BOUNDARY CONCENTRATION IN SERIES COIN FLIP. THEOREM "ON EQUALITY OF EVENTS SUM OF THE FIRST TO GUES THE NUMBER OF SERIES"

Filatov O.

КРАЕВЫЕ УПЛОТНЕНИЯ В СЕРИЯХ ПОДБРАСЫВАНИЙ МОНЕТЫ. ТЕОРЕМА «О РАВЕНСТВЕ СУММЫ ПЕРВЫХ УГАДАННЫХ СОБЫТИЙ ЧИСЛУ СЕРИЙ»

Филатов О. В.

Филатов Олег Владимирович / Filatov Oleg - инженер-программист,
Закрытое акционерное общество «Научно технический центр «Модуль», г. Москва

Аннотация: *показывается, как при любом снятии информации со случайной бинарной последовательности образуется неопределённость, аналогичная принципу неопределённости Гейзенберга в физике, дана попытка объяснить эту неопределённость эффектом экранирования, более известным как парадокс Пенни, игра Пенни; используя свойство средней длины составных событий, образующих бинарную последовательность, рассчитано число образующих эту последовательность серий; дана техника изменения вероятности угадываний, основанная на эффекте краевых уплотнений в коротких сериях.*

Abstract: *it shows how for any receiving of information from random binary sequence generated uncertainty similar to the Heisenberg uncertainty principle in physics, an attempt to explain given the uncertainty of the effect of screening, better known as the Penny paradox, game Penny; using medium length composite event property, forming a binary sequence, calculated the number sequence forming this series; given the changes in the probability of guessing technique is base on the effect of edge condensations in short series.*

Ключевые слова: *элементарное событие, составное событие, игра Пенни, парадокс Пенни, краевые уплотнения, бинарная последовательность, НТЦ Модуль.*

Keywords: *elementary event, a composite event, game Penny, Penny paradox, binary sequence, random binary sequence, edge condensations.*

Введение

Для объяснения комбинаторного эффекта экранирования, известного под именами: «Игра Пенни», «Парадокс Пенни», необходимо ввести понятие «Составное событие» [1-3, 5]. «Составные события» - это серии одинаковых элементарных бинарных событий, иначе говоря: непрерывные серии либо из «0», либо из «1». Из составных событий образованы случайные бинарные последовательностей. Примером такой последовательности служит последовательность из результатов выпадений «честной» монеты (одна сторона монеты «0», другая «1»). То есть, составными событиями называют группы выпавших одинаковых элементарных событий (подчёркнуты), которые выделены спереди и сзади инверсными событиями: «1<u>0</u>1», «0<u>1</u>0», «1<u>00</u>1», «0<u>11</u>0», «1<u>000</u>1», «0<u>1111</u>0», … Причём, выделяющие инверсные события не являются частью этого составного события, но они являются частями других составных событий: «…0 <u>1111</u> 0000 <u>11</u>0…».

В своих постулатах Соломон Голомб описал получаемые при работе сдвигового регистра составные события (которые он называл «сериями»). Но количественный расчёт составных событий стал возможен после нахождения формулы ф.1 [5], связывающей число составных событий ${}^{n}S$ с числом бросков монеты N (бинарные последовательности) [1-3],:

$$ {}^{n}S = \frac{N}{2^{n+1}} \qquad \text{(ф. 1)} $$

n - длина составного события, например: «000», n=3; «111», n=3; «0000», n = 4; «1111», n = 4.

${}^{n}S$ – ожидаемое число составных событий длины n при N бросках «монеты».

Формула 1 описывает количественные распределения нового типа вероятностных объектов – составных событий, из которых образуются случайные бинарные последовательности (число элементарных бинарных событий N). На основе открытых законов распределения составных событий в случайных бинарных пос-ях [1 - 6] был разработан алгоритм генерации псевдослучайных последовательностей. Этот алгоритм, по аналогии с открытыми Голомбом регистровыми последовательностями «Максимальной длины», можно назвать «Алгоритмом генерации псевдослучайной бинарной последовательности бесконечной длины» [2, 6]. Сгенерированная при помощи этого алгоритма псевдослучайная последовательность «Бесконечной длины» не имеет периодов повторения. В то время как Голомбовские последовательности «Максимальной длины» образуют периоды повторения.

При изучении характеристик составных событий случайных бинарных последовательностей был сделан ряд открытий [7 - 9], эта статья посвящена одному из них. Оказалось, что края фрагментов любой длины обладают более плотным содержанием составных событий, а в центральных областях плотность составных событий меньше, чем в краевых областях.

Основная часть.

В таблице 1 «Равновероятные отрезки» приведены восемь возможных бинарных комбинаций, которые можно построить из трёх бит. Рассмотрим самые простые составные события единичной длины: «0», «1», в таблице 1 они подчёркнуты. Замечаем, что составные события единичной длины распределены не равномерно.

В столбце №2, таблицы 1, только два составных события единичной длины, это: «1» - в строке № 2, и «0» - в строке № 5. Центральное единичное составное событие «1» (таблица 1, строка № 2) должно быть с двух сторон обозначено выделяющими его нулями. Поэтому, комбинации центральных составных событий выглядит так: «0<u>1</u>0»; «1<u>0</u>1».

Краевое одинарное составное событие – то же, либо единица «1», либо ноль «0». Но для выделения краевого события «**<u>0</u>**», смотри таблицу 1, строка № 2, столбец № 1, требуется одно выделяющие событие: «край фрагмента **<u>0</u>**1». Для единичных составных событий расположенных на правом крае фрагмента: «0**<u>1</u>** край фрагмента» (таблица 1, строки № 1, 5) требуется спереди один выделяющий ноль «0».

Таблица 1. Равновероятные отрезки

	Составные события одинарной длины		
	№ 1	№ 2	№ 3
0	0	0	0
1	0	0	**<u>1</u>**
2	**<u>0</u>**	**<u>1</u>**	**<u>0</u>**
3	**<u>0</u>**	1	1
4	**<u>1</u>**	0	0
5	**<u>1</u>**	**<u>0</u>**	**<u>1</u>**
6	1	1	**<u>0</u>**
7	1	1	1

Таблица 2. Одинарные составные события

Левый край	Центр серии	Правый край
		1
0	1	0
0		
1		
1	0	1
		0
Sum = 4	Sum = 2	Sum = 4

Таблица 2 «Одинарные составные события» демонстрирует позиционную неравномерность распределения единичных составных событий внутри фрагментов из трёх бит. Подсчитаны суммы «Sum» одинарных составных событий для каждого из трёх разрядов (столбцов). Число составных событий на «Левом крае» (первый разряд) и на «Правом крае» (третий разряд) одинакова: «Sum = 4». Но в центре серий, (второй разряд) составные события единичной длины выпадают в два раза реже чем по краям: «Sum = 2». В каждом краевом столбце таблицы 2 (левом и правом), содержится составных событий в два раза больше, чем в центральном столбце. Что позволяет говорить об уплотнении составных событий («корочках») на краях серий (отрезков, фрагментов) и о менее плотном наполнении составными событиями внутренних (ядерных) областей фрагментов заданной длины.

В таблицах 1 и 2 была показана комбинаторная раскладка для ${}^{n=1}S$ - составных событий единичной длины (n=1) в фрагментах длиной L=3. Рассмотрим комбинаторную раскладку составных событий ${}^{n}S$ иной длины (n>1). Покажем, что эффект «корочки» - повышенная концентрация составных событий по краям серий (отрезков), не зависит от длин фрагментов L и длин составных событий n. В таблице 3 дана комбинаторная раскладка составных событий ${}^{n}S$ длин $n = 1, \dots, 6$, на длине фрагментов: L=6 (X_1; X_2; X_3; X_4; X_5; X_6). А так же формулы расчёта численности ${}^{n}S$.

Для упрощения понимание таблицы 3, рассмотрим её строку №4, в которой показано, что на длине L = 6 возможны 10 вариантов ${}^{n}_{L}Sum$ размещения составных событий $^{n=4}S$: «0000»; «1111». Составные события $^{n=4}S$ можно разместить на длине L = 6, только в трёх позиция: в левой краевой позиции, в центральной позиции, и в правой краевой позиции.

