

**Практическое занятие №2.**  
**ТОЧЕЧНЫЕ ОЦЕНКИ ОСНОВНЫХ ЧИСЛОВЫХ ХАРАКТЕРИСТИК И**  
**ПАРАМЕТРОВ ГЕНЕРАЛЬНОЙ СОВОКУПНОСТИ**

**Цель занятия.** Уметь рассчитывать точечные оценки числовых характеристик и параметров генеральной совокупности в MS Excel.

**Учебные вопросы:**

1. Точечная оценка математического ожидания.
2. Точечная оценка дисперсии и среднего квадратичного отклонения.
3. Точечная оценка коэффициента корреляции.
4. Метод моментов оценки параметров распределения

В MS Excel для расчета точечных оценок можно использовать встроенные функции (категории статистические), расчет по формулам в электронных ячейках (эти возможности продемонстрированы в задании 1) и надстройку **Пакет анализа. Описательная статистика** (задание 2).

Для удобства работы с диапазонами ячеек рекомендуется (как и в предыдущей лабораторной работе) присвоить диапазону имя.

**Задание 1**

Для заданной выборки из генеральной совокупности рассчитать точечные оценки параметров закона распределения вероятностей, используя встроенные функции и расчет по формулам в электронных ячейках. Проанализировать полученные результаты.

*Решение.*

Наблюдения помещены в диапазон ячеек Excel A1:E27. Для наглядности результаты представлены в табл. 1.5.

Точечные оценки параметров закона распределения можно разделить на 4 группы показателей:

положения – максимальное, минимальное значения, средние величины (средняя арифметическая, гармоническая, геометрическая); вышеназванные показатели характеризуют расположение данных на числовой оси;

структуры – мода и медиана, которые отражают более полные особенности совокупности;

разброса – дисперсия, стандартное отклонение, размах выборки, среднее линейное отклонение, коэффициенты вариации, осцилляции и линейный коэффициент вариации, они описывают степень изменчивости данных относительно своего центра;

вида закона распределения – асимметрия и эксцесс, характеризующие такие его особенности, как симметричность, сходство с нормальным законом распределения вероятностей.

Таблица 1.5

	A	B	C	D	E
1	Числовые характеристики выборки				
2	<i>Выборка</i>	<i>Характеристики</i>			
3	7,36	Количество элементов			25
4	6,605	<i>Показатели положения</i>			
5	11,029	Максимальное значение			86,068
6	14,653	Минимальное значение			6,605
7	16,237	Средняя арифметическая			37,072
8	19,626	Средняя гармоническая			22,510
9	13,493	Средняя геометрическая			29,411
10	19,672	<i>Показатели структуры</i>			
11	24,79	Мода			27,679
12	23,067	Медиана			27,679
13	26,68	<i>Показатели разброса</i>			
14	27,679	Дисперсия			589,195
15	30,546	Стандартное отклонение			24,273
16	35,645	Размах выборки			79,463
17	39,593	Среднее линейное отклонение			20,105
18	40,624	Коэффициент вариации			65,477
19	47,624	Коэффициент осцилляции			2,143
20	52,032	Линейный коэффициент вариации			0,542
21	27,679	<i>Показатели вида закона распределения</i>			
22	60,809	Коэффициент асимметрии			0,720
23	65,506	Коэффициент эксцесса			-0,669
24	70,22				
25	78,625				
26	80,929				
27	86,068				

В табл. 1.6 приведены математические формулы и соответствующие функции Excel (или формулы Excel) для расчета числовых характеристик выборки (диапазону ячеек с данными A3:A27 присвоено имя Y). В ячейке F3 находится формула =ABS(A3-\$E\$7),..., в ячейке F27 – формула =ABS(A27-\$E\$7).

