КИНЕМАТИКА
И ФИЗИКА
НЕБЕСНЫХ
ТЕЛ том 28 № 1 2012

УДК 521.9:528.111:519.24

И. В. Джунь

Международный экономико-гуманитарный университет
4, корп. 2, ул. С. Демьянчука, г. Ровно, 33018

Какими должны быть разности «observation – calculation» при постановке современных экспериментов в астрометрии?

Обосновывается важность закона ошибок Джеффриса для анализа разностей O – C. На основе этого закона показано, что эксперимент можно считать проведенным корректно только при условии, когда значения O – C следуют t-распределению с числом степеней свободы ν между 5 и 9 или джеффрисовой форме распределения Пирсона VII типа с показателем степени m между 3 и 5 при постоянстве метрологической ситуации. Если последняя нестабильна, то для распределения разностей O – C левая граница для ν может смещаться не менее чем до 3, а показатель m — не менее чем до 2.

ЯКИМИ МАЮТЬ БУТИ РІЗНИЦІ «OBSERVATION – CALCULATION» ПРИ ПРОВЕДЕННІ СУЧАСНИХ ЕКСПЕРИМЕНТІВ В АСТРОМЕТРІЇ? Джунь Й. В. — Обгрунтовується значення закону похибок Джеффріса в аналізі різниць O – C. На основі цього закону показано, що експеримент можна вважати проведеним коректно лише за умови, якщо значення O – C підкоряються t-розподілу з числом ступенів свободи ν між 5 і 9, чи джеффрісовій формі розподілу Пірсона VII типу з показником степеня m між 3 і 5 за сталої метрологічної ситуації. Якщо остання нестабільна, то ліва межа для ν може зміщуватись не менше ніж до 3, а показник степеня m — не менше ніж до 2.

WHAT ARE DIFFERENCES “OBSERVATION – CALCULATION” BOUND TO BE DURING MODERN EXPERIMENTS IN ASTROMETRY?, by Dzhun I. V. — The importance of the Jeffreys Errors Law to the analysis of differences O – C is justified. Based on the law, it is shown that an experiment may be considered as performed correctly only when the differences O – C follow the t-distribution with the number of degrees of freedom ν between 5 and 9 or Pearson’s VII type distribution with the index

of degree m between 3 and 5 in stable metrological situation. When metrological situation is unstable, the left limit can shift no less than to 3 for v and no less than to 2 for m.

Главной особенностью современных астрономических экспериментов есть невиданное, по сравнению с классическими, увеличение числа наблюдений, обусловленное автоматизацией измерений. Это обстоятельство коренным образом меняет устоявшиеся подходы к обработке и анализу данных, так как нынешние гигантские по объему выборки неизбежно опровергают заложенные еще Гауссом фундаментальные принципы классической теории ошибок. Это и не удивительно, ведь эта теория создавалась на практической основе микровыборок, а главным критерием ее разработки была простота вычислений. В связи с вышеизложенным особый интерес представляет рассмотрение современных подходов к тому, какими вероятностными свойствами должны обладать разности *O – C* при постановке нынешних астрометрических экспериментов.

Как известно, обработка данных базируется на математическом моделировании, которое опирается на две основные концепции:

— математическое представление изучаемого процесса;

— вероятностное представление о том, какими свойствами должны обладать ошибки наблюдений (в нашем случае — разности *O – C*).

Если исходить из принципа, подтверждаемого историей науки, что любая теория существует для того, чтобы со временем пострадать от фактов, то, в этом случае, исследователь должен уделять самое пристальное внимание анализу разностей «практика минус теория», т. е. ошибкам *O – C*. По поводу анализа этих разностей академик А. Мигдал выразился так: «Очень сложный вопрос — сравнение теории с экспериментом. Для меня, как для физика-теоретика, это очень волнующий вопрос... Несовпадение хорошей теории с опытом, как правило, означает, что произошло «какое-то малое или большое открытие...» [16]. Эта проблема особенно интересовала основателя математической статистики К. Пирсона, который еще в 1902 г. в работе [32], закладывая основы неклассической теории ошибок, на с. 283 посвящает этому вопросу целый раздел под названием: «Согласие между теорией и наблюдениями определяется в каждом случае по особым свойствам генерального распределения ошибок».

