Учреждение образования
«Гомельский государственный университет
имени Франциска Скорины»

И. В. КУРАЧЕНКО, С. А. ЗЯТЬКОВ, Г. Г. ГОНЧАРЕНКО

БИОМЕТРИЯ

Практическое руководство

для студентов специальности 1-31 01 01-02
«Биология (научно-педагогическая деятельность)»

Гомель
ГГУ им. Ф. Скорины
2022

УДК 57.087.1(076)
ББК 28с51я73
К93

Рецензенты:

доктор биологических наук В. В. Потенко,
кандидат биологических наук Д. И. Каган

Рекомендовано к изданию научно-методическим советом
учреждения образования «Гомельский государственный
университет имени Франциска Скорины»

Кураченко, И. В.

К93 Биометрия : практическое руководство / И. В. Кураченко, С. А. Зятьков, Г. Г. Гончаренко ; Гомельский гос. ун-т им. Ф. Скорины. – Гомель : ГГУ им. Ф. Скорины, 2022. – 47 с.
ISBN 978-985-577-888-3

Практическое руководство посвящено описанию статистических методов, которые наиболее широко используются для обработки экспериментальных данных в области биологии. Издание содержит темы лабораторных занятий курса, приложения и список литературы. Может быть использовано как на лабораторных занятиях по соответствующим темам курса «Биометрия», так и для самостоятельной подготовки.

Адресовано студентам биологического факультета специальности 1-31 01 01-02 «Биология (научно-педагогическая деятельность)», а также аспирантам и научным сотрудникам, работающим в области биологии и медицины.

УДК 57.087.1(076)
ББК 28с51я73

ISBN 978-985-577-888-3

ОГЛАВЛЕНИЕ

ВВЕДЕНИЕ

Важным показателем научно-исследовательской работы является её эффективность. Высокие значения могут быть достигнуты в тех случаях, когда результативность работы способствует получению новой информации. Эффективность проводимых исследований во многом определяется качеством планирования, постановки и проведения эксперимента, а также глубинной анализа полученных результатов. В проведении исследований немаловажную роль играют все перечисленные этапы в равной степени.

В настоящее время уровень математической подготовки биолога определяется умением применять различные методы математической статистики. Однако одни и те же методы математической статистики, применяемые к исходным материалам, дают совершенно неодинаковые значения, отличающиеся по уровню достоверности и зависящие от личности «обработчика», что является недопустимым.

Принципиальное значение в качестве анализа материалов имеют ошибки, выявить которые очень сложно. К такому типу ошибок можно отнести неправильное определение достоверности отдельных показателей из числа представляемых в таблицах, графиках, доверительных интервалах и т. д.

Информативность результатов экспериментальных исследований в немалой степени определяется глубиной статистической обработки и качеством их представления. Применение статистических приемов позволяет исследователю выявить недостатки в постановке опытов и интерпретации результатов ещё до их публикации.

Математическая обработка данных, полученных в ходе биологического исследования, – важное условие при выполнении курсовых и дипломных работ. Для того чтобы сделать достоверное предположение об изучаемом явлении, используют только части – выборки, так как всю генеральную совокупность обследовать практически невозможно (тысячи, сотни тысяч, миллионы испытуемых). Выборка должна отражать все свойства генеральной совокупности, т. е. должна быть репрезентативной. Основные способы достижения этого условия – случайная выборка и моделирование выборки по свойствам генеральной совокупности.

Существенным при организации выборки является вопрос о необходимом и достаточном объеме. Малое количество ($n < 30$) вариант не обеспечит точности результатов. Большое количество ($n > 100$) приведет к увеличению трудоемкости (времени и стоимости) исследования.

Статистическая достоверность, или статистическая значимость, результатов исследования определяется при помощи методов статистического вывода, который предъявляет определенные требования к численности, или объему выборки.

Наибольший объем выборки необходим при разработке диагностической методики – от 200 до 1 000–2 500 вариант. Если необходимо сравнивать две выборки, их общая численность должна быть не менее 50 вариант; численность сравниваемых выборок должна быть приблизительно одинаковой. Если изучается взаимосвязь между какими-либо свойствами, то объем выборки должен быть не меньше 30–35 вариант. Чем больше изменчивость изучаемого свойства, тем больше должен быть объем выборки. Поэтому изменчивость можно уменьшить, увеличивая однородность выборки, например, по полу, возрасту и т. д.

Руководство содержит темы лабораторных занятий курса, приложение и список литературы, может быть использовано как на лабораторных занятиях по соответствующим темам курса «Биометрия», так и для самостоятельной подготовки.

Адресовано студентам биологического факультета, а также аспирантам и научным сотрудникам, работающим в области биологии и медицины.

ЛАБОРАТОРНАЯ РАБОТА 1
ПЕРВИЧНАЯ И ВТОРИЧНАЯ ГРУППИРОВКА ЭКСПЕРИМЕНТАЛЬНЫХ ДАННЫХ

Материалы и оборудование: данные замеров статистических величин; миллиметровка; калькуляторы.

Цель: образование выборочных, сгруппированных статистических совокупностей и их графическое отображение.

Пояснения к заданиям

Статистической совокупностью называют некоторое множество относительно однородных предметов или объектов, объединяемых по выбранному признаку. Теоретически бесконечно большую или приближающуюся к бесконечности совокупность всех единиц или членов вариационного ряда называют *генеральной.* Генеральная совокупность может состоять из такого большого количества единиц, что изучить их все не представляется возможным. Поэтому приходится иметь дело со сравнительно небольшими, *выборочными* совокупностями. Выборочная совокупность, наиболее полно отражающая свойства генеральной, называется *репрезентативной.* Например, при изучении морфометрических показателей рептилий исключаются рептилии с анатомированными хвостами; при изучении роста деревьев в высоту исключаются деревья, сломанные бурей, поврежденные огнем и т. д. При образовании выборки используется *метод случайного отбора,* то есть выдерживается принцип объективности.

Статистическая совокупность подвергается ранжированию, которое заключается в следующем:

а) находят минимальные и максимальные варианты (лимиты);

б) определяют вариационный размах: $\rho = x_{\max} - x_{\min}$;

в) при $\rho \leq 11$ проводят первичную группировку, т. е. $i = 1$;

г) при $\rho \geq 11$ весь диапазон значений признака разбивают на «классовые промежутки» и величину интервала определяют по формуле:

$$i = \frac{x_{\max} - x_{\min}}{1 + 3{,}32 \log n};$$

д) находят нижнюю границу первого интервала: $1 = x_{\max} - \frac{i}{2}$;

е) производят ранжирование данных и заполняют рабочую таблицу.

Ход работы

Задание 1. Провести группировку данных по качественным признакам по следующей схеме: объект, предмет, вариация, объём совокупности, число классов.

Задача. На звероводческой ферме выращивают норок: стандартных коричневых – 120 особей, сапфировых – 180 особей, серебристо-голубых – 160 особей, черных – 40 особей. Определить долю особей каждого из окрасов, изобразить диаграмму распределения норок по окрасу. Сделать обоснованный вывод.

Задание 2. Провести первичную группировку по экспериментальным данным о длине левого уха (в см) у 84 кроликов-мериносов.

12	11	13	14	10	13	13	13	14	12	12	12	14	13	13	14	13	14	12	15	12
11	13	10	12	13	12	11	12	14	11	10	15	12	11	11	13	13	12	15	11	12
13	11	12	12	14	16	12	14	12	11	14	12	14	11	13	12	14	11	14	12	14
11	10	16	11	12	12	12	13	14	12	12	12	12	13	13	11	12	13	14	11	12

Данные представить в виде рабочей таблицы (таблица 1).

Таблица 1 – Первичная группировка кроликов-мериносов

Классы, x_i	Разноска	Частота, f_{xi}

Составить безынтервальный вариационный ряд, найти моду и медиану. Вариационный ряд отобразить графически (на оси OX отметить значения классов, на оси OY – значения частот) и сделать обоснованный вывод.

Задание 3. Провести вторичную группировку по экспериментальным данным о длине тела у 80 экземпляров густеры озера Швакшта (в мм).

143	143	128	130	143	127	143	157	120	119	94	145	138	118	134
95	148	144	120	140	140	120	138	142	153	130	138	153	135	124
130	148	150	138	130	137	135	134	135	136	142	124	114	142	139
111	133	165	164	127	126	145	126	145	125	132	134	172	139	137
138	137	137	133	151	139	139	117	141	131	100	107	140	129	132
125	120	142	158	141										

Данные представить в виде рабочей таблицы (таблица 2).

Таблица 2 – Вторичная группировка густеры

Границы классов	Срединные значения классов, x_i	Разноска	Частота, f_{xi}

Составить интервальный вариационный ряд. Найти моду и медиану. Вариационный ряд отобразить графически (на оси ОХ отметить нижние значения классов, на оси OY – значения частот) и сделать обоснованный вывод.

Задание 4. Изучен живой вес 70 телят ярославских помесей при рождении (в кг):

27 32 32 31 32 28 37 35 26 28 32 28 35 36 28 39 43 28 33 36 34
26 32 33 36 30 35 36 28 37 43 32 32 23 26 26 36 28 27 35 37 34
40 32 33 32 35 32 28 26 37 27 31 35 37 31 29 30 26 29 29 31 32
35 41 40 31 36 29 33

Составить вариационный ряд. Найти моду и медиану.

Вопросы для самоконтроля

1. Что такое выборочная и сгруппированная статистическая совокупность?
2. Как образуется классовый интервал?
3. Что такое кумулята, огива? Как они отображаются графически?
4. Как построить полигон распределения статистических частот, гистограмму, диаграмму?
5. Что такое репрезентативность?
6. Перечислите типы отбора выборок.
7. Сформулируйте закон больших чисел.
8. В каких случаях проводится вторичная группировка?
9. Назовите закономерности распределения вариант в вариационном ряду.

ЛАБОРАТОРНАЯ РАБОТА 2
ЗАКОНОМЕРНОСТИ РАСПРЕДЕЛЕНИЙ

Материалы и оборудование: мешочки с семенами кукурузы (бобовых), калькуляторы.

Цель: экспериментальное доказательство теорем сложения и умножения вероятностей.

Пояснения к заданиям

Для того чтобы выяснить, произойдет или не произойдет событие при заданном комплексе факторов, нужно осуществить этот комплекс, т. е. провести испытание. *Испытанием* является любой эксперимент, в результате которого производят наблюдения. События, происходящие при одном и том же комплексе факторов, называются *однородными*. Установлено, что однородные случайные события в большой их массе подчиняются некоторым закономерностям. Эти закономерности получили название *вероятностных*. События с одинаковыми возможностями осуществления называются *равновозможными*. Числовая характеристика случайного события, при которой в любой достаточно большой серии испытаний частота события лишь незначительно отличается от этой характеристики, называется *вероятностью события*. Исходы испытания являются простейшими случайными событиями. Вероятностью случайного события называется отношение числа отходов, благоприятствующих событию, к числу всех возможных исходов.

