Лабораторная работа 1
Вариант 62
Выборочный метод

Даны три выборки А,Б и Д (объемом N =54) образцов горных пород, в каждом из которых определено содержание Al2O3 в процентах.

1) Рассмотрим выборку А.

I. Требуется построить эмпирические дифференциальную и интегральную функции распределения по выборкам А, Б и Д.

Результаты измерений приведены в таблице 1.

Таблица 1–Выборка А

9,47	9,99	10,59	9,49	8,52	11,22	10,46	9,54	10,19
11,55	11,28	10,79	8,59	8,53	10,77	9,18	11,64	8,76
9,29	9,54	10,82	8,15	9,67	10,38	8,64	12,42	7,38
10,73	9,9	9,86	10,38	8,58	10,22	9,72	9,65	10,37
11,66	8,52	9,52	11,2	9,73	9,44	10,53	10,28	10,19
10,72	8,77	10,94	10,75	8,22	9,36	8,63	12,27	8,47

Решение поставленной задачи разобьем на этапы:

1. Определим интервал группирования по формуле Стирлинга:

$$h = \frac{Xmax - Xmin}{1 + 3.22 * lgN} = \frac{Xmax - Xmin}{7}$$

$$h_A = \frac{12{,}42 - 7{,}38}{7} = 0{,}72 \approx 1$$

2. Разобьем весь интервал изменения измеренного признака на частичные интервалы длиной $h_A \approx 1$ для выборки А.

3. Найдем середины интервалов x_i по выборке А по формуле:

$$x_i = \frac{a_i - b_i}{2},$$

Где: Где a_i- левая граница соответствующего интервала группирования; b_i- правая граница соответствующего интервала группирования.

4. Вычислим частоту попадания измеряемой величины в каждый интервал n_i по выборке А.

5. Вычислим относительные частоты w_i в выборке А по формуле:

$$w_i = \frac{n_i}{N}$$

Где: N – объем выборки, n_i– число попадания элементов в каждом интервал.

6. Вычислим плотность относительных частот v_i в выборке А по формуле:

$$v_i = \frac{w_i}{h}$$

Где: w_i– относительная частота, h – интервал группирования.

7. Вычислим накопленные относительные частоты F_i в каждой выборке по формуле:

$$F_i = \sum_{j=1}^{i} w_i$$

Где: w_i– относительная частота.

Результаты представим в таблиц 2

8. На основе полученных данных по выборке А построим гистограммы частот, гистограммы плотности относительных частот и гистограммы плотности накопленных частот.

Для построения эмпирической интегральной функции распределения, соединим плавной линией правые границы площадок построенной гистограммы накопленных относительных частот. Для построения эмпирической дифференциальной функции распределения, соединим плавной линией середины площадок, построенной гистограммы плотности относительных частот.

Таблица 2 –Результат по выборке А

Интервал	x_i	n_i	w_i	v_i	F_i
7…8	7,5	1	0,02	0,02	0,02
8…9	8,5	12	0,22	0,22	0,24
9…10	9,5	16	0,30	0,30	0,54
10…11	10,5	17	0,31	0,31	0,85
11…12	11,5	6	0,11	0,11	0,96
12…13	12,5	2	0,04	0,04	1,00

