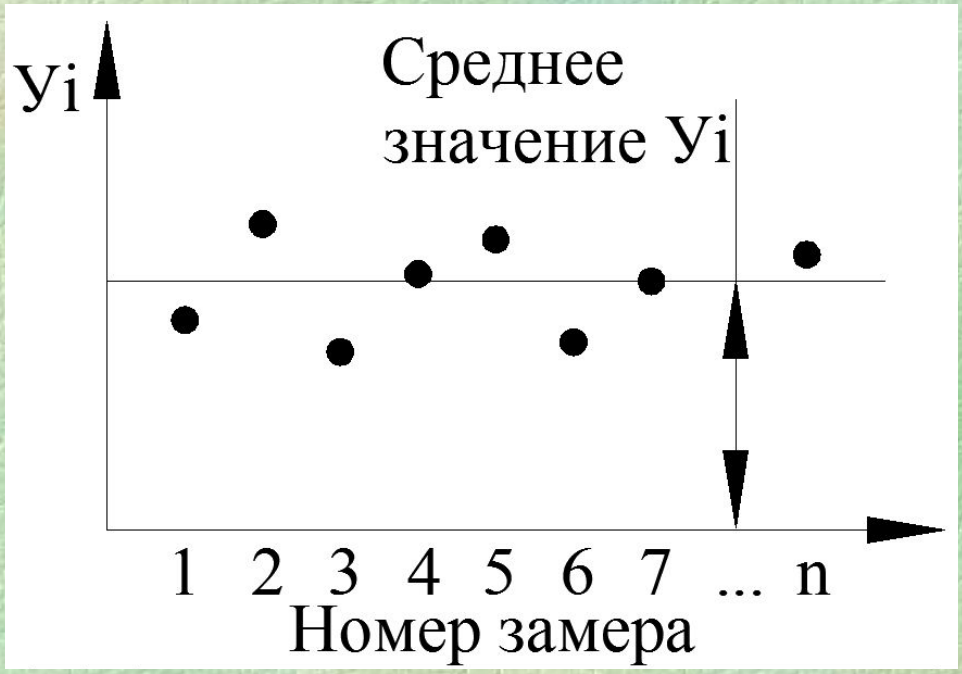
Статистическая обработка результатов измерений.
Нормальный закон распределения.

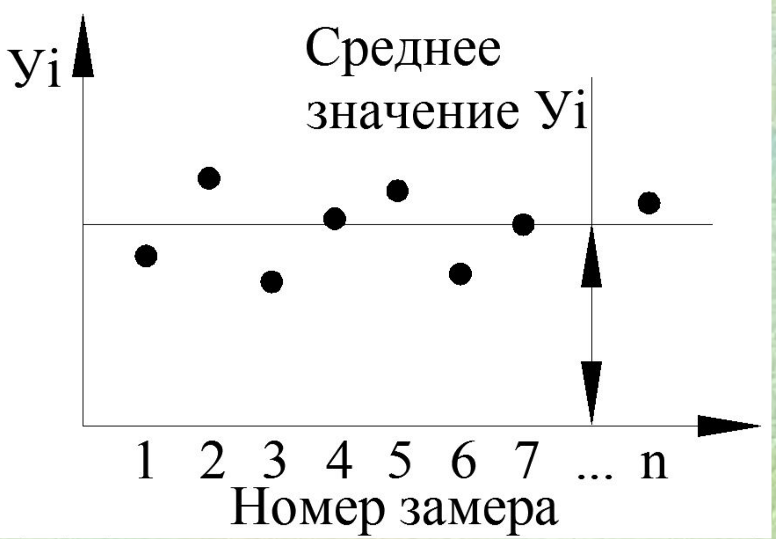
Нормальный закон распределения

Для получения закона распределения любой случайной величины У, ее необходимо неоднократно измерить. Пусть в эксперименте проведено n-ое количество замеров выходного параметра Уi, который зависит от одного, либо от нескольких входных параметров-аргументов Хi.

