Aktuárské vědy

Maurice Fréchet

Sur les précisions comparées de la moyenne et de la médiane

Aktuárské vědy, Vol. 5 (1935), No. 1, 29–34

Persistent URL: http://dml.cz/dmlcz/144620

der letztere gleich $(1-\gamma_i)\,U_i$ angesetzt, ist die quadratische Gleichung (39a)

$$0=(T_i^2-S_i^2+U_i^2)\,\mathfrak{w}_i^2-(T_i^2-S_i^2+2\gamma_i U_i^2)\,\mathfrak{w}_i+\gamma_i^2 U^2 \quad (58)$$

aufzulösen und, jenach $\gamma_i>\frac{1}{2}$ oder $\gamma_i<\frac{1}{2}$, die kleinere Wurzel $\mathfrak{w}_i$ oder die größere $\overline{\mathfrak{w}}_i$ zu wählen, um den Sicherheitskoeffizienten Θ_i des Vertrages zu bestimmen

$$\Theta_i=S_i/R_i^*,\quad R_i^*=U_i\,(\gamma_i-\mathfrak{w}_i)-S_i\mathfrak{w}_i \text{ bei } \gamma_i>\tfrac{1}{2}$$
$$R_i^*=U_i\,(\overline{\mathfrak{w}}_i-\gamma_i)-S_i\,(1-\overline{\mathfrak{w}}_i) \text{ bei } \gamma_i<\tfrac{1}{2}. \quad (59)$$

Einen mittleren Sicherheitskoeffizienten Θ findet man wieder nach (57), aber die Berechnung der Hauptkonstanten (56a) erfordert noch für die dort auftretende Größe γ die Verwendung eines Mittelwertes aus den γ_i der einzelnen Verträge

$$\gamma=\frac{\Sigma\gamma_i^* U_i}{\Sigma U_i},\quad \gamma_i^*=\gamma_i \text{ oder } 1-\gamma_i \text{ jenach } \gamma_i>\tfrac{1}{2} \text{ oder } \gamma_i<\tfrac{1}{2}. \quad (59a)$$

Bei einer anderen Art der Durchschnittsrechnung wäre von den Werten $R_i S_i T_i \gamma_i$ der einzelnen Verträge auszugehen und für jene Verträge, die nicht von einer einzigen Wahrscheinlichkeit abhängen, die entsprechenden Hoffnungswerte $\mathfrak{U}_i$ gemäß (43) zu rechnen, dann bietet der Betrag $\varepsilon=\Sigma S_i/\Sigma\mathfrak{U}_i$ einen geeigneten Mittelwert für das relative Überentgelt der Vertragskombination (Für Verträge einfachster Art wird $\mathfrak{U}_i$ mit U_i identisch.) Anderseits ist als durchschnittlicher Sicherheitskoeffizient, analog (57),

$$\vartheta=\frac{\Sigma S_i}{\Sigma R_i^*} \text{ oder auch } \vartheta=\frac{\sqrt{\Sigma S_i^2}}{\sqrt{\Sigma R_i^2}}$$

zu nehmen, endlich für γ der Durchschnitt

$$\gamma=\frac{\Sigma\gamma_i^*\,\mathfrak{U}_i}{\Sigma\,\mathfrak{U}_i}\left\{\begin{array}{l}\gamma_i^*=\gamma_i \text{ für } R_i=\mathfrak{U}_i\,(\gamma_i-\mathfrak{w}_i)-S_i\,\mathfrak{w}_i\\ \gamma_i^*=1-\gamma_i \text{ für } R_i=\mathfrak{U}_i\,(\mathfrak{w}_i-\gamma_i)-S_i\,(1-\mathfrak{w}_i).\end{array}\right.$$

(Fortsetzung folgt.)

Sur les précisions comparées de la moyenne et de la médiane.

Par *Maurice Fréchet* (Université de Paris).

Dans bien des cas, la valeur moyenne est à préférer à la médiane comme valeur typique d'un ensemble de valeurs. Mais il n'en est pas toujours ainsi: dans d'autres cas, la médiane fournit une valeur plus représentative (vie médiane, revenu médian) ou plus facile à déterminer numériquement.

Cependant certains statisticiens opposent cette objection au choix de la médiane, que sa précision serait moindre. Nous voulons montrer que, selon les lois de probabilités considérées, l'avantage de la précision peut aller à la médiane comme à la moyenne.

