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 интервальных оценок параметров распределения. Проверка статистической гипотезы о нормальном законе распределения.

Студентка гр. 8382	___________________	Звегинцева Е.Н.
Студент гр. 8382	___________________	Мирончик П.Д.
Преподаватель	___________________	Середа А.-В.И.

Санкт-Петербург
2022

Цель работы.

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

Основные теоретические положения.

Доверительным называют интервал, который с заданной надежностью γ покрывает заданный параметр.

Интервальной оценкой математического ожидания по выборочной средней $\overline{x_B}$ при неизвестном среднем квадратическом отклонении σ генеральной совокупности служит доверительный интервал:

$$\mu \in (\overline{x_B} - \frac{S}{\sqrt{n}} t_\gamma, \overline{x_B} + \frac{S}{\sqrt{n}} t_\gamma),$$

где

$\overline{x_B}$ – статистическая оценка математического ожидания;

S – исправленная выборочная дисперсия;

n – объём выборки;

t_γ – из таблицы.

Интервальной оценкой среднеквадратического отклонения σ по исправленной выборочной дисперсии служит доверительный интервал:

$$\sigma \in (S(1-q), S(1+q)),$$

где

S – исправленная выборочная дисперсия;

q – из таблицы.

Критерий Пирсона, или критерий χ^2(Хи-квадрат), применяют для проверки гипотезы о соответствии эмпирического распределения предполагаемому теоретическому распределению $F(x)$.

Метод позволяет оценить статистическую значимость различий двух или нескольких относительных показателей (частот, долей).

Теоретические частоты вычисляются по формуле:

$$n_i^{'} = p_i * N,$$

где $p_i = \int f(x)dx$

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-a)^2}{2\sigma^2}\right).$$

Следует привести теоретические частоты к функции Лапласа.

Если $z = \frac{x-a}{\sigma}$, то $f(x)$ примет следующий вид:

$$f(z) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{z^2}{2}\right)$$

Для данной задачи $z_i = \frac{x_i - x_в}{S}$. Преобразуя формулу $p(i)$, получим:

$$p_i = \Phi(z_{i+1}) - \Phi(z_i),$$

где $\phi(z_i) = \frac{1}{\sqrt{2\pi}} \int_0^{z_i} \exp\left(-\frac{z^2}{2}\right) \mathrm{dx}$ – функция ошибок.

Если $\chi^2_{obs} \leq \chi^2_{crit}$ - гипотеза принимается, иначе ($\chi^2_{obs} > \chi^2_{crit}$) – гипотезу отвергают.

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 χ^2. Дать содержательную интерпретацию полученным результатам.

Выполнение работы.

При выполнении предыдущих лабораторных работ были получены выборочные данные(см. рис.1), такие как статистические оценки математического ожидания, дисперсии, СКО и исправленная выборочная дисперсия. Для их получения использовался сгенерированный интервальны ряд из первой лабораторной работы.

```
    Интервал  Середина  Частоты
0  [321,350)     335.5        2
1  [350,379)     364.5       10
2  [379,408)     393.5       10
3  [408,437)     422.5       13
4  [437,466)     451.5       36
5  [466,495)     480.5       19
6  [495,524)     509.5       13
7  [524,553)     538.5       11
N: 114
Матожидание: 453.280701754386
Дисперсия: 2541.9606802092953
СКО: 50.417860726227715
Исправленная дисперсия: 2564.4559074677845
Исправленное СКО: 50.640457220169175
```

Рисунок 1 – данные из первых лабораторных работ

Определим доверительный интервал для мат. ожидания по формуле:

$$\mu \in (\overline{x_B} - \frac{S}{\sqrt{n}} t_\gamma, \overline{x_B} + \frac{S}{\sqrt{n}} t_\gamma), \text{ где}$$

$\overline{x_B}$ – статистическая оценка математического ожидания, S – исправленная выборочная дисперсия, n – объём выборки;

Для доверительной точности 0.95 и объема выборки n=114 было выбрано значение t_γ =1.980 по приложению 6.

При γ=0.95: $\mu \in$ (443.8897332769508, 462.67167023182117)

Для доверительной точности 0.99 и объема выборки n=114 было выбрано значение t=2.617 по приложению 6.

