

Основные понятия и определения планирования эксперимента

ГОСТ 24026–80 задаёт любой **эксперимент** как систему операций, воздействий и (или) наблюдений, направленных на получение информации об объекте при исследовательских испытаниях.

Для нас это метод исследования явления в контролируемых и управляемых (исследователь активно взаимодействует с объектом) условиях.

Обычно проводится в рамках научного исследования и служит для проверки гипотезы, а также установления причинных связей между феноменами.

Включает в себя:

- постановку задачи, которая должна решаться;
- выбор **отклика** y – зависимой переменной (выходной параметр, выход, целевая функция, по ГОСТу недопустимо реакция, параметр). **Пример.** При определении зависимости стойкости режущего инструмента от температуры рабочей зоны и способа подвода СОЖ откликом будет стойкость. Откликов может быть несколько. **Пример.** При разработке оптимальной рецептуры нового пластифицированного полимерного материала качество готовой продукции оценивается семью откликами: термостойкость композиции, блеск, морозостойкость, модуль упругости при заданной температуре, предел прочности при растяжении, относительное удлинение при разрыве, число перегибов до разрушения;
- выбор **фактора** x – независимой переменной (входной параметр, вход, по ГОСТу недопустимо параметр). Факторов может быть несколько. **Пример.** При оптимизации процесса изготовления резистора экспериментатора должны интересовать одиннадцать факторов: давление при прессовке, температура при прессовке, время выдержки под давлением, температура в муфеле при прессовке, время температурной выдержки, дисперсность наполнителя, соотношение флюса и наполнителя, давление при шамотировании, дисперсность сажи, время выдержки при шамотировании, дисперсность флюса;
- выбор **уровня фактора** – его величина будет поддерживаться постоянной (**фиксированный уровень**) или усредняться (**случайный уровень**);
- подбор сочетания уровней факторов.

Планирование эксперимента – это:

- в широком понимании наука о его проведении;
- в узком понимании процедура выбора числа и условий проведения опытов, необходимых и достаточных для решения поставленной задачи с требуемой точностью.

ГОСТ задаёт любой **опыт** как воспроизведение исследуемого явления в определённых условиях проведения эксперимента при возможности регистрации его результатов.

Планирование эксперимента служит для повышения эффективности научного исследования, поскольку любой эксперимент эффективен.

Пример. Во время Второй мировой войны английские торговые суда оснащали зенитными орудиями для защиты от вражеских бомбардировщиков, что требовало больших финансовых затрат. Для оценки эффективности предпринимаемых мер предлагали два критерия: число сбитых бомбардировщиков или потери судов, оснащённых орудиями, по сравнению с судами без орудий. Остановились на втором, поскольку основной задачей была защита судов.

При планировании эксперимента **существенно** следующее:

- стремление к минимизации числа опытов;
- использование математического аппарата, формализующего многие действия экспериментатора;
- выбор чёткой стратегии, позволяющей принимать обоснованные решения по окончании каждого опыта.

При решении поставленной задачи применяется **математическая модель** объекта исследования – уравнение, связывающее отклик с фактором: $y = f(x)$.

Окончательный этап – **анализ**, включающий в себя:

- сбор и обработку данных с упорядочением по логическим критериям;
- вычисление статистик для проверки гипотезы;
- **интерпретацию** (в словесной, графической или табличной форме) результатов для экспериментатора.

Магический и латинский квадраты

Первая мысль о том, что эксперимент можно планировать, предположительно пришла в голову первобытного человека (случайно убедившись, что острым камнем можно убить мамонта, он стал выдвигать дальнейшие гипотезы, которые после экспериментальной проверки привели к созданию копья, а затем и лука со стрелами), однако он не пользовался статистическими методами.

Постепенно **натурный** эксперимент уступил место **числовому** – расположению чисел таким образом, чтобы выполнялись строгие условия:

- суммы по всем строкам, столбцам и диагоналям квадратной таблицы равны – **магический квадрат**, которому принадлежит первенство в планировании эксперимента. Например,

4 9 2

квадрат Ю 3 5 7

8 1 6

16 3 2 13

5 10 11 8

квадрат Дюрера 9 6 7 12

4 15 14 1

В настоящее время магические квадраты применяются в кодировании и декодировании, экономических расчётах и диетологии;

- по ГОСТу каждый символ встречается один раз в каждой строке и в каждом столбце квадратной таблицы – **латинский квадрат**. Например,

A E D C B

C D A B E

квадрат Эйлера B C E A D

E A B D C

D B C E A

В настоящее время латинские квадраты применяются как средство сокращения перебора в комбинаторных задачах оптимизации, поскольку во многих из них нереален **полный перебор** – поиск решения математической задачи путём исчерпывания всех возможных вариантов.

Точные планы и их характеристики

ГОСТ задаёт любой **план эксперимента** как совокупность данных, определяющих число, условия и порядок реализации опытов.

Пусть имеем n измерений, которые можно выполнить в любых точках t замкнутой области Ω_t , причём в каждой точке t_i ($i = 1..n$) можно реализовать несколько чисел повторных измерений r_i . Тогда план эксперимента состоит из t_i (**спектр плана**, который ГОСТ задаёт как совокупность всех точек плана, отличающихся уровнями хотя бы одного фактора) и соответствующих r_i :

$$\left\{ \begin{matrix} t_1, & t_2, & \dots, & t_n \\ r_1, & r_2, & \dots, & r_n \end{matrix} \right\}, \text{ где } \sum_{i=1}^n r_i = n.$$

При фиксированном n получаем **точный план**, но на практике удобнее пользоваться **нормированным точным планом**, состоящим из t_i и соответствующих относительных чисел повторных измерений $p_i = r_i / n$:

$$\varepsilon_n = \left\{ \begin{array}{l} t_1, t_2, \dots, t_n \\ p_1, p_2, \dots, p_n \end{array} \right\}.$$

p_i – рациональные числа, удовлетворяющие **нормировке**

$$\sum_{i=1}^n p_i = 1, \quad p_i \geq 0.$$

В качестве **координат точек плана** выступают только те переменные, значения которых можно варьировать при выборе плана.

