Критерии согласия. Сложная гипотеза

Критерии Колмогорова и ω^2 для сложной гипотезы. Задача проверки *простой гипотезы* о виде закона распределе-ния случайной величины X на практике встречается довольно редко. Гораздо чаще бывает необходимо проверить по *слу- чайной выборке* $\vec{X}_n$ из *генеральной совокупности* X *сложную гипотезу* о принадлежности функции распределения $F(t)$ слу-чайной величины X заданному *параметрическому множеству* распределений $\{F(t;\theta), \theta \in \Theta\}, \Theta \subset \mathbb{R}^d$:

$$H_0\colon F(t) = F_0(t;\theta), \theta \in \Theta.$$

Кажется естественным сначала каким-то образом постро-ить оценку $\widehat{\theta}(\vec{X}_n)$ параметра θ, а затем применить *критерии Колмогорова* и ω^2 для проверки *гипотезы*

$$H_0\colon\ F(t) = F_0(t;\widehat{\theta}(\vec{x}_n)),$$

где $\widehat{\theta}(\vec{x}_n)$ — значение оценки $\widehat{\theta}(\vec{X}_n)$ по данным *выборки* $\vec{x}_n$ -К сожалению, при таком подходе эти *критерии* уже не будут

непараметрическими — при гипотезе H_0 распределение модифицированных *статистик* $\widehat{D}(\vec{X}_n)$ и $\widehat{\omega}^2(\vec{X}_n)$, где

$$\widehat{D}(\vec{x}_n) = \sup_t |F_n(t) - F_0(t;\widehat{\theta}(\vec{x}_n))|,$$

$$\widehat{\omega}^2(\vec{x}_n) = \frac{1}{12n^2} + \frac{1}{n}\sum_{i=1}^{n}\left(F_0(x_{(i)};\widehat{\theta}(\vec{x}_n)) - \frac{2i-1}{2n}\right)^2,$$

вообще говоря, зависит от F_0 и от метода нахождения оценки $\widehat{\theta}(\vec{X}_n)$, что требует составления большого количества таблиц распределений.

Однако если $\widehat{\theta}(\vec{X}_n)$ — *оценки максимального правдоподобия* параметра θ, а элементы $F(t;\theta)$ параметрического множества $\{F(t;\theta),\ \theta \in \Theta\}$ функций распределений получаются при помощи преобразования сдвига и масштаба какого-нибудь одного своего представителя $F(t;\theta_0)$, т.е.

$$F(t;\theta) = F(\frac{t-a}{b}, \theta_0),$$

то для критериев Колмогорова и ω^2 достаточно иметь только одну таблицу для каждого семейства. К таким семействам относятся все важные типы распределений, и, в частности, нормальное. Более того, при небольшой модификации статистик $\widehat{D}(\vec{X}_n)$ и $\widehat{\omega}^2(\vec{X}_n)$ их распределение при $n \geqslant 5$ практически перестает зависеть* от n.

Критерий χ^2 для сложной гипотезы. Пусть функция распределения дискретной случайной величины X, принимающей конечное множество значений $u_1, \ldots, u_r$, зависит от d-мерного вектора параметров θ. Тогда вероятность p_k того, что X примет возможное значение u_k, зависит от θ, т.е. $p_k = p_k(\theta), k = 1, r$. А так как вероятности $p_1(\theta), \ldots, p_r(\theta)$ полностью определяют функцию распределения случайной величины X, то в рассматриваемом случае *основная гипотеза* принимает следующий вид:

$$H_0: \; PrX = u_k = p_k(\theta), \quad k = 1, r, \quad \theta \in \Theta \subset \mathbb{R}^d.$$

Эту сложную гипотезу можно проверить при помощи модификации критерия χ^2 Пирсона.

Пусть $\widehat{\theta}(\vec{x}_n)$ — значение оценки $\widehat{\theta}(\vec{X}_n)$ максимального правдоподобия для θ, а $n_k(\vec{x}_n)$ — количество элементов выборки $\vec{x}_n$, равных u_k, $k = 1, r$. Оценку $\widehat{\theta}(\vec{X}_n)$ получают в результате минимизации логарифма *функции правдоподобия*

$$L(\vec{X}_n;\theta) = \frac{n!}{n_1! \ldots n_r!} \prod_{k=1}^{r} p_i^{n_k(\vec{X}_n)}(\theta), \quad \sum_{i=1}^{r} n_i(\vec{X}_n) = n,$$

