КРАТКОЕ СОДЕРЖАНИЕ ЗАНЯТИЯ 5

Доц. Гарбузова
Таисия Георгиевна

Рекомендуемая литература:

- **1.** М.Г. Назаров. Общая теория статистики. Учебник. [Электронный ресурс] : Учебники — Электрон. дан. — М. : Омега-Л, 2010. — 410 с. — Режим доступа: http://e.lanbook.com/book/5534 . Раздел «Экономика и менеджмент».
- **2.** Годин, А.М. Статистика: Учебник. [Электронный ресурс] : Учебники — Электрон. дан. — М. : Дашков и К, 2011. — 460 с. — Режим доступа: http://e.lanbook.com/book/967 . Раздел «Экономика и менеджмент».
- **3.**Балдин, К.В. Общая теория статистики: Учебное пособие. [Электронный ресурс] : Учебные пособия / К.В. Балдин, А.В. Рукосуев. — Электрон. дан. — М. : Дашков и К, 2010. — 312 с. — Режим доступа: http://e.lanbook.com/book/955 .Раздел «Экономика и менеджмент».

8.3.Методика исчисления показателей, характеризующих тенденцию динамики

Задание: *Определите все показатели, характеризующие тенденцию развития данного явления во времени:*

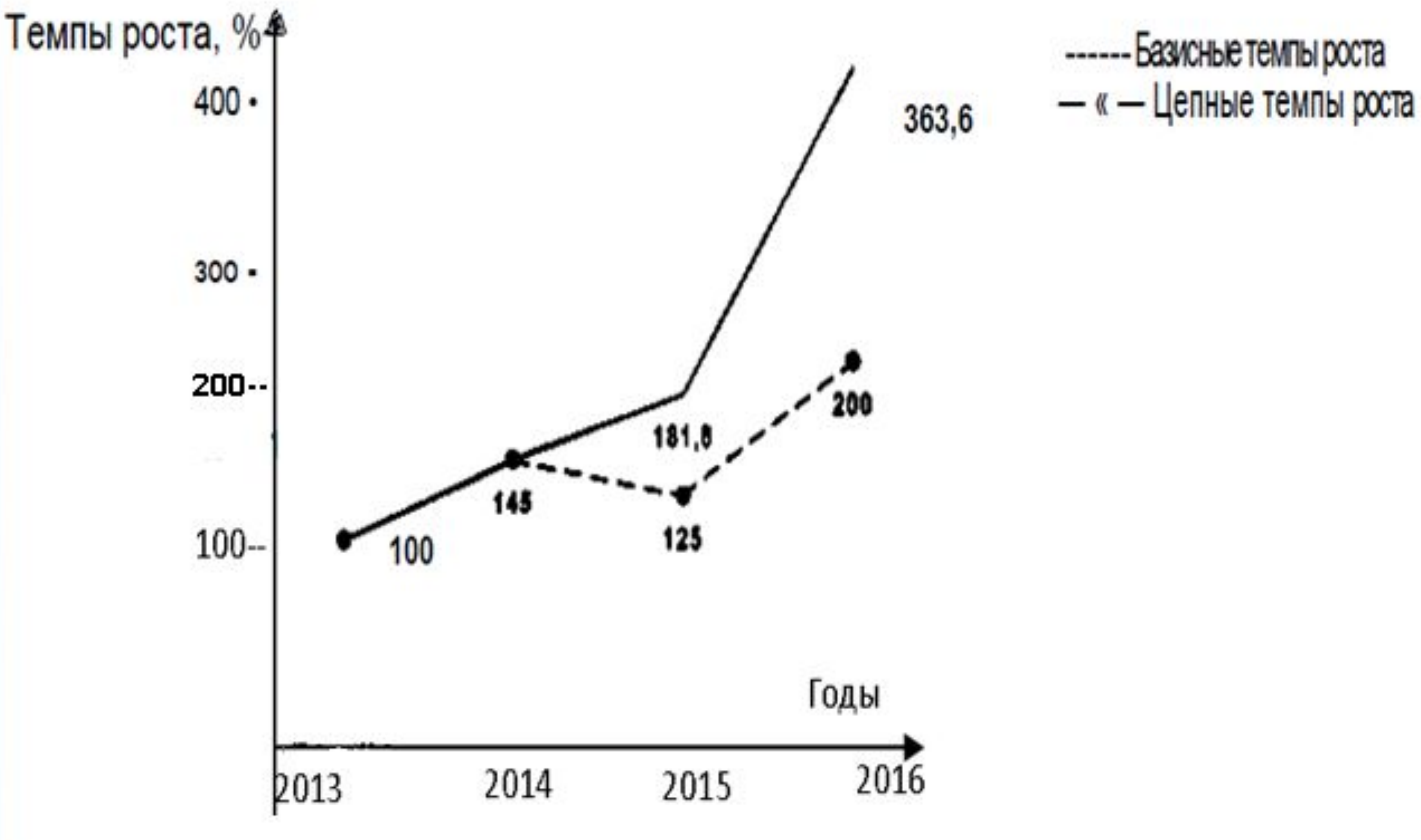
1. *Абсолютные приросты базисные (накопленные) и цепные (годовые).*
2. *Темпы роста базисные и цепные.*
3. *Темпы прироста базисные и цепные.*
4. *Абсолютное значение одного процента прироста; темп наращивания одного процента.*
5. *Средний абсолютный прирост; средний темп роста; средний темп прироста.*
6. *Постройте график базисных и цепных темпов роста.*
7. *Сделайте выводы на основании расчетов.*

8.3.Методика исчисления показателей, характеризующих тенденцию динамики

Основные показатели динамики розничного товарооборота торгового дома

Годы	Розничный товарооборот, млн руб.	Абсолютный прирост, млн руб.		Темпы роста, %		Темпы прироста, %		Абсолютные значения 1% прироста, млн руб.	Темп наращивания 1%
		базисный (накопленный)	цепной (годовой)	базисный	цепной	базисный	цепной		
А	1	2	3	4	5	6	7	8	9
2013	1100	-	-	100	-	-	-	-	-
2014	1600	1600-1100=500	1600-1100=500	$\frac{1600}{1100}\cdot 100=145$	$\frac{1600}{1100}\cdot 100=145$	45	45	$\frac{500}{45}=11$	$\frac{500100}{1100}=45$
2015	2000	2000-1100=900	2000-1600=400	$\frac{2000}{1100}\cdot 100=181{,}8$	$\frac{2000}{1600}\cdot 100=125$	81,8	25	$\frac{400}{25}=16$	$\frac{400100}{1100}=36$
2016	4000	4000-1100=2900	4000-2000=2000	$\frac{4000}{1100}\cdot 100=363{,}6$	$\frac{4000}{2000}\cdot 100=200$	263,6	100	$\frac{2000}{100}=20$	$\frac{2000100}{1100}=181{,}8$

График динамики товарооборота торгового дома за 2013-2016 годы.



9.ОПИСАТЕЛЬНЫЙ АНАЛИЗ: НОРМАЛЬНОЕ РАСПРЕДЕЛЕНИЕ

Графиком нормального распределения является симметричная колоколообразная кривая, которая задается уравнением:

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

9.ОПИСАТЕЛЬНЫЙ АНАЛИЗ: НОРМАЛЬНОЕ РАСПРЕДЕЛЕНИЕ

Кривая нормального распределения полностью определяется средней арифметической и стандартным отклонением – $N\ \bar{x},\sigma$.

