Задача про викторину на ЕГЭ по математике

Открытый банк задач ЕГЭ по математике содержит несколько задач про викторину. Первая задача этого набора клонов (номер 508866 в банке) решена Анной Малковой на сайте ЕГЭ-Студии (под номером 10). Мы попробуем рассмотреть общую ситуацию.

Задача. В викторине участвуют n команд. Все команды разной силы, и в каждой встрече выигрывает та команда, которая сильнее. В первом раунде встречаются две случайно выбранные команды. Ничья невозможна. Проигравшая команда выбывает из викторины, а победившая команда играет со следующим случайно выбранным соперником. Известно, что в первых k играх победила команда А. Какова вероятность того, что эта команда выиграет $(k+1)$-й раунд?

Немного о понятии вероятностного пространства. Чтобы в той или иной задаче математически строго говорить о вероятности и чётко осознавать, что именно мы в качестве вероятности вычисляем, необходимо описать *вероятностное пространство* данной задачи. Для этого в общем случае следует сделать три вещи.

1. Нужно ввести *множество элементарных событий* Ω, элементами которого служат исходы эксперимента, выпадения жребия и т. д.

   Что именно будет элементарным событием — определяется конкретной задачей. Если, например, мы подбрасываем монету и нас интересует вероятность выпадения орла, то элементарное событие — это результат отдельного броска (О или Р); в данном случае $\Omega = \{О, Р\}$ и $|\Omega| = 2$. Если же мы задаёмся вопросом, с какой вероятностью выпадут ровно три решки при пяти подбрасываниях, то множество Ω состоит из всевозможных пятибуквенных слов (скажем, ОРООР), составленных из букв О и Р; в данном случае $|\Omega| = 2^5 = 32$. (Напомним, что через $|X|$ обозначается число элементов конечного множества X.)

2. Нужно договориться, что мы считаем *событием*. Вообще, событие — это некоторая совокупность элементарных событий; иными словами, событие A есть некоторое подмножество множества Ω. В ряде ситуаций любое подмножество $A \subset \Omega$ является событием, но так бывает не всегда. Например, в вышеописанном случае пяти подбрасываний монеты любое пятибуквенное слово из букв О и Р служит событием; а вот в задачах на геометрическую вероятность (где вероятность определяется как отношение площадей) событиями могут служить лишь *измеримые* множества (у которых в принципе можно вычислить площадь).

3. Нужно указать, каким образом мы вычисляем вероятности событий, то есть каким образом мы сопоставляем произвольному событию A число $p(A) \in [0; 1]$. Разумеется, вероятность достоверного события Ω должна равняться единице, вероятность невозможного события $\varnothing$ — нулю, а вероятность объединения двух (и даже счётного числа) непересекающихся событий — сумме вероятностей этих событий[1].

Простым примером вероятностного пространства является так называемая *классическая вероятностная модель*. В рамках этой модели вышеперечисленные три пункта конкретизируются следующим образом.

[1]Похоже на площадь, не правда ли? Эта аналогия не случайна. На самом деле и вероятность, и площадь являются частными случаями общего понятия *меры*; оно изучается в разделе математики, который называется *теория меры*. Один знакомый математик любил говорить, что «нет никакой теории вероятностей, есть теория меры».

1. Множество элементарных событий Ω конечно: $\Omega = \{\omega_1, \omega_2, \ldots, \omega_n\}$.

2. Событием является любое подмножество A множества Ω.

3. Все элементарные события ω_i в условиях данной задачи можно считать равновероятными, и им приписывается вероятность $\frac{1}{n} = \frac{1}{|\Omega|}$. Если событие A состоит из m элементарных событий, то вероятность $p(A)$ события A определяется по формуле
$$p(A) = \frac{m}{n} = \frac{|A|}{|\Omega|}.$$

Так, в задаче о вероятности выпадения ровно трёх решек при пяти подбрасываниях работает классическая вероятностная модель. Про множество Ω этой задачи мы уже сказали. Интересующее нас событие A состоит из всех пятибуквенных слов, в которых три буквы Р и две буквы О. Таких слов $C_5^3 = 10$, и потому вероятность события A равна $\frac{C_5^3}{2^5} = \frac{10}{32} = \frac{5}{16}$.

Немного о понятии условной вероятности. Условная вероятность понадобится нам далее при решении задачи. Для лучшего понимания условной вероятности стоит прежде всего разобрать отличный пример из Википедии про две шестигранные кости, в котором данное понятие иллюстрируется весьма наглядным образом.

