КИНЕМАТИКА
И ФИЗИКА
НЕБЕСНЫХ
ТЕЛ том 27 № 5 2011

УДК 521.95:519

И. В. Джунь

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

Метод диагностики математических моделей в теоретической астрономии и астрометрии

На основе теории весовой функции определены области допустимого и недопустимого моделирования. Принадлежность к той или иной области определяется значениями пирсоновских характеристик асимметрии и эксцесса для разностей «observation – calculation».

МЕТОД ДІАГНОСТИКИ МАТЕМАТИЧНИХ МОДЕЛЕЙ В ТЕОРЕТИЧНІЙ АСТРОНОМІЇ І АСТРОМЕТРІЇ, Джунь Й. В. — На основі теорії вагової функції визначені області допустимого і недопустимого моделювання. Належність до тієї чи іншої області визначається значеннями пірсонівських характеристик асиметрії і ексцесу для різниць «observation – calculation».

A METHOD FOR DIAGNOSTICS OF MATHEMATICAL MODELS IN THEORETICAL ASTRONOMY AND ASTROMETRY, by Dzhun I. V. — The areas of termissible limit and imtermissible limit are defined on the basis of the theory of weight function. The belonging to that or another area is determined by values of Pearson's characteristics of the asymmetry and excess for the differences "observation – calculation".

Суть теории математического диагностирования моделей в астрономии, которую мы предлагаем, сводится к статистическому анализу разностей «observation calculation» ($O - C$) на основе весовой функции [2, 6], которая позволяет достаточно строго, просто и наглядно оценить корректность моделирования. Обычно эти разности получают по формуле

$$x_i \quad y_i - Y_i, \tag{1}$$

где y_i — результаты наблюдений, Y_i — модельные значения.

Впервые идею оценки согласия теории и практики на основе изучения особых свойств разностей $O - C$ высказал К. Пирсон [18]. Например, разности полученных из наблюдений и эфемеридных координат астероида будут статистически очень плохими, если математическая модель его орбиты ущербна с теоретической точки зрения, скажем по причине неучета возмущения его орбиты неизвестным телом.

Обширные исследования, проведенные научной школой Е. П. Фёдорова, показали, что в большинстве случаев разности $O - C$ не являются гауссовыми [3—5, 7—11, 14—16]. Поэтому для оценки значений Y_i в (1) мы должны применить метод максимального правдоподобия. Без потери общности в модели (1) мы оценим только одно значение Y на основании n наблюдений, минимизируя функцию правдоподобия

$$L = \prod_{i=1}^{n} f(x_i), \tag{2}$$

где $f(x_i)$ — плотность распределения в точке x_i. В даном случае Y — искомый параметр закона плотности.

Логарифмируя (2) и полагая в первом приближении, что L зависит только от Y, имеем

$$\frac{\partial \ln L}{\partial Y} = \sum_{i=1}^{n} \frac{f'(x_i)}{f(x_i)} = 0. \tag{3}$$

Умножая числитель и знаменатель в (3) на x_i, с учетом (1) получим оценку искомой величины Y, предполагая применение метода последовательных приближений:

$$Y = \frac{\sum y_i p(x_i)}{\sum p(x_i)}, \tag{4}$$

где весовая функция $p(x_i)$ определяется формулой

$$p(x_i) = -\frac{f'(x_i)}{x_i f(x_i)}. \tag{5}$$

Заметим, что теорию весовой функции в виде (5) впервые предложили астрономы Х. Р. Хюльме и Л. С. Т. Симс в 1939 г. [17]. Однако они не нашли аналитического выражения для вычисления весовой функции и вычисляли веса $p(x_i)$ при помощи транспортира и линейки по графику сглаженного эмпирического распределения ошибок.

