И. ВЫСОЦКИЙ,
г. Москва

Задача 16

ЗАДАЧИ С УЛИЦЫ

Задача о дискотеке и шляпах

Задача о дискотеке восьмидесятых

Давно задуманная статья о распродаже неожиданно потребовала больше внимания, чем автор полагал прежде. В результате задача о распродаже стала частью более красивой задачи про двудольные графы — с сюжетом, в котором кавалеры и дамы кружатся в танце. Автор благодарен Максиму Евгеньевичу Жуковскому за живое участие в обсуждении вероятностной модели.

Представьте себе теплый июньский вечер. В сельском клубе устроена дискотека, привлекающая окрестную молодежь. Начинает играть медленная романтическая музыка, и юноши устремляются к кружку щебечущих о своем девушек. Необходимым условием успешного приглашения является взаимная симпатия, хотя бы минимальная. Образуется некоторое количество танцующих пар, а те, кто по какой-то причине не нашел себе пару, проводят время за приятной беседой. И вот объявляется белый танец. Инициатива теперь исходит от девушек, но формирование пар подчиняется тем же правилам — антипатия даже с одной стороны делает образование пары невозможным.

Наблюдая наяву, или хотя бы умозрительно, процесс, тот, кто обладает пытливым умом, не может не задать себе вопрос: зависит ли среднее число танцующих пар от того, кто именно вводит вероятностную меру — дамы или кавалеры. Другой вопрос: какова вероятность, что определенный кавалер будет танцевать или что, более того, он будет танцевать с определенной дамой.

Инициацию приглашения на танец придется формализовать. Абстрагируемся от всего, что связано с возможной нерешительностью, стеснительностью или, напротив, гусарским куражом участников, и ограничимся единственным фактором — взаимными симпатиями. Можно представить всю совокупность присутствующих на дискотеке как *двудольный* граф, а взаимные симпатии как ребра, соединяющие некоторые вершины одной доли (кавалеры) с какими-то вершинами другой доли (дамы). Будем считать, что кавалеров ровно n, что дам ровно k и что вероятность образования взаимной симпатии в любой паре «кавалер-дама» равна p и не зависит от прочих имеющихся симпатий. Таким образом, мы приходим к *случайному* двудольному графу. Пример показан на рисунке 1. Танцующие пары в таком случае образуют подграф, состоящий из отдельных, не смежных друг с другом ребер. Такой подграф называется *паросочетанием*.

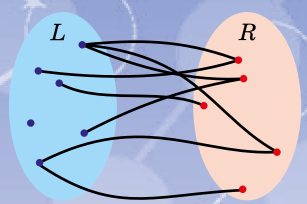


Рис. 1. Двудольный граф

Двудольные графы и максимальные паросочетания

Пусть множество вершин графа разбито на два множества, L (левое) и R (правое), и каждое ребро соединяет одну из вершин множества L с какой-то вершиной множества R. Такой граф называется *двудольным* (см. рис. 1). Может случиться, что какие-то вершины изолированы. Применительно к танцплощадке это значит, что кто-то из присутствующих не имеет ни одной потенциальной пары. Может быть, он и испытывает к кому-то симпатию, но она невзаимная.

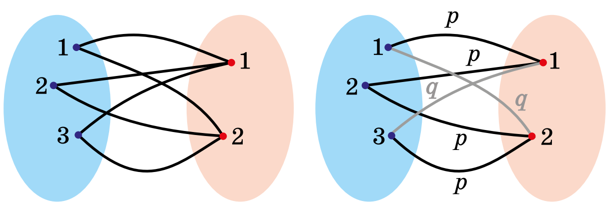
Есть дополнительные материалы на сайте raum.math.ru.

Если множество L состоит из n вершин, а множество R содержит ровно k вершин, то наибольшее возможное число ребер в графе равно nk. Такой наибольший по числу ребер граф называется *полным двудольным графом* и обычно обозначается $K_{n,\,k}$.

На рисунке 2,*а* показан полный граф $K_{3,\,2}$. В нем 6 ребер: (1; 1), (1; 2), (2; 1), (2; 2), (3; 1) и (3; 2).

Предположим, что каждое отдельное ребро появляется с вероятностью p, а с вероятностью q не появляется. Дополнительно предположим, что каждое ребро появляется или не появляется независимо от других ребер. Получается *биномиальный случайный двудольный граф* $D \subset K_{n,\,k}$ на том же множестве вершин, в котором каждое отдельное ребро присутствует с вероятностью p, а с вероятностью $q = 1 - p$ отсутствует.

