МИНИСТЕРСТВО СЕЛЬСКОГО ХОЗЯЙСТВА
РОССИЙСКОЙ ФЕДЕРАЦИИ
ФГБОУ ВО «Кубанский государственный
аграрный университет имени И.Т. Трубилина»

Н. Х. Ворокова,
А.Е. Жминько, А. Е. Сенникова

Статистика

Методические рекомендации
для контактной и самостоятельной работы обучающихся
по направлению подготовки
05.03.06 Экология и природопользование

Краснодар
КубГАУ
2021

УДК 519.22
ББК 22.172

С о с т а в и т е л и :
Н. Х. Ворокова, А. Е. Жминько, А. Е. Сенникова

Статистика: метод. рекомендации для контактной и самостоятельной работы обучающихся по направлению подготовки 05.03.06 «Экология и природопользование» / сост. Н. Х. Ворокова, А. Е. Жминько, А. Е. Сенникова. – Краснодар: КубГАУ, 2021. – 83 с.

В методических рекомендациях с представлены теоретические аспекты, рекомендуемая литература для изучения дисциплины, контрольные вопросы, задачи для самостоятельного решения, тесты.

Предназначены для обучающихся по направлению подготовки 05.03.06 «Экология и природопользование». Соответствует ФГОС ВО.

УДК 519.22
ББК 22.172

Оглавление

Введение

Одним из основных подходов к обоснованию и последующему принятию решений является статистический, основанный на использовании статистических методов и приемов анализа.

Статистические методы обработки данных можно разделить на следующие группы.

а) По способу получения экспериментальных данных:

- активный эксперимент;
- пассивный эксперимент (выборочное или сплошное наблюдение).

б) По цели обработки данных:

- описательные (получение и сравнение числовых характеристик экспериментальных данных) – анализ вариационных рядов, выборочный метод, проверка статистических гипотез и другие;
- аналитические (количественная оценка и анализ зависимостей, описывающих изучаемые объекты (процессы) – дисперсионный анализ, регрессионный анализ, анализ рядов динамики и другие).

Цель учебного пособия – оказать помощь обучающимся в овладении приемами и методами статистико-математического исследования; в закреплении теоретических знаний, полученных на лекциях и при самостоятельной работе во внеучебное время.

Значительная часть задач учебного пособия составлена на основе фактических данных Краснодарстата и сельскохозяйственных организаций Краснодарского края.

Обучающийся, на основании изучения рекомендуемой литературы, самостоятельно выполняет задания по темам в соответствии с индивидуальным вариантом. Для облегчения выполнения самостоятельного задания, по всем темам изложены необходимые краткие методические указания и приводится решение типовых задач.

Тема 1 Абсолютные и относительные статистические величины

Абсолютные и относительные величины выступают как ключевые элементы в статистической науке. Они используются для определения количественных характеристик, динамики их изменения. Абсолютные и относительные величины отражают разные характеристики, но без одних не могут существовать другие.

Абсолютными статистическими величинами называются показатели, выражающие размеры, объемы и уровни общественных явлений и процессов.

Относительными величинами называются обобщающие показатели, характеризующие количественные соотношения сопоставляемых статистических величин.

Относительные величины структуры характеризуют состав изучаемой совокупности и показывают, какой удельный вес (какую долю) в общем итоге составляет каждая ее часть. Они получаются в результате деления значения каждой части совокупности на их общий итог.

Относительные величины координации характеризуют соотношение отдельных частей целого, одна из которых принимается за базу сравнения. К таким показателям относится число сельских жителей на 100 городских, число женщин на 100 мужчин, площадь посева технических культур на 100 га зерновых и т.п.

Относительная величина выполнения плана выражает степень выполнения планового задания за определенный период времени и исчисляется как отношение фактически достигнутого уровня ($У_1$) к плановому заданию ($У_п$):

$$К_{в.п.} = \frac{У_1}{У_п}. \tag{1.1}$$

Относительная величина планового задания показывает степень напряженности плана по сравнению с базисным периодом и определяется как отношение планового уровня на предстоящий период к фактически достигнутому уровню за предшествующий период ($У_о$):

$$К_{п.з} = \frac{У_п}{У_0}. \tag{1.2}$$

Относительная величина динамики характеризует изменение одноименного явления во времени, получается в результате сопоставления фактического уровня в текущем периоде с базисным:

$$К_д = \frac{У_1}{У_0}. \qquad (1.3)$$

Относительные величины динамики, планового задания и выполнения плана взаимосвязаны:

$$К_д = К_{в.п.} \cdot К_{п.з}. \qquad (1.4)$$

Относительные величины интенсивности показывают степень распространения данного явления в определенной среде.

Относительные величины сравнения представляют собой соотношение одноименных показателей, относящихся к различным объектам или территориям, но за один и тот же период или момент времени.

Пример 1.1 На основании имеющихся данных изучить структуру посевных площадей сельскохозяйственных культур и рассчитать различные виды относительных величин. Структуру посевных площадей изобразить графически. Сделать вывод.

Решение. Построим вспомогательную таблицу 1.1 и проведем расчет показателей.

Таблица 1.1 – Динамика и структура посевных площадей

Культура	Площадь посева, га			Структура, %			Относительные величины		
	2018г	План на 2019г	2019г	2018г	План на 2019г	2019г	динамики	выполнения плана	планового задания
Озимая пшеница	1265	1300	1350	26,1	27,6	27,7	1,067	1,038	1,028
Кукуруза	459	320	350	9,5	6,8	7,2	0,763	1,094	0,697
Овощи открытого грунта	500	510	480	10,3	10,8	9,8	0,960	0,941	1,020
Овес	1300	1280	1300	26,8	27,2	26,6	1,000	1,016	0,985
Подсолнечник	1320	1300	1400	27,3	27,6	28,7	1,061	1,077	0,985
Итого	4844	4710	4880	100,0	100,0	100,0	1,007	1,036	0,972

Для наглядного представления статистических данных, структуру посевных площадей изобразим графически (рисунок 1).

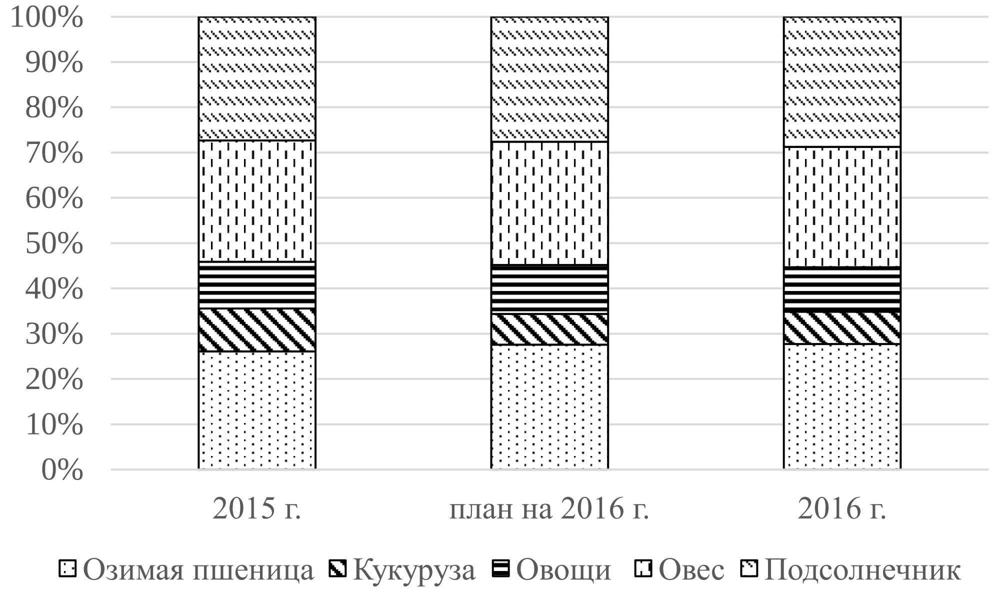

Рисунок 1.1 – Структура посевных площадей, %

Вывод. В отчетном году (2019 г.) по сравнению с базисным (2018 г.) площадь посева озимой пшеницы увеличилась на 6,7 % или на 85 га. При этом план посева перевыполнен на 3,8 %, что составляет 50 га. Посевы кукурузы в отчетном году сократились на 23,7 % (109 га), но план был перевыполнен на 9,4 % (30 га). Площадь посадки овощей открытого грунта в отчетном году снизилась на 4,0 % или на 20 га, план посадки не довыполнен на 5,9 % или 30 га. Площадь посева овса за два года не изменилась, план посева перевыполнен на 1,6 %, что составляет 20 га. Посевы подсолнечника в отчетном году по сравнению с базисным выросли на 6,1 % (80 га), план посева перевыполнен на 7,7 % (100 га). В целом отмечается расширение посевных площадей по сравнению с базисным периодом на 0,7 %, что составляет 36 га. План посева перевыполнен на 3,6 % (170 га).

Задачи для самостоятельного решения

Задача 1.1 По данным приложения Д определить относительные величины структуры и динамики посевных площадей. Структуру площадей изобразить графически. Сделать выводы.

Задача 1.2 По данным приложения Д по одному варианту рассчитать относительные величины динамики, выполнения плана и планового задания. Сделать выводы.

Задача 1.3 В организации объем производства зерна в 2019 г. составил 194137 ц. По плану в 2017 г. предусматривалось увеличить объем производства зерна на 7,5 % по сравнению с 2019 г. Плановое задание было перевыполнено в 2017 г. на 5,8 %. Определить объем производства в 2017 г. по плану и фактически, рассчитать коэффициент динамики. Сделать вывод.

Задача 1.4 По имеющимся данным о внесении минеральных удобрений на 1 га посева сельскохозяйственных культур в сельскохозяйственных организациях Краснодарского края за 2017–2019 гг. рассчитать относительные величины сравнения, приняв за базу средний уровень показателя по Краснодарскому краю. Сделать вывод.

Таблица 1.2 – Внесение минеральных удобрений

Район	Внесено удобрений, кг		
	2017 г.	2018 г.	2019 г.
Брюховецкий	97	113	128
Выселковский	134	131	134
Динской	87	89	99
Кавказский	128	168	149
Тимашевский	109	108	128
Усть-Лабинский	83	82	88
В среднем по краю	109	109	117

Задача 1.5 Себестоимость производства 1 ц зерна озимой пшеницы в 2018 г. составила 448 руб. На 2019 г. планировалось снизить себестоимость на 4,8 %, при этом фактическая себестоимость в 2019 г. по сравнению с предыдущим годом выросла на 1,5 %. Рассчитать фактический, плановый уровни себестоимости в 2019 г. и коэффициент выполнения плана. Сделать вывод.

Задача 1.6. На основании данных таблицы определить различные относительные величины.

Таблица 1.3 – Численность населения региона (на начало года)

Показатель	2018 г.	2019 г.
Среднегодовая численность населения, тыс. чел.	5513,8	5570,9
в том числе городского	2994,9	3041,9
сельского	2518,9	2529,0
Число родившихся, чел.	74117	72986
в том числе городского	45344	45538
сельского	28773	27448
Число умерших, чел.	71378	71550
в том числе городского	38909	39190
сельского	32469	32360

Контрольные тесты для проверки знаний

1 Абсолютные величины могут быть:
   а) базисные;
   б) индивидуальные;
   в) суммарные;
   г) агрегатные.

2 Натуральные единицы измерения используются для характеристики показателей
   а) явлений в свойственной для них форме;
   б) в денежном выражении;
   в) использования трудовых ресурсов;
   г) относительных.

3 Абсолютные величины могут выражаться в ... .
   а) натуральных единицах измерения;
   б) процентах
   в) трудовых единицах измерения;
   г) промилле.

4 Относительной статистической величиной называют показатель, выражающий
   а) количественные соотношения различных явлений;
   б) размеры общественных явлений;
   в) развитие явления в пространстве;
   г) объем общественных явлений.

5 Относительная величина планового задания рассчитывается как отношение следующих уровней
   а) текущего к базисному;
   б) планового к базисному;
   в) запланированного к фактическому;
   г) фактического к установленному плану.

6 Относительная величина выполнения плана рассчитывается как отношение следующих уровней
   а) фактического к установленному плану;
   б) базисного к плановому;
   в) запланированного к фактическому;
   г) планового к базисному.

7 Относительная величина структуры характеризует динамику явления
   а) соотношение частей явления;
   б) состав изучаемого явления;
   в) развитие явления в пространстве;
   г) объем явления.

8 Относительная величина координации характеризует состав явления
   а) развитие явления во времени
   б) соотношение частей явления;
   в) развитие явления в пространстве;
   г) объем явления.

9 Существует взаимосвязь между относительными величинами
   а) $К_{пл.з.} = К_{д} \cdot К_{в.пл.}$
   б) $К_{д} = К_{в.пл.} + К_{пл.з.}$
   в) $К_{в.пл.} = К_{д} \cdot К_{пл.з.}$

г) $К_{д.} = К_{пл.з.} \cdot К_{в.пл.}$

10 Относительные величины сравнения характеризуют соотношение различных явлений или процессов
- а) одноименных величин, относящихся к разным объектам;
- б) частей изучаемого явления;
- в) разноименных абсолютных величин;
- г) фактического уровня к запланированному.

11 Относительная величина интенсивности характеризует состав явления
- а) соотношение частей явления;
- б) развитие явления во времени;
- в) соотношение между разноименными величинами;
- г) развитие явления в определенной среде.

12 В относительных величинах используют различные базы сравнения. Если база сравнения 1 – получают
- а) проценты;
- б) продецемилле;
- в) коэффициенты;
- г) промилле.

13 В относительных величинах используют различные базы сравнения. Если база сравнения 100 – получают
- а) проценты;
- б) продецемилле;
- в) коэффициенты;
- г) промилле.

14 В относительных величинах используют различные базы сравнения. Если база сравнения 1000 – получают
- а) проценты;
- б) продецемилле;
- в) коэффициенты;
- г) промилле.

15 Относительные величины динамики подразделяются на
- а) цепные;
- б) однородные;
- в) разнородные;
- г) базисные.

Вопросы для самоподготовки

1. Дайте определение абсолютных величин, назовите их виды и единицы измерения.
2. Назовите виды относительных величин.
3. Приведите примеры взаимосвязи относительных величин.
4. Назовите единицы измерения относительных величин.
5. Приведите примеры относительных величин, характеризующих наличие и использование ресурсов в организации.
6. Опишите способы расчета различных видов относительных величин.

Тема 2 Вариационные ряды

Вариационные ряды строятся по количественным признакам. По способу построения они бывают дискретными (прерывными) и непрерывными. Дискретный ряд распределения основан на прерывной вариации, при которой значения признака выражены целыми числами (тарифный разряд рабочих, число зерен в початке, число раскрытых преступлений и т. д.). Если признак непрерывный, т. е. на определенном промежутке может принимать любое значение, или если число значений дискретного признака велико, то строится интервальный ряд распределения.

Вариационные ряды состоят из двух элементов: вариант и частот. Варианта (x_i) – это отдельное значение варьирующего признака, которое он принимает в ряду распределения. Частота (n_i) – это численность отдельных вариант или каждой группы вариационного ряда.

Характеристиками вариационного ряда являются его наибольшее ($x_{\max}$), наименьшее ($x_{\min}$) значения и размах вариации (R):

$$R = x_{\max} - x_{\min}\ . \tag{2.1}$$

Сумма всех частот называется объемом вариационного ряда (n):

$$\sum_{i=1}^{k} n_i = n_1 + n_2 + \dots n_k = n. \tag{2.2}$$

Отношение частоты данного варианта к объему совокупности называется *относительной частотой* ($\hat{p}_i$) или *частостью* этого варианта:

$$\hat{p}_i = \frac{n_i}{\sum_{i=1}^{k} n_i} = \frac{n_i}{n}, \quad \Sigma \hat{p}_i = 1. \tag{2.3}$$

Если признак дискретный, то вариационный ряд представляет упорядоченную совокупность значений признака и соответствующих им частот или частостей.

Число равных по длине интервалов (k), на которое разбивается совокупность, если она нормально распределена, можно определить по формуле Стерджесса

$$k=1+3{,}322lgn. \tag{2.4}$$

Величина интервала определяется по формуле

$$h = \frac{x_{\max} - x_{\min}}{k} \tag{2.5}$$

Вариационные ряды позволяют получить первое представление об изучаемом распределении. Далее необходимо определить числовые характеристики: положения (средняя арифметическая, мода, медиана); рассеяния (дисперсия, среднее квадратическое отклонение, коэффициент вариации); меры скошенности (коэффициент асимметрии) и островершинности (эксцесс) распределения.

Средней арифметической $(\bar{x})$ вариационного ряда называется отношение суммы произведений вариант на соответствующие частоты к объему совокупности:

$$\bar{x} = \frac{\sum x_i n_i}{\sum n_i} = \frac{\sum x_i n_i}{n}. \tag{2.6}$$

Часто применяется формула простой средней арифметической:

$$\bar{x} = \frac{\Sigma x_i}{n}. \tag{2.7}$$

Модой (M_0) дискретного вариационного ряда называется вариант, имеющий наибольшую частоту. Ряды могут быть одно- и многомодальными.

Медианой (M_e) дискретного вариационного ряда называется вариант, делящий ряд на две равные части.

Если дискретный вариационный ряд имеет *2n* членов:

$x_1,\ x_2, \ldots, x_n,\ x_{n+1}, \ldots,\ x_{2n}$, то

$$M_e = \frac{x_n + x_{n+1}}{2}. \tag{2.8}$$

Если же дискретный вариационный ряд имеет (*2n+1*) членов:

$x_1, x_2, \ldots,\ x_{n-1}, x_n, x_{n+1}, \ldots,\ x_{2n+1}$,

Тогда $M_e = x_{n+1}$.

В интервальных вариационных рядах мода и медиана находятся по формулам:

$$M_0 = x_{Mo} + h\frac{n_{Mo} - n_{Mo-1}}{(n_{Mo} - n_{Mo-1}) + (n_{Mo} - n_{Mo+1})}, \tag{2.9}$$

где x_{Mo} – начало модального интервала;

h – длина (величина) модального интервала;

n_{Mo} – частота модального интервала;

n_{Mo-1} – частота предмодального интервала;

n_{Mo+1} – частота послемодального интервала;

$$M_e = x_{Me} + h \cdot \frac{0{,}5n - S_{Me-1}}{n_{Me}}, \tag{2.10}$$

где x_{Me} – начало медианного интервала;

h – длина медианного интервала;

n – объем совокупности;

S_{Me-1} – накопленная частота интервала, предшествующего медианному;

n_{Me} – частота медианного интервала.

При расчете средней арифметической в интервальном ряду в качестве вариант x_i принимаются середины соответствующих интервалов.

Мода и медиана используются в качестве характеристик среднего положения в случае, если границы ряда нечеткие или если ряд несимметричен.

Дисперсия ряда распределения определяется по формулам:

$$\sigma^2 = \frac{\sum (x_i - \bar{x})^2}{n}, \qquad \sigma^2 = \frac{\sum (x_i - \bar{x})^2 n_i}{n} \qquad (2.11)$$

и характеризует средний квадрат отклонения x_i от $\bar{x}$.

Среднее квадратическое отклонение вариационного ряда

$$\sigma = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n}}, \quad \sigma = \sqrt{\frac{\sum (x_i - \bar{x})^2 n_i}{n}} \qquad (2.12)$$

выражается в тех же единицах, что и x_i.

Коэффициент вариации характеризует относительную колеблемость изучаемого признака и обычно служит для сравнения вариации разных показателей по одной и той же совокупности, или вариацию одного показателя по разным совокупностям:

$$V = \frac{\sigma}{\bar{x}} \cdot 100 \ \%. \qquad (2.13)$$

Пример 2.1 По списку в организации числится 110 рабочих, которые имеют следующие разряды: 3, 5, 6, 4, 3, 4, 6, 4, 5, 3, 2, 2, 3, 4, 5, 3, 4, 5, 4, 1, 4, 5, 5, 4, 3, 4, 6, 4, 2, 4, 4, 4, 3, 5, 6, 4, 3, 3, 2, 3, 4, 3, 1, 2, 4, 4, 5, 6, 1, 3, 4, 5, 3, 4, 4, 3, 2, 6, 1, 2, 4, 5, 3, 3, 2, 3, 6, 4, 3, 4, 5, 4, 3, 3, 2, 6, 3, 3, 4, 5, 4, 4, 3, 3, 2, 1, 2, 1, 6, 5, 4, 3, 2, 3, 4, 4, 3, 5, 6, 1, 5, 6, 4, 3, 4, 5, 6, 4, 3, 5.

Составить ряд распределения рабочих по разрядам. Найти накопленные частоты и частости. Определить средний разряд рабочего, модальный и медианный разряд, дисперсию и среднее квадратическое отклонение. Вариационный ряд изобразить графически.

Решение. Подсчитаем число рабочих, имеющих определенный разряд, и запишем в таблицу 2.1. Определим накопленные частоты и частости.

Таблица 2.1 – Вспомогательная таблица для расчета показателей дискретного вариационного ряда

Разряд рабочего	Число рабочих n_i	Накопленное число рабочих S_i	Относительная частота $\hat{p}_i$	xn	$\lvert x - \bar{x} \rvert n$	$(x - \bar{x})^2 n$
1	7	7	0,064	7	18,9	51,0
2	12	19	0,109	24	20,4	34,7
3	29	48	0,264	87	20,3	14,2
4	33	81	0,300	132	9,9	3,0
5	17	98	0,154	85	22,1	28,7
6	12	110	0,109	72	27,6	63,5
Сумма	110	-	1,000	407	119,2	195,1

Дискретный ряд распределения можно изобразить графически в виде полигона частот или частостей, а также кумуляты. В этом случае по оси абсцисс откладываются значения признака, а по оси ординат – соответствующие им частоты, частости или накопленные частоты. Полученные точки соединяются отрезками. Полигон и кумулята распределения рабочих по разрядам изображены на рисунках 2.1 и 2.2.

