà défini à la section 1.2.

Suite à un appel à l'un ou l'autre des simulateurs, l'utilisateur obtient N variates Pearson type 3 non centrées réduites, X , au coefficient d'asymétrie $C_s = 2(\lambda^{**} .5)$ et pour N probabilités P déterminées au hasard.

Certains énoncés sont essentiels à l'intérieur du programme ou de la sous-routine fournie par l'utilisateur:

PROGRAM nom (output, <u>Tape 10</u>)	L'unité 10 est réservée à la lecture, à l'intérieur des simulateurs, des fichiers PBCOEF ou PROB.
DIMENSION P(), X()	Afin de réserver suffisamment d'espace de mémoire, insérer un nombre égal ou supérieur à N .
COMMON/TRANS/A, B, C, N, INIT, G1	Assure le transfert des paramètres d'entrée vers les simulateurs.
CALL Prepare (P, X)	Exécution du simulateur

Voici un exemple d'utilisation des simulateurs sur CDC.

1

↓

Get, Genere 1, tape 10 = PROB. ou Get, Genere 2, tape 10 = PBCOEF

FTN

Load, LG0

Satisfy, Genere 1 ou Satisfy, Genere 2

Execute

7/8/9

7

↓

Program princ (input, output, tape 10)

Dimension P(50), X(50)

Common/TRANS/A, B, C, N, INIT, G1

N = 50

Read *, A, B, C, INIT

Call prepare (P, X)

Print *, P, X

STOP

End

7/8/9

2 4 6 2134

6/7/8/9

Dans cet exemple, 50 probabilités sont simulées et 50 variates sont générées. Les probabilités (P) et les variates de Pearson 3 (K) sont ensuite imprimées.

La liste des sous-routines des deux simulations est fournie dans les 2 pages suivantes.

GENERE 1

```

SUBROUTINE PREPARE (P,VAL)
  DIMENSION P(1),VAL(1)
  COMMON /TRANS/A,B,C,N,INIT,G1
  DATA 11/2147483647/,X1/16614000000000000000B/
  DO 1 I=1,N
10 INIT=MOD(16807*INIT,11)
  P(I)=FLOAT(INIT)*X1
  IF(P(I).LT..0001.OR.P(I).GT..9999)GO TO 10
  1 CONTINUE
  G1=2./B**.5
  U=C+B/A
  S=B**.5/ABS(A)
  CALL GERE (P,VAL)
  DO 30 I=1,N
  VAL(I)=U+VAL(I)*S
30 CONTINUE
  RETURN
  END

```

```

SUBROUTINE GERE (P,VAL)
  DIMENSION TAB(27,41),P(1),VAL(1)
  COMMON /DONNE/TAB,CS/TRANS/A,B,C,N,INIT,G1
  DATA IPASS/0/
  C*****LECTURE DES VALEURS DE LA TABLE DE HARTER
  IF(IPASS.GT.0)GO TO 1
  READ(10,*)((TAB(I,J),I=1,27),J=1,41)
  IPASS=1
  1 CALL INDICEG (L,G1)
  CS=(L-1)/10.
  DO 9 IJ=1,N
  CALL INDICEP (J,P(IJ))
  C*****CALCUL DE LA VALEUR AU POINT (P,G1)=TAB(J,L)
  CALL CALCUL(J,L,P(IJ),G1,VAL(IJ))
  9 CONTINUE
  99 RETURN
  END

```

```

SUBROUTINE INDICEP (J,P)
  DIMENSION Z(27)
  COMMON /PROB/Z
  DATA Z/.0001,.0005,.001,.005,.01,.02,.025,.04,.05,.1,.2,.3,.4,
  .5,.6,.7,.8,.9,.95,.96,.975,.98,.99,.995,.999,.9995,.9999/
  C*****LA PROBABILITE EST EXPRIMEE EN TERME DE J VARIANT DE 2 A 26.
  JJ=0
  IF(P.GT..5)JJ=1
  KK=1+13*JJ
  P1=P+.000001-.000002*JJ
  DO 10 K=KK,27
  J=K
  IF(P1-Z(K).LE.0.)GO TO 12
10 CONTINUE
12 J=J-JJ
  RETURN
  END

