

Задание.

4.) *Построение гистограммы и графика эмпирической функции распределения.*

- По имеющимся значениям случайной величины построить вариационный ряд. Найти x_{\min} и x_{\max} .
- Выбрать промежуток $[a, b]$, в котором принимает значения случайная величина. При этом лучше взять значение $a \leq x_{\min}$, $a \in \mathbf{Z}$, значение $b \geq x_{\max}$, $b \in \mathbf{Z}$, близкое к x_{\max} .
- Разбить $[a, b]$ на 10 равных частей Δ , точками a_i ; $a = a_1 < a_2 < \dots < a_{11} = b$. Найти длину промежутков Δ_i , $h = \frac{b-a}{10}$.
- Составить таблицу:

№ интервала i	Границы интервала. $\Delta_i = (a_i, a_{i+1})$	Середина интервала. $x_i^* = \frac{a_i + a_{i+1}}{2}$	Подсчет числа значений ξ , попавших в Δ_i .	Число значений ξ , попавших в Δ_i , ν_i
--------------------	---	--	--	---

- По результатам таблицы 1 построить гистограмму и график эмпирической функции распределения.

5.) Найти выборочное среднее \bar{x} .

6.) Найти несмещенную оценку дисперсии s_1^2 .

7.) По виду гистограммы и имеющимся точечным оценкам выдвинуть гипотезу о характере распределения.

Данные:

Вариант 5									
42	48	45	45	55	40	32	61	28	45
70	46	43	41	49	37	36	51	28	52
27	61	40	32	46	42	53	40	50	36
49	29	39	68	38	43	54	46	64	72
32	36	50	23	33	47	42	47	42	53
43	55	43	36	35	59	30	64	35	56
29	53	69	35	52	48	60	56	37	64
28	42	32	35	48	62	36	41	51	45
49	42	37	33	49	75	40	58	47	45
44	46	41	40	35	66	41	37	46	50

Решение.

Пункт 4. Внесем данные в таблицу для расчета:

42	48	45	45	55	40	32	61	28	45
70	46	43	41	49	37	36	51	28	52
27	61	40	32	46	42	53	40	50	36
49	29	39	68	38	43	54	46	64	72
32	36	50	23	33	47	42	47	42	53
43	55	43	36	35	59	30	64	35	56
29	53	69	35	52	48	60	56	37	64
28	42	32	35	48	62	36	41	51	45
49	42	37	33	49	75	40	58	47	45

44	46	41	40	35	66	41	37	46	50
----	----	----	----	----	----	----	----	----	----

Объем выборки $n = 100$. Упорядочим данные по возрастанию, получим:

23
27
28
28
28
29
29
30
32
32
32
32
33
33
35
35
35
35
35
36
36
36
36
36
36
37
37
37
37
38
39
40
40
40
40
40
41
41
41
41
42
42
42
42
42
42
42
43
43
43
43

44
45
45
45
45
45
46
46
46
46
46
46
47
47
47
48
48
48
49
49
49
49
50
50
50
51
51
52
52
53
53
53
54
55
55
56
56
58
59
60
61
61
62
64
64
64
66
68
69
70
72
75

Подсчитывая значения вариант, построим вариационный ряд:

варианта	частота
23	1
27	1
28	3
29	2
30	1
32	4
33	2
35	5
36	5
37	4
38	1
39	1
40	5
41	4
42	6
43	4
44	1
45	5
46	5
47	3
48	3
49	4
50	3
51	2
52	2
53	3
54	1
55	2
56	2
58	1
59	1
60	1
61	2
62	1
64	3
66	1
68	1
69	1
70	1
72	1
75	1
Сумма	100

Найдем $x_{\min} = 23$, $x_{\max} = 75$.

Значения целые, поэтому выбираем промежуток $[a; b] = [23; 75]$. Размах вариации (длина промежутка) будет равна $75 - 23 = 52$, разбиваем на 10 равных частей. Длина каждого частичного интервала будет $h = \frac{52}{10} = 5,2$.

Составляем таблицу 1:

№ интервала	начало	конец	середина интервала	подсчет числа значений	число значений, попавших в интервал
1	23	28,2	25,6	1+1+1+1+1	5
2	28,2	33,4	30,8	1+1+1+1+1+1+1+1+1	9
3	33,4	38,6	36	1+1+1+1+1+ 1+1+1+1+1+ 1+1+1+1+1	15
4	38,6	43,8	41,2	1+1+1+1+1+ 1+1+1+1+1+ 1+1+1+1+1+ 1+1+1+1+1	20
5	43,8	49	46,4	1+1+1+1+1+ 1+1+1+1+1+ 1+1+1+1+1+ 1+1+1+1+1+1	21
6	49	54,2	51,6	1+1+1+1+1+ 1+1+1+1+1+1	11
7	54,2	59,4	56,8	1+1+1+1+1+1	6
8	59,4	64,6	62	1+1+1+1+1+1+1	7
9	64,6	69,8	67,2	1+1+1	3
10	69,8	75	56,8	1+1+1	3

По результатам таблицы построим гистограмму и график эмпирической функции распределения.

