МИНИСТЕРСТВО СЕЛЬСКОГО ХОЗЯЙСТВА РОССИЙСКОЙ ФЕДЕРАЦИИ

ФГБОУ ВПО «КУБАНСКИЙ ГОСУДАРСТВЕННЫЙ АГРАРНЫЙ УНИВЕРСИТЕТ»

МАТЕМАТИЧЕСКАЯ СТАТИСТИКА

Практикум

Кафедра статистики и прикладной математики

Краснодар
КубГАУ
2015

УДК 519.22(076.5)
ББК 22.172
М34

Р е ц е н з е н т :

Е. И. Артемова – доктор экономических наук, профессор
(Кубанский государственный аграрный университет)

Коллектив авторов:
Т. А. Курнякова, И. А. Кацко, А. Е. Жминько, Е. В. Кремянская,
А. Е. Сенникова, Н. Х. Ворокова

М 34 **Математическая статистика**: практикум / Т. А. Курнякова [и др.]. – Краснодар : КубГАУ, 2015. – 103 с.

Практикум содержит краткое теоретическое изложение основных положений дисциплины «Математическая статистика» в тематическом разрезе, а также задания к практическим занятиям, позволяющие сформировать и закрепить умения и навыки математической обработки и анализа статистических показателей.

Предназначен для студентов-бакалавров по направлениям подготовки «Агрономия», «Садоводство».

УДК 519.22(076.5)
ББК 22.172

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 практикума – 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 практикума составлена на основе фактических данных Краснодарстата и сельскохозяйственных организаций Краснодарского края.

Настоящий практикум предназначен для студентов-бакалавров по направлениям подготовки «Агрономия», «Садоводство».

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.

1 Абсолютные и относительные величины

Абсолютными статистическими величинами называются величины, выражающие размеры, объемы и уровни общественных явлений и процессов. Они имеют определенные единицы измерения: натуральные (килограммы, штуки, центнеры, гектары, кубометры, километры); стоимостные (рубли, доллары, евро) и трудовые (человеко-часы, человеко-дни). Часто применяются комбинированные (кВт-час, машино-день, тонно-километр), а также условные (эталонные гектары, условные банки консервов и т. п.) единицы измерения.

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

Для выражения результата сопоставления одноименных величин используются коэффициенты (если база сравнения принимается за 1), проценты (если база сравнения принимается за 100), промилле (если база сравнения принимается за 1000), продецимилле (если база сравнения принимается за 10000). Относительные величины могут быть выражены именованными числами. Например, плотность населения на 1 км^2, количество произведенного молока на 100 га сельскохозяйственных угодий и т.п.

По характеру, назначению и сущности выражаемых количественных соотношений различают следующие виды относительных величин: структуры; координации; выполнения плана; планового задания; динамики; интенсивности; сравнения.

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

$$К_д = \frac{У_1}{У_0}. \tag{1.3}$$

Относительные величины динамики, планового задания и выполнения плана взаимосвязаны

$$K_{д} = K_{в.п.} \cdot K_{п.з}. \quad (1.4)$$

Относительные величины интенсивности показывают степень распространения данного явления в определенной среде. Обычно это отношение двух различных, но связанных между собой абсолютных величин (численность населения к площади территории, на которой оно проживает; фондообеспеченность (среднегодовая стоимость основных фондов на 100 га сельхозугодий)).

Относительные величины сравнения представляют собой соотношение одноименных показателей, относящихся к различным объектам или территориям, но за один и тот же период или момент времени. При помощи этих величин сопоставляются показатели по разным странам, регионам, предприятиям, объектам.

Пример 1.1. На основании имеющихся данных изучить структуру посевных площадей сельскохозяйственных культур и рассчитать различные виды относительных величин (коэффициент выполнения плана, коэффициент планового задания и коэффициент динамики). Структуру посевных площадей изобразить графически. Сделать вывод.

Решение.

Построим вспомогательную таблицу 1.1 и проведем расчет показателей.

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

Таблица 1.1 – Динамика и структура посевных площадей

Культура (группа культур)	Площадь посева, га			Структура, %			Относительные величины		
	2013г.	План на 2014г.	2014г.	2013г.	План на 2014г.	2014г.	планового задания	выполнения плана	динамики
Озимая пшеница	1265	1300	1350	26,1	27,6	27,7	1,028	1,038	1,067
Кукуруза	459	320	350	9,5	6,8	7,2	0,697	1,094	0,763
Овощи открытого грунта	500	510	480	10,3	10,8	9,8	1,020	0,941	0,960
Овес	1300	1280	1300	26,8	27,2	26,6	0,985	1,016	1,000
Подсолнечник	1320	1300	1400	27,3	27,6	28,7	0,985	1,077	1,061
Итого	4844	4710	4880	100,0	100,0	100,0	0,972	1,036	1,007

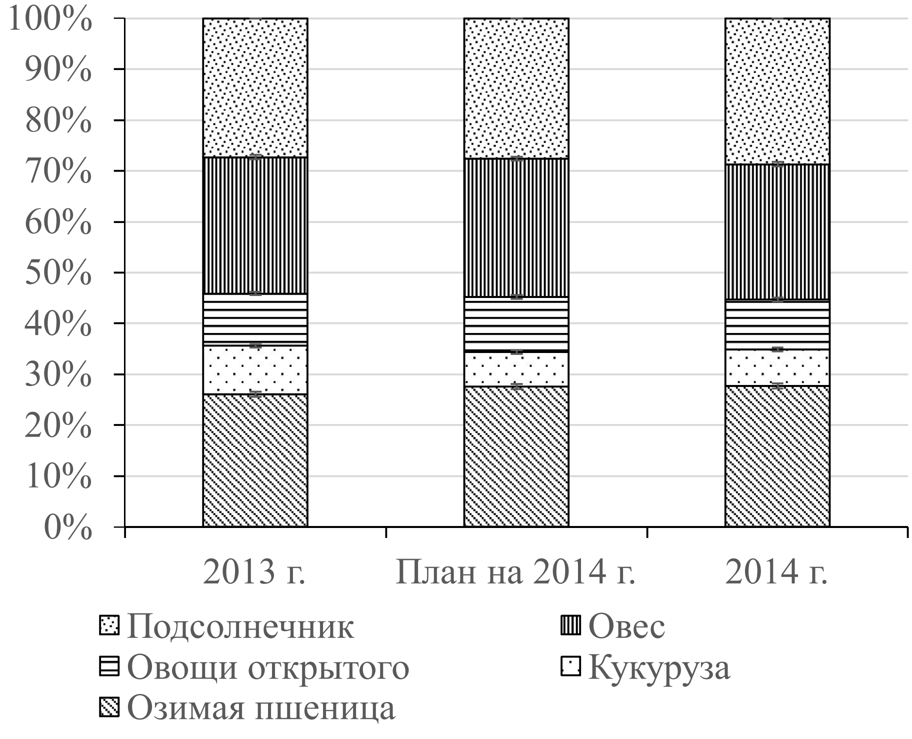


Рисунок 1.1 – Структура посевных площадей, процентов

Вывод. В отчетном году (2014 г.) по сравнению с базисным (2013 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 недовыполнен 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 0,7 % (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. На основании данных приложения А по одному варианту определить относительные величины структуры, планового задания, выполнения плана и динамики посевных площадей. Структуру посевов изобразить графически. Сделать вывод.

Задача 1.2. В организации объем производства зерна в 2013 г. составил 194137 ц. По плану в 2014 г. предусматривалось увеличить объем производства зерна на 7,5 % по сравнению с 2013 г. Плановое задание было перевыполнено в 2014 г. на 5,8 %. Определить объем производства в 2014 г. по плану и фактически, рассчитать коэффициент динамики. Сделать вывод.

Задача 1.3. Объем реализации маслосемян подсолнечника в организации в 2014 г. по сравнению с 2013 г. увеличился на 14,9 %

или на 1870 ц. Договорные обязательства по реализации данной продукции перевыполнены на 4,3 %. Определить уровень реализации маслосемян в 2013 и 2014 гг., размер договорных обязательств по реализации и степень напряженности договорных обязательств. Сделать вывод.

Задача 1.4. Себестоимость производства 1 ц зерна озимой пшеницы в 2013 г. составила 448 руб. На 2014 г. планировалось снизить себестоимость на 4,8 %, при этом фактическая себестоимость в 2014 г. по сравнению с предыдущим годом выросла на 1,5 %. Рассчитать фактический, плановый уровни себестоимости в 2014 г. и коэффициент выполнения плана. Сделать вывод.

Задача 1.5. По имеющимся данным о внесении минеральных удобрений на 1 га посева сельскохозяйственных культур в сельскохозяйственных организациях Краснодарского края за 2012 – 2014 гг. рассчитать относительные величины сравнения, приняв за базу средний уровень показателя по Краснодарскому краю. Сделать вывод.

Район	Внесено удобрений, кг		
	2012 г.	2013 г.	2014 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 1.6. По данным двух районов Краснодарского края провести сравнение валовых сборов продукции растениеводства. Изучить изменение валовых сборов в динамике. Сделать вывод.

Показатель	Динской район			Кавказский район		
	2012 г.	2013 г.	2014 г.	2012 г.	2013 г.	2014 г.
Валовой сбор, тыс. т: зерна зерновых и зернобобовых культур	152,1	275,8	257,2	253,0	374,4	386,3
маслосемян подсолнечника	25,2	29,4	31,2	23,8	25,1	25,7
сахарной свеклы	114,2	113,7	124,5	379,0	211,3	218,9
картофеля	18,0	17,0	19,2	12,4	10,5	12,0
овощей	75,5	82,5	80,3	14,3	9,7	11,0

Задача 1.7. По имеющимся за 2014 г. данным определить относительные показатели структуры, сравнения, координации и интенсивности. Сделать вывод.

Показатель	ОАО «Родина» Каневский район	ОАО АФ ПЗ «Победа» Каневский район
Общая земельная площадь, га	9561	21567
в том числе: площадь сельскохозяйственных угодий	8898	19837
площадь древесно-кустарниковых растений	437	735
пруды и водоемы	40	152
болота	20	122
прочие земли	166	721
Среднегодовая численность персонала, чел.	390	935
Среднегодовая стоимость основных средств, тыс. руб.	1016775	1434210

Вопросы для самоподготовки

1. Дайте определение абсолютных величин, назовите их виды и единицы измерения.

2. Назовите виды относительных величин.

3. Какова взаимосвязь между относительными величинами выполнения плана, планового задания и динамики?

4. Назовите единицы измерения относительных величин.

5. Приведите примеры относительных величин, характеризующих наличие и использование ресурсов в организации.

2 Статистические ряды распределения

Статистика изучает количественную сторону массовых явлений и процессов по варьирующим признакам.

Признак – это особенность или характерная черта, присущая отдельным единицам совокупности.

Результаты статистического наблюдения оформляются в виде статистических рядов распределения. Статистический ряд распределения представляет собой упорядоченное распределение единиц изучаемой совокупности на группы по определенному признаку. Они характеризуют состав (структуру) изучаемого явления, позволяют судить об однородности совокупности, границах ее изменения, закономерностях развития наблюдаемого объекта. В соответствии с характером выражения признаков, статистические ряды распределения подразделяются на атрибутивные (качественные) и вариационные (количественные).

Атрибутивные ряды образуются по качественным (описательным) признакам, которыми могут выступать занимаемая должность работников, профессия, пол, образование, национальность, сорт и т. п.

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):

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

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

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

Частость часто выражается в процентах, тогда $\Sigma \hat{p}_i = 100$.

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

x_i	x_1	x_2	x_3	…	x_k
n_i	n_1	n_2	n_3	…	n_k
$\hat{p}_i$	$\hat{p}_1$	$\hat{p}_2$	$\hat{p}_3$.	…	$\hat{p}_k$

Если признак непрерывный, или число значений дискретного признака велико, то в этом случае строится интервальный вариационный ряд, в котором значения признака заданы интервалами. Интервалы могут иметь как равную, так и неравную длину. Неравные интервалы применяются, когда изменение группировочного признака по единицам совокупности происходит неравномерно и в значительных пределах. Для построения такого ряда весь промежуток изменения признака разбивается на ряд отдельных интервалов, длина которых по группам прогрессивно возрастает или убывает.

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

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

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:

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

$$M_e = x_{n+1}.$$

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 модального интервала;

n_{Mo} – частота модального интервала;

n_{Mo-1} – частота предмодального интервала;

n_{Mo+1} – частота послемодального интервала;

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

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:

$$V = \frac{\sigma}{\bar{x}} \cdot 100\%. \tag{2.13}$$

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. В результате получим дискретный вариационный ряд.

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.

