Les plus grands âges en Suisse

Par E. J. Gumbel, Université de Lyon

Page

Le Bureau Fédéral Statistique de la Suisse a mis à ma disposition ses observations sur les plus grands âges [1]). Le cours de ces valeurs est représenté dans la figure (1), qui contient pour chaque année de l'intervalle 1879—1930 et pour les deux sexes l'âge en années entières après lequel le dernier survivant est mort. Les traits entre les diverses observations n'ont aucune signification statistique. Ils sont mis seulement pour simplifier l'aspect.

Cette série est précieuse puisque des données exactes sur les hauts âges sont assez rares. Car les statistiques démographiques ne donnent en général les décès survenus après l'âge de 100 ans que d'une façon globale. Les statisticiens ne se sont occupés des hauts âges que rarement. Car l'opinion, suivant laquelle il s'agit ici de cas individuels qui ne peuvent pas être soumis à des règles statistiques, est assez répandue. Il s'en suivrait que le traitement de telles observations ne présente aucun intérêt scientifique et ne peut aboutir à aucun résultat sérieux. Le but de cet article est de refuter cette opinion et de montrer que cette série suit bien les règles du calcul des probabilités.

Il est juste qu'il s'agit ici de cas individuels, mais ce fait ne s'oppose pas en principe à un traitement statistique, pourvu qu'ils soient basés sur de nombreuses observations. C'est évidemment le cas, puisque dans les 52 ans le nombre des décès était de 1.508.880 hommes et de 1.454.821 femmes. Donc notre problème entre dans la ligne des recherches de Bortkiewicz, qui ont abouti à la loi des événements rares. Car le nombre d'observations est très grand tandis que la probabilité de l'événement envisagé est très petite.

Notre direction de recherches sera la suivante: d'abord nous traitons ces observations du point de vue élémentaire. Puis nous montrons le schéma du calcul des probabilités capable de donner un sens statistique à ces valeurs. Nous établissons trois méthodes de comparaison qui nous mèneront à un accord satisfaisant entre la théorie et les observations. Nous allons faire ces recherches sans employer trop de formules mathématiques et sans choisir une formule

---

[1]) C'était par l'intermédiaire du professeur Friedli, Berne. Je profite de cette occasion pour le remercier.

analytique pour les fonctions biométriques. Le second point distingue ce travail de quelques travaux précédents [1] [2] où nous avons supposé que la distribution des décès suivant les âges soit gaussienne pour les hauts âges [3]).

1. Observations

Nous traitons l'âge, en années entières, après lequel le dernier survivant d'une année est mort. Pour simplifier, cette grandeur sera appelée le plus grand âge d'une année déterminée. Il est évident que la série représentée dans la figure (1) n'a aucune tendance, qu'elle est irrégulière du point de vue du temps. Donc il est légitime de supposer qu'elle est stationnaire, c'est-à-dire qu'on peut la traiter comme distribution. Dans ce but on prend le plus grand âge comme variable statistique et l'on énumère combien de fois les différents âges sont arrivés.

C'était par exemple une fois, dans l'année 1890, que le plus vieux homme est mort à 96 ans, une autre fois, en 1904, que cet âge était 104 ans, tandis qu'il y avait deux ans (1920 et 1926) où le plus vieux était mort après 102 ans. Les observations classées de cette manière se trouvent dans la table I. La première colonne contient les âges, les quatre autres le nombre absolu des années dans lesquelles cet âge était le plus grand. La deuxième et la troisième colonne traitent chaque sexe, la quatrième les deux sexes réunis, la dernière se rapporte au plus vieux des deux sexes. Cette colonne ne peut pas être réduite aux deux premières, car elle ne contient pour chaque année qu'une observation, l'âge après lequel le plus vieux des deux sexes est mort.

I. Distribution des plus grands âges

Ages	Hommes	Femmes	Ensemble	Le plus âgé des deux sexes
96	1	—	1	—
97	8	3	11	1
98	9	6	15	3
99	9	9	18	6
100	10	12	22	13
101	5	10	15	11
102	2	8	10	7
103	7	1	8	7
104	1	2	3	3
105	—	—	—	—
106	—	1	1	1
Nombre d'années N_1	52	52	104	52
Age limite $\bar{\omega}$	99,596	100,250	99,923	100,942
Ecart type σ	2,040	1,838	1,970	1,791

[1]) L'âge limite. Bulletin de la Société mathématique de France. Tome LX, fasc. 3/4. Paris 1932.

