УДК 677.03.001.5
к.т.н. Литовский С.М.

МИНИСТЕРСТВО ОБРАЗОВАНИЯ РЕСПУБЛИКИ БЕЛАРУСЬ

ВИТЕБСКИЙ ГОСУДАРСТВЕННЫЙ ТЕХНОЛОГИЧЕСКИЙ УНИВЕРСИТЕТ

МЕТОДИЧЕСКИЕ УКАЗАНИЯ

к лабораторным работам по курсу

Методы и средства исследований механико-технологических процессов текстильной промышленности

часть 1

ВИТЕБСК
1996

СОДЕРЖАНИЕ

Введение

Современный технический прогресс текстильной промышленности связан с развитием ее техники и технологии. Для успешного управления технологическими процессами и их оптимизации с целью повышения производительности оборудования и качества продукции уже недостаточно знать отдельные качественные стороны процесса.

Для анализа сложных технологических процессов широко применяются методы экспериментального математического моделирования. Использование методов планирования эксперимента позволяет получать математические модели исследуемого процесса в реализованном диапазоне изменения многих факторов, влияющих на процесс, наиболее экономичным и эффективным способом.

Данные методические указания разработаны с целью освоения методов экспериментальных исследований и являются, по сути, кратким обобщением различных методик, изложенных в ряде специализированных изданий по математическому планированию экспериментов. Основное внимание уделено корректной обработке данных активных и пассивных экспериментов.

Определение основных числовых характеристик совокупности случайных величин

Основные сведения

При измерении свойств продуктов текстильных производств и технологических параметров, как правило, получается совокупность случайных величин, которая может быть определена числовыми характеристиками: средним (математическим ожиданием), дисперсией, коэффициентом вариации, квадратической неровнотой. Известно, что числовые характеристики меняются от выборки к выборке и являются также случайными величинами, которые варьируют с заданной доверительной вероятностью в определенном интервале. Чем больше ошибка числовой характеристики, тем шире интервал. Точность каждой числовой характеристики определяется ее ошибкой, а надежность - доверительной вероятностью. Задаваясь точностью и надежностью при известной дисперсии случайной величины, можно определить доверительный объем испытаний для оценки числовой характеристики.

Получение совокупности случайных величин

Перед непосредственной реализацией опытов по анализу случайной величины исследователь должен осуществить ряд организационных и технических мероприятий, от тщательности выполнения которых зависит в большой мере успех эксперимента, а именно:

- проверить свойства сырья и материалов и установить их соответствие задачам исследования;
- проверить состояние оборудования (стендов, приборов и т.д.);
- при необходимости провести тарировку и определить точность показаний измерительной техники;
- при использовании аналоговой характеристики исследуемого параметра (непрерывной регистрации в виде диаграммы, осциллограммы и т.п.) осуществить ее дискретизацию с целью получения совокупности случайных величин;
- проведение одной серии опытов поручать только одному исполнителю, т.к. замена исполнителей может привести к наложению субъективных погрешностей наблюдения.

Для ознакомления с методикой определения основных числовых характеристик совокупности случайных величин необходимо получить данную совокупность. Она может быть получена на разрывной машине (прочность, удлинение), весах (масса отрезков пряжи, полосок ткани или трикотажа), круткомере (крутка пряжи) и других приборах. Можно воспользоваться совокупностями, приведенными в таблице (приложение 6).

Расчет оценок математического ожидания, дисперсии и среднего квадратического отклонения

Математическое ожидание $\overline{Y}$ (среднее значение) определяет центр распределения случайных величин, около которого группируется большая их часть. Абсолютными характеристиками рассеяния случайной величины Y около центра распределения $\overline{Y}$ является дисперсия $S^2\{Y\}$ и среднее квадратическое отклонение $S\{Y\}$.

Расчет оценок математического ожидания, дисперсии и среднего квадратического отклонения для анализируемой выборки осуществляется по следующим формулам:

$$\overline{Y} = \frac{1}{m}\sum_{i=1}^{m} Y_i ;$$

$$S^2\{Y\} = \frac{1}{m-1}\sum_{i=1}^{m}(Y_i - \overline{Y})^2 ;$$

$$S\{Y\} = \sqrt{S^2\{Y\}} .$$

Исключение резко выделяющихся экспериментальных данных

Совокупность полученных экспериментальных данных часто имеет значения, резко выделяющиеся относительно других, что приводит к постановке вопроса об их исключении из дальнейшей обработки. Причиной появления таких данных может быть изменение условий проведения опыта в момент наблюдения, ошибочная регистрация параметра (по вине оператора) и т.п. Независимо от причин получения резко выделяющихся данных они могут существенно исказить числовые характеристики. С другой стороны, при необоснованном исключении таких данных числовые характеристики также будут искажены.

Самый надежный метод определения возможности исключения резко выделяющихся данных - это анализ условий, при которых они были получены. Если условия существенно отличаются от стандартных (или установленных по плану эксперимента), то данные необходимо исключить из дальнейшей обработки независимо от их величины.

Если определение существенности изменения условий эксперимента невозможно или представляет большие трудности, то используют статистический метод исключения данных, сущность которого заключается в следующем:

- находят в совокупности максимальную и минимальную величины и определяют расчетные значения критерия Смирнова-Грабса

$$V_{R\ \max} = \frac{Y_{i\ \max} - \overline{Y}}{S\{Y\}}\sqrt{\frac{m}{m-1}} ;$$

$$V_{R\ \min} = \frac{\overline{Y} - Y_{i\ \min}}{S\{Y\}}\sqrt{\frac{m}{m-1}} ;$$

- сравнивают полученные значения с табличным V_T (приложение 1); если $V_{R\ \max}$ или $V_{R\ \min}$ больше V_T, то соответствующее значение Y_i необходимо исключить из совокупности, а затем повторить расчет оценок $\overline{Y}$, $S^2\{Y\}$ и $S\{Y\}$;
- процедуру повторяют до полного исключения резко выделяющихся значений из совокупности.

Расчет относительных характеристик рассеяния случайной величины

Такой характеристикой является коэффициент вариации $CV\{Y\}$:

$$CV\{Y\} = \frac{S\{Y\}}{\overline{Y}} .$$

Если данная величина выражается в процентах, то она называется квадратической неровнотой $C\{Y\}$:

$$C\{Y\} = \frac{S\{Y\}}{\overline{Y}} 100.$$

Определение ошибки среднего и границ доверительного интервала

В результате измерений исследуемого параметра возникают ошибки (погрешности измерения), для описания которых введены оценки абсолютной ε_i и относительной δ_i погрешности. Абсолютная и относительная доверительные ошибки, допущенные при оценке математического ожидания, определяются по формулам:

$$\varepsilon\{\overline{Y}\} = \frac{2S\{Y\}}{\sqrt{m}};$$

$$\delta\{\overline{Y}\} = \frac{2C\{Y\}}{\sqrt{m}}.$$

Двусторонним доверительным интервалом называется интервал, который покрывает неизвестный параметр распределения с заданной доверительной вероятностью p_D:

$$\overline{Y} - \varepsilon\{Y\} \le \overline{Y} \le \overline{Y} + \varepsilon\{Y\}.$$

В практике текстильных исследований при статистической обработке обычно принимают $p_D = 0{,}95$. Величина, равная $\alpha = 1 - p_D$, называется уровнем значимости и иногда выражается в процентах.

