

## Лекция 1. Интервальные оценки.

### Основные понятия математической статистики.

Задачи математической статистики – обратные по отношению к теории вероятностей. В теории вероятностей задается некоторая вероятностная модель и нужно посчитать вероятности тех или иных исходов или какие-нибудь характеристики тех или иных вероятностных событий. В математической статистике уже известно, какие события произошли и на основе этих событий нужно построить математическую модель – вероятностное пространство, или какие-нибудь характеристики этого вероятностного пространства, например, моменты распределения случайных величин.

Метода математической статистики делятся на две большие группы – параметрические и непараметрические. В параметрических методах считается, что задан некоторый класс распределения случайных величин (нормальное, пуассоновское, равномерное и т.д.), то есть известен параметрический класс распределений. Однако, параметры этих распределений неизвестны. Например, для нормального распределения неизвестно математическое ожидание или дисперсия, для распределения Пуассона – параметр Пуассона и так далее. На основе наблюдений за случайной величиной, имеющей это заданное распределение, нужно что-то сказать про параметры распределения.

При непараметрических методах статистики неизвестен класс распределений, зависящих от параметра и нам нужны свойства распределений как таковых. Например, неизвестен класс распределений, но на основании значений случайной величины требуется восстановить функцию распределения или плотность распределения.

### Методы параметрической статистики.

В качестве событий будем рассматривать событие, заключающееся в том, что наблюдается значение некоторой случайной величины. Эту случайную величину мы будем обозначать  $\xi$ . Будем считать, что функция распределения этой случайной величины  $\xi$  зависит от некоторого параметра  $\theta$ .  $\theta$  может быть как скалярной, так и векторной величиной, например  $\theta = \{\theta_1, \dots, \theta_n\} \in \mathcal{R}_m$ , где  $\mathcal{R}_m$  –  $m$ -мерное евклидово пространство. Параметр  $\theta$  неизвестен, но задано множество его возможных значений  $\Theta \subset \mathcal{R}_m$ .

Известно, что чем больше число наблюдений за случайной величиной, тем более точно можно сказать о ее значении. Далее мы будем рассматривать ситуации, в которых есть возможность много раз проводить независимые эксперименты по наблюдению за случайной величиной  $\xi$  и на их основании судить о законах распределения случайной величины. В связи с этим вводится следующее понятие.

**Определение.** Выборка объема  $n$  из распределения случайной величины  $\xi$  – это набор значений  $\xi_1, \dots, \xi_n$  случайной величины  $\xi$ , полученный в серии из  $n$  независимых

испытаний по наблюдению  $\xi$ .

Таким образом, выборка рассматривается как набор чисел, которые мы получили в эксперименте. При этом мы можем рассматривать ее как случайный вектор, с  $n$  координатами. Каждая его координата – случайная величина. Все координаты независимы и одинаково распределены.

## Функция правдоподобия.

**Определение.** Плотность распределения вектора  $\xi_1, \dots, \xi_n$  (для абсолютно непрерывной случайной величины) или вероятность получения выборки  $\xi_1, \dots, \xi_n$  (для дискретной случайной величины  $\xi$ ), рассматриваемая как функция параметра  $\theta \in \Theta$ , называется **функцией правдоподобия случайной величины  $\xi$**  и обозначается  $L(\xi_1, \dots, \xi_n, \theta)$ . Таким образом,

$$L(\xi_1, \dots, \xi_n, \theta) = \begin{cases} \prod_{i=1}^n p(\xi_i, \theta), & \text{абс. непр.} \\ \prod_{i=1}^n P(\xi_i, \theta), & \text{дискр.} \end{cases} \quad (1.1)$$

где  $p(\cdot, \cdot)$  – плотность абсолютно непрерывной случайной величины,  $P(\xi_i, \theta)$  – вероятность дискретной случайной величине принять значение  $\xi_i$ , если ее функция распределения определена значением параметра  $\theta \in \Theta$ .

**Замечание.** В теории вероятностей параметр  $\theta$  известен. Тогда по плотности можно узнать вероятность попадания в малую окрестность точки. Однако, в математической статистике параметр  $\theta$  неизвестен, но известны значения. Поэтому, если в плотности распределения вектора  $\xi$  подставим вместо  $x$  те значения, которые были получены в эксперименте, а параметр  $\theta$  оставим неизвестным, то получим формулу для определения плотности с другим смыслом. Плотность будет зависеть от параметра  $\theta$  при фиксированной выборке. То есть формально функция правдоподобия записывается либо как плотность (для абсолютно непрерывной случайной величины), либо как вероятность принять заданное выборочное значение (в случае дискретной случайной величины). И эта функция рассматривается как функция параметра  $\theta$  при фиксированном  $\xi$ .

Почему плотность как функция параметра  $\theta$  называется функцией правдоподобия? Возьмем некоторую нормально распределенную случайную величину. Есть выборка объема 1. Плотность этой случайной величины – гауссова кривая. Если рассматривать ее как функцию  $x$ , то  $p(x)dx$  – вероятность попасть в малый интервал вокруг точки  $x$ . Если рассматривать эту функцию как функцию параметра  $\theta$ , график функции не изменится, но теперь функция будет характеризовать вероятность появления выборки как функции параметра  $\theta$ . Если функция правдоподобия большая в некоторой точке  $\theta$ , то вероятность получения выборки при заданном параметре большая (и наоборот). Это характеризует достаточно большое правдоподобие заданного значения параметра  $\theta$ .

**Пример.** Рассмотрим функцию правдоподобия случайной величины, равномерно распределенной на  $[0, \theta]$ ,  $n = 1$ . Плотность равномерно распределённой случайной величины изображена на рис. 1.1. Если в качестве  $x$  выбрать значение, полученное в эксперименте, то функция правдоподобия – рис. 1.2.

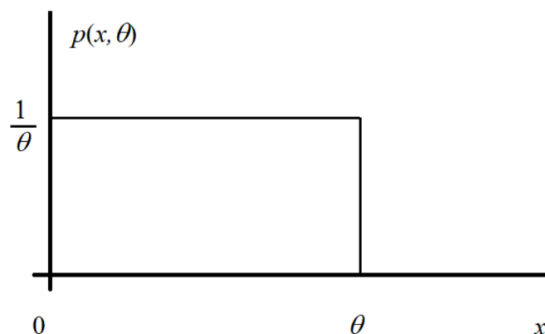


Рис. 1.1: Плотность вероятности.

