

МИНИСТЕРСТВО СЕЛЬСКОГО ХОЗЯЙСТВА
РОССИЙСКОЙ ФЕДЕРАЦИИ

Федеральное государственное бюджетное образовательное
учреждение высшего образования
«КУБАНСКИЙ ГОСУДАРСТВЕННЫЙ АГРАРНЫЙ
УНИВЕРСИТЕТ ИМЕНИ И. Т. ТРУБИЛИНА»

СИСТЕМНАЯ ЭКОЛОГИЯ

Учебное пособие

Под общей редакцией И. С. Белюченко

Краснодар
КубГАУ
2017

УДК 504:311.1 (075.8)

ББК 20.1

С40

Рецензенты:

А. В. Смагин – доктор биологических наук, профессор
(Московский государственный университет
имени М. В. Ломоносова);

Э. А. Сиротюк – доктор биологических наук, профессор
(Майкопский государственный технологический университет)

Коллектив авторов:

Л. Б. Попок, О. А. Мельник, Ю. Ю. Никифоренко,
Д. А. Антоненко, А. А. Теучеж

С40 **Системная экология** : учеб. пособие / Л. Б. Попок [и др.];
под общ. ред. И. С. Белюченко. – Краснодар : КубГАУ, 2017. –
164 с.

ISBN 978-5-00097-282-3

В учебном пособии в доступной форме рассматриваются вопросы организации сбора и обработки данных экологических исследований. В основу положены технологические приемы обработки экологической информации с использованием инструмента статистического анализа в электронных таблицах MS EXCEL как наиболее доступного и не требующего значительных временных затрат при освоении дисциплины «Системная экология».

Пособие предназначено для студентов по направлению подготовки 05.03.06 «Экология и природопользование» факультета экологии и других биологических специальностей, а также аспирантов и преподавателей, занимающихся обработкой экологической информации.

УДК 504: 311.1 (075.8)

ББК 20.1

© ФГБОУ ВО «Кубанский
государственный аграрный
университет имени
И. Т. Трубилина», 2017

ISBN 978-5-00097-282-3

ВВЕДЕНИЕ

В процессе изучения дисциплины «Системная экология» студенты приобретают навыки обобщения полученных ими за годы обучения в вузе знаний по экологии на основе системного подхода к теоретическим вопросам общей экологии и применения системного анализа к решению экологических проблем, дают с единых позиций характеристику всех основных компонентов экосистем, описывая их взаимосвязи друг с другом и с внешней средой.

Целью освоения дисциплины «Системная экология» является выработка у студентов системного мышления, умения составлять физические и математические модели, описывающие функционирование экологических систем, и использовать методы системного анализа при исследовании этих экосистем. Для достижения поставленной цели решаются следующие задачи: поиск информации по полученному заданию, сбор и анализ данных, необходимых для проведения конкретных экологических исследований; построение стандартных теоретических моделей исследуемых процессов, анализ, оценка, интерпретация полученных результатов и обоснование выводов.

Основным методом исследований в «Системной экологии» является системный анализ, на основании которого разрабатываются способы исследования разнообразных сложных систем или экологических ситуаций. При системном подходе используются математический аппарат теории исследования операций, методы многомерной статистики, компьютерное моделирование. Существенной частью исследования систем является выбор способа описания происходящих в них изменений и формализация такого описания. Сложность формализации определяется сочетанием разнотипных факторов, характеризующих систему.

Разработка методов системного анализа как научной дисциплины ведется по нескольким направлениям. Одним из

важнейших является создание принципов построения и использования моделей, имитирующих протекание реальных процессов, способов их объединения в системы и такого представления в ЭВМ, которое обеспечивало бы простоту их использования без потери адекватности. Другое направление связано с изучением организационных структур, и прежде всего систем, обладающих иерархической организацией. Прежде чем рассмотреть основные принципы применения системного анализа в экологии, определим ее место среди биологических наук и ее основные разделы. Для успешного решения теоретических и практических задач многие разделы биологии и экологии следует рассматривать с системных позиций и применять соответствующие методы исследований.

Настоящее учебное пособие предназначено для студентов, а также аспирантов и преподавателей биологических специальностей для проведения ими статистической обработки данных, полученных в ходе научных исследований и экспериментов, с помощью программы MS EXCEL. В пособии уделяется особое внимание решению задач практической направленности с тем, чтобы на конкретных примерах, близких реальной производственной или научно-исследовательской профессиональной деятельности, научить студентов собирать, подготавливать и проводить статистическую обработку полученных данных. Рассматриваются возможности электронных таблиц MS EXCEL, занимающих устойчиво лидирующее положение среди программ статистической обработки данных.

РАЗДЕЛ 1. СТАТИСТИЧЕСКИЙ АНАЛИЗ РЕЗУЛЬТАТОВ ИССЛЕДОВАНИЙ

Тема 1. Шкалы измерения признаков

При проведении экологических исследований необходимо знать правила сбора описательного материала и его статистической обработки.

Полевой этап экологического исследования и этап камеральной обработки позволяет получить данные по количественному и видовому составу животных или растений, химическому или физическому анализу проб почвы, воды, воздуха, определить биометрические характеристики особей (вес, длина, высота, количество лепестков) или их качественные характеристики (цвет шкурки, состояние хвои или листьев) и т. д.

Чтобы систематизировать собранный материал и сделать предварительные выводы об изучаемой системе, необходимо привлечь знания из математической статистики.

Описательная статистика позволяет обобщать первичные результаты, полученные при наблюдении или в эксперименте. Процедуры здесь сводятся к группировке данных, построению распределения их частот, выявлению центральных тенденций распределения (например, средней арифметической), к оценке разброса данных по отношению к найденной центральной тенденции.

Математическая статистика оперирует такими понятиями как: генеральная совокупность, выборка, выборочный метод исследования.

Генеральная совокупность – множество всех мыслимых значений наблюдений (объектов) однородных относительно некоторого признака. Число всех наблюдений, составляющих генеральную совокупность, называется ее объемом (N).

Выборка – это совокупность случайно отобранных наблюдений. Объем выборки – это количество исследуемых элементов (n). При $n < 30$ выборка малая, при $30 < n < 100$ – средняя, при $n > 100$ – большая.

Способ, когда изучается лишь небольшая группа объектов и по его результатам делаются выводы обо всей генеральной совокупности, называется **выборочным методом исследования**.

1.1 Типы шкал измерения признаков

Статистические данные – это сведения о том, какие значения приняли интересующие исследователя признаки в статистической совокупности.

Признаки бывают **количественные** и **качественные**.

Качественные признаки – это признаки, которые характеризуются некоторым свойством или состоянием элементов совокупности. Это может быть цвет лепестков, окраска шкурки животного, вид или род особей, качество почвы и т. д.

Количественные признаки – это те признаки, которые выражаются числами.

Чтобы данные были наглядны, применяют группировку данных, т. е. отображают их в той или иной шкале. При упорядочении собранных данных легко обработать их математически и вывести статистические показатели, которые будут характеризовать изучаемую совокупность.

Для **качественных** признаков применяются **номинальные** и **порядковые (ранговые) шкалы**. Для **количественных** признаков – **интервальные шкалы**.

1. Номинальные шкалы. Состояние номинального признака называется **модальностью**. Например, признак «окраска шкурки норки» имеет несколько модальностей: белая, бежевая, коричневая, черная, голубая.

Для характеристики номинальных данных часто используется пропорция или процентное отношение. Например, частота встречаемости определенного вида выражается в процентах.

Чтобы рассчитать частоту встречаемости определенного вида, необходимо его численность поделить на общую чис-

ленность особей всех видов в пробе, т. е. найти долю по численности каждого вида в общей совокупности.

Арифметические операции над величинами, измеренными в номинальной шкале, лишены смысла.

Графически номинальные данные представляются в виде столбиковой диаграммы.

В таблице 1 приводится пример расчета частоты встречаемости видов в пробе, а на рисунке 1 – графическое изображение распределения этих видов.

Таблица 1 – Частота встречаемости видов в пробе

№ п/п	Вид или род	Количество, экз	Частота встречаемости	Частота встречаемости, %
1	<i>Daphnia longispina</i>	235	0,338	33,81
2	<i>Daphnia cucullata</i>	114	0,164	16,40
3	<i>Daphnia pulex</i>	197	0,283	28,35
4	<i>Simacephalus</i>	84	0,121	12,09
5	<i>Ceriodaphnia</i>	53	0,076	7,63
6	<i>Moina macroscopa</i>	12	0,017	1,73
Сумма		695	1	100

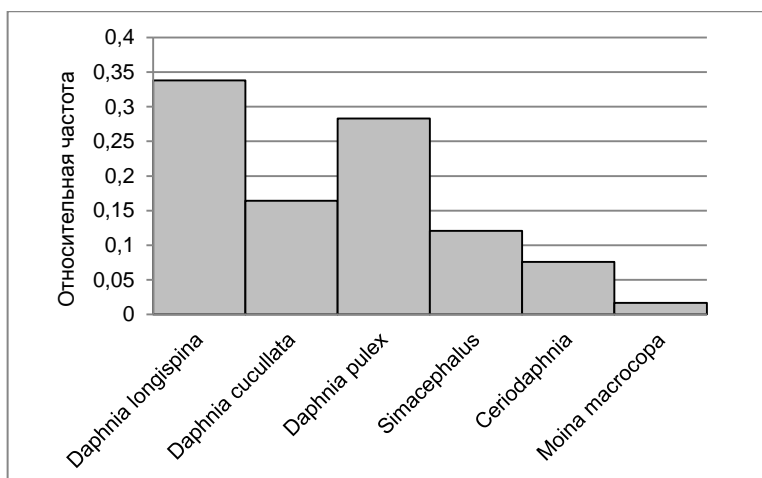


Рисунок 1 – График распределения видов в пробе

2. Ранговые шкалы. В отличие от номинальных шкал порядковые шкалы соответствуют таким качественным переменным, для которых характерна упорядоченность, направленность или степень важности. Например, выраженные в баллах устойчивость к загрязняющим веществам, категория жизненного состояния деревьев, качество воды и т. д.

Состояние порядкового признака называется **рангом**.

Рассчитываются ранги исходя из следующего определения.

Рангом R_i наблюдения X_i среди величин X_1, \dots, X_n называют тот порядковый номер, который получит значение X_i при расстановке чисел X_1, \dots, X_n в порядке возрастания или убывания, если значения чисел X_i не совпадают по величине. В противном случае каждому из совпадающих чисел присваивается ранг, равный среднему арифметическому их порядковых номеров.

Сумма рангов должна быть равна сумме порядковых номеров. Для ранговой шкалы в качестве показателя средней тенденции используют **медиану** (значение признака, который находится по середине ряда).

Пример 1. Исходный ряд оценок признака из 8 объектов в порядковой шкале имеет вид: 4, 6, 10, 3, 10, 7, 7, 8. Ранжировать ряд.

Решение. Расположим объекты по возрастанию и присвоим сначала порядковые номера, а затем ранги:

По возрастанию	3	4	6	7	7	8	10	10	Сумма
Порядковый номер	1	2	3	4	5	6	7	8	36
Ранг	1	2	3	4,5	4,5	6	7,5	7,5	36

При совместном рассмотрении данных, измеренных в разных шкалах, их необходимо перевести в одну шкалу. Чаще переводят количественные данные в «качественную шкалу».

Пример 2. Урожайность с.-х. культуры и степень пораженности культуры мучнистой росой на 10 участках перевести в ранговую шкалу.

Номер участка	1	2	3	4	5	6	7	8	9	10
Урожайность (ц/га)	10,3	12,0	8,9	14,7	8,7	13,4	6,7	5,7	18,6	11,2
Пораженность (балл)	3	5	1	0	2	1	5	4	1	3

Решение. Количественным данным можно сразу присвоить ранг по величине признака, а качественным данным сначала присвоим порядковый номер, а затем, применяя выше изложенное правило, присвоим ранги. При этом надо не забывать, что признаки с одинаковыми значениями имеют одинаковые ранги, равные среднему арифметическому их порядковых номеров.

Номер участка	1	2	3	4	5	6	7	8	9	10
Урожайность (ц/га)	10,3	12,0	8,9	14,7	8,7	13,4	6,7	5,7	18,6	11,2
Ранг R_1	5	7	4	9	3	8	2	1	10	6
Пораженность (балл)	3	5	1	0	2	1	5	4	1	3
Порядковый номер	6	9	2	1	5	3	10	8	4	7
Ранг R_2	6,5	9,5	3	1	5	3	9,5	8	3	6,5

В экологических исследованиях при учете организмов часто используются специфические шкалы, разработанные для выделения классов доминирования по численности или по биомассе. Например, шкала Любарского, основанная на относительном обилии видов, шкала Браун-Бланке, основанная на проективном покрытии, логарифмические шкалы оценки относительного обилия видов. Они представляют собой таблицы, в которых обилие видов разграничивается на несколько классов (5–6), каждому из которых присваивается свой балл. А поскольку обилие вида может представлять очень большую

величину, для удобства вводятся логарифмические шкалы, в которых вместо обилий фигурируют логарифмы этих чисел.

3. Интервальные шкалы. Шкала, в которой можно отразить, насколько по степени выраженности заданного свойства один из объектов отличается от другого, называется **интервальной**. Для того чтобы задать интервальную шкалу, надо определить начальную точку и единицу измерения. Далее каждому объекту ставят в соответствие число, показывающее, на сколько единиц измерения этот объект отличается от объекта, принятого за начальную точку (например, температура в градусах Цельсия или масса в г и т.п.). Количественные шкалы допускают арифметические преобразования.

1.2 Группировка данных по количественной вариации

Вариационным рядом выборки называется способ ее записи, при котором элементы располагаются в порядке возрастания или убывания с соответствующими им весами. В качестве весов выступают частоты или относительные частоты.

Элементы выборки x_i называются **вариантами**.

Число n_i , показывающее, сколько раз элемент x_i встречается в данной совокупности, называется **частотой**.

Сумма всех частот равна объему выборки (n):

$$n_1 + n_2 + \dots + n_i + \dots + n_k = n .$$

Относительной частотой варианты называется отношение соответствующей ей частоты к объему выборки: $w_i = n_i/n$.

Вариационные ряды бывают дискретные и интервальные.

Дискретные ряды строятся в том случае, когда различия между вариантами выражаются целыми числами, между которыми нет, и не может быть переходов. Например, количество щенков у лисиц, количество лепестков на цветке, число лучей в плавнике у рыб, число позвонков у птиц и т. д.

Интервальные вариационные ряды строятся, когда значения признаков отличаются друг от друга на сколь угодно малую величину и заранее оговаривается степень точности измерения признака. Обычно признаки выражены дробными числами и выражают измерения физических величин (кг, м, см, мм).

В том случае, если признак сильно варьирует, интервальные ряды могут быть построены и для вариант, выраженных целыми числами.

Для характеристики вариационных рядов, наряду с частотами и относительными частотами, применяются накопительные частоты (S_i).

Накопительная частота – это сумма частот всех классов, не превышающих данный. Рассчитывается по формуле:

$$S_i = n_i + n_{i-1} + n_{i-2} + \dots + n_1. \quad (1)$$

1.3 Алгоритм построения интервального вариационного ряда

1. Найти число интервалов или классов, на которое разбивается выборка значений признака. Строгих правил при выборе желаемого количества интервалов нет. Его можно выбрать произвольно в зависимости от объема выборки (таблица 2) или по формуле $K \sim 1 + 3,322 \cdot \lg(n)$. Значение K округляется до ближайшего целого числа.

Таблица 2 – Определение числа классов по объему выборки

Объем выборки (от – до)	Число классов
25–40	5–6
40–60	6–8
60–100	7–10
100–200	8–12
> 200	10–15

2. Определить длину интервала $i = (x_{max} - x_{min}) / K$ и округлить для удобства вычислений.

3. Наметить границы классов. Для этого определить нижнюю границу первого классового интервала $l = x_{min} - i/2$. Прибавив к этой величине величину интервала, получить верхнюю границу 1-го класса. Так поступать до тех пор, пока не получится класс, в который попадает максимальная варианта. Для того чтобы разграничить классы, верхние границы уменьшают на величину, равную принятой точности измерения признака.

4. Подсчитать частоту попадания признака в каждый i -ый интервал.

5. Полученные данные оформить в таблицу.

Пример 3. Построить вариационный ряд для выборки объемом $n = 25$ при количественной дискретной вариации.

39	41	40	42	41
40	42	44	40	43
42	41	43	39	42
41	42	39	41	39
43	41	39	43	42

Решение. Расположим варианты в порядке возрастания от минимального $x_{min} = 39$ до максимального $x_{max} = 44$. Определим, сколько раз каждое значение встретилось в полученном ряду, результаты занесем в таблицу:

x_i	39	40	41	42	43	44
n_i	5	3	6	6	4	1

Пример 4. Распределить данные по весу 25 кроликов (кг) в вариационный ряд.

3,2	4,5	5,2	5,6	6,0
3,8	4,7	5,2	5,7	6,3
4,1	4,9	5,3	5,8	6,4
4,3	5,0	5,3	5,8	6,7
4,3	5,1	5,4	5,9	7,3

Решение. Здесь нет естественных классов, как в случае с дискретной вариацией.

Размах вариации составляет $x_{\max} - x_{\min} = 7,3 - 3,2 = 4,1$.

Весь этот интервал необходимо разбить на определенное число классов, при выборе которых необходимо одновременно принять во внимание размеры классового промежутка. Лучше, если i будет удобным числом, например, 0,5; 1; 5; 10, а не 0,45; 1,1; 6; 11. Количество классов может быть несколько меньше или больше, чем рассчитанное по формуле $K \sim 1 + 3,32 \cdot \lg(n)$. Так, например, воспользовавшись этой формулой или таблицей 2, определим количество классов в нашем примере, равное 6. В таком случае величина классового интервала будет равна $4,1 / 6 \approx 0,7$. Это не удобное число. Но если принять желаемое количество классов 8 или 9, то размеры классов будут 0,5 кг.

В таком случае можно наметить следующие классы:

3,0–3,5–4,0–4,5–5,0–5,5–6,0–6,5–7,0–7,5.

Ряд получился несколько растянутым. Его можно сделать более сжатым, если принять $i = 1$.

Разграничим классы, уменьшив их верхние границы на величину, равную точности измерения признака: 0,1.

Подсчитав количество вариантов, попавшее в каждый класс, получим следующие вариационные ряды:

Классы	Частоты	Классы	Частоты
3,0–3,4	1	3,0–3,9	2
3,5–3,9	1	4,0–4,9	6
4,0–4,4	3	5,0–5,9	12
4,5–4,9	3	6,0–6,9	4
5,0–5,4	7	7,0–7,9	1
5,5–5,9	5		
6,0–6,4	3		
6,5–6,9	1		
7,0–7,5	1		
$I = 0,5$	$n = 25$	$i = 1,0$	$n = 25$

1.4 Графическое изображение вариационных рядов

Графическое представление вариационного ряда позволяет представить в наглядной форме закономерности распределения изучаемой величины и служит для анализа распределения вариантов в совокупности. Обычно скошенность и многовершинность распределения говорит о недостаточном количестве вариантов или о неоднородности выборки – смешение двух качественно различных совокупностей. Графически вариационные ряды можно представить в виде полигона, гистограммы и кумуляты.

Принято полигон строить для безынтервального вариационного ряда, а гистограмму для интервального.

Полигоном частот называют ломаную линию, соединяющую точки, соответствующие срединным значениям классов (x_i) и частотам этих разрядов (n_i). Срединные значения откладываются по оси абсцисс (X), а частоты – по оси ординат (Y). Слева и справа полигон доводят до нулевых классов.

Пример 5. Распределить данные в безынтервальный вариационный ряд и построить полигон частот.

25	26	27	27	27
28	27	26	28	29
26	29	25	26	27
29	27	28	30	25
27	30	27	26	28

Решение. Расположим варианты в порядке возрастания от минимального $x_{\min} = 25$ до максимального $x_{\max} = 30$. Определим, сколько раз каждое значение встретилось в полученном ряду, результаты занесем в таблицу:

x_i	25	26	27	28	29	30
n_i	3	5	8	4	3	2

Для построения полигона по оси абсцисс (X) откладываем варианты (x_i), а по оси ординат (Y) соответствующие им частоты (n_i). Слева и справа полигон доводим до нуля (рисунок 2).

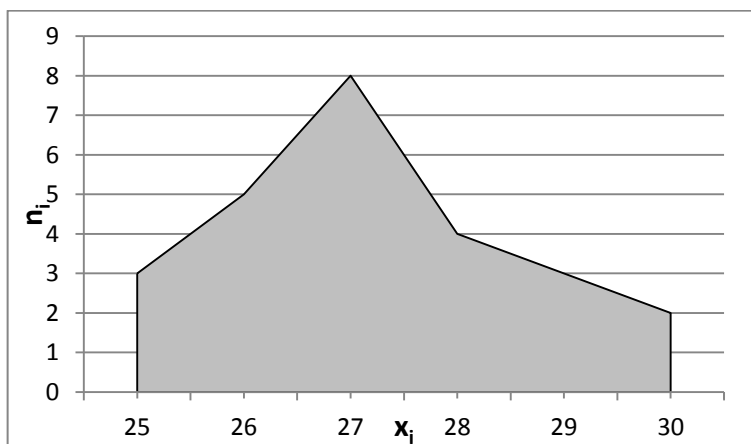


Рисунок 2 – Полигон распределения данных

Гистограмма частот – это ступенчатая фигура, состоящая из смежных прямоугольников, построенных на одной прямой, основания которых равны ширине классов, а высота – частоте n_i или относительной частоте w_i .

Кумулята – это изображение распределения, ординаты которого пропорциональны накопительным частотам. Кумуляту применяют, если необходимо сравнить одновременно несколько эмпирических распределений, неравных по объему, особенно в тех случаях, когда частоты рядов выражены в процентах от общего числа наблюдений.

Пример 6. При проведении анализа зависимости между площадью листьев у древесных растений и загрязненностью атмосферы в одном из районов города Краснодара исследовалась выборка из 50 листьев тополя. После измерения их площадей были получены следующие данные (см^2):

43	32	29	38	35	36	42	26	36	35
38	38	38	38	34	36	44	36	40	34
22	41	40	53	40	33	32	20	27	15
21	24	23	27	25	34	17	29	28	10
25	18	40	29	27	43	26	26	31	31

Для рассчитанных площадей листьев выполнить следующие задания:

1. Определить размах выборки.
2. Построить вариационный ряд.
3. Построить гистограмму и полигон распределения.
4. Построить кумуляту распределения.

Сделать выводы о характере распределения изучаемого показателя.

Решение.

1. Размах вариации равен: $x_{\max} - x_{\min} = 53 - 10 = 43$.
2. Так как размах вариации велик, строим интервальный вариационный ряд. Если количество классов принять равное 7, то классовый интервал составит примерно 6 см^2 .

Наметим классовые интервалы:

7–13–19–25–31–37–43–49–55.

Разграничим классы, уменьшая верхнюю границу классовых интервалов на величину точности измерения признака, и подсчитаем соответствующие каждому классу частоты. Получится следующий вариационный ряд:

x_i	7–12	13–18	19–24	25–30	31–36	37–42	43–48	49–55
n_i	1	3	5	12	14	11	3	1

3. Для построения гистограммы по оси X откладываем нижние границы классов, а по оси Y – соответствующие им частоты (рисунок 3).

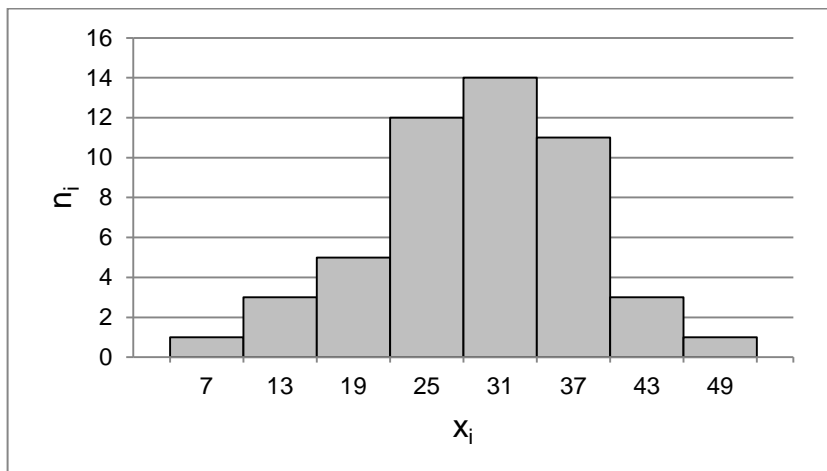


Рисунок 3 – Гистограмма распределения данных

Для построения полигона по оси абсцисс (X) откладываем срединные значения классов (x_i), а по оси ординат (Y) соответствующие им частоты (n_i). Слева и справа полигон доводим до нулевых классов (рисунок 4).

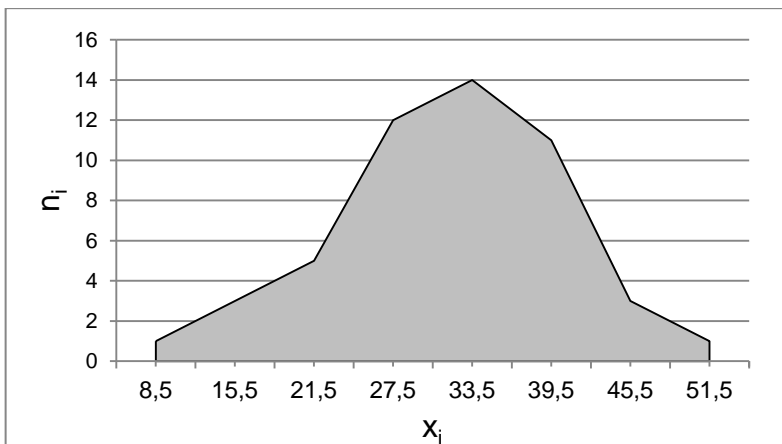


Рисунок 4 – Полигон распределения данных

4. Для построения кумуляты предварительно рассчитаем накопленные частоты:

x_i	7–12	13–18	19–24	25–30	31–36	37–42	43–48	49–55
n_i	1	3	5	12	14	11	3	1
S_i		4	9	21	35	46	49	50

По оси X откладываем верхние границы классовых интервалов, а по оси Y – накопительные частоты. Полученные точки соединяем плавной линией (рисунок 5).

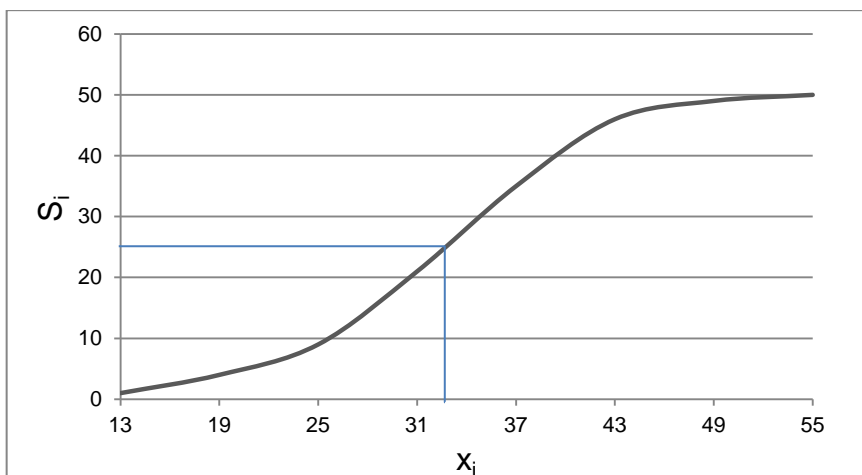


Рисунок 5 – Кумулята распределения данных

Вывод. По полученным графикам (по симметричности гистограммы) можно сделать вывод о том, что выборка представляет собой однородную совокупность и является репрезентативной, т. е. будет отражать по выбранному показателю (площадь листьев) состояние деревьев на выбранном объекте. По кумуляте можно определить, что 50 % промеренных листьев будут иметь площадь около 33 см².