[2]) Das Alter des Methusalem. Zeitschrift für Schweizerische Statistik und Volkswirtschaft, 69. Band, Heft 4, 1933.

[3]) Die Verteilung der Gestorbenen um das Normalalter. Aktuarske Vedy, Bd. IV Nr. 2. Prag 1933.

Pour caractériser ces distributions on calcule les moyennes arithmétiques et les écarts types. La variable étant le plus grand âge, la moyenne arithmétique sera appelée l'âge limite $\bar{\omega}$. L'écart type σ mesure la précision de cette valeur. Plus grand est l'écart type, moins précise sera la détermination de la moyenne. Le calcul de ces deux valeurs ne présente aucune difficulté arithmétique. On le fait par des méthodes bien connues à partir de l'âge arbitrairement choisi de 100 ans. Ces valeurs sont données dans les dernières lignes de la table I.

L'âge limite des hommes est moindre, son écart type plus grand que celui des femmes. Les valeurs pour les deux sexes réunis se tiennent naturellement entre les valeurs des sexes séparés. L'âge limite pour le plus vieux des deux sexes est plus grand, son écart type plus petit que les valeurs pour les trois autres distributions.

Ces distributions ne sont pas symétriques. Envisageons les valeurs extrêmes: la distribution des hommes varie entre 96 et 104 ans, mais celle des femmes entre 97 et 106 ans. Les mêmes limites valent pour la quatrième distribution. Enfin la distribution pour les deux sexes varie de 96 à 106 ans.

Pour classer de telles observations on emploie dans la statistique encore un autre procédé: la fréquence cumulative. C'est le nombre relatif des cas où la variable, le plus grand âge, est inférieur ou égal à une valeur donnée. Par exemple, pour les hommes c'est seulement une fois sur 52 que le plus grand âge était inférieur ou égal à 96 ans. Pour l'âge 97 ce nombre est 9 sur 52, etc. Enfin tous les âges étaient inférieurs ou égaux à 106 ans. Ces fréquences cumulatives sont données dans la table II.

II. Fréquences cumulatives des plus grands âges

Ages	Hommes	Femmes	Ensemble	Plus âgé des deux
96	0,019	—	0,010	—
97	0,173	0,058	0,115	0,019
98	0,346	0,173	0,260	0,077
99	0,519	0,346	0,433	0,192
100	0,712	0,577	0,644	0,442
101	0,808	0,769	0,788	0,654
102	0,846	0,923	0,885	0,788
103	0,981	0,942	0,962	0,923
104	1,000	0,981	0,990	0,981
105	1,000	0,981	0,990	0,981
106	1,000	1,000	1,000	1,000

Les différences successives de ces grandeurs sont naturellement les valeurs relatives de la table I. Le cours discontinu de ces valeurs est donné par les traits orthogonaux des figures 2—5.

Enfin il existe une troisième manière élémentaire de traiter ces séries. Considérons toutes les observations faites jusqu'à une certaine année. Alors il existe une influence très intéressante du temps représentée dans la figure (1). Pour les hommes le plus grand âge commence en 1879 avec 101 ans. Le plus grand âge observé jusqu'à une certaine époque monte en deux étapes jusqu'à 104 ans. Pour les femmes ces valeurs montent de 99 ans en trois étapes jusqu'à 106 ans. En conséquence le plus grand âge augmente avec la durée observée, c'est-à-dire avec le nombre d'observations. Ce phénomène est absolument inexplicable, si l'on croit à l'existence d'un âge limite fixe.

2. Théorie de la plus grande valeur

Il s'agit maintenant de rechercher si l'on peut donner une raison statistique pour ces distributions, si l'on peut savoir pourquoi elles sont asymétriques, si l'on peut expliquer le champ de variation, enfin pourquoi les plus grands âges augmentent avec le nombre d'observations. Dans cette recherche nous nous basons sur un fait assez élémentaire: les plus grands âges sont les plus grandes valeurs des âges au moment de la mort. Cela nous mène à comparer ces observations aux valeurs extrêmes d'une variable statistique: l'âge au moment de la mort. Nous supposons que la distribution des décès suivant les âges est illimitée vers la droite, c'est-à-dire qu'un âge limite fixe n'existe pas. Cette supposition que nous avons déjà traitée dans notre article précédent paraît hardie. Pourtant c'est là la clef d'une explication de ces observations.

