Труды по дискретной математике
Том 3, с. 139–146
М.: ФИЗМАТЛИТ, 2000

ВЕРОЯТНОСТЬ СОВМЕСТНОСТИ ОДНОЙ СИСТЕМЫ СЛУЧАЙНЫХ УРАВНЕНИЙ СПЕЦИАЛЬНОГО ВИДА

В. Ф. КОЛЧИН

Рассмотрим систему уравнений относительно неизвестных $x_1, \ldots, x_n$, имеющую вид

$$\delta(x_{i(t)}, x_{j(t)}) = a_t, \qquad t = 1, \ldots, T, \tag{1}$$

где $x_1, \ldots, x_n$ принимают значения $0, 1, \ldots, q-1$,

$$\delta(x, y) = \begin{cases} 0, & \text{если} \quad x = y, \\ 1, & \text{если} \quad x \neq y, \end{cases}$$

индексы $i(t)$, $j(t)$, $t = 1, \ldots, T$, — независимые случайные величины, принимающие значения $1, \ldots, n$ с равными вероятностями, правые части $a_1, \ldots, a_T$ принимают значения 0 и 1, относительно них справедлива одна из гипотез H_0 и H_1.

При гипотезе H_0 существует вектор

$$x^* = (x_1^*, \ldots, x_n^*)$$

с компонентами, принимающими значения $0, 1, \ldots, q-1$, и

$$a_t = \delta\left(x_{i(t)}^*, x_{j(t)}^*\right), \qquad t = 1, \ldots, T.$$

Пусть n_i — число компонент вектора x^*, принявших значение i, $i = 0, 1, \ldots, q-1$. Положим

$$p = \frac{1}{n^2} \sum_{i=0}^{q-1} n_i^2. \tag{2}$$

Нетрудно видеть, что при гипотезе H_0 правые части $a_1, \ldots, a_T$ являются независимыми одинаково распределенными случайными величинами, зависящими от левой части системы (1), и

$$\mathbf{P}\{a_1 = 0\} = p, \qquad \mathbf{P}\{a_1 = 1\} = 1 - p. \tag{3}$$

© В. Ф. Колчин, 2000

При гипотезе H_1 предполагается, что правые части $a_1, \ldots, a_T$ не зависят от левой части системы (1), независимы между собой и имеют распределение (3).

Таким образом, при гипотезе H_0 система (1) совместна при любом T, поскольку она имеет решение $x_1^*, \ldots, x_n^*$. При гипотезе H_1 система (1) может оказаться несовместной.

Обозначим $S_{n,T}$ событие, состоящее в том, что система (1) совместна. Для различения гипотез H_0 и H_1 можно взять критическое множество $S_{n,T}$ и принимать гипотезу H_0, если система (1) совместна, и принимать гипотезу H_1, если система (1) несовместна.

Ясно, что

$$\mathbf{P}\{S_{n,T} \mid H_0\} = 1,$$

а ошибка

$$\beta = \mathbf{P}\{S_{n,T} \mid H_1\}$$

равна вероятности совместности системы (1) при случайных правых частях с распределением (3).

Таким образом, основной характеристикой предложенного критерия является вероятность совместности системы (1) при гипотезе H_1.

Для наглядности приведем одну интерпретацию системы (1). Пусть в урне имеется n предметов q сортов, причем n_i — число предметов сорта i,

$$i = 0, 1, \ldots, q-1, \qquad n_0 + n_1 + \ldots + n_{q-1} = n.$$

Проводится T независимых экспериментов, в каждом из которых путем случайной выборки с возвращением извлекаются два предмета и определяется, совпадают их сорта, или они различны. Ясно, что результаты этих T экспериментов можно записать в виде системы (1) при справедливости гипотезы H_0. Если сравнения предметов производятся так, что результат не зависит от выбранных предметов и значения 0 и 1 выбираются с вероятностями p и $1-p$ соответственно, то этим T экспериментам отвечает система (1) при справедливости гипотезы H_1.

