

ЧАСТНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ
ПРОФЕССИОНАЛЬНОГО ОБРАЗОВАНИЯ
«СТАВРОПОЛЬСКИЙ МНОГОПРОФИЛЬНЫЙ КОЛЛЕДЖ»

Методические указания

к практическим занятиям и практической подготовке
по дисциплине «**Статистика**»

для обучающихся по специальности
38.02.07 Банковское дело

Ставрополь, 2022

Настоящие методические указания составлены в соответствии с требованиями Федерального государственного образовательного стандарта для обучающихся по специальности 38.02.07 Банковское дело и программой дисциплины «Статистика»

Составитель: Астафьев В. А.

Рассмотрено на заседании методического объединения укрупненных групп специальностей 38.00.00 Протокол № 6 от 25.05.2022 г.

Рекомендовано к использованию в учебном процессе Методическим советом СМК, протокол № 6 от 26.05.2022 г.

ВВЕДЕНИЕ

Программа учебной дисциплины «Статистика» предназначена для реализации требований ФГОС ПО к минимуму содержания и уровню подготовки выпускников по специальности профессионального образования «Банковское дело».

Изучение дисциплины позволит иметь представление: об общих основах статистической науки, искусстве организации и проведения статистических исследований, анализа и обобщении их результатов, навыков прогнозирования.

Результатом изучения курса должно быть знание принципов современной организации статистических служб, категорий и понятий статистики, методов организации сбора, обработки материалов статистического наблюдения, их анализа с помощью обобщающих показателей. Студенты, изучившие курс, обязаны уметь организовать сплошное и не сплошное наблюдение, строить статистические графики и таблицы, анализировать статистические данные, формулировать выводы, вытекающие из проведенного анализа.

В дисциплине рассматриваются методы сбора статистической информации, вопросы методологии и практика осуществления статистической сводки и группировки построения статистических таблиц и графиков, расчета абсолютных, относительных и средних величин, их использование в анализе социально-экономических явлений, анализ рядов динамики, показателей вариации, методологию индексного анализа, вопросы теории и практики выборочного наблюдения, оценки взаимосвязи признаков.

Цели и задачи учебной дисциплины – требования к результатам освоения дисциплины:

В результате освоения учебной дисциплины обучающийся должен знать:

- предмет, метод и задачи статистики;
- общие основы статистической науки;
- принципы организации государственной статистики;
- современные тенденции развития статистического учета;
- основные способы сбора, обработки, анализа и наглядного представления информации;
- основные формы и виды действующей статистической отчетности;
- технику расчета статистических показателей, характеризующих социально-экономические явления.

Уметь:

- собирать и регистрировать статистическую информацию;
 - проводить первичную обработку и контроль материалов наблюдения;
- выполнять расчеты статистических показателей и формулировать основные выводы;

- осуществлять комплексный анализ изучаемых социально-экономических явлений и процессов, в том числе с использованием средств вычислительной техники.

СОДЕРЖАНИЕ

Практическое занятие 1. Предмет, метод и задачи статистики	6
Практическое занятие 2. Статистическое наблюдение	7
Практическая подготовка 1. Сводка и группировка статистических данных	9
Практическая подготовка 2. Ряды распределения и их графическое изображение	14
Практическая подготовка 3. Абсолютные и относительные величины в статистике	21
Практическая подготовка 4. Виды средних величин и их исчисление в статистике	29
Практическая подготовка 5. Показатели вариации	37
Практическая подготовка 6. Ряды динамики. Показатели анализа динамических рядов....	40
Практическая подготовка 7. Способы обработки динамических рядов	45
Практическая подготовка 8. Статистические индексы	47
Практическая подготовка 9. Статистическое изучение взаимосвязей	52
Практическое занятие 3. Выборочное наблюдение. Понятие о выборочном наблюдении .	59

Практическое занятие 1. Предмет, метод и задачи статистики

Теоретическая часть

Термин «статистика» был введен в науку немецким ученым Готфридом Ахенвалем в 1746 г. и первоначально он означал государственное управление, т.е. под статистикой понимали совокупность сведений о государстве. Сам же термин «статистика» происходит от латинского слова «status», которое означает состояние (положение) вещей, а в средние века означало политическое состояние государства.

Многовековое развитие статистической науки привело к изменению понятия «статистика». В наше время под статистикой понимают отрасль науки, которая изучает количественные закономерности массовых общественных явлений и процессов в неразрывной связи с их качественной стороной, т.е. занимается получением, обработкой и анализом информации, характеризующей количественные закономерности развития общества.

Отрасль статистической науки, которая изучает количественные показатели деятельности правоохранительных органов и органов юстиции (милиции, судов, прокуратуры, арбитража, нотариата, адвокатуры, исправительных заведений и др.), т.е. тех учреждений, где решаются правовые и юридически значимые вопросы, называется правовой статистикой.

Предметом правовой статистики служит количественная сторона качественно однородных массовых правовых и иных юридически значимых явлений и процессов, тенденции и закономерности их развития в конкретных условиях места и времени.

Объектом изучения правовой статистики служит деятельность всех государственных органов (прокуратуры, милиции, судов, исправительно-трудовых учреждений, арбитражного суда, нотариата и др.), осуществляющих уголовно-правовую, административную и гражданско-правовую охрану общественного и государственного строя, всех форм собственности, прав и интересов граждан и организаций.

Задания к практическому занятию

Подготовить сообщения по вопросам:

1. Что такое статистика?
2. Какие составные элементы статистической науки Вы знаете?
3. Что является предметом общей теории статистики?
4. Объект изучения экономической статистики.
5. Какие задачи решает статистика?
6. Перечислите методологические задачи статистики.

Практическое занятие 2. Статистическое наблюдение

Теоретическая часть

Всякая новая статистическая работа начинается со статистического наблюдения, представляющего собой планомерное, научно-организованное собирание данных явлениях и процессах общественной жизни. Тема имеет своей целью дать представление о значении статистического наблюдения как этапа статистического наблюдения как этапа статистического исследования и источника всех статистических сведений, о принципах и правилах его научной организации и проведения. Необходимо уяснить, что от качества проведения статистического наблюдения зависит качество результатов всех последующих этапов статистического исследования.

Чтобы провести статистическое наблюдение, необходимо разработать его план, который состоит из программно-методологических и организационных вопросов. Основное содержание программно-методологического раздела плана составляют вопросы об объекте и единице наблюдения, о программе наблюдения и принципах её разработки, о статистическом формуляре и инструкции. Важнейшими вопросами организационного раздела плана являются следующие: о месте, времени и сроках наблюдения, о подборе и инструктировании кадров, о мерах по обеспечению точности наблюдения.

Уяснив основные понятия статистического наблюдения, следует переходить к изучению вопроса об организационных формах статистического наблюдения, о его видах и способах получения сведений. При этом необходимо разобраться в особенностях классификации статистического наблюдения по различным признакам (например, по степени полноты охвата наблюдением изучаемого объекта, по времени регистрации или по периодичности проведения статистических обследований, по источнику сведений, на основе которого заполняется формуляр наблюдения).

Особое внимание необходимо уделить вопросам точности наблюдения, выяснить, какие ошибки могут возникать в процессе статистического наблюдения и что является их источником, каковы способы регистрации данных.

Задания для выполнения на практическом занятии

1. Составьте перечень важнейших признаков, характеризующих как единицу статистического наблюдения: а) промышленное предприятие; б) кредитную организацию (банк); в) школу; г) библиотеку; д) театр; ж) студента.
2. Назовите 10 признаков, характеризующих коммерческий банк как единицу статистического наблюдения.
3. Какие признаки нужно выделить для статистического обследования:
а) культурного уровня населения; б) деятельности банковского сектора; в) студентов вузов с целью изучения бюджета времени.

4. Сформулируйте вопросы для регистрации при статистическом наблюдении следующих признаков по некоторым объектам наблюдения:
- а) занятие человека; б) его семейное положение; в) успеваемость студента.
5. Спроектируйте статистическое обследование (сформулируйте цель наблюдения, формуляр и инструкцию) :
- а) переписи населения; б) переписи банков; в) переписи школ; г) переписи жилищного фонда.

Практическая подготовка 1. Сводка и группировка статистических данных

Алгоритм выполнения работы

Группировкой называется расчленение единиц изучаемой совокупности на однородные группы по определенным, существенным для них признакам. Группировочным признаком называется признак, по которому проводится разбиение единиц совокупности на отдельные группы.

В задании 1 основанием группировки служит величина уставного капитала.

Величина равного интервала определяется по формуле:

$$h = \frac{X_{\max} - X_{\min}}{n},$$

где h – величина интервала;

X_{\max} – максимальное значение признака в совокупности;

X_{\min} – минимальное значение признака в совокупности;

n – число групп.

Результаты необходимо оформить в таблице 1:

Таблица 1 - Распределение предприятий по величине уставного капитала

Величина уставного капитала, тыс. руб.	Количество предприятий	Средняя величина уставного капитала, тыс. руб.

Для характеристики структуры полученного ряда распределения необходимо рассчитать моду и медиану.

Мода - значение признака, наиболее часто встречающееся в совокупности. Для интервальных вариационных рядов мода определяется по формуле:

$$M_o = X_{M_o} + h \frac{f_{M_o} - f_{M_o-1}}{(f_{M_o} - f_{M_o-1}) + (f_{M_o} - f_{M_o+1})},$$

где M_o – мода;

X_{M_o} – нижняя граница значения интервала, содержащего моду;

h – величина модального интервала;

f_{M_o} – частота модального интервала;

f_{M_o-1} – частота интервала, предшествующего модальному;

f_{M_o+1} – частота интервала, следующего за модальным.

Медиана – значение признака, приходящееся на середину упорядоченной совокупности. Для ее определения необходимо подсчитать сумму накопленных частот ряда. Наравивание продолжается до получения накопленной суммы частот, впервые превышающей половину. Медиана интервального ряда распределения определяется по формуле:

$$Me = x_{Me} + h \frac{\frac{1}{2} \sum f - S_{Me-1}}{f_{Me}},$$

где Me – медиана;

X_{me} – нижняя граница значения интервала, содержащего медиану;

h – величина медианного интервала;

$\sum f$ – сумма частот;

S_{me-1} – сумма накопленных частот, предшествующих медианному интервалу;

f_{me} – частота медианного интервала.

Пример 1. Для изучения распределения предприятий по величине уставного капитала (тыс. рублей) необходимо:

1) построить интервальный вариационный ряд, выделив 4 группы с равными интервалами, результаты оформить в таблице;

2) рассчитать среднюю величину уставного капитала на одно предприятие в целом по совокупности и отдельно по каждой группе;

3) определить моду, медиану, сделать выводы.

Исходные данные: величина уставного капитала – 96, 79, 138, 76, 105, 132, 97, 67, 126, 65, 89, 84, 90, 67, 150, 93, 124, 102, 106, 85, 78, 103, 115, 76, 79, 96, 79, 74, 70, 80.

1. Определим шаг интервала: $h = \frac{X_{\max} - X_{\min}}{n}$,

где $X_{\min} = 65$, $X_{\max} = 150$, $n = 4$, значит $h = 21,25$.

2. Построим группировку, оформив ее в таблицу 2:

Таблица 2 - Распределение предприятий по величине уставного капитала

Величина уставного капитала, тыс. руб.	Количество предприятий	Средняя величина уставного капитала, тыс. руб.
65 – 86,25	14	75,6
86,25 – 107,5	10	97,7
107,5 – 128,75	3	121,67
128,75 – 150	3	140
итого	30	94

3. Рассчитаем среднюю величину уставного капитала на одно предприятие в целом по совокупности и отдельно по каждой группе:

$$X_{cp} = \frac{\sum x}{n}$$

$X_{cp} = 2821 \setminus 30 = 94$ тыс. руб – средняя величина уставного капитала в целом по совокупности;

$X_1 = (79+76 + 67 + 65 + 84 + 67 + 80 + 85 + 78 + 76 + 79 + 79 + 74 +70) / 14 = 75,6$ тыс. руб – средняя величина уставного капитала по первой группе;

Аналогично рассчитываются средние величины по 2, 3 и 4 группам:

$X_2 = 97,7$ тыс. руб – средняя величина уставного капитала по второй группе;

$X_3 = 121,67$ тыс. руб – средняя величина уставного капитала по третьей группе;

$X_4 = 140$ тыс. руб – средняя величина уставного капитала по четвертой группе.

