Задача

Дана выборка объёмом $n = 30$:

6.28	6.31	6.23	6.35	6.32	6.36	6.33	6.31	6.26	6.21
6.31	6.38	6.34	6.25	6.28	6.39	6.27	6.32	6.29	6.30
6.24	6.32	6.26	6.35	6.32	6.31	6.29	6.28	6.33	6.36

1. Найти статистический ряд и построить полигон частот.

2. Составить интервальный статистический ряд, взяв $7 - 10$ интервалов, и построить гистограмму частот.

3. Найти: оценки математического ожидания $\bar{x}$, выборочную дисперсию $D_{\text{в}}$, исправленную выборочную дисперсию s^2, выборочное среднее квадратическое отклонение $\sigma_{\text{в}}$, исправленное среднее квадратическое отклонение s.

4. С доверительной вероятностью $\gamma = 0.99$ найти доверительный интервал:

а) для математического ожидания $M(X)$ в случае известной дисперсии, предполагая $D(X) = s^2$,

б) для математического ожидания $M(X)$ в случае неизвестной дисперсии.

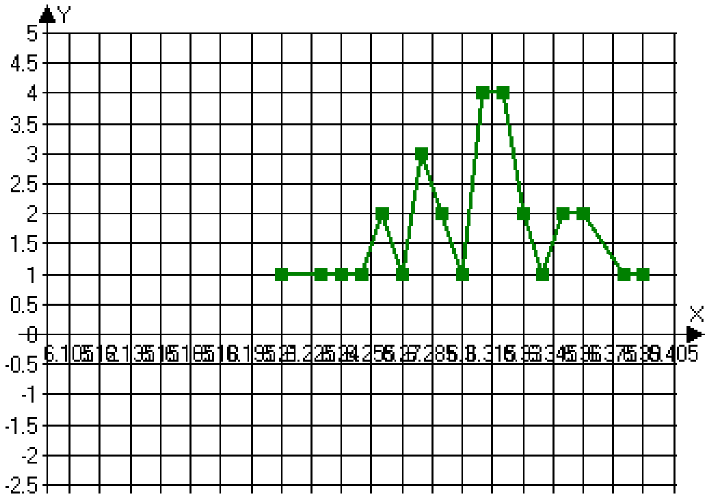
5. По выборке объёма $n = 9$, извлечённой из нормальной генеральной совокупности с известным средним квадратическим отклонением $\sigma = 4$, найдена выборочная средняя $\bar{x} = 16.5$. При уровне значимости 0.05 проверить нулевую гипотезу $H_0: M(X) = 15$ при конкурирующей гипотезе $H_1: M(X) > 15$.

Решение

1. Построим статистический ряд:

X	6.21	6.23	6.24	6.25	6.26	6.27	6.28	6.29	6.30	6.31	6.32	6.33	6.34	6.35	6.36	6.38	6.39
n	1	1	1	1	2	1	3	2	1	4	4	2	1	2	2	1	1

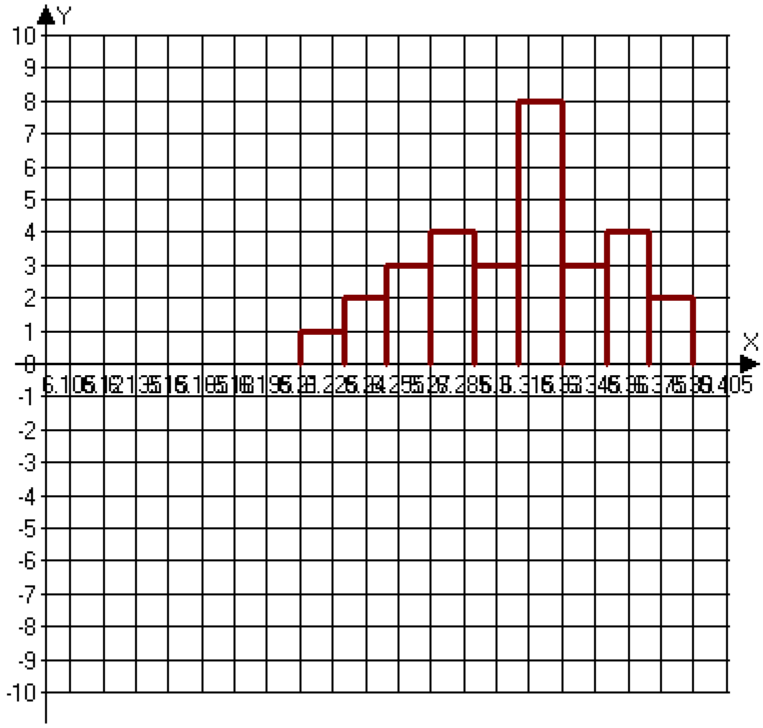
Построим полигон частот. Для этого на оси Ox отмечает варианты, а на оси Oy соответствующие частоты. Полученные точки соединяем точками:



2. Разобьём выборку на 9 интервалов. Получим:

Интервал	[6.21; 6.23)	[6.23; 6.25)	[6.25; 6.27)	[6.27; 6.29)	[6.29; 6.31)	[6.31; 6.33)	[6.33; 6.35)	[6.35; 6.37)	[6.37; 6.39]
Частота	1	2	3	4	3	8	3	4	2

Построим гистограмму частот. Для этого по оси Ox откладываем полученные интервалы, а по оси Oy – соответствующие частоты. Получим:



3. Дополним таблицу серединами интервалов:

Интервал	[6.21; 6.23)	[6.23; 6.25)	[6.25; 6.27)	[6.27; 6.29)	[6.29; 6.31)	[6.31; 6.33)	[6.33; 6.35)	[6.35; 6.37)	[6.37; 6.39]

