



Рис. 1

Задача 7

И. ВЫСОЦКИЙ,
г. Москва

ЗАДАЧИ С УЛИЦЫ

Задача об узком шоссе

По очень узкому и очень длинному шоссе один за другим едут автомобили. Автомобилей много, и каждый водитель хочет ехать со своей, удобной ему, скоростью. Двух совершенно одинаковых скоростей нет, поэтому через некоторое время те, кто едет побыстрее, догоняют тех, кто едет помедленнее, и машины сбиваются в группы (рис. 1). Будем называть эти группы *вереницами*. Можно считать, что автомобилей n штук и что шоссе достаточно длинное для того, чтобы вереницы полностью сформировались.

Сколько образуется верениц? Разумеется, речь идет о математическом ожидании и о дисперсии этой случайной величины.

Задача должна представить интерес для служб организации движения, поскольку решение дает ясный ответ на вопрос о числе обгонных полос, которые нужно устроить на не очень оживленных однополосных шоссе. Другое приложение — организация эвакуационных путей, устройство систем хранения и транспортировки грузов на автоматизированных складах и т.п.

Несмотря на кажущуюся нехватку данных, задача даже проще тех, что мы обсуждали в предыдущих статьях. Поэтому мы поставим сразу несколько вопросов, но изучим только случайную величину «*общее число верениц*». Число автомобилей в веренице и число верениц определенной длины оставим на потом.

Среднее и дисперсия числа верениц

Мы любим задачи в некоторой сюжетной интерпретации. Поясню: гораздо приятнее решать задачу про пешехода, чем про закон движения материальной точки; про землекопов, чем про абстрактный автомат, обладающий заданной производительностью непонятного труда. Так мы поступали в задаче коллекционера и в задаче о такси: вместо безликой случайной последовательности букв конечного алфавита мы говорили о коллекции принцесс или о повторяющейся машине такси. Более того, даже если речь идет о чисто математическом объекте, ради образности мы стараемся придумать привлекательный сюжет.



Есть дополнительные материалы на сайте raum.math.ru.

Здесь мы поступим наоборот. Формализуем задачу о шоссе до конца, то есть сведем ее к изучению вероятностных свойств сугубо математического объекта — случайной перестановки. Вместо автомобилей у нас будут натуральные числа $1, 2, \dots, n$, а вместо их скоростей — перестановка тех же чисел в случайном порядке. Так можно сделать, поскольку нам важны не сами скорости, а только какая из них выше или ниже. Например, если автомобилей три, а их водители хотят ехать со скоростями 70, 60 и 80 км/ч, то мы получаем перестановку¹

$$\begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}.$$

Здесь самый медленный (скорость 1) едет вторым, самый быстрый — третьим. В этом случае образуется ровно две вереницы. Первая состоит из одного-единственного автомобиля № 1, поскольку автомобиль № 2 отстанет, двигаясь с меньшей скоростью. Он возглавит вторую вереницу, потому что автомобиль № 3 догонит машину № 2, упрется в нее и будет ехать за ней со скоростью 60 км/ч.

Для большей ясности дадим еще один пример с $n = 8$ автомобилями и обведем рамками вереницы² во второй строке:

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ \boxed{5} & \boxed{4} & \boxed{3} & \boxed{8} & \boxed{6} & \boxed{1} & \boxed{2} & \boxed{7} \end{pmatrix}. \quad (1)$$

Первые элементы в вереницах образуют убывающую подпоследовательность в перестановке, которая получается по простому закону. В нее входит первый элемент, а затем каждый элемент, который меньше, чем все предыдущие. Это хорошо видно, если изобразить перестановку не таблицей, а диаграммой (рис. 2).

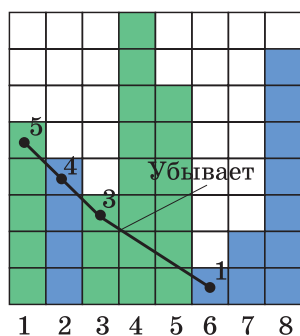


Рис. 2. Представление перестановки (1). Здесь четыре вереницы: две зеленые и две синие. Лидеры — числа 5, 4, 3 и 1.

¹ Напомним, что перестановка порядка (длины) n — это взаимно однозначное отображение множества натуральных чисел от 1 до n на себя. Или иначе — расположение этих чисел в произвольном порядке.

