И. ВЫСОЦКИЙ,
г. Москва

Задача 2

ЗАДАЧИ С УЛИЦЫ

Задача о варенье

Как правильно сливать остатки варенья

Мария Петровна варит варенье в огромном медном тазу, доставшемся ей еще от бабушки. Готовое варенье она разливает по литровым банкам строго доверху и тут же их закатывает на зиму.

Проблема в том, что варенья практически никогда не хватает на последнюю банку: его в тазу всегда оказывается разное количество, как Мария Петровна ни старается. Остатки, которых не хватает на целую банку, Мария Петровна сливает в специальную литровую банку, которую любовно называет *заветной банкой*. Ее Мария Петровна держит отдельно и не предназначает для зимнего хранения. Когда эта банка наполняется, Мария Петровна тут же заводит новую заветную банку, а полную немедленно подает на стол со словами «остатки сладки».

Задача 1. Сколько будет сварено тазов к моменту, когда наполнится первая заветная банка? Разумеется, речь идет о математическом ожидании этой величины.

В этой задаче трудно усмотреть связь с задачей о настольной игре (журнал «Математика» № 6). Тем не менее связь есть, и она очень тесная: можно считать, что остатки варенья в тазу — это случайная величина от 0 до 1 л. Тогда возникает вопрос: сколько таких случайных величин нужно, чтобы их сумма достигла или превысила 1 л, то есть как раз объем заветной банки.

Решение. Пусть J — то самое число тазов сваренного варенья и нужно найти $\mathrm{E}\,J$. Обозначение мы взяли, помня, что речь идет о варенье (англ. jam). Попробуем свести задачу к задаче о настольной игре.

Разобьем 1 литр на n маленьких объемчиков по $\frac{1}{n}$ литра. Для образности можно предположить, что Мария Петровна перекладывает варенье из таза в заветную банку посредством маленькой ложечки как раз такого объема.

Возникает случайная величина X_n «число тазов, которые будут сварены к моменту, когда суммарный объем остатков окажется не менее чем n ложечек, то есть не менее 1 л».

Нечто очень похожее мы уже делали, когда решали задачу о количестве бросков кубика, которые нужно сделать, чтобы сумма достигла заданного значения. Введем *индикаторы* событий «в первый раз слито ровно k ложечек» по следующему правилу:

$$I_1 = \begin{cases} 1, \text{ если в первом тазу оставалась 1 ложечка,} \\ 0, \text{ если иначе,} \end{cases}$$

$$I_2 = \begin{cases} 1, \text{ если в первом тазу оставалось 2 ложечки,} \\ 0, \text{ если иначе} \end{cases}$$

Есть дополнительные материалы на сайте raum.math.ru.

и так далее вплоть до индикатора I_n. Больше не нужно: если бы осталось больше, чем n ложечек, то есть больше литра, Мария Ивановна закатала бы эту банку на зиму.

Тут у внимательного читателя должен возникнуть вопрос: а что если остатков в первом тазу не хватит даже на одну ложечку? Этой трудности можно избежать двумя способами.

1. Включить в рассмотрение и этот случай тоже, введя индикатор I_0. Тогда возникнут небольшие технические сложности в предстоящих выкладках.

2. Увеличить n, тем самым уменьшая объем ложечки $\frac{1}{n}$ таким образом, чтобы хотя бы одна ложечка наверняка бы наскреблась.

Мы пойдем вторым путем. По сути, мы переформулировали задачу: имеется последовательность случайных слагаемых (остатков), каждое из которых с равными вероятностями принимает натуральные значения от 1 до n (число ложечек), и нужно найти математическое ожидание случайной величины X_n «число слагаемых к моменту, когда их сумма впервые достигла или превысила n». В этих терминах индикатор I_k является индикатором события «первое слагаемое в точности равно k».

