

МИНИСТЕРСТВО СЕЛЬСКОГО ХОЗЯЙСТВА РФ  
ДЕПАРТАМЕНТ НАУЧНО-ТЕХНОЛОГИЧЕСКОЙ ПОЛИТИКИ И ОБРАЗОВАНИЯ  
ФГБОУ ВО КОСТРОМСКАЯ ГСХА

Кафедра  
частной зоотехнии, разведения и генетики

## **ГЕНЕТИКА И БИОМЕТРИЯ**

**Часть I. Биометрические методы  
анализа качественных и количественных признаков**

Учебное пособие  
для студентов направления подготовки  
36.03.02 «Зоотехния»  
очной и очно-заочной форм обучения

КАРАБАЕВО  
Костромская ГСХА  
2015

УДК 578.087.1 : 636

ББК 45.31

Г 35

*Составитель:* к.с.-х.н., доцент кафедры частной зоотехнии, разведения и генетики Костромской ГСХА С.Г. Белокуров.

*Рецензент:* д.б.н., профессор, директор ГНУ Костромского НИИСХ А.В. Баранов.

*Рекомендовано к изданию методической комиссией факультета ветеринарной медицины и зоотехнии, протокол № 6 от 10 декабря 2014 года.*

**Г 35      Генетика и биометрия. Часть 1. Биометрические методы анализа качественных и количественных признаков :** учебное пособие для студентов направления подготовки 36.03.02 «Зоотехния» очной и очно-заочной форм обучения / сост. С.Г. Белокуров. — Караваево : Костромская ГСХА, 2015. — 46 с.

В издании изложены методики вычисления основных статистических параметров и коэффициентов, применяемых в зоотехнии животных на примере больших и малых выборок.

Учебное пособие предназначено для студентов направления подготовки 36.03.02 «Зоотехния» очной и очно-заочной форм обучения, может быть полезным магистрантам, аспирантам и преподавателям биологических дисциплин.

УДК 578.087.1 : 636

ББК 45.31

## ОГЛАВЛЕНИЕ

Введение .....	4
1. Условные обозначения основных биометрических параметров.....	6
2. Понятие о генеральной и выборочной совокупности .....	7
3. Статистический анализ в многочисленных выборках ( $n > 30$ ).....	8
3.1. Вариационный ряд и вариационная кривая.....	8
3.2. Средняя арифметическая $\bar{X}$ .....	12
3.3. Среднее квадратическое (стандартное) отклонение $\sigma$ .....	13
3.4. Коэффициент вариации $C_v$ .....	14
3.5. Взаимосвязь между признаками .....	15
3.6. Оценка достоверности статистических показателей.....	20
4. Статистический анализ в малочисленных выборках ( $n = 3 \div 30$ ) .....	24
5. Статистический анализ признаков с альтернативной изменчивостью .....	27
6. Основы дисперсионного анализа.....	30
7. Задания для самостоятельной работы.....	32
8. Вопросы для самоконтроля.....	39
9. Краткий словарь терминов, применяемых в биометрии.....	41
Список рекомендуемых источников .....	43
Приложения .....	44

## ВВЕДЕНИЕ

Развитие современного животноводства сопровождается накоплением большого количества информации по многим вопросам общей, прикладной генетики и селекции. При этом остро стоит проблема классификации данных, их упорядочение и систематизация, а также научно-практический анализ, завершающийся формулировкой реальных предложений для дальнейшего развития той или иной отрасли.

Специалистам (зоотехникам, ветеринарным врачам) в своей повседневной производственной деятельности приходится иметь дело с большими группами животных: вид, порода, стадо, линии, семейства, которые характеризуются большим разнообразием признаков, позволяющим отличать их друг от друга. В то же время отдельные животные, входящие в состав этих групп, также различаются между собой, но в целом им свойственны сходные признаки, типичные для вида, породы, стада, линии, семейства. Таким образом, при изучении и работе с большими группами животных выясняется значительная изменчивость признаков, что создает трудности для их характеристики и анализа. Наряду с этим установлено, что все биологические явления, свойства и признаки у множества особей одного вида имеют статистическую природу, то есть подчиняются математическим закономерностям. Это вызвало необходимость их приложения к изучению биологических объектов.

