

Методические материалы к лабораторному практикуму по курсу  
“Эконометрика”

*Лабораторная работа № 1*

## Статистические характеристики случайных величин

**Цель занятия** – освоение навыков работы с основными статистическими характеристиками в среде **Microsoft Excel**. В результате выполнения заданий практикума студенты должны познакомиться и освоить технологию работы с показателями описательной статистики – положения, разброса, асимметрии и показатели, описывающие закон распределения (таблицы частот, гистограммы, полигоны, кумуляты), а также давать содержательную интерпретацию полученных оценок.

### *1. Описательная статистика*

Пример выполнения задания. Выборка стоимости набора из 30 продуктов питания по отдельным городам центрального региона РФ по состоянию на 1.01. 99 г. представлена в таблице:

Города	Стоимость набора из 30 продуктов, руб.
1. Владимир	389,04
2. Вологда	417,78
3. Иваново	394,00
4. Кострома	371,96
5. Москва	525,96
6. Нижний Новгород	405,12
7. Рязань	419,52
8. Тверь	401,93
9. Ярославль	418,97

Необходимо рассчитать основные показатели описательной статистики и сделать соответствующие выводы.

### **Порядок выполнения задания.**

1. Сформировать задание на рабочем листе среды **Microsoft Excel**.
2. В меню «Сервис» выбрать раздел «Анализ данных», в котором необходимо выбрать режим «Описательная статистика». После вызова режима «Описательная статистика» в появившемся диалоговом окне необходимо задать параметры и выделить режим «Итоговая статистика», установить флажок «Метки в первой строке», выбрать уровень надежности 95%, определить выдаваемые к-ый наибольший и к-ый наименьший элементы выборки (здесь будем полагать  $k=1$ ).
3. После нажатия кнопки **ОК** на экране появляется таблица с показателями описательной статистики. Необходимо сделать анализ полученных оценок, привлекая также для анализа гистограмму (полигон), и кумуляту (интегральный процент).

Среднее	416,0311111
Стандартная ошибка	14,71123677
Медиана	405,12
Мода	#Н/Д
Стандартное отклонение	44,13371031
Дисперсия выборки	1947,784386
Эксцесс	6,063066502
Асимметричность	2,259263465
Интервал	154
Минимум	371,96
Максимум	525,96
Сумма	3744,28
Счет	9
Наибольший(1)	525,96
Наименьший(1)	371,96
Уровень надежности(95,0%)	33,92419477

Уровень надежности  $p$ , выраженный здесь в процентах, соответствует уровню доверительной вероятности  $0,95$  и уровню значимости  $\alpha = 1 - p$ .

Проведенное выборочное исследование позволяют сделать следующий вывод. С вероятностью  $0,95$  средняя стоимость набора находится в пределах от  $382,11$  руб. до  $449,95$  руб. Этот вывод следует из того, что предельная ошибка составляет  $\Delta_{\bar{x}} = 33,95$  (показатель представлен в строке соответствующей *уровню надежности 95%*) и, следовательно, доверительный интервал для генеральной средней, рассчитываемый по формуле  $[\bar{x} - \Delta_{\bar{x}}; \bar{x} + \Delta_{\bar{x}}]$ , составляет  $- [382,11; 449,95]$ . В нашем примере  $\bar{x}$  – выборочное среднее, а  $s$  – стандартная ошибка

«Надежность» средней в выборке также подтверждается незначительным уровнем коэффициента вариации  $v = \frac{s}{\bar{x}} \cdot 100\% = \frac{44,13}{416,03} \cdot 100\% = 10,6\%$  и незначительным отклонением среднего от медианы  $Me$ :  $\bar{x} - Me = 416,03 - 405,12 = 10,91$ .

Значительные положительные значения коэффициентов асимметрии и эксцесса позволяют говорить о том, что данное эмпирическое распределение существенно отличается от нормального и характеризуется скоплением членов ряда в центре распределения («островершинное» распределение) и имеет правостороннюю асимметрию.

Для построения гистограммы и кумуляты («интегральный процент») необходимо в меню «Анализ данных» выбрать режим «Гистограмма». Зададим в появившемся окне необходимые параметры: «Входной интервал», «Выходной интервал», «Интегральный процент», «Вывод графика» и нажмем кнопку **ОК**.