Если некоторое событие может произойти при n испытаниях и числе исходов, которые благоприятствуют (a) или не благоприятствуют (b) наступлению события, то вероятность того, что событие произойдет, может быть определена как $p = \frac{a}{n}$. Вероятность того, что событие не произойдет, будет выглядеть, как $q = \frac{b}{n}$, $p + q = 1$.

Основные теоремы теории вероятностей.

1. Вероятность суммы двух несовместных, независимых событий равна сумме их вероятностей

$$P\,(A + B) = P\,(A) + P\,(B).$$

2. Вероятность сложного события (т. е. наступления двух событий независимых одно от другого) равна произведению вероятностей отдельных событий

$$P\,(A*B) = P(A) * P(B).$$

3. Вероятности отдельных возможных исходов даются последовательными членами разложения Бинома Ньютона, например:

$$(p + q)^3 = p^3 + 3p^2q + 3pq^2 + q^3.$$

Ход работы

Задание 1. Для доказательства теоремы сложения двух независимых случайных событий провести 20 извлечений по одному семени из мешочка, в котором находится по 10 зеленых, белых и желтых семян кукурузы. Результаты испытаний занести в таблицу 3.

Таблица 2 – Вторичная группировка густеры

События	Разноска	Число случаев
Желтое		m_1
Зеленое		m_2
Белое		m_3

Задание 2. Определить эмпирические вероятности данных событий. Найти вероятность того, что вынутое наугад семя окажется окрашенным.

Задание 3. Определить теоретические вероятности данных событий и найти отклонение $P_{\text{эмп}} - P_{\text{теор}}$.

Задание 4. Решить упражнения:

1. В урне m белых и n черных шаров. Какова вероятность того, что вынутый наугад шар окажется: а) белым, б) черным?
2. Стрелок, сделав 200 выстрелов, попал в цель 190 раз. Какова вероятность попадания в цель? Сколько будут попаданий в цель, если стрелок сделает 300 выстрелов?
3. Какова вероятность выпадения суммы цифр при бросании двух кубиков, равной 7?
4. Какова вероятность выпадения «решкой» одной монеты (двух, трех, не менее одной) при бросании трех монет?
5. Какова вероятность выпадения суммы цифр, равная семи, при бросании трех кубиков?
6. В городе N с населением в 100 000 жителей родилось 8 000 новорожденных. Какова вероятность рождения детей: абсолютная, удельная?
7. В корзине три красных и семь зелёных яблок. Из корзины вынимают одно яблоко. Найти вероятность того, что оно будет красным.
8. В мешочке пять букв: М, О, Л, О, Т. Какова вероятность того, что доставая по одной карточке наугад, получится слово «молот», «том»?
9. В колоде 36 карт. Какова вероятность того, что вынутая карта окажется тузом?

10. В лотерее 4 выигрышных и 96 безвыигрышных билета. Какова вероятность, что два билета окажутся выигрышными?

11. В пассажирском поезде 12 вагонов. Какова вероятность того, что двое друзей независимо друг от друга окажутся в одном вагоне?

12. На полке в случайном порядке расставлено 40 книг, среди которых находится трехтомник Пушкина. Найти вероятность того, что эти тома стоят в порядке возрастания слева направо, но не обязательно рядом.

13. На контроль качества медицинских препаратов поступила партия из 6 штук. Вероятность того, что препарат окажется некачественным, равна 0,25.

Найти вероятности Pn(k) того, что число некачественных препаратов k в партии составляет 0, 1, 2, 3, 4, 5, 6.

Построить ломаную линию с вершинами в точках Pn(k).

Найти наивероятнейшее число некачественных препаратов.

Задание 5. С помощью треугольника Паскаля и бинома Ньютона получить вероятности распределения различных комбинаций детей по полу в семьях, имеющих четырех детей.

Вопросы для самоконтроля

1. Что такое вероятность?
2. В чем сущность теоремы сложения вероятностей? Приведите примеры.
3. В чем сущность теоремы умножения вероятностей? Приведите примеры.
4. В чем сущность биномиального распределения?
5. Приведите примеры достоверных, независимых, несовместимых событий.

ЛАБОРАТОРНАЯ РАБОТА 3
СРЕДНИЕ ВЕЛИЧИНЫ УРОВНЯ РЯДА

Материалы и оборудование: данные статистических величин и расчетов по лабораторным работам 1, 2; калькуляторы.

Цель: определение статистических показателей выборочной совокупности.

Пояснения к заданиям

Наиболее распространенным и используемым показателем является *среднеарифметическая величина*. Она рассчитывается по формуле:

$$\overline{X} = \frac{(x_1 + x_2 + ... + x_n)}{n} = \frac{\sum x_i}{n},$$

где n – общее число вариант;

$\sum$ – знак суммирования;

x_i – значения вариант.

Средняя квадратическая величина используется при вычислении средней из величин объема, запаса, площади. Рассчитывается по формуле:

$$\overline{X}_q = \frac{\sqrt{\left(\sum x_i^{\,2}\right)}}{n},$$

где $x_i^{\,2}$ – квадраты замеряемых величин – объем, площадь и т. п.;

n – общее число вариант;

x_i – значения вариант.

Средняя геометрическая M_g (или $\overline{X_g}$) используется для расчета среднего темпа роста изучаемого признака. Она известна также как средняя логарифмическая, так как ее логарифм есть арифметическая средняя логарифмов составляющих величин.

Вычисляется по формуле

$$\overline{X_g} = \sqrt{x_{\mathrm{i}} \cdot x_2 \cdot x_{3...} ... x_n},$$

где $x_{\mathrm{i}} \cdot x_2 \cdot x_{3...} ... x_n$ – темпы роста (величины, показывающие, во сколько раз увеличивался признак от периода к периоду);

n – число периодов.

При $n > 2$ формулу удобнее применять в логарифмическом виде:

$$\ln \overline{X_g} = \frac{1}{n}(\ln x_1 + \ln x_2 + ... + \ln x_n) = \left(\frac{1}{n}\right) \cdot (\sum \ln x_i).$$

Откуда $\overline{X}_g = e^{\ln}$, где e – основание натуральных логарифмов, равно 2,72.

Средняя гармоническая используется для вычисления средней величины отношений двух варьирующих величин. Определяется по формуле:

$$\overline{X_h} = \frac{n}{\sum \frac{1}{x_i}},$$

где n – число значений;

x_i – значения соотношений величин.

Выборочная совокупность довольно точно воспроизводит свойства и соотношения в генеральной совокупности, но не абсолютно точно вследствие вариации изучаемых признаков. Поэтому между статистическими показателями выборочной совокупности и действительными значениями этих показателей генеральной совокупности всегда будут некоторые расхождения, которые являются случайными ошибками выборки (иначе – случайными ошибками репрезентативности) и называются *основными ошибками* того или иного статистического показателя. *Ошибка среднего значения* (или ошибка стандартная, выборочности, статистическая) вычисляется по формулам:

$$m_x = \frac{\sigma}{\sqrt{n}},$$

$$m_x = \frac{\sigma}{\sqrt{n-1}}$$

при малых объемах ($n \leq 30$).

Ход работы

Задание 1. Рассчитать средние величины для данных о весе 11 поросят при рождении: 1,0 1,2 1,5 1,3 1,3 1,4 1,4 1,1 1,0 0,9 2,5. Вариационный ряд не составлять, произвести расчеты прямым способом.

Задание 2. Рассчитать средние величины для данных о длине левого уха (в см) у 70 кроликов-мериносов.

12 11 13 14 10 13 13 13 14 12 12 12 14 13 13 14 13 14 12 15
12 11 13 10 12 13 12 11 12 14 11 10 15 12 11 11 13 13 12 15
11 12 13 11 12 12 14 16 12 14 12 11 14 12 14 11 13 12 14 11
14 12 14 11 10 16 11 12 12 12.

Проведя первичную группировку, рассчитать средние величины следующими способами:

а) прямым способом через значения вариант по таблице 4.

Таблица 4 – Расчет средних величин через значения вариант

x_i	Разноска	f	$f\,x_i$	$f\,x_i^{\,2}$
			$\sum f\,x_i =$	$\sum f\,x_i^{\,2} =$

б) прямым способом через центральные отклонения по таблице 5.

Таблица 5 – Расчет средних величин через центральные отклонения

x_i	Разноска	f	$f\,x_i$	$x_i - X$	$f\,(x_i - X)$	$f\,(x_i - X)^2$
			$\sum f\,x_i =$			$\sum f\,(x_i - X)^2 =$

в) способом условной средней ($A = Mo$, тогда $a = x_i - A$) по таблице 6.

Таблица 6 – Расчет средних величин способом условной средней

x_i	Разноска	f	a	fa	fa^2
				$\sum fa =$	$\sum fa^2 =$

Задание 3. Изучен живой вес 70 телят ярославских помесей при рождении (в кг).

27 32 32 31 32 28 37 35 26 28 32 28 35 36 28 39 43 28 33 36
34 26 32 33 36 30 35 36 28 37 43 32 32 23 26 26 36 28 27 35
37 34 40 32 33 32 35 32 28 26 37 27 31 35 37 31 29 30 26 29
29 31 32 35 41 40 31 36 29 33

Данные задания 3 представить в виде рабочей таблицы и произвести расчеты способом условной средней (таблица 7)

Таблица 7 – Расчет средних величин способом условной средней при вторичной группировке

Границы классов	Срединные значения классов, x_i	x_i	Разноска	f	α	$f\,\alpha$	$f\,\alpha^2$
						$\sum f\,\alpha =$	$\sum f\,\alpha^2 =$

Задание 4. Рассчитать статистические показатели и ошибку выборочности по экспериментальным данным о длине тела у 85 экземпляров густеры озера Швакшта (в мм).

143	143	128	130	143	127	143	157	120	119	94	145	138	118	134
95	148	144	120	140	140	120	138	142	153	130	138	153	135	124
130	148	150	138	130	137	135	134	135	136	142	124	114	142	139
111	133	165	164	127	126	145	126	145	125	132	134	172	139	137
138	137	137	133	151	139	139	117	141	131	100	107	140	129	132
125	120	142	158	141	124	154	154	139	117					

Данные задания 4 представить в виде рабочей таблицы и произвести расчеты способом условной средней представить в виде рабочей таблицы (таблица 7).

Вопросы для самоконтроля

1. Перечислите статистические показатели, относящиеся к группе величин, характеризующих центральную тенденцию вариационного ряда.
2. На какие группы делятся статистические показатели?
3. Приведите доказательства свойств средней арифметической.
4. В каких исследованиях применяется средняя гармоническая?
5. В каких исследованиях применяется средние квадратическая и геометрическая?

ЛАБОРАТОРНАЯ РАБОТА 4
СРЕДНИЕ ВЕЛИЧИНЫ ВАРИАБЕЛЬНОСТИ ПРИЗНАКА

Материалы и оборудование: данные статистических величин и расчетов по лабораторным работам 1, 2; калькуляторы.

Цель: Определение статистических показателей выборочной совокупности.

Пояснения к заданиям

Рассеяние вариант выборки относительно средней характеризуется:
– центральным отклонением;
– дисперсией;
– среднеквадратическим отклонением;
– коэффициентом вариации.

