И. ВЫСОЦКИЙ,
г. Москва

Задача 14

ЗАДАЧИ С УЛИЦЫ

Задача о чемоданах

Случайные опыты с равновозможными элементарными событиями в природе практически не встречаются, почти все такие опыты — дело рук человеческих. Существует огромное количество жребиев и лотерей, часть из которых не отличается избыточной честностью. Но все же иногда лотереи возникают естественным образом. Например, выдача чемоданов в аэропорту часто превращается в лотерею. Не в том смысле, получите вы свой чемодан или нет, а в смысле, когда же он, наконец, покажется на ленте транспортера.

Если в самолете есть бизнес-класс или привилегированные пассажиры, то их багаж обычно выдается в первую очередь. Да и то не всегда: в некоторых аэропортах не предусмотрено приоритетное обслуживание чемоданов с VIP-ярлыками.

Будем считать, что чемоданы сдаются в случайном порядке, перемешиваются на линиях сортировки, затем грузчики устраивают беспредел при загрузке багажного отсека самолета, а их коллеги в аэропорту назначения порядка не добавляют. И тогда на ленту транспортера в зале выдачи багажа чемоданы поступают в совершенном беспорядке.

Предположим, что Петровы — папа, мама и двое детей — летят на отдых именно таким рейсом. У них a чемоданов, и они, разумеется, не рассчитывают на то, что все их чемоданы вернутся к ним так же одномоментно, как они их сдавали. Но вот как они вернутся? Когда ждать первого? Как долго ждать последнего? И когда уже начинать волноваться?

Прежде чем формулировать и решать задачи, обратим внимание вдумчивого читателя на то, что вероятностная модель здесь в точности такая же, какая возникает при раздаче карт «на туза» (см. статью «Задача о жребии на туза», журнал «Математика», № 8, 2023). Роль тузов играют чемоданы семьи Петровых. Существенное отличие в том, что на сей раз нас не интересуют остатки от деления: пассажиры не картежники, а чемоданы не раздаются им по кругу. Сейчас нам интересны средние значения и возможные отклонения от этих средних.

Задача о первом чемодане

Задача 1. Всего из самолета на транспортер в случайном порядке выгружают n чемоданов. Каким по счету приедет первый чемодан семьи Петровых? Каковы математическое ожидание и дисперсия этой случайной величины?

Закодируем чемоданы на ленте с помощью двоичной последовательности. Чемоданы семьи Петровых будем отмечать символом 1, а чужие чемоданы — символом 0. Например, последовательность 0010101100 получится, если всего чемоданов $n = 10$, чемоданов Петровых $a = 4$ и они оказались по счету третьим, пятым, седьмым и восьмым.

Есть дополнительные материалы на сайте raum.math.ru.

Интуитивное решение задачи о среднем. Вообразим себе, что у Петровых есть еще один чемодан — виртуальный, или *призрачный*, о котором они даже не знают, и что этот чемодан всегда выдается последним. Тогда двоичная последовательность разбивается на (a + 1) фрагментов, каждый из которых оканчивается единицей. В приведенном примере последовательность 00101011001 получит красную жирную единицу в конце, символизирующую призрачный чемодан Петровых, и получится 5 фрагментов:

(001), (01), (01), (1), (001).

Интуитивная часть решения состоит в том, чтобы сообразить, что ни один фрагмент не хуже и не лучше других, а потому длины этих (a + 1) фрагментов суть одинаково распределенные случайные величины, сумма которых равна (n + 1). Поэтому средняя длина каждого фрагмента будет $\frac{n+1}{a+1}$.

Пример. Если в самолете было 200 чемоданов, а у Петровых было 4 чемодана, то математическое ожидание случайной величины «номер первого чемодана Петровых в последовательности чемоданов на ленте» равно $\frac{201}{5}=40{,}2$.

Это решение, заметим, совершенно верное, но немного скользкое: в нем присутствуют не слишком убедительные рассуждения о справедливости и непонятный призрачный чемодан. Дадим точное решение, которое позволит нам найти не только ожидание, но и дисперсию. Введем индикаторы: будем метить не все чемоданы, а только чужие.

