История опытов с бросанием монеты
Часть 4. Опыт Романовского[1]

Г.И.Фалин, д.ф.м.н., проф.
кафедра теории вероятностей
механико-математический факультет
МГУ им М.В.Ломоносова (Москва)
http://mech.math.msu.su/~falin

В четвёртой части статьи (части 1-3 указаны в списке литературы) мы кратко рассказываем о жизни и научной работе русского математика В.И.Романовского. Используя оригинальную публикацию этого выдающегося учёного, мы описываем его опыт, который заключался в бросании комплекта из четырёх монет. Строгий статистический анализ результатов этого опыта показывает, что его нельзя интерпретировать как серию независимых бросаний правильной монеты. В связи с этим мы кратко обсуждаем проблему поиска идеального стохастического эксперимента, который можно рассматривать как статистическую копию эксперимента с бросанием монеты.

Ключевые слова: бросание монет, опыты, устойчивость частот,

[1] Г.И.Фалин. История опытов с бросанием монеты. Часть 4. Опыт Романовского. *Математика в профильной школе. Фрактал*, 2015, №1, стр.42-49.

Всеволод Иванович Романовский (22.11/5.12н.ст.1879 – 6.12.1954) – русский, советский математик. Его жизни и научной деятельности посвящена книга [1] из серии «Научно-биографическая литература», которую издаёт Академия Наук для читателей, интересующихся историей отечественной науки.

В.И.Романовский родился в г.Верном (с 1921 г. – Алма-Ата), где после службы в армии жил его отец, Иван Бернардович Романовский – польский дворянин, сосланный в Туркестан за участие в польских волнениях. Вскоре семья переехала в Ташкент, который стал одним из самых важных городов Средней Азии после присоединения её к России во второй половине 19 века. В Ташкенте В.И.Романовский окончил реальное училище и в 1900 году поступил в Петербургский технологический институт.

Вскоре он понял, что хотел бы заниматься математикой, и решил поступить в Петербургский университет. Для этого ему пришлось вернулся в Ташкент, чтобы сдать экзамены в гимназии и получить аттестат зрелости. Поступив в Петербургский университет в 1901 году, он окончил его в 1905 году и в 1906 году был оставлен в университете «по кафедре чистой математики … для приготовления к профессорской и преподавательской деятельности …» (цитата из представления ректора). В 1908 году материальные проблемы вынудили В.И.Романовского вернуться в Ташкент, где он стал преподавателем в родном реальном училище.

В 1911 году он получил должность доцента чистой математики Варшавского университета, а после защиты в 1912 году магистерской диссертации стал профессором. В В 1912 году В.И.Романовский опубликовал первую научную работу по теории верятностей [2], посвящённую закону больших чисел, в которой, в частности, изложены результаты проведённого им эксперимента с бросанием монеты. Об этом эксперименте В.И.Романовский пишет и в широко известном довоенном учебнике по статистике [3], стр. 47. Именно об этом эксперименте идёт речь в школьных учебниках. Отметим, что в книге [3] Романовский приводит и данные опыта Джевонса, о котором мы рассказали в третьей части нашей статьи. При этом допущена опечатка – в столбце с данными первой серии опытов Джевонса указано, что 4 орла выпало 210 раз, хотя в книге Джевонса стоит число 201 (210 – это ожидаемое из теории число опытов с 4 орлами).

Через год после начала в 1914 году Первой мировой войны Варшавский университет был эвакуирован, сначала в Москву, а затем в Ростов-на-Дону, и преобразован в Донской (Ростовский) университет. В 1918 при активном участии В.И.Романовского был открыт Ташкентский университет и он опять вернулся в город, где прошли его молодые годы, уже навсегда. В.И.Романовский внёс выдающийся вклад в развитие науки и образования в Узбекистане; он был деканом физико-математического факультета Ташкентского университета, его избрали в Академию Наук и Верховный Совет Узбекской ССР и т.д. Он воспитал много учеников и некоторые из них стали не только крупными математиками, но и государственными деятелями: С.Х.Сираждинов был ректором Ташкентского университета, вице-президентом АН Узбекистана, а с 1967 по 1980 – Председателем Верховного Совета Узбекской ССР, Т.А.Сарымсаков был ректором Ташкентского университета, вице-президентом и президентом АН Узбекистана. Именем В.И.Романовского названа улица в Ташкенте и институт математики АН Узбекистана.

