

Ковальський А.А., Тулученко Г.Я.

*Національний технічний університет «Харківський політехнічний інститут», вул. Кирпичова, 2, м. Харків,
email: artem.kovalskyi@ieee.khpi.edu.ua*

РОЛЬ ФРІДРІХА РОБЕРТА ГЕЛЬМЕРТА У РОЗВИТКУ РОЗПОДІЛУ ХІ-КВАДРАТ

Анотація. *Розглянуто внесок Фрідріха Роберта Гельмерта у розвиток розподілу Хі-квадрат, а також вплив його досліджень на сучасний статистичний аналіз.*

Abstract. *The paper examines Friedrich Robert Helmert's contribution to the development of the chi-square distribution, his role in statistics and geodesy, and the impact of his research on modern statistical analysis.*

Ключові слова: *розподіл Хі-квадрат, статистика, геодезія, Гельмерт, Пірсон.*

Key words: *Chi-square distribution, statistics, geodesy, Helmert, Pearson.*

Глибокий аналіз генезису фундаментальних концепцій, таких як розподіл χ^2 -квадрат, розробку якого започаткував Фрідріх Роберт Гельмерт, є необхідним для повноцінного розуміння сучасного статистичного інструментарію. Попри засадничу роль цього розподілу в теорії та практиці статистики, історичний контекст його виникнення та внесок окремих особистостей, зокрема Гельмерта, часто залишаються поза фокусом сучасної математичної історіографії та дидактики.

Дослідження наукової спадщини Гельмерта дозволяє уточнити етапи розвитку статистики та виявити потенційно евристичні аспекти його методології. Аналіз першоджерел розкриває мотивацію досліджень, зв'язок із тогочасними науковими проблемами (напр., теорією похибок) та еволюцію математичного апарату, що призвела до формалізації критерію χ^2 . Розуміння цього генезису поглиблює сприйняття теоретичних основ сучасної статистики та може стимулювати розробку нових підходів в аналізі даних і стохастичному моделюванні.

Одним із важливих попередників Карла Пірсона в розробці критерію Хі-квадрат є німецький математик і геодезист Фрідріх Роберт Гельмерт (1843–1917) [1]. У 1876 році, задовго до робіт Карла Пірсона (рис. 1) [2], Гельмерт опублікував роботу «Про ймовірність степеневих сум похибок спостережень та деякі пов'язані з цим питання» (рис. 2) [3].

X. On the Criterion that a given System of Deviation from the Probable in the Case of a Correlated System of Variables is such that it can be reasonably supposed to have arisen from Random Sampling. By KARL PEARSON, F.R.S. University College, London*.

THE object of this paper is to investigate a criterion of the probability on any theory of an observed system of errors and to apply it to the determination of goodness of fit in the case of frequency curves.

(1) Preliminary Proposition. Let $x_1, x_2, \dots, x_n$ be a system of deviations from the means of n variables with standard deviations $\sigma_1, \sigma_2, \dots, \sigma_n$ and with correlations $r_{12}, r_{13}, r_{23}, \dots, r_{n-1, n}$.

Then the frequency surface is given by

$$Z = Z_0 e^{-\frac{1}{2} \left\{ S_1 \left(\frac{R_{pp}}{R} \frac{x_p^2}{\sigma_p^2} \right) + 2S_2 \left(\frac{R_{pq}}{R} \frac{x_p x_q}{\sigma_p \sigma_q} \right) \right\}}$$

where R is the determinant

$$\begin{vmatrix} 1 & r_{12} & r_{13} & \dots & r_{1n} \\ r_{21} & 1 & r_{23} & \dots & r_{2n} \\ r_{31} & r_{32} & 1 & \dots & r_{3n} \\ \dots & \dots & \dots & \dots & \dots \\ r_{n1} & r_{n2} & r_{n3} & \dots & 1 \end{vmatrix}$$

and R_{pp}, R_{pq} the minors obtained by striking out the p th row and p th column, and the p th row and q th column. S_1 is the sum for every value of p , and S_2 for every pair of values of p and q .

Now let

$$\chi^2 = S_1 \left(\frac{R_{pp}}{R} \frac{x_p^2}{\sigma_p^2} \right) + 2S_2 \left(\frac{R_{pq}}{R} \frac{x_p x_q}{\sigma_p \sigma_q} \right) \dots \dots \dots (ii)$$

Then: $\chi^2 = \text{constant}$, is the equation to a generalized "ellipsoid," all over the surface of which the frequency of the system of errors or deviations $x_1, x_2, \dots, x_n$ is constant. The values which χ must be given to cover the whole of space are from 0 to ∞ . Now suppose the "ellipsoid" referred to its principal axes, and then by squeezing reduced to a sphere $X_1, X_2, \dots, X_n$ being now the coordinates; then the chance of a system of errors with as great or greater frequency than

* Communicated by the Author.