Каждое значение Уi, в силу разных причин, может отличаться от других его значений.



Важнейшими характеристиками закона распределения являются **математическое ожидание Му и дисперсия σ^2.**



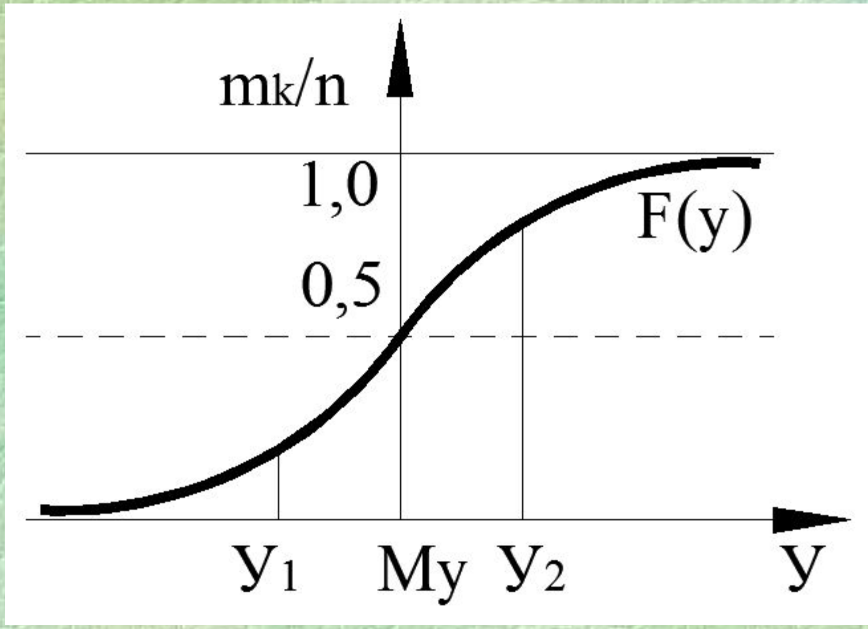
Математическим ожиданием Му называется наиболее вероятное значение величины У при **n** → ∞ :

$$My = \frac{1}{n}\sum_{i=1}^{n} y_i = \bar{y}$$

Дисперсией σ^2 называют характеристику, которая определяет кучность (разброс) значений Уi относительно Му. При **n** → ∞ σ^2 можно рассчитать по формуле:

$$\sigma^2 = \frac{1}{n}\sum_{i=1}^{n} (y_i - My)^2$$

По значениям Уi можно построить график **функции распределения F(у)**. Для этого по горизонтальной оси отложим значения Уi, а по вертикали – **относительное количество опытов m_k/n**, в которых замеренное значение Уi оказалось меньше заданного значения Уk.



Функции распределения F(у) является интегральной функцией.

При Уk = -∞ m_k/n = 0;

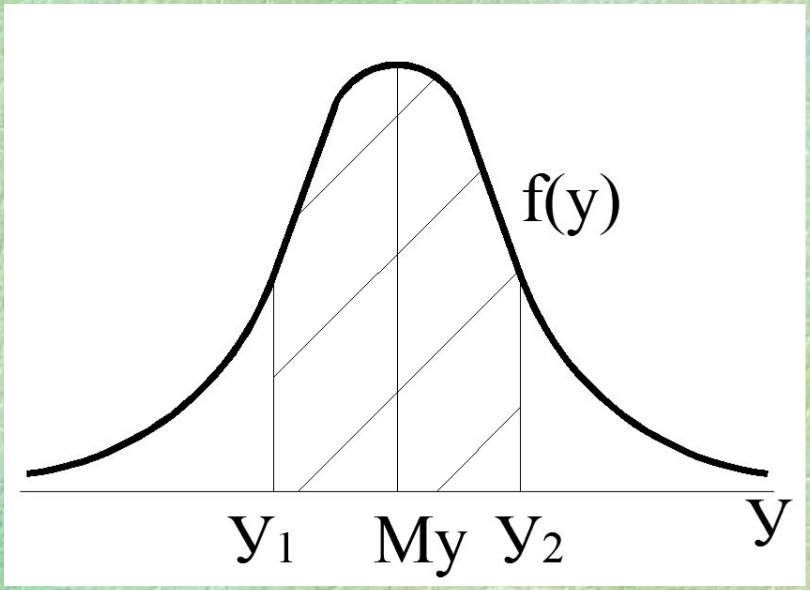
при Уk = ∞ m_k/n = 1.

