ПАРАМЕТРИЧЕСКАЯ ИДЕНТИФИКАЦИЯ

ФН-12. Магистры - 3 семестр

Лекция 1. (1-2.) ИДЕНТИФИКАЦИЯ МАТЕМАТИЧЕСКИХ МОДЕЛЕЙ. ВВЕДЕНИЕ. ЭЛЕМЕНТЫ ТЕОРИИ ВЕРОЯТНОСТИ И МАТЕМАТИЧЕСКОЙ СТАТИСТИКИ

1.1. Задачи и методы идентификации математических моделей.

Литература.

1. Гроп Д. Методы идентификации систем. - М.: Мир, 1979. -302 с.
2. Льюнг Л. Идентификация систем. Теория для пользователя. - М.: Наука, 1991. - 432 с.
3. Закс Ш. Теория статистических выводов - М.: Мир, 1975. -570 с.
4. Дилигенская А.Н. Идентификация объектов управления (учебное пособие). ГОУВПО "Самарский государственный технический университет".
5. Бар-Шалом Я., Ли Х-Р,(перевод Дмитриев Д. Д.). Траекторная обработка. Принципы, способы и алгоритмы. МГТУ, 2011.
6. Зарубин В.С. Математическое моделирование в технике. - М.: МГТУ им. Н.Э. Баумана, 2003.
7. Коровин С. К., Фомичев В. В. Наблюдатели состояния для линейных систем с неопределенностью. - М.: "Физматлит", 2007.
8. Isermann R.,Münchhof M. Identification of Dynamic Systems.Springer-Verlag, Berlin. 2011.
9. Волков И.К., Зуев С.М.,Цветкова Г.М. Случайные процессы. - М.: МГТУ им. Н.Э. Баумана, 2003.
10. Горяинов В.Б., Павлов И.В., Цветкова Г.М.,Тескин О.И. Математическая статистика- М.: МГТУ им. Н.Э. Баумана, 2001.

Термин ”идентификация” стал широко применяться в качестве одного из базовых терминов теории управления в шестидесятых годах XX века (identifico - отождествляю,распознаю).

Модель объекта управления — информация о наиболее существенных характеристиках ОУ.

Модели объектов управления:

- физические или математические;
- статические или динамические;
- одномерные или многомерные;
- линейные или нелинейные;
- детерминированные или стохастические;
- непрерывные или дискретные;
- стационарные или нестационарные;
- с сосредоточенными или распределенными параметрами;
- структурированные или агрегированные;
- параметрические или непараметрические;
- т.д.

Решение задачи построение модели динамической системы по данным наблюдений за их поведением составляет основной предмет теории идентификации (Льюнг Л.).

Математическая модель (образ) представляет собой абстрактное отражение реального объекта (оригинала, прообраза).

Свойства математических моделей.

Полнота ММ позволяет отразить в достаточной мере именно те характеристики и особенности ТО, которые интересуют нас с точки зрения поставленной цели проведения вычислительного эксперимента.
Точность ММ дает возможность обеспечить приемлемое совпадение реальных и найденных при помощи ММ значений выходных параметров ТО.
Адекватность ММ - это способность ММ описывать выходные параметры ТО с относительной погрешностью не более некоторого заданного значения.
Экономичность ММ оценивают затратами на вычислительные ресурсы, необходимые для реализации ММ на ЭВМ.
Робастность ММ характеризует ее устойчивость по отношению к погрешностям исходных данных, способность нивелировать эти погрешности.
Наглядность ММ является ее желательным, но необязательным свойством.

ММ объектов, рассматриваемых в курсе можно представить в виде

$$y = f(x, a, u),$$

или

$$\begin{cases} \dot{x} = h(x, a, u) \\ y = f(x, a, u), \end{cases}$$

где f — скалярная или векторная функция векторного аргумента;
$x \in R^n$ — вектор состояния объекта управления;
$a \in R^m$ — внутренние параметры объекта управления;
$y \in R^k$ — выходы объекта управления;
$u \in R^l$ — управление.
Построение ММ объекта может производиться несколькими методами: аналитическим, экспериментальным и экспериментально-аналитическим.

Основные задачи идентификации.

1. Выбор общей структуры модели.
2. Определение оценок **неизвестных параметров модели объекта** на основе информации об управлении и о выходе объекта, и информации о действующих на объект случайных помех.
3. Получении оценки **вектора состояния системы** по данным о входе и выходе при наличии случайных помех.

Задача получения оценки полного вектора состояния на основе информации о прилагаемой управлении и выходе системы может решаться с использованием **наблюдателей**.
Параметры моделей системы полагаются известными, вся неизвестная информация о изучаемой системе моделируется с помощью неопределенной помехи, представляемой некоторой неизвестной функцией (например, липшецевой или ограниченной).

С точки зрения **теории наблюдателей** можно выделить две задачи

1. Задача наблюдения — задача оценки состояния системы в момент времени t_0 по известным входу и выходу системы $u(t)$ и $y(t)$ при $t \geq t_0$, т.е задача восстановления начального значения фазового вектора по будущим измерениям входа и выхода.

2. Задача идентификации — задача оценки состояния системы в момент времени t_* по данным о входе и выходе при $t \leq t^*$, т.е задача восстановления фазового вектора в текущий момент времени t_* по измерениям входа и выхода в прошлом.

Замечание: часто не делают различия между задачами наблюдением и идентификацией, объединяя эти понятия.

Динамическая система характеризуется своим начальным состоянием и законом, по которому система переходит из начального состояние в другое. В рамках курса под математической моделью ДС будем рассматривать модель представленную системой ОДУ с управлением и выходом вида

$$\dot{x} = f(x, u, t),$$
$$y = \varphi(x, u, t),$$

или разностных уравнений вида

$$\Delta x_{k+1} = f_k(x_k, u_k, t_k),$$
$$y_{r+1} = \varphi_k(x_k, u_k, t_k).$$

Идентификация — это нахождение оптимальной (в смысле выбранных критериев) модели, построенной по результатам наблюдений над входными и выходными переменными объекта.

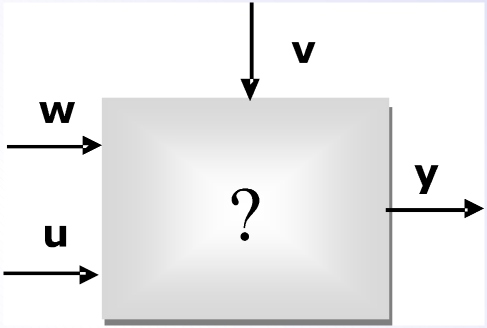

Рис. 1. Динамическая система

На рисунке u — контролируемый (управляемый) входной сигнал, y — выходной (наблюдаемый) сигнал. w и v — помехи (возмущения), w — измеряемая непосредственно помеха (возмущение), v — неизмеряемая, или измеряемые косвенно, по воздействию, оказанному ими на выходной сигнал.
Характер помех может изменяться в широком диапазоне.
Помехи можно разбить на два класса: измеряемые и измеряемые косвенно, по воздействию, оказанному ими на выходной сигнал.
Воздействие помех (как правило) обнаруживается лишь по тому совокупному эффекту, который порождает действие помех в выходном сигнале.

”Черный ящик”. ”Серый ящик”. ”Белый ящик”.

Идентификация динамических объектов в общем случае состоит в переходе от ”черного ящика” к ”белому”.

1. Структурная идентификация;
2. Параметрическая идентификация.

”Черный ящик” (black-box model)

Априорных знаний об объекте мало. Проводилось только экспериментальное моделирование (измерения входных/выходных сигналов).

Есть предположения о структуре модели и/или используются стандартные (линейные) математические модели. Параметры моделей — средства подстройки моделей к экспериментальным данным.

”Белый ящик”. (white-box model)

Известны теоретические (физические, химические, биологические) законы, на основе которых строится модель, структура и параметры математической модели.

”Серый ящик” (gray-box model)

Известны физические законы, на основе которых строится модель. Проводилось экспериментальное моделирование (измерения сигналов).

1. Известна структура математической модели, оценки параметров. Часто задается системой ДУ (light-gray-box model).
2. Параметры и структура математической модели неизвестны (dark-gray-box model).

Общая схема идентификации модели

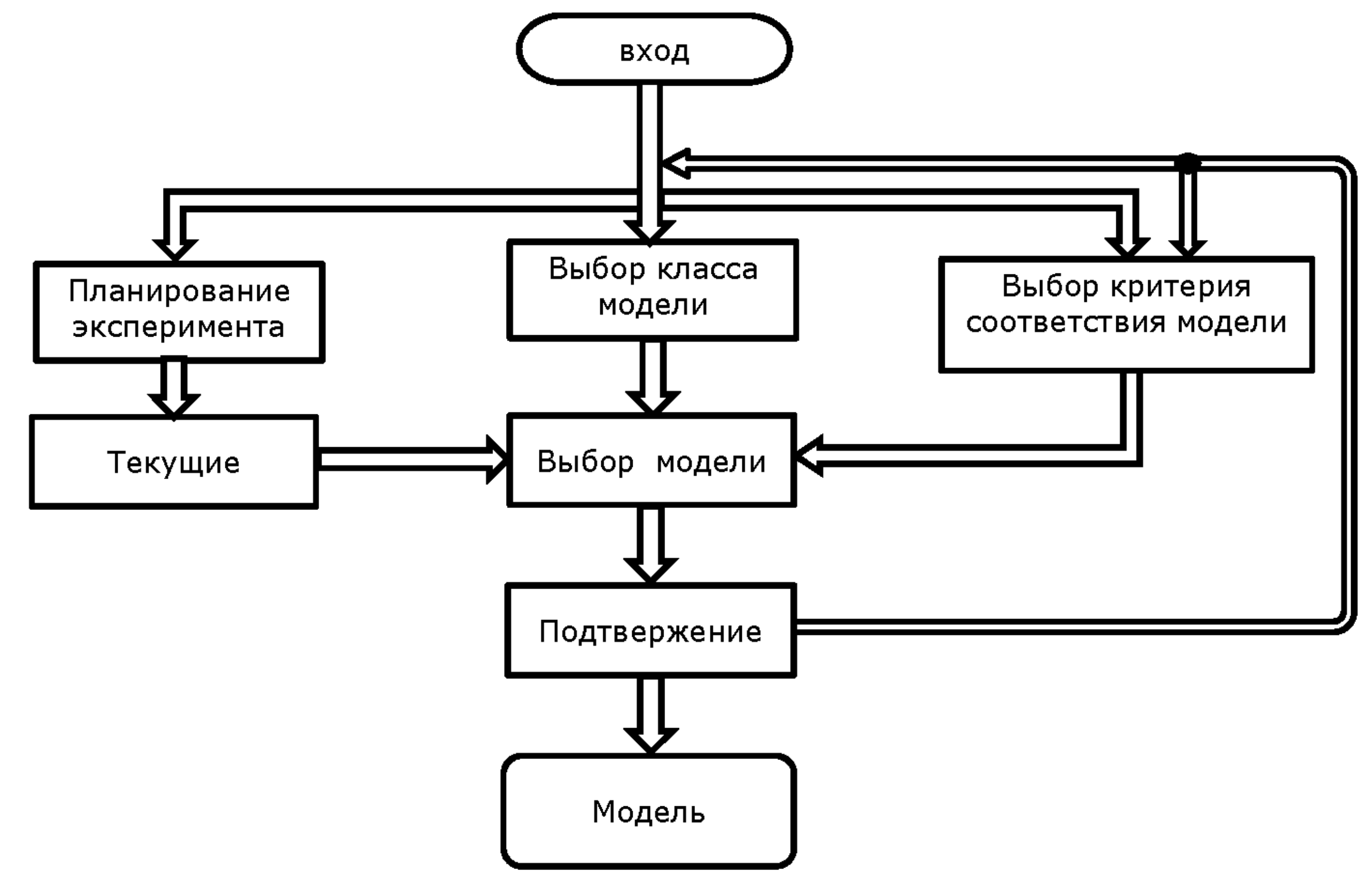

Рис. 2

Структурная идентификация

Объект управления представляет собой "черный ящик". Задачей структурной идентификации является представление реального объекта в виде математической модели.