4. Определим моду и медиану.

Мода вычисляется по формуле:

$$Mo = X_{Mo} + h \cdot \frac{f_{Mo} - f_{Mo-1}}{(f_{Mo} - f_{Mo-1}) + (f_{Mo} - f_{Mo+1})}.$$

Мода находится в интервале от 65 до 86,25 тыс. руб

Так как $X_{Mo} = 65$, $h = 21,25$, $f_{Mo} = 14$, $f_{Mo-1} = 0$, $f_{Mo+1} = 10$, то $Mo = 81,5$.

Следовательно, наиболее часто встречающаяся величина уставного капитала 81,5 тыс. рублей.

Чтобы определить в каком интервале лежит медиана необходимо построить ряд накопленных частот, для этого заполним таблицу 3:

Таблица 3 - Распределение предприятий по величине уставного капитала

Величина уставного капитала, тыс. руб.	Количество предприятий	Средняя величина уставного капитала	Накопленная частота
65 – 86,25	14	75,6	14
86,25 - 107,5	10	97,7	24
107,5 – 128,75	3	121,67	27
128,75 – 150	3	140	30
Итого:	30	94	

По ряду накопленных частот определяем, что медиана находится в интервале от 86,25 до 107,5 тыс. руб

Подставив значения в формулу:

$$Me = X_{Me} + h \frac{\frac{1}{2} \sum f - S_{me-1}}{f_{Me}},$$

где $X_{Me} = 86,25$, $h = 21,25$, $S_{Me-1} = 14$, $f_{Me} = 10$, получим $Me = 88,4$.

Таким образом, предприятие, имеющее величину уставного капитала 88,4 тыс. рублей делит имеющуюся совокупность на две равные части.

ЗАДАНИЕ 1

По данным о величине уставного капитала 25 предприятий (тыс. рублей) произвести интервальную группировку, выделив 4 группы с равными

интервалами, определить среднюю величину уставного капитала в целом по совокупности и отдельно по каждой группе, результаты оформить в таблице, рассчитать моду и медиану. Сделать краткие выводы.

Исходные данные

Номер предприятия	Номер варианта									
	1	2	3	4	5	6	7	8	9	10
1	85	90	95	104	96	117	98	102	190	97
2	90	86	80	87	79	80	63	192	94	74
3	85	99	81	88	138	84	95	183	108	86
4	104	95	65	65	76	87	66	176	87	67
5	65	68	69	67	105	60	58	184	79	69
6	87	101	122	94	132	86	87	187	118	85
7	98	113	117	83	97	99	75	198	87	94
8	70	48	68	54	67	53	84	183	68	43
9	85	64	92	79	126	84	73	187	79	105
10	87	88	78	98	65	87	66	161	65	89
11	105	85	116	78	89	110	76	190	98	67
12	52	63	102	74	84	69	60	173	53	52
13	73	51	126	42	90	50	56	169	42	65
14	116	100	73	96	67	96	68	200	53	85
15	63	73	81	77	150	99	99	194	61	47
16	68	49	58	57	93	63	68	187	68	64
17	64	67	86	58	124	59	103	174	60	53
18	118	98	135	98	102	135	54	174	109	104
19	52	63	76	49	106	63	108	154	45	56
20	95	103	84	98	85	96	109	195	94	93
21	46	45	73	83	78	45	98	191	28	54
22	70	72	94	85	103	99	82	166	57	63
23	85	80	75	72	115	58	67	164	65	73
24	74	83	78	65	76	73	73	180	37	57
25	65	84	63	66	79	107	86	170	87	74

Номер предприятия	Номер варианта									
	11	12	13	14	15	16	17	18	19	20
1	61	109	48	66	67	75	75	66	54	47
2	91	133	80	83	101	46	47	101	40	82
3	75	123	64	75	108	74	67	108	74	89
4	87	132	75	96	103	89	82	93	95	74
5	74	95	63	35	70	57	79	103	94	84
6	46	147	35	66	114	22	77	70	62	51
7	102	120	91	94	89	57	78	114	93	95
8	77	130	66	57	73	87	38	89	84	70
9	88	119	77	75	94	71	106	138	93	119
10	63	154	62	85	122	73	77	80	34	61
11	65	116	54	50	135	55	38	93	81	84
12	66	141	55	68	105	87	47	96	62	104
13	107	145	96	115	104	57	89	123	102	125
14	104	74	93	77	104	75	104	139	53	86
15	99	114	88	99	100	68	58	105	76	85
16	65	124	54	65	90	28	89	104	72	84
17	87	137	76	83	120	37	39	104	55	81
18	86	103	75	75	77	57	67	100	71	74
19	53	117	44	93	159	84	59	93	61	49
20	59	144	48	76	80	65	30	68	52	104
21	98	100	87	97	109	60	96	120	91	64
22	49	150	38	93	90	84	67	83	54	139
23	113	119	92	62	120	77	111	159	119	68
24	67	110	56	82	71	50	62	86	51	50
25	71	113	60	77	75	76	70	69	62	74

Практическая подготовка 2. Ряды распределения и их графическое изображение

Алгоритм выполнения работы

Ряд распределения является одним из видов группировок.

Ряд распределения — представляет собой упорядоченное распределение единиц изучаемой совокупности на группы по определенному варьирующему признаку.

В зависимости от признака, положенного в основу образования ряда распределения различают атрибутивные и вариационные ряды распределения:

Атрибутивными — называют ряды распределения, построенные по качественными признакам.

Ряды распределения, построенные в порядке возрастания или убывания значений количественного признака называются вариационными.

В первом столбце приводятся количественные значения варьирующегося признака, которые называются вариантами и обозначаются X_i . Дискретная варианта — выражается целым числом. Интервальная варианта находится в пределах от и до. В зависимости от типа варианты можно построить дискретный или интервальный вариационный ряд.

Во втором столбце содержится количество конкретных вариантов, выраженное через частоты или частоты:

Частоты — это абсолютные числа, показывающие столько раз в совокупности встречается данное значение признака, которые обозначают f_i . Сумма всех частот равна должна быть равна численности единиц всей совокупности.

Частоты (W_i) — это частоты, выраженные в процентах к итогу. Сумма всех частотей, выраженных в процентах, должна быть равна 100%, в долях—единице.

Графическое изображение рядов распределения

Наглядно ряды распределения представляются при помощи графических изображений.

Ряды распределения изображаются в виде:

Полигона

Гистограммы

Кумуляты

Огивы

Полигон

При построении полигона на горизонтальной оси (ось абсцисс) откладывают значения варьирующего признака, а на вертикальной оси (ось ординат) — частоты или частоты.

Полигон на рис. 1 построен по данным микропереписи населения России.

Домохозяйства,	одного	двух	трех	5	или	всего
----------------	--------	------	------	---	-----	-------

состоящие из:	человека	человек	человек	более	
Число домохозяйств в %	19,2	26,2	22,6	20,5	100,0



Рис. 1. Распределение домохозяйств по размеру

Условие: Приводятся данные о распределении 25 работников одного из предприятий по тарифным разрядам: 4; 2; 4; 6; 5; 6; 4; 1; 3; 1; 2; 5; 2; 6; 3; 1; 2; 3; 4; 5; 4; 6; 2; 3; 4

Задача: Построить дискретный вариационный ряд и изобразить его графически в виде полигона распределения.

Решение:

В данном примере вариантами является тарифный разряд работника. Для определения частот необходимо рассчитать число работников, имеющих соответствующий тарифный разряд.

Тарифный разряд X_i	Число работников f_i
1	3
2	5
3	4
4	6
5	3
6	4
Итого:	25

Полигон используется для дискретных вариационных рядов.

Для построения полигона распределения (рис 1) по оси абсцисс (X) откладываем количественные значения варьирующего признака — варианты, а по оси ординат — частоты или частоты.



Если значения признака выражены в виде интервалов, то такой ряд называется интервальным.

Интервальные ряды распределения изображают графически в виде гистограммы, кумуляты или огивы.

Статистическая таблица

Условие: Приведены данные о размерах вкладов 20 физических лиц в одном банке (тыс.руб) 60; 25; 12; 10; 68; 35; 2; 17; 51; 9; 3; 130; 24; 85; 100; 152; 6; 18; 7; 42.

Задача: Построить интервальный вариационный ряд с равными интервалами.

Решение:

Исходная совокупность состоит из 20 единиц ($N = 20$).

По формуле Стерджесса определим необходимое количество используемых групп:
 $n = 1 + 3,322 \cdot \lg 20 = 5$

Вычислим величину равного интервала: $i = (152 - 2) / 5 = 30$ тыс.руб

Расчленим исходную совокупность на 5 групп с величиной интервала в 30 тыс.руб.

Результаты группировки представим в таблице:

Размер вкладов тыс.руб X_i	Число вкладов f_i	Число вкладов в % к итогу W_i
2 — 32	11	55
32 — 62	4	20
62 — 92	2	10
92 — 122	1	5
122 — 152	2	10
Итого:	20	100

При такой записи непрерывного признака, когда одна и та же величина встречается дважды (как верхняя граница одного интервала и нижняя граница другого интервала), то эта величина относится к той группе, где эта величина выступает в роли верхней границы.

Гистограмма

Для построения гистограммы по оси абсцисс указывают значения границ интервалов и на их основании строят прямоугольники, высота которых пропорциональна частотам (или частостям).

На рис. 2. изображена гистограмма распределения населения России в 1997 г. по возрастным группам.

Все население	В том числе в возрасте								
	до 10	10-	20-	30-	40-	50-	60-	70 и	Всего

		20	30	40	50	60	70	старше	
Численность населения	12,1	15,7	13,6	16,1	15,3	10,1	9,8	7,3	100,0



Рис. 2. Распределение населения России по возрастным группам
Условие: Приводится распределение 30 работников фирмы по размеру месячной заработной платы

Размер заработной платы руб. в месяц	Численность работников чел.
до 5000	4
5000 — 7000	12
7000 — 10000	8
10000 — 15000	6
Итого:	30

Задача: Изобразить интервальный вариационный ряд графически в виде гистограммы и кумуляты.

Решение:

Неизвестная граница открытого (первого) интервала определяется по величине второго интервала: $7000 - 5000 = 2000$ руб. С той же величиной находим нижнюю границу первого интервала: $5000 - 2000 = 3000$ руб.

Для построения гистограммы в прямоугольной системе координат по оси абсцисс откладываем отрезки, величины которых соответствуют интервалам вариационного ряда. Эти отрезки служат нижним основанием, а соответствующая частота (частость) — высотой образуемых прямоугольников.

Построим гистограмму:



Для построения кумуляты необходимо рассчитать накопленные частоты (частоты). Они определяются путем последовательного суммирования частот (частостей) предшествующих интервалов и обозначаются S . Накопленные частоты показывают, сколько единиц совокупности имеют значение признака не больше, чем рассматриваемое.

Кумулята

Распределение признака в вариационном ряду по накопленным частотам (частостям) изображается с помощью кумуляты.

Кумулята или кумулятивная кривая в отличие от полигона строится по накопленным частотам или частостям. При этом на оси абсцисс помещают значения признака, а на оси ординат — накопленные частоты или частоты (рис. 3).



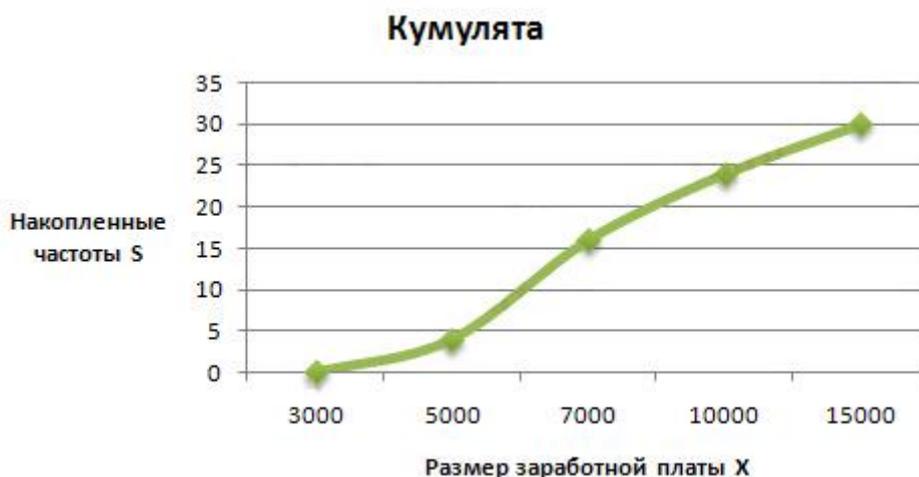
Рис. 3. Кумулята распределения домохозяйств по размеру

4. Рассчитаем накопленные частоты:

Накопленная частота первого интервала рассчитывается следующим образом: $0 + 4 = 4$, для второго: $4 + 12 = 16$; для третьего: $4 + 12 + 8 = 24$ и т.д.

