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

ИНСТИТУТ КИБЕРНЕТИКИ

КАФЕДРА ВЫСШЕЙ МАТЕМАТИКИ

Лабораторная работа 3

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

ВАРИАНТ 67

Тема: **«Проверка статистических гипотез с помощью критерия χ 2 и критерия Колмогорова»**

Выполнил:
Студент 3-го курса
Мусатов Д. Ю

Группа: КМБО-03-18

МОСКВА – 2021

Содержание

1 Задания

Задание I. Проверка гипотезы о нормальном распределении с помощью критерия χ^2. В соответствии с номером варианта взять из файла **D1_2021** выборку $\{x_1, ..., x_N\}$.
Построить группированную выборку (интервальный вариационный ряд).
Найти оценки математического ожидания $\widetilde{a}$ и дисперсии $\widetilde{\sigma}^2$.

Построить
1) таблицу 1.1 следующего вида

k	a_k	$\frac{a_k-\widetilde{a}}{\widetilde{\sigma}}$	$\frac{1}{\widetilde{\sigma}}\varphi\left(\frac{a_k-\widetilde{a}}{\widetilde{\sigma}}\right)$	$\Phi\left(\frac{a_k-\widetilde{a}}{\widetilde{\sigma}}\right)$	p_k^*
0	a_0	$\frac{a_0-\widetilde{a}}{\widetilde{\sigma}}$	$\frac{1}{\widetilde{\sigma}}\varphi\left(\frac{a_0-\widetilde{a}}{\widetilde{\sigma}}\right)$	$\Phi\left(\frac{a_0-\widetilde{a}}{\widetilde{\sigma}}\right)$	–
1	a_1	$\frac{a_1-\widetilde{a}}{\widetilde{\sigma}}$	$\frac{1}{\widetilde{\sigma}}\varphi\left(\frac{a_1-\widetilde{a}}{\widetilde{\sigma}}\right)$	$\Phi\left(\frac{a_1-\widetilde{a}}{\widetilde{\sigma}}\right)$	p_1^*
2	a_2	$\frac{a_2-\widetilde{a}}{\widetilde{\sigma}}$	$\frac{1}{\widetilde{\sigma}}\varphi\left(\frac{a_2-\widetilde{a}}{\widetilde{\sigma}}\right)$	$\Phi\left(\frac{a_2-\widetilde{a}}{\widetilde{\sigma}}\right)$	p_2^*
...	...	...	...	...	...
$m-1$	a_{m-1}	$\frac{a_{m-1}-\widetilde{a}}{\widetilde{\sigma}}$	$\frac{1}{\widetilde{\sigma}}\varphi\left(\frac{a_{m-1}-\widetilde{a}}{\widetilde{\sigma}}\right)$	$\Phi\left(\frac{a_{m-1}-\widetilde{a}}{\widetilde{\sigma}}\right)$	p_{m-1}^*
m	a_m	$\frac{a_m-\widetilde{a}}{\widetilde{\sigma}}$	$\frac{1}{\widetilde{\sigma}}\varphi\left(\frac{a_m-\widetilde{a}}{\widetilde{\sigma}}\right)$	$\Phi\left(\frac{a_m-\widetilde{a}}{\widetilde{\sigma}}\right)$	p_m^*
					$\sum_{k=1}^{m} p_k^*$

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

3) таблицу 1.2 следующего вида

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

где значения p_k^* находятся в соответствии с указаниями к **Заданию I**.

Задание II. Проверка гипотезы о равномерном распределении с помощью критерия χ^2.
В соответствии с номером варианта взять из файла **D2_2021** выборку $\{x_1, ..., x_N\}$. и значения a и b.

Построить:
1) группированную выборку и таблицу 2.1 следующего вида

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^*	$\|w_m - p_m^*\|$	$\frac{N(w_m - p_m^*)^2}{p_m^*}$
		$\sum_{k=1}^{m} w_k$	$\sum_{k=1}^{m} p_k^*$	$max\|w_k - p_k^*\|$	$\sum_{k=1}^{m} \frac{N(w_k - p_k^*)^2}{p_k^*}$

где значения p_k^* находятся в соответствии с указаниями к **Заданию II**;

2) 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 0,05.

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

В соответствии с номером варианта взять из файла **D2_2021** выборку $\{x_1, ..., x_N\}$. и значения a и b.

Построить:

1) на одном рисунке график эмпирической функции распределения $F_N(x)$ данной выборки и график функции распределения $F(x)$ равномерного закона на отрезке $[a, b]$;

2) таблицу 3.1 следующего вида

a	b	N	D_N	$D_N\sqrt{N}$	x^*	$F(x^*)$	$F_N(x^*)$	$F_N(x^* - 0)$

где $D_N = \max\limits_{1 \leqslant j \leqslant N}(max(|F_N(x_{(j)}) - F(x_{(j)})|, |F_N(x_{(j)} - 0) - F(x_{(j)})|))$,

$x^* = x_{(j)}$, если $D_N = max(|F_N(x_{(j)}) - F(x_{(j)})|, |F_N(x_{(j)} - 0) - F(x_{(j)})|)$.

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

Вычисления проводить с точностью до 0,00001.

2 Краткие теоретические сведения

2.1 Нормальное распределение(a, σ^2)

Плотность распределения: $f_\xi(x) = \frac{1}{\sigma} \cdot \varphi_0\left(\frac{x-a}{\sigma}\right)$, где $\varphi_0(x) = \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{x^2}{2}}$

Функция распределения: $F_\xi(x) = \Phi\left(\frac{x-a}{\sigma}\right)$, где $\Phi(x) = \int\limits_{-\infty}^{x} \varphi_0(t)dt$

Математическое ожидание: $M(x) = \widetilde{x} = a$

Дисперсия: $D(x) = s_B^2 = \sigma^2$

Среднее квадратическое отклонение: $\sigma = \sqrt{s_B^2} = \sigma$

2.2 Равномерное распределение

Плотность распределения: $f_\xi(x) = \begin{cases} 0, & x \notin (a,b) \\ C, & x \in (a,b) \end{cases}, \quad C = \frac{1}{b-a}$

Функция распределения: $F_\xi(x) = \int\limits_{-\infty}^{x} f_\xi(t)dt = \begin{cases} 0, & x \leqslant a \\ \frac{x-a}{b-a}, & x \in (a,b) \\ 1, & x \geqslant b \end{cases}$

Математическое ожидание: $M = \widetilde{x} = \frac{a+b}{2}$

Дисперсия: $D = s_B^2 = \frac{(b-a)^2}{12}$

Среднее квадратическое отклонение: $\sigma = \sqrt{D} = \sqrt{s_B^2} = \frac{b-a}{2\sqrt{3}}$

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

Когда наблюдаемые случайные величины ξ и η дискретны, можно применить критерий однородности, основанный на распределении χ^2 .

Пусть дискретные случайные величины ξ и η принимают значения $x_1^* < x_2^* < ...$

Схема проверки гипотезы H_0= $\{F_\xi(x) = F_\eta(x)\}$ при конкурирующей гипотезе H_1= $\{F_\xi(x) \neq F_\eta(x)\}$ по выборкам $x = (x_1, x_2, ..., x_N)$ и $y = (y_1, y_2, ..., y_M)$, полученных при наблюдениях случайных величин ξ и η соответственно, при уровне значимости α:

1. По числовым выборкам $x = (x_1, x_2, ..., x_N)$ и $y = (y_1, y_2, ..., y_M)$, находим частоты n_{i_1} и n_{i_2} (n_{i_1} - число значений x_i^*, встречающихся в выборке $x = (x_1, x_2, ..., x_N)$, n_{i_2} - число значений y_i^*, встречающихся в выборке $y = (y_1, y_2, ..., y_M)$).

2. Находим выборочное значение критерия
$$\xi_B^2 = (N+M)\left[\sum_{i=1}^{m}\left(\frac{n_{i_1}^2}{N(n_{i_1}+n_{i_2})} + \frac{n_{i_2}^2}{M(n_{i_1}+n_{i_2})}\right) - 1\right]$$

3. По заданному значению уровня значимости α берем по функции распределения $\chi^2(l)$ с число степеней свободы $l = m - 1$ критическое значение $\chi^2_{kp,\alpha}(l)$.

Далее делается вывод о справедливости гипотезы:
если $\chi^2(l) \leqslant \chi^2_{kp,\alpha}(l)$, то при уровне значимости α принимается основная гипотеза H_0;
если $\chi^2(l) > \chi^2_{kp,\alpha}(l)$, то при уровне значимости α принимается альтернативная гипотеза H_1 .

