

Тема.

# НОРМАЛЬНЫЙ ЗАКОН РАСПРЕДЕЛЕНИЯ

Составила: М.П. Филиппова  
доцент кафедры высшей математики ИМИ СВФУ

# ПЛАН ЛЕКЦИИ

**&1. Основные характеристики  
нормального распределения**

**&2. Алгоритм построения кривой  
нормального распределения**

**&3. Проверка нормальности  
распределения результативного  
признака**

# &1. Основные характеристики нормального распределения

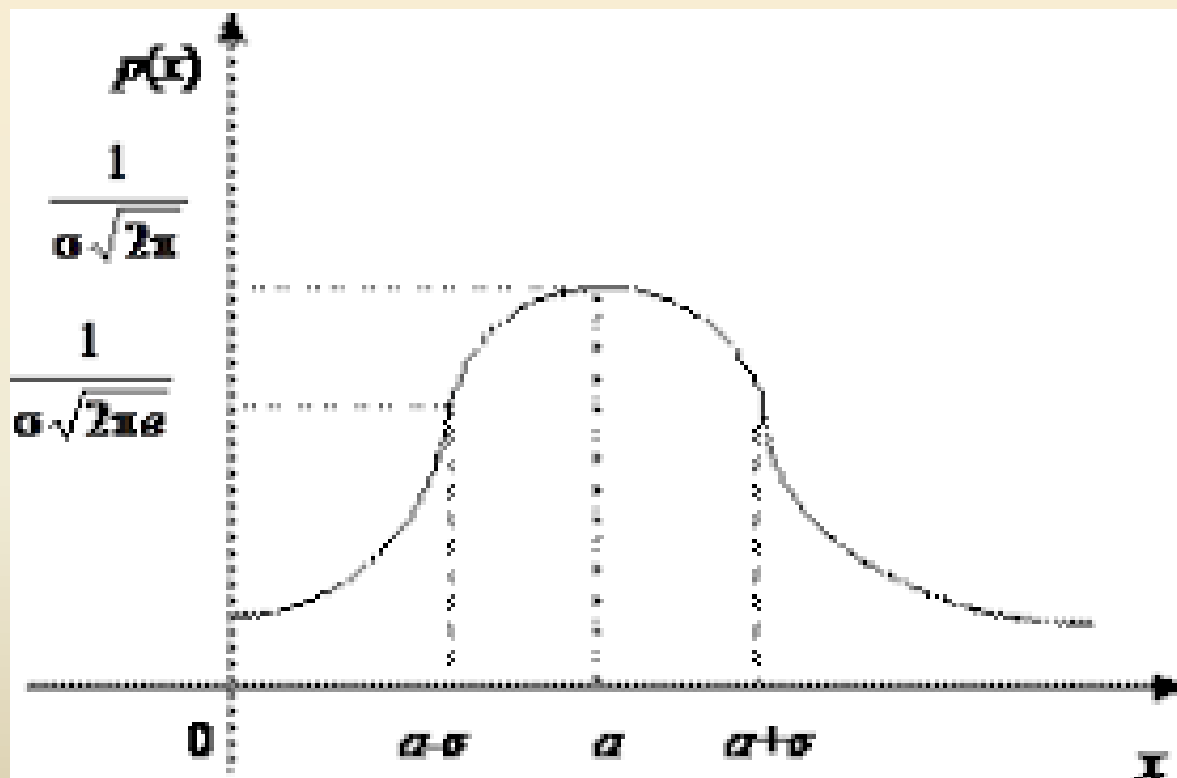
Нормальный закон распределения играет важнейшую роль в применении численных методов в психологии. Он лежит в основе измерений, разработки тестовых шкал, методов проверки гипотез.