Гистограмма частот:

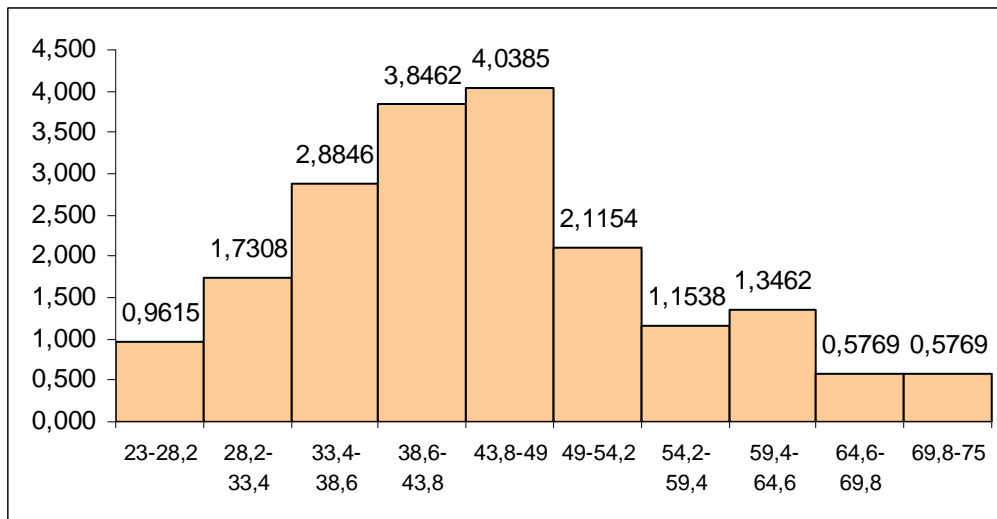
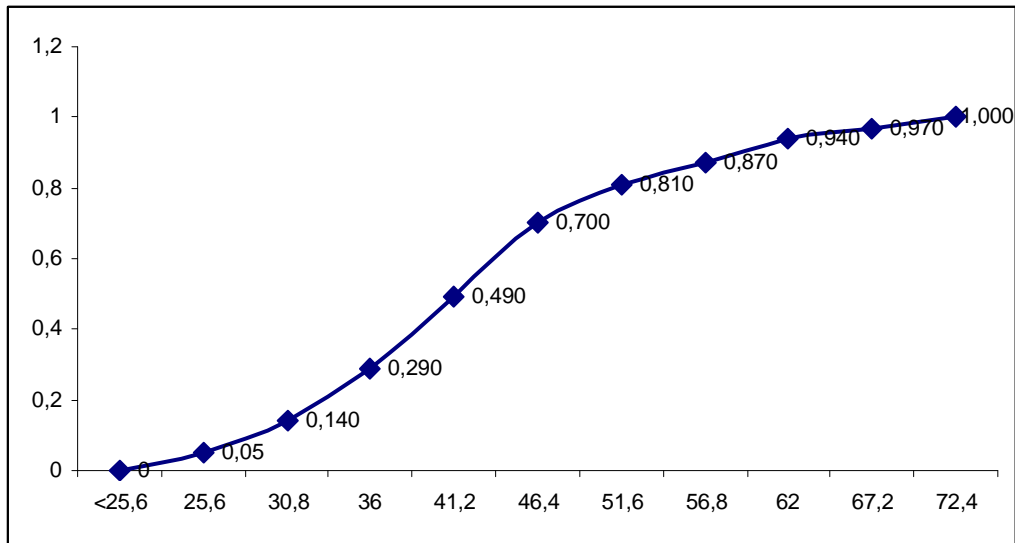


График эмпирической функции:



Дополнительные расчеты для построения в таблице 2:

	начало	конец	частота	плотность частоты	отн. частота	накопл. отн. частота
1	23	28,2	5	0,962	0,050	0,05
2	28,2	33,4	9	1,731	0,090	0,140
3	33,4	38,6	15	2,885	0,150	0,290
4	38,6	43,8	20	3,846	0,200	0,490
5	43,8	49	21	4,038	0,210	0,700
6	49	54,2	11	2,115	0,110	0,810
7	54,2	59,4	6	1,154	0,060	0,870
8	59,4	64,6	7	1,346	0,070	0,940
9	64,6	69,8	3	0,577	0,030	0,970
10	69,8	75	3	0,577	0,030	1,000

Пункт 5. Найдем выборочное среднее $\bar{\xi}$

$$\bar{\xi} = \frac{1}{n} \sum_{i=1}^n x_i n_i = \frac{1}{100} 4504,8 = 45,048.$$

Пункт 6. Найдем несмещенную оценку дисперсии s^2

$$s^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{\xi})^2 n_i = \frac{1}{99} 12309,69 \approx 124,340.$$

Исправленное выборочное среднее квадратическое отклонение

$$s = \sqrt{s^2} \approx 11,151$$

Расчеты в таблице 3:

x_i	n_i	$x_i n_i$	$(x_i - \bar{\xi})^2 n_i$
25,6	5	128	1891,124
30,8	9	277,2	1827,05
36	15	540	1227,995

41,2	20	824	296,1421
46,4	21	974,4	38,38598
51,6	11	567,6	472,2157
56,8	6	340,8	828,657
62	7	434	2011,592
67,2	3	201,6	1472,133
72,4	3	217,2	2244,396

Сумма 100 4504,8 12309,69

Пункт 7. По виду гистограммы и имеющимся точечным оценкам выдвинем гипотезу о характере распределения.

Исходя из внешнего вида гистограммы, можно с большой определенностью выдвинуть гипотезу о то, что генеральная совокупность имеет нормальное распределение с параметрами $a = 45,048$ и $\sigma = 11,151$. Предполагаемая плотность распределения

$$f(x) = \frac{1}{11,151\sqrt{2\pi}} \exp\left(-\frac{(x-45,048)^2}{2 \cdot 124,34}\right).$$

Графическое изображение кривой и гистограммы:

