E. J. Gumbel

Les centenaires

*Aktuárské vědy*, Vol. 7 (1938), No. 1, 10–17

Persistent URL: http://dml.cz/dmlcz/144681

bilden nach Formel (18) die Ausdrücke

$$\bar{A}_{\underbrace{x_{i_1}, x_{i_2}, \ldots x_{i_{l-1}}}_{1,2,\ldots l-1}, \overset{l}{x_l} x_{i_{l+1}} \ldots x_{i_n}}$$

Summieren wir alle so erhaltenen Gleichungen, so erhalten wir wegen

$$\sum Z^i_{l,\, l+1 \ldots n} = \binom{n-l+i}{i} Z_l^{n- \ +i}$$

$$\bar{A}_{x_1 x_2 \ldots \overset{l}{x_l} \ldots x_n} = \sum_{i=0}^{l-1} (-1)^i \binom{n-l+i}{i} Z_l^{n-l+i}.^{3)} \tag{19}$$

Bilden wir alle Gleichungen, die wir erhalten, wenn wir der Reihe nach $x_1, x_2, \ldots, x_n$ mit $x_l$ vertauschen und addieren wir diese Gleichungen, so erhalten wir

$$A^{\underline{n-l}}_{x_1 \ldots x_n} = \sum_{i=0}^{l-1} (-1)^i \binom{n-l+i}{i} Z^{n-l+i+1}. \tag{20}$$

Dabei bezeichnen wir mit $Z^{l+1} = \sum_{k=1}^{n} Z_k^l$. Die Formel (20) ist die bekannte $Z$-Formel.

---

# Les centenaires.

par *E. J. Gumbel,* Université de Lyon.

Les observations portant sur les âges élevés sont nécessairement rares, donc insuffisantes pour les calculs compliqués que demande la construction des tables de mortalité. Voila pourquoi les propriétés asymptotiques des fonctions biométriques sont encore objets de controverses. On prétend même assez souvent que ces observations sont tout à fait irrégulières. Le but de cet article est de prouver qu'au contraire elles suivent bien les règles du calcul des probabilités pourvu qu'on prenne une série d'années suffisamment longue.

Considérons les centenaires, c'est-à-dire les personnes décédése après avoir atteint 100 ans. Le tableau I contient, pour la Suisse et pour chaque année de 1879 à 1930, le nombre $n$ des centenaires. Ces valeurs ont été mises à ma disposition par le Bureau Fédéral de Statistique.[1])

---

[3]) Eine Formel für die Sterbenswahrscheinlichkeiten, die der Gleichung (19) entspricht, findet man bei Vajda: Über Wahrscheinlichkeiten in der Theorie der Versicherung verbundener Leben. (Assekuranz Jahrbuch, Bd. 54, Seite 43.)

[1]) Je profite de cette occasion pour le remercier.

## I. Nombre des centenaires, Suisse 1879—1930.

| Année | Hommes | Femmes | Têtes |
|---|---|---|---|
| 1879 | 1 | — | 1 |
| 80 | 1 | 1 | 2 |
| 81 | 1 | 1 | 2 |
| 82 | — | 3 | 3 |
| 83 | 1 | 1 | 2 |
| 85 | 2 | 1 | 3 |
| 88 | — | 1 | 1 |
| 91 | 1 | 1 | 2 |
| 92 | 1 | 1 | 2 |
| 93 | — | 1 | 1 |
| 95 | 1 | — | 1 |
| 97 | 1 | 2 | 3 |
| 98 | 1 | — | 1 |
| 99 | — | 2 | 2 |
| 1901 | — | 1 | 1 |
| 02 | — | 1 | 1 |
| 03 | — | 1 | 1 |
| 04 | 4 | 1 | 5 |
| 06 | 1 | — | 1 |
| 07 | 2 | 2 | 4 |
| 08 | — | 2 | 2 |
| 09 | — | 1 | 1 |
| 1910 | 1 | — | 1 |
| 12 | 1 | — | 1 |
| 13 | 1 | 2 | 3 |
| 14 | 2 | 4 | 6 |
| 15 | — | 2 | 2 |
| 16 | — | 2 | 2 |
| 17 | — | 1 | 1 |
| 18 | — | 1 | 1 |
| 19 | 1 | 1 | 2 |
| 20 | 1 | 2 | 3 |
| 21 | 1 | — | 1 |
| 22 | — | 2 | 2 |
| 23 | 1 | — | 1 |
| 24 | 2 | 3 | 5 |
| 25 | 1 | 2 | 3 |
| 26 | 1 | 2 | 3 |
| 27 | — | 1 | 1 |
| 28 | — | 3 | 3 |
| 29 | 2 | 3 | 5 |
| 30 | — | 1 | 1 |
| | 33 | 56 | 89 |

Dans de nombreuses années, il n'y eut aucun centenaire. Pour simplifier, ces années n'ont pas été mises dans ce tableau. Le maximum pour les hommes se trouve dans l'année 1904, où ce nombre était 4. Pour les femmes, le même maximum se trouve dans l'année 1914, où il y avait six centenaires en tout, ce qui représente le maximum pour la population entière.