Considérons une variable aléatoire X et soit $F(x)$ sa fonction des probabilités totales:

$$F(x) = \text{Probab. } \{X < x\}.$$

La valeur moyenne, v, de X sera

$$v = \int_{-\infty}^{+\infty} x \, \mathrm{d}F(x) \tag{1}$$

et la valeur médiane, m, de X sera solution de

$$F(m) = \tfrac{1}{2}. \tag{2}$$

Pour déterminer empiriquement v et m, on peut substituer à v et m la moyenne arithmétique V et la médiane M d'un nombre fini, N, de valeurs de X observées au cours d'épreuves indépendantes quelconques. On pourra estimer les erreurs à craindre sur v et m en calculant les écarts quadratiques moyens μ', μ'' de V et de M avec v et m respectivement. Tout se réduit à montrer qu'il existe des lois de probabilité pour lesquelles $\mu' > \mu''$. Nous en donnerons deux exemples, et, pour simplifier les calculs, nous prendrons le cas de $N = 3$: celui où V et M sont calculés au moyen des valeurs X_1, X_2, X_3, observées pour X au cours de trois épreuves.

On voit facilement que si

$$P(x) = \text{Probabilité } \{M < x\}$$

et si $F(x)$ est différentiable, on a

$$\mathrm{d}P(x) = 6\, F(x)\, [1 - F(x)]\, \mathrm{d}F(x). \tag{3}$$

Car $F\,[1 - F]\,\mathrm{d}F$ est la probabilité élémentaire pour que

$$X_1 \leqq x \leqq X_2 = M \leqq x + \mathrm{d}x \leqq X_3$$

et il y a six permutations de X_1, X_2, X_3.

D'où:

$$P(x) = F^2(x)\, [3 - 2\, F(x)]. \tag{4}$$

On vérifie d'abord que la valeur médiane, m, de la variable aléatoire X, étant telle que $F(m) = \frac{1}{2}$, est telle que $P(m) = \frac{1}{2}$; c'est donc aussi la valeur médiane de M.

D'autre part si l'on pose $v = \overline{X}$, on voit que

$$v = \tfrac{1}{3}\, (\overline{X}_1 + \overline{X}_2 + \overline{X}_3) = \overline{V};$$

de sorte que v est aussi la valeur moyenne de V. Ainsi se trouve justifié l'emploi de V et de M comme valeurs empiriques de v et de m.

Ceci étant, on sait que si μ est l'écart quadratique moyen de X à partir de v, on a $\mu' = \mu/\sqrt{3}$. Calculons μ''. Soient μ''_1, μ''_2 les écarts quadratiques moyens de M à partir de $\overline{M}$ et de zéro. On a

$$\mu''^2 = \mu''^2_1 + (\overline{M} - m)^2; \quad \mu''^2_2 = \mu''^2_1 + (\overline{M})^2$$

avec

$$\mu''^2_2 = \int_{-\infty}^{+\infty} x^2 \, \mathrm{d}P(x).$$

D'où:

$$\mu''^2 = \int_{-\infty}^{+\infty} x^2 \, \mathrm{d}P(x) + (\overline{M} - m)^2 - (\overline{M})^2. \tag{5}$$

On a de même:

$$\mu^2 = \int_{-\infty}^{+\infty} x^2 \, \mathrm{d}F(x) - v^2. \tag{6}$$

D'où:

$$\mu''^2 - \mu'^2 = \int_{-\infty}^{+\infty} x^2 \, \mathrm{d}[P(x) - \tfrac{1}{3}F(x)] + (\overline{M} - m)^2 - (\overline{M})^2 + \frac{v^2}{3},$$

$$\mu''^2 - \mu'^2 = \int_{-\infty}^{+\infty} x^2 \left\{6\, F(x)\, [1 - F(x)] - \tfrac{1}{3}\right\} \mathrm{d}F(x) + \\ + (\overline{M} - m)^2 - (M)^2 + \frac{v^2}{3}. \tag{7}$$

La formule se simplifie dans le cas où la loi de probabilité de X est symétrique, c'est à dire où

$$F(-x) = 1 - F(x);$$

alors $m = v = 0$. De plus, dans ce cas, la loi de probabilité de M sera aussi symétrique, de sorte que la valeur moyenne $\overline{M}$ de m est aussi nulle. Finalement, on a, dans ce cas particulier

$$\mu''^2 - \mu'^2 = \int_{-\infty}^{+\infty} x^2 \left\{6\, F(x)\, [1 - F(x)] - \tfrac{1}{3}\right\} \mathrm{d}F(x). \tag{8}$$

Appliquons les formules (7), (8) à nos deux exemples.

