И. ВЫСОЦКИЙ,
i_r_vysotsky@hotmail.com,
г. Москва

Рисунки Николая Крутикова

ДВЕ ЗАДАЧИ ИЗ УЧИТЕЛЬСКОЙ ОЛИМПИАДЫ

Последние два года Департамент образования Москвы проводит многопредметную олимпиаду для учителей «Московский учитель», в подготовке заданий для которой принимают участие члены ассоциации учителей математики г. Москвы. В математическом разделе олимпиады оба года представлена теория вероятностей. В этой статье мы с вами решим задачи по вероятности обеих олимпиад и обсудим некоторые их обобщения и развитие.
Материалы этой статьи можно полностью или частично использовать при планировании уроков по вероятности в 9–11-х общеобразовательных и специализированных классах.

Для решения обеих задач требуются только общие сведения о серии испытаний Бернулли:

> Если проводится n одинаковых и независимых испытаний с вероятностью успеха p в каждом испытании, то вероятность ровно k успехов вычисляется по формуле
> $$C_n^k p^k (1-p)^k,$$
> а математическое ожидание числа успехов равно np.

Задача 1 (2016). Стрелок стреляет по 10 мишеням. У него 20 патронов, и на каждую мишень стрелку дается не более двух попыток. Вероятность попадания при каждом отдельном выстреле 0,7. Известно, что стрелок израсходовал 17 патронов. Найдите математическое ожидание числа пораженных мишеней.

Решение. По каждой мишени стрелок сделал хотя бы один выстрел. Итого, 10 выстрелов. Еще 7 выстрелов пришлось сделать по тем мишеням, по которым стрелок промахнулся при первой попытке.

Схематически ситуация показана на рисунке 1. Пустые кружки — удачный выстрел, кружки с крестиком — промах. Для простоты три мишени, пораженные сразу, изображены слева.

	1	2	3	4	5	6	7	8	9	10
1-я попытка. Сбито три мишени	○	○	○	⊗	⊗	⊗	⊗	⊗	⊗	⊗
2-я попытка. Сбито Y мишеней				○	⊗	○	⊗	⊗	○	○

Рис. 1

Обозначим X общее число сбитых мишеней, то есть удачных выстрелов: $X = 3 + Y$,
где Y — число мишеней, сбитых со второй попытки. Вторые попытки (всего их 7) образуют серию испытаний Бернулли с вероятностью успеха 0,7 в каждом. Значит,
$$\mathrm{E}Y = 7 \cdot 0{,}7 = 4{,}9.$$
Поэтому
$$\mathrm{E}X = 3 + \mathrm{E}Y = 3 + 4{,}9 = 7{,}9.$$

Задача 2 (2015). В день города на свет появилось 40 младенцев. Мэр города решил каждому младенцу вручить медаль Юного Горожанина. Мальчикам полагается медаль на голубой ленточке, а девочкам — на розовой. Помощнице мэра велели купить 40 ленточек, и только тут сообразили, что забыли уточнить, сколько каких ленточек нужно. Поразмыслив, помощница купила 20 голубых ленточек и 20 розовых. Считая, что рождение мальчика и рождение девочки равновозможны, найдите наиболее вероятное число младенцев, которым достанется ленточка «не того цвета».

Решение. Решим задачу в общем виде, считая, что младенцев не 40, а $2n$. Число «обиженных» обозначим X, а число мальчиков — Y.

Обиженных нет в одном-единственном случае: если $Y = n$. Вероятность этого:

$$\mathrm{P}(X=0)=C_{2n}^{n}\cdot 0{,}5^{2n}.$$

Будет ровно k обиженных (при $1 \leqslant k \leqslant n$), если

$$Y = n - k \text{ или } Y = n + k,$$

то есть

$$\mathrm{P}(X = k) = \mathrm{P}(Y = n - k) + \mathrm{P}(Y = n + k) =$$

$$= C_{2n}^{n-k}\cdot 0{,}5^{2n} + C_{2n}^{n+k}\cdot 0{,}5^{2n} = 2C_{2n}^{n-k}\cdot 0{,}5^{2n}.$$

Следовательно, задача сводится к поиску наибольшего из чисел

$$2C_{2n}^{0},\ 2C_{2n}^{1},\ 2C_{2n}^{2},\ \ldots,\ 2C_{2n}^{n-1},\ C_{2n}^{n}.$$

Очевидно,

$$C_{2n}^{0} < C_{2n}^{1} < 2C_{2n}^{2} < \ldots < C_{2n}^{n-1}.$$

Осталось сравнить $2C_{2n}^{n-1}$ и C_{2n}^{n}:

$$\frac{2C_{2n}^{n-1}}{C_{2n}^{n}} = \frac{2\cdot(2n)!\cdot n!\cdot n!}{(n-1)!\cdot(n+1)!\cdot(2n)!} = \frac{2n}{n+1} > 1$$

(если, конечно, $n > 1$).

