БЕЛОРУССКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ

Биологический факультет

Кафедра общей экологии и методики преподавания биологии

СОГЛАСОВАНО
Председатель учебно-методической комиссии биологического факультета
Поликсенова В.Д.

____

«10» ноября 2015 г.

СОГЛАСОВАНО
Декан биологического факультета
Лысак В.В.

____

«10» ноября 2015 г.

Регистрационный номер № УД-420

УЧЕБНО-МЕТОДИЧЕСКИЙ КОМПЛЕКС ПО УЧЕБНОЙ ДИСЦИПЛИНЕ

Биометрия

для специальностей
1-31 01 01 Биология (по направлениям);
1-31 01 02 Биохимия;
1-31 01 03 Микробиология;
1-33 01 01 Биоэкология

Составители: канд. биол. наук, доцент Жукова А.А.,
канд. биол. наук, доцент Еремова Н.Г.,
ст. преп. Минец М.Л.

Рассмотрено и утверждено
на заседании
Научно-методического совета БГУ

«11» ноября 2015 г.

протокол № 2

РЕЦЕНЗЕНТЫ:

Кафедра общей биологии и ботаники Учреждения образования «Белорусский государственный педагогический университет имени М. Танка;

В.П. Семенченко, заведующий лабораторией гидробиологии Государственного научно-производственного объединения «Научно-производственный центр НАН Беларуси по биоресурсам», доктор биологических наук, член-корреспондент НАН Беларуси

СОДЕРЖАНИЕ

ПОЯСНИТЕЛЬНАЯ ЗАПИСКА

Учебно-методический комплекс (УМК) по учебной дисциплине «Биометрия» создан в соответствии с требованиями Положения об учебно-методическом комплексе на уровне высшего образования и предназначен для студентов специальностей 1-31 01 01 Биология (по направлениям); 1-31 01 02 Биохимия; 1-31 01 03 Микробиология; 1-33 01 01 Биоэкология. Содержание разделов УМК соответствует образовательным стандартам высшего образования данных специальностей. Главная цель УМК – оказание методической помощи студентам в систематизации учебного материала в процессе подготовки к итоговой аттестации по курсу «Биометрия».

Структура УМК включает:

1. Учебно-методическое обеспечение дисциплины

1.1. Теоретический раздел (материалы для теоретического изучения дисциплины в объеме, установленном учебным планом по специальности).

1.2. Практический раздел (материалы для проведения лабораторных занятий по дисциплине в соответствии с учебным планом).

2. Контроль самостоятельной работы студентов (материалы текущей и итоговой аттестации, позволяющие определить соответствие учебной деятельности обучающихся требованиям образовательных стандартов высшего образования и учебно-программной документации, в т.ч. вопросы для подготовки к зачету, задания, тесты, вопросы для самоконтроля и др.).

3. Вспомогательный раздел.

3.1. Учебно-программные материалы (типовая учебная программа, учебные программы (рабочий вариант) для студентов дневной и заочной форм получения образования).

3.2. Информационно-аналитические материалы (список рекомендуемой литературы, перечень электронных образовательных ресурсов и их адреса и др.).

Работа с УМК должна включать на первом этапе ознакомление с тематическим планом дисциплины, представленным в типовой учебной программе. С помощью рабочего варианта учебной программы по дисциплине можно получить информацию о тематике лекций и лабораторных занятий, перечнях рассматриваемых вопросов и рекомендуемой для их изучения литературы.

Для подготовки к лабораторным занятиям и зачету необходимо, в первую очередь, использовать материалы, представленные в разделе учебно-методическое обеспечение дисциплины, а также материалы для текущего контроля самостоятельной работы.

В ходе подготовки к итоговой аттестации рекомендуется ознакомиться с требованиями к компетенциям по дисциплине, изложенными в типовой учебной программе, структурой рейтинговой системы, а также перечнем вопросов для подготовки к зачету.

1. ТЕОРЕТИЧЕСКИЙ РАЗДЕЛ

Содержание курса и учебно-методическая карта представлены в учебной программе по дисциплине «Биометрия» для специальностей 1-31 01 01 Биология (по направлениям); 1-31 01 02 Биохимия; 1-31 01 03 Микробиология; 1-33 01 01 Биоэкология которая доступна по адресу

http://elib.bsu.by/handle/123456789/134345

Там же приведен список литературы для освоения программы курса. Для скачивания и ознакомления в электронной библиотеке доступны лекция 1 "Значение статистических методов в научных исследованиях" **http://elib.bsu.by/handle/123456789/109899**, лекции 10, 11 "Регрессионный анализ" **http://elib.bsu.by/handle/123456789/109901, http://elib.bsu.by/handle/123456789/109902.**

Целью курса является формирование у студентов целостной системы знаний о современных подходах статистического анализа данных; освоение методов, позволяющих выявлять количественные закономерности в биологических явлениях; ознакомление с принципами построения математических моделей биологических явлений и процессов; формирование навыков и умений компьютерной обработки экспериментальных данных; ознакомление с правилами корректного представления результатов исследований; формирование способности к критическому анализу представляемых в публикациях данных.

План-конспект
лекций по курсу «Биометрия»

ЗНАЧЕНИЕ СТАТИСТИЧЕСКИХ МЕТОДОВ В НАУЧНЫХ ИССЛЕДОВАНИЯХ. ТИПЫ ДАННЫХ

Основная цель данного курса – ознакомить вас с основными современными методами статистики, которые могут использоваться в биологических измерениях. Что это вам даст:

1. *Возможность познания количественных закономерностей в биологических явлениях и процессах;*
2. *Основы обработки экспериментальных данных и правила корректного представления их коллегам;*
3. *Избавление от боязни математически оформленных научных статей и критический анализ представленных в них данных;*
4. *Принципы построения математических моделей биологических явлений и процессов.*

Под статистикой будем понимать ***науку, занимающуюся изучением данных, отражающих естественную (природную) изменчивость***.

В настоящее время статистика является обширной и очень бурно развивающейся областью науки. Ее методы сегодня в той или иной мере применяются практически во *всех научных дисциплинах, включая гуманитарные*.

Статистика всегда имеет дело *не с отдельными объектами, а с их совокупностями*. Поэтому если мы, пытаясь определить, например, средний возраст учеников в школе, выясним возраст только у одного из них, это нам ничего не даст. Такое наблюдение не будет представлять для нас интерес. И только *несколько* таких наблюдений (скажем, 100), выполненных для *разных* учеников одним и тем же способом, составят совокупность, называемую *данными*. *Данные – это совокупность измерений (наблюдений), выполненных на объектах одной категории по одинаковой схеме*.

