МИНОБРНАУКИ РОССИИ

Федеральное государственное бюджетное образовательное учреждение высшего образования
«Юго-Западный государственный университет»
(ЮЗГУ)

Кафедра дизайна и индустрии моды

УТВЕРЖДАЮ
Проректор по учебной работе
________________ О.Г. Локтионова
«16» 05 ________________ 2023 г.

МЕТОДЫ АНАЛИЗА ВЛИЯНИЯ ФАКТОРОВ

Методические указания по выполнению лабораторной и самостоятельной работы

Курск 2023

УДК 658.5

Составитель: С.В. Ходыревская

Рецензент
Доктор технических наук, профессор В.В. Куц

Методы анализа влияния факторов: методические указания по выполнению лабораторной и самостоятельной работы / Минобрнауки России, Юго-Зап. гос. ун-т; сост.: С.В. Ходыревская. – Курск, 2023. – 20 с.:– Библиогр.: с. 20.

Излагаются сведения о методах анализа влияния факторов. Рассмотрены примеры применения методов анализа влияния факторов для решения задач в области контроля и управления качеством. Приведены задания для самостоятельного выполнения, вопросы для самопроверки и подготовки, а также тест для самоконтроля.

Методические указания предназначены для бакалавров и специалистов всех направлений подготовки и специальностей и для всех форм обучения.

Текст печатается в авторской редакции

Подписано в печать . Формат 60х84 1/16.
Усл.печ. л. 1,16. Уч.-изд. л. 1,05.
Тираж 100 экз. Заказ 434 Бесплатно.
Юго-Западный государственный университет.
305040, г. Курск, ул. 50 лет Октября, 94.

1 Цель работы: получить основные сведения о методах анализа влияния факторов и получить практические навыки их использования, а также навыки расчетов и построения графиков при проведении анализа в среде *LibreOffice*.

2 Задания для самостоятельного выполнения:

Задание 1. Выяснить, есть ли существенные различия в работе двух смен, на основании данных выданных преподавателем.

Какой графический метод следует выбрать для представления результатов анализа?

Задание 2. Выяснить, как влияет замена шлифовального круга на результаты измерений на основании данных выданных преподавателем. Какой графический метод следует выбрать для представления результатов анализа?

Задание 3. Выяснить, как влияет замена шлифовального круга на долю брака в первую и вторую смены на основании данных выданных преподавателем. Какой графический метод следует выбрать для представления результатов анализа?

Задание 4. Необходимо оценить работу двух лаборантов, на основании данных выданных преподавателем. Какой графический метод следует выбрать для представления результатов анализа?

Задание 5. На основании данных выданных преподавателем проанализировать данные экспертов и проверить согласованность их мнений.

Задание 6. Выяснить, оказало ли влияние изменение технологии обработки поверхности изделия на контролируемую характеристику, на основании данных выданных преподавателем. Какой графический метод следует выбрать для представления результатов анализа?

Задание 7. Оценить использование трех станков разной модификации на точность изготовления детали на основании данных выданных преподавателем.

Задание 8. Выяснить, как влияет замена шлифовального круга на количество брака (шт.) полученного в первую и вторую смены на основании данных выданных преподавателем. Какой графический метод следует выбрать для представления результатов анализа?

Задание 9. Была разработана новая методика. Для оценки ее эффективности были проведены контрольные опыты (данные выдаются преподавателем). Можно ли рекомендовать применять новую методику вместо общепринятой? Какой графический метод следует выбрать для представления результатов анализа?

Задание 10. Два лаборанта проводили измерения. На основании данных выданных преподавателем определить можно ли объединить результаты лаборантов в одну выборку? Какой графический метод следует выбрать для оценки результатов работы лаборантов?

3. Краткие теоретические сведения

Существует довольно много подходов, позволяющих определить, оказывает ли некоторый фактор влияние на качество процесса. Например, мы говорим, что уровень квалификации персонала оказывает влияние на качество выпускаемой продукции. Так ли это? Как доказать наличие данного влияния? И если влияние есть, то насколько сильное? Или, как доказать, что качество сырья оказывает влияние на качество продукции? Конечно, есть мнения экспертов, но хотелось бы получить более объективную доказательную базу для выводов. Для решения подобных вопросов активно используют статистические методы.

Методы, позволяющие подтвердить влияние фактора можно классифицировать следующим образом (рисунок 1):

Рисунок 1 – Классификация методов анализа влияния факторов

С помощью аналитических методов анализа можно решить большинство практических задач по анализу факторных влияний на качество продукции. Все методы объединены одной целью – выяснить оказывает ли влияние некоторый фактор на качество продукции или процесса. Обратите внимание, что методы не отвечают на вопрос о том, какое влияние оказывает фактор. Чтобы оценить силу влияния фактора можно использовать графический инструментарий.

Чтобы выбрать адекватный метод анализа в любой практической ситуации следует использовать схему выбора метода анализа влияния факторов (рисунок 2).