² Распространим термин «вереница» и на перестановки тоже. В англоязычной литературе встречается слово *gip*, которое можно перевести как угодно.

Рассмотрение перестановки на рисунке 2 и краткое размышление в общем виде приводят к двум очевидным выводам.

1. Верениц столько, сколько *лидеров*, то есть чисел в убывающей подпоследовательности, возглавляющих вереницы (в приведенном примере это числа 5, 4, 3 и 1), а потому можно пересчитывать не сами вереницы, а только лидеров. Слово *лидер* мы позаимствовали у автомобильной интерпретации.

2. Лидером является число, которое меньше всех предыдущих. В терминах автомобильных это значит, что возглавляет вереницу тот, чья скорость ниже, чем скорость любого едущего перед ним.

Итак, задача сводится к перечислению лидеров. Обозначим их количество (то есть число верениц) X . Это случайная величина, математическое ожидание и дисперсию которой мы хотим найти.

Введем индикаторы I_k , где $k = 1, \dots, n$. Индикатор I_k равен 1 или 0 в зависимости от того, осуществилось или нет событие A_k « k -й элемент перестановки — лидер». Очевидно, математическое ожидание первого индикатора равно

$$EI_1 = P(A_1) = 1,$$

поскольку первое число является лидером в любой перестановке.

Второй индикатор равен единице не всегда, а только если второе число меньше первого. Вероятность этого $\frac{1}{2}$. Поэтому

$$EI_2 = P(A_2) = \frac{1}{2}.$$

Рассуждая таким же образом дальше, найдем, что

$$EI_k = P(A_k) = \frac{1}{k}$$

для всех k . Переходя в равенстве

$$X = I_1 + I_2 + \dots + I_n \quad (2)$$

к математическим ожиданиям, получаем:

$$EX = 1 + \frac{1}{2} + \frac{1}{3} + \dots + \frac{1}{n} = H_n. \quad (3)$$

Символ H_n означает n -е гармоническое число. В статье о задаче коллекционера (журнал «Математика», 2022, № 9) мы показали, что

$$\ln n < H_n < \ln n + 1$$

при $n > 1$. На самом деле верно равенство

$$H_n = \ln n + \gamma_n,$$

где $\lim_{n \rightarrow \infty} \gamma_n = \gamma = 0,577\dots$ Число γ называется константой Эйлера–Маскерони. При больших n можно считать, что

$$EX \approx \ln n + 0,577. \quad (4)$$

Перейдем к вопросу о дисперсии. Найдя ее, мы сможем понять, насколько сильно в реальности число верениц в перестановке может отличаться от своего среднего.

D104		f _к = D103+(B104-1)/B104/B104								
	A	B	C	D	E	F	G	H	I	J
1	Задача об узком шоссе									
2										
3			Реш.правило k=	3						
4										
5		Точный расчет						Прибл. формулы		
6		n=	EX=H_n	DX=H_n-H_2,n	Ст.откл	Лев.гр	Прав.гр	EX=lnn+0,577	DX=lnn-1,068	
103		97	5,157072	3,522395	1,876804	-0,473341	10,787486	5,151711	3,506711	
104		98	5,167277	3,532495	1,879493	-0,471203	10,805756	5,161967	3,516967	
105		99	5,177378	3,542494	1,882151	-0,469076	10,823831	5,172120	3,527120	
106		100	5,187378	3,552394	1,884779	-0,466961	10,841716	5,182170	3,537170	
107		101	5,197279	3,562197	1,887378	-0,464856	10,859413	5,192121	3,547121	
108		102	5,207082	3,571904	1,889948	-0,462762	10,876927	5,201973	3,556973	
109		103	5,216791	3,581519	1,892490	-0,460679	10,894262	5,211729	3,566729	
110		104	5,226407	3,591042	1,895004	-0,458607	10,911420	5,221391	3,576391	
111		105	5,235930	3,600475	1,897492	-0,456545	10,928406	5,230960	3,585960	