Точное решение. Пронумеруем числами от 1 до (n – a) чемоданы, не принадлежащие семье Петровых, то есть чужие. Пусть I_k — индикатор события A_k «k-й чужой чемодан появился прежде всех чемоданов Петровых», то есть:

I_k = 1, если так и случилось,

I_k = 0, если перед k-м чужим чемоданом вклинился хотя бы один чемодан Петровых.

Тогда сумма всех индикаторов равна числу чужих чемоданов, появившихся перед первым чемоданом Петровых. Если добавить единицу, то получится случайная величина X_1 «номер первого чемодана Петровых»:

$$X_1 = 1 + I_1 + I_2 + \ldots + I_{n-a}.$$

Чтобы найти вероятность события A_k, то есть события I_k = 1, вообразим себе только часть двоичной последовательности, а именно подпоследовательность, состоящую из a единиц (чемоданов Петровых) и одного нуля (k-го чужого чемодана).

$$\ldots\underbrace{0\ldots\overbrace{1\ldots1\ldots1\ldots1}^{a\text{ единиц}}}_{\text{Всего }(a+1)\text{ элементов}}\ldots$$

Символ 0 среди этих (a + 1) элементов окажется на первом месте с вероятностью $\frac{1}{a+1}$. Поэтому

$$\mathrm{E}I_k=\frac{1}{a+1},$$

а

$$\mathrm{E}X_1=\mathrm{E}I_1+\ldots+\mathrm{E}I_{n-a}+1=\frac{n-a}{a+1}+1=\frac{n+1}{a+1}. \qquad (1)$$

Дисперсию найдем, используя этот же подход. Сначала найдем $\mathrm{E}X_1^2$. Имеем:

$$X_1^2=\left(1+I_1+I_2+\ldots+I_{n-a}\right)^2=$$

$$=1+I_1^2+I_2^2+\ldots+I_{n-a}^2+2\left(I_1+I_2+\ldots+I_{n-a}\right)+$$

$$+2\left(I_1I_2+\ldots+I_jI_k+\ldots+I_{n-a-1}I_{n-a}\right).$$

В скобках сумма всевозможных попарных произведений I_jI_k, где $1\leqslant j<k\leqslant n-a$.

Учитывая, что $I_k^2=I_k$, получаем:

$$\mathrm{E}X_1^2=$$

$$=\mathrm{E}\left(X_1+2\left(X_1-1\right)+2\left(I_1I_2+\ldots+I_jI_k+\ldots+I_{n-a-1}I_{n-a}\right)\right)=$$

$$=3\mathrm{E}X_1-2+$$
$$+2(\mathrm{E}(I_1I_2)+\ldots+\mathrm{E}(I_jI_k)+\ldots+\mathrm{E}(I_{n-a-1}I_{n-a})).$$

Осталось понять, чему равно $\mathrm{E}(I_jI_k)$. Случайная величина I_jI_k равна 1, только если чужие чемоданы с номерами j и k появились на ленте транспортера прежде, чем первый из чемоданов Петровых. Вероятность этого равна $\frac{2}{(a+2)(a+1)}$. Это можно найти тем же рассуждением, каким мы нашли, что $\mathrm{P}\left(I_k=1\right)=\frac{1}{a+1}$, а можно более формально:

$$\mathrm{P}\left(I_jI_k=1\right)=\frac{C_2^2C_a^0}{C_{a+2}^2}=\frac{2}{(a+2)(a+1)}.$$