Плотность вероятности:

$$p(x, \theta) = \begin{cases} \frac{1}{\theta}, & 0 < x < \theta \\ 0, & \text{иначе} \end{cases} \quad (1.2)$$

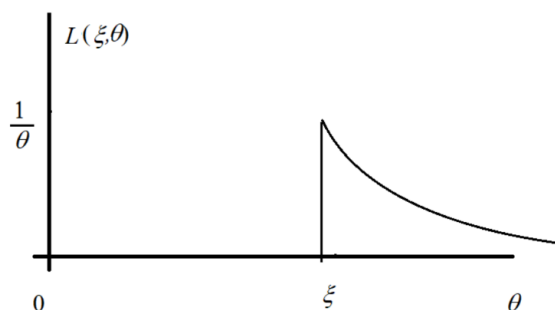


Рис. 1.2: Функция правдоподобия.

Функция правдоподобия:

$$L(\xi, \theta) = \begin{cases} \frac{1}{\theta}, & 0 < \xi < \theta \\ 0, & \text{иначе} \end{cases} \quad (1.3)$$

## Задачи параметрической статистики.

**Определение.** Статистикой называется любая (измеримая) функция выборки, то есть любая случайная величина вида  $t_n(\xi_1, \dots, \xi_n)$ , определенная в  $R^n$  и принимающая числовые

значения.

### Примеры.

Выборочное среднее (среднее арифметическое всех выборочных значений):

$$\bar{\xi} = \frac{1}{n} \sum_{j=1}^n \xi_j \quad (1.4)$$

Выборочный  $k$ -й момент:

$$\overline{\xi^k} = \frac{1}{n} \sum_{j=1}^n (\xi_j)^k \quad (1.5)$$

Аналогично можно определить выборочную дисперсию

$$S = \overline{\xi^2} - (\bar{\xi})^2 = \frac{1}{n} \sum_{j=1}^n \left( \xi_j - \frac{1}{n} \sum_{j=1}^n \xi_k \right)^2 \quad (1.6)$$

Однако далее мы будем пользоваться другой формулой для выборочной дисперсии:

$$S_0 = \frac{1}{n-1} \sum_{j=1}^n \left( \xi_j - \frac{1}{n} \sum_{k=1}^n \xi_k \right)^2 \quad (1.7)$$

### Задачи параметрической статистики:

- задана выборка  $\xi_1, \dots, \xi_n$  объема  $n$  из распределена случайной величины  $\xi$
- задана функция правдоподобия  $L(\xi_1, \dots, \xi_n, \theta)$
- конкретное значение параметра  $\theta$  неизвестно, но задано множество  $\Theta$  его возможных значений

При этих условиях требуется построить решающее правило, которое по выборке  $\xi_1, \dots, \xi_n$  делает вывод о значении  $\theta \in \Theta$ .

Можно выделить несколько важнейших задач о параметре  $\theta$  в зависимости от того, что нужно найти:

- некоторое значение  $\hat{\theta}(\xi_1, \dots, \xi_n)$ , в определенном смысле "близкое" к истинному значению  $\theta$  (**точечное оценивание**)
- интервал со случайными границами  $(t_1(\xi_1, \dots, \xi_n), t_2(\xi_1, \dots, \xi_n))$  который с заданной вероятностью покрывает (содержит) истинное значение параметра  $\theta$  (**интервальное оценивание**)

- решение о том, верно или неверно наше предположение о том или ином значении параметра  $\theta$  (**проверка статистических гипотез**) (например, есть предположение, что математическое ожидание нормального распределения принимает неотрицательные значения. Мы получили некоторую выборку, на основе которой нужно сказать, правдоподобно ли наше предположение.)
- Можно рассматривать и ситуации, в которых требуется сформулировать некоторое формальное решающее правило, следуя которому, можно добиться минимальных потерь (рисков). (**оптимальные статистические решения**)

## Методы непараметрической статистики.

**Определение.** Если все  $n$  элементов выборки  $\xi_1, \dots, \xi_n$  расположены в порядке неубывания их величины, и члены такой неубывающей последовательности обозначены  $\xi_{(k)}$ ,  $\xi_{(1)} \leq \xi_{(2)} \leq \dots \leq \xi_{(n)}$ , то каждое из  $\xi_{(k)}$  называется **порядковой статистикой**, а последовательность  $\xi_{(1)}, \dots, \xi_{(n)}$  называется **вариационным рядом**, построенным по выборке  $\xi_1, \dots, \xi_n$ .

Значение  $\xi_{(k)}$ , стоящее на  $k$ -м месте, называется  $k$ -й **порядковой статистикой**,  $\xi_{(1)}$  - минимальным членом вариационного ряда,  $\xi_{(n)}$  - максимальным членом вариационного ряда  $\xi_{(1)}, \dots, \xi_{(n)}$ .

## Интервальные оценки параметра распределения.

**Определение.** Назовем **интервальной оценкой параметра**  $\theta$  интервал  $(t_1(\xi_1, \dots, \xi_n), t_2(\xi_1, \dots, \xi_n))$  такой, что вероятность того, что истинное значение параметра  $\theta$  попадет в этот интервал, задана, равна  $\gamma$  :

$$\gamma = P_{\theta} (t_1(\xi_1, \dots, \xi_n) \leq \theta \leq t_2(\xi_1, \dots, \xi_n)) \quad (1.8)$$

не зависит от  $\theta \in \Theta$ .

**Интерпретация:** случайный интервал с вероятностью  $\gamma$  содержит истинное значение параметра  $\theta$ .

Оценка тем лучше, чем меньше длина интервала (точность оценки) и чем выше вероятность  $\gamma$  (надежность, уровень доверия).

**Пример.** Интервальная оценка математического ожидания нормально распределенной случайной величины при известной дисперсии.