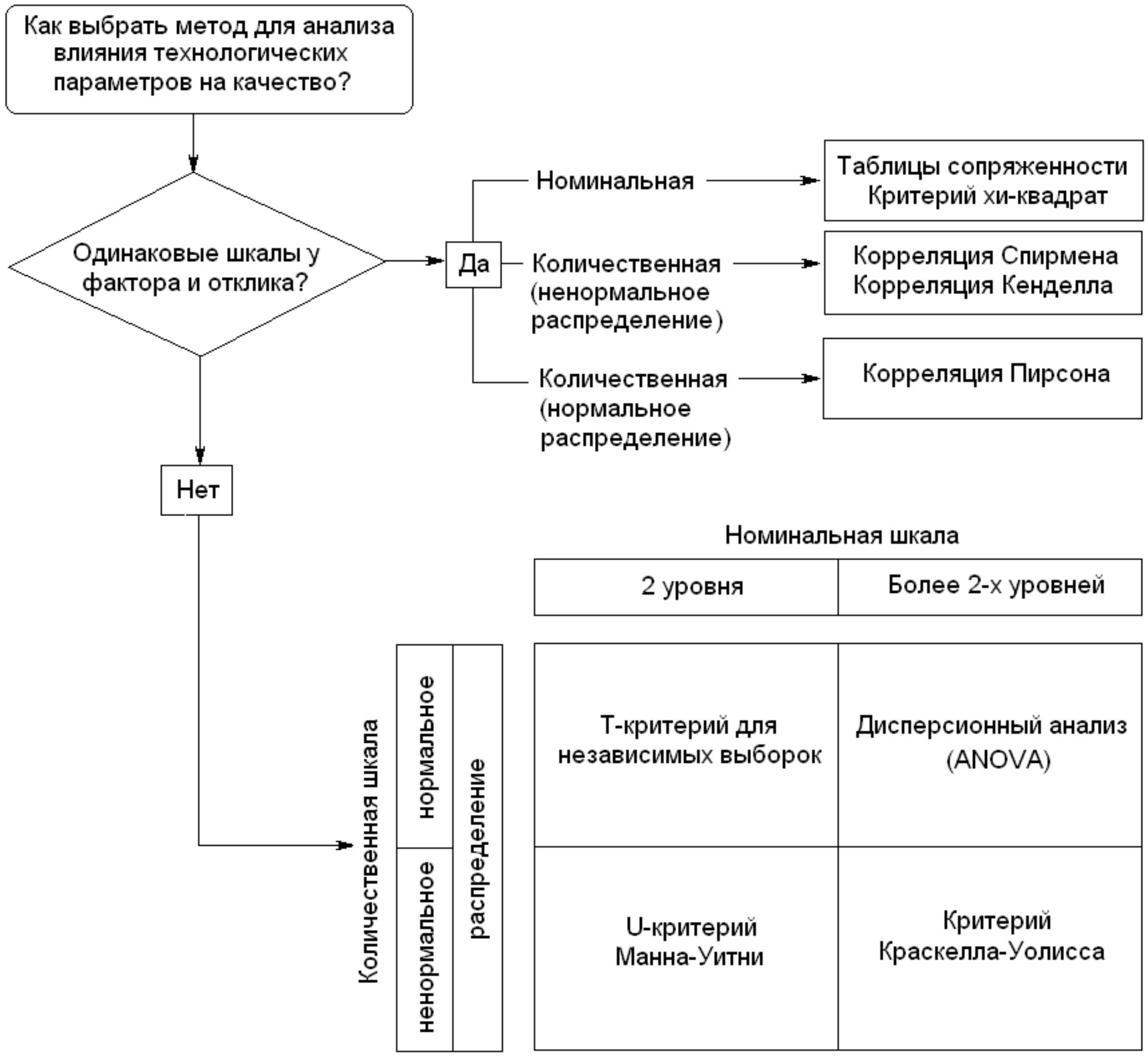


Рисунок 2 – Схема выбора метода анализа влияния фактора

Использование графических методов для представления результатов анализа, представлено в таблице 1.

Таблица 1

Использование графических методов для представления результатов анализа

Метод	Вид графика
Т – критерий Стьюдента	«Ящик с усами» (box plot)
Корреляция Спирмена	Диаграмма рассеяния
Критерий Краскела-Уоллиса	«Ящик с усами» (box plot)
Критерий Хи-квадрат	Круговая диаграмма

3.1. Критерий «хи-квадрат» для анализа таблиц сопряженности

Исходные данные для анализа:

1) Любая ячейка таблицы сопряженности должна содержать не менее 5 наблюдений;

2) Количество уровней номинальной шкалы не должно превышать 4. В противном случае результаты анализа будут утрачивать информативность.

Критерий хи-квадрат определяется по формуле:

$$\chi^2 = \sum \frac{\left(f_0 - f_e\right)^2}{f_e}, \qquad (1)$$

где f_0 – наблюдаемая частота в каждой ячейке; f_e – ожидаемые частоты в каждой ячейке.

Ожидаемые частоты определяются исходя из данных таблицы сопряженности наблюдаемых частот по формуле:

$$f_e = \frac{n_r \cdot n_c}{n}, \qquad (2)$$

где n_r – накопительная частота по строке; n_c – накопительная частота по столбцу; n – общий размер выборки.

