



МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РОССИЙСКОЙ ФЕДЕРАЦИИ
Федеральное государственное бюджетное образовательное учреждение
высшего образования
«МИРЭА – Российский технологический университет»

ИНСТИТУТ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА
КАФЕДРА ВЫСШЕЙ МАТЕМАТИКИ

Типовой расчет 2

по курсу «Теория вероятностей и математическая статистика, часть 2»

ВАРИАНТ 31

Тема: _____ **Проверка статистических гипотез** _____

Выполнил:
Студент 3-го курса
Бредихин В.А.
Группа: КМБО-02-21

МОСКВА – 2024

Задание

Задание 1. Проверка гипотезы о нормальном распределении с помощью критерия χ^2 .

В соответствии с номером варианта взять из файла **MC_D_Norm** выборку $\{y_1, \dots, y_N\}$ объемом $N=200$. Построить Таблицу 1.1 интервального ряда, положив $a_0 = \min y_i$, $a_m = \max y_i$, число интервалов находится по формуле Стерджеса $m = 1 + [\log_2 N]$, длины всех интервалов $(a_{k-1}, a_k]$, $k=1, \dots, m$, одинаковы.

Таблица 1.1. Интервальный ряд.

Интервалы	n_k	w_k
$[a_0, a_1]$	n_1	w_1
$(a_1, a_2]$	n_2	w_2
...
$(a_{m-1}, a_m]$	n_m	w_m
	$\sum_{k=1}^m n_k$	$\sum_{k=1}^m w_k$

Найти оценки математического ожидания \hat{a} , дисперсии $\hat{\sigma}^2$ и среднего квадратического отклонения $\hat{\sigma}$ по формулам

$$\hat{a} = \frac{1}{N} \sum_{i=1}^N y_i, \quad \hat{\sigma}^2 = \frac{1}{N} \sum_{i=1}^N (y_i)^2 - (\hat{a})^2, \quad \hat{\sigma} = \sqrt{\hat{\sigma}^2}.$$

Построить Таблицу 1.2.

Таблица 1.2. Вычисление p_k^* .

k	a_k	$\frac{a_k - \hat{a}}{\hat{\sigma}}$	$\frac{1}{\hat{\sigma}} \varphi_0\left(\frac{a_k - \hat{a}}{\hat{\sigma}}\right)$	$\Phi\left(\frac{a_k - \hat{a}}{\hat{\sigma}}\right)$	p_k^*
0	a_0	$\frac{a_0 - \hat{a}}{\hat{\sigma}}$	$\frac{1}{\hat{\sigma}} \varphi_0\left(\frac{a_0 - \hat{a}}{\hat{\sigma}}\right)$	$\Phi\left(\frac{a_0 - \hat{a}}{\hat{\sigma}}\right)$	—
1	a_1	$\frac{a_1 - \hat{a}}{\hat{\sigma}}$	$\frac{1}{\hat{\sigma}} \varphi_0\left(\frac{a_1 - \hat{a}}{\hat{\sigma}}\right)$	$\Phi\left(\frac{a_1 - \hat{a}}{\hat{\sigma}}\right)$	p_1^*
...
m	a_m	$\frac{a_m - \hat{a}}{\hat{\sigma}}$	$\frac{1}{\hat{\sigma}} \varphi_0\left(\frac{a_m - \hat{a}}{\hat{\sigma}}\right)$	$\Phi\left(\frac{a_m - \hat{a}}{\hat{\sigma}}\right)$	p_m^*
					$\sum_{k=1}^m p_k^*$

Формулы для вычисления $\varphi_0(x)$, $\Phi(x)$, p_k^* имеются в указаниях к **Заданию 1**.

Построить график плотности нормального распределения $N(\hat{a}, \hat{\sigma}^2)$, наложенный на гистограмму относительных частот.

Построить Таблицу 1.3.

Таблица 1.3. Вычисление выборочного значения критерия χ_B^2 .

k	Интервал	w_k	p_k^*	$ w_k - p_k^* $	$\frac{N(w_k - p_k^*)^2}{p_k^*}$
1	$[a_0, a_1]$	w_1	p_1^*	$ w_1 - p_1^* $	$\frac{N(w_1 - p_1^*)^2}{p_1^*}$
2	$(a_1, a_2]$	w_2	p_2^*	$ w_2 - p_2^* $	$\frac{N(w_2 - p_2^*)^2}{p_2^*}$
...
m	$(a_{m-1}, a_m]$	w_m	p_m^*	$ w_m - p_m^* $	$\frac{N(w_m - p_m^*)^2}{p_m^*}$
		$\sum_{k=1}^m w_k$	$\sum_{k=1}^m p_k^*$	$\max w_k - p_k^* $	$\sum_{k=1}^m \frac{N(w_k - p_k^*)^2}{p_k^*}$

Проверить с помощью критерия χ^2 гипотезу о соответствии выборки нормальному распределению $N(\hat{a}, \hat{\sigma}^2)$ при уровне значимости $\alpha = 0,05$.

Задание 2. Проверка гипотезы о равномерном распределении с помощью критерия χ^2 .

В соответствии с номером варианта взять из файла **MC_D_Unif** выборку $\{y_1, \dots, y_N\}$ и значения a и b . Построить Таблицу 2.1 интервального ряда, аналогичную Таблице 1.1., положив $a_0 = a$, $a_m = b$.

Построить Таблицу 2.2.

Таблица 2.2. Вычисление p_k^* .

k	a_k	$f(a_k)$	$F(a_k)$	p_k^*
0	a_0	$f(a_0)$	0	—
1	a_1	$f(a_1)$	$F(a_1)$	p_1^*
...
m	a_m	$f(a_m)$	$F(a_m)$	p_m^*
				$\sum_{k=1}^m p_k^*$

Формулы для вычисления значений $f(a_k)$, $F(a_k)$, p_k^* имеются в указаниях к

Заданию 2.

