Геометрическая вероятность в причинно-следственных связях.

Филатов Олег Владимирович / Filatov Oleg - инженер-программист

НТЦ Модуль, г. Москва. fil_post@rambler.ru

Ключевые слова: Уолтер Пенни, игра Пенни, геометрическая частота, геометрическая вероятность, КоД Пи, постулаты Голомба.

Keywords: Walter Penny, Penny game, a geometric rate, geometric probability, CoD Pi, the postulates of Golomb.

Введение.

Данная статья предназначена для широкой публики. В статье, в популярной форме (в форме описания эксперимента по подтверждению особых способностей экстрасенса) сделана попытка рассказать о новом, ещё не признанном, математическом направлении на стыке программирования, комбинаторики и вероятностей. А именно, о «Комбинаторике Длинных Последовательностей», сокращённое название – «КоД Пи». По сравнению с обычной комбинаторикой, которая работает на коротких базовых отрезках, длиной в 10 – 50 событий (из-за того, что в комбинаторных формулах используются факториальные операции), Код Пи работает на отрезках длиной в миллионы и миллиарды событий, так как формулы Кода Пи не содержат факториалов. Код Пи является, на данном этапе, преимущественно экспериментальной наукой, в которой эксперименты производятся на большом количестве случайных бинарных событий при помощи компьютерных поисковых алгоритмов [1, 2, 3, 8]. Хотя доказана центральная теорема для данного научного направления: «О амплитудно-частотной характеристике идеальной бинарной случайной последовательности» [2, 5, 8]. Компьютерный поиск закономерностей в больших массивах случайных бинарных данных появился только недавно, предыдущие поколения исследователей не обладали компьютерами для проведения подобных поисковых работ.

Легенда для сюжета статьи. Экстрасенс утверждает, что связан с глубинной сутью вещей, и он может чувствовать объединения одинаковых событий (групп) монет, накрытых сверху непроницаемыми экранами. Но он может чувствовать не любые выкладки из монет, а только истинно случайные, которые возникают при честном подбрасывании монеты и посылаются самой сутью Вселенной.

Дадим перевод с экстрасенсорной терминологии на человеческий язык. Экстрасенс утверждает, что будет попадать в цепочки одинаковых событий («0000», «111111», …), образованные по результатам выпадений честной монеты, с вероятностью иной, чем приписывает формула равновероятностных исходов, ф.1:

$$^{n}p = \frac{1}{2^{n}} \qquad \text{Ф.1}$$

где n - длина цепочки событий.

Ниже описанный эксперимент с «экстрасенсом» является реальным компьютерным экспериментом $[2, 4, 8]$, подтверждённый результатами. Допуски на полученные результаты в базовых экспериментах для этой статьи (для широкого круга читателей), не вводятся по той же причине, по которой они не вводятся при обсуждении выпадения монеты в статьях для широкого круга читателей.

Основная часть.

Утверждение экстрасенса (что он может менять вероятность обнаружения серий монет), является очень смелым, ведь всем хорошо известно, ещё со школьных времён, что выпадения монеты никак не связаны между собой, и не влияют друг на друга. Нельзя по анализу выпавших событий предсказать результат будущего выпадения монеты, в котором вероятность выпадения любого из двух возможных событий равна 0,5. А вероятность выпадения цепочки из n бинарных событий, в любых комбинациях, считается по ф.1, и зависит только от количества событий n и ни как не зависит от их величин.

Утверждение экстрасенса проверили экспериментально. Специально для него взяли несколько тысяч монет, поочерёдно каждую из них честно подбрасывали и клали в продолжение цепочки ранее выпавших монет, на краю дороги, накрывая сверху непроницаемым экраном. В результате, по дороге можно было свободно ходить вдоль всей цепочки с выпавшими монетами. Причём, у учёных проводящих эксперимент, сохранилась уверенность, что вероятность обнаружения состояния любой монеты цепочки осталась раной 0,5. А цепочку из n любых результатов можно угадать, в среднем, с вероятностью, рассчитываемой по ф.1.