Пусть  $\xi_1, \dots, \xi_n$  — выборка из нормального распределения  $\mathcal{N}(\mu, \sigma^2)$ . Построим случайный интервал, накрывающий значение математического ожидания  $\mu$  (то есть в этом примере  $\theta = \mu$ ) с заданной вероятностью. Для этого построим случайную величину,

зависящую от выборочных значений и от неизвестного параметра, но имеющее известное распределение:

$$\frac{1}{n} \sum_{j=1}^n (\xi_j - \mu) \sim \mathcal{N}\left(0, \frac{\sigma^2}{n}\right) \quad (1.9)$$

Сделаем так, чтобы указанная величина имела единичную дисперсию (для красоты)

$$T(\xi_1, \dots, \xi_n, \mu) = \frac{1}{n} \sum_{j=1}^n (\xi_j - \mu) \frac{\sqrt{n}}{\sigma} \sim \mathcal{N}(0, 1) \quad (1.10)$$

Теперь мы можем указать интервал, в котором эта случайная величина содержится с заданной вероятностью. Плотность распределения указана на рис. 1.3. Нас интересует интервал, в который случайная величина попадает с вероятностью  $1 - \alpha$ . Следовательно,

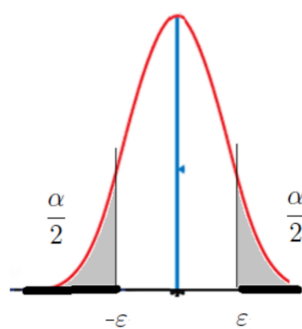


Рис. 1.3: Плотность распределения.

$$P(|T(\xi_1, \dots, \xi_n, \mu)| < \varepsilon) = \frac{1}{\sqrt{2\pi}} \int_{-\varepsilon}^{\varepsilon} e^{-\frac{u^2}{2}} du = \Phi(\varepsilon) - \Phi(-\varepsilon) = 1 - \alpha \quad (1.11)$$

Подставим явное выражение для функции  $T$

$$P\left(-\varepsilon < \frac{\sqrt{n}}{\sigma} \left(\frac{1}{n} \sum_{j=1}^n (\xi_j - \mu)\right) < \varepsilon\right) = 1 - \alpha \quad (1.12)$$

Разрешим это неравенство относительно  $\mu$

$$P\left(\frac{1}{n} \sum_{j=1}^n \xi_j - \varepsilon \frac{\sigma}{\sqrt{n}} < \mu < \frac{1}{n} \sum_{j=1}^n \xi_j + \varepsilon \frac{\sigma}{\sqrt{n}}\right) = 1 - \alpha \quad (1.13)$$

Таким образом, мы получили интервальную оценку параметра  $\mu$  при известной дисперсии.

## Общий метод построения интервальной оценки.

Обобщим этот пример, описав один из способов построения интервальной оценки. Этот метод состоит в поиске функции  $T(\xi_1, \dots, \xi_n, \theta)$ , такой, что

- распределение случайной величины  $T(\xi_1, \dots, \xi_n, \theta)$  не зависит от  $\theta \in \Theta$
- случайная величина  $T(\xi_1, \dots, \xi_n, \theta)$  обладает известной функцией распределения  $F_T(z)$
- неравенство

$$a_1 \leq T(\xi_1, \dots, \xi_n, \theta) \leq a_2 \quad (1.14)$$

выполненное с вероятностью  $\gamma = F_T(a_2) - F_T(a_1)$ , можно решить относительно  $\theta$  и переписать в виде

$$t_1(\xi_1, \dots, \xi_n) \leq \theta \leq t_2(\xi_1, \dots, \xi_n) \quad (1.15)$$

Здесь мы считаем, что случайная величина  $T$  имеет абсолютно непрерывное распределение, поэтому можно писать нестрогие неравенства.

## Некоторые утверждения из линейной алгебры.

Чтобы строить интервальные оценки, нужно иметь хороший набор функций с известными распределениями. Чтобы строить такие функции, нам нужны некоторые знания линейной алгебры.

**Определение.** Линейный оператор

$$U : \mathcal{R} \rightarrow \mathcal{R} \quad (1.16)$$

называется **ортогональным** (унитарным), если это преобразование сохраняет скалярное произведение

$$(Ux, Uy) = (x, y) \quad \forall x, y \in \mathcal{R} \quad (1.17)$$

В любом ортонормированном базисе  $\{e_i\} \subset \mathcal{R}$  матрица  $\tilde{U}$  оператора  $U$ ,  $\tilde{U} = \{(e_i, Ue_j)\}$ , – ортогональная:

$$\tilde{U}^{tr} \equiv \tilde{U}^* = \tilde{U}^{-1}, \quad \det \tilde{U} = \pm 1 \quad (1.18)$$

**Теорема об ортогональном разложении евклидова пространства.** Пусть  $\mathcal{L}$  — линейное подпространство  $\mathcal{R}$ ,  $\mathcal{L}^\perp$  — ортогональное дополнение. Тогда для любого

$$x \in \mathcal{R} \quad x = x_1 + x_2, \quad x_1 \in \mathcal{L}, \quad x_2 \in \mathcal{L}^\perp, \quad \mathcal{R} = \mathcal{L} \oplus \mathcal{L}^\perp \quad (1.19)$$

Это разложение единственно, так как

$$x = x_1 + x_2 = x'_1 + x'_2 \quad (1.20)$$

следовательно

$$0 = (x_1 - x'_1) + (x_2 - x'_2) \quad (1.21)$$

где первое слагаемое принадлежит  $\mathcal{L}$ , а второе —  $\mathcal{L}^\perp$ . В силу ортогональности этих слагаемых

$$\|x_1 - x'_1\|^2 + \|x_2 - x'_2\|^2 = 0 \quad (1.22)$$

откуда

$$x_1 = x'_1, \quad x_2 = x'_2 \quad (1.23)$$

Данное утверждение можно проиллюстрировать (рис. 1.4). Пусть есть некоторое трехмерное пространство,  $x$  – вектор,  $\mathcal{L}$  – двумерное подпространство. Тогда любой вектор  $x$  можно представить в виде ортогонального разложения на два ортогональных вектора.

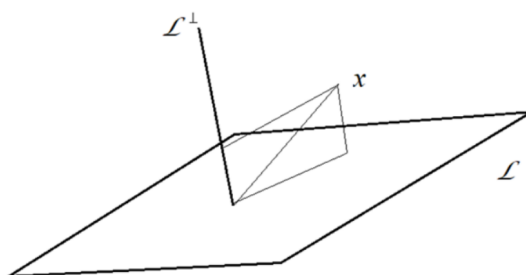


Рис. 1.4: Иллюстрация к объяснению.