I. Afin qu'on ne puisse dire que les lois de probabilités pour lesquelles $\mu' > \mu''$ sont exceptionnelles, nous prendrons comme premier exemple le cas où X satisfait à la première loi de Laplace, c'est à dire où $\mathrm{d}F(x)$ est proportionnel à $e^{-|x|}\,\mathrm{d}x$. On a, en effet, constaté que la loi de probabilité des erreurs d'observation satisfait presque aussi bien à la première loi de Laplace qu'à la seconde, celle où $\mathrm{d}F(x)$ est proportionnel à $e^{-x^2}\,\mathrm{d}x$ (dans chacun des deux cas, en adoptant une unité de mesure convenable).

On a alors, dans notre premier exemple:

$$F(x) = \tfrac{1}{2}e^x \quad \text{pour} \quad x \leqq 0, \qquad F(x) = 1 - \tfrac{1}{2}e^{-x} \quad \text{pour} \quad x \geqq 0.$$

Comme, ici, la loi est symétrique, on peut appliquer la formule (8) ou plus simplement

$$\mu''^2 - \mu'^2 = 2\int_{-\infty}^{0} x^2 \{6\,F(x)\,[1 - F(x)] - \tfrac{1}{3}\}\,\mathrm{d}F(x),$$

soit ici

$$\mu''^2 - \mu''^2 = \int_{-\infty}^{0} x^2 \{\tfrac{3}{2}e^x\,[2 - e^x] - \tfrac{1}{3}\}\,e^x\,\mathrm{d}x =$$

$$= \int_{-\infty}^{0} \left[3x^2e^{2x} - \tfrac{3}{2}x^2e^{3x} - \frac{x^2\,e^x}{3}\right] \mathrm{d}x$$

et puisque

$$\int_{-\infty}^{0} x^2e^{kx}\,\mathrm{d}x = \frac{2}{k^3} \quad \text{pour} \quad k > 0,$$

on aura

$$\mu''^2 - \mu'^2 = -\tfrac{1}{36}$$

ou

$$\mu'^2 = \mu''^2 + \tfrac{1}{36}.$$

On a bien $\mu' > \mu''$.

II. En choisissant artificiellement une loi convenable, nous montrerons même qu'il peut arriver que l'écart quadratique moyen de V avec v soit infini alors que celui de M avec m est fini (m, v, $\overline{M}$ étant finis). Il suffit de prendre pour X la loi de probabilité (non symétrique) donnée par

$$F(x) = \begin{cases} 0 & \text{si } x \leqq 1 \\ 1 - x^{-\alpha} & \text{si } x \geqq 1 \end{cases}$$

avec $1 < \alpha \leqq 2$.

On a en effet dans ce cas, en vertu de (2),

$$m = \sqrt[\alpha]{2}$$

et en vertu de (1)

$$v = \int_{1}^{+\infty} x\alpha x^{-\alpha-1}\,\mathrm{d}x = \alpha\int_{1}^{+\infty} x^{-\alpha}\,\mathrm{d}x,$$

d'où puisque $\alpha \geqq 1$:

$$v = \frac{\alpha}{\alpha - 1}.$$

Enfin

$$\bar{M} = \int_{-\infty}^{+\infty} x \, \mathrm{d}P(x) = \int_{1}^{\infty} x \left(1 - x^{-\alpha}\right) x^{-\alpha} \left(\alpha x^{-\alpha-1}\right) \mathrm{d}x =$$

$$= \int_{1}^{+\infty} \left(\alpha x^{-2\alpha} - \alpha x^{-3\alpha}\right) \mathrm{d}x = \frac{\alpha^2}{(3\alpha - 1)(2\alpha - 1)}.$$

Ainsi dans les formules (5), (6): v, m, M sont finis. D'autre part, l'intégrale

$$\int_{-\infty}^{+\infty} x^2 \, \mathrm{d}P(x) = \int_{-\infty}^{+\infty} x^2 . \, 6F\,(1 - F)\, \mathrm{d}F = 6 \int_{1}^{\infty} x^2 \left(1 - x^{-\alpha}\right) x^{-\alpha} \, \alpha x^{-\alpha-1} \, \mathrm{d}x =$$

$$= 6\alpha \int_{1}^{\infty} \left(x^{1-2\alpha} - x^{1-3\alpha}\right) \mathrm{d}x = \frac{3\alpha^2}{(\alpha - 1)(3\alpha - 2)}$$

est finie, et alors en vertu de (5), μ'' est fini. Au contraire l'intégrale

$$\int_{-\infty}^{+\infty} x^2 \, \mathrm{d}F(x) = \int_{1}^{+\infty} \alpha x^{1-\alpha} \, \mathrm{d}x = \left[\frac{\alpha x^{2-\alpha}}{2 - \alpha}\right]_{1}^{+\infty}$$

qui est finie pour $\alpha > 2$, est infinie pour $\alpha \leqq 2$.

