Метрология, стандартизация и сертификация

Практика 2

Нормальное распределение, обработка экспериментальных данных

Теоретическая часть

1. Статистические величины

Математическое ожидание *M(x)* — среднее вероятностное значение случайной величины
Математическое ожидание — теоретическая величина, к которой приближается среднее значение случайной величины при большом числе испытаний.

Математическое ожидание квадрата отклонения случайной величины от ее математического ожидания *M (x)* называется дисперсией величины x и обозначается σ^2

$$\sigma^2 = M[x - M(x)]^2 = M(x^2) - M^2(x)$$

Дисперсия суммы независимых случайных величин равна сумме их дисперсий.

Если появление некоторого события в каждом испытании имеет вероятность p, то математическое ожидание частоты m этого события при n испытаниях равно:

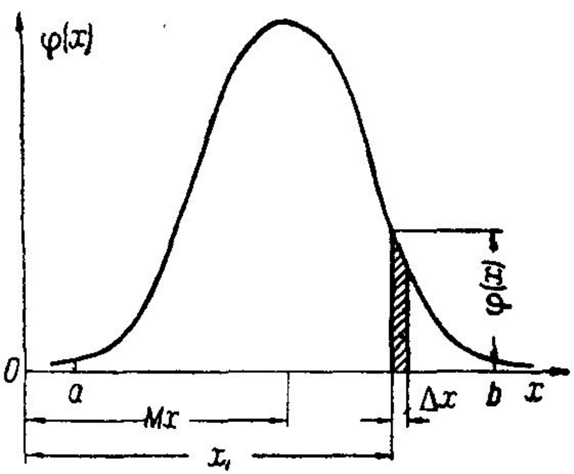
$$M(m) = np$$

Дисперсия частоты

$$\sigma^2 = np(1-p)$$

Положительное значение квадратного корня из дисперсии называется средним квадратическим отклонением

Формулы, приведенные выше формулы для средних значений случайной величины, ее математического ожидания и дисперсии относились к случаю, когда случайная величина дискретна и число возможных ее значений конечно.



Для определения понятий математического ожидания и дисперсии непрерывной случайной величины нужно ввести новое понятие — плотности распределения.
Обозначим через *X* некоторую непрерывную случайную величину, которая может принимать любые числовые значения из промежутки *(a, b)*.
Пусть *x* есть некоторое число из этого промежутка.
Определим вероятность *dP* того, что величина *X* принимает значения, заключенные между *x* и *x* + *dx*.

Эта вероятность пропорциональна *dx* (при бесконечно малом *dx)* и зависит от *x*.
Поэтому положим: dP = ϕ(x)dx.

Функция ϕ(x) называется плотностью распределения вероятностей случайной величины *X,* произведение ϕ(x)dx — элементом вероятности.
Кривая $y = \phi(x)$ называется кривой распределения вероятностей случайной величины.

Если известна плотность распределения ϕ(x) случайной величины, то вероятность того, что значения, принимаемые этой величиной, будут заключены в промежутке между x_1 и x_2, равна следующему интегралу:

$$P(x_1 < X < x_2) = \int_{x_1}^{x_2} \varphi(x)dx \qquad \int_{-\infty}^{\infty} \varphi(x)dx = 1$$