Доверительный объем испытаний

Анализируя точность оценки среднего значения, можно решить, является ли она достаточной или требуется увеличение объема измерений. Задаваясь требуемой величиной относительной ошибки (например, $\delta = 3\%$) и приняв квадратическую неровноту по данным предыдущих опытов или другой априорной информации, можно рассчитать доверительный объем выборки:

$$m\{\overline{Y}\} \ge \left(\frac{u\{p_D\} \cdot C\{Y\}}{\delta\{\overline{Y}\}} \right)^2,$$

где $u\{p_D\}$ - квантиль нормального распределения случайной величины (при $p_D = 0{,}954$ $u\{p_D\} = 2$).

Определение вида дифференциального закона распределения случайной величины

Основные сведения

Наиболее полной характеристикой совокупности случайных величин является дифференциальная или интегральная функции распределения. Для определения вида распределения в исследуемой совокупности используется критерий Пирсона.

Совокупность случайных величин может быть получена на разрывной машине (прочность, удлинение), весах (масса отрезков пряжи, полосок ткани или трикотажа), крутомере (крутка пряжи) и других приборах (можно воспользоваться совокупностями, приведенными в приложении 7).

Формирование частотной таблицы

Полученный ряд экспериментальных значений делят на классы (интервалы). Исходя из количества элементов совокупности m, число классов k определяют по формуле (с округлением до целого):

$$k = 3{,}332 \cdot \lg m + 1 \text{ при } 50 < m < 200;$$

$$k = 4 \cdot \sqrt[5]{0{,}75 \cdot (m-1)^2} \text{ при } m > 200.$$

Например, для $m = 50$ принимаем $k = 7$.

Находим в анализируемой выборке максимальное $Y_{\max}$ и минимальное $Y_{\min}$ значения и определяем величину интервала:

$$\Delta Y = \frac{Y_{\max} - Y_{\min}}{k}.$$

Составляем таблицу (Таблица 1) и разносим все значения анализируемой совокупности по соответствующим классам.

Таблица 1

№ класса	1	2	3	…	k
Границы класса	$Y_{\min}$ - $(Y_{\min} + \Delta Y)$	$(Y_{\min} + \Delta Y)$ - $(Y_{\min} + 2 \cdot \Delta Y)$	$(Y_{\min} + 2 \cdot \Delta Y)$ - $(Y_{\min} + 3 \cdot \Delta Y)$	… …	$(Y_{\min} + (k-1) \cdot \Delta Y)$ - $Y_{\max}$
Значения Y_i					
Частота m_i					
Среднее Y_i^*					

Количество случайных величин в каждом классе m_i называется частотой. После сортировки значений определяем частоту m_i и математическое ожидание (среднее) Y_i^* в каждом классе.

Дальнейшие расчеты сводим в таблицу (Таблица 2).

Таблица 2

i	Границы классов	m_i	Y_i^*	y_i	$m_i \cdot y_i$	y_i^2	$m_i \cdot y_i^2$	m_i^T	$\frac{(m_i - m_i^T)^2}{m_i}$
1.				-3					
2.				-2					
3.				-1					
4.				0					
5.				1					
6.				2					
7.				3					
$\sum$	-		-	-		-		-	

Значение Y_i^* в том классе, где m_i принимает максимальное значение, называется условным нулем выборки Y_0^*.

Значения y_i находятся по формуле (и округляются до ближайшего целого):

$$y_i = \frac{Y_i^* - Y_0^*}{\Delta Y}.$$

Определение оценок математического ожидания, среднего квадратического отклонения и квадратической неровноты

По способу отсчета от условного нуля находим среднее значение выборки:

$$\overline{Y} = Y_0^* + \frac{\Delta Y}{m} \sum_{i=1}^{k} m_i \cdot y_i \ .$$

Находим среднее квадратическое отклонение и квадратическую неровноту:

$$S\{Y\} = \frac{\Delta Y}{\sqrt{m}} \sqrt{\sum_{i=1}^{k} m_i \cdot y_i^2 - \frac{1}{m} \cdot \left(\sum_{i=1}^{k} m_i \cdot y_i \right)^2} \ ;$$

$$C\{Y\} = \frac{S\{Y\}}{\bar{Y}} \cdot 100\% \ .$$

Определение закона распределения исследуемой величины

Задаемся видом предполагаемой дифференциальной или интегральной функции распределения. Как правило, случайные величины, являющиеся предметом анализа при исследовании технологических процессов текстильной промышленности, отвечают нормальному закону распределения:

$$\varphi\{Y\} = \frac{1}{\sqrt{2\pi}} \cdot \exp\left[-\frac{\left(Y_i^* - \bar{Y} \right)^2}{2 \cdot S^2\{Y\}} \right].$$

Вычисляем теоретические частоты m_i^T в каждом классе:

$$m_i^T = \frac{m \cdot \Delta Y}{S\{Y\}} \cdot \varphi\{Y\} \ .$$

Полученные значения заносим в таблицу (Таблица 2) и определяем наблюдаемое значение критерия Пирсона:

$$\chi^2_{набл} = \sum_{i=1}^{k} \frac{\left(m_i - m_i^T \right)^2}{m_i} \ .$$

Из таблицы (приложение 2) определяем критическое значение критерия Пирсона $\chi^2_{крит}$ при условии, что доверительная вероятность $p_D = 0{,}95$ и число степеней свободы $f = k - 2$.

Если $\chi^2_{набл} \leq \chi^2_{крит}$, то анализируемую величину можно считать распределенной по нормальному закону. Если $\chi^2_{набл} \geq \chi^2_{крит}$, то необходимо использовать другие функции (лог-нормальную, экспоненциальную, показательно-степенную и др. [1, стр.30]) до нахождения распределения, адекватного исследуемой случайной величине.

Построение графика функции распределения

Наглядное представление о различиях между экспериментальными значениями и теоретической функцией распределения можно получить путем построения частотного полигона (Рисунок 1).

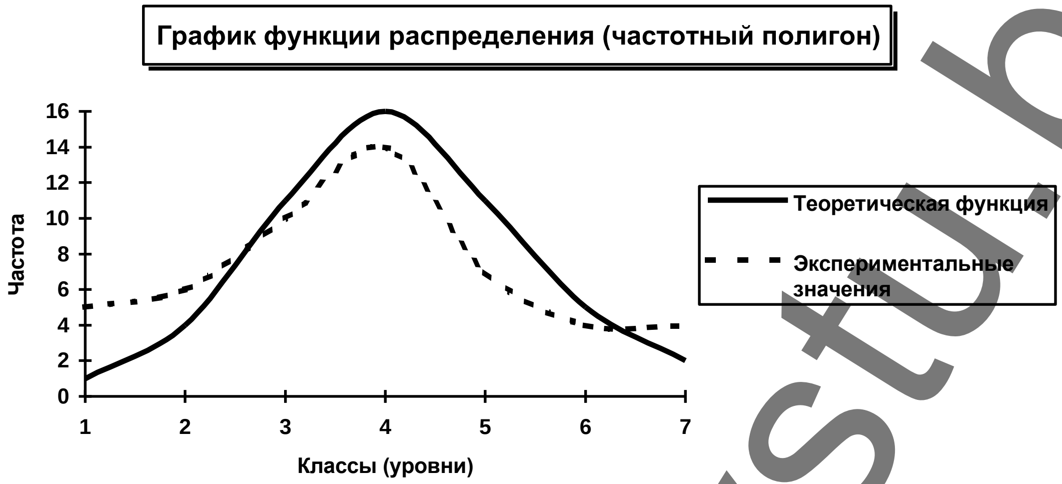

Рисунок 1

Определение статических корреляционных однофакторных моделей по данным пассивного эксперимента

