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:

$$\chi^2_B = \sum_{i=1}^{k} \frac{(n_i - np_i)^2}{np_i}. \quad (1)$$

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

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	p_i	np_i	$\frac{(n_i - np_i)^2}{np_i}$
$(-\infty,\ -0.33)$	2	0.11	2.19	0.02
$[-0.33,\ 0.31)$	5	0.24	4.72	0.02
$[0.31,\ 0.95)$	7	0.29	5.88	0.21
$[0.95,\ 1.60)$	3	0.21	4.25	0.37
$[1.60,\ \infty)$	3	0.09	1.77	0.85

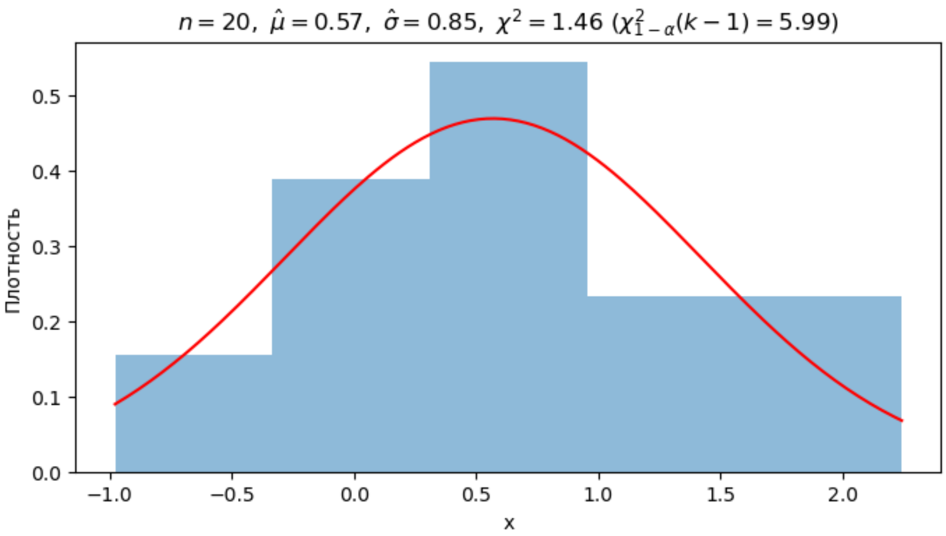

Рис. 1: Гистограмма и теоретическая плотность для нормальной выборки, $n = 20$.

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	p_i	np_i	$\frac{(n_i - np_i)^2}{np_i}$
$(-\infty,\ -2.07)$	1	0.02	1.58	0.21
$[-2.07,\ -1.59)$	5	0.04	3.86	0.34
$[-1.59,\ -1.11)$	10	0.08	7.70	0.69
$[-1.11,\ -0.62)$	13	0.13	12.53	0.02
$[-0.62,\ -0.14)$	18	0.17	16.63	0.11
$[-0.14,\ 0.34)$	13	0.18	18.00	1.39
$[0.34,\ 0.82)$	14	0.16	15.88	0.22
$[0.82,\ 1.31)$	13	0.11	11.43	0.22
$[1.31,\ 1.79)$	5	0.07	6.71	0.43
$[1.79,\ \infty)$	8	0.03	3.21	7.15

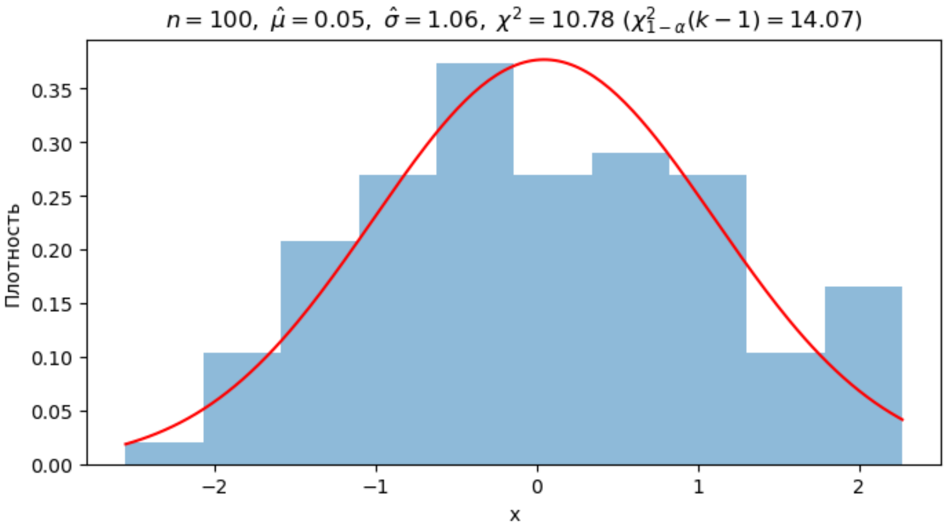

Рис. 2: Гистограмма и теоретическая плотность для нормальной выборки, $n = 100$.

