Министерство образования и науки Российской Федерации
Федеральное государственное бюджетное образовательное учреждение
высшего профессионального образования
**«Владимирский государственный университет имени
Александра Григорьевича и Николая Григорьевича Столетовых»
(ВлГУ)**

Институт Машиностроения и Автомобильного транспорта

Кафедра Автотранспортная и техносферная безопасность

Курс лекций по дисциплине

«АНАЛИТИЧЕСКИЕ ЧИСЛЕННЫЕ МЕТОДЫ В ПЛАНИРОВАНИИ ЭКСПЕРИМНЕТОВ И ИНЖЕНЕРНОМ АНАЛИЗЕ»

Направление подготовки 23.04.01 «Технология транспортных процессов»

Программа подготовки: «Организация автомобильных перевозок и безопасность движения»

Уровень высшего образования : магистратура
Форма обучения : очная

Составитель
Ф.П. Касаткин

Владимир 2016 г.

Планирование эксперимента при выборе типа автомобилей
для перевозки сыпучих грузов

В качестве критерия выбора берем производительность автомобиля – это планируемый или выполненный объем перевозок в единицу времени (более часто используют т/ч). Задача – выбрать из имеющихся в наличии автомобиль, обеспечивающий наибольшую производительность.

Величину производительности можно определить по формуле:

$$W_Q = \frac{q_н \cdot \gamma_{ст}}{\frac{l_n}{V_т \cdot \beta} + t_{пр}} = \frac{q_н \cdot \gamma_{ст} \cdot V_т \cdot \beta}{l_{ге} + t_{пр} \cdot V_т \cdot \beta}$$

где Q_e - масса груза перевезенного за ездку, т; q_i - номинальная грузоподъемность автомобиля, т; $\gamma_{ст}$ - коэффициент статического использования грузоподъемности; $V_т$ техническая скорость автомобиля, км/ ч., β - коэффициент использования пробега; $l_{ге}$ - расстояние перевозки. км., -время выполнения погрузо-разгрузочных работ.

Величину производительности можно также определить, разделив суточный объем перевозок на время работы автомобиля в течение суток .

Из приведенных рассуждений видно, что на величину производительности влияют следующие параметры: При планировании эксперимента принимаем, β = 0,5 - автомобиль-самосвал в одну сторону едет с грузом – в другую –без груза; $\gamma_{ст}$ = 1 – загрузка автомобиль соответствует его грузоподъемности . Остальные переменные выбираем исходя из технических характеристик имеющихся на предприятии автомобилей. Для разных расстояний перевозки проводим свой эксперимент. Следовательно, проводим планирование полного трехфакторного эксперимента.

Проводим дисперсионный анализ результатов эксперимента.

1). Проверка измерений на равноточность. Равноточность измерений или однородность полученного ряда оценок дисперсии S_{y1}^2; $S_{y2}^2 \ldots S_{yN}^2$, определяется по критерию Кохрена.

$$S_{yn}^2 = \frac{\sum(\bar{y}_{ni} - y_{ni})^2}{\gamma - 1}; \bar{y}_{ni} = \frac{\sum y_{ni}}{\gamma}; G^Э = \frac{\max\{S_{yn}^2\}}{\sum S_{yn}^2}$$

$G^Э$ сравнивается с табличным значением $G_{f_1 f_2}^T$, если при выбранном уровне значимости α имеет место неравенство $G^Э \le G_{f_1 f_2}^T$, то измерения во всех γ опытах равноточны, если же $G^Э > G_{f_1 f_2}^T$, то измерения при выбранном уровне значимости не равноточны и дальнейшие расчеты продолжать не имеет смысла. Тогда необходимо опыт, оценка дисперсии которого наибольшая, повторить более точно и скорректировать исходные данные, либо увеличить число параллельных измерений γ в каждом опыте.

Число степеней свободы критерия Кохрена определяется по формулам: f_1 число степеней свободы числителя $f_1 = \gamma - 1$, f_2 число степеней свободы знаменателя $f_2 = N$.

2). Определение ошибки эксперимента. Ошибка вычисляется по формуле:

$$S_{\bar{y}} = \sqrt{\frac{\sum S_{yn}^2}{N \cdot \gamma}}$$

3). Определение коэффициентов уравнения регрессии. Коэффициенты уравнения регрессии определяются по следующим формулам: $b_i = \frac{\sum a_i \bar{y}_n}{N}; b_{ij} = \frac{\sum a_{ij} \bar{y}_n}{N}; b_{ijk} = \frac{\sum a_{ijk} \bar{y}_n}{N}$ где a_i и $a_{ij} = \pm 1$ берутся из матриц планирования эксперимента.

4). Дисперсионный анализ уравнения регрессии:

а) определение ошибки при вычислении коэффициентов уравнения регрессии.

Ошибки при вычислении коэффициентов для ПФЭ типа 2^n одинаковы и определяются по формуле: $S^2 = \{b_n\} = S^2_{\bar{y}} \frac{1}{N}; или \;\; S = \{b_n\} = S_{\bar{y}} \frac{1}{\sqrt{N}};$

б) проверка значимости коэффициентов уравнения регрессии. Проверку значимости коэффициентов уравнения регрессии проводят по t критерию Стьюдента. Для каждого коэффициента определяют $t^Э$ критерий по формуле: $t_n^Э = \frac{|b_n|\sqrt{N}}{S_{\bar{y}}}$ и сравнивают с табличным значением $t^T_{\alpha,f}$ выбранного для заданного уровня значимости и числа степеней свободы, которое равно: $f = N(\gamma - 1)$.

Если для данного коэффициента выполняется неравенство $t^Э \geq t^T_{\alpha,f}$, то при выбранном уровне значимости соответствующий коэффициент уравнения регрессии значим, если неравенство не выполняется, то коэффициент не значим и членом уравнения регрессии с этим коэффициентом можно пренебречь.

в) проверка гипотезы об адекватности полученного уравнения действительному процессу можно рассчитать двумя способами используя F критерий Фишера:

по формуле вычисляют экспериментальное значение статистики:

$$F^Э = \frac{S^2_{yag}}{S^2_{\bar{y}}}; S^2_{yag} = \frac{\sum \bar{y}_n^2 - N\sum b_n^2}{N - (k+1)}; F^T_{f_1, f_2}$$

где k число значимых коэффициентов регрессии, не считая свободного члена, и сравнивают с табличным значением, где f_1 число степеней свободы числителя отношения, $f_1 = N - (k+1)$; f_2 число степеней свободы знаменателя отношения, $f_2 = N(\gamma - 1)$, γ - число параллельных измерений.

Если при выбранном уровне значимости выполняется неравенство $F^Э \leq F^T_{f_1, f_2}$ то полученное уравнение регрессии адекватно действительному уравнению, если неравенство не выполняется, то полученное уравнение регрессии неадекватно описывает полученный процесс;

опытное значение критерия Фишера может быть взятым равным отношению дисперсии неадекватности к общей дисперсии или дисперсии воспроизводимости:

$$F_{опыт} = \frac{S^2 Y_{н.а.}}{S^2 Y_{воспр.}}; \;\; S^2 Y_{н.а.} = \frac{\gamma \cdot S^2_{ОСТ}}{N - k} = \frac{\gamma \cdot \sum_{i=1}^{N} (Y_{i\,рас} - \bar{Y}_{i\,опыт.})^2}{N - k};$$

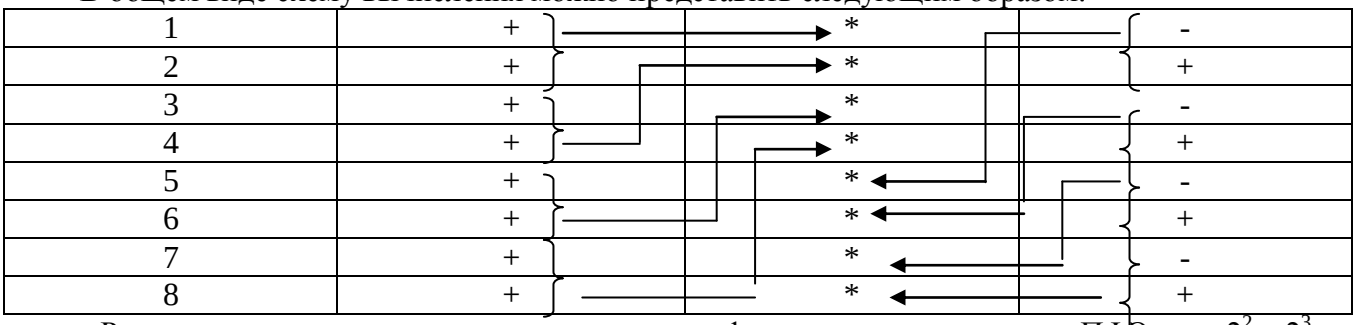
$$S^2 Y_{воспр.} = \frac{\sum_{i=1}^{N} \sum_{j=1}^{k} (Y_{ij\,рас.} - Y_{i\,опыт.})^2}{N(k-1)}$$

г) определение границ доверительных интервалов для коэффициентов регрессии. Границы доверительных интервалов, в которых с заданной вероятностью $P = 1 - \alpha$ находятся коэффициенты уравнения регрессии $\{\beta\}$ определяются по формуле:

$b_n - t^T_{\alpha,f} S\{b_n\} \le \beta_n \ge b_n + t^T_{\alpha,f} S\{b_n\}$ где $t^T_{\alpha,f}$ табличное значение t - критерия Стьюдента при заданном уровне значимости и числе степеней свободы f □$N(\tilde{\gamma}$ ı).

Вычислять коэффициенты регрессии можно по методу, предложенному Йетсом. Процедура состоит в попарном сложении и вычитании компонент вектора-выхода. Число циклов равно числу факторов, участвующих в эксперименте.

В общем виде схему вычисления можно представить следующим образом:



Рассмотрим правила вычисления для полного факторного эксперимента ПФЭ типа 2^2 и 2^3.

№ опыта	Вектор выхода	Цикл 1	Цикл 2	Коэффициенты b
1	$\bar{y}_1$	$\bar{y}_1 + \bar{y}_2 = \bar{y}_1'$	$\bar{y}_1' + y_2' = y_1''$	$b_0 = \frac{1}{2^2} y_1''$
2	$\bar{y}_2$	$\bar{y}_3 + \bar{y}_4 = \bar{y}_2'$	$\bar{y}_3' + y_4' = y_2''$	$b_1 = \frac{1}{2^2} y_2''$
3	$\bar{y}_3$	$\bar{y}_2 - \bar{y}_1 = \bar{y}_3'$	$\bar{y}_2' - y_1' = y_3''$	$b_2 = \frac{1}{2^2} y_3''$
4	$\bar{y}_4$	$\bar{y}_4 - \bar{y}_3 = \bar{y}_4'$	$\bar{y}_4' - y_3' = y_4''$	$b_{12} = \frac{1}{2^2} y_4''$

Схема вычисления коэффициентов регрессии для ПФЭ типа 2^3 методом Йетса.

№ опыта	Вектор выхода	Цикл 1	Цикл 2	Цикл 3	Коэффициенты b
1	$\bar{y}_1$	$\bar{y}_1 + \bar{y}_2 = \bar{y}_1'$	$\bar{y}_1' + y_2' = y_1''$	$\bar{y}_1'' + y_2'' = y_1'''$	$b_0 = \frac{1}{2^3} y_1'''$
2	$\bar{y}_2$	$\bar{y}_3 + \bar{y}_4 = \bar{y}_2'$	$\bar{y}_3' + y_4' = y_2''$	$\bar{y}_3'' + y_4'' = y_2'''$	$b_1 = \frac{1}{2^3} y_2'''$
3	$\bar{y}_3$	$\bar{y}_5 + \bar{y}_6 = \bar{y}_3'$	$\bar{y}_5' + y_6' = y_3''$	$\bar{y}_5'' + y_6'' = y_3'''$	$b_2 = \frac{1}{2^3} y_3'''$
4	$\bar{y}_4$	$\bar{y}_7 + \bar{y}_8 = \bar{y}_4'$	$\bar{y}_7' + y_8' = y_4''$	$\bar{y}_7'' + y_8'' = y_4'''$	$b_{12} = \frac{1}{2^3} y_4'''$
5	$\bar{y}_5$	$\bar{y}_2 - \bar{y}_1 = \bar{y}_5'$	$\bar{y}_2' - y_1' = y_5''$	$\bar{y}_2'' - y_1'' = y_5'''$	$b_3 = \frac{1}{2^3} y_5'''$
6	$\bar{y}_6$	$\bar{y}_4 - \bar{y}_3 = \bar{y}_6'$	$\bar{y}_4' - y_3' = y_6''$	$\bar{y}_4'' - y_3'' = y_6'''$	$b_{13} = \frac{1}{2^3} y_6'''$
7	$\bar{y}_7$	$\bar{y}_6 - \bar{y}_5 = \bar{y}_7'$	$\bar{y}_6' - \bar{y}_5' = \bar{y}_7''$	$\bar{y}_6'' - \bar{y}_5'' = \bar{y}_7'''$	$b_{23} = \frac{1}{2^3} y_7'''$

8	$\bar{y}_8$	$\bar{y}_8 - \bar{y}_7 = \bar{y}_8'$	$\bar{y}_8' - \bar{y}_7' = \bar{y}_8''$	$\bar{y}_8'' - \bar{y}_7'' = \bar{y}_8'''$	$b_{123} = \frac{1}{2^3} y_8'''$

При проведении экспериментов, чаще всего используют полиномиальные модели, которые могут быть линейными, квадратичными и более высоких порядков.

