МИНОБРНАУКИ РОССИИ

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

Кафедра дизайна и технологии изделий легкой промышленности

УТВЕРЖДАЮ
Проректор по учебной работе
_____________ О.Г. Локтионова
«15» 12 _____________ 2017 г.

АПРИОРНОЕ РАНЖИРОВАНИЕ ФАКТОРОВ

Методические указания
по выполнению лабораторных работ
для студентов направления подготовки 29.03.05

Курск 2017

УДК 687.01:004.9

Составитель: Т.А. Добровольская

Рецензент
Кандидат технических наук, доцент *Т.М. Ноздрачева*

Априорное ранжирование факторов: методические указания по выполнению лабораторных работ / Юго-Зап. гос. ун-т; сост.: Т.А. Добровольская. - Курск, 2017. - 7 с.: ил. 1, табл. 1.- Библиогр.: с. 7.

Излагаются основные сведения о методике проведения психологического эксперимента (априорного ранжирования факторов) и этапах обработки его результатов. Указывается порядок выполнения лабораторной работы

Предназначены для студентов направления подготовки 29.03.05 «Конструирование изделий легкой промышленности» дневной и заочной форм обучения

Текст печатается в авторской редакции

Подписано в печать . Формат 60×84 1/16.
Усл.печ.л. . Уч.-изд.л. . Тираж 25 экз. Заказ. Бесплатно.
Юго-Западный государственный университет.
305040, г. Курск, ул. 50 лет Октября, 94

АПРИОРНОЕ РАНЖИРОВАНИЕ ФАКТОРОВ

Цель работы: Изучение методики проведения и обработки результатов психологического эксперимента (априорного ранжирования факторов).

Задание:
1. Произвести выбор факторов, влияющих на параметр оптимизации (согласно заданию).
2. Произвести анкетирование и определить согласованность мнений экспертов.
3. Построить диаграмму рангов.
4. Сделать выводы.

ОСНОВНЫЕ ТЕОРЕТИЧЕСКИЕ СВЕДЕНИЯ

На стадии предварительного изучения объекта исследования при формализации априорных сведений иногда полезно проведение психологического эксперимента, заключающегося в объективной обработке данных, полученных в результате опроса специалистов или из исследований, опубликованных в литературе. Такой эксперимент позволяет более правильно принять или отвергнуть некоторые предварительные гипотезы, дать сравнительную оценку влияния различных факторов на параметры оптимизации и тем самым правильно отобрать факторы для последующего активного эксперимента, обоснованно исключить некоторые из них из дальнейшего рассмотрения.

При решения подобных задач можно использовать метод априорного ранжирования факторов.

Сущность этого метода заключается в следующем: все факторы, которые согласно априорной информации могут иметь существенное влияние, ранжируются в порядке убывания вносимого ими вклада. Вклад каждого фактора оценивается по величине ранга - места, которое отведено исследователем (специалистом при опросе, автором статьи и так далее) данному фактору при ранжировании всех факторов с учетом их предполагаемого (количественно неизвестного) влияния на параметры оптимизации. При сборе мнений путем опроса специалистов каждому из них предлагается заполнить анкету, в которой перечислены факторы, их размерность, предполагаемые интервалы варьирования. Заполняя анкету, специалист определяет место факторов в ранжированном ряду. Одновременно он может включить дополнительные факторы или мнение об изменении интервалов варьирования.

Обработка результатов ранжирования производится в следующем порядке:

1. Определяется сумма рангов по факторам:

$$r_j = \sum_{1}^{m} a_{ij}, \tag{1}$$

где m – число исследователей;

a_{ij} – ранг каждого i-го фактора у j-го исследователя.

2. Определяется разность Δ_j между суммой рангов каждого фактора и средней суммой рангов:

$$\Delta_j = r_j - R, \tag{2}$$

где $R = \frac{1}{k} \cdot \sum_{j=1}^{k} r_j$ - средняя сумма рангов;

k - число факторов.

3. Определяют сумму квадратов отклонений s:

$$S = \sum_{j=1}^{k} \Delta_j^2. \tag{3}$$

4. Оценивается степень согласованности мнений всех исследователей с помощью коэффициента конкордации ω:

$$\omega = \frac{12 \cdot S}{m^2 \cdot \left(k^3 - k\right) - m \cdot \sum_{1}^{m} T_j}, \tag{4}$$

где $T_j = \sum \left(t_j^3 - t_j\right)$;

t_j – число одинаковых рангов в j – ом ранжировании.

Использовать коэффициент конкордации можно после оценки его значимости, которая определяется с помощью χ^2 – критерия (Пирсона):

$$\chi^2 = m \cdot (k-1) \cdot \omega. \tag{5}$$

Гипотеза о наличии согласия исследователей может быть принята, если при заданном числе степеней свободы (f=k-1), табличное значение $\chi^2_{табл}$ меньше расчетного χ^2 для 5 %-го уровня значимости.

5. Оценив согласованность мнений всех исследователей, строят среднюю диаграмму рангов, откладывая по одной оси факторы, а по другой – соответствующие суммы рангов (рис.1).

Чем меньше сумма рангов данного фактора, тем выше его место на диаграмме. С помощью последней оценивается значимость факторов.

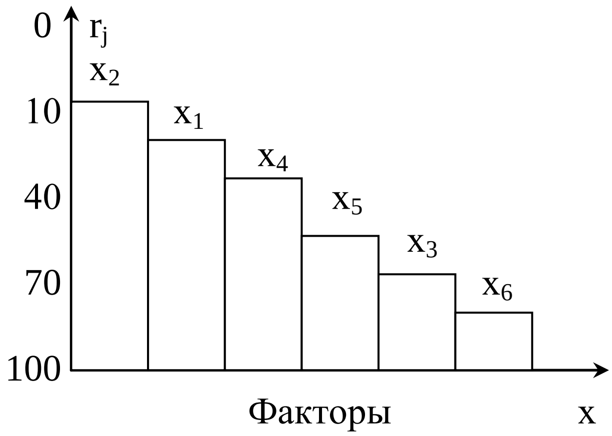


Рисунок 1- Распределение значимости факторов
на диаграмме рангов

В ситуациях с очень большим числом факторов, кроме общей согласованности мнений исследователей, рассматривают с помощью χ^2- распределения и согласованность по каждому фактору в отдельности.

6. Для оценки весомости свойства можно использовать формулу:

$$c_j = \frac{\frac{1}{r_j}}{\sum_{j=1}^{k} \frac{1}{r_j}}, \qquad (6)$$

7. При очень низком коэффициенте согласованности ($\omega < 0,5$) для каждого эксперта рассчитывают *коэффициент ранговой корреляции Спирмена*, который показывает тесноту связи мнения конкретного эксперта с общим мнением группы:

$$r_s = 1 - \frac{6 \cdot \sum_{i=1}^{n} \left(d_i^2\right)}{k(k^2 - 1)} \ , \qquad (7)$$

где d_i^2 - квадраты разности рангов;

k – число пар рангов.

Коэффициент ранговой корреляции может изменяться в пределах от 0 до 1. Мнение того эксперта, который имеет наименьший коэффициент ранговой корреляции в группе, может быть исключено из дальнейшей обработки данных. В этом случае коэффициент согласованности рассчитывается заново

МЕТОДИКА ВЫПОЛНЕНИЯ РАБОТЫ

Для оценкм качества одежды предложен список из семи показателей: x_1 – уровень обработки изделия; x_2 – современность и качество применяемых материалов; x_3 - современность модели изделия; x_4 – стоимость изделия; x_5 – функциональность изделия; x_6 – антропологические характеристики; x_7 – выразительность фирменного знака и совершенство упаковки.

Необходимо определить какие из этих показателей оказывают наиболее существенон влияние.

Для принятия решения список показателей был роздан 14 специалистам, каждый их которых отметил показатели по степени их важности. Результаты опроса мнения специалистов привелены в табл. 1.

Таблица 1- Матрица рангов

Специалист	Показатели							r_S
	x_1	x_2	x_3	x_4	x_5	x_6	x_7	
1	3	2	1	5	4	6	7	
2	2	4	1	3	5	6	7	
3	3	2	1	7	5	4	6	
4	3	2	1	7	5	4	6	
5	2	3	1	5	4	6	7	
6	2	3	1	6	4	5	7	
7	2	3	1	5	4	6	7	
8	4	2	1	6	3	5	7	
9	1	3	2	6	4	5	7	
10	3	2	1	5	4	6	7	
11	2	3	1	6	4	5	7	
12	3	4	1	5	2	6	7	
13	3	2	1	4	5	7	6	
14	4	2	1	6	3	5	7	
r_j								
a_{ij}\								
Δ_j								
Δ^2_j								
$1/r_j$								
c_j								

Обработка результатов исследования ведется по п. 1-7, описанным в теоретической части.

Библиографический список

1. Добровольская, Т.А. Методы и средства исследования технологических процессов и объектов легкой промышленности [Текст]: учебное пособие / Т.А. Добровольская, Т.И. Леонтьева; Курск.гос. техн. ун-т. Курск: ЮЗГУ, 2006 г. - 190 с.

2. Гмурман, В. Е. Теория вероятностей и математическая статистика [Текст] : учебное пособие для бакалавров / В. Е. Гмурман. - 12-е изд. - М. : Юрайт, 2012. - 479 с.

МИНОБРНАУКИ РОССИИ

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

Кафедра дизайна и технологии изделий легкой промышленности

УТВЕРЖДАЮ
Проректор по учебной работе
_______ О.Г. Локтионова
«15» 12 2017 г.

СТАТИСТИЧЕСКОЕ ИССЛЕДОВАНИЕ СОВОКУПНОСТИ СЛУЧАЙНЫХ ВЕЛИЧИН

Методические указания
по выполнению лабораторных работ
для студентов направления подготовки 29.03.05

КУРСК 2017

УДК 687.01:51-74

Составитель: Т.А. Добровольская

Рецензент
Кандидат технических наук, доцент *Т.М. Ноздрачева*

Статистическое исследование совокупности случайных величин: методические указания по выполнению лабораторных работ / Юго-Зап. гос. ун-т; сост.: Т.А. Добровольская. - Курск, 2017. - 19 с.: табл. 3.- Библиогр.: с. 19.

Излагаются основные сведения о числовых характеристиках случайных величин и рекомендации по овладению методикой определения основных характеристик совокупности случайных величин, их доверительных интервалов, абсолютных и относительных ошибок, применению критериев согласия для определения вида распределения случайных величин.

Предназначены для студентов направления подготовки 29.03.05 «Конструирование изделий легкой промышленности» дневной и заочной форм обучения

Текст печатается в авторской редакции

Подписано в печать . Формат 60×84 1/16.
Усл.печ.л. . Уч.-изд.л. . Тираж 25 экз. Заказ. Бесплатно.
Юго-Западный государственный университет.
305040, г. Курск, ул. 50 лет Октября, 94

СТАТИСТИЧЕСКОЕ ИССЛЕДОВАНИЕ СОВОКУПНОСТИ СЛУЧАЙНЫХ ВЕЛИЧИН

Цель работы: изучение методов расчета числовых характеристик случайных величин, их достоверности и доверительного объема выборки. Определение вида распределения случайных величин с использованием критерия согласия χ^2.

Задание:

1. По результатам предварительного эксперимента рассчитать основные числовые характеристики случайных величин, их доверительные интервалы, абсолютные и относительные ошибки. Определить доверительный объем измерений.
2. Определить вид закона распределения случайных величин.

Отчет должен содержать:

1. Обработку экспериментальных данных, представленных в виде вариационной таблицы.
2. Все расчеты, указанные в п. 1-2 задания.
3. Выводы по полученным результатам.

ТЕОРЕТИЧЕСКИЕ СВЕДЕНИЯ

Основные понятия

Случайной величиной называют переменную величину, которая в результате испытания принимает только одно из возможных значений, ранее неизвестное и зависящее от случайных причин.

Случайные величины подразделяют на непрерывные и дискретные.

Непрерывные случайные величины могут принимать все значения из некоторого конечного или бесконечного промежутка.

Прерывные (дискретные) случайные величины могут принимать только отдельные изолированные значения. Например, число дефектов в выборке, число остановов ткацкого станка в заданный период времени выражаются только целочисленными значениями.

Статистическая величина, принимая разные значения при испытании, образует *статистическую совокупность* из однородных элементов, обладающих качественной общностью. Статистические совокупности подразделяются на генеральные и выборочные.

Генеральная совокупность состоит из множества однородных элементов, имеющих общие свойства, оцениваемые статистическим характеристиками и охватывает все массовое явление (например, вся продукция, все сырье некоторой фабрики, все население, которое обеспечивается одеждой, изготовляемой на швейной фабрике).

Выборочная совокупность состоит из части объектов наблюдений генеральной совокупности, отобранной определенным образом. Из выборочной совокупности определяют выборочные характеристики с тем, чтобы по ним судить о генеральных характеристиках. Число элементов выборки называется ее *объемом.*

Всякое статистическое исследование начинается с отбора какого-то количества членов генеральной совокупности, испытания их по какому-либо признаку и записи результатов измерений. В результате получается выборочная совокупность в виде первоначальной таблицы вариантов.

Вторым этапов статистического исследования является упорядочение первоначальной таблицы. Достигается это путем составления вариационного ряда, то есть переписи всех членов первоначальной таблицы в порядке возрастания (убывания) вариантов или путем составления таблицы распределения частот.

Частотой m_i вариантов при дискретном изменении признака называется число одинаковых вариантов в выборочной совокупности, а при непрерывном изменении признака – число вариантов, попадающих на тот или иной из частных интервалов ΔY_i, на которые разбивается общий интервал изменения признака в выборке.

Таблицей распределения частот при дискретном изменении признака называется таблица, состоящая из отличных один от другого вариантов, записанных в порядке возрастания с указанием их частот; в случае непрерывного изменения признака это таблица, состоящая из частных интервалов ΔY_i или их середин Y_i с указанием частот вариантов, приходящихся на каждый из этих интервалов.

