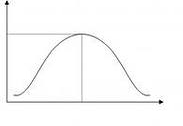
**Теоретический материал «Законы распределение случайной величины» диск с заданиями Назина Н.Б\Статистические методы и модели управления\Законы распределение случайной величины\законы распределения случайной величины..pdf**

**Лабораторная работа № 6.**

**Тема: Свойства нормального закона распределения. Проверка распределения на «нормальность».**

**Для непрерывной случайной величины наиболее популярным является нормальный закон**

**распределения.**

****

**Для нормального закона распределения характерно:**

1. **Среднее арифметическое, мода и медиана равны.**
2. **Асимметрия и эксцесс равны (приблизительно) 0.**
3. **Наложенная на гистограмму линия тренда приближена к графику нормального распределения.**
4. **68% всех его наблюдений лежат в диапазоне  (стандартное отклонение), а в диапазоне  содержит 95% значений, в диапазоне  - 98% значений.**

**Если хотя бы три из четырех свойств выполняются, распределение можно считать нормальным.**

**Пример**

**Теория**

**Выборка** – это группа элементов, выбранная для исследования из всей совокупности элементов (из генеральной совокупности). Задача выборочного метода состоит в том, чтобы сделать правильный вывод относительно всего собрания элементов (генеральной совокупности). Для этого выборка должна быть репрезентативной. Обычно используют случайный метод выбора элементов из генеральной совокупности.

**Пример**. Имеются данные о количественном составе 60 семей. Описать дискретную случайную величину (ДСВ): «количественный состав семьи» (найти основные выборочные характеристики).

Занести данные в документ.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| |  | | --- | | Количественный состав семьи | | 5 | | 3 | | 3 | | 3 | | 5 | | 3 | | 4 | | 4 | | 4 | | 6 | | 5 | | 3 | | 6 | | 3 | | 3 | | 4 | | 3 | | 5 | | 3 | | 4 | | 5 | | 3 | | 4 | | 4 | | 3 | | 4 | | 3 | | 3 | | 4 | | 3 | | 5 | | 3 | | 2 | | 4 | | 3 | | 4 | | 4 | | 5 | | 3 | | 1 | | 6 | | 4 | | 6 | | 4 | | 4 | | 5 | | 4 | | 6 | | 4 | | 2 | | 2 | | 2 | | 3 | | 5 | | 4 | | 2 | | 4 | | 4 | | 5 | | 4 | | Для описания случайной величины необходимо найти основные выборочные характеристики:   * 1. среднее значение (математическое ожидание),   2. границы изменения случайной величины (минимум и максимум),   3. наиболее часто встречаемое значение (мода).   4. Интервал, в который попадает наибольшее значение случайной величины (границы считаются: среднее – стандартное отклонение; среднее + стандартное отклонение)   Это можно сделать при помощи функций СРЗНАЧ, МИН, МАКС, СТАНДОТКЛОН, МОДА, ЭКСЦЕСС, СКОС. В том случае, если выборки расположены не в один столбец (строку).  Такие же результаты можно получить с использованием **Описательной статистики в Пакете анализа.**  Для установления Пакет анализа нажмите кнопку Office – Параметры Excel (рис.1)- Надстройки – Управление –Надстройки Excel кнопка Перейти – Пакет анализа (поставить галочку).      В пункте Данные выберите Анализ данных.  Для определения характеристик выборки используется процедура **Описательная статистика**. В появившемся диалоговом окне указать входной диапазон, т. е. вести ссылку на ячейки, содержащие анализируемые данные при помощи мыши;   * Если в выделенном входном диапазоне присутствуют заголовки столбцов или строк, необходимо поставить флажок в **Метки в первой строке (столбце).** * указать выходной диапазон, в который будут выведены результаты анализа. Для этого следует поставить переключатель в положение **Выходной диапазон**, навести указатель мыши в поле ввода **Выходной диапазон** и щелкнуть левой кнопкой мыши, затем указатель мыши навести на верхнюю левую ячейку выходного диапазона и щелкнуть левой кнопкой мыши (т.е. выделить ***одну ячейку на пустом месте листа***); * в разделе **Группировка** переключатель установить в положение **по столбцам**; * установить флажок в поле **Итоговая статистика**; * нажать **ОК**.  |  |  | | --- | --- | | *Количественный состав семьи* | | | Среднее | 3,816667 | | Стандартная ошибка | 0,145539 | | Медиана | 4 | | Мода | 4 | | Стандартное отклонение | 1,127344 | | Дисперсия выборки | 1,270904 | | Эксцесс | -0,15392 | | Асимметричность | 0,080898 | | Интервал | 5 | | Минимум | 1 | | Максимум | 6 | | Сумма | 229 | | Счет | 60 | |

**Вывод: Характеристики ДСВ «Количественный состав семьи»:**

**Среднее значение – 3,8±0,14 (±ошибка среднего), минимальный размер семьи 1 человек, максимальный – 6, чаще всего встречаются семьи с 4 членами, большая часть семей имеют в своем составе от 2,7 (3,8-1,13 (среднее – стандартное отклонение) до 4,9 человек (среднее + стандартное отклонение).**