Donc nous avons besoin de la distribution de la plus grande valeur d'une variable statistique illimitée. Cette fonction a été traitée dans plusieurs travaux précédents [1]). Nous nous limitons à résumer les résultats.

Soit $w(x)$ une distribution initiale illimitée vers la droite et $W(x)$ la probabilité d'une valeur inférieure à x. Soit N le nombre des observations faites sur cette distribution. Parmi eux il y aura une qui sera la plus grande. On demande sa valeur. Mais il faut préciser cette question. Car il y a une probabilité

$$\mathfrak{W}(x) = \overset{N}{W(x)} \tag{1}$$

pour que la plus grande valeur parmi N observations soit moindre que x. Donc la plus grande valeur est elle-même une variable statistique ayant la probabilité

$$\mathfrak{w}(x) = N\, \overset{N-1}{W(x)}\, w(x), \tag{2}$$

fonction de N pour qu'elle soit contenue entre x et $x + dx$. Cette distribution aura une dominante $\tilde{u}$, c'est-à-dire une valeur qui est la plus probable, une espérance mathématique $\bar{u}$ et un écart type σ. Pour le cas spécial d'une distribution initiale de Gauss nous avons montré dans notre travail antérieur qu'on peut déterminer la dominante par

$$W(\tilde{u}) = 1 - \frac{1}{N} \tag{3}$$

---

[1]) La plus petite valeur parmi les plus grandes. Comptes Rendus 196, p. 1857. Paris 1933.
La plus petite valeur parmi les plus grandes et la plus grande valeur parmi les plus petites. Comptes Rendus 197, p. 965. Paris 1933.

Cette équation vaut non seulement pour cette distribution initiale, mais pour maintes autres. Elle demande qu'il soit légitime de poser pour de grands nombres d'observations

$$\frac{w(\tilde{u})}{1 - W(\tilde{u})} \longrightarrow - \frac{w'(\tilde{u})}{w(\tilde{u})} \tag{4}$$

La dominante de la dernière valeur augmentera avec les observations. La manière dont elle augmente dépend de la distribution initiale et en espèce des constantes qui y existent.

Nous précisons maintenant notre question en cherchant la forme vers laquelle la distribution (2) tend [1]), si N tend vers de très grandes valeurs. Pour résoudre cette question on doit demander quelques propriétés analytiques de la distribution initiale. Nous supposons que ses dérivées successives $w^{(\nu)}(x)$ aient pour la dominante, pour de très grandes valeurs de N et pour $\nu = 1, 2, 3 \ldots$ la propriété, conséquence de (4)

$$\frac{w^{(\nu)}(\tilde{u})}{w(\tilde{u})} \longrightarrow (-)^{\nu} \left(\frac{w(\tilde{u})}{1 - W(\tilde{u})} \right)^{\nu} \tag{4'}$$

C'est d'abord une condition analytique, mais en même temps c'est une condition numérique. Car, si cette relation est vérifiée du point de vue analytique, on peut fixer une précision de sorte que cette condition vaut à partir d'un certain nombre d'observations.

