**МИНОБРНАУКИ РОССИИ**

**Санкт-Петербургский государственный**

**электротехнический университет**

**«ЛЭТИ» им. В.И. Ульянова (Ленина)**

**Кафедра математического обеспечения и применения ЭВМ**

отчет

**по лабораторной работе №3**

**по дисциплине «Статистические методы обработки**

**экспериментальных данных»**

Тема: Обработка выборочных данных.

Нахождение оценок параметров распределения.

Проверка статистической гипотезы о нормальном законе распределения.

|  |  |  |
| --- | --- | --- |
| Студент гр. 6381 |  | Фиалковский М.С. |
| Преподаватель |  | Середа В.И. |

Санкт-Петербург

2020

**Цель работы.**

Получение практических навыков вычисления интервальных статистических оценок параметров распределения выборочных данных и проверки «справедливости» статистических гипотез.

**Постановка задачи.**

Для заданной надежности определить (на основании выборочных данных и результатов выполнения лабораторной работы №2) границы доверительных интервалов для математического ожидания и среднеквадратического отклонения случайной величины. Проверить гипотезу о нормальном распределении исследуемой случайной величины с помощью критерия Пирсона . Дать содержательную интерпретацию полученным результатам.

**Ход работы.**

Найдем границы доверительного интервала для математического ожидания случайной величины, предположительно имеющей нормальное распределение, в случае неизвестной дисперсии.

Воспользуемся формулой подставив туда значения из прошлой работы:

Используя таблицу и параметры и получим   
 Получим:

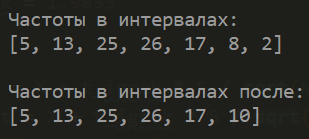
Далее будем находить границы доверительного интервала для СКО. Для этого по таблице найдем для и . Получим , проинтерполировав значение для 90 и 100. Будем использовать формулу

и получим*:*

Теперь проверим гипотезу о согласии с нормальным распределением с параметрами (459.91, 59.42). Для этого воспользуемся критерием для проверки сложной гипотезы. При этом проверяемые гипотезы будут формулироваться следующим образом:

выборка взята из нормального распределения с параметрами и .

:выборка взята не из нормального распределения с параметрами и .

Для использования критерия для проверки гипотезы сформируем интервалы таким образом, чтобы ожидаемые частоты в них были не менее Объединим последние 2 интервала. Получим:

На основании данных, полученных при расчёте значений критерия, построим таблицу для 6 интервалов:

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |
|  |  |  | 1.29 |  |  |  |  |
| 536.71 | 623.00 | 1.29 |  | 10 | 0.10 | 9.41 | 0.0364 |

Статистика критерия:

Критическое значение статистики критерия на уровне значимости  
 и при степенях свободы равно .

Получаем, что . Поэтому принимаем гипотезу .

**Выводы.**

В ходе выполнения лабораторной работы были изучены теоретические и получены практические навыки нахождения интервальных статистических оценок для параметров нормального распределения – математического ожидания и среднеквадратического отклонения.

Получили интервальную оценку математического ожидания генеральной совокупности с 95% доверительной вероятностью:

Получили интервальную оценку среднеквадратического отклонения генеральной совокупности с 95% доверительной вероятностью:  
.

Проверили гипотезу согласия с нормальным распределением с использованием критерия на уровне значимости . На основании произведённых расчётов и полученного и критического значений статистического критерия это гипотеза не была опровергнута, что может говорить о том, что выборка действительно согласована с нормальным законом распределения.

**ПРИЛОЖЕНИЕ А**

**ИСХОДНЫЙ КОД**

#!/usr/bin/env python

import lab1

import lab2

from scipy.stats import t, chi2

from scipy.special import erf

Ф = lambda x: erf(x/2\*\*0.5)/2 # функция лапласа

from math import sqrt

def print\_2digits(list, string):

print(string)

for elem in list:

print("{0:.2f}".format(elem), end=" ")

print(end="\n")

N = lab1.selection\_size # 96

n = N-1 # 95

general\_population = lab1.read\_data(filename=lab1.data\_file\_name)

sample = lab1.get\_sample(general\_population, lab1.selection\_size)

sample\_density = [pair.density for pair in sample]

table = lab2.build\_table(sample\_density)