Их вычисляют по формулам:
– центральное отклонение

$$\alpha = x_i - \bar{x};$$

– дисперсия

$$\sigma^2 = \frac{1}{N-1}\sum_{i=1}^{k}(x_i - \bar{x})^2;$$

– среднеквадратическое отклонение

$$\sigma^2 = \sqrt{\frac{1}{N-1}\sum_{i=1}^{k}(x_i - \bar{x})^2};$$

– коэффициент вариации

$$C_\chi = \frac{\sigma_\chi \cdot 100}{\bar{\chi}}.$$

При C_χ < 30 % выборка имеет большую степень концентрации вариант возле величины. При 30 % ≤ C_χ ≤ 100 % степень концентрации допустимая. При C_χ > 100 % делается вывод о неоднородности выборки.

Ход работы

Задание 1. Рассчитать средние величины для данных о весе 10 поросят при рождении: 1,1 1,3 1,2 1,3 1,5 1,4 1,4 1,1 1,0 1,5. Определить показатели вариации признака.

Задание 2. Рассчитать показатели вариации признака для данных о длине хвоста (в см) у 50 валахских овец.

12 11 13 14 10 13 13 13 14 12 12 12 14 13 13 14 13 14 12 13
11 12 14 11 10 15 12 11 11 13 13 12 15 11 12 12 14 16 12 14
11 14 12 14 11 13 12 14 11 14

Проведя первичную группировку, рассчитать показатели вариации признака прямым способом через значения вариант по таблице 4.

Задание 3. Изучена высота 40 растений пшеницы (в см).

28 32 28 35 36 28 39 43 28 33 36 34 26 27 35 37 34 40 32 33
35 32 27 31 35 37 31 29 30 26 29 29 31 32 35 41 40 31 36 29

Данные представить в виде рабочей таблицы и произвести расчеты способом условной средней (таблица 7).

Задание 4. Рассчитать статистические показатели вариации признака по экспериментальным данным о высоте ушной раковины лисицы фенека (в мм).

141	143	128	130	143	127	143	157	120	119	94	145	138	118	134
135	148	144	120	140	140	120	138	142	153	130	138	153	135	124
130	148	150	138	130	137	135	134	135	136	142	124	114	142	139
111	133	165	164	127	126	145	126	145	125	132	134	172	139	137
138	137	137	133	151	139	139	117	141	131	100	107	140	129	132

Данные представить в виде рабочей таблицы и произвести расчеты способом условной средней (таблица 7).

Вопросы для самоконтроля

1. Перечислите средние величины, характеризующие вариабельность признака.
2. Укажите свойства среднего квадратического отклонения.
3. В каких случаях применяется коэффициент вариации?
4. Назовите форму расчеты дисперсии.
5. Какой вывод можно сделать при значении коэффициента вариации, равному 35 %?

ЛАБОРАТОРНАЯ РАБОТА 5
СТАТИСТИЧЕСКИЙ АНАЛИЗ ЧАСТОТ РАСПРЕДЕЛЕНИЙ

Материал и оборудование: таблицы, калькуляторы.

Цель: определение нормальности распределения данных с помощью критерия χ^2.

Пояснения к заданиям

При сравнении наблюдаемых и ожидаемых результатов применяются особые критерии оценки, в частности критерий хи-квадрат (χ^2). Критерий предложен Карлом Пирсоном и представляет собой сумму отношений между квадратами разностей эмпирических и вычисленных или ожидаемых частот к ожидаемым частотам:

$$\chi^2 = \sum \frac{(p - p')^2}{p'},$$

где $\sum$ – знак суммирования,

p – эмпирическая частота,

p' – ожидаемая или теоретически вычисленная частота. Использование χ^2-теста необходимо для того, чтобы узнать, подтверждается ли гипотеза экспериментом, т. е. насколько верны условия эксперимента, позволяют ли они с высокой степенью достоверности подтвердить или опровергнуть исходное предположение. Если бы фактические данные полностью совпадали с теоретическими, значение критерия было бы равно нулю. По мере увеличения разницы между этими показателями значение критерия будет возрастать. Каждому значению χ^2 соответствует определенная вероятность его появления (таблица 8).

Таблица 8 – Критические значения χ^2 для трех степеней доверительной вероятности

Число степеней свободы, U	Уровень значимости			Число степеней свободы, U	Уровень значимости		
	0,95	0,99	0,999		0,95	0,99	0,999
1	2	3	4	5	6	7	8
1	3,8	6,6	10,8	26	38,9	45,6	54,1
2	6,0	9,2	13,8	27	40,1	47,0	55,5
3	7,8	11,3	16,3	28	41,3	48,3	56,9
4	9,5	13,3	18,5	29	42,6	49,6	58,3
5	11,1	15,1	20,5	30	43,8	50,9	59,7
6	12,6	16,8	22,5	32	46,2	53,5	62,4
7	14,1	18,5	24,3	34	48,6	56,0	65,2
8	15,5	20,1	26,1	36	51,0	58,6	67,9
9	16,9	21,7	27,9	38	53,4	61,1	70,7
10	18,3	23,2	29,6	40	55,8	63,7	73,4
11	19,7	24,7	31,3	42	58,1	66,2	76,1
12	21,0	26,2	32,9	44	60,5	68,7	78,7
13	22,4	27,7	34,5	46	62,8	71,2	81,4
14	23,7	29,1	36,1	48	65,2	73,7	84,0
15	25,0	30,6	37,7	50	67,5	76,2	86,7
16	26,3	32,0	39,3	55	73,3	82,3	93,2
17	27,6	33,4	40,8	60	79,1	88,4	99,6
18	28,9	34,8	42,3	65	89,8	94,4	106,0
19	30,1	36,2	43,8	70	90,5	100,4	112,3
20	31,4	37,6	45,3	75	96,2	106,4	118,5
21	32,7	38,9	46,8	80	101,9	112,3	124,8

Окончание таблицы 8

1	2	3	4	5	6	7	8
22	33,9	40,3	48,3	85	107,5	118,2	131,0
23	35,2	41,6	49,7	90	113,1	124,1	137,1
24	36,4	43,0	51,2	95	118,7	130,0	143,3
25	37,7	44,3	52,5	100	124,3	135,8	149,4

Значение χ^2 в таблице указывают те границы, до которых полученные значения критерия не дают оснований сомневаться в высказанном предположении с определенной степенью вероятности, Значения χ^2, превышающие табличные, будут указывать на несостоятельность гипотезы, т. е. признается тот факт, что различие между фактическими и теоретически ожидаемыми результатами является достоверным, значимым (Приложение Г).

Ход работы

Задание. Для предложенных задач произвести расчет критериев и оценить их величину:

– достоверности;

– Фишера;

– хи-квадрат.

Задача 1. Из 100 вакцинированных заболело 8 человек. Определить уровень заболеваемости в исследуемой группе людей и сравнить с теоретической 0,12. Провести анализ следующим образом: составить диаграмму заболеваемости, вычислить статистические характеристики (p, δ_p, m_p, P, C_v), дать оценку достоверности (доверительный интервал при трех уровнях значимости; H_0; критерий Стьюдента). Сделать обоснованный вывод.

Задача 2. Получены данные о распределении бычков и телочек в совхозе «Восток» за 2002 год: бычков – 1 857, телочек – 1 256. Соответствует ли распределение бычков и телочек соотношению 1:1. Провести анализ следующим образом: составить диаграмму заболеваемости, вычислить статистические характеристики (p, δ_p, m_p, P, C_v), дать оценку достоверности (доверительный интервал при трех уровнях значимости; H_0; критерий Стьюдента). Сделать обоснованный вывод.

Задача 3. Получены данные о распределении самок и самцов плодовой мушки: самок – 126, самцов – 250. Соответствует ли данное распределение соотношению 1:1. Провести анализ следующим образом:

составить диаграмму заболеваемости, вычислить статистические характеристики (p, $\delta_р$, $m_р$, P, C_v), дать оценку достоверности (доверительный интервал при трех уровнях значимости; H_0; критерий Стьюдента). Сделать обоснованный вывод.

Вопросы для самоконтроля

1. В каких случаях применяется попарный метод сравнения данных?
2. Критерий Пирсона как сумма отношений между эмпирическим и наблюдаемым явлениями.
3. Что является критерием оценки при анализе частот?
4. Особенности проведения статистического анализа частей распределений.
5. Как определить нормальность распределения данных с помощью χ^2?

ЛАБОРАТОРНАЯ РАБОТА 6
ПОКАЗАТЕЛИ ВАРИАЦИИ

Материалы и оборудование: таблицы; калькуляторы.

Цель: Оценить асимметрию и эксцесс, подтвердить правило трех сигм.

Пояснения к заданиям

Отклонение распределений фактических данных от нормального типа характеризуется основными моментами – r_3, r_4, которые показывают асимметричность (коэффициент асимметрии – A) и крутость (коэффициент эксцессов распределений – E):

$$A = r_3 = \frac{\mu_3}{\left(\sqrt{\mu_2}\right)^3}; \qquad E = r_4 - 3 = \frac{\mu_4}{\left(\mu_2\right)^2} - 3$$

или

$$A = \frac{\sum f (x_i - X)^3}{n \cdot \sigma^3}; \qquad E = \frac{\sum f (x_i - X)^4}{n \cdot \sigma^4} - 3.$$

Ошибка показателя асимметрии проводится по формуле:

$$m_A = \sqrt{\frac{6}{N}}.$$

Ошибка показателя эксцесса равна удвоенной ошибке показателя асимметрии. Оценку достоверности асимметрии и эксцесса проводят по формулам:

$$t = \frac{A}{m_a} \geq 3 \text{ и } t = \frac{E}{m_e} \geq 3.$$

Ход работы

Задание 1. Произвести расчет средних величин, нормированных отклонений по данным о длине правого уха (в см) у 60 серебристо-черных лисиц.

12	10	14	14	13	12	12	12	15	13	11	12	12	14	12
12	13	14	11	13	14	12	13	12	12	14	12	14	13	13
12	13	12	12	13	12	11	11	12	13	14	12	14	12	14
10	11	10	11	15	11	16	11	16	11	11	11	12	15	14

Применить прямой способ через центральные отклонения по таблице 9.

Таблица 9 – Расчет нормированных отклонений для доказательства правила «трех сигм»

x_i	Раз-носка	f	$f x_i$	$x_i - X$	$f(x_i - X)$	$f(x_i - X)^2$	$f(x_i - X)^3$	$f(x_i - X)^4$	t	$\frac{f}{n}$, %
			$\sum f x_i =$			$\sum =$	$\sum =$	$\sum =$		

Задание 2. Построить полигон распределения, отметив моду, медиану и среднее арифметическое на графике. Сделать вывод о характере распределения изучаемого признака.

Задание 3. Для доказательства правила трех сигм построить кривую нормального распределения, отметив на оси абсцисс – нормированные отклонения, на оси ординат относительные частоты в процентах.

Задание 4. Рассчитать коэффициенты асимметрии и эксцесса, дав им оценку достоверности.

Задание 5. Определить долю вариант под кривой нормального распределения, используя Приложения А, Б, в пределах: а) от 0 до +2,15σ;

б) от –1,09σ до –0,08σ; в) за пределами ±2,31σ; г) от *X* до 2,34σ; д) между ±1,98σ.