Как мы говорили, детали упоминаемого в школьных учебниках опыта Романовского легко найти на стр.47 его книги [3] (на рис.2 приведён фрагмент этой страницы с описанием опыта). В отличие от того, что пишут в школьных учебниках, Романовский не бросал монету 80640 раз и не подсчитывал число выпадений орла. На самом деле он бросал n=20160 раз комплект из четырёх монет и ставил своей целью сравнить частоты («частости» в терминологии Романовского) пяти возможных элементарных исходов A_k=«k орлов и $4-k$ решек», $k=0,1,2,3,4,$ с их теоретическими вероятностями. Эти вероятности образуют биномиальное распределение:

$$P(A_0)=\frac{1}{16},\ P(A_1)=\frac{4}{16},\ P(A_2)=\frac{6}{16},\ P(A_3)=\frac{4}{16},\ P(A_4)=\frac{1}{16}.$$

Поэтому ожидаемое согласно теории среднее число появлений элементарных исходов A_0, A_1, A_2, A_3, A_4 в серии из n=20160 независимых экспериментов равно $N_0^*=nP(A_0)=1260,$ $N_1^*=nP(A_1)=5040$, $N_2^*=nP(A_2)=7560$, $N_3^*=nP(A_3)=5040$, $N_4^*=nP(A_4)=1260$ (отметим, что Романовский говорил о наиболее вероятном числе появления того или иного события, а не о его среднем значении). Реально события A_0, A_1, A_2, A_3, A_4 наблюдались в N_0=1402, N_1=5085, N_2=7583, N_3=4909, N_4=1181 экспериментах соответственно. Общее число выпавших орлов равно $N_1+2N_2+3N_3+4N_4=39702$, что составляет примерно 49.23% от числа $4n=80640$ (общего числа брошенных монет). На первый взгляд, соответствие между теоретическими и экспериментальными данными очень хорошее.

Однако, как и в аналогичном опыте Джевонса, бросание n=20160 раз комплекта из четырёх монет можно интерпретировать как бросание $4n=80640$ раз одной монеты *только при условии, что каждая из этих четырёх монет выпадает «орлом» вверх независимо от других* (иначе говоря, число «орлов» при одном бросании имеет приведённое выше биномиальное распределение). На самом деле результаты опыта В.И.Романовского с большой степенью достоверности противоречат гипотезе о выполнении этого условия.

Как мы отмечали при анализе опыта Джевонса, для проверки гипотезы о значениях вероятностей событий A_0, A_1, A_2, A_3, A_4 нужно подсчитать суммарное взвешенное квадратичное отклонение реальных чисел появления этих событий (т.е. N_k) от их значений $N_k^*=nP(A_k)$, ожидаемых из теории: $X^2=\sum_{k=0}^{4}\frac{1}{N_k^*}\left(N_k-N_k^*\right)^2$. Если это суммарное отклонение слишком велико, то проверяемую гипотезу следует отвергнуть. Вычисление величины X^2 удобно проводить с помощью Таблицы 1 (рис.1) по следующему алгоритму:

(1) в первом столбце этой таблицы нужно указать номера интересующих нас событий A_0, A_1, A_2, A_3, A_4.