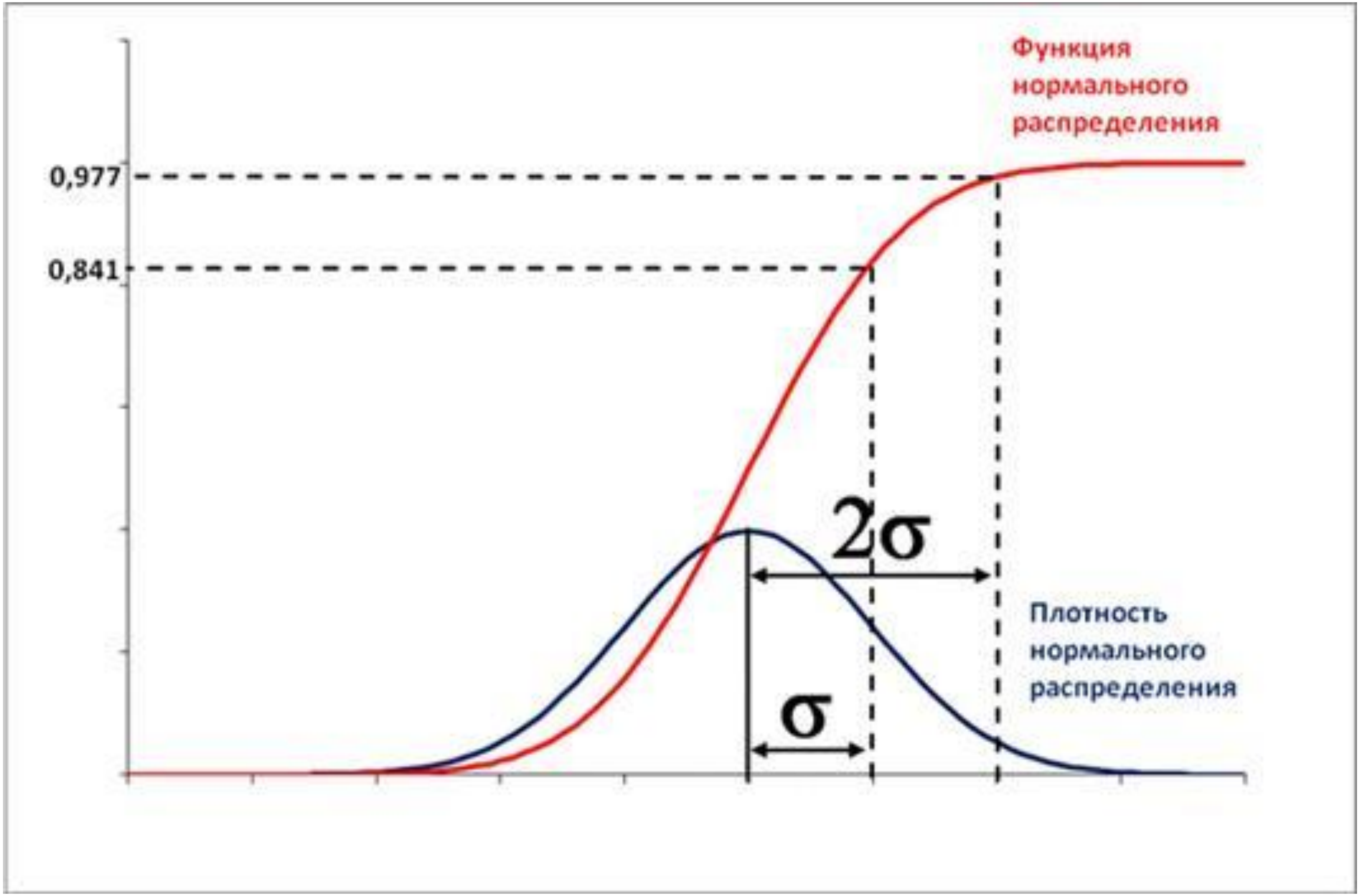
Математическое ожидание *M(x)* непрерывной случайной величины, распределенной равномерно от a до b равно: M(x)=(a+b)/2