Критическое значение $\chi^2_{kp,\alpha}(l) = F_{\chi^2(l)}(1-\alpha)$, где $F_{\chi^2(l)}(1-\alpha)$ - функция распределения закона $\chi^2(l)$ с плотностью

$$f(x) = \begin{cases} 0, \quad x \leqslant 0; \\ \\ \dfrac{x^{\frac{l}{2}-1}}{2^{\frac{l}{2}}\Gamma\left(\dfrac{l}{2}\right)}\, e^{-\frac{x}{2}}, \quad x > 0 \end{cases}$$

2.4 Общая схема проверки статистических гипотез с помощью критерия Колмогорова

Критерий Колмогорова-Смирнова для проверки однородности двух случайных выборок применяется в случае, когда функции распределения наблюдаемых случайных величины ξ и η непрерывны.

При проверке гипотезы H_0= $\{F_\xi(x) = F_\eta(x)\}$ против гипотезы H_1= $\{F_\xi(x) \neq F_\eta(x)\}$ в критерии Колмогорова-Смирнова рассматривается статистика

$D_{N,M}(X,Y) = sup\{ \; |F_N(x,X) - F_M(x,Y)| : -\infty < x < +\infty \; \}$,

для которой выполняется свойство

$$P\left(D_{N,M}(X,Y)\sqrt{\frac{NM}{N+M}} \leqslant z\right) \underset{N\longrightarrow\infty}{\longrightarrow} K(z), \text{ где } \quad K(z) = \begin{cases} 0, \quad z \leqslant 0; \\ \sum\limits_{j=-\infty}^{+\infty}(-1)^j e^{-2j^2z^2}, \quad z > 0 \end{cases}$$

Схема проверки гипотезы H_0 при конкурирующей гипотезе H_1 по выборкам $x = (x_1, x_2, ..., x_N)$ и $y = (y_1, y_2, ..., y_M)$, полученных при наблюдениях случайных величин ξ и η соответственно, при уровне значимости α:

1. По числовым выборкам $x = (x_1, x_2, ..., x_N)$ и $y = (y_1, y_2, ..., y_M)$, строим вариационные ряды $x_{(1)} \leqslant x_{(2)} \leqslant ... \leqslant x_{(N-1)} \leqslant x_{(N)}$ и $y_{(1)} \leqslant y_{(2)} \leqslant ... \leqslant y_{(M-1)} \leqslant y_{(M)}$ соответственно.

2. Находим значение

$$\begin{aligned} D_{N,M} = \max_{j,k}\{ \; &|F_N(x_{(j)}) - F_M(x_{(j)})|, \; |F_N(x_{(j)-0}) - F_M(x_{(j)})|, \\ &|F_N(y_{(k)}) - F_M(y_{(k)})|, \; |F_N(y_{(k)}) - F_M(y_{(k)} - 0)| \; \} \end{aligned},$$

где $F_N(x_{(j)}) = \dfrac{j}{N}, \;\; F_N(x_{(j)} - 0) = \dfrac{j-1}{N}, \;\; F_M(y_k) = \dfrac{k}{M}, \;\; F_M(y_k - 0) = \dfrac{k-1}{M}$

$$F_M(x(j)) = \begin{cases} 0, x_{(j)} < y_{(1)}; \\ \dfrac{k}{M}, \; y_{(k)} \leqslant x_{(j)} < y_{(k+1)}; \\ 1, \; x_{(j)} \geqslant y_{(M)}; \end{cases} \qquad F_N(y(k)) = \begin{cases} 0, y_{(k)} < x_{(1)}; \\ \dfrac{j}{N}, \; y_{(j)} \leqslant y_{(k)} < x_{(j+1)}; \\ 1, \; y_{(k)} \geqslant x_{(N)}; \end{cases}$$

Затем находим $K_{N,M} = D_{N,M}\sqrt{\dfrac{NM}{N+M}}$

3. По заданному значению уровня значимости α берем по функции распределения Колмогорова критическое значение k_α.

Далее делается вывод о справедливости гипотезы:
если $K_{N,M} \leqslant k_\alpha$, то при уровне значимости α принимается основная H_0;
если $K_{N,M} > k_\alpha$, то при уровне значимости α принимается альтернативная гипотеза H_1.

Иногда рассматривают конкурирующие гипотезы
$H_1^+ = \{ \; sup\{M[F_N(x,X) - F_M(x,Y)] : \; -\infty < x < +\infty\}\}$ и
$H_1^- = \{ \; inf\{M[F_N(x,X) - F_M(x,Y)] : \; -\infty < x < +\infty\}\}$

При проверке гипотезы H_0 против гипотезы H_1^+ берут статистику

$$K^+_{N,M}(X,Y) = D^+_{N,M}(X,Y)\sqrt{\frac{NM}{N+M}},$$

где $D^+_{N,M}(X,Y) = sup\{M[F_N(x,X) - F_M(x,Y)] : \; -\infty < x < +\infty\}$.

При проверке гипотезы H_0 против гипотезы H_1^- берут статистику

$$K_{N,M}^-(X,Y) = D_{N,M}^-(X,Y)\sqrt{\frac{NM}{N+M}},$$

где $D_{N,M}^-(X,Y) = -inf\{M[F_N(x,X) - F_M(x,Y)]:\ -\infty < x < +\infty\}$.

Статистики $K_{N,M}^+(X,Y)$, $K_{N,M}^-(X,Y)$, а также $K_{M,N}^+(X,Y)$, $K_{M,N}^-(X,Y)$ имеют одинаковое распределение и для них выполняется свойство

$$\lim_{N\to\infty,\ M\to\infty} P(K_{N,M}^+(X,Y) \leqslant z) = \lim_{N\to\infty,\ M\to\infty} P(K_{N,M}^-(X,Y) \leqslant z) =$$

$$= \lim_{N\to\infty,\ M\to\infty} P(K_{M,N}^+(X,Y) \leqslant z) = \lim_{N\to\infty,\ M\to\infty} P(K_{M,N}^-(X,Y) \leqslant z) = 1 - e^{-2z^2},\ 0 \leqslant z < \infty.$$

Схема проверки гипотезы H_0 при конкуриющей гипотезе H_1^+ по выборкам $x = (x_1, x_2, ..., x_N)$ и $y = (y_1, y_2, ..., y_M)$, полученных при наблюдениях случайных величин ξ и η соответственно, при уровне значимости α :

1. По числовым выборкам $x = (x_1, x_2, ..., x_N)$ и $y = (y_1, y_2, ..., y_M)$, строим вариационные ряды $x_{(1)} \leqslant x_{(2)} \leqslant ... \leqslant x_{(N-1)} \leqslant x_{(N)}$ и $y_{(1)} \leqslant y_{(2)} \leqslant ... \leqslant y_{(M-1)} \leqslant y_{(M)}$ соответственно.

2. Находим значение

$$\begin{aligned} D_{N,M} = \max_{j,k}\{\ & |F_N(x_{(j)}) - F_M(x_{(j)})|,\ |F_N(x_{(j)-0}) - F_M(x_{(j)})|, \\ & |F_N(y_{(k)}) - F_M(y_{(k)})|,\ |F_N(y_{(k)}) - F_M(y_{(k)} - 0)|\ \} \end{aligned},$$

где $F_N(x_{(j)}) = \dfrac{j}{N}$, $F_N(x_{(j)} - 0) = \dfrac{j-1}{N}$, $F_M(y_k) = \dfrac{k}{M}$, $F_M(y_k - 0) = \dfrac{k-1}{M}$

$$F_M(x(j)) = \begin{cases} 0, x_{(j)} < y_{(1)}; \\ \dfrac{k}{M},\ y_{(k)} \leqslant x_{(j)} < y_{(k+1)}; \\ 1,\ x_{(j)} \geqslant y_{(M)}; \end{cases} \qquad F_N(y(k)) = \begin{cases} 0, y_{(k)} < x_{(1)}; \\ \dfrac{j}{N},\ y_{(j)} \leqslant y_{(k)} < x_{(j+1)}; \\ 1,\ y_{(k)} \geqslant x_{(N)}; \end{cases}$$

Затем находим $K_{N,M}^+ = D_{N,M}\sqrt{\dfrac{NM}{N+M}}$

3. По заданному значению уровня значимости α берем критическое значение $s_\alpha = \sqrt{-\dfrac{1}{2}\,ln\,\alpha}$.