Таблица 2.1 – Распределение рабочих по разрядам

Разряд рабочего	Число рабочих	Накопленное число рабочих	Относительная частота
	n_i	S_i	$\hat{p}_i$
1	7	7	0,064
2	12	19	0,109
3	29	48	0,264
4	33	81	0,300
5	17	98	0,154
6	12	110	0,109
Сумма	110	-	1,000

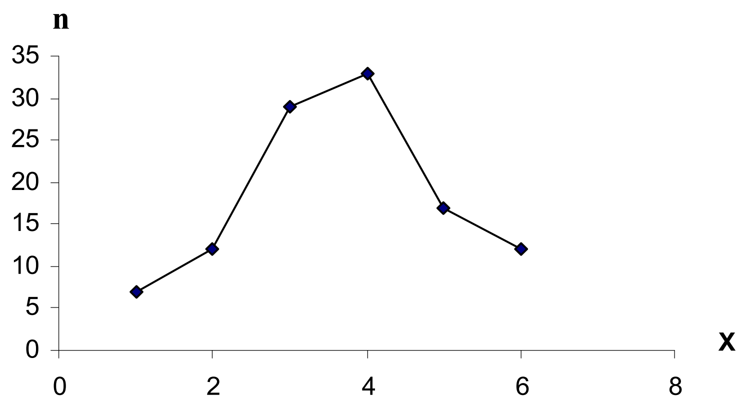


Рисунок 2.1 - Полигон распределения рабочих по разрядам

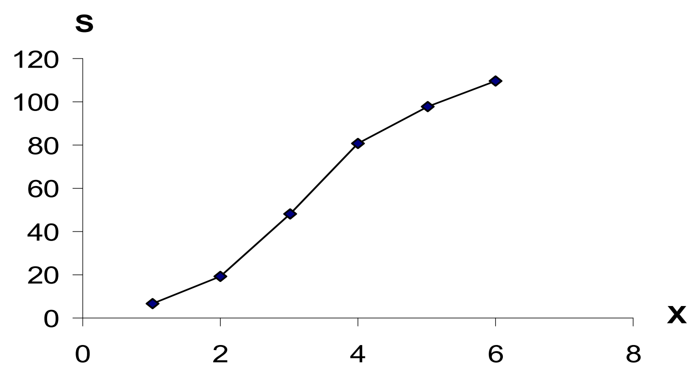


Рисунок 2.2 - Кумулята распределения рабочих по разрядам

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

$$\overline{x} = \frac{1\cdot 7 + 2\cdot 12 + 3\cdot 29 + 4\cdot 33 + 5\cdot 17 + 6\cdot 12}{110} = \frac{407}{110} = 3{,}7\,.$$

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.

Далее определим дисперсию по формуле 2.11:

$$\sigma^2 = \frac{(1-3{,}7)^2\cdot 7 + (2-3{,}7)^2\cdot 12 + (3-3{,}7)^2\cdot 29 + (4-3{,}7)^2\cdot 33 + (5-3{,}7)^2\cdot 17 +}{110}$$

$$\frac{+\,(6-3{,}7)^2\cdot 12}{} = \frac{195{,}1}{110} = 1{,}774.$$

Среднее квадратическое отклонение определяется как квадратный корень из дисперсии:

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

Коэффициент вариации:

$$V = \frac{\sigma}{\overline{x}}\cdot 100 = \frac{1{,}33}{3{,}7}\cdot 100 = 35{,}9\,\%..$$

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 (рис. 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

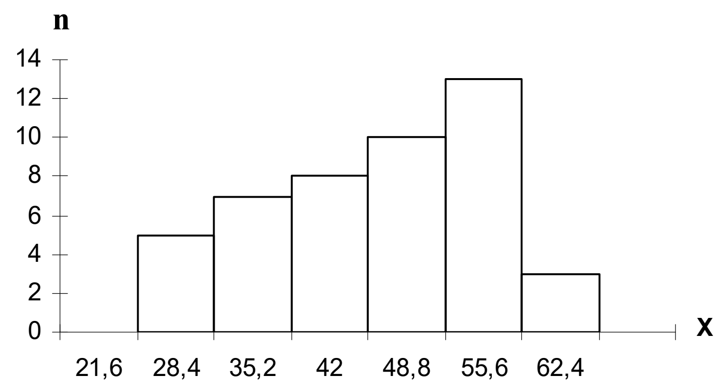


Рисунок 2.3 - Гистограмма распределения сельскохозяйственных организаций по урожайности кукурузы на зерно, ц/га

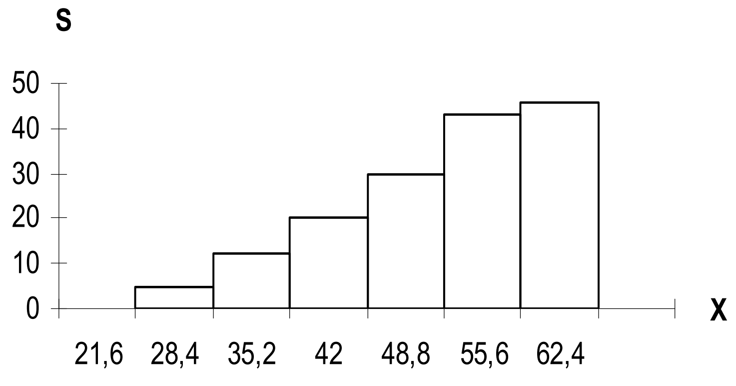


Рисунок 2.4 - Кумулята распределения сельскохозяйственных организаций по урожайности кукурузы на зерно, ц/га

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:

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

Средняя урожайность кукурузы на зерно составит:

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

Определим показатели вариации:

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

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

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

г) коэффициент вариации: $V = \frac{10{,}0}{42{,}7} \cdot 100 = 23{,}4\%$.

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 квадратического отклонения, урожайность колебалась в среднем в границах (42,7 ± 10,0) ц/га, т.е. от 32,7 до 52,7 ц/га. Коэффициент вариации свидетельствует 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 Б). 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. Составить интервальный ряд распределения организаций по продолжительности уборки озимой пшеницы (приложение Б). 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 2.6. Составить интервальный ряд распределения сельскохозяйственных организаций по количеству внесенных минеральных удобрений на 1 га посева (приложение Б). Найти накопленные частоты. Определить среднее, модальное и медианное значения, а также показатели вариации. Ряд распределения изобразить графически с помощью гистограммы и кумуляты распределения.

Задача 2.7. Составить интервальный вариационный ряд распределения организаций по балльной оценке качества почв (приложение Б). Определить среднее, модальное и медианное значения, а также показатели вариации. Ряд распределения изобразить графически с помощью гистограммы и кумуляты распределения.

Задача 2.8. 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 2.9. Кукурузные початки распределились следующим образом по числу зерен:

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

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.

Задача 2.10. На опытной делянке измерялась длина стеблей растений озимой пшеницы. Были получены следующие результаты (см): 80,7; 80,0; 82,4; 83,0; 81,9; 82,2; 82,9; 81,1; 80,6; 82,9; 81,4; 82,8; 80,8; 81,6; 82,1; 80,9; 81,8; 82,0; 81,0; 80,5.

Составить вариационный ряд с равными интервалами, изобразить его графически. Определить среднюю, модальную, медианную длину стебля и показатели вариации.

Задача 2.11. Имеются следующие данные о диаметре соцветий-корзинок подсолнечника на опытном поле (см):
24,2; 25,8; 30,6; 32,0; 28,1; 29,6; 33,2; 34,0; 32,9; 29,9; 26,0; 28,1; 24,8; 27,3; 24,5; 30,1; 32,2; 33,6; 31,2; 31,9; 33,9; 25,6; 26,3; 27,4; 28,9; 32,0; 32,1; 33,3; 31,5; 24,9.

Составить вариационный ряд с равными интервалами, изобразить его графически. Определить средний, модальный, медианный диаметры соцветия и показатели вариации.

Задача 2.12. Кукурузные початки распределились следующим образом по числу рядов зерен:

Число рядов зерен в початке, шт.	Число початков по вариантам, шт.				
	I	II	III	IV	V
14	10	4	6	2	16
16	8	12	11	10	6
18	3	17	5	13	9
20	4	15	8	15	10
22	12	8	9	3	7

По одному варианту найти накопленные частоты и частости. Определить модальное, медианное и среднее число рядов зерен, среднее квадратическое отклонение и коэффициент вариации. Ряд распределения изобразить графически с помощью полигона и кумуляты распределения.

Задача 2.13. Используя имеющиеся данные о распределении колосьев озимой пшеницы по числу колосков, по одному варианту найти накопленные частоты и частости. Определить модальное, медианное и среднее число колосков в колосе, среднее квадратическое отклонение и коэффициент вариации. Ряд распределения изобразить графически с помощью полигона и кумуляты распределения.

Число колосков в колосе, шт.	Число колосьев по вариантам опыта, шт.				
	I	II	III	IV	V
24	16	11	10	5	12
25	8	4	14	9	4
26	3	7	8	18	2
27	5	12	2	3	17
28	10	15	9	6	7

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.

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:

$$\Delta = t \cdot \mu, \qquad (3.2)$$

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:

$$\mu_{\bar{x}} = \frac{\sigma}{\sqrt{n}}, \qquad (3.3)$$

б) если отбор случайный бесповторный:

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

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:

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

где $\widetilde{x}$ - выборочная средняя.

В малых выборках ($n \leq 30$):

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

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:

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

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:

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

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.

а) 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:

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

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.

Пример 3.2. Для определения всхожести приготовленных для посева семян случайным образом было отобрано 1000 зерен. В результате оказалось, что 90 % отобранных зерен всхожи. С доверительной вероятностью 0,954 определить, в каких пределах будет находиться всхожесть семян во всей партии.

Решение.

Так как объем выборочной совокупности очень мал по сравнению с генеральной совокупностью, для решения задачи необходимо воспользоваться формулой предельной ошибки выборки для случайного повторного отбора для доли.

При γ =0,954 и n = 1000, по таблице t = 2.

$$\Delta_p = t\sqrt{\frac{w(1-w)}{n}} = 2\sqrt{\frac{0{,}9 \cdot 0{,}1}{1000}} = 0{,}019 \text{ или } 1{,}9\ \%.$$

Тогда всхожесть семян во всей партии будет находиться в пределах: $w - \Delta_p \le P \le w + \Delta_p$; $0{,}881 \le P \le 0{,}919$, т. е. от 88,1 до 91,9 %.

Пример 3.3. Новый сорт подсолнечника был высеян на 8 делянках одинаковой площади. Получена следующая урожайность на делянках, ц/га: 25,1; 28,4; 24,4; 27,6; 29,4; 27,8; 26,5; 28,8. При уровне доверительной вероятности 0,95 оценить границы, в которых будет находиться средняя урожайность подсолнечника в генеральной совокупности.

Решение.

По условию задачи имеем: n=8; γ =0,95.

Найдем среднюю урожайность и среднее квадратическое отклонение урожайности (таблица 3.1).

Таблица 3.1. – Вспомогательная таблица для расчета средней и дисперсии

Номер делянки	Урожайность, ц/га x_i	$x_i - \widetilde{x}$	$(x_i - \widetilde{x})^2$
1	25,1	-2,15	4,6225
2	28,4	1,15	1,3235
3	24,4	-2,85	8,1225
4	27,6	0,35	0,1225
5	29,4	2,15	4,6225
6	27,8	0,55	0,3025
7	26,5	-0,75	0,55625
8	28,8	1,55	2,4025
Итого	218,0	–	22,08

$$\widetilde{x} = \frac{\sum x_i}{n} = \frac{218,0}{8} = 27,25; \quad S^2 = \frac{\sum (x_i - \widetilde{x})^2}{n-1} = \frac{22,08}{8-1} = 3,1543; \quad S = \sqrt{3,1543} = 1,78.$$

Отбор участков – случайный бесповторный, но так как объем генеральной совокупности неизвестен, а отношение n / N очень мало, то расчет ошибок выборки производится по формуле для случайного повторного отбора:

$$\mu_{\bar{x}} = \frac{S}{\sqrt{n}} = \frac{1,78}{\sqrt{8}} = 0,63; \quad P\% = \frac{\mu_{\bar{x}}}{\widetilde{x}} \cdot 100 = \frac{0,63}{27,25} \cdot 100 = 2,3\%.$$

Число степеней свободы k=n–1=8–1=7. При $\gamma = 0,95$ и k=7 по таблице t - Стьюдента t=2,36.

Следовательно, $\Delta_{\bar{x}} = t \cdot \mu_{\bar{x}} = 2,36 \cdot 0,63 = 1,49.$

Вывод. Средняя урожайность данного сорта подсолнечника по результатам опыта составила 27,25 ц/га, и по делянкам урожайность в среднем колебалась в границах (27,25 ± 1,78) ц/га. Относительная средняя ошибка выборки, называемая точностью опыта, составила 2,3 %, что свидетельствует о достаточно высокой точности опыта. С доверительной вероятностью 0,95 можно утверждать, что средняя урожайность подсолнечника в генеральной совокупности будет находиться в интервале (27,25-1,49; 27,25+1,49), т.е. от 25,76 до 28,74 ц/га.

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 3.2. 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.

Задача 3.9. В районе имеется 1000 индивидуальных хозяйств. Сколько хозяйств необходимо отобрать для выборочного обследования, чтобы предельная ошибка урожайности овощей была не больше 10 ц/га. На основании предварительного исследования установлено, что среднее квадратическое отклонение урожайности овощей составляет 40 ц/га. Уровень доверительной вероятности принять равным: а) 0,954, б) 0,99.

Задача 3.10. Для определения влажности зерна случайным образом было взято 100 проб. Средний процент влажности составил 15 %, при среднем квадратическом отклонении 3 %. Считая отношение n / N незначительным, определить с вероятностью 0,997, в каких пределах будет заключена влажность зерна во всей партии.

Задача 3.11. Для определения потерь при уборке зерна было наложено 100 метровок. Средняя величина потерь составила 1,4 ц/га при среднем квадратическом отклонении 0,3 ц/га. Определить, в каких границах будет находиться величина потерь зерна со всей площади в 500 га с доверительной вероятностью 0,99.