3.2. Корреляция Спирмена и Кендалла

Для ранговых данных используются другие коэффициенты корреляции, такие как коэффициенты Спирмена и Кендалла.

Коэффициент ранговой корреляции Спирмена определяется по формуле:

$$K_C = 1 - 6\sum_{i=1}^{n} \frac{d_i^2}{n\left(n^2 - 1\right)} \qquad (3)$$

где n – количество наблюдений; d_i – разность ранговых показателей в i-том наблюдении.

В отличие от коэффициента Пирсона данный коэффициент не является мерой линейной связи. Коэффициент $К_С$ лежит в промежутке от -1 до 1. При совпадении рангов $К_С$ =1, при противоположных рангах $К_С$ = -1.

Значимость коэффициента Спирмена может быть проверена по t-критерию

$$t = K_C \sqrt{\frac{n-2}{1-K_C^2}}\,, \tag{4}$$

имеющему распределение Стьюдента с числом степеней свободы $n - 2$.

Коэффициент ранговой корреляции Кендалла определяется по формуле:

$$\tau = 1 - \frac{4Q}{n(n-1)}, \tag{5}$$

где Q – минимальное число обменов соседних элементов одной из ранжировок для ее приведения (совпадения) с другой.

Как и для коэффициента Спирмена при совпадении рангов $\tau = 1$, при противоположных рангах $\tau = -1$.

Статистика для проверки значимости этого коэффициента

$$Z = \sqrt{\frac{9n(n-1)}{2(2n+5)}}\tau \tag{6}$$

имеет нормальное распределение $N(0, 1)$.

Коэффициент ранговой корреляции Кендалла имеет некоторые преимущества перед коэффициентом Спирмена – в частности, он может использоваться и для многофакторного анализа. При достаточно большом числе объектов ($n \geq 10$) между значениями ранговых коэффициентов существует простая связь $К_С = 1{,}5\tau$.

3.3. Т-критерий для независимых выборок

Исходные данные для анализа:

1) Номинальная шкала должна иметь ровно 2 уровня;
2) Данные внутри сравниваемых подгрупп должны быть нормально распределены;
3) Каждый уровень номинальной шкалы должен иметь хотя бы 30 значений.

Анализ проводится по Т-критерию Стьюдента:

$$T = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_{\bar{x}_1}^2}{n_1} + \frac{s_{\bar{x}_2}^2}{n_2}}}\,, \tag{7}$$

где n_1, n_2 – размеры сравниваемых выборок;

$\bar{x}_1$, $\bar{x}_2$ – среднее значение по выборкам;

$s_{\bar{x}_1}, s_{\bar{x}_2}$ – выборочные стандартные ошибки и определяются по формуле:

$$s_{\bar{x}} = \frac{\sigma}{\sqrt{n}} \quad (8)$$

Расчетное значение Т-критерия Стьюдента сравниваем с табличным.

Определяем $T_{табл}$ по значению числа степеней свободы

$$\text{число степеней свободы} = n_1 + n_2 - 2. \quad (9)$$

Если расчетное значение больше табличного, то фактор оказывает существенное влияние на зависимую величину.

3.4. Дисперсионный анализ

Во многих практических ситуациях представляет интерес влияние того или иного качественного фактора на рассматриваемый показатель. Влияет ли квалификация наладчиков на качество обработки поверхности? Влияет ли метод обработки на точность изготовления детали? Зависит ли прочность болта из стекловолокнита от температуры при прессовании? Ответ на эти и аналогичные вопросы дается методами однофакторного дисперсионного анализа.

Пусть, например, оценка качества поверхности детали проводится с помощью k приборов и необходимо исследовать, влияет ли фактор «прибор» на результат измерений. Если приборов два, то проверка гипотезы о равенстве средних показаний приборов проводится рассмотренными выше методами. Если же приборов больше двух, используются методы дисперсионного анализа. Проверяется нулевая гипотеза H_0: $\mu_1 = \mu_2 = \ldots = \mu_k$ об отсутствии влияния на результативный признак X (результат измерений) фактора А («прибор»), имеющего k уровней A_j, $j = 1, 2, \ldots, k$.

Основная идея дисперсионного анализа состоит в том, чтобы сопоставить дисперсию за счет воздействия фактора А с дисперсией, обусловленной случайными причинами. Если различие между ними несущественно, то влияние фактора А на признак X незначительно. Если же различие между факторной и остаточной дисперсиями значимо, то это говорит о влиянии фактора А на рассматриваемый признак X.

Предполагается, что случайная величина X имеет нормальное распределение с математическим ожиданием μ_j, зависящим от уровня фактора A_j, и постоянной дисперсией σ^2. В качестве исходных данных используются выборочные значения величины X, полученные для каждого уровня фактора A; число элементов выборки на каждом уровне равно n, тогда общее число наблюдений пk, обозначим через x_{ij} результат i-го наблюдения (i = 1, 2, ..., п) за j-м фактором.