Далее делается вывод о справедливости гипотезы:
если $K_{N,M}^+ \leqslant s_\alpha$, то при уровне значимости α принимается основная гипотеза H_0;
если $K_{N,M}^+ \geqslant s_\alpha$, то при уровне значимости α принимается альтернативная гипотеза H_1^+;

Схема проверки гипотезы H_0 при конкурирующей гипотезе H_1^+ проводится аналогично.

3 Результаты расчётов

В программе расчёта был использован язык программирования Python.
Работа осуществлялась в среде Jupyter Notebook.

3.1 Задание №1

Исходная выборка

0,57449	0,03156	-0,37043	0,17178	1,44855	-0,32346	-0,29277	0,23479	-0,75252	-0,10496
1,77142	0,56115	1,4909	1,49796	0,17993	0,7799	1,32721	0,60618	0,00642	0,02607
0,97563	0,06541	2,1449	0,41022	0,04321	0,27108	-0,00354	-0,34842	0,72862	1,55716
-0,77904	0,57806	-0,33951	-0,47123	0,07041	0,70973	0,25956	-0,95821	-0,71237	1,43153
0,6307	0,44395	-0,68234	-0,32134	0,12111	1,94839	0,56118	0,43551	0,21878	-1,24623
0,25832	0,60141	1,41768	0,52711	1,472	0,10061	0,38725	1,70926	0,91447	0,80336
-0,04137	-0,93922	-0,26222	-1,81064	1,86124	0,70078	1,19596	-0,02966	0,40458	1,3365
0,80947	-0,34906	1,65231	1,661	0,20538	0,2126	0,30023	0,15962	2,2281	-0,15093
-0,37135	-1,99111	0,42467	-0,38438	-0,47199	-0,57173	-0,88295	1,27701	-1,11478	-0,17734
-0,75236	-0,90651	2,17084	-0,25015	0,29631	0,98042	-0,55342	1,35258	-1,10275	0,34195
-0,45308	1,73779	0,65766	1,18183	-0,38429	-0,03274	0,06211	1,92864	0,17461	0,71345
-0,11517	0,87593	-1,04093	-0,27874	0,60467	0,17479	-0,72073	0,61575	0,18401	0,33681
-1,06277	-0,5331	-0,24172	0,33456	0,89634	0,50277	0,67348	0,32683	-0,125	-0,20868
0,24373	0,89785	0,38912	-0,59241	-0,24913	1,10598	2,42994	-0,65016	2,095	0,59435
0,78888	-0,14191	-1,10848	0,07334	1,7372	0,0085	0,93942	1,06237	-0,50819	0,77735
1,00513	1,85915	0,65868	-0,14702	0,77537	-0,70454	1,29008	-0,54021	1,15592	1,24858
0,88847	0,42821	-0,52983	1,38144	1,18535	0,70942	0,90209	0,67157	-0,53932	-0,45475
0,84618	1,09132	1,89393	2,04818	0,3525	0,42235	0,12583	-1,29217	-0,53414	1,13713
-0,17522	0,43283	-0,50829	-1,2331	-0,49711	-0,08304	0,39196	0,50895	1,34122	0,2668
0,62106	-0,48537	-0,06912	1,15618	-0,09977	-1,19084	1,55921	2,0015	0,69147	-1,50157

Упорядоченная выборка

-1,99111	-1,81064	-1,50157	-1,29217	-1,24623	-1,2331	-1,19084	-1,11478	-1,10848	-1,10275
-1,06277	-1,04093	-0,95821	-0,93922	-0,90651	-0,88295	-0,77904	-0,75252	-0,75236	-0,72073
-0,71237	-0,70454	-0,68234	-0,65016	-0,59241	-0,57173	-0,55342	-0,54021	-0,53932	-0,53414
-0,5331	-0,52983	-0,50829	-0,50819	-0,49711	-0,48537	-0,47199	-0,47123	-0,45475	-0,45308
-0,38438	-0,38429	-0,37135	-0,37043	-0,34906	-0,34842	-0,33951	-0,32346	-0,32134	-0,29277
-0,27874	-0,26222	-0,25015	-0,24913	-0,24172	-0,20868	-0,17734	-0,17522	-0,15093	-0,14702
-0,14191	-0,125	-0,11517	-0,10496	-0,09977	-0,08304	-0,06912	-0,04137	-0,03274	-0,02966
-0,00354	0,00642	0,0085	0,02607	0,03156	0,04321	0,06211	0,06541	0,07041	0,07334
0,10061	0,12111	0,12583	0,15962	0,17178	0,17461	0,17479	0,17993	0,18401	0,20538
0,2126	0,21878	0,23479	0,24373	0,25832	0,25956	0,2668	0,27108	0,29631	0,30023
0,32683	0,33456	0,33681	0,34195	0,3525	0,38725	0,38912	0,39196	0,40458	0,41022
0,42235	0,42467	0,42821	0,43283	0,43551	0,44395	0,50277	0,50895	0,52711	0,56115
0,56118	0,57449	0,57806	0,59435	0,60141	0,60467	0,60618	0,61575	0,62106	0,6307
0,65766	0,65868	0,67157	0,67348	0,69147	0,70078	0,70942	0,70973	0,71345	0,72862
0,77537	0,77735	0,7799	0,78888	0,80336	0,80947	0,84618	0,87593	0,88847	0,89634
0,89785	0,90209	0,91447	0,93942	0,97563	0,98042	1,00513	1,06237	1,09132	1,10598
1,13713	1,15592	1,15618	1,18183	1,18535	1,19596	1,24858	1,27701	1,29008	1,32721
1,3365	1,34122	1,35258	1,38144	1,41768	1,43153	1,44855	1,472	1,4909	1,49796
1,55716	1,55921	1,65231	1,661	1,70926	1,7372	1,73779	1,77142	1,85915	1,86124
1,89393	1,92864	1,94839	2,0015	2,04818	2,095	2,1449	2,17084	2,2281	2,42994

Группированная выборка(интервальный вариационный ряд)

Интервалы	x*_k	n_k	w_k
[-1.99111 , -1.43848]	-1.7147943749999999	3	0.015
(-1.43848 , -0.88585]	-1.162163124999997	12	0.06
(-0.88585 , -0.33322]	-0.6095318749999998	32	0.16
(-0.33322 , 0.21942]	-0.05690062499999976	45	0.225
(0.21942 , 0.77205]	0.49573062500000026	48	0.24
(0.77205 , 1.32468]	1.0483618750000003	29	0.145
(1.32468 , 1.87731]	1.6009931250000005	21	0.105
(1.87731 , 2.42994]	2.1536243750000006	10	0.05
		200	1

Математическое ожидание: $\widetilde{a} = 0.34099$

Выборочная дисперсия: $\widetilde{\sigma}^2 = 0.87348$

Таблица 1.1:

k	a_k	$\frac{a_k - \widetilde{a}}{\widetilde{\sigma}}$	$\frac{1}{\widetilde{\sigma}}\varphi\left(\frac{a_k - \widetilde{a}}{\widetilde{\sigma}}\right)$	$\Phi\left(\frac{a_k - \widetilde{a}}{\widetilde{\sigma}}\right)$	p_k^*
0	-1,99111	-2,66989	0,01293	0,00379	-
1	-1,43848	-2,03721	0,05734	0,02081	0,02081
2	-0,88585	-1,40454	0,17033	0,08008	0,05926
3	-0,33322	-0,77186	0,33907	0,2201	0,14002
4	0,21942	-0,13919	0,45232	0,44465	0,22455
5	0,77205	0,49349	0,40437	0,68917	0,24451
6	1,32468	1,12616	0,24225	0,86995	0,18079
7	1,87731	1,75884	0,09725	0,9607	0,09075
8	2,42994	2,39151	0,02616	0,99161	0,0393
					1