Правила для составления таблицы распределения частот:

1. Разделить общий интервал на k частных интервалов (классы). Число классов k зависит от объема выборки m (табл.1)

Таблица 1

m	40-60	61-100	101-200	201-300	301-500
k	5-7	7-10	10-13	13-15	15-18

2. Определить классовый интервал по формуле:

$$\Delta Y = \frac{Y_{max} - Y_{min}}{k}, \qquad (1)$$

где Y_{max} – наибольшее значение признака в выборке;

Y_{min} – наименьшее значение признака в выборке

3. Определить принадлежность каждого из вариантов к тому или иному частному интервалу.
4. Найти частоту m_i вариантов, приходящихся на каждый интервал путем подсчета точек.

Таблицы распределения частот представляют собой таблично-заданные функции, связывающие частоты m_i с частными интервалами ΔY_i изменения признака Y. Указанные функции, характеризующие распределение признака Y в выборочных совокупностях, называются эмпирическими (полученными из опыта) законами распределения признака.

Эмпирические законы распределения могут быть изображены графически в виде полигонов или гистограмм.

Теоретические законы распределения строят по известным функциональным зависимостям. Сопоставление эмпирических и теоретических кривых позволяет более глубоко изучить выборочные статистические совокупности. Из теоретических кривых чаще всего принимают нормальную кривую. Графически закон нормального распределения выражается симметричной, одновершинной плавной кривой.

Уравнение нормальной кривой описывается уравнением:

$$f(y) = -\frac{1}{\sigma\sqrt{2\pi}} \cdot e^{\frac{-(y-M)^2}{2\sigma^2}}, \qquad (2)$$

где f(y) – частота встречаемого признака;

M – средняя арифметическая величина;

σ и σ^2 – среднее квадратическое отклонение и дисперсия признака.

Если принять $\sigma = 1$ и вместо $\frac{x - M}{\sigma}$ подставить u, то уравнение нормальной кривой примет вид

$$\varphi(u) = -\frac{1}{\sqrt{2\pi}} \cdot e^{\frac{-u^2}{2}}.$$

В таком виде уравнение носит название функции нормированного отклонения, а величина u является нормированным отклонением. Таблицы значения $\varphi(u)$ называются таблицами ординат кривой нормального распределения. По этим таблицам можно рассчитать теоретическую кривую нормального распределения для любого эмпирического вариационного ряда.

Сводные выборочные характеристики

Средняя величина признака в выборке

Среднее значение определяет центр распределения случайных величин, около которого группируется большая их часть. Этот центр характеризуется *средней арифметической (начальным моментом)*:

$$M\{Y\} = \mu_0 = \frac{1}{m} \cdot \sum_{i=1}^{m} Y_i, \qquad (3)$$

где Y_i - случайная величина;

m – число наблюдений.

В практике исследований текстильной и легкой промышленности для характеристики среднего уровня применяют помимо средней арифметической величины моду и медиану.

Модой (M_o) называется наиболее часто встречающаяся величина, а модальным классом вариационного ряда такой класс, на который приходится наибольшее число случаев.

Медианой (M_e) называется такое значение признака, которое делит всю совокупность на две равные части и представляет собой центральную величину.

Мера рассеяния признака в выборке

Существует несколько видов меры рассеяния.

Размах $R = Y_{max} - Y_{min}$

Абсолютной характеристикой рассеяния случайной величины Y около центра распределения M{Y} является *дисперсия* $S^2\{Y\}$ *(центральный момент второго порядка)* или *среднее квадратическое* отклонение S{Y}.

$$S^2\{Y\} = \mu_2 = \frac{1}{m} \cdot \sum_{i=1}^{m} (Y_i - M\{Y\})^2 . \qquad (4)$$

$$S\{Y\} = \sqrt{S\{Y\}^2} . \qquad (5)$$

При нормальном распределении практически вся совокупность признаков заключена в пределах $M\{Y\} \pm 3S\{Y\}$.

При малом числе наблюдений ($m \leq 30$) дисперсия определяется по следующей формуле:

$$\bar{S}^2\{Y\} = \frac{1}{m-1} \cdot \sum_{i=1}^{m} (Y_i - M\{Y\})^2 . \qquad (6)$$

Коэффициент вариации является относительной характеристикой рассеяния случайной величины:

$$CV\{Y\} = \frac{S\{Y\}}{M\{Y\}} . \qquad (7)$$

Если коэффициент вариации выражается в процентах, то эта величина называется *квадратической неровнотой*:

$$C\{Y\} = \frac{S\{Y\}}{M\{Y\}} \cdot 100\% . \qquad (8)$$

Показатель асимметрии. Эксцесс

Эмпирическое распределение, как правило, не имеет точного совпадения с теоретическим. Погрешность, возникающая из-за несоответствия эмпирических и теоретических кривых распределения, является следствием наличия в эмпирическом распределении асимметрии и эксцесса.

Величиной *эксцесса* характеризуют вершинность кривой нормального распределения. Кривые распределения разделяют на высоковершинные (положительный эксцесс), низковершинные и многовершинные (отрицательный эксцесс). Коэффициент эксцесса представляет собой отношение центрального момента четвертого порядка к среднему квадратическому отклонению в четвертой степени и определяется по формуле:

$$E = \frac{\mu_4}{S\{Y\}^4} = \frac{\sum_{i=1}^{m}(Y_i - M\{Y\})^4}{m \cdot S\{Y\}^4} - 3. \tag{9}$$

При *асимметричном* распределении наблюдается увеличение частот в правой (правосторонняя или положительная асимметрия) или левой (левосторонняя или отрицательная асимметрия) половине кривой. Показателем асимметрии является коэффициент асимметрии A, определяемый отношением центрального момента третьего порядка к среднему квадратическому отклонению в кубе:

$$A = \frac{\mu_3}{S\{Y\}^3} = \frac{\sum_{i=1}^{m}(Y_i - M\{Y\})^3}{m \cdot S\{Y\}^3}. \tag{10}$$

Статистическая проверка гипотез о законе распределения

Расхождение между эмпирическим и теоретическим распределениями обычно оценивается с помощью критериев согласия.

Критерий согласия – критерий гипотезы о том, что генеральная совокупность имеет распределение предполагаемого типа (например, нормальное). Наиболее распространенный критерий согласия χ^2 (Пирсона).

Функция χ^2 представляет собой сумму квадратов отклонений эмпирических численностей от теоретической, поделенных на теоретическую численность:

$$\chi^2 = \sum \frac{m_i - \overline{m}_i}{\overline{m}_i},$$

где m_i – эмпирическая частота в каждом i-том классе значений признака;

$\overline{m}_i$ - теоретическая частота в каждом i-том классе значений признака.

Проверку гипотезы о виде функции распределения с помощью этого критерия проводят следующим образом:

1. а) по выборке строят гистограмму.

Если в каком-либо интервале число наблюдений окажется меньше 5, то его объединяют с соседним интервалом.

б) задаются видом гипотетической функции распределения и устанавливают от каких параметров она зависит.

в) определяют теоретическую вероятность попадания в каждый из интервалов случайной величины с заданным распределением:

$$p_i = F(Y_i, r_1, r_2, \ldots) - F(Y_{i-1,} r_1, r_2 \ldots). \quad (11)$$

г) вычисляют

$$\chi^2_q = \sum_{i=1}^{k} \cdot \frac{(m_i - mp_i)^2}{mp_i} = \frac{1}{m} \cdot \sum_{i=1}^{m} \frac{m_i^2}{p_i} - m. \quad (12)$$

2. Величина χ^2_q для данного критерия согласия имеет χ^2 распределение с ν=m-1-r степенями свободы, где r – число неизвестных параметров теоретического распределения, определяемых по данным выборки. Задавшись уровнем значимости q, по таблице χ^2 распределения находят критическое значение $\chi^2_{кр}$.
3. Сравнивают $\chi^2_{кр}$ и χ^2_q и выносят решение о принятии ($\chi^2_q \le \chi^2_{кр}$) или отклонении ($\chi^2_q > \chi^2_{кр}$) рассматриваемой гипотезы о функции закона распределения.

Критерий χ^2 используется при количестве измерений > 50

Упрощенные способы вычисления статистических характеристик

Для вычисления различных статистических характеристик пользуются моментами второго, третьего, четвертого порядка. Целесообразно для вычисления этих моментов пользоваться упрощенными способами вычислений, одним из которых является *способ моментов (произведений).*

В соответствии с этим методом основные статистические характеристики определяют: среднее арифметическое по формуле (13), дисперсию по формуле (14), коэффициент асимметрии по формуле (15), коэффициент эксцесса по формуле (16).

$$M_В \approx a_0 + \Delta Y \cdot \frac{S_1}{m}; \quad (13)$$

$$S_B^2 \approx \frac{\Delta Y^2}{m} \cdot (S_2 - \frac{S_1^2}{m}); \tag{14}$$

$$A_B \approx \frac{\Delta Y^3}{m^3 \cdot S_B^3} \cdot (m^2 \cdot S_3 - 3 \cdot m \cdot S_1 \cdot S_2 + 2 \cdot S_1^3); \tag{15}$$

$$E_B \approx \frac{\Delta Y^4}{m^4 \cdot S_B^4} \cdot (m^3 \cdot S_4 - 4 \cdot m^2 \cdot S_3 \cdot S_1 + 6 \cdot m \cdot S_2 \cdot S_1^2 - 3 \cdot S_1^4) - 3, \tag{16}$$

где

$$S_1 = \sum_{i=1}^{k} m_i \cdot a_i; \quad S_2 = \sum_{i=1}^{k} m_i \cdot a_i^2; \quad S_3 = \sum_{i=1}^{k} m_i \cdot a_i^3; \quad S_4 = \sum_{i=1}^{k} m_i \cdot a_i^4;$$

$a_i = \dfrac{Y^* - a_0}{\Delta Y}$ - нормированное значение параметра,

Y^* – середины частных интервалов ΔY;

a_0 – условный нуль, то есть середина частного интервала с наибольшей частотой.

Ошибки и доверительные интервалы оценок числовых характеристик нормального распределения.

В результате измерения параметров технологического процесса или свойств продукции возникают погрешности.

Абсолютной погрешностью (ошибкой) измерения ε_i называют разность между результатом измерения Y_i и действительным значением Y_0 измеряемой величины:

$$\varepsilon_i = Y_i - Y_0. \tag{17}$$

Относительной погрешность измерения δ_i называют отношение абсолютной ошибки к результату измерения:

$$\delta_i = \frac{\varepsilon_i}{Y_i}. \tag{18}$$

Погрешность появляется вследствие изменения параметров объекта во времени, ошибок оператора, инструментальных ошибок, методических ошибок.

Для исследователя важно знать точность и надежность оценки каждого определяемого параметра. Представление о точности и надежности оценок параметра распределения дают доверительные интервалы. Для генеральной совокупности случайных величин выборочная оценка Т любого параметра θ распределения есть случайная величина.

Двусторонним доверительным интервалом называют интервал от $\mathrm{T}-\varepsilon_{\text{Д}}$ до $\mathrm{T}+\varepsilon_{\text{Д}}$, который покрывает неизвестный параметр распределения с заданной доверительной вероятностью $\mathrm{P}_{\text{Д}}$.

Доверительной вероятностью $P_{Д}$ (надежностью) называется вероятность того, что истинное значение числовых характеристик (особенно при $m \geq 60$) лежит в этом интервале, т.е.

$$\mathrm{P}\{\theta_{\text{Н}} = \mathrm{T}-\varepsilon_{\text{Д}} \leq \theta \leq \mathrm{T}+\varepsilon_{\text{Д}} = \theta_{\text{В}}\} = \mathrm{P}_{\text{Д.}} \tag{19}$$

Величину $\varepsilon_{\text{Д}}$ называют *доверительной гарантийной ошибкой*. Она характеризует случайную ошибку параметра θ распределения и связана со средней квадратической ошибкой $\mathrm{S}^2\{\mathrm{Y}\}$. Чем меньше значение $\varepsilon_{\text{Д}}$, тем больше точность оценки Т.

С увеличением доверительной вероятности $\mathrm{P}_{\text{Д}}$ увеличивается доверительная ошибка $\varepsilon_{\text{Д.}}$ Достоверность получаемой числовой характеристики определяется доверительной вероятность и доверительной ошибкой, которые нужно всегда рассматривать в совокупности.

Обычно значение $\mathrm{P}_{\text{Д}}$ принимают следующее: для поисковых работ $0{,}9 \div 0{,}95$; для исследования процессов и машин $0{,}95 \div 0{,}9$; для контроля качества продукции 0,95-0,99.

В практике исследований параметров процессов и продуктов текстильной промышленности при надежности $\mathrm{P}_{\text{Д}} = 0{,}954$ точность измерения считается высокой, если $\delta\{\mathrm{T}\} \leq 2\%$, средней – если $2\% < \delta\{\mathrm{T}\} \leq 5\%$, низкой – если $5\% < \delta\{\mathrm{T}\} \leq 10\%$ и очень низкой – если $10\% < \delta\{\mathrm{T}\}$.

Абсолютная и относительная ошибки среднего. Необходимое число испытаний при определении среднего.

Абсолютная доверительная ошибка, допущенная при оценке среднего значения генеральной совокупности, когда объем выборки большой ($m \geq 100$) и случайная величина Y распределена по нормальному закону, определяется так:

$$\varepsilon\{\overline{\mathrm{Y}}\} = \mathrm{U}\{\mathrm{P}_{\text{Д}}\} \cdot \mathrm{S}\{\mathrm{Y}\} \cdot \frac{1}{\sqrt{\mathrm{m}}}, \tag{20}$$

где m- число измерений;

$U\{P_Д\}$- абсцисса нормального распределения случайной величины, определяемая доверительной вероятностью $P_Д$ из следующего соотношения:

$$P_Д = 2\Phi(U\{P_Д\}), \tag{21}$$

где $\Phi(U\{P_Д\})$ – функция Лапласа.