Волнует ли проблема такого согласия и астрономов? Часто ли они изучают свойства разностей *O – C*? Например после получения регрессионной модели, хотя бы с позиции классического способа наименьших квадратов. К сожалению, этого не наблюдается. Астроном, применяя для оценивания параметров математической модели, как правило, классический метод наименьших квадратов (МНК), обычно считает, что К. Ф. Гаусс, создавая его, основательно во всём разобрался, и каких-либо особых подозрений относительно разностей *O – C* быть не может. Весь парадокс состоит в том, что Гаусс не исследовал в полной мере границы применимости МНК. Об этом он открыто пишет в [2]: «Какие средства может требовать вычислитель от теории вероятностей при обработке наблюдений, которые не вполне свободны от систематических ошибок? Опубликование особого исследования об этих средствах мы откладываем до другого случая». Этого другого случая ему так и не предоставилось. В то же время каждый астроном знает, что полностью исключить систематические ошибки из результатов наблюдений нельзя. Можно только загнать их в некоторые границы. По этому поводу И. Г. Колчинский писал [14]: «Источники систематических погрешностей нужно изучать и исключать в процессе обработки данных, но никогда нет гарантии, что это можно сделать достаточно хорошо». В то же время, главное математическое условие моделирования — это полное отсутствие систематических ошибок в результатах экспериментов [2]. Понятно, эти неслучайные ошибки как-то будут влиять на распределение разностей *O* – С. Вопросы анализа такого влияния являются достаточно сложными не только для астрометриста, но и для математика [1]. Да и астроном, как правило, не знаком с современными представлениями о том, какими наиболее желательными вероятностными свойствами должны обладать разности *O* – *C*.

С позиций классической теории ошибок наиболее желательно, чтобы разности *O* – *C* следовали закону Гаусса.

Со временем проблема анализа разностей *O* – *C* существенно усложнилась после того, как Г. Джеффрис в работах [29—31] показал теоретическую и практическую несостоятельность закона Гаусса, если наблюдений более 500. Джерри, которого цитирует Д. Тьюки в работе [34], высказался еще более критически: «Нормальный закон — это миф. В реальном мире никогда не было и никогда не будет нормального распределения».

Практически оказалось, что даже ошибки массовых технических и производственных измерений в 92 % случаев не являются гауссовыми [18]. Если при числе наблюдений больше 500 разности *O* – *C* негауссовы, то какими они должны быть для хорошо поставленного эксперимента и при идеальном согласии теории с практикой? Пионером решения этой проблемы является Г. Джеффрис, который начал реализовывать упомянутую выше идею К. Пирсона. Опираясь на результаты его известного эксперимента [32], Джеффрис в работе [30] приходит к выводу, что идеально хаотические ошибки при однородной метрологической ситуации и при отсутствии систематических влияний следуют распределению Пирсона VII типа:

$$f(x) = \frac{\Gamma(m-1)}{\sqrt{2\pi(m-0.5)}\,\Gamma(m-0.5)}\frac{1}{\sigma}\left[1+\frac{m^2}{2(m-0.5)^3}\left(\frac{x}{\sigma}\right)^2\right]^{-m}, \quad (1)$$

с показателем степени *m* в пределах

$$3 \le m \le \infty, \quad (2)$$

что соответствует следующим границам для эксцесса:

$$6 \qquad . \qquad\qquad (3)$$

В выражении (1) , , m — параметры распределения, которое можно назвать законом ошибок Джеффриса или просто законом Джеффриса. Правильность рекомендации Джеффриса (1) была подтверждена фундаментальными исследованиями, выполнеными в Киеве под руководством академика Е. П. Фёдорова [3, 4, 9, 15] и академика Я. С. Яцкива [13, 21, 23].