Задача 3.12. На основании 50 наблюдений установлено, что на выполнение производственной операции в среднем затрачивается 5 мин., при среднем квадратическом отклонении 1 мин. Считая время выполнения операции нормально распределенной случайной величиной, определить границы, в которых будет находиться истинное значение времени выполнения операции с доверительной вероятностью 0,9; 0,95; 0,99.

Задача 3.13. Из 500 отобранных клубней картофеля 90 % соответствовало первому сорту. Определить среднюю и предельную ошибку выборки, границы, в которых находится доля картофеля первого сорта во всей партии с вероятностью 0,954. Определить необходимый объем выборки, чтобы снизить ошибку выборки в 2 раза.

Задача 3.14. Выборочным методом изучалась всхожесть семян овощного гороха. Из партии семян 25 проб (серий), по которым оказалось, что всхожесть семян составила 70 % при среднем весе 10 семян 11 г. Среднее квадратическое отклонение веса составило 2 г на 10 шт. семян. С вероятностью 0,954 определить границы, в которых будет находиться всхожесть и средний вес семян овощного гороха во всей партии.

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. Перечислите виды и способы формирования выборочной совокупности.
5. Назовите характеристики выборочной и генеральной совокупности.
6. Как определяются средняя и предельная ошибки выборки при различных способах отбора?

7. Как определяется необходимая численность выборки?

8. Какие существуют способы распространения данных выборочного наблюдения на генеральную совокупность?

4 Проверка статистических гипотез

Статистической гипотезой называется всякое предположение о генеральной совокупности, проверяемое по выборке. Статистические гипотезы делятся на:

- параметрические – гипотезы, сформулированные относительно параметров (среднего значения, дисперсии и т.д.) распределения известного вида;

- непараметрические – гипотезы, сформулированные относительно вида распределения (например, определение по выборке нормальности генеральной совокупности).

Процесс использования выборки для проверки гипотезы называется статистическим доказательством. Выдвинутую гипотезу называют основной или нулевой, и обозначают H_0. Гипотезу, противоположную нулевой, называют конкурирующей или альтернативной и обозначают H_1. Например, выдвигается гипотеза, что средняя урожайность культуры в данных условиях составляет 50 ц/га, тогда $H_0 : \bar{x} = 50$. Альтернативная гипотеза может иметь вид: $H_1 : \bar{x} \neq 50$; $H_1 : \bar{x} > 50$ или $H_1 : \bar{x} < 50$ (средняя урожайность не равна, или больше, или меньше 50 ц/га).

Так как проверка статистических гипотез обычно осуществляется по выборочным данным, то возникает возможность принятия ошибочных решений. Выбор между нулевой и конкурирующей гипотезой может сопровождаться ошибками двух родов.

Ошибка первого рода заключается в том, что будет отвергнута правильная гипотеза, т. е. когда в действительности верна H_0 гипотеза, а в результате проверки она была отвергнута и принята гипотеза H_1. Вероятность ошибки первого рода называется уровнем значимости и обозначается α.

$$\alpha = P(H_1 / H_0). \tag{4.1}$$

Ошибка второго рода состоит в том, что будет принята неправильная гипотеза, т. е. если в действительности верна некоторая альтернативная гипотеза, а по выборочным данным была принята неверная гипотеза H_0. Вероятность ошибки второго рода обозначается β.

$$\beta = P(H_0 / H_1). \tag{4.2}$$

Существует правильное решение двух видов: $P\ (H_0 / H_0) = 1 - \alpha$ 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 - верная	$P\ (H_0 / H_0) = 1 - \alpha$	$P\ (H_1 / H_0) = \alpha$
H_0 - неверная	$P\ (H_0 / H_1) = \beta$	$P\ (H_1 / H_1) = 1 - \beta$

Правило, по которому принимается решение о том, что верна или неверна 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.

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 распределениях: $u, \chi^2, t, F.$ Непараметрические гипотезы проверяют с помощью критериев согласия, использующих распределения: χ^2, Колмогорова, Смирнова.

Пусть проверяется нулевая гипотеза H_0: $\bar{x} = a$.

В зависимости от вида альтернативной гипотезы рассматривают три случая:

а) Если H_1: $\bar{x} \neq a$, то в этом случае рассматривают двустороннюю критическую область и используют дифференциальную функцию f (K/H_0) для определения соответствующих критических значений (или иначе, квантилей – границ области принятия гипотезы: левой $(K_{1-a/2})$и правой $(K_{a/2})$. Площадь под криволинейной трапецией дифференциальной функции слева от $K_{1-a/2}$ и справа от $K_{a/2}$ равна *α*/2. Критическая точка *К* разделяет область допустимых значений критерия и критическую область.

Общая площадь, ограниченная криволинейной трапецией дифференциальной функции, квантилями и осью абсцисс, равна (1-*α*) (рисунок 4.1).

$$P\ (K_{1-a/2} < K < K_{a/2}) = 1 - \alpha. \tag{4.3}$$

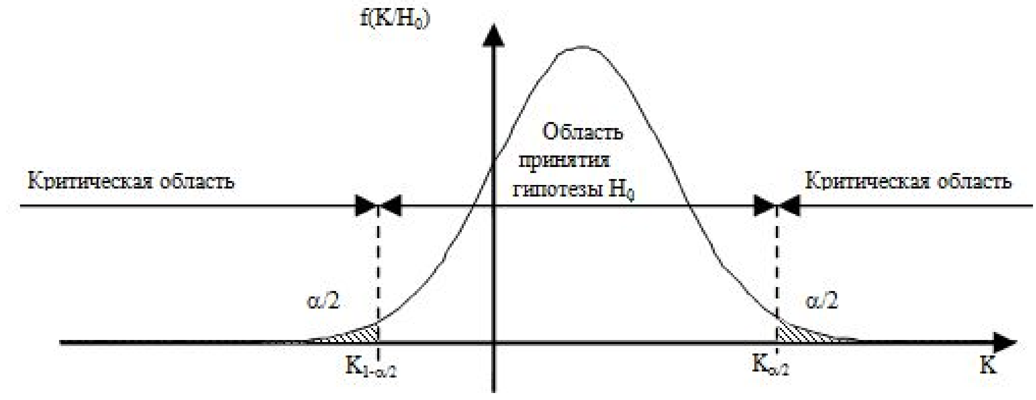
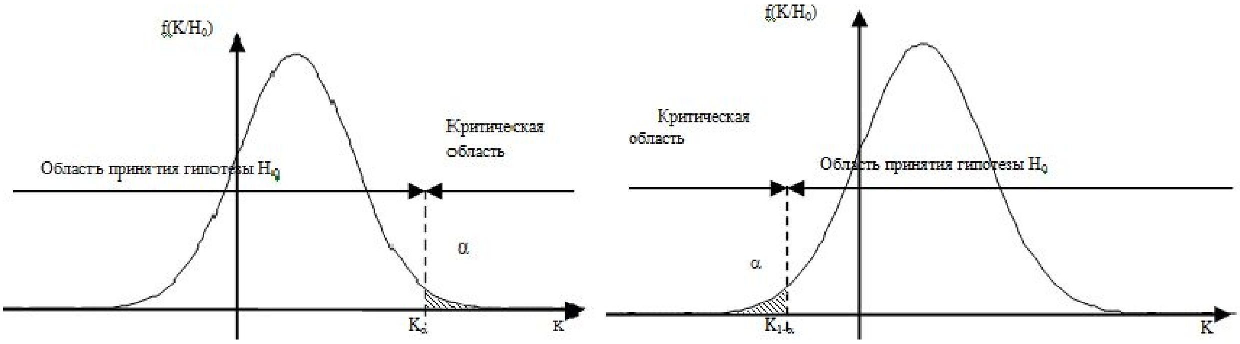


Рисунок 4.1 - Двусторонняя критическая область

б) Если H_1: $\bar{x} > a$, то рассматривается правосторонняя критическая область (площадь под криволинейной трапецией справа от K_a равна α) (рисунок 4.2 а).

$$P(K>K_a)=\int_{K_a}^{+\infty} f\left(K / H_0\right) dK = \alpha. \qquad (4.4)$$



а) б)

Рисунок 4.2 - Критические области: а) правосторонняя, б) левосторонняя

в) Если H_1: $\bar{x} < a$, то рассматривается левосторонняя критическая область (площадь под криволинейной трапецией слева от K_{1-a} равна α) (рисунок 4.2 б).

$$\text{P } (K<K_{1-a})=\int_{-\infty}^{K_{1-a}} f\left(K / H_0\right) dK = \alpha\,. \qquad (4.5)$$

Проверка статистических гипотез обычно осуществляется в определенной последовательности.

1. Располагая выборочными данными $\left(x_1, x_2, \ldots, x_n\right)$, формулируют нулевую гипотезу H_0 и конкурирующую гипотезу H_1.

2. Задают уровень значимости α (обычно принимают α=0,1; 0,01; 0,05; 0,001).

3. Выбирают критерий *К*, по которому будет проверяться выдвинутая гипотеза. Наиболее часто используют следующие распределения критериев:

u – нормальное распределение;

χ^2 - распределение Пирсона (хи-квадрат);

t – распределение Стьюдента;

F – распределение Фишера-Снедекора.

4. По выборочным данным $x_1, x_2, ..., x_n$ определяют фактически наблюдаемое значение критерия $K_н$.

5. В зависимости от вида альтернативной гипотезы выбирают по соответствующей таблице (приложения В–Е) критические значения критерия для двусторонней $\left(\mathrm{K}_{1-\alpha/2} \text{ и } \mathrm{K}_{\alpha/2}\right)$ или односторонней области $\left(\mathrm{K}_{1-\alpha} \text{ или } \mathrm{K}_{\alpha}\right)$. Сравнивается фактически наблюдаемое значение критерия с критическим. Если фактически наблюдаемое значение критерия попадает в критическую область, то H_0 отвергается, в противном случае принимается гипотеза H_0 и считается, что H_0 не противоречит выборочным данным (при этом существует возможность ошибки с вероятностью, равной α).

При проверке гипотезы о нормальном распределении генеральной совокупности сравниваются эмпирические (наблюдаемые) и теоретические (вычисленные в предложении распределения) частоты. Для этого используется критерий χ^2 –Пирсона с $v = k - r - 1$ степенями свободы (*k* – число групп, на которые разбит вариационный ряд распределения, *r* – число оцениваемых параметров). При нормальном распределении оценивается математическое ожидание и среднее квадратическое отклонение, следовательно, *r* = 2. Если $\chi_н^2 \ge \chi_{кр}^2$, то нулевая гипотеза отвергается и считается, что предположение о нормальности распределения не согласуется с эмпирическими данными. В противном случае $\left(\chi^2_{н.} < \chi^2_{кр.}\right)$ нулевая гипотеза принимается.

Теоретические вероятности (p_i), попадания случайной величины в частичные интервалы $\left(x_i; x_{i+1}\right)$ находятся по формуле

$$p_i = P\left(x_i \le X \le x_{i+1}\right) = \Phi\left(u_{i+1}\right) - \Phi\left(u_i\right), \left(i = 1, 2, \dots k\right), \tag{4.6}$$

где $u_i = \dfrac{x_i - \bar{x}}{\sigma}$.

Значения $\Phi(u_i)$ находятся по данным приложения Г.

Пример 4.1. Имеется распределение сельскохозяйственных организаций Краснодарского края по урожайности озимой пшеницы:

Группа хозяйств по урожайности озимой пшеницы, ц/га	Число хозяйств в группе
20,01–26,7	6
26,71–33,4	9
33,41–40,1	11
40,11–46,8	13
46,81–53,5	6
53,51–60,2	5
Итого	50

Требуется проверить нулевую гипотезу, что совокупность организаций по урожайности озимой пшеницы распределяется по нормальному закону. Уровень значимости принять равным 0,05.

Решение.

Из условия следует, что точные параметры гипотетического нормального закона нам неизвестны, поэтому нулевую гипотезу можно сформулировать следующим образом: $f(x)$ является функцией нормального распределения с параметрами $M(X)=\alpha=\bar{x}$ и $D(X)=\sigma^2$.

Для проверки этой нулевой гипотезы необходимо найти точечные оценки математического ожидания и среднего квадратического отклонения нормально распределенной случайной величины:

$$\bar{x}=\frac{\sum x_i n_i}{n}=39{,}3 \text{ ц/га}; \qquad \sigma=\sqrt{\frac{1}{n}\sum (x_i-\bar{x})^2 n_i}=9{,}8 \text{ ц/га}.$$

Вычисления, необходимые для определения наблюдаемого значения критерия χ^2, проведем в таблицах 4.2 и 4.3.

Таблица 4.2 – Определение теоретических частот распределения хозяйств по урожайности озимой пшеницы

Группа хозяйств по урожайности озимой пшеницы, ц/га ($x_i - x_{i+1}$)	Число хозяйств, n_i	$u_i=\frac{x_i-39{,}3}{9{,}8}$	$u_{i+1}=\frac{x_{i+1}-39{,}3}{9{,}8}$	$\Phi(u_i)$	$\Phi(u_{i+1})$	P_i	$n_i'=np_i$
20,0–26,7	6	$-\infty$	-1,28	-0,5	-0,4007	0,0993	5
26,7–33,4	9	-1,28	-0,60	-0,4007	-0,2264	0,1743	9
33,4–40,1	11	-0,60	0,08	-0,2264	0,0325	0,2589	13
40,1–46,8	13	0,08	0,76	0,0325	0,2779	0,2454	12
46,8–53,5	6	0,76	1,45	0,2779	0,4263	0,1484	7
53,5–60,2	5	1,45	$+\infty$	0,4263	0,5	0,0737	4
Итого	50	-	-	-	-	1,0000	50

Так как нормально распределенная случайная величина изменяется по всей числовой оси, то в данном примере крайние границы первого и последнего интервалов берутся равными $-\infty$ и $+\infty$.