Il s'agit de rechercher la distribution théorique du nombre annuel des centenaires. Le tableau II contient les répartitions observées. Pour sa construction, on énumère le nombre d'années avec $n = 0, 1, 2, 3, 4$ centenaires. La première colonne contient les valeurs $n$, la deuxième et troisième le nombre d'années où il y avait $n$ centenaires hommes et femmes. La colonne 4 a trait à la population entière. Cette répartition dérive de la série originale par le même procédé que les colonnes 2 et 3. Enfin la colonne 5 contient le nombre d'années dans lesquelles soit $n$ hommes soit $n$ femmes sont morts centenaires. On l'obtient par la sommation des colonnes 2 et 3. Puisque le nombre d'années observées est 52 pour chaque sexe on a 104 observations pour les deux sexes réunis. Pour pouvoir tracer les quatre distributions sur la même échelle, ces nombres sont divisés par deux. Les cinq lignes suivantes du tableau II contiennent le nombre moyen des centenaires par an $\overline{n}$, le deuxième moment $\overline{n^2}$, la dispersion $\sigma^2$ et deux grandeurs qui seront expliquées tout à l'heure.

II. Répartitions observées des années à centenaires.

| Nombre des centenaires | Hommes | Femmes | Têtes | Hommes ou Femmes |
|---|---|---|---|---|
| 0 | 27 | 18 | 10 | 22,5 |
| 1 | 19 | 18 | 18 | 18,5 |
| 2 | 5 | 11 | 11 | 8 |
| 3 | 0 | 4 | 8 | 2 |
| 4 | 1 | 1 | 1 | 1 |
| 5 | 0 | 0 | 3 | 0 |
| 6 | 0 | 0 | 1 | 0 |
| $\overline{n}$ | 0,63462 | 1,07692 | 1,71154 | 0,85577 |
| $\overline{n^2}$ | 1,05769 | 2,19231 | 5,01923 | 1,62500 |
| $\sigma^2$ | 0,65495 | 1,03255 | 2,08986 | 0,89266 |
| $\sigma(\sigma^2)$ | 0,116 | 0,255 | 0,384 | 0,186 |
| $52e^{-\overline{n}}$ | 27,56712 | 17,71344 | 9,39055 | 22,09775 |

Le nombre moyen annuel des centenaires est plus élevé pour les femmes que pour les hommes. Autrement dit: il y a plus de femmes qui atteignent 100 ans, fait qui est dû à la supériorité générale des femmes du point de vue biométrique. En même temps la dispersion de la répartition qui a trait aux femmes, est plus élevée.

On pourrait s'étonner de ce que la dispersion de la cinquième colonne est un peu supérieure à la demi-somme des dispersions des deux premières populations (colonnes 2 et 3). Ce point mérite d'être éclairci quoique le théorême de l'addition des repartitions ne se rapporte pas directement à la mortalité.

Etablissons donc la relation entre les moyennes d'une répartition totale (colonne 5) et des deux répartitions composantes (colonnes 2 et 3), ayant le même poids. Pour deux répartitions $\varphi_1$ et $\varphi_2$ portant sur la même variable, la répartition totale $\varphi$ est d'après la condition de l'aire

$$\varphi = \frac{\varphi_1 + \varphi_2}{2}.$$

Désignons par $M_{r,1}$ et $M_{r,2}$ les $r^{\text{ièmes}}$ moments des deux répartitions calculées autour de l'origine. Alors le $r^{\text{ième}}$ moment $M_r$ de la distribution totale sera

$$M_r = \frac{M_{r,1} + M_{r,2}}{2}.$$

L'espérance mathématique de la distribution totale est la demi-somme des valeurs des deux distributions ajoutées, relation qui est vérifiée par les valeurs $\overline{n}$. Pour le deuxième moment on aura

$$M_2 = \frac{M_{2,1} + M_{2,2}}{2}.$$

On obtient la dispersion de la distribution totale par

$$\sigma^2 = M_2 - M_1^2.$$

Introduisons les dispersions des deux distributions

$$\sigma_1^2 = M_{2,1} - M_{1,1}^2$$
$$\sigma_2^2 = M_{2,2} - M_{1,2}^2$$

et l'on obtient

$$\sigma^2 = \tfrac{1}{2}(M_{2,1} - M_{1,1}^2) + \tfrac{1}{2}(M_{2,2} - M_{1,2}^2) + \tfrac{1}{4}(M_{1,1}^2 - 2M_{1,1}\,M_{1,2} + M_{1,2}^2) =$$
$$= \frac{\sigma_1^2 + \sigma_2^2}{2} + \left(\frac{M_{1,1} - M_{2,1}}{2}\right)^2.$$