Вопросы для самоконтроля

1. Как осуществить расчет нормального распределения?
2. Как определяется тип кривой при помощи критерия Пирсона?
3. Для каких целей применяются статистические критерии?
4. Дайте определение *ошибки выборочности*?
5. Как определяются основные ошибки статистических показателей?

ЛАБОРАТОРНАЯ РАБОТА 7 РЕПРЕЗЕНТАТИВНОСТЬ ВЫБОРОЧНЫХ ПОКАЗАТЕЛЕЙ

Материал и оборудование: данные статистических величин и расчетов по лабораторным работам 5, 6; таблицы; калькуляторы.

Цель: оценить степень достоверности статистических показателей.

Пояснения к заданиям

Точность опыта, или *процент ошибки наблюдения* – это процент расхождения между генеральной и выборочной средней, который вычисляется по формулам:

$$p = \frac{100 \cdot m_x}{\overline{X}}$$

или

$$p = \frac{C}{\sqrt{n}}.$$

Точность опыта показывает, насколько процентов можно ошибиться, если утверждать, что генеральная средняя равна полученной выборочной средней.

Полученный процент ошибки сопоставляется с заданным: если он не больше заданного, точность достаточная, а если больше, то точность результата является неудовлетворительной; значит, следует увеличить число наблюдений.

После вычисления того или иного статистического показателя необходимо проверить степень его надежности или достоверности путем деления величины данного показателя на величину его основной ошибки:

$$t = \frac{X}{m_x},$$

где X – величина любого статистического показателя;

m_x – величина ошибки любого статистического показателя.

Если частное t получится равным или больше трех, то значение показателя является надежным, достоверным, и им можно пользоваться для разных сопоставлений и выводов. Если же это отношение будет меньше трех, то данный показатель оказывается ненадежным, величина его недостоверна и является лишь в той или иной мере вероятной. Такие показатели нельзя сопоставлять между собой или производить на основе их заключения. Нередко приходится решать вопрос, насколько существенно различие в значениях показателей какого-либо признака, вычисленных для разных совокупностей. С этой целью находится основная ошибка разницы чисел и доказывается ее достоверность по вышеописанному принципу. Ошибка разности вычисляется как квадратный корень из суммы квадратов основных ошибок исследуемого показателя, то есть

$$m_s = \sqrt{m_1^2 + m_2^2}.$$

Полученную разность показателей делят на его ошибку. Находится показатель существенности различия средних значений

$$t = \frac{\left|\overline{X}_1 - \overline{X}_2\right|}{m_s} = \frac{\left|\overline{X}_1 - \overline{X}_2\right|}{\sqrt{m_1^2 + m_2^2}}.$$

Если этот показатель получится больше трех, то различие существенно, доказано, и данное мероприятие вызвало существенное изменение; обе сравниваемые выборочные совокупности являются представителями качественно разных генеральных совокупностей. Если же он получится меньше трех, то можно утверждать, что расхождение оказалось случайным, недостоверным и во всяком случае целесообразность данного мероприятия осталась недоказанной.

Расчет доверительных интервалов статистических показателей. Среднее значение, основное (квадратическое) отклонение, коэффициент изменчивости, асимметрия, эксцесс дают представление о величине и форме распределений наблюдений. Однако они не дают представления о возможных значениях случайной величины. Оно заключается в вычислении вероятности того, что значение величины будет заключаться в определенных границах. Считается, что границы достоверно определены если вероятность близка к единице, например, 0,99 или 0,999 (99,0% или 99,9%) (таблица 10 и Приложение В).

Таблица 10 – Значения показателя t (критерия Стьюдента)

Число степеней свободы	Доверительные уровни		
	95 %	99 %	99,9 %
9	2,3	3,2	4,8
10	2,2	3,2	4,6
11–14	2,2	3,0	4,3
15–20	2,1	2,9	3,9
21–30	2,1	2,8	3,7
31–60	2,0	2,7	3,5
61–120	2,0	2,6	3,4
∞	1,96	2,58	3,29

Соответствующие границы называются *доверительными*. В зависимости от типа распределения данных доверительный интервал рассчитывается двумя способами. В симметричных распределениях, близких к нормальному, размах отклонений данных от средней арифметической обычно равен приблизительно 3σ в обе стороны от значения средней (так называемый закон трех сигм).

При расчете доверительного интервала для трех стандартных доверительных уровней (95 %, 99 %, 99,9 %) *t* выбирается по числу степеней свободы из таблицы критерия Стьюдента.

Доверительный интервал статистического показателя, например, средней арифметической строится по формуле:

$$\overline{X} - t \cdot m_{\overline{x}} < \mu < \overline{X} + t \cdot m_{\overline{x}}.$$

Подставляя в формулу величины среднего арифметического, коэффициента вариации, коэффициента асимметрии, эксцесса и их ошибок, определяются доверительные интервалы для этих показателей.

Ход работы

Задание 1. Получены данные о количестве хвостовых щитков у змей.

42	58	44	54	41	50	46	46	54	48	43	49
50	48	46	46	45	53	48	48	53	53	48	41
46	40	50	43	49	51	52	46	42	44	48	45
47	46	43	50	47	45	48	40	44	42	48	45
54	50	56	48	45	45	51	42	44	47	46	45

Провести анализ следующим образом:

– вычислить статистические характеристики (Mo, Me, X, σ, m_x, P, C_v);

– дать оценку достоверности (доверительный интервал при трех уровнях значимости; H_0; критерий Стьюдента);

– сделайте обоснованный вывод.

Задание 2. Используя критерий Стьюдента разницы двух средних, ответить на вопрос: отличаются ли по температуре тела самцы и самки тушканчиков?

Самцы:	37,5	35,9	37,5	37,8	37,2	37,9	36,9
Самки:	37,5	35,4	37,1	37,9	37,0	37,7	37,9

Задание 3. Исследовалось влияние лекарственного препарата на величину некоторого параметра.

Х (опыт)	160	120	140	180	130	160	
У (контроль)	180	160	220	180	160	200	170

По критерию Стьюдента для уровня значимости *p* = 0,05 проверить, эффективен ли препарат.

Вопросы для самоконтроля

1. В чем сущность нулевой гипотезы?
2. Как определяется доверительный интервал при нормальном распределении статистических показателей?
3. Что такое уровень значимости?
4. Что такое доверительная вероятность?

ЛАБОРАТОРНАЯ РАБОТА 8
ОДНОФАКТОРНЫЙ ДИСПЕРСИОННЫЙ АНАЛИЗ

Материалы и оборудование: калькуляторы, таблицы.

Цель: Определение значимости различия признака биологического объекта методом однофакторного дисперсионного анализа.

Пояснения к заданиям

Дисперсионный, или *вариансный анализ* (analysis of variance) предполагает установление роли отдельных факторов в изменчивости того или иного признака, при котором общая дисперсия как количественных, так и качественных признаков раскладывается на отдельные составляющие. У изучаемых признаков в эксперименте имеется не одно, а несколько значений, которые называют *градациями* или *уровнями фактора А*. Число наблюдений (вариант) в каждой группе обозначается как «n». Схема обозначения членов вариационного ряда при однофакторном дисперсионном анализе представлена в таблице 11.

Таблица 11 – Схема однофакторного дисперсионного анализа

Число градаций фактора A	Повторности			n	$\sum x_i$ / $\sum x_i^2$	$\frac{\left(\sum x_i\right)^2}{n}$
	1	2	3			
Контроль	x_i / x_i^2					
Смесь 1						
Смесь 2						
Примечание: по данным таблицы рассчитать суммы $\sum\sum x_i$; $\sum\sum x_i^2$; $\frac{\left(\sum x_i\right)^2}{n}$						

Различают 3 типа варьирования:

а) σ_y^2 – общее варьирование вариант, независимо от того, в какой группе они находятся, вокруг общей средней $\bar{x}$;

б) σ_x^2 – варьирование средних каждого уровня данного изучаемого фактора вокруг общей средней $\bar{x}$;

в) σ_z^2 – варьирование вариант внутри каждой группы вокруг каждой групповой средней $\bar{x}_i$ (так называемая остаточная).

Между ними существует соотношение:

$$\sigma_y^2 = \sigma_x^2 + \sigma_z^2.$$

Для каждого типа варьирования вычисляются суммы квадратов отклонений по следующим формулам:

– общая сумма квадратов:

$$C_y = \sum\sum x_i^2 - \frac{\left(\sum\sum x_i\right)^2}{N};$$

– сумма квадратов для групповых средних (факторальная):

$$C_x = \sum\frac{\left(\sum x_i\right)^2}{n} - \frac{\left(\sum\sum x_i\right)^2}{N};$$

– сумма квадратов для межгрупповых средних (случайная):

$$C_z = \sum\sum x_i^2 - \sum\frac{\left(\sum x_i\right)^2}{n},$$

где $\sum x_i$ – для каждой группы (уровня фактора A);

n_i – число наблюдений в каждой группе;

N – общее число вариант;

r – число уровней (градаций) фактора.

При делении сумм квадратов, обозначаемых C_x и C_z, на число степеней свободы получаются *средние квадраты (вариансы)* – σ^2 непосредственно измеряющие суммарную вариацию.

Оценка дисперсии каждой из групп связана со степенью свободы, при этом необходимо учитывать большую вариансу (например, если $\sigma^2_x \geq \sigma^2_z$, то за $df_1 = df_x = r - 1$. $df_2 = df_z = N - r$. Далее проводится проверка гипотезы $H_0 : \overline{X}_1 = \overline{X}_2 = \ldots = \overline{X}_к$, т. е. утверждения, что все групповые средние не зависят от влияния фактора A. Если верна H_0, то межгрупповая дисперсия (в генеральной совокупности) должна быть равна внутригрупповой, т. е. $H_0 : \sigma^2_x = \sigma^2_z$. При этом вычисленное значение F меньше табличного при уровне значимости a. Следовательно, гипотезу об отсутствии влияния фактора A не отклоняют (Приложения Д, Е).

Ход работы

Задание 1. Провести дисперсионный анализ по предложенной в таблице 11 схеме. Изучен процент гемоглобина в крови кур разных пород. Влияет ли породность на процент гемоглобина:

Породы	Повторности
– итальянские	87 92 86 91
– куропатчатые	91 90 88 89
– минорки	85 82 85 86
– бентамы	82 82 85 84

Задание 2. Провести дисперсионный анализ по предложенной схеме. Задача. При кормлении тушканчиков получены данные о средних температурах тела. Влияет ли пол на изменчивость температуры тела?

Пол	Повторности
– самки	36,9 36,8 37,0 36,6
– самцы	36,7 36,7 36,8 36,6

Задание 3. Проводились опыты по удобрению карповых прудов негашёной известью (600 кг/га) и суперфосфатом (72,8 кг/га), а также их смесью. Четвертый пруд в каждом блоке не удобрялся. В четвертом пруду продуктивность составила: 58 84 39; при применении фосфатов – 72 72 64; при применении извести – 49 55 48; при смешивании – 74 74 85. Влияют ли известь (Са), фосфат (Р) и их смеси на продуктивность пруда.