Построить Таблицу 2.3, аналогичную Таблице 1.3.

Построить график плотности равномерного распределения на отрезке $[a, b]$, наложенный на гистограмму относительных частот.

Проверить с помощью критерия χ^2 гипотезу о соответствии выборки равномерному распределению на отрезке $[a, b]$ при уровне значимости $\alpha = 0,05$.

Заданию 3.

Проверить гипотезу о соответствии выборки равномерному распределению на отрезке $[a, b]$ при уровне значимости $\alpha = 0,05$ с помощью критерия Колмогорова.

Результаты вычислений приводить в отчете с точностью до 0,00001.

Указания

В разделе отчета **Краткие теоретические сведения** следует привести:

- выражения для плотности и функции распределения, математического ожидания, дисперсии, среднего квадратичного отклонения для нормального распределения и равномерного распределения на отрезке $[a, b]$;
- общую схему проверки статистических гипотез с помощью критерия χ^2 ;
- таблицу критических значений $\chi^2_{кр,\alpha}(l)$ при $\alpha = 0,05$

l	4	5	6	7	8
$\chi^2_{кр,\alpha}(l)$	9,487729	11,070498	12,591587	14,067140	15,507313

- общую схему проверки статистических гипотез с помощью критерия Колмогорова;
- таблицу критических значений распределения Колмогорова

α	0,01	0,02	0,03	0,05	0,07	0,1
k_α	1,627624	1,517427	1,449086	1,358099	1,294675	1,223848

В этом разделе должны быть описаны функции языка программирования, которые использованы в программе расчета, в частности, связанные с вычислением значений плотностей, функций распределений.

Краткие теоретические сведения

Нормальное распределение:

Плотность распределения	$\frac{1}{\sigma} \varphi\left(\frac{x-a}{\sigma}\right) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-a)^2}{2\sigma^2}}$
Функция распределения	$\Phi\left(\frac{x-a}{\sigma}\right) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{(t-a)^2}{2\sigma^2}} dt$
Математическое ожидание	a
Дисперсия	σ^2
Среднеквадратическое отклонение	σ

Равномерное распределение на отрезке $[a, b]$:

Плотность распределения	$\frac{1}{b-a}, a \leq x \leq b$
Функция распределения	$\frac{x-a}{b-a}, a \leq x < b$
Математическое ожидание	$\frac{a+b}{2}$
Дисперсия	$\frac{(b-a)^2}{12}$
Среднеквадратическое отклонение	$\frac{b-a}{2\sqrt{3}}$

Общая схема проверки статистических гипотез с помощью критерия χ^2 :

При проверке гипотезы с помощью критерия χ^2 найденное значение критерия $\chi_B^2 = \sum_{k=1}^m \frac{N(w_k - p_k^*)^2}{p_k^*}$ сравнивается с критическим значением $\chi_{кр, \alpha}^2(l)$, где

α – уровень значимости, l – число степеней свободы.

Если $\chi_B^2 \leq \chi_{кр, \alpha}^2(l)$, то гипотеза о соответствии выборки заданному распределению не противоречит экспериментальным данным при уровне значимости α .

Если $\chi_B^2 > \chi_{кр, \alpha}^2(l)$, то гипотеза о соответствии выборки заданному распределению противоречит экспериментальным данным при уровне значимости α .

Таблица критических значений $\chi^2_{кр,\alpha}(l)$ при $\alpha = 0,05$:

l	4	5	6	7	8
$\chi^2_{кр,\alpha}(l)$	9,487729	11,070498	12,591587	14,067140	15,507313

Общая схема проверки статистических гипотез с помощью критерия

Колмогорова:

При проверке гипотезы с помощью критерия Колмогорова найденное значение критерия $D_N\sqrt{N}$ сравнивается с критическим значением k_α при уровне значимости α .

Если $D_N\sqrt{N} \leq k_\alpha$, то гипотеза о соответствии выборки заданному распределению не противоречит экспериментальным данным при уровне значимости α .

Если $D_N\sqrt{N} > k_\alpha$, то гипотеза о соответствии выборки заданному распределению противоречит экспериментальным данным при уровне значимости α .

Таблица критических значений распределения Колмогорова:

α	0,01	0,02	0,03	0,05	0,07	0,1
k_α	1,627624	1,517427	1,449086	1,358099	1,294675	1,223848

В программе расчёта был использован интерпретируемый язык программирования Python. Использовались следующие функции библиотеки SciPy для языка программирования Python:

Нахождение значения функции плотности нормального распределения в точке x .

```
import scipy.stats as sps
sps.norm(loc, scale).pdf(x)
```

Нахождение значения функции нормального распределения в точке x .

```
import scipy.stats as sps
sps.norm(loc, scale).cdf(x)
```

Параметры: loc – среднее значение случайной величины;
 $scale$ – стандартное отклонение случайной величины.