Размер заработной платы руб в месяц X_i	Численность работников чел. f_i	Накопленные частоты S
до 5000	4	4
5000 — 7000	12	16
7000 — 10000	8	24
10000 — 15000	6	30
Итого:	30	-

При построении кумуляты накопленная частота (частость) соответствующего интервала присваивается его верхней границе:



Оги́ва

Оги́ва строится аналогично кумуляте с той лишь разницей, что накопленные частоты помещают на оси абсцисс, а значения признака — на оси ординат.

Разновидностью кумуляты является кривая концентрации или график Лоренца. Для построения кривой концентрации на обе оси прямоугольной системы координат наносится масштабная шкала в процентах от 0 до 100. При этом на оси абсцисс указывают накопленные частоты, а на оси ординат — накопленные значения доли (в процентах) по объему признака.

Равномерному распределению признака соответствует на графике диагональ квадрата (рис. 4). При неравномерном распределении график представляет собой вогнутую кривую в зависимости от уровня концентрации признака.

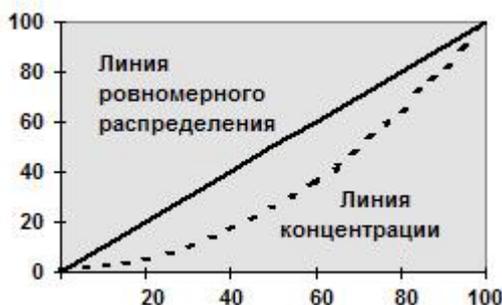


Рис. 4. Кривая концентрации

ЗАДАНИЯ ДЛЯ ВЫПОЛНЕНИЯ

Задание 1. Имеются следующие данные об успеваемости 30 студентов группы по теории статистики в летнюю сессию прошлого учебного года: 5,4, 3, 3, 5,4,4, 4, 3, 4, 4, 5, 4, 4, 3, 2, 5, 3, 4, 4, 4, 3, 2, 5, 2, 5, 5, 2, 3, 3.

Постройте:

а) ряд распределения студентов по оценкам, полученным в сессию, и изобразите его графически;

б) ряд распределения студентов по уровню успеваемости, выделив в нем две группы студентов: неуспевающих (2 балла), успевающих (3 балла и выше);

в) укажите, каким видом ряда распределения (вариационным или атрибутивным) является каждый из этих двух рядов.

Задание 2. Известны следующие данные о результатах тестирования работников банка в текущем году (баллов):

108	106	122	118	109	120	117
117	102	118	130	108	119	108
118	116	115	114	114	117	116
106	114	119	112	119	106	105

Выявить структуру работников банка по результатам тестирования, выделив четыре группы с равными интервалами и графически отобразить её с помощью круговой диаграммы.

Задание 3. На основе имеющихся данных о стаже работы и среднемесячной заработной плате сотрудников кредитной организации постройте ряд распределения по признаку стаж работы, образовав 5 групп с равными интервалами. Отобразить построенный ряд в виде гистограммы. Сделайте выводы.

№ сотрудника банка п/п	Стаж работы, лет	Месячная зарплата, руб.	№ сотрудника банка п/п	Стаж работы, лет	Месячная зарплата, руб.
1	1,0	15022	11	12,0	16561
2	6,5	15424	12	10,5	16309
3	9,0	16531	13	9,0	16255
4	4,5	15490	14	5,0	16093
5	6,0	15677	15	10,6	16392
6	2,5	15096	16	5,0	15962
7	2,7	15425	17	5,4	15851
8	16,0	17304	18	7,5	16120
9	14,0	17033	19	8,0	16398
10	11,0	18271	20	8,5	16847

Практическая подготовка 3. Абсолютные и относительные величины в статистике

Алгоритм выполнения работы

Относительный показатель динамики (ОПД) представляет собой отношение уровня исследуемого процесса или явления за данный период времени (по состоянию на данный момент времени) к уровню этого же процесса или явления в прошлом:

$$\text{ОПД} = \frac{\text{Текущий показатель}}{\text{Предыдущий или базисный показатель}}$$

Относительный показатель структуры (ОПС) представляет собой соотношение структурных частей изучаемого объекта и их целого:

$$\text{ОПС} = \frac{\text{Показатель, характеризующий часть совокупности}}{\text{Показатель по всей совокупности в целом}}$$

Относительный показатель структуры выражается в долях единицы или в процентах. Рассчитанные величины, соответственно называемые долями или удельными весами, показывают, какой долей обладает или какой удельный вес имеет *i*-я часть в общем итоге.

Пример 2. Основные показатели деятельности малых предприятий в Ставропольском крае по отраслям экономики края представлены следующими данными:

	единиц	
	Базисный период	Отчетный период
Всего, в том числе	7924	7391
Промышленность	1165	1188
Сельское хозяйство	201	333
Транспорт и связь	145	153
Строительство	1018	855
Торговля и общественное питание	4584	4127
Финансы и кредит	22	49
прочие	789	686

Определить относительные показатели структуры и динамики, оформить в таблицу.

ОПС характеризует какой долей обладает или какой удельный вес имеет та или иная часть в общем итоге:

$$\text{ОПС} = \frac{\text{Показатель, характеризующий часть совокупности}}{\text{Показатель по всей совокупности в целом}}$$

1. Определим относительный показатель структуры для базисного периода:

$$\begin{aligned} \text{ОПС}_{\text{пром}} &= 1165/7924 * 100\% = 0,148 * 100\% = 14,8\% \\ \text{ОПС}_{\text{с/х}} &= 201/7924 * 100\% = 0,025 * 100\% = 2,5\% \\ \text{ОПС}_{\text{тран}} &= 145/7924 * 100\% = 0,018 * 100\% = 1,8\% \\ \text{ОПС}_{\text{стр}} &= 1018/7924 * 100\% = 0,128 * 100\% = 12,8\% \\ \text{ОПС}_{\text{торг}} &= 4584/7924 * 100\% = 0,578 * 100\% = 57,8\% \\ \text{ОПС}_{\text{фин}} &= 22/7924 * 100\% = 0,003 * 100\% = 0,3\% \\ \text{ОПС}_{\text{прочие}} &= 789/7924 * 100\% = 0,10 * 100\% = 10\% \end{aligned}$$

Аналогично рассчитываются относительные показатели структуры для отчетного периода.

ОПД показывает во сколько раз текущий уровень превышает предшествующий (базисный). Относительный показатель динамики это отношение уровня исследуемого процесса или явления за данный период времени к уровню этого же процесса или явления в прошлом:

$$\text{ОПД} = \frac{\text{Текущий показатель}}{\text{Предыдущий или базисный показатель}}$$

$$\begin{aligned} \text{ОПД}_{\text{пром}} &= 1188/1165 * 100\% = 102\% \\ \text{ОПД}_{\text{с/х}} &= 333/201 * 100\% = 166\% \\ \text{ОПД}_{\text{тран}} &= 153/145 * 100\% = 105,5\% \\ \text{ОПД}_{\text{стр}} &= 855/1018 * 100\% = 84\% \\ \text{ОПД}_{\text{торг}} &= 4127/4584 * 100\% = 90\% \\ \text{ОПД}_{\text{фин}} &= 49/22 * 100\% = 223\% \\ \text{ОПД}_{\text{пр}} &= 686/789 * 100\% = 87\% \end{aligned}$$

Оформим полученные результаты в таблицу 4:

Таблица 4 - Основные показатели деятельности малых предприятий в Ставропольском крае по отраслям экономики края представлены следующими данными

	Базисный период		Отчетный период		Отчетный период в % к базисному
	единиц	%	единиц	%	
Всего,	7924		7391		93,3
в том числе:					
промышленность	1165	14,8	1188	16	102
сельское хозяйство	201	2,5	333	4,5	166
транспорт и связь	145	1,8	153	2,1	105,5
строительство	1018	12,8	855	11,6	84
торговля	4584	57,8	4127	55,8	90
финансы и кредит	22	0,3	49	0,7	223
прочие	789	10	686	9,3	87

Вывод: по полученным результатам можно судить о том, что как в базисном так и в отчетном периоде преобладает торговля. Ее доля составляет более 50%. Показатели деятельности малых предприятий в Ставропольском крае по отраслям экономики в отчетном периоде по отношению к базисному увеличились: промышленность на 2%, сельское хозяйство на 60%, финансы и кредит более чем в 2 раза. В тоже время уменьшились строительство и торговля.

ЗАДАНИЕ 2

Для изучения формы выражения статистических показателей и их экономической интерпретации по имеющимся данным необходимо:

1. рассчитать относительные показатели структуры;
2. относительные показатели динамики;
3. сделать краткие выводы.

Расчеты относительных показателей структуры и динамики результаты необходимо представить в виде таблицы 4:

Таблица 4 - Анализ производства продукции по отраслям всей промышленности Ставропольского края

Отрасли промышленности	2001		2002		2002 г. в % к 2001 г.
	млн. руб.	%	млн. руб.	%	
Машиностроение					
...					

Исходные данные

Вариант 1. Производство продукции по отраслям всей промышленности Ставропольского края характеризуется следующими данными:

	млн. руб.	
	предыдущий	отчетный
Всего по краю, в том числе:		
электроэнергетика и топливная промышленность	38282,2	48241,0
химическая и нефтехимическая промышленность	12697,8	14097,0
машиностроение, деревообрабатывающая и целлюлозно-бумажная промышленность	7914,5	7374,2
промышленность строительных материалов	4340,6	5826,4
легкая промышленность	1088,3	1233,4
пищевая промышленность	1408,4	1590,1
прочая промышленность	7938,7	9697,5
	2893,9	8422,4

Вариант 2. Данные о производстве потребительских товаров в Ставропольском крае приводятся ниже:

	млн. руб.	
	предыдущий	отчетный
Потребительские товары, всего в том числе:	8758,3	11103,6

пищевые продукты	6575,5	8391,9
винно-водочные изделия и пиво	389,4	472,0
непродовольственные товары	1793,4	2239,7

Вариант 3. Основные показатели работы Минераловодского отделения Северо-Кавказской железной дороги характеризуются грузооборотом:

	тыс. тонн	
	предыдущий	отчетный
Отправлено грузов всего, в том числе:	12684,3	14491,8
нефтегрузы	2326,1	3024,2
лесные грузы	10,9	11,4
минеральные удобрения	1968,0	1810,7
хлебные грузы	1320,0	2341,6
цемент	596,4	846,5
металлолом	288,2	453,4
прочие грузы	6174,7	6004,0

Вариант 4. Количество оказанных организациями почтовой связи услуг (по видам) в Ставропольском крае менялось следующим образом:

	тыс. шт.	
	предыдущий	отчетный
Отправлено всего, в том числе:	70212	71866
печатных изданий	46553	44673
писем, карточек, бандеролей	17308	20581
посылок	150	131
денежных переводов	767	809
телеграмм	1006	946
пенсионных выплат	4428	4726

Вариант 5. Данные о посевных площадях основных сельскохозяйственных культур в хозяйствах всех категорий Ставропольского края приводятся ниже:

	тыс. га	
	предыдущий	отчетный
Зерновые культуры	1863,3	1999,2
Подсолнечник	206,9	223,2
Сахарная свекла	16,9	18,0
Картофель	37,7	38,8
Овощи	23,9	22,2
Плоды и ягоды	22,3	20,4
Виноград	9,2	8,9
Итого:	2180,2	2330,7

Вариант 6. Перевозки пассажиров крупными и средними организациями транспорта Ставропольского края характеризуются следующими данными:

	млн. пасс.-км	
	предыдущий	отчетный
Всего, в том числе:	398,1	368,0
железнодорожный	11,1	13,2

автомобильный	232,0	207,1
воздушный	0,5	0,6
трамвайный	58,6	53,7
троллейбусный	95,9	93,5

Вариант 7. Данные о валовом сборе зерновых культур Ставропольского края по видам приводится ниже:

	тыс. тонн	
	предыдущий	отчетный
Зерновые культуры всего, в том числе:	4720,8	6103,8
пшеница	3448,4	4570,0
ячмень	982,7	1116,5
овес	104,3	89,5
просо	39,3	42,4
гречиха	5,5	5,3
кукуруза на зерно	50,4	169,1
горох	90,2	111,0

Вариант 8. Данные заключительного учета заготовки кормов в Ставропольском крае приведены в таблице:

	тыс. тонн	
	предыдущий	отчетный
Заготовлено кормов, всего в том числе:	1138,1	976,0
сена	334,8	335,3
сенажа	352,4	347,5
соломы	431,0	256,9
кормовых корнеплодов	19,9	36,3

Вариант 9. Производство продукции животноводства в сельхозпредприятиях Ставропольском крае характеризуются следующими данными:

	тыс. тонн	
	предыдущий	отчетный
Реализовано скота и птицы на убой в живом весе, в том числе:	56,9	57,8
крупного рогатого скота	17,2	16,5
свиней	8,3	9,9
овец и коз	6,6	5,0
птицы	24,6	26,0
прочее мясо	0,2	0,4

Вариант 10. Объемы реализации сельскохозяйственных продуктов всеми сельхозпроизводителями в Ставропольском крае составили:

	тыс. тонн	
	предыдущий	отчетный
Зерно	2846,9	4665,2
Подсолнечник	133,8	162,1
Картофель	47,3	58,4
Овощи	49,5	59,2
Скот и птица	89,3	95,6
Молоко	262,9	271,8

Итого:	3429,7	5312,3
--------	--------	--------

Вариант 11. Число хозяйствующих субъектов, учтенных в ЕГРПО в Ставропольском крае:

единиц

	предыдуци й	отчетны й
Сельское хозяйство	22756	22973
Торговля и общественное питание	16159	16871
Промышленность	5516	5620
Строительство	3747	3854
Здравоохранение, физкультура и социальное обеспечение	3791	3923
Транспорт и связь	1025	1057
Наука и научное обслуживание	1263	1417
Всего:	54257	55715

Вариант 12. Распределение учтенных в ЕГРПО хозяйствующих субъектов по организационно-правовым формам в Ставропольском крае характеризуют следующие данные:

единиц

	предыдущий	отчетный
Всего учтено субъектов, в том числе:	62498	64606
коммерческие организации	48888	50277
некоммерческие организации	8127	8286
организации без права юридического лица	2491	2723
индивидуальные предприниматели	2992	3320

Вариант 13. Распределение учтенных в ЕГРПО хозяйствующих субъектов по формам собственности в Ставропольском крае представлено ниже:

единиц

	предыдущий	отчетны й
Учтенных объектов всего, из них:	62498	64606
сельское хозяйство	22756	22973
торговля и общественное питание	16159	16871
промышленность	5516	5620
строительство	3747	3854
транспорт и связь	1025	1057
наука и образование	2417	2524
управление	4965	5183

здравоохранение, физкультура и социальное обеспечение и прочее	5913	6524
--	------	------

Вариант 14. Основные показатели деятельности малых предприятий в Ставропольском крае по отраслям экономики края представлены следующими данными:

	единиц	
	предыдущий	отчетный
Всего, в том числе	7924	7391
промышленность	1165	1188
сельское хозяйство	201	333
транспорт и связь	145	153
строительство	1018	855
торговля и общественное питание	4584	4127
финансы и кредит	22	49
прочие	789	686

Вариант 15. Данные о продаже основных продуктов сельского хозяйства сельхозпредприятиями края в Ставропольском крае приводятся ниже:

	тыс. тонн	
	предыдущий	отчетный
Зерновые культуры	2762,8	3873,9
Подсолнечник	108,5	126,6
Сахарная свекла	205,5	382,8
Картофель	3,9	4,6
Овощи, плоды и ягоды	52,2	57,3
Скот и птица	68,2	70,1
Молоко	140,6	145,1
Всего:	3341,7	4660,4

Задания 3 .

Задание 1. Имеются следующие данные о величине капитала банка:

	2015	2016	2017	2018
Совокупный капитал, млн. руб.	2453	2968	3326	3415

Вычислите относительные показатели динамики с переменной и постоянной базой сравнения. Проверьте их взаимосвязь.

Задание 2. Объем продаж акций ПАО «Газпромбанк» в 2018 г. в сопоставимых ценах вырос по сравнению с предшествующим годом на 5% и составил 146 млн руб. Определите объем продаж в 2017 г.

Задание 3. Коммерческий банк планировал в 2018 г. по сравнению с 2017 г. увеличить выдачу автокредитов на 14,5%. Выполнение установленного плана составило 102,7%. Определите относительный показатель динамики выдачи автокредитов.

Задание 4. Восстановите недостающие цифры в таблице, используя абсолютные и относительные показатели.

Таблица 1 – Структура дебетовых и кредитных карт, выпущенных кредитными организациями на 01.04.2018-2019г.г. (источник:ЦБ РФ)

	2018		2019	
	трлн. руб.	%	трлн. руб.	%
Кредитные карты		11,8	41 483	
Дебетовые карты в том числе		88,2	582002	
дебетовые карты с «овердрафтом»		31,9	368472	
дебетовые карты без «овердрафта»		56,3	213530	
Итого	518942	100,0	623485	

Сделайте необходимые выводы.

Практическая подготовка 4. Виды средних величин и их исчисление в статистике

Алгоритм выполнения работы

Средняя величина – обобщающая характеристика совокупности однотипных явлений по какому-либо варьирующему признаку, который показывает уровень признака, отнесенный к единице совокупности. Средняя величина отражает общее и типичное для всей совокупности в конкретных условиях места и времени.

Степенные средние.

Гармоническая

Геометрическая

Квадратическая

Задача 1. Дан ряд чисел: 15; 15; 12; 14; 13. Найдите размах, среднее арифметическое, медиану и моду этого ряда.

Решение

1) Размах ряда чисел – это разность между наибольшим и наименьшим из этих чисел. В данном случае размах равен $R = 15 - 12 = 3$

2) Среднее арифметическое данного ряда находим по формуле средней арифметической простой. $X_{ср} = (15 + 15 + 12 + 14 + 13) / 5 = 13,8$

3) Для определения медианы необходимо предложенный ряд упорядочить – расположить числа, например, в порядке возрастания: 12; 13; 14; 15; 15. Медиана нечетного количества чисел в дискретном ряду – это число, записанное посередине. Медиана четного количества чисел – это среднее арифметическое двух чисел, находящихся посередине.

Поскольку в нашем случае количество чисел ряда нечетное, то $Me = 14$.

4) Мода дискретного ряда чисел – это число, которое встречается в данном ряду чаще других. Так как число 15 встречается в нашем ряду чаще других, то $Mo = 15$.

Задача 2. Имеется информация о численности студентов ВУЗов города и удельном весе (%) обучающихся студентов на коммерческой основе:

ВУЗы города	Общее число студентов (тыс. чел.)	Из них удельный вес, обучающихся на коммерческой основе (%)
УрФУ	15	15
УрГЭУ	3	10
УрГЮА	7	20

Определить: 1) средний удельный вес студентов ВУЗов, обучающихся на коммерческой основе; 2) число этих студентов.

Решение

Для решения расширим предложенную таблицу:

ВУЗы города	Общее число студентов (тыс. чел.) f	Из них удельный вес, обучающихся на коммерческой основе (%) x	Число студентов, обучающихся на коммерческой основе (тыс. чел.)
УрФУ	15	15	$15:100 \times 15 = 2,25$
УрГЭУ	3	10	$3:100 \times 10 = 0,30$
УрГЮА	7	20	$7:100 \times 20 = 1,4$
Итого:	25	-	3,95

Средний удельный вес студентов ВУЗов, обучающихся на коммерческой основе определим по формуле средней арифметической взвешенной: $X_{ср} = (15 \times 15 + 3 \times 10 + 7 \times 20) / (15 + 3 + 7) = 15,8\%$.

Ответ. Средний удельный вес студентов ВУЗов, обучающихся на коммерческой основе равен 15,8%, число этих студентов – 3 950 человек.

Задача 3. Сумма невыплаченной своевременно задолженности по кредитам на 1 июля составила 92,4 млн. денежных единиц. По отдельным отраслям экономики она распределялась следующим образом:

Отрасль народного хозяйства	Сумма невыплаченной задолженности, млн. денежных единиц	Удельный вес невыплаченной задолженности в общем объеме кредитов, %
А	32,0	20
В	14,0	28
С	46,4	16

Определить средний процент невыплаченной своевременно задолженности. О

боснуйте выбор формы средней.

Решение

Поскольку на различных предприятиях сумма задолженности по кредитам разная при разных удельных весах, то применим формулу средней гармонической взвешенной.

$$X_{\text{ср}} = \Sigma W / \Sigma(W/x) = (32+14+46,4)/(32/20+14/28+46,4/16) = 92,4/5 = 18,48 \%$$

Ответ. Средний процент невыплаченной своевременно задолженности равен 18,48%.

Задания для выполнения

Задание 1. Имеются следующие данные о ценах на акции банков в одном из городов РФ:

Цена за одну акцию, руб.	Стоимость пакета акций, тыс. руб.
300-400	29,4
400-500	20,5
500-600	7,3
600-700	7,0
700-800	4,0

Рассчитайте среднюю цену одной акции.

Задание 2. Качество обслуживания клиентов банка характеризуется следующими данными (за месяц):

Вид продукции	Процент претензий по качеству обслуживания клиентов	Общая сумма операций по обслуживанию клиентов банка, руб.
А	1,3	2135
В	0,9	3560
С	2,4	980

Определите средний процент претензий по качеству обслуживания клиентов в целом по банку.

Вид степенной средней	Показатель степени средней (k)	Формула расчета	
		Простая	Взвешенная
Гармоническая	- 1	$\bar{x} = \frac{n}{\sum_{i=1}^n \frac{1}{x_i}}$	$\bar{x} = \frac{\sum_{i=1}^n x_i f_i}{\sum_{i=1}^n x_i}$
Геометрическая	1	$x = \sqrt[n]{x_1 x_2 \dots x_n}$	$x = \sqrt[\sum f_i]{x_1^{f_1} x_2^{f_2} \dots x_n^{f_n}}$
Арифметическая	0	$\bar{x} = \frac{\sum_{i=1}^n x_i}{n}$	$\bar{x} = \frac{\sum_{i=1}^n x_i f_i}{\sum_{i=1}^n f_i}$
Квадратическая	2	$\bar{x} = \sqrt{\frac{\sum_{i=1}^n x_i^2}{n}}$	$\bar{x} = \sqrt{\frac{\sum_{i=1}^n x_i^2 f_i}{\sum_{i=1}^n f_i}}$

Средняя обозначается через \bar{x} . Черта сверху символизирует процесс осреднения индивидуальных значений. Частота – повторяемость отдельных значений признака – обозначается буквой f.

Вопрос о выборе средней решается в каждом отдельном случае, исходя из задач исследования и наличия исходной информации.

Средняя арифметическая простая используется в тех случаях, когда варианты или варьирующие признаки встречаются только по одному разу и имеют одинаковый вес в совокупности. Средняя арифметическая взвешенная используется, когда данные сгруппированы, а отдельные значения признака встречаются неодинаковое число раз.

Средняя гармоническая – это величина, обратная средней арифметической из обратных значений признака. Средняя гармоническая вычисляется в тех случаях, когда в качестве весов применяются не единицы совокупности, а произведения этих единиц на значения признака (то есть $M=x \cdot f$).

Средняя гармоническая простая исчисляется в тех случаях, когда веса одинаковы, то есть равны между собой.

Средняя геометрическая простая используется при вычислении среднего коэффициента роста (темпа роста) в рядах динамики.

Средняя квадратическая используется для расчетов среднего квадратического отклонения (?) при изучении темы «Показатели вариации».

Для вычисления средней в дискретных рядах варианты нужно умножить на частоты и сумму произведений разделить на сумму частот, то есть по

средней арифметической взвешенной
$$\bar{x} = \frac{\sum xf}{\sum f}$$
.

Для вычисления средней в интервальных рядах нужно перейти к дискретному ряду, то есть по каждой группе вычислить значение интервала, заменить интервал его средним значением и вычислить по формуле

$$\bar{x} = \frac{\sum \frac{\sum x}{2} \cdot f}{\sum f}$$

Для того чтобы проверить правильность выбора формул, надо учитывать:

- среднее значение признака не должно выходить за пределы минимального и максимального значений признака совокупности;
- среднее значение ближе к тому значению признака, которому соответствует большая частота.