Следовательно, $2C_{2n}^{n-1} > C_{2n}^{n}$. Поэтому самое вероятное, что будет ровно один младенец, «обиженный» не своим цветом ленточки, независимо от общего числа младенцев. Результат неожиданный.

Задачу можно решить, исходя из графических соображений: «согнем пополам» биномиальное распределение числа мальчиков (рис. 2) и получим распределение числа обиженных (рис. 3): средний столбик не изменяется, но становится столбиком 0 (обиженных нет). Два любых симметричных друг другу столбика «накладываются друг на друга», образуя столбик удвоенной высоты.

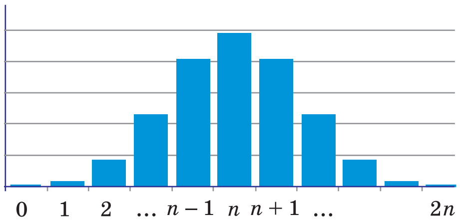

Рис. 2. Распределение Y

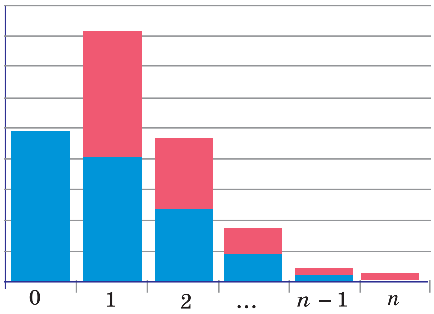

Рис. 3. Распределение X

На рисунке 3 видно, что наибольшая вероятность у значения 1. Такое графическое решение дали многие участники олимпиады.

Ответ: 1.

Обсуждение задачи про стрелка

У внимательного читателя должен был возникнуть вопрос, а откуда известна вероятность попадания при каждом выстреле? Более естественна постановка задачи, когда известно число пораженных мишеней или число потраченных патронов, но не вероятность попадания. И как раз эту неизвестную вероятность хорошо бы оценить.

При этом оценка неизвестной вероятности зависит не только от данных задачи, но и от метода оценивания. Итак, первая задача.

Задача 3. Стрелок стреляет по 10 мишеням. У него 20 патронов, и на каждую мишень стрелку дается не более двух попыток. Известно, что стрелок израсходовал 17 патронов. Оцените вероятность попадания при каждом выстреле. (Метод оценивания мы не указали, оставляя тем самым некоторую свободу выбора.)

Решение. Сперва следует выбрать метод оценивания — от этого зависят дальнейшие действия.

Будем использовать метод моментов, который состоит в том, что математическое ожидание случайной величины, зависящее от некоторого параметра (в нашем случае — вероятности), приравнивается к математическому ожиданию этой случайной величины. Математическое ожидание называют первым моментом, отсюда название метода. Обоснование метода использует тот факт, что случайная величина с конечной дисперсией не может очень далеко отклоняться от своего математического ожидания.

В нашем случае ожидание случайной величины Z числа неизрасходованных патронов приравняем к ее наблюдаемому значению: $20 - 17 = 3$.

Обозначим неизвестную вероятность p. Стрелок «сэкономил» столько патронов из своих двадцати, сколько у него было успешных первых попыток. Поэтому

$$\mathrm{E}Z = 10p$$

($10p$ — это математическое ожидание числа попаданий при первых попытках). Метод моментов дает

$$10\hat{p} = 20 - 17, \quad \hat{p} = 0{,}3.$$

Заметим, что мы нашли не саму неизвестную вероятность, а лишь некоторую ее оценку. Чтобы отличать оценку $\hat{p}$ от самой вероятности p, мы ставим над буквой «крышку».

Похоже, что нас с самого начала ввели в заблуждение по поводу меткости стрелка: вероятность попадания 0,7 плохо сочетается с тем, что потрачено 17 патронов.

Примечание. Если бы мы использовали не метод моментов, а какой-нибудь другой, то результаты, вообще говоря, могли бы быть другими. Важно понимать, что оценка любой величины неотделима от метода, которым эта оценка сделана.

Решим задачу с другими данными. Пусть теперь нам известно число пораженных мишеней.