**Определение.** Оператор

$$P : \mathcal{R} \rightarrow \mathcal{R} \quad (1.24)$$

называется **ортогональным проектором** на  $\mathcal{L}$ , если

$$Px = x_1 \quad (1.25)$$

**Свойства ортогональных проекторов:**

1.  $P$  – линейный оператор. Действительно, пусть

$$x = x_1 + x_2, \quad y = y_1 + y_2 \quad (1.26)$$

Тогда линейная комбинация

$$ax + by = (ax_1 + by_1) + (ax_2 + by_2) \quad (1.27)$$

Вычислим значение ортогонального проектора линейной комбинации

$$\Pi(\alpha x + \beta y) = \alpha x_1 + \beta y_1 = \alpha \Pi x + \beta \Pi y \quad (1.28)$$

2.  $\Pi$  — самосопряженный оператор

$$\Pi = \Pi^* \quad (1.29)$$

Действительно

$$(\Pi x, y) = (x, y) = (x_1, y_1) = (x, y_1) = (x, \Pi y) \quad (1.30)$$

3. Квадрат проектора равен самому проектору

$$\Pi^2 = \Pi \quad (1.31)$$

Действительно

$$\Pi x = x_1 = \Pi x_1 = \Pi^2 x, \quad x \in \mathcal{R} \quad (1.32)$$

**Замечание.** Оператор **сопряженный**, если:

$$(Ax, y) = (x, A^*y) \quad (1.33)$$

**Самосопряженный** оператор:

$$(Ax, y) = (x, Ay) \quad (1.34)$$

то есть

$$A = A^* \quad (1.35)$$

**Лемма.** Условия 1,2,3 необходимы и достаточны, чтобы  $\Pi$  был ортогональным проектором. Он ортогонально проецирует на

$$\mathcal{L} = \{x \in \mathcal{R}, \Pi x = x\} \quad (1.36)$$

А ортогональным дополнением к этому линейному подпространству является

$$\mathcal{L}^\perp = \{x \in \mathcal{R}, \Pi x = 0\} \quad (1.37)$$

Доказательство. Обозначим

$$\mathcal{L} = \{x \in \mathcal{R}\}, \quad \mathcal{N} = \{x \in \mathcal{R}, \Pi x = 0\} \quad (1.38)$$

Тогда

1.  $\mathcal{L} \perp \mathcal{N}$ , так как  $x \in \mathcal{L}$ ,  $y \in \mathcal{N}$ , следовательно,

$$(x, y) = (\Pi x, y) = (x, \Pi y) = 0 \quad (1.39)$$

2. Для любого  $x \in \mathcal{R}$

$$x = \Pi x + (I - \Pi)x \quad (1.40)$$

и значит,  $\Pi x \in \mathcal{L}$ , так как

$$\Pi(\Pi x) = \Pi^2 x = \Pi x \quad (1.41)$$

а второе слагаемое  $(I - \Pi)x \in \mathcal{N}$ , так как

$$\Pi(I - \Pi)x = \Pi x - \Pi^2 x = 0 \quad (1.42)$$

3. Значит (см. предыдущий пункт),  $\mathcal{R}$  представляется в виде двух ортогональных подпространств, причем любой  $x$  представляется в виде суммы (1.40).

$$\mathcal{R} = \mathcal{L} \oplus \mathcal{N} \quad (1.43)$$

то есть  $\mathcal{N} = \mathcal{L}^\perp$ .  $\Pi$  ортогонально проецирует на  $\mathcal{L}$ ,  $(I - \Pi)$  – на  $\mathcal{L}^\perp$ .

Лемма доказана.

Рассмотрим способы построения ортогональных проекторов. Известно, что проекция вектора  $x$  на плоскость есть основание перпендикуляра, опущенного из  $x$  на  $\mathcal{L}$  (см. рис. 1.5). Эта точка основания перпендикуляра – ближайшая к  $x$ . Фактически, чтобы найти проекцию  $x$  на  $\mathcal{L}$ , нужно найти точку из  $\mathcal{L}$ , ближайшую к  $x$ .

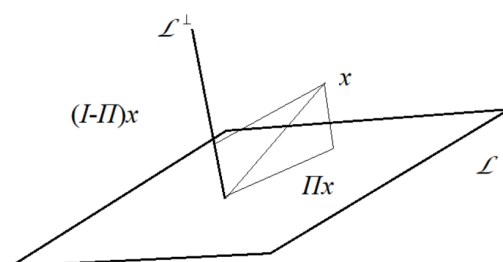


Рис. 1.5: Иллюстрация к объяснению.

**Теорема 1.** Пусть

$$\rho(x, \mathcal{L}) = \inf \{ \|x - y\|, y \in \mathcal{L} \} \quad (1.44)$$

Тогда ближайшая точка из  $\mathcal{L}$  есть проекция на  $\mathcal{L}$

$$\rho(x, \mathcal{L}) = \|x - \Pi x\| \quad (1.45)$$

где  $\Pi$  ортогонально проецирует на  $\mathcal{L}$ .

Доказательство. Точная нижняя грань нормы достигается там же, где и точная нижняя грань квадрата нормы. Для  $y \in \mathcal{L}$  :

$$\|x - y\|^2 = \|(x - \Pi x) + (\Pi x - y)\|^2 = \|x - \Pi x\|^2 + \|\Pi x - y\|^2 \geq \|x - \Pi x\|^2 \quad (1.46)$$

**Пример.** Проектор на множество значений  $\mathcal{R}(A)$  линейного оператора,  $A : \mathcal{R}_N \rightarrow \mathcal{R}_n$ . Пространство значений линейного оператора – это множество тех элементов, которые имеют вид  $Ay$ . Чтобы найти проекцию, нужно решить задачу

$$\inf_{y \in \mathcal{R}_N} \|x - Ay\| = \|x - \Pi_{\mathcal{R}(A)} x\| \quad (1.47)$$