По определению, условная вероятность события A при условии события B равна
$$p(A|B) = \frac{p(A \cap B)}{p(B)}.$$

Поясним на примере классической вероятностной модели, откуда такая формула берётся. Говоря об условной вероятности A при условии B, мы «сужаем» множество элементарных событий с Ω до B (ведь остальные элементарные события помимо B нас уже не интересуют) и рассматриваем событие $A \cap B$ в новом «суженном» вероятностном пространстве B (в примере из Википедии хорошо видно это сужение с 36 случаев до 10). Тогда
$$p(A|B) = \frac{|A \cap B|}{|B|} = \frac{\frac{|A \cap B|}{|\Omega|}}{\frac{|B|}{|\Omega|}} = \frac{p(A \cap B)}{p(B)}.$$

Решение задачи. Пусть силы команд равны 1, 2, ..., n (чем больше сила команды, тем она сильнее). Опишем множество элементарных событий Ω.

Вначале выберем случайно одну команду; пусть её сила равна m. Затем определим ей по жребию *турнирный путь*, то есть последовательность команд-соперников во всех раундах: в первом раунде команда m играет с командой силой s_1, во втором — с командой s_2, ..., в последнем $(n-1)$-м раунде — с командой s_{n-1}. Разумеется, команда m не обязательно пройдёт намеченный жребием турнирный путь до конца: если окажется, что $m < s_i$, то команда m проигрывает в i-м раунде и выбывает (её турнирный путь на этом прерывается).

Таким образом, турнирный путь команды m можно представить как упорядоченный набор $(m, s_1, s_2, \ldots, s_{n-1})$, который является перестановкой чисел 1, 2, ..., n. Элементарными событиями служат всевозможные турнирные пути всех команд, то есть всевозможные упорядоченные наборы $(m, s_1, s_2, \ldots, s_{n-1})$, то есть просто все перестановки чисел 1, 2, ..., n.

Итак, Ω есть множество всех перестановок чисел 1, 2, ..., n, так что $|\Omega| = n!$. Все эти перестановки, очевидно, равновероятны, так что задача решается в рамках классической вероятностной модели.

Обозначим через V_k следующее событие: «случайно выбранная вначале команда побеждает в k-м раунде». Событие V_k состоит из всех турнирных путей $(m, s_1, s_2, \ldots, s_{n-1})$, для которых выполнены неравенства $m > s_1$, $m > s_2$, ..., $m > s_k$. В задаче требуется найти условную вероятность

$$p(V_{k+1}|V_k) = \frac{p(V_{k+1} \cap V_k)}{p(V_k)}.$$

При этом, очевидно, $V_{k+1} \cap V_k = V_{k+1}$, поскольку $V_{k+1} \subset V_k$ (ведь победа в $(k+1)$-м раунде возможна только при условии победы в предыдущем k-м раунде). Следовательно,

$$p(V_{k+1}|V_k) = \frac{p(V_{k+1})}{p(V_k)}. \tag{1}$$

По формуле полной вероятности имеем:

$$p(V_k) = \sum_{m=1}^{n} p(m)p(V_k|m),$$

где $p(m) = \frac{1}{n}$ — вероятность того, что случайно выбранная вначале команда имеет силу m, $p(V_k|m)$ — вероятность победы в k-м раунде команды с силой m. Ясно, что при $m \leqslant k$ у команды m нет шансов победить в k-м раунде (то есть $p(V_k|m) = 0$), так что

$$p(V_k) = \frac{1}{n} \sum_{m=k+1}^{n} p(V_k|m). \tag{2}$$

Вычислим вероятность $p(V_k|m)$ при $m \geqslant k+1$. Соответствующее событие состоит из всех турнирных путей

$$(m, s_1, \ldots, s_k, s_{k+1}, \ldots, s_{n-1}) \tag{3}$$

с фиксированным m, для которых выполнены неравенства $m > s_1$, $m > s_2$, ..., $m > s_k$ (не путать с событием V_k, в котором m не фиксировано). Подсчитаем количество таких путей.