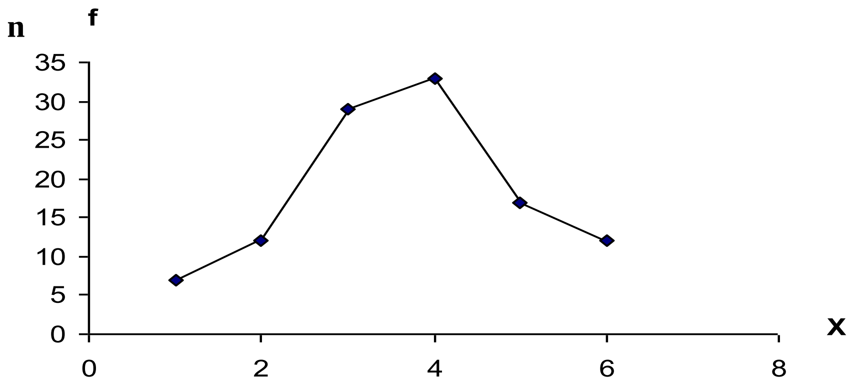

Рисунок 2.1 – Полигон распределения рабочих по разрядам

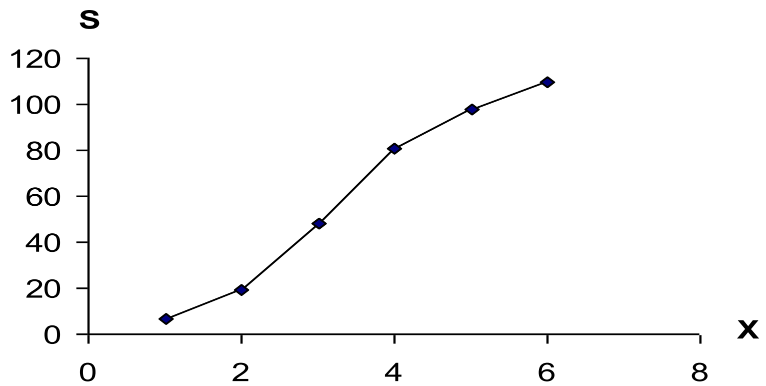

Рисунок 2.2 – Кумулята распределения рабочих по разрядам

Средний разряд рабочих определим по формуле средней арифметической взвешенной (2.6):

$$\bar{x} = \frac{407}{110} = 3{,}7$$

Наибольшее число рабочих имеет четвертый разряд, значит Mo = 4. Так как всего в организации 110 рабочих (четное число), то медиана равна средней арифметической из разрядов 55 и 56 рабочего в ранжированном ряду, т.е. четвертому разряду: Me = 4.

Определим среднее линейное отклонение:

$$L = \frac{\sum|x-\bar{x}|n}{\sum n} = \frac{119,2}{110} = 1,08.$$

Далее определим дисперсию по формуле 2.11:

$$\sigma^2 = \frac{\sum(x-\bar{x})^2 n}{\sum n} = \frac{195,1}{110} = 1,774.$$

Среднее квадратическое отклонение определяется как квадратный корень из дисперсии:

$$\sigma = \sqrt{1,774} = 1,33 .$$

Коэффициент вариации:

$$V_L = \frac{L}{\bar{x}} \cdot 100 = \frac{1,08}{3,7} = 29,2 \text{ \%},$$

$$V_\sigma = \frac{\sigma}{\bar{x}} \cdot 100 = \frac{1,33}{3,7} = 35,9 \text{ \%}.$$

Вывод. В организации наиболее часто встречается четвертый разряд рабочего. Половина рабочих имеет разряд до четвертого, а другая половина – четвертый и выше. Средний разряд рабочего по организации составляет 3,7. Разряд рабочих в среднем варьирует в границах от 2,4 до 5, а с учетом округления результатов – от 2 до 5. Коэффициент вариации показывает, что имеются довольно большие различия в квалификации рабочих.

Пример 2.2 По 46 сельскохозяйственным организациям Краснодарского края за 2014 г. имеются следующие данные об урожайности кукурузы на зерно (ц/га): 44,0; 37,1; 24,8; 37,9; 51,5; 52,5; 50,3; 47,5; 30,7; 39,0; 56,9; 62,3; 51,9; 53,9; 46,6; 32,0; 50,7; 50,5; 37,4; 54,4; 47,5; 52,1; 48,4; 50,0; 28,5; 57,8; 33,8; 24,4; 48,6; 47,5; 21,6; 38,9; 52,3; 54,4; 37,1; 36,5; 47,2; 47,9; 22,5; 43,0; 29,1; 53,7; 25,0; 30,5; 28,5; 38,6.

Составить вариационный ряд с равными интервалами. Найти накопленные частоты. Вариационный ряд изобразить графически. Определить среднюю урожайность кукурузы на зерно, модальное и медианное значения, а также показатели вариации.

Решение. По формуле 2.4 найдем число групп, на которое необходимо разбить вариационный ряд:

$$k = 1 + 3,322 \lg 46 = 6,52 .$$

Учитывая небольшой объем вариационного ряда, примем k = 6. По формуле 2.5 определим величину интервала:

$$h = \frac{62,3 - 21,6}{6} = 6,8 .$$

Границы интервалов составят: 21,6 – 28,4; 28,4 – 35,2; 35,2 – 42,0; 42,0 – 48,8; 48,8 – 55,6; 55,6 – 62,4.

Подсчитав число организаций в каждой группе, получим вариационный ряд. Все промежуточные расчеты проведем в таблице 2.2.

Интервальный вариационный ряд изображается графически с помощью гистограммы и кумуляты распределения. На оси абсцисс откладываются границы интервалов варьирующего признака, а на оси ординат – частоты. Каждому интервалу соответствует прямоугольник, по высоте равный частоте или частости (рисунок 2.3 и 2.4).

Таблица 2.2 – Вспомогательная таблица для расчета показателей вариационного ряда

Группа организаций по урожайности кукурузы на зерно, ц/га	Число организаций в группе (n_i)	Накопленное число организаций (S_i)	Среднее значение интервала (x_i)	xn	$\lvert x-\bar{x}\rvert n$	$(x-\bar{x})^2 n$
21,6–28,4	5	5	25,0	125,0	88,5	1566,45
28,4–35,2	7	12	31,8	222,6	76,3	831,67
35,2–42,0	8	20	38,6	308,8	32,8	134,48
42,0–48,8	10	30	45,4	454,0	27,0	72,90
48,8–55,6	13	43	52,2	678,6	123,5	1173,25
55,6–62,4	3	46	59,0	177,0	48,9	797,07
Итого:	46	-	-	1966,0	397,0	4575,81

Средняя урожайность кукурузы на зерно составит:

$$\bar{x} = \frac{1966}{46} = 42,7 \text{ ц/га.}$$

Найдем моду вариационного ряда, используя формулу 2.9:

$$Mo = 48,8 + 6,8\frac{13-10}{(13-10)+(13-3)} = 50,4 \text{ ц/га}.$$

Медиана определяется по формуле 2.10:

$$Me = 42,0 + 6,8\frac{\frac{46}{2}-20}{10} = 44,0 \text{ ц/га.}$$

Определим показатели вариации:

а) размах вариации: $R = 62,3 - 21,6 = 40,7$ ц/га;

б) среднее линейное отклонение: $L = \frac{\sum|x-\bar{x}|n}{\sum n} = \frac{397,0}{46}$=8,6 ц/га;

в) дисперсия: $\sigma^2 = \frac{4575,82}{46} = 99,474$;

г) среднее квадратическое отклонение: $\sigma = \sqrt{99,474} \approx 10,0$ ц/га;

д) коэффициент вариации:

$$V_L = \frac{L}{\bar{x}} \cdot 100 = \frac{8,6}{42,7} \cdot 100 = 20,1\ \%,$$

$$V_\sigma = \frac{\sigma}{\bar{x}} \cdot 100 = \frac{10,0}{42,7} \cdot 100 = 23,4\ \%,$$

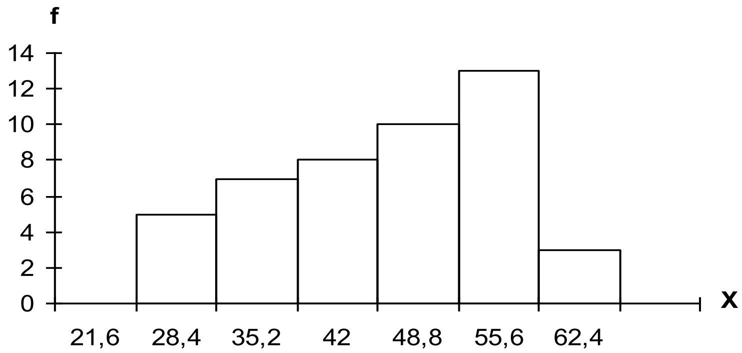

Рисунок 2.3 - Гистограмма распределения сельскохозяйственных организаций по урожайности кукурузы на зерно, ц/га

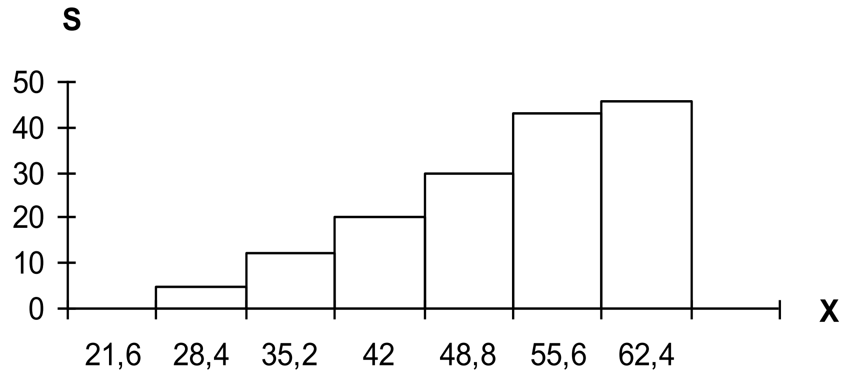

Рисунок 2.4 - Кумулята распределения сельскохозяйственных организаций по урожайности кукурузы на зерно, ц/га

Вывод. Расчеты показали, что в хозяйствах наиболее часто встречается урожайность кукурузы на зерно 50,4 ц/га *(Мо)*, половина организаций имеет урожайность кукурузы на зерно до 44,0 ц/га, а половина – выше *(Ме)*. Средняя урожайность кукурузы на зерно по организациям составила 42,7 ц/га. Согласно среднего линейного отклонения урожайность в организациях изменяется в пределах 42,7±8,6 ц/га, а согласно среднего квадратического отклонения, урожайность колебалась в среднем в границах (42,7 ± 10,0) ц/га, т.е. от 32,7 до 52,7 ц/га. Коэффициенты вариации, рассчитанные с учетом среднего линейного и среднего квадратического отклонения свидетельствуют о небольшой колеблемости урожайности кукурузы на зерно в хозяйствах края.

Задачи для самостоятельного решения

Задача 2.1 При сдаче экзамена студентами были получены следующие оценки:

Балл сдачи экзамена (x)	2	3	4	5
Число студентов (n)	4	13	5	3

Найти накопленные частоты и частости. Определить среднее, модальное и медианное значения, а также показатели вариации. Ряд распределения изобразить графически с помощью полигона и кумуляты распределения.

Задача 2.2 Распределение студентов факультета характеризуется следующими данными:

Возраст студентов, лет (x)	17	18	19	20	21
Число студентов (n)	5	50	70	10	25

Определить среднее, модальное и медианное значения возраста студентов, а также показатели вариации. Ряд распределения изобразить графически с помощью полигона и кумуляты распределения.

Задача 2.3 Рабочие 10 бригад распределяются следующим образом по разрядам:

Разряд рабочего	Число рабочих по бригадам									
	I	II	III	IV	V	VI	VII	VIII	IX	X
2	2	4	2	3	4	3	2	1	5	2
3	8	5	3	5	5	6	5	4	6	4
4	10	8	5	7	8	10	9	10	8	11
5	5	6	10	5	4	6	3	6	5	7
6	3	4	4	3	2	3	2	3	4	5

По одной бригаде найти накопленные частоты и частости. Определить модальный, медианный и средний разряд, среднее квадратическое отклонение и коэффициент вариации. Ряд распределения изобразить графически с помощью полигона и кумуляты распределения.

Задача 2.4 Составить интервальный ряд распределения организаций по урожайности озимой пшеницы (приложение Е). Найти накопленные частоты и частости. Определить среднее, модальное и медианное значения, а также показатели вариации. Ряд распределения изобразить графически с помощью гистограммы и кумуляты распределения.

Задача 2.5 Дан интервальный вариационный ряд распределения крестьянских хозяйств по площади посевов на одно хозяйство:

Группа хозяйств по площади посевов, га	22 – 26	26 – 30	30 – 34	34 – 38	38 – 42	42 – 46
Число хозяйств	5	8	16	11	10	5

Вариационный ряд изобразить графически. Определить: моду и медиану; среднюю площадь посевов на одно хозяйство; среднее квадратическое отклонение и коэффициент вариации.

Задача 2.6 Кукурузные початки распределились следующим образом по числу зерен. По одному варианту найти накопленные частоты и частости. Определить модальное, медианное и среднее число зерен, среднее квадратическое отклонение и коэффициент вариации. Ряд распределения изобразить графически.

Число зерен в початке, шт.	Число початков по вариантам, шт.				
	I	II	III	IV	V
820 – 850	6	11	8	14	2
850 – 880	2	4	12	9	6
880 – 910	13	2	3	4	3
910 – 940	7	8	5	7	15
940 – 970	3	12	10	11	7

Контрольные тесты для проверки знаний

1. Наиболее часто встречающееся значение признака у единиц данной совокупности называется …
   а) медианой;
   б) вариацией;
   в) модой;
   г) частостью.
2 Если все варианты значений признака уменьшить в 3 раза, то дисперсия …
   а) не изменится;
   б) уменьшится в 9 раз;
   в) уменьшится в 3 раза;
   г) увеличится в 3 раза.
3 Среднегодовые остатки оборотных средств предприятия, при наличии информации за каждый месяц, определяется по формуле средней …
   а) геометрической;
   б) гармонической;
   в) арифметической;
   г) хронологической.
4 Если коэффициент вариации составляет менее 33,3 %, то совокупность …
   а) неоднородная;
   б) средней однородности;
   в) умеренной однородности;
   г) однородная.
5 Уровень дохода, наиболее часто встречающийся в совокупности, называется _________ доходом.
   а) реальным
   б) модальным
   в) децильным
   г) медианным.
6 Если частоты всех значений признака умножить на 8 единиц, то средняя арифметическая величина …
   а) увеличится на 8 единиц;
   б) останется неизменной;
   в) увеличится в 8 раз;
   г) уменьшится в 8 раз.
7 Размахом вариации называется __________ максимального и минимального значений признака.
   а) разность;
   б) сумма;
   в) частное от деления;

г) произведение.

8 Вариацию признака по всей совокупности под влиянием всех факторов, обусловивших эту вариацию, измеряет …

а) общая дисперсия;
б) средняя из внутригрупповых дисперсий;
в) межгрупповая дисперсия;
г) внутригрупповая дисперсия.

9 Ряды распределения делят на:

а) вариационные;
б) динамические;
в) атрибутивные;
г) территориальные.

10 При непрерывной вариации признака целесообразно построить

а) дискретный вариационный ряд;
б) интервальный вариационный ряд;
в) ряд распределения;
г) невозможно построить.

11 Накопленные частоты используются при построении

а) полигона;
б) гистограммы;
в) кумуляты;
г) не используются.

12 Ряд, построенный по количественному признаку, называется

а) атрибутивным;
б) качественным;
в) группировочным;
г) вариационным.

13 Ряд, построенный по качественному признаку, называется

а) атрибутивным;
б) вариационным.
в) группировочным;
г) качественным;

14 Показатели структуры вариационного ряда

а) простая средняя арифметическая;
б) мода;
в) медиана;
г) дисперсия.

15 Ряд распределения может характеризоваться

а) двумя модами;
б) одной модой;
г) ни одной модой.

Вопросы для самоподготовки

1. Дайте определение вариационного ряда и назовите виды вариационных рядов.

2. Как определяется число групп и величина интервала при построении интервального вариационного ряда?

3. Способы графического изображения вариационных рядов.

4. Средняя арифметическая и ее свойства.

5. Абсолютные и относительные показатели вариации.

6. Дисперсия, ее свойства и способы расчета.

Тема 3 Выборочный метод

Сбор данных для статистического изучения явлений может проводиться сплошным и выборочным методами. При сплошном наблюдении обследуются все единицы изучаемой совокупности. При выборочном наблюдении отбирается часть единиц генеральной совокупности, а показатели, найденные по отобранной части единиц, должны достаточно точно характеризовать показатели всей совокупности единицы.

По процедуре отбора различают два вида отбора:

– повторный, при котором отобранная единица возвращается назад в генеральную совокупность и может попасть в выборку более чем один раз;

– бесповторный, когда каждая отобранная из совокупности единица один раз участвует в процессе отбора.

При проведении выборочного наблюдения возникают ошибки регистрации и ошибки репрезентативности (представительности). Ошибки репрезентативности – это расхождения между обобщающими характеристиками выборочной и генеральной совокупности, возникающие вследствие несплошного характера наблюдения. Желательно, чтобы величина ошибок была небольшой. Так как численное значение ошибки не известно, то ее возможная оценка дается с помощью расчета средней и предельной ошибок выборки. Обычно величина ошибок определяется для средней арифметической и для доли единиц, обладающих определенным признаком.

Предельная ошибка выборки находится как предел отклонения выборочной характеристики от генеральной, гарантируемой с заданной, обычно близкой к единице, вероятностью, называемой доверительной вероятностью.

Для средней арифметической предел отклонения имеет вид:

$$\lim_{n \to \infty} P\left(\left|\bar{x} - \tilde{x}\right| \le \Delta\right) = \gamma , \tag{3.1}$$

где $\bar{x}$ – генеральная средняя;

$\tilde{x}$ – выборочная средняя;

Δ – предельная ошибка выборки,

γ – уровень доверительной вероятности.

Предельная и средняя ошибки выборки связаны соотношением:

$$\Delta = t \cdot \mu , \tag{3.2}$$

где μ – средняя ошибка выборки;

t – коэффициент, зависящий от уровня доверительной вероятности.

Обычно уровень доверительной вероятности равен 0,9; 0,95 или 0,99. При большом объеме выборочной совокупности для этих уровней доверительной вероятности t равно 1,65; 1,96 или 2,58 соответственно.

Средняя ошибка выборки находится в зависимости от вида и способа отбора. Различают следующие способы отбора: собственно-случайный; механический; типический (районированный); серийный (гнездовой); комбинированный; многоступенчатый; многофазный; взаимопроникающий и другие.

При простой случайной выборке отбор единиц производится из генеральной совокупности путем жеребьевки или с помощью таблицы случайных чисел. При этом способе единица наблюдения совпадает с единицей отбора.

Средняя ошибка выборки ($\mu_{\bar{x}}$) находится по формуле

а) если отбор случайный повторный:

$$\mu_{\bar{x}} = \frac{\sigma}{\sqrt{n}}, \tag{3.3}$$

б) если отбор случайный бесповторный:

$$\mu_{\bar{x}} = \sqrt{\frac{\sigma^2}{n}\left(1 - \frac{n}{N}\right)}, \tag{3.4}$$

где n – объем выборочной совокупности;

N – объем генеральной совокупности;

σ^2 – дисперсия генеральной совокупности. Так как ее значение обычно неизвестно, то в формулах берется значение выборочной дисперсии $\left(\sigma_в^2\right)$.

В больших выборках (n>30) выборочная дисперсия определяется по формуле:

$$\sigma_в^2 = \frac{\sum (x_i - \tilde{x})^2 n_i}{n}, \tag{3.5}$$

где $\tilde{x}$ - выборочная средняя.

В малых выборках ($n \leq 30$):

$$\sigma^2{}_в = \frac{\sum (x_i - \tilde{x})^2 n_i}{n - 1}. \tag{3.6}$$

Выборочная дисперсия в малых выборках обычно обозначается S^2.

Значения коэффициента t для больших выборок находятся по таблице интеграла вероятностей в соответствии с выбранным уровнем доверительной вероятности. Для малых выборок t находят по таблице критических значений t – Стьюдента в соответствии с уровнем доверительной вероятности и числом степеней свободы $k=n-1$ (приложение В).

Доверительный интервал, который покрывает неизвестное значение генеральной средней с заданной доверительной вероятностью, определяется неравенством:

$$\tilde{x} - \Delta_{\bar{x}} \leq \bar{x} \leq \tilde{x} + \Delta_{\bar{x}}, \tag{3.7}$$

где $\Delta_{\bar{x}} = t \cdot \mu_{\bar{x}}$.

При случайном отборе средняя ошибка выборки для доли (P) находится по формуле

а) если отбор повторный:

$$\mu_p = \sqrt{\frac{w(1-w)}{n}}, \tag{3.8}$$

б) если отбор бесповторный:

$$\mu_p = \sqrt{\frac{w(1-w)}{n}\left(1-\frac{n}{N}\right)}\ . \tag{3.9}$$

В формулах w – это выборочная доля единиц, обладающих данным признаком.

Доверительный интервал для генеральной доли определяется следующим неравенством:

$$w - \Delta_p \le P \le w + \Delta_p, \tag{3.10}$$

где $\Delta_p = t \cdot \mu_p$.

При проведении выборочного наблюдения важным является обеспечение достаточно большого объема выборки, чтобы достигалась необходимая точность результатов и были приемлемы затраты средств и труда на проведение исследования.

Необходимый объем выборки выводится из формул предельной ошибки выборки.

При собственно-случайном повторном отборе:

$$n = \frac{t^2\sigma_в^2}{\Delta^2}\ . \tag{3.11}$$

При собственно-случайном бесповторном отборе:

$$n = \frac{t^2\sigma_в^2 N}{N\Delta^2 + t^2\sigma_в^2}\ . \tag{3.12}$$

Пример 3.1 Считая полученные числовые характеристики ($\bar{x}$; σ^2) интервального ряда распределения в примере 2.2 результатом случайной бесповторной 10 % выборки, определить с доверительной вероятностью 0,95:

а) границы доверительного интервала для средней урожайности кукурузы на зерно по всей совокупности хозяйств;

б) необходимый объем выборки, если предельная ошибка будет уменьшена в 2 раза.

Решение. Средняя урожайность по выборке n = 46 хозяйств составила $\bar{x}$ = 42,7 ц/га, дисперсия σ^2 = 99,47.

Объем генеральной совокупности: $N = \frac{n}{0,1} = \frac{46}{0,1}$ = 460 (организаций).

При доверительной вероятности 0,95 значение t=1,96.

Тогда предельная ошибка выборки составит:

$$\Delta_{\bar{x}} = 1,96\sqrt{\frac{99,47}{46}\left(1-\frac{46}{460}\right)} = 2,734.$$

Вывод. Средняя урожайность кукурузы на зерно на одно хозяйство во всей генеральной совокупности при доверительной вероятности 0,95 определяется промежутком 42,7 ± 2,7 ц/га, т.е. покрывается интервалом от 40,0 до 45,4 ц/га.

б) Необходимый объем выборки при предельной ошибке, уменьшенной в два раза, будет равен:

$$n = \frac{t^2 \sigma_в^2 N}{N\Delta^2 + t^2 \sigma_в^2} = \frac{1{,}96^2 \cdot 99{,}47 \cdot 460}{1{,}368^2 \cdot 460 + 1{,}96 \cdot 99{,}47} = \frac{175777{,}018}{1242{,}344} = 141{,}5 \approx 142 .$$

Вывод. Необходимый объем выборки n = 142 организации, т. е. при уменьшении предельной ошибки в 2 раза, объем выборки увеличивается в 3 раза.

Задачи для самостоятельного решения

Задача 3.1 Считая числовые характеристики интервального ряда распределения в соответствии со своим вариантом результатами случайной бесповторной 20 % выборки (см. задачи 2.4 – 2.7), с доверительной вероятностью 0,95 определить:

а) границы доверительного интервала для генеральной средней;

б) необходимый объем выборки, если предельная ошибка выборки будет уменьшена в 2 раза.