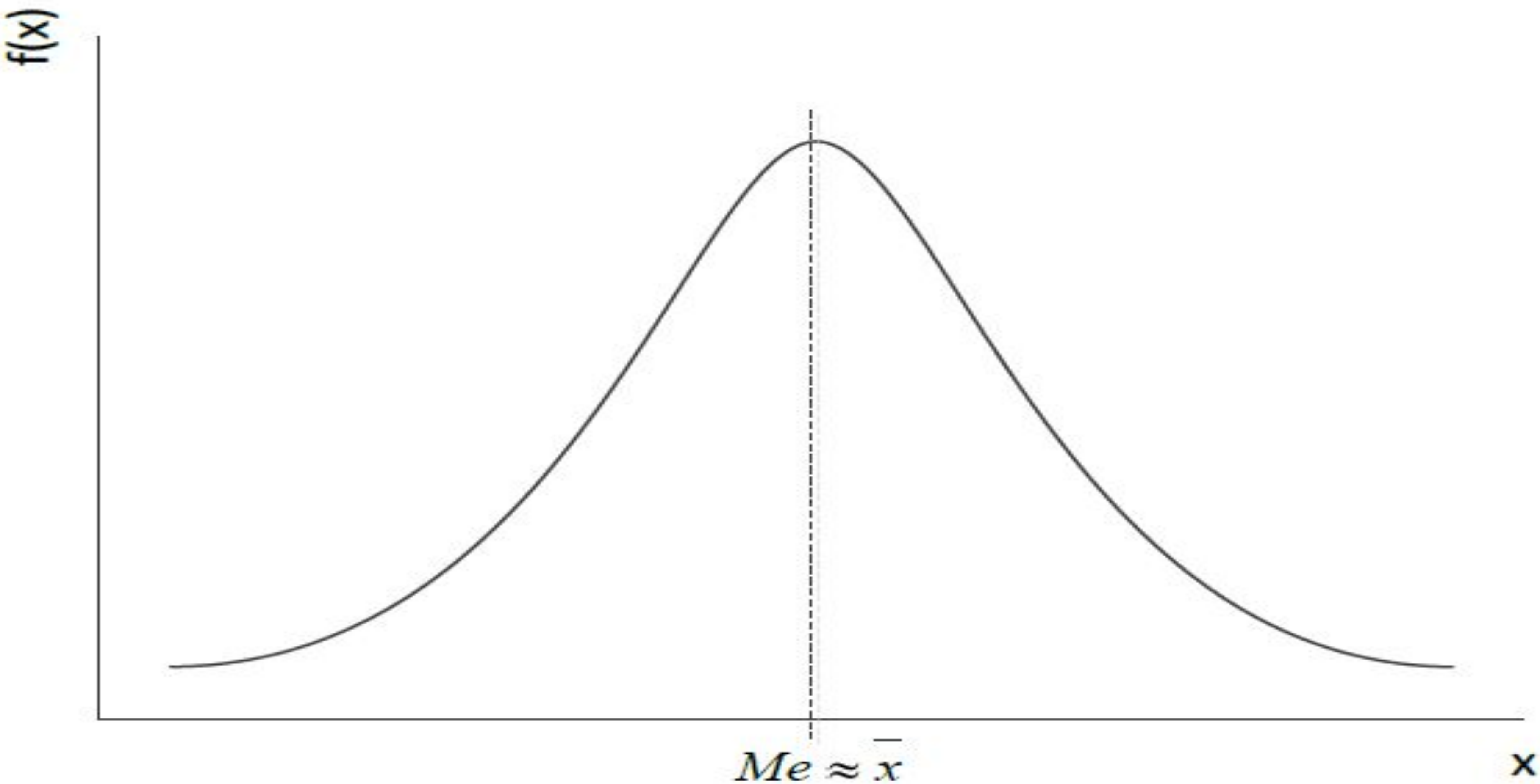
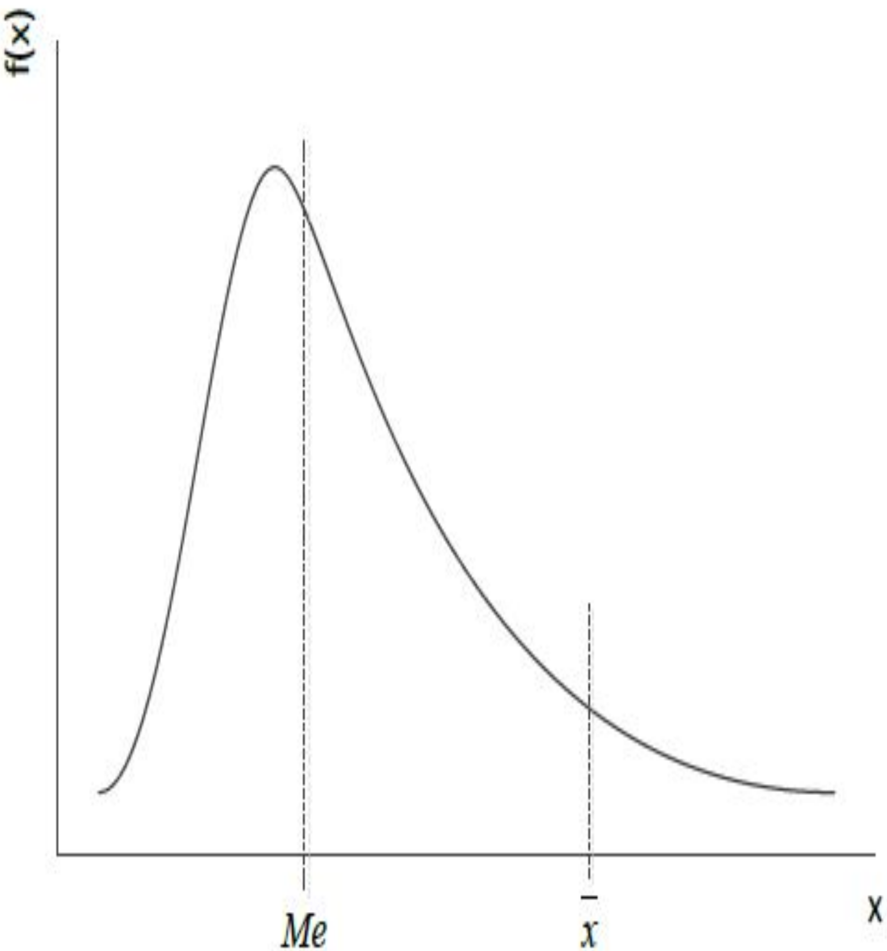
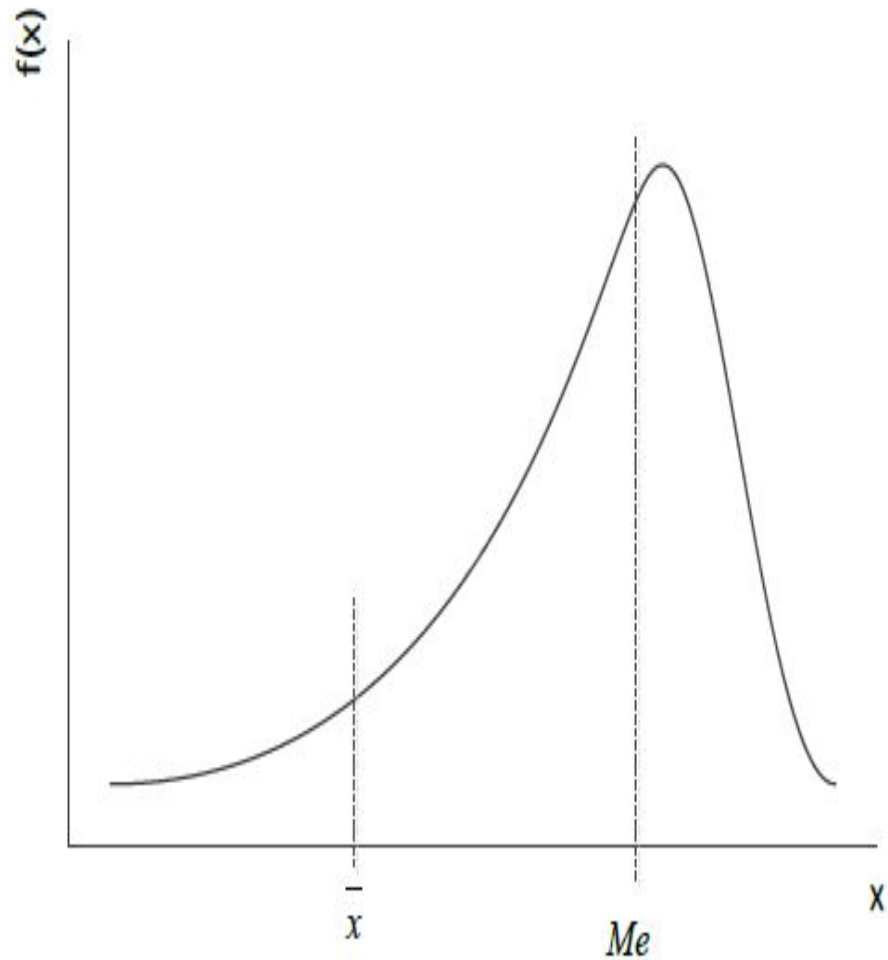


График нормального распределения

9.1.Ненормальное распределение



Правосторонняя асимметрия



Левосторонняя асимметрия

Среднее и медиана для ненормальных распределений

9.2.ПРОВЕРКА НА НОРМАЛЬНОСТЬ

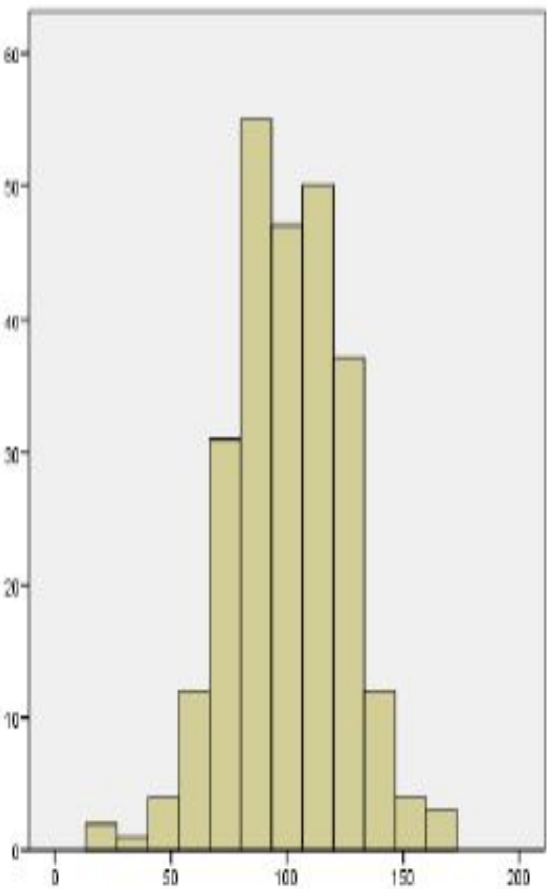
Для проверки количественных данных на нормальность используются следующие методы:

1. **Графические**

2. **Аналитические .**

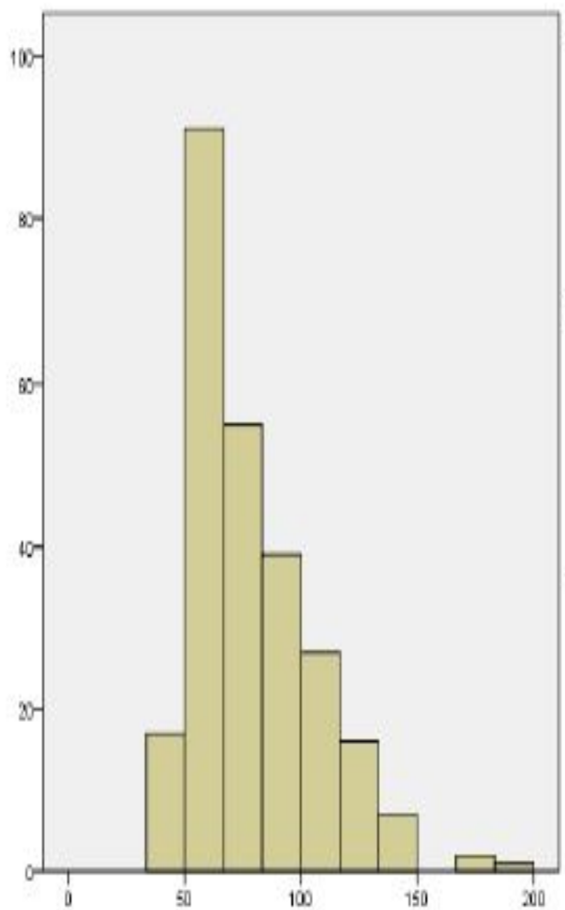
9.2.1. Графические методы
9.2.1.1.Гистограмма