Для обозначения любого составного элемента такой совокупности употребляют термины «*единичное наблюдение*», «*отдельное наблюдение*» или «*варианта*» (но не «вариант»!). **(S)** Математически единичные наблюдения обозначаются как x_i, где i – порядковый номер. Количество наблюдений, составляющих совокупность, называют *объемом* и обозначают латинской буквой n. В случае с учениками $n = 100$, а i, естественно, изменяется от 1 до 100.

То фактическое свойство объекта, которое мы измеряем у него в ходе исследования, называется *переменной*, или *признаком.* В случае с учениками переменной был возраст. Переменными будут также социальное положение человека, его рост, вес, психо-эмоциональный тип и т.д. Заметьте, что в слове «переменная» отражено важнейшее свойство *любых* наблюдений или измерений, а именно то, что они *редко принимают одинаковые значения*. Различие между отдельными наблюдениями называют *вариацией* (от лат. variatio – изменение, колебания). Степень же различий между наблюдениями носит название «*вариабельность*».

Часть вариации любого признака обусловлена *естественными причинами*, например, генетическими различиями организмов, разницей в воспитании, различиями в склонности к заболеваниям и т.п. Однако наряду с естественной изменчивостью на численных значениях признаков сказываются также неточности, которые *неизбежно* возникают при проведении измерений и наблюдений. Разницу между результатом измерения и реальной величиной признака называют *погрешностью* или *ошибкой*. Выделяют две группы ошибок:

- *Случайные ошибки*: они не зависят от воли человека и вызываются причинами, не поддающимися устранению (пример).
- *Систематические ошибки*: обусловлены неисправностями или неточностью измерительных приборов и личными качествами исследователя (недостаточный опыт, усталость, невнимательность). Систематические ошибки поддаются исправлению или, по меньше мере, снижению частоты их появления путем наладки или модернизации оборудования, а также за счет тренинга исследовательского персонала.

Все множество переменных можно классифицировать на перечисленные ниже типы:

А. *Количественные переменны: объединяют такие свойства изучаемых объектов, разные состояния которых можно выразить при помощи чисел*. Существует два вида таких переменных – *непрерывные* и *дискретные*.

Б. *Порядковые* (*ранговые*) *переменные.* Очень многие признаки невозможно измерить количественно, однако при этом некоторые из них *поддаются выстраиванию (ранжированию) в порядке возрастания степени проявления*. С. *Качественные* (= *атрибутивные, номинальные*) *переменные*

В. Качественными называют признаки, которые *невозможно измерить количественно* (черный или белый, живой или мертвый, самец или самка).

Г. Существует и еще один самостоятельный класс переменных, которые называют *индексами* или *производными переменными*. Они являются результатом выполнения *определенных математических операций* над двумя или более независимо измеренными переменными. К производным переменным следует отнести также *удельные скорости* протекания процессов, например, количество вопросов в тесте, пройденных за единицу времени, прирост среднего балла успеваемости за год и т.п.

ОПИСАТЕЛЬНАЯ СТАТИСТИКА

В ряде случаев обработку данных полезно начать с их *группировки*. ***Группировка*** *– это систематизация первичных биометрических данных, направленная на извлечение заключенной в них информации и выявление общих закономерностей, которым подчиняется изучаемое явление или объект*.

Допустим, была изучена плодовитость у 10 крольчих. Варианты x_1, $x_2...x_{10}$, обозначающие количество крольчат у отдельных самок, имели следующие значения:

4 5 3 4 4 2 4 5 3 4

Расположим приведенные данные в порядке возрастания числовых значений, т.е. ***ранжируем*** их, и далее подсчитаем, сколько раз каждая варианта встречается в данной совокупности. Так как плодовитость варьирует от 2 до 5, варианты нашей выборки распределятся следующим образом:

Количество крольчат (x_i):	2	3	4	5
Число вариант (f_i):	1	2	5	2

Этот двойной ряд чисел, показывающий, каким образом числовые значения признака связаны с их *повторяемостью*, называется *вариационным рядом.*

Число, показывающее, сколько раз отдельная варианта встречается в совокупности, называется ***частотой варианты***. Частоты обозначаются буквой f с индексом i, где i – значение варианты. Общая сумма частот всегда равна объему выборки, т.е. $\Sigma f_i=n$.

Из разобранного примера следует, что *при небольшом объеме выборки и незначительной вариации признака, количественные данные достаточно сгруппировать по значениям вариант*. Однако, такой способ разбивки данных *не будет целесообразным при большом количестве вариант и значительной вариации признака*. В подобных случаях прибегают к разнесению данных по *классам*, *которые охватывают сразу несколько значений вариант (****интервалы****)*.

Вариационный ряд показывает исследователю, как часто встречаются значения отдельных вариант и как варьирует изучаемый признак. *Вариационный ряд* позволяет выяснить следующие свойства выборки:

1) *Границы изменчивости признака*: минимальное и максимальное значения вариант, или ***лимиты***. *Разница между лимитами называется* ***размахом выборки.***

2) Исследователь может установить, *какой класс встречается чаще всего, т.е. т.н.* ***модальный класс***. Наиболее часто встречающееся значение называется ***модой (Мо)***. У кроликов, как мы выяснили, мода равна 4. У некоторых распределений может быть больше, чем одна мода.

Для удобства изучения свойств вариационного ряда, *распределение* его вариант в соответствии с частотами встречаемости обычно *изображают графически*. *Графическое изображение вариационного ряда называется* ***кривой распределения, или вариационной кривой.*** Есть два способа изображения вариационных рядов:

1) В случаях, когда данные группируются по значениям отдельных вариант, графическое изображение ряда называется ***полигоном распределения***.

2) Если при группировке данные разносятся по классам, объединяющим несколько вариант, то графическое изображение такого вариационного ряда называется ***гистограммой***.

Существует большое количество ***параметров описательной статистики***. Все они распадаются на две большие группы:

1) *показатели, характеризующие центральную тенденцию в изучаемой совокупности*.
2) *показатели, характеризующие степень изменчивости (=вариабельности) изучаемого признака*.

*С*имметричное колоколообразное распределение называется ***нормальным***. Нормальные распределения встречаются очень часто и имеют огромное значение в статистике. Оказывается, что любое нормальное распределение полностью можно описать при помощи всего двух параметров – *среднего значения* и *стандартного отклонения*.