Левая краевая позиция строки 4 содержит: «0_1; 0_2; 0_3; 0_4; 1_5; X_6 »; «1_1; 1_2; 1_3; 1_4; 0_5; X_6 », где X_6 принимает значения: «0», «1» - всего четыре комбинации: «…1_4; 0_5; 0_6 », «…1_4; 0_5; 1_6 », «…0_4; 1_5; 0_6 », «…0_4; 1_5; 1_6».

Центральная позиция строки 4 содержит две комбинации: «1_1; 0_2; 0_3; 0_4; 0_5; 1_6»; «0_1; 1_2; 1_3; 1_4; 1_5; 0_6».

Правая краевая позиция строки 4 содержит: «X_1; 1_2; 0_3; 0_4; 0_5; 0_6»; «X_1; 0_2; 1_3; 1_4; 1_5; 1_6» , где X_1 принимает значения: «0», «1» - всего возможны четыре комбинации, а значит четыре краевых составных события $^{n=4}S$.

В обеих краевых позициях равная численность составных событий $^{n=4}S$. Все остальные позиции - центральные (принадлежат ядру фрагмента L). В строке 4 видим преобладание краевых составных событий ${}_{edge}^{\ n}S_L^{\ *}$=4 (они помечены «*»), над ядерными составными событиями: ${}_{core}^{\ n}S_L$=2.

Таблица 3. Комбинаторные уплотнения в сериях длин L

n	${}_{edgeX}^{\ \ n}S_L^{\ *} = 2^{L-n}$	${}^{L>n}_{core}S_L = 2^{L-n-1}$					${}^{n}_{L}Sum$	Формулы
1	32*	16	16	16	16	32*	128	${}_{edge}^{\ n}S_L = 2\cdot {}_{edgeX}^{\ \ n}S_L^{\ *} = 2^{L-n+1}$
2	16*	8	8	8	16*		56	${}^{core}_{\ n}S_L = \sum_{n=1}^{n<L} {}_{core}^{\ n}S_L = \frac{L-n-1}{2^{n+1-L}}$
3	8*	4	4	8*			24	
4	4*	2	4*				10	
5	2*	2*					4	${}^{n}_{L}Sum = \begin{cases} {}_{edge}^{\ n}S_L \\ L \ge n \end{cases} + \begin{cases} {}^{core}_{\ n}S_L \\ L > n \end{cases} \to \frac{L-n+3}{2^{n+1-L}}$
6	***2***						2	
						$S = (L+1)\cdot 2^{L-1} =$	224	
В таблице: L=6		$T_N(L) = L \cdot 2^L = 384$						${}^{edge}S_L = \sum_{n=1}^{n=L} {}_{edge}^{\ n}S_L = 2^{L+1}-2$

В таблице 3, численности краевых составных событий имеют в правом верхнем углу символ «*». Краевые события длин n на левом и правом крае фрагмента длины L обозначены: ${}_{edge1}^{\ \ n}S_L^{\ *}$. Формулы для расчёта краевых и центральных составных событий ${}^{L>n}_{core}S_L$, даны в столбце «Формулы».

Сумма краевых событий ${}_{edge}^{\ n}S_L$ (без «*») по строке n представляет сумму левых ${}_{edge1}^{\ \ n}S_L^{\ *}$ и правых краевых событий ${}_{edge2}^{\ \ n}S_L^{\ *}$, ф.2.1:

$$ {}_{edge}^{\ n}S_L = {}_{edge1}^{\ \ n}S_L^{\ *} + {}_{edge2}^{\ \ n}S_L^{\ *} = 2^{L-n+1} \qquad \text{(ф.2.1)} $$

Пример расчета числа краевых событий длины n=2, в фрагментах L=6: ${}_{edgeX}^{\ \ n=2}S_{L=6}^{\ *} = 2^{6-2} = 16^{*}$, (таблица 3, строка n=2). Общая сумма левых и правых краевых составных событий: ${}_{edge}^{n=2}S_{L=6} = 2^{6-2+1}$=32 (нет в таблице 3).

Общее число краевых событий всех длин ${}^{edge}S_L$ есть сумма $\sum_{n=1}^{n=L} {}_{edge}^{\ n}S_L$ всех краевых составных событий по всем длинам $n = 1, \dots, L$, ф.2.2:

$$ {}^{edge}S_L = \sum_{n=1}^{n=L} {}_{edge}^{\ n}S_L = 2 + \sum_{n=1}^{n=L} 2\cdot\left({}_{edgeX}^{\ \ n}S_L^{\ *}\right) = 2^{L+1}-2 \qquad \text{(ф.2.2)} $$

Причём два самых длинных краевых события будут равны в своей длине, длине самой серии: $n = L$, эти, самые длинные краевые события, одновременно существуют на двух краях сразу, поэтому их можно назвать как дважды краевыми. Действительно, сумма значений со звездой «*» в таблице 3, плюс два состояния из столбца: ${}_{edge1}^{\ \ n}S_L^{\ *}$, строка 6 - равна 126: ${}^{edge}S_{L=6} = 2^{6+1} - 2 = 126$.

Между краевыми событиями ${}_{edgeX}^{\ \ n}S_L^{\ *}$ располагаются центральные (ядерные) события ${}_{core}^{\ n}S_L$ длины n, которые находятся в i -х позициях внутри фрагмента L, где $1 < i < L$. В i - ой позиции находится X_i центральных события. Число позиций ${}_{core}^{\ n}X_L$, в которых размещены центральные события, зависит от длины n центрального события ${}_{core}^{\ n}S_L$ и считается по ф.2.3.1:

$$ {}_{core}^{\ n}X_L = \sum X_i = L - n - 1 \qquad \text{(ф.2.3.1)} $$

Рассчитаем по ф.2.3.1 значения для L=6 и сравним полученные результаты с данными из таблицы 3: ${}_{core}^{n=1}X_6$=6-1-1=4; ${}_{core}^{n=2}X_6$=3; ${}_{core}^{n=5}X_6$=0 .

Число центральных событий ${}_{core}^{L>n}S_L$ длины n в любой из допустимых позиций X_i , рассчитывается по ф.2.3.2:

$${}_{core}^{L>n}S_L = 2^{L-n-1} \qquad (ф.2.3.2)$$

Сумма центральных составных событий ${}^{core}S_L$ всех длин n из всех возможных позиций ядра, фрагмента L, ф.2.4:

$${}^{core}_{n}S_L = \sum_{n=1}^{n<L} {}_{core}^{n}S_L = (L-n-1)\cdot\ 2^{L-n-1} \qquad (ф.2.4)$$

В любой строке n таблицы 3 сумма краевых ${}_{edge}^{n}S_L$ и центральных ${}_{core}^{n}S_L$ составных событий равна ${}_{L}^{n}Sum$, ф.2.5:

$${}_{L}^{n}Sum = {}_{edge}^{n}S_L + {}^{core}_{n}S_L = \begin{cases} 2^{L-n+1} \\ L \ge n \end{cases} + \begin{cases} (L-n-1)\cdot 2^{L-n-1} \\ L > n \end{cases} \qquad (ф.2.5)$$

Примеры расчёта столбца ${}_{L}^{n}S$, таблицы 3 (строки 1- 6), по формуле 2.5, для L=6: ${}_{L=6}^{n=1}S = 2^{6-1+1} + (6-1-1)2^{6-1-1} = 128$ (строка 1, таблицы 3); ${}_{L=6}^{n=2}S$ =56 (строка 2); ${}_{L=6}^{n=3}S$ =24 (строка 3); ${}_{L=6}^{n=4}S$ =10 (строка 4); ${}_{L=6}^{n=5}S$ =4 (строка 5). При n=6 не выполняется условие $L > n$, поэтому нет слагаемого: $(L-n-1)\cdot 2^{L-n-1}$, и ${}_{L=6}^{n=6}S$ =2.

Из состоящего из L бит слова, можно создать 2^L комбинаций (пример для L=3 приведён в таблице 1). Причём, каждая комбинация будет состоять из L бит. Назовем число L бит - длиной фрагмента (длиной отрезка). Число элементарных событий N, из которых образованы все 2^L комбинации, можно сравнить с периодом $T_N(L)$, после которого комбинации начнут повторяться. Период $T_N(L)$ рассчитывается по ф.2.6:

$$N = T_N(L) = L\cdot 2^L \qquad (ф.2.6)$$

Где N, для комбинаторного случая, не подвержено случайным вероятностным флуктуациям, таблицы 1 – 3, оно дискретно. А в потоковых пос-тях N принимает любые значения (равные числу выпадений реальных бинарных событий), а не только дискретные значения, получаемые из ф.2.6.