```

GENERE 1 (suite)

```

SUBROUTINE INDICEG (L,G1)
C*****LA VALEUR DE G1 EST EXPRIMEE EN TERME DE L VARIANT DE 2 A 40.
JJ=0
IF(G1.GT.2.)JJ=1
L=G1*10.+2.-2.*JJ
RETURN
END

SUBROUTINE CALCUL(J,L,P,G1,VAL)
DIMENSION TAB(27,41),VALI(3)
DIMENSION Z(27)
COMMON /PROB/Z/DONNE/TAB,CS
DATA Z/.0001,.0005,.001,.005,.01,.02,.025,.04,.05,.1,.2,.3,.4,
..5,.6,.7,.8,.9,.95,.96,.975,.98,.99,.995,.999,.9995,.9999/
T2(Y1,Y2,Y3,X1,X2,X3)=((Y3-Y2)*(X2**2.-X1**2.)-(Y2-Y1)*
1(X3**2.-X2**2.))/(X3-X2)*(X2-X1)*(X1-X3)
T1(Y1,Y2,Y3,X1,X2,X3)=((Y2-Y1)-T2(Y1,Y2,Y3,X1,X2,X3)*
1(X2-X1))/(X2**2.-X1**2.)
T3(Y1,Y2,Y3,X1,X2,X3)=Y2-T1(Y1,Y2,Y3,X1,X2,X3)*X2**2.-T2(Y1,
1Y2,Y3,X1,X2,X3)*X2
DO 60 I=1,3
JJ=I-2
VALI(I)=T1(TAB(J+JJ,L-1),TAB(J+JJ,L),TAB(J+JJ,L+1),
1CS-.1,CS,CS+.1)*G1**2.+T2(TAB(J+JJ,L-1),TAB(J+JJ,L),TAB
2(J+JJ,L+1),CS-.1,CS,CS+.1)*G1+T3(TAB(J+JJ,L-1),TAB
3(J+JJ,L),TAB(J+JJ,L+1),CS-.1,CS,CS+.1)
60 CONTINUE
VAL=T1(VALI(1),VALI(2),VALI(3),Z(J-1),Z(J),Z(J+1))*P**2.+T2
1(VALI(1),VALI(2),VALI(3),Z(J-1),Z(J),Z(J+1))*P+T3(VALI(1),
2VALI(2),VALI(3),Z(J-1),Z(J),Z(J+1))
RETURN
END

```

GENERE 2

```

SUBROUTINE PREPARE (P,VAL)
DIMENSION P(1),VAL(1)
COMMON /TRANS/A,B,C,N,INIT,G1
DATA I1/2147483647/,X1/16614000000000000000B/
DO 1 I=1,N
10 INIT=MOD(16807*INIT,I1)
P(I)=FLOAT(INIT)*X1
IF(P(I).LT..0001.OR.P(I).GT..9999)GO TO 10
1 CONTINUE
G1=2./B**.5
CALL GERE (P,VAL)
U=C+B/A
S=B**.5/ABS(A)
DO 30 I=1,N
VAL(I)=U+VAL(I)*S
30 CONTINUE
RETURN
END

```

```

SUBROUTINE GERE (P,VAL)
DIMENSION P(1),VAL(1),COEF(4,8,8),AC(8)
COMMON /TRANS/A,B,C,N,INIT,G1/CAL/COEF,AC
DATA IPASS/0/
C*****LECTURE DES COEFFICIENTS (COEF)
IF(IPASS.GT.0)GO TO 1
READ(10,*)((COEF(IN,JA,I),I=1,8),IN=1,4),JA=1,8)
IPASS=1
1 DO 9 IJ=1,N
CALL SECTION (P(IJ))
CALL VALEUR (P(IJ),G1,VAL(IJ))
9 CONTINUE
RETURN
END

```

```

SUBROUTINE SECTION (P)
COMMON /INTER/IN,PP
PP=P
IF(PP.GT..5)PP=1.-PP
PP=-(ALOG10(PP))
IN=1
IF((P.LT..5).AND.(P.GE..04))IN=2
IF((P.GE..5).AND.(P.LT..95))IN=3
IF(P.GE..95)IN=4
RETURN
END

```

```

SUBROUTINE VALEUR (P,G,VAL)
DIMENSION COEF(4,8,8),AC(8)
COMMON /CAL/COEF,AC/INTER/IN,PP
C*****CALCUL DES 8 COEFFICIENTS POUR LA PROBABILITE P
DO 30 JA=1,8
AC(JA)=0.
DO 32 I=1,8
32 AC(JA)=AC(JA)+COEF(IN,JA,I)*PP**(I-1)
IF(G.LT..00001) GO TO 33
30 CONTINUE
IF(P.GE..5.AND.P.LT..55)AC(1)=AC(1)-.0001433298
VAL=0.
DO 40 JA=1,8
C*****SIMULATION DE LA VALEUR AU POINT (P,G1)
VAL=VAL+AC(JA)*G**(JA-1)
40 CONTINUE
GO TO 9
33 CONTINUE
IF(P.GE..5.AND.P.LT..55)AC(1)=AC(1)-.0001433298
VAL=AC(1)
9 RETURN
END

```