Относительная доверительная (гарантийная) ошибка среднего значения определяется по формуле:

$$\delta\{\overline{Y}\} = \frac{\varepsilon\{\overline{Y}\}}{\overline{Y}} \cdot 100 = U\{P_Д\} \cdot C\{Y\} \cdot \frac{1}{\sqrt{m}}. \tag{22}$$

Пользуясь формулой (22), можно рассчитать доверительный объем измерений при определении выборочного среднего значения:

$$m\{\overline{Y}\} = \left(\frac{U\{P_Д\} \cdot CV\{Y\}}{\delta\{\overline{Y}\}} \right)^2. \tag{23}$$

Задаваясь величиной относительной ошибки $\delta\{\overline{Y}\}$ и доверительной вероятностью $P_Д$ и приняв квадратическую неровноту (%) по данным других опытов или другой информации, можно рассчитать объем выборки.

Абсолютная и относительная ошибки среднего квадратического отклонения. Объем выборки при определении этой характеристики.

При большом объеме выборки доверительная абсолютная ошибка среднего квадратического отклонения:

$$\varepsilon\{S\} = U\{P_Д\} \cdot S\{Y\} \cdot \frac{1}{\sqrt{2m}}, \tag{24}$$

доверительная относительная ошибка:

$$\delta\{S\} = U\{P_Д\} \cdot \frac{100}{\sqrt{2m}}. \tag{25}$$

Доверительный объем выборки при определении $S\{Y\}$:

$$m\{S\} = \frac{1}{2} \cdot \left(\frac{U\{P_Д\} \cdot 100}{\delta\{S\}} \right)^2. \tag{26}$$

Абсолютная и относительная ошибки коэффициента вариации. Объем выборки при определении этой характеристики

При большом объеме выборки $(m \geq 30)$ абсолютную и относительную ошибки коэффициента вариации случайной величины Y из нормальной генеральной совокупности определяют следующим образом:

$$\varepsilon\{\mathrm{CV}\} = \mathrm{U}\{\mathrm{P_Д}\} \cdot \mathrm{CV}\{\mathrm{Y}\} \cdot \sqrt{\frac{1 + 2 \cdot \mathrm{CV}^2\{\mathrm{Y}\}}{2 \cdot \mathrm{m}}}, \tag{27}$$

$$\delta\{\mathrm{CV}\} = 100 \cdot \mathrm{U}\{\mathrm{P_Д}\} \cdot \sqrt{\frac{1 + 2 \cdot \mathrm{CV}^2\{\mathrm{Y}\}}{2 \cdot \mathrm{m}}}. \tag{28}$$

Преобразуя формулу (28), получим выражение для определения объема выборки при определении коэффициента вариации с необходимой точностью:

$$\mathrm{m}\{\mathrm{CV}\} = \frac{100^2}{2} \cdot \frac{\mathrm{U}^2\{\mathrm{P_Д}\}}{\delta^2\{\mathrm{CV}\}} \cdot \sqrt{1 + 2 \cdot \mathrm{CV}^2\{\mathrm{Y}\}}. \tag{29}$$

МЕТОДИКА ВЫПОЛНЕНИЯ РАБОТЫ

Исследовалась прочность на разрыв полосок сатина по утку (в *дан*). Проведено 90 измерений

41,1	43,0	39,7	40,4	43,5	41,6	44,5	42,8	42,0
42,9	39,5	40,3	43,4	41,5	38,9	42,7	41,9	39,9
39,3	40,2	40,6	43,3	45,8	42,6	41,8	39,8	40,5
40,1	43,2	40,5	41,5	42,5	41,7	39,7	40,4	43,3
43,1	41,4	41,2	42,4	41,6	39,6	40,3	43,2	42,0
41,3	38,3	42,3	41,5	39,5	40,2	43,1	41,9	42,5
44,1	42,2	41,5	39,4	40,1	43,9	41,8	42,4	41,0
42,1	41,3	39,3	40,9	43,8	41,7	42,3	40,9	41,3
41,2	39,2	40,9	43,7	41,6	42,2	40,9	41,2	42,8
39,1	40,7	43,6	41,5	42,1	40,7	41,1	42,7	39,8

1. Определим основные числовые характеристики для данной выборки, их абсолютные и относительные ошибки, значения гра-

ниц доверительных интервалов при доверительной вероятности $P_Д$ = 0,954, доверительный объем измерений при определении числовых характеристик для обеспечения относительной ошибки этих характеристик 5%.

1.1. Составим таблицу распределения частот.

Определим величину классового интервала по формуле (1).

$$\Delta Y = \frac{45{,}8 - 38{,}3}{12} = 0{,}63$$

Определим границы классов, распределим все значения по классам и определим частоту значений по классам. Результаты сведем в табл. 2

Таблица 2

Граница классов	Частота m	Y^*	a	m·a	$m \cdot a^2$	$m \cdot a^3$	$m \cdot a^4$
38,3-38,93	2	38,62	-5	-10	50	-250	1250
38,93-39,56	7	39,25	-5	-28	112	-448	1792
39,56-40,19	8	39,88	-3	-24	72	-216	648
40,19-40,82	11	40,5	-2	-22	44	-88	176
40,82-41,45	14	41,14	-1	-14	14	-14	14
41,45-42,08	16	41,77	0	0	0	0	0
42,08-42,71	14	42,4	1	14	14	14	14
42,71-43,34	7	43,03	2	14	28	56	112
43,34-43,97	6	43,66	3	18	54	162	486
43,97-44,6	3	44,29	4	12	48	192	768
44,6-45,23	0	44,9	5	0	0	0	0
45,23-45,86	1	45,55	6	6	36	216	1296
$\sum$	90			$S_1 = -34$	$S_2 = 472$	$S_3 = -376$	$S_4 = 6556$

1.2. Рассчитаем основные числовые характеристики с использованием метода произведений по формулам (13,14).

$$M_В \approx 41{,}77 + 0{,}63 \cdot \frac{-34}{90} \approx 41{,}532;$$

$$S_В^2 \approx \frac{0{,}63^2}{90} \cdot (472 - \frac{34^2}{90}) \approx 2{,}025; \quad S_В = \sqrt{S_В^2} \approx 1{,}423$$

По формулам (7,8) определим коэффициент вариации и квадратическую неровноту:

$$CV\{Y\}=\frac{1,423}{41,532}=0,035;$$

$$C\{Y\}=0,035\cdot 100=3,5\%.$$

1.3. Определим доверительный интервал и ошибки среднего.

В соответствии с формулой (21) и пользуясь значениями нормированной функции Лапласа, для $\Phi(U\{P_Д\})=\frac{P_Д}{2}=\frac{0,954}{2}=0,477$ находим $U\{P_Д\}$=2.

Используя формулы (20,22), определим абсолютную и относительную ошибки:

$$\varepsilon\{\overline{Y}\}=2\cdot 1,423\cdot\frac{1}{\sqrt{90}}=0,304;$$

$$\delta\{\overline{Y}\}=2\cdot 3,5\cdot\frac{1}{\sqrt{90}}=0,731\%.$$

Границы доверительного интервала для среднего значения будут следующими:

$$41,228 = 41,532-0,304 \le \eta \le 41,532+0,304 = 41,836$$

1.4. Определим доверительный интервал и ошибки среднего квадратического отклонения.

По формулам (24,25) имеем:

$$\varepsilon\{S\}=2\cdot 1,423\cdot\frac{1}{\sqrt{2\cdot 90}}=0,215,$$

$$\delta\{S\}=2\cdot\frac{100}{\sqrt{2\cdot 90}}=14,9\%.$$

Границы доверительного интервала для среднего квадратического отклонения будут следующими:

$$1,119 = 1,423-0,215 \le \sigma \le 1,423+0,215 = 1,727$$

Определим объем выборки по (26) для обеспечения относительной ошибки среднего квадратического отклонения 5%:

$$m\{S\} = \frac{1}{2} \cdot \left(\frac{2 \cdot 100}{5}\right)^2 = 800.$$

1.5. Определим доверительный интервал и ошибки коэффициента вариации

По формулам (27,28) имеем:

$$\varepsilon\{CV\} = 2 \cdot 0{,}035 \cdot \sqrt{\frac{1 + 2 \cdot 0{,}035^2}{2 \cdot 90}} = 5{,}175 \cdot 10^{-3},$$

$$\delta\{CV\} = 100 \cdot 2 \cdot \sqrt{\frac{1 + 2 \cdot 0{,}035^2}{2 \cdot 90}} = 14{,}925\%.$$

Границы доверительного интервала для коэффициента вариации будут следующими:

$$0{,}03 = 0{,}035 - 5{,}175 \cdot 10^{-3} \le \gamma \le 0{,}035 + 5{,}175 \cdot 10^{-3} = 0{,}04$$

Таким образом, можно утверждать, что квадратическая неровнота генеральной совокупности находится в интервале $3\% \le \gamma \le 4\%$.

Определим объем выборки по (29) для обеспечения относительной ошибки коэффициента вариации 5%:

$$m\{CV\} = \frac{100^2}{2} \cdot \frac{2}{5} \cdot \sqrt{1 + 2 \cdot 0{,}035^2} = 801.$$

2. Определим вид распределения с использованием критерия согласия Пирсона.

Исследуем соответствует ли распределения экспериментальной выборки нормальному закону распределения.

При нормальном распределении выражение (11) примет вид:

$$p_i = \Phi_0((Y_i - M)/\sigma) - \Phi_0((Y_{i-1} - M)/\sigma).$$

Значения $\Phi_0\left(\frac{Y_i - M}{\sigma}\right)$ определяются с помощью таблиц нормированной функции Лапласа.

Результаты расчетов сведем в табл. 3

Таблица 3

Граница классов		m	Y_i - M	$\dfrac{Y_i - M}{\sigma}$	$\Phi_0\left(\dfrac{Y_i - M}{\sigma}\right)$	p_i	$\dfrac{m_i^2}{p_i}$
Y_{i-1}	Y_i						
38,93	39,56	9	-1,972	-1,386	-0,4171	0,096	843,75
39,56	40,19	8	-1,342	-0,943	-0,3272	0,069	927,536
40,19	40,82	11	-0,712	-0,5	-0,1915	0,136	889,706
40,82	41,45	14	-0,082	-0,058	-0,0231	0,168	1167
41,45	42,08	16	0,548	0,385	0,1499	0,173	1480
42,08	42,71	14	1,178	0,828	0,2962	0,148	1321
42,71	43,34	7	1,808	1,271	0,3881	0,092	532,609
43,34	43,97	6	2,438	1,713	0,4566	0,069	521,739
43,97	44,6	4	3,068	2,158	0,4845	0,028	571,429
$\sum$		90				1	8576

Используя формулу (12) определим:

$$\chi_q^2 = \frac{8576}{90} - 90 = 5{,}289$$

Для ν=9-1-2=6 степеней свободы и q=0,05 по таблице χ^2 распределения определяем $\chi^2_{кр}$ = 12,6.

Так $\chi^2_q < \chi^2_{кр}$, то гипотезу о нормальном распределении случайной величины принимаем.

Определим коэффициенты асимметрии и эксцесса для эмпирического распределения, используя формулы (15,16) и данные табл. 2.

$$A_B \approx \frac{0{,}63^3}{90^3 \cdot 1{,}423^3} \cdot (90^2(-376) - 3 \cdot 90 \cdot (-34) \cdot 472 + 2 \cdot (-34)^3) \approx 0{,}091;$$

$$E_B \approx \frac{0{,}63^4}{90^4 \cdot 1{,}423^4}(90^3 \cdot 6556 - 4 \cdot 90^2 \cdot (-376) \cdot (-34) + 6 \cdot 90 \cdot 472 \cdot 34^2 - 3 \cdot 34^4) - 3$$

$$E_B \approx -0{,}274$$

Таким образом, распределение имеет небольшую положительную правую асимметрию и отрицательный эксцесс, то есть кривая эмпирического распределения имеет более плоскую вершину, чем кривая теоретического нормального распределения.

Погрешность эмпирического распределения может быть вычислена:

$$П=(|0,125 \cdot A|+|0,058 \cdot E|) \cdot 100.$$

Погрешность, превышающая 5 % считается значимой.

$$П=(|0,125 \cdot 0,091|+|0,058 \cdot (-0,274|) \cdot 100=2,7\%$$

Коэффициенты асимметрии и эксцесса для данного вариационного ряда, а также погрешность (П<5%) показывают, что отклонение эмпирического распределения от нормального несущественно.

Контрольные вопросы:

1. Понятие случайной величины.
2. Какие случайные величины относятся к непрерывным, а какие к дискретным?
   Привести примеры.
3. Понятие генеральной и выборочной совокупности.
4. Какие функции относятся к эмпирическим законам распределения?
5. Что показывает среднее значение случайной величины? Как оно определяется?
6. Чем выражается мера рассеяния случайной величины?
7. Как определяется дисперсия и среднее квадратическое отклонение случайной величины?
8. Что такое коэффициент вариации? Его сущность.
9. Что характеризует асимметрия и эксцесс?
10. Понятие абсолютной и относительной погрешности?
11. Что такое доверительный интервал числового значения характеристик случайной величины?
12. Что такое доверительный объем измерений? Как его рассчитать при определении различных характеристик случайной величины (среднего, среднего квадратического отклонения, коэффициента вариации)?
13. Сущность проверки гипотезы о виде распределения с помощью критерия согласия Пирсона.

Библиографический список

1. Добровольская, Т.А. Методы и средства исследования технологических процессов и объектов легкой промышленности [Текст]: учебное пособие / Т.А. Добровольская, Т.И. Леонтьева; Курск.гос. техн. ун-т. Курск: ЮЗГУ, 2006 г. - 190 с.
2. Гмурман, В. Е. Теория вероятностей и математическая статистика [Текст] : учебное пособие для бакалавров / В. Е. Гмурман. - 12-е изд. - М. : Юрайт, 2012. - 479 с.
3. Сизенов, Л.К. Моделирование и оптимизация точности технологических процессов [Текст]: учебное пособие / Л.К. Сизенов. - М.: МГТУ им. А.Н. Косыгина, 2001. -330 с.