Линейная модель: $Y = b_0 x_0 + b_1 x_1 + b_2 x_2 + \ldots + b_n x_n = \sum_{i=0}^{n} B_i X_i$;

Квадратичная модель: $Y = \sum_{i=1}^{n} b_i x_i + \sum_{i<j} b_{ij} x_i x_j + \sum_{i=1}^{n} b_{ii} x_i^2$;

Где b_i коэффициенты модели, показывающие степень влияния каждого *i*-го фактора на функцию отклика; x_0 фиктивная переменная, равная 1, используемая для нахождения свободного члена уравнения регрессии b_0; b_0 остаточный член, характеризующий среднее значение функции отклика при $X_i = 0$. Аналогично составляются модели третьего и более высоких порядков. Данные модели еще носят название факторные.

Фактором называется измеряемая переменная величина, принимающая в некоторый момент времени определенное значение. Факторы соответствуют способам воздействия на объект исследования. Каждый фактор имеет область определения. Фактор считается заданным, если вместе с его названием указан область его определения, то есть все значения, которые может принимать данный фактор. Факторы могут быть количественными, то есть их можно измерить, взвесить или качественными – это разные вещества, аппараты и т.д.

При активном планировании эксперимента уровни факторов выбираются симметрично относительно центра плана. Совокупность уровней факторов и образует собой факторное пространство. Факторное пространство иногда называется решеткой планирования.

Число точек плана при ортогональном планировании определяется по формуле: $N = K^n$, где K - число уровней для каждого из факторов, n - число факторов. Так для двухфакторного эксперимента на двух уровнях число точек плана $N = 2^2 = 4$; для трех факторного на двух уровнях $N = 2^3 = 8$; для восьми факторного $N = 2^8 = 256$. То есть по мере увеличения числа точек план быстро увеличивается.

Каждой точке факторного пространства отвечает истинное и опытное значение функции отклика $У_{опытн}$. Совокупность всех истинных значений функции отклика, отвечающих точкам факторного пространства, называется поверхностью функции отклика. ***Основные обозначения:***

X_i *i*-й фактор в натуральном виде; x_i *i*-й фактор в кодированном виде, величина безразмерная; $+1$ верхняя граница (верхний уровень) изменения *i*- го фактора x_i -1 нижняя граница (нижний уровень) изменения *i*- го фактора x_i X_i^0 середина выбранного диапазона измерений *i*- го фактора (основной уровень); λ_i шаг изменения (варьирования) *i*- го фактора в процессе экспериментирования; Y функция отклика; N количество опытов в плане; n количество варьируемых факторов (переменных); γ количество параллельных замеров в каждом опыте; f число степеней свободы той или иной величины; α уровень значимость (вероятность ошибки принятия той или иной гипотезы), $\alpha = 0{,}05$ для $P = 0{,}95$.

Основной (нулевой) уровень является исходной точкой для построения плана эксперимента, обычно это оптимальное (наилучшее значение параметра). Если это значение неизвестно, то выбирается одно из области значений, возможно среднее значение. **Шаг варьирования факторов** – некоторое число (свое для каждого фактора), прибавление которого к основному уровню дает верхний, а вычитание – нижний уровни факторов. В качестве верхнего уровня может быть выбрано максимальное значение фактора, в качестве нижнего минимальное, а в качестве основного – среднее значение.

Совокупность уровней факторов с отвечающими им опытными значениями функции отклика, записанные в таблицу, называют матрицей планирования. Для ее составления необходимо задать границы уровни и шаг варьирования независимых переменных в натуральном и кодированном виде.

Факторы / Уровни	X_1		…	X_n	
	Физ. Значение	код		Физ. значение	код
Верхний	$X_1^в$	+1	…	$X_n^в$	+1
Нижний	$X_1^н$	1	…	$X_n^н$	1
Основной X_i^0	$\frac{X_1^в + X_1^н}{2}$	0	…	$\frac{X_n^в + X_n^н}{2}$	0
Шаг варьирования λ	$\frac{X_1^в - X_1^н}{2}$		…	$\frac{X_n^в - X_n^н}{2}$	

Например, для двухфакторного эксперимента на двух уровнях матрица планирования представляется в виде таблицы, где $X_1^н$ и $X_1^в$ нижний и верхний уровни первого фактора, $X_2^н$ и $X_2^в$ нижний и верхний уровни второго фактора.

№ опыта	Уровни факторов		Функция отклика		№ опыта	Уровни факторов		Кодовое обозначение	Функция отклика
	X_1	X_2				X_1	X_2		
1	$X_1^н$	$X_2^н$	Y_1		1	-1	-1	1	Y_1
2	$X_1^в$	$X_2^н$	Y_2		2	+1	-1	A	Y_2
3	$X_1^н$	$X_2^в$	Y_3		3	-1	+1	B	Y_3
4	$X_1^в$	$X_2^в$	Y_4		4	+1	+1	AB	Y_4

Для упрощения записи 1 может быть опущена. Все возможные комбинации из 2-х факторов, варьируемых на 2-х уровнях, будут исчерпаны, если мы поставим 4 опыта:

№ опыта	Матрица планирования X				Вектор выхода $Y_n = (\bar{y}_n)$
	x_0	x_1	x_2	x_1x_2	
1	+	-	-	+	Y_1
2	+	+	-	-	Y_2
3	+	-	+	-	Y_3
4	+	+	+	+	Y_4

Такое планирование позволяет представить изучаемую зависимость неполным квадратным уравнением: $Y = b_0 + b_1x_1 + b_2x_2 + b_{12}x_1x_2$

Переход от натуральных значений переменных к кодированным и наоборот осуществляется по формуле: $x_i = \frac{X_i - X_i^0}{\lambda_i}$

При увеличении числа факторов X_i например, до 3-х, чтобы исчерпать все возможные комбинации факторов, варьируемых на 2-х уровнях, необходимо поставить 8 опытов:

№ опыта	Матрица планирования X								Вектор выхода
	x_0	x_1	x_2	x_3	x_1x_2	x_1x_3	x_2x_3	$x_1x_2x_3$	
1	+	-	-	-	+	+	+	-	Y_1
2	+	+	-	-	-	-	+	+	Y_2
3	+	-	+	-	-	+	-	+	Y_3
4	+	+	+	-	+	-	-	-	Y_4
5	+	-	-	+	+	-	-	+	Y_5
6	+	+	-	+	-	+	-	-	Y_6
7	+	-	+	+	-	-	+	-	Y_7
8	+	+	+	+	+	+	+	+	Y_8

Как видно из таблицы план проведения экспериментов для трех переменных получили из плана двух переменных, повторив его дважды: один раз при значениях x_3, находящихся на нижнем, а второй раз на верхнем уровнях. Если можно включить в рассмотрение четвертый фактор x_4, то аналогичным образом повторяют планирование и т.д. для сколько угодно числа независимых переменных. После составления матрицы планирования проводится сам эксперимент и его анализ.

1. Дисперсионный анализ результатов эксперимента.

а) проверка измерений на равноточность. Равноточность измерений или однородность полученного ряда оценок дисперсии $S_{y1}^2;\ S_{y2}^2 \ldots\ S_{yN}^2$, определяется по критерию Кохрена.

$$S_{yn}^2 = \frac{\sum(\bar{y}_{ni} - y_{ni})^2}{\gamma - 1};\ \bar{y}_{ni} = \frac{\sum y_{ni}}{\gamma};\ G^{Э} = \frac{\max\{S_{yn}^2\}}{\sum S_{yn}^2}$$

$G^{Э}$ сравнивается с табличным значением $G_{f_1 f_2}^{T}$, если при выбранном уровне значимости α имеет место неравенство $G^{Э} \le G_{f_1 f_2}^{T}$, то измерения во всех γ опытах равноточны, если же $G^{Э} > G_{f_1 f_2}^{T}$, то измерения при выбранном уровне значимости не равноточны и дальнейшие расчеты продолжать не имеет смысла. Тогда необходимо опыт, оценка дисперсии которого наибольшая, повторить более точно и скорректировать исходные данные, либо увеличить число параллельных измерений γ в каждом опыте.

Число степеней свободы критерия Кохрена определяется по формулам: f_1 число степеней свободы числителя $f_1 = \gamma - 1$, f_2 число степеней свободы знаменателя $f_2 = N$.

б) определение ошибки эксперимента. Ошибка вычисляется по формуле:

$$S_{\bar{y}} = \sqrt{\frac{\sum S_{yn}^2}{N \cdot \gamma}}$$

в) определение коэффициентов уравнения регрессии. Коэффициенты уравнения регрессии определяются по следующим формулам: $b_i = \frac{\sum a_i \bar{y}_n}{N};\ b_{ij} = \frac{\sum a_{ij} \bar{y}_n}{N};\ b_{ijk} = \frac{\sum a_{ijk} \bar{y}_n}{N}$ где a_i и $a_{ij} = \pm 1$ берутся из матриц планирования эксперимента.

Вычислять коэффициенты регрессии можно по методу, предложенному Йетсом. Процедура состоит в попарном сложении и вычитании компонент вектора-выхода. Число циклов равно числу факторов, участвующих в эксперименте.

В общем виде схему вычисления можно представить следующим образом:

1	+	*	-
2	+	*	+
3	+	*	-
4	+	*	+
5	+	*	-
6	+	*	+
7	+	*	-
8	+	*	+

Рассмотрим правила вычисления для полного факторного эксперимента ПФЭ типа 2^2 и 2^3.

№ опыта	Вектор выхода	Цикл 1	Цикл 2	Коэффициенты b
1	$\bar{y}_1$	$\bar{y}_1 + \bar{y}_2 = \bar{y}_1'$	$\bar{y}_1' + y_2' = y_1''$	$b_0 = \frac{1}{2^2} y_1''$
2	$\bar{y}_2$	$\bar{y}_3 + \bar{y}_4 = \bar{y}_2'$	$\bar{y}_3' + y_4' = y_2''$	$b_1 = \frac{1}{2^2} y_2''$

3	$\bar{y}_3$	$\bar{y}_2 - \bar{y}_1 = \bar{y}_3'$	$\bar{y}_2' - y_1' = y_3''$	$b_2 = \frac{1}{2^2} y_3''$
4	$\bar{y}_4$	$\bar{y}_4 - \bar{y}_3 = \bar{y}_4'$	$\bar{y}_4' - y_3' = y_4''$	$b_{12} = \frac{1}{2^2} y_4''$

Схема вычисления коэффициентов регрессии для ПФЭ типа 2^3 методом Йетса.

№ опыта	Вектор выхода	Цикл 1	Цикл 2	Цикл 3	Коэффициенты b
1	$\bar{y}_1$	$\bar{y}_1 + \bar{y}_2 = \bar{y}_1'$	$\bar{y}_1' + y_2' = y_1''$	$\bar{y}_1'' + y_2'' = y_1'''$	$b_0 = \frac{1}{2^3} y_1'''$
2	$\bar{y}_2$	$\bar{y}_3 + \bar{y}_4 = \bar{y}_2'$	$\bar{y}_3' + y_4' = y_2''$	$\bar{y}_3'' + y_4'' = y_2'''$	$b_1 = \frac{1}{2^3} y_2'''$
3	$\bar{y}_3$	$\bar{y}_5 + \bar{y}_6 = \bar{y}_3'$	$\bar{y}_5' + y_6' = y_3''$	$\bar{y}_5'' + y_6'' = y_3'''$	$b_2 = \frac{1}{2^3} y_3'''$
4	$\bar{y}_4$	$\bar{y}_7 + \bar{y}_8 = \bar{y}_4'$	$\bar{y}_7' + y_8' = y_4''$	$\bar{y}_7'' + y_8'' = y_4'''$	$b_{12} = \frac{1}{2^3} y_4'''$
5	$\bar{y}_5$	$\bar{y}_2 - \bar{y}_1 = \bar{y}_5'$	$\bar{y}_2' - y_1' = y_5''$	$\bar{y}_2'' - y_1'' = y_5'''$	$b_3 = \frac{1}{2^3} y_5'''$
6	$\bar{y}_6$	$\bar{y}_4 - \bar{y}_3 = \bar{y}_6'$	$\bar{y}_4' - y_3' = y_6''$	$\bar{y}_4'' - y_3'' = y_6'''$	$b_{13} = \frac{1}{2^3} y_6'''$
7	$\bar{y}_7$	$\bar{y}_6 - \bar{y}_5 = \bar{y}_7'$	$\bar{y}_6' - \bar{y}_5' = \bar{y}_7''$	$\bar{y}_6'' - \bar{y}_5'' = \bar{y}_7'''$	$b_{23} = \frac{1}{2^3} y_7'''$
8	$\bar{y}_8$	$\bar{y}_8 - \bar{y}_7 = \bar{y}_8'$	$\bar{y}_8' - \bar{y}_7' = \bar{y}_8''$	$\bar{y}_8'' - \bar{y}_7'' = \bar{y}_8'''$	$b_{123} = \frac{1}{2^3} y_8'''$

2. Дисперсионный анализ уравнения регрессии:

а) определение ошибки при вычислении коэффициентов уравнения регрессии.

Ошибки при вычислении коэффициентов для ПФЭ типа 2^n одинаковы и определяются по формуле: $S^2 = \{b_n\} = S_{\bar{y}}^2 \frac{1}{N}; или \;\; S = \{b_n\} = S_{\bar{y}} \frac{1}{\sqrt{N}};$

б) проверка значимости коэффициентов уравнения регрессии. Проверку значимости коэффициентов уравнения регрессии проводят по t критерию Стьюдента. Для каждого коэффициента определяют $t^Э$ критерий по формуле: $t_n^Э = \frac{|b_n|\sqrt{N}}{S_{\bar{y}}}$ и сравнивают с табличным значением $t_{\alpha,f}^T$ выбранного для заданного уровня значимости и числа степеней свободы, которое равно: $f = N(\gamma - 1)$.

Если для данного коэффициента выполняется неравенство $t^Э \geq t_{\alpha,f}^T$, то при выбранном уровне значимости соответствующий коэффициент уравнения регрессии значим, если неравенство не выполняется, то коэффициент не значим и членом уравнения регрессии с этим коэффициентом можно пренебречь.

в) проверка гипотезы об адекватности полученного уравнения действительному процессу можно рассчитать двумя способами используя F критерий Фишера:

по формуле вычисляют экспериментальное значение статистики:

$$F^{Э} = \frac{S^2_{yag}}{S^2_{\bar{y}}}; S^2_{yag} = \frac{\sum \bar{y}^2_n - N\sum b^2_n}{N-(k+1)}; F^T_{f_1, f_2}$$

где k число значимых коэффициентов регрессии, не считая свободного члена, и сравнивают с табличным значением, где f_1 число степеней свободы числителя отношения, $f_1 = N - (k+1)$; f_2 число степеней свободы знаменателя отношения, $f_2 = N(\gamma - 1)$, γ - число параллельных измерений.

Если при выбранном уровне значимости выполняется неравенство $F^{Э} \le F^T_{f_1, f_2}$ то полученное уравнение регрессии адекватно действительному уравнению, если неравенство не выполняется, то полученное уравнение регрессии неадекватно описывает полученный процесс;

опытное значение критерия Фишера может быть взятым равным отношению дисперсии неадекватности к общей дисперсии или дисперсии воспроизводимости:

$$F_{опыт} = \frac{S^2 Y_{н.а.}}{S^2 Y_{воспр.}}; \quad S^2 Y_{н.а.} = \frac{\gamma \cdot S^2_{ОСТ}}{N-k} = \frac{\gamma \cdot \sum_{i=1}^{N} (Y_{i\,рас} - \bar{Y}_{i\,опыт.})^2}{N-k};$$

$$S^2 Y_{воспр.} = \frac{\sum_{i=1}^{N}\sum_{j=1}^{k} (Y_{ij\,рас.} - Y_{i\,опыт.})^2}{N(k-1)}$$

г) определение границ доверительных интервалов для коэффициентов регрессии. Границы доверительных интервалов, в которых с заданной вероятностью $P = 1 - \alpha$ находятся коэффициенты уравнения регрессии $\{\beta\}$ определяются по формуле: $b_n - t^T_{\alpha, f} S\{b_n\} \le \beta_n \ge b_n + t^T_{\alpha, f} S\{b_n\}$ где $t^T_{\alpha, f}$ табличное значение t - критерия Стьюдента при заданном уровне значимости и числе степеней свободы $f = N(\gamma - 1)$.

ПЛАНИРОВАНИЕ ЭКСПЕРИМЕНТА

ОЧЕРК ПО ИСТОРИИ ПЛАНИРОВАНИЯ ЭКСПЕРИМЕНТА

Планирование эксперимента — продукт нашего времени, однако истоки его теряются в глубине веков. Давайте попытаемся проследить за течением этих ручейков, слившихся сегодня в реку планирования эксперимента.

Существует известный произвол в определении того, относится ли то или иное событие или явление к истокам планирования эксперимента или нет. Поэтому отбор материала в известной мере определяется вкусами и информированностью авторов.

Понимая существующие трудности, начнем все же рассмотрение истории, ибо как иначе можно узнать о будущем?

1. Предыстория

Истоки планирования эксперимента уходят в глубокую древность и связаны с числовой мистикой, пророчествами и суевериями. Это собственно не планирование физического эксперимента, а планирование числового эксперимента, т. е. расположение чисел так, чтобы выполнялись некоторые строгие условия, например, на равенство сумм по строкам, столбцам и диагоналям квадратной таблицы, клеточки которой заполнены числами натурального ряда. Такие условия выполняются в магических квадратах, которым, по-видимому, принадлежит первенство в планировании эксперимента.

Согласно одной легенде примерно в 2200 г. до н. э." китайский император Ю выполнял мистические вычисления с помощью магического квадрата, который был изображен на панцире божественной черепахи.

Квадрат императора Ю

4	9	2
3	5	7
8	1	6

Клетки этого квадрата заполнены числами от 1 до 9, и суммы чисел **по** строкам, столбцам и главным диагоналям равны 15.

В 1514 г. немецкий художник Альбрехт Дюрер изобразил магический квадрат в правом углу своей знаменитой гравюры аллегории «Меланхолия». Два числа в нижнем горизонтальном ряду (15 и 14) составляют год создания гравюры. В этом состояло своеобразное «приложение» магического квадрата.

Квадрат Дюрера

16	3	2	13
5	10	11	8
9	6	7	12
4	15	14	1

В течение нескольких веков построение магических квадратов занимало умы индийских, арабских, немецких, французских математиков.

В настоящее время магические квадраты используются при планировании эксперимента в условиях линейного дрейфа, при планировании экономических расчетов и составлении рационов питания, в теории кодирования и т. д.

Построение магических квадратов является задачей комбинаторного анализа, основы которого в его современном понимании заложены Г. Лейбницем. Он не только рассмотрел и решил основные комбинаторные задачи, но и указал на большое практическое применение комбинаторного анализа: к кодированию и декодированию, к играм и статистике, к логике изобретений и логике геометрии, к военному искусству, грамматике, медицине, юриспруденции, технологии и к комбинации наблюдений. Последняя область применения наиболее близка к планированию эксперимента.

Г. Лейбниц проектировал создание «обобщенной математики», используя которую можно свести понятия к символам, символы к числам и благодаря этому подвергнуть понятие механическому вычислению. Основная роль при этом отводилась комбинаторике.

Этот проект во времена Лейбница казался абсурдным и фантастическим, но в настоящее время, благодаря развитию вычислительной техники и математической логики, часть этого плана ужи реализуется. И в его реализации немалую роль сыграло планирование эксперимента.

Одной из комбинаторных задач, имеющей прямое отношение к планированию эксперимента, занимался известный петербургский математик Л. Эйлер. В 1779 г. он предложил задачу о 36 офицерах как некоторый математический курьез.

Он поставил вопрос, можно ли выбрать 36 офицеров 6 рангов из 6 полков по одному офицеру каждого ранга от каждого полка и расположить их в каре так, чтобы в каждом ряду и в каждой шеренге было бы по одному офицеру каждого ранга и по одному от каждого полка. Задача эквивалентна построению парных ортогональных 6x6 квадратов. Оказалось, что эту задачу решить невозможно. Эйлер высказал предположение, что не существует пары ортогональных квадратов порядка $n \equiv 2 \pmod 4$.

Задачей Эйлера, в частности, и латинскими квадратами вообще занимались впоследствии многие математики, однако почти никто из них не задумывался над практическим применением латинских квадратов. Тем больший интерес могут представить работы немецкого математика Э. Шредера, который в конце XIX в. применил латинские квадраты в алгебре логики при разработке теории алгоритмов и исчислений.

Для того чтобы полностью выписать все формулы, входящие в какой-либо алгоритм, Шредер поставил задачу отыскать все возможные исчисления и получить их полный перебор. При этом ему стало ясно, что сделать полный перебор в большинстве случаев практически невозможно и нужно думать о том, как его уменьшить, как доказать, что данная система функциональных уравнений замкнута относительно логических следствий из них и т. д. Для сокращения перебора Шредер использовал латинские квадраты, которыми он занимался с большим увлечением. Эта работа использовалась Шредером для получения чисто логических результатов — для доказательства независимости закона дистрибутивности конъюнкции относительно дизъюнкции от системы более простых логических законов.

В 1900 г. П. Терри подтвердил путем систематического пересчета справедливость предположения Эйлера для n ≡ 2 = 6. Р. Боуз, С. Шрикханде и Е. Паркер доказали теорему о существовании пары ортогональных квадратов порядка n ≡ 2 (mod 4) для n ≠ 6, и предположение Эйлера оказалось справедливым только для случая, когда n = 6. Так Эйлеровская задача о 36 офицерах способствовала интенсивному изучению теории латинских квадратов многими математиками в различных странах в течение последних 150 лет.

В настоящее время латинские квадраты являются одним из наиболее популярных способов ограничения на рандомизацию при наличии источников неоднородностей дискретного типа в планировании эксперимента. Группировка элементов латинского квадрата, благодаря своим свойствам (каждый элемент появляется один и только один раз в каждой строке и в каждом столбце квадрата), позволяет защитить главные эффекты от влияния источника неоднородностей. Широко используются латинские квадраты и как средство сокращения перебора в комбинаторных задачах.

Говоря об истоках планирования эксперимента с ограничением на рандомизацию, уместно вспомнить имя еще одного немецкого математика XIX в., положившего начало теории блок-схем или неполноблочных планов, как принято говорить в планировании эксперимента. Блок-схемы впервые изучались И. Штейнером в 1850 г. с точки зрения комбинаторики как тактические конфигурации. В планировании эксперимента блок-схемы стали использоваться значительно позже, о чем расскажем далее.

Мы остановились на предыстории планирования эксперимента, связанной с комбинаторным анализом. Обратимся теперь к другим областям знаний. В европейской философии основы методологии научных исследований закладывались Р. Декартом, Ф. Бэконом и Дж. Миллем. Так, Милль в 1843 г., рассматривая причинные связи явлений, формулировал известные методы опытного исследования: метод сходства, метод различия, метод сопутствующих явлений и др. Эти методы представляли собой простейшие правила, которые помогали устанавливать взаимосвязи между причинами и следствиями (и наоборот). Теория Милля была основана на весьма сильных упрощающих предположениях (постулировался набор факторов, игнорировались взаимодействия между ними и случайные воздействия). При этом не использовался математический аппарат для количественных оценок.

В математике закладывались основы количественных методов исследования. Работа шла в нескольких направлениях. П. Ферми и Б. Паскаль в переписке обсуждали задачи теории вероятностен (XVH в.). Тейлор, Маклорен и Л. Эйлер (снова Эйлер!) (XVIIT п.) разработали аппарат теории приближения функций степенными рядами. А. Коши (XIX в.) развил метод градиента для решении систем совместных уравнений. Ю. Кели создал матричные обозначения.

А. Лежандр и К. Гаусс создали и обосновали с вероятностных позиций метод наименьших квадратов. Позже все эти направления соединились в планировании эксперимента.

Обсуждая нашу тему, нельзя не коснуться положения дел в физике. Здесь появляются работы Дж. Максвелла — метод градиента проникает в физику. Дж. Гиббс вводит представление о поверхности отклика и вместе с Л. Больцманом утверждает, что природе свойственны скорее статистические, чем детерминированные закономерности.

Начало XX в. знаменуется созданием английской статистической школы (К. Пирсон), получением важных результатов Стьюдента (В. Госсета — английского статистика) по статистике мплых выборок [1—4].

2. Становление

Возникновение современных статистических методов планирования эксперимента связано с именем Р. Фишер. С 1918 г. он начал свою известную серию работ на Рочемстедской агробиологической станции в Англии. В 1935 г. появилась его монография «Design of Experiments», давшая название всему направлению.

Среди методов планирования первым был дисперсионный анализ (кстати, Фишеру принадлежит и термин «дисперсия»). Фишер создал основы этого метода, описав полные классификации дисперсионного анализа (однофакторный и многофакторный эксперименты) и неполные классификации дисперсионного анализа без ограничения и с ограничением на рандомизацию. При этом он широко использовал латинские квадраты и блок-схемы. Вместе с Ф. Йетсом он описал их статистические свойства. В 1942 г. А. Кишен рассмотрел планирование по латинским кубам, которое явилось дальнейшим развитием теории латинских квадратов. Затем Р. Фишер независимо опубликовал сведения об ортогональных гипер-греко-латинских кубах и гипер-кубах. Вскоре после этого (1946—1947 гг.) Р. Рао рассмотрел их комбинаторные свойства. Дальнейшему развитию теории латинских квадратов посвящены работы Х. Манна (1947—1950 гг.). Здесь интересно отметить, что начиная с работ Р. Муфанга (1935 г.) развивается теория квазигрупп, т. е. множеств E с бинарной операцией на E, для которой существуют обе обратные ей операции. Теория квазигрупп имеет прямую связь с теорией латинских квадратов. Так, если область определения конечна и задана списком ее элементов, то эта операция может быть эффективно определена квадратной таблицей с двумя входами, в каждом столбце и в каждой строке которой без повторений помещены все элементы из области E. Такая таблица является латинским квадратом.

Первое глубокое математическое исследование блок-схем выполнено Р. Боузом в 1939 г. Вначале была разработана теория сбалансированных неполно блочных планов (BIB-схем). Затем Р. Боуз, К. Нер и Р. Рао обобщили эти планы и разработали теорию частично сбалансированных неполноблочных планов (PBIB-схем). С тех пор изучению блок-схем уделяется большое внимание как со стороны специалистов по планированию эксперимента (Ф. Йетс, Г. Кокс, В. Кохрен, В.Федерер,К. Гульден, О. Кемпт-горн и многие др.), так и со стороны специалистов по комбинаторному анализу (Боуз, Ф. Шимамото, В. Клатсворси, С. Шрикханде, А. Гоффман и др.). Их усилия были направлены на разработку способов построения блок-схем, составления каталогов планов и на решение вопроса классификации PBIB-схем по ассоциативным схемам.

Исследования Р. Фишера, проводившиеся в связи с работами по агробиологии, знаменуют начало первого этапа развития методов планирования эксперимента. Фишер разработал метод факторного планирования. Йетс предложил для этого метода простую вычислительную схему. Факторное планирование получило широкое распространение. Особенностью полного факторного эксперимента является необходимость ставить сразу большое число опытов. Это вполне пригодно в агробиологии, но связано со значительными трудностями в технических приложениях.

В 1945 г. Д. Финни ввел дробные реплики от факторного эксперимента. Это позволило резко сократить число опытов и открыло дорогу техническим приложениям планирования. Другая возможность сокращения необходимого числа опытов была показана в 1946 г. Р. Плакеттом и Д. Берманом, которые ввели насыщенные факторные планы.

Г. Хотеллинг в 1941 г. обратил внимание на то, что многие практические, особенно технические, задачи можно рассматривать как экстремальные. Он предложил находить экстремум по экспериментальным данным с использованием степенных разложений и градиента (подробно для случая одной независимой переменной). Проблемы, поставленные второй мировой войной, привели к интенсификации исследований в этом направлении.

Следующим важным этапом было введение принципа последовательного шагового экспериментирования. Этот принцип, высказанный в 1947 г. М. Фридманом и Л. Сэвиджем, позволил распространить на экспериментальное определенно экстремума известный в математике прием — итерацию. Идейно их работа примыкает работам А. Вальда по последовательному анализу, сформировавшемуся в 1943—1950 гг.

Чтобы построить современную теорию планирования эксперимента, не хватало одного звена — формализации объекта исследования. Это звено появилось в 1947 г. после создания Н. Винером теории кибернетики. Кибернетическое понятие «черный ящик», играет в планировании важную роль.

Кроме того, кибернетика стимулировала развитие вычислительной техники, дав в руки исследователей инструмент, позволяющий быстро решать возникающие при планировании математические задачи [3, 5].

3. Развитие

В 1951 г. работой американских ученых Дж. Бокса и К. Уилсона начался новый этап развития планирования эксперимента. Эта работа подытожила предыдущие. В ней ясно сформулирована и доведена до практических рекомендаций идея последовательного экспериментального определения оптимальных условий проведения процессов с использованием оценки коэффициентов степенных разложений методом наименьших квадратов, движения по градиенту и отыскания интерполяционного полинома (степенного ряда) в области экстремума функции отклика («почти стационарной» области).

В 1954—1955 гг. Дж. Бокс, а затем Дж. Бокс и П. Юл показали, что планирование эксперимента можно использовать при исследовании физико-химических механизмов процессов, если априори высказаны одна или несколько возможных гипотез. Здесь планирование эксперимента пересекалось с исследованиями по химической кинетике. Интересно отметить, что кинетику можно рассматривать как метод описания процесса с помощью дифференциальных уравнений, традиции которого восходят к И. Ньютону. Описание процесса дифференциальными уравнениями, называемое детерминистическим, нередко противопоставляется статистическим моделям. Указанные работы наметили конкретный путь сближения этих подходов. Если кинетическая модель уже известна, возникает вопрос о планировании эксперимента для уточнения полученных ранее констант. Это направление также получило развитие в последние годы, в частности в работах Н. П. Клепикова и С. Н. Соколове, а затем В. В. Федорова в связи с задачами ядерной физики.

Если для первого этапа развития планирования эксперимента характерны полевые и лабораторные агробиологические исследования, то второй этап — это этап лабораторных, главным образом химических, исследований. Хотя развитие этого направления еще не закончено, но уже можно говорить о третьем этапе — этапе промышленных экспериментов. Он начался в 1957 г., когда Бокс модифицировал свой метод, приспособив его для использования в промышленности. Этот метод стал называться «эволюционным планированием».

Почти одновременно с эволюционным планированием Бокс и Дж. Хантер сформулировали принцип ротатабельности для описания «почти стационарной» области, развивающейся в настоящее время в важную ветвь теории планирования эксперимента. В той же работе показана возможность планирования с разбиением на ортогональные блоки, указанная ранее независимо де Бауном. Дальнейшим развитием этой идеи было планирование, ортогональное к неконтролируемому временному дрейфу, которое следует рассматривать как важное открытие в экспериментальной технике — значительное увеличение возможностей экспериментатора.

Слабым местом было отсутствие метода оценки риска, связанного с тем, что в рассмотрение не включен какой-либо важный фактор. Для преодоления этой трудности в 1957—1959 гг. Ф. Сатерзвайтом был развит метод «случайного баланса» (первый план со случайными уровнями факторов (для 30 факторов) был, вероятно, реализован Е. Шроком в 1949 г. Первый такой эксперимент Ф. Сатерзвайт провел в 1952 г. (для 14 факторов), а теоретические основы развил в 1955—1956 гг.) Этот метод возник не в связи с общим ходом развития статистических идей, а скорее как прием формализации психо-физиологических методов решения сложных задач человеком. Поэтому он вызвал острую дискуссию.

В 1958 г. Г. Шеффе предложил новый метод планирования эксперимента для изучения физико-химических диаграмм состав— свойство. В этом методе, названном метод ом «симплексной решетки», используется то обстоятельство, что состав многокомпонентной системы дается точкой в правильном симплексе. Здесь произошло пересечение с химической топологией, созданной Н. С. Курнаковым [б].

Симплексы оказались пригодными и для построения ротатабельных планов, что было показано в 1960 г. Дж. Боксом и Д. Бен-киным. С. Адельманом в 1961 г. обобщена идея Финни о дробных репликах. Им построены нерегулярные дробные реплики, которые в ряде случаев позволяют существенно сократить число опытов.

С развитием средств автоматизации и вычислительной техники возник вопрос о возможности использования какого-либо алгоритма планирования при автоматическом управлении процессом с помощью вычислительной машины. Этот вопрос был теоретически решен в 1962 г. В. Спендлейем, Дж. Хекстом и Ф. Химснорстом, Боксом и Дж. Дженкинсом, создавшими основы теории адаптационной оптимизации, в которой сочетаются идеи планировании и теория случайных функций.

Следует отметить, что ряд научных концепций, развивающихся параллельно с планированием эксперимента, оказывает влиянии на его развитие: к их числу, кроме уже упоминавшихся, относятся теория принятия решений, теория игр, линейное, нелинейное и динамическое программирование, правдоподобные рассуждения, введенные Д. Пойа. Планирование эксперимента входит как составная часть в более общие концепции: исследование операций и системотехнику, теорию управляющих систем.

Современному состоянию планирования эксперимента и развитию этого направления в СССР посвящены две статьи в Информационных материалах научного совета по комплексной проблеме «Кибернетика» № 10 1969 г. Это — статья В. В. Налимова «Решенные и нерешенные проблемы планирования эксперимента» и статья Ю. П. Адлера, Ю. В. Грановского, Е. В. Марковой «Некоторые аспекты развития исследований по планированию эксперимента».

Отметим, однако, что наиболее характерной чертой, пожалуй, является объединение подхода Дж. Бокса с подходом американского математика Дж. Кифера, рассматривавшего формальные аспекты теории планирования. Значительные усилия для такого объединения делаются в СССР под руководством В. В. Налимова.

Растут возможности методов планирования эксперимента, расширяется сфера их приложения. Надо полагать, что развитие уже достигло такого состояния, когда можно писать с большой буквы — «Математическая теория эксперимента», когда накоплен значительный практический опыт.

Общие сведения о математической теории планирования эксперимента

Познание свойств материального мира включает следующие этапы: первичное изучение некоторых явлений, процессов при помощи опытов, создание гипотезы (теории) и ее проверка на опыте или эксперименте.

Во всех отраслях производства и науки менеджерам приходится решать те или иные проблемы. Они проходят при своем разрешении некоторым цикл: от постановки проблемы к процессу ее решения и далее от результата решения к воплощению его в практике. Исследования строятся по принципу «гипотеза — информация (наблюдение, эксперимент) — статистическая обработка (проверка гипотез на основе вероятностной)». Часто они вновь возникают на качественно более высоком уровне развития и вновь решаются с использованием уже накопленного опыта (рис. 1.1).

Из определения эмпирического уровня познания ясно, что на этом уровне решение любой задачи всегда опирается на и**нформацию** о конкретной технико-экономической ситуации. Эта информация может быть получена при наблюдении (изучение объекта без **вмешательства** в его нормальное функционирование) и эксперименте (изучение объекта при **целенаправленном воздействии** на его параметры), а также при изучении опыта других исследователей и производственного опыта (I_1). Накопленная к началу новых работ информация, получаемая из печатных, актовых, рукописных или не документальных источников, позволяет также до постановки собственных наблюдений и экспериментов ответить на ряд важных вопросов, в частности:

- на каких исходных методических предпосылках (методы сбора и обработки информации, обеспечение точности измерений, инструменты, организация работ и т. д.) строить дальнейшее эмпирическое исследование;
- как учесть достоинства и недостатки предыдущих работ в данной отрасли; что можно использовать в исследовании из методик «смежных» (а может быть, и «далеких») областей и т. д;
- каковы общие совпадающие выводы в ранее выполненных исследованиях и каковы (что наиболее важно) противоречия в этих выводах и у разных авторов; в чем возможные причины таких противоречий;
- каковы теоретические взгляды на решение поставленной технико-экономической задачи; какие «смежные» теории можно привлечь к объяснению ее результатов и т. д.

На основе анализа информации (I_1). исследователь выбирает один из возможных путей исследования (А—Г на рис. 1.1).

Путь А. Если исходной информации (I_1) достаточно, то нужно сформулировать технико-экономические гипотезы и проверить их вероятностно-статистическими методами (этапы VIII—IX). Эти методы объединяются единой философской концепцией, основанной на материалистической категории случайности. Фундаментом исследований является теория вероятностей, в частности, ее аксиомы и теоремы, законы распределения вероятностей, теория зависимости между случайными величинами, случайные процессы и функции.

Математическая статистика объединяет методы, которые позволяют по результатам испытаний, содержащихся в информации (в частности, в I_1) делать технико-экономические **выводы с определенной достоверностью.** Наиболее часто применяются такие ее ветви, как описательная статистика и выборочный метод, распределение выборочных характеристик и теория эффективных оценок, проверка гипотез и непараметрические методы, последовательный анализ, корреляционный и регрессионный анализы, анализ временных рядов и случайных процессов.

Пути Б и В. Если исходной информации недостаточно, то необходимо провести наблюдение за объектом или эксперимент над ним. Как видно из рис. 1.1, результаты наблюдений (I_2) и экспериментов (I_3) также должны проходить математико-статистическую обработку.

Слово "эксперимент" происходит от латинского "experimentum" и означает опыт, проба. Под экспериментом понимают изучение природных явлений путем научного опыта при точно учитываемых условиях, которые дают возможность следить за ходом исследуемого процесса и повторять его каждый раз при воспроизведении тех же условий.

Путь Г. Проведение и анализ экспериментов будут существенно оптимизироваться, если привлечь **математическую теорию эксперимента**. Следует сразу обратить внимание на три особенности такого пути (заштрихованное поле на рис. 1.1):

математический аппарат используется не только на обычном этапе после проведения эксперимента (Э.2), но и на новом этапе— предшествующем (Э.1);исследователь планирует эксперимент исходя из целей данного этапа работ (выбор наиболее существенных факторов, поиск оптимума, получение интерполяционной зависимости, описание явления и т. д.) так, чтобы будущая математическая модель заранее обладала некоторым набором свойств, оптимальным с точки зрения этих целей (минимум экспериментальных точек для ее построения или независимость оценок коэффициентов модели, или максимальная точность оценки координат оптимума, или нечувствительность результатов к грубым ошибкам в опытах и др. 119]);

все экспериментальное исследование **разделяется на этапы,** на каждом из которых проводится специфическое планирование: эксперимент—построение модели—интерпретация модели и принятие решения о дальнейшем направлении исследования; после каждого шага можно принять строго обоснованное решение о том, что делать дальше; систематическое применение в исследованиях такой последователь-поп шаговой процедуры способствует значительному умень-шению общего объема экспериментальных работ;

математическое планирование экспериментов дает нам набор четких, заранее составленных **алгоритмов,** которые могут быть применены к широкому классу объектов (от управления технологическими процессами до поиска лучших форм лечения); математическое многофакторное моделированиие на основе таких планов может быть осуществлено и без ЭВМ (например, при расчете на специальных бланках 18] с помощью арифмометров, см. гл. 3).

Как бы ни было построено исследование до принятия решения о реализации его результатов (этап XII) в практике данной отрасли производства или науки (этап XIII), ему нужно пройти технико-экономический анализ этих результатов (этап X) и экспериментально-производственную проверку (этап XI).

Этап X является обязательным, поскольку на нем производится обратный **перевод** математических символов (числовых результатов, моделей и. т. д.) в термины данной отрасли с расшифровкой полученных выводов и их тщательной профессионально-логической проверкой. Математический результат служит для того, чтобы исследователь на его основании воссоздал в своем мышлении необходимые технико-экономические понятия и принял правильное решение, отвечающее на поставленные в исходной проблеме задачи.

Планирование эксперимента – это процедура выбора числа и условий проведения опытов, необходимых и достаточных для решения поставленной задачи с требуемой точностью.

При исследовании различных процессов экономики и техники экспериментатор, как правило, сталкивается со случайными событиями и случайными величинами. Это объясняется тем обстоятельством, что всегда имеет место разброс параметров в пределах заданных допусков. Кроме того, измерения производятся с ошибками. Поэтому обработка полученных результатов может быть произведена только с помощью статистических методов. Однако статистические методы позволяют лишь в среднем оценить полученные результаты и определить разброс оцениваемых процессов относительно их средних значений. Математическая теория планирования эксперимента позволяет экспериментатору спланировать опыт так, чтобы при минимальной затрате времени и средств получить максимальную информацию. При этом основной задачей математической теории планирования эксперимента является разработка методов получения математических моделей адекватно описывающих изучаемые процессы или изучаемые явления. Эта задача называется задачей идентификации.

Если модель достаточно точно описывает объект, то эксперимент на объекте можно заменить экспериментом на модели. В настоящее время большое распространение наряду с физическими моделями получили абстрактные математические модели.

Моделирование ~ метод анализа экспериментальной информации. Математическая модель ~ математическое отображение наиболее существенных сторон процесса. Математическая модель ~ записанная в математических символах абстракция реального явления, так сконструированная, чтобы анализ ее давал возможность проникнуть в сущность явления. Математическая модель устанавливает соотношение между совокупностью переменных, которые являются параметрами управления явления.

Математические модели могут иметь различные аналитические выражения. Они могут быть выражены с помощью полиномов 0-го, 1-го, 2-го и более высоких порядков, а также с помощью показательных, логарифмических и других функций. Определение конкретного вида модели адекватно описывающей на основе опытных данных изучаемый процесс, зависит от наличия у экспериментатора априорной информации о виде модели. Чем выше степень информирования, тем лучше решается поставленная задача.

0-й степени: $Y = b_0$; 1-й степени: $Y = b_0 + b_1 x_1 + b_2 x_2$;

2-й $Y = b_0 + b_1 x_1 + b_2 x_2 + b_{12} x_1 x_2 + b_{11} x_1^2 + b_{22} x_2^2;$

$$Y = b_0 + b_1 x_1 + b_2 x_2 + b_{12} x_1 x_2 + b_{11} x_1^2 + b_{22} x_2^2 + b_{112} x_1^2 x_2 + b_{122} x_1 x_2^2 + b_{111} x_1^3 + b_{222} x_2^3$$

где $\tilde{y}$ зависимая переменная; $\tilde{x}$ управляемые независимые переменные; *в* - коэффициенты.

Обычно все объекты реального мира выражаются дифференциальными уравнениями. Это самая распространенная модель, так как законы природы описываются математически дифференциальными уравнениями.

$$a_n y^{(n)} + a_{n-1} y^{(n-1)} + \ldots + a_1 y^1 + a_0 y = B_e U^{(e)} + B_{e-1} U^{(e-1)} + \ldots + B_1 U^1 + B_0 U$$

где *A* и *B* функции времени *a(t), в(t).*

Математические модели универсальны, то есть модели созданные для исследования одних процессов и систем, затем с успехом могут использоваться для изучения других.

Часто во многих случаях в задачах необходимо установить зависимость случайной величины *Y* от одной или нескольких случайных величин *X*. Эти зависимости могут быть статистическими и корреляционными. Статистической называется зависимость, при которой изменение одной из величин влечет изменение распределения другой. Статистическая зависимость проявляющееся в том, что при изменении одной из величин изменяется среднее значение другой, называется корреляционной.

Основные положения математической теории планирования эксперимента

При планировании эксперимента существенным являются следующие положения:

- минимизации общего числа опытов;
- одновременным изменением (варьированием) всеми параметрами, определяющими процесс по специальным правилам – алгоритмам;

- использование математического аппарата, формализующего действия экспериментатора;
- выбора четкой стратегии, позволяющей принимать обоснованные решения после каждой серии эксперимента.

Планирование эксперимента можно разделить на два направления:

~ планирование с задачей выявления механизма явления или процесса, т.е. выявления степени влияния различных факторов на исследуемый параметр оптимизации. Полученная при этом модель служит интерполяционной формулой. Интерполяция (изменение, переделка) ~ в математике и статистике отыскание промежуточных значений величины по некоторым известным ее значениям. Например, отыскание значения функции *f(x)* в точках *x*, лежащих между точками $x_0<x_1<\ldots<x_n$, по известным значениям $y_i = f(x_i)$ где $i = 0, 1, 2 \ldots n$). Если *x* лежит вне интервала (x_0, x_n), аналогичная процедура называется экстраполяция. С помощью интерполяционной формулы представляется возможность выявить силу влияния различных факторов на изучаемый параметр.

~ планирование так называемых экстремальных экспериментов (оптимизационных задач) с задачей определения таких значений уровней факторов, при которых исследуемый параметр оптимизации будет иметь наибольшее (наименьшее) значение.

При решении указанных задач обычно устанавливается следующая последовательность действий.

1. Вначале делается описание эксперимента и формулируется его основная задача; выбирается параметр оптимизации (или его общий вид), с помощью которого будет происходить оценка качества процесса; определяются факторы, оказывающие влияние на параметр оптимизации, и определяются границы изменения факторов.

2. Если на функцию отклика оказывают значительное влияние несколько факторов, то вначале, применяя экспертный метод или метод случайного баланса, производится отсев всех незначащих факторов.

3. Выбирается центральная точка эксперимента, для определения которой диапазон изменения каждого из факторов делится пополам. Этим самым определяется центр плана и шаг варьирования по каждому из аргументов.

4. Выбирается вид математической модели, которой предполагается описывать рассматриваемое явление. Как правило, вначале это явление описывается линейной моделью.

5. В зависимости от условий и задач эксперимента выбирается вид планирования.

6. Определяется возможность использования дробных реплик.

7. Производится эксперимент, на основе которого определяется ММ процесса.

8. Производится проверка воспроизводимости эксперимента.

9. Отсеиваются незначащие эффекты с помощью критерия Стьюдента.

10. Полученная модель проверяется на адекватность с помощью критерия Фишера.

11. Если полученная линейная модель окажется адекватной, то производится движение по вектору-градиенту в область экстремума, по достижении которой, уже с помощью квадратичной модели, определяется экстремум, т.е. значения факторов, при которых параметр оптимизации будет иметь оптимальное значение.

В самом общем виде математическую модель или зависимость параметров оптимизации от факторов можно записать в виде уравнения, называемого уравнением регрессии: $Y = \varphi(x_1,\ldots x_n; z_1,\ldots z_m, w)$, где Y~ параметр оптимизации, функция отклика, реакция, т.е. зависимая переменная; $x_1, x_2 \ldots, x_n$~ факторы, т.е. управляемые, независимые переменные (каждый фактор может принимать в опыте несколько значений или уровней), *z* и w~ переменные и постоянные влияющие на параметр оптимизации, не поддающиеся управлению.

Уравнение регрессии и ограничения, накладываемые на изменение варьируемых факторов, называют ММ.

При проведении экспериментов, чаще всего используют полиномиальные модели, которые могут быть линейными, квадратичными и более высоких порядков.

Линейная модель: $Y = b_0x_0 + b_1x_1 + b_2x_2 + \ldots + b_nx_n = \sum_{i=0}^{n} B_i X_i;$

Квадратичная модель: $Y = \sum_{i=1}^{n} b_i x_i + \sum_{i<j} b_{ij} x_i x_j + \sum_{i=1}^{n} b_{ii} x_i^2;$

Где b_i коэффициенты модели, показывающие степень влияния каждого *i*-го фактора на функцию отклика; x_0 фиктивная переменная, равная 1, используемая для нахождения свободного члена уравнения регрессии b_0; b_0 остаточный член, характеризующий среднее значение функции отклика при $X_i = 0$. Аналогично составляются модели третьего и более высоких порядков. Данные модели еще носят название факторные.

ПАРАМЕТР ОПТИМИЗАЦИИ

При планировании экстремального эксперимента очень важно определить параметр, который нужно оптимизировать. Сделать это совсем не так просто, как кажется на первый взгляд. Цель исследования должна быть сформулирована очень четко и допускать количественную оценку. Будем называть характеристику цели, заданную количественно, параметром оптимизации. Параметр оптимизации является реакцией (откликом) на воздействие факторов, которые определяют поведение выбранной вами системы. Реакция объекта многогранна, многоаспектна. Выбор того аспекта, который представляет наибольший интерес, как раз и задается целью исследования.

При традиционном нематематическом подходе исследователь стремится как-то учесть разные аспекты, взвесить их и принять *согласованное* решение о том, какой опыт *лучше.* Однако разные экспериментаторы проведут сравнение опытов неодинаково. Различия, если хотите, одно из проявлений таланта исследователя или его бездарности.

Прежде чем сформулировать требования к параметрам оптимизации и рекомендации по их выбору, познакомимся с различными видами параметров.

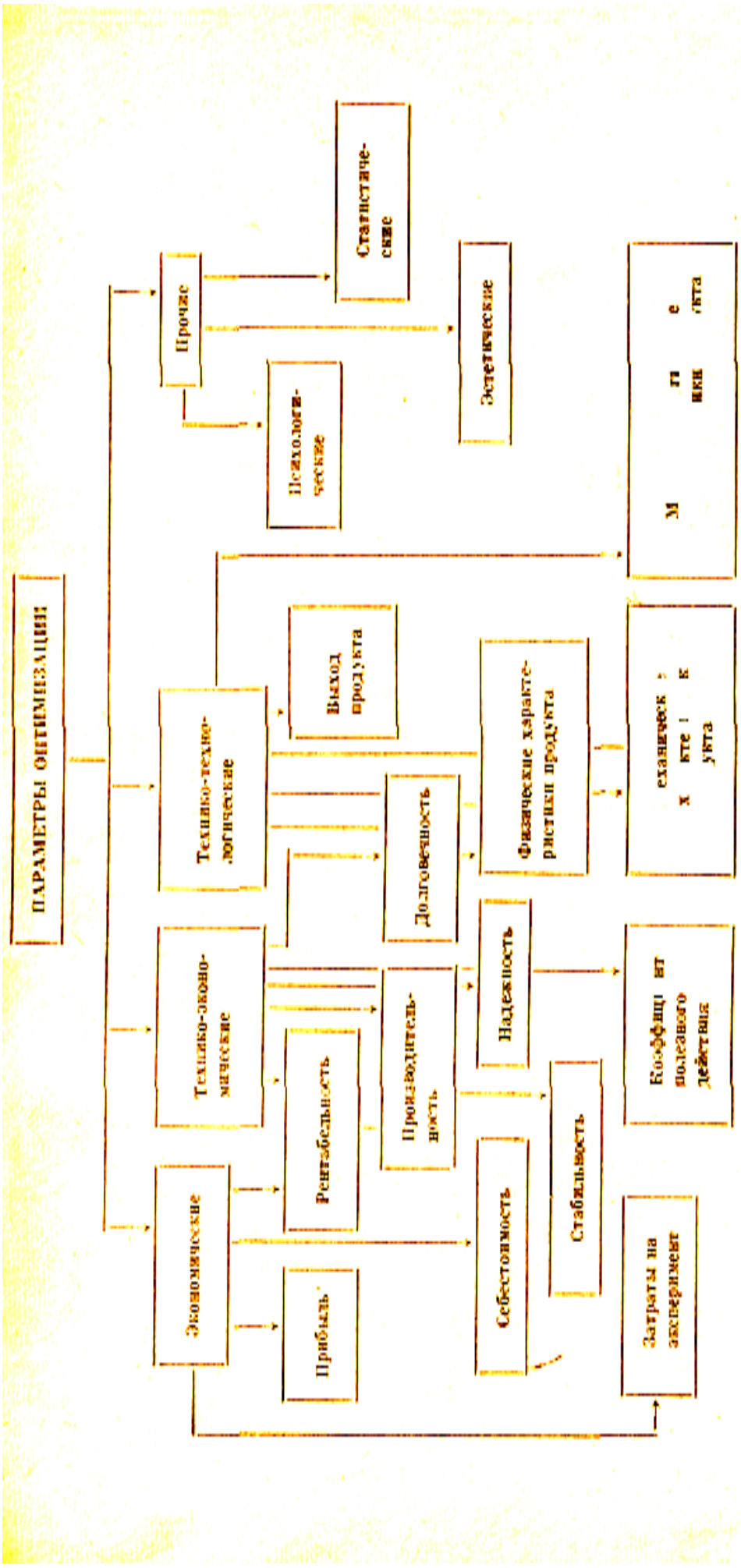
2.1. Виды параметров оптимизации

В зависимости от объекта и цели исследования параметры оптимизации могут быть весьма разнообразными. Чтобы ориентироваться в этом многообразии, введем некоторую классификацию (рис. 2). Мы не стремимся к созданию полной и детальной классификации. Наша задача — построить такую условную схему, которая включала бы ряд практически важных случаев и помогала экспериментатору ориентироваться в реальных ситуациях.

Реальные ситуации, как правило, сложны. Они часто требуют одновременного учета нескольких, иногда очень многих, параметров. В принципе каждый объект может характеризоваться сразу всей совокупностью параметров, приведенных на рис. 2, или любым подмножеством из этой совокупности. Движение к оптимуму возможно, если выбран один-единственный параметр оптимизации. Тогда прочие характеристики процесса уже не выступают в качестве параметров оптимизации, а служат ограничениями. Другой путь — построение обобщенного параметра оптимизации как некоторой функции от множества исходных [1—3].

Прокомментируем некоторые элементы схемы.

Экономические параметры оптимизации, такие, как прибыль, себестоимость и рентабельность, обычно используются при исследовании действующих промышленных объектов, тогда как затраты на эксперимент имеет смысл оценивать в любых исследованиях, в том числе и лабораторных. Если цена опытов одинакова (см. «Ограничения»), затраты на эксперимент пропорциональны числу опытов, которые необходимо поставить для решения данной задачи. Это в значительной мере определяет выбор плана эксперимента.


ПАРАМЕТРЫ ОПТИМИЗАЦИИ
Прочие
Статистические
Психологические
Эстетические
Технико-технологические
Выход продукта
Долговечность
Физические характеристики продукта
Технико-экономические
Надежность
Производительность
Рентабельность
Коэффициент полезного действия
Экономические
Себестоимость
Стабильность
Прибыль
Затраты на эксперимент

Среди технико-экономических параметров наибольшее распространение имеет производительность. Такие параметры, как долговечность, надежность и стабильность, связаны с длительными наблюдениями. Имеется некоторый опыт их использования при изучении дорогостоящих ответственных объектов, например
радиоэлектронной аппаратуры.

Почти во всех исследованиях приходится учитывать количество и качество получаемого продукта. Как меру количества продукта используют выход, например, процент выхода химической реакции, выход годных изделий.

Показатели качества чрезвычайно разнообразны. В нашей схеме они сгруппированы по видам свойств. Характеристики количества и качества продукта образуют группу технико-технологических параметров.

Под рубрикой «прочие» сгруппированы различные параметры, которые реже встречаются, но не являются менее важными. Сюда попали статистические параметры, используемые для улучшения характеристик случайных величин или случайных функций. В качестве примеров назовем задачи на минимизацию дисперсии случайной величины, на уменьшение числа выбросов случайного процесса за фиксированный уровень и т. д. Последняя задача возникает, в частности, при выборе оптимальных настроек автоматических регуляторов или при улучшении свойств нитей (проволока, пряжа, искусственное волокно и др.).

С ростом сложности объекта возрастает роль психологических аспектов взаимодействия человека или животного с объектом. Так, при выборе оптимальной организации рабочего места оператора параметром оптимизации может служить число ошибочных действий в различных возможных ситуациях. Сюда относятся задачи выработки условных рефлексов типа задачи «крысы в лабиринте».

При решении задачи технической эстетики или сравнении произведений искусства возникает потребность в эстетических параметрах. Они основаны на ранговом подходе, который будет рассмотрен ниже.

Таковы некоторые виды параметров оптимизации.

Давайте рассмотрим следующий пример [I].

Пример 1. Во время второй мировой войны несколько сот английских торговых судов на Средиземном море были вооружены зенитными орудиями для защиты от вражеских бомбардировщиков. Поскольку это мероприятие было достаточно дорогим (требовалось иметь на каждом судне боевую команду), через несколько месяцев решили оценить его эффективность. Какой из параметров оптимизации более подходит для этой цели?

Число сбитых самолетов.

Потери в судах, оснащенных орудиями, по сравнению с судами без орудий.

Если Вы считаете, что эффективность установления орудий на торговые суда можно оценить числом сбитых самолетов, то Вы вряд ли смогли бы занять пост командующего английским флотом на Средиземном море. Выбранный Вами параметр оптимизации оценивает эффективность уничтожения самолетов. В то же время ясно, что значения параметра оптимизации в этом случае будут низкими, так как существуют куда более эффективные средства для этой цели (авиация, боевой флот), чем зенитные орудия на торговых судах.

Если же Вы полагаете, что эффективность установки орудий на торговые суда можно оценить сопоставлением потерь в судах, оснащенных орудиями, с потерями в судах без орудий, то это разумный выбор параметра оптимизации, потому что основной задачей при установке орудий была защита судов. Самолеты вынуждены были теперь использовать противозенитные маневры и бомбометание с большой высоты, что .уменьшало потери. Из числа атакованных самолетами торговых судов с зенитными орудиями было потоплено 10% судов, а потери в судах без орудий составили 25%. Затраты на установку орудий и содержание боевых расчетов окупились очень быстро.

2.2. Требования к параметру оптимизации

Параметр оптимизации — это признак, по которому мы хотим оптимизировать процесс. Он должен быть *количественным,* задаваться числом. Мы должны уметь его измерять при любой возможной

комбинации выбранных уровней факторов. Множество значений, которые может принимать параметр оптимизации, будем называть областью его определения. Области определения могут быть непрерывными и дискретными, ограниченными и неограниченными. Например, выход реакции — это параметр оптимизации с непрерывной ограниченной областью определения. Он может изменяться в интервале от 0 до 100%. Число бракованных изделий, число зерен на шлифе сплава, число кровяных телец в пробе крови — вот примеры параметров с дискретной областью определения, ограниченной снизу.

Уметь измерять параметр оптимизации — это значит располагать подходящим прибором. В ряде случаев такого прибора может по существовать или он слишком дорог. Если нет способа количественного измерения результата, то приходится воспользоваться приемом, называемым ранжированием (ранговым подходом). При этом параметрам оптимизации присваиваются оценки — ранги по заранее выбранной шкале: двухбалльной, пятибалльной и т. д. Ранговый параметр имеет дискретную ограниченную область определения. В простейшем случае область содержит два значения (да, нет; хорошо, плохо). Это может соответствовать, например, годной продукции и браку.

Ранг — это количественная оценка параметра оптимизации, но она носит условный (субъективный) характер. Мы ставим в соответствие качественному признаку некоторое число — ранг.

Для каждого физически измеряемого параметра оптимизации можно построить ранговый аналог. Потребность в построении такого аналога возникает, если имеющиеся в распоряжении исследователя численные характеристики неточны или неизвестен способ построения удовлетворительных численных оценок. При прочих равных условиях всегда нужно отдавать предпочтение физическому измерению, так как ранговый подход менее чувствителен и с его помощью трудно изучать тонкие эффекты.

Пример 2. Ваша жена решила испечь яблочный пирог по новому рецепту (аналогичный пример рассмотрен в литературе [4]). Вам, конечно, трудно остаться в стороне, и вы предлагаете ей свои услуги по оптимизации этого процесса. Цель процесса — получение вкусного пирога, но такая формулировка цели еще не дает возможности приступить к оптимизации:
необходимо выбрать количественный критерий, характеризующий степень достижения цели. Можно принять следующее решение: очень вкусный пирог получает отметку 5, просто вкусный пирог — отметку 4 и т. д.

Как вы полагаете, можно ли после такого решения переходить к оптимизации процесса?

Давайте разберемся. Нам важно количественно оценить результат оптимизации. Решает ли отметка эту задачу? Конечно, потому что, как мы договорились, отметка 5 соответствует очень вкусному пирогу и т. д. Другое дело, что этот подход, называемый ранговым, часто оказывается грубым, нечувствительным. Но возможность такой" количественной оценки результатов не должна вызывать сомнений.

Другие примеры рангового подхода: определение чемпиона мира по фигурному катанию или гимнастике, дегустация вин, сравнение произведений искусства и т. д. Или, если хотите, из области химии: сравнение продуктов по цвету, прозрачности, форме кристаллов.

Следующее требование: параметр оптимизации должен выражаться *одним числом.* Иногда это получается естественно, как регистрация показания прибора. Например, скорость движения машины определяется числом на спидометре. Чаще приходится производить некоторые вычисления. Так бывает при расчете выхода реакции. В химии часто требуется получать продукт с заданным отношением компонентов, например, *A : B=3 : 2.* Один из возможных вариантов решения подобных задач состоит в том, чтобы выразить отношение одним числом (1,5) и в качестве параметра оптимизации пользоваться значениями отклонений (или квадратов отклонений) от этого числа.
Еще одно требование, связанное с количественной природой параметра оптимизации, — *однозначность* в статистическом смысле. Заданному набору значений факторов должно соответствовать одно с точностью до ошибки эксперимента значение параметра оптимизации. (Однако обратное неверно: одному и тому же значению параметра могут соответствовать разные наборы значений факторов.)

Для успешного достижения цели исследования необходимо, чтобы параметр оптимизации действительно оценивал эффективность функционирования системы в заранее выбранном смысле. Это требование является главным, определяющим корректность постановки задачи. «Если мы требуем победы и не знаем, что подразумеваем под этим, мы встретимся с призраком, стучащимся к нам в дверь» [5].

Представление об эффективности не остается постоянным в ходе исследования. Оно меняется по мере накопления информации и в зависимости от достигнутых результатов. Это приводит к последовательному подходу при выборе параметра оптимизации. Так, например, на первых стадиях исследования технологических процессов в качестве параметра оптимизации часто используется выход продукта. Однако в дальнейшем, когда возможность повышения выхода исчерпана, нас начинают интересовать такие параметры, как себестоимость, чистота продукта и т. д.

Говоря об оценке эффективности функционирования системы, важно помнить, что речь идет о системе в целом. Часто система состоит из ряда подсистем, каждая из которых может оцениваться своим локальным параметром оптимизации. При этом оптимальность каждой из подсистем по своему параметру оптимизации «не исключает возможности гибели системы в целом» [б].

Пример 3. При флотации сульфидной руды в лабораторных условиях изучалась эффективность применения нового реагента-пенообразователя по схеме рис. 3 [7]. В качестве параметра оптимизации выбрано извлечение (при заданном качестве) концентрата в основной флотации. После проведения эксперимента выяснилось, что реагент дает более высокий выход концентрата по сравнению с прежним пенообразователем. Как вы считаете, обоснованно ли выбран параметр оптимизации, если ставилась задача оптимизации всего процесса флотационного обогащения руды?

Параметр оптимизации — извлечение концентрата в основной флотации — при решении задачи оптимизации всего процесса флотационного обогащения руды выбран не совсем обоснованно. Это правильный ответ, потому что существенно достижение конечной цели — получение готового концентрата (после перечистной флотации), а выбранный параметр оптимизации характеризует эффективность достижения промежуточной цели. Промежуточная цель — повышение выхода концентрата после основной флотации — была достигнута, но при промышленных испытаниях снизились показатели контрольной флотации, что привело к снижению извлечения и качества концентрата по всему циклу. Параметр оптимизации оказался неэффективным с точки зрения достижения конечной цели.

На эту сторону параметра оптимизации обращается внимание в книге Бира [6]: «Отличительной особенностью любой кибернетической системы можно считать полную бессмысленность рассмотрения ее иначе, как единого организма».

Мало иметь эффективный параметр оптимизации. Надо еще, чтобы он был эффективный в статистическом смысле. Понятие статистической эффективности достаточно сложное, и мы не будем здесь заниматься точными формулировками. Фактически это требование сводится к выбору параметра оптимизации, который

определяется с наибольшей возможной точностью. (Если и эта точность недостаточна, тогда приходится обращаться к увеличению числа повторных опытов.)

Пусть, например, нас интересует исследование прочностных характеристик некоторого сплава. В качестве меры прочности можно использовать как прочность на разрыв, так и макротвердость. Поскольку эти характеристики функционально связаны, то с точки зрения эффективности они эквивалентны. Однако точность измерения первой характеристики существенно выше, чем второй. Требование статистической эффективности заставляет отдать предпочтение прочности на разрыв.

Следующее требование к параметру оптимизации — требование *универсальности* или *полноты*. Под универсальностью параметра оптимизации понимается его способность всесторонне характеризовать объект. В частности, технологические параметры оптимизации недостаточно универсальны: они не учитывают экономику. Универсальностью обладают, например, обобщенные параметры оптимизации, которые строятся как функции от нескольких частных параметров [3].