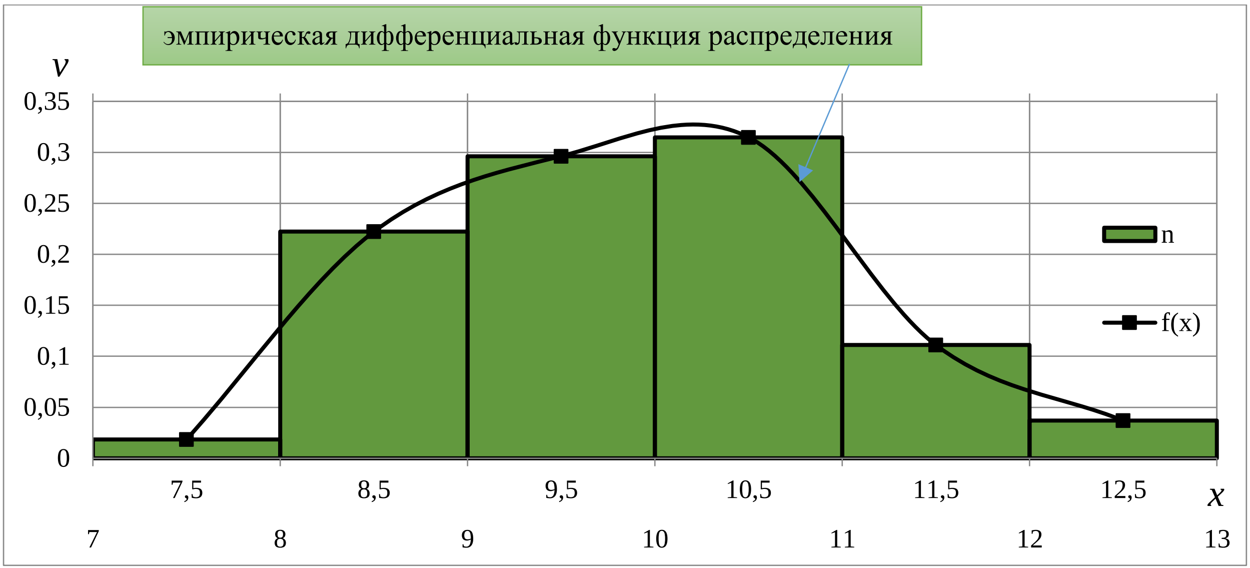

Рис.1 – Гистограмма плотности относительных частот

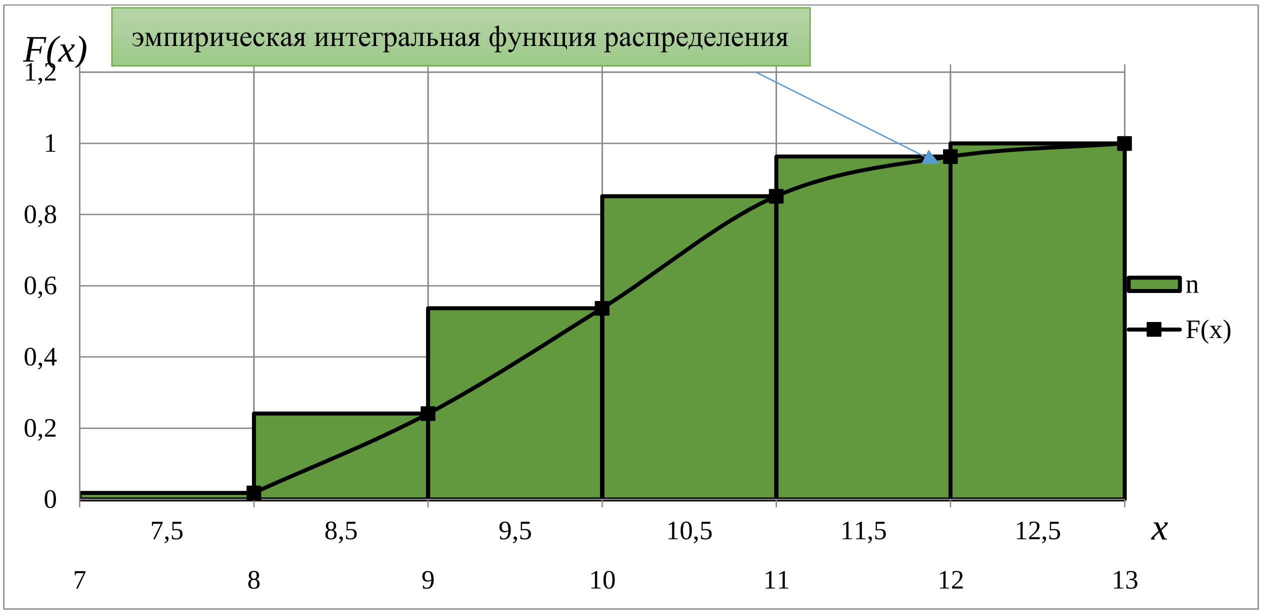

Рис.2 – Гистограмма накопленных относительных частот

II. Вычислить точечные оценки параметров распределения для выбороки А.

А. Оценка математического ожидания - выборочная средняя, мода и медиана для выборок А,Б и Д.

Выборочная средняя – среднее значение, вокруг которого находятся все изучаемые величины:

$$\overline{X_B} = \frac{\sum_{i=1}^{m} x_i n_i}{N}$$

$$\overline{X_B}\text{А} = \frac{7{,}5 * 1 + 8{,}5 * 12 + 9{,}5 * 16 + 10{,}5 * 17 + 11{,}5 * 6 + 12{,}5 * 2}{54} = 9{,}89$$

<u>Медиана (Me)</u> – это абсцисса линии, которая делит площадь ограниченную дифференциальной функцией на две равные части.

$$\mathrm{Me_A} = \frac{9{,}5 + 10{,}5}{2} = 10$$

<u>Мода</u> – точка максимума плотности распределения, в дискретном случае - точка максимума функции вероятности. Чтобы найти моду, необходимо в первую очередь определить модальный интервал. Модальным интервалом будет тот, которому соответствует наибольшая частота.

$$\mathrm{Mo} = x_o + h * \frac{\Delta n_1}{\Delta n_1 + \Delta n_2} = 10 + \frac{17 - 16}{(17 - 16) + (17 + 6)} = 10{,}08$$

Где: x_o — нижняя граница модального интервала; h — величина модального интервала; $\Delta n_1 = n_{Mo} - n_{Mo-1}$ разность частот модального и предмодального интервалов; $\Delta n_2 = n_{Mo} - n_{Mo+1}$ разность частот модального и послемодального интервалов.