Контрольные вопросы

1. Дать понятия генеральной совокупности данных и выборки данных.
2. Дать определение качественным и количественным признакам в биологических исследованиях. Привести примеры.
3. Какие шкалы используют при характеристике качественных признаков?

4. Какие шкалы используют при характеристике количественных признаков?
5. Дать понятие номинальным шкалам.
6. Дать понятие ранговым шкалам.
7. Дать понятие интервальным шкалам.
8. Что называется вариационным рядом?
9. Дать понятие дискретного и интервального вариационного ряда.
10. Перечислить основные этапы построения интервального вариационного ряда.
11. Как графически можно представить вариационные ряды?
12. Что называют полигоном частот?
13. Что называют гистограммой частот?
14. Что называют кумулятой?

Домашнее задание

1. Соберите по 25–30 листьев с определенной породы деревьев, растущих в разных экологических условиях (2–3 выборки).
2. Ознакомьтесь с методикой расчета площадей листьев. (Федорова А. И., Никольская А. Н. Практикум по экологии и охране окружающей среды. 2003, С. 123. Лабораторная работа № 34).
3. Постройте вариационные ряды для полученных выборок площадей листьев.
4. Нарисуйте графики изменчивости площадей листьев в разных экологических условиях.

Задания к самостоятельной работе

1. Исходный вариационный ряд оценок признака (например, степень повреждения деревьев в 10-балльной шкале) у 10 объектов имеет вид: 2, 8, 4, 3, 5, 9, 2, 10, 6, 6. Ранжировать ряд.

2. Постройте полигон распределение частот для следующего ряда.

x_i	2	3	5	7	9
n_i	10	15	25	33	12

3. Исходный вариационный ряд оценок признака у 9 объектов в порядковой шкале имеет вид: 3, 5, 8, 2, 2, 10, 6, 7, 4. Ранжируйте ряд.

4. Постройте гистограмму следующего распределения.

Классы (x_i)	8,6– 9,3	9,4– 10,1	10,2– 10,9	11,0– 11,7	11,8– 12,5	12,6– 13,3	13,4– 14,1	14,2– 14,9
Частоты (n_i)	2	6	15	23	25	17	7	5

5. Наблюдаемые значения полуденной температуры мая разбиты на 5 интервалов. Постройте гистограмму распределения.

x_i	0–4	5–9	10–14	15–19	20–24	25–29
n_i	4	5	8	7	4	3

6. Постройте кумюляту для следующего распределения:

x_i	1,0–1,9	2,0–2,9	3,0–3,9	4,0–4,9	5,0–5,9	6,0–7,0
n_i	0	15	23	5	6	1

7. Распределить следующие данные в вариационный ряд и построить гистограмму распределения.

3,5	2,3	2,9	3,1	3,2	3,5
3,1	2,9	4,0	2,2	3,1	2,4
3,6	2,5	4,1	4,4	3,2	3,3
3,8	3,2	3,0	2,8	2,4	2,5

8. Постройте полигон распределение частот для следующего ряда.

x_i	2	3	5	7	9	11	13	15
n_i	10	15	25	33	12	8	10	21

Тема 2. Обзор основных статистических характеристик

Статистические показатели разделяются на 2 группы: показатели, отражающие наиболее типичное в ряде, и показатели, измеряющие степень варьирования признака.

К первой группе относятся различные средние величины: среднее арифметическое, мода, медиана, среднее геометрическое, среднее гармоническое и т.д.

Ко второй группе – вариационный размах, абсолютное среднее, дисперсия (варианса), среднее квадратическое отклонение, коэффициенты вариации.

2.1 Средние величины

Простая средняя арифметическая – сумма значений всех вариантов, отнесенная к числу вариант:

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i = \frac{x_1 + x_2 + x_3 + \dots + x_n}{n}. \quad (2)$$

Взвешенная средняя арифметическая – сумма произведений значений отдельных вариантов на их частоты, отнесенная к сумме частот:

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i n_i = \frac{x_1 \cdot n_1 + x_2 \cdot n_2 + x_3 \cdot n_3 + \dots + x_i \cdot n_i}{n}. \quad (3)$$

Средняя арифметическая имеет смысл только для качественно однородной совокупности. Так, например, нельзя вычислить средний вес животного для совокупности животных разных возрастных групп или среднюю площадь для листьев разных пород деревьев.

Мода (M_o) – это наиболее часто встречаемая варианта в данной совокупности.

Например, мода распределения равна 18.

x_i	16	17	18	20
n_i	5	1	20	6

Использование моды в биологических и экологических исследованиях особенно актуально, когда данные не являются чисто количественными. Тогда мода достаточно хорошо характеризует типичное.

Для определения моды интервальных рядов применяют формулу

$$M_o = x_{\text{ниж}} + i \cdot \left(\frac{n_{M_o} - n_{M_o-1}}{2n_{M_o} - n_{M_o-1} - n_{M_o+1}} \right), \quad (4)$$

где $x_{\text{ниж}}$ – нижняя граница модального класса, т. е. класса с наибольшей частотой встречаемости;

n_{M_o} – частота модального класса;

n_{M_o-1} – частота класса, предшествующего модальному;

n_{M_o+1} – частота класса, следующего за модальным;

i – величина классового интервала.

Пример 1. Распределение проективного покрытия определенного вида в фитоценозе представлено следующей таблицей:

Покрытие	0	0,5	1	5	10	20
Частота	13	4	3	3	1	1

Здесь проективное покрытие измеряется в процентах. Такая информация не является чисто количественной.

Мода этого распределения равна 0, а среднее арифметическое 2 %.

$$(13 \cdot 0 + 0,5 \cdot 4 + 1 \cdot 3 + 5 \cdot 3 + 10 \cdot 1 + 20 \cdot 1) / (13 + 4 + 3 + 3 + 1 + 1) = (2 + 3 + 15 + 10 + 20) / 25 = 50 / 25 = 2 \%$$

Если бы случайно в выборку не попала площадка с 20 % покрытием, то среднее покрытие в этом фитоценозе было бы равно 1 %, а мода осталась бы прежней. Это говорит о том, что при качественной вариации мода является более устойчивым показателем, чем среднее арифметическое и лучше отражает типичное в ряде.

Медиана (M_e) – это значение признака, относительно которого ряд распределения делится на две равные по объему части. Например, в распределении

12	14	16	18	20	22	24	26	28
----	----	----	----	-----------	----	----	----	----

медианой будет центральная варианта, т. е. $M_e = 20$. По обе стороны от нее стоит по 4-е варианты.

Для ряда с четным числом членов медиана равна полусумме ее центральных вариантов. Например, для ряда

6	8	10	12	14	16	18	20	22	24
---	---	----	----	-----------	-----------	----	----	----	----

медианой будет полусумма его центральных членов, т.е.

$$M_e = \frac{14 + 16}{2} = 15.$$

Для вариационного ряда, сгруппированного в классы, медиану находят по формуле:

$$M_e = x_{M_e} + i \cdot \left(\frac{n/2 - \sum n_s}{n_{M_e}} \right) \quad (5)$$

где x_{M_e} – нижняя граница медианного интервала или полусумма соседних классовых вариантов, если ряд безынтер-

вальный; n – объем выборки; i – величина классового интервала; $\sum n_s$ – накопительная частота для класса, предшествующего медианному; n_{M_e} – частота медианного класса.

Пример 2. Найдем медиану для ряда покрытий определенного вида в 200 площадках по 1 м² в изучаемом фитоценозе.

№ п/п	Покрытие	Частота	Накопительная частота
1	0–4	11	11
2	5–14	35	46
3	15–24	28	74
4	25–34	32	106
5	35–44	25	131
6	45–54	27	158
7	55–64	16	174
8	65–74	19	193
9	75–85	7	200

Медиана находится между 100 и 101-м числами ряда и приходится на класс с покрытием от 25 до 35 %. Подставив необходимые данные в формулу (5), получим:

$$M_e = 25 + 10 \cdot \left(\frac{100 - 74}{32} \right) = 25 + 10 \cdot 0,8125 \approx 33,125\% .$$

Медиана, как и мода, мало зависит от крайних значений признака.

2.2 Меры вариации признака

Учет вариации того или иного признака в совокупности имеет очень большое значение в экологических исследованиях, так как всякая вариация в популяциях животных или растений отражает различия между организмами – в их наслед-

ственной природе, в тех условиях, в которых они обитают. Без оценки вариации невозможно сравнение двух совокупностей.

Дисперсия. Дисперсия – это **средний квадрат отклонений вариант данной совокупности от их средней величины.**

Дисперсия генеральной совокупности обычно обозначается σ^2 , а выборочная дисперсия обозначается S_x^2 .

В случае $n \geq 30$ дисперсию рассчитывают по формуле 6.

$$\sigma^2 = \frac{\sum_{i=1}^k (x_i - \bar{x})^2}{n}. \quad (6)$$

Исправленная выборочная дисперсия рассчитывается для простой выборки по следующим формулам:

$$S_x^2 = \frac{\sum_{i=1}^k (x_i - \bar{x})^2}{n}, \quad S_x^2 = \frac{1}{n} \left(\sum_{i=1}^k x_i^2 - \frac{(\sum_{i=1}^k x_i)^2}{n} \right).$$

А для выборки, распределенной в вариационный ряд, по формулам:

$$S_x^2 = \frac{\sum_{i=1}^k (x_i - \bar{x})^2 n_i}{n}, \quad S_x^2 = \frac{1}{n} \left(\sum_{i=1}^k n_i x_i^2 - \frac{(\sum_{i=1}^k x_i n_i)^2}{n} \right).$$

При $n < 30$ дисперсию рассчитывают по формуле 7 (квадраты отклонений делят на $(n-1)$).

$$\sigma^2 = \frac{\sum_{i=1}^k (x_i - \bar{x})^2}{n-1}. \quad (7)$$

Исправленная выборочная дисперсия рассчитывается для простой выборки по следующим формулам:

$$S_x^2 = \frac{\sum_{i=1}^k (x_i - \bar{x})^2}{n-1}, \quad S_x^2 = \frac{1}{n-1} \left(\sum_{i=1}^k x_i^2 - \frac{(\sum_{i=1}^k x_i)^2}{n} \right).$$

А для выборки, распределенной в вариационный ряд, по формулам:

$$S_x^2 = \frac{\sum_{i=1}^k (x_i - \bar{x})^2 n_i}{n-1}, \quad S_x^2 = \frac{1}{n-1} \left(\sum_{i=1}^k n_i x_i^2 - \frac{(\sum_{i=1}^k x_i n_i)^2}{n} \right).$$

Среднее квадратическое отклонение (стандартное отклонение). Среднее квадратическое отклонение – это квадратный корень из дисперсии, обозначается S_x или σ . Выражается в тех же единицах, что и признак. Чем больше варьирует признак, тем больше величина этого показателя. Рассчитывается по формуле:

при $n < 30$

$$S_x = \sqrt{\frac{\sum_{i=1}^k n_i (x_i - \bar{x})^2}{n-1}}, \quad (8)$$

при $n \geq 30$

$$S_x = \sqrt{\frac{\sum_{i=1}^k n_i (x_i - \bar{x})^2}{n}}. \quad (9)$$

Коэффициент вариации. Дисперсия и среднее квадратическое отклонение – показатели «привязанные» к признаку, т.е. выражаются в тех же единицах что и сам признак. Если требуется сравнить разнородные показатели (например, рост и вес), то применяется относительный показатель – коэффициент вариации. Коэффициент вариации анализирует ту же характеристику, что и дисперсия со стандартным отклонением (а именно – вариабельность признака), но в относительных (безразмерных) единицах. Обозначается – V . Значения его изменяются от 0 до 100 %, но в некоторых случаях, при сильно асимметричном распределении может превышать 100 %.

$$V = \frac{S_x}{\bar{x}} \cdot 100\% \quad (10)$$

Если коэффициент вариации больше 35 %, то делается вывод, что изучаемая совокупность является неоднородной. В таких случаях в качестве характеристики наиболее типичного признака лучше вместо среднего арифметического использовать моду или медиану.

Пример 3. Выборочная совокупность задана распределением:

x_i	1	2	3	4
n_i	20	15	10	5

Найти выборочную дисперсию, среднее квадратическое отклонение, коэффициент вариации.

Решение. По формуле взвешенной средней найдем среднее арифметическое:

$$\bar{x} = \frac{\sum x_i \cdot n_i}{n} = \frac{1 \cdot 20 + 2 \cdot 15 + 3 \cdot 10 + 4 \cdot 5}{20 + 15 + 10 + 5} = \frac{100}{50} = 2.$$

По формуле для взвешенной дисперсии определим дисперсию:

$$\begin{aligned} S_x^2 &= \frac{\sum_{i=1}^k (x_i - \bar{x})^2 n_i}{n} = \\ &= \frac{(1-2)^2 \cdot 20 + (2-2)^2 \cdot 15 + (3-2)^2 \cdot 10 + (4-2)^2 \cdot 5}{50} = \frac{50}{50} = 1. \end{aligned}$$

Среднее квадратическое отклонение рассчитаем как корень квадратный из дисперсии:

$$S_x = \sqrt{S_x^2} = \sqrt{1} = 1.$$

Далее рассчитаем коэффициент вариации:

$$V = \frac{S_x}{\bar{x}} \cdot 100\% = \frac{1}{2} \cdot 100\% = 50\%.$$

2.3 Оценка параметров генеральной совокупности

Статистическая ошибка выборочной средней (ошибка репрезентативности выборочной средней). При вычислении среднего значения выборки допускается возможность неправильного отражения характера генеральной совокупности (по причине недостаточности материала или неправильно организованного исследования). Степень этой возможности показывает статистическая ошибка. Она не связана с неточностью измерения величины признака и выражается в тех же единицах, что и сама средняя. Обозначается: по старому – m , по новому – $S_{\bar{x}}$. Записывается в виде: $\bar{x} \pm S_{\bar{x}}$.

Вычисляется по следующей формуле:

$$S_{\bar{x}} = \frac{S_x}{\sqrt{n}} = \sqrt{\frac{S_x^2}{n}}, \quad (11)$$

где S_x – среднее квадратическое отклонение.

Точность определения средней величины позволяет оценить показатель C_s , который вычисляется по формуле:

$$C_s = \frac{S_{\bar{x}}}{\bar{x}} \cdot 100\% . \quad (12)$$

Точность средних показателей, которыми оценивают результаты наблюдений, считается удовлетворительной, если коэффициент C_s не превышает 3–5 %.

Одна из задач математической статистики – оценить параметры большого массива данных по исследованной ее части.

Статистическое оценивание может выполняться двумя способами:

1. Точечная оценка – оценка для некоторой определенной точки;

2. Интервальная оценка – по данной выборке оценивается интервал, в котором лежит истинное значение с заданной вероятностью.

Выборочное среднее \bar{x} является точечной оценкой генеральной средней μ .

Генеральная дисперсия имеет две точечных оценки: выборочная дисперсия и исправленная выборочная дисперсия.

Интервальная оценка для генеральной средней. По известным значениям выборочных характеристик можно установить интервал, в котором с той или иной вероятностью находится величина генерального параметра.

Вероятности, признанные достаточными для уверенного суждения о генеральных параметрах на основании выборочных показателей, называются **доверительными**.

Обычно в качестве доверительных используют три порога вероятностей. Выбор той или иной доверительной вероятности производится исследователем, исходя из практических соображений о той мере ответственности, с какой делаются выводы о генеральных параметрах. Чем выше мера ответственности, тем более высокий уровень доверительной вероятности: 99 % или 99,9 %.

Доверительная вероятность 0,95 (95 %) считается достаточной для исследований в области биологии и экологии.

Доверительным вероятностям соответствуют величины нормированных отклонений, которые приводятся в специальной статистической таблице (приложение А).

Доверительная вероятность (P)	0,95	0,99	0,999
Нормированное отклонение (t)	1,96	2,58	3,29

Границы доверительного интервала для генеральной средней, устанавливаемые по величине выборочной средней, определяются по формуле:

$$\bar{x} - \frac{t \cdot S_x}{\sqrt{n}} \leq \mu \leq \bar{x} + \frac{t \cdot S_x}{\sqrt{n}}, \quad (13)$$

где t – нормированный показатель распределения Стьюдента с $(n-1)$ степенями свободы. Он находится в специальной таблице распределения Стьюдента по соответствующему уровню значимости и числу степеней свободы (приложение А). Обычно в биологических исследованиях принимается 5%-й уровень значимости, т.е. допускается вероятность ошибки 0,05.

Следует помнить, что применение критерия Стьюдента возможно только для нормально распределенной совокупности.

Величину $\frac{t \cdot S_x}{\sqrt{n}}$ называют погрешностью средней величины, и среднее значение обычно записывают вместе с величиной погрешности:

$$\bar{x} \pm tS_{\bar{x}}. \quad (14)$$

Доверительный интервал для доли. Доверительный интервал для генеральной доли (P) устанавливается таким же способом, как и для генеральной средней, т.е.

$$p - \frac{t \cdot S_{pj}}{\sqrt{n}} \leq P \leq p + \frac{t \cdot S_{pj}}{\sqrt{n}}, \quad (15)$$

где S_{pj} – ошибка выборочной доли, определяемая по формуле:

$$(16)$$

$$S_{pj} = \sqrt{\frac{p(1-p)}{n}}.$$

Пример 4. Среднее значение прироста сосны в выборке из 20 деревьев равно 32,5 см. Определить доверительный интервал для среднего значения прироста сосны на изучаемом фитоценозе, если $S_x = 7,3$ см.

Решение. По таблице распределения Стьюдента найдем t .

Для числа степеней свободы $n - 1 = 20 - 1 = 19$ и уровня значимости $\alpha = 0,05$ $t = 2,09$. Тогда погрешность среднего равна

$$tS_x / \sqrt{n} = 2,09 \cdot 7,3 / \sqrt{20} = 15,26 / 4,47 = 3,4.$$

Среднее значение генеральной совокупности можно записать как $32,5 \pm 3,4$.

Нижняя граница доверительного интервала равна:

$$32,5 - 3,4 = 29,1$$

Верхняя граница доверительного интервала равна:

$$32,5 + 3,4 = 35,9.$$

Контрольные вопросы

1. Перечислить статистические показатели, относящиеся к средним величинам.

2. Перечислить статистические показатели, измеряющие степень варьирования признака.

3. Дать понятие простой и взвешенной средним арифметическим.

4. Что называется модой и медианой признака?

5. Что называется дисперсией?

6. Дать понятие среднему квадратическому отклонению.

7. Дать понятие коэффициенту вариации. В каких единицах он измеряется?

8. Дать понятие статистической ошибке выборочной средней.

Задания к самостоятельной работе

1. Переведите количественно и качественно измеряемые признаки в ранговые.

Урожайность (ц/га)	22,0	18,9	4,7	18,7	13,4	16,7	15,7	18,6	21,2
Качество почв (балл.)	5	1	0	2	1	5	4	2	3

2. При анализе эфиппий с Кубанского водохранилища в двадцати пробах были получены следующие данные для *Ceriodaphnia* (экз.):

20	19	21	18	18
19	19	18	17	19
20	21	18	17	19
19	18	19	20	18

Распределите данные в вариационный ряд, постройте полигон распределения и рассчитайте среднее и медиану.

3. Распределите следующие данные в вариационный ряд, постройте гистограмму и кумуляту распределения. Рассчитайте моду.

3,2	2,8	2,9	3,0	3,1	3,3	2,9
3,1	2,7	3,4	2,9	3,0	2,9	2,8
2,6	3,0	2,8	3,0	3,1	2,9	3,0

4. Определите медиану для ряда покрытий *Cladonia sylvatica* в 189 площадках по 1 м².

Классы	0–	5–	15–	25–	35–	45–	55–	65–	75–
	4	14	24	34	44	54	64	74	85
Частоты (n _i)	35	28	32	25	27	16	16	9	1

5. При анализе эфиппий с Кубанского водохранилища в разных пробах были получены следующие данные для *Ceriodaphnia* (экз):

20	19	21	18	18	20	21	18	17	19
19	19	18	17	19	19	18	19	20	18

Определите степень варьирования эфиппий в пробах.

6. Найти выборочное среднее, выборочную дисперсию и выборочное среднее квадратическое отклонение, если совокупность задана таблицей распределения:

x _i	0	1	2	3	4	5	7
n _i	8	7	16	10	6	2	1

7. Какой признак варьирует сильнее? Первый характеризуется средней $\bar{x}_1 = 2,5$ кг и средним квадратическим отклонением $S_{X1} = 0,62$ кг. Второй – величинами $\bar{x}_2 = 9,2$ см и $S_{X2} = 1,47$ см.

8. Рассчитайте доверительный интервал для высоты растений (см), если $S_x = 0,92$ см.

Высота, см	6,5	6,8	7,2	8,0	7,5	6,8	9	7,6	5,9	6,2
------------	-----	-----	-----	-----	-----	-----	---	-----	-----	-----

9. Найдите среднее арифметическое распределения:

x_i	1,0–1,9	2,0–2,9	3,0–3,9	4,0–4,9	5,0–5,9	6,0–7,0
n_i	0	15	23	5	6	1

10. В выборке из 50 листьев 18 поражены некрозом. Определите доверительный интервал для доли пораженных листьев.

11. На пяти площадках из 22 обследованных встречается вид клевер луговой (*Trifolium pratense* L.). Оцените доверительный интервал для встречаемости этого вида в биоценозе.

12. Оценить погрешность средней величины прироста сосны в 20 деревьях, выбранных случайно из 100:

39, 37, 30, 48, 36, 38, 22, 33, 21, 30, 25, 43, 28, 36, 20, 35, 20, 41, 36.

Тема 3. Предварительная статистическая обработка экологических данных с помощью программы Microsoft Excel

С помощью средств описательной статистики, представленных в пакете прикладных программ Microsoft Excel, рассмотрим три варианта статистической обработки данных: на основе вспомогательной таблицы и формул (ручной способ с промежуточными расчетами); на основе Мастера функций и с помощью надстройки Пакет анализа.

Пример 1. Используя электронные таблицы Microsoft Excel, найти выборочное среднее, выборочную дисперсию, выборочное среднее квадратическое отклонение и коэффициент вариации, если совокупность задана таблицей распределения:

x _i	0	1	2	3	4	5	6
n _i	8	7	16	10	6	2	1

Для расчета необходимых статистических характеристик будем пользоваться следующими формулами:

$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i n_i$	Среднее арифметическое
$S_x^2 = \frac{1}{n} \left(\sum x_i^2 n_i - \frac{(\sum x_i n_i)^2}{n} \right)$	Выборочная дисперсия
$S_x = \sqrt{S_x^2}$	Среднее квадратическое отклонение
$V = \frac{S_x}{\bar{x}} \cdot 100\%$	Коэффициент вариации

Решение:

1. Запустите программу EXCEL.

2. Заполните шапку расчетной таблицы и введите данные, как показано на рисунке 6.

3. Пропустив после таблицы 1 строку, последовательно введите название рассчитываемых показателей или их условное обозначение: среднее (\bar{x}), дисперсия (SS), среднее квадратическое отклонение (S), коэффициент вариации (V) (рисунок 6).

	A	B	C	D	E	F
1	N	x	n	xn	xx	xxn
2	1	0	8			
3	2	1	7			
4	3	2	16			
5	4	3	10			
6	5	4	6			
7	6	5	2			
8	7	6	1			
9	суммы					
10						
11	среднее					
12	дисперсия					
13	ср.кв.отклон.					
14	коэф. вар.					

Рисунок 6 – Подготовка данных к расчету статистических характеристик

4. Для промежуточных расчетов заполните в ячейках, D2, E2, F2 формулы для произведений соответственно: $x \cdot n$, $x \cdot x$, $x \cdot x \cdot n$. Для этого не забывайте указать знак «=» и мышкой отмечать необходимые аргументы формулы в той же строке, где набирается формула.

5. После того, как все формулы набраны и введены, выделите мышкой формульную строку, поставьте курсор мыши на маленький черный квадратик в правом нижнем углу ячейки-маркер заполнения и, не отпуская ее, протащите вниз до строки, содержащей последний ряд данных.

6. Выделите блок всей расчетной таблицы (за исключением «шапки»). На панели инструментов отметьте знак суммы (Σ). Получатся все необходимые промежуточные данные для подстановки в формулы для расчета соответствующих статистических характеристик.

7. Напротив названия статистической характеристики наберите ее расчетную формулу, указывая мышкой на ячейку, содержащую необходимые числа. В результате получим таблицу как на рисунке 7.

	A	B	C	D	E	F
1	N	x	n	xn	xx	xxn
2	1	0	8	0	0	0
3	2	1	7	7	1	7
4	3	2	16	32	4	64
5	4	3	10	30	9	90
6	5	4	6	24	16	96
7	6	5	2	10	25	50
8	7	6	1	6	36	36
9	суммы		50	109	91	343
10						
11	среднее		2,18			
12	дисперсия		2,1076			
13	ср. кв. отклон.		1,451758			
14	коэф. вар		0,665944	66,59438	%	

Рисунок 7 – Расчет статистических характеристик по формулам

8. Сохраните результаты расчетов в файле под именем СТАТИСТИКА в своей папке.

Пример 2. Наблюдаемые значения полуденной температуры мая заданы таблицей:

4	2	1	5	6	7
10	15	14	10	10	12
15	16	17	18	15	15
27	24	25	26	6	7
13	14	15	20	25	21
6					

С помощью пакета прикладных программ Microsoft Excel найти выборочное среднее, моду, медиану, выборочную дисперсию, выборочное среднее квадратическое отклонение, коэффициент вариации, статистическую ошибку выборочной средней, точность определения средней, доверительный интервал для генеральной средней ($\alpha = 0,05$).

Решение.

1. На втором листе файла СТАТИСТИКА введите данные, помещая каждое число в отдельную ячейку.

2. Пропустив после данных свободную строку, в первом столбике таблицы последовательно наберите название статистических характеристик, которые необходимо рассчитать.

3. Рассчитайте среднее значение полуденной температуры мая. Для этого в ячейку напротив наименования статистической характеристики поместите пустую рамку. В строке меню щелкните левой кнопкой мыши по слову Вставка. В выпавшем диалоговом окне выберите пункт Функция. Среди категорий функции выберите Статистические, а в правом окне с перечислением имен функций выберите функцию СРЗНАЧ (среднее значение аргумента). Нажмите кнопку ОК.

4. В выпавшем диалоговом окне укажите границы массива аргументов в окне под названием Число1. Для этого выделите мышью ячейки листа, содержащие сведения о распределении полуденной температуры мая – A1:F5.

5. При нажатии кнопки ОК рассчитанное значение средней величины помещается в обозначенную вами ячейку, где ранее стояла рамка.

6. Аналогично, вызывая соответствующие функции, рассчитайте моду, медиану, дисперсию, среднее квадратическое отклонение.

7. Остальные статистические характеристики рассчитываются по соответствующим формулам с использованием

предварительно рассчитанных величин (на них необходимо указывать мышкой). Если в формуле используется квадратный корень, то его можно вычислить, набирая с клавиатуры слово «корень», не забывая при этом ставить перед ним знак «=» и в скобках помещать аргумент функции, указывая на него мышкой.

8. Для расчета t-критерия используется функция СТЬЮДРАСПРОБР. Для нее необходимо указать вероятность (0,05) и число степеней свободы (n-1).

9. Сохраните расчеты (рисунок 8) в файле СТАТИСТИКА.

The screenshot shows the Microsoft Excel interface with a menu bar (Файл, Правка, Вид, Вставка, Формат, Сервис, Данные, Окно, Справка) and a toolbar. The active window is titled "Microsoft Excel - СТАТИСТИКА". The spreadsheet contains data in columns A through H and rows 1 through 19. Row 4 is highlighted. Below the data table, rows 8 through 19 contain statistical calculations.