Нахождение значения функции плотности равномерного распределения на отрезке $[a, b]$ в точке x .

```
import scipy.stats as sps
sps.uniform(loc, scale).pdf(x)
```

Нахождение значения функции равномерного распределения на отрезке $[a, b]$ в точке x .

```
import scipy.stats as sps
sps.uniform(loc, scale).cdf(x)
```

Параметры: $loc = a$; $scale = b - a$

Результаты расчетов

Задание 1. Проверка гипотезы о нормальном распределении с помощью критерия χ^2_B

Исходные данные: вариант 31

Неупорядоченная выборка:

0,19246	2,50718	0,48842	-0,64335	1,13986	-0,83374	-0,00543	-1,01062	-0,2675	1,81427
0,80313	-2,10011	-1,69252	-1,17174	0,04062	1,21214	-0,15503	-0,01923	-0,39158	-0,26278
-3,48576	-0,24848	-1,37462	0,51209	0,45522	-0,75527	-1,11888	0,82193	0,3198	0,035
-1,44455	0,18723	-1,45271	-2,196	1,34416	-0,60553	0,19637	-1,65996	-1,55406	-0,03702
-1,00083	1,23301	-0,36834	-2,57496	-0,84309	0,12664	1,25903	-0,37929	-1,43049	0,557
-2,43754	-0,36249	-0,07659	-2,81298	-0,22463	-1,03946	-0,28941	-2,65289	-1,17698	-0,72634
1,03367	-0,29843	-1,07685	-0,57961	0,62144	-0,36482	-0,22634	-2,14024	-1,99505	-0,52249
-3,00016	0,05747	-0,98226	-0,94451	-0,45254	-0,18782	-0,01519	-2,03009	-1,85127	-0,76626
0,24126	-0,97998	-3,27611	0,68675	-1,15395	0,66579	0,5771	-0,42936	-2,08443	-1,12185
1,26235	-1,35962	0,22539	-1,84506	-1,27376	-0,23521	-0,7322	-1,80748	-0,2668	0,48225
-0,80451	-1,51108	-2,28142	-0,29567	-1,54421	-0,71845	-0,80734	-0,72826	-0,44158	-0,3058
-0,96092	-2,14852	-0,60312	-1,87268	-0,04441	-1,24396	-2,3728	-0,31385	-2,83995	-1,09129
-2,06365	-1,09816	-1,63113	0,35897	-1,84191	0,48095	-0,32613	0,30875	0,17399	-0,83632
0,53289	0,1505	-1,55066	-1,33178	0,42111	0,48713	-1,87991	-0,3564	-1,31472	-0,27495
-1,05892	-0,59223	-2,10569	-0,19574	-0,25139	-1,85589	-1,68775	1,01949	-0,90667	-0,74056
0,19837	-2,36987	-0,84297	-0,34504	-2,22153	-2,32131	-0,72785	-0,97034	-0,56135	-0,84604
-2,19243	0,29972	-2,77571	-0,59727	-0,61462	0,85916	-0,26241	-1,70304	-0,7322	-1,13333
-0,29219	-1,64364	-0,30763	-2,15429	-1,19716	-0,79949	-0,46311	0,84735	0,27475	-1,07222
-0,23144	-0,4156	-1,9949	-2,09081	-0,80145	-2,09474	-3,22127	0,00302	-0,72342	-0,79797
-0,91219	0,4666	-2,51264	-1,19945	0,48462	-0,36119	1,29751	0,9435	0,18866	0,264

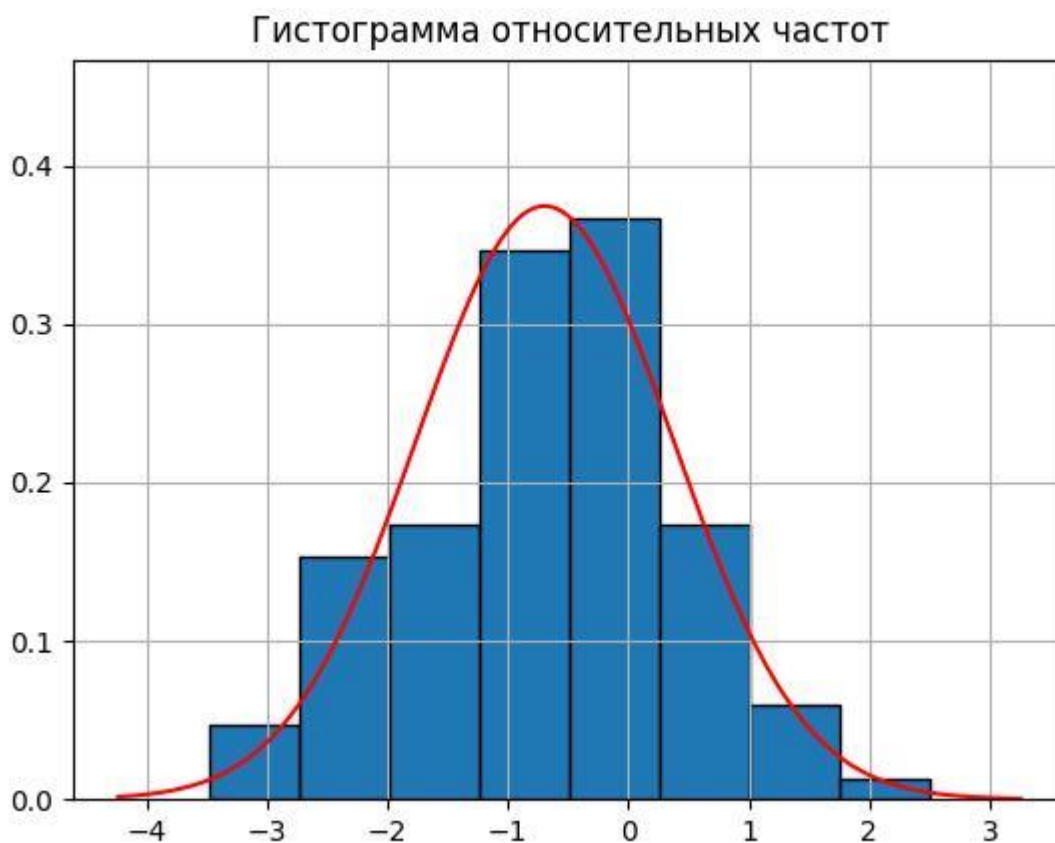
Упорядоченная выборка:

-3,48576	-3,27611	-3,22127	-3,00016	-2,83995	-2,81298	-2,77571	-2,65289	-2,57496	-2,51264
-2,43754	-2,3728	-2,36987	-2,32131	-2,28142	-2,22153	-2,196	-2,19243	-2,15429	-2,14852
-2,14024	-2,10569	-2,10011	-2,09474	-2,09081	-2,08443	-2,06365	-2,03009	-1,99505	-1,9949
-1,87991	-1,87268	-1,85589	-1,85127	-1,84506	-1,84191	-1,80748	-1,70304	-1,69252	-1,68775
-1,65996	-1,64364	-1,63113	-1,55406	-1,55066	-1,54421	-1,51108	-1,45271	-1,44455	-1,43049
-1,37462	-1,35962	-1,33178	-1,31472	-1,27376	-1,24396	-1,19945	-1,19716	-1,17698	-1,17174
-1,15395	-1,13333	-1,12185	-1,11888	-1,09816	-1,09129	-1,07685	-1,07222	-1,05892	-1,03946
-1,01062	-1,00083	-0,98226	-0,97998	-0,97034	-0,96092	-0,94451	-0,91219	-0,90667	-0,84604
-0,84309	-0,84297	-0,83632	-0,83374	-0,80734	-0,80451	-0,80145	-0,79949	-0,79797	-0,76626
-0,75527	-0,74056	-0,7322	-0,7322	-0,72826	-0,72785	-0,72634	-0,72342	-0,71845	-0,64335
-0,61462	-0,60553	-0,60312	-0,59727	-0,59223	-0,57961	-0,56135	-0,52249	-0,46311	-0,45254
-0,44158	-0,42936	-0,4156	-0,39158	-0,37929	-0,36834	-0,36482	-0,36249	-0,36119	-0,3564
-0,34504	-0,32613	-0,31385	-0,30763	-0,3058	-0,29843	-0,29567	-0,29219	-0,28941	-0,27495
-0,2675	-0,2668	-0,26278	-0,26241	-0,25139	-0,24848	-0,23521	-0,23144	-0,22634	-0,22463
-0,19574	-0,18782	-0,15503	-0,07659	-0,04441	-0,03702	-0,01923	-0,01519	-0,00543	0,00302
0,035	0,04062	0,05747	0,12664	0,1505	0,17399	0,18723	0,18866	0,19246	0,19637
0,19837	0,22539	0,24126	0,264	0,27475	0,29972	0,30875	0,3198	0,35897	0,42111
0,45522	0,4666	0,48095	0,48225	0,48462	0,48713	0,48842	0,51209	0,53289	0,557
0,5771	0,62144	0,66579	0,68675	0,80313	0,82193	0,84735	0,85916	0,9435	1,01949
1,03367	1,13986	1,21214	1,23301	1,25903	1,26235	1,29751	1,34416	1,81427	2,50718

Таблица 1.1. Интервальный ряд

Интервал	n_k	w_k
$[-3.48576, -2.73664]$	7	0,035
$(-2.73664, -1.98752]$	23	0,115
$(-1.98752, -1.23841]$	26	0,13
$(-1.23841, -0.48929]$	52	0,26
$(-0.48929, 0.25983]$	55	0,275
$(0.25983, 1.00894]$	26	0,13
$(1.00894, 1.75806]$	9	0,045
$(1.75806, 2.50718]$	2	0,01
	200	1

Рисунок 1. Гистограмма относительных частот для нормального распределения



Оценка среднего значения: -0.69528

Оценка дисперсии: 1.13493

Оценка среднего квадратического отклонения: 1.06533

Таблица 1.2. Вычисление p_k^*

k	a_k	$\frac{a_k - \hat{a}}{\hat{\sigma}}$	$\frac{1}{\hat{\sigma}} \varphi_0\left(\frac{a_k - \hat{a}}{\hat{\sigma}}\right)$	$\Phi\left(\frac{a_k - \hat{a}}{\hat{\sigma}}\right)$	p_k^*
0	-3,48576	-2,61935	0,01212	0,00440	0,00000
1	-2,73664	-1,91617	0,05972	0,02767	0,02767
2	-1,98753	-1,21300	0,17944	0,11257	0,08489
3	-1,23841	-0,50982	0,32884	0,30509	0,19252
4	-0,48929	0,19336	0,36754	0,57666	0,27157
5	0,25983	0,89654	0,25055	0,81502	0,23836
6	1,00895	1,59972	0,10417	0,94517	0,13015
7	1,75806	2,30290	0,02641	0,98936	0,04419
8	2,50718	3,00608	0,00408	0,99868	0,01064
					1

Таблица 1.3. Вычисление выборочного значения критерия χ_B^2

k	Интервал	w_k	p_k^*	$ w_k - p_k^* $	$\frac{N(w_k - p_k^*)^2}{p_k^*}$
1	[-3.48576, -2.73664]	0,03500	0,02767	0,00733	0,38818
2	(-2.73664, -1.98752]	0,11500	0,08489	0,03011	2,13524
3	(-1.98752, -1.23841]	0,13000	0,19252	0,06252	4,06108
4	(-1.23841, -0.48929]	0,26000	0,27157	0,01157	0,09863
5	(-0.48929, 0.25983]	0,27500	0,23836	0,03664	1,12672
6	(0.25983, 1.00894]	0,13000	0,13015	0,00015	0,00004
7	(1.00894, 1.75806]	0,04500	0,04419	0,00081	0,00298
8	(1.75806, 2.50718]	0,01000	0,01064	0,00064	0,00775
		1,00000	1,00000	0,06252	7,82062

Вывод: значение $\chi_B^2 = 7,82062$, а значение $\chi_{кр,\alpha}^2(5) = 11,0705$ при $\alpha = 0,05$.