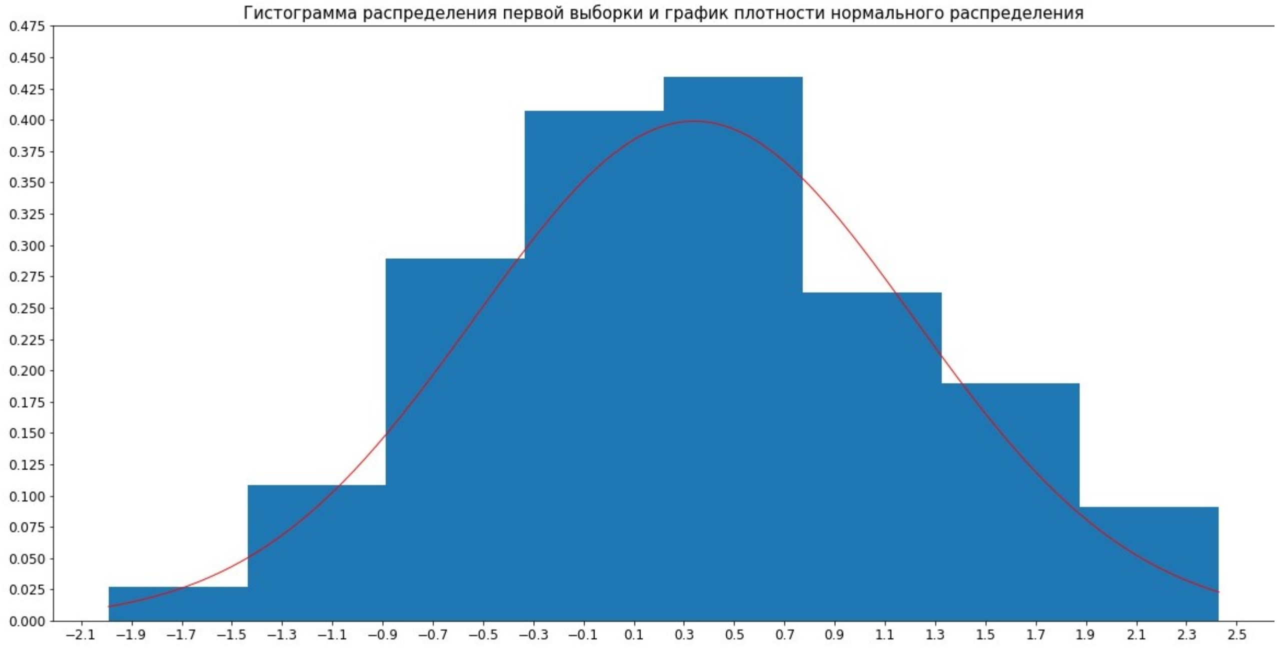


Таблица 1.2:

k	Интервал	w_k	p_k^*	$\lvert w_k - p_k^*\rvert$	$\frac{N(w_k - p_k^*)^2}{p_k^*}$
1	[-1.99111 , -1.43848]	0,015	0,02081	0,00581	0,32442
2	(-1.43848 , -0.88585]	0,06	0,05926	0,00074	0,00185
3	(-0.88585 , -0.33322]	0,16	0,14002	0,01998	0,5702
4	(-0.33322 , 0.21942]	0,225	0,22455	0,00045	0,00018
5	(0.21942 , 0.77205]	0,24	0,24451	0,00451	0,01664
6	(0.77205 , 1.32468]	0,145	0,18079	0,03579	1,41703
7	(1.32468 , 1.87731]	0,105	0,09075	0,01425	0,44752
8	(1.87731 , 2.42994]	0,05	0,0393	0,0107	0,58265
		1	1	0,03579	3,36049

Из таблицы 1.2 значение выборочного критерия $\chi^2_B = 3.36049$

3.2 Задание №2

Исходная выборка

1,82676	1,11916	0,95513	1,48516	1,98652	0,27193	0,44206	-0,00891	3,88508	2,98723
2,86516	-0,83333	0,80208	3,55245	3,35499	0,49812	3,97689	3,96353	3,40892	3,8591
-0,5972	-0,04325	0,29231	-0,00165	1,29094	-0,61537	3,10636	1,41709	3,05054	-0,67723
3,6337	2,86254	0,80153	1,30167	-0,60476	2,75841	-0,6826	1,47998	2,83601	-1,24538
0,91579	4,54419	2,1536	0,40467	4,55834	4,58079	3,53397	2,2099	3,12082	4,50137
-1,27991	0,94751	0,37514	-0,13598	4,18899	3,19334	1,63662	0,21752	3,58673	0,48647
4,23065	3,97152	-0,89134	-1,50591	3,82439	4,43701	-0,04838	3,74155	2,6125	-0,00525
1,99921	-0,62995	2,23918	1,49791	3,8053	3,50719	3,06872	2,87828	-1,04335	4,53547
-0,63892	0,42651	4,37461	-1,38336	4,07431	-0,2803	2,9122	1,93046	4,34338	2,4109
-0,1924	1,6631	3,39403	0,2737	3,19963	1,84653	0,78714	-0,28664	-1,0478	1,4873
1,5242	1,19187	4,20039	0,87077	-0,06637	1,44911	-0,03789	0,89096	-0,34453	2,51966
4,36308	3,87282	-0,14952	1,50462	1,91838	-0,98009	-0,37589	1,73313	-1,1989	1,65682
-0,70706	2,02141	4,05186	4,44397	1,1141	0,84192	0,56382	2,25242	1,12307	-1,42143
1,91692	2,84558	4,15373	4,07333	1,21524	0,63629	0,22588	2,72761	3,28283	4,10987
2,81417	-0,00055	0,99972	2,85095	1,65627	0,7007	-0,44195	1,20261	1,98725	-0,67253
-1,18383	2,7464	3,95633	-0,42261	1,55153	-1,24587	1,92619	-0,52869	4,15196	3,19402
-0,21338	1,03431	3,12246	1,87422	3,5397	0,10357	4,27097	1,37121	4,25474	-0,39053
1,62424	0,39015	1,43923	3,34121	1,6448	3,83476	0,60286	0,97093	2,56041	2,12109
-0,5573	3,44259	0,37838	0,45255	0,08045	-1,44382	1,68244	3,99513	0,62311	2,6976
3,56185	3,15607	-0,34722	3,43271	0,18306	-0,10962	1,72947	3,02395	0,69472	2,17141

Упорядоченная выборка

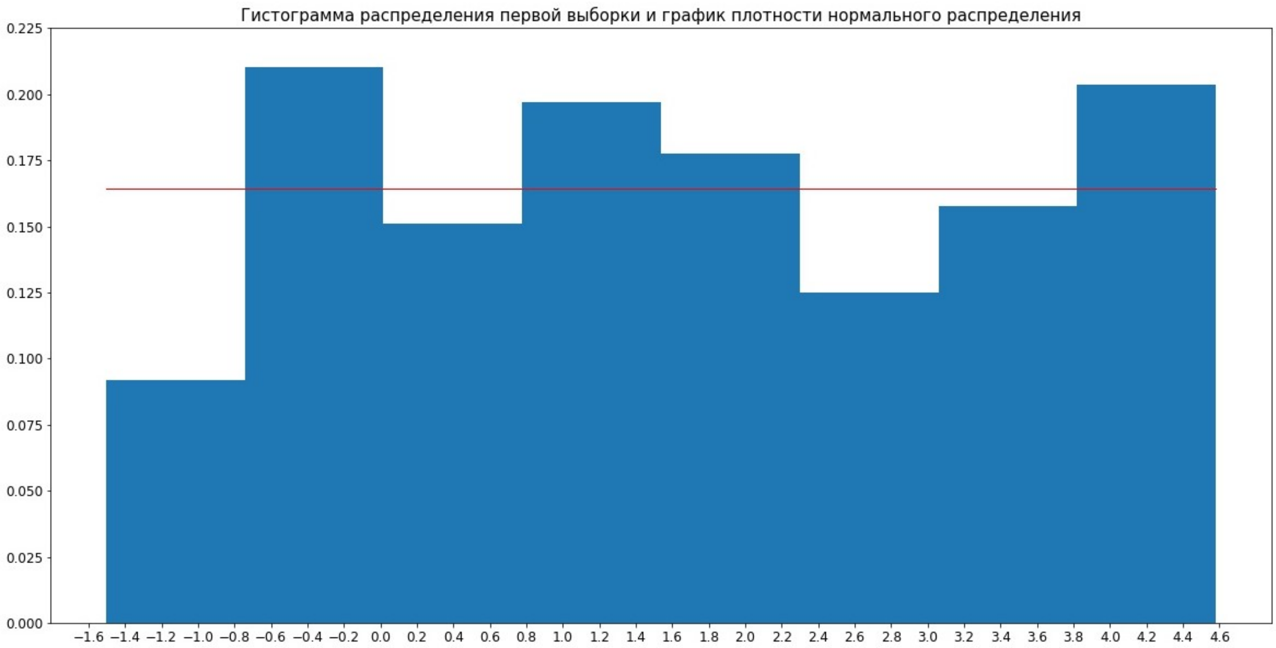
-1,50591	-1,44382	-1,42143	-1,38336	-1,27991	-1,24587	-1,24538	-1,1989	-1,18383	-1,0478
-1,04335	-0,98009	-0,89134	-0,83333	-0,70706	-0,6826	-0,67723	-0,67253	-0,63892	-0,62995
-0,61537	-0,60476	-0,5972	-0,5573	-0,52869	-0,44195	-0,42261	-0,39053	-0,37589	-0,34722
-0,34453	-0,28664	-0,2803	-0,21338	-0,1924	-0,14952	-0,13598	-0,10962	-0,06637	-0,04838
-0,04325	-0,03789	-0,00891	-0,00525	-0,00165	-0,00055	0,08045	0,10357	0,18306	0,21752
0,22588	0,27193	0,2737	0,29231	0,37514	0,37838	0,39015	0,40467	0,42651	0,44206
0,45255	0,48647	0,49812	0,56382	0,60286	0,62311	0,63629	0,69472	0,7007	0,78714
0,80153	0,80208	0,84192	0,87077	0,89096	0,91579	0,94751	0,95513	0,97093	0,99972
1,03431	1,1141	1,11916	1,12307	1,19187	1,20261	1,21524	1,29094	1,30167	1,37121
1,41709	1,43923	1,44911	1,47998	1,48516	1,4873	1,49791	1,50462	1,5242	1,55153
1,62424	1,63662	1,6448	1,65627	1,65682	1,6631	1,68244	1,72947	1,73313	1,82676
1,84653	1,87422	1,91692	1,91838	1,92619	1,93046	1,98652	1,98725	1,99921	2,02141
2,12109	2,1536	2,17141	2,2099	2,23918	2,25242	2,4109	2,51966	2,56041	2,6125
2,6976	2,72761	2,7464	2,75841	2,81417	2,83601	2,84558	2,85095	2,86254	2,86516
2,87828	2,9122	2,98723	3,02395	3,05054	3,06872	3,10636	3,12082	3,12246	3,15607
3,19334	3,19402	3,19963	3,28283	3,34121	3,35499	3,39403	3,40892	3,43271	3,44259
3,50719	3,53397	3,5397	3,55245	3,56185	3,58673	3,6337	3,74155	3,8053	3,82439
3,83476	3,8591	3,87282	3,88508	3,95633	3,96353	3,97152	3,97689	3,99513	4,05186
4,07333	4,07431	4,10987	4,15196	4,15373	4,18899	4,20039	4,23065	4,25474	4,27097
4,34338	4,36308	4,37461	4,43701	4,44397	4,50137	4,53547	4,54419	4,55834	4,58079

