Министерство науки и высшего образования Российской Федерации

Федеральное государственное бюджетное образовательное учреждение
высшего образования
«Оренбургский государственный университет»

Кафедра механики материалов, конструкций и машин

Н.А. Морозов

ПРОВЕРКА СТАТИСТИЧЕСКИХ ГИПОТЕЗ

Методические указания

Рекомендовано к изданию редакционно-издательским советом федерального государственного бюджетного образовательного учреждения высшего образования «Оренбургский государственный университет» для обучающихся по образовательным программам высшего образования по направлениям подготовки 15.04.01 Машиностроение, 20.04.01 Техносферная безопасность

Оренбург
2019

УДК 62-192
ББК 30.40
М 80

Рецензент – кандидат технических наук А.А. Гаврилов

Морозов, Н.А.
М 80 Проверка статистических гипотез: методические указания / Н.А. Морозов; Оренбургский гос. ун-т. – Оренбург: ОГУ, 2019. – 31 с.

Методические указания содержат варианты заданий и пример решения задачи по проверке статистических гипотез.

Методические указания предназначены для практических занятий и самостоятельной работы обучающихся направлений подготовки 15.04.01 Машиностроение, 20.04.01 Техносферная безопасность по дисциплинам «Надежность механических систем», «Методы расчета надежности механических систем», а также будут полезны для самостоятельной работы и практических занятий обучающихся по техническим направлениям бакалавриата и магистратуры.

УДК 62-192
ББК 30.40

Содержание

1 Краткие теоретические сведения

Для обучающихся по направлению подготовки 15.04.01 Машиностроение результаты освоения материала, представленного в методических указаниях, ориентированы на формирование компетенций ОПК-12 (обладать способностью подготавливать научно-технические отчеты, обзоры, публикации по результатам выполненных исследований в области машиностроения) и ПК-4 (обладать способностью подготавливать заявки на изобретения и промышленные образцы, организовывать работы по осуществлению авторского надзора при изготовлении, монтаже, наладке, испытаниях и сдаче в эксплуатацию выпускаемых изделий и объектов машиностроения), для обучающихся по направлению подготовки 20.04.01 Техносферная безопасность – ПК-13 (обладать способностью применять методы анализа и оценки надежности и техногенного риска).

На разных этапах статистического исследования возникает необходимость в формулировании и экспериментальной проверке некоторых предположительных утверждений (гипотез). При установлении законов распределения той или иной случайной величины, либо ее характеристик, как правило, получаем выборку из генеральной совокупности, то есть ограниченную последовательность опытных данных. По выборке нельзя дать точное определение параметров генеральной совокупности, можно только сделать предположение об их значениях, то есть предложить гипотезу. Эти гипотезы проверяются путем обработки выборки случайной выборки и называются статистическими.

Задача проверки гипотез состоит в том, чтобы установить, противоречит принятая гипотеза экспериментальным данным или нет.

Наряду с выдвинутой гипотезой, которая называется нулевой (основной) и обозначается H_0, рассматривают и противоречащую ей гипотезу, которая может быть принята в том случае, если первая не подтвердится. Эту гипотезу называют альтернативной (конкурирующей) гипотезой и обозначают H_1.

Допустим, проверяется предположение о том, что исследуемая характеристика подчиняется конкретному, к примеру, нормальному закону распределения. Данное предположение выдвигается как нулевая гипотеза H_0. Конкурирующих к ней гипотез может быть выдвинуто несколько.

С целью краткости записи гипотез применяют специальное обозначение. Допустим, нулевая гипотеза заключается в предложении, что математические ожидания двух нормально распределенных совокупностей m_x и m_y равны, а альтернативная гипотеза заключается в том, что они не равны. Записать это можно следующим образом:

$$H_0 : m_x = m_y, \ \ H_1 : m_x \neq m_y .$$

Гипотезы делятся на простые и сложные. Простой называют гипотезу, содержащую только одно предположение. Сложной называют гипотезу, содержащую конечное и бесконечное число предложений.

Так как нулевая гипотеза выражает предварительно выбранную точку знания, то есть обязательно следует проверить, не противоречит ли высказанная гипотеза существующей в нашем распоряжении выборке. Данная проверка осуществляется с помощью статистического критерия и называется статистической проверкой гипотез.

Статистический критерий – строгое математическое правило, по которому принимается или отвергается та или иная статистическая гипотеза.

Например, для партии произведенных деталей осуществляется выборочный контроль, по итогам которого следует принять решение: или целую партию рассматривать годной (с возможным уровнем брака), или целую партию рассматривать негодной. За нулевую возьмем гипотезу, что вся партия годна, а за альтернативную – что партия содержит брак выше допустимого.

Из выбранных для контроля n деталей бракованных может оказаться z штук. Контролеры бракуют всю партию, если из выбранных n деталей оказалось k бракованных деталей (или больше) и считают годной всю партию, если бракованных деталей меньше k $(0 \le z < n)$. То есть, если $z < k$, то принимается гипотеза H_0, если $z \ge k$, то гипотеза H_0 отвергается.

В разных выборках n деталей из одной партии значение z будет различным. То есть z является случайной величиной со своим законом распределения.

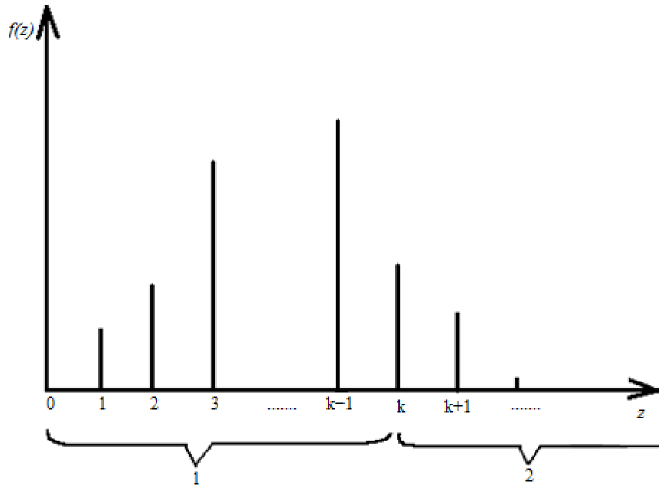
В этом примере величина z может рассматриваться как признак или функция брака. В общем случае такой признак называется показателем согласованности (критической статистикой или статистической характеристикой) [1].

Величина k служит в качестве критерия проверки (критерия согласия, критерия соответствия), то есть свода правил, указывающих, при каких значениях величины показателя согласованности гипотеза отвергается, а при каких не отвергается.