На место чисел $s_1, \ldots, s_k$ можно поставить любую упорядоченную выборку из множества $\{1, 2, \ldots, m-1\}$; количество таких выборок равно

$$(m-1)(m-2)\ldots(m-k) = C_{m-1}^k \cdot k!$$

(в справедливости последнего равенства нетрудно убедиться непосредственно, но и комбинаторный смысл его тоже ясен: взяли неупорядоченную выборку C_{m-1}^k способами и превратили её в упорядоченную $k!$ способами). На позиции $s_{k+1}, \ldots, s_{n-1}$ ставим оставшиеся $n-1-k$ чисел в произвольном порядке; это можно сделать $(n-1-k)!$ способами. Итого искомое количество путей (ведущих команду m к выигрышу в k-м раунде) равно

$$C_{m-1}^k \cdot k!(n-1-k)!.$$

Общее количество путей (3) с какими угодно s_1, ..., s_{n-1} равно $(n-1)!$, поэтому

$$p(V_k|m) = \frac{C_{m-1}^k \cdot k!(n-1-k)!}{(n-1)!} = \frac{C_{m-1}^k}{C_{n-1}^k}. \tag{4}$$

Этот результат интуитивно очевиден (если вдуматься). Подставляем его в (2):

$$p(V_k) = \frac{1}{n} \sum_{m=k+1}^{n} \frac{C_{m-1}^k}{C_{n-1}^k} = \frac{1}{nC_{n-1}^k} \left(C_k^k + C_{k+1}^k + \ldots + C_{n-1}^k \right). \tag{5}$$

Полученная формула (5) упрощается с помощью тождества[2]

$$C_k^k + C_{k+1}^k + \ldots + C_{n-1}^k = C_n^{k+1} \tag{6}$$

[2]Изначально автор упустил из виду это тождество. На него указали Александр Моркотун и Артём Заневский.

(доказательство которого будет приведено в конце статьи) и приобретает вид:

$$p(V_k) = \frac{C_n^{k+1}}{nC_{n-1}^k}.$$

Данная формула решает задачу. Действительно, остаётся написать

$$p(V_{k+1}) = \frac{C_n^{k+2}}{nC_{n-1}^{k+1}}$$

и воспользоваться формулой (1) для получения окончательного результата:

$$p(V_{k+1}|V_k) = \frac{C_n^{k+2}}{C_n^{k+1}}\frac{C_{n-1}^k}{C_{n-1}^{k+1}} =$$
$$= \frac{n!}{(k+2)!(n-k-2)!} \cdot \frac{(k+1)!(n-k-1)!}{n!} \cdot \frac{(n-1)!}{k!(n-1-k)!} \cdot \frac{(k+1)!(n-k-2)!}{(n-1)!} = \frac{k+1}{k+2}.$$

Итак, искомая вероятность равна $\frac{k+1}{k+2}$ и не зависит от n.

Примечание 1. Формулу (4) можно получить несколько иначе. Перед первым раундом у команды m имеется $n-1$ соперников, из которых $m-1$ слабее её; поэтому вероятность победы команды m в первом раунде равна

$$p(V_1|m) = \frac{m-1}{n-1}.$$

Если команда m победила в первом раунде и прошла во второй, то у неё теперь $n-2$ соперников, из которых $m-2$ слабее её. Вероятность победить во втором раунде *при условии победы в первом раунде* равна $\frac{m-2}{n-2}$; умножая на вероятность прохода во второй раунд, получим

$$p(V_2|m) = \frac{m-1}{n-1} \cdot \frac{m-2}{n-2}.$$

Продолжая эти рассуждения, приходим к формуле

$$p(V_k|m) = \frac{(m-1)(m-2)\ldots(m-k)}{(n-1)(n-2)\ldots(n-k)},$$

что как раз и равно C_{m-1}^k/C_{n-1}^k.

Примечание 2. Осталось убедиться в справедливости тождества (6):

$$C_k^k + C_{k+1}^k + \ldots + C_{n-1}^k = C_n^{k+1}.$$

Докажем его по индукции. База $n = k+1$ очевидна: $C_k^k = C_{k+1}^{k+1}$. Проверим переход от $n-1$ к n:

$$C_k^k + C_{k+1}^k + \ldots + C_{n-1}^k + C_n^k = C_n^{k+1} + C_n^k = \frac{n!}{(k+1)!(n-k-1)!} + \frac{n!}{k!(n-k)!} =$$
$$= \frac{n!}{(n-k-1)!k!}\left(\frac{1}{k+1} + \frac{1}{n-k}\right) = \frac{n!}{(n-k-1)!k!} \cdot \frac{n+1}{(k+1)(n-k)} = \frac{(n+1)!}{(k+1)!(n-k)!} = C_{n+1}^{k+1},$$

что и требовалось.