Posons pour le second membre

$$\frac{w(\tilde{u})}{1 - W(\tilde{u})} = \alpha \tag{5}$$

et introduisons une nouvelle variable statistique

$$y = \alpha\,(x - \tilde{u}). \tag{6}$$

Alors la distribution finale de la dernière valeur sera

$$\mathfrak{w}(x) = \alpha\, e^{-y-e^{-y}} \tag{2'}$$

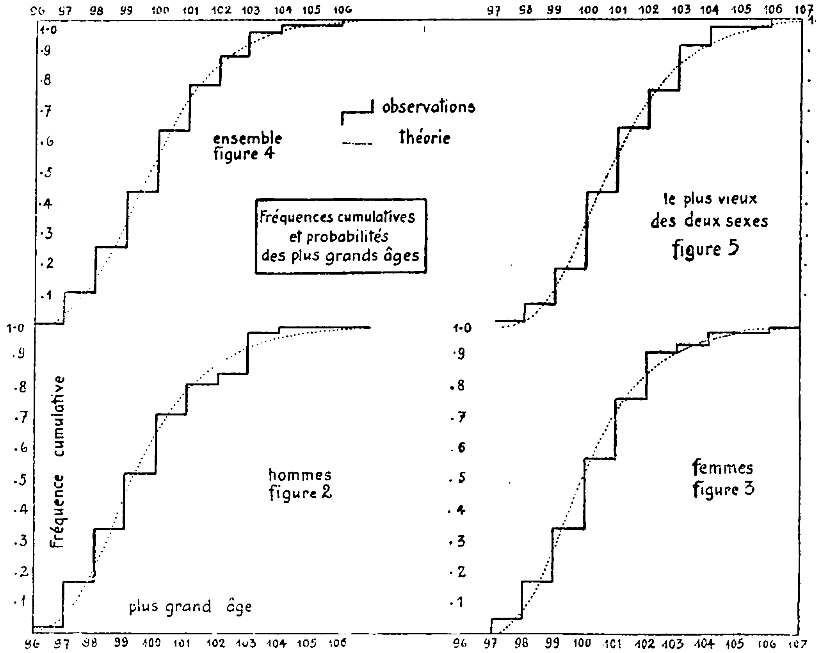
et la probabilité pour que la dernière valeur soit inférieure à x sera

$$\mathfrak{W}(x) = e^{-e^{-y}} \tag{1'}$$

La table III donne les valeurs de $\mathfrak{W}(x)$. On obtient à l'aide de (2′) l'espérance mathématique de la plus grande valeur [2])

[1]) La distribution limite de la plus petite valeur parmi les plus grandes. Comptes Rendus 197, p. 1082. Paris 1933.

[2]) L'espérance mathématique de la m^e valeur. Comptes Rendus 198, p. 33. Paris 1934.



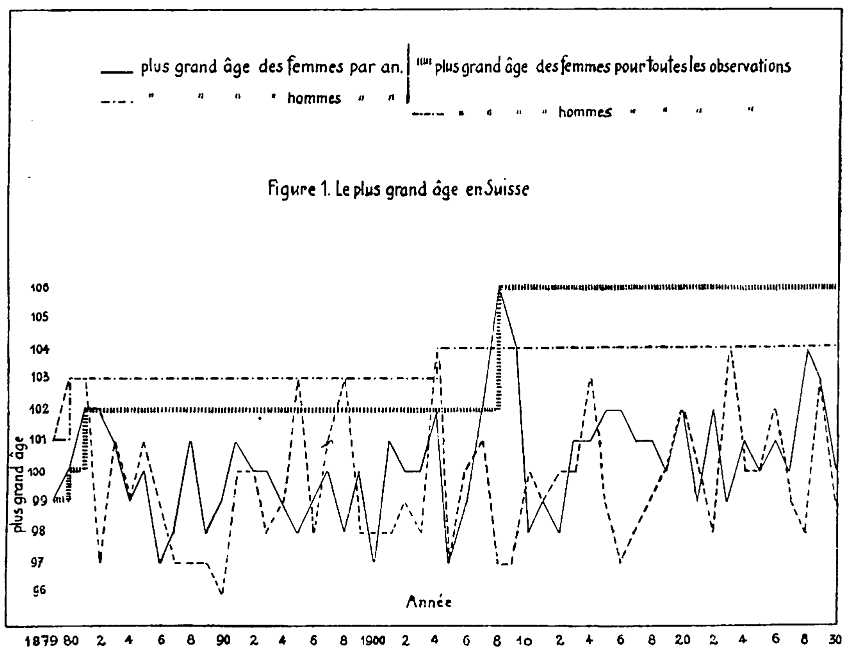


Figure 1. Le plus grand âge en Suisse

$$u = \tilde{u} + \frac{0{,}57722}{\alpha} \tag{7}$$

et l'écart type de cette distribution [1]) par

$$\alpha\,\sigma = 1{,}28255. \tag{8}$$

Donc pour des nombres d'observations croissants l'espérance mathématique de la plus grande valeur augmente. On peut spécifier la manière pour une distribution initiale donnée.

Dans notre travail précédent nous avons traité spécialement la distribution initiale de Gauss et nous avons supposé que la distribution finale est elle-même gaussienne. Du point de vue analytique cette opinion assez répandue n'est pas juste. Elle mène à un écart type

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

Pour ce calcul les deux théories ne conduisent qu'à une différence de 15 %.

