



МИНОБРНАУКИ РОССИИ

Федеральное государственное бюджетное образовательное
учреждение высшего образования
«МИРЭА – Российский технологический университет»
РТУ МИРЭА

Институт кибернетики
Кафедра информационной безопасности

Долгосрочная работа №1
по дисциплине
«Математическая статистика»

Студент: Никишина А.А., ККСО-03-19

Преподаватель: Пастухова Ю.И.

Москва 2021

Содержание

1. Описание данных	4
2. Вариационный ряд	4
3. Характеристики вариационного ряда	4
3.1. Мода	4
3.2. Медиана	5
3.3. Выборочная средняя	5
3.4. Выборочная дисперсия	5
3.5. Среднее квадратическое отклонение	5
4. Полигон	5
5. Группировка данных	6
6. Гистограмма	7
7. Выборочный коэффициент асимметрии и эксцесса	8
7.1. Коэффициент асимметрии	8
7.2. Коэффициент эксцесса	8
8. Характеристики интервального ряда	9
8.1. Выборочная средняя	9
8.2. Выборочная дисперсия	9
8.3. Среднее квадратическое отклонение	9
9. Проверка 2 гипотез критерием Пирсона	9
9.1. Нормальное распределение.	9
9.2. Равномерное распределение	11

10	Критерий Колмогорова	12
10.1	Нормальное распределение	12
10.2	Равномерное распределение	13
11	Доверительные интервалы	14
11.1	Интервал для матожидания при известной $\sigma_{\text{в}}$	14
11.2	Интервал для матожидания при неизвестной $\sigma_{\text{в}}$. . .	15
11.3	Интервал для дисперсии	15
12	Выводы	16

1. Описание данных

В работе был измерен ИМТ (индекс массы тела человека) $n=58$ людей имеющих гипертанию и болезни сердца. Самый низкий ИМТ из выборки составил 27, самый высокий 46 . Вариационный ряд: x_i варианты, количество человек имеющих данный рост n_i , и относительная частота w_i

2. Вариационный ряд

$1 \leq i \leq 8$	1	2	3	4	5	6	7	8
x_i	21	22	24	25	26	27	28	29
n_i	2	1	3	1	3	6	5	2
w_i	0,03	0,02	0,05	0,02	0,05	0,10	0,09	0,03
$9 \leq i \leq 16$	9	10	11	12	13	14	15	16
x_i	30	31	32	33	34	35	36	37
n_i	5	5	4	3	2	4	1	3
w_i	0,09	0,09	0,07	0,05	0,03	0,07	0,02	0,05
$17 \leq i \leq 23$	17	18	19	20	21	22	23	\sum_j
x_i	38	40	41	42	44	45	46	
n_i	1	1	1	2	1	1	1	58
w_i	0,02	0,02	0,02	0,03	0,02	0,02	0,02	1

Таблица 1 – Вариационный ряд

3. Характеристики вариационного ряда

3.1. Мода

Мода дискретного вариационного ряда – это варианта с максимальной частотой.

Модой данного ряда является значение 27, т. к. частота этого значения максимальна;

$$M_0 = 27$$

3.2. Медиана

Медиана вариационного ряда – это значение, которая делит его на две равные части (по количеству вариантов).

Медианой является значение 33 т. к. это значение делит выборку на 2 равные части (одна часть выборки не больше 33, другая не меньше 33).

$$M_e = 33$$

3.3. Выборочная средняя

Выборочной средней называется среднее арифметическое всех значений выборки:

$$\bar{x}_в = \frac{1}{n} \sum_i (x_i \cdot n_i) = \frac{1}{58} \sum_1^{23} (x_i \cdot n_i) = \frac{1}{58} \cdot (21 \cdot 2 + 22 \cdot 1 + 24 \cdot 3 + 25 \cdot 1 + 26 \cdot 3 + 27 \cdot 6 + 28 \cdot 5 + 29 \cdot 2 + 30 \cdot 5 + 31 \cdot 5 + 32 \cdot 4 + 33 \cdot 3 + 34 \cdot 2 + 35 \cdot 4 + 36 \cdot 1 + 37 \cdot 3 + 38 \cdot 1 + 40 \cdot 1 + 41 \cdot 1 + 42 \cdot 2 + 44 \cdot 1 + 45 \cdot 1 + 46 \cdot 1) = 31,45$$