Степенные средние дают обобщающую характеристику совокупности и являются абстрактными величинами, полученными расчетным путем, в то же время эти средние не отражают всех особенностей совокупности, они могут быть различными для одинаковых совокупностей или иметь одинаковое значение для совокупности с различным строением.

Задания для выполнения на практическом занятии

Задание 1. Сумма резервов, сформированных под возможные потери по ссудам, характеризуется следующими данными:

Годы	2013	2014	2015	2016	2017	2018
Сумма резервов под возможные потери по ссудам, млн. руб.	100	110	118	121	130	136

Определите средний уровень резерва за исследуемый период.

Задание 2. Оборотные средства банковского сектора характеризуются следующими данными:

Годы	На 01. 04. 2018	На 01.05 2018	На 01.06. 2018	На 01.07. 2018	На 01.08. 2018
Оборотные средства, млрд. руб.	434,8	483,8	523,2	559,9	599,4

Определите среднюю стоимость оборотных средств за представленный период.

Кроме степенных средних в статистике для относительной характеристики величины варьирующего признака и внутреннего строения рядов распределения пользуются структурными средними, которые представлены ,в основном, **модой и медианой**.

Мода — это наиболее часто встречающийся вариант ряда. Мода применяется, например, при определении размера одежды, обуви, пользующейся наибольшим спросом у покупателей. Модой для дискретного ряда является варианта, обладающая наибольшей частотой. При вычислении моды для интервального вариационного ряда необходимо сначала

определить модальный интервал (по максимальной частоте), а затем — значение модальной величины признака по формуле:

$$M_0 = x_0 + n \frac{f_m - f_{m-1}}{(f_m - f_{m-1}) + (f_m - f_{m+1})},$$

где:

M_0 — значение моды

x_0 — нижняя граница модального интервала

h — величина интервала

f_m — частота модального интервала

f_{m-1} — частота интервала, предшествующего модальному

f_{m+1} — частота интервала, следующего за модальным

Медиана — это значение признака, которое лежит в основе ранжированного ряда и делит этот ряд на две равные по численности части.

Для определения медианы в дискретном ряду при наличии частот

сначала вычисляют полусумму частот $\frac{\sum f_i}{2}$, а затем определяют, какое значение варианта приходится на нее. (Если отсортированный ряд содержит нечетное число признаков, то номер медианы вычисляют по формуле:

$$Me = (n * (\text{число признаков в совокупности}) + 1) / 2,$$

в случае четного числа признаков медиана будет равна средней из двух признаков находящихся в середине ряда).

При вычислении медианы для интервального вариационного ряда сначала определяют медианный интервал, в пределах которого находится медиана, а затем — значение медианы по формуле:

$$M_e = x_0 + h \frac{\frac{\sum f_i}{2} - S_{m-1}}{f_m},$$

где:

M_e — искомая медиана

x_0 — нижняя граница интервала, который содержит медиану

h — величина интервала

$\sum f_i$ — сумма частот или число членов ряда

S_{m-1} — сумма накопленных частот интервалов, предшествующих медианному

f_m — частота медианного интервала

Пример. Найти моду и медиану.

Возрастные группы	Число студентов	Сумма накопленных частот $\sum S$
До 20 лет	346	346
20 — 25	872	1218
25 — 30	1054	2272
30 — 35	781	3053

35 — 40	212	3265
40 — 45	121	3386
45 лет и более	76	3462
Итого	3462	

Решение:

В данном примере модальный интервал находится в пределах возрастной группы 25-30 лет, так как на этот интервал приходится наибольшая частота (1054).

Рассчитаем величину моды:

$$M_0 = x_0 + h \frac{f_m - f_{m-1}}{(f_m - f_{m-1}) + (f_m - f_{m+1})} = 25 + 5 \frac{1054 - 872}{(1054 - 872) + (1054 - 781)} = 27 \text{ лет.}$$

Это значит что модальный возраст студентов равен 27 годам.

Вычислим медиану. Медианный интервал находится в возрастной группе 25-30 лет, так как в пределах этого интервала расположена варианта, которая делит совокупность на две равные части ($\sum f_i / 2 = 3462 / 2 = 1731$). Далее подставляем в формулу необходимые числовые данные и получаем значение медианы:

$$M_e = x_0 + h \frac{\frac{\sum f_i}{2} - S_{m-1}}{f_m} = 25 + 5 \frac{\frac{3462}{2} - 1218}{1054} = 27,4 \text{ года.}$$

Это значит что одна половина студентов имеет возраст до 27,4 года, а другая свыше 27,4 года.

Кроме моды и медианы могут быть использованы такие показатели, как квартили, делящие ранжированный ряд на 4 равные части, децили - 10 частей и перцентили — на 100 частей.

Задания для выполнения

Задание 1. Рассчитать моду и медиану по данным таблицы.

Группы предприятий по стоимости ОПФ, у.е.	Число предприятий, f	Середина интервалов, х	Накопленная частота, S
14-16	2	15	2
16-18	6	17	8
18-20	10	19	18
20-22	4	21	22
22-24	3	22	25
Итого	25	-	-

Задание 2. Имеются данные о распределении рабочих 5 участков по их квалификации (тарифному разряду). Рассчитать моду и медиану по данным таблицы, используя программу Excel.

Тарифный разряд, X_i	Число рабочих f	Частость, w	Накопленная частота, S
2	1	0,05	1
3	5	0,25	6
4	8	0,4	14
5	4	0,2	18
6	2	0,1	20
Итого:	20	1,0	20

Практическая подготовка 5. Показатели вариации

Алгоритм выполнения работы

Вариацией называется колеблемость, многообразие, изменяемость величины признака. Вариация зависит от различных факторов и их сочетаний в каждом конкретном случае, например, успеваемость. Вариация бывает случайной и систематической.

Измерение вариации дает возможность оценить степень воздействия на данный признак других варьирующих признаков, определение вариации необходимо при организации выборочного наблюдения, построения статистических моделей и т.д. Вариация существует в пространстве и времени. Под вариацией в пространстве понимается колеблемость значений признака по различным территориям. Вариация во времени подразумевает изменение значений признака в различные периоды или моменты времени.

Показатели вариации делятся на две группы: абсолютные и относительные.

К абсолютным показателям относятся: размах вариации, среднее линейное отклонение, дисперсия и среднее квадратическое отклонение. К относительным показателям относятся: коэффициент осцилляции, линейный коэффициент вариации, коэффициент вариации.

Размах вариации (R) показывает наибольшее различие между единицами совокупности и рассчитывается как разность между наибольшими (x_{\max}) и наименьшими (x_{\min}) значениями варьирующего признака. Размах вариации выражается именованными числами:

$$R = x_{\max} - x_{\min}$$

Размах вариации - важный, но не единственный показатель колеблемости признака.

Для анализа вариации используется величина, вокруг которой происходят колебания и рассеяния значений признака. При обобщении этих колебаний снова применяется метод средних, чтобы найти среднюю величину этих отклонений. Такая средняя называется средним линейным отклонением (\bar{d}), которое определяется как средняя арифметическая из абсолютных значений отклонений вариант от средней:

простые

$$\bar{d} = \frac{\sum |x - \bar{x}|}{n}$$

взвешенные

$$\bar{d} = \frac{\sum [|x - \bar{x}| \cdot f]}{\sum f}$$

Средний квадрат отклонений (дисперсия):

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

$$\sigma^2 = \frac{\sum [(x - \bar{x})^2 \cdot f]}{\sum f}$$

Среднее квадратическое отклонение:

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

$$\sigma = \sqrt{\frac{\sum [(x - \bar{x})^2 \cdot f]}{\sum f}}$$

Относительные показатели:

1. Коэффициент осцилляции.

$$V_R = \frac{R}{\bar{x}} \cdot 100\%$$

2. Линейный коэффициент вариации.

$$V_d = \frac{d}{\bar{x}} \cdot 100\%$$

3. Коэффициент вариации.

$$V_\sigma = \frac{\sigma}{\bar{x}} \cdot 100\%$$

Если $V_\sigma \leq 33\%$, то совокупность однородна, \bar{x} типична для данной совокупности.

Задания для выполнения на практическом занятии

Задание 1. В целях контроля качества обслуживания клиентов банка выполнен мониторинг операционного дня деятельности кредитного отдела по обслуживанию 500 клиентов. Его результаты приведены в таблице.

Сумма выданных кредитов, тыс. руб.	1700	1800	1900	2000	2100
Количество кредитных операций, шт.	46	87	164	135	68

Определите: 1) Размах вариации; 2) дисперсию; 3) среднее квадратическое отклонение; 4) среднее линейное отклонение; 5) коэффициент вариации. Используйте для расчетов ресурс программы Excel.

Задание 2. Согласно приведенным данным в банке за один операционный день было выданы дебетовые карты различных видов

Виды карт	Visa Classic	Eurocard/MasterCard Mass	VisaGold	VisaBusiness
Сумма выданных средств, тыс. руб.	28 020	14 145	51 238	85 479
Количество	23	21	10	42

кредитных операций по пластиковым картам, шт.				
---	--	--	--	--

Определите: 1) Размах вариации; 2) дисперсию; 3) среднее квадратическое отклонение; 4) среднее линейное отклонение; 5) коэффициент вариации. Используйте для расчетов ресурс программы Excel.

Практическая подготовка 6. Ряды динамики. Показатели анализа динамических рядов

Алгоритм выполнения работы

Ряды динамики — это ряды статистических показателей, характеризующих развитие явлений природы и общества во времени. Публикуемые Госкомстатом России статистические сборники содержат большое количество рядов динамики в табличной форме. Ряды динамики позволяют выявить закономерности развития изучаемых явлений.

Ряды динамики содержат два вида показателей. Показатели времени (годы, кварталы, месяцы и др.) или моменты времени (на начало года, на начало каждого месяца и т.п.). Показатели уровней ряда. Показатели уровней рядов динамики могут быть выражены абсолютными величинами (производство продукта в тоннах или рублях), относительными величинами (удельный вес городского населения в %) и средними величинами (средняя заработная плата работников отрасли по годам и т. п.). В табличной форме ряд динамики содержит два столбца или две строки.

Правильное построение рядов динамики предполагает выполнение ряда требований:

- все показатели ряда динамики должны быть научно обоснованными, достоверными;
- показатели ряда динамики должны быть сопоставимы по времени, т.е. должны быть исчислены за одинаковые периоды времени или на одинаковые даты;
- показатели ряда динамики должны быть сопоставимы по территории;
- показатели ряда динамики должны быть сопоставимы по содержанию, т.е. исчислены по единой методологии, одинаковым способом;
- показатели ряда динамики должны быть сопоставимы по кругу учитываемых хозяйств. Все показатели ряда динамики должны быть приведены в одних и тех же единицах измерения.

Уровни интервального ряда характеризуют результат изучаемого процесса за период времени: производство или реализация продукции (за год, квартал, месяц и др. периоды), число принятых на работу, число родившихся и т.п. Уровни интервального ряда можно суммировать. При этом получаем такой же показатель за более длительные интервалы времени.

Средний уровень в интервальных рядах динамики (\bar{y}) исчисляется по формуле средней арифметической простой:

$$\bar{y} = \frac{\sum y}{n}$$

y — уровни ряда (y_1, y_2, \dots, y_n),

n — число периодов (число уровней ряда).

Уровни моментных рядов динамики характеризуют состояние изучаемого явления на определенные моменты времени. Каждый последующий уровень включает в себя полностью или частично

предыдущий показатель. Так, например, число работников на 1 апреля 1999 г. полностью или частично включает число работников на 1 марта.

Если сложить эти показатели, то получим повторный счет тех работников, которые работали в течение всего месяца. Полученная сумма экономического содержания не имеет, это расчетный показатель.

В моментных рядах динамики с равными интервалами времени средний уровень ряда исчисляется по формуле средней хронологической:

$$\bar{y} = \frac{y_1 + y_2 + y_3 + \dots + y_{n-1} + y_n}{n-1}$$

y - уровни моментного ряда;

n - число моментов (уровней ряда);

$n - 1$ — число периодов времени (лет, кварталов, месяцев).

Анализ рядов динамики

Для обоснованной оценки развития явлений во времени необходимо исчислить аналитические показатели: абсолютный прирост, коэффициент роста, темп роста, темп прироста, абсолютное значение одного процента прироста.