Pour appliquer cette théorie aux observations précédentes il faut préciser les significations biométriques des symboles qui entrent dans la théorie des plus grandes valeurs. La variable statistique traitée ici est l'âge au moment de la mort. La distribution initiale $w(x)\,dx$ est la probabilité pour que cet âge soit situé entre x et $x + dx$. Appelons cette grandeur $\vartheta(x)\,dx$. La probabilité $1 - W(x)$ d'une valeur supérieure à x sera la probabilité $l\,(x)$ d'atteindre cet âge. La plus grande valeur sera le plus grand âge, sa dominante $\tilde{u}$ le plus grand âge dominant $\tilde{\omega}$. L'espérance mathématique de la plus grande valeur u sera l'âge limite ω. Enfin le coefficient α sera l'intensité de mortalité au plus grand âge dominant $\mu\,(\tilde{\omega})$.

En appliquant la théorie des plus grandes valeurs aux plus grands âges nous ne choisissons pas une formule déterminée pour les tables de mortalité. Au lieu de cela nous supposons seulement que les dérivées successives de la distribution des décès suivant les âges vérifient pour un nombre d'observations suffisamment grand la condition (4') que l'on peut écrire

$$\frac{\vartheta^{(\nu)}(\tilde{\omega})}{\vartheta\,(\tilde{\omega})} = (-1)^{\nu}\,\mu^{\nu}(\tilde{\omega}). \tag{9}$$

Cette condition est remplie parmi d'autres pour la distribution de Gauss et pour celle de Gompertz, formules de mortalité en commun usage. Donc cette supposition est légitime quoique la vraie nature des dérivées successives de la distribution des décès soit inconnue.

---

[1]) Les moments des distributions finales de la première et de la dernière valeur. Comptes Rendus 198, p. 141. Paris 1934.

III. La somme $\mathfrak{W}(x)$ de la distribution finale

y	$\mathfrak{W}(x)$	y	$\mathfrak{W}(x)$
— 2,0	0,00062	4,5	0,98895
— 1,5	0,01132	5,0	0,99328
— 1,0	0,06599	5,5	0,99592
— 0,5	0,19230	6,0	0,99752
0,0	0,36788	6,5	0,99850
0,5	0,54525	7,0	0,99909
1,0	0,69220	7,5	0,99945
1,5	0,80003	8,0	0,99967
2,0	0,87345	8,5	0,99980
2,5	0,92120	9,0	0,99987
3,0	0,95143	9,5	0,99993
3,5	0,97025	10,0	0,99996
4,0	0,98185		

3. Comparaison entre la théorie et les observations

La première vérification de notre hypothèse est le fait que le plus grand âge augmente avec le nombre d'observations. Car on peut identifier le plus grand âge observé jusqu'à une certaine année à l'espérance mathématique de la plus grande valeur de la distribution des décès pour le nombre d'observations correspondant. Alors le plus grand âge doit augmenter avec le temps puisque le nombre des décès survenus jusqu'à une certaine année augmente. D'ailleurs il serait possible de spécifier cette influence en calculant comment l'âge limite augmente avec le temps. Mais cette recherche supposerait un choix spécial d'une formule biométrique, par exemple celle de Gompertz, ce qui sortirait du cadre assigné à cet article.

Pour continuer l'analogie entre le plus grand âge et la plus grande valeur d'une variable statistique, il s'agit de comparer les observations avec la distribution théorique du plus grand âge. Nous nous limitons d'abord aux trois premières distributions. Cette comparaison est vite faite. En écrivant les formules nécessaires dans les désignations biométriques la fréquence cumulative du plus grand âge doit être d'après (1')

$$\mathfrak{W}(x) = e^{-e^{-\mu(\tilde{\omega})(x-\tilde{\omega})}} \tag{10}$$

On n'a qu'à déterminer les deux constantes $\mu(\tilde{\omega})$ et $\tilde{\omega}$ par les observations. Cela mène à la résolution des équations (8) et (7) qui seront

$$\sigma = \frac{1,28255}{\mu(\tilde{\omega})} \tag{11}$$

et

$$\bar{\omega} = \tilde{\omega} + \frac{0,57722}{\mu(\tilde{\omega})} \tag{12}$$

Les parties gauches sont la moyenne arithmétique et l'écart type, valeurs observées et données à la fin de la table I. On obtient l'intensité de mortalité pour l'âge dominant par