Начался эксперимент. Экстрасенс медленно шёл от начала к концу цепочки, держа правую руку ладонью вниз – сканируя ладонью её монеты. Останавливался, указывал с какой монеты надо начать открывать монеты спереди и сзади от неё. Для простоты изложения будем обозначать одну стороны монеты как «0», а другую сторону монеты как «1». Открываемые монеты составляли цепочки событий: «11», «111», «1111», «11111», … . Ну и соответственно цепочки из: «00», «000», «0000», «0000», … . Длина цепочки событий ограничивалась выпадением другой стороны, например цепочку «0000» на самом деле получали из такой последовательности сторон монеты: «100001». Но, ограничивающие длину цепочки инверсные события (единицы), спереди и сзади, отбрасывались. Учитывались только численность одинаковых событий в цепочках. Конечно же, экстрасенсу попадались и цепочки единичной длины: «0», «1».

Поскольку экстрасенс утверждал, что он чувствует группировку одинаковых событий вместе, поэтому учитывались только длины цепочек, а не образующие их значения «0», «1». То есть цепочки из пяти нулей «00000» и пяти единиц «11111» учитывались только по их длине пять и не делились на цепочки из нулей и на цепочки из единиц. Так же и для всех цепочек других длин.

Экстрасенс завершил проход вдоль ряда монет. Статистика была собрана. Оказалось, что формула, описывающая численность цепочек ${}^{n}S_{\mathrm{g}}$, в зависимости от длины n цепочки выглядит так, ф.2 [2, 4, 7, 8]:

$$ {}^{n}S_{\mathrm{g}} = Z_{\mathrm{g}} \cdot \frac{n}{2^{n+1}} \qquad \text{Ф.2} $$

где Z_{g} – число монет, на которые указывал экстрасенс (число серий).

Действительно, внешний вид формулы неожиданно сильно отличается от общепризнанной ф.1. А это значит, что экстрасенс действительно имеет особые отношения с причинно-следственной связью. А особые отношения с причинно-следственной связью могут приводить к особым отношениям со временем. Например, экстрасенс, может быть, обладает ясновиденьем, и может заглядывать в будущее. Ведь причинно-следственная связь – это первый шаг к освоению времени.

Перепишем полученную экспериментальным путём формулу ф.2 в виде частотной зависимости ${}^{n}f_{\mathrm{g}}$, ф.3. В частотной формуле ф.3 устранена зависимость от числа серий Z_{g} [2, 4, 7, 8].

$$ {}^{n}f_{\mathrm{g}} = \frac{n}{2^{n+1}} \qquad \text{Ф.3} $$

Переход к частотному виду ф.3, позволит сравнивать получаемую формулу с классической формулой вероятности, ф.1, так как вероятность p является то же частотой.