Следовательно,

$$\mathrm{E}X_1^2=3\frac{n+1}{a+1}-2+2C_{n-a}^2\frac{2}{(a+2)(a+1)}=$$

$$=3\frac{n+1}{a+1}-2+2\frac{(n-a)(n-a-1)}{(a+2)(a+1)}=$$

$$=\frac{3(n+1)(a+2)-2(a+1)(a+2)+2(n-a)(n-a-1)}{(a+2)(a+1)}=$$

$$=\frac{2n^2-an+4n-a+2}{(a+2)(a+1)}.$$

Теперь легко найти дисперсию:

$$\mathrm{D}X_1=\mathrm{E}X_1^2-\mathrm{E}^2X_1=\frac{2n^2-an+4n-a+2}{(a+2)(a+1)}-\frac{(n+1)^2}{(a+1)^2}=$$

$$=\frac{an+n^2a-a^2n-a^2}{(a+2)(a+1)^2}=\frac{a(n+1)(n-a)}{(a+2)(a+1)^2}. \qquad (2)$$

Для вычислений, как обычно, составим электронную таблицу. Поскольку для ожидания и дисперсии получились явные выражения без всяких рекурсий, таблицу можно сделать очень лаконичной (рис. 1):

C7 fx =C4*(C3+1)*(C3-C4)/(C4+2)/(C4+1)^2

	A	B	C	D	E	F
1		Задача о чемоданах. Задача 1				
2						
3		n=	200		m =	2
4		a=	4		Интервал	
5					от EX-ms	до EX+ms
6		EX =	40,200		-24,625	105,025
7		DX =	1050,560			
8		Ст.откл s=	32,412			

Рис. 1. Вычисления в MS Excel по формулам (1) и (2)

Помимо расчета математического ожидания и дисперсии, добавим в таблицу вычисление интервала значений величины X_1 «номер первого чемодана» шириной $2m$ стандартных отклонений с центром в $\mathrm{E}X_1$. Например, при $m = 2$ получается интервал от $\mathrm{E}X_1 - 2\sqrt{\mathrm{D}X_1}$ до $\mathrm{E}X_1 + 2\sqrt{\mathrm{D}X_1}$, то есть от $-24{,}625$ до $105{,}025$. Разумеется, левую границу нужно поднять до 1, поскольку порядковый номер чемодана может быть только натуральным числом. Таким образом, разумно ждать, что первый чемодан Петровых объявится среди первых 105 чемоданов.

О последующих чемоданах

Предположение, что чемоданы совершенно случайно перемешаны, было более или менее оправданным, пока речь шла только о первом чемодане Петровых. Если вспомнить, что Петровы сдавали в аэропорту багаж на одной и той же стойке регистрации, то равномерное рассеивание их чемоданов по всей совокупности чемоданов становится сомнительным. Можно было бы попробовать уточнить математическую модель, введя в нее функцию, описывающую искажение равномерности, при условии, что известна позиция первого из их чемоданов. Можно даже пофантазировать на тему, как эта функция устроена. Однако такое усложнение модели сделает невозможным ее изучение элементарными средствами, хотя не препятствует компьютерному моделированию.

Мы же будем, как и прежде, предполагать случайность и равномерность распределения всех чемоданов на ленте, даже если речь идет не о первом чемодане Петровых, а о первой паре, первой тройке и т.п.

Задача 2. Когда приедут остальные чемоданы Петровых?

Пусть речь о чемодане Петровых номер j. Обозначим номер этого чемодана на ленте через X_j.

Модифицируем индикаторы: пусть $I_k = 1$, только если k-й чужой чемодан появился на ленте прежде, чем j-й Петровых. Теперь нам подходит не одно, а несколько взаимных расположений нуля и a единиц в двоичной подпоследовательности. Символ 0 теперь может занять не обязательно первое место, но также и второе, третье, вплоть до j-го. Вероятность этого $\frac{j}{a+1}$. Значит,

$$\mathrm{E}X_j = \frac{(n-a)j}{a+1} + 1 + (j-1) = \frac{j(n+1)}{a+1},$$

что и так было понятно из наших интуитивных рассуждений.

Таким образом, если у Петровых 4 чемодана из 200, находившихся в самолете, то номер последнего их чемодана в среднем равен $\frac{4 \cdot 201}{5} = 160{,}8$, что бы это ни значило.

Когда начинать беспокоиться?