Вопросы для самоконтроля

1. В чем состоит сущность метода дисперсионного анализа?
2. Как проводится оценка варьирования при дисперсионном анализе?
3. Что такое параметрические и непараметрические критерии? Как они вычисляются?
4. Для каких целей применяются статистические критерии?

ЛАБОРАТОРНАЯ РАБОТА 9
КОРРЕЛЯЦИОННЫЙ АНАЛИЗ

Материал и оборудование: таблицы; калькуляторы.

Цель: Определение простого коэффициента корреляции и оценки его достоверности.

Пояснения к заданиям

Коэффициент корреляции. Отличительной особенностью биологических объектов является многообразие признаков, характеризующих каждый из них. Часто наблюдается связь между вариациями по различным признакам. В простейшем случае связь между двумя переменными величинами строго однозначна. Например, вес образцов, сделанных из одного и того же материала, определяется их объемом. Такого рода зависимость принято называть *функциональной.* Для биологических объектов связь обычно бывает менее жесткой: объекты с одинаковым значением одного признака имеют, как правило, разные значения по другим признакам. Такую связь между вариациями разных признаков называют *корреляцией* (дословный перевод: соотношение) между признаками. Заполняется рабочая таблица (таблица 12).

Таблица 12 – Схема корреляционного анализа

X	Y	X^2	Y^2	$X \cdot Y$
Σ	Σ	Σ	Σ	Σ

Затем вычисляется коэффициент корреляции, его ошибка и достоверность.

При малых объемах выборок (до 30–50) коэффициент корреляции вычисляется по формуле

$$r = \frac{\dfrac{\sum x \cdot y - (\sum x \cdot \sum y)}{N}}{\sqrt{\left[\dfrac{\sum x^2 - (\sum x)^2}{N}\right] \cdot \left[\dfrac{\sum y^2 - (\sum y)^2}{N}\right]}}.$$

Достоверность коэффициента корреляции можно оценить по формуле

$$t = \left(0{,}5 \cdot \ln\left(\frac{1+r}{1-r}\right)\right) \cdot \sqrt{N-3},$$

где t – критерий Стьюдента при числе свободы $\nu = N - 2$.

Можно применить метод Z, определить критические значения r, представленные в Приложениях Ж, И.

По результатам расчетов делается вывод:

– связи между признаками: прямая ($r > 0$) или обратная ($r < 0$);

– тесноте связи:

$r = 1{,}0$ (близка к функциональной);

$r = 0{,}901 – 0{,}999$ (очень высокая);

$r = 0{,}701 – 0{,}900$ (высокая);

r = 0,501 – 0,700 (значительная);
r = 0,301 – 0,500 (слабая);
r = 0 – 0,300 (отсутствует).

Затем определяется достоверность вычисленных показателей и сравниваются их значения, полученные по способу смешанных моментов и по приведенной формуле для малой выборки.

Ход работы

Задание 1. У окуня озера Баторино измерены длина головы (x) и длина грудного плавника (y).

x	66	61	67	73	51	59	48	47	58	44	41	54	52	47	51	45
y	38	31	36	43	29	33	28	25	36	26	21	30	28	27	28	26

Вычислите коэффициент корреляции и определить его достоверность.

Задание 2. У 15 серебристо-черных лисиц (совхоз «Белорусский») были измерены (в см) длина туловища (x) и длина хвоста (y).

x	70	65	66	65	71	68	64	57	66	65	67	62	67	62	63
y	40	40	40	40	40	42	39	38	41	43	39	45	43	38	40

Вычислите коэффициент корреляции и определите его достоверность.

Вопросы для самоконтроля

1. Что такое коэффициент корреляции? Как он вычисляется?
2. Что такое метод Z?
3. В чем преимущество числа z перед коэффициентом корреляции r Можно ли переводить r в Z и обратно?
4. Какова сфера применения корреляционного анализа в биологических исследованиях?

ЛАБОРАТОРНАЯ РАБОТА 10
ЛИНЕЙНЫЙ РЕГРЕССИОННЫЙ АНАЛИЗ

Материалы и оборудование: таблицы; калькуляторы.

Цель: определение величины коэффициентов уравнений регрессии методом наименьших квадратов и их статистической оценки.

Пояснения к заданиям

Расчет линии регрессии. Регрессионный анализ предполагает аналитическое выражение вероятностной связи между признаками различного вида.

Регрессионные модели обычно используют для выражения разного рода связей в лесной таксации, лесоводстве и в других лесных дисциплинах. Чаще всего они применяются для нахождения общей зависимости по экспериментальным данным. Выведенное уравнение сглаживает (выравнивает) полученные (экспериментальные) данные. В этом случае сохраняется главная тенденция изменения функции в зависимости от изменения аргументов, и устраняются случайные отклонения.

По форме различают линейную регрессию и не линейную. По направлению связи различают прямую (т. е. с увеличением признака x увеличивается признак y) и обратную (т. е. с увеличением x уменьшается y). Наиболее точная оценка принадлежности к виду связи производится с помощью *метода наименьших квадратов* (МНК). При минимальном МНК сумма квадратов отклонений эмпирических значений y от теоретических полученных по выбранному уравнению регрессии стремится к минимуму.

Простейшей теоретической линией регрессии является прямая линия, или парабола первого порядка, которая имеет вид:

$$Y' = a_0 + a_1 \cdot x,$$

где Y' – теоретические значения функции или зависимой переменной;

x – аргумент или независимая переменная;

a_0, a_1 – коэффициенты уравнения, имеющие различное значение в зависимости от специфики изучаемого явления;

y – эмпирические значения зависимой переменной.

Коэффициенты определяются по формулам:

$$a_0 = \overline{Y} - a_1 \cdot \overline{X};$$

$$a_1 = \frac{\sum(x - \overline{X}) \cdot (y - \overline{Y})}{\sum(x - \overline{X})^2},$$

где $\overline{X}$ и $\overline{Y}$ средние арифметические рядов аргументов (X) и функции (Y).

Ход вычислений приведен в таблице 13.

Таблица 13 – Схема регрессионного анализа

x	y	$x-\overline{X}$	$(x-\overline{X})^2$	$(y-\overline{Y})$	$(x-\overline{X})\cdot(y-\overline{Y})$	$(y-\overline{Y})^2$	Y'
Σ	Σ	Σ	Σ		Σ	Σ	

В качестве исходных данных используются вариационные ряды частичной совокупности или средние значения высот и диаметров при вычислении смешанных моментов.

Если известны среднеквадратические отклонения для рядов x и y и найден коэффициент корреляции между ними (см. лабораторную работу 9), то величина a_1 вычисляется по формуле:

$$a_1 = r_{xy} \cdot \left(\frac{\sigma_y}{\sigma_x}\right),$$

где σ_y, σ_x – дисперсии выборочных (или усредненных) рядов;

r_{xy} – мера корреляционной взаимосвязи.

Точки пересечения с осями ординат и абсцисс равны соответственно

$$y = a_0,\ x = -\left(\frac{a_0}{a_1}\right).$$

Оценка регрессионных уравнений. Поскольку в определении линий регрессии участвуют несколько параметров, то необходимо оценить пределы изменчивости каждого из них.

Наиболее вероятная область расположения линии прямой регрессии по отношению к оси абсцисс определяется величиной коэффициента a_1 и тангенсом угла, геометрическим смыслом которого является коэффициент корреляции. При отсутствии регрессии $r = 0$, линия регрессии y по x располагается горизонтально по отношению к оси абсцисс, а линия регрессии x по y – вертикально. Место их пересечения соответствует средним значениям обоих признаков.

Второй коэффициент определяет величину отрезка, отсекаемого на оси y линией регрессии. Его величина определяет границы регрессии по ординате, которая расширяется в обе стороны от средней точки $(\overline{x},\ \overline{y})$.

Поскольку опытные данные всегда имеют определенную величину изменчивости, то и все показатели, в том числе и уравнения регрессии, определяются с некоторой степенью достоверности.

Определение величины ошибки найденных уравнений и оценка достоверности полученных коэффициентов уравнения прямой проводится по формулам:

$$m_{y\cdot x}=\sqrt{\frac{\Sigma(y-y')^2}{N-n}},$$

или

$$\sigma_{y\cdot x}=\sqrt{\frac{\Sigma(\mathrm{y}_i-y_i')^2}{n-2}},$$

где $\sigma_{y\cdot x}$, $m_{y\cdot x}$ – ошибка уравнения;

y – эмпирические значения функции;

y' – теоретические значения функции;

N – число точек эмпирической линии регрессии, по которым вычислялось уравнение регрессии;

n – число коэффициентов уравнения, включая свободный член.

Здесь величина $\sigma_{y\cdot x}$ имеет такое же значение как и σ в вариационном ряду. В пределах одной $\sigma_{y\cdot x}$ отклонения распределяются вверх и вниз от линии регрессии в 68 % случаев. В 95 % они лежат в пределах $2\sigma_{y\cdot x}$, а в 99,7 % случаев отклонения от теоретической линии регрессии составляют величину $3\sigma_{y\cdot x}$. Ошибку уравнения регрессии можно определить и по формуле:

$$m_{y\cdot x}=\sigma_y\sqrt{1-r^2},$$

где $m_{y\cdot x}$ – ошибка теоретических значений функции;

σ_y – среднее квадратическое отклонение ряда y;

r – коэффициент корреляции между x и y (можно использовать и корреляционное отношение при наличии криволинейности связи между признаками).

Эта формула представляет упрощенный вариант вычислений и применяется для больших выборок.

В таблицах по вычислению коэффициентов уравнений в последней колонке рассчитываются теоретические значения функции. Получаем попарно разности $(y-y')$, возводим все разности в квадрат и получаем их сумму: $\sum(y-y')^2$. Применив формулу $m_{y\cdot x}=\sqrt{\dfrac{\sum(y-y')^2}{N-n}}$ к уравнению прямой параболы, определим их ошибку.

Достоверность найденного коэффициента a_1 определяется по формуле

$$t = \frac{a_1 \cdot \sigma_x \cdot \sqrt{N-1}}{m_{yx}},$$

где t – величина критерия Стьюдента, сравниваемая с критической при числе степеней свободы $\nu = N - 2$;

σ_x – среднее квадратическое отклонение ряда аргументов;

m_{yx} – ошибка уравнения;

N – объем выборки.

Если вычисленная величина меньше табличной, то связь между *x, y* и значением a_1 достоверны, а если вычисленная величина будет больше табличной, то связь данных признаков и значение первого коэффициента недостоверны.

Достоверность отличия от нуля коэффициента a_0 можно оценить по формуле:

$$t = \frac{a_0}{m_{y\cdot x}\sqrt{\dfrac{1}{N}+\dfrac{1}{N-1}\cdot\left(\dfrac{\overline{O''}}{\sigma_x}\right)^2}},$$

где t – величина критерия Стьюдента, сравниваемая с критической при числе степеней свободы $\nu = N - 2$;

σ_x – среднее квадратическое отклонение ряда аргументов;

m_{yx} – ошибка уравнения;

N – объем выборки.