Необходимо заметить, что некоторые объемные ряды высокого качества, например распределение относительных значений биржевых индексов, исследованные в работах [27, 28, 33], действительно имеют значение 2, которое попадает в границы, указанные Джеффрисом.

Самый неожиданный результат состоял в том, что не только ошибки астрономических наблюдений подчинялись закону Джеффриса (1) [3—5, 15, 25, 24, 30], но также и ошибки геодезических [10], геофизических [12], гравиметрических [5, 11], экономических [27, 28, 33] и других рядов наблюдений. Какой бы ряд ошибок не исследовался не важно, исторических или современных наблюдений — во всех случаях для его ошибок вырисовывалась «магическая» кривая распределения Пирсона VII типа, т. е. тот самый закон Джеффриса (1).

Следует сказать, что до сих пор еще не оценено в должной мере значение того прорыва в теории ошибок и анализе данных, который осуществил Джеффрис в своих трудах [29—31], обосновывая фундаментальную форму (1). В университетах США, Европы, не говоря уже о вузах Украины или России, обычно даже не упоминается об этом достижении, или оно скрывается за расплывчатыми терминами робастных процедур. Современные математики только удостоили Джеффриса похвалы «за совершенно отчетливое понимание идеи робастности» [22]. Ими, скорее всего, так и не осознано фундаментальное и синтетическое значение закона ошибок Джеффриса (1). Даже после появления фундаментального труда Г. Джеффриса [31], в котором теоретически безукоризненно были рассмотрены проблемы нормального закона и показаны преимущества формы (1), некоторые математики все ещё продолжали тиражировать работы о примитивных моделях смеси нескольких распределений, упуская из виду то обстоятельство, что для такой «смеси» невозможно построить границы для эффективных оценок, ввиду её нерегулярности.

Авторитет Гаусса в теории ошибок в университетах остается незыблемым, а идеи Джеффриса воспринимаются не как прорыв к новому, а, скорее, как занимательный научный курьез, хотя они наиболее полно выражают суть современных высоких математико-статистических технологий [19].

Для демонстрации того, что закон Джеффриса (1) является естественной эволюцией и обобщением классической теории ошибок, приведем для него нижние границы неравенства Рао — Крамера для выборочных оценок дисперсий его параметров:

$$ {}_{2}\quad \frac{{}^{2}}{n}\ \frac{(m\quad 0.5)^{2}(m\quad 1)}{m^{3}}, \tag{4}$$

$$ {}_{2}\quad \frac{{}^{2}}{2n}\ \frac{m\quad 1}{m\quad 0.5}, $$

$$ {}^{2}_{m}\quad n\quad (m\quad 0.5)\quad (m)\ \frac{m\quad 1}{2m^{2}(m\quad 0.5)}\ {}^{1}, \tag{5}$$

где (m) — тригамма-функция.

В реальном случае оценки параметров закона плотности (1) будут эффективными, если эмпирические распределения разностей *O* – *C* следуют этому закону. Действительно ли это так? В качестве доказательства мы приведем вычисленные выборочные значения моментных отношений:

$$ {}_{1}\quad {}^{2}_{3}/\ {}^{3}_{2},\qquad {}_{2}\ 3\quad {}_{4}/\ {}^{2}_{2}\quad 3 $$

для семи серий разностей *O* – *C*, полученных по проекту MERIT в работе [8].

Как видно из таблицы, все распределения разностей *O* – *C* имеют несущественную асиметрию и, за исключением одного случая, — высокодостоверный положительный эксцес, т. е. в большинстве своем принадлежат области симметричных распределений Пирсона VII типа, т. е. области с эксцессами в пределах 0 . Но, как показал Джеффрис в работе [30], симметричные распределения даже с эксцессом > 1.5 удовлетворительно аппроксимируются распределением (1).

