Проверка статистических гипотез

Тема 11

План лекции:

- 11.1. Понятие статистической гипотезы
- 11.2. Классический метод проверки гипотез
- 11.3. Сущность метода проверки нулевой гипотезы
- 11.4. Алгоритм проверки нулевой гипотезы
- 11.5. Проверка гипотез о законе распределения
- 11.6. Проверка гипотезы о нормальном распределении генеральной совокупности по критерию Пирсона

11.1. Понятие статистической гипотезы

Статистической гипотезой называется всякое высказывание о генеральной совокупности (случайной величине), проверяемое по выборке (то есть по результатам наблюдений).

Различают гипотезы **простые**, содержащие только одно предположение, и **сложные**, содержащие более одного предположения.

Процедура сопоставления гипотезы с выборочными данными называется проверкой гипотезы. Для проверки гипотез используют аналитические и статистические методы.

11.2. Классический метод проверки гипотез

В соответствии с поставленной задачей и на основании выборочных данных формулируется (выдвигается) гипотеза H_0, которая называется **основной** или ***нулевой***. Одновременно с выдвинутой гипотезой H_0, рассматривается противоположная ей гипотеза H_1, которая называется **конкурирующей** или ***альтернативной***.

Для проверки нулевой гипотезы вводят специально подобранную случайную величину K, распределение которой известно и называют ее ***критерием***.

Поскольку гипотеза H_0 для генеральной совокупности принимается по выборочным данным, то она может быть ошибочной. При этом возможны ошибки двух родов.

Ошибка первого рода состоит в том, что отвергается гипотеза H_0, когда она на самом деле верна.

Ошибка второго рода состоит в том, что отвергается альтернативная гипотеза H_1, когда она на самом деле верна.

1) Для определения вероятности ошибки первого рода вводится параметр α:

$\alpha = P_{H_0}(H_1)$ - вероятность того, что будет принята гипотеза H_1, при условии, что H_0 верна. Величину α называют *уровнем значимости*. Обычно α выбирают в пределах $0{,}001 - 0{,}1$.

2) Вероятность ошибки второго рода определяется параметром β:

$\beta = P_{H_1}(H_0)$ - вероятность того, что будет принята гипотеза H_0, при условии, что H_1 верна. Величину $(1-\beta)$, то есть недопустимость ошибки второго рода (отвергнуть неверную и принять верную гипотезу H_1) называют *мощностью критерия*.

11.3. Сущность метода проверки нулевой гипотезы

Множество всех значений критерия разбивают на два непересекающихся подмножества: одно из них содержит значения критерия, при которых нулевая гипотеза H_0 отвергается; другое – при которых она принимается.

Критической областью называется совокупность значений критерия, при которых нулевую гипотезу отвергают.

Областью принятия гипотезы (областью допустимых значений) называется совокупность значений критерия, при которых нулевую гипотезу принимают.