МИНОБРНАУКИ РОССИИ

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

Кафедра дизайна и технологии изделий легкой промышленности

УТВЕРЖДАЮ
Проректор по учебной работе
_______________ О.Г. Локтионова
«_1_» _02_ _______ 2018 г.

ОПРЕДЕЛЕНИЕ КОРРЕЛЯЦИОННЫХ МОДЕЛЕЙ ПРИ ИССЛЕДОВАНИИ ОБЪЕКТОВ ЛЕГКОЙ ПРОМЫШЛЕНОСТИ

Методические указания
по выполнению лабораторных работ
для студентов направления подготовки 29.03.05

Курск 2018

УДК 687.02

Составитель: Т.А. Добровольская

Рецензент
Кандидат технических наук, доцент *Т.М. Ноздрачева*

Определение корреляционных моделей при исследовании объектов легкой промышленности: методические указания по выполнению лабораторных работ / Юго-Зап. гос. ун-т; сост.: Т.А. Добровольская. Курск, 2018. 18 с. Библиогр.: с. 12.

Излагаются основные сведения о методике обработки результатов пассивного эксперимента при построении однофакторных и многофакторных корреляционных моделей. Указывается порядок оформления лабораторной работы.

Предназначены для студентов направления подготовки 29.03.05 «Конструирование изделий легкой промышленности» дневной и заочной форм обучения.

Текст печатается в авторской редакции

Подписано в печать . Формат 60×84 1/16.
Усл.печ.л. . Уч.-изд.л. . Тираж 25 экз. Заказ . Бесплатно
Юго-Западный государственный университет.
305040, г. Курск, ул. 50 лет Октября, 94.

Лабораторная работа

ОПРЕДЕЛЕНИЕ КОРРЕЛЯЦИОННЫХ МОДЕЛЕЙ ПРИ ИССЛЕДОВАНИИ ОБЪЕКТОВ ЛЕГКОЙ ПРОМЫШЛЕНОСТИ

Цель работы: изучение методики обработки данных пассивного эксперимента и получения однофакторных и многофакторных корреляционных зависимостей.

Задание:

I. Построить однофакторную корреляционную модель:

1. Провести пассивный эксперимент, представить данные в виде двух последовательностей сопряженных случайных чисел.
2. Построить корреляционную таблицу.
3. Рассчитать парный коэффициент корреляции и определить корреляционную модель.
4. Проверить коэффициент корреляции и коэффициент регрессии в корреляционной модели на значимость.
5. Сделать выводы.

II. Построить многофакторную корреляционную модель:

1. Провести пассивный эксперимент, представить данные в виде трех последовательностей сопряженных случайных чисел.
2. Рассчитать множественный коэффициент корреляции.
3. Определить коэффициенты регрессии в двухфакторной корреляционной модели.
4. Проверить множественный коэффициент корреляции и коэффициенты регрессии на значимость.
5. Рассчитать частные коэффициенты корреляции.
6. Сделать выводы.

МЕТОДИКА ВЫПОЛНЕНИЯ РАБОТЫ

I. По замерам обхвата груди X и роста Y у 20 11-летних мальчиков найти однофакторную корреляционную модель зависимости Y(X).

x	62,6	64,1	67,6	66	69,2	67,6	67,3	64,6	65	70	69	72,4	70	70	71,2	67,8	67,4	68,5	65,9	65,5
y	129	130	135	135	136	137	138	139	139	139	139	142	140	140	139	134	140	141	132	132

1. Составим корреляционную таблицу (табл. 1).

1.1 Определим величину классовых интервалов:

$$\Delta \tilde{o} = \frac{\max X - \min X}{5}; \qquad \Delta y = \frac{\max Y - \min Y}{5}$$

1.2. Сформируем таблицу двумерного распределения (выделенная часть табл. 1): разделим совокупность двух величин на классы, определим среднее значение каждого класса ($\overline{Y}_i^*$, $\overline{X}_j^*$), найдем частоту попадания в каждый классовый интервал пары значений из приведенной выше двумерной совокупности.

1.3. В строке №1 и столбце №1 таблицы 1 подсчитываем число попаданий значений совокупности в каждый классовый интервал.

1.4. В строку №2 вносим условные средние значения Y:

$$\overline{y}_{ix} = \frac{1}{m_{xj}} \cdot \sum m_{ji} \overline{Y}_i^*, \text{ т.е.}$$

$$\overline{y}_{1x} = \frac{130{,}3 \cdot 2}{2} = 130{,}3$$

$$\overline{y}_{2x} = \frac{132{,}9 \cdot 2 + 135{,}5 \cdot 1 + 138{,}1 \cdot 2}{5} = 135{,}5$$

…………………………………………………

В столбец №2 вносим условные средние значения X:

$$\overline{\tilde{o}}_{iy} = \frac{1}{m_{yj}} \cdot \sum m_{ji} \overline{X}_j^*, \text{ т.е.}$$

$$\overline{\tilde{o}}_{1y} = \frac{63{,}6 \cdot 2}{2} = 63{,}6$$

$$\overline{x}_{2y} = \frac{65{,}6 \cdot 2 + 67{,}6 \cdot 1}{3} = 66{,}26$$

……………………………………………………

1.5. При заполнении строки и столбца №3 напротив класса с наибольшей частотой необходимо поставить 0. Затем вправо или вниз ставится +1, +2 и т.д.; влево или вверх -1, -2 и т.д.

Далее заполняем строки №4, №5 и столбцы №4 и №5.

1.6. Для заполнения столбца №6 необходимо выполнить следующие расчеты:

y_1=130,3 $\sum^{x} mx_y \alpha_x = 2 \cdot \alpha_{\tilde{o}1}$

y_2=132,9 $\sum^{x} mx_y \alpha_x = 2 \cdot \alpha_{\tilde{o}2} + 1 \cdot \alpha_{\tilde{o}3}$

………………………………………..

Для заполнения строки №6 необходимо выполнить следующие расчеты:

x_1=63,6 $\sum^{y} mx_y \alpha_y = 2 \cdot \alpha_{y1}$

x_2=65,6 $\sum^{y} mx_y \alpha_y = 2 \cdot \alpha_{y2} + 1 \cdot \alpha_{y3} + 2 \cdot \alpha_{y4}$

……………………………………………….

1.7. Заполним строки и столбцы №8 и №9. Вычислим необходимые суммы по строкам и столбцам.

2. Определим коэффициент корреляции. При использовании корреляционной таблицы он вычисляется по формуле:

$$r_{yx} = \frac{\sum^{xy} m_{xy} \alpha_x \alpha_y - n \cdot \overline{\alpha_x} \cdot \overline{\alpha_y}}{n \cdot \sigma(\alpha_x) \cdot \sigma(\alpha_y)}, \qquad (1)$$

где $\overline{\alpha_x} = \frac{S_{1x}}{n}$; $\overline{\alpha_y} = \frac{S_{1y}}{n}$;

$$\sigma^2(\alpha_x) = \frac{1}{n}\cdot(S_{2x} - \frac{S_{1x}^2}{n}); \qquad \sigma^2(\alpha_y) = \frac{1}{n}\cdot(S_{2y} - \frac{S_{1y}^2}{n}).$$

Определим коэффициент детерминации: $KD_r = r_{yx}^2$. Этот коэффициент показывает, сколько % изменений выходного параметра обусловлено изменениями входного параметра.

3. Определим значимость коэффициента корреляции.

   Для этого воспользуемся критерием Стьюдента, расчетное значение которого определяется по формуле:

$$t_R = \frac{r_{yx}}{\sqrt{1 - r_{yx}^2}} \cdot \sqrt{n-2}. \tag{2}$$

Табличное значение: t_T [p_D = 0,95; f = 20-2 = 18) = 2,1. Если $t_R > t_T$, то коэффициент корреляции значим и гипотеза о наличии корреляционной взаимосвязи между X и Y не отвергается.

4. Определим дисперсионное и корреляционное отношение по данным таблицы 1:

$$h_{yx}^2 = \frac{n\cdot\sum\frac{1}{m_x}\cdot\left(\sum m_{xy}\alpha_y\right)^2 - \left(\sum m_{xy}\alpha_y\right)^2}{n\cdot\sum m_y\alpha_y^2 - \left(\sum m_y\alpha_y\right)^2}; \tag{3}$$

$$h_{yx} = \sqrt{h_{yx}^2}.$$

Таблица 1

y \ *x*	*Δx*	*62,6-64,6*	*64,6-66,6*				*№1*	*№2*	*№3*	*№4*	*№5*	*№6*	*№7*	*№8*	*№9*
Δy	$\overline{X}_j^*$ \ $\overline{Y}_i^*$	$x_1 =$ 63,6	$x_2 =$ 65,6	$x_3 =$	$x_4 =$	$x_5 =$	m_y	$\overline{x}_y$	α_y	$m_y\alpha_y$	$m_y\alpha_y^2$	$\sum\limits^{x} m_{xy}\alpha_x$	$\alpha_y\sum\limits^{x} m_{xy}\alpha_x$	$\left(\sum\limits^{x} m_{xy}\alpha_x\right)^2$	$\dfrac{\left(\sum\limits^{x} m_{xy}\alpha_x\right)^2}{m_y}$
129-131,6	$y_1 =130,3$	••					*2*	*63,6*							
131,6-134,2	$y_2 =132,9$		••	•			*3*	*66,3*							
	$y_3 =$		•												
	$y_4 =$		••												
	$y_5 =$														
№1	m_x	*2*	*5*				*20*			$S_{1y}=$	$S_{2y}=$	$\sum =$	$\sum =$	$\sum =$	$\sum =$
№2	$\overline{y}_x$	*130,3*	*135,5*												
№3	α_x														
№4	$m_x\alpha_x$						$S_{1x}=$								
№5	$m_x\alpha_x^2$						$S_{2x}=$								
№6	$\sum\limits^{y} m_{xy}\alpha_y$						$\sum =$								
№7	$\alpha_x\sum\limits^{y} m_{xy}\alpha_y$						$\sum =$						$\sum =$		
№8	$\left(\sum m_{xy}\alpha_y\right)^2$						$\sum =$								
№9	$\dfrac{\left(\sum m_{xy}\alpha_y\right)^2}{m_x}$						$\sum =$								

5. Проверим гипотезу о линейной связи между Y и X. Расчетное значение критерия Фишера определяем по формуле:

$$F_R = \frac{\dfrac{(h_{yx}^2 - r_{yx}^2)}{(k-2)}}{\dfrac{(1-h_{yx}^2)}{(n-k)}},$$

(4)

где k – число классов.

Табличное значение критерия Фишера: F_T [p_D =0,95; $f_{числ}$ = k – 2 = 5 - 2 = 3; $f_{знам}$ = n – k = 20 – 5 = 15] = 3,58. Если $F_R < F_T$, то гипотеза о линейной взаимосвязи между случайными величинами X и Y не отвергается.

6. Получим корреляционные уравнения согласно выражению:

$$Y = \overline{Y} + \rho_y \cdot (X - \overline{X}), \tag{5}$$

где $\rho_y = \dfrac{\sigma_y}{\sigma_x} r_{yx}$ - коэффициент регрессии.

Для этого вначале необходимо определить средние значения и средние квадратические отклонения величин X и Y. С использованием табл. 1 эти величины можно определить по формулам:

$$\overline{x} = a_0 + \frac{S_{1x}}{n} \cdot \Delta x; \qquad \overline{y} = b_0 + \frac{S_{1y}}{n} \cdot \Delta y.$$

(a_0 и b_0 в этих формулах – это середина соответствующего класса с наибольшей частотой)

$$\sigma^2(x) = \frac{\Delta x^2}{n} \cdot \left(S_{2x} - \frac{S_{1x}^2}{n} \right); \qquad \sigma^2(y) = \frac{\Delta y^2}{n} \cdot \left(S_{2y} - \frac{S_{1y}^2}{n} \right).$$

7. Определим значимость коэффициента регрессии в корреляционном уравнении, используя критерий Стьюдента, расчетное значение которого определяется по формуле:

$$t_R = \frac{|\rho_y|}{S\{\rho_y\}} = \frac{|\rho_y| \cdot \sigma_x \cdot \sqrt{n-2}}{\sigma_y \cdot \sqrt{1-r_{yx}^2}}. \tag{6}$$

Табличное значение: t_T [p_D = 0,95; f = 20-2 = 18) = 2,1. Если $t_R\{\rho_x\} > t_T$, то коэффициент значим.

Расчет по корреляционной таблице является приближенным и достаточно трудоемким. Для более точной и быстрой обработки результатов пассивного эксперимента рекомендуется применение ЭВМ. В приложении 1 представлена технология получения однофакторной корреляционной модели с использованием программы Mathcad.

II. По результатам 20 измерений трех сопряженных случайных величии (Y – рост мальчиков; X_1 – обхват груди; X_2 – высота головы) были определены следующие статистические характеристики: $\bar{Y} = 136{,}8$, $\bar{X}_1 = 67{,}62$, $\bar{X}_2 = 25{,}5$; S{Y} = 3,67, S{X_1} = 2,43, S{X_2} = 0,9; $r_{yx_1} = 0{,}7$, $r_{yx_2} = 0{,}56$, $r_{x_1x_2} = 0{,}46$. Требуется определить двухфакторную корреляционную модель Y = f(X_1, X_2).

1. Определим множественный коэффициент корреляции по формуле:

$$R_{YX_1,X_2} = \sqrt{\frac{r_{YX_1}^2 + r_{YX_2}^2 - 2r_{YX_1}r_{YX_2}r_{X_1X_2}}{1-r_{X_1X_2}^2}}. \tag{7}$$

Коэффициент множественной детерминации $KD_R = R_{yx_1x_2}^2$.