Кривая нормального распределения случайной величины

$$y = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-a)^2}{2\sigma^2}}$$

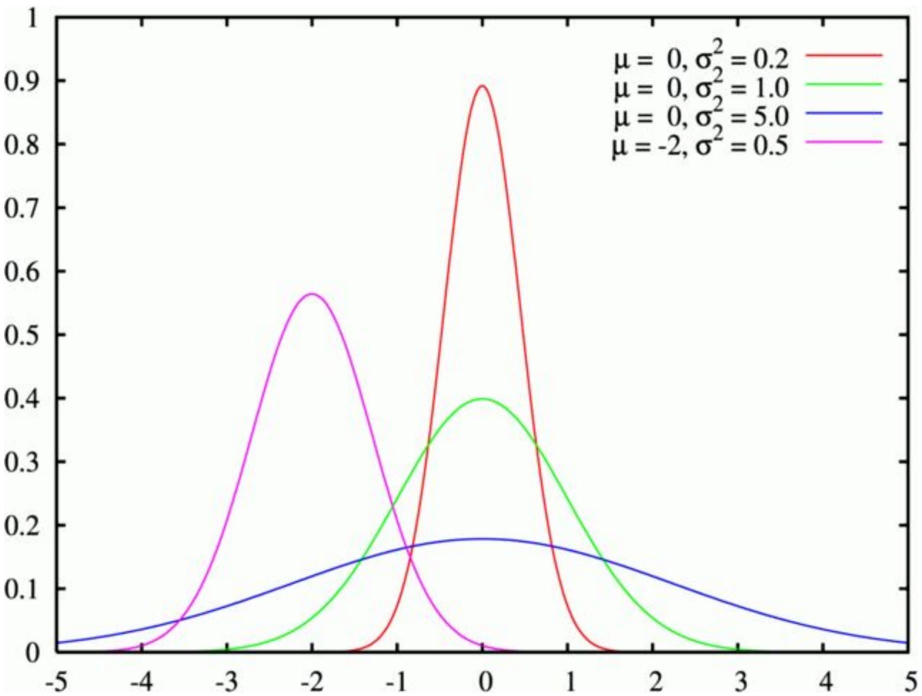
где a - математическое ожидание, σ^2 - дисперсия, σ - среднее квадратическое отклонение