*Нормальное распределение* выражается следующей формулой:

$$f_{\text{отн}} = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(x_i - \bar{x}_B)^2}{2\sigma^2}} \quad (1),$$

где  $f_{\text{отн.}}$  — относительные частоты появления каждого конкретного значения случайной величины  $x_i$ . Предполагается, что переменная  $x_i$ , может принимать бесконечно большие и бесконечно малые значения, количество измерений бесконечно, а интервал квантования мал. По этой формуле при различных значениях среднего арифметического ( $\bar{x}_B$ ) и среднего квадратического отклонения ( $\sigma$ ) получается семейство нормальных кривых.

Нормальное распределение имеет *колоколообразную форму* (см. Рис. 3), асимптотически приближается к оси  $X$  (то есть может принимать сколь угодно малые значения по ординате при стремлении икс-значений к плюс или минус бесконечности), значения моды, медианы и среднего арифметического равны между собой.



см. Рис. 1

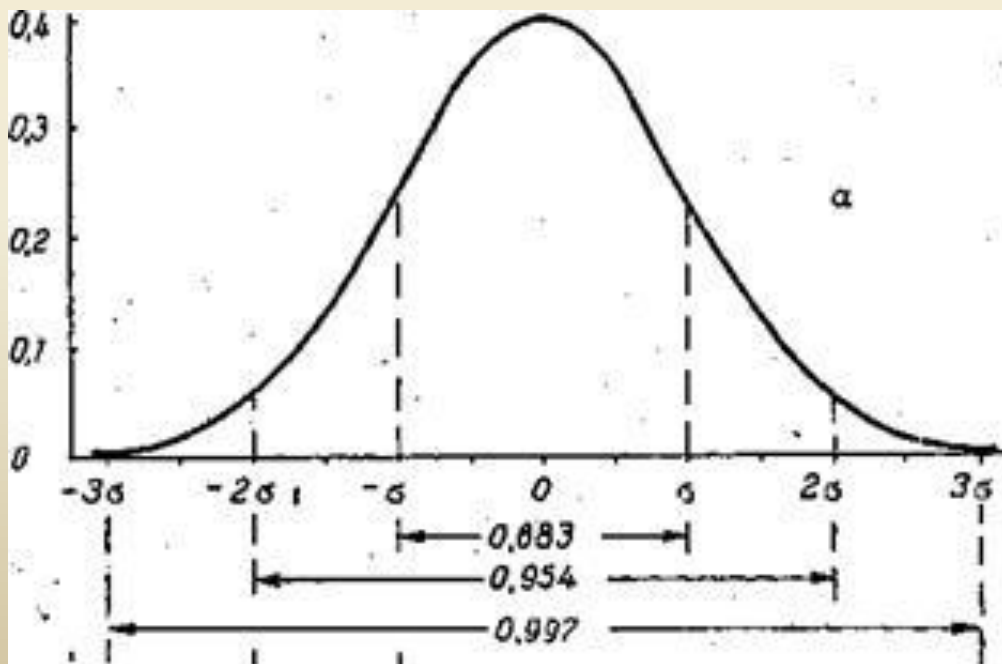
Параметры распределения – это его числовые характеристики, указывающие, где “в среднем” располагаются значения признака, насколько эти значения изменчивы и наблюдается ли преимущественное появление определенных значений признака. Наиболее практически важными параметрами являются математическое ожидание, дисперсия, среднее квадратическое отклонение, показатели асимметрии и эксцесса.

В реальных психологических исследованиях мы оперируем не параметрами, а их приближенными значениями, так называемыми оценками параметров. Это объясняется ограниченностью обследованных выборок. Чем больше выборка, тем ближе может быть оценка параметра к его истинному значению. В дальнейшем, говоря о параметрах, мы будем иметь в виду их оценки.