Частота	1	2	3	4	3	8	3	4	2
Середина	6.22	6.24	6.26	6.28	6.30	6.32	6.34	6.36	6.38

Вычислим требуемые оценки по соответствующим формулам:

$$\bar{x} = \frac{1}{N} \cdot \sum_{i=1}^{9} x_i n_i =$$

$$= \frac{1}{30} \cdot (6.22 \cdot 1 + 6.24 \cdot 2 + 6.26 \cdot 3 + 6.28 \cdot 4 + 6.30 \cdot 3 + 6.32 \cdot 8 + 6.34 \cdot 3 + 6.36 \cdot 4 + 6.38 \cdot 2) = \frac{1}{30} \cdot 189.28 = 6.31$$

$$D_{\text{в}} = \frac{1}{N} \cdot \sum_{i=1}^{9} (x_i - x^*)^2 n_i =$$

$$= \frac{1}{30} \cdot ((6.22 - 6.31)^2 \cdot 1 + (6.24 - 6.31)^2 \cdot 2 + (6.26 - 6.31)^2 \cdot 3 + (6.28 - 6.31)^2 \cdot 4 + (6.30 - 6.31)^2 \cdot 3 + (6.32 - 6.31)^2 \cdot 8 + (6.34 - 6.31)^2 \cdot 3 + (6.36 - 6.31)^2 \cdot 4 + (6.38 - 6.31)^2 \cdot 2) = \frac{1}{30} \cdot 0.0526 = 0.0017$$

$$s^2 = \frac{n}{n-1} \cdot D_{\text{в}} = \frac{30}{29} \cdot 0.0017 = 0.0018$$

$$\sigma_{\text{в}} = \sqrt{D_{\text{в}}} = \sqrt{0.0017} = 0.041$$

$$s = \sqrt{s^2} = \sqrt{0.0018} = 0.042$$

4. а) Если дисперсия известна, то доверительный интервал будет равен:

$$\bar{x} - \frac{t \cdot \sigma}{\sqrt{n}} < a < \bar{x} + \frac{t \cdot \sigma}{\sqrt{n}}$$

По условию $n = 30$, $\sigma = 0.041$. t найдём из условия: $2\Phi(t) = \gamma = 0.99 \rightarrow \Phi(t) = 0.495 \rightarrow t = 2.58$. Тогда доверительный интервал будет равен:

$$6.31 - \frac{2.58 \cdot 0.041}{\sqrt{30}} < M(X) < 6331 + \frac{2.58 \cdot 0.041}{\sqrt{30}}$$

$$6.291 < M(X) < 6.329$$

б) Если средне квадратическое отклонение и дисперсия генеральной совокупности не известно, то доверительный интервал для математического ожидания ищется по формуле:

$$\bar{x} - t_\alpha \cdot \frac{D_{\text{в}}}{\sqrt{n}} < M(X) < \bar{x} + t_\alpha \cdot \frac{D_{\text{в}}}{\sqrt{n}}$$

t_α находится по таблице при заданном объёме выборки и доверительной вероятности (доверительная вероятность равна 0.99). Для данных задачи $t_\alpha = 2.75$. Получим:

$$6.31 - 2.75 \cdot \frac{0.0017}{\sqrt{30}} < M(X) < 6.31 + 2.75 \cdot \frac{0.0017}{\sqrt{30}}$$
$$6.309 < M(X) < 6.311$$

5. Для решения данной задачи необходимо проверить гипотезу о том, что неизвестная генеральная средняя нормальной совокупности точно равна числовому значению, когда дисперсия генеральной совокупности известна.

Сформулируем нулевую и конкурирующую гипотезы согласно условию задачи.

$H_0: a = M(X) = 15$ – неизвестная генеральная средняя нормально распределённой совокупности с известной дисперсией равна числовому значению.

$H_1: M(X) > 15$ – неизвестная генеральная средняя нормально распределённой совокупности с известной дисперсией больше числового значения.

Так как конкурирующая гипотеза – правосторонняя, то и критическая область – правосторонняя.

В качестве критерия для сравнения выборочной средней с гипотетической генеральной средней нормальной совокупности, когда дисперсия генеральной совокупности известна, используется случайная величина u.

Его наблюдаемое значение $u_{\text{набл}}$ рассчитывается по формуле:

$$u_{\text{набл}} = \frac{\bar{x} - M(X)}{\sigma_{\text{ген}}} \cdot \sqrt{n}$$

$\bar{x}$ – выборочная средняя, $M(X)$ – числовое значение генеральной средней, $\sigma_{\text{ген}}$ – выборочное среднее квадратическое отклонение, n – объём выборки. Подставляя данные задачи, получим:

$$u_{\text{набл}} = \frac{\bar{x} - M(X)}{\sigma_{\text{ген}}} \cdot \sqrt{n} = \frac{16.5 - 15}{4} \cdot \sqrt{9} = 1.125$$

Так как конкурирующая гипотеза – правосторонняя, то критическое значение и следует находить по таблице функции Лапласа из равенства

$\Phi_0(u_{\text{кр}}) = \frac{1-2\alpha}{2}$.

По условию $\alpha = 0.05$, тогда $\frac{1-2\alpha}{2} = \frac{1-2\cdot 0.05}{2} = 0.45$. Откуда

$$\Phi_0(u_{\text{кр}}) = 0.45 \to u_{\text{кр}} = 1.65$$

Т.к. $1.125 < 1.65$, то $u_{\text{набл}} < u_{\text{кр}}$, поэтому на данном уровне значимости нулевая гипотеза принимается. Это означает, что на уровне значимости 0.05 нулевая гипотеза H_0 принимается.