Затем для каждой из пяти строк, соответствующих конкретному значению $k=0,1,2,3,4$, нужно:

(2) выписать вероятность $P(A_k)$ наступления события A_k (и тем самым сформировать второй столбец таблицы);

(3) умножением этой вероятности на число n=20160 экспериментов вычислить теоретическое среднее число наступлений события A_k (и тем самым сформировать третий столбец таблицы);

(4) после этого в четвёртый столбец таблицы записать, сколько раз в эксперименте наблюдалось событие A_k;

(5) вычитая из числа, записанного в четвёртом столбце, число, записанное в третьем столбце, найти отклонение наблюдаемого числа появлений события A_k от теоретического среднего (и тем самым сформировать пятый столбец таблицы);

(6) в последнем, шестом, столбце таблицы нужно записать число $\left(N_k - N_k^*\right)^2 \Big/ N_k^*$ – для этого следует число из пятого столбца возвести в квадрат и результат разделить на число из третьего столбца (конечно, той же строки).

Сложив все пять чисел (для $k = 0,1,2,3,4$) из последнего, шестого, столбца таблицы, в правом нижнем углу мы получим значение X^2; для опыта Романовского $X^2 = 24.83307$. Для контроля полезно сложить числа во втором, третьем, четвёртом и пятом столбцах. В результате мы должны получить 1 и 0 во втором и пятом столбцах соответственно, и 20160 (общее число экспериментов) в третьем и четвёртом столбцах.

Таблица 1. Результаты опыта Романовского

k	$P(A_k)$	$N_k^* = nP(A_k)$	N_k	$N_k - N_k^*$	$\left(N_k - N_k^*\right)^2 \Big/ N_k^*$
0	1/16	1260	1402	142	16.00317
1	4/16	5040	5085	45	0.401786
2	6/16	7560	7583	23	0.069974
3	4/16	5040	4909	-131	3.40496
4	1/16	1260	1181	-79	4.953175
Итого	1	20160	20160	0	$X^2 = 24.83307$

Рис.1.

Чтобы понять, велико или нет значение $X^2 = 24.83307$, применим знаменитый результат Карла Пирсона, который мы использовали в статье про опыт Джевонса.

Допустим, что в некотором стохастическом эксперименте выделена полная система событий $A_0, A_1, \ldots, A_r$. Предположим, далее, что мы повторили этот эксперимент n раз, независимо от друга, и для каждого события A_k подсчитали общее число N_k его появлений. Числа $N_0, N_1, \ldots, N_r$, конечно, являются случайными, так что случайной величиной будет и суммарное относительное отклонение $X^2 = \sum_{k=0}^{r} \frac{\left(N_k - nP(A_k)\right)^2}{nP(A_k)}$ этих чисел от их средних значений $nP\left(A_0\right), nP\left(A_1\right), \ldots, nP\left(A_r\right)$ (величину X^2 называют статистикой Пирсона).

Замечательный результат Пирсона заключается в том, что при $n \to \infty$ вероятность события $X^2 < x$ зависит только от x и r и не зависит от вероятностей $P\left(A_k\right)$: $\lim_{n\to\infty} P\left(X^2 < x\right) = F(x;r)$. Функция $F(x;r)$ называется функцией распределения χ^2 («хи-квадрат») с r степенями свободы; при $x \le 0$ она, очевидно, равна 0, а при $x > 0$ задаётся следующей формулой:

$$F(x;r) = \frac{1}{\Gamma\left(\frac{r}{2}\right)} \int_0^{x/2} u^{\frac{r}{2}-1} e^{-u} du ,$$

где $\Gamma(\alpha)=\int_0^{+\infty} u^{\alpha-1}e^{-u}du$ – гамма-функция Эйлера.

Распределение «хи-квадрат» удобно характеризовать так называемыми верхними квантилями; это числа $z_\varepsilon(r)$ такие, что $\lim_{n\to\infty} P\left(X^2 \geq z_\varepsilon(r)\right)=\varepsilon$. При заданных ε и r вычислить число $z_\varepsilon(r)$ можно, например, с помощью стандартной функции CHIINV пакета Microsoft Excel: $z_\varepsilon(r)=\mathrm{CHIINV}(\varepsilon,r)$. При большом n вероятность события $\left\{X^2 \geq z_\varepsilon(r)\right\}$ приближённо равна ε. Если взять число ε достаточно маленьким (например, 0.1%, 0.01% и т.п.) и пренебрегать возможностью наступления событий, вероятность которых меньше или равна ε, то событие $\left\{X^2 \geq z_\varepsilon(r)\right\}$ следует считать невозможным (с практической точки зрения). На этом основан знаменитый критерий «хи-квадрат» проверки гипотезы о характере проводимого эксперимента:

если по результатам эксперимента оказалось, что величина X^2 больше, чем $z_\varepsilon(r)$, то проверяемую гипотезу следует отвергнуть (с уровнем значимости ε).