2. Среднее квадратическое отклонение множественного коэффициента корреляции определяют по формуле:

$$S\{R_{YX_1X2}\} = \frac{(1-R_{YX_1X2}^2)}{\sqrt{n-2-1}}. \tag{8}$$

3. Определим значимость множественного коэффициента корреляции, используя критерий Стьюдента, расчетное значение которого определяется формуле:

$$t_R\{S_{YX_1X2}\} = \frac{R_{YX_1X_2}}{S\{R_{YX_1X2}\}}. \qquad (9)$$

Табличное значение критерия Стьюдента равно $t_T[p_D = 0{,}95;$ $f = 20 - 2 - 1 = 17] = 2{,}11$. Если $t_R > t_T$, то множественный коэффициент корреляции значим.

4. Коэффициенты двухфакторной корреляционной модели в стандартизированной форме определяются по формулам:

$$q_1 = \frac{r_{YX_1} - r_{YX_2} \cdot r_{X_1X_2}}{1 - r^2_{X_1X_2}}; \qquad (10)$$

$$q_2 = \frac{r_{YX_2} - r_{YX_1} \cdot r_{X_1X_2}}{1 - r^2_{X_1X_2}}. \qquad (11)$$

5. Дисперсии найденных коэффициентов одинаковы и определяются по формуле:

$$S^2\{q_1\} = S^2\{q_2\} = \frac{1 - R^2_{YX_1X_2}}{(1 - r^2_{X_1X_2})(n-2)}. \qquad (12)$$

6. Для определения значимости коэффициентов регрессии определяем t_R:

$$t_R\{q_1\} = \frac{|q_i|}{S\{q_i\}}. \qquad (13)$$

Табличное значение критерия Стьюдента равно $t_T[p_D = 0{,}95;$ $f = 20 - 2 = 18] = 2{,}1$. Если $t_R\{q_1\} > t_T$ и $t_R\{q_2\} > t_T$ то коэффициенты значимы.

7. Коэффициенты многофакторной корреляционной модели с натуральными значениями факторов определяют по формулам:

$$a_i = q_i \frac{S\{Y\}}{S\{X_i\}}; \qquad (14)$$

$$a_0 = \overline{Y} - \sum_{i=1}^{N} a_{ij} \overline{X}_i. \qquad (15)$$

8. Частные коэффициенты корреляции определим по формулам:

$$r_{YX_1(X_2)} = \frac{r_{YX_1} - r_{YX_2} \cdot r_{X_1X_2}}{\sqrt{(1-r^2_{YX_2})(1-r^2_{X_1X_2})}}\;; \qquad (16)$$

$$r_{YX_2(X_1)} = \frac{r_{YX_2} - r_{YX_1} . r_{X_1X_2}}{\sqrt{(1-r^2_{YX_1})(1-r^2_{X_1X_2})}}. \qquad (17)$$

В приложении 2 представлена технология получения двухфакторной корреляционной модели с применением ПЭВМ в программе Mathcad.

Контрольные вопросы

1. Виды корреляционных зависимостей.
2. Что характеризует ковариация и коэффициент корреляции двух статистических величин? В чем их отличие?
3. В каких пределах изменяется парный коэффициент корреляции? О чем говорит его величина?
4. Что показывает коэффициент регрессии в корреляционных уравнениях?
5. Что такое корреляционное отношение? В чем его отличие от коэффициента корреляции?
6. Каким образом можно определить меру линейности корреляционной связи между двумя величинами?
7. Множественный коэффициент корреляции. Что он показывает? В каких пределах изменяется?
8. Для чего коэффициенты регрессии в многофакторных корреляционных уравнениях определяют в стандартизированном виде?
9. Каким образом проверяется значимость коэффициентов регрессии в корреляционных уравнениях?
10. Что показывает частный коэффициент корреляции? В каких пределах он изменяется?

Библиографический список

1. Добровольская Т.А. Методы и средства исследования технологических процессов и объектов легкой промышленности: учебное пособие / Т.А. Добровольская, Т.И. Леонтьева; Курск.гос. техн. ун-т. Курск, 2006 г. 206 с.
2. Добровольская Т.А. Решение инженерно-технических задач на ПЭВМ с использованием программы «Mathcad»: учеб. пособие/ Курск. гос. техн. ун-т. Курск, 2002. 96 с.
3. Кудрявцев Е.М. Mathcad 8. М.: ДМК, 2000. 320 с.
4. Кирьянов Д.В. Самоучитель Mathcad 2001.СПб.: БХВ-Петербург, 2002. 544с.

Приложение 1

Определение однофакторной корреляционной модели

1. Ввод статистических данных

Y - рост мальчиков, X - обхват груди

$i := 0..19 \qquad n := 20$

$Y_i :=$	$X_i :=$
129	62.6
130	64.1
135	67.6
135	66
136	69.2
137	67.6
138	67.3
139	64.6
139	65
139	70
139	69
142	72.4
140	70
140	70
139	71.2
134	67.8
141	67.4
141	68.5
132	65.9
132	65.5

2. Определим парный коэффициент корреляции

$r := \mathrm{corr}(X, Y) \qquad r = 0.703$

3. Определим коэффициент детерминации

$D := r^2 \qquad D = 0.494$

4. Определим значимость коэффициента корреляции

$$tr := \frac{r}{\sqrt{1-r^2}} \cdot \sqrt{n-2} \qquad tr = 4.195$$

5. Определим средние значения

$$My := \mathrm{mean}(Y) \qquad My = 136.85$$

$$Mx := \mathrm{mean}(X) \qquad Mx = 67.585$$

6. Определим средние квадратические отклонения

$$Sy := \mathrm{stdev}(Y) \qquad Sy = 3.719$$

$$Sx := \mathrm{stdev}(X) \qquad Sx = 2.466$$

7. Определим корреляционное отношение

$Syr := Sy \cdot \sqrt{1-r^2}$ - условное среднеквадратическое отклонение

$$Syr = 2.644$$

$$h := \frac{Syr}{Sy} \qquad h = 0.711$$

8. Проверим гипотезу о линейной связи

$$\xi := h^2 - r^2$$

$$\sigma\xi := 2 \cdot \sqrt{\frac{\xi}{n}} \cdot \sqrt{\left(1-h^2\right)^2 - \left(1-r^2\right)^2 + 1}$$

$$t := \frac{\xi}{\sigma\xi} \qquad t = 0.24$$

Если t > 3, то корреляционная зависимость линейная

9. Определим коэффициент регрессии в линейной корреляционной модели

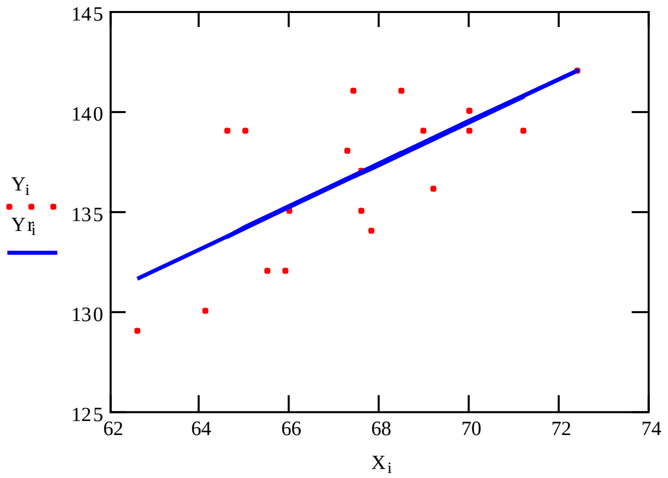
$$\rho := \frac{Sy}{Sx} \cdot r \qquad \rho = 1.06$$

10. Построим корреляционную однофакторную модель

$$Yr := \rho \cdot (X - Mx) + My$$

Yr =

	0
0	131.565
1	133.156
2	136.866
3	135.17
4	138.562
5	136.866
6	136.548
7	133.686
8	134.11
9	139.41
10	138.35
11	141.954
12	139.41
13	139.41
14	140.682
15	137.078
16	136.654
17	137.82
18	135.064



11. Проверим значимость коэффициента регрессии в корреляционном уравнении

$$tqr := \frac{|\rho| \cdot Sx \cdot \sqrt{n-2}}{Sy \cdot \sqrt{1-r^2}} \qquad tqr = 4.195$$

Приложение 2

Определение двухфакторной корреляционной модели

1. Ввод статистических данных

Y - рост мальчиков, X1 - обхват груди, X2 - высота головы

$i := 0..19 \quad n := 20$

$Y_i :=$	$X1_i :=$	$X2_i :=$
129	62.6	23.9
130	64.1	26
135	67.6	26
135	66	24.8
136	69.2	26
137	67.6	24.7
138	67.3	25.8
139	64.6	25.9
139	65	26.2
139	70	27.1
139	69	26.5
139	71.4	25.9
140	70	25.6
140	70	25.4
140	71.2	25.9
134	67.8	25.4
141	67.4	25.3
142	70.2	25.5
132	65.9	23.8
132	65.5	23.4

1. Определим средние значения

$My := mean(Y) \qquad My = 136.8$

$Mx1 := mean(X1) \qquad Mx1 = 67.62$

$Mx2 := mean(X2) \qquad Mx2 = 25.455$

2. Определим средние квадратические отклонения

$Sy := stdev(Y)$ $Sy = 3.669$

$Sx1 := stdev(X1)$ $Sx1 = 2.439$

$Sx2 := stdev(X2)$ $Sx2 = 0.905$

3. Определим парные коэффициенты корреляции

$r1 := corr(Y, X1)$ $r1 = 0.702$

$r2 := corr(Y, X2)$ $r2 = 0.56$

$r12 := corr(X1, X2)$ $r12 = 0.466$

4. Определим множественный коэффициент корреляции

$$R := \sqrt{\frac{r1^2 + r2^2 - 2 \cdot r1 \cdot r2 \cdot r12}{1 - r12^2}}$$ $R = 0.75$

5. Определим множественную детерминацию

$Dr := R^2$ $Dr = 0.562$

6. Определим среднее квадратическое отклонение множественного коэффициента корреляции

$$SR := \frac{1 - R^2}{\sqrt{n - 2 - 1}}$$ $SR = 0.106$

7. Определим значимость множественного коэффициента корреляции

$$tR := \frac{R}{SR}$$ $tR = 7.065$

8. Определим коэффициенты двухфакторной модели в стандартизованной форме

$$q1 := \frac{r1 - r2 \cdot r12}{1 - r12^2} \qquad q2 := \frac{r2 - r1 \cdot r12}{1 - r12^2}$$

$$q1 = 0.563 \qquad q2 = 0.298$$

9. Определим среднее квадратическое отклонение найденных коэффициентов

$$Sq1 := \frac{1 - R^2}{\left(1 - r12^2\right) \cdot (n-2)} \qquad Sq1 = 0.031$$

$$Sq := \sqrt{Sq1} \qquad Sq = 0.176$$

10. Определим значимость коэффициентов регрессии

$$trq1 := \frac{q1}{Sq} \qquad trq1 = 3.196$$

$$trq2 := \frac{q2}{Sq} \qquad trq2 = 1.688$$

11. Определим коэффициенты двухфакторной модели с натуральными значениями факторов

$$a1 := q1 \cdot \frac{Sy}{Sx1} \qquad a2 := q2 \cdot \frac{Sy}{Sx2}$$

$$a0 := My - (a1 \cdot Mx1 + a2 \cdot Mx2)$$

$$a1 = 0.847 \qquad a2 = 1.206 \qquad a0 = 48.799$$

12. Определим частные коэффициенты корреляции

$$Ryx1 := \frac{r1 - r2 \cdot r12}{\sqrt{\left(1 - r2^2\right) \cdot \left(1 - r12^2\right)}} \qquad Ryx1 = 0.602$$

$$Ryx2 := \frac{r2 - r2 \cdot r12}{\sqrt{\left(1 - r1^2\right) \cdot \left(1 - r12^2\right)}} \qquad Ryx2 = 0.475$$

МИНОБРНАУКИ РОССИИ

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

Кафедра дизайна и технологии изделий легкой промышленности

УТВЕРЖДАЮ
Проректор по учебной работе
__________ О.Г. Локтионова
«__»__________ 2018 г.

ОБРАБОТКА ДАННЫХ ОДНОФАКТОРНОГО ЭКСПЕРИМЕНТА ПРИ ИССЛЕДОВАНИЯХ В ЛЕГКОЙ ПРОМЫШЛЕННОСТИ

Методические указания
по выполнению лабораторных работ
для студентов направления подготовки 29.03.05

Курск 2018

УДК 687.02

Составитель: Т.А. Добровольская

Рецензент
Кандидат технических наук, доцент *Т.М. Ноздрачева*

Обработка данных однофакторного эксперимента при исследованиях в легкой промышленности: методические указания по выполнению лабораторных работ / Юго-Зап. гос. ун-т; сост.:Т.А. Добровольская. Курск, 2018. 14 с.: табл. 3. - Библиогр.: с. 9.

Излагаются основные сведения о методике обработки результатов однофакторного эксперимента при построении линейной регрессионной модели. Указывается порядок оформления лабораторной работы.

Предназначены для студентов направления подготовки 29.03.05 «Конструирование изделий легкой промышленности» дневной и заочной форм обучения.

Текст печатается в авторской редакции

Подписано в печать . Формат 60×84 1/16.
Усл.печ.л. . Уч.-изд.л. . Тираж 25 экз. Заказ . Бесплатно
Юго-Западный государственный университет.
305040, г. Курск, ул. 50 лет Октября,94.

Лабораторная работа

ОБРАБОТКА ДАННЫХ ОДНОФАКТОРНОГО ЭКСПЕРИМЕНТА ПРИ ИССЛЕДОВАНИЯХ В ЛЕГКОЙ ПРОМЫШЛЕННОСТИ

Цель работы: изучение методики построения линейной однофакторной регрессионной модели по данным активного эксперимента.

Задание:

1. Провести активный эксперимент.
2. Произвести первичную обработку результатов эксперимента.
3. Определить вид модели.
4. Рассчитать коэффициенты уравнения регрессии.

