УДК 311+ 311.214

М. М. Таенват[1]

Владивостокский государственный университет экономики и сервиса
Владивосток. Россия

Степенная регрессия на примере третьего закона Кеплера

В статье исследуется степенная регрессия. Представлены четыре способа вывода функционального уравнения из нормальных уравнений; получение нормальных уравнений из функционального уравнения путём дифференцирования; дисперсионный анализ; экстраполяция в степенной регрессии; повышение точности формулы третьего закона Кеплера; число Эйлера; статистическая надёжность, F-статистика; интерполирование вероятностей.

Ключевые слова и словосочетания: формула третьего закона Кеплера, функциональное уравнение, нормальные уравнения, дифференцирование, дисперсионный анализ, экстраполяция, число Эйлера, статистическая надёжность, F-статистика, интерполирование вероятностей.

M. M. Taenvat

Vladivostok State University of Economics and Service
Vladivostok. Russia

Power regression as an example of the third law of Kepler Planets

All about Power regression: four ways to display a functional equation of the normal equations; obtain the normal equations from the functional equation by differentiating; analysis of variance; extrapolation of the power of the regression improve the accuracy of the formula Kepler's third law; Euler number; statistical reliability, F-statistics; interpolation of probabilities.

Keywords: formula of Kepler's third law, functional equation, normal equation, differentiation, analysis of variance, extrapolation, Euler number, statistical reliability, F-statistics, interpolation of probabilities.

Использование регрессии в физике небесных тел. Устойчивое отношение к статистике как к скучному предмету поможет изменить данная статья. В сознании большинства статистика предстает лишь одной стороной – подсчётом изготовленных изделий, съеденных продуктов и т.п. Но когда такие подсчёты могут привести к открытию мирового значения, статистика захватывает дух!

[1] Таенват Михаил Михайлович – аспирант кафедры государственного и муниципального управления института права и управления ВГУЭС; e-mail: hh.dd.00@mail.ru.

Кеплер Джон, немецкий математик, астроном, механик, оптик и астролог, первооткрыватель законов движения планет Солнечной системы, открыл третий закон движения планет, за что ныне мог бы получить Нобелевскую премию… Сейчас мы можем за полчаса повторить его путь, уяснив, как это было.

Автор статьи «Законы Кеплера» Крис Импи [10] отметил: «Законы Кеплера применяются к любому орбитальному движению, будь то планеты вокруг Солнца, Луны вокруг Земли или звезды вокруг центра галактики. Второй и третий законы были результатом попыток Кеплера найти закономерности в орбитах планет. Второй и третий законы Кеплера изучают математические отношения между расстоянием планеты от Солнца и скорости её движения вокруг Солнца. Оба они являются последствиями применения закона всемирного тяготения Ньютона и закона сохранения момента импульса объекта, движущегося по эллиптической траектории, но Кеплер удивительно смог получить их без любого из этих понятий!».

Но суть и в том, что количественные действия в любых предметах после математической обработки могут обернуться важным открытием и для Вас. Каких же любых предметах? Л. Закс производил подсчёты в медико-биологической лаборатории, Вы же можете считать в любой иной лаборатории. Л. Закс пишет: «Для оценивания параметров по выборочным данным разработаны многочисленные методы. Особое значение имеет метод максимального правдоподобия (Р.А. Фишер); это универсальный метод максимального оценивания неизвестных параметров, применимый в случаях, когда вид функции распределения известен; оценки неизвестных параметров в этом случае равны значениям, при которых полученная выборка имеет максимальную вероятность появления, т. е. в качестве оценок отыскиваются значения, максимизирующие функцию максимального правдоподобия для параметров, в предположении, что эти параметры существуют. Этот метод построения точечных оценок параметров находится в тесной связи с методом наименьших квадратов» [2. С. 71–72].

Мэтьюз и Финк [3] приводят прекрасный пример использования уравнения регрессии: «Техника численного анализа в науке и в инженерном деле часто применяется для вычерчивания графиков экспериментальных данных. В 1601 г. немецкий астроном Джон Кеплер (Johannes Kepler) сформулировал третий закон движения планет, где x – расстояние до Солнца, измеряемое в миллионах километров, T – период прохождения по орбите в днях и C – постоянная. Наблюдения пар данных (x, T) для первых четырёх планет, Меркурия, Венеры, Земли и Марса дали итоги (58; 88), (108; 225), (150; 365), (228; 687); *коэффициент С*, полученный методом наименьших квадратов, равен С=0,199769… В этом случае функция имеет вид».

Авторы приводят степенную подгонку: «Предположим, что точек с различными абсциссами».

$$E(A)=\sum_{k=1}^{N}(Ax^{M}-y)^{2}.$$

Поскольку имеем лишь одну переменную *A*, взятие частных производных не требуется.

$E(A)=\sum_{k=1}^{N}u^{2}$, *если* $u=Ax^{M}-y$. *Отсюда* $\frac{du}{dA}=\sum_{k=1}^{N}1\times A^{1-1}\times x^{M}-0=\sum_{k=1}^{N}x^{M}$; $\frac{dE(A)}{du}=\sum_{k=1}^{N}2u^{1}$.

Соответственно $\frac{dE(A)}{dA}=\frac{dE(A)}{du}\times\frac{du}{dA}=\sum_{k=1}^{N}2u^{1}\times\sum_{k=1}^{N}x^{M}=\sum_{k=1}^{N}2(Ax^{M}-y)^{1}\times(x^{M})=$

$=2\sum_{k=1}^{N}(Ax^{M+M}-yx^{M})=2(Ax^{2M}-yx^{M})$; $E^{'}(A)=0$; $0=\sum_{k=1}^{N}(Ax^{2M}-yx^{M})=A\sum_{k=1}^{N}x^{2M}-\sum_{k=1}^{N}yx^{M}$;

$A\sum_{k=1}^{N}x^{2M}=\sum_{k=1}^{N}yx^{M}$.

