ИДЗ 19.1. Вариант 18. Агаев Хамза Р3234 (поток 1.5).

В результате эксперимента получены следующие данные, записанные в виде статистического ряда:

2,1	2,3	1,5	3,1	2,7	1,9	2,4	0,9	2,5	1,1
1,3	2,9	2,3	3,9	2,4	3,6	1,6	3,2	2,9	2,0
2,1	3,3	0,8	3,5	1,7	2,6	4,1	2,8	1,2	2,5
1,1	2,4	1,5	3,2	2,7	1,5	3,7	1,9	3,1	4,0
4,1	2,9	2,0	2,0	1,1	0,7	3,3	2,5	1,6	2,4
2,1	3,2	0,9	2,8	4,2	2,8	1,9	1,2	1,7	3,5
2,7	3,9	2,4	1,7	3,6	2,5	0,8	3,1	2,1	1,3
3,2	1,6	0,7	2,6	1,3	2,0	3,7	2,9	4,0	3,1
2,8	4,1	1,9	3,6	3,3	2,9	0,6	1,5	1,2	2,4
1,1	3,5	1,6	2,4	3,9	2,7	2,5	1,9	2,6	3,2

Решение:

а и б) Построим интервальный вариационный ряд распределения. По статистическим данным находим: $min\, x_i = 0,6;$ $max\, x_i = 4,2.$

Размах вариации: $\omega = 4,2 - 0,6 = 3,6;$

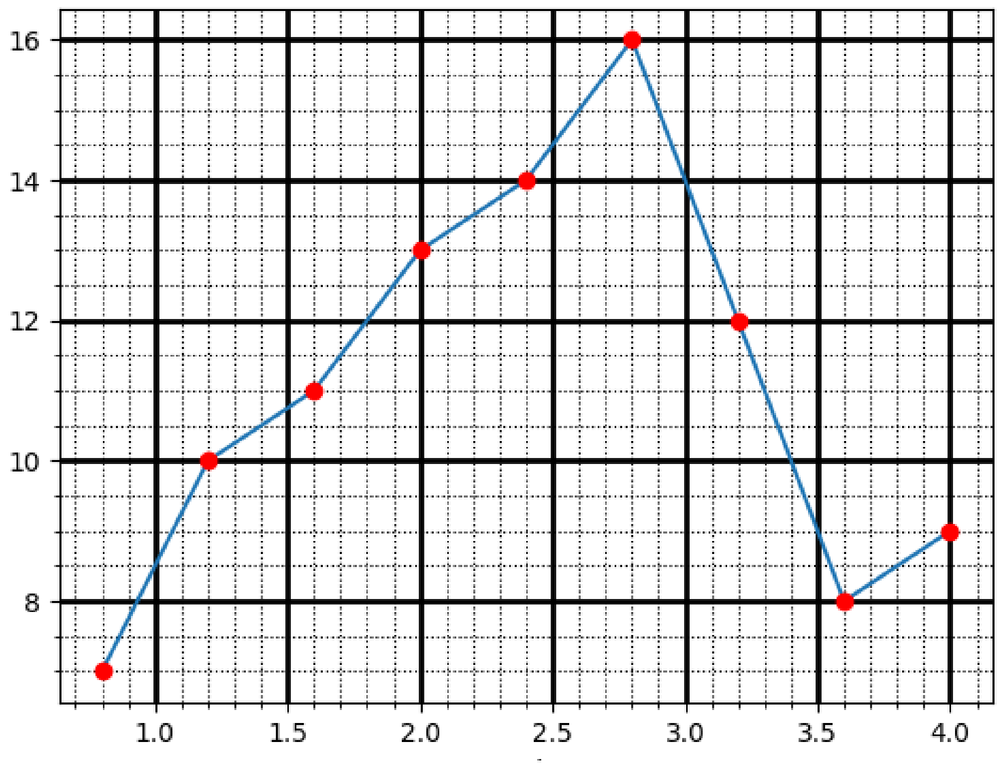
Выборку разобьем на 9 равных интервалов. Величина отдельного интервала: $h = 3,6/9 = 0,4.$ Подсчитаем частоту m_i по каждому интервалу и запишем в таблицу. Перейдем к дискретному ряду распределения, выбрав в качестве вариант x_i середины интервалов.