3.4. Выборочная дисперсия

Выборочная дисперсия – это среднее арифметическое квадратов отклонений всех вариантов выборки от её средней:

$$D_в = \frac{1}{n} \sum_i (x_i - \bar{x}_в)^2 \cdot n_i = \frac{1}{n} \sum_1^{23} (x_i^2 \cdot n_i) - \bar{x}_в^2 = \frac{1}{58} (21^2 \cdot 2 + 22^2 \cdot 1 + 24^2 \cdot 3 + 25^2 \cdot 1 + 26^2 \cdot 3 + 27^2 \cdot 6 + 28^2 \cdot 5 + 29^2 \cdot 2 + 30^2 \cdot 5 + 31^2 \cdot 5 + 32^2 \cdot 4 + 33^2 \cdot 3 + 34^2 \cdot 2 + 35^2 \cdot 4 + 36^2 \cdot 1 + 37^2 \cdot 3 + 38^2 \cdot 1 + 40^2 \cdot 1 + 41^2 \cdot 1 + 42^2 \cdot 2 + 44^2 \cdot 1 + 45^2 \cdot 1 + 46^2 \cdot 1) - 31,45^2 = 34,04$$

3.5. Среднее квадратическое отклонение

Среднее квадратическое отклонение характеризует разброс значений относительно среднего выборочного

$$\sigma_в = \sqrt{D_в} = \sqrt{34,04} = 5,8344$$

4. Полигон

Полигон – это ломаная, соединяющая точки $(x_1, n_1), (x_2, n_2), \dots, (x_k, n_k)$, где x_i – это варианты или наблюдаемые значения, а n_i – частота вариантов.

Полигон (относительных частот) $(x_1, w_1), (x_2, w_2), \dots, (x_k, w_k)$ это ломаная, соединяющая точки, где x_i – это варианты или наблюдаемые значения, а w_i – относительная частота вариантов.

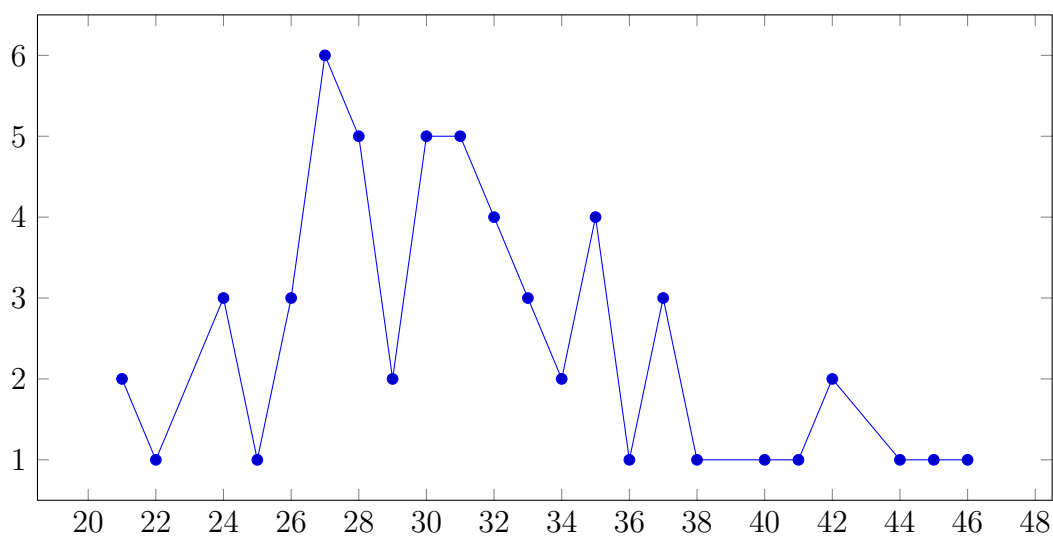


Рисунок 1 – Полигон с частотами вариантов

Полигон относительных частот выглядит точно так же, разница только в том что в оси у будут находится относительные частоты (и у графика будет другой масштаб).

