

АКТИВНЫЙ И ПАССИВНЫЙ ЭКСПЕРИМЕНТ

Получение регрессионных уравнений может быть распространена с использованием множественной линейной регрессии и активного эксперимента на основе планов экспериментов второго порядка.

Обычно процессы химической технологии являются сложными, поэтому на процесс влияет ряд факторов.

Возможны два подхода к исследованию многофакторных систем.

Первый можно сформулировать следующим образом: «Изменяй факторы по одному». Исследование системы разбивается на серии, в пределах каждой из которых изменяется (варьируется) лишь один фактор, а остальные неизменны. В следующей серии изменяется второй фактор и т. д.

Идея другого подхода - построение плана эксперимента, предусматривающего изменение всех влияющих факторов, чтобы этот он обеспечивал максимум точности.

Второй подход значительно эффективнее первого. При одном объеме эксперимента и одинаковой точности опытов получается большая точность результатов.

Слайд 2

4.1. Проверка воспроизводимости эксперимента

При планировании эксперимента целесообразно, если это не связано с дорогостоящими испытаниями, убедиться в том, что опыты воспроизводимы. Для этого проводят N_p серий из k параллельных опытов в рассматриваемой области изменения влияющих факторов. Результаты опытов заносят в таблицу

Таблица 4.1- Проверка воспроизводимости опытов

Номер серии опытов	Результаты параллельных опытов				$Y_{i,j}$, средн.	s_j^2
	Y_{11}	Y_{12}	...	Y_{1k}		
1	Y_{11}	Y_{12}	...	Y_{1k}	$Y_{1, \text{средн.}}$	s_1^2
2	Y_{21}	Y_{22}	...	Y_{2k}	$Y_{2, \text{средн.}}$	s_2^2
3	Y_{31}	Y_{32}	...	Y_{3k}	$Y_{3, \text{средн.}}$	s_3^2
			...			
j	Y_{j1}	Y_{j2}	...	Y_{jk}	$Y_{j, \text{средн.}}$	s_j^2
			...			
N_p	Y_{N_p1}	Y_{N_p2}	...	Y_{N_pk}	$Y_{N_p, \text{средн.}}$	$S_{N_p}^2$

Слайд 3

Процедура проверки воспроизводимости состоит из следующих шагов

1) Вычисляется среднее арифметическое значение функции отклика для каждой серии опытов j ($j = 1 \dots N_p$) по формуле

$$Y_{j, \text{средн.}} = \frac{1}{k} \sum_{i=1}^k y_{ji} \cdot$$

2) Оценивается дисперсия для каждой серии параллельных опытов по формуле

$$s_j^2 = \frac{1}{k-1} \sum_{i=1}^k (Y_{ji} - Y_{j,\text{средн.}})^2 .$$

3) Определяется расчётное значение критерия Кохрена по формуле

$$G_p = \max s_j^2 / \sum_{j=1}^{N_p} s_j^2 .$$

По специальным таблицам (см. Приложение Б) определяют табличное значение критерия Кохрена – $G_{\text{табл.}}$. Оно зависит от доверительной вероятности P (как правило, $P=0.95$) или от значения $p = 1 - P$, которое называется уровнем значимости, от N_p и $k - 1$.

Условие воспроизводимости опытов $G_p < G_{\text{табл.}}$; при этом считается, что оценки дисперсий однородны.

4) Определяется погрешность эксперимента. Вычисляется оценка дисперсии воспроизводимости (погрешность экспериментальных данных) по формуле

$$s_Y^2 = \frac{1}{N_p} \sum_j^{N_p} s_j^2 .$$

с ней связано число степеней свободы $N_p (k - 1)$

Слайд 4

Пример 4.1. Оценка воспроизводимости опытов

Номер серии опытов	Результаты параллельных опытов				$Y_{j,\text{средн.}}$	s_j^2
1	14.8	14.7	14.4	14.5	14.6	0.10
2	45.5	45.7	45.1	45.1	45.3	0.28
3	12.6	12.8	12.2	12.0	12.4	0.40
4	50.4	50.6	49.6	50.0	50.2	0.60
						$\sum_j s_j^2 = 1,38$

По данным таблицы вычисляется

$$G_p = \max s_j^2 / \sum_{j=1}^{N_p} s_j^2 = 0.6/1.38; G_p = 0.43.$$

Оценка дисперсии воспроизводимости (характеризует погрешность экспериментальных данных)

$$S_Y^2 = \frac{1}{N_p} \sum_j^{N_p} s_j^2; S_Y^2 = \frac{1}{4} \sum_j^4 s_j^2; S_Y^2 = 0.35;$$

Здесь $j = 1, \dots, 4$; $G_{\text{табл.}} = 0.684$ для $N_p = 4, k - 1 = 3$.

Фрагмент таблицы значений критерия Кохрена для уровня значимости 0.05:

N_p	Число степеней свободы $f = k - 1$		
	1	2	3
2	0.999	0.975	0.939
3	0.967	0.871	0.798
4	0.907	0.768	0.684
5	0.841	0.684	0.598

Так как $G_p < G_{\text{табл.}}$, то опыты считаются воспроизводимыми, а оценки дисперсий – однородными.

Следует отметить, что существующие программные продукты ориентированы на построение уравнений регрессии, как при наличии, так и при отсутствии параллельных опытов. Построение уравнений регрессии в этих случаях имеет ряд особенностей, которые рассмотрим далее.

Слайд 6

Полный факторный эксперимент

План полного факторного эксперимента (ПФЭ) предусматривает условия всех опытов, которые следует провести. План эксперимента задается в виде матрицы планирования - прямоугольной таблицы, каждая строка которой содержит данные определенного опыта, а каждый столбец - значения независимых переменных в разных опытах (они называются факторами) и соответствующие значения зависимой переменной, называемой целевой функцией, или функцией отклика.

При использовании полного факторного эксперимента математическое описание исследуемого процесса получается в виде регрессионных уравнений: части ряда Тейлора для функций отклика. При этом ограничиваются линейной частью разложения функций отклика и членами, содержащими произведения факторов в первой степени.

В этом случае удается находить уравнения локального участка поверхности функции отклика, определяемого интервалами варьирования факторов для исследуемого процесса

На основании полного факторного эксперимента математическая модель

процесса строится в виде уравнений регрессии

$$y_{расч,k} = \beta_0 + \beta_1 x_{1,k} + \dots + \beta_n x_{n,k} + \beta_{12} x_{1,k} x_{2,k} + \dots + \beta_{(n-1)n} x_{(n-1),k} x_{n,k}, k = 1, \dots, N$$

где n - количество факторов; N - число экспериментальных точек. $k = 1, \dots, N$.

Полный факторный эксперимент обеспечивает возможность получить математическую модель процесса в некоторой локальной области, содержащейся в окрестности точки с координатами $(x_{0,1}, x_{0,2}, \dots, x_{0,n})$. Эта точка называется центром плана.

Для удобства обработки экспериментальных данных вводят кодированные переменные по формуле

$$X_i = \frac{x_i - x_{0,i}}{\Delta x_i}$$

где x_i - экспериментальные значения факторов, Δx_i - интервал варьирования соответствующего фактора, $i = 1, 2, \dots, n$; $x_{0,i}$ - значение фактора на базовом уровне.

Выбор интервалов варьирования и базовых уровней факторов является сложной задачей. Интервалы варьирования факторов определяются в результате предварительных экспериментальных исследований. Общие рекомендации по выбору интервалов варьирования факторов следующие: интервал варьирования должен быть достаточно велик по сравнению с ошибкой эксперимента.

На практике интервал варьирования факторов принимают в размере 15-35% от значения фактора на нулевом (базовом) уровне; в качестве базового уровня фактора часто выбирают его значение в отработанном режиме функционирования химико-технологического процесса.

Уравнения регрессии в кодированных переменных приобретают вид

$$y_{расч,k} = b_0 + b_1 X_{1,k} + \dots + b_n X_{n,k} + b_{12} X_{1,k} X_{2,k} + \dots + b_{(n-1)n} X_{(n-1),k} X_{n,k}, k = 1, \dots, N$$

Классическая задача регрессионного анализа состоит в вычислении коэффициентов уравнений регрессии $b_0, b_1, b_2, \dots, b_n, b_{12}, \dots, b_{(n-1)n}$, определяющих взаимосвязь зависимой переменной $y_{расч,k}$ и множества независимых переменных X_{ik} в рассматриваемой точке k .

Эти уравнения являются линейными относительно коэффициентов регрессии.

Слайд 7

На рисунке 4.1 показаны 4 точки, для которых следует проводить опыты при полном факторном эксперименте, для двух факторов, каждый из которых варьируется на двух уровнях - верхнем(+1) и нижнем(-1).

Слайд 8

Матрица планирования ПФЭ

Для удобства вычисления коэффициентов уравнений регрессии все факторы в ходе полного факторного эксперимента варьируют на двух уровнях: нижнем (-1) и верхнем (+1).

В таблице 4.2 приведены условия проведения опытов полного трехфакторного эксперимента 2³. Эти опыты соответствуют в факторном пространстве вершинам куба.

Общее число опытов при ПФЭ $N = 2^n$, где 2-число уровней варьирования значения каждого фактора, n - число факторов.

Таблица 4.2 -Матрица планирования ПФЭ $N = 2^3$

Номер опыта	Факторы			Функция отклика
	X ₁	X ₂	X ₃	
1	-1	-1	-1	y ₁
2	+1	-1	-1	y ₂
3	-1	+1	-1	y ₃
4	+1	+1	-1	y ₄
5	-1	-1	+1	y ₅
6	+1	-1	+1	y ₆
7	-1	+1	+1	y ₇
8	+1	+1	+1	y ₈

Пример 4.2. В восьми опытах исследуется влияние трёх факторов (температуры T , К, давления p , МПа и времени t , с) на выход продукта.

Здесь в любом из опытов температура либо 1000 К (нижний уровень), либо 1200 К (верхний уровень); аналогично варьируются p и t . Выбор центра плана и интервалов варьирования представлены в таблице:

Слайд 9

	Температура, К	Давление, МПа	Время, с
Основной уровень	1000	750	50
Интервал варьирования	100	250	10
Верхний уровень	1200	1000	60
Нижний уровень	1100	500	40

Матрицу планирования практического эксперимента в физических переменных для этого случая можно представить в виде:

Слайд 10

Номер опыта	T, К	P, МПа	t, с
1	1000	500	40
2	1200	500	40
3	1000	1000	40
4	1200	1000	40
5	1000	500	60
6	1200	500	60
7	1000	1000	60
8	1200	1000	60

Из этой матрицы видны основные принципы построения матриц планирования полного факторного эксперимента:

- частота смены уровней варьирования первого фактора чередуется от опыта к опыту,
- для следующих факторов частота смены уровней варьирования в два раза меньше, чем у предыдущего фактора

Слайд 11

Для обработки результатов эксперимента и определения коэффициентов регрессии по простым формулам, которые будут приведены ниже, используется матрица в кодированных переменных, которая для трёх факторов может быть представлена в одном из следующих видов

№ опыта	X ₁	X ₂	X ₃	y
1	-1	-1	-1	y ₁
2	1	-1	-1	y ₂
3	-1	1	-1	y ₃
4	1	1	-1	y ₄
5	-1	-1	1	y ₅
6	1	-1	1	y ₆
7	-1	1	1	y ₇
8	1	1	1	y ₈

Слайд 12

Матрица планирования ПФЭ в кодированных переменных обладает, кроме указанных выше, следующими свойствами

$$\sum_{j=1}^N X_{ji} = 0; \quad \sum_{j=1}^N X_{ji}^2 = N; \quad \sum_{j=1}^N X_{ji} X_{ji} = 0.$$

где j - номер опыта; i, j - номера факторов, причем $i \neq j$.

Последняя формула выражает свойство ортогональности матрицы