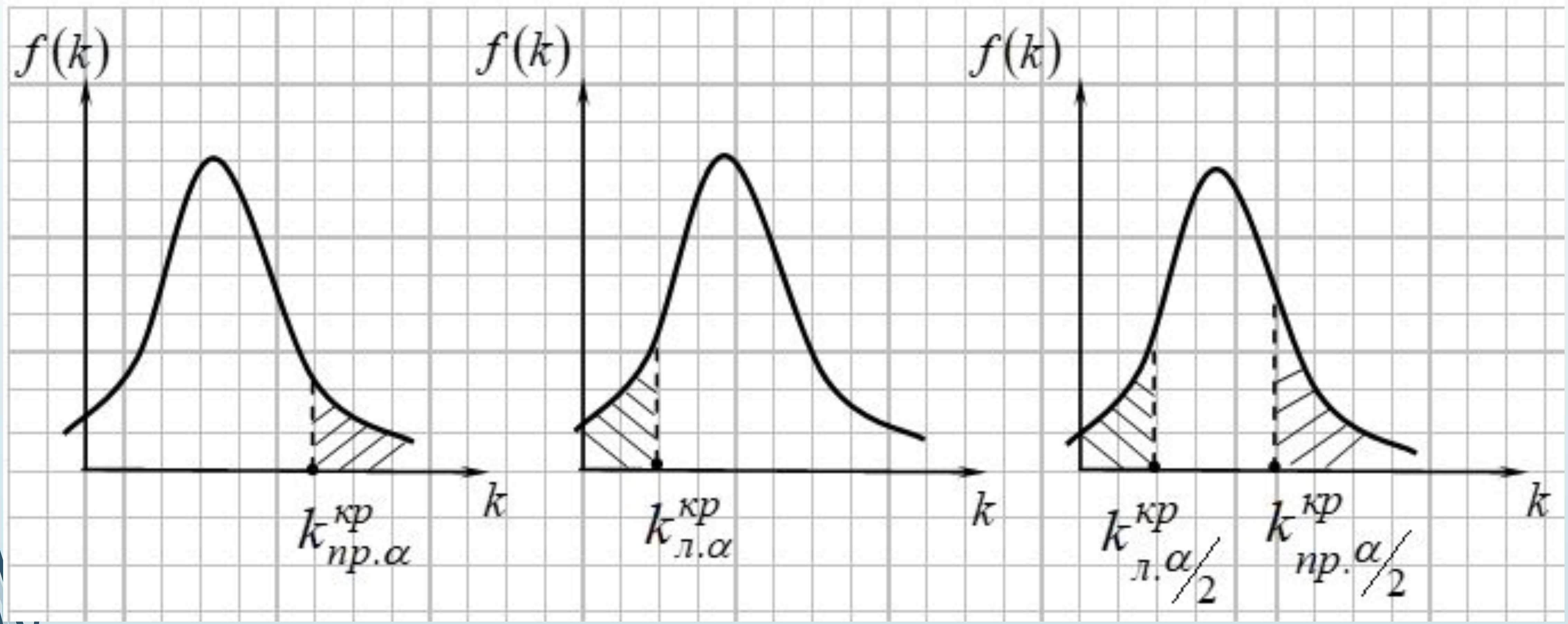
Обозначим критическую область ω. Если вычисленное по выборке значение критерия K попадает в критическую область ω, то гипотеза H_0 отвергается и принимается гипотеза H_1. В этом случае можно совершить ошибку первого рода, вероятность которой равна α. Иначе, вероятность того, что критерий K примет значение из критической области ω, должна быть равна заданному значению α, то есть $P(K \in \omega) = \alpha$.

Критическая область ω определяется неоднозначно. Возможны три случая расположения ω. Они определяются видом нулевой и альтернативной гипотез и законом распределения критерия K.

Правосторонняя критическая область состоит из интервала $\left(k_{пр.\alpha}^{кр}; +\infty\right)$, где $k_{пр.\alpha}^{кр}$ определяется из условия $P\left(K > k_{пр.\alpha}^{кр}\right) = \alpha$ и называется правосторонней точкой, отвечающей уровню значимости α.

Левосторонняя критическая область состоит из интервала $\left(-\infty; k_{л.\alpha}^{кр}\right)$, где $k_{л.\alpha}^{кр}$ определяется из условия $P\left(K < k_{л.\alpha}^{кр}\right) = \alpha$ и называется левосторонней точкой, отвечающей уровню значимости α.

Двусторонняя критическая область состоит из следующих двух интервалов: $\left(-\infty; k_{л.\alpha/2}^{кр}\right)$ и $\left(k_{пр.\alpha/2}^{кр}; +\infty\right)$, где точки $k_{л.\alpha/2}^{кр}$ и $k_{пр.\alpha/2}^{кр}$ определяются из условий $P\left(K < k_{л.\alpha/2}^{кр}\right) = \frac{\alpha}{2}$ и $P\left(K > k_{пр.\alpha/2}^{кр}\right) = \frac{\alpha}{2}$ и называются двусторонними критическими точками.



f(k)
k
$k^{кр}_{пр.\alpha}$
f(k)
k
$k^{кр}_{л.\alpha}$
f(k)
k
$k^{кр}_{л.\alpha/2}$
$k^{кр}_{пр.\alpha/2}$

11.4. Алгоритм проверки нулевой гипотезы

1. Располагая выборкой, формулируют нулевую гипотезу H_0 и альтернативную гипотезу H_1.

