

САНКТ-ПЕТЕРБУРГСКИЙ ПОЛИТЕХНИЧЕСКИЙ  
УНИВЕРСИТЕТ им. ПЕТРА ВЕЛИКОГО

ИНСТИТУТ ПРИКЛАДНОЙ МАТЕМАТИКИ И МЕХАНИКИ

КАФЕДРА ПРИКЛАДНОЙ МАТЕМАТИКИ

ЛАБОРАТОРНАЯ РАБОТА №7  
ПРОВЕРКА ГИПОТЕЗЫ О ЗАКОНЕ  
РАСПРЕДЕЛЕНИЯ ГЕНЕРАЛЬНОЙ  
СОВОКУПНОСТИ. МЕТОД ХИ-КВАДРАТ.

Студент группы 3630102/70301

Камянский Д.В.

Преподаватель

Баженов А. Н.

САНКТ-ПЕТЕРБУРГ  
2020 г.

# Содержание

<b>1. Список таблиц .....</b>	<b>3</b>
<b>2. Постановка задачи .....</b>	<b>4</b>
<b>3. Теория.....</b>	<b>4</b>
3.1. Метод максимального правдоподобия .....	4
3.2. Критерий согласия Пирсона .....	4
<b>4. Реализация.....</b>	<b>5</b>
<b>5. Результаты .....</b>	<b>5</b>
5.1. Метод максимального правдоподобия .....	5
5.2. Критерий Пирсона .....	5
<b>6. Выводы .....</b>	<b>5</b>
<b>7. Список литературы .....</b>	<b>6</b>
<b>8. Приложения .....</b>	<b>6</b>

# 1 Список таблиц

1	Таблица вычислений $\chi^2$ .....	5
---	-----------------------------------	---

## 2 Постановка задачи

Необходимо сгенерировать выборку объемом 100 элементов для нормального распределения  $N(x; 0, 1)$ . По сгенерированной выборке оценить параметры  $\mu$  и  $\sigma$  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  $\chi$ . В качестве уровня значимости взять  $\alpha = 0,05$ . Привести таблицу вычислений  $\chi^2$ .

## 3 Теория

### 3.1 Метод максимального правдоподобия

Метод максимального правдоподобия – метод оценивания неизвестного параметра путём максимизации функции правдоподобия.

$$\hat{\theta}_{\text{МП}} = \operatorname{argmax} \mathbf{L}(x_1, x_2, \dots, x_n, \theta) \quad (1)$$

Где  $\mathbf{L}$  это функция правдоподобия, которая представляет собой совместную плотность вероятности независимых случайных величин  $X_1, x_2, \dots, x_n$  и является функцией неизвестного параметра  $\theta$

$$\mathbf{L} = f(x_1, \theta) \cdot f(x_2, \theta) \cdot \dots \cdot f(x_n, \theta) \quad (2)$$

Оценкой максимального правдоподобия будем называть такое значение  $\hat{\theta}_{\text{МП}}$  из множества допустимых значений параметра  $\theta$ , для которого функция правдоподобия принимает максимальное значение при заданных  $x_1, x_2, \dots, x_n$ .

Тогда при оценивании математического ожидания  $m$  и дисперсии  $\sigma^2$  нормального распределения  $N(m, \sigma)$  получим:

$$\ln(\mathbf{L}) = -\frac{n}{2} \ln(2\pi) - \frac{n}{2} \ln(\sigma^2) - \frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - m)^2 \quad (3)$$

Отсюда находятся выражения для оценок  $m$  и  $\sigma^2$ :

$$\begin{cases} m = \bar{x} \\ \sigma^2 = s^2 \end{cases} \quad (4)$$

### 3.2 Критерий согласия Пирсона

Разобьём генеральную совокупность на  $k$  непересекающихся подмножеств  $\Delta_1, \Delta_2, \dots, \Delta_k$ ,  $\Delta_i = (a_i, a_{i+1}]$ ,  $p_i = P(X \in \Delta_i)$ ,  $i = 1, 2, \dots, k$  – вероятность того, что точка попала в  $i$ ый промежуток.

Так как генеральная совокупность это  $\mathbb{R}$ , то крайние промежутки будут бесконечными:  $\Delta_1 = (-\infty, a_1]$ ,  $\Delta_k = (a_k, \infty)$ ,  $p_i = F(a_i) - F(a_{i-1})$

$n_i$  – частота попадания выборочных элементов в  $\Delta_i$ ,  $i = 1, 2, \dots, k$ .

В случае справедливости гипотезы  $H_0$  относительно частоты  $\frac{n_i}{n}$  при больших  $n$  должны быть близки к  $p_i$ , значит в качестве меры имеет смысл взять:

$$Z = \sum_{i=1}^k \frac{n}{p_i} \left( \frac{n_i}{n} - p_i \right)^2 \quad (5)$$

Тогда

$$\chi_B^2 = \sum_{i=1}^k \frac{n}{p_i} \left( \frac{n_i}{n} - p_i \right)^2 = \sum_{i=1}^k \frac{(n_i - np_i)^2}{np_i} \quad (6)$$

Для выполнения гипотезы  $H_0$  должны выполняться следующие условия:

$$\chi_B^2 < \chi_{1-\alpha}^2(k-1) \quad (7)$$

где  $\chi_{1-\alpha}^2(k-1)$  – квантиль распределения  $\chi^2$  с  $k-1$  степенями свободы порядка  $1-\alpha$ , где  $\alpha$  заданный уровень значимости.

## 4 Реализация

Работы была выполнена на языке *Python3.8.2* Для генерации выборок использовался модуль `.random`. Для генерации выборок и обработки функции распределения использовалась библиотека *scipy.stats*.

## 5 Результаты

### 5.1 Метод максимального правдоподобия

При подсчете оценок параметров закона нормального распределения методом максимального правдоподобия были получены следующие значения:

$$\begin{aligned} \hat{m}_{\text{МП}} &= 0.090527 \\ \hat{\sigma}_{\text{МП}}^2 &= 0.963167 \end{aligned} \quad (8)$$

### 5.2 Критерий Пирсона

Таблица 1: Таблица вычислений  $\chi^2$

i	$\Delta_i$	$n_i$	$p_i$	$\frac{(n_i - np_i)^2}{np_i}$
1	$(-\infty, -1.0]$	15	0.1288	0.3501
2	$(-1.0, -0.5)$	10	0.1411	1.1988
3	$(-0.5, 0.0)$	24	0.1927	1.1634
4	$(0.0, 0.5)$	19	0.2021	0.0721
5	$(0.5, 1.0)$	13	0.1629	0.6626
6	$(1.0, \infty)$	19	0.1725	0.1771
$\Sigma$		100	1	3.6241

$$\chi_B^2 = 3.6241$$

## 6 Выводы

Табличное значение квантиля  $\chi_{1-\alpha}^2(k-1) = \chi_{0.95}^2(5) = 11.0705$ . Полученное значение критерия согласия Пирсона  $\chi_B^2 = 3.6241 < \chi_{0.95}^2(5)$ , следовательно основная гипотеза  $H_0$  на исходной выборке не может быть отвергнута на уровне значимости  $\alpha = 0.05$ .

## 7 Список литературы

- [1] Модуль numpy - <https://physics.susu.ru/vorontsov/language/numpy.html>
- [2] Модуль scipy - <https://docs.scipy.org/doc/scipy/reference/>
- [3] Таблица значений  $\chi^2$  - <http://statsoft.ru/home/textbook/modules/sttable.html#chi>

## 8 Приложения

Код лабораторной: <https://github.com/dkamianskii/MatStatLabs/tree/master/Lab7>