Гипотеза о соответствии выборки нормальному распределению не противоречит экспериментальным данным (может быть принята) при уровне значимости $\alpha = 0.05$, т.к. $\chi_B^2 \leq \chi_{кр,\alpha}^2(5)$.

Задание 2. Проверка гипотезы о равномерном распределении с помощью критерия χ^2_B

Исходные данные: вариант 31, a = -1.25, b = 3.07

Неупорядоченная выборка:

2,223	0,68079	-0,41847	-0,38948	-0,0061	1,4978	1,73507	1,2225	2,62407	1,35805
-0,58221	0,3545	-0,68985	2,45699	-0,32404	-0,63709	-0,18827	-1,02058	1,39019	2,72189
0,02257	2,64161	0,21917	2,46652	2,98705	2,36196	1,14139	1,15361	-0,17097	1,69604
-1,07053	0,48882	2,97694	1,55299	-1,08696	2,57366	0,49315	2,71155	-0,42735	-0,53103
0,43484	0,16253	-0,49261	2,49925	1,82486	0,74941	2,72323	0,44696	-1,06707	0,65086
-0,84294	1,16817	2,55987	1,71082	-0,1978	2,32953	-0,59671	1,98862	0,21868	0,00287
2,04559	0,05322	-0,16837	1,25526	0,89619	0,05577	1,1118	3,03053	0,67121	0,41857
1,92665	-1,06789	-0,8294	2,98316	2,44883	0,06379	-0,18015	-0,99176	-1,21709	2,15711
1,76785	0,74922	0,87397	1,05939	-0,31902	1,34497	0,98422	2,18461	1,41099	1,12726
2,72915	1,25802	2,57945	-0,9389	3,02769	0,79229	0,64093	-0,65592	1,64365	2,21153
0,43095	2,16852	0,25979	0,26705	1,22651	0,32204	-0,92617	1,26485	2,76961	-0,20211
1,69153	1,80659	0,35496	-0,73917	1,61653	2,17387	-0,27543	2,61418	-0,66519	1,89947
0,41075	0,91459	2,16853	0,74066	2,11074	0,16096	2,12535	-0,12826	1,95187	-1,12592
2,64454	-0,23501	1,64481	-1,24532	1,9597	-0,14029	1,70201	-0,62801	-0,8905	1,70221
1,5634	-0,46227	2,93114	2,49316	-0,57771	2,1829	-0,7349	-0,96593	-1,04144	-0,98025
2,08066	0,12713	0,03995	1,84615	2,02167	2,37626	0,61515	1,62266	-1,01054	2,96446
-1,13966	0,68708	-1,16844	-0,43135	2,58039	1,48709	-0,33943	-1,14835	0,73995	2,51401
2,223	0,68079	-0,41847	-0,38948	-0,0061	1,4978	1,73507	1,2225	2,62407	1,35805
-0,58221	0,3545	-0,68985	2,45699	-0,32404	-0,63709	-0,18827	-1,02058	1,39019	2,72189
0,02257	2,64161	0,21917	2,46652	2,98705	2,36196	1,14139	1,15361	-0,17097	1,69604

Упорядоченная выборка:

-1,24532	-1,21709	-1,16844	-1,14835	-1,13966	-1,12592	-1,10772	-1,08696	-1,07053	-1,06789
-1,06707	-1,04144	-1,02058	-1,01054	-0,99176	-0,98418	-0,98025	-0,96593	-0,9389	-0,92617
-0,90174	-0,8905	-0,88027	-0,84294	-0,8294	-0,73917	-0,7349	-0,72435	-0,68985	-0,67581
-0,66519	-0,65967	-0,65592	-0,63709	-0,62801	-0,6116	-0,59671	-0,58221	-0,57771	-0,53103
-0,49261	-0,46227	-0,43135	-0,42735	-0,41847	-0,38948	-0,33943	-0,32404	-0,31902	-0,27543
-0,23501	-0,20211	-0,1978	-0,18827	-0,18015	-0,17097	-0,16837	-0,14029	-0,12826	-0,01001
-0,0061	0,00287	0,02257	0,03063	0,03995	0,04958	0,05322	0,05468	0,05577	0,06379
0,09791	0,12713	0,16096	0,16253	0,21868	0,21917	0,22074	0,25979	0,26521	0,26705
0,32204	0,3545	0,35496	0,41075	0,41857	0,43095	0,43484	0,44696	0,48882	0,49315
0,61515	0,64093	0,65086	0,67121	0,68079	0,68708	0,73224	0,73995	0,74066	0,74922
0,74941	0,79229	0,82353	0,87397	0,89619	0,91459	0,98422	1,05939	1,1118	1,12726
1,14139	1,15361	1,16191	1,16817	1,2225	1,22651	1,25526	1,25802	1,26485	1,27465
1,34497	1,35805	1,39019	1,41099	1,46366	1,48709	1,4978	1,55299	1,5634	1,61653
1,62266	1,6249	1,64365	1,64481	1,69153	1,69604	1,70201	1,70221	1,71082	1,73507
1,76785	1,80659	1,82486	1,84615	1,86882	1,89947	1,92665	1,95187	1,95525	1,9597
1,98862	2,02167	2,03229	2,04559	2,0786	2,08066	2,08303	2,11074	2,12535	2,15711
2,16852	2,16853	2,17387	2,1829	2,18461	2,21153	2,223	2,28384	2,32461	2,32953
2,36196	2,37626	2,44883	2,45699	2,46182	2,46652	2,49316	2,49925	2,51401	2,55987
2,57366	2,5755	2,57945	2,58039	2,61418	2,62407	2,64161	2,64454	2,71155	2,72189
2,72323	2,72915	2,76961	2,93114	2,96446	2,97694	2,98316	2,98705	3,02769	3,03053