	A	B	C	D	E	F	G	H
1	4	2	1	5	6	7		
2	10	15	14	10	10	12		
3	15	16	17	18	15	15		
4	27	24	25	26	6	7		
5	13	14	15	20	25	21		
6	6							
7								
8	Среднее				13,58065			
9	Мода				15			
10	Медиана				14,29032			
11	Дисперсия				53,51828			
12	Ср кв. отклонение				7,315619			
13	Козффци. вариации				0,53868			
14	Ошибка средней				1,313924			
15	Точность определения средней				0,09675			
16	t критерий				2,042272			
17	Погрешность средней				2,727747			
18	Нижняя граница дов интервала				10,8529			
19	Верхняя граница дов. интервала				16,30839			

Рисунок 8 – Расчет статистических характеристик с помощью Мастера функций MS EXCEL

Пример 3. Сравниваются уровни загрязнения катионами Cu^{2+} (мг/л) на двух участках водоема с эталонным. Рассчитайте основные статистические характеристики для этих показателей с помощью надстройки **Пакет Анализа**.

Эталонный участок	0,05	0,06	0,01	0,04	0,11	0,15	0,09	0,08
Участок 1	0,08	0,12	0,15	0,14	0,09	0,13	0,11	0,08
Участок 2	0,11	0,13	0,15	0,15	0,09	0,08	0,10	0,15

Решение.

1. Введите данные на 3-й лист файла **СТАТИСТИКА**, располагая их по столбцам.

2. Выберите в меню **Данные пункт Анализ данных**, появится окно с одноименным названием.

В области **Инструменты анализа** представлен список реализованных в MS Excel методов статистической обработки (рисунок 9).

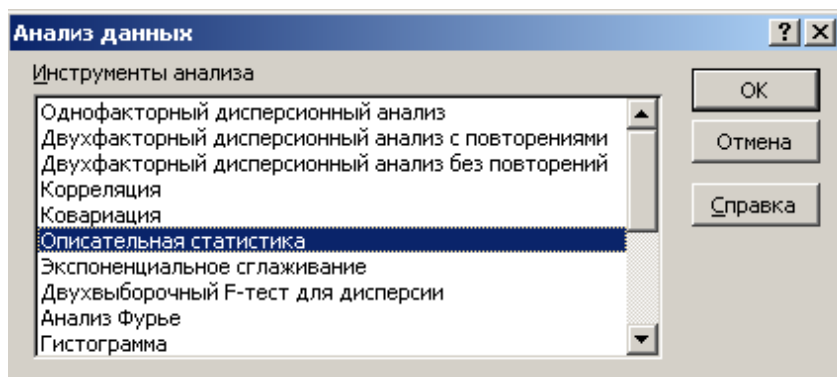


Рисунок 9 – Окно Анализа данных

3. Выберите пункт **Описательная статистика**.

4. Поставьте птички на параметрах **Итоговая статистика** и **Уровень надежности** и укажите на переключатель **по столбцам**.

5. В поле **Входной интервал** введите ссылку на ячейки, содержащие анализируемые данные и нажмите **ENTER** (рисунок 10).

6. В поле **Выходной интервал** укажите на левую верхнюю ячейку выходного диапазона.

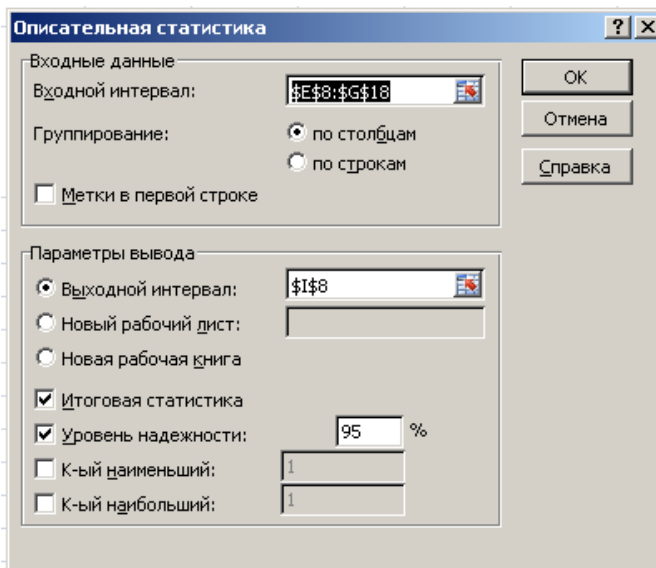


Рисунок 10 – Ввод данных для анализа

7. Нажмите **ОК**. Итоговая статистика будет выведена в указанный вами диапазон (рисунок 11).

<i>Столбец1</i>		<i>Столбец2</i>		<i>Столбец3</i>	
Среднее	0,081818182	Среднее	0,110909091	Среднее	0,121818182
Стандартная о	0,01212197	Стандартная о	0,011318091	Стандартная о	0,008719694
Медиана	0,08	Медиана	0,11	Медиана	0,13
Мода	0,11	Мода	0,08	Мода	0,15
Стандартное о	0,040204025	Стандартное о	0,03753786	Стандартное о	0,028919952
Дисперсия вы	0,001616364	Дисперсия вы	0,001409091	Дисперсия вы	0,000836364
Эксцесс	-0,242968785	Эксцесс	-0,293333333	Эксцесс	-1,708049779
Асимметрично	-0,128329317	Асимметрично	0,273959972	Асимметрично	-0,129793547
Интервал	0,14	Интервал	0,13	Интервал	0,08
Минимум	0,01	Минимум	0,05	Минимум	0,08
Максимум	0,15	Максимум	0,18	Максимум	0,16
Сумма	0,9	Сумма	1,22	Сумма	1,34
Счет	11	Счет	11	Счет	11
Уровень надеж	0,027009431	Уровень надеж	0,025218277	Уровень надеж	0,019428688

Рисунок 11 – Итоговая статистика по трем показателям

Тема 4. Совместный анализ нескольких выборок

4.1 Сущность параметрических и непараметрических методов сравнения выборок

Следующей задачей статистического анализа, решаемой после определения основных (выборочных) характеристик и анализа одной выборки, является совместный анализ нескольких выборок.

Ни одно исследование не обходится без сравнений. Сравниваются данные ключевых участков мониторинга, эмпирически полученные данные с теоретическими.

Методы сравнения выборок делятся на **параметрические** и **непараметрические**.

Параметрические связаны с необходимостью вычисления выборочных характеристик – средней величины и показателей вариации и применяются в случаях, когда изучаемая совокупность подчиняется закону нормального распределения. В этой ситуации задача оценки различий двух групп независимых наблюдений решается с использованием **параметрических критериев** статистики: критерия Стьюдента (t), если сравнение выборок ведется по средним значениям (\bar{x}_1 и \bar{x}_2), или критерия Фишера (F), если сравнение выборок ведется по их дисперсиям.

В действительности не все биологические объекты распределяются нормально. Кроме того, исследователю часто приходится иметь дело не только с количественными, но и с качественными, или атрибутивными признаками, которые могут выражаться порядковыми номерами, индексами, баллами. В таких случаях применяются **непараметрические, или порядковые критерии**. К ним относятся критерий знаков, двухвыборочный критерий Вилкоксона, критерий Ван дер Вардена, критерий Уайта. Выбор этих критериев не требует большого числа членов выборки и знаний вида распределения, но все же зависит от целого ряда условий.

4.2 Проверка гипотезы о нормальном законе распределения

Предположение о нормальном законе распределения используется во многих статистических методах обработки информации, например, в регрессионном анализе при установлении зависимостей между случайными величинами, в дисперсионном анализе при проверке статистических гипотез. Поэтому выполнимость предположения о нормальности при обработке данных рекомендуется обязательно проверять.

Рассмотрим несколько методов, которые применяются для проверки, согласуются ли экспериментальные данные с нормальным распределением.

1. По выборочным значениям из генеральной совокупности визуально оценивается закон распределения данной совокупности в предположении, что **кривая нормального распределения** имеет колоколообразную форму и строго симметрична.

2. Вычисляется среднее арифметическое и медиана. Если они друг от друга значительно не отличаются, считается, что распределение подчиняется нормальному закону. В противном случае мы имеем дело с асимметричной выборкой.

3. Для малых выборок используется свойство, что показатели асимметрии и эксцесса в нормально распределенной совокупности равны нулю. Асимметрия показывает «скошенность» кривой распределения относительно нормальной кривой, а эксцесс замеряет «заостренность» кривой (положительный – заостренная кривая, отрицательный – «тупая»). Рассчитываются асимметрия, эксцесс и их ошибки по следующим формулам:

$$A = \frac{\sum (x_i - \bar{x})^3}{nS_x^3}, \quad (17)$$

$$m_A = \sqrt{6/n}, \quad (18)$$

$$E = \frac{\sum (x_i - \bar{x})^4}{nS_x^4} - 3, \quad (19)$$

$$m_E = 2\sqrt{6/n}, \quad (20)$$

где А и Е – асимметрия и эксцесс, а m_A и m_E – их ошибки репрезентативности.

Нулевая гипотеза, или предположение, что в генеральной совокупности показатели асимметрии и эксцесса равны нулю, опровергается, если абсолютные величины асимметрии и эксцесса превышают более чем втрое свои ошибки репрезентативности.

4. Проверяется однородность совокупности оценкой соотношения V (коэффициент вариации) $\leq 33\%$. Если совокупность неоднородна, следует исключить из нее самые "аномальные" наблюдения, поскольку они, скорее всего, нетипичны для данного исследования. Для устранения аномальных наблюдений используется правило «трех сигм»: Наблюдение признается аномальным и отбрасывается, если его отклонение от выборочной средней более чем в 3 раза превышает среднеквадратическое отклонение выборки. То есть для нормально распределенных данных должно выполняться неравенство $|\bar{x} - x_i| \leq 3\sigma$. Безусловно, любые операции с исходной совокупностью, в том числе и связанные с изменением ее объема, должны быть обоснованными и поясняемыми.

5. При размере выборки больше 100 наиболее известным критерием, используемым для проверки распределения на нормальность, является критерий согласия хи-квадрат Пирсона. Его наблюдаемое значение рассчитывается по формуле:

$$\chi^2_{\text{набл}} = \sum_{i=1}^k \frac{(n_{\text{эмпир}} - n_{\text{теор}})^2}{n_{\text{теор}}} . \quad (21)$$

Принимается уровень значимости α (или p – другое обозначение, часто используемое в пакетах ПП) – это вероятность справедливости нулевой гипотезы при условии ее отвержения. Теоретическое значение критерия находят по специальной таблице по уровню значимости α и числу степеней свободы f . В общем случае $f = k - m - 1$, где k – число классов, m – число неизвестных параметров предполагаемой модели, оцениваемых по данным выборки. Если все параметры предполагаемой модели известны, $m = 0$. Так как для нормального закона распределения известно два параметра (средняя величина и стандартное отклонение), то $m = 2$ и $f = k - 3$.

Если $\chi^2_{\text{набл}} < \chi^2_{\text{крит.}}$ ($\alpha \leq 0,05$, $f = k - 3$), то принимается нулевая гипотеза, что данное распределение подчиняется нормальному закону. Если $\chi^2_{\text{набл}} \geq \chi^2_{\text{крит.}}$ ($\alpha \leq 0,05$, $f = k - 3$), то принимается альтернативная гипотеза, что данное распределение не подчиняется нормальному закону.

Критерий предполагает, что результаты наблюдений сгруппированы в вариационный ряд и разбиты на классы, количество которых должно быть не менее 8. Количество вариантов в классах также должно быть не менее 8, кроме крайних интервалов, в которых число вариантов может быть меньше.

Пример 1. Имеются данные по содержанию меди в почве в 30 пробах (мг/кг). Проверить, подчиняется ли распределение меди в пробах нормальному закону.

3,72	3,92	4,07	2,86	4,25	3,72
5,04	5,38	4,42	4,24	3,16	4,35
4,77	4,72	4,18	3,92	3,64	4,85
4,82	4,16	4,47	3,94	3,28	3,52
2,85	4,82	5,17	4,82	3,27	3,72

Решение. Проверку на соответствие нормальному закону распределения будем проводить на основе коэффициентов эксцесса и асимметрии. Нулевая гипотеза будет заключаться в утверждении, что данные в генеральной совокупности подчиняются нормальному закону распределения. Вычисление проводится стандартным способом с помощью мастера функций MS EXCEL по следующему алгоритму:

1. Введите данные. Поставьте пустую рамку в ячейку, куда будет помещаться результат.

2. Зайдите в Мастер функций. Среди категорий Статистические выберите функцию ЭКСЦЕСС. Нажмите кнопку ОК. В выпавшем диалоговом окне укажите границы массива аргументов в строке под названием Число1. Для этого выделите мышью ячейки листа, содержащие сведения о содержании меди в пробах и нажмите ENTER. Рассчитанное значение коэффициента помещается в ячейку, где ранее стояла рамка.

Для расчета асимметрии вызовите функцию СКОС и введите необходимые значения переменной.

Рассчитайте ошибки эксцесса (18) и асимметрии (16) по соответствующим формулам.

Расчетное значение эксцесса получилось $-0,71428$, а ошибка коэффициента эксцесса равна $0,89427$. Отношение коэффициента эксцесса к ошибке по модулю равно $0,79859$.

Расчетное значение асимметрии равно $-0,15967$, а ее ошибка $0,447214$. Отношение этих показателей по модулю равно $0,35704$.

Так как в обоих случаях отношения коэффициентов к их ошибкам не превосходят 3, нулевую гипотезу отбросить нельзя, что дает основание говорить о нормальном характере распределения содержания меди в генеральной совокупности.

Пример 2. Используя критерий Пирсона, при уровне значимости $0,05$ проверьте гипотезу о нормальном распределении совокупности площадей листьев ($n = 200$).

x_i	5	7	9	11	13	15	17	19	21
n_i	2	15	29	50	35	28	24	12	5

Решение. Схема решения приведена на рисунке 12. Сначала рассчитаем по известной схеме параметры распределения. $(\bar{x}; S)$. Затем сформируем таблицу для расчета теоретических частот нормального закона. При помощи встроенной функции НОРМРАСП с параметрами $(x; \bar{x}; S; 0)$ для каждого значения x найдем значение функции плотности нормального распределения $P(t)$. Например, ячейка F2 содержит формулу =НОРМРАСП (A2; среднее; станд. откл; 0). В следующем столбике рассчитаем теоретические частоты нормального распределения по формуле $P(t) \cdot i \cdot n$, где i – классовый интервал (например, ячейка G3 содержит формулу =(F3·2·объем).

В столбике H определяются округленные теоретические частоты с помощью функции =ОКРУГЛ(G2; 0). Затем рассчитаем статистику $\chi^2_{\text{эмпир.}}$ как сумму всех элементов последнего столбца.

Теоретическое значение $\chi^2_{\text{теор.}}$ рассчитаем с помощью функции ХИ2ОБР (уровень значимости; число степеней свободы). В нашем примере ХИ2ОБР (0,05; 9 – 2 – 1). Схема расчета приведена на рисунке 12. Так как $\chi^2_{\text{набл.}} < \chi^2_{\text{теор.}}$, то нулевая гипотеза о соответствии данных по площади листьев нормальному закону принимается.

	A	B	C	D	E	F
1	ж	пз	$t=(x-\bar{x})/S/\sigma$	$P(t)$	пт	$(пт-пз)^2/пт$
2	5	2	-2,153	0,01	4	1,299
3	7	15	-1,595	0,03	12	0,502
4	9	29	-1,036	0,07	26	0,333
5	11	50	-0,478	0,10	40	2,635
6	13	35	0,081	0,11	44	1,999
7	15	28	0,640	0,09	36	1,908
8	17	24	1,198	0,05	22	0,295
9	19	12	1,757	0,02	10	0,643
10	21	5	2,315	0,01	3	1,239
11	сумма	200		Хи квадрат эмп.		10,790
12				Хи квадрат теор.		12,592
13						
14	объем	200,00				
15	среднее	12,71				
16	SS	12,82				
17	S	3,58				

Рисунок 12 – Проверка на нормальный закон распределения по критерию χ^2

4.3 Критерий Фишера сравнения дисперсий

В некоторых экологических исследованиях важно сравнить показатели рассеяния. Доказательство достоверности различий дисперсий в опытных группах позволяет установить влияние анализируемых факторов (удобрения, корма, лекарства, химического вещества и т. д.) на изучаемые признаки или биологические свойства (урожайность, устойчивость к загрязнениям, к заболеваниям и т. д.).

В качестве критерия проверки нулевой гипотезы о равенстве генеральных дисперсий используют случайную величину F (критерий Фишера), равную отношению большей исправленной выборочной дисперсии к меньшей.

$$F_{\text{набл}} = \frac{S_{x \text{ большая}}^2}{S_{x \text{ меньшая}}^2}. \quad (22)$$

По таблице распределения Фишера (приложение Б) находим F критическое $F_{\text{крит}}(\alpha, f_1 = n_1 - 1, f_2 = n_2 - 1)$ со степенями свободы f_1 и f_2 , где n_1 – объем выборки, по которой вычислена большая выборочная дисперсия. В таблице f_1 располагается по горизонтали.

Если $F_{\text{набл}} < F_{\text{крит}}(\alpha, f_1, f_2)$ то принимается нулевая гипотеза, что генеральные дисперсии различаются незначимо. В противном случае принимается альтернативная гипотеза о статистически значимом различии генеральных дисперсий.

Пример 3. Имеется две выборки объемом $n_1 = 11$ и $n_2 = 28$ из популяций обыкновенных и овальных прудовиков, для которых отношение высоты к ширине имеют дисперсии $S_{x1}^2 = 0,59$ и $S_{x2}^2 = 0,38$. Необходимо проверить гипотезу о равенстве дисперсий для изучаемых популяций при уровне значимости $\alpha = 0,05$.

Решение.

По формуле $F_{набл} = \frac{S_{x\text{ большая}}^2}{S_{x\text{ меньшая}}^2}$ найдем $F_{набл} = \frac{0,59}{0,38} = 1,55$.

С помощью таблиц для критерия Фишера или с помощью функции FРАСПОБР (вероятность; степень свободы 1; степень свободы 2) найдем $F_{крит}$:

$$F_{крит}(\alpha \leq 0,05; f_1 = n_1 - 1 = 11 - 1 = 10, f_2 = n_2 - 1 = 28 - 1 = 27) = 2,2.$$

Так как $F_{набл} < F_{крит}$, нет оснований отвергнуть нулевую гипотезу о равенстве генеральных дисперсий.

Пример 4. Два лаборанта выполняли определение кальция в воде методом комплексонометрического титрования. Было выполнено по 5 параллельных определений. Дисперсия результатов первого лаборанта составила 0,00045, второго – 0,00225. Значимо ли различие между дисперсиями?

Решение. Вычисляем отношение дисперсий:

$F_{набл} = 0,00225 / 0,00045 = 5$. Сравниваем с табличным значением критерия Фишера: $F_{крит}(\alpha = 0,05, f_1 = 4, f_2 = 4) = 6,39$. Вывод: так как $F_{набл} < F_{крит}$, дисперсии отличаются незначимо.

4.4 Критерий Стьюдента (t-критерий) сравнения средних величин

Чтобы применить этот критерий, необходимо быть уверенным, что генеральные совокупности, из которых взяты сравниваемые выборки, распределены нормально и что их генеральные дисперсии равны между собой.

При использовании критерия можно выделить два случая: для независимых и для зависимых выборок.

4.4.1 T-критерий для независимых выборок

Независимые выборки получаются, если мы из генеральной совокупности случайным образом выбрали какие-либо элементы и поделили их на две группы, либо также случайно, либо относительно некоторого признака.

Выясняется, различаются ли средние некоторой переменной, вычисленные в разных подгруппах наблюдений. Например, по одной и той же переменной сравниваются данные с фонового участка и с загрязненного, с опытной группы и с контрольной.

Когда обе сравниваемые группы обладают большой численностью ($n > 30$) не обязательного одинакового объема, $t_{набл}$ рассчитывается по формуле:

$$t_{набл} = \frac{|\bar{x}_1 - \bar{x}_2|}{\sqrt{\frac{S_{x_1}^2}{n_1} + \frac{S_{x_2}^2}{n_2}}}. \quad (23)$$

При сравнении двух групп с малым n ($n < 30$) величину критерия Стьюдента находим по формуле:

$$t_{набл} = \frac{|\bar{x}_1 - \bar{x}_2|}{\sqrt{(n_1 - 1)S_{x_1}^2 + (n_2 - 1)S_{x_2}^2}} \cdot \sqrt{\frac{n_1 \cdot n_2 (n_1 + n_2 - 2)}{n_1 + n_2}}. \quad (24)$$

По специальной таблице (приложение А) по принятому уровню значимости и числу степеней свободы $f = n_1 + n_2 - 2$ находят t критическое.

Если $|t_{набл}| < t_{крит}(\alpha, f)$, то принимается нулевая гипотеза о том, что нет различий между сравниваемыми параметрами. В противном случае, при $|t_{набл}| > t_{крит}(\alpha, f)$, принимается

альтернативная гипотеза, что сравниваемые средние величины различны на уровне значимости α .

Пример 4. При исследовании местности в течении двух сезонов случайным образом были отобраны участки и установлено число пораженных болезнью деревьев. Получены следующие результаты: $n_1 = 16$; $\bar{x}_1 = 350$; $S_{x_1}^2 = 9$; $n_2 = 20$; $\bar{x}_2 = 320$; $S_{x_2}^2 = 16$. Проверить, является расхождение среднего числа пораженных деревьев случайным или же болезнь пошла на убыль.

Решение. По формуле

$$t_{набл} = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{(n_1 - 1)S_{x_1}^2 + (n_2 - 1)S_{x_2}^2}} \cdot \sqrt{\frac{n_1 \cdot n_2 (n_1 + n_2 - 2)}{n_1 + n_2}}$$

определим

$$t_{набл} = \frac{350 - 320}{\sqrt{15 \cdot 9 + 19 \cdot 16}} \cdot \sqrt{\frac{16 \cdot 20 \cdot (16 + 20 - 2)}{16 + 20}} = 24,9.$$

Табличное значение $t_{крит}(\alpha; f = n_1 + n_2 - 2) = t_{крит}(0,05; 34) = 2,03$.

Так как $|t_{набл}| > t_{крит}(\alpha, f)$, различия в среднем числе пораженных растений, измеренном в разные моменты времени статистически значимо на 5 % уровне значимости. То есть болезнь пошла на убыль.

Пример 5. При исследовании качества атмосферного воздуха по площади листовой пластинки в двух парках, расположенных на разном расстоянии от источника загрязнения, были рассчитаны площади листьев тополя. Получены следующие результаты: $n_1 = 50$; $\bar{x}_1 = 35$; $S_{x_1}^2 = 2,5$; $n_2 = 60$; $\bar{x}_2 = 32$;

$S_{x_2}^2 = 1,6$. Влияет ли источник загрязнения на площадь листовой пластинки?

Решение. Воспользуемся формулой (23).

$$t_{набл} = \frac{35 - 32}{\sqrt{\frac{2,5}{50} + \frac{1,6}{60}}} = 10,83.$$

Найдем

$$t_{крит}(\alpha; f = n_1 + n_2 - 2) = f(0,05; 50 + 60 - 2 = 108) = 1,98.$$

Так как $|t_{набл}| > t_{крит}(\alpha, f)$, то средние значения площадей листьев различимы на 5% уровне значимости.

4.4.2 Парный t -критерий для зависимых, связанных выборок

Выборка называется парной, когда один элемент в первой группе по какому-либо содержательному признаку сопоставляется с соответствующим элементом во второй группе. Зависимые связанные выборки можно получить, если рассматривать одну переменную до и после воздействия внешнего фактора или при двух разных воздействиях. Процедуры сравнения зависимых выборок применяются для обнаружения результата какого-либо воздействия, либо, напротив, для подтверждения его отсутствия.

Рассмотрим пример использования t -критерия Стьюдента для связанных и, очевидно, равных по численности выборок. В случае связанных выборок с равным числом измерений в каждой вычисление значения t осуществляется по формуле:

$$t_{эмт} = \frac{\bar{d}}{S_d}, \quad (25)$$

где S_d – ошибка разницы, которая определяется по формуле:

$$S_d = \sqrt{\frac{\sum d_i^2 - \frac{(\sum d_i)^2}{n}}{n \cdot (n-1)}}, \quad (26)$$

где $d_i = x_i - y_i$ – разности между соответствующими значениями переменной X и переменной Y, а \bar{d} – среднее этих разностей. Число степеней свободы $df = n - 1$.

Если $t_{эмп} < t_{крит}$, то нулевая гипотеза принимается, в противном случае принимается альтернативная гипотеза.

Пример 6. Сравнить остаточную плотность популяции колорадского жука (%) при использовании инсектицидов «Карате» и «Децис».

Температура °С	10	12	14	16	18	20	22	24	26	28
«Карате»	98	97	97	96	90	85	74	70	64	62
«Децис»	100	95	96	90	85	70	64	65	60	60

Решение. Выдвигаем содержательную гипотезу: препараты «Карате» и «Децис» по-разному влияют на плотность колорадского жука.

Решим задачу с помощью электронных таблиц MS EXCEL (рисунок 13), применив для вычисления статистических показателей указанные выше формулы.

Так как выборки статистически различаются на 5 % уровне значимости, то можно сделать вывод, что эффективность препаратов различная.

	A	B	C	D	E
1					
2	Температура °С	“Карате” X1	“Децис” X2	d	dd
3	10	98	100	-2	4
4	12	97	95	2	4
5	14	97	96	1	1
6	16	96	90	6	36
7	18	90	85	5	25
8	20	85	70	15	225
9	22	74	64	10	100
10	24	70	65	5	25
11	26	64	60	4	16
12	28	62	60	2	4
13	суммы			48	440
14					
15	n	10			
16	d среднее		4,8		
17	Sd	1,526069752		$S_d = \sqrt{\frac{\sum d_i^2 - \frac{(\sum d_i)^2}{n}}{n \cdot (n-1)}}$	
18	tэмп	3,145334604			
19	tтеор(0,05;9)	2,262157158			
20	Выборки статистически различаются на 5% уровне значимости				

Рисунок 13 – Сравнение попарных данных по t -критерию

4.5 Непараметрические методы сравнения выборок

Критерий знаков. Сравнивая выборки с попарно связанными вариантами, наблюдаемые между ними различия обозначают знаками плюс (положительный эффект) и минус (отрицательный эффект воздействующего на признак фактора). Считается, что если попарно сравниваемые выборки не различаются, то число плюсовых и минусовых разностей окажется одинаковым. Если же налицо заметное преобладание плюсов или минусов, то это может быть следствием воздействия на признак учитываемого фактора. Нулевая гипотеза проверяется с помощью критерия знаков (z). Величина критерия определяется большим числом однозначных разностей. При этом нулевые разности в расчет не принимаются. Значимость проверяется с помощью специальной таблицы (прило-

жение В), в которой содержатся критические точки критерия (z_{st}) для 5 и 1%-го уровня значимости и числа парных наблюдений n (без нулевых разностей). Нулевая гипотеза опровергается, если $z_{\phi} \geq z_{st}$.

Пример 8. Для выявления тенденции изменения состояния деревьев на изучаемой территории исследовалась жизнённость 15 деревьев по пятибалльной шкале с временным интервалом в 3 года (чем выше балл, тем хуже состояние деревьев).

Результаты двукратного выполнения работы (в баллах) оформлены в таблице.

№ дерева	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
2001 г	2	2	2	2	2	3	3	3	3	3	2	2	3	3	3
2003 г	2	3	3	4	3	2	3	4	4	3	4	3	2	4	4
Знак разности	0	-	-	-	-	+	0	-	-	0	-	-	+	-	-

Проверяется гипотеза H_0 : жизненное состояние деревьев не изменилось за 3 года. Альтернативная гипотеза: состояние деревьев ухудшилось за 3 года. Подсчитаем значение статистики критерия z_{ϕ} , равное числу отрицательных разностей баллов. Согласно табличным данным $z_{\phi} = 10$, $n = 12$.