Отчет должен содержать:

1. Матрицу планирования с результатами эксперимента.
2. Все этапы обработки данных однофакторного эксперимента.
3. Полученную регрессионную модель.
4. Анализ результатов эксперимента.

МЕТОДИКА ВЫПОЛНЕНИЯ РАБОТЫ

Рассмотрим операции, которые совершает исследователь при обработке данных однофакторного эксперимента на примере, в котором исследовалось изменение стойкости бельевых трикотажных полотен к истиранию после опытной носки.

В качестве параметра оптимизации (Y) выбрана стойкость полотна к истиранию; в качестве фактора (X) – число циклов носки.

Было проведено 5 опытов в 5-ти повторностях. По результатам активного эксперимента составлена матрица планирования однофакторного эксперимента (табл. 1).

Таблица 1

Матрица планирования

u	X_u	v							
		Y_{uv}					$\overline{Y}_u$	$S_u^2\{Y\}$	W_R
		1	2	3	4	5			
1	5	2530	2525	2550	2540	2555	2540	162,5	3,774
2	10	2175	2185	2210	2180	2200			
3	15	1865	1855	1840	1855	1835			
4	20	1500	1515	1520	1495	1520			
5	25	1135	1140	1150	1160	1165			

Вначале определяем средние значение и дисперсии выходного параметра для каждого опыта матрицы:

$$\overline{Y}_1 = \frac{2530 + 2525 + 2550 + 2540 + 2555}{5} = 2540$$

$$S_1^2 = \frac{1}{5-1} \cdot [(2530 - 2540)^2 + (2525 - 2540)^2 + \ldots + (2555 - 2540)^2 = 162{,}5 .$$

Для остальных опытов данные величины определяются аналогично. Результаты вычислений заносятся в табл.1

1. Исключение резко выделяющихся данных.

Расчетное значение критерия Смирнова—Грабса определяют по формулам:

при подозрении резко выделяющегося максимального значения $Y_{i\,max}$

$$V_{Rmax} = \frac{(Y_{i\,max} - \overline{Y})}{S\{Y\}} \sqrt{\frac{n}{n-1}} ; \qquad (1)$$

при подозрении резко выделяющегося минимального значения $Y_{i\,min}$

$$V_{R\,min} = \frac{(\overline{Y} - Y_{i\,min})}{S\{Y\}} \sqrt{\frac{n}{n-1}} . \qquad (2)$$

Рассмотрим эту операцию при анализе первого опыта матрицы:

$$V_{R\,max} = \frac{2555 - 2540}{\sqrt{162{,}5}} \cdot \sqrt{\frac{5}{5-1}} = 1{,}316;$$

$$V_{R\,min} = \frac{240 - 2525}{\sqrt{162{,}5}} \cdot \sqrt{\frac{5}{5-1}} = 1{,}316.$$

По таблице критических значений находим V_T. V_T [p_D = 0,95; m = 5] = 1,869. Так как значения $V_{Rmax} < V_T$ и $V_{Rmin} < V_T$, то рассматриваемые значения не являются резко выделяющимися и остаются для дальнейшей статистической обработки.

Аналогично данная операция производится для остальных опытов.

2. Проверка гипотезы о нормальном распределении случайных величин.

Для определения вида распределения в малой выборке используют критерий Шапиро – Уикла, который для каждого u-го опыта матрицы включает определение расчетного значения W_R:

$$W_R = \frac{Q^2}{S_u^2\{Y\}}, \qquad (3)$$

где $Q = q_m(Y_m - Y_1) + \ldots + q_{m-k+1}(Y_{m-k+1} - Y_k)$;

$k = m/2$ – при четном числе m; $k = (m-1)/2$ – при нечетном числе m;

$$Y_m \geq Y_{m-1} \geq \ldots \geq Y_1.$$

В рассматриваемом примере находим: 2555 > 2550 > 2540 > 2530 > 2525

$$Q_1 = 0{,}6646 \cdot (2555 - 2525) + 0.2413 \cdot (2550 - 2530) = 24{,}77$$

$$W_{R1} = \frac{25{,}77^2}{162{,}5} = 3{,}774$$

Аналогично производятся вычисления для других опытов. Результаты заносятся в табл.1.

Расчетное значение W_T [p_D = 0,95; m =5] = 0,762. Так как $W_{R1} > W_T$, то гипотезу о нормальном распределении не отвергаем.

3. Проверка гипотезы об однородности дисперсий в опытах матрицы.

Так как число повторностей одинаково, то используем критерий Кочрена, расчетное значение которого определяется по формуле:

$$G_R = \frac{S_u{}^2{}_{max}\{Y\}}{\sum\limits_{u=1}^{N} S_u^2\{Y\}}. \qquad (4)$$

Табличное значение критерия G_T [p_D = 0,95; N = 5; f = 5-1 = 4] = 0,544. Если $G_R < G_T$, то гипотеза об однородности дисперсий не отвергается.

4. Определение средней дисперсии выходного параметра в опытах матрицы:

$$S_{(1)}^2\{Y\} = \frac{1}{N}\sum_{u=1}^{N} S_u^2\{Y\}. \qquad (5)$$

5. Определение подходящего вида регрессионной модели.

Поскольку интервал варьирования постоянный, то необходимо определить неразделенные разности первого порядка.

Далее нужно сравнить разницу между неразделенными разностями с удвоенной среднеквадратической ошибкой:

Если разница между неразделенными разностями не превышает удвоенной среднеквадратической ошибки, то считаем неразделенные разности первого порядка тождественными и условно принимаем уравнение прямой линии:

$$Y_R = d_0 + d_1 \cdot (X - \overline{X}). \qquad (6)$$

6. Определение коэффициентов регрессии:

$$d_0 = \frac{1}{N}\sum_{u=1}^{N} \overline{y}_u = \overline{y} \quad ; \qquad d_1 = \frac{\sum\limits_{u=1}^{N} (x_u - \overline{x})\overline{y}_u}{\sum\limits_{u=1}^{N} (x_u - \overline{x})^2}.$$

Для удобства сведем расчеты в табл.2.

Таблица 2

u	X_u	$X_u - \overline{X}$	$(X_u - \overline{X})^2$	$\overline{Y_u}$	$(X_u - \overline{X}) \cdot \overline{Y}_u$
1					
2					
3					
4					
5					
$\sum\limits_{u=1}^{N}$					

7. Проверка адекватности полученного уравнения. Расчетное значение критерия Фишера определяется по формуле:

$$F_R = \frac{S^2_{(2)}\{y\}}{S^2_{(1)}\{y\}}, \tag{7}$$

где $S^2_{(1)}\{y\}$-средняя дисперсия или дисперсия воспроизводимости, определяемая в п.4.

$S^2_{(2)}\{y\}$-дисперсия, характеризующая рассеивание средних экспериментальных значений $\bar{y}_u$ относительно прямой линии:

$$S^2_{(2)}\{y\} = \frac{m}{N-2}\sum_{u=1}^{N}\left(\bar{y}_u - y_{ru}\right)^2. \tag{8}$$

Для определения дисперсии, характеризующей рассеивание экспериментальных значений около прямой линии составим табл.3.

Таблица 3

u	X_u	Y_{Ru}	$\overline{Y_u}$	$\overline{Y}_u - Y_{Ru}$	$\left(\overline{Y}_u - Y_{Ru}\right)^2$
1					
2					
3					
4					
5					
$\sum\limits_{u=1}^{N}$					

Используя данные табл. 3 находим $S^2_{(2)}$ и затем определяем F_R.

Табличное значение F_T [p_D = 0,95; $f\left\{S_{(1)}^2\right\} = 5 \cdot 4 = 20$;
$f\left\{S_{(2)}^2\right\} = 5 - 2 = 3$] = 8,66. Если $F_R < F_T$, то гипотезу об адекватности модели не отвергаем.

8. Определение значимости коэффициентов регрессии и их доверительных интервалов.

Для оценки значимости коэффициентов регрессии используется критерий Стьюдента, расчетное значение которого определяют по формуле:

$$t_R\{d_i\} = \frac{|d_i|}{S\{d_i\}}\,, \tag{9}$$

где $S\{d_i\}$-оценка среднего квадратического отклонения коэффициента регрессии d_i.

Для оценки дисперсий коэффициентов регрессии d_0 и d_1 используют формулы:

$$S^2\{d_0\} = \frac{S^2\{y\}}{mN} = \frac{S^2\{\bar{y}\}}{N}; \qquad S^2\{d_1\} = \frac{S^2}{m\sum\limits_{u=1}^{N}\left(x_u - \bar{x}\right)^2} = \frac{S^2\{\bar{y}\}}{\sum\limits_{u=1}^{N}\left(x_u - \bar{x}\right)^2}.$$

Дисперсия $S^2\{y\}$ определяется по формуле:

$$S^2\{y\} = \frac{(m-1)\cdot N \cdot S_{(1)}^2\{y\} + (N-2)\cdot S_{(2)}^2\{y\}}{m \cdot N - 2}.$$

Табличное значение t_T [p_D = 0,95; f = 5·5 -2 = 23] = 2,07. Если $t_R > t_T$, то полученные коэффициенты значимы и, следовательно, связь между X и Y значима.

В приложении 1 представлена технология получения однофакторной регрессионной модели на ПЭВМ с использованием программы Mathcad.

Контрольные вопросы

1. Сущность традиционного однофакторного планирования.
2. Что называется матрицей планирования эксперимента?
3. Для чего каждый опыт в матрице планирования проводят в нескольких повторностях?
4. Какие параметры называют входными, а какие – выходными?
5. Что такое интервал варьирования? Как он определяется?
6. Что включает в себя предварительная обработка результатов эксперимента?
7. С помощью какого критерия можно исключить резко выделяющиеся данные?
8. Основные операции обработки данных однофакторного эксперимента.
9. Как можно определить вид математической модели, получаемой на основе данных однофакторного эксперимента?
10. Что такое неразделенные и разделенные разности? Чем они отличаются?
11. Какой метод лежит в основе определения коэффициентов регрессии однофакторной модели?
12. Понятие адекватности математической модели. Как она проверяется?

Библиографический список

1. Добровольская Т.А. Методы и средства исследования технологических процессов и объектов легкой промышленности [Текст]: учебное пособие / Т.А. Добровольская, Т.И. Леонтьева; Курск.гос. техн. ун-т. Курск, 2006 г. 206 с.
2. Добровольская Т.А. Решение инженерно-технических задач на ПЭВМ с использованием программы «Mathcad»: учеб.пособие/ Курск. Гос. Техн. Ун-т. Курск, 2002. 96 с.
3. Кудрявцев Е.М. Mathcad 8. М.: ДМК, 2000. 320с.
4. Глушаков С.В., Жакин И.А., Хачиров Т.С. Математическое моделирование Mathcad 2000, Matlab 5. Харьков: Фолио; М.: ООО «Издательство АСТ», 2001. 524с.

Приложение 1

Получение линейной однофакторной регрессионной модели

1.Ввод данных активного эксперимента

Количество наблюдений $m := 5 - 1 \quad n := 5 - 1 \quad u := 0..m \quad v := 0..n$

Значения Xu

Значения Yuv

$X_u :=$ $\quad$ $Y_{u,0} := \; Y_{u,1} := \; Y_{u,2} := \; Y_{u,3} := \; Y_{u,4} :=$

X_u	$Y_{u,0}$	$Y_{u,1}$	$Y_{u,2}$	$Y_{u,3}$	$Y_{u,4}$
5	2530	2525	2550	2540	2555
10	2175	2185	2210	2180	2200
15	1865	1855	1840	1855	1835
20	1500	1515	1520	1495	1520
25	1135	1140	1150	1160	1165

$$X = \begin{pmatrix} 5 \\ 10 \\ 15 \\ 20 \\ 25 \end{pmatrix} \qquad Y = \begin{pmatrix} 2.53\times 10^3 & 2.525\times 10^3 & 2.55\times 10^3 & 2.54\times 10^3 & 2.555\times 10^3 \\ 2.175\times 10^3 & 2.185\times 10^3 & 2.21\times 10^3 & 2.18\times 10^3 & 2.2\times 10^3 \\ 1.865\times 10^3 & 1.855\times 10^3 & 1.84\times 10^3 & 1.855\times 10^3 & 1.835\times 10^3 \\ 1.5\times 10^3 & 1.515\times 10^3 & 1.52\times 10^3 & 1.495\times 10^3 & 1.52\times 10^3 \\ 1.135\times 10^3 & 1.14\times 10^3 & 1.15\times 10^3 & 1.16\times 10^3 & 1.165\times 10^3 \end{pmatrix}$$

2.Определение средних значений в каждом опыте матрицы

$$MY_u := \mathrm{mean}\left[\left(Y^T\right)^{\langle u\rangle}\right] \qquad MY = \begin{pmatrix} 2.54\times 10^3 \\ 2.19\times 10^3 \\ 1.85\times 10^3 \\ 1.51\times 10^3 \\ 1.15\times 10^3 \end{pmatrix}$$

3.Определение дисперсий в каждом опыте матрицы

$$S_u := \frac{1}{m}\cdot\left[\sum_v \left(Y_{u,v} - MY_u\right)^2\right] \qquad S = \begin{pmatrix} 162.5 \\ 212.5 \\ 150 \\ 137.5 \\ 162.5 \end{pmatrix}$$

4.Исключение резко выделяющихся значений

$$VRmax_u := \frac{max\left[\left(Y^T\right)^{\langle u\rangle}\right] - MY_u}{\sqrt{S_u}} \cdot \sqrt{\frac{m+1}{m}}$$

$$VRmax = \begin{pmatrix} 1.316 \\ 1.534 \\ 1.369 \\ 0.953 \\ 1.316 \end{pmatrix}$$

$$VRmin_u := \frac{MY_u - min\left[\left(Y^T\right)^{\langle u\rangle}\right]}{\sqrt{S_u}} \cdot \sqrt{\frac{m+1}{m}}$$

$$VRmin = \begin{pmatrix} 1.316 \\ 1.15 \\ 1.369 \\ 1.43 \\ 1.316 \end{pmatrix}$$

5.Проверка гипотезы о нормальном распределении

$k := \frac{m}{2}$ $\quad k = 2$ $\quad Ys_u := sort\left[\left(Y^T\right)^{\langle u\rangle}\right]$

$$Q_u := 0.6646\left[\left(Ys_u\right)_4 - \left(Ys_u\right)_0\right] + 0.2413\left[\left(Ys_u\right)_3 - \left(Ys_u\right)_1\right] \qquad Q = \begin{pmatrix} 24.764 \\ 28.087 \\ 23.557 \\ 21.441 \\ 24.764 \end{pmatrix}$$

$$Wr_u := \frac{\left(Q_u\right)^2}{S_u} \qquad Wr = \begin{pmatrix} 3.774 \\ 3.712 \\ 3.7 \\ 3.343 \\ 3.774 \end{pmatrix}$$

Если Wru > 0.762, то гипотеза о нормальном распределении Yv не отвергается.

