Наш семинар: математические сюжеты

Среднее число случайных слагаемых в растущей сумме, достигшей заданного значения

И. Р. Высоцкий

Сумма независимых одинаковых случайных величин, имеющих дискретное или непрерывное равномерное распределение на каком-то промежутке, распределена по закону, близкому к нормальному. Это факт хорошо известен. Если дано число слагаемых, то указать интервал, в который попала такая сумма с любой наперёд заданной вероятностью, несложно. Ошибка при этом обусловлена заменой истинного распределения нормальным, и, как правило, ничтожно мала даже при не очень большом числе слагаемых.

Интересно поставить обратную задачу: каково должно быть число случайных слагаемых определённого вида, чтобы их сумма *впервые* достигла некоторого наперёд заданного значения.

Мы начнём издалека. Сначала будем бросать обычную игральную кость и складывать выпадающие очки. Нас будет интересовать вероятность того, что в какой-то момент сумма выпавших очков станет равна некоторому числу n, а также математическое ожидание числа сделанных бросков к моменту, когда число n будет достигнуто.

Затем перейдём к непрерывному случаю, складывая независимые случайные величины, равномерно распределённые на интервале $(0; 1)$.

Основная задача — найти математическое ожидание числа слагаемых в такой сумме в момент, когда она стала больше или равна наперёд заданного числа x. А именно, мы покажем, что это математическое ожидание равно

$$m(x)=\sum_{j=0}^{[x]}\frac{(-1)^j(x-j)^je^{x-j}}{j!}$$

и что на бесконечности функция $m(x)$ асимптотически приближается к линейной функции $y=2x+2/3$.

§ 1. Стоп-числа в игре с правильной костью

Все в детстве любили (многие любят и во взрослом состоянии) незамысловатые игры: на столе картонное поле с извилистой тропинкой. Игроки по очереди бросают кость. Сколько очков выпало, на столько шагов игрок продвигает свою фишку, с которой по дороге случаются разные приятные и неприятные события. Фишка может отъехать назад, попав на несчастливое поле, или наоборот — проскочить несколько полей, заработать для хозяина бросок вне очереди и т. п.

Предположим, что мы играем в такую игру, но без счастливых или несчастливых полей. Самый простой и естественный вопрос: с какой вероятностью фишка в какой-то момент остановится на поле с номером n? Все ли поля равновероятны? Если кость симметричная, то задача легко формализуется.

Задача 1. Какова вероятность, что при последовательных бросаниях правильной игральной кости сумма очков в какой-то момент станет в точности равна n?

Решение. Назовём *стоп-числом* число, которое в какой-то момент стало суммой. Интуиция подсказывает, что поскольку каждый бросок увеличивает сумму в среднем на 3,5, то стоп-числами окажутся в среднем 2 из каждых 7 чисел. Это рассуждение и даёт результат, весьма близкий к точному.

Дадим точное решение задачи, приводящее к рекуррентной формуле. Рассмотрим по очереди возможности, возникающие после первого броска. Если первый бросок дал 1, то n будет стоп-числом, только если $n-1$ тоже стоп-число. Так же обстоит дело со всеми исходами первого броска: если первый бросок дал k очков ($k=1,\ldots,6$), то n будет стоп-числом, только если $n-k$ также стоп-число, образовавшееся при последующих бросках. Обозначим через p_n вероятность того, что n — стоп-число. Тогда

$$p_n=\frac{1}{6}\cdot p_{n-1}+\frac{1}{6}\cdot p_{n-2}+\ldots+\frac{1}{6}\cdot p_{n-6}=\frac{1}{6}\sum_{k=1}^{6}p_{n-k}. \qquad (1)$$

Из свойств линейной рекурсии известно, что последовательность p_n представима в виде суммы геометрических прогрессий со знаменателями, модуль которых не превосходит 1, при этом хотя бы один из них в точности равен 1 (поскольку сумма коэффициентов в правой части равна 1). Поэтому рекурсия имеет пределом некоторое число, которое, как мы понимаем, должно равняться $2/7$. Получить явное выражение для p_n непросто, даже если удастся решить соответствующее характеристическое уравнение 6-й степени.

Чтобы понять, как ведут себя вероятности p_n, зададим начальные условия: $p_{-5} = \ldots = p_{-1} = 0$, $p_0 = 1$. Это естественное предположение: отрицательной сумма быть не может, а нулевой она является с вероятностью $p_0 = 1$ до того, как кость брошена первый раз. Тогда

$$p_1 = \frac{1}{6} = \frac{7^0}{6}, \quad p_2 = \frac{1}{6}\left(1 + \frac{1}{6}\right) = \frac{1}{6} \cdot \frac{7}{6} = \frac{7^1}{6^2}$$

и так далее: при $n \leqslant 6$

$$p_n = \frac{1}{6}\left(\frac{7^{n-2}}{6^{n-1}} + \frac{7^{n-3}}{6^{n-2}} + \ldots + \frac{7^0}{6} + 1\right) = \frac{1}{6} \cdot \left(\frac{1}{6} \cdot \frac{(7/6)^{n-1} - 1}{7/6 - 1} + 1\right) = \frac{7^{n-1}}{6^n}.$$

Такая закономерность сохраняется, пока последняя ненулевая вероятность p_{n-6} в сумме (1) равна единице. Таким образом, поначалу вероятности p_n образуют растущую геометрическую прогрессию, которая достигает наибольшего значения при $n = 6$. При $n > 6$ эта закономерность нарушается, поскольку при $n = 7$ в сумме (1) последнее слагаемое $p_0 = 1$ уступает место слагаемому $p_1 = 1/6$. Затем снова рост до $n = 11$. Далее происходят затухающие колебания вероятностей с периодом 5–6 шагов. Локальные максимумы вероятностей выделены в таблице 1 жирным шрифтом, при этом разумно считать, что уже при $n \geqslant 20$ вероятность p_n неотличима от $2/7$.

Если вы составляете детскую игру с кубиком и фишками (см. с. 104) и хотите сделать её повеселее, устройте какую-нибудь первую неприятность (или наоборот) как раз на поле 6.