Ход работы

Задание 1. На основе основных положений темы и расчетов по лабораторной работе 9 определить взаимообусловленность признаков X и Y аналитическими уравнениями.

Задание 2. Определить ошибки регрессионных уравнений и достоверность коэффициентов линейных уравнений.

Вопросы для самоконтроля

1. Приведите формулу общего уравнения регрессии.
2. В чем сущность коэффициента регрессии?
3. Приведите алгоритм построения эмпирических рядов регрессии.
4. В чем сущность коэффициента корреляции.
5. Укажите типы связи между переменными.

ЛАБОРАТОРНАЯ РАБОТА 11
ИТОГОВОЕ ЗАНЯТИЕ (ЗАЧЕТНАЯ ЗАДАЧА)

Материалы и оборудование: таблицы; калькуляторы.

Цель: Определение статистических величин и их статистической оценки.

Ход работы

Задание. Решить зачетную задачу описательной статистики по усвоенной ранее схеме и ответить на теоретический вопрос.

Задача 1. Изучена плодовитость (число щенков) самок серебристо-черных лисиц.

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

Провести анализ:

1. Составить вариационный ряд (ВР) и изобразить его графически.
2. Вычислить статистические характеристики (Mo, Me, X, ∂, m_x, P, Cv).
3. Дать оценку достоверности (доверительный интервал при трех уровнях значимости; H_0; критерий Стьюдента). Сделать обоснованный вывод. Ответ.
4. Понятие корреляции. Привести формулу коэффициента корреляции прямым способом через центральные отклонения.

Задача 2. Получены данные о количестве хвостовых щитков у змей.

42	58	44	54	41	50	46	46	54	48	43	49
50	48	46	46	45	53	48	48	53	53	48	41
46	40	50	43	49	51	52	46	42	44	48	45
47	46	43	50	47	45	48	40	44	42	48	45
54	50	56	48	45	45	51	42	44	47	46	45

Провести анализ:

1. Составить ВР и изобразить его графически.
2. Вычислить статистические характеристики (Mo, Me, X, ∂, m_x, P, Cv).
3. Дать оценку достоверности (доверительный интервал при трех уровнях значимости; H_0; критерий Стьюдента). Сделать обоснованный вывод.
4. Что такое вероятность? Привести формулу расчета вероятности случайного события.

Задача 3. Имеются данные о весе кроликов (в кг).

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

Провести анализ:

1. Составить ВР и изобразить его графически.
2. Вычислить статистические характеристики (Mo, Me, X, ∂, m_x, P, Cv).
3. Дать оценку достоверности (доверительный интервал при трех уровнях значимости; H_0; критерий Стьюдента). Сделать обоснованный вывод.
4. Привести алгоритм дисперсионного анализа (ДА I).

ЛИТЕРАТУРА

1. Васильева, Л. А. Статистические методы в биологии, медицине и сельском хозяйстве : учебное пособие / Л. А. Васильева. – Новосибирск : Институт цитологии и генетики СО РАН, 2007. – 124 с.

2. Красильников, В. В. Высшая математика. Вероятность. Статистика. Исследование операций : учебное пособие / В. В. Красильников. – Набережные Челны : Печатный двор, 1996. – 225 с.

3. Петри, А. Наглядная статистика в медицине / А. Петри, К. Сэбин. – М. : Издательский дом Гэотар-мед, 2003. – 143 с.

4. Тукшаитов, Р. Х. Основы оптимального представления статистических показателей на графиках, диаграммах и в таблицах / Р. Х. Тукшаитов. – Киев : Типография КГЭУ, 2006. – 227 с.

5. Рокицкий, П. Ф. Биологическая статистика / П. Ф. Рокицкий. – Минск : Высшая школа, 1973. – 320 с.

ПРИЛОЖЕНИЕ А
(обязательное)

Таблица вероятностей при нормальном распределении

Таблица А.1 – Таблица вероятностей при нормальном распределении. Доли площади под нормальной кривой в пределах от $-t$ до $+t$

t	Сотые доли t									
	0	1	2	3	4	5	6	7	8	9
0,0	0000	0080	0160	0239	0319	0399	0478	0558	0638	0717
0,1	0797	0876	0955	1034	1113	1192	1271	1350	1428	1507
0,2	1585	1663	1741	1819	1897	1974	2051	2128	2205	2282
0,3	2358	2434	2510	2586	2661	2737	2812	2886	2961	3035
0,4	3108	3182	3255	3328	3401	3473	3545	3616	3688	3759
0,5	3829	3899	3969	4039	4108	4177	4245	4313	4381	4448
0,6	4515	4581	4647	4713	4778	4843	4907	4971	5035	5098
0,7	5161	5223	5285	5346	5407	5467	5527	5587	5646	5705
0,8	5763	5821	5878	5935	5991	6047	6102	6157	6211	6265
0,9	6319	6372	6424	6476	6528	6579	6629	6680	6729	6778
1,0	6827	6875	6923	6970	7017	7063	7109	7154	7199	7243
1,1	7287	7330	7373	7415	7457	7499	7540	7580	7620	7660
1,2	7699	7737	7775	7813	7850	7887	7923	7959	7995	8029
1,3	8064	8098	8132	8165	8198	8230	8262	8293	8324	8355
1,4	8385	8415	8444	8473	8501	8529	8557	8584	8611	8638
1,5	8664	8690	8715	8740	8764	8789	8812	8836	8859	8882
1,6	8904	8926	8948	8969	8990	9011	9031	9051	9070	9090
1,7	9109	9127	9146	9164	9181	9199	9216	9233	9249	9265
1,8	9281	9297	9312	9327	9342	9357	9371	9385	9399	9412
1,9	9426	9439	9451	9464	9476	9488	9500	9512	9523	9534
2,0	9545	9556	9566	9576	9586	9596	9606	9616	9625	9634
2,1	9643	9651	9660	9668	9676	9684	9692	9700	9707	9715
2,2	9722	9729	9736	9743	9749	9756	9762	9768	9774	9780
2,3	9786	9791	9797	9802	9807	9812	9817	9822	9827	9832
2,4	9836	9840	9845	9849	9853	9857	9861	9865	9869	9872
2,5	9876	9879	9883	9886	9889	9892	9895	9898	9901	9904
2,6	9907	9909	9912	9915	9917	9920	9922	9924	9926	9929
2,7	9931	9933	9935	9937	9939	9940	9942	9944	9946	9947
2,8	9949	9960	9952	9953	9955	9956	9958	9959	9960	9961
2,9	9963	9964	9965	9966	9967	9968	9969	9970	9971	9972
3,0	9973	9981	9986	9990	9993	9995	9997	9998	9999	9999

ПРИЛОЖЕНИЕ Б
(обязательное)

t-распределение по Стьюденту

Таблица Б.1 – Доли площади под кривой *t*-распределения по Стьюденту в пределах от $-t$ до $+t$ для малого числа наблюдений *n*

t \ n	2	3	4	5	6	7	8	9	10	12	14	16	18	20	∞
0,1	063	071	073	075	076	076	077	077	077	078	078	078	078	079	080
0,2	126	140	146	149	151	152	153	154	154	155	155	156	156	156	158
0,3	186	208	216	221	224	226	227	228	229	230	231	232	232	233	236
0,4	242	272	284	290	294	297	299	300	302	303	304	305	306	306	311
0,5	295	333	347	357	362	365	368	369	371	373	375	376	377	377	383
0,6	344	391	409	419	425	430	433	435	437	439	441	433	444	444	452
0,7	389	444	466	417	485	490	493	496	498	502	504	505	507	508	516
0,8	430	492	518	531	540	546	550	553	556	558	562	564	565	566	576
0 9	467	537	537	581	591	597	602	606	608	613	616	618	619	621	632
1,0	500	577	609	626	637	644	649	653	657	661	664	667	669	670	683
1,1	530	614	648	667	679	687	692	697	700	705	709	711	713	715	729
1,2	558	647	684	704	716	725	731	736	739	745	748	751	753	755	770
1,3	583	677	716	737	750	759	765	770	774	780	784	788	789	791	806
1,4	605	704	744	766	780	789	796	801	805	811	815	818	821	822	838
1,5	626	728	769	792	806	816	823	828	832	838	842	846	848	850	866
1,6	644	749	792	815	830	839	846	852	856	862	866	870	872	874	890
1,7	661	769	812	836	850	860	867	872	877	883	887	890	893	895	911
1,8	677	786	830	854	868	878	885	890	895	901	905	908	910	912	928
1,9	692	802	846	870	884	894	901	906	910	916	920	923	925	927	943
2,0	705	816	861	884	898	908	914	919	923	929	933	936	938	940	954
2,1	717	829	873	896	910	920	926	931	935	940	944	947	949	951	964
2 2	728	841	885	907	921	930	936	941	945	950	954	956	958	960	972

Окончание таблицы Б.1

$t \backslash n$	2	3	4	5	6	7	8	9	10	12	14	16	18	20	∞
2,3	739	852	895	917	930	939	945	950	953	958	961	964	966	967	979
2,4	749	862	904	926	938	947	953	957	960	965	968	970	972	973	984
2,5	758	870	912	933	946	953	959	963	966	970	973	975	977	978	988
2,6	766	878	920	940	952	959	965	968	971	975	978	980	981	982	991
2,7	774	886	926	946	957	964	969	973	976	979	982	984	985	986	993
2,8	782	893	932	951	962	969	973	977	979	983	985	987	988	989	995
2,9	789	899	937	956	966	973	977	980	982	986	988	989	990	991	996
3,0	795	905	942	960	970	976	980	983	985	988	990	991	992	993	997
3,1	801	910	947	964	973	979	983	985	987	990	992	993	994	994	998
3,2	807	915	951	967	976	981	985	987	989	992	993	994	995	995	999
3,3	813	919	954	970	979	984	987	989	991	993	994	995	996	996	999
3,4	818	923	958	973	981	986	989	991	992	994	995	996	997	997	999
3,5	823	927	961	975	983	987	990	992	993	995	996	997	997	998	1
3,6	828	931	963	977	984	989	991	993	994	996	997	997	998	998	
3,7	832	934	966	979	986	990	992	994	995	996	997	998	998	998	
3,8	836	937	968	981	987	991	993	995	996	997	998	998	999	999	
3,9	840	940	970	982	989	992	994	995	996	998	998	999	999	999	
4,0	844	943	972	984	990	993	995	996	997	998	998	999	999	999	
4,1	848	945	974	985	991	994	995	997	997	998	999	999	999	999	
4,2	851	948	975	986	992	994	996	997	998	999	999	999	999	1	
4,3	855	950	977	987	992	995	996	997	998	999	999	999	999		
4,4	858	952	978	988	993	995	997	998	998	999	999	999	1		
4,5	861	954	980	989	994	996	997	998	999	999	999	1			
4,6	864	956	981	990	994	996	998	998	999	999	1				
4,7	867	958	982	991	995	997	998	998	999	999					
4,8	869	959	983	991	995	997	998	999	999	999					
4,9	872	961	984	992	996	997	998	999	999	1					

ПРИЛОЖЕНИЕ В
(обязательное)

Критерий Стьюдента

Таблица В.1 – Значения *t* при различных уровнях значимости (Р)