На рисунке 2,*б* показан случайный двудольный подграф D графа $K_{3,2}$, который содержит 4 ребра (для наглядности два отсутствующих ребра показаны бледно-серым цветом). Вероятность образования такого графа равна p^4q^2.



а) Полный двудольный граф $K_{3,\,2}$ б) Случайный граф $D \subset K_{3,\,2}$

Рис. 2

Будем по очереди выбирать ребра, то есть пары соединенных вершин, но так, чтобы каждая вершина не могла участвовать в двух парах. Образуется подграф, который называется *паросочетанием*. Число ребер в паросочетании назовем мощностью паросочетания. Паросочетание называется *максимальным*, если к нему нельзя добавить ни одного ребра. Почему нас интересуют именно максимальные паросочетания? Если симпатизирующие друг другу кавалер (вершина в левой доле) и дама (дамы образуют правую долю) оказались свободны, то они присоединятся к танцующим, образовав еще одну пару. По крайней мере, будем считать, что танцуют все, кто может.

У одного и того же двудольного графа максимальные паросочетания могут отличаться числом ребер. На рисунке 3 показаны два максимальных паросочетания в некотором двудольном графе. Первое состоит из единственного ребра (2; 2), а второе содержит два ребра. Ребра, не вошедшие в паросочетание, показаны бледно. В этом графе есть еще и третье максимальное паросочетание. Найдите его.

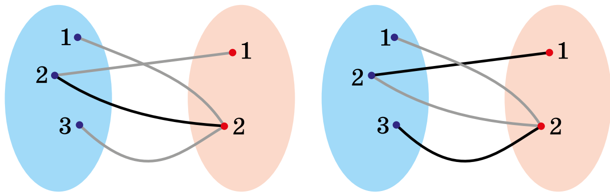


Рис. 3. Два максимальных паросочетания в двудольном графе

Как получаются вероятности максимальных паросочетаний

Случайный опыт, который мы рассматриваем, — образование максимального паросочетания в двудольном графе. Элементарные события — возможные максимальные паросочетания. Вероятности этих элементарных событий определить непросто. Но если граф невелик, это можно сделать непосредственно. Для примера рассмотрим граф D на рисунке 2,*б*. Он имеет три максимальных паросочетания: {(1; 1), (2; 2)}, {(2; 1), (3; 2)} и {(1; 1), (3; 2)} (рис. 4).

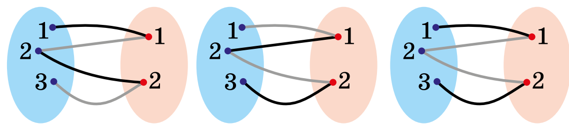


Рис. 4. Три возможных максимальных паросочетания в графе D

Многое зависит от того, кто из кавалеров приглашает первым. Если порядок кавалеров случаен и каждый выбирает одну из дам тоже случайным образом, то возникает процесс *выбора слева*, который проиллюстрирован с помощью дерева на рисунке 5.

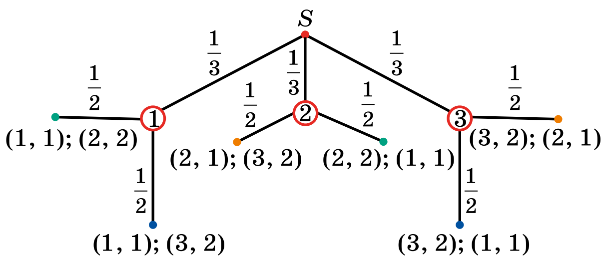


Рис. 5. Случайный выбор ребер слева в графе D (от вершин левой доли, то есть от кавалеров)

Если первым приглашает даму кавалер 1 (вероятность равна $\frac{1}{3}$), то он обязательно пригласит даму 1 (в паросочетание попадет ребро (1, 1)). В этом случае кавалер 2 с необходимостью пригласит даму 2, кавалер 3 останется без пары, и получится паросочетание (1, 1), (2, 2). Если первым приглашает кавалер 2, то возможны варианты. Рассматривая дерево на рисунке 5, находим вероятности всех трех максимальных паросочетаний. Они оказываются одинаковыми.

Максимальное паросочетание	● (1, 1), (2, 2)	● (1, 1), (3, 2)	● (2, 1), (3, 2)
Вероятность	$\frac{1}{3}\cdot\frac{1}{2}+\frac{1}{3}\cdot\frac{1}{2}=\frac{1}{3}$	$\frac{1}{3}\cdot\frac{1}{2}+\frac{1}{3}\cdot\frac{1}{2}=\frac{1}{3}$	$\frac{1}{3}\cdot\frac{1}{2}+\frac{1}{3}\cdot\frac{1}{2}=\frac{1}{3}$

Мы рассмотрели выбор слева. Аналогично можно устроить случайный *выбор справа*, выбирая в случайном порядке вершины в правой доле и случайные инцидентные ей ребра (объявить белый танец). Получится другая вероятностная мера на том же множестве элементарных событий.