Отсюда коэффициент *A* кривой, построенной методом наименьших квадратов (табл. 1), $y=Ax^{M}$, равен

$$A=\frac{\sum_{k=1}^{N}x_k^{M}y_k}{\sum_{k=1}^{N}x_k^{2M}}.$$

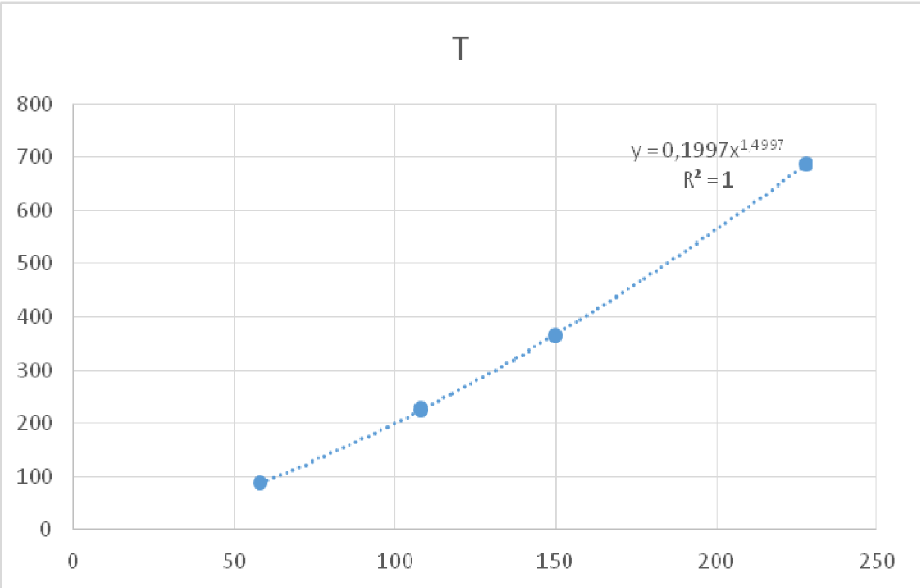

Рис. 1. Кривая, полученная методом наименьших квадратов, $T=0{,}199769x^{3/2}$, вычерчена для первых четырёх планет, которые использовал Кеплер для выведения третьего закона движения планет [3]

Таблица 1

Исходные данные для коэффициента А

x	y	x^3/2	(x^(3/2))*y	x^(2*(3/2))
58	88	441,7148	38870,906	195112
108	225	1122,369	252533,01	1259712
150	365	1837,117	670547,82	3375000
228	687	3442,725	2365151,7	11852352

3327103,5/16682176=0,199440616.

В описанном случае выводится только коэффициент *A*, коэффициент *M* уже известен. Это сберегает время, когда *A* – уже известная физическая, иная постоянная. Но нас интересует, как получен и коэффициент *M*. И здесь приходит на помощь таблица готовых нормальных уравнений (табл. 2) для наиболее важных функциональных уравнений книги по медико-биологической статистике, написанной Л. Заксом.

Таблица 2

Нормальные уравнения для наиболее важных функциональных уравнений

Функциональные уравнения	Нормальные уравнения
$y = a + b \times x$	$a \times n + b\sum x = \sum y$ $a\sum x + b\sum x^2 = \sum (x \times y)$
$\lg_{10} y = a + b \times x$	$a \times n + b\sum x = \sum \lg_{10} y$ $a\sum x + b\sum x^2 = \sum (x \times \lg_{10} y)$
$y = a + b \times \lg_{10} x$	$a \times n + b\sum \lg_{10} x = \sum y$ $a\sum \lg_{10} x + b\sum (\lg_{10} x)^2 = \sum (y \times \lg_{10} x)$
$\lg_{10} y = a + b \times \lg_{10} x$	$a \times n + b\sum \lg_{10} x = \sum \lg_{10} y$ $a\sum \lg_{10} x + b\sum (\lg_{10} x)^2 = \sum (\lg_{10} x \times \lg_{10} y)$
$y = a \times b^x$ $\lg_{10} y = \lg_{10} a + x \times \lg_{10} b$	$n \times \lg_{10} a + \lg_{10} b\sum x = \sum \lg_{10} y$ $\lg_{10} a\sum x + \lg_{10} b\sum x^2 = \sum (x \times \lg_{10} y)$
$y = a + b \times x + c \times x^2$	$a \times n + b\sum x + c\sum x^2 = \sum y$ $a\sum x + b\sum x^2 + c\sum x^3 = \sum xy$ $a\sum x^2 + b\sum x^3 + c\sum x^4 = \sum (x^2 y)$

Окончание табл. 2

$y = a + b \times x + c \times \sqrt{x}$	$n \times \lg_{10} a + b\sum x + c\sqrt{x} = \sum y$ $a\sum x + b\sum x^2 + c\sum\sqrt{x^3} = \sum(xy)$ $a\sum\sqrt{x} + b\sum\sqrt{x^3} + c\sum x = \sum\sqrt{xy}$
$y = a \times b^x \times c^{x^2}$ $\lg_{10} y = \lg_{10} a + x \times \lg_{10} b + x^2 \lg_{10} c$	$n \times \lg_{10} a + \lg_{10} b\sum x + \lg_{10} c\sum x^2 = \sum \lg_{10} y$ $\lg_{10} a\sum x + \lg_{10} b\sum x^2 + \lg_{10} c\sum x^3 = \sum(x \times \lg_{10} y)$ $\lg_{10} a\sum x^2 + \lg_{10} b\sum x^3 + \lg_{10} c\sum x^4 = \sum(x^2 \times \lg y)$

Дополним таблицу Закса:

Таблица 3

Дополненения к таблице 2

Функциональные уравнения	Нормальные уравнения
$\ln_e y = a + b \times \ln_e x$	$an + b\sum \ln_e x = \sum \ln_e y;$ $a\sum \ln_e x + b\sum(\ln_e x)^2 = \sum(\ln_e x \times \ln_e y)$
$lb_2 y = a + b \times lb_2 x$	$an + b\sum lb_2 x = \sum lb_2 y;$ $a\sum lb_2 x + b\sum(lb_2 x)^2 = \sum(lb_2 x \times lb_2 y)$

Логарифмы по разным основаниям взаимно переходят друг в друга. Приводим абсолютно точные пересчёты и их формулы:

$\log_{10} x = \log_{10} e \times \ln_e x = \log_{10} 2{,}718281828459 04 \times \ln_e x \cong 0{,}434294482 \times \ln_e x$

4,060443×0,434294482=1,763427989;

$\ln_e x = \ln_e 10 \times \log_{10} x \cong 2{,}302585093 \times \log_{10} x$

1,763428×2,302585=4,060443025.