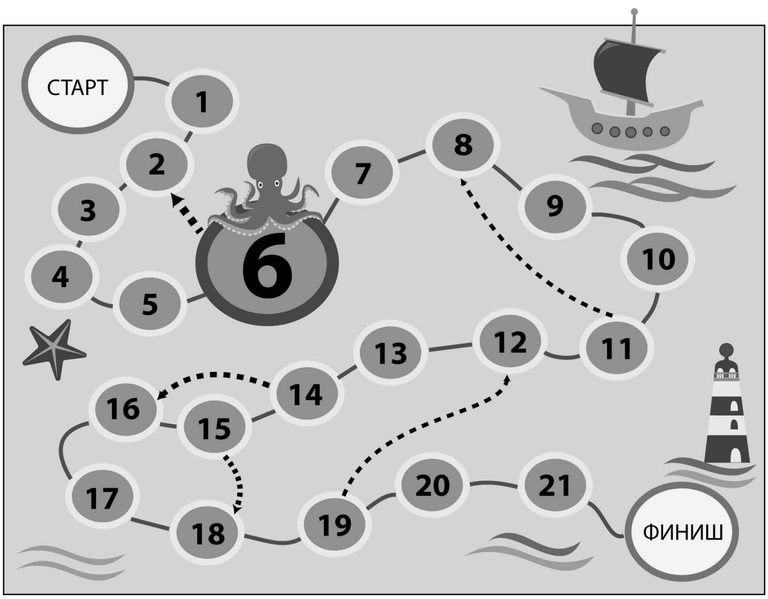
Интересно посмотреть на «трудозатраты», нужные для того, чтобы достичь поля n. В случае с обычным кубиком для этого потребуется от $n/6$ бросков (если невероятно повезёт) до n бросков (при столь же невероятном невезении).

Задача 2. Найти математическое ожидание числа бросков, сделанных к моменту, когда сумма очков впервые оказалась больше или равна n.

Решение. Пусть случайная величина S_n — число бросков, которые пришлось сделать, чтобы сумма очков достигла n. Введём для чисел

Таблица 1

n	p_n	n	p_n	n	p_n	n	p_n
−5	0,000000000000	15	0,286113932137	35	0,285710193751	55	0,285714292240
−4	0,000000000000	**16**	**0,287071429920**	36	0,285711733841	56	0,285714283629
−3	0,000000000000	17	0,286701924733	37	0,285715051280	57	0,285714280890
−2	0,000000000000	18	0,285586725149	**38**	**0,285716315054**	58	0,285714284162
−1	0,000000000000	19	0,284712810463	39	0,285714800621	59	0,285714287275
0	**1,000000000000**	20	0,285621080152	40	0,285713653567	**60**	**0,285714287468**
1	0,166666666667	**21**	**0,285967983759**	41	0,285713624686	61	0,285714285944
2	0,194444444444	22	0,285943659029	42	0,285714196508	62	0,285714284895
3	0,226851851852	23	0,285755697214	**43**	**0,285714606953**	63	0,285714285106
4	0,264660493827	24	0,285597992628	44	0,285714532898	64	0,285714285808
5	0,308770576132	25	0,285599870541	45	0,285714235872	**65**	**0,285714286083**
6	**0,360232338820**	26	0,285747713887	46	0,285714141747	66	0,285714285884
7	0,253604395290	**27**	**0,285768819510**	47	0,285714223111	67	0,285714285620
8	0,268094016728	28	0,285735625468	48	0,285714322848	68	0,285714285566
9	0,280368945441	29	0,285700953208	**49**	**0,285714343905**	69	0,285714285678
10	0,289288461040	30	0,285691829207	50	0,285714300064	**70**	**0,285714285773**
11	**0,293393122242**	31	0,285707468637	51	0,285714261258	71	0,285714285767
12	0,290830213260	**32**	**0,285725401653**	52	0,285714265489	72	0,285714285715
13	0,279263192334	33	0,285721682947	53	0,285714286112	73	0,285714285686
14	0,283539658507	34	0,285713826853	**54**	**0,285714296613**	74	0,285714285697



$k = 1, \ldots, 6$ шесть бинарных случайных величин I_k — индикаторов событий «первый бросок дал ровно k очков»:

$$I_k = \begin{cases} 1, & \text{если первый бросок дал ровно } k \text{ очков}, \\ 0, & \text{если первый бросок дал не } k \text{ очков}. \end{cases}$$

Тогда

$$S_n = I_1 S_{n-1} + I_2 S_{n-2} + \ldots + I_6 S_{n-6} + 1,$$

где случайная величина S_{n-k} означает число бросков, не считая первого, потребовавшихся для того, чтобы достичь суммы очков $n-k$. Единица в конце нужна, чтобы учесть первый бросок. Величины I_k и S_{n-k} независимы, поскольку индикатор I_k относится только к первому броску. Введём для ожиданий $\mathrm{E}S_n$ краткое обозначение m_n, перейдём в полученном равенстве к математическому ожиданию и получим линейное неоднородное рекуррентное уравнение

$$m_n = 1 + \frac{1}{6}\sum_{k=1}^{6} m_{n-k}, \tag{2}$$

отличающееся от (1) только ненулевым свободным членом. Неудивительно, что частным решением уравнения (2) является последовательность $m_n = \frac{2}{7}n$ $(n \in \mathbb{N})$ — это легко проверить непосредственно. Общее решение (2) имеет вид

$$m_n = \frac{2}{7}n + a_n,$$

где a_n — некоторая сходящаяся последовательность, которая даёт решение соответствующего однородного уравнения; она зависит от начальных условий.

Взяв для рекурсии (2) естественные начальные условия $m_n = 0$ при $n \leqslant 0$ и $m_1 = 1$ (сумма достигнет единицы обязательно при первом же броске), можно провести расчёт. Он показывает, что

$$m_n \approx \frac{2}{7}n + \frac{10}{21}.$$

Если кость брошена ровно 10 раз, то ожидание суммы выпавших очков равно 35. Обратим ситуацию (см. таблицу 2) и видим, что для достижения суммы 35, требуется в среднем не 10 бросков, а 10 «с хвостиком»: больше на примерно 10/21. Природу хвостика понять нетрудно: 35 окажется стоп-числом с вероятностью около 2/7. А с вероятностью 5/7 сумма очков перескочит число 35, не остановившись на нём. Вот на это перескакивание и расходуются лишние примерно $\frac{10}{21}$ броска.