Таким образом, множество возможных значений показателя согласованности z в соответствии с принятым критерием k разбивается на два подмножества ($z<k$ и $z\ge k$, $0\le z \le n$), поэтому попадание возможного значения показателя согласованности критической статистики в подмножество ($z<k$) означает принятие гипотезы H_0, а в другое ($z \ge k$) – отклонение гипотезы H_0.

В общем случае областью отклонения гипотезы (критической областью) называется область, при попадании в которую статистической характеристики z гипотеза H_0 отвергается, а областью принятия – область, при попадании в которую статистической характеристики z нулевая гипотеза H_0 принимается.

Приведенные положения иллюстрированы на рисунке 1, где по оси абсцисс отложена критическая статистика z, а по оси ординат – плотность ее распределения $f(z)$.

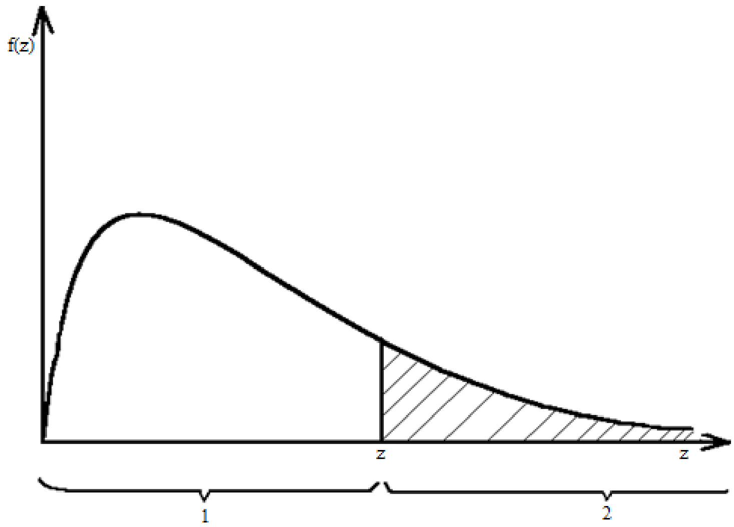


1 – область принятия нулевой гипотезы;

2 – область отклонения нулевой гипотезы.

Рисунок 1– Области отклонения и принятия гипотезы

В случае, если анализировать не дискретные случайные величины, а непрерывные (например, размер детали), получим непрерывную плотность распределения критической статистики (рисунок 2).



1 – область принятия нулевой гипотезы;

2 – область отклонения нулевой гипотезы.

Рисунок 2 – Непрерывная плотность распределения критической статистики

Так как критическая статистика считается случайной величиной, то постоянно имеется риск сформировать настолько неудачную выборку, что сведения, содержащиеся в ней, окажутся абсолютно ошибочными. Теория статистической проверки гипотез дает возможность определить вероятность такой ошибки. Такую вероятность называют уровнем значимости α (на рисунке 2 это площадь заштрихованной фигуры). Тогда область принятия гипотезы будет являться доверительным интервалом для статической характеристики z с доверительной вероятностью $\gamma = 1 - \alpha$.

На первый взгляд может показаться, что необходимо выбирать уровень значимости α всегда как можно меньше, но в действительности это не так. Дело в том, что при проверке гипотезы можно совершить одну из двух ошибок:

– ошибка первого рода (риск поставщика) заключается в отклонении верной гипотезы;

– ошибка второго рода (риск заказчика) заключается в принятии ошибочной гипотезы.

При уменьшении вероятности ошибки первого рода α всегда увеличивается вероятность ошибки второго рода ($1 - \alpha$). На практике обычно принимают α =0,05 или α =0,10.

При решении некоторых задач по проверке статистических гипотез необходимо определить не одностороннее ограничение (рисунок 2), а двухстороннее (рисунок 3). В этом случае всю область значений случайной величины z разделяют на три области: 1 – область неправдоподобно малых значений; 2 – область естественных (вероятных) значений; 3 – область неправдоподобно больших значений. Обычно z_{min} и z_{max} выбирают из условия, чтобы площади слева от z_{min} и справа от z_{max} под кривой плотности распределения $f(z)$ были бы равны между собой и составляли бы $\alpha/2$.

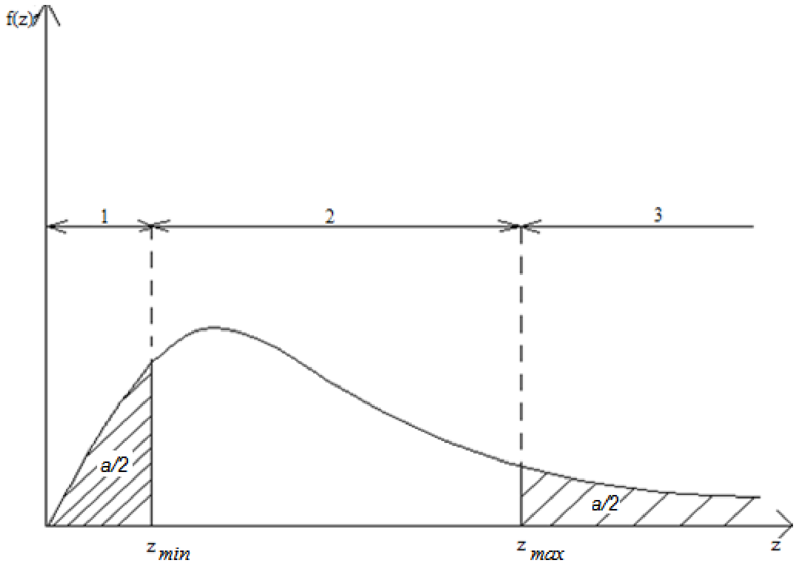


Рисунок 3 – Двухстороннее ограничение критической статистики

Простая статистическая характеристика z не всегда является достаточной характеристикой для решения задач статистической проверки гипотез. В действительности используют специально разработанные статистические характеристики, которые различаются от типа решаемых задач: например, статистические характеристики Стьюдента, Фишера, Пирсона, Смирнова [2].

Все гипотезы можно проверять, используя алгоритм:

1 Выдвигаем предположение (гипотезу) H_0.

2 Задаем величину уровня значимости α.

3 Выбирают критическую статистику z (некоторую функцию от результатов измерений).

4 Из таблиц критической статистики находят 100(1-α/2)-процентную и 100(α/2) – процентную точки критической статистики: $z_{\frac{\alpha}{2}}^{\min}$ и $z_{\frac{\alpha}{2}}^{\max}$ в случае двухстороннего ограничения, или $100(1-\alpha)$ - процентную точку $z_{\alpha}^{\min}$ для одностороннего ограничения.