как (см. **(3.2)**) решение системы уравнений

$$\sum_{k=1}^{r} \frac{n_k(\vec{X}_n)}{p_k(\theta)} \frac{\partial p_k(\theta)}{\partial \theta_j} = 0, \quad j = \overline{1, d}.$$

—

Пример 5.4. Пусть X и Y — непрерывные случайные величины с функциями распределения $F(t)$ и $G(t)$ соответственно. Даны выборка $\vec{x}_{10}$ с элементами

$$-0{,}15;\ 8{,}60;\ 5{,}00;\ 3{,}71;\ 4{,}29;\ 7{,}74;\ 2{,}48;\ 3{,}25;\ -1{,}15;\ 8{,}38$$

и выборка $\vec{y}_{10}$ с элементами

$$2{,}55;\ 12{,}07;\ 0{,}46;\ 0{,}35;\ 2{,}69;\ -0{,}94;\ 1{,}73;\ 0{,}73;\ -0{,}35;\ -0{,}37.$$

Проверим на уровне значимости $\alpha = 0{,}05$ гипотезу (5.10) против альтернативной гипотезы (5.11).

Выписываем значения объединенного вариационного ряда заданных выборок

$$-1{,}15;\ -0{,}94;\ -0{,}37;\ -0{,}35;\ 0{,}46;\ 0{,}73;\ 1{,}73;\ 2{,}48;\ 2{,}55;\ 2{,}69;\ 3{,}25;\ 3{,}71;\ 4{,}29;\ 5{,}00;\ 7{,}74;\ 8{,}38;\ 8{,}60;\ 12{,}07$$

и последовательность чисел δ_i, $i = 1, 20$,

$$1;\ 0;\ 0;\ 0;\ 1;\ 0;\ 0;\ 0;\ 0;\ 1;\ 0;\ 0;\ 1;\ 1;\ 1;\ 1;\ 1;\ 1;\ 1;\ 0.$$

Вычислив по формуле (5.16) значения величин s_j, $j = 1, 20$, и подставив их в (5.17), определим, что $D(\vec{x}_{10}, \vec{y}_{10}) = 6$. В таблице квантилей распределения статистики* $D(\vec{X}_m, \vec{Y}_n)$ квантили $D_{1-\alpha} = D_{0,95}$ нет, но есть квантиль $D_{0,9476} = 6$. Поэтому гипотезу (5.10) следует отклонить в пользу альтернативной гипотезы (5.11) на уровне значимости $\alpha = 0{,}0524$.

Критерии независимости

Критерий Спирмена. Пусть имеется *случайная выборка* $(X_1, Y_1), \ldots, (X_n, Y_n)$ из *генеральной совокупности* двумерной непрерывной случайной величины (X, Y) с функцией распределения $F(t, \tau)$, а $F_X(t)$ и $F_Y(\tau)$ — функции распределения случайных величин X и Y соответственно. Если случайные величины X и Y имеют нормальные распределения, то для проверки *статистической гипотезы* об их независимости

$$H_0\colon\ F(t,\tau) = F_X(t)\,F_Y(\tau) \tag{5.18}$$

можно использовать процедуру, связанную с вычислениями *выборочного коэффициента корреляции* (см. формулу (6.12)).

Если же о распределениях непрерывных случайных величин X и Y ничего не известно, то для проверки *основной гипотезы* (5.18) при *альтернативной гипотезе*

$$H_1\colon\ F(t,\tau) \neq F_X(t)\,F_Y(\tau)\ \text{ для некоторых } (t,\,\tau) \in \mathbb{R}^2$$

используют ***ранговый критерий Спирмена***, основанный на следующем понятии.

Определение 5.1. ***Рангом*** $R_i(\vec{z}_N)$ ***элемента*** z_i числовой ***последовательности*** $\vec{z}_N = (z_1,\,\ldots,\,z_N)$ называют его порядковый номер в *вариационном ряду* $z_{(1)},\,\ldots,\,z_{(N)}$.

Согласно определению, $R_i(\vec{z}_N)$ — это число элементов последовательности $z_1,\,\ldots,\,z_N$, не бóльших чем z_i, которое можно записать следующим образом:

$$R_i(\vec{z}_N) = 1 + \sum_{k=1} \eta(z_i - z_k),$$

где $\eta(t)$ — *функция Хевисайда*. Ранг любого элемента последовательности $\vec{z}_N$ — это натуральное число в диапазоне от 1 до N, причем ранг наименьшего элемента последовательности равен 1, а ранг наибольшего — N.

