Конспект по курсу

Математическая статистика

Contributors:
Андрей Степанов
Алексей Журавлев

Лектор:
Шабанов

МФТИ

Последнее обновление: 25 февраля 2015 г.

Содержание

1 Введение. Сходимости векторов.

1.1 Введение.

Математическая статистика – это раздел теории вероятностей, который решает обратные задачи к классическим задачам в теории вероятностей.

Типичная задача в теории вероятностей – это найти или оценить характеристики случайного эксперимента, зная его природу случайности.

Типичная задача в математической статистике – по данным результатов случайного эксперимента выяснить природу его случайности.

Пример (Классический пример)**.** В городе есть n жителей, m из которых болеют. Считаем, что n дано заранее.

- Задача ТВ: с какой вероятностью при известном m в случайной выборке из a жителей будет b заболевших

- Задача МС: известно, что в выборке из a жителей оказалось b заболевших. Как в этом случае можно оценить m

Пример (Выборка)**.** Предположим, что мы проводим эксперимент. Пусть дан какой-то физический прибор, и пусть ξ – случайная величина, описывающая результат измерения этим прибором, $\xi \sim P_\xi$ (ξ имеет распределение P_ξ). Например, если прибор – это счетчик Гейгера, то ξ – это уровень радиации, им зарегестрированный. Давайте также считать, что на время эксперимента распределение ξ не меняется, и результат измерения прибора не зависит от предыдущих измерений. Пусть $X_1, \ldots, X_n$ – эти результаты измерения в какие-то моменты времени. На языке теории вероятностей это можно переформулировать так: $X_1, \ldots, X_n$ – *реализации независимых одинаково распределенных случайных величин* $\xi_1, \ldots, \xi_n$.

Задача состоит в том, чтобы оценить $E\xi$ по этим самым $X_1, \ldots, X_n$.

Пример (Регрессионная модель)**.** Пусть материальная точка движется по прямой, стартовав из точки x_0 с постоянной скоростью, равной v_0. Мы их не знаем, и будем считать, что это случайные величины. Пусть $x_1, \ldots, x_n$ – это измеренные нами положения этой материальной точки в моменты времени $t_1, \ldots, t_n$ соответственно. Или, по другому, $x_1, \ldots, x_n$ – это реализации случайных величин $\xi_1, \ldots, \xi_n$, причем ξ_i отвечает за измеренный нами результат положения точки в момент времени t_i. Понятно, что эти случайные величины уже будут зависимы. Дополнительно положим, что *погрешность измерения подчиняется нормальному распределению.* То есть: $\xi_i = x_0 + v_0 \cdot t_i + \varepsilon_i$, где ε_i – нормально распределенная случайная величина, отвечающая за ошибку i-того измерения.

Задача заключается в том, чтобы оценить x_0 и v_0 по этим данным $(x_0, \ldots, x_n, t_0, \ldots, t_n)$

Пример (Проверка на однородность)**.** Пусть $X_1, \ldots, X_n$ – это результаты эксперимента в условиях A, а $Y_1, \ldots, Y_m$ – результаты того-же самого эксперимента в условиях B. Нужно выяснить, влияют ли эти условия на результат. (Если для сокращения записи отождествить результат эксперимента со случайной величиной, реализацией которой он является, а также считать, что $X_i \sim X$, $Y_i \sim Y$ то можно записать так: $X \overset{d}{\sim} Y$?)

Замечание. Как мы видим, задача матстатистики – представить оптимальное решение на основе статистических данных. Типичная характерная черта задач – это довольно большое количество дополнительных ограничений на природу явлений (независимость и одинаковая распределенность результата, нормальное распределение погрешностей и т.д.). Такие ограничения в реальных условиях иногда бывает трудно проверить, поэтому нужно быть крайне внимательным при применении какого-либо результата из матстатистики в реальных задачах. Однако такое требование к внимательности компенсируется тем, что результаты из матстатистики находят широкое применение в экспериментальной физике, машинном обучении, data mining и прочих облостях науки.