$$\frac{1}{\mu(\tilde{\omega})} = 0{,}77970\,\sigma \tag{11'}$$

Puis on aura le plus grand âge dominant

$$\tilde{\omega} = \bar{\omega} - 0{,}45005\,\sigma \tag{12'}$$

On n'a qu'à introduire ces valeurs dans (10) pour pouvoir comparer les fréquences cumulatives observées aux probabilités théoriques. Cette comparaison est faite dans le paragraphe suivant.

Enfin il y a un troisième critère pour juger l'accord entre la théorie et les observations. Si les plus grands âges peuvent être interprétés comme plus grandes valeurs, on doit se demander si les valeurs extrêmes des plus grands âges sont justifiées par notre théorie.

Le calcul des valeurs théoriques des maxima est simple. Car ce sera la plus grande valeur qu'on doit attendre pour la distribution (2′) de la plus grande valeur pour N_1 observations. Cette grandeur doublement extrême aura, comme la plus grande valeur elle-même, une distribution à laquelle on arrive en répétant les opérations qui nous ont mené à la distribution (2′). Si l'on traite (2′) comme distribution initiale, la distribution finale de la plus grande valeur est pour N_1 observations

$$\mathfrak{w}_1(y) = e^{-(y - lg\,N_1) - e^{-(y - lg\,N_1)}},$$

dont on tire la dominante

$$\tilde{y}_1 = lg N_1.$$

La dominante du maximum du plus grand âge sera d'après (6)

$$\tilde{\omega}_1 = \tilde{\omega} + \frac{lg\,N_1}{\mu(\tilde{\omega})} \tag{13}$$

En choisissant pour les trois distributions les valeurs respectives de $\tilde{\omega}$, $\mu(\tilde{\omega})$ et N_1 on aura les valeurs théoriques des maxima des plus grands âges comparables aux valeurs observées.

Le calcul du minimum du plus grand âge est un peu plus compliqué. La plus petite valeur [1]) d'une distribution initiale $\mathfrak{w}(y)$ possède pour N_1 observations la densité de probabilité

$${}_1\mathfrak{w}(y) = N_1\,(1 - \mathfrak{W}(y))^{N_1 - 1}\,\mathfrak{w}(y).$$

La dominante ${}_1\tilde{y}$ de la plus petite valeur sera la solution de

[1]) La distribution limite de la plus grande valeur parmi les plus petites. Comptes Rendus 197, p. 1381. Paris 1933.

$$\frac{N_1 - 1}{1 - \mathfrak{W}}\,\mathfrak{w} = \frac{\mathfrak{w}'}{\mathfrak{w}}$$

En introduisant (1') et (2') et en posant $e^{-{}_1y} = z$, on obtient

$$\frac{N_1 - 1}{1 - e^{-z}}\, z\, e^{-z} = z - 1.$$

Alors il s'agit de résoudre

$$N_1 - 1 = \left(1 - \frac{1}{z}\right)(e^z - 1) \tag{14}$$

pour trouver la dominante de la plus petite valeur

$${}_1\tilde{y} = -\,lg\, z \tag{15}$$

On en tire d'après (6) la dominante ${}_1\tilde{\omega}$ du minimum du plus grand âge

$${}_1\tilde{\omega} = \tilde{\omega} + \frac{{}_1\tilde{y}}{\mu(\tilde{\omega})} \tag{16}$$

A l'aide des trois valeurs $\tilde{\omega}$, $\mu(\tilde{\omega})$ et N_1 on peut comparer ces grandeurs avec les observations de la table I.

4. Valeurs numériques

Nous possédons à présent l'outil mathématique nécessaire pour juger si les observations des plus grands âges s'accordent à la théorie des probabilités. Chacun des 104 chiffres est une valeur observée du plus grand âge. D'après notre théorie les fréquences cumulatives (table II) doivent s'accorder aux probabilités (1') des plus grandes valeurs. Calculons d'abord à l'aide de (11') et (12') les constantes nécessaires pour cette comparaison pour les trois distributions. La quatrième sera traitée plus tard. Les valeurs sont données dans la table suivante.