Задача 3.2 В агрохолдинге имеется 360 комбайнов. В результате случайного бесповторного отбора было обследовано 60 комбайнов, из которых со сроком эксплуатации свыше 8 лет оказалось 30 %. При уровне доверительной вероятности 0,997 определить долю и количество комбайнов со сроком эксплуатации свыше 8 лет в целом по агрохолдингу.

Задача 3.3 В области имеется 3,6 тыс. фермерских хозяйств. В результате случайного бесповторного отбора 12 % хозяйств оказалось, что средняя урожайность зерновых составила 36 ц/га при среднем квадратическом отклонении 10 ц/га. Известно, что 85 % общей площади посевов зерновых культур занято озимой пшеницей. С вероятностью 0,95 определить границы, в которых будет находиться урожайность зерновых культур и доля посевов озимой пшеницы во всех фермерских хозяйствах. Оценить величину валового сбора зерна во всех фермерских хозяйствах, если общая площадь посева зерновых культур в области составила 300 тыс. га.

Задача 3.4 При проверке качества семян сахарной свеклы было отобрано 30 проб в случайном порядке. Средний процент всхожести семян составил 77 % при среднем квадратическом отклонении 5 %. Определить границы, в которых будет находиться среднее значение процента всхожести семян во всех образцах. Расчеты произвести с вероятностью 0,954.

Задача 3.5 Проводилось испытание 10 сортов озимой пшеницы. Каждый сорт высевался на 6 делянках опытного поля одинаковой площади в равных условиях. По одному сорту определить среднюю урожайность, среднюю и предельную ошибку выборки. Уровень доверительной вероятности принять 0,95.

Таблица 3.1 – Урожайность озимой пшеницы, ц/га

Номер делянки	Номер сорта									
	1	2	3	4	5	6	7	8	9	10
1	55,1	49,4	60,2	41,2	55,6	66,5	78,1	49,6	66,4	54,2
2	50,6	51,1	61,3	44,1	54,8	68,8	78,2	55,4	67,2	60,3
3	54,2	45,7	59,8	49,6	49,7	70,2	70,3	53,2	75,4	62,1
4	58,7	47,9	64,1	43,5	52,3	64,7	79,2	56,7	73,1	59,6
5	53,4	50,4	65,4	46,7	54,2	68,4	80,6	54,2	66,2	58,6
6	56,8	51,2	63,8	45,2	50,7	67,1	81,2	50,8	69,1	61,4

Задача 3.6 В населенном пункте имеется 1200 хозяйств населения. В результате случайного бесповторного отбора 10 % хозяйств оказалось, что средняя урожайность овощей составила 250 ц/га при среднем квадратическом отклонении 56 ц/га. Известно, что 42 % посевов овощей занимают помидоры. С вероятностью 0,954 определить границы, в которых будет находиться средняя урожайность овощей и доля посевов помидор в общей площади посевов во всех хозяйствах населения. Определить возможный валовой сбор овощей, если известно, что площадь посева овощей во всех хозяйствах составляет 50 га. Как изменится предельная ошибка выборки, если число отобранных хозяйств населения увеличить в 1,5 раза?

Задача 3.7 В районе имеется 670 крестьянских хозяйств. Сколько хозяйств необходимо взять для обследования, если известно, что средний размер земельного участка составляет 43 га, при среднем квадратическом отклонении 18 га. Уровень вероятности принять 0,95, точность 5 %.

Задача 3.8 Способом случайного бесповторного отбора произведено выборочное обследование урожайности подсолнечника в 20 % хозяйств. Средняя урожайность составила 18 ц/га при среднем квадратическом отклонении 4 ц/га. Определить: среднюю и предельную ошибку выборки; границы, в которых находится урожайность подсолнечника во всех хозяйствах.

Контрольные тесты для проверки знаний

1 Для использования выборочной совокупности для дальнейшего развития социально-экономического явления необходимо, чтобы разница между средним значением генеральной совокупности и средним значением выборочной совокупности была не больше __________ ошибки выборки.

а) средней;
б) генеральной;
в) предельной;
г) индивидуальной.

2 Способ наблюдения, при котором дается подобное описание отдельных единиц наблюдения в статистической совокупности, называется…

а) выборочным наблюдением;
б) сплошным наблюдением;

в) обследованием основного массива;
г) монографическим обследованием.

3 Для использования выборочной совокупности для дальнейшего анализа развития социально-экономического явления необходимо, чтобы разница между средним значением генеральной совокупности и средним значением выборочной совокупности была не больше __________ ошибки выборки.
а) индивидуальной;
б) генеральной;
в) предельной;
г) средней.

4 Выборка называется малой в том случае, если ее объем составляет менее __________ единиц.
а) 30;
б) 40;
в) 50;
г) 100.

5 Если вероятность, гарантирующую результат, увеличить с 0,954 до 0,997, то объем повторной случайной выборки увеличится в ... раз.
а) 2,25;
б) 2,30;
в) 4,50;
г) 4,25.

6 Способы отбора единиц в выборочную совокупность:
а) механический;
б) типический;
в) аналитический;
г) серийный.

7 Недостающим элементом в формуле расчета объема выборки при бесповторном случайном отборе: $n = \dfrac{t^2 \times N \times ...}{N \times \Delta^2 + t^2 \times \sigma^2}$
а) σ;
б) σ^2;
в) Δ^2;
г) (1 – n/N).

8 Недостающим элементом в формуле расчета объема выборки при бесповторном случайном отборе : $n = \dfrac{t^2 \times N \times \sigma^2}{N \times ... + t^2 \times \sigma^2}$
а) σ;
б) σ^2;
в) (1 – n/N).
г) Δ^2.

9 Недостающим элементом в формуле расчета объема выборки при бесповторном случайном отборе: $n = \dfrac{t^2 \times N \times \sigma^2}{N \times \Delta^2 + t^2 \times ...}$
а) σ;
б) σ^2;
в) (1 – n/N).
г) Δ^2.

10 Репрезентативность результатов выборочного наблюдения зависит от ... .

а) вариации признака
б) способа формирования выборочной совокупности
в) объема выборки
г) определения границ объекта исследования

11 Для расчета средней ошибки выборки используют формулу:

$$\mu = \frac{\sigma}{\sqrt{n-1}}$$ при ...

а) наличии высокого уровня вариации признака
б) изучении качественных характеристик явлений
в) малой выборке
г) уточнении данных сплошного наблюдения

12 Средняя ошибка случайной повторной выборки ... , если ее объем увеличить в 4 раза.

а) уменьшится в 2 раза
б) увеличится в 4 раза
в) уменьшится в 4 раза
г) не изменится

13 Недостающим элементом формулы предельной ошибки случайной выборки при бесповторном отборе является: $\ldots \times \sqrt{\frac{\sigma^2}{n}(1-\frac{n}{N})}$

а) t
б) t^2
в) n^2
г) n

14 Средняя ошибка выборки (μ) характеризует:

а) вариацию признака
б) тесноту связи между двумя факторами
в) величину предельной ошибки выборки при t=1
г) величину предельной ошибки выборки при t→∞

15 Под выборочным наблюдением понимают ... .

а) сплошное наблюдение всех единиц совокупности
б) несплошное наблюдение части единиц совокупности
в) несплошное наблюдение части единиц совокупности, отобранных случайным способом
г) обследование наиболее крупных единиц совокупности

Вопросы для самоподготовки

1. В чем состоит сущность выборочного метода?
2. В каких областях применяется выборочный метод?
3. Какие ошибки выборочного наблюдения вы знаете?
4. Назовите виды и способы формирования выборочной совокупности.
5. Назовите характеристики выборочной и генеральной совокупности.
6. Как определяется необходимая численность выборки?
7. Какие существуют способы распространения данных выборочного наблюдения на генеральную совокупность?

Тема 4 Проверка статистических гипотез

Статистической гипотезой называется всякое высказывание о генеральной совокупности, проверяемое по выборке. Она может касаться вида неизвестного распределения, отдельных параметров распределений, связей между случайными величинами и т.п. Например, малые предприятия по выручке от реализации продукции распределяются по нормальному закону. Известно, что совокупность безработных территориального образования распределяется по нормальному закону, выдвигается гипотеза, что средний стаж работы составляет 10 лет. Если сравниваются две или более генеральных совокупностей, имеющих один и тот же закон распределения, то могут быть выдвинуты гипотезы о равенстве средних значений или дисперсий этих совокупностей.

Процесс использования выборки для проверки гипотезы называется *статистическим доказательством*. Выдвигаемую гипотезу называют *нулевой или основной* H_0. Наряду с нулевой гипотезой рассматривают ей противоположную гипотезу, которая называется *альтернативной или конкурирующей* и обозначается H_1. Например, H_0: $M(X)=1$, математическое ожидание генеральной совокупности, распределенной по показательному закону, равно единице. Тогда конкурирующая гипотеза может иметь вид: H_1: $M(X)>1$, или $M(X)<1$, или $M(X)\neq 1$ (математическое ожидание больше 1, или меньше 1, или не равно 1). Выдвигаемая гипотеза может содержать одно или несколько предположений. Если параметрическая гипотеза содержит одно утверждение, то она называется простой, а если множество утверждений – то сложной. Простой будет гипотеза, что среднее время безотказной работы холодильника определенной марки составляет 45 тыс. часов, а сложной – что среднее время безотказной работы составит менее 45 тыс. часов.

Так как проверка статистических гипотез осуществляется по выборочным данным, то нельзя быть уверенным об истинности или ложности выдвинутой гипотезы. Выбор между гипотезами H_0 и H_1 может сопровождаться ошибками двух родов. Ошибка первого рода заключается в том, что будет отвергнута верная нулевая гипотеза. Ошибка *первого рода* α означает вероятность принятия H_1, если верна гипотеза H_0, т.е. $\alpha = P(H_1/H_0)$. Ошибка второго рода состоит в том, что будет принята неправильная гипотеза. Ошибка *второго рода* означает вероятность принятия H_0, если верна гипотеза H_1: $\beta = P(H_0/H_1)$. Существует правильное решение двух родов: $P(H_0/H_0) = 1-\alpha$ и $P(H_1/H_1) = 1-\beta$ (таблица 4.1).

Таблица 4.1 – Ошибки первого и второго рода

Принятая гипотеза	H_0	H_1
H_0 - верна	$P(H_0/H_0)=1-\alpha$	$P(H_1/H_0)=\alpha$
H_0 – неверна	$P(H_0/H_1)=\beta$	$P(H_1/H_1)=1-\beta$

Правило, по которому принимается решение о том, что верна или не верна гипотеза H_0, называется *критерием,* где:

$\alpha = P(H_1/H_0)$ – *уровень значимости критерия*;

$M = 1-\beta = P(H_1/H_1)$ – *мощность критерия*.

Чем меньше уровень значимости, тем меньше будет вероятность отклонить верную нулевую гипотезу. Уровень значимости обычно задается заранее. Чем больше мощность критерия, тем меньше вероятность ошибки второго рода. Нельзя уменьшить одновременно величину ошибки первого и второго рода при заданном объеме выборки.

Статистическим критерием К называют случайную величину, с помощью которой принимают решение о принятии или отклонении H_0.

Для проверки параметрических гипотез используют *критерии значимости*, основанные на статистиках: u, χ^2, t, F (приложения 5-7). Непараметрические гипотезы проверяют с помощью *критериев согласия*, использующих статистики распределений: χ^2, Колмогорова-Смирнова и т.д.

После выбора статистического критерия, все множество его значений разбивается на два непересекающихся подмножества. Одно подмножество содержит значения критерия, при которых нулевая гипотеза принимается. Это подмножество называется областью допустимых значений критерия или областью принятия гипотезы. Второе подмножество содержит значения критерия, при которых нулевая гипотеза отвергается и называется критической областью. Точка, разделяющая эти подмножества, называется критической точкой, которая обозначается $К_{кр}$. По данным выборки находится наблюдаемое значение критерия. Затем наблюдаемое значение критерия сравнивается с критическим значением. Если наблюдаемое значение критерия принадлежит критической области, то нулевую гипотезу отвергают, а если принадлежит области допустимых значений – то гипотезу принимают.

Пример 4.1 Урожайность озимой пшеницы определенного сорта по совокупности крестьянских хозяйств распределяется по нормальному закону с известным средним квадратическим отклонением $\sigma = 6{,}4$ ц/га и генеральной средней $\bar{X}_г = 60{,}0$ ц/га. По выборочной совокупности 50 крестьянских хозяйств найдена выборочная средняя урожайность, составившая 63 ц/га. При уровне значимости α=0,05 проверить нулевую гипотезу H_0: $\bar{X} = \bar{X}_г = 60$, при конкурирующей гипотезе H_1: $\bar{X} \neq 60$.

Решение. Необходимо рассмотреть критерий $К = u$, где

$$U_{набл} = \frac{(\bar{X} - \bar{X}_г)\sqrt{n}}{\sigma} = \frac{(63{,}0 - 60{,}0)\sqrt{50}}{6{,}4} = 3{,}31.$$

а) По условию конкурирующая гипотеза имеет вид $\bar{X} \neq 60{,}0$, поэтому критическая область двусторонняя. Найдем критическую точку из равенства $\Phi(u_{кр,\alpha/2}) = (1-\alpha)/2 = (1-0{,}05)/2 = 0{,}475$. Согласно приложения 1: $u_{кр} = 1{,}96$.

$|u_{набл}| > u_{кр}$, поэтому следует отклонить нулевую гипотезу, то есть выборочная и гипотетическая генеральная средняя различаются значимо.

б) Проверим нулевую гипотезу H_0: $\bar{X} = a_0$, при H_1: $\bar{X} > a_0$.

При конкурирующей гипотезе H_1: $\bar{X} > a_0$ критическая область является правосторонней. Критическую точку находят из равенства

$$\Phi(u_{\text{кр}}) = \frac{1-2\alpha}{2}. \tag{4.1}$$

Если $U_{\text{н}} > U_{\text{кр}}$,то выдвинутая нулевая гипотеза отвергается, выборочная средняя значимо отличается от генеральной средней.

Если $U_{\text{н}} < U_{\text{кр}}$, то нулевая гипотеза принимается, нет оснований отвергнуть нулевую гипотезу.

По условию примера 13.1 конкурирующая гипотеза имеет вид H_1: $\bar{x}$ >60,0, поэтому критическая область правосторонняя. Найдем критическую точку из равенства $\Phi(u_{\text{кр.},\alpha/2}) = (1–2\alpha)/2 = (1–2\cdot 0{,}05)/2 = 0{,}45$. Согласно приложения 1: $u_{\text{кр.}}= 1{,}645$. Так как $|u_{н.}| > u_{кр.}$, то следует отклонить нулевую гипотезу.

в) Проверим нулевую гипотезу H_0: $\bar{X} = a_0$, при H_1: $\bar{X} < a_0$.

Наблюдаемое значение находится по формуле (13.4). Критическую точку находят из равенства (13.6), учитывая, что критическая точка левосторонняя.

Если $U_{\text{н}} < U_{\text{кр}}$,то выдвинутая нулевая гипотеза отвергается, выборочная средняя значимо отличается от генеральной средней.

Если $U_{\text{н}} > U_{\text{кр}}$, то нулевая гипотеза принимается, нет оснований отвергнуть нулевую гипотезу.

Учитывая, что $U_{\text{кр. левост.}} =- U_{\text{кр. правост.}}$, то вывод можно формулировать, как и в пункте (б).

Если по условию примера 13.1 средняя выборочная урожайность составила 58 ц/га, то конкурирующая гипотеза имеет вид H_1: $\bar{x} < 60{,}0$, критическая область левосторонняя. При $\alpha = 0{,}05$, $u_{\text{кр.}}= – 1{,}64$.

$$U_{\text{н}} = \frac{(\bar{X}-\bar{X}_{\text{г}})\sqrt{n}}{\sigma} = \frac{(58{,}0-60{,}0)\sqrt{50}}{6{,}4} = -2{,}21.$$

Так как $U_{\text{н}} < U_{\text{кр}}$,то выдвинутая нулевая гипотеза отвергается, выборочная средняя значимо отличается от генеральной средней.

Пусть генеральная совокупность распределена по нормальному закону, но числовые характеристики непосредственно неизвестны. Несмещенными их оценками служат выборочная средняя и «исправленная» выборочная дисперсия s^2.

Тогда в качестве критерия проверки нулевой гипотезы используется

$t -$ распределение Стьюдента. Наблюдаемое значение критерия находится по формуле:

$$t_{\text{н}} = \frac{\bar{X}-a_0}{s/\sqrt{n}} = \frac{(\bar{X}-a_0)\sqrt{n}}{s}. \tag{4.2}$$

Критическая область строится в зависимости от вида конкурирующей гипотезы, как рассмотрено выше. При заданном уровне значимости α и числе степеней свободы $k = n - 1$ по таблице распределения Стьюдента (приложение 3) находится критическое значение критерия для односторонней или двухсторонней критической области. Сравнивая наблюдаемое значение критерия с критическим значением, формулируется вывод.

Пример 4.2 Ожидаемая урожайность подсолнечника может составить 30 ц с 1 га. В результате посева на 6 делянках одинаковой площади была получена урожайность с 1 га: 31,8; 28,8; 29,4; 30,2; 32,2;30,6. При уровне значимости 0,05 проверить гипотезу о равенстве средней урожайности подсолнечника 30 ц/га.

Решение. Считая распределение урожайности по участкам нормально распределенной случайной величиной, проверим нулевую гипотезу H_0: $\bar{X} = a_0$, при H_1: $\bar{X} \neq a_0$. По условию a_0 =30. Дисперсия генеральной совокупности неизвестна, выборка малая по объему, найдем «исправленную» выборочную дисперсию, для чего составим таблицу 4.2.

$$\bar{X} = \frac{\sum x_i}{n} = \frac{183{,}0}{6} = 30{,}5;\ s^2 = \frac{\sum(x_i - \bar{X})^2}{n-1} = \frac{8{,}78}{6-1} = 1{,}756;\ s = \sqrt{1{,}756} =$$
$$= 1{,}325.$$

$$t_н = \frac{(\bar{X} - a_0)\sqrt{n}}{s} = \frac{(30{,}5 - 30{,}0) \cdot \sqrt{6}}{1{,}325} = 0{,}92.$$

Средняя урожайность подсолнечника составила 30,5 ц/га. По участкам урожайность в среднем колебалась в границах 30,5$\pm$1,3 ц/га, т. е. от 29,2 до 31,8 ц/га.

По таблице t–распределения Стьюдента при уровне значимости α = 0,05 и числе степеней свободы $k = n - 1 =$ 6–1 = 5, $t_{кр}$ = 2,57.

Таблица 4.2 – Урожайность подсолнечника с 1 га, ц

№ п/п	Урожайность, ц/га	$x_i - \bar{X}$	$(x_i - \bar{X})^2$
1	31,8	1,3	1,69
2	28,8	-1,7	2,89
3	29,4	-1,1	1,21
4	30,2	-0,3	0,09
5	32,2	1,7	2,89
6	30,6	0,1	0,01
Итого	183,0	0,0	8.78

Сравниваем наблюдаемое значение критерия с критическим значением. Так как $t_н < t_{кр.}$, то нулевая гипотеза принимается, средняя урожайность подсолнечника на всех участках может составить 30 ц с 1 га.

Гипотеза может быть проверена с использованием доверительного интервала средней арифметической.

$$\bar{X} - t_{кр} \frac{s}{\sqrt{n}} \leq a_0 \leq \bar{X} + t_{кр} \frac{s}{\sqrt{n}},$$

30,5 –2,57$\cdot\frac{1{,}325}{\sqrt{6}} \leq a_0 \leq$ 30,5 + 2,57$\cdot\frac{1{,}325}{\sqrt{6}}$, $29{,}1 \leq a_0 \leq 31{,}9$.

С доверительной вероятностью 0,95 можно утверждать, что средняя урожайность на всех участках может находиться в пределах от 29,1 до 31,9 ц/га.

Так как этот интервал покрывает значение a_0 =30, то нулевая гипотеза принимается.

Проверим нулевую гипотезу H_0: $\bar{X} = a_0$, при H_1: $\bar{X} > a_0$.

Критическая область является правосторонней. При уровне значимости α = 0,05 и числе степеней свободы $k = n - 1$ =6 – 1 = 5, $t_{кр}$ = 2,01. Так как $t_н < t_{кр.}$, то нулевая гипотеза принимается.

Пример 4.3 Оценить существенность различий в средней урожайности двух сортов озимой пшеницы. Для первого сорта средняя урожайность $\bar{X}_1 = 35{,}6$ ц/га и выборочная дисперсия $S_1^2 = 8{,}05$, для второго сорта средняя урожайность $\bar{X}_2 = 48{,}4$ ц/га и выборочная дисперсия $S_2^2 = 14{,}31$. Объемы выборок n_1=5 и n_2=5 соответственно.

Решение. Выдвигаем нулевую гипотезу о том, что средние урожайности двух сортов пшеницы не отличаются друг от друга, т.е. $H_0: \bar{X}_1 = \bar{X}_2$, при альтернативной гипотезе $H_1: \bar{X}_1 \neq \bar{X}_2$ – урожайности существенно различны. Примем уровень значимости $\alpha = 0{,}05$.

Так как выборки независимые, причем $n_1 = n_2$, то применим –критерий Стьюдента с $k = n_1 + n_2 - 2$ степенями свободы.

$$t_н = \frac{|\bar{X}_1 - \bar{X}_2|}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}} = \frac{48{,}4 - 35{,}6}{\sqrt{\frac{8{,}05}{5} + \frac{14{,}31}{5}}} = \frac{12{,}8}{2{,}11} = 6{,}07.$$

По приложению 3 определим критическое значение t-распределения:

$$t_{кр} = t_{0{,}05;10} = 2{,}31.$$

При числе степеней свободы $k = 5 + 5 - 2 = 10$.

Так как $t_н > t_{кр}$, то нулевую гипотезу следует отклонить. Следовательно, два сорта озимой пшеницы отличаются статистически значимо по величине средней урожайности.

Задачи для самостоятельного решения

Задача 4.1 Проводилось испытание 6 сортов озимой пшеницы. Каждый сорт высевался на 6 делянках одинаковой площади. При 5 % уровне значимости проверить гипотезу о существенности различий в средней урожайности двух сортов озимой пшеницы (комбинация сортов предлагается студенту преподавателем).

Таблица 4.3 – Урожайность озимой пшеницы по сортам, ц/га

Повторение	Урожайность озимой пшеницы по сортам, ц/га					
	«Батько»	«Крошка»	«Пал-Пич»	«Таня»	«Ласточка»	«Москвич»
1	52,3	49,6	53,3	63,2	53,3	64,2
2	53,8	49,4	55,8	65,8	54,8	65,2
3	51,0	50,1	54,0	64,7	55,7	64,3
4	53,1	50,3	53,9	62,9	53,9	63,9
5	52,8	49,9	52,8	66,0	54,6	65,0
6	53,0	50,0	55,4	66,3	55,8	65,5

Задача 4.2 Проводилось испытание 4 сортов озимого ячменя. Каждый сорт высевался на 5 делянках одинаковой площади (таблица 4.6). При 5 % уровне значимости оценить существенность различий в средней урожайности двух сортов озимого ячменя (комбинация сортов предлагается студенту преподавателем).