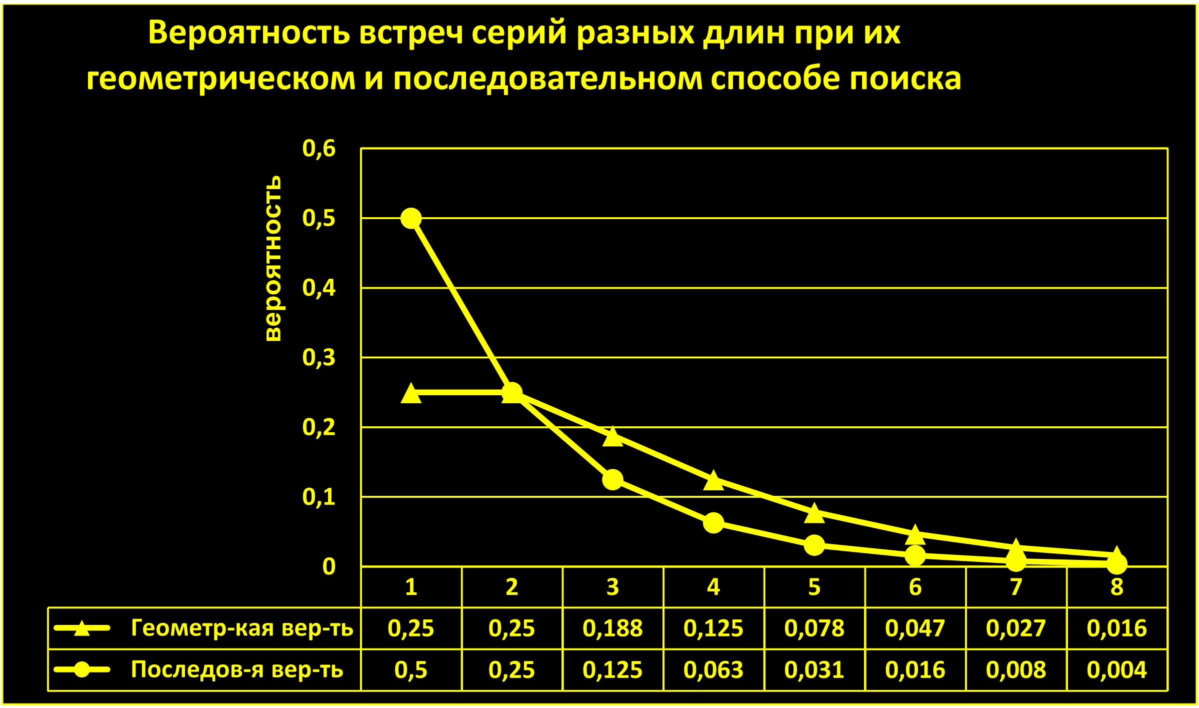
Две частотные формулы: ф.1 и ф.3 сильно отличаются друг от друга, что по мысли экстрасенса доказывает его особые способности. Чем длиннее цепочка, тем отчётливее она им ощущается, утверждал экстрасенс, поэтому, говорил он, ему легче угадывать длинные, истинно случайные цепочки.

Для расчёта того, насколько чаще положенного по ф.1, экстрасенс угадывал цепочки событий, обозначим вероятность обнаружения цепочки длины n, по ф.1, через: ${}^{n}p$. Разделив ф.1 на ф.3, получим по ф.4 коэффициент k, который показывает, во сколько раз чаще экстрасенс угадывает цепочку длины n, чем следовало бы по ф.1.

$$ k = \frac{{}^{n}f_{\mathrm{g}}}{{}^{n}p} = \frac{n}{2^{n+1}} : \frac{1}{2^{n}} = \frac{n}{2} \qquad \text{Ф.4} $$

Из ф.4 видно, что экстрасенс совсем плохо, в два раза хуже, чем по ф.1, угадывает (или хорошо не угадывает) цепочки из одиночных событий («0», «1»). Угадывает ровно столько же цепочек длины два («00», «11»), сколько и предсказывает ф.1 – это единственная точка совпадения эксперимента и общепринятой теории вероятности. А вот все остальные цепочки он угадывает чаще и лучше, чем следует из ф.1.

Давайте, для наглядности, построим график вероятности угадываний цепочек из n событий подряд, по общепринятой ф.1 («Последов-ая вер-ть» на графике 1). А также график частот, с которыми экстрасенс обнаруживал цепочки событий («Геометр-кая вер-ть» на графике 1).



	1	2	3	4	5	6	7	8
Геометр-кая вер-ть	0,25	0,25	0,188	0,125	0,078	0,047	0,027	0,016
Последов-я вер-ть	0,5	0,25	0,125	0,063	0,031	0,016	0,008	0,004

График 1

Научное объяснение результатов эксперимента с экстрасенсом при помощи геометрической вероятности.

Определение геометрической вероятности. Геометрической вероятностью называется вероятность попадания точки в пространственную область, которая пропорциональна размеру этой области (длине, площади).