Таблица 2.1

k	Интервал	w_k	p_k^*	$\lvert w_k - p_k^* \rvert$	$\frac{N(w_k - p_k^*)^2}{p_k^*}$
1	[-1.51 , -0.74507]	0,07	0,125	0,055	4,84
2	(-0.74507 , 0.01577]	0,16	0,125	0,035	1,96
3	(0.01577 , 0.7766]	0,115	0,125	0,01	0,16
4	(0.7766 , 1.53744]	0,15	0,125	0,025	1
5	(1.53744 , 2.29828]	0,135	0,125	0,01	0,16
6	(2.29828 , 3.05912]	0,095	0,125	0,03	1,44
7	(3.05912 , 3.81995]	0,12	0,125	0,005	0,04
8	(3.81995 , 4.59]	0,155	0,125	0,03	1,44
		1	1	0,055	11,04

Из таблицы 2.1 значение выборочного критерия $\chi_B^2 = 11.04$

Интервалы	x*_k	n_k	w_k
[-1.51 , -0.74507]	-1.12549125	14	0.07
(-0.74507 , 0.01577]	-0.36465375	32	0.16
(0.01577 , 0.7766]	0.39618375000000006	23	0.115
(0.7766 , 1.53744]	1.1570212500000001	30	0.15
(1.53744 , 2.29828]	1.9178587500000002	27	0.135
(2.29828 , 3.05912]	2.6786962500000002	19	0.095
(3.05912 , 3.81995]	3.4395337500000003	24	0.12
(3.81995 , 4.59]	4.20037125	31	0.155
		200	1


Гистограмма распределения первой выборки и график плотности нормального распределения
0.225
0.200
0.175
0.150
0.125
0.100
0.075
0.050
0.025
0.000
−1.6 −1.4 −1.2 −1.0 −0.8 −0.6 −0.4 −0.2 0.0 0.2 0.4 0.6 0.8 1.0 1.2 1.4 1.6 1.8 2.0 2.2 2.4 2.6 2.8 3.0 3.2 3.4 3.6 3.8 4.0 4.2 4.4 4.6

3.3 Задание №3

Исходная выборка

1,82676	1,11916	0,95513	1,48516	1,98652	0,27193	0,44206	-0,00891	3,88508	2,98723
2,86516	-0,83333	0,80208	3,55245	3,35499	0,49812	3,97689	3,96353	3,40892	3,8591
-0,5972	-0,04325	0,29231	-0,00165	1,29094	-0,61537	3,10636	1,41709	3,05054	-0,67723
3,6337	2,86254	0,80153	1,30167	-0,60476	2,75841	-0,6826	1,47998	2,83601	-1,24538
0,91579	4,54419	2,1536	0,40467	4,55834	4,58079	3,53397	2,2099	3,12082	4,50137
-1,27991	0,94751	0,37514	-0,13598	4,18899	3,19334	1,63662	0,21752	3,58673	0,48647
4,23065	3,97152	-0,89134	-1,50591	3,82439	4,43701	-0,04838	3,74155	2,6125	-0,00525
1,99921	-0,62995	2,23918	1,49791	3,8053	3,50719	3,06872	2,87828	-1,04335	4,53547
-0,63892	0,42651	4,37461	-1,38336	4,07431	-0,2803	2,9122	1,93046	4,34338	2,4109
-0,1924	1,6631	3,39403	0,2737	3,19963	1,84653	0,78714	-0,28664	-1,0478	1,4873
1,5242	1,19187	4,20039	0,87077	-0,06637	1,44911	-0,03789	0,89096	-0,34453	2,51966
4,36308	3,87282	-0,14952	1,50462	1,91838	-0,98009	-0,37589	1,73313	-1,1989	1,65682
-0,70706	2,02141	4,05186	4,44397	1,1141	0,84192	0,56382	2,25242	1,12307	-1,42143
1,91692	2,84558	4,15373	4,07333	1,21524	0,63629	0,22588	2,72761	3,28283	4,10987
2,81417	-0,00055	0,99972	2,85095	1,65627	0,7007	-0,44195	1,20261	1,98725	-0,67253
-1,18383	2,7464	3,95633	-0,42261	1,55153	-1,24587	1,92619	-0,52869	4,15196	3,19402
-0,21338	1,03431	3,12246	1,87422	3,5397	0,10357	4,27097	1,37121	4,25474	-0,39053
1,62424	0,39015	1,43923	3,34121	1,6448	3,83476	0,60286	0,97093	2,56041	2,12109
-0,5573	3,44259	0,37838	0,45255	0,08045	-1,44382	1,68244	3,99513	0,62311	2,6976
3,56185	3,15607	-0,34722	3,43271	0,18306	-0,10962	1,72947	3,02395	0,69472	2,17141

