

МИНИСТЕРСТВО СЕЛЬСКОГО ХОЗЯЙСТВА
РОССИЙСКОЙ ФЕДЕРАЦИИ

ФГБОУ ВПО «КУБАНСКИЙ ГОСУДАРСТВЕННЫЙ
АГРАРНЫЙ УНИВЕРСИТЕТ»

Факультет экологии
Кафедра прикладной экологии

МАТЕМАТИЧЕСКОЕ МОДЕЛИРОВАНИЕ В ЭКОЛОГИИ

Учебно-методическое пособие для практических занятий

Краснодар КубГАУ 2015

Составители: Горковенко Н.Е.

Пособие предназначено для оказания методической помощи при подготовке к семинарам по дисциплине **«Математическое моделирование в экологии»**, содержит тематику, цели и задачи каждого занятия, методику его проведения, задания для подготовки к семинарам, список рекомендуемой литературы.

Издание предназначено для обучающихся по направлению подготовки: 05.06.01 – Науки о Земле.

Рассмотрено и одобрено методической комиссией факультета экологии 29.06.2015 г., протокол № 10.

© Горковенко Н.Е., 2015
© ФГБОУ ВПО «Кубанский
государственный аграрный
университет», 2015

Методические указания для аспирантов
по дисциплине Б1.В.ДВ.1 Математическое моделирование
в экологии
к практическому занятию № 1

Тема: Моделирование роста популяции.

Цель занятия: Сформировать у аспирантов представление о математических моделях, описывающих изменение популяционных характеристик.

Задачи: Освоить методы построения дискретных моделей популяционной динамики.

Аспирант должен знать:

1) до изучения темы:

- Понятие математических моделей;
- Виды математических моделей биологических процессов;

2) после изучения темы:

- Дискретные модели неограниченной одиночной популяции;
- Дискретные модели ограниченной популяции.

Изучение темы занятия направлено на формирование профессиональных компетенций: способность использовать современные методы обработки, анализа и синтеза полевой и лабораторной биологической информации, знание принципов составления научно-технических проектов и отчетов (ПК-5).

Методика проведения занятия.

1. Определение темы занятия. Преподаватель поясняет цели и задачи занятия, значение полученных знаний для будущей работы по специальности.

2. Теоретическая часть.

1) **Построение дискретной модели неограниченной одиночной популяции.** Пусть скорость роста популяции в период времени n пропорциональна размеру популяции в начале этого периода

$$\Delta x_n = x_{n+1} - x_n = a \cdot x_n \quad (1)$$

Тогда численность популяции в следующий момент времени определится по формуле

$$x_{n+1} = (1 + a) \cdot x_n \quad (2)$$

Согласно (2) можно записать

$$x_1 = (1 + a) \cdot x_0$$
$$x_2 = (1 + a) x_1 = (1 + a)(1 + a)x_0 = (1 + a)^2 x_0$$

$$x_3 = (1+a)x_2 = (1+a)(1+a)(1+a)x_0 = (1+a)^3 x_0$$

$$\dots\dots$$

$$x_n = (1+a)^n x_0$$

При известном начальном значении x_0 можно рассчитать динамику популяции во времени. В зависимости от коэффициента роста $r = a + 1$ возможны следующие ситуации:

- 1) $a > 0, (1+a) > 1 \Rightarrow (1+a)^n \rightarrow \infty$ при $n \rightarrow \infty$ – неограниченный рост;
- 2) $a = 0, 1+a=1$ – численность популяции не меняется;
- 3) $-1 < a < 0, 0 < 1+a < 1 \Rightarrow x_n \rightarrow 0$ – вымирание популяции;
- 4) $a = -1$ – вымирание за один период времени;
- 5) $a < -1$ – отрицательные численности (нереальная ситуация).

Уравнение (1) обычно записывается в форме

$$x_{n+1} = r x_n$$

Уравнение (2) называется дискретным аналогом экспоненциальной модели одиночной популяции.

- 2) **Построение дискретной модели ограниченной популяции: логистическое уравнение.** Одним из классических примеров дискретных моделей является логистическое разностное уравнение для одиночной лимитированной популяции. Получаемая при этом динамика системы включает в себя все многообразие типов поведения реальных популяционных систем, как простых или упорядоченных, так и сложных или хаотических. Переход от простого поведения к хаосу обладает схожими закономерностями, присущими различным моделям популяций или различным сложным динамическим системам. Исследуем разностную модель, задаваемую логистическим уравнением.

Как было отмечено ранее, дискретный аналог экспоненциальной модели (2) одиночной нелимитированной популяции при коэффициенте прироста предсказывает неограниченный рост численности популяции. В реальности ни одна популяция не может увеличиваться бесконечно вследствие ограниченности пищевых ресурсов и других ограничивающих внешних факторов. Для учета этого обстоятельства введем условие ограничения роста. Пусть коэффициент прироста r будет зависеть от численности популяции, а именно, будет убывать по мере роста численности популяции по закону $r \sim r(1 - x_n)$. Тогда уравнение (2) примет вид

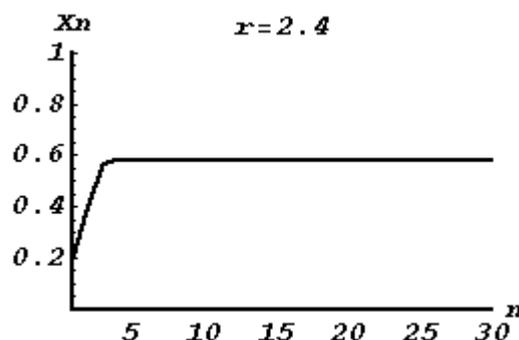
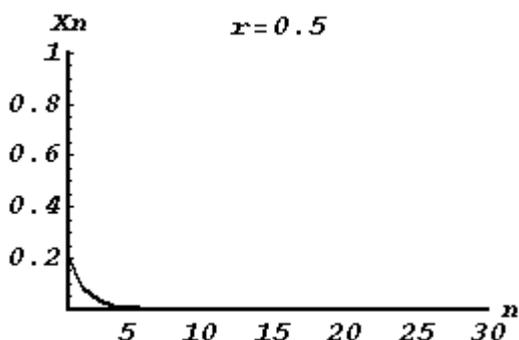
$$x_{n+1} = x_n r (1 - x_n), n=0,1,2,3,\dots \quad (3)$$

Уравнение (3), называемое логистическим уравнением или дискретным аналогом модели Ферхюльста–Пирла, может описывать не только динамику популяций, но и многие другие явления в природе и обществе. Отметим, что величина x в уравнении (3) меняется от 0 до 1, а r от 0 до 4. При других значениях x и r логистическое уравнение дает отрицательные значения численности популяции.

Задавая различные значения параметра r естественной скорости роста и начальной численности x_0 популяции, можно получить качественно различные типы поведения переменной, удовлетворяющие разностному уравнению (3). Разностное уравнение наряду с равновесием и циклами может иметь хаотические решения, не стремящиеся ни к какому притягивающему решению.

На рис. 1 показаны графики зависимости от номера периода времени для разных значений параметра r ($x_0=0,2$). Для значения $r = 0,5$ популяция за несколько периодов времени приходит к вымиранию (нулевое значение x_n при $n \rightarrow \infty$). При $r = 2,4$ имеем стационарное значение численности $x_n=0.5833$. Из рис. 1 видно, что при $r=3,33$ конечное значение численности популяции начинает осциллировать между двумя уровнями, которые соответствуют значениям 0.829635 и 0.470666, то есть мы имеем цикл с периодом 2. С ростом r динамика системы усложняется.

Для $r=3,5$ процесс приходит к устойчивым периодическим колебаниям с периодом 4 (установившиеся значения численности $x_n = 0.874997; 0.500887; 0.826939; 0.382818$). И, наконец, при $r=3,9$ можно наблюдать, что процесс перестал быть периодическим. При увеличении значения номера периода времени численность популяции принимает новые неповторяющиеся значения. Такое поведение называется нерегулярным или хаотическим.



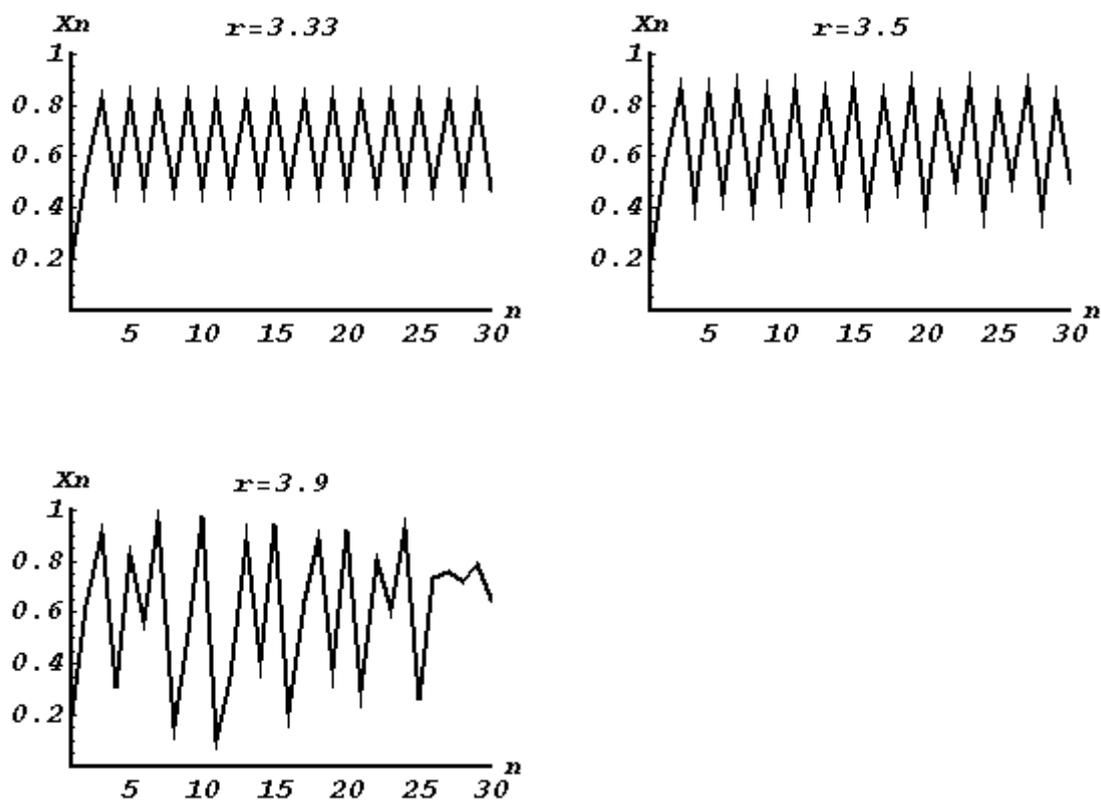


Рис.1. Зависимость x_n от периода времени n для различных значений параметра r

3) Практическая работа:

Решение предложенных задач по построению дискретных моделей роста популяции.

Рекомендуемая литература:

Основная литература:

1. Ризниченко Г.Ю., Рубин А.Б. Математические модели биологических продукционных процессов. М. МГУ. 1993.
2. Кипятков В.Е. Практикум по математическому моделированию в популяционной экологии (учебное пособие). СПб, 2002.

Дополнительная литература:

1. Быков А.А. Моделирование природоохранной деятельности. Учебн. пособие. М., Изд-во НУМЦ Госкомэкологии России, 2008.
2. Вольтерра В. Математическая теория борьбы за существование / В. Вольтерра. – М. : Наука, 1976. – 285 с.

Методические указания для аспирантов
по дисциплине Б1.В.ДВ.1 Математическое моделирование
в экологии
к практическому занятию № 2

Тема: Моделирование процесса межвидовой конкуренции.

Цель занятия: Сформировать у аспирантов представление о математических моделях, описывающих межвидовую конкуренцию.

Задачи: Изучить особенности моделей, описывающих межвидовую конкуренцию, их основные свойства и область применения.

Аспирант должен знать:

1) до изучения темы:

- Понятие математических моделей;
- Виды математических моделей биологических процессов;

2) после изучения темы:

- особенности моделей, описывающих межвидовую конкуренцию;
- основные свойства и область применения моделей, описывающих межвидовую конкуренцию.

Изучение темы занятия направлено на формирование профессиональных компетенций: способность использовать современные методы обработки, анализа и синтеза полевой и лабораторной биологической информации, знание принципов составления научно-технических проектов и отчетов (ПК-5).

Методика проведения занятия.

1. Определение темы занятия. Преподаватель поясняет цели и задачи занятия, значение полученных знаний для будущей работы по специальности.

2. Теоретическая часть.

Системы, изученные Вольтерра, состоят из нескольких биологических видов и запаса пищи, который используют некоторые из рассматриваемых видов. О компонентах системы формулируются следующие допущения.