Пример выбора параметра оптимизации, обладающего полнотой, рассмотрен в работе [8] для процессов зонной перекристаллизации. Обычно применяемый для этой цели коэффициент распре-

деления, представляющий отношение концентраций примесей в твердой и жидкой фазах, излишне специфичен. Предложен более полный параметр оптимизации — энтропийная функция *S*

$$S = -\sum_{i=1}^{m}\sum_{j=1}^{n} c_{ij} \log c_{ij}$$ где с,

. — концентрация г-й примеси (при их числе *m)* в у-м участке слитка (при их числе *n).*

Желательно, чтобы параметр оптимизации имел *физический смысл, был простым и легко вычисляемым.*

Требование физического смысла связано с последующей интерпретацией результатов эксперимента. Не представляет труда объяснить, что значит максимум извлечения, максимум содержания ценного компонента. Эти и подобные им технологические параметры оптимизации имеют ясный физический смысл, но иногда для них может не выполняться, например, требование статистической эффективности. Тогда рекомендуется переходить к преобразованию параметра оптимизации. Преобразование, например типа arcsin $\sqrt{y}$, может сделать параметр оптимизации статистически эффективным (например, дисперсии становятся однородными), но остается неясным: что же значит достигнуть экстремума

этой величины?

Второе требование часто также оказывается весьма существенным. Для процессов разделения термодинамические параметры

оптимизации более универсальны. Однако на практике ими пользуются мало: их расчет довольно труден.

Пожалуй, из этих двух требований первое является более существенным, потому что часто удается найти идеальную характеристику системы и сравнить ее с реальной характеристикой. Иногда при этом целесообразно нормировать параметр с тем, чтобы он принимал значения от нуля до единицы.

Кроме высказанных требований и пожеланий при выборе параметра оптимизации нужно еще иметь в виду, что параметр оптимизации в некоторой степени оказывает влияние на вид математической модели исследуемого объекта. Экономические параметры, в силу их аддитивной природы, легче представляются простыми функциями, чем физико-химические показатели. Не случайно методы линейного программирования, основанные на простых моделях, получили широкое распространение именно в экономике. Температура плавления сплава является, как известно, сложной, многоэкстремальной характеристикой состава, тогда как стоимость сплава зависит от состава линейно.

Итак, вы наверное уже поняли, что найти параметр оптимизации, удовлетворяющий всем требованиям, все равно, что поймать жар-птицу.

Факторное пространство.

Фактором называется измеряемая переменная величина, принимающая в некоторый момент времени определенное значение. Факторы соответствуют способам воздействия на объект исследования. Каждый фактор имеет область определения. Фактор считается заданным, если вместе с его названием указан область его определения, то есть все значения, которые может принимать данный фактор. Факторы могут быть количественными, то есть их можно измерить, взвесить или качественными – это разные вещества, аппараты и т.д.

При планировании эксперимента факторы должны быть **управляемыми**. Экспериментатор, выбрав нужное значение фактора, должен иметь возможность поддерживать его постоянным в течение всего опыта, т.е. управлять фактором. В этом состоит особенность активного эксперимента. Если влиять на уровни факторов невозможно (сезонные, температурные изменения), то проводится пассивный эксперимент, так как такими факторами управлять невозможно, за ними можно только наблюдать.

Чтобы точно определить фактор, нужно указать последовательность действий, с помощью которых устанавливается конкретное его значение. Такое определение называется **операциональным**. Введение операционального определения обеспечивает однозначное понимание фактора. Например, если фактором является температура системы, то необходимо указать в какой точке, каким прибором и с какой точностью необходимо проводить измерения и как устанавливать измерительную аппаратуру. Точность замера факторов должна быть по возможность максимальной.

Факторы должны быть **однозначными**, так как трудно управлять фактором, который является функцией других факторов.

При планировании эксперимента обычно одновременно измеряется несколько факторов, то есть используется совокупность факторов. К совокупности факторов предъявляются следующие требования.

Совместность факторов – все комбинации осуществимы и безопасны (разрушение объекта испытания – очень далеко от оптимизации).

Независимость факторов – установка фактора на любом уровне вне зависимости от уровня других факторов. Это значит, что между факторами должна отсутствовать корреляция. Требование некоррелированности не означает, что между факторами нет никакой связи. Достаточно, чтобы эта связь не была линейной.

При активном планировании эксперимента уровни факторов выбираются симметрично относительно центра плана. Совокупность уровней факторов и образует собой факторное пространство. Факторное пространство иногда называется решеткой планирования.

Число точек плана при ортогональном планировании определяется по формуле: $N = K^n$, где K - число уровней для каждого из факторов, n - число факторов. Так для двухфакторного эксперимента на двух уровнях число точек плана $N = 2^2 = 4$; для трех факторного на двух уровнях $N = 2^3 = 8$; для восьми факторного $N = 2^8 = 256$. То есть по мере увеличения числа точек план быстро увеличивается.

Каждой точке факторного пространства отвечает истинное и опытное значение функции отклика $У_{опытн}$. Совокупность всех истинных значений функции отклика, отвечающих точкам факторного пространства, называется поверхностью функции отклика. ***Основные обозначения:***

X_i ~ i-й фактор в натуральном виде; x_i ~ i-й фактор в кодированном виде, величина безразмерная; +1~ верхняя граница (верхний уровень) изменения i- го фактора x_i -1~ нижняя граница (нижний уровень) изменения i- го фактора x_i X_i^0 середина выбранного диапазона измерений i- го фактора (основной уровень); λ_i~ шаг изменения (варьирования) i- го фактора в процессе экспериментирования; Y ~ функция отклика; N~ количество опытов в плане; n~ количество варьируемых факторов (переменных); γ~ количество параллельных замеров в каждом опыте; f~ число степеней свободы той или иной величины; α ~ уровень значимость (вероятность ошибки принятия той или иной гипотезы), $\alpha = 0{,}05$ для $P = 0{,}95$.

Основной (нулевой) уровень является исходной точкой для построения плана эксперимента, обычно это оптимальное (наилучшее значение параметра). Если это значение неизвестно, то выбирается одно из области значений, возможно среднее значение. **Шаг варьирования факторов** – некоторое число (свое для каждого фактора), прибавление которого к основному уровню дает верхний, а вычитание – нижний уровни факторов. В качестве верхнего

уровня может быть выбрано максимальное значение фактора, в качестве нижнего минимальное, а в качестве основного – среднее значение.

Совокупность уровней факторов с отвечающими им опытными значениями функции отклика, записанные в таблицу, называют матрицей планирования. Для ее составления необходимо задать границы уровни и шаг варьирования независимых переменных в натуральном и кодированном виде.

Факторы / Уровни	X_1		…	X_n	
	Физ. Значение	код		Физ. значение	код
Верхний	$X_1^в$	+1	…	$X_n^в$	+1
Нижний	$X_1^н$	−1	…	$X_n^н$	−1
Основной X_i^0	$\frac{X_1^в + X_1^н}{2}$	0	…	$\frac{X_n^в + X_n^н}{2}$	0
Шаг варьирования λ	$\frac{X_1^в - X_1^н}{2}$		…	$\frac{X_n^в - X_n^н}{2}$	

Например, для двухфакторного эксперимента на двух уровнях матрица планирования представляется в виде таблицы, где $X_1^н$ и $X_1^в$ нижний и верхний уровни первого фактора, $X_2^н$ и $X_2^в$ нижний и верхний уровни второго фактора.

№ опыта	Уровни факторов		Функция отклика		№ опыта	Уровни факторов		Кодовое обозначение	Функция отклика
	X_1	X_2				X_1	X_2		
1	$X_1^н$	$X_2^н$	Y_1		1	−1	−1	1	Y_1
2	$X_1^в$	$X_2^н$	Y_2		2	−1	−1	A	Y_2
3	$X_1^н$	$X_2^в$	Y_3		3	−1	+1	B	Y_3
4	$X_1^в$	$X_2^в$	Y_4		4	+1	+1	AB	Y_4

Для упрощения записи 1 может быть опущена. Все возможные комбинации из 2-х факторов, варьируемых на 2-х уровнях, будут исчерпаны, если мы поставим 4 опыта:

№ опыта	Матрица планирования X				Вектор выхода $Y_n = (\bar{y}_n)$
	x_0	x_1	x_2	x_1x_2	
1	+	−	−	+	Y_1
2	+	+	−	−	Y_2
3	+	−	+	−	Y_3
4	+	+	+	+	Y_4

Такое планирование позволяет представить изучаемую зависимость неполным квадратным уравнением: $Y = b_0 + b_1x_1 + b_2x_2 + b_{12}x_1x_2$

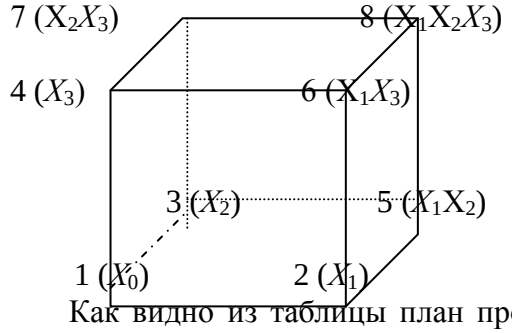
Переход от натуральных значений переменных к кодированным и наоборот осуществляется по формуле: $x_i = \frac{X_i - X_i^0}{\lambda_i}$

При увеличении числа факторов X_i например, до 3-х, чтобы исчерпать все возможные комбинации факторов, варьируемых на 2-х уровнях, необходимо поставить 8 опытов:

№ опыта	Матрица планирования X								Вектор выхода
	x_0	X_1	x_2	x_3	x_1x_2	x_1x_3	x_2x_3	$X_1x_2x_3$	
1	+	−	−	−	+	+	+	−	Y_1
2	+	+	−	−	−	−	+	+	Y_2
3	+	−	+	−	−	+	−	+	Y_3

4	+	+	+		+				Y_4
5	+			+	+			+	Y_5
6	+	+		+		+			Y_6
7	+			+			+		Y_7
8	+	+	+	+	+	+	+	+	Y_8

Факторное пространство можно представить графически:



Как видно из таблицы план проведения экспериментов для трех переменных получили из плана двух переменных, повторив его дважды: один раз при значениях x_3, находящихся на нижнем, а второй раз на верхнем уровнях. Если можно включить в рассмотрение четвертый фактор x_4, то аналогичным образом повторяют планирование и т.д. для сколько угодно числа независимых переменных. После составления матрицы планирования проводится сам эксперимент и его анализ.

1. Дисперсионный анализ результатов эксперимента.

а) проверка измерений на равноточность. Равноточность измерений или однородность полученного ряда оценок дисперсии $S^2_{y1};\ S^2_{y2} \ldots\ S^2_{yN}$, определяется по критерию Кохрена.

$$S^2_{yn} = \frac{\sum(\bar{y}_{ni} - y_{ni})^2}{\gamma - 1}; \bar{y}_{ni} = \frac{\sum y_{ni}}{\gamma}; G^{Э} = \frac{\max\{S^2_{yn}\}}{\sum S^2_{yn}}$$

$G^{Э}$ сравнивается с табличным значением $G^T_{f_1 f_2}$, если при выбранном уровне значимости α имеет место неравенство $G^{Э} \le G^T_{f_1 f_2}$, то измерения во всех γ опытах равноточны, если же $G^{Э} > G^T_{f_1 f_2}$, то измерения при выбранном уровне значимости не равноточны и дальнейшие расчеты продолжать не имеет смысла. Тогда необходимо опыт, оценка дисперсии которого наибольшая, повторить более точно и скорректировать исходные данные, либо увеличить число параллельных измерений γ в каждом опыте.

Число степеней свободы критерия Кохрена определяется по формулам: f_1 число степеней свободы числителя $f_1 = \gamma$, f_2 число степеней свободы знаменателя $f_2 = N$.

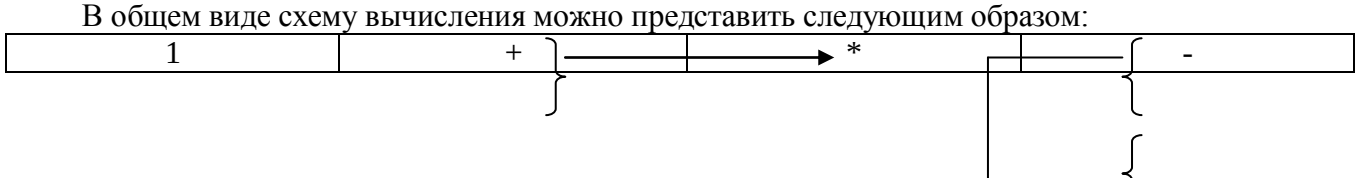
б) определение ошибки эксперимента. Ошибка вычисляется по формуле:

$$S_{\bar{y}} = \sqrt{\frac{\sum S^2_{yn}}{N \cdot \gamma}}$$

в) определение коэффициентов уравнения регрессии. Коэффициенты уравнения регрессии определяются по следующим формулам: $b_i = \frac{\sum a_i \bar{y}_n}{N}; b_{ij} = \frac{\sum a_{ij} \bar{y}_n}{N}; b_{ijk} = \frac{\sum a_{ijk} \bar{y}_n}{N}$ где a_i и $a_{ij} = \pm 1$ берутся из матриц планирования эксперимента.

Вычислять коэффициенты регрессии можно по методу, предложенному Йетсом. Процедура состоит в попарном сложении и вычитании компонент вектора-выхода. Число циклов равно числу факторов, участвующих в эксперименте.

В общем виде схему вычисления можно представить следующим образом:



2	+	*	+
3	+	*	-
4	+	*	+
5	+	*	-
6	+	*	+
7	+	*	-
8	+	*	+

Рассмотрим правила вычисления для полного факторного эксперимента ПФЭ типа 2^2 и 2^3.

№ опыта	Вектор выхода	Цикл 1	Цикл 2	Коэффициенты b
1	$\bar{y}_1$	$\bar{y}_1 + \bar{y}_2 = \bar{y}_1'$	$\bar{y}_1' + y_2' = y_1''$	$b_0 = \frac{1}{2^2} y_1''$
2	$\bar{y}_2$	$\bar{y}_3 + \bar{y}_4 = \bar{y}_2'$	$\bar{y}_3' + y_4' = y_2''$	$b_1 = \frac{1}{2^2} y_2''$
3	$\bar{y}_3$	$\bar{y}_2 - \bar{y}_1 = \bar{y}_3'$	$\bar{y}_2' - y_1' = y_3''$	$b_2 = \frac{1}{2^2} y_3''$
4	$\bar{y}_4$	$\bar{y}_4 - \bar{y}_3 = \bar{y}_4'$	$\bar{y}_4' - y_3' = y_4''$	$b_{12} = \frac{1}{2^2} y_4''$

Схема вычисления коэффициентов регрессии для ПФЭ типа 2^3 методом Йетса.

№ опыта	Вектор выхода	Цикл 1	Цикл 2	Цикл 3	Коэффициенты b
1	$\bar{y}_1$	$\bar{y}_1 + \bar{y}_2 = \bar{y}_1'$	$\bar{y}_1' + y_2' = y_1''$	$\bar{y}_1'' + y_2'' = y_1'''$	$b_0 = \frac{1}{2^3} y_1'''$
2	$\bar{y}_2$	$\bar{y}_3 + \bar{y}_4 = \bar{y}_2'$	$\bar{y}_3' + y_4' = y_2''$	$\bar{y}_3'' + y_4'' = y_2'''$	$b_1 = \frac{1}{2^3} y_2'''$
3	$\bar{y}_3$	$\bar{y}_5 + \bar{y}_6 = \bar{y}_3'$	$\bar{y}_5' + y_6' = y_3''$	$\bar{y}_5'' + y_6'' = y_3'''$	$b_2 = \frac{1}{2^3} y_3'''$
4	$\bar{y}_4$	$\bar{y}_7 + \bar{y}_8 = \bar{y}_4'$	$\bar{y}_7' + y_8' = y_4''$	$\bar{y}_7'' + y_8'' = y_4'''$	$b_{12} = \frac{1}{2^3} y_4'''$
5	$\bar{y}_5$	$\bar{y}_2 - \bar{y}_1 = \bar{y}_5'$	$\bar{y}_2' - y_1' = y_5''$	$\bar{y}_2'' - y_1'' = y_5'''$	$b_3 = \frac{1}{2^3} y_5'''$
6	$\bar{y}_6$	$\bar{y}_4 - \bar{y}_3 = \bar{y}_6'$	$\bar{y}_4' - y_3' = y_6''$	$\bar{y}_4'' - y_3'' = y_6'''$	$b_{13} = \frac{1}{2^3} y_6'''$
7	$\bar{y}_7$	$\bar{y}_6 - \bar{y}_5 = \bar{y}_7'$	$\bar{y}_6' - \bar{y}_5' = \bar{y}_7''$	$\bar{y}_6'' - \bar{y}_5'' = \bar{y}_7'''$	$b_{23} = \frac{1}{2^3} y_7'''$
8	$\bar{y}_8$	$\bar{y}_8 - \bar{y}_7 = \bar{y}_8'$	$\bar{y}_8' - \bar{y}_7' = \bar{y}_8''$	$\bar{y}_8'' - \bar{y}_7'' = \bar{y}_8'''$	$b_{123} = \frac{1}{2^3} y_8'''$

2. Дисперсионный анализ уравнения регрессии:

а) определение ошибки при вычислении коэффициентов уравнения регрессии.

Ошибки при вычислении коэффициентов для ПФЭ типа 2^n одинаковы и определяются по формуле: $S^2 = \{b_n\} = S_{\bar{y}}^2 \frac{1}{N}; или \;\; S = \{b_n\} = S_{\bar{y}} \frac{1}{\sqrt{N}};$

б) проверка значимости коэффициентов уравнения регрессии. Проверку значимости коэффициентов уравнения регрессии проводят по $\tilde{t}$ критерию Стьюдента. Для каждого

коэффициента определяют $t^{Э}$ критерий по формуле: $t_n^{Э} = \frac{|b_n|\sqrt{N}}{S_{\bar{y}}}$ и сравнивают с табличным значением $t_{\alpha,f}^{T}$ выбранного для заданного уровня значимости и числа степеней свободы, которое равно: $f = N(\gamma - 1)$.

Если для данного коэффициента выполняется неравенство $t^{Э} \geq t_{\alpha,f}^{T}$,то при выбранном уровне значимости соответствующий коэффициент уравнения регрессии значим, если неравенство не выполняется, то коэффициент не значим и членом уравнения регрессии с этим коэффициентом можно пренебречь.

в) проверка гипотезы об адекватности полученного уравнения действительному процессу можно рассчитать двумя способами используя F критерий Фишера:

~ по формуле вычисляют экспериментальное значение статистики:

$$F^{Э} = \frac{S_{yag}^2}{S_{\bar{y}}^2}; S_{yag}^2 = \frac{\sum \bar{y}_n^2 - N\sum b_n}{N-(k+1)}; F_{f_1,f_2}^{T}$$

где k число значимых коэффициентов регрессии, не считая свободного члена, и сравнивают с табличным значением, где f_1~ число степеней свободы числителя отношения, f_1 □N~ $(k+1)$; f_2~ число степеней свободы знаменателя отношения, f_2 □$N(\gamma$~ ı), γ - число параллельных измерений.

Если при выбранном уровне значимости выполняется неравенство $F^{Э} \leq F_{f_1,f_2}^{T}$ то полученное уравнение регрессии адекватно действительному уравнению, если неравенство не выполняется, то полученное уравнение регрессии неадекватно описывает полученный процесс;

~ опытное значение критерия Фишера может быть взятым равным отношению дисперсии неадекватности к общей дисперсии или дисперсии воспроизводимости:

$$F_{опыт} = \frac{S^2 Y_{н.а.}}{S^2 Y_{воспр.}}; \quad S^2 Y_{н.а.} = \frac{\gamma \cdot S^2_{ОСТ}}{N-k} = \frac{\gamma \cdot \sum_{i=1}^{N} (Y_{i\text{рас}} - \bar{Y}_{i\text{опыт.}})^2}{N-k};$$

$$S^2 Y_{воспр.} = \frac{\sum_{i=1}^{N}\sum_{j=1}^{k} (Y_{ij\text{рас.}} - Y_{i\text{опыт.}})^2}{N(k-1)}$$

г) определение границ доверительных интервалов для коэффициентов регрессии. Границы доверительных интервалов, в которых с заданной вероятностью P = 1~ α находятся коэффициенты уравнения регрессии $\{\beta\}$ определяются по формуле: $b_n - t_{\alpha,f}^{T} S\{b_n\} \leq \beta_n \geq b_n + t_{\alpha,f}^{T} S\{b_n\}$ где $t_{\alpha,f}^{T}$ табличное значение t - критерия Стьюдента при заданном уровне значимости и числе степеней свободы f □$N(\gamma$~ ı).

Рассмотрим пример решения технологической задачи с помощью ПФЭ типа 2^n. Необходимо провести исследование уравнения регрессии: $Y = R_{изд} = f(X_1, X_2, X_3)$

В рассмотренном примере были учтены все возможные взаимодействия факторов для полного факторного эксперимента 2^3. Столбцы в матрице планирования иногда называют эффектами. Эффект взаимодействия двух факторов называется эффектом взаимодействия первого порядка (X_1X_2), а эффект взаимодействия трех факторов называется эффектом взаимодействия второго порядка ($X_1X_2X_3$). Вообще эффект взаимодействия максимального порядка в полном факторном эксперименте имеет порядок на единицу меньше числа факторов. В литературе часто встречаются названия парных эффектов (X_1X_2) и тройных эффектов ($X_1X_2X_3$). Полное число всех возможны эффектов, включая b_0, линейные эффекты и взаимодействия всех порядков, равно числу

опытов полного факторного эксперимента. Чтобы найти число всех возможных взаимодействий некоторого порядка, можно воспользоваться формулой числа сочетаний:

$C_m^k = \frac{k!}{m!(k-m)!}$ где k – число факторов, а m – число элементов во взаимодействии.

Например, для планов 2^3 и 2^4 число парных взаимодействий:

$$C_3^2 = \frac{3!}{2!(3-2)!} = 3; C_4^2 = \frac{4!}{2!(4-2)!} = 6;$$

Матрицы планирования независимо от числа факторов обладают общими свойствами. Эти свойства определяют качества модели, так как эксперимент планируется для того, чтобы получить модель, обладающую некоторыми оптимальными свойствами. На практике это означает, что оценки коэффициентов модели должны быть наилучшими.

Два свойства следуют непосредственно из построения матрицы и являются свойствами отдельных столбцов матрицы планирования. Первое из них – симметричность относительно центра эксперимента формулируется следующим образом: алгебраическая сумма элементов вектор-столбца каждого фактора равна нулю, или:

$\sum_{i,j,k=1}^{N} x_{ijk} = 0$ где I, j, k – номер фактора, а N – число факторов.

Второе свойство – так называемое условие нормировки – формулируется следующим образом: сумма квадратов элементов каждого столбца равна числу опытов:

$\sum_{i,j,k=1}^{N} x_{ijk}^2 = 0$. Это следует из того, что значения факторов в матрице = +,- 1.

Два следующих свойства относятся к совокупности столбцов.

Третье свойство: сумма почленных произведений двух вектор столбцов матрицы равна нулю, это свойство называется ортогональностью матрицы планирования:

$$\sum_{i=1}^{N} x_{ij} x_{uk} = 0.$$

Последние, четвертое свойство называется ротатабельностью, т.е. точки в матрице планирования подбираются так, что точность предсказаний значений параметров оптимизации одинакова на равных расстояниях от центра эксперимента и не зависит от направления

ЛИТЕРАТУРА

1. Ю.П. Адлер, Е.В. Маркова, Ю.В. Грановский «Планирование эксперимента при поиске оптимальных решений», Наука, 1976 г., 279 с. 22.18./А31
2. С.М. Ермаков «Математическая теория планирования эксперимента», М.: Наука, 1983 г., 390 с. 22.18/М34
3. В.И. Плескунин, Е.Д. Воронина «Теоретические основы организации и анализа выборочных данных в эксперименте», Издательства Ленинградского университета, Ленинград, 1979 г., 232 с. 22.18/П38
4. А.П. Кондрашов, Е.В. Шестопалов «Основы физического эксперимента и математическая обработка результатов измерений», М.: Атомиздат, 1977 г., 200с.
5. В.Г. Гродский, Ю.П. Адлер «Планирование промышленных экспериментов», М.: Металлургия, 1974 г., 264 с.
6. Ю.В. Завадский «Планирование эксперимента в задачах автомобильного транспорта», М.: МАДИ 1978 г., 156 с.

Вопросы по курсу:

«Планирование и организация эксперимента»

1. История появления планирования эксперимента.
2. Общие сведения о математической теории планирования эксперимента. Научный и промышленный эксперимент.
3. Основные положения математической теории планирования эксперимента.
4. Этапы проведения и анализа эксперимента.
5. Факторное пространство. Требования, предъявляемые к факторам.
6. Факторное пространство. Требования, предъявляемые к совокупности факторов.
7. Математическая модель объекта исследования (черный ящик, функция отклика).
8. Полный факторный эксперимент. Основной уровень, шаг варьирования, матрица планирования.
9. Основные свойства матрицы планирования.
10. Обработка результатов эксперимента.

Дисперсионный анализ результатов эксперимента (оценка равноточности и ошибки эксперимента).

Определение коэффициентов уравнения регрессии.

Дисперсионный анализ уравнения регрессии.

11. Эффекты взаимодействия.
12. Дробно-факторное планирование.
13. Неполные планы. Планы выборочного контроля.
14. Полуреплика 2^{3-1}. Определяющий контраст, эффект смешивания, генерирующее соотношение.
15. Полуреплика 2^{4-1}. Определяющий контраст, эффект смешивания, генерирующее соотношение.
16. Полуреплика 2^{5-1}. Определяющий контраст, эффект смешивания, генерирующее соотношение.
17. ¼ реплика или реплика 2^{5-2}. обобщающий определяющий контраст, эффект смешивания, генерирующее соотношение.
18. Рандомизация.
19. Определение области экстремума. Движение по вектор-градиенту.
20. Ортогональное планирование 2-го порядка. Корректирование квадратичных переменных. Расчет коэффициентов.
21. Определение координат экстремальной точки.
22. Планирование эксперимента с качественными факторами.
23. Обобщенный параметр оптимизации