Нормальное распределение:
плотность вероятности и функция распределения

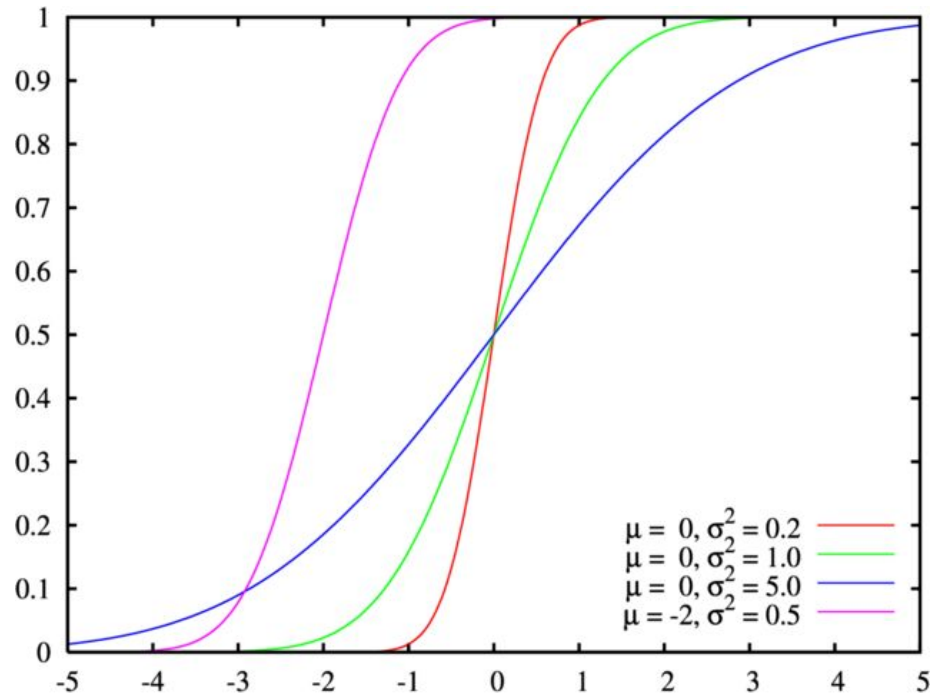


Нормальное распределение

Плотность вероятности



Функция распределения



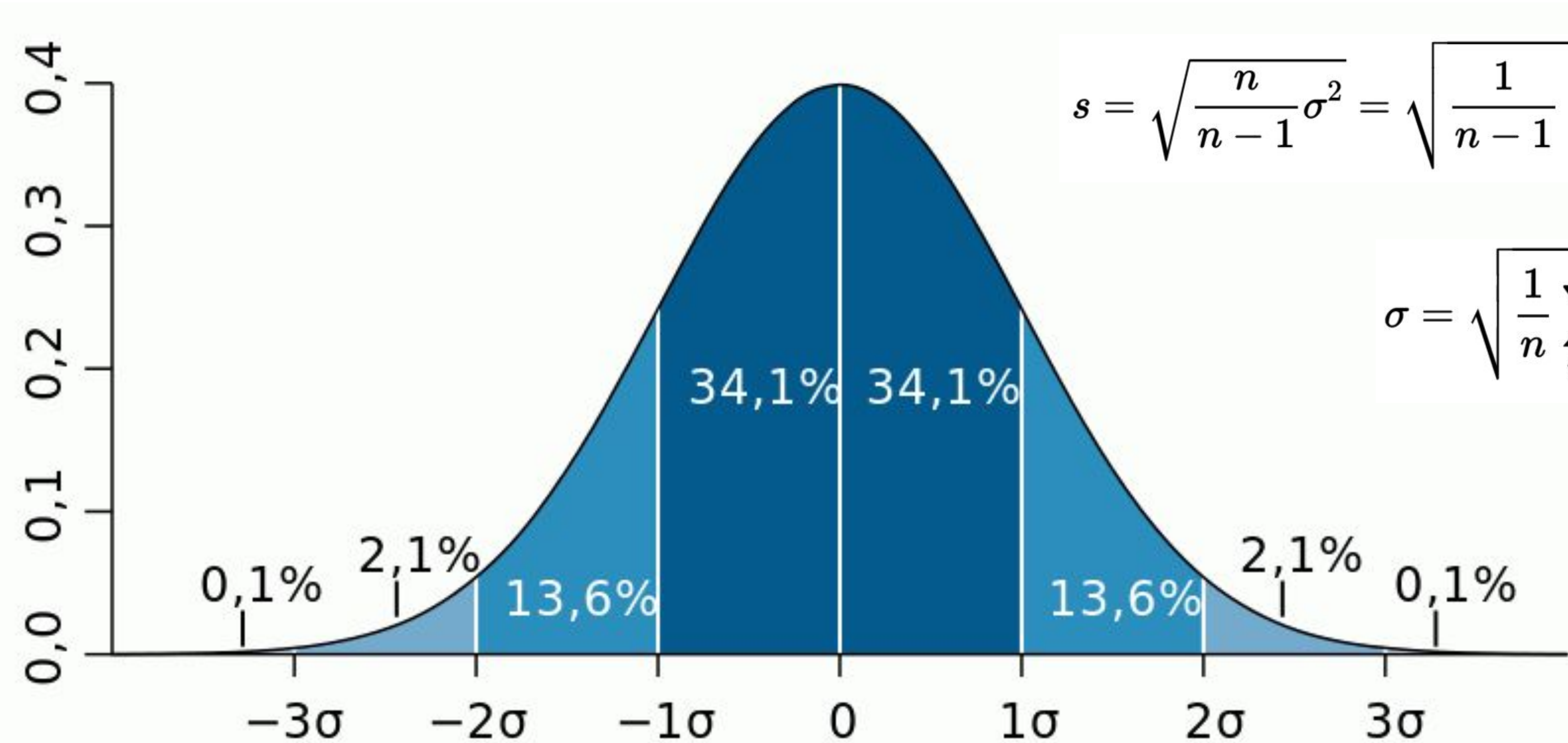
μ — математическое ожидание (среднее значение), медиана и мода распределения,
σ — среднеквадратическое отклонение (σ^2 — дисперсия) распределения.

Зеленая линия соответствует стандартному нормальному распределению

Плотность вероятности нормального распределения

Правило трёх сигм

https://ru.wikipedia.org/wiki/Среднеквадратическое_отклонение



$$s = \sqrt{\frac{n}{n-1}\sigma^2} = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2};$$

$$\sigma = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2}.$$

Практически все значения нормально распределённой случайной величины лежат в интервале $(\bar{x} - 3\sigma;\ \bar{x} + 3\sigma)$