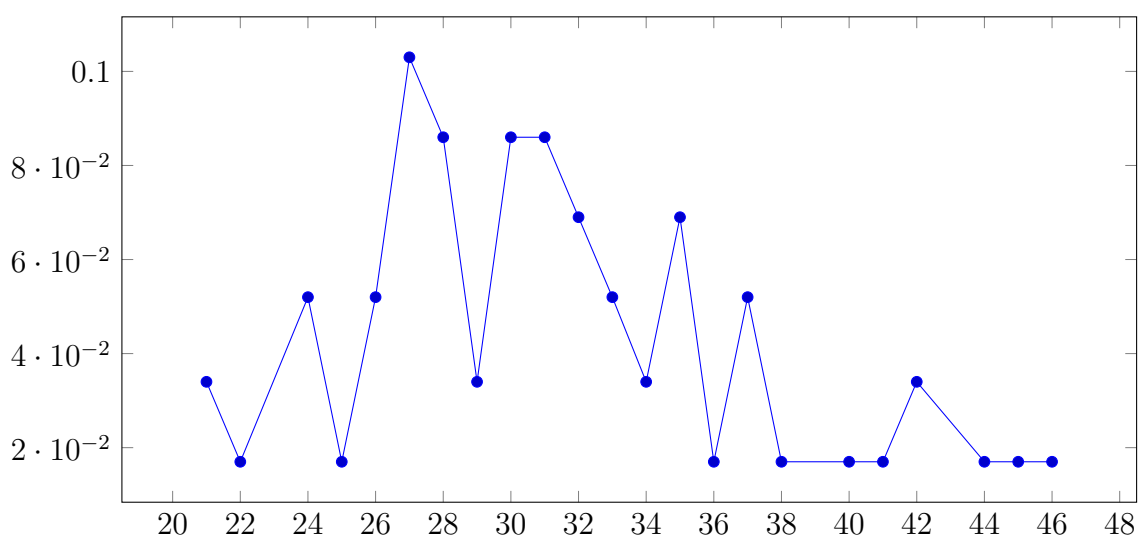


Рисунок 2 – Полигон относительных частот

5. Группировка данных

Сгруппируем данные по 9 интервалам длиной 3.

i	Интервал	Середина интервала	n_i	w_i
1	[19, 22]	20,5	3	0.05
2	[22, 25]	23,5	4	0.07
3	[25, 28]	26,5	14	0,24
4	[28, 31]	29,5	12	0,21
5	[31, 34]	32,5	9	0,16
6	[34, 37]	35,5	8	0,14
7	[37, 40]	38,5	2	0,03
8	[40, 43]	41,5	3	0,05
9	[43, 46]	44,5	3	0,05

Таблица 2 – Интервальный вариационный ряд

6. Гистограмма

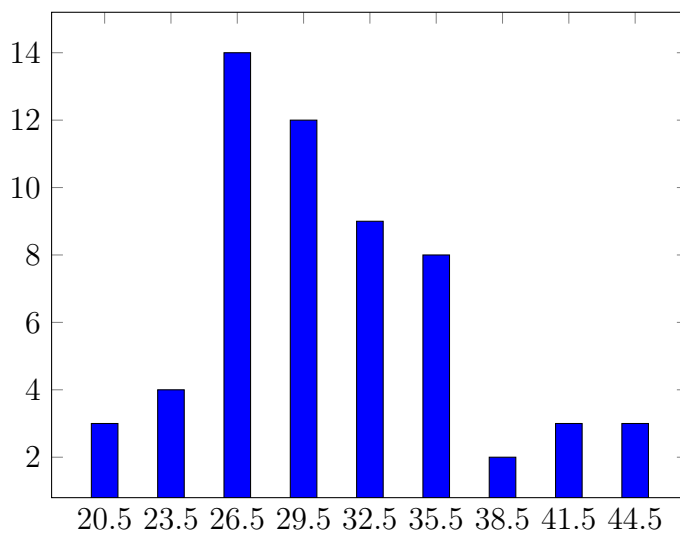


Рисунок 3 – Гистограмма

Судя по виду гистограммы мы можем сделать предварительный вывод: распределение скорее всего нормальное.

7. Выборочный коэффициент асимметрии и эксцесса

7.1. Коэффициент асимметрии

Коэффициент асимметрии характеризует скошенность распределения по отношению к математическому ожиданию. Асимметрия положительна, если «длинная часть» кривой распределения расположена справа от математического ожидания; асимметрия отрицательна, если «длинная часть» кривой расположена слева от математического ожидания.

