Санкт-Петербургский Политехнический Университет
Петра Великого

Институт прикладной математики и механики

Кафедра "Прикладная математика"

Отчёт
Лабораторная работа №7
по дисциплине
"Математическая статистика"

Выполнил студент:
Салихов С.Р.
группа: 3630102/70401

Проверил:
к.ф-м.н., доцент
Баженов Александр Николавич

Санкт-Петербург
2020 г.

Содержание

Список иллюстраций

Список таблиц

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.

2 Теория

2.1 Метод максимального правдоподобия

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

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

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

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

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 $x_1, x_2, \ldots, x_n$.

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

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

2.2 Хи-квадрат

Разобьём генеральную совокупность на k непересекающихся подмножеств $\Delta_1, \Delta_2, \ldots, \Delta_k,\ \Delta_i = (a_i, a_{i+1}],\ p_i = P(X \in \Delta_i),\ i = 1, 2, \ldots, k$ – 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 $\Delta_i,\ i = 1, 2, \ldots, k$.

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:

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

Тогда

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

Теорема К.Пирсона. Статистика критерия χ^2 асимптотически распределена по закону χ^2 с k - 1 степенями свободы.

3 Реализация

Для генерации выборки был использован *Python* 3.7 и модуль numpy. Для отрисовки графиков использовался модуль matplotlib. scipy.stats для обработки функций распределений.

4 Результаты

Метод максимального правдоподобия:

$$\hat{\mu} \approx 0.03, \hat{\sigma} \approx 0.95$$

Критерий согласия χ^2:

Таблица 1: Вычисление χ^2_B при проверке гипотизы H_0 о нормальном законе распределения $\mathrm{N}(x, \hat{\mu} \approx 0.16, \hat{\sigma} \approx 1.04)$

i	$Delta_i$, a_{i-1}, a_i	n_i	p_i	np_i	$n_i - np_i$	$\frac{(n_i-np_i)^2}{np_i}$
1	[-∞, -1.01]	14	0.1562	15.62	-1.62	0.17
2	[-1.01, -0.37]	19	0.2004	20.04	-1.04	0.05
3	[-0.37, 0.28]	23	0.2517	25.17	-2.17	0.19
4	[0.28, 0.92]	19	0.2122	21.22	-2.22	0.23
5	[0.92, 1.56]	14	0.1201	12.01	1.99	0.33
6	[1.56, ∞]	11	0.0594	5.94	5.06	4.32
$\sum$	-	100	1.0000	100.00	0.00	5.29 = χ^2_B

Количество промежутков k = 6.
Уровень значимости $\alpha = 0.05$.

5 Обсуждение

По результатам работы, значение критерия согласия Пирсона:
$\chi^2_B = 5.29$. Табличное значение квартиля $\chi^2_{1-\alpha}(k-1) = \chi^2_{0.95}(5) = 11.07$.

Таким образом, $\chi^2_B < \chi^2_{0.95}(5)$, из этого следует, что основная гипотеза H_0(о нормальном законе распределения $\mathrm{N}(x, \hat{\mu}, \hat{\sigma})$), на уровне зависимости $\alpha = 0.05$, соотносится с выборкой.

При распределении Лапласа(n = 30),

$$\hat{\mu} \approx 0.53, \hat{\sigma} \approx 1.52$$

По результатам работы, значение критерия согласия Пирсона:
$\chi^2_B = 12.94$.
Таким образом, $\chi^2_B > \chi^2_{0.95}(5)$, из этого следует, что основная гипотеза H_0(о нормальном законе распределения $\mathrm{N}(x, \hat{\mu}, \hat{\sigma})$), на уровне зависимости $\alpha = 0.05$, при распределении Лапласа не верна.

6 Литература

Модуль numpy

Модуль matplotlib

Модуль scipy

Таблица значений χ^2

7 Приложения

Код лабораторной
Код отчёта