Выборочное среднее, соответствующее j-му уровню фактора A (групповое среднее), вычисляется по формуле

$$\bar{x}_j = \frac{1}{n}\sum_{i=1}^{n} x_{ij}, \quad (10)$$

а общее среднее

$$\bar{\bar{x}} = \frac{1}{nk}\sum_{j=1}^{k}\sum_{i=1}^{n} x_{ij} = \frac{1}{k}\sum_{j=1}^{k}\bar{x}_j \quad (11)$$

Общая сумма квадратов — это сумма квадратов отклонений наблюдаемых значений x_{ij} от общего среднего:

$$Q = \sum_{j=1}^{k}\sum_{i=1}^{n}\left(x_{ij} - \bar{x}\right)^2 = \sum_{j=1}^{k}\sum_{i=1}^{n} x_{ij}^2 - nk\bar{\bar{x}}^2 \quad (12)$$

Факторная сумма квадратов (обусловленная влиянием фактора A) — это сумма квадратов отклонений групповых средних от общей средней:

$$Q_A = n\sum_{j=1}^{k}\left(\bar{x}_j - \bar{\bar{x}}\right)^2 = n\sum_{j=1}^{k}\bar{x}_j^2 - nk\bar{\bar{x}}^2 \quad (13)$$

Остаточная сумма квадратов характеризует рассеяние внутри группы:

$$Q_e = \sum_{j=1}^{k}\sum_{i=1}^{n}\left(x_{ij} - \bar{x}_j\right)^2 \quad (14)$$

На практике эта сумма определяется из основного тождества дисперсионного анализа, в соответствии с которым

$$Q = Q_A + Q_e \quad (15)$$

Соответствующие числа степеней свободы

$$v = nk - 1;\ v_A = k - 1;\ v_e = k(n - 1) \quad (16)$$

а дисперсии

$$s^2 = Q/\nu;\ s_A^2 = Q_A/\nu_A;\ s_e^2 = Q_e/\nu_e \tag{17}$$

Если нулевая гипотеза о равенстве средних справедлива, то эти дисперсии являются несмещенными оценками дисперсий генеральной совокупности. Значительное превышение дисперсии s_A^2 над дисперсией s_e^2 можно объяснить различием средних в группах. Поэтому для проверки нулевой гипотезы используется статистика

$$F = \frac{s_A^2}{s_e^2} = \frac{Q_A/(k-1)}{Q_e/k(n-1)}, \tag{18}$$

которая имеет распределение Фишера с числами степеней свободы *(k – 1) и k(n – 1).* Нулевая гипотеза не противоречит результатам наблюдений на заданном уровне значимости α, если

$$F < F_{1-\alpha}\,(k-1,\ k(n-1)). \tag{19}$$

В этом случае считается, что фактор А не оказывает существенного влияния на показатель X.

Результаты расчета сводятся в таблицу 2.

Таблица 2

Результаты проведения дисперсионного анализа

Источник дисперсии	Сумма квадратов	Число степеней	Дисперсия	Статистика Фишера
Фактор А Остаток	Q_A Q_e	ν_A ν_e	s_A^2 s_e^2	F
Общая	Q	ν	s^2	

3.5. U-критерий Манна-Уитни

U-критерий Манна-Уитни (англ. Mann-Whitney U test) — непараметрический статистический критерий, используемый для оценки различий между двумя выборками по уровню какого-либо признака, измеренного количественно. Позволяет выявлять различия в значении параметра между малыми выборками.

Этот метод определяет, достаточно ли мала зона перекрещивающихся значений между двумя рядами (ранжированным рядом значений параметра в первой выборке и таким же во второй выборке). Чем меньше значение критерия, тем вероятнее, что различия между значениями параметра в выборках достоверны.

Ограничения применимости критерия

1. В каждой из выборок должно быть не менее 3 значений признака. Допускается, чтобы в одной выборке было два значения, но во второй тогда не менее пяти.

2. В каждой выборке должно быть не более 60 значений параметра, но уже при выборках в 20 и более единиц ранжирование становится довольно трудоемким.

Использование критерия

Для применения U-критерия Манна-Уитни нужно произвести следующие операции.

1. Составить единый ранжированный ряд из обоих сопоставляемых выборок, расставив их элементы по степени нарастания признака и приписав меньшему значению меньший ранг. Общее количество рангов получится равным: $N = n_1 + n_2$, где n_1 — количество единиц в первой выборке, а n_2 — количество единиц во второй выборке.

2. Разделить единый ранжированный ряд на два, состоящие соответственно из единиц первой и второй выборок. Подсчитать отдельно сумму рангов, пришедшихся на долю элементов первой выборки, и отдельно — на долю элементов второй выборки. Определить большую из двух ранговых сумм (T_x), соответствующую выборке с n_x единиц.

3. Определить значение U-критерия Манна-Уитни по формуле:

$$U = n_1 \cdot n_2 + \frac{n_x \cdot (n_x + 1)}{2} - T_x.$$

4. По таблице определить критические значения критерия для данных n_1 и n_2. Если полученное значение U меньше табличного или равно ему для избранного уровня статистической значимости, то признается наличие существенного различия между уровнем признака в рассматриваемых выборках (принимается альтернативная гипотеза). Если же полученное значение U больше табличного, принимается нулевая гипотеза. Достоверность различий тем выше, чем меньше значение U.

