Федеральное государственное автономное образовательное учреждение высшего образования
Санкт-Петербургский Политехнический университет Петра Великого
Физико-Механический институт

Лабораторная 5

Выполнил студент гр. 5030102/20101: Бугайцев М.В. ______________
Преподаватель: Баженов А. Н. ______________
Работа принята: Дата ______________

Санкт-Петербург
2025

Содержание

1 Постановка задачи

Сгенерировать выборку объёмом 100 элементов для нормального распределения $N(x, 0, 1)$. По сгенерированной выборке оценить параметры μ и σ нормального закона методом максимального правдоподобия. В качестве основной гипотезы H_0 будем считать, что сгенерированное распределение имеет вид $N(x, \hat{\mu}, \hat{\sigma})$. Проверить основную гипотезу, используя критерий согласия χ^2. В качестве уровня значимости взять $\alpha = 0.05$. Привести таблицу вычислений χ^2. Исследовать точность (чувствительность) критерия χ^2 – сгенерировать выборки равномерного распределения и равномерного распределения объёма 20 элементов. Проверить их на нормальность.

2 Теоретическое обоснование

2.1 Основные понятия

Гипотеза H_0 - это предположение о том, что исследуемая выборка принадлежит генеральной совокупности с заданным законом распределения $F(x)$. В данной работе:

- в качестве H_0 выдвинута гипотеза о нормальном распределении $N(x, \mu, \sigma)$;
- альтернативная гипотеза H_1 - распределение не является нормальным.

Критерий согласия χ^2 - метод проверки гипотезы о законе распределения, основанный на сравнении наблюдаемых частот n_i с теоретическими частотами np_i, где p_i - вероятность попадания в интервал Δ_i при условии истинности H_0.

2.2 Метод максимального правдоподобия (ММП)

Для оценки параметров нормального распределения μ и σ используются:

- Оценка для μ:
$$\hat{\mu} = \frac{1}{n} \sum_{i=1}^{n} x_i.$$
- Оценка для σ (несмещённая):
$$\hat{\sigma} = \sqrt{\frac{1}{n-1} \sum_{i=1}^{n} (x_i - \hat{\mu})^2}.$$

Эти оценки обеспечивают наилучшее соответствие выборки предполагаемому нормальному распределению.

2.3 Правило проверки гипотезы о законе распределения по методу χ^2

1. Выбираем уровень значимости α.
2. По таблице находим квантиль $\chi^2_{1-\alpha}(k-1)$ распределения хи-квадрат с k$-$1 степенями свободы порядка $1 - \alpha$.
3. С помощью гипотетической функции распределения $F(x)$ вычисляем вероятности $p_i = P(X \in \Delta_i)$, $i = 1, ..., k$.
4. Находим частоты n_i попадания элементов выборки в подмножества Δ_i, $i = 1, ..., k$.
5. Вычисляем выборочное значение статистики критерия χ^2:
$$\chi^2_B = \sum_{i=1}^{k} \frac{(n_i - np_i)^2}{np_i}. \tag{1}$$

6. Сравниваем χ_B^2 и квантиль $\chi_{1-\alpha}^2(k-1)$.
   а) Если $\chi_B^2 < \chi_{1-\alpha}^2(\text{k} - 1)$, то гипотеза H_0 на данном этапе проверки принимается.
   б) Если $\chi_B^2 >= \chi_{1-\alpha}^2(k-1)$, то гипотеза H_0 отвергается, выбирается одно из альтернативных распределений, и процедура проверки повторяется.

2.4 Исследование чувствительности критерия χ^2

Для проверки точности критерия генерируются:

1. Нормальное распределение $N(0,1)$:
   - Объём выборки $n = 100$
   - Ожидается, что H_0 будет приниматься (p-value > 0.05)
2. Равномерное распределение:
   - Объём выборки $n = 20$
   - Ожидается, что H_0 будет отвергаться (p-value ≤ 0.05)
   - Для равномерного распределения с дисперсией 1 используется интервал $[-\sqrt{3}, \sqrt{3}]$, так как:
     $$\sigma^2 = \frac{(\sqrt{3} - (-\sqrt{3}))^2}{12} = 1$$

3 Практическая часть

3.1 Нормальное распределение, $n = 20$

Таблица 1: Таблица расчёта статистики χ^2 для проверки нормальности выборки $n = 20,\ \alpha = 0.05$

Интервал	n_i	np_i	$\frac{(n_i-np_i)^2}{np_i}$
$[-0.98,\ -0.66)$	2	0.81	1.76
$[-0.66,\ -0.33)$	0	1.39	1.39
$[-0.33,\ -0.01)$	3	2.06	0.43
$[-0.01,\ 0.31)$	2	2.66	0.16
$[0.31,\ 0.63)$	5	2.98	1.36
$[0.63,\ 0.95)$	2	2.90	0.28
$[0.95,\ 1.28)$	1	2.45	0.86
$[1.28,\ 1.60)$	2	1.80	0.02
$[1.60,\ 1.92)$	2	1.14	0.64
$[1.92,\ 2.24)$	1	0.63	0.22