Союз математики и биологии составляет содержание биометрии — науки о статистическом анализе групповых свойств биологических объектов. Термин «биометрия» (от лат. *bios* — жизнь и *metron* — мера) введен в 1889 г. английским ученым Ф. Гальтоном, впервые применившим математический анализ при изучении явлений изменчивости и наследственности.

*Современная биометрия* — это особый раздел биологии и генетики, позволяющий использовать не только присущий им комплекс специфических методов исследований, но и в значительной мере опирающийся на различные математические приемы анализа, оценки и характеристики признаков и свойств живых объектов. Прежде всего, это относится к оценке продуктивных и племенных качеств животных, доли влияния на признак различных факторов, в том числе наследственности и среды, при вычислении величин, характеризующих связи между признаками, к оценке эффективности методов селекции и прогнозирования ее результатов, используя для этого сравнительно небольшие выборки, а выводы, полученные при их изучении, с достаточной точностью и надежностью распространить на большие по численности линии, семейства, стада, породы, виды.

Цель биометрического анализа: получение комплекса статистических параметров и коэффициентов, характеризующих изучаемую совокупность по одному или нескольким признакам.

### *Задачи биометрического анализа*

1. Выявить особенности и типы варьирования количественных и качественных признаков, характер распределения особей с разным уровнем признака.

2. Определить степень фенотипического уровня проявления признаков у особей в совокупности путем вычисления средних величин ( $\bar{X}$ ,  $G$ ,  $S$ ,  $H$ ,  $M_o$ ,  $M_e$ ).

3. Установить степень фенотипической и генотипической изменчивости признаков с помощью параметров  $\sigma$ ,  $C_v$ ,  $\sigma^2$ ,  $\lim$ ,  $t$ .

4. Изучить величину фенотипической и генотипической взаимосвязи между различными признаками и ее направление с использованием  $r$ ,  $R$ .

5. Сравнить группы по величине средних с применением метода статистических ошибок, критериев достоверности ( $t_d$ ,  $X^2$ ).

6. Определить долю влияния различных факторов на фенотипическую и генотипическую изменчивость признака с использованием дисперсионного и факторного анализа.

7. Дать характеристику изучаемой популяции по комплексу генетических и статистических параметров, степени гомо- и гетерозиготности, проявлению гетерозиса и инбредной депрессии, типу наследования признаков.

Предмет биометрического анализа — изменчивость изучаемого признака в совокупности.

*Источники информации.* Наиболее распространенным источником, характеризующим практически все массивы сельскохозяйственных животных, являются первичные данные зоотехнического, племенного и ветеринарного учета, которые систематически заполняются специалистами хозяйств по установленной форме в различных журналах, ведомостях, племенных карточках, базе данных ПЭВМ.

Другим источником информации служат данные научно-производственных опытов, проводимые на ограниченном числе животных, а также результаты специальных экспериментов на лабораторных животных при изучении различных биологических показателей. При этом информация записывается в протокольных книгах, дневниках наблюдений, ведомостях и специальных бланках.

Первичным материалом могут служить фотографии животных; микрофотографии клеток тканей, кариотипов; мазков крови, срезов тканей.