На экране получим рассчитанные частоты, накопленные частоты и частоты, а также изображение гистограммы и кумуляты (графа «Карман» означает выбранный, как правило, автоматически интервал накопления частот, задаваемый полуоткрытым снизу числовым интервалом, т.е.  $a < x \leq b$ ).

табл. 1

Карман	Частота	Интегральный %
371,96	1	11,11%
423,2933333	7	88,89%
474,6266667	0	88,89%
Еще	1	100,00%

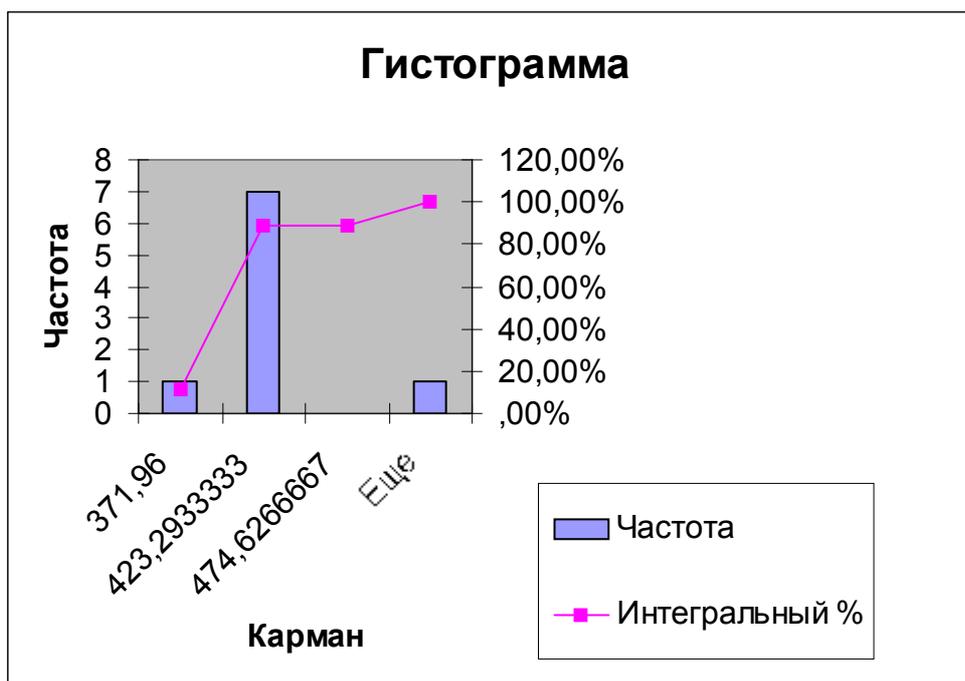


Табл. 2

<i>Частота</i>	<i>Накопленная частота</i>	<i>Частость</i>	<i>Частость, %</i>
1	1	0,0588	5,88
9	10	0,5882	58,82
5	15	0,8824	88,24
1	16	0,9412	94,12
1	17	1	100

**Примечание:**

- Порядок расчета накопленных частостей в графе «интегральный процент, %» табл.1:  
 На основании частот (см. кол. 1 табл. 2) рассчитываются накопленные частоты (см. кол. 2 табл. 2). Далее, каждое значение накопленной частоты делится на максимальное накопленное значение, в результате чего получаются частости, выраженные в долях единицы (см. кол. 3 табл. 2). После преобразования частостей в процентный формат получаем окончательный результат (см. кол. 4 табл.2).
- Как правило, гистограмму предпочтительнее изображать в виде смежных прямоугольных областей, поэтому столбики гистограммы целесообразно расширить до соприкосновения друг с другом. Для этого следует подвести стрелку на гистограмму и двойным щелчком мыши вызвать экранную форму «Формат рядов данных». В появившемся диалоговом окне необходимо активизировать вкладку «Параметры» и в поле «Ширина зазора» установить значение 0. После указанных преобразований гистограмма примет стандартный вид.

Самостоятельно выполните пункты 1-3 для уровней надежности – 99% и 90%.и проведите статистический анализ в среде **Microsoft Excel**.