Число степеней свободы	Уровень значимости *P*				
	0,1	0,05	0,02	0,01	0,001
1	6,31	12,7	31,82	63,66	—
2	2,92	4,30	6,97	9,93	31,60
3	2,35	3,18	4,54	5,84	12,94
4	2,13	2,78	3,75	4,60	8,61
5	2,02	2,57	3,37	4,03	6,86
6	1,94	2,45	3,14	3,71	5,96
7	1,90	2,37	3,00	3,50	5,41
8	1,86	2,31	2,90	3,36	5,04
9	1,83	2,26	2,82	3,25	4,78
10	1,81	2,23	2,76	3,17	4,59
11	1,80	2,20	2,72	3,11	4,44
12	1,78	2,18	2,68	3,06	4,32
13	1,77	2,16	2,65	3,01	4,22
14	1,76	2,15	2,62	2,98	4,14
15	1,75	2,13	2,60	2,95	4,07
16	1,75	2,12	2,58	2,92	4,02
17	1,74	2,11	2,57	2,90	3,97
18	1,73	2,10	2,55	2,88	3,92
19	1,73	2,09	2,54	2,86	3,88
20	1,73	2,09	2,53	2,85	3,85
21	1,72	2,08	2,52	2,83	3,82
22	1,72	2,07	2,51	2,82	3,79
23	1,71	2,07	2,50	2,81	3,77
24	1,71	2,06	2,49	2,80	3,75
25	1,71	2,06	2,49	2,79	3,73
26	1,71	2,06	2,48	2,78	3,71
27	1,70	2,05	2,47	2,77	3, 69
28	1,70	2,05	2,47	2,76	3,67
29	1,70	2,05	2,46	2,76	3,66
30	1,70	2,04	2,46	2,75	3,65
∞	1,64	1,96	2,33	2,58	3,29

ПРИЛОЖЕНИЕ Г
(обязательное)

Критерий Пирсона

Таблица Г.1 – χ^2-распределение

df	Вероятности значения χ^2, превышающего табличное									
	0,99	0,95	0,90	0,75	0,50	0,25	0,10	0,05	0,025	0,01
1	....	....	0,02	0,10	0,45	1,32	2,71	3,84	5,02	6,63
2	0,02	0,10	0,21	0,58	1,39	2,77	4,61	5,99	7,38	9,21
3	0,11	0,35	0,58	1,21	2,37	4,11	6,25	7,81	9,35	11,34
4	0,30	0,71	1,06	1,92	3,36	5,39	7,78	9,49	11,14	13,28
5	0,55	1,15	1,61	2,67	4,35	6,63	9,24	11,07	12,83	15,09
6	0,87	1,64	2,20	3,45	5,35	7,84	10,64	12,59	14,45	16,81
7	1,24	2,17	2,83	4,25	6,35	9,04	12,02	14,07	16,01	18,48
8	1,65	2,73	3,49	5,07	7,34	10,22	13,36	15,51	17,53	20,09
9	2,09	3,33	4,17	5,90	8,34	11,39	14,68	16,92	19,02	21,67
10	2,56	3,94	4,87	6,74	9,34	12,55	15,99	18,31	20,48	23,21
11	3,05	4,57	5,58	7,58	10,34	13,70	17,28	19,68	21,92	24,72
12	3,57	5,23	6,30	8,44	11,74	14,85	18,55	21,03	23,34	26,22
13	4,11	5,89	7,04	9,30	12,34	15,98	19,81	22,36	24,74	27,69
14	4,66	6,57	7,79	10,17	13,34	17,12	21,06	23,68	26,12	29,14
15	5,23	7,26	8,55	11,04	14,34	18,25	22,31	25,00	27,49	30,58
16	5,81	7,96	9,31	11,91	15,34	19,37	23,54	26,30	28,85	32,00
17	6,41	8,67	10,09	12,79	16,34	20,49	24,77	27,59	30,19	33,41
18	7,01	9,39	10,86	13,68	17,34	21,60	25,99	28,87	31,53	34,81
19	7,63	10,12	11,65	14,56	18,34	22,72	27,20	30,14	32,85	36,19
20	8,26	10,85	12,44	15,45	19,34	23,83	28,41	31,41	34,17	37,57
21	8,90	11,59	13,24	16,34	20,34	24,93	29,62	32,67	35,48	38,93
22	9,54	12,34	14,04	17,24	21,34	26,04	30,81	33,92	36,78	40,29
23	10,20	13,09	14,85	18,14	22,34	27,14	32,01	35,17	38,08	41,64
24	10,86	13,85	15,66	19,04	23,34	28,24	33,20	36,42	39,36	42,98
25	11,52	14,61	16,47	19,94	24,34	29,34	34,38	37,65	40,65	44,31
26	12,20	15,38	17,29	20,84	25,34	30,43	35,56	38,89	41,92	45,64
27	12,88	16,15	18,11	21,75	26,34	31,53	36,74	40,11	43,19	46,96
28	13,56	16,93	18,94	22,66	27,34	32,62	37,92	41,34	44,46	48,28
29	14,26	17,71	19,77	23,57	28,34	33,71	39,09	42,56	45,72	49,59
30	14,95	18,49	20,60	24,48	29,34	34,80	40,26	43,77	46,98	50,89
40	22,16	26,51	29,05	33,66	39,34	45,62	51,80	55,76	59,34	63,69
50	29,71	34,76	37,69	42,94	49,33	56,33	63,17	67,50	71,42	76,15
60	37,48	43,19	46,46	52,29	59,33	66,98	74,40	79,08	83,30	88,38
70	45,44	51,74	55,33	61,70	69,33	77,58	85,53	90,53	95,02	100,42
80	53,54	60,39	64,28	71,14	79,33	88,13	96,58	101,88	106,63	112,33
90	61,75	69,13	73,29	80,62	89,33	98,64	107,56	113,14	118,14	124,12
100	70,06	77,93	82,36	90,13	99,33	109,14	118,50	124,34	129,56	135,81

ПРИЛОЖЕНИЕ Д

(обязательное)

Критерий Фишера (P = 0,05)

Таблица Д.1 – Значения F при уровне значимости 0,05 (df_1 – число степеней свободы для большей варианси, которая берется числителем)

df_2 \ df_1	1	2	3	4	5	6	7	8	9	10	12	15	20	30	∞
1	161	200	216	225	230	234	237	239	241	242	244	246	248	250	254
2	18,51	19,00	19,16	19,25	19,30	19,33	19,35	19,37	19,38	19,40	19,41	19,43	19,45	19,46	19,50
3	10,13	9,55	9,28	9,12	9,01	8,94	8,89	8,85	8,81	8,79	8,74	8,70	8,66	8,62	8,53
4	7,71	6,94	6,59	6,39	6,26	6,16	6,09	6,04	6,00	5,94	5,91	5,86	5,80	5,75	5,63
5	6,61	5,79	5,41	5,19	5,05	4,95	4,88	4,82	4,77	4,74	4,68	4,62	4,56	4,50	4,36
6	5,99	5,14	4,76	4,53	4,39	4,28	4,21	4,15	4,10	4,06	4,00	3,94	4,87	3,81	3,67
7	5,59	4,74	4,35	4,12	3,97	3,87	3,79	3,73	3,68	3,64	3,57	3,51	3,44	3,38	3,23
8	5,32	4,46	4,07	3,84	3,69	3,58	3,50	3,44	3,39	3,35	3,28	3,22	3,15	3,08	2,93
9	5,12	4,26	3,86	3,63	3,48	3,37	3,29	3,23	3,18	3,14	3,07	3,01	2,94	2,86	2,71
10	4,96	4,10	3,71	3,48	3,33	3,22	3,14	3,07	3,02	2,98	2,91	2,85	3,77	2,70	2,54
11	4,84	3,98	3,59	3,36	3,20	3,09	3,01	2,95	2,90	2,85	2,79	2,72	2,65	2,57	2,40
12	4,75	3,89	3,49	3,26	3,11	3,00	2,91	2,85	2,80	2,75	2,69	2,62	2,54	2,47	2,30
13	4,67	3,80	3,41	3,18	3,02	2,92	2,83	2,77	2,71	2,67	2,60	2,53	2,46	2,38	2,21
14	4,60	3,74	3,34	3,11	2,96	2,85	2,76	2,70	2,65	2,60	2,53	2,46	2,39	2,31	2,13
15	4,54	3,68	3,29	3,06	2,90	2,79	2,71	2,64	2,59	2,54	2,48	2,40	2,33	2,25	2,07

Окончание таблицы Д.1

df_2 \ df_1	1	2	3	4	5	6	7	8	9	10	12	15	20	30	∞
16	4,49	3,63	3,24	3,01	2,85	2,74	2,66	2,59	2,54	2,49	2,42	2,35	2,28	2,19	2,01
17	4,45	3,59	3,20	2,96	2,81	2,70	2,61	2,55	2,49	2,45	2,38	2,31	2,23	2,15	1,96
18	4,41	3,55	3,16	2,93	2,77	2,66	2,58	2,51	2,46	2,41	2,34	2,27	2,19	2,11	1,92
19	4,38	3,52	3,13	2,90	2,74	2,63	2,54	2,48	2,42	2,38	2,31	2,23	2,16	2,07	1,88
20	4,35	3,49	3,10	2,87	2,71	2,60	2,51	2,45	2,39	2,35	2,28	2,20	2,12	2,04	1,84
21	4,32	3,47	3,07	2,84	2,68	2,57	2,49	2,42	2,37	2,32	2,25	2,18	2,10	2,01	1,81
22	4,30	3,44	3,05	2,82	2,66	2,55	2,46	2,40	2,34	2,30	2,23	2,15	2,07	1,98	1,78
23	4,28	3,42	3,03	2,80	2,64	2,53	2,44	2,37	2,32	2,27	2,20	2,13	2,05	1,96	1,76
24	4,26	3,40	3,01	2,78	2,62	2,51	2,42	2,36	2,30	2,25	2,18	2,11	2,03	1,94	1,73
25	4,24	3,39	2,99	2,76	2,60	2,49	2,40	2,34	2,28	2,24	2,16	2,09	2,01	1,92	1,71
26	4,22	3,37	2,98	2,74	2,59	2,47	2,39	2,32	2,27	2,22	2,15	2,07	1,99	1,90	1,69
27	4,21	3,35	2,96	2,73	2,57	2,46	2,37	2,31	2,25	2,20	2,13	2,06	1,97	1,88	1,67
28	4,20	3,34	2,95	2,71	2,56	2,45	2,36	2,29	2,24	2,19	2,12	2,04	1,96	1,87	1,65
29	4,18	3,33	2,93	2,70	2,55	2,43	2,35	2,28	2,22	2,18	2,10	2,03	1,94	1,85	1,64
30	4,17	3,32	2,92	2,69	2,53	2,42	2,33	2,27	2,21	2,16	2,09	2,01	1,93	1,84	1,62
40	4,08	3,23	2,84	2,61	2,45	2,34	2,25	2,18	2,12	2,08	2,00	1,92	1,84	1,74	1,51
60	4,00	3,15	2,76	2,53	2,37	2,25	2,17	2,10	2,04	1,99	1,92	1,84	1,75	1,65	1,39
120	3,92	3,07	2,68	2,45	2,29	2,17	2,09	2 02	1,96	1,91	1,83	1,75	1,66	1,55	1,25
∞	3,84	3,00	2,60	2,37	2,21	2,10	2,01	1,94	1,88	1,83	1,75	1,67	1,57	1,46	1,00