Характеристика положения распределения на числовой оси называется ***средним значением***. Важная роль средних значений заключается в их свойстве *нивелировать индивидуальные различия единиц совокупности и характеризовать совокупность в целом*.

Существует несколько видов средних. Наиболее обычна в биологических исследованиях ***средняя арифметическая***. Рассчитывается путем суммирования всех вариант выборки и деления этой суммы на объем выборки:

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

Необходимо понимать, что *средняя арифметическая* как таковая имеет смысл только тогда, когда она рассчитывается *для качественно однородной совокупности*. Для группы, которая включает неодинаковое число и молодых, и взрослых особей, нужно подсчитать среднюю арифметическую для каждой возрастной группы, а затем, если все же нужно найти усредненный показатель для всей совокупности, найти ***взвешенную среднюю***:

$$\bar{x} = \frac{\bar{x}_1 n_1 + \bar{x}_2 n_2 + ... + \bar{x}_k n_k}{\sum n_k}$$

где x_1, x_2 … x_k – средние арифметические частных выборок, а n_1, n_2 … n_k – ***объемы*** (=***веса***) этих частных совокупностей. Всего совокупностей – k.

Для вычисления средних величин различных *приростов* (веса, размеров тела, численности популяции и т.п.) за определенные промежутки времени следует использовать с***реднюю геометрическую*** (обычная средняя в таких случаях дает завышенные результаты):

$$\bar{x}_g = \sqrt[n]{x_1 \times x_2 \ldots \times x_n}$$

Далее, чтобы получить обобщенную оценку разброса данных, можно использовать параметр, который называется ***дисперсия***. С ее помощью уже можно охарактеризовать степень разброса данных в совокупности. *Чем больше разброс значений, тем больше дисперсия*. Однако, дисперсия измеряется в единицах, равных *квадрату единицы измерения* исходного признака. Поэтому чаще используют корень квадратный из дисперсии, который называют ***стандартным отклонением***:

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

Буквой s будем обозначать выборочное стандартное отклонение. В случае с генеральной совокупностью используют греческую букву σ.

При работе с *небольшими выборками* применение выведенной нами формулы приводит к *смещенной оценке* стандартного отклонения. Поэтому в знаменатель вносят поправку:

$$s^2 = \frac{\sum (x_i - \bar{x})^2}{n-1}; s = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n-1}}$$

Если *распределение* **ассиметрично,** полагаться на среднее и стандартное отклонение нельзя, иначе мы получим превратное представление о совокупности, не подчиняющейся *нормальному распределению*. Для описания таких данных лучше подходит ***медиана***. *Она представляет собой такое значение, которое делит распределение ровно пополам: половина значений больше медианы, половина – не больше*. Обозначается как *Me*. Для характеристики разброса значений найдем *значения, не выше которых оказались 25% и 75% всех результатов измерений* – эти величины называются 25-м и 75-м *процентилями*. Медиана и процентили, в отличие от среднего и стандартного отклонения, не дают полного описания распределения. Для описания распределения чаще всего применяют 25-й и 75-й процентили. Однако можно рассчитывать и любые другие процентили. Например, в качестве границ нормы лабораторных показателей часто используют 5-й и 95-й процентили.

Однако очень многие признаки невозможно измерить численно. Например, можно быть только мужчиной или женщиной, мертвым или живым, и т.п. Здесь мы имеем дело с *качественными признаками*. Единственный способ описания качественных признаков состоит в том, чтобы подсчитать *число* объектов, имеющих одно и то же значение. Кроме того, можно подсчитать, какая *доля* от общего числа объектов приходится на то или иное значение. Для характеристики совокупности, которая состоит из двух классов, *достаточно указать численность одного из них p*. Заметим еще, что *p* есть еще и *вероятность* того, что случайно выбранный марсианин окажется розовым. При этом доля *p* в некотором смысле аналогична среднему по совокупности значений количественного признака.

Не совсем ясно, что в данном случае понимать под разбросом, если значений признака всего два. Не вдаваясь в математическую сторону вопроса, отметим, что стандартное отклонение при качественной изменчивости вычисляется по следующей формуле:

$$s = \sqrt{pq}\ ,$$

или, если есть необходимость выразить встречаемость каждой из альтернативы в абсолютных величинах, то по следующей формуле:

$$s = \sqrt{npq}\ ,\ \text{где } n \text{ – объем выборки.}$$

Таким образом, стандартное отклонение полностью определяется величиной *p*. Этим оно принципиально отличается от стандартного отклонения для нормального распределения, которое не зависит от среднего значения.

Иногда в биологических исследованиях возникает необходимость сравнить степень вариабельности качественно разнородных признаков. Одним из таких показателей является ***коэффициент вариации***, предложенный Карлом Пирсоном. Вычисляется как отношение стандартного отклонения к средней арифметической, выраженное в процентах:

$$CV = \frac{s}{\bar{x}}100\%$$

В большинстве случаев величины *коэффициента вариации* в биологических исследованиях *не превышают 20*%. Если коэффициент превышает это значение, мы можем быть уверены, что имеем дело с очень разнородной выборкой.

Для характеристики точности выборочных оценок используют ***стандартную ошибку***. Этот показатель можно рассчитать для любого статистического показателя, но мы остановимся на *стандартной ошибке средней арифметической*, поскольку именно в этой величине чаще всего возникает необходимость при биологических исследованиях. ***Центральная предельная теорема*** гласит следующее: *выборочные средние имеют нормальное распределение независимо от распределения исходной совокупности и объема выборок*. Поскольку распределение нормальное, его можно охарактеризовать с помощью *средней арифметической* и *стандартного отклонения*. *Среднее выборочных средних совпадает с генеральным средним значением*. $S_{\bar{x}}$ является мерой точности, с которой выборочное среднее оценивает генеральную среднюю μ. Поэтому $S_{\bar{x}}$ носит название *стандартной ошибки средней арифметической*.

Выборочные средние варьируют в √n раз меньше, чем варианты генеральной совокупности:

$$s_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$$

Поскольку на практике генеральные параметры нам зачастую неизвестны, мы пользуемся выборочными оценками.

$$s_{\bar{x}} = \frac{s}{\sqrt{n}}$$

Итак, мы получили формулу, с помощью которой и рассчитывается стандартная ошибка средней арифметической при выполнении биологических исследований. Из формулы расчета стандартной ошибки хорошо видно*, что по мере увеличения числа наблюдений величина стандартной ошибки будет уменьшаться.* Свои стандартные ошибки имеют и другие статистические показатели, поскольку все они являются *выборочными*: можно было бы высчитать стандартную ошибку коэффициента вариации, дисперсии и т.д.