# 1. УСЛОВНЫЕ ОБОЗНАЧЕНИЯ ОСНОВНЫХ БИОМЕТРИЧЕСКИХ ПАРАМЕТРОВ

$v$	значения признака у отдельного объекта (варианта)
$N$	объем генеральной совокупности
$n$	объем выборочной совокупности
$\tilde{M}, \tilde{X}$	средняя арифметическая генеральной совокупности
$\bar{M}, \bar{X}$	средняя арифметическая выборочной совокупности
$\bar{M}_{636}, \bar{X}_{636}$	средняя взвешенная выборочной совокупности
$G$	средняя геометрическая
$S$	средняя квадратическая
$H$	средняя гармоническая
$M_o$	Мода
$M_e$	Медиана
$p, q$	доля — отношение числа объектов с признаком (а) к общему объему выборки
$\sigma$	среднее квадратическое (стандартное) отклонение
$\lim$	лимиты — указания минимального и максимального значения признака в совокупности
$C_v$	коэффициент вариации (изменчивости)
$C$	дисперсия – сумма квадратов центральных отклонений
$\sigma^2$	Варианса
$t$	нормированное отклонение
$r$	коэффициент корреляции
$R_{x/y}$	коэффициент регрессии
$r_w$	коэффициент повторяемости
$h^2$	коэффициент наследуемости
$\gamma_x^2$	квадрат корреляционного отношения, основной показатель силы влияния
$S_{\bar{X}}; S_{\sigma}; S_{Cv}$	статистическая ошибка с символом, соответствующим параметру
$t_{\bar{X}}; t_{\sigma}; t_{Cv}$	критерий достоверности безошибочности прогноза выборочного параметра
$t_d$	критерий достоверности разности между средними величинами двух выборок
$\nu$	число степеней свободы
$p$	уровень вероятности безошибочного прогноза
$\alpha$	уровень значимости
$\chi^2$	критерий соответствия распределений
$\Sigma$	знак суммирования

## 2. ПОНЯТИЕ О ГЕНЕРАЛЬНОЙ И ВЫБОРОЧНОЙ СОВОКУПНОСТИ

*Совокупностью* называется любое множество отличающихся друг от друга и в то же время сходных в некоторых существенных отношениях биологических объектов.

Наблюдения, проводимые над биологическими объектами, могут охватывать всех членов изучаемой совокупности, но могут и ограничиваться лишь некоторой ее частью. В первом случае достигается исчерпывающая и достоверная информация об объекте. Однако к этому методу прибегают не всегда из-за невозможности и нецелесообразности сопряженных со значительными затратами труда, средств и времени. Поэтому в подавляющем большинстве изучению подвергают какую-то часть совокупности выборку. Общая совокупность называется *генеральной* и представляет собой большой массив биологических объектов. Практически ее объем ограничен и может быть различным в зависимости от объекта наблюдений и от задачи, поставленной перед исследователем. Так весь крупный рогатый скот, обитающий на земном шаре, представляет собой генеральную совокупность. Если требуется изучить какую-то породу, то в этом качестве будут выступать все животные, входящие в нее, независимо от места нахождения. Структурные единицы породы (линии, семейства, типы) также могут рассматриваться как генеральная совокупность.

Под *выборочной совокупностью* подразумевается часть генеральной совокупности, предназначенной для ее характеристики. Объем выборки может быть сравнительно большим ( $n > 30 \div \infty$ ) и малым ( $n > 3 \div 30$ ). Это необходимо учитывать, так как приемы статистической обработки данных будут разными.

Для того чтобы по выборочной совокупности можно было судить о генеральной совокупности, она должна обладать следующими свойствами:

- *однородностью* (включение животных только одного пола, возраста, породы);
- *рендомичностью* (случайность, то есть любой член генеральной совокупности имеет одинаковую вероятность попасть в нее);
- *репрезентативностью* (отражать с заданной точностью и надежностью по параметрам выборочной совокупности аналогичных показателей в генеральной совокупности).

Все единицы совокупности нужно характеризовать с той степенью точности (в тех единицах измерения), которая подходит для данного конкретного признака.