Следующая задача наиболее естественная: в какой момент Петровы должны начать волноваться, если их чемоданов все нет и нет? Мы по-прежнему считаем, что все чемоданы распределены равномерно и случайно, хотя понимаем, что в реальности это не совсем так. Сначала найдем вероятность того, что среди большого количества приехавших чемоданов нет ни одного чемодана Петровых.

Задача 3. Какова вероятность того, что ни одного из нужных чемоданов еще нет, хотя почти все пассажиры уже получили багаж и разошлись?

По-прежнему будем считать, что всего чемоданов n. Пусть k из них уже появились на ленте и только $(n - k)$ чемоданов еще томятся в застенках службы выдачи багажа. Найдем вероятность события A, которое состоит в том, что среди этих $(n - k)$ чемоданов находятся все a чемоданов семьи Петровых. Эта задача стандартная: $(n - k)$ скрытых до сих пор чемоданов образуют случайную выборку из совокупности объемом n, в которой ровно a чемоданов Петровых. Поэтому

$$\mathrm{P}(A) = \frac{C_a^a C_{n-a}^{n-k-a}}{C_n^{n-k}} = \frac{C_{n-a}^k}{C_n^k}. \qquad (3)$$

Задача 4. В какой момент следует решить, что равномерность распределения нарушена или что какой-то из чемоданов Петровых вообще не прилетел, если считать маловероятными события с вероятностью меньше, чем наперед заданное $\varepsilon > 0$?

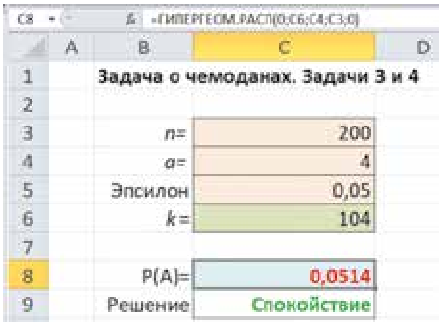
Эту задачу мы уже практически решили на языке доверительных интервалов. Если взять доверительный интервал шириной $4\sqrt{\mathrm{D}X_1}$, то верхняя граница интервала, в который должен попасть первый чемодан, равна, как мы помним, 105 (см. рис. 1). Это так, если мы выбрали именно такое решающее правило: плюс-минус два стандартных отклонения. Однако мы не очень понимаем, насколько малой вероятности ε это правило соответствует, поскольку пока не знаем, как распределена величина X_1.

Попробуем решить задачу, задав значение допустимой вероятности ε. Именно она будет устанавливать границу между спокойствием и волнением Петровых. Как только число появившихся на ленте чемоданов (k штук) станет настолько большим, что вероятность события A окажется меньше ε, глава семьи должен предположить, что что-то пошло не так. То есть нужно найти наименьшее целое решение неравенства

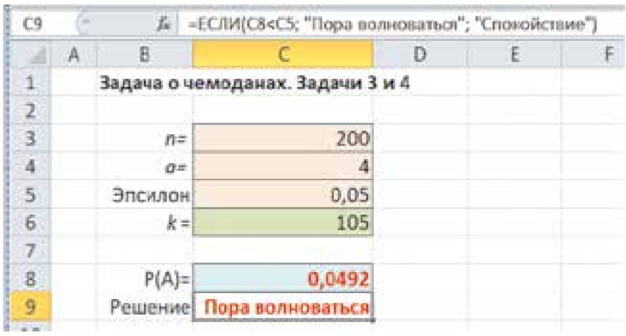
$$\frac{C_{n-a}^{k}}{C_{n}^{k}} < \varepsilon \qquad (4)$$

относительно k, считая числа a и n известными и постоянными.

Решение с помощью электронной таблицы показано на рисунке 2 для $n = 200$, $a = 4$ и $\varepsilon = 0{,}05$. Число k в ячейке C6 находится подбором. Нужно начать, например, с $k = 50$ и повышать его до тех пор, пока число в ячейке C8 не станет меньше заданного значения ε в ячейке C5.