La dispersion de la distribution totale sera donc plus grande que la demi-somme des dispersions, puisque les deux moyennes diffèrent. On constate aisément que les dispersions du tableau II verifient cette relation.

La théorie des distributions des années à centenaires est immédiate. Soit $l(x)$ la probabilité d'un décès après l'âge $x$, $1 - l(x)$ la probabilité contraire. Alors la probabilité pour que, parmi $N$ décédés d'une année, il y ait $n$ après l'âge $x$, est, d'après la formule de Bernoulli

$$w(n, x) = \binom{N}{n} l(x)^n (1 - l(x))^{N-n}.$$

Pour l'âge $x = 100$, la probabilité $l(x)$ est très petite tandis que $N$ est très grand. Le produit $N\,l(x)$ tend vers une valeur finie $\overline{n}$, le nombre moyen des centenaires par an. Alors la probabilité $w(n, x)$ tend vers la loi des événements rares établie par Poisson et dont la portée statistique a été prouvée par L. von Bortkiewicz

$$w(n) = \frac{\overline{n}^n e^{-\overline{n}}}{n!}.$$

Sur $k = 52$ années, on doit s'attendre à ce qu'il y ait eu $k\,w(n)$ années avec $n$ centenaires. Il existe un critère pour savoir si cette théorie est juste. Car pour cette loi on doit avoir $\overline{n} = \sigma^2$. Mais cette identité ne vaut que pour la distribution théorique qui suppose un nombre infini d'observations. Pour les répartitions observées, où le nombre d'observations est nécessairement fini, la moyenne arithmétique et la dispersion sont elles-mêmes des aléatoires qui dependent du nombre d'observations. Cette relation ne peut donc être vérifiée que d'une façon approximative.

Les dernières lignes du tableau II montrent un bon accord. pour la répartition des hommes et pour la catégorie hommes ou femmes, tandis que les déviations sont sensibles pour les femmes et même assez grandes pour l'ensemble. On peut donc prévoir que la théorie sera bien vérifiée par les observations pour les deux premières répartitions, tandis que l'accord sera moins bon pour les autres.

D'une façon numérique on observe que l'écart type $\sigma(\sigma^2)$ de la dispersion est pour la loi des événements rares

$$\sigma(\sigma^2) = \sqrt{\frac{\bar{n}(2\bar{n}+1)}{k}}$$

valeur contenue dans l'avant-dernière linne du tableau II. On constate que toutes les différences entre les moyennes $\bar{n}$ et les dispersions $\sigma^2$ sont inférieures aux écarts type des dispersions. C'est seulement pour la catégorie „Têtes" que cette différence atteint la valeur $\sigma(\sigma^2)$.

Les répartitions pour les femmes et pour la population entière possèdent une dominante pour $n = 1$ tandis que les deux autres diminuent toujours avec $n$. Cette différence s'explique facilement puisque la loi des événements rares possède une dominante distincte de zéro, si l'espérance mathématique est égale ou plus grande que 1 ce qui a lieu pour les répartitions des femmes et de la population entière.