**Свойства НЗР.** О чем же свидетельствует среднее квадратическое отклонение? Оно позволяет сказать, что большая часть исследуемой выборки располагается в пределах  $\sigma$  от средней. Что это значит? Статистики показали, что при нормальном распределении «большая часть» результатов, располагающаяся в пределах одного среднего квадратического отклонения по обе стороны от средней, в процентном отношении всегда одна и та же и не зависит от величины среднего арифметического отклонения: она соответствует 68% популяции (т.е. 34% ее элементов располагается слева и 34% - справа от средней):

Точно так же рассчитали, что 94,45% элементов популяции при нормальном распределении не выходит за пределы двух стандартных отклонений от средней:

В пределах трех стандартных отклонений уместается почти вся популяция- 99,73%.



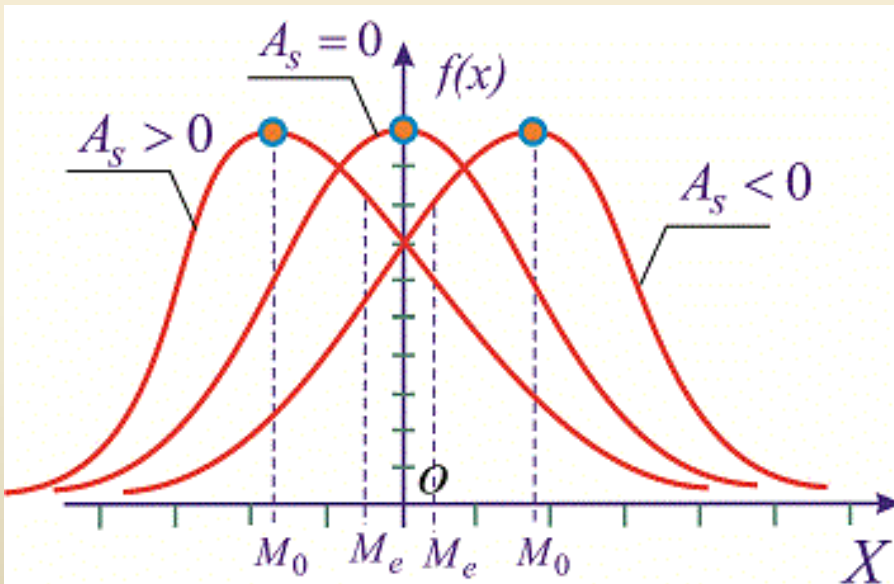
см. Рис. 2

# Характеристики асимметрии и эксцесса

Мера *асимметрии* – коэффициент асимметрии ( $A_s$ ), рассчитываемый по формуле

$$A_s = \frac{\sum_{i=1}^k (x_i - \bar{x}_B)^3}{n\sigma^3} \quad (1)$$

Асимметрия характеризует степень асимметричности распределения. Коэффициент асимметрии изменяется от минус до плюс бесконечности ( $-\infty < A_s < +\infty$ ), для симметричных распределений  $A_s = 0$ .