C8 =ГИПЕРГЕОМ.РАСП(0;C6;C4;C3;0)

	A	B	C	D
1		Задача о чемоданах. Задачи 3 и 4		
2				
3		n=	200	
4		a=	4	
5		Эпсилон	0,05	
6		k=	104	
7				
8		P(A)=	0,0514	
9		Решение	Спокойствие	


C9 =ЕСЛИ(C8<C5; "Пора волноваться"; "Спокойствие")

	A	B	C	D	E	F
1		Задача о чемоданах. Задачи 3 и 4				
2						
3		n=	200			
4		a=	4			
5		Эпсилон	0,05			
6		k=	105			
7						
8		P(A)=	0,0492			
9		Решение	Пора волноваться			

Рис. 2. Расчет в MS Excel. Для расчета по формуле (3) использована функция ГИПЕРГЕОМ.РАСП

Как видим, 104 чужих чемодана еще не должны переполнить чашу терпения Петровых, а 105 уже должны. Результат поистине удивительным образом совпал с тем, что мы получили, исходя из правила «плюс-минус два стандартных отклонения».

Как сделать оценку без электронной таблицы

Можно обойтись без компьютера, если применить немножко математического анализа. Например, известно, что вблизи нуля функции $y = x + 1$ и $y = e^x$ эквивалентны в том смысле, что график первой является касательной к графику второй, поэтому их значения близки в малой окрестности точки $x = 0$ (рис. 3). При этом $e^x > x + 1$. Разумеется, эта близость лишь первого порядка, вторые производные в нуле у этих функций уже не равны. Но часто эквивалентности первого порядка хватает для многих приложений.

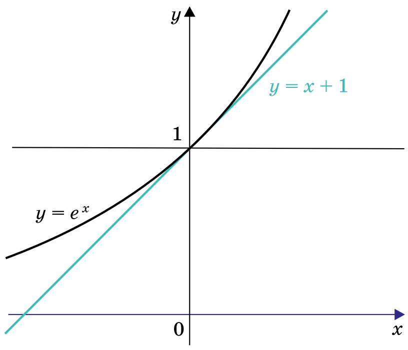


Рис. 3

Решим неравенство (4) приближенно: найдем, может быть, чуть более слабую оценку для k, но такую, при которой это неравенство будет заведомо выполнено. Если на ленту выгружено уже k чемоданов, то не выгружено $(n - k)$ чемоданов. Для краткости назовем это число m. Неравенство (4) примет вид

$$\frac{C_{n-a}^{n-m}}{C_{n}^{n-m}} < \varepsilon,$$

откуда непосредственным переходом к факториалам в левой части получаем:

$$\frac{(n-a)!\,m!}{n!\,(m-a)!} = \frac{(n-a)(n-a-1)\cdot\ldots\cdot(m-a+1)}{n(n-1)\cdot\ldots\cdot(m+1)} =$$

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

Разумно считать, что m намного больше числа a, ибо в противном случае волноваться следовало бы уже давно. Тогда все числа от $\frac{a}{n}$ до $\frac{a}{m+1}$ малы, и разумно сделать оценку:

$$\left(1-\frac{a}{n}\right)\left(1-\frac{a}{n-1}\right)\cdot\ldots\cdot\left(1-\frac{a}{m+1}\right) <$$

$$< e^{-\frac{a}{n}} \cdot e^{-\frac{a}{n-1}} \cdot \ldots \cdot e^{-\frac{a}{m+1}} = e^{-\left(\frac{a}{n}+\frac{a}{n-1}+\ldots+\frac{a}{m+1}\right)} = e^{-a(H_n - H_m)},$$

где

$$H_n = 1 + \frac{1}{2} + \ldots + \frac{1}{n}$$

— гармоническое число (см. статьи «Задача коллекционера» и «Задача об узком шоссе»). Мы знаем, что гармоническое H_n близко к $\ln n$. Этим обстоятельством можно плодотворно воспользоваться, но лучше мы оценим не слагаемые по отдельности, а разность $(H_n - H_m)$ целиком.