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	p_i	np_i	$\frac{(n_i - np_i)^2}{np_i}$
$[-\sqrt{3},\ -0.85)$	5	0.09	1.84	5.44
$[-0.85,\ -0.20)$	1	0.17	3.44	1.73
$[-0.20,\ 0.44)$	3	0.23	4.58	0.54
$[0.44,\ 1.08)$	4	0.22	4.32	0.02
$[1.08,\ \sqrt{3}]$	7	0.15	2.90	5.78

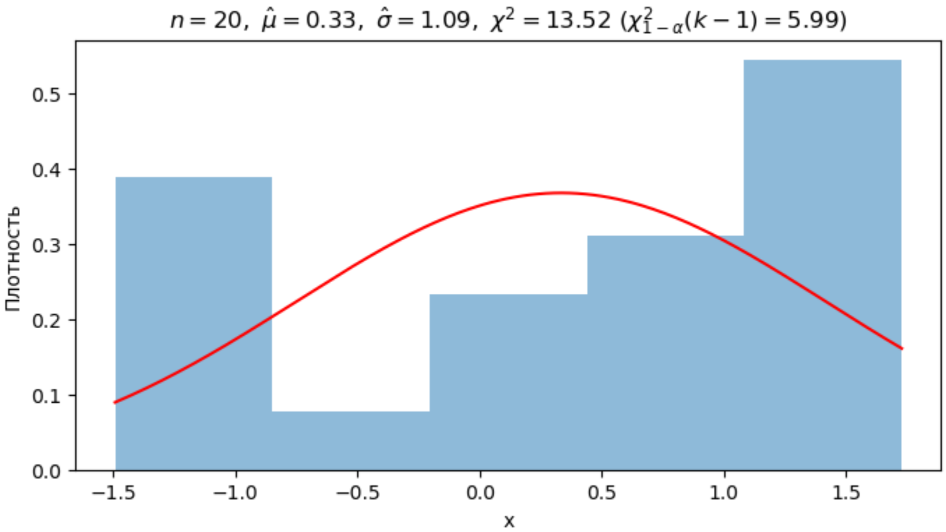

Рис. 3: Гистограмма и теоретическая плотность для равномерной выборки, $n = 20$.

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	p_i	np_i	$\frac{(n_i - np_i)^2}{np_i}$
$[-\sqrt{3},\ -1.35)$	12	0.05	5.13	9.21
$[-1.35,\ -1.01)$	10	0.08	7.75	0.65
$[-1.01,\ -0.67)$	15	0.10	10.46	1.98
$[-0.67,\ -0.34)$	11	0.13	12.57	0.20
$[-0.34,\ 0.00)$	6	0.13	13.49	4.16
$[0.00,\ 0.34)$	12	0.13	12.90	0.06
$[0.34,\ 0.68)$	8	0.11	11.01	0.82
$[0.68,\ 1.02)$	11	0.08	8.37	0.82
$[1.02,\ 1.36)$	4	0.06	5.68	0.50
$[1.36,\ \sqrt{3}]$	11	0.03	3.44	16.62

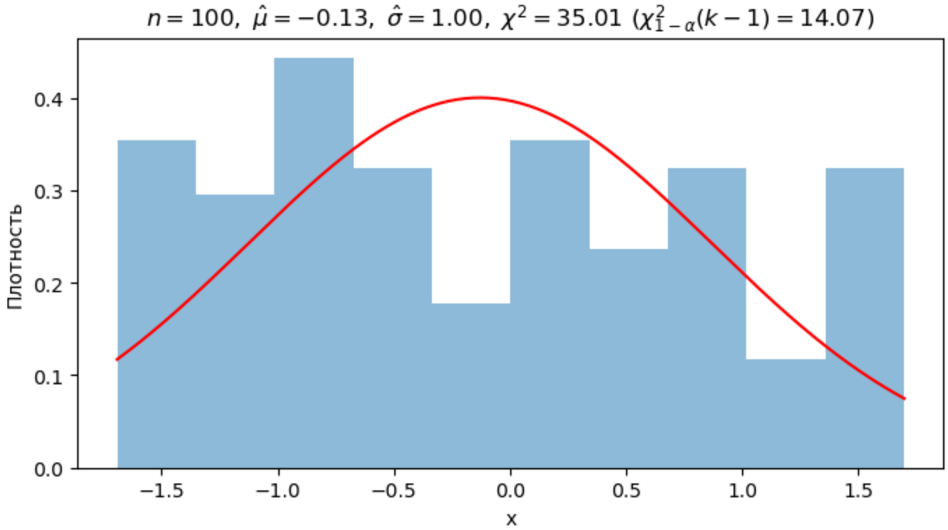

Рис. 4: Гистограмма и теоретическая плотность для равномерной выборки, $n = 100$.

4 Заключение

Проведённое исследование показало, что при большом объёме выборки ($n = 100$) критерий χ^2 надёжно подтверждает нормальность распределения, а оценки ММП дают $\hat{\mu} \approx 0$, $\hat{\sigma} \approx 1$. Для малых выборок ($n = 20$) статистическая нестабильность частот может приводить к ложным выводам. Проверка равномерно распределённых выборок ($n = 20, 100$) отвергает гипотезу нормальности.