Из формул (4) также видим, что при $m =$ (закон Гаусса), ^{2} и ^{2} превращаются в классические формулы для вычисления дисперсии среднего арифметического и стандарта. Что же касается параметра *m*, то гауссова форма распределения (1) постулирует значение *m* наперед известным и равным бесконечности. На самом же деле для каждого ряда наблюдений характерно свое конкретное значение *m*, очень дале-

Значения статистик $_1$ и для семи серий разностей *O* — *C* (международный проект MERIT)

Номер распределения	Дуга	Объем выборки, *n*	$_1$	
1	466.5—471.5	658	0.017 0.025	1.44 0.19
2	471.5—476.5	579	0.059±0.050	1.36 0.20
3	476.5—481.5	813	0.081 0.049	1.64 0.17
4	481.5—486.5	817	0.067 0.044	1.03 0.17
5	486.5—491.5	732	0.008 0.016	0.18 0.18
6	491.5—496.5	880	0.010 0.017	3.74 0.16
Суммарное	466.5—496.5	4475	0.0048 0.0050	1.858 0.073

кое от бесконечности, которое, как нами было отмечено выше, является ключевой характеристикой ошибок этого ряда. Заметим, что границы (4) получены Джеффрисом в работе [30], а граница (5) — автором в работе [7]. Мы уделяем столько внимания закону Джеффриса (1), исходя из того, что им была представлена не только новая концепция идеального хаоса, адекватного практике наблюдений, но и положено начало нового витка в совершенствовании методов обработки данных.

Чтобы показать значение прорыва Джеффриса вследствие введения закона (1), сообщим то, что мало или совершенно неизвестно астрометристам.

1. К закону (1) Джеффрис пришёл, исходя из классической кривой Пирсона VII типа, которую он модифицировал так, чтобы информационная матрица распределения (1) была столь же безупречной, как информационная матрица закона Гаусса.

2. Форма (1) есть обобщением двух наиважнейших в теории ошибок и анализе данных распределений —

а) закона Гаусса, который мы получаем из (1), если m ;

б) распределения Стьюдента, которому соответствуют в (1) значения m, кратные 0.5;

в) соотношение между параметром m в (1) и количеством степеней свободы распределения Стьюдента, как показано в работе [26], следующее:

$$2m \quad 1; \tag{6}$$

г) закон ошибок Джеффриса (1) может иметь и дробные значения степеней свободы.

Таким образом, ключевые для анализа разностей $O - C$ границы (2) с учетом выражения (6) можно представить следующим образом:

$$5 \qquad 9. \tag{7}$$

3. Параметр m распределения (1) зависит исключительно от эксцесса в интервале 0 и его можно расматривать как меру уклонения более гибкого закона ошибок Джеффриса от классического нормального закона.

Предложение Джеффрисом формы (1) знаменует собой необходимую эволюцию в теории ошибок.

Была сделана попытка [20] передвинуть зону идеального джеффрисового хаоса несколько левее от левой границы в (2): «Джеффрис с несколько чрезмерной осторожностью высказал предположение, что независимые данные при однородных условиях аппроксимировались бы t-распределением с 9—5 степенями свободы. Ряды наблюдений широты, корреляционная структура которых, надо полагать, не исследовалась, но которые представляют собой ряды действительных наблюдений при неоднородных условиях, подгоняются (Джеффрисом в [30] — авт.) под t-распределение с 3.5 и 4.5 степенями свободы. Если бы в них тоже проявлялась корреляция, чего можно было бы ожидать, то экстраполяция на независимый случай потребовала бы для ряда меньшего числа степеней свободы. Это согласуется в какой-то мере с результатом Хьюбера (устное сообщение), который для выглядящих примерно так же данных высокого качества тоже назвал в качестве подходящего примера распределение t_3» [20, с. 41].

