

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 нормаль-
ном законе распределения.

Студентка гр. 7381

Алясова А.Н.

Студент гр. 7381

Кортев Ю.В.

Преподаватель

Середа А.-В.И.

Санкт-Петербург

2021

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 \bar{x}_B при неизвестном среднем квадратическом отклонении σ генеральной совокупности служит доверительный интервал:

$$\bar{x}_B - \frac{s}{\sqrt{n}} t_\gamma \leq a \leq \bar{x}_B + \frac{s}{\sqrt{n}} t_\gamma,$$

где

\bar{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:

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

где

S – исправленная выборочная дисперсия;

q – из таблицы.

$$0 \leq \sigma \leq S(1 + q), q > 1$$

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 $z_i = \frac{x_i - x_e}{s}$. Преобразуя формулу $p(i)$, получим:

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

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

Если $\chi^2_{\text{наб}} \leq \chi^2_{\text{крит}}$ – гипотеза принимается, иначе ($\chi^2_{\text{наб}} > \chi^2_{\text{крит}}$) – 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 лабораторной работы №2 были получены выборочные данные, представленные в табл. 1.

Количество интервалов определено по формуле Стерджесса:

$$k = 1 + 3.322 * \log(n) = 8,$$

где n – объем выборки.

Ширина интервала:

$$h \approx 34.$$

Размер выборки:

$$n = 117.$$

Таблица 1

Интервал	Абсолютная частота	Относительная частота
[320, 354)	9	0.07692
[354,388)	4	0.03419
[388,422)	27	0.23077
[422,456)	25	0.21368
[456,490)	24	0.20513
[490,524)	17	0.14530
[524,558)	7	0.05983
[558,592)	3	0.02564
[592,593]	1	0.00855

Статистические оценки математического ожидания:

$$\bar{x} = \frac{1}{N} \sum_{i=1}^k x_i n_i = 448.8803$$

Статистические оценки дисперсии:

$$D_B = \frac{1}{N} \sum_{i=1}^k (x_i - \bar{x})^2 n_i = 3380.2592$$

Среднеквадратическое отклонение:

$$\sigma = \sqrt{D_B} = \sqrt{3380.2592} = 58,1400$$

Исправленная выборочная дисперсия:

$$S^2 = \frac{N}{N-1} D_B = \frac{117}{116} * 3380.2592 = 3409,3994$$

$$S = \sqrt{S^2} = \sqrt{3409,3994} = 58,3901$$

Определим доверительный интервал для мат. ожидания по формуле:

$$\bar{x}_B - \frac{S}{\sqrt{n}} t_\gamma \leq a \leq \bar{x}_B + \frac{S}{\sqrt{n}} t_\gamma, \text{ где}$$

\bar{x}_B – статистическая оценка математического ожидания;

S – исправленная выборочная дисперсия;

n – объём выборки;

$t_\gamma = 1,984$ – из таблицы (при уровне значимости $\alpha = 0,05$ и $n = 117$).

$$\bar{x}_B - \frac{S}{\sqrt{n}} t_\gamma = 448.8803 - \frac{58,3901}{\sqrt{117}} * 1,984 = 438,1920$$

$$\bar{x}_B + \frac{S}{\sqrt{n}} t_\gamma = 448.8803 + \frac{58,3901}{\sqrt{117}} * 1,984 = 459,5687$$

Из полученных результатов можно сделать вывод, что статистическая оценка математического ожидания попадет в интервал (438,1920; 459,5687) с вероятностью 95%.

Определим доверительный интервал для среднеквадратического отклонения по формуле:

$$S(1 - q) \leq \sigma \leq S(1 + q)$$

S – исправленная выборочная дисперсия;

$q = 0,14$ – из таблицы (при уровне значимости $\alpha = 0,05$ и $n = 117$).

$$S(1 - q) = 58,3901 * 0.86 = 50.2155$$

$$S(1 + q) = 58,3901 * 1,14 = 66.5647$$

Из полученных результатов можно сделать вывод, что полученный интервал (50.2155, 66.5647) покрывает величину σ с вероятностью 95%.

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

Вычислим теоретические вероятности и частоты попадания в каждый интервал. Результаты представлены в табл. 2-3.