3.6. Критерий Краскела-Уоллиса

Если мы не можем сказать что-либо определенное об альтернативах к H_0, можно воспользоваться для ее проверки свободным от распределения критерием Краскела–Уоллиса. Для этого заменим наблюдения x_{ij} их рангами r_{ij} , упорядочивая всю совокупность $\|x_{ij}\|$ в порядке возрастания (для определенности). Затем для каждой обработки j (т.е. для каждого столбца исходной таблицы) надо вычислить

$$R_j = \sum_{i=1}^{n_j} r_{ij}$$

где R_j — это средний ранг, рассчитанный по столбцу. Если между столбцами нет систематических различий, средние ранги R_j , j = 1, . . . , k не должны значительно отличаться от среднего ранга, рассчитанного по всей совокупности $\|r_{ij}\|$. Ясно, что последний равен (N + 1)/2. Поэтому величины

$$\left(R_1 - \frac{N+1}{2}\right)^2, \; \dots \; , \left(R_k - \frac{N+1}{2}\right)$$

при H_0 в совокупности должны быть небольшими. Составляя общую характеристику, разумно учесть различия в числе наблюдений для разных обработок и взять в качестве меры отступления от чистой случайности величину

$$H = \frac{12}{N(N+1)} \sum_{j=1}^{k} \frac{R_j^2}{n_j} - 3(N+1)$$

Эта величина называется статистикой Краскела–Уоллеса. Множитель 12/[N(N + 1)] нужен для стабилизации ее распределения при большом числе наблюдений.

4. Примеры выполнения заданий

Пример 1: По данным количества продукции 1 и 2 сорта в разных цехах составляем таблицу сопряженности наблюдаемых частот (таблица 3)

Таблица 3

Таблица сопряженности наблюдаемых частот

Сорт	Номер цеха		Сумма
	Цех №1	Цех №2	
Сорт 1	34	48	82
Сорт 2	5	23	28
Сумма	39	71	110

По формуле (2) рассчитываем ожидаемые частоты и составляем таблицу сопряженности ожидаемых теоретических частот (таблица 4).

Таблица 4

Таблица сопряженности ожидаемых теоретических частот

Сорт	Номер цеха		Сумма
	Цех №1	Цех №2	
Сорт 1	29,1	52,9	82
Сорт 2	9,9	18,1	28
Сумма	39	71	110

$$f_{e_{11}} = \frac{82 \cdot 39}{110} = 29{,}1 \qquad f_{e_{12}} = \frac{28 \cdot 39}{110} = 9{,}9$$

$$f_{e_{21}} = \frac{82 \cdot 71}{110} = 52{,}9 \qquad f_{e_{22}} = \frac{28 \cdot 71}{110} = 18{,}1$$

И на основании данных таблиц 3 – 4 рассчитываем критерий хи-квадрат по формуле (1):

$$\chi^2_{расч} = \frac{(34 - 29{,}1)^2}{29{,}1} + \frac{(48 - 52{,}9)^2}{52{,}9} + \frac{(23 - 18{,}1)^2}{18{,}1} + \frac{(5 - 9{,}9)^2}{9{,}9} = 5{,}22$$

Расчетное значение критерия хи-квадрат сравниваем с табличным.

Определяем $\chi^2_{табл}$ по значению числа степеней свободы

число степеней свободы = (число строк – 1)(число столбцов – 1),

в нашем случае число степеней свободы равно 1, и при уровне значимости 5%, $\chi^2_{табл}$= 3,841.

Если расчетное значение больше табличного, то фактор оказывает существенное влияние на зависимую величину с вероятностью 95%.

Как и в нашем примере, $\chi^2_{расч} > \chi^2_{табл}$, следовательно, цех №2 работал хуже, так как выпустил больше продукции 2 сорта.

Для графического представления данных используется круговая диаграмма (рисунок 3).

Пример 2. Два эксперта проводят органолептический анализ жемчуга: ранжируют по убыванию качества 10 жемчужин. Результаты представлены в первых трех строках таблицы 5.