1.2 Сходимости случайных векторов

Определение 1.1. Пусть $\{\xi_n : n \in \mathbb{N}\}$ – последовательность случайных векторов из $\mathbb{R}^m$. ξ – случайный вектор из $\mathbb{R}^m$. Говорят что:

- ξ_n сходится к ξ почти наверное (обозначение: $\xi_n \overset{\text{п.н.}}{\to} \xi$), если :
$$P(\{\omega : \lim_{n\to\infty} \xi_n(\omega) = \xi(\omega)\}) = 1$$
- ξ_n сходится к ξ по вероятности (обозначение $\xi_n \overset{P}{\to} \xi$), если:
$$\forall \varepsilon > 0 : \lim_{n\to\infty} P(\{\omega : \|\xi_n(\omega) - \xi(\omega)\| > \varepsilon\}) = 0$$
- ξ_n сходитіся к ξ по распределению (обозначение $\xi_n \overset{d}{\to} \xi$), если:
$$\forall f \text{ – ограниченной непрерывной функции } \mathbb{R}^m \mapsto \mathbb{R} : \lim_{n\to\infty} Ef(\xi_n) = Ef(\xi)$$

Утверждение 1.1. *Пусть* $\xi_n = (\xi_n^1, \dots, \xi_n^m), \xi = (\xi^1, \dots, \xi^m)$. *Тогда:*

- $\xi_n \overset{\text{п.н.}}{\to} \xi \Leftrightarrow \forall j : \xi_n^j \overset{\text{п.н.}}{\to} \xi^j$
- $\xi_n \overset{P}{\to} \xi \Leftrightarrow \forall j : \xi_n^j \overset{P}{\to} \xi^j$

Замечание. Для сходимости по распределению это утверждение не верно

Определение 1.2. Функции распределения F_{ξ_n} называются слабо сходящимися к F_ξ (обозначение: $F_{\xi_n} \overset{w}{\to} F_\xi$), если:

$$\forall f \text{ – ограниченной непрерывной функции } \mathbb{R} \mapsto \mathbb{R} : \int_\Omega f(x) dF_{\xi_n} \to \int_\Omega f(x) dF_\xi$$

Теорема 1.2 (Александрова). *Пусть $\xi, \xi_1, \xi_2, \dots$ – случайные величины. Тогда следующие утверждения эквивалентны:*

1. $\xi_n \xrightarrow{d} \xi$
2. $F_{\xi_n} \xrightarrow{w} F_\xi$
3. $\forall x$ – *точка непрерывности* $F_\xi : \lim_{n\to\infty} F_{\xi_n}(x) = F_\xi(x)$

Теорема 1.3 (многомерный случай, более слабая). *Пусть $\xi, \xi_1, \xi_2, \dots$ – случайные векторы из $\mathbb{R}^m$. Пусть F_ξ непрерывна. Тогда $\xi_n \xrightarrow{d} \xi \Leftrightarrow \forall x \in \mathbb{R}^m : \lim_{n\to\infty} F_{\xi_n}(x) = F_\xi(x)$*

1.3 Предельные теоремы

Теорема 1.4 (Закон больших чисел). *Пусть $\xi_1, \xi_2, \dots$ – попарно-некоррелированные одинаково распределенные случайные величины, $D\xi_i$ конечна, $S_n = \xi_1 + \dots + \xi_n$. Тогда:*

$$\frac{S_n - ES_n}{n} \xrightarrow{P} 0$$

, причем $ES_n = n \cdot a$, $a = E\xi_i$.

Теорема 1.5 (Усиленный закон больших чисел). *Пусть $\xi_1, \xi_2, \dots$ – независимые одинаково распределнные случайные величины (или векторы), $E\xi_i = a$ – конечно, $S_n = \xi_1 + \dots + \xi_n$. Тогда:*

$$\frac{S_n}{n} \xrightarrow{п.н.} a$$

Теорема 1.6 (Центральная предельная теорема). *Пусть $\xi_1, \xi_2, \dots$ – независимые одинаково распределнные случайные величины, $E\xi_i = a$ – конечно, $0 < D\xi_i = \sigma^2$ – тоже конечно, $S_n = \xi_1 + \dots + \xi_n$. Тогда:*

$$\sqrt{n}\left(\frac{S_n}{n} - a\right) \xrightarrow{d} \mathcal{N}(0, \sigma^2)$$

Теорема 1.7 (Многомерная центральная предельная теорема). *Пусть $\xi_1, \xi_2, \dots$ – независимые одинаково распределнные случайные векторы, $E\xi_i = a$ – конечно, $0 < D\xi_i = \Sigma$ – матрица ковариаций, тоже конечна, $S_n = \xi_1 + \dots + \xi_n$. Тогда:*

$$\sqrt{n}\left(\frac{S_n}{n} - a\right) \xrightarrow{d} \mathcal{N}(0, \Sigma)$$

Утверждение 1.8. *Пусть $\xi, \xi_1, \xi_2, \dots$ – случайные векторы. Тогда:*

1. $\xi_n \xrightarrow{п.н.} \xi \Rightarrow \xi_n \xrightarrow{P} \xi$
2. $\xi_n \xrightarrow{P} \xi \Rightarrow \xi_n \xrightarrow{d} \xi$

Доказательство.