Таблица 2

n	m_n	Отличие m_n от $\frac{2}{7}n+\frac{10}{21}$
32	9,6190338922	−0,00001373
33	9,9047592939	−0,00000261
34	10,1904809768	0,00000479
35	**10,4761948037**	**0,00000433**
36	10,7619049974	0,00000024
37	11,0476167313	−0,00000232
38	11,3333317826	−0,00000155

§ 2. Непрерывный случай. Точное решение задачи

Если слагаемые — не очки на кубике, а непрерывные случайные величины, то ситуация должна быть примерно такой же, как в случае с кубиком: математическое ожидание числа слагаемых должно приблизительно линейно зависеть от наперёд заданного стоп-числа x. При этом тоже должен быть какой-то «хвостик», но, вероятно, уже не 10/21, а какой-то другой. Можно поэкспериментировать на компьютере, бросая разные n-гранные кости и увеличивая n, но мы попробуем решить задачу в общем виде. Рассмотрим теперь последовательность одинаковых и независимых случайных величин ξ_i $(i=1,2,\ldots)$, равномерно распределённых на интервале $(0;1)$.

Задача 3. Найти математическое ожидание числа независимых равномерно распределённых на интервале $(0;1)$ величин ξ_i к моменту, когда их сумма впервые достигла числа x.

Обозначим через $S(x)$ или просто S число слагаемых в сумме

$$\tau=\tau_S=\xi_1+\xi_2+\ldots+\xi_S,$$

которая впервые достигла числа x. Очевидно, $S(x)>x$. Величина S является случайной функцией аргумента x, поскольку каждое x порождает распределение $p_s=\mathrm{P}(S=s)=\mathrm{P}(\tau_{s-1}<x\leqslant\tau_S)$. Природа вероятностей p_s известна — они выражаются через плотности f_s распределений[1] сумм τ_s:

$$p_s(x)=\int\limits_{x-1}^{x} f_{s-1}(t)(t+1-x)\,dt=\int\limits_{0}^{1} t f_{s-1}(t+x-1)\,dt \qquad (3)$$

[1] Распределения Ирвинга — Холла. Функции f_s являются многочленами.

(для общности положим $f_0(x)$ равной дельта-функции $\delta(x)$). Используя соотношение (3), видимо, можно получить вероятности p_s в явном виде. Но мы применим для поиска $\mathrm{E}S(x)$ иной метод.

Для краткости записи введём функцию $m(x) = \mathrm{E}S(x)$, которая непрерывна при $x > 0$. Это интуитивно ясно — при малом изменении x дискретная целочисленная величина S с вероятностью, стремящейся к единице, не изменяется. Строгое доказательство непрерывности функции $m(x)$ при $x > 0$ можно провести непосредственно — это несложное упражнение в математическом анализе. Другой путь: непрерывность следует из того, что функции f_s и, следовательно, вероятности p_s непрерывны.

Вопрос о математическом ожидании случайной величины S приводит к непрерывному аналогу уравнения (2):

$$\mathrm{E}S = m(x) = 1 + \int\limits_0^1 m(x-t)\,dt. \tag{4}$$

Получился частный (для равномерно распределённых слагаемых) случай уравнения, которое называют *уравнением восстановления*[2)]. Причину такого названия обсудим ниже. Перепишем уравнение иначе:

$$m(x) = 1 + \int\limits_{x-1}^{x} m(t)\,dt. \tag{5}$$

При $x < 0$ следует считать, что $m(x) = 0$. При этом можно доопределить функцию в нуле: из уравнения (5) получается

$$m(0) = 1 + \int\limits_{-1}^{0} m(t)\,dt = 1.$$

Это согласуется с соображением, что требуется одно слагаемое, чтобы сумма стала положительной, т. е. $\lim_{x\to 0+} m(x) = 1$. Приняв эти соглашения, получим начальные условия:

$$\begin{cases} m(x) = 0, & \text{если } x < 0, \\ m(0) = 1. \end{cases} \tag{6}$$

Таким образом, мы ищем непрерывное при $x > 0$ решение задачи (5)–(6).

Если функция $y = m(x)$ непрерывна в точках x и $x-1$, то из (5) следует:

$$m'(x) = m(x) - m(x-1), \tag{7}$$

[2)] Общий вид и вывод уравнения восстановления можно найти, например, в статье [5]. См. также [4, 6].

т. е. функция $m(x)$ дифференцируема в точке x. Значит, функция $m(x)$ дифференцируема во всех точках, кроме точки 0 (где она имеет разрыв) и точки 1 (поскольку имеет разрыв в точке 0).

Сделаем небольшое техническое упрощение — введём вспомогательную функцию $g(x)=e^{-x}m(x)$. Тогда уравнение (7) и условия (6) преобразуются в задачу

$$g'(x)=-\frac{1}{e}g(x-1),\qquad \begin{cases} g(x)=0 & \text{при } x<0,\\ g(0)=1. \end{cases} \tag{8}$$

Уравнение (8), впрочем, как и уравнение (7), — линейное дифференциальное уравнение с *запаздывающим аргументом*[3)].

Обычно такие уравнения решаются по шагам на последовательных промежутках, длина которых равна запаздыванию. Сначала найдём решение на промежутке $[0;1)$, потом на промежутке $[1;2)$ и так далее. На каждом промежутке получается обычная задача Коши, поскольку решение на предыдущем промежутке уже известно. На $[0;1)$ получаем задачу

$$g'(x)=0,\quad g(0)=1,$$

откуда $g(x)=1$.

Должно выполняться равенство

$$g(1)=\lim_{x\to 1-}g(x)=1.$$

Поэтому на промежутке $[1;2)$ получается задача

$$g'(x)=-\frac{1}{e},\quad g(1)=1,$$

имеющая единственное решение $g(x)=1-(x-1)e^{-1}$ при $1\leqslant x<2$.