Левой части системы (1) соответствует случайный граф $G_{n,T}$ с n занумерованными вершинами и T ребрами

$$e_t = (i(t),\, j(t)), \qquad t = 1, \ldots, T.$$

Можно считать, что ребра графа $G_{n,T}$ образуются в результате T независимых испытаний, причем в каждом испытании ребро соединяет любую из $\binom{n}{2}$ пар различных вершин с вероятностью $2n^{-2}$ и образует петлю в любой из n вершин с вероятностью n^{-2}. Таким образом, реализации графа $G_{n,T}$ могут иметь петли и кратные ребра.

Строение случайного графа $G_{n,T}$ хорошо изучено [2], особенно в случае, когда $n, T \to \infty$ так, что $2T/n \to \lambda$, $1 < \lambda < 1$. В этом случае связными компонентами графа $G_{n,T}$ с вероятностью, стремящейся к единице, являются деревья и компоненты с одним циклом. В соответствии с этим система (1) разлагается на подсистемы с непересекающимися множествами аргументов. Система (1) совместна тогда и только тогда, когда совместна

каждая из этих подсистем, при этом вероятность совместности $P_{n,T}$ системы (1) есть произведение вероятностей совместности этих подсистем.

Легко видеть, что вероятность совместности подсистемы, граф которой является деревом, равна единице. По тем же соображениям вероятность совместности подсистемы, граф которой содержит цикл, зависит только от строения подсистемы, соответствующей этому циклу.

Найдем вероятность совместности системы уравнений вида (1), граф которой представляет собой цикл длины r. Пусть система имеет вид

$$\delta(x_{i(t)}, x_{j(t)}) = b_t, \qquad t = 1, \ldots, r, \tag{4}$$

и ребра

$$e_t = (x_{i(t)}, x_{j(t)}), \qquad t = 1, \ldots, r,$$

образуют цикл без самопересечений, а правые части $b_1, \ldots, b_r$ не зависят от левой части системы, независимы между собой и имеют распределение (3).

Чтобы увидеть, чему равна вероятность совместности системы (4), потребуется следующее утверждение.

Пусть $\xi_1, \ldots, \xi_m$ — независимые одинаково распределенные случайные величины с распределением

$$\mathbf{P}\{\xi_1 = k\} = \frac{1}{q-1}, \qquad k = 1, \ldots, q-1. \tag{5}$$

Рассмотрим сумму

$$S_m = \xi_1 + \ldots + \xi_m,$$

и пусть

$$p_m = \mathbf{P}\{S_m \equiv 0 \pmod q\}.$$

Лемма 1. *При $m \geq 1$ и $q \geq 3$*

$$p_m = \frac{1}{q-2}\left(1 - \frac{1}{(q-1)^{m-1}}\right). \tag{6}$$

Доказательство. По формуле полной вероятности

$$p_m = \sum_{k=1}^{q-1} \mathbf{P}\{\xi_1 = k\}\,\mathbf{P}\{S_{m-1} \equiv q-k \pmod q\} =$$

$$= \frac{1}{q-1}\sum_{k=1}^{q-1} \mathbf{P}\{S_{m-1} \equiv k \pmod q\} = \frac{1}{q-1}(1 - p_{m-1}).$$

Полагая $p_1 = 0$ и

$$f(x) = \sum_{k=1}^{\infty} p_k x^k,$$

из рекуррентного соотношения находим, что

$$f(x) = \frac{x^2}{(q-1)(1-x)} - \frac{xf(x)}{q-1},$$

откуда

$$f(x) = \frac{x^2}{(q-1)(1-x)(1-x/(q-1))} = \frac{x^2}{q-2}\left(\frac{q-1}{1-x} - \frac{1}{1-x/(q-1)}\right).$$

Разлагая простые дроби в ряд по степеням x и беря коэффициент при x^m, получаем формулу, указанную в лемме.

Заметим, что если $q \geq 3$ то $p_m > 0$ при $m > 1$ и $p_1 = 0$.