Задача 1. Найти вероятности всех трех максимальных паросочетаний в графе D, если объявлен белый танец, то есть если дамы в случайном порядке случайным образом выбирают кавалеров.

Ответ:

Максимальное паросочетание	● (1, 1), (2, 2)	● (1, 1), (3, 2)	● (2, 1), (3, 2)
Вероятность	$\frac{3}{8}$	$\frac{1}{4}$	$\frac{3}{8}$

Иная картина возникает, если выбор пар доверяется стороннему лицу. Предположим, некто видит граф D (см. рис. 2,*б*) и строит паросочетание, каждый раз выбирая с равными шансами одно из доступных ребер. Такой метод построения паросочетания можно назвать *реберным*, имея в виду, что сразу выбирается случайное ребро целиком, а не его концы по очереди.

Задача 2. Найдите вероятности всех трех максимальных паросочетаний в графе D при реберном выборе.

Ответ: вероятности совпадают с вероятностями из задачи 1.

Мы рассмотрели лишь три из великого множества возможных способов определения вероятностной меры: случайный выбор слева, случайный выбор справа и случайный реберный выбор. Ограничимся изучением первых двух, когда кавалеры в случайном порядке выбирают дам и наоборот. Поставим вопросы о вероятностях того, что полученное максимальное паросочетание имеет определенную мощность, и о математическом ожидании этой мощности.

Мы видим, что если двудольный граф сформирован, то вероятность образования того или иного паросочетания зависит от того, как производится выбор — справа или слева. Если же граф случайный и каждое ребро в нем возникает с вероятностью p независимо от других ребер, то ситуация иная. Мы покажем, что вероятности формирования паросочетаний в биномиальном случайном двудольном графе D не зависят от того, из какой доли производится выбор — справа или слева.

Средняя мощность максимального паросочетания (среднее число танцующих пар)

Пусть имеется биномиальный случайный двудольный граф $D_{n,k} \subset K_{n,k}$, то есть он содержит n вершин в левой доле и k вершин в правой доле. Будем строить максимальное паросочетание. Предположим, что случайный выбор пар производится слева. Иными словами, кавалеры приглашают дам. Назовем $X_{n,k}$ число ребер в получившемся максимальном паросочетании (мощность паросочетания) и для краткости обозначим через $m_{n,k}$ математическое ожидание этой величины:

$$m_{n,k} = \mathrm{E}X_{n,k}.$$

Ребер в паросочетании не больше, чем вершин в наименьшей из долей. Поэтому

$$0 \le m_{n,k} \le \min(n, k).$$

В частности, $m_{n,0} = m_{0,n} = 0$ для любого n, то есть паросочетание пусто, если пуста хотя бы одна из долей графа.

Задача 3. Найти числа $m_{n,k}$.

Решение. Выберем случайную вершину в левой доле (назовем ее первой). Возможны два случая.

1. Первая вершина имеет степень 0, то есть из нее не выходит ни одного ребра. Вероятность этого равна q^k. Удалив первую вершину, получим случайный двудольный граф $D_{n-1,k}$ (рис. 6) и новую случайную величину $X_{n-1,k}$ «мощность максимального паросочетания в графе $D_{n-1,k}$». Поскольку ни одно ребро не было выбрано, в этом случае

$$X_{n-1,k} = X_{n,k}$$

и потому

$$m_{n,k} = m_{n-1,k}.$$

2. Из первой вершины выходит хотя бы одно ребро. Тогда при выборе ребра получается пара вершин, связанных этим ребром. Удалив обе эти вершины и все связанные с ними ребра, мы получим случайный двудольный граф $D_{n-1,k-1}$ (рис. 7) и случайную величину $X_{n-1,k-1}$ «мощность максимального паросочетания в $D_{n-1,k-1}$». Очевидно, в этом случае

$$X_{n-1,k-1} = X_{n,k} - 1,$$

и потому

$$m_{n,k} = 1 + m_{n-1,k-1}.$$

На рисунках 6 и 7 показано, как происходит формирование графа после удаления первой вершины в случае, когда она изолированная, и в случае, когда она неизолированная.

