Проверка статистических гипотез

1. Понятие статистической гипотезы
2. Статистический критерий проверки нулевой гипотезы
3. Определение критической области
4. Мощность критерия
5. Примеры проверки статистических гипотез

1. Статистической называют гипотезу о виде неизвестного распределения или о параметрах известных распределений.

Примеры гипотез:

- математическое ожидание нормально распределенной случайной величины равно 10;
- элемент откажет в следующем испытании;
- генеральная совокупность распределена по закону Пуассона;
- дисперсии двух генеральных совокупностей равны;
- на Марсе есть жизнь.

Все ли они являются статистическими?

Наряду с выдвинутой гипотезой рассматривают и противоречащую ей гипотезу. Если выдвинутая гипотеза при проверке будет отвергнута, то имеет место противоречащая гипотеза.

Нулевой (основной) называется выдвинутая гипотеза H_0.

Конкурирующей (альтернативной) называется гипотеза H_1, которая противоречит основной.

Нулевая гипотеза $H_0: a = 10$. *В чем состоит альтернативная?*

Простой называют гипотезу, содержащую только одно предположение.

Сложной называют гипотезу, которая состоит из конечного или бесконечного числа простых гипотез.

Примеры:

$H_0: \lambda = 10$, где λ – параметр показательного распределения;

$H_0: \lambda > 10$, λ – параметр показательного распределения;

$H_0: a = 10$, где a – параметр нормального распределения (σ известно);

$H_0: a = 10$, a – параметр нормального распределения (σ неизвестно);

Выдвинутая гипотеза может быть правильной или неправильной, поэтому возникает необходимость ее проверки. Поскольку проверку производят статистическими методами, ее называют *статистической*.

В итоге статистической проверки гипотезы в двух случаях могут быть допущены ошибки двух родов, последствия которых важно анализировать.

Ошибка первого рода – будет отвергнута правильная гипотеза.

Ошибка второго рода – будет принята неправильная гипотеза.

Правильное решение может быть принято тоже в двух случаях:
1) гипотеза принимается, причем и в действительности она правильная;
2) гипотеза отвергается, причем и в действительности она неверна.

Вероятность совершить ошибку первого рода называется *уровнем значимости* α.

Наиболее часто уровень значимости принимают равным 0,05 или 0,01.

---

2. Статистический критерий проверки нулевой гипотезы

Для проверки нулевой гипотезы используют специально подобранную случайную величину, распределение которой известно:

- *Z*, если она распределена нормально
- *F* – по закону Фишера-Снедекора
- *T* – по закону Стьюдента
- χ^2 – по закону «хи квадрат»

Пусть эта величина обозначена в целях общности через *K*.

Статистическим критерием (*критерием*) называется случайная величина *K*, которая служит для проверки нулевой гипотезы.

Пример 2.1 При проверке гипотезы о равенстве дисперсий двух нормальных генеральных совокупностей в качестве критерия *K* принимается отношение исправленных выборочных дисперсий:

$$F = \frac{s_1^2}{s_2^2}$$

Наблюдаемым значением $K_{\text{набл}}$ называется значение критерия, вычисленное по выборке.

После выбора критерия множество его значений разбивается на два непересекающихся подмножества: одно из них содержит значения критерия, при которых нулевая гипотеза отвергается, другое – при которых принимается

Критической областью называется совокупность значений критерия, при которых нулевую гипотезу отвергают.

Область принятия гипотезы (область допустимых значений) – совокупность значений критерия, при которых гипотезу принимают.

Критические точки (границы) $k_{\text{кр}}$ – точки, разделяющие критическую область и область принятия гипотезы.

Правосторонней называют критическую область, определяемую неравенством $K > k_{\text{кр}}$ где $k_{\text{кр}}$ – положительное число.

Левосторонней называют критическую область, определяемую неравенством $K < k_{\text{кр}}$ где $k_{\text{кр}}$ – отрицательное число.

Односторонней называют правостороннюю или левостороннюю критическую область.