Лемма 2. *При $q \geq 3$, $r \geq 1$ вероятность совместности системы* (4) *равна*

$$p_{q,r} = 1 - r(1-p)\,p^{r-1}.$$

Доказательство. Наряду с системой (4) рассмотрим систему над кольцом вычетов по модулю q

$$x_{i(t)} - x_{j(t)} = c_t, \qquad t = 1, \ldots, r, \tag{7}$$

где $c_t = 0$, если $b_t = 0$, и c_t принимает значения $1, \ldots, q-1$ с равными вероятностями, то есть имеет распределение (5), если $b_t = 1$. Поскольку системе (4) соответствует цикл, система (7) совместна тогда и только тогда, когда

$$\sum_{t=1}^{r} c_t \equiv 0 \pmod q.$$

Итак, если

$$\mathbf{P}\left\{\sum_{t=1}^{r} c_t \equiv 0 \pmod q\right\} > 0, \tag{8}$$

то существует реализация $c_1, \ldots, c_r$, при которой система (7) имеет решение. Согласно лемме 1 соотношение (8) выполняется для всех сумм с числом слагаемых, большим единицы. Поэтому система (4) совместна с вероятностью единица, если число единиц в цикле не равно единице. Отсюда следует утверждение леммы, так как $r(1-p)\,p^{r-1}$ есть вероятность того, что в цикле ровно одна единица.

Прежде чем перейти к нахождению асимптотической формулы для вероятности $P_{n,T}$ совместности системы (1) при гипотезе H_1, приведем еще два вспомогательных утверждения.

Обозначим α_r число циклов длины r в графе $G_{n,T}$, соответствующем левой части системы (1), $r = 1, \ldots, n$. Хорошо известно [1]–[5], что если $n, T \to \infty$ так, что $2T/n \to \lambda$, $0 < \lambda < 1$, то с вероятностью, стремящейся к единице, граф $G_{n,T}$ разлагается на связные компоненты, являющиеся деревьями или компонентами, содержащими ровно один цикл. При этом общее число циклов

$$\nu_{n,T} = \alpha_1 + \ldots + \alpha_n$$

имеет в пределе распределение Пуассона с параметром

$$\Lambda = -\frac{1}{2}\ln(1-\lambda).$$

Одним из результатов статьи [4] является следующее утверждение.

Лемма 3. *Если $n, T \to \infty$, $2T/n \to \lambda$, $0 < \lambda < \infty$, то для любого фиксированного R случайные величины $\alpha_1, \ldots, \alpha_R$ асимптотически независимы и число α_r циклов длины r распределено в пределе по закону Пуассона с параметром*

$$\lambda_r = \frac{\lambda^r}{2r}, \qquad r = 1, \ldots, R.$$

Если $\lambda > 1$, то некоторые циклы могут оказаться в так называемой гигантской компоненте графа $G_{n,T}$, появляющейся на этой стадии эволюции случайного графа. Для нашей задачи важно, чтобы были независимы события, состоящие в том, что подсистемы, соответствующие различным циклам, совместны. Для этого достаточно, чтобы циклы не имели общих ребер. Покажем, что с вероятностью, стремящейся к единице, это требование выполняется.

Лемма 4. *Если $n, T \to \infty$, $2T/n \to \lambda$, $0 < \lambda < \infty$, то для любого фиксированного R в графе $G_{n,T}$ с вероятностью, стремящейся к единице, нет двух циклов длины, не превосходящей R, имеющих хотя бы одно общее ребро.*

Доказательство. Если в графе $G_{n,T}$ такие два цикла существуют, то существует цикл длины, не большей $2R$, имеющий перемычку длины, не большей R. Пусть $\xi_r^{(t)}$ — число циклов длины r, имеющих перемычку длины t. Представим $\xi_r^{(t)}$ в виде суммы индикаторов. Занумеруем все $\binom{T}{r+t}$ наборов по $r+t$ строк системы (1) и с i-м набором свяжем случайную величину, равную единице, если эти строки образуют в $G_{n,T}$ цикл без самопересечений длины r с перемычкой длины t. Нетрудно видеть, что математическое ожидание $\xi_r^{(t)}$ удовлетворяет неравенству

$$\mathrm{M}\,\xi_r^{(t)} \le \binom{T}{r+t}\binom{n}{r+t-1}\binom{r+t-1}{r} \times$$

$$\times\,(r-1)!\,r(r-1)(t-1)!\,(r+t)!\left(\frac{2}{n^2}\right)^{r+t} \le \frac{2r}{n}\left(\frac{2T}{n}\right)^{r+t}.$$