Для отображения используется **матрица плана**, которую ГОСТ задаёт как стандартную форму записи условий проведения экспериментов в виде прямоугольной таблицы, строки которой отвечают опытам, столбцы – факторам.

Ключевую роль в планировании эксперимента играет **информационная матрица** M , но на практике удобнее пользоваться **нормированной информационной матрицей**

$$M(\varepsilon_n) = M / n.$$

Рассмотрим её для линейных математических моделей. В общем виде **однооткликовую математическую модель** можно записать как

$$y(t) = x_1(t, \theta) + v(t),$$

где $y(t)$ – отклик;

$x_1(t, \theta)$ – **переменная состояния** (известная с точностью до параметров θ функция, отражающая зависимость от факторов и параметров);
 $v(t)$ – ошибка измерений.

При линейной зависимости переменной состояния $x_1(t, \theta) = f^T(t) \theta$,

где $f(t)$ – известная регрессионная вектор-функция;

T – операция транспонирования.

Нормированная информационная матрица $M(\varepsilon_n) = \sum_{i=1}^n f(t_i) \sigma^{-2}(t_i) p_i f^T(t_i)$,

где $\sigma^2(t_i)$ – дисперсия отклика в точке t_i .

Многооткликовая математическая модель является обобщением однооткликовой на случай нескольких откликов.

При линейной зависимости переменной состояния $x_1(t, \theta) = F^T(t) \theta$,
где $F(t)$ – матрица, составленная из $f(t)$.

Нормированная информационная матрица $M(\varepsilon_n) = \sum_{i=1}^n F(t_i) D^{-1}(t_i) p_i F^T(t_i)$,

где $D(t_i)$ – дисперсионная матрица откликов в точке t_i .

Ещё одно важное понятие – **информационная матрица однократных измерений в точке плана**:

для однооткликовой математической модели $M(t) = f(t) \sigma^{-2}(t) f^T(t)$,

для многооткликовой математической модели $M(t) = F(t) D^{-1}(t) F^T(t)$.

Нормированная информационная матрица при некоррелированных измерениях представляет собой взвешенную сумму $M(t)$:

$$M(\varepsilon_n) = \sum_{i=1}^n p_i M(t_i).$$

Для характеристики точности оценок параметров математической модели на практике удобнее пользоваться **нормированной дисперсионной матрицей оценок параметров** $D(\varepsilon_n, \hat{\theta}) = n D(\hat{\theta}) = M^{-1}(\varepsilon_n)$,

где $\hat{\theta}$ – оценка параметра θ ;

$D(\hat{\theta}) = M^{-1}$ – **дисперсионная матрица оценок параметров** (упорядоченная совокупность дисперсий и ковариаций оценок параметров).

Для характеристики точности прогноза по однооткликовой математической модели на практике удобнее пользоваться **нормированной дисперсией оценки отклика** $d(\varepsilon_n, t) = n d(t) = f^T(t) D(\varepsilon_n, \hat{\theta}) f(t)$,

где $d(t) = \sigma^2[y(t, \theta)] = \sigma^2[\hat{x}_1(t)]$ – дисперсия оценки отклика.

Для характеристики точности прогноза по многооткликовой математической модели на практике удобнее пользоваться **нормированной дисперсионной матрицей оценки отклика** $D(\varepsilon_n, t) = n D(t) = F^T(t) D(\varepsilon_n, \hat{\theta}) F(t)$,

где $D(t) = D[y(t, \hat{\theta})] = D[\hat{x}_1(t)]$ – дисперсионная матрица оценки отклика.

Выражения $d(\varepsilon_n, t)$ и $D(\varepsilon_n, t)$ справедливы и для линейных, и для нелинейных математических моделей.

Обобщённые планы и их характеристики

Нормированный непрерывный план задаётся замкнутой областью Ω_t и вероятностной мерой $P(t)$, определённой на этой области и удовлетворяющей **нормировке**

$$\int_{\Omega_t} dP(t) = 1, \quad P(t) \geq 0, \quad t \in \Omega_t.$$

$P(t)$ может быть сосредоточена и в конечном числе точек, тогда получаем **нормированный дискретный план** и пользуемся **дельта-функцией Дирака** – обобщённой функцией, позволяющей выразить точечное воздействие:

$$P(t) = \sum_i p_i \delta(t - t_i).$$

Кстати, Поль Дирак нашёл способ выразить любое натуральное число всего через три двойки и математические операции: $z = -\log_2 \left(\log_2 \left(\sqrt{\sqrt{\dots \sqrt{2}}} \right) \right)$,

где количество знаков квадратного корня равняется z .

Нормированный дискретный план $\tilde{\epsilon}$ записывается так же, как нормированный точный план ϵ_n , однако здесь снимается требование рациональности p_i .

Непрерывные и дискретные планы вместе создают **обобщённые планы**.

Рассмотрим их характеристики. **Нормированная информационная матрица нормированного непрерывного плана**:

для однооткликовой математической модели $M(\epsilon) = \int_{\Omega_t} f(t) \sigma^{-2}(t) f^T(t) p(t) dt,$

для многооткликовой математической модели $M(\epsilon) = \int_{\Omega_t} F(t) D^{-1}(t) F^T(t) p(t) dt,$

где $p(t) dt = dP(t)$.

Нормированная информационная матрица нормированного дискретного плана вплоть до обозначений совпадает с точным планом.

Для любого нормированного непрерывного плана ϵ всегда найдётся нормированный дискретный план $\tilde{\epsilon}$, чей спектр содержит не более $n_0 = [p(p+1)/2] + 1$ точек и чья нормированная информационная матрица $M(\tilde{\epsilon})$ совпадает с нормированной информационной матрицей $M(\epsilon)$ плана ϵ . Это значит, что **нормиро-**

ванную информационную матрицу любого нормированного непрерывного плана можно выразить как

$$M(\varepsilon) = \sum_{i=1}^n p_i M(t_i), \quad n \leq n_0,$$

где $M(t)$ – информационная матрица однократных измерений в точке плана.