Таблица 4.4 – Урожайность озимого ячменя по сортам, ц/га

Номер испытания	Урожайность озимого ячменя по сортам, ц/га			
	«Атаман»	«Виват»	«Визит»	«Маргрет»
1	55,3	49,6	42,6	40,9
2	56,8	48,4	41,8	44,9
3	62,0	44,3	43,4	45,7
4	60,4	46,6	44,0	44,3
5	59,2	49,8	42,9	41,6

Задача 4.3 Проведено выборочное обследование 10 % личных подсобных хозяйств населения 8 районов случайным бесповторным способом. Получены следующие результаты об урожайности овощей.

Таблица 4.5 – Показатели производства овощей

Район	Средняя урожайность, ц/га	Среднее квадратическое отклонение, ц/га	Доля овощей в площади участков, %	Число обследованных дворов
1	210	30	30	100
2	246	35	35	80
3	305	32	40	150
4	220	24	50	120
5	164	20	36	60
6	280	23	65	70
7	340	40	45	90
8	316	36	53	100

При уровне значимости 0,05 по двум районам проверить гипотезы о равенстве: средних выборочных урожайностей, долей посевов овощей в площади приусадебных участков.

Задача 4.4

Определённые сорта озимой пшеницы (приложение И) испытывались на одинаковом числе участков на протяжении семи лет. При уровне значимости α = 0,05 проверить нулевую гипотезу о существенности различий в урожайности двух сортов озимой пшеницы по индивидуальным вариантам.

Вариант	Сорта	**Вариант**	Сорта
1	I и II	**14**	III и IV
2	I и III	**15**	III и V
3	I и IV	**16**	III и VI

Вариант	Сорта	**Вариант**	Сорта
4	I и V	**17**	III и VII
5	I и VI	**18**	VI и VII
6	I и VII	**19**	VI и VIII
7	I и VIII	**20**	VII и VIII
8	II и III	**21**	IV и VI
9	II и IV	**22**	IV и VII
10	II и V	**23**	V и VII
11	II и VI	**24**	V и VIII
12	II и VII	**25**	IV и VIII
13	II и VIII	-	-

Задача 4.5 Проверить гипотезу о равенстве средних урожайностей овощей в двух совокупностях хозяйств, если в результате случайной выборки получены следующие результаты:

1 совокупность		2 совокупность	
Урожайность, т с 1 га	Число хозяйств	Урожайность, т с 1 га	Число хозяйств
x_i	n_i	y_i	m_i
25-35	15	15-25	22
35-45	30	25-35	30
45-55	24	35-45	41
	n=69	45-55	17
			m=110

Задача 4.6 По двум независимым выборкам объема n_1 и n_2, извлеченным из нормальных генеральных совокупностей, проверить гипотезу о равенстве средних, при уровне значимости α=0,01, если:

а) $\bar{x} = 50$; $\bar{y} = 45$; $D(X) = 1200$; $D(Y) = 2025$; $n_1 = 35$; $n_2 = 45$;

б) $\bar{x} = 70$; $\bar{y} = 60$; $D(X) = 1470$; $D(Y) = 1320$; $n_1 = 60$; $n_2 = 40$.

Задача 4.7 Провести две случайные выборки по одному из показателей приложения 8, объемами n_1 и n_2. Проверить нулевую гипотезу о равенстве выборочных средних, при уровне значимости 0,05 (предполагается, что дисперсии неизвестны и одинаковы): а) n_1=n_2=20; б) n_1=20; n_2=10.

Контрольные тесты для проверки знаний

1 Для проверки гипотезы о равенстве генеральных дисперсий нормальных совокупностей наблюдаемое зна5чение критерия находят как:

а) $F_{\text{набл}} = \dfrac{S_{\text{Б}}^2}{S_{\text{М}}^2}$

б) $F_{\text{набл}} = \frac{S_{\text{М}}^2}{S_{\text{Б}}^2}$;

в) $F_{\text{набл}} = \frac{D_{\text{Б}}^2}{D_{\text{М}}^2}$;

г) $F_{\text{набл}} = \frac{D_{\text{М}}^2}{D_{\text{Б}}^2}$

2 Для проверки гипотезы о равенстве неизвестной генеральной дисперсии σ^2 гипотетическому значению σ_0^2 надо вычислить

а) $\chi_{\text{набл}}^2 = \frac{(n-1)S^2}{\sigma_0^2}$

б) $\chi_{\text{набл}}^2 = \frac{(n-1)\sigma^2}{\sigma_0^2}$;

в) $\chi_{\text{набл}}^2 = \frac{(n-1)S^2}{\sigma}$

г) $\chi_{\text{набл}}^2 = \frac{nS^2}{\sigma_0^2}$;

3 Из нормальной генеральной совокупности извлечена выборка объёма $n = 21$ и по ней найдена исправленная выборочная дисперсия $S^2 = 16{,}2$. Тогда наблюдаемое значение критерия $\chi_{\text{набл}}^2$ при $\sigma_0^2 = 15$, равна:

а) 216;

б) 2,16;

в) 21,6;

г) 0,216.

4 Для проверки гипотезы о существенности различий генеральных средних в случае больших независимых выборок, при условии, что известны соответствующие дисперсии, надо вычислить:

а) $Z_{\text{набл}} = \frac{\overline{x}-\overline{y}}{\sqrt{D(X)+D(Y)}}$;

б) $Z_{\text{набл}} = \frac{\overline{x}-\overline{y}}{\sqrt{D(X)/n+D(Y)/m}}$;

в) $Z_{\text{набл}} = \frac{D(X)-D(Y)}{\sqrt{D(X)/n+D(Y)/m}}$;

г) $Z_{\text{набл}} = \frac{n-m}{\sqrt{D(X)/n+D(Y)/m}}$.

5 Наблюдаемое значение критерия для проверки гипотезы о равенстве генеральных средних в случае независимых выборок при $n = 40, m = 50; D(X) = 80, D(Y) = 100; \overline{x} = 130, \overline{y}$=140, равно:

а) 5;

б) -5;

в) -0,5;

г) 0,5.

6 При проверке гипотезы о равенстве математических ожиданий в случае малых независимых выборок наблюдаемое значение критерия равно:

а) $T_{\text{набл}} = \frac{\overline{x}-\overline{y}}{\sqrt{(n-1)S_X^2+(m-1)S_Y^2}}\sqrt{\frac{nm-2}{n+m}}$;

б) $T_{\text{набл}} = \frac{\overline{x}-\overline{y}}{S_X^2+S_Y^2}\sqrt{\frac{nm(n+m-2)}{n+m}}$

в) $T_{\text{набл}} = \frac{\overline{x}-\overline{y}}{\sqrt{(n-1)S_X^2+(m-1)S_Y^2}}\sqrt{\frac{nm(n+m-2)}{n+m}}$;

г) $T_{\text{набл}} = \frac{\overline{x}-\overline{y}}{(n-1)+(m-1)}\sqrt{\frac{nm(n+m-2)}{n+m}}$.

7 Наблюдаемое значение критерия для проверки гипотезы о равенстве математиче-

ских ожиданий при $n = 12, m = 18; S_X^2 = 0{,}84;\ S_Y^2 = 0{,}40;\ \overline{x} = 31{,}2;\ \overline{y} = 29{,}2$, равно:

а) 7,1;
б) 0,71;
в) -7,1;
г) -0,71.

8 При проверке гипотезы о равенстве выборочной средней и генеральной средней в случае больших независимых выборок наблюдаемое значение критерия равно:

а) $U_{\text{набл}} = \frac{(\overline{x}-a_0)\sqrt{n}}{\sigma}$;
б) $U_{\text{набл}} = \frac{(\overline{x}-a_0)\sqrt{\sigma}}{n}$;
в) $U_{\text{набл}} = \frac{(\overline{x}+a_0)\sqrt{n}}{\sigma}$
г) $U_{\text{набл}} = \frac{(\overline{x}-a_0)}{\sigma\sqrt{n}}$

9 Наблюдаемое значение критерия для проверки гипотезы о равенстве выборочной средней и генеральной средней при $n = 100; \overline{x} = 27{,}56; a_0 = 26; \sigma = 5{,}2$, равно:

а) 3;
б) 15,6;
в) -3;
г) -15,6.

10 При проверке гипотезы о существенности различий средних нормальных генеральных совокупностей в случае малых зависимых выборок наблюдаемое значение критерия равно:

а) $T_{\text{набл}} = \overline{d}\sqrt{(n-1)}/S_d$;
б) $T_{\text{набл}} = \overline{d}\sqrt{n}/S_d$;
в) $T_{\text{набл}} = \overline{d}\sqrt{n}/\sigma$;
г) $T_{\text{набл}} = \sqrt{n}/S_d$

11 Наблюдаемое значение критерия для проверки гипотезы о существенности различий средних нормальных генеральных совокупностей в случае малых зависимых выборок при

x_i	2	3	5	6	8	10
y_i	10	3	6	1	7	4

равна:

а) 2,4;
б) -2,4;
в) 0,24;
г) -0,24

12 При проверке гипотезы о существенности различий между наблюдаемой относительной частотой и гипотетической вероятностью наблюдаемое значение критерия равно:

а) $U_{\text{набл}} = \frac{|(m/n)-p_0|\sqrt{x}}{\sqrt{p_0q_0}}$;
б) $U_{\text{набл}} = \frac{|(m/n)-p_0|\sqrt{x}}{\sqrt{p_0}}$;
в) $U_{\text{набл}} = \frac{|(m/n)-p_0|\sqrt{p_0}}{\sqrt{q_0}}$;
г) $U_{\text{набл}} = \frac{|(m/n)-p_0|\sqrt{n}}{\sqrt{p_0q_0}}$.

13 Наблюдаемое значение критерия для проверки гипотезы о существенности различий между наблюдаемой относительной частотой и гипотетической вероятностью

при n=100; $m/n = 0{,}14;\ p_0 = 0{,}20,$ равно:

а) 1,5;

б) 0,15;

в) -1,5;

г) 0,15.

14 При проверке гипотезы о существенности различий двух вероятностей биномиальных распределений наблюдаемое значение критерия равно:

а) $U_{\text{набл}} = \frac{m_1/n_1 - m_2/n_2}{\sqrt{\frac{m_1+m_2}{n_1+n_2}(1-\frac{m_1+m_2}{n_1+n_2})(\frac{1}{n_1}+\frac{1}{n_2})}};$

б) $U_{\text{набл}} = \frac{m_1/n_1 - m_2/n_2}{\sqrt{\frac{m_1+m_2}{n_1+n_2}(1-\frac{m_1+m_2}{n_1+n_2})}}$

в) $U_{\text{набл}} = \frac{m_1/n_1 - m_2/n_2}{\sqrt{(1-\frac{m_1+m_2}{n_1+n_2})(\frac{1}{n_1}+\frac{1}{n_2})}};$

г) $U_{\text{набл}} = \frac{m_1/n_1 - m_2/n_2}{\sqrt{\frac{m_1+m_2}{n_1+n_2}(1-\frac{m_1+m_2}{n_1+n_2})(\frac{1}{n_1}+\frac{1}{n_2})}}$

15 Наблюдаемое значение критерия для проверки гипотезы о существенности различий двух вероятностей биномиальных распределений при $m_1 = 15;\ n_1 = 800;\ m_2 = 25;\ n_2 = 1000$ равно:

а) 8,9;

б) 0,89;

в) -0,89;

г) -8,9.

Вопросы для самоподготовки

1. Что называется статистической гипотезой?
2. Какие существуют виды статистических гипотез?
3. В чем заключаются ошибки первого и второго рода?
4. В чем заключается сущность понятия «статистический критерий»?
5. Раскройте понятия «критическая область», «область принятия гипотезы», «критическая точка».
6. Назовите этапы проверки статистических гипотез.

Тема 5 Дисперсионный анализ

Задача дисперсионного анализа – количественная оценка влияния тех или иных факторов (или уровней факторов) на изменчивость средних значений наблюдаемых случайных величин.

Сущность дисперсионного анализа. Дисперсионный анализ состоит в выделении и оценке отдельных факторов, вызывающих изменчивость. С этой целью производят разложение общей дисперсии ($\boldsymbol{S}^{\mathbf{2}}_{\text{общ}}$) наблюдаемой совокупности (общей дисперсии признака), вызванной всеми источниками изменчивости, на составляющие дисперсии, порожденные независимыми факторами ($\boldsymbol{S}^{\mathbf{2}}_{\text{факт}}$ и $\boldsymbol{S}^{\mathbf{2}}_{\text{ост}}$**)**. Каждая из этих составляющих дает оценку дисперсии, вызванную конкретным источником изменчивости, в общей совокупности. Для проверки значимости этих составляющих оценок дисперсии их сравнивают с общей дисперсией в общей совокупности (по критерию Фишера).

Если уже установлено, что фактор существенно влияет на X, и требуется выяснить, какой из уровней оказывает наибольшее воздействие, то производится дополнительное попарное сравнение средних.

Рассмотрим случай однофакторного анализа, когда на X воздействует один фактор, имеющий p постоянных уровней.

Пусть на нормально распределённый количественный признак X воздействует фактор F с p постоянными уровнями. Число наблюдений на каждом уровне одинаково и равно q.

Пусть наблюдалось $n=pq$ значений x_{ij} признака X, где i– номер испытания ($i = 1, 2, \ldots, q$), j – номер уровня фактора ($j = 1, 2, \ldots, p$).

Номер испытания	Уровни фактора F_j			
	F_1	F_2	…	F_p
1	x_{11}	x_{12}	…	x_{1p}
2	x_{21}	x_{22}	…	x_{2p}
…	…	…	…	…
q	x_{q1}	x_{q2}	…	x_{qp}
Групповая средняя	$\overline{x}_{\text{гр }1}$	$\overline{x}_{\text{гр }2}$	…	$\overline{x}_{\text{гр }p}$

Практически остаточную сумму находят как $S_{\text{ост}} = S_{\text{общ}} - S_{\text{факт}}$, где

$\boldsymbol{S}_{\text{общ}} = \sum_{j=1}^{p}\sum_{i=1}^{q}(\boldsymbol{x}_{ij} - \overline{\boldsymbol{x}})^{\mathbf{2}}$ – общая сумма квадратов отклонений наблюдаемых значений от общей средней $\overline{x}$;

$\boldsymbol{S}_{\text{факт}} = \boldsymbol{q}\sum_{j=1}^{p}(\overline{\boldsymbol{x}}_{\text{гр}j} - \overline{\boldsymbol{x}})^{\mathbf{2}}$ – факторная сумма квадратов отклонений групповых средних от общей средней, которая характеризует рассеяние между группами;

$\boldsymbol{S}_{\text{ост}} = \sum_{i=1}^{q}(\boldsymbol{x}_{i1} - \overline{\boldsymbol{x}}_{\text{гр}1})^2 + \sum_{i=1}^{q}(\boldsymbol{x}_{i2} - \overline{\boldsymbol{x}}_{\text{гр}2})^2 + \cdots + \sum_{i=1}^{q}(\boldsymbol{x}_{ip} - \overline{\boldsymbol{x}}_{\text{гр}p})^2$ – остаточная сумма квадратов отклонений наблюдаемых значений группы от своей групповой средней, которая характеризует рассеяние внутри групп.

Пример 5.1 Двумя приборами произведены по два измерения физической величины, истинный размер которой равен x. Рассматривая в качестве фактора систематическую ошибку C, а в качестве его уровней - систематические ошибки C_1 и C_2 соответственно первого и второго прибора, показать, что $S_{\text{факт}}$ определяется систематическими, а $S_{\text{ост}}$ – случайными ошибками измерений.

Решение. Введём обозначения: α_1, α_2 – случайные ошибки первого и второго измерений первым прибором; β_1, β_2 – случайные ошибки первого и второго измерений вторым прибором.

Тогда наблюдаемые значения результатов измерений соответственно равны (первый индекс при x указывает на номер измерения, а второй – на номер прибора):

$x_{11} = x + C_1 + \alpha_1$, $x_{21} = x + C_1 + \alpha_2$; $x_{12} = x + C_2 + \beta_1$, $x_{22} = x + C_2 + \beta_2$.

Средние значения измерений первым и вторым приборами соответственно равны:

$\overline{x}_{\text{гр }1} = x + C_1 + [(\alpha_1 + \alpha_2)/2] = x + C_1 + \alpha$,

$\overline{x}_{\text{гр }2} = x + C_2 + [(\beta_1 + \beta_2)/2] = x + C_2 + \beta$.

Общая средняя $\overline{x} = \frac{(\overline{x}_{\text{гр }1} + \overline{x}_{\text{гр }2})}{2} = x + \left[\frac{C_1 + C_2}{2}\right] + [(\alpha + \beta)/2]$,

факторная сумма $S_{\text{факт}} = (\overline{x}_{\text{гр }1} - \overline{x})^2 + (\overline{x}_{\text{гр }2} - \overline{x})^2$.

Подставив величины, заключённые в скобках, получим

$S_{\text{факт}} = [(C_1 - C_2)^2/2] + (C_1 - C_2)(\alpha - \beta) + [(\alpha - \beta)^2/2]$.

Мы видим, что $S_{\text{факт}}$ определяется в основном, первым слагаемым (так как случайные ошибки измерений малы) и, следовательно, действительно отражает влияние фактора C.

Остаточная сумма $S_{\text{ост}} = (x_{11} - \overline{x}_{\text{гр }1})^2 + (x_{21} - \overline{x}_{\text{гр }1})^2 + (x_{12} - \overline{x}_{\text{гр }2})^2 + (x_{22} - \overline{x}_{\text{гр }2})^2$.

Подставив величины, заключённые в скобках, получим $S_{\text{ост}} = (\alpha_1 - \alpha)^2 + (\alpha_2 - \alpha)^2 + (\beta_1 - \beta)^2 + (\beta_2 - \beta)^2$.

Мы видим, что $S_{\text{ост}}$определяется случайными ошибками измерений и, следовательно, действительно отражает влияние случайных причин.

Покажем, что $\boldsymbol{S}_{\text{общ}} = \boldsymbol{S}_{\text{факт}} + \boldsymbol{S}_{\text{ост}}$. Для упрощения вывода ограничимся двумя уровнями и двумя испытаниями на каждом уровне.

Номер испытания	Уровни фактора	
i	F_1	F_2
1 2	X_{11} X_{21}	X_{12} X_{22}
$\overline{x}_{\text{гр } j}$	$\overline{x}_{\text{гр }1}$	$\overline{x}_{\text{гр }2}$

Тогда $S_{\text{общ}} = (x_{11} - \overline{x})^2 + (x_{21} - \overline{x})^2 + (x_{12} - \overline{x})^2 + (x_{22} - \overline{x})^2$.

Вычтем и прибавим к каждому наблюдаемому значению на первом уровне групповую среднюю $\overline{x}_{\text{гр }1}$, а на втором - $\overline{x}_{\text{гр }2}$. Выполнив возведение в квадрат и учитывая, что сумма всех удвоенных произведений равна нулю, получим $S_{\text{общ}} = 2\left[\left(\overline{x}_{\text{гр }1} - \overline{x}\right)^2 + \left(\overline{x}_{\text{гр }2} - \overline{x}\right)^2\right] + (x_{11} - \overline{x}_{\text{гр }1})^2 + (x_{21} - \overline{x}_{\text{гр }1})^2 + (x_{12} - \overline{x}_{\text{гр }2})^2 + (x_{22} - \overline{x}_{\text{гр }2})^2 = S_{\text{факт}} + S_{\text{ост}}$.

Разделив суммы квадратов на соответствующее число степеней свободы, получим общую, факторную и остаточную дисперсии:

$\boldsymbol{S}^{\mathbf{2}}_{\text{общ}} = \frac{\boldsymbol{S}_{\text{общ}}}{\boldsymbol{pq-1}}$, $\boldsymbol{S}^{\mathbf{2}}_{\text{факт}} = \frac{\boldsymbol{S}_{\text{факт}}}{\boldsymbol{p-1}}$, $\boldsymbol{S}^{\mathbf{2}}_{\text{ост}} = \frac{\boldsymbol{S}_{\text{ост}}}{\boldsymbol{p(q-1)}}$, где p – число уровней фактора, q – число наблюдений на каждом уровне; $(pq – 1)$ – число степеней свободы общей дисперсии; $(p – 1)$ – число степеней свободы факторной дисперсии; $p(q – 1)$ – число степеней свободы остаточной дисперсии.

Если нулевая гипотеза о равенстве средних справедлива, то все эти дисперсии являются несмещёнными оценками генеральной дисперсии. Например, учитывая, что объём выборки $n=pq$, заключаем, что

$\boldsymbol{S}^{\mathbf{2}}_{\text{общ}} = \frac{\boldsymbol{S}_{\text{общ}}}{\boldsymbol{pq-1}} = \frac{\boldsymbol{S}_{\text{общ}}}{\boldsymbol{n-1}}$ – исправленная выборочная дисперсия, которая является несмещённой оценкой генеральной дисперсии.

Пример 5.2 На 4 опытных участках произведены испытания трёх сортов винограда. Получены данные об урожайности винограда в расчете на 1 га посевной площади Результаты испытаний приведены в таблице. Методом дисперсионного анализа при уровне значимости 0,05 проверить нулевую гипотезу о влиянии фактора сорта на урожайность.

Номер испытания	Уровни фактора F_j		
i	F_1	F_2	F_3
1	51	52	42
2	52	54	44
3	56	56	50
4	57	58	52
$\overline{x}_{\text{гр }j}$	54	55	47

Решение. Составим расчётную таблицу. Найдём общую и факторную суммы квадратов отклонений.

$$S_{\text{общ}} = \sum_{j=1}^{p} P_j - \frac{\left(\sum_{j=1}^{p} R_j\right)^2}{pq} = 32714 - 32448 = 266;$$

$$S_{\text{факт}} = \frac{\sum_{j=1}^{p} R_j^2}{q} - \frac{\left(\sum_{j=1}^{p} R_j\right)^2}{pq} = \frac{130400}{4} - 32448 = 152.$$

Найдём остаточную сумму квадратов отклонений: $S_{\text{ост}} = S_{\text{общ}} - S_{\text{факт}} = 266 - 152 = 114$.

Номер испытания	Уровни фактора F_j						Итоговый столбец
i	F_1		F_2		F_3		
	x_{i1}	x_{i1}^2	x_{i2}	x_{i2}^2	x_{i3}	x_{i3}^2	
1 2 3 4	51 52 56 57	2601 2704 3136 3249	52 54 56 58	2704 2916 3136 3364	42 44 50 52	1764 1936 2500 2704	
$P_j = \sum_{i=1}^{4} x_{ij}^2$		11690		12120		8904	$\sum P_j$ =32714
$R_j = \sum_{i=1}^{4} x_{ij}$	216		220		188		$\sum R_j = 624$
R_j^2	46656		48400		35344		$\sum R_j^2$ =130400

Найдём факторную и остаточную дисперсии:

$S_{\text{факт}}^2 = \frac{S_{\text{факт}}}{p-1} = \frac{152}{3-1} = 76$; $S_{\text{ост}}^2 = \frac{S_{\text{ост}}}{p(q-1)} = \frac{114}{3(4-1)} = \frac{114}{9} = 12{,}67.$

Найдём наблюдаемое значение критерия:

$F_{\text{набл}} = \frac{S_{\text{факт}}^2}{S_{\text{ост}}^2} = \frac{76}{12{,}67} = 6.$

Учитывая, что число степеней свободы числителя k_1=2, а знаменателя k_2=9 и уровень значимости α=0,05, находим критическое значение критерия по таблице приложений $F_{\text{набл}}$ (0,05; 2; 9) =4,26.