Для примера рассчитаем комбинаторный период N: где, $N = T_N(L)$, для таблиц 1 и 3: $T_N(3) = 3\cdot 2^3$=24; $T_N(6) = 6\cdot 2^6$=384.

Рассмотрим аномальную ситуацию для суммы краевых фрагментов ${}_{edge}^{n}S_L = 2\cdot {}_{edgeX}^{n}S_L{}^*$, при: $L = n$. Аномалия возникает при равенстве длин фрагмента L и составного события n, которое должно содержаться внутри фрагмента L. В этом случае, $L = n$, фрагмент является L одновременно и краевым событием n. Эта аномалия является вопросом терминологий, и требует отдельной проработки, но ф.2.7 и для этого случая выдаёт правильный результат:

$${}_{edge}^{n}S_L = 2\cdot {}_{edgeX}^{n}S_L{}^* = 2^{L-n+1} \qquad (ф.2.7)$$

Действительно, при $L = n$: ${}_{edge}^{n=L}S_L = 2^{L-L+1} = 2$.

Результаты экспериментов по обнаружению составных событий на краях и в области ядра L - фрагментов потоковой последовательности.

В комбинаторных таблицах (пример для L=3 дан в таблице 1), в рамках одного периода $T_N(L)$ любое состояние, из 2^L возможных, встречается один раз. В реальных потоках случайных бинарных событий выпадение комбинаций происходит случайным образом. Наберём статистику выпадений для каждого из 2^L возможных состояний. Перейдём от комбинаторных таблиц (типа таблица 1) к потокам случайных бинарных событий с произвольным числом элементарных событий N в них. Рассмотрим результаты двух компьютерных экспериментов, таблица 4, по поиску составных событий единичной длины («**1**» , «**0**») во фрагментах разных длин: $L_1 = 6$ и $L_2 = 11$. В каждом из двух потоков элементарных событий было одинаковое количество случайных бинарных событий $N = 2\cdot 10^7$. В первом бинарном потоке, последовательно выпадающие случайные бинарные события делились на последовательные фрагменты по шесть ($L_1 = 6$) элементарных событий («0», «1»). Во втором бинарном потоке, последовательно выпадающие случайные бинарные события делились на последовательные фрагменты по одиннадцать

($L_2 = 11$) элементарных событий. Во всех фрагментах производился поиск как краевых, так и ядерных единичных составных событий.

В таблице 4 даны количества найденных краевых и ядерных составных событий (в скобках даны составные события, рассчитанные по: ф.3.1, ф.3.3). В таблице 4 символ L_X, обозначает элементарное событие («0», «1») и его номер в фрагменте длины L, с которого начинается рассматриваемое составное событие ${}^{n=1}S_{L=6}$, примеры: L_1 = ($L_1$23456); L_2 = (1L_2 3456); L_4 = (123 L_4 56); L_6 = (12345 L_6).

Таблица 4. Количества найденных краевых и ядерных составных событий

1	2	3	4	5	6	7
L_X	${}_{edgeX}^{n=1}S_{L=6}{}^*$	${}_{core}^{n=1}S_{L=6}$	${}_{edgeX}^{n=1}S_{L=11}{}^*$	${}_{core}^{n=1}S_{L=11}$	${}_{edgeX}^{n=3}S_{L=6}{}^*$	${}_{core}^{n=3}S_{L=6}$
L_1	1667732 (1666666)	no	908979 (909091)	no	416297 (416667)	no
L_2	no	833883 (833333)	no	455251 (454545)	no	208774 (208333)
L_3	no	833956	no	454672	no	208356
L_4	no	832189	no	453999	417296	no
L_5	no	832843	no	454053	no	no
L_6	1665935	no	no	454563	no	no
…			no	…		
L_{11}			908999	no		
Sum	(3333333)	3332871 (3333333)	1817978 (1818182)	(4090909)	833593 (833333)	417130 (416667)
${}_L^nSum$	6666538		5909416		1250723	
Z	3333333		1818181		3333333	
N	$2 \cdot 10^7$		$2 \cdot 10^7$		$2 \cdot 10^7$	
В скобках даны теоретически рассчитанные величины						

При делении последовательности N элементарных событий на фрагменты (отрезки, серии) по L событий в каждом фрагменте, получим $Z = \frac{N}{L}$ таких фрагментов. В строке «Z» таблицы 4 даны численности фрагментов в N пос-ти, для L = 6 и L = 11.

Таблица 4 составлена для составных событий (пример не полярных ${}^nS$: «111», «000», «0», «1»). Z_L - число фрагментов длиной L в пос-ти из N событий. Элементарное событие левого края фрагмента это: «1» или «0». Вероятность его повторения («1», «0») равна 0,5. В половине случаев ($Z/2$) это элементарное событие повторится («11», «00»), а в половине случаев не повторится («10», «01»). Это значит, что с левого края находится $Z/2$ составных событий длины один: ${}_L^1S = \frac{Z}{2} = \frac{N}{2 \cdot L}$. По соображениям симметрии на правых краях фрагментов L будет столько же (с точностью до случайной флуктуации) составных событий ${}_L^1S$. Симметрия существует для составных событий любых длин $n < L$, то есть: ${}_L^{n<L}S$.

По формуле 3.1 рассчитываются краевые составные события на одном из двух краёв фрагмента ${}_{edgeX}^{n}S_L{}^*$:

$$ {}_{edgeX}^{n}S^* = \frac{N}{T_N(L)} \cdot {}_{edgeX}^{L>n}S_L{}^* = \frac{N}{L \cdot 2^L} \cdot 2^{L-n} = \frac{N}{L \cdot 2^n} \qquad (ф.3.1) $$

Число краевых событий единичной длины ($n = 1$): ${}_{edgeX}^{n=1}S_L{}^* = \frac{N}{L \cdot 2}$.

По ф. 3.2 рассчитываются краевые составные события на двух краях фрагментов, включая описанную выше аномальную ситуацию, ${}_{edgeX}^{n}S_L{}^*$:

$$ {}_{edge}^{n}S_L(N) = 2 \cdot {}_{edgeX}^{n}S_L{}^* = \frac{N \cdot 2}{L \cdot 2^n} \qquad (ф.3.2) $$

В таблице 4 даны экспериментально полученные значения «${}_{edgeX}^{n=1}S_{L=6}{}^*$» для длин фрагментов $L = 6$: ${}_{edgeX}^{n=1}S_{L=6}(N)^*$={1667732, 1665935}, они хорошо совпадают с рассчитанным по формуле значением: ${}_{edgeX}^{n=1}S_{L=6}(N)^*$= 1666666. Экспериментально полученные числа краевых единичных событий в фрагментах длины $L = 11$: ${}_{edgeX}^{n=1}S_{L=11}(N)^*$={908979, 908999}, расчёт: ${}_{edgeX}^{n=1}S_{L=11}(N)^*$=909091. Краевые составные события $n = 3$: ${}_{edgeX}^{n=3}S_{L=11}(N)^*$={416297, 417296}, расчёт: ${}_{edgeX}^{n=3}S_{L=11}(N)^*$=416667.

Численность внутренних (ядерных) единичных составных событий ${}_{core}^{n=1}S_L$ следует из численности ${}_{edgeX}^{n=1}S_L(N)^*$. Действительно, численность элементарных событий на позиции L_2 , образующих своим выпадением единичные левые краевые составные события, равна численности этих краевых событий: ${}_{edgeX}^{n=1}S_{L=11}(N)^*$=909091. А так как половина выпадающих элементарных событий в следующей позиции L_3

будет не равна выпавшим событиям в L_2, то они фактом своего выпадения в L_3 образуют ядерное составное событие в позиции L_2. Иными словами, так как вероятность (не) повторения события L_2 в позиции L_3 равна 0.5, то половина L_2, образовавших ${}_{edgeX}^{n=1}S_L(N)^*$, превратятся в ядерные события единичной длины: ${}_{core}^{n=1}S_L = 0.5 \cdot {}_{edgeX}^{n=1}S_L(N)^*$. Следствием равной вероятности существования ядерных составных событий равной длины, являются их одинаковые количества внутри всех фрагментов L.