Для определения критических значений статистики критерия z_{st} смотреть приложение С. Для уровня значимости $\alpha = 0,05$ при $n = 12$ значение $z_{st} = 10$. Следовательно, выполняется неравенство $z_{\phi} \geq z_{st}$ ($10 = 10$). Поэтому в соответствии с правилом принятия решения нулевая гипотеза отклоняется на уровне значимости 0,05 и принимается альтернативная гипотеза, что позволяет сделать вывод об ухудшении состояния деревьев за 3 года.

Т-критерий Уайта. Алгоритм применения этого критерия сводится к следующему: все члены сравниваемых выборок располагаются в возрастающем порядке в один ранжиру-

емый ряд. Затем каждой варианте присваивается ранг. Если сравниваемые выборки не отличаются друг от друга, то и суммы их рангов должны быть равны. Чем значительнее расхождения между выборками, тем больше будет разница между суммами их рангов. Случайная разница оценивается с помощью Т-критерия Уайта, критическое значение которого (T_{st}) для 5%-го и 1%-го уровня значимости с учетом n_1 и n_2 сравниваемых выборок приводятся в специальной таблице (см. таблицу XI, Г. Ф. Лакин, с. 282). Если фактически найденное значение Т-критерия окажется меньше табличного для n_1 и n_2 и принятого уровня значимости, то нулевая гипотеза отвергается, и расхождения считаются случайными.

Пример 9. Имеются данные по плотности определенного вида растений на двух точках мониторинга. Требуется проверить гипотезу о том, что нет расхождений в средней величине плотности этого растения на изучаемых точках.

1-я точка (экз/м ²)	7	10	14	15	12	16	12
2-я точка (экз/м ²)	11	12	16	13	18	15	

Решение. Расположим обе выборки в один ряд по возрастанию значений признаков и определим их ранги:

Плотность (экз/м ²)	7	10	11	12	12	12	13	14	15	15	16	16	18
Ранг	1	2	3	5	5	5	7	8	9,5	9,5	11,5	11,5	13

Отдельно для каждой выборки найдем суммы рангов:

$$1\text{-я точка: } 1 + 2 + 8 + 9,5 + 5 + 11,5 + 5 = 42,$$

$$2\text{-я точка: } 3 + 5 + 11,5 + 7 + 13 + 9,5 = 49.$$

Меньшую сумму рангов (T_{ϕ}) сравниваем с табличным значением этого критерия для $n_1 = 7$ и $n_2 = 6$ и 5%-го уровня

значимости, которое равно $T_{st} = 7$ (см. Г. Ф. Лакин, таблица XI, с. 283). Так как $T_{\phi} \geq T_{st}$ нулевую гипотезу отвергать нельзя. Разница между плотностью изучаемого вида растений на выбранных точках статистически не достоверна.

Контрольные вопросы

1. На какие группы делят методы сравнения выборок?
2. Какие вы знаете параметрические критерии статистики?
3. Какие вы знаете непараметрические критерии статистики?
4. Какие методы применяются для проверки, согласуются ли экспериментальные данные с нормальным распределением?
5. Когда в экологических исследованиях используют критерий Фишера?
6. Когда в экологических исследованиях используют критерий Стьюдента (t-критерий)?
7. Какая выборка называется парной?
8. Перечислить непараметрические методы сравнения выборок.

Задания к самостоятельной работе

1. Совокупность 267 датских угрей распределилась по числу позвонков следующим образом:

X	111	112	113	114	115	116	117	118	119
N	3	9	31	71	82	46	19	5	1

Характеристики этого распределения следующие: $\bar{x} = 114,74; S_x = 1,35$. Рассчитать теоретические частоты нормального распределения.

2. При уровне значимости $\alpha \leq 0,05$ проверить гипотезу о нормальном распределении генеральной совокупности, если известны эмпирические и теоретические частоты:

Nэ	6	13	38	74	106	85	30	14
Nт	3	14	42	82	99	76	37	13

3. Определить, отличается ли динамика биомассы активного ила в очистных сооружениях, $\text{кг}/\text{м}^2$ при изменении уровня разбавления сточных вод. Получить доверительные границы изменения генерального показателя.

Месяц	1	2	3	4	5	6	7	8	9	10	11	12
Разб. в 2 раза	3,5	3,7	4,0	3,6	3,5	4,0	4,3	4,3	4,0	3,9	3,8	3,5
Разб. в 3 раза	4,0	3,5	3,0	3,1	4,5	4,5	3,7	3,9	3,9	2,5	2,0	3,1

4. Существуют ли различия в уровне загрязнения Pb^{2+} $\text{мг}/\text{кг}$ растений на эталонном участке (лес в районе г. Горячий Ключ) и в агроэкосистеме, расположенной вблизи автодороги Горячий Ключ – Краснодар? Получить доверительные границы изменения генерального показателя.

Эталонный участок	Автотрасса
0,002	0,0021
0,0015	0,0024
0,001	0,0031
0,0009	0,003
0,0007	0,0015
0,0025	0,0013
0,0021	0,0024
0,0014	–

5. Определить, отличается ли концентрация окислов азота на границе санитарно-защитной зоны от ПДК = 0,085 мг/м³. Получить доверительные границы изменения генерального показателя.

Протяженность границы СЗЗ, км		0,5	1,3	1,2	1,0	1,2	1,3	0,7
Концентрация окислов азота, мг/м ³	Вар.1	0,09	0,1	0,06	0,08	0,12	0,05	0,09
	Вар.2	0,05	0,06	0,11	0,12	0,03	0,02	0,09

6. При измерении величины газообмена в опытной ($n_1 = 10$) и контрольной ($n_2 = 10$) группах животных были получены соответственно следующие величины дисперсий – $S_{x_1}^2 = 163,9$ и $S_{x_2}^2 = 89,3$. Можно ли считать существенным различие в изменчивости процесса газообмена в опытной и контрольной группах животных?

7. Численность изучаемого вида в 35 выборках распределилась следующим образом.

42	40	43	41	50
52	45	42	52	52
40	42	43	41	39
50	50	49	40	45
41	41	39	32	28

Определите тип распределения изучаемого вида.

8. Два лаборанта выполняли определение кальция в воде методом комплексонометрического титрования. Было выполнено по 5 параллельных определений. Дисперсия результатов первого лаборанта составила 0,00045, второго – 0,00225. Значимо ли различие между дисперсиями?

Тема 5. Оценка плотности популяции

5.1 Оценка пространственного распределения особей

При оценке состояния экологических систем очень важно знать численность или плотность определенных видов организмов. Например, о состоянии экосистемы можно судить по плотности вида – индикатора. В принципе, с плотностью может быть связана и любая другая однозначно и четко регистрируемая величина, линейно отражающая численность особей: содержание ДНК (РНК), число фекалий животных на единицу маршрута, количество силуэтов птиц, появляющихся на фоне диска Луны за единицу времени и др.

На основе статистического анализа численностей организмов, встреченных в пробах, в экологических исследованиях основываются важные экологические заключения о характере среды обитания и типе биотических взаимодействий между особями. Кроме того, решается вопрос о размере учетной площадки и количестве отбираемых проб для адекватного отображения реальной ситуации на изучаемом объекте.

Обычно до того, как определяют численность или плотность определенного вида, устанавливают тип его распределения в пространстве.

5.1.1 Индекс Одума

Простейший критерий оценки пространственного распределения особей индекс Одума базируется на отношении выборочной дисперсии (S_x^2) к ее среднему (\bar{x}):

$$I_o = \frac{S_x^2}{\bar{x}}, \quad (27)$$

где S_x^2 – дисперсия, а \bar{x} – среднее значение.

Равенство $I_o = 1$ соответствует случайному размещению организмов в пространстве, что свидетельствует об однородности среды обитания и отсутствии сильных биотических взаимодействий между ними. Это предположение принимается в качестве нулевой гипотезы. Расселение организмов по биотопу признается равномерным в случае $I_o < 1$, что влечет вывод о сильной конкуренции между особями за жизненно необходимые ресурсы. Статистически значимое неравенство $I_o > 1$ указывает на агрегированные размещения организмов в пространстве, вызванные неоднородностями биотопа и/или сильными положительными взаимодействиями между особями, приводящие к образованию стай или колоний.

Однако в природе существуют псевдоагрегированные размещения организмов, вызванные случайными процессами расселения организмов по биотопу (Василевич, 1969), что затрудняет выявление биотических и абиотических факторов, определяющих тип распределения особей.

Уровень значимости индекса Одума можно оценить, сравнивая рассчитанное значение I_o с табличным значением F-критерия Фишера-Снедекора с числом степеней свободы $df_1 = df_2 = n - 1$.

В тех случаях, когда $I_o < 1$, с табличным значением F-критерия Фишера-Снедекора сравнивается обратная величина: \bar{x} / S_x^2 . Если рассчитанное значение превышает табличное, то нулевая гипотеза о случайном распределении опровергается.

Более содержательный анализ пространственной структуры популяции основан на проверке согласованности эмпирических распределений численностей организмов с известными вероятностными моделями. Положительное биномиальное распределение соответствует равномерному размещению, Пуассона – случайному, отрицательное биномиальное – агре-

гированному и модель Неймана (Neuman, 1939) – псевдоагрегированному.

Индекс Одума характеризует распределение особей в каждый момент сбора, поэтому может быть использован для анализа изменения пространственной структуры популяции во времени.

Пример 1. С помощью индекса Одума определите тип распределения пиявок в 35 пробах.

1	2	1	1	1	3	2	1	1
3	2	2	2	2	2	0	1	2
0	1	1	1	0	0	1	2	1
1	2	0	0	2	6	2	1	

Решение. С помощью таблиц EXCEL определим среднее число пиявок на одну пробу и дисперсию распределения пиявок в пробах: $\bar{x} = 1,4$; $S_x^2 = 1,31$; $\text{Io} = 1,09$. Проверим статистическую значимость индекса Одума.

По таблице F-критерия Фишера-Снедекора (приложение Б) или с помощью функции ФРАСПРОБР (вероятность; число степеней свободы 1; число степеней свободы 2) найдем значение F-критерия Фишера-Снедекора с числом степеней свободы $f_1 = f_2 = 35 - 1 = 34$. $F_{\text{крит.}}(\alpha; 34; 34) = 1,77$.

Так как расчетное значение (1,03) меньше $F_{\text{крит.}}(1,77)$, то на принятом уровне значимости нулевая гипотеза остается в силе, т. е. распределение носит случайный характер.

Контрольные вопросы

1. Дать понятие индексу Одума.
2. Для чего используется в экологических исследованиях индекс Одума?

РАЗДЕЛ 2. СТРУКТУРНЫЙ АНАЛИЗ ЭКОЛОГИЧЕСКИХ СИСТЕМ

Тема 6. Анализ видového разнообразия

6.1 Индексы разнообразия

Альфа-разнообразие – это разнообразие внутри местообитания или одного сообщества. Видовое разнообразие оценивают: по отдельным таксономическим группам (по семействам растений, насекомых, птиц, млекопитающих); по трофическим группам (фотосинтезирующим растениям, растительноядным видам, хищникам и т. д.); по пространственным группам (число видов по ярусам леса, распределение бентоса по поверхности дна и в слое грунта и т. д.); по размерным группам (микро-, мезо- и макрофауна почв или илов) и т. д.

При сравнении различных биоценозов по разнообразию необходимо, чтобы размерность видového пространства сравниваемых комплексов была примерно одинаковой. Поэтому следует учитывать только 10–15 преобладающих видов, составляющих приблизительно 90 % общей численности или биомассы, так как они определяют структуру и продуктивность сообщества.

Рассмотрим наиболее популярные индексы при оценке α -разнообразия.

1. Индекс Маргалефа. Рассчитываются по формуле:

$$D_{Mg} = (S - 1) / \ln N, \quad (28)$$

где N – общее число особей всех видов,

S – число выявленных видов.

Он оценивает видовое богатство и принимает максимальное значение, если все особи принадлежат к разным ви-

дам ($S = N$) и равен 0, когда все особи принадлежат к одному виду ($S = 1$). Достоинство этого индекса – легкость расчетов.

2. Индекс доминирования Симпсона:

$$D = \sum_{i=1}^S p_i^2, \quad (29)$$

где p_i – относительное обилие i -го вида.

Величина индекса варьирует от 0 до 1. Чем больше D , тем выше доминирование одного или нескольких видов.

3. Индекс полидоминантности Симпсона:

$$S_\lambda = 1 / \sum_{i=1}^S p_i^2. \quad (30)$$

Чем ближе S_λ к числу видов S , тем выше разнообразие.

4. Индекс Шеннона:

$$H = - \sum_{i=1}^k (p_i \cdot \ln(p_i)), \quad (31)$$

где p_i – либо удельная численность (частное от деления численности i -го вида на общую численность всех видов), либо удельная биомасса (частное от деления его биомассы на общую биомассу всех видов в пробе).

Индекс Шеннона обычно варьирует в пределах от 1,5 до 3,5, редко превышая 4,5.

Он наиболее популярен при сравнении видовых комплексов по разнообразию.

5. Индекс выравнивания Пиелу. На основе индекса Шеннона можно вычислить показатель **выравнивания Е** (отношение наблюдаемого разнообразия к максимальному):

$$E = \frac{H}{\ln S} \quad (32)$$

$E \in [0, 1]$, причем $E = 1$ при равном обилии видов.

Он показывает степень равномерного распределения видов по их количеству.

6.2 Т-критерий сравнение индексов Шеннона

Для сравнения индексов Шеннона применяется критерий Стьюдента, который рассчитывается по формуле:

$$t_{\text{набл}} = \frac{|H_1 - H_2|}{\sqrt{|\text{var } H_1 - \text{var } H_2|}}, \quad (33)$$

где $\text{var } H_1$ и $\text{var } H_2$ – дисперсии H , получаемые по формуле:

$$\text{var } H = \frac{\sum p_i (\ln p_i)^2 - (\sum p_i \ln p_i)^2}{N} + \frac{S-1}{2N^2}. \quad (34)$$

Число степеней свободы определяется по формуле:

$$df = \frac{(\text{var } H_1 + \text{var } H_2)^2}{\frac{(\text{var } H_1)^2}{N_1} + \frac{(\text{var } H_2)^2}{N_2}}, \quad (35)$$

где H_1 и H_2 – индексы Шеннона, рассчитанные для 1-го и 2-го видового комплекса; $\text{var } H_1$ и $\text{var } H_2$ – их дисперсии; N_1 и N_2 – число особей (групп в выборках).

Пример 1. Для следующего сообщества насекомых рассчитать индексы видового богатства Маргалефа, Симпсона, Шеннона и выравненности Пиелу. Интерпретировать результаты.

№ п/п	Вид	Число особей
1	Стрекоза-коромысло	1
2	Кузнечик зеленый	25
3	Улитка-янтарка	6
4	Клоп травяной	18
5	Тля гороховая	112
6	Долгоносик клеверный	42
7	Червь дождевой	58
Сумма (N)		262

Решение.

1. Индекс Маргалефа:

$$D_{Mg} = (S - 1) / \ln N = \frac{7 - 1}{\ln(262)} = 1,0775$$

2. Расчет индексов Шеннона и Симпсона произведем с помощью пакета прикладных программ MS EXCEL, построив вспомогательную таблицу:

№ п/п	Вид	Число особей	P_i	P_i^2	$P_i \ln P_i$
1	Стрекоза-коромысло	1	0,0038	0,0000	-0,0213
2	Кузнечик зеленый	25	0,0954	0,0091	-0,2242
3	Улитка-янтарка	6	0,0229	0,0005	-0,0865
4	Клоп травяной	18	0,0687	0,0047	-0,1840
5	Тля гороховая	112	0,4275	0,1827	-0,3633
6	Долгоносик клеверный	42	0,1603	0,0257	-0,2935
7	Червь дождевой	58	0,2214	0,0490	-0,3338
	Сумма	262	1,0000	0,2718	-1,5065
Количество особей					
		N	262,0000		
Количество видов					
		S	7,0000		
Видовое богатство					
		M_G	1,0775		
Доминирование					
		D	0,2718		
Индекс полидоминантности					
		S_λ	3,6791		
Индекс Шеннона					
		H	1,5065		
Выравненность					
		E	0,7742		

Полученные расчеты позволяют сделать вывод о том, что видовое богатство рассматриваемого сообщества близко к среднему, индекс полидоминантности указывает на то, что видовое разнообразие невелико, виды в сообществе распределены относительно равномерно.

Пример 2. Сравните по разнообразию древесные растения двух парков города Краснодара.

1-й парк		2-й парк	
Вид	Число	Вид	Число
Клен	5	Береза	150
Тополь	10	Дуб	25
Береза	20	Орешник	30
Дуб	15	Ель	28
Орешник	20	Клен	11
Ель	11	Черемуха	4
Черемуха	25	Рябина	10
Ива	7	Ива	20
Сосна	10	Тополь сер.	20
Лиственница	15	Тополь	5
Рябина	4		
Сирень	6		
Акация	20		
Осина	10		

Решение.

1. С помощью электронных таблиц MS EXCEL рассчитаем индексы Шеннона для обоих видовых комплексов (рисунок 14).

2. Рассчитаем дисперсии индексов Шеннона по формуле (34).

3. Рассчитаем наблюдаемое значение t-критерия по формуле (33).

4. Рассчитаем число степеней свободы для t-критерия Стьюдента по формуле (35).

5. Примем уровень значимости $\alpha = 0,05$ и рассчитаем по таблице распределения Стьюдента (приложение А) критическое значение критерия.

6. Сравним $t_{\text{набл}}$ и $t_{\text{крит}}$.

7. Сделаем вывод о сходстве или различии видовых комплексов по видовому разнообразию.

Вывод: Так как $t_{\text{набл}} > t_{\text{крит}}$, видовые комплексы отличаются по разнообразию на 5%-м уровне значимости. Парк 1-й разнообразнее парка 2-го.

	А	В	С	Д	Е	Г	Н	І	Ј	К	Л	
1	Сравнение видовых комплексов по разнообразию											
2												
3	N п/п	Вид	Число	P_i	$P_i \ln P_i$	$P_i / (\ln P_i)^2$	N п/п	Вид	Число	P_i	$P_i \ln P_i$	$P_i / (\ln P_i)^2$
4	1	Клен	5	0,02809	-0,100347	0,358473	1	Береза	150	0,49505	-0,34807	0,244726
5	2	Тополь	10	0,05618	-0,161753	0,465718	2	Дуб	25	0,082508	-0,20585	0,513557
6	3	Береза	20	0,11236	-0,245624	0,536946	3	Орешник	30	0,09901	-0,22896	0,529487
7	4	Дуб	15	0,08427	-0,208461	0,515676	4	Ель	28	0,082409	-0,22008	0,524115
8	5	Орешник	20	0,11236	-0,245624	0,536946	5	Клен	11	0,036304	-0,12038	0,39915
9	6	Ель	11	0,061798	-0,172038	0,478935	6	Черемуха	4	0,013201	-0,05713	0,247217
10	7	Черемуха	25	0,140449	-0,275689	0,541153	7	Рябина	10	0,033003	-0,11258	0,384024
11	8	Ива серебри	7	0,039326	-0,127253	0,411776	8	Ива я	20	0,066007	-0,17941	0,487626
12	9	Сосна	10	0,05618	-0,161753	0,465718	9	Тополь сер.	20	0,066007	-0,17941	0,487626
13	10	Лиственница	15	0,08427	-0,208461	0,515676	10	Тополь	5	0,016502	-0,06773	0,277974
14	11	Рябина	4	0,022472	-0,085292	0,323724	суммы		303		-1,71958	4,095503
15	12	Сирень	6	0,033708	-0,11427	0,38738						
16	14	Акация	20	0,11236	-0,245624	0,536946						
17	15	Осина	10	0,05618	-0,161753	0,465718						
18	суммы		178		-2,513941	6,540786						
19												
20	N_1	178	df	465,76336	N_2	303						
21	H_1	2,51394066	$t_{\text{набл}}$	16,758865	H_2	1,71957705						
22	VAR_{H_1}	0,00146188	$t_{\text{крит}}$	1,9650786	VAR_{H_2}	0,0037086						

Рисунок 14 – Расчетная таблица для сравнения видовых комплексов по разнообразию

Контрольные вопросы

1. Какие вы знаете индексы видового разнообразия?
2. Дать понятие индексу Маргалефа.
3. Дать понятие индексу доминирования Симпсона и по-лидоминантности Симпсона.
4. Дать понятие индексу Шеннона.
5. Дать понятие индексу выравненности Пиелу.

Задания к самостоятельной работе

1. Количество деревьев в парке распределилось следующим образом:

№ п/п	Вид	Кол-во особей
1	Конский каштан	36
2	Ель колючая	8
3	Кипарисовик Лавсона	3
4	Клен татарский	14
5	Катальпа яйцевидная	2
6	Тополь пирамидальный	19
7	Ясень высокий	10
8	Липа сердцевидная	24

Рассчитать индексы видового богатства Маргалефа, индексы Шеннона и Симпсона, а также индекс выравненности Пиелу.

2. Получены следующие данные по видовому богатству насекомых в сообществе.

№ п/п	Вид	Число особей
1	Стрекоза-коромысло	3
2	Кузнечик зеленый	25
3	Клоп травяной	20
4	Тля гороховая	120
5	Долгоносик клеверный	42

Рассчитайте индекс видового богатства Маргалефа.

Оцените разнообразие сообщества при помощи индекса Шеннона и рассчитайте индекс выровненности Шеннона.

2. Получены следующие данные по видовому богатству птиц в дубовом лесу.

№ п/п	Вид	Число гнездовых территорий
1	Зяблик	15
2	Зарянка	12
3	Лазоревка	10
4	Королек	5
5	Крапивник	3
6	Московка	3
7	Серая мухоловка	1

Рассчитайте индекс видового богатства Маргалефа.

Оцените разнообразие птиц в сообществе с помощью индекса Шеннона и рассчитайте индекс выровненности Шеннона.

3. Рассчитать индексы видового богатства Маргалефа, индексы Шеннона и Симпсона, а также индекс выравненности Пиелу для древесной растительности городского сквера по следующим данным:

Название вида	Количество особей
1. Вяз обыкновенный	412
2. Липа мелколистная	328
3. Клен ясеневидный	129
4. Береза бородавчатая	54
5. Тополь черный	3
5 видов	число особей N = 926

РАЗДЕЛ 3. ОЦЕНКА ВЗАИМОСВЯЗИ В СИСТЕМЕ

Тема 7. Корреляционный анализ

7.1. Сущность и задачи корреляционного анализа

Интерес экологов к потенциалу корреляционного анализа обусловлено тем, что он позволяет:

- 1) получить за короткое время обобщение большого количества данных;
- 2) осуществить достоверное прогнозирование экологических процессов и явлений;
- 3) предсказать возможные значения одного показателя, зная величину другого;
- 4) оптимизировать процесс режимных наблюдений и минимизировать количество определяемых компонентов.

Например, по паре кальций-магний коэффициент корреляции составил 0,84. Следовательно, определив только кальций всегда можно найти содержание магния по пересчетному коэффициенту.

В качестве переменных могут быть данные наблюдений, экспериментов, тестирований.

Задача корреляционного анализа состоит в том, чтобы определить характер связи между сопряженными признаками и убедиться в статистической достоверности найденного количественного значения связи.

Предварительный анализ структуры связи осуществляется с помощью графического анализа. Если попарно нанести на координатную плоскость точки, соответствующие значениям x_i и y_i , то они займут определенную область, которая называется корреляционным полем.

На основании анализа корреляционного поля легко решить вопрос о наличии или отсутствии связи, проследить ее характер (линейная, нелинейная, функциональная или стохастическая) и ее тенденцию (положительная, отрицательная).

В случае наличия корреляционной связи разброс точек (корреляционное поле) имеет вид более или менее правильно-го эллипса со сгущением точек вокруг главной оси и сравнительно редким их расположением на периферии (рисунок 15, а). Если связь слабо выражена, то разброс точек носит случайный характер (рисунок 15, б).

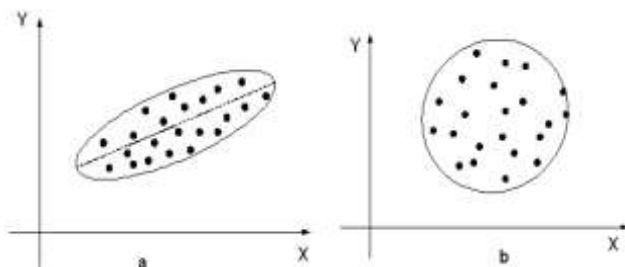


Рисунок 15 – Корреляционное поле

7.2 Коэффициент корреляции Пирсона

Когда сравниваются два признака, которые выражены количественно и распределены по нормальному закону, наличие **линейной** связи между ними можно оценить с помощью **коэффициента корреляции Пирсона**. Это параметрический показатель, для вычисления которого рассчитывают средние и стандартные отклонения результатов двух измерений. Коэффициент корреляции может принимать значения от -1 до $+1$. Знак « $-$ » означает, что связь обратная, « $+$ » – прямая. Чем ближе коэффициент к $|1|$, тем теснее линейная связь.

Если большие значения из одного набора данных связаны с большими значениями другого набора, то связь называется положительной линейной связью (положительная корреляция). Если малые значения одного набора связаны с большими значениями другого – отрицательная линейная связь (отрицательная корреляция). Если данные двух диапазонов

никак не связаны – отсутствие линейной связи (нулевая корреляция).

Качественная оценка коэффициента корреляции представлена в таблице 3.

Таблица 3 – Качественная оценка коэффициента корреляции

Коэффициент корреляции	Теснота связи
±0,91–1,00	Очень сильная
±0,81–0,90	Весьма сильная
±0,65–0,80	Сильная
±0,45–0,64	Умеренная
±0,25–0,44	Слабая
до ±0,25	Очень слабая

Формула для расчета r :

$$r_{xy} = \frac{\overline{xy} - \bar{x} \cdot \bar{y}}{S_x \cdot S_y}, \quad (36)$$

где S_x и S_y – среднеквадратические отклонения соответственно по переменным x и y , рассчитываемые по формулам:

$$S_x = \sqrt{\frac{\sum x^2}{n} - (\bar{x})^2}, \quad S_y = \sqrt{\frac{\sum y^2}{n} - (\bar{y})^2}, \quad (37)$$

а \bar{x}, \bar{y} – средние значения по признаку x и y .

Достоверность коэффициента корреляции проверяется по критерию Стьюдента.

При $n > 100$ формула для расчета t -наблюдаемого следующая:

$$t_{набл} = \frac{|r|\sqrt{n}}{\sqrt{1-r^2}}. \quad (38)$$

При $n \leq 100$ t -наблюдаемое рассчитывается по формуле:

$$t_{\text{набл}} = \frac{|r| \cdot \sqrt{n-2}}{\sqrt{1-r^2}}. \quad (39)$$

Критическое значение t -критерия Стьюдента рассчитывается по таблице или с помощью встроенной функции СТЬЮДРАСПОБР в соответствии с принятым уровнем значимости α и числом степеней свободы $f = n - 2$: $t_{\text{крит}}(\alpha, f = n - 2)$.

При $t_{\text{набл}} > t_{\text{крит}}$ нулевая гипотеза о равенстве нулю коэффициента корреляции между изучаемыми признаками в генеральной совокупности отвергается, и r статистически значим на принятом уровне значимости.

Рассмотрим пример применения корреляционного анализа в биоиндикационных исследованиях.

Пример 1. Чтобы знать, какой орган одуванчика лекарственного лучше использовать для индикации тяжелых металлов, необходимо выяснить, в каких частях растения они в большей степени аккумулируются. Для этого проводится исследование, в результате которого выясняется связь между содержанием тяжелого металла в листьях и в корнях.

Результаты сбора эмпирического материала приведены в первых двух столбцах рисунка 16.

Перед расчетом коэффициента корреляции делается графический анализ взаимосвязи, т. е. строится корреляционное поле (рисунок 17). Это проще всего сделать с помощью мастера диаграмм Ms Excel, выбрав точечную диаграмму.