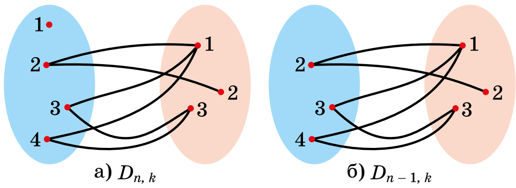


Рис. 6

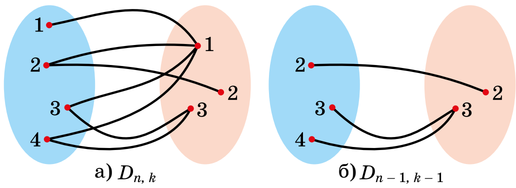


Рис. 7

Введем индикатор I события «первая вершина имеет степень 0», то есть $I = 1$, если первой вершине не инцидентно ни одно ребро, и $I = 0$, если хотя бы одно ребро есть. Тогда

$$X_{n,\,k} = I \cdot X_{n-1,\,k} + (1 - I) \cdot (1 + X_{n-1,\,k-1}).$$

Индикатор I связан только с первой вершиной, а случайная величина $X_{n-1,\,k}$ образуется при выборе паросочетания из графа $D_{n-1,\,k}$, в котором первой вершины уже нет. Значит, случайные величины I и $X_{n-1,\,k}$ независимы. Аналогично находим, что величины $(1 - I)$ и $1 + X_{n-1,\,k-1}$ тоже независимы. Поэтому, переходя к математическим ожиданиям, получаем:

$$\mathrm{E}X_{n,\,k} = \mathrm{E}I \cdot \mathrm{E}X_{n-1,\,k} + \mathrm{E}(1 - I) \cdot \mathrm{E}(1 + X_{n-1,\,k-1}).$$

Число $\mathrm{E}I$ равно вероятности события «первая вершина имеет степень 0»: $\mathrm{E}I = q^k$. Следовательно,[1]

$$m_{n,\,k} = q^k m_{n-1,\,k} + (1 - q^k)(1 + m_{n-1,\,k-1}). \qquad (1)$$

Для начала рекурсии следует положить $m_{n,\,0} = m_{0,\,n} = 0$ для любого $n \geq 0$. Этого достаточно, чтобы построить таблицу чисел $m_{n,\,k}$. Возьмем $q = 0{,}9$, то есть мы будем считать, что взаимные симпатии возникают довольно редко: в среднем одна из десяти возможных (рис. 8).

Неожиданная симметрия

Мы видели (см. задачу 1), что вероятность образования определенного максимального паросочетания в фиксированном графе зависит от того, с какой доли начинается случайный выбор ребер (задачи 1 и 2). Кроме того, формула (1) не отличается симметрией по n и k. Поэтому несколько удивляет симметричность полученной таблицы: среднее число танцующих пар, похоже, не зависит от того, кто кого приглашает, то есть какой выбор мы используем — слева или справа. Действительно ли $m_{n,\,k} = m_{k,\,n}$ для любых n и k? Или эта симметричность кажущаяся и при больших n и k пропадает?

Задача 4. Доказать, что для любых $n,\ k \geq 0$ имеется симметрия $m_{n,\,k} = m_{k,\,n}$, то есть что среднее число танцующих пар не зависит от того, кто

[1] Формулу (1) можно получить без индикаторов как аналог формулы полной вероятности для условных математических ожиданий.

Задача о дискотеке

Среднее количество пар (средняя мощность максимального паросочетания)