Представим искомое число

$$H_n - H_m = \frac{1}{m+1} + \frac{1}{m+2} + \ldots + \frac{1}{n}$$

как площадь многоугольника, составленного из столбиков единичной ширины на отрезке от $(m + 1)$ до $(n + 1)$ (рис. 4).

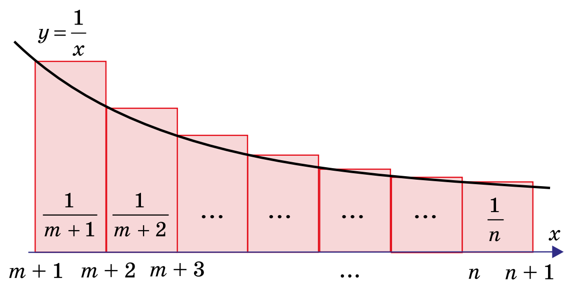


Рис. 4

Площадь многоугольника больше, чем площадь криволинейной трапеции, заключенной под графиком функции $y = \frac{1}{x}$ на отрезке $[m + 1; n + 1]$:

$$H_n - H_m > \int_{m+1}^{n+1} \frac{dx}{x} = \ln\frac{n+1}{m+1}. \qquad (5)$$

Поэтому

$$\frac{C_{n-a}^{n-m}}{C_n^{n-m}} < e^{-a(H_n - H_m)} < e^{-a\ln\frac{n+1}{m+1}} = \left(\frac{m+1}{n+1}\right)^a.$$

Если эту величину сделать меньше, чем число ε, то неравенство (4) будет заведомо выполнено:

$$\left(\frac{m+1}{n+1}\right)^a < \varepsilon,$$

откуда

$$m+1 < (n+1)\sqrt[a]{\varepsilon}.$$

Тогда

$$k > (n+1)(1-\sqrt[a]{\varepsilon}). \qquad (6)$$

Для проверки положим $n = 200$, $a = 4$ и $\varepsilon = 0{,}05$, как делали при непосредственном подсчете (см. рис. 2). Тогда получилось $k = 105$. Посмотрим, что после всех сделанных оценок и ослаблений даст неравенство (6):

$$k > 201(1-\sqrt[4]{0{,}05}) \approx 105{,}95,$$

то есть наименьшее k получается 106. Оценка (6) почти совпадает с оценкой (4), при этом оценку (6) можно найти с помощью карандаша и клочка бумаги, не говоря о калькуляторе.

Сделанные выкладки базируются на двух последовательных оценках: на каждом шаге левую часть неравенства (4) мы оценивали сверху более удобным выражением, при этом мы все время предполагали, что число m чемоданов, еще не попавших на транспортер, существенно больше, чем число a чемоданов Петровых. Без этого оценка (6) получится грубой. Интересно было бы поисследовать вопрос, насколько при данном n широк интервал значений a, для которых между (4) и (6) наблюдается удовлетворительное согласие.

Окончательное закрытие вопроса о первом чемодане

Мы уже много знаем о том, каким по счету появится первый чемодан: знаем, математическое ожидание этой величины, ее дисперсию; знаем, насколько вероятно то, что он не появится среди первых k чемоданов, и даже знаем, когда следует заподозрить пропажу. Обидно не решить задачу о первом чемодане до конца.

Задача 5. Найти распределение случайной величины X_1.

Чтобы случилось событие $X_1 = k$, где $k = 1, 2, \ldots, n - a + 1$, нужно, чтобы на позицию k попал один из a чемоданов Петровых (вероятность этого $\frac{a}{n}$) и чтобы при этом условии среди первых $(k - 1)$ чемоданов не оказалось их чемоданов (условная вероятность этого $\frac{C_{n-a}^{k-1}}{C_{n-1}^{k-1}}$).

Перемножив эти вероятности, находим:

$\mathrm{P}(X_1 = k) = \frac{aC_{n-a}^{k-1}}{nC_{n-1}^{k-1}}$, или, что то же самое $\mathrm{P}(X_1 = k) = \frac{aC_{n-a}^{k-1}}{kC_n^k}$.

