ПЕРСОНАЛІЇ

До 120-річчя від дня народження
одного з найвидатніших вчених
XX ст. – ***Гарольда Джеффріса,***
геофізика, математика, астронома

Й.В. Джунь

ГАРОЛЬД ДЖЕФФРІС І ЙОГО ЗАКОН ПОХИБОК

Гарольд Джеффріс, або сер Гарольд, як наполегливо рекомендувала називати його в листуванні з автором леді Джеффріс, народився 22 квітня 1891 р. у містечку Фетфілд, графства Дарем у Великобританії. Навчався в Армстронг, потім Сент-Джонс коледжах Кембриджського університету, після закінчення якого залишився там працювати. В 1917–1922 рр. працював у Метеорологічній службі Англії, а в 1922–1958 рр. – знову в знаменитому Кембриджському університеті (з 1946 року – професором астрономії і філософії). В 1953 р. королева Великобританії удостоїла Гарольда Джеффріса лицарського титулу за видатні наукові досягнення, які звеличують Великобританію. Помер сер Гарольд у Кембриджі 18 березня 1989 р.

Вражає різнобічність інтересів Гарольда Джеффріса. Це видно з повного семитомного видання його праць [21], де, крім рафіновано бездоганних математичних чи геофізичних робіт, є висохудожнє есе про драму Гріга "Пер Гюнт" і тривожні нотатки на екологічні теми.

Основні праці Джеффріса стосуються нашої планети [10]. Його фундаментальна монографія "Земля, її походження, історія і будова" [25] є найкращою науковою роботою про Землю і походження Сонячної системи за всю історію нашої цивілізації. Разом з нідерландським астрономом В. де Сіттером Джеффріс визначив стиснутість Землі за прецесією її осі. Він також склав криву часу пробігу сейсмічних хвиль, яка застосовується для координування епіцентрів далеких землетрусів. Після Дж. Дарвіна Гарольд Джеффріс істотно розвинув теорію припливної еволюції системи Земля–Місяць, визначив енергію припливного тертя на океанських узбережжях. Разом з відомим космологом Дж. Джинсом розробив теорію припливної еволюції Сонячної системи, оцінив її вік.

Проте найбільший рейтинг за кількістю посилань у монументальної роботи Гарольда Джеффріса "The Probability" [26], яка після 1939 р. витримала 9 перевидань. В чому ж секрет такого постійного інтересу до цієї праці Джеффріса, доволі незвичного, як для простого посібника з теорії ймовірностей? Цей інтерес визначається життєвістю, неперехідною актуальністю трьох його концепцій, яскраво виражених у вищезгаданій книзі.

По-перше, Джеффріс у [26] дає нове, після Мізеса [30] та Колмогорова [29], узагальненіше поняття ймовірності, чи не найкраще в сучасній математиці і для сучасної експериментальної науки, про яке ще Пуанкаре в [14] писав: "Ми користуємось цим поняттям, але не знаємо, що це таке". Це означення ймовірності, яке ґрунтується на теорії ступенів правдоподібності [26], відкриває усі глибини імовірнісного бачення Джеффрісом науки як сукупності більш чи менш правдоподібних гіпотез. Можливо, такий погляд не всім подобався, проте такий підхід є найадекватнішим, динамічним і корисно-критичним до всіх наших наукових досягнень.

По-друге, Джеффріс [26] краще ніж усі інші вчені, реалізує наукову заповідь Гаусса, що математика повинна бути передусім подругою практики [1]. Тому видання [26], незважаючи на його дуже високий математичний і філософський рівень, істотно наближає нас до вирішення конкретних проблем сучасної науки в найрізноманітніших її галузях.

По-третє, Джеффріс створив нову, другу після Гаусса, імовірнісну концепцію ідеально хаотичних похибок, тобто дав імовірнісну модель дійсно незалежних випадкових похибок, адекватну практиці багаторазових спостережень. До цієї моделі Джеффріс прийшов, проаналізувавши в роботі [28] результати експерименту К. Пірсона [31]. Серед геодезистів мало кому відомо, що цей експеримент є одним з найоб'ємніших досліджень дійсних розподілів похибок, який виконав засновник математичної статистики. Результати згаданого вище експерименту опубліковано в одному з найпрестижніших видань Великобританії [31]. Спираючись на цей експеримент, Гарольд Джеффріс показав, що незалежні випадкові похибки спостережень, якщо їх кількість перевищує 500, явно підтверджують свій негауссовий характер і задовільно можуть бути описані розподілом Пірсона VII типу:

$$f(x)=\frac{\Gamma(m+1)}{\sqrt{2\pi(m-0,5)}\Gamma(m+0,5)}\cdot\frac{1}{\sigma}\cdot\left\{1+\frac{m^2}{2(m-0,5)^3}\left(\frac{x-\lambda}{\sigma}\right)^2\right\}^{-m} \tag{1}$$