Логарифм по основанию 2 – lb (binär):

$$lbx = \frac{\log_{10} x}{\log_{10} 2} = \frac{1}{0{,}301029996}\log_{10} x = 3{,}321928095\log_{10} x$$

1,763428×3,321928=5,857981017;

$$lbx = \frac{\ln_e x}{\ln_e 2} = \frac{1}{0{,}693147181}\ln_e x = 1{,}442695041\ln_e x$$

4,060443×1,442695=5,85798098.

Таблица 4

Первый вариант исходных данных для коэффициентов нормальных уравнений

x	y	lnx	lny	ln(x^2)	lnx*lny
1	2	3	4	5	6
58	88	4,060443	4,477337	16,4872	18,17997

Окончание табл. 4

1	2	3	4	5	6
108	225	4,682131	5,4161	21,92235	25,35889
150	365	5,010635	5,899897	25,10646	29,56223
228	687	5,429346	6,532334	29,4778	35,4663

$an + b\sum \ln_e x = \sum \ln_e y;$

$a\sum \ln_e x + b\sum (\ln_e x)^2 = \sum (\ln_e x \times \ln_e y).$

4a+19,18256b=22,32567;

19,18256a+92,99381b=108,5674.

Систему уравнений на выбор можно решить методами исключения Гаусса, разложения на треугольные матрицы либо матричным, но, экономя место, мы используем калькулятор уравнений: Ответ: a = -64515504413/40046318464=-1,611022109, b = 3753809803/2502894904=1,499787225.

EXP(-1,611022109)=0,199683412

Используя дифференцирование, покажем, как выведена система уравнений:

$$E(a,b) = \sum_{k=1}^{N} (a + b \times \lg_{10} x - \lg_{10} y)^2.$$

Поскольку имеем две переменных *a* и *b*, возьмём частные производные:

$$E(a,b) = \sum_{k=1}^{n} u^2, \text{ если } u = a + b \times \lg_{10} x - \lg_{10} y. \text{ Отсюда } \frac{\partial u}{\partial a} = \sum_{k=1}^{n} 1 \times a^{1-1} + 0 - 0 = \sum_{k=1}^{n} 1 \quad \text{и}$$

$$\frac{\partial E(a,b)}{\partial u} = \sum_{k=1}^{n} 2u^{2-1} = \sum_{k=1}^{n} 2u;$$

$$\text{Соответственно } \frac{\partial E(a,\, b)}{\partial a} = \frac{\partial E(a,\, b)}{\partial u} \times \frac{\partial u}{\partial a} = \sum_{k=1}^{n} 2u^1 \times \sum_{k=1}^{n} 1 = \sum_{k=1}^{n} 2(a + b \times \lg_{10} x - \lg_{10} y) \ \times (1) =$$

$$= 2\sum_{k=1}^{N} (a + b \times \lg_{10} x - \lg_{10} y);\ E^{'}(A) = 0;\ 0 = \sum_{k=1}^{n} (a + b \times \lg_{10} x - \lg_{10} y) = a \times n + b\sum_{k=1}^{n} \lg_{10} x - \sum_{k=1}^{n} y;$$

$$a \times n + b\sum_{k=1}^{n} \lg_{10} x = \sum_{k=1}^{n} \lg_{10} y.$$

Зафиксируем *a* и продифференцируем *E (a, b)* по *b*.

Получим

$$E(a,b) = \sum_{k=1}^{n} u^2, \text{ если } u = a + b \times \lg_{10} x - \lg_{10} y. \text{ Отсюда } \frac{\partial u}{\partial b} = \sum_{k=1}^{n} 0 + 1 \times b^{1-1} \times \lg_{10} x - 0 = \sum_{k=1}^{n} \lg_{10} x;$$

$$\text{Соответственно}: \frac{\partial E(a,\, b)}{\partial b} = \frac{\partial E(a,\, b)}{\partial u} \times \frac{\partial u}{\partial b} = \sum_{k=1}^{n} 2u^1 \times \sum_{k=1}^{n} \lg_{10} x =$$

$$= \sum_{k=1}^{n} 2(a + b \times \lg_{10} x - \lg_{10} y) \ \times (\lg_{10} x) = \sum_{k=1}^{n} 2(a \lg_{10} x + b \lg_{10} x^{1+1} - \lg_{10} y \times \lg_{10} x);$$

$$E'(A)=0;\quad 0=\sum_{k=1}^{n}(a+b\times\lg_{10}x-\lg_{10}y)=a\times n+b\sum_{k=1}^{n}\lg_{10}x-\sum_{k=1}^{n}y;$$

$$a\times n+b\sum_{k=1}^{n}\lg_{10}x=\sum_{k=1}^{n}\lg_{10}y.$$

Таблица 5

Второй вариант исходных данных для коэффициентов нормальных уравнений

lgx	lgy	lg(x^2)	lgx*lgy
1,763428	1,944483	3,109678	3,428955
2,033424	2,352183	4,134812	4,782984
2,176091	2,562293	4,735373	5,575783
2,357935	2,836957	5,559857	6,689359

$an+b\sum\lg_{10}x=\sum\lg_{10}y;$

$a\sum\lg_{10}x+b\sum(\lg_{10}x)^2=\sum(\lg_{10}x\times\lg_{10}y).$

4a+8,330878b=9,695915;

8,330878a+17,53972b=20,47708.

Используем калькулятор уравнений: Ответ: *a* = -132105258110/ 188837937279=-0,699569483, *b* = 566417518315/377675874558= 1,499745037. $10^{-0,699569483}=0,19972412$.