### 3. СТАТИСТИЧЕСКИЙ АНАЛИЗ В МНОГОЧИСЛЕННЫХ ВЫБОРКАХ ( $n > 30$ )

#### 3.1. Вариационный ряд и вариационная кривая

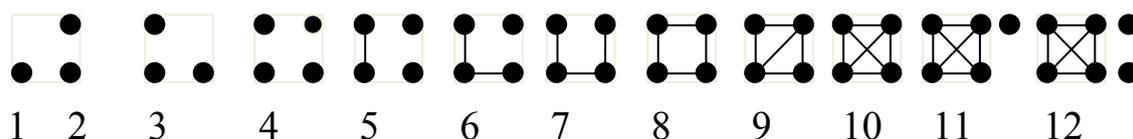
При характеристике количественных признаков и большом числе вариантов первичные данные необходимо упорядочить, то есть сгруппировать в определенном порядке. Допустим, что для характеристики свиноматок крупной белой породы по плодовитости на ферме зарегистрировано 100 опоросов, имеющих следующее количество поросят в помете:

9, 11, 15, 12, 11, 14, 11, 10, 11, 13, 8, 10, 8, 10, 11, 8, 11, 7, 9, 10, 9, 7, 8, 7, 9, 11, 12, 12, 12, 11, 12, 14, 13, 11, 15, 10, 8, 7, 8, 5, 9, 5, 10, 9, 6, 10, 10, 6, 13, 6, 10, 13, 12, 9, 6, 10, 11, 9, 9, 10, 12, 7, 8, 7, 11, 10, 8, 7, 9, 10, 11, 9, 9, 13, 15, 10, 7, 8, 10, 9, 9, 12, 14, 13, 9, 10, 11, 10, 7, 9, 10, 11, 13, 14, 8, 9, 8, 11, 8, 14.

Значения вариантов представляют собой ряд цифр. При внимательном рассмотрении нетрудно обнаружить, что каждая цифра встречается не один раз. В свою очередь это означает, что в данной выборке имеются животные с одинаковой плодовитостью. Возникает необходимость упорядочения вариантов, то есть расположить их в порядке возрастания. Для чего следует найти минимальное ( $\min$ ) и максимальное ( $\max$ ) значения признака. Наименьшее число поросят в помете — 5, а наибольшее — 15 голов.

Ранжируем выборку в пределах от 5 до 15 вариантов и находим их частоты. Вначале определяем величину классового промежутка, который обозначается латинской буквой  $k$ . В данном примере признак варьирует дискретно и величина классового промежутка равна 1 поросенку. Остается выяснить, сколько свиноматок дали 5, 6, 7, 8, 9, 10, 11, 12, 13, 14 и 15 поросят, т.е. разнести варианты по классам. Разноску вариант по классам следует делать, не выбирая из всей выборки свиноматок только с указанной плодовитостью, это трудоемко и нецелесообразно, а начинать с первой варианты, подряд, отмечая каждую из них с помощью условных знаков — точек и черточек.

Удобным оказывается шифр частот, представленный на рисунке 1.



*Рис. 1. Шифр частот*

Закончив разноску вариант по классам, шифр частот переводят в арабские цифры (табл. 1).

Таблица. 1. Техника разности вариант по классам

Класс		Шифр частоты	Частота $f$
№	Значение		
1	5	•	2
2	6	••	4
3	7	•• ┌───┐ │   │ └───┘	9
4	8	•• •• ┌───┐ │   │ └───┘	12
5	9	•• •• ┌───┐ │   │ └───┘	17
6	10	•• •• ┌───┐ │   │ └───┘	18
7	11	•• •• ┌───┐ │   │ └───┘	15
8	12	•• •• ┌───┐ │   │ └───┘	8
9	13	•• •• ┌───┐ │   │ └───┘	7
10	14	•• •• ┌───┐ │   │ └───┘	5
11	15	•	3

Таким образом, сгруппировав значения изучаемого признака в классы, получаем два ряда цифр (классы и частоты), представляющих собою вариационный ряд, который показывает, как изменяется признак от минимальной до максимальной величины, а также частоту вариант в каждом классе.