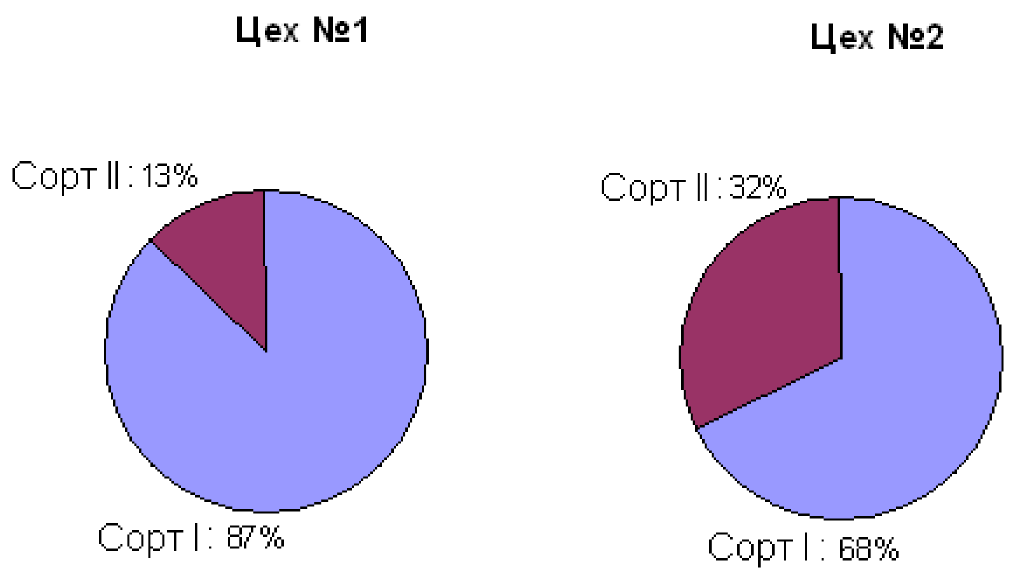


Рисунок 3 – Круговая диаграмма

Таблица 5

№ п/п	1	2	3	4	5	6	7	8	9	10
Эксперт 1	2	1	3	4	6	5	8	7	10	9
Эксперт 2	3	2	1	4	6	7	5	9	10	8
d_i	-1	-1	2	0	0	-2	3	-2	0	1
d_i^2	1	1	4	0	0	4	9	4	0	1

Проверить согласованность мнений экспертов по коэффициенту ранговой корреляции Спирмена.

Решение:

Найдем входящие в формулу (3) величины d_i - разность между рангами, присвоенные экспертами i-му объекту, а также квадраты этих величин; результаты вычислений приведены в нижних двух строках таблицы. Тогда

$$K_C = 1 - 6 \cdot 24 / 990 = 0{,}85$$

Для проверки значимости найдем значение статистики Стьюдента по формуле (4):

$$t_{расч} = 0{,}85\sqrt{\frac{10-2}{1-0{,}85^2}} = 4{,}56$$

Критическое значение при уровне значимости 0,05 по справочной таблице $t_{кр} = 1{,}86$. Так как $t_{расч} > t_{кр}$ - корреляция значима. Учитывая достаточно высокое значение коэффициента корреляции и его значимость, можно считать, что степень близости ранжировок экспертов высокая.

Пример 3. Необходимо оценить влияние технологии чистовой обработки (три вида технологий) на точность изготовления детали. При каждом виде технологии проводится по четыре замера отклонения размера детали от номинала в мкм (см. таблицу 6).

Таблица 6

Номер замера	Вид технологии		
	1	2	3
1	1	2	3
2	2	1	2
3	2	3	2
4	1	2	3

Решение:

Имеем $n = 4$, $k = 3$.

Групповые средние

$\bar{x}_1 = (1+2+2+1)/4 = 1{,}5$; $\bar{x}_2 = (2+1+3+2)/4 = 2$; $\bar{x}_3 = (3+2+2+3)/4 = 2{,}5$, а общая средняя $\bar{x} = (1{,}5+2+2{,}5)/3 = 2$.

Общая сумма квадратов

$$Q = 1^2 + 2^2 + 2^2 + 1^2 + 2^2 + 1^2 + 3^2 + 2^2 + 3^2 + 2^2 + 2^2 + 3^2 - 4 \cdot 3 \cdot 2^2 = 6,$$

факторная сумма квадратов

$$Q_A = 4\,(1{,}5^2 + 2^2 + 2{,}5^2) - 4 \cdot 3 \cdot 2^2 = 2,$$

остаточная сумма квадратов

$$Q_e = Q - Q_A = 4.$$

Результаты расчета сводим в таблицу 7.

Таблица 7

Источник дисперсии	Сумма квадратов	Число степеней	Дисперсия	Статистика Фишера
Фактор А	2	2	1	2,25
Остаток	4	9	0,44	
Общая	6	11	0,55	

Из справочной таблицы находим квантиль распределения Фишера

$$F_{1-\alpha}\,(k - 1,\, k(n - 1)) = F_{0{,}95}(2{,}9) = 4{,}26.$$

Так как выборочное значение статистики оказалось меньше критического (рисунок 4), нулевая гипотеза принимается: в данном случае влияние технологии изготовления на точность детали несущественно.

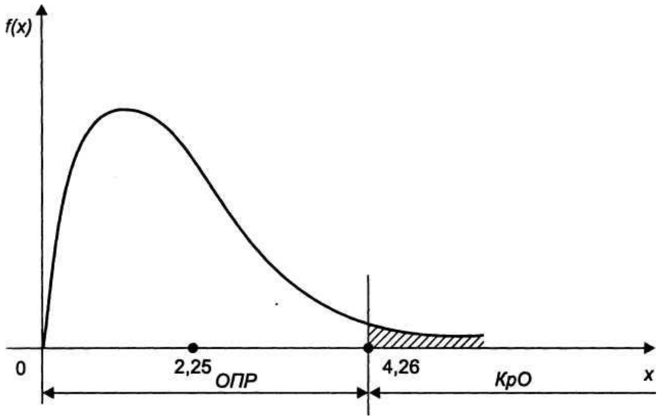


Рисунок 4 – Положение критической области при проведении дисперсионного анализа

Подобным образом в двухфакторном дисперсионном анализе оценивается влияние двух факторов A и B, а также их взаимодействия AB на результативный признак X: проверяются три соответствующие нулевые гипотезы. В трехфакторном анализе по аналогии исследуется влияние на признак X трех факторов A, B и C, их парных взаимодействий AB, BC и AC, а также общего взаимодействия ABC.