Так как $F_{\text{набл}} > F_{\text{кр}}$ – нулевую гипотезу об отсутствии влияния фактора сорта на урожайность винограда отвергается. **Вывод:** ***имеется статистически значимая разность в средней урожайности между сортами винограда.***

Задачи для самостоятельного решения

Задача 5.1 Провести дисперсионный анализ полевого опыта по сортоиспытанию при уровне значимости 0,05. Оценить существенность различий в урожайности 4 сортов озимой пшеницы с помощью наименьшей разности в сравнении с первым вариантом, взятым за стандарт.

Номер испытания	Уровни фактора			
i	F_1	F_2	F_3	F_4
1	36	56	52	39
2	47	61	57	57
3	50	64	59	63
4	58	66	58	61
5	67	66	79	65
$\overline{x}_{\text{гр } j}$	51,6	62,6	61	57

Задача 5.2 Произведено испытание эффективности шести видов минеральных удобрений в восьми фермерских хозяйствах. Методом дисперсионного анализа проверить гипотезу о незначимости влияния применения минеральных удобрений при уровне значимости 0,01. Результаты испытаний приведены в следующей таблице.

Номер испытания	Уровни фактора					
i	F_1	F_2	F_3	F_4	F_5	F_6
1	100	92	74	68	64	69
2	101	102	87	80	83	71
3	126	104	88	83	83	80
4	128	115	93	87	84	80
5	133	119	94	96	90	81
6	141	122	101	97	96	82
7	147	128	102	106	101	86
8	148	146	105	127	111	99
$\bar{x}_{\text{гр } j}$	128	116	93	93	89	81

Задача 5.3 Произведено всего 10 испытаний урожайности яровой пшеницы, из них на 4 участках высажен первый сорт пшеницы, ещё на 4 – второй сорт, на 2 участках – третий (таблица). Методом дисперсионного анализа проверить гипотезу о несущественности различий средней урожайности сортов при уровне значимости 0,05.

Номер испытания	Уровни фактора F_j		
i	F_1	F_2	F_3
1	35	30	21
2	32	24	22
3	31	26	34
4	30	20	31
$\bar{x}_{\text{гр } j}$	32	25	27

Задача 5.4 Произведено по семь испытаний на каждом из четырёх уровней фактора F, где фактор F – различные виды минеральных удобрений. Методом дисперсионного анализа проверить гипотезу о несущественности различий урожайности при применении четырёх видов удобрений при уровне значимости 0,05. Результаты испытаний приведены в следующей таблице.

Номер испытания	Уровни фактора			
i	F_1	F_2	F_3	F_4
1	51	52	56	54
2	59	58	56	58
3	53	66	58	62
4	59	69	58	64
5	63	70	70	66
6	69	72	74	67
7	72	74	78	69
$\bar{x}_{\text{гр } j}$	60,9	65,9	64,3	62,9

Задача 5.5 Произведены испытания в четырёх районов Краснодарского края на каждом из трёх уровней фактора F, где F – урожайность овощей на

приусадебных участках. Получены следующие результаты об урожайности овощей. Методом дисперсионного анализа проверить гипотезу о значимости различий урожайности овощей при уровне значимости 0,05.

Номер испытания	Уровни фактора F_j		
i	F_1	F_2	F_3
1	27	24	22
2	23	20	21
3	29	26	36
4	29	30	37
$\overline{x}_{\text{гр } j}$	27	25	29

Задача 5.6 По данным таблицы оценить значимость различий в средней урожайности сортов яровой пшеницы. Уровень значимости равен 0,05.

Номер испытания	Уровни фактора				
i	F_1	F_2	F_3	F_4	F_5
1	42	66	35	64	70
2	55	91	50	70	79
3	67	96	60	79	88
4	67	98	69	81	90
$\overline{x}_{\text{гр } j}$	57,75	87,75	53,5	73,5	81,75

Задача 5.7 По данным следующей таблицы оценить существенность влияния дозы минеральных удобрений на число зерен в колосе озимой пшеницы. Уровень значимости принять 0,05.

Номер испытания	Уровни фактора			
i	F_1	F_2	F_3	F_4
1	6	6	9	7
2	7	7	12	9
3	8	11	13	10
4	11	12	14	10
$\overline{x}_{\text{гр } j}$	8	9	12	9

Задача 5.8 По данным следующей таблицы определить, является ли существенным влияние предшественника на урожайность риса.

Предшественник	Урожайность риса сорта «Вираж» в зависимости от предшественника				
	1	2	3	4	5
Рис	42,1	41,6	35,6	34,2	30,9
Озимая пшеница	49,7	43,2	45,8	43,0	45,2
Многолетние травы	56,2	57,0	54,0	51,4	52,1

Контрольные тесты для проверки знаний

1 Задача дисперсионного анализа состоит:

а) количественная оценка влияния неучтённых факторов на изменчивость средних значений наблюдаемых случайных величин;

б) количественная оценка влияния тех или иных факторов на изменчивость средних квадратических отклонений наблюдаемых случайных величин;
в) качественная оценка влияния уровней факторов на изменчивость средних значений наблюдаемых случайных величин;
г) количественная оценка влияния тех или иных факторов (или уровней факторов) на изменчивость средних значений наблюдаемых случайных величин.

2 На практике остаточную сумму находят как:
а) $S_{\text{ост}} = S_{\text{общ}} + S_{\text{факт}}$;
б) $S_{\text{общ}} = S_{\text{факт}} - S_{\text{ост}}$;
в) $S_{\text{ост}} = S_{\text{общ}} - S_{\text{факт}}$;
г) $S_{\text{общ}} = S_{\text{ост}} - S_{\text{факт}}$.

3 Общая сумма квадратов отклонений наблюдаемых значений от общей средней $\overline{x}$ равна:
а) $S_{\text{общ}} = \sum_{j=1}^{p}\sum_{i=1}^{q}(x_{ij} - \overline{x})$;
б) $S_{\text{факт}} = \sum_{j=1}^{p}\sum_{i=1}^{q}(x_{ij} + \overline{x})^2$;
в) $S_{\text{ост}} = \sum_{j=1}^{p}\sum_{i=1}^{q}(x_{ij} - \overline{x})^2$;
г) $S_{\text{общ}} = \sum_{j=1}^{p}\sum_{i=1}^{q}(x_{ij} - \overline{x})^2$.

4 Факторная сумма квадратов отклонений групповых средних от общей средней, которая характеризует рассеяние между группами равна:
а)
$$S_{\text{факт}} = q\sum_{j=1}^{p}(\overline{x}_{\text{гр}j} - \overline{x})$$
б)
$$S_{\text{факт}} = q\sum_{j=1}^{p}(\overline{x}_{\text{гр}j} - \overline{x})^2$$
в)
$$S_{\text{факт}} = p\sum_{j=1}^{q}(\overline{x}_{\text{гр}j} - \overline{x})^2$$
г)
$$S_{\text{факт}} = q\sum_{j=1}^{p}(\overline{x}_{\text{гр}j} + \overline{x})^2$$

5 Остаточная сумма квадратов отклонений наблюдаемых значений группы от своей групповой средней, которая характеризует рассеяние внутри групп, равна:
а)
$$S_{\text{ост}} = \sum_{i=1}^{q}\left(x_{i1} + \overline{x}_{\text{гр}1}\right)^2 + \sum_{i=1}^{q}\left(x_{i2} + \overline{x}_{\text{гр}2}\right)^2 + \cdots + \sum_{i=1}^{q}(x_{ip} + \overline{x}_{\text{гр}p})^2;$$
б)
$$S_{\text{ост}} = \sum_{i=1}^{q}\left(x_{i1} - \overline{x}_{\text{гр}1}\right)^2 + \sum_{i=1}^{q}\left(x_{i2} - \overline{x}_{\text{гр}2}\right)^2 + \cdots + \sum_{i=1}^{q}(x_{ip} - \overline{x}_{\text{гр}p})^2;$$
в)
$$S_{\text{ост}} = \sum_{i=1}^{q}\left(x_{i1} - \overline{x}_{\text{гр}1}\right) + \sum_{i=1}^{q}\left(x_{i2} - \overline{x}_{\text{гр}2}\right) + \cdots + \sum_{i=1}^{q}(x_{ip} - \overline{x}_{\text{гр}p}) ;$$
г)
$$S_{\text{общ}} = \sum_{i=1}^{q}\left(x_{i1} - \overline{x}_{\text{гр}1}\right)^2 + \sum_{i=1}^{q}\left(x_{i2} - \overline{x}_{\text{гр}2}\right)^2 + \cdots$$

$+\sum_{i=1}^{q}(x_{ip}-\overline{x}_{\text{гр}p})^2.$

6 В формуле $S_{\text{общ}}^2 = \frac{S_{\text{общ}}}{pq-1}$ параметр p означает:

а) число уровней фактора;
б) число наблюдений на каждом уровне;
в) число степеней свободы общей дисперсии;
г) число степеней свободы остаточной дисперсии.

7 Общую дисперсию находят по формуле:

а) $S_{\text{общ}}^2 = \frac{S_{\text{общ}}}{p(q-1)}$;
б) $S_{\text{общ}}^2 = \frac{S_{\text{общ}}}{p-1}$;
в) $S_{\text{общ}}^2 = \frac{S_{\text{общ}}}{pq-1}$;
г) $S_{\text{общ}}^2 = \frac{S_{\text{общ}}}{q-1}$.

8 Остаточную дисперсию находят по формуле

а) $S_{\text{ост}}^2 = \frac{S_{\text{ост}}}{p(q-1)}$;
б) $S_{\text{ост}}^2 = \frac{S_{\text{ост}}}{p-1}$;
в) $S_{\text{ост}}^2 = \frac{S_{\text{ост}}}{pq-1}$;
г) $S_{\text{ост}}^2 = \frac{S_{\text{ост}}}{q-1}$.

9 Факторную дисперсию находят по формуле:

а) $S_{\text{факт}}^2 = \frac{S_{\text{факт}}}{p(q-1)}$;
б) $S_{\text{факт}}^2 = \frac{S_{\text{факт}}}{p-1}$
в) $S_{\text{факт}}^2 = \frac{S_{\text{факт}}}{pq-1}$
г) $S_{\text{факт}}^2 = \frac{S_{\text{факт}}}{q-1}$

10 В формуле $S_{\text{общ}}^2 = \frac{S_{\text{общ}}}{pq-1}$ параметр q означает:

а) число уровней фактора
б) число наблюдений на каждом уровне;
в) число степеней свободы общей дисперсии;
г) число степеней свободы факторной дисперсии.

11 В формуле $S_{\text{общ}}^2 = \frac{S_{\text{общ}}}{pq-1}$ величина $(pq-1)$ означает:

а) число уровней фактора
б) число наблюдений на каждом уровне;
в) число степеней свободы общей дисперсии;
г) число степеней свободы факторной дисперсии

12 В формуле $\boldsymbol{S}_{\text{факт}}^{\mathbf{2}} = \frac{\boldsymbol{S}_{\text{факт}}}{\boldsymbol{p-1}}$ величина $(p-1)$ означает:

а) число уровней фактора
б) число наблюдений на каждом уровне;
в) число степеней свободы общей дисперсии;
г) число степеней свободы факторной дисперсии

13 В формуле $\boldsymbol{S}_{\text{ост}}^{\mathbf{2}} = \frac{\boldsymbol{S}_{\text{ост}}}{\boldsymbol{p(q-1)}}$ величина $p(q-1)$ означает:

а) число уровней фактора
б) число наблюдений на каждом уровне

в) число степеней свободы факторной дисперсии;
г) число степеней свободы остаточной дисперсии.

14 $S_{общ} = 428$, $S_{факт} = 224$, тогда $S_{ост}$ равно:
а) 652;
б) 0;
в) 204
г) 1,91.

15 Наблюдаемое значение критерия Фишера равно:
а) $F_{набл} = \frac{S^2_{факт}}{S^2_{ост}}$;
б) $F_{набл} = \frac{S^2_{ост}}{S^2_{факт}}$;
в) $F_{набл} = \frac{S^2_{общ}}{S^2_{ост}}$;
г) $F_{набл} = \frac{S^2_{факт}}{S^2_{общ}}$.

Вопросы для самоподготовки

1. На чем основан дисперсионный анализ?

2. В чем заключается основная задача дисперсионного анализа?

3. Назовите виды дисперсионного анализа. Перечислите этапы проведения дисперсионного анализа.

5. По какой формуле определяется наблюдаемое значение критерия Фишера-Снедекора?

Тема 6 Статистическое изучение связей

Изучаемые явления (процессы) характеризуются признаками или показателями, которые подразделяются на факторные и результативные. Признаки, оказывающие влияние на другие признаки, называются факторными (независимыми, экзогенными). Признаки, изменяющиеся под влиянием факторных признаков, называются результативными (зависимыми, эндогенными). Если на изменение результативного признака оказывает влияние один, доминирующий фактор, то связь между ними называется парной и тогда строится модель парной регрессии. Неявно предполагается, что этот фактор (x) является причиной изменения результативного признака (y).

Частным случаем статистической связи является корреляционная, которая проявляется в среднем, в массе наблюдений, как статистическая закономерность. При корреляционной связи с изменением факторного признака на определенную величину изменяется среднее значение результативного признака. Обычно корреляционная зависимость представляется как функциональная зависимость между переменными в виде уравнения регрессии.

Задачами корреляционно-регрессионного анализа являются:

- установление типа уравнения регрессии;
- определение параметров уравнения регрессии и оценка значимости параметров;
- оценка тесноты и направления связи между переменными;
- оценка значимости уравнения регрессии;
- определение прогнозных значений зависимой переменной и оценка полученного прогноза.

При изучении связей между признаками устанавливают ее аналитическое выражение в виде линейного и нелинейного уравнения связи. Линейная связь описывается уравнением $y = a + bx$, которое на графике имеет вид прямой линии. При нелинейной зависимости используются параболическая, степенная, показательная и другие функции.

Применение корреляционно-регрессионного анализа предполагает проведение исследований в несколько этапов.

Первый этап: подбор факторных и результативных признаков, между которыми изучается причинно-следственная связь.

Второй этап: определение формы связи и подбор математического уравнения, которое наиболее полно отражает характер взаимосвязи между признаками. Для этого используют графический метод. В прямоугольной системе координат на оси абсцисс откладывают значения факторного признака (x), на оси ординат – результативного (y).

Третий этап: рассчитываются параметры уравнения связи с целью установления количественного влияния факторных признаков на результат. При парной линейной связи параметр a – свободный член уравнения, b – коэффициент регрессии, который показывает, на сколько единиц в натуральном выражении изменится результативный признак при изменении факторного на единицу.

Параметры линейного уравнения определяют методом наименьших квадратов, путем составления и решения системы уравнений:

$$\begin{cases} \sum y = na + b\sum x \\ \sum xy = a\sum x + b\sum x^2 \end{cases}, \text{ откуда } b = \frac{n\sum xy - \sum x\sum y}{n\sum x^2 - \left(\sum x\right)^2}, \; a = \frac{\sum y}{n} - b\frac{\sum x}{n}. \quad (6.1)$$

Четвертый этап: оценка и анализ полученных результатов при помощи коэффициентов корреляции, детерминации, эластичности и других.

Коэффициент корреляции (r) характеризует направление и тесноту связи, он изменяется от -1 до 1.

Чем ближе абсолютное значение коэффициента корреляции к нулю, тем линейная связь между признаками слабее. Чем ближе к единице, тем связь сильнее. Если $r = \pm 1$, то связь линейная функциональная. Если $r = 0$, то связь между признаками отсутствует, т. е. они линейно независимы.

Знак коэффициента корреляции показывает направление связи между признаками. Если $r > 0$, то связь прямая, обе переменные изменяются в одном направлении. Если $r < 0$, то связь обратная, переменные изменяются в противоположных направлениях (рисунок 1.3).

Коэффициент корреляции рассчитывается по формуле:

$$r = \frac{\overline{xy} - \overline{x} \cdot \overline{y}}{\sigma_x \cdot \sigma_y}. \quad (6.2)$$

где $\overline{x} = \frac{\sum x}{n}$ – среднее значение x

$\overline{y} = \frac{\sum y}{n}$ - среднее значение y;

$\overline{xy} = \frac{\sum xy}{n}$ - среднее значение произведения xy;

σ_x и σ_y –среднее квадратическое отклонение x и y соответственно:

$$\sigma_x = \sqrt{\frac{\sum x^2}{n} - \left(\overline{x}\right)^2}, \; \sigma_y = \sqrt{\frac{\sum y^2}{n} - \left(\overline{y}\right)^2}.$$

Квадрат коэффициента корреляции называется коэффициентом детерминации. Коэффициент детерминации D показывает долю влияния фактора x на результативную переменную y, а $(1 - r^2)$ – долю влияния других, неучтенных в модели факторов.

$$D = r^2 \cdot 100 \quad \%. \quad (6.3)$$

Он показывает, какая часть колеблемости результативного признака объясняется вариацией факторного признака.

Относительное влияние факторного признака на результативный признак определяется с помощью коэффициента эластичности:

$$Э = b \cdot \frac{\overline{x}}{\overline{y}}. \quad (6.4)$$

Он показывает, на сколько процентов в среднем изменится результативный признак с изменением факторного на 1 %.

Для определения тесноты связи между двумя признаками, измеренными в порядковых шкалах, применяются менее точные, но более простые по расчету непараметрические показатели, в частности коэффициенты корреляции рангов (или ранговые коэффициенты корреляции) Спирмена (ρ) и Кендалла (τ).

Оба показателя основаны на корреляции не самих значений изучаемых признаков, а их рангов. Ранг – это порядковый номер, присваиваемый каждому индивидуальному значению x и y (отдельно) в ранжированном ряду. Оба признака необходимо ранжировать (нумеровать) в одном и том же порядке: от меньших значений к большим или наоборот. Если встречается несколько одинаковых значений x (или y), то каждому из них присваивается ранг, равный частному от деления суммы рангов, приходящихся на эти значения, на число равных значений. Ранги признаков x и y обозначают соответственно символами R_x и R_y.

Для расчета коэффициента корреляции рангов Спирмена значения признаков x и y нумеруют от 1 до n. Затем для каждой пары рангов находят их разность $d_i = R_{x_i} - R_{y_i}$, а коэффициент корреляции рангов определяют по формуле:

$$\rho = 1 - \frac{6\sum d_i^2}{n(n^2 - 1)}, \tag{6.5}$$

где n – число наблюдаемых пар значений x и y.

Коэффициент корреляции рангов Кендалла определяется по формуле:

$$\tau = \frac{2S}{n(n-1)}, \text{ где } S=P-Q. \tag{6.6}$$

Порядок расчета коэффициента корреляции рангов Кендалла:

1) значения x и y ранжируют, т. е. определяют R_{xi} и R_{yi};

2) значения R_{xi} располагают в порядке возрастания и параллельно записывают соответствующее каждому R_{xi} значение R_{yi};

3) для каждого значения R_{yi} подсчитывают число следующих за ним рангов более высокого порядка и число следующих за ним рангов, меньших по значению; находят соответствующие суммы P и Q;

4) определяют разность S и коэффициент корреляции рангов Кендалла.

Как и линейный коэффициент корреляции, коэффициенты корреляции рангов могут изменяться в пределах от -1 до +1. Чем ближе их значения по модулю к 1, тем теснее связь между x и y.

Коэффициент Кендалла всегда меньше по значению, чем коэффициент Спирмена, причем $\tau \approx \frac{2}{3}\rho$.

В статистике выделяют классификационные переменные, которые позволяют разбить совокупность наблюдений на непересекающиеся множества, которые трудно или невозможно упорядочить по какому-либо признаку (например: качество жилья – плохое, удовлетворительное, хорошее, отличное; пол особи – мужской и женский; вид и род в биологии и т. д.).

В общем случае основным инструментом исследования является изучение зависимостей между классификационными переменными при помощи построения таблиц сопряженности.

Таким образом, при исследовании тесноты связи между качественными признаками строят таблицы сопряженности.

Если каждый из двух качественных признаков принимает только альтернативные значения, то таблица сопряженности имеет вид:

Признак А	Признак В	
	да	нет
Да	a	b
Нет	c	d

Каждая из клеток данной таблицы соответствует известной альтернативе того или другого признака. Для оценки тесноты связи рассчитывают коэффициенты ассоциации (K_a) и контингенции ($K_к$):

$$K_a = \frac{ad - bc}{ad + bc} \tag{6.7}$$

$$K_к = \frac{ad - bc}{\sqrt{(a+b)(c+d)(a+c)(b+d)}}. \tag{6.8}$$

Данные коэффициенты изменяются от -1 до +1 , причем коэффициент контингенции всегда меньше коэффициента ассоциации. Связь считается значимой и подтвержденной, если $|K_a| \geq 0{,}5$, а $|K_к| \geq 0{,}3$.

Если каждый из качественных признаков, измеренных в номинальной шкале, имеет более двух значений, то таблица сопряженности примет вид:

Значения признака А	Значения признака В						Всего
	B_1	B_2	…	B_j	…	B_p	
A_1	f_{11}	f_{12}	...	f_{1j}	...	f_{1p}	n_1
A_2	f_{21}	f_{22}	...	f_{2j}	...	f_{2p}	n_2
…	...	...	...	...	...	...	...
A_i	f_{i1}	f_{i2}	...	f_{ij}	...	f_{ip}	n_i
…	...	...	...	...	...	...	...
A_s	f_{s1}	f_{s2}	...	f_{sj}	...	f_{sp}	n_s
Итого	m_1	m_2	...	m_j	...	m_p	$\Sigma n_i = \Sigma m_j$

Для определения тесноты связи применяют коэффициенты взаимной сопряженности Пирсона и Чупрова. Эти коэффициенты неотрицательны и принимают нулевое значение тогда, когда признаки статистически независимы.

Коэффициент Пирсона исчисляется по формуле:

$$C = \sqrt{\frac{\varphi^2}{1+\varphi^2}}, \tag{6.9}$$

где φ^2 – показатель взаимной сопряженности.

$$\varphi^2 = \sum_{i=1}^{s} \sum_{j=1}^{p} \frac{f_{ij}^2}{n_i m_j} - 1, \tag{6.10}$$

где s – число категорий (значений) признака А;
p – число категорий признака В.

Коэффициент Чупрова исчисляется по формуле:

$$K_ч = \sqrt{\frac{\varphi^2}{\sqrt{(s-1)(p-1)}}}. \tag{6.11}$$

Чем ближе C и $Kч$ к 1, тем связь теснее. Коэффициент взаимной сопряженности Чупрова более точен, поскольку учитывает число значений по каждому признаку.