Таблица 1.6

Наименование характеристики	Математическая запись или формула	Функция (формула для вычислений) в MS Excel
Количество элементов	$n$	=СЧЁТ(Y)
Максимальное значение	$\max(Y)$	=МАКС(Y)
Минимальное значение	$\min(Y)$	=МИН(Y)
Средняя арифметическая	$\bar{X} = \frac{1}{n} \sum_{i=1}^n x_i$	=СРЗНАЧ(Y) или =СУММ(Y)/СЧЕТ(Y)

Средняя гармоническая	$\bar{x}_{гар} = \sum n_i / \sum \frac{n_i}{x_i} = n / \sum \frac{n_i}{x_i}$ ,	=СРГАРМ(Y)
Средняя геометрическая	$\bar{x}_{геом} = \sqrt[n]{x_1 \cdot x_2 \dots x_n}$ ,	=СРГЕОМ(Y)
Мода	$N_{M_0} = \max\{n_i\}$	=МОДА(Y) или =МОДА.ОДН(Y)
Медиана	$N_{M_e} = \frac{n+1}{2}$ ,	=МЕДИАНА(Y)
Дисперсия	$\tilde{S}_x^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{X})^2$	=ДИСП(Y) или =ДИСП.В(Y)
Стандартное отклонение	$\tilde{S}_x = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{X})^2} = \sqrt{\tilde{S}_x^2}$	=СТАНДОТКЛОН(Y) или =СТАНДОТКЛОН.В(Y) или =КОРЕНЬ(ДИСП(Y))
Размах выборки	$R = \max(Y) - \min(Y)$	= МАКС(Y) – МИН(Y)
Среднее линейное отклонение	$d = \frac{\sum  x - \bar{x} }{n}$	=СРОТКЛ(Y) или =СУММ(F3:F27)/ СЧЕТ(Y)
Коэффициент вариации	$V = \frac{\tilde{S}}{\bar{x}} \cdot 100\%$ .	=СТАНДОТКЛОН(Y)/СРЗНАЧ(Y)
Коэффициент осцилляции	$V_R = \frac{R}{\bar{x}}$	=R/СРЗНАЧ(Y)
Линейный коэффициент вариации	$V_d = \frac{d}{\bar{x}}$	=d/ СРЗНАЧ(Y)
Коэффициент асимметрии	$\hat{k}_a = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{X})^3 / S_x^3$ ,	=СКОС(Y)
Коэффициент эксцесса	$\hat{k}_3 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{X})^4 / S_x^4 - 3$	=ЭКСЦЕСС(Y)

Полученные результаты расчета соответствуют правилу мажорантности средних, т. е.  $\bar{x}_{гарм.} \leq \bar{x}_{геом.} \leq \bar{x}_{арифм.}$ .

Коэффициент вариации  $V_0$ , больший 60 %, что свидетельствует о большой колеблемости результатов наблюдений в выборочной совокупности (если  $V_0 \leq 33$  %, то совокупность считается однородной). Ненадежность средней арифметической в выборке подтверждается также и ее значительным отклонением от медианы:  $37,077 - 27,68 = 9,39$  (25,34 %).

Для данной выборки  $\bar{x} > M_e \geq M_0$ , коэффициент асимметрии положителен, но его значение невелико, поэтому имеет место незначительная правосторонняя асимметрия. Коэффициент эксцесса отрицателен, поэтому распределение является немного более островершинным по сравнению с ПРВ нормального распределения.

## Задание 2

Для данных предыдущего задания (*Выборка*. Ячейки A2:A27) определить точечные оценки параметров закона распределения вероятностей, используя надстройку MS Excel **Пакет анализа. Описательная статистика**.

*Решение.*

В меню **Данные** выбрать **Пакет анализа. Описательная статистика**. В открывшемся диалоговом окне следует:

1. Задать входной диапазон, т. е. выделить диапазон ячеек A3:A27 (или указать диапазон Y); для наглядности результатов можно вместе с диапазоном задать название столбца из ячейки A2 (метка в 1-й строке).

2. Задать группирование – для одного столбца можно не указывать, для нескольких столбцов следует указать, будет ли группирование осуществляться по столбцам или строкам.

3. Поставить при необходимости флажок «метки в первой строке или метки в первом столбце». Для рассматриваемого примера, если диапазон ячеек A3:A27, то флажок не ставится, если A2:A27, то ставится.

4. Для указания выходного диапазона нужно активировать переключатель «Выходной диапазон», поставить указатель манипулятора «Мышь» в поле ввода «Выходной диапазон», указать левую верхнюю ячейку выходного диапазона; можно выбрать «Новый рабочий лист/Новая рабочая книга».