Действительно,

$$(x - Ay, x - Ay) = (x, x) - 2(x, Ay) + (Ay, Ay) \quad (1.48)$$

приравняем нулю производную по  $y$

$$-2(A^*x, \delta y) + 2(A^*Ay, \delta y) \quad (1.49)$$

что дает при невырожденном  $A$

$$(A^*A)y^* = A^*x \quad (1.50)$$

$$(A^*A)y^* = (A^*A)^{-1} A^*x \quad (1.51)$$

то есть

$$\Pi_{\mathcal{R}(A)} = A(A^*A)^{-1}A^* \quad (1.52)$$

**Замечание к примеру.** Как вычислить производную?

$$\Phi(y + \partial y) = \Phi(y) + (\Phi'(y), \partial y) + o(\|\partial y\|) \quad (1.53)$$

В нашем случае

$$\Phi(y) = \|x - Ay\|^2 \quad (1.54)$$

Значит, нужно проверить, можно ли представить функционал (1.54) в виде (1.53). Запишем

приращение

$$\begin{aligned}\|x - A(y + \partial y)\|^2 &= (x - Ay, x - Ay) + 2(x - A(y), A\partial y) + (A\partial y, A\partial y) = \\ &= \Phi(y) + (2A^*(x - Ay), \partial y) + o(\|\partial y\|)\end{aligned}\quad (1.55)$$

Отсюда

$$\Phi'(y) = 2A^*(x - Ay) \quad (1.56)$$

Тогда, приравняв производную к нулю, получаем условие минимума

$$\Phi'(y) = 0 \implies y = (A^*A)^{-1} A^*x \quad (1.57)$$

Тогда

$$Ay = A(A^*A)^{-1} A^*x \quad (1.58)$$

## Невырожденное многомерное нормальное распределение.

Рассмотрим невырожденное многомерное распределение. Невырожденность означает, что матрица ковариации этого распределения обратима. Определим многомерное нормальное распределение как распределение с плотностью

$$\begin{aligned}p_\xi(x) &= \frac{(\det \Sigma^{-1})^{\frac{1}{2}}}{(2\pi)^{\frac{n}{2}}} \exp\left(-\frac{1}{2}(\Sigma^{-1}(x - \mu), x - \mu)\right) = \\ &= \frac{1}{[(2\pi)^n \det \Sigma]^{\frac{1}{2}}} \exp\left(-\frac{1}{2}(\Sigma^{-1}(x - \mu), x - \mu)\right)\end{aligned}\quad (1.59)$$

Для нормального распределения

$$M\xi = \mu, \quad cov(\xi_i, \xi_j) = \Sigma_{ij} : M(\xi - \mu)(\xi - \mu)^* = \Sigma \quad (1.60)$$

– матрица ковариаций.

Здесь

$$M\xi = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} x p_\xi(x) dx \quad (1.61)$$

$$\Sigma_{ij} = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} (x - \mu)(x - \mu)^* p_\xi(x) dx \quad (1.62)$$

## Свойство нормального распределения.

**Теорема 2.** Пусть  $\xi$  — случайный вектор  $\mathcal{R}_n$ , координаты которого  $\xi_1, \dots, \xi_n$  в некотором ортонормированном базисе  $\{e_i\}$  независимы в совокупности и нормальны  $\mathcal{N}(0, \sigma^2)$ . Если  $U$  — ортогональный оператор  $\mathcal{R}_n \rightarrow \mathcal{R}_n$ , то координаты  $(U\xi)_1, \dots, (U\xi)_n$  вектора  $U\xi$  распределены так же, как координаты, вектора  $\xi$ . То есть, распределение вектора  $\xi$  не меняется при преобразовании поворота и отражения.

Доказательство. По условию теоремы

$$p_\xi(x) = \prod_{i=1}^n \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{x_i^2}{2\sigma^2}} = \left( \frac{1}{\sqrt{2\pi\sigma}} \right)^n \exp \frac{\|x\|^2}{2\sigma^2} \quad (1.63)$$

Из теории вероятностей известен следующий факт. Пусть  $\xi$  – вектор с известным распределением,  $A$  – некоторый обратимый линейный оператор, а вектор  $\eta$  получается линейным преобразованием:

$$\eta = A\xi, \quad \eta \in \mathcal{R}_n, \quad \xi \in \mathcal{R}_n \quad (1.64)$$

Дана плотность  $p_\xi(x)$ ,  $x \in \mathcal{X}$ . Чтобы определить  $p_\eta(\cdot)$ , запишем равенства

$$P(\eta \in Q) = \int_{y \in Q} p_\eta(y) dy = P(A\xi \in Q) = \int_{Ax \in Q} p_\xi(x) dx = \int_{y \in Q} p_\xi(A^{-1}y) |det A^{-1}| dy \quad (1.65)$$

Таким образом, получили равенство двух интегралов по одинаковым областям, под знаками которых стоят различные выражения с плотностями. Интегралу равны, если равны их подынтегральные значения. Поэтому

$$p_{U\xi}(x) = p_\xi(U^{-1}x) |det \tilde{U}^{-1}| = p_\xi(x) \quad (1.66)$$

то есть

$$\xi \sim \mathcal{N}(0, \sigma^2 I) \quad (1.67)$$

**Следствие.** Пусть  $\xi$  – случайный вектор, координаты которого в некотором ортонормированном базисе удовлетворяют условиям теоремы 2. Тогда координаты вектора  $\xi$  в любом базисе удовлетворяют этим условиям.

В данном случае мы преобразуем матрицей  $U$  не вектор, а базис.

Доказательство. Действительно, любой ортонормированный базис можно записать в виде

$$\tilde{e}_i = Ue_i, \quad i = 1, \dots, n \quad (1.68)$$

где  $\{e_i\}$  – некоторый ортонормированный базис  $\mathcal{R}_n$ . Поэтому

$$\tilde{\xi}_i = (\xi, \tilde{e}_i) = (\xi, Ue_i) = (U^{-1}\xi, e_i) \quad (1.69)$$

$U^{-1}$  – ортогональный оператор

$$U^* = U^{-1} \quad (1.70)$$

то есть  $\tilde{\xi}_1, \dots, \tilde{\xi}_n$  – координаты  $\tilde{U}\xi$  в ортонормированном базисе  $\{e_i\}$ .

$$(U^{-1}x, U^{-1}y) = (x, U^{-1} * U^{-1}y) = (x, (U^{-1})^{-1} U^{-1}y) = (x, y) \quad (1.71)$$

**Теорема 2\*.** Пусть  $\xi$  — случайный вектор  $\mathcal{R}_n$ , координаты которого  $\xi_1, \dots, \xi_n$  в некотором ортонормированном базисе  $\{e_i\}$  удовлетворяют условиям:

$$M\xi_j = 0, \quad M\xi_i\xi_j = \sigma^2\delta_{ij}, \quad i, j = 1, \dots, n \quad (1.72)$$

Тогда таким условиям удовлетворяют координаты вектора  $\xi$  в любом ортонормированном базисе.