Параметрами распределения являются числа n и a. Построим распределение случайной величины X_1 при $n = 200$, $a = 4$ (рис. 5). Даже больше: построим универсальное распределение величины X_j для j-го чемодана Петровых при $j = 1, \ldots, a$. Параметр j указан в ячейке C5 (рис. 6).

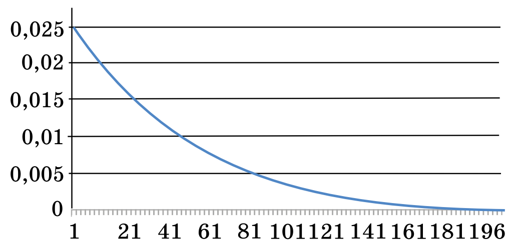
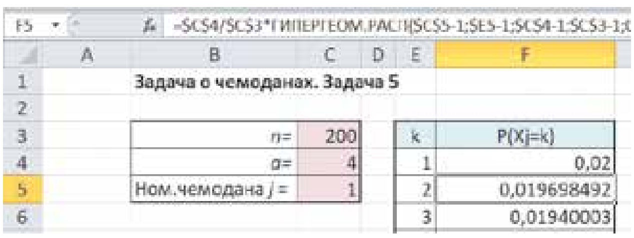


Рис. 5. Диаграмма распределения X_1 при $a = 4$ и $n = 200$



F5 =$C$4/$C$3*ГИПЕРГЕОМ.РАСП($C$5-1;$E5-1;$C$4-1;$C$3-1;...

	A	B	C	D	E	F
1		Задача о чемоданах. Задача 5				
2						
3		n=	200		k	P(Xj=k)
4		a=	4		1	0,02
5		Ном.чемодана j =	1		2	0,019698492
6					3	0,01940003

Рис. 6. Расчет распределения X_j при $j = 1$

При взгляде на рисунок 5 возникает подозрение, что последовательность вероятностей $\mathrm{P}(X_1 = k)$ близка к геометрической прогрессии. Конечно, это не геометрическая прогрессия (мы уже упоминали в № 8, что последовательности такого вида называются гипергеометрическими), но отношение последующего члена к предыдущему поначалу — при малых k — действительно меняется медленно:

$$\frac{P(X_1=k+1)}{P(X_1=k)}=\frac{C_{n-a}^{k}C_{n-1}^{k-1}}{C_{n-1}^{k}C_{n-a}^{k-1}}=1-\frac{a-1}{n-k}.$$

Это неудивительно. Ведь мы имеем дело с разновидностью серии испытаний до первого успеха, а от классического случая ситуация отличается только тем, что совокупность чемоданов конечна. Распределение должно быть тем ближе к геометрическому, чем больше чемоданов всего и чем меньше среди них чемоданов Петровых.

При $1 < j < a$ распределение приобретает двускатную форму. Ясно, что распределения величин X_j и X_{a-j+1} симметричны друг другу. Заготовки для расчетов находятся в электронной таблице.

Задачи для самостоятельного решения

1. По свистку учителя физкультуры 16 мальчиков и 14 девочек моментально выстроились в шеренгу в случайном порядке. Какой по счету слева стоит первая девочка? Найдите математическое ожидание и дисперсию этой величины.

2. После последнего экзамена 250 студентов курса сдали зачетные книжки на оформление в деканат в случайном порядке. В деканате книжки не сортировали по группам, а просто проштамповали и сложили в стопки случайным образом. Через день за зачетками пришла староста первой группы, в которой 20 студентов. Среди нескольких зачеток, лежавших сверху, она не нашла ни одной из своей группы. Тогда она стала искать дальше. После какой по счету чужой зачетки староста должна считать, что с вероятностью 0,98 с зачетками ее группы случилось что-то неправильное?

3. В лотерее участвует 49 пронумерованных шаров. Во время тиража шесть выигрышных шаров выпадают из барабана по очереди в случайном порядке. Участник лотереи ставит на шесть номеров заранее. Какова вероятность того, что он не угадает ни одного номера?