Таблица 6

Третий вариант исходных данных для коэффициентов нормальных уравнений

log2(x)	log2(y)	log2(x^2)	log2(x)*log2(y)
5,857981	6,459432	34,31594	37,83923
6,754888	7,813781	45,62851	52,78121
7,228819	8,511753	52,25582	61,52992
7,83289	9,424166	61,35417	73,81846

$an+b\sum\log_2 x=\sum\log_2 y;$

$a\sum\log_2 x+b\sum(\log_2 x)^2=\sum(\log_2 x\times\log_2 y).$

4a+27,67458b=32,20913;

19,18256a+193,5544b=225,9688.

Используем калькулятор уравнений:

Ответ: *a* = -48432003580/20838054559=-2,324209462,

b= 62505275423/41676109118=1,49978673,

$2^{-2,324209462}=0,199683985.$

Далее проведём обработку той же выборки с помощью программы Microsoft Office Excel, существо некоторых её показателей мы рассмотрим ниже. Суть прочих показателей хорошо описана в [1].

Таблица 7

Сводная регрессионного анализа зависимости периода прохождения первых четырёх планет, которые использовал Кеплер, от расстояния до Солнца (экспоненциальная модель)

ВЫВОД ИТОГОВ	
Регрессионная статистика	
Множественный R	0,999991
R-квадрат	0,999981
Нормированный R-квадрат	0,999972
Стандартная ошибка	0,004598
Наблюдения	4

ДИСПЕРСИОННЫЙ АНАЛИЗ

–	df	SS	MS	F	Значимость F
Регрессия	1	2,251954	2,251954	106513,2	9,39E-06
Остаток	2	4,23E-05	2,11E-05		
Итого	3	2,251996			

Окончание табл. 7

-	Коэффициенты	Стандартная ошибка	t-статистика	P-значение	Нижние 95%
Y-пересечение	-1,61083	0,022157	-72,7004	0,000189	-1,70617
ln(x)	1,499748	0,004595	326,3635	9,39E-06	1,479976

ВЫВОД ОСТАТКА

Наблюдение	Предсказанное ln(y)	Остатки
1	4,478809	-0,00147
2	5,411184	0,004916
3	5,903858	-0,00396
4	6,531818	0,000516

Таблица 8

Предельные значения F и t-статистик

Односторонний критерий	
Тестовая F-статистика	18,51282051
Двусторонний t-критерий	
Нижняя доверительная граница	-4,30265273
Верхняя доверительная граница	4,30265273

Поскольку 106513,2>18,51282051 и 9,39E-06<0,05, с 95% надёжностью нулевая гипотеза отклоняется. И далее, поскольку 326,3635>4,30265273, а 9,39E-06<0,05, нулевая гипотеза отклоняется окончательно.

Рис. 2. График остатков

Гомоскедастичность не выявлена.

Математическая модель примет вид:

$$\ln \hat{Y}_{i_e} = -1{,}61083 + 1{,}499748 \times \ln_e X_i;$$

$$\hat{b}_0 = 2{,}71828182845904^{-1{,}61083} = 0{,}199721776;$$

$$\hat{b}_1 = 1{,}499748;$$

$$\ln \hat{Y}_{i_e} = b_0 + b_1 \times X_i;$$

$$\hat{Y}_i = \hat{b}_0 \hat{X}_i^{b_1};$$

$$\hat{Y}_i = 0{,}199721776 \times \hat{X}_i^{1{,}499748}.$$

Проверка:

$$\hat{Y}_{Земли} = 0{,}199721776 \times 150^{1{,}499748} = 366{,}4493 \textit{дня}$$

Предсказание периода прохождения по орбите Юпитера:

$$\hat{Y}_{Юпитера} = 0{,}199721776 \times 778{,}33^{1{,}499748} = 4329{,}547 \textit{дней};$$

$$4329{,}547 / 365 = 11{,}8617726 \textit{ Земных лет.}$$

4329,547/365=11,8617726 лет, вмещающих 2 високосных года, а поскольку 2 дня составляют 0,005479452 года, то 11,8617726+-+0,005479452=11,86725205 года.

Сайт «Большая энциклопедия школьника» сообщает, что период обращения по орбите самой большой планеты Юпитер составляет 11,867 лет! Кеплер и для сегодняшнего дня сделал прекрасный расчёт.

Мэтьюз и Финк [3. С. 290] в упражнениях к главе «Построение кривой по точкам» приводят современные данные, которые мы и пустим в обработку. Авторы дают указание: «Ниже приведены измерения расстояний девяти планет от Солнца и их звёздный период в днях. Найдите степенную подгонку вида для (a) первых четырёх планет и (b) всех девяти планет».

Таблица 9

Измерения расстояний девяти планет от Солнца и их звёздный период, дн.

Планета	Расстояние от Солнца (км 10^6)	Звёздный период, дни
Меркурий	57,59	87,99
Венера	108,11	224,7
Земля	149,57	365,26
Марс	227,84	686,98
Юпитер	778,14	4332,4
Сатурн	1427	10759
Уран	2870,3	30684
Нептун	4499,9	60188
Плутон	5909	90710

Примечание: сост. по [3. С. 290].

