G.-H. KNIBBS

Applications statistiques de la série de Fourier

Journal de la société statistique de Paris, tome 53 (1912), p. 409-410
<http://www.numdam.org/item?id=JSFS_1912__53__409_0>

VIII

VARIÉTÉ

APPLICATIONS STATISTIQUES DE LA SÉRIE DE FOURIER

(Illustrées par l'analyse des chiffres de mariage et de suicide et de la température)

Je résume brièvement un article paru dans le volume XLV du Journal de la Société royale des Nouvelle-Galles du Sud, dans lequel j'ai indiqué la technique de l'emploi de la série de Fourier dans les calculs statistiques. En général, il est possible pour cet usage de mettre la série sous une des deux formes suivantes :

$$y = f(x) \{ a_0 + \sum_{k=1}^{k=n} a_k \sin k (x + \alpha_k) \}$$

ou

$$y = f(x) \{ a_0 + \sum_{k=1}^{k=n} (a_k \sin kx + b_k \cos kx) \}$$

où y représente la valeur momentanée de la fonction. Les résultats statistiques périodiques sont ordinairement des valeurs de groupes, c'est-à-dire les intégrales de l'expression ci-dessus entre les limites qui répondent à la position de l'argument dans la période de fluctuation (par exemple, au commencement et à la fin des mois, quand l'année entière est représentée par la période 2π).

Généralement, on peut supposer que l'amplitude d'une fluctuation est proportionnée à la grandeur de y. Puisque, dans bien des résultats statistiques, l'élément séculaire de l'accroissement qui est représenté par $f(x)$ est presque indépendant de la fluctuation elle-même, il faut égaliser les valeurs de groupes en les rapportant à une valeur fixe, préférablement à celle du milieu de la période entière. Cette égalisation s'exécute avec l'aide des facteurs $f(x_m)$ et $f(x_k)$, où $f(x_m)$ représente la valeur momentanée à l'époque considérée, et $f(x_k)$ la valeur au milieu de la période mineure. Une solution peut alors se trouver en faisant usage des termes ci-dessus.

Généralement, il faut d'abord égaliser les mois ou les autres périodes de groupes et il est facile de trouver dans chaque cas la meilleure méthode pour arriver à un résultat pratique. J'ai calculé les solutions pour tous les termes périodiques jusqu'à ceux en $6x$, et j'ai pu les illustrer par les chiffres de suicide et de mariage et la température. Ainsi, l'influence sinistre qu'a la date de Pâques se trouve nettement démontrée par un calcul assez simple. Quant à la connexion qu'il y a en Australie entre le suicide et la température, jai trouvé que le résultat moyen peut être exprimé par la formule suivante :

$$g = 0{,}33 + 0{,}003\,(\mathrm{T} - 62^\circ \text{ Fahrenheit})$$

où g représente le nombre quotidien de suicides par million et T la température en degrés Fahrenheit. Pour les degrés Celsius, la formule peut s'écrire de la manière suivante :

$$g = 0{,}33 + 0{,}0017\,(t - 18{,}5^\circ \text{ C.}).$$

L'intervalle entre la valeur maximum et celle qui la suit est de 9 jours à peu près.

Enfin, dans cette note, j'ai indiqué un nombre considérable 'de formules destinées à faciliter les solutions pratiques, et elle me paraît aussi de nature à faciliter les travaux des calculateurs qui ont à déduire des coefficients ou des angles épocaux dans tous les cas où les données originales sont représentées par des valeurs de groupes.

G.-H. Knibbs, C. M. G., F. R. A. S., F. S. S., etc.
(*Statistician of the Commonwealth of Australia.*)