Aktuárské vědy

E. J. Gumbel

La plus grande valeur. I

Aktuárské vědy, Vol. 5 (1935), No. 2, 83–89

Persistent URL: http://dml.cz/dmlcz/144628

Remarque générale.

Au point de vue plus générale, notre méthode A montre qu'il ne suffit pas — en appliquant le procédé d'itération — de donner à l'équation à résoudre la forme $i = \varphi(i)$, et de prouver la convergence, c'est aussi la distribution de la fonction $\varphi(i)$ qui joue un rôle important, parfois même prépondérant.

La plus grande valeur.

E. J. Gumbel, Université de Lyon.

1. La distribution finale de la plus grande valeur.
2. Calcul des moments.
3. Application à quelques distributions initiales.

Pour une distribution non limitée et pour un nombre fini d'observations, il y en aura certainement au moins une qui sera la plus grande. Sa valeur sera finie puisque nous ne pouvons observer que des valeurs finies. Il s'agit de préciser cette valeur et la manière dont elle dépend du nombre des observations.

Des parties importantes de ce problème sont déjà résolues. Car il est reconnu que cette valeur a elle-même le caractère d'une variable statistique, c'est à dire qu'il existe une distribution de la plus grande valeur qui dépend du nombre des observations et de la distribution dite initiale que l'on traite. M. von Bortkiewicz[1]) a calculé son espérance mathématique pour la distribution initiale de Gauss et pour un petit nombre d'observations. M. L. H. C. Tippett,[2]) qui s'est limité aux mêmes conditions, a calculé les premiers moments. M. von Mises[3]) a donné une approximation de l'espérance mathématique valable pour un grand nombre d'observations et pour une répartition quelconque satisfaisant à une condition spéciale. M. Tricomi[4]) a calculé directement l'espérance

[1]) L. von Bortkiewicz: Variationsbreite und mittlerer Fehler. Sitzungsberichte der Berliner mathematischen Gesellschaft, 21. Jahrg. No. 1, 1922.
Die Variationsbreite beim Gaussschen Fehlergesetz. Nordisk Statistisk Tidskrift Bd. I, 1922.

[2]) L. H. C. Tippett: On the extreme individuals and the range of samples taken from a normal population. Biometrika, vol. XVII, parts 3 and 4, Cambridge 1925.

[3]) R. von Mises: Über die Variationsbreite einer Beobachtungsreihe. Sitzungsberichte der Berliner Mathematischen Gesellschaft 5. Okt. 1922.

[4]) F. Tricomi: Determinazione del valore asintotico di un certo integrale. Rendiconti della R. Accademia nazionale dei Lincei, vol. XVII, ser. 6a, 1. sem. fasc. 2, Roma 1933.

mathématique pour la loi de Gauss et pour un grand nombre d'observations. C'est pour les mêmes conditions que M. de Finetti[5]) a établi des tables numériques de la valeur centrale. M. R. A. Fisher[6]) a calculé les premiers moments en employant la notion féconde de la distribution finale. Il y est arrivé par la définition très restreinte de la plus grande valeur comme fonction linéaire de la variable statistique elle-même.

Un travail important est dû à M. E. L. Dodd[7]) qui a calculé l'espérance mathématique de la dernière valeur pour les distributions initiales de Gauss et sa transformation logarithmique, pour les distributions de Makeham et de Pearson, et pour la loi des événements rares. Enfin M. M. Fréchet,[8]) antérieurement à M. R. A. Fisher, a cherché la distribution finale stable, qui reproduit la distribution initiale en partant des valeurs de la variable rapportée à une valeur moyenne. Ce procédé mène à une distribution finale pour laquelle les moments d'ordre élevé divergent.

Dans ce qui suit nous allons systématiser et élargir ces résultats et établir des conditions moins rigides qui permettront des applications à maintes distributions initiales. Dans ce but nous commencerons par la distribution de la plus grande valeur pour un nombre quelconque d'observations, puis nous établirons sa forme finale pour un très grand nombre d'observations par le développement usuel. Enfin nous allons calculer tous ces moments à l'aide d'une formule de récurrence qui permettra des applications simples à des distributions initiales données.

Dans ce but nous allons nous limiter à des distributions initiales continues, telles que tous les moments existent.