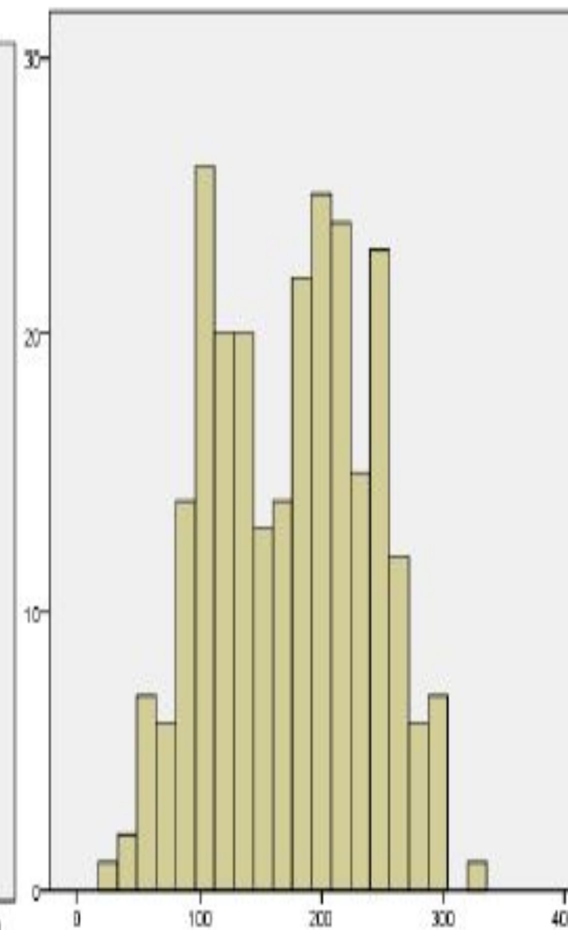
Нормальное распределение



Симметричная гистограмма

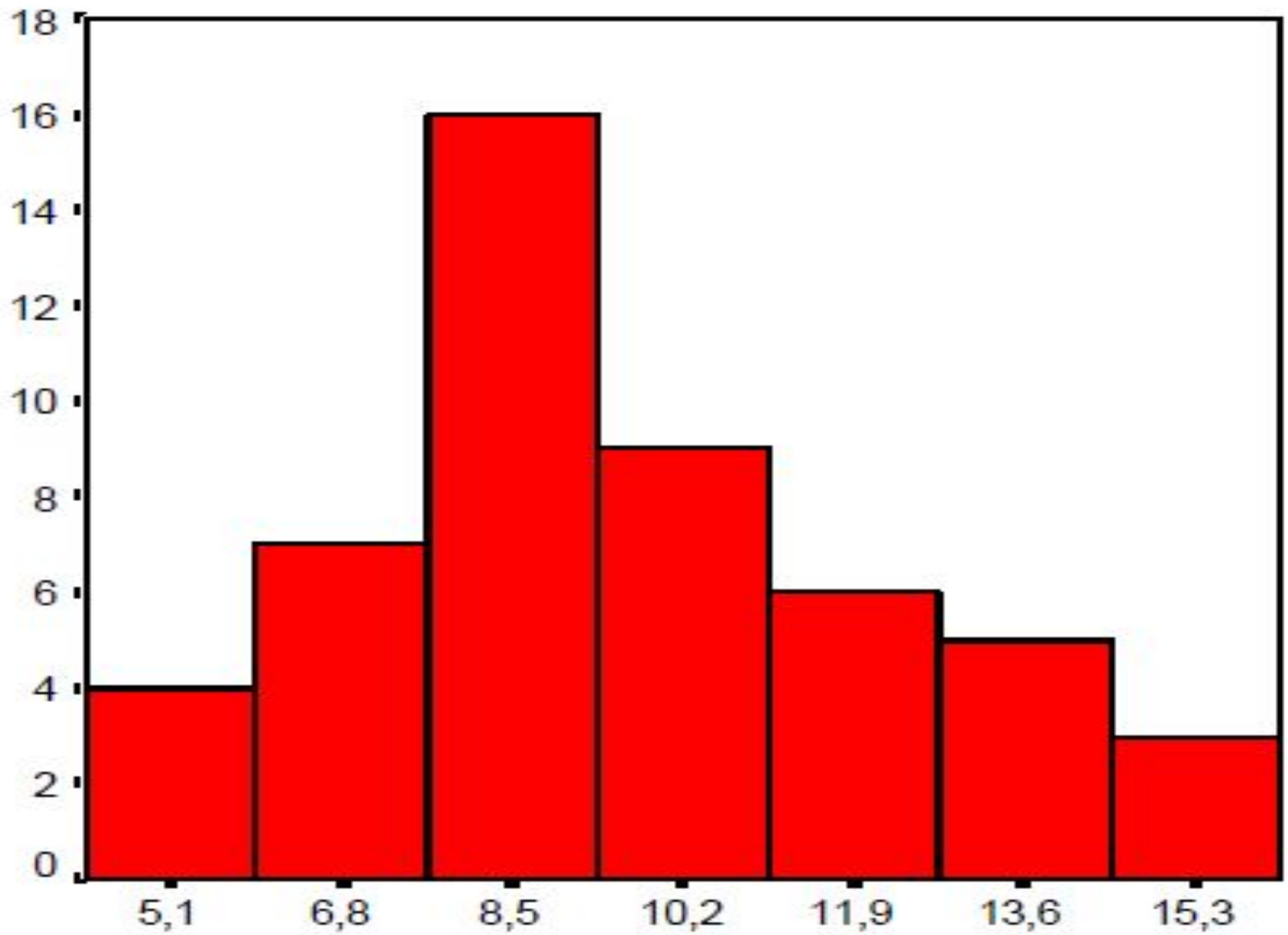
Ненормальное распределение





Асимметричная гистограмма

9.2.1. Графические методы
9.2.1.1.Гистограмма



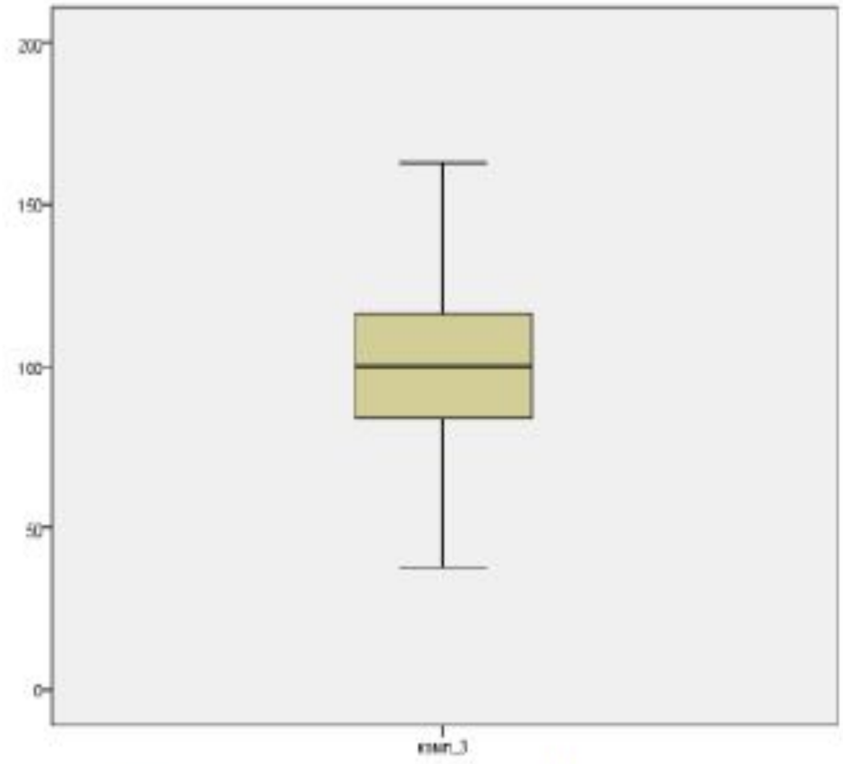
Гистограмма по продажам

9.2.1. Графические методы
9.2.1.2. «Ящик с усами»

Идея построения ящика с усами

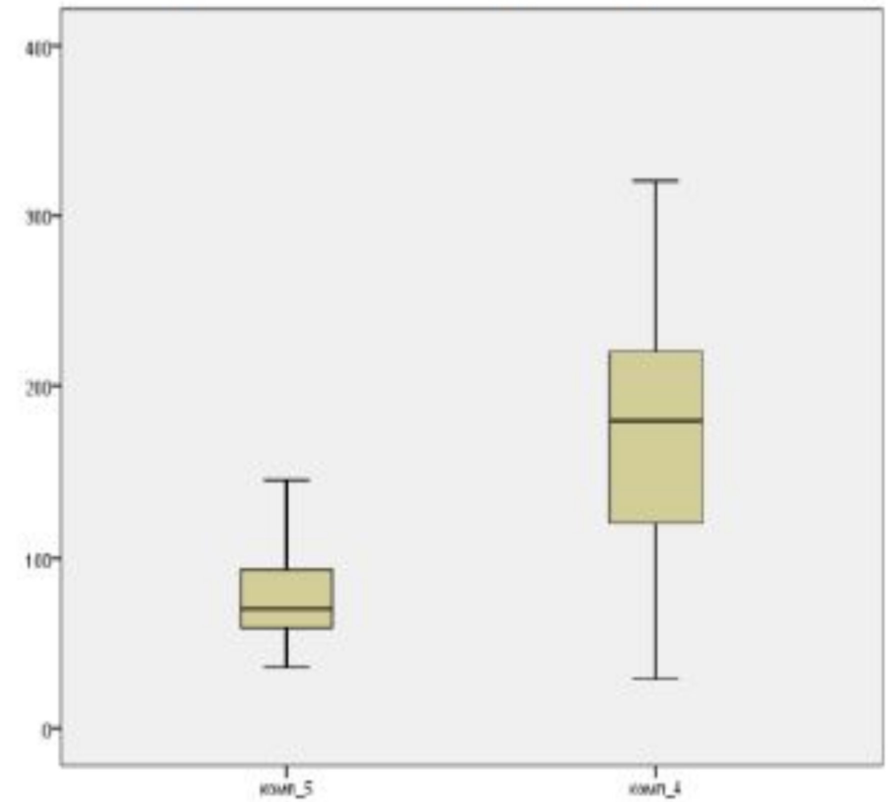
9.2.1. Графические методы
9.2.1.2. «Ящик с усами»

Нормальное распределение



Симметричный ящик

Ненормальное распределение



Асимметричный ящик

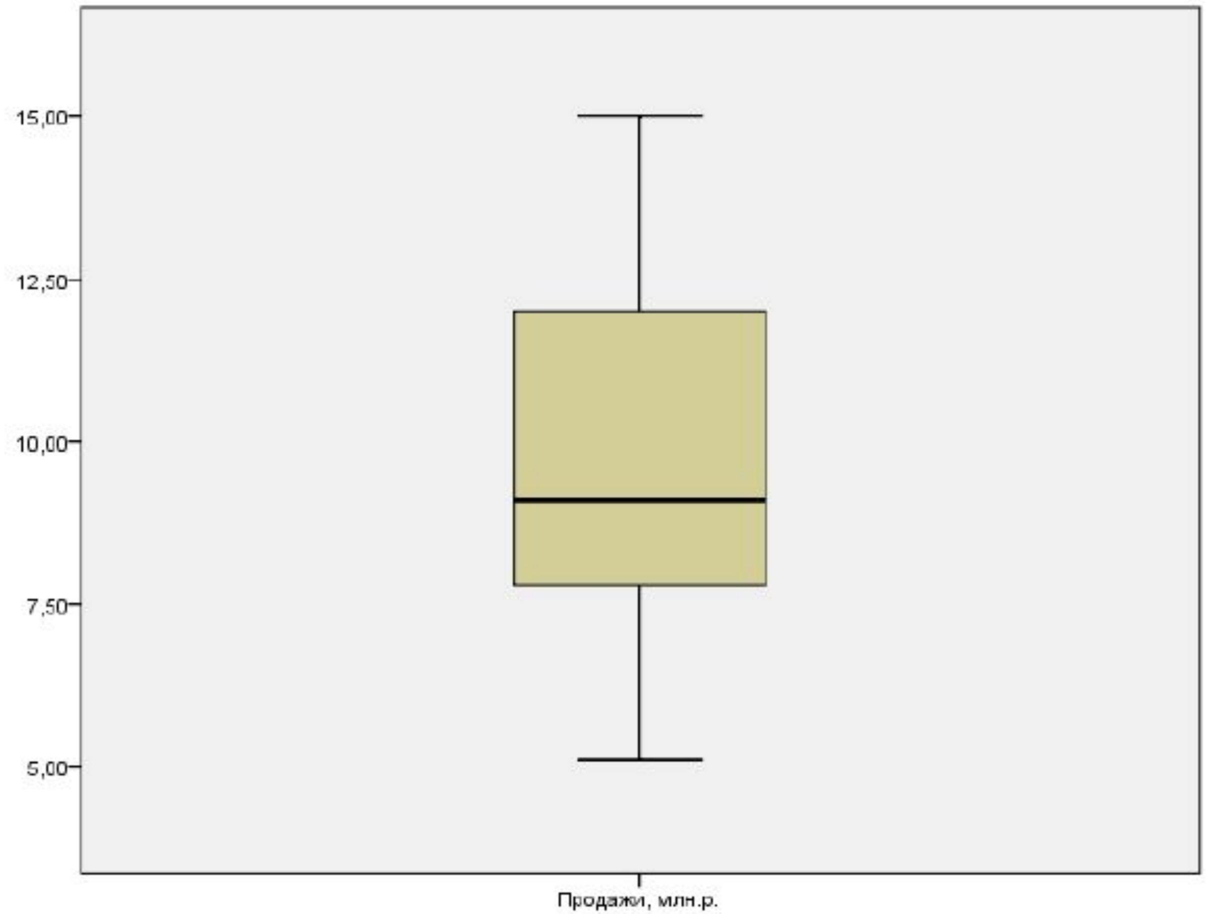
9.2.1. Графические методы
9.2.1.2. «Ящик с усами»

Построим ящик с усами по данным Задачи 1:

- Максимальное значение **15**
- Верхняя квартиль **11.75**
- Медиана **8.9**
- Нижняя квартиль **7.55**
- Минимальное значение **5.1**

9.2.1. Графические методы
9.2.1.2. «Ящик с усами»

Ящик будет иметь вид



Ящик с усами по продажам

Условия задачи 1.

9,4	8,0	6,3	10,0	15,0	8,2	7,3	9,2	5,8	8,7
5,2	13,2	8,1	7,5	11,8	14,6	8,5	7,8	10,5	6,0
5,1	6,8	8,3	7,7	7,9	9,0	10,1	8,0	12,0	14,0
8,2	9,8	13,5	12,4	5,5	7,9	9,2	10,8	12,1	12,4
12,9	12,6	6,7	9,7	8,3	10,8	15,0	7,0	13,0	9,5

Данные об объемах продаж пиломатериалов по месяцам, млн.руб.