Будем искать не $\mathrm{E}\,X_n$, а $\mathrm{E}\,X_m$, где $1 \le m \le n$. Можно написать равенство

$$X_m = 1 + I_1X_{m-1} + I_2X_{m-2} + I_3X_{m-3} + \ldots \\ \ldots + I_{n-1}X_{m-(n-1)} + I_nX_{m-n}. \quad (1)$$

Случайная величина X_k означает число слагаемых, не считая первого, потребовавшихся для того, чтобы сумма стала больше или равна k. Нас не смущает, что индекс может оказаться нулевым или отрицательным: чтобы сумма натуральных слагаемых превзошла отрицательное число, достаточно взять 0 слагаемых. Поэтому пусть $X_k = 0$ при $k \le 0$. Аналогично $X_1 = 1$, поскольку сумма, состоящая из одного-единственного натурального слагаемого, обязательно не меньше, чем 1.

Единица в начале суммы нужна, чтобы учесть первое слагаемое. Перейдем в равенстве к математическим ожиданиям:

$$\mathrm{E}\,X_m = 1 + \mathrm{E}\,(I_1X_{m-1}) + \mathrm{E}\,(I_2X_{m-2}) + \mathrm{E}\,(I_3X_{m-3}) + \ldots + \\ + \mathrm{E}\,(I_nX_{m-n}). \quad (2)$$

Величины I_k и X_{m-k} независимы, поскольку индикатор I_k относится только к первому слагаемому, а случайная величина X_{m-k} — ко всем последующим, но не к первому. Для независимых величин математическое ожидание произведения равно произведению ожиданий. Значит,

$$\mathrm{E}\,(I_kX_{n-k}) = \mathrm{E}\,I_k \cdot \mathrm{E}\,X_{n-k}.$$

Чтобы найти математическое ожидание индикаторов, составим их распределение. Все они распределены одинаково по закону Бернулли[1]:

$$I_k \sim \begin{pmatrix} 0 & 1 \\ \frac{n-1}{n} & \frac{1}{n} \end{pmatrix},$$

значит,

$$\mathrm{E}\,I_k = 0\cdot\frac{n-1}{n} + 1\cdot\frac{1}{n} = \frac{1}{n}.$$

Введем для $\mathrm{E}\,X_k$ краткое обозначение r_k, и равенство примет вид

$$r_m = 1 + \frac{1}{n}\left(r_{m-1} + r_{m-2} + \ldots + r_{m-n}\right) = 1 + \frac{1}{n}\sum_{k=1}^{n} r_{m-k}. \quad (3)$$

Чтобы разобраться в этой рекурсии, возьмем следующие начальные условия:

$$r_k = \mathrm{E}\,X_k = 0 \text{ при } k \le 0 \text{ и } r_1 = \mathrm{E}\,X_1 = 1.$$

Почему нужно взять именно такие условия, мы обсудили в абзаце после формулы (1). Попробуем подметить закономерность.

$$r_2 = 1 + \frac{1}{n}\Big(1 + \underbrace{0 + 0 + \ldots + 0}_{n-1 \text{ слагаемых}}\Big) = 1 + \frac{1}{n},$$

$$r_3 = 1 + \frac{1}{n}\Big(1 + \frac{1}{n} + 1 + \underbrace{0 + 0 + \ldots + 0}_{n-2 \text{ слагаемых}}\Big) =$$

$$= 1 + \frac{1}{n}\left(2 + \frac{1}{n}\right) = 1 + 2\cdot\frac{1}{n} + \frac{1}{n^2} = \left(1 + \frac{1}{n}\right)^2.$$

Закономерность проявляется ярко: при $1 \le m \le n$ верно равенство

$$r_m = \left(1 + \frac{1}{n}\right)^{m-1}. \quad (4)$$

Докажем его с помощью индукции. Мы знаем, что (4) верно для $m = 0$, 1 и 2. Таким образом, база индукции у нас замечательная. Докажем, что (4) верно для произвольного $m \le n$, если оно верно для всех предыдущих значений:

$$r_m = 1 + \frac{1}{n}\left(r_{m-1} + r_{m-2} + \ldots + r_{m-n}\right) =$$

$$= 1 + \frac{1}{n}\left(\left(1 + \frac{1}{n}\right)^{m-2} + \left(1 + \frac{1}{n}\right)^{m-3} + \ldots \right.$$

$$\left. \ldots + \left(1 + \frac{1}{n}\right)^0 + r_0 + r_{-1} + \ldots + r_{m-n}\right).$$