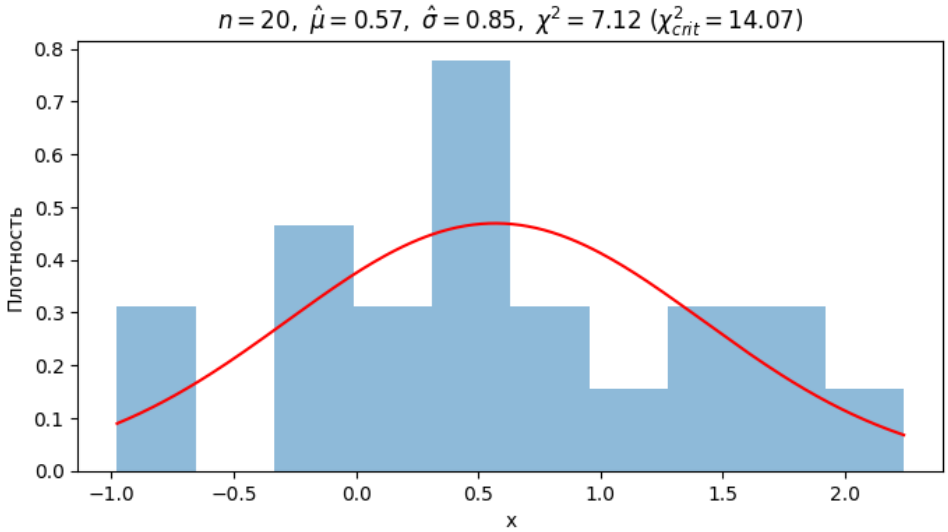

Рис. 1: Гистограмма и теоретическая плотность для нормальной выборки, $n = 20$.

3.2 Нормальное распределение, $n = 100$

Таблица 2: Таблица расчёта статистики χ^2 для проверки нормальности выборки $n = 100,\ \alpha = 0.05$

Интервал	n_i	np_i	$\frac{(n_i - np_i)^2}{np_i}$
$[-2.55,\ -2.07)$	1	1.58	0.21
$[-2.07,\ -1.59)$	5	3.86	0.34
$[-1.59,\ -1.11)$	10	7.70	0.69
$[-1.11,\ -0.62)$	13	12.53	0.02
$[-0.62,\ -0.14)$	18	16.63	0.11
$[-0.14,\ 0.34)$	13	18.00	1.39
$[0.34,\ 0.82)$	14	15.88	0.22
$[0.82,\ 1.31)$	13	11.43	0.22
$[1.31,\ 1.79)$	5	6.71	0.43
$[1.79,\ 2.27)$	8	3.21	7.15

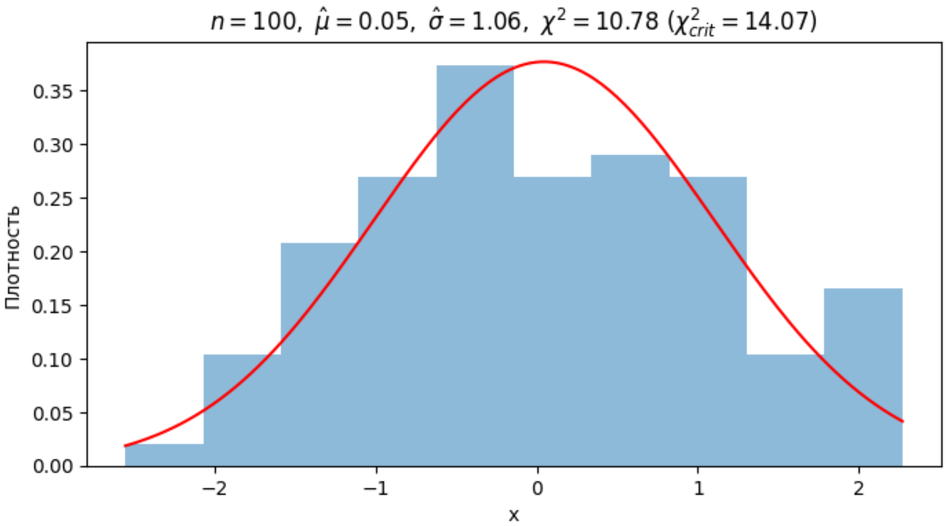

Рис. 2: Гистограмма и теоретическая плотность для нормальной выборки, $n = 100$.