9.2.2. Аналитические методы

Для аналитической проверки на нормальность существует ***различные тесты***:

- критерий Хи-квадрат,
- критерий Колмогорова,
- критерий Шапиро-Уилка
- критерий Жарка-Бера и другие.

9.2.2. Аналитические методы
9.2.2.1.Критерий Жарка-Бера (Jarque-Bera)

Суть критерия :

По данным выборки оценивается ***скошенность*** (асимметрия) и ***«вытянутость»*** фактического распределения и сравнивается с *нормальным*.

9.2.2. Аналитические методы

9.2.2.1.Критерий Жарка-Бера (Jarque-Bera)

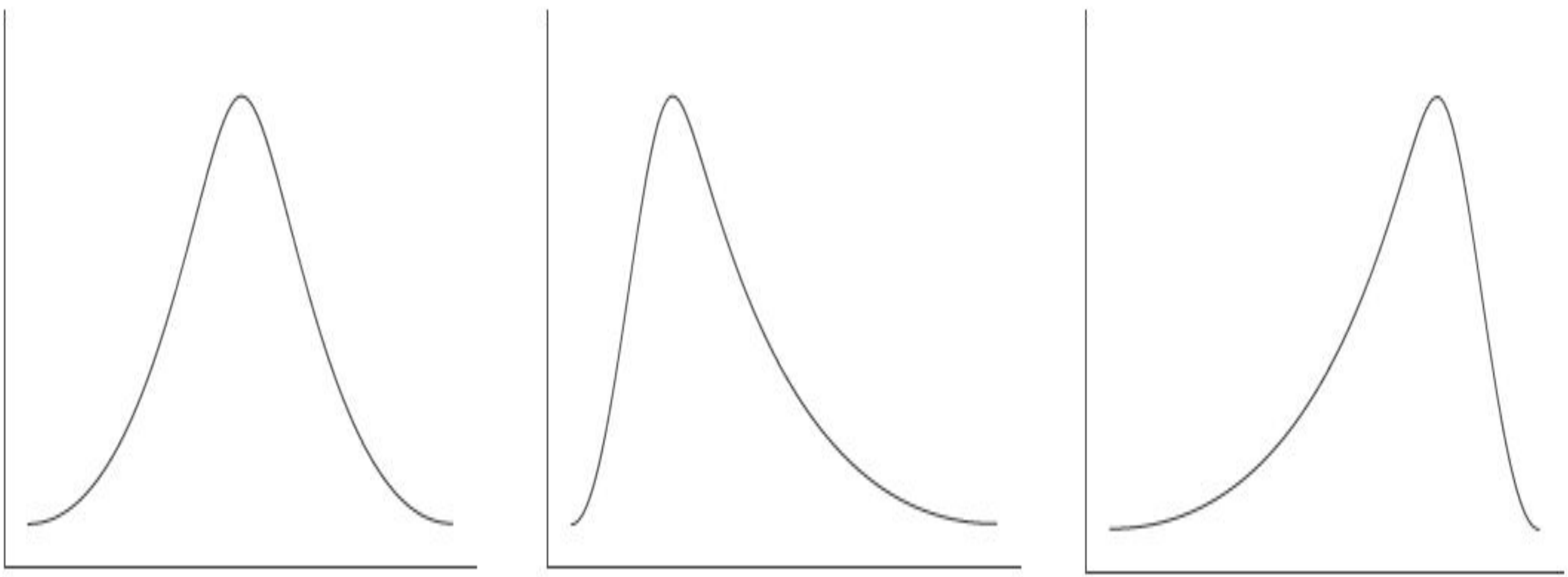
За оценку *асимметрии* распределения отвечает ***коэффициент асимметрии***

$$Sk = \frac{1}{N} \cdot \sum \left[\nu_i \cdot \left(\frac{x_i - \bar{x}}{\tilde{\sigma}} \right)^3 \right]$$

x_i – середины интервалов группировки, $\tilde{\sigma} = \sigma \cdot \sqrt{\frac{N-1}{N}}$

9.2.2.1.Критерий Жарка-Бера (Jarque-Bera) *Коэффициент асимметрии (Sk)*

Нормальное распределение | Ненормальное распределение



Симметричное, $Sk = 0$

Правосторонняя асимметрия, $Sk > 0$

Левосторонняя асимметрия, $Sk < 0$

9.2.2.1.Критерий Жарка-Бера (Jarque-Bera)

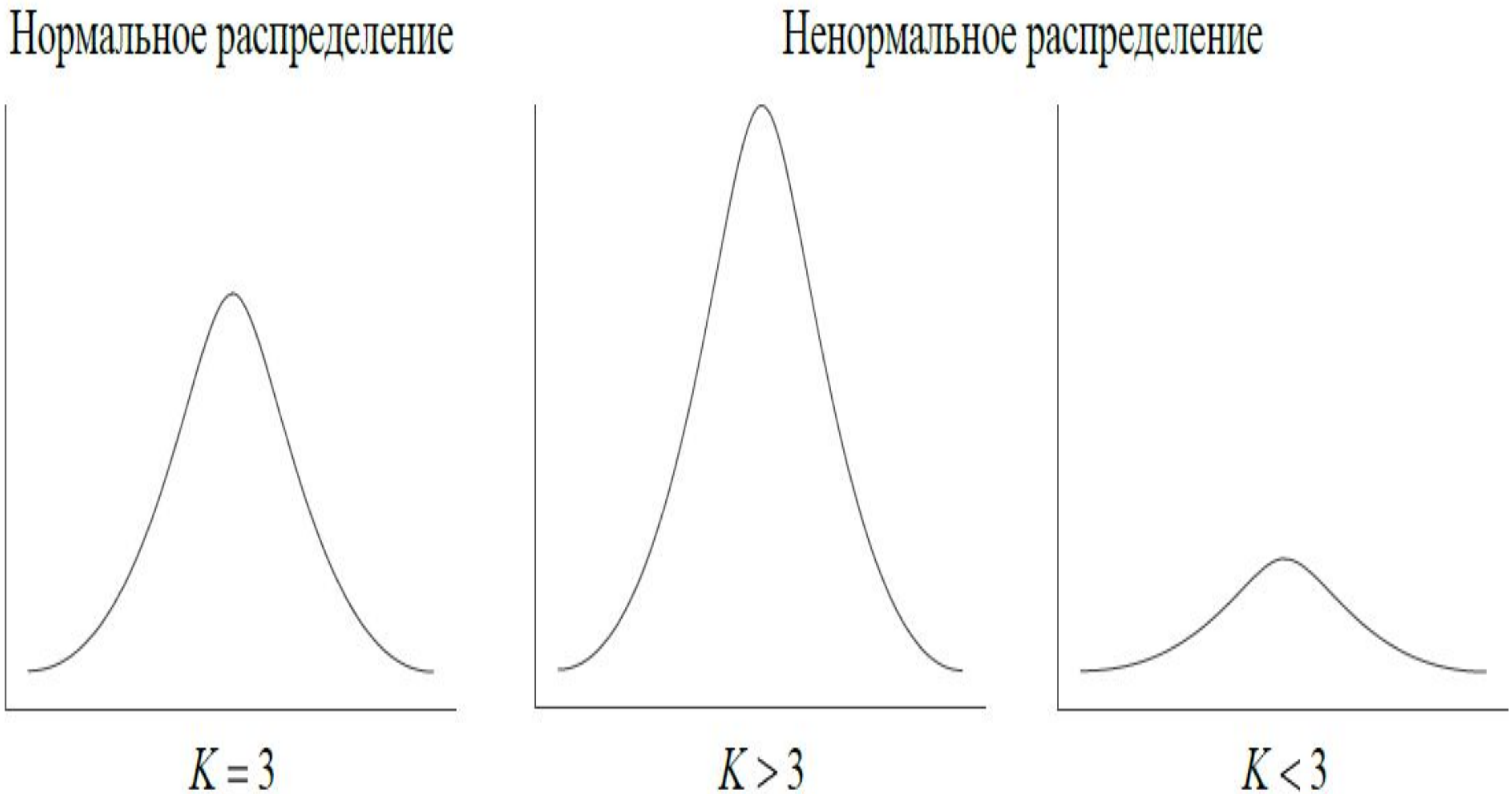
Эксцесс (К)

- За оценку «вытянутости» распределения отвечает ***эксцесс (К):***

$$K = \frac{1}{N} \cdot \sum \left[\nu_i \cdot \left(\frac{x_i - \bar{x}}{\tilde{\sigma}} \right)^4 \right]$$

x_i – середины интервалов группировки, $\tilde{\sigma} = \sigma \cdot \sqrt{\frac{N-1}{N}}$

9.2.2.1.Критерий Жарка-Бера (Jarque-Bera)
Эксцесс (K)



9.2.2.1.Критерий Жарка-Бера (Jarque-Bera)

Алгоритм критерия Жарка-Бера:

1) Выдвинуть гипотезу ***H₀*** о нормальном распределении выборки.