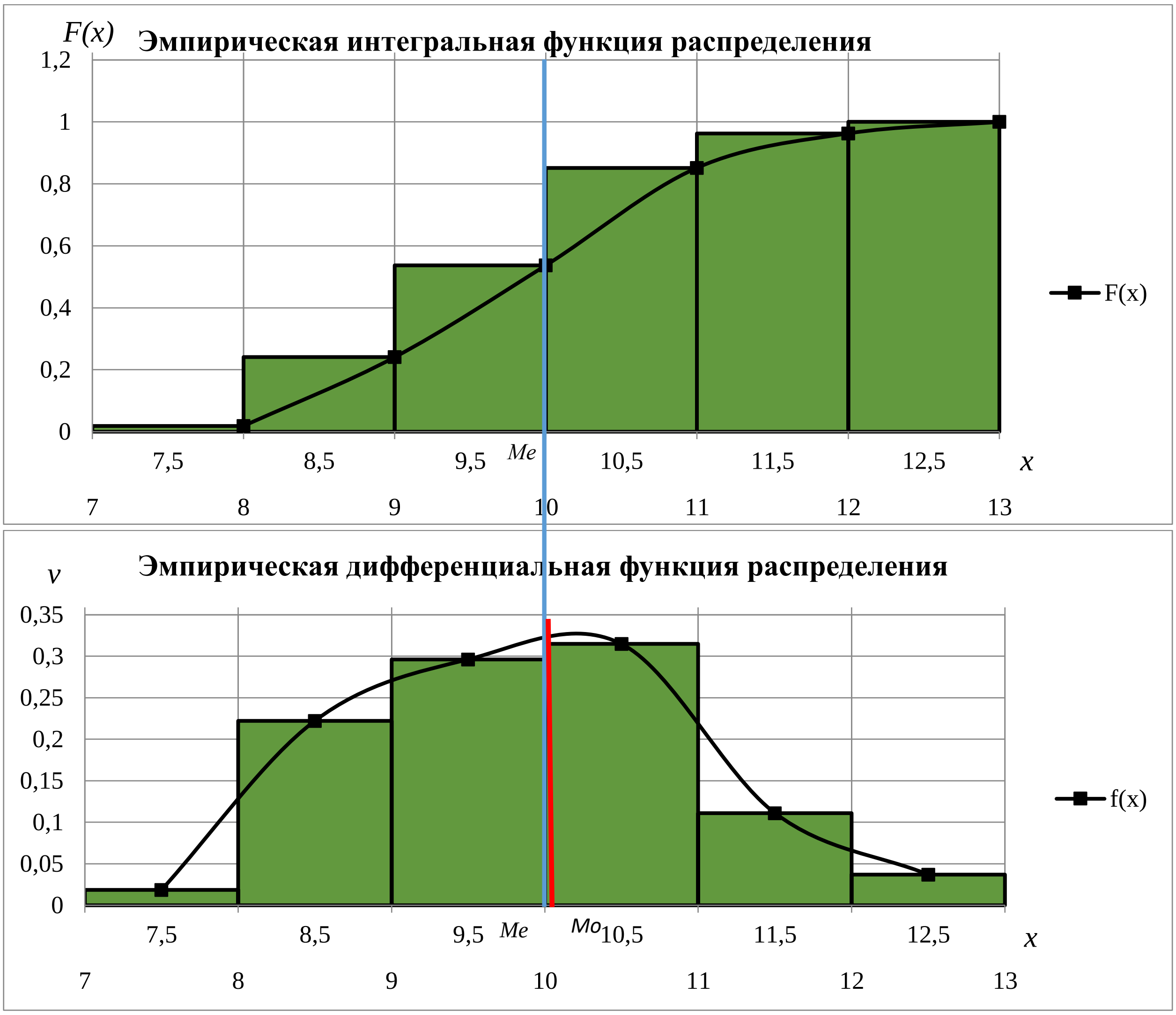

Рис.3 – Медина и Мода выборки А

3. <u>Дисперсия</u> - разброс случайных величин вокруг своего математического ожидания

<u>Оценка дисперсии</u> – выборочная дисперсия и Dв и квадрат стандарта ($S^2[x]$). Выборочная дисперсия является смещѐнной оценкой дисперсии, несмещѐнной оценкой дисперсии является квадрат стандарта, поэтому мы будем вычислять несмещѐнной оценкой дисперсии, и, следовательно, Оценкой среднеквадратического отклонения является стандарт s для выбороки А по формулам:

$$S^2[x] = \frac{\sum_{i=1}^{m}(X_i - \overline{X_B})n_i}{N-1}$$

$$s = \sqrt{S^2[x]}$$

Таблица 3 – Вспомогательная таблица по выборке А

x_i	$t_i = x_i - \overline{x_{\text{в}}}$	${t_i}^2$	n_i
7,5	−2,39	5,71	1
8,5	−1,39	1,93	12
9,5	−0,39	0,15	16
10,5	0,61	0,37	17
11,5	1,61	2,60	6
12,5	2,61	6,82	2

$$S^2[x] = \frac{\sum_{i=1}^{m}(X_i - \overline{X_B})^2 n_i}{N-1} = \frac{5{,}71 * 1 + 1{,}93 * 12 + 0{,}15 * 16 + 0{,}37 * 17 + 2{,}6 * 6 + 6{,}82 * 2}{53} = 1{,}26$$

$$s = \sqrt{S^2[x]} = \sqrt{1{,}26} = 1{,}12$$

1) Рассмотрим выборку Б.

I. Требуется построить эмпирические дифференциальную и интегральную функции распределения по выборке Б.

Дана выборка Б (объемом N =54) образцов горных пород, в каждом из которых определено содержание Al_2O_3 в процентах.

Результаты измерений приведены в таблице 4:

Таблица 4–Выборка Б

16,36	13,79	15,24	16,01	14,55	12,5	10,73	13,19	17,08
14,78	13,39	16,08	14,84	14,23	18,22	15,83	17,05	13,96
15,53	16,89	16,62	15,06	14,94	16,07	14,67	13,35	13,49
12,57	16,41	14,77	13,21	17,21	16,2	14,46	13,94	14,5
16,63	17,34	16,8	16,96	15,06	12,1	16,34	16,48	15,88
16,07	15,91	14,11	15,63	15,55	15,71	14,03	14,02	16,51

Решение поставленной задачи разобьем на этапы:

1. Определим интервал группирования по формуле Стирлинга:

$$h = \frac{Xmax - Xmin}{1 + 3.22 * lgN} = \frac{Xmax - Xmin}{7}$$

$$h_{Б} = \frac{18,22 - 10,73}{7} = 1,07 \approx 1$$

2. Разобьем весь интервал изменения измеренного признака на частичные интервалы длиной $h_{Б} \approx 1$ для выборки и длиной $h_{В} \approx 2$.