Пример 6.1 По данным 10 организаций (таблица 6.1) об энерговооруженности рабочей силы и выручке от продажи продукции на одного работника определить: форму связи между признаками; параметры уравнения связи; коэффициенты корреляции, детерминации и эластичности. Оценить значимость коэффициента корреляции.

Таблица 6.1 – Энерговооруженность и выручка от продажи продукции на одного среднегодового работника

№ организации	Энергетические мощности на среднегодового работника, л.с. (x)	Выручка от продажи продукции на одного работника, тыс. руб. (y)	x^2	xy	y^2	Теоретическое значение выручки на одного работника, тыс. руб. $y_x=a+bx$
1	54,9	762,3	3014,01	41850,27	581101,29	870,7
2	32,8	438,9	1075,84	14395,92	192633,21	455,2
3	55,2	672,1	3047,04	37099,92	451718,41	876,3
4	43,9	636,2	1927,21	27929,18	404750,44	663,9
5	45,6	908,4	2079,36	41423,04	825190,56	695,8
6	59,3	1027,4	3516,49	60924,82	1055550,76	953,4
7	45,9	550,9	2106,81	25286,31	303490,81	701,5
8	52,2	1054,9	2724,84	55065,78	1112814,01	819,9
9	43,2	613,5	1866,24	26503,20	376382,25	650,7
10	50,6	812,6	2560,36	41117,56	660318,76	789,8
Итого	483,6	7477,2	23918,20	371596,00	5963950,50	7477,2

В данном случае результативный признак y – выручка от продажи продукции в расчете на одного работника, а факторный признак x – энерговооруженность труда.

Решение.

1) Исходные данные нанесем на график (рисунок 6.1). По расположению точек на графике видно, что связь можно выразить уравнением $y_x = a + bx$.

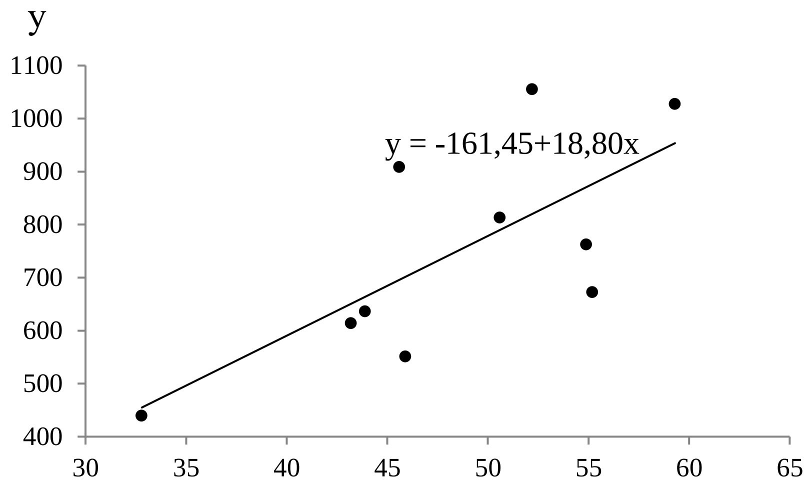

Рисунок 6.1 – Зависимость между энерговооруженностью рабочей силы и выручкой на работника

2) Для нахождения параметров линейного уравнения составим систему уравнений, используя данные таблицы 6.1. Система уравнений имеет вид:

$$\begin{cases} 7477{,}2 = 10a + 483{,}6b, \\ 371596{,}0 = 483{,}6a + 23918{,}2b. \end{cases}$$

Решив систему, получим a = -161,45; b = 18,80.

Следовательно, уравнение связи энерговооруженности рабочей силы и выручки от реализации продукции на одного работника имеет вид:

$$y_x = -161{,}45 + 18{,}80x.$$

Нанесем полученные точки на график, проведем прямую, выраженную уравнением регрессии.

3) Определим средние значения показателей:

$$\bar{x} = \frac{483{,}6}{10} = 48{,}36, \qquad \bar{y} = \frac{7477{,}2}{10} = 747{,}72,$$

$$\overline{xy} = \frac{371596{,}00}{10} = 37159{,}60.$$

4) Найдем среднее квадратическое отклонение x и y:

$$\sigma_x = \sqrt{\frac{23918{,}2}{10} - 48{,}36^2} = 7{,}29;$$

$$\sigma_y = \sqrt{\frac{5963950,50}{10} - 747,72^2} = 193,16;$$

5) Рассчитаем коэффициент корреляции:

$$\text{в)}\ r = \frac{\overline{xy} - \bar{x} \cdot \bar{y}}{\sigma_x \sigma_y} = \frac{37159,60 - 48,36 \cdot 747,72}{7,29 \cdot 193,16} = 0,71.$$

Следовательно, связь между изучаемыми признаками прямая и довольно тесная.

6) Коэффициент детерминации равен:

$Д = r^2 \cdot 100 = 0,71^2 \cdot 100 = 50,4$ %.

Таким образом, колеблемость выручки на среднегодового работника по совокупности хозяйств на 50,4 % обусловлена вариацией энерговооруженности рабочей силы. При этом увеличение энерговооруженности рабочей силы на 1 л.с. приводит к росту выручки в расчете на одного работника на 18,80 тыс. руб.

7) Оценим значимость коэффициента корреляции, учитывая, что совокупность малая.

$$t_н = |r| \cdot \sqrt{\frac{n-2}{1-r^2}} = 0,71\sqrt{\frac{10-2}{1-0,71^2}} = 2,85.$$

При уровне значимости α=0,05 и числе степеней свободы k=10-2=8 критическое значение $t_{кр.}$=2,31. Т.к. $t_н > t_{кр}$, коэффициент корреляции признается статистически значимым.

8) Коэффициент эластичности составит:

$$Э = 18,80 \cdot \frac{48,36}{747,72} = 1,22.$$

Таким образом, рост энерговооруженности труда на 1 % в данной совокупности организаций приводит к росту выручки на одного работника в среднем на 1,22 %.

Задачи для самостоятельного решения

Задача 6.1 По данным 20 организаций об урожайности озимой пшеницы и качестве почв в баллах (приложение Е) определить: форму связи между признаками; параметры уравнения регрессии; тесноту связи между признаками. Сделать вывод.

Задача 6.2 По данным 20 организаций об урожайности и продолжительности уборки озимой пшеницы (приложение Е) определить: форму связи между признаками; параметры уравнения регрессии; коэффициенты корреляции, детерминации и эластичности. Сделать вывод.

Задача 6.3 По данным 20 организаций об урожайности озимой пшеницы и дозе внесения удобрений (приложение Е) определить: форму связи между признаками; параметры уравнения регрессии; тесноту связи между признаками. Сделать вывод.

Задача 6.4 По данным приложения Е выявить влияние качества почв, продолжительности уборки и количества внесенных минеральных удобрений на урожайность озимой пшеницы. Определить: множественные коэффициенты регрессии и эластичности; парные и множественные коэффициенты корреляции и детерминации. Оценить значимость коэффициентов регрессии и корреляции. Сделать вывод.

Задача 6.5 По данным 15 организаций (приложение Ж) определить: форму связи между признаками; параметры уравнения связи; коэффициенты корреляции, детерминации и эластичности. Оценить значимость коэффициента корреляции.

Исследования провести по следующим парам признаков:

1) энерговооруженность рабочей силы и выручка от продажи продукции на одного работника;

2) фондовооруженность рабочей силы и выручка от продажи продукции на одного работника;

3) фондообеспеченность и выручка от продажи продукции на 1 га сельхозугодий;

4) среднегодовая заработная плата одного работника и выручка от продажи продукции на одного работника;

5) затраты на проданную продукцию и выручка от продажи продукции на 1 га сельхозугодий.

Контрольные тесты для проверки знаний

1 Наиболее тесную связь показывает коэффициент корреляции r_{xy} = … .
   а) r_{xy} = 0,982;
   б) r_{xy} = 0,991;
   в) r_{xy} = 0, 990;
   г) r_{xy} = 0,871.

2 Обратную связь между признаками показывают коэффициенты корреляции :
   а) r_{xy} = - 0,982;
   б) r_{xy} = 0, 990;
   в) r_{xy} = 0,991;
   г) r_{xy} = - 0,871.

3 Прямую связь между признаками показывают коэффициенты корреляции
   а) r_{xy} = - 0,982;
   б) r_{xy} = 0, 990;
   в) r_{xy} = 0,991;
   г) r_{xy} = - 0,871.

4 Простейшим приемом выявления корреляционной связи между двумя признаками яв-

ляется ... .

а) расчет коэффициента корреляции;
б) расчет коэффициента эластичности;
в) построение уравнения корреляционной связи;
г) корреляционное поле.

5 Теснота связи двух признаков при нелинейной зависимости определяется по формуле ... .

а) $\dfrac{\overline{xy} - \overline{x} \cdot \overline{y}}{\sigma_x \sigma_y}$;

б) $\sqrt{\dfrac{\delta^2}{\sigma^2}}$;

в) $\dfrac{\sum (x - \overline{x})(y - \overline{y})}{\sqrt{\sum (x - \overline{x})^2 \sum (y - \overline{y})^2}}$;

6 Для корреляционных связей характерно ... .

а) разным значениям одной переменной соответствуют различные средние значения другой;
б) с изменением значения одной из переменных, другая изменяется строго определенным образом;
в) связь двух величин возможна лишь при условии, что вторая из них зависит только от первой и ни от чего более.

7 Тесноту связи между двумя альтернативными признаками можно измерить с помощью коэффициента:

а) корреляции рангов Спирмена;
б) ассоциации;
в) контингенции;
г) конкордации.

8 Парный коэффициент корреляции показывает тесноту ... .

а) линейной зависимости между двумя признаками на фоне действия остальных, входящих в модель;
б) линейной зависимости между двумя признаками при исключении влияния остальных, входящих в модель;
в) тесноту нелинейной зависимости между двумя признаками;
г) связи между результативным признаком и остальными, включенными в модель.

9 Частный коэффициент корреляции показывает тесноту ... .

а) линейной зависимости между двумя признаками на фоне действия остальных, входящих в модель;
б) нелинейной зависимости;
в) связи между результативным признаком и остальными, включенными в модель;
г) линейной зависимости между двумя признаками при исключении влияния остальных, входящих в модель.

10 Парный коэффициент корреляции может принимать значения ... .

а) от 0 до 1;
б) от -1 до 0;
в) от -1 до 1;
г) любые положительные.

11 Множественный коэффициент корреляции может принимать значения ...

а) от 0 до 1;
б) от -1 до 0;
в) от -1 до 1;
г) любые положительные.

12 Коэффициент детерминации может принимает значения ... .
а) от 0 до 1;
б) от -1 до 0;
в) от -1 до 1;
г) любые положительные.

13 Коэффициент детерминации равен ... коэффициента корреляции
а) квадрату множественного;
б) квадратному корню из множественного;
в) квадрату парного;
г) квадрату частного.

14 Коэффициент детерминации характеризует ... .
а) долю дисперсии результативной переменной, обусловленной влиянием независимых переменных, входящих в модель;
б) дисперсию результативной переменной;
в) долю дисперсии результативной переменной, обусловленной влиянием всех неучтенных в модели факторов;
г) долю дисперсии результативной переменной, обусловленной влиянием наиболее весомого в модели фактора.

15 Прямолинейная связь между факторами исследуется с помощью уравнения регрессии ... .
а) $\overline{y}_x = a_0 + a_1 x$;
б) $\overline{y}_x = a_0 + \frac{a_1}{x}$;
в) $\overline{y}_x = a_0 + a_1 x + a_2 x^2$;
г) $\overline{y}_x = a_0 x^{a_1}$.

Вопросы для самоподготовки

1. Перечислите основные этапы проведения корреляционно-регрессионного анализа.
2. Как определить параметры линейного уравнения регрессии?
3. Как производится расчет показателей тесноты связи между признаками?
4. Что характеризуют коэффициенты регрессии и эластичности?
5. Что характеризуют коэффициенты корреляции рангов?
6. По каким формулам рассчитываются коэффициенты корреляции рангов Спирмена и Кендалла?
7. Что характеризуют и как рассчитываются коэффициенты ассоциации и контингенции?

Тема 7 Ряды динамики

Рядом динамики в статистике называется ряд последовательно расположенных в хронологическом порядке показателей, которые характеризуют состояние, и изменение явлений во времени.

Он состоит из двух элементов: уровней ряда (y) и периодов или моментов времени (t).

Условием правильного построения ряда динамики является сопоставимость уровней ряда во времени. При этом оно достигается едиными правилами сбора данных (обработки данных), единой методикой расчета показателей.

Различают моментные и интервальные ряды динамики. В моментном ряду уровни выражают размер явления на момент времени, а в интервальном – за определенный период времени. Примерами моментного ряда могут служить динамика площадей сельскохозяйственных угодий на 1 января каждого года, остатки удобрений на складе на начало каждого месяца; примерами интервального ряда – динамика валовых сборов сельскохозяйственных культур и их урожайности, количества внесенных удобрений, высеянных семян и т.п.

Ряды динамики по полноте времени могут быть:

- полные, в которых даты или периоды следуют друг за другом в хронологическом порядке (месяц, квартал, полугодие и т.д.)
- неполные, в которых хронологическая последовательность в периодах времени не соблюдается.

Уровни в рядах динамики могут быть выражены абсолютными, относительными и средними величинами.

Абсолютные величины – это форма количественного выражения статистических показателей, характеризующая объем или размер того или иного общественного явления в определенное время на определенной территории.

Относительные величины получаются как частное от деления одной величины (текущей) на другую величину (базисную).

Средние величины – показатели характеризующие типичный размер варьирующего (изменяющегося) признака в расчете на единицу совокупности в данных условиях места и времени.

Графически ряды динамики изображаются линейными, либо столбиковыми диаграммами. По оси абсцисс откладываются показатели времени, а по оси ординат - уровни ряда (либо базисные темпы роста).

Основные показатели рядов динамики могут быть рассчитаны цепным или базисным способом.

Если каждый последующий уровень сравнивается с одним из уровней взятым за базу сравнения, то показатель называется базисным.

Если каждый последующий уровень динамического ряда сравнивается с предыдущим, то показатель называется цепным.

Абсолютный прирост представляет разность двух уровней, он показывает, насколько единиц изменяется текущий уровень ряда по сравнению с предыдущим или базисным.

Коэффициент роста представляет отношение двух уровней, показывает во сколько раз текущий уровень больше базисного или предыдущего, или какую часть его составляет.

Темп роста характеризует относительную скорость развития изучаемого явления во времени. Он показывает, сколько процентов составляет текущий уровень по сравнению с базисным или предыдущим.

Темп прироста показывает, на сколько процентов изменяется текущий уровень по сравнению с базисным или предыдущим.

Значение 1% прироста представляет собой отношение абсолютного прироста цепного к темпу прироста цепному и показывает, сколько абсолютных единиц содержится в 1 % росте или снижения изучаемого показателя.

Условные обозначения:

y_i – текущий (сравниваемый) уровень, i = 1, 2, 3,…, n;

y_1 – уровень, принятый за постоянную базу сравнения (обычно начальный);

y_n – конечный уровень.

Средний уровень ряда динамики определяется в зависимости от имеющихся данных, а также от конкретного вида ряда динамики.

Для интервальных рядов динамики с равными периодами времени средний уровень определяется по формуле простой арифметической.

Таблица 7.1- Расчет показателей ряда динамики

Показатель	Метод расчета	
	базисный	цепной
Абсолютный прирост (A)	$A_{б_i} = y_i - y_1$ (7.1)	$A_{ц_i} = y_i - y_{i-1}$ (7.2)
Коэффициент роста (K_p)	$K_{рб_i} = \frac{y_i}{y_1}$ (7.3)	$K_{рц_i} = \frac{y_i}{y_{i-1}}$ (7.4)
Темп роста (T_p)	$T_{рб_i} = K_{рб_i} \cdot 100\ \%$ (7.5)	$T_{рц_i} = K_{рц_i} \cdot 100\ \%$ (7.6)
Темп прироста ($T_{пр}$)	$T_{прб_i} = T_{рб_i} - 100\ \%$ (7.7)	$T_{прц_i} = T_{рц_i} - 100\ \%$ (7.8)
Абсолютное значение 1 % прироста ($Зн.1\%$)	$Зн.1\% = 0{,}01 \cdot y_{i-1}$ или $Зн.1\% = \frac{A_ц}{T_{пр\ ц}}$ (7.9)	

В моментных рядах динамики, если имеются полные исчерпывающиеся данные о развитии явлений, то средний уровень определяется по формуле средней арифметической взвешенной.

Если промежутки времени между датами одинаковы, то средний уровень ряда динамики рассчитывается по формуле средней хронологической.

Для характеристики интенсивности развития явления за длительный период времени рассчитываются средние показатели динамики (таблица 7.2).

Средние показатели динамики исчисляются одинаково для интервальных и моментных рядов, исключение составляет лишь расчет среднего уровня ряда.

Основным условием построения и анализа ряда динамики является сопоставимость уровней во времени.

К несопоставимости приводит изменение состава или территориальных границ изучаемой совокупности, переход к другим единицам измерения, инфляционные процессы. Несопоставимыми ряды динамики являются и в том случае, если они составлены из неодинаковых по продолжительности времени периодов.

При обнаружении несопоставимости уровней ряда должна применяться процедура смыкания, если невозможен их прямой пересчет.

Таблица 7.2 – Расчет средних показателей ряда динамики

Показатель	Метод расчета
Средний уровень ($\bar{y}$) а) интервального ряда	$\bar{y} = \frac{\Sigma y}{n}$ (7.10)
б) моментного ряда с равными интервалами	$\bar{y} = \frac{\frac{1}{2}y_1 + y_2 + y_3 ... + y_{n-1} + \frac{1}{2}y_n}{n-1}$ (7.11)
в) моментного ряда с неравными интервалами	$\bar{y} = \frac{\Sigma yt}{\Sigma t}$ (7.12)
Средний абсолютный прирост ($\bar{A}$)	$\bar{A} = \frac{y_n - y_1}{n-1}$ или $\bar{A} = \frac{\Sigma A_ц}{n-1}$ (7.13)
Средний коэффициент роста ($\bar{K}_р$)	$\overline{K_р} = \sqrt[n-1]{\Pi K_ц}$ или $\overline{K_р} = \sqrt[n-1]{\frac{y_n}{y_1}}$ (7.14)
Средний темп роста ($\overline{T_р}$), %	$\overline{T_р} = \overline{K_р} \cdot 100\ \%$ (7.15)
Средний темп прироста ($\overline{T_{пр}}$), %	$\overline{T_{пр}} = \overline{T_р} - 100\ \%$ или $\overline{T_{пр}} = (\overline{K_р} - 1) \cdot 100\%$ (7.16)
Среднее значение 1% прироста ($\overline{Зн.1\%}$)	$\overline{Зн.1\%} = \frac{\bar{A}}{\overline{T_{пр}}}$ (7.17)

Смыкание может быть произведено двумя способами.

1 способ. Данные за предшествующие периоды умножаются на коэффициент перехода, который определяется как отношение показателей на тот момент времени, когда произошло изменение условий формирования уровней ряда.

2 способ. Уровень переходного периода принимается для второй части ряда за 100 % и от этого уровня определяются соответствующие показатели. При этом получается сопоставимый ряд относительных величин.

Иногда в динамических рядах отсутствуют промежуточные или последующие уровни. Их можно исчислить с помощью методов интерполяции (нахождение промежуточного неизвестного уровня при наличии известных соседних уровней) и экстраполяции (нахождение уровней за пределами изучаемого ряда, т. е. продление в будущее тенденции, наблюдавшейся в прошлом, или в прошлое на основании текущих уровней).

Часто требуется выявить общую тенденцию изменения явления во времени. Если по уровням ряда динамики она четко не просматривается, то для этого применяют различные методы: укрупнения временных интервалов (периодов); скользящих средних; аналитического выравнивания.

Метод укрупнения временных интервалов (периодов) заключается в том, что вместо первоначальных уровней рассчитываются и сравниваются средние уровни за укрупненные периоды времени (месячные – в квартальные, квартальные – в годовые и т.д.).

Метод сглаживания ряда динамики с помощью скользящей средней заключается в том, что фактические уровни заменяются средними арифметическими уровнями за укрупненные периоды. Расчет средних ведется способом скольжения, т. е. постепенным исключением из принятого периода скольжения первого уровня и включением следующего. При этом, посредством осреднения эмпирических данных, индивидуальные колебания погашаются, общая тенденция развития явления выражается в виде плавной линии (теоретические уровни).

Метод аналитического выравнивания состоит в подборе для данного ряда динамики такой теоретической линии, которая выражает основные черты или закономерности изменения уровней явления. Чаще всего при выравнивании используют линейное уравнение

$$\hat{y}_t = a + bt, \tag{7.18}$$

где a – свободный член уравнения;

b – коэффициент;

t – порядковый номер года.

Параметры уравнения определяются методом наименьших квадратов, путем составления и решения системы уравнений

$$\begin{cases} \sum y = na + b\sum t, \\ \sum yt = a\sum t + b\sum t^2. \end{cases} \tag{7.19}$$

Далее проводится оценка надежности уравнения тренда с помощью F – критерия Фишера.

Пример 7.1 По имеющимся данным об урожайности чайного листа рассчитать показатели ряда динамики. Сделать вывод.

Решение. Проведем расчет показателей ряда динамики по соответствующим формулам. Результаты расчетов представлены в таблице 7.3.

Таблица 7.3 - Расчет показателей ряда динамики урожайности чайного листа

Год	Урожайность, ц/га	Абсолютный прирост, ц		Коэффициент роста		Темп роста, %		Темп прироста, %		Значение 1% прироста, ц
		базисный	цепной	базисный	цепной	базисный	цепной	базисный	цепной	
		$А_б$	$А_ц$	$К_{р\,б}$	$К_{р\,ц}$	$Т_{р\,б}$	$Т_{р\,ц}$	$Т_{пр\,б}$	$Т_{пр\,ц}$	$Зн.1\%$
2015	1,0	–	–	1,000	1,000	100,0	100,0	–	–	–
2016	0,9	-0,1	-0,1	0,900	0,900	90,0	90,0	-10,0	-10,0	0,010
2017	3,6	2,6	2,7	3,600	4,000	360,0	400,0	260,0	300,0	0,009
2018	4,8	3,8	1,2	4,800	1,333	480,0	133,3	380,0	33,3	0,036
2019	5,0	4,0	0,2	5,000	1,042	500,0	104,2	400,0	4,2	0,048
В среднем	3,06	1,0		1,495		149,5		49,5		0,020

Вывод: расчеты показали, что средняя урожайность чайного листа в динамике за пять лет составила 3,06 ц/га. При этом ежегодно она повышалась в среднем на 1,0 ц/га или на 49,5 %. Один процент прироста соответствовал в среднем 0,02 ц/га.

Пример 7.2 Определить тенденцию изменения урожайности чайного листа в хозяйствах Краснодарского края различными методами. Сделать вывод.