5 В функцию z подставляют значения выборки и подсчитывают ее численное значение. Если окажется, что значения принадлежит области вероят-

ностных значений, то нулевая гипотеза с доверительной вероятностью $(1-\alpha)$ подтверждается. В противном случае гипотеза не подтверждается.

При решении практических задач часто необходимо определить какому закону распределения соответствует полученная выборка. Рассмотрим метод Н.В. Смирнова. Этот метод применяется, если выборка представлена отдельными экспериментальными точками. Согласно данному методу в качестве меры согласования теоретического и статического законов распределения используется функция разности статической и теоретической функции распределения.

Проверка гипотезы осуществляется в соответствии с алгоритмом:

1 По внешнему виду (на графике) функции распределения выдвигается нулевая гипотеза, что распределение подчиняется какому-то закону $F(x)$ (например, нормальному).

2 Задаем величину уровня значимости α.

3 Вычисляем значение критической статистики по формуле (1):

$$u=\frac{1}{12\cdot n}+\sum_{i=1}^{n}\left[F(x_i)-\frac{2\cdot i-1}{2\cdot n}\right]^2, \qquad (1)$$

где n – количество экспериментальных точек;

$F(x_i)$ – теоретическое значение исследуемой функции распределения при параметрах распределения (m_x и D_x), найденных из эксперимента;

i – порядковый номер наблюдения в вариационном ряду.

Если проверяется нормальный закон распределения, то вместо функции $F(x_i)$ берется табличное значение нормированной нормальной функции распределения (в этой формуле D_x – среднее квадратическое отклонение):

$$F(x_i)=\Phi\left(\frac{x_i-m_x}{D_x}\right).$$

Из таблицы 1 для уровня значимости α находят критерии u_α.

Таблица 1 – Значения критической границы по методу Смирнова

α	0,5	0,1	0,05	0,02	0,01
u_α	0,118	0,347	0,461	0,620	0,744

4 Проверяем условие $u < u_\alpha$. Если оно выполняется, то нулевая гипотеза принимается.

Рассмотрим еще один метод проверки гипотез о законе распределения – метод К. Пирсона. Согласно этому методу мерой расхождения теоретического и статистического законов распределения служит сумма квадратов разностей между частотой и вероятностью попадания случайной величины в интервале, на которые разбиваются множество возможных значений этой величины [3].

Проверка гипотезы применительно к нормальному закону распределения осуществляется в соответствии с алгоритмом:

1 Выдвигаем гипотезу H_0, что статистический закон подчиняется нормальному закону распределения.

2 Выбираем уровень значимости α.

3 Наблюдаемые n значений случайной величины разбиваем на l интервалов и вычисляем показатель согласованности по следующей формуле:

$$u = \sum_{j=1}^{l} \frac{n(h_j - p_j)^2}{p_j}, \tag{2}$$

где h_j – частота попадания случайной величины в j-й интервал;

p_j – вероятность попадания случайной величины в j-й интервал.

Так же можно использовать формулу (3):

$$u = \sum_{j=1}^{l} \frac{m_j^2}{np_j} - n, \tag{3}$$

где m_j – число наблюдений, попавших в j-й интервал.

Вероятность P_j попадания случайной величины x в j-й интервал вычисляется следующим образом:

$$P_j = F(x_{j+1}) - F(x_j), \tag{4}$$

где $F(x)$ – теоретический закон распределения с параметрами m_x и D_x.

Для нормального закона распределения P_j будет равно:

$$P_j = \Phi\left(\frac{x_{j+1} - m_x}{D_x}\right) - \Phi\left(\frac{x_j - m_x}{D_x}\right). \tag{5}$$

4 Критическую границу $u_\alpha = \chi^2$ для нормального закона распределения определяем по таблице квантилей функции распределения Пирсона для (l-3) степеней свободы (приложение Б).

5 Проверяем условие $u<u_\alpha$. Если оно выполняется, то расхождение между экспериментальными данными и гипотезой полагают несущественным. В противном случае гипотеза H_0 отвергается.

Прежде, чем сравнивать между собой две выборки (например, на равенство средних значений), необходимо убедиться, что дисперсия этих двух выборок равны между собой. Такая проверка делается с помощью критерия Фишера по следующему алгоритму:

1 Выдвигаем гипотезу $H_0 : D_x^{\ 2} = D_y^{\ 2}$.

2 Выбираем уровень значимости α.

3 Критической статистикой служит функция F, определяемая по фор-

муле:

$$\text{при } D_x{}^2 > D_y{}^2 \quad F = \frac{D_x{}^2}{D_y{}^2}, \quad \text{при} \quad D_x{}^2 < D_y{}^2 \quad F = \frac{D_y{}^2}{D_x{}^2}. \tag{6}$$

4 Критическое значение статистики F_α соответствует распределению Фишера (приложение В) с $v_1 = n_1 - 1$ и $v_2 = n_2 - 1$ степенями свободы.

5 Проверяем условие $F < F_\alpha$. Если оно выполняется, то гипотеза H_0 принимается. В противном случае гипотеза H_0 отвергается.

На практике часто возникает необходимость сравнить параметры объектов из двух партий. Значения этих параметров будут представлены в виде двух выборок с объемами n_x и n_y. После расчетов получим статистические характеристики m_x, m_y и $D_x{}^2$, $D_y{}^2$.

Проверка гипотезы о равенстве двух средних значений производится следующим образом:

1 Выдвигаем гипотезу H_0: m_x=m_y.

2 Выбираем уровень значимости α .

3 Критической статистикой служит функция:

$$t = \frac{m_x - m_y}{\sqrt{\dfrac{(n_x - 1)\cdot D_x^2 + (n_y - 1)\cdot D_y^2}{n_x + n_y - 2}}} \cdot \sqrt{\frac{n_x \cdot n_y}{n_x + n_y}} . \tag{7}$$

Если объем выборок одинаков, можно использовать формулу:

$$t = \frac{m_x - m_y}{\sqrt{D_x^2 + D_y^2}} \cdot \sqrt{n} . \tag{8}$$

4 Критическое значение статистики t_α ищем по таблице квантилей распределения Стьюдента (приложение Г) с $\nu = n_x + n_y - 2$ степенями свободы для уровня значимости α. При пользовании этой таблицей следует учесть, что вместо уровня значимости α нужно использовать доверительную вероятность $\gamma = 1- \alpha$. То есть, например, отыскать критическое значение t при α = 0,05, можно в столбце с обозначением γ = 0,95.