При одновременном бросании четырёх монет мы имеем 5 возможных элементарных исходов A_k=«k орлов и $4-k$ решек», $k=0,1,2,3,4$. Поэтому число степеней свободы равно 4. Выберем уровень значимости $\varepsilon=10^{-4}$. С помощью формулы =CHIINV(0.0001,4) программы Microsoft Excel мы можем вычислить, что $z_{0.0001}(4)\approx 23.51274$. Таким образом, величина X^2 может превысить уровень 23.51274 лишь в одном случае из 10000. Если бы кто-то каждый день n=20160 раз бросал комплект из четырёх монет (реально на такое число повторений эксперимента потребуется около месяца), то за тысячу лет будет лишь около 36 дней, когда статистика Пирсона $X^2=\sum_{k=0}^{4}\frac{1}{N_k^*}\left(N_k-N_k^*\right)^2$ превысит уровень 23.51274. Вряд ли в единичном эксперименте Романовского наблюдалось столь маловероятное событие – гораздо разумнее признать, что проверяемая гипотеза о характере проведённого эксперимента ошибочна. Иначе говоря, результаты опыта В.И.Романовского противоречат гипотезе о том, что бросаемые им четыре монеты были идеально правильными и выпадали орлом вверх независимо друг от друга. Соответственно, опыт В.И.Романовского с бросанием n=20160 раз комплекта из четырёх монет *нельзя интерпретировать* как бросание $4n=80640$ раз одной монеты. Основанный на результатах этой *интерпретации* вывод о том, что полученная Романовским частота выпадения орла, равная 49.23%, подтверждает факт устойчивости частот, является результатом произвольного манипулирования опытными данными.

Из данных таблицы 1 с расчётами величины X^2 хорошо видно, что большое число орлов (3 или 4), так же как и большое число решек (3 или 4), получалось слишком часто. Видимо, это связано с тем, что монеты плохо перемешивались. Это совсем не удивительно, если отметить следующее обстоятельство. Допустим, что на одно бросание комплекта из четырёх монет (с последующей записью результата) уходит 1 минута. Тогда на n=20160 повторений эксперимента потребуется 336 часов, т.е. 14 дней непрерывной работы, без перерывов на обед, отдых или сон.

Это замечание и проведённый анализ опытов Джевонса и Романовского хорошо показывают, что бросание монеты нельзя считать хорошей моделью стохастического эксперимента.

Вот как об этом пишет В.Феллер на стр. 38 своего знаменитого учебника [4]: «применение утончённых статистических методов к фактическим результатам опытов с бросанием монет неизменно показывало, что выпадение герба и выпадение решётки *не* являются одинаково вероятными событиями… отклонения от нашей схемы всегда связаны с такими явлениями, как, например, несовпадение центра тяжести монеты с её геометрическим центром». На стр.40 он продолжает: «Бросание любой настоящей монеты приводит к искажённым результатам, из-за действия не поддающихся учёту факторов».

Поэтому уже давно появилась идея заменить бросание монеты другим стохастическим экспериментом, который *статистически неотличим* от эксперимента с бросанием монеты, является его *статистической копией*. Под статистической копией бросания монеты мы понимаем стохастический эксперимент, в результате которого наблюдается появление или нет события A, вероятность которого равна 0.5. Поскольку вероятность появления интересующего нас события A такая же, как и вероятность выпадения орла при бросании правильной монеты, появление события A в этом эксперименте *можно интерпретировать* как выпадение орла (или решки) в опыте с бросанием одной монеты.