Решение. Для упрощения расчетов начало отсчета времени t было перенесено в середину ряда динамики. Так как $\Sigma t = 0$, то параметры *«а»* и *«b»* найдем следующим образом:

$$a = \frac{\sum y}{n} = \frac{30{,}6}{9} = 3{,}4\,; \qquad b = \frac{\sum yt}{\sum t^2} = \frac{-2{,}0}{60} = -0{,}033\,.$$

Тогда уравнение прямой имеет вид: $\hat{y}_t = 3{,}4 - 0{,}033\,t$.

Подставим в данное уравнение значения t и найдем теоретические (выравненные) уровни урожайности чайного листа (последний столбец таблицы 7.4). Фактические и теоретические значения урожайности изобразим графически (рисунок 7.1).

Таблица 7.4 – Вспомогательная таблица для выявления общей тенденции изменения урожайности чайного листа

Год	Урожайность, ц/га	Метод укрупнения периодов	Метод скользящей средней		Метод аналитического выравнивания			
			за три года		номер года	квадрат номера года	произведение параметров	выравненные значения
	y	средняя по трёхлетиям	сумма	средняя	t	t^2	yt	$\hat{y}_t = a + bt$
2011	5,7	4,37			-4	16	-22,8	3,53
2012	4,7		13,1	4,4	-3	9	-14,1	3,50
2013	2,7		9,6	3,2	-2	4	-5,4	3,47
2014	2,2	1,37	5,9	2,0	-1	1	-2,2	3,43
2015	1,0		4,1	1,4	0	0	0,0	3,40
2016	0,9		5,5	1,8	1	1	0,9	3,37
2017	3,6	4,47	9,3	3,1	2	4	7,2	3,33
2018	4,8		13,4	4,5	3	9	14,4	3,30
2019	5,0		-	-	4	16	20,0	3,27
Итого	30,6	-	-	-	0	60	-2,0	30,60

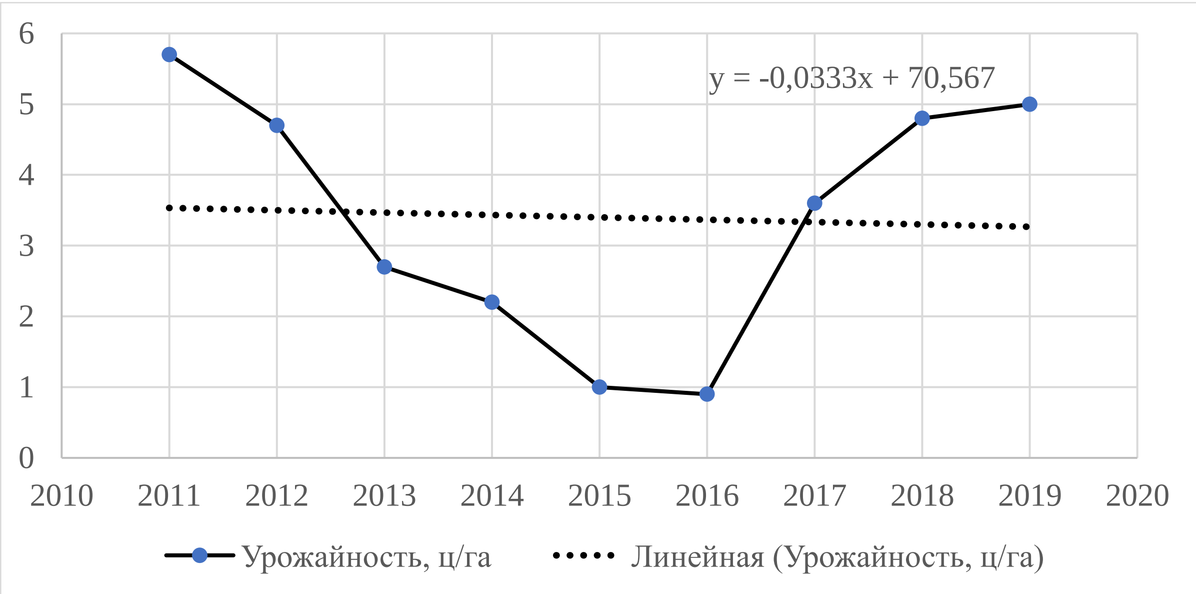

Рисунок 7.1 -Динамика урожайности чайного листа, ц/га

Вывод: расчеты показали, что средняя урожайность чайного листа за 2011 – 2019 гг. составила 3,4 ц/га. В среднем она ежегодно снижалась на 0,033 ц/га. На графике наглядно видна выраженная тенденция к снижению урожайности исследуемой культуры.

Пример 7.3 В 2016 г. к агрообъединению примкнуло два хозяйства, что привело к несопоставимости ряда динамики (таблица 7.5). Привести его к сопоставимому виду, применив смыкание динамического ряда.

Таблица 7.5 – Динамика объемов производства продукции в агрообъединении, млрд руб.

Объем производства продукции	Год					
	2014	2015	2016	2017	2018	2019
При старом составе	19,7	20,2	21,2			
При новом составе			22,8	24,6	25,2	26,1
Смыкание по 1 способу	21,2	21,7	22,8	24,6	25,2	26,1
Смыкание по 2 способу, %	92,9	95,3	100,0	107,9	110,5	114,4

Смыкание по 1 способу расчета:

$$\frac{22,8}{21,2} = 1,0755; \qquad 20,2 \cdot 1,0755 = 21,7; \qquad 19,7 \cdot 1,0755 = 21,2;$$

Смыкание по 2 способу расчета:

$$\frac{20,2}{21,2} \cdot 100 = 95,3; \qquad \frac{19,7}{21,2} \cdot 100 = 92,9; \qquad \frac{24,6}{22,8} \cdot 100 = 107,9;$$

$$\frac{25,2}{22,8} \cdot 100 = 110,5; \qquad \frac{26,1}{22,8} \cdot 100 = 114,4.$$

Вывод: расчеты показали, что изменение состава агрообъединения привело к росту объема производства продукции. При этом в динамике за шесть лет он увеличился на 4,9 млрд руб. или на 21,5 процентных пункта.

Задачи для самостоятельного решения

Задача 7.1 По данным приложения З по своему варианту за 5 лет определить базисные и цепные показатели ряда динамики, показатели динамики в среднем за период. Расчеты представить в табличной форме, сделать вывод.

Задача 7.2 По данным приложения З выявить общую тенденцию урожайности сельскохозяйственных культур, используя приемы укрупнения периодов, трехлетней скользящей средней и аналитического выравнивания. Изобразить на графике фактические и выравненные (теоретические) уровни. Сделать вывод по результатам расчетов.

Задача 7.3 По данным задачи 7.2 спрогнозировать урожайность на предстоящие два года, применив среднегодовой абсолютный прирост, среднегодовой темп роста и результаты аналитического выравнивания.

Задача 7.4 Остатки минеральных удобрений на складе на 01.01 составили 220 т, на 01.02 – 190 т, на 01.03 – 200 т, на 01.04 – 260 т. Определить средние остатки удобрений за первый квартал.

Задача 7.5 Поголовье коров в сельскохозяйственной организации на 01.01 составляло 800 гол., 15.01 было выбраковано 30 гол., 05.02 переведено из нетелей в основное стадо 55 гол., 24.02 куплено 10 гол., 12.03 продано 15 гол., 21.03 выбраковано 25 гол. Определить среднее поголовье коров за первый квартал.

Задача 7.6 Площадь земель, предоставленная крестьянским (фермерским) хозяйствам в 2012 г., возросла по сравнению с предыдущим годом на 1,9 %, в 2013 г. – на 6,6 и в 2014 г. – на 6,1 %. Определить средний процент прироста земель и их площадь в 2011, 2012 и 2013 гг., если известно, что в 2014 г. общая площадь крестьянских хозяйств составила 13845 тыс. га.

Задача 7.7 До 2016 г. в состав производственного объединения входило 20 организаций. В 2016 г. в него влилось еще 4 организации, и оно стало объединять 24 организации. Провести смыкание ряда динамики, используя данные таблицы.

Объем реализации, млн руб.	2014 г.	2015 г.	2016 г.	2017 г.	2018 г.	2019 г.
По 20 организациям	491,6	501,1	510,2	-	-	-
По 24 организациям	-	-	580,5	610,0	612,9	615,5

Задача 7.8 Используя взаимосвязь показателей динамики, определить уровни ряда динамики и недостающие в таблице базисные показатели динамики по следующим данным об урожайности озимой пшеницы.

Год	Урожайность озимой пшеницы, ц/га	Базисные показатели динамики			Значение 1% прироста, ц/га
		абсолютный прирост, ц	темп роста, %	темп прироста, %	
2013	55,1	-		-	-
2014		- 2,8			
2015			110,3		
2016					
2017				17,1	0,633
2018			121,1		
2019		13,5			

Задача 7.9 По данным таблицы провести смыкание динамического ряда.

Валовой сбор зерна, тыс. ц	2013г.	2014г.	2015г.	2016г.	2017 г.	2018г.	2019г.
В старых границах	44,5	45,0	48,2	-	-	-	-
В новых границах	-	-	60,0	63,6	61,1	64,2	65,6

Контрольные тесты для проверки знаний

1 Среднегодовой коэффициент роста в рядах динамики определяется по формуле средней…

а) арифметической;

б) квадратической;
в) геометрической;
г) хронологической.

2 Если за два анализируемых периода времени темп роста объемов производства продукции составил 140%, то это значит, что объем производства увеличился…
а) на 40%;
б) в 4 раза;
в) на 140%;
г) в 14 раз.

3 В рядах динамики, для расчета среднего темпа роста применяется формула средней …
а) арифметической;
б) квадратической;
в) геометрической;
г) Хронологической.

4 При расчете среднего коэффициента роста с помощью средней геометрической подкоренное выражение представляет собой __________ цепных коэффициентов роста.
а) частное;
б) разность;
в) сумму;
г) произведение

5 Ряд динамики, характеризующий экспорт страны по каждому году за период с 2010 по 2017 годы, по виду относится к __________ рядам динамики.
а) произвольным;
б) интервальным
в) моментным

6 Ряд динамики, показатели которого характеризуют наличие на предприятии остатков оборотных средств на первое число каждого месяца 2017 г., называется …
а) моментным с равными интервалами;
б) моментным с не равными интервалами;
в) интервальным с неравными интервалами;
г) интервальным с равными интервалами;

7 Абсолютный прирост в рядах динамики исчисляется как __________ уровень ряда.
а) частное;
б) разность;
в) сумму;
г) произведение

8 Ежеквартальные темпы прироста должны быть в среднем равны ... % (с точностью до 0,1 %), чтобы выручка от реализации продукции в четвертом квартале текущего года по сравнению с четвертым кварталом предыдущего года возросла с 600 тыс. руб. до 798,6 тыс. руб. 7,4;

9 По формуле $T_p = \frac{y_i}{y_0}$ определяется … .
а) базисный темп роста;
б) цепной темп роста;
в) базисный темп прироста;
г) абсолютное значение 1% прироста.

10 По формуле $T_p = \dfrac{y_i}{y_{i-1}}$ определяется…..

а) базисный темп роста;
б) цепной темп роста;
в) базисный темп прироста;
г) абсолютное значение 1% прироста.

11 Средний уровень моментного ряда динамики с равными временными промежутками исчисляется по формуле средней ... .

а) арифметической простой;
б) арифметической взвешенной;
в) хронологической простой;
г) гармонической взвешенной.

12 Средний уровень моментного ряда динамики с неравными временными промежутками исчисляется по формуле средней ...

а) арифметической простой;
б) арифметической взвешенной;
в) хронологической простой;
г) хронологической взвешенной.

13 Средний уровень интервального ряда динамики с равными временными промежутками исчисляется по формуле средней ... .

а) арифметической простой;
б) арифметической взвешенной;
в) хронологической простой;
г) гармонической взвешенной.

14 Средний уровень интервального ряда динамики с неравными временными промежутками исчисляется по формуле средней ... .

а) арифметической простой;
б) арифметической взвешенной;
в) хронологической простой;
г) гармонической взвешенной.

15 Ряд динамики – это ряд чисел, характеризующий

а) состав изучаемого явления;
б) состояние и изменение явлений во времени;
в) изменение явлений в пространстве;
г) современное состояние явлений.

Вопросы для самоподготовки

1. Ряды динамики, их элементы и правила построения.
2. Виды рядов динамики.
3. Показатели ряда динамики и порядок их расчета.
4. Средние показатели ряда динамики и порядок их расчета.
5. Приемы выявления основной тенденции развития в рядах динамики.
6. Что понимается под интерполяцией и экстраполяцией ряда динамики?
7. Как осуществляется статистическое прогнозирование уровней ряда динамики?
8. Как проводится смыкание рядов динамики?

Список литературы

1.Бондаренко П.С. Теория вероятностей и математическая статистика (для бакалавров) [Электронный ресурс]: учебное пособие: учебное пособие /П.С. Бондаренко, Г.В. Горелова, И.А. Кацко. — Электрон. текстовые данные. — Москва: КноРус, 2019. — 389 с. — 978-5-406-06704-8. – Режим доступа: https://www.book.ru/book/930219

2. Статистические методы анализа данных: учеб. пособие / И. А. Кацко, Н. Х. Ворокова, А. Е. Жминько, А. Е. Сенникова – Краснодар: Краснодарский ЦНТИ – филиал ФГБУ «РЭА» Минэнерго России, 2017. – 203 с. Режим доступа: https://edu.kubsau.ru/file.php/120/Statisticheskie_metody_analiza_dannykh_423865_v1_.PDF

3. Блатов И.А. Теория вероятностей и математическая статистика [Электронный ресурс]: учебное пособие / И.А. Блатов, О.В. Старожилова. — Электрон. текстовые данные. — Самара: Поволжский государственный университет телекоммуникаций и информатики, 2017. — 276 с. — 2227-8397. — Режим доступа: http://www.iprbookshop.ru/75412.html

4. Колемаев В.А. Теория вероятностей и математическая статистика [Электронный ресурс] : учебник для вузов / В.А. Колемаев, В.Н. Калинина. — 2-е изд. — Электрон. текстовые данные. — М. : ЮНИТИ-ДАНА, 2017. — 352 с. — 5-238-00560-1. — Режим доступа: http://www.iprbookshop.ru/71075.html

5. Логинов В.А. Теория вероятностей и математическая статистика [Электронный ресурс]: сборник задач / В.А. Логинов. — Электрон. текстовые данные. — М.: Московская государственная академия водного транспорта, 2017. — 72 с. — 2227-8397. — Режим доступа: http://www.iprbookshop.ru/76719.html

6. Теория вероятностей и математическая статистика [Электронный ресурс]: учебник-практикум / А.В. Браилов [и др.]. — Электрон. текстовые данные. — Ижевск: Регулярная и хаотическая динамика, Институт компьютерных исследований, 2016. — 414 с. — 978-5-4344-0415-0. — Режим доступа: http://www.iprbookshop.ru/69368.html

Приложения

Приложение А

Значения функций $\varphi(x) = \dfrac{1}{\sqrt{2\pi}} \cdot e^{-\frac{1}{2}x^2}$ **и** $\Phi(x) = \dfrac{1}{\sqrt{2\pi}} \int\limits_0^x e^{-\frac{1}{2}z^2} dz$.

x	$\varphi(x)$	$\Phi(x)$	x	$\varphi(x)$	$\Phi(x)$	x	$\varphi(x)$	$\Phi(x)$
0,00	0,3989	0,0000	0,40	0,3683	0,1554	0,80	0,2897	0,2881
01	3989	0040	41	3668	1591	81	2874	2910
02	3989	0080	42	3653	1628	82	2850	2939
03	3988	0120	43	3637	1664	83	2827	2967
04	3986	0160	44	3621	1700	84	2803	2995
05	3984	0199	45	3605	1736	85	2780	3023
06	3982	0239	46	3589	1772	86	2756	3051
07	3980	0279	47	3572	1808	87	2732	3078
08	3977	0319	48	3555	1844	88	2709	3106
09	3973	0359	49	3538	1879	89	2685	3133
0,10	0,3970	0,0398	0,50	0,3521	0,1915	0,90	0,2661	0,3159
11	3965	0438	51	3503	1950	91	2637	3186
12	3961	0478	52	3485	1985	92	2613	3212
13	3956	0517	53	3467	2019	93	2589	3238
14	3951	0557	54	3448	2054	94	2565	3264
15	3945	0596	55	3429	2088	95	2541	3289
16	3939	0636	56	3410	2123	96	2516	3315
17	3932	0675	57	3391	2157	97	2492	3340
18	3925	0714	58	3372	2190	98	2468	3365
19	3918	0753	59	3352	2224	99	2444	3389
0,20	0,3910	0,0793	0,60	0,3332	0,2257	1,00	0,2420	0,3413
21	3902	0832	61	3312	2291	01	2396	3438
22	3894	0871	62	3292	2324	02	2371	3461
23	3885	0910	63	3271	2357	03	2347	3485
24	3876	0948	64	3251	2389	04	2323	3508
25	3867	0987	65	3230	2422	05	2299	3531
26	3857	1026	66	3209	2454	06	2275	3554
27	3847	1064	67	3187	2486	07	2251	3577
28	3836	1103	68	3166	2517	08	2227	3599
29	3825	1141	69	3144	2549	09	2203	3621
0,30	0,3814	0,1179	0,70	0,3123	0,2580	1,10	0,2179	0,3643
31	3802	1217	71	3101	2611	11	2155	3665
32	3790	1255	72	3079	2642	12	2131	3686
33	3778	1293	73	3056	2673	13	2107	3708
34	3765	1331	74	3034	2703	14	2083	3729
35	3752	1368	75	3011	2734	15	2059	3749
36	3739	1406	76	2989	2764	16	2036	3770
37	3726	1443	77	2966	2794	17	2012	3790
38	3712	1480	78	2943	2823	18	1989	3810

x	$\varphi(x)$	$\Phi(x)$	x	$\varphi(x)$	$\Phi(x)$	x	$\varphi(x)$	$\Phi(x)$
39	3697	1517	79	2920	2852	19	1965	3830
1,20	0,1942	0,3849	1,70	0,0940	0,4554	2,40	0,0224	0,4918
21	1919	3869	71	0925	4564	42	0213	4922
22	1895	3888	72	0909	4573	44	0203	4927
23	1872	3907	73	0893	4582	46	0194	4931
24	1849	3925	74	0878	4591	48	0184	4934
25	1826	3944	75	0863	4599	50	0175	4938
26	1804	3962	76	0848	4608	52	0167	4941
27	1781	3980	77	0833	4616	54	0158	4945
28	1758	3997	78	0818	4625	56	0151	4948
29	1736	4015	79	0804	4633	58	0143	4951
1,30	0,1714	0,4032	1,80	0,0790	0,4641	2,60	0,0136	0,4953
31	1691	4049	81	0775	4649	62	0129	4956
32	1669	4066	82	0761	4656	64	0122	4959
33	1647	4082	83	0748	4664	66	0116	4961
34	1626	4099	84	0734	4671	68	0110	4963
35	1604	4115	85	0721	4678	70	0104	4965
36	1582	4131	86	0707	4686	72	0099	4967
37	1561	4147	87	0694	4693	74	0093	4969
38	1539	4162	88	0681	4699	76	0088	4971
39	1518	4177	89	0669	4706	78	0084	4973
1,40	0,1497	0,4192	1,90	0,0656	0,4713	2,80	0,0079	0,4974
41	1476	4207	91	0644	4719	82	0075	4976
42	1456	4222	92	0632	4726	84	0071	4977
43	1435	4236	93	0620	4732	86	0067	4979
44	1415	4251	94	0608	4738	88	0063	4980
45	1394	4265	95	0596	4744	90	0060	4981
46	1374	4279	96	0584	4750	92	0056	4982
47	1354	4292	97	0573	4756	94	0053	4984
48	1334	4306	98	0562	4761	96	0050	4985
49	1315	4319	99	0551	4767	98	0047	4986
1,50	0,1295	0,4332	2,00	0,0540	0,4772	3,00	0,00443	0,49865
51	1276	4345	02	0519	4783			
52	1257	4357	04	0498	4793	3,10	00327	49903
53	1238	4370	06	0478	4803	3,20	00238	49931
54	1219	4382	08	0459	4812			
55	1200	4394	10	0440	4821	3,30	00172	49952
56	1182	4406	12	0422	4830	3,40	00123	49966
57	1163	4418	14	0404	4838			
58	1145	4429	16	0387	4846	3,50	00087	49977
59	1127	4441	18	0371	4854			
1,60	0,1109	0,4452	2,20	0,0355	0,4861	3,60	00061	499841
61	1092	4463	22	0339	4868	3,70	00042	49989
62	1074	4474	24	0325	4875	3,80	00029	499928

x	$\varphi(x)$	$\Phi(x)$	x	$\varphi(x)$	$\Phi(x)$	x	$\varphi(x)$	$\Phi(x)$
63	1057	4484	26	0310	4881			
64	1040	4495	28	0297	4887	3,90	00020	49995
65	1023	4505	30	0283	4893	4,00	0,00013	499968
66	1006	4515	32	0270	4898		38	
67	0989	4525	34	0258	4904	4,50		499997
68	0973	4535	36	0246	4909	5,00	0000160	4999999
69	0957	4545	38	0235	4913		0000015	

Приложение Б

Критические точки распределения t-Стьюдента

Число степеней свободы k	Уровень значимости α (двусторонняя критическая область)					
	0,10	0,05	0,02	0,01	0,002	0,001
1	6,31	12,71	31,82	63,7	318,31	636,62
2	2,92	4,30	6,97	9,92	22,33	31,60
3	2,35	3,18	4,54	5,84	10,22	12,92
4	2,13	2,78	3,75	4,60	7,17	8,61
5	2,01	2,57	3,37	4,03	5,89	6,87
6	1,94	2,45	3,14	3,71	5,21	5,96
7	1,90	2,37	3,00	3,50	4,79	5,41
8	1,86	2,31	2,90	3,36	4,50	5,04
9	1,83	2,26	2,82	3,25	4,30	4,78
10	1,81	2,23	2,76	3,17	4,14	4,59
11	1,80	2,20	2,72	3,11	4,03	4,44
12	1,78	2,18	2,68	3,06	3,93	4,32
13	1,77	2,16	2,65	3,01	3,85	4,22
14	1,76	2,15	2,62	2,98	3,79	4,14
15	1,75	2,13	2,60	2,95	3,73	4,07
16	1,75	2,12	2,58	2,92	3,69	4,02
17	1,74	2,11	2,57	2,90	3,65	3,97
18	1,73	2,10	2,55	2,88	3,61	3,92
19	1,73	2,09	2,54	2,86	3,58	3,88
20	1,73	2,09	2,53	2,85	3,55	3,85
21	1,72	2,08	2,52	2,83	3,53	3,82
22	1,72	2,07	2,51	2,82	3,51	3,79
23	1,71	2,07	2,50	2,81	3,49	3,77
24	1,71	2,06	2,49	2,80	3,47	3,75
25	1,71	2,06	2,49	2,79	3,45	3,73
26	1,71	2,06	2,48	2,78	3,44	3,71
27	1,71	2,05	2,47	2,77	3,42	3,69
28	1,70	2,05	2,46	2,76	3,41	3,67
29	1,70	2,05	2,46	2,76	3,40	3,66
30	1,70	2,04	2,46	2,75	3,39	3,65
40	1,68	2,02	2,42	2,70	3,31	3,55
60	1,67	2,00	2,39	2,66	3,23	3,46
120	1,66	1,98	2,36	2,62	3,16	3,37
∞	1,65	1,96	2,33	2,58	3,09	3,29
	0,05	0,025	0,01	0,005	0,001	0,0005
	Уровень значимости α (односторонняя критическая область)					