Двусторонней называют критическую область, определяемую неравенствами $K < k_1,\ K > k_2,\ k_1 < k_2$.

Как определяется критическая область, если критические точки симметричны относительно нуля?

Основной принцип проверки статистических гипотез: если наблюдаемое значение критерия принадлежит критической области, то гипотезу отвергают, если наблюдаемое значение критерия принадлежит области принятия гипотезы – гипотезу принимают.

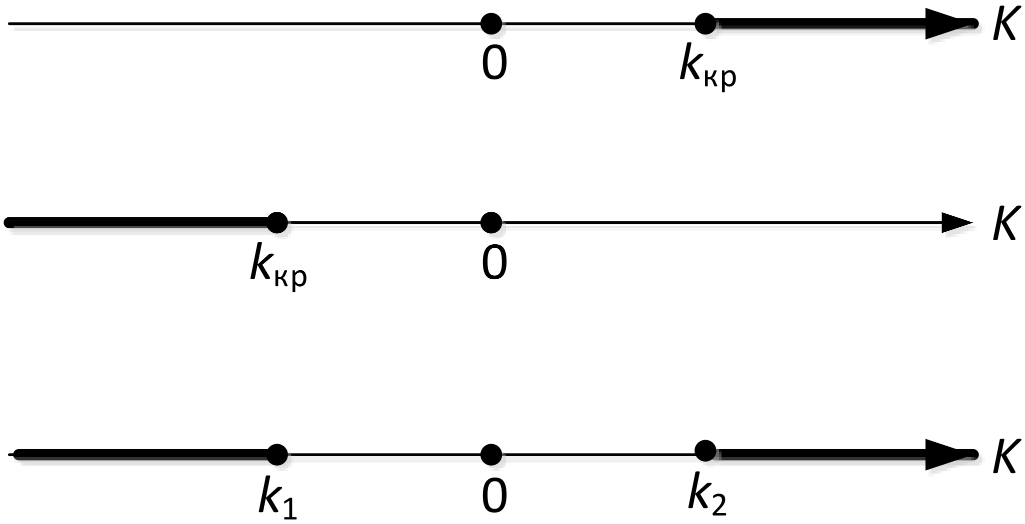

Рисунок 2.1 – Правосторонняя, левосторонняя и двусторонняя критические области

3. Определение критической области

Пусть надо найти правостороннюю критическую область. Для этого:

1) задается уровень значимости α.

2) по таблицам определяется критическая точка, исходя из требования: при условии справедливости нулевой гипотезы вероятность того, что критерий *K* примет значение, большее $k_{кр}$, равна уровню значимости:

$$P(K > k_{кр}) = \alpha$$

3) когда критическая точка найдена, вычисляется наблюдаемое значение критерия по данным выборок. Если $K_{набл} > k_{кр}$, нулевая гипотеза отвергается, в противном случае нет оснований ее отвергнуть.

Наблюдаемое значение критерия может оказаться большим $k_{кр}$ в силу ряда причин (малый объем выборки, недостатки метода проведения эксперимента). В этом случае, отвергнув правильную нулевую гипотезу, совершают ошибку первого рода. Чему равна вероятность этой ошибки?

Приняв нулевую гипотезу, ошибочно полагать, что она доказана, поэтому корректнее говорить «данные наблюдений согласуются с нулевой гипотезой». Отвергают гипотезу более категорично, чем принимают.

Определение левосторонней критической области проводится аналогично, взяв за основу выражение

$$P\left(K < k_{\text{кр}}\right) = \alpha$$

Двусторонняя критическая область определяется двумя неравенствами. Критические точки находятся исходя из требования:

$$P(K < k_1) + P(K > k_2) = \alpha$$

Если распределение критерия симметрично относительно нуля, то

$$P\left(K < -k_{\text{кр}}\right) = P\left(K > k_{\text{кр}}\right)$$

$$P\left(|K| > k_{\text{кр}}\right) = \frac{\alpha}{2}$$

Критические точки находятся по соответствующим таблицам.