Это свойство имеет большую практическую ценность: всегда можно заменить нормированный непрерывный план, обладающий теми или иными экстремальными показателями информационной матрицы, столь же эффективным нормированным дискретным планом, чей спектр содержит не более n_0 точек.

Критерии оптимальности эксперимента

В зависимости от априорной информации выделяют следующие **критерии**:

- **оптимальности дискриминирующего эксперимента** – направлены на ранжирование и отсеивание гипотез. Используются редко;
- **оптимальности уточняющего эксперимента** – направлены на максимально точное оценивание факторов, переменных состояния и откликов математической модели, форма которой известна с точностью до параметров. Чаще всего выражаются в виде функционала от нормированной дисперсионной матрицы оценок параметров и в свою очередь делятся на:
 - **связанные с точностью оценок параметров.** Рассмотрим главные критерии этой группы:
 - **D-оптимальности** – оптимальный план ε^* удовлетворяет условию

$$|D(\varepsilon^*, \hat{\theta})| = \min_{\varepsilon} |D(\varepsilon, \hat{\theta})|; \quad |M(\varepsilon^*)| = \max_{\varepsilon} |M(\varepsilon)|,$$

где минимизация (максимизация) проводится на допустимом множестве планов. Чёрточками, окаймляющими символ матрицы, обозначается **детерминант (определитель)** – многочлен, комбинирующий элементы квадратной матрицы так, что его значение сохраняется при транспонировании и линейных комбинациях строк или столбцов;

- **усечённой D-оптимальности** – стремятся максимально точно определить не все параметры, а только часть ($p' < p$). В этом случае минимизируют соответствующий минор p' -го порядка матрицы $D(\varepsilon, \hat{\theta})$;
- **A-оптимальности** – минимизируют на допустимом множестве планов след нормированной дисперсионной матрицы оценок параметров

$$sp [D(\varepsilon^*, \hat{\theta})] = sp [M^{-1}(\varepsilon^*)] = \min_{\varepsilon} sp [D(\varepsilon, \hat{\theta})] = \min_{\varepsilon} sp [M^{-1}(\varepsilon)],$$

где $sp [...]$ – **след** (сумма элементов главной диагонали) матрицы;

- **E -оптимальности** – оптимальный план удовлетворяет условию

$$\lambda_{\max}[D(\varepsilon^*, \hat{\theta})] = \min_{\varepsilon} \max_i \lambda_i [D(\varepsilon, \hat{\theta})],$$

где $\lambda[D]$ – **собственное значение** матрицы D (числовой коэффициент собственного вектора матрицы).

Пример. Возьмём произвольную квадратную матрицу $A = \begin{pmatrix} -1 & -6 \\ 2 & 6 \end{pmatrix}$ и умножим её справа на вектор $\bar{a} = \begin{pmatrix} 2 \\ -1 \end{pmatrix}$, тогда

$$A\bar{a} = \begin{pmatrix} -1 & -6 \\ 2 & 6 \end{pmatrix} \begin{pmatrix} 2 \\ -1 \end{pmatrix} = \begin{pmatrix} -1 \cdot 2 + (-6) \cdot (-1) \\ 2 \cdot 2 + 6 \cdot (-1) \end{pmatrix} = \begin{pmatrix} 4 \\ -2 \end{pmatrix} = 2 \begin{pmatrix} 2 \\ -1 \end{pmatrix};$$

- **связанные с точностью оценок откликов.** Рассмотрим главные критерии этой группы:

- **G -оптимальности** – оптимальный план удовлетворяет условию

$$\max_t d(\varepsilon^*, t) = \min_{\varepsilon} \max_t d(\varepsilon, t), \quad t \in \Omega_t,$$

где минимизируют на допустимом множестве планов максимальную по области нормированную дисперсию оценки отклика однооткликовой математической модели. Если же модель многооткликовая, минимизируют максимальный по области детерминант или след нормированной дисперсионной матрицы оценки векторного отклика;

- **Q -оптимальности** – отличается от предыдущего критерия тем, что минимизируют не максимальные, а средние по области величины. Например, для однооткликовой математической модели

$$\int_{\Omega_t} d(\varepsilon^*, t) dt = \min_{\varepsilon} \int_{\Omega_t} d(\varepsilon, t) dt.$$

Теорема эквивалентности

Непосредственное построение планов по критерию D -оптимальности затруднительно, поскольку связано с необходимостью вычислять детерминант и решать экстремальную задачу большой размерности. Это обстоятельство привело к созданию более простых, итерационных по своей природе алгоритмов, в основе которых лежит теорема эквивалентности. **Утверждения:**

- план ε^* максимизирует $|M(\varepsilon)|$;
- план ε^* минимизирует $\max_t \delta(\varepsilon, t)$, $t \in \Omega_t$, $\delta(\varepsilon, t) = \text{sp}[M^{-1}(\varepsilon) M(t)]$;
- $\max_t \delta(\varepsilon^*, t) \equiv p$.

Теорема: нормированные информационные матрицы всех планов, удовлетворяющих этим утверждениям, совпадают между собой.

Следствие: в точках оптимального плана функция $\delta(\varepsilon^*, t)$ достигает максимального значения, равного p . Это условие часто используется для **проверки D-оптимальности плана**, однако следует иметь в виду, что оно является необходимым, но не достаточным.

Теорема эквивалентности в обобщённой форме, данной выше, справедлива как для однооткликовой, так и многооткликовой линейной математической модели, поскольку функция $\delta(\varepsilon, t)$ легко выражается привычными формулами: в первом случае $\delta(\varepsilon, t) = \sigma^{-2}(t) d(\varepsilon, t)$, во втором случае $\delta(\varepsilon, t) = sp [D^{-1}(t) d(\varepsilon, t)]$.