1. La distribution finale de la plus grande valeur.

Soit $w(x)$ la distribution initiale d'une variable statistique x. C'est à dire la probabilité d'une valeur comprise entre x et $x + dx$ est $w(x)\,dx$. Nous la supposons limitée à gauche par une valeur $x = i$ que nous laissons indéterminée et illimitée vers la droite de la sorte qu'elle s'approche asymptotiquement de zéro pour des grandes valeurs de la variable. Cette supposition est nécessaire puisque la distribution

---

[5]) B. de Finetti: Sulla legge di probabilità degli estremi. Metron, vol. IX. N. 3—4, Roma 1932.

[6]) R. A. Fisher: Limiting forms of the Frequency distribution of the largest or smallest member of a sample. Proc. Cambridge Philos. Soc. Vol. XXIV, Part. 2, 1928.

[7]) E. L. Dodd: The greatest and the least variate under general laws of error. Transactions of the American Mathematical Society, vol. 25, No. 4, 1923.

[8]) M. Fréchet: Sur la loi de probabilité de l'écart maximum. Annales de la Société polonaise de Mathématique. T. VI, 1927.

doit satisfaire à la condition dite de l'aire

$$\int_i^\infty w(z)\,\mathrm{d}z = 1.$$

La probabilité d'une valeur inférieure à x sera

$$W(x) = \int_i^x w(z)\,\mathrm{d}z. \tag{1}$$

C'est une fonction monotone et croissante de x, qui remplit les conditions initiales $W(i) = 0$; $W(\infty) = 1$.

La probabilité $\mathfrak{W}(x)$ pour que, parmi N observations ou échantillons, aucune ne dépasse une valeur x, sera

$$\mathfrak{W}(x) = W^N(x). \tag{2}$$

Donc la densité de probabilité de la plus grande valeur sera

$$\mathfrak{w}(x) = \frac{\mathrm{d}}{\mathrm{d}x} W^N(x) \tag{3}$$

$$= N W^{N-1}(x)\, w(x). \tag{4}$$

Pour N observations et pour une distribution initiale illimitée $w(x)$ la plus grande valeur est elle-même une variable statistique dont la distribution $\mathfrak{w}(x)$ dépend de la distribution initiale $w(x)$ et du nombre N des observations. Pour fixer les idées, calculons, d'abord les trois moyennes de cette distribution, c'est à dire la valeur dominante, la valeur centrale et l'espérance mathématique.

La valeur dominante u de la distribution (3), c'est à dire la plus grande valeur qui a le maximum de probabilité, sera calculée par le procédé usuel. C'est la solution $x = u$ de l'équation

$$\mathfrak{w}'(x) = 0$$

ou

$$\frac{\mathfrak{w}'(x)}{\mathfrak{w}(x)} = 0.$$

On aura à l'aide de (4)

$$\frac{N-1}{W(x)}\, w(x) + \frac{w'(x)}{w(x)} = 0.$$

Cette équation, qui a la forme d'une équation différentielle nous interesse seulement pour la valeur $x = u$ de la variable. Pour trouver une solution pour des grandes valeurs de N, traitons le quotient

$$\frac{w(x)}{1 - W(x)} = \alpha. \tag{5}$$

Aux grandes valeurs de N correspondront de grandes valeurs de x. Car pour un nombre illimité d'observations la plus grande valeur dépassera toute grandeur donnée. Donc le dénominateur et le numérateur tendront vers zéro et le quotient sera indéterminé. D'après la règle de L'Hôpital on aura

$$\lim_{N=\infty} \frac{w(x)}{1 - W(x)} = - \frac{w'(x)}{w(x)}. \tag{6}$$

En supposant que N soit assez grand, la valeur u sera aussi grande, de telle sorte qu'il sera légitime d'employer cette équation pour le calcul de la dominante. Elle est d'abord une condition sur la nature de la distribution initiale $w(x)$, puis sur le nombre N des observations faites. Donc on aura, en introduisant (6) dans l'équation qui précède (5),

$$\frac{N-1}{W(x)} w(x) = \frac{w(x)}{1 - W(x)}$$

où $w(x)$ sera très petit, mais positif. L'équation donne

$$N - 1 = N W(x)$$

c'est à dire

$$W(u) = 1 - \frac{1}{N} \tag{7}$$

équation dont on peut calculer la dominante de la plus grande valeur à l'aide de (1), pourvu que N soit suffisamment grand de façon à justifier l'égalité (6).