Вероятность того, что измеряемое значение Уi окажется в интервале от $У_1$ до $У_2$, можно определить по формуле:

$$p(У_1 \leq Уi \leq У_2) = F(У_2) - F(У_1)$$

Более наглядно закон распределения можно представить с помощью **плотности распределения f(У)**, которая является дифференциальной функцией и связана с F(У) зависимостью:

$$F(y) = \int_{-\infty}^{y_i} f(y)\,dy$$

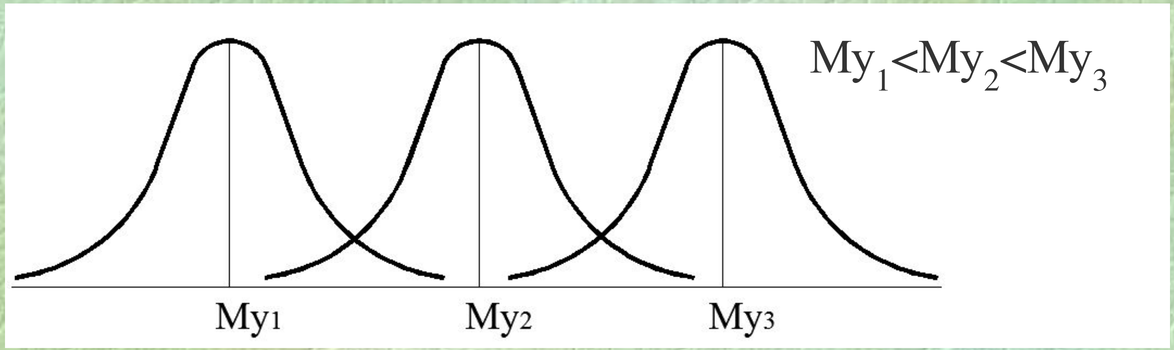


Если **выходнои параметр Уi** (функция) можно рассматривать как **сумму достаточно большого числа случайных величин Xi** (аргументов), то данная величина также является **случайной** и обычно **подчиняется нормальному закону распределения**.

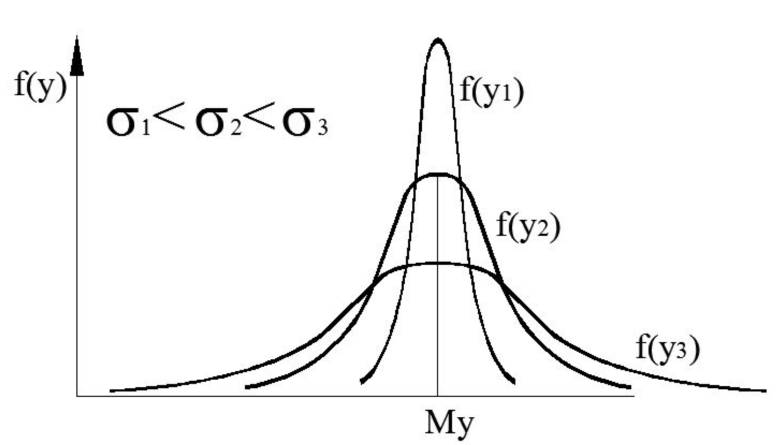
Кривую f(У) для нормального закона распределения можно построить с помощью уравнения Гаусса:

$$f(y) = \frac{1}{\sigma\sqrt{2\pi}} exp\left(-\frac{(y-My)^2}{2\sigma^2}\right)$$