3. Найдем середины интервалов x_i по выборке Б по формуле:

$$x_i = \frac{a_i - b_i}{2},$$

Где: Где a_i- левая граница соответствующего интервала группирования; b_i- правая граница соответствующего интервала группирования.

4. Вычислим частоту попадания измеряемой величины в каждый интервал n_i по выборке Б.

5. Вычислим относительные частоты w_i в выборке Б по формуле:

$$w_i = \frac{n_i}{N}$$

Где: N – объем выборки, n_i– число попадания элементов в каждом выборке Б по формуле:

$$v_i = \frac{w_i}{h}$$

Где: w_i– относительная частота, h – интервал группирования.

6. Вычислим накопленные относительные частоты F_i в выборке Б по формуле:

$$F_i = \sum_{j=1}^{i} w_i$$

Где: w_i– относительная частота.

Результаты представим в таблицы 5.

7. На основе полученных данных по выборке Б построим гистограммы частот, гистограммы плотности относительных частот и гистограммы плотности накопленных частот. Для построения эмпирической интегральной функции распределения, соединим плавной линией правые границы площадок построенной гистограммы накопленных относительных частот. Для построения эмпирической дифференциальной функции распределения, соединим плавной линией середины площадок, построенной гистограммы плотности относительных частот.

Таблица 5 –Результат по выборке Б

Интервал	x_i	n_i	w_i	v_i	F_i
10…11	10,5	1	0,02	0,02	0,02
11…12	11,5	3	0,06	0,06	0,07
12…13	12,5	8	0,15	0,15	0,22
13…14	13,5	12	0,22	0,22	0,44
14…15	14,5	10	0,19	0,19	0,63
15…16	15,5	15	0,28	0,28	0,91
16…17	16,5	4	0,07	0,07	0,98
17…18	17,5	1	0,02	0,02	1,00