Необходимо получить описание структуры математической модели (переход от "черного ящика" к "серому ящику").

1. Выбор класса моделей-кандидатов из которых будет выбрана подходящая модель, т.е. выбор общей структуры модели и класс уравнений, которыми предполагается описывать объект.
2. Выбор точек приложения внешних воздействий и сбора информации о реакциях объекта (размещение управляющих устройств и датчиков).
3. Приложение внешних воздействий и анализ реакции объекта на них.

Выбор класса ММ зависит от типа объекта. Модели основывающиеся на теории множеств (инфологические, семантические, синтаксические...)

Модели в виде систем дифференциальных или разностных уравнений используются для описания динамических систем.

Проверка модели

Проверка модели путем сравнения выходных данных модели и соответствующих выходных сигналов объекта. Критерий соответствия модели и объекта выбирается в зависимости от наличия априорной информации об объекте, требуемой точности, выбраного метода идентификации, времени, выделенного на проведение эксперимента и др.

Параметрическая идентификация

Структура ММ объекта заранее известна, параметры неизвестны, имеются данные экспериментов (данные наблюдений). Переход от "серого ящика" к "белому ящику".

К данным наблюдений обычно относятся входные и выходные сигналы.

Классификация методов идентификации по способу получения данных.

1) активная идентификация — в процессе эксперимента, эксперимент спланирован заранее;

2) пассивная идентификация — в процессе эксплуатации (невозможно провести эксперимент или влиять на процедуру регистрации (изменения) данных);

2) идентификация смешанного типа — в процессе эксплуатации к рабочим управляющим сигналам добавляются тестовые сигналы.

Активная идентификация

Объект исследования выводится из нормального режима эксплуатации. Исследования проводятся в специализированных лабораторных условиях. На входы объекта подаются тестовые сигналы специального вида.

Виды сигналов:ступенчатые, импульсные, гармонические, случайные воздействия с заданными параметрами.

Активная идентификация ОУ может производиться во временной области и/или в частотной области.

Активную идентификацию используют при разработке новых технологий применительно к действующим промышленным объектам, в изучении новых явлений, в первоначальной разработке математической модели.

В результате экспериментов получают частотные характеристики системы как ответ на гармонические воздействия (АФЧХ, ЛАЧХ, ЛФЧХ и др.), или переходные (временные) характеристики системы (ступенчатое изменение управления, управление в виде ”узкого” импульса и др.)

Пассивная идентификация

Объект функционирует в режиме нормальной эксплуатации.
На его входы поступают рабочие сигналы управления.
Возможна идентификация динамических систем без управления.
Пассивную идентификацию используют для уточнения математической модели, для слежения за изменениями в объекте. Информация оперативно используется в системе управления объектом.

Идентификация смешанного типа

Объект не выводится из нормального режима эксплуатации, к управляющим сигналам добавляются тестовые воздействия.

Условия проведения эксперимента, с учетом всевозможных ограничений, определяются заранее, при планировании эксперимента.

Определяются: диапазон изменения входных сигналов, возможность их изменения, количество опытов, частоту регистрации выходных сигналов необходимых для решения поставленной задачи с требуемой точностью.

1.2. Элементы теории вероятности

Теория вероятностей является разделом математики, в котором изучают математические модели **случайных экспериментов**, т.е. экспериментов, исходы которых нельзя определить однозначно условиями проведения опыта.

Важнейший принцип ТВ состоит в **статистической устойчивости частот** возможных исходов испытаний. При неограниченном возрастании количества случайных испытаний доля испытаний, приведших к заданному исходу, в ряду всех проведенных испытаний стабилизируется около некоторого предельного числа.

При построении вероятностной модели выделяют такой набор исходов данного случайного испытания, который удовлетворяет двум условиям.

1. При проведении случайного испытания должен наступить один из (элементарных) исходов ω в выбранном наборе исходов.

2. Все исходы в наборе являются взаимно исключающими, т.е. случайное испытание не должно завершаться одновременно двумя исходами.

Теоретико-множественная трактовка основных понятий и аксиоматическое построение ТВ

Аксиоматический подход был введен А.Н. Колмогоровым (в 30-х гг.).

Ω — множество всех возможных исходов испытания. Каждый $\omega \in \Omega$ называют элементарным исходом (событием).

Любое событие A рассматривается как некоторое подмножество универсального множества Ω ($A \subseteq \Omega$).

$\varnothing \subseteq \Omega$ — невозможное событие.

Ω — событие, происходящее всегда (при любом элементарном исходе), называется достоверным событием.

Суммой нескольких событий $A_1, A_2, \dots, A_n$ называется объединение множеств $A_1 \cup A_2 \cup \dots \cup A_n$.

Произведением нескольких событий $A_1, A_2, \dots, A_n$ называется пересечение множеств $A_1 \cap A_2 \cap \dots \cap A_n$.

Событием $\bar{A}$, противоположным событию A, называется дополнение множества A до Ω, $\Omega \setminus A$.

Под операциями над событиями понимаются операции над соответствующими множествами.

Определение 1.1. Пусть Ω — некоторое множество. Некоторое семейство $\mathcal{A}$ подмножеств из Ω называют σ **-алгеброй**, если:

1. $\Omega \in \mathcal{A}$ (множество Ω содержится в $\mathcal{A}$;
2. $(\forall B) \in \mathcal{A} \Rightarrow \overline{B} \in \mathcal{A}$
3. $(\forall B) B_k \in \mathcal{A}$, $(\forall k) k \in \mathbb{N}$:

$$\bigcup_{k=1}^{\infty} B_k \in \mathcal{A} \qquad \text{и} \qquad \bigcap_{k=1}^{\infty} B_k \in \mathcal{A}.$$

σ -алгебра замкнута относительно теоретико-множественных операций.

Определение 1.2. Пару $(\Omega, \mathcal{A})$ из множества Ω и σ -алгебры $\mathcal{A}$ подмножеств из Ω называют **измеримым пространством**.

Определение 1.3. **Вероятностью (вероятностной мерой)** на измеримом пространстве $(\Omega, \mathcal{A})$ называют скалярную функцию $\mathbf{P} : \mathcal{A} \to \mathbb{R}$, определенную на σ-алгебре $\mathcal{A}$, которая удовлетворяет трем условиям:
а) $\mathbf{P}[A] \geq 0$, $A \in \mathcal{A}$;
б) $\mathbf{P}[\Omega] = 1$;
в) для любых попарно не пересекающихся множеств $B_k \in \mathcal{A}$, $k \in \mathbb{N}$, имеет место равенство

$$\mathbf{P}\left[\bigcup_{k=1}^{\infty} B_k\right] = \sum_{k=1}^{\infty} \mathbf{P}[B_k].$$

Множество Ω, на котором задана σ-алгебра $\mathcal{A}$, — это пространство элементарных исходов, элементы σ-алгебры $\mathcal{A}$ — это события, вероятностная мера $\mathbf{P}$ — это функция, задающая вероятности всевозможных событий.

Определение 1.4. Тройку $(\Omega, \mathcal{A}, \mathbf{P})$ из непустого множества Ω, заданной на Ω σ-алгебры $\mathcal{A}$ и определенной на σ-алгебре $\mathcal{A}$ вероятностной меры $\mathbf{P}$ называют **вероятностным пространством**.

Определение 1.5. Пусть $(\Omega, \mathcal{A}, \mathbf{P})$ — вероятностное пространство и $A_1,\ A_2 \in \mathcal{A}$ — события, причем $\mathbf{P}[A_1] > 0$. Под **условной вероятностью** события A_2 относительно события A_1 понимают число

$$\mathbf{P}[A_2 \mid A_1] \triangleq \frac{\mathbf{P}[A_1 \cap A_2]}{\mathbf{P}[A_1]}.$$

Определение 1.6. Пусть $(\Omega, \mathcal{A}, \mathcal{A}P)$ — вероятностное пространство. События $A,\ B \in \mathcal{A}$, для которых $\mathbf{P}[A] > 0$ и $\mathbf{P}[B] > 0$, называют **независимыми**, если выполнены равенства

$$\mathbf{P}[A \mid B] = \mathbf{P}[A] \quad \text{и} \quad \mathbf{P}[B \mid A] = \mathbf{P}[B].$$

В противном случае события A и B называют **зависимыми**.

Определение 1.7. Отображение $\xi\colon \Omega \to X$ измеримого пространства $(\Omega, \mathcal{A})$ в измеримое пространство $(X, \mathcal{B})$ называют **измеримой функцией,** если для любого $B \in \mathcal{B}$ прообраз множества B, т.е. множество $\xi^{-1}(B) = \{\omega \in \Omega\colon \xi(\omega) \in B\}$, принадлежит A. ▮

Измеримое пространство $(\Omega, \mathcal{A})$ входит в состав вероятностного пространства $(\Omega, \mathcal{A}, \mathbf{P})$ (дополнительно задана вероятностная мера).
Измеримая функция $\xi(\omega)$, $\omega \in \Omega$, преобразует вероятностную меру $\mathbf{P}$ на $(\Omega, \mathcal{A})$ в вероятностную меру $\mathbf{P}_\xi$, определяемую равенством $\mathbf{P}_\xi[B] = \mathbf{P}[\xi^{-1}(B)]$ и превращает измеримое пространство $(X, \mathcal{B})$ в вероятностное пространство $(X, \mathcal{B}, \mathcal{P}_\xi)$. Таким способом можно заменить вероятностное пространство, „данное от природы“, на более удобное и простое.