q= 0,9

k\n	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1	0	0,100	0,190	0,271	0,344	0,410	0,469	0,522	0,570	0,613	0,651	0,686	0,718	0,746	0,771	0,794
2	0	=E3^$B8*C8+(1-$E$3^$B8)*(1+C7)					0,909	1,016	1,112	1,199	1,277	1,348	1,413	1,471	1,523	1,570
3	0	0,271	0,520	0,748	0,958	1,149	1,323	1,482	1,627	1,758	1,878	1,986	2,084	2,173	2,254	2,327
4	0	0,344	0,663	0,958	1,230	1,480	1,710	1,921	2,114	2,290	2,451	2,598	2,731	2,853	2,963	3,063
5	0	0,410	0,792	1,149	1,480	1,787	2,071	2,332	2,573	2,795	2,998	3,183	3,353	3,508	3,649	3,778
6	0	0,469	0,909	1,323	1,710	2,071	2,406	2,718	3,006	3,272	3,517	3,742	3,949	4,138	4,311	4,470
7	0	0,522	1,016	1,482	1,921	2,332	2,718	3,077	3,411	3,721	4,008	4,274	4,518	4,743	4,949	5,138
8	0	0,570	1,112	1,627	2,114	2,573	3,006	3,411	3,790	4,144	4,473	4,778	5,060	5,321	5,561	5,782
9	0	0,613	1,199	1,758	2,290	2,795	3,272	3,721	4,144	4,540	4,910	5,255	5,575	5,872	6,147	6,401
10	0	0,651	1,277	1,878	2,451	2,998	3,517	4,008	4,473	4,910	5,320	5,704	6,063	6,397	6,706	6,993
11	0	0,686	1,348	1,986	2,598	3,183	3,742	4,274	4,778	5,255	5,704	6,127	6,523	6,893	7,239	7,560
12	0	0,718	1,413	2,084	2,731	3,353	3,949	4,518	5,060	5,575	6,063	6,523	6,956	7,363	7,744	8,099
13	0	0,746	1,471	2,173	2,853	3,508	4,138	4,743	5,321	5,872	6,397	6,893	7,363	7,806	8,221	8,611
14	0	0,771	1,523	2,254	2,963	3,649	4,311	4,949	5,561	6,147	6,706	7,239	7,744	8,221	8,672	9,096
15	0	0,794	1,570	2,327	3,063	3,778	4,470	5,138	5,782	6,401	6,993	7,560	8,099	8,611	9,096	9,553

Рис. 8. Таблица чисел $m_{n,\,k}$ при $q = 0{,}9$

кого приглашает — кавалеры дам или дамы кавалеров.

Доказательство. Будем считать, что $n \ge k$ и вести доказательство по индукции по k. Базис индукции здесь весьма экзотический.

Во-первых, очевидно, $m_{n,n} = m_{n,n}$, то есть равенство $m_{n,k} = m_{k,n}$ верно при $n = k$.

Во-вторых, если $k = 0$, то $m_{n,k} = m_{k,n} = 0$ по определению, то есть и в этом случае утверждение верно.

В-третьих, нам потребуется равенство $m_{n,k} = m_{k,n}$ при $k = 1$. Оно почти очевидно. Предположим, что одна доля графа состоит из n вершин, а другая — из единственной вершины. Независимо от того, из какой доли ведется выбор, паросочетание будет пустым, если ребер нет вовсе (вероятность этого q^n), и будет состоять из единственного ребра, если есть хотя бы одно ребро (вероятность этого $1 - q^n$). И в том, и в другом случае среднее число ребер в паросочетании равно

$$m_{n,1} = m_{1,n} = q^n \cdot 0 + (1 - q^n) \cdot 1 = 1 - q^n.$$

Базис индукции: равенство $m_{n,k} = m_{k,n}$ выполняется при $k = 0$, при $k = 1$ и при $k = n$. При $n \ge k \ge 0$ точки, которые образуют базис индукции, отмечены на координатной плоскости зеленым цветом на рисунке 9. Розовым цветом показаны пары $(n; k)$, для которых равенство $m_{n,k} = m_{k,n}$ предстоит доказать.

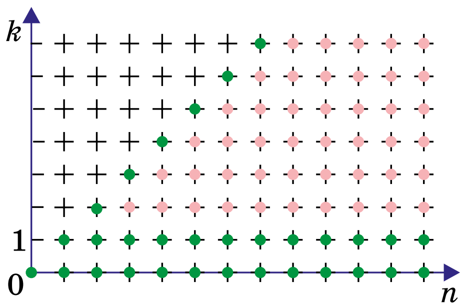


Рис. 9

Теперь сформулируем *предположение индукции*. Будем считать, что $2 \le k_0 < n_0$ и что равенство $m_{n,k} = m_{k,n}$ верно для всех пар (n, k), где $0 \le n < n_0$, и для всех пар (n_0, k), где $0 \le k < k_0$.

Совершим *индукционный переход*, то есть докажем равенство $m_{n,k} = m_{k,n}$ для пары $n = n_0$, $k = k_0$. Обозначим $\Delta_{n,k}$ разность $(m_{n,k} - m_{k,n})$ и покажем, что она равна нулю.