содержание Pb					
станция	корни (X)	листья (Y)	XY	X ²	Y ²
1	9,47	9,49	89,8703	89,6809	90,0601
2	2,62	4,8	12,576	6,8644	23,04
3	3,49	6,12	21,3588	12,1801	37,4544
4	8,09	9,82	79,4438	65,4481	96,4324
5	9,38	9,8	91,924	87,9844	96,04
6	2,54	4,18	10,6172	6,4516	17,4724
7	7,77	17,05	132,4785	60,3729	290,7025
8	2,55	8,12	20,706	6,5025	65,9344
9	5,24	13,02	68,2248	27,4576	169,5204
10	11,4	12,1	137,94	129,96	146,41
11	2,44	3,76	9,1744	5,9536	14,1376
12	2,39	5,41	12,9299	5,7121	29,2681
13	13,01	18,1	235,481	169,2601	327,61
14	3,12	16,2	50,544	9,7344	262,44
15	5,91	12,94	76,4754	34,9281	167,4436
16	9,63	13,59	130,8717	92,7369	184,6881
17	4,97	19,9	98,903	24,7009	396,01
18	11,49	12,17	139,8333	132,0201	148,1089
19	3,98	6,4	25,472	15,8404	40,96
20	1,39	9,16	12,7324	1,9321	83,9056
21	4,14	5,96	24,6744	17,1396	35,5216
22	0,92	3,72	3,4224	0,8464	13,8384
23	3,56	3,74	13,3144	12,6736	13,9876
24	1,27	3,51	4,4577	1,6129	12,3201
25	1,84	4,81	8,4824	3,3856	21,2521
26	6,69	7,42	49,6398	44,7561	55,0564
27	3,27	3,97	12,9819	10,6929	15,7609
28	2,22	2,86	6,3492	4,9284	8,1796
29	5,47	16,7	91,349	29,9209	278,89
30	6,93	9,87	68,3991	48,0249	97,4169
31	8,57	8,81	75,5017	73,4449	77,6161
32	0,76	3	2,28	0,5776	9
33	0,44	1,5	0,66	0,1936	2,25
34	1,2	1,8	2,16	1,44	3,24
35	0,5	0,7	0,35	0,25	0,49
36	1,2	3,9	4,68	1,44	15,21
37	0,53	3,8	2,014	0,2809	14,44
38	0,2	3,6	0,72	0,04	12,96
39	0,13	3,3	0,429	0,0169	10,89
40	0,26	2,6	0,676	0,0676	6,76
суммы	161,51	298,01	1830,098	1237,454	3392,718
средние	4,03775	7,45025	45,75244	30,93635	84,81796
стандартное отклонение x			3,825301		
стандартное отклонение y			5,414031		
коэффициент корреляции			0,756638		

Рисунок 16 – Расчет коэффициента корреляции между содержанием свинца в листьях и корнях одуванчика лекарственного

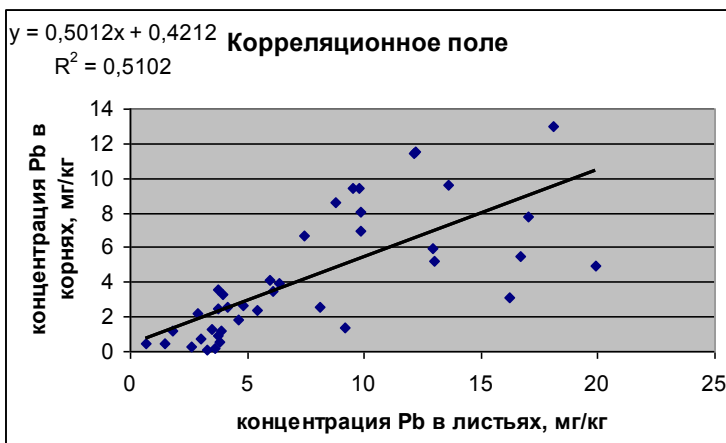


Рисунок 17 – Графический анализ корреляционной взаимосвязи между концентрацией свинца в корнях и листьях одуванчика

Анализ корреляционного поля указывает на линейную и прямую зависимость (точки концентрируются вокруг прямой линии и располагаются на плоскости ХОУ слева направо).

В соответствии с формулой (36) рассчитываем коэффициент корреляции, дополнив таблицу столбцами для расчета необходимых промежуточных величин (рисунок 16).

В результате получилось, что коэффициент корреляции между концентрацией свинца в разных частях одуванчика равен 0,75, что говорит о том, что связь между переменными (листья и корни) существует и является сильной.

Проверим статистическую значимость коэффициента.

$$t_{набл} = \frac{|r| \cdot \sqrt{n-2}}{\sqrt{1-r^2}} = \frac{0,75 \cdot \sqrt{40-2}}{\sqrt{1-0,75^2}} = 6,99.$$

$$t_{крит}(\alpha, f = n - 2) = t(0,05; 38) = 2,02.$$

Так как $t_{набл} > t_{крит}$, коэффициент корреляции статистически значим на 5%-м уровне значимости.

Мы рассмотрели пример расчета коэффициента корреляции «ручным способом», т. е. так, как это можно было бы рассчитывать на калькуляторе. Кроме того, мы должны знать, что в случае сгруппированных данных, техника расчета коэффициента корреляции несколько иная. В этом случае строятся корреляционные решетки (см. Л. Ф. Лакин, с. 190).

EXCEL имеет набор специальных функций, которые при определенных условиях более удобные в использовании, чем примененный нами способ.

Рассчитать коэффициент корреляции Пирсона между двумя массивами (X и Y) можно при помощи функции MS EXCEL ПИРСОН (массив 1; массив 2) или функции КОРРЕЛ (массив 1; массив 2), где *массив 1* – ссылка на диапазон ячеек первой выборки (X); *массив 2* – ссылка на диапазон ячеек второй выборки (Y).

7.3 Множественная корреляция

При большом числе наблюдений, когда коэффициенты корреляции необходимо последовательно вычислять для нескольких выборок, получаемые коэффициенты для удобства сводят в таблицы, называемые **корреляционными матрицами**.

Корреляционная матрица – это квадратная таблица, в которой на пересечении строки и столбца находится коэффициент корреляции между соответствующими параметрами.

В MS EXCEL для вычисления корреляционных матриц используется процедура **Корреляция** из пакета **Анализ данных**. Процедура позволяет получить корреляционную матрицу, содержащую коэффициенты корреляции между различными параметрами.

Для реализации процедуры необходимо:

- 1) выполнить команду **Сервис - Анализ данных**;
- 2) в появившемся списке **Инструменты анализа** выбрать строку **Корреляция** и нажать кнопку **ОК**;

3) в появившемся диалоговом окне указать **Входной интервал**, то есть ввести ссылку на ячейки, содержащие анализируемые данные. Входной интервал должен содержать не менее двух столбцов;

4) в разделе **Группировка** переключатель установить в соответствии с введенными данными (по столбцам или по строкам);

5) указать **выходной интервал**, то есть ввести ссылку на ячейку, с которой будут показаны результаты анализа. Размер выходного диапазона будет определен автоматически, и на экран будет выведено сообщение в случае возможного наложения выходного диапазона на исходные данные. Нажать кнопку **ОК**.

В выходной диапазон будет выведена корреляционная матрица, в которой на пересечении каждой строки и столбца находится коэффициент корреляции между соответствующими параметрами. Ячейки выходного диапазона, имеющие совпадающие координаты строк и столбцов, содержат значение 1, так как каждый столбец во входном диапазоне полностью коррелирует сам с собой.

Рассматривается отдельно каждый коэффициент корреляции между соответствующими параметрами.

Пример 2. Постройте корреляционную матрицу и оцените взаимосвязь между урожайностью (y), мощностью гумусового слоя (x_1) и количеством сложного состава минерального удобрения (x_2). Эмпирические данные представлены в следующей таблице:

№ п/п	1	2	3	4	5	6	7	8	9
x_1	1,00	1,00	1,00	2,00	2,00	2,00	2,50	2,50	2,50
x_2	0,20	0,20	0,20	0,40	0,40	0,40	0,30	0,30	0,30
y	18,20	18,60	18,70	21,60	23,40	23,70	22,00	23,00	22,50

Решение. Придерживаясь алгоритма, изложенного в п. 7.3, построим следующую корреляционную матрицу:

		Строка 2	Строка 3
Строка 1	1		
Строка 2	0,654654	1	
Строка 3	0,875964	0,865801827	1

Отсюда видно, что между урожайностью (Строка 3) и мощностью гумусового слоя (Строка 1) наблюдается высокая корреляционная связь (0,8759); чуть меньше (0,8658) взаимосвязь между урожайностью и количеством минерального удобрения (Строка 2). Умеренная связь между содержанием гумуса и количеством внесенных удобрений.

7.4 Ранговая корреляция

Если переменные принадлежат к порядковой шкале или не подчиняются нормальному распределению, вместо коэффициента Пирсона рассчитывается ранговая корреляция по Спирмэну. Это непараметрический показатель, с помощью которого пытаются выявить связь между рангами соответственных величин в двух рядах измерений.

Этот коэффициент рассчитывать проще, однако результаты получаются менее точными, чем при использовании r .

Рассмотрим конкретный пример. Предположим, мы имеем n описаний, расположенных по обилию 1-го вида от наименьшего обилия к наибольшему. Обозначим место в ряду каждого описания (его ранг) через x_i ($i = 1, 2, 3, \dots, n$).

Аналогично место описания в ряду, ранжированном по обилию 2-го вида обозначим через y_i ($i = 1, 2, 3, \dots, n$). Если между этими двумя видами имеется положительная корреляция, то малому обилию 1-го вида будет соответствовать малое обилие второго вида и наоборот. Оценить такую связь и позволяет коэффициент ранговой корреляции Спирмэна.

Формула коэффициента корреляции Спирмэна выглядит так:

$$r = 1 - \frac{6 \sum_{i=1}^n d_i^2}{n^3 - n}, \quad (40)$$

где d_i – разность рангов по обоим признакам i -го объекта;
 n – число объектов.

Достоверность рангового коэффициента корреляции проверяется по t -критерию:

$$t_{набл} = r_s \sqrt{\frac{n-2}{1-r_s^2}}. \quad (41)$$

Критическое значение критерия для эмпирического коэффициента корреляции Спирмэна для 5% -го уровня значимости можно установить по формуле:

$$t_{кр} = \frac{1,96}{\sqrt{n-1}} \cdot \left(1 - \frac{0,16}{n-1}\right). \quad (42)$$

Нулевая гипотеза отвергается, если $t_{набл} > t_{крит}$.

Пример 3. Рассмотрим пример вычисления рангового коэффициента корреляции. Предположим, что группа городов ранжирована по степени урбанизации и уровню загрязнения окружающей среды. Обозначим города буквами от А до К и выпишем для них следующие две последовательности рангов.

Города	А	Б	В	Г	Д	Е	Ж	З	И	К
Степень урбанизации	3	7	5	9	1	8	6	10	4	2
Уровень загрязненности	2	4	3	5	1	9	8	10	7	6
Разности d_i	1	3	2	4	0	-1	-2	0	-3	-4
Разности d_i^2	1	9	4	10	0	1	4	0	9	16

$$\text{Тогда } r = 1 - 6 \cdot \frac{1+9+4+16+0+1+4+0+9+16}{10^3 - 10} = 0,636.$$

Значение полученного коэффициента указывает на довольно высокую степень согласованности между степенью урбанизации и уровнем загрязненности среды в городах.

Проверим статистическую значимость коэффициента:

$$t_{\text{набл}} = r_s \sqrt{\frac{n-2}{1-r_s^2}} = 0,636 \cdot \sqrt{\frac{10-2}{1-0,636^2}} = 2,33,$$

$$t_{\text{крит}}(\alpha = 0,05; n = 10) = \frac{1,96}{\sqrt{10-1}} \cdot \left(1 - \frac{0,16}{10-1}\right) = 0,64.$$

Так как $t_{\text{набл}} > t_{\text{крит}}$, то на 5%-м уровне значимости можно полагать, что связь между уровнем урбанизации и уровнем загрязненности городов умеренная.

Контрольный вопросы

1. Перечислить основные достоинства корреляционного анализа.
2. Каковы задачи корреляционного анализа?
3. Представить качественную оценку коэффициента корреляции.
4. Дать понятие корреляционной матрице.
5. В каких случаях используется ранговая корреляция?

Задания к самостоятельной работе

1. Установить тесноту связи между содержанием фосфора в почве и содержанием фосфора в растениях по следующим данным:

Почва (x)	1	4	5	9	13	11	23	23	28
Растения (y)	64	21	54	81	93	76	77	95	109

2. Оцените тесноту связи между жизненным состоянием древостоя (балл) и степенью загрязненности территории (балл) по следующим данным:

ЖСП	1	2	1	2	1	4	5	5	3
СЗ	2	4	3	5	1	2	3	4	1

3. Существует ли зависимость между средним уровнем воды (в метрах) в период нереста в Кубанском водохранилище и средней численностью личинок леща (шт/м³) в прибрежной части.

Н, м	51,5	52	52,5	53	53,5	54	54,5	55
Лещ	14,6	15,8	18,6	22,4	24,8	26,4	39,6	34,8

4. Определите тесноту связи между суммой годовых осадков в крае и уровнем загрязнения аллохтонным веществом прибрежных экосистем Кубанского водохранилища, оцененному по БПК, мг/л.

Ос, мм	350	320	230	150	55	130	150	210	310	205	120	105
БПК	5,8	4	5,4	4,2	1,4	3,5	2,1	3,4	4,3	2,4	1,1	1,3

5. Определите тесноту связи между проективными покрытиями двух видов растений, оцененными в баллах:

1-ый вид (балл)	1	2	1	2	1	4	5	5	3
2-ой вид (балл)	2	4	3	5	1	8	10	7	6

6. Даны данные для двух показателей X и Y.

Вычислить коэффициент корреляции. Оценить статистическую значимость коэффициента корреляции.

X	10	12	11	14	16	21	18	22	25	27	25	26
Y	1	3	2	4	6	3	8	9	9	11	12	10

7. Оценить тесноту связи между количеством заболеваний культуры мучнистой росой и количеством осадков. Вариация по обоим признакам выражена в рангах.

X	1	8	5	2	3,5	9	10	7	6	3,5
Y	3	6	10	5	2	7	9	4	1	8

8. Даны данные для двух показателей X и Y.

Вычислить коэффициент корреляции.

Оценить статистическую значимость коэффициента корреляции.

X	12	7	10	5	7	8	11	8	6	9
Y	4,6	2,6	4,3	2,4	3,1	3,8	4,2	2,9	2,7	3,4

9. Оцените взаимосвязь между степенью урбанизации городов и уровнем их загрязненности. Данные выражены в рангах.

Города	А	Б	В	Г	Д	Е	Ж	З	И	К
Степень урбанизации	3	7	5	9	1	8	6	10	4	2
Уровень загрязненности	2	4	3	5	1	9	8	10	7	6

Тема 8. Регрессионный анализ

8.1 Этапы линейного регрессионного анализа

Классический регрессионный анализ достаточно широко используется в экологических исследованиях. С его помощью, например, проводится оценка некоторых факторов среды по растительности (методы биоиндикации) или моделируются зависимости параметров экосистем от окружающей среды.

При одновременном наблюдении над любыми двумя переменными, например, плотностью определенного вида растений и содержанием влаги в почве, результаты оказываются возможным выразить в виде графика, где оси x и y представляют соответственно плотность растений и содержание почвенной влаги. Если экспериментальные точки достаточно хорошо ложатся на кривую, то характер этой кривой можно определить на глаз. Построение кривой регрессии заключается в отыскании такой кривой определенной формы (прямой, параболы и т. д.), на которую лучше всего ложатся экспериментальные точки.

Различают **простую (парную) и множественную регрессию** линейного и нелинейного типа.

Пример простой линейной регрессии: $y = b_1x + b_0$

Пример множественной линейной регрессии:

$$y = b_1x + b_2x_2 + b_3x_3 + b_4x_4 + \dots + b_0. \quad (43)$$

Линейный регрессионный анализ обычно включает четыре стандартных этапа:

1) оценка по заданной выборке $\{x_i; y_i\}$, $i = 1 \dots k$ коэффициентов b_1 и b_0 линейных моделей $y = b_0 + b_1x$ или $y = b_1x$.

Обычно «наилучшая» функция подбирается методом наименьших квадратов (МНК), который предполагает, что сумма квадратов разностей между эмпирическими точками и

точками, рассчитанными по теоретической модели, была бы минимальной, т. е.

$$\sum_{i=1} (y_i - \hat{y}_i)^2 \rightarrow \min ;$$

- 2) оценка остаточной дисперсии, характеризующей качество аппроксимации исходных данных;
- 3) оценка значимости коэффициентов b_1 и b_0 ;
- 4) проверка адекватности модели экспериментальным данным.

8.2 Процедура построения линейной регрессионной модели

Применение линейного регрессионного анализа предполагает, что данные в выборках распределены по нормальному закону и зависимость между переменными прямолинейна.

1. Коэффициенты парной линейной регрессии рассчитываются по формулам:

$$b_1 = \frac{\overline{xy} - \bar{x}\bar{y}}{x^2 - (\bar{x})^2} \text{ или } b_1 = r \frac{S_y}{S_x}; b_0 = \bar{y} - b_1 \cdot \bar{x}. \quad (44)$$

Положительный знак при коэффициенте регрессии b_1 говорит о прямой связи между признаками, знак « \rightarrow » говорит об обратной связи между признаками.

Чтобы рассчитать коэффициенты множественной регрессионной модели пользуются пакетами прикладных программ.

2. Чтобы проверить, насколько хорошо независимая переменная предсказывает зависимую переменную, необходимо произвести дисперсионный анализ, т. е. рассчитать ряд мер вариации:

$SS_{общ} = \sum (y_i - \bar{y})^2$ – сумма квадратов разностей между фактическими значениями y и средним значением \bar{y} , которая называется общей суммой квадратов.

$SS_{рез} = \sum (\hat{y} - \bar{y})^2$ – сумма квадратов разностей между прогнозируемым значением и средним значением \bar{y} , которая называется регрессионная сумма.

$SS_{ост} = \sum (y - \hat{y})^2$ – сумма квадратов разности между прогнозируемым значением \hat{y} и фактическим значением y .

Одной из наиболее эффективных оценок адекватности регрессионной модели, мерой качества подгонки регрессионной модели к наблюдаемым значениям, является *коэффициент детерминации* ($0 \leq R^2 \leq 1$), определяемый по формуле:

$$R^2 = \frac{SS_{рез}}{SS_{общ}}. \quad (45)$$

В случае парной регрессии это квадрат коэффициента корреляции Пирсона. При значении $R^2 > 0,7$ считается, что вариация результативного признака обусловлена в основном влиянием включенных в регрессионную модель факторов.

3. Оценка значимости коэффициентов регрессии производится с помощью t -критерия Стьюдента путем сопоставления их значений с величиной случайной ошибки. В случае парной регрессии они рассчитываются по формулам:

$$t_{b_1} = \frac{b_1}{m_{b_1}}; \quad t_{b_0} = \frac{b_0}{m_{b_0}}; \quad t_{крит}(\alpha; f = n - 2). \quad (46)$$

4. Адекватность построенного уравнения генеральным данным проводится путем проверки статистической значимости коэффициента детерминации R^2 на основе F -критерия Фишера:

$$F_{\text{набл.}} = \frac{R^2(n-m-1)}{(1-R^2)m} > F_{\alpha; f_1; f_2}, \quad (47)$$

где n – размер выборки (количество экспериментов); m – число факторов в уравнении регрессии. Нулевая гипотеза заключается в том, что $R^2 = 0$.

Чтобы найти значение F -критическое, используют функцию ФРАСПОБР с параметром $\alpha = 0,05$ и числами степеней свободы $f_1 = m$ и $f_2 = n - (m + 1)$, где m – это число переменных, а n – число точек данных.

Для парной линейной регрессии

$$F_{\text{факт}} = \frac{r^2(n-2)}{1-r^2}, \text{ а } F_{\text{крит.}}(\alpha = 0,05; f_1 = m; f_2 = n - m - 1). \quad (48)$$

Числа степеней свободы в этом случае рассчитываются по формуле; $f_1 = 1$; $f_2 = n - 2$. Если $F_{\text{набл}} > F_{\text{крит.}}$, то гипотеза о незначимости коэффициента детерминации R^2 отвергается.

5. Качество модели можно также оценить с помощью средней ошибки аппроксимации по формуле:

$$\bar{A} = \frac{1}{n} \sum_{i=1}^n A_i = \frac{1}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right| \cdot 100\%. \quad (49)$$

Допустимый предел значений средней ошибки аппроксимации – не более 8–10 %. В этом случае модель оценивается как достаточно точная, в противном случае говорят о плохом качестве построенной модели.

6. Прогнозное значение Y_p определяется путем подстановки в уравнение регрессии $Y = b_0 + b_1x$ прогнозного значения x_p . Интерпретировать коэффициент b_1 можно следующим образом: увеличение x на одну единицу (в единицах измерения переменной x) приведет к увеличению значения y на b_1 единиц (в единицах измерения переменной y).

8.3 Построение парной линейной регрессионной модели с помощью MS EXCEL

Построим линейную регрессионную модель с помощью MS EXCEL на примере оценки взаимосвязи содержания свинца в листьях и корнях одуванчика (тема 7.2, рисунок 16).

В EXCEL линейный регрессионный анализ осуществляется с помощью функции **ЛИНЕЙН**. После вызова этой функции стандартным способом (**Вставка > Функция**), заполняются поля соответствующей формы (рисунок 18).

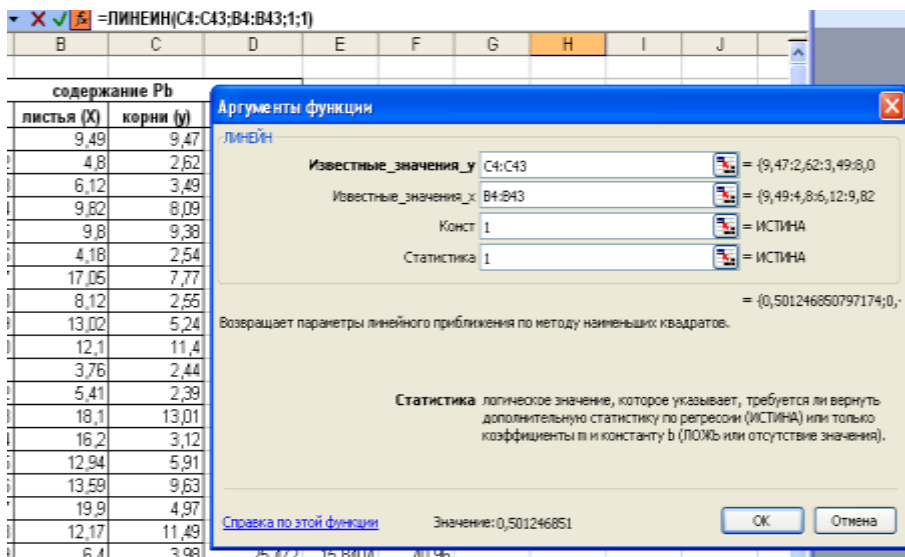


Рисунок 18 – Запуск функции ЛИНЕЙН

В форме ЛИНЕЙН заполняются четыре поля:

- в поле **Изв. знач y** вводится ссылка на массив значений Y – ячейки C4: C43, которые видны слева от формы;
- в поле **Изв. знач x** вводится ссылка на массив значений X – ячейки B4: B43, которые видны в первом столбце таблицы Excel;

– в поле **Константа** вводится логическое значение ИСТИНА (1) или ЛОЖЬ (0). При значении ИСТИНА анализ проводится для модели: $Y = b_0 + b_1X$, а при значении ЛОЖЬ – для модели $Y = b_1X$;

– в поле **Статистика** вводится логическое значение ИСТИНА (1), если надо выводить результаты статистического анализа, и значение ЛОЖЬ (или ничего не вводится), если эти результаты не нужны.

Чтобы вывести результаты расчетов, перед вызовом функции **ЛИНЕЙН** необходимо выделить ячейки, в которые они будут записаны (на рисунке 19 эти ячейки выделены желтым цветом). После заполнения аргументов функции (рисунок 18), необходимо одновременно нажать клавиши **CTRL+SHIFT+ENTER**.

	А	В	С	Д	Е	Ф
1						
2	содержание Pb					
3	станция	листья (X)	корни (y)		РЕЗУЛЬТАТЫ	
4	1	9,49	9,47			
5	2	4,8	2,62		0,501247	0,421165
6	3	6,12	3,49		0,079664	0,733683
7	4	9,82	8,09		0,51024	2,555246
8	5	9,8	9,38		39,58896	38
9	6	4,18	2,54		258,4874	248,1126
10	7	17,05	7,77			
11	8	8,12	2,55			
12	9	13,02	5,24			

Рисунок 19 – Результаты регрессионного анализа зависимости содержания свинца в листьях одуванчика от его содержания в корнях

Расшифровка результатов расчета функции **ЛИНЕЙН** представлена в таблице 4.

Таблица 4 – Результаты регрессионного анализа на основе функции ЛИНЕЙН

$0,5012 - b_1$	$0,4212 - b_0$
$0,0797 - mb_1$ – Ошибка коэффициента b_1	$0,7337 - mb_0$ – Ошибка коэффициента b_0
$0,5102 - r^2$ – коэффициент детерминации	$2,5552 - S_{\bar{y}}$ – Стандартная ошибка для оценки y
$39,5890 - F$ – F-критерий	$38,0000 - f_v$ число степеней свободы
$258,4874 - SS_{рез}$	$248,1126 - SS_{ост}$

Полученные расчеты позволяют представить зависимость содержания свинца в корнях одуванчика от его содержания в листьях в виде следующей линейной функции: $y = 0,5012x + 0,4212$.

Зная эту зависимость, можно прогнозировать содержание свинца в корнях, если известно его содержание в листьях и наоборот.

Используя функцию ФРАСПОБР (0,05; 1; 40 – 2), найдем критическое значение критерия: $F_{крит}(0,05; 1; 38) = 4,098172$. Так как $F_{набл} = 39,58 > F_{кр.} = 4,09$, гипотеза о значимости уравнения регрессии принимается.

Регрессионное уравнение можно получить, построив в EXCEL точечный график по данным и добавив линию тренда.

Чтобы добавить линию тренда, необходимо уже на готовом графике щелкнуть правой кнопкой мыши по любой точке данных и указать в выпавшем окне «добавить линию тренда». Затем необходимо указать тип функции, щелкнуть по закладке «параметры» и поставить «птички» на «показать уравнение на диаграмме» и «поместить на диаграмме величину достоверности аппроксимации». В результате получится диаграмма с уравнением и его оценкой (рисунок 20).

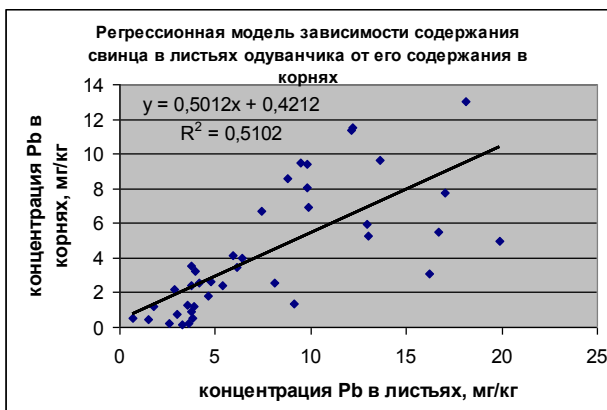


Рисунок 20 – Регрессионная модель зависимости содержания свинца в листьях одуванчика от его содержания в корнях

8.4 Построение множественной линейной регрессионной модели с помощью MS EXCEL

Для построения множественной линейной регрессионной модели необходимо:

1) подготовить список из n строк и m столбцов, содержащий экспериментальные данные (столбец, содержащий выходную величину y должен быть либо первым, либо последним в списке);

2) обратиться к меню ***Данные/Анализ данных/Регрессия***.