Рис. 3

К счастью, индикаторы в сумме (2) независимы, поскольку k -й индикатор зависит только от того, меньше k -е число всех предыдущих или нет, но не от порядка предыдущих чисел. Поэтому дисперсию X можно найти как сумму дисперсий всех индикаторов³. Значит, нужно найти дисперсию каждого индикатора. Индикатор I_k имеет распределение

$$I_k \sim \begin{pmatrix} 0 & 1 \\ \frac{k-1}{k} & \frac{1}{k} \end{pmatrix}.$$

Поэтому дисперсия индикатора равна

$$DI_k = EI_k^2 - E^2 I_k = EI_k - E^2 I_k = \frac{1}{k} - \frac{1}{k^2} = \frac{k-1}{k^2}.$$

Следовательно,

$$DX = 1 + \frac{1}{2} + \frac{1}{3} + \dots + \frac{1}{n} - \left(1 + \frac{1}{4} + \frac{1}{9} + \dots + \frac{1}{n^2} \right) = H_n - H_{n,2}. \quad (5)$$

Напомним, что символом $H_{n,2}$ обозначается сумма обратных квадратов от 1 до $\frac{1}{n^2}$ и это число называется *гармоническим числом второго порядка*. Известно⁴, что $H_{n,2} \rightarrow \frac{\pi^2}{6}$ при $n \rightarrow \infty$.

Таким образом, при больших n для вычисления нужд можно смело считать, что дисперсия числа верениц равна

$$\ln n + \gamma - \frac{\pi^2}{6} = \ln n - 1,068... \quad (6)$$

Выбирая в качестве решающего правила три стандартных отклонения, можно считать, что

³ Используется теорема о том, что дисперсия суммы независимых случайных величин равна сумме их дисперсий.

⁴ Этот факт мы упоминали в статье о задаче коллекционера, но не доказывали, сославшись на то, что доказательство сугубо не школьное. Но поскольку факт встретился уже дважды, автор обещает разместить почти честное доказательство в конце этой статьи, если, конечно, ему не будет лень.

с весьма высокой вероятностью число автомобилей верениц на шоссе будет находиться в пределах

$$\begin{aligned} &\text{от } \ln n - 3\sqrt{\ln n - 1,068} + 0,577 \\ &\text{до } \ln n + 3\sqrt{\ln n - 1,068} + 0,577. \end{aligned} \quad (7)$$

Остается заметить, что приближения по формулам (4) и (6) работают удовлетворительно (скажем, до второго знака после запятой), начиная где-то с $n = 350$. Сходимость довольно медленная, в чем, разумеется, виновата неторопливость логарифма.

Настало время запустить электронную таблицу MS Excel и посмотреть на результаты наших трудов. Мы возьмем $n \leq 500$, проведем точные вычисления по формулам (3) и (5), а приближенные вычисления по формулам (4) и (6) поместим в дополнительных столбцах I и J просто для сравнения.

На рисунке 3 видно, что, например, при $n = 100$, то есть если автомобилей 100, число верениц будет, вероятно, находиться в пределах от $-0,467$ до $10,842$, то есть в реальности от 1 до 10. При этом среднее число верениц равно $5,187$. Поэтому на участке шоссе, где в среднем одновременно находится 100 автомобилей, разумно иметь хотя бы 5–6 расширений с обгонными полосами.

Распределение числа верениц

Дав ответ на вопрос о среднем и о дисперсии числа верениц, мы вправе замахнуться на большее: а вдруг удастся найти вероятность того, что количество верениц X окажется равно ровно k , где $k = 1, \dots, n$, то есть найти распределение случайной величины X .

Как упоминалось при обсуждении задачи о совпадении («Математика», 2023, № 1), найти распределение случайной величины часто сложнее, чем его математическое ожидание или дисперсию. Здесь тот самый случай. Но все же

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1	Числа Стирлинга 1 рода														
2															
3		n\k	1	2	3	4	5	6	7	8	9	10	11	12	
4		1	1												
5		2	1	1											
6		3	2	=D5+(\$B6-1)*E5											
7		4	6	11	6	1									
8		5	24	50	35	10	1								
9		6	120	274	225	85	15	1							
10		7	720	1764	1624	735	175	21	1						
11		8	5040	13068	13132	6769	1960	322	28	1					
12		9	40320	109584	118124	67284	22449	4536	546	36	1				
13		10	362880	1026576	1172700	723680	269325	63273	9450	870	45	1			
14		11	3628800	10628640	12753576	8409500	3416930	902055	157773	18150	1320	55	1		
15		12	39916800	120543840	150917976	105258076	45995730	13339535	2637558	357423	32670	1925	66	1	

Рис. 4. Таблица чисел $c(n, k)$ (числа Стирлинга 1-го рода)

это можно сделать, и даже не перенапрячь при этом наш усталый мозг.