Особый случай: $X = \mathbb{R}$, а $\mathcal{B}$ — **борелевская σ-алгебра**, т.е. минимальная σ-алгебра, содержащая все промежутки (открытые, замкнутые, полуоткрытые) на числовой оси.
В этом случае, измеримая функция определяется, как функция $\xi(\omega)$, $\omega \in \Omega$, для которой прообраз любого множества B вида $B = \{x \in \mathbb{R}\colon x < a\}$ принадлежит σ-алгебре $\mathcal{A}$, т.е. является событием.
Такую измеримую функцию в теории вероятностей называют **случайной величиной**.
На множестве $\mathbb{R}^n$ можно задать σ-алгебру, отталкиваясь от множеств вида

$$T_1 \times T_2 \times \ldots \times T_n = \left\{(x_1, \ldots, x_n) \in \mathbb{R}^n\colon x_i \in T_i,\ i = \overline{1,\, n}\right\},$$

где T_i — промежутки числовой оси.
Борелевская σ-алгебра $\mathcal{B}$ в $\mathbb{R}^n$ — это наименьшая σ-алгебра, содержащая указанные множества.
Измеримая функция, отображающая Ω в $\mathbb{R}^n$, представляет собой совокупность n случайных величин — ее координатных функций.
Такую функцию называют n-мерным **случайным вектором**.
Случайный вектор — многомерное обобщение понятия случайной величины.
Случайная величина — n-мерный случайный вектор с $n = 1$.

Случайный вектор можно определить как n-мерную векторную функцию $\xi(\omega)$, $\omega \in \Omega$, для которой при любом $x \in \mathbb{R}^n$ множество

$$A_x \triangleq \{\omega \in \Omega \colon \xi(\omega) < x\}$$

принадлежит $\mathcal{A}$ (является событием).
Здесь

$$\xi(\omega) \triangleq \begin{pmatrix} \xi_1(\omega) \\ \vdots \\ \xi_n(\omega) \end{pmatrix}, \qquad x \triangleq \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix},$$

неравенство $\xi(\omega) < x$ означает, что $\xi_k(\omega) < x_k$, $k = \overline{1, n}$. ■
Изучение случайной величины $\xi(\omega)$, $\omega \in \Omega$, сводится к анализу вероятностной меры $\mathbf{P}_\xi$, которая порождается этой функцией на борелевской σ-алгебре.
Исходное вероятностное пространство $(\Omega, \mathcal{A}, \mathbf{P})$ перестает играть роль.
Вероятностная мера $\mathbf{P}_\xi$ полностью определяется по своим значениями на промежутках вида $(-\infty, a)$. Можно найти значение вероятности на любом множестве борелевской σ-алгебры, комбинируя такие промежутки и используя свойства аддитивности и счетной аддитивности вероятности.

Определение 1.8. Пусть $\xi(\omega)$, $\omega \in \Omega$, — случайная величина (случайный вектор). Скалярную функцию $F_\xi(x) \triangleq \mathbf{P}_\xi[A_x]$, где $x \in \mathbb{R}$ ($x \in \mathbb{R}^n$) и $A_x \triangleq \{\omega \in \Omega\colon \xi(\omega) < x\}$, называют **функцией распределения (вероятностей) случайной величины (случайного вектора)** $\xi(\omega)$, $\omega \in \Omega$.

Случайные величины $\xi_i(\omega)$, $\omega \in \Omega$, $i = \overline{1,\, n}$, можно рассматривать как координаты n-мерного случайного вектора. Иногда функцию распределения случайного вектора называют **совместной функцией распределения**. Сокращенные формы записи:

$$F_\xi(x) = \mathbf{P}[\xi(\omega) < x], \quad F_\xi(x_1, \ldots, x_n) = \mathbf{P}[\xi_k(\omega) < x_k,\ k = \overline{1,\, n}\,].$$

Определение 1.9. Случайную величину (случайный вектор), принимающую не более чем счетное множество значений, называют **дискретной случайной величиной (дискретным случайным вектором)**.

Для дискретного случайного вектора $\xi(\omega)$, $\omega \in \Omega$, множество значений можно записать в виде конечной или бесконечной последовательности $\{x_{(k)}\}_{k=1}^{N} \subset \mathbb{R}^n$, где $N \leq \infty$. Функция распределения случайного вектора имеет вид

$$F_\xi(x) \equiv \sum_{x_{(k)}<x} \mathbf{P}[\xi(\omega) = x_{(k)}].$$

Определение 1.10. Непрерывным случайным вектором (СВ) называют n -мерный случайный вектор $\xi(\omega)$, функцию распределения которого можно представить в виде

$$F_\xi(x) = \int_{-\infty}^{x} f_\xi(y)\, dy,$$

где $x \triangleq \left(\eta^{\mathbb{T}}(x_1 \ldots x_n)\right)^{\mathbb{T}}$, $y \triangleq (y_1 \ldots y_n)^{\mathbb{T}}$, или, что то же самое,

$$F_\xi(x_1, \ldots, x_n) = \int_{-\infty}^{x_1} \ldots \int_{-\infty}^{x_n} f_\xi(y_1, \ldots, y_n)\, dy_1 \ldots dy_n.$$

Функцию $f_\xi(x) = f_\xi(x_1, \ldots, x_n)$ называют **плотностью распределения (вероятностей)**. В частном случае при $n = 1$ непрерывный СВ $\xi(\omega)$ называют **непрерывной случайной величиной**.

Свойства плотности распределения $f_\xi(x)$ n-мерного НСВ $\xi(\omega)$.

1. $f_\xi(x) \geq 0$, $x \in \mathbb{R}^n$;

2. Вероятность $\mathbf{P}[\xi(\omega) \in G]$ попадания в произвольную n-мерную область G: $\mathbf{P}[\xi(\omega) \in G] = \int_G f_\xi(x)\, dx$;

$\int_{\mathbb{R}^n} f_\xi(x)\, dx = 1$;

вероятность попадания значения СВ в множество нулевой площади в $\mathbb{R}^n$ (в фиксированную точку) всегда равна нулю;

3. Плотность распределения $f_\xi(x)$ в своих точках непрерывности равна:

$$f_\xi(x_1, \ldots, x_n) = \frac{\partial^n F_\xi(x_1, \ldots, x_n)}{\partial x_1 \ldots \partial x_n};$$

4. Если $\xi(\omega) \triangleq \left(\eta^{\mathbb{T}}(\omega)\, \varepsilon^{\mathbb{T}}(\omega)\right)^{\mathbb{T}}$, где $\xi(\omega)$ — n-мерный, $\eta(\omega)$ — m-мерный ($0 < m < n$), а $\varepsilon(\omega)$ — $(n-m)$-мерный непрерывные СВ с плотностями распределения $f_\xi(x)$, $f_\eta(y)$, $f_\varepsilon(z)$, то

$$f_\eta(y) = \int_{\mathbb{R}^{n-m}} f_\xi(y, z)\, dz, \qquad f_\varepsilon(z) = \int_{\mathbb{R}^m} f_\xi(y, z)\, dy,$$

где $y \in \mathbb{R}^m$, $z \in \mathbb{R}^{n-m}$, $x = \left(y^{\mathbb{T}}\, z^{\mathbb{T}}\right)^{\mathbb{T}} \in \mathbb{R}^n$.

Обобщенная плотность распределения вероятностей дискретного СВ $\xi(\omega)$ с множеством возможных значений $\{x_{(k)}\}_{k=1}^{N} \subset \mathbb{R}^n$, $N \leq \infty$:

$$f_\xi(x) = \sum_{k=1}^{N\leq\infty} \mathbf{P}[\xi(\omega) = x_{(k)}]\delta(x - x_{(k)}),$$

где $\delta(x) \equiv \delta(x_1)\delta(x_2)\ldots\delta(x_n)$ — δ *-функция Дирака.*

Функцию и плотность распределения вероятностей СВ $\xi(\omega) = (\xi_1(\omega)\ldots\xi_n(\omega))^{\mathbb{T}}$ называют **совместной функцией распределения вероятностей** и **совместной плотностью распределения вероятностей** случайных величин $\xi_1(\omega), \ldots, \xi_n(\omega)$.

Для СВ задан **закон распределения случайного вектора** (при $n = 1$ **случайной величины**), если задана функция или плотность распределения вероятностей случайного вектора $\xi(\omega)$.

Закон распределения СВ $\xi(\omega) = (\xi_1(\omega)\ldots\xi_n(\omega))^{\mathbb{T}}$ называют также **совместным законом распределения вероятностей** случайных величин $\xi_1(\omega), \ldots, \xi_n(\omega)$.

Определение 1.11. Случайные векторы $\xi_k(\omega)$, $k = \overline{1,\ N}$, заданные на одном и том же вероятностном пространстве $(\Omega, \mathcal{A}, \mathbf{P})$, называют **независимыми в совокупности**, если функция распределения $F_\xi(x_1, \ldots, x_n)$ блочного случайного вектора $\xi(\omega) = (\xi_1(\omega) \ldots \xi_n(\omega))^{\mathbb{T}}$ имеет вид

$$F_\xi(x_1, \ldots, x_n) = \prod_{k=1}^{N} F_{\xi_k}(x_k),$$

где $F_{\xi_k}(x_k)$, $k = \overline{1,\ N}$, — функции распределения случайных векторов $\xi_k(\omega)$. Если $\xi_k(\omega)$, $k = \overline{1,\ N}$, — случайные величины, то говорят о **независимых случайных величинах**.

Определение 1.12. Пусть $\eta(\omega)$ — n-мерный, $\varepsilon(\omega)$ — m-мерный, $\xi(\omega) \triangleq \left(\eta^{\mathbb{T}}(\omega)\ \varepsilon^{\mathbb{T}}(\omega)\right)^{\mathbb{T}}$ — $(n+m)$-мерный случайные векторы с плотностями распределения (возможно, обобщенными) $f_\eta(x)$, $f_\varepsilon(y)$, $f_\xi(x,y)$. **Условной плотностью распределения (условным законом распределения)** случайного вектора $\eta(\omega)$ при условии, что случайный вектор $\varepsilon(\omega)$ принял некоторое фиксированное значение $y \in \mathbb{R}^n$, называют функцию

$$f(x\,|\,y) = \frac{f_\xi(x,y)}{f_\varepsilon(y)}.$$

Пусть на вероятностном пространстве $(\Omega, \mathcal{A}, \mathbf{P})$ задан n-мерный случайный вектор $\xi(\omega)$, S — множество возможных значений $\xi(\omega)$, $S \subset \mathbb{R}^n$ и измеримая (относительно борелевской алгебры $\mathcal{B}$ в $\mathbb{R}^n$) векторная функция $\varphi\colon S \to \mathbb{R}^m$.

Определение 1.13. Векторной (скалярной, $m = 1$**) функцией случайного вектора (случайной величины,** $n = 1$) $\xi(\omega)$ называют композицию функций

$$\eta(\omega) = \varphi(\xi(\omega)).$$

В качестве измеримой функции $\varphi(x)$ можно взять любую непрерывную, кусочно непрерывную или интегрируемую векторную функцию.