$A_s = \frac{m_3}{\sigma_B^3}$, где $m_k = \frac{\sum_i (x_i - \bar{x}_B)^k}{n}$ - центральный эмпирический момент третьего порядка, куб стандартного выборочного отклонения.

$$A_s = \frac{1}{58} \cdot \frac{1}{5.8344^3} \cdot ((21 - 31, 45)^3 + (22 - 31, 45)^3 + (24 - 31, 45)^3 + (25 - 31, 45)^3 + (26 - 31, 45)^3 + (27 - 31, 45)^3 + (28 - 31, 45)^3 + (29 - 31, 45)^3 + (30 - 31, 45)^3 + (31 - 31, 45)^3 + (32 - 31, 45)^3 + (33 - 31, 45)^3 + (34 - 31, 45)^3 + (35 - 31, 45)^3 + (36 - 31, 45)^3 + (37 - 31, 45)^3 + (38 - 31, 45)^3 + (40 - 31, 45)^3 + (41 - 31, 45)^3 + (42 - 31, 45)^3 + (44 - 31, 45)^3 + (45 - 31, 45)^3 + (46 - 31, 45)^3) = 0.68$$

7.2. Коэффициент эксцесса

Коэффициент эксцесса — число характеризующее степень остроты пика распределения случайной величины.

$$E_k = \frac{m_4}{\sigma_B^4} - 3$$
$$E_k = \frac{1}{58} \cdot \frac{1}{5.8344^4} \cdot ((21 - 31, 45)^4 + (22 - 31, 45)^4 + (24 - 31, 45)^4 + (25 - 31, 45)^4 + (26 - 31, 45)^4 + (27 - 31, 45)^4 + (28 - 31, 45)^4 + (29 - 31, 45)^4 + (30 - 31, 45)^4 + (31 - 31, 45)^4 + (32 - 31, 45)^4 + (33 - 31, 45)^4 + (34 - 31, 45)^4 + (35 - 31, 45)^4 + (36 - 31, 45)^4 + (37 - 31, 45)^4 + (38 - 31, 45)^4 + (40 - 31, 45)^4 + (41 - 31, 45)^4 + (42 - 31, 45)^4 + (44 - 31, 45)^4 + (45 - 31, 45)^4 + (46 - 31, 45)^4) - 3 = -0.634$$

8. Характеристики интервального ряда

8.1. Выборочная средняя

$$\bar{x}_B = \frac{1}{n} \sum_{(i)} (x_i \cdot n_i)$$
$$\bar{x}_B = \frac{1}{58} (20,5 \cdot 3 + 23,5 \cdot 4 + 26,5 \cdot 14 + 29,5 \cdot 12 + 32,5 \cdot 9 + 35,5 \cdot 8 + 38,5 \cdot 2 + 41,5 \cdot 3 + 44,5 \cdot 3) = 30,90$$

8.2. Выборочная дисперсия

$$D_B = \frac{1}{n} \sum_{(i)} (x_i - \bar{x}_B)^2 \cdot n_i = \frac{1}{n} \sum_{(i)} (x_i^2 \cdot n_i) - \bar{x}_B^2$$
$$D_B = \frac{1}{58} \cdot (20,5^2 \cdot 3 + 23,5^2 \cdot 4 + 26,5^2 \cdot 14 + 29,5^2 \cdot 12 + 32,5^2 \cdot 9 + 35,5^2 \cdot 8 + 38,5^2 \cdot 2 + 41,5^2 \cdot 3 + 44,5^2 \cdot 3) - 30,90^2 = 35,14$$

8.3. Среднее квадратическое отклонение

$$\sigma_B = \sqrt{D_B} = \sqrt{35,14} = 5,9276$$

9. Проверка 2 гипотез критерием Пирсона

Предположим что распределение нормальное.

9.1. Нормальное распределение.

Гипотеза: генеральная совокупность имеет нормальное распределение.

Посчитаем теоретическую вероятность попадания данных в каждый интервал:

$$p_i = \Phi\left(\frac{\beta_i - \bar{x}_B}{\sigma}\right) - \Phi\left(\frac{\alpha_i - \bar{x}_B}{\sigma}\right)$$
$$p_1 = \Phi\left(\frac{22 - 30,90}{5,9276}\right) - \Phi\left(\frac{19 - 30,90}{5,9276}\right) = -0,4332 + 0,4778 = 0,0446$$
$$p_2 = \Phi\left(\frac{25 - 30,90}{5,9276}\right) - \Phi\left(\frac{22 - 30,90}{5,9276}\right) = -0,3389 + 0,4332 = 0,0943$$
$$p_3 = \Phi\left(\frac{28 - 30,90}{5,9276}\right) - \Phi\left(\frac{25 - 30,90}{5,9276}\right) = -0,1879 + 0,3389 = 0,1510$$
$$p_4 = \Phi\left(\frac{31 - 30,90}{5,9276}\right) - \Phi\left(\frac{28 - 30,90}{5,9276}\right) = 0,0080 + 0,1879 = 0,1959$$
$$p_5 = \Phi\left(\frac{34 - 30,90}{5,9276}\right) - \Phi\left(\frac{31 - 30,90}{5,9276}\right) = 0,1985 - 0,0080 = 0,1905$$

