Aktuárské vědy

E. J. Gumbel

La plus grande valeur. III

Aktuárské vědy, Vol. 5 (1935), No. 4, 145–160

Persistent URL: http://dml.cz/dmlcz/144636

Terms of use:

Institute of Mathematics of the Czech Academy of Sciences provides access to digitized documents strictly for personal use. Each copy of any part of this document must contain these *Terms of use*.

This document has been digitized, optimized for electronic delivery and stamped with digital signature within the project *DML-CZ: The Czech Digital Mathematics Library* http://dml.cz

La plus grande valeur.

E. J. Gumbel, Université de Lyon.

(Fin.)

Le tableau suivant I contient pour des valeurs equidistantes de la variable y (colonne 1) la distribution finale dans la forme $\mathfrak{w}(y)/\alpha$ (colonne 3) et ses différences (colonne 4). Pour des grandes valeurs négatives de y la distribution tend vers zéro comme $e^{-e^{-y}}$, pour des grandes valeurs positives comme e^{-y}. Donc du côté gauche elle tend plus vite vers zéro que du côté droit, et elle sera pratiquement contenue dans des limites assez étroites. La deuxième colonne du tableau I contient la valeur (14)

$$\mathfrak{W}(y) = \int_{-\infty}^{y} \mathfrak{w}(z)\, dz = e^{-e^{-y}} \qquad (39)$$

c'est à dire la probabilité pour que la plus grande valeur soit inférieure à une grandeur y donnée. Nous avons employé pour ces calculs les valeurs numériques de la fonction exponentielle donnée par Gau[10]) et Coolidge[11]).

I. La distribution finale et sa somme.

y	$\mathfrak{W}(y)$	$\mathfrak{w}(y)/\alpha$	Δ
— 2,5	—	0,00006	
			0,00451
— 2,0	0,00062	0,00457	
			0,04614
— 1,5	0,01132	0,05071	
			0,12865
— 1,0	0,06599	0,17936	
			0,13769
— 0,5	0,19230	0,31705	
			0,05083
0,0	0,36788	0,36788	
			— 0,03717
0,5	0,54525	0,33071	
			— 0,07606

[10]) Emile Gau: Calculs numériques et graphiques. Collection Armand Colin. Paris 1925.

[11]) J. L. Coolidge: Einführung in die Wahrscheinlichkeitsrechnung. Deutsche Ausgabe von F. M. Urban, Leipzig-Berlin 1927.

y	$\mathfrak{W}(y)$	$\mathfrak{w}(y)/x$	Δ
1,0	0,69220	0,25465	— 0,07614
1,5	0,80003	0,17851	— 0,06030
2,0	0,87345	0,11821	— 0,04259
2,5	0,92120	0,07562	— 0,02731
3,0	0,95143	0,04831	— 0,01901
3,5	0,97025	0,02930	— 0,01132
4,0	0,98185	0,01798	— 0,00699
4,5	0,98895	0,01099	— 0,00429
5,0	0,99328	0,00670	— 0,00263
5,5	0,99592	0,00407	— 0,00160
6,0	0,99752	0,00247	— 0,00097
6,5	0,99850	0,00150	— 0,00059
7,0	0,99909	0,00091	— 0,00036
7,5	0,99945	0,00055	— 0,00022
8,0	0,99967	0,00033	— 0,00013
8,5	0,99980	0,00020	— 0,00007
9,0	0,99987	0,00013	— 0,00006
9,5	0,99993	0,00007	— 0,00003
10,0	0,99996	0,00004	

Pour des grandes valeurs positives de y, il suffit de calculer une des valeurs $\mathfrak{W}(y)$ ou $\mathfrak{w}(y)$. Car on aura d'après le procédé (6)

$$\frac{1 - \mathfrak{W}(y)}{\mathfrak{w}(y)} = \frac{1}{1 - e^{-y}}.$$

On peut donc poser approximativement pour ces valeurs

$$\mathfrak{W}(y) = 1 - \mathfrak{w}(y)$$

pourvu que

$$e^{y} >> 1$$

ce qui est certainement le cas pour $y \geqq 6$.