Обозначим $c(n; k)$ число перестановок порядка n , содержащих ровно k верениц. Очевидно,

$$P(X = k) = \frac{c(n; k)}{n!},$$

так что вопрос сводится к поиску чисел $c(n; k)$.

Попробуем рекурсию, то есть выразим число $c(n; k)$ через такие же числа c , но с меньшими индексами. Рассмотрим отдельно разбиения на вереницы, где элемент n — лидер и где n — не лидер.

Если элемент n — лидер, то он образует отдельную первую вереницу. За ней следует разбиение перестановки $(1, 2, \dots, n-1)$ на $(k-1)$ верениц, а таких разбиений $c(n-1; k-1)$.

Если элемент n — не лидер, то нужно взять любое из $c(n-1; k)$ разбиений перестановки $(1, 2, \dots, n-1)$ на k верениц и «подселить» в любую из верениц число n , только не на первую позицию, а после любого из уже стоящих чисел, которых всего $(n-1)$. Получается $(n-1)c(n-1; k)$ способов получить k верениц, где n — не лидер.

Объединяя эти варианты, получаем равенство $c(n; k) = c(n-1; k-1) + (n-1)c(n-1; k)$ (8) при $n \geq 2$. Для начала рекурсии нужно взять $c(n; 0) = 0$ при любом n и $c(1; 1) = 1$. Первое — поскольку перестановок без верениц не бывает, а второе — поскольку перестановка длины 1 одна-единственная и содержит ровно 1 вереницу⁵.

Таблицу чисел $c(n; k)$ построим, пользуясь MS Excel и формулой (8) (рис. 4).

У автора уже нет ни сил, ни желания скрывать, что числа $c(n; k)$ известны как *числа Стирлинга первого рода*⁶ и хорошо изучены. Они

⁵ Иногда для общности и удобства считают, что $c(0; 0) = 1$.

⁶ Наряду с обозначением $c(n; k)$ для чисел Стирлинга 1-го рода часто встречается обозначение $[n, k]$. Автор всегда думал, что любители подобных обозначений сознательно стараются израсходовать побольше бумаги.

обычно определяются иначе — не через вереницы, но об этом мы поговорим позже, когда будем решать задачу о циклах в перестановках.

Ясно, что числа в первом столбце равны факториалам:

$$c(n; 1) = (n-1)!.$$

Ясно также, что сумма чисел в каждой строке равна $n!$. Эти практически очевидные равенства мы не обсуждаем здесь, но вынесем в задачи для самостоятельного решения.

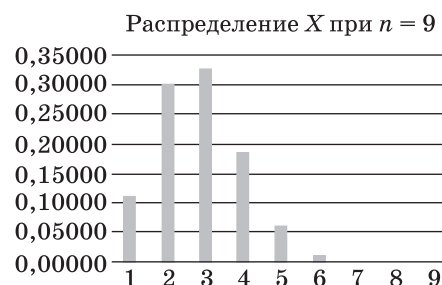
Имея таблицу чисел Стирлинга 1-го рода, легко построить таблицу вероятностей $P(X = k)$. Достаточно каждое число разделить на $n!$ при соответствующем n (рис. 5).

Итак, вероятность того, что случайная перестановка длины n содержит ровно k верениц, равна

$$P(X = k) = \frac{c(n; k)}{n!}. \quad (9)$$

Это равенство и таблица чисел $c(n; k)$ потребуются нам не только здесь, но и в других задачах, связанных с перестановками, например, в задаче о циклах.

Для наглядности построим диаграмму распределения при каком-нибудь n . Скажем, при $n = 9$ (рис. 6). Вероятности сначала растут, но затем стремительно падают и при $k > 7$ неотличимы от 0.