Доказательство. В силу линейности математического ожидания

$$M\xi = 0 \quad (1.73)$$

получаем

$$MU\xi = 0 \quad (1.74)$$

Для доказательства второго соотношения заметим, что матрицу ковариаций можно записать в виде

$$M\xi\xi^* = \sigma^2 I \quad (1.75)$$

Перепишем

$$MU\xi\xi^*U^* = \sigma^2 UU^* = \sigma^2 I \quad (1.76)$$

Таким образом, поворот случайного вектора с нулевым средним, координаты которого некоррелированы и имеют одинаковые дисперсии, не меняет распределение в случае нормальности этого распределения и не меняет моменты, если про распределение ничего не известно.

## Распределение хи-квадрат с $n$ степенями свободы.

**Определение.** Пусть  $\xi = (\xi_1, \dots, \xi_n) \sim \mathcal{N}(0, I)$ . Статистика

$$\chi_k^2 = \sum_{j=1}^k \xi_j^2 \quad (1.77)$$

называется **статистикой хи-квадрат с  $k$  степенями свободы.**

Заметим, что  $\xi \in \mathcal{R}_n$ ,  $\xi_1, \dots, \xi_n \sim \mathcal{N}(0, 1)$ ,  $M\chi_k^2 = k$ ,  $D\chi_k^2 = 2k$ .

Графики плотности этой статистики в зависимости от степеней свободы указаны на рис. 1.6. Если степень свободы равна  $k = 1$ , то мы получаем квадрат нормального распределения. В случае  $k = 2$  плотность начинается в точке 0.5, после чего убывает. Для всех остальных случаев максимум плотности попадает в точку  $k - 2$ .

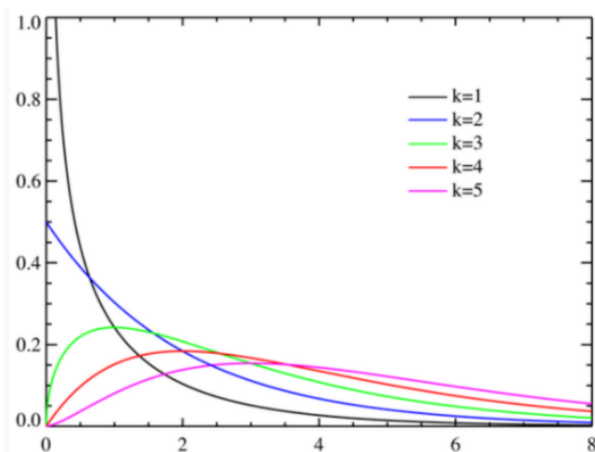


Рис. 1.6: Плотности статистики хи-квадрат.

**Лемма 2.** Пусть  $\mathcal{R}_k$  —  $k$ -мерное линейное подпространство  $\mathcal{R}_n$ . Если  $\xi \sim \mathcal{N}(0, I)$ , то квадрат проекции этого вектора имеет распределение хи-квадрат с  $k$ -степенями свободы

$$\|\Pi_k \xi\|^2 = \chi_k^2 \quad (1.78)$$

**Доказательство.** Выберем ортонормированный базис  $e_1, \dots, e_k, \dots, e_n$  пространства  $\mathcal{R}_n$  так, чтобы  $e_1, \dots, e_k$  образовывал ортонормированный базис  $\mathcal{R}_k$ . Тогда, согласно теореме 2, все случайные величины  $(\xi, e_1), \dots, (\xi, e_k)$  распределены согласно  $\mathcal{N}(0, 1)$  и независимы в совокупности, и, следовательно,

$$\|\Pi_k \xi\|^2 = \sum_{i=1}^k \xi_i^2 = \chi_k^2 \quad (1.79)$$

**Замечание.** Если  $\xi_j \sim \mathcal{N}(0, \sigma^2)$ ,  $j = 1, \dots, n$ , то

$$\|\Pi_k \xi\|^2 \sim \sigma^2 \chi_k^2 \quad (1.80)$$

## Распределение Снедекора-Фишера с $k$ и $m$ степенями свободы.

**Определение.** Пусть  $\xi_1, \dots, \xi_k, \xi_{k+1}, \dots, \xi_{k+m}$  распределены нормально  $\mathcal{N}(0, 1)$  и независимы в совокупности. Тогда случайная величина

$$\varphi_{k,m} = \frac{\frac{1}{k} \sum_{j=1}^k \xi_j^2}{\frac{1}{m} \sum_{j=k+1}^{k+m} \xi_j^2} \quad (1.81)$$

называется **статистикой Снедекора-Фишера с  $k$  и  $m$  степенями свободы.**

Иначе:

$$\varphi_{k,m} = \frac{\frac{\chi_k^2}{k}}{\frac{\chi_m^2}{m}} \quad (1.82)$$

где  $\chi_k^2$  и  $\chi_m^2$  независимы.