Для получения общего аналитического выражения весовой функции (5) воспользуемся дифференциальным представлением семейств распределений Пирсона [1]:

$$\frac{f'(x)}{f(x)} = -\frac{x + c_1}{c_0 + c_1 x + c_2 x^2}, \tag{6}$$

где началом отсчета для x служит среднее значение, а величины c_0, c_1, c_2 связаны простыми соотношениями с асимметрией A распределения и моментным отношением β_2 [1]:

$$c_0 \quad \frac{^2(4 \ _2 - 3 \ _1)}{b},$$
$$c_1 \quad \frac{\sqrt{\ _1} (\ _2 \quad 3)}{b}, \qquad (7)$$
$$c_2 \quad \frac{2 \ _2 - 3 \ _1 - 6}{b}.$$

Здесь

$$_1 \quad A^2 \quad ^2_3 / \ ^3_2, \qquad _2 \quad _4 / \ ^2_2,$$
$$b \quad 2(5 \ _2 - 6 \ _1 - 9), \qquad _r \quad \int_{l_1}^{l_2} x^r f(x)dx,$$

r = 2, 3, 4, $_0$ = 1, $_1$ = 0, а l_1 и l_2 — нижняя и верхняя границы естественной области определения плотности $f(x)$.

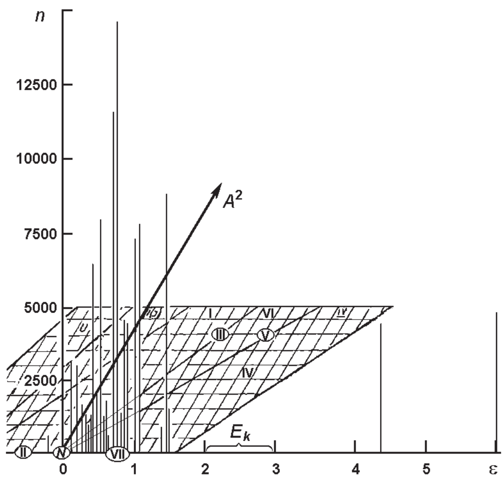

Рис. 1. Расположение эмпирических распределений ошибок и разностей O – – C астрономических, космических, гравиметрических, геодезических, экономических рядов (область E_k) на оси эксцесса (линия VII, правее точки N, соответствует семейству распределений Пирсона VII типа)

Вопрос о том, насколько адекватно представление разностей O — C семейством распределений Пирсона был рассмотрен нами в работе [9]. Для такой проверки был использован график для идентификации типов распределений Пирсона, взятый из таблиц [1] (рис. 1). В плоскости A^2O различным типам кривых Пирсона соответствуют области, кривые, прямые и точки. Значения статистик $_1$ и $_2$ позволяют определить по этому графику тип функции распределения Пирсона и практически проверить гипотезу о совпадении той или иной пирсоновской кривой с гистограммой. Каждое эмпирическое распределение ошибок и разностей O – C на рис. 1 характеризуется тремя координатами: эксцессом , квадратом коэффициента асимметрии A^2 и обьемом выборки n. Нормальному распределению соответствует точка N начало координат. Видно, что практически все эмпирические распределения ошибок и разностей O – C находятся в области кривых Пирсона. Преобладающая масса распределений группируется справа от точки N на линии кривой Пирсона VII типа; два распределения на графике слабо смещены от линии VII в область распределений Пирсона IV типа и одно в область распределений Пирсона I типа.

Только два ряда ошибок гринвичских определений широты на зенит-телескопе Куксона в 1927—1931 гг. и 1932—1936 гг. соответственно имеют эксцессы 4.3 и 6.0 и не попадают в область кривых распределений Пирсона. Однако эти наблюдения не являются типичными для классической астрометрии, так как этот телескоп плавает на ртути и очень чувствителен к малейшому движению воздуха. Точность наблюдений на этом приборе подвержена значительным флюктуациям во времени, что и приводит к высоким эксцессам [1].

Подставляя (7) в (6), получаем следующее общее аналитическое выражение для весовой функции семейств распределений Пирсона:

$$p(x) \quad \frac{x \quad c_1}{x(c_0 \quad c_1 x \quad c_2 x^2)} \quad \frac{1}{c_0 \quad c_1 x \quad c_2 x^2} \quad \frac{c_1}{x(c_0 \quad c_1 x \quad c_2 x^2)}. \qquad (8)$$

В классическом случае для закона Гаусса ($_1 = 0$, $_2 = 3$) имеем из (7) $c_0 = {}^2$, $c_1 = 0$, $c_2 = 0$, и весовая функция приобретает вид константы: $p(x) = {}^2$.