Теоретические частоты n_i' находятся по формуле $n_i'=np_i$.

Таблица 4.3 – Определение наблюдаемого значения критерия Пирсона

$x_i - x_{i+1}$	Частоты		$n_i - n_i'$	$(n_i - n_i')^2$	$\frac{(n_i - n_i')^2}{n_i'}$
	n_i	n_i'			
20,0–26,7	6	5	1	1	0,2
26,7–33,4	9	9	0	0	0
33,4–40,1	11	13	-2	4	0,308
40,1–46,8	13	12	1	1	0,083
46,8–53,5	6	7	-1	1	0,143
53,5–60,2	5	4	1	1	0,25
Итого	50	50	0	-	0,984

В результате вычисления получили $\chi^2_{н.} = 0{,}984$. Найдем по таблице χ^2 – распределения (приложение Д), при заданном уровне значимости α=0,05 и числе степеней свободы $v=k-r-1=6-2-1=3$, критическое значение: $\chi^2_{кр.} = \chi^2_{0,5} = 7{,}815$.

Так как $\chi^2_н < \chi^2_{кр.}$ (0,984<7,815), то нулевая гипотеза принимается. Сельскохозяйственные организации по урожайности озимой пшеницы распределяются по нормальному закону.

Пример 4.2. Из нормальной генеральной совокупности сельскохозяйственных предприятий, рассматриваемых по показателю урожайности пшеницы, с известным средним квадратическим отклонением $\sigma = 9{,}4$ и генеральной средней $\bar{x}_0 = 48{,}1$, извлечена выборка объема n=50. По ней найдена выборочная средняя $\bar{x}$ =52. Требуется при уровне значимости α=0,05 проверить нулевую гипотезу H_0:

а) $\bar{x} = \bar{x}_0 = 48{,}1$, при конкурирующей гипотезе $H_1 : \bar{x} \neq 48{,}1$;

б) $\bar{x} = \bar{x}_0 = 48{,}1$, при конкурирующей гипотезе $H_1 : \bar{x} < 48{,}1$;

в) $\bar{x} = \bar{x}_0 = 48{,}1$, при конкурирующей гипотезе $H_1 : \bar{x} > 48{,}1$.

Решение.

Необходимо рассмотреть критерий $K=u$,

где $u_н = \dfrac{(\bar{x}-\bar{x}_0)\cdot\sqrt{n}}{\sigma} = \dfrac{(52-48{,}1)\cdot\sqrt{50}}{9{,}4} = 2{,}93.$

а) По условию конкурирующая гипотеза имеет вид $\bar{x} \neq 48{,}1$, поэтому критическая область двусторонняя. Найдем критическую точку из равенства $Ф(u_{кр.а/2}) = (1-\alpha)/2 = (1-0{,}05)/2 = 0{,}475.$

Согласно приложения Г, $u_{кр.}$=1,96. Так как $u_{н.} > u_{кр.}$, то следует отклонить нулевую гипотезу, значит выборочная и гипотетическая генеральная средняя статистически различаются значимо.

б) По условию конкурирующая гипотеза имеет вид $\bar{x}$<38,1, поэтому критическая область левосторонняя. Найдем критическую точку из равенства $Ф(u_{кр.,а}) = (1-2\alpha)/2 = (1-0{,}1)/2 = 0{,}45.$

Согласно приложения Г, $u_{кр.} = +1{,}65.$ Так как $u_{н.} > u_{кр.}$, то следует отклонить нулевую гипотезу H_0, то есть выборочная и гипотетическая генеральная средняя статистически различаются значимо.

Замечание. Вид альтернативной гипотезы существенно влияет на статистический вывод, поэтому ее нужно выдвигать, исходя из реального смысла рассматриваемой задачи. Практически в качестве альтернативной гипотезы следует выбрать такую, которая побуждает лицо, принимающее решение, к действиям. Например, при проверке гипотезы о том, что средний вес взятого продукта соответствует объявленному номиналу (H_0), в качестве альтернативной гипотезы (H_1) следует принять гипотезу – средний вес меньше номинала (покупателя обвешивают).

Пример 4.3. Оценить существенность различий в средней урожайности двух сортов озимой пшеницы. Для первого сорта средняя урожайность $\bar{x}_1$=45,6 ц/га и выборочная дисперсия S_1^2=8,05, а для второго сорта средняя урожайность $\bar{x}_2 = 55{,}4$ ц/га и выборочная дисперсия $S_2^2 = 14{,}31$. Объемы выборки $n_1 = 5$ 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 независимы, причем, $n_1 = n_2$, то применим критерий t – Стьюдента с $k = n_1 + n_2 - 2$ степенями свободы.

$$t_н = \frac{|\bar{x}_1 - \bar{x}_2|}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}} = \frac{55{,}4 - 45{,}6}{\sqrt{\frac{8{,}05}{5} + \frac{14{,}31}{5}}} = \frac{9{,}8}{2{,}11} = 4{,}64.$$

По приложению В определим критическое значение:

$$k = 5 + 5 - 2 = 8,\ t_{кр} = t_{0{,}05;8} = 2{,}31.$$

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

Замечание. Если выборки малые и неравные ($n_1 \neq n_2$), то нулевую гипотезу о равенстве двух выборочных средних ($H_0 : \bar{x}_1 = \bar{x}_2$), проверяют с использованием критерия t - Стьюдента с ($n_1 + n_2 - 2$) степенями свободы:

$$t_н = \frac{|\bar{x}_1 - \bar{x}_2|}{\sqrt{\frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2} \cdot \frac{n_1 + n_2}{n_1 \cdot n_2}}}.$$

Пример 4.4. По двум районам проверить гипотезу о равенстве двух средних выборочных урожайностей $H_0 : \bar{x}_1 = \bar{x}_2$ при уровне значимости 0,05. По первому району проведено 60 наблюдений, по второму 70. Средняя урожайность по первому району 38,6 ц/га при среднем квадратическом отклонении 5 ц/га, по второму району средняя урожайность 43 ц/га при среднем квадратическом отклонении 10 ц/га.

Решение.

$$u_н = \frac{|\bar{x}_1 - \bar{x}_2|}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} = \frac{|38{,}6 - 43|}{\sqrt{\frac{25}{60} + \frac{100}{70}}} = 3{,}24.$$

Рассмотрим конкурирующую гипотезу $H_1 : \bar{x}_1 \neq \bar{x}_2$, следовательно, критическая область двусторонняя. Найдем критическую точку из равенства $\Phi(u_{кр.a/2}) = (1-\alpha)/2 = (1-0{,}05)/2 = 0{,}475.$ Согласно приложения Г, $u_{кр.}$=1,96. Так как $u_н > u_{кр}$, то следует отклонить нулевую гипотезу, то есть выборочные средние урожайности различаются статистически значимо.

Пример 4.5. Два сорта озимой пшеницы испытывались на одинаковом числе участков на протяжении семи лет (таблица 4.4). При уровне значимости α = 0,05 проверить нулевую гипотезу о существенности различий в урожайности двух сортов озимой пшеницы.

Решение.

Так как имеются две зависимости выборки, т.е. существует определенная корреляция между урожайностью сортов по годам, то необходимо оценить значимость не разности двух выборочных средних, а средней разности $\bar{d} = \overline{x_1 - x_2}$.

Выдвигаем нулевую гипотезу, что средняя величина различий в урожайности пшеницы равна нулю: $H_0 : \bar{d} = 0$ при $H_1 : \bar{d} \neq 0$.

По данным таблицы 4.4 найдем среднюю разность $\bar{d}$ и ошибку средней разности $S_{\bar{d}}$:

$$\bar{d} = \frac{\sum d_i}{n} = \frac{28}{7} = 4; \quad S_{\bar{d}} = \sqrt{\frac{\sum (d_i - \bar{d})^2}{n(n-1)}} = \sqrt{\frac{48}{7(7-1)}} = 1{,}07; \; t_н = \frac{\bar{d}}{S_{\bar{d}}} = \frac{4}{1{,}07} = 3{,}74,$$

где $d_i = x_{1i} - x_{2i}$, n – число пар наблюдений.

Таблица 4.4 - Вспомогательная таблица для расчета ошибки средней разности

Год	Урожайность, ц/га		Разность $d_i = x_{1i} - x_{2i}$	$(d_i - \bar{d})$	$(d_i - \bar{d})^2$
	x_{1i}	x_{2i}			
2008	53	47	6	2	4
2009	49	48	1	-3	9
2010	45	46	-1	-5	25
2011	56	51	5	1	1
2012	58	52	6	2	4
2013	55	50	5	1	1
2014	59	53	6	2	4
Сумма	-	-	28	-	48

При $\alpha = 0{,}05$; $k = n - 1 = 7 - 1 = 6$, $t_{0,05;6} = 2{,}45$.

Сопоставив наблюдаемое значение t с критическим, можно сделать вывод, что два сорта существенно различаются по уровню урожайности.

Задачи для самостоятельного решения

Задача 4.1. Проводилось испытание 6 сортов озимой пшеницы. Каждый сорт высевался на 6 делянках одинаковой площади (таблица 4.5). При 5 % уровне значимости проверить гипотезу о существенности различий в средней урожайности двух сортов озимой пшеницы (комбинация сортов предлагается студенту преподавателем).

Таблица 4.5 – Урожайность озимой пшеницы по сортам, ц/га

Повторение	Сорт					
	«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 4.6 – Урожайность озимого ячменя по сортам, ц/га

Номер испытания	Сорт			
	«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 4.7).

Таблица 4.7 – Урожайность и доля овощей по районам

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. Два сорта озимой пшеницы испытывались на одинаковом числе участков в течение 8 лет. При уровне значимости 0,05 проверить гипотезу о существенности различий в урожайности двух сортов озимой пшеницы.

Таблица 4.8 – Урожайность озимой пшеницы, ц/га

Год	Номер варианта (сорта)							
	I	II	III	IV	V	VI	VII	VIII
2007	41	38	36	43	45	54	42	65
2008	53	48	45	56	58	50	54	60
2009	55	50	47	61	65	49	60	62
2010	42	40	46	48	49	57	50	59
2011	46	42	41	49	50	55	49	64
2012	58	54	55	63	68	53	64	63
2013	46	41	40	49	52	56	49	58
2014	57	53	60	50	62	52	61	62

Задача 4.5. Определённые сорта озимой пшеницы (приложение Ж) испытывались на одинаковом числе участков на протяжении семи лет. При уровне значимости $\alpha = 0{,}05$ 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
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 4.6. По данным таблицы 2.2 (пример 2.2) проверить нулевую гипотезу о нормальном распределении сельскохозяйственных организаций по урожайности кукурузы на зерно. Уровень значимости принять равным 0,05.

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.

5 Дисперсионный анализ

В практике обработки результатов наблюдений часто возникает вопрос о том, насколько существенное влияние оказывает изменение одного или нескольких факторов на измеряемую величину. Например, можно исследовать влияние количества вносимых удобрений на урожайность сельскохозяйственной культуры. При более детальном исследовании причин, влияющих на урожайность, возможно, потребуется проверить влияние двух факторов, например, количества удобрений и сортов культуры. Возможно рассмотрение влияния и большего числа факторов.

Многообразие различий в природно-климатических и хозяйственных условиях затрудняет точное определение результатов воздействия того или иного фактора на урожайность. В связи с этим в сельском хозяйстве широкое применение получил полевой опыт, который проводится с определенной повторностью. Введение повторности делянок дает возможность определить размер случайных колебаний, соответственно сравнить случайные отклонения с различиями, обусловленными определенными факторами, а случайную (остаточную) дисперсию – с систематической, вызванной факторами опыта. Для этого осуществляется предварительная оценка результатов опыта, путем проведения дисперсионного анализа с использованием критерия F – Фишера-Снедекора.

Дисперсионный анализ основан на разложении общей дисперсии результативного признака на составляющие (компоненты), сравнивая которые друг с другом посредством критерия F, можно определить долю общей вариации изучаемого (результативного) признака, обусловленную действием на него как регулируемых, так и не регулируемых в опыте факторов. Критерий Фишера представляет собой отношение двух дисперсий (межгрупповой к остаточной) и служит критерием оценки влияния на результативный признак изучаемых в опыте факторов

$$F_{н} = \frac{S_v^2}{S_z^2}. \quad (5.1)$$

В зависимости от количества факторов, дисперсионный анализ подразделяется на однофакторный и многофакторный.

Алгоритм дисперсионного анализа зависит от модели и схемы постановки опыта или эксперимента.

Предположим, что на нормально распределенный результативный признак оказывает влияние какой-то один фактор А. Исследователь выделил λ уровней этого фактора. Каждый уровень повторен *n* раз. Результаты опыта по уровням фактора и повторениям независимы. Обозначим: i=1,2,……, λ; j=1,2,….,n; x_{ij} - значение результативного признака по i-тому уровню фактора и j-тому повторению. Тогда

$$x_{ij} = \bar{x} + A_i + \varepsilon_{ij}, \quad (5.2)$$

где $\bar{x}$ – общая средняя результативного признака по опыту или эксперименту;

A_i – эффект влияния фактора А;

ε_{ij} – влияние случайных факторов.