Давайте посмотрим на фрагмент бинарной случайной последовательности: «11100010111**000000**110**111111**000111010111…». В этом фрагменте мы замечаем группы событий (выделены жирным) у которых длина больше, чем у других групп. Значит, по определению геометрической вероятности, их длина работает на увеличение вероятности попадания в эти цепочки. Именно это мы видим в ф.3.

В ф.3 частота попадания в цепочку длины n прямо зависит от этой длины (n), точно так, как того требует определение геометрической вероятности.

Действительно, обладая свободным перемещением вдоль цепочки случайных бинарных событий, и производя случайное вскрытие той или иной монеты, мы полностью удовлетворяем условию на случайное попадание точки в пространственную область пропорционально размеру этой области, в нашем случае – длине.

Таким образом, все сверхъестественные способности, на самом деле, оказались проявлением давно известного и хорошо изученного закона геометрической вероятности.

Любой заядлый скептик (отрицающий сверх способности), идя вдоль цепочки с монетами, и случайным образом указываю на любую из них, по закону геометрической вероятности получит распределение частот, описываемое в ф.3, (кривая «Геометр-кая вер-ть» на графике 1).

А если этот скептик будет по очереди (от начала цепочки к её концу) просматривать каждую монету, и набирать статистику то у него получится распределение частот, описываемое в ф.1 (кривая «Последов-ая вер-ть» на графике 1).

Следовательно, совершенно очевидно, что вероятность обнаружения цепочки событий всего на всего зависит от способа их поиска.

При последовательном учёте всех составных событий, будет получена частота обнаружения $f = \frac{1}{2^n}$.

При случайном погружении зонда в бинарную последовательность будет получена частота обнаружения $f = \frac{n}{2^n}$.

Стоп. Как это, вероятность обнаружения цепочки событий всего на всего зависит от способа их поиска? <u>Этого не может быть никогда</u>. Потому, что выпадения монеты никак не связаны между собой, и не влияют друг на друга. Нельзя по анализу выпавших событий предсказать результат будущего выпадения монеты, в котором вероятность выпадения любого из двух возможных событий равна 0,5. А вероятность выпадения цепочки из n бинарных событий, в любых комбинациях, считается по ф.1, и зависит только от количества событий n и ни как не зависит от их величин.

В работах [6, 8] даны ещё несколько способов изменяющих вероятность обнаружения искомых серий.

Поскольку верхний объём статьи, установленный Журналом, уже практически достигнут, а переход от геометрических частот к геометрическим вероятностям ещё не сделан, то читатель может его развитие проследить в работах [7, 8]. Напишем в этой статье, без вывода, что геометрическая вероятность P_{g}^{A} рассчитывается по ф.5 [7, 8].

В качестве справки, упомянем, что впервые изменение вероятностей при обнаружении цепочек событий, достиг Уолтер Пенни в 1969 году (смотри в интернете «Игра Пенни»). Довольно быстро было дано качественное объяснение процессов приводящих к парадоксу Пенни, но в этих объяснениях не возникло понимание того, что наблюдается процесс изменения вероятности (управление вероятностью). Только когда удалось произвести количественные расчёты для пар цепочек в игре Пенни [8, 9], и появились другие поисковые правила R, приводящие к возможности менять (управлять) процентные соотношения находимых цепочек [6, 8], то стало понятно, что правила игры Пенни являются одной из разновидностей поисковых правил R меняющих вероятность обнаружения искомых цепочек.

На первый взгляд покажется странным, но постулаты Голомба так же являются поисковыми правилами, но вывернутыми на оборот. По постулатам Голомба создаётся некоторое подобие реальной бинарной последовательности (постулаты требуют многократной, периодической повторяемости пос-ти, к сожалению, при росте числа повторов периодов голомбовской пос-ти, не ограниченно растёт число не компенсировано угаданных в ней нулей и единиц). Поэтому постулаты Голомба, как и правила игры Пенни легко выводятся из формул «Комбинаторики длинных последовательностей» [8, 10].