4. Мощностью критерия называется вероятность попадания критерия в критическую область при условии справедливости альтернативной гипотезы (вероятность того, что нулевая гипотеза будет отвергнута, если верна альтернативная гипотеза).

Пусть для проверки гипотезы принят определенный уровень значимости и выборка имеет фиксированный объем. Обозначим вероятность ошибки второго рода через β, тогда мощность равна $1-\beta$. Если мощность возрастает, вероятность β совершить ошибку второго рода уменьшается. Таким образом, чем мощность больше, тем вероятность ошибки второго рода меньше.

Критическую область следует строить так, чтобы мощность критерия была максимальной.

Мощность критерия – вероятность того, что не будет допущена ошибка второго рода.

Очевидно, что чем меньше вероятности ошибок первого и второго рода, тем критическая область «лучше». Однако при заданном объеме выборки если уменьшить α, то β будет возрастать. Например, при $\alpha = 0$ будут приниматься все гипотезы, в том числе и неправильные, то есть возрастает вероятность β ошибки второго рода.

Выбор α зависит от «тяжести последствий» ошибок для каждой конкретной задачи. Например, если ошибка первого рода повлечет большие потери, а второго рода – малые, следует принять возможно меньшее α.

Увеличение объема выборок является способом одновременного уменьшения вероятностей ошибок первого и второго рода.

5.1 Сравнение двух дисперсий нормальных генеральных совокупностей

Пусть генеральные совокупности *X* и *Y* распределены нормально. По извлеченным из этих совокупностей независимым выборкам с объемами, соответственно равными n_1 и n_2, найдены исправленные выборочные дисперсии s_1^2 и s_2^2. Требуется по исправленным дисперсиям при заданном уровне значимости α проверить нулевую гипотезу:

$$H_0: D(X) = D(Y)$$

В качестве критерия проверки нулевой гипотезы о равенстве генеральных дисперсий принимается отношение бóльшей исправленной дисперсии к меньшей, то есть случайная величина

$$F = \frac{s_1^2}{s_2^2}$$

Величина *F* при условии справедливости нулевой гипотезы имеет распределение Фишера – Снедекора со степенями свободы $k_1 = n_1 - 1$, $k_2 = n_2 - 1$, где n_1 и n_2 – соответственно объемы выборок, по которым вычислены бóльшая и меньшая исправленные дисперсии.

Критическая область строится в зависимости от вида конкурирующей гипотезы.

а) конкурирующая гипотеза $H_1: D(X) > D(Y)$

Строится <u>правосторонняя</u> критическая область, исходя из требования, чтобы вероятность попадания критерия *F* в эту область в предположении справедливости нулевой гипотезы была равна уровню значимости α:

$$P\left[F > F_{\text{кр}}(\alpha, k_1, k_2)\right] = \alpha$$

Критическая область определяется неравенством $F > F_{\text{кр}}$, а область принятия нулевой гипотезы – неравенством $F < F_{\text{кр}}$.

Критическая точка $F_{\text{кр}}(\alpha, k_1, k_2)$ находится по таблице критических точек распределения Фишера – Снедекора.

По данным выборок находится значение $F_{\text{набл}}$.

Если $F_{\text{набл}} < F_{\text{кр}}$, нет оснований отвергнуть нулевую гипотезу, в противном случае гипотеза отвергается.

Если отношение ${s_1^2}/{s_2^2}$ больше критической точки, то дисперсии различаются между собой с вероятностью $p = 1 - \alpha$.