2) Вычислить фактическое значение критерия по формуле:

$$J-B=\frac{N}{6}\cdot\left(Sk^2+\frac{(K-3)^2}{4}\right)$$

9.2.2.1.Критерий Жарка-Бера (Jarque-Bera)

Алгоритм критерия Жарка-Бера.

3) Определить табличное значение критерия на основе специальных таблиц критических значений Пирсона на уровне значимости 0,05 и числе степеней свободы 2. Это значение равно 5,991.

Уровень значимости – это вероятность ошибиться, утверждая, что распределение ненормальное. Общепринятым является вероятность ошибки не превышающая 5%.

Число степеней свободы, в данном случае, отвечает за количество параметров в формуле критерия: там участвуют асимметрия и эксцесс.

9.2.2.1.Критерий Жарка-Бера (Jarque-Bera)

Алгоритм критерия Жарка-Бера.

4) Если $J-B>5{,}991$, то гипотеза H_0 о нормальном распределении выборки отклоняется, т.е. распределение не является нормальным.

Если $J-B<5{,}991$, то гипотеза H_0 о нормальном распределении выборки принимается, т.е. распределение является нормальным.

Произведем вычисления по действиям по данным предыдущего примера:

1. *Составим расчетную таблицу для вычисления асимметрии и эксцесса:*

№	Середина интервала x_i	Частота ν_i	$x_i - \bar{x}$	$\dfrac{x_i - \bar{x}}{\tilde{\sigma}}$	$\nu_i \cdot \left(\dfrac{x_i - \bar{x}}{\tilde{\sigma}}\right)^3$	$\nu_i \cdot \left(\dfrac{x_i - \bar{x}}{\tilde{\sigma}}\right)^4$
1						
2						
3						
4						
5						
6						
7						
Сумма:	—	50	—	—		

Произведем вычисления по действиям:

2. По данным расчетной таблицы произведем расчет *асимметрии и эксцесса:*

$$Sk = \frac{1}{N} \cdot \sum \left[\nu_i \cdot \left(\frac{x_i - \bar{x}}{\tilde{\sigma}} \right)^3 \right]$$

$$K = \frac{1}{N} \cdot \sum \left[\nu_i \cdot \left(\frac{x_i - \bar{x}}{\tilde{\sigma}} \right)^4 \right]$$

Произведем вычисления по действиям:

3. Выдвинем гипотезу ***Но*** о нормальном распределении выборки.
4. Вычислим фактическое значение критерия Жарка-Бера:

$$J-B=\frac{N}{6}\cdot\left(Sk^2+\frac{(K-3)^2}{4}\right)$$

Произведем вычисления по действиям:

5. Произведем проверку гипотезы и сделаем выводы.

$J-B < 5{,}991$, значит гипотеза H_0 о нормальном распределении выборки принимается,

т.е. распределение является нормальным.

9.3. СХЕМА ВЫБОРА АДЕКВАТНЫХ ОПИСАТЕЛЬНЫХ СТАТИСТИК

9.4.АНАЛИЗ ВЛИЯНИЯ ФАКТОРОВ

•

9.4.АНАЛИЗ ВЛИЯНИЯ ФАКТОРОВ

9.4.1. *Выбор метода анализа влияния факторов*

Фактор	Отклик	Метод
Количественный (нормальный)	Количественный (нормальный)	**Корреляция Пирсона**
Количественный (ненормальный)	Количественный (ненормальный)	**Корреляция Спирмена**
Качественный (2 значения)	Количественный (нормальный)	**Т-критерий Стьюдента**
Качественный (2 значения)	Количественный (ненормальный)	**U-критерий Манна-Уитни**
Качественный (3 и более значений)	Количественный (нормальный)	**Дисперсионный анализ**
Качественный (3 и более значений)	Количественный (ненормальный)	**Н-критерий Краскела-Уоллиса**
Качественный	Качественный	**Критерий Хи-квадрат**

9.4.2. КОЭФФИЦИЕНТЫ КОРРЕЛЯЦИИ

Рассмотрим влияние **количественного фактора на количественный отклик.**

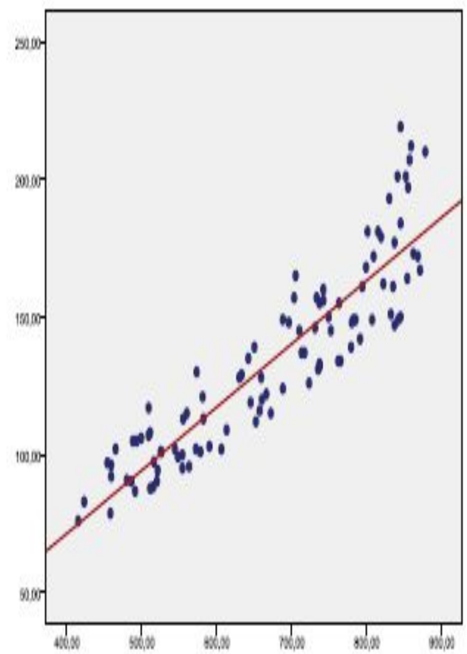
Корреляция – мера линейной связи между двумя количественными признаками.

Диаграмму рассеяния – график, на котором по горизонтальной оси отмечаются значения фактора (x), а по вертикальной – отклика (y). Расположение точек говорит и о силе связи, и о ее характере.

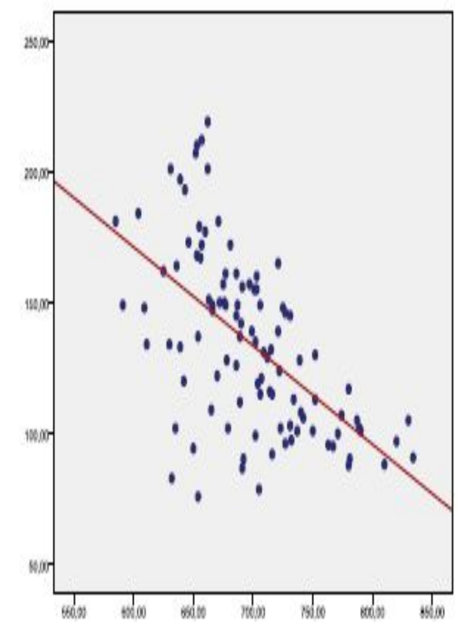
9.4.2. КОЭФФИЦИЕНТЫ КОРРЕЛЯЦИИ

Диаграмма рассеяния

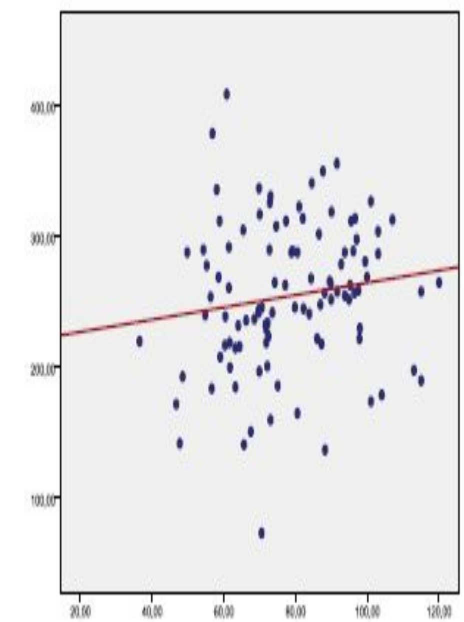
Сильная положительная корреляция



Умеренная отрицательная корреляция



Отсутствие корреляции



9.4.2. КОЭФФИЦИЕНТЫ КОРРЕЛЯЦИИ

***Свойства коэффициента корреляции*.**

1) изменяется в пределах от **-1 до 1**;
2) если коэффициент корреляции менее **0,3**, то значимая статистическая связь *отсутствует*;
3) если коэффициент корреляции равен **1 или -1**, то связь не корреляционная, а *функциональная (полная*);
4) если коэффициент больше **0**, то связь называется *прямой,* если меньше **0** – то *обратной*;