IV. Constantes pour les distributions du plus grand âge

	Hommes	Femmes	Ensemble
Dominante $\tilde{\omega}$	98,678	99,423	99,041
Intensité de mortalité à l'âge dominant $\mu(\tilde{\omega})$.	1:1,591	1:1,433	1:1,536

La dominante pour les hommes est inférieure à celle pour les femmes. Les valeurs pour l'ensemble se tiennent entre les valeurs pour les deux sexes séparés. La même hiérarchie vaut pour l'intensité de mortalité. Les dominantes sont toutes proches des âges limites observés (table I), ce qui convient au fait que ces distributions sont assez serrées.

Pour la comparaison entre la théorie (1') et les observations (table II) il faudrait calculer les trois séries des valeurs $\mathfrak{W}(x)$. Au lieu de cela nous inversons la question. Pour des valeurs choisies de y on connaît par la table III la probabilité $\mathfrak{W}(x)$ des valeurs inférieures à y. Nous calculons pour des valeurs choisies de $\mathfrak{W}(x)$ les plus grands âges correspondants par (6) ce qui mène dans les désignations biométriques à

$$x = \tilde{\omega} + \frac{y}{\mu(\tilde{\omega})} \tag{17}$$

La table V donne les probabilités des âges inférieurs à x, grandeurs comparables à celles de la colonne 1 de la table II. La première colonne de la table V donne les valeurs y. La seconde, troisième et quatrième contiennent d'après (17) les valeurs correspondantes des plus grands âges pour les trois distributions envisagées, enfin la cinquième donne la probabilité pour que ces âges soient les plus grands.

Les figures 2—4 comparent les fréquences cumulatives observées des plus grands âges pour les hommes, les femmes et l'ensemble des deux sexes (table II) aux probabilités, valeurs théoriques de la table V. L'accord est tout à fait satisfaisant, puisque les courbes continues passent très bien à travers les valeurs discontinues. Une coïncidence n'est pas possible vu que les observations sont nécessairement discontinues, tandis que la théorie prévoit un cours continu.

Reste encore à traiter la distribution des plus âgés des deux sexes. La probabilité du plus grand âge des femmes d'être inférieur à x (colonne 3) est $\mathfrak{W}(x)$ (colonne 5). Soit $\mathfrak{W}_1(x)$ la probabilité pour les hommes et pour les mêmes âges. Alors la probabilité du plus grand âge du plus vieux des deux sexes d'être inférieur à x est $\mathfrak{W}(x) \cdot \mathfrak{W}_1(x)$ d'après le problème des probabilités composées. Mais la colonne 5 de la table V contient les valeurs $\mathfrak{W}(x)$ et $\mathfrak{W}_1(x_1)$ pour des âges différents. Pour former le produit pour les mêmes âges nous avons interpolé les probabilités des plus grands âges pour les hommes. La table V, colonne 6, donne les probabilités $\mathfrak{W}_1(x)$ qu'un âge x choisi pour les femmes (colonne 3) soit aussi le plus grand pour les hommes. En multipliant $\mathfrak{W}_1(x)$ avec la probabilité $\mathfrak{W}(x)$ que le même âge soit le plus grand pour les femmes (colonne 5), on obtient la probabilité pour que cet âge soit le plus grand pour le plus vieux des deux sexes (colonne 7). Il suffit de faire ce calcul à trois décimales. (Voir tableau à la page suivante.)

La figure 5 compare les fréquences cumulatives observées du plus grand âge du plus vieux des deux sexes (table II, colonne 5) aux valeurs théoriques de la table V, colonne 7. Cet accord aussi est tout à fait satisfaisant. L'application de la théorie de la plus grande valeur au plus grand âge est légitime.

On demandera enfin si la plus grande valeur théorique du plus grand âge est justifiée par les observations. En introduisant dans la formule (13) $N_1 = 52$ pour les deux sexes et $N_1 = 104$ pour l'ensemble on obtient à l'aide de la table IV les valeurs théoriques du maximum du plus grand âge pour les trois distributions.