ЗАКОНЫ РАСПРЕДЕЛЕНИЯ ВЕРОЯТНОСТЕЙ СЛУЧАЙНЫХ ВЕЛИЧИН

В биологических исследованиях все предсказания и суждения строят на основе определенных ***теоретических распределений вероятностей***. В каждом конкретном случае выбирается свое теоретическое распределение вероятностей. Если экспериментальные данные плохо соответствуют *теоретическому распределению*, мы должны сильно усомниться в том, действительно ли на изучаемое явление влияют те факторы, которые мы предполагаем.

Вспомним несколько понятий из теории вероятностей.

***Событие** – результат, или исход отдельного испытания.*

*Несколько событий называются **несовместимыми**, если в условиях испытания каждый раз возможно наступление только одного из них.* Иначе события будут ***совместимыми***.

*Два события называются **противоположными**, если наступление любого из них исключает появление другого.*

*Если при каждом испытании событие наступает неизбежно, то оно называется **достоверным**. Если в заданных условиях событие произойти не может – оно называется **невозможным**. Если же событие в каждом отдельном испытании может произойти, а может и не произойти, его называют **случайным**.*

***Вероятностью** называется отношение числа случаев, благоприятствующих наступлению ожидаемого события, к числу всех возможных исходов:*

$$\mathrm{P}(A) = m/n\ ,$$

По определению, вероятность может изменяться от 0 до 1. Очевидно, что нулю равна вероятность *невозможного* события, тогда как *достоверное* событие наступает с вероятностью, равной 1. Очевидно также, что вероятность *случайного события* лежит между 0 и 1. Вероятность ожидаемого события принято обозначать p, а противоположного ему – q. $p + q = 1$.

По нормальному закону распределяются вероятности многих *непрерывных биологических признаков*, т.е. признаков, числовые значения которых могут принимать не только целые, но и дробные числовые значения.

Вероятность определенного значения непрерывно распределяющейся случайной величины выражается следующей формулой:

$$P(x_i) = \frac{1}{\sigma\sqrt{2\pi}} e^{-1/2[(x_i-\mu)/\sigma]^2}$$

Уравнение, описывающее нормальное распределение вероятностей, содержит две постоянные – $\pi = 3{,}14$ и e - основание натурального логарифма = 2,72. Остается только *два параметра, от которых зависит форма нормального распределения – это стандартное отклонение σ и величина μ.*

Колоколообразную кривую, которой изображается нормальное распределение, называют еще *нормальной кривой*. В идеале нормальная кривая строго симметрична относительно ее центра – *генеральной средней*. Однако *эмпирические кривые* могут быть *скошены* в ту или иную сторону. Степень этой «скошенности» выражает т.н. ***коэффициент асимметрии***:

$$g_1 = \frac{\sum (x_i - \bar{x})^3}{ns^3}$$

Есть и еще один коэффициент, характеризующий форму эмпирической нормальной кривой – это так называемый ***коэффициент эксцесса***. Он показывает *степень пологости* нормальной кривой и вычисляется по формуле:

$$g_2 = \frac{\sum (x_i - \bar{x})^4}{ns^4} - 3$$

Большинство классических статистических методов исходят из нормальности распределения данных. Учитывая важность этого распределения, необходимо знать некоторые его основные свойства:

1) *Для нормального распределения характерно совпадение по абсолютной величине средней арифметической, медианы и моды*.

2) Другое важное свойство нормального распределения заключается в том, что *практически все варианты (более 99%) укладываются в интервал плюс-минус три стандартных отклонения относительно среднего значения*. Это свойство известно как ***правило трех сигм***.

Закон Пуассона описывает вероятность появления случайных и редких событий. Примером таких событий может быть частота возникновения мутаций у бактерий за одну генерацию, количество случаев гриппа, зарегистрированное за месяц в летнее время, частота рождения троен у людей и т.д.

Математическое выражение *закона Пуассона* можно представить следующим образом:

$$P_n(m) = \frac{a^m}{m!e^a},$$

где *m* – частота ожидаемого события в *n* независимых испытаниях, $a \cong np$ – средняя арифметическая редкого события, $e = 2{,}7183$ – основание натурального логарифма, *m*! – факториал частоты ожидаемого события (произведение натуральных чисел от 0 до *m*).

ДИСПЕРСИОННЫЙ АНАЛИЗ

Методы, позволяющие оценивать статистическую значимость различий между выборками, называют *тестами*. В ходе реализации того или иного теста обычно рассчитывается некая величина, называемая *статистическим критерием*. На основе критерия судят о величине разницы между сравниваемыми группами. Существуют две большие группы статистических критериев:

1) ***параметрические***: их расчет основан на параметрах, характеризующих распределение выборочных единиц (отсюда и название), т.е. они представляют собой функции от этих параметров. Применимы в тех случаях, когда генеральные совокупности, из которых взяты анализируемые выборки, распределяются *по нормальному закону*.
2) ***непараметрические***: указных выше ограничений не имеют, т.е. они не требуют, чтобы анализируемые данные подчинялись нормальному закону распределения.

В сравнении с непараметрическими, параметрические критерии обладают большей мощностью, т.е. способностью выявлять различия между сравниваемыми выборками, и поэтому им следует отдавать предпочтение.

Дисперсионный анализ (ANOVA – *analysis of variance*) предназначен для *одновременного сравнения средних арифметических нескольких выборок* (2 и более). Анализ полученных данных при использовании любого статистического критерия начинается с формулировки т.н. *нулевой гипотезы* H_0. Она называется так потому, что предполагает, что между сравниваемыми параметрами разница крайне несущественна, равна нулю. Чтобы оценить величину различий, нужно каким-то образом *сравнить разброс выборочных средних с разбросом значений внутри групп*. Дисперсию совокупности можно оценить двумя способами. Дисперсию генеральной совокупности можно оценить на основании групповых дисперсий. Такая оценка никоим образом не будет зависеть от различий групповых средних. *С другой стороны, разброс выборочных средних тоже позволяет оценить дисперсию генеральной совокупности.* Понятно, что такая оценка уже будет зависеть от различий выборочных средних.

Оценим дисперсию совокупности *по выборочным средним*. Эта оценка называется *межгрупповой дисперсией*:

$$s^2_{\text{меж}} = n\, s^2_x,$$

где s^2_x – квадрат стандартного отклонения выборки из выборочных средних.

Межгрупповую дисперсию называют еще *факториальной дисперсией*, т.е. обусловленной влиянием изучаемого фактора.