При γ=0.99: $\mu \in$ (440.8684974587052, 465.69290605006677)

Определим доверительный интервал для среднеквадратического отклонения по формуле:

$$\sigma \in (S(1-q), S(1+q)), \text{ где}$$

S – исправленная выборочная дисперсия

Для доверительной точности 0.95 и объема выборки n=114 было выбрано значение q=0.143 по приложению 7

При γ=0.95: $\sigma \in$ (43.39887183768498, 57.88204260265337)

Для доверительной точности 0.99 и объема выборки n=114 было выбрано значение q=0.198 по приложению 7

При γ=0.99: $\sigma \in$ (40.61364669057568, 60.66726774976267)

По результатам видно, что большая доверительная вероятность обеспечивается за счет увеличения доверительного интервала, что делает оценку менее точной в обоих случаях (при вычислении доверительного интервала и для мат.ожидания, и для СКО).

Проверим гипотезу о нормальном распределении исследуемой случайной величины с помощью критерия Пирсона χ^2.

Формулируем нулевую гипотезу: величина ν имеет нормальное распределение с математическим ожиданием $\bar{x}$ и среднеквадратическим отклонением S;

Вычислим теоретические вероятности и частоты попадания в каждый интервал. Результаты представлены в табл.1.

Таблица 1

x_i	x_{i+1}	n_i	z_i	z_{i+1}	$F(z_i)$	$F(z_{i+1})$	p_i	n_i'
321	350	2	-2,61215	-2,03949	-0,4955	-0,4793	0,016202	1,847017
350	379	10	-2,03949	-1,46683	-0,4793	-0,42879	0,050511	5,758279
379	408	10	-1,46683	-0,89416	-0,42879	-0,31438	0,114406	13,04231
408	437	13	-0,89416	-0,3215	-0,31438	-0,12608	0,188299	21,46612
437	466	36	-0,3215	0,251169	-0,12608	0,099158	0,225241	25,67746
466	495	19	0,251169	0,823833	0,099158	0,294983	0,195825	22,32402
495	524	13	0,823833	1,396498	0,294983	0,418718	0,123735	14,10577
524	553	11	1,396498	1,969163	0,418718	0,475533	0,056815	6,476919

С использованием полученных частот вычислим χ^2_{obs} по формуле:

$$\chi^2_{obs} = \sum_{i=1}^{8} \frac{(n_i - n_i^{'})^2}{n_i^{'}}$$

Вычисленные результаты представлены в табл.2

Таблица 2

$\boldsymbol{n_i}$	$\boldsymbol{n_i}'$	$(\boldsymbol{n_i} - \boldsymbol{n_i'})^2/\boldsymbol{n_i'}$
2	1,847017	0,012671199
10	5,758279	3,124578747
10	13,04231	0,709661414
13	21,46612	3,338991969
36	25,67746	4,149742667
19	22,32402	0,49494192
13	14,10577	0,08668296
11	6,476919	3,158640428
		15,0759113

Сравним полученные значения с табличным значением $\chi^2_{\text{крит}}$.

$$\chi^2_{obs} = 15{,}0759113$$

Число степеней свободы: $k = 8 - 3 = 5$

$$\chi^2_{crit} = 11{,}07$$

Нулевая гипотеза принимается в случае, когда $\chi^2_{obs} \le \chi^2_{crit}$

В нашем случае мы отвергаем нулевую гипотезу, так как $15{,}07 > 11{,}07$, т.е. $\chi^2_{obs} > \chi^2_{crit}$. Следовательно величина ν не имеет нормальное распределение с математическим ожиданием $\bar{x}$ и среднеквадратическим отклонением S.

Выводы.

В ходе выполнения лабораторной работы были получены границы доверительных интервалов для математического ожидания и СКО случайной величины с доверительными вероятностями 0,95 и 0,99. Из полученных результатов можно сделать вывод, что размер доверительного интервала увеличивается при увеличении доверительной вероятности.

Также была проверена гипотеза о нормальном распределении исследуемой случайной величины с помощью критерия Пирсона χ^2. Был сделан вывод, что нулевая гипотеза отвергается, т.к. $\chi^2_{obs} > \chi^2_{crit}$, следовательно, исследуемая случайная величина не принадлежит нормальному закону распределения.