Продвигаясь таким образом далее, получаем:

$$g(x)=1-\frac{x-1}{e}+\frac{(x-2)^2}{2e^2}-\frac{(x-3)^3}{6e^3}+\ldots+\frac{(x-j)^j}{j!\,e^j}$$

при $j\leqslant x<j+1$. Тогда

$$m(x)=\sum_{j=0}^{[x]}\frac{(-1)^j(x-j)^je^{x-j}}{j!}. \tag{9}$$

Интересно посмотреть на результат при всех натуральных $x=n\in\mathbb{N}$. Формула (9) принимает вид

$$m(n)=\sum_{j=0}^{n}\frac{(-1)^j(n-j)^j}{j!}e^{n-j}=\sum_{k=0}^{n}\frac{(-1)^{n-k}k^{n-k}}{(n-k)!}e^k. \tag{10}$$

3) См., например, неоднократно переизданную монографию [3].

Получается некоторый многочлен степени n, вычисленный в точке e. В частности, чтобы достичь единицы, требуется $m(1) = e$ слагаемых; чтобы сумма достигла двух, потребуется в среднем

$$m(2) = e^2 - (2-1)e^{2-1} = e^2 - e$$

слагаемых и так далее:

$$m(3) = e^3 - 2e^2 + \frac{1}{2}e, \quad m(4) = e^4 - 3e^2 + 2e - \frac{1}{6}e, \quad \ldots$$

Результат удобно проиллюстрировать таблицей, в которую столбиками выпишем члены рядов для последовательных отрицательных степеней e: e^{-1}, e^{-2}, e^{-3}, … При этом каждый следующий столбик сместим на одну строку вниз по отношению к предыдущему. Во всех столбцах каждое число, кроме верхней единицы, получается умножением сверху стоящего числа на множитель $-k/(n-k)$.

x \ e^{-k}	e^{-1}	e^{-2}	e^{-3}	e^{-4}	e^{-5}	e^{-6}	e^{-7}	…	e^{-k}	…
1	1									
2	-1	1								
3	$1/2$	-2	1							
4	$-1/6$	2	-3	1						
5	$1/24$	$-4/3$	$9/2$	-4	1					
6	$-1/120$	$2/3$	$-9/2$	8	-5	1				
7	$1/720$	$-4/15$	$27/8$	$-32/3$	$25/2$	-6	1			
…	…	…	…	…	…	…	…	…		
n	$\frac{(-1)^{n-1}}{(n-1)!}$	$\frac{(-2)^{n-2}}{(n-2)!}$	$\frac{(-3)^{n-3}}{(n-3)!}$	…	…	…	…	…	$\frac{(-k)^{n-k}}{(n-k)!}$	…
…	…	…	…	…	…	…	…	…	…	…

Сумма чисел в k-м столбце равна e^{-k}. Если умножить числа k-го столбца на e^k, сумма в каждом столбце станет равна 1, а сумма чисел в n-й строке окажется равна $\mathrm{E}S(n)$. Приведём результаты с точностью до четырёх знаков в таблице 3.

Суммы быстро приближаются к $2n + 2/3$. Возникает предположение, что

$$m(x) = 2x + \frac{2}{3} + o(1). \tag{11}$$

Таблица 3

Степени e^k	2,7183	7,3891	20,0855	54,5982	148,4132	403,429	1096,63	2980,96	8103,08	22026,5	ES: сумма чисел строки, умноженных на соответствующие степени e^k
k / n	1	2	3	4	5	6	7	8	9	10	
1	1										**2,7183**
2	−1	1									**4,6708**
3	0,5	−2	1								**6,6666**
4	−0,1667	2	−3	1							**8,6667**
5	0,0417	−1,3333	4,5	−4	1						**10,6667**
6	−0,0083	0,6667	−4,5	8	−5	1					**12,6667**
7	0,0014	−0,2667	3,375	−10,6667	12,5	−6	1				**14,6667**
8	−0,0002	0,0889	−2,025	10,6667	−20,8333	18	−7	1			**16,6667**
9	0,0002	−0,0254	1,0125	−8,5333	26,0417	−36	24,5	−8	1		**18,6667**
10	−0,0000	0,0063	−0,4339	5,6889	−26,0417	54	−57,1667	32	−9	1	**20,6667**
	…………………………										

§ 3. Асимптотическая оценка

Докажем равенство (11). Из общей теории линейных уравнений с запаздывающим аргументом следует, что непрерывное при $x > 0$ решение задачи (6)–(7) приближается к некоторой линейной функции. Если подставить частное решение $m = ax + b$ в уравнение (5), то несложно найти, что $a = 2$, т. е. решение (9) с ростом x асимптотически приближается к линейной функции $y = 2x + b$. Обычно формулируют более слабое утверждение, которое в наших терминах имеет вид

$$\frac{m(x)}{x} \xrightarrow[x\to\infty]{} \frac{1}{\mathrm{E}\xi_1} = 2,$$

и получается из простой и важной леммы Вальда (Wald’s equation), которую мы докажем позже.

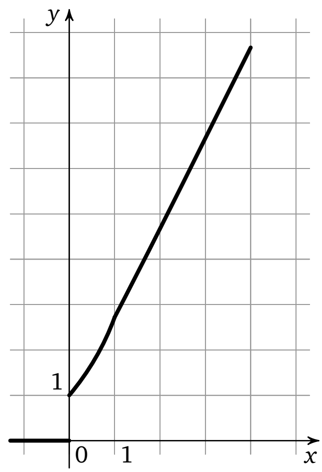


Рис. 1

Этот результат согласуется с естественным интуитивным предположением, что увеличение суммы на 1 требует в среднем двух дополнительных слагаемых, поскольку каждое в среднем равно 0,5.

На рис. 1 показан график функции $y = m(x)$. Уже при $2 < x \leqslant 3$ график функции $m(x)$ невозможно на глаз отличить от прямой.

Пока что решённая задача относится к тем, где точное решение, доставив эстетическое удовольствие, не снимает главный вопрос: а всё же, сколько это будет? Расчёт по формуле (9) затруднителен, а асимптотика

$$m(x) \sim 2x + \frac{2}{3}$$

требует доказательства, хотя ясно видна в таблице 3.

Сначала докажем две леммы, причём первая потребуется позже, когда мы будем обсуждать связь нашей задачи с числами Эйлера.