По ф.3.3 считают числа ядерных составных событий находящихся в любой одной позиции из возможных ядерных позиций, для любых длин n:

$$ {}_{core}^{n}S_L = 0.5 \cdot {}_{edgeX}^{n}S_L(N)^* = \frac{N}{L \cdot 2^{n+1}} \qquad (ф.3.3)$$

Для примера рассмотрим распределение ядерных составных событий длиной n=3, внутри фрагментов длины L = 6, таблица 4, столбец 7. Покажем через звёздочки «***» их все возможные положения внутри фрагментов L (1,2,3,4,5,6): два краевых положения ${}_{edgeX}^{n=3}S_{L=6}{}^*$: «***,4,5,6»; «1,2,3,***»; два ядерных (центральных) положения ${}_{core}^{n=3}S_{L=6}$: «1,***,5,6»; «1,2,***,6».

Численности краевых событий ${}_{edgeX}^{n=3}S_{L=6}{}^*$ левых краёв равны числу фрагментов $Z_{L=6} = \frac{N}{L}$. Поскольку, вероятность повторения элементарного события (находящегося с левого края фрагмента «1» или «0»), равна 0.5, то в половине случаев $(Z/2)$ это элементарное событие повторится («11» или «00»), а в половине случаев не повторится («10» или «01»). Что означает, что с левого края находится $Z/2$ составных событий длины один: ${}_{L}^{1}S = \frac{Z}{2} = \frac{N}{2 \cdot L}$. Далее из состояний: «00», «11», в половине случаев будет достигнуты состояния: «000», «111» (а в другой половине будут состояния: «001», «110» - не подлежащие дальнейшему анализу). Далее, в половине случаев из состояний: «000», «111» будут достигнуты состояния: «1110», «0001», что соответствует: ${}_{L}^{3}S = \frac{Z}{2 \cdot 2 \cdot 2} = \frac{N}{8 \cdot L}$. Математическое ожидание для ${}_{edge1}^{n}S_L{}^*$ получаем по ф.3.1: ${}_{edgeX}^{n}S_L{}^* = \frac{N}{L \cdot 2^n} =$ 416667, оно хорошо совпадает с экспериментально полученными величинами: 416297; 417296 (столбец 6, таблицы 4).

Для получения числа ядерных составных событий, в какой либо из позиций фрагмента L, проведём следующий расчёт. Всего $Z = \frac{N}{L}$ фрагментов. Поэтому число цепочек («0<u>111</u>0»; «1<u>000</u>1» - ядерные составные события n=3) выпавших с начала L фрагментов, будет (ф.3.3): ${}_{core}^{n=3}S_{L=6} = 2 \cdot \frac{N}{L \cdot 2^{n+2}} = \frac{N}{L \cdot 2^{n+1}}$ =208333, что хорошо совпадает с экспериментально полученными величинами: 208774; 208356 (смотри столбец 7, таблицы 4).

Выразим формулу ядерных составных событий (3.3) для каждой одной возможной позиции в ядре фрагмента (а не сумму событий во всех возможных позициях ядра) через составные события потоковой пос-ти, то есть через ф.1. Получаемая ф.3.4.1 устанавливает связь между числом составных событий длины n в потоковой пос-ти из N элементарных событий (эл) и числом ядерных событий внутри $b = \frac{N}{L}$ фрагментов:

$$ {}_{core}^{n}S(N)_L = \frac{N}{L \cdot 2^{n+1}} = \frac{b}{N} \cdot {}^{n}S(N) = \frac{1}{L} \cdot {}^{n}S(N) \qquad (ф.3.4.1)$$

Рассчитаем, по ф.3.4.1, значения и запишем их в таблицу 4, к экспериментально полученным данным: ${}_{core}^{n=3}S_{L=6} = \frac{1}{6} \cdot \frac{2 \cdot 10^6}{2^{3+1}}$ =208333; ${}_{core}^{n=1}S_{L=11} = \frac{1}{11} \cdot \frac{2 \cdot 10^6}{2^{1+1}} = 454545$; ${}_{core}^{n=1}S_{L=6} = \frac{1}{6} \cdot \frac{2 \cdot 10^6}{2^{1+1}}$ =833333.

Приравняв N из ф.3.4.1 к взятому из ф.2.6 периоду $T_N(L) = L \cdot 2^L$, получим ф.2.3.2, смотри ф.3.4.2:

$$ {}_{core}^{n}S(T_N)_L = \frac{T_N(L)}{L \cdot 2^{n+1}} = \frac{L \cdot 2^L}{L \cdot 2^{n+1}} = 2^{L-n-1} \qquad (ф.3.4.2)$$

Исключив N из ф.3.4.1 получим мизесовскую частоту ${}_{core}^{n}f_L$, ф.3.4.3:

$$ {}_{core}^{n}f_L = \frac{1}{L \cdot 2^{n+1}} \qquad (ф.3.4.3)$$

Заметим - отношение частот соседних длин: n_1; n_2 равно 2.

Выразим формулу суммы краевых составных событий с обоих краёв фрагментов (3.2) ,через составные события потоковой последовательности, то есть через ф.1. Получаемая ф.3.5.1 устанавливает связь между числом составных событий длины n потоковой пос-ти из N элементарных событий (эл) и числом b, краевых событий фрагментов: $b = \frac{N}{L}$

$$
{}_{edge}^{\;\;n}S_L(N) = \frac{N \cdot 2}{L \cdot 2^n} = \frac{b}{2^{n-1}} = \frac{4}{L} \cdot {}^{n}S(N) = \frac{4}{L} \cdot \frac{N}{2^{n+1}} \qquad (ф.3.5.1)
$$

Рассчитаем по ф.3.5.1 *сумму* краевых составных событий длинны n с обоих краёв (в таблице 4 экспериментально полученные данные по каждому краю в 2 раза меньше, так как ${}_{edge}^{\;\;n}S_L(N)$ их сумма): ${}_{edge}^{n=1}S_{L=6}(N) = \frac{4}{6} \cdot \frac{2 \cdot 10^6}{2^{1+1}} = 3333333$; ${}_{edge}^{n=1}S_{L=11}(N)$ =1818182; ${}_{edge}^{n=3}S_{L=6}(N)$ =833333.

Приравняв N из ф.3.5.1 к взятому из ф.2.6 периоду $T_N(L) = L \cdot 2^L$, получим ф.2.1, смотри ф.3.5.2:

$$
{}_{edge}^{\;\;n}S_L(T_N) = \frac{4}{L} \cdot \frac{L \cdot 2^L}{2^{n+1}} = 2^{L-n+1} \qquad (ф.3.5.2)
$$

Исключив N из ф.3.5.1 получим мизесовскую частоту ${}_{edge}^{\;\;n}f_L$, ф.3.5.3:

$$
{}_{edge}^{\;\;n}f_L = \frac{4}{L \cdot 2^{n+1}} \qquad (ф.3.5.3)
$$

Заметим - отношение частот соседних длин: n_1; n_2 равно 2.