1. Следует из того, что векторная сходимость эквивалентна покоординатной, а для координат верны одномерные аналоги.

2. Доказывается аналогично одномерному случаю.

□

Лемма 1.9 (о сходящейся подпоследовательности)**.** *Пусть* $\xi_n \xrightarrow{P} \xi$. *Тогда* $\exists$ *подпоследовательность* $\xi_{n_k} : \xi_{n_k} \xrightarrow{п.н.} \xi$

Теорема 1.10 (о наследовании сходимостей)**.** *Пусть* $\xi, \xi_1, \xi_2, \dots$ *– случайные векторы из* $\mathbb{R}^m$. *Пусть также* $h : \mathbb{R}^m \mapsto \mathbb{R}^s$ *– функция, непрерывная почти всюду относительно распределения* ξ *(то есть* $\exists B \in \mathcal{B}(\mathbb{R}^m), P_\xi(B) = 1 : h$ *непрерывна на* B. *Тогда:*

1. $\xi_n \xrightarrow{п.н.} \xi \Rightarrow h(\xi_n) \xrightarrow{п.н.} h(\xi)$

2. $\xi_n \xrightarrow{P} \xi \Rightarrow h(\xi_n) \xrightarrow{P} h(\xi)$

3. *Если дополнительно* h *непрерывна всюду, а не почти всюду, то:* $\xi_n \xrightarrow{d} \xi \Rightarrow h(\xi_n) \xrightarrow{d} h(\xi)$

Доказательство.

1. $1 = P(\{\omega : \xi_n(\omega) \to \xi(\omega)\}) \le P(\{\omega : h(\xi_n)(\omega) \to h(\xi)(\omega)\})$, так как $\{\omega : \xi_n(\omega) \to \xi(\omega)\} \subset \{\omega : h(\xi_n)(\omega) \to h(\xi)(\omega)\}$

2. Пусть не выполнено, что $h(\xi_n) \xrightarrow{P} h(\xi)$. Тогда $\exists \varepsilon, \delta$ а так же подпоследовательность ξ_{n_k}:
$$\forall k : P(\|h(\xi_{n_k}) - h(\xi)\| \ge \varepsilon) \ge \delta$$
. Так как $\xi_{n_k} \xrightarrow{P} \xi$, то существует подпоследовательность $\xi_{n_{k_l}} : \xi_{n_{k_l}} \xrightarrow{п.н.} \xi$. Тогда согласно первому пункту $h(\xi_{n_{k_l}}) \xrightarrow{п.н.} h(\xi)$. Но такого быть не может, т.к. $\forall k : P(\|h(\xi_{n_k}) - h(\xi)\| \ge \varepsilon) \ge \delta$

3. Пусть $f : \mathbb{R}^m \mapsto \mathbb{R}$ – непрерывная ограниченная функция. Тогда $g = f \circ h$ – тоже непрерывная ограниченная функция. Так как $\xi_n \xrightarrow{d} \xi$, то $Eg(\xi_n) \to Eg(\xi)$. А значит, $Ef(h(\xi_n)) \to Ef(h(\xi))$

□

Лемма 1.11 (Слуцкого)**.** *Пусть* $\xi_n \xrightarrow{d} \xi$ *– случайные величины, а* $\eta_n \xrightarrow{d} C = const$. *Тогда* $\xi_n + \eta_n \xrightarrow{d} \xi + C$ *и* $\xi_n \cdot \eta_n \xrightarrow{d} \xi \cdot C$.