Pour établir la distribution théorique, on calcule d'abord

$$k\,w(0) = ke^{-\bar{n}}$$

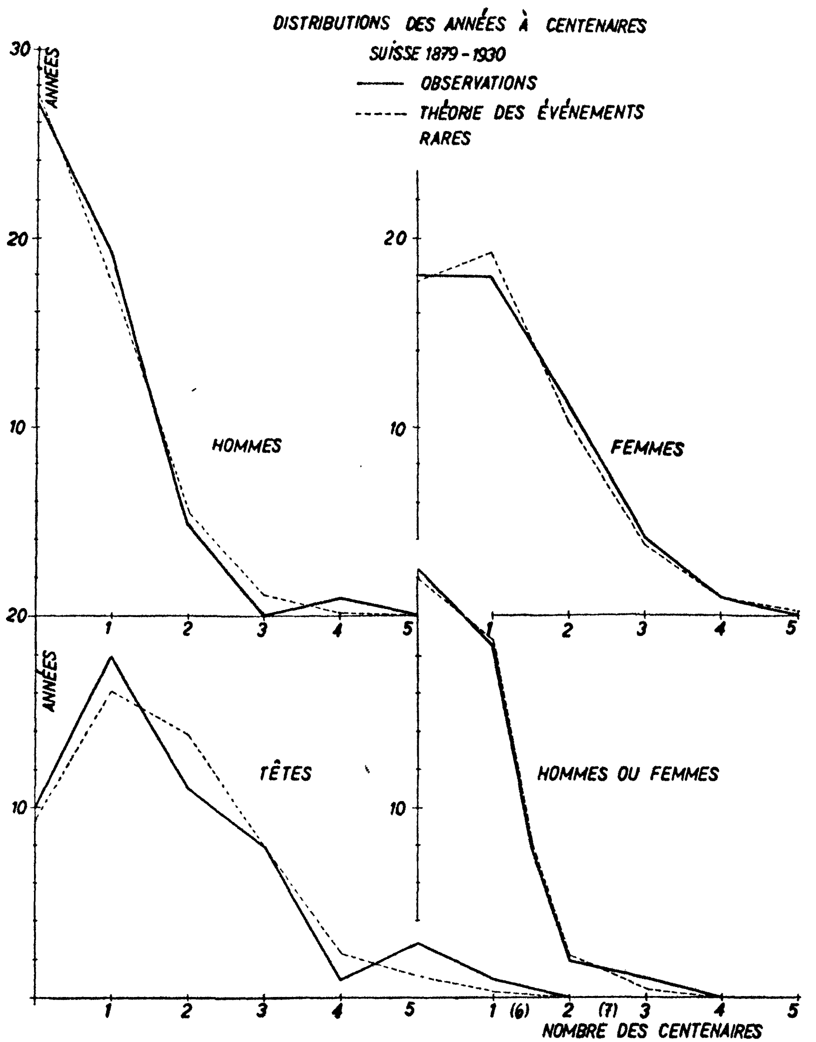
valeur contenue dans la dernière ligne du tableau II. On aura toutes les autres valeurs par la formule recurrente

$$k\,w(n) = k\,w(n-1)\,\frac{\bar{n}}{n}$$

dont on tire les séries théoriques contenues dans le tableau III. Chaque chiffre y est analogue au chiffre correspondant du tableau II.

III. Distributions théoriques d'années à centenaires Suisse 1879—1930.

| Nombre des centenaires | Hommes | Femmes | Têtes | Hommes ou Femmes |
|---|---|---|---|---|
| 0 | 27,57 | 17,71 | 9,39 | 22,10 |
| 1 | 17,50 | 19,08 | 16,07 | 18,91 |
| 2 | 5,55 | 10,27 | 13,75 | 8,09 |
| 3 | 1,17 | 3,69 | 7,85 | 2,31 |
| 4 | 0,19 | 0,99 | 3,36 | 0,49 |
| 5 | 0,02 | 0,21 | 1,15 | 0,09 |
| 6 | — | 0,04 | 0,33 | 0,01 |
| 7 | — | 0,01 | 0,08 | — |
| 8 | — | — | 0,02 | — |




Les figures montrent que les distributions théoriques s'accordent fort bien avec les observations. Le nombre de ceux qui survivent à 100 ans suit la loi des événements rares. Pour compléter cette étude, recherchons si les nombres d'années $k(n)$, où il y avait le même nombre $n = 0, 1, 2, 3$ d'hommes et de femmes centenaires, sont bien tels qu'il faut les attendre.

IV. Nombres d'années ayant le même nombre de centenaires pour les deux sexes.

| $n$ | $k(n)$ observé | $k(n)$ théoriqiue |
|---|---|---|
| 0 | 10 | 9,39 |
| 1 | 6 | 6,42 |
| 2 | 1 | 1,10 |
| 3 | 0 | 0,08 |
| | 17 | 16,99 |

Le tableau IV compare les valeurs $k(n)$ observés (colonne 2) et leur somme avec les valeurs théoriques (colonne 3) et leur somme. On obtient ces grandeurs en divisant les produits de la deuxième et troisième colonne du tableau III par 52. Cet accord aussi ne laisse rien à désirer.