Если опыт поставлен способом организованных повторений (рендомизированных блоков), то

$$x_{ij} = \bar{x} + A_i + P_j + \varepsilon_{ij}, \quad (5.3)$$

где P_j - эффект повторений.

В дисперсионном анализе расчеты проводят в следующей последовательности:

- рассчитываются средние значения признака по каждому варианту опыта и каждому повторению;
- определяются необходимые суммы квадратов отклонений по результативному признаку;
- находится число степеней свободы вариации по каждому источнику;
- рассчитываются средние квадраты отклонений (дисперсии);
- определяются наблюдаемые и критические значения критерия Фишера-Снедекора;
- оценивается значимость различий групповых средних по вариантам опыта.

Таблица 5.1 – Схема однофакторного дисперсионного анализа

Источник вариации	Суммы квадратов отклонений, SS	Число степеней свободы, k	Средний квадрат отклонений (дисперсии), S^2	F – критерий
Вариантов (фактора А)	$SS_v = n\left(\sum_{i=1}^{\lambda}(\bar{x}_i - \bar{x})^2\right)$	$k_v = \lambda - 1$	$S_v^2 = \frac{SS_v}{k_v}$	$F_н = \frac{S_v^2}{S_z^2}$
Повторений	$SS_p = \lambda\left(\sum_{j=1}^{n}(\bar{x}_j - \bar{x})^2\right)$	$k_p = n - 1$	$S_p^2 = \frac{SS_p}{k_p}$	$F_н = \frac{S_p^2}{S_z^2}$
Остаточная	$SS_z = SS_o - SS_p - SS_v$	$k_z = (n-1)(\lambda-1)$	$S_z^2 = \frac{SS_z}{k_z}$	
Общая	$SS_0 = \sum_{i=1}^{\lambda}\sum_{j=1}^{n}(x_{ij} - \bar{x})^2$	$k_0 = n\lambda - 1$	$S^2 = \frac{SS_o}{k_o}$	

В приведенных в таблице 5.1 формулах n – число повторений; λ - число групп (вариантов опыта); $N = \lambda \cdot n$ – общее число наблюдений.

Критическое значение F находится по таблице распределения F – Фишера-Снедекора (приложение Е) при уровне значимости α и числе степеней свободы k_v и k_z .

Если $F_н > F_{кр}$, то утверждение о существенности различий результативного признака, обусловленных влиянием фактора на него, признается статистически достоверным, т. е. имеются существенные различия между средними по вариантам опыта.

Нулевая гипотеза (H_0) сводится к предположению, что генеральные групповые средние равны между собой: $H_o : \bar{x}_1 = \bar{x}_2 = ... = \bar{x}_\lambda$. Это предположение ($H_0$) опровергается, если $F_н > F_{кр}$ для принятого уровня значимости (α) и числа степеней свободы k_v и k_z . В противном случае ($F_н < F_{кр}$) нулевая гипотеза принимается.

Пример 5.1. Провести дисперсионный анализ полевого опыта по сортоиспытанию (таблица 5.2). Оценить существенность различий в урожайности 4 сортов озимой пшеницы с помощью наименьшей разности в сравнении с первым вариантом, взятым за стандарт.

Решение.

1. Найдем среднюю урожайность озимой пшеницы по вариантам опыта (сортам), повторностям и в целом по опыту (таблица 5.2).

Таблица 5.2 – Урожайность озимой пшеницы, ц/га

Номер сорта	Повторности					Сумма Σx_i	Средняя урожайность, $\bar{x}_i$
	1	2	3	4	5		
1	41	38	36	43	40	198	39,6
2	57	52	51	55	54	269	53,8
3	51	50	47	49	52	249	49,8
4	42	40	36	48	46	212	42,4
Сумма, Σx_j	191	180	170	195	192	928	-
Средняя урожайность, $\bar{x}_j$	47,75	45	42,5	48,75	48	-	46,4

2. Рассчитаем суммы квадратов отклонений индивидуальных значений урожайности от общей средней. Для этого построим вспомогательную таблицу 5.3.

Таблица 5.3 – Квадраты отклонений урожайности

Номер сорта	$(x_{ij} - \bar{x})^2$					Сумма
	1	2	3	4	5	
1	29,16	70,56	108,16	11,56	40,96	260,4
2	112,36	31,36	21,16	73,96	57,76	296,6
3	21,16	12,96	0,36	6,76	31,36	72,6
4	19,36	40,96	108,16	2,56	0,16	171,2
Сумма	182,04	155,84	237,84	94,84	130,24	800,8

3. Определим суммы квадратов отклонений вариантов (сортов):

$$SS_v = n\sum_{i=1}^{\lambda}(x_i - \bar{x})^2 = 5((39{,}6 - 46{,}4)^2 + (53{,}8 - 46{,}4)^2 + (49{,}8 - 46{,}4)^2 + (42{,}4 - 46{,}4)^2) = 642{,}8;$$

повторений $SS_p = \lambda\sum_{j=1}^{n}(\bar{x}_j - \bar{x})^2 = 4((47{,}75 - 46{,}4)^2 + (45 - 46{,}4)^2 + (42{,}5 - 46{,}4)^2 +$

$+(48{,}75-46{,}4)^2+(48-46{,}4)^2)=108{,}3;$

общая $SS_0=\sum_{i=1}^{\lambda}\ \sum_{j=1}^{n}(x_{ij}-\overline{x})^2=800{,}8;$

остаточная $SS_z=SS_0-SS_p-SS_v=800{,}8-108{,}3-642{,}8=49{,}7$.

4. Найдем число степеней свободы:

вариантов $K_v=\lambda-1=4-1=3$;

повторений $K_p=n-1=5-1=4$;

остаточное $K_z=(n-1)(\lambda-1)=4\cdot 3=12$;

общее $K_o=n\lambda-1=4\cdot 5-1=19.$

5. Составим таблицу дисперсионного анализа 5.4, в которой средние квадраты определяются, путем деления сумм квадратов отклонений на соответствующее число степеней свободы.

Таблица 5.4 – Результаты дисперсионного анализа

Вариация	Суммы квадратов отклонений SS	Число степеней свободы, k	Средние квадраты отклонений (дисперсии, S^2)	Наблюдаемые значения $F_н$
Вариантов (сортов)	642,8	3	214,27	51,73
Повторений	108,3	4	27,08	6,54
Остаточная	49,7	12	4,14	–
Общая	800,8	19	–	–

6. Используя данные таблицы 5.4, определим фактическое отношение дисперсий:

$$F_н=\frac{S_v^2}{S_z^2}=\frac{214{,}27}{4{,}14}=51{,}73.$$

При вероятности γ=0,95 (уровне значимости 0,05) и числе степеней свободы k_v и k_z, по таблице F – распределения (см.

приложение Е) $F_{кр} = F_{0,05}(k_1 = 3, k_2 = 12) = 3{,}49.$ Так как $F_н > F_{кр}$, то нулевая гипотеза о равенстве урожайности по сортам озимой пшеницы отвергается. Имеется хотя бы одна статистически существенная разность в средней урожайности.

Для оценки существенности частных различий вычисляют:

а) *среднюю ошибку опыта*

$$S_{\bar{x}} = \sqrt{\frac{s_z^2}{n}} = \sqrt{\frac{4{,}14}{5}} = 0{,}91;$$

б) *ошибку разности средних*

$$S_d = S_{\bar{x}} \cdot \sqrt{2} = 0{,}91 \cdot 1{,}414 = 1{,}29;$$

в) *наименьшую существенную разность*

$$НСР_{a,k_z} = t_{a,k_z} \cdot S_d.$$

Здесь значение t_{a,k_z} находится по таблицам распределения Стьюдента при уровне значимости α и числе степеней свободы $k_{z \cdot 0,05;12} = 2{,}18.$

$НСР_{0,05} = 2{,}18 \cdot 1{,}29 = 2{,}81.$

Рассчитаем разности в средней урожайности по сортам и сопоставим их с наименьшей существенной разностью (таблица 5.5).

Таблица 5.5 – Средняя урожайность по сортам озимой пшеницы, ц/га

Сорт	Средняя урожайность	Разность по стандартам
1	39,6	–
2	53,8	14,2
3	49,8	10,2
4	42,4	2,8
$НСР_{0,05}$	–	2,81

Значит, первый и четвертый сорта статистически незначимо отличаются друг от друга, остальные (второй и третий) превосходят стандарт по урожайности, так как их разности с первым сортом больше НСР.

Задачи для самостоятельного решения

Задача 5.1. По данным таблицы 5.6 провести дисперсионный анализ полевого опыта. Оценить существенность различий в урожайности риса по различным дозам удобрений с помощью $HCP_{0,05}$.

Таблица 5.6 – Урожайность риса при различных дозах удобрений, ц/га

Доза удобрений	Повторения							
	1	2	3	4	5	6	7	8
Без удобрений	41,8	43,1	42,7	41,6	38,7	40,5	42,3	45,3
$N_0P_{90}K_{60}$	43,3	45,2	44,7	44,8	42,4	41,6	46,4	44,8
$N_{60}P_{90}K_{60}$	45,3	47,5	46,8	46,0	46,2	44,3	45,8	47,3
$N_{90}P_{90}K_{60}$	47,3	49,8	48,8	48,5	49,3	47,9	48,6	49,4
$N_{120}P_{90}K_{60}$	51,5	53,5	52,7	51,7	51,6	51,0	49,8	50,0
$N_{150}P_{90}K_{60}$	59,3	57,3	56,5	55,3	54,1	58,7	55,8	54,4
$N_{180}P_{90}K_{60}$	48,6	49,8	50,1	49,3	47,4	50,2	49,8	49,0

Задача 5.2. По данным таблицы 5.7 определить, является ли существенным влияние предшественника на урожайность риса.

Таблица 5.7 – Урожайность риса сорта «Вираж» в зависимости от предшественника

Предшественник	Повторения				
	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

Задача 5.3. Определить, является ли существенным влияние способов основной обработки почвы на урожайность сахарной свеклы (таблица 5.8).

Таблица 5.8 – Урожайность сахарной свеклы при разных способах обработки почвы, т/га

Вариант опыта	Повторения			
	1	2	3	4
Контроль (дисковое лущение стерни + вспашка в августе)	31,5	31,7	31,0	32,0
Полупар с перепашкой	36,8	37,0	36,0	37,5
Дисковое лущение стерни + мелкая вспашка в июле + вспашка в сентябре	32,5	32,0	32,7	33,0
Дисковое лущение + вспашка в сентябре	34,3	34,5	34,0	35,0
Комбинированная обработка	35,5	35,0	36,0	35,7

Задача 5.4. По данным таблицы 5.9 оценить значимость различий в средней урожайности сортов яровой пшеницы. Уровень значимости равен 0,05.

Таблица 5.9 – Урожайность яровой пшеницы, ц/га

Номер сорта	Повторения			
	1	2	3	4
1	28,1	29,2	30,0	27,3
2	25,0	24,3	28,5	29,2
3	27,2	26,4	31,0	26,4
4	23,6	27,2	25,2	24,8
5	30,0	33,0	36,0	29,8
6	23,0	26,0	26,0	24,8

Задача 5.5. По данным таблицы 5.10 провести дисперсионный анализ полевого опыта. Оценить существенность различий в

урожайности картофеля по различным дозам удобрений с помощью $HCP_{0,05}$.

Таблица 5.10 – Урожайность картофеля сорта «Невский» по вариантам опыта, ц/га

Доза удобрений	Повторения				
	1	2	3	4	5
$N_{67}P_{50}K_{62}$	263	259	268	272	283
$N_{47}P_{50}K_{72}$	231	243	229	251	230
$N_{93}P_{100}K_{140}$	225	229	223	231	219
$N_{140}P_{150}K_{215}$	214	208	218	210	211

Задача 5.6. По данным таблицы 5.11 оценить существенность влияния дозы минеральных удобрений на число продуктивных стеблей озимой пшеницы на 1 м^2. Уровень значимости принять 0,05.

Таблица 5.11 – Количество продуктивных стеблей озимой пшеницы на 1 м^2, шт.

Доза удобрений	Повторения				
	1	2	3	4	5
$N_{60}P_{60}K_{30}$	373	382	369	395	378
$N_{68}P_{44}K_{24}$	404	400	411	415	409
$N_{90}P_{67}K_{40}$	425	429	421	419	427
$N_{126}P_{80}K_{72}$	432	435	439	441	433

Задача 5.7. По данным таблицы 5.12 оценить существенность влияния дозы минеральных удобрений на число зерен в колосе озимой пшеницы. Уровень значимости принять 0,05.

Таблица 5.12 – Количество зерен в колосе озимой пшеницы, шт.

Доза удобрений	Повторения			
	1	2	3	4
$N_{60}P_{60}K_{30}$	25	23	22	24
$N_{60}P_{34}K_{34}$	27	29	28	26
$N_{110}P_{82}K_{51}$	30	33	34	32

Задача 5.8. По данным таблицы 5.13 оценить значимость различий в средней урожайности отдельных сортов ярового ячменя. Уровень значимости принять 0,05.