Выполним задание «а»:

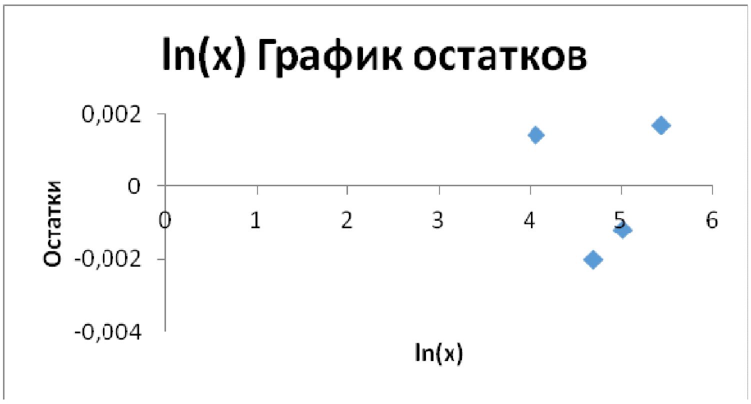

Рис. 3. График остатков

Таблица 10

Сводная регрессионного анализа зависимости периода прохождения первых четырёх планет, которые использовал Мэтьюз, от расстояния до Солнца (экспоненциальная модель)

ВЫВОД ИТОГОВ	
Регрессионная статистика	
Множественный R	0,999998
R-квадрат	0,999995
Нормированный R-квадрат	0,999993
Стандартная ошибка	0,00227
Наблюдения	4

Дисперсионный анализ

-	df	SS	MS	F	Значимость F
Регрессия	1	2,253078	2,253078	437091,8	2,29E-06
Остаток	2	1,03E-05	5,15E-06		
Итого	3	2,253088			

Окончание табл. 10

-	Коэф.	Стандартная ошибка	t-статистика	P-значение	Нижние 95%	Верхние 95%	Нижние 95,0%	Верхние 95,0%
Y-пересечение	-1,58024	0,010892	-145,09	4,75E-05	-1,62711	-1,53338	-1,62711	-1,53338
ln(x)	1,494081	0,00226	661,1292	2,29E-06	1,484358	1,503805	1,484358	1,503805

ВЫВОД ОСТАТКА

Наблюдение	Предсказанное ln(y)	Остатки
1	4,475789	0,001435
2	5,416761	-0,002
3	5,901763	-0,00115
4	6,530591	0,001714

Таблица 11

Предельные значения F и t-статистик

Односторонний критерий	
Тестовая F-статистика	18,51282051
Двусторонний критерий	
Нижняя доверительная граница	-4,30265273
Верхняя доверительная граница	4,30265273

Поскольку 437091,8>5,591447848 и 2,29E-06<0,05, то с 95%-й надёжностью нулевая гипотеза отклоняется. И далее, поскольку 661,1292>->2,364624251, а 2,29E-06<0,05, нулевая гипотеза отклоняется окончательно.

Итак, как и ожидалось, более, чем было возможно Кеплеру, точная формула: EXP(-1,58024)= 0,20592567; $T = 0{,}20592567\, x^{1{,}494081}$.

Выполним задание «b» (рис. 4):

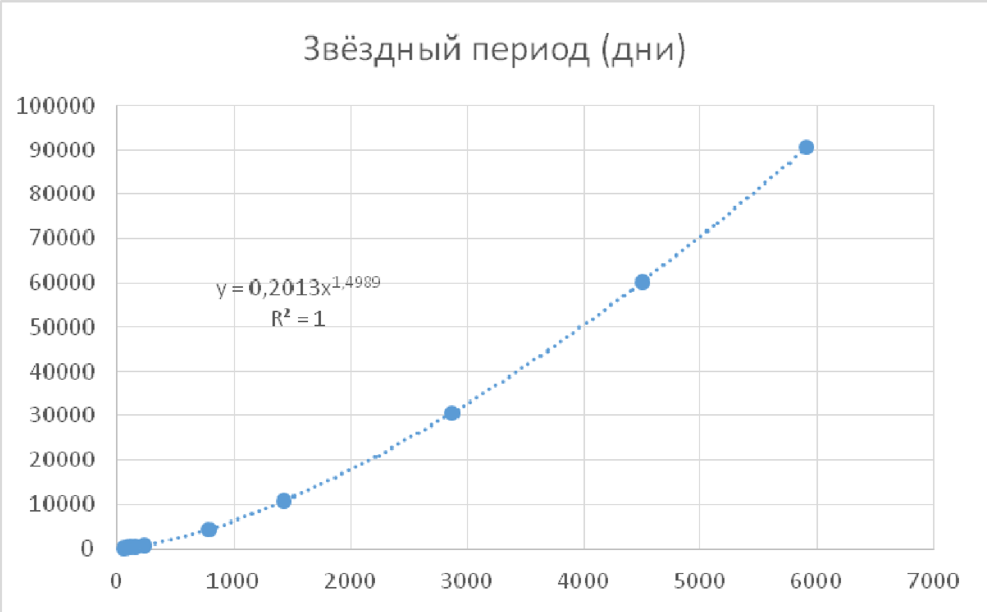

Рис. 4. Кривая, полученная методом наименьших квадратов, вычерчена для всех девяти планет, которые использовали Мэтьюз и Финк для выведения третьего закона движения планет

Таблица 12

Сводная регрессионной зависимости периода прохождения всех 9 планет по орбите от расстояния до Солнца (экспоненциальная модель) с использованием современных данных

ВЫВОД ИТОГОВ	
Регрессионная статистика	
Множественный R	0,9999996
R-квадрат	0,9999993
Нормированный R-квадрат	0,9999992
Стандартная ошибка	0,0023366
Наблюдения	9

Дисперсионный анализ

-	df	SS	MS	F	Значимость F
Регрессия	1	53,51453	53,51453	9801780	8,96E-23
Остаток	7	3,82E-05	5,46E-06		
Итого	8	53,51457			

Окончание табл. 12

	Коэффициенты	Стандартная ошибка	t-статистика	P-Значение	Нижние 95%
Y-пересечение	-1,603165	0,00319	-502,567	3,26E-17	-1,61071
ln(x)	1,4988974	0,000479	3130,779	8,96E-23	1,497765

ВЫВОД ОСТАТКА

Наблюдение	Предсказанное ln(y)	Остатки
1	4,4723886	0,004835
2	5,4163946	-0,00163
3	5,9029596	-0,00235
4	6,5338142	-0,00151
5	8,3748542	-0,00098
6	9,2838202	-0,00032
7	10,331313	0,000184
8	11,005275	-4,7E-05
9	11,413607	0,001816

Таблица 13

Предельные значения F и t-статистик

Односторонний критерий	
Тестовая F-статистика	5,591447848
Двусторонний критерий	
Нижняя доверительная граница	-2,364624251
Верхняя доверительная граница	2,364624251

Поскольку 9801780>5,591447848 и 8,96E-23<0,05, то с 95%-й надёжностью нулевая гипотеза отклоняется. И далее, поскольку 3130,779>2,364624251, а 8,96E-23<0,05, то нулевая гипотеза отклоняется окончательно.