Основные сведения

При исследовании технологических процессов и объектов часто оказывается, что выходной параметр и фактор (входной параметр) оказываются случайными величинами. В результате дискретных измерений фактора X (например, массы 500-миллиметрового отрезка пряжи) и выходного параметра Y (например, разрывной нагрузки вышеупомянутого отрезка) получают две последовательности сопряженных случайных чисел:

$$X_1, X_2, \ldots, X_m;$$
$$Y_1, Y_2, \ldots, Y_m.$$

Каждой паре полученных значений соответствует определенная точка в корреляционном поле точек. Для оценки степени взаимосвязи двух случайных величин X и Y рассчитывают числовую характеристику r_{YX}, называемую коэффициентом корреляции. Для корреляционной взаимосвязи двух случайных величин характерно наличие двух зависимостей: $\overline{Y}(X)$ и $\overline{X}(Y)$, которые в корреляционном поле точек изображаются в виде сопряженных прямых. Причем, чем меньше разброс точек в корреляционном поле, тем сильнее теснота связи между случайными величинами и тем меньше угол φ (Рисунок 2) между сопряженными прямыми.

В практике текстильных исследований корреляционная связь между случайными величинами считается:

- слабой при $0{,}3 \le |r_{YX}| \le 0{,}4$
- средней при $0{,}4 < |r_{YX}| \le 0{,}7$
- сильной при $0{,}7 < |r_{YX}| \le 0{,}9$
- очень сильной при $0{,}9 < |r_{YX}|$

Для ознакомления с методикой коэффициента корреляции и построения однофакторной корреляционной модели необходимо получить две совокупности сопряженных случайных величин (т.е. совокупность пар случайных значений). Они могут быть получены на разрывной машине, весах, круткомере и других приборах (можно воспользоваться совокупностями, приведенными в приложении 6).

Расчет основных статистических характеристик

Определяем средние значения ($\overline{X}$ и $\overline{Y}$) и дисперсии ($S^2\{X\}$ и $S^2\{Y\}$) для совокупностей:

$$\overline{X} = \frac{1}{m}\sum_{i=1}^{m} X_i \ ;$$

$$\overline{Y} = \frac{1}{m}\sum_{i=1}^{m} Y_i \ ;$$

$$S^2\{X\} = \frac{1}{m-1}\sum_{i=1}^{m}(X_i - \overline{X})^2 \ ;$$

$$S^2\{Y\} = \frac{1}{m-1}\sum_{i=1}^{m}(Y_i - \overline{Y})^2 \ .$$

Расчет коэффициента корреляции и определение его значимости

Значение коэффициента корреляции рассчитываем по формуле:

$$r_{YX} = \frac{\sum_{i=1}^{m}(X_i - \overline{X})\cdot(Y_i - \overline{Y})}{(m-1)\cdot S\{X\}\cdot S\{Y\}} \ .$$

По значению коэффициента делаем вывод о тесноте корреляционной взаимосвязи между X и Y.

Для определения значимости коэффициента корреляции определяем расчетное значение критерия Стьюдента:

$$t_R\{r_{YX}\} = \frac{r_{YX}\sqrt{m-2}}{\sqrt{1-r_{YX}^2}}$$

Теоретическое значение критерия t_T определяем по таблице (приложение 3) при условии, что $p_D = 0{,}95$ и $f = m-2$. Если $t_R\{r_{YX}\} > t_T$, то гипотеза о наличии корреляционной взаимосвязи между X и Y не отвергается.

Определение линейной модели корреляционной взаимосвязи

Рассчитываем значения коэффициентов линейных уравнений сопряженных прямых:

$$d_{0X} = \overline{Y} \ ; \ \ d_{1X} = \frac{r_{YX}\cdot S\{Y\}}{S\{X\}};$$

$$d_{0Y} = \overline{X} \ ; \ \ d_{1Y} = \frac{r_{YX}\cdot S\{X\}}{S\{Y\}}.$$

Подставляем полученные значения в соответствующие уравнения:

$$Y = d_{0X} + d_{1X}(X - \overline{X})\,;$$

$$X = d_{0Y} + d_{1Y}(Y - \overline{Y})\,.$$

Раскрываем скобки и получаем уравнения прямых. Строим оси координат, наносим корреляционное поле точек, а затем строим сопряженные прямые (показываем угол φ между ними).

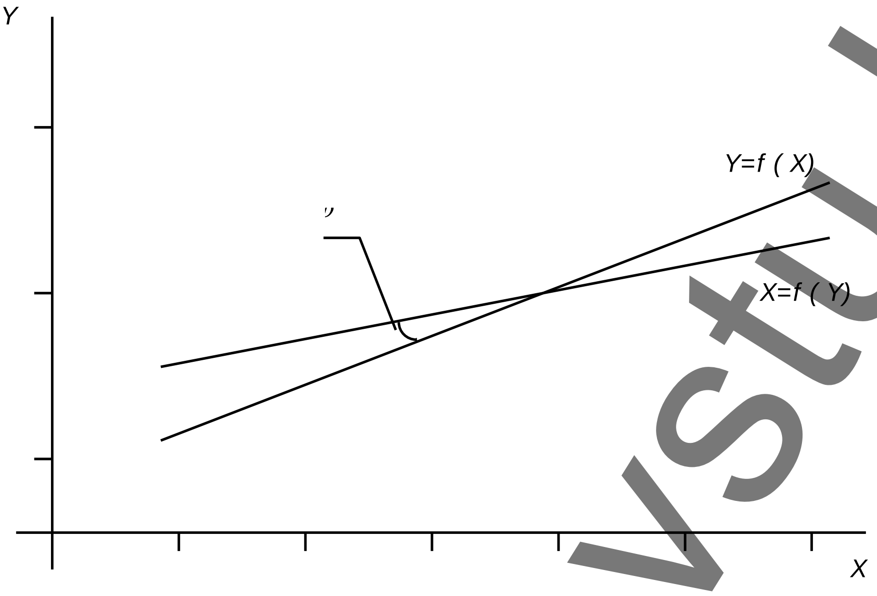

Рисунок 2

Определение статических корреляционных многофакторных моделей по данным пассивного эксперимента

Основные сведения

В том случае, если требуется проанализировать зависимость одной случайной величины (Y) от нескольких случайных величин $X_1, X_2, \ldots, X_M$, необходимо определить корреляционную многофакторную модель [1, стр.197]:

$$Y = a_0 + a_1 X_1 + a_2 X_2 + \ldots + a_M X_M .$$

Методику рассмотрим на примере разработки двухфакторной корреляционной модели.

В результате дискретных измерений факторов X_1, X_2 и выходного параметра Y получают совокупность сопряженных случайных чисел (можно воспользоваться совокупностями, приведенными в приложении 6).

Расчет основных статистических характеристик

Определяем средние значения и дисперсии:

$$\overline{X}_1 = \frac{1}{m}\sum_{i=1}^{m} X_{1i} \; ;$$

$$\overline{X}_2 = \frac{1}{m}\sum_{i=1}^{m} X_{2i} \; ;$$

$$\overline{Y} = \frac{1}{m}\sum_{i=1}^{m} Y_i \; ;$$

$$S^2\{X_1\} = \frac{1}{m-1}\sum_{i=1}^{m} (X_{1i} - \overline{X}_1)^2 \; ;$$

$$S^2\{X_2\} = \frac{1}{m-1}\sum_{i=1}^{m}(X_{2i} - \overline{X}_2)^2 \;;$$

$$S^2\{Y\} = \frac{1}{m-1}\sum_{i=1}^{m}(Y_i - \overline{Y})^2 \;.$$

Расчет парных коэффициентов корреляции

Значения парных коэффициентов корреляции отражают тесноту взаимосвязи двух параметров и определяются для каждых двух переменных:

$$r_{YX_1} = \frac{\sum_{i=1}^{m}(X_{1i} - \overline{X}_1)\cdot(Y_i - \overline{Y})}{(m-1)\cdot S\{X_1\}\cdot S\{Y\}};$$

$$r_{YX_2} = \frac{\sum_{i=1}^{m}(X_{2i} - \overline{X}_2)\cdot(Y_i - \overline{Y})}{(m-1)\cdot S\{X_2\}\cdot S\{Y\}};$$

$$r_{X_1X_2} = \frac{\sum_{i=1}^{m}(X_{1i} - \overline{X}_1)\cdot(X_{2i} - \overline{X}_2)}{(m-1)\cdot S\{X_1\}\cdot S\{X_2\}}.$$

Расчет множественного коэффициента корреляции и определение его значимости

Теснота линейной связи между случайными величинами X_1, X_2 и Y определяется полным (совокупным) или множественным коэффициентом корреляции. Этот коэффициент определяет силу совместного влияния всех факторов на выходной параметр и для двухфакторной модели имеет вид:

$$R_{YX_1X_2} = \sqrt{\frac{r_{YX_1}^2 + r_{YX_2}^2 - 2r_{YX_1}r_{YX_2}r_{X_1X_2}}{1 - r_{X_1X_2}^2}}\;.$$

Используя критерий Стьюдента, определяем значимость найденного коэффициента:

$$t_R\{R_{YX_1X_2}\} = \frac{R_{YX_1X_2}}{S\{R_{YX_1X_2}\}}\;.$$

Среднее квадратическое отклонение определяется по формуле:

$$S\{R_{YX_1X_2}\} = \frac{(1 - R_{YX_1X_2}^2)}{\sqrt{m - M - 1}}\,,$$

где $M = 2$ (число факторов) и $m = 10$ (количество испытаний).

Теоретическое значение критерия Стьюдента t_T определяется из таблицы (приложение 3) при условии, что $p_D = 0{,}95$ и $f = m - M - 2$. Если $t_R\{R_{YX_1X_2}\} > t_T$, то множественный коэффициент корреляции значим.

Определение линейной модели корреляционной взаимосвязи

Искомая модель имеет вид:

$$Y = a_0 + a_1X_1 + a_2X_2\,,$$

или в стандартизированной форме:

$$t = q_1t_1 + q_2t_2\,,\text{ где}$$

a_0, a_1, a_2 - коэффициенты с натуральными значениями факторов;

q_1, q_2 - стандартизированные значения коэффициентов. Они определяются по следующим формулам:

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

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

Натуральные значения определяются по формулам:

$$a_1 = q_1 \frac{S\{Y\}}{S\{X_1\}};$$

$$a_2 = q_2 \frac{S\{Y\}}{S\{X_2\}};$$

$$a_0 = \overline{Y} - \sum_{i=1}^{M} a_i \cdot \overline{X}_i = \overline{Y} - a_1\overline{X}_1 - a_2\overline{X}_2$$

Подставляем значения полученных коэффициентов в уравнение и получаем корреляционную модель в натуральных значениях.

Определение регрессионной однофакторной модели (модели первого порядка) по данным активного эксперимента

Основные сведения

При решении многих инженерных задач возникает необходимость в установлении связи между независимыми переменными X_1, X_2, … X_k и зависящей от них величиной Y. В случае если Y является случайной величиной, а X_1, X_2, … X_k - величины неслучайные, для разработки искомой математической модели вида $Y = f(X_1, X_2, \ldots, X_k)$ используется метод регрессионного анализа (метод наименьших квадратов).

Применение регрессионного анализа правомерно при выполнении следующих условий:

Значения выходного параметра Y в каждом опыте матрицы планирования эксперимента представляют собой независимые, нормально распределенные случайные величины.

Дисперсии выходного параметра в различных опытах матрицы однородны.

Значения уровней факторов не являются линейной комбинацией от уровней остальных факторов.

Точность определения значений выходного параметра значительно ниже точности определения величины уровня фактора.

Если одно из приведенных выше условий не будут выполняться, эффективность анализа значительно снижается и по найденной модели могут быть получены неверные технологические выводы.

В данной работе метод регрессионного анализа рассмотрен на примере построения линейной однофакторной модели (модели первого порядка).

Условия проведения активного эксперимента

Цель данной работы - получить модель вида:

$$Y_R = a_0 + a_1 X. \qquad (1)$$

Для получения данной модели проводят активный эксперимент в широком диапазоне изменения фактора X. Обычно применяют число уровней фактора, т.е. число опытов в матрице планирования

N≥5. Для повышения точности определения выходного параметра Y каждый опыт матрицы повторяется несколько раз (m≥2).

По результатам эксперимента заполняем таблицу (Таблица 3). Можно воспользоваться значениями, приведенными в приложении 8.

Таблица 3

X_u	u	Y_{ui}					$\sum_{i=1}^{m} Y_{ui}$	$\overline{Y_u}$	$S_u^2\{Y\}$
		i=1	i=2	i=3	i=4	i=5			
	1								
	2								
	3								
	4								
	5								

Итого:

Нахождение статистических характеристик

Находят средние значения функции отклика по строкам $\overline{Y}_u$ и построчные дисперсии $S_u^2\{Y_u\}$ по формулам:

$$\overline{Y}_u = \frac{1}{m}\sum_{i=1}^{m} Y_{u_i}, \quad S_u^2\{Y_u\} = \frac{1}{m-1}\sum_{i=1}^{m}(Y_{u_i} - \overline{Y}_u)^2, \text{ где}$$

m - число повторностей опыта.

Проверка гипотезы об однородности дисперсий

Если число повторных опытов m одинаково для всех опытов матрицы, то для проверки однородности дисперсий применяется критерий Кочрена, расчетное значение которого определяется по формуле:

$$G_R = \frac{S_{u\ \max}^2\{Y\}}{\sum_{u=1}^{N} S_u^2\{Y\}}$$

Расчетное значение G_R сравнивают с табличным значением G_T, которое определяют по таблице (приложение 5) в зависимости от числа опытов в матрице N и числа степеней свободы дисперсии $f\{S_u^2\} = m - 1$ для заданной доверительной вероятности.

Если $G_R < G_T$, то гипотеза об однородности дисперсий принимается, если нет - следует применить методику исключения резко выделяющихся величин или найти причину возникновения большой дисперсии в u-м опыте, а затем повторить (полностью или частично) экспериментальную часть работы.

Если число повторных опытов m различно для разных опытов матрицы, то для проверки гипотезы об однородности дисперсий в опытах матрицы применяется критерий Бартлета [1, стр.55].