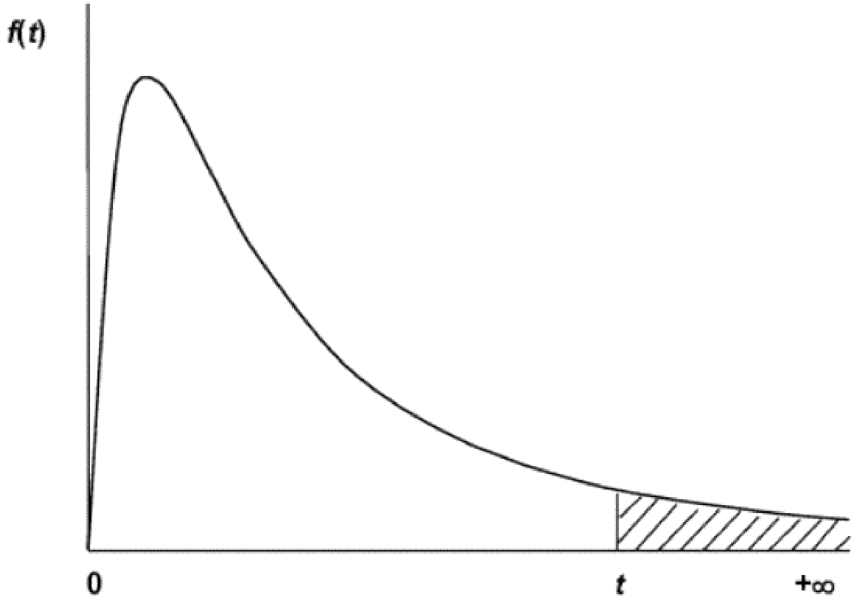

Рисунок 5.1 – Критическая точка распределения Фишера

б) конкурирующая гипотеза $H_1: D(X) \neq D(Y)$

Строится двусторонняя критическая область, исходя из требования, чтобы вероятность попадания критерия в эту область в предположении справедливости нулевой гипотезы была равна уровню значимости α.

Левая F_1 и правая F_2 границы определяются из соотношения:

$$P(F < F_1) = \frac{\alpha}{2}, \qquad P(F > F_2) = \frac{\alpha}{2}$$

Вероятности попадания критерия как в «правую», так и «левую» части критической области равны α/2. Вероятность попадания критерия во всю двустороннюю критическую область равна α/2+α/2=α. Почему?

Для проверки гипотезы H_0 при конкурирующей гипотезе $H_1: D(X) \neq D(Y)$ достаточно найти правую критическую точку $F_2 = F_{\text{кр}}(\frac{\alpha}{2}, k_1, k_2)$.

Если $F_{\text{набл}} < F_{\text{кр}}$, нет оснований отвергнуть нулевую гипотезу, в противном случае гипотеза отвергается.

Пример 5.1. По двум независимым выборкам объемов $n_1 = 12$, $n_2 = 15$, извлеченным из нормальных генеральных совокупностей *X* и *Y*, найдены исправленные выборочные дисперсии $s_1^2 = 11{,}41$, $s_2^2 = 6{,}52$. При уровне значимости 0,05 проверить нулевую гипотезу $H_0 \colon D(X) = D(Y)$ при конкурирующей гипотезе а) $H_1 \colon D(X) > D(Y)$; б) $H_1 \colon D(X) \neq D(Y)$.

Решение.

$$F_{\text{набл}} = \frac{s_1^2}{s_2^2} = \frac{11{,}41}{6{,}52} = 1{,}75$$

Уровень значимости $\alpha = 0{,}05$, степени свободы: $k_1 = 11,\ k_2 = 14$

а) критическая точка для уровня значимости $\alpha = 0{,}05$:

$$F_{\text{кр}}(0{,}05; 11; 14) = 2{,}565$$

б) критическая точка для уровня значимости $\alpha = 0{,}025$:

$$F_{\text{кр}}(0{,}025; 11; 14) = 3{,}095$$

В обоих случаях $F_{\text{набл}} < F_{\text{кр}} \Longrightarrow$ нет оснований отвергнуть гипотезу H_0.

Решить задачу при $s_1^2 = 17{,}8$.

5.2 Проверка гипотезы о значимости выборочного коэффициента корреляции

Пусть двумерная генеральная совокупность *(X, Y)* распределена нормально. Из этой совокупности извлечена выборка объема *n* и по ней найден выборочный коэффициент корреляции $r_{xy} \neq 0$.