Доказательство. Пусть t – точка непрерывности функции распределения $F_{\xi+C}$ случайной величины $\xi + C$. Докажем только для суммы, для произведения аналогично. Пусть $\varepsilon > 0$ такое,что $t \pm \varepsilon$ – тоже точки непрерывности

$F_{\xi+C}$. Мы хотим показать, что $F_{\xi_n+\eta_n}(t) \to F_{\xi+C}(t)$. Будем для этого пользоваться тем, что $\eta_n \xrightarrow{d} C \Leftrightarrow \eta_n \xrightarrow{P} C$.

$$P(\xi_n + \eta_n \le t) = P(\xi_n + \eta_n \le t, \eta_n \ge C - \varepsilon) + P(\xi_n + \eta_n \le t, \eta_n < C - \varepsilon) \le$$
$$P(\xi_n \le t - C + \varepsilon) + P(\|\eta_n - C\| > \varepsilon) = F_{\xi_n+c}(t+\varepsilon) + P(\|\eta_n - C\| > \varepsilon)$$

Но $\xi_n \xrightarrow{d} \xi \Rightarrow \xi_n + C \xrightarrow{d} \xi + C$. Кроме того, $t+\varepsilon$ – точка непрерывности $F_{\xi+C}$

$$\limsup_{n\to\infty} F_{\xi_n+\eta_n}(t) \le \lim_{n\to\infty} F_{\xi_n+c}(t+\varepsilon) + \lim_{n\to\infty} P(\|\eta_n - C\| > \varepsilon) = F_{\xi+C}(t+\varepsilon)$$

Аналогично,

$$1 - f_{\xi_n+\eta_n}(t) \le P(\|\eta_n - C\| > \varepsilon) + 1 - F_{\xi_n+C}(t-\varepsilon)$$

Откуда

$$\liminf_{n\to\infty} F_{\xi_n+\eta_n}(t) \ge F_{\xi+C}(t-\varepsilon)$$

В силу произвольности $\varepsilon > 0$ и того факта, что t – точка непрерывности функции распределения $F_{\xi+c}$ получаем, что

$$\lim_n F_{\xi_n+\eta_n} = F_{\xi+C}(t)$$

□

Утверждение 1.12 (применение леммы Слуцкого)**.** *Пусть $\xi_n \xrightarrow{d} \xi$ – случайные величины. $h(x)$ – дифференцируема в точке $a \in \mathbb{R}$. $\{b_n > 0\}$ – числовая последовательность, $b_n \to 0$. Тогда:*

$$\frac{h(a+\xi_n b_n) - h(a)}{b_n} \xrightarrow{d} h'(a)\xi$$

Доказательство. По лемме Слуцкого $\xi_n b_n \xrightarrow{d} \xi_n \cdot 0 = 0$. Рассмотрим

$$H(x) = \begin{cases} \frac{h(x+a)-h(a)}{x}, x \ne 0 \\ h'(a), x = 0 \end{cases}$$

$H(x)$ непрерывна в 0 по определению, а также непрерывна на $\mathbb{R} \setminus 0$ как композиция непрерывных функций. По теореме о наследовании сходимостей $H(\xi_n b_n) \xrightarrow{d} H(0) = h'(a)$. По лемем слуцкого $\xi_n H(\xi_n b_n) \xrightarrow{d} h'(a)\xi$ □

Теорема 1.13 (многомерный вариант)**.** *Пусть $\xi_n \xrightarrow{d} \xi$ – случайные векторы. $h(x)$ – дифференцируема в точке $a \in \mathbb{R}$ (то есть существует матрица частных производных, или матрица Якоби $J(h)$). $\{b_n > 0\}$ – числовая последовательность, $b_n \to 0$. Тогда:*

$$\frac{h(a+\xi_n b_n) - h(a)}{b_n} \xrightarrow{d} J(h)(a)\xi$$

2 Вероятностно-статистические модели и выборки

2.1 Вероятностно-статистическая модель

Определение 2.1. Множество всех возможных значений наблюдения называется выборочным пространством и обозначается $\mathfrak{X}$

Определение 2.2. Наблюдение X – это результат случайного выбора элемента из выборочного пространства. Наша цель – по наблюдению X сделать выводы о его распределении P.

Определение 2.3. Если $X = (X_1, \dots, X_n)$ – набор независимых одинаково-распределенных случайных величин имеющих распределению P, то X называется выборкой размера n из неизвестного распределения P

Замечание. В дальнейшем будем обозначать: $X = (X_1, \dots, X_n)$ – выборка на вероятностном пространстве $(\Omega, \mathcal{F}, P)$ с неизвестным распределением P на $(\mathfrak{X}, \mathcal{B}_x)$ (например, $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$).

Определение 2.4. Для каждого множества $B \in \mathcal{B}_x$) введем $P_n^*(B) = \frac{\nu_n(B)}{n}$, где $\nu_n(B)$ – это количество элементов из $X_1, \dots, X_n$, которые попали в B. То есть формально:

$$P_n^*(B) = \frac{1}{n}\sum_{i=1}^{n} I\{X_i \in B\}$$

. Такая величина называется эмпирическим распределением.