2. Выбирают критерий проверки гипотезы H_0, зависящий от выборочных данных и условий рассматриваемой задачи. Наиболее часто используют случайные величины, имеющие следующие законы распределения: нормальный, Стьюдента, Фишера-Снедекора, хи-квадрат.

3. Задают уровень значимости выбранного критерия и определяют соответствующую ему критическую область. Для определения критической области достаточно найти критическую точку $t_{кр}$ - ее границу. Для каждого критерия имеются таблицы, по которым находят критическую точку.

4. Вычисляют значение критерия по результатам произведенных измерений и сравнивают с критической точкой.

5. Нулевую гипотезу *отвергают*, если вычисленное значение критерия попадает в критическую область, или считают *справедливой*, если оно окажется внутри области допустимых значений.

11.5. Проверка гипотез о законе распределения

Во многих случаях закон распределения изучаемой случайной величины X неизвестен, но есть основания предположить, что он имеет вполне определенный вид: нормальный, экспоненциальный или какой-либо другой.

Пусть выдвинута гипотеза H_0 о каком-либо законе распределения. Для проверки этой гипотезы H_0 требуется по выборке сделать заключение, согласуются ли результаты наблюдений с высказанным предположением.

Статистический критерий проверки гипотезы о предполагаемом законе неизвестного распределения называется *критерием согласия.*

Он используется для проверки согласия предполагаемого вида распределения с опытными данными на основании выборки.

Существуют различные критерии согласия: Пирсона, Колмогорова, Фишера и другие. Наиболее часто применяется критерий Пирсона.

11.6. Проверка гипотезы о нормальном распределении генеральной совокупности по критерию Пирсона

Пусть выборка из генеральной совокупности X задана в виде статистического интервального ряда:

Таблица 1

$[x_1, x_2)$	$[x_2, x_3)$	$\dots$	$[x_m, x_{m+1})$
n_1	n_2	$\dots$	n_m

где n_i - интервальные частоты, $\sum_{i=1}^{m} n_i = n$ - объем выборки,

m - число интервалов, h - длина интервала, x_i - середина интервала.

Требуется проверить гипотезу H_0 о том, что генеральная совокупность X распределена по нормальному закону, применяя критерий Пирсона.

Правило проверки

1. Вычисляем $\overline{x_в}$ и $\sigma_в$.

2. Находим теоретические частоты n_i'.

Их можно вычислить двумя способами.

<u>Первый способ</u>

$$n_i' = \frac{n \cdot h}{\sigma_в} \cdot \varphi(t_i),$$

где n - объем выборки, h - шаг, $t_i = \frac{x_i - \overline{x_в}}{\sigma_в}$; $\varphi(x) = \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{x^2}{2}}$ - функция Гаусса, значение которой в точке t_i находим по таблице.

$P_i = \frac{\varphi(t_i)}{\sigma_в} \cdot h$ - вероятность попадания значений случайной величины X в i - й интервал.

Для вычисления n_i' составляем табл. 2.

Таблица 2

i	x_i	n_i	$x_i - \overline{x_в}$	t_i	$\varphi(t_i)$	P_i	$n_i' = P_i \cdot n$
1	x_1	n_1	$x_1 - \overline{x_в}$	t_1	$\varphi(t_1)$	P_1	$n_1' = P_1 \cdot n$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
m	x_m	n_m	$x_m - \overline{x_в}$	t_m	$\varphi(t_m)$	P_m	$n_m' = P_m \cdot n$
$\sum$		n				1	n

<u>*Второй способ:*</u>

$$n_i' = P_i \cdot n$$

где n - объем выборки, $z_i = \dfrac{x_i - \overline{x_в}}{\sigma_в}$, $P_i = \Phi(z_{i+1}) - \Phi(z_i)$ - вероятность попадания X в i - й интервал, $\Phi(z)$ - значение функции Лапласа.