Лемма 1 о случайном остатке. *Пусть η — произвольная действительная случайная величина, а $\xi \sim U(0;1)$ — стандартная равномерно распределённая случайная величина, и эти величины независимы. Тогда дробная часть суммы $\{\eta + \xi\}$ равномерно распределена на промежутке $[0;1)$, при этом величины $\{\eta + \xi\}$ и η независимы.*

Доказательство. Нужно показать, что на промежутке $[0;1)$ функция распределения суммы $\eta + \xi$ тождественно равна x: $F_{\eta+\xi}(x) = x$ независимо от распределения величины η. Не ограничивая общности, можно считать, что $0 \leqslant \eta < 1$. Рассмотрим два случая: $0 \leqslant x \leqslant \eta$ и $\eta < x < 1$:

$$F_{\eta+\xi}(x) = \mathrm{P}(\eta + \xi \leqslant x) =$$
$$= \begin{cases} \mathrm{P}(1 \leqslant \eta + \xi < x + 1), & \text{если } 0 \leqslant x < \eta, \\ \mathrm{P}(\eta \leqslant \eta + \xi < x) + \mathrm{P}(1 \leqslant \eta + \xi < \eta + 1), & \text{если } \eta \leqslant x < 1 \end{cases} =$$
$$= \begin{cases} \mathrm{P}(1 - \eta \leqslant \xi < x + 1 - \eta), & \text{если } 0 \leqslant x < \eta, \\ \mathrm{P}(0 \leqslant \xi < x - \eta) + \mathrm{P}(1 - \eta \leqslant \xi < 1), & \text{если } \eta \leqslant x < 1 \end{cases} =$$
$$= \begin{cases} x, & \text{если } 0 \leqslant x < \eta, \\ x - \eta + \eta, & \text{если } \eta \leqslant x < 1 \end{cases} = x.$$

Отсюда следует утверждение леммы. □

Иллюстрируют эту лемму стрелочные часы: если кто-то в какой-то момент случайным образом покрутит минутную стрелку, то после этого часы будут показывать совершенно случайное время безо всякой связи с тем, что они показывали до того.

Если имеется дополнительная информация о слагаемых, то ситуация меняется. Например, если известно, что $\eta < 0$ и $\eta + \xi \geqslant 0$, то сумма $\eta + \xi$ может иметь уже вовсе не равномерное распределение.

Лемма 2 о случайном остатке. *Пусть случайные величины ξ и η независимы, причём величина ξ равномерно распределена на интервале* $(0; 1)$, *а величина η равномерно распределена на интервале* $(-1; 0)$. *Тогда*

$$\mathrm{E}(\xi + \eta \mid \xi + \eta \geqslant 0) = \frac{1}{3}.$$

Доказательство. На координатной плоскости $\xi O \eta$ условия

$$T = \{0 < \xi < 1,\ -1 < \eta < 0,\ \xi + \eta \geqslant 0\}$$

определяют треугольник (см. рис. 2), на котором равномерно распределён случайный вектор (ξ, η). Поэтому

$$\mathrm{E}(\xi + \eta \mid T) = \frac{\iint\limits_T (x + y)\,dx\,dy}{\iint\limits_T dx\,dy}.$$

Добавим третью координатную ось (см. рис. 3). Числитель дроби равен объёму пирамиды с основанием T и с вершиной в точке $(1; 0; 1)$, а знаменатель равен площади треугольника T. Поэтому

$$\mathrm{E}(\xi + \eta \mid T) = \frac{1/6}{1/2} = \frac{1}{3}.$$

Лемма 2 доказана. □

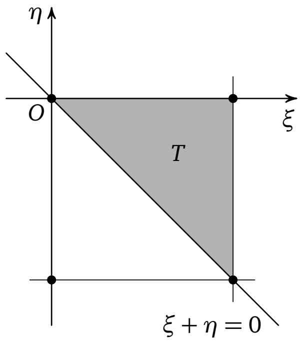


Рис. 2

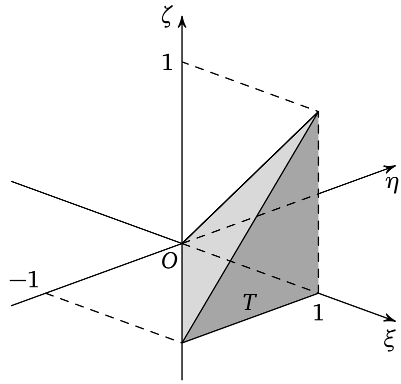


Рис. 3

Вернёмся к суммам равномерных слагаемых и математическому ожиданию их числа. Пусть $x > 0$. Рассмотрим последовательность независимых случайных величин $\xi_1, \xi_2, \xi_3, \ldots, \xi_k, \ldots$, имеющих стандартное равномерное распределение $U(0, 1)$. Будем составлять суммы

$$\tau_1 = \xi_1, \quad \tau_2 = \xi_1 + \xi_2, \quad \tau_s = \tau_{s-1} + \xi_s, \quad \ldots$$

до тех пор, пока очередная сумма τ не окажется больше или равна числа x. Как и прежде, число слагаемых в сумме τ обозначим S:

$$\tau = \tau_S = \sum_{k=1}^{S} \xi_k.$$

Сперва докажем ещё одну важную лемму, которую ещё называют леммой или равенством Вальда[4] [5].

Лемма 3. $\mathrm{E}\tau = \mathrm{E}\xi_1 \cdot \mathrm{E}S$ (*в нашем случае* $\mathrm{E}\tau = \frac{1}{2}\mathrm{E}S$).

Доказательство. Введём индикаторы I_s событий «$s-1$ слагаемого оказалось недостаточно для достижения суммы x» при каждом натуральном $s > 0$:

$$I_s = \begin{cases} 0, & \text{если } \tau_{s-1} \geqslant x, \\ 1, & \text{если } \tau_{s-1} < x. \end{cases}$$

Тогда $S = I_1 + I_2 + \ldots$ и $\tau = I_1\xi_1 + I_2\xi_2 + \ldots$

Ясно, что $I_1 = 1$ и что каждый следующий индикатор I_s зависит только от τ_{s-1}, но не зависит от ξ_s. Поэтому, переходя к математическим ожида-

[4] «Мы знаем, что это равенство важное, поскольку носит чьё-то имя», — пишет Петер Небрес в своей статье [5]. Присоединимся к его мнению.