Упорядоченная выборка

-1,50591	-1,44382	-1,42143	-1,38336	-1,27991	-1,24587	-1,24538	-1,1989	-1,18383	-1,0478
-1,04335	-0,98009	-0,89134	-0,83333	-0,70706	-0,6826	-0,67723	-0,67253	-0,63892	-0,62995
-0,61537	-0,60476	-0,5972	-0,5573	-0,52869	0,44195	-0,42261	-0,39053	-0,37589	-0,34722
-0,34453	-0,28664	-0,2803	-0,21338	-0,1924	-0,14952	-0,13598	-0,10962	-0,06637	-0,04838
-0,04325	-0,03789	-0,00891	-0,00525	-0,00165	-0,00055	0,08045	0,10357	0,18306	0,21752
0,22588	0,27193	0,2737	0,29231	0,37514	0,37838	0,39015	0,40467	0,42651	0,44206
0,45255	0,48647	0,49812	0,56382	0,60286	0,62311	0,63629	0,69472	0,7007	0,78714
0,80153	0,80208	0,84192	0,87077	0,89096	0,91579	0,94751	0,95513	0,97093	0,99972
1,03431	1,1141	1,11916	1,12307	1,19187	1,20261	1,21524	1,29094	1,30167	1,37121
1,41709	1,43923	1,44911	1,47998	1,48516	1,4873	1,49791	1,50462	1,5242	1,55153
1,62424	1,63662	1,6448	1,65627	1,65682	1,6631	1,68244	1,72947	1,73313	1,82676
1,84653	1,87422	1,91692	1,91838	1,92619	1,93046	1,98652	1,98725	1,99921	2,02141
2,12109	2,1536	2,17141	2,2099	2,23918	2,25242	2,4109	2,51966	2,56041	2,6125
2,6976	2,72761	2,7464	2,75841	2,81417	2,83601	2,84558	2,85095	2,86254	2,86516
2,87828	2,9122	2,98723	3,02395	3,05054	3,06872	3,10636	3,12082	3,12246	3,15607
3,19334	3,19402	3,19963	3,28283	3,34121	3,35499	3,39403	3,40892	3,43271	3,44259
3,50719	3,53397	3,5397	3,55245	3,56185	3,58673	3,6337	3,74155	3,8053	3,82439
3,83476	3,8591	3,87282	3,88508	3,95633	3,96353	3,97152	3,97689	3,99513	4,05186
4,07333	4,07431	4,10987	4,15196	4,15373	4,18899	4,20039	4,23065	4,25474	4,27097
4,34338	4,36308	4,37461	4,43701	4,44397	4,50137	4,53547	4,54419	4,55834	4,58079

График эмпирической функции распределения выборки $F_N(x)$ и график функции распределения $F(x)$ равномерного закона на отрезке $[a, b]$ на одном рисунке:

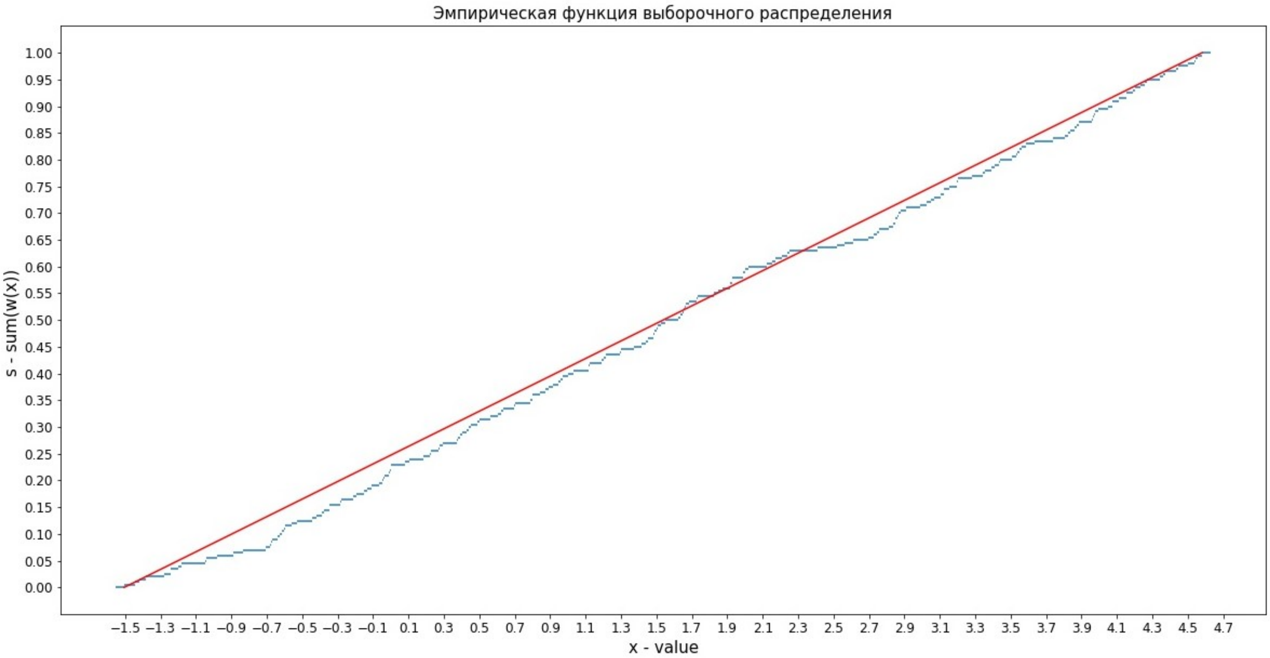


Таблица 3:

a	b	N	D_N	$D_N\sqrt{N}$	x^*	$F(x^*)$	$F_N(x^*)$	$F_N(x^* - 0)$
-1.51	4.59	200	0.06125	0.86614	-0.70706	0.13125	0.075	0.1105

Из таблицы 3 значение выборочного критерия $D_N\sqrt{N} = 0.86614$

4 Анализ результатов и выводы

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

l	4	5	6	7	8
$\chi^2_{kp,\alpha}(l)$	$9,5$	$11,1$	$12,6$	$14,1$	$15,5$

Таблица критических значений распределения Колмогорова k_α

α	$0,01$	$0,02$	$0,05$	$0,1$	$0,2$
k_α	$1,63$	$1,57$	$1,36$	$1,22$	$1,07$

Задание I.

Гипотеза о соответствии выборки нормальному распределению $N(\widetilde{a}, \widetilde{\sigma}^2)$ не противоречит экспериментальным данным (может быть принята) при уровне значимости $\alpha = 0,05$, так как $\chi_B = 3.36049 < \chi^2_{kp,\alpha}(l) = 11,1$ при $l = m - 3 = 5$

Задание II.

Гипотеза о соответствии выборки равномерному распределению на отрезке [-1.50591, 4.58079] не противоречит экспериментальным данным (может быть принята) при уровне значимости $\alpha = 0,05$, так как $\chi_B = 11.04 < \chi^2_{kp,\alpha}(l) = 11,1$ при $l = m - 1 = 7$

Задание III.

Гипотеза о соответствии выборки равномерному распределению на отрезке [2.54, 8.84] не противоречит экспериментальным данным (может быть принята) при уровне значимости $\alpha = 0,05$, так как $D_N\sqrt{N}0.86614 < k_\alpha = 1,36$.

5 Список использованной литературы

1. Математическая статистика [Электронный ресурс]: метод. указания по выполнению лаб. работ / А.А. Лобузов — М.: МИРЭА, 2017.

2. Боровков А. А. Математическая статистика. — СПб.: Лань, 2010.-704 с.

3. Гмурман В.Е. Теория вероятностей и математическая статистика. — М.: Юрайт, 2013. — 479 с.

4. Гмурман В.Е. Руководство к решению задач по теории вероятностей и математической статистике. — М.: Юрайт, 2013. — 404 с.

5. Емельянов Г.В.Скитович В.П. Задачник по теории вероятностей и математической статистике. — СПб.: Лань, 2007. — 336 с.

6. Ивченко Г. И., Медведев Ю. И. Введение в математическую статистику. — М.: Изд-во ЛКИ, 2010. — 599 с.

7. Кибзун А.И., Горяинова Е.Р., Наумов А.В. Теория вероятностей и математическая статистика. Базовый курс с примерами и задачам. Учебное пособие — М.:ФИЗМАТЛИТ, 2005. — 232 с.

8. Кобзарь А.И. Прикладная математическая статистика: Для инженеров и научных работников — М.: ФИЗМАТЛИТ, 2006. — 816 с.

9. Монсик В.Б., Скрынников А. А. Вероятность и статистика.— М. : БИНОМ, 2015 — 384 с.

10. Сборник задач по теории вероятностей, математической статистике и теории случайных функций: Учеб. пособие для вузов / Под ред. А. А. Свешникова. — СПб.: Лань, 2012. — 472 с.

11. Письменный Д.Т. Конспект лекций по теории вероятностей, математической статистике и случайным процессам: учеб. пособие для вузов. — М.: Айрис-пресс, 2013. — 288 с.

12. Ramachandran Kandethody M., Tsokos Chris P. Mathematical Statistics with Applications in R. — N-Y.: Academic Press, 2009. — 826 p.

