*Федеральное государственное бюджетное образовательное учреждение* *высшего образования*

|  |  |
| --- | --- |
| **Изображение выглядит как эмблема, герб, нашивка, символ  Автоматически созданное описание** | ***«Московский государственный технический университет  имени Н.Э. Баумана»***  ***(национальный исследовательский университет)***  ***(МГТУ им. Н.Э. Баумана)*** |

**ФАКУЛЬТЕТ** ФУНДАМЕНТАЛЬНЫЕ НАУКИ\_\_\_\_\_\_\_\_

**КАФЕДРА** ВЫЧИСЛИТЕЛЬНАЯ МАТЕМАТИКА И МАТЕМАТИЧЕСКАЯ ФИЗИКА (ФН11)\_

**НАПРАВЛЕНИЕ ПОДГОТОВКИ** МАТЕМАТИКА И КОМПЬЮТЕРНЫЕ НАУКИ (02.03.01)

**Отчет**

**по лабораторной работе № \_8\_**

**Название лабораторной работы: Применение критерия χ^2 Пирсона к проверке гипотезы о виде функции распределения.**

**Вариант № 2**

**Дисциплина:** Теория вероятности и математическая статистика

Студент группы ФН11-52Б **\_\_\_\_\_\_\_\_\_\_\_\_\_ Кожемякин Г.А.**

(Подпись, дата) (И.О. Фамилия)

Преподаватель  **\_\_\_\_\_\_\_\_\_\_\_\_\_ Облакова Т.В.**

(Подпись, дата) (И.О. Фамилия)

Москва, 2024

**Задание**

1. Используя группированную выборку из задачи №1, проверьте на уровне гипотезу : выборка взята из генеральной совокупности, распределенной по закону
2. Неизвестные параметры распределения , если это необходимо, найдите методом моментов или методом максимального правдоподобия по выборке.
3. Постройте совмещенные графики гистограммы относительных частот и плотности, соответствующей функции распределения
4. Дайте анализ полученного решения.

**Исходные данные**

|  |  |  |
| --- | --- | --- |
| Номер варианта | Закон распределения | Уровень значимости |
| 2 |  | 0,05 |

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 8.432 | 6.077 | 8.982 | 2.897 | 5.784 | 29.159 | 14.564 | 9.869 | 11.621 | 10.405 |
| 11.589 | 15 | 8.659 | 15.142 | 5.345 | 7.69 | 16.504 | 8.742 | 2.834 | 9.846 |
| 13.697 | 13.167 | 6.846 | 6.727 | 12.614 | 3.655 | 19.627 | 1.572 | 11.4 | 8.384 |
| 4.807 | 7.897 | 13.881 | 6.244 | 13.188 | 12.419 | 8.279 | 9.37 | 7.233 | 15.503 |
| 13.47 | 8.511 | 10.134 | 18.831 | 9.422 | 3.715 | 8.63 | 6.99 | 10.187 | 17.491 |
| 8.262 | 11.019 | 13.02 | 7.619 | 5.1 | 7.757 | 7.772 | 16.063 | 8.736 | 12.214 |
| 7.998 | 5.651 | 5.225 | 8.525 | 12.935 | 9.312 | 5.12 | 6.146 | 19.032 | 15.049 |
| 9.643 | 8.633 | 4.367 | 3.91 | 5.957 | 6.345 | 14.577 | 9.878 | 13.658 | 4.104 |
| 5.752 | 5.693 | 15.207 | 6.776 | 5.055 | 4.96 | 35.182 | 14.399 | 13.383 | 14.144 |
| 3.77 | 6.394 | 11.695 | 8.474 | 13.196 | 7.636 | 11.219 | 10.463 | 6.802 | 10.135 |

**Решение задачи**

1. Группировка выборки

Группировку выборки осуществляемые аналогично задаче из лабораторной работы №1.

Находим крайние члены вариационного ряда и размах выборки.

Максимум: 35,182

Минимум: 1,572

Размах выборки: w = Максимум – Минимум = 33,61

Находим число интервалов группировки по правилу Стерджиса:

Ширина интервала: h = Размах выборки / число интервалов = w / l 4,8

Таким образом, была проведена группировка статистических данных. Результатом группировки является интервальный вариационный ряд.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Интервал | Середина интервала | Частота | Относительная частота pi | pi/h |
| [1,57;6,37) | 3,97 | 25 | 0,25 | 0.0521 |
| [6,37;11,17) | 8,77 | 40 | 0,4 | 0.0833 |
| [11,17;15,98) | 13,58 | 27 | 0,27 | 0.0563 |
| [15,98;20,78) | 18,38 | 6 | 0,06 | 0.0125 |
| [20,78;25,58) | 23,18 | 0 | 0 | 0 |
| [25,58;30,38) | 27,98 | 1 | 0,01 | 0.0021 |
| [30,38;35,18) | 32,78 | 1 | 0,01 | 0.0021 |

2. Оценка параметров

Для распределения Хи-квадрат с параметром функция правдоподобия имеет вид:

Найти оценку методом максимального правдоподобия не удается вследствие сложного вида уравнения (требуется найти производную от гамма-функции).