Окно «Анализ данных» изображено на рисунке 21.

Если пункт «***Анализ данных***» в меню «***Данные***» отсутствует, то следует обратиться к пункту «***Надстройки***» и установить флажок «***Пакет анализа***».

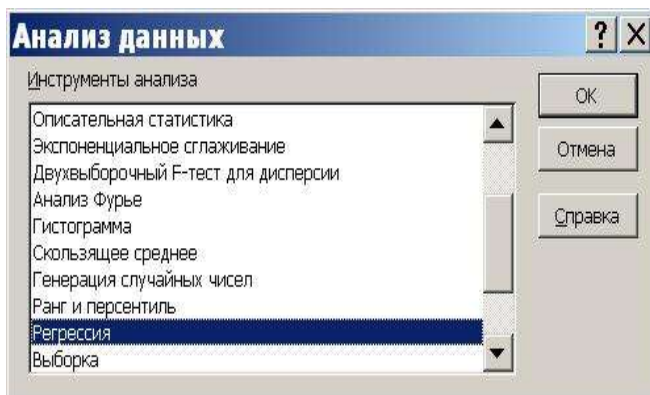


Рисунок 21 – Окно надстройки «Анализ данных»

3) в диалоговом окне «Регрессия» (рисунок 22) задать:

– входной интервал Y;

– входной интервал X;

– выходной интервал – верхняя левая ячейка интервала, в который будут помещаться результаты вычислений (рекомендуется разместить на новом рабочем листе);

4) нажать «ОК» и проанализировать результаты.

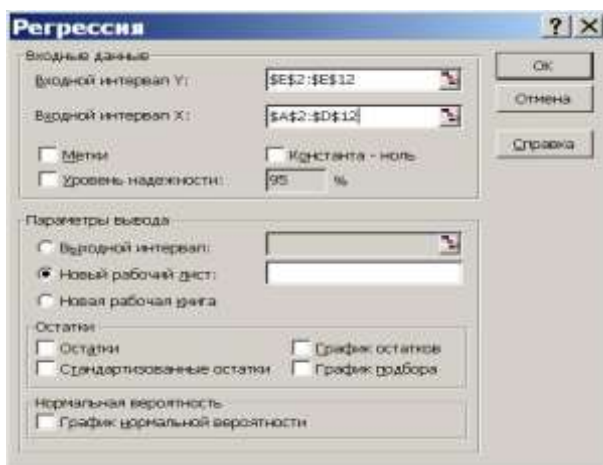


Рисунок 22 – Окно «Регрессия»

Пример 1. По данным о зависимости урожайности (y) от величины гумусового слоя (x_1) и количества минеральных удобрений (x_2) построить множественную регрессионную модель.

Решение. Внесем данные, как показано на рисунке 23. Для выполнения регрессионного анализа введите в диапазон B2:D10 исходные данные. Затем в меню **Данные** выберите пункт **Анализ данных** и далее укажите строку **Регрессия**. В появившемся диалоговом окне укажите **Входной интервал Y:**(D2:D10), **Входной интервал X:** (B2:C10). Укажите выходной диапазон (A11) и нажмите кнопку **ОК**.

В результате получим **ВЫВОД ИТОГОВ**, который содержит три таблицы (рисунок 23).

	A	B	C	D	E	F	G	H	I
1	№ п/п	X₁	X₂	Y					
2	1	1	0,2	18,2					
3	2	1	0,2	18,6					
4	3	1	0,2	18,7					
5	4	2	0,4	21,6					
6	5	2	0,4	23,4					
7	6	2	0,4	23,7					
8	7	2,5	0,3	22					
9	8	2,5	0,3	23					
10	9	2,5	0,3	22,5					
11	ВЫВОД ИТОГОВ								
12									
13	Регрессионная статистика								
14	Множественный R	0,9575394							
15	R-квадрат	0,9168818							
16	Нормированный R-т	0,8891757							
17	Стандартная ошибка	0,7325754							
18	Наблюдения	9							
19									
20	Дисперсионный анализ								
21		<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Значимость F</i>			
22	Регрессия	2	35,52	17,76	33,09317	0,000574234			
23	Остаток	6	3,22	0,53667					
24	Итого	8	38,74						
25									
26	Коэффициенты статистики, P-Значение				Верхние 95%	Нижние 95%	Верхние 95%	Нижние 95%	Средние 95,0%
27	У-пересечение	14,1	0,94575	14,9088	5,73E-06	11,78583133	16,4142	11,7858	16,4142
28	Переменная X1	1,8	0,51801	3,47484	0,013225	0,532477619	3,06752	0,53248	3,06752
29	Переменная X2	13	3,95636	3,26565	0,0167	3,31913791	22,6809	3,31914	22,6809

Рисунок 23 – Результаты множественного регрессионного анализа

В 1-ой таблице значение R^2 (0,9168) указывает на то, что около 92 % дисперсии функции отклика (урожайности) объясняется вариацией линейной комбинации факторов x_1 и x_2 . (содержания гумуса и количеством внесенных удобрений).

Во второй таблице значимость F-критерия (0,00057) указывает на высокую точность аппроксимации исследуемой зависимости линейной зависимостью, так как $p < 0,05$ (т.е. вероятность нулевой гипотезы о равенстве R^2 меньше, чем 0,05).

В третьей таблице указаны коэффициенты полученной регрессионной модели при соответствующих переменных. В нашем случае: $b_1 = 1,8$; $b_2 = 13$; $b_0 = 14,1$.

Таким образом, получено следующее регрессионное уравнение: $y = 14,1 + 1,8 \cdot x_1 + 13x_2$.

Контрольные вопросы

1. Что позволяет оценить регрессионный анализ?
2. Какие стандартные этапы обычно включает линейный регрессионный анализ?
3. Процедура построения множественной линейной регрессионной модели с помощью MS EXCEL.

Задания к самостоятельной работе

1. По данным о зависимости роста и массы животного построить уравнение линейной регрессии. Найти прогнозное значение y при $x = 29$ см. Оценить адекватность модели при помощи средней ошибки аппроксимации.

n	1	2	3	4	5	6	7	8	9	10
X, см	31	32	33	34	35	35	40	41	42	46
Y, кг	7,8	8,3	7,6	9,1	9,6	9,8	11,8	12,1	14,7	13

2. Вывести уравнение для описания зависимости между суммой годовых осадков в крае и уровнем загрязнения аллохтонным веществом прибрежных экосистем Кубанского водохранилища, оцененному по БПК, мг/л и выяснить адекватность полученной модели.

Ос, мм	220	170	64	310	270	210	130	110	55	320	370	220
БПК	3,4	1,3	1,5	4,4	3,3	3,7	1,4	1,3	1,7	4,5	5,6	2,4

3. Существует ли зависимость между количеством выпасаемого на лугу скота и загрязнением находящегося рядом водоема органическим веществом. В случае обнаружения зависимости определить силу связи данных факторов и рассчитать по полученной функции, какая численность стада не приведет к увеличению содержания органического вещества выше ПДК. Оценить адекватность полученной модели. $ПДК_{БПК} = 3$ мг/л.

Кол-во	200	500	1000	1500	1700	1900	2000	2500	2700	3000
мг/л	0,8	0,7	1,4	1,2	1,8	1,9	2,2	3,5	3,2	4,4

4. Получены следующие данные, отображающие влияющие всхожести семян кукурузы на урожай.

Всхожесть x, %	51	73	88	97
Урожай y, ц/га	17,0	26,0	29,7	39,4

Постойте уравнение регрессии, предполагая, что между урожаем и всхожестью имеется линейная связь. Оцените адекватность модели при помощи средней ошибки аппроксимации.

5. Скорость потребления кислорода колорадским жуком при различных температурах среды представлена в таблице.

t	10	15	20	25	30
Q, мкл/г час	88	128	200	285	358

Предполагая, что потребление кислорода жуком возрастает по степенному закону, ($Q = at^b$), рассчитать по этим данным коэффициенты а и b, построить эмпирическую и сглаженную линию регрессии в прямоугольных координатах с помощью средств MS EXCEL.

6. Изучена зависимость скорости питания в мг/экз/сут *Diaphanosoma brachyurum* от биомассы водорослей в среде. Результаты представлены в таблице:

Биомасса X, мг/мл	0,5	1,0	1,5	2,0	3,0	5,0	7,0
У, мг/экз/сут	3	3,5	4,5	4,7	5,3	5,8	5,9

Методом регрессионного анализа показать, как зависит скорость питания бактерий от биомассы водорослей. Построить графики эмпирической и выровненной линии регрессии, используя средства MS EXCEL.

7. Подвергните регрессионному анализу данные по динамике накопления азота в листьях картофеля представленные в следующей таблице. Предварительно постройте графики динамики накопления азота. По результатам анализа сделайте заключение.

Органы растений (листья)	Варианты	Дата взятия проб					
		июль			август		
		4	14	24	4	14	24
	контроль	4,9	18,9	28,5	32,7	16,7	15,3
	опыт	4,6	16,1	31,0	37,7	35,6	28,1

8. В таблице даны среднедекадные температуры почвы на глубине 10 см за 4 летних месяца 2000 г.

Агроклимат. показатель	май			июнь			июль			август		
	I	II	III	I	II	III	I	II	III	I	II	III
t° почвы, 13 часов	1,2	8,0	13,3	18,3	19,1	19,2	20,9	22,5	20,4	19,0	16,0	12,8

Постройте график изменения температуры почвы на глубине 10 см по декадам летних месяцев. Сделайте регрессионный анализ. Найдите число месяца с максимальной температурой почвы.

Тема 9. Дисперсионный анализ

Дисперсионный анализ – это анализ изменчивости признака под влиянием каких-либо контролируемых факторов.

В отличие от корреляционного анализа, в дисперсионном анализе исследователь исходит из предположения, что одни переменные выступают как влияющие (именуемые *факторами* или *независимыми переменными*), а другие (*результативные признаки* или *зависимые переменные*) – подвержены влиянию этих факторов.

Можно, например, изучить состояние водоема в зависимости от способа очистки поступающих в него сточных вод или метода измерения параметров состояния этого водоема. Можно исследовать урожайность какой-либо сельскохозяйственной культуры в зависимости от факторов места, времени года, количества загрязняющих веществ, метеорологических условий и т. д.

Дисперсионный анализ позволяет дать ответ на следующие вопросы:

– относятся ли измерения, проведенные в разных условиях к одной генеральной совокупности, или эти группы измерений принципиально отличаются друг от друга;

– если различия между группами наблюдений существенны, имеется ли возможность агрегировать наблюдения некоторых групп;

– можно ли вычленить долю дисперсии, обусловленную влиянием фактора, на общем фоне случайных флуктуаций.

9.1 Понятие об однофакторном дисперсионном анализе

Мы рассмотрим простейший случай, когда проверяется действие только одного фактора, влияющего на результат эксперимента.

Пример 1. Пусть требуется проверить, насколько влияет и влияет ли вообще на урожай (продуктивность экосистемы) источник загрязнения. В этом случае качество почвы может меняться в зависимости от удаленности от источника загрязнения и давать разную урожайность (продуктивность). Данные для проведения однофакторного дисперсионного анализа представляются следующей таблицей:

Качество почвы	Номер эксперимента				
	1	2	3	4	5
T1	X ₁₁	X ₁₂	X ₁₃	...	X _{1n}
T2	X ₂₁	X ₂₂	X ₂₃	...	X _{2n}
T3	X ₃₁	X ₃₂	X ₃₃	...	X _{3n}
...
m	X _{m1}	X _{m2}	X _{m3}	...	X _{mn}

Строки таблицы соответствуют градациям фактора. В нашем случае – это качество почвы. Столбцы – экспериментам. Причем, m – полное число градаций фактора, а n – число экспериментов.

Обозначим через \bar{x}_1 среднее арифметическое из n урожайностей на почве первой категории, через \bar{x}_2 – среднее урожайностей на почве второй категории и т. д.

Общее среднее арифметическое всех nm измерений обозначим: $\bar{x} = \frac{1}{nm} \sum_{i=1}^m \sum_{k=1}^n x_{ik}$.

Здесь суммирование по k при постоянном i дает сумму по всем наблюдениям i -й серии (т. е. по почве i -й категории). Дальнейшее суммирование по i дает итог по всем категориям почвы.

Математически можно доказать следующее тождество:

$$\sum_{i=1}^m \sum_{k=1}^n (x_{ik} - \bar{x})^2 = n \sum_{i=1}^m (\bar{x}_i - \bar{x})^2 + \sum_{i=1}^m \sum_{k=1}^n (x_{ik} - \bar{x}_i)^2,$$

или в сокращенном виде $\theta_{\text{общ}} = \theta_{\text{межгрупп.}} + \theta_{\text{внутригрупп.}}$

Таким образом, общая сумма квадратов распадается на две составные части, первая из которых связана с оценкой дисперсии между категориями почвы, а вторая – с оценкой дисперсии внутри всех категорий почвы. Компактно процедуру однофакторного дисперсионного анализа можно записать в виде таблицы (таблица 5)

Таблица 5 – Схема однофакторного дисперсионного анализа

Компонента дисперсии	Сумма квадратов	Число степеней свободы	Выборочная дисперсия
Между группами	$n \sum_{i=1}^m (\bar{x}_i - \bar{x})^2$	$m - 1$	S_{x1}^2
Внутри групп	$\sum_{i=1}^m \sum_{k=1}^n (x_{ik} - \bar{x}_i)^2$	$mn - m$	S_{x2}^2
Полная (общая)	$\sum_{i=1}^m \sum_{k=1}^n (x_{ik} - \bar{x})^2$	$mn - 1$	S_x^2

Сравнивая дисперсию урожайности между категориями почвы с дисперсией урожайности «внутри» категорий почвы, по величине их отношения судят, насколько рельефно проявляется влияние такого фактора, как качество почвы. В этом сравнении как раз и заключается основная идея дисперсионного анализа.

Проверку нулевой гипотезы о том, что нормальные распределения всех величин $X_1, X_2, X_3, \dots, X_n$ (урожайности) тождественны, т.е. все mn наблюдений можно рассматривать как выборку из одной и той же нормальной совокупности, производят с помощью F – критерия. Для этого находят отношение

$$F_{набл} = \frac{S_{x1\text{ межгрупповая}}^2}{S_{x2\text{ внутригрупповая}}^2}, \quad (50)$$

где S_{x1}^2 – дисперсия между группами;

S_{x2}^2 – дисперсия внутри групп.

F-критическое определяют по таблице распределения Фишера для принятого уровня значимости α и чисел степеней свободы k_1 (для большей дисперсии) и k_2 (для меньшей дисперсии):

$$F_{крит}(\alpha = 0,05; k_1 = m - 1; k_2 = mn - m). \quad (51)$$

Если $F_{набл} > F_{крит}$ нулевая гипотеза отклоняется, и следует считать, что среди средних значений имеются хотя бы два не равных друг другу.

EXCEL имеет набор специальных функций, которые при определенных условиях более удобные в использовании, чем примененный нами способ.

9.2 Проведение дисперсионного анализа с помощью электронных таблиц EXCEL

В MS EXCEL для проведения однофакторного дисперсионного анализа используется процедура **Однофакторный дисперсионный анализ**.

Для ее реализации необходимо:

1) ввести данные в таблицу, так чтобы в каждой строке (или столбце) оказались данные, соответствующие одному значению исследуемого фактора, а строки (столбцы) располагались в порядке возрастания (убывания) величины исследуемого фактора;

2) выполнить команду **Данные > Анализ данных**;

3) в появившемся диалоговом окне **Анализ данных** в списке **Инструменты анализа** выбрать процедуру **Однофакторный дисперсионный анализ**, затем нажать кнопку **ОК**;

4) в появившемся диалоговом окне задать **Входной интервал**, то есть ввести ссылку на диапазон анализируемых данных, содержащий все столбцы данных;

5) в разделе **Группировка** переключатель установить в положение *по строкам (по столбцам)*;

6) указать **выходной интервал**, то есть ввести ссылку на ячейку, с которой будут показаны результаты анализа. Размер выходного диапазона будет определен автоматически, и на экран будет выведено сообщение в случае возможного наложения выходного диапазона на исходные данные. Нажать кнопку **ОК**.

Выходной диапазон будет включать в себя результаты дисперсионного анализа: средние, дисперсии, критерий Фишера и другие показатели. Влияние исследуемого фактора определяется по величине значимости критерия Фишера, которая находится в таблице **Дисперсионный анализ** на пересечении строки **Между группами** и столбца **Р-значение**. В случаях, когда Р-значение < 0,05, критерий Фишера значим, и влияние исследуемого фактора можно считать доказанным.

Пример 1. Рассмотрим данные пятикратного ($n = 5$) измерения урожайности на трех ($m = 3$) категориях почвы. В таблице 6 приведены данные не фактического, а условного эксперимента. Требуется проверить, влияет ли категория почвы (ее загрязненность) на величину урожая.

Таблица 6 – Результаты измерения урожайности в относительных единицах

Номер категории почвы	Номер эксперимента					Выборочное среднее
	1	2	3	4	5	
i	x_{i1}	x_{i2}	x_{i3}	x_{i4}	x_{i5}	\bar{x}
1	12	15	17	13	16	$\bar{x}_1 = 14,6$
2	20	17	16	25	14	$\bar{x}_2 = 18,4$
3	10	12	11	13	8	$\bar{x}_3 = 10,8$

Из таблицы 6 имеем:

$$\bar{x} = 1/15 \sum_{i=1}^3 \sum_{m=1}^5 x_{im} = 14,6; \theta_1 = 5 \sum_{i=1}^3 (x_i - \bar{x})^2 = 144,4;$$

$$\theta_2 = \sum_{i=1}^3 \sum_{k=1}^5 (x_{ik} - \bar{x}_i)^2 = 105,2; \quad \theta = \sum_{i=1}^3 \sum_{k=1}^5 (x_{ik} - \bar{x})^2 = 249,6.$$

Найдем дисперсии, поделив суммы на соответствующее число степеней свободы.

$$S_{x_1}^2 = \frac{\theta_1}{2} = 72,2; \quad S_{x_2}^2 = \frac{\theta_2}{12} = 8,76; \quad S_x^2 = \frac{\theta}{14} = 17,82.$$

Заполним следующую таблицу дисперсионного анализа урожайности.

Таблица 7 – Дисперсионный анализ урожайности на почвах разного качества

Компонента дисперсии	Суммы квадратов	Число степеней свободы	Выборочная дисперсия
Между категориями почв	$\Theta_1 = 144,4$	2	$S_{x_1}^2 = 72,2$
Внутри категорий почв	$\Theta_2 = 105,2$	12	$S_{x_2}^2 = 8,76$
Полная (общая)	$\Theta = 249,6$	14	$S_x^2 = 17,82$

Произведем проверку нулевой гипотезы с помощью F-критерия:

$$F_{набл} = \frac{S_{x_1}^2}{S_{x_2}^2} = \frac{72,2}{8,76} = 8,24.$$

При двух степенях свободы большей дисперсии ($k_1 = 2$) и двенадцати степенях свободы меньшей дисперсии ($k_2 = 12$) по таблице критерия Фишера находим критические границы для F, равные при 5%-м уровне значимости 3,88 и 1%-м – 6,93.

Полученное нами из наблюдений $F_{набл}$ превышает указанные границы, и поэтому нулевая гипотеза должна быть от-

вергнута, т. е. урожайность на рассматриваемых категориях почвы не одинакова, а это значит, что загрязнения влияют на урожайность.

Пример 2. Решим тот же пример на компьютере, используя технологию, изложенную выше (см. п. 9.2).

Таблица 8 – Дисперсионный анализ влияния качества почвы на урожайность с.-х. культуры

ИТОГИ						
<i>Группы</i>	<i>Счет</i>	<i>Сумма</i>	<i>Сред- нее</i>	<i>Дис- персия</i>		
Строка 1	5	73	14,6	4,3		
Строка 2	5	92	18,4	18,3		
Строка 3	5	54	10,8	3,7		
Дисперсионный анализ						
<i>Источник вариации</i>	<i>SS</i>	<i>df</i>	<i>MS</i>	<i>F</i>	<i>P- значение</i>	<i>F крит.</i>
Между группами	144,4	2	72,2	8,2357	0,0056	3,8852
Внутри групп	105,2	12	8,7666			
Итого	249,6	14				

Контрольные вопросы

1. Дать понятие дисперсионному анализу.
2. Чем дисперсионный анализ отличается от корреляционного?
3. Дать понятие однофакторному дисперсионному анализу.
4. Изложите процедуру однофакторного дисперсионного анализа в электронных таблицах MS EXCEL.

Задания к самостоятельной работе

1. При оценке влияния автомобильной магистрали на состояние древесной растительности были проведены исследования по определению площади листовой пластинки у тополя, произрастающего на разном расстоянии от дороги. Влияет ли удаленность от дороги на площадь листовой пластинки тополя (см^2), если получились следующие результаты измерений?

Расстояние от дороги (м)	Повторности				
	1	2	3	4	5
100–500	39	39	37	41	36
500–1000	41	40	39	40	38
1500–2000	42	40	39	42	40
2000–2500	42	45	43	42	40

2. Предположим, что изучается влияние возрастающих доз удобрений определенного типа на урожайность какой-либо культуры. Пусть имеется четыре дозы удобрения (A_1, \dots, A_4 , причем $A_1 < A_2 < A_3 < A_4$), которое использовали на пяти делянках по каждой дозе ($m = 4, n = 5$). Требуется выяснить, влияет ли повышение дозы удобрения на урожайность и если да, то достоверен ли этот вывод настолько, чтобы можно было рекомендовать этот опыт сельскому хозяйству. Результаты наблюдений приведены в таблице:

Доза удобрения	Урожайность ц/га				
	Номера делянок				
	1	2	3	4	5
A_1	150	140	150	145	150
A_2	190	150	170	150	165
A_3	200	170	200	170	180
A_4	230	190	210	190	200

3. В таблице представлены данные содержания подвижных форм марганца, меди и цинка в пахотном слое чернозема выщелоченного при различных вариантах агротехнологии. Методом однофакторного дисперсионного анализа выявить, влияет ли способ обработки на содержание тяжелых металлов в пахотном слое.

Варианты	Способы обработки								
	Марганец (Mn)			Медь (Cu)			Цинк (Zn)		
	Д ₁	Д ₂	Д ₃	Д ₁	Д ₂	Д ₃	Д ₁	Д ₂	Д ₃
000	13,5	17,5	15,4	0,98	0,71	0,69	0,99	1,06	1,21
111	28,0	15,8	15,1	1,42	0,79	1,45	1,63	1,05	1,06
222	18,6	18,6	12,5	1,34	0,88	1,32	1,24	0,76	0,76
333	26,1	25,8	21,6	2,00	1,18	1,40	1,33	1,57	1,30
среднее	21,6	19,4	16,2	1,44	0,89	1,22	1,30	1,20	1,08

4. Решите предыдущую задачу для свинца, кобальта, кадмия.

Варианты	Способы обработки								
	Свинец (Pb)			Кобальт (Co)			Кадмий (Cd)		
	Д ₁	Д ₂	Д ₃	Д ₁	Д ₂	Д ₃	Д ₁	Д ₂	Д ₃
000	1,04	0,91	0,94	0,17	0,15	0,26	0,089	0,071	0,049
111	0,65	1,03	0,53	0,17	0,07	0,13	0,078	0,072	0,082
222	0,76	1,31	0,74	0,08	0,04	0,03	0,065	0,054	0,051
333	0,82	1,66	1,02	0,10	0,07	0,08	0,060	0,076	0,062
среднее	0,82	1,23	0,82	0,13	0,08	0,12	0,068	0,076	0,061

ТЕСТЫ

Тема: *Основы теории систем и системного анализа*

1. Для оценки состояния системы больше информации несет ...

- : биомасса
- : поток энергии
- +: структура

2. Методология системной экологии заключается в ...

- : стремлении изучать живую природу путем перечисления отдельных ее форм
- +: анализе закономерностей функционирования природных образований
- +: использовании принципа целостности природных образований
- : детальном изучении природных образований

3. По своим целям и задачам системная экология наиболее близка к ...

- +: теоретической экологии
- : прикладной экологии
- : экологии микроорганизмов

4. Прикладной аспект системной экологии заключается в ...

- +: количественном анализе фактического материала
- +: компьютерной обработке информации
- : формулировке теоретических законов экологии

5. Системная экология изучает проблемы ...

- : охраны окружающей среды
- : социальной экологии
- : инженерной экологии
- +: познания специфических закономерностей взаимоотношений организмов со средой на уровне особи, популяции, биоценоза

6. Теория систем это ...

- + : наука, изучающая общие свойства сложных систем, методы их исследования, создания и управления
- : наука, изучающая основы мироздания
- : наука, изучающая экологические закономерности

7. Основоположником общей теории систем является ...

- : В. И. Вернадский
- : Н. Винер
- : Ю. Одум
- + : Л. Бергаланфи
- : В. Н. Сукачев

8. Синергетика изучает ...

- + : процессы самоорганизации самой разной природы
- + : пространственные и временные структуры самой разной природы
- : методы обработки экспериментальных данных
- : поведение закрытых систем

9. Устойчивой является экосистема, в которой ...

- + : максимальное количество симбиотических связей между организмами
- + : на единицу имеющегося потока энергии приходится максимальная биомасса
- : минимум информационных потоков
- : минимальное количество симбиотических связей

10. Конечная цель системного анализа заключается в ...

- : установлении связей между элементами системы
- : построении модели системы
- : статистической обработке результатов исследования
- + : решении конкретной проблемы

11. Поведение системы это ...

- + : последовательность принимаемых ею состояний, ведущее к цели системы
- : смена количества элементов в системе
- : смена количества связей в системе
- : стремление к объединению с другими системами

12. Границы системы являются открытыми ...

+: при наличии связей системы с внешней средой

-: при отсутствии связей с внешней средой

-: при обмене информацией между элементами внутри системы

13. По отношению к массе лучшими индикаторными ви-

дами являются ...

+: крупные

-: мелкие

14. Научные разработки, направленные на обследование и изучение систем различного рода называются ...

+: проектами

-: диссертациями

-: курсовыми работами

-: квалификационными работами

-: рефератами

15. Математическое определение системы включает в себя ...

+: множество всех внутренних элементов системы

+: множество систем, находящихся во взаимодействии с данной системой

+: множество всех связей элементов системы между собой, а также с внешней средой

+: закон функционирования экосистемы

-: часть внутренних элементов системы

+: множество моментов времени, для которых измерены мгновенные значения переменных

16. Строение системы характеризуют следующие понятия ...

+: элемент

+: подсистема

+: компоненты

+: внешняя среда

+: связь

-: равновесие

-: устойчивость

17. Функционирование системы характеризуют следующие понятия ...

- + : связь
- + : равновесие
- + : устойчивость
- + : поведение
- + : развитие
- : элемент
- : подсистема
- : компонент

18. Способность системы возвращаться в состояние равновесия после того, как под влиянием внешних воздействий она была выведена из этого состояния называется ...

- + : устойчивостью
- : равновесием
- : поведением

19. Устойчивые закономерные связи между элементами системы, отражающие пространственное и временное расположение элементов и характер их взаимодействия называются ...

- + : структурой системы
- : поведением системы
- : устойчивостью системы
- : закономерностью системы

20. Устойчивость экосистемы – это ...

- + : способность экосистемы восстанавливаться после того, как ее структура и функции были нарушены
- : способность экосистемы противостоять допустимым нагрузкам
- : свойства системы реагировать на изменения в окружающей среде

21. Структура системы характеризует ...

- + : строение системы
- + : совокупность связей в системе
- : количество элементов системы
- : качественные характеристики элементов

22. Внешняя среда это ...

- + : совокупность всех объектов, чьи свойства меняются в результате изменения системы
- + : совокупность всех объектов, изменение свойств которых влияет на систему
- : совокупность всех объектов системы

23. Иерархия – это ...