Учитывая, что $r, t \le 2R$, находим, что математическое ожидание общего числа таких циклов с перемычками не превосходит c/n, где c — некоторая постоянная, зависящая от R и λ.

Теорема 1. *Если $n, T \to \infty$ так, что $2T/n \to \lambda$, и $p\lambda < 1$, то*

$$P_{n,T} \longrightarrow \exp\left\{-\frac{\lambda(1-p)}{2(1-\lambda p)}\right\}.$$

Доказательство. Оценим вероятность несовместности системы (1). Занумеруем все $\binom{T}{r}$ наборов по r строк и обозначим $A_i^{(r)}$ событие,

состоящее в том, что i-й набор состоит из цикла без самопересечений, дающего несовместную подсистему, $i = 1, \ldots, \binom{T}{r}$. Ясно, что вероятность несовместности системы (1) можно записать в виде

$$1 - P_{n,T} = \mathbf{P}\left(\bigcup_{r=1}^{n}\bigcup_{i=1}^{\binom{T}{r}} A_i^{(r)}\right).$$

Поскольку $A_i^{(r)} \cap A_j^{(s)} = \emptyset$ при $R \neq s$ и любых i, j, эту вероятность можно переписать в виде

$$1 - P_{n,T} = P_1 + P_2, \tag{9}$$

где

$$P_1 = \mathbf{P}\left(\bigcup_{r=1}^{R}\bigcup_{i=1}^{\binom{T}{r}} A_i^{(r)}\right), \qquad P_2 = \mathbf{P}\left(\bigcup_{r>R}\bigcup_{i=1}^{\binom{T}{r}} A_i^{(r)}\right).$$

Обозначим $B_i^{(r)}$, $i = 1, \ldots, \binom{T}{r}$, событие, состоящее в том, что выбранные r строк системы образуют цикл без самопересечений. Нетрудно видеть, что

$$\mathbf{P}\left(B_i^{(r)}\right) = \binom{n}{r}\frac{(r-1)!\,r!}{2}\left(\frac{2}{n^2}\right)^r. \tag{10}$$

Действительно, цикл, соответствующий r выбранным строкам, может быть составлен из различных вершин различными способами. Выбрать r вершин, из которых строится цикл, можно $\binom{n}{r}$ способами. Из r вершин можно построить $(r-1)!/2$ различных циклов. Ребра, образующие каждый из этих циклов, можно упорядочить $r!$ способами при размещении их по r выбранным строкам. Поэтому событие $B_i^{(r)}$ может реализоваться любым из

$$\binom{n}{r}\frac{(r-1)!\,r!}{2}$$

способов, и каждый из них имеет вероятность $(2/n^2)^r$ при $r \geq 3$. Заметим, что, как нетрудно проверить, формула (10) верна и при $r = 1$ и $r = 2$. Согласно лемме 2

$$\mathbf{P}\left(A_i^{(r)}\right) = \mathbf{P}\left(B_i^{(r)}\right)(1 - p_{q,r}) = \binom{n}{r}\frac{r!\,r!}{2}\left(\frac{2}{n^2}\right)^r(1-p)\,p^{r-1}.$$

Поэтому из (10) следует, что

$$P_2 \leq \sum_{r=R+1}^{\infty}\binom{T}{r}\mathbf{P}\left(A_1^{(r)}\right) \leq \sum_{r=R+1}^{\infty}\binom{T}{r}\binom{n}{r}\frac{r!\,r!}{2}\left(\frac{2}{n^2}\right)^r(1-p)\,p^{r-1} \leq$$

$$\leq \sum_{r=R+1}^{\infty}\frac{T^r n^r}{2}\left(\frac{2}{n^2}\right)^r(1-p)\,p^{r-1} = \sum_{r=R+1}^{\infty}\frac{T(1-p)}{n}\left(\frac{2Tp}{n}\right)^{r-1}.$$

Учитывая, что $2T/n \to \lambda$ и $\lambda p < 1$, находим, что выбором достаточно большого R правая часть этой оценки может быть сделана сколь угодно малой.

