

Конспекты к экзамену по математической статистике

June 1, 2016

Contents

0.0	Список вопросов к экзамену по математической статистике . . .	2
0.1	Случайная выборка, генеральная совокупность, функция распределения выборки	3
0.2	Эмпирическая функция распределения, гистограмма	4

0.0 Список вопросов к экзамену по математической статистике

1. 1. Случайная выборка и генеральная совокупность
2. Функция распределения выборки
2. 1. Эмпирическая функция распределения
2. Гистограмма
3. Выборочные характеристики. Выборочные моменты
4. Точечные оценки и их свойства
5. Функция правдоподобия. Неравенство Крамера-Рао
6. Метод максимального правдоподобия, свойства оценок максимального правдоподобия
7. Метод моментов для точечных оценок
8. Достаточные статистики
9. Интервальные оценки. Доверительные интервалы
10. Интервальные оценки.
Доверительные интервал для дисперсии нормальной генеральной совокупности
11. Асимптотические свойства оценки максимального правдоподобия.
Асимптотический доверительный интервал
12. Проверка статистических гипотез.
Критерий Неймана-Пирсона проверки простых гипотез
13. Наиболее мощный критерий. Теорема Неймана-Пирсона
14. Проверка статистических гипотез о параметрах нормального распределения
15. Критерии для сложных гипотез
16. Функция мощности при альтернативе
17. Критерий согласия χ^2 -Пирсона
18. Критерий согласия Колмогорова
19. Критерий однородности Колмогорова-Смирнова
20. Критерий однородности χ^2

0.1 Случайная выборка, генеральная совокупность, функция распределения выборки

Def. 1. Выборка Пусть эксперимент состоит в проведении n испытаний, результат j -го из которых является случайной величиной $X_j : \Omega_j \rightarrow \mathcal{X}_j$.

Кортеж из этих случайных величин (случайный вектор) $X = (X_1, \dots, X_n)$ называется (случайной) выборкой, а г.в. X_j называются элементами выборки

А значение $x = (x_1, \dots, x_n) = X(\omega)$ называется реализацией выборки

Далее всегда, если не указано иное, случайные величины будут обозначаться заглавными буквами, а их реализации соответствующими строчными

Далее X_j полагаются независимыми

Def. 2. Выборочное пространство Выборочным пространством называется измеримое пространство $(\mathcal{X}, \mathcal{A})$, где $\mathcal{X} = \{X(\omega); \omega \in \Omega\}$ есть множество возможных значений выборки, а \mathcal{A} — σ -алгебра

Особенно важен случай, когда случайные величины X_j являются независимыми и имеют распределение одной случайной величины ξ . Этот случай соответствует повторению n раз одного эксперимента, описываемого случайной величиной ξ

Def. 3. Генеральная совокупность Генеральной совокупностью называют распределение $\mathcal{L}(\xi)$ случайной величины ξ

Оно может быть задано, например, множеством возможных значений г.в. ξ и её функцией распределения

При этом X называют выборкой из (генеральной совокупности) $\mathcal{L}(\xi)$

Def. 4. Функция распределения выборки $X \in \mathcal{L}(\xi)$

$$F_X(x) = \mathbb{P}\{X \leq x\} = \prod \mathbb{P}\{X_j \leq x_j\} = \prod F_{X_j}(x_j)$$

0.2 Эмпирическая функция распределения, гистограмма

Пусть $A \subset \Omega_0$ событие, происходящее в ходе испытания с вероятностью $\mathbb{P}A = p$, и пусть эксперимент состоит в проведении n таких независимых испытаний

Тогда

$$\Omega = \prod_{j=1}^n \Omega_0$$

A случайная величина

$$X_j = I_{\{\omega; \omega_j \in A\}} = \begin{cases} 1; & \omega_j \in A \\ 0; & \omega_j \notin A \end{cases}$$

является индикатором того, что в ходе j -го испытания случилось событие A

Пусть г.в. $k = \sum_{j=1}^n X_j$ — число проявлений A в ходе эксперимента.

Введём г.в. $p_n^* = \frac{1}{n} \sum_{j=1}^n X_j$.

Очевидно $\mathbb{E}p_n^* = p$.

Кроме того, из ЗБЧ в форме Бернулли следует $\lim_{n \rightarrow \infty} \mathbb{P}\{|p_n^* - p| < \varepsilon\} = 1 \quad \forall \varepsilon > 0$

Таким образом, значение случайной величины p_n^* можно считать приближённой оценкой величины p

Пусть теперь $X = (X_1, \dots, X_n)$ — выборка объёма n из генеральной совокупности $\mathcal{L}(\xi)$, $x = (x_1, \dots, x_n)$ — реализация.

Def. 5. Порядковые статистики Каждой реализации x можно сопоставить в соответствие его перестановку $x_{(1)} \leq \dots \leq x_{(n)}$,

j -й порядковой статистикой называется случайная величина $X_{(j)}$, при каждой реализации $X(\omega) = x$, принимает значение $X_{(j)}(\omega) = x_{(j)}$

Def. 6. Вариационный ряд Случайный вектор $(X_{(1)}, \dots, X_{(n)})$ называется вариационным рядом

Def. 7. Эмпирическая функция распределения Для каждого $t \in \mathbb{R}$ зададим случайную величину $\mu_n(x)$, равную количеству элементов выборки X , значения которых не превосходят t :

$$\mu_n(x) = \sum I_{\{X_j \leq t\}}$$

Эмпирической функцией распределения, построенной по выборке X , называют случайную функцию $t \mapsto F_n(t)$

$$F_n(x) = \frac{1}{n} \mu_n(t)$$

Её значение в точке t является случайной величиной, сходящейся по вероятности к значению $F(t)$ теоретической функции распределения

EDF можно переписать с помощью функции Хевисайда (Heaviside):

$$H(t) = \begin{cases} 0; & t < 0 \\ 1; & t \geq 0 \end{cases}$$

$$F_n(t) = \frac{1}{n} \sum_{j=1}^n H(t - X_{(k)})$$

Def. 8. Гистограмма Разобьём область значений г.в. ξ на равные интервалы Δ_i , и для каждого Δ_i подсчитаем число n_i элементов x_j вектора x , попавших в Δ_i , $n = \sum n_i$.

Построим график ступенчатой функции

$$t \mapsto \frac{n_i}{nh_i}, \quad t \in \Delta_i, h_i = |\Delta_i|$$

Полученный график (при желании, само отображение) называется Гистограммой, построенной по данной реализации выборки

Соединим середины смежных отрезков этого графика. Полученная ломанная называется полигоном частот

С уменьшением $\max\{h_i\}$, гистограмма и полигон частот всё более точно приближают вероятности попадания в каждый из интервалов разбиения