Эта статья является вводной частью для рассмотрения геометрического поиска случайных цепочек. В этой статье были рассмотрена только геометрическая частота попадания в составные события разных длин. Если эта статья будет напечатана, то в последующей статье будет получена формула для геометрической вероятности P_{g}^{A}, через комбинаторную вероятность P_p, ф.5:

$$P_{\mathrm{g}}^{A} = P_p \cdot (n-1) = \frac{n-1}{2^n} \qquad \text{Ф.5}$$

где n - длина цепочки событий.

Обсуждения

Перемещение экстрасенса вдоль ряда с карточками даёт ему возможность работы с геометрическими вероятностями. Но эти же перемещения можно сравнить и с перемещением во времени.

Интересно понять следующую модель. Ряд карточек закрывает события – пос-ть выпадений стороны монеты. Перемещение над пос-тью событий одного и того же предмета можно осуществлять только во времени. Таким образом, перемещение над рядом из событий пос-го выпадения монеты, накрытых карточками, является моделью перемещения во времени. Но одновременно это же перемещение является и геометрической вероятностью. В данной модели перемещения над карточками время и пространство объединяются в одно.

Так как при одинаковом количестве попыток найти заданные серии одинаковой длины, геометрический алгоритм находит больше серий, чем последовательный, то эти алгоритмы имеют разную поисковую мощь. Можно выдвинуть гипотезу, что различие поисковой мощи у алгоритмов определяется различием в поисковых степенях свободы. Чем больше степеней поисковой свободы, тем большей мощью обладает поисковый алгоритм [6, 7, 8].

Последовательный поисковый алгоритм обладает одной степенью свободы, он может искать только справа от зонда.

Геометрический поисковый алгоритм обладает двумя степенями поисковой свободы, он ищет слева и справа от зонда.

Для увеличения вероятности нахождения искомых серий алгоритм геометрического поиска требует существования не рассматриваемых элементарных событий, которые пропускаются алгоритмом без определения их содержимого («0» или «1»). Не рассматриваемые элы можно сравнить с своеобразной «тёмной материей» но не вселенной, а вероятностного пространства. «Тёмная материя» из элементарных событий, принципиально должна существовать для работы геометрического алгоритма.

Приведём ещё несколько коротких, но не менее эффектных экспериментальных открытий сделанных в рамках «Комбинаторики длинных последовательностей». Описание экспериментов и алгоритм программы даны в [1, 2, 3, 8].

Неожиданная асимметрия 1. Мы все привыкли к мысли, что чётных и нечётных составных событий («1», «0», «11», «00», «111», «000», «1111», «0000», …) при подбрасывании монеты должно быть поровну. Это не так. Нечётных событий больше. Формула ф.6, связывает число составных событий длиной n с общим числом бросков монеты N:

$$^{n}S = \frac{N}{2^{n+1}} \qquad \text{Ф.6}$$

Поскольку по ф.6 самыми многочисленными составными событиями являются события единичной длины ^{1}S («1», «0»), то это и определяет численное преимущество встречаемых нечётных серий при подбрасывании монеты над чётными сериями.

Ещё большей числовой асимметрии подвержены цуги - цепочки, образуемые составными событиями одинаковой длины. Пример цуг: «01010», «11001100», «111000». Отметим, что составное событие, окружённое спереди и сзади (слева и справа) составными событиями других длин, тоже является цугой, вернее единичной цугой. Пример единичной цуги: «11**0**11», «10**1111**001».

Неожиданная асимметрия 2. Сравнение численностей составных событий и цуг приводит к ещё одному факту, в который просто так поверить нельзя. Так в популярных лекциях по вероятностям, лектор любит задавать вопрос, о том, что чаще встречаются десять (сто , тысяча, …) одинаковых событий подряд (составные события: «0000000000», «1111111111») или чередования такой же длины (цуги: «1010101010», «0101010101»). Все дружно отвечают, что одинаково. И этот ответ не правильный, для подразумевающейся по умолчанию поисковой модели (последовательно просматриваем все результаты выпадений монеты). Цуг указанной длины (цуги: «1010101010», «0101010101») в четыре раза будет найдено меньше, чем составных событий («0000000000», «1111111111»).