Определение 1.14. **Математическим ожиданием** n-мерного случайного вектора $\xi(\omega) = (\xi_1(\omega) \dots \xi_n(\omega))^{\mathbb{T}}$ с функцией плотности вероятностей $f_\xi(x) = f_\xi(x_1, \dots, x_n)$ называют n-мерный вектор

$$\mathbf{M}[\xi(\omega)] \triangleq \int_{\mathbb{R}^n} x f_\xi(x)\, dx,$$

компонентами которого являются числа

$$\mathbf{M}[\xi_k(\omega)] = \int_{\mathbb{R}^n} x_k f_\xi(x)\, dx, \quad k = \overline{1,\, n}.$$

Числа $\mathbf{M}[\xi_k(\omega)]$ представляют собой математические ожидания координатных случайных функций $\xi_k(\omega)$.
Математическое ожидание случайного вектора не определено, если хотя бы один несобственный интеграл $\int_{\mathbb{R}^n} x_k f_\xi(x)\, dx$ расходится.

Свойства математического ожидания

Математическое ожидание случайного вектора составляется из математических ожиданий координатных случайных функций, рассмотрим свойства математического ожидания для одномерного случая.

Пусть $\xi(\omega)$, $\eta(\omega)$ — случайные величины, имеющие математические ожидания, а $\alpha, \beta \in \mathbb{R}$ — произвольные постоянные.

1. Если $\xi(\omega) \geq 0$ (т.е. принимает только неотрицательные значения), то $\mathbf{M}[\xi(\omega)] \geq 0$;
2. $\mathbf{M}[\alpha\xi(\omega) + \beta\eta(\omega)] = \alpha\mathbf{M}[\xi(\omega)] + \beta\mathbf{M}[\eta(\omega)]$;
3. $|\mathbf{M}[\xi(\omega)]| \leq \mathbf{M}[|\xi(\omega)|]$;
4. Если $\xi(\omega)$, $\eta(\omega)$ — независимые случайные величины, то $\mathbf{M}[(\xi(\omega)\eta(\omega)] = \mathbf{M}[\xi(\omega)]\mathbf{M}[\eta(\omega)]$;
5. $(\mathbf{M}[\xi(\omega)\eta(\omega)])^2 \leq \mathbf{M}[\xi^2(\omega)]\mathbf{M}[\eta^2(\omega)]$.

Условное математическое ожидание

Рассмотрим математическое ожидание части компонент случайного вектора с использованием их условной плотности распределения вероятностей относительно других компонент.

Например, пусть $\eta(\omega)$ — n-мерный, $\varepsilon(\omega)$ — m-мерный случайные векторы и $f(x \,|\, y)$ — условная плотность распределения $\eta(\omega)$ при условии, что $\varepsilon(\omega)$ принял значение y.

Определение 1.15. Значением $\mathbf{M}[\eta(\omega) \,|\, y]$ условного математическим ожидания случайного вектора $\eta(\omega)$ при условии $\varepsilon(\omega) = y$ называют m-мерный вектор

$$\mathbf{M}[\eta(\omega) \,|\, y] \triangleq \int_{\mathbb{R}^n} x\, f(x \,|\, y)\, dx.$$

Значение условного математического ожидания $\mathbf{M}[\eta(\omega) \,|\, y]$ является функцией $\varphi(y)$ переменного y, меняющегося в области значений случайного вектора $\varepsilon(\omega)$.

Определение 1.16. Функцию $\varphi(\varepsilon(\omega)) \equiv \mathbf{M}[\eta(\omega) \,|\, \varepsilon(\omega)]$ от случайного вектора $\varepsilon(\omega)$ называют **условным математическим ожиданием** случайного вектора $\eta(\omega)$ при условии $\varepsilon(\omega)$.

Определение 1.17. **Дисперсией** $\mathbf{D}[\xi(\omega)]$ скалярной случайной величины $\xi(\omega)$ называют математическое ожидание случайной величины $(\xi(\omega) - \mathbf{M}[\xi(\omega)])^2$, т.е. число

$$\mathbf{D}[\xi(\omega)] \triangleq \mathbf{M}[(\xi(\omega) - \mathbf{M}[\xi(\omega)])^2].$$

Для случайной величины $\xi(\omega)$ с плотностью распределения $f_\xi(x)$ дисперсия может быть вычислена:

$$\mathbf{D}[\xi(\omega)] = \int_{-\infty}^{\infty} (x - \mathbf{M}[\xi(\omega)])^2 f_\xi(x)\, dx.$$

Свойства дисперсии

1. $\mathbf{D}[\xi(\omega)] \geq 0$;
2. $\mathbf{D}[\alpha\xi(\omega)] = \alpha^2\mathbf{D}[\xi(\omega)]$, $\alpha \in \mathbb{R}$;
3. $\mathbf{D}[\xi(\omega) + \alpha] = \mathbf{D}[\xi(\omega)]$, $\alpha \in \mathbb{R}$;
4. если случайные величины $\xi(\omega)$ и $\eta(\omega)$ независимые, то $\mathbf{D}[\xi(\omega) + \eta(\omega)] = \mathbf{D}[\xi(\omega)] + \mathbf{D}[\eta(\omega)]$;
5. $\mathbf{D}[\alpha] = 0$, т.е. дисперсия случайной величины, имеющей постоянное значение (детерминированной величины), равна нулю;
6. $\mathbf{D}[\xi(\omega)] = \mathbf{M}[\xi^2(\omega)] - (\mathbf{M}[\xi(\omega)])^2$;

Определение 1.18. **Ковариацией** двух случайных величин $\xi(\omega)$ и $\eta(\omega)$ называют число

$$\mathbf{cov}[\xi(\omega),\eta(\omega)] \triangleq \mathbf{M}[(\xi(\omega)-\mathbf{M}[\xi(\omega)])(\eta(\omega)-\mathbf{M}[\eta(\omega)])].$$

$$\mathbf{cov}[\xi(\omega),\eta(\omega)] = \int\int_{\mathbb{R}^2}(x-\mathbf{M}[\xi(\omega)])(y-\mathbf{M}[\eta(\omega)])f_{\xi\eta}(x,y)\,dxdy,$$

где $f_{\xi\eta}(x,y)$ совместная плотность распределения случайных величин $\xi(\omega)$ и $\eta(\omega)$.

Свойства ковариации

а) $\mathbf{cov}[\xi(\omega),\eta(\omega)] \equiv \mathbf{cov}[\eta(\omega),\xi(\omega)]$;
б) $\mathbf{cov}[\xi(\omega),\xi(\omega)] \equiv \mathbf{D}[\xi(\omega)]$;
в) если случайные величины $\xi(\omega)$ и $\eta(\omega)$ независимые, то $\mathbf{cov}[\xi(\omega),\eta(\omega)] = 0$;
г) $\mathbf{cov}[\xi(\omega),\eta(\omega)] = \mathbf{M}[\xi(\omega)\eta(\omega)] - \mathbf{M}[\xi(\omega)]\,\mathbf{M}[\eta(\omega)]$;
д) $2\mathbf{cov}[\xi(\omega),\eta(\omega)] = \mathbf{D}[\xi(\omega)+\eta(\omega)] - \mathbf{D}[\xi(\omega)] - \mathbf{D}[\eta(\omega)]$.

Определение 1.19. **Коэффициентом корреляции** скалярных случайных величин $\xi(\omega)$ и $\eta(\omega)$ называют число

$$\rho[\xi(\omega), \eta(\omega)] \triangleq \frac{\mathbf{cov}[\xi(\omega), \eta(\omega)]}{\sqrt{\mathbf{D}[\xi(\omega)]\,\mathbf{D}[\eta(\omega)]}}.$$

Две случайные величины $\xi(\omega)$ и $\eta(\omega)$ называют **некоррелированными**, если $\rho[\xi(\omega), \eta(\omega)] = 0$, или, что то же самое, $\mathbf{cov}[\xi(\omega), \eta(\omega)] = 0$. ∎

Свойства коэффициента корреляции

1. $|\rho[\xi(\omega), \eta(\omega)]| \leq 1$;
2. $|\rho[\xi(\omega), \eta(\omega)]| = 1$ тогда и только тогда, когда $\eta(\omega) = \alpha\xi(\omega) + \beta$, где $\alpha,\ \beta \in \mathbb{R}$, при $\rho[\xi(\omega), \eta(\omega)] = 1$ $\alpha > 0$
при $\rho[\xi(\omega), \eta(\omega)] = -1$ $\alpha < 0$.

Определение 1.20. **Ковариационной матрицей случайного вектора** $\xi(\omega) \triangleq (\xi_1(\omega) \ldots \xi_n(\omega))^{\mathbb{T}}$ называют матрицу

$$\mathbf{cov}[\xi(\omega)] \triangleq (\mathbf{cov}\,[\xi_i(\omega), \xi_j(\omega)])\,.$$

Ковариационная матрица n -мерного случайного вектора представляет собой квадратную матрицу порядка n из ковариаций пар координатных случайных функций. ▮

Свойства ковариационной матрицы

1. $\mathbf{cov}[\xi(\omega)]$ — симметрическая неотрицательно определенная квадратная матрица порядка n ;

2. если $\eta(\omega) = A\xi(\omega) + b$, где $A \in M_{mn}(\mathbb{R})$ — фиксированная матрица, $b \in \mathbb{R}^m$ — фиксированный вектор, то $\mathbf{cov}[\eta(\omega)] = A\mathbf{cov}[\xi(\omega]A^{\mathbb{T}}$.

Определение 1.21. **Характеристической функцией** n -мерного **случайного вектора** ξ называют функцию

$$\varphi_\xi(\lambda) \triangleq \mathbf{M}[\exp(i\lambda^{\mathbb{T}}\xi(\omega))],$$

где $\lambda \in \mathbb{R}^n$, а i — мнимая единица. ▮

Характеристическая функция является изображением *экспоненциального интегрального преобразования Фурье*.

$$\varphi_\xi(\lambda) = \int_{\mathbb{R}^n} \exp(i\lambda^{\mathbb{T}}x) f_\xi(x)\, dx,$$

В классе *функций, интегрируемых с квадратом*, существует взаимно однозначное соответствие между характеристическими функциями и плотностями распределения.

Свойства характеристической функции

1. $|\varphi_\xi(\lambda)| \leq 1$.
2. $\varphi_\xi(\mathbf{O}) = \mathbf{1}$.
3. $\varphi_\xi(-\lambda) = \varphi_\xi^*(\lambda)$, при изменении знака аргумента значение характеристической функции меняется на комплексно сопряженное.
4. Если $\eta(\omega) = A\xi(\omega) + b$, где матрица $A \in M_{mn}(\mathbb{R})$ и вектор $b \in \mathbb{R}^m$ фиксированы, то $\varphi_\eta(\mu) = \exp(i\mu^{\mathbb{T}}b)\varphi_\xi(A^{\mathbb{T}}\mu)$.
5. Если случайные величины $\xi_k(\omega)$, $k = \overline{1,\ N}$, независимы в совокупности, $\eta(\omega) = \xi_1(\omega) + \ldots + \xi_n(\omega)$, то $\varphi_\eta(\lambda) = \prod\limits_{k=1}^{N} \varphi_{\xi_k}(\lambda)$, где $\lambda = (\lambda_1 \ldots \lambda_n)^{\mathbb{T}}$.