Таблица 5.13 – Урожайность озимого ячменя, ц/га

Сорт	Повторения			
	1	2	3	4
«Атаман»	55,3	61,9	59,0	58,6
«Виват»	46,3	46,9	48,9	47,5
«Визит»	43,3	42,5	41,8	42,9
«Маргрет»	43,5	45,3	45,6	44,8

Задача 5.9. По данным таблицы 5.14 оценить значимость различий в средней урожайности отдельных сортов озимой пшеницы. Уровень значимости принять 0,05.

Таблица 5.14 – Урожайность озимой пшеницы, ц/га

Сорт	Повторения			
	1	2	3	4
«Восторг»	41,3	47,3	45,9	47,4
«Красота»	47,6	48,9	49,5	50,1
«Фишт»	57,5	52,0	53,7	63,4
«Фортуна»	64,8	57,1	66,1	63,2

Задача 5.10. По данным таблицы 5.15 провести дисперсионный анализ полевого опыта. Оценить существенность различий в урожайности озимой пшеницы по различным дозам удобрений с помощью $НСР_{0,05}$.

Таблица 5.15 – Урожайность озимой пшеницы, ц/га

Доза удобрений	Повторения			
	1	2	3	4
$N_{60}P_{60}K_{30}$	45,9	43,0	46,0	42,8
$N_{60}P_{34}K_{34}$	38,2	41,5	37,2	39,4
$N_{68}P_{44}K_{24}$	41,8	43,9	39,3	42,1
$N_{105}P_{60}K_{60}$	52,2	46,3	43,4	48,5
$N_{90}P_{67}K_{40}$	46,3	49,7	42,1	48,8
$N_{126}P_{80}K_{72}$	58,6	60,2	59,1	60,4
$N_{110}P_{82}K_{51}$	56,8	58,0	56,1	57,3

Вопросы для самоподготовки

1. На чем основан дисперсионный анализ?
2. В чем заключается основная задача дисперсионного анализа?
3. Назовите виды дисперсионного анализа.
4. Перечислите этапы проведения дисперсионного анализа.
5. По какой формуле определяется наблюдаемое значение критерия Фишера-Снедекора?
6. Что такое наименьшая существенная разность и для чего она используется?

6 Корреляционно-регрессионный анализ

Корреляционно - регрессионный анализ – это совокупность статистических и математических методов, используемых для количественного анализа связей между различными явлениями и процессами. При корреляционной связи изменение результативного признака (y) обусловлено влиянием факторных признаков ($x_1, x_2, ..., x_n$). В зависимости от числа признаков, между которыми изучается связь, различают парную и множественную связь. Если изучается связь между результативным признаком, двумя и более факторными признаками, то она называется множественной.

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). На поле графика отмечают точки, соответствующие индивидуальным значениям признаков, и по характеру их расположения судят о форме и направлении связи.

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:

$$\begin{cases} \Sigma y = an + b\Sigma x, \\ \Sigma yx = a\Sigma x + b\Sigma x^2. \end{cases} \tag{6.1}$$

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. Если имеет место прямая связь, то $0 \le r \le 1$, если обратная, то $-1 \le r \le 0$. Коэффициент корреляции рассчитывается по формуле

$$r = \frac{\overline{xy} - \bar{x} \cdot \bar{y}}{\sigma_x \cdot \sigma_y}. \tag{6.2}$$

При возведении коэффициента корреляции в квадрат и выражении этого результата в процентах получают коэффициент детерминации:

$$Д = r^2 \cdot 100\ \%. \tag{6.3}$$

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

Коэффициент эластичности определяется по формуле

$$Э = b \cdot \frac{\bar{x}}{\bar{y}}. \tag{6.4}$$

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:

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

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 $\tau \approx \frac{2}{3}\rho$.

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 ($K_к$):

$$K_a = \frac{ad - bc}{ad + bc} \qquad (6.7),$$

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

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 $|K_к| \geq 0,3$.

Пример 6.1. По данным об урожайности озимой пшеницы и количестве внесенных минеральных удобрений на один гектар посева определить: форму связи между признаками, параметры уравнения регрессии, тесноту связи между признаками. Исчислить коэффициенты корреляции, детерминации и эластичности. Сделать вывод.

Решение.

Урожайность озимой пшеницы зависит от количества внесенных удобрений, поэтому факторным признаком является количество внесенных удобрений на 1 га посева, а результативным признаком – урожайность (рисунок 6.1). Нанесем на график пары значений x и y, представленные в таблице 6.1.

По характеру расположения точек на графике видно, что зависимость можно выразить уравнением: $y = a + bx$.

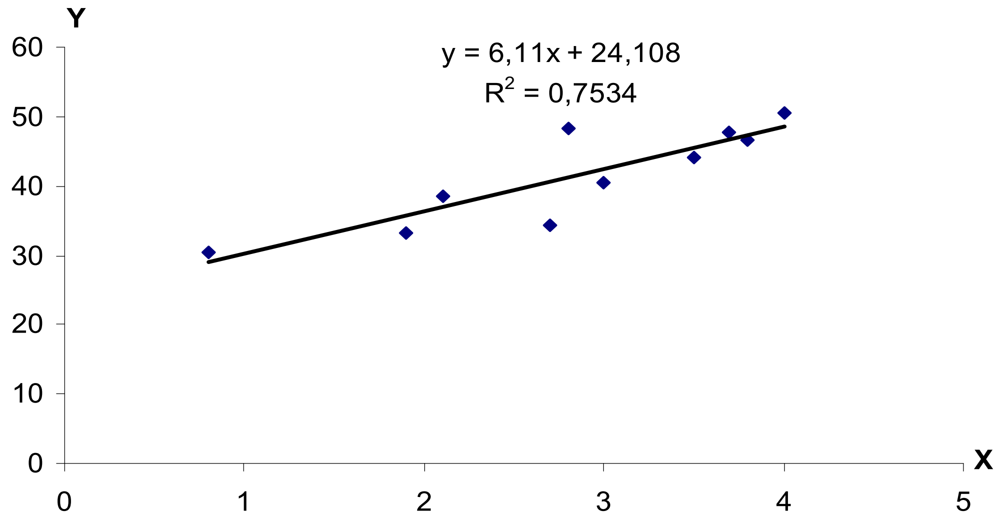


Рисунок 6.1 – Зависимость между урожайностью озимой пшеницы и количеством внесенных удобрений

Заполним вспомогательную таблицу 6.1.

Таблица 6.1 – Урожайность озимой пшеницы и количество внесенных минеральных удобрений на 1 га посева

№ организации	Урожайность, ц/га (y)	Внесено удобрений на 1 га, ц д. в. (x)	y^2	x^2	xy	Теоретическое значение $\hat{y}=a+bx$
1	44,2	3,5	1953,64	12,25	154,70	45,49
2	40,4	3,0	1632,16	9,00	121,20	42,44
3	48,3	2,8	2332,89	7,84	135,24	41,22
4	33,1	1,9	1095,61	3,61	62,89	35,72
5	46,5	3,8	2162,25	14,44	176,70	47,33
6	50,5	4,0	2550,25	16,00	202,00	48,55
7	30,4	0,8	924,16	0,64	24,32	29,00
8	47,6	3,7	2265,76	13,69	176,12	46,71
9	38,6	2,1	1489,96	4,41	81,06	36,94
10	34,4	2,7	1183,36	7,29	92,88	40,60
Итого	414,0	28,3	17590,04	89,17	1227,11	414,00

Найдем параметры уравнения, составив систему двух уравнений по данным таблицы 6.1. Система уравнений будет иметь вид:

$$\begin{cases} 414{,}0 = 10a + 28{,}3b, \\ 1227{,}11 = 28{,}3a + 89{,}17b. \end{cases}$$

Решив систему, получим: a = 24,1;
b = 6,11.

Тогда, уравнение связи между урожайностью и количеством внесенных удобрений будет выглядеть следующим образом: *y = 24,1+6,11x*.

Значит, при увеличении количества вносимых минеральных удобрений на 1 га посева на 1 ц д. в., урожайность озимой пшеницы в среднем увеличивается на 6,11 ц/га.

Рассчитаем коэффициент корреляции.

$$\bar{x} = \frac{\Sigma x}{n} = \frac{28{,}3}{10} = 2{,}83; \qquad \bar{y} = \frac{\Sigma y}{n} = \frac{414}{10} = 41{,}4;$$

$$\overline{xy} = \frac{\Sigma xy}{n} = \frac{1227{,}11}{10} = 122{,}711;$$

$$\sigma_x = \sqrt{\frac{\Sigma x^2}{n} - \left(\bar{x}^2\right)} = \sqrt{\frac{89{,}17}{10} - 2{,}83^2} = 0{,}953;$$

$$\sigma_y = \sqrt{\frac{\Sigma y^2}{n} - \left(\bar{y}^2\right)} = \sqrt{\frac{17590{,}04}{10} - 41{,}4^2} = 6{,}711.$$

$$r = \frac{\overline{xy} - \bar{x} \cdot \bar{y}}{\sigma_x \cdot \sigma_y} = \frac{122{,}711 - 2{,}83 \cdot 41{,}4}{0{,}953 \cdot 6{,}711} = 0{,}868.$$

Следовательно, между изучаемыми признаками существует прямая очень тесная связь.

Коэффициент детерминации составит:

$$Д = r^2 \cdot 100 = 0{,}868^2 \cdot 100 \approx 75{,}3\ \%.$$

Таким образом, вариация урожайности озимой пшеницы на 75,3 % обусловлена вариацией количества внесенных на 1 га минеральных удобрений.

Коэффициент эластичности равен:

$$Э = b\frac{\bar{x}}{\bar{y}} = 6{,}11\frac{2{,}83}{41{,}4} = 0{,}42.$$

Изменение количества вносимых минеральных удобрений на 1 га на 1 % приводит к изменению урожайности на 0,42 %.

Так как зависимость между урожайностью и количеством минеральных удобрений изучалась по выборочным данным, то необходимо оценить значимость коэффициента корреляции. Выдвигаем нулевую гипотезу – величина коэффициента корреляции в генеральной совокупности равна нулю $H_o : Z_r = 0$, при альтернативной $H_1 : Z_r \neq 0$. Проверку нулевой гипотезы проведем с помощью критерия *t*-Стьюдента.

Найдем наблюдаемое значение критерия:

$$t_н = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}} = \frac{0{,}868 \cdot \sqrt{10-2}}{\sqrt{1-0{,}753}} \approx 4{,}94.$$

При α = 0,05 и числе степеней свободы $k = n - 2 = 8$ по таблице $t_{0,05;8} = 2{,}31$.

Так как $t_н > t_{0,05;8}$, то нулевая гипотеза отвергается, коэффициент корреляции существенно отличен от нуля. Значит, применение минеральных удобрений оказывает статистически существенное влияние на урожайность озимой пшеницы.

Пример 6.2. Используя данные предыдущей задачи, с помощью коэффициентов корреляции рангов Спирмена и Кендалла, измерить тесноту связи между количеством внесенных минеральных удобрений на 1 га посева и урожайностью озимой пшеницы.

Решение.

1) Для удобства расчета все вспомогательные вычисления проведем в таблице 6.2, расположив организации в порядке возрастания факторного признака.

Таблица 6.2 – Вспомогательная таблица для расчета коэффициентов корреляции рангов Спирмена и Кендалла

x_i	y_i	R_{x_i}	R_{y_i}	$d_i = R_{x_i} - R_{y_i}$	d_i^2	Число рангов R_{yi}	
						большего порядка	меньшего порядка
0,8	30,4	1	1	0	0	9	0
1,9	33,1	2	2	0	0	8	0
2,1	38,6	3	4	-1	1	6	1
2,7	34,4	4	3	1	1	6	0
2,8	48,3	5	9	-4	16	1	4
3,0	40,4	6	5	1	1	4	0
3,5	44,2	7	6	1	1	3	0
3,7	47,6	8	8	0	0	1	1
3,8	46,5	9	7	2	4	1	0
4,0	50,5	10	10	0	0	–	–
x	x	x	x	x	Σd_i^2=24	$P = 39$	$Q = 6$

2) Определим коэффициент корреляции рангов Спирмена:

$$\rho = 1 - \frac{6 \cdot 24}{10 \cdot (10^2 - 1)} = 0{,}855.$$

3) Рассчитаем коэффициент корреляции рангов Кендалла:

$$\tau = \frac{2 \cdot (39 - 6)}{10 \cdot (10 - 1)} = 0{,}733.$$

Следовательно, между количеством внесенных минеральных удобрений и урожайностью озимой пшеницы существует тесная связь.

Задачи для самостоятельного решения

Задача 6.1. По данным 20 организаций об урожайности и продолжительности уборки озимой пшеницы (приложение Б) определить: форму связи между признаками; параметры уравнения регрессии; коэффициенты корреляции, детерминации и эластичности. Сделать вывод.

Задача 6.2. По данным 20 организаций об урожайности озимой пшеницы и качестве почв в баллах (приложение Б) определить: форму связи между признаками; параметры уравнения регрессии; тесноту связи между признаками. Сделать вывод.

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

Задача 6.4. По данным приложения Б 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. По результатам расчетов написать вывод.

Задача 6.5. По данным 20 организаций (приложение Б) с помощью коэффициентов корреляции рангов Спирмена и Кендалла измерить тесноту связи между:

1) продолжительностью уборки и урожайностью озимой пшеницы;

2) качеством почв и урожайностью озимой пшеницы;

3) количеством внесенных минеральных удобрений и урожайностью озимой пшеницы.