Статистической копией бросания одной монеты можно считать эксперимент с бросанием игральной кости, если в качестве основного события A рассматривать событие «выпала чётная цифра: 2, 4 или 6». Ясно, что с равным успехом можно брать событие «выпала цифра 1, 2 или 3» и, вообще, выпадение цифры из любого заранее фиксированного набора из трёх цифр. Однако, к бросанию игральной кости применимы все замечания В.Феллера, так что этот эксперимент мало чем отличается от бросания монеты в смысле своего качества.

Один из основоположников современной статистики, выдающийся британский учёный Карл Пирсон (Karl Pearson, 27.03.1857 – 27.04.1936), предположил, что идеальным стохастическим экспериментом является игра в рулетку в казино Монте Карло. Вот как он мотивировал этот выбор: «Считается, что весь этот аппарат изготовлен исключительно точно и перед использованием отрегулирован с величайшей тщательностью. Принимая во внимание механическую точность этого инструмента, проницательные и бдительные глаза многочисленных игроков, трудно представить машину, которая бы лучше подходила для иллюстрации законов случайности, чем рулетка Монте Карло.» (стр. 47 книги [5]).

При игре в рулетку, как известно, на вращающемся колесе равномерно расположены 37 лунок с цветными номерами от 0 до 36; цифра 0 имеет зелёный цвет, а номера от 1 до 36 – или красный, или чёрный. Номера стоят не по порядку, а следующим образом (в направлении вращения часовой стрелки): 0 (зелёный), $\underline{32}$, 15, $\underline{19}$, 4, $\underline{21}$, 2, $\underline{25}$, 17, $\underline{34}$, … , $\underline{12}$, 35, $\underline{3}$, 26. Подчёркнутые номера от 1 до 36 (их ровно 18) – красные, а остальные (их тоже ровно 18) – чёрные. Крупье (сотрудник казино, который ведёт игру) раскручивает колесо рулетки и бросает шарик так, чтобы он двигался по окружности колеса в направлении, противоположном вращению колеса. Когда и колесо, и шарик останавливаются, шарик оказывается в одной из 37 лунок – её номер и цвет определяют результат игры. Для идеальной рулетки вероятность появления конкретного номера равна $\frac{1}{37}$. При игнорировании номера 0 вероятность появления номера красного цвета равна

$\frac{18}{36}=\frac{1}{2}$, чёрного цвета – $\frac{18}{36}=\frac{1}{2}$. Можно сказать, что при игре в рулетку (с игнорированием номера 0) появление красного (чёрного) цвета статистически неотличимо от появления орла (решки) при бросании монеты.

После детального анализа реальных статистических данных, записанных в казино, К.Пирсон заключил: «Можно с уверенностью сказать, что, если мы интересуемся только общим числом красных и чёрных результатов, рулетка Монте Карло подчиняется математическим законам случайности» (стр.51 книги [5]). В этой цитате К.Пирсон особо подчёркнул, что его вывод базировался *только* на анализе *общего числа* результатов каждого цвета. Иначе говоря, К.Пирсон ясно понимал, что сделал лишь первый шаг в процессе анализа гипотезы о полной случайности результатов игры в рулетку.

Поэтому К.Пирсон решил рассмотреть и частоты появления более сложных событий: частоты появления отдельных чисел, включая и 0, а также частоты появления последовательностей чисел одного цвета заданной длины. В результате оказалось, что «коротких последовательностей номеров одного цвета недостаточно много, а изменение цвета происходит гораздо чаще, чем предписывают законы случайности…» (стр.57 книги [5]). Примерно так же обстояло дело с опытами Джевонса и Романовского – если интересоваться только общим числом орлов, всё нормально, но если изучать более сложные события, становится ясно, что о полной случайности вести речь нельзя.