Почему же раньше этого не заметили? Потому что искали при помощи других поисковых правил (которые не требуют компьютеров). А именно, для упрощения ручной обработки данных, длинная случайная последовательность из нулей и единиц не образовывалась вообще. А из результатов выпадений монеты изначально образовывали отрезки равной длины (например, по десять событий), и количества фрагментов с цугами (цуги: «1010101010», «0101010101») при таком поиске равно количеству фрагментов с составными событиями («0000000000», «1111111111»).

Так что, лектору, на самом деле, прежде чем задавать вопрос о том, каких цепочек больше, надо уточнять, каким образом производится поиск этих цепочек.

Очень простые отношения. В любой случайной последовательности из N результатов подбрасываний монеты ($N/1$), будет $N/2$ составных событий и $N/3$ цуг, $N/4$ - число всех последовательных составных событий [1], и $N/8$ - число всех инверсных составных событий [1].

Выводы

Открытие зависимости частоты обнаружения искомой цепочки от способа её поиска - является самым значимым открытием в современной теории вероятности.

Зависимость вероятности обнаружения от способа поиска цепочки является большой философская проблемой, большой статистической проблемой для гуманитарных, технических и фундаментальных наук.

Модная теория одновременного существования множества мировых историй приобретает косвенную поддержку в явлении зависимости вероятности обнаружения равновероятных случайных бинарных событий от способа их поиска. Когда один из возможных будущих вариантов мира (в смысле альтернативных историй) начинает приобретать черты предопределённости и начинает реализовываться, при выдерживании в течение достаточного времени конкретного способа (системы) поиска событий. Но существуют миры (альтернативные события) все сразу, просто остальные миры не проявляются, из-за того, что при поиске данной физической моделью они не просматриваются.

Библиографический список

1. Филатов О. В., Филатов И.О., Макеева Л.Л. и др. «Потоковая теория: из сайта в книгу». Москва, «Век информации», 2014. С.200.
2. Филатов О. В., Филатов И.О. «Закономерность в выпадении монет – закон потоковой последовательности». Германия, Издательский Дом: LAPLAMBERT Academic Publishing, 2015, с. 268.
3. Филатов О. В., Филатов И.О., статья «О закономерностях структуры бинарной последовательности», «Журнал научных публикаций аспирантов и докторантов», №5, 2014.
4. Филатов О. В., Филатов И.О., Статья «О закономерностях структуры бинарной последовательности (продолжение)», «Журнал научных публикаций аспирантов и докторантов», № 6, 2014.
5. Филатов О. В., статья «Теорема «О амплитудно-частотной характеристике идеальной бинарной случайной последовательности», «Проблемы современной науки и образования», № 1 (31), 2015 г.

6. Филатов О. В., статья «Описание схем управления вероятностью выпадения независимых составных событий», «Проблемы современной науки и образования», №2 (44), 2016 г.
7. Филатов О. В., статья «The use of geometric probability to change the probability of finding a series of random deposition coins. / Применение геометрической вероятности для изменения вероятности нахождения серий случайных выпадений монеты.», журнал «Проблемы современной науки и образования / Problems of modern science and education», № 22 (64), 2016 г.
8. Авторский сайт со статьями: http://kodpi.net/
9. Филатов О. В., статья «Расчёт численностей поисковых шаблонов в парадоксе Пенни», «Проблемы современной науки и образования», № 11 (41), 2015 г.
10. Филатов О. В., статья «Derivation of formulas for Golomb postulates. A method for creating pseudo-random sequence of frequencies Mises. Basics "Combinatorics of long sequences." / Вывод формул для постулатов Голомба. Способ создания псевдослучайной последовательности из частот Мизеса. Основы "Комбинаторики длинных последовательностей"», журнал «Проблемы современной науки и образования / Problems of modern science and education», № 17 (59), 2016 г.