Из (1) следует:

$$\Delta_{n,k} = m_{n,k} - m_{k,n} =$$
$$= q^k m_{n-1,k} + (1 - q^k)(1 + m_{n-1,k-1}) -$$
$$- q^n m_{k-1,n} - (1 - q^n)(1 + m_{k-1,n-1}).$$

По предположению индукции

$$m_{n-1,k-1} = m_{k-1,n-1},$$

поэтому

$$\Delta_{n,k} = q^k m_{n-1,k} - q^n m_{k-1,n} + (q^n - q^k)(1 + m_{n-1,k-1}).$$

Еще раз воспользуемся предположением

$$m_{k-1,n} = m_{n,k-1} \text{ и } m_{n-1,k} = m_{k,n-1}:$$
$$\Delta_{n,k} = q^k m_{k,n-1} - q^n m_{n,k-1} + (q^n - q^k)(1 + m_{n-1,k-1}),$$

и применим рекурсию (1) теперь к числам

$$m_{k,n-1},\ m_{n,k-1},\ m_{n-1,k-1}:$$
$$\Delta_{n,k} = q^k(q^{n-1}m_{k-1,n-1} +$$
$$+ (1 - q^{n-1})(1 + m_{k-1,n-2})) - q^n(q^{k-1}m_{n-1,k-1} +$$
$$+ (1 - q^{k-1})(1 + m_{n-1,k-2})) +$$
$$+ (q^n - q^k)(1 + q^{k-1}m_{n-2,k-1} +$$
$$+ (1 - q^{k-1})(1 + m_{n-2,k-2})).$$

Преобразуем полученное выражение, стремясь уменьшить количество слагаемых и вынося за скобки общие множители:

$$\Delta_{n,k} = (q^k - q^{k+n-1})(1 + m_{k-1,n-2}) -$$
$$- (q^n - q^{n+k-1})(1 + m_{n-1,k-2}) +$$
$$+ (q^n - q^k)(1 + q^{k-1}m_{n-2,k-1} +$$
$$+ (1 - q^{k-1})(1 + m_{n-2,k-2})) =$$
$$= (1 - q^{k-1})q(q^{k-1}m_{k-1,n-2} - q^{n-1}m_{n-1,k-2} +$$
$$+ (q^{n-1} - q^{k-1})(1 + m_{n-2,k-2})) =$$
$$= q(1 - q^{k-1})\ \Delta_{n-1,k-1}.$$

По предположению индукции $\Delta_{n-1,k-1} = 0$. Значит, $\Delta_{n,k} = 0$. Утверждение доказано.

В ходе преобразований использовались равенства:

$$m_{k-1,n} = m_{n,k-1},\ m_{n-1,k} = m_{k,n-1},$$
$$m_{n-1,k-1} = m_{k-1,n-1},\ m_{k-1,n-2} = m_{n-2,k-1},$$

которые верны в силу предположения индукции. Это можно увидеть на рисунке 10: при доказательстве равенства $m_{n,k} = m_{k,n}$ для красной точки использовалась истинность этого равенства для четырех синих точек.

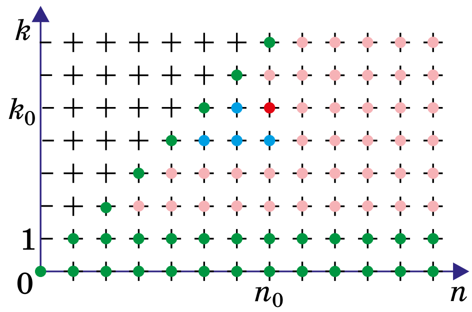


Рис. 10

У читателя мог возникнуть вопрос: зачем потребовалось отдельно доказывать, что $m_{n,1} = m_{1,n}$? Дело в том, что в доказательстве фигурирует число $m_{n-2,k-2}$, поэтому наименьшее k, для которого проводится доказательство по индукции, должно быть не меньше 2. Поэтому для $k = 1$ утверждение пришлось доказывать отдельно и включать его в базис. Можно было пойти другим путем, заявив, что $m_{n,-1} = 0$, но это еще бóльшая экзотика.

О вероятностях

Теперь, когда мы умеем вычислять среднее число ребер $m_{n,k}$ в максимальном паросочетании, можно перейти к более сложному вопросу о вероятностях всяких событий.

Задача 5. Найти вероятность того, что танцует определенная пара или определенный кавалер.

В терминах двудольных графов задача формулируется так: найти вероятность того, что в образовавшемся паросочетании участвует определенное ребро или определенная вершина левой доли.

Решение. Пусть l и r — две какие-то вершины из левой и правой долей случайного двудольного графа $D_{n,k}$. Кратко скажем, что l — левая вершина, а r — правая. Обозначим $A_{l,r}$ событие «ребро (l, r) участвует в паросочетании». Введем индикатор этого события: $I_{l,r} = 1$, если событие $A_{l,r}$ осуществилось, и $I_{l,r} = 0$ в противном случае. Тогда $\mathrm{E}I_{l,r} = \mathrm{P}(A_{l,r})$. Ни одно ребро в графе $D_{n,k}$ не лучше и не хуже никакого другого, поэтому эту вероятность можно обозначить буквой ρ без индексов: $\mathrm{E}I_{l,r} = \rho$ для любого ребра (l, r).