$$\begin{aligned}
p_6 &= \Phi\left(\frac{37 - 30,90}{5,9276}\right) - \Phi\left(\frac{34 - 30,90}{5,9276}\right) = 0,3485 - 0,1985 = 0,1500 \\
p_7 &= \Phi\left(\frac{40 - 30,90}{5,9276}\right) - \Phi\left(\frac{37 - 30,90}{5,9276}\right) = 0,4382 - 0,3485 = 0,0897 \\
p_8 &= \Phi\left(\frac{43 - 30,90}{5,9276}\right) - \Phi\left(\frac{40 - 30,90}{5,9276}\right) = 0,4793 - 0,4382 = 0,0411 \\
p_9 &= \Phi\left(\frac{46 - 30,90}{5,9276}\right) - \Phi\left(\frac{43 - 30,90}{5,9276}\right) = 0,4946 - 0,4793 = 0,0153
\end{aligned}$$

i	$I_i = [\alpha, \beta]$	n_i	$\Phi\left(\frac{\alpha_i - \bar{x}_B}{\sigma}\right)$	$\Phi\left(\frac{\beta_i - \bar{x}_B}{\sigma}\right)$	p_i	$m_i = n \cdot p_i$	$(n_i - m_i)^2$	$\frac{(n_i - m_i)^2}{m_i}$
1	[19, 22]	3	-0,4778	-0,4332	0,0446	2,5868	0,1707	0,0660
2	[22, 25]	4	-0,4332	-0,3389	0,0943	5,4694	2,1591	0,3948
3	[25, 28]	14	-0,3389	-0,1879	0,1510	8,7580	27,4786	3,1375
4	[28, 31]	12	-0,1879	0,0080	0,1959	11,3622	0,4068	0,0358
5	[31, 34]	9	0,0080	0,1985	0,1905	11,0490	4,1984	0,3800
6	[34, 37]	8	0,1985	0,3485	0,1500	8,7000	0,4900	0,0563
7	[37, 40]	2	0,3485	0,4382	0,0897	5,2026	10,2566	1,9714
8	[40, 43]	3	0,4382	0,4793	0,0411	2,3838	0,3797	0,1593
9	[43, 46]	3	0,4793	0,4946	0,0153	0,8874	4,4631	5,0294
Σ		58				56,3992	50,0031	11,2305

Таблица 3

При $\alpha = 0.05$

$df = k - 1 - r = 9 - 1 - 2 = 6$, где k количество интервалов, r количество параметров распределения (для нормального и равномерного распределений $r = 2$)

$$\chi_{\text{табл}}^2 = 12.592$$

$$\chi_{\text{набл}}^2 = \sum_i \frac{(n_i - m_i)^2}{m_i} = 11.2305$$

$\chi_{\text{табл}}^2 > \chi_{\text{набл}}^2$ значит можно считать, что генеральная совокупность имеет нормальное распределение и мы принимаем гипотезу о том, что генеральная совокупность имеет нормальное распределение.

Проверим гипотезу о равномерном распределении.

9.2. Равномерное распределение

$$\bar{x}_B = 30,90; D_B = 35,14.$$

$$\left\{ \begin{array}{l} \frac{a+b}{2} = \bar{x}_B \\ \frac{(b-a)^2}{12} = D_B \\ f(x) = 0, x < a \\ f(x) = \frac{1}{b-a}, a \leq x \leq b \\ f(x) = 0, x > b \end{array} \right. \quad \left\{ \begin{array}{l} a = 20,6326 \\ b = 41,1674 \\ f(x) = 0, x < a \\ f(x) = 0,0487, a \leq x \leq b \\ f(x) = 0, x > b \end{array} \right.$$

Посчитаем теоретическую вероятность попадания данных в каждый интервал:

$\int_{\alpha}^{\beta} f(x)dx$, где α, β — крайние точки интервала.