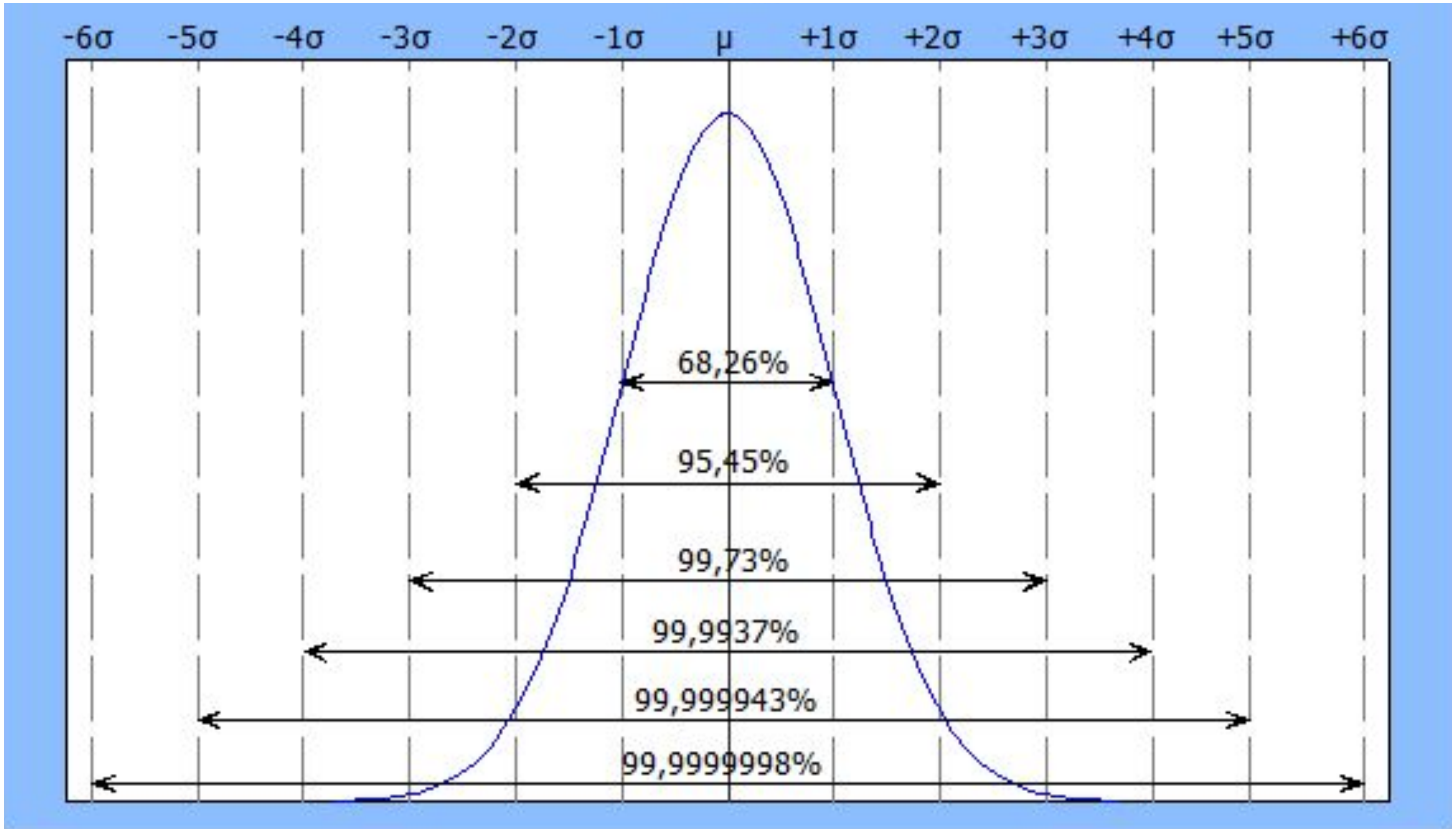
Более строго — приблизительно с вероятностью 0,9973 значение нормально распределённой случайной величины лежит в указанном интервале (при условии, что величина $\bar{x}$ истинная, а не полученная в результате обработки выборки).
Если же истинная величина $\bar{x}$ неизвестна, то следует пользоваться не s a s.