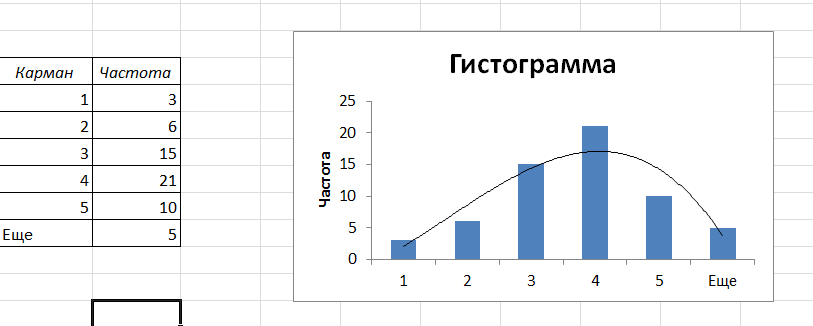
**Задание. По данным примера лабораторной работы проверить подчиняется ли случайная величина (количество человек в семье) нормальному закону распределения.**

**Данные (количество людей в семье) необходимо скопировать в новый документ и провести их обработку при помощи описательной статистики. Результат обработки представлен в таблице 1.**

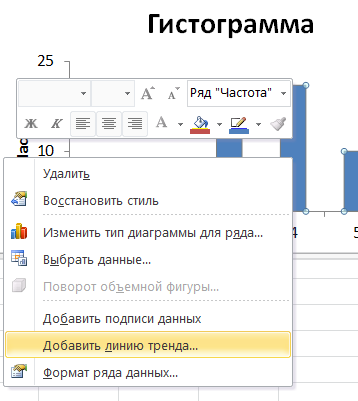
**Таблица 1**

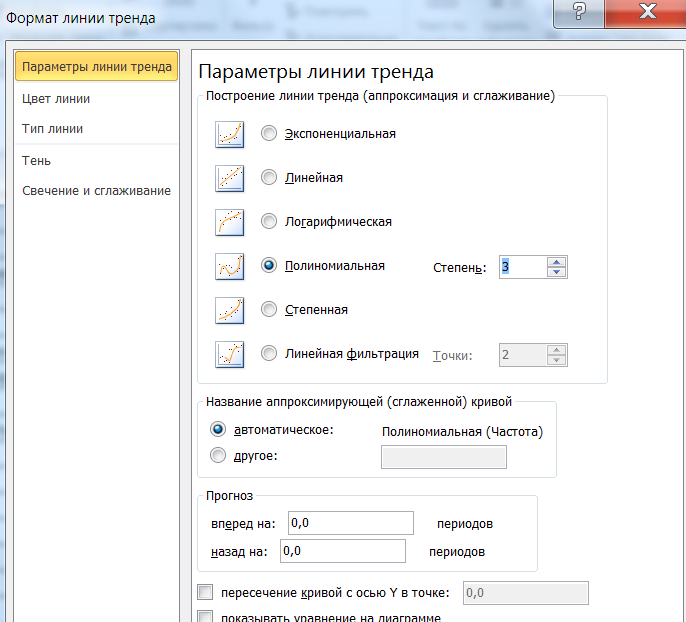
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| |  |  | | --- | --- | | ***Количественный состав семьи*** |  | | **Среднее** | **3,816667** | | **Стандартная ошибка** | **0,145539** | | **Медиана** | **4** | | **Мода** | **4** | | **Стандартное отклонение** | **1,127344** | | **Дисперсия выборки** | **1,270904** | | **Эксцесс** | **-0,15392** | | **Асимметричность** | **0,080898** | | **Интервал** | **5** | | **Минимум** | **1** | | **Максимум** | **6** | | **Сумма** | **229** | | **Счет** | **60** | | **Вывод по пунктам 1 и 2. Среднее, медиана и мода практически равны. Асимметрия и эксцесс отличаются от 0, но не существенно.** |

**Пункт 3. Построить гистограмму (воспользоваться инструментом анализа Гистограмма из Пакета анализа). Интервал карманов заполнить самостоятельно (1 2 3 4 5 потому что мы обрабатываем число людей в семье). Если интервал карманов не будет заполнен, Excel «сделает» это сам.**

****

**На гистограмму поместить линию тренда полиномиальную степени 3.**

****

****

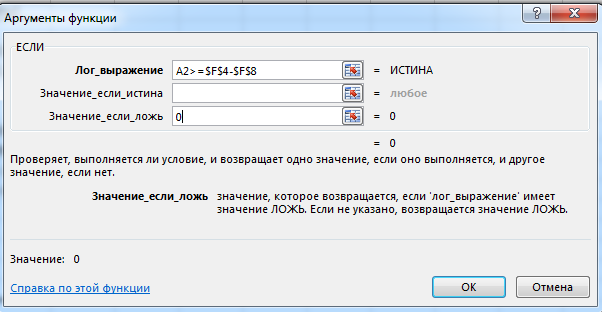
**Вывод по пункту 3 Линия тренда напоминает график нормального распределения.**