Если верна нулевая гипотеза, то как внутригрупповая, так и межгрупповая дисперсии служат оценками одной и той же дисперсии генеральной совокупности и должны быть приближенно равны. Посмотрим, так ли это, путем вычисления т.н. *F-критерия,* или *критерия Фишера*:

F = [Дисперсия совокупности, оцененная по выборочным средним] / [Дисперсия совокупности, оцененная по выборочным дисперсиям], или

$$F = s^2_{\text{меж}} \,/\, s^2_{\text{вну}}$$

Поскольку и числитель, и знаменатель этого отношения – это оценки одной и той же величины – дисперсии совокупности σ^2, поэтому значение F должно быть близко к 1. Если F значительно превышает 1, *нулевую гипотезу следует отвергнуть и принять альтернативную. Значение любого статистического критерия, начиная с которого мы отвергаем нулевую гипотезу, называется* <u>*критическим значением*</u>.

Вероятность справедливости нулевой гипотезы обозначается буквой P. В большинстве биологических исследований считают достаточным, чтобы эта вероятность не превышала 5%. Эта максимальная приемлемая вероятность ошибочно отвергнуть нулевую гипотезу называется *критическим уровнем значимости* и обозначается α. Ее сравнивают с *критическим уровнем значимости* α. Если P оказывается больше, чем α, то нулевая гипотеза сохраняется. И наоборот.

Критическое значение F однозначно определяется критическим уровнем значимости α и еще двумя параметрами, которые называются *внутригрупповым* и *межгрупповым числом степеней свободы* υ. Межгрупповое число степеней свободы определяется по уже известной нам формуле: $\upsilon_{\text{меж}} = m - 1$. Внутригрупповое число степеней свободы: $\upsilon_{\text{вну}} = m(n - 1)$.

СРАВНЕНИЕ ДВУХ ГРУПП И МНОЖЕСТВЕННЫЕ СРАВНЕНИЯ

Критерий Стьюдента, или *t*-критерий:

$$t = \frac{\bar{x}_1 - \bar{x}_2}{s_{\bar{x}_1 - \bar{x}_2}}$$

Для двух случайных выборок, извлеченных из одной генеральной совокупности, это отношение, *как правило, будет близко к нулю, т.к. близка к нулю будет разница между средними. Чем меньше (по абсолютной величине) t, тем больше вероятность нулевой гипотезы*. Для нахождения величины t нужно знать разность выборочных средних и стандартную ошибку их разности. Стандартная ошибка разности двух средних вычисляется следующим образом:

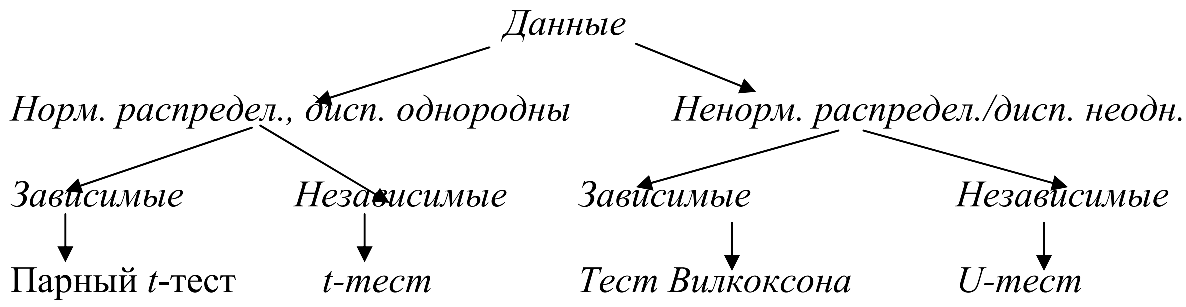
$$s_{\bar{x}_1 - \bar{x}_2} = \sqrt{s^2_{\bar{x}_1} + s^2_{\bar{x}_2}}$$

Если две выборки извлечены из одной и той же совокупности, то вероятность получить значение t, большее +2.1 или меньшее -2.1, составляет всего 5%. Следовательно, *если значение t находится вне предела интервала от* -

2.1 *до* +2.1, *нулевую гипотезу следует отклонить, а наблюдаемые различия признать статистически значимыми*.

Критические значения t сведены в специальные таблицы и определяются, исходя из *уровня значимости* и *числа степеней свободы*. Число степеней свободы для t-критерия определяется как $n_1 + n_2 - 2$. *Чем больше объем выборок, тем меньше критическое значение t*.

Два непараметрических теста, позволяющих сравнить две выборки – это *U-тест Манна-Уитни (Mann-Whitney U-test) и тест Вилкоксона (Wilcoxon matched pairs test)*. Оба эти критерия являются ранговыми. Алгоритм выбора корректного теста:



АНАЛИЗ ЧАСТОТ

Для сравнения двух выборочных долей используют z-критерий:

$$z = \frac{p_1 - p_2}{s_{p_1 - p_2}}$$

Стандартная ошибка разности долей рассчитывается аналогично ошибке разности средних **(S)**:

$$s_{p_1 - p_2} = \sqrt{s_{p_1}^2 + s_{p_2}^2}$$

Если сравниваемые выборки имеют объемы n_1 и n_2, то дисперсии этих выборок равны **(S)**:

$$s_{p_1} = \sqrt{\frac{p_1(1-p_1)}{n_1}} \text{ и } s_{p_2} = \sqrt{\frac{p_2(1-p_2)}{n_2}}$$

Развернутую формулу для вычисления z-критерия мы можем записать следующим образом:

$$z = \frac{p_1 - p_2}{\sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}}}$$

Критерий χ^2 (читается «хи-квадрат») вычисляется следующим образом:

$$\chi^2 = \sum \frac{(F-T)^2}{T}$$

где F – фактическая (= наблюдаемая), а T - теоретически ожидаемая частота. 3.84 – критическое значение χ^2 для 5%-ного уровня значимости. Применение χ^2 правомерно, если *ожидаемое число в любой из ячеек таблицы больше или равно 5* (иначе мы должны прибегнуть к точному критерию

Фишера). Критическое значение χ^2 зависит от размеров таблицы сопряженности, то есть от числа строк и столбцов. Размер таблицы выражается числом степеней свободы: $\upsilon = (R-1)(C-1)$, где R – количество строк, C – количество столбцов в таблице. Легко подсчитать, что для 2х2 таблицы имеем $\upsilon = 1$. Критические значения критерия для других значений числа степеней свободы можно найти в специальных таблицах.