Таблица 2.1. Интервальный ряд

Интервал	n_k	w_k
$[-1.25, -0.71000]$	28	0,14
$(-0.71000, -0.17000]$	28	0,14
$(-0.17000, 0.37000]$	27	0,135
$(0.37000, 0.91000]$	22	0,11
$(0.91000, 1.45000]$	19	0,095
$(1.45000, 1.99000]$	27	0,135
$(1.99000, 2.53000]$	28	0,14
$(2.53000, 3.07000]$	21	0,105
	200	1

Рисунок 2. Гистограмма относительных частот для равномерного распределения

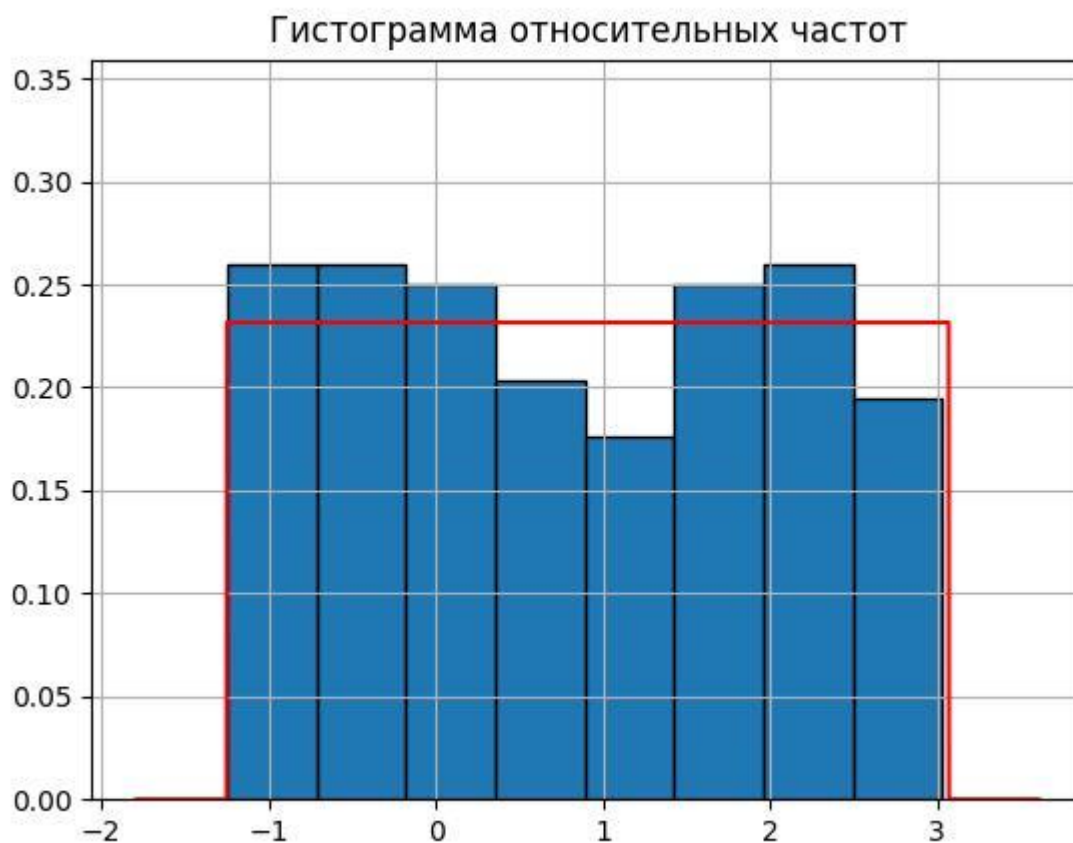


Таблица 2.2. Вычисление p_k^*

k	a_k	$f(a_k)$	$F(a_k)$	p_k^*
0	-1,25000	0,23148	0,00000	0,00000
1	-0,71000	0,23148	0,12500	0,12500
2	-0,17000	0,23148	0,25000	0,12500
3	0,37000	0,23148	0,37500	0,12500
4	0,91000	0,23148	0,50000	0,12500
5	1,45000	0,23148	0,62500	0,12500
6	1,99000	0,23148	0,75000	0,12500
7	2,53000	0,23148	0,87500	0,12500
8	3,07000	0,23148	1,00000	0,12500
				1,00000

Таблица 2.3. Вычисление выборочного значения критерия χ_B^2

k	Интервал	w_k	p_k^*	$ w_k - p_k^* $	$\frac{N(w_k - p_k^*)^2}{p_k^*}$
1	[-1.25, -0.71]	0,140	0,125	0,015	0,360
2	(-0.71, -0.17]	0,140	0,125	0,015	0,360
3	(-0.17, 0.37]	0,135	0,125	0,010	0,160
4	(0.37, 0.91]	0,110	0,125	0,015	0,360
5	(0.91, 1.45]	0,095	0,125	0,030	1,440
6	(1.45, 1.99]	0,135	0,125	0,010	0,160
7	(1.99, 2.53]	0,140	0,125	0,015	0,360
8	(2.53, 3.07]	0,105	0,125	0,020	0,640
		1,000	1,000	0,030	3,840

Вывод: значение $\chi_B^2 = 3,840$, а значение $\chi_{кр,\alpha}^2(7) = 14,06714$ при $\alpha = 0,05$.

Гипотеза о соответствии выборки равномерному распределению не противоречит экспериментальным данным (может быть принята) при уровне значимости $\alpha = 0.05$, т.к. $\chi_B^2 < \chi_{кр,\alpha}^2(7)$.