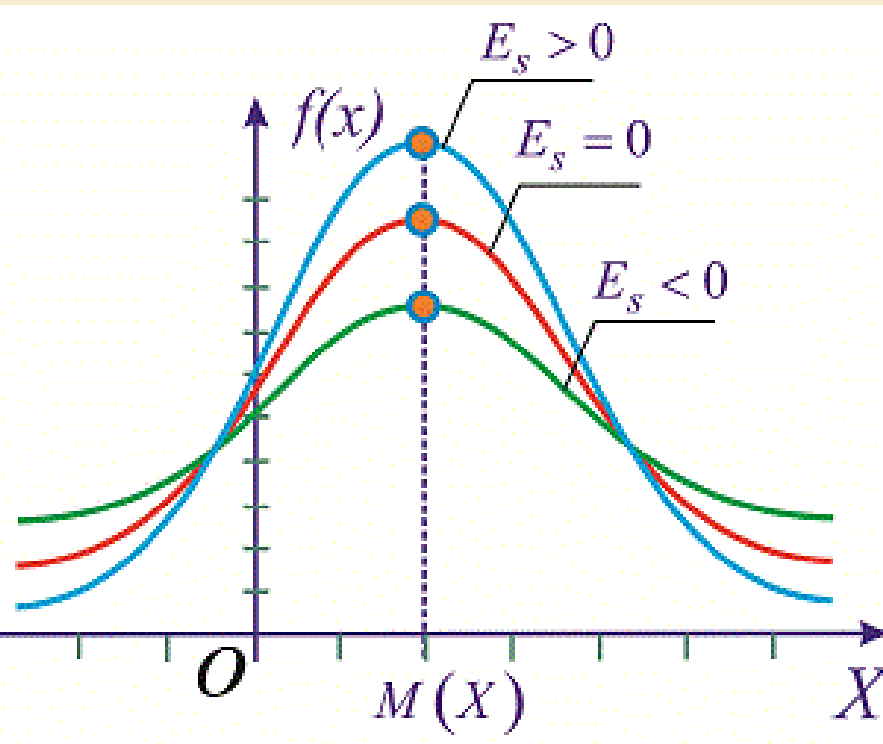


см. Рис. 3

В тех случаях, когда какие-нибудь причины благоприятствуют более частому появлению значений, которые выше или, наоборот, ниже среднего, образуются асимметричные распределения. При левосторонней, или положительной, асимметрии в распределении чаще встречаются более низкие значения признака, а при правосторонней, или отрицательной – более высокие (см. Рис. 3). Для симметричных распределений  $A=0$ ; где  $\bar{x}_B = M(X)$

Мера *экссесса* (островершинности) – коэффициент экссесса ( $E_x$ ), рассчитываемый по формуле:

$$E_x = \frac{\sum_{i=1}^k (x_i - \bar{x}_B)^4}{n\sigma^4} - 3 \quad (2)$$



(см. Рис. 4)

В тех случаях, когда какие-либо причины способствуют преимущественному появлению средних или близких к средним значений, образуется распределение с положительным эксцессом.

Если же в распределении преобладают крайние значения, причем одновременно и более низкие, и более высокие, то такое распределение характеризуется отрицательным эксцессом и в центре распределения может образоваться впадина, превращающая его в двувершинное (см. Рис. 4).  $\bar{x}_B = M(X)$



## &2. Алгоритм построения кривой нормального распределения

Есть несколько способов построения кривой нормального распределения по эмпирическим данным, если есть основания предположить близость данного распределения к нормальному. По одному из этих способов теоретические частоты ( $m'$ ) отыскиваются по формуле:

$$m' = \frac{Nh}{\sigma} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}},$$

где  $\frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} = \varphi(t)$  — табулированная величина, отыскиваемая

по отклонениям  $t$ , а  $Nh/\sigma$  — константа, на которую умножаются значения  $\varphi(t)$  и которая определяет теоретические частоты исходя из общей численности единиц совокупности и числа выделяемых групп.

## Алгоритм расчета теоретических частот

- 1) данные сортируются по убыванию;
- 2) подсчитывается частота повторения каждого значения, это и есть эмпирическая частота ( $m$ );
- 3) рассчитывается средняя арифметическая ряда ( $\bar{M}$ );
- 4) рассчитывается среднее квадратическое отклонение ( $\sigma$ );
- 5) находится нормированное отклонение каждого варианта от средней

$$x - \bar{M}$$

арифметической, т. е.  $t = \frac{\text{---}}{\sigma}$ ;

- 6) для найденных  $t$  по таблицам определяется  $\varphi(t)$  (см Приложение 3.2);
- 7) рассчитывается константа  $Nh/\sigma$ ,  $N$ - объем выборки,

$$h = \frac{x_{\max} - x_{\min}}{1 + 3.322 \cdot \lg N}$$

- 8) каждое значение  $\varphi(t)$  умножается на константу.

Результаты умножения (после округления до целых чисел) будут искомыми частотами ( $m'$ ) теоретической кривой нормального распределения.