Ainsi, dans le cas considéré où $1 < \alpha \leqq 2$, μ est infini d'après (6) et par suite, $\mu' = \mu/\sqrt{3}$ est aussi infini.

Note. Après avoir rédigé l'article ci-dessus, nous retrouvons dans nos papiers une note qui illustre du point de vue pratique les considérations théoriques présentées ci-dessus.

Cette note[1]) concerne une série d'observations, réparties sur 24 jours à raison de 500 par jour et publiées par C. S. Peirce.

Ces observations avaient été sur ma demande soumises au calcul par un de mes élèves, M. Samama. Celui-ci avait representé les observations de chaque journée successivement au moyen de la première et de la seconde loi de Laplace. Il avait trouvé que la majorité des meilleures représentations s'effectuaient par la seconde loi de Laplace (souvent dite normale) mais qu'une proportion non négligeable restait à l'actif de la première loi.

Or, dans la note de Washington, on peut lire les lignes suivantes, au § 8: „The ordinary statement based on the normal law is that the determination of the median is 25% worse than that of the mean. A comparison of the standard deviations of the median and mean in columns (1) and (2) shows that for these observations the median is better

[1]) Note on C. S. Peirce's experimental discussion of the law of errors, by E. B. Wilson and M. M. Hilferty, Proc. Nat. Acad. of Sc., vol. 15, 1929, p. 124, Washington.

determined than the mean on 13 days, worse determined on 9 days and equally well determined on 2 days. Roughly speaking this means that mean and median are on the whole about equally well determined."

On observera d'ailleurs que notre étude concerne le cas de 3 observations et qu'il s'agit ici de la médiane et de la moyenne empirique calculées 24 fois pour 500 observations.

Zum Problem der Massenfabrikation.

Otomar Pankraz.

In jeder entwickelten Wirtschaft wirken verschiedene wirtschaftliche „Kräfte", die sich teilweise in ihren Wirkungen gegenseitig stören, sodaß manchmal als Resultat eine wirtschaftliche Erscheinung eintritt, die weder erwartet noch beabsichtigt war. Um diese unbeabsichtigte und gewöhnlich schädliche Wirkungen zu vermeiden oder zu milden, strebt jede Wirtschaft zur Planwirtschaft.

Ich will mich nicht mit der Frage der Definitionen des Wirtschaftsplanes befassen. Für unsere Zwecke genügt, wenn wir einige Punkte hervorheben, die in der Regel bei jedem Wirtschaftsplan erfüllt sind:

1. Man verlangt, daß es möglich ist entweder alle oder wenigstens einen größeren Teil der wirtschaftlichen Zahlen als Funktionen der zeitlichen Parameter auszudrücken.

2. In jedem Augenblicke resp. in jedem Zeitintervalle von gegebener Länge soll ein Gleichgewicht zwischen dem Güterverbrauch und der Güterproduktion stattfinden. Dabei muß man eine besondere Rücksicht auf das Bilden der Vorräte (Lager) nehmen.

3. Durch Einflüsse, die bewußt oder unbewußt bei dem Feststellen der Bedingungen des Gleichgewichtes [siehe Punkt 2] vernachlässigt wurden, können Abweichungen vom Gleichgewichtszustande entstehen.

4. Es wird die Bedeutung des menschlichen Willens betont und man läßt zu, daß dieser Wille bewußt und zweckmäßig die wirtschaftliche Tätigkeit als ein Ganzes lenken und so (wenigstens teilweise) die Entstehung der Kontraste und Unstimmigkeiten in dieser Tätigkeit vermeiden kann. Es bleibt dabei die Frage offen, inwieweit überhaupt dieses Eingreifen des menschlichen Willens möglich ist.

5. Infolge der Punkte sub 2. und 4. stehen bei der Herstellung eines wirtschaftlichen Planes die deduktiven Methoden im Vordergrund. Das bedeutet aber keine Abweichung oder sogar einen Gegensatz zu den induktiven Methoden, schon deswegen nicht, weil die Bestimmung der Hilfsgrößen, die das Gleichgewicht sub 2. erfordert, sich an empirischen Daten stützen muß.

Im folgenden will ich nur solche wirtschaftliche Erscheinungen ins Auge fassen, welche sich kvantitativ ausdrücken lassen; speziell will ich