Оценка этой группы свиноматок по живой массе, кг: 197, 247, 207, 223, 215, 229, 232, 226, 228, 220, 226, 238, 236, 226, 225, 247, 195, 220, 252, 236, 220, 259, 240, 245, 219, 252, 256, 223, 245, 235, 256, 234, 228, 222, 262, 268, 219, 217, 224, 215, 259, 253, 247, 209, 222, 226, 230, 211, 220, 223, 257, 227, 274, 230, 266, 247, 258, 228, 254, 261, 259, 228, 247, 225, 226, 215, 211, 226, 226, 229, 247, 228, 238, 225, 232, 223, 220, 236, 247, 226, 222, 230, 229, 238, 236, 233, 230, 239, 228, 245, 215, 238, 220, 233, 219, 229, 217, 227, 223.

Составим вариационный ряд. Для этого вначале определим размах изменчивости — лимит ( $\lim$ ), то есть разницу между максимальным и минимальным значением признака. В нашем примере  $\max = 274$  кг,  $\min = 195$  кг,  $\lim = 274 - 195 = 79$  кг.

При этом величину классового промежутка  $k$  рассчитывают по формуле

$$k = \frac{\text{max} - \text{min}}{\text{желательное количество классов}}.$$

Количество классов в вариационном ряду зависит от степени изменчивости изучаемого признака и объема выборки  $n$ . Чем больше выборка, тем больше классов нужно брать при составлении вариационного ряда. Обычно придерживаются следующих границ, указанных в таблице 2.

Таблица 2. Определение количества классов в зависимости от объема выборки

Количество вариант в выборке	31-40	41-60	61-100	101-200	> 201
Количество классов	5-6	6-8	7-10	8-12	10-15

При этом величина классового промежутка должна быть как можно ближе к единице измерения изучаемого признака. В данном примере, если возьмем 7 классов, то величина  $k = 79/7 = 11,3$  кг, 8 классов соответственно  $79/8 = 9,9$  кг, 9 классов  $79/9 = 8,8$  кг, 10 классов  $79/10 = 7,9$  кг.

Последние две величины классового промежутка близки к единице измерения, то есть 1 кг. Следует помнить, что чем больше будет взято количество классов, тем точнее расчеты.

Отсюда берем величину  $k = 79/10 = 7,9$  кг, и ее можно округлить до 8 кг.

Далее нужно определить границы классов. Они должны быть удобными для разности вариант и обработки вариационного ряда. Нижней границей первого класса является минимальное значение признака (195 кг), а верхней — величина, большая на  $k$  ( $195 + 8 = 203$  кг). Затем определяем только верхние границы каждого последующего класса, постоянно прибавляя величину  $k$ , в последний класс войдет max значение признака. Разница между верхней границей предыдущего класса и нижней границей следующего класса равна единице измерения (1 кг). Вариационный ряд будет выглядеть следующим образом (табл. 3).

Таблица 3. Пример определения средней живой массы свиноматок ( $n = 100$ )

№	Класс		среднее значение, $W_i$	Частота $f$	Отклонение $a = \frac{w_i - A}{k}$	$f_a$	$f_a^2$
	граница						
	нижняя	верхняя					
1	195	203	199	2	-3	-6	18
2	204	211	207,5	4	-2	-8	16
3	212	219	215,5	9	-1	-9	9
4	220	227	$A = 223,5$	28	0	0	0
5	228	235	231,5	21	+1	21	21
6	236	243	239,5	10	+2	20	40
7	244	251	247,5	10	+3	30	90
8	252	259	255,5	11	+4	44	176
9	260	267	263,5	3	+5	15	75
10	268	275	271,5	2	+6	12	72
				$\Sigma = n = 100$		$\Sigma f_a = 119$	$\Sigma f_a^2 = 517$

Графическое изображение вариационного ряда представлено на рисунках 2, 3.