**Пример.** На группе из 30 добровольцев-студентов, был проделан опыт по изучению глазодвигательной координации. Задача заключалась в том, чтобы поражать предъявляемые на дисплее движущиеся мишени. Были предъявлены 10 последовательностей из 25 мишеней. Построить кривую распределения величины, отражающей количество пораженных мишеней.

### **РЕШЕНИЕ**

Исп- емые	к-во пораж. мишеней
1	12
2	21
3	10
4	15
5	15
6	19
7	17
8	14
9	13
10	11
11	20
12	15
13	15
14	14
15	17

до воз.	m	$x_i - \bar{x}$	t	$\varphi(t)$	$8,5^*(t)$	m'
10	1	5,2	1,34	0,1626	1,3824	1,4
11	1	4,2	1,08	0,2227	1,8933	1,9
12	1	3,2	0,82	0,285	2,423	2,4
13	1	2,2	0,57	0,3391	2,8829	2,9
14	2	1,2	0,31	0,3802	3,2323	3,2
15	4	0,2	0,05	0,3984	3,3871	3,4
17	2	-1,8	-0,46	0,3589	3,0513	3,1
19	1	-3,8	-0,98	0,2468	2,0982	2,1
20	1	-4,8	-1,24	0,1849	1,572	1,6
21	1	-5,8	-1,49	0,1315	1,118	1,1

ср.зн= 15,2 15

Дисп= 15,067  $n=15$   $h=(x - \bar{x}) / n = 2,2$

ст.откл.= 3,88  $\lg 15 = 1,18$   $\text{Const} = Nh/\sigma = 8,5$

### кривая нормального распределения



(см. Рис. 5)

## **&3. Проверка нормальности распределения результативного признака**

Если мы применяем параметрические методы (к примеру, формулу для расчета коэффициента корреляции Браве-Пирсона или дисперсионный анализ) которые следует применять только тогда, когда известно или доказано, что распределение признака является нормальным (Суходольский Г.В., 1972; Шеффе Г., 1980 и др.), то в этом случае нам необходимо убедиться в нормальности распределения результативного признака. Нормальность распределения результативного признака можно проверить путем расчета показателей асимметрии и эксцесса и сопоставления их с критическими значениями (Пустыльник Е.И., 1968, Плохинский Н.А., 1970 и др.). Рассмотрим применение метода Е.И. Пустыльника на примере.

Действовать будем по следующему алгоритму:

- 1) рассчитаем критические значения показателей асимметрии и эксцесса по формулам Е.И.Пустыльника и сопоставим с ними эмпирические значения;
- 2) если эмпирические значения показателей окажутся ниже критических, сделаем вывод о том, что распределение признака не отличается от нормального.

Расчет эмпирических показателей асимметрии и эксцесса будем производить по формулам данным ранее.

Сначала сделаем расчет промежуточных значений, который удобно выполнять поэтапно, занося данные в таблицу

## Расчет промежуточных значений

№	$x_i$	$(x_i - \bar{x})$	$(x_i - \bar{x})^2$	$(x_i - \bar{x})^3$	$(x_i - \bar{x})^4$
1	11	0,94	0,884	0,831	0,781
2	13	2,94	8,644	25,412	74,712
3	12	1,94	3,764	7,301	14,165
4	9	-1,06	1,124	-1,191	1,262
5	10	-0,06	0,004	-0,000	0,000
6	11	0,94	0,884	0,831	0,781
7	8	-2,06	4,244	-8,742	18,009
8	10	-0,06	0,004	-0,000	0,000
9	15	4,94	24,404	120,554	595,536
10	14	3,94	15,524	61,163	240,982
11	8	-2,06	4,244	-8,742	18,009
12	7	-3,06	9,364	-28,653	87,677
13	10	-0,06	0,004	-0,000	0,000
14	10	-0,06	0,004	-0,000	0,000
15	5	-5,06	25,604	-129,554	655,544
16	8	-2,06	4,244	-8,742	18,009
Суммы	161		102,944	30,468	1725,467

Для расчетов в таблице, необходимо значение среднего арифметического, которое вычисляется по формуле:

$$\bar{x} = \frac{\sum x_i}{n}$$

где  $x_i$  - каждое наблюдаемое значение признака;  
 $n$  - количество наблюдений.