Задание 3. Проверка гипотезы о равномерном распределении с помощью критерия Колмогорова

Исходные данные: вариант 31

Рисунок 3. График эмпирической функции распределения (синий) и график функции распределения (красный) равномерного закона



Таблица 3.1. Вычисление выборочного значения критерия Колмогорова

a	b	N	D_N	$\sqrt{N}D_N$	y^*	$F(y^*)$	$F_N(y^*)$	$F_N(y^* - 0)$
-1.25	3.07	200	0.048831	0.690575	0.26705	0.351169	0.4	0.395

Вывод: значение $\sqrt{N}D_N = 0.690575$, а значение $k_\alpha = 1,358099$ при $\alpha = 0,05$.

Гипотеза о соответствии выборки равномерному распределению не противоречит экспериментальным данным (может быть принята) при уровне значимости $\alpha = 0.05$, т.к. $\sqrt{N}D_N \leq k_\alpha$.

Список литературы

1. Математическая статистика [Электронный ресурс]: метод. указания по выполнению лаб. работ / А.А. Лобузов — М.: МИРЭА, 2017.
2. Гмурман В.Е., Теория вероятностей и математическая статистика. – М.: Юрайт, 2020.
3. Ивченко Г.И., Медведев Ю.И. Математическая статистика. – М.: URSS, 2020.

Приложение

```
import matplotlib.pyplot as plt
import scipy.stats as sps
import pandas as pd
import numpy as np
import math

norm = pd.read_excel(r"D:\norm_data.xlsx").to_numpy()[0]
unif = pd.read_excel(r"D:\unif_data.xlsx").to_numpy()[0]

norm_sort = np.sort(norm)
unif_sort = np.sort(unif)

# ## Задание 1
# #### Интервальный ряд для нормального распределения

m = 1 + int(math.log(len(norm), 2))
norm_h = abs(norm_sort[-1] - norm_sort[0])/m

N = pd.DataFrame(columns=['n', 'w'])
i = 0

for k in range(1, m+1):

    if k == 1:
        interval = f'[{(norm_sort[0]+(k-1)*norm_h):.5f},
{(norm_sort[0]+k*norm_h):.5f}]'
    else:
        interval = f'({(norm_sort[0]+(k-1)*norm_h):.5f},
{(norm_sort[0]+k*norm_h):.5f})'

    count = 0
    while i < len(norm_sort):
        if norm_sort[i] <= norm_sort[0]+k*norm_h:
            i, count = i+1, count+1
        else:
            N.loc[interval] = [count, count/len(norm_sort)]
            break

N.loc[interval] = [count, count/len(norm_sort)]
N.loc[' '] = [N['n'].sum(), N['w'].sum()]

N
```

```

N.to_excel(r'D:\tables\table1_1.xlsx')

# #### Расчет оценок некоторых характеристик нормального распределения

norm_mean = 0 # математическое ожидание

for i in norm_sort:
    norm_mean += i/len(norm_sort)

norm_var = 0 # дисперсия

for i in norm_sort:
    norm_var += i*i/len(norm_sort)

norm_var -= norm_mean*norm_mean

norm_std = norm_var**0.5 # среднее квадратическое отклонение

print(f'Оценка математического ожидания: {norm_mean:.5f}')
print(f'Оценка дисперсии: {norm_var:.5f}')
print(f'Оценка среднего квадратического отклонения: {norm_std:.5f}')

# #### Таблица 1.2

table1_2 = pd.DataFrame(columns=['a_k', '(a_k-mean)/std', 'f/std', 'F', 'p_k'])

for k in range(m+1):

    z = norm_sort[0]+k*norm_h
    y = (z-norm_mean)/norm_std
    f = sps.norm.pdf(y)/norm_std
    F = sps.norm.cdf(y)

    if k > 1:
        table1_2.loc[k] = [z, y, f, F, F - sps.norm.cdf((norm_sort[0]+(k-1)*norm_h-
norm_mean)/norm_std)]
    elif k == 1:
        table1_2.loc[k] = [z, y, f, F, F]
    elif k == m:
        table1_2.loc[k] = [z, y, f, F, 1 - sps.norm.cdf((norm_sort[0]+(m-1)*norm_h-
norm_mean)/norm_std)]
    else:
        table1_2.loc[k] = [z, y, f, F, 0]

```

```
table1_2.loc[' '] = [' ', ' ', ' ', ' ', table1_2['p_k'].sum()]
table1_2
```

```
table1_2.to_excel(r'D:\tables\table1_2.xlsx')
```

```
# ### Гистограмма относительных частот и график плотности нормального
распределения
```

```
height_N = []
```

```
plt.grid(True)
plt.title('Гистограмма относительных частот')
n, bins, rects = plt.hist(norm_sort.tolist(), bins=8, ec='k')
```

```
i = 0
for r in rects:
    r.set_height(N['w'].iloc[i]/norm_h)
    height_N.append(N['w'].iloc[i]/norm_h)
    i+=1
```

```
x = np.arange(norm_sort[0]-norm_h, norm_sort[-1]+norm_h, 0.001)
plt.plot(x, sps.norm(norm_mean, norm_std).pdf(x), color='red')
```

```
plt.ylim(0, max(height_N)+0.1)
plt.savefig(r'D:\tables\plot1.jpg')
```

```
plt.show()
```

```
# ### Таблица 1.3
```

```
table1_3 = pd.DataFrame(columns=['Интервал', 'w_k', 'p_k', '|w_k - p_k|',
'something'])
```

```
for k in range(1, m+1):
```

```
    a = N.iloc[k-1]['w']
    b = table1_2.iloc[k]['p_k']
```

```
    if k == 1:
        interval = f'[{(norm_sort[0]+(k-1)*norm_h):.5f},
{(norm_sort[0]+k*norm_h):.5f}]'
    else:
        interval = f'[{(norm_sort[0]+(k-1)*norm_h):.5f},
{(norm_sort[0]+k*norm_h):.5f}]'
```

```

table1_3.loc[k] = [interval, a, b, abs(a-b), len(norm_sort)*(a-b)*(a-b)/b]

table1_3.loc[' '] = [' ', table1_3['w_k'].sum(), table1_3['p_k'].sum(), table1_3['|w_k
- p_k|'].max(), table1_3['something'].sum()]
table1_3

table1_3.to_excel(r'D:\tables\table1_3.xlsx')

print(f'Число степеней свободы: {m-3}')

### Задание 2
#### Интервальный ряд для равномерного распределения

a = -1.25; b = 3.07
unif_h = abs(b - a)/m

U = pd.DataFrame(columns=['n', 'w'])
i = 0

for k in range(1, m+1):

    if k == 1:
        interval = f'[{(a+(k-1)*unif_h)}, {(a+k*unif_h):.3f}]'
    else:
        interval = f'({(a+(k-1)*unif_h):.3f}, {(a+k*unif_h):.3f})'

    count = 0
    while i < len(unif_sort):
        if unif_sort[i] <= a+k*unif_h:
            i, count = i+1, count+1
        else:
            U.loc[interval] = [count, count/len(unif_sort)]
            break

U.loc[interval] = [count, count/len(unif_sort)]
U.loc[' '] = [U['n'].sum(), U['w'].sum()]

U.to_excel(r'D:\tables\table2_1.xlsx')

#### Таблица 2.2

table2_2 = pd.DataFrame(columns=['a_k', 'f(a_k)', 'F(a_k)', 'p_k'])