При нелинейной параметризации математической модели теорема эквивалентности тоже справедлива, но нормированная информационная матрица плана и информационная матрица однократных измерений зависят от неизвестных параметров. По этой причине план ε^* при нелинейной параметризации называется **локально D-оптимальным**.

Замечание: теорема эквивалентности строго справедлива только для обобщённых планов, а для точных соблюдается лишь приближённо.

Подготовка к РК1

- Точность оценок параметров, получаемых из экспериментальных данных, зависит от расположения точек в факторном пространстве. Разнесение точек приводит к повышению точности;

- **Соответствие планов.** Пусть точный план задан матрицей $X = \begin{pmatrix} -1 & -1 \\ 1 & -1 \\ -1 & 1 \\ 1 & 1 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}$.

Точки плана $x_1(-1; -1)$, $x_2(1; -1)$, $x_3(-1; 1)$, $x_4(1; 1)$, $x_5 \equiv x_6(0; 0)$,

тогда нормированный непрерывный план $\varepsilon = \left\{ \frac{x_1}{6}, \frac{x_2}{6}, \frac{x_3}{6}, \frac{x_4}{6}, \frac{x_5}{3} \right\}$;

$$\begin{pmatrix} 1 & -1 & -1 \\ 1 & 1 & -1 \\ 1 & -1 & 1 \\ 1 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}.$$

- Вычисление матриц.** Пусть точный план задан матрицей $F =$

Нормированная информационная матрица

$$M(\varepsilon_n) = \frac{1}{N} F^T F = \frac{1}{6} \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ -1 & 1 & -1 & 1 & 0 & 0 \\ -1 & -1 & 1 & 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & -1 & -1 \\ 1 & 1 & -1 \\ 1 & -1 & 1 \\ 1 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2/3 & 0 \\ 0 & 0 & 2/3 \end{pmatrix},$$

нормированная дисперсионная матрица оценок параметров

$$D(\varepsilon_n, \hat{\theta}) = M^{-1}(\varepsilon_n) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 3/2 & 0 \\ 0 & 0 & 3/2 \end{pmatrix}.$$

Непрерывные D -оптимальные планы на отрезке

Рассмотрим случай, когда математическая модель задаётся полиномом произвольной степени m в одномерном пространстве, где областью планирования является отрезок $[-1; 1]$ с фактором $-1 \leq x \leq 1$. Математическая модель имеет вид

$$y(x, \theta) = \sum_{i=0}^m \theta_i x^i.$$

Непрерывные D -оптимальные планы для этого случая построены для любой m :

m	x	D
2	-1	$\begin{pmatrix} 3 & 0 & -3 \end{pmatrix}$
	0	$\begin{pmatrix} 0 & 1,5 & 0 \end{pmatrix}$
	1	$\begin{pmatrix} -3 & 0 & 4,5 \end{pmatrix}$
3	-1	$\begin{pmatrix} 3,248 & 0 & -3,748 & 0 \end{pmatrix}$
	-0,447	$\begin{pmatrix} 0 & 15,757 & 0 & -16,257 \end{pmatrix}$
	0,447	$\begin{pmatrix} -3,748 & 0 & 6,247 & 0 \end{pmatrix}$
	1	$\begin{pmatrix} 0 & -16,257 & 0 & 18,756 \end{pmatrix}$

	-1	$\begin{pmatrix} 5 & 0 & -16,653 & 0 & 11,654 \\ 0 & 19,284 & 0 & -21,163 & 0 \\ -16,653 & 0 & 98,541 & 0 & -83,766 \\ 0 & -21,163 & 0 & 25,541 & 0 \\ 11,654 & 0 & -83,766 & 0 & 76,491 \end{pmatrix}$
	-0,655	
	0	
	0,655	
4	1	

Пример. Для управления процессом синтеза аммиака важно знать максимальную температуру в катализаторной зоне. Она зависит от режима работы колонны, определяемого степенью открытия вентиля, задающей газовый поток, давлением контактного газа, расходом контактного газа, температурой контактного газа и отношением водорода к азоту. Для контроля температуры в катализаторной зоне установим три термопары, тем самым зададим три точки плана, тогда $m = 3 - 1 = 2$. Из таблицы координаты точек плана $x_1 = -1$, $x_2 = 0$, $x_3 = 1$. Математическая модель имеет вид $y(x, \theta)|_{m=2} = \theta_0 + \theta_1 x + \theta_2 x^2$, что позволяет записать **регрессионную матрицу плана** – матрицу значений функций независимых переменных в точках плана:

$$F = \begin{pmatrix} 1 & -1 & 1 \\ 1 & 0 & 0 \\ 1 & 1 & 1 \end{pmatrix}, \text{ соответственно } F^T = \begin{pmatrix} 1 & 1 & 1 \\ -1 & 0 & 1 \\ 1 & 0 & 1 \end{pmatrix}.$$

В каждой из выбранных точек плана проведём два измерения температуры:

i	y_1	y_2	$y_{cp} = \frac{y_1 + y_2}{2}$
1	454,3	451,6	452,95
2	522,3	521,1	521,7
3	463,5	466,6	465,05

Вектор измерений $\bar{y} = \begin{pmatrix} y_{cp1} \\ y_{cp2} \\ y_{cp3} \end{pmatrix}$,

умножим на него матрицу справа: $F^T \bar{y} = \begin{pmatrix} 1 & 1 & 1 \\ -1 & 0 & 1 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 452,95 \\ 521,7 \\ 465,05 \end{pmatrix} = \begin{pmatrix} 1439,7 \\ 12,1 \\ 918 \end{pmatrix}$.

Вектор оценок параметров $\bar{\theta} = \frac{1}{3} D F^T \bar{y} = \frac{1}{3} \begin{pmatrix} 3 & 0 & -3 \\ 0 & 1,5 & 0 \\ -3 & 0 & 4,5 \end{pmatrix} \begin{pmatrix} 1439,7 \\ 12,1 \\ 918 \end{pmatrix} = \begin{pmatrix} 521,7 \\ 6,05 \\ -62,7 \end{pmatrix}$,

тогда уравнение математической модели $y(x, \hat{\theta}) = 521,7 + 6,05x - 62,7x^2$.