Задача 6.6. По данным таблицы 6.3 с помощью коэффициентов ассоциации и контингенции оценить наличие связи между окраской плодов томатов в третьем и во втором поколении.

Таблица 6.3 – Число растений томатов с разной окраской плодов

Во втором поколении	В третьем поколении	
	красная и малиновая	желтая и желто-красная
Красная и красно-желтая	38	72
Желтая и желто-красная	10	78

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?

7 Ряды динамики

Рядом динамики называется последовательный ряд чисел, характеризующий изменение уровней явления или процесса во времени. Он состоит из двух элементов: уровней ряда (y); периодов или моментов времени (t). При построении ряда динамики необходимо обеспечить сопоставимость уровней за весь изучаемый период времени.

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 1января 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.

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

Для характеристики развития явления во времени определяют показатели: абсолютный прирост, темп роста, темп прироста базисным и цепным способом, значение одного процента прироста.

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.

Таблица 7.1- Расчет показателей ряда динамики

Показатель	Метод расчета	
	базисный (с постоянной базой)	цепной (с переменной базой)
Абсолютный прирост (A)	$A_{б_i} = y_i - y_1$ (7.1)	$A_{ц_i} = y_i - y_{i-1}$ (7.2)
Коэффициент роста ($К_р$)	$К_{рб_i} = \frac{y_i}{y_1}$ (7.3)	$К_{рц_i} = \frac{y_i}{y_{i-1}}$ (7.4)
Темп роста ($Т_р$)	$Т_{рб_i} = К_{рб_i} \cdot 100\%$ (7.5)	$Т_{рц_i} = К_{рц_i} \cdot 100\%$ (7.6)
Темп прироста ($Т_{пр}$)	$Т_{прб_i} = Т_{рб_i} - 100\%$ (7.7)	$Т_{прц_i} = Т_{рц_i} - 100\%$ (7.8)
Абсолютное значение 1 % прироста ($Зн.1\%$)	$Зн.1\% = 0{,}01 \cdot y_{i-1}$ или $Зн.1\% = \frac{А_ц}{Т_{пр\ ц}}$ (7.9)	

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.

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 ($\overline{\text{А}}$)	$\bar{A} = \frac{y_n - y_1}{n-1}$ или $\bar{A} = \frac{\Sigma A_ц}{n-1}$ (7.13)
Средний коэффициент роста ($\overline{\text{К}}_\text{р}$)	$\overline{K_p} = \sqrt[n-1]{ПK_ц}$ или $\overline{K_p} = \sqrt[n-1]{\frac{y_n}{y_1}}$ (7.14)
Средний темп роста ($\overline{\text{Т}_\text{р}}$), %	$\overline{T_p} = \overline{K_p} \cdot 100$ % (7.15)
Средний темп прироста ($\overline{\text{Т}_\text{пр}}$), %	$\overline{T}_{пр} = \overline{T}_p - 100$ % или $\overline{T}_{пр} = (\overline{K}_p - 1) \cdot 100\%$ (7.16)
Среднее значение 1% прироста ($\overline{Зн.1\%}$)	$\overline{Зн.1\%} = \frac{\bar{A}}{\overline{\overline{T}}_{пр}}$ (7.17)

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.

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 в Краснодарском крае (приложение И) рассчитать показатели ряда динамики. Сделать вывод.

Решение.

Расчеты оформим в таблице 7.3.

Вывод: расчеты показали, что средняя урожайность чайного листа в динамике за пять лет составила 2,3 ц/га. При этом ежегодно она снижалась в среднем на 0,95 ц/га или на 33,8 %. Один процент прироста соответствовал в среднем 0,028 ц/га.

Таблица 7.3 - Расчет показателей ряда динамики урожайности чайного листа

Год	Урожай-ность, 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%*
2009	4,7	-	-	-	-	100	100	-	-	-
2010	2,7	-2	-2	0,574	0,574	57,4	57,4	-42,6	-42,6	0,047
2011	2,2	-2,5	-0,5	0,468	0,815	46,8	81,5	-53,2	-18,5	0,027
2012	1,0	-3,7	-1,2	0,213	0,455	21,3	45,5	-78,7	-54,5	0,022
2013	0,9	-3,8	-0,1	0,191	0,900	19,1	90,0	-80,9	-10,0	0,010
Средние показатели	2,3	-0,95		0,662		66,2		-33,8		0,028

Пример 7.2. Определить тенденцию изменения урожайности чайного листа в хозяйствах Краснодарского края методом укрупнения периодов, скользящей средней и аналитического выравнивания. Сделать вывод.

Решение.

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

$$a = \frac{\sum y}{n} = \frac{38,7}{9} = 4,3\,; \qquad b = \frac{\sum yt}{\sum t^2} = \frac{-60,9}{60} = -1,02.$$

Тогда уравнение прямой имеет вид: $\hat{y}_t = 4,3 - 1,02\ t.$

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					
		средняя по трёхлетиям	сумма	средняя	номер года	квадрат номера года	произведение параметров	выравненные значения
	y				t	t^2	yt	$\hat{y}_t = a + bt$
2005	8,9	7,2			-4	16	-35,6	8,36
2006	8,1		21,5	7,2	-3	9	-24,3	7,35
2007	4,5		18,3	6,1	-2	4	-9,0	6,33
2008	5,7	4,4	14,9	5,0	-1	1	-5,7	5,32
2009	4,7		13,1	4,4	0	0	0,0	4,30
2010	2,7		9,6	3,2	1	1	2,7	3,29
2011	2,2	1,4	5,9	2,0	2	4	4,4	2,27
2012	1,0		4,1	1,4	3	9	3,0	1,26
2013	0,9		-	-	4	16	3,6	0,24
Итого	38,7	-	-	-	0	60	-60,9	38,70

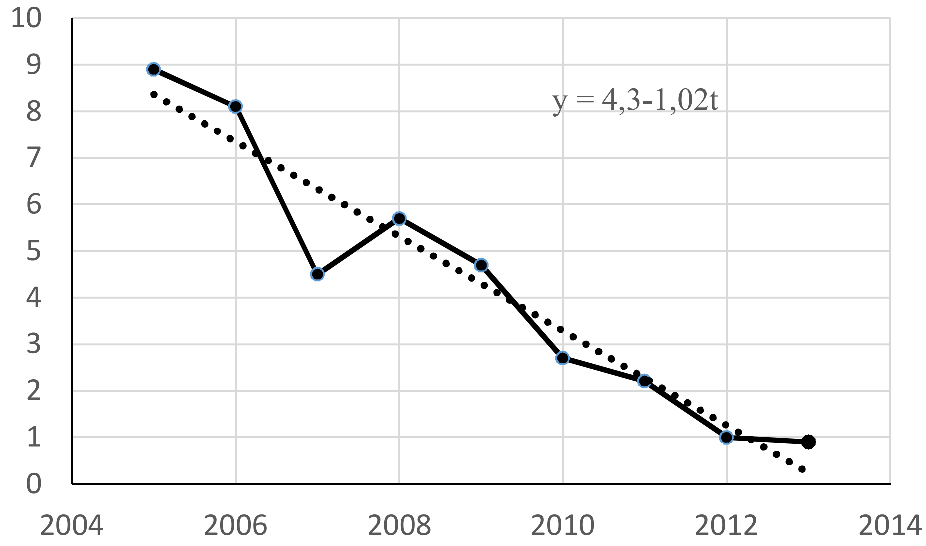


Рисунок 7.1 -Динамика урожайности чайного листа, ц/га

Вывод: расчеты показали, что средняя урожайность чайного листа за 2005-2013 гг. составила 4,3 ц/га. В среднем она ежегодно снижалась на 1,02 ц/га. На графике наглядно видна четко выраженная тенденция к снижению урожайности исследуемой культуры.

Пример 7.3. В 2011 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

Объем производства продукции, млрд руб.	Год					
	2009	2010	2011	2012	2013	2014
При старом составе	19,7	20,2	21,2			
При новом составе			22,8	24,6	25,2	26,1
Смыкание по 1 способу, млн. руб.	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

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

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

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

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 И 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 И 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. Имеются данные о численности сельскохозяйственных машин в организации на 1 января соответствующего года (таблица 7.6).

Таблица 7.6 – Динамика наличия техники в организации

Наименование машин	2009 г.	2010 г.	2011 г.	2012 г.	2013 г.	2014 г.
Тракторы гусеничные	48	45	44	41	43	45
Тракторы колесные	54	56	52	50	49	51
Комбайны зерноуборочные	20	18	17	21	19	20
кукурузоуборочные	6	6	5	4	5	4
силосоуборочные	15	13	12	14	16	18
Грузовые автомобили	22	24	26	31	30	28

Определить среднегодовую численность машин за 2009–2014 гг.

Задача 7.5. 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 7.6. 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 7.7. 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 7.8. Используя взаимосвязь показателей динамики, определить уровни ряда динамики и недостающие в таблице базисные показатели динамики по следующим данным об урожайности озимой пшеницы (таблица 7.7).

Таблица 7.7 – Динамика урожайности озимой пшеницы в организации

Год	Урожайность озимой пшеницы, ц/га	Базисные показатели динамики			Значение 1% прироста, ц/га
		абсолютный прирост, ц	темп роста, %	темп прироста, %	
2006	55,1	-		-	-
2007		- 2,8			
2008			110,3		
2009					
2010				17,1	0,633
2011			121,1		
2012		13,5			
2013					
2014				20,4	0,691

Задача 7.9. До 2011 г. в состав производственного объединения входило 20 организаций. В 2011 г. в него влилось еще 4 организации, и оно стало объединять 24 организации. Провести смыкание ряда динамики, используя данные таблицы 7.8.

Таблица 7.8 - Динамика объема реализации продукции объединения

Объем реализации, млн руб.	2009г.	2010г.	2011г.	2012г.	2013г.	2014г.
По 20 организациям	491,6	501,1	510,2	-	-	-
По 24 организациям	-	-	580,5	610,0	612,9	615,5

Задача 7.10. По данным таблицы 7.9 провести смыкание динамического ряда.

Таблица 7.9 - Динамика валового сбора зерна в фермерских хозяйствах района

Валовой сбор зерна, тыс. ц	2008г.	2009г.	2010г.	2011г.	2012г.	2013г.	2014 г.
В старых границах	44,5	45,0	48,2	-	-	-	-
В новых границах	-	-	60,0	63,6	61,1	64,2	65,6

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?

Рекомендуемая литература

1. Бондаренко, П. С. Практикум по эконометрике: учеб.-практ. пособие для бакалавров / П. С. Бондаренко, И. А. Кацко, В. И. Перцухов, А. Е. Сенникова, А. Е. Жминько, Т. В. Соловьева, Е. Д. Стеганцова, Т. Ю. Чернобыльская; под ред. П. С. Бондаренко. – Краснодар: Кубанский ГАУ, 2013. – 164 с., ил. (Серия: Вероятность, статистика и прикладные исследования в аграрном университете).

2. Бондаренко, П. С. Теория вероятностей и математическая статистика: учеб. пособие для бакалавров / П. С. Бондаренко, Г. В. Горелова, И. А. Кацко. – Краснодар: Кубанский ГАУ, 2013.– 340 с., ил. (Серия: Вероятность, статистика и прикладные исследования в аграрном университете).

3. Гмурман, В. Е. Руководство к решению задач по теории вероятностей и математической статистике: учеб. пособие. – 11-е изд., перераб. и доп. / В. Е. Гмурман. – М.: Юрайт, 2011. – 405 с.

4. Гмурман, В. Е. Теория вероятностей и математическая статистика: учеб. пособие для студентов вузов (бакалавров) / В. Е. Гмурман. – 12-е изд. – М.: Юрайт, 2012. – 480 с.

5. Доспехов, Б. А. Методика полевого опыта (с основами статистической обработки результатов исследований). – 5-е изд., перераб. и доп. / Б. А. Доспехов. – М.: Агропромиздат, 1985. – 351 с.

6. Ляховецкий, А. М. Статистика: учеб. пособие для бакалавров / А. М. Ляховецкий, Е. В. Кремянская, Н. В. Климова; под ред. В. И. Нечаева. – Краснодар: Кубанский ГАУ, 2013. – 359 с., ил. (Серия: Вероятность, статистика и прикладные исследования в аграрном университете).