Le tableau I permet le calcul de l'écart probable ϱ. Il sera défini par la différence

$$x'' - x' = 2\varrho$$

où les valeurs x'' et x' correspondent à

$$\mathfrak{W}(x'') = \tfrac{3}{4},$$
$$\mathfrak{W}(x') = \tfrac{1}{4}.$$

Les équations (39)

$$e^{-e^{-y''}} = \tfrac{3}{4},$$
$$e^{-e^{-y'}} = \tfrac{1}{4}$$

mènent à

$$y'' = \quad 1{,}24590,$$
$$y' = -0{,}32663$$

dont on tire à l'aide de la transformation finale (15)

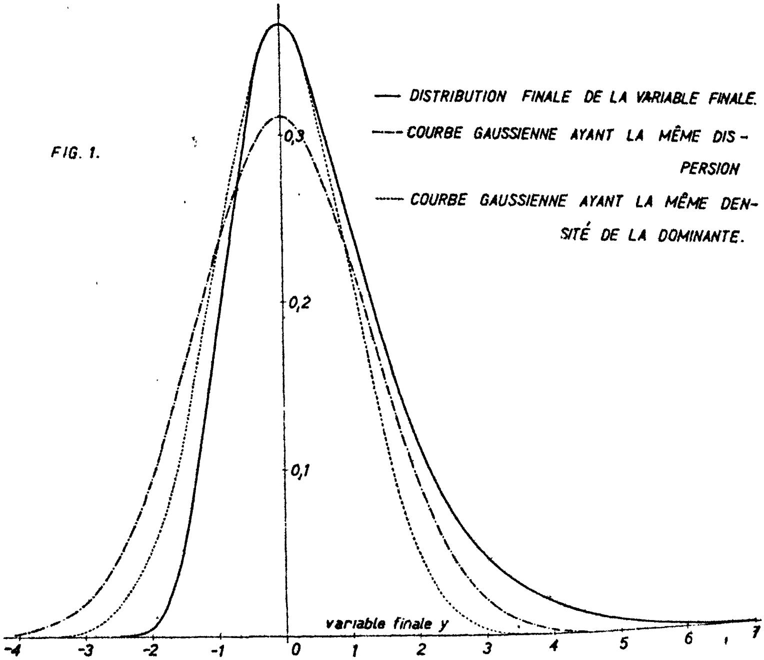
$$\alpha\varrho = 0{,}78626 \tag{38''}$$

La probabilité pour que la plus grande valeur soit contenue dans l'intervalle

$$u - \frac{0{,}32663}{\alpha} < x < u + \frac{1{,}24590}{\alpha}$$

est égale à $\frac{1}{2}$, de même la probabilité d'une valeur extérieure à cet intervalle.

FIG. 1.



La distribution finale $\mathfrak{w}(y)/\alpha$ est représentée dans la figure 1. Par comparaison cette figure montre aussi la distribution de Gauss ayant le même maximum. Donc l'écart type de cette courbe sera donné par

$$\sigma = 1{,}08444.$$

Sa forme analytique sera

$$\varphi_1(y) = \frac{1}{e}\, e^{-\frac{y^2}{2\sigma^2}}.$$

Pour le calcul de ces valeurs il suffit de multiplier les valeurs de l'exposant données par Davenport[12]) en fonction de y/σ par $1/e$.

[12]) Davenport, C. B.: Statistical Methods, London-New York, 1914.

En outre nous avons dessiné la courbe de Gauss ayant le même écart type que la distribution finale de la dernière valeur. Sa forme analytique

$$\varphi_2(y) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{y^2}{2\sigma^2}}$$

devient par (38)

$$\varphi_2(y) = \frac{1}{\pi}\sqrt{\frac{3}{\pi}}\, e^{-\frac{y^2}{2}\frac{6}{\pi^2}}.$$

Pour le calcul de ces valeurs on multiplie les abscisses dans le tableau de Davenport par $\frac{\pi}{\sqrt{6}}$ et les ordonnées correspondantes à y par $\frac{1}{\pi}\sqrt{\frac{3}{\pi}}$.

Le sommet de cette courbe ne sera que $\frac{1}{\pi}\sqrt{\frac{3}{\pi}} = 0{,}31105$, c'est à dire de 15,4% inférieure à la valeur exacte. La distribution finale est plus resserrée du côté gauche et plus ample du côté droit que les deux courbes de Gauss.

3. Application à quelques distributions initiales.

Appliquons nos résultats au calcul de la plus grande valeur et de l'écart type de la distribution initiale classique de Gauss

$$u(x) = \frac{h}{\sqrt{\pi}} e^{-t^2} \tag{40}$$

où

$$t = h\,(x - \overline{x}) \tag{41}$$

valeur distincte de la variable introduite dans les intégrales Gamma est la variable réduite. L'espérance mathématique de cette distribution est $\overline{x}$ et l'écart type $1/h\sqrt{2}$. La probabilité d'une valeur inférieure à x est

$$W(x) = \tfrac{1}{2}[1 + \Phi(t)] \tag{42}$$

où

$$\Phi(t) = \frac{2}{\sqrt{\pi}}\int_0^t e^{-t^2}\,\mathrm{d}t \tag{43}$$

est l'intégrale de Gauss. La condition (13) y est remplie. L'équation (6) qui sera

$$\frac{2h}{\sqrt{\pi}}\,\frac{e^{-t^2}}{1-\Phi(t)} = 2ht \tag{44}$$

est verifiée pour toute valeur de t suffisamment grande de la sorte que

$$2t^2 >> 1. \tag{45}$$

En effet on obtient par la règle de L'Hopital

$$\frac{e^{-t^2}:(t\sqrt{\pi})}{1-\Phi(t)}=\frac{-\dfrac{e^{-t^2}}{\sqrt{\pi}}\left(2+\dfrac{1}{t^2}\right)}{-\dfrac{2}{\sqrt{\pi}}e^{-t^2}}=1+\frac{1}{2t^2}.$$

Donc l'égalité (6) se transforme ici à la condition (45). Cette condition étant remplie, la dominante τ de la plus grande valeur calculée par (7) sera donnée d'après (42) par

$$\Phi(\tau)=1-\frac{2}{N}. \tag{46}$$

La dominante elle-même sera d'après (41)

$$u=\bar{x}+\frac{\tau}{h}. \tag{47}$$

Elle augmentera avec l'espérance mathématique et avec l'écart type initial. Donc une augmentation de la densité de probabilité de l'espérance mathématique de la variable initiale diminuera la dominante de la plus grande valeur.

Il s'agit encore de montrer comment les conditions finales sont réalisées. Mais pour les dérivées de la distribution de Gauss on a, comme c'est bien connu, les polynomes d'Hermite

$$\frac{\mathrm{d}^\nu w(x)}{\mathrm{d}x^\nu}=h^\nu w(x)\,H_\nu$$

où les H_ν sont des sommes de puissances de $(-2t)^{\nu-2\lambda}$; $\lambda=0,1\ldots$ Conséquemment à notre hypothèse centrale (45) c'est seulement le plus grand membre des polynomes d'Hermite qui nous interesse. Donc on peut poser

$$w^{(\nu)}(u)=w(u)\,(-2)^\nu\tau^\nu h^\nu.$$

Mais pour des grandes valeurs de t on aura d'après (44)

$$(-1)^\nu\left(\frac{w(u)}{1-W(u)}\right)^\nu=(-2)^\nu\,\tau^\nu h^\nu.$$

Donc l'égalité (13) est bien verifiée et il est légitime d'employer la forme finale de la distribution de la plus grande valeur (16) pour le calcul de l'espérance mathématique (25) et des moments (37).

La condition I est identique à un théorème employé dans le développement en série d'une distribution observée d'après les dérivées successives de la distribution de Gauss.

Pour la détermination de l'espérance mathématique de la plus grande valeur il faut traiter d'après (25) la grandeur α donnée par (5) pour la valeur dominante. On aura d'après (44)

$$x = 2h\tau \tag{44'}$$

ce qui remplit la condition II.

Donc l'espérance mathématique de la plus grande valeur de la variable réduite sera

$$\bar{t} = \tau + \frac{\gamma}{2\tau}. \tag{48}$$

On peut négliger le deuxième facteur puisque toute notre approximation repose sur la supposition (45), argument qui sera renforcé par la valeur numérique de la constante γ. On aura donc

$$\bar{t} = \tau. \tag{49}$$

Ce calcul est légitime pour des valeurs de N tel que (45) soit verifiée. Avec de la bonne volonté on peut admettre que cette condition est déjà remplie pour $t \geqq 3$.

Pour les calculs pratiques de $\bar{t}$ comme fonction de N on va d'abord changer le sens de l'égalité (46) et calculer, pour des valeurs de $t = \bar{t}$ choisies, le nombre N correspondant, d'après

$$\frac{N}{2} = \frac{1}{1 - \Phi(t)}. \tag{50}$$

Pour les nombres $3 \leqq t \leqq 3{,}46$, on a employé les valeurs à 7 décimales données par Markoff.[13] On obtient

$$\frac{N}{2} = \frac{\sqrt{\pi}}{2} : \int_t^\infty e^{-t^2}\, dt.$$

Pour l'intervalle $3{,}46 \leqq t \leqq 4$, on a employé les valeurs de Czuber[14] à 11 décimales. Comme la fonction se rapproche très vite de 1, on n'emploie pour le calcul qu'autant de décimales qu'il en existe après les 9 de cette fonction. Aussi la table de Czuber ne permet de poursuivre le calcul que jusqu'à $t = 4$, puisque à ce moment le nombre neuf existe déjà huit fois dans la fonction $\Phi(t)$. Pour de très grandes valeurs de t on emploie la valeur limite (44) qui conduit à

$$\frac{N}{2} = t e^{t^2} \sqrt{\pi}. \tag{51}$$

Du point de vue stricte c'est seulement l'emploi de cette formule qui est légitime. Donc on aurait dû calculer de nouveau toutes les valeurs

---

[13]) A. A. Markoff: Tables des valeurs de l'intégrale $\int_x^\infty e^{-t^2} dt$. St. Pétersbourg 1888.

[14]) E. Czuber: Wahrscheinlichkeitsrechnung. Leipzig-Berlin 1924.

de t en fonction de N ou se limiter aux valeurs $t \geq 4$. Mais du point de vue pratique il nous semblait légitime d'user les tables de l'intégrale de Gauss et d'employer pour cela la formule simple (50), quoique ces valeurs soient pour ainsi dire trop précises pour notre théorie. Les résultats sont portés dans le tableau II.

Pour un nombre donné N d'observations on fait une interpolation pour trouver les valeurs correspondantes de $\bar{t}$. La figure 2 montre la dominante τ de la plus grande valeur comme fonction du logarithme du nombre d'observations. Si N devient dixmille fois plus grand que sa valeur initiale, τ n'augmente que de 40%. Car pour $\frac{1}{2}N = 45268$ on aura $\tau = 3$, tandis que pour $\frac{1}{2}N = 45268 \,.\, 10^4$ on trouvera par interpolation logarithmique la valeur $\tau = 4{,}23$. Donc notre condition (45) se transforme à la question de savoir si le nombre d'observations est de l'ordre de grandeur

$$N \geq 10^5. \tag{45'}$$

II. La dominante de la plus grande valeur de la variable réduite pour la distribution initiale de Gauss.

τ	$\frac{1}{2}N \,.\, 10^{-6}$	Δ	τ	$\frac{1}{2}N \,.\, 10^{-6}$	Δ
3,00	0,045268		3,25	0,232408	15222
3,01	0,048218	2950	3,26	0,248745	16337
3,02	0,051371	3153	3,27	0,266281	17536
3,03	0,054739	3368	3,28	0,285110	18829
3,04	0,058340	3601	3,29	0,305327	20217
3,05	0,062190	3850	3,30	0,327043	21716
3,06	0,066306	4116	3,31	0,350369	23326
3,07	0,070708	4402	3,32	0,375431	25062
3,08	0,075417	4709	3,33	0,402364	26933
3,09	0,080456	5039	3,34	0,431313	28949
3,10	0,085847	5391	3,35	0,462432	31119
3,11	0,091617	5770	3,36	0,495894	33462
3,12	0,097793	6176	3,37	0,531879	35985
3,13	0,104406	6613	3,38	0,570585	38706
3,14	0,111488	7082	3,39	0,612225	41640
3,15	0,119073	7585	3,40	0,657034	44809
3,16	0,127198	8125	3,41	0,705258	47224
3,17	0,135904	8706	3,42	0,757161	51903
3,18	0,145234	9330	3,43	0,813045	55884
3,19	0,155234	10000	3,44	0,87322	6017
3,20	0,165954	10720	3,45	0,93804	6482
3,21	0,177449	11495	3,46	1,0079	699
3,22	0,189777	12328	3,47	1,0831	752
3,23	0,203000	13223	3,48	1,1642	811
3,24	0,217186	14186	3,49	1,2515	873

τ	$\frac{1}{2}N \cdot 10^{-6}$	Δ	τ	$\frac{1}{2}N \cdot 10^{-6}$	Δ
3,50	1,3457	942	3,81	14,05	106
3,51	1,4473	1016	3,82	15,20	115
3,52	1,5568	1095	3,83	16,45	125
3,53	1,6749	1181	3,84	17,80	135
3,54	1,8024	1265	3,85	19,28	148
3,55	1,9399	1375	3,86	20,87	159
3,56	2,0883	1484	3,87	22,60	173
3,57	2,2486	1603	3,88	24,48	188
3,58	2,4215	1729	3,89	26,53	205
3,59	2,6083	1868	3,90	28,74	221
3,60	2,8102	2018	3,91	31,15	241
3,61	3,0280	2179	3,92	33,77	262
3,62	3,2635	2355	3,93	36,50	283
3,63	3,5179	2544	3,94	39,71	311
3,64	3,7931	2752	3,95	43,07	336
3,65	4,0901	2970	3,96	46,73	366
3,66	4,4117	3216	3,97	50,71	398
3,67	4,7596	3479	3,98	55,04	433
3,68	5,1356	3760	3,99	59,77	473
3,69	5,5423	4067	4,0	64,85	508
3,70	5,9827	4404	4,1	149,5	846
3,71	6,4587	4760	4,2	350,6	2011
3,72	6,9750	5163	4,3	839,0	4884
3,73	7,5335	5585	4,4	2047	1208
3,74	8,1380	6045	4,5	5092	3045
3,75	8,7943	6563	4,6	1292 . 10	7830
3,76	9,5030	7087	4,7	3418 . 10	2126
3,77	10,27	77	4,8	8816 . 10	5398
3,78	11,10	83	4,9	$2372 \cdot 10^2$	1490
3,79	12,01	91	5,0	$6512 \cdot 10^2$	4140
3,80	12,99	98			

Du point de vue analytique la dépendance de la dominante du nombre d'observations est donnée par

$$\frac{\mathrm{d}N}{\mathrm{d}u} = h\frac{\mathrm{d}N}{\mathrm{d}t}$$

ce qui sera d'après (51) et la condition (44)

$$\frac{\mathrm{d}u}{\mathrm{d}N} = \frac{1}{2h\bar{t}N}. \tag{52}$$

La dominante et l'espérance mathématique de la plus grande valeur augmenteront donc d'une façon toujours plus lente avec le nombre d'observations.

La limite (51) permet de trouver une approximation asymptotique de la plus grande valeur. Pour des valeurs de t suffisamment grandes on peut en effet l'écrire sous la forme

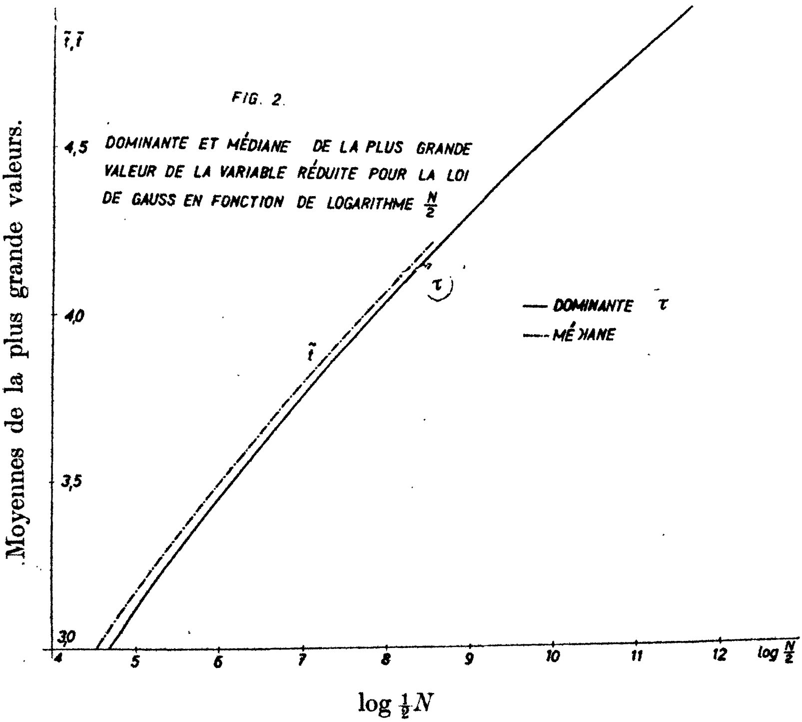
$$\lg N = \lg 2\sqrt{\pi} + \lg \bar{t} + \bar{t}^2.$$

On en déduit que t croît plus lentement que la racine du logarithme usuel de N, donc d'une manière extrêmement lente. Pour plus de simplicité on peut écrire en première approximation pour de très grandes valeurs de N

$$\begin{aligned} \bar{t} &= \sqrt{\lg N} \\ &= 1{,}51743 \sqrt{\log N} \end{aligned} \tag{53}$$

formule qui donne une valeur trop grande, mais à l'aide de laquelle on peut calculer simplement une seconde approximation.

Cette formule a déjà été employée par M. Ph. Frank.[15]) Elle a été prouvée de manières différentes par MM. von Mises,[3]) Dodd[7]) et



[15]) Ph. Frank: Über die Fortpflanzungsgeschwindigkeit. Physikalische Zeitschrift vol. 19, p. 516, 1918.

Tricomi.[4]) Le fait (48) que l'éspérance mathématique de la dernière valeur surpasse la dominante, relation qui dérive d'une facon générale de (25), a été prouvé pour la distribution de Gauss par M. J. Splawa Neyman.[16])

La figure (2) montre la dominante τ de la dernière valeur comme fonction du logarithme de $\frac{1}{2}N$. Elle s'applique à une distribution de Gauss ayant l'espérance mathématique zéro et un écart type $\frac{1}{\sqrt{2}}$ M. de Finetti[5]) avait calculé la médiane de la dernière valeur pour une distribution ayant l'écart type 1. Donc ces valeurs seront comparables aux nôtres après division par $\sqrt{2}$. Ces valeurs de t sont portées aussi dans la figure (2). On voit que la médiane est toujours un peu supérieure à la dominante, mais que la différence décroît avec le nombre d'observations comme la théorie le prescrit. Puisque les deux conditions finales sont remplies il est légitime d'appliquer la transformation finale (15) qui sera d'après (44') et (47)

$$y = 2\bar{t}\,(t - \bar{t}).$$

La distribution finale de la dernière valeur sera d'après (16)

