S. Fogelson
Sur la méthode d'inversion en statistique mathématique

Časopis pro pěstování matematiky a fysiky, Vol. 64 (1935), No. 6, 207--209

Persistent URL: http://dml.cz/dmlcz/123556

Konstanty A_n se stanoví podle ortogononálních vlastností (6)

$$A_n = \frac{\sum_{x=0}^{\infty} \mathrm{H}(x)\, \mathrm{p}_n(x)}{S_{nn}}.$$

Označíme-li $m'_{(i)}$ faktoriální moment funkce $\mathrm{H}(x)$ i-tého řádu a $m_{(i)}$ moment funkce $\mathrm{f}(x)$, máme, použijeme-li rovnic (5) a (6),

$$A_n = \frac{\sum_{i=0}^{n} (-)^i \binom{n}{i} \frac{m'_{(i)}}{m_{(i)}}}{n!\,(1+d)^n}. \tag{10}$$

Stanovíme-li parametry h a d ve funkci $\mathrm{f}(x)$ podle Pearsonovy metody, totiž

$$h = m_{(1)}, \quad d = \frac{m_{(2)}}{m_{(1)}} - m_{(1)},$$

jest

$$A_1 = A_2 = 0.$$

Sur la méthode d'inversion en statistique mathématique.

S. Fogelson, Warszawa.

On sait, que la méthode la plus générale de description d'une distribution des fréquences (ou des probabilités) consiste en emploi de la fonction-somme (Summenfunktion, probability Integral), $\mathrm{V}(x)$, exprimant — selon le cas — la fréquence des valeurs observés ou la probabilité totale des valeurs d'une variable aléatoire, qui n'excèdent pas x. L'introduction de l'intégrale de Stieltjes permet alors d'analyser les distributions continues et discontinues par une méthode uniforme et générale.

Les propriétés de la fonction $\mathrm{V}(x)$, qui est une fonction non décroissante, et qui admet des limites finies 0 et 1 pour $x = \mp\infty$, permettent d'effectuer une inversion uniforme de cette fonction: en posant

$$w = \mathrm{V}(x) \tag{1}$$

on a inversement

$$x = \mathrm{f}(w), \tag{2}$$

f étant une fonction non décroissante, définie dans l'intervalle (0, 1). Les propriétés de cette fonction se déduisent facilement de celles de la fonction $\mathrm{V}(x)$. La fonction $\mathrm{f}(w)$ représente la distribution donnée des fréquences (ou des probabilités) aussi bien que la fonction $\mathrm{V}(x)$ et, dans un nombre de cas, elle se prête mieux aux différents calculs que cette dernière et facilite ainsi la résolution de certains problèmes. Ceci tient à ce que certaines grandeurs statistiques, qui s'expriment au moyen des intégrales de Stieltjes de la forme

$$\int_{-\infty}^{+\infty} \varphi(x)\, \mathrm{dV}(x)$$

deviennent, après la substitution donnée par la formule (2), des intégrales simples (parfois généralisés) de la forme

$$\int_0^1 \Phi(w)\, \mathrm{d}w.$$

Comme premier exemple de l'emploi de la méthode d'inversion citons le problème suivant: soit x une variable aléatoire et soient $x_1, x_2, \ldots, x_N$ les valeurs de cette variable, observés dans un grand nombre N des épreuves répétées, et cherchons l'espérance mathématique e_N de la valeur observée ξ telle, qu'une partie déterminée des valeurs observées — soit αN avec $0 < \alpha < 1$ — restent inférieures à ξ.

Ce problème est une généralisation évidente du problème des „partition values" d'une série statistique (médiane, quartiles, déciles etc.). Dans le cas le plus simple, où la distribution initiale est continue, $\alpha = p/q$ et $N \equiv q$ (mod 1), la distribution exacte des probabilités de ξ est donnée par la formule

$$\mathrm{dV}_N(\xi) = \frac{N!}{m!\, n!} \mathrm{V}^n(\xi)\, [1 - \mathrm{V}(x)]^m\, \mathrm{dV}(x)$$

avec $n = N\, p/q$ et $m = N - n - 1$. La méthode d'inversion, appliquée aux intégrales, exprimant e_N et $s_N{}^2$ (le carré de l'erreur moyenne), ramène ces intégrales à une forme remarquablement facile aux évaluations asymptotiques et permet ainsi à trouver des formules asymptotiques

$$\lim_{N \to \infty} e_N = \mathrm{f}(\alpha), \quad \lim_{N \to \infty} s_N{}^2 = \frac{c}{\sqrt{N}}. \tag{3}$$

Pour le cas de la médiane ($\alpha = \frac{1}{2}$) ces formules sont démontrées au moyen de la méthode d'inversion dans la „Revue Trimestrielle de Statistique publiée par l'Office Central de Statistique de la République Polonaise", 1930, fasc. 2.

La théorie des courbes de concentration peut servir comme un second exemple de la méthode d'inversion. En appelant m la valeur moyenne de la distribution initiale

$$m = \int_0^1 \mathrm{f}(w)\, \mathrm{d}w$$

on aura immédiatement l'équation de la courbe de concentration correspondante

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

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