Абсолютные приросты (Δy) показывают, на сколько единиц изменился последующий уровень ряда по сравнению с предыдущим (или по сравнению с начальным уровнем). Формулы расчета можно записать следующим образом:

$$\Delta y^{\sigma} = y_n - y_0$$

$$\Delta y^4 = y_n - y_{n-1} \text{ , где}$$

Δy — абсолютный прирост (Δy^4 — цепной, Δy^{σ} — базисный);

y_n — уровень ряда за отчетный период;

y_{n-1} — уровень ряда предыдущего периода;

y_0 — уровень ряда начальный.

При уменьшении абсолютных значений ряда будет соответственно "уменьшение", "снижение".

Коэффициент роста показывает, во сколько раз изменился уровень ряда по сравнению с предыдущим или по сравнению с начальным уровнем. Формулы расчета можно записать следующим образом:

$$K_p^4 = \frac{y_n}{y_{n-1}} ; K_p^{\sigma} = \frac{y_n}{y_0}$$

Темпы роста показывают, сколько процентов составляет последующий уровень ряда по сравнению с предыдущим или по сравнению с начальным уровнем. Формулы расчета можно записать следующим образом:

$$T_p^4 = \frac{y_n}{y_{n-1}} \times 100\% \text{ или } T_p^4 = K_p^4 \times 100\%$$

Темпы прироста показывают, на сколько процентов увеличился уровень отчетного периода по сравнению с предыдущим или по сравнению с начальным уровнем. Формулы расчета можно записать следующим образом:

$T_{пр} = T_p - 100\%$ или $T_{пр} = \text{абсолютный прирост} / \text{уровень предшествующего периода} * 100\%$

Если абсолютные уровни в ряду уменьшаются, то темп будет меньше 100% и соответственно будет темп снижения (темп прироста со знаком минус).

Абсолютное значение 1% прироста показывает, сколько единиц надо произвести в данном периоде, чтобы уровень предыдущего периода возрос на 1%. Определить величину абсолютного значения 1% прироста можно двумя способами:

уровень предшествующего периода разделить на 100;

цепные абсолютные приросты разделить на соответствующие цепные темпы прироста.

$$= \frac{\Delta y}{T_{пр}} = \frac{y_{n-1}}{100}$$

Абсолютное значение 1% прироста =

В динамике, особенно за длительный период, важен совместный анализ темпов прироста с содержанием каждого процента прироста или снижения.

Заметим, что рассмотренная методика анализа рядов динамики применима как для рядов динамики, уровни которых выражены абсолютными величинами (т, тыс. руб., число работников и т.д.), так и для рядов динамики, уровни которых выражены относительными показателями (% брака, % зольности угля и др.) или средними величинами (средняя урожайность в ц/га, средняя заработная плата и т.п.).

Средний годовой темп роста и средний годовой темп прироста

Среднегодовой темп роста исчисляется в следующей последовательности:

сначала по формуле средней геометрической исчисляют среднегодовой коэффициент роста (снижения) — $\overline{K_p}$

на базе среднегодового коэффициента определяют среднегодовой темп роста ($\overline{T_p}$) путем умножения коэффициента на 100%:

$$\overline{T_p} = \overline{K_p} * 100\%$$

Среднегодовой темп прироста ($\overline{T_{np}}$) определяется путем вычитания из темпа роста 100%.

$$\overline{T_{np}} = \overline{T_p} - 100\%$$

Среднегодовой коэффициент роста (снижения) по формулам средней геометрической может быть исчислен двумя способами:

1) на базе абсолютных показателей ряда динамики по формуле:

$$\overline{K_p} = \sqrt[n-1]{\frac{Y_n}{Y_1}} = \sqrt[4]{\frac{234}{200}} = \sqrt[4]{1,17} = 1,040,$$

n — число уровней;

n — 1 — число лет в период;

2) на базе ежегодных коэффициентов роста по формуле

$$\overline{K_p} = \sqrt[m]{k_1 \cdot k_2 \dots k_m} = \sqrt[4]{1,05 \times 1,038 \times 1,055 \times 1,017} = \sqrt[4]{1,17} = 1,040.$$

m — число коэффициентов.

Результаты расчета по формулам равны, так как в обеих формулах показатель степени — число лет в периоде, в течение которого происходило изменение. А подкоренное выражение — это коэффициент роста показателя за весь период времени.

Среднегодовой темп прироста определяется путем вычитания из среднегодового темпа роста 100%. В нашем примере среднегодовой темп прироста равен

$$\overline{T_{np}} = \overline{T_p} - 100\% = 104.0\% - 100\% = 4.0\%$$

Среднегодовой темп роста (прироста) позволяет сравнивать динамику развития взаимосвязанных явлений за длительный период времени (например, среднегодовые темпы роста численности работающих по отраслям экономики, объема производства продукции и др.), сравнивать динамику какого-либо явления по разным странам, исследовать динамику какого-либо явления по периодам исторического развития страны.

Задания для выполнения

Задание 1. Вычислите цепные и базисные абсолютные приросты, темпыроста и прироста, а также абсолютное значение 1 % прироста по следующим данным :

годы	2013	2014	2015	2016	2017	2018
Выполнено расчетных операций с использованием пластиковых карт, млн. руб.	140,1	223,8	195,7	237,4	179,3	189,1

Рассчитайте средние показатели для данного ряда динамики.

Задание 2. Формирование резерва на возможные потери по ссудам характеризуется следующими данными:

Годы	2013	2014	2015	2016	2017	2018
Сумма резервов под возможные потери по ссудам, млн. руб.	100	110	118	121	130	136

Определите показатели анализа рядов динамики базисным методом:

- абсолютный прирост (млн. руб.);
- темп роста (в %);
- темп прироста (в %).

Полученные показатели представить в виде таблицы.

Вычислите средний уровень ряда. Произведите обработку данного ряда динамики способом аналитического выравнивания по прямой, используя программу Excel.

Сделайте вывод.

Задание 3. Оборотные средства банковского сектора характеризуются следующими данными:

Годы	2014	2015	2016	2017	2018
Оборотные средства, млрд. руб.	434,8	483,8	523,2	559,9	599,4

Определите базисным способом:

а) абсолютный прирост (млрд. руб.);

б) темп роста (в %);

в) темп прироста (в %);

г) среднегодовой уровень оборотных средств (млрд. руб.).

Полученные результаты представьте в таблице, проанализируйте их и сделайте вывод.

Практическая подготовка 7. Способы обработки динамических рядов

Алгоритм выполнения работы

В ходе обработки динамического ряда важнейшей задачей является выявление основной тенденции развития явления (тренда) и сглаживание случайных колебаний. Для решения этой задачи в статистике существуют особые способы, которые называют методами выравнивания.

Выделяют три основных способа обработки динамического ряда:

а) укрупнение интервалов динамического ряда и расчет средних для каждого укрупненного интервала;

б) метод скользящей средней;

в) аналитическое выравнивание (выравнивание по аналитическим формулам).

Выравнивание рядов динамики по методу наименьших квадратов, как и выравнивание, посредством других приёмов, должно осуществляться в пределах качественно однородных периодов. Если в динамическом ряду есть качественно неоднородные периоды, то выявлять тенденцию целесообразно в пределах каждого из них.

Расчёт параметров можно значительно упростить, если отсчёт времени $t = 0$ осуществлять с середины динамического ряда. Тогда значения t , размещённые выше середины, будут отрицательными, а ниже - положительными. В обоих случаях $\sum t = 0$. Для этого уровень, который будет пребывать в середине ряда динамики, берут за условное начало отсчёта или нулевое значение.

Для того чтобы сумма показателей времени равнялась нулю, условные обозначения нужно давать таким образом:

при нечётном числе уровней ряда динамики, чтобы выполнить условие $\sum t = 0$, уровень, который будет прибывать в середине ряда, приравнивают к нулю, а уровни, расположенные выше его, помечают числами со знаком «минус» (-1; -2; -3 и т. п.), а ниже - числами со знаком «плюс» (+1; +2; +3 и т. д.);

при парном числе уровней ряда динамики уровни, которые лежат выше среднего значения (оно находится в середине между двумя средними датами), помечают натуральными числами со знаком «минус» (-1; -3; -5 и т. п.), а уровни, которые лежат ниже среднего значения, — натуральными числами со знаком «плюс» (+1; +2; +3 и т. д.).

При условии, что $\sum t = 0$, система нормальных уравнений упрощается, приобретая в случае линейной зависимости такой вид:

$$\begin{cases} a_0 n = \sum y \\ a_1 \sum t^2 = \sum ty \end{cases},$$

Откуда

$$a_0 = \frac{\sum y}{n},$$

$$a_1 = \frac{\sum yt}{\sum t^2},$$

(3.54)

Парабола второго порядка ($Y_t = a_0 + a_1 t + a_2 t^2$) используется для описания рядов динамики, в которых меняется направление развития: со снижения показателей на их рост и наоборот.

Параметр a_2 называется коэффициентом регрессии и характеризует изменение интенсивности развития в единицу времени. При $a_2 > 0$ наблюдается ускоренное развитие, при $a_2 < 0$ – замедленное.

$$a_0 n + a_2 \sum t^2 = \sum y$$

$$a_1 \sum t^2 = \sum yt$$

$$a_0 \sum t^2 + a_2 \sum t^4 = \sum yt^2$$

Отсюда

$$a_1 = \frac{\sum yt}{\sum t^2}$$

$$a_0 = \frac{\sum t^4 \cdot \sum y - \sum t^2 \cdot \sum yt^2}{n \sum t^4 - (\sum t^2)^2}$$

$$a_2 = \frac{n \sum yt^2 - \sum t^2 \cdot \sum y}{n \sum t^4 - (\sum t^2)^2}$$

Задания для выполнения

Задание 1. Формирование резерва на возможные потери по ссудам характеризуется следующими данными:

Годы	2013	2014	2015	2016	2017	2018
Сумма резервов под возможные потери по ссудам, млн. руб.	100	110	118	121	130	136

С целью выявления тенденции развития представленного динамического ряда произведите аналитическое выравнивание по прямой.

Задание 2. По имеющимся данным в таблице об объеме совершенных депозитных операций банка за операционный день

Годы	2014	2015	2016	2017	2018
Депозитные операции, тыс. руб.	885,7	932,6	980,1	1028,7	1088,4

произведите аналитическое выравнивание по прямой целью выявления тенденции развития представленного динамического ряда.

Практическая подготовка 8. Статистические индексы

Алгоритм выполнения работы

Индекс – это относительная величина, показывающая, во сколько раз уровень изучаемого явления в данных условиях отличается от уровня того же явления в других условиях.

Буквой i обозначают индивидуальные индексы, а буквой I – общие индексы. Знак внизу справа означает период: 0 – базисный, 1 – отчетный.

Все экономические индексы классифицируются по следующим признакам:

- степени охвата явления или процесса;
- базе сравнения;
- виду весов;
- форме построения;
- характеру объекта исследования;
- объекту исследования;
- составу явления;
- периоду исчисления;

С помощью индексов решаются следующие задачи:

- измерение динамики массового процесса или явления за два и более периодов времени;
- измерение динамики среднего показателя;
- измерение соотношения показателей по разным территориям;
- определение степени влияния изменений значений одних показателей на динамику других.

Индивидуальные и общие индексы

В зависимости от экономического назначения индивидуальные индексы бывают: физического объема продукции, себестоимости, цен, трудоемкости и т.д.

Индекс физического объема продукции рассчитывается по формуле:

$$i_q = \frac{q_1}{q_0},$$

где q_1 – количество какого-либо товара в отчетном периоде;

q_0 – количество какого-либо товара в базисном периоде.

Индивидуальные индексы других показателей строятся аналогично. Это могут быть индексы цен, себестоимости продукции, производительности труда и т.п.

Общие индексы строятся для количественных и качественных показателей.

Основной формой общего индекса является агрегатный индекс.

В агрегатной форме сравниваются две суммы одноименных показателей. Числитель и знаменатель представляют собой сумму

произведений двух величин, одна из которых меняется (индексируемая величина), другая остается неизменной (вес индекса).

Индексируемой величиной называется признак, изменение которого изучается (цена товара, курс акций, затраты рабочего времени, урожайность и пр.)

Вес индекса – величина, служащая для целей соизмерения индексируемых величин.