Вспомним, что

$$r_0 = r_{-1} = \ldots = r_{m-n} = 0,$$

а для суммирования прочих слагаемых применим формулу суммы геометрической прогрессии со знаменателем $1 + \frac{1}{n}$:

$$r_m = 1 + \frac{1}{n}\left(\left(1 + \frac{1}{n}\right)^{m-2} + \left(1 + \frac{1}{n}\right)^{m-3} + \ldots + \left(1 + \frac{1}{n}\right)^0\right) =$$

$$= 1 + \frac{1}{n}\left(\frac{\left(1 + \frac{1}{n}\right)^{m-1} - 1}{\left(1 + \frac{1}{n}\right) - 1}\right) = 1 + \left(\left(1 + \frac{1}{n}\right)^{m-1} - 1\right) = \left(1 + \frac{1}{n}\right)^{m-1}.$$

[1] Распределение бинарной случайной величины часто называют распределением Бернулли и обозначают буквой B. Можно записать, что $I_k \sim B\left(\frac{1}{n}\right)$. Число в скобках — параметр распределения — равен вероятности того, что $I_k = 1$.

Мы почти достигли цели. Осталось вспомнить, что нас интересует на самом деле только случай $m = n$. Получается:

$$\mathrm{E}X_n = r_n = \left(1+\frac{1}{n}\right)^{n-1}. \qquad (5)$$

Чтобы избавиться от предположения, что Мария Петровна черпает остатки варенья из таза маленькой ложечкой, устремим n к бесконечности. Тем самым мы позволяем Марии Петровне свободно переливать остатки из таза в заветную банку непрерывной струйкой:

$$\mathrm{E}J = \lim_{n\to\infty}\mathrm{E}X_n = \lim_{n\to\infty}\left(1+\frac{1}{n}\right)^{n-1} =$$

$$= \lim_{n\to\infty}\left(1+\frac{1}{n}\right)^{n}\cdot\lim_{n\to\infty}\frac{n}{n+1} = e\cdot 1 = e.$$

Чтобы наполнить заветную банку, Марии Петровне в среднем придется сварить e = 2,7182818284... тазов, то есть среднее число тазов равно основанию натурального логарифма.

Случай, когда заветных банок больше одной

У вас уже наверняка появился вопрос, как быстро будет наполнена вторая заветная банка, а за ней третья. Возникает подозрение, что вторую банку Мария Петровна наполнит, перелив в нее остатки из $2e$ тазов (в среднем). Увы, не все так просто. Дело в том, что когда первая заветная банка отправляется на стол, вторая, скорее всего, не пуста: в нее слито немножко варенья, которое не поместилось в первую. Поэтому выражения для второй, третьей и последующих заветных банок сложнее:

$$\mathrm{E}\,J_2 = e^2 - e,$$

$$\mathrm{E}\,J_3 = e^3 - 2e^2 + \frac{1}{2}e,$$

$$\mathrm{E}\,J_4 = e^4 - 3e^3 + 2e^2 - \frac{1}{6}e.$$

К сожалению, вывести эти равенства с помощью разбиения на маленькие ложечки очень сложно, если вообще возможно: замечательная закономерность (4), которой мы с успехом воспользовались, после первой заветной банки нарушается и становится не такой замечательной.

Общее рассуждение можно провести непосредственно с непрерывными функциями, решая дифференциальное уравнение вместо рекурсии. Результат следующий. Если нужно, чтобы сумма случайных слагаемых (остатков варенья) заполнила $N \in \boldsymbol{L}$ заветных банок, то среднее число слагаемых равно

$$\mathrm{E}\,J_N = \sum_{k=0}^{N}\frac{(-1)^{N-k}k^{N-k}}{(N-k)!}e^k. \qquad (6)$$

Структура этого выражения крайне интересная. Выпишем члены рядов последовательных отрицательных степеней числа e в виде таблицы (рис. 1). При этом каждый следующий столбец сместим на одну строку вниз по отношению к предыдущему. Во всех столбцах каждое число, кроме верхней единицы, получается умножением сверху стоящего числа на множитель $-\frac{k}{N-k}$. Тогда в строках таблицы чудесным образом получаются коэффициенты при e^k из равенства (6).