5. Переключатель «Итоговая статистика» установить в активное состояние.

6. Задать требуемый уровень надежности оценок. По умолчанию задано  $\beta = 95 \%$ :

7. «K-й наибольший» – установить в активное состояние, если в выходную таблицу необходимо включить строку для  $k$ -го наибольшего (начиная с максимума  $x_{\max}$ ) значения элемента выборки. В поле, расположенное напротив флажка, ввести число  $k$ . Если  $k = 1$ , то строка будет содержать максимальное значение элемента выборки.

8. «K-й наименьший» – установить в активное состояние, если в выходную таблицу необходимо включить строку для  $k$ -го наименьшего (начиная с минимума  $x_{\min}$ ) значения элемента выборки. В поле, расположенное напротив флажка, ввести число  $k$ . Если  $k = 1$ , то строка будет содержать минимальное значение элемента выборки.

В результате работы инструмента «Описательная статистика» (табл. 1.7) в заданном выходном диапазоне для каждого столбца данных будут выведены статистические характеристики: среднее, стандартная ошибка среднего, равная  $S/\sqrt{n}$ , медиана, мода, стандартное отклонение, дисперсия выборки, эксцесс, асимметричность, интервал = максимум-минимум, сумма, счет, наибольшее, наименьшее, уровень надежности.

Таблица 1.7

	А	В	С
1	Описательная статистика		
2	<i>Выборка</i>	<i>Столбец 1</i>	
3	7,36		
4	6,605	Среднее	37,072
5	11,029	Стандартная ошибка	4,855

6	14,653	Медиана	27,679
7	16,237	Мода	27,679
8	19,626	Стандартное отклонение	24,273
9	13,493	Дисперсия выборки	589,195
10	19,672	Экссесс	-0,669
11	24,79	Асимметричность	0,720
12	23,067	Интервал	79,463
13	26,68	Минимум	6,605
	A	B	C
14	27,679	Максимум	86,068
15	30,546	Сумма	926,791
16	35,645	Счет	25,000
17	39,593	Наибольший (1)	86,068
18	40,624	Наименьший (1)	6,605
19	47,624	Уровень надежности (95,0 %)	10,020
20	52,032		
21	27,679		
22	60,809		
23	65,506		
24	70,22		
25	78,625		
26	80,929		
27	86,068		

Анализ полученных результатов показывает, что аналогичные характеристики, полученные двумя рассмотренными методами, совпадают. На основании рассчитанных по выборке показателей описательной статистики с уровнем надежности 95 % можно предположить, что среднее значение выборки находится в пределах от 27,052 до 47,092 (левая граница доверительного интервала – 37,072 – 10,020, правая граница – 37,072 + 10,020).

## ЗАДАНИЯ ДЛЯ САМОСТОЯТЕЛЬНОГО РЕШЕНИЯ ПО ТЕМЕ «ТОЧЕЧНЫЕ ОЦЕНКИ»

### Задание 1.

Найти эмпирическую ИФР по заданной таблице частот. Построить график:

а) 

$x_i$	0	2	4	6	8	10	12
$n_i$	5	10	15	20	18	12	7

б) 

$x_i$	-1	1	2	4	7	9	11
$n_i$	6	7	2	2	3	4	5

### Задание 2.

Число пятерок, четверок, троек и двоек, полученных на экзамене, распределилось следующим образом:

$x_i$ (балл)	2	3	4	5
$n_i$ (число оценок)	3	13	7	2

Найти эмпирическую ИФР балла и построить ее график. Найти частоту того, что балл будет в границах от 3,5 до 4,8.

### Задание 3.

Бизнесмен с целью увеличения прибыли занялся анализом статистического материала, касающегося продажи обуви:

Размер обуви	36	37	38	39	40	41	42	43	44
Число проданных пар в %	6	8	15	25	20	10	8	5	3

Построить эмпирическую ИФР, гистограмму и полигон частот размера проданной обуви. Найти частоту продажи обуви 42-44 размеров.