Интервал $(\bar{x} - 2\sigma;\ \bar{x} + 2\sigma)$ дает вероятность 0.954

Правило шести сигм

https://ru.wikipedia.org/wiki/Шесть_сигм
https://en.wikipedia.org/wiki/Six_Sigma

Название происходит от статистического понятия среднеквадратичного отклонения, обозначаемого греческой буквой σ. Зрелость производственного процесса в этой концепции описывается как σ -рейтинг отклонений, или процентом бездефектной продукции на выходе, так, процесс управления качеством 6σ на выходе даёт 99,99966 % выходов без дефектов, или не более 3.4 дефектных выходов на 1 млн операций.



Задачи

Создать файл в Excel: *Фамилия*_МСС_Пр02

Задача 1. Рассчитать кривую нормального распределения случайной величины

$$y = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-a)^2}{2\sigma^2}}$$

где a - математическое ожидание, σ^2 - дисперсия, σ - среднее квадратическое отклонение

- величину **a** взять в файле ***МСС_Пр02_Распределение (…).xls***
- на интервале -5+**a** ≤ x+**a** ≤ 5 +**a** с шагом 0.05 и σ = 1.0, 0.3, 3.0
- построить диаграммы (оформление как в **Практика 1**)

Подсказка

1) Лист с расчетами назвать **Норм**
2) Для удобства сделать отдельный столбец с вычислением показателя экспоненты
3) **При вычислениях не забывать приоритет операций и ставить скобки !**
4) Диаграммы точечные (сглаженные линии без маркеров), легенда внизу, подписи осей с одним знаком после запятой
5) Диаграммы размещать на отдельном листе, линии подписывать как s=1.0, s=0.3, s=3.0
6) Лист с диаграммами назвать **D-Норм**

Задача 2

Рассчитать нормальные распределения с использованием функции Excel **НОРМРАСП** при для -5+a ≤ x +a ≤ 5 +a, шаг 0.05 и σ = 1

Вычислить столбец разностей самостоятельно вычисленной функции (Задача 1) и с использованием функции **НОРМРАСП.**

Найти сумму по столбцу разностей.

Нулевая или околонулевая сумма (меньше 1E-10) - признак правильных вычислений

Подсказка

1) величину **a** взять в файле ***МСС_Пр02_Распределение.xls***

2) Расчеты выполнять на том же листе **Норм** в соседних столбцах

Задача 3

Рассчитать и построить графики при - 4≤ x ≤ 4, шаг 0.05

- стандартного нормального интегрального распределения с использованием функции Excel **НОРМСТРАСП**

Это распределение имеет среднее равное нулю и стандартное отклонение равное единице. Эта функция используется вместо таблицы для стандартной нормальной кривой.

- стандартного нормального интегрального распределения с использованием функции **НОРМРАСП**
- плотности стандартного нормального распределения: с использованием функции **НОРМРАСП**

$$f(x; 0; 1) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$$

Подсказка

1) Новый лист с расчетами назвать **Норм2**
2) Диаграмма точечная (сглаженные линии без маркеров), подписи осей с одним знаком после запятой
3) Диаграмму размещать на отдельном листе
4) Лист с диаграммой назвать **D-Норм2**

Теоретическая часть

2. Обработка экспериментальных данных

Обработку серии измерений следует проводить в следующем порядке:

1) определить среднее арифметическое;

2) найти среднюю квадратическую ошибку отдельного измерения (т.к. работаем с выборкой, а не с генеральной совокупностью)

$$\sigma = \sqrt{\frac{\sum_{i=1}^{n}(\bar{x} - x_i)^2}{n-1}}$$

3) определить наибольшую возможную ошибку А отдельного измерения и убедиться, что среди результатов измерений нет таких, которые отличались бы от среднего арифметического более чем на А. Если бы таковые оказались, их следует отбросить и начать обработку сначала;

4) определить среднюю квадратическую ошибку σ_0 среднего арифметического.

$$\sigma_0 = \sqrt{\frac{\sum_{i=1}^{n}(\bar{x} - x_i)^2}{n(n-1)}}$$

Задача 4

Обработать шестнадцать измерений, представляющих собой результаты анализа раствора на содержание в нем $MgCl_2$. Каждый вычисляет свой ряд на основании значений а (величину а брать из 1 задачи)