```

```

for k in range(m+1):

    a_k = a+k*unif_h
    f = sps.uniform(a, b-a).pdf(a_k)
    F = sps.uniform(a, b-a).cdf(a_k)

    table2_2.loc[k] = [a_k, f, F, F - sps.uniform(a, b-a).cdf(a+(k-1)*unif_h)]

table2_2.loc[' '] = [' ', ' ', ' ', table2_2['p_k'].sum()]
table2_2

table2_2.to_excel(r'D:\tables\table2_2.xlsx')

# #### Гистограмма относительных частот и график плотности равномерного
# распределения

height_U = []

plt.grid(True)
plt.title('Гистограмма относительных частот')
n, bins, rects = plt.hist(unif_sort.tolist(), bins=8, ec='k')

i = 0
for r in rects:
    r.set_height(U['w'].iloc[i]/unif_h)
    height_U.append(U['w'].iloc[i]/unif_h)
    i+=1

x = np.arange(a-unif_h, b+unif_h, 0.001)
plt.plot(x, sps.uniform(a, b-a).pdf(x), color='red')

plt.ylim(0, max(height_U)+0.1)
plt.savefig(r'D:\tables\plot2.jpg')

plt.show()

# #### Таблица 2.3

table2_3 = pd.DataFrame(columns=['Интервал', 'w_k', 'p_k', '|w_k - p_k|',
'something'])

for k in range(1, m+1):

```

```

x = U.iloc[k-1]['w']
y = table2_2.iloc[k]['p_k']

if k == 1:
    interval = f'[{(a+(k-1)*unif_h)}, {(a+k*unif_h):.2f}]'
else:
    interval = f'({(a+(k-1)*unif_h):.2f}, {(a+k*unif_h):.2f})'

table2_3.loc[k] = [interval, x, y, round(abs(x-y),5), round(len(unif_sort)*(x-
y)*(x-y)/y, 5)]

table2_3.loc[' '] = [' ', table2_3['w_k'].sum(), table2_3['p_k'].sum(), table2_3['|w_k
- p_k|'].max(), table2_3['something'].sum()]
table2_3

table2_3.to_excel(r'D:\tables\table2_3.xlsx')

print(f'Число степеней свободы: {m-1}')

# ## Задание 3
# #### Графики функций распределения

plt.grid(True)
plt.title('Функции распределения')

for i in range(len(unif_sort)-1):
    plt.plot([unif_sort[i], unif_sort[i+1]], [(i+1)/len(unif_sort), (i+1)/len(unif_sort)],
color='blue')

plt.plot([unif_sort[0]-unif_h, unif_sort[0]], [0, 0], color='blue')
plt.plot([unif_sort[-1], unif_sort[-1]+unif_h], [1, 1], color='blue')

x = np.arange(a, b, 0.001)
plt.plot(x, (x-a)/(b-a), color='red')

plt.plot([a-unif_h, a], [0, 0], color='red')
plt.plot([b, b+unif_h], [1, 1], color='red')

plt.savefig(r'D:\tables\plot3.jpg')
plt.show()

# #### Таблица 3.1

def F(x):

```

```

    if x < a:
        return 0
    elif x > b:
        return 1
    return (x-a)/(b-a)

def F_N(x):
    if x < unif_sort[0]:
        return 0
    elif x >= unif_sort[-1]:
        return 1
    return unif_sort.tolist().index(x)+1/len(unif)

max_D, max_y = -1000, -1000

for k in unif_sort:
    D_N = abs(F_N(k)-F(k)) if abs(F_N(k)-F(k)) > abs(F_N(k)-F(k)-1/len(unif))
    else abs(F_N(k)-F(k)-1/len(unif))
    if D_N > max_D:
        max_D, max_y = D_N, k

table3_1 = pd.DataFrame(columns=['a', 'b', 'N', 'D_N', 'D_N_N', 'y*', 'F(y*)',
'F_N(y*)', 'F_N(y*-0)'])
table3_1.loc[' '] = [a, b, len(unif), max_D, max_D*len(unif)**0.5, max_y,
F(max_y), F_N(max_y), F_N(max_y)-1/len(unif)]

table3_1.to_excel(r'D:\tables\table3_1.xlsx')

```