9.4.2. КОЭФФИЦИЕНТЫ КОРРЕЛЯЦИИ

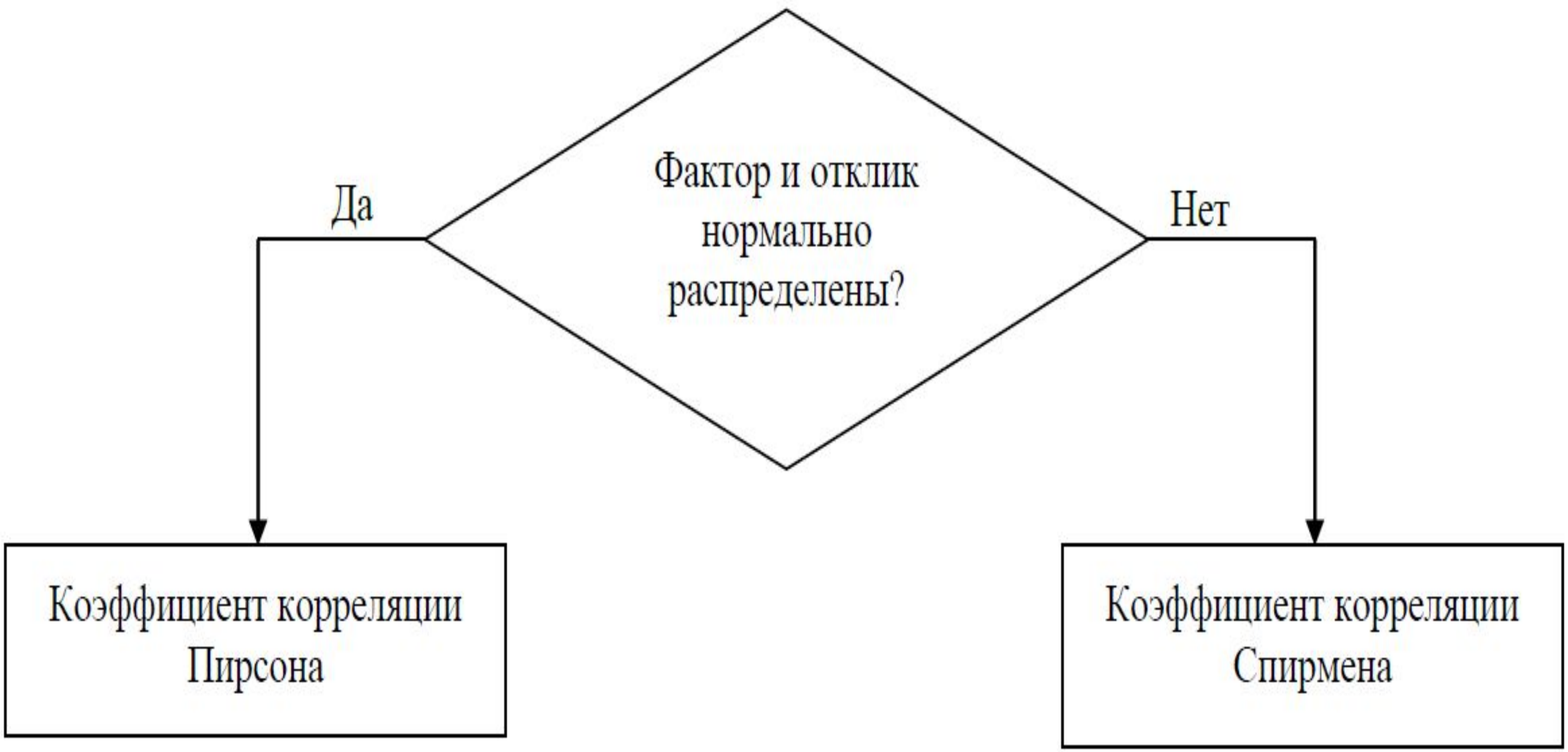
Свойства коэффициента корреляции.

5) сила связи может быть охарактеризована, например, следующим образом:

Теснота связи	Абсолютное значение коэффициента корреляции
отсутствует	Менее 0,3
слабая	0,3 – 0,5
умеренная	0,5 – 0,7
высокая	0,7 и более

9.4.2. КОЭФФИЦИЕНТЫ КОРРЕЛЯЦИИ

В зависимости от *нормальности распределения* фактора и отклика используют ***следующие коэффициенты корреляции:***



9.4.2.1. Коэффициент корреляции Пирсона

Особенности использования

1) число наблюдений фактора и отклика должно быть равно (обозначим N);
2) фактор и отклик должны иметь распределение, близкое к нормальному (или извлечены из нормально распределенных выборок бо́льшего объема);

9.4.2.1. Коэффициент корреляции Пирсона. Особенности использования.

3) формула для расчета (пусть x – фактор, а y – отклик):

$$r_{xy} = \frac{\overline{xy} - \overline{x} \cdot \overline{y}}{\sigma_x \cdot \sigma_y}, \text{ где}$$

$\overline{x}$ – среднее по фактору;

$\overline{y}$ – среднее по отклику;

$\overline{xy}$ – среднее произведений фактора на отклик;

9.4.2.1. Коэффициент корреляции Пирсона. Особенности использования.
Для выборочной совокупности

σ_x – стандартное отклонение для фактора: $\sigma_x = \sqrt{\dfrac{\sum (x_i - \tilde{x})^2}{n-1}}$;

σ_y – стандартное отклонение для фактора: $\sigma_y = \sqrt{\dfrac{\sum (y_i - \tilde{y})^2}{n-1}}$;

Рассмотрим на предыдущем примере.

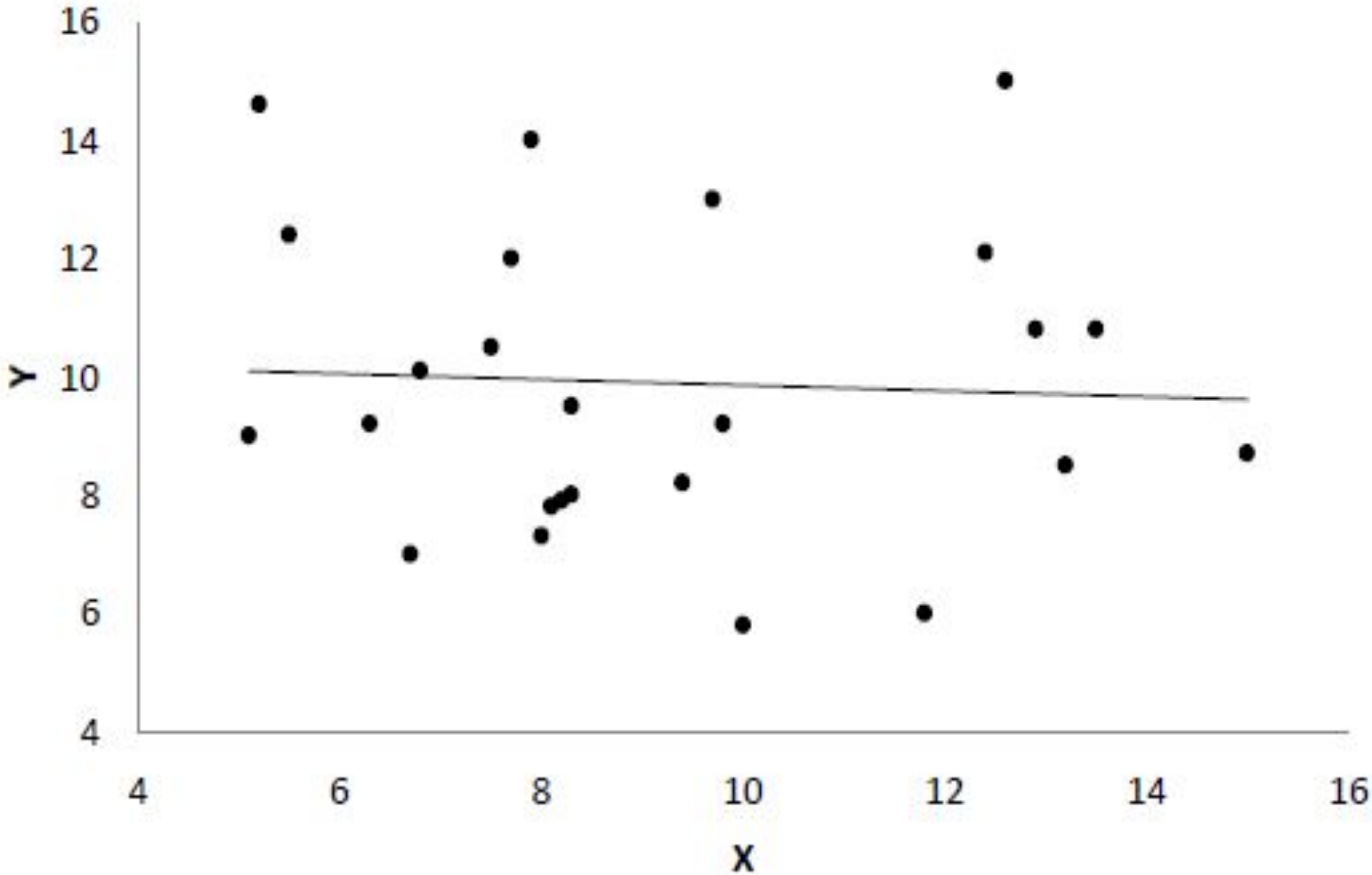
9,4	8,0	6,3	10,0	15,0	8,2	7,3	9,2	5,8	8,7
5,2	13,2	8,1	7,5	11,8	14,6	8,5	7,8	10,5	6,0
5,1	6,8	8,3	7,7	7,9	9,0	10,1	8,0	12,0	14,0
8,2	9,8	13,5	12,4	5,5	7,9	9,2	10,8	12,1	12,4
12,9	12,6	6,7	9,7	8,3	10,8	15,0	7,0	13,0	9,5

Данные об объемах продаж пиломатериалов по месяцам, млн.руб.