1.3. Элементы математической статистики

Математическая статистика — раздел математики, который занимается разработкой методов получения научно обоснованных выводов о массовых явлениях и процессах по данным наблюдений или экспериментов.

Определение 1.22. В математической статистике множество возможных значений случайной величины X называют **генеральной совокупностью** случайной величины X или просто генеральной совокупностью X.

Исходным материалом для изучения свойств генеральной совокупности (т.е. некоторой случайной величины) являются **экспериментальные (статистические) данные**, под которыми понимают значения случайной величины, полученные в результате повторений случайного эксперимента (наблюдений над случайной величиной)

Эксперимент (хотя бы теоретически) может быть повторен сколько угодно раз в одних и тех же условиях.

Распределение случайной величины X_i, $i = 1, 2, \ldots$, заданной на множестве исходов i-го эксперимента, не зависит от номера испытания и совпадает с распределением генеральной совокупности X (**независимые повторные эксперименты**, **испытания**, **независимые повторные наблюдения**).

Определение 1.23. Случайная выборка из генеральной совокупности X $\vec{X}_n = (X_1, \ldots, X_n)$ (X_1 , …, X_n). — это совокупность независимых случайных величин X_1 , …, X_n , каждая из которых имеет то же распределение, что и случайная величина X , где n — **объем случайной выборки**, случайные величины X_i — **элементы случайной выборки**.

Любое возможное значение $\vec{x}_n = (x_1, \ldots, x_n)$ случайной выборки $\vec{X}_n$ — **выборка** из генеральной совокупности X (**реализация случайной выборки** $\vec{X}_n$).

Выборка $\vec{x}_n$ — совокупность n чисел x_1 , …, x_n , полученных в результате проведения n повторных независимых наблюдений над случайной величиной X .

Основой любых выводов о вероятностных свойствах генеральной совокупности X , т.е. **статистических выводов**, является **выборочный метод**, суть которого заключается в том, что свойства случайной величины X устанавливаются путем изучения тех же свойств на случайной выборке.

Множество возможных значений случайной выборки $\vec{X}_n$ называют **выборочным пространством** и обозначают $\mathcal{X}_n$ (n -мерное линейное арифметическое пространство $\mathbb{R}^n$, или его подмножество).
Если X — дискретная случайная величина, то выборочное пространство — конечное или счетное.
Элементы X_i , $i = \overline{1,\ n}$, случайной выборки $\vec{X}_n$ независимы и имеют то же распределение, что и генеральная совокупность X . Функция распределения $F_{\vec{X}}(t_1, \ldots, t_n)$ случайной выборки $\vec{X}_n$ имеет вид

$$F_{\vec{X}}(t_1, \ldots, t_n) = \mathbf{P}\left\{X_1 < t_1,\ \ldots,\ X_n < t_n\right\} = \\ = \prod_{i=1}^{n} \mathbf{P}\left\{X_i < t_i\right\} = \prod_{i=1}^{n} F(t_i),$$

где $F(t)$ — функция распределения случайной величины X (генеральной совокупности X). ▮
О распределении случайной величины X могут быть самые общие представления. Например, X является непрерывной случайной величиной, о распределении практически ничего не известно. Или функция распределения известна, но не известны параметры, от которых она зависит.
Часто можно говорить только о семействе (классе) $\mathcal{P}$ распределений случайной выборки, в котором содержится **априорная информация**.

Статистическая модель — выборочное пространство, на котором задан класс распределений $\mathcal{P}$, обозначается $\{F(x)\}$.

Статистическую модель называют **параметрической**, если функция (плотность) распределения вероятности задана с точностью до неизвестного вектора параметров $\vec{\theta} = (\theta_1 \dots \theta_r)$ множеством возможных значений Θ, т.е. $\vec{\theta} \in \Theta$. Обозначение: $\left\{F(x;\vec{\theta});\ \vec{\theta} \in \Theta\right\}$.

Множество Θ называют **параметрическим множеством**.

Статистическую модель называют **непрерывной** (**дискретной**), если случайная величина X является непрерывной (дискретной).

Если генеральная совокупность X является дискретной случайной величиной, распределение X задают в виде таблицы (ряда распределений).

Если генеральная совокупность X является непрерывной случайной величиной распределение X задают в виде плотности распределения $p_X(x)$.

Любую функцию $g(\vec{X}_n)$ случайной выборки в математической статистике называют **статистикой (выборочной характеристикой)**.

Распределение этой случайной величины называют **выборочным распределением**.

Выборочное распределение однозначно определяется совместным распределением случайных величин $X_1, \ldots, X_n$, т.е. распределением случайной выборки $\vec{X}_n$.

Значение $g(\vec{x}_n)$ выборочной характеристики $g(\vec{X}_n)$, определенное по реализации $\vec{x}_n$ случайной выборки $\vec{X}_n$, называют ее **выборочным значением**.

Сходимость по вероятности и сходимость по распределению (слабая сходимость).

Определение 1.24. Последовательность $\{Y_n\}$ случайных величин называют сходящейся по вероятности к Y ($Y_n \xrightarrow{\mathbf{P}} Y$), если для любого $\varepsilon > 0$

$$\lim_{n\to\infty} \mathbf{P}|Y_n - Y| > \varepsilon \to 0 .$$

Определение 1.25. Последовательность $F_{Y_n}(x)$ функций распределения называют сходящейся по распределению (слабо сходящейся) $F_{Y_n}(x) \underset{n\to\infty}{\Rightarrow} F_Y(x)$, если имеет место равенство

$$\lim_{n\to\infty} F_{Y_n}(x) = F_Y(x)$$

в каждой точке непрерывности $F_Y(x)$.

Основные задачи математической статистики

Источниками информации для решении задачи математической статистики.

1. Результаты наблюдений (эксперимента) в виде выборки из некоторой генеральной совокупности скалярной (векторной) случайной величины.
2. Априорная информация о свойствах изучаемого объекта.

1. Задача оценивания неизвестных параметров

Функция распределения генеральной совокупности известна с точностью до параметра θ;

Определение 1.26. **Оценкой** $\tilde{\theta}$ параметра θ называют всякую функцию результатов наблюдений над случайной величиной(СВ) X (статистику), с помощью которой судят о значении θ. $\tilde{\theta}_n = \tilde{\theta}_n(X_1, X_2, ..., X_n)$

$X_1, X_2, ..., X_n$ — СВ, $\tilde{\theta}$ — СВ, закон распределения $\tilde{\theta}$ зависит от закона распределения случайной величины X и от n.

2. Проверка статистических гипотез

Определение 1.27. Статистической гипотезой называют любое предположение о распределении вероятностей (законе распределения и/или параметрах закона распределения вероятностей) наблюдаемой случайной величины (скалярной или векторной).

3. Установление формы и степени связи между случайными величинами (корреляционный, дисперсионный, регрессионный анализ и др.)

Задача оценивания неизвестных параметров

Точечные оценки

Необходимо найти такую **статистику** $\widehat{\theta}(\vec{X}_n)$, **выборочное значение** $\widehat{\theta} = \widehat{\theta}(\vec{x}_n)$ которой для рассматриваемой **реализации** $\vec{x}_n$ **случайной выборки** можно было бы считать приближенным значением параметра $\vec{\theta}$. Статистику $\widehat{\theta}(\vec{X}_n)$, выборочное значение $\widehat{\theta}$ которой для любой реализации $\vec{x}_n$ принимают за приближенное значение неизвестного параметра θ, называют его **точечной оценкой (оценкой)**, а $\widehat{\theta}$ — **значением точечной оценки (оценки)**.

Требование: выборочное значение $\widehat{\theta}$ ($\widetilde{\theta}(\vec{X}_n)$, $\theta^*(\vec{X}_n)$) соответствовало истинному значению параметра $\vec{\theta}$.

Свойства точечных оценок

Состоятельные, несмещенные и эффективные оценки.

Определение 1.28. Статистику $\widehat{\theta}(\vec{X}_n)$ называют **состоятельной оценкой** параметра $\theta \in \Theta$, если с ростом *объема выборки* n она сходится по вероятности к оцениваемому параметру θ, т.е.

$$\widehat{\theta}(\vec{X}_n) \xrightarrow[n\to\infty]{\mathbf{P}} \theta.$$

Для состоятельной оценки $\widehat{\theta}(\vec{X}_n)$ отклонение ее от θ на величину ε и более становится маловероятным при большом объеме выборки.

Определение 1.29. Статистику $\widehat{\theta}(\vec{X}_n)$ называют **несмещенной оценкой** параметра θ, если ее математическое ожидание совпадает с θ, т.е. $\mathbf{M}\widehat{\theta}(\vec{X}_n) = \theta$ для любого фиксированного n.

Оценка не дает систематической погрешности относительно истинного значения параметра θ.

оценка $\widehat{\theta}(\vec{X}_n)$ является **асимптотически несмещенной**, если при $n \to \infty$ она сходится по вероятности к своему математическому ожиданию, т.е. для любого $\varepsilon > 0$

$$\lim_{n\to\infty} \mathbf{P}\left|\widehat{\theta}(\vec{X}_n) - \mathbf{M}\widehat{\theta}(\vec{X}_n)\right| < \varepsilon = 1.$$

Определение 1.30. Если в некотором классе несмещенных оценок параметра θ, имеющих конечную дисперсию, существует такая оценка $\widehat{\theta}(\vec{X}_n)$, что для всех оценок $\widetilde{\theta}(\vec{X}_n)$ из этого класса, выполняется неравенство

$$\mathbf{D}\widehat{\theta}(\vec{X}_n) \le \mathbf{D}\widetilde{\theta}(\vec{X}_n),$$

то **оценку** $\widehat{\theta}(\vec{X}_n)$ называют **эффективной в данном классе оценок**.

Дисперсия эффективной оценки параметра в некотором классе является минимальной среди дисперсий всех оценок из рассматриваемого класса несмещенных оценок.

Теорема 1. Оценка $\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ (**выборочное среднее**) математического ожидания $\theta = \mathbf{M}X = \mu$ генеральной совокупности X с конечной дисперсией является несмещенной, состоятельной и эффективной в классе всех **линейных оценок**, т.е. оценок вида $\widetilde{\theta}(\vec{X}_n) = \sum_{i=1}^{n} \alpha_i X_i$, где $\sum_{i=1}^{n} \alpha_i = 1$, для произвольной параметрической модели.

Теорема 2. Если $\vec{X}_n$ — случайная выборка из генеральной совокупности X с конечной дисперсией σ^2, то выборочная дисперсия $\widehat{\sigma}^2(\vec{X}_n)$ — смещенная состоятельная оценка σ^2.

Исправленная выборочная дисперсия

Статистика

$$S^2(\vec{X}_n) = \frac{1}{n-1}\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2$$

называется **исправленной выборочной дисперсией**.