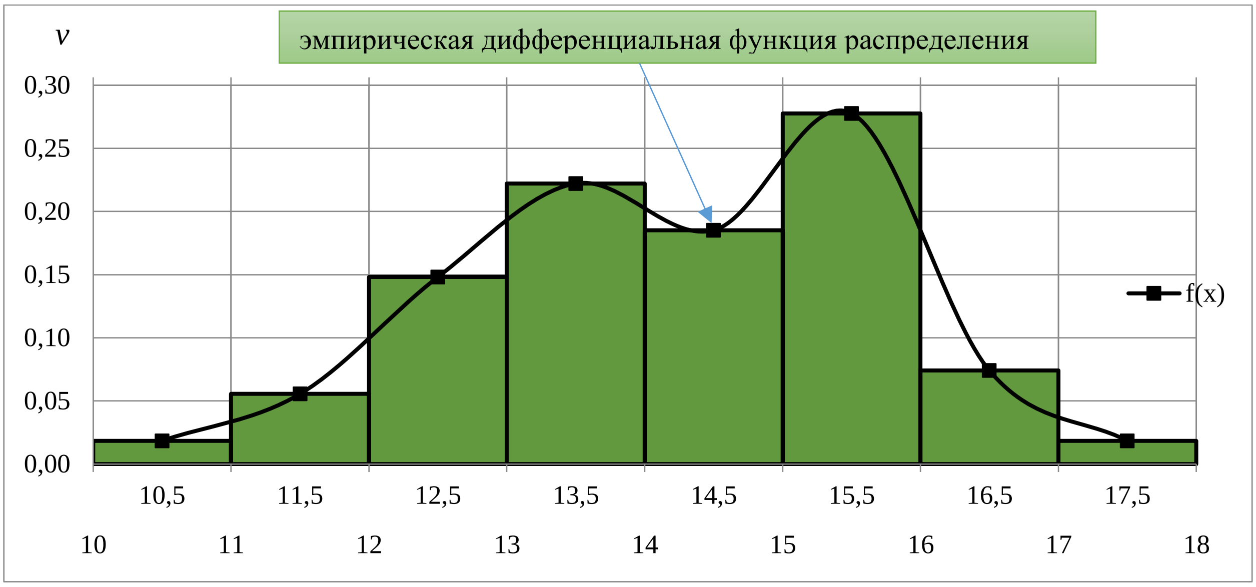

Рис.4 – Гистограмма плотности относительных частот выборки Б

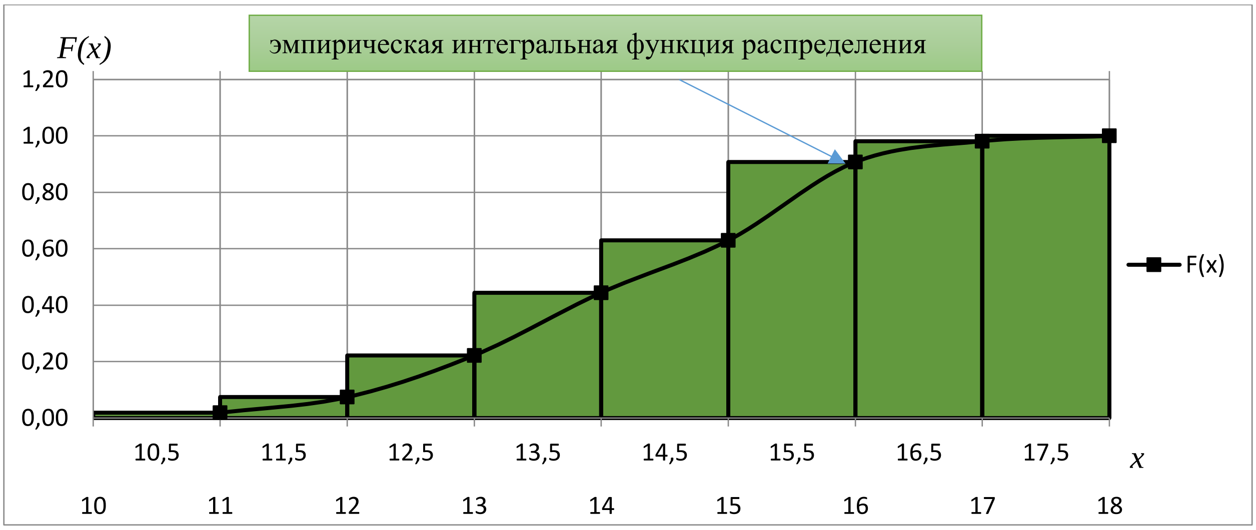

Рис.5 Гистограмма накопленных относительных частот выборки Б

III. Вычислить точечные оценки параметров распределения для выборок А, Б и Д.

В. Оценка математического ожидания - выборочная средняя, мода и медиана для выбороки Б .

Выборочная средняя – среднее значение, вокруг которого находятся все изучаемые величины:

$$\overline{X_B} = \frac{\sum_{i=1}^{m} x_i n_i}{N}$$

$$\overline{X_B}Б = \frac{10{,}5*1+11{,}5*3+12{,}5*8+13{,}5*12+14{,}5*10+15{,}5*15+16{,}5*4+17{,}5*1}{54} = 14{,}22$$

Медиана (Ме) – это абсцисса линии, которая делит площадь ограниченную дифференциальной функцией на две равные части.

$$Me_Б = \frac{13{,}5+14{,}5}{2} = 14$$

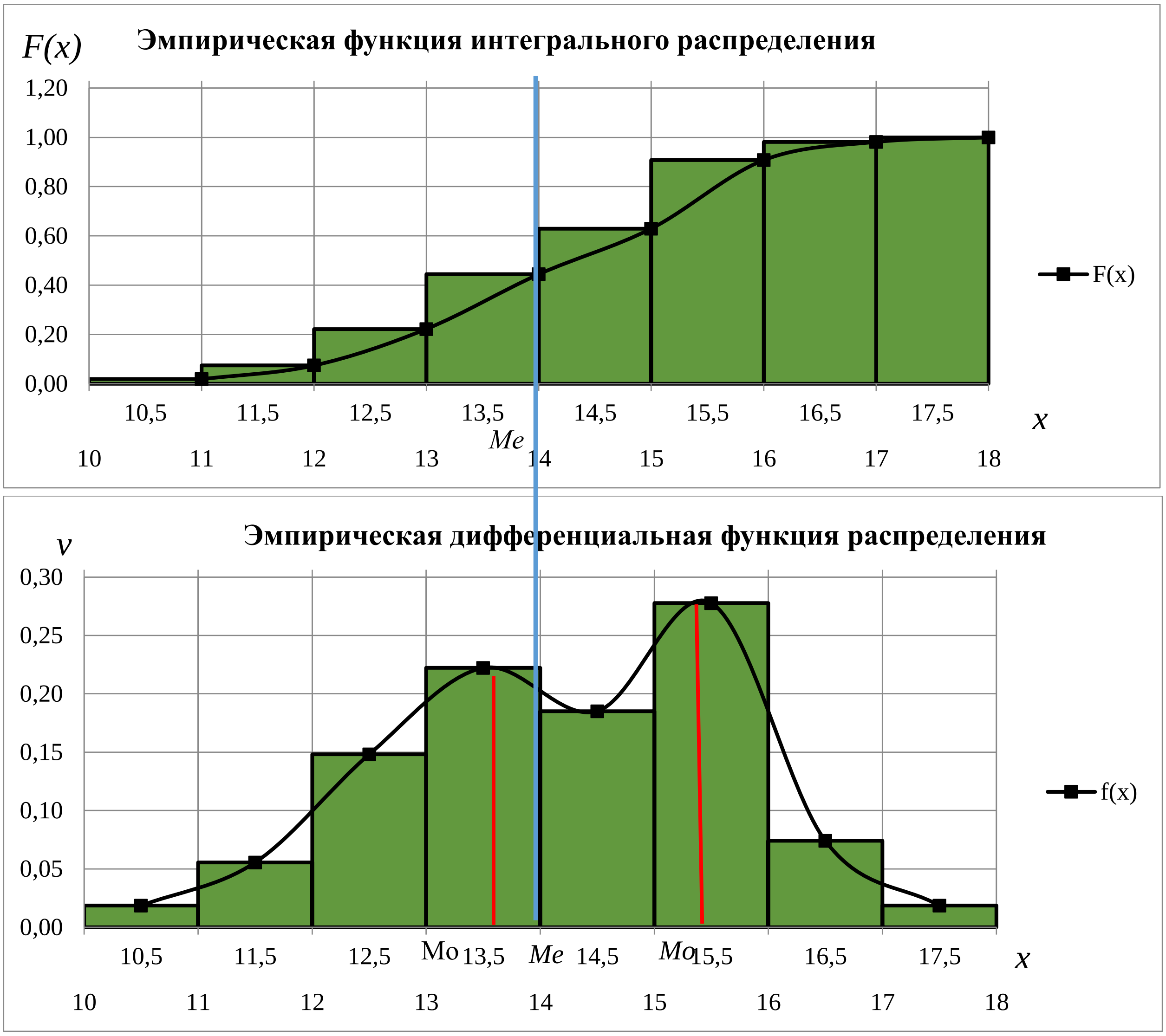

Рис.6 – Медиана и Мода выборки Б

Мода – точка максимума плотности распределения, в дискретном случае - точка максимума функции вероятности. Чтобы найти моду, необходимо в первую очередь определить модальный интервал. Модальным интервалом будет тот, которому соответствует наибольшая частота.