5 Рассчитанное значение t сравнивается с критическим t_α. При $|t| > t_\alpha$ гипотеза H_0 отвергается. В противном случае говорят, что средние значения выборок значимо не отличаются в статистическом смысле.

На практике при статистических расчетах часто возникает необходимость проверки подозрительных членов выборки, которые могут являться промахами, то есть появиться вследствие ошибки.

Определим наличие промахов с помощью r-критерия по алгоритму:

1 Построим из членов выборки, содержащей n членов, вариационный ряд и выдвинем гипотезу, что один из крайних членов данного ряда x_1 или x_n принадлежит выборке, а не является случайным промахом.

2 Выбираем уровень значимости α .

3 Критической статистикой служит функция:

$$r = \frac{m_x - x_1}{D_x\sqrt{\frac{(n-1)}{n}}} \text{ или } r = \frac{x_n - m_x}{D_x\sqrt{\frac{(n-1)}{n}}}. \qquad (9)$$

4. Критическое значение r_α выбирают по таблице в приложении Д для принятого уровня значимости и числа степеней свободы $\nu = n$-2.

5. Если $r \geq r_\alpha$, гипотеза отвергается, то есть исследуемый член выборки является промахом.

2 Постановка задачи и исходные данные

Даны варианты результатов замеров случайных величин несущей способности *R*(кН) и нагрузки *N* (кН) различных элементов объектов. Убедиться, что крайние члены вариационных рядов нагрузки и несущей способности принадлежат генеральным совокупностям. Также проверить гипотезы о равенстве дисперсий и средних значений нагрузки и несущей способности. Провести проверку гипотезы о нормальном распределении данных выборок. Проверку всех указанных гипотез рекомендуется проводить при уровне значимости 0,05.

Исходные данные.

Вариант 1. *N*: 425, 483, 532, 610, 553, 751, 526, 673, 581, 480; *R*: 664, 766, 688, 843, 711, 769, 700, 774, 722, 747, 721, 805, 739, 744.

Вариант 2. *N*: 831, 956, 1063, 1204, 1110, 1503, 1053, 1357, 1163, 987; *R*: 1349, 1527, 1378, 1681, 1437, 1491, 1413, 1543, 1443, 1487, 1418, 1609, 1473, 1484.

Вариант 3. *N*: 1661, 1842, 2125, 2430, 2202, 3006, 2102, 1702, 2333, 1969; *R*: 2695, 3053, 2754, 3430, 2871, 2983, 2836, 3082, 2881, 2969, 2886, 3202, 2943, 2961.

Вариант 4. *N*: 622, 823, 702, 872, 736, 909, 781, 1012, 781, 1122; *R*: 1016, 1113, 1032 1112, 1063, 1122, 1081, 1146, 1085, 1152, 1083, 1261, 1106, 1201.

Вариант 5. *N*: 1243, 1652, 1411, 1751, 1472, 1803, 1579, 2021, 1592, 2253; *R*: 2034, 2226, 2068, 2221, 2125, 2241, 2162, 2293, 2165, 2312, 2163, 2321, 2213, 2406.

Вариант 6. *N*: 1864, 2472, 2123, 2621, 2202, 2701, 2364, 3035, 2386, 3373; *R*: 3043, 3332, 3091, 3338, 3184, 3365, 3244, 3436, 3242, 2461, 3242, 3783, 3311, 3610.

Вариант 7. *N*: 101, 136, 123, 142, 111, 155, 136, 168, 131, 185; *R*: 162, 181, 173, 182, 171, 186, 185, 199, 171, 192, 181, 210, 187, 213.

Вариант 8. *N*: 223, 252, 322, 272, 333, 284, 372, 295, 415, 306; *R*: 376, 408, 379, 402, 381, 412, 396, 421, 395, 422, 393, 441, 402, 463.

Вариант 9. *N*: 242, 285, 355, 285, 370, 316, 402, 313, 440, 335; *R*: 406, 445, 412, 441, 422, 449, 438, 454, 438, 462, 431, 485, 442, 506.

Вариант 10. *N*: 265, 306, 375, 312, 391, 343, 431, 341, 485, 356; *R*: 433, 382, 445, 481, 465, 482, 461, 498, 467, 501, 461, 523, 475, 544.

Вариант 11. *N*: 281, 332, 403, 431, 422, 362, 471, 373, 521, 386; *R*: 473, 512, 488, 514, 496, 525, 506, 532, 505, 539, 504, 560, 515, 588.

Вариант 12. *N*: 311, 354, 438, 368, 450, 384, 506, 397, 565, 417; *R*: 506, 559, 512, 552, 532, 553, 535, 576, 546, 574, 548, 609, 551, 632.

Вариант 13. *N*: 492, 581, 673, 522, 756, 555, 602, 531, 472, 411; *R*: 741, 731, 802, 723, 743, 722, 775, 706, 747, 719, 847, 684, 765, 673.

Вариант 14. *N*: 989, 1162, 1354, 1055, 1506, 1102, 1201, 1065, 944, 833; *R*: 1483, 1474, 1605, 1444, 1481, 1443, 1546, 1417, 1493, 1434, 1683,1371, 1522, 1345.

Вариант 15. *N*: 1967, 2337, 1707, 2107, 3007, 2207, 2408, 2128, 1898, 1668; *R*: 2968, 2941, 3201, 2881, 2961, 2881, 3087, 2834, 2984, 2874, 3464, 2749, 3059, 2697.

Вариант 16. *N*: 1126, 797, 1018, 789, 906, 736, 877, 709, 823, 629; *R*: 1202, 1106, 1262, 1088, 1154, 1088, 1144, 1089, 1127, 1061, 1111, 1034, 1119, 1019.

Вариант 17. *N*: 2254, 1596, 2028, 1570, 1807, 1477, 1759, 1411, 1653, 1242; *R*: 2410, 2212, 2329, 2165, 2291, 2167, 2249, 2121, 2220, 2062, 2229, 2033, 2317, 2167.

Вариант 18. *N*: 125, 134, 104, 129, 156, 124, 111, 130, 109; *R*: 333, 351, 402, 388, 360, 362, 357, 395, 349, 355.

Вариант 19. *N*: 188, 131, 168, 132, 151, 117, 145, 123, 138, 102; *R*: 215, 173, 210, 170, 183, 169, 181, 182, 189, 171, 180, 179, 181, 161.

Вариант 20. *N*: 303, 411, 281, 379, 281, 333, 272, 322, 256, 224; *R*: 463, 406, 448, 399, 421, 395, 427, 391, 414, 382, 401, 372, 404, 374.

3 Пример выполнения задачи

Нагрузка N (кН): 227, 258, 271, 288, 292, 301, 322, 331, 372, 413; несущая способность R (кН): 371, 379, 388, 396, 395, 406, 460, 408, 409, 411, 419, 424, 441, 397. Наработка до отказа (тыс. циклов): 312, 329, 248, 321, 418, 267, 193, 375, 358.

Убедиться, что крайние члены вариационных рядов нагрузки и несущей способности принадлежат генеральным совокупностям. Также проверить гипотезы о равенстве дисперсий и средних значений нагрузки и несущей способности. Провести проверку гипотезы о нормальном распределении выборки наработки до отказа. Проверку всех указанных гипотез рекомендуется проводить при уровне значимости 0,05.

Решение задачи.

Вариационный ряд нагрузки N (кН): 227, 258, 271, 288, 292, 301, 322, 331, 372, 413.

Выдвигаем нулевую гипотезу, что последний член вариационного ряда нагрузки (N_{10} =413 кН) принадлежит генеральной совокупности. Назначаем уровень значимости α =0,05.

Найдем математическое ожидание, дисперсию и среднее квадратическое отклонение нагрузки:

$$m_N = \frac{\sum_{i=1}^{10} N_i}{n} = \frac{227+258+271+288+292+301+322+331+372+413}{10} = 307{,}5\ кН,$$

$$D_N^2 = \frac{\sum_{i=1}^{10} (N_i - m_N)^2}{n-1} = \frac{6480{,}25+2450{,}25+1332{,}25+380{,}25+240{,}25}{9} +$$
$$+\frac{42{,}25+210{,}25+552{,}25+4160{,}25+11130{,}25}{9} = 2697{,}85\ кН^2,$$

$$D_N = 51{,}94\ кН.$$

Вычисляем характеристику r по формуле (9):

$$r = \frac{413 - 307{,}5}{51{,}94\sqrt{\frac{(10-1)}{10}}} = 2{,}14.$$

По таблице приложения Д находим для числа степеней свободы (10 – 2 = 8) критическое значение $r_\alpha = 2{,}294$.

Поскольку r = 2,14 < $r_\alpha = 2{,}294$, то нулевая гипотеза принимается, и последний член ряда нагрузки не является промахом.

Выдвигаем нулевую гипотезу, что первый член вариационного ряда нагрузки ($N_1 = 227$ кН) принадлежит генеральной совокупности.

Вычисляем характеристику r по формуле (9):

$$r = \frac{307{,}5 - 227}{51{,}94\sqrt{\frac{(10-1)}{10}}} = 1{,}63.$$

Поскольку r = 1,63 < r = 2,294 , то гипотеза принимается, и первый член ряда нагрузки не является промахом.

Вариационный ряд несущей способности R (кН): 371, 379, 388, 395, 396, 397, 406, 408, 409, 411, 419, 424, 441, 460.

Выдвигаем нулевую гипотезу, что последний член вариационного ряда несущей способности ($R_{14} = 460$ кН) принадлежит генеральной совокупности. Назначаем уровень значимости α =0,05. Вычисляем математическое ожидание, дисперсию и среднее квадратическое отклонение:

$$m_R = \frac{\sum_{i=1}^{14} R_i}{n} = \frac{371+379+388+396+395+397+406+460+408+409}{14} +$$

$$+\frac{411+419+424+441+397}{14} = 407,4\ кН,$$

$$D_R^2 = \frac{\sum_{i=1}^{14}(R_i - m_R)^2}{n-1} = \frac{1324,96+806,56+379,36+129,96+153,76+1,96+0,36+2,56}{13} +$$

$$+\frac{12,96+134,56+275,56+1128,96+2766,76+108,16}{13} = 555,88\ кН^2,$$

$$D_R = 23,58\ кН.$$

Вычисляем характеристику r по формуле (9):

$$r = \frac{460-407,4}{23,58\sqrt{\frac{(14-1)}{14}}} = 2,32.$$

По таблице приложения Д находим для числа степеней свободы (14 – 2 = 12) критическое значение r = 2,461.

Поскольку r = 2,32 < r = 2,461 , то нулевая гипотеза принимается, и последний член ряда несущей способности не является промахом.

Выдвигаем нулевую гипотезу, что первый член вариационного ряда несущей способности ($N_1 = 371$ кН) принадлежит генеральной совокупности.

Вычисляем характеристику r по формуле (9):

$$r = \frac{407,4-371}{23,58\sqrt{\frac{(14-1)}{14}}} = 1,61.$$

Поскольку r = 1,61 < r = 2,461, то гипотеза принимается, и первый член ряда нагрузки не является промахом.

Выдвинем гипотезу о равенстве дисперсий двух имеющихся выборок H_0: $D_R^2 = D_N^2$, назначаем уровень значимости $\alpha = 0{,}05$. Вычисляем по формуле (6) статистику F:

$$F = \frac{2697{,}85}{555{,}88} = 4{,}85.$$

По таблице приложения В для $\alpha = 0{,}05$ и чисел степеней свободы $\nu_1 = 10 - 1 = 9$ и $\nu_2 = 14 - 1 = 13$ находим для $\nu_2 = 10$ $F_{0,05}$ = 3,02, для $\nu_2 = 20$ $F_{0,05}$ = 2,39, следовательно, для $\nu_2 = 13$:

$$F_{0,05} = 3{,}02 - \frac{3{,}02 - 2{,}39}{10} \cdot 3 = 2{,}83.$$

Поскольку F = 4,85 > $F_{0,05}$ = 2,83, то гипотеза H_0 отвергается; значения, полученные по выборкам, противоречат утверждению о статистическом равенстве дисперсий.

Проверим гипотезу: значимо ли отличаются средние значения двух выборок при уровне значимости 0,05.

Определяем статистику Стьюдента по формуле (7):

$$t = \frac{407{,}4 - 307{,}5}{\sqrt{\frac{(10-1)\cdot 2697{,}85 + (14-1)\cdot 555{,}88}{10+14-2}}} \cdot \sqrt{\frac{10 \cdot 14}{10+14}} = 7{,}03.$$

Критическое значение t_α при α = 0,05 и числе степеней свободы ν = 22 находим из таблицы (приложение Г) для ν = 20 $t_{0,05}$ = 2,086; для ν = 25 $t_{0,05}$ = 2,06; следовательно, для $\nu = 22$:

$$t_{0,05} = 2{,}086 - \frac{2{,}086 - 2{,}06}{5} \cdot 2 = 2{,}076.$$

Так как t = 7,03 > 2,076 = $t_{0,05}$, то делаем вывод, что средние значения выборок отличаются значимо.

Дан вариационный ряд наработки до отказа (тыс. циклов): 193, 248, 267, 312, 321, 329, 358, 375, 418.

Требуется проверить гипотезу о соответствии этой выборки нормальному закону распределения.

Выбираем уровень значимости 0,05. Рассчитываем математическое ожидание и дисперсию:

$$m_t = \frac{\sum_{i=1}^{9} t_i}{n} = \frac{193 + 248 + 267 + 312 + 321 + 329 + 358 + 375 + 418}{9} =$$
$$= 313{,}4 \text{ } \textit{тыс. циклов},$$

$$D_t^2 = \frac{\sum_{i=1}^{9} (t_i - m_t)^2}{n-1} = \frac{14506{,}85 + 4282{,}97 + 2157{,}08 + 2{,}09 + 57{,}09}{8} +$$
$$+ \frac{241{,}98 + 1985{,}2 + 3789{,}09 + 10931{,}87}{8} = 4744{,}28 \text{ } \textit{тыс. циклов}^2;$$

$$D_t = 68{,}88 \text{ } \textit{тыс. циклов.}$$

По таблице нормального распределения (приложение А) находим теоретические значения функции распределения:

$$\Phi\left(\frac{193 - 313{,}4}{68{,}88}\right) = \Phi(-1{,}75) = 0{,}04006,$$

$$\Phi\left(\frac{248-313{,}4}{68{,}88}\right)=\Phi(-0{,}95)=0{,}1711,$$

$$\Phi\left(\frac{267-313{,}4}{68{,}88}\right)=\Phi(-0{,}67)=0{,}2514,$$

$$\Phi\left(\frac{312-313{,}4}{68{,}88}\right)=\Phi(-0{,}02)=0{,}492,$$

$$\Phi\left(\frac{321-313{,}4}{68{,}88}\right)=\Phi(0{,}11)=0{,}5438,$$

$$\Phi\left(\frac{329-313{,}4}{68{,}88}\right)=\Phi(0{,}23)=0{,}591,$$

$$\Phi\left(\frac{358-313{,}4}{68{,}88}\right)=\Phi(0{,}65)=0{,}7422,$$

$$\Phi\left(\frac{375-313{,}4}{68{,}88}\right)=\Phi(0{,}89)=0{,}8133,$$

$$\Phi\left(\frac{418-313{,}4}{68{,}88}\right)=\Phi(1{,}52)=0{,}93574.$$

По формуле (1):

$$u=\frac{1}{12\cdot 9}+\left[0{,}04006-\frac{2\cdot 1-1}{2\cdot 9}\right]^2+\left[0{,}1711-\frac{2\cdot 2-1}{2\cdot 9}\right]^2+\left[0{,}2514-\frac{2\cdot 3-1}{2\cdot 9}\right]^2+$$

$$+\left[0{,}492-\frac{2\cdot 4-1}{2\cdot 9}\right]^2+\left[0{,}5438-\frac{2\cdot 5-1}{2\cdot 9}\right]^2+\left[0{,}591-\frac{2\cdot 6-1}{2\cdot 9}\right]^2+\left[0{,}7422-\frac{2\cdot 7-1}{2\cdot 9}\right]^2+$$

$$+\left[0{,}8133-\frac{2\cdot 8-1}{2\cdot 9}\right]^2+\left[0{,}93574-\frac{2\cdot 9-1}{2\cdot 9}\right]^2=0{,}00926+0{,}00024+0{,}00002+0{,}0007+$$

$$+0{,}01063+0{,}00192+0{,}0004+0{,}0004+0{,}0004+0{,}00008=0{,}02405.$$

В таблице 1 находим $u_{0,05}$ = 0,461.

Поскольку u = 0,02405 < $u_{0,05}$ = 0,461, то принимается нулевая гипотеза о том, что выборка не противоречит нормальному закону распределения с вероятностью 95 %.

4 Вопросы и задания для самоконтроля

1 Что такое уровень значимости?

2 Что такое ошибка первого и ошибка второго рода?

3 Что будет происходить с ошибкой второго рода при уменьшении уровеня значимости?

4 Что такое односторонние и двухсторонние ограничения критических статистик?

5 Перечислите критические статистики.

6 Приведите алгоритм проверки статистических гипотез.

7 С какой целью делается проверка гипотезы о резко выделяющихся членах выборки?

8 С какой целью производится проверка гипотез о законе распределения?

9 Для каких типов выборок применяются критерии Н.В. Смирнова и Пирсона?

10 С какой целью производится проверка гипотезы о равенстве двух средних значений и с помощью какого критерия?

11 Что такое нулевая и альтернативная гипотезы?

12 Что такое статистический критерий?

Список использованных источников

1 Куренков, В.И. Надежность изделий и систем ракетно-космической техники: лабораторный практикум / В.И. Куренков, В.В. Волоцуев. – Самара: Самарский государственный аэрокосмический университет, 2010. – 116 с.

2 Зорин, В.А. Надежность механических систем: учебник / В.А. Зорин. – М. : ИНФРА-М, 2015. – 380 с.

3 Половко, А.М. Основы теории надежности. Практикум: учеб. пособие для вузов / А.М. Половко, С. В. Гуров. – СПб. : БВХ-Петербург, 2006. – 560 с.

Приложение А
(обязательное)

Квантили функции нормального распределения Ф(x)

