**Лабораторная работа № 2**

**Первичный анализ данных в системе Excel** **и STATISTICA**

Первичный анализ статистических данных в EXCEL

*Расчёт данных*

Формат представления исходных данных для выполнения первичного статистического анализа выполняется записью информации в 1 строку или 1 столбец (таблица 1).

Таблица 1 – Исходные данные

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| xi | 7.1 | 7.7 | 3,6 | 8,3 | 8,8 | 10,4 | 8,9 | 9 | 8,9 | 14 | 9,7 | 9,4 | 8,5 | 15,9 | 12,6 | 9,1 | 6,2 | 10,7 | 13,8 | 13,6 | 15,2 | 3,4 | 9,3 | 13,3 | 6,7 | 7,9 | 4,9 | 4,5 | 8 | 17,1 |

Для корректной статистической обработки необходимо определить величину класса по формуле:

**i = (Xmax – Xmin )/(1+3,322\*lgN)**, где N–число наблюдений.

В данном случае: i = (17,1-3,4)/(1+3,322\*lg30)=2,32.

Таблица 2 – Размер интервалов

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| классы | 3,4 | 5,7 | 8,0 | 10,4 | 12,7 | 15,0 | 17,3 |

**Шаг 1**. В системе Excel в меню **Сервис** открыть модуль **Анализ данных** (рис. 1).



*Рис. 1*

**Шаг 2.** В модуле **Анализ данных** выбрать **Описательная статистика**, после чего щелкнуть мышкой **OK** (рис. 2). В появившемся окне выполнить операции и установки, как показано на рис. 3. Щелкнуть мышкой **OK**. Результат обработки появится в указанном поле (выходной интервал $C$1). В таблице 3 показаны результаты статистической обработки.



*Рис. 2 – Окно Анализа данных*



*Рис. 3 – Стартовая панель*

Таблица 3 – Описательная статистика (результат обработки)

|  |  |
| --- | --- |
| **Параметры** | **Значения** |
| Стандартная ошибка | 0,650247521 |
| Мода | 8,9 |
| Дисперсия выборки | 12,68465517 |
| Асимметричность | 0,338895597 |
| Минимум | 3,4 |
| Сумма | 286,5 |
| Среднее | 9,55 |
| Медиана | 8,95 |
| Стандартное отклонение | 3,561552354 |
| Эксцесс | -0,37736441 |
| Интервал | 13,7 |
| Максимум | 17,1 |
| Счет | 30 |

*Построение гистограммы*

Вернуться в модуль **Анализ данных** выбрать опцию **Гистограмма**, после чего щелкнуть мышкой **OK**. В появившемся окне выполнить операции и установки, как показано на рис.4. Щелкнуть мышкой **OK**. Результат обработки появится в указанном поле (выходной интервал $G$1, рис.5).



*Рис. 4 – Стартовая панель*



*Рис. 5 – Гистограмма*

 **Задание 1.** Измерили длину хвоста у оленьих мышей в возрате одного года:

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 58 | 57 | 64 | 61 | 56 | 65 | 63 | 58 | 63 |
| 60 | 59 | 61 | 54 | 58 | 66 | 67 | 63 | 63 |
| 61 | 60 | 58 | 57 | 65 | 61 | 60 | 68 | 64 |
| 63 | 56 | 59 | 64 | 61 | 64 | 57 | 60 | 63 |
| 58 | 52 | 60 | 59 | 57 | 61 | 54 | 58 | 64 |
| 62 | 59 | 60 | 63 | 60 | 60 | 64 | 59 | 63 |
| 63 | 59 | 62 | 63 | 61 | 65 | 61 | 64 | 57 |
| 59 | 54 | 64 | 63 | 57 | 59 | 59 | 58 | 63 |
| 62 | 63 | 62 | 62 | 60 | 62 | 57 | 56 | 60 |
| 63 | 57 | 63 | 61 | 59 | 61 | 59 | 60 |  |

Рассчитайте параметры описательной статистки и постройте гистограмму.

**Задание 2.** Ширина верхнего последнео коренного зуба у ископаемого млекопитающего была следующей:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 5,8 | 6,2 | 6,3 | 6,1 | 6,1 | 5,7 |
| 6,5 | 6,0 | 6,1 | 5,8 | 6,3 | 62 |
| 62 | 5,4 | 5,9 | 6,0 | 5,7 | 5,9 |
| 6,1 | 6,7 | 6,2 | 6,5 | 6,2 | 6,1 |
| 6,2 | 5,7 | 6,1 | 5,7 | 5,9 | 6,0 |
| 5,7 | 5,9 | 6,1 | 5,9 | 6,0 | 6,1 |

Рассчитайте параметры описательной статистки и постройте гистограмму.

Первичный анализ данных в системе STATISTICA

Запустите модуль **Basic Statistics/Tables (основные статистики/таблицы)** из **Переключателя модулей**. Высветите в стартовой панели модуля Basic Statistics/Tables (основные статистики/таблицы) строку Descriptive statistics (описательная статистика) (рис. 6, 7):

 

|  |  |
| --- | --- |
| *Рис. 6 – Основные статистики/таблицы* | *Рис. 7 – Описательная статистика* |

**Шаг 1**. Нажмите кнопку **OK**. Перед вами откроется окно **Descriptive statistics (описательная статистика)** (рис. 8, 9).