Когда число наблюдений невелико, используют *точный критерий Фишера*. Он основан на *переборе всех возможных вариантов заполнения таблицы* сопряженности при данной численности групп, поэтому чем она меньше, тем проще его применить.

$$P = \frac{\dfrac{R_1!R_2!C_1!C_2!}{N!}}{O_{11}!O_{12}!O_{21}!O_{22}!}$$

Где R_1 и R_2 – *маргинальные* строковые суммы, C_1 и C_2 – *маргинальные* суммы по столбцам, O_{11}, O_{12}, O_{21} и O_{22} – числа в клетках, N – общее число наблюдений. Восклицательный знак, как и всегда в математике, обозначает факториал (факториал числа – произведение целых чисел от этого числа до единицы. Факториал нуля равен 1). Построив все остальные варианты заполнения таблицы, возможные при данных суммах по строкам и столбцам, по этой же формуле рассчитывают и их вероятность. *Вероятности, которые не превосходят вероятность исходной таблицы (включая саму эту вероятность), суммируют*. Полученная сумма – это и есть величина P для *двустороннего варианта* точного критерия Фишера.

ДОВЕРИТЕЛЬНЫЕ ИНТЕРВАЛЫ. РАСЧЕТ РЕПРЕЗЕНТАТИВНОГО ОБЪЕМА ВЫБОРКИ

Мы выяснили, что истинное среднее при нормальном распределении в 95% случаев лежит на расстоянии не больше двух стандартных ошибок среднего от выборочного среднего. Этот промежуток длиной в 4 стандартные ошибки и есть доверительный интервал. *Смысл доверительного интервала* из этого примера достаточно ясен: мы не знаем точно, чему равна некоторая величина, но с заданной вероятностью можем указать интервал, в котором она находится. 95%-ный доверительный интервал для разности средних определяется неравенством

$$(\bar{x}_1 - \bar{x}_2) - t_{0,05} s_{\bar{x}_1 - \bar{x}_2} < \mu_1 - \mu_2 < (\bar{x}_1 - \bar{x}_2) + t_{0,05} s_{\bar{x}_1 - \bar{x}_2}$$

В этот интервал разность истинных средних попадает в 95% случаев, т.е. с вероятностью 95%. Значения левой и правой частей приведенного неравенства называют, соответственно, *нижним* и *верхним доверительным пределами* и обозначают L_1 и L_2.

Формула для 100(1 – α)-процентного доверительного интервала для средней арифметической:

$$\bar{x} - t_\alpha s_{\bar{x}} < \mu < \bar{x} + t_\alpha s_{\bar{x}},$$

где t_α - критическое значение t для уровня значимости α и числа степеней свободы
$\upsilon = n - 1$ (n – объем выборки).

Аналогичную же формулу мы можем записать и для нахождения доверительного интервала доли:

$$p - t_\alpha s_p < p < p + t_\alpha s_p$$

Смысл доверительного интервала для среднего (доли) совершенно аналогичен смыслу доверительного интервала для разности средних (долей).

Если объем выборки достаточно велик (от 20 и выше), то можно воспользоваться «правилом двух стандартных ошибок» и в качестве t_α использовать $t_{0.05}$, приблизительно равное двум.

Установлено, что необходимая численность выборки n, отвечающая определенной точности, *зависит от величины стандартной ошибки средней арифметической*:

$$n = \frac{t_\alpha{}^2 s^2}{\Delta^2}$$

Здесь t – критерий Стьюдента, связанный с определенным уровнем значимости, s^2 - дисперсия, а дельта – желаемая точность, определяемая по формуле:

$$\Delta = t_\alpha s_{\bar{x}}$$

Для определения необходимого объема выборки нужно выполнить *предварительные наблюдения или измерения*, в ходе которых будет оценена величина стандартной ошибки.

Чтобы уменьшить ошибку выборочно средней в К раз, объем выборки нужно увеличить в К^2 раз. При определении необходимого объема выборки в случае с качественными признаками используется формула:

$$n = \frac{t^2 p(1-p)}{\Delta^2}$$

КОРРЕЛЯЦИОННЫЙ АНАЛИЗ

В природе существуют два основных типа зависимостей между переменными:

Функциональная зависимость. При такого рода зависимости определенному значению одной переменной (*аргументу*) соответствует одно единственное значение другой переменной (*функции*). *Статистическая (корреляционная) зависимость*, когда *определенному значению одного признака соответствует не одно, а целая гамма значений другого признака*. Корреляционные зависимости выявляются *только на групповых* объектах с использованием методов математической статистики.

Корреляционная связь между признаками может быть линейной и криволинейной (нелинейной), положительной и отрицательной.

Показатель, который обозначается буквой *r* называется *коэффициентом корреляции*:

$$r = \frac{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{s_x \cdot s_y} \Rightarrow$$

$$r = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum (x_i - \bar{x})^2 \sum (y_i - \bar{y})^2}}$$

Коэффициент корреляции – очень удобный показатель степени взаимосвязи между признаками, получивший *широкое применение* на практике. Это безразмерное число, которое изменяется от -1 до +1. При независимом варьировании признаков, когда связь между ними отсутствует, $r = 0$. Чем сильнее связь между признаками, тем больше и величина коэффициента корреляции. При этом в пределах от 0 до +1 связь между признаками будет в той или иной степени положительной, а в пределах от 0 до -1 – отрицательной.

О тесной корреляции можно говорить только когда r не ниже 0,7. Коэффициенты корреляции порядка 0,5-0,6 - средние, ниже 0,5 - указывают на слабую связь.

Получаемый коэффициент корреляции *r* всегда является выборочным показателем, поскольку вычисляется на основе двух выборок из двух генеральных совокупностей. Следовательно, выборочный коэффициент корреляции служит лишь оценкой генерального параметра ρ. Степень расхождения между выборочным коэффициентом корреляции *r* и генеральным параметром ρ можно, как всегда, оценить *при помощи стандартной ошибки*. Стандартная ошибка коэффициента корреляции вычисляется по формуле:

$$s_r = \frac{1 - r^2}{\sqrt{n}}$$

Данная формула применима для больших выборок. Если же объем выборки не превышает 100 наблюдений, то используют другую формулу:

$$s_r = \sqrt{\frac{1 - r^2}{n - 2}}$$

По определению, чем меньше стандартная ошибка, тем точнее выборочный коэффициент корреляции оценивает генеральный параметр ρ.