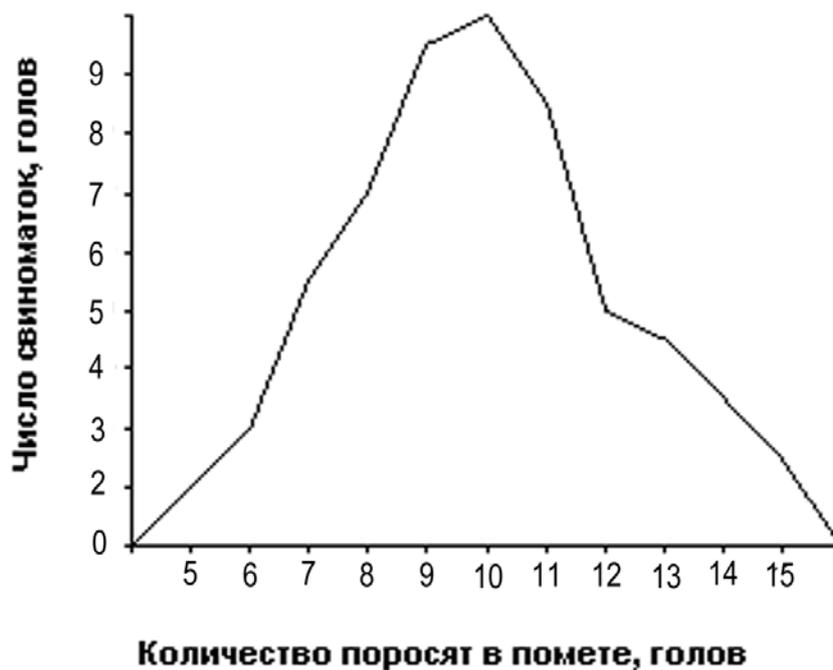
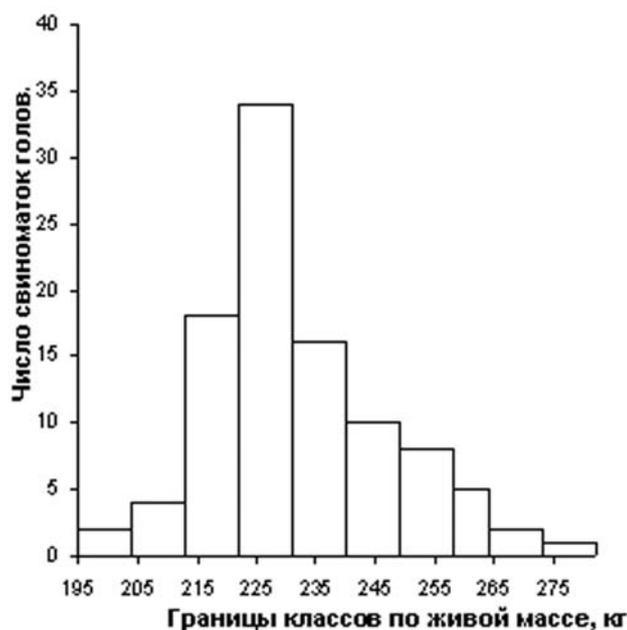


Рис. 2 Полигон распределения свиноматок по плодовитости



*Рис. 3. Гистограмма  
распределения свиноматок  
по живой массе*

Из гистограммы можно получить полигон, соединив середины верхних сторон столбиков линиями. В практике животноводства вариационную кривую можно использовать для выяснения результативности отбора.

### 3.2. Средняя арифметическая $\bar{X}$

Это основной параметр, который показывает, какое значение признака наиболее характерно в целом для какой-либо совокупности. Она используется для сравнения пород, стад, линий, семейств и других групп, а также характеризует средний фенотип и генотип популяции.

Все расчеты удобнее вести с помощью вариационного ряда. Рассмотрим это на примере анализа живой массы свиноматок, который был составлен ранее (см. табл. 3).

Среднюю арифметическую вычисляют по формуле

$$\bar{X} = A \pm kb,$$

где  $A$  — условная средняя;

$k$  — классовой промежуток;

$b$  — поправка к условной средней, которую определяем по формуле

$$b = \pm \frac{\sum fa}{n},$$

где  $f$  — число вариантов в классе;

$a$  — отклонение от условной средней;

$n$  — число вариантов в выборке.

Если она получается со знