Časopis pro pěstování matematiky a fysiky

Maurice Fréchet

Sur deux relations simples entre le «coefficient» de corrélation et le «rapport» de corrélation

Časopis pro pěstování matematiky a fysiky, Vol. 64 (1935), No. 6, 209--210

Persistent URL: `http://dml.cz/dmlcz/123623`

Terms of use:

© Union of Czech Mathematicians and Physicists, 1935

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

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

$$z = \frac{1}{m} \int_0^w \mathrm{f}(w)\, \mathrm{d}w; \tag{4}$$

cette forme très simple permet non seulement une analyse plus facile des propriétés des courbes de concentration, mais aussi la résolution de certains problèmes nouveaux, comme, par exemple, la recherche de la distribution initiale, correspondant à une courbe de concentration donnée. Soit, dans le cas d'une distribution continue

$$z = \psi(w) \tag{5}$$

l'équation d'une courbe de concentration; m étant une constante arbitraire, ou aura immédiatement

$$\mathrm{f}(w) = m\, \mathrm{d}\psi/\mathrm{d}w \tag{6}$$

La méthode d'inversion fut employée récemment, dans un ordre d'idées un peu différent, par M. Robert Schmidt, de Kiel (Allemagne) — v. The Annals of Mathematical Statistics, Vol. 5, N. 1 (March, 1934), pp. 30—72: Statistical Analysis of One-Dimensional Distributions.

Sur deux relations simples entre le «coefficient» de corrélation et le «rapport» de corrélation.

Maurice Fréchet, Paris.

Notre attention a été récemment attirée par M. Karel Tříska (de Prague) sur l'effet d'un changement d'échelle sur la valeur du coefficient de corrélation. A cette occasion, nous avons pu obtenir une formule qui est peut-être nouvelle.

Soient deux variables aléatoires X, Y et soit b_X la valeur moyenne de Y quand X est donné. Si b_X est une fonction toujours croissante (ou toujours décroissante) de X, on peut, par un changement d'échelle de X seul, transformer la courbe des moyennes b_X en une droite donnée quelconque non parallèle aux axes. Il suffit de prendre pour nouvelle variable $X_1 = Pb_x + Q$ où P, Q sont des nombres certains convenablement choisis. Soient η le „rapport" de corrélation de Pearson de Y à X avant cette transformation, R le „coefficient" de corrélation de Bravais-Galton, après ce changement d'échelle, c'est à dire entre Y et X_1: On a

$$\boxed{|R| = \eta} \tag{1}$$

On peut démontrer cette formule directement. Mais on peut aussi l'établir immédiatement en utilisant la formule que nous avons indiquée dans le premier volume de la Revue de l'Institut international de Statistique et qui lie: le „rapport“ η ci-dessus, le „coefficient“ de corrélation r de Y et X (c'est à dire sans changement d'échelle) et la valeur ϱ que prend r quand on y remplace chaque système de valeurs de X et de Y par X et b_X. On voit facilement que

$$\boxed{r = \varrho\eta} \tag{2}$$

Le changement d'échelle ci-dessus transforme respectivement r, ϱ en R et ± 1 sans changer η: la formule (2) devient (1).

La formule (2) met en évidence que la valeur de r dépend non seulement, comme η, du resserrement des points observés autour de la ligne des moyennes, c'est à dire de la rigueur de la dépendance, mais aussi, comme ϱ, de la forme de la ligne des moyennes, circonstance indifférente à la mesure de la dépendance.

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

Maurice Fréchet, Paris.

Soient X une variable aléatoire et

$$\mathrm{F}(x) = \text{Probabilité } \{X < x\};$$

soient v la valeur moyenne et m la valeur probable de X. Si l'on effectue, par exemple, trois épreuves indépendantes donnant pour X les valeurs X_1, X_2, X_3, on peut prendre pour valeurs empiriques de v et de m, la moyenne arithmétique V et la médiane M de X_1, X_2, X_3. Les ordres de grandeur des erreurs à craindre dans cette détermination peuvent être estimés au moyen des écarts quadratiques moyens μ' de V à partir de v, μ'' de M à partir de m. Contrairement à une opinion assez répandue, la détermination de la valeur médiane n'est pas nécessairement moins précise que celle de la valeur moyenne. C'est à dire qu'on peut citer des lois de probabilités $\mathrm{F}(x)$ pour lesquelles $\mu' > \mu''$. Tel est le cas de la première loi que Laplace avait proposée comme loi des erreurs d'observation, à savoir celle pour laquelle $d\mathrm{F}(x) = \frac{1}{2} e^{-|x|} dx$.

On peut même citer des lois pour lesquelles v, m, μ'' sont finis et μ' infini; il suffit de prendre, par exemple

$$\begin{aligned} &\mathrm{F}(x) = 0 \quad \text{pour } x \leqq 1 \\ &\mathrm{F}(x) = 1 - x^{-\alpha} \quad \text{pour } x \geqq 1 \\ &\text{avec} \quad 1 \leqq \alpha \leqq 2. \end{aligned}$$