Рис. 5. График остатков

Поскольку присутствует определённая гомоскедастичность, выводы требуют осторожности.

Итак, самая точная формула, что нам удалось рассчитать:

EXP(-1,603165)= 0,201258526;

$T = 0{,}201258526 \times x^{1{,}4988974}$

Значение F-статистики при 1 и 7 степенях свободы и величине ошибки равно 9819961,394. Но при $\alpha = 8{,}96 \times 10^{-23}$ оно равно уже 9801128,085; 1 неуспех приходится на $1{,}11607 \times 10^{22}$ успехов. Поскольку 9801128,085<9801780<<9819961,394 и 8,96E-23< $8{,}9 \times 10^{-23}$, с 0,999999999999999999999911 надёжностью нулевая гипотеза отклоняется. И если ещё мы используем первые 1000 знаков после запятой числа *e* (рис. 6), большинство из которых программа Microsoft Office Excel не использует, допуская, как пишет Мэтьюз, ошибку округления [3. С. 39], можно говорить о хороших шансах на успех в межпланетных полётах.

Третий закон Кеплера [9]: «Квадраты периодов обращения планет вокруг Солнца относятся, как кубы больших полуосей орбит планет» – справедлив не только для планет, но и для их спутников.

$$\frac{T^2}{a^3} = const;$$

где T — периоды обращения планет вокруг Солнца, a — длина больших полуосей их орбит.

Итак, формула третьего закона с нарастающей точностью:

$T = 0{,}199721776x^{1{,}499748}$;

$T = 0{,}20592567\, x^{1{,}494081}$;

$T = 0{,}201258526x^{1{,}4988974}$.

Произведём преобразования, получив формулу (2):

$$T = 0{,}201258526x^{1{,}4988974} = 0{,}201258526x^{1{,}4988974\times\frac{2}{2}} = 0{,}201258526 \times \sqrt{x^{2{,}9977948}};$$

$$T^2 = (0{,}201258526 \times \sqrt{x^{2{,}9977948}})^2 = 0{,}040504994 \times x^{2{,}9977948}; \quad \frac{T^2}{a^{2{,}9977948}} = 0{,}040504994.$$

Мэтьюз отмечает: «Важно сознавать, что численное решение – это не точное математическое решение. Точность численного решения уменьшается по многим причинам… Понимание этих трудностей часто может привести профессионала к правильному выполнению и/или усовершенствованию численного алгоритма» [3. С. 37-38].

Найдём абсолютную и относительную ошибки:

$$E_x = \left|x - \hat{x}\right| = 3 - 2{,}9977948 = 0{,}0022052;$$

$$R_x = \frac{\left|x - \hat{x}\right|}{x} = \frac{3 - 2{,}9977948}{3} = 0{,}000735067.$$

Итак, численно ошибка не подавляет. Как указывает Мэтьюз: «Ошибка может распространиться в последующих вычислениях». Далее мы можем полностью устранить эту ошибку, использовав формулу, но за это мы должны будем заплатить искажением коэффициента *A*.

e — математическая константа, основание натурального логарифма, трансцендентное число. Иногда число *e* называют *числом Эйлера или числом Непера* [8]. Натуральные логарифмы (ln) в качестве основания имеют постоянную

$$e \cong 2{,}718281 \ (\textit{предел ряда } e = 1 + \frac{1}{1} + \frac{1}{1\times 2} + \frac{1}{1\times 2\times 3} + \frac{1}{1\times 2\times 3\times 4} + \ldots = 2{,}708333333).$$

Что означает 99,9 999 999 999 999 999 999 911% надёжность? Это означает, что при бросании монеты 76-кратное появление герба ещё допускается как вероятное, в то время как семидесяти семикратное уже рассматривается как «сверхслучайное».

2,7182818284 5904523536 0287471352 6624977572 4709369995
9574966967 6277240766 3035354759 4571382178 5251664274
2746639193 2003059921 8174135966 2904357290 0334295260
5956307381 3232862794 3490763233 8298807531 9525101901
1573834187 9307021540 8914993488 4167509244 7614606680
8226480016 8477411853 7423454424 3710753907 7744992069
5517027618 3860626133 1384583000 7520449338 2656029760
6737113200 7093287091 2744374704 7230696977 2093101416
9283681902 5515108657 4637721112 5238978442 5056953696
7707854499 6996794686 4454905987 9316368892 3009879312
7736178215 4249992295 7635148220 8269895193 6680331825
2886939849 6465105820 9392398294 8879332036 2509443117
3012381970 6841614039 7019837679 3206832823 7646480429
5311802328 7825098194 5581530175 6717361332 0698112509
9618188159 3041690351 5988885193 4580727386 6738589422
8792284998 9208680582 5749279610 4841984443 6346324496
8487560233 6248270419 7862320900 2160990235 3043699418
4914631409 3431738143 6405462531 5209618369 0888707016
7683964243 7814059271 4563549061 3031072085 1038375051
0115747704 1718986106 8739696552 1267154688 9570350354

Первые 1000 знаков после запятой числа e[1]

(последовательность A001113 в OEIS)

Рис. 6. Первые 1000 знаков после запятой числа *e*

По теореме вероятностей для независимых событий вероятности равны:

$$P_{4x} = (\frac{1}{2})^4 = 0{,}06250 \rangle 0{,}05;$$

$$P_{5x} = (\frac{1}{2})^5 = 0{,}03125 \langle 0{,}05;$$

$$P_{16x} = (\frac{1}{2})^{16} = 0{,}000015288;$$

$$P_{17x} = (\frac{1}{2})^{17} = 0{,}00000762939;$$

$$P_{76x} = (\frac{1}{2})^{76} = 1{,}32349 \times 10^{-23};$$

$$P_{77x} = (\frac{1}{2})^{77} = 6{,}61744 \times 10^{-24}.$$

Другими словами, примерно 0, 0015 и 0, 00076%. Итак, статистическая надёжность 95% означает, что случайное появление обстоятельства столь же невероятно, как и событие, состоящее в выпадании герба подряд 17 раз. Вероятность того, что при n-кратном бросании монеты каждый раз будет выпадать герб, равна $(1/2)^n$ и приведена в табл. 14. Об этом хорошо сказано в [2].