С высказанными в [20] выводами можно было бы полностью согласиться, если бы не следующие обстоятельства. Во-первых, ряды К. Пирсона [32], на основании анализа которых Джеффрис определил пределы (2), получены в однородных условиях наблюдений. Во-вторых, два ряда наблюдений широты, на которые ссылаются в [20], получены не только в крайне неоднородных условиях наблюдений (на открытом воздухе в течение всех сезонов на протяжении 1927—1931 и 1932—1936 гг.), но и на очень чувствительном инструменте (плавающем на ртути зенит-телескопе Куксона, который дрожал, реагируя на малейшее движение ветра) [30]. Это обстоятельство не было известно авторам работы [20]. Описаная особенность устройства телескопа вызывала существенные флюктуации точности наблюдений во времени, что, в соответствии с теоремой Эддингтона — Огородникова [17], неизбежно приводит к значительному увеличению положительного эксцесса распределения ошибок наблюдений. Учитывая вышесказанное, случай с телескопом Куксона не является типичным. В третьих, Джеффрис сообщает о пределах (2), свойственных наблюдениям при однородной метрологической обстановке. Если флюктуации точности наблюдений реальны и статистически будут доказаны, то, естественно, левая граница для   в (7) может приближаться к   3, а левая граница для m в (2) — к $m = 2$, что характерно для ошибок гринвичских широтных наблюденый ($m = 2.71$ для ряда 1927—1931 гг. и $m = 2.26$ для ряда 1932—1936 гг.) [30].

Учитывая сказанное, можно сделать следующий вывод: разности $O - C$ в идеале (если теория адекватна практике), при постоянной метрологической ситуации, должны следовать распределению Пирсона VII типа (1) со значениями показателя m в пределах (2) или t-распределению со степенями свободы в пределах (7).

При наличии флюктуаций точности наблюдений левая граница в (2) может сдвигаться не ниже $m = 2$, а левая в (7) не меньше чем к   = 3. Если же значение m или , вычисленные для конкретного ряда значений $O - C$, не попадают в указанные пределы, то это есть неоспоримым свидетельством плохой теории изучаемого явления, или грубо поставленного астрометрического эксперимента. Астрометрист, который провел тот или иной эксперимент, может быть спокоен только тогда, когда значения m или , вычисленные для разностей $O - C$, попадают в указанные пределы, или близки к ним. Если же этого не наблюдается, то это есть очень весомый повод для беспокойства относительно правильности используемой теории, или же по поводу не совсем качественно поставленного эксперимента.

1. *Бородачев Н. А.* Основные вопросы теории точности производства / Под ред. А. Н. Колмогорова. — М.-Л.: Изд-во АН СССР, 1950.—360 с.

2. *Гаусс К. Ф.* Избранные геодезические сочинения: Пер. с лат. и нем. / Под ред. Г. В. Багратуни. — М.: Изд-во геодез. лит., 1957.—Т. 1. Способ наименьших квадратов.—234 с.

3. *Джунь И. В.* Распределение Пирсона VII типа в ошибках наблюдений над колебаниями широт // Астрометрия и астрофизика.—1969.—Вып. 2.— С. 101—115.

4. *Джунь И. В.* О назначении весов астрономическим наблюдениям // Астрометрия и астрофизика.—1970.—Вып.10.—С. 26—34.

5. *Джунь И. В.* Флюктуации веса индивидуальных измерений ускорения силы тяжести и способ их учета при обработке баллистических наблюдений // Повторные гравиметрические наблюдения. — М.: Изд-во МГК при Презид. АН СССР и НПО «Нефтегеофизика», 1983.—С. 43—52.

6. *Джунь И. В.* Некоторые аспекты практического использования *Lp* и эксцесс-оценок при обработке геодезических измерений // Изв. вузов. Геодезия и аэрофотосъемка.—1986.—№ 4.—С. 43—48.

7. *Джунь И. В.* О границах неравенства Рао — Крамера для дисперсий оценок параметров распределения Пирсона VII типа // Кинематика и физика небес. тел.—1988.—**4**, № 1.—С. 85—87.

8. *Джунь И. В.* Распределение Пирсона VII типа ошибок лазерных наблюдений ИСЗ // Кинематика и физика небес. тел.—1991.—**7**, № 3.—С. 82—91.

9. *Джунь И. В.* Математическая обработка астрономической и космической информации при негауссовых ошибках наблюдений: Автореф. дис. ... д-ра физ.-мат. наук. — Киев: ГАО АН Украины, 1992.—46 с.

10. *Джунь И. В.* Устарел ли способ наименьших квадратов? // Кинематика и физика небес. тел.—2000.—**16**, № 3.—С. 281—288.