Оценим теперь P_1. Вместо P_1 удобнее рассматривать вероятность дополнительного события, состоящего в том, что в $G_{n,T}$ нет цикла длины, не превосходящей R, которому соответствовала бы несовместная подсистема. Поскольку случайные величины $\alpha_1, \ldots, \alpha_R$ асимптотически независимы и циклы длины, не превосходящей R, с вероятностью, стремящейся к единице, не имеют общих ребер, рассматриваемая вероятность P_1 в пределе равна произведению вероятностей совместности подсистем, соответствующих каждому из циклов.

Учитывая асимптотическую независимость величин $\alpha_1, \ldots, \alpha_R$, можно отдельно рассмотреть вклад в это произведение циклов каждой фиксированной длины. В силу леммы 2 часть этого произведения, соответствующая одному циклу длины r, равна $p_{q,r}$, а часть произведения, соответствующая всем циклам длины r, равна

$$\sum_{m=0}^{\infty} \mathbf{P}\{\alpha_r = m\}\, p_{q,r}^m.$$

В условиях теоремы

$$\sum_{m=0}^{\infty} \mathbf{P}\{\alpha_r = m\}\, p_{q,r}^m \longrightarrow \sum_{m=0}^{\infty} \frac{\lambda_r^m e^{-\lambda_r}}{m!}\, p_{q,r}^m =$$

$$= \exp\{-\lambda_r(1 - p_{q,r})\} = \exp\left\{-\frac{\lambda^r (1-p)\, p^{r-1}}{2}\right\}.$$

Для любого фиксированного R вклад в вероятность совместности от циклов длины, не превосходящей R, равен

$$\prod_{r=1}^{R} \exp\left\{-\frac{\lambda^r (1-p)\, p^{r-1}}{2}\right\} = \exp\left\{-\frac{1}{2}\sum_{r=1}^{R} \lambda^r (1-p)\, p^{r-1}\right\}.$$

Поскольку $\lambda p < 1$, выбором достаточно большого R хвост

$$\sum_{r>R} \lambda^r (1-p)\, p^{r-1}$$

может быть сделан сколь угодно малым.

Учитывая, что P_2 в (9), как было показано, также может быть сделано сколь угодно малым выбором достаточно большого R, находим, что

$$1 - P_{n,T} \longrightarrow 1 - \exp\left\{-\frac{1}{2}\sum_{r=1}^{R} \lambda^r (1-p)\, p^{r-1}\right\} =$$

$$= 1 - \exp\left\{-\frac{\lambda(1-p)}{2(1-\lambda p)}\right\}.$$

Отсюда следует утверждение теоремы.

СПИСОК ЛИТЕРАТУРЫ

1. *Колчин В. Ф.* Системы случайных уравнений. — М.: МИЭМ, 1988.
2. *Erdos P., Renyi A.* On evolution of random graphs. — Magyar Tud. Akad. Mat. Kut. Int. Kozl., 1960, V. 5, p. 17–61.
3. *Колчин В. Ф.* О числе циклов в случайном графе. — В сб.: Вероятностные задачи дискретной математики. — М.: МИЭМ, 1990, с. 3–8.
4. *Колчин В. Ф., Хохлов В. И.* О числе циклов в неравновероятном случайном графе. — Дискрет. матем., 1990, т. 2, №3, с. 137–145.
5. *Takacs L.* On the limit distribution of the number of cycles in a random graph. — J. Appl. Probab., 1988, v. 25, p. 359–375.