Пример 5.5. Рассмотрим выборку $\vec{z}_4 = (3{,}8,\,4{,}7,\,-2{,}6,\,17{,}3)$. Ее вариационный ряд имеет вид $-2{,}6;\ 3{,}8;\ 4{,}7;\ 17{,}3$. Поэтому $R_1(\vec{z}_4) = 2$, $R_2(\vec{z}_4) = 3$, $R_3(\vec{z}_4) = 1$, $R_4(\vec{z}_4) = 4$. #

Определение 5.2. ***Рангом элемента*** Z_i ***случайной выборки*** $\vec{Z}_N = (Z_1,\,\ldots,\,Z_N)$ называют случайную величину $R_i(\vec{Z}_N)$, реализация которой $R_i(\vec{z}_N)$ есть ранг реализации z_i случайной величины Z_i в вариационном ряду $z_{(1)},\,\ldots,\,z_{(N)}$.

Обозначим через $R_i = R_i(\vec{X}_n)$ — ранг элемента X_i случайной выборки $X_1,\,\ldots,\,X_n$, а через $S_i = S_i(\vec{Y}_n)$ — ранг элемента Y_i случайной выборки $Y_1,\,\ldots\,,Y_n$.

Ранговым коэффициентом корреляции Спирмена назовем случайную величину

$$\rho(\vec{X}_n, \vec{Y}_n) = \frac{\sum\limits_{i=1}^{n}(R_i - \overline{R})(S_i - \overline{S})}{\sqrt{\sum\limits_{i=1}^{n}(R_i - \overline{R})^2 \sum\limits_{i=1}^{n}(S_i - \overline{S})^2}}, \tag{5.19}$$

где

$$\overline{R} = \frac{1}{n}\sum_{i=1}^{n} R_i, \qquad \overline{S} = \frac{1}{n}\sum_{i=1}^{n} S_i.$$

Статистика (5.19) является выборочным коэффициентом корреляции последовательностей рангов $R_1, \ldots, R_n$ и $S_1, \ldots, S_n$.

Согласно определению рангов R_i, S_i, $i = \overline{1, n}$,

$$\overline{R} = \overline{S} = \frac{1}{n}\sum_{i=1}^{n} i = \frac{n+1}{2},$$

и можно показать, что

$$x_1 < x_2 < \ldots < x_n. \tag{5.20}$$

В этом случае реализация r_i ранга R_i равна i, $i = \overline{1, n}$, и значение $\rho(\vec{x}_n, \vec{y}_n)$ статистики $\rho(\vec{X}_n, \vec{Y}_n)$ можно вычислить по формуле

Без ограничения общности можно считать, что значения пар наблюдений (x_i, y_i), $i = \overline{1, n}$, занумерованы в порядке возрастания их первых элементов, т.е. так, что выполняются неравенства (5.21)

$$\rho(\vec{X}_n, \vec{Y}_n) = 1 - \frac{6}{n(n^2-1)}\sum_{i=1}^{n}(R_i - S_i)^2.$$

$$\rho(\vec{x}_n, \vec{y}_n) = 1 - \frac{6}{n(n^2-1)}\sum_{i=1}^{n}(i - s_i)^2,$$

где s_i — реализация ранга S_i, $i = \overline{1, n}$.

Можно показать, что при истинности основной гипотезы (5.18)

$$\mathbf{M}\rho(\vec{X}_n, \vec{Y}_n) = 0, \qquad \mathbf{D}\rho(\vec{X}_n, \vec{Y}_n) = \frac{1}{n-1}, \tag{5.22}$$

и, следовательно, при этом *выборочные значения* статистики $\rho(\vec{X}_n, \vec{Y}_n)$ невелики и группируются около нуля. Поэтому (и это кажется достаточно естественным) ранговый критерий Спирмена отклоняет H_0 на *уровне значимости* α, если

$$|\rho(\vec{x}_n, \vec{y}_n)| > \rho_{1-\alpha/2},$$

где $\rho_{1-\alpha/2}$ — квантиль уровня $1-\alpha/2$ распределения случайной величины $\rho(\vec{X}_n, \vec{Y}_n)$ при истинности основной гипотезы (5.18). При небольших n это распределение табулировано*. Известно, что при $n \to \infty$ и при истинности основной гипотезы (5.18)

$$\lim_{n\to\infty} \mathbf{P}\left\{ \frac{\rho(\vec{X}_n, \vec{Y}_n) - \mathbf{M}\rho(\vec{X}_n, \vec{Y}_n)}{\sqrt{\mathbf{D}\rho(\vec{X}_n, \vec{Y}_n)}} < t \right\} = \Phi_0(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{t} e^{-\frac{u^2}{2}}\, du,$$

т.е. квантили случайной величины $\rho(\vec{X}_n, \vec{Y}_n)$ можно приближенно вычислять при помощи таблиц квантилей стандартного нормального распределения.