N	e^{-k}									
	e^{-1}	e^{-2}	e^{-3}	e^{-4}	e^{-5}	e^{-6}	e^{-7}	...	e^{-k}	...
1	1									
2	−1	1								
3	$\frac{1}{2}$	−2	1							
4	$-\frac{1}{6}$	2	−3	1						
5	$\frac{1}{24}$	$-\frac{4}{3}$	$\frac{9}{2}$	−4	1					
6	$-\frac{1}{120}$	$\frac{2}{3}$	$-\frac{9}{2}$	8	−5	1				
7	$\frac{1}{720}$	$-\frac{4}{15}$	$\frac{27}{8}$	$-\frac{32}{3}$	$\frac{25}{2}$	−6	1			
...	...	...	...	...	...	...	...	...		...

Рис. 1

Выражение (6) может быть продолжено до непрерывной функции, распространяясь на действительный аргумент x вместо N, и имеет линейное приближение:

$$\mathrm{E}\,J_x = 2x + \frac{2}{3} + \varphi(x),$$

где функция φ очень быстро стремится к нулю (на рисунке 2 график функции $\mathrm{E}\,J_x$ невозможно на глаз отличить от прямой уже при $x > 2$).

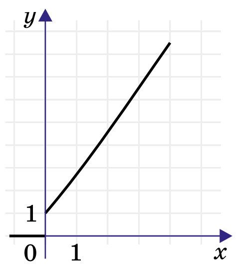

Рис. 2

Подробнее об этой задаче в непрерывной форме, о том, как решить эту задачу комбинаторными средствами, и о том, как она связана с числами Эйлера первого рода, можно прочесть в журнале «Математическое просвещение»,

СУММ =ЕСЛИ(G$6=$C11;1;ЕСЛИ(G$6<$C11;G10*(-G$6/($C11-G$6));""))

Задача о варенье

Математическое ожидание числа случайных слагаемых,

равномерно распределенных на отрезке [0;1], в сумме, впервые достигшей числа N<=12

	k	1	2	3	4	5	6	7	8	9	10	11	12	EJ_N
	e^k	2,7183	7,3891	20,0855	54,598	148,413	403,429	1096,63	2980,96	8103,08	22026,5	59874,1	162755	
N	1	1,0000												2,718282
	2	-1,0000	1,0000											4,670774
	3	0,5000	-2,0000	1,0000										6,666566
	4	-0,1667	2,0000	-3,0000	=ЕСЛИ(G$6=$C11;1;ЕСЛИ(G$6<$C11;G10*(-G$6/($C11-G$6));""))									8,666604
	5	0,0417	-1,3333	4,5000	-4,0000	1,0000								10,66666
	6	-0,0083	0,6667	-4,5000	8,0000	-5,0000	1,0000							12,66667
	7	0,0014	-0,2667	3,3750	-10,6667	12,5000	-6,0000	1,0000						14,66667
	8	-0,0002	0,0889	-2,0250	10,6667	-20,8333	18,0000	-7,0000	1,0000					16,66667
	9	0,0000	-0,0254	1,0125	-8,5333	26,0417	-36,0000	24,5000	-8,0000	1,0000				18,66667
	10	0,0000	0,0063	-0,4339	5,6889	-26,0417	54,0000	-57,1667	32,0000	-9,0000	1,0000			20,66667
	11	0,0000	-0,0014	0,1627	-3,2508	21,7014	-64,8000	100,0417	-85,3333	40,5000	-10,0000	1,0000		22,66667
	12	0,0000	0,0003	-0,0542	1,6254	-15,5010	64,8000	-140,0583	170,6667	-121,5000	50,0000	-11,0000	1,0000	24,66667

Рис. 3

сер. 3, вып. 24, 2019 (с. 101–120). К сожалению, автор ничего не знает о дисперсии случайной величины J даже в случае одной заветной банки. Здесь может помочь численное моделирование.

Решение задачи в MS Excel

Задача 2. Написать таблицу в MS Excel для вычисления среднего числа слагаемых при достижении суммы N по формуле (6) .

На рисунке 3 представлено решение.