Рис. 1 Перша сторінка роботи [2] Карла Пірсона

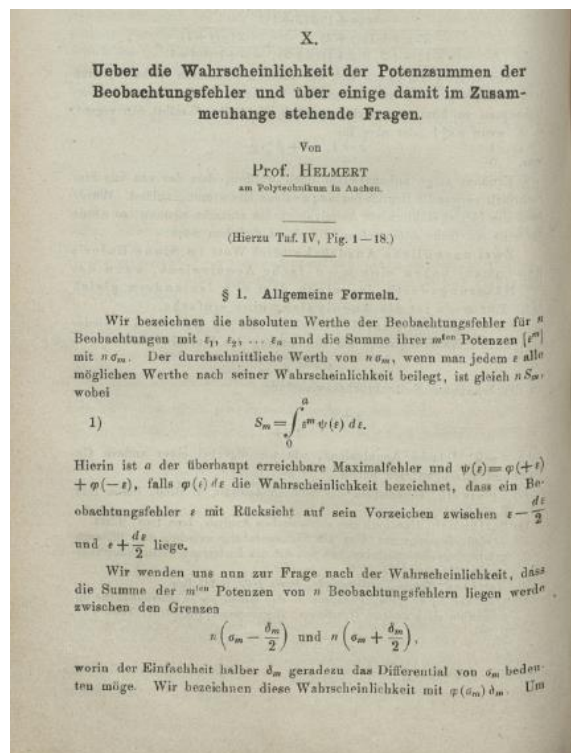


Рис. 2 Перша сторінка роботи [3] Фрідріха Роберта Гельмерта

У цій роботі він математично довів, якщо $Z_1, Z_2, \dots, Z_n$ - незалежні випадкові величини зі стандартним нормальним розподілом $(N(0, 1))$, то сума їх квадратів:

$$\chi^2 = \sum_{i=1}^n Z_i^2$$

має розподіл, який нині відомий як розподіл Хі-квадрат з n ступенями свободи.

Гельмерт також розглянув важливий окремий випадок, тісно пов'язаний зі практичним застосуванням критерію Хі-квадрат. Він показав, якщо $X_1, X_2, \dots, X_n \in$ незалежними випадковими величинами, що мають нормальний розподіл із математичним сподіванням μ та дисперсією σ^2 $(N(\mu, \sigma^2))$, тоді статистика:

$$\frac{(n-1)S^2}{\sigma^2} = \frac{1}{\sigma^2} \sum_{i=1}^n (X_i - \bar{X}), \quad (1)$$

де $\bar{X}$ - вибіркове середнє, S^2 - вибіркова дисперсія, має розподіл Хі-квадрат з $(n-1)$ ступенями свободи.

Хоча розподіл Хі-квадрат нині асоціюється з ім'ям Карла Пірсона, його справжнім першовідкривачем був Фрідріх Роберт Гельмерт. Він застосував описані вище теоретичні результати до аналізу похибок вимірювань у геодезії, що стало основою для подальшого використання цього розподілу в статистичному аналізі. Зазначені дослідження Гельмерт здійснював під час роботи викладачем (а згодом професором) на кафедрі практичної геометрії та геодезії у Рейнсько-Вестфальському технічному університеті Аахена (Німеччина) [1].

Свої наукові результати Гельмерт публікував у німецьких виданнях [4], що обмежило коло науковців, ознайомлених із його працями. Відсутність перекладів англійською мовою спричинила те, що згодом його результати були незалежно повторно відкриті англomовними статистиками: розподіл хі-квадрат – Карлом Пірсоном (1900), а розподіл вибіркової дисперсії для вибірки з нормального розподілу – Стьюдентом (1908) та Фішером (1915, 1925) [5].

ЛІТЕРАТУРА

- [1] Helmert, F. R. (n.d.). *RWTH Aachen University GIA – English*. RWTH Aachen University. <https://www.gia.rwth-aachen.de/cms/gia/das-institut/geschichte/~bfgbfx/helmert/?lidx=1>
- [2] Pearson, K. (1900). X. On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 50(302), 157–175. <https://doi.org/10.1080/14786440009463897>
- [3] Helmert, F. R. (1876). Ueber die Wahrscheinlichkeit der Potenzsummen der Beobachtungsfehler und über einige damit im Zusammenhang stehende Fragen. *Zeitschrift für Mathematik und Physik*, 21, 192–219. http://gdz.sub.uni-goettingen.de/dms/load/img/?PPN=PPN599415665_0021&DMDID=DMDLOG_0018
- [4] Helmert, J. H. (n.d.). *JH Helmert Bibliographie*. GFZ Potsdam. https://gfzpublic.gfz-potsdam.de/rest/items/item_117035/component/file_117034/content
- [5] Plackett, R. L. (1983). Karl Pearson and the chi-squared test. *International Statistical Review*, 61(1), 1–14. <https://doi.org/10.2307/1402731>