3.3 Равномерное распределение, $n = 20$

Таблица 3: Таблица расчёта статистики χ^2 для проверки нормальности выборки $n = 20,\ \alpha = 0.05$

Интервал	n_i	np_i	$\frac{(n_i - np_i)^2}{np_i}$
$[-1.49,\ -1.17)$	4	0.73	14.53
$[-1.17,\ -0.85)$	1	1.10	0.01
$[-0.85,\ -0.53)$	0	1.52	1.52
$[-0.53,\ -0.20)$	1	1.92	0.44
$[-0.20,\ 0.12)$	1	2.22	0.67
$[0.12,\ 0.44)$	2	2.36	0.05
$[0.44,\ 0.76)$	2	2.29	0.04
$[0.76,\ 1.08)$	2	2.04	0.00
$[1.08,\ 1.41)$	4	1.66	3.29
$[1.41,\ 1.73)$	3	1.24	2.49

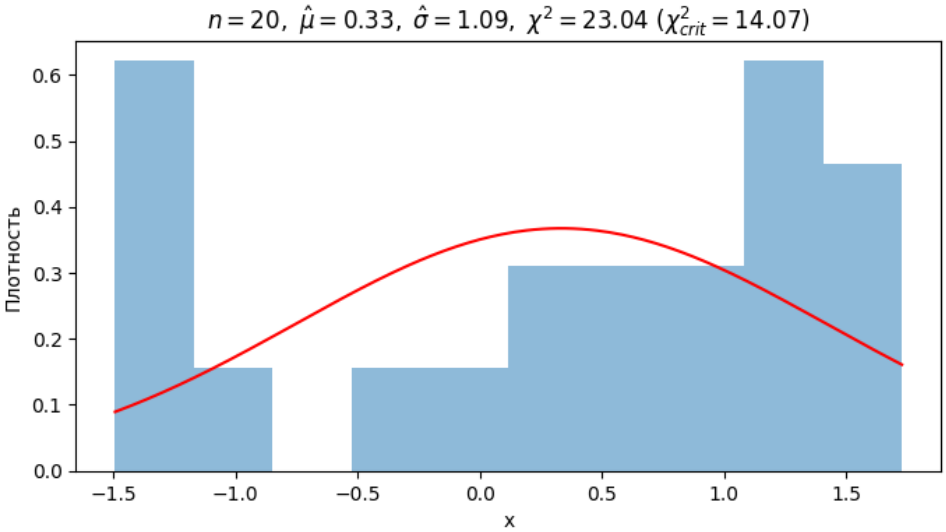

Рис. 3: Гистограмма и теоретическая плотность для равномерной выборки, $n = 20$.

3.4 Равномерное распределение, $n = 100$

Таблица 4: Таблица расчёта статистики χ^2 для проверки нормальности выборки $n = 100,\ \alpha = 0.05$

Интервал	n_i	np_i	$\frac{(n_i - np_i)^2}{np_i}$
$[-1.69,\ -1.35)$	12	5.13	9.21
$[-1.35,\ -1.01)$	10	7.75	0.65
$[-1.01,\ -0.67)$	15	10.46	1.98
$[-0.67,\ -0.34)$	11	12.57	0.20
$[-0.34,\ 0.00)$	6	13.49	4.16
$[0.00,\ 0.34)$	12	12.90	0.06
$[0.34,\ 0.68)$	8	11.01	0.82
$[0.68,\ 1.02)$	11	8.37	0.82
$[1.02,\ 1.36)$	4	5.68	0.50
$[1.36,\ 1.70)$	11	3.44	16.62

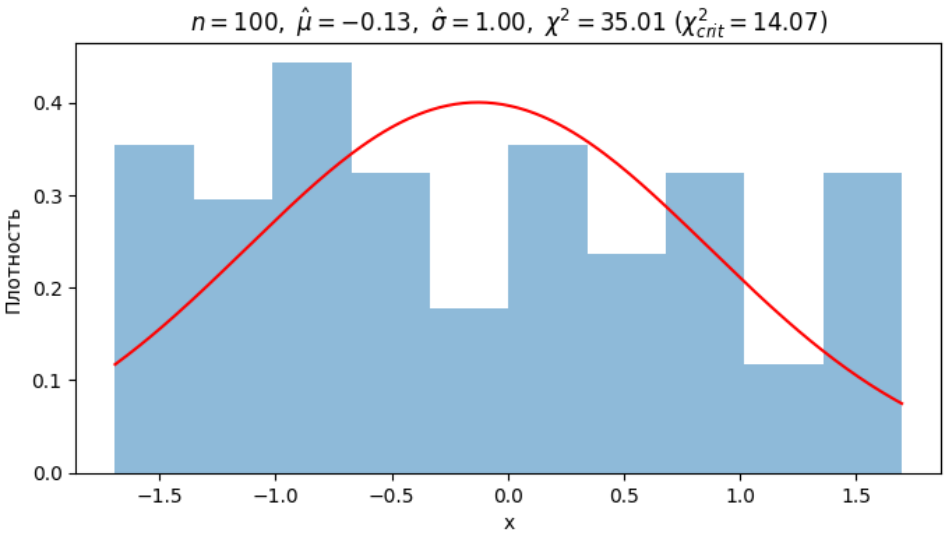

Рис. 4: Гистограмма и теоретическая плотность для равномерной выборки, $n = 100$.