V. Probabilités des plus grands âges

Variable	Plus grand âge x			Probabilités		
y	Hommes	Femmes	Ensemble	$\mathfrak{W}(x)$	$\mathfrak{W}_1(x)$	$\mathfrak{W}(x) \cdot \mathfrak{W}_1(x)$
1	2	3	4	5	6	7
— 1,5	96,29	97,29	96,74	0,01132	0,098	0,001
— 1,0	97,09	98,00	97,50	0,06599	0,219	0,014
— 0,5	97,88	98,71	98,27	0,19230	0,375	0,072
0	98,68	99,42	99,04	0,36788	0,534	0,196
0,5	99,47	100,13	99,81	0,54525	0,668	0,364
1,0	100,27	100,85	100,58	0,69220	0,770	0,533
1,5	101,06	101,56	101,34	0,80003	0,846	0,677
2,0	101,86	102,27	102,11	0,87345	0,898	0,784
2,5	102,65	102,98	102,88	0,92120	0,934	0,860
3,0	103,45	103,69	103,65	0,95143	0,957	0,911
3,5	104,25	104,40	104,42	0,97025	0,973	0,944
4,0	105,04	105,11	105,18	0,98185	0,982	0,965
4,5	105,84	105,83	105,95	0,98895	0,989	0,978
5,0	106,63	106,54	106,72	0,99328	0,993	0,986

Ces grandeurs portées dans la table VI sont presque identiques aux observations: le plus grand âge atteint pendant les 52 ans était 106 ans, tandis que notre théorie prévoit 106,17 ans. On ne peut pas demander plus, puisque l'âge du dernier survivant nous est connu seulement en nombre entier. Le calcul correspondant pour les minima du plus grand âge est fait à l'aide de la formule (14). Pour les distributions des plus grands âges des hommes et des femmes on a $N_1 - 1 = 51$. Une interpolation simple mène à $z = 4{,}2177$, dont on tire à l'aide de (15), ${}_1\tilde{y} = -1{,}4393$. Pour la distribution des deux sexes on a $N_1 - 1 = 103$. On obtient par la même méthode $z = 4{,}8729$ et ${}_1\tilde{y} = -1{,}5837$.

Les valeurs numériques des minima calculées par (16) à l'aide de la table IV sont portées dans la table VI. L'accord est frappant. Les différences entre la théorie et les observations sont toutes moindres qu'un an.

VI. Les extrèmes des plus grands âges

	Minimum		Maximum		Champ de variation	
	Observations	Théorie	Observations	Théorie	Observations	Théorie
	${}_1\omega'$	${}_1\tilde{\omega}$	ω_1'	$\tilde{\omega}_1$	$\omega_1' - {}_1\omega'$	$\tilde{\omega}_1 - {}_1\tilde{\omega}$
Hommes. .	96	96,39	104	104,96	8	8,57
Femmes . .	97	97,36	106	105,08	9	8,72
Ensemble .	96	96,61	106	106,17	10	9,56

Les deux dernières colonnes de la table VI contiennent les distances entre le maximum et le minimum. Les espérances mathématiques des extrêma seront pour les minima un peu inférieures, pour les maxima un peu supérieures aux valeurs calculées des dominantes. Mais ce calcul supplémentaire de même que celui de l'écart type ne vaut pas la peine.

5. Conclusion

Dans notre travail précédent nous avons calculé l'âge limite en analogie avec l'espérance mathématique de la plus grande valeur pour une distribution illimitée et pour un grand nombre d'observations en traitant la distribution des décès suivant les âges comme gaussienne. Nous avons supposé que la distribution finale de la plus grande valeur était de même gaussienne.

Dans cet article la connaissance de la forme finale de cette distribution nous a permis de libérer notre théorie de ces suppositions et de l'appliquer pourtant aux observations.

Pour les trois critères envisagés, la dépendance du plus grand âge de la durée envisagée, les fréquences cumulatives des plus grands âges et leurs valeurs extrêmes, notre théorie est bien vérifiée. Cet accord est possible parce que ces cas apparemment individuels reposent sur de nombreuses observations. Ceci rattache nos recherches aux travaux de Bortkiewicz. Lorsque cet auteur a prouvé que le nombre des soldats tués chaque année par un cheval suivait cette loi, il ne voulait guère augmenter nos connaissances hippiques. De même nous avons suivi le sort de la femme qui a atteint 106 ans dans le but de montrer que les valeurs extrêmes, considérées en général comme arbitraires, suivent bien les règles du calcul des probabilités, et qu'on peut même vérifier cela — pourvu que les observations sont bonnes.