***Пункт 4.* Определить какой процент всех наблюдений лежит в диапазоне ,  . Для этого воспользоваться логической функцией ЕСЛИ. Примерный синтаксис.**

**«=ЕСЛИ(A2>=$F$4-$F$8;ЕСЛИ(A2<$F$4+$F$8;1;0);0)», где А2 – адрес первого значения исходных данных, в F4 – среднее, в F8- стандартное отклонение**

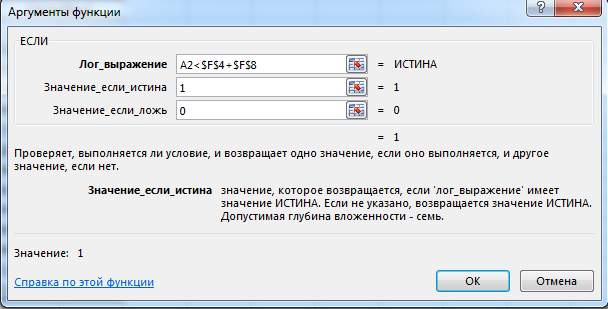
**Здесь необходимо использовать вложение функции ЕСЛИ:**

**Первое окно**

****

**Ставим курсор в значение\_если\_истина и снова вызываем функцию ЕСЛИ**

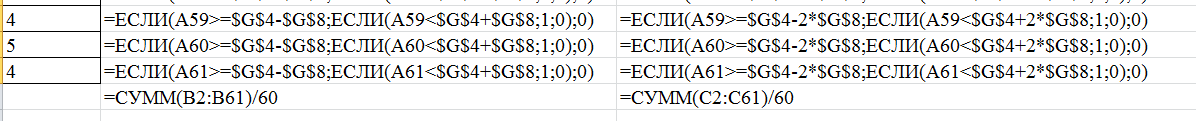
**Второе окно**

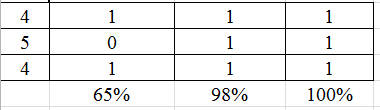
****

**Результатом работы функции будет получение 1, если значение попадает в интервал и 0, если не попадает.**

**Далее посчитать сумму единиц и найти процентное соотношение =СУММ(B2:B61)/60.**

**Так же просчитать для интервалов  и **

****

****

**Вывод по пункту 4. В результате получилось, что в интервал  попадает 65% данных, в интервал  - 98% данных, а в интервал  100% данных.**

**Сделать вывод о правильности вашего первоначального предположения о близости полученного распределения к нормальному.**

**Итоговый вывод: Данные приближаются к нормальному закону распределения, т.к.**

* **среднее=моде=медиане;**
* **асимметрия и эксцесс близки к нулю,**
* **наложенная на гистограмму линия тренда имеет вид графика нормального распределения**
* **в интервал  попадает 65% данных, в интервал  - 98% данных, а в интервал  100% данных**

**Самостоятельная работа.**

1. **Проверьте, подчиняются ли нормальному закону распределения данные по весу мужчин (Для обработки\БД\_ЛПУ).**
2. **Проверьте, подчиняются ли нормальному закону распределения данные по росту мужчин (Для обработки\БД\_ЛПУ).**
3. **Проверить, подчиняются ли данные нормальному закону распределения данные о выбросе загрязняющих веществ из 50 источников (см. таблица).**

**(замечание: переформатируйте данные в один столбец)**

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **10,4** | **18,6** | **10,3** | **26** | **45** | **18,2** | **17,3** | **19,2** | **25,8** | **18,7** |
| **28,2** | **25,2** | **18,4** | **17,5** | **41,8** | **14,6** | **10** | **37,8** | **10,5** | **16** |
| **18,1** | **16,8** | **38,5** | **37,7** | **17,9** | **29** | **10,1** | **28** | **12** | **14** |
| **14,2** | **20,8** | **13,5** | **42,4** | **15,5** | **17,9** | **19** | **10,8** | **12,1** | **12,4** |
| **12,9** | **12,6** | **16,8** | **19,7** | **18,3** | **36,8** | **15** | **37** | **13** | **19,5** |

1. **Проверьте, подчиняются ли нормальному закону распределения данные поля «ИМТ», заполненных в работе №1.**
2. **Для данных, заполненных в работе №1 проверьте, подчиняются ли нормальному закону распределения поле «АД диаст».**
3. **Проверьте, подчиняются ли нормальному закону распределения данные по весу женщин (Для обработки\БД\_ЛПУ).**
4. **Проверьте, подчиняются ли нормальному закону распределения данные по росту женщин (Для обработки\БД\_ЛПУ).**
5. **Для данных, заполненных в работе №1 проверьте, подчиняются ли нормальному закону распределения поле «АД сист».**
6. **Проверьте, подчиняются ли нормальному закону распределения данные по температуре воздуха за 2012 год по г. Сургут (Для обработки\с rp5.ru).**
7. **Проверьте, подчиняются ли нормальному закону распределения данные индексу Кердо для мальчиков (Для обработки\Индекс Кердо).**
8. **Проверьте, подчиняются ли нормальному закону распределения данные индексу Кердо для девочек (Для обработки\Индекс Кердо).**