13. Ивановский Р.И. Теория вероятностей и математическая статистика: Основы, прикладные аспекты с примерами и задачами в среде Mathcad:Учеб. пособие для вузов — СПб.: БХВ-Петербург, 2008. — 528 с.

```
import matplotlib.pyplot as plt
import numpy as np
import math
import plotly.graph_objects as go
import copy
import xlsxwriter
import pandas as pd

df=pd.read_excel("first_selection.xlsx","Лист1")
df_2=pd.read_excel("second_selection.xlsx","Лист1")
Selection_1_unsort = []
Selection_2_unsort = []
for i in range(len(df.columns)):
    Selection_1_unsort.extend(df[df.columns[i]].tolist())
    Selection_2_unsort.extend(df_2[df_2.columns[i]].tolist())

print(df.columns[0])
print(len(Selection_2_unsort))

Selection_1 = copy.deepcopy(Selection_1_unsort)
Selection_1.sort()
Selection_2 = copy.deepcopy(Selection_2_unsort)
Selection_2.sort()

N=len(Selection_1_unsort)
m=1+int(math.log(N,2))
n_0=min(Selection_1_unsort)
n_m=max(Selection_1_unsort)
d_n=n_m-n_0
a_n=[n_0]
for i in range(m-1):
    a_n.append(a_n[i]+d_n/m)
a_n.append(n_m)

print(len(a_n))
for i in range(len(a_n)):
    print(a_n[i])

n_n=[]
counter = 0
k=1
helper = a_n[k]
for i in range(len(Selection_1)):
    if Selection_1[i]<=helper:
        counter += 1
    else:
        n_n.append(counter)
        k += 1
        helper = a_n[k]
        if (Selection_1[i]<=helper):
            counter=1
        else: counter=0
n_n.append(counter)

w_n=[]
for i in range(len(n_n)):
    w_n.append(n_n[i]/N)
```

```
x_n = []
for i in range(len(a_n)-1):
    x_n.append((a_n[i]+a_n[i+1])/2)

math_exp_n=0
for i in range(len(x_n)):
    math_exp_n+=(x_n[i]*w_n[i])
print(math_exp_n)

h_n=(n_m-n_0)/m

Disp_n=0
for i in range(len(x_n)):
    Disp_n+=w_n[i]*((x_n[i])**2)
Disp_n-=((h_n**2)/12+math_exp_n**2)
print(Disp_n)

stand_dev_n=math.sqrt(Disp_n)

print(stand_dev_n)

#функция распределения
def normal_prob(x, miu, stdev):
    return 0.5 * (1 + math.erf((x-miu)/(stdev * 2**0.5)))

def c_norm_prob(x, miu, stdev):
    return (x-miu)/(stdev)

def density_normal_prob_t(t):
    return math.exp(-(t**2)/2)/math.sqrt(2*math.pi)

P_n = []
C_n_prob = []
n_prob = []
F_n_prob = []
P_n.append(" - ")
for i in range (2):
    C_n_prob.append(c_norm_prob(a_n[i], math_exp_n,stand_dev_n))
    n_prob.append(density_normal_prob_t(C_n_prob[i])/stand_dev_n)
    F_n_prob.append(normal_prob(a_n[i], math_exp_n,stand_dev_n))

P_n.append(F_n_prob[1])
for i in range(2,len(a_n)-1):
    C_n_prob.append(c_norm_prob(a_n[i], math_exp_n,stand_dev_n))
    n_prob.append(density_normal_prob_t(C_n_prob[i])/stand_dev_n)
    F_n_prob.append(normal_prob(a_n[i], math_exp_n,stand_dev_n))
    P_n.append(F_n_prob[i]-F_n_prob[i-1])

C_n_prob.append(c_norm_prob(a_n[len(a_n)-1], math_exp_n,stand_dev_n))
n_prob.append(density_normal_prob_t(C_n_prob[len(a_n)-1])/stand_dev_n)
F_n_prob.append(normal_prob(a_n[len(a_n)-1], math_exp_n,stand_dev_n))
P_n.append(1-F_n_prob[len(a_n)-2])

k=list(range(0, m+1))

P_n_h = []
```

```
C_n_prob_h = []
n_prob_h = []
F_n_prob_h = []
a_n_h=[]

P_n_h.append(P_n[0])
C_n_prob_h.append(round(C_n_prob[0],5))
n_prob_h.append(round(n_prob[0],5))
F_n_prob_h.append(round(F_n_prob[0],5))
a_n_h.append(round(a_n[0],5))

for i in range(1,len(P_n)):
    P_n_h.append(round(P_n[i],5))
    C_n_prob_h.append(round(C_n_prob[i],5))
    n_prob_h.append(round(n_prob[i],5))
    F_n_prob_h.append(round(F_n_prob[i],5))
    a_n_h.append(round(a_n[i],5))

a_n_h.append(' ')
C_n_prob_h.append(' ')
n_prob_h.append(' ')
F_n_prob_h.append(' ')
P_n_h.append(round(sum(P_n[1:len(P_n)]),5))

fig = go.Figure(data=[go.Table(header=dict(values=["k", "a_k" , "(a_k-a)/sig", "fi((a_k-a)/sig)/sig",
"F((a_k-a)/sig)", "p_k" ]),
cells=dict(values=[k, a_n_h, C_n_prob_h, n_prob_h, F_n_prob_h, P_n_h ]))
])
fig.show()

i_n=[]
w_n_h=[]
abs_n_w_p_h = []
P_n_h_1=P_n_h[1:len(P_n_h)]
n_w_p=[]
k_h=[]

h="["+str(round(a_n[0],5))+" , "+str(round(a_n[1],5))+"]"
i_n.append(h)
for i in range(1,len(a_n)-1):
    i_n.append("("+str(round(a_n[i],5))+" , "+str(round(a_n[i+1],5))+"]")
i_n.append(' ')


for i in range(len(w_n)):
    w_n_h.append(round(w_n[i], 5))
w_n_h.append(sum(w_n))

for i in range(len(w_n)):
    abs_n_w_p_h.append(round((abs(w_n[i]-P_n_h_1[i])), 5))
abs_n_w_p_h.append(max(abs_n_w_p_h))

for i in range(len(w_n)):
    n_w_p.append(round(N*((w_n[i]-P_n_h_1[i])**2)/P_n_h_1[i],5))
n_w_p.append(sum(n_w_p))

k_h=k[1:len(k)]
```

```
fig = go.Figure(data=[go.Table(header=dict(values=["k", "Интервалы", "w_k" , "p_k", "|w_k - p_k|" ,
"N(w_k-p_k)^2/p_k"]),
cells=dict(values=[k_h, i_n, w_n_h, P_n_h_1, abs_n_w_p_h, n_w_p]))
])
fig.show()

density_x_n = np.linspace(n_0, n_m, 200)
density_y_n = []
for i in range(len(density_x_n)):

density_y_n.append(density_normal_prob_t(c_norm_prob(density_x_n[i],math_exp_n,stand_dev_n)))

fig = plt.figure(figsize=(20, 10))
plt.hist(Selection_1,bins=m, density=True)
plt.plot( a_n, n_prob,'r', linewidth=1)
plt.title("Гистограмма распределения первой выборки и график плотности нормального
распределения",fontsize =15)
plt.yticks(np.arange(0.0, 0.5, step=0.025), fontsize =12)
plt.xticks(np.arange(round(Selection_1[0],1)-0.1, round(Selection_1[len(Selection_1)-1]+0.2,1),
step=0.2), fontsize =12)
plt.savefig('selection_1_hist_t.jpg')

fig = plt.figure(figsize=(20, 10))
plt.hist(Selection_1,bins=m, density=True)
plt.plot( density_x_n, density_y_n,'r', linewidth=1)
plt.title("Гистограмма распределения первой выборки и график плотности нормального
распределения",fontsize =15)
plt.yticks(np.arange(0.0, 0.5, step=0.025), fontsize =12)
plt.xticks(np.arange(round(Selection_1[0],1)-0.1, round(Selection_1[len(Selection_1)-1]+0.2,1),
step=0.2), fontsize =12)
plt.savefig('selection_1_hist.jpg')

x_n_h =[]
n_n_h =[]

for i in range(len(x_n)):
    x_n_h.append(round(x_n[i],5))
x_n_h.append(' ')

for i in range(len(n_n)):
    n_n_h.append(round(n_n[i], 5))
n_n_h.append(sum(n_n))

fig = go.Figure(data=[go.Table(header=dict(values=["Интервалы", "x*_k", "n_k" , "w_k"]),
cells=dict(values=[i_n, x_n, n_n_h, w_n_h]))
])
fig.show()

N_2=len(Selection_2)