On peut se demander enfin, quelle est la relation entre cette théorie et la théorie du plus grand âge établie par l'auteur dans quelques travaux précédents.[1]) L'une et l'autre sont des distributions finales où l'on suppose que le nombre d'observations $N$, à savoir des décès, augmente sans limite. De même, on étudie les âges pour lesquels la probabilité de survivre tend vers zéro. Aussi dans les deux problèmes ne choisit-on pas une formule analytique pour le cours de la mortalité. Mais en regard de ces accords il y a des différences. La distribution des centenaires part de la probabilité discontinue pour qu'on ait $n$ décès à des âges superieurs à $x$ parmi $N$ décès, et l'on emploie les conditions limites $x = 100$; $N\,l(x) \rightarrow \bar{n}$ pour arriver à une distribution finale portant sur une variable discontinue, le nombre d'années.

La distribution finale du plus grand âge part de la densité de probabilité $\mathfrak{w}_N(x)$ pour que $x$ soit le plus grand âge au décès parmi $N$, à savoir

$$\mathfrak{w}_N(x) = -N(1 - l(x))^{N-1}\, d\, l(x).$$

On applique les conditions limites $x \rightarrow \infty$; $N\,l(x) \sim 1$ et l'on arrive à la distribution portant sur une variable continue, un âge

$$\mathfrak{w}(x) = \mu(\tilde{\omega})\, e^{-\mu(\tilde{\omega})(x-\tilde{\omega}) - e^{-\mu(\tilde{\omega})(x-\tilde{\omega})}}$$

où $\tilde{\omega}$ est le plus grand âge dominant, et $\mu(\tilde{\omega})$ son intensité de mortalité.

---

[1]) Les plus grands âges en Suisse. Journal de Statistique et Revue économique suisse. 70e année, fasc. 4, Bâle 1934.

Les piu alte eta in Suezia. Giornale dell'Istituto Italiano degli Attuari, Anno VI, No 4. Rome 1935.

L'âge limite. Aktuárské Vědy T VI, No 1, 1936.

La durée extrême de la vie humaine. Actualités scientifiques et nd ustrielles, Hermann & Cie., Paris 1937.

Les résultats atteints par les deux théories montrent que les observations portant sur les âges élevés suivent bien les règles du calcul des probabilités pourvu que ces observations soient bonnes et suffisamment nombreuses, et que la théorie soit bien établie.

---

# Diskuse o pensijním pojištění.

Dr. *Antonín Zelenka.*

V č. 185 Pojistného Obzoru byla uveřejněna neslušná a nevěcná polemika anonyma, přikládajícího si titul a jméno „Pojistný technik M. Fitz" s mým článkem „Diskuse o pensijním pojištění", kterým jsem uváděl na pravou míru jeho článek v č. 183 uvedeného časopisu. Ačkoliv tón p. Fitze předem mne zbavoval povinnosti odpověděti, neboť tomu pánovi nejde o věcné diskutování, ale o pomluvy, přece jen jsem vzhledem k závažnosti tématu zaslal redakci Pojistného Obzoru další článek. Redakce byla ochotna tak učiniti; protože ale před uveřejněním článku dala jej p. Fitzovi k přečtení a vyžádala si ihned jeho odpověď, hodlajíc ji uveřejniti současně s mým článkem, projevil jsem nesouhlas s jejím postupem, který nebyl nestranný a požádal jsem, aby můj článek neuveřejnila s tím, že požádám o jeho publikaci redakci „Aktuárských Věd". *A. Z.*

Můj článek v č. 183 Pojistného obzoru, kterým jsem uvedl na pravou míru vývody pojistného technika M. Fitze v čísle předcházejícím, byl podnětem k jeho odpovědi v č. 184 P. O. Odpověď po meritorní stránce nepřináší nic nového. Psána je však způsobem, který nesvědčí o dobré vůli diskusí dospěti k vyjasnění věci; odpovídám-li přece, je to jen proto, aby u čtenářů méně informovaných nevznikl snad dojem, že jeho kritika je správná. Těm, jimž jde o zjištění pravdy, postačí srovnání a rozbor všech tří článků, zejména v bodech, které obsahují konkrétní fakta. Pouze v takových bodech je možná polemika odborné úrovně. Řečnické otázky snad obvyklé v útocích neodpovědného tisku nejsou přesvědčivým argumentem, zejména jde-li o anonyma. Přes všechnu mou snahu se mi totiž nepodařilo zjistiti osobnost pana pojistného technika M. Fitze a proto soudím, že jde o anonyma, který patrně ze skromnosti si přeje zůstati nepoznán.

K jednotlivým bodům připomínám toto:

1. Pan Fitz tvrdí, že ve svém prvém článku napsal v podstatě, že V. P. Ú. neuveřejňuje běžně a pravidelně statistiky o frekvenci