1. Пища либо имеется в неограниченном количестве, либо ее поступление с течением времени жестко регламентировано. 2. Особи каждого вида отмирают так, что в единицу времени погибает постоянная доля существующих особей. 3. Хищные виды поедают жертвы, причем в единицу времени количество съеденных жертв всегда пропорционально вероятности встречи особей этих двух видов, т.е. произведению количества хищников на количество жертв. 4. Если имеются пища в неограниченном количестве и несколько видов, которые способны ее потреблять, то доля

пищи, потребляемая каждым видом в единицу времени, пропорциональна количеству особей этого вида, взятого с некоторым коэффициентом, зависящим от вида (модели межвидовой конкуренции). 5. Если вид питается пищей, имеющейся в неограниченном количестве, прирост численности вида за единицу времени пропорционален численности вида. 6. Если вид питается пищей, имеющейся в ограниченном количестве, то его размножение регулируется скоростью потребления пищи, т.е. за единицу времени прирост пропорционален количеству съеденной пищи.

Перечисленные гипотезы позволяют описывать сложные живые системы при помощи систем обыкновенных дифференциальных уравнений, в правых частях которых имеются суммы линейных и билинейных членов. Как известно, такими уравнениями описываются и системы химических реакций. Такое сходство уравнений в химических и экологических моделях позволяет применить для математического моделирования кинетики популяций те же методы исследований, что и для систем химических реакций. Вольтерровские уравнения могут быть получены не только из локального "принципа встреч", ведущего свое происхождение из статистической физики, но и исходя из баланса масс каждого из компонентов биогеоценоза и энергетических потоков между этими компонентами.

3) Практическая работа:

Решение предложенных задач.

Рекомендуемая литература:

Основная литература:

3. Ризниченко Г.Ю., Рубин А.Б. Математические модели биологических продукционных процессов. М. МГУ. 1993.
4. Кипятков В.Е. Практикум по математическому моделированию в популяционной экологии (учебное пособие). СПб, 2002.

Дополнительная литература:

3. Быков А.А. Моделирование природоохранной деятельности. Учебн. пособие. М., Изд-во НУМЦ Госкомэкологии России, 2008.
4. Вольтерра В. Математическая теория борьбы за существование / В. Вольтерра. – М. : Наука, 1976. – 285 с.

Методические указания для аспирантов
по дисциплине Б1.В.ДВ.1 Математическое моделирование
в экологии
к практическому занятию № 3

Тема: Моделирование взаимодействия популяций хищника и жертвы.

Цель занятия: Сформировать у аспирантов представление математических моделях, описывающих взаимодействие популяций хищника и жертвы.

Задачи: Освоить методы построения моделей, описывающих взаимодействие популяций хищника и жертвы.

Аспирант должен знать:

1) до изучения темы:

- Основные виды математических моделей популяционной динамики;

2) после изучения темы:

- основные особенности моделей, описывающих взаимодействие хищника и жертвы;
- свойства и ограничения моделей, описывающих взаимодействие хищника и жертвы.

Изучение темы занятия направлено на формирование профессиональных компетенций: способность использовать современные методы обработки, анализа и синтеза полевой и лабораторной биологической информации, знание принципов составления научно-технических проектов и отчетов (ПК-5).

Методика проведения занятия.

1. Определение темы занятия. Преподаватель поясняет цели и задачи занятия, значение полученных знаний для будущей работы по специальности.

2. Теоретическая часть.

Допустим, что на некоторой территории обитают два вида животных: кролики (питающиеся растениями) и лисы (питающиеся кроликами). Пусть число кроликов x , число лис y . Используя Модель Мальтуса с необходимыми поправками, учитывающими поедание кроликов лисами, приходим к следующей системе, носящей имя модели Вольтерры — Лотки:

$$\begin{cases} \dot{x} = (\alpha - cy)x; \\ \dot{y} = (-\beta + dx)y. \end{cases}$$

Эта система имеет равновесное состояние, когда число кроликов и лис постоянно. Отклонение от этого состояния приводит к колебаниям численности кроликов и лис, аналогичным колебаниям гармонического осциллятора. Как и в случае гармонического осциллятора, это поведение не является структурно устойчивым: малое изменение модели (например, учитывающее ограниченность ресурсов, необходимых кроликам) может привести к качественному изменению поведения. Например, равновесное состояние может стать устойчивым, и колебания численности будут затухать. Возможна и противоположная ситуация, когда любое малое отклонение от положения равновесия приведет к катастрофическим последствиям, вплоть до полного вымирания одного из видов.

Поскольку в реальных популяциях присутствует много возмущающих факторов, не учтенных в модели Лотки — Вольтерры, эта модель вряд ли может претендовать на адекватное описание реальности. Этому недостатка лишена модель Холлинга — Тэннера, учитывающая большее число реальных факторов. В этой модели скорость изменения популяции хищников задается выражением $ax - bx^2/y = x(a - bx/y)$. Оно выбрано из следующих соображений. Когда пищи (жертв) много ($y \approx +\infty$), популяция хищников растет по правилу Мальтуса с показателем a . С уменьшением числа жертв скорость роста популяции хищников падает и при $y < bx/a$ становится отрицательной (последнее, грубо говоря, является следствием предположения, что для поддержания жизни одного хищника необходимо $k = b/a$ жертв).

Скорость изменения популяции жертв состоит из трех компонент. Первый член cy соответствует закону Мальтуса, второй $-dy^2$ описывает внутривидовую конкуренцию и вызван ограниченностью ресурсов экологической ниши, занимаемой популяцией жертв. При отсутствии хищников жертвы подчиняются уравнению

$$y' = y(c - dy).$$

Наконец, третий компонент скорости изменения популяции жертв в модели Холлинга-Тэннера описывает ее взаимодействие с хищниками и имеет вид $-pxy/(q + y)$ ($p, q > 0$). Это выражение правдоподобнее описывает межвидовое взаимодействие, нежели соответствующий член $-dxy$ модели Лотки — Вольтерры. В последней число жертв, убиваемых одним хищником за единицу времени, равно dy и растет пропорционально числу жертв, что неправдоподобно. В модели Холлинга — Тэннера коэффициент хищничества равен $px/(q + y)$. Он не может превышать величины p/q и при неограниченном росте популяции жертв стремится, монотонно возрастая, к числу p/q ,

выражающему естественную потребность хищников в пище. В результате получается следующая система уравнений (*модель Холлинга-Тэннера*)

$$x' = (a - bx/y)x, \quad 3)$$

$$y' = [c - dy - px/(q + y)]y.$$

3) Практическая работа:

Решение предложенных задач.

Рекомендуемая литература:

Основная литература:

1. Ризниченко Г.Ю., Рубин А.Б. Математические модели биологических продукционных процессов. М. МГУ. 1993.
2. Кипятков В.Е. Практикум по математическому моделированию в популяционной экологии (учебное пособие). СПб, 2002.

Дополнительная литература:

1. Быков А.А. Моделирование природоохранной деятельности. Учебн. пособие. М., Изд-во НУМЦ Госкомэкологии России, 2008.
2. Вольтерра В. Математическая теория борьбы за существование / В. Вольтерра. – М. : Наука, 1976. – 285 с.

Методические указания для аспирантов
по дисциплине Б1.В.ДВ.1 Математическое моделирование
в экологии
к практическому занятию № 4

Тема: Модели для фитопланктонных сообществ.

Цель занятия: Сформировать у аспирантов представление о математических моделях, описывающих взаимодействие организмов в водной экосистеме.

Задачи: Освоить методы построения моделей водных экосистем.

Аспирант должен знать:

1) до изучения темы:

- Понятие математических моделей;
- Виды математических моделей биологических процессов;

2) после изучения темы:

- Дискретные модели неограниченной одиночной популяции;
- Дискретные модели ограниченной популяции.

Изучение темы занятия направлено на формирование профессиональных компетенций: способность использовать современные методы обработки, анализа и синтеза полевой и лабораторной биологической информации, знание принципов составления научно-технических проектов и отчетов (ПК-5).

Методика проведения занятия.

1. Определение темы занятия. Преподаватель поясняет цели и задачи занятия, значение полученных знаний для будущей работы по специальности.

2. Теоретическая часть.

Модели водных экосистем занимают большое место в математической экологии, в первую очередь потому, что водная среда гораздо более гомогенна, чем суша, ее легче изучать и моделировать. Значительная доля гидробионтов, в первую очередь фитопланктон, являются микроорганизмами, к ним применимы многие методы математического моделирования разработанные и экспериментально проверенные на микробных популяциях.

Водные системы дают людям, животным, сельскому хозяйству и промышленности воду. Океаны, моря и реки обеспечивают в разных странах от 20% до 80% потребности людей в белковой пище. Однако качество воды в водоемах и их продуктивность неожиданно и резко падает. Это связано в первую очередь с тем, что водоемы традиционно использовались людьми как бесплатные системы по переработке отходов, что привело к их существенному загрязнению, нарушению естественных биологических и химических процессов. Потребности оптимизации использования водных систем и понимания происходящих в них процессов привели к быстрому развитию математического моделирования водных систем. В настоящее время насчитываются тысячи моделей разной степени сложности и подробности. Планирование любого водохозяйственного мероприятия сопровождается и предваряется построением математической модели водной системы.

В 70-80 годы особенно активно развивались модели озерных экосистем. (Jorgensen S.E. Lake management. Oxford, 1980). Одной из важнейших задач была выработка борьбы с эвтрификацией - "цветением" озер в связи с увеличением количества поступающего в них органического вещества, а также биогенных веществ, в первую очередь азота, вместе со стоками вод из сельскохозяйственных угодий. Озеро представляет собой

относительно замкнутую экосистему, поэтому моделирование потоков вещества и энергии в ней обычно проводится путем выделения нескольких круговоротов, обладающих различными характерными временами. Это быстрый первичный кругооборот (фито- и бактериопланктон, легко окисляющиеся органические вещества и минеральный субстрат); вторичный кругооборот – бактерии, мирный зоопланктон (фильтраторы) и некоторые виды хищного зоопланктона, высшие трофические уровни - консументы. Наконец, медленный кругооборот представляют относительно консервативные компоненты: трудноокисляемое органическое вещество (водный гумус), донные отложения, популяции долгоживущих гидробионтов.

Учет иерархии времен отдельных круговоротов позволяет представить озерную экосистему в виде своеобразной "матрешки" - вложенных друг в друга процессов. При этом определяющим является первичный кругооборот, систему которого на малых временах можно считать замкнутой. Более медленные процессы можно рассматривать как его возмущения.

Математические модели помогают разработать оптимальную стратегию управления водными ресурсами, в том числе рыбным хозяйством. Дело в том, что наряду с ухудшением состояния воды причиной падения продуктивности водоемов являются систематические переловы. В биологическом смысле они приводят к такому состоянию рыбного стада, когда воспроизводительная способность популяции не может компенсировать убыль в результате вылова. Перелов в экономическом смысле – это сокращение поголовья рыбного стада настолько, что промысел становится нерентабельным.

Решение задачи оптимизации систематического лова рыбы восходит к работам Баранова (1918). Представив коэффициенты общей смертности в виде суммы коэффициентов естественной и промысловой гибели в формуле численности рыбного стада, Баранов оценил величину улова и смог подойти к постановке задачи оптимального вылова. Значительный шаг в решении этой проблемы сделали Риккер (1958) и Бивертон и Холт (1957), связавшие модели с конкретным статистическим материалом рыбоводства и ихтиологии и предложившие методики решения задач управления.

Особенно большой вклад в моделирование рыбных популяций внес В.В.Меншуткин, ("Математическое моделирование популяций и сообществ водных животных", Л.,1971), который представил схему взаимодействий в водной экосистеме как контур с обратными связями. Такая система может обладать устойчивым стационарным состоянием, в ней могут возникать колебательные или квазистохастические режимы. Подобные схемы, часто

весьма детальные, были положены в основу моделей рыбного стада многих озер и морей.

Научную базу **описания обменных процессов водных экосистем** дали работы основоположника математической экологии Алексея Андреевича Ляпунова, крупнейшего русского ученого, стоящего у истоков также и других областей математического моделирования, в том числе математической лингвистики. В работах Ляпунова впервые в одной модели были объединены физические (гидродинамические) и биологические (хищничество) процессы. А.А. Ляпунов подчеркивал важность для экосистем как физической (поглощение энергии света), так и биологической (образование биомассы) роли фотосинтеза. Впервые идеология такого моделирования была разработана в модели экосистемы пелагиали тропических вод океана для 44 рейса научно-исследовательского судна Витязь, задача которого состояла в изучении продуктивности этой системы в рамках международной биологической программы. (Ляпунов, 1972).

В настоящее время аналогичные модели потоков вещества и энергии используются для анализа процессов океанизации и эвтрофикации окраинных морей и описания распределения планктона в различных районах мирового океана. Последние десятилетия для океанологов и лимнологов стали доступными результаты дистанционного зондирования вод океанов и морей в видимой части спектра. Такие наблюдения дают возможность оценить концентрацию хлорофилла в поверхностном слое и на основе статистических методов оценить пространственную концентрацию фитопланктона. Таким образом, теоретические модели, основу которых заложил А.А. Ляпунов, развиваются и наполняются конкретным содержанием. Так среднемесячные поля концентрации хлорофилла рассчитаны для периода 1978-1984 гг. (Esaías et.al. 1986 и более поздние работы) на основе измерений, проведенных сканером CZCS, установленным на борту спутника "Нимбус-7" (США). Использование этих данных позволило провести статистический анализ сезонного цикла первичной продукции для северной части Атлантического океана и оценить величины глобальной фотосинтетической первичной продукции.