- + : упорядоченность компонентов по степени важности
- : составная часть системы
- : компонент системы
- : поведение системы

24. К свойствам всех систем относятся следующие характеристики ...

- + : целостность
- + : делимость
- + : разнообразие
- + : связность
- + : иерархичность
- + : целенаправленность
- : деградация
- : эволюция

25. По степени сложности системы делятся на ...

- + : простые
- + : сложные
- + : большие
- : статические
- : динамические

26. По происхождению системы делятся на ...

- + : искусственные
- + : естественные
- + : виртуальные
- + : смешанные
- : иерархические
- : открытые
- : замкнутые

27. Избежать затруднений в моделировании большой системы можно, если ...

+ : разбить ее на ряд подсистем меньшей размерности

- : выделить главные элементы

- : изменить цель моделирования

28. Какие из этих систем относятся к самому высокому уровню сложности?

+ : мозг человека

- : химическая реакция

- : язык (в лингвистике)

- : экологическая система

- : биологическая клетка

- : компьютер

- : автомобиль

- : часы

29. Классификационные характеристики экосистемы «Озеро» ...

- : простая

+ : сложная

+ : динамическая

- : искусственная

+ : иерархическая

- : закрытая

+ : открытая

30. Гомеостаз – это ...

+ : способность системы к саморегулированию при изменении условий окружающей среды

- : способность системы сохранять свой состав при изменении условий окружающей среды

- : свойства системы реагировать на изменения в окружающей среде

31. Что означает состояние гомеостаза экологической системы?

+ : Состояние внутреннего динамического равновесия

- : Состояние активно протекающих процессов сукцессии

- : Состояние деструкции растительных сообществ экосистемы

Тема: *Экосистема – предмет системного анализа*

32. Как следует понимать климаксное состояние экосистемы?

- + : Как состояние динамического равновесия
- : Как состояние деградации в результате эндогенной сукцессии
- : Как состояние резкой изменчивости экосистемы под влиянием внешних факторов
- : Как состояние активного протекания сукцессионных процессов

33. Термин «экосистема» ввел ...

- + : Генсли
- : Вернадский
- : Геккель
- : Сукачев

34. Закон пирамиды энергии утверждает, что с одного трофического уровня экологической пирамиды переходит на другой:

- : не более 1 % энергии
- : более 50 %;
- + : не более 10 %;
- : 99 %.

35. Принципы конкурентного исключения изложил ...

- + : Г.Ф. Гаузе
- : Р. Линдеман
- : В.В. Станчинский
- : Г. Г. Винберг

36. Балансовый подход к изучению роста и развития животных и к исследованиям экосистем развил ...

- + : В.С. Ивлев
- : В.В. Станчинский
- : Р. Линдеман
- : А. Лотки

37. Г. Г. Винберг величину первичной продукции экспериментально оценил с помощью ...

+: метода темных и светлых склянок

-: теоремы косинусов

-: метода проб и ошибок

38. Впервые о функциональной, а не структурной основе экосистем заявил ...

-: В.С. Ивлев

-: В.В. Станчинский

+: Р. Линдеман

-: А. Лотки

39. Что такое экосистема?

-: Ассоциация растительности, занимающая определенное положение в пространстве, отличающаяся от смежных ассоциаций

+: Единый природный комплекс, образованный живыми организмами и средой их обитания, в котором живые и косные компоненты взаимосвязаны обменом вещества, энергии и информации

-: Единый природный комплекс, включающий растительность, почву и подстилающие горные породы

40. Сообщество ... включает в себя автотрофные элементы

-: всегда

+: не всегда

41. Основным энергетическим блоком экосистемы является ...

+: фитоценоз

-: зооценоз

-: микробоценоз

42. Границы биогеоценоза определяются по горизонтали ...

+: границами входящего в него фитоценоза

-: высотой надземных органов растений

-: глубиной проникновения подземных органов растений

-: глубиной проникновения микроорганизмов

43. Границы биогеоценоза определяются по вертикали

...

- : границами входящего в него фитоценоза
- +: высотой надземных органов растений
- +: глубиной проникновения подземных органов растений
- +: глубиной проникновения микроорганизмов

44. Сообщество, находящееся в равновесии с окружающей средой, называется ...

- +: климаксным
- : равновесным
- : развивающимся

45. Различия между сообществом, биоценозом, биогеоценозом заключаются в ...

- +: определении границ этих систем в природе
- : количестве особей
- : количестве взаимосвязей между видами

46. Смена одних биоценозов другими на определенном участке земной поверхности за некоторый период времени называется ...

- +: сукцессией
- : экотоном
- : естественным отбором
- : саморегуляцией экосистемы

47. Биоценоз считается устойчивым, если он сохраняется в течение ...

- +: нескольких десятилетий
- : нескольких лет
- : нескольких месяцев

48. Консорция – это ...

- +: структурная единица биоценоза, объединяющая автотрофные и гетеротрофные организмы на основе пространственных и пищевых связей вокруг центрального ядра.
- : структурная часть в вертикальном расчленении биоценоза, ограниченная в пространстве или во времени

-: структурная часть в горизонтальном расчленении биоценоза, отличающаяся от других частей составом и свойствами компонентов

49. Синузия – это ...

+: структурная часть в вертикальном расчленении биоценоза, ограниченная в пространстве или во времени

-: структурная часть в горизонтальном расчленении биоценоза, отличающаяся от других частей составом и свойствами компонентов

-: территория или акватория, занимаемая популяцией

-: совокупность всех факторов среды, в пределах которых возможно существование вида

50. Парцелла – это ...

+: структурная часть в горизонтальном расчленении биоценоза, отличающаяся от других частей составом и свойствами компонентов.

-: группы видов в сообществе, обладающие сходными функциями и нишами одинакового размера

-: основная единица классификации объектов в видовом пространстве в фитоценологии

-: статистически устойчивая комбинация видов, обитающих приблизительно в одинаковых условиях среды и сбалансированная по своим трофическим связям.

51. Периодом современной теоретической экологии считается ...

-: последние 10 лет XX в.

+: последние тридцать лет XX в.

-: последние 50 лет XX в.

52. Элементарной биохорологической единицей биосферы является ...

-: экосистема

+: биогеоценоз

-: ландшафт

-: биом

-: биогеографическая область

53. Элементарной единицей абиотической составляющей экосистемы считается ...

-: химический элемент

+: фактор среды

-: вода

-: температура воздуха

54. Элементарной единицей биотической составляющей экосистемы считается ...

-: популяция

-: гильдия

+: особь

-: когорта

55. Экологическими эквивалентами называются ...

+: организмы, занимающие одинаковые или сходные экологические ниши в разных географических областях

-: виды живых организмов, проживающих на различных территориях с одинаковыми климатическими условиями

-: организмы, занимающие разные экологические ниши в одинаковых географических областях

56. Популяция – это ...

-: совокупность связанных и взаимозависимых организмов

+: множество особей или групп особей, обладающих свойством самовоспроизведения и способностью совместно адаптивно реагировать на изменение внешней среды

-: различные группы особей, имеющие одну среду обитания

-: совокупность взаимосвязанных и взаимозависимых организмов

57. Правильная последовательность иерархических уровней морфологического строения экосистем

1: особь

2: группа (когорта)

3: популяция

4: трофический уровень

5: экосистема

58. Основная ячейка трансформации энергии в экосистеме, являющаяся ее структурной частью называется

- + : консорция
- : популяция
- : гильдия
- : ассоциация

59. Эмерджентными называют ...

- : свойства биологических объектов противостоять вредным веществам
- : свойства общие для различных иерархических уровней
- + : качественно новые свойства, возникающие по мере объединения компонентов в более сложные функциональные единицы и отсутствующие на предыдущем уровне
- : способность биологических объектов не реагировать на изменения, происходящие в окружающей среде

60. Зона перехода между двумя конкурирующими сообществами называется ...

- : сукцессией
- + : экотоном
- : саморегуляцией экосистемы
- : естественным отбором

61. Тенденция к увеличению разнообразия и плотности живых организмов на границах сообществ называется ...

- : экотоном
- : ассоциацией
- + : краевым эффектом

62. Признаки разграничения экосистем бывают ...

- + : физиономические
- + : таксономические
- + : экологические
- : статистические
- : морфологические

63. Основными количественными показателями для экосистем являются ...

- + : плотность видов

+: плотность организмов

+: плотность биомассы

-: численность видов

-: биомасса видов

64. Объективными интегральными характеристиками трофического уровня являются ...

+: количество энергии на единицу площади

+: количество информации на единицу площади

-: плотность видов

-: плотность организмов

65. Границы экосистемы определяются ...

+: целями исследования

-: материальными ресурсами

-: количеством видов

-: дифференциальными видами

66. Способ разграничения экосистемы по экологическим признакам называется ...

+: градиентным анализом

-: классификацией

-: дискретным анализом

-: экологическим анализом

67. Континуум – это ...

+: непрерывная мозаика популяционных распределений, связанных условиями среды

-: тенденция к увеличению разнообразия и плотности живых организмов на границах сообществ

-: зона перехода между двумя конкурирующими сообществами

68. Таксономические критерии разграничения экосистем базируются на ...

+: численно преобладающих одиночных видах

+: совокупности характерных видов

-: совокупности редких видов

-: совокупности исчезающих видов

69. Первостепенными экологическими проблемами являются ...

- : борьба с различными формами загрязнения биосферы
- +: проблема нормирования допустимого уровня антропогенной нагрузки на конкретные экосистемы.
- +: регуляция численности популяций
- +: управление продукционными процессами
- +: проблема устойчивости природных и антропогенных ценозов

70. Основные методы исследований в экологии включают в себя ...

- +: полевые наблюдения
- +: эксперименты в поле и лаборатории
- +: моделирование
- +: прогноз
- : космический мониторинг

71. Исследования, связанные с вмешательством в состав или структуру экосистемы называются ...

- +: экспериментальными методами
- : моделированием
- : прогнозированием
- : полевыми наблюдениями

72. Эксперименты бывают ...

- +: активные
- +: пассивные
- +: однофакторные
- +: многофакторные
- +: полевые
- +: лабораторные
- : экстремальные
- : теоретические

73. Эксперимент, в котором исследователь выступает как бы в роли наблюдателя, не вмешиваясь в ход эксперимента, называется...

+: пассивным

-: активным

-: многофакторным

74. К какого рода экспериментам относятся задачи экологического мониторинга?

+: к пассивным

-: к активным

75. Правильная последовательность этапов системного анализа ...

1: Выбор проблемы

2: Содержательная постановка задачи и ограничение степени ее сложности.

3: Установление иерархии целей и задач.

4: Выбор путей решения задач.

5: Моделирование.

6: Оценка возможных стратегий.

7: Внедрение результатов.

76. Изоморфизм – это ...

+: взаимно однозначное отображение соответствия между оригиналом и моделью в области изучаемых свойств.

-: отображение части свойств оригинала на модель.

-: любая совокупность связанной информации

77. Гомоморфизм – это ...

-: взаимно однозначное отображение соответствия между оригиналом и моделью в области изучаемых свойств.

+: отображение части свойств оригинала на модель.

-: математическое описание оригинала, отражающее его целостность

78. Правильная последовательность этапов построения модели ...

- 1: формулирование целей моделирования
- 2: качественный анализ экосистемы
- 3: формулировка законов и гипотез относительно структуры экосистемы
- 4: идентификацию модели (определение ее параметров)
- 5: верификацию модели
- 6: исследование модели

79. Любая совокупность связанной информации, объединенной вместе по определенному признаку называется ...

- +: базой данных
- : формулой
- : графической моделью

80. Краевой эффект в экологии – это ...

- : упорядоченность в расположении видов популяций и сообществ по градиенту
- : резкий переход между двумя и более различными сообществами.
- : вытеснение более слабых особей к краям сообщества
- +: тенденция к увеличению разнообразия и плотности органического вещества на границе сообществ

81. Теоретические исследования А. Лотки и В. Вольтера экспериментально подтвердил ...

- +: Гаузе
- : Л.Г.Раменский
- : Дж. Харпер

82. Основными количественными показателями экосистем являются следующие ...

- +: плотность видов
- +: плотность организмов
- +: плотность биомассы
- : численность видов
- : численность организмов

83. Группа видов, которые делят один и тот же ресурс и потому связана отношениями наиболее острой конкуренции называется ...

+: гильдия

-: когорта

-: стадо

84. Гипотеза консортивных связей Беклемишева-Раменского заключается в том, что ...

+: основной ячейкой трансформации энергии в экосистеме является консорция

-: основной ячейкой трансформации энергии в экосистеме является почва

-: конкурирующие виды являются основной ячейкой трансформации энергии в экосистеме

85. Может ли одна и та же модель одновременно использоваться для прогноза, анализа и описания взаимосвязи между переменными?

+: да

-: нет

86. Модель, которая представляют собой формализацию с помощью ЭВМ любых эмпирических сведений об объекте называется...

+: имитационной

-: аналитической

-: регрессионной

87. Схематические, математические и компьютерные модели относятся к ...

+: информационным моделям

-: вербальным моделям

-: материальным моделям

88. К абиотическим экологическим факторам относятся ...

-: Фитоценозы, определяющие ход биологической продуктивности

-: Почва, включая почвенных микроорганизмов и почвенную влагу

- + : Почвенная влага, воздух и подстилающие горные породы
- : Солнечная радиация и продуценты, использующие ее для производства

89. Как соотносятся между собой понятия биогеоценоз и экосистема?

- : как синонимы
- + : биогеоценоз – объективно существующая реальность, тогда как экосистема – есть отражение этой реальности в нашем сознании, определяемое целями исследования
- : экосистема представляет собой частный случай биогеоценоза
- + : биогеоценоз представляет собой частный случай экосистемы

90. Дайте определение биоценоза

- : Совокупность растительных организмов, занимающих определенную территорию
- : Совокупность почвенных микроорганизмов, определяющих формирование плодородного гумусового слоя
- : Совокупность животных, образующих трофические цепи
- + : Совокупность, взаимодействующих между собой организмов, населяющих экосистему

91. Равенство средних величин для нормально распределенных совокупностей в случае равенства их дисперсий можно определить с помощью ...

- : F-критерия
- : критерия хи-квадрат
- + : t-критерия Стьюдента

92. Согласованность эмпирического распределения с теоретическим можно определить с помощью ...

- : F-критерия
- + : критерия хи-квадрат
- : t-критерия Стьюдента

93. Равенство дисперсий двух эмпирических распределений можно проверить с помощью ...

+: F-критерия

-: критерия хи-квадрат

-: t-критерия Стьюдента

94. Гипотеза о равенстве средних величин двух эмпирических распределений подтверждается на принятом уровне значимости, если ...

-: $t_{набл} > t_{крит}$

+: $t_{набл} < t_{крит}$

95. Соответствие эмпирического распределения особей в пробах закону Пуассона указывает на то, что особи распределены в пространстве ...

+: случайно

-: равномерно

-: контагиозно

Тема: Методы исследования, эмпирико-статистические модели экосистем

96. Применение в экологических исследованиях методов математической статистики определяется тем, что ...

+: экосистемы являются стохастически-детерминированными системами

-: экосистемы являются динамическими системами

-: экосистемы являются саморегулирующимися системами

97. Первичная статистическая обработка данных включает в себя следующие процедуры ...

+: отображение переменных в той или иной шкале

+: статистическое описание исходных совокупностей (определение пределов варьирования, построение эмпирических распределений)

+: восстановление пропущенных наблюдений

+: унификация типов переменных (перевод признаков в одну шкалу)

+: анализ законов распределений

-: построение динамической модели

-: проверка модели на адекватность

98. Для характеристики номинальных данных наиболее часто используются...

+: пропорция и процентное отношение

-: абсолютные величины

-: логарифмы чисел

99. Порядковые шкалы соответствуют таким качественным переменным, для которых характерна ...

+: упорядоченность,

-: непрерывность

-: дискретность

100. Флуктуирующая асимметрия – это ...

+: мелкие ненаправленные нарушения гомеостаза развития, являющиеся ответом организма на состояние окружающей среды.

-: точное соответствие между правой и левой половинами тела

-: отклонения в промерах между правой и левой половинами тела

101. Для анализа состояния наземных экосистем по методу флуктуирующей асимметрии используются следующие виды-индикаторы ...

+: береза повислая

-: тополь

-: ель

+: полевая мышь

-: белка

102. Для состояния водных экосистем по методу флуктуирующей асимметрии используются следующие виды-индикаторы ...

+: плотва

+: окунь

- +: лещ
- +: лягушка
- : сом
- : судак
- : уж

103. Для оценки качества среды по методу флуктуирующей асимметрии используется ...

- +: пятибалльная шкала
- : десятибалльная шкала
- : восьмибалльная шкала

104. Статистическая значимость различий между выборками по величине интегрального показателя стабильности развития определяется по ...

- +: по t-критерию Стьюдента
- : по критерию Фишера
- : по критерию хи-квадрат

105. Правильные значения, оценивающие показатель плотности популяции следующие ...

- +: кг/га
- : экз./м
- +: экз./л
- +: экз./м²
- +: экз./м³

106. Ошибка при оценке плотности популяции должна составлять не более ...

- : 5 %
- : 10 %
- +: 20 %
- : 30 %

107. Правильные значения, оценивающие показатель обилия популяции следующие ...

- +: %
- +: экз./ч
- : шт.
- : кг/м²

108. Встречаемость – это ...

- + : процент пробных площадок, на которых встречается данный вид
- : процентное содержание вида на данной площадке
- + : процент пробных площадок, на которых встречается набор видов
- : процентное содержание группы видов на данной площадке

109. Встречаемость вида считается высокой, если он встречается в ...

- + : 51 % площадок
- + : 70 % площадок
- : 45 % площадок
- + : 90 % площадок

110. Репрезентативностью выборки – это ...

- + : Степень соответствия характеристик выборки истинным характеристикам биологического объекта
- : Степень не соответствия характеристик выборки истинным характеристикам биологического объекта
- : Точность исследования
- : Способ, когда изучается лишь небольшая группа объектов

111. Фиксированный размер площадок важен при...

- + : оценке биоразнообразия
- : при описании растительности на фитоценоотическом уровне
- : при описании мелких неподвижных организмов
- : при описании деревьев

112. Назовите автора следующего принципа: Скопление особей популяции, с одной стороны, усиливает конкуренцию между ними за пищевые ресурсы и жизненное пространство, с другой – приводит к повышению способности группы в целом к выживанию

- + : Олли
- : Барри Коммонер
- : Шельфорд
- : Дарвин

113. Гипотезе Вильямса заключается в том, что...

- + : при возрастании обследуемой площади увеличивается числа видов
- : при возрастании обследуемой площади числа видов уменьшается
- : зависимость между числом видов и обследованной площадью выражается экспоненциальной кривой

114. Математическим описанием связи число видов - площадь чаще всего является ...

- + : ступенчатая функция
- : экспоненциальная функция
- : парабола
- : синусоидальная кривая

115. К качественным измерениям признаков относятся следующие шкалы ...

- + : номинальная шкала
- + : ранговая шкала
- : интервальная шкала

116. Из всех перечисленных статистических характеристик параметров популяции меры вариации следующие ...

- : средняя арифметическая
- + : дисперсия
- + : коэффициент вариации
- + : стандартное отклонение
- : мода
- : медиана

117. Величина ошибки измерения параметров популяции зависит от ...

- + : изменчивости признака
- + : размеров выборки
- : численности популяции
- : величины генеральной совокупности

118. Какое из этих исследований должно проводиться раньше?

+ : Исследование особенностей распределения особей в пределах сообщества

- : Определения численности

119. Больше проб или их размеров требуется в учете ...

- : случайного распределения организмов

+ : контагиозного распределения организмов

120. Адекватность моделей распределения популяций проверяют с помощью критерия ...

- : Стьюдента

+ : хи-квадрат

- : Фишера

121. Наиболее пригодной для сравнения видов в сообществе является ...

- : плотность

+ : продукция

- : проектное покрытие

- : встречаемость

- : биомасса

122. Ранговое распределение устанавливает зависимость между ...

+ : числом особей каждого вида и числом видов

- : обилием и рангом вида

+ : обилием вида и рангом вида по обилию

- : числом особей каждого вида и числом видов

123. В стабильных ненарушенных сообществах кривая для численности лежит ... кривой для биомассы

- : выше

+ : ниже

124. Постоянство – это ...

+ : отношение числа содержащих изучаемый вид выборок (p) к общему числу выборок (P), выраженное в процентах

- : отношение числа особей одного вида (n) к общей численности особей (N)

- : число особей на единицу площади или объема

125. Нумерическое видовое богатство – это ...

- +: число видов на строго оговоренное число особей или на определенную биомассу
- : число составляющих его видов
- : число составляющих его особей

126. Нумерическое видовое богатство применяется если ...

- +: не имеется полного списка видов сообщества
- +: исследуются экологические воздействия на сообщества рыб
- +: исследуются водные объекты
- : имеется полный список видов сообщества

127. Лучшим среди индексов, оценивающим доминирование видов в сообществе считается индекс ...

- : Индекс Маргалефа
- : Индекс Шеннона
- : Индекс выровненности
- +: Индекс Симпсона
- : Индекс Менхиника

128. Правильная последовательность процедур для анализа данных по разнообразию видов следующая ...

- 1: Формирование выборок
- 2: Графический анализ данных
- 3: Проверка эмпирических данных теоретической модели
- 4: Расчет индексов разнообразия
- 5: Проверка статистических гипотез значимости различий между сообществами

129. Для сравнения с результатами исследований по разнообразию сообществ других авторов предпочтительнее использовать ...

- +: индекса Шеннона
- : индекс Симпсона
- : индекс Маргалефа

130. Индекс Шеннона может принять следующее значение ...

- : 1,1
- +: 1,6
- : 10
- +: 3,8

131. Индекс выравненности принимает значения в промежутке ...

+: [0; 1]

-: [0; 2]

-: [-1; 1]

-: [0,5; 1]

-: [1; 3]

Тема: *Модели факториальной экологии*

132. Выравненность – это ...

+: равномерность распределения видов по их обилию в сообществе

-: число видов, отнесенное к определенной площади

-: число видов на строго оговоренное число особей

133. Законы толерантности относятся к ...

+: физиологическим факторам

-: к ресурсным факторам

134. Можно ли утверждать следующее высказывание: «Лимитирующие факторы для сообщества всегда потребляются из окружающей среды полностью»

+: да

-: нет

135. Укажите из перечисленных факторов антропогенные стрессоры

+: химическое загрязнение

-: температурный режим

+: шум

-: ветер

-: паразитизм

136. Модель роста растения Полетаева формализует следующий закон:

+: Закон Либиха

-: Закон Ферхюльста

-: Закон Митчерлиха

137. Кто автор принципа совокупного действия факторов?

+: Митчерлих

-: Алехин

-: Полетаев

-: Одум

138. Функция зависимости количественных оценок характеристик экосистемы от важнейших экологических факторов называется ...

-: функцией правдоподобия

+: функцией отклика

-: функцией благополучия

139. Математическим аппаратом метода функции отклика является ...

+: нелинейный регрессионный анализ

-: системы дифференциальных уравнений

-: матричная алгебра

140. В математической формулировке лимитирующим называется тот фактор, по которому для достижения заданного относительного изменения функции отклика необходимо....

+: минимальное относительное изменение значения фактора

-: максимальное относительное изменение значения фактора

141. При отсутствии (или дефиците) кальция многие моллюски при строительстве своих раковин заменяют его на стронций. Какую гипотезу факториальной экологии подтверждает это наблюдение?

+: Гипотезу компенсации экологических факторов

-: Гипотезу незаменимости фундаментальных факторов

-: Принцип совокупности экологических факторов

142. Одним из вариантов формализации закона минимума Либиха является ...

+: модель роста растения Полетаева.

-: модель функции отклика

-: модель Ферхюльста-Пирла

143. Основные процессы превращения вещества и энергии в растениях, учтенные в модели роста растения Полетаева следующие ...

+: процесс возрастания биомассы (роста), идущий за счет фотосинтеза и поглощения веществ из почвы и атмосферы

+: процесс основного обмена (дыхания)

-: выделение продуктов обмена

144. Частная функция отклика представляет собой ...

+: функцию одной переменной

-: функцию двух переменных

-: функцию многих переменных

145. Лучшими биоиндикаторами являются ...

+: Стенотопные виды

-: Эвриотопные виды

146. Обобщенной функцией отклика F_k называется функция ...

+: зависимости значений k -го показателя или процесса от всех рассматриваемых экологических факторов

+: представленная как комбинация частных функций отклика

-: зависимости значений k -го показателя или процесса от одного из рассматриваемых экологических факторов

147. Что является более надежным индикатором?

-: Численность одного вида

+: Численное отношение разных видов

148. Конкуренция между двумя популяциями разных видов возрастает при условии, когда:

-: одна популяция более многочисленна, чем другая

-: два вида тесно взаимосвязаны

+: экологические ниши конкурирующих видов перекрываются в больших масштабах

-: у видов сходная плодовитость

149. Укажите все категории экологических ниш

+: пространственная

+: трофическая

+: многомерная

+ : фундаментальная

+ : реализованная

- : прямоугольная

150. Заполнение ниш происходит по следующей схеме:

- : Крупный сменяет мелкого

+ : Мелкий сменяет крупного

+ : Менее организованный сменяет высокоорганизованного

- : Высокоорганизованный сменяет менее организованного

+ : Генетически изменчивый сменяет менее изменчивого

151. Концепция одномерной иерархической ниши более актуальна по отношению к ...

+ : Животному миру

- : Растительному миру

152. Принцип классификации сообществ по Браун-Бланке заключается ...

+ : В преобразовании матрицы данных к блочной форме

- : В анализе матрицы корреляций между видами

153. Законы толерантности соответствуют ...

+ : физиологическим факторам

- : ресурсным факторам

Тема: *Динамические модели экосистем*

154. К динамическим показателям популяций относятся следующие ...

+ : рождаемость

+ : смертность

+ : скорость роста популяции

- : плотность

- : масса

- : встречаемость

155. Кривая скорости роста популяции имеет вид ...

- + : одновершинной параболы
- : гиперболы
- : логарифмической кривой
- : синусоиды

156. Логистический рост численности популяции графически выражается ...

- + : S-образной кривой
- : одновершинной параболой
- : гиперболой
- : логарифмической кривой
- : синусоидой
- : экспонентой

157. Можно ли утверждать следующее высказывание: «Лимитирующие факторы для сообщества всегда потребляются из окружающей среды полностью»

- + : да
- : нет

158. Физиологическая продолжительность жизни – это ...

- + : продолжительность жизни, без влияния на особей лимитирующих факторов
- : продолжительность жизни в реальных условиях среды
- : среднее арифметическое продолжительности жизни всех особей

159. Законы толерантности соответствуют ...

- + : физиологическим факторам
- : ресурсным факторам

160. Скорость роста популяции обратно пропорциональна ...

- + : длительности поколения
- : чистой скорости воспроизводства
- : ожидаемому числу потомков

161. Модель роста растения Полетаева формализует следующий закон:

+: Закон Либиха

-: Закон Ферхюльста

-: Закон Митчерлиха

162. Согласно гипотезе Николсона контролирующие стабильность популяции факторы управляются ...

+: плотностью

-: численностью

-: пространственной структурой

163. Принцип экспоненциального роста используется в экологии, прежде всего, для того, чтобы ...

+: количественно охарактеризовать потенциальные возможности популяции к росту

+: выявить факторы, ограничивающие рост изучаемой популяции

-: определить максимальную численность популяции

164. В модели Мальтуса зависимость скорости роста популяции от ее численности представляет собой ... функцию

+: линейную

-: параболическую

-: гиперболическую

165. Согласно закона ограниченного роста Дарвина ...

+: Окружающая среда действует как лимитирующий фактор на биоценотический потенциал популяции (прежде всего на потенциал размножения).