з показником степеня m в межах

$$3\le m\le 5, \tag{2}$$

що відповідає такому діапазону для куртозисів

$$4,2\le\beta_2\le 9,0 \tag{3}$$

В (1) λ, σ, m – параметри розподілу, який можна назвати законом похибок Джеффріса, або просто законом Джеффріса; $\Gamma(m)$ – гамма-функція. Зазначимо також, що щільність ймовірності (1) не є ідентичною класичній кривій Пірсона VII типу. Це – спеціальна форма цієї кривої, модифікована Джеффрісом так, щоб забезпечити її найбажаніші математичні властивості.

Варто зазначити, що досі відсутні переклади книги [26] українською чи російською мовами. Тому широкому колу геодезистів розподіл (1) невідомий, як і його унікальні математичні властивості, які полягають у такому:

1) якщо $m=\infty$, розподіл (1) є законом похибок Гаусса;

2) якщо m кратне 0,5, розподіл (1) ідентичний розподілу Стьюдента. Співвідношення між параметром m розподілу (1) і кількістю ступенів свободи t-розподілу є таким [22]:

$$\nu=2m-1. \tag{4}$$

Отже, закон похибок Джеффріса (1) є узагальненням двох найважливіших розподілів, які використовують при математичній обробці геодезичних спостережень – нормального закону і розподілу Стьюдента;

3) закон (1) є гнучкішим в апроксимаціях емпіричних розподілів похибок, ніж t-розподіл, оскільки він може мати дробові значення ступенів свободи;

4) параметр m в (1) залежить лише від ексцесу ε [7] $m=2,5+\frac{3}{\varepsilon}$ і відображає найхарактернішу особливість розподілів похибок великого обсягу – їх додатний ексцес. Крім того, параметр m є чутливою мірою відхилення реальних розподілів похибок від закону Гаусса;

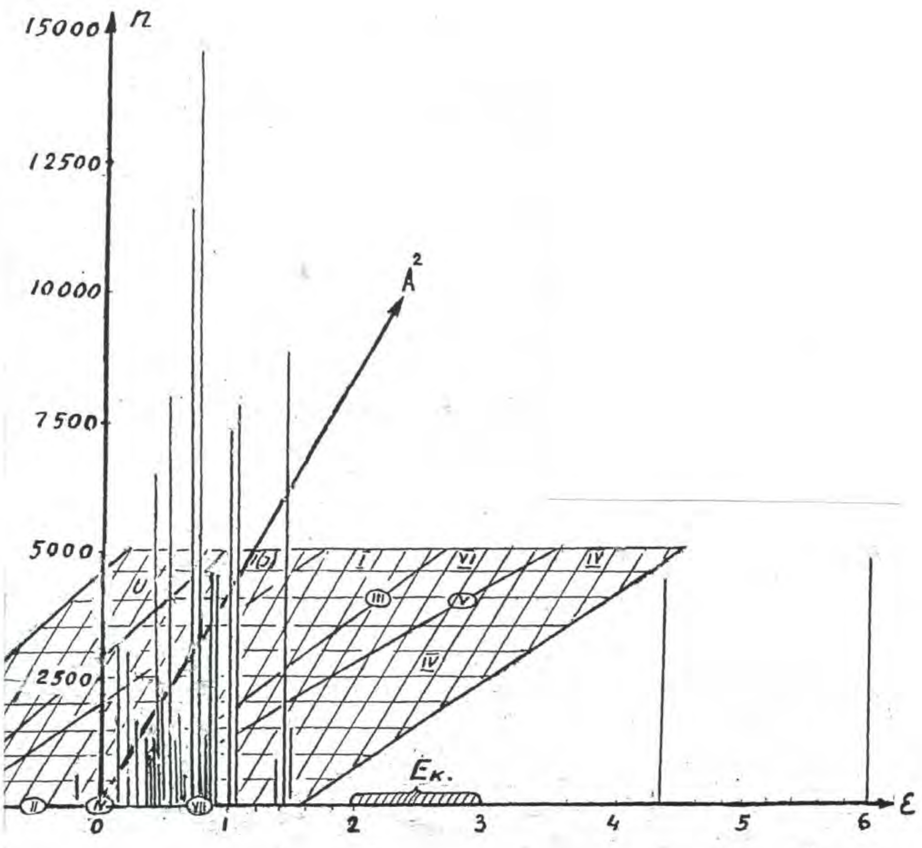
5) найважливішим є те, що інформаційна матриця закону похибок Джеффріса (1) є діагональною, як і у нормального закону [7];

6) закон (1) є регулярним в інтервалі $(-\infty;+\infty)$.