La densité de probabilité de la dominante de la plus grande valeur sera d'après (4) et (7)

$$\mathfrak{w}(u) = \frac{N w(u)}{e}. \tag{8}$$

Pour des valeurs croissantes de N les distributions de la dernière valeur $\mathfrak{w}(x, N)$ forment un système de courbes avec des valeurs consécutives croissantes de dominantes $u_N, u_{N+1}, \ldots$

La distribution de la dernière valeur pour $N + 1$ observations sera d'après (4)

$$\mathfrak{w}(x, N + 1) = \frac{N+1}{N} W(x)\, \mathfrak{w}(x, N).$$

On aura donc

$$\mathfrak{w}(x, N + 1) \gtreqless \mathfrak{w}(x, N), \tag{9}$$

si

$$W(x) \gtreqless \frac{N}{N+1} \approx 1 - \frac{1}{N},$$

c'est à dire si

$$x \gtreqless u_N \tag{9'}$$

Pour des valeurs a) antérieures, b) postérieures à la dominante de la dernière valeur pour N observations, la densité de probabilité qu'une valeur x soit la dernière pour N observations est a) plus grande, b) moindre que la densité de probabilité qu'elle le soit pour $N + 1$ observations. La courbe de la distribution de la dernière valeur pour $N + 1$ observations coupe celle de la distribution pour N observations à sa dominante.

Calculons en outre la médiane v ou valeur centrale de la distribution (4) qui sera définie par

$$W^N(v) = \tfrac{1}{2}.$$

Donc

$$W(v) = e^{-\frac{1}{N}\lg 2}.$$

Pour les valeurs suffisamment grandes de N, on obtient la médiane de la plus grande valeur par

$$W(v) = 1 - \frac{1}{N}\lg 2. \tag{10}$$

La médiane de la plus grande valeur sera donc supérieure à la dominante. Mais la différence disparaît pour des valeurs croissantes de N.

La dominante et la médiane de la dernière valeur augmenteront avec le nombre des observations. Car, en différentiant (7) par rapport à N, on aura

$$w(u)\frac{\mathrm{d}u}{\mathrm{d}N} = \frac{1}{N^2},$$

c'est à dire

$$\frac{\mathrm{d}u}{\mathrm{d}N} = \frac{1}{N^2 w(u)}$$

ce qui est toujours positif. En stricte analogie on aura

$$\frac{\mathrm{d}v}{\mathrm{d}N} = \frac{\lg 2}{N^2 w(v)}.$$

Mais en outre du nombre des observations ces valeurs dépendront des constantes qui existent dans la distribution initiale.

On peut facilement préciser la manière dont la plus grande valeur dépend du nombre d'observations. En introduisant (7) dans l'équation (5) on aura pour la valeur dominante

$$\alpha = Nw(u). \tag{5'}$$

En introduisant cette valeur dans l'équation traitée à l'instant on voit que la dominante de la dernière valeur est liée au logarithme du nombre d'observations par une relation très simple. Car on obtient

$$\frac{\mathrm{d}u}{\mathrm{d}N} = \frac{1}{N\alpha}.$$

Si α est indépendant de N ou tend, soit en augmentant, soit en diminuant, pour des valeurs croissantes de N vers une constante $1/c$, on aura après intégration

$$u = \xi + c \log N \tag{7'}$$

où ξ est la dominante de la distribution initiale et c joue le rôle d'une mesure des écarts. Car pour une seule observation la distribution $w(x)$ est d'après (4) identique à la distribution initiale $w(x)$. Pour N observations la plus grande valeur calculée à partit de dominante initiale ξ sera un multiple d'une mesure des écarts. Dans ce cas α sera la valeur inverse d'une mesure des écarts.