Подсказка

1) Новый лист с расчетами назвать **Среднее**

2) Исходную информацию взять из файла ***МСС_Пр02_Распределение (...).xls***

3) Вычислить свой набор данных. Величина **а** такая же как и в задачах 1,2, см. файл ***МСС_Пр02_Распределение.xls***

3) Вычислить среднее арифметическое. $\bar{x}$

- для определения n (числа значений) использовать функцию Excel **СЧЕТ**
- не забывать закреплять нужные ячейки при вычислениях

4) формат вывода результата должен соответствовать исходным данным: среднее арифметическое - 1 знак после запятой, σ и σ_0 - два знака после запятой

5) Вычислить столбец ошибок отдельных измерений как $\bar{x} - x_i$

6) Вычислить среднюю квадратическую ошибку

отдельного измерения по формуле:

$$\sigma = \sqrt{\frac{\sum_{i=1}^{n}(\bar{x} - x_i)^2}{n-1}}$$

6) Сравнить 3σ и $\left|\bar{x} - x_i\right|$.

Если $\left|\bar{x} - x_i\right| > 3\sigma$, то x_i наблюдение отбрасывается как ошибочное и расчеты по пунктам выполнить заново
При выкидывании значений создать новый столбец с данными и для них вести снова вести вычисления по пунктам 5-6

6) Вычислить среднюю квадратическую ошибку среднего арифметического по формуле

$$\sigma_0 = \sqrt{\frac{\sum_{i=1}^{n}(\bar{x} - x_i)^2}{n(n-1)}}$$

7) Записать ответ в виде

$$\bar{x} = \bar{x}(\textit{значение}) \pm \sigma_0(\textit{значение})$$

Задача 5

Испытаниями установлено, что относительная ошибка прибора равна 12%. Сколько дублирующих приборов надо поставить, чтобы обеспечить относительную точность результатов в 10, 5, 3 и 1%?

Подсказка

1) Новый лист с расчетами назвать **Среднее2.** В ячейке А1 написать **Задача 5**

2) Использовать соотношения

$$\sigma = \sqrt{\frac{\sum_{i=1}^{n}(\bar{x}-x_i)^2}{n-1}} \qquad \sigma_0 = \sqrt{\frac{\sum_{i=1}^{n}(\bar{x}-x_i)^2}{n(n-1)}} \qquad \sigma_0 = \frac{\sigma}{\sqrt{n}}$$

3) σ=12%, σ_0 = 10%, 5%, 3%, 1%. Найти n, округляя полученное значение до целого в большую сторону.

Задача 6

Точность прибора составляет 6%. Сколько раз надо повторить измерение, чтобы точность среднего арифметического полученных измерений была равна 2%?

Подсказка

1) Вычисления вести на листе **Среднее2.** Сделать надпись **Задача 6**

2) Использовать формулы из задачи 6

Дополнение

Соответствие встроенных функций Excel и OpenOffice Calc

http://www.oivt.ru/blog/sootvetstvie-vstroennyh-funkciy-excel-i-openoffice-calc

Сравнение функциональности LibreOffice и MS Office

https://wiki.documentfoundation.org/Feature_Comparison:_LibreOffice_-_Microsoft_Office/ru

Соответствие команд MS Excel и Calc

http://wiki.harlamenkov.ru/wiki/RU/kb/00000427
http://inf-w.ru/?page_id=67

Функция ФОШ

https://support.office.com/ru-ru/article/Функция-ФОШ-c53c7e7b-5482-4b6c-883e-56df3c9af349

Справочное руководство LibreOffice

https://help.libreoffice.org/Main_Page/ru
https://help.libreoffice.org/Calc/Welcome_to_the_Calc_Help/ru

Статистические функции LibreOffice

https://help.libreoffice.org/Calc/Statistics_Functions/ru

Соответствие английских и русских названий функций в Excel

http://brusentsov.com/2009/12/27/3519
http://sirexcel.ru/sootvetstvie-funkcij-na-anglijskom-i-russkom-yazykake-v-excel/