Пример 4. Необходимо установить оказало ли влияние замена шлифовального круга на контролируемую характеристику, если известно, что «до замены» была взята выборка объемом 60 единиц, среднее арифметическое которой равно 24,99960 мм, стандартное отклонение – 0,000924 мм; «после замены» была взята выборка объемом 90 единиц при этом среднее арифметическое стало равно 24,99952 мм, а дисперсия – 0,000986. Какой графический метод следует выбрать для представления результатов анализа?

Решение:

По полученным данным выдвигаем гипотезу о том, что замена шлифовального круга повлияла на процесс шлифования колец подшипников. Данную гипотезу проверим, рассчитав T-критерий Стьюдента.

Выборочные стандартные ошибки определяем по формуле (8):

$$s_{\bar{x}_1}^2 = \frac{\sigma^2}{n} = \frac{0{,}000924^2}{60} = 1{,}42 \cdot 10^{-8}$$

$$s_{\bar{x}_2}^2 = \frac{\sigma^2}{n} = \frac{0{,}000986^2}{90} = 1{,}08 \cdot 10^{-8}$$

Подставим полученные характеристики в формулу (7) для расчета критерия Стьюдента:

$$T_{расч} = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_{\bar{x}_1}^2}{n_1} + \frac{s_{\bar{x}_2}^2}{n_2}}} = \frac{24{,}9996 - 24{,}99952}{\sqrt{\frac{1{,}42 \cdot 10^{-8}}{60} + \frac{1{,}08 \cdot 10^{-8}}{90}}} = \frac{0{,}8 \cdot 10^{-4}}{0{,}189 \cdot 10^{-4}} = 4{,}23$$

По справочной таблицы значений Т-критерию Стьюдента определяем табличное значение критерия при уровне значимости 0,05 и числе степеней свободы 148 по формуле (9), $Т_{табл} = 1{,}9759$.

Поскольку $Т_{расч} > Т_{табл}$, то фактор не оказывает существенное влияние на зависимую величину, следовательно, замена шлифовального круга оказывает существенное влияние на замену шлифовального круга при шлифовании колец подшипников.

Для графического представления данных используем «ящик с усами» или диаграмму размаха (рисунок 5).

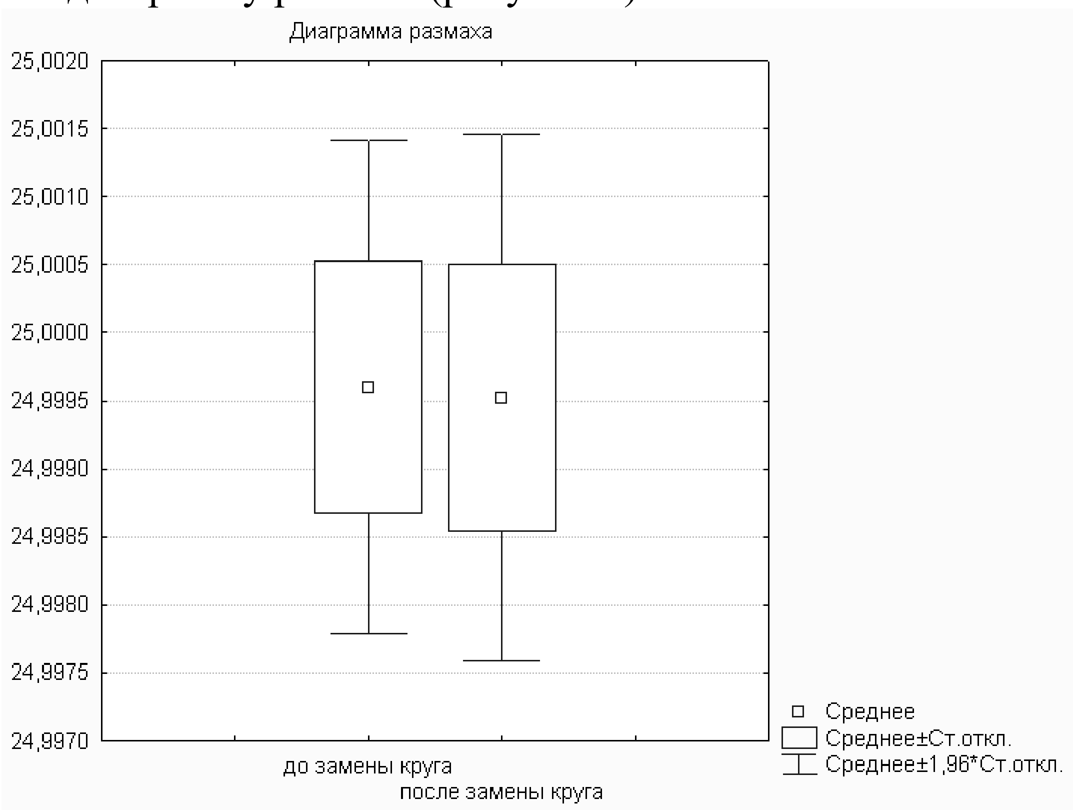


Рисунок 5 – «Ящик с усами» для оценки размаха диаметра колец подшипников до и после замены шлифовального круга

5 Порядок выполнения работы

Получив у преподавателя исходные данные для выполнения лабораторной работы, студент изучает теоретические сведения согласно пункту 3. Далее выполняет на компьютере задания на основании исходных данных выданных преподавателем в программе *LibreOffice*.