Приложение В

Критические точки распределения F Фишера – Снедекора
(k_1 – число степеней свободы большей дисперсии,
k_2 – число степеней свободы меньшей дисперсии)

Уровень значимости a=0,05										
	k_1									
k_2	1	2	3	4	5	6	8	12	24	∞
1	161,4	199,5	215,7	224,6	230,2	234,0	238,9	243,9	249,1	254,3
2	18,51	19,00	19,16	19,25	19,30	19,33	19,37	19,41	19,45	19,50
3	10,13	9,55	9,28	9,12	9,01	8,94	8,85	8,74	8,64	8,53
4	7,71	6,94	6,59	6,39	6,26	6,16	6,04	5,91	5,77	5,63
5	6,61	5,79	5,41	5,19	5,05	4,95	4,82	4,68	4,53	4,37
6	5,99	5,14	4,76	4,53	4,39	4,28	4,15	4,00	3,84	3,67
7	5,59	4,74	4,35	4,12	3,97	3,87	3,73	3,57	3,41	3,23
8	5,32	4,46	4,07	3,84	3,69	3,58	3,44	3,28	3,12	2,93
9	5,12	4,26	3,86	3,63	3,48	3,37	3,23	3,07	2,90	2,71
10	4,96	4,10	3,71	3,48	3,33	3,22	3,07	2,91	2,74	2,54
11	4,84	3,98	3,59	3,36	3,20	3,09	3,95	2,79	2,61	2,40
12	4,75	3,89	3,49	3,26	3,11	3,00	2,85	2,69	2,51	2,30
13	4,67	3,81	3,41	3,18	3,03	2,92	2,77	2,60	2,42	2,21
14	4,60	3,74	3,34	3,11	2,96	2,85	2,70	2,53	2,35	2,13
15	4,54	3,68	3,29	3,06	2,90	2,79	2,64	2,48	2,29	2,07
16	4,49	3,63	3,24	3,01	2,85	2,74	2,59	2,42	2,24	2,01
17	4,45	3,59	3,20	2,96	2,81	2,70	2,55	2,38	2,19	2,96
18	4,41	3,55	3,16	2,93	2,77	2,66	2,51	2,34	2,15	1,92
19	4,38	3,52	3,13	2,90	2,74	2,63	2,48	2,31	2,11	1,88
20	4,35	3,49	3,10	2,87	2,71	2,60	2,45	2,28	2,08	1,84
21	4,32	3,47	3,07	2,84	2,68	2,57	2,42	2,25	2,05	1,81
22	4,30	3,44	3,05	2,82	2,66	2,55	2,40	2,23	2,03	1,78
23	4,28	3,42	3,03	2,80	2,64	2,53	2,37	2,20	2,01	1,76
24	4,26	3,40	3,01	2,78	2,62	2,51	2,36	2,18	1,98	1,73
25	4,24	3,39	2,99	2,76	2,60	2,49	2,34	2,16	1,96	1,71
26	4,23	3,37	2,98	2,74	2,59	2,47	2,32	2,15	1,95	1,69
27	4,21	3,35	2,96	2,73	2,57	2,46	2,31	2,13	1,93	1,67
28	4,20	3,34	2,95	2,71	2,56	2,45	2,29	2,12	1,91	1,65
29	4,18	3,33	2,93	2,70	2,55	2,43	2,28	2,10	1,90	1,64
30	4,17	3,32	2,92	2,69	2,53	2,42	2,27	2,09	1,89	1,62
60	4,00	3,15	2,76	2,53	2,37	2,25	2,10	1,92	1,70	1,39
∞	3,84	3,00	2,60	2,37	2,21	2,10	2,94	1,75	1,52	1,00

Приложение Г

Критические точки распределения χ^2

Число степеней свободы k	Уровень значимости α							
	0,99	0,975	0,95	0,9	0,1	0,05	0,025	0,01
1	0,00016	0,00098	0,00393	0,0158	2,71	3,84	5,02	6,63
2	0,0201	0,0506	0,103	0,211	4,61	5,99	7,38	9,21
3	0,115	0,216	0,352	0,584	6,25	7,81	9,35	11,34
4	0,297	0,484	0,711	1,064	7,78	9,49	11,14	13,28
5	0,554	0,831	1,145	1,61	9,24	11,07	12,83	15,09
6	0,872	1,24	1,64	2,20	10,64	12,59	14,45	16,81
7	1,24	1,69	2,17	2,83	12,02	14,07	16,01	18,48
8	1,65	2,18	2,73	3,49	13,36	15,51	17,53	20,09
9	2,09	2,70	3,33	4,17	14,68	16,92	19,02	21,67
10	2,56	3,25	3,94	4,87	15,99	18,31	20,48	23,21
11	3,05	3,82	4,57	5,58	17,28	19,68	21,92	24,73
12	3,57	4,40	5,23	6,30	18,55	21,03	23,34	26,22
13	4,11	5,01	5,89	7,04	19,81	22,36	24,74	27,69
14	4,66	5,63	6,57	7,79	21,06	23,68	26,12	29,14
15	5,23	6,26	7,26	8,55	22,31	25,00	27,49	30,58
16	5,81	6,91	7,96	9,31	23,54	26,30	28,85	32,00
17	6,41	7,56	8,67	10,09	24,77	27,59	30,19	33,41
18	7,01	8,23	9,39	10,86	25,99	28,87	31,53	34,81
19	7,63	8,91	10,12	11,65	27,20	30,14	32,85	36,19
20	8,26	9,59	10,85	12,44	28,41	31,41	34,17	37,57
21	8,90	10,28	11,59	13,24	29,62	32,67	35,48	38,93
22	9,54	10,98	12,34	14,04	30,81	33,92	36,78	40,29
23	10,20	11,69	13,09	14,85	32,01	35,17	38,08	41,64
24	10,86	12,40	13,85	15,66	33,20	36,42	39,36	42,98
25	11,52	13,12	14,61	16,47	34,38	37,65	40,65	44,31
26	12,20	13,84	15,38	17,29	35,56	38,89	41,92	45,64
27	12,88	14,57	16,15	18,11	36,74	40,11	43,19	46,96
28	13,56	15,31	16,93	18,94	37,92	41,34	44,46	48,28
29	14,26	16,05	17,71	19,77	39,09	42,56	45,72	49,59
30	14,95	16,79	18,49	20,60	40,26	43,77	46,98	50,89
40	22,16	24,43	26,51	29,05	51,81	55,76	59,34	63,69
50	29,71	32,36	34,76	37,69	63,17	67,50	71,42	76,15
60	37,48	40,48	43,19	46,46	74,40	79,08	83,30	88,38
70	45,44	48,76	51,74	55,33	85,53	90,53	95,02	100,4
80	53,54	57,15	60,39	64,28	96,58	101,9	106,6	112,3
90	61,75	65,61	69,13	73,29	107,6	113,1	118,1	124,1
100	70,06	74,22	77,93	82,36	118,5	124,3	129,6	135,8

Приложение Д

Посевные площади сельскохозяйственных культур
в организациях Краснодарского края, га

№ варианта	2018 г.			2019 г.			План на 2019 г.		
	зерновые	технические	кормовые	зерновые	технические	кормовые	зерновые	технические	кормовые
1	4500	2600	2900	4800	2200	2800	4600	2400	2900
2	3400	1300	2600	3500	1800	2000	3600	1500	2100
3	3200	900	1500	3300	1000	1400	3200	1100	1300
4	4300	2200	2700	4500	2000	2600	4600	2000	2500
5	2800	900	1600	3000	900	1500	2900	1000	1400
6	3400	1300	2000	3200	1500	2100	3300	1400	2000
7	3600	1700	2200	3600	1900	2100	3700	1800	2300
8	5000	2000	3000	5100	2000	2800	5000	2200	3100
9	4800	1500	2700	4700	1600	2800	4800	1700	2600
10	3000	1100	1900	3200	1000	1900	3300	1100	1800
11	4600	1400	2700	4800	1500	2500	4700	1600	2500
12	3400	1300	2900	3500	1500	2800	3400	1500	2800
13	5400	2200	2900	5300	2300	3000	5400	2500	2800
14	2500	800	1400	2400	1000	1500	2500	900	1500
15	3100	1600	1900	3300	1500	2000	3200	1600	1800
16	4900	2000	3300	5000	2100	3200	5100	2100	3100
17	4100	6100	2700	4200	1600	2800	4100	1800	2600
18	5200	2200	2900	5000	2300	2800	5100	2300	2700
19	2400	1100	1700	2600	1200	1600	2500	1300	1700
20	3100	1200	1700	3000	1500	1600	3000	1400	1700
21	4900	1900	2900	5000	1800	3200	5100	1900	3000
22	4700	2000	3000	4600	2100	3100	4600	2200	3200
23	3500	1400	2300	3600	1500	2200	3600	1400	1600
24	5100	2400	3200	5000	2500	3300	5100	2500	3400
25	4600	2100	2900	4700	2200	3000	4600	2000	3000

Приложение Е

Показатели производства озимой пшеницы

№ организации	Качество почв, балл	Продолжительность уборки, дней	Внесено минеральных удобрений на 1 га, кг д. в.	Урожайность с 1 га, ц
1	68	15	156	42,0
2	80	9	156	53,0
3	55	14	158	40,0
4	45	13	84	31,0
5	87	11	149	60,1
6	88	13	145	61,2
7	90	9	280	62,0
8	78	13	134	46,1
9	65	15	163	42,0
10	70	14	115	45,3
11	64	17	97	28,4
12	61	15	157	45,5
13	51	18	81	34,0
14	63	16	103	38,0
15	66	13	115	40,5
16	88	11	300	68,0
17	48	9	164	48,1
18	80	11	280	66,0
19	94	10	320	69,5
20	76	12	250	64,0
21	53	17	97	36,5
22	64	7	97	38,9
23	80	10	140	56,0
24	86	12	260	61,0
25	70	15	115	44,0
26	77	13	130	52,5
27	81	9	290	62,4
28	92	10	280	66,0
29	75	13	255	66,4
30	58	14	75	33,5
31	66	15	160	45,0
32	55	16	102	32,6
33	58	16	108	39,5
34	75	11	146	56,5
35	60	10	188	42,0

№ организации	Качество почв, балл	Продолжительность уборки, дней	Внесено минеральных удобрений на 1 га, кг д. в.	Урожайность с 1 га, ц
36	45	13	105	32,0
37	80	12	260	57,1
38	89	13	230	55,0
39	90	9	275	63,0
40	78	14	134	52,0
41	65	9	172	43,4
42	68	14	165	45,2
43	67	7	101	29,9
44	61	12	157	35,8
45	52	16	101	30,0
46	63	15	102	39,0
47	65	8	115	40,5
48	86	10	300	66,5
49	48	14	156	46,5
50	80	9	275	62,7
51	94	8	320	68,5
52	75	10	245	62,1
53	50	17	159	35,2
54	64	19	98	38,0
55	80	11	145	57,4
56	85	13	260	60,1
57	69	15	110	45,0
58	75	15	130	50,5
59	79	8	290	63,0
60	92	10	285	68,0
61	74	9	245	60,7
62	65	15	160	56,4
63	52	17	122	44,1
64	68	15	156	42,0
65	80	9	156	53,0
66	55	14	158	40,0
67	45	13	84	31,0
68	87	11	149	60,1
69	88	13	145	61,2
70	90	9	280	62,0

Приложение Ж

Показатели деятельности организаций

№ п/п	Площадь, га		Выручка от продажи продукции, тыс. руб.			Среднегодовая численность работников, чел.	Энергетические мощности, л. с.	Затраты на проданную продукцию, тыс. руб.	Среднегодовая заработная плата на работника, тыс. руб.	Среднегодовая стоимость основных фондов, тыс. руб.	
	сельхозугодий	пашни	2015 г.	2016 г. по плану	2016 г. фактически					всего	на 1 га сельхозугодий
1	4785	4739	258888	277290	271380	356	19560	253986	158,1	299652	62,6
2	2556	2511	62568	70830	77247	176	5780	87414	132,0	140859	55,1
3	5553	5516	235914	241437	225150	335	18494	213234	171,0	313695	56,5
4	6276	6250	238770	248313	231579	364	15974	229701	143,4	320241	51,0
5	5660	5613	210423	226185	248001	273	12452	244512	181,8	325962	57,6
6	12737	12674	870399	908658	814764	793	47005	699261	222,9	916617	72,0
7	6768	6552	304101	330678	394431	716	32880	348396	135,9	378969	56,0
8	5429	4701	206754	287460	263733	250	13051	239967	190,8	262290	48,3
9	6081	5594	235629	240633	244776	399	17252	204159	165,3	246513	40,5
10	6453	6139	320664	331410	226704	279	14117	206118	167,4	265749	41,2
11	11228	11069	687732	752610	704730	1100	60861	592848	173,4	810489	72,2
12	1956	1751	121812	137352	169947	338	10295	140244	143,4	171366	87,6
13	15138	13974	715800	781560	839946	1040	64010	749331	171,6	731220	48,3
14	3331	3169	119178	137328	165090	246	11748	145899	175,5	164928	49,5
15	5967	5867	211794	259500	329349	353	20331	307896	201,0	331170	55,5
16	4350	4324	408726	430701	336555	299	20435	284490	199,5	396996	91,3
17	9041	8985	425934	585333	501192	615	30336	447531	182,4	42177	46,7

№ п/п	Площадь, га		Выручка от продажи продукции, тыс. руб.			Средне-годовая числен-ность ра-ботни-ков, чел.	Энергетиче-ские мощности, л. с.	Затраты на про-данную продук-цию, тыс. руб.	Среднегодо-вая заработ-ная плата на работни-ка, тыс. руб.	Среднегодовая стоимость ос-новных фон-дов, тыс. руб.	
	сельхо-зугодий	пашни	2015 г.	2016 г. по плану	2016 г. фактиче-ски					всего	на 1 га сель-хоз-угодий
										9	
18	3764	3147	105690	111798	125460	107	5941	119694	198,6	136008	36,1
19	9025	7723	281169	300384	416718	273	17146	375021	255,6	435351	48,2
20	7878	7203	196692	255957	243738	358	21078	256686	171,6	274938	34,9
21	4706	4422	201384	298386	396708	797	26314	324714	108,3	425598	90,4
22	8629	7832	469440	661800	356940	360	21103	335088	191,4	362955	42,1
23	10582	10289	792243	862263	601143	1147	45278	590475	146,7	488889	46,2
24	6699	6621	391659	407586	375210	437	28540	339639	180,9	465366	69,5
25	1974	1912	127719	143691	145860	260	9046	129099	146,1	133914	67,8
26	4236	4171	178785	198630	197376	289	14912	179742	158,7	263055	62,1
27	15214	14947	275793	361020	492399	607	30272	456420	164,1	673716	44,3
28	4420	4392	183669	247593	214035	174	11188	174051	205,5	178719	40,4
29	3786	3686	174738	182145	186714	165	8514	159474	214,8	209442	55,3
30	7397	7286	381465	500610	450111	426	24815	439380	200,4	470304	63,6
31	2009	1909	91779	105744	109371	123	7294	101856	176,1	61467	30,6
32	3532	3317	126372	135369	150432	215	12569	152961	154,8	114423	32,4
33	9090	8987	321537	345747	332532	522	24012	302697	179,7	497922	54,8
34	10180	9578	661728	691470	720258	768	43343	699366	199,8	971094	95,4
35	6697	6488	349851	376419	278775	647	21459	253254	135,3	355611	53,1
36	12664	12478	1553805	1652895	1588125	1512	88452	1109367	193,8	1322121	104,4
37	13518	10941	601122	751347	716709	1206	57738	644316	139,8	717807	53,1

№ п/п	Площадь, га		Выручка от продажи продукции, тыс. руб.			Средне-годовая численность работников, чел.	Энергетические мощности, л. с.	Затраты на проданную продукцию, тыс. руб.	Среднегодовая заработная плата на работника, тыс. руб.	Среднегодовая стоимость основных фондов, тыс. руб.	
	сельхозугодий	пашни	2015 г.	2016 г. по плану	2016 г. фактически					всего	на 1 га сельхозугодий
38	2635	2526	213747	230760	248070	360	14820	189720	158,1	196044	74,4
39	11727	10430	390504	437730	475332	610	33731	455880	158,7	459891	39,2
40	6883	6668	202530	218673	227142	427	20794	225075	123,9	278133	40,4
41	3566	3502	101625	107892	109131	218	12016	134001	119,1	210327	59,0
42	6589	6214	285465	301350	334116	566	21902	282669	136,2	384765	58,4
43	7439	7389	320718	374100	481842	343	18235	441876	226,8	443670	59,6
44	5820	5413	186762	282849	349014	307	17112	326502	198,3	388620	66,8
45	10630	9978	535212	782250	939510	530	30768	771213	254,1	847470	79,7
46	6174	5540	231414	241260	312957	430	13170	323283	172,5	457728	74,1
47	4875	4748	223440	286005	283086	332	11697	261876	171,9	344082	70,6
48	12238	11319	779454	860934	501699	666	26985	410001	173,1	380649	31,1
49	4112	3875	157242	199368	232035	268	16217	207246	168,6	261474	63,6
50	2951	2861	150525	167820	212910	238	10361	182310	170,4	323271	109,5
51	2433	2328	88764	101766	183606	232	15050	156288	168,0	266370	109,5
52	7198	6977	310470	314136	298407	476	21998	256884	156,3	334179	46,4
53	3769	3593	183825	211860	254580	249	12770	227271	204,6	251016	66,6
54	2897	2868	112536	121500	137736	102	6400	126156	216,9	127626	44,1

Приложение 3

Урожайность сельскохозяйственных культур, ц с 1 га посевной площади

№	Культура	2005 г.	2006 г.	2007 г.	2008 г.	2009 г.	2010 г.	2011 г.	2012 г.	2013 г.	2014 г.	2015 г.	2016г.	2017 г.	2018 г.
1	Зерновые культуры	29,6	41,0	42,5	41,0	38,3	51,9	43,2	46,1	52,6	40,8	50,4	53,4	55,9	56,4
2	Пшеница озимая	33,4	43,1	46,7	42,7	45,1	55,3	45,7	49,7	55,1	39,8	50,1	54,7	57,5	58,4
3	Пшеница яровая	13,7	23,6	27,3	23,8	22,5	34,2	26,8	30,2	31,9	26,3	27,7	31,9	34,1	3938
4	Рожь озимая	28,5	23,6	30,3	22,2	26,4	44,7	32,1	38,5	44,7	31,2	48,5	2,2	34,9	38,7
5	Кукуруза на зерно	31,4	43,9	40,3	40,2	21,8	49,5	33,8	33,8	47,7	41,9	53,0	53,2	53,5	55,0
6	Ячмень озимый	35,0	44,9	41,2	43,5	47,5	51,4	46,9	49,2	53,8	37,1	53,1	49,8	59,5	53,7
7	Ячмень яровой	16,0	22,3	23,6	25,8	18,7	36,9	27,0	25,0	33,8	27,4	30,5	32,7	34,8	36,4
8	Овес	13,3	25,5	25,5	25,2	21,0	33,7	23,3	24,7	30,5	25,2	26,8	31,0	31,6	33,9
9	Просо	6,7	10,9	13,2	14,4	12,5	21,2	16,4	16,6	24,0	16,8	18,8	17,1	23,8	21,4
10	Гречиха	3,8	4,6	6,3	6,6	4,5	10,2	4,9	9,8	7,1	7,0	4,7	5,5	7,7	14,7
11	Рис	32,9	39,8	44,3	47,1	48,3	50,7	60,3	62,1	61,0	64,3	57,6	62,9	63,0	59,3
12	Зернобобовые	8,8	23,1	19,5	22,6	14,4	32,9	23,1	23,7	27,9	21,1	21,0	24,2	27,7	32,2
13	Горох	8,7	23,5	19,6	23,3	14,6	33,9	23,6	24,0	28,1	21,9	21,8	25,6	27,6	32,9
14	Сахарная свекла	215,8	390,1	323,1	359,6	262,4	438,6	381,1	361,2	438,1	423,0	517,1	490,3	461,3	534,5
15	Масличные культуры	13,3	17,3	18,8	18,7	16,4	22,1	20,2	19,4	21,9	21,4	24,1	22,1	21,8	23,6
16	Подсолнечник	14,1	17,3	20,0	20,7	18,9	23,3	20,9	20,8	23,3	23,2	25,7	24,3	24,1	25,1
17	Соя	10,1	17,6	14,5	12,8	9,1	15,8	18,0	15,1	18,5	18,0	20,4	16,9	16,1	20,3
18	Картофель	69,2	92,8	88,4	89,6	78,9	96,8	93,9	89,0	96,4	98,3	100,0	107,4	108,2	112,2
19	Овощи	59,1	75,3	85,4	93,4	79,7	102,8	106,3	98,8	111,7	106,2	106,8	111,8	121,1	116,5
20	Кукуруза на силос	123,8	186,9	155,5	171,3	126,3	189,3	166,1	140,2	193,3	155,3	204,4	203,0	217,2	214,9
22	Сено многолетних трав	33,5	37,6	33,7	26,9	22,4	38,2	32,6	46,2	43,7	36,4	36,0	41,4	42,3	55,3
23	Сено однолетних трав	29,8	18,9	27,6	34,9	21,9	30,6	25,0	23,8	30,4	25,9	24,9	29,7	26,7	30,6
24	Плоды и ягоды	70,2	41,0	57,8	49,1	55,5	70,4	69,6	60,5	74,4	87,3	107,1	96,6	99,9	127,9
25	Виноград	84,9	55,8	68,6	48,6	80,7	74,0	84,5	79,5	113,4	75,6	103,6	99,9	85,9	117,0
26	Чайный лист	6,1	8,7	8,9	8,1	4,5	5,7	4,7	2,7	2,2	1,0	0,9	3,6	4,8	5,0

Приложение И

Урожайность озимой пшеницы по сортам, ц/га

Год	Сорт							
	I	II	III	IV	V	VI	VII	VIII
2013	53	46	50	44	55	39	58	49
2014	43	48	41	46	49	40	50	44
2015	45	46	43	43	48	42	49	43
2016	56	51	52	50	59	46	53	52
2017	58	52	56	51	61	44	55	55
2018	55	48	43	47	60	38	51	54
2019	59	52	61	49	64	41	56	59

Учебное издание

Ворокова Нодира Хасановна
Жминько Альбина Евгеньевна
Сенникова Алина Евгеньевна

Статистика

Методические рекомендации

В авторской редакции