По формуле 3.6 рассчитывается общая сумма ${}^{core}S_L$ (таблица 4, строка «Sum» столбцы 3 и 7) ядерных составных событий длин n по всем позициям их размещения внутри фрагмента длины L. В каждой отдельной внутренней позиции фрагментов L находится ${}_{core}^{\;\;n}S(N)_L$ ядерных событий, а общее число событий является суммой по всем внутренним позициям: ${}^{core}S_L = \sum {}_{core}^{\;\;n}S_L$, где ${}_{core}^{\;\;n}S_L$- рассчитывается по ф.3.3. Комбинаторное число внутренних, позиций X_i зависит от длины фрагмента L и от длины ядерного события n, ф.2.3.1: ${}_{core}^{\;\;n}X_L = L - n - 1$. Из выше сказанного и ф. 2.4, получаем ф.3.6:

$$
{}^{core}S_L = (L - n - 1) \cdot\ 2^{L-n-1} \cdot \frac{N}{L} = (L - n - 1) \cdot \frac{{}^{n}S(N)}{L} \qquad (ф.3.6)
$$

Действительно, учитывая, что $T_N(L) = L \cdot 2^L$, ф.2.6, получаем:

$$
{}_{\;\;n}^{core}S_L(N) = \frac{L-n-1}{2^{n+1-L}} \cdot \frac{N}{T_N(L)} = \frac{L-n-1}{L \cdot 2^{L-L}} \cdot \frac{N}{2^{n+1}} = \frac{L-n-1}{L} \cdot\ {}^{n}S(N)
$$

Общее число составных событий ${}_{L}^{n}Sum$ (всех краевых и всех центральных) будет являться суммой: ${}^{core}S_L$ - ф.3.5 и ${}^{core}S_L$ - ф.3.6: ${}_{L}^{n}Sum = {}_{edge}^{\;\;n}S_L + {}_{\;\;n}^{core}S_L = \begin{cases} 2^{L-n+1} \\ L \geq n \end{cases} + \begin{cases} \frac{L-n-1}{2^{n+1-L}} \\ L > n \end{cases}$. По ф.3.7 считаются составные события длины n (краевые плюс центральные): ${}_{L}^{n}Sum = {}_{edge}^{\;\;n}S_L(N) + {}_{\;\;n}^{core}S_L(N)$.

$$
{}_{L}^{n}Sum = 4 \cdot \frac{{}^{n}S(N)}{L} + (L - n - 1) \cdot \frac{{}^{n}S(N)}{L} = \frac{L-n+3}{L} \cdot\ {}^{n}S(N) \qquad (ф.3.7)
$$

Для примера рассчитаем по ф.3.7 суммы ${}_{L}^{n}Sum$ краевых и ядерных составных событий из всех позиций фрагмента, таблицы 4, имея в виду, что ${}^{n}S(N)$ рассчитывается по ф.1: ${}_{L=6}^{n=1}Sum = \frac{8}{6} \cdot \frac{N}{2^{1+1}}$ =6666667; ${}_{L=11}^{n=1}Sum = \frac{13}{11} \cdot \frac{N}{2^{1+1}}$ =5909091; ${}_{L=6}^{n=3}Sum = \frac{6}{6} \cdot \frac{N}{2^{3+1}}$ =1250000.

Техника управления числом угаданных серий.

Приведём пример техники управления числом угаданных серий основанной на работе с краевыми уплотнениями и эффектом экранирования (игра Пенни - более известное его название). Так как центральных событий меньше чем краевых (смотри таблицы 1, 2), то вероятность выпадения подряд двух центральных событий «0**1**0» + «0**1**0» = «0**1**00**1**0» меньше вероятности угадывания подряд трёх краевых событий «01» + «01» + «01» = «010101». Поделив потоковую последовательность на серии равной длины: L=3. И, начав учитывать в последовательных выпадениях таких серий (L=3) образуемые ими цепочки «0**1**00**1**0» и «0**1**0**1**0**1**», согласно таблицам 1 и 2, найдём численное превышение цепочек «0**1**0**1**0**1**» над «0**1**00**1**0» цепочками.

Так как длина центрального события «0**1**0» не равна длине краевого события «01», то нужно оговорить условия проведения поискового эксперимента, которые с классической точки зрения уравняют вероятность обнаружения цепочек «0**1**00**1**0» и «0**1**0**1**0**1**» друг с другом. Для выравниваний вероятностей предлагаются две формулы набора: для «0**1**00**1**0» - (I) «C1_1, C2_2, C3_3» + (II) «C1_4, C2_5, C3_6» + (III) «*,*,*», а для «0**1**0**1**0**1**» - (I) «*,E1_1, E2_2» + (II) «*,E1_3, E2_4» + (III) «*,E1_5, E2_6».

Поисковый алгоритм 1

Серия «0**1**0**1**0**1**» набирается по формуле: (I) «*,E1_1, E2_2» + (II) «*,E1_3, E2_4» + (III) «*,E1_5, E2_6». Где: буква «E» - первая буква от слова $edge$, обозначает работу с краевым составным событием; Ek_i – предсказываемые выпадающие элементарные бинарные события (стороны монеты); «*» - обозначает выпадение события, значение которого не предсказывается и не учитывается ни каким образом. Серия «010101» может быть набрана только в случае выпадения трёх серий подряд: (I) «*,0,1» + (II) «*,0, 1» + (III) «*, 0, 1». В случае угадываний значений: E1(=0), E2(=1), в трёх сериях подряд, происходит учёт серии «010101», и начинается поиск новой пос-ти серий: (I) + (II) + (III). Если после выпадения серий (I) не выпала серия (II) , то начался новый поиск серии (I). Если после выпадения серий (I) + (II) не выпала серия (III), то начался новый поиск серии (I).

Поисковый алгоритм 2

Серия «010010» набирается по формуле: (I) «$C1_1$, $C2_2$, $C3_3$» + (II) «$C1_4$, $C2_5$, $C3_6$» + (III) «*,*,*». Где: C1, C2, C3 – обозначают сделанные предсказания для выпадающих величин (сторон монеты). Серия «010010» может быть набрана только в случае выпадения двух серий подряд: (I) «010» + (II) «010». В случае угадывания значений C1(=0), C2(=1), C3(=0) в двух сериях подряд происходит зачёт угаданной серии «010010» и начинается новые предсказания выпадений последовательности серий: (I) + (II). Если после выпадения угаданной серии (I) не выпала предсказанная серия «010», то поиск серий: (I) + (II) начинается заново, с поиска (предсказания выпадения) серии (I). В серии (III): «*,*,*» все три выпадающих события являются не угадываемые и не рассматриваемые, её смысл заключается в том, что бы уравнять числа выпадающих событий в набираемой серии «010010» с числом выпадающих элементарных событий в наборе «010101». Скорее всего, присутствие серии (III) не обязательно.

Результаты работ поисковых алгоритмов

В таблице 5 даны результаты поиска, по вышеописанным правилам, серий: «010101»; «010010», в пос-ти случайных бинарных событий.

Таблица 5. Поиск серий из краевых и центральных составных событий

Искомая серия	Найдено серий	Формула набора серии	Тип событий	Игрок
«010101»	79165	0, 1 + 0, 1 + 0,1	Краевой	Btn585
«010010»	61586	0, 1, 0 + 0, 1, 0	Центральный	Btn586
Число элементарных бинарных событий на искомую серию: $N = 2 \cdot 10^7$				

Таблица 5 демонстрирует, что было сделано 13333333 предсказания выпадения монеты, учитывая существование краевых составных событий, в рамках этих предсказаниях было угадано 79165 серии «010101»; и что было сделано 13333335 других предсказаний, учитывающих существование центральных составных событий, в рамках этих предсказаниях было угадано 61586 серии «010010». При одинаковых количествах предсказаний, по технике набора краевых событий было найдено 79165 серии «010101», по технике набора центральных событий было найдено 61586 серии «010010».

Одна и та же формула описывает распределение инверсий внутри фрагментов длин n и угадываний внутри фрагментов длин $(n-1)$.

Инверсией называется два стоящих вместе неравных бинарных элементарных события: «01» или «10». Из фрагмента (слова) длины L можно создать $I(C_i^L)$ комбинации с i инверсиями, ф.4.1 [14]:

$$I(C_i^L) = \frac{2 \cdot (L-1)!}{i! \cdot (L-1-i)!}\ ; \text{ где } L = 1,2,3, \dots;\ i \le L-1 \qquad (\text{Ф. } 4.1)$$

Где, $I(C_i^L)$– число комбинаций содержащих внутри себя i инверсии в любых комбинациях; L –длина слова; i – число инверсий внутри слова L.

Ф.4.2 – формула расчёта мат. ожидания инверсий ${}^{L}I(i)$ для слов длиной n, содержащих $0 \le i \le L-1$ инверсий, в файле из N эл (элементарных событий). Ф. 4.2 получается умножением комбинаторного распределения $I(C_i^L)$ на множители: $\frac{N}{L} \cdot \frac{1}{2^L}$. Где: $\frac{N}{L}$ - число фрагментов длины L в бинарной последовательности из N элементарных событий; множитель $\frac{1}{2^L}$ является весом для каждого уникального инверсионного состояния.

$${}_{i}^{L}I = I(C_i^L) \cdot \frac{N}{L} \cdot \frac{1}{2^L} = \frac{2 \cdot (L-1)!}{i! \cdot (L-1-i)!} \cdot \frac{N}{L} \cdot \frac{1}{2^L} \qquad (\text{Ф.}4.2)$$

где $L = 1,2,3, \dots;\ 0 \le i \le L-1$

В таблице 6, в столбце $I(C_i^L)$ приведены количества инверсионных состояний с числом инверсий i в каждом из них, комбинаторно возможных во фрагментах из L элементарных бинарных событий (эл), (ф.4.1). В столбце ${}_{i}^{L}I$ даны количества фрагментов из L эл, найденные в пос-ти из $5 \cdot 10^8$ бинарных событий, в которых содержатся по i инверсии (ф.4.2).