Проте значення закону похибок Джеффріса (1) не лише в тому, що він має вищезгадані дуже цінні математичні властивості. Головна перевага функції (1) в тому, що емпіричні похибки зазвичай підпорядковуються саме їй, якщо кількість спостережень переважає 500. Цей факт підтверджено фундаментальними дослідженнями, виконаними в Києві у 1969–1992 рр. для астрономічних і космічних спостережень під керівництвом академіка Є.П. Федорова [3, 4, 7], потім – академіка Я.С. Яцківа [11, 15, 17].

На рисунку подано результати експерименту, виконаного автором [7] з метою перевірки ідеї Джеффріса про відповідність емпіричних розподілів похибок великих обсягів, розподілу Пірсона

VII типу. В основу такої перевірки, з метою забезпечення наочності, покладено графік для ідентифікації типів розподілів Пірсона, взятий із таблиць [2]. Кожний емпіричний розподіл на рисунку має три координати: ексцес ε, квадрат асиметрії A^2, обсяг n. Нормальному закону на рисунку відповідає точка N, яка і є початком координат. Як видно з рисунка, основною особливістю переважної більшості емпіричних розподілів є їх практично відсутня асиметрія і високозначущий додатний ексцес, близький до 1. Саме це є характерними ознаками симетричного розподілу Пірсона VII типу.



Розміщення емпіричних розподілів похибок астрономічних, космічних, гравіметричних, геодезичних, економічних рядів (область E_k) на осі ексцесу ε (на графіку лінія VII, праворуч від точки N, відповідає сім'ї Пірсона VII типу)

На рисунку практично всі вертикальні прямі розміщені на лінії цього розподілу. Ці прямі групуються не навколо точки N, а в околі іншої точки, яка істотно зміщена праворуч в область додатних ексцесів, і розташована на лінії кривих Пірсона VII типу. Доречно зазначити, що до такого самого висновку прийшов відомий російський вчений Н. І. Ідельсон, який ще в 1947 р. писав: “Ніколи ще, наскільки нам відомо, не зустрічались ряди похибок з від'ємним ексцесом” [12]. Проте цей висновок був забутий.

Несподіваним результатом досліджень автора було і те, що не лише похибки астрономічних і космічних спостережень підпорядковувались розподілу Пірсона VII типу (1) [3, 4, 7, 18, 19], але і похибки геодезичних, неспотворених допусками спостережень [9], гравіметричних [5, 6], геофізичних [8], економічних [23, 32]. Який би ряд похибок не досліджував автор, не важливо – класичних, чи сучасних спостережень, – в усіх випадках похибки відповідали всеохопному закону Джеффріса (1).

Показавши в роботах [25–28] теоретичну і практичну неспроможність закону Гаусса за кількості спостережень понад 500, Джеффріс не ставив мети позбавити цей закон його високого статусу. Сер Гарольд лише врахував наявні реалії. Крім того, всі варіанти запропонованих ним оцінок можна отримати лише після того, як ми спочатку застосуємо прості процедури оцінювання за Гауссом. Джеффріс всіляко намагався сприяти еволюції класичних підходів до математичної обробки даних. Щоб показати це, розглянемо границі нерівності Рао–Крамера для оцінок параметрів закону похибок Джеффріса (1):

$$\sigma_\lambda^2 \ge \frac{\sigma^2}{n} \cdot \frac{(m-0,5)^2 \cdot (m+1)}{m^3}; \quad \sigma_\sigma^2 \ge \frac{\sigma^2}{2n} \cdot \frac{m+1}{m+5}; \tag{5}$$

$$\sigma_m^2 \ge \left\{ n \left[\psi'(m-0,5) - \psi'(m) - \frac{m+1}{2m^2(m-0,5)} \right] \right\}^{-1}, \tag{6}$$

де $\psi'(m)$ – псі-функція.