Si au contraire α augmente indéfiniment avec N, la dominante de la plus grande valeur augmentera plus lentement que le logarithme de N, de sorte que l'on pourra écrire

$$u = \xi + c\,f(\lg N). \tag{7''}$$

C'est de ces propriétés de α que dépend aussi la question de savoir, si la distribution de la plus grande valeur se resserre pour un nombre croissant d'observations. En effet si α est indépendant de N ou tend vers une constante, il resulte de (8) que la densité de probabilité de la dominante devient indépendante du nombre d'observations. Donc la distribution de la plus grande valeur ne change pas de forme avec N. Si au contraire α augmente indéfiniment avec N, la distribution finale de la plus grande valeur se resserre pour un nombre croissant d'observations. Le cas où α diminue indéfiniment avec N sans qu'il n'existe une limite ne se présentera pas pourvu que nous fassions un choix raisonnable parmi les distributions initiales. Considérons une variable positive, et limitons nous aux grandes valeurs. Alors chaque valeur de x pourra être traitée comme une valeur dominante pour un nombre convenablement choisi N. La probabilité totale sera d'après (5)

$$W(x) = 1 - e^{-\int_0^x \alpha\, dz}.$$

Or, les moments d'ordre élevé de la distribution initiale n'existent qu'à condition que $1 - W(x)$ tend plus vite vers zéro que toute puissance négative de x. Supposons maintenant que α diminue indéfiniment avec N, c'est-à-dire avec x et posons

$$\alpha \to \frac{k}{x^n},$$

où k et n seront des constantes positives, alors les moments supérieurs divergeront. Puisque nous nous limitons aux distributions initiales telles que tous les moments existent, il n'y aura que deux cas: ou α tend vers une constante ou α augmente indéfiniment. Ces deux cas décideront des propriétés de la plus grande valeur pour savoir si elle augmente comme le logarithme du nombre d'observations ou plus lentement.

Pour l'espérance mathématique de la plus grande valeur, M. von Mises[3]) a prouvé qu'elle tend vers la dominante pour un nombre croissant d'observations. Il demande qu'on ait pour une valeur fixe et positive de k

$$\frac{1 - W(x+k)}{1 - W(x)} \to 0. \tag{11}$$

Si cette condition est remplie l'espérance mathématique de la plus grande valeur peut être calculée en première approximation par la formule simple (7). En outre la distribution de la plus grande valeur se resserre alors autour de l'espérance mathématique pour un nombre croissant d'observations. Le sens analytique de cette condition est que la distribution initiale tend plus vite vers zéro qu'une fonction exponentielle.

M. de Finetti[5]) a prouvé ce théorème pour la valeur centrale. En suivant sa méthode nous allons prouver que la distribution se resserre autour de la dominante, ce qui n'est qu'une autre forme d'énoncer le théorème de Mises. Dans ce but, la probabilité pour que la plus grande valeur soit moindre que $u + \varepsilon$ (ε positif) doit tendre vers 1 et la probabilité pour que la plus grande valeur soit moindre que $u - \varepsilon$ doit tendre vers zéro pour un nombre croissant d'observations. Donc nous demandons

$\mathfrak{W}(u+\varepsilon) \to 1$ ou $\lg \mathfrak{W}(u+\varepsilon) \to -0$ ou d'après (2) $N \lg W(u+\varepsilon) \to -0$

et

$\mathfrak{W}(u-\varepsilon) \to 0$ ou $\lg \mathfrak{W}(u-\varepsilon) \to -\infty$ ou d'après (2) $N \lg W(u-\varepsilon) \to -\infty$.

En introduisant (7) ces postulats seront

$$\frac{\lg W(u+\varepsilon)}{1 - W(u)} \to -0; \quad \frac{\lg W(u-\varepsilon)}{1 - W(u)} \to -\infty.$$

Les valeurs de ces quotients indéterminés seront d'après L'Hôpital

$$-\frac{w(u+\varepsilon)}{W(u+\varepsilon)\, w(u)} \quad \text{et} \quad -\frac{w(u-\varepsilon)}{W(u-\varepsilon)\, w(u)}.$$

Puisque les probabilités $W(u+\varepsilon)$ et $W(u-\varepsilon)$ tendent vers 1, la distribution de la plus grande valeur se resserre autour de la dominante si

$$\frac{w(u+\varepsilon)}{w(u)} \to 0$$

ce qui implique

$$\frac{w(u-\varepsilon)}{w(u)} \to \infty.$$

Mais pour les grandes valeurs de x il y a certainement un nombre d'observations N, tel que $x = u$. Donc la condition est tout simplement

$$\frac{w(x+\varepsilon)}{w(x)} \to 0 \tag{12}$$

ce qui est identique à la condition de Mises. (La suite.)