6.Проверка гипотезы об однородности дисперсии в опытах матрицы

$$\sum_{i=0}^{m} S_i = 825 \qquad Gr := \frac{max(S)}{\sum_{i=0}^{m} S_i} \qquad Gr = 0.258$$

Если Gr < 0.544, то гипотеза об однородности дисперсий не отвергается

7 . Определяем среднюю дисперсию выходного параметра в опытах матрицы

$$SY := \frac{1}{m+1} \cdot \sum_{i=0}^{m} S_i \qquad SY = 165$$

8.Определяем подходящий вид регрессионной модели по неразделенным разностям первого порядка

$i := 0..m-1 \qquad \Delta R_i := \left|MY_{i+1} - MY_i\right| \qquad \Delta R = \begin{pmatrix} 350 \\ 340 \\ 340 \\ 360 \end{pmatrix}$

$\Delta Ri := \max(\Delta R) - \min(\Delta R)$

$\Delta Ri = 20 \qquad 2\cdot\sqrt{SY} = 25.69$

Если ΔRi не превышает значения 2S, то разности можно считать тождественными

9.Определяем коэффициент регрессии для Yr=d0+d1(X-MX)

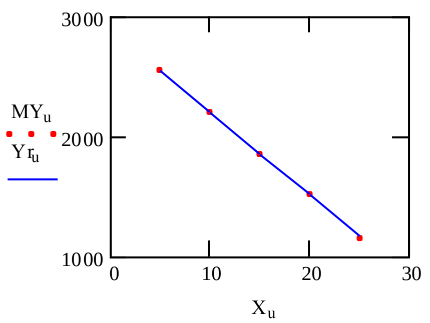
$$d_0 := \frac{1}{m+1}\cdot\sum_{u=0}^{m} MY_u \qquad d_0 = 1.848\times 10^3$$

$$d_1 := \frac{\sum_{u=0}^{m} (X_u - \mathrm{mean}(X))\cdot MY_u}{\sum_{u=0}^{m} (X_u - \mathrm{mean}(X))^2} \qquad d_1 = -69.2$$

10.Построим однофакторную модель Yr=d0+d1(X-MX)

$MX := \mathrm{mean}(X)$

$Yr_u := d_0 + d_1\cdot(X_u - MX)$



11.Определяем адекватность полученного уравнения

$$S1Y := \frac{m+1}{m-1}\cdot\sum_{u=0}^{m} (MY_u - Yr_u)^2 \qquad S1Y = 200$$

$Fr := \frac{S1Y}{SY}$ $\quad$ $Fr = 1.212$

12.Определяем значимость коэффициентов регрессии и их доверительные интервалы

$S2Y := \frac{m \cdot (m+1) \cdot SY + (m-1) S1Y}{(m+1) \cdot (m+1) - 2}$ $\quad$ $S2Y = 169.565$

$S2d_0 := \frac{S2Y}{(m+1) \cdot (m+1)}$ $\quad$ $S2d_0 = 6.783$

$S2d_1 := \frac{S2Y}{(m+1) \cdot \sum_{u=0}^{m} (X_u - MX)^2}$ $\quad$ $S2d_1 = 0.136$

$i := 0..1$

$Sd_i := \sqrt{S2d_i}$ $\quad$ $Sd = \begin{pmatrix} 2.604 \\ 0.368 \end{pmatrix}$

$tRd_i := \frac{d_i}{Sd_i}$ $\quad$ $tRd = \begin{pmatrix} 709.584 \\ -187.885 \end{pmatrix}$

Т.к. tRdi >> tT = 2,07, то коэффициенты значимы и следовательно связь Y и X значима

$\varepsilon d_i := Sd_i \cdot 2.07$ $\quad$ $\varepsilon d = \begin{pmatrix} 5.391 \\ 0.762 \end{pmatrix}$

$d_i - \varepsilon d_i \le \delta_i \le d_i + \varepsilon d_i$

$d - \varepsilon d = \begin{pmatrix} 1.843 \times 10^3 \\ -69.962 \end{pmatrix}$ $\blacksquare \le \delta \le \blacksquare$ $d + \varepsilon d = \begin{pmatrix} 1.853 \times 10^3 \\ -68.438 \end{pmatrix}$

13.Определяем границы доверительных интервалов средних значений выходного параметра

$Sm_u := \sqrt{S2d_0 + S2d_1 \cdot (X_u - MX)^2}$ $\quad$ $Sm = \begin{pmatrix} 4.511 \\ 3.19 \\ 2.604 \\ 3.19 \\ 4.511 \end{pmatrix}$

$\varepsilon m_u := 2.07 \cdot Sm_u$ $\qquad$ $\varepsilon m = \begin{pmatrix} 9.337 \\ 6.603 \\ 5.391 \\ 6.603 \\ 9.337 \end{pmatrix}$

$Ymu_u := Yr_u - \varepsilon m_u$ $\qquad$ $Ym0_u := Yr_u + \varepsilon m_u$

$Ymu = \begin{pmatrix} 2.531 \times 10^3 \\ 2.187 \times 10^3 \\ 1.843 \times 10^3 \\ 1.495 \times 10^3 \\ 1.147 \times 10^3 \end{pmatrix}$ $\qquad$ $Ym0 = \begin{pmatrix} 2.549 \times 10^3 \\ 2.201 \times 10^3 \\ 1.853 \times 10^3 \\ 1.509 \times 10^3 \\ 1.165 \times 10^3 \end{pmatrix}$

14.Определяем границы доверительных интервалов для индивидуальных значений выходного параметра

$Se_u := \sqrt{(Sm_u)^2 + S2Y}$ $\qquad$ $Se = \begin{pmatrix} 13.781 \\ 13.407 \\ 13.28 \\ 13.407 \\ 13.781 \end{pmatrix}$

$Yeu_u := Yr_u - Se_u \cdot 2.07$ $\qquad$ $Ye0_u := Yr_u + Se_u \cdot 2.07$

$Yeu = \begin{pmatrix} 2.511 \times 10^3 \\ 2.166 \times 10^3 \\ 1.821 \times 10^3 \\ 1.474 \times 10^3 \\ 1.127 \times 10^3 \end{pmatrix}$ $\qquad$ $Ye0 = \begin{pmatrix} 2.569 \times 10^3 \\ 2.222 \times 10^3 \\ 1.875 \times 10^3 \\ 1.53 \times 10^3 \\ 1.185 \times 10^3 \end{pmatrix}$

МИНОБРНАУКИ РОССИИ

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

Кафедра дизайна и технологии изделий легкой промышленности

УТВЕРЖДАЮ
Проректор по учебной работе
_______________ О.Г. Локтионова
«1» 02 _______________ 2018 г.

ПОЛУЧЕНИЕ РЕГРЕССИОННОЙ МОДЕЛИ ПО ДАННЫМ ПОЛНОГО ФАКТОРНОГО ЭКСПЕРИМЕНТА ПРИ ИССЛЕДОВАНИИ ТЕХНОЛОГИЧЕСКИХ ПРОЦЕССОВ ЛЕГКОЙ ПРОМЫШЛЕННОСТИ

Методические указания
по выполнению лабораторных работ
для студентов направления подготовки 29.03.05

Курск 2018

УДК 687.02

Составитель: Т.А. Добровольская

Рецензент
Кандидат технических наук, доцент *Т.М. Ноздрачева*

Получение регрессионной модели по данным полного факторного эксперимента при исследовании технологических процессов легкой промышленности: методические указания по выполнению лабораторных работ / Юго-Зап. гос. ун-т; сост.: Т.А. Добровольская. Курск, 2018. 12 с.: табл. 3. - Библиогр.: с. 8.

Излагаются основные сведения об этапах обработки результатов полного факторного эксперимента, получении и анализе многофакторных регрессионных моделей. Указывается порядок выполнения лабораторной работы.

Предназначены для студентов направления подготовки 29.03.05 «Конструирование изделий легкой промышленности» дневной и заочной форм обучения.

Текст печатается в авторской редакции

Подписано в печать . Формат 60×84 1/16.
Усл.печ.л. . Уч.-изд.л. . Тираж 25 экз. Заказ . Бесплатно
Юго-Западный государственный университет.
305040, г. Курск, ул. 50 лет Октября,94.

Лабораторная работа

ПОЛУЧЕНИЕ РЕГРЕССИОННОЙ МОДЕЛИ ПО ДАННЫМ ПОЛНОГО ФАКТОРНОГО ЭКСПЕРИМЕНТА ПРИ ИССЛЕДОВАНИИ ТЕХНОЛОГИЧЕСКИХ ПРОЦЕССОВ ЛЕГКОЙ ПРОМЫШЛЕННОСТИ

Цель работы: изучение основных теоретических положений планирования полного факторного эксперимента, получение многофакторной регрессионной модели исследуемого процесса.

Задание:

1. Изучить основные положения планирования эксперимента.
2. Осуществить выбор факторов, их уровней и интервалов варьирования.
3. Получить уравнение регрессии и проверить его на адекватность.
4. Сделать выводы по полученной модели.

Отчет должен содержать:

1. Краткое обоснование выбора факторов и параметра оптимизации.
2. Матрицу планирования ПФЭ.
3. Математическую обработку данных эксперимента, которая включает
   - 3.1. Определение дисперсии в каждом опыте.
   - 3.2. Проверку воспроизводимости процесса.
   - 3.3. Определение коэффициентов уравнения регрессии.
   - 3.4. Проверку коэффициентов уравнения на значимость.
   - 3.5. Определение адекватности модели.
4. Анализ полученной математической модели.

МЕТОДИКА ВЫПОЛНЕНИЯ РАБОТЫ

В качестве параметра оптимизации (Y) выбрана стойкость трикотажного полотна к истиранию.

В качестве факторов:

X_1 – линейная плотность пряжи, текс

X_2 – плотность полотна по вертикали

Таблица 1

Условия проведения эксперимента

Факторы	Уровни варьирования			Интервал варьирования
	-1	0	+1	
X_1	15,5	17	18,5	1,5
X_2	55	60	65	5

Таблица 2

Матрица планирования ПФЭ 2^2

№ оп	Факторы				Параметр оптимизации				S^2
	Натуральн.		Кодирован.						
	X_1	X_2	X_1	X_2	Y_1	Y_2	Y_3	$Y_{ср}$	
1	18,5	65	+	+	4640	4650	4630	4640	
2	15,5	65	-	+	4080	4070	4120	4090	
3	18,5	55	+	-	4000	3960	3980	3980	
4	15,5	55	-	-	3440	3470	3460	3457	

Вначале определяем средние значения и дисперсии в опытах матрицы, результаты заносим в табл.2.

1. Исключение резко выделяющихся данных. Расчетное значение критерия Смирнова - Грабса определяют по формулам:

при подозрении резко выделяющегося максимального значения $Y_{i\,max}$

$$V_{Rmax}=\frac{(Y_{i\,max}-\overline{Y})}{S\{Y\}}\sqrt{\frac{n}{n-1}}; \quad (1)$$

при подозрении резко выделяющегося минимального значения $Y_{i\,min}$

$$V_{R\,min} = \frac{(\bar{Y} - Y_{i\,min})}{S\{Y\}}\sqrt{\frac{n}{n-1}}. \qquad (2)$$

По таблице критических значений необходимо найти V_T. V_T [p_D = 0,95; m = 3] = 1,412. Если $V_{Rmax} < V_T$ и $V_{Rmin} < V_T$, то рассматриваемые значения не являются резко выделяющимися и остаются для дальнейшей статистической обработки.

2. Проверка гипотезы об однородности дисперсий по критерию Кочрена:

$$G_R = \frac{S_u{}^2{}_{max}\{Y\}}{\sum_{u=1}^{N} S_u^2\{Y\}}. \qquad (3)$$

Табличное значение критерия Кочрена G_T[p_D = 0,95; f = 2; N = 4] = 0,7679. Если $G_R < G_T$, то дисперсии однородны.

3. Определение дисперсии воспроизводимости.

Если доказано, что дисперсии $S^2\{Y\}$ однородны, то дисперсия воспроизводимости равна:

$$S^2\{Y\} = \frac{1}{N}\sum_{u=1}^{N} S_u^2\{Y\}. \qquad (4)$$

4. Определение оценок коэффициентов регрессии.

Общий вид двухфакторной модели: $y = b_0 + b_1 \cdot x_1 + b_2 \cdot x_2 + b_{12} \cdot x_1 \cdot x_2$

Коэффициенты регрессии в этом уравнении определяются по формулам:

$$b_i = \frac{1}{N}\sum_{u=0}^{N} x_{iu}\bar{Y}_u\,; \qquad (5)$$

$$b_{ij} = \frac{1}{N}\sum_{u=1}^{N} x_{iu}x_{ju}\bar{Y}_u\,; \qquad (6)$$

$$b_{ije} = \frac{1}{N}\sum_{u=1}^{N} x_{iu}x_{ju}x_{eu}\bar{Y}_u\,. \qquad (7)$$

В нашем случае:

$$b_0 = \frac{1}{4}\cdot(4640+4090+3980+3457) = 4041{,}75$$

$$b_1 = \frac{1}{4}\cdot(4640-4090+3980-3457) = 268{,}25$$

$$b_2 = \frac{1}{4}\cdot(4640+4090-3980-3457) = 323{,}25$$

$$b_{12} = \frac{1}{4}\cdot(4640-4090-3980+3457) = 6{,}75$$

В результате получаем регрессионную двухфакторную модель:

$$Y_R = 4042 + 268\cdot x_1 + 323\cdot x_2 + 6{,}75\cdot x_1 x_2.$$

Данная модель не является окончательной и после проверки значимости коэффициентов уточняется.