Вычисление дисперсии воспроизводимости выходного параметра в опытах матрицы

Если в опытах матрицы дисперсии однородны и число повторных опытов одинаково, то средняя дисперсия определяется по формуле:

$$S_{\text{восп}}^2\{Y\} = \frac{1}{N}\sum_{u=1}^{N} S_u^2\{Y\}.$$

Вычисление коэффициентов искомого уравнения (модели) и их дисперсий

Для получения искомого уравнения (1) предварительно находят коэффициенты уравнения:

$$Y_R = d_0 + d_1(X - \overline{X}), \tag{2}$$

где

$$\overline{X} = \frac{1}{N}\sum_{u=1}^{N} X_u;$$

$$d_0 = \frac{1}{N}\sum_{u=1}^{N} \overline{Y}_u = \overline{Y};$$

$$d_1 = \frac{\sum_{u=1}^{N}(X_u - \overline{X})\overline{Y}_u}{\sum_{u=1}^{N}(X_u - \overline{X})^2}.$$

Расчеты необходимых сумм сводим в таблицу (Таблица 4).

Таблица 4

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

Преобразуем уравнение (2) в уравнение (1).

Проверка адекватности полученной модели

Вначале определяется дисперсия неадекватности:

$$S^2_{н\,ад} = \frac{m\sum_{u=1}^{N}(\overline{Y}_u - Y_{Ru})^2}{N-2}, \text{ где}$$

Y_{Ru} - возвращаемые моделью расчетные значения выходного параметра, которые определяют для каждого опыта путем подстановки в полученное уравнение соответствующих значений входных параметров.

Расчеты необходимых сумм сводим в таблицу (Таблица 5).

Таблица 5

u	X_u	d_1X_u	Y_{Ru}	$\overline{Y}_u$	$\overline{Y}_u - Y_{Ru}$	$(\overline{Y}_u - Y_{Ru})^2$
1						
2						
3						
4						
5						
$\sum_{u=1}^{N}$						

Определяют расчетное значение критерия Фишера:

$$F_R = \frac{S^2_{\text{í àä}}\{Y\}}{S^2_{\text{âî ñï}}\{\overline{Y}\}}, \text{ если } S^2_{\text{í àä}}\{Y\} > S^2_{\text{âî ñï}}\{\overline{Y}\};$$

$$\text{или } F_R = \frac{S^2_{\text{âî ñï}}\{\overline{Y}\}}{S^2_{\text{í àä}}\{Y\}}, \text{ если } S^2_{\text{âî ñï}}\{\overline{Y}\} > S^2_{\text{í àä}}\{Y\}.$$

Расчетное F_R значение критерия сравнивают с табличным F_T, которое определяют по таблице (приложение 4) при условии, что P_D=0,95, $f\{S^2_{\text{âî ñï}}\} = N(m-1)$, $f\{S^2_{\text{í àä}}\} = N-2$.

Если $F_R < F_T$, то с вероятностью P_D гипотеза об адекватности полученной модели принимается.

Если гипотеза об адекватности отвергается, необходимо переходить к описанию процесса моделью более высокого порядка на базе другого вида эксперимента.

Оценка значимости полученных коэффициентов регрессии

Значимость полученных коэффициентов оценивается с помощью критерия Стьюдента, расчетное значение которого (для каждого коэффициента) определяется по формуле:

$$t_R\{d_i\} = \frac{|d_i|}{\sqrt{S^2\{d_i\}}},$$

где

$$S^2\{d_0\} = \frac{S^2\{Y\}}{mN};$$

$$S^2\{d_1\} = \frac{S^2\{Y\}}{m\sum_{u=1}^{N}(X_u - \overline{X})^2};$$

$$S^2\{Y\} = \frac{(m-1)NS^2_{\text{âî ñï}}\{Y\} + (N-2)S^2_{\text{í àä}}\{Y\}}{mN-2}$$

Полученное расчетное значение t_R сравнивается с табличным t_T, которое определяют по таблице (приложение 3) при условии, что P_D=0,95 и число степеней свободы $f\{S^2\} = Nm-2$.

Если $t_R\{d_i\} > t_T$ выполняется для обоих коэффициентов разработанной модели, то линейная связь между X и Y значима.

Определение регрессионных многофакторных математических моделей по данным активного эксперимента

Основные сведения

В настоящее время в научных исследованиях широкое применение получили математико-статистические методы планирования экспериментов, в которых математический аппарат играет активную роль, диктуя исследователю определенную схему постановки эксперимента и последовательность анализа результатов.

В задачу планирования эксперимента входит:

- выбор необходимых для эксперимента опытов, т.е. построение матрицы планирования;
- выбор методов математической обработки результатов эксперимента.

Матрица планирования эксперимента представляет собой таблицу, в которой указаны значения уровней факторов в различных сериях опытов. Матрицы планирования должны удовлетворять ряду требований:

- *ортогональность* - независимость получаемых коэффициентов регрессии и возможность исключения членов модели с незначимыми коэффициентами без последующего пересчета значимых коэффициентов;
- *ротатабельность* - постоянство дисперсии выходного параметра на равных расстояниях от центра эксперимента;
- *униформность* - постоянство дисперсии выходного параметра в некоторой области вокруг центра эксперимента.

Эксперимент, реализующий все возможные неповторяющиеся комбинации уровней исследуемых факторов, называется полным факторным экспериментом (ПФЭ). Он применяется для получения регрессионной многофакторной модели (РМФМ) при исследовании локального участка факторного пространства, не соответствующего его экстремальной части. РМФМ, получаемая по результатам ПФЭ, имеет вид линейного полинома

$$Y_R = b_0 + b_1x_1 + \ldots + b_ix_i + \ldots + b_Mx_M \quad (1)$$

или неполного полинома второго порядка

$$Y_R = b_0 + b_1x_1 + \ldots + b_ix_i + \ldots + b_Mx_M + b_{12}x_1x_2 + \ldots + b_{ij}x_ix_j + \ldots + b_{M-1,M}x_{M-1}x_M, \quad (2)$$

где Y_R - расчетное значение выходного параметра; x_i - кодированные значения уровней факторов; b_i, b_{ij} - значения коэффициентов регрессии; $i = 1,\ldots, M$, $j = 1,\ldots, M$ - номер фактора.

При факторном планировании в отличие от традиционного (однофакторного) по величине коэффициентов регрессии b_i, b_{ij} в РМФМ можно судить о влиянии на выходной параметр не только каждого фактора x_i, но и их взаимодействия x_ix_j, т.е. изменения влияния одного фактора при переходе второго фактора на другой уровень.

Разработка матрицы планирования

Для составления матрицы планирования необходимо определить требуемое количество опытов:

$$N = k^M,$$

где *k* - число уровней варьирования каждого фактора, изменяя которое можно уменьшать или увеличивать *N*.

Необходимо учесть, что для вычисления коэффициентов регрессии искомого уравнения (2) должно соблюдаться условие $N \geq N_k$ (N_k - число коэффициентов регрессии в РМФМ), а для оценки адекватности полученной модели это условие усиливается, т.е. $N > N_k$.

В матрице планирования используются кодированные значения уровней фактора:

(-) - нижний уровень фактора (равен -1);

(+) - верхний уровень фактора (равен +1);

Например, для двухуровневого трехфакторного эксперимента (2^3) матрица ПФЭ содержит восемь опытов (Таблица 6). Для изучения описываемой методики можно воспользоваться значениями, приведенными в приложении 9.