Число ребер в паросочетании равно $X_{n,k}$. С другой стороны, число ребер в паросочетании равно сумме всех индикаторов

$$X_{n,k} = \sum_{l\in L,\, r\in R} I_{l,r},$$

где суммирование ведется по всем nk возможным ребрам (l, r) графа $D_{n,k}$. Перейдем в этом равенстве к математическим ожиданиям:

$$m_{n,k} = \mathrm{E}X_{n,k} = \sum_{l\in L,\, r\in R} \mathrm{E}I_{l,r} = nk\rho = nk\mathrm{P}\left(A_{l,r}\right),$$

откуда

$$\mathrm{P}\left(A_{l,r}\right) = \frac{m_{n,k}}{nk}. \qquad (2)$$

Чтобы решить вторую часть задачи, заметим, что число вершин левой доли, участвующих в паросочетании, тоже равно $X_{n,k}$. Поэтому вероятность события B_l «левая вершина l участвует в паросочетании» вычисляется аналогично:

$$\mathrm{P}\left(B_l\right) = \frac{m_{n,k}}{n}. \qquad (3)$$

Задача о неликвидном товаре

Беглый взгляд фиксирует артефакты: шляпу с мягкими полями кустарной работы из синтетической соломки, махровый халат с мордой тигра на спине, белую вазу с зернистой фотопечатью «Афродита, выходящая из морской пены» и плюшевую корову в треть натуральной величины.

Главный парадокс вещевого рынка имеет психологическую природу: неужели это кто-то в здравом уме купит?

Но зайдем в эту кунсткамеру через неделю и с удивлением увидим, что коровы нет. Не надо спрашивать продавца, куда она делась: не дай бог, он решит, что вырос спрос на мягкую игрушку крупных форм, и закажет поставщику еще три такие коровы.

Рассмотрим более простую модель, ограничившись одной товарной группой. Пусть речь идет, например, только о шляпах.

Предположим, мы имеем n потенциальных покупателей в некотором относительно изолированном сообществе (деревня, поселок) и k единиц однородного товара (шляпы, для определенности). Шляпы могут быть одной модели, а могут отличаться потребительскими свойствами (цвет, фасон, размер, ширина тесемки и т.п.). Наша цель — оценить число проданных единиц товара X или, наоборот, количество неликвида.

Будем считать, что каждому случайному покупателю каждая случайно взятая единица товара нравится с вероятностью p или не нравится с вероятностью $q = 1 - p$. Будем также считать, что предпочтения покупателей независимы.

Если какие-то шляпы приглянулись очередному покупателю, то он купит одну из них (одной голове две шляпы не нужны). Если клиенту ничего не понравилось, то он уходит без покупки. Есть еще такие, которые ничего не покупают, даже если им что-то нравится, но их мы исключим из рассмотрения, ибо они не покупатели, а праздно шатающиеся граждане.

Получается знакомая модель — биномиальный двудольный случайный граф. Отличие лишь в том, что инициатива может исходить только от покупателей и не исходит от шляп.

Мы знаем, что в среднем будет продано $m_{n,k}$ шляп, и умеем находить эти числа. Сейчас нас интересует, сколько шляп останется непроданными. Попробуем оценить остаток товара сверху — насколько эта величина зависит от исходного количества товара. Разумно считать, что шляп не больше, чем покупателей, то есть что $k \le n$.

Задача 6. В предположении, что $n \ge k$, оценить среднее количество вершин правой доли, не участвующих в максимальном паросочетании (среднее число нераспроданных шляп).

Решение. Перепишем (1) в виде

$$m_{n,k} = m_{n-1,k-1} + 1 + q^k(m_{n-1,k} - m_{n-1,k-1} - 1). \qquad (4)$$

Увеличение правой доли графа на одну вершину (добавление одной шляпы) не может снизить продажи, но и не может увеличить их больше, чем на единицу. Поэтому

$$0 \le m_{n-1,k} - m_{n-1,k-1} \le 1.$$

Значит, последнее слагаемое в (4) отрицательно, но не меньше $-q^k$, то есть для любого j

$$m_{n,j} = m_{n-1,j-1} + (1-\varepsilon),$$

где $0 \le \varepsilon_j \le q^j$.