11. *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-во МГК при Презид. АН СССР и НПО «Нефтегеофизика», 1984.—С. 87—100.

12. *Джунь И. В., Сомов В. И.* О некоторых фундаментальных вопросах математической обработки геофизической информации // Геодинамические исследования в Украине. — Киев, 1996.—С. 167—178.

13. *Иванов Г. А., Сергеева Т. П., Яценко А. И.* Выбор режима измерений программы ФОН на автоматической машине ПАРСЕК // Кинематика и физика небес. тел.—1990.—**6**, № 1.—С 79—83.

14. *Колчинский И. Г.* Наблюдение и факт в астрономии. — Киев: Наук. думка, 1982.—104 с.

15. *Король А. К.* Склонения ярких и слабых звезд в единой системе. — Киев: Наук. думка, 1969.—236 с.

16. *Мигдал А.* Симметрично ли пространство? // Наука и жизнь.—1971.—№ 9.— С. 53—57.

17. *Огородников К. Ф.* Метод обработки наблюдений введением средних весов с применением к статистическому изучению звездных движений // Астрон. журн.—1928.—**5**, вып. 1.—С. 1—21.

18. *Орлов А. И.* Часто ли распределение результатов наблюдений является нормальным? // Заводская лаборатория.—1991.—№ 7.—С. 64—66.

19. *Орлов А. И.* Высокие статистические технологии // Заводская лаборатория.— 2003.—**69**, № 11.—С. 55—60.

20. *Хампель Ф., Рончетти Э., Рауссеу П., Штаэль В. М.* Робастность в статистике.

Подход на основе функций влияния / Пер. с англ. — М.: Мир, 1989.—512 с.

21. *Харин А. С., Яцкив Я. С.* Изучение ошибок наблюдений Голосеевского каталога звезд широтных программ // Астрометрия и астрофизика.—1970.—Вып. 10.— С. 34—43.
22. *Хьюбер П.* Робастность в статистике. — М.: Мир, 1984.—304 с.
23. *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.— Вып. 37.—С. 56—60.
24. *Branham R. L.* Techniques for dealling with discordant observations // Relavity in Celestial Mechanics and Astronometry. — Reidel, 1986.—P. 229—230.
25. *Broslavets D. G., Dzhun' I. V., Gorel G. K., Gudkova L. A.* Research on statistical distributions of observation errors of minor panets // Extension and Connection of Reference Frames using Ground Based CCD Technique. International astronomical conference. — Nikolaev: Atoll, 2001.—P. 150—156.
26. *Dzhun' I. V., Novitskij P. V.* Comments upon the use of the Pearson law of type VII in astronometry // Kinematics and Physics of Celestial Bodies.—1992.—**8**, N 5.— P. 78—81.
27. *Gazda V.* Normal probability distribution in financial theory — false assumption and consequences // Department of Economics, University of Economics, Faculty of Business Ekonomics, Kosice, 1999.—P. 3—6.
28. *Gazda V., Dzhun' I. V.* About distribution of Stock index returns fluctuations // Business Review: Scientific journal of the Faculty of Business Economics of the University of Economics in Bratislava with a seat in Kosice, 2002.—**1**, N 2.— P. 20—27.
29. *Jeffreys H.* The law of error and the combination of observations // Phil. Trans. Roy. Soc. London A.—1937.—**237**.—P. 231—271.
30. *Jeffreys H.* The law of error in the greenwich variation of latitude observations // Mon. Notic. Roy Astron. Soc.—1939.—**99**, N 9.—P. 703—709.
31. *Jeffreys H.* Theory of probability. — Sec. Edition. — Oxford, 1940.—468 p.
32. *Pearson K.* On the mathematical theory of errors of judgment, with special reference to the Personal equation // Phil. Trans. Roy. Soc. London A.—1902.—**198**.— P. 253— 296.
33. *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.
34. *Tukey J. W.* The future of data analysis // Ann. Math. Stat.—1967.—**33**.—P. 1—67.

Поступила в редакцию 02.08.10