

Метрология, стандартизация и сертификация

Практика 4

Статистические критерии

Сортировки, ранги и персентили (процентили)

https://ru.wikipedia.org/wiki/Квантили_распределения_Стьюдента <https://ru.wikipedia.org/wiki/Квантиль>
<http://excel2.ru/articles/rang-i-persentil-v-nadstroyke-paket-analiza-ms-excel>
<http://excel2.ru/articles/procentili-v-ms-exce> 1 <http://excel2.ru/articles/kvantili-raspredeleniy-ms-excel>

Для упорядочивания значений в Excel имеются следующие возможности:

- 1) Сортировка через меню «Данные» и назначение номеров (рангов) в полученной последовательности как членов арифметической прогрессии от 1 с шагом 1
- 2) С помощью функции «**РАНГ**» (без сортировки исходной последовательности)
РАНГ(число;ссылка;порядок) (ранг - номер позиции числа в отсортированном списке)

Число - это число, для которого определяется ранг.

Ссылка - это массив или ссылка на список чисел. Нечисловые значения в ссылке игнорируются.

Порядок - это число, определяющее способ упорядочения.

Если порядок равен 0 (нулю) или опущен, то Ранг=1 у наибольшего числа. Если порядок - это любое ненулевое число, то Ранг=1 у наименьшего числа.

- 3) С помощью надстройки «Пакет анализа» инструмент «Ранг и персентиль»
Excel 2007 и позже: Вкладка Данные -> Анализ -> Анализ данных
Excel 2003 и ранее: Меню Сервис -> Анализ данных

Квантиль в математической статистике — значение, которое заданная случайная величина не превышает с фиксированной вероятностью

Пусть P - вероятность для случайной величины X . Пусть $\alpha \in (0,1)$.

α -квантилью распределения P называют такое число x_α , что

$$P(X \leq x_\alpha) \geq \alpha$$

$$P(X \geq x_\alpha) \geq 1 - \alpha$$

0,25-квантиль - первый (или нижний) квартиль (от лат. quarta — четверть);

0,5-квантиль - медиана (от лат. mediana — середина) или второй квартиль;

0,75-квантиль - третий (или верхний) квартиль

Интерквартильный размах - называется разность между третьим и первым квартилями. Интерквартильный размах - характеристика разброса распределения величины. Медиана и интерквартильный размах могут быть использованы вместо математического ожидания и дисперсии в случае распределений с большими выбросами, либо при невозможности вычисления последних.

Значения дециля делят рассматриваемую совокупность на десять равных частей.

Первый дециль отделяет 10 % наименьших величин, лежащих ниже дециля, от 90 % наибольших величин, лежащих выше дециля.

Персентили также называются процентиями или центиями и делят совокупность на 100 частей.

<https://ru.wikipedia.org/wiki/Квантиль>

В Excel есть функция **Персентиль (Процентиль)**

Оформление и сдача работы

1. Создать файл в Excel: **Фамилия_МСС_Пр04.xls** (фамилия и все прочее в названии - русскими буквами)

2. Тип файла - «Книга Excel 97-2003»

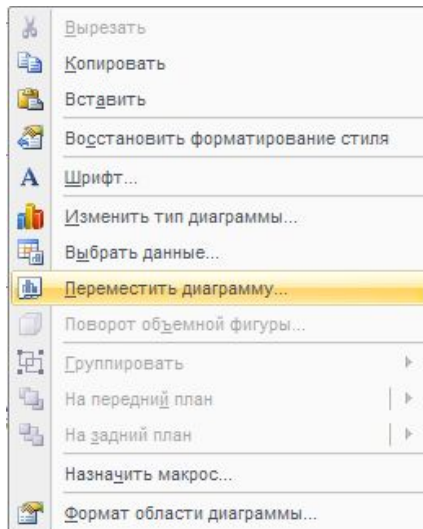
3. В задачах:

В начале - записывать «**Задача N**», ячейку выделять желтым цветом

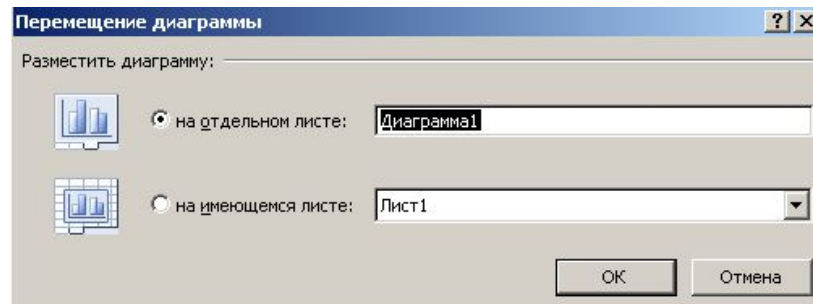
В конце - записывать полный ответ задачи, ячейки с ответом выделять жирной рамкой

4. Рисунки всегда находятся на отдельном листе

щелкнуть право клавишей мыши на диаграмме и задать команду «Переместить диаграмму» как показано на рисунках ниже.



Если в задаче указано несколько параметров, например стандартное отклонение 1, 3, 0.3, то все три графика на одной диаграмме (если прямо не указано иное)



5. За ненадлежащее оформление будут снижаться баллы

6. При наличии в папке файла вида **МСС_Пр04_Вопросы и ответы (...).doc** быть готовым отвечать при сдаче работы на указанные вопросы

Задача 1

Для исходных данных по добыче нефти (см. **МСС_Пр04_Распределение (...).xls** и лист Table 3.5) вычислить ранг и персентиль (процентиль)

Подсказка

- 1) Данные по добыче выделить на листе Table 3.5, скопировать и сделать специальную вставку в свой файл с транспонированием (так чтобы строка данных стала столбцом). Сделать заголовок столбца «Добыча, tb/d»
- 2) Вычислить ранг и персентиль (процентиль) с использованием надстройки «Пакет анализа»
- 3) Выделить строки в полученной таблице (красным полужирным шрифтом), содержащие: первый, второй и третий квартили, вычислить интерквартильный размах

Задача 2 (просто использование формул)

(использование стандартного нормального распределения, где среднее равно нулю и стандартное отклонение σ равно единице) <https://statanaliz.info/excel/formuly/111-normalnoe-raspredelenie-v-excel>

Найти чему равна вероятность попадания случайной величины в пределы $\pm 1\sigma$, $\pm 2\sigma$ и $\pm 3\sigma$ (по квантили найти вероятность).

Подсказка:

- 1) искомая вероятность равна разности: $\Phi(z) - \Phi(-z)$, где $\Phi(z)$ - функция **НОРМСТРАСП** или **НОРМРАСП** (с соответствующими параметрами и типом интегральная), либо, что даст тот же результат, $2\Phi(z) - 1$
- 2) Формат результата (вероятность) - выразить в процентах, 6 знаков после запятой

Задача 3 (просто использование формул)

(обратная к задаче 2: по заданной вероятности найти в интервал какой ширины (полуширины) попадет случайная величина) <https://statanaliz.info/excel/formuly/111-normalnoe-raspredelenie-v-excel>

Найти чему равна полуширина интервала, в который с 50%, 90%, 95% и 99% вероятностью попадет случайное число (по вероятности найти квантиль)

Подсказка:

- 1) искомая полуширина (или верхняя граница) определяется с помощью функции **НОРМСТОБР** $((1+p)/2)$, где p - вероятность в долях единицы
- 2) Формат результата (вероятность) - 6 знаков после запятой

Задача 4 (использование формул применительно к ряду данных)

Решить задачи 2 и 3 с теми же параметрами применительно к имеющейся выборке (см. задача 1) - уровни добычи нефти.

Подсказка:

- 1) Рассчитать описательную статистику (как в МСС_Пр01), разместив результат на отдельном листе **ОпСтат**, взять из описательной статистики **«среднее»** и **«стандартное отклонение»**
- 2) Для аналога задачи 2 использовать функцию **НОРМРАСП** (с соответствующими параметрами - **«среднее»** и **«стандартное отклонение»**)
Для аналога задачи 3 использовать функцию **НОРМОБР** (с соответствующими параметрами - **«среднее»** и **«стандартное отклонение»**)
- 3) Вычисления вести на новом листе **«Добыча»**