Пример 5.6. В табл. 5.1 представлены $n = 10$ значений (x_i, y_i), $i = \overline{1, 10}$, непрерывной двумерной случайной величины (X, Y). Проверим на уровне значимости $\alpha = 0{,}05$ гипотезу H_0 о независимости случайных величин X и Y.

Таблица 5.1

x_i	−1,63	1,11	1,15	−1,93	0,38	−1,08	−0,31	0,60	0,12	0,92
y_i	0,54	0,88	−1,21	0,89	−0,64	−0,21	0,08	−0,74	0,79	0,14

Строим последовательность рангов (табл. 5.2). По формуле (5.20) вычисляем реализацию статистики $\rho(\vec{X}_n, \vec{Y}_n)$

$$\rho(\vec{x}_n, \vec{y}_n) = 1 - \frac{6}{10(10^2-1)}\Big((2-7)^2 + (9-9^2) + (10-1)^2 + (1-10)^2 +$$

$$+ (6-3)^2 + (3-4)^2 + (4-5)^2 + (7-2)^2 + (5-8)^2 + (8-6)^2\Big) =$$

$$= 1 - \frac{6}{990}\Big(25 + 0 + 81 + 81 + 9 + 1 + 1 + 25 + 9 + 2\Big) \approx -0{,}4118.$$

Таблица 5.2

r_i	2	9	10	1	6	3	4	7	5	8
s_i	7	9	1	10	3	4	5	2	8	6

По таблицам распределения статистики $\rho(\vec{X}_n, \vec{Y}_n)$ рангового критерия Спирмена* находим квантили

$$\rho_{0,952} = 0{,}6726, \qquad \rho_{0,97} = 0{,}7374, \qquad \rho_{0,983} = 0{,}80223, \quad (5.23)$$

а квантили $\rho_{1-\alpha/2} = \rho_{0,975}$ нет, так как $\rho(\vec{X}_n, \vec{Y}_n)$ — дискретная случайная величина. Тем не менее, из значений квантилей (5.23) заключаем, что $|\rho(\vec{x}_n, \vec{y}_n)| < \rho_{0,952}$ и H_0 не отклоняется даже на большем уровне значимости.

Таблицы сопряженности признаков и критерий χ^2. Пусть имеется случайная выборка

$$(\vec{X}_n, \vec{Y}_n) = \big((X_1, Y_1), \ldots, (X_n, Y_n)\big)$$

из генеральной совокупности двумерной дискретной случайной величины (X, Y), где случайная величина X может принимать значения $u_1, \ldots, u_r$, а случайная величина Y — значения $v_1, \ldots, v_s$. Определим случайную величину $n_{ij}(\vec{X}_n, \vec{Y}_n)$, реализация n_{ij} которой равна количеству элементов выборки $(\vec{x}_n, \vec{y}_n) = \big((x_1, y_1), \ldots, (x_n, y_n)\big)$, совпадающих с элементом (u_i, v_j), $i = 1, r$, $j = 1, s$.

Введем случайные величины $n_{i\cdot}(\vec{X}_n, \vec{Y}_n)$ и $n_{\cdot j}(\vec{X}_n, \vec{Y}_n)$, значения $n_{i\cdot}$ и $n_{\cdot j}$ которых определим по формулам

$$n_{i\cdot} = \sum_{j=1}^{n} n_{ij}, \quad n_{\cdot j} = \sum_{i=1}^{n} n_{ij}.$$

При этом $n_{i\cdot}$ — количество элементов выборки $(\vec{x}_n, \vec{y}_n)$, в которых встретилось значение u_i, а $n_{\cdot j}$ — количество элементов выборки $(\vec{x}_n, \vec{y}_n)$, в которых встретилось значение v_j. Кроме того, имеют место очевидные равенства

$$\sum_{i=1}^{r} n_{i\cdot} = \sum_{j=1}^{s} n_{\cdot j} = \sum_{i=1}^{r}\sum_{j=1}^{s} n_{ij} = n.$$

В рассматриваемом случае результаты наблюдений удобно оформлять в виде таблицы, называемой ***таблицей сопряженности признаков*** (табл. 5.3).