Непрерывные D -оптимальные планы на гиперкубе

Рассмотрим случай, когда математическая модель имеет вид

$y(x, \theta) = \theta_0 + \theta_1 x_1 + \cdots + \theta_n x_n + \theta_{n+1} x_1^2 + \cdots + \theta_{2n} x_n^2 + \theta_{2n+1} x_1 x_2 + \cdots + \theta_k x_{n-1} x_n$,
то есть уравнение регрессии включает линейные эффекты, парные взаимодействия и квадраты независимых переменных. Областью планирования является **гиперкуб** (обобщение куба на произвольное число измерений) в n -мерном пространстве, определяемый неравенствами $-1 \leq x_i \leq 1$, $i = 1 \dots n$. Непрерывные D -оптимальные планы для этого случая построены аналитически для различных размерностей пространства n .

Планы Кифера:

n	Вершины, E_0		Середины рёбер, E_1		Центры граней, E_2		Число точек плана r
	α_0	Число точек	α_1	Число точек	α_2	Число точек	
1	0,6666	2	0,3333	1	0	0	3
2	0,5832	4	0,3206	4	0,0962	1	9
3	0,5758	8	0,2274	12	0,1968	6	26
4	0,5928	16	0,1228	32	0,2844	24	72
5	0,617	32	0,02548	80	0,358	80	192

Здесь E – множество точек плана; α – частота точек.

Планы Коно:

n	Вершины, E_0		Середины рёбер, E_1		Центр, E_n		Число точек плана r
	α_0	Число точек	α_1	Число точек	α_n	Число точек	
1	0,6666	2	0,3333	1	0	0	3
2	0,5832	4	0,3206	4	0,0962	1	9
3	0,5103	8	0,4242	12	0,06547	1	21
4	0,4506	16	0,5021	32	0,04738	1	49
5	0,4021	32	0,5622	80	0,03675	1	113

Частота всех точек плана, принадлежащих одному множеству, одинакова.

Пример. Пусть задан непрерывный план в двухмерном пространстве. $n = 2$, что идентично для планов Кифера и Коно, но пользуемся последними. Из соответствующей таблицы число точек плана $r = 9$, но на практике удобнее пользоваться нормированными планами, поэтому переходим к относительным числам повторных измерений p . Тогда план имеет вид

$$\varepsilon = \left\{ \begin{array}{l} x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8, x_9 \\ p_1, p_2, p_3, p_4, p_5, p_6, p_7, p_8, p_9 \end{array} \right\}.$$

Точки плана не повторяются (разные координаты по осям), поэтому $p = \frac{1}{r} = \frac{1}{9}$.

Тогда план принимает вид $\varepsilon = \left\{ \begin{array}{l} x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8, x_9 \\ \frac{1}{9}, \frac{1}{9}, \frac{1}{9}, \frac{1}{9}, \frac{1}{9}, \frac{1}{9}, \frac{1}{9}, \frac{1}{9}, \frac{1}{9} \end{array} \right\}$, что неин-

формативно, поэтому переходим к частотам точек, считая $p = \alpha/\text{число точек}$.

Тогда план принимает вид

$$\varepsilon = \left\{ \begin{array}{l} x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8, x_9 \\ 0,1458, 0,1458, 0,1458, 0,1458, 0,08015, 0,08015, 0,08015, 0,08015, 0,0962 \end{array} \right\}.$$

Пример. Требуется найти лучший по критерию D -оптимальности план эксперимента, содержащего около 10 опытов. Примем размерность пространства $n = 3$, тогда из таблицы планов Коно число точек плана $r = 21$, относительные числа повторных измерений

$$p_1 = \dots = p_8 = 0,0637875, \quad p_9 = \dots = p_{20} = 0,03535, \quad p_{21} = 0,06547.$$

Числа измерений в точках плана

$$r_1 = \dots = r_8 = 10 \quad p_1 = 0,637875 \rightarrow 1,$$

$$r_9 = \dots = r_{20} = 10 \quad p_9 = 0,3535 \rightarrow 0,$$

$$r_{21} = 10 \quad p_{21} = 0,06547 \rightarrow 1,$$

оптимальное число точек плана $r^* = 8r_1 + 12r_9 + r_{21} = 8 \cdot 1 + 12 \cdot 0 + 1 = 9$, поэтому заданный D -оптимальный план содержит 9 опытов.

Число измерений оценивается не только в точках плана, но и в множестве точек плана. **Пример.** Требуется найти лучший по критерию D -оптимальности план эксперимента, содержащего 6 опытов. Примем размерность пространства $n = 3$, тогда из таблицы планов Коно частоты точек

$$\alpha_0 = 0,5103, \quad \alpha_1 = 0,4242, \quad \alpha_n = 0,06547.$$

Число измерений в множестве точек плана

$$r(E_0) = 6 \quad \alpha_0 = 3,0618 \rightarrow 3, \quad r(E_1) = 6 \quad \alpha_1 = 2,5452 \rightarrow 3, \quad r(E_n) = 6 \quad \alpha_n = 0,39282 \rightarrow 0.$$

Полный факторный эксперимент типа 2^k

В ходе эксперимента учитывают не все, а лишь **существенные** (влияющие на процесс) факторы, поскольку флюктуация любого неучтённого фактора сильно увеличит ошибку опыта. Однако при числе факторов $k > 15$ следует обратиться к **методам отсеивания несущественных факторов** (например, воспользоваться формализацией априорной информации).

Фактор считается **заданным**, если вместе с его названием указана **область определения** – совокупность всех значений, которые в принципе может принимать рассматриваемый фактор.