$$\mathfrak{w}(x) = 2h\bar{t}e^{-y-e^{-y}}. \tag{54}$$

La densité de probabilité de la dominante sera

$$\mathfrak{w}(u) = \frac{2h^2(u-\bar{x})}{e}.$$

Donc

$$\frac{d\mathfrak{w}(u)}{dN} = \frac{2h^2}{e}\frac{du}{dN} =$$

$$= \frac{h}{N\tau e} \tag{55}$$

d'après (52). La densité de probabilité de la dominante augmentera avec le nombre d'observations mais d'une façon extrêmement lente. Cela veut dire que les distributions des plus grandes valeurs se resserrent pour des nombres croissants d'observations, et l'on peut augmenter, mais d'une façon extrèmement lente, la précision de la dernière valeur en augmentant le nombre d'observations.

En effet la dispersion de la distribution de la plus grande valeur sera d'après (38) et (44')

$$\sigma^2 = \frac{\pi^2}{6}\frac{1}{4\tau^2h^2}$$

[16]) J. Splawa-Neyman: Sur les valeurs théoriques de la plus grande de n erreurs. Prac Matematyczno-Fizycznych, t. 33.

dont on tire l'écart type

$$\sigma = \frac{0{,}64128}{h\tau}.$$

De même l'écart probable sera d'après (38″)

$$\varrho = \frac{0{,}39313}{h\tau}. \tag{56'}$$

Le fait que l'écart probable est inverse à t, a déjà été trouvé par M. Finetti. Ainsi, plus l'espérance mathématique de la dernière valeur est grande, plus l'écart type et l'écart probable sont petits. Les deux écarts de la distribution de la plus grande valeur augmentent avec l'écart type initial et décroissent avec le nombre d'observations. Puisque dans les expressions (37′) il y a le facteur multiplicatif α du coté gauche, tous les moments diminuent, mais d'une façon très lente, avec le nombre d'observations. Ce n'est qu'une autre forme du fait que la distribution finale de la distribution de Gauss se resserre pour un nombre croissant d'observations.

Dans ce qui suit nous allons traiter deux distributions initiales dont les distributions finales de la plus grande valeur se comportent tout à fait autrement.

Pour la distribution exponentielle

$$w(x) = \beta e^{-\beta x} \tag{57}$$

ayant l'espérance mathématique

$$\bar{x} = 1/\beta. \tag{58}$$

la condition (12) n'est pas remplie. Car

$$\frac{w(x+k)}{w(x)} = e^{-\beta k}$$

ne disparait pas pour les valeurs croissantes de x. La probabilité d'une valeur inférieure à x sera

$$W(x) = 1 - e^{-\beta x}. \tag{59}$$

La dominante sera d'après (7)

$$u = x \lg N$$

calcul qui est légitime puisque l'équation (6) est verifiée pour toute valeur de N. On aura

$$Nw(u) = \beta \tag{60}$$

constant, indépendant de N. Ce fait simplifiera tous nos calculs. Il suffit naturellement à la condition II des propriétés finales. L'égalité (13) que nous voulons verifiée au moins approximativement pour des grandes valeurs de la variable, est exacte ici. Car

$$\frac{w^{(\nu)}(u)}{w(u)} = (-1)^\nu \beta^\nu$$

et de même

$$(-1)^\nu \left(\frac{w(u)}{1 - W(u)}\right)^\nu = (-1)^\nu \beta^\nu.$$

Donc la distribution finale sera exactement

$$\mathfrak{w}(x) = \beta\, e^{-y - e^{-y}} \tag{61}$$

où la variable finale est

$$y = \beta \left(x - \frac{\lg N}{\beta}\right)$$
$$= \beta x - \lg N. \tag{62}$$

En introduisant (62) on peut écrire la distribution finale

$$\mathfrak{w}(x) = \beta N e^{-\beta x - N e^{-\beta x}}. \tag{63}$$

D'ailleurs, on peut arriver à ce résultat sans faire usage du développement de la probabilité $W(x)$ autour de la dominante u. En introduisant l'identité

$$\frac{1}{N}\, e^{\beta u} = 1$$

dans la distribution (4) de la plus grande valeur on obtient à l'aide de (59)

$$W^N(x) = \left(1 - \frac{1}{N}\, e^{-\beta(x-u)}\right)^N.$$

Pour des valeurs suffisamment grandes de N on aura

$$W^N(x) = e^{-N e^{-\beta x}}$$

dont on tire (63) à l'aide de (3).

L'espérance mathématique de la plus grande valeur sera d'après (25)

$$\bar{u} = \bar{x}\, (\lg N + \gamma). \tag{64}$$

Donc pour des valeurs croissantes d'observations l'espérance mathématique de la plus grande valeur tendra, d'ailleurs assez lentement, vers la dominante quoique la condition (12) ne soit pas remplie.