4. Найдите функцию вероятностей $P(X_j = k)$ распределения случайной величины X_j «порядковый номер j-го чемодана Петровых на багажной ленте», где $k = j, \ldots, n - a + j$.

Ответы к задачам из статьи «Задача о жребии "на туза"»

1. 0,0689. **2.** 0,198. **3.** Примерно в 1,5 раза.

4. *Доказательство*. Воспользуемся (3):

$$p_{m+1}=\sum_{i=1}^{\frac{n-m-1}{k}} s_{ik+m+1}=s_{m+1}+s_{k+m+1}+s_{2k+m+1}+\ldots+s_{qk+m+1},$$

где q — наибольшее целое, при котором

$$qk+m+1\le n.$$

Из (1) следует, что если $a \ge 2$, то $s_j < s_{j-1}$ для всех натуральных $j \le n-1$. Значит,

$$p_{m+1} < s_m + s_{k+m} + s_{2k+m} + \ldots + s_{qk+m} + s_{(q+1)k+m} = p_m.$$

Последнее слагаемое $s_{(q+1)k+m}$ добавлено на случай, когда $(q+1)k+m=n$. Если это не так, то $(q+1)k+m>n$. Ничего страшного нет, будем считать, что $s_{(q+1)k+m}=0$. Во всяком случае, это неотрицательное слагаемое не уменьшает сумму.

5. Пусть в колоде n карт и среди них a тузов. Найдем вероятность $s_{(b),j}$ события «b-й по счету туз окажется j-й картой в колоде» ($b \le a$). Вероятность того, что j-й картой окажется туз, равна $\frac{a}{n}$.

При этом условии среди первых $(j-1)$ карт должно быть $(b-1)$ тузов и $(j-b)$ нетузов. В колоде осталось $(n-1)$ карт, из них можно составить C_{n-1}^{j-1} равновозможных выборок объемом $(j-1)$, из которых ровно $C_{a-1}^{b-1}C_{n-a}^{j-b}$ удовлетворяют нужным условиям. Значит, вероятность того, что среди первых $(j-1)$ карт окажется $(b-1)$ тузов, равна $\frac{C_{a-1}^{b-1}C_{n-a}^{j-b}}{C_{n-1}^{j-1}}$. Умножая это число на $\frac{a}{n}$, получаем равенство (7):

$$s_{(b),j}=\frac{aC_{a-1}^{b-1}C_{n-a}^{j-b}}{nC_{n-1}^{j-1}}.$$

Равенство (8) получается похожим образом. Нужно собрать подходящую комбинацию первых j карт. Для этого нужно взять выборку из b тузов для карт с номерами $i_1, i_2, \ldots, i_{b-1}, j$ и скомбинировать ее с выборкой из $(j-b)$ нетузов для оставшихся $(j-b)$ карт.

Другой способ — разделить (7) на число способов зафиксировать карты для $(b-1)$ тузов среди первых $(j-1)$ карт в колоде: $s_{i_1,i_2,\ldots,i_{b-1},j}=\frac{s_{(b),j}}{C_{j-1}^{b-1}}$. Преобразования приводят к (8).

6. *Доказательство*. Чтобы первого туза получил m-й игрок, нужно, чтобы $(m-1)$ игроков перед ним не получили туза при первой раздаче. Вероятность этого гипергеометрическая:

$$\underbrace{\frac{n-a}{n}\cdot\frac{n-a-1}{n-1}\cdot\ldots\cdot\frac{n-a-m+2}{n-m+2}}_{(m-1)\text{ множителей}}=\frac{C_{n-a}^{m-1}}{C_{n}^{m-1}}.$$

Теперь m-й игрок становится первым при раздаче колоды, в которой осталось $(n-m+1)$ карт. Вероятность того, что именно он получит туза, равна $p_{1,(n-m+1)}$. Перемножая эти вероятности, приходим к требуемому равенству.

7. 0,0528.

Комментарий. Ссылки на формулы (1)-(8) в ответах к задачам из статьи «Задачи о жребии "на туза"» см. в № 8 за 2023 г.