Поэтому оценка получается в результате суммирования геометрической прогрессии:

$$m_{n,k} = m_{n-1,k-1} + (1-\varepsilon_k) =$$
$$= m_{n-2,k-2} + 2 - \varepsilon_{k-1} - \varepsilon_k = \ldots$$
$$\ldots = m_{n-k,0} + k - \sum_{j=1}^{k}\varepsilon_j = k - \sum_{j=1}^{k}\varepsilon_j > k - \sum_{j=1}^{\infty} q^j =$$
$$= k - \frac{q}{1-q} = k - \frac{q}{p}.$$

Получается, что средняя мощность максимального паросочетания велика: она меньше максимально возможной на величину, ограниченную сверху числом $\frac{q}{p}$, которое не зависит ни от n, ни от k. Иными словами, среднее число непроданных шляп меньше $\frac{q}{p}$ к моменту, когда число зашедших покупателей сравнивается с числом шляп.

Естественно было бы ждать, что шляп останется довольно много, если товар так себе. Но мы выяснили, что это не так: в среднем непроданных шляп мало, причем в некоторых случаях неправдоподобно мало, поскольку их среднее число оценивается сверху константой. Например, если $q = 0{,}9$ (случайному покупателю в среднем нравится лишь одна шляпа из 10), то среднее число непроданных шляп не превосходит $\frac{0{,}9}{0{,}1} = 9$ независимо от того, сколько их было в продаже, 10, 50 или 200. Интересное наблюдение.

Задачи для самостоятельного решения

1. В соответствии с сюжетом известной песни, на танцах 9 ребят и 10 девчат. Взаимные симпатии в каждой паре устанавливаются независимо от других с вероятностью $q = 0{,}2$. С помощью электронной таблицы найдите среднее количество танцующих пар. Верно ли, что если увеличить вдвое число девчат и ребят, то среднее число танцующих также вырастет вдвое?

2. В условиях предыдущей задачи на дискотеке присутствуют Петр и Юля. Найдите вероятность того, что:

а) Петр будет танцевать с кем-то;

б) Петр будет танцевать именно с Юлей.

3*. Обозначим $p_{n,k}(s)$ вероятность того, что при случайном выборе слева в биномиальном двудольном графе $D_{n,k}$ образуется паросочетание, в котором ровно s ребер. Покажите, что при $n, k \ge 1$ и $1 \le s \le \min(k, n)$ верно соотношение

$$p_{n,k}(s) = q^k p_{n-1,k}(s) + (1-q^k)p_{n-1,k-1}(s-1).$$

Ответы к задачам из статьи «Задача о разборчивой невесте»

1. Приблизительно 0,331.

2. *Указание.* В равенстве $p_{n,k} = \frac{k}{n}\left(H_{n-1} - H_{k-1}\right)$ нужно представить H_{k-1} как $H_{k-2} + \frac{1}{k-1}$.

3. *Указание.* Нужно воспользоваться равенствами (2) и (3) или формулой из задачи 4*.

4. *Доказательство.* Равенство можно получить алгебраически из $p_{n,k} = \frac{k}{n}\left(H_{n-1} - H_{k-1}\right)$ или составить из комбинаторных соображений. Пойдем вторым путем. Заменим числа a_j их рангами — натуральными числами от 1 до n, расположенными по величине в том же порядке, что и числа a_j.

Всего равновозможных перестановок чисел от 1 до n ровно $n!$. Пусть число i равно количеству чисел в последней веренице, которые больше, чем лидер предпоследней вереницы. Нужно найти число перестановок, где среди первых k чисел наибольшее число равно $n - i$, а среди чисел n, $n-1$, ..., $n-i+1$ первым встречается число n, причем нужно это сделать для каждого $i = 1, \ldots, n-k$ и сложить полученные выражения. Применим правило умножения.

1. Число $(n-i)$ поставим на одно из k первых мест — k способов это сделать.

2. Выберем среди $(n-k)$ последних мест (последняя вереница) i мест для самых больших чисел n, $n-1$, ..., $n-i+1$ — C_{n-k}^{i} способов. Число n поставим на первое из выбранных мест, а остальные $(i-1)$ чисел от $n-1$ до $n-i+1$ разместим на прочих местах $(i-1)!$ способами.

3. Оставшиеся $(n-i-1)$ чисел расположим на оставшихся местах $(n-i-1)!$ способами. Получаем:

$$p_{n,k} = \frac{1}{n!}\sum_{i=1}^{n-k} k \cdot C_{n-k}^{i} \cdot (i-1)! \cdot (n-i-1)!.$$

Осталось выполнить очевидные преобразования.