Заполним расчетную таблицу:

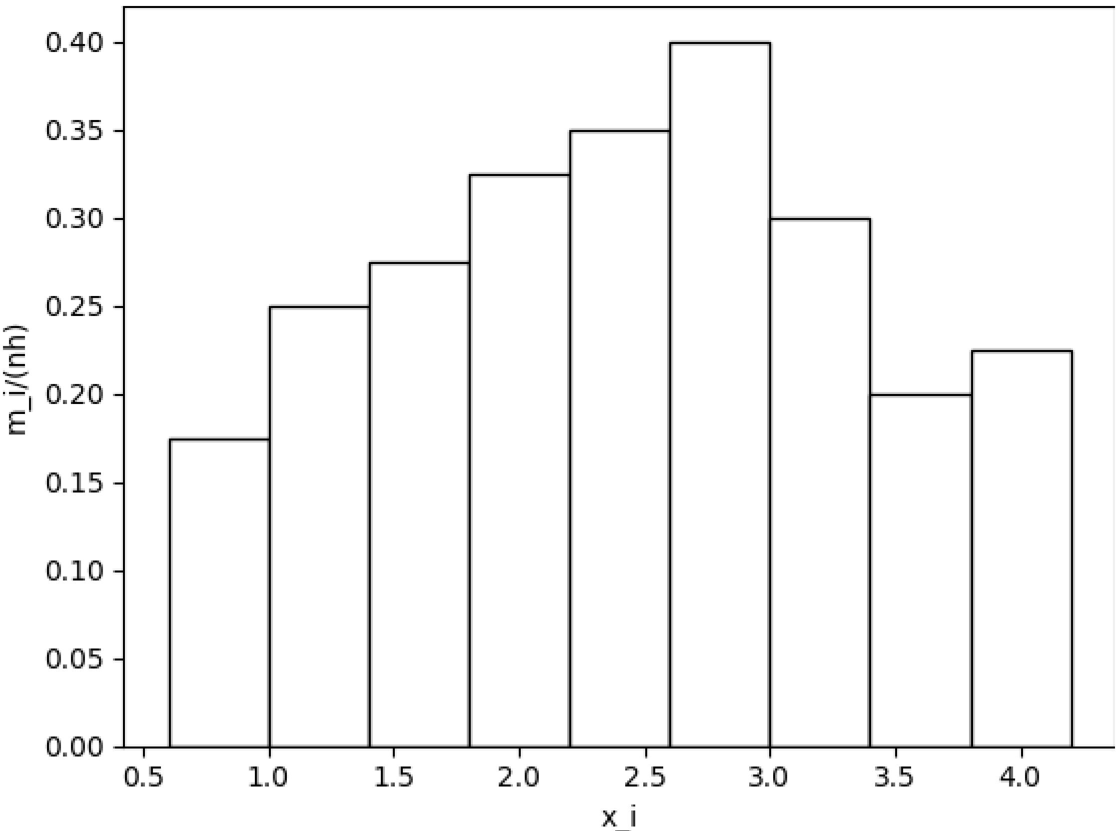
Интервал	x_i	m_i	m_i/n	Накопленная относительная	$x_i * m_i$	$x^2_i * m_i$

					частота		
0,6	1	0,8	7	0,07	0,07	5,6	4,48
1	1,4	1,2	10	0,1	0,17	12	14,4
1,4	1,8	1,6	11	0,11	0,28	17,6	28,16
1,8	2,2	2	13	0,13	0,41	26	52
2,2	2,6	2,4	14	0,14	0,55	33,6	80,64
2,6	3	2,8	16	0,16	0,71	44,8	125,44
3	3,4	3,2	12	0,12	0,83	38,4	122,88
3,4	3,8	3,6	8	0,08	0,91	28,8	103,68
3,8	4,2	4	9	0,09	1	36	144
			100			242,8	675,68