Методика построения агрегатного индекса предусматривает решение трех вопросов:

- какая величина будет индексируемой;
- по какому составу разнородных элементов явления необходимо вычислить индекс;
- что будет служить весом при расчете индекса.

При выборе веса индекса руководствуются следующим правилом: если строится индекс количественного показателя, то веса берутся за базисный год, при построении индекса качественного показателя используются веса отчетного периода.

В качестве меры соизмерения разнородных продуктов используют цену, трудоемкость продукции, себестоимость и др.

Рассмотрим особенности построения агрегатных индексов для наиболее часто встречающихся в статистике показателей.

Агрегатный индекс физического объема. Предположим, нужно показать изменение объема выпускаемой продукции на мебельной фабрике в отчетном периоде по сравнению с базисным периодом. Фабрика выпускает столы, шкафы, диваны. Сложить эту различную несоизмеримую продукцию в физических единицах нельзя. Но если представить всю продукцию в стоимостном выражении (приняв цены в качестве соизмерителя), тогда можно сравнивать стоимость продукции одного года со стоимостью продукции другого года. А чтобы изменение цен не влияло на величину стоимостного показателя, продукцию двух лет надо оценить в одних и тех же ценах. Если выпуск продукции условно обозначить через q , а цены - через p , то формула агрегатного индекса физического объема выразится следующим образом:

$$I_q = \frac{\sum q_1 p_0}{\sum q_0 p_0},$$

где q_1 и q_0 - количество продукции соответственно в отчетном и базисном периодах;

p_1 и p_0 - цены соответственно сопоставимые и базисного периода.

Разность между числителем и знаменателем $\sum q_1 p_0 - \sum q_0 p_0$ - покажет изменение стоимости продукции (в абсолютном выражении) за счет изменения объема продукции.

При построении агрегатного индекса физического объема могут использоваться и другие соизмерители. Так, например, если принять в качестве соизмерителей себестоимость единицы продукции в базисном

периоде (z_0), то агрегатный индекс физического объема можно записать как

$$I_q = \frac{\sum q_1 z_0}{\sum q_0 z_0}.$$

Тогда разность $\sum q_1 z_0 - \sum q_0 z_0$ покажет, как изменились общие затраты (издержки) на производство в связи с изменением выпуска продукции.

Если в качестве соизмерителей принять затраты времени на единицу продукции в базисном периоде t_0 , то формула агрегатного индекса физического объема будет иметь вид:

$$I_q = \frac{\sum q_1 t_0}{\sum q_0 t_0},$$

а разность $\sum q_1 t_0 - \sum q_0 t_0$ будет характеризовать изменение общих затрат времени на производство продукции за счет изменения объема выпуска.

По аналогии с индексом физического объема для определенного набора товаров (продуктов) может быть построен и агрегатный индекс цен (индекс качественного показателя) и агрегатные индексы для многих других показателей.

С помощью индексов можно определить влияние отдельных фактов на изменение динамики сложного явления. Используя взаимосвязь индексов, можно установить, например, в какой мере выпуск продукции вырос за счет увеличения численности работников и в какой мере – за счет повышения производительности труда.

Пример 4. Имеются следующие данные за два периода о ценах и объемах реализации трех видов товаров по одному из торговых предприятий (табл. 6):

Таблица 6 - Данные о ценах и объемах реализации товаров по одному из торговых предприятий

Вид товара	Базисный период		Текущий период	
	цена за единицу, руб.	продано товаров, шт.	цена за единицу, руб.	продано товаров, шт.
	P_0	q_0	P_1	q_1
А	45	2500	87	1700
Б	27	830	35	2300
В	12	610	14	1000

Рассчитать:

- 1) индивидуальные индексы физического объема реализации товаров;
- 2) общий индекс физического объема реализации;
- 3) общий индекс цены;
- 4) индекс товарооборота (стоимость товаров). Показать взаимосвязь между индексами. Сделать выводы.

- 1) товар А: $i = \frac{1700}{2500} = 0,68$;
- 2) товар Б: $i = \frac{2300}{830} = 2,77$;
- 3) товар В: $i = \frac{1000}{610} = 1,64$.

Индекс физического объема определяется по формуле:

$$I_q = \frac{\sum q_1 p_0}{\sum q_0 p_0} = \frac{1700 \cdot 45 + 2300 \cdot 27 + 1000 \cdot 12}{2500 \cdot 45 + 830 \cdot 27 + 610 \cdot 12} = 1,059 * 100\% = 105,9\%.$$

Индекс цены:

$$I_p = \frac{\sum q_1 p_1}{\sum q_1 p_0} = \frac{1700 \cdot 87 + 2300 \cdot 35 + 1000 \cdot 14}{1700 \cdot 45 + 2300 \cdot 27 + 1000 \cdot 12} = 1,609 * 100\% = 160,9\%.$$

Индекс товарооборота:

$$I_{pq} = \frac{\sum q_1 p_1}{\sum q_0 p_0} = \frac{1700 \cdot 87 + 2300 \cdot 35 + 1000 \cdot 14}{2500 \cdot 45 + 830 \cdot 27 + 610 \cdot 12} = 1,704 * 100\% = 170,4\%$$

Взаимосвязь между индексами выражается следующей формулой:

$$I_{pq} = I_p \cdot I_q ,$$

$$I_{pq} = 105,9 * 160,9 = 170,4\%.$$

Изменение товарооборота за счет изменения количества проданного товара составило 105,9%, т.е. увеличилось на 5,9%; за счет изменения цены товарооборот увеличился на 60,9%, а за счет изменения и цены и количества проданного товара составил 170,4%.

ЗАДАНИЕ 4

По данным о себестоимости и объемах производства продукции предприятия необходимо определить:

1. индивидуальные и сводные индексы себестоимости;
2. сводный индекс физического объема производства продукции;
3. сводный индекс затрат на производство продукции;
4. показать взаимосвязь сводных индексов и сделать выводы.

Исходные данные

Номер вида продукции	Произведено, ц		Себестоимость, руб./кг	
	Базисный период	Отчетный период	Базисный период	Отчетный период
1	80	120	35	30
2	450	600	28	27
3	100	150	40	38
4	300	360	32	30
5	75	90	30	25

6	400	620	28	26
7	50	70	46	42
8	260	350	40	35
9	70	140	50	45
10	480	550	46	32
11	90	150	48	40
12	320	400	42	38
13	60	100	20	18
14	420	560	17	15
15	120	180	45	42
16	680	800	44	40
17	200	250	80	82
18	360	450	70	68
19	400	500	60	62
20	100	120	55	50

На основании исходных данных таблицы необходимо в соответствии с индивидуальным вариантом студента отобрать только значения по указанным в нижеследующей таблице номерам видов продукции.

Вариант	1	2	3	4	5
Номера видов продукции	1,2,3	4,5,6	7,8,9	10,11,12	13,14,15
Вариант	6	7	8	9	10
Номера видов продукции	4,13,14	3,15,16	2,17,18	1,19,20	16,17,18
Вариант	11	12	13	14	15
Номера видов продукции	5,11,12	1,7,14	2,8,15	3,9,16	4,10,17
Вариант	16	17	18	19	20
Номера видов продукции	5,11,18	6,12,19	7,13,20	1,10,20	5,9,19

Практическая подготовка 9. Статистическое изучение взаимосвязей

Задачи корреляционного анализа сводятся к измерению тесноты связи между варьирующими признаками, определению неизвестных причинных связей и оценке факторов, оказывающих наибольшее влияние на результативный признак.

Задачи регрессионного анализа лежат в сфере установления формы зависимости, определения функции регрессии, использования уравнения для оценки неизвестных значений зависимой переменной.

Корреляционный и регрессионный анализ как общее понятие включает в себя измерение тесноты, направления связи и установление аналитического выражения (формы) связи (регрессионный анализ).

По *форме зависимости* различают:

- линейную регрессию, которая выражается уравнением прямой (линейной функцией) вида: $y = b \cdot x + a$;

- нелинейную регрессию, которая выражается уравнениями вида параболы, гиперболы и т.д.

По *направлению связи* различают:

- прямую регрессию (положительную), возникающую при условии, если с увеличением (уменьшением) независимой величины значения зависимой также увеличиваются (уменьшаются);

- обратную (отрицательную) регрессию, появляющуюся при условии, что с увеличением (уменьшением) независимой величины зависимая уменьшается (увеличивается).

Для характеристики влияния изменений X на вариацию Y служат методы регрессионного анализа. В случае парной линейной зависимости уравнение регрессии примет вид:

$$\hat{y}_x = b \cdot x + a,$$

где \hat{y}_x - рассчитанное значение результативного признака после подстановки в уравнение регрессии.

Формулы для вычисления параметров a и b получены с помощью *метода наименьших квадратов*:

$$b = \frac{\overline{y \cdot x} - \bar{y} \cdot \bar{x}}{\overline{x^2} - \bar{x}^2},$$
$$a = \bar{y} - b \cdot \bar{x},$$

Параметр b - это коэффициент регрессии, характеризующий влияние, которое оказывает изменение X на Y . Он показывает, на сколько единиц в среднем изменится Y при изменении X на одну единицу. Если $b > 0$, то наблюдается положительная связь. Если $b < 0$, то увеличение X на единицу влечет за собой уменьшение Y в среднем на b .

На практике для количественной оценки тесноты связи широко используется *линейный коэффициент корреляции* (иногда его называют

просто коэффициентом корреляции). Линейный коэффициент корреляции вычисляется по формуле:

$$r_{xy} = b \cdot \frac{\sigma_x}{\sigma_y},$$

где b – коэффициент регрессии;

σ_x, σ_y – средние квадратические отклонения по переменным X на Y соответственно.

Причем $-1 \leq r_{xy} \leq 1$. Чем ближе значение коэффициента корреляции по модулю к единице, тем связь между изучаемыми показателями теснее.

Можно использовать и другие формулы, но результат должен быть одинаковым для всех вариантов расчета.

Количественные критерии оценки тесноты связи можно записать в виде таблицы 7.

Таблица 7 - Количественные критерии оценки тесноты связи	
Величина коэффициента корреляции r_{xy}	Теснота связи между x и y
до $ \pm 0,3 $	практически отсутствует
$ \pm 0,3 - \pm 0,5 $	слабая
$ \pm 0,5 - \pm 0,7 $	умеренная
$ \pm 0,7 - \pm 1,0 $	сильная

Квадрат r_{xy} равен так называемому *коэффициенту детерминации* (D или R^2), который показывает, какая часть вариации зависимого признака объясняется включенным в модель фактором.

Статистически оценить существенность параметров и построенного уравнения регрессии можно с помощью коэффициента аппроксимации. Средняя ошибка аппроксимации рассчитывается по формуле:

$$\bar{A}_i = \frac{1}{n} \sum \left| \frac{y - \hat{y}_x}{y} \right| \cdot 100\% .$$

Модель считается подобранной достаточно хорошо, если средняя ошибка аппроксимации не превышает 8-10%.

Пример 5. Найти с помощью средств MS Excel линейную зависимость между объемом выпуска и затратами на производство. Рассчитать коэффициент корреляции, коэффициенты детерминации и аппроксимации.

Данные сведены в таблицу 8.

Таблица 8 - Данные затрат на производство и объем выпуска продукции

Номер наблюдения	Затраты на производство y ,	Объем выпуска x ,
		тыс. ед.

	тыс. руб.	
1	68,80	45,10
2	61,20	41,30
3	59,90	38,70
4	56,70	36,50
5	55,00	36,20
6	54,30	32,40
7	49,30	28,10
Итого	405,20	258,30

1. Построение уравнения линейной регрессии

Так как для линейной регрессии вида $\hat{y}_x = b \cdot x + a$ параметры можно оценить с помощью формул:

$$b = \frac{\overline{y \cdot x} - \bar{y} \cdot \bar{x}}{\overline{x^2} - \bar{x}^2}, \quad (1)$$

$$a = \bar{y} - b \cdot \bar{x},$$

то расчетная таблица создается в виде, показанном на рисунке 1. На этом рисунке для расчета параметров уравнения регрессии используется диапазон (A1:E14).

Запустив электронные таблицы **MS Excel**, заполним столбец \hat{A} значениями величины y , а столбец C - значениями x (рис. 1). Заголовок столбца \hat{A} , где должно быть x^2 , отметим как $x^{\wedge}2$.