При сравнении коэффициентов корреляции, вычисленных на независимых выборках, *нулевая гипотеза заключается в том, что обе эти выборки происходят из одной и той же генеральной совокупности с параметром ρ*. Эта нулевая гипотеза проверяется, опять-таки, с помощью *t*-критерия Стьюдента:

$$t = \frac{r_1 - r_2}{\sqrt{s_{r_1}^2 + s_{r_2}^2}}$$

Ошибки сравниваемых коэффициентов вычисляются по одной из приведенных выше формул (в зависимости от объема выборки, см. выше).

Более точная оценка разности между коэффициентами, полученными на малых выборках, получается при использовании *z*-преобразования:

$$t = \frac{z_1 - z_2}{\sqrt{\frac{1}{n_1 - 3} + \frac{1}{n_2 - 3}}}$$

Существуют непараметрические коэффициенты, которые позволяют измерять степень сопряженности между признаками *независимо от закона распределения и формы связи*. Одним из таких показателей, наиболее часто используемых на практике, является ранговый *коэффициент корреляции Спирмена*:

$$r_s = 1 - \frac{6\sum d^2}{n(n^2 - 1)}$$

где d – разность между рангами сопряженных значений признаков X и Y, а n – число парных наблюдений, или объем выборки.

Установить тесноту связи между двумя качественными признаками позволяет ***коэффициент ассоциации*** (*фи*):

$$\varphi = \frac{ad - bc}{\sqrt{(a+b)(c+d)(a+c)(b+d)}}$$

где *a, b, c, d* – частоты вариант, распределенные по клеткам 2х2-таблицы.

РЕГРЕССИОННЫЙ АНАЛИЗ

Уравнение $y = a + bx$ является наиболее общим типом прямолинейной зависимости. При $y = bx$ мы полагаем, что $a = 0$, а в случае с уравнением $y = x$, $a = 0$ и $b = 1$.

В биометрии величину b называют ***коэффициентом регрессии***, а само уравнение $y = a + bx$ – ***регрессионным уравнением***, или просто ***регрессией***. Величина x в биометрии называется ***независимой переменной***, а величина y – ***зависимой переменной***.

Задача регрессионного анализа состоит в том, чтобы любую форму корреляционной связи *выразить уравнением* определенной функции. Непосредственно же в ходе регрессионного анализа определяются величины коэффициента регрессии и свободного члена регрессионного уравнения и оценивается их статистическая значимость.

Выборочная оценка разброса точек вокруг прямой регрессии определяется по формуле:

$$s_{y|x} = \sqrt{\frac{\sum[y_i - (a + bx_i)]^2}{n-2}}$$

где $(a + bx_i)$ – среднее значение признака y в точке x_i, $y_i - (a + bx_i)^2$ - расстояние от точки до регрессионной прямой.

Величина $s_{y|x}$ называется *остаточным стандартным* отклонением. Соответственно, $s^2_{y|x}$ называется *остаточной дисперсией*.

2. ПРАКТИЧЕСКИЙ РАЗДЕЛ

Лабораторные занятия по курсу проводятся в компьютерных классах. Перечень лабораторных занятий включает 7 разделов по 4 часа каждое.

С темами лабораторных занятий можно ознакомится в учебной программе (рабочий вариант) по дисциплине «Биометрия» по специальностям 1-31 01 01 Биология (по направлениям); 1-31 01 02 Биохимия; 1-31 01 03 Микробиология; 1-33 01 01 Биоэкология, которая доступна по адресу

http://elib.bsu.by/handle/123456789/20653

Материалы, необходимые для подготовки к выполнению лабораторных работ, содержатся в Методическом пособии по использованию программы STATISTICA при обработке данных биологических исследований / Мастицкий С.Э. Мн., 2009. 76 с., которое также доступно по адресу **http://bio.bsu.by/ecology/files/courses/biometrics/labinstr.pdf**

3. КОНТРОЛЬ САМОСТОЯТЕЛЬНОЙ РАБОТЫ СТУДЕНТОВ

Тесты для самоконтроля

1. Объем выборочной совокупности обозначается (выберите правильный ответ):
а) α
б) n
в) m
г) p

2. Какой из перечисленных ниже биологических признаков является ранговым?
а) длина конечности
б) число лепестков в цветке
в) вкус яблока
г) скорость роста популяции

3. Какие из перечисленных ниже показателей являются мерой центральной тенденции в исследуемой совокупности?
а) стандартная ошибка средней арифметической
б) дисперсия
в) медиана
г) стандартное отклонение
д) коэффициент эксцесса

4. Генеральное среднее значение обозначается:

а) α
б) μ
в) σ
г) р

5. Стандартное отклонение генеральной совокупности обозначается:
а) α
б) μ
в) σ
г) р

6. Какие из перечисленных ниже показателей являются мерой разброса значений изучаемого признака?
а) стандартная ошибка средней арифметической
б) дисперсия
в) медиана
г) стандартное отклонение
д) коэффициент эксцесса

7. Стандартная ошибка средней отражает:
а) степень различия между сравниваемыми выборками
б) неточность измерительного прибора
в) точность средней
г) отсутствие необходимого опыта по выполнению измерений у персонала лаборатории

8. Среднее значение характеризует:
а) разброс значений изучаемого признака
б) положение распределения на числовой оси
в) ширину основания распределения
г) какое значение признака обладает наибольшей частотой встречаемости

9. У нормального распределения коэффициент асимметрии равен:
а) 0
б) +1
в) -1
г) >1
д) <1

10. У нормального распределения коэффициент эксцесса равен:
а) 0
б) +1
в) -1
г) >1

д) <1

11. Какой процент наблюдений укладывается в интервал от $-\sigma$ до $+\sigma$ у нормального распределения?
а) 68%
б) 99%
в) 95 %
г) 97%
12. Какие из тестов реже выявляют различия между сравниваемыми выборками?
а) параметрические
б) непараметрические

13. Дисперсионный анализ предназначен для нахождения разницы между:
а) коэффициентами вариации двух выборок
б) арифметическими средними двух и более выборок
в) стандартными ошибками двух и более выборок
г) стандартными отклонениями двух выборок

14. Нулевую гипотезу следует отклонить, если рассчитанный в ходе теста статистический критерий:
а) не попадает в заданную область критических значений
б) попадает в заданную область критических значений
в) равен 0
г) меньше 0

15. Значение статистического критерия, начиная с которого нулевая гипотеза отклоняется, называется:
а) альтернативным
б) достоверным
в) статистически значимым
г) критическим
д) нулевым