Будем понимать под «Совпадением» равенство величин выпавшего и предсказанного бинарного события (например, предсказывается перед подбрасыванием монеты выпадение орла и подброшенная монета выпала стороной с орлом вверх). Обозначим событие совпадения буквами ${}^{L_e}_{n}OV$ (от английского overlap), в которых n - число совпадений предсказанного события с элементарными событиями фрагмента, L_e – длина фрагмента в элементарных событиях (в элах).

Процедура поиска совпадений ${}^{L_e}_{n}OV$ заключается в следующем. Так как вероятность выпадения любого предсказанного значения(«0», «1») равна 0.5, то можно произвольно зафиксировать в качестве предсказательной величины либо «0», либо «1». И сравнивать с зафиксированной величиной выпадающие случайные бинарные величины. Пос-ть $F0.5(5 \cdot 10^8)$ - состоит из $5 \cdot 10^8$ случайных равновероятных бинарных событий, делится на последовательные фрагменты по L_e событий в каждом. В каждом фрагменте подсчитывается число элементарных событий n, равных выбранному предсказанному значению. Найденные в L_e совпадения суммировались в раздельных счётчиках (каждый счётчик учитывает только фрагменты с определённым числом совпадений в них). Так как существует такое явление, как «блуждание около нулевой отметки», которое заключается в том, что число выпавших «0» и «1» бывает крайне редко равно друг другу при конкретном числе бинарных событий N, то, обычно: $f_0(N) \neq f_1(N)$. Описание явления «блуждании» выходит за рамки статьи, но отметим, что для уменьшения дисперсии вызванного этим явлением автор предлагает способ «Привязки к выпавшим совпадениям». Суть «Привязки к выпавшим совпадениям» в том, что насчитанные счётчиками события делятся на показания счётчика фиксировавшего полностью не угаданные состояния. Или наоборот, делить на показания счётчика, который фиксировал число фрагментов, в которых угаданы все предсказания ${}^{L_e}_{n=0}OV$.

В таблице 6, значения столбца «$2 \cdot {}^{L_e}_{n}OV / {}^{L_e}_{n=0}OV$» получены делением показаний счётчиков на показания счётчика фиксировавшего полностью не угаданные состояния ${}^{n=0}_{L_e}OV$, с умножением частного на два. Формулы 4.1 и 4.2 оперируют сразу двумя инверсиями: и «01» и «10» (поэтому в них присутствует множитель «2»), что эквивалентно одновременному поиску угадываний и по «0» и по «1». Поисковый эксперимент искал совпадения только с шаблоном «1», а для учёта событий равных «0», надо удвоить найденное число совпадений. Этим объясняется множитель два в заголовке: «$2 \cdot {}^{L_e}_{n}OV / {}^{L_e}_{n=0}OV$».

Таблица 6. Инверсии и совпадения в фрагментах $F0.5(2 \cdot 10^7)$

		Распределение инверсий в L фрагментах		Распределение совпадений в фрагментах длины $L_e = L - 1$		
L	i, n	$I(C_i^L)$	${}^{L}_{i}I_{\text{эксп}}$	${}^{L_e}_{n}OV$	${}^{L_e}_{i=n}OV = \frac{L \cdot {}^{L}_{i}I}{L-1}$	$2 \cdot {}^{L_e}_{n}OV / {}^{L_e}_{n=0}OV$
	0	2	19531	22514	22321	2
	1	14	136718	156499	156250	13,9935
	2	42	410156	468412	468750	41,9390
	3	70	683593	780475	781250	69,8825
8	4	70	683593	781916	781250	69,9238
	5	42	410156	468838	468750	41,9519
	6	14	136718	156273	156250	13,9836
	7	2	19531	22216	22321	1,9950
	8	-------	-------	-------	-------	-------
	0	2	44642	52376	52083	2
	1	12	267857	313137	312500	11,9738
	2	30	669642	780120	781250	29,9179
7	3	40	892856	1041053	1041667	39,8967
	4	30	669642	781944	781250	29,9320
	5	12	267857	312781	312500	11,9753
	6	2	44642	51923	52083	1,9940
	7	-------	-------	-------	-------	-------
		2) Btn293	14)Btn345	Btn594		15_2)Btn594

В столбце ${}^{L}_{i}I_{\text{эксп}}$ таблицы 6 даны экспериментально найденные численности инверсий i в L фрагментах.

В столбце «${}^{L_e}_{i=n}OV = \frac{L \cdot {}^{L}_{i}I}{L-1}$ » таблицы 6 даны рассчитанные по ф. 4.3 математические ожидания выпадений фрагментов, в которых будет обнаружено по n выпадений либо «0» либо «1».

В столбце «$2 \cdot {}^{L_e}_{n}OV / {}^{L_e}_{n=0}OV$» таблицы 6 рассчитаны коэффициенты, полученные делением найденных в поисковом эксперименте удвоенных значений ${}^{L_e}_{n}OV$ на значение ${}^{L_e}_{n=0}OV$.

Заметим, что при $n = L$ сумма всех возможных бинарных комбинаций (состояний) в фрагменте (слове): $\sum C_n^m = 2^n$, равна сумме всех инверсных состояний в этого же фрагмента: $\sum I(C_i^L) = 2^L$, и: $\sum C_n^m = \sum I(C_i^L)$.

В таблице 6 теоретически рассчитанные по ф.4.1 комбинаторные значения в столбце $I(C_i^L)$ хорошо совпадают с полученными в результате поискового эксперимента значениями в столбце «$2 \cdot {}_{n}^{L_e}OV / {}_{n=0}^{L_e}OV$».

Число фрагментов ${}_{n}^{L_e}OV$ длиной L_e с числом n угаданных событий в них рассчитывается по ф.4.3:

$$ {}_{n}^{L_e}OV = \frac{L_e!}{n! \cdot (L_e - n)!} \cdot \frac{N}{L_e \cdot 2^{L_e}} \text{ ; где } L_e = L - 1 \qquad (\text{Ф.4.3}) $$

Заметим, что величина полученная по ф.4.2 не равна величине полученной по ф.4.3: ${}_{i}^{L}I \neq {}_{n}^{L_e}OV$.

По ф.4.4 рассчитываются комбинаторные численности:

$$ OV\left(C_n^{L_e}\right) = \frac{2 \cdot L_e!}{n! \cdot (L_e - n)!} \qquad (\text{Ф.4.4}) $$

Рассчитаем для примера по ф.4.4: $OV\left(C_{n=3}^{L_e=6}\right) = \frac{2 \cdot 6!}{3! \cdot (6-3)!} = 40$, что хорошо совпадает с экспериментальным значением - 39,897.

Для примера рассчитаем ${}_{n=3}^{L_e=6}OV$ из таблицы 6: $I(C_{i=3}^{L=7}) \cdot \frac{N}{L} \cdot \frac{1}{2^L} \cdot \frac{L}{L-1} = \frac{2 \cdot (L-1)!}{i! \cdot (L-1-i)!} \cdot \frac{N}{L-1} \cdot \frac{1}{2^L} = \frac{L_e!}{n! \cdot (L_e - n)!} \cdot \frac{N}{L_e \cdot 2^{L_e}} = \frac{6!}{3! \cdot (6-3)!} \cdot \frac{N}{6 \cdot 2^6} = 1041667.$

Обсуждение

Первая часть статьи, по сути, поднимает вопрос об уравнивании шансов для выпадения сложных составных (наборных) событий. Ожидается, что правильно уравненные шансы приводят к статистически одинаковым количествам выпадений (обнаружений) составных (сложных) событий. Интуитивные представления о выравнивании возможностей игроков, первая часть статьи, привела к систематическому неравенству выпадений их серий. На демонстрации этого систематического неравенства шансов игроков первая часть была закончена (таблица 5). Формат статьи не позволил уйти в обсуждение условий для уравнивания шансов составных событий в равновероятных бинарных (потоковых) последовательностях $F0.5(N)$.