$$\mathrm{Mo}_1 = x_o + h * \frac{\Delta n_1}{\Delta n_1 + \Delta n_2} = 13 + \frac{12 - 10}{(12 - 10) + (12 + 4)} = 13{,}67$$

$$\mathrm{Mo}_2 = 15 + \frac{15 - 10}{(15 - 10) + (15 + 4)} = 15{,}31$$

Где: x_o — нижняя граница модального интервала; h — величина модального интервала; $\Delta n_1 = n_{Mo} - n_{Mo-1}$ разность частот модального и предмодального интервалов; $\Delta n_2 = n_{Mo} - n_{Mo+1}$ разность частот модального и послемодального интервалов.

4. Дисперсия - разброс случайных величин вокруг своего математического ожидания

Оценка дисперсии – выборочная дисперсия и Dв и квадрат стандарта (S^2[x]). Выборочная дисперсия является смещѐнной оценкой дисперсии, несмещѐнной оценкой дисперсии является квадрат стандарта, поэтому мы будем вычислять несмещѐнной оценкой дисперсии, и, следовательно, Оценкой среднеквадратического отклонения является стандарт s для выбороки Б по формулам:

$$S^2[x] = \frac{\sum_{i=1}^{m} (X_i - \overline{X_B}) n_i}{N - 1}$$

$$s = \sqrt{S^2[x]}$$

Таблица 6 – Вспомогательная таблица по выборке Б

x_i	$t_i = x_i - \overline{x_\text{в}}$	${t_i}^2$	n_i
10,5	-3,72	13,85	1
11,5	-2,72	7,41	3
12,5	-1,72	2,97	8
13,5	-0,72	0,52	12
14,5	0,28	0,08	10
15,5	1,28	1,63	15
16,5	2,28	5,19	4
17,5	3,28	10,74	1

$$S^2[x] = \frac{13,85*1+7,41*3+2,97*8+0,52*12+0,08*10+1,63*15+5,19*4+10,74*1}{53} = 2,32$$

$$s = \sqrt{S^2[x]} = \sqrt{2,32} = 1,52$$

2) Рассмотрим выборку Д.

Дана выборка Д (объемом N =54) образцов горных пород, в каждом из которых определено содержание Al2O3 в процентах.

I. Требуется построить эмпирические дифференциальную и интегральную функции распределения по выборке Д.

Результаты измерений приведены в таблице 7:

Таблица 7 – Выборка Д

10,02	11,21	16,45	9,64	6,33	11,48	10,9	11,84	12,88
6,62	12,25	11,77	12,41	8,65	11,96	10,48	10,92	12,15
14,35	8,31	10,69	9,44	11,42	10,12	9,47	11,73	8,22
12,57	16,41	14,77	13,21	17,21	16,2	14,46	13,94	14,5
16,63	17,34	16,8	16,96	15,06	12,1	16,34	16,48	15,88
16,07	15,91	14,11	15,63	15,55	15,71	14,03	14,02	16,51

Решение поставленной задачи разобьем на этапы:

1. Определим интервал группирования по формуле Стирлинга:

$$h = \frac{Xmax - Xmin}{1 + 3.22 * lgN} = \frac{Xmax - Xmin}{7}$$

$$h_{В} = \frac{17,34 - 6,33}{7} = 1,57 \approx 2$$

2. Разобьем весь интервал изменения измеренного признака на частичные интервалы длиной $h_{В} \approx 2$ для выборки Д.

3. Найдем середины интервалов x_i по к выборке Д по формуле:

$$x_i = \frac{a_i - b_i}{2},$$

Где: Где a_i- левая граница соответствующего интервала группирования; b_i- правая граница соответствующего интервала группирования.

4. Вычислим частоту попадания измеряемой величины в каждый интервал n_i по выборке Д.

5. Вычислим относительные частоты w_i в выборке Д по формуле:

$$w_i = \frac{n_i}{N}$$

Где: N – объем выборки, n_i– число попадания элементов в каждом интервал.

6. Вычислим плотность относительных частот v_i в выборке Д по формуле:

$$v_i = \frac{w_i}{h}$$

Где: w_i– относительная частота, h – интервал группирования.

7. Вычислим накопленные относительные частоты F_i в выборке Д по формуле:

$$F_i = \sum_{j=1}^{i} w_i$$

Где: w_i– относительная частота.

Результаты представим в таблицы 8

8. На основе полученных данных по выборке Д построим гистограммы частот, гистограммы плотности относительных частот и гистограммы плотности накопленных частот.