Рис. 6. Диаграмма распределения случайной величины X при $n = 9$

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1	Распределение числа верениц X														
2															
3		n\k	1	2	3	4	5	6	7	8	9	10	11	12	
4		1	1,00000												
5		2	0,50000	0,50000											
6		3	0,33333	=Числа Стирлинга 1!E6/ФАКТР(Вероятности\$B6)											
7		4	0,25000	0,45833	0,25000	0,04167									
8		5	0,20000	0,41667	0,29167	0,08333	0,00833								
9		6	0,16667	0,38056	0,31250	0,11806	0,02083	0,00139							
10		7	0,14286	0,35000	0,32222	0,14583	0,03472	0,00417	0,00020						
11		8	0,12500	0,32411	0,32569	0,16788	0,04861	0,00799	0,00069	0,00002					
12		9	0,11111	0,30198	0,32552	0,18542	0,06186	0,01250	0,00150	0,00010	0,00000				
13		10	0,10000	0,28290	0,32316	0,19943	0,07422	0,01744	0,00260	0,00024	0,00001	0,00000			
14		11	0,09091	0,26627	0,31950	0,21068	0,08560	0,02260	0,00395	0,00045	0,00003	0,00000	0,00000		
15		12	0,08333	0,25166	0,31507	0,21974	0,09602	0,02785	0,00551	0,00075	0,00007	0,00000	0,00000	0,00000	

Рис. 5. Вероятности $P(X = k) = \frac{c(n; k)}{n!}$. Если на шоссе 10 автомобилей, то вероятность того, что они образуют ровно 3 вереницы, равна 0,32316

Как обычно, все использованные электронные таблицы и диаграмма размещены на сайте <http://ptlab.mcsme.ru> на странице «Задачи с улицы».

Еще несколько задач в этой связи

Сформулируем еще три довольно естественные задачи, связанные с вереницами.

Задача 1. Какова вероятность того, что в случайной перестановке отдельная вереница (скажем, первая или вторая) имеет определенную длину? Поставив себя на место водителя уазика на рисунке 1, можно задать себе вопрос о вероятности того, что за мной в хвосте плетутся еще ровно 3 автомобиля или ровно 5 автомобилей. Каково математическое ожидание случайной величины «длина одной конкретной вереницы»?

Мы не будем решать эту задачу здесь и сейчас, но вернемся к ней, когда будем обсуждать циклы или орбиты в случайных перестановках в одной из следующих статей.

Задача 2. Каково наиболее вероятное число верениц? На диаграмме (см. рис. 6) при $n = 9$ оно равно 3. Что будет в общем случае, зависит от того, при каком k число Стирлинга $c(n; k)$ наибольшее.

Автор не знает ответа. Вероятно, с ростом n наиболее вероятное значение растет медленно, примерно как математическое ожидание EX , то есть близко к $\ln n$. Эту гипотезу можно опровергнуть или найти ей косвенное подтверждение с помощью численного моделирования.

Задача 3. Какой по счету элемент перестановки с наибольшей вероятностью попадает в предпоследнюю вереницу?

Это очень интересная и даже знаменитая задача, которую мы обсудим тоже не сейчас.

Обещанное дополнение

Равенство

$$\lim_{n \rightarrow \infty} H_{n,2} = 1 + \frac{1}{4} + \frac{1}{9} + \frac{1}{16} + \dots = \frac{\pi^2}{6}.$$

Доказательство. Разложим синус в ряд Тейлора–Маклорена:

$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \dots \quad (10)$$

Справа мы видим степенной ряд, то есть многочлен бесконечной степени. По аналогии с обычными многочленами он должен раскладываться на множители вида $(x - a_k)$, где a_k — корни, то есть точки, в которых функция $\sin x$ обращается в 0. Эти точки нам хорошо известны: $a_0 = 0$ и $a_k = \pm \pi k$, где $k \in \mathbb{N}$. Все множители, кроме $x - 0 = x$, можно нормировать, представив в виде $1 - \frac{x}{a_k}$.

Значит, должно быть верно равенство, в правой части которого бесконечное произведение:

$$\sin x = Cx \left(1 - \frac{x}{\pi}\right) \left(1 + \frac{x}{\pi}\right) \left(1 - \frac{x}{2\pi}\right) \left(1 + \frac{x}{2\pi}\right) \dots \dots \left(1 - \frac{x}{k\pi}\right) \left(1 + \frac{x}{k\pi}\right) \dots, \quad (11)$$

то есть

$$\sin x = Cx \left(1 - \frac{x^2}{\pi^2}\right) \left(1 - \frac{x^2}{4\pi^2}\right) \dots \left(1 - \frac{x^2}{k^2\pi^2}\right) \dots$$