Таблица 5.3

X	Y				
	v_1	v_2	...	v_s	
u_1	n_{11}	n_{12}	...	n_{1s}	$n_{1\cdot}$
u_2	n_{21}	n_{22}	...	n_{2s}	$n_{2\cdot}$
...	...	...	...	...	...
u_r	n_{r1}	n_{r2}	...	n_{rs}	$n_{r\cdot}$
	$n_{\cdot 1}$	$n_{\cdot 2}$	...	$n_{\cdot s}$	n

Пусть далее

$$p_{ij} = \mathbf{P}\{X = u_i, Y = v_j\}, \quad p_{i\cdot} = \mathbf{P}\{X = u_i\}, \quad p_{\cdot j} = \mathbf{P}\{Y = v_j\},$$
$$i = \overline{1, r}, \quad j = \overline{1, s}.$$

Дискретные случайные величины X и Y независимы тогда и только тогда, когда

$$\mathbf{P}\{X = u_i, Y = v_j\} = \mathbf{P}\{X = u_i\}\,\mathbf{P}\{Y = v_j\}, \quad i = \overline{1, r}, \quad j = \overline{1, s}.$$

Поэтому основную гипотезу о независимости дискретных случайных величин X и Y можно представить в следующем виде:

$$H_0\colon\ p_{ij} = p_{i\cdot}p_{\cdot j},\ \ i = 1, r,\ \ j = 1, s. \tag{5.24}$$

При этом, как правило, в качестве альтернативной используют гипотезу

$$H_1\colon\ p_{ij} \neq p_i p_j \ \text{ для некоторых } i = \overline{1, r},\ j = \overline{1, s}. \tag{5.25}$$

Для проверки основной гипотезы (5.24) при альтернативной гипотезе (5.25) К. Пирсон предложил использовать статистику $\widehat{\chi}^2(\vec{X}_n, \vec{Y}_n)$, называемую ***статистикой Фишера — Пирсона***, реализация $\widehat{\chi}^2(\vec{x}_n, \vec{y}_n)$ которой определяется формулой

$$\widehat{\chi}^2(\vec{x}_n, \vec{y}_n) = n \sum_{i=1}^{r} \sum_{i=1}^{s} \frac{\left(n_{ij} - \dfrac{n_{i\cdot} n_{\cdot j}}{n}\right)^2}{n_{i\cdot} n_{\cdot j}}. \tag{5.26}$$

Из закона больших чисел следует, что при $n \to \infty$

$$\frac{n_{ij}(\vec{X}_n, \vec{Y}_n)}{n} \to p_{ij}, \quad \frac{n_{i\cdot}(\vec{X}_n, \vec{Y}_n)}{n} \to p_{i\cdot}, \quad \frac{n_{\cdot j}(\vec{X}_n, \vec{Y}_n)}{n} \to p_{\cdot j},$$

$$i = \overline{1, r}, \quad j = \overline{1, s}.$$

Поэтому при истинности гипотезы H_0 и больших объемах выборки $(\vec{x}_n, \vec{y}_n)$ должно выполняться приближенное равенство

$$n_{ij} \approx n_{i\cdot} n_{\cdot j}, \quad i = \overline{1, r},\ \ j = \overline{1, s},$$

и, следовательно, значения (5.26) статистики $\widehat{\chi}^2(\vec{X}_n, \vec{Y}_n)$ должны быть „не слишком велики“. „Слишком большие“ значения должны свидетельствовать о том, что H_0 неверна.

Ответ на вопрос о том, какие значения нужно считать слишком большими, а какие — нет, дает следующая теорема.