Утверждение 2.1. *Пусть P – неизвестное распределение X_i. Тогда $\forall B \in \mathcal{B}_x : P_n^*(B) \overset{п.н.}{\to} P(B)$.*

Доказательство. Фиксируем $B \in \mathcal{B}_x$. Тогда для фиксированного B индикаторы $I\{X_i \in B\}$ будут являтся случайными величинами, причем независимыми и одинаково распределенными, поскольку исходные случайные величины были независимыми и одинаково распределенными. Введем

$$S_n = \sum_{i=1}^{n} I\{X_i \in B\}$$

По усиленному закону больших чисел:

$$\frac{1}{n}S_n \overset{\text{п.н.}}{\to} P(X_1 \in B) = P(B)$$

□

Определение 2.5. Пусть $X = (X_1, \dots, X_n)$ – выборка.

$$F_n(x) = \frac{1}{n}\sum_{i=1}^{n} I\{X_i \le x\}$$

называется эмпирической функцией распределения (она является функцией распределения для эмпирического распределения P_n^*).

Замечание.

$$\forall x \in \mathbb{R} : F_n(x) \overset{\text{п.н.}}{\to} F(x)$$

Теорема 2.2 (Гливенко - Кантелли)**.** *Если $F(x)$ – функция распределения элементов выборки $X_1, \ldots, X_n$ (то есть функция распределения для распределения P), то:*

$$\sup_{x\in\mathbb{R}} |F_n(x) - F(x)| \overset{п.н.}{\to} 0$$

Доказательство. Зафиксируем элементарный исход $\omega \in \Omega$. Тогда случайные величины $X_1, \ldots, X_n$ превращаются в числа. Посмотрим на функцию распределения $F_n(x)$. Она является непрерывной справа, так как является конечной суммой индикаторов вида $I\{X_i \le x\}$, а $F(x)$ непрерывна справа как функция распределения. Модуль их разности тоже непрерывен справа. Тогда $\sup_{x\in\mathbb{R}} |F_n(x)-F(x)| = \sup_{x\in\mathbb{Q}} |F_n(x)-F(x)|$ – это супремум счетного числа случайных величин. Поэтому $|F_n(x)-F(x)|$ тоже является случайной величиной.

Пусть $N \in \mathbb{N}$ – достаточно большое. Для каждого $k \in \{1, \ldots, N-1\}$ введем $x_{k,N} = \inf\{x : F(x) \ge \frac{k}{N}\}$. Полагаем также $x_{0,N} = -\infty, x_{N,N} = +\infty$. Оценим $F_n(x) - F(x)$ для $x \in [x_{k,N}, x_{k+1,N})$:

$$\begin{aligned} F_n(x) - F(x) \le F_n(x_{k+1,N} - 0) - F(x_{k,N}) = \\ = F_n(x_{k+1,N} - 0) - F(x_{k+1,N} - 0) + F(x_{k+1,N} - 0) - F(x_{k,N}) \le \\ F_n(x_{k+1,N} - 0) - F(x_{k+1,N} - 0) + \frac{1}{n} \end{aligned}$$

Совершенно аналогично показывается, что $F_n(x)-F(x) \ge F_n(x_{k,N})-F(x_{k,N})-\frac{1}{N}$ Откуда получаем, что:

$$\sup_{x\in\mathbb{Q}} |F_n(x)-F(x)| \le \max_{k,l}\{|F_n(x_{k,N}-0)-F(x_{k,N}-0)|, |F_n(x_{l,N})-F(x_{l,N})|\}+\frac{1}{N}$$

Зафиксируем $\varepsilon > 0$ и возьмем $N : \frac{1}{N} < \varepsilon$. По усиленному закону больших чисел $F_n(x) \overset{\text{п.н.}}{\to} F_(x), F_n(x-0) \overset{\text{п.н.}}{\to} F(x-0)$. То есть $\forall x \in \mathbb{Q} : P(\limsup_n |F_n(x)-F(x)| > \varepsilon) = 0$. А значит, $P(\sup_{x\in\mathbb{Q}} \limsup_n |F_n(x)-F(x)| > \varepsilon) = 0$. Поменяв местами sup и lim sup и считая $\varepsilon = \frac{1}{m}$, получаем:

$$\forall m : P(\limsup_n \sup_{x\in\mathbb{Q}} |F_n(x) - F(x)| > \frac{1}{m}) = 0$$