Принцип неопределённости Гейзенберга, введённый для физики, пришло время перенести в теорию вероятности. Так как дестабилизация вызываемая процедурой инструментального измерения, описываемая принципом неопределённости, так же в полной мере работает при организации составных событий в бинарных пос-тях. Так как любая попытка формирования частей составных событий и объединение этих частей в логически целое событие (одну сущность) приводит к искажению естественного состояния (пропорций) бинарной последовательности. Что собственно и было показано в начале статьи. Причём, интуитивно очевидные меры по выравниванию шансов игроков не смогли в реальности уравновесить их шансы. Объявляя искомую серию, например «10101», мы тем самым искажаем краями этой серии те составные события, которые содержат эти самые придуманные нами края. Задавая только одну длину серии, не оговаривая её вида, мы опять, как показано в начале этой статьи, деформируем как края тех последовательностей, которые принимаем за серии, так и локальную структуру последовательности, в которой ищем эти серии. Как показано в моих работах, от того что мы ищем, и как мы ищем, зависит эффективность и количество того, что мы находим. И интуитивная вера в равновероятные исходы обнаружения серий одинаковой длины в последовательностях $F0.5(N)$ абсолютно ошибочна. В наследство, после «кончины» этой интуитивной веры в равные вероятности выпадений серий одинаковой длины, возник вопрос о том, как создать равные вероятности обнаружения для каждой из конкурирующих друг с другом серий одной длины, при их поиске в потоковой пос-ти $F0.5(N)$?

Теорема «О равенстве суммы первых угаданных полярных событий числу серий».

Для демонстрации особенности работы с сериями (фрагментами) произвольной длины опишем эффектный физико-математический фокус на современных компьютерах. Фокусник показывает, что человеческий мозг работает с гигантским набором данных не медленнее компьютера.

Секрет фокуса в том, что сумма элементарных событий, которые образуют первые краевые события, в каждом фрагменте равна, с точностью до случайной флуктуации, числу этих случайных фрагментов. Доказываемая теорема, в свою очередь, базируется на теореме: «О амплитудно-частотной характеристике идеальной бинарной случайной последовательности» [2, 5].

Опишем этот фокус. Фокусник просит создать любое количество серий (фрагментов) произвольно меняющейся от фрагмента к фрагменту длины, но большей некоторой минимальной длины L_{min}, от физического генератора случайных бинарных равновероятных величин («0», «1»). Число фрагментов должно быть больше 10^3 (лучше больше 10^9). Принцип расчёта минимальной длины L_{min} будет дан ниже.

Фокусник угадывает по сумме элементарных событий, которые образуют первые краевые события, в каждом фрагменте, число фрагментов, которые создал генератор случайных чисел, с точностью превышающей 5%. Суммируемые элементарные события, образующие первые краевые события отбираются следующим образом. Фрагменты, начинающиеся с «0», в нарастающую сумму вносят нулевой вклад. В фрагментах начинающихся с «1» считают сколько «1» выпало подряд, начиная с первой единицы. Например: k-й фрагмент начинается так: «$1_1 1_2 0_3 0_4 \ldots e_{50}$» - две единицы подряд, фрагмент k+1 начинается так: «$1_1 0_2 0_3 0_4 \ldots e_{50}$» - одна единица, фрагмент k+2 начинается так: «$1_1 1_2 1_3 0_4 \ldots e_{50}$» - три единицы подряд, и т.д. Компьютер должен просуммировать все единицы стоящие в начале этих фрагментов (2+1+3 = 6), и выдать эту сумму. Фокусник тут же её повторяет с заявлением, что с точностью до 5% это число равно числу фрагментов с произвольно меняющейся длинной, полученных от генератора («0», «1»). Естественно, он прав.

Резюмируем сказанное. Минимальная длина фрагмента зависит от количества элементарных событий пос-ти N, [1-3]. И искомые в фрагментах единицы «1» можно заменить на поиск нулей «0». Если первое элементарное событие фрагмента не совпало с выбранным шаблоном (не угадано), то вклад фрагмента в сумму нуль. Иначе прибавляем к сумме число единиц в фрагменте (до первого выпавшего «0»).

Доказательство теоремы «О равенстве суммы первых угаданных полярных событий числу серий».

Замечаем, что выпадающие подряд в начале фрагмента одинаковые элементарные события (до первого выпадения события другой величины) являются полярными составными событиями [1-3]. К ним применимы все теоремы о составных событиях [2, 5]. Заметим, что численности выпадающих «0» и «1» равны с точностью до случайной флуктуации. Выберем в качестве значения поискового шаблона «0». С выбранным «0» будем сравнивать первое элементарное событие каждого фрагмента случайной длины $L > L_{min}$. Просуммируем все первые «0» события (так как они равны выбранному шаблону), до первой выпавшей «1», обозначив сумму как: X_{el}.

Из теоремы «О амплитудно-частотной характеристике идеальной бинарной случайной пос-ти» [2,5] известно, что численность составных событий уменьшается в два раза с ростом длины на единицу. Множество составных событий $S0$, образованных из «0» в началах фрагментов, в сумме содержит X_{el} нулей. Из теоремы следует, что средняя длина составного события равна двум элементарным событиям. Следовательно, число фрагментов начавшихся с «0» равно $\frac{1}{2}X_{el}$. Обозначим их: $S0 = \frac{1}{2}X_{el}$.

Но существует, с точностью до случайной флуктуации, столько же фрагментов $\frac{1}{2}X_{el}$, которые начинаются с составных событий образованных из «1». Обозначим эти составные события: $S1 = \frac{1}{2}X_{el}$.

Поскольку в начале каждого фрагмента находится составное событие: либо $S0$, либо $S1$, то, следовательно, полное число фрагментов созданных генератором: $S0 + S1 = \frac{1}{2}X_{el} + \frac{1}{2}X_{el} = X_{el}$.

«Теорема доказана».

Следует отметить, что тестовая серия – искомое составное событие не обязательно должно начинаться с первого элементарного события каждого фрагмента. Тестовая серия может набираться в любых местах фрагментов, если она оканчивается до последнего события фрагмента. Набранная сумма Sum элементарных событий, содержащихся в таких сериях, вследствие теоремы [5], так же однозначно связана с числом найденных составных событий, которое будет равно: Sum /2. А следовательно, если поисковым шаблоном была «1», то учитывая половину составных событий (Sum /2) отброшенных по причине того, что они образованы из «0», общее число фрагментов (отброшенных и учтённых) равно набранной сумме единиц (Sum).

Заметим, что составные события, как из нулей, так и из единиц, в начале рассмотренных выше фрагментов, являются областями непрерывных угадываний (предсказаний выпадений). Первое не совпадение между предсказанным и выпавшим событием приводит к оформлению, определению длины составного события. Отсюда - сумма непрерывно угаданных элементарных событий равна (с точностью до случайной флуктуации) числу содержащих их серий (фрагментов).

Представленное доказательство верно для фрагментов, длина которых в подавляющих количествах превосходит длины начальных составных событий. Поясним это на формулах. Так как зависимость числа составных событий ${}^{n}S$, длины n, от числа бросков монеты N: ${}^{n}S = \frac{N}{2^{n+1}}$, то для каждого числа бросков монеты N существует такая максимальная длина составного события n, математическое ожидание выпадения которого ${}^{n}S$ меньше единицы, но больше, чем у более длинных составных событий k: ${}^{k}S(N) < {}^{n}S(N) < 1$. Связанную пару значений: n и N, найдём из ${}^{n}S$ неравенства: $\frac{N}{2^{n+1}} < 1$. Отсюда: $2^{n+1} < N$, и: $n < \log_2(N) - 1$. Так для $N = 2 \cdot 10^7$ наименьшая длина составного события, соответствующая неравенству, равна n = 24. Действительно: $\frac{2 \cdot 10^7}{2^{24+1}} < 1$. Значит, если описанный выше фокус по угадыванию числа фрагментов проводится на фрагментах длиной в 24 события, то надо остановить генератор на $N = 2^{24+1}$=33554432 событии, для уменьшения числа выпадений событий ${}^{24}S$ и более длинных.