В данном случае:

$$\bar{x} = \frac{161}{16} = 10,06$$

Стандартное отклонение (сигма) вычисляется по формуле:

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

где  $x_i$  - каждое наблюдаемое значение признака;  
 $\bar{x}$  - среднее значение (среднее арифметическое);  
 $n$  - количество наблюдений.

В данном случае:

$$\sigma = \sqrt{\frac{102,944}{16 - 1}} = \sqrt{6,893} = 2,62$$

Подставляя в формулы для расчета  $A$  и  $E$  полученные значения  $n$ ,  $\sigma$  и соответствующие значения из таблицы, получаем:

$$A = \frac{+30,468}{16 \cdot 2,62^3} = +0,106$$

$$E = \frac{1725,467}{16 \cdot 2,62^4} - 3 = -0,711$$

Теперь рассчитаем критические значения для показателей  $A$  и  $E$  по формулам Е.И. Пустыльника:

$$A_{\text{кр}} = 3 \cdot \sqrt{\frac{6 \cdot (n-1)}{(n+1) \cdot (n+3)}}$$

$$E_{\text{кр}} = 5 \cdot \sqrt{\frac{24 \cdot n \cdot (n-2) \cdot (n-3)}{(n+1)^2 \cdot (n+3) \cdot (n+5)}}$$

где  $n$  - количество наблюдений.



В данном случае:

$$A_{кр} = 3 \cdot \sqrt{\frac{6 \cdot (16 - 1)}{(16 + 1) \cdot (16 + 3)}} = 3 \cdot \sqrt{\frac{90}{323}} = 1,58$$

$$E_{кр} = 5 \cdot \sqrt{\frac{24 \cdot 16 \cdot (16 - 2) \cdot (16 - 3)}{(16 + 1)^2 \cdot (16 + 3) \cdot (16 + 5)}} = 5 \cdot \sqrt{\frac{69888}{115311}} = 3,89$$

$$A_{эмп} = 0,106$$

$$A_{эмп} < A_{кр}$$

$$E_{эмп} = -0,711$$

$$E_{эмп} < E_{кр}$$

Так как эмпирические значения  $A$  и  $E$  меньше критических значений, то можно сделать следующий вывод: распределение результативного признака в данном примере не отличается от нормального распределения.

Значения функции  $\varphi(t) = \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}}$

Значения функции  $\varphi(t) = \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}}$