Таблица 14

Вероятность *P* того, что при n-кратном бросании монеты каждый раз она выпадает одной и той же стороной, как модель случайного события

n	2^n	P	
		2^{-n}	Уровень
1	2	0,5	
2	4	0,25	
3	8	0,125	
4	16	0,0625	‹10%
5	32	0,03125	‹5%
6	64	0,01562	
7	128	0,00781	‹1%
8	256	0,00391	‹0,5%
9	512	0,00195	
10	1024	0,00098	≈0,1%
11	2048	0,00049	≈0,05%
12	4096	0,00024	
13	8192	0,00012	
14	16384	0,00006	‹0,01%
15	32768	0,00003	

Примечание: сост. по [2. С. 111].

Таблица 15

Дополнение вероятностей для третьего закона Кеплера

n	2^n	P	
		2^{-n}	Уровень
16	65536	0,000015288	
17	131072	0,00000762939	
76	$1{,}55579\times10^{22}$	$1{,}32349\times10\text{-}^{23}$	
77	$1{,}51116\times10^{23}$	$6{,}61744\times10^{-24}$	

И здесь мы можем показать, насколько важно знание бинара:

$lb_2\, 0{,}0000076293\,9 = -17.$

Лотар Закс указывает: «Если мы выбираем коэффициент таким, что высказывание в 95% окажется правильным и только в 5% неправильным, то мы говорим: со статистической надёжностью S в 95% доверительный интервал выборочной статистики содержит параметр генеральной совокупности». Подводя итог, нужно заметить: мы имеем 5%-ные шансы отвергнуть верный коэффициент уравнения и 95%-ные – принять тоже верный коэффициент.

Интерполирование вероятностей. Для чего нужен способ интерполирования значения уровня F-критерия для $v_1\, u\, v_2$ степеней свободы? [2. С. 152] В особых случаях, прежде всего тогда, когда исследуемое представляет опасность для жизни человека, необходимо принимать меньшие, чем α=0,001 вероятности ошибок. Так, например, при изготовлении вакцины требуется предельная константа сыворотки. Небезупречные измерения должны быть обнаружены и исключены.

Л. Закс отмечает: «Необоснованное принятие нуль-гипотезы «сыворотка в норме» означает опасную ошибку» [2. С. 114]. Нуль-гипотеза – гипотеза о том, что две совокупности, рассматриваемые с точки зрения одного или нескольких признаков, одинаковы, т. е. действительное различие равно нулю, а найденное из опыта отличие от нуля носит случайный характер. Среднее значение μ генеральной совокупности, оцениваемое на основании случайной выборки, не отличается от желаемого значения μ_0.

И далее Л. Закс пишет, что наука делает ячейки сети всё меньше с тем, чтобы постоянно выдвигать и проверять всё новые гипотезы, наиболее точно и наиболее правдоподобно объясняющие этот мир. Получающиеся при этом выводы и заключения никогда не будут абсолютно надёжны, но они ведут от предварительных гипотез ко всё более общим и строгим теориям, выдерживающим тщательную проверку, приводят к лучшему познанию мира и двигают науку [2. С. 112].

Определённое сводной таблицей значение F=106513,2 $(\nu_1 = 1 u \nu_2 = 2)$ расположим таким образом между двумя табличными значениями (F_1, F_2) с вероятностями ошибок $\alpha u \alpha \times m, что\, F_1 \langle F \langle F_2$ и подберём вероятность того, что это значение будет превышено.

$$(m.e.\alpha = 0000090909\,1, m = \frac{000952381}{000909091} = 1{,}047619048):$$

$$F_1 = 104998{,}4947 \langle F = 106513{,}2 \langle 109998{,}489;$$

$$k = \frac{F_2 - F_1}{F_2 - F_1} = \frac{109998{,}489 - 106513{,}2}{109998{,}489 - 104998{,}4947} = 0{,}697058594;$$

$$P = \alpha \times m^k = 0{,}0000090909\,1^{1{,}047619048} = 0{,}0000093905\,3.$$

Наблюдаемое F-значение лежит между границами 0,000909091 и 0,000952381%. Составим таблицу статистик:

Таблица 16

Таблица статистик

Надёжность S, %	Вероятн. ошибки, α %	Число станд. откл., δ	Предельные значения F-статистики	Предел. знач. t-статистики	Предельные значения t-статистики
1	2	3	4	5	6
90	10	1,56	8,53	-2,92	2,92
95	5	1,96	18,51	-4,3	4,3
95,44	4,56	2	20,44	-4,5	4,5
99	1	2,58	98,5	-9,9	9,9
99,73	0,27	3	368,9	-19,2	19,2
99,9	0,1	3,29	998,5	-31,6	31,6
99,99	0,01		9998,5	-99,9	99,9

Окончание табл. 16

1	2	3	4	5	6
99,9936	0,0064	4	15623,50001	124,994	124,994
99,999047619	0,000952381		104998,4947	-324,0347123	324,0347123
0,99999090909	0,000909091		109998,489	-331,6602011	331,6602011

Определим число стандартных отклонений для S=90%. Итак, P=0,1/2=0,05, z=1,56.

Поскольку в таблице мы получили t=326,3635, то получаем подтверждение:
-331,66‹326,3635‹331,66.