Так как выборка сформирована случайным образом, это еще не означает, что коэффициент корреляции генеральной совокупности $\rho_{xy} \neq 0$

При заданном уровне значимости α проверить нулевую гипотезу:

$$H_0: \rho_{xy} = 0$$

при конкурирующей гипотезе:

$$H_1: \rho_{xy} \neq 0$$

Если нулевая гипотеза отвергается, то выборочный коэффициент корреляции значимо отличается от нуля (или значим), а *X* и *Y* коррелированы, то есть связаны линейной зависимостью.

В качестве критерия проверки нулевой гипотезы принимается случайная величина

$$T = \frac{r_{xy}\sqrt{n-2}}{\sqrt{1-{r_{xy}}^2}}$$

Величина *T* при справедливости нулевой гипотезы имеет распределение Стьюдента с $k = n - 2$ степенями свободы.

Поскольку конкурирующая гипотеза имеет вид $\rho_{xy} \neq 0$, критическая область – двусторонняя.

Критическая точка $t_{\text{кр}}(\alpha, k)$ для заданного уровня значимости α и числа степеней свободы $k = n - 2$ определяется по таблице критических точек распределения Стьюдента.

По выборочным данным вычисляется значение критерия $T_{\text{набл}}$.

Если $|T_{\text{набл}}| < t_{\text{кр}}$, то нет оснований отвергнуть нулевую гипотезу, в противном случае гипотеза отвергается.

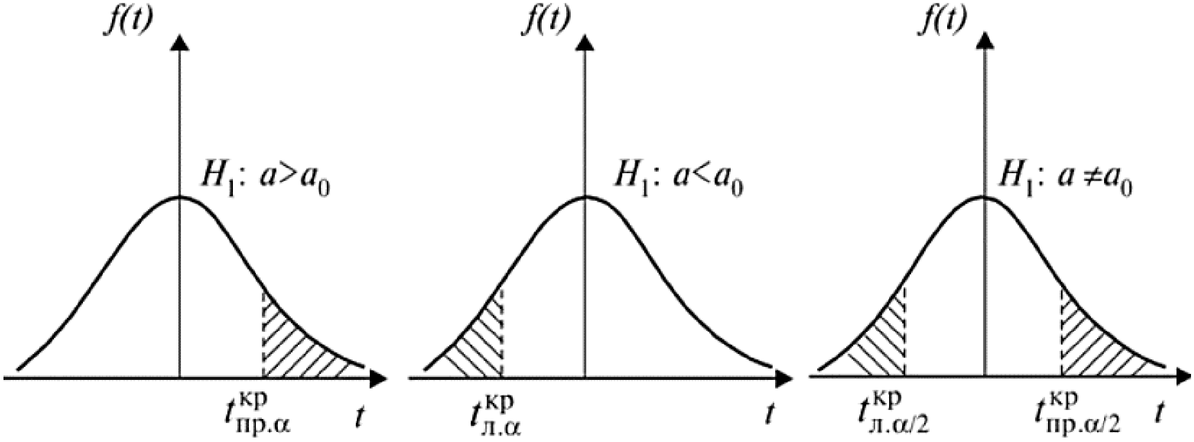

Рисунок 5.2 – Распределение Стьюдента (правосторонняя, левосторонняя и двусторонняя критическая область)

Пример 5.2 По выборке объема $n = 122$, извлеченной из нормальной двумерной совокупности, найден выборочный коэффициент корреляции $r_{xy} = 0{,}4$. При $\alpha = 0{,}05$ проверить нулевую гипотезу $H_0 \colon \rho_{xy} = 0$ при конкурирующей гипотезе $H_1 \colon \rho_{xy} \neq 0$.