-: Окружающая среда не действует как лимитирующий фактор на биоценотический потенциал популяции.

166. Чистая скорость воспроизводства показывает ...

+: во сколько раз увеличится численность популяции за одно поколение

-: на сколько особей увеличится популяция за одно поколение

167. R-стратегия это ...

-: Низкий потенциал роста, высокая способность конкурировать за ресурсы и использовать их.

-: Низкое сопротивление условиям окружающей среды при высокой упругости

+: Высокая плодовитость, способность быстро распространяться в новые места обитания, короткий жизненный цикл

-: Низкий потенциал роста за счет высокой сопротивляемости воздействию окружающей среды

168. К-стратегия – это ...

+: Низкий потенциал роста за счет высокой сопротивляемости воздействию окружающей среды

-: Низкое сопротивление условиям окружающей среды при высокой упругости

-: Высокая плодовитость, способность быстро распространяться в новые места обитания. короткий жизненный цикл

-: Низкий потенциал роста, высокая способность конкурировать за ресурсы и использовать их

169. Коэффициент K в логистическом уравнении означает ...

-: Величину, показывающую во сколько раз возрастет численность популяции при экспоненциальном росте

-: Коэффициент внутривидовой конкуренции

-: Чистую скорость размножения, характеризующую общее число производимых потомков

+: Плотность популяции, которая в данных условиях среды может поддерживаться на постоянном уровне

170. Ордината точки перегиба логистической кривой соответствует ...

+: половине максимальной численности

-: одной трети максимальной численности

171. Коэффициент внутривидовой конкуренции представляет собой ...

+: отношение биотического потенциала к емкости среды

-: отношение биотического потенциала к численности популяции в данный момент времени

172. Согласно принципу Гаузе ...

- + : два вида не могут существовать в одной и той же экологической нише, если их экологические потребности совпадают
- : два вида могут существовать в одной и той же экологической нише, если их экологические потребности совпадают

173. Конкуренция представляет собой фактор ...

- + : зависящий от плотности
- : не зависящий от плотности

174. В модели Вольтерра жертва увеличивает число своих особей пропорционально ...

- + : внутренней скоростью роста (r_1)
- + : плотности жертвы (N_1)
- : числу встреч особей хищника и жертвы $N_1 N_2$

175. В модели Вольтерра хищник увеличивает число своих особей пропорционально ...

- + : произведению плотностей хищника и жертвы $N_1 N_2$
- : смертности (d_2), умноженной на плотность хищника

176. Кто автор принципа совокупного действия факторов?

- + : Митчерлих
- : Алехин
- : Полетаев
- : Одум

177. Функция зависимости количественных оценок характеристик экосистемы от важнейших экологических факторов называется ...

- : функцией правдоподобия
- + : функцией отклика
- : функцией благополучия

178. Математическим аппаратом метода функции отклика является ...

- + : нелинейный регрессионный анализ
- : системы дифференциальных уравнений
- : матричная алгебра

179. В математической формулировке лимитирующим называется тот фактор, по которому для достижения заданного относительного изменения функции отклика необходимо....

+: минимальное относительное изменение значения фактора

-: максимальное относительное изменение значения фактора

180. Одним из вариантов формализации закона минимума Либиха является ...

+: модель роста растения Полетаева.

-: модель функции отклика

-: модель Ферхюльста-Пирла

181. Основные процессы превращения вещества и энергии в растениях, учтенные в модели роста растения Полетаева следующие ...

+: процесс возрастания биомассы (роста), идущий за счет фотосинтеза и поглощения веществ из почвы и атмосферы

+: процесс основного обмена (дыхания)

-: выделение продуктов обмена

182. Частная функция отклика представляет собой ...

+: функцию одной переменной

-: функцию двух переменных

-: функцию многих переменных

183. Лучшими биоиндикаторами являются ...

+: Стенотопные виды

-: Эвриотопные виды

184. Обобщенной функцией отклика F_k называется функция ...

+: зависимости значений k -го показателя или процесса от всех рассматриваемых экологических факторов

+: представленная как комбинация частных функций отклика

-: зависимости значений k -го показателя или процесса от одного из рассматриваемых экологических факторов

185. Какие по отношению к массе виды являются лучшими индикаторами?

+: Более крупные

-: Более мелкие

186. Что является более надежным индикатором?

- : Численность одного вида
- +: Численное отношение разных видов

187. Укажите все категории экологических ниш

- +: пространственная
- +: трофическая
- +: многомерная
- +: фундаментальная
- +: реализованная
- : прямоугольная

188. Укажите из перечисленных факторов антропогенные стрессоры

- +: химическое загрязнение
- : температурный режим
- +: шум
- : ветер
- : паразитизм

189. Можно ли утверждать следующее высказывание: «Лимитирующие факторы для сообщества всегда потребляются из окружающей среды полностью»

- +: да
- : нет

190. Модель роста растения Полетаева формализует следующий закон:

- +: Закон Либиха
- : Закон Ферхюльста
- : Закон Митчерлиха

191. Кто автор принципа совокупного действия факторов?

- +: Митчерлих
- : Алехин
- : Полетаев
- : Одум

192. Функция зависимости количественных оценок характеристик экосистемы от важнейших экологических факторов называется ...

-: функцией правдоподобия

+: функцией отклика

-: функцией благополучия

193. Математическим аппаратом метода функции отклика является ...

+: нелинейный регрессионный анализ

-: системы дифференциальных уравнений

-: матричная алгебра

194. В математической формулировке лимитирующим называется тот фактор, по которому для достижения заданного относительного изменения функции отклика необходимо....

+: минимальное относительное изменение значения фактора

-: максимальное относительное изменение значения фактора

195. При отсутствии (или дефиците) кальция многие моллюски при строительстве своих раковин заменяют его на стронций. Какую гипотезу факториальной экологии подтверждает это наблюдение?

+: Гипотезу компенсации экологических факторов

-: Гипотезу незаменимости фундаментальных факторов

-: Принцип совокупности экологических факторов

196. Одним из вариантов формализации закона минимума Либиха является ...

+: модель роста растения Полетаева.

-: модель функции отклика

-: модель Ферхюльста-Пирла

197. Частная функция отклика представляет собой ...

+: функцию одной переменной

-: функцию двух переменных

-: функцию многих переменных

198. Основные процессы превращения вещества и энергии в растениях, учтенные в модели роста растения Полетаева следующие ...

+: процесс возрастания биомассы (роста), идущий за счет фотосинтеза и поглощения веществ из почвы и атмосферы

+: процесс основного обмена (дыхания)

-: выделение продуктов обмена

199. Лучшими биоиндикаторами являются ...

+: Стенотопные виды

-: Эвриотопные виды

200. бобщенной функцией отклика F_k называется функция ...

+: зависимости значений k -го показателя или процесса от всех рассматриваемых экологических факторов

+: представленная как комбинация частных функций отклика

-: зависимости значений k -го показателя или процесса от одного из рассматриваемых экологических факторов

201. Какие по отношению к массе виды являются лучшими индикаторами?

+: Более крупные

-: Более мелкие

202. Что является более надежным индикатором?

-: Численность одного вида

+: Численное отношение разных видов

203. Укажите все категории экологических ниш

+: пространственная

+: трофическая

+: многомерная

+: фундаментальная

+: реализованная

-: прямоугольная

Тема: *Экологическое моделирование и прогнозирование*

204. Проведение учета количественных характеристик измеряемых величин и разнообразных взаимосвязей между ними в природных условиях, а так же учета реакций природных объектов на различные внешние воздействия стали возможными при возникновении ...

-: экологического контроля

+: экологического прогнозирования и моделирования

-: экологического контроля и прогнозирования

205. Созданием моделей объектов мониторинга, которые позволяют объективно выявлять уровень их чувствительности к различным природным и антропогенным воздействиям занимается ...

+: экологическое моделирование

-: экологическое прогнозирование

-: экологический контроль

206. Возможность прогнозировать будущие изменения и оценивать устойчивость всей биосферы различным воздействия человека позволяют модели ... уровня

-: локального

-: регионального

+: глобального

207. Отрезок времени, который система затрачивает на переход из неравновесного состояния, обусловленного различными причинами, в новое равновесное, называется временем ...

-: перехода

+: релаксации

-: восстановления

208. Прогностические модели, отражающие различные стороны человеческого воздействия на природу, могут быть ...

-: только формализованными

+: формализованными и неформализованными

-: только неформализованными

209. Формализованные модели системы и ее частей описываются ...

- : описательными средствами
- +: математическими средствами

210. ... представляет собою способ кодирования как известных, так и неизвестных свойств, изучаемых объектов.

- : Описательная модель
- +: Математическая модель
- : Неформализованная модель

211. Методы экологического моделирования можно разделить на ...

- +: физические и математические
- : математические и географические
- : физические и географические

212. ... представляет собою определенное упрощение той или иной экосистемы.

- +: Экологическая модель
- : Математическая модель
- : Физическая модель

213. Модель определяет ... развития экосистемы на основе оценки ее качественных изменений.

- : закономерности
- +: динамику

214. Модель ... направления создается с целью оценки влияния внешних факторов на переменные, описывающие экосистему, в обозримом будущем

- : экологического
- : описательного
- +: прогностического

215. ... модели зависят от набора показателей и числа связей в ней

- : Оригинальность
- +: Сложность
- : Актуальность

216. Важным информационным показателем при моделировании любой экосистемы является ее ...

- : развитие
- +: устойчивость
- : стагнация

217. ... – это способность экосистемы противостоять разным уровням воздействия среды.

- +: Устойчивость
- : Функционирование
- : Динамичность

218. Фаза стагнации в развитии экосистемы характеризуется периодом ...

- +: застоя
- : гибели
- : роста

219. Фаза перехода экосистемы в новое состояние может проходить вместо фазы ...

- : роста
- +: гибели
- : стагнации

*Тема: Информационное обеспечение и организация
экологического мониторинга*

220. Назовите данные, не входящий в состав справочного информационного блока ...

- : многолетние климатические данные
- : данные характеристики почв
- +: данные лабораторных анализов, проведенных в процессе мониторинга

221. На первом этапе организации мониторинга

- +: формируется цель, и определяются основные задачи
- : определяются оцениваемые параметры
- : проводят наблюдения и отбирают пробы

222. Картографические сведения о природных условиях и хозяйственном использовании территории отражены в ... форме картографической информации

+: базовой

-: оперативной

-: прогностической

223. ... форма картографических сведений позволяет определить изменения в природных или агроландшафтных системах при влиянии хозяйственной деятельности на эти системы и на здоровье человека

-: Базовая

-: Оперативная

+: Оценочная

224. ... карты включают сведения о распределении температуры, влажности, направлении и скорости ветра по отдельным полигонам в различные сезоны года

-: Базовые

+: Прогностические

-: Оценочные

225. Третий уровень информационного обеспечения мониторинга включает

+: сбор информации полевых наблюдений и лабораторных анализов в процессе проведения мониторинга

-: только сбор информации полевых наблюдений в процессе проведения мониторинга

-: только сбор информации лабораторных анализов в процессе проведения мониторинга

226. Управление сбором информации, а также функционированием системы мониторинга и его отдельных блоков осуществляет ...

+: экологический центр

-: блок моделирования и прогноза

-: банк обработанных данных

227. Блок ... выполняет сбор сведений на стационарных полигонах экспедициями

+ : сбора и передачи информации

- : анализа материалов и разработки программ по управлению природными процессами

- : блок обработки и хранения информации

228. Блок ... включает различные базы данных и их первичный анализ

- : сбора и передачи информации

- : анализа материалов и разработки программ по управлению природными процессами

+ : блок обработки и хранения информации

229. Блок ... включает базу моделей, характеризующих перемещение отдельных загрязнений при различных режимах природно-климатических условий и используемых при разработке прогнозов

- : сбора и передачи информации

+ : прогностического моделирования

- : блок обработки и хранения информации

230. ... включает сбор всей уже имеющейся информации по интересующей проблеме и ее анализ

- : Сбор информации полевых наблюдений в процессе проведения мониторинга

+ : Рекогносцировочное исследование

- : Сбор информации лабораторных анализов в процессе проведения мониторинга

231. Процесс проведения экологического мониторинга сопровождается документацией ...

- : от отбора проб до их анализа

+ : от отбора проб до разработки прогностической модели

- : с момента обработки данных до разработки прогностической модели

232. Каждая отобранная проба сопровождается ...

- : ведомостью
- +: протоколом
- : сводной таблицей данных

233. Результаты лабораторных исследований записываются в ...

- +: лабораторный журнал
- : протокол
- : полевой дневник

234. Все первичные результаты (протоколы, рабочие журналы и др. документы) хранятся в течение ...

- : одного года
- : двух лет
- +: всего периода мониторинга по данному полигону

235. Аналитическая информация должна отличаться ... достоверностью и реально отражать наличие определенных веществ в различных объектах

- +: качественно и количественно
- : качественно
- : количественно

236. Первостепенным фактором, оказывающим большое влияние на достоверность анализа, является ...

- : ведение лабораторного журнала
- +: отбор проб
- : обобщение результатов

237. Во избежание ошибок необходимо использовать ... методы отбора проб

- : современные
- : зарубежные
- +: стандартные

238. На основе ... вырабатываются рекомендации по стабилизации (или улучшению) экологической ситуации в определенном районе

- : полевых исследований
- +: анализа и обработки полученных данных

-: справочной информации

239. Ежегодное ведение мониторинга завершается ...

+: составлением научного отчета

-: сбором данных полевых и лабораторных исследований

-: обработкой результатов

240. В ... годового отчета излагаются важнейшие результаты выполненных исследований

-: обзоре литературы

+: аннотации

-: приложения

СПИСОК ЛИТЕРАТУРЫ

1. Анализ данных и математическое моделирование в экологии и природопользовании : учеб. пособие / И. С. Белюченко, А. В. Смагин, Л. Б. Попок, Л. Е. Попок. – Краснодар : КубГАУ, 2015. – 265 с.

2. Белюченко И. С. Организация экологического мониторинга биоразнообразия при изменении окружающей среды: практическое пособие для магистров / И. С. Белюченко, О. А. Мельник, Ю. Ю. Никифорова. – Краснодар : КубГАУ, 2012. – 70 с.

3. Белюченко И. С. Практикум по экологии : учеб.-метод. пособие / И. С. Белюченко, Л. Б. Попок. – Краснодар : КубГАУ, 2010. – 293 с.

4. Биомониторинг состояния окружающей среды : учеб. пособие / под. ред. И. С. Белюченко, Е. В. Федоненко, А. В. Смагина. – Краснодар : КубГАУ, 2014. – 153 с.

5. Биоразнообразие : курс лекций / сост. : Б. В. Кабельчук, И. О. Лысенко, А. В. Емельянов, А. А. Гусев. – Ставрополь : АГРУС, 2013. – 156 с.

6. Бродский А. К. Биоразнообразие : учебник / А. К. Бродский. – М. : Академия, 2012. – 207 с.

7. Василевич В. И. Статистические методы в геоботанике / В. И. Василевич. – М. : Наука, 1969. – 232 с.

8. Гринин А. С. Математическое моделирование в экологии : учеб. пособие для вузов / А. С. Гринин, Н. А. Орехов, В. Н. Новиков. – М. : ЮНИТИ-ДАНА, 2003. – 269 с.

9. Дулепов В. И. Системная экология : учеб. пособие / В. И. Дулепов, О. А. Лескова, И. С. Майорова. – Владивосток : Изд.-во ВГУЭС, 2004. – 252 с.

10. Елина Е. Е. Биоразнообразие / Е. Е. Елина. – Оренбург : Экспресс-печать, 2016. – 36 с.

11. Лакин Г. Ф. Биометрия : учеб. пособие для биологич. спец. вузов / Г. Ф. Лакин. – М. : Высшая школа, 1980. – 293 с.

12. Лебедева Н. В. Биологическое разнообразие : учеб. пособие для студ. высш. учеб. заведений / Н. В. Лебедева,

Н. Н. Дроздов, Д. А. Кривошукский. – М. : Гуманит. изд. центр ВЛАДОС, 2004. – 432 с.

13. Методическое пособие по статистической обработке данных экологического мониторинга / И. С. Белюченко, О. А. Мельник, Ю. Ю. Петух [и др.] – Краснодар : КубГАУ, 2010. – 60 с.

14. Мэгарран Э. Экологическое разнообразие и его измерение / Э. Мэгарран. – М. : Мир, 1992. – 161 с.

15. Негробов О. П. Учение о биоразнообразии / О. П. Негробов, С. О. Негробов, О. О. Маслова. – Воронеж : Изд. полиграфический центр Воронежского государственного университета, 2012. – 124 с.

16. Основы экологического мониторинга : практ. пособие / И. С. Белюченко, А. В. Смагин, Г. В. Волошина [и др.]. – Краснодар : КубГАУ, 2012. – 252 с.

17. Попок Л. Б. Основы компьютерных технологий в экологии и природопользовании : практ. пособие / Л. Б. Попок, Л. Е. Попок; под общ. ред. И. С. Белюченко. – Краснодар : КубГАУ, – 2012. – 53 с.

18. Пузаченко Ю. Г. Математические методы в экологических и географических исследованиях : учеб. пособие для студ. вузов / Ю. Г. Пузаченко. – М. : Издательский центр «Академия», 2004. – 416 с.

19. Розенберг Г. С. Количественная гидроэкология / Г. С. Розенберг, В. К. Шитиков, Д. П. Мозговой. – СамНЦ РАН, 2003. – 396 с.

20. Розенберг Г. С. Модели в фитоценологии / Г. С. Розенберг. – М. : Наука, 1984. – 240 с.

21. Розенберг Г. С. Экологическая информатика : учеб. пособие / Г. С. Розенберг, В. К. Шитиков, Д. П. Мозговой. – Самара : Изд-во Самар. ун-та, 1993. – 151 с.

22. Системная экология : практикум / И. С. Белюченко, Е. И. Муравьев, Е. В. Попок, Л. Б. Попок. – Краснодар : КубГАУ, 2007. – 184 с.

ПРИЛОЖЕНИЯ

Приложение А

Таблица А1 – Значение t-критерия Стьюдента при уровне
значимости 0,05; 0,01

Число степеней свободы (к)	Р		Число степеней свободы (к)	Р	
	0,05	0,01		0,05	0,01
1	12,706	63,657	18	2,1009	2,8784
2	4,3027	9,9248	19	2,0930	2,8609
3	3,1825	5,8409	20	2,0860	2,8453
4	2,7764	4,0641	21	2,0796	2,8314
5	2,5706	4,0321	22	2,0739	2,8188
6	2,4469	3,7074	23	2,0687	2,0873
7	2,3646	3,4995	24	2,0639	2,7969
8	2,3060	3,3554	25	2,0595	2,7874
9	2,2622	3,2498	26	2,0555	2,7787
10	2,2281	3,1693	27	2,0518	2,7707
11	2,2010	3,1058	28	2,0484	2,7633
12	2,1788	3,0545	29	2,0452	2,7564
13	2,1604	3,0123	30	2,0423	2,7500
14	2,1448	2,9768	40	2,0211	2,7045
15	2,1315	2,9467	60	2,0003	2,6603
16	2,1199	2,9208	120	1,9799	2,6174
17	2,1098	2,8982	140	1,9600	2,5758

Приложение Б

Таблица Б1 – F-критерий Фишера (уровень значимости 0,05)

f ₁	f ₂						
	1	2	3	4	5	6	7
1	161	200	216	225	230	234	237
2	18,5	19,0	19,2	19,2	19,3	19,3	19,4
3	10,1	9,55	9,28	9,12	9,10	8,94	8,89
4	7,71	6,94	6,59	6,39	6,26	6,16	6,09
5	6,61	5,79	5,41	5,19	5,05	4,95	4,88
6	5,99	5,14	4,76	4,53	4,39	4,28	4,21
7	5,59	4,74	4,35	4,12	3,97	3,87	3,79
8	5,42	4,46	4,07	3,84	3,69	3,58	3,50
9	5,12	4,26	3,86	3,63	3,48	3,37	3,29
10	4,96	4,10	3,71	3,48	3,33	3,22	3,14
11	4,84	3,98	3,59	3,36	3,20	3,09	3,01
12	4,75	3,89	3,49	2,26	3,11	3,00	2,91
13	4,67	3,81	3,41	3,18	3,03	2,92	2,83
14	4,60	3,74	3,34	3,11	2,96	2,85	2,76
15	4,54	3,68	3,29	3,06	2,90	2,79	2,71
16	4,49	3,63	3,24	3,01	2,85	2,74	2,66

Приложение В

Таблица В1 – Критические значения z -критерия знаков
при разных уровнях значимости α
и объеме выборки n

n	$\alpha, \%$		n	$\alpha, \%$		n	$\alpha, \%$		n	$\alpha, \%$	
	5	1		5	1		5	1		5	1
6	6	–	30	21	23	54	35	37	78	49	51
7	7	–	31	22	24	55	36	38	79	49	52
8	8	8	32	23	24	56	36	39	80	50	52
9	8	9	33	23	25	57	37	39	81	50	53
10	9	10	34	24	25	58	37	40	82	51	54
11	10	11	35	24	26	59	38	40	83	51	54
12	10	11	36	25	27	60	39	41	84	52	55
13	11	12	37	25	27	61	39	41	85	53	55
14	12	13	38	26	28	62	40	42	86	53	56
15	12	13	39	27	28	63	40	43	87	54	56
16	13	14	40	27	29	64	41	43	88	54	57
17	13	15	41	28	30	65	41	44	89	55	58
18	14	15	42	28	30	66	42	44	90	55	58
19	15	16	43	29	31	67	42	45	91	56	59
20	15	17	44	29	31	68	43	46	92	56	59
21	16	17	45	30	32	69	44	46	93	57	60
22	17	18	46	31	33	70	44	47	94	57	60
23	17	19	47	31	33	71	45	47	95	58	61
24	18	19	48	32	34	72	45	48	96	59	62
25	18	20	49	32	34	73	46	48	97	59	62
26	19	20	50	33	35	74	46	49	98	60	63
27	20	21	51	33	36	75	47	50	99	60	63
28	20	22	52	34	36	76	48	50	100	61	64
29	21	22	53	35	37	77	48	51	–	–	–
P	0,05	0,01	–	0,05	0,01	–	0,05	0,01	–	0,05	0,01

Приложение Г

Таблица Г1 – Таблица критических точек распределения χ^2

Число степеней свободы (k)	Уровень значимости						
	0,01	0,02	0,025	0,05	0,95	0,975	0,99
1	6,6	5,4	5,0	3,8	0,0039	0,00098	0,00016
2	9,2	7,8	7,4	6,0	0,103	0,051	0,020
3	11,3	9,8	9,3	7,8	0,352	0,216	0,115
4	13,3	11,7	11,1	9,5	0,711	0,484	0,297
5	15,1	13,4	12,8	11,1	1,15	0,831	0,554
6	16,8	15,0	14,4	12,6	1,64	1,24	0,872
7	18,5	16,6	16,0	14,1	2,17	1,69	1,24
8	20,1	18,2	17,5	15,5	2,73	2,18	1,65
9	21,7	19,7	19,0	16,9	3,33	2,70	2,09
10	23,2	21,2	20,5	18,3	3,94	3,25	2,56
11	24,7	22,6	21,9	19,7	4,57	3,82	3,05
12	26,2	24,1	23,3	21,0	5,23	4,40	3,57
13	27,7	25,5	24,7	22,4	5,89	5,01	4,11
14	29,1	26,9	26,1	23,7	6,57	5,63	4,66
15	30,6	28,3	27,5	25,0	7,26	6,26	5,23
16	32,0	29,6	28,8	26,3	7,96	6,91	5,81
17	33,4	31,0	30,2	27,6	8,67	7,56	6,41
18	34,8	32,3	31,5	28,9	9,39	8,23	7,01
19	36,2	33,7	32,9	30,1	10,1	8,91	7,63
20	37,6	35,0	34,2	31,4	10,9	9,59	8,26
21	38,9	36,3	35,5	32,7	11,6	10,3	8,90
22	40,3	37,7	36,8	33,9	12,3	11,0	9,54
23	41,6	39,0	38,1	35,2	13,1	11,7	10,2
24	43,0	40,3	39,4	36,4	13,8	12,4	10,9
25	44,3	41,6	40,6	37,7	14,6	13,1	11,5
26	45,6	42,9	41,9	38,9	15,4	13,8	12,2
27	47,0	44,1	43,2	40,1	16,2	14,6	12,9
28	48,3	45,4	44,5	41,3	16,9	15,3	13,6
29	49,6	46,7	45,7	42,6	17,7	16,0	14,3
30	50,9	48,0	47,0	43,8	18,5	16,8	15,0

ОГЛАВЛЕНИЕ

ВВЕДЕНИЕ.....	3
РАЗДЕЛ 1. СТАТИСТИЧЕСКИЙ АНАЛИЗ РЕЗУЛЬТАТОВ ИССЛЕДОВАНИЙ.....	5
Тема 1. Шкалы измерения признаков.....	5
1.1 Типы шкал измерения признаков.....	6
1.2 Группировка данных по количественной вариации.....	10
1.3 Алгоритм построения интервального вариационного ряда.....	11
1.4 Графическое изображение вариационных рядов	14
Тема 2. Обзор основных статистических характеристик.....	23
2.1 Средние величины.....	23
2.2 Меры вариации признака.....	26
2.3 Оценка параметров генеральной совокупности...	31
Тема 3. Предварительная статистическая обработка экологических данных с помощью программы Microsoft Excel.....	38
Тема 4. Совместный анализ нескольких выборок.....	46
4.1 Сущность параметрических и непараметрических методов сравнения выборок.....	46
4.2 Проверка гипотезы о нормальном законе распределения.....	47
4.3 Критерий Фишера сравнения дисперсий	52
4.4 Критерий Стьюдента (t-критерий) сравнения средних величин.....	53
4.4.1 t-критерий для независимых выборок.....	54
4.4.2 Парный t-критерий для зависимых, связных выборок.....	56
4.5 Непараметрические методы сравнения выборок.....	58

Тема 5. Оценка плотности популяции.....	64
5.1 Оценка пространственного распределения особей	64
5.1.1 Индекс Одумана.....	64
РАЗДЕЛ 2. СТРУКТУРНЫЙ АНАЛИЗ ЭКОЛОГИЧЕСКИХ СИСТЕМ.....	67
Тема 6. Анализ видового разнообразия.....	67
6.1 Индексы разнообразия.....	67
6.2 t-критерий сравнения индексов Шеннона.....	69
РАЗДЕЛ 3. ОЦЕНКА ВЗАИМОСВЯЗИ В СИСТЕМЕ.....	75
Тема 7. Корреляционный анализ.....	75
7.1 Сущность и задачи корреляционного анализа.....	75
7.2 Коэффициент корреляции Пирсона.....	76
7.3 Множественная корреляция.....	81
7.4 Ранговая корреляция.....	83
Тема 8. Регрессионный анализ.....	88
8.1 Этапы линейного регрессионного анализа.....	88
8.2 Процедура построения линейной регрессионной модели.....	89
8.3 Построение парной линейной регрессионной модели с помощью MS EXCEL.....	92
8.4 Построение множественной линейной регрессионной модели с помощью MS EXCEL.....	95
Тема 9. Дисперсионный анализ.....	102
9.1 Понятие об однофакторном дисперсионном анализе.....	102
9.2 Проведение дисперсионного анализа с помощью электронных таблиц EXCEL.....	105
ТЕСТЫ.....	111
СПИСОК ЛИТЕРАТУРЫ.....	155
ПРИЛОЖЕНИЯ.....	157

Учебное издание

Попок Людмила Борисовна, **Мельник** Ольга Александровна,
Никифоренко Юлия Юрьевна и др.

СИСТЕМНАЯ ЭКОЛОГИЯ

Учебное пособие

В авторской редакции

Подписано в печать 15.04.2017. Формат 60 × 84 ¹/₁₆.

Усл. печ. л. – 9,5. Уч. изд. л. – 7,5.

Тираж 80 экз. Заказ № ____ .

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