6 Содержание отчета

Отчет по лабораторной работе должен содержать следующие пункты:

– название лабораторной работы;

– цель работы;

– краткие теоретические сведения;

– краткое описание хода выполнения работы;

– индивидуальные задания для выполнения лабораторной работы;

– результаты выполнения работы: таблицы с исходными данными, графики, итоговая таблица расчета.

– выводы.

Вопросы для самопроверки и подготовки

1. Перечислите основные этапы алгоритма проверки гипотез о параметрах распределения.

2. Какие основные распределения, используемые в статистических расчетах Вы знаете?

3. Как определяют квантили распределений?

4.При решении, каких задач используются методы дисперсионного анализа?

5. Что понимают под адекватностью регрессионной модели?

6. Что характеризует коэффициент детерминации?

7. Какова зависимость между значениями ранговых коэффициентов Кендела и Спирмена?

8. Для анализа влияния, каких факторов используется критерий Стьюдента?

Тест для самоконтроля

1. Эксперимент, в процессе которого исследуется стохастическая зависимость одной величины Y от нескольких других величин X_i, называется ________________ .

2. Установите последовательность этапов проверки статистических гипотез:

1 – формулировка основной гипотезы
2 – расчёт статистики критерия
3 – задание уровня значимости
4 – построение критической области
5 – выводы об истинности гипотезы

3. Установите соответствие

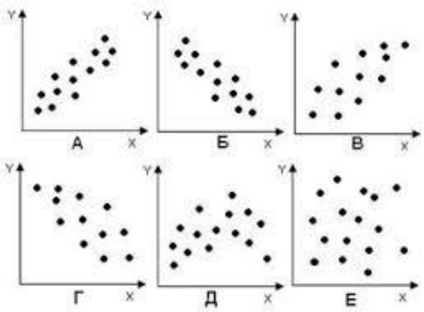
1	Коэффициент регрессии	А	$R = \dfrac{\operatorname{cov}(X,Y)}{\sigma_x \cdot \sigma_y}$
2	Линейный коэффициент корреляции	Б	$R_{y/x} = \dfrac{\sigma_y}{\sigma_x} r$
3	Выборочный линейный коэффициент корреляции	В	$r_{xy}^{p} = \dfrac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum (x_i - \bar{x})^2 \cdot \sum (y_i - \bar{y})^2}}$
4	Коэффициента корреляции Пирсона		$r = \dfrac{\sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n} (x_i - \bar{x})^2 \sum_{i=1}^{n} (y_i - \bar{y})^2}}$

4. Установите соответствие

Метод	Вид графика
Т – критерий Стьюдента	«Ящик с усами» (box plot)
Корреляция Спирмена	«Ящик с усами» (box plot)
Критерий Краскела-Уоллиса	Круговая диаграмма
Критерий Хи-квадрат	Диаграмма рассеяния

5. Установите соответствие:

1 – сильная положительная корреляция
2 – слабая положительная корреляция
3 – сильная отрицательная корреляция
4 – слабая отрицательная корреляция
5 – отсутствие корреляции
6 – криволинейная корреляция



6. Чертеж, на котором статистическая информация изображается посредством геометрических фигур или символических знаков называется ________________ .

7. Установите последовательность основных этапов статистического управления качеством:

1 – статистическое обследование
2 – наладка процесса
3 – статистическое управление

Библиографический список

1. Всеобщее управление качеством: учебник для вузов / О.П. Глудкин [и др.] – М.: Горячая линия – Телеком, 2001. – 599 с.

2. Клячкин, В. Н. Статистические методы в управлении качеством: компьютерные технологии : учебное пособие / В. Н. Клячкин. - Москва : Финансы и статистика, 2009. - 304 с. : ил.

3. Ефимов, В. В. Статистические методы в управлении качеством продукции : учебное пособие / В. В. Ефимов, Т. В. Барт. - М. : КноРус, 2006. - 240 с.

4. Статистические методы контроля и управления качеством: методические указания по выполнению практических работ и самостоятельной работы по дисциплине «Статистические методы контроля и управления качеством» для студентов направления подготовки 27.03.01 Стандартизация и метрология / Юго-Зап. гос. ун-т; сост.: С.В. Ходыревская. Курск, 2018. 144 с.: прилож. 13. Библиогр.: с. 119.

5. Пономарев С. В. Управление качеством продукции. Введение в системы менеджмента качества: Учебное пособие / С. В. Пономарев, С. В. Мищенко, В. Я. Белобрагин. - М. : Стандарты и качество, 2004. - 248 с. (гриф УМО)