Таблица 8 –Результат по выборке Д

Интервал	x	n	w	v	F
6…8	7	2	0,04	0,02	0,04
8…10	9	6	0,11	0,06	0,15
10…12	11	13	0,24	0,12	0,39
12…14	13	8	0,15	0,07	0,54
14…16	15	13	0,24	0,12	0,78
16…18	17	12	0,22	0,11	1,00

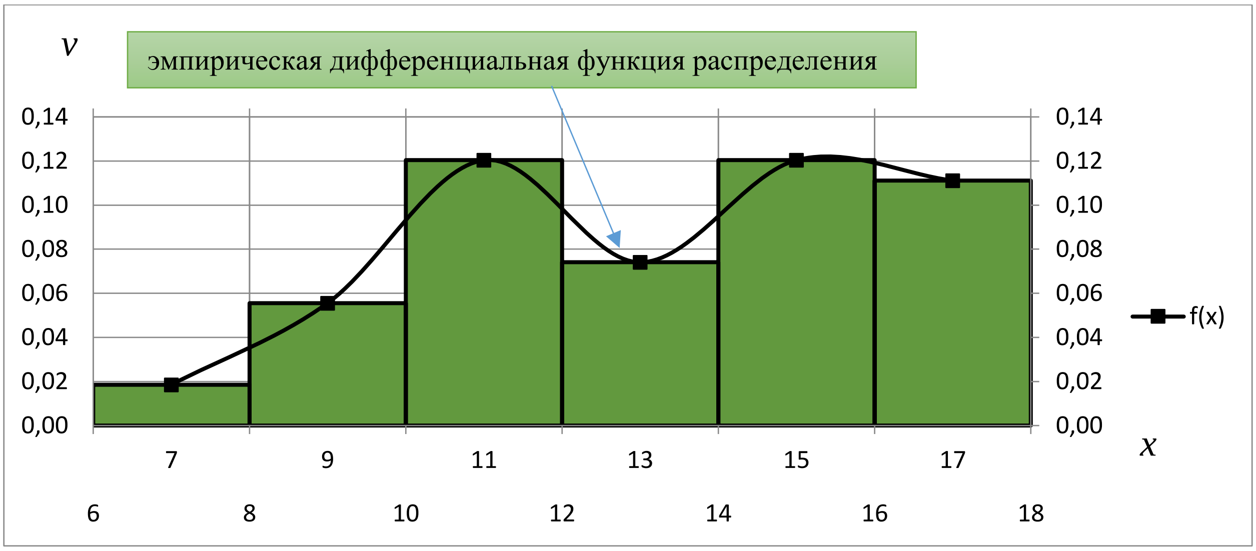

Рис.7–Гистограмма плотности относительных частот выборки Д

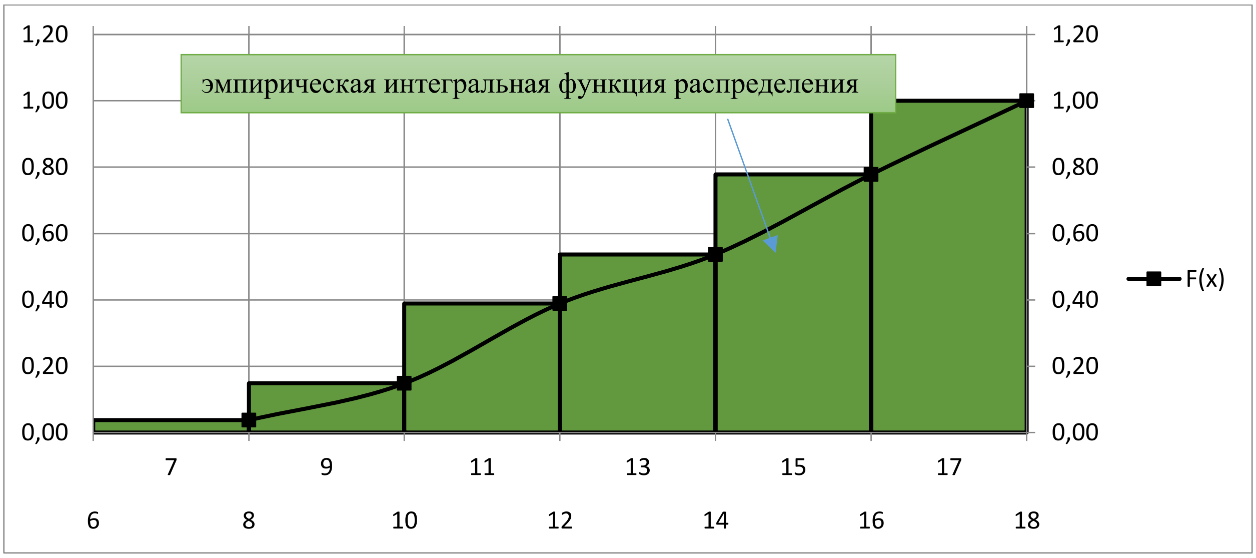

Рис.8 – Гистограмма накопленных относительных частот выборки Д

II. Вычислить точечные оценки параметров распределения для выборок А, Б и Д.

C. Оценка математического ожидания - выборочная средняя, мода и медиана для выборки Д.

Выборочная средняя – среднее значение, вокруг которого находятся все изучаемые величины:

$$\overline{X_B} = \frac{\sum_{i=1}^{m} x_i n_i}{N}$$

$$\overline{X_B}Д = \frac{7*2+9*6+11*13+13*8+15*13+17*12}{54} = 13{,}22$$

Медиана (Ме) – это абсцисса линии, которая делит площадь ограниченную дифференциальной функцией на две равные части.