Оптическая активность пигментов, содержащихся в клетках фитопланктона во многом формирует свойства гидрооптических полей. Это явление служит основой разработки оптических методов исследований распределения и свойств полей фитопланктона с помощью дистанционного зондирования с борта судна, самолетов или спутниковых платформ.

Совокупность методов исследования экосистем, в первую очередь водных, с помощью оптических методов принято называть биооптикой. С

ней тесно связаны методы анализа влияния биологических объектов на формирование полей освещенности с использованием методов гидрооптики и гидрофизики, с одной стороны, и гидробиологии, биофизики, теории популяций, с другой. На примере этого быстро развивающегося направления науки особенно наглядно проявляется междисциплинарность современного знания о сложных системах, включающих в себя процессы физической, химической и биологической природы.

Экодинамические модели, аккумулирующие данные, полученные с помощью дистанционных методов наблюдений, обычно содержат подмодели: а) популяционной динамики, включающая в себя алгоритм для вычисления фотосинтетической первичной продукции; б) физическую (гидродинамическую модель переноса и диффузии и в) модель формирования подводного оптического поля. Для решения гидродинамической задачи в идеале строится специальная трехмерная гидродинамическая модель течений и горизонтального и вертикального перемешивания. Для моделирования динамики органического вещества, в частности, численности фитопланктона, важным является учет конкуренции и хищничества

В моделях выделены биологические виды фитопланктона и группы минеральных питательных веществ. Фитопланктон представлен m видами, их содержание в среде обозначено y_i для вида i . Минеральное питание растительных организмов разбивается на n групп сходных веществ (на основе азота, фосфора, кремния и т.п.). В рассматриваемых моделях питательные вещества предполагаются не взаимозаменяемыми. Содержание веществ группы j в среде обозначается z_j . Рост биомассы клеток вида i происходит с удельной скоростью $\mu_i(z)$ в зависимости от содержания z биогенов во внешней среде. Под y понимается вектор с компонентами y_i , а под z - вектор с компонентами z_j . Одна из моделей сообщества в хемостате имеет вид:

$$\begin{cases} \frac{dy_i}{dt} = (\mu_i(z) - D)y_i \\ \frac{dz_j}{dt} = D(z_{j0} - z_j) - \sum_{i=1}^m v_{ij}(z_j)y_i \end{cases} \quad \text{для } i = 1, \dots, m; j = 1, \dots, n. \quad (1)$$

Через D обозначена скорость протока вещества в системе, через z_0 - содержание минеральных питательных веществ во входящем потоке, через $v_{ij}(z_j)$ - удельные скорости поглощения вещества группы j организмами вида i .

Сравним свойства решений в модели (1) с замкнутой моделью. Эта модель описывает динамику биомасс основных групп фитопланктона, минеральных веществ и отмершей органики. Сначала рассмотрим модель без различия минеральных веществ по группам, эти вещества и отмершая органика представлены одной группой каждая. Блок отмершей органики с массой s введен для описания замкнутого цикла преобразования веществ. Функция $r(s)$ описывает скорость преобразования органики в минеральные соединения при бактериальном разложении. Удельные скорости элиминации микроорганизмов задаются функциями $e_i(y_i)$ их содержания в среде.

Элиминация включает в себя процессы естественной смертности, внутривидовой конкуренции и изъятия особей из системы по иным причинам. Но вместе с тем принято предположение, что отмершая органика остается внутри системы (замкнутость по веществу). Остальные обозначения соответствуют предыдущей модели, опущены индексы там, где они не нужны. Тогда замкнутая модель приобретает вид системы дифференциальных уравнений:

$$\begin{cases} \frac{dy_i}{dt} = \mu_i(z)y_i - e_i(y_i)y_i \\ \frac{dz}{dt} = r(s) - \sum_{i=1}^m \mu_i(z)y_i \\ \frac{ds}{dt} = \sum_{i=1}^m e_i(y_i)y_i - r(s) \end{cases} \quad \text{для } i = 1, \dots, m. \quad (2)$$

3) Практическая работа:

Решение предложенных задач.

Рекомендуемая литература:

Основная литература:

1. Ризниченко Г.Ю., Рубин А.Б. Математические модели биологических продукционных процессов. М. МГУ. 1993.
2. Кипятков В.Е. Практикум по математическому моделированию в популяционной экологии (учебное пособие). СПб, 2002.

Дополнительная литература:

1. Быков А.А. Моделирование природоохранной деятельности. Учебн. пособие. М., Изд-во НУМЦ Госкомэкологии России, 2008.
2. Вольтерра В. Математическая теория борьбы за существование / В. Вольтерра. – М. : Наука, 1976. – 285 с.

Методические указания для аспирантов
по дисциплине Б1.В.ДВ.1 Математическое моделирование
в экологии
к практическому занятию № 5

Тема: Модели продукционного процесса растений.

Цель занятия: Сформировать у аспирантов представление о математических моделях, описывающих фитоценозы.

Задачи: Освоить методы построения моделей растительных сообществ.

Аспирант должен знать:

1) до изучения темы:

- Понятие математических моделей;
- Виды математических моделей биологических процессов;

2) после изучения темы:

- Методы управления агробиоценозом с помощью математического моделирования.

Изучение темы занятия направлено на формирование профессиональных компетенций: способность использовать современные методы обработки, анализа и синтеза полевой и лабораторной биологической информации, знание принципов составления научно-технических проектов и отчетов (ПК-5).

Методика проведения занятия.

1. Определение темы занятия. Преподаватель поясняет цели и задачи занятия, значение полученных знаний для будущей работы по специальности.

2. Теоретическая часть.

Одной из наиболее продвинутых областей в математической экологии является моделирование продукционного процесса растений. Это определяется практической значимостью таких моделей для оптимизации агрокультуры и тепличного хозяйства. Здесь математические модели используются для выбора оптимальной стратегии проведения сельскохозяйственных мероприятий: орошения, полива, внесения удобрений, выбора сроков посева или посадки растений с целью получения максимального урожая. Для полностью контролируемого тепличного хозяйства возможно построение модели, описывающей весь цикл процессов при заданных условиях. Тогда с помощью модели оптимальный "рецепт" управления культурой может быть задан полностью на все время вегетации.

Если же моделируется посев в открытом грунте, на который оказывают влияние непредсказуемые погодные условия, агробиоценоз нуждается в оперативном управлении, для него используются динамические модели, допускающие оперативное изменение параметров и, возможно, структуры модели в соответствии с изменениями погодных условий.

Всю систему происходящих в агробиоценозе процессов обычно представляют в виде блочной иерархической структуры. Выделяются биотический и абиотический блоки. Среди биотических процессов выделяют в отдельные блоки рост и развитие посева, функционирование почвенной микрофлоры, развитие энтомофауны, развитие болезней сельскохозяйственных культур, взаимодействие посева с сорняками и др.

Абиотические блоки включают в себя модели, описывающие формирование теплового, водного режима почвы и приземных слоев воздуха, концентрации и передвижения биогенных и токсических солей, различных остатков распада пестицидов, ростовых веществ и метаболитов в почве, концентрации углекислого газа в посевах. Пример блок-схемы модели продуктивности агроэкосистемы приведен на рис. 1.

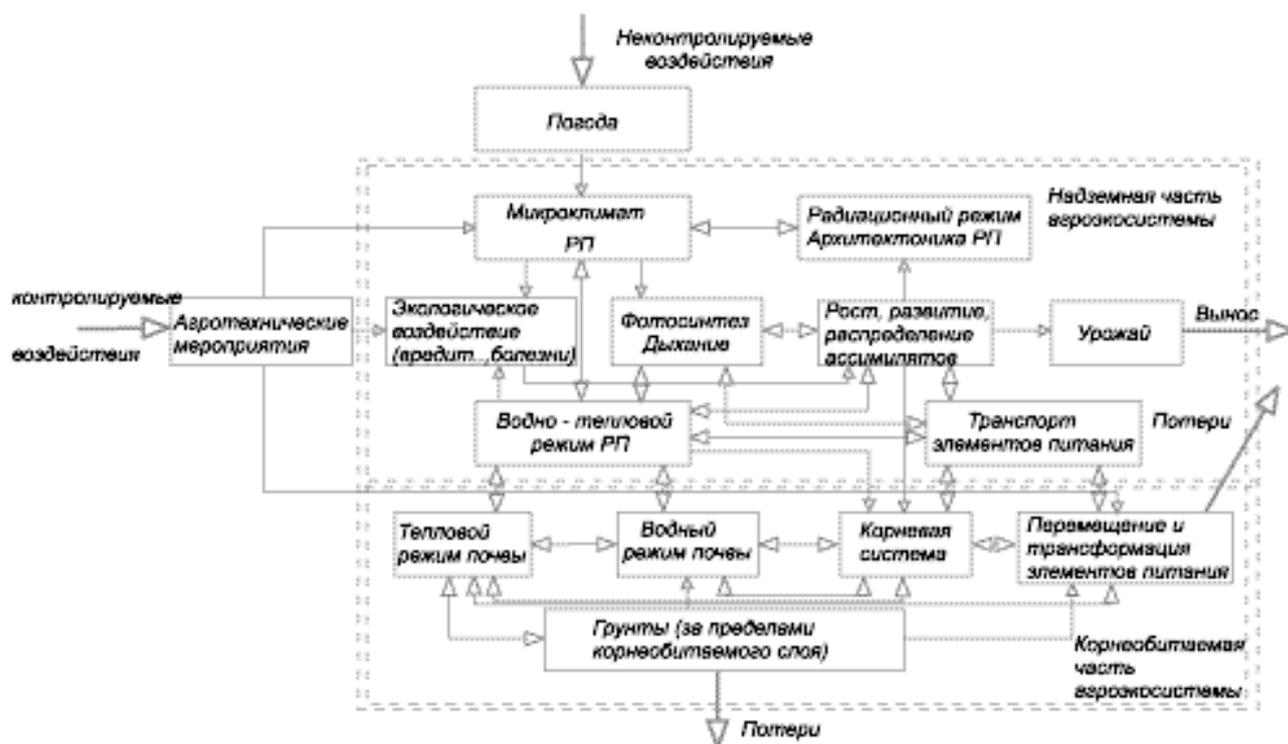


Рис.1. Схема взаимодействия процессов в агробиоценозе пшеницы.

Блочная структура позволяет изучать, изменять и детализировать одни блоки, не меняя других. Как правило, число параметров внутри блоков существенно больше числа параметров, которыми блоки соединяются между собой. На основе блоков синтезируются целостные динамические модели,

способные прогнозировать изменение во времени ряда характерных параметров растений, в первую очередь биомассу всего растения и отдельных органов, начиная от всходов (иногда от момента посева) до завершения вегетации (созревания). Первые такие модели были разработаны коллективом американских авторов (SPAM - Soil-Plant-Atmosphere Model, Schawerofft et.al., 1974) и де Витом и его группой (BESCROP - Basic Crop Simulation, De Wit, 1978). В настоящее время имеется несколько десятков такого типа моделей формирования урожая, разработанных с разной степенью детализации для сои, пшеницы, трав, кукурузы, хлопчатника и других культур.

Структура и сложность модели продукционного процесса растений, степень ее детализации, форма представления процессов, происходящих в растении, определяются двумя обстоятельствами: предметом и целью моделирования. Модель роста травы, биомассу которой можно считать однородной, предназначенной для корма скота, может быть существенно проще, чем модели культур, урожай которых заключен в репродуктивных органах (злаки, бобовые) или корнеплодах. Для практических целей удобнее простая модель, позволяющая давать прогноз урожая при определенных погодных условиях или рекомендации по оптимальному режиму полива и внесения минеральных удобрений. Изучение физиологических особенностей растений и их реакций на почвенные и погодные условия требует построения сложных моделей с блочной структурой.

На рис. 2 представлена блок схема модели посева сои, которая представляет имитационное описание роста, развития и урожайности сои и считается наиболее подробной из разработанных к настоящему времени моделей сельскохозяйственных культур. В ней объединены несколько субмоделей и большое количество входных данных. В модели SOYMOD четыре: категории сухого вещества: структурные углеводы, доступные углеводы (неустойчивые соединения, которые могут передвигаться по растению), крахмал и белок. Эти вещества распределяются между различными морфологическими частями: пластинками и черешками листьев, плодами и корнями. Предполагается, что необходимые для процессов роста и жизнедеятельности растения материалы – это азот и углерод. Они перемещаются между морфологическими частями растения и реализуются в этих частях для роста, дыхания, образования новых органов, транспортных процессов. Соотношение между углеродом и азотом используют в качестве функций контроля за ростом различных частей растения. В этом смысле модель сои относится к моделям углеродно-азотного типа.

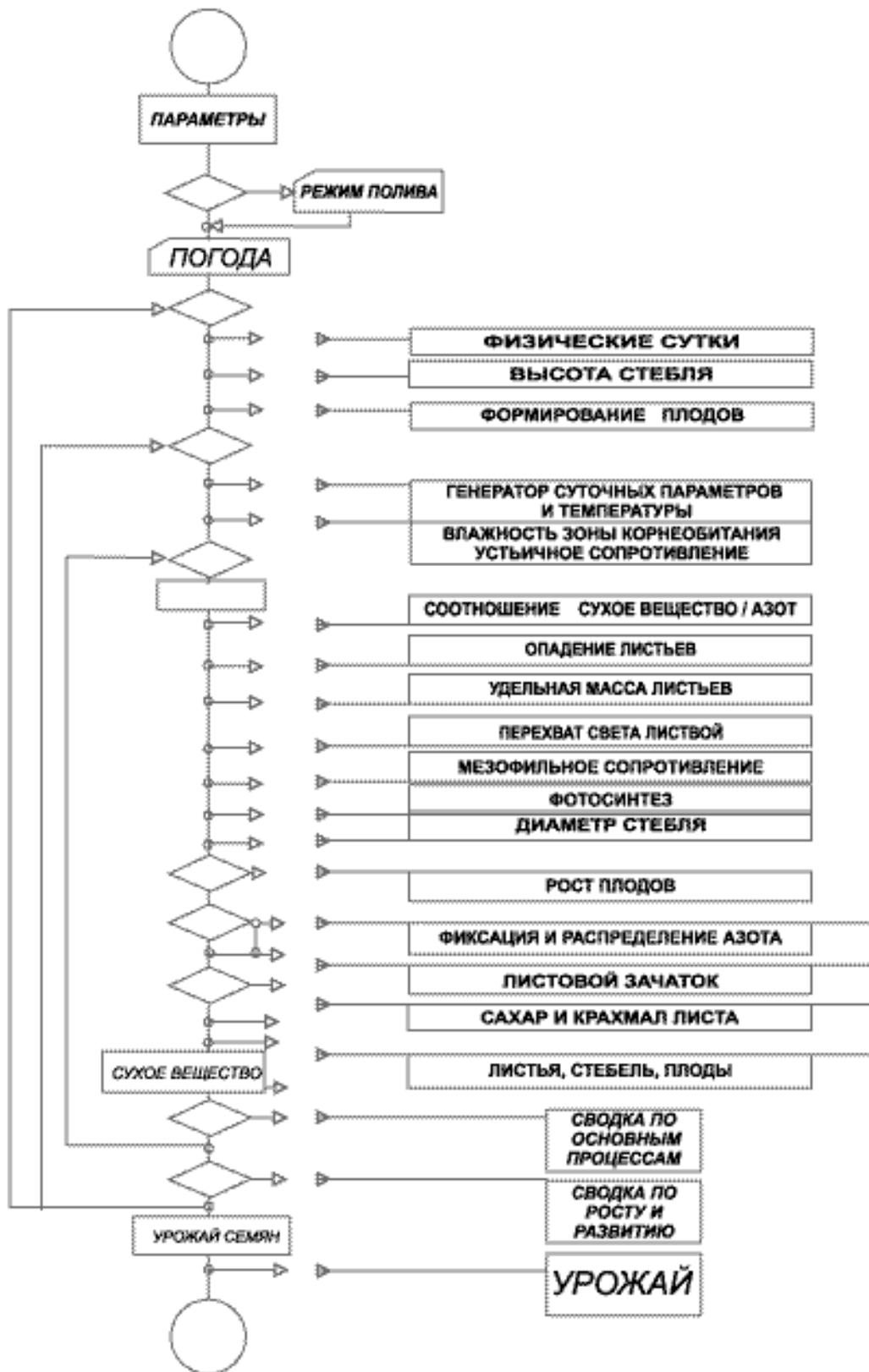


Рис. 2. Блок-схема модели сои.

Модель агрофитоценоза пшеницы (система Симона) - наиболее детальная отечественная имитационная модель продукционного процесса сельскохозяйственных растений разработана под руководством Р.А. Полуэктова в Петербургском агрофизическом институте. С математической

точки зрения модель представляет собой систему из нескольких уравнений в частных производных параболического типа и нескольких десятков обыкновенных дифференциальных уравнений. При переходе к численной схеме выбирают шаг интегрирования по координате x и по времени t . Базовый временной шаг модели выбран равным одному часу, это позволяет имитировать суточный ход как абиотических (энергообмен), так и биотических (фотосинтез, метаболизм) процессов. Состояние абиотической части системы характеризуется набором вертикально распределенных переменных: радиации, температуры, и влажности воздуха в посевах, температуры и влажности почвы и др. Биологическая часть системы представлена переменными: плотность ассимилирующей поверхности фитоэлементов, поглощающей поверхности корней, плотность отдельных составляющих биомассы (углеводы, аминокислоты, белки) и фитомассы в целом и др.

В модель включено описание процессов трех типов (рис. 1): 1) энерго- и массообмен, происходящий в среде обитания растений (в почве и приземном воздухе) и в самих растениях; 2) совокупность биофизических и физиологических процессов в растительном покрове, определяющих прирост биомассы, рост и развитие отдельных органов растения и формирования урожая; 3) экологическое взаимодействие культурных растений с сорняками, болезнетворными микроорганизмами и вредителями.

В качестве входных переменных выступают контролируемые (агротехника) и неконтролируемые (погода) внешние воздействия. Динамика погодных условий представлена реализациями многомерного случайного процесса. Отдельные блоки модели подробно описаны в монографиях: (Бондаренко и др., 1982; Заславский, Полуэктов, 1988). Каждый из блоков представляет собой описание группы однородных физических, биофизических, биохимических или физиологических процессов в отдельных частях системы почва-растение-атмосфера. Каждый из блоков решает свою математическую задачу, и может быть верифицирован (т.е. проведена проверка правильности его работы) на независимых массивах экспериментальных данных. Объединение блоков в целостную систему и возможная последующая редукция (упрощение) этой системы представляет собой также самостоятельную задачу из области теории сложных систем.

Модели типа SOYMOD или СИМОНА слишком сложны для использования в практике. В своем полном объеме такие модели, действительно, служат исследовательским целям, причем они непрерывно развиваются, их структура и значения параметров уточняются с использованием новейших данных о характеристиках моделируемой

системы. Их тщательный анализ открывает путь и для практических применений. Потребитель может вести диалог для решения конкретных практических вопросов сельскохозяйственного производства. Например, запрашивать возможные пределы изменения урожая для конкретного поля, задавая величину и сроки выпадения осадков или решать оптимизационные задачи о сроках и дозах внесения удобрений.

3) Практическая работа:

Решение предложенных задач.

Рекомендуемая литература:

Основная литература:

1. Ризниченко Г.Ю., Рубин А.Б. Математические модели биологических продукционных процессов. М. МГУ. 1993.
2. Кипятков В.Е. Практикум по математическому моделированию в популяционной экологии (учебное пособие). СПб, 2002.

Дополнительная литература:

1. Быков А.А. Моделирование природоохранной деятельности. Учебн. пособие. М., Изд-во НУМЦ Госкомэкологии России, 2008.
2. Вольтерра В. Математическая теория борьбы за существование / В. Вольтерра. – М. : Наука, 1976. – 285 с.

Методические указания для аспирантов
по дисциплине Б1.В.ДВ.1 Математическое моделирование
в экологии
к практическому занятию № 6

Тема: Моделирование экосистем с учетом лимитирующего фактора.

Цель занятия: Сформировать у аспирантов представление о математических моделях, описывающих системы с лимитирующими факторами.

Задачи: Освоить методы выделения лимитирующего звена.

Аспирант должен знать:

- 1) до изучения темы:
 - Понятие лимитирующих факторов;
 - Виды математических моделей биологических процессов;
- 2) после изучения темы:

- Математическую теорию описания систем с лимитирующими факторами;
- Понятие об Л-системах.

Изучение темы занятия направлено на формирование профессиональных компетенций: способность использовать современные методы обработки, анализа и синтеза полевой и лабораторной биологической информации, знание принципов составления научно-технических проектов и отчетов (ПК-5).

Методика проведения занятия.

1. Определение темы занятия. Преподаватель поясняет цели и задачи занятия, значение полученных знаний для будущей работы по специальности.

2. Теоретическая часть.

Концепция лимитирующих факторов принадлежит немецкому агрохимику Юстусу Либиху, который предложил знаменитый закон минимума: "Каждое поле содержит одно или несколько питательных веществ в минимуме и одно или несколько других в максимуме. Урожай находится в соответствии с этим минимумом питательных веществ". Либих понимал под этим относительный минимум питательного вещества по сравнению с содержанием других веществ. Позже в экологической литературе фактор, находящийся в минимуме, стали называть лимитирующим фактором. Закон "лимитирующего фактора" для фотосинтетических процессов в 1905 г. предложил Ф.Блэкман, а в 1965 г. Н.Д. Иерусалимский сформулировал этот закон для ферментативных процессов. Естественно, что при изменении соотношений факторов, лимитирующий фактор может изменяться.

Математическая теория описания систем с лимитирующими факторами разработана И.А. Полетаевым и его школой (Полетаев, 1966, Гильдерман, 1970). Обобщение принципа лимитирующего фактора с использованием теоретико-множественных представлений и разработка методов выделения лимитирующего звена с энтропийных экстремальных принципов дана в работах А.П. Левича (1988, 1994). Разработан класс дискретно непрерывных моделей, описываемых системами дифференциальных уравнений, структура которых меняется в определенные моменты времени, а последние в свою очередь зависят от значения решений самих уравнений. И.А. Полетаевым эти системы названы системами с лимитирующим фактором, или Л-системами.

Самым простым примером Л-системы служит химическая реакция соединения нескольких веществ в одно, происходящая практически мгновенно. Вещества-реагенты могут взаимодействовать лишь в строго

определенных стехиометрических соотношениях, и поэтому скорость рассматриваемой реакции определяется самой медленной в данный момент времени из скоростей притока каждого из исходных веществ в сферу реакции. Иными словами, свойства Л-системы определяются наличием узкого места и его сменой в процессе функционирования.

Метод описания элемента экологической модели как Л-системы можно проиллюстрировать на примере простой модели роста растения (Полетаев, 1973). Основные процессы превращения вещества и энергии в растениях, это – процесс возрастания биомассы (роста), идущий за счет фотосинтеза и поглощения веществ из почвы и атмосферы, и процесс основного обмена (дыхания), поставляющий свободную энергию для жизнедеятельности биомассы за счет частичного расходования вещества, содержащегося в самой биомассе. Блок-схема растения изображена на рис. 1.

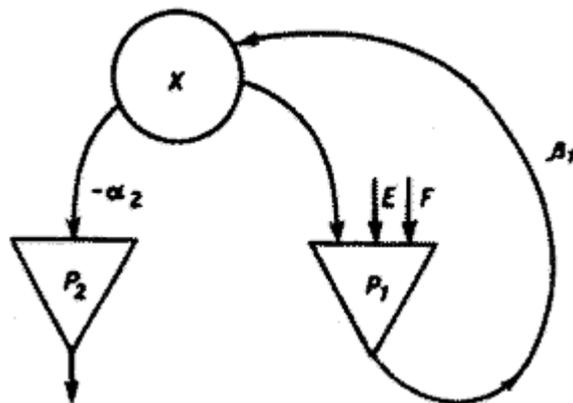


Рис. 1. Схема растения как системы с лимитирующими факторами.

На этой схеме кружком обозначен запас вещества – биомассы, а треугольниками – процессы. Стрелки указывают направление течения веществ; процесс P_1 свет E и некоторое вещество F , дает прирост биомассы x ; процесс дыхания P_2 потребляет часть биомассы, поставляя свободную энергию для жизнедеятельности оставшейся части, а выделяет продукты, которые в данной модели не учитываются. Процесс роста биомассы растения описывается уравнением:

$$\frac{dx}{dt} = \beta_1 P_1 - \alpha_2 P_2$$

$x(t)$ - значение биомассы в момент времени t . По смыслу задачи эта функция непрерывна и имеет производную. Пусть P_1 и P_2 - две действительные неотрицательные переменные – интенсивности роста и дыхания.

Величина интенсивности процессов P_j в момент времени t определяется в Л-системе лимитирующей из всех входных компонент, т.е. той, которая обеспечивает наименьшую интенсивность P_j . Все остальные компоненты, потоки которых превышают минимальный, оказываются в избытке и потребляются лишь частично. Лимитирующий фактор j -го процесса меняется по величине со временем вместе с соответствующей ему интенсивностью P_j . Вследствие изменения со временем состояния системы компонента i^* лимитирующая процесс в момент t_1 , может с течением времени перестать ею быть. Тогда начиная с момента t_2 лимитирующей станет уже другая компонента i^{**} . В модели растения интенсивности процессов P_1, P_2 запишутся в виде

$$P_1 = \min [x, E, F], \quad P_2 = \min (x).$$

Величина потока компоненты i на вход процесса j в момент t зависит от имеющегося в системе количества этой компоненты x , а для поступающих извне компонент (E, F) - от состояния среды. Поток внешней компоненты (вода, растворенные вещества) можно считать постоянным, пропорциональным наличию этой компоненты во внешней среде. Поток света E предполагается пропорциональным произведению интенсивности света E_0 на величину σ -поглощающей поверхности растения. Если растение в процессе роста не меняет формы, т.е. сохраняет геометрическое подобие, то величину поверхности можно считать пропорциональной квадрату линейных размеров или квадрату корня кубического из величины биомассы (биомасса пропорциональна объему растений, или кубу линейного размера). Таким образом, величина потока света запишется в виде

$$E = E_0 k(x) x^{2/3},$$

где $k(x)$ - коэффициент, отражающий форму растения.

В зависимости от соотношений величины и параметров системы k, σ, E_0, F величина P_1 может быть равна наименьшему из трех значений. Следовательно, решение распадается на три области, в каждой из которой рост лимитируется одним из факторов.

В последующих работах по моделированию конкретных экологических систем аппарат Л-систем нашел широкое применение. Рассмотрение биогеоценозов как систем с лимитирующими факторами позволяет "автоматически" следить за лимитированием каждого процесса. Особенно широкое применение нашел этот способ для описания микробных сообществ, где типы взаимодействия видов могут быть описаны сравнительно просто

при помощи обыкновенных дифференциальных уравнений (разных, в различных областях лимитирования).

Микробные сообщества являются классическими модельными экспериментальными системами для изучения экосистемных взаимодействий. Теоретический анализ видовой структуры трофического уровня одноклеточных автотрофных и гетеротрофных организмов с точки зрения Л-систем показал, что условием сосуществования конкурирующих видов является различие определяющих факторов развития, что является подтверждением принципа Гаузе о существовании лишь одного вида в каждой экологической нише. Это условие позволяет каждому виду зависеть в большей степени от одного определенного фактора среды и в меньшей степени - от других. В случае конкуренции только за взаимозаменяемые элементы питания каждый вид имеет вообще один лимитирующий трофический фактор, отличающийся от лимитирующих факторов других видов. Эти свойства поликультур микроорганизмов учитывают при совместном культивировании нескольких видов, которое дает возможность путем изменения процентного содержания видов в сообществе регулировать химический состав суммарной биомассы и ассимиляционный коэффициент поликультуры в целом.

Л-системы являются эффективным аппаратом для изучения популяций в экологическом окружении с ограничениями, накладываемыми на их развитие наличием корма, местообитания и прочим. Устойчивые состояния и кинетика переходных процессов определяются при этом для популяции каждого вида и каждый момент времени конечным, и притом достаточно малым, числом критических ограничений для компонент узкого места.

В различных фазах жизненного цикла и в различные сезоны давление критических условий не будет одинаковым. Опознавание критических компонент и ограничений в каждом отдельном случае и их количественное изучение являются задачей экспериментальной биологии. Применение формализованного принципа лимитирующих факторов позволяет при этом формализовать классификацию явлений в биогеоценозе и более эффективно использовать результаты количественных измерений.

Концепция лимитирующих факторов оказалась весьма плодотворной в инженерной экологии (*biomanipulation*). Управляющими для биогеоценоза могут быть потоки только тех ресурсов, которые лимитируют рост сообщества. Поэтому разработка концепции и методов как экспериментального выявления лимитирующих факторов для многовидовых сообществ и адекватных модельных подходов представляет собой весьма актуальную практическую задачу.

3) Практическая работа:

Решение предложенных задач.

Рекомендуемая литература:

Основная литература:

1. Ризниченко Г.Ю., Рубин А.Б. Математические модели биологических продукционных процессов. М. МГУ. 1993.
2. Кипятков В.Е. Практикум по математическому моделированию в популяционной экологии (учебное пособие). СПб, 2002.

Дополнительная литература:

1. Быков А.А. Моделирование природоохранной деятельности. Учебн. пособие. М., Изд-во НУМЦ Госкомэкологии России, 2008.
2. Вольтерра В. Математическая теория борьбы за существование / В. Вольтерра. – М. : Наука, 1976. – 285 с.

Методические указания для аспирантов
по дисциплине Б1.В.ДВ.1 Математическое моделирование
в экологии
к практическому занятию № 7

Тема: Статистическая обработка результатов измерений радиоактивности атмосферных выпадений.

Цель занятия: Сформировать у аспирантов общее представление о математической обработке результатов измерений.

Задачи: Освоить методы математической обработки цифровых данных.

Аспирант должен знать:

- 1) до изучения темы:
 - Понятие о статистических методах обработки данных;
 - Общие правила вычислений;
- 2) после изучения темы:
 - Порядок работы с цифровыми результатами исследований;
 - Построение вариационного ряда.

Изучение темы занятия направлено на формирование профессиональных компетенций: способность использовать современные методы обработки, анализа и синтеза полевой и лабораторной биологической

информации, знание принципов составления научно-технических проектов и отчетов (ПК-5).

Методика проведения занятия.

1. Определение темы занятия. Преподаватель поясняет цели и задачи занятия, значение полученных знаний для будущей работы по специальности.

2. Теоретическая часть.

Математическую обработку результатов анализов начинают с проверки полученного вариационного ряда на присутствие так называемых «выскакивающих» значений (вариант). Они обычно являются следствием грубой ошибки или получены в других условиях и не относятся к данному ряду измерений.

Наиболее простой способ выявления «выскакивающих» значений основан на оценке различий вариантов в упорядоченном ряду (т.е. в ряду, члены которого расположены в порядке возрастания) по формулам:

$$- \text{ для проверки максимального значения } - \frac{X_n - X_{n-1}}{X_n - X_1} ;$$

$$- \text{ для проверки минимального значения } - \frac{X_2 - X_1}{X_n - X_1} ,$$

где X_1, X_2, X_{n-1}, X_n – соответственно первый, второй, предпоследний и последний члены упорядоченного ряда.

Пример. Радиоактивность молока по стронцию-90 в пределах одного географического района выражена следующими величинами:

№1	№2	№3	№4	№5	№6	№7	№8	№9
6	8	8	9	10	11	12	12	20

пикокюри/л.

Анализ данного ряда позволяет предположить, что «выскакивающей» является наибольшая величина радиоактивности. Рассчитаем теперь

отношение: $\frac{X_n - X_{n-1}}{X_n - X_1} = \frac{X_9 - X_8}{X_9 - X_1} = \frac{20 - 12}{20 - 6} = \frac{8}{14} = 0,572$

Полученный результат оценим с помощью таблицы и, если он будет больше табличного, то вариант № 9 является «выскакивающим» и его следует исключить из всех последующих операций по математической обработке ряда.

Такое заключение можно сделать с высокой (99%) и обычной (95%) достоверностью. В таблице №1 для 9 членов ряда (n=9) указаны критические величины 0,437 (для 95%-ной достоверности) и 0,555 (для 99%-ной достоверности). Результат 0,572 превышает обе эти величины и, следовательно, вариант №9 (20 пикокюри/л) необходимо отбросить.

При проверке наименьшего варианта расчет проводят аналогично:

$$\frac{X_2 - X_1}{X_n - X_1} = \frac{8 - 6}{12 - 6} = 0,33$$

Это значение ниже табличного (при $n=8$) и, следовательно, для исключения из ряда наименьшего значения (6 пикокюри/л) нет оснований.

В таблице №1 представлены пограничные значения и для двух других приемов оценки «выскакивающих» вариант. Они отличаются от описанного не в принципе, а некоторыми особенностями, повышающими

продуктивность метода в отдельных ситуациях. Так отношение $\frac{X_n - X_{n-1}}{X_n - X_2}$ позволяет эффективнее выявлять «выскакивающую» наибольшую варианту в условиях, когда есть основания опасаться, что и первая, наименьшая,

варианта тоже является «выскакивающей». Отношение же $\frac{X_n - X_{n-2}}{X_n - X_1}$ позволяет эффективнее решать задачу в условиях, когда предполагаются «выскакивающими» сразу две наибольшие варианты. Иногда возникает положение, когда для проверяемого ряда вычислены три описанных вида отношений и лишь одно из них превышает пограничное для уровня достоверности 99%. Следует подчеркнуть, что это уже дает право на исключение крайней варианты. Неэффективность двух других отношений показывает лишь, что в данной ситуации их «мощность» оказалось недостаточной.

Пример. Радиоактивность комбикорма по цезию-137 представлена рядом величин: 34, 44, 49, 51, 55, 59, 67, 96, 139 пикокюри/кг.

Проверим, не является ли последний член данного ряда «выскакивающим»:

$$\frac{X_n - X_{n-1}}{X_n - X_1} = \frac{139 - 96}{139 - 34} = 0,410,$$

$$\frac{X_n - X_{n-1}}{X_n - X_2} = \frac{139 - 96}{139 - 44} = 0,453,$$

$$\frac{X_n - X_{n-2}}{X_n - X_1} = \frac{139 - 67}{139 - 34} = \underline{\underline{0,686}}$$

Последний результат (0,686) выше табличных значений для 9-ти членов ряда, что позволяет исключить величину радиоактивности комбикорма 139пк/кг как «выскакивающую».

Статистические показатели.

Для общей характеристики случайных величин используются различные показатели: среднее значение (средняя арифметическая), стандартная и относительная ошибки среднего арифметического, коэффициент корреляции и т.д.

Средняя арифметическая величина.

Согласно теории ошибок наиболее близким к истинному значению измеряемой величины является среднее арифметическое значение ряда отдельных измерений, т.е.

$$M = \frac{a_1 + a_2 + a_3 + \dots + a_n}{n},$$

где M – средняя арифметическая;

a – члены вариационного ряда (значения отдельных измерений радиоактивности);

n – число измерений.

Например: Радиоактивность атмосферных выпадений, измеренная в десяти хозяйствах одной области, составила: 10, 12, 9, 11, 8, 13, 12, 10, 11, 9 милликюри на 1 км^2 в месяц (мКи/ км^2 в мес.). Требуется определить средний уровень радиоактивности выпадений.

Средняя арифметическая величина радиоактивных выпадений в этом случае будет равна:

$$M = \frac{10 + 12 + 9 + 11 + 8 + 13 + 12 + 10 + 11 + 9}{10} = 10,5 \text{ мКи/км}^2.$$

Таким образом, в среднем на территории данной области радиоактивность атмосферных выпадений составляла $10,5 \text{ мКи/км}^2$ в месяц.

В случае неравномерного распределения обрабатываемых данных наиболее приемлемой является не простая средняя арифметическая, а средняя арифметическая взвешенная величина. Эта величина наиболее точно учитывает вклад (вес) каждого измерения исследуемого вариационного ряда и определяется по формуле:

$$M_{\text{взв}} = \frac{M_1 \times P_1 + M_2 \times P_2 + M_3 \times P_3 + \dots + M_n \times P_n}{P_1 + P_2 + P_3 + \dots + P_n},$$

где: $M_{\text{взв}}$ – взвешенная средняя арифметическая величина;

$M_1, M_2, M_3, \dots, M_n$ – средние значения (средние арифметические) исследуемых рядов или интервалов одного ряда;

$P_1, P_2, P_3, \dots, P_n$ – количества (математический вес) исследуемого продукта (кг, л).

3) Практическая работа:

Решение предложенных задач.

Рекомендуемая литература:

Основная литература:

1. Игнатенко Г.К. Первичная обработка данных экологического мониторинга. Обнинск: ИАТЭ. – 2006. – 76 с.
2. Халафян А.А. Статистический анализ данных. М.: Бином-Пресс. – 2008. – 512 с.

Дополнительная литература:

1. Джефферс Дж. Введение в системный анализ: применение в экологии. М.: Мир, 1981. – 256 с.
2. Кендалл М., Стюарт А. Многомерный статистический анализ и временные ряды. – М.: Наука. 1976. – 736 с.
3. Рокицкий П.Ф. Биологическая статистика. Минск. Высшейш. шк. – 1973. – 320 с.

Методические указания для аспирантов
по дисциплине Б1.В.ДВ.1 Математическое моделирование
в экологии
к практическому занятию № 8

Тема: Расчет коэффициента вариации и стандартной ошибки средней арифметической величины.

Цель занятия: Сформировать у аспирантов общее представление об ошибках измерений и их математическом вычислении.

Задачи: Освоить методы вычисления статистических ошибок измерений.

Аспирант должен знать:

1) до изучения темы:

- Понятие о статистических методах обработки данных;
- Общие правила вычислений;

2) после изучения темы:

- Понятие о систематических, случайных и статистических ошибках измерения;
- Способы вычисления статистических ошибок измерения.

Изучение темы занятия направлено на формирование профессиональных компетенций: способность использовать современные методы обработки, анализа и синтеза полевой и лабораторной биологической информации, знание принципов составления научно-технических проектов и отчетов (ПК-5).

Методика проведения занятия.

1. Определение темы занятия. Преподаватель поясняет цели и задачи занятия, значение полученных знаний для будущей работы по специальности.

2. Теоретическая часть.

Общее представление об ошибках измерений.

Различают ошибки систематические, случайные и промахи.

Систематические ошибки появляются постоянно при повторных измерениях и возникают, как правило, ввиду неисправности измерительных приборов или недостатков самого метода измерения. Систематические ошибки всегда односторонне влияют на результаты измерения, увеличивая или уменьшая их, и могут быть всегда устранимы путем совершенствования радиометрической аппаратуры.

Случайные ошибки возникают при изменении напряжения электротока в сети во время проведения радиометрических измерений, недостаточной чувствительности аппаратуры и т.п. Все это приводит к тому, что несколько измерений одной и той же величины дают различные результаты.

Исключить случайные ошибки, возникающие при измерении, нельзя, но, используя закон теории вероятностей, можно оценить их и сделать правильный вывод о достоверности полученных результатов.

Необходимо помнить, что увеличение числа измерений уменьшает влияние случайных ошибок. Если же случайная ошибка больше систематической, то именно случайную ошибку нужно уменьшить в первую очередь.

Промахи. Источником подобных ошибок является невнимательность исследователя. Для устранения промахов нужно соблюдать аккуратность, тщательность в работе и в записях результатов измерений. Многократное измерение одной и той же величины в одних и тех же условиях, как правило, позволяет обнаружить и устранить промахи.

Статистические ошибки.

Абсолютной ошибкой данного измерения называется разность между средним арифметическим значением и отдельным измерением. Абсолютная ошибка обозначается греческой буквой Δ .

Точность каждого значения данного ряда измерений характеризуется средней квадратической ошибкой, которую обозначают греческой буквой ς (сигма малая) и определяют по формуле:

$$\varsigma = \sqrt{\frac{\sum(a - M)^2}{n - 1}},$$

где $\sum(a - M)^2$ – сумма квадратов отклонений всех измерений ряда от среднего арифметического значения;

$n-1$ – число членов вариационного ряда, уменьшенное на единицу.

Средний результат определения радиоактивности записывают в виде формулы: $A=M\pm\zeta$. Здесь средняя квадратическая ошибка (ζ) показывает усредненную величину отклонения каждого измерения от своей средней арифметической.

В так называемых нормальных вариационных рядах весь размах изменчивости (варьирования), ограниченный максимальным и минимальным значением изучаемого показателя, включает в себе шестикратную величину среднего квадратического отклонения. При этом максимальный вариант отличается от средней арифметической на значение $+3\zeta$, а минимальный вариант – на значение -3ζ . Поэтому принято весь размах изменчивости выражать такой записью: $M\pm 3\zeta$.

На основании этой особенности изменчивости в нормальных рядах можно осуществлять некоторые ориентировочные расчеты. Например, по величине средней арифметической и значению ζ можно определить величину максимального и минимального членов в данной совокупности измерений.

Как будет показано в дальнейшем, по размаху изменчивости можно рассчитать приближенное значение среднего квадратического отклонения.

Среднее квадратическое отклонение (ζ) – величина именованная, она выражается в тех же единицах, что и среднее арифметическое значение: кюри/км², мкюри/кг, пкюри/л и т.д.

Пример. Для определения ζ воспользуемся величинами радиоактивности зернофуража по стронцию-90, приведенными в таблице:

Радиоактивность объекта, пКи/кг	a-M	(a-M) ²
41	4	16
38	1	1
35	-2	4
40	3	9
39	2	4
38	1	1
37	0	0
36	-1	1
32	-5	25
34	-3	9
Сумма 370		70
Среднее 37		

Подставив полученные величины в формулу $\zeta = \sqrt{\frac{\sum(a-M)^2}{n-1}}$, определим среднюю квадратичную ошибку отдельного измерения данного вариационного ряда:

$$\zeta = \sqrt{\frac{70}{9}} = \sqrt{7,77} = 2,8 \text{ пк/кг.}$$

Итоговый результат определения радиоактивности объекта после проведенной математической обработки записывают в виде: $A = M \pm \zeta = 37 \pm 2,8 \text{ пк/кг.}$

Методы статистической обработки, требующие при определении усредненных ошибок возведения в квадрат каждого отклонения от средней арифметической величины, называются квадратическими. Эти методы при практическом применении, как видно, довольно громоздки.

При большом числе измерений количество арифметических действий достигает иногда десятков, что не редко ведет к так называемым ошибкам внимания.

Поэтому приходится пользоваться более простыми методами, которые хотя и менее точны, однако дают вполне удовлетворительные результаты (ошибка вычисления статистических показателей при этом обычно не превышает 5%).

Так, среднюю квадратическую ошибку (ζ) можно определить по разности максимального и минимального значения измерений в соответствии с формулой:

$$\zeta = \frac{a_{\text{макс.}} - a_{\text{мин.}}}{K_{\zeta}},$$

где: K_{ζ} – величина коэффициента, определяемого по таблице №2 приложения.

При большом количестве измерений (например, более сотни) коэффициент – « K_{ζ} » можно брать равным 6, поскольку максимальный размах вариабильности ряда близок к $M \pm 3\zeta$, т.е. укладывается в пределы 6ζ .

Подставив в приведенную формулу данные из предшествующего примера, получим:

$$\zeta = \frac{41 - 32}{3,08} \approx 3,0 \text{ пк/кг.}$$

Как видно, результаты достоверно близки, а бóльшая простота второго способа вычислений очевидна.

Процентное отношение величины средней квадратичной ошибки отдельного измерения к своей средней арифметической называется коэффициентом вариации. Он определяется по формуле:

$$K_v = \frac{\zeta}{M} \times 100\% .$$

Пример. Скорость счета от одной пробы молока, исследуемого на УМФ-1500 при многократном измерении была выражена следующим рядом (имп/мин): 10, 13, 8, 12, 10, 11, 14, 12, 11, 9. Средняя арифметическая

величина данного ряда $M=11$. Средняя квадратическая ошибка отдельного измерения, определяемая по размаху изменчивости, будет:

$$\zeta = \frac{14 - 8}{3,08} = 2 \text{ имп/мин},$$

Это значит, что величина каждого члена ряда отклоняется от средней арифметической в среднем на 2, т.е. $M \pm \zeta = 11 \pm 2$ имп/мин.

Коэффициент вариации будет: $K_v = \frac{2}{11} \times 100 = 18\%$. Эта величина в данном случае характеризует не только изменчивость признака у членов ряда (разброс), но и погрешность самого метода измерения (радиоактивность пробы молока мы измеряли с точностью $\pm 18\%$).

Очень важным статистическим показателем является средняя квадратическая (стандартная ошибка среднего арифметического значения. Она характеризует соответствие полученной средней арифметической величины истинному значению радиоактивности исследуемого объекта.

Стандартная ошибка определяется по формуле: $m = \sqrt{\frac{\sum(a - M)^2}{n \times (n - 1)}}$.

Пример. Определить стандартную ошибку средней арифметической величины, исходя из данных, приведенных в таблице (концентрация стронция-90 в свекле, пк/кг):

a	a-M	(a-M) ²
13,3	-6,1	37,21
14,7	-4,7	22,09
16,2	-3,2	10,24
16,8	2,6	6,76
19,0	0,4	0,16
19,6	0,2	0,04
20,1	0,7	0,49
21,4	2,0	4,00
21,4	2,0	4,00
22,5	3,1	9,61
23,0	3,6	12,96
24,8	5,4	29,16

Отсюда: $\sum a = 232,8$; $M = \frac{232,8}{12} = 19,4$; $\sum(a-M)^2 = 136,72$.

Подставив полученные значения в формулу, определим стандартную ошибку всего вариационного ряда: $m = \sqrt{\frac{136,72}{12 \times 11}} = 1,02$

Если известна величина средней квадратичной ошибки отдельного измерения, то стандартную ошибку вариационного ряда определяют по формуле: $m = \frac{\zeta}{\sqrt{n}}$.

Воспользуемся теперь оценкой ζ через размах варьирования. В нашем примере $R=24,8-13,3=11,5$; а K_ζ (по таблице №2 для двенадцати измерений) $=3,26$. Отсюда $\zeta = \frac{11,5}{3,26} = 3,53$. Стандартная ошибка $m = \frac{3,53}{\sqrt{12}} = 1,02$.

Как видно, оба результата совпали.

Стандартную ошибку средней арифметической можно определить константным методом по формуле: $m = K \times \Sigma(a - M)$. В таблице №3 приложения даны константы для различного количества вариантов.

Ход вычисления: после обычного определения средней арифметической величины ($M = \frac{\Sigma a}{n}$) находят отклонения от нее для каждого варианта $(a - M)$. Последние суммируют без учета арифметических знаков и полученную сумму умножают на константу, взятую из строки соответствующего количества вариантов (таблица №3 приложения).

Пример. Вычислить среднюю арифметическую величину и ее стандартную ошибку следующего вариационного ряда (округление до десятков произведены для упрощения расчетов).

Варианты (а): 130, 130, 120, 140, 130, 120, 110, 110, 150, 160.

Отклонения $(a - M)$: 0, 0, 10, 10, 0, 10, 20, 20, 20, 30.

Отсюда: $\Sigma a = 1300$; $M = \frac{1300}{10} = 130$; $\Sigma(a - M) = 120$.

Коэффициент K (при десяти вариантах) $= 0,0418$. Стандартная ошибка средней арифметической величины будет $m = 0,0418 \times 120 = 5,016 \approx 5$.

Если в состав вариационного ряда входит не более десяти членов, то стандартную ошибку средней арифметической величины можно ориентировочно определить путем деления величины размаха варьирования на число членов ряда по формуле:

$$m = \frac{a_{\text{макс.}} - a_{\text{мин.}}}{n}$$

Для расчетов воспользуемся вариационным рядом из предыдущего примера:

$$m = \frac{160 - 110}{10} = \frac{50}{10} = 5.$$

Как видно, результат примерно тот же самый, что и при определении стандартной ошибки константным методом.

Более точно величину стандартной ошибки по размаху варьирования определяют по формуле:

$$m = \frac{a_{\text{макс.}} - a_{\text{мин.}}}{K_m},$$

где K_m – величина коэффициента, определяемого по таблице №4 приложения.

Из того же примера имеем:

$$m = \frac{160 - 110}{9,73} = \frac{50}{9,73} = 5,14.$$

При определении этой же ошибки обычным (квадратическим) методом она оказывается равной 5,164. То есть результаты практически совпадают.

Полученный при математической обработке вариационного ряда результат записывают в виде средней арифметической величины со своей стандартной ошибкой, т.е. $M \pm m$; в нашем случае - 130 ± 5 .

Важно строго разграничивать применение средней квадратической ошибки отдельного измерения (ζ) и средней квадратической ошибки среднего арифметического (m), исходя из того, что ζ выражает величину отклонения каждого измерения от своей средней арифметической (M), в то время как m характеризует величину отклонения самой средней арифметической от истинного значения радиоактивности объекта ($M_{\text{истин.}}$).

При сравнении средних арифметических величин двух рядов часто приходится определять их суммарную ошибку по формуле:

$$m_e = \sqrt{m_1^2 + m_2^2},$$

где: m_1 и m_2 – стандартные ошибки средних арифметических двух рядов.

Необходимо помнить, что при нахождении суммарной ошибки складываются не сами стандартные ошибки, а их квадраты и из их суммы квадратов извлекается квадратный корень.

Отсюда вытекают два вывода:

1. Значение отдельных ошибок быстро попадает по мере уменьшения величины этих ошибок. Так, допустим M_1 и M_2 – две средние арифметические, а m_1 и m_2 – их средние квадратические (стандартные) ошибки, соответственно равные в одном случае 4 и 4, а в другом – 8 и 0, т.е. арифметическая сумма ошибок будет равна в обоих случаях 8. Однако суммарная ошибка в первом случае будет равна $\sqrt{4^2 + 4^2} = \sqrt{32} = 5,65$; а во втором случае $\sqrt{8^2 + 0^2} = \sqrt{64} = 8$.

Если мы хотим повысить точность измерения, то в первую очередь должны стремиться уменьшить ту ошибку измерения, которая больше, ибо доля ее в суммарной ошибке является решающей.

2. Стандартная ошибка среднего арифметического значения всегда меньше ошибки результата каждого отдельного измерения; т.е. m меньше ζ , поскольку m равна ζ , деленной на квадратный корень из числа измерений ($m = \frac{\zeta}{\sqrt{n}}$).

Отсюда следует, что точность результата будет возрастать, при увеличении числа измерений. Например, чтобы повысить точность определения радиоактивности в объектах ветнадзора в 3 раза, число измерений нужно увеличить в 9 раз.

Процентное отношение стандартной ошибки к средней арифметической величине называется относительной ошибкой среднего арифметического. Она определяется по формуле: $\zeta = \frac{m}{M} \times 100\%$.

Например, радиоактивность клубней картофеля по стронцию-90 равна $10 \pm 1,1$ пКи/кг. Относительная ошибка среднего арифметического при этом будет: $\zeta = \frac{1,1}{10} \times 100\% = 11\%$.

3) Практическая работа:

Решение предложенных задач.

Рекомендуемая литература:

Основная литература:

1. Игнатенко Г.К. Первичная обработка данных экологического мониторинга. Обнинск: ИАТЭ. – 2006. – 76 с.
2. Халафян А.А. Статистический анализ данных. М.: Бином-Пресс. – 2008. – 512 с.

Дополнительная литература:

1. Джефферс Дж. Введение в системный анализ: применение в экологии. М.: Мир, 1981. – 256 с.
2. Кендалл М., Стюарт А. Многомерный статистический анализ и временные ряды. – М.: Наука. 1976. – 736 с.
3. Рокицкий П.Ф. Биологическая статистика. Минск. Вышейш. шк. – 1973. – 320 с.
4. Шеффе Г. Дисперсионный анализ. М.: Наука. – 1980. – 512 с.

Методические указания для аспирантов
по дисциплине Б1.В.ДВ.1 Математическое моделирование
в экологии
к практическому занятию № 9

Тема: Определение достоверности изменения величин по t-критерию Стьюдента. Определение критерия Пирсона.

Цель занятия: Сформировать у аспирантов общее представление о достоверности различий между показателями.

Задачи: Освоить методы математической обработки цифровых данных.

Аспирант должен знать:

1) до изучения темы:

- Понятие об ошибках измерений;
- Общие правила вычислений;

2) после изучения темы:

- Методы определения степени достоверности между изучаемыми величинами.

Изучение темы занятия направлено на формирование профессиональных компетенций: способность использовать современные методы обработки, анализа и синтеза полевой и лабораторной биологической информации, знание принципов составления научно-технических проектов и отчетов (ПК-5).

Методика проведения занятия.

1. Определение темы занятия. Преподаватель поясняет цели и задачи занятия, значение полученных знаний для будущей работы по специальности.

2. Теоретическая часть.

Определение достоверности разности между средними арифметическими двух вариационных рядов.

Для сравнения уровней радиоактивной загрязненности продуктов сельского хозяйства в разных зонах в пределах одного промежутка времени или в одном районе (области) в разные годы требуется определить не только разницу средних арифметических двух рядов измерений, но и установить достоверность этого различия, т.е. подтвердить его неслучайный характер. Методика оценки достоверности разности средних арифметических двух статистических рядов изложена в следующем примере.

Предположим, что при анализе проб молока из пяти хозяйств было установлено, что радиоактивность этого продукта по стронцию-90 в 1976г. в

данной местности была равна $M_1 \pm m_1 = 14.5 \pm 1.8$ пКи/л и в 1977г. – $M_2 \pm m_2 = 10 \pm 1.4$ пКи/л (всего было проведено десять радиохимических анализов). Снижение радиоактивности молока за год составило $14.5 - 10 = 4.5$ пк/л, т.е. радиоактивность уменьшилась примерно на одну треть. Насколько достоверно это уменьшение радиоактивности? Ответить на это можно только после определения коэффициента достоверности (t – критерий Стьюдента) по формуле:

$$t = \frac{M_1 - M_2}{\sqrt{m_1^2 + m_2^2}}$$

Подставив вместо буквенных обозначений соответствующие цифровые данные, получим: $t = \frac{14.5 - 10}{\sqrt{3.24 + 1.96}} = \frac{4.5}{\sqrt{5.2}} = \frac{4.5}{2.28} = 1.97$.

Сравнивая полученный результат с табличным значением критерия Стьюдента для десяти измерений по двум строкам (число степеней свободы $f = n_1 + n_2 - 2 = 5 + 5 - 2 = 8$) видим, что наш результат (1,97) меньше указанного в таблице №5 (см. приложение) не только для высокой (99%), но и для обычной достоверности (95%), ибо табличные коэффициенты соответственно равны 3,36 и 2,31. Поэтому следует сделать вывод об отсутствии достоверной разницы в уровнях радиоактивности молока между 1976 и 1977гг. Разница считается действительной, если ее достоверность равна или больше 95% (т.е. $P \geq 95\%$ или $P \geq 0,95$).

Таблицей №5 пользуются также при определении уровня достоверности отдельно взятой средней арифметической. Например, если средняя арифметическая величина и ее стандартная ошибка равны 11 ± 2 , а число вариант 5, то, определив критерий достоверности $t = \frac{11}{2} = 5.5$ и количество степеней свободы ($f = 5 - 1 = 4$), находим по таблице, что уровень достоверности средней арифметической более 99% (но меньше 99,9%), ибо соответствующие табличные значения критерия Стьюдента при указанном количестве степеней свободы равны 4,60 и 8,61, а найденное нами значение t – критерия занимает промежуточное положение.

По формуле: $t = \frac{M_1 - M_2}{\sqrt{m_1^2 + m_2^2}}$ достоверность различий между средними

арифметическими определяют лишь в том случае, если варианты одной совокупности изменяются независимо от вариантов другой совокупности. Если же между вариантами обеих совокупностей имеется взаимосвязь (сопряженность), то достоверность различий между рядами целесообразно вычислять методом прямой разности, который заключается в следующем. Варианты сравниваемых совокупностей (a_1 и a_2) выписывают в рядом

расположенные строки (вертикальные или горизонтальные) и в каждой паре вариант определяют разность ($a_1 - a_2$). Затем суммируют полученные разности и находят их среднее арифметическое значение. После этого по размаху варьирования вычисляют стандартную ошибку этого среднего арифметического, делят на нее значение самой средней арифметической и получают величину t – критерия, характеризующую достоверность различий между сравниваемыми рядами.

Пример. Содержание цезия-137 в пробах свеклы, отобранных с одних и тех же участков в семи контрольных пунктах области в 1975 и 1976 гг., характеризуется данными, приведенными в таблице (графы 2 и 3).

Номера контрольных пунктов	Уровни радиоактивности свеклы (пКи/кг)		Разность ($a_1 - a_2$)
	1975г (a_1)	1976г (a_2)	
1	19,5	16,9	2,6
2	20,2	16,2	4,0
3	22,9	19,4	3,5
4	27,7	25,6	2,1
5	30,5	29,3	1,2
6	35,6	31,4	4,2
7	31,9	28,5	3,4
Сумма	188,3	167,3	21,0
Среднее	26,9	23,9	3,0

Поскольку пробы свеклы в 1975 и 1976гг. брали на одних и тех же участках, можно предположить наличие зависимости между значениями из обоих сравниваемых рядов. Это дает основание провести вычисление достоверности разности между средними арифметическими этих рядов методом «сопряженных пар». Для этого среднюю арифметическую разностей ($a_1 - a_2$) делят на ее стандартную ошибку, определяемую по формуле:

$$m = \frac{a_{\text{макс.}} - a_{\text{мин.}}}{n} = \frac{4,2 - 1,2}{7} = 0,43.$$

В итоге получаем величину t – критерия: $\frac{3,0}{0,43} = 7,0$.

В данном случае число степеней свободы $f = n - 1 = 7 - 1 = 6$.

При этом значении числа степеней свободы в таблице №5 для уровня достоверности 99,9% имеет величину 5,96, а наша величина t – критерия, равно 7, значительно превосходит табличный показатель. Следовательно, различие между рядами является высокодостоверным. Оно свидетельствует о

существенном снижении радиоактивности исследуемого объекта в 1976г. по сравнению с 1975г.

Если бы в данном примере применялся обычный критерий (т.е. не учитывающий сопряженность пар), то различие не было бы обнаружено, ибо большой разброс величин радиоактивности в пределах каждого ряда из-за резкого различия в характере почв привел бы к завышенным значениям стандартной ошибки средних арифметических и, следовательно, к заниженной величине t – критерия.

Для подтверждения этого положения расчетами, находим величину стандартных ошибок m_1 и m_2 по размаху варьирования:

$$m_1 = \frac{35,6 - 19,5}{7} = \frac{16,1}{7} = 2,3$$
$$m_2 = \frac{31,4 - 16,2}{7} = \frac{15,2}{7} = 2,17$$

и определим t – критерий разности между средними арифметическими по формуле:

$$t = \frac{M_1 - M_2}{\sqrt{m_1^2 + m_2^2}} = \frac{26,9 - 23,9}{\sqrt{2,3^2 + 2,17^2}} = \frac{3}{\sqrt{9,9989}} = \frac{3}{3,16} = 0,95,$$

что значительно ниже табличного 95-процентного уровня достоверности (при $f=n_1+n_2-2=7+7-2=12$ в таблице №5 указана величина t – критерия, равная 2,18), несмотря на то, что в этом случае используют вдвое большее число степеней свободы.

Поэтому методом «сопряженных пар» пользуются только тогда, когда взаимосвязь пар несомненна (в этом случае влияние уменьшения числа степеней свободы перекрывается влиянием уменьшением стандартной ошибки разности) и в случаях одинакового количества вариантов в сравниваемых рядах.

Сравнение двух альтернативных распределений.

Для оценки разницы между признаками используют критерий различия «хи-квадрат» (χ^2). Его можно применять для изучения многих биологических процессов (при анализе влияния на организм различных факторов).

Данные наблюдений (опытов) группируют в таблицы, состоящие из нескольких полей. Наиболее простой является четырехпольная таблица. Значение χ^2 в этом случае определяют по формуле:

$$\chi^2 = \frac{([p_1 \times p_4 - p_2 \times p_3] - 1/2n)^2 \times n}{(p_1 + p_2) \times (p_3 + p_4) \times (p_1 + p_3) \times (p_2 + p_4)},$$

где p – частоты (их размещение показано в нижеследующей таблице, в которой в качестве примера приведены результаты эксперимента по профилактике лучевой болезни у овец с помощью химического радиопротектора);

n – общее число наблюдений.

Прямые скобки, в которых находятся $p_1 \times p_4 - p_2 \times p_3$, показывают, что берется только абсолютное, т.е. положительное численное значение этой разности.

Пример. При проверке эффективности противолучевого препарата при введении его подкожно до облучения были получены следующие данные:

Группа овец	Исход лучевой болезни		Сумма
	Выздоровело животных	Пало животных	
Леченные	$p_1=25$	$p_2=23$	$p_1+p_2=48$
Контрольные	$p_3=7$	$p_4=29$	$p_3+p_4=36$
Сумма	$p_1+p_3=32$	$p_2+p_4=52$	$n=84$

Для выяснения значимости действия радиопротектора использована четырехпольная таблица. В ней буквами p_1 и p_2 обозначены частоты для двух верхних полей таблицы, а буквами p_3 и p_4 – для двух нижних. В самой нижней строке таблицы – суммы частот по вертикальным столбцам (p_1+p_3) и (p_2+p_4), в правом крайнем столбце таблицы – суммы частот по горизонтальным строчкам (p_1+p_2) и (p_3+p_4). Общая сумма $n = p_1+p_2+ p_3+p_4$. В нашем примере $n=84$ овцам.

Подставив в вышеприведенную формулу соответствующие числовые значения и произведя необходимые вычисления, получим величину χ^2 :

$$\chi^2 = \frac{([25 \times 29 - 23 \times 7] - 84/2)^2 \times 84}{48 \times 36 \times 32 \times 52} = \frac{22888656,0}{2875392,0} \approx 8,0.$$

По таблице критических значений χ^2 (таблица №6 приложения) для уровня достоверности 95% и 99% находим граничные величины этого критерия, применительно к нашему случаю. Они соответственно равны 3,85 и 6,63 (при одной степени свободы для четырехпольной таблицы). Поскольку полученная нами величина $\chi^2=8,0$ намного превосходит оба табличных значения, то можно считать эффективность противолучевого препарата значимой ($P>0,99$).

Доверительный интервал.

Зная среднюю арифметическую величину (M) ряда измерений и ее стандартную ошибку (m), можно установить с определенной точностью те границы, в которых находится истинная средняя арифметическая ($M_{\text{ист.}}$). Для этого используют формулу:

$$M - tm \leq M_{\text{ист.}} \leq M + tm,$$

где t – критерий по Стьюденту.

Эти границы получили название доверительных; интервал, т.е. разница между максимумом и минимумом, также называемые доверительным.

Заранее устанавливают тот или иной уровень достоверности (P), при котором желают получить доверительные границы для истинного значения средней арифметической величины ($M_{\text{ист.}}$).

В интервале $M \pm m$ истинная величина среднего арифметического значения содержится в 68% случаев. Это имеет следующий смысл: пусть взято 100 выборок по « n » измерений каждая и, следовательно, получен 100 доверительных интервалов. Все они будут несколько различаться между собой как величиной средних арифметических, так величиной стандартных ошибок, но только 68 из этих интервалов покроют $M_{\text{ист.}}$ (или, иными словами, 68 этих интервалов будут содержать $M_{\text{ист.}}$).

Однако такая вероятность (68%) недостаточна для надежного определения истинного значения средней арифметической величины. В биологических применениях статистики достаточно надежным считается 95%-ный доверительный интервал.

Пример. Определяли концентрацию стронция-90 в молоке (в пКи/л). При девяти измерениях был получен результат $A = M \pm m = 31 \pm 2$. По таблице находим значение t – критерия для уровня достоверности $P = 0,95$. При числе степеней свободы $f = 9 - 1 = 8$ величина t – критерия равна 2,31. Доверительный интервал определяем по вышеприведенной формуле: $31 - 2 \times 2,31 \leq M_{\text{ист.}} \leq 31 + 2 \times 2,31$ или $26,38 \leq M_{\text{ист.}} \leq 35,62$. Таким образом, вероятность того, что истинное значение средней арифметической радиоактивности молока находится в пределах от 26,38 до 35,62 пКи/л, равна или больше 95%. Результат записывают так: $M_{\text{ист.}} = 31 \pm 4,62$ пКи/л ($P \geq 0,95$ или $P \leq 0,05$).

Обычно используют следующие три уровня достоверности: $P = 0,95$ (95%), 0,99 (99%) и 0,999 (99,9%), что соответствует уровням значимости $P = 0,05$ (5%), 0,01 (1%) и 0,001 (0,1%).

3) Практическая работа:

Решение предложенных задач.

Рекомендуемая литература:

Основная литература:

1. Игнатенко Г.К. Первичная обработка данных экологического мониторинга. Обнинск: ИАТЭ. – 2006. – 76 с.
2. Халафян А.А. Статистический анализ данных. М.: Бином-Пресс. – 2008. – 512 с.

Дополнительная литература:

1. Джефферс Дж. Введение в системный анализ: применение в экологии. М.: Мир, 1981. – 256 с.
2. Кендалл М., Стюарт А. Многомерный статистический анализ и временные ряды. – М.: Наука. 1976. – 736 с.
3. Рокицкий П.Ф. Биологическая статистика. Минск. Высшейш. шк. – 1973. – 320 с.

Методические указания для аспирантов

по дисциплине Б1.В.ДВ.1 Математическое моделирование

в экологии

к практическому занятию № 10

Тема: Корреляционный и регрессионный анализ.

Цель занятия: Сформировать у аспирантов представление о корреляционной зависимости между показателями и ее количественной оценке.

Задачи: Освоить методы корреляционного и регрессионного анализа.

Аспирант должен знать:

1) до изучения темы:

- Понятие о статистических методах обработки данных;
- Понятие об оценке достоверности различия между показателями;

2) после изучения темы:

- Сущность корреляционного анализа;
- Сущность регрессионного анализа.

Изучение темы занятия направлено на формирование профессиональных компетенций: способность использовать современные методы обработки, анализа и синтеза полевой и лабораторной биологической информации, знание принципов составления научно-технических проектов и отчетов (ПК-5).

Методика проведения занятия.

1. Определение темы занятия. Преподаватель поясняет цели и задачи занятия, значение полученных знаний для будущей работы по специальности.

2. Теоретическая часть.

Корреляционный анализ.

Для прогнозирования радиоактивной загрязненности продуктов сельского хозяйства большое значение имеет изучение и измерение степени зависимости уровней радиоактивности звеньев цепи: почва–растение–животное (продукты животноводства). Эту проблему решают с помощью корреляционного анализа, который позволяет определить изменяются ли два признака самостоятельно, независимо друг от друга, или вариация одного признака в какой-то степени связана с вариацией другого.

Между признаками существуют функциональные и корреляционные связи. При функциональных связях каждому значению одного признака соответствует одно определенное значение другого признака. При корреляционных связях численному значению одного признака соответствует много значений другого признака.

Например, существует бесспорная зависимость между количеством выпавших на местности радиоактивных веществ и содержанием их в продуктах растениеводства (в фураже). Но это не значит, что определенному количеству РВ соответствует строго определенная радиоактивная загрязненность сельскохозяйственной продукции. Под влиянием типа и структуры почвы, способа внесения удобрений, методов посева и т.д. при одном и том же содержании РВ в почве радиоактивная загрязненность урожая может колебаться в значительных пределах.

Различают прямую и обратную корреляцию. При прямой корреляции увеличение одного признака ведет к увеличению другого, при обратной – увеличение одного признака соответственно связано с уменьшением другого.

Для количественной характеристики корреляции используют коэффициент, который вычисляют по формуле:

$$r = \frac{\sum (a_1 - M_1) \times (a_2 - M_2)}{n \times \zeta_1 \times \zeta_2},$$

где r – коэффициент корреляции;

Σ – знак суммирования;

M_1 и M_2 – средние арифметические величины двух статистических рядов;

a_1 и a_2 – члены (варианты) этих рядов;

$(a_1 - M_1)$ – отклонения от средней величины ряда 1;

$(a_2 - M_2)$ – отклонения от средней величины ряда 2;

ζ_1 и ζ_2 – средние квадратические ошибки отдельных измерений;

n – число наблюдений (измерений).

Пример. Требуется определить величину и направление корреляционной связи между содержанием кальция в почве г/кг (ряд a_1) и радиоактивной загрязненностью травы в нКи/кг (ряд a_2). Необходимые для вычисления коэффициента корреляции данные приведены в таблице:

a_1	a_2	$a_1 - M_1$	$a_2 - M_2$	$(a_1 - M_1)(a_2 - M_2)$
2,2	18	-0,3	-3,0	0,9
<u>1,5</u>	24	-1,0	3,0	-3,0
<u>4,0</u>	<u>12</u>	1,5	-0,9	-13,5
2,5	19	0	-2,0	0
3,5	15	1,0	-6,0	-6,0
2,0	27	-0,5	6,0	-3,0
1,6	<u>32</u>	-0,9	11,0	-9,9
2,0	21	-0,5	0	0
1,7	25	-0,8	4,0	-3,2
4,0	17	1,5	-4,0	-6,0
$M_1=2,5$	$M_2=21$			$=-43,7$

Определяем ζ_1 и ζ_2 по формуле: $\zeta = \frac{a_{\text{макс.}} - a_{\text{мин.}}}{K_\zeta}$.

$$\zeta_1 = \frac{4,0 - 1,5}{3,08} = \frac{2,5}{3,08} = 0,8 \quad \zeta_2 = \frac{32 - 12}{3,08} = \frac{20}{3,08} = 6,5$$

Подставив в формулу коэффициента корреляции соответствующие вычисленные значения, получим:

$$r = \frac{-43,7}{10 \times 6,5 \times 0,8} = \frac{-43,7}{52} = -0,8$$

Сравнивая полученную величину коэффициента корреляции с табличным значением для десяти измерений, видим, что наш коэффициент (-0,8) превосходит табличное значение коэффициента корреляции не только для достоверности 95 %, но и для достоверности 99 %, т.е. нами установлена высоко достоверная зависимость между концентрацией кальция в почве и выносом радиоактивных веществ растениями. Отрицательный знак полученного коэффициента корреляции указывает на наличие обратной связи между признаками: с увеличением содержания кальция в почве усвояемость радиоактивных веществ растениями уменьшается.

Регрессионный анализ (метод наименьших квадратов).

Коэффициент корреляции указывает лишь на степень связи в вариации двух попеременных величин, но не дает возможности судить о том, как количественно меняется одна величина по мере изменения другой. Этот вопрос решают с помощью регрессии. При простой корреляции изучают

зависимость между изменчивостью двух признаков. С помощью регрессии ставят задачу установить количественную взаимосвязь между признаками. Одним из приемов решений этой задачи при наличии прямолинейной зависимости является метод наименьших квадратов.

Зависимость между величинами в этом случае выражается уравнением регрессии, которое имеет вид обычного уравнения прямой линии:

$$Y=aX+v,$$

где Y и X – коррелирующие величины;

a – коэффициент пропорциональности, который показывает степень зависимости Y от X ;

v – первоначальное значение Y при $X=0$.

Для того, чтобы определить численное значение « a » и « v » в уравнении $Y=aX+v$, надо решить систему двух уравнений:

- 1) $Y=axX+vx_n$
- 2) $XU=axX^2+vxX$

Составление этих уравнений основано на методе наименьших квадратов, который позволяет вычислить такие параметры для уравнений регрессии, при которых сумма квадратов отклонений эмпирических значений « Y » от критически вычисленных окажется наименьшей.

Пример. В результате проведенных испытаний ядерного оружия загрязненность почвы по стронцию-90 постепенно возросла до 7,8 мКи/км². По мере увеличения радиоактивности почвы повышалось содержание стронция-90 в фураже и продуктах животноводства, в частности в молоке. Динамика этого процесса представлена в таблице (X – радиоактивность почвы в десятках мКи/км², Y – радиоактивность молока в десятках пКи/л):

X^2	X	Y	XU
1	2	3	4
0,81	0,9	0,43	0,4
0,81	0,9	0,42	0,4
1,7	1,3	0,51	0,7
5,3	2,3	0,87	2,0
9,0	3,0	1,06	3,2
12,2	3,5	1,17	4,1
19,4	4,4	1,37	6,0
17,6	4,2	1,37	5,7
23,0	4,8	1,44	6,9
23,0	4,8	1,47	7,0
1	2	3	4

30,2	5,5	1,54	8,5
28,1	5,3	1,50	7,9
42,2	6,5	1,76	11,4
42,2	6,5	1,62	10,5
51,8	7,2	1,93	13,9
49,0	7,0	1,82	12,7
64,0	8,0	1,94	15,5
44,9	6,7	2,48	16,6
60,8	7,8	2,79	21,8
$X^2=540$	$X=90$	$Y=27,5$	$XY=155$

В левом и правом столбцах таблицы, а так же в нижней строке, приведены данные, полученные в результате обработке двух средних рядов. После подстановки итоговых значений этой таблицы в уравнения 1 и 2 они приобретут следующий вид:

$$1) 27,5=90a+19б$$

$$2) 155=540a+90б$$

Для решения их обычными алгебраическими методами надо умножить коэффициенты уравнения 1 на 6 и вычесть уравнение 2 из уравнения 1.

$$165=540a+114б$$

$$\underline{155=540a+90б}$$

$$10=24б$$

Отсюда «б» равно $\frac{10}{24}=0,42$. После подстановки значения «б» в уравнение 1 получим значение «а»: $27,5=90a+8$, отсюда, $a=0,22$.

В окончательном виде уравнение регрессии будет следующим:

$$Y=0,22X+0,42$$

Если подставить в это уравнение различные значения уровней загрязненности почвы, можно получить соответствующие этим уровням количества стронция-90, содержащегося в молоке. Предположим, что загрязнение почвы составляет 30 (три десятка) пКи остронция-90 на 1км^2 . Тогда содержание радиоизотопа в молоке будет:

$$Y=0,22 \times 30 + 0,42 = 1,08 \text{ (т.е. } 10,8 \text{ пКи/л)}.$$

Расхождение этой величины с приведенным выше в таблице значением не превышает нескольких процентов: $\frac{1,08 - 1,06}{1,08} \times 100\% = 1,85\%$, что подтверждает приемлемость пользования полученным уравнением для определения количественных значений одной величины при измерении другой.

Определение количественной зависимости между варьирующими величинами методом наименьших квадратов может быть использовано для ориентировочного прогнозирования радиоактивной загрязненности продукции сельского хозяйства на следе радиоактивного облака.

Пример. После глобального выпадения радиоактивных веществ загрязненность почвы по стронцию-90 составила 50 (5 десятков) мКи/км². Требуется определить возможное содержание этого нуклида в молоке у коров на данной территории.

Пользуясь уже полученными нами уравнением регрессии, находим величину радиоактивной загрязненности молока:

$$Y=0,22 \times 5 + 0,42 = 1,1 + 0,42 = 1,52 \text{ (т.е. 15,2 пКи/л).}$$

Общая схема статистического анализа.

Математическая обработка результатов исследований имеет целью выявить направленность наблюдаемых сдвигов, оценить их достоверность и определить общие закономерности их изменения.

При оценке результатов измерений необходима их предварительная подготовка с целью ускорения и упрощения дальнейших расчетов. Для этого необходимо выполнить следующие условия:

- результаты измерений расположить в возрастающем или убывающем порядке, т.е. произвести их ранжирование;
- цифровые данные, если они выражены в виде десятичных дробей, привести к целым числам путем умножения на какую-то избранную постоянную величину (10, 100 и т.п.);
- исключить заведомо сомнительные значения исходных данных (как правило, это крайние варианты).

Статистический анализ материала начинают с вычислений средней арифметической величины, описывающей одним числом результаты ряда измерений. Однако, вычисление одного этого статистического показателя, характеризующего ряд наблюдений через их среднее значение, недостаточно для описания полученных результатов измерений в связи с тем, что не учитывается отклонение отдельных членов ряда от средней арифметической величины и самой средней от истинного значения радиоактивности объекта. В связи с этим возникает необходимость оценить погрешность полученных величин и самой средней арифметической, для чего определяют среднее квадратическое отклонение (ς) и стандартную ошибку средней арифметической величины (m), позволяющие судить о степени рассеивания (разброса) наблюдаемых значений.

Наиболее просто значение среднего квадратического отклонения отдельных измерений и стандартной ошибки средней арифметической величины определяют по размаху варьирования, т.е. по разности значений крайних вариантов в каждом ранжированном ряду (эту разность делят соответственно на коэффициенты K_c или K_m).

Одна из задач статистической обработки – оценка достоверности различий средних значений – решается путем вычисления t-критерия по Стьюденту. Разницу между показателями считают достоверной при уровне значимости $P \geq 95\%$.

В тех случаях, когда необходимо установить связь между исследуемыми признаками и оценить ее степень, применяется корреляционный анализ. Количественным выражением степени корреляции является коэффициент корреляции (коэффициент по абсолютной величине не может быть больше единицы).

Применение регрессионного анализа (метод наименьших квадратов) позволяет количественно оценить биологическую закономерность, связывающую признаки между собой и их влияние друг на друга.

Применение критерия χ^2 (хи-квадрат) дает возможность сравнить эмпирические распределения качественных признаков и установить достоверность различий между ними.

Заключительным этапом исследования является обобщение и анализ полученных результатов измерений. От того, насколько всесторонне и глубоко выполнена эта работа с учетом материалов математической обработки цифровых данных, зависит объективность оценок, правомерность и убедительность выводов и рекомендаций.

3) Практическая работа:

Решение предложенных задач.

Рекомендуемая литература:

Основная литература:

1. Игнатенко Г.К. Первичная обработка данных экологического мониторинга. Обнинск: ИАТЭ. – 2006. – 76 с.
2. Халафян А.А. Статистический анализ данных. М.: Бином-Пресс. – 2008. – 512 с.

Дополнительная литература:

1. Джефферс Дж. Введение в системный анализ: применение в экологии. М.: Мир, 1981. – 256 с.
2. Кендалл М., Стюарт А. Многомерный статистический анализ и временные ряды. – М.: Наука. 1976. – 736 с.

3. Рокицкий П.Ф. Биологическая статистика. Минск. Высшейш. шк. – 1973. – 320 с.
4. Шеффе Г. Дисперсионный анализ. М.: Наука. – 1980. – 512 с.