$$\left\{ \begin{array}{l} \hat{a} + \hat{b} = 2 \cdot \bar{x}_B \\ \hat{b} - \hat{a} = 2\sqrt{3} \cdot \sigma_B \end{array} \right. \quad \left\{ \begin{array}{l} \hat{a} = 20,6297 \\ \hat{b} = 24,0938 \end{array} \right.$$

$$\begin{aligned} p_1 &= \int_{19}^{22} f(x)dx = \int_{20,6297}^{22} f(x)dx = (22 - 20,6297) \cdot 20,6297 = 0,0667 \\ p_2 &= \int_{22}^{25} f(x)dx = \int_{22}^{25} f(x)dx = (25 - 22) \cdot 20,6297 = 0,1461 \\ p_3 &= \int_{25}^{28} f(x)dx = \int_{25}^{28} f(x)dx = (28 - 25) \cdot 20,6297 = 0,1461 \\ p_4 &= \int_{28}^{31} f(x)dx = \int_{28}^{31} f(x)dx = (31 - 28) \cdot 20,6297 = 0,1461 \\ p_5 &= \int_{31}^{34} f(x)dx = \int_{31}^{34} f(x)dx = (34 - 31) \cdot 20,6297 = 0,1461 \\ p_6 &= \int_{34}^{37} f(x)dx = \int_{34}^{37} f(x)dx = (37 - 34) \cdot 20,6297 = 0,1461 \\ p_7 &= \int_{37}^{40} f(x)dx = \int_{37}^{40} f(x)dx = (40 - 37) \cdot 20,6297 = 0,1461 \\ p_8 &= \int_{40}^{43} f(x)dx = \int_{40}^{43} f(x)dx = (43 - 40) \cdot 20,6297 = 0,1461 \\ p_9 &= \int_{43}^{46} f(x)dx = \int_{43}^{46} f(x)dx = (46 - 43) \cdot 20,6297 = 0,1461 \end{aligned}$$

i	$I_i = [\alpha, \beta]$	n_i	$p_i = \int_{\alpha}^{\beta} f(x)dx$	$m_i = n \cdot p_i$	$(n_i - m_i)^2$	$\frac{(n_i - m_i)^2}{m_i}$
1	[19, 22]	3	0,0667	3,8703	0,7574	0,1957
2	[22, 25]	4	0,1461	8,4734	20,0113	2,3617
3	[25, 28]	14	0,1461	8,4734	30,5433	3,6046
4	[28, 31]	12	0,1461	8,4734	12,4369	1,4678
5	[31, 34]	9	0,1461	8,4734	0,2773	0,0327
6	[34, 37]	8	0,1461	8,4734	0,2241	0,0264
7	[37, 40]	2	0,1461	8,4734	41,9049	4,9455
8	[40, 43]	3	0,1461	8,4734	29,9581	3,5355
9	[43, 46]	3	0,1461	8,4734	29,9581	3,5355
Σ		58		71,6575	166,0715	19,7055

Таблица 4

При $\alpha = 0.05$

$$df = k - 1 - r = 9 - 1 - 2 = 6$$

$$\chi^2_{\text{табл}} = 12.592$$

$$\chi^2_{\text{набл}} = \sum_i \frac{(n_i - m_i)^2}{m_i} = 19,7055$$

$$\chi^2_{\text{табл}} < \chi^2_{\text{набл}} \text{ значит гипотеза отвергается.}$$

10. Критерий Колмогорова

Функция распределения показывает вероятность того, что случайно выбранное из выборки значение будет меньше аргумента функции.

F_n – эмпирическая функция распределения. Для её построения будем использовать w_i , т.е. F_n – сумма частот на тех интервалах, которых существуют точки не большие чем x .

F^* – теоретические функции распределения, для их построения используем p_i , т.е. F^* – сумма вероятностей на тех интервалах, в которых существуют точки не большие чем x .