Полагают $z_1 = (-\infty)$, $z_{m+1} = +\infty$.

Для вычисления n_i' составляем табл. 3.

Таблица 3

i	Границы интервала		n_i	Границы интервала		$\Phi(z_i)$	$\Phi(z_{i+1})$	P_i	n_i'
	x_i	x_{i+1}		z_i	z_{i+1}				
1	x_1	x_2	n_1	$-\infty$	z_2	-0,5	$\Phi(z_2)$	P_1	n_1'
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
m	x_m	x_{m+1}	n_m	z_m	$+\infty$	$\Phi(z_m)$	0,5	P_m	n_m'
$\sum$			n					1	n

3. Сравниваем эмпирические (n_i) и теоретические (n_i') частоты с помощью критерия Пирсона.

Для этого:

1) составляем расчетную табл.4 , по которой находим

$\chi^2_{набл}$ - наблюдаемое значение критерия $\chi^2_{набл} = \sum_{i=1}^{m} \frac{(n_i - n_i')^2}{n_i'}$

Таблица 4.

i	n_i	n_i'	$n_i - n_i'$	$(n_i - n_i')^2$	$\frac{(n_i - n_i')^2}{n_i'}$	n_i^2	$\frac{n_i^2}{n_i'}$
1	n_1	n_1'	$n_1 - n_1'$	$(n_1 - n_1')^2$	$\frac{(n_1 - n_1')^2}{n_1'}$	n_1^2	$\frac{n_1^2}{n_1'}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
m	n_m	n_m'	$n_m - n_m'$	$(n_m - n_m')^2$	$\frac{(n_m - n_m')^2}{n_m'}$	n_m^2	$\frac{n_m^2}{n_m'}$
$\sum$	n				$\chi^2_{набл}$		

Контроль: $\chi^2_{набл} = \sum_{i=1}^{m} \frac{n_i^2}{n_i'} - n$.

2) Находим число степеней свободы k: $k = m - r - 1$

где m - число интервалов; r - число параметров предполагаемого распределения,

Для нормального распределения $k = m - 3$, так как $r = 2$ (нормальный закон распределения характеризуется двумя параметрами a и σ).

4. В таблице критических точек (*квантилей*) распределения χ^2 по заданному уровню значимости α и числу степеней свободы находим $\chi^2_{кр}(\alpha; k)$ правосторонней критической области.

Если $\chi^2_{набл} < \chi^2_{кр}$ - нет оснований отвергнуть гипотезу H_0 о нормальном распределении генеральной совокупности.

Если $\chi^2_{набл} > \chi^2_{кр}$ - гипотезу отвергаем.

Замечание.

1) Объем выборки должен быть достаточно велик $(n \geq 50)$.

2) Малочисленные частоты $(n_i < 5)$ следует объединить. В этом случае и соответствующие им теоретические частоты также надо сложить.

Если производилось объединение частот, то при определении числа степеней свободы по формуле $k = m - 3$ следует в качестве m принять число интервалов, оставшихся после объединения частот.

Пример. Пусть из генеральной совокупности X задана выборка объемом 50. Требуется проверить гипотезу H_0 о нормальном распределении генеральной совокупности по данной выборке.

1. Из рассмотренных выше примеров известно:

- интервальный ряд табл. 5

Таблица 5

Интервалы	$[-2{,}06;-1{,}46)$	$[-1{,}46;-0{,}86)$	$[-0{,}86;-0{,}26)$	$[-0{,}26;0{,}34)$
Частоты n_i	2	6	11	15
Интервалы	$[0{,}34;0{,}94)$	$[0{,}94;1{,}54)$	$[1{,}54;2{,}14)$	
Частоты n_i	11	3	2	

$\sum_{i=1}^{7} n_i = 50$.