З формул (5) видно, що при $m = \infty$ (закон Гаусса) σ_λ^2 і σ_σ^2 перетворюються на добре відомі геодезистам формули для обчислення дисперсій похибок середньої арифметичної та середньоквадратичної похибки.

Обмеженість закону Гаусса полягає в тому, що він не враховує найістотнішої особливості дійсних розподілів помилок – їхній додатний ексцес. Гауссова форма розподілу (1) постулює m наперед відомим і таким, що дорівнює ∞. Джеффріс показав небезпечність такого припущення, а також те, що таке припущення блокує еволюцію класичних процедур обробки даних. Насправді кожен ряд похибок має конкретне значення m, яке дуже далеке від ∞ і яке є його найголовнішою характеристикою. Тому сера Гарольда завжди насамперед цікавили значення m, отримані для того чи іншого ряду помилок. Зазначимо, що границі для σ_λ^2 та σ_σ^2 отримані Джеффрісом, а границі для σ_m^2 – автором у роботі [7].

Для апроксимації похибок рядів спостережень великих обсягів багато дослідників запропонували різні функції, які, проте, не досліджувались на такому високому рівні, як це зробив Джеффріс для розподілу (1). Наприклад, розрекламований і відомий геодезистам L_p-розподіл, як виявилось, має недіагональну інформаційну матрицю Фішера і до того ж є нерегулярним при $p \ne 2$ [7]. Досі ще не оцінено належно значення досягнень в аналізі спостережень великих обсягів, які започаткував Джеффріс. На геодезичних факультетах ВНЗ України, як правило, не згадуються його досягнення, або ж приховуються за невиразним терміном "робастні процедури". Математики ще не створили чогось адекватного, принаймні подібного за своїми властивостями, до закону похибок (1). Сучасні математики не бачать блискучих математичних властивостей цього закону. Вони лише удостоїли Джеффріса похвали за "досконало чітке розуміння ідеї робастності" [16]. Швидше за все, вони так і не усвідомили фундаментального значення його закону похибок. Навіть після появи роботи [26], де теоретично бездоганно розглянуто проблеми нормального закону, гнучкість і переваги форми (1), деякі математики все ще продовжували тиражувати роботи про примітивні моделі суміші кількох розподілів з різними дисперсіями [20], упускаючи з виду те, що для таких сумішей, внаслідок їхньої нерегулярності, неможливо знайти границі для ефективних оцінок. Гаусс у теорії обробки геодезичних даних і врівноваження залишається єдиним загальновизнаним авторитетом. Ідеї і досягнення Джеффріса все ще залишаються поза увагою науковців-геодезистів. Головна причина цього навіть не в тому, що сучасна вища освіта відстає від вимог практики – вона в тому, що приблизно 90 % наукових публікацій, як свідчить дослідження [13], є науковим непотребом. У такій масі публікацій, на жаль, швидко забулись навіть вишукані математичні праці Джеффріса з дослідження похибок спостережень. Проте, зважаючи на те, що його фундаментальна робота "The Probability" [26] і в наш час регулярно перевидається, можна вважати, що ідеї Джеффріса переживають ренесанс.

1. Гаусс К.Ф. Избранные геодезические сочинения. Том 1. Способ наименьших квадратов / Под ред. Г.В. Багратуни. Пер. с лат. и нем. – М.: Изд. геодезич. лит., 1957. 2. Большев Л. Н., Смирнов А. В.