Определим значимость произвольного эмпирического F-критерия [2. С. 152], в особенности для значений с P›0,1, воспользовавшись аппроксимацией, предложенной Паульсоном, которая справедлива для числа степеней свободы, не меньшего трёх (чем больше число степеней свободы, тем лучше аппроксимация), причём значимая вероятность определяется как площадь, соответствующая z-границам на обоих концах нормального распределения. Находим новое F-значение, равное 602187,5.

$$\hat{z} = \frac{(1+\frac{2}{9\times 6})\times 7372017^{\,1/3} - (1-\frac{2}{9\times 6})}{\sqrt{\frac{2}{9\times 6}\times 7372017^{\,2/3} + \frac{2}{9\times 1}}} = 5{,}188614\,.$$

По таблице площадей [2. С. 68] под кривой стандартного нормального распределения Фишера [11] от z до ∞ вероятность того, что переменная z примет значение ⩾ z P=0,0001. Поскольку:

$\mu \pm 1{,}96\sigma, \textit{или } z = \pm 1{,}96 \textit{ накрывают } 95\% \textit{ всей площади}$
$(P = 0{,}025; 0{,}025\times 2 = 0{,}05; 1-0{,}05 = 0{,}95)\,\textit{и}$

$\mu \pm 3\sigma, \textit{или } z = \pm 3 \textit{ накрывают } 99{,}73\% \textit{ всей площади}$
$(P = 0{,}0013; 0{,}0013\times 2 = 0{,}0026; 1-0{,}0026 = 0{,}9974)\,\textit{и}$
$\mu \pm 5{,}188614\sigma, \textit{или } z = \pm 5{,}188614 \textit{ накрывают не менее } 99{,}98\% \textit{ всей площади}$
$(0{,}0001\times 2 = 0{,}0002;\ 1-0{,}0002 = 0{,}9998).$

Поскольку $\hat{z} = 5{,}188614$, а максимальное табличное значение $\hat{z} = 3{,}7$ (*при котором вероятность ошибки* $P = 0{,}0001$). Вероятность ошибки $\alpha = \hat{z}\times 2 = 0{,}0001\times 2 = 0{,}0002;$ отсюда статистическая надёжность S=1-0,0002=0,0098. Вероятность ошибки, разумеется, значительно ниже, поскольку возможности таблицы Фишера ограничены потолком в 3,7. Подтвердилось ожидание, что надёжность выше. Рассмотрим связь F и 1/F и ν_1 *и* ν_2 [2. С. 150]:

$$F(\nu_1,\nu_2;1-\alpha) = \frac{1}{F(\nu_2,\nu_1;\alpha)}.$$

По соотношению 1 легко вычислить значение $F_{0,95}$ при известном $F_{0,05}$. Если даны $\nu_1 = 1\,\textit{и}\,\nu_2 = 2; \alpha = 0{,}05, F = 18{,}51.$ Найти F для: $\nu_1 = 1\,\textit{и}\,\nu_2 = 2; \alpha = 0{,}95.$ Определяем для $\nu_1 = 2\,\textit{и}\,\nu_2 = 1; \alpha = 0{,}05, F = 199{,}5$ [2. С. 138–149], откуда искомое значение равно 1/199,5=0,00501. Программа Microsoft Office Excel даёт ответ F=0,005012531. Способ получения данных вручную всё же остаётся нужным ввиду различных обстоятельств.

В заключение отметим, что надёжность S=99,999047619%, полученная для выведенного Кеплером уравнения, и сегодня актуальна, даже сегодня звучит, поскольку по умолчанию используется S=95%. Запуская ракету к Марсу, Вы получите ошибку α=9,52381E-06 – шансы у ракеты не будут отличными, но хорошими будут. Почему всё же не отличными? По мнению Мэтьюза: «Множество реальных данных содержат неопределённость или ошибку. Этот вид ошибки рассматривается как шум. Она воздействует на точность любого численного вычисления, основой которого являются данные. Улучшение точности не достигается при успешных вычислениях, использующих зашумленные данные» [3. С. 49–50].

Представленные Мэтьюзом исходные данные, как и ожидалось, в задании «a» значительно уменьшили ошибку; в задании «b» ошибка по существу исчезла. Таким образом, получен идеальный итог. Уотшем и Паррамоу [4] пишут, что новая литература и новые методы задействовали количественные приёмы, ранее применявшиеся только в физике, в то же время развив и приспособив технику количественного анализа к экономике. Напомним, экспоненциальные модели необходимы для расчётов в экономике.

1. Винстон, У. Microsoft Excel: анализ данных и построение бизнес-моделей / У. Винстон; пер. с англ. – М.: Русская редакция, 2005. – 576 с.: ил.+CD-ROM.
2. Закс, Л. Статистическое оценивание /Л. Закс; пер. с нем. – М.: Статистика, 1976. –598 с.
3. Мэтьюз, Джон Г. Численные методы. Использование MATLAB / Д. Г. Мэтьюз, К. Д. Финк; пер. с англ. – 3-е изд. – М.: Издательский дом «Вильямс», 2001. – 720 с.
4. Уотшем, Т. Дж. Количественные методы в финансах: учеб. пособие для вузов / Т. Дж. Уотшем, К. Паррамоу; пер. с англ. – М.: Финансы, ЮНИТИ, 1999. – 527 с.
5. NASA Научно-исследовательский центр Эймса [Электронный ресурс]. – Режим доступа: Johannes Kepler, NASA.com (дата обращения 15.03.2014).
6. Свободная энциклопедия Википедия [Электронный ресурс]. – Режим доступа: ru.wikipedia.org›Солнечная система (дата обращения 15.03.2014).
7. Большая энциклопедия школьника [Электронный ресурс]. – Режим доступа: http://www. for-schoolboy.ru›YUpiter-374.html (дата обращения 15.02.2014).
8. Свободная энциклопедия Википедия [Электронный ресурс]. – Режим доступа: ru.wikipedia.org›E (число) (дата обращения 20.09.2014).
9. Свободная энциклопедия Википедия [Электронный ресурс]. – Режим доступа: ru.wikipedia.org›Законы Кеплера (дата обращения 08.10.2014).
10. Импи К. Законы Кеплера / К. Импи, И. Даубар-Спитайл, П. Гей // Учить астрономию [Электронный ресурс]. – Режим доступа: http://m.teachastronomy.-com/astropedia/article/Keplers-Laws (дата обращения 25.09.2014).