Если же есть достаточно много составных событий, длины которых превышают длину фрагментов (так как продолжение за пределами фрагментов не возможно, то это проявляется в угадывании всех элементарных событий фрагментов), то сумма элементарных событий в непрерывно угаданных составных событиях определённой полярности (например «1») меньше числа фрагментов.

Рассмотрим пример, когда длины выпадающих составных событий массово превышают длины фрагментов. В таблице 7 даны результаты поискового эксперимента. В качестве предсказательного шаблона была выбрана «1». Бинарные события последовательно создавал генератор случайных событий. Случайные числа, по мере их создания, формировались во фрагменты указанной длины L.

Допустим, программа предсказала, что в первом фрагменте, первое созданное генератором случайное значение равно «1». Если первым значением выпал «0», то этот фрагмент увеличивал показание счётчика фрагментов с не угаданным первым значением, таблица 7, строка «0».

Если первое полученное число во фрагменте L совпадало с поисковым шаблоном «1» (равно единице), то считаются все «1» за ней, пока не выпадет «0». Затем увеличивается на единицу значение счётчика фрагментов для найденного числа «1». Затем, счётчик «1» устанавливают в ноль. Из генератора получают недостающие до L события. Осуществляется переход к следующему фрагменту.

В таблице 7 представлен результат работы вышеописанной поисковой программы, которая запускалась четыре раза для обработки фрагментов длин: L=3; L=4; L=5; L=6. При каждом включении программы создавалось $N = 2 \cdot 10^7$ случайных бинарных событий, которые в каждом из экспериментов, делились на Z = N / L фрагментов. Для каждого включения предсказательный шаблон равен «1».

Таблица 7. Распределение головных S1 – событий в коротких фрагментах

Длина головных ${}^{n}S$ событий	Длины фрагментов L			
	L = 3	L = 4	L = 5	L = 6
0	3334250	2500739	2000990	1667908
1	1666785	1250563	999748	833552
2	**833703**	624209	500223	416496
3	**831929**	**313009**	250139	208306
4	нет	**311480**	**124463**	103507
5	нет	нет	**124437**	**51713**
6	нет	нет	нет	**51852**
Сумма ${}^{n}S$ событий	3332417	2499261	1999010	1665426
Сумма: $n \cdot {}^{n}S$; % от Z	5829978 87,45	4683928 93,68	3870648 96,77	3275167 98,25
Число фрагментов Z	6666667	5000000	4000000	3333334
$N = 2 \cdot 10^7$; Btn592				

Из таблицы 7 видно, что численности угаданных событий зависят от их длины n, длины фрагмента L, числа бросков монеты N, ф.5.1:

$$ {}_{edge}^{n}SX_L = \begin{cases} \dfrac{Z}{2^{n+1}} = \dfrac{1}{2^{n+1}} \cdot \dfrac{N}{L}; \text{ где } 0 \le n < L \\ \\ \dfrac{Z}{2^{Z-1}} \end{cases} \qquad (\text{ф.5.1}) $$

На концах фрагментов длин L образуются уплотнения из угаданных составных событий длин $n = L$. В таблице 7 строчки, между которыми пропадает отношение равное двум, выделены жирным шрифтом. Например, при длине фрагмента L = 3, числа составных событий ${}^{n=2}S(833703)$ и ${}^{n=3}S(831929)$ равны друг другу, с точностью до случайной флуктуации. Что, на первый взгляд, выражается в нарушении закона двойного уменьшения числа событий с ростом длины на единицу, но на самом деле является его следствием: $\sum_{n=L}^{\infty}({}^{n}S) = 2 \cdot {}^{L}S$. То есть, сумма выпадающих обрезанных (не поместившихся в фрагментах) составных событий ${}^{n}S$, начиная с длины $L + 1$ и всех последующих длин, равна численности выпавших составных событий ${}^{n=L}S$ длины L. Составные события ${}^{n=L}S$ невозможно отличить от ${}^{L<n}S$, их общая численность: ${}^{n=L}S + {}^{L<n}S = 2 \cdot {}^{L}S$. Этот эффект объясняет краевые уплотнения описанные в начале статьи.

Интересно заметить, что при движении наблюдателя от начала фрагмента к его концу (последовательная генерация случайных бинарных событий) уплотнений нет в началах фрагментов, а есть только на концах фрагментов. Но последовательно записав все бинарные события в файл, и начав движение по фрагментам от конца файла к его началу, мы обнаружим, что в бывших концах фрагментов (а теперь, при обратном движении, их началах) уже нет уплотнений. А в том, что стало теперь концом фрагмента (а раньше было его началом), появились уплотнения. То есть, уплотнения являются наведённым эффектом, зависящим от направления движения внутри фрагментов, в случайной бинарной пос-ти.

То есть, при статической нарезке фрагментов (начало статьи) уплотнения наблюдаются на обоих концах фрагментов. А при «включении» пошагового движения вдоль фрагмента, головные уплотнения во фрагментах превращаются в величину, которая уменьшается в два раза с каждым шагом по фрагменту, таблица 7, формула 5.1, а численность составных событий на конце фрагментов требует особого расчёта. Наблюдатель, двигающийся вдоль бинарной последовательности, видит в последовательности, наведённые его движением эффекты. Причём эти эффекты имеют векторную (импульсную) природу, так как зависят от направления движения (исчезают и появляются от наличия движения наблюдателя вдоль последовательности).

Выводы.

Края фрагментов любой длины обладают более плотным содержанием составных событий ${}^{n}S$, а в центральных областях плотность составных событий ${}^{n}S$ меньше, чем в краевых областях.

При организации набора серий длины L из составных событий центральных областей и составных событий краевых областей, при равных условиях, число серий созданных из центральных составных событий будет ощутимо меньше, чем число серий созданных из краевых составных событий. Это объясняется эффектом экранирования, более известным как: парадокс Пенни, игра Пенни.

Показано, что формула, ранее применяемая для расчёта числа инверсий i внутри фрагментов длиной L, так же служит для получения комбинаторного распределения случайных угадываний элементарных событий внутри фрагментов длины $(\mathrm{L}-1)$.

Доказана теорема «О равенстве суммы первых угаданных полярных событий числу серий», выявила пропорциональную связь между некоторой частью случайных бинарных событий, содержащейся внутри серий произвольно меняющейся длины, и числом самих этих серий. Тестовая серия берётся из любых мест фрагментов, если она оканчивается до последнего события фрагмента. Тестовые серии, превышающие длину фрагмента, в который они проецируются, не подходят под действие доказанной теоремы.

На экспериментальных данных показано, что процедура сбора информации, заключающаяся в последовательном просмотре данных, приводит к появлению искажений. Эти искажения зависят от направления просмотра (перебора событий) в бинарной последовательности и не позволяют воссоздать истинную структуру последовательности. Что является аналогом принципа неопределённости Гейзенберга для статистической обработки случайных бинарных последовательностей.

Литература

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.* «О закономерностях структуры бинарной последовательности», «Журнал научных публикаций аспирантов и докторантов». № 5, 2014.
4. *Филатов О. В., Филатов И. О.* «О закономерностях структуры бинарной последовательности (продолжение)». «Журнал научных публикаций аспирантов и докторантов». № 6, 2014.
5. *Филатов О. В.* «Теорема «О амплитудно-частотной характеристике идеальной бинарной случайной последовательности». «Проблемы современной науки и образования». № 1 (31), 2015 г.
6. *Филатов О. В.* «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. «Проблемы современной науки и образования». № 17 (59), 2016 г.
7. *Филатов О. В.* «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». «Проблемы современной науки и образования». № 22 (64), 2016 г.
8. *Филатов О. В.* «Описание схем управления вероятностью выпадения независимых составных событий». «Проблемы современной науки и образования». № 2 (44), 2016 г.
9. *Филатов О. В.* Managed probability of Penny series against classical probability series of equal length. Not a typical conversion Mises. / Управляемая вероятность выпадения серий Пенни против классической вероятности выпадения серий равной длины. Не типичное преобразование Мизеса». «Проблемы современной науки и образования». № 29 (71), 2016 г.