Пользуясь теоремой о непрерывности вероятностной меры, имеем:

$$P(\lim_n \sup_{x\in\mathbb{Q}} |F_n(x) - F(x)| > 0) = 0$$

$\square$

Замечание. Пусть на $\mathfrak{X}$ задана σ-алгебра $\mathcal{B}_x$. Как правило $(\mathfrak{X}, \mathcal{B}_x) = (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ Наблюдение X – это, формально, тождественная случайная величина на $(\mathfrak{X}, \mathcal{B}_x, P)$. Обычно про распределение $P = P_X$ известно, что оно принадлежит некому классу распределений $\mathcal{P}$, например, классу нормальных распределений

Определение 2.6. Тройка $(\mathfrak{X}, \mathcal{B}_x, \mathcal{P})$, где $\mathcal{P}$ – это класс вероятностных мер на $(\mathfrak{X}, \mathcal{B}_x)$, называется вероятностно-статистической моделью

Замечание. $\forall P \in \mathcal{P} : (\mathfrak{X}, \mathcal{B}_x, P)$ – вероятностное пространство

Определение 2.7. Вероятностно-статистическая модель $(\mathfrak{X}, \mathcal{B}_x, \mathcal{P})$ называется параметрической, если класс $\mathcal{P}$ параметризован, то есть $\mathcal{P} = \{P_\theta : \theta \in \Theta\}$. Также считаем, что $P_{\theta_1} \neq P_{\theta_2}$, если $\theta_1 \neq \theta_2$

2.2 Моделирование выборки

2.2.1 Конечная выборка

Мы хотим смоделировать конечную выборку $(X_1, \ldots, X_n)$ в терминах веротяностно-статистической модели. Пусть X_i принимает значения из $\mathfrak{X}$ и имеет неизвестное распределение $P \in \mathcal{P}$. В этом случае удобно рассмотреть следующую статистическую модель: $(\mathfrak{X}^n, \mathcal{B}_x^n, \mathcal{P}^n)$, где $\mathfrak{X}^n = \mathfrak{X} \times \ldots \times \mathfrak{X}$, $\mathcal{B}_x^n = \sigma(B_1 \times \ldots \times B_n : \forall i : B_i \in \mathcal{B}_x)$, $\mathcal{P}^n = \{P^n : P \in \mathcal{P}\}$, $P^n(B_1 \times \ldots \times B_n) = P(B_1) \cdot \ldots \cdot P(B_n)$. То есть в качестве выборочного пространства мы берем декартову степень, в качестве сигма алгебры, как и в случае с $\mathcal{B}(\mathbb{R}^n)$, минимальную сигма-алгебру, порожденную декартовыми произведениями всех измеримых множеств, а в качестве класса мер – класс степеней всех мер, где степень меры означает естественное продолжение одномерной меры на многомерное пространство. Какими же выбрать $(X_1, \ldots, X_n)$? Это просто: $X_i : \mathfrak{X}^n \mapsto \mathfrak{X}$ такое что $X_i((x_1, \ldots, x_n)) = x_i$.

2.2.2 Счетная выборка

TO BE CONTINUED...

Замечание. Мы будем опускать индексы у $\mathfrak{X}$, $\mathcal{B}_x$, $\mathcal{P}$ в целях удобства

Замечание. В параметрической модели вопрос о неизвестном распределении P_θ сводится в вопросу о значении $\theta \in \Theta$

2.3 Статистики и оценки

Пусть X – наблюдение со значениями из $\mathfrak{X}$ и неизвестным распределением P_θ, где $\theta \in \Theta$

Определение 2.8. Статистикой $S(X)$ называется измеримая функция от результатов наблюдения, то есть: $S : \mathfrak{X} \mapsto E$, где $(E, \mathcal{E})$ – измеримое пространство, а S является $\mathcal{E}|\mathcal{B}_x$ измеримой.

Если $S : \mathfrak{X} \mapsto \Theta$, то S называется оценкой параметра из Θ.

Пример.

1. Если $g : \mathbb{R} \mapsto \mathbb{R}$ – борелевская функция, то $\overline{g(x)} = \frac{1}{n}\sum_{i=1}^{n} g(X_i)$ называется выборочной характеристикой функции $g(x)$ (среднее значение по элементам выборки)