X, S = lab2.get\_main\_values\_from\_table(table)

print("С прошлой работы получим значения:\nX = {0:.2f}\nS = {1:.2f}".format(X, S))

igrek = 0.95

t\_igrek = 1.9855

a\_left = X - t\_igrek \* S / sqrt(N)

a\_right = X + t\_igrek \* S / sqrt(N)

print("Доверительный интервал для математического ожидания: ", end="")

print("({0:.4f}; {1:.4f})\n".format(a\_left, a\_right))

gamma = 0.95

q = 0.147

sd\_left = S\*(1-q)

sd\_right = S\*(1+q)

print("Доверительный интервал для среднеквадратического отклонения: ", end="")

print("({0:.4f}; {1:.4f})\n".format(sd\_left, sd\_right))

borders, buckets = lab1.get\_interval\_sample(sample\_density)

# print(borders, buckets)

vals = [len(bucket) for bucket in buckets]

# print(borders)

# print(buckets)

print("Частоты в интервалах:")

print(vals)

print()

# Объединяем последние интервалы

borders[-2] = (borders[-2][0], borders[-1][1])

del borders[-1]

buckets[-2] += buckets[-1]

del buckets[-1]

vals = [len(bucket) for bucket in buckets]

# print(borders)

# print(buckets)

print("Частоты в интервалах после:")

print(vals)

print()

# границы интервалов в интервальном ряду

bins = [elem[0] for elem in borders] + [borders[-1][1]]

print\_2digits(bins, "Границы интервалов:")

# # стандартизуем границы интервалов

z = [(e - X)/S for e in bins]

z[0] = -float('inf')

z[-1] = float('inf')

print\_2digits(z, "Значения Z:")

p = []

for i in range(1, len(z)):

p.append(Ф(z[i])-Ф(z[i-1]))

print\_2digits(p, "Вероятности попадания в интервалы:")

print(end="\n")

# freq\_theor = list(map(lambda x: x\*n, p))

freq\_theor = [x\*N for x in p]

print\_2digits(freq\_theor, "Теоретические частоты по интервалам:")

print(end="\n")

freq\_expect = vals

print\_2digits(freq\_expect, "Ожидаемые частоты по интервалам:")

print(end="\n")

# посчитаем статистику критерия Хи-квадрат

stats\_crit = []

for i in range(len(freq\_expect)):

stats\_crit.append((freq\_expect[i] - freq\_theor[i])\*\*2/freq\_theor[i])

chi2\_calculated = sum(stats\_crit)

print("Cтатистики критерия Хи-квадрат:")

for elem in stats\_crit:

print("{0:.4f}".format(elem), end=" ")

print(end="\n")

print('Общее значение: {}'.format(chi2\_calculated))

print(end="\n")

k = 5 # число интервалов

l = 2 # число параметров

alpha = 0.05 # уровень значимости

# критическое значение статистики критерия

chi2\_criterion = chi2.ppf(1-alpha, df=k-l-1)

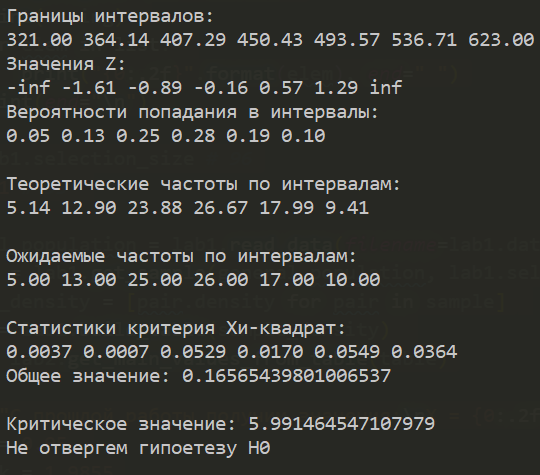
print('Критическое значение: {}'.format(chi2\_criterion))

# принимаем решение

if chi2\_calculated > chi2\_criterion:

print('Отвергаем гипотезу H0')

else:

 print('Не отвергем гипоетезу H0')