Все значения в ячейках рабочей таблицы, начиная со столбца D , рассчитаны с помощью формул. Например, в ячейку $D2$ введена формула $=B2 * C2$. Остальные клетки этого столбца, до $D8$ включительно, заполняются формулами автоматически. С этой целью необходимо сделать активной ячейку $D2$. Затем подвести указатель мыши к нижнему правому углу. Когда указатель мыши с белого крестика изменится на черный, удерживая нажатой левую кнопку мыши, скопировать формулу в следующие ячейки этого столбца (методом протягивания). То есть, в диапазоне ($D2:\hat{A}8$) все формулы вводятся именно таким образом – копированием в столбце.

	A	B	C	D	E
1		y	x	xy	x²
2	1	68,8	45,1	=B2*C2	=C2^2
3	2	61,2	41,3	=B3*C3	=C3^2
4	3	59,9	38,7	=B4*C4	=C4^2
5	4	56,7	36,5	=B5*C5	=C5^2
6	5	55	36,2	=B6*C6	=C6^2
7	6	54,3	32,4	=B7*C7	=C7^2
8	7	49,3	28,1	=B8*C8	=C8^2
9	Итого	=СУММ(B2:B8)	=СУММ(C2:C8)	=СУММ(D2:D8)	=СУММ(E2:E8)
10	Среднее значение	=СРЗНАЧ(B2:B8)	=СРЗНАЧ(C2:C8)	=СРЗНАЧ(D2:D8)	=СРЗНАЧ(E2:E8)
11	Среднее квадратическое отклонение	=СТАНДОТКЛОНП(B2:B8)	=СТАНДОТКЛОНП(C2:C8)	=СТАНДОТКЛОНП(D2:D8)	=СТАНДОТКЛОНП(E2:E8)
12					
13				b= =(D10-C10*B10)/(E10-C10^2)	
14				a= =B10-D13*C10	
15				r= =D13*C11/B11	
16				R^2= =D15^2	

Рисунок 1 - Исходные данные и используемые формулы для решения задачи

В диапазоне (B9:A11) копирование формул выполняется по строкам. Для этого достаточно ввести формулу в ячейку A9 и методом протягивания распространить формулу в восьмой строке до столбца A включительно.

Напомним, что функции СУММ, СРЗНАЧ, СТАНДОТКЛОНП можно ввести, выполнив команду основного меню: Вставка – Функция. В появившемся диалоговом окне Мастер Функции, в категории Статистические выбрать нужную функцию (рис. 2).

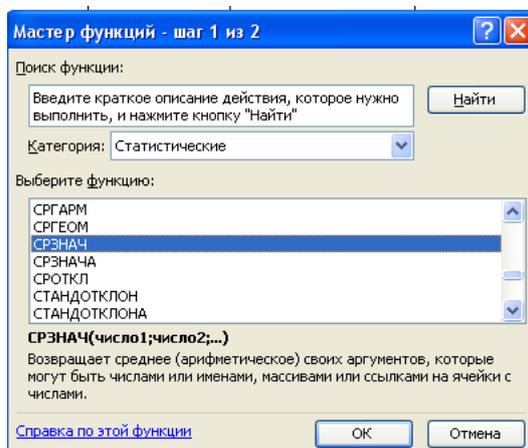


Рисунок 2 - Мастер функций

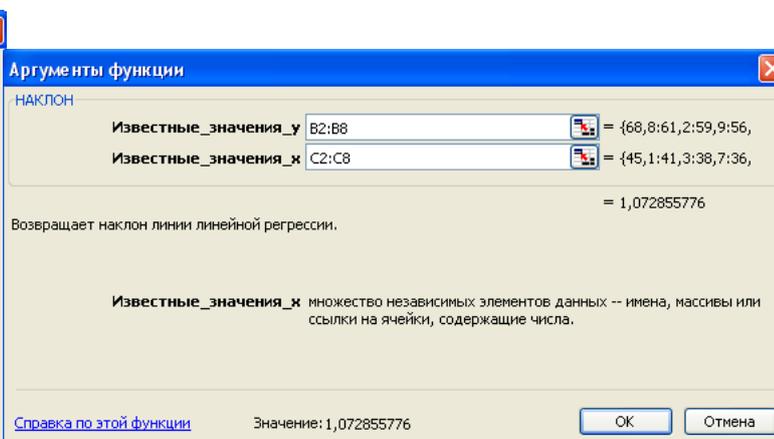


Рисунок 3 - Окно функции НАКЛОН

Далее в ячейки D13, D14 записываем формулы (1) для вычисления параметров уравнения.

Итак, имеем уравнение регрессии:

$$\hat{y}_x = 1,07 \cdot x + 18,30,$$

т.е. при увеличении объема выпускаемой продукции на 1 тыс. ед. затраты на производство возрастут на 1070 руб.

В отличие от коэффициента b , параметр a в данном уравнении – это значение результата y при факторе $x = 0$. Если признак-фактор x не имеет

или не может иметь нулевого значения, то такая трактовка свободного члена бессмысленна. У него нет экономического содержания.

2. Вычисление коэффициента корреляции, коэффициента детерминации

Уравнение регрессии всегда дополняется показателем тесноты связи. Поэтому в ячейку D15 внесем формулу $=D13*C11/B11$ для вычисления линейного коэффициента корреляции r_{xy} . На языке математики она выглядит

$$\text{как: } r_{xy} = b \cdot \frac{\sigma_x}{\sigma_y}.$$

В данном случае $r_{xy} = 0,97$ (рис. 5). Следовательно, имеется сильная связь между результатом и фактором.

Подсчитать коэффициент корреляции в MS Excel можно и при помощи функции КОРРЕЛ, в окнах панели которой необходимо ввести адреса диапазонов, как ряда коэффициента x , так и ряда y (рис. 6).

Для оценки качества подбора линейного уравнения регрессии находят коэффициент детерминации $R^2 = (r_{xy})^2$. Он отражает долю вариации результативного признака, объясненную с помощью уравнения регрессии (долю дисперсии результата, объясненную регрессией), в общей дисперсии y . Для рассмотренного примера $R^2 = 0,94$ (рис. 5). Следовательно, уравнением регрессии объясняется 94% дисперсии результативного признака, а прочими, неучтенными в модели факторами – 6%.

	A	B	C	D	E
1		y	x	xy	x^2
2	1	68,80	45,10	3102,88	2034,01
3	2	61,20	41,30	2527,56	1705,69
4	3	59,90	38,70	2318,13	1497,69
5	4	56,70	36,50	2069,55	1332,25
6	5	55,00	36,20	1991	1310,44
7	6	54,30	32,40	1759,32	1049,76
8	7	49,30	28,10	1385,33	789,61
9	Итого	405,2	258,3	15153,77	9719,45
10	Среднее значение	57,89	36,90	2164,82	1388,49
11	Среднее квадратическое	5,74	5,18	513,58	380,20
12					
13			b=	1,07	
14			a=	18,30	
15			r=	0,97	
16			R^2=	0,94	

Рисунок 5 - Результат первой части решения задачи

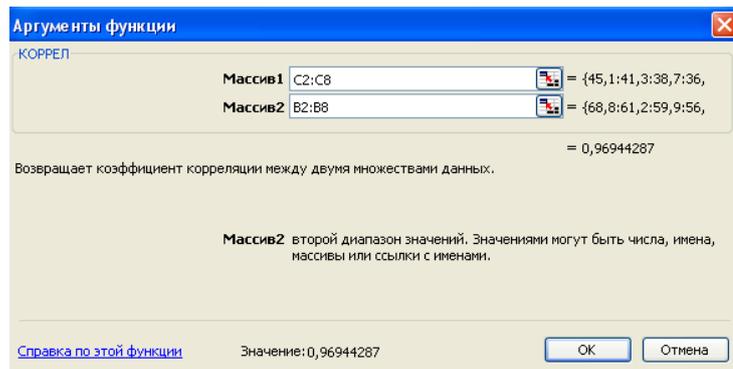


Рисунок 6 - Окно функции КОРРЕЛ

ЗАДАНИЕ 5

Найти линейную зависимость между X и Y . Рассчитать коэффициент корреляции, коэффициенты детерминации. Сделать выводы.

Исходные данные

	Вариант 1									
Y	22	24	18	17	25	28	23	26	32	35
X	17	31	14	14	28	28	32	23	27	33
	Вариант 2									
Y	25	28	27	31	26	34	29	32	36	32
X	26	30	31	28	27	35	36	38	40	44
	Вариант 3									
Y	25	18	16	23	20	21	19	17	15	14
X	26	32	24	20	38	32	25	17	22	50
	Вариант 4									
Y	18	22	16	14	15	20	21	24	23	27
X	35	32	27	26	44	21	23	14	20	32
	Вариант 5									
Y	12	15	13	14	17	19	22	20	24	26
X	32	33	28	27	31	30	26	22	21	20
	Вариант 6									
Y	22	25	18	12	26	30	40	25	32	28
X	23	29	21	22	36	22	30	19	26	42
	Вариант 7									
Y	30	26	25	28	24	27	22	329	18	21
X	18	27	16	13	31	25	38	16	28	30
	Вариант 8									
Y	40	36	38	32	26	28	30	25	22	23
X	25	26	15	17	27	24	26	20	31	35
	Вариант 9									
Y	28	32	30	33	35	38	42	50	40	32
X	40	36	38	32	26	28	30	25	22	23
	Вариант 10									
Y	28	34	20	19	35	27	28	21	30	38
X	24	28	22	15	30	26	36	18	33	32
	Вариант 11									
Y	28	34	20	19	35	27	28	21	30	38
X	32	36	23	24	40	18	22	15	21	40
	Вариант 12									
Y	18	27	16	13	31	25	38	16	28	30

X	35	32	27	26	44	21	23	14	20	32
				Вариант 13						
Y	17	31	14	14	28	28	32	23	27	33
X	28	32	30	33	35	38	42	50	40	32
				Вариант 14						
Y	32	36	23	24	40	18	22	15	21	40
X	22	24	18	17	25	28	23	26	32	35
				Вариант 15						
Y	26	30	31	28	27	35	36	38	40	44
X	18	22	16	14	15	20	21	24	23	27

Практическое занятие 3. Выборочное наблюдение. Понятие о выборочном наблюдении

Теоретическая часть

Выборочное наблюдение представляет собой такой вид несплошного наблюдения, при котором обследованию подвергается часть единиц исследуемой совокупности, отобранная на основе специально разработанных научных принципов, позволяющих по отобранной части совокупности получить данные для характеристики всей совокупности в целом.

Вся изучаемая совокупность явлений, из которой производится отбор некоторой части единиц для выборочного наблюдения, называется **генеральной совокупностью**. Та же часть единиц, которая отобрана из генеральной совокупности для выборочного наблюдения, называется **выборочной совокупностью**.

Средняя или относительная величина (доля) признака в генеральной совокупности называются **генеральными**, а средняя или относительная величина (доля) признака в выборочной совокупности называются **выборочными**.

\bar{X} – генеральная средняя

\bar{x} – выборочная средняя,

P – относительная величина (доля) признака в генеральной совокупности,

W – относительная величина (доля) признака в выборочной совокупности.

По сравнению с генеральной совокупностью характеристики выборочной совокупности могут иметь некоторые неточности. Подобные неточности выборочного наблюдения представляют собой ошибки репрезентативности.

Ошибками репрезентативности называют расхождения между средними величинами или долями признака выборочной и генеральной совокупностей. Ошибки репрезентативности могут быть систематическими и случайными.

Систематическими называют ошибки репрезентативности, которые возникают из-за нарушения научного принципа отбора единиц в выборочную совокупность.

Случайные ошибки репрезентативности – это неточности, которые возникают из-за того, что выборочная совокупность не совсем правильно воспроизводит структуру генеральной совокупности.

Задания для выполнения на практическом занятии

Задача 1 При проверке импортирования груза на таможне методом случайной выборки было обработано 200 изделий. В результате был установлен средний вес изделия 30г., при СКО=4г с вероятностью 0,997. Определите пределы в которых находится средний вес изделий генеральной совокупности.

Задача 2

В городе проживает 250тыс. семей. Для определения среднего числа детей в семье была организована 2%-я бесповторная выборка семей. По ее результатам было получено следующее распространение семей по числу детей:

$P=0,954$. Найти пределы в которых будет находиться среднее число детей в генеральной совокупности.

Число детей в семье, x_i	0	1	2	3	4	5
Кол-во детей в семье	1000	2000	1200	400	200	200