- числовые характеристики выборки $\overline{x_в} = -0{,}032$, $\sigma_в = 0{,}8195$, $A_S^* = 0{,}0759$, $E_k^* = 0{,}3112$.

2. *Проверим гипотезу* H_0 *с помощью средних квадратических отклонений коэффициентов* A_S^* *и* E_k^*.

Критерием распределения выборки по нормальному закону является равенство нулю коэффициентов A_S^* и E_k^*.

Если они отличны от нуля, то для предварительного выбора закона распределения вычисляют средние квадратические отклонения для A_S^* и E_k^*

$$\sigma^*_{A_S} = \sqrt{\frac{6(n-1)}{(n+1)\cdot(n+3)}} \qquad \sigma^*_{E_k} = \sqrt{\frac{24n\cdot(n-2)\cdot(n-3)}{(n-1)^2\cdot(n+3)\cdot(n+5)}}$$

Если A_S^* и E_k^* отличаются по модулю от нуля не более чем на удвоенные средние квадратические отклонения, то есть $\left|A_S^*\right| \le 2\sigma_{A_S^*}$ и $\left|E_k^*\right| \le 2\sigma_{E_k^*}$, то можно предположить, что данная выборка распределена по нормальному закону.

Рассчитаем $\sigma^*_{A_S} = \sqrt{\dfrac{6(50-1)}{(50+1)\cdot(50+3)}} = \sqrt{\dfrac{294}{2703}} = \sqrt{0{,}1087} = 0{,}3297$

$$\sigma^*_{E_k} = \sqrt{\frac{1200\cdot 48\cdot 47}{2401\cdot 53\cdot 55}} = \sqrt{\frac{2707200}{6998915}} = \sqrt{0{,}3868} = 0{,}6219 .$$

Для A^*_S условие критерия выполняется: $|0{,}0759| < 2\cdot 0{,}3297$.

Для E^*_k условие критерия выполняется: $|-0{,}3112| < 2\cdot 0{,}6219$.

Гипотезу H_0 принимаем, то есть можно предположить, что генеральная совокупность X распределена по нормальному закону.

3. *Проверим гипотезу* H_0 *по критерию Пирсона.*

1) $\overline{x_в} = -0{,}032$, $\sigma_в = 0{,}8195$.

2) Найдем теоретические частоты n_i' вторым способом.

Интервальный ряд (табл.5) содержит интервалы с частотами меньшими 5. Следовательно, два первых и два последних интервала объединяем, при этом соответствующие частоты суммируем.

Составим расчетную табл.6

Таблица 6

i	Границы интервала		n_i	Границы интервала		$\Phi(z_i)$	$\Phi(z_{i+1})$	P_i	n_i'
	x_i	x_{i+1}		z_i	z_{i+1}				
1	-2,06	-0,86	8	$-\infty$	-1,01	-0,5	-0,3438	0,1562	7,81
2	-0,86	-0,26	11	-1,01	-0,28	-0,3438	-0,1103	0,2335	11,675
3	-0,26	0,34	15	-0,28	0,45	-0,1103	0,1736	0,2839	14,195
4	0,34	0,94	11	0,45	1,19	0,1736	0,3830	0,2094	10,47
5	0,94	2,14	5	1,19	$+\infty$	0,3830	0,5	0,1170	5,85
$\sum$								**1**	**50**

3) Сравним эмпирические (n_i) и теоретические (n_i') частоты. Для этого составляем расчетную табл.7

Таблица 7

i	n_i	n_i'	$n_i - n_i'$	$(n_i - n_i')^2$	$\frac{(n_i - n_i')^2}{n_i'}$	n_i^2	$\frac{n_i^2}{n_i'}$
1	8	7,810	0,190	0,0361	0,0046	64	8,1946
2	11	11,675	-0,675	0,4556	0,0390	121	10,3640
3	15	14,195	0,805	0,6480	0,0457	225	15,8507
4	11	10,470	0,530	0,2809	0,0268	121	11,5568
5	5	5,850	-0,850	0,7225	0,1235	25	4,2735
$\sum$					**0,2396**		**50,2396**

Контроль: $\sum_{i=1}^{5} \frac{n_i^2}{n_i'} - n = 50,2396 - 50 = 0,2396$

$\chi^2_{набл} = 0,2396$. Расчеты проведены верно.