Таблица А.1 – Значения функции нормального распределения

x		0	1	2	3	4	5	6	7	8	9
0,0	0	5000	5040	5080	5120	5160	5199	5239	5279	5319	5359
0,1	0	5398	5438	5478	5517	5557	5596	5636	5675	5714	5753
0,2	0	5793	5832	5871	5910	5948	5987	6026	6064	6103	6141
0,3	0	6179	6217	6255	6293	6331	6368	6406	6443	6480	6517
0,4	0	6554	6591	6628	6664	6700	6736	6772	6808	6844	6879
0,5	0	6915	6950	6985	7019	7054	7088	7123	7157	7190	7224
0,6	0	7257	7291	7324	7357	7389	7422	7454	7486	7517	7549
0,7	0	7580	7611	7642	7673	7704	7734	7764	7794	7823	7852
0,8	0	7881	7910	7939	7967	7995	8023	8051	8078	8106	8133
0,9	0	8159	8186	8212	8238	8264	8289	8315	8340	8365	8389
1,0	0	8413	8438	8461	8485	8508	8531	8554	8577	8599	8621
1,1	0	8643	8665	8686	8708	8729	8749	8770	8790	8810	8830
1,2	0	8849	8869	8888	8907	8925	8944	8962	8980	8997	9015
1,3	0,9	0320	0490	0658	0824	0988	1149	1308	1466	1621	1774
1,4	0,9	1924	2073	2220	2364	2507	2647	2785	2922	3056	3189
1,5	0,9	3319	3448	3574	3699	3822	3943	4062	4179	4295	4408
1,6	0,9	4520	4630	4738	4845	4950	5053	5154	5254	5352	5449
1,7	0,9	5543	5637	5728	5818	5907	5994	6080	6164	6246	6327
1,8	0,9	6407	6485	6562	6637	6712	6784	6856	6926	6995	7062
1,9	0,9	7128	7193	7257	7320	7381	7441	7500	7558	7615	7670
2,0	0,9	7725	7778	7831	7882	7932	7982	8030	8077	8124	8169
2,1	0,9	8214	8257	8300	8341	8382	8422	8461	8500	8537	8574
2,2	0,9	8610	8645	8679	8713	8745	8778	8809	8840	8870	8899
2,3	0,9	8928	8956	8983	9010	9036	9061	9086	9111	9134	9158
2,4	$0,9_2$	1802	2024	2240	2451	2656	2857	3053	3244	3431	3613
2,5	$0,9_2$	3790	3963	4132	4297	4457	4614	4766	4915	5060	5201
2.6	$0,9_2$	5339	5473	5603	5731	5855	5975	6093	6207	6319	6427
2,7	$0,9_2$	6533	6636	6736	6833	6928	7020	7110	7197	7282	7365
2,8	$0,9_2$	7445	7523	7599	7673	7744	7814	7882	7948	8012	8074
2,9	$0,9_2$	8134	8193	8250	8305	8359	8411	8462	8511	8559	8605
3,0	$0,9_2$	8650	8694	8736	8777	8817	8856	8893	8930	8965	8999

Продолжение таблицы А.1

x		0	1	2	3	4	5	6	7	8	9
3,1	$0,9_3$	0324	0646	0957	1260	1553	1836	2102	2378	2636	2886
3,2	0.9_3	3129	3363	3590	3810	4024	4230	4429	4623	4810	4991
3,3	$0,9_3$	5166	5335	5499	5658	5811	5959	6103	6242	6376	6505
3,4	$0,9_3$	6631	6752	6869	6982	7091	7197	7299	7398	7493	7585
3,5	$0,9_3$	7674	7760	7842	7922	7999	8074	8146	8215	8282	8347
3,6	$0,9_3$	8409	8469	8527	8583	8637	8689	8739	8787	8834	8879
3,7	$0,9_3$	8922	8964	9004	9043	9080	9116	9150	9184	9216	9247
3,8	$0,9_4$	2765	3052	3327	3593	3848	4094	4331	4558	4777	4988
3,9	$0,9_4$	5190	5385	5573	5753	5926	6092	6252	6406	6554	6696
4,0	$0,9_4$	6833	6964	7090	7211	7327	7439	7546	7649	7748	7843
4,1	$0,9_4$	7934	8022	8106	8186	8264	8338	8409	8477	8542	8605
4.2	$0,9_4$	8665	8723	8778	8832	8882	8931	8978	9023	9066	9107
4,3	$0,9_5$	1460	1837	2198	2544	2876	3193	3497	3788	4066	4332
4,4	$0,9_5$	4588	4832	5065	5288	5502	5706	5902	6089	6268	6439
4,5	$0,9_5$	6602	6759	6908	7051	7187	7318	7442	7561	7675	7784
4,6	$0,9_5$	7888	7987	8081	8172	8258	8340	8419	8494	8566	8634
4,7	$0,9_5$	8699	8761	8821	8877	8931	8983	9032	9079	9124	9166
4,8	$0,9_6$	2067	2454	2822	3173	3508	3827	4131	4420	4696	4958
4,9	$0,9_6$	5208	5446	5673	5888	6094	6289	6475	6652	6821	6981
5,0	$0,9_6$	7134	7278	7416	7548	7672	7791	7904	8011	8113	8210
5,1	$0,9_6$	8302	8389	8472	8551	8626	8698	8765	8830	8891	8949
5,2	$0,9_7$	004	056	105	152	197	240	280	318	354	388
5,3	$0,9_7$	421	452	481	509	539	560	584	606	628	648
5,4	$0,9_7$	667	685	702	718	734	748	762	775	787	799
5,5	$0,9_7$	810	821	831	840	849	857	865	873	880	886
5,6	$0,9_7$	893	899	905	910	915	920	924	929	933	936
5,7	$0,9_8$	40	44	47	50	53	55	58	60	63	65
5,8	$0,9_8$	67	69	71	72	74	75	77	78	79	81
5,9	$0,9_8$	82	83	81	85	86	87	87	88	89	90

Примечания

1 Для отрицательных значений аргумента $F(-x) = 1 - F(x)$.

2 Индекс у цифры 9 означает ее повторение, например, при $x = 3,95$ имеем $F(x)=0,9_4 6092=0,99996092$.

Приложение Б
(обязательное)

Квантили функции распределения Пирсона

Таблица Б.1 – Значения квантилей функции распределения Пирсона

$v = n - 1$	α					
	0,995	0,975	0,95	0,05	0,025	0,005
1	0,000	0,001	0,004	3,841	50,24	7,879
2	0,010	0,051	0,103	5,991	7,378	10,597
3	0,0717	0,216	0,352	7,815	9,348	12,838
4	0,207	0,484	0,711	9,488	11,143	14,860
5	0,412	0,831	1,145	11,070	12,832	16,750
6	0,676	1,237	1,635	12,592	14,449	18,548
7	0,989	1,690	2,167	14,067	13,013	20,278
8	1,344	2,180	2,733	15,507	17,535	21,955
9	1,735	2,700	3,325	16,919	19,023	23,589
10	2,156	3,247	3,940	18,307	20,483	25,188
11	2,603	3,816	4,575	19,675	21,920	26,757
12	3,074	4,404	5,226	21,026	23,336	28,300
13	3,565	5,009	5,892	22,362	24,736	29,819
14	4,075	5,629	6,571	23,685	26,119	31,319
15	4,601	6,262	7,261	24,996	27,448	32,801
16	5,142	6,908	7,962	26,296	28,845	34,267
17	5,697	7,564	8,762	27,587	30,191	35,718
18	6,265	8,231	9,930	28,869	31,526	37,156
19	6,844	8,907	10,117	30,144	32,852	38,582
20	7,434	9,591	10,851	31,410	34,170	39,997
21	8,034	10,283	11,591	32,671	35,479	41,401
22	8,643	10,982	12,338	33,924	36,781	42,796
23	9,260	11,688	13,091	35,172	38,076	44,181
24	9,886	12,401	13,848	36,415	39,364	45,558
25	10,520	13,120	14,611	37,652	40,646	46,928
26	11,160	13,844	15,379	38,885	41,923	48,920
27	11,808	14,573	16,151	40,113	43,194	49,645
28	12,461	15,308	16,928	41,337	44,461	50,993
29	13,121	16,047	17,708	42,557	45,722	52,336
30	13,787	16,791	18,493	43,773	46,979	53,672