Математическое ожидание и дисперсия распределения Снедекора-Фишера с  $k$  и  $m$  степенями свободы равны, соответственно,

$$M\varphi_{k,m} = \frac{m}{m-2}, \quad m > 2 \quad (1.83)$$

$$D\varphi_{k,m} = \frac{2m^2 \left(1 + \frac{m-2}{k}\right)}{(m-2)^2(m-4)^2}, \quad m > 4 \quad (1.84)$$

**Лемма 3.** Пусть  $\mathcal{R}_k$  и  $\mathcal{R}_m$  — линейные подпространства  $\mathcal{R}_n$ , причем  $\mathcal{R}_k \perp \mathcal{R}_m$ . Пусть  $\Pi_k$  и  $\Pi_m$  — ортогональные проекторы на  $\mathcal{R}_k$  и  $\mathcal{R}_m$ . Тогда при условиях Леммы 2

$$\varphi_{k,m} = \frac{\|\Pi_k \xi\|^2/k}{\|\Pi_m \xi\|^2/m} \quad (1.85)$$

**Доказательство.** Выберем ортонормированный базис  $e_1, \dots, e_k, \dots, e_n$  пространства  $\mathcal{R}_n$  так, чтобы  $e_1, \dots, e_k$  образовывал ортонормированный базис  $\mathcal{R}_k$ , а  $e_{k+1}, \dots, e_n$  — ортонормированный базис  $\mathcal{R}_m$ . Тогда, согласно теореме 2, все случайные величины  $e_1, \dots, e_k$  распределены согласно  $\mathcal{N}(0, 1)$  и независимы в совокупности, и, следовательно,

$$\frac{\|\Pi_k \xi\|^2/k}{\|\Pi_m \xi\|^2/m} \sim \varphi_{k,m} \quad (1.86)$$

## Распределение Стьюдента с $n$ степенями свободы.

**Определение.** Пусть случайная величина  $\xi$  распределена нормально  $\mathcal{N}(0, 1)$ , а случайная величина  $\chi_k^2$  — согласно распределению хи-квадрат с  $k$  степенями свободы и они независимы. Тогда случайная величина

$$\tau_k = \frac{\xi}{\sqrt{\frac{\chi_k^2}{k}}} \quad (1.87)$$

называется **статистикой Стьюдента с  $k$  степенями свободы.**

Плотности статистики Стьюдента изображены на рис. 1.7.

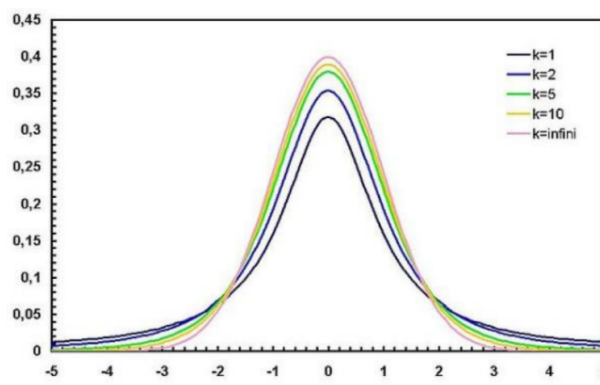


Рис. 1.7: Плотности статистики Стьюдента.

Математическое ожидание и дисперсия распределения Стьюдента с  $k$  степенями свободы равны, соответственно,

$$M\tau_k = 0, \quad k \geq 2 \quad (1.88)$$

$$D\tau_k = \frac{k}{k-2}, \quad k > 2 \quad (1.89)$$

**Лемма 4.** Пусть  $\xi \sim \mathcal{N}(0, I)$   $\Pi\xi = (\xi, e)e$ ,  $e \in \mathcal{R}_n$  — фиксированный вектор единичной нормы:  $\|e\| = 1$ . Тогда

$$\tau_{n-1} = \frac{(\xi, e)}{\left[ \|(I - \Pi)\xi\|^2 / (n-1) \right]^{\frac{1}{2}}} \quad (1.90)$$

– статистика Стьюдента, с  $n - 1$  степенью свободы.

**Доказательство.** Выберем ортонормированный базис  $e, e_1, \dots, e_k, \dots, e_{n-1}$  пространства  $\mathcal{R}_n$ . Тогда

$$\|(I - \Pi)\xi\|^2 = \sum_{j=1}^{n-1} \xi_j^2 = \chi_{n-1}^2 \quad (1.91)$$

и не зависит от  $(\xi, e) \sim \mathcal{N}(0, 1)$ .

**Следствие.** Пусть  $e_1, \dots, e_n$  — ортонормированный базис  $\mathcal{R}_n$ . Возьмем вектор  $e$  в виде диагонали первого координатного угла, образованного этим базисом

$$e = \frac{1}{\sqrt{n}} \sum_{j=1}^n e_j \in \mathcal{R}_n \quad (1.92)$$

Тогда скалярное произведение

$$(\xi, e) = \frac{1}{\sqrt{n}} \sum_{j=1}^n \xi_j \quad (1.93)$$

$k$ -я координата вектора проекции

$$(\Pi\xi)_k = \frac{1}{n} \sum_{j=1}^n \xi_j \quad (1.94)$$

Тогда разность

$$\|\xi - \Pi\xi\|^2 = \sum_{i=1}^n \left( \xi_i - \frac{1}{n} \sum_{j=1}^n \xi_j \right)^2 \quad (1.95)$$

то есть

$$\tau_{n-1} = \frac{\frac{1}{\sqrt{n}} \sum_{j=1}^n \xi_j}{\left[ \frac{\sum_{j=1}^n \left( \xi_j - \frac{1}{n} \sum_{j=1}^n \xi_j \right)^2}{n-1} \right]^{\frac{1}{2}}} \quad (1.96)$$

Полученная конструкция будет иметь распределение Стьюдента.

**Замечание.** Если математическое ожидание отлично от нуля, то есть  $\xi_j \sim \mathcal{N}(\mu, \sigma^2)$  и независимы, то

$$\tau_{n-1} = \frac{\frac{1}{\sqrt{n}} \sum_{j=1}^n (\xi_j - \mu)}{\left[ \frac{\sum_{j=1}^n \left( \xi_j - \frac{1}{n} \sum_{j=1}^n \xi_j \right)^2}{n-1} \right]^{\frac{1}{2}}} = \frac{\frac{1}{n} \sum_{j=1}^n (\xi_j - \mu)}{\sqrt{\frac{1}{n(n-1)} \sum_{j=1}^n \left( \xi_j - \frac{1}{n} \sum_{j=1}^n \xi_j \right)^2}} \quad (1.97)$$

Это замечание относится и к леммам 2 и 3.

## Интервальные оценки параметров нормального распределения.

Теперь мы можем строить интервальные оценки параметров распределения и в других случаях. Пусть  $\xi_1, \dots, \xi_n$  — выборка из нормального распределения  $\mathcal{N}(\mu, \sigma^2)$ . Построим случайные интервалы, накрывающие значения параметров распределения с заданной вероятностью (уровнем доверия).