Целые и десятичные доли <i>t</i>	Сотые доли <i>t</i>									
	0	1	2	3	4	5	6	7	8	9
0,0	0,3989	0,3989	0,3989	0,3988	0,3986	0,3984	0,3982	0,3980	0,3977	0,3973
0,1	3970	3965	3961	3956	3951	3945	3939	3932	3925	3918
0,2	3910	3902	3894	3885	3876	3876	3857	3847	3836	3825
0,3	3814	3802	3790	3778	3765	3752	3739	3726	3712	3697
0,4	3683	3668	3653	3637	3621	3605	3589	3572	3555	3538
0,5	3521	3503	3485	3467	3448	3429	3410	3391	3372	3352
0,6	3332	3312	3292	3271	3251	3230	3209	3187	3166	3144
0,7	3123	3101	3079	3056	3034	3011	2989	2966	2943	2920
0,8	2897	2874	2850	2827	2803	2780	2756	2732	2709	2685
0,9	2661	2637	2613	2589	2565	2541	2516	2592	2468	2444
1,0	0,2420	0,2396	0,2371	0,2347	0,2323	0,2299	0,2275	0,2251	0,2227	0,2203
1,1	2179	2155	2131	2107	2083	2059	2036	2012	1989	1965
1,2	1942	1919	1895	1872	1849	1926	1804	1781	1758	1736
1,3	1714	1691	1669	1647	1626	1604	1582	1561	1539	1518
1,4	1497	1476	1456	1435	1415	1394	1374	1354	1334	1315
1,5	1295	1276	1257	1238	1219	1200	1182	1163	1145	1127
1,6	1109	1092	1074	1057	1040	1023	1006	0989	0973	0957
1,7	0940	0925	0909	0893	0878	0863	0848	0833	0818	0804
1,8	0790	0775	0761	0748	0734	0721	0707	0694	0681	0669
1,9	0656	0644	0632	0620	0608	0596	0584	0573	0562	0551
2,0	0,0540	0,0529	0,0519	0,0508	0,0498	0,0488	0,0478	0,0468	0,0459	0,0449
2,1	0440	0431	0422	0413	0404	0396	0387	0379	0371	0363
2,2	0355	0347	0339	0332	0325	0317	0310	0241	0297	0290
2,3	0283	0277	0270	0264	0258	0252	0246	0303	0235	0229
2,4	0224	0219	0213	0208	0203	0198	0194	0189	0184	0180
2,5	0175	0171	0167	0163	0158	0154	0151	0147	0143	0139
2,6	0136	0132	0129	0126	0122	0119	0116	0113	0110	0107
2,7	0104	0101	0099	0096	0093	0091	0088	0086	0084	0081
2,8	0079	0077	0075	0073	0071	0069	0067	0065	0063	0061
2,9	0060	0058	0056	0055	0053	0051	0050	0048	0047	0046
3,0	0,0044	0,0043	0,0042	0,0040	0,0039	0,0038	0,0037	0,0036	0,0035	0,0034
3,1	0033	0032	0031	0030	0029	0028	0,0037	0026	0025	0025
3,2	0024	0023	0022	0022	0021	0020	0020	0019	0018	0018
3,3	0017	0017	0016	0016	0015	0015	0014	0014	0013	0013
3,5	0009	0008	0008	0008	0011	0010	0010	0010	0009	0009
3,6	0006	0006	0006	0005	0008	0007	0007	0005	0007	0006
3,7	0004	0004	0004	0004	0005	0005	0005	0007	0005	0004
3,8	0003	0003	0003	0003	0004	0004	0003	0003	0003	0003
3,9	0002	0002	0002	0002	0003	0002	0002	0002	0002	0002
4,0	0,0001	0,0001	0,0001	0,0001	0002	0002	0002	0002	0001	0001
4,1	0,0001338	—	—	—	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001
4,5	0,0000160	—	—	—	—	—	—	—	—	—
5,0	0,0000015	—	—	—	—	—	—	—	—	—

# Основные формулы

$$f_{\text{ОТН}} = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(x_i - \bar{x}_B)^2}{2\sigma^2}}$$

**(1) Нормальный закон  
распределения**

$$A_s = \frac{\sum_{i=1}^k (x_i - \bar{x}_B)^3}{n\sigma^3}$$

**(2) Асимметрия**

$$E_x = \frac{\sum_{i=1}^k (x_i - \bar{x}_B)^4}{n\sigma^4} - 3$$

**(3) Экцесс**

# Вопросы для самоконтроля

1. Дайте описание нормального закона распределения
2. Смысл параметров нормального закона распределения
3. Алгоритм построения кривой нормального распределения
4. Проверка нормальности распределения результативного признака

# Задача для самоконтроля

При определении степени выраженности некоторого психического свойства в опытной группе были получены следующие результаты.

Опытная группа – 18, 15, 16, 11, 14, 15, 16, 16, 16, 22, 17, 12, 11, 12, 18, 19, 20

Построить кривую распределения признака и дать заключение об отклонении данного распределения от нормального.

**СПАСИБО ЗА ВНИМАНИЕ!**