Таблица 2

x_i	$x_i - \bar{x}_B$	v_i	$\Phi(v_i)$
320	-128,88034	-2,2072	-0,4864
354	-94,8803	-1,6249	-0,4479
388	-60,8803	-1,0426	-0,3514
422	-26,8803	-0,4604	-0,1774
456	7,1197	0,1219	0,0485
490	41,1197	0,7042	0,2594
524	75,1197	1,2865	0,4009
558	109,1197	1,8688	0,4692
592	143,1197	2,4511	0,4929
626	177,1196	3,0334	0,4988

Таблица 3

\bar{v}_i	$f(\bar{v}_i)$	$p_i = f(v_i) * \frac{h}{S}$	n'_i	$p_i = \Phi(v_{i+1}) - \Phi(v_i)$	n'_i
-1.9161	0.0636	0.0371	4,3352	0.0384	4,4973
-1.3338	0.1639	0.0954	11,1668	0.0965	11,2867
-0.7515	0.3008	0.1752	20,4928	0.1741	20,3667
-0.1692	0.3933	0.2290	26,7928	0.2259	26,4296
0.4131	0.3663	0.2133	24,9565	0.2108	24,6670
0.9954	0.2431	0.1415	16,5613	0.1415	16,5573
1.5777	0.1149	0.0669	7,8299	0.0683	7,9919
2.1600	0.0387	0.0225	2,6373	0.0237	2,7734
2.7422	0.0093	0.0054	0,6329	0.0059	0,6917

Вычислим $\chi^2_{\text{наб}}$ с использованием полученных частот по формуле:

$$\chi^2_{\text{наб}} = \sum_{i=1}^7 (n_i - n'_i)^2 / n'_i$$

Результаты представлены в табл. 4-5.

1-й способ.

Таблица 4

n_i	n_i'	$n_i - n_i'$	$(n_i - n_i')^2$	$(n_i - n_i')^2 / n_i'$
9	4,3352	4,6648	21,7607	5,0196
4	11,1668	-7,1668	51,3632	4,5996
27	20,4928	6,5072	42,3442	2,0663
25	26,7928	-1,7928	3,2142	0,1200
24	24,9565	-0,9565	0,9149	0,0367
17	16,5613	0,4387	0,1924	0,0116
7	7,8299	-0,8299	0,6887	0,0880
3	2,6373	0,3627	0,1315	0,0499
1	0,6329	0,3671	0,1348	0,2130
Сумма				12.2046

2-й способ.

Таблица 5

n_i	n_i'	$n_i - n_i'$	$(n_i - n_i')^2$	$(n_i - n_i')^2 / n_i'$
9	4,4973	4,5027	20,2740	4,5080
4	11,2867	-7,2867	53,0964	4,7043
27	20,3667	6,6333	44,0010	2,1604
25	26,4296	-1,4296	2,0438	0,0773
24	24,6670	-0,6670	0,4449	0,0180
17	16,5573	0,4427	0,1960	0,0118
7	7,9919	-0,9919	0,9839	0,1231
3	2,7734	0,2266	0,0514	0,0185
1	0,6917	0,3083	0,0950	0,1374
Сумма				11.7590

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

$$\chi^2_{\text{крит}} = 11,070$$

$$\chi^2_{\text{наб}_1} = 12.2046 > \chi^2_{\text{крит}}$$

$$\chi^2_{\text{наб}_2} = 11.7590 > \chi^2_{\text{крит}}$$

Из полученных результатов можно сделать вывод, что данные отвергаются гипотезой χ^2 и не имеют нормального распределения, так как $\chi^2_{\text{наб}_1} > \chi^2_{\text{крит}}$ в обоих способах.

Выводы.

В ходе выполнения лабораторной работы были получены границы доверительных интервалов для математического ожидания и среднеквадратического отклонения случайной величины. Из полученных результатов можно сделать вывод, что интервал (438,1920; 459,5687) покрывает математическое ожидание и интервал (50.2155, 66.5647) покрывает величину σ с вероятностью 95%.

Также была проверена гипотеза о нормальном распределении исследуемой случайной величины с помощью критерия Пирсона χ^2 . Из полученного результата можно сделать вывод, что гипотеза отвергается, т.к. $\chi^2_{\text{наб}_1} > \chi^2_{\text{крит}}$, соответственно, исследуемая случайная величина не принадлежит нормальному закону распределения.