16. Уровень значимости обозначается:
а) α
б) β
в) σ
г) φ

17. Коэффициент ассоциации обозначается:
а) α
б) β
в) σ
г) φ

18. Тест Стьюдента предназначен для нахождения разницы между:
а) коэффициентами вариации двух выборок
б) арифметическими средними двух выборок
в) стандартными ошибками двух выборок
г) стандартными отклонениями двух выборок

19. В уравнении $y = a + bx$ коэффициент b называют:
а) коэффициентом регрессии
б) уровнем значимости
в) коэффициентом детерминации
г) стандартной ошибкой коэффициента регрессии
д) свободным членом уравнения

20. В уравнении $y = a + bx$ коэффициент a называют:
а) коэффициентом регрессии
б) уровнем значимости
в) коэффициентом детерминации
г) стандартной ошибкой коэффициента регрессии
д) свободным членом уравнения

21. Какое из приведенных уравнений является уравнением множественной регрессии?
а) Y = a + bX + cZ
б) Y = a + bX
в) Y = bX
г) Y = a

22. Оценивается степень связи между двумя нормально распределенными признаками. Какой коэффициент корреляции для этого лучше использовать?
а) коэффициент Пирсона
б) коэффициент Спирмена

23. Вероятность справедливости нулевой гипотезы обозначается:
а) α
б) μ
в) σ
г) p

24. 95%-ная доверительная область линии регрессии показывает:
а) где с вероятностью 0.95 проходит истинная линия регрессии
б) разброс 95% выборочных точек относительно линии регрессии
в) разброс 95% значений зависимой переменной, которые предсказываются регрессионной моделью

25. Какое из перечисленных ниже условий должно выполняться в отношении линии регрессии?
а) сумма квадратов расстояний от выборочных точек до линии регрессии должна быть минимальной
б) сумма квадратов расстояний от выборочных точек до линии регрессии должна быть максимальной

26. Кластерный анализ, основан на измерении «расстояния» между «объектами» в многомерном пространстве. Чем больше это «расстояние», тем...
а) меньше объекты различаются между собой
б) больше объекты различаются между собой

27. Статистическая ошибка I рода состоит в:
а) сохранении ложной нулевой гипотезы
б) отклонении верной нулевой гипотезы

28. Статистическая ошибка II рода состоит в:
а) сохранении ложной нулевой гипотезы
б) отклонении верной нулевой гипотезы

29. Размах это:
а) разница между максимальным и минимальным значениями признака
б) интервал значений признака между 25-м и 75-м процентилями
в) расстояние между минимальным значением признака и центром распределения
г) расстояние между максимальным значением признака и центром распределения

30. С помощью какого из перечисленных методов можно проверить, подчиняются ли анализируемые данные закону нормального распределения?
а) критерий хи-квадрат
б) критерий Уилкоксона
в) критерий знаков
г) критерий Тьюки

Вопросы для подготовки к зачету

1. Статистика как наука. Значение статистических методов в научных исследованиях.
2. Краткая история развития статистики
3. Данные и их типы
4. Генеральная совокупность и выборка.

5. Случайный отбор объектов из генеральной совокупности. Способы рандомизации.
6. Группировка данных
7. Среднее значение и стандартное отклонение
8. Медиана и процентили
9. Показатели описательной статистики при качественной изменчивости
10. Коэффициент вариации
11. Стандартная ошибка
12. Представление средних величин, мер разброса и стандартных ошибок в научных публикациях
13. Эмпирические и теоретические распределения вероятностей случайных величин
14. Вероятности и их свойства
15. Закон нормального распределения вероятностей
16. Биномиальное распределение
17. Негативное биномиальное распределение
18. Закон Пуассона
19. Параметрические и непараметрические критерии
20. Дисперсионный анализ (ANOVA): постановка задачи
21. Две оценки дисперсии в ANOVA
22. Критическое значение F-критерия
23. Статистические ошибки I и II рода
24. Трансформация данных
25. Принцип теста Стьюдента (t-теста)
26. Критическое значение t
27. Типичные ошибки в использовании критерия Стьюдента
28. Непараметрические методы сравнения двух выборок
29. Методы множественных сравнений
30. Анализ частот: z-критерий
31. Таблицы сопряженности: критерий $\chi 2$
32. Способы определения нормальности распределения данных
33. Точный критерий Фишера (Fisher’s exact test)
34. Доверительный интервал для разности средних и долей
35. Доверительный интервал для средней арифметической и доли
36. Проверка гипотез с помощью доверительных интервалов
37. Расчет репрезентативного объема выборки
38. Основные типы зависимостей между переменными
39. Коэффициент корреляции
40. Статистическая значимость коэффициента корреляции
41. z-преобразование Фишера
42. Минимальное число наблюдений для планируемой точности коэффициента корреляции
43. Сравнение двух коэффициентов корреляции
44. Коэффициент корреляции Спирмена

45. Корреляция между качественными признаками
46. Общее представление о регрессии
47. Оценка параметров регрессионного уравнения по выборке
48. Разброс значений вокруг регрессионной прямой
49. Стандартные ошибки коэффициентов регрессионного уравнения
50. Оценка статистической значимости регрессии
51. Оценка значимости регрессии с помощью доверительных интервалов
52. Доверительная область для линии регрессии
53. Дисперсионный анализ регрессии
54. Анализ остатков
55. Связь регрессии и корреляции
56. Понятие о множественной и нелинейной регрессии
57. Понятие о многомерной совокупности
58. Кластерный анализ
59. Дискриминантный анализ
60. Анализ главных компонент

4. ВСПОМОГАТЕЛЬНЫЙ РАЗДЕЛ

Учебно-программные материалы

Учебная программа по дисциплине «Биометрия» по специальностям 1-31 01 01 Биология (по направлениям); 1-31 01 02 Биохимия; 1-31 01 03 Микробиология; 1-33 01 01 Биоэкология доступна по адресу

http://elib.bsu.by/handle/123456789/134345

Список рекомендуемой литературы и Интернет-ресурсов

Список рекомендуемой литературы и Интернет-ресурсов приведен в учебной программе по дисциплине «Биометрия» по специальностям 1-31 01 01 Биология (по направлениям); 1-31 01 02 Биохимия; 1-31 01 03 Микробиология; 1-33 01 01 Биоэкология, которая доступна по адресу:

http://elib.bsu.by/handle/123456789/134345

Учебно-методический комплекс по учебной дисциплине «Биометрия» для специальностей 1-31 01 01 Биология (по направлениям); 1-31 01 02 Биохимия; 1-31 01 03 Микробиология; 1-33 01 01 Биоэкология доступен по адресу:

http://elib.bsu.by/handle/123456789/16214