5. Проверка значимости коэффициентов регрессии по критерию Стьюдента:

$$t_R\{b_i\} = \frac{|b_i|}{S\{b_i\}}, \qquad (8)$$

где

$$S^2\{b_i\} = \frac{1}{N} S^2\{\overline{Y}\}, \qquad (9)$$

$$S^2\{\overline{Y}\} = \frac{1}{m} S^2\{Y\}. \qquad (10)$$

Табличное значение критерия Стьюдента $t_T[p_D = 0{,}95;\ f = (4\cdot(3\text{-}1) = 8] = 2{,}306$. Сопоставляя расчетные и табличные значения, делаем вывод, значимы ли коэффициенты b_1 и b_2.

6. Проверка адекватности математической модели по критерию Фишера:

$$F_R = \frac{S^2\text{aä}\{Y\}}{S^2\{Y\}} = \frac{S^2\text{îàä}\ \{Y\}}{S^2\{\overline{Y}\}}, \qquad (11)$$

где

$$S^2\text{àä}\{Y\} = m\cdot S^2\text{îàä}\ \{Y\} = \frac{m\sum\limits_{u=1}^{N}(\overline{Y}_U - Y_{RU})^2}{N - \text{Ì}\ \ - 1}. \qquad (12)$$

Для расчета дисперсии, обусловленной неадекватностью математической модели, составим таблицу.

Таблица 3

u	Y_{Ru}	$\overline{Y}_u$	$\overline{Y}_u - Y_{Ru}$	$\left(\overline{Y}_u - Y_{Ru}\right)^2$
1				
2				
3				
4				
$\sum$				

Значения Y_{Ru} рассчитываются по полученной модели с использование данных табл. 2 (значения x подставляются в модель в кодированных единицах) т.е.

$$Y_{Ru1} = 4035 + 275 + 330; \qquad Y_{Ru2} = 4035 - 275 + 330 \text{ и т.д.}$$

Табличное значение критерия Фишера F_T = [p_D = 0,95; $f\{S^2\{Y\}\}$ = 4·(3-1) = 8; $f\{S^2_{над}\}$ = 4-2-1=1] = 5,32. Если $F_R < F_T$, то модель можно считать адекватной.

В приложении 1 представлена технология получения многофакторной модели по данным ПФЭ с использованием программы Mathcad.

Контрольные вопросы:

1. Сущность многофакторного планирования.
2. Какой эксперимент называется полным факторным?
3. Какой вид имеет многофакторная регрессионная модель, получаемая по данным ПФЭ.
4. Какие основные этапы включает в себя определение многофакторной регрессионной модели на базе ПФЭ?
5. Свойства ПФЭ.
6. Основные этапы обработки результатов ПФЭ.
7. Каким образом определяется значимость коэффициентов регрессии?

8. Что такое адекватность математической модели? Как она определяется?
9. Что необходимо сделать, если гипотеза об адекватности модели отвергается?
10. Что включает в себя анализ полученной математической модели?

Библиографический список

1. Добровольская Т.А. Методы и средства исследования технологических процессов и объектов легкой промышленности [Текст]: учебное пособие / Т.А. Добровольская, Т.И. Леонтьева; Курск.гос. техн. ун-т. Курск, 2006 г. 206 с.
2. Добровольская Т.А. Решение инженерно-технических задач на ПЭВМ с использованием программы «Mathcad»: учеб. пособие/ Курск. гос. техн. ун-т. Курск, 2002. 96 с.
3. Кудрявцев Е.М. Mathcad 8. М.: ДМК, 2000. 320 с.

Приложение 1

Определение линейной многофакторной регрессионной модели по данным ПФЭ

1. Ввод данных эксперимента

$M := 2$ - количество факторов

$m := 3$ - число повторностей

$N := 4$ - число опытов

$i := 0..N-1$

$X_{i,0} :=$ $X_{i,1} :=$ $Y_{i,0} :=$ $Y_{i,1} :=$ $Y_{i,2} :=$

$X_{i,0}$	$X_{i,1}$	$Y_{i,0}$	$Y_{i,1}$	$Y_{i,2}$
1	1	4640	4650	4630
–1	1	4080	4070	4120
1	–1	4000	3960	3980
–1	–1	3440	3470	3460

$$X = \begin{pmatrix} 1 & 1 \\ -1 & 1 \\ 1 & -1 \\ -1 & -1 \end{pmatrix} \qquad Y = \begin{pmatrix} 4.64\times 10^3 & 4.65\times 10^3 & 4.63\times 10^3 \\ 4.08\times 10^3 & 4.07\times 10^3 & 4.12\times 10^3 \\ 4\times 10^3 & 3.96\times 10^3 & 3.98\times 10^3 \\ 3.44\times 10^3 & 3.47\times 10^3 & 3.46\times 10^3 \end{pmatrix}$$

2. Определение средних значений параметра оптимизации

$k := 0..N-1$

$$Ysr_k := \mathrm{mean}\left[\left(Y^T\right)^{\langle k\rangle}\right] \qquad Ysr = \begin{pmatrix} 4.64\times 10^3 \\ 4.09\times 10^3 \\ 3.98\times 10^3 \\ 3.457\times 10^3 \end{pmatrix}$$

3. Определение дисперсий параметра оптимизации

$l := m-1$

$$Sy_i := \frac{\sum_{j=0}^{l} \left(Ysr_i - Y_{i,j}\right)^2}{m-1} \qquad Sy = \begin{pmatrix} 100 \\ 700 \\ 400 \\ 233.333 \end{pmatrix}$$

4. Исключение резко выделяющихся значений

$$VRmax_i := \frac{\max\left[\left(Y^T\right)^{\langle i\rangle}\right] - Ysr_i}{\sqrt{Sy_i}} \cdot \sqrt{\frac{m}{m-1}}$$

$$VRmax = \begin{pmatrix} 1.225 \\ 1.389 \\ 1.225 \\ 1.069 \end{pmatrix}$$

$$VRmin_i := \frac{Ysr_i - \min\left[\left(Y^T\right)^{\langle i\rangle}\right]}{\sqrt{Sy_i}} \cdot \sqrt{\frac{m}{m-1}}$$

$$VRmin = \begin{pmatrix} 1.225 \\ 0.926 \\ 1.225 \\ 1.336 \end{pmatrix}$$

5. Проверка однородности дисперсий

$$Gr := \frac{\max(Sy)}{\sum Sy} \qquad Gr = 0.488$$

6. Определение дисперсии воспроизводимости

$$SY := \frac{\sum Sy}{N} \qquad SY = 358.333$$

7. Расчет коэффициентов регрессии

$$a_0 := \frac{\sum_i Ysr_i}{N} \qquad a_0 = 4.042 \times 10^3$$

$j := 1..M \qquad u := 0..M-1$

$$a_j := \frac{\sum_i \left(X^{\langle j-1\rangle}\right)_i \cdot Ysr_i}{N}$$

$$C := \frac{M \cdot (M-1)}{2} \qquad C = 1$$

$C1 := C + M \qquad C1 = 3 \qquad j := M + 1..C1$

$$a_j := \frac{\sum_i \left(X^{\langle 0\rangle}\right)_i \cdot \left(X^{\langle 1\rangle}\right)_i \cdot Ysr_i}{N}$$

$$a = \begin{pmatrix} 4.042 \times 10^3 \\ 268.333 \\ 323.333 \\ 6.667 \end{pmatrix}$$

8. Определение значимости коэффициентов регрессии

$SSY := \frac{SY}{m}$ $\qquad Sb := \frac{SSY}{N}$

$SSY = 119.444$ $\qquad Sb = 29.861$

$Nk := M + 1 + C$ $\qquad Nk = 4$

$j := 0 .. Nk - 1$

$tr_j := \frac{|a_j|}{\sqrt{Sb}}$ $\qquad tr = \begin{pmatrix} 739.618 \\ 49.105 \\ 59.169 \\ 1.22 \end{pmatrix}$

9. Определение адекватности математической модели

$Yr_i := a_0 + a_1 \cdot \left(X^{\langle 0 \rangle}\right)_i + a_2 \cdot \left(X^{\langle 1 \rangle}\right)_i$

$Yr = \begin{pmatrix} 4.633 \times 10^3 \\ 4.097 \times 10^3 \\ 3.987 \times 10^3 \\ 3.45 \times 10^3 \end{pmatrix}$ $\qquad Ysr = \begin{pmatrix} 4.64 \times 10^3 \\ 4.09 \times 10^3 \\ 3.98 \times 10^3 \\ 3.457 \times 10^3 \end{pmatrix}$

$S\acute{\imath}\grave{a}\ddot{a} := \frac{\left[\sum_i \left(Yr_i - Ysr_i\right)^2\right] \cdot m}{N - M - 1}$ $\qquad S\acute{\imath}\grave{a}\ddot{a} = 533.333$

$Fr := \frac{S\acute{\imath}\grave{a}\ddot{a}}{SSY}$ $\qquad Fr = 4.465$

10. Построение поверхности отклика

$N1 := 20$ $\qquad Step := \frac{1 - (-1)}{N1}$ $\qquad Step = 0.1$

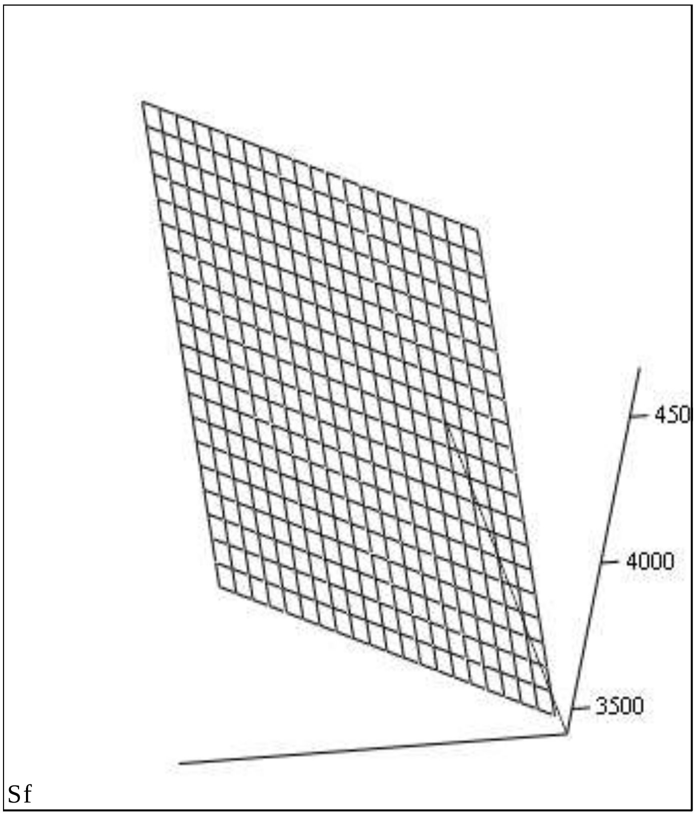
$i := 0, 1 .. N1$ $\qquad j := 0, 1 .. N1$

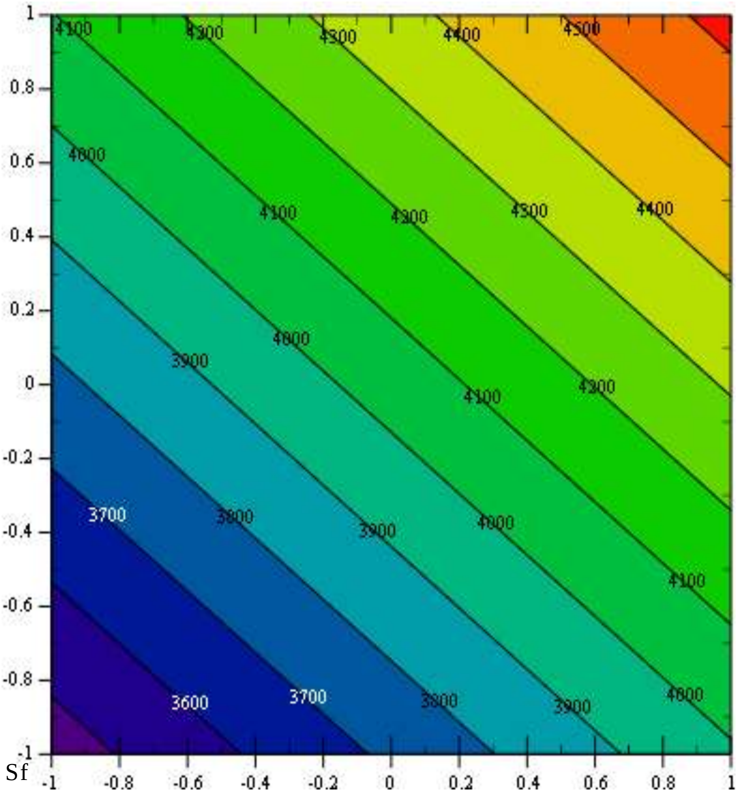
$x1_i := -1 + Step \cdot i$ $\qquad x2_j := -1 + Step \cdot j$

$f(x1, x2) := a_0 + a_1 \cdot x1 + a_2 \cdot x2$

$Sf_{i,j} := f\left(x1_i, x2_j\right)$

$f(0, 0) = 4.042 \times 10^3$


4500
4000
3500
Sf


4100
4200
4300
4400
4500
4000
3900
3800
3700
3600
Sf
-1
-0.8
-0.6
-0.4
-0.2
0
0.2
0.4
0.6
0.8
1