Решение.

$$T_{\text{набл}} = \frac{r_{xy}\sqrt{n-2}}{\sqrt{1-{r_{xy}}^2}} = \frac{0{,}4 \cdot \sqrt{122-2}}{\sqrt{1-0{,}16}} = 4{,}78$$

Уровень значимости $\alpha = 0{,}05$, число степеней свободы $k = 120$.

$$t_{\text{кр}}(0{,}05; 120) = 1{,}98$$

Так как $T_{\text{набл}} > t_{\text{кр}}$, нулевая гипотеза отвергается.

Иначе говоря, выборочный коэффициент корреляции r_{xy} значимо отличается от нуля. Это означает, что величины *X* и *Y* коррелированы.

5.3 Проверка гипотезы о нормальном распределении генеральной совокупности. Критерий согласия Пирсона

Критерием согласия называют критерий проверки гипотезы о предполагаемом законе неизвестного распределения.

Пусть по выборке объема *n* получено эмпирическое распределение. Допустим, что в предположении нормального распределения генеральной совокупности вычислены теоретические частоты n_i'

Варианты, x_i	x_1	x_2	$\ldots$	x_s
Эмпирические частоты, n_i	n_1	n_2	$\ldots$	n_s
Теоретические частоты, n_i'	n_1'	n_2'	$\ldots$	n_s'

При уровне значимости *α* требуется проверить нулевую гипотезу H_0: генеральная совокупность распределена нормально.

В качестве критерия проверки используется случайная величина

$$\chi^2 = \sum_{i=1}^{s} \frac{(n_i - n_i')^2}{n_i'}$$

При $n \to \infty$ закон распределения указанной случайной величины стремится к закону распределения χ^2 с k степенями свободы независимо от того, какому закону распределения подчинена генеральная совокупность. Поэтому критерий называется критерием согласия «хи квадрат».

Число степеней свободы определяется по формуле:

$$k = s - 1 - r$$

s – число групп (частичных интервалов) выборки,

r – число параметров предполагаемого распределения.

Строится правосторонняя критическая область исходя из требования, чтобы вероятность попадания критерия в эту область в предположении справедливости нулевой гипотезы была равна α:

$$P\left[\chi^2 > \chi^2{}_{\text{кр}}(\alpha, k)\right] = \alpha$$

Критическая точка $\chi^2{}_{\text{кр}}(\alpha, k)$ для заданного уровня значимости α и числа степеней свободы k определяется по таблице критических точек распределения χ^2.

По выборочным данным вычисляется значение критерия $\chi^2{}_{\text{набл}}$.

Если $\chi^2{}_{\text{набл}} < \chi^2{}_{\text{кр}}$, нет оснований отвергнуть нулевую гипотезу, в противном случае нулевая гипотеза отвергается.

<u>Замечания</u>

1. Объем выборки должен быть достаточно велик (не менее 50). Каждая группа должна содержать не менее 5-8 вариант; малочисленные группы следует объединять в одну, суммируя частоты.

2. Поскольку возможны ошибки первого и второго рода, особенно если теоретические и эмпирические частоты подозрительно близки, следует проявлять осторожность. Можно повторить опыт, увеличить число наблюдений, воспользоваться другими критериями, построить график распределения, вычислить асимметрию и эксцесс.

3. Для контроля вычислений используют формулу

$$\chi^2{}_{\text{набл}} = \left(\sum_{i=1}^{s} \frac{{n_i}^2}{n_i'}\right) - n$$

5.4 Методика вычисления теоретических частот нормального распределения

1. Весь интервал наблюдаемых значений X (выборки объема n) делят на s частичных интервалов (x_i, x_{i+1}) одинаковой длины. Находят середины частичных интервалов

$$x_i^* = \frac{(x_i + x_{i+1})}{2}$$

В качестве частоты n_i варианты x_i^* принимают число вариант, которые попали в *i*-й интервал. Получается последовательность равноотстоящих вариант и соответствующих им частот:

x_1^*	x_2^*	$\ldots$	x_s^*
n_1	n_2	$\ldots$	n_s

$$\sum_{i=1}^{s} n_i = n$$

2. Вычисляют выборочную среднюю $\bar{x}^*$ и выборочное среднее квадратическое отклонение σ^*.