В столбце P результаты:

$$\mathrm{E}\,J = e = 2{,}71828\ldots,$$

$$\mathrm{E}J_2 = e^2 - e = 4{,}670774\ldots \text{ и т. д.}$$

Задача 3. Написать таблицу в MS Excel для моделирования математического ожидания и дисперсии числа слагаемых при достижении суммы 1.

Вариант решения на рисунке 4.

В таблице на рисунке 4 для получения случайных чисел на интервале от 0 до 1 использована функция СЛЧИС().

В ячейках нечетных строк (вторая строка в каждом эксперименте) проверяется, достигла ли единица суммы случайных слагаемых предыдущей строки, и если да, то в ячейке выводится номер слагаемого.

В столбце R вычисляется номер первого слагаемого, когда это случилось. Столбец R и дает нуж-

СУММ =ЕСЛИ(СУММ($D6:E6)>1;E$5;"")

Задача о варенье

Численное моделирование. Среднее и дисперсия числа равномерных стандартных слагаемых,

сумма которых впервые достигла 1.

	Ном.слаг.	1	2	3	4	5	6	7	8	9	10	11	12	13	14	Чис.слаг	Среднее	Медиана
Номер эксперимента	1	0,5484	0,24971	0,76565	0,16176	0,40279	0,2088	0,46857	0,60471	0,73879	0,26411	0,99064	0,79948	0,46035	0,98171	3	2,727273	2,5
			=ЕСЛИ(СУММ($D6:E6)>1;E$5;"")				6	7	8	9	10	11	12	13	14		Мин	Макс
	2	0,7077	0,64982	0,50789	0,85745	0,00955	0,73694	0,43084	0,33928	0,37164	0,90629	0,42999	0,62862	0,61381	0,81376	2	2	5
			2	3	4	5	6	7	8	9	10	11	12	13	14		Диспр	Ст.откл
	3	0,7636	0,34418	0,53336	0,98142	0,06734	0,5265	0,55575	0,17268	0,62135	0,36863	0,14679	0,07986	0,34742	0,47766	2	0,743802	0,862439
			2	3	4	5	6	7	8	9	10	11	12	13	14			
	4	0,2613	0,07705	0,38381	0,33223	0,5572	0,91058	0,36294	0,76925	0,68257	0,80198	0,12836	0,77096	0,95455	0,7551	4		
					4	5	6	7	8	9	10	11	12	13	14			
	5	0,8629	0,39707	0,11827	0,34671	0,63909	0,33852	0,24917	0,26904	0,53803	0,66759	0,40986	0,48627	0,34513	0,012	2		
			2	3	4	5	6	7	8	9	10	11	12	13	14			
	6	0,1463	0,95601	0,15554	0,86953	0,66666	0,41297	0,1064	0,3299	0,49292	0,64866	0,11262	0,66123	0,04368	0,14865	2		
			2	3	4	5	6	7	8	9	10	11	12	13	14			
	7	0,2866	0,45875	0,97009	0,46958	0,36574	0,96081	0,49257	0,54081	0,99201	0,69784	0,26308	0,0876	0,30313	0,74349	3		
				3	4	5	6	7	8	9	10	11	12	13	14			
	8	0,3801	0,59878	0,15508	0,4503	0,62313	0,14678	0,92491	0,42727	0,84746	0,80699	0,69678	0,10844	0,82679	0,08682	3		
				3	4	5	6	7	8	9	10	11	12	13	14			
	9	0,6618	0,70602	0,37434	0,24613	0,77891	0,48952	0,35218	0,77296	0,46323	0,70337	0,53123	0,02465	0,2111	0,86325	2		
			2	3	4	5	6	7	8	9	10	11	12	13	14			
	10	0,746	0,31662	0,99569	0,03951	0,24098	0,88643	0,94417	0,46481	0,83405	0,30807	0,23708	0,13743	0,83968	0,15609	2		
			2	3	4	5	6	7	8	9	10	11	12	13	14			
	11	0,0508	0,74073	0,08427	0,2265	0,56245	0,07082	0,12056	0,60557	0,34956	0,42723	0,73029	0,94856	0,96381	0,20113			

Рис. 4

ное моделирование. В столбцах S и T вычисляются описательные характеристики массива R.