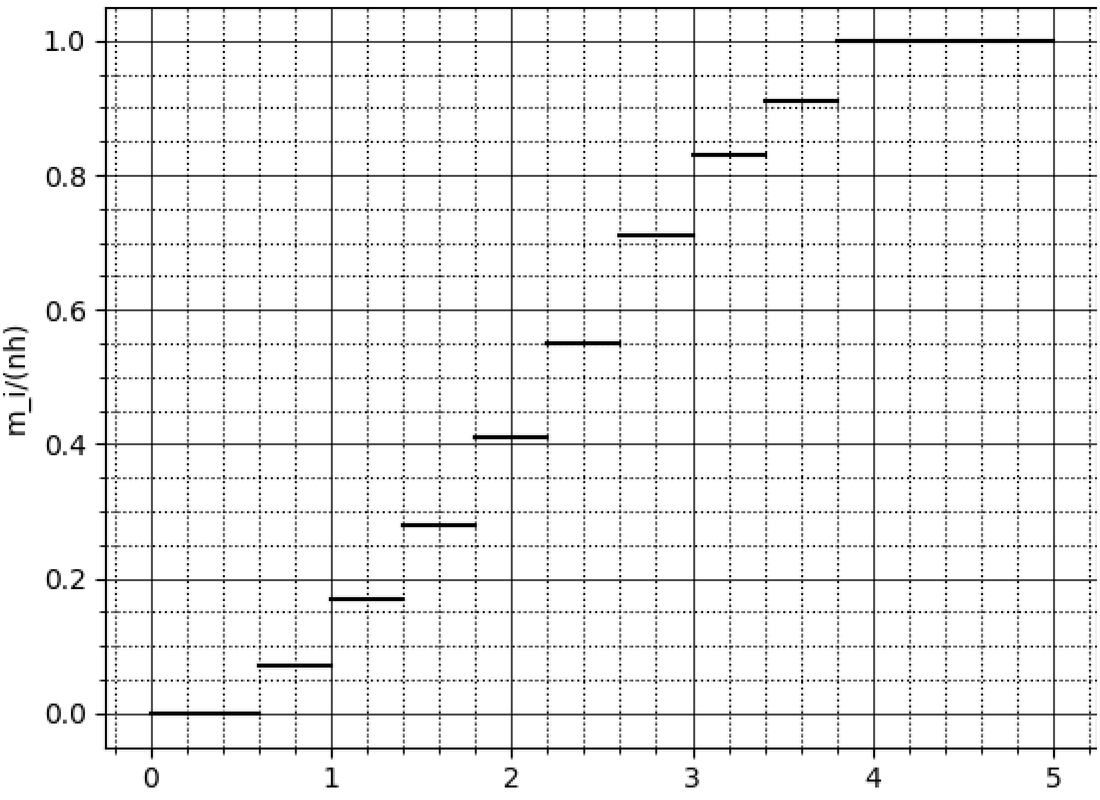
в) Полигон частот:

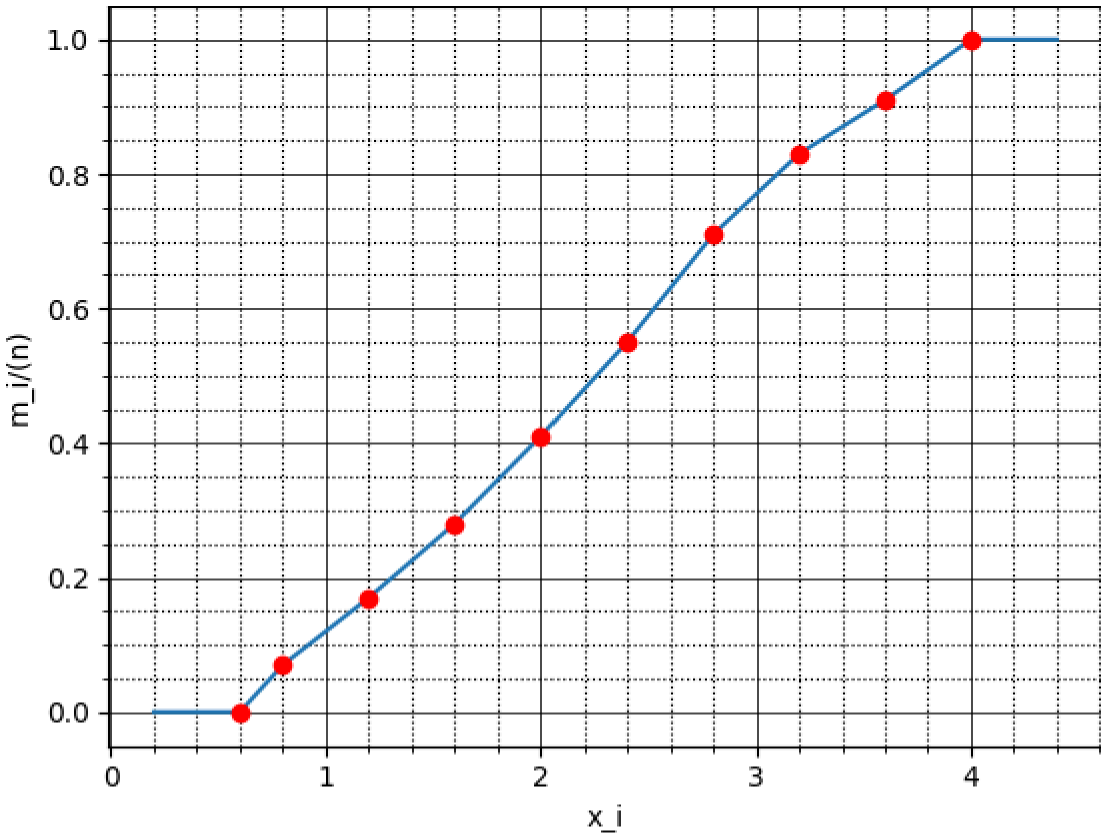


Гистограмма относительных частот:



Эмпирическая функция распределения:





г) Числовые характеристики выборки:

Выборочное среднее: $\overline{x} = 1/n * \sum_{i=1}^{k} m_i x_i$ = 242,8/100 = 2,428

Выборочная дисперсия:

$$D = \sum_{i=1}^{k} m_i x_i^2/n - (\overline{x})^2 = 675,68/100 - 2,428^2 = 0,861616$$

д) Проведем проверку гипотезы H_0: генеральная совокупность распределена по нормальному закону; конкурирующая гипотеза H_1 : генеральная совокупность не распределена по нормальному закону.

Используем критерий Пирсона; уровень значимости 0,025.

Найдем значения теоретических частот. Используем формулу:

$m'_i = nh * \phi(u_i)/\sigma_B$, где $u_i = (x_i - \overline{x})/\sigma_B$, $\phi(u) = exp(-u^2/2)/\sqrt{2\pi}$.

В данном случае среднее квадратическое отклонение: $\sigma_B = \sqrt{D} \approx 0,928$.

Заполним расчетную таблицу:

x_i	u_i	$\phi(u_i)$	m'_i
0,8	-1,754310345	0,085628173	3,690869526
1,2	-1,323275862	0,1662158496	7,164476274
1,6	-0,8922413793	0,2679416965	11,54921105
2	-0,4612068966	0,3586908391	15,46081203
2,4	-0,03017241379	0,3987607283	17,18796243
2,8	0,400862069	0,3681430357	15,8682343
3,2	0,8318965517	0,2822493244	12,16591916
3,6	1,262931034	0,1797054905	7,745926313
4	1,693965517	0,09501712694	4,095565817

Расчетное значение критерия вычислим по формуле:

$\chi^2_H = \sum_i (m_i - m'_i)^2 / m'_i$. Заполним расчетную таблицу:

m_i	m'_i	$(m_i - m'_i)^2 / m'_i$
7	3,690869526	2,966873909
10	7,164476274	1,122230641
11	11,54921105	0,02611717636
13	15,46081203	0,391673855
14	17,18796243	0,5912919854
16	15,8682343	0,001094148226
12	12,16591916	0,002262810204
8	7,745926313	0,00833385649
9	4,095565817	5,873052891

Итого получаем, что $\chi^2_H = 10,98293127$.

По соответствующей таблице найдем теоретическое значение критерия: $\chi^2_{кр}(\alpha,\ k)$. В данном случае $\alpha = 0,025$, количество степеней свободы: $k = s - r - 1$, где s - количество групп выборки, r - число параметров, оцениваемых по выборке. Получаем $k = 9 - 2 - 1 = 6$. По соответствующей таблице $\chi^2_{кр}(\alpha,\ k) = 16$.

Таким образом $\chi^2_H < \chi^2_{кр}(\alpha,\ k)$, то есть на уровне значимости $\alpha = 0,025$, принимаем нулевую гипотезу о том, что генеральная совокупность, из которой извлечена выборка, имеет нормальное распределение.

e) Доверительный интервал для оценки истинного значения генеральной средней $\tilde{x}$измеряемой величины вычислим по формуле:

$$\bar{x} - (t_\gamma \sigma)/\sqrt{n} < \tilde{x} < \bar{x} + (t_\gamma \sigma)/\sqrt{n}.$$

Уровень надежности $\gamma = 0,9$. По количеству степеней свободы $f = n - 1 = 9 - 1 = 8$ и уровню значимости $\gamma = 0,9$ по таблице распределения Стьюдента находим: t_γ=1,86.

Вычислим точность оценки: $(t_\gamma \sigma)/\sqrt{n} = 0,172608$.

Получаем следующий доверительный интервал:
$2,255392 < \tilde{x} < 2,600608$. С вероятностью 90% данный интервал накроет истинное значение $\tilde{x}$ генеральной средней.

При размере выборки $n > 30$ доверительный интервал для генерального среднего квадратического отклонения $\tilde{\sigma}$ определяется по формуле:

$$\sqrt{2n}/(\sqrt{2n - 3} + t_\gamma) * \sigma < \tilde{\sigma} < \sqrt{2n}/(\sqrt{2n - 3} - t_\gamma) * \sigma.$$

Ввиду достаточно большого количества наблюдений смещенностью найденного значения σ пренебрегаем. Таким образом, искомый доверительный интервал:

$$0,889685 < \tilde{\sigma} < 1,16151.$$

С вероятностью 90% данный интервал накроет истинное значение генерального среднего квадратического отклонения.