## **2. Ковариация и корреляция**

При работе в среде **Microsoft Excel** в меню «Сервис» выбрать раздел «Анализ данных», в котором необходимо выбрать режимы «Ковариация» и «Корреляция». После вызова соответствующего режима, в появившемся диалоговом окне необходимо выделить (активизировать) режим «Метки в первой строке» для получения семантической наглядности получаемых результатов (при условии, что в первой строке /первом столбце/ вводимых данных присутствует заголовок).

Руководствуясь приведенной ниже шкалой Челдока определить взаимосвязь показателей в следующих задачах:

Теснота связи	Значение коэффициента корреляции при наличии:	
	прямой связи	обратной связи
Слабая	0,1...0,3	-(0,1...0,3)
Умеренная	0,3...0,5	-(0,3...0,5)
Заметная	0,5...0,7	-(0,5...0,7)
Высокая	0,7...0,9	-(0,7...0,9)
Весьма высокая	0,9...0,99	-(0,9...0,99)

1. Требуется определить зависимость между величиной уставного капитала предприятий X и количеством выставленных акций Y для предприятий города, выставивших акции на чековый аукцион, данные о предприятиях приведены в следующей таблице:

Номер предприятия	Уставной капитал, млн. руб. (X)	Число выставленных акций (Y)
1	2954	856
2	1605	930
3	4102	1563
4	2350	682
5	2625	616
6	1795	495
7	2813	815
8	1751	858
9	1700	467
10	2264	661

2. Показатели уровня образования, уровня преступности, а также отношение числа безработных к числу вакансий в некоторых центральных областях РФ в 1995 г. (по данным Госкомстата РФ) приведены в следующей таблице:

Область	Уровень образования	Отношение числа безработных к числу вакансий	Уровень преступности
Брянская	735	22,3	908
Владимирская	788	10,8	791
Ивановская	779	52,9	804
Калужская	795	2,2	701
Костромская	740	10,4	685
Г. Москва	902	0,4	496
Московская	838	2,4	536
Нижегородская	736	5,4	936
Орловская	762	4,1	662
Рязанская	757	4,1	671
Смоленская	772	1,0	920
Тверская	764	4,2	1040
Тульская	764	2,1	809
Ярославская	755	25,1	882

По представленным выборочным данным требуется установить наличие взаимосвязи между указанными показателями в центральном регионе РФ.

**Примечание:**

1. Уровень образования рассчитывается как численность лиц с высшим и средним специальным образованием на 1000 жителей области.
2. Уровень преступности рассчитывается как число совершенных преступлений на 100000 жителей области.

**3.Задания для самостоятельной работы.**

Работая в средах **Microsoft Excel** и **Statgraphics Plus for Windows** рассчитать основные показатели описательной статистики и сделать соответствующие выводы для следующих заданий:

- а) Провести анализ курса доллара по итогам торгов на российских валютных биржах 06.02. 95 г. (доверительная вероятность – 0,9 и 0,95)

Валютная биржа	Курс, руб./ долл.
1. Московская межбанковская	4133
2. Санкт-Петербургская	4165
3. Сибирская межбанковская	4126
4. Уральская межрегиональная	4130
5. Азиатско-Тихоокеанская межбанковская	4115
6. Ростовская межбанковская	4127
7.Нижегородская валютно-фондовая	4133

- б) Проанализировать педагогический стаж преподавателей вуза

(с уровнем надежности 95% и 99%)

Фамилия преподавателя	Педагогический стаж, лет
1. Орлов	15
2. Грачев	8
3. Петухов	10
4. Голубев	7
5. Курочкин	5
6. Соловьев	10
7. Синицин	5
8. Воробьев	2
9. Ласточкин	11

в) Провести анализ результатов сбора зерна по некоторым хозяйствам района (с уровнем значимости 0,05 и 0.01)

Наименование хозяйства	Валовой сбор зерна, ц.
1. Ивановка	600
2. Петровское	520
3. Ивантеевское	400
4. Семеновское	600
5. Андреевка	500
6. Мележа	380

г) Провести анализ индивидуального жилищного строительства по отдельным районам Ярославской области в 1998 г.

Районы	Площадь, м <sup>2</sup>
1. Большесельский	718
2. Борисоглебский	1319
3. Гаврилов-Ямский	919
4. Даниловский	1321
5. Любимский	437
6. Мышкинский	218
7. Некрасовский	2121
8. Переславский	8872
9. Пешехонский	3011
10. Рыбинский	1389
11. Тутаевский	730
12. Угличский	4728