10.1. Нормальное распределение

Гипотеза: генеральная совокупность имеет нормальное распределение.

i	$I_i = [\alpha, \beta]$	n_i	w_i	$F_n(x)$	p_i	$F^*(x)$	$ F_n(x) - F^*(x) $
1	[19, 22]	3	0,0517	0,0517	0,0446	0,0446	0,0071
2	[22, 25]	4	0,0690	0,1207	0,0943	0,1389	0,0182
3	[25, 28]	14	0,2414	0,3621	0,1510	0,2899	0,0722
4	[28, 31]	12	0,2069	0,5690	0,1959	0,4858	0,0832
5	[31, 34]	9	0,1552	0,7241	0,1905	0,6763	0,0478
6	[34, 37]	8	0,1379	0,8621	0,1500	0,8263	0,0358
7	[37, 40]	2	0,0345	0,8966	0,0897	0,9160	0,0194
8	[40, 43]	3	0,0517	0,9483	0,0411	0,9571	0,0088
9	[43, 46]	3	0,0517	1,0000	0,0153	0,9724	0,0276
max							0,0832

Таблица 4

$$\lambda = \sqrt{n} \cdot \max |F_n - F^*| = 7,6158 \cdot 0,0832 = 0,6334$$

$P(\lambda = 0,6334) = 0,8222$, значит вероятность того, что гипотеза принята верно, равна 0,8222

10.2. Равномерное распределение

Гипотеза: генеральная совокупность имеет нормальное распределение.

i	$I_i = [\alpha, \beta]$	n_i	w_i	$F_n(x)$	p_i	$F^*(x)$	$ F_n(x) - F^*(x) $
1	[19, 22]	3	0,0517	0,0517	0,0667	0,0667	0,0150
2	[22, 25]	4	0,0690	0,1207	0,1461	0,2128	0,0921
3	[25, 28]	14	0,2414	0,3621	0,1461	0,3589	0,0032
4	[28, 31]	12	0,2069	0,5690	0,1461	0,5050	0,0640
5	[31, 34]	9	0,1552	0,7241	0,1461	0,6511	0,0730
6	[34, 37]	8	0,1379	0,8621	0,1461	0,7972	0,0649
7	[37, 40]	2	0,0345	0,8966	0,1461	0,9433	0,0467
8	[40, 43]	3	0,0517	0,9483	0,1461	1,0894	0,1411
9	[43, 46]	3	0,0517	1,0000	0,1461	1,2355	0,2355
max							0,2355

Таблица 5

$$\lambda = \sqrt{n} \cdot \max |F_n - F^*| = 7,6158 \cdot 0,2355 = 1,7933$$

$P(\lambda = 1,7933) = 0,0033$, значит вероятность того, что гипотеза принята верно, равна 0,0033

11. Доверительные интервалы

11.1. Интервал для матожидания при известной $\sigma_{\text{в}}$

Доверительный интервал для математического ожидания вычисляется по формуле:

$$\bar{x}_{\text{в}} - \frac{t\sigma}{\sqrt{n}} < MX < \bar{x}_{\text{в}} + \frac{t\sigma}{\sqrt{n}}$$

где $\bar{x}_{\text{в}}$ - выборочное среднее, n - размер выборки, $t = \Phi^{-1}(\frac{\gamma}{2})$, т.е. t удовлетворяет условию $\Phi(t) = (\frac{\gamma}{2})$

Для расчёта доверительного интервала по Стьюденту t берётся из таблицы Стьюдента при $df = n - 1 = 57$, $\alpha = 1 - \gamma$

При $\gamma = 0,95$

$$t = \Phi^{-1}(0,475) = 1,96$$

$$\bar{x}_{\text{в}} - \frac{t\sigma}{\sqrt{n}} = 30,90 - \frac{1,9600 * 5,9276}{\sqrt{58}} = 29,3710$$

$$\bar{x}_{\text{в}} + \frac{t\sigma}{\sqrt{n}} = 30,90 + \frac{1,9600 * 5,9276}{\sqrt{58}} = 32,4221$$

Значит доверительный интервал при $\gamma = 0,95$

$$29,3710 < MX < 32,4221$$

При $\gamma = 0,99$

$$t = \Phi^{-1}(0,475) = 2,58$$

$$\bar{x}_{\text{в}} - \frac{t\sigma}{\sqrt{n}} = 30,90 - \frac{2,5800 * 5,9276}{\sqrt{58}} = 28,8885$$

$$\bar{x}_{\text{в}} + \frac{t\sigma}{\sqrt{n}} = 30,90 + \frac{2,5800 * 5,9276}{\sqrt{58}} = 32,9046$$

Значит доверительный интервал при $\gamma = 0,99$

$$28,8885 < MX < 32,9046$$

11.2. Интервал для матожидания при неизвестной $\sigma_{\text{в}}$

При неизвестной $\sigma_{\text{в}}$ вместо неё используется:

$s = \sqrt{D_{\text{исправ}}} = \sqrt{\frac{n}{n-1} D_{\text{в}}} = 5,9793$ Для расчёта доверительного интервала по Стьюденту t берётся из таблицы Стьюдента при $df = n - 1 = 57, \alpha = 1\gamma$

При $\gamma = 0,95$

$$t = \Phi^{-1}(0,475) = 2,0034$$

$$\bar{x}_{\text{в}} - \frac{t\sigma}{\sqrt{n}} = 30,90 - \frac{2,0034 * 5,9793}{\sqrt{58}} = 29,3236$$

$$\bar{x}_{\text{в}} - \frac{t\sigma}{\sqrt{n}} = 30,90 + \frac{2,0034 * 5,9793}{\sqrt{58}} = 32,4695$$

Значит доверительный интервал при $\gamma = 0,95$

$$29,3236 < MX < 32,4695$$

При $\gamma = 0,99$

$$t = \Phi^{-1}(0,475) = 2,6672$$

$$\bar{x}_{\text{в}} - \frac{t\sigma}{\sqrt{n}} = 30,90 - \frac{2,6672 * 5,9793}{\sqrt{58}} = 28,8024$$

$$\bar{x}_{\text{в}} - \frac{t\sigma}{\sqrt{n}} = 30,90 + \frac{2,6672 * 5,9793}{\sqrt{58}} = 32,9907$$

Значит доверительный интервал при $\gamma = 0,99$

$$28,8024 < MX < 32,9907$$

11.3. Интервал для дисперсии

Доверительный интервал для дисперсии считается следующим образом:

$$\frac{(n-1)s^2}{\chi_{\alpha_1,k}^2} < D < \frac{(n-1)s^2}{\chi_{\alpha_2,k}^2}$$

$$\text{где } \alpha_1 = \frac{1-\gamma}{2}; \alpha_2 = \frac{1-\gamma}{2}; k = n - 1 = 57$$

При $\gamma = 0,95$

$$\alpha_1 = \frac{1-0,95}{2} = 0,025; \alpha_2 = \frac{1-0,95}{2} = 0,975;$$

$$\chi_{0,025,57}^2 = 82.12$$

$$\chi_{0,975,57}^2 = 39.66$$

$$\frac{(n-1)s^2}{\chi_{\alpha_1, k}^2} < D < \frac{(n-1)s^2}{\chi_{\alpha_2, k}^2}$$

$$\frac{57 \cdot 35,7523}{82,12} < D < \frac{57 \cdot 35,7523}{39,66}$$

$$24,8159 < D < 51,3837$$

Значит доверительный интервал при $\gamma = 0.95$

$$24,8159 < D < 51,3837$$

При $\gamma = 0,99$

$$\alpha_1 = \frac{1 - 0,99}{2} = 0,005; \alpha_2 = \frac{1 - 0,99}{2} = 0,995;$$

$$\chi_{0.005, 57}^2 = 90.72$$

$$\chi_{0.995, 57}^2 = 34.77$$

$$\frac{(n-1)s^2}{\chi_{\alpha_1, k}^2} < D < \frac{(n-1)s^2}{\chi_{\alpha_2, k}^2}$$

$$\frac{57 \cdot 35,7523}{90,72} < D < \frac{57 \cdot 35,7523}{34,77}$$

$$22,4634 < D < 58,6103$$

Значит доверительный интервал при $\gamma = 0.99$

$$22,4634 < D < 58,6103$$

12. Выводы

В итоге опираясь на проделанные ранее вычисления, мы можем сделать вывод:

По форме гистограммы и полигона уже можно было сделать предварительный вывод о том, что случайная величина скорее всего распределена нормально. Что подтверждает критерий Пирсона и критерий Колмогорова. Критерий Пирсона показал, что случайная величина скорее распределена нормально, а не равномерно (гипотеза о равномерном распределении была отвергнута). Критерий Колмогорова тоже показал что распределение скорее нормальное, но вероятность равномерного распределения согласно критерию Колмогорова тоже высока. Разный результат для гипотезы о равномерном распределении объясняется тем, что эти критерии оценивают разные характеристики случайной величины.

Так как вычисленные в работе математическое ожидание и дисперсия являются оценками, для них был вычислен доверительный интервал. Математическое ожидание находится в пределах от 28 до 32, а дисперсия от 22 до 58.