Раскроем скобки. При больших степенях x коэффициенты устроены сложно, но нам нужны не они, а только коэффициенты при x и x^3 :

$$\sin x = Cx - C \frac{1}{\pi^2} \left(\frac{1}{1} + \frac{1}{4} + \frac{1}{9} + \dots\right) x^3 + A. \quad (12)$$

Слагаемое A содержит все степени x , начиная с пятой, которые нам не нужны. Достаточно того, что тождества (10) и (12) по-разному представляют одну и ту же функцию, и потому можно приравнять коэффициенты при x и x^3 :

$$C = 1 \text{ и } C \frac{1}{\pi^2} \left(1 + \frac{1}{4} + \frac{1}{9} + \dots \right) = \frac{1}{3!},$$

откуда

$$1 + \frac{1}{4} + \frac{1}{9} + \frac{1}{16} + \dots = \frac{\pi^2}{6}.$$

Разумеется, в этом рассуждении есть логический пробел. Но оно показывает, что если синус представим в виде (11), а бесконечная сумма $1 + \frac{1}{4} + \frac{1}{9} + \frac{1}{16} + \dots$ существует, то она обязана

равняться $\frac{\pi^2}{6}$. Аккуратное доказательство того,

что эта сумма существует, является простым упражнением по математическому анализу. Позвольте его здесь не приводить.

Задачи для самостоятельного решения

- Предположим, что на шоссе 50 автомобилей.
 - Найдите среднее число образовавшихся верениц.
 - Оцените интервалом число верениц, используя правило трех стандартных отклонений.
- а) Восстановите числа Стирлинга, которых не видно на рисунке 4, в ячейках Е6 и F6.
 - Восстановите соответствующие закрытые вероятности в таблице на рисунке 5.
- Какова вероятность того, что 13 машин на длинной дороге сойдутся ровно в 5 верениц?
- Докажите, что $c(n; 1) = (n-1)!$.
- Докажите, что сумма чисел Стирлинга $c(n; k)$, взятая по всем k от 1 до n , равна $n!$.

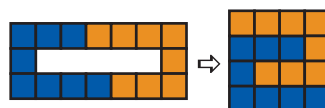
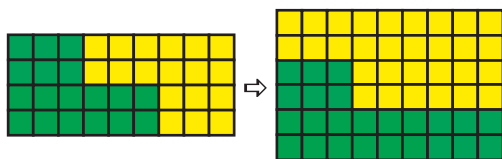
Ответы к задачам из статьи «Задача о совпадениях»

- $\frac{1}{56} \approx 0,018$.
- а) $\frac{2}{3}$; б) $\frac{n-1}{n}$.
- а) $\frac{7}{13} \approx 0,538$;
- а) $\frac{n-1}{2}$.

Продолжение 4. Начало на с. 20.

Тема: «Сравнение натуральных чисел»

№ 4	1	2	3	4
А	2011 и 2010	856 и 658	500 и 3003	4412 и 4421
Б	11 648 и 9999	1892 и 989	88 898 и 89 898	70 088 и 70 885
В	2053 и 1988	56 567 и 6657	12 112 и 12 121	5555 и 50 151
Г	36 300 и 3630	9723 и 2378	37 444 и 73 444	4507 и 4570
Д	5065 и 6056	10 004 и 590	2070 и 4050	92 415 и 92 416
Е	9187 и 12 435	6909 и 6099	5101 и 4901	804 и 488
Ж	70 140 и 70 142	5114 и 4225	33 333 и 3333	66 555 и 65 556
З	48057 и 84057	4012 и 4008	80 808 и 80 090	10 524 и 10 254
И	5060 и 50600	11 001 и 999	1001 и 909	7198 и 7189



Тема: «Округление натуральных чисел»

№ 5	1	2	3	4	5
А	70 834	856	5003	1888	56 052
Б	60 962	7892	8 801 124	88 885	81 534
В	182 057	665 597	46 012	5999	5717
Г	36 309	2378	11 234	4507	6 255 132
Д	551 904	500 704	450 270	900 241	540 081
Е	30 988	69 096	5 103 401	80 435	85 900 020
Ж	77 142	5445	333 333	3169	755 470
З	405 717	89 882	874 001	60 754	51 111 969
И	50 993	8263	1 400 000	700 960	5 205 025

Окончание на сайте raum.math.ru.