m_2=1+int(math.log(N,2))
u_0=min(Selection_2)
u_m=max(Selection_2)
d_u=u_m-u_0
a_u=[u_0]
for i in range(m-1):
    a_u.append(a_u[i]+d_u/m_2)
a_u.append(u_m)
```

```
n_u=[]
counter = 0
k=1
helper = a_u[k]
for i in range(len(Selection_2)):
    if Selection_2[i]<=helper:
        counter += 1
    else:
        n_u.append(counter)
        k += 1
        helper = a_u[k]
        if (Selection_2[i]<=helper):
            counter=1
        else: counter=0
n_u.append(counter)

w_u=[]
for i in range(len(n_u)):
    w_u.append(n_u[i]/N_2)

x_u = []
for i in range(len(a_u)-1):
    x_u.append((a_u[i]+a_u[i+1])/2)

math_exp_u=0
for i in range(len(x_u)):
    math_exp_u+=(x_u[i]*w_u[i])
print(math_exp_u)

h_u=(u_m-u_0)/m_2

Disp_u=0
for i in range(len(x_u)):
    Disp_u+=w_u[i]*((x_u[i])**2)
Disp_u-=((h_u**2)/12+math_exp_u**2)
print(Disp_u)

stand_dev_u=math.sqrt(Disp_u)

n_u_h =[]

for i in range(len(n_u)):
    n_u_h.append(round(n_u[i], 5))
n_u_h.append(sum(n_u))

u_w_p=[]
p_k_u=1/m_2

for i in range(len(w_u)):
    u_w_p.append(round(N*((w_u[i]-p_k_u)**2)/p_k_u,5))
u_w_p.append(sum(u_w_p))

k_u=list(range(1, m+1))

i_u = []
w_u_h = []
abs_u_w_p_h = []
```

```
h="["+str(round(a_u[0],5))+" , "+str(round(a_u[1],5))+"]"
i_u.append(h)
for i in range(1,len(a_u)-1):
    i_u.append("("+str(round(a_u[i],5))+" , "+str(round(a_u[i+1],5))+"]")
i_u.append(' ')

for i in range(len(w_u)):
    w_u_h.append(round(w_u[i], 5))
w_u_h.append(sum(w_u))

for i in range(len(w_u)):
    abs_u_w_p_h.append(round((abs(w_n[i]-p_k_u)), 5))
abs_u_w_p_h.append(max(abs_u_w_p_h))

P_u_h=[p_k_u for i in range(m_2)]
P_u_h.append(sum(P_u_h))

fig = go.Figure(data=[go.Table(header=dict(values=["Интервалы", "x*_k", "n_k" , "w_k"]),
cells=dict(values=[i_u, x_u, n_u_h, w_u_h]))
])
fig.show()

fig = go.Figure(data=[go.Table(header=dict(values=["k", "Интервалы", "w_k" , "p_k", "|w_k - p_k|" ,
"N(w_k-p_k)^2/p_k"]),
cells=dict(values=[k_u, i_u, w_u_h, P_u_h, abs_u_w_p_h, u_w_p]))
])
fig.show()

h_x=[u_0, u_m]
h_y=[1/(u_m-u_0), 1/(u_m-u_0)]

fig = plt.figure(figsize=(20, 10))
plt.hist(Selection_2,bins=m, density=True)
plt.plot( h_x, h_y,'r', linewidth=1)
plt.title("Гистограмма распределения первой выборки и график плотности нормального
распределения",fontsize =15)
plt.yticks(np.arange(0.0, 0.25, step=0.025), fontsize =12)
plt.xticks(np.arange(round(Selection_2[0],1)-0.1, round(Selection_2[len(Selection_2)-1]+0.2,1),
step=0.2), fontsize =12)
plt.savefig('selection_2_hist.jpg')

S=[1/N_2]
now_h= 1/(u_m-u_0)
for i in range(1,N_2):
    S.append(S[i-1]+(1/N_2))

x_2=np.linspace(u_0, u_m,200)
y_2 =[]
for i in range(len(x_2)):
    y_2.append((x_2[i]-u_0)*now_h)

fig = plt.figure(figsize=(20, 10))
plt.hlines(0, Selection_2[0]-0.05, Selection_2[0])
for i in range(len(S)-1):
    plt.hlines(S[i], Selection_2[i], Selection_2[i+1])
plt.hlines(S[i+1], Selection_2[i+1], Selection_2[i+1]+0.05)
```

```
plt.plot(x_2, y_2, color = 'r')
plt.title('Эмпирическая функция выборочного распределения',fontsize =15)
plt.ylabel('s - sum(w(x)) ', fontsize =15)
plt.yticks(np.arange(0.0, 1.05, step=0.05), fontsize =12)
plt.xticks(np.arange(round(Selection_2[0],1), round(Selection_2[len(Selection_2)-1]+0.2,1), step=0.2),
fontsize =12)
plt.xlabel('x - value', fontsize =15)
plt.savefig('uniform_emp.jpg')

matrix_1_s=[[0] * 20 for i in range(10)]
matrix_2_s=[[0] * 20 for i in range(10)]
for i in range(10):
    for j in range(20):
        matrix_1_s[i][j]=round(Selection_1[i+j*10],5)
        matrix_2_s[i][j]=round(Selection_2[i+j*10],5)

workbook = xlsxwriter.Workbook('Selection_1_sort.xlsx')
worksheet = workbook.add_worksheet()

row = 0

for col, data in enumerate(matrix_1_s):
    worksheet.write_column(row, col, data)

workbook.close()

workbook = xlsxwriter.Workbook('Selection_2_sort.xlsx')
worksheet = workbook.add_worksheet()

row = 0

for col, data in enumerate(matrix_2_s):
    worksheet.write_column(row, col, data)

workbook.close()

k_h_h=copy.deepcopy(k_h)
k_h_h.append(" ")
k_u_h=copy.deepcopy(k_u)
k_u_h.append(" ")

matrix_1_t=[[0] * 9 for i in range(6)]
matrix_2_t=[[0] * 9 for i in range(6)]
for i in range(9):
        matrix_1_t[0][i] = k_h_h[i]
        matrix_1_t[1][i] = i_n[i]
        matrix_1_t[2][i] = w_n_h[i]
        matrix_1_t[3][i] = P_n_h_1[i]
        matrix_1_t[4][i] = abs_n_w_p_h[i]
        matrix_1_t[5][i] = n_w_p[i]
        matrix_2_t[0][i] = k_u_h[i]
        matrix_2_t[1][i] = i_u[i]
        matrix_2_t[2][i] = w_u_h[i]
        matrix_2_t[3][i] = P_u_h[i]
        matrix_2_t[4][i] = abs_u_w_p_h[i]
        matrix_2_t[5][i] = u_w_p[i]
```

```
workbook = xlsxwriter.Workbook('1_table_1_2.xlsx')
worksheet = workbook.add_worksheet()

row = 0

for col, data in enumerate(matrix_1_t):
    worksheet.write_column(row, col, data)

workbook.close()

k=list(range(0, m+1))
k.append(" ")

matrix_1_t_1=[[0] * 10 for i in range(6)]
for i in range(10):
        matrix_1_t_1[0][i] = k[i]
        matrix_1_t_1[1][i] = a_n_h[i]
        matrix_1_t_1[2][i] = C_n_prob_h[i]
        matrix_1_t_1[3][i] = n_prob_h[i]
        matrix_1_t_1[4][i] = F_n_prob_h[i]
        matrix_1_t_1[5][i] = P_n_h[i]

workbook = xlsxwriter.Workbook('1_table_1_1.xlsx')
worksheet = workbook.add_worksheet()

row = 0

for col, data in enumerate(matrix_1_t_1):
    worksheet.write_column(row, col, data)

workbook.close()

workbook = xlsxwriter.Workbook('2_table_2_1.xlsx')
worksheet = workbook.add_worksheet()

row = 0

for col, data in enumerate(matrix_2_t):
    worksheet.write_column(row, col, data)

workbook.close()

def helper_func( x, u_0, u_m):
    return ((x-u_0)/(u_m-u_0))

x_j = abs(helper_func(Selection_2[0],u_0, u_m)-S[0])
index_j = 0
for i in range(1,len(S)):
    max_now_f_0=max(abs(helper_func(Selection_2[i],u_0, u_m)-S[i-1]),
abs(helper_func(Selection_2[i],u_0, u_m)-S[i]))
    if(x_j<max_now_f_0):
        x_j = max_now_f_0
        index_j = i

index_j

Selection_2[index_j]
```

```
S[index_j]

helper_func(Selection_2[index_j],u_0, u_m)

helper_func(Selection_2[index_j-1],u_0, u_m)

x_j

x_j*math.sqrt(200)
```