Приложения

Приложение А

Посевные площади сельскохозяйственных культур в организациях Краснодарского края, га

№ варианта	2013 г.			2014 г.			План на 2014 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 Б

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

Продолжение приложения Б

№ организации	Качество почв, балл.	Продолжительность уборки, дней	Внесено минеральных удобрений на 1 га, кг д. в.	Урожайность с 1 га, ц
32	55	16	102	32,6
33	58	16	108	39,5
34	75	11	146	56,5
35	60	10	188	42,0
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

Окончание приложения Б

№ организации	Качество почв, балл.	Продолжительность уборки, дней	Внесено минеральных удобрений на 1 га, кг д. в.	Урожайность с 1 га, ц
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 В

Критические точки t - распределения 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	63,66	318,31	636,62
2	2,92	4,30	6,97	9,93	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	1,70	2,05	2,47	2,77	3,42	3,69
28	1,70	2,05	2,47	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 Г

Значения функций $\varphi(x)=\frac{1}{\sqrt{2\pi}}\cdot e^{-\frac{1}{2}x^2}$ **и** $\Phi(x)=\frac{1}{\sqrt{2\pi}}\int_0^x e^{-\frac{1}{2}t^2}dt.$

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	3652	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
39	3697	1517	79	2920	2852	19	1965	3830

Окончание приложения Г

x	$\varphi(x)$	$\Phi(x)$	x	$\varphi(x)$	$\Phi(x)$	x	$\varphi(x)$	$\Phi(x)$
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	3883	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
63	1057	4484	26	0310	4881			
64	1040	4495	28	0297	4887	3,90	00020	49995
65	1023	4505	30	0283	4893	4,00	0,0001338	499968
66	1006	4515	32	0270	4898			
67	0989	4525	34	0258	4904	4,50	0000160	499997
68	0973	4535	36	0246	4909	5,00	0000015	49999997
69	0957	4545	38	0235	4913			

Приложение Д

Критические точки χ^2 - распределения Пирсона

ν \ α	0,20	0,10	0,05	0,02	0,01	0,001
1	1,642	2,706	3,841	5,412	6,635	10,827
2	3,219	4,605	5,991	7,824	9,210	13,815
3	4,642	6,251	7,815	9,837	11,345	16,266
4	5,989	7,779	9,488	11,668	13,277	18,467
5	7,289	9,236	11,070	13,388	15,086	20,515
6	8,558	10,645	12,592	15,033	16,812	22,457
7	9,803	12,017	14,067	16,622	18,475	24,322
8	11,030	13,362	15,507	18,168	20,090	26,125
9	12,242	14,684	16,919	19.679	21,666	27,877
10	13,442	15,987	18,307	21,161	23,209	29,588
11	14,631	17,275	19,675	22,618	24,725	31,264
12	15,812	18,549	21,026	24,054	26,217	32,909
13	16,985	19,812	22,362	25,472	27,688	34,528
14	18,151	21,064	23,685	26,783	29,141	36,123
15	19,311	22,307	24,996	28,259	30,578	37,697
16	20,465	23,542	26,296	29,633	32,000	39,252
17	21,615	24,769	27,587	30,995	33,409	40,790
18	22,760	25,989	28,869	32,346	34,805	42,312
19	23,900	27,204	30,144	33,687	36,191	43,820
20	25,038	28,412	31,410	35,020	37,566	45.315
21	26,171	29,615	32,671	36,343	38,932	46,797
22	27,301	30,813	33,924	37,659	40,289	48,268
23	28,429	32,007	35,172	38,968	41,638	49,728
24	29,553	33,196	36,415	40,270	42,980	51,179
25	30,675	34,382	37,652	41,566	44,314	52,620
26	31,795	35,563	38,885	42,856	45,642	54,052
27	32,912	36,741	40,113	44,140	46,963	55,476
28	34,027	37,916	41,337	45,419	48,278	56,893
29	35,139	39,087	42,557	46,693	49,588	58,302
30	36,250	40,256	43,773	47,962	50,892	59,703

Приложение Е

Критические точки распределения F Фишера – Снедекора
(ν_1 – число степеней свободы большей дисперсии, ν_2 – число степеней свободы меньшей дисперсии), уровень значимости α=0,05

ν_2 \ ν_1	1	2	3	4	5	6	7	8	9	10	12	16	20	24	30	50	100	∞
1	161	200	216	225	230	234	237	239	241	242	244	246	248	249	250	252	253	254
2	18,51	19,00	19,16	19,25	19,30 •	19,33	19,36	19,37	19,38	19,39	19,41	19,43	19,44	19,45	19,46	19,47	19,49	19,50
3	10,13	9,55	9,28	9,12	9,01	8,94	8,88	8,84	8,81	8,78	8,74	8,69	8,66	8,64	8,62	8,58	8,56	8,53
4	7,71	6,94	6,59	6,39	6,26	6,16	6,09	6,04	6,00	5,96	5,91	5,84	5,80	5,77	5,74	5,70	5,66	5,63
5	6,61	5,79	5,41	5,19	5,05	4,95	4,88	4,82	4,78	4,74	4,68	4,60	4,56	4,53	4,50	4,44	4,40	4,36
6	5,99	5,14	4,76	4,53	4,39	4,28	4,21	4,15	4,10	4,06	4,00	3,92	3,87	3,84	3,81	3,75	3,71	3,67
7	5,59	4,74	4,35	4,12	3,97	3,87	3,79	3,73	3,68	3,63	3,57	3,49	3,44	3,41	3,38	3,32	3,28	3,23
8	5,32	4,46	4,07	3,84	3,69	3,58	3,50	3,44	3,39	3,34	3,28	3,20	3,15	3,12	3,08	3,03	2,98	2,93
9	5,12	4,26	3,86	3,63	3,48	3,37	3,29	3,23	3,18	3,13	3,07	2,98	2,93	2,90	2,86	2,80	2,76	2,71
10	4,96	4,10	3,71	3,48	3,33	3,22	3,14	3,07	3,02	2,97	2,91	2,82	2,77	2,74	2,70	2,64	2,59	2,54
11	4,84	3,98	3,59	3,36	3,20	3,09	3,01	2,95	2,90	2,86	2,79	2,70	2,65	2,61	2,57	2,50	2,45	2,40
12	4,75	3,88	3,49	3,26	3,11	3,00	2,92	2,85	2,80	2,76	2,69	2,60	2,54	2,50	2,46	2,40	2,35	2,30
13	4,67	3,80	3,41	3,18	3,02	2,92	2,84	2,77	2,72	2,67	2,60	2,51	2,46	2,42	2,38	2,32	2,26	2,21
14	4,60	3,74	3,34	3,11	2,96	2,85	2,77	2,70	2,65	2,60	2,53	2,44	2,39	2,35	2,31	2,24	2,19	2,13
15	4,54	3,68	3,29	3,06	2,90	2,79	2,70	2,64	2,59	2,55	2,48	2,39	2,33	2,29	2,25	2,18	2,12	2,67
16	4,49	3,63	3,24	3,01	2,85	2,74	2,66	2,59	2,54	2,49	2,42	2,33	2,28	2,24	2,20	2,13	2,07	2,01
17	4,45	3,59	3,20	2,96	2,81	2,70	2,62	2,55	2,50	2,45	2,38	2,29	2,23	2,19	2,15	2,08	2,02	1,96
18	4,41	3,55	3,16	2,93	2,77	2,66	2,58	2,51	2,46	2,41	2,34	2,25	2,19	2,15	2,11	2,04	1,98	1,92
19	4,38	3,52	3,13	2,90	2,74	2,63	2,55	2,48	2,43	2,38	2,31	2,21	2,15	2.11	2,07	2,00	1,94	1,88
20	4,35	3,49	3,10	2,87	2,71	2,60	2,52	2,45	2,40	2,35	2,28	2,18	2,12	2,08	2,04	1,96	1,90	1,84
22	4,30	3,44	3,05	2,82	2,66	2,55	2,47	2,40	2,35	2,30	2,23	2,13	2,07	2,03	1,98	1,91	1,84	1,78
24	4,26	3,40	3,01	2,78	2,62	2,51	2,43	2,36	2,30	2,26	2,18	2,09	2,02	1,98	1,94	1,86	1,80	1,73
26	4,22	3,37	2,98	2,74	2,59	2,47	2,39	2,32	2,27	2,22	2,15	2,05	1,99	1,95	1,90	1,82	1,76	1,69
28	4,20	3,34	2,95	2,71	2,56	2,44	2,36	2,29	2,24	2,19	2,12	2,02	1,96	1,91	1,87	1,78	1,72	1,65
32	4,15	3,30	2,90	2,67	2,51	2,40	2,32	2,25	2,19	2,14	2,07	1,97	1,91	1,86	1,82	1,74	1,67	1,59
36	4,11	3,26	2,86	2.63	2,48	2,36	2,28	2,21	2,15	2,10	2,03	1,93	1,87	1,82	1,78	1,69	1,62	1,55
40	4,08	3,23	2,84	2,61	2,45	2,34	2,25	2,18	2,12	2,07	2,00	1,90	1,84	1,79	1,74	1,66	1,59	1,51
60	4,00	3,15	2.76	2,52	2,37	2,25	2,17	2,10	2,04	1,99	1,92	1,81	1,75	1,70	1,65	1,56	1,48	1,39
100	3,94	3,09	2,70	2,46	2,30	2,19	2,10	2,03	1,97	1,92	1,85	1,75	1,68	1,63	1,57	1,48	1,39	1,28
200	3,89	3,04	2,65	2,41	2.26	2,14	2,05	1,98	1,92	1,87	1,80	1,69	1,62	1,57	1,52	1,42	1,32	1,19
∞	3,84	2,99	2,60	2,37	2,21	2,09	2,01	1,94	1,88	1,83	1,75	1,64	1,57	1,52	1,46	1,35	1,24	1,00

Приложение Ж

Урожайность озимой пшеницы по сортам, ц/га

Год	Сорт							
	I	II	III	IV	V	VI	VII	VIII
2008	53	46	50	44	55	39	58	49
2009	43	48	41	46	49	40	50	44
2010	45	46	43	43	48	42	49	43
2011	56	51	52	50	59	46	53	52
2012	58	52	56	51	61	44	55	55
2013	55	48	43	47	60	38	51	54
2014	59	52	61	49	64	41	56	59

Приложение И

Урожайность сельскохозяйственных культур в хозяйствах всех категорий Краснодарского края (центнеров с одного гектара посевной площади)

Вариант	Культура	Год									
		2004	2005	2006	2007	2008	2009	2010	2011	2012	2013
1	Пшеница озимая	44,2	48,2	42,7	45,1	55,3	45,7	49,7	55,1	39,8	50,1
2	Пшеница яровая	22,5	27,5	23,8	22,5	34,2	26,8	30,2	31,9	26,3	27,7
3	Ячмень озимый	46,5	42,0	43,5	47,5	51,4	46,9	49,2	53,8	37,1	53,1
4	Ячмень яровой	23,4	24,2	25,8	18,7	36,9	27,0	25,0	33,8	27,4	30,5
5	Рожь озимая	23,3	29,5	22,2	26,4	44,7	32,1	38,5	44,7	31,2	48,5
6	Кукуруза на зерно	48,1	44,1	40,2	21,8	49,5	33,8	33,8	47,7	41,9	53,0
7	Овес	26,3	26,6	25,2	21,0	33,7	23,3	24,7	30,5	25,2	26,8
8	Просо	9,9	13,4	14,4	12,5	21,2	6,4	16,6	24,0	16,8	18,8
9	Гречиха	4,3	6,4	6,6	4,5	10,2	4,9	9,8	7,1	7,0	4,7
10	Рис	39,7	44,4	47,1	48,3	50,7	60,3	62,1	61,0	64,3	57,6
11	Зернобобовые – всего	23,3	19,7	22,9	14,4	32,9	23,1	23,7	27,9	41,1	21,0
12	Горох	23,6	19,8	23,3	14,6	33,9	23,6	24,0	28,1	21,9	21,8
13	Сахарная свекла	396	328	359,6	262,4	438,6	381,1	361,2	438,1	423,0	517,1
14	Масличные культуры - всего	18,0	19,3	18,7	16,4	22,1	20,2	19,4	21,9	21,4	24,1
15	Подсолнечник	18,2	20,9	20,7	18,9	23,3	20,9	20,8	23,3	23,2	25,7
16	Соя	18,0	12,6	12,8	9,1	15,8	18,0	15,1	18,5	18,0	20,4
17	Рапс озимый	16,6	14,7	16,9	16,2	19,1	17,8	19,5	20,1	16,1	23,1
18	Картофель	92,8	88,4	89,6	78,9	96,8	93,9	89,0	96,4	98,6	100,0
19	Овощи	74,3	85,4	93,4	79,7	102,8	106,3	98,8	111,7	106,2	106,8
20	Бахчи продовольственные	43,2	63,2	66,2	49,9	68,7	86,5	66,6	63,7	84,8	97,6

Окончание приложения И

Вариант	Культура	Год									
		2004	2005	2006	2007	2008	2009	2010	2011	2012	2013
21	Силосные культуры (без кукурузы)	73,2	64,0	80,9	65,5	96,6	91,6	78,5	66,9	81,5	122,7
22	Однолетние травы на сено	19,2	27,6	34,9	21,9	30,6	25,0	23,8	30,4	25,9	24,9
23	Однолетние травы на зеленый корм	95,5	86,1	101,1	79,9	113,2	87,3	93,1	91,1	69,4	88,3
24	Многолетние травы на сено	26,4	23,6	26,9	16,5	30,7	32,6	46,2	43,7	36,4	33,3
25	Плоды и ягоды	45,7	57,8	49,1	55,5	70,4	69,9	60,5	74,4	87,3	107,1
26	Виноград	56,1	68,6	48,6	80,7	74,0	84,5	79,5	113,4	75,6	103,6
27	Чайный лист	8,7	8,9	8,1	4,5	5,7	4,7	2,7	2,2	1,0	0,9

Оглавление

Учебное издание

Курнякова Татьяна Александровна
Кацко Игорь Александрович
Жминько Альбина Евгеньевна
Кремянская Елена Владимировна
Сенникова Алина Евгеньевна
Ворокова Нодира Хасановна

МАТЕМАТИЧЕСКАЯ СТАТИСТИКА

Практикум

В авторской редакции

Подписано в печать _______. Формат бумаги 60×84 1/16
Усл. печ. л. – 6,0. Уч.-изд. л. – 4,7.
Тираж ____ Заказ №______

Типография Кубанского государственного
аграрного университета.
350044, г. Краснодар, ул. Калинина, 13