Задача 4. Стрелок стреляет по 10 мишеням. У него 20 патронов, и на каждую мишень стрелку дается не более двух попыток. Известно, что стрелок поразил 5 мишеней. Методом моментов оцените вероятность попадания при каждом выстреле.

Решение. Снова обозначим неизвестную вероятность p и выразим через нее математическое ожидание числа пораженных мишеней X. Каждая пораженная мишень поражена либо при первой попытке, либо при второй, если первая неудачна. Поэтому вероятность того, что каждая отдельная мишень поражена, равна

$$p + (1 - p)p = 2p - p^2.$$

Всего мишеней 10, поэтому

$$\mathrm{E}X = 10(2p - p^2).$$

Метод моментов дает:

$$10\left(2\hat{p} - \hat{p}^2\right) = 5,$$

откуда

$$2\hat{p}^2 - 4\hat{p} + 1 = 0, \quad \hat{p} = 2 - \sqrt{2} \approx 0{,}586$$

(корень, больший единицы, отбросим сразу).

Напоследок решим задачу, в которой у нас есть вся информация о прошедшей стрельбе.

Задача 5. Стрелок стреляет по 10 мишеням. У него 20 патронов, и на каждую мишень ему дается не более двух попыток. Известно, что стрелок поразил всего 5 мишеней, потратив на это 17 патронов. Оцените вероятность попадания при каждом отдельном выстреле методом наибольшего правдоподобия.

Решение. Из условия следует, что из 10 первых попыток удачными оказались 3, а из 7 вторых попыток — еще 2 попытки. Вероятность такого стечения обстоятельств равна

$$C_{10}^3 p^3 (1 - p)^7 \cdot C_7^2 p^2 (1 - p)^5.$$

Используем метод наибольшего правдоподобия, который состоит в том, чтобы считать вероятность наступившего события наибольшей возможной. Обычно в обоснование метода приводят соображения, связанные с условной вероятностью. Поступим проще — нет оснований считать осуществившееся событие одним из маловероятных, поскольку в маловероятные события мы не верим. Следовательно, разумно считать случившееся весьма и даже наиболее правдоподобным. Отсюда и название метода, и его суть — оценка подбирается так, чтобы вероятность осуществившегося события была наибольшей.

Иными словами, мы ищем точку $\hat{p}$, в которой функция

$$y = p^3(1 - p)^7 \cdot p^2(1 - p)^5 = p^5(1 - p)^{12}$$

на отрезке $0 \leqslant p \leqslant 1$ принимает наибольшее значение. Это можно сделать с помощью производной, а можно хитро воспользоваться неравенством Коши о средних, согласно которому среднее геометрическое неотрицательных чисел не превосходит их среднего арифметического:

$$\sqrt[n]{a_1 a_2 \ldots a_n} \le \frac{a_1 + a_2 + \ldots + a_n}{n};$$

при этом равенство достигается, только если все числа равны между собой.

Найдем наибольшее значение функции

$$\sqrt[17]{12^5 \cdot 5^{12} y} = \sqrt[17]{(12p)^5 (5(1-p))^{12}},$$

которая принимает наибольшее значение в той же точке, что и функция y.

Неравенство Коши дает

$$\sqrt[17]{(12p)^5 (5(1-p))^{12}} \le \frac{5 \cdot 12p + 12 \cdot 5(1-p)}{17} = \frac{60}{17}.$$

Нам интересно не столько само наибольшее значение y, сколько факт, что оно достигается при обязательном условии

$$12p = 5(1-p),$$

откуда

$$\hat{p} = \frac{5}{17} \approx 0{,}294.$$

Это и есть *оценка наибольшего правдоподобия* неизвестной вероятности p.

Обратите внимание — в трех разных задачах мы искали оценки неизвестной вероятности p и получили три разных числа. Ничего удивительного. В разных задачах была разная исходная информация. Нельзя даже ставить вопрос о том, какая оценка наилучшая: они сделаны в разных условиях. Кстати, если вы решите задачи 3 и 4 методом наибольшего правдоподобия, а задачу 5 — методом моментов, то в каждой задаче оценка совпадет со сделанной нами. Факт интересный и не всеобщий – во многих задачах оценки, сделанные разными методами, различаются.

Строгий и искушенный специалист сейчас скажет, что в наших рассказах нет важных вещей. Например, нет ничего о свойствах оценок. Конечно, хотелось бы, чтобы сделанные оценки $\hat{p}$ были удовлетворительно близки к истинному значению p, да еще и неплохо бы обсудить, в каком смысле близки и что такое «удовлетворительно». Это правда важно, но я думаю, что в школьной науке важнее сделать первый шаг и достичь достижимого, чем безнадежно рассуждать о высоких материях. Важнее решать содержательные задачи на интуитивном уровне, чем копаться в тонких вопросах оценивания, так ничему и не научившись. Всему свое время.