Из теоремы 2 следует, что она является несмещенной и состоятельной оценкой дисперсии σ^2 генеральной совокупности.

Теорема 3 (о единственности эффективной оценки). Пусть $\widehat{\theta}(\vec{X}_n)$ и $\widetilde{\theta}(\vec{X}_n)$ — две эффективные оценки для параметра θ рассматриваемой параметрической модели. Тогда

$$\widehat{\theta}(\vec{X}_n) = \widetilde{\theta}(\vec{X}_n),$$

где равенство следует понимать в вероятностном смысле:

$$\mathbf{P}\left\{\vec{X}_n \in \{\vec{x}_n : \widehat{\theta}(\vec{x}_n) \neq \widetilde{\theta}(\vec{x}_n)\}\right\} = 0.$$

Теорема 4 (неравенство Рао — Крамера). Пусть параметрическая модель является *регулярной* и $\widehat{\theta}(\vec{X}_n)$ — несмещенная оценка неизвестного параметра θ . Тогда имеет место неравенство

$$\mathbf{D}\widehat{\theta}(\vec{X}_n) \geq \frac{1}{nI(\theta)}, \tag{1.1}$$

где $I(\theta) = \mathbf{M}\left(\dfrac{\partial \ln p(X;\theta)}{\partial\theta}\right)^2$ — **количество информации по Фишеру** в одном наблюдении, $p(t;\theta)$ — плотность распределения генеральной совокупности X (*непрерывная стат. модель*) или вероятность события $\{X = t\}$ (*дискретная стат. модели*).

$e(\theta) = \dfrac{1}{nI(\theta)\,\mathbf{D}\widehat{\theta}(\vec{X}_n)}$ — **показатель эффективности по Рао — Крамеру**.

Определение 1.31. Несмещенную **оценку** $\widehat{\theta}(\vec{X}_n)$ параметра $\theta \in \Theta \subset \mathbb{R}$ называют **эффективной по Рао — Крамеру**, если показатель эффективности $e(\theta) = 1$.

Понятие достаточных статистик

$\vec{X}_n$ — *случайная выборка* из *генеральной совокупности* X с функцией распределения $F(x;\theta)$, где θ — неизвестный параметр.

$T = T(\vec{X}_n)$ — некоторая *статистика*. Значение $T(\vec{x}_n) = t$ статистики T — известно, *выборка* $\vec{x}_n$ — *реализация случайной выборки* $\vec{X}_n$ не известна.

Определение 1.32. Статистику $T(\vec{X}_n)$ называют **достаточной** для параметра θ, если условная функция распределения $F_{\vec{X}_n}(z_1, \ldots, z_n | T(\vec{X}_n) = t)$ случайной выборки $\vec{X}_n$ при условии $T(\vec{X}_n) = t$ не зависит от параметра θ при любом возможном значении t.

Критерий достаточности статистики

Функцию $L(X_1, \ldots, X_n; \theta) = p(x_1;\theta) \ldots p(x_n;\theta)$, называют **функцией правдоподобия**.

Здесь $p(x;\theta)$ — ПРВ НСВ или вероятность события $\{X = x\}$ в случае ДСВ, X_i, $i = \overline{1, n}$, — *элементы случайной выборки* $\vec{X}_n$.

Теорема 5 (критерий факторизации Неймана — Пирсона). Статистика $T = T(x_1, \dots, x_n)$ является достаточной для параметра θ тогда и только тогда, когда для любой реализации $(x_1, \dots, x_n)$ случайной выборки $(X_1, \dots, X_n)$ *выборочное значение* функции правдоподобия имеет вид

$$L(X_1, \dots, X_n; \theta) = g(T(x_1, \dots, x_n), \theta)\, h(x_1, \dots, x_n), \tag{1.2}$$

т.е. может быть представлено в виде произведения двух сомножителей,первый зависит от результатов наблюдений x_1, ..., x_n только через статистику $T = T(x_1, \dots, x_n)$, второй не зависит от θ.

Всякая *эффективная по Рао — Крамеру оценка* $\widehat{\theta} = \widehat{\theta}(\vec{X}_n)$ параметра θ является достаточной статистикой.

Методы оценивания неизвестных параметров по данным сл. выборки:
(наиболее часто использующиеся)

– *метод моментов*;

– *метод максимального правдоподобия*;

– *метод наименьших квадратов*.

Интервальные оценки

Интервальная оценка позволяет получить вероятностную характеристику точности оценивания неизвестного параметра.

Определение 1.33. Интервал $(\underline{\theta}(\vec{X}_n),\ \overline{\theta}(\vec{X}_n))$ называют **доверительным интервалом** для θ с **коэффициентом доверия (уровнем доверия)** γ, если выполнено неравенство $\mathbf{P}\{\underline{\theta}(\vec{X}_n) \leq \theta \leq \overline{\theta}(\vec{X}_n)\} = \gamma$.

Интервальная оценка $(\underline{\theta}(\vec{X}_n),\ \overline{\theta}(\vec{X}_n))$ представляет собой интервал со случайными границами, который с заданной вероятностью γ накрывает неизвестное истинное значение параметра θ.
Для любой реализации $\vec{x}_n$ случайной выборки $\vec{X}_n$, вероятностной характеристикой точности оценивания параметра θ является длина интервала, т.е. случайная величина $l(\vec{X}_n) = \overline{\theta}(\vec{X}_n) - \underline{\theta}(\vec{X}_n)$.
Статистику $\underline{\theta}(\vec{X}_n)$ ($\overline{\theta}(\vec{X}_n)$) называют **односторонней нижней (односторонней верхней)** γ-доверительной границей, если выполнено неравенство $\mathbf{P}\{\underline{\theta}(\vec{X}_n) < \theta\} = \gamma$ ($\mathbf{P}\{\theta < \overline{\theta}(\vec{X}_n)\} = \gamma$).

Алгоритм построения доверительного интервала

– построение центральной статистики $T(\vec{X}_n,\theta)$ с известной функцией распределения $F_T(t)$;

– представление заданного коэффициента доверия γ в виде $\gamma=1-\alpha-\beta$;

– нахождение квантилей h_α и $h_{1-\beta}$ уровня α и $1-\beta$ функции распределения $F_T(t)$;

– нахождение значений *нижней* $\underline{\theta}(\vec{x}_n)$ и *верхней* $\overline{\theta}(\vec{x}_n)$ *границ* искомой интервальной оценки.

Если $T(\vec{x}_n,\theta)$ — возрастающая функция параметра θ, то путем решения уравнений

$$T(\vec{x}_n,\underline{\theta})=h_\alpha,\qquad T(\vec{x}_n,\overline{\theta})=h_{1-\beta} \tag{1.3}$$

Если $T(\vec{x}_n,\theta)$ — убывающая функция параметра θ, то путем решения уравнений

$$T(\vec{x}_n,\underline{\theta})=h_{1-\beta},\qquad T(\vec{x}_n,\overline{\theta})=h_\alpha \tag{1.4}$$

2. Проверка статистических гипотез

Статистической гипотезой называют любое предположение о виде или параметрах неизвестного законе распределении.

Проверяемая гипотеза H_0 (нулевая). Альтернативная гипотеза H_1.

Если гипотеза не противоречит статистическим данным и ее следует принять, если гипотеза противоречит — отклонить

Если имеются предположения об известном законе распределения генеральной совокупности, то статистические гипотезы относительно неизвестного истинного значения параметра θ называют **параметрическими гипотезами**. Никакой *априорной информацией* относительно параметра θ нет.

Если θ — скаляр, то гипотеза называется **однопараметрической**, а если θ — вектор, то **многопараметрическая**.

Если закон распределения генеральной совокупности неизвестен, то гипотезы называют **непараметрическими**.

Основные типы гипотез:

1) о равенстве числовых характеристик генеральных совокупностей ($\mu_1 = \mu_2$, $\sigma_1^2 = \sigma_2^2$);
2) о числовых значениях ($\mu = \mu_0$, $\sigma^2 = \sigma_0^2$);
3) о виде закона распределения ($F(x) = F_T(x)$, $F(x)$ — неизвестная ФР СВ X, $F_T(x)$ — теоретическая ФР);
4) проверка гипотез об однородности выборок;
5) о стохастической независимости выборок.

Основные понятия

Статистическую гипотезу H называют **простой**, если она имеет вид $H:\vec{\theta}=\vec{\theta}_0$, где $\vec{\theta}_0$ — некоторое заданное значение параметра.

Статистическую гипотезу называют **сложной**, если она имеет вид $H:\vec{\theta}\in D$, где D — некоторое множество значений параметра θ, состоящее более чем из одного элемента.

Проверка двух простых гипотез

Проверяются две *простые статистические гипотезы* вида $H_0:\theta=\theta_0,\qquad H_1:\theta\neq\theta_0$, где θ_0, θ_1 — два заданных (различных) значения параметра. H_0 — основная, H_1 — альтернативная гипотеза.

По данным *выборки* $\vec{x}_n$ необходимо принять решение о справедливости одной из указанных гипотез.

Критерием (статистическим критерием), проверки гипотез называют правило, по которому по данным выборки $\vec{x}_n$ принимается решение о справедливости либо гипотезы H_0, либо гипотезы H_1.

Критерий задают с помощью **критического множества** W являющегося подмножеством *выборочного пространства* $\mathcal{X}_n$ *случайной выборки* $\vec{X}_n$.

Все *выборочное пространство* разбивается на два непересекающихся подмножества: W —критическую область (область отклонения гипотезы H_0) и $\overline{W}$ (дополнение множества W) —область допустимых значений (область принятия гипотезы H_0).

Решение :

1) выборка $\vec{x}_n$ принадлежит W — принимают альтернативную гипотезу H_1, основную гипотезу H_0 отвергают;
2) выборка $\vec{x}_n$ принадлежит $\overline{W}$ — принимают основную гипотезу H_0, альтернативную гипотезу H_1 отвергают,.

Ошибки

1) **ошибка первого рода** — принять гипотезу H_1, но верна H_0;
2) **ошибка второго рода** — принять гипотезу H_0, но верна H_1.

Вероятность совершения ошибок:

1) первого рода $\alpha = \mathbf{P}\{\vec{X}_n \in W \mid H_0\} = \int\limits_{W}\!\!\cdots\!\!\int \prod\limits_{k=1}^{n} p(t_k;\theta_0)\,dt_1\ldots dt_n$,

2) второго рода $\beta = \mathbf{P}\{\vec{X}_n \in \overline{W} \mid H_1\} = \int\limits_{\overline{W}}\!\!\cdots\!\!\int \prod\limits_{k=1}^{n} p(t_k;\theta_1)\,dt_1\ldots dt_n$,

где $\mathbf{P}\{A \mid H_j\}$ — вероятность события A при условии, что справедлива гипотеза H_j, $j = 0,\,1$, $\prod_{k=1}^{n} p(t_k;\theta_0)$ — ФПРВ случайной выборки $\vec{X}_n$
Уровень значимости критерия — вероятность совершения ошибки первого рода α.