ниям, получим:

$$\mathrm{E}\tau = \sum_{s=1}^{\infty} \mathrm{E}(I_s\xi_s) = \sum_{s=1}^{\infty} \mathrm{E}I_s \mathrm{E}\xi_s = \mathrm{E}\xi_1 \sum_{s=1}^{\infty} \mathrm{E}I_s = \frac{1}{2}\sum_{s=1}^{\infty} \mathrm{E}I_s = \tfrac{1}{2}\mathrm{E}\sum_{s=1}^{\infty} I_s = \tfrac{1}{2}\mathrm{E}S. \qquad \square$$

Лемма Вальда выявляет связь между $\mathrm{E}S$ и $\mathrm{E}\tau$, а лемма 2 будет нужна при доказательстве равенства $\mathrm{E}\tau = x + 1/3 + o(1)$. Собрав эти результаты воедино, мы получим $\mathrm{E}S = 2x + 2/3 + o(1)$.

Чтобы получить равенство $\mathrm{E}\tau = x + 1/3 + o(1)$, достаточно убедиться, что в сумме

$$\tau = \left(\sum_{k=1}^{S-1} \xi_k - x\right) + \xi_S + x$$

слагаемое, взятое в скобки, имеет асимптотически равномерное распределение на $(-1; 0)$:

$$\sum_{k=1}^{S-1} \xi_k - x \xrightarrow[x\to\infty]{} \eta,$$

где η — некоторая случайная величина, имеющая равномерное распределение $U(-1, 0)$. Это, в свою очередь, следует из того, что сумма

$$\tau_{s-1} = \sum_{k=1}^{s-1} \xi_k$$

имеет распределение, очень близкое к нормальному $N\left(\frac{s-1}{2}, \frac{s-1}{12}\right)$, которое в своей средней части, будучи сужено на любой единичный интервал, с ростом s приближается к равномерному на этом интервале (см. рис. 4).

На хвостах нормального распределения это не так, но поведение на хвостах не важно, поскольку при подходящем выборе границ средней части

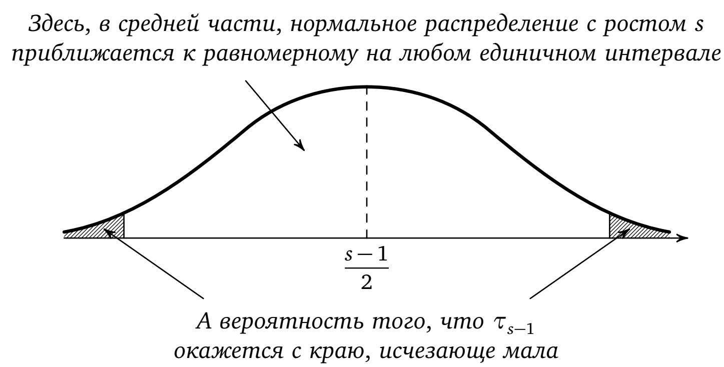


Рис. 4

вероятность того, что τ_{s-1} окажется на каком-то из хвостов, стремительно приближается к 0 при росте x, а значит, при росте s, поскольку $s > x$.

Проведём эти рассуждения теперь аккуратно и в правильном порядке.

Утверждение. *$\mathrm{E}\tau = 2x + 2/3 + o(1)$ при $x \to \infty$.*

Доказательство. Центрируя и нормируя случайную величину τ_{s-1}, получаем случайную величину

$$\zeta_{s-1} = \frac{\tau_{s-1} - (s-1)/2}{\sqrt{(s-1)/12}} = 2\sqrt{3}\frac{\tau_{s-1} - (s-1)/2}{\sqrt{s-1}},$$

значения которой при условии $x - 1 < \tau_{s-1} < x$ принадлежат интервалу

$$A = \left(2\sqrt{3}\,\frac{x-1-(s-1)/2}{\sqrt{s-1}};\ 2\sqrt{3}\,\frac{x-(s-1)/2}{\sqrt{s-1}}\right).$$

Длина интервала A равна $\dfrac{2\sqrt{3}}{\sqrt{s-1}}$. Увеличивая x, мы тем самым увеличиваем s, и последовательность величин ζ_{s-1} сходится к стандартной нормальной случайной величине, а её плотности $f_{\zeta_{s-1}}(t)$ равномерно сходятся к стандартной нормальной плотности

$$\varphi(t) = \frac{1}{\sqrt{2\pi}}\exp\left(-\frac{t^2}{2}\right).$$

Разобьём $\mathrm{E}\tau$ в сумму условных математических ожиданий при условиях

$$H = \{A \subset (-s^{1/4}; s^{1/4})\} \quad \text{и} \quad \overline{H} = \{A \not\subset (-s^{1/4}; s^{1/4})\}.$$

Получим:

$$\mathrm{E}\tau = \mathrm{E}(\tau|H)\mathrm{P}(H) + \mathrm{E}(\tau|\overline{H})\mathrm{P}(\overline{H}).$$

Поскольку $s \xrightarrow[x\to\infty]{} \infty$, получаем $\mathrm{P}(H) \xrightarrow[x\to\infty]{} 1$ и $\mathrm{P}(\overline{H}) \xrightarrow[x\to\infty]{} 0$. Поэтому

$$\mathrm{E}(\tau|H) - \mathrm{E}\tau \xrightarrow[x\to\infty]{} 0, \tag{12}$$

и вместо $\mathrm{E}\tau$ можно искать $\mathrm{E}(\tau|H)$, ограничиваясь гипотезой H. При этом условии рассмотрим относительный размах функции плотности $f_{\zeta_{s-1}}$ на интервале A:

$$\frac{\sup_A f_{\zeta_{s-1}} - \inf_A f_{\zeta_{s-1}}}{\inf_A f_{\zeta_{s-1}}} = \frac{\sup_A f_{\zeta_{s-1}}}{\inf_A f_{\zeta_{s-1}}} - 1.$$