Воспользуемся методом моментов для нахождения оценки параметра .

Математическое ожидание равно:

Таким образом, оценкой неизвестного параметра является выборочное среднее.

3. Вычисление статистики и квантиля

Теоретические вероятности равны:

,

где – функция распределения случайной величины Х, распределенной по закону Хи-квадрат.

Таблица теоретических частот и вероятностей:

|  |  |  |  |
| --- | --- | --- | --- |
| Интервал | Частота | pti | npti |
| [1,57;6,37) | 25 | 0.217567 | 21.757 |
| [6,37;11,17) | 40 | 0.439158 | 43.916 |
| [11,17;15,98) | 27 | 0.242844 | 24.284 |
| [15,98;20,78) | 6 | 0.076736 | 7.674 |
| [20,78;25,58) | 0 | 0.018082 | 1.808 |
| [25,58;30,38) | 1 | 0.003546 | 0.355 |
| [30,38;35,18) | 1 | 0.000614 | 0.061 |

Объединим три последних строки с четвертой, так как для них npti<5.

|  |  |  |  |
| --- | --- | --- | --- |
| Интервал | Частота | pti | npti |
| [1,57;6,37) | 25 | 0.217567 | 21.757 |
| [6,37;11,17) | 40 | 0.439158 | 43.916 |
| [11,17;15,98) | 27 | 0.242844 | 24.284 |
| [15,98;35,18) | 8 | 0.098978 | 9.898 |

Вычислим статистику по формуле:

Далее, найдем квантиль . Так как для распределения была произведена оценка одного параметра, число степеней свободы будет равно l – 1 – 3 – 1 = 2.

Получаем, что так как , то основная гипотеза Н0 принимается.

Изображение выглядит как График, диаграмма, линия, текст

Автоматически созданное описание

**Выводы**

В ходе проделанной лабораторной работы осуществлена группировка заданной выборки из генеральной совокупности, была осуществлена проверка на уровне гипотезы : выборка взята из генеральной совокупности, распределенной по закону , а также был найден неизвестный параметр данного распределения при помощи метода моментов. Полученные результаты проиллюстрированы графиком с совмещенной гистограммой относительных частот и плотностью распределения

**Приложение**

В приложении представлен программный код на языке Python.

**import** scipy.stats **as** stats  
**import** numpy **as** np  
**import** matplotlib.pyplot **as** plt  
**import** scipy.stats **as** stats  
k = 9.97   
intervals = [  
 (1.57, 6.37),  
 (6.37, 11.17),  
 (11.17, 15.98),  
 (15.98, 20.78),  
 (20.78, 25.58),  
 (25.58, 30.38),  
 (30.38, 35.18)  
]  
  
**def** calculate\_theoretical\_frequencies(k, intervals, n=100):  
 frequencies = []  
 **for** a, b **in** intervals:  
 P = stats.chi2.cdf(b, k) - stats.chi2.cdf(a, k)  
 frequency = P \* n  
 frequencies.append(frequency)  
 **return** frequencies  
  
n = 100  
theoretical\_frequencies = calculate\_theoretical\_frequencies(k, intervals, n)  
**for** i, (interval, freq) **in** enumerate(zip(intervals, theoretical\_frequencies)):  
 print(**f"Интервал {**interval**}: Теоретическая частота = {**freq**:.6f}"**)  
  
intervals = [(1.57, 6.37), (6.37, 11.17), (11.17, 15.98), (15.98, 20.78),   
 (20.78, 25.58), (25.58, 30.38), (30.38, 35.18)]  
frequencies = [0.0521,0.0833,0.0563,0.0125,0,0.0021,0.0021]  
left\_edges = [interval[0] **for** interval **in** intervals]  
right\_edges = [interval[1] **for** interval **in** intervals]  
k = 9.97  
x = np.linspace(0, 35, 1000)   
y = stats.chi2.pdf(x, k)   
plt.figure(figsize=(10, 6))  
plt.bar(left\_edges, frequencies, width=[right\_edges[i] - left\_edges[i] **for** i **in** range(len(intervals))],   
 align=**'edge'**, alpha=0.6, edgecolor=**'black'**)  
plt.plot(x, y, color=**'red'**)  
plt.title(**'Совмещённая гистограмма относительных частот и плотность распределения χ²'**)  
plt.grid(**True**)  
plt.show()