Теорема 5.3. Если истинна гипотеза H_0, то распределение статистики $\widehat{\chi}^2(\vec{X}_n, \vec{Y}_n)$ при $n \to \infty$ слабо сходится к случайной

величине, имеющей χ^2-распределение с числом степеней свободы $k=(r-1)(s-1)$:

$$\lim_{n\to\infty}\mathbf{P}\{\widehat{\chi}^2(\vec{X}_n,\vec{Y}_n)<z\}=\int_0^z\frac{t^{\frac{k}{2}-1}}{2^{\frac{k}{2}}\Gamma\left(\frac{k}{2}\right)}e^{-\frac{t}{2}}\,dt,\quad z>0.\quad\#$$

В соответствии с теоремой 5.3 ***критерий независимости*** χ^2 отклоняет гипотезу H_0 на уровне значимости $1-\alpha$, если

$$\widehat{\chi}^2(\vec{x}_n,\vec{y}_n)>\chi^2_{1-\alpha}((r-1)(s-1)),$$

где $\chi^2_{1-\alpha}((r-1)(s-1))$ — квантиль уровня значимости $1-\alpha$ χ^2-распределения с числом степеней свободы $(r-1)(s-1)$. При этом считается*, что критерий χ^2 можно использовать, если $n_{i\cdot}n_{\cdot j}/n\geqslant 5$.

Правую часть равенства (5.26) можно преобразовать к форме, более удобной для практического использования:

$$\widehat{\chi}^2(\vec{x}_n,\vec{y}_n)=n\Big(\sum_{i=1}^{r}\sum_{j=1}^{s}\frac{n_{ij}^2}{n_{i\cdot}n_{\cdot j}}-1\Big).\tag{5.27}$$

В частном, но очень распространенном случае таблиц сопряженности при $r=s=2$ формула (5.26) для вычисления $\widehat{\chi}^2(\vec{x}_n,\vec{y}_n)$ имеет еще более простой вид:

$$\widehat{\chi}^2(\vec{x}_n,\vec{y}_n)=\frac{n(n_{11}n_{22}-n_{12}n_{21})^2}{n_{1\cdot}n_{2\cdot}n_{\cdot 1}n_{\cdot 2}}.\tag{5.28}$$

Заметим, что для таблиц сопряженности при $r=s=2$, как правило, используют статистику $\widetilde{\chi}^2(\vec{X}_n,\vec{Y}_n)$ с реализациями

$$\widetilde{\chi}^2(\vec{x}_n,\vec{y}_n)=\frac{\big(n|n_{11}n_{22}-n_{12}n_{21}|-n/2\big)^2}{n_{1\cdot}n_{2\cdot}n_{\cdot 1}n_{\cdot 2}},\tag{5.29}$$

называемую ***статистикой Фишера — Пирсона с поправкой Йейтса на непрерывность***, распределение которой лучше согласуется с χ^2-распределением.

Пример 5.7. В табл. 5.4 приведены результаты 145 наблюдений двумерного дискретного случайного вектора (X, Y). Проверим на уровне $\alpha = 0{,}05$ гипотезу H_0 о независимости случайных величин X и Y.

Таблица 5.4

X	Y			
	3	4	5	
0	45	25	15	85
1	11	11	13	35
2	9	9	7	25
	65	45	35	145

В рассматриваемом случае $r = 3$, $s = 3$, т.е. случайные величины X и Y принимают по три различных значения. Вычислим по формуле (5.27) значение $\widehat{\chi}^2(\vec{x}_n, \vec{y}_n)$ величины $\widehat{\chi}^2(\vec{X}_n, \vec{Y}_n)$:

$$\widehat{\chi}^2(\vec{x}_n, \vec{y}_n) = 145\left(\frac{45^2}{65\cdot 85} + \frac{25^2}{45\cdot 85} + \frac{15^2}{35\cdot 85} + \frac{11^2}{65\cdot 35} + \right.$$

$$\left. + \frac{11^2}{45\cdot 35} + \frac{13^2}{35\cdot 35} + \frac{9^2}{65\cdot 25} + \frac{9^2}{45\cdot 25} + \frac{7^2}{35\cdot 25} - 1\right) =$$

$$= 145\Big(0{,}3665 + 0{,}1634 + 0{,}0756 + 0{,}0532 +$$

$$+ 0{,}0768 + 0{,}1380 + 0{,}0498 + 0{,}072 + 0{,}056 - 1\Big) =$$

$$= 145 \cdot 0{,}0513 = 7{,}4385.$$

По таблице квантилей χ^2-распределения (см. табл. П.3) с числом степеней свободы $(r-1)(s-1) = 4$ находим

$$\chi^2_{1-\alpha}((r-1)(s-1)) = \chi^2_{0{,}95}(4) = 9{,}49.$$

Таким образом, оснований для отклонения гипотезы H_0 о независимости случайных величин X и Y недостаточно.