При достаточно большом x можно заменить в этом выражении сложно устроенную плотность $f_{\zeta_{s-1}}$ стандартной нормальной плотностью φ (в силу равномерной сходимости f_{ζ_s} к φ), при этом расширив для удобства интервал A до его замыкания $[A]$:

$$\frac{\sup_A f_{\zeta_{s-1}}}{\inf_A f_{\zeta_{s-1}}}-1+o(1)=\frac{\max_{[A]}\varphi}{\min_{[A]}\varphi}-1=\frac{\exp(-t_1^2/2)}{\exp(-t_2^2/2)}-1=\exp\left(\frac{t_2^2-t_1^2}{2}\right)-1,$$

где t_1 и t_2 — точки, в которых φ принимает соответственно наибольшее и наименьшее значения на отрезке $[A]$. Очевидно,

$$|t_1|,|t_2|\leqslant s^{1/4}\quad\text{и}\quad|t_2-t_1|\leqslant\frac{2\sqrt{3}}{\sqrt{s-1}}.$$

Тогда

$$\exp\left(-\frac{2\sqrt{3}\,s^{1/4}}{\sqrt{s-1}}\right)-1\leqslant\frac{\max\varphi}{\min\varphi}-1=\frac{\exp(-t_1^2/2)}{\exp(-t_2^2/2)}-1\leqslant\exp\left(\frac{2\sqrt{3}\,s^{1/4}}{\sqrt{s-1}}\right)-1.$$

Левая и правая части этого неравенства стремятся к 0 при $s\to\infty$. Следовательно,

$$\frac{\max_{[A]}\varphi}{\min_{[A]}\varphi}-1\xrightarrow[x\to\infty]{}0$$

и поэтому

$$\frac{\sup_A f_{\zeta_{s-1}}-\inf_A f_{\zeta_{s-1}}}{\inf_A f_{\zeta_{s-1}}}\xrightarrow[x\to\infty]{}0.$$

Таким образом, какое бы значение ни приняла случайная величина $S>x$, с ростом x распределение случайной величины

$$\tau-\xi_S=\sum_{k=1}^{S-1}\xi_k$$

приближается к равномерному на интервале $(x-1;x)$, а случайная величина $\tau-\xi_S-x$ приближается к случайной величине η, равномерно распределённой на интервале $(-1;0)$. С помощью леммы 2 получаем:

$$\begin{aligned}\mathrm{E}(\tau|H,\tau\geqslant x)&=\mathrm{E}(\xi_S+(\tau-\xi_S-x)\mid\tau\geqslant x)+x+o(1)=\\&=\mathrm{E}(\xi_S+\eta\mid\xi_S+\eta\geqslant 0)+x+o(1)=x+\frac{1}{3}+o(1).\end{aligned}$$

Учитывая (12), видим, что

$$\mathrm{E}\tau=x+\frac{1}{3}+o(1).$$

Отсюда и из леммы 3 следует:

$$\mathrm{E}S=2x+\frac{2}{3}+o(1).$$

Чтобы сумма независимых стандартных равномерных слагаемых достигла числа x, в среднем потребуется приблизительно $2x + 2/3$ слагаемых. Добиться с помощью случайных слагаемых суммы, в точности равной x, практически невозможно. Почти наверняка сумма превзойдёт число x на величину, в среднем равную примерно $1/3$. Вот на эту «лишнюю» треть и расходуются в среднем лишних $2/3$ случайного слагаемого. □

Рассмотренная задача — частный случай одной из основных задач так называемой *теории восстановления* (Renewal Theory), о которой в российской математической литературе практически нет упоминаний. Известно переводное издание [2], адресованное преимущественно инженерам и, вероятно, поэтому вышедшее в издательстве «Советское радио».

В общем случае под процессом обновления или восстановления (Renewal Process) имеется в виду случайный процесс, состоящий из последовательных сумм независимых неотрицательных случайных величин, имеющих какой-то определённый смысл. Процесс продолжается до тех пор, пока сумма этих величин не достигнет наперёд заданного значения.

Классический пример — работа прибора, который должен функционировать до определённого момента (например, до окончания гарантийного срока или до выработки ресурса). Однако прибор может ломаться и до истечения гарантии, и в таких случаях подлежит ремонту (восстановлению). Случайные слагаемые здесь — периоды времени между поломками, сумма которых даёт полное время работы прибора.

Возникают естественные вопросы — каково среднее количество этих случайных слагаемых, т. е. поломок, случившихся к моменту окончания гарантийного срока, и во сколько в среднем это обойдётся. Более серьёзный вопрос — не следует ли в какой-то момент из соображений экономии заменить исправный прибор новым ещё до окончания ресурса?

Из предыдущего абзаца ясно, что породило математическую теорию восстановления. Исследователей в приложениях больше интересовали процессы восстановления, где каждый интервал между двумя поломками имеет показательное или близкое к показательному распределение, что породило представление о процессе восстановления как об обобщении пуассоновского процесса.

§ 4. Связь задачи о числе слагаемых с числами Эйлера

Подойдём к рассмотренной задаче с комбинаторной стороны. Правда, придётся пожертвовать произвольностью достигаемой суммы. Поставим вопрос только о целых суммах и временно забудем всё, сделанное прежде.

Начнём с простого случая $n=1$: сколько в среднем случайных слагаемых нужно, чтобы достичь единицы? Этот случай поможет понять смысл и логику общего построения. Возьмём случайное слагаемое ξ_1 и прибавим к нему ξ_2. Если сумма меньше единицы, добавим ξ_3, и так до тех пор, пока полученная сумма не достигнет $n=1$. Присмотримся к остаткам — дробные части последовательно получающихся сумм τ_s равны

$$Z_1 = \{\tau_1\} = \{\xi_1\} = \xi_1, \quad Z_2 = \{\tau_1 + \xi_2\}, \quad Z_3 = \{\tau_2 + \xi_3\}, \quad \ldots$$

Найдём вероятность события $S > s$, т. е. события $\tau_s < 1$ ($m \in \mathbb{N}$).

Во-первых, в силу леммы 1 все остатки Z_i равномерно распределены на интервале $(0;1)$: $Z_i \sim U(0,1)$, во-вторых, все они независимы попарно и в совокупности. В-третьих, остатки растут, пока $\tau_s < 1$, но как только сумма достигает 1, очередной остаток оказывается меньше предыдущего. Таким образом, событие $\tau_{s-1} < 1 \leqslant \tau_s$ случается только тогда, когда с остатками происходит событие $Z_1 < Z_2 < Z_3 < \ldots < Z_{s-1} > Z_s$, а событие $\tau_s < 1$ эквивалентно событию $Z_1 < Z_2 < Z_3 < \ldots < Z_{s-1} < Z_s$.