Обсуждение задачи про младенцев

Задача про младенцев больше известна в другой формулировке.

Задача 6. На борту авиалайнера $2n$ пассажиров. Авиакомпания погрузила на борт n порций обеда с курицей и n порций с рыбой. С вероятностью 0,5 пассажир предпочитает рыбу, а с вероятностью 0,5 — курицу. Если пассажиру не хватает того, что он выбрал, пассажир делается недовольным. Каково наиболее вероятное число недовольных пассажиров?

Здесь вместо младенцев — пассажиры, вместо розовых и голубых ленточек — курица и рыба. И мы знаем ответ: наиболее вероятно, что будет единственный недовольный пассажир (скорее всего, он сидит где-то в хвосте самолета).

Интересно, а должны ли мы (или кто-нибудь другой) полагаться на эту величину как на ориентир? Ну, скажем, какой-нибудь менеджер в авиакомпании может сказать: «Отлично, раз он такой один, мы возьмем для него одну запасную курицу (и одну рыбу), и в большинстве случаев все будет в ажуре!» Или так должен рассуждать сам пассажир?

Удивительно даже не то, что мода случайной величины «число недовольных» X не зависит от числа пассажиров. Удивительнее то, что эта величина не отражает среднее, то есть математическое ожидание числа недовольных пассажиров. Вот тут интуиция протестует: каким это образом самое вероятное значение может не быть ожидаемым?

А происходит это из-за несимметричности распределения (см. рис. 3). Математическое ожидание числа недовольных больше единицы, если на борту лайнера хотя бы 8 человек. И оно увеличивается с ростом числа пассажиров, хотя не очень быстро — пропорционально $\sqrt{n}$. Можно показать, что

$$\mathrm{E}X = \frac{nC_{2n}^{n}}{2^{2n}} \approx \sqrt{\frac{n}{\pi}} \approx 0{,}564\sqrt{n}$$

(см., например, решения задач V Заочной олимпиады по теории вероятностей на http://ptlab.mccme.ru/node/1702). Что все это значит с практической точки зрения? И с чьей практической точки зрения?

С точки зрения пассажира, мода важнее. Пассажир рассуждает так: «Недовольных, скорее всего, мало (один или вообще нет). Это утверждение применимо к одному полету (моему), и я вправе рассчитывать, что не окажусь неудачником и получу желаемое. А для этого постараюсь не оказаться в самом хвосте самолета».

С точки зрения руководства авиакомпании, ситуация иная. Службе пассажирских перевозок важнее знать среднее число недовольных — ведь полетов много. Здорово, конечно, что в каждом конкретном случае недовольным будет, скорее всего, кто-то один, но для планирования питания это неважно.

Руководство могло бы рассуждать так: «Мы имеем лайнер, в котором 200 пассажиров. Если запасных порций не брать, то среднее число недовольных на каждом рейсе с 200 пассажирами будет примерно

$$\mathrm{E}X \approx 0{,}564\sqrt{100} = 5{,}64,$$

а стандартное отклонение примерно (см. тот же источник)

$$\sigma = \sqrt{\frac{\pi - 2}{2\pi} \cdot 100} \approx 0{,}426\sqrt{100} = 4{,}26.$$

Теория вероятностей утверждает, что число недовольных пассажиров почти наверняка (с вероятностью больше 0,95) не превысит

$$\mathrm{E}X + 2\sigma \approx 5{,}64 + 8{,}52 = 14{,}16.$$

Значит, нужно взять на борт 30 дополнительных порций — 15 с рыбой и 15 с курицей (точный расчет показывает, что меньше: по 12 каждого вида). Тогда стюардессам придется извиняться менее чем в 5% случаев (что приемлемо в соответствии с политикой компании), и мы создадим имидж компании, заботящейся о своих пассажирах, получим дополнительные очки в рейтинге и поборемся за звание лучшей авиакомпании Восточной Европы или даже мира. А оставшиеся порции съедят особо голодные пассажиры и экипаж».

Не знаю, используют ли крупные российские авиакомпании вероятностные расчеты при планировании пассажирских перевозок (хотелось бы верить). Знаю только, что статью эту пишу в Иркутске, куда летел самолетом одной крупной компании, сидел в предпоследнем ряду и мне не хватило порции с говядиной. Пришлось есть курицу (справедливости ради отмечу, что было вкусно).