Гораздо удобнее было бы сделать таблицу со случайными числами на одном листе, а все проверки и вычисления — на другом. Но в этом случае таблица была бы менее наглядна. Школьник, занявшийся подобным проектом, легко модифицирует таблицу наиболее удобным образом.

Объем выборки небольшой — всего 22 эксперимента. Для получения более представительных данных таблицу нужно расширить хотя бы до 200 экспериментов, а лучше — больше. Это не составит труда.

Обратите внимание: в ячейке S6 полученное среднее равняется 2,727273..., что весьма близко к числу e.

Для оценки и изучения дисперсии можно пользоваться ячейкой S10.

Электронная таблица размещена на сайтах raum.math.ru и на сайте ptlab.mccme.ru.

Задачи для самостоятельного решения

1. Можно ли при постановке задачи заменить таз кастрюлей объемом 40 л? Как изменятся решение и ответ?

2. Предположим, что Мария Петровна вычерпывает остатки из тазов с вареньем и накладывает их в заветную банку объемом 1 л с помощью чайной ложечки объемом 2 мл. Сколько тазов варенья ей придется сварить, чтобы заветная банка наполнилась хотя бы наполовину?

3. Сколько тазов варенья нужно сварить, чтобы наполнить пять заветных банок?

Ответы к задачам из статьи «Задача о настольной игре»

1. 16. **2.** $p_n = \frac{1}{8}\sum_{k=1}^{n-k} p_{n-k}$; $p_8 = \frac{9^7}{8^8} = 0{,}2850\ldots$; $p_{256} \approx \frac{2}{9} = 0{,}(2)$. **3.** 578,19. **4.** $\left(1+\frac{1}{n}\right)^{n-1} \approx e.$

Из интернета

Нейробиологи объяснили, почему не всем детям дается математика

Наталья Панасенко, rg.ru

Немецкие специалисты обнаружили ген, «отвечающий» за математические способности. По их данным, основную роль в развитии математических способностей у детей играет ген ROBO1.

Как утверждают ученые из Института изучения когнитивных процессов и мозга человека им. Макса Планка (Германия), результаты математических работ выше среднего уровня наблюдаются у детей «с ранними индивидуальными различиями в правой теменной доле мозга, что связано с геном ROBO1».

«Математические способности формируются благодаря сложному взаимодействию генов и окружающей среды, — говорит Майкл Скейде, руководитель работы. — Исследования показывают, что 60% различий связаны с генетической предрасположенностью. Вот почему способности к математике часто передаются по наследству».

Однако пока мало известно, какие гены способствуют развитию математических способностей и как они влияют на мозг. «Поэтому нашей целью было выявить связи между известными генами-кандидатами, влияющими на развитие мозга, и насколько их вариации критичны для проявления математических способностей в раннем возрасте, до того, как дети начнут обучение математике», — пишут нейробиологи.

В исследовании Института им. Макса Планка (Германия) участвовали 178 обычных детей от трех до шести лет (без какой-либо математической подготовки). Ученые провели анализ генов, которые предположительно отвечают за математические способности у детей, и с помощью магнитно-резонансной томографии сравнили объем серого вещества во всех отделах детского мозга, особенно в тех областях, в которых объем коррелировал с результатами тестов по математике.

Когда этим детям было уже от семи до девяти лет, нейробиологи сравнили полученные данные с результатами стандартного теста по математике. Результаты математических работ выше среднего уровня были у детей с отчетливо представленными генетическими изменениями. И эти изменения были связаны с геном ROBO1. Исследователи обнаружили, что у детей, несущих ген ROBO1, значительно больше серого вещества в задней части правой теменной доли — в областях мозга, которые считаются центром обработки числовой информации и визуально-пространственной рабочей памяти.

Майкл Скейде утверждает, что математический талант проявляется уже в раннем детстве, поэтому предвидеть успехи детей можно задолго до того, как ребенок сможет решить свои первые арифметические задачи. Дело в том, что у некоторых детей, очевидно, в мозгу есть «более развитый математический центр». И если со стороны родителей и школы будет соответствующая поддержка, этот талант может полностью развиться.