Нам удалось переформулировать задачу: имеются s последовательных независимых случайных чисел, равномерно распределённых на интервале $(0;1)$, и нужно найти вероятность того, что они случайным образом расположились по возрастанию в порядке их появления. Вероятность этого, очевидно, равна $1/s!$.

Для события $\tau_s < 1$ введём индикатор

$$I_s = \begin{cases} 1, & \text{если } \tau_s < 1, \\ & \text{т. е. если } s \text{ слагаемых не хватило для достижения суммы } n = 1, \\ 0, & \text{если } \tau_s \geqslant 1. \end{cases}$$

Случайная величина S легко выражается через эти индикаторы:

$$S = 1 + I_1 + I_2 + I_3 + \ldots = 1 + \sum_{s=1}^{\infty} I_s.$$

В этой сумме каждый следующий индикатор I_s равен единице, если s слагаемых не хватило, чтобы сумма достигла числа 1. Как только сумма 1 достигнута, все последующие индикаторы равны нулю.

Учитывая, что

$$\mathrm{E}I_s = \mathrm{P}(I_s = 1) = \mathrm{P}(\tau_s < 1) = \frac{1}{s!},$$

и переходя к математическим ожиданиям, получаем:

$$\mathrm{E}S = 1 + \sum_{s=1}^{\infty} \mathrm{E}I_s = 1 + \sum_{s=1}^{\infty} \frac{1}{s!} = \sum_{s=0}^{\infty} \frac{1}{s!} = e.$$

Получился уже известный результат: среднее количество слагаемых, необходимых для достижения в сумме единицы, равно e.

Возможно обобщение на другие целые суммы. Принцип тот же. Сумма n будет впервые достигнута при s слагаемых, если $\tau_{s-1} < n \leqslant \tau_s$. Так же как раньше, найдём вероятности события «s слагаемых недостаточно для достижения суммы n», т. е. события $\tau_s < n$. Это событие эквивалентно тому, что в последовательности независимых остатков $Z_1, Z_2, \ldots, Z_s$ ровно $n-1$ число меньше предыдущего. Заменив случайные числа Z_i их *рангами* $\rho(i)$ (ранг — номер в упорядоченном ряду), вместо случайной последовательности остатков получаем случайную перестановку их рангов. Таким образом, задача сводится к вопросу о том, какова вероятность того, что в случайной перестановке $(\rho(1), \rho(2), \rho(3), \ldots, \rho(s))$ длины s наблюдается ровно $n-1$ *падение*, т. е. ровно $n-1$ пара $\rho(i)$, $\rho(i+1)$, где $\rho(i) > \rho(i+1)$. Ответ на этот вопрос дают *числа Эйлера*[5)], см., например, [1, с. 297–300].

Определение. Числом Эйлера I рода $A(s, i)$ называется количество перестановок длины s, которые содержат ровно i падений (или подъёмов).

Очевидно, $A(s, 0) = 1$ при $s > 0$ — этим равенством мы пользовались в случае $n = 1$. Для общности обычно полагают $A(0, 0) = 1$, $A(s, i) = 0$ при $i < 0$ или при $s < 0$, а также при $0 < s \leqslant i$. При $s > 0$, $i \geqslant 0$ справедливо основное рекуррентное соотношение

$$A(s, i) = (i+1)A(s-1, i) + (s-i)A(s-1, i-1).$$

Имеется и явная формула:

$$A(s, i) = \sum_{j=0}^{i} (-1)^j C_{s+1}^j (i+1-j)^s.$$

Вывод этих соотношений остаётся за рамками статьи.

С помощью чисел Эйлера можно записать вероятность события $A_s < n$:

$$\mathrm{P}(A_s < n) = \sum_{i=0}^{n-1} \frac{A(s, i)}{s!}.$$

Снова введём индикаторы

$$I_s = \begin{cases} 1, & \text{если } A_s < n, \\ 0, & \text{если } A_s \geqslant n, \end{cases}$$

[5)] Числа Эйлера (или числа Эйлера I рода) не следует путать с эйлеровыми числами, которые возникают при разложении в степенной ряд гиперболического секанса.

составим сумму

$$S = 1 + \sum_{s=1}^{\infty} I_s$$

и перейдём к математическим ожиданиям:

$$\mathrm{E}S = 1 + \sum_{m=1}^{\infty} \mathrm{P}(I_m = 1) = 1 + \sum_{m=1}^{\infty} \mathrm{P}(A_m < n) = \\ = 1 + \sum_{m=1}^{\infty}\sum_{i=0}^{n-1} \frac{A(m,i)}{m!} = \sum_{m=0}^{\infty}\sum_{i=0}^{n-1} \frac{A(m,i)}{m!}. \qquad (13)$$

Последнее преобразование сделано с учётом соглашения

$$A(0,0) = 1, \quad A(0,i) = 0 \quad \text{при } i > 0.$$

Читатель может ради любопытства и удовольствия самостоятельно из равенства (13) получить равенство (10). Потребуются несколько утомительные, но естественные комбинаторные преобразования.

Список литературы

[1] *Кнут Д., Грэхем Р., Паташник О.* Конкретная математика. Основание информатики. М.: Мир, 1998.

[2] *Кокс Д. Р., Смит В. Л.* Теория восстановления. М.: Советское радио, 1967.

[3] *Мышкис А. Д.* Линейные дифференциальные уравнения с запаздывающим аргументом. М.: Наука, 1972.

[4] *Doob J. L.* Renewal theory from the point of view of the theory of probability // Trans. AMS. 1948. Vol. 63. P. 422–438.

[5] *Nebres P.* Renewal theory and its applications. The University of Chicago, 2011. http://math.uchicago.edu/~may/VIGRE/VIGRE2011/REUPapers/Nebres.pdf

[6] *Ross Sh. M.* The inspection paradox // Probab. Engrg. Inform. Sci. 2003. Vol. 17, № 1. P. 47–51.

Иван Ростиславович Высоцкий, МЦНМО

i_r_vysotsky@hotmail.com