ПРИЛОЖЕНИЕ Е
(обязательное)

Критерий Фишера (P = 0,01)

Таблица Е.1 – Значения F при уровне значимости 0,01 (df_1 – число степеней свободы для большей варианса, которая берется числителем)

df_2 \ df_1	1	2	3	4	5	6	7	8	9	10	12	15	20	30	
1	4052	4999	5403	5625	5764	5859	5928	5982	6022	6056	6106	6157	6209	6261	6366
2	98,50	99,00	99,17	99,25	99,30	99,33	99,36	99,37	99,39	99,40	99,42	99,43	99,45	99,47	99,50
3	34,12	30,82	29,46	28,71	28,42	27,91	27,67	27,49	27,35	27,23	27,05	26,87	26,69	26,50	26,13
4	21,20	18,00	16,69	15,98	15,52	15,21	14,98	14,80	14,66	14,55	14,37	14,20	14,02	13,84	13,46
5	16,26	13,27	12,06	11,39	10,97	10,67	10,46	10,29	10,16	10,05	9,89	9,72	9,55	9,38	9,02
6	13,75	10,92	9,78	9,15	8,75	8,47	8,26	8,10	7,98	7,87	7,72	7,56	7,40	7,23	6,88
7	12,25	9,55	8,47	7,85	7,46	7,19	6,99	6,84	6,72	6,62	6,47	6,31	6,16	5,99	5,65
8	11,26	8,65	7,59	7,01	6,63	6,37	6,18	6,03	5,91	5,81	5,67	5,52	5,36	5,20	4,86
9	10,56	8,02	6,99	6,42	6,06	5,80	5,61	5,47	5,35	5,26	5,11	4,96	4,81	4,65	4,31
10	10,04	7,56	6,55	5,99	5,64	5,39	5,20	5,06	4,94	4,85	4,71	4,56	4,41	4,25	3,91
11	9,65	7,21	6,22	5,67	5,32	5,07	4,89	4,74	4,63	4,54	4,40	4,25	4,10	3,94	3,60
12	9,33	6,93	5,95	5,41	5,06	4,82	4,64	4,50	4,39	4,30	4,16	4,01	3,86	3,70	3,36
13	9,07	6,70	5,74	5,21	4,86	4,62	4,44	4,30	4,19	4,10	3,96	3,82	3,66	3,51	3,16
14	8,86	6,51	5,56	5,04	4,69	4,46	4,28	4,14	4,03	3,94	3,80	3,66	3,51	3,35	3,00
15	8,68	6,36	5,42	4,89	4,56	4,32	4,14	4,00	3,89	3,80	3,67	3,52	3,37	3,21	2,87

Окончание таблицы Е.1

df_2 \ df_1	1	2	3	4	5	6	7	8	9	10	12	15	20	30	∞
16	8,53	6,23	5,29	4,77	4,44	4,20	4,03	3,89	3,78	3,69	3,55	3,41	3,26	3,10	2,75
17	8,40	6,11	5,18	4,67	4,34	4,10	3,93	3,79	3,68	3,59	3,46	3,31	3,16	3,00	2,65
18	8,29	6,01	5,09	4,58	4,25	4,01	3,84	3,71	3,60	3,51	3,37	3,23	3,08	2,92	2,57
19	8,18	5,93	5,01	4,50	4,17	3,94	3,77	3,61	3,52	3,43	3,30	3,15	3,00	2,84	2,49
20	8,10	5,85	4,94	4,43	4,10	3,87	3,70	3,56	3,46	3,37	3,23	3,09	2,94	2,78	2,42
21	8,02	5,78	4,87	4,37	4,04	3,81	3,64	3,51	3,40	3,31	3,17	3,03	2,88	2,72	2,36
22	7,95	5,72	4,82	4,31	3,99	3,76	3,59	3,45	3,35	3,26	3,12	2,98	2,83	2,67	2,31
23	7,88	5,66	4,76	4,26	3,94	3,71	3,54	3,41	3,30	3,21	3,07	2,93	2,78	2,62	2,26
24	7,82	5,61	4,72	4,22	3,90	3,67	3,50	3,36	3,26	3,17	3,03	2,89	2,74	2,58	2,21
25	7,77	5,57	4,68	4,18	3,85	3,63	3,46	3,32	3,22	3,13	2,99	2,85	2,70	2,54	2,17
26	7,72	5,53	4,64	4,14	3,82	3,59	3,42	3,29	3,18	3,09	2,96	2,81	2,66	2,50	2,13
27	7,68	5,49	4,60	4,11	3,78	3,56	3,39	3,26	3,15	3,06	2,93	2,78	2,63	2,47	2,10
28	7,64	5,45	4,57	4,07	3,75	3,53	3,36	3,23	3,12	3,03	2,90	2,75	2,60	2,44	2,06
29	7,60	5,42	4,54	4,04	3,73	3,50	3,33	3,20	3,09	3,00	2,87	2,73	2,57	2,41	2,03
30	7,56	5,39	4,51	4,02	3,70	3,47	3,30	3,17	3,07	2,98	2,84	2,70	2,55	7,39	2,01
40	7,31	5,18	4,31	3,85	3,51	3,29	3,12	2,99	2,89	2,80	2,66	2,52	2,37	2,20	1,80
60	7,08	4,98	4,13	3,63	3,34	3,12	2,95	2,82	2,72	2,63	2,50	2,35	2,20	2,03	1,60
120	6,85	4,79	3,95	3,48	3,17	2,96	2,79	2,66	2,56	2,47	2,34	2,19	2,03	1,86	1,3 9
∞	6,63	4,61	3,78	3,32	3,02	2,80	2,64	2,51	2,41	2,32	2,18	2,04	1,88	1,70	1,00

ПРИЛОЖЕНИЕ Ж
(обязательное)

Коэффициент корреляции

Таблица Ж.1 – Необходимые значения коэффициента корреляции *r* при различных уровнях значимости *P* и разном числе степеней свободы *df* (*df* = *n* – 2)

df	*P*		*df*	*P*	
	0,05	0,01		0,05	0,01
5	0,75	0,87	27	0,37	0,47
6	0,71	0,83	28	0,36	0,46
7	0.67	0,80	29	0,36	0,46
8	0,63	0,77	30	0,35	0,45
9	0,60	0,74	35	0,33	0,42
10	0,58	0,71	40	0,30	0,39
11	0,55	0,68	45	0,29	0,37
12	0,53	0,66	50	0,27	0,35
13	0,51	0,64	60	0,25	0,33
14	0,50	0,62	70	0,23	0,30
15	0,48	0,61	80	0,22	0,28
16	0,47	0,59	90	0,21	0,27
17	0,46	0,58	100	0,20	0,25
18	0,44	0,56	125	0,17	0,23
19	0,43	0,55	150	0,16	0,21
20	0,42	0,54	200	0,14	0,18
21	0,41	0,53	300	0,11	0,15
22	0,40	0,52	400	0,10	0,13
23	0,40	0,51	500	0,09	0,12
24	0,39	0,50	700	0,07	0,10
25	0,38	0,49	900	0,06	0,09
26	0,37	0,48	1000	0,06	0,09

ПРИЛОЖЕНИЕ И

(обязательное)

Метод Z

Таблица И.1 – Значения *r* при различных величинах *z* (от 0 до 2,99). Для краткости ноль перед коэффициентом корреляции опущен, поэтому 0997 надо читать как 0,0997

z	Сотые доли *z*									
	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0,0	0000	0100	0200	0300	0400	0500	0599	0699	0798	0898
0,1	0997	1096	1194	1293	1391	1489	1586	1684	1781	1877
0,2	1974	2070	2165	2260	2355	2449	2543	2636	2729	2821
0,3	2913	3004	3095	3185	3275	3364	3452	3540	3627	3714
0,4	3800	3885	3969	4053	4136	4219	4301	4382	4462	4542
0,5	4621	4699	4777	4854	4930	5005	5080	5154	5227	5299
0,6	5370	5441	5511	5580	5649	5717	5784	5850	5915	5980
0,7	6044	6107	6169	6231	6291	6351	6411	6469	6527	6584
0,8	6640	6696	6751	6805	6858	6911	6963	7014	7064	7114
0,9	7163	7211	7259	7306	7352	7398	7443	7487	7531	7574
1,0	7616	7658	7699	7739	7779	7818	7857	7895	7932	7969
1,1	8005	8041	8076	8110	8144	8178	8210	8243	8275	8306
1,2	8337	8367	8397	8426	8455	8483	8511	8538	8565	8591
1,3	8617	8643	8658	8692	8717	8741	8764	8787	8810	8832
1,4	8854	8875	8896	8917	8937	8957	8977	8996	9015	9033
1,5	9051	9069	9087	9104	9112	9138	9154	9170	9186	9201
1,6	9217	9232	9246	9261	9275	9289	9302	9316	9329	9341
1,7	9354	9366	9379	9391	9402	9414	9425	9436	9447	9458
1,8	9468	9478	9488	9498	9508	9517	9527	9536	9545	9554
1,9	9562	9571	9579	9587	9595	9603	9611	9618	9626	9633
2,0	9640	9647	9654	9661	9668	9674	9680	9687	9693	9699
2,1	9705	9710	9716	9721	9727	9732	9738	9743	9748	9753
2,2	9757	9762	9767	9771	9776	9780	9785	9789	9793	9797
2,3	9801	9805	9809	9812	9816	9820	9823	9827	9830	9834
2,4	9837	9840	9843	9846	9849	9852	9855	9858	9861	9863
2,5	9866	9869	9871	9874	9876	9879	9881	9883	9886	9888
2,6	9890	9892	9894	9897	9899	9901	9903	9904	9906	9908
2,7	9910	9912	9914	9915	9917	9919	9920	9922	9923	9925
2,8	9926	9928	9929	9931	9932	9933	9935	9936	9937	9938
2,9	9940	9941	9942	9943	9944	9945	9946	9947	9948	9949

Производственно-практическое издание

Кураченко Ирина Витальевна,
Зятьков Сергей Александрович,
Гончаренко Григорий Григорьевич

БИОМЕТРИЯ

Практическое руководство

Редактор А. А. Банчук
Корректор В. В. Калугина

Подписано в печать 01.11.2022. Формат 60x84 1/16.
Бумага офсетная. Ризография.
Усл. печ. л. 2,8. Уч.-изд. л. 3,1.
Тираж 10 экз. Заказ 535.

Издатель и полиграфическое исполнение:
учреждение образования
«Гомельский государственный университет имени Франциска Скорины».
Свидетельство о государственной регистрации издателя, изготовителя,
распространителя печатных изданий № 3/1452 от 17.04.2017.
Специальное разрешение (лицензия) № 02330 / 450 от 18.12.2013.
Ул. Советская, 104, 246028, Гомель.