Таблица 6

u	Факторы				Сочетания				Y_{ui}		$\overline{Y}_u$	$S_u^2\{Y\}$
	x_0	x_1	x_2	x_3	$x_1\ x_2$	$x_1\ x_3$	$x_2\ x_3$	$x_1\ x_2\ x_3$	Y_{u1}	Y_{u2}		
1	+	-	-	-								
2	+	+	-	-								
3	+	-	+	-								
4	+	+	+	-								
5	+	-	-	+								
6	+	+	-	+								
7	+	-	+	+								
8	+	+	+	+								
k_i									-	-	$\sum$	$\sum$

Зная квадратическую неровноту выходного параметра *C{Y}* по данным предварительного эксперимента или на основании другой априорной информации, а также задаваясь величиной относительной ошибки $\delta\{\overline{Y}\}$ и доверительной вероятностью $u(P_D)$, можно рассчитать требуемое число повторностей каждого опыта:

$$m(\overline{Y}) = \left(\frac{u\{P_D\}C\{Y\}}{\delta\{\overline{Y}\}}\right)^2$$

Обработку результатов ПФЭ проводят в следующем порядке.

Нахождение статистических характеристик

Находят средние значения функции отклика по строкам $\overline{Y}_u$ и построчные дисперсии $S_u^2\{Y_u\}$ по формулам:

$$\overline{Y}_u = \frac{1}{m}\sum_{i=1}^{m} Y_{u_i}, \quad S_u^2\{Y_u\} = \frac{1}{m-1}\sum_{i=1}^{m}(Y_{u_i} - \overline{Y}_u)^2, \text{ где}$$

m - число повторностей опыта.

Проверка гипотезы об однородности дисперсий

Если число повторных опытов *m* одинаково для всех опытов матрицы, то для проверки однородности дисперсий применяется критерий Кочрена, расчетное значение которого определяется по формуле:

$$G_R = \frac{S_{u\ \max}^2\{Y\}}{\sum_{u=1}^{N} S_u^2\{Y\}}$$

Расчетное значение G_R сравнивают с табличным значением G_T, которое определяют (приложение 5) в зависимости от числа опытов в матрице *N* и числа степеней свободы дисперсии $f\{S_u^2\} = m - 1$ для заданной доверительной вероятности.

Если $G_R < G_T$, то гипотеза об однородности дисперсий принимается, если нет - следует применить методику исключения резко выделяющихся величин или найти причину возникновения большой дисперсии в *u*-м опыте, а затем повторить (полностью или частично) экспериментальную часть работы.

Если число повторных опытов *m* различно для разных опытов матрицы, то для проверки гипотезы об однородности дисперсий в опытах матрицы применяется критерий Бартлета [1, стр.55].

Вычисление дисперсии воспроизводимости выходного параметра в опытах матрицы

Если в опытах матрицы дисперсии однородны и число повторных опытов одинаково, то средняя дисперсия определяется по формуле:

$$S^2_{\text{âî ñï}}\{Y\} = \frac{1}{N}\sum_{u=1}^{N} S_u^2\{Y\}.$$

Вычисление коэффициентов искомого уравнения (модели)

Коэффициенты регрессии определяются по следующим формулам:

$$b_i = \frac{1}{N}\sum_{u=1}^{N} x_{iu}\overline{Y}_u (i = 0,1,\ldots, M);$$

$$b_{ij} = \frac{1}{N}\sum_{u=1}^{N} x_{iu}x_{ju}\overline{Y}_u (i \neq j);$$

$$b_{ijl} = \frac{1}{N}\sum_{u=1}^{N} x_{iu}x_{ju}x_{lu}\overline{Y}_u (i \neq j \neq l).$$

В результате подстановки найденных коэффициентов в уравнение (2) получается регрессионную многофакторную модель, которая, однако, не является окончательной моделью изучаемого процесса.

Оценка значимости полученных коэффициентов регрессии

Значимость полученных коэффициентов оценивается с помощью критерия Стьюдента, расчетное значение которого (для каждого коэффициента) определяется по формуле:

$$t_R\{b_i\} = \frac{|b_i|}{\sqrt{S^2\{b_i\}}}$$

где

$$S^2\{b_i\} = \frac{1}{N} S^2\{\overline{Y}\},$$

в свою очередь

$$S^2\{\overline{Y}\} = \frac{1}{m} S^2_{\text{âî ñï}}\{Y\}.$$

Полученное расчетное значение t_R сравнивается с табличным t_T, которое определяют по таблице (приложение 3) при условии, что P_D=0,95 и число степеней свободы $f\{S_u^2\} = N(m-1)$.

Если $t_R\{b_i\} > t_T$, то коэффициент b_i значим. Если $t_R\{b_i\} < t_T$, то коэффициент b_i незначим, и его необходимо приравнять нулю, т.е. исключить член b_iX_i из модели.

Необходимо учитывать, что значимость коэффициентов зависит не только от удельного влияния данного фактора на выходной параметр, но и от интервала варьирования уровней фактора. Незначимость может быть обусловлена малым интервалом варьирования фактора, большой дисперсией воспроизводимости вследствие наличия неуправляемых и неконтролируемых факторов, а также расположением основного уровня фактора близко к точке частного экстремума выходного параметра по этому фактору. После исключения незначимых коэффициентов записывается искомая модель.

Проверка адекватности полученной модели

Проверку адекватности модели можно проводить только при условии, что число проведенных опытов больше числа коэффициентов модели.

Вначале определяется дисперсия неадекватности:

$$S^2_{\text{í àä}} = \frac{\sum\limits_{u=1}^{N} (\overline{Y}_u - Y_{Ru})^2}{N - N_{\text{çí . êî ýô.}}}\text{, где}$$

$N_{зн.коэф.}$ - число значимых (оставшихся) коэффициентов в модели.

Y_{Ru} - возвращаемые моделью расчетные значения выходного параметра, которые определяют для каждого опыта путем подстановки в полученное уравнение соответствующих значений входных параметров.

Определяют расчетное значение критерия Фишера:

$$F_R = \frac{S^2_{\text{í àä}}\{Y\}}{S^2\{\overline{Y}\}}\text{, если } S^2_{\text{í àä}}\{Y\} > S^2\{\overline{Y}\};$$

$$\text{или } F_R = \frac{S^2\{\overline{Y}\}}{S^2_{\text{í àä}}\{Y\}}\text{, если } S^2\{\overline{Y}\} > S^2_{\text{í àä}}\{Y\}.$$

Расчетное F_R значение критерия сравнивают с табличным F_T, которое определяют по таблице (приложение 4) при условии, что P_D=0,95, $f\{S_u^2\} = N(m-1)$, $f\{S^2_{\text{í àä}}\} = N - N_{\text{çí . êî ýô.}}$.

Если $F_R < F_T$, то с вероятностью P_D гипотеза об адекватности полученной модели принимается.

Если гипотеза об адекватности отвергается, необходимо переходить к описанию процесса полиномом второго порядка на базе другого вида эксперимента или, если это возможно, проводить эксперимент с меньшим интервалом варьирования уровней факторов. Однако неоправданное уменьшение интервала варьирования может обусловить статистическую незначимость коэффициентов регрессии.

Исследование полученной регрессионной многофакторной модели

Получив математическую модель, исследователь проводит ее анализ. Вклад фактора в величину выходного параметра при переходе от нижнего к верхнему уровню называется *эффектом фактора*. Чем больше коэффициент регрессии, тем выше эффект этого фактора, т.е. тем сильнее влияние фактора на выходной параметр. Таким образом, по величине коэффициентов регрессии в модели можно осуществить ранжирование факторов по силе их влияния на Y.

Наиболее наглядным является графическое построение (например, с помощью “STATGRAPHICS”) поверхности отклика для двухфакторной регрессионной модели путем изображения линий одинакового уровня выходного параметра (изолиний). Каждая линия представляет собой проекцию сечения поверхности отклика плоскостью, параллельной плоскости чертежа.

Анализируя вид полученной поверхности легко определить влияние каждого фактора на выходной параметр.