В зависимости от способа оценки все факторы делятся на **количественные** (например, время, давление, объём, температура) и **качественные** (технологические способы, аппараты, исполнители, вещества). Первым соответствует чи- словая шкала, а вторым – порядковая, которая условно сопоставляет уровни фактора числам натурального ряда (производит кодирование).

Ко всем факторам предъявляют следующие **требования**:

- **управляемость** – планирование эксперимента возможно лишь в том случае, когда уровни фактора подчиняются воле экспериментатора. **Управлять фактором** – установить нужное (определенное или выбранное) значение и поддерживать его постоянным в течение всего опыта или менять по заданной программе. **Пример.** Колонна для синтеза аммиака установлена на открытой площадке. Является ли температура воздуха фактором, включаемым в планирование эксперимента? Нет, поскольку неуправляема;
- **однозначность** – фактор должен задаваться в явном виде и иметь точное определение для чёткого понимания;
- **операциональность** – для однозначности фактора указывают последовательность действий (операций), с помощью которых устанавливают его уровни. **Пример.** Если фактором является давление в некоем аппарате, следует указать, в какой точке и каким прибором измерять его в ходе эксперимента;
- **точность замера** – должна быть максимально высокой. Степень точности определяется диапазоном изменения фактора: при исследовании процесса, длящегося десятки часов, нет необходимости учитывать доли минуты, а в реактивных процессах могут потребоваться доли секунды;
- **непосредственное воздействие** – фактор должен влиять на процесс напрямую, поскольку трудно управлять фактором, являющимся функцией других. **Сложные факторы** (например, соотношения между компонентами, их логарифмы) участвуют в планировании эксперимента лишь в отдельных случаях, когда динамические особенности объекта исследования представлены в статической форме. **Пример.** Исследуется диффузионный процесс. Является ли скорость диффузии фактором, включаемым в планирование эксперимента? Нет, поскольку зависит от концентрации, величины поверхности соприкосновения двух фаз и коэффициента растворения, в свою очередь зависящего от коэффициента диффузии, являющегося функцией нескольких переменных, и толщины диффузионного слоя, которая неодинакова в разных зонах.

Поскольку обычно одновременно изменяется несколько факторов, важны **требования к совокупности факторов**:

- **совместимость** – все комбинации факторов осуществимы (физически реализуемы) и безопасны. Однако несовместимость может наблюдаться на границах областей определения факторов или внутри них;
- **независимость** – отсутствие корреляции между факторами, возможность установления фактора на любом уровне независимо от уровней других факторов. Некоррелированность не означает, что между значениями факторов нет никакой связи: достаточно, чтобы связь не была линейной. **Пример.** Исследуется заданная термодинамическая система. Можно ли включить в планирование эксперимента сразу три фактора: давление P , объём V и температуру T ? Нет, поскольку для этой системы справедливо уравнение Менделеева-Клапейрона $PV = vRT$, где v – количество вещества, R – универсальная газовая постоянная, то есть два фактора задаются, а третий вычисляется.

Полный факторный эксперимент – эксперимент, реализующий все возможные комбинации уровней факторов.

Пример. При двух факторах ($k = 2$) число опытов $N = 2^k = 2^2 = 4$ описывается следующей **матрицей планирования эксперимента**, где каждая строка называется вектор-строкой, а каждый столбец – вектор-столбцом:

№	x_1	x_2	y	
1	-1	-1	y_1	$\begin{pmatrix} -1 & -1 \end{pmatrix}$
2	1	-1	y_2	$\begin{pmatrix} 1 & -1 \end{pmatrix}$
3	-1	1	y_3	$\begin{pmatrix} -1 & 1 \end{pmatrix}$
4	1	1	y_4	$\begin{pmatrix} 1 & 1 \end{pmatrix}$

Такая запись громоздка, особенно при многих факторах, и для сокращения вводят условные **буквенные обозначения строк**: порядковый номер фактора ставят в соответствие строчной букве латинского алфавита $x_1 \rightarrow a$, $x_2 \rightarrow b$. Однако у факторов два уровня, как быть с требованием однозначности? Если выписать буквы лишь для факторов на верхних уровнях, условия проведения опыта будут заданы однозначно. Опыт со всеми факторами на нижних уровнях обозначают (1). Тогда вышеупомянутая матрица принимает компактный вид

$$\begin{pmatrix} -1 & -1 \\ 1 & -1 \\ -1 & 1 \\ 1 & 1 \end{pmatrix} \quad (1)$$

$$\begin{matrix} a \\ b \\ ab \end{matrix} \rightarrow (1), a, b, ab$$

Процедура построения матрицы планирования эксперимента независимо от её размерности основана на **правиле чередования знаков**: в первом столбце знаки меняются поочерёдно, во втором – чередуются через 2, в третьем – через 4 и далее по степеням двойки.

Свойства полного факторного эксперимента типа 2^k

Все матрицы планирования эксперимента обладают следующими свойствами:

- **симметричность** – сумма элементов каждого вектор-столбца нулевая;
- **нормировка** – сумма квадратов элементов каждого вектор-столбца равняется числу опытов;
- **ортогональность** – сумма почлененных произведений любых двух вектор-столбцов нулевая;
- **ротатабельность** – по ГОСТу свойство плана, при котором дисперсия оценки отклика зависит только от расстояния от центра плана. Точки в матрице планирования эксперимента должны подбираться так, чтобы точность предсказания параметра оптимизации не зависела от направления движения в факторном пространстве.

Пример. Данна матрица b , a , b , a . Выполняются ли свойства?

$$\begin{pmatrix} -1 & 1 \\ 1 & -1 \\ -1 & 1 \\ 1 & -1 \end{pmatrix}.$$

Запишем матрицу в привычном виде

Если четыре строки, число опытов $N = 4$. Если два столбца, два фактора x_1 , x_2 .

Проверим симметричность, $x_1: -1 + 1 - 1 + 1 = 0$, $x_2: 1 - 1 + 1 - 1 = 0 \Rightarrow$ да.

Проверим нормировку, $x_1: 1 + 1 + 1 + 1 = 4 = N$, $x_2: 1 + 1 + 1 + 1 = 4 = N \Rightarrow$ да.

Проверим ортогональность, $x_1 x_2: -1 - 1 - 1 - 1 = -4 \neq 0 \Rightarrow$ нет.

Вывод: матрица планирования эксперимента составлена неоптимально.

Минимизация числа опытов

Полный факторный эксперимент обладает избыточностью опытов, их число сокращают за счёт информации, не очень существенной при построении математической модели. При этом матрица планирования эксперимента не должна лишаться своих оптимальных свойств.

Рассмотрим самый простой случай – полный факторный эксперимент 2^2 , чьи результаты можно представить в виде неполного квадратного уравнения

$$y(x, \theta) = \theta_0 + \theta_1 x_1 + \theta_2 x_2 + \theta_3 x_1 x_2$$

с матрицей планирования эксперимента

№	x_0	x_1	x_2	$x_1 x_2$	y
1	1	-1	-1	1	y_1
2	1	1	-1	-1	y_2
3	1	-1	1	-1	y_3
4	1	1	1	1	y_4

Если исследуемый процесс можно описать линейной математической моделью, то $\theta_3 \rightarrow 0$ и достаточно определить лишь θ_0 , θ_1 и θ_2 . Остаётся одна степень свободы, которую употребим для минимизации числа опытов: вектор-столбец $x_1 x_2$ используем для нового фактора x_3 , чтобы поставить не восемь, а всего четыре опыта при изучении трёх факторов.

Правило минимизации: для сокращения числа опытов необходимо новому фактору присвоить вектор-столбец, принадлежащий тому эффекту взаимодействия, которым можно пренебречь.

Пример. Взамен полного факторного эксперимента 2^3 предлагается матрица планирования эксперимента

№	x_0	x_1	x_2	x_3	y
1	1	-1	-1	1	y_1
2	1	1	1	-1	y_2
3	1	-1	1	-1	y_3
4	1	1	-1	1	y_4

Адекватна ли такая замена? Проверим симметричность, $x_1: -1 + 1 - 1 + 1 = 0$, $x_2: -1 + 1 + 1 - 1 = 0$, $x_3: 1 - 1 - 1 + 1 = 0 \Rightarrow$ да. Проверку нормировки пропустим, поскольку факторный эксперимент неполный. Проверим ортогональность, $x_1 x_2: 1 + 1 - 1 - 1 = 0$, $x_1 x_3: -1 - 1 + 1 + 1 = 0$, $x_2 x_3: -1 - 1 - 1 - 1 = -4 \neq 0 \Rightarrow$ нет. Вывод: такая замена неадекватна.

Дробная реплика

Поставив четыре опыта при изучении трёх факторов, мы воспользовались половиной полного факторного эксперимента 2^3 – **полурепликой** 2^{3-1} .

В общем виде имеем **условное обозначение реплик** 2^{k-p} ,

где 2 – число уровней фактора;

k – число факторов;

p – число линейных эффектов, приравненных к эффектам взаимодействия.

Для удобства сведём условные обозначения основных реплик в таблицу:

k	Название реплики	Условное обозначение реплики	N	
			Для дробного факторного эксперимента	Для полного факторного эксперимента
3	Полуреплика 2^3	2^{3-1}	4	8
4	Полуреплика 2^4	2^{4-1}	8	16
5	Четверть-реплика 2^5	2^{5-2}	8	32
6	1/8-реплика 2^6	2^{6-3}	8	64
7	1/16-реплика 2^7	2^{7-4}	8	128
5	Полуреплика 2^5	2^{5-1}	16	32
6	Четверть-реплика 2^6	2^{6-2}	16	64
7	1/8-реплика 2^7	2^{7-3}	16	128
8	1/16-реплика 2^8	2^{8-4}	16	256
9	1/32-реплика 2^9	2^{9-5}	16	512
10	1/64-реплика 2^{10}	2^{10-6}	16	1024

Замечания:

- с ростом числа факторов увеличивается дробность применяемых реплик;
- предельное число факторов для 8 опытов равняется 7, для 16 опытов – 15;
- план эксперимента с предельным числом факторов для данного числа опытов называется **насыщенным**;
- при максимальном числе факторов дробный факторный эксперимент эффективнее полного факторного эксперимента в 2048 раз.

Организация эксперимента

Для формализации (логического упорядочивания, связного изложения) имеющихся сведений об объекте исследования рекомендуется пользоваться стандартной **анкетой для сбора априорной информации**, где содержится:

- постановка задачи и выбор параметров оптимизации:**
 - краткое описание процесса и объекта исследования;
 - формулировка цели исследования (если решаемых задач несколько, они ранжируются по степени важности);
 - выбор параметров оптимизации (указываются порядковый номер, название, размерность, область определения, точность);
 - желаемый результат (число, точность);
 - оценка результата (какой будет считаться отличным, хорошим, удовлетворительным, неудовлетворительным);

выбор факторов:

- список всех подозреваемых (способных влиять на процесс) факторов;
- список факторов, включаемых в реальный эксперимент (указываются по-рядковый номер, название, размерность, область определения, область интереса, точность);
- возможность установления значения каждого фактора на любом уровне;
- возможность сохранения заданного уровня в течение всего опыта;
- вероятность остановки процесса;

число опытов:

- желаемое число опытов и ограничение на число опытов;
- желаемый срок проведения эксперимента;
- приблизительная длительность одного опыта;
- стоимость одного опыта как **технологическая себестоимость**, определяемая суммой затрат на осуществление процесса. В общем виде

$$C_m = C_M + C_{зп} + C_u,$$

где C_M – стоимость основных материалов, израсходованных на опыт;

$C_{зп}$ – заработка платы исполнителей со всеми начислениями;

C_u – цеховые расходы на ремонт и амортизацию оборудования и оснастки, электроэнергию, а также вспомогательные материалы.

В укрупнённом виде $C_m = I_{пост} + \frac{I_{пер}}{N}$,

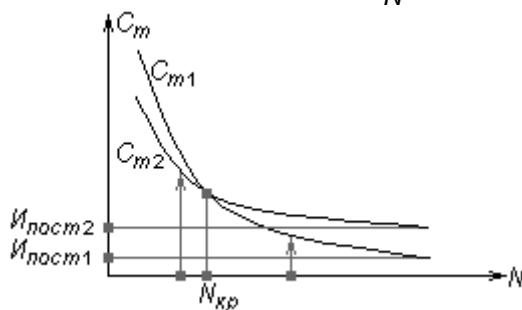
где $I_{пост}$ – постоянные издержки при проведении одного опыта (текущие расходы на материалы, заработную плату исполнителей, электроэнергию, отопление);

$I_{пер}$ – переменные издержки при проведении N опытов (единовременные расходы на текущий ремонт и новую оснастку).

Пусть технологическая себестоимость первого плана эксперимента

$$C_{m1} = I_{пост1} + \frac{I_{пер1}}{N},$$

а второго – соответственно $C_{m2} = I_{пост2} + \frac{I_{пер2}}{N}$.



Из этого графика видно, что при $N > N_{kp}$ (критическое число опытов) с экономической точки зрения целесообразно выбрать первый план эксперимента, а при $N < N_{kp}$ – второй;

- желаемое число уровней фактора;
- возможность проведения повторных измерений в одном опыте;
- желаемая стратегия проведения опытов;
- **учёт априорной информации:**
 - сводка условий и результатов, достигнутых при исследовании аналогичных процессов;
 - результаты предварительного (затравочного) эксперимента;
 - взаимодействие факторов.

К реализации плана эксперимента приступают лишь после заполнения анкеты.

К проведению опытов тщательно **готоятся**: нужно собрать опытную установку или привести в рабочее состояние оборудование, проверить и откалибровать измерительные устройства, подготовить исходное сырьё (образцы), чётко изложить последовательность действий (операций).

Определение брака измерений

Самый простой вариант – использовать **критерий Стьюдента**. Опыт считается бракованым, если экспериментальное значение превышает табличное:

$$\frac{|y - y_{cp}|}{\sigma} > t,$$

где y – проверяемое измерение;

y_{cp} – среднее остальных измерений;

$$\sigma = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (y_i - y_{cp})^2} \quad \text{– стандартное отклонение;}$$

n – число остальных измерений, дающее степень свободы $n-1$;

i – порядковый индекс измерения;

t – распределение Стьюдента, табличное значение из метода 10.

Пример. Имеем измерения 2,37 2,71 2,76 3,58, где последнее вызывает сомнение, не **грубое** ли оно, поскольку резко отклоняется от прочих, выделяется на их фоне. Проверим измерение $y = 3,58$, исключив его из расчёта:

число остальных измерений $n = 3$,

$$\text{среднее остальных измерений } y_{cp} = \frac{2,37 + 2,71 + 2,76}{3} = 2,61,$$

$$\text{стандартное отклонение } \sigma = \sqrt{\frac{(-0,24)^2 + 0,1^2 + 0,15^2}{3-1}} = 0,21,$$

$$\frac{|y - y_{cp}|}{\sigma} = \frac{3,58 - 2,61}{0,21} = 4,62, \text{ степень свободы } n-1 = 3-1 = 2,$$

из метода 10 распределение Стьюдента $t = 4,3$,

$4,62 > 4,3$, поэтому четвёртый опыт бракованный.

Для определения брака измерений можно использовать и другие критерии – например, **r-распределение**. Опыт считается бракованным, если экспериментальное значение превышает табличное:

$$\frac{|y - y_{cp}|}{\sigma \sqrt{\frac{n-1}{n}}} > r,$$

где y – проверяемое измерение;

y_{cp} – среднее измерений;

$$\sigma = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (y_i - y_{cp})^2} \text{ – стандартное отклонение;}$$

n – число измерений, дающее степень свободы $n-2$;

i – порядковый индекс измерения;

r – r-распределение, табличное значение:

Степень свободы	$r(5\%)$	Степень свободы	$r(5\%)$
1	1,412	6	2,172
2	1,689	7	2,237
3	1,869	8	2,294
4	1,996	9	2,343
5	2,093	10	2,387

Пример. Имеем измерения 10,85 12,15 16, где последнее вызывает сомнение, не грубое ли оно. Проверим измерение $y = 16$, не исключая его из расчёта: число измерений $n = 3$,

$$\text{среднее измерений } y_{cp} = \frac{10,85 + 12,15 + 16}{3} = 13,$$

$$\text{стандартное отклонение } \sigma = \sqrt{\frac{(-2,15)^2 + (-0,85)^2 + 3^2}{3-1}} = 2,68,$$

$$\frac{|y - y_{cp}|}{\sigma \sqrt{\frac{n-1}{n}}} = \frac{16 - 13}{2,68 \sqrt{\frac{3-1}{3}}} = 1,38, \text{ степень свободы } n-2 = 3-2 = 1,$$

из таблицы выше t -распределение $t = 1,412$,
 $1,38 < 1,412$, поэтому третий опыт небракованный.

Пример. Проверим полученный в предыдущем примере результат по критерию Стьюдента, исключив измерение $y = 16$ из расчёта:

число остальных измерений $n = 2$,

$$\text{среднее остальных измерений } y_{cp} = \frac{10,85 + 12,15}{2} = 11,5,$$

$$\text{стандартное отклонение } \sigma = \sqrt{\frac{(-0,65)^2 + 0,65^2}{2-1}} = 0,92,$$

$$\frac{|y - y_{cp}|}{\sigma} = \frac{16 - 11,5}{0,92} = 4,89, \text{ степень свободы } n-1 = 2-1 = 1,$$

из метода 10 распределение Стьюдента $t = 12,7$,
 $4,89 < 12,7$, поэтому третий опыт небракованный.