Таким образом, и рулетка не может служить хорошей моделью стохастического эксперимента. Как же найти идеальный стохастический эксперимент? Об этом мы расскажем в нашей следующей статье, посвящённой опыту В.Феллера.

Литература

1. Боголюбов А.Н., Матвиевская Г.П. Всеволод Иванович Романовский. 1879-1954. М.:.Наука, 1997.

2. Романовский В.И. Закон больших чисел и теорема Якова Бернулли. Протоколы заседаний Общества естествоиспытателей при Варшавском университете за 1911 г. 1912, №4, стр. 39-63.

3. Романовский В.И. Элементарный курс математической статистики. 2-е изд. Госпланиздат, Москва, 1939.

4. В.Феллер. Введение в теорию вероятностей и её приложения. Том 1. М.:Мир, 1984.

5. K.Pearson. The chances of death and other studies in evolution. Vol.1. Edvard Arnold, London, 1897.

6. Г.И.Фалин. История опытов с бросанием монеты. Часть 1. Опыт де Бюффона. *Математика в школе,* 2014, №9, стр.55-60.

7. Г.И.Фалин. История опытов с бросанием монеты. Часть 2. Опыт де Моргана. *Математика в школе*, 2014, №10, стр.52-57.

8. Г.И.Фалин. История опытов с бросанием монеты. Часть 3. Опыт Джевонса. *Математика в школе*, 2015, №1, стр.52-58.

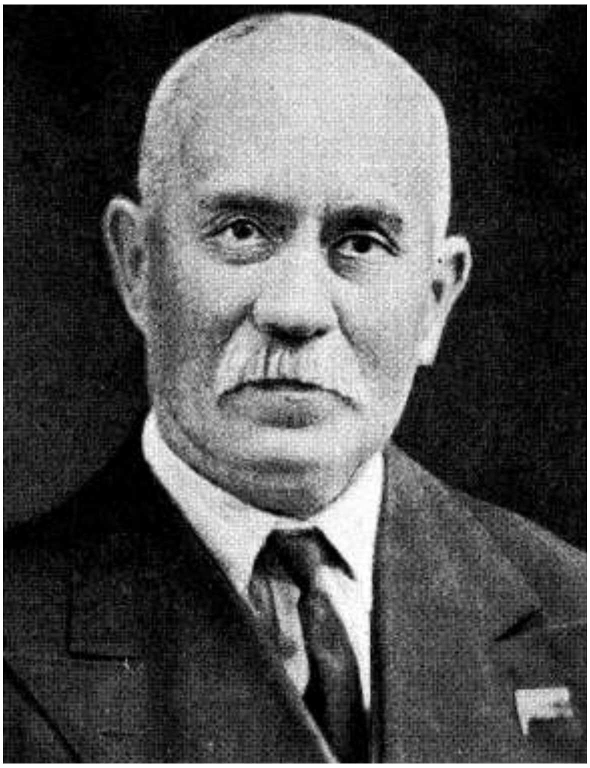

Рис. 1 Всеволод Иванович Романовский

Из других опытов приведем результаты опытов автора настоящего курса, предпринятых с целью проверки другой части теоремы Якова Бернулли, — проверки опытного подтверждения различных вероятностей различного приближения относительно частости события к его вероятности, и опубликованных в Варшавских университетских известиях за 1912 г. Автор бросил 20 160 раз 4 монеты. Комбинации: 4 орла; 3 орла и 1 решетка; 2 орла и 2 решетки; 1 орел и 3 решетки; 4 решетки имеют вероятности: $\frac{1}{16}$, $\frac{4}{16}$, $\frac{6}{16}$, $\frac{4}{16}$, $\frac{1}{16}$, и потому, умножив на эти числа 20 160, мы будем иметь следующие наивероятнейшие частости этих комбинаций:

1 260, 5 040, 7 560, 5 040, 1 260.

Бросания же дали частости:

1 181, 4 909, 7 583, 5 085, 1 402.

Рис. 2 Часть страницы из «Элементарного курса математической статистики» В.И.Романовского, где описан его опыт