Таблицы математической статистики. – М.: ВЦ АН СССР, 1968. – 474 с. 3. Джунь И.В. Распределение Пирсона VII типа в ошибках наблюдений над колебаниями широт // Астрометрия и астрофизика. К., 1969. – Вып. 2. – С. 101–115. 4. Джунь И.В О назначении весов астрономическим наблюдениям // Астрометрия и астрофизика. – К., 1970. – Вып. 10. – С. 26–34. 5. Джунь И.В., Арнаутов Г.П., Стусь Ю.Ф, Щеглов С.Н. Особенность закона распределения баллистических измерений ускорения силы тяжести // Повторные гравиметрические наблюдения. – М.: Изд МГК при Презид. АН СССР и НПО "Нефтегеофизика", 1984. – С. 87–100. 6. Джунь И.В. Флюктуации веса индивидуальных измерений ускорения силы тяжести и способ их учета при обработке баллистических наблюдений // Повторные гравиметрические наблюдения. – М.: Изд МГК при Презид. АН СССР и НПО "Нефтегеофизика", 1983. – С. 43–52. 7. Джунь И.В. Математическая обработка астрономической и космической информации при негауссовых ошибках наблюдений. Автореферат дис. доктора физ-мат. наук. – К.: ГАО АНУ, 1992. – 46 с. 8. Джунь И.В., Сомов В.И. О некоторых фундаментальных вопросах математической обработки геофизической информации.// Геодинамические исследования в Украине. – К., 1996. – С. 167–178. 9. Джунь И.В. Устарел ли способ наименьших квадратов? // Кинем. и физ. неб. тел, 2000. – Т. 16, № 3. – С. 281–288. 10. Козенко А.В. Гарольд Джеффріс, 1891–1989. – М.: Наука, 2008. – 248 с. 11. Иванов Г.А., Сергеева Т.П., Яценко А.И. Выбор режима измерений программы ФОН на автоматической машине ПАРСЕК // Кинем. и физ. неб. тел., 1990, Т. 6, № 1. – С. 79–83. 12. Идельсон Н. И. Способ наименьших квадратов и теория математической обработки наблюдений. – М.: Геодезиздат, 1947. – 359 с. 13. Орлов А.И. Высокие статистические технологии // Заводская лаборатория, 2003. – Т. 69. – № 11. – С. 55–60. 14. Пуанкаре А. О науке. Пер. с франц. – М.: Наука, 1983. – 560 с. 15. Харин А.С., Яцкив Я.С. Изучение ошибок наблюдений Голосеевского каталога звезд широтных программ // Астрометрия и астрофизика. – К., 1970. – Вып. 10. – С. 34–43. 16. Хьюбер П. Робастность в статистике. – М.: Мир, 1984. – 304 с. 17. Яцкив Я.С., Молотай А.А. Об использовании оптимальных линейных оценок математического ожидания и стандартного отклонения при обработке результатов астрономических наблюдений // Астрометрия и астрофизика, 1979. – № 37. – С. 56–60. 18. Branham R.L. Techniques for dealing with discordant observations // Relativity in Celestial Mechanics and Astrometry. Reidel, 1986. – P. 229–230. 19. Broslavets D.G., Dzhun' I.V., Gorel G.K., Gudkova L.A. Research on statistical distributions of observation errors of minor planets // Extension and Connection of Reference Frames using Ground Based CCD Technique. International astronomical conference- Nikolaev; Atoll,-2001, p. 150–156. 20. Cohen C.A. Estimation in Mixtures of Two Normal Distributions// Tehnometrics, 1967, Vol. 9, № 1. 21. Collected Papers of sir Harold Jeffreys on Geophysics and Others Sciences. In six volumes. Edited by sir Harold Jeffreys and Berta Swirles (lady Jeffreys) Gordon and Breach, Science Publishers Ltd. London-Paris_New York, 1977. 22. Dzhun' I.V., Novitskij P.V. Comments upon the Use of the Pearson Law of Type VII in Astrometry.// Kinematics and Physics of Celestial Bodies, New York, Allerton Press, 1992, vol.8, №5, p.78–81. 23. Gazda V. Normal probability Distribution in financial Theory – false Assumption and Consequences // Department of Economics, University of Economics, Faculty of Business Economics, Kosice, 1999. 24. Gentleman W. M. Robust estimation of multivariate location by minimizing p-th powers deviations – Dissertation – Princeton University and Memorandum: MM 65-1215-16, 1965. 25. Jeffreys H. The Earth: its origin, history and physical constitution. Cambridge University Pres, 1959. – 420 p. 26. Jeffreys H. Theory of Probability. Sec. Edition, Oxford, 1940. 27. Jeffreys H. The Law of Errors and the Combination of Observations // Philosophical Transactions of the Royal Society, series A, 1937, №237, p. 231-271. 28. Jeffreys H. The Law of Error in the Greenwich Variation of Latitude Observations // Monthly. Notices. of the RAS, 1939, v. 99, № 9, p. 703–709. 29. Kolmogoroff A. Grundbegriffe der Wahrscheinlichkeitsrechnung. – Berlin, 1933. 30. Mises R. On the fundamentions of probability and statistics. AMS, 1941, № 12. 31. Pearson K. On the Mathematical Theory of Errors of Judgment, with Special Reference to the Personal Equation // Philosophical Transactions of the Royal Society of London. Ser. A., 1902, vol. 198, p. 253-296. 32. Peters E. E. Fractal Market Analysis. Applying Chaos Theory to Investment and Economics. New York, Chichester, Brisbane, Toronto, Singapore: John Willey and Sons, INC, 1981, p. 18–53.