Приложение В

(обязательное)

Значения статистики Фишера

Таблица В.1 – Значения статистики $F_{0,05}$

ν_2	ν_1													
	1	2	3	4	5	6	7	8	9	10	20	50	100	
1	161	200	216	225	230	234	237	239	241	242	243	252	253	254
2	18,51	19,00	19,10	19,25	19,30	19,33	19,35	19,17	19,38	19,39	19,44	19,48	19,49	19,50
3	10,13	9,55	9,28	9,12	9,01	8,94	8,89	8,85	8,81	8,79	8,66	8,58	8,55	8,53
4	7,71	6,94	6,59	6,39	6,26	6,16	6,09	6,04	6,00	5,96	5,80	5,70	5,66	5,63
5	6,01	5,79	5,41	5,19	5,05	4,95	4,88	4,82	4,77	4,74	4,56	4,44	4,41	4,36
7	5,59	4,74	4,35	4,13	3,97	3,89	3,79	3,67	3,64	3,64	3.44	3,32	3,27	3,23
10	4,96	4,10	3,71	3,48	3,33	3,22	3,14	3,07	3,02	2,98	2,77	2,64	2,59	2,54
20	4,35	3,49	3,10	2,87	2,71	2,60	2,51	2,45	2,39	2,35	2,12	1,97	1,91	1,84
50	4,03	3,18	2,79	2,56	2,40	2.29	2,20	2,13	2,07	2,03	1,78	1,60	1,52	1,44
100	3,94	3,09	2,70	2,46	2,31	2,19	2,10	2,03	1,97	1,93	1,68	1,48	1,39	1,28
	3,84	3,00	2,60	2,37	2,21	2,10	2,01	1,94	1,88	1,83	1,57	1,35	1,24	1,0

Примечание – Степень свободы выборки с большей дисперсией – ν_1, степень свободы выборки с меньшей дисперсией – ν_2.

Приложение Г
(обязательное)

Квантили функции распределения Стьюдента

Таблица Г.1 – Значения квантилей функции распределения Стьюдента

Число степеней свободы ν	Вероятность γ											
	0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9	0,95	0,98	0,99
1	0,16	0,32	0,51	0,73	1,00	1,38	1,96	3,08	6,31	12,71	31,82	63,66
2	14	29	44	62	0,82	06	34	1,89	2,92	4,30	6,96	9,92
3	14	28	42	58	76	0,98	25	64	35	3,18	4,54	5,84
4	13	27	41	57	74	94	19	53	13	2,78	3,75	4,60
5	13	27	41	56	73	92	16	48	01	57	36	03
6	0,13	0,26	0,40	0,55	1,72	1,91	1,13	1,44	1,94	2,45	3,14	3,71
7	13	26	40	55	71	90	12	41	89	36	00	50
8	13	26	40	55	70	89	11	40	86	31	2,90	35
9	13	26	40	54	70	88	10	38	83	26	82	25
10	13	26	40	54	70	88	09	37	81	23	76	17
11	0,13	0,26	0,40	0,54	0,70	0,88	1,09	1,36	1,80	2,20	2,72	3,11
12	13	26	39	54	69	87	08	36	78	18	68	05
13	13	26	39	54	69	87	08	35	77	16	65	01
14	13	26	39	54	69	87	08	34	76	14	62	2,98
15	13	26	39	54	69	87	07	34	75	13	60	95
16	0,13	0,26	0,39	0,53	0,69	0,86	1,07	1,34	1,75	2,12	2,58	2,92
17	13	26	39	53	69	86	07	33	74	11	57	90
18	13	26	39	53	69	86	07	33	73	10	55	88
19	13	26	39	53	69	86	07	33	73	09	54	86
20	13	26	39	53	69	86	06	32	72	09	53	84
21	0,13	0,26	0,39	0,53	0,69	0,86	1,06	1,32	1,72	2,08	2,52	2,83
22	13	26	39	53	69	86	06	32	72	07	51	82
23	13	26	39	53	68	86	06	32	71	07	50	81
24	13	26	39	53	68	86	06	32	71	06	49	80
25	13	26	39	53	68	86	06	32	71	06	48	79
26	0,13	0,26	0,39	0,53	0,68	0,86	1,06	1,31	1,71	2,06	2,48	2,78
27	13	26	39	53	68	85	06	31	70	05	47	77
28	13	26	39	53	68	85	06	31	70	05	47	76
29	13	26	39	53	68	85	05	31	70	04	46	76
30	13	26	39	53	68	85	05	31	70	04	46	75
40	0,13	0,25	0,39	0,53	0,68	0,85	1,05	1,30	1,68	2,02	2,42	2,70
60	13	25	39	53	68	85	05	30	67	00	39	66
120	0,13	0,25	0,39	0,53	0,68	0,84	1,04	1,29	1,66	1,98	2,36	2,62
∞	13	25	38	52	67	84	04	28	64	96	33	58

Приложение Д
(обязательное)

Значения r-статистики

Таблица Д.1 – Значения r-статистики

Число степеней свободы	α			
	0,10	0,05	0,025	0,01
1	1,406	1,414	1,414	1,414
2	1,645	1,689	1,710	1,723
3	1,791	1,869	1,917	1,955
4	1,894	1,996	2,067	2,130
5	1,974	2,093	2,182	2,265
6	2,041	2,172	2,273	2,374
7	2,097	2,237	2,349	2,464
8	2,146	2,294	2,414	2,540
9	2,190	2,343	2,470	2,606
10	2,229	2,387	2,519	2,663
11	2,264	2,426	2,562	2,714
12	2,297	2,461	2,602	2,759
13	2,326	2,493	2,638	2,800
14	2,354	2,523	2,670	2,837
15	2,380	2,551	2,701	2,871
16	2,404	2,577	2,728	2,903
17	2,426	2,600	2,754	2,932
18	2,447	2,623	2,778	2,959
19	2,467	2,644	2,801	2,984
20	2,486	2,664	2,823	3,008
21	2,504	2,683	2,843	3,030
22	2,520	2,701	2,862	3,051
23	2,537	2.717	2.880	3,071