Для трехфакторной модели строят три семейства изолиний для двух факторов, используя три стабилизации третьего фактора (на нижнем, основном и верхнем уровне).

При $M > 3$ наглядное представление о геометрическом образе поверхности отклика становится невозможным из-за отсутствия у человека интуиции в многомерных пространствах.

Определение регрессионной многофакторной модели второго порядка по *D*-оптимальным матрицам

Основные сведения

В последнее время появились матрицы, которые удовлетворяют требованиям оптимальности оценок коэффициентов модели и выходного параметра при уменьшенном числе опытов. Матрицу, которая обеспечивает получение минимума обобщенной дисперсии, т.е. минимума дисперсии всех коэффициентов регрессии ($S^2\{b\}\to\min$) называют *D*-оптимальной. Одним из достоинств данных матриц является то, что факторы варьируются только на трех уровнях.

Выбор матрицы планирования

Описываемая методика позволяет получить модель вида:

$$Y_R = b_0 + \sum_{i=1}^{M} b_i x_i + \sum_{\substack{i=j=1 \\ j \# i}}^{M} b_{ij} x_i x_j + \sum_{i=1}^{M} b_{ii} x_i^2$$

Ниже приведены некоторые наиболее известные матрицы, имеющие хорошие статистические характеристики и включающие небольшое число опытов. При этом используются следующие обозначения:

- M - число факторов (входных параметров);
- N - общее число опытов в матрице;
- $N_ц$ - число опытов в центре эксперимента;
- В условном обозначении строк в матрице используются следующие сокращения:
- a, b, c, d, e - факторы (соответственно X_1, X_2, X_3, X_4, X_5) на верхнем уровне;
- a(0),b(0),c(0),d(0),e(0) - факторы (соответственно X_1, X_2, X_3, X_4, X_5) на основном уровне;
- (1) - все факторы в данной строке на нижнем уровне.

Матрица Коно (Ko_2):

M	N	$N_ц$	Условное обозначение строк в матрице
2	9	1	ab,b,a,(1),ab(0),b(0),a(0)b,a(0),a(0)b(0)

Матрица Бокса (B_3):

M	N	$N_ц$	Условное обозначение строк в матрице
3	14	0	abc,bc,ac,c,ab,b,a,(1),ab(0)c(0),b(0)c(0),a(0)bc(0),a(0)c(0),a(0)b(0)c,a(0)b(0)

Матрица Бокса (B_4):

M	N	$N_ц$	Условное обозначение строк в матрице
4	24	0	abcd,bcd,acd,cd,abd,bd,ad,d,abc,bc,ac,c,ab,b,a,(1),ab(0)c(0)d(0),b(0)c(0)d(0),a(0)bc(0)d(0),a(0)c(0)d(0),a(0)b(0)cd(0),a(0)b(0)d(0),a(0)b(0)c(0)d,a(0)b(0)c(0)

Матрица Бокса (B_5):

M	N	$N_ц$	Условное обозначение строк в матрице
5	42	0	abcde,bcde,acde,cde,abde,bde,ade,de,abce,bce,ace,ce,abe,be,ace,abcd,bcd,acd,cd,abd,bd,ad,d,abc,bc,ac,c,ab,b,a,(1),ab(0)c(0)d(0)e(0),b(0)c(0)d(0)e(0),a(0)bc(0)d(0)e(0),a(0)c(0)d(0)e(0),a(0)b(0)cd(0)e(0),a(0)b(0)d(0)e(0),a(0)b(0)c(0)de(0),a(0)b(0)c(0)e(0),a(0)b(0)c(0)d(0)e,a(0)b(0)c(0)d(0)

Матрица Хартли (Ha_5):

M	N	$N_ц$	Условное обозначение строк в матрице
5	27	1	abcde,bcd,acd,cde,abd,bde,ade,d,abc,bce,ace,c,abe,b,a,e,ab(0)c(0)d(0)e(0),b(0)c(0)d(0)e(0),a(0)bc(0)d(0)e(0),a(0)c(0)d(0)e(0),a(0)c(0)d(0)e(0),a(0)b(0)cd(0)e(0),a(0)b(0)d(0)e(0),a(0)b(0)c(0)de(0),a(0)b(0)c(0)e(0),a(0)b(0)c(0)d(0)e,a(0)b(0)c(0)d(0),a(0)d(0)c(0)d(0)e(0)

При проведения эксперимента по одной из вышеперечисленных матриц необходимо прибегать к рандомизации опытов.

Для изучения описываемой методики можно воспользоваться значениями, приведенными в приложении 10.

Зная квадратическую неровноту выходного параметра *C{Y}* по данным предварительного эксперимента или на основании другой априорной информации, а также задаваясь величиной относительной ошибки $\delta\{\overline{Y}\}$ и доверительной вероятностью $u(P_D)$, можно рассчитать требуемое число повторностей каждого опыта:

$$m(\overline{Y}) = \left(\frac{u\{P_D\}C\{Y\}}{\delta\{\overline{Y}\}}\right)^2$$

Обработку полученных данных проводят в следующем порядке.

Нахождение статистических характеристик

Находят средние значения функции отклика по строкам $\overline{Y}_u$ и построчные дисперсии $S_u^2\{Y_u\}$ по формулам:

$$\overline{Y}_u = \frac{1}{m}\sum_{i=1}^{m} Y_{u_i}, \quad S_u^2\{Y_u\} = \frac{1}{m-1}\sum_{i=1}^{m}(Y_{u_i} - \overline{Y}_u)^2, \text{ где}$$

m - число повторностей опыта.

Проверка гипотезы об однородности дисперсий

<u>Если число повторных опытов *m* одинаково</u> для всех опытов матрицы, то для проверки однородности дисперсий применяется критерий Кочрена, расчетное значение которого определяется по формуле:

$$G_R = \frac{S_{u\,\max}^2\{Y\}}{\sum_{u=1}^{N} S_u^2\{Y\}}$$

Расчетное значение G_R сравнивают с табличным значением G_T, которое определяют (приложение 5) в зависимости от числа опытов в матрице *N* и числа степеней свободы дисперсии $f\{S_u^2\} = m - 1$ для заданной доверительной вероятности.

Если $G_R < G_T$, то гипотеза об однородности дисперсий принимается, если нет - следует применить методику исключения резко выделяющихся величин или найти причину возникновения большой дисперсии в *u*-м опыте, а затем повторить (полностью или частично) экспериментальную часть работы.

<u>Если число повторных опытов *m* различно</u> для разных опытов матрицы, то для проверки гипотезы об однородности дисперсий в опытах матрицы применяется критерий Бартлета [1, стр.55].

