Санкт-Петербургский Политехнический
Университет им. Петра Великого

Институт прикладной математики и механики

Кафедра прикладной математики

Лабораторная работа №7

3 курс, группа 33631/2

Студент Д. А. Плаксин

Преподаватель Баженов А. Н.

Санкт-Петербург
2019 г.

Содержание

1 Список таблиц

2 Постановка задачи

Необходимо сгенерировать выборку объемом 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. В качестве ровня значимости взять $\alpha = 0,05$. Привести таблицу вычислений χ^2.

3 Теория

3.1 Метод максимального правдоподобия

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

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

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 θ

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

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

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

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

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 iый 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:

$$Z = \sum_{i=1}^{k}\frac{n}{p_i}\left(\frac{n_i}{n} - p_i\right)^2 \tag{4}$$

Тогда

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

Для выполнения гипотезы H_0 должны выполняться следующие условия [4]:

$$\chi_B^2 < \chi_{1-\alpha}^2(k-1) \tag{6}$$

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

4 Реализация

Работы была выполнена на языке $Python3.7$. Для генерации выборок использовался модуль [1]. Для построения графиков использовалась библиотека matplotlib [2]. Функции распределения обрабатывались при помощи библиотеки scipy.stats [3]

5 Результаты

5.1 Метод максимального правдоподобия

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

$$\hat{m}_{\text{МП}} = -0.1235$$
$$\hat{\sigma}^2_{\text{МП}} = 0.9877 \tag{7}$$

5.2 Критерий Пирсона

Таблица 1: Таблица вычислений χ^2

i	Δ_i	n_i	p_i	np_i	$n_i - np_i$	$\frac{(n_i-np_i)^2}{np_i}$
1	$(-\infty, -1.6605]$	5	0.0517	5.1670	-0.1670	0.0054
2	(-1.6605, -0.7749)	17	0.1950	19.4966	-2.4966	0.3197
3	(-0.7749, 0.1108)	40	0.3390	33.8962	6.1038	1.0991
4	(0.1108, 0.9965)	27	0.2778	27.7824	-0.7824	0.0220
5	$(0.9965, \infty)$	11	0.1284	12.8411	-1.8411	0.2640
$\sum$		100	1	100	0	1.8665

$$\chi^2_B = 1.8665$$

6 Выводы

В данной работе получено значение критерия согласия Пирсона $\chi^2_B = 1.8665$. Табличное значение квантиля $\chi^2_{1-\alpha}(k-1) = \chi^2_{0.95}(4) = 9,4877$ [5].

Значит $\chi^2_B < \chi^2_{0.95}(4)$, из этого следует, что основная гипотеза H_0 соотносится с выборкой на уровне $\alpha = 0.05$.

7 Список литературы

[1] Модуль numpy - https://physics.susu.ru/vorontsov/language/numpy.html

[2] Модуль matplotlib - https://matplotlib.org/users/index.html

[3] Модуль scipy - https://docs.scipy.org/doc/scipy/reference/

[4] https://en.wikipedia.org/wiki/Pearson%27s_chi-squared_test

[5] Таблица значений χ^2 - https://ru.wikipedia.org/wiki/%D0%9A%D0%B2%D0%B0%D0%BD%D1%82%D0%B8%D0%B8_%D1%80%D0%B0%D1%81%D0%BF%D1%80%D0%B5%D0%B4%D0%B5%D0%BB%D0%B5%D0%BD%D0%B8%D1%8F_%D1%85%D0%B8-%D0%BA%D0%B2%D0%B0%D0%B4%D1%80%D0%B0%D1%82%D0%B2%D0%B0%D0%B4%D1%80%D0%B0%D1%82

8 Приложения

Код отчёта: https://github.com/MisterProper9000/MatStatLabs/blob/master/MatStatLab5/MatStatLab7.tex

Код лаборатрной: https://github.com/MisterProper9000/MatStatLabs/blob/master/MatStatLab5/MatStatLab7.py