Разделим исходные данные на две части по 25 наблюдений и выясним, влияет ли фактор (*x) на отклик (y)*

№	x	y
1	9,4	8,2
2	5,2	14,6
3	5,1	9
4	8,2	7,9
5	12,9	10,8
6	8	7,3
7	13,2	8,5
8	6,8	10,1
9	9,8	9,2
10	12,6	15
11	6,3	9,2
12	8,1	7,8
13	8,3	8
14	13,5	10,8
15	6,7	7
16	10	5,8
17	7,5	10,5
18	7,7	12
19	12,4	12,1
20	9,7	13
21	15	8,7
22	11,8	6
23	7,9	14
24	5,5	12,4
25	8,3	9,5

Построим диаграмму рассеяния



Расчетная таблица для вычисления коэффициента корреляции Пирсона

№	x	y	xy	(x - □x)^2	(y - □y)^2
1	9.4	8.2	77.08	0.04	2.88
2	5.2	14.6	75.92	15.97	22.13
3	5.1	9	45.9	16.78	0.80
4	8.2	7.9	64.78	0.99	3.98
…	….	…..	….	….	…..
25	8.3	9.5	78.85	0.80	0.16
Сумма	**229,9**	**247,4**	**2265,92**	**190,69**	**159,49**

Решение

Средние значения:

$$\overline{x}=\frac{\sum x_i}{N}=\frac{229{,}9}{25}=9{,}20,\ \overline{y}=\frac{\sum y_i}{N}=\frac{247{,}4}{25}=9{,}90,\ \overline{xy}=\frac{\sum x_i y_i}{N}=\frac{2265{,}92}{25}=90{,}64.$$

Стандартные отклонения:

$$\sigma_x=\sqrt{\frac{\sum\left(x_i-\overline{x}\right)^2}{N-1}}=\sqrt{\frac{190{,}69}{25-1}}=2{,}82,\ \sigma_y=\sqrt{\frac{\sum\left(y_i-\overline{y}\right)^2}{N-1}}=\sqrt{\frac{159{,}49}{25-1}}=2{,}58.$$

Корреляция Пирсона:

$$r_{xy}=\frac{\overline{xy}-\overline{x}\cdot\overline{y}}{\sigma_x\cdot\sigma_y}=\frac{90{,}64-9{,}20\cdot 9{,}90}{2{,}82\cdot 2{,}58}=-0{,}06.$$

Полученный коэффициент корреляции по абсолютному значению не превышает 0,3, значит, существенной статистической связи между фактором и откликом не наблюдается.

9.4.2.2. Коэффициент корреляции Спирмена. Особенности использования.

1) фактор и отклик могут иметь ненормальное распределение;

2) число наблюдений фактора и отклика должно быть равно (обозначим N);

3) формула для расчета:

$$r_S = 1 - \frac{6 \cdot \sum_{i=1}^{N} d_i^2}{N\left(N^2 - 1\right)},$$ где d_i^2 - квадрат разности рангов фактора и отклика.

9.4.2.2. Коэффициент корреляции Спирмена. Ранжирование

- Процедура ***ранжирования*** представляет собой упорядочивание значений по возрастанию.
- Примеры ранжирования, в том числе и

Выборка 1	Ранги
5	3
3	2
2	1
9	4
20	5

Выборка 2	Ранги
5	4
3	1
4	2,5
4	2,5
9	5

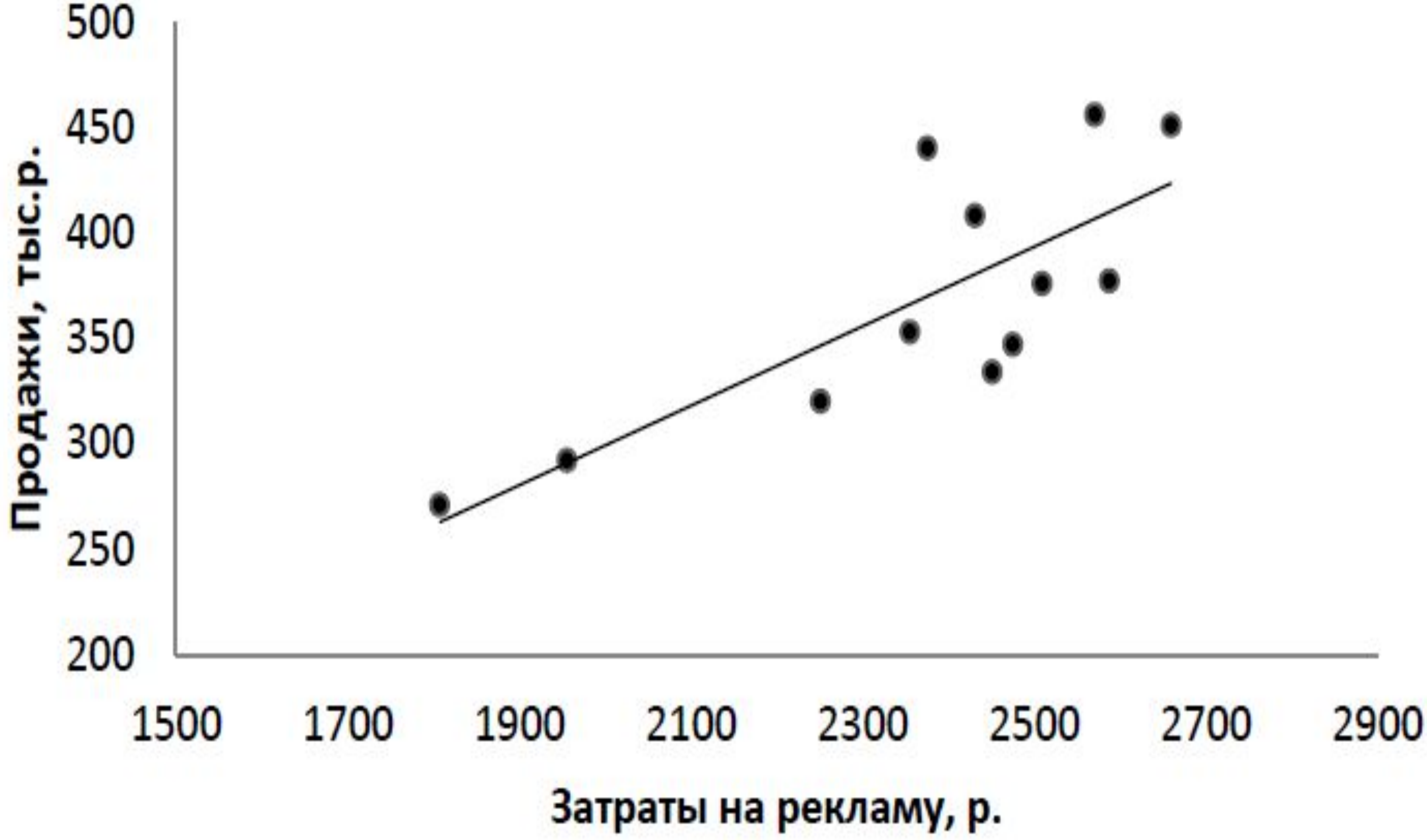
Выборка 3	Ранги
4	3
2	1
4	3
4	3
10	5

Пример расчета корреляции Спирмена

Задание: Выяснить, влияют ли затраты на рекламу на объем продаж по данным

Месяц	Затраты на рекламу, р	Продажи, тыс.р
Январь	2510	376
Февраль	2588	377
Март	2452	334
Апрель	2476	347
Май	1956	292
Июнь	2252	320
Июль	1808	271
Август	2356	353
Сентябрь	2660	451
Октябрь	2571	456
Ноябрь	2432	408
Декабрь	2376	440

Решение
1.Постоим диаграмму рассеяния.



Решение
2.Вычислим коэффициент Спирмена

Месяц	Затраты на рекламу, р, (x)	Продажи, тыс.р (y)	Ранг x	Ранг y	d	d^2
Январь	2510	376	9	7	2	4
Февраль	2588	377	11	8	3	9
Март	2452	334	7	4	3	9
Апрель	2476	347	8	5	3	9
Май	1956	292	2	2	0	0
Июнь	2252	320	3	3	0	0
Июль	1808	271	1	1	0	0
Август	2356	353	4	6	-2	4
Сентябрь	2660	451	12	11	1	1
Октябрь	2571	456	10	12	-2	4
Ноябрь	2432	408	6	9	-3	9
Декабрь	2376	440	5	10	-5	25
Сумма	—	—	—	—	—	74

Решение
2.Вычислим коэффициент Спирмена

$$r_S = 1 - \frac{6 \cdot \sum_{i=1}^{N} d_i^2}{N\,(N^2 - 1)} = 1 - \frac{6 \cdot 74}{12 \cdot (12^2 - 1)} = 0{,}741.$$

Полученный коэффициент корреляции свидетельствует о прямой высокой связи между затратами на рекламу и продажами.