1. Дисперсия  $\sigma^2$  неизвестна, а математическое ожидание  $\mu$  — известно. Тогда

$$\frac{1}{\sigma^2} \sum_{j=1}^n (\xi_j - \mu)^2 \sim \chi_n^2 \quad (1.98)$$

Таким образом, мы сконструировали функцию, которая при любом параметре  $\mu$  имеет одно и то же распределение. тогда, зная плотность распределения, мы можем найти интервал, который содержит эту случайную величину с вероятностью  $1 - \alpha$ :

$$P\left(\chi_n^2 < \varepsilon_1\right) = P\left(\chi_n^2 > \varepsilon_2\right) = \frac{\alpha}{2} \quad (1.99)$$

Отсюда

$$P\left(\varepsilon_1 \leq \chi_n^2 \leq \varepsilon_2\right) = 1 - \alpha \quad (1.100)$$

Если мы хотим, чтобы правее точки  $\varepsilon_2$  случайная величина попадала с вероятностью  $\frac{\alpha}{2}$ , то, задав  $\frac{\alpha}{2}$ , мы можем найти  $\varepsilon_2$  по таблице распределения хи-квадрат (см. рис. 1.8). Для  $\varepsilon_1$  аналогично.

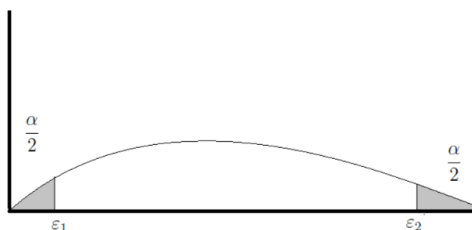


Рис. 1.8: Иллюстрация к пояснению.

Отсюда

$$P\left(\varepsilon_1 \leq \frac{1}{\sigma^2} \sum_{j=1}^n (\xi_j - \mu)^2 \leq \varepsilon_2\right) = 1 - \alpha \quad (1.101)$$

Разрешим неравенство относительно  $\sigma^2$

$$P\left(\frac{\sum_{j=1}^n (\xi_j - \mu)^2}{\varepsilon_2} \leq \sigma^2 \leq \frac{\sum_{j=1}^n (\xi_j - \mu)^2}{\varepsilon_1}\right) = 1 - \alpha \quad (1.102)$$

2. Дисперсия  $\sigma^2$  неизвестна, требуется оценить математическое ожидание  $\mu$ .

В этом случае мы можем воспользоваться известным результатом (1.97), который мы получили для распределения Стьюдента. Тогда получим

$$\frac{\frac{1}{\sqrt{n}} \sum_{j=1}^n (\xi_j - \mu)}{\left[ \frac{\|(I-\Pi)\xi\|^2}{n-1} \right]^{\frac{1}{2}}} = \tau_{n-1} \quad (1.103)$$

В силу симметричности данного распределения, задавая  $\alpha$ , мы можем вычислить  $\varepsilon$  (рис. 1.9).

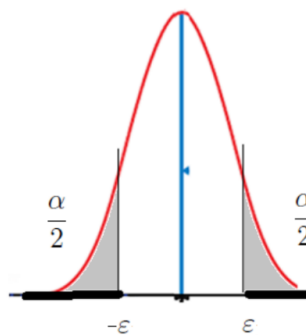


Рис. 1.9: Иллюстрация к объяснению.

Тогда

$$P(|\tau_{n-1}| < \varepsilon) = 1 - \alpha \quad (1.104)$$

Подставим вместо  $\tau$  дробь

$$P\left(-\varepsilon \leq \frac{\frac{1}{\sqrt{n}} \sum_{j=1}^n (\xi_j - \mu)}{\left[ \frac{\|(I-\Pi)\xi\|^2}{n-1} \right]^{\frac{1}{2}}} < \varepsilon\right) = 1 - \alpha \quad (1.105)$$

Решим неравенство относительно  $\mu$

$$P\left(\frac{1}{n} \sum_{j=1}^n \xi_j - \varepsilon \frac{\|(I-\Pi)\xi\|}{\sqrt{n(n-1)}} < \mu < \frac{1}{n} \sum_{j=1}^n \xi_j + \varepsilon \frac{\|(I-\Pi)\xi\|}{\sqrt{n(n-1)}}\right) = 1 - \alpha \quad (1.106)$$

3. Требуется оценить дисперсию  $\sigma^2$ , если математическое ожидание  $\mu$  — неизвестно. В этом случае в качестве случайной величины, зависящей от неизвестного параметра  $\sigma^2$  будем использовать дробь

$$\frac{\|(I-\Pi)\xi\|^2}{\sigma^2} = \chi_{n-1}^2 \quad (1.107)$$

Ее распределение всегда одинаково. Аналогично предыдущему случаю выбираем  $\varepsilon_1$  и  $\varepsilon_2$

по значению  $\alpha$  (рис. 1.10).

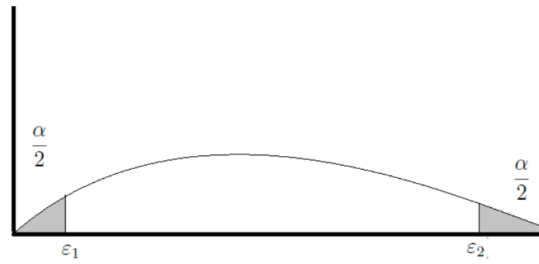


Рис. 1.10: Иллюстрация к объяснению.

Полученные  $\epsilon$  подставим в неравенство

$$P\left(\chi_{n-1}^2 < \epsilon_1\right) = P\left(\chi_{n-1}^2 > \epsilon_2\right) = \frac{\alpha}{2} \quad (1.108)$$

Тогда получим

$$P\left(\epsilon_1 \leq \chi_{n-1}^2 \leq \epsilon_2\right) = 1 - \alpha \quad (1.109)$$

Разрешаем дробь относительно  $\sigma^2$

$$P\left(\frac{\|(I - \Pi)\xi\|^2}{\epsilon_2} \leq \sigma^2 \leq \frac{\|(I - \Pi)\xi\|^2}{\epsilon_1}\right) = 1 - \alpha \quad (1.110)$$