Для симметричных семейств Пирсона $_1 = 0$, $c_1 = 0$ и на основании (8) имеем:

$$p(x) \quad \frac{1}{c_0 \quad c_2 x^2} \quad \frac{5 \;_2 - 9}{2 \;_2{}^2 \quad (\;_2 - 3)x^2} = \frac{5 \quad 6}{2 \;{}^2 \;_2 \quad x^2}, \qquad (9)$$

где эксцесс .

График весовой функции (9) для симметричных распределений Пирсона представлен на рис. 2. Классическая прямая $p(x) = {}^2$ делит на этом рисунке пространство весовой функции на две области.

Область А — область допустимого моделирования, для которой весовая функция является либо постоянной, либо неограниченно убывает с увеличением $|x|$.

Область В — область недопустимого моделирования, так как весовая функция, приписывает больший вес всё большим ошибкам x, что является метрологическим абсурдом. В данном случае весовая

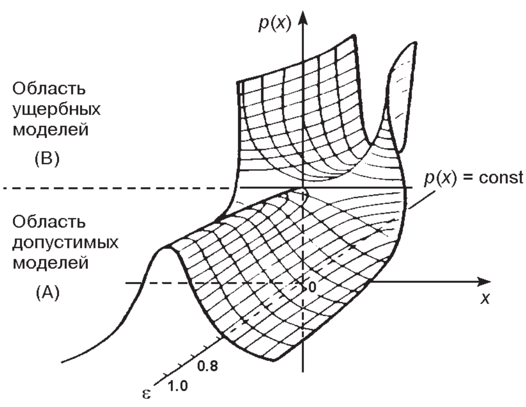

Рис. 2. Фрагмент поверхности весовой функции $p(x)$ для симметричных распределений Пирсона. Линия $p(x)$ = = const делит пространство весовой функции на две различные области

функция свидетельствует о нашем вторжении в опасную область несостоятельного моделирования, так как веса становятся бесконечно большими с увеличением ошибки наблюдения.

При наличии обоих слагаемых в формуле (8) имеем множество несимметричных семейств весовых функций. Если ошибка наблюдения $x = 0$, мы получаем бесконечные веса. Следовательно, ассимметрия распределения ошибок приводит нас в область недопустимого моделирования *C*, которая определяется наличием члена c_1 в формуле (8). Асимметрия значений *O* – *C* свидетельствует о крайне опасной области моделирования, делающей невозможной какое-либо оценивание, так как вблизи значений $x = 0$ мы получаем сверхэффективные веса и асимметричность функции (8) (даже ее разрыв при < 3). Уместно заметить, что ещё Д. И. Менделеев интуитивно считал отсутствие асимметрии распределения ошибок хорошим критерием согласованности ряда измерений и отсутствия в них систематических ошибок [13].

Основной вывод настоящей работы состоит в том, что модель может быть признана математически состоятельной только в том случае, если весовая функция попадает в область А и только в А, когда распределение разностей *O* – *C* имеет эксцесс 0 и $_1 = 0$. Если же фактическое значение эксцесса существенно меньше нуля, или же $_1$ существенно отличается от нуля, или имеет место то и другое, то либо эксперимент поставлен плохо, либо модель ущербна и не учитывает некоторые существенные факторы наблюдений.

Основным результатом настоящей работы есть следующий вывод: статистики $_1$ и $_2$, определяющие вид весовой функции (8), играют решающую роль в диагностике математического моделирования на современном этапе развития астрономии. Математическое моделирование является состоятельным только в том случае, когда весовая функция разностей *O* – *C* не попадает в области B и C, т. е. когда она принадлежит области А на рис. 2:

$$= 0. \qquad (10)$$

Области недопустимого моделирования характеризуются условиями

$$\text{и} \quad = 0, \qquad (11)$$

или ситуацией, когда при любом эксцессе

Практически процедура диагностики математического моделирования сводится к построению доверительных интервалов к статистикам и *A*, дисперсии которых известны [12, с. 423]:

$$_2 \quad \sqrt{\frac{24n(n \quad 1)^2}{(n \quad 3)(n \quad 2)(n \quad 3)(n \quad 5)}}, \qquad (13)$$

$$_A \quad \sqrt{\frac{6n(n \quad 1)}{(n \quad 2)(n \quad 1)(n \quad 3)}}, \qquad (14)$$

Для диагностики моделирования достаточно найти 90 % доверительные интервалы для   и A по формулам

$$1.645 \quad {}_{2}, \qquad A \quad 1.645 \quad {}_{A}, \tag{15}$$

где квантиль $t_{10\,\%} = 1.645$.