При изменении параметра Му форма нормальной кривой не изменяется. В этом случае, если математическое ожидание **Му уменьшилось или увеличилось, график нормальной кривой сдвигается влево или вправо .**



При изменении параметра **σ изменяется форма нормальной кривой**. Если **σ увеличивается, то максимальное значение функции *f(x)* убывает, и наоборот**, так как **площадь, ограниченная кривой распределения и осью *Ох*, должна быть постоянной и равной 1.**



σ – среднеквадратичное отклонение

Для реального эксперимента, т.е для случая, когда количество замеров Уi значительно меньше ∞, **имеют дело с выборкой значений Уi**. В этом случае вместо математического ожидания Му используется **среднее арифметическое** от $\sum_{i=1}^{n} y_i$:

$$\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$$

а **дисперсия** обозначается символом S^2 и определяется по формуле:

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(y_i - \bar{y})^2$$

Знаменатель **(n-1)** называется **числом степеней свободы** и обозначается символом **f**.

Числитель $\sum_{i=1}^{n}(y_i - \bar{y})^2$ называется **суммой квадратов отклонений** и обозначается **SS**. Тогда:

$$S^2 = SS/f$$

Рассмотрим пример. Пусть задана выборка значений роста группы студентов:

№	Рост, см
1	159
2	163
3	165
4	168
5	168
6	169
7	170
8	172
9	173
10	174
11	175
12	175
13	177
14	177
15	178
16	179
17	180
18	181
19	183
20	192

Требуется построить для этой выборки функцию распределения и плотность распределения, а также рассчитать ее дисперсию.

Для наглядности отобразим выборку графически:

Рост студентов является исследуемой функцией. Разобьем диапазон значений представленной выборки на 10 одинаковых интервалов. Для этого:

1. Определим размах диапазона:

$$R = y_{max} - y_{min} = 192 - 159 = 33.$$

2. Рассчитаем шаг интервала: $h = R/10 = 3{,}3$.
3. Заполним таблицу:

№	Угрi	Sn_i	n_i	F(y)	f(y)
1	162,3	5	5	0,25	0,25
2	165,6	6	1	0,3	0,05
3	168,9	7	1	0,35	0,05
4	172,2	13	6	0,65	0,3
5	175,5	14	1	0,7	0,05
6	178,8	17	3	0,85	0,15
7	182,1	18	1	0,9	0,05
8	185,4	19	1	0,95	0,05
9	188,7	19	0	0,95	0
10	192	20	1	1	0,05

Где $\mathbf{y_{грi}}$ – верхняя граница i-того интервала;

Σni – количество студентов, чей рост меньше $y_{грi}$;

ni – количество студентов, чей рост соответствует i-той группе;

$F(y) = \Sigma n_i/n$ – функция распределения;

$f(y) = n_i/n$ – плотность распределения.

4. Построим графики:

5. Заполним таблицу:

№	Рост, см	(У-Усрi)^2
1	159	222,01
2	163	118,81
3	165	79,21
4	168	34,81
5	168	34,81
6	169	24,01
7	170	15,21
8	172	3,61
9	173	0,81
10	174	0,01
11	175	1,21
12	175	1,21
13	177	9,61
14	177	9,61
15	178	16,81
16	179	26,01
17	180	37,21
18	181	50,41
19	183	82,81
20	192	327,61
	173,9	1095,8

6. Из таблицы следует, что:

$$\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i = 173{,}9 \text{см} \; ; \; SS = \sum_{i=1}^{n}(y_i - \bar{y})^2 = 1095{,}8 \text{см}^2$$

7. Рассчитаем дисперсию S^2, учитывая, что
$f = (n-1) = 20 - 1 = 19$:

$$S^2 = SS/f = 1095{,}8/19 = 57{,}67 \text{см}^2$$