### Задание 4.

По таблице частот определить оценки математического ожидания, дисперсии и среднего квадратического отклонения:

$x_i$	1	3	6	16	26
$n_i$	8	40	10	2	1

### Задание 5.

В результате 5 измерений получены следующие данные: 91, 94, 103, 105, 108. Найти оценки математического ожидания, дисперсии и среднего квадратического отклонения.

### Задание 6.

Найти оценки математического ожидания, выборочной и исправленной дисперсии по таблице частот:

$x_i$	186-192	192-198	198-204	204-210	210-216	216-220
$n_i$	2	5	8	11	7	4

**Задание 10.**

При тестировании обучаемых на внимательность 5 человек набрали более 11 баллов (оказались удивительно внимательными и проницательными), 10 человек получили от 5-ти до 10-ти баллов (оказались достаточно внимательными) и 3 человека оказались очень рассеянными и набрали 4 балла и менее.

Построить ИФР, гистограмму и полигон частот балла.

**Задание 11.**

В следующей таблице приведены статистические данные о поголовье скота в крестьянских хозяйствах Тверской губернии в XIX веке в пересчете на одно крестьянское хозяйство.

Год	1800	1820	1840	1860	1880	1900
Поголовье	2,5	2,3	2,0	1,9	1,6	1,5

Найти оценки математического ожидания, дисперсии и СКО поголовья скота, приходящегося на одно крестьянское хозяйство в период с 1800-ого по 1900-ый год.

**Задание 12.**

При социологическом опросе в г. Энске оказалось, что 15 человек получают заработную плату в размере 2500, 30 – 3500 рублей, 15 – 6000 рублей, 15 – 9000 рублей, 10 – 11000 рублей, 10 – 12000 рублей, 5 – 16000 рублей.

Оценить среднюю зарплату, ее дисперсию и СКО.

**Задание 13.**

В следующей таблице приведены данные в % числа участвующих в революционном движении в 70-х годах XIX века в России согласно возрастному показателю.

Возраст (лет)	16-20	20-25	25-30	30-35	35-40	40-45	Более 45
Число участников в %	37	45	13	3	1	1	0

Построить ИФР, гистограмму и полигон частот возраста участников. Какова частота участия молодых людей в возрасте до 25 лет?

**Задание 14.**

В следующей таблице приведены статистические данные о поголовье скота в крестьянских хозяйствах Тверской губернии в XIX веке в пересчете на одно крестьянское хозяйство.

Год	1800	1820	1840	1860	1880	1900
Поголовье	2,5	2,3	2,0	1,9	1,6	1,5

Найти оценки математического ожидания, дисперсии и СКО поголовья скота, приходящегося на одно крестьянское хозяйство в период с 1800-ого по 1900-ый год.

**Задание 15.**

При социологическом опросе в г. Энске оказалось, что 15 человек получают заработную плату в размере 2500, 30 – 3500 рублей, 15 – 6000 рублей, 15 – 9000 рублей, 10 – 11000 рублей, 10 – 12000 рублей, 5 – 16000 рублей.

Оценить среднюю зарплату, ее дисперсию и СКО.

**Задание 16.**

По выборке объема  $N=41$  найдена смещенная оценка генеральной дисперсии. Найти несмещенную оценку дисперсии генеральной совокупности.

**Задание 17.**

Имеются данные о еженедельном количестве проданных компьютеров одной из фирм: 398, 412, 560, 474, 544, 690, 587, 600, 613, 457, 504, 477, 530, 641, 359, 566, 452, 633, 474, 499, 580, 606, 344, 455, 505, 396, 347, 441, 390, 632, 400, 582. Составьте вариационный ряд. Найдите среднее количество проданных компьютеров. Рассчитайте показатели вариации.

**Задание 18.**

Число пассажиров одного из рейсов за 30 дней составило: 128, 121, 134, 118, 123, 109, 120, 116, 125, 128, 121, 129, 130, 131, 127, 119, 114, 124, 110, 126, 134, 125, 128, 123, 128, 133, 132, 136, 134, 129. Составьте вариационный ряд. Найдите среднее число пассажиров в рейсе. Рассчитайте показатели вариации. Сделайте анализ полученных результатов.