Для большей точности вычисления   и A мы рекомендуем в выборочные моменты m порядка $r = 2$, $r = 4$ вводить поправки Шеппарда [12, с. 396]:

$$M_2^R \quad m_2 \quad \frac{h^2}{12}, \tag{16}$$

$$M_4^R \quad m_4 \quad \frac{1}{2} m_2 h^2 \quad \frac{7}{240} h^4, \tag{17}$$

где h — интервал группирования.

Таким образом, математическая модель может быть признана состоятельной лишь в том случае, если доверительный интервал для   накрывает точку   , или же он весь смещён вправо от этой точки в область положительных эксцессов. В то же время доверительный интервал для A **обязательно** должен накрывать точку $A = 0$.

Несостоятельной математическая модель является в случае, если имеет место хотя бы одна из следующих ситуаций: 1) весь доверительный интервал для   смещен влево от точки   , т. е. его верхняя граница   + 1.645 $_2$ < 0; 2) доверительный интервал для A не накрывает точку $A = 0$.

Что делать, если модель оказалась несостоятельной? Есть два пути решения этой проблемы:1) создание более адекватных математических моделей исследуемых явлений, 2) усовершенствование наблюдений путем исключения систематических ошибок и увеличения их точности.

1. *Большев Л. Н., Смирнов Н. В.* Таблицы математической статистики. — М.: ВЦ АН СССР, 1968.—474 с.

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

3. *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 Президиуме АН СССР и НПО «Нефтегеофизика», 1983.—С. 46—52.

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

5. *Джунь И. В.* Новый класс распределений для аппроксимации эмпирических рядов ошибок астрономических наблюдений // Современная астрометрия: Тр. 23-й астрометр. конф. СССР. — Л.: Наука, 1987.—С. 373—378.

6. *Джунь И. В.* Теория веса геодезического измерения, построенная на принципе правдоподобия // Геодезия, картография и аэрофотосъемка.—1988.— № 47.— С. 9—13.

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

8. *Джунь И. В.* Теория больших выборок и её значение в современной астрометрии // Вивчення геодинамічних процесів методами астрономії, геодезії і геофізики: Тез. доп. Міжнар. конф., Полтава, 9—12 жовтня 2001 р. — Полтава, 2001.— С. 21—22.

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., Арнаутов Г. П., Стусь Ю. Ф., Щеглов С. Н.* Особенность закона распределения результатов баллистических измерений ускорения силы тяжести // Повторные гравиметрические наблюдения. — М.: Изд-во МГК при Президиуме АН СССР и НПО «Нефтегеофизика», 1984.—С. 87—100.

11. *Джунь И. В., Славинская А. А.* Обработка наблюдений на астролябии Данжона с учетом эксцесса закона ошибок остаточных погрешностей // Изучение Земли как планеты методами геофизики, геодезии и астрономии: Тр. II Орловской конф. (Полтава, 29.09—03.10.1986 г.). — Киев: Наук. думка, 1988.— С. 222—226.

12. *Крамер Г.* Математические методы статистики. — М.: Мир, 1975.—648 с.

13. *Маликов М. В.* Основы метрологии. — М.: Комитет по делам мер и измерительных приборов при СМ СССР, 1979.—480 с.

14. *Харин А. С., Яцкив Я. С.* Изучение ошибок наблюдений Голосеевского каталога звезд широтных программ // Астрометрия и астрофизика.—1970.— Вып. 10.— С. 34—43.

15. *Branham R. L.* Techniques for dealing with discordant observations // Relativity in Celestial Mechanics and Astrometry. — Reidel, 1986.—P. 229—230.

16. *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.

17. *Hulme H. R., Syms L. S. T.* The law of errors and the combinations of observations // Mon. Notic. Roy Astron. Soc.—1939.—**99**, N 8.—P. 642—658.

18. *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. 235—296.

Поступила в редакцию 02.08.10