Il s'agit maintenant de fixer, à partir de quelle valeur de N on veut négliger la différence entre ces deux valeurs. Si l'on choisit la même précision comme pour la loi de Gauss, en d'autres termes quand on considère

$$18 >> 1 \tag{65}$$

ou une différence de 5,6% comme négligeable, le nombre d'observations nécessaire pour identifier l'espérance mathématique de la plus grande valeur avec la dominante devient

$$N \geqq \tfrac{1}{3} \cdot 10^5. \tag{65'}$$

Puisque la constante γ est calculée par

$$\gamma = \lim_{N=\infty} \left(\sum_{\nu=1}^{N} \frac{1}{\nu} - \lg N \right)$$

on peut écrire pour des valeurs de N extrêmement grandes

$$\bar{u} = \bar{x} \cdot \sum_{\nu=1}^{N} \frac{1}{\nu}.$$

Pour l'écart type on aura d'après (38) et (60)

$$\sigma = 1{,}28255\bar{x} \tag{66}$$

indépendant de N. Donc à partir d'un certain nombre d'observations on ne peut plus augmenter la précision en augmentant ce nombre. D'ailleurs tous les moments sont indépendants du nombre d'observations à l'exception de l'espérance mathématique. Si le nombre d'observations est suffisamment grand, la distribution de la plus grande valeur ne change plus la forme, elle reste la même et ne fait que se déplacer vers la droite pour des nombres croissants d'observations.

Cette qualité tout à fait différente de celle de la distribution finale pour la loi de Gauss tient à ce que la condition (12) n'est pas remplie pour la distribution exponentielle. D'ailleurs le traitement de la distribution exponentielle nous mène à une autre forme pour énoncer la condition I des propriétés finales: La distribution initiale doit se comporter à l'infini comme si elle était une distribution exponentielle.

Comme dernier exemple nous allons traiter la distribution

$$w(x) = a^2 x e^{-ax} \tag{67}$$

pour le domaine positif. Elle représente un cas spécial très simple de la distribution III de Pearson. L'espérance mathématique est

$$\bar{x} = \frac{2}{a}. \tag{68}$$

La condition (12) n'est pas remplie. Car pour des valeurs croissantes de x on aura

$$\frac{w(x+k)}{w(x)} = \frac{x+k}{x} e^{-ak}$$
$$\to e^{-ak}$$

grandeur indépendante de x qui ne tend pas vers zéro. La probabilité d'une valeur inférieure à x devient

$$W(x) = \int_0^{ax} te^{-t}\,dt$$
$$= 1 - axe^{-ax} - e^{-ax}. \quad (69)$$

La condition (6) nous permet le calcul de la dominante. L'égalité (6)

$$\frac{a^2u}{1+au} = -\frac{1}{u} + a$$

sera réalisée dès que

$$a^2u^2 >> 1 \quad (70')$$

Alors la dominante de la plus grande valeur calculée par (7) sera la solution de

$$\frac{1}{N} = \alpha u e^{-\alpha u} \quad (71)$$

ou de

$$\alpha u - \lg \alpha u = \lg N. \quad (72)$$

Si nous nous limitons à la même précision employée dans la distribution de Gauss, ce calcul sera légitime à partir de $a^2u^2 = 18$, c'est à dire à partir de $N = 13$. Pour de grandes valeurs de N on peut poser

$$a = \alpha$$

équation qui sera légitime lorsque

$$au > 18 \quad (70)$$

c'est-à-dire

$$N > \tfrac{1}{3} \,.\, 10^7.$$

A partir de ce nombre élevé la dominante u est déjà une fonction linéaire de $\lg N$. Car en différentiant (72) on obtient

$$\frac{du}{d \lg N} = \frac{1}{\alpha\left(1 - \frac{1}{\alpha u}\right)}$$

ce qui devient d'après la condition (70)

$$\frac{du}{d \lg N} = \frac{1}{\alpha}.$$

Donc l'augmentation de la valeur dominante avec le logarithme des nombres d'observations est proportionelle à l'espérance mathématique de la distribution initiale. En première approximation on peut poser

$$u = \frac{\bar{x}}{2} \lg N. \quad (73)$$

Pour le calcul de l'espérance mathématique il faut traiter la condition (13).

obtient pour les dérivées de la distribution initiale d'après la formule connue pour la $\nu^{\text{ième}}$ dérivée d'un produit

$$\mathfrak{w}^{(\nu)}(x) = \alpha^2 e^{-\alpha x} [(-1)^\nu \alpha^\nu x + (-1)^{\nu-1} \nu \alpha^{\nu-1}].$$

En introduisant la valeur de la dominante par (71) on obtient

$$w^{(\nu)}(u) = (-1)^\nu \alpha^{\nu+1} \left(\frac{1}{N} - \nu e^{-\alpha u}\right)$$

ou à l'aide de (71)

$$w^{(\nu)}(u) = \frac{(-1)^\nu \alpha^{\nu+1}}{N}\left(1 - \frac{\nu}{\alpha u}\right).$$

De l'autre côté on aura

$$(-1)^\nu \frac{w^{\nu+1}(u)}{[1-W(u)]^\nu} = \frac{(-1)^\nu \alpha^{\nu+1}(\alpha u)^{\nu+1} e^{-(\nu+1)\alpha u}}{e^{-\nu \alpha u}(1+\alpha u)^\nu} = \frac{(-1)^\nu \alpha^{\nu+1}}{N} \frac{1}{\left(1 + \frac{1}{\alpha u}\right)^\nu}$$

en introduisant (71). Puisque

$$w^{(\nu)}(u) = (-1)^\nu \frac{w^{\nu+1}(u)}{[1-W(u)]^\nu}$$

l'égalité (13) peut être considérée comme verifiée par la condition (70) comme nous l'avons prévu dans le cas général.

Cela nous permet la construction de la distribution finale. On aura d'après (67), (69) et (70)

$$Nw(u) = a \tag{74}$$

indépendant du nombre des observations comme pour la distribution initiale exponentielle. Donc, l'espérance mathématique sera d'après (25)

$$\bar{u} = u + \frac{\gamma}{a}. \tag{75}$$

Plus l'espérance mathématique de la distribution initiale sera grande, plus l'espérance mathématique de la dernière valeur s'éloignera de sa dominante. Mais pour un nombre croissant d'observations l'espérance mathématique de la plus grande valeur tendra vers la dominante. Car on peut poser en première approximation

$$a\bar{u} = \lg N + \gamma$$

où γ peut certainement être négligé vis à vis de $\lg N$ pour le nombre d'observations nécessaire à la justification de toute la construction. En conséquence de (74) la distribution finale sera

$$\mathfrak{w}(x) = a e^{-(x-u)a - e^{-(x-u)a}}. \tag{76}$$

Tous les moments, à l'exception de l'espérance mathématique, ne dépendront plus du nombre d'observations. Donc cette distribution possède à peu près les mêmes propriétés que la distribution exponentielle. Seulement le nombre d'observations nécessaire pour arriver à la distribution finale est beaucoup plus grand et l'espérance mathématique de la plus grande valeur ne sera que la moitié de celle de la distribution exponentielle.

Conclusion.

Pour une distribution initiale illimitée vers la droite et pour N observations il existe une distribution de la plus grande valeur. Sa dominante peut être calculée à l'aide de la formule (7) pourvu que le nombre des observations soit assez grand pour permettre l'introduction de la règle de l'Hopital sous la forme (6). La condition (13) identique à (6) nous permet la construction de la forme finale de la distribution de la plus grande valeur pour un très grand nombre d'observations. Cette distribution finale est doublement exponentielle. C'est pour elle que nous avons calculé l'espérance mathématique de la plus grande valeur (25), qui se rapproche sous certaines conditions de la dominante. En outre nous avons trouvé une formule de recurrence (37) qui permet le calcul de tous les moments à l'aide des sommes S_ν connues par la théorie des fonctions Gamma.

Pour la distribution exponentielle la distribution finale est atteinte même pour de très petits nombres d'observations. Par contre on a besoin de 100 000 observations pour la distribution de Gauss.

Pour une distribution initiale exponentielle et pour le cas spécial traité de la distribution III de Pearson, tous les moments à l'exception de l'espérance mathématique seront indépendants du nombre d'observations. Donc les distributions finales se deplaceront pour des valeurs croissantes d'observations vers la droite suivant l'axe de la variable. Pour la distribution de Gauss les distributions finales se resserrent pour des valeurs croissantes d'observations. En choisissant la même précision pour les trois distributions traitées on obtient: L'identification entre l'espérance mathématique de la plus grande valeur et la dominante est légitime si le nombre des observations surpasse l'ordre de grandeur $\frac{1}{8}10^5$ pour la distribution exponentielle, 10^5 pour la distribution de Gauss et $\frac{1}{3}10^7$ pour le cas spécial traité de la distribution III de Pearson. Ce nombre élevé provient du fait que cette distribution finale ne se resserre pas.

La distribution finale de la dernière valeur pour toute distribution initiale qui se comporte à l'infini comme une exponentielle, tend vers la distribution finale de la distribution initiale exponentielle. Son espérance mathématique tend vers la dominante. La distribution finale de la dernière valeur se propage à droite pour des nombres d'observations croissants. Si la valeur α, définie par (5), tend vers une constante, la distribution ne change plus de forme. Si la valeur α augmente indéfiniment, elle se resserre. Dans le premier cas, la dominante de la plus grande valeur augmente comme le logarithme du nombre d'observations, dans l'autre cas, elle augmente plus lentement.