$$\text{Ме}_\text{Д} = \frac{11+13}{2} = 12$$

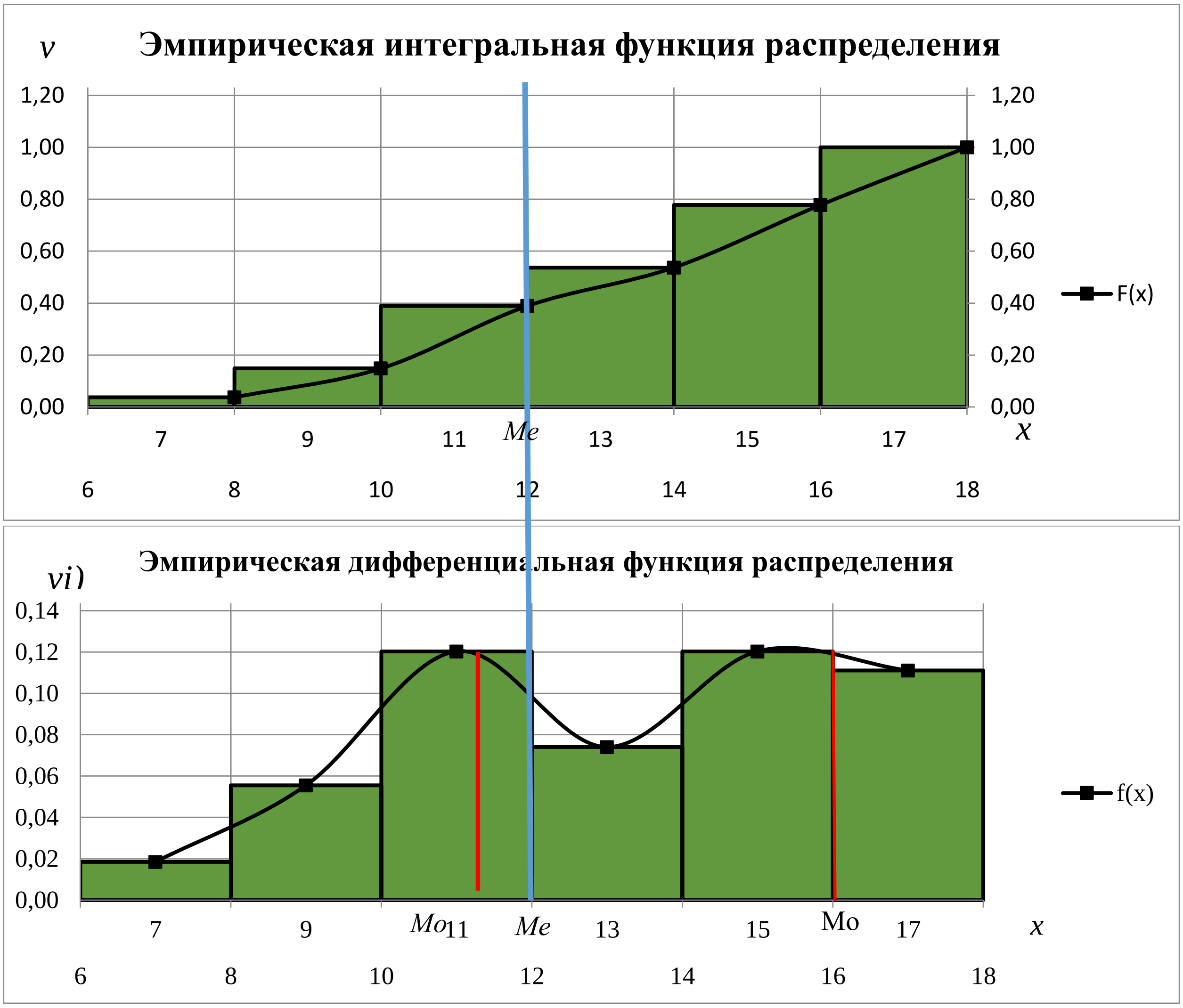

Рис.9 – Медиана и мода выборки Д

Мода – точка максимума плотности распределения, в дискретном случае - точка максимума функции вероятности. Чтобы найти моду, необходимо в первую очередь определить модальный интервал. Модальным интервалом будет тот, которому соответствует наибольшая частота.

$$\text{Mo}_1 = x_o + h * \frac{\Delta n_1}{\Delta n_1 + \Delta n_2} = 10 + \frac{13-6}{(13-8)+(13+8)} = 11{,}17$$

$$\text{Mo}_2 = 14 + \frac{13-8}{(13-8)+(13+12)} = 16{,}00$$

Где: x_o — нижняя граница модального интервала; h — величина модального интервала; $\Delta n_1 = n_{Mo} - n_{Mo-1}$ разность частот модального и предмодального интервалов; $\Delta n_2 = n_{Mo} - n_{Mo+1}$ разность частот модального и послемодального интервалов.

5. Дисперсия - разброс случайных величин вокруг своего математического ожидания

Оценка дисперсии – выборочная дисперсия и Dв и квадрат стандарта (S^2[x]). Выборочная дисперсия является смещѐнной оценкой дисперсии, несмещѐнной оценкой дисперсии является квадрат стандарта, поэтому мы будем вычислять несмещѐнной оценкой дисперсии, и, следовательно, Оценкой среднеквадратического отклонения является стандарт s для выборок А,Б и Д по формулам:

$$S^2[x] = \frac{\sum_{i=1}^{m}(X_i - \overline{X_B})n_i}{N-1}$$

$$s = \sqrt{S^2[x]}$$

Таблица 9 – Вспомогательная таблица по выборке Д

x_i	$t_i = x_i - \overline{x}_в$	${t_i}^2$	n_i
7	-6,22	38,72	2
9	-4,22	17,83	6
11	-2,22	4,94	13
13	-0,22	0,05	8
15	1,78	3,16	13
17	3,78	14,27	12

$$S^2[x] = \frac{38{,}72 * 2 + 17{,}83 * 6 + 4{,}94 * 13 + 0{,}05 * 8 + 3{,}16 * 13 + 14{,}27 * 12}{53} = 8{,}70$$

$$s = \sqrt{S^2[x]} = \sqrt{8{,}70} = 2{,}95$$