Вычисление дисперсии воспроизводимости

Вычисляют дисперсию воспроизводимости по формуле:

$$S^2\{\overline{Y}\} = \frac{\sum_{u=1}^{N} S_u^2\{Y\}}{m(N - N_{\ddot{o}} + 1)}$$

Вычисление коэффициентов искомого уравнения (модели) и их дисперсий

$$b_0 = g_1\sum_{u=1}^{N}\overline{Y}_u - g_2\sum_{i=1}^{M}\sum_{u=1}^{N} x_{iu}^2\overline{Y}_u;$$

$$b_i = g_3 \sum_{u=1}^{N} x_{iu} \overline{Y}_u ;$$

$$b_{ij} = g_4 \sum_{u=1}^{N} x_{iu} x_{ju} \overline{Y}_u ;$$

$$b_{ii} = g_5 \sum_{u=1}^{N} x_{iu}^2 \overline{Y}_u + g_6 \sum_{i=1}^{M} \sum_{u=1}^{N} x_{iu}^2 \overline{Y}_u - g_2 \sum_{u=1}^{N} \overline{Y}_u ;$$

$$S^2\{b_0\} = g_1 S^2\{\overline{Y}\} ;$$

$$S^2\{b_i\} = g_3 S^2\{\overline{Y}\} ;$$

$$S^2\{b_{ij}\} = g_4 S^2\{\overline{Y}\} ;$$

$$S^2\{b_{ii}\} = g_7 S^2\{\overline{Y}\} ;$$

Значения постоянных коэффициентов g_i приведены в таблице (Таблица 7)

Таблица 7

Матрица	M	g_1	g_2	g_3	g_4	g_5		g_1
$Ко_2$	2	0,55556	0,33333	0,16666	0,25000	0,5	0,00000	0,50000
$В_3$	3	0,40625	0,15625	0,10000	0,12500	0,5	-0,09375	0,40625
$В_4$	4	0,22917	0,06250	0,05556	0,06250	0,5	-0,10417	0,39583
$В_5$	5	0,15821	0,03320	0,02941	0,03125	0,5	-0,09180	0,40820
$На_5$	5	0,13804	0,03030	0,55560	0,06250	0,5	-0,09091	0,40909

Оценка значимости полученных коэффициентов регрессии

Значимость полученных коэффициентов оценивается с помощью критерия Стьюдента, расчетное значение которого (для каждого коэффициента) определяется по формуле:

$$t_R\{b_i\} = \frac{|b_i|}{\sqrt{S^2\{b_i\}}}$$

Полученное расчетное значение t_R сравнивается с табличным t_T, которое определяют по таблице (приложение 3) при условии, что P_D=0,95 и число степеней свободы $f\{S_u^2\} = N(m-1)$.

Если $t_R\{b_i\} > t_T$, то коэффициент b_i значим. Если $t_R\{b_i\} < t_T$, то коэффициент b_i незначим, и его необходимо приравнять нулю, т.е. исключить член b_iX_i из модели.

Необходимо учитывать, что значимость коэффициентов зависит не только от удельного влияния данного фактора на выходной параметр, но и от интервала варьирования уровней фактора. Незначимость может быть обусловлена малым интервалом варьирования фактора, большой дисперсией воспроизводимости вследствие наличия неуправляемых и неконтролируемых факторов, а также расположением основного уровня фактора близко к точке частного экстремума выходного параметра по этому фактору.

Следует отметить, что исключение членов модели с коэффициентами b_{ii} в случае их незначимости без пересчета значимых коэффициентов b_{jj} и b_0 является некорректным приемом, хотя его часто применяют.

При числе факторов более трех (M > 3) с целью повышения адекватности модели рекомендуется проводить последовательное исключение членов модели с незначимыми коэффициентами b_{ii} (начиная с минимального) с пересчетом оставшихся коэффициентов [1, стр.178].

Проверка адекватности полученной модели

Проверку адекватности модели можно проводить только при условии, что число проведенных опытов больше числа коэффициентов модели.

Вначале определяется дисперсия неадекватности:

$$S^2_{íàä} = \frac{\sum_{u=1}^{N} (\overline{Y}_u - Y_{Ru})^2}{N - N_{çí.êîýô.} - (N_ö - 1)^2}\text{, где}$$

$N_{зн.коэф.}$ - число значимых (оставшихся) коэффициентов в модели.

Y_{Ru} - возвращаемые моделью расчетные значения выходного параметра, которые определяют для каждого опыта путем подстановки в полученное уравнение соответствующих значений входных параметров.

Определяют расчетное значение критерия Фишера:

$$F_R = \frac{S^2_{íàä}\{Y\}}{S^2\{\overline{Y}\}}\text{, если } S^2_{íàä}\{Y\} > S^2\{\overline{Y}\};$$

$$\text{или } F_R = \frac{S^2\{\overline{Y}\}}{S^2_{íàä}\{Y\}}\text{, если } S^2\{\overline{Y}\} > S^2_{íàä}\{Y\}.$$

Расчетное F_R значение критерия сравнивают с табличным F_T, которое определяют по таблице (приложение 4) при условии, что P_D=0,95, $f\{S^2_u\} = N(m-1)$, $f\{S^2_{íàä}\} = N - N_{çí.êîýô.}$.

Если $F_R < F_T$, то с вероятностью P_D гипотеза об адекватности полученной модели принимается.

Если гипотеза об адекватности отвергается, необходимо переходить к описанию процесса полиномом второго порядка на базе другого вида эксперимента или, если это возможно, проводить эксперимент с меньшим интервалом варьирования уровней факторов. Однако неоправданное уменьшение интервала варьирования может обусловить статистическую незначимость коэффициентов регрессии.

Исследование полученной регрессионной многофакторной модели

Используя полученную модель необходимо определить характер изменения поверхности отклика в экстремальном участке и определить комбинацию уровней факторов, обеспечивающих экстремальное значение выходного параметра, т.е. оптимальные условия исследуемого процесса.

Наиболее наглядным является графическое построение (например, с помощью «STATISTICA FOR WINDOWS») поверхности отклика для двухфакторной регрессионной модели путем изображения линий одинакового уровня выходного параметра (изолиний). Каждая линия представляет собой проекцию сечения поверхности отклика плоскостью, параллельной плоскости чертежа.

Анализируя вид полученной поверхности легко определить влияние каждого фактора на выходной параметр.

Для трехфакторной модели строят три семейства изолиний для двух факторов, используя три стабилизации третьего фактора (на нижнем, основном и верхнем уровне).

При $M > 3$ наглядное представление о геометрическом образе поверхности отклика становится невозможным из-за отсутствия у человека интуиции в многомерных пространствах.

Для нахождения координат точки оптимума (оптимальных значений факторов) можно воспользоваться одним из множества методов, реализованных на ЭВМ (например, «STATISTICA FOR WINDOWS»).

Литература

Севостьянов А.Г. Методы и средства исследований механико-технологических процессов текстильной промышленности. М., «Легкая индустрия», 1980.

Спиридонов А.А. Планирование эксперимента при исследовании технологических процессов. М., «Машиностроение», 1981.

Тойберг П. Оценка точности результатов измерений. Пер. с немецк. М., «Энергоатомиздат», 1988.

Севостьянов А.Г., Кудинов А.В., Литвинов М.С. и др. Методы и средства исследований механико-технологических процессов текстильной промышленности. Лабораторный практикум. М., «Легкая промышленность и бытовое обслуживание», 1986.

Айвазян С.А., Енюков И.С., Мешалкин Л.Д. Прикладная статистика: Основы моделирования и первичная обработка данных. М., «Финансы и статистика», 1983.

Айвазян С.А., Енюков И.С., Мешалкин Л.Д. Прикладная статистика: Исследование зависимостей. М., «Финансы и статистика», 1985.

Кокс Д., Снелл Э. Прикладная статистика: Принципы и примеры. М., «Мир», 1984.

Ким Дж.-О., Мьюллер Ч.У., Клекка У.Р., Олдендерфер М.С., Блэшфилд Р.К. Факторный, дискриминантный и кластерный анализ: Пер. С англ. Под ред. И.С.Енюкова. М., «Финансы и статистика», 1989.

Григорьев С.Г., Перфилов А.М., Левандовский В.В., Юнкеров В.И. STATGRAPHICS на персональном компьютере. Санкт-Петербург., «ПО-3», 1992.