4) Зададим $\alpha = 0,05$.

Вычислим число степеней свободы $k = 5 - 3 = 2$ и найдем $\chi^2_{кр}(0,05; 2) = 6,0$. Получим $\chi^2_{набл} < \chi^2_{кр}$.

Следовательно, нет оснований отвергать гипотезу H_0 о нормальном распределении генеральной совокупности X.

Другими словами различие между эмпирическими (n_i) и теоретическими (n_i') частотами незначительное (случайное), которое можно объяснить малым объемом выборки.

Построим нормальную кривую. Для этого составим табл.8.

Таблица 8

Середины интервалов	-1,76	-1,16	-0,56	0,04	0,64	1,24	1,84
$\frac{p_i}{h}$	0,05	0,19	0,39	0,52	0,34	0,14	0,03

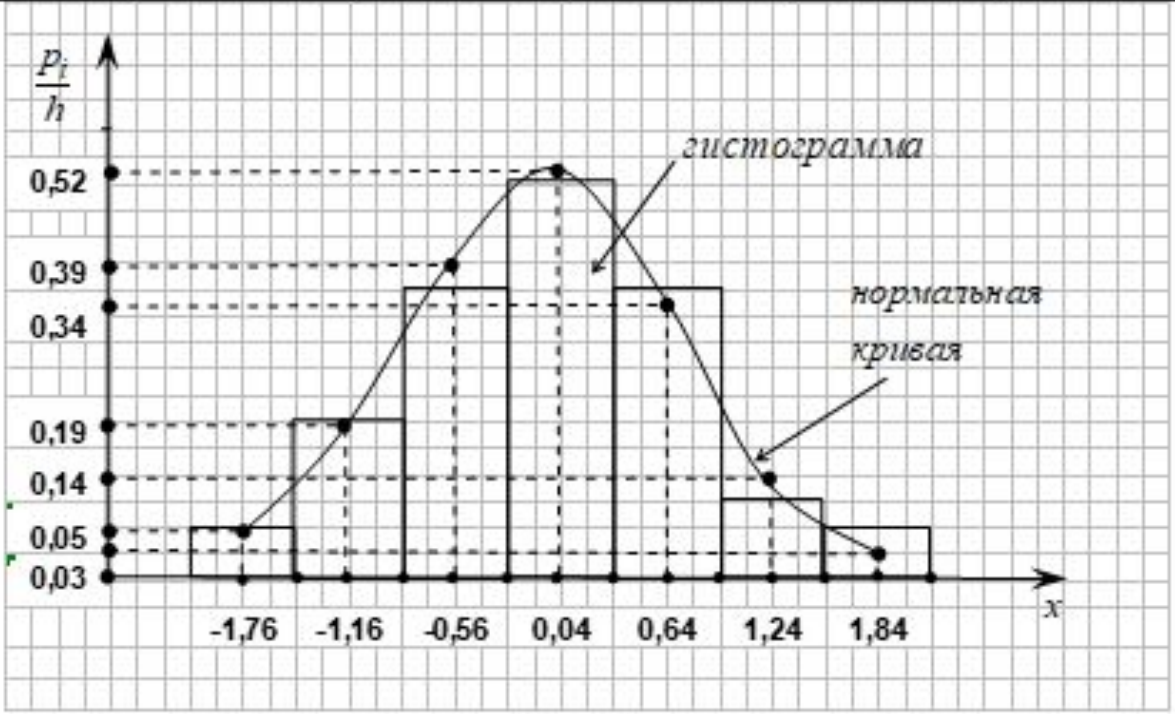


Рис.2

Так как гипотеза о нормальном распределении не отвергается, то нормальная кривая хорошо сглаживает гистограмму.