3. Нормируют случайную величину X:

$$Z = \frac{(X - \bar{x}^*)}{\sigma^*}$$

и вычисляют концы интервалов (z_i, z_{i+1}):

$$z_i = \frac{(x_i - \bar{x}^*)}{\sigma^*}, \qquad z_{i+1} = \frac{(x_{i+1} - \bar{x}^*)}{\sigma^*},$$

причем её наименьшее значение $z_1 = -\infty$, а наибольшее $z_s = \infty$,

4. Вычисляют вероятности p_i попадания X в интервалы (x_i, x_{i+1}):

$$p_i = \Phi(z_{i+1}) - \Phi(z_i)$$

где Φ – функция Лапласа

5. Находят искомые теоретические частоты:

$$n_i' = np_i$$

Пример 5.3 Найти теоретические частоты по заданному интервальному распределению выборки объема *n* = 200, предполагая, что генеральная совокупность распределена нормально.

Номер интервала	**Границы интервала**		**Частота**	**Номер интервала**	**Границы интервала**		**Частота**
i	x_i	x_{i+1}	n_i	i	x_i	x_{i+1}	n_i
1	4	6	15	*6*	14	16	21
2	6	8	26	*7*	16	18	24
3	8	10	25	*8*	18	20	20
4	10	12	30	*9*	20	22	13
5	12	14	26				$\Sigma n_i = 200$

Решение

1. Определение середин частичных интервалов:

x_i^*	5	7	9	11	13	15	17	19	21
n_i	15	26	25	30	26	21	24	20	13

2. Вычисление средней и СКО: $\bar{x}^* = 12{,}63,\ \sigma^* = 4{,}695$

3. Нормирование случайной величины X:

i	x_i	x_{i+1}	$x_i - \bar{x}^*$	$x_{i+1} - \bar{x}^*$	$z_i = \frac{(x_i - \bar{x}^*)}{\sigma^*}$	$z_{i+1} = \frac{(x_{i+1} - \bar{x}^*)}{\sigma^*}$
1	4	6	–	– 6,63	–∞	–1,41
2	6	8	– 6,63	– 4,63	–1,41	–0,99
3	8	10	– 4,63	– 2,63	–0,99	–0,56
4	10	12	– 2,63	– 0,63	–0,56	–0,13
5	12	14	– 0,63	1,37	–0,13	0,29
6	14	16	1,37	3,37	0,29	0,72
7	16	18	3,37	5,37	0,72	1,14
8	18	20	5,37	7,37	1,14	1,57
9	20	22	7,37	–	1,57	+∞

4. Определение вероятностей p_i и теоретических частот n_i':

i	z_i	z_{i+1}	$\Phi(z_i)$	$\Phi(z_{i+1})$	$p_i = \Phi(z_{i+1}) - \Phi(z_i)$	$n_i' = np_i$	n_i
1	–∞	–1,41	– 0,5	–0,421	0,079	15,79	15
2	–1,41	–0,99	–0,421	–0,338	0,083	16,62	26
3	–0,99	–0,56	–0,338	–0,212	0,126	25,13	25
4	–0,56	–0,13	–0,212	–0,053	0,159	31,79	30
5	–0,13	0,29	–0,053	0,115	0,168	33,63	26
6	0,29	0,72	0,115	0,264	0,149	29,76	21
7	0,72	1,14	0,264	0,374	0,110	22,02	24
8	1,14	1,57	0,374	0,442	0,068	13,63	20
9	1,57	+∞	0,442	0,5	0,058	11,65	13

5. Критерий Пирсона:

$$\chi^2_{\text{набл}} = \sum_{i=1}^{s} \frac{(n_i - n_i')^2}{n_i'} = 13{,}07, \qquad \chi^2_{\text{кр}} = 12{,}59$$

$\chi^2_{\text{набл}} < \chi^2_{\text{кр}}$, то есть данные наблюдений согласуются с нулевой гипотезой.