*Рис. 8 – Предварительные установки*

Выполните установки, как показано на рис. 8 и 9.

**Шаг 2**. Загрузите в систему STATISTICA исходные данные (рис. 10)



|  |  |
| --- | --- |
| *Рис. 9 – Предварительные установки* | *Рис. 10* |

**Шаг 3.** Возвратитесь в окно **Descriptive statistics (описательная статистика)**. В появившемся окне выберите переменные для анализа (рис. 11). Выбор переменных осуществляется с помощью кнопки **Variables (переменные)**. Нажав кнопку **Variables** в окне (рис. 11) выберите переменные (в данном случае X и Y) (рис. 12).

 

|  |  |
| --- | --- |
| *Рис. 11 – Диалоговое окно результатов* | *Рис. 12 – Окно выбора переменных* |

**Шаг 4.** В диалоговом окне результатов последовательно нажмите кнопки:

− **Summary Descriptive statistics (описательная статистика)**;

− **Frequency Tables (таблица частот)**;

− **Histograms (гистограммы)**.

Результаты статистической обработки представлены на рис. 13–17. Из представленной обработки следует, что массив данных X описывается функцией нормального распределения, а Y – нет, что достаточно хорошо видно из рис. 16, 17. Более подробную информацию можно извлечь из анализа результатов обработки, представленных в таблицах (рис. 13–15).



*Рис. 13*



*Рис. 14*



*Рис. 15*



*Рис. 16*



*Рис. 17*

**Задание 3**. Количество птенцов в гнездах лесной ласточки *Iridoprocne bicolor* было следующим:

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 4 | 5 | 4 | 5 | 5 | 4 | 5 | 4 | 3 | 5 | 4 | 5 |
| 6 | 1 | 6 | 4 | 4 | 4 | 5 | 5 | 3 | 5 | 5 | 4 |
| 6 | 4 | 6 | 2 | 3 | 4 | 5 | 5 | 5 | 5 | 5 | 5 |
| 4 | 5 | 5 | 6 | 4 | 6 | 2 | 5 | 5 | 3 | 5 | 5 |
| 5 | 4 | 6 | 4 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 |
| 5 | 5 | 4 | 6 | 7 | 6 | 3 | 5 | 5 | 6 | 5 | 5 |
| 5 | 4 | 4 | 2 | 4 | 4 | 6 | 2 | 6 | 5 | 4 | 5 |
| 5 | 5 | 5 | 5 | 4 | 5 | 4 | 6 | 5 | 4 | 7 | 5 |
| 5 | 5 | 6 | 6 | 4 | 4 | 4 | 6 | 5 | 4 | 3 | 5 |
| 5 | 7 | 5 | 5 | 5 | 5 | 4 | 3 | 7 | 6 | 4 | 4 |

Рассчитайте параметры описательной статистки и постройте гистограмму.

**Задание 4**. Изучен живой вес 63 телят холмогорских помесей при рождении (в кг):

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 27 | 32 | 32 | 31 | 32 | 28 | 37 | 35 | 26 | 28 |
| 32 | 39 | 34 | 30 | 37 | 26 | 27 | 40 | 35 | 37 |
| 28 | 43 | 26 | 35 | 45 | 26 | 35 | 32 | 32 | 35 |
| 35 | 28 | 32 | 36 | 32 | 36 | 37 | 33 | 23 | 31 |
| 36 | 33 | 33 | 28 | 23 | 26 | 34 | 32 | 36 | 27 |
| 32 | 39 | 30 | 30 | 36 | 38 | 24 | 32 | 30 | 31 |
| 28 | 36 | 36 |  |  |  |  |  |  |  |

Рассчитайте параметры описательной статистки и постройте гистограмму.

**Задание 5.** Имеются следующие данные о росте (в см) взрослых мужчин:

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 162 | 151 | 161 | 170 | 167 | 164 | 166 | 164 | 173 | 172 |
| 165 | 153 | 164 | 169 | 170 | 154 | 163 | 159 | 161 | 167 |
| 168 | 164 | 170 | 166 | 176 | 157 | 159 | 158 | 160 | 161 |
| 167 | 155 | 166 | 167 | 173 | 165 | 175 | 165 | 174 | 167 |
| 170 | 169 | 159 | 159 | 160 | 156 | 161 | 162 | 161 | 181 |
| 159 | 169 | 160 | 169 | 161 | 161 | 166 | 164 | 170 | 180 |
| 158 | 167 | 169 | 165 | 166 | 172 | 168 | 171 | 178 | 178 |
| 171 | 165 | 161 | 162 | 182 | 164 | 171 | 169 | 176 | 177 |
| 170 | 169 | 171 | 160 | 165 | 165 | 179 | 161 | 178 | 173 |

Рассчитайте параметры описательной статистки и постройте гистограмму.

**Задание 6.** Обработайте следующие данные о длине третьего верхнего предкоренного зуба у 21 экземпляра ископаемого млекопитающего *Ptilodus montanus* (в мм):

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 3,2 | 2,8 | 2,9 | 3,0 | 3,1 | 3,3 | 2,9 |
| 3,1 | 2,7 | 3,4 | 2,9 | 3,0 | 2,9 | 2,8 |
| 2,6 | 3,0 | 2,8 | 3,0 | 3,1 | 2,9 | 3,0 |

Рассчитайте параметры описательной статистки и постройте гистограмму.