Критерий Неймана — Пирсона

Вероятность $1-\beta$ **не** допустить ошибку второго рода (отвергнуть H_0, когда она неверна) называется **мощностью** критерия.

Требования к критической области

$$\mathbf{P}(\theta_n \in W/H_0) = \alpha \qquad \mathbf{P}(\theta_n \in W/H_1) = \max$$

Вероятность попадания статистики θ в критическую область W — минимальна и равна α, если верна H_0 и максимально в противоположном случае (при заданном уровне значимости α мощность критерия $1-\beta$ — максимальна).

Лемма 1. Критерий Неймана — Пирсона. Среди всех критериев заданного уровня значимости α, проверяющих простую гипотезу $H_0 : \theta = \theta_0$ против альтернативной гипотезы $H_1 : \theta = \theta_1$, критерий отношения правдоподобия $\varphi(\vec{X}_n) = \dfrac{L(\vec{X}_n;\theta_1)}{L(\vec{X}_n;\theta_0)}$, где $L(\vec{X}_n;\theta) = \prod_{i=1}^n p(X_i;\theta)$, является наиболее мощным.

Для построения **оптимального** (наиболее мощного) при заданном уровне значимости α **критерия Неймана — Пирсона** в *критическое множество* W включают элементы $\vec{x}_n$ *выборочного пространства* $\mathcal{X}_n$ случайной выборки $\vec{X}_n$, для которых выполняется неравенство $\varphi(\vec{x}_n) \geq C_\varphi$, константу C_φ выбирают из условия $\mathbf{P}\{\varphi(\vec{X}_n) \geq C_\varphi | H_0\} = \int \underset{\varphi(t_1,\ldots,t_n)\geq C_\varphi}{\ldots} \int L(t_1, \ldots, t_n; \theta_0)\, dt_1 \ldots dt_n = \alpha$, которое обеспечивает заданное значение уровня значимости α.

Вероятность ошибки второго рода не может быть уменьшена при данном значении вероятности ошибки первого рода α.

Определение объема выборки

Определить объем n^* случайной выборки, при котором может быть построен критерий для проверки двух простых гипотез $H_0: \theta = \theta_0$ и $H_1: \theta = \theta_1$ с заданными или меньшими значениями вероятностей α и β совершения *ошибок первого* и *второго рода* соответственно.

Величину n^* определяют как минимальное целое значение n, для которого система неравенств 1.5 может быть выполнена при некотором значении константы $C = C^*$.

$$\begin{cases} \mathbf{P}\{\varphi(X_1, \ldots, X_n) \geq C_\varphi | \theta = \theta_0\} \leq \alpha, \\ \mathbf{P}\{\varphi(X_1, \ldots, X_n) < C_\varphi | \theta = \theta_1\} \leq \beta. \end{cases} \tag{1.5}$$

Критическое множество определяемое в соответствии с оптимальным критерием Неймана — Пирсона, с заданными значениями α, β будет задаваться неравенством $\varphi(x_1, \ldots, x_n) \geq C^*_\varphi$.

Сложные параметрические гипотезы

Даны две *сложные гипотезы*

$$H_0 : \theta \in \Theta_0, \qquad H_1 : \theta \in \Theta_1, \tag{1.6}$$

где Θ_0, Θ_1 — некоторые непересекающиеся области значений параметра θ. Например, $\Theta_0 : \theta \leq \theta_0$ и $\Theta_1 : \theta \geq \theta_1$, где θ_0 и θ_1 — некоторые фиксированные значения параметра, удовлетворяющие неравенству $\theta_0 < \theta_1$.

Критерий проверки сложных гипотез (1.6) задается с помощью *критического множества* W *реализаций случайной выборки* $\vec{X}_n$.

Решение :

– реализация $\vec{x}_n$ сл. выборки $\vec{X}_n \in W$, H_0 отвергают, H_1 принимают;

– реализация $\vec{x}_n$ сл. выборки $\vec{X}_n \notin W$, H_1 отвергают, H_0 принимают.

Вероятность совершения ошибок:

1) первого рода $\alpha(\theta) = \mathbf{P}\{(X_1, \ldots, X_n) \in W|\theta\}, \quad \theta \in \Theta_0;$,

2) второго рода $\beta(\theta) = \mathbf{P}\{(X_1, \ldots, X_n) \in \overline{W}|\theta\}, \quad \theta \in \Theta_1$,

Величины $\alpha(\theta)$, $\beta(\theta)$ являются функциями от параметра θ.

Максимально возможное значение вероятности совершения ошибки первого рода $\alpha = \max\limits_{\theta \in \Theta_0} \alpha(\theta)$ называют **размером критерия**.

Непараметрические гипотезы

Предположения о виде статистической модели можно сформулировать как статистические гипотезы и проверить при помощи *статистических критериев* на основании *статистических данных*.

Критерии согласия. Простая гипотеза

Критериями согласия называют статистические критерии, предназначенные для обнаружения расхождений между гипотетической статистической моделью и реальными данными (насколько предположения о распределении случайных величин соответствуют экспериментальным данным).

Предположения о распределении случайных величин (виде закона распределения) может быть выдвинуто исходя из теоретических предпосылок, результатов экспериментов, гистограмм и др.

Стандартные гипотезы. Нулевая гипотеза $H_0 : F_n(x) = F(x)$.
Альтернативная гипотеза $H_1 : F_n(x) \neq F(x)$.
$F(x)$ — теоретическая ФРВ, $F_n(x)$ — эмпирическая ФРВ.

Теорема 6. Пусть $\widehat{R}(t, \vec{X}_n)$ — выборочная ФРВ, построенная по сл. выборке $\vec{X}_n$ объема n из ГС с равномерным законом распределения на $[0, 1]$. Если H_0 — истина ФРВ СВ, то $D(\vec{X}_n)$ совпадает с ФРВ СВ $\sup\limits_{0 \le t \le 1} |\widehat{R}(t, \vec{X}_n) - t|$.

Критерий Колмогорова.

$\vec{X}_n$ — случайная выборка объема n из генеральной совокупности (ГС) **непрерывной** СВ X. $H_0 : F_n(x) = F(x)$, $H_1 : F_n(x) \neq F(x)$, где $F(x)$ — теоретическая ФРВ, $F_n(x)$ — эмпирическая ФРВ.

Мера расхождения между теоретическим и эмпирическим распределением — максимальное значение абсолютной разности между $F(x)$ и $F_n(x)$:

$$D_n = \sup_{-\infty<x<\infty} |F(x) - F_n(x)| .$$

Распределение статистики критерия D_n при гипотезе H_0 не зависит от вида ФРВ $F(x)$ (теор.6). Если $F(x)$ **непрерывна**, то при $\forall\lambda > 0$

$$\lim_{n\to\infty} P\left(\sqrt{n}D_n \le \lambda\right) = K(\lambda) = \sum_{j=-\infty}^{\infty} \left((-1)^j \exp(-2j^2\lambda^2)\right) .$$

Величина $\sup\limits_{-\infty<x<\infty} |F(x) - F_n(x)|$ имеет порядок $1/\sqrt{n}$, в качестве статистики используют величину $\sqrt{n}D_n$.

Предельную ФРВ $K(\lambda)$ можно использовать при $n \ge 20$, критическую границу полагать равной $\lambda_\alpha/\sqrt{n}$. Критическое значение λ_α получают из соотношения $1 - K(\lambda_\alpha) = \alpha$.

Критерий Колмогорова ”хорошо различает” сильно отличающиеся друг от друга ФРВ $F(x)$ и $F_n(x)$ (даже на небольшом интервале).

Критерий ω^2

$\vec{X}_n$ — случайная выборка объема n из ГС **непрерывной** СВ X.
Если $F_n(x) \neq F(x)$ на достаточно большом промежутке, но максимальное значение абсолютной разности между $F(x)$ и $F_n(x)$ невелико, то для проверки стандартных гипотез H_0 и H_1 применяют ω^2, использующий статистику

$$\omega^2(\vec{X}_n) = \frac{1}{12n^2} + \frac{1}{n}\sum_{i=1}^{n}\left(F_0(X_{(i)}) - \frac{2i-1}{2n}\right)^2, \qquad (1.7)$$

где $X_{(i)}$, $i = \overline{1,\, n}$, — элементы вариационного ряда случайной выборки $X_1, \ldots, X_n$. Гипотеза H_0 отклоняется в пользу H_1 на уровне значимости α, если $\omega^2(\vec{x}_n) > \omega^2_{1-\alpha}$, где $\omega^2_{1-\alpha}$ — квантиль уровня $1-\alpha$ распределения статистики $\omega^2(\vec{X}_n)$ при условии истинности гипотезы H_0.
Распределение статистики $\omega^2(\vec{X}_n)$ при истинности основной гипотезы H_0 не зависит от F_0.
Таблицы квантилей статистики $\omega^2(\vec{X}_n)$ для малых n существуют. Для больших n — таблицы предельного распределения статистики $n\omega^2(\vec{X}_n)$.

Критерий согласия χ^2 (критерий Пирсона).

Пусть X — дискретная случайная величина, принимающая r различных значений $u_1, \ldots, u_r$ с положительными вероятностями $p_1, \ldots, p_r$:

$$\mathbf{P}\{X = u_k\} = p_k, \quad k = \overline{1, r}, \quad \sum_{k=1}^{r} p_k = 1.$$

Пусть в выборке $\vec{x}_n = (x_1, \ldots, x_n)$ число u_k встретилось $n_k(\vec{x}_n)$ раз, $k = \overline{1, r}$, случайные величины $n_1(\vec{X}_n), \ldots, n_r(\vec{X}_n)$ зависимы $\left(\sum_{k=1}^{r} n_k(\vec{x}_n) = n\right)$.

Теорема 7 (теорема Пирсона). Распределение случайной величины $\sum_{k=1}^{r} \dfrac{(n_k(\vec{X}_n) - np_k)^2}{np_k}$ при $n \to \infty$ слабо сходится к χ^2 *-распределению* с $r-1$ степенями свободы.

Простые гипотезы:

$H_0 : p_1 = p_{10}, \ldots, p_r = p_{r0}$,

H_1 : существуют такие k, что $p_k \neq p_{k0}$, $k = \overline{1, r}$

где $p_{10}, \ldots, p_{r0}$ — известные величины.

1. Если H_0 —**истинна**, при $n \to \infty$ распределение СВ (теор. 7)

$$\chi^2(\vec{X}_n) = \sum_{k=1}^{r} \frac{(n_k(\vec{X}_n) - np_{k0})^2}{np_{k0}} = n \sum_{k=1}^{r} \frac{\left(\frac{n_k(\vec{X}_n)}{n} - p_{k0}\right)^2}{p_{k0}} \qquad (1.8)$$

стремится к распределению χ^2 с $r-1$ степенями свободы.

2. Если H_0 — **ложная**, то при $n \to \infty$ $\dfrac{n_k(\vec{X}_n)}{n} \to p_k, \quad k = \overline{1, r}$ (по закону больших чисел).

При $n \to \infty$

$$\frac{n_k(\vec{X}_n)}{n} - p_{k0} = \left(\frac{n_k(\vec{X}_n)}{n} - p_k\right) + (p_k - p_{k0}) \to p_k - p_{k0}.$$

Если $\exists k = \overline{1, r} | p_k - p_{k0} \neq 0$, то статистика $\chi^2(\vec{X}_n)$ принимает большие значения, чем в случае **1.**

Критерий согласия χ^2.

При больших n на уровне значимости α гипотезу H_0 **отклоняют** в пользу гипотезы H_1, если $\chi^2(\vec{x}_n) > \chi^2_{1-\alpha}(r-1)$,

где $\chi^2_{1-\alpha}(r-1)$ — квантиль уровня $1-\alpha$ χ^2-распределения с $r-1$ степенями свободы, $\chi^2(\vec{x}_n)$ — реализация случайной величины (1.8).

Гипотезу H_0 **принимают**, если $\chi^2(\vec{x}_n) \leq \chi^2_{1-\alpha}(r-1)$,

Критерии согласия. Сложная гипотеза

Критерии Колмогорова и ω^2 для сложной гипотезы

Задача:проверить по случайной выборке $\vec{X}_n$ из генеральной совокупности X сложную гипотезу о принадлежности ФРВ $F(x)$ СВ X заданному **параметрическому множеству (семейству)** распределений $F(x;\theta)$, $\theta \in \Theta$, $\Theta \subset \mathbb{R}^d$:

$H_0 : F_n(x) = F(x;\theta), \quad \theta \in \Theta$.

Для критериев Колмогорова и ω^2 достаточно иметь только одну таблицу для каждого семейства (нормальное, t-распределение...). При небольшой модификации статистик $\widehat{D}(\vec{X}_n)$ и $\widehat{\omega}^2(\vec{X}_n)$ их распределение при $n \geq 5$ практически перестает зависеть от n.

$H_0 : F_n(x) = F(x;\widehat{\theta}(\vec{x}_n)$,

где $\widehat{\theta}(\vec{x}_n)$ — значение оценки $\widehat{\theta}(\vec{X}_n)$ параметра θ полученные **методом максимального правдоподобия (ММП)**, элементы $F(x;\theta)$ параметрического множества $\{F(x;\theta), \ \theta \in \Theta\}$ ФРВ получаются при помощи преобразования сдвига и масштаба какого-нибудь одного своего представителя $F(x;\theta)$, т.е. $F(x;\theta) = F(\dfrac{x-a}{b}, \theta_0),$.

Критерий χ^2 Пирсона для сложной гипотезы

Пусть ФРВ ДСВ $X = \{u_1, \dots, u_r\}$, зависит от d-мерного вектора параметров θ. $p_k = p_k(\theta)$, $k = \overline{1,\, r}$ — вероятность того, что X примет значение u_k.

Вероятности $p_1(\theta), \dots, p_r(\theta)$ полностью определяют ФРВ СВ X. Основная гипотеза: $H_0 : P(X = u_k) = p_k(\theta), \quad k = \overline{1,\, r}, \quad \theta \in \Theta \subset \mathbb{R}^d$.

Пусть $\widehat{\theta}(\vec{x}_n)$ — значение оценки параметра $\theta(\vec{X}_n)$, полученная ММП для θ, $n_k(\vec{x}_n)$ — количество элементов выборки $\vec{x}_n$, равных u_k, $k = \overline{1,\, r}$.

Оценку $\widehat{\theta}(\vec{X}_n)$ получают в результате минимизации логарифма *функции правдоподобия*

$$L(\vec{X}_n;\theta) = \frac{n!}{n_1! \dots n_r!} \prod_{k=1}^{r} p_i^{n_k(\vec{X}_n)}(\theta), \quad \sum_{i=1}^{r} n_i(\vec{X}_n) = n,$$

как решение системы уравнений

$$\sum_{k=1}^{r} \frac{n_k(\vec{X}_n)}{p_k(\theta)} \frac{\partial p_k(\theta)}{\partial \theta_j} = 0, \quad j = \overline{1,\, d}.$$

Распределение СВ при $n \to \infty$

$$\chi^2(\vec{X}_n) = \sum_{i=1}^{r} \frac{(n_i(\vec{X}_n) - np_i(\widehat{\theta}(\vec{X}_n)))^2}{np_i(\widehat{\theta}(\vec{X}_n))}$$

слабо сходится к СВ, имеющей χ^2 *-распределение* с $r - d - 1$ степенями свободы.

Если X — НСВ с ФРВ $F(x)$, то множество возможных значений X разбивают на конечное число непересекающихся подмножеств и переходят к ДСВ X'.

$$H_0 : F(x) \in \{F(t;\theta),\ \theta \in \Theta \subset \mathbb{R}^d\}.$$

Оценку $\widehat{\theta}(\vec{X}_n)$ строят ММП по значениям частот $n_1(\vec{x}\,'_n), \ldots, n_r(\vec{x}\,'_n)$ СВ X' (не по наблюдениям $X_1, \ldots, X_n$ СВ X).

3. Установление формы и степени связи между случайными величинами

Пусть Y — СВ, поведение которой необходимо определить по значениям СВ X_1 и X_2.

Например, Y — это степень шума двигателя автомашины, а X_1 и X_2 — соответственно величина пробега автомобиля и вес груза в нем.

Корреляционный и дисперсионный анализ позволяет дать ответ на вопрос: есть ли связь между X_1, X_2 и Y и насколько она существенна.

На основе регрессионного анализа по имеющимся статистическим данным можно построить регрессионную модель .

$y = \varphi(x_1, x_2)$,

где y — среднее значение шума Y в зависимости от значений x_1 и x_2 случайных величин X_1 и X_2.

Такая модель позволяет в дальнейшем различные задачи.

Задача

Трехмерный случайный вектор $\vec{X} = (X_1, X_2, X_3)$ распределен по нормальному закону с вектором средних значений $\vec{m} = (2{,}5,\ 1,\ 2)$ и матрицей ковариаций

$$\Sigma = \begin{pmatrix} 2 & 1 & 0{,}5 \\ 1 & 3 & 2 \\ 0{,}5 & 2 & 4 \end{pmatrix}.$$

Найти:

а) совместную плотность распределения случайного вектора $\vec{X}$;

б) одномерные плотности распределения случайных величин X_1, X_2 и X_3.

а.Вычислим определитель матрицы Σ и матрицу $\widetilde{\Sigma} = \Sigma^{-1}$:

$$\begin{vmatrix} 2 & 1 & 0{,}5 \\ 1 & 3 & 2 \\ 0{,}5 & 2 & 4 \end{vmatrix} = 13{,}25,$$

$$\widetilde{\Sigma} = \Sigma^{-1} = \frac{1}{13{,}25}\begin{pmatrix} 8 & -3 & 0{,}5 \\ -3 & 7{,}75 & -3{,}5 \\ 0{,}5 & -3{,}5 & 5 \end{pmatrix},$$

Тогда

$$(\vec{x} - \vec{m})\widetilde{\Sigma}(\vec{x} - \vec{m})^{\mathbb{T}} = (x_1 - 2{,}5 \quad x_2 - 1 \quad x_3 - 2) \times$$
$$\times \frac{1}{13{,}25}\begin{pmatrix} 8 & -3 & 0{,}5 \\ -3 & 7{,}75 & -3{,}5 \\ 0{,}5 & -3{,}5 & 5 \end{pmatrix}\begin{pmatrix} x_1 - 2{,}5 \\ x_2 - 1 \\ x_3 - 2 \end{pmatrix} =$$
$$= \frac{1}{13{,}25}\Big[8(x_1 - 2{,}5)^2 - 6(x_1 - 2{,}5)(x_2 - 1) + 7{,}75(x_2 - 1)^2 +$$
$$+(x_1 - 2{,}5)(x_3 - 2) - 7(x_2 - 1)(x_3 - 2) + 5(x_3 - 2)^2\Big],$$

Совместная плотность распределения случайного вектора $\vec{X}$, имеет вид

$$p_{\vec{X}}(\vec{x}) = \frac{1}{(\sqrt{2\pi})^3\sqrt{13{,}25}} \times$$
$$\times \exp\Big[-\frac{1}{26{,}5}\Big(8(x_1-2{,}5)^2 - 6(x_1-2{,}5)(x_2-1) +$$
$$+ 7{,}75(x_2-1)^2 + (x_1 - 2{,}5)(x_3 - 2) - 7(x_2 - 1)(x_3 - 2) + 5(x_3 - 2)^2\Big)\Big].$$

б.Случайные величины X_1, X_2 и X_3 распределены по нормальному закону с параметрами $m_1 = 2{,}5$ и $\sigma_1 = \sqrt{2}$, $m_2 = 1$ и $\sigma_2 = \sqrt{3}$, $m_3 = 2$ и $\sigma_3 = \sqrt{2}$. Поэтому

$$P_{X_1}(x) = \frac{1}{2\sqrt{\pi}}e^{-(x-2{,}5)^2/4}, \quad P_{X_2}(x) = \frac{1}{\sqrt{6\pi}}e^{-(x-1)^2/6},$$
$$P_{X_3}(x) = \frac{1}{2\sqrt{2\pi}}e^{-(x-2)^2/8}.$$

Многомерное нормальное распределение

Плотность (невырожденного) нормального распределения для *случайного вектора* $\vec{X} = (X_1, \ldots, X_n)$ произвольной размерности $n > 2$.

Многомерный нормальный закон определяется вектором средних значений $\vec{m} = (m_1, \ldots, m_n)$ и положительно определенная симметрическая **матрицей ковариаций (ковариационной матрицей)** $\Sigma = (\sigma_{i\,j}), \quad i, j = \overline{1, n}$.

Диагональный элемент σ_{ii} — **дисперсия** СВ X_i, $\sigma_i = \sqrt{\sigma_{ii}}$ — **СКО** X_i.

Элемент σ_{ij}, $i \neq j$ называют **ковариацией** СВ X_i и X_j ($\sigma_{ij} = \rho_{ij}\sigma_i\sigma_j$, ρ_{ij} — **коэффициент корреляции** X_i и X_j).

Плотность многомерного (n-мерного) нормального распределения (матричная запись)

$$p_{\vec{X}}(\vec{x}) = \frac{1}{(\sqrt{2\pi})^n (\det \Sigma)^{\frac{1}{2}}} e^{-\frac{1}{2}(\vec{x}-\vec{m})\widetilde{\Sigma}(\vec{x}-\vec{m})^{\mathbb{T}}},$$

где $\widetilde{\Sigma} = \Sigma^{-1}$.

Плотность стандартного многомерного $\Sigma = I$, $\vec{m} = (0, \ldots, 0)$ **(** n **-мерного) нормального распределения**.

$$p_{X_1,\ldots,X_n}(x_1, \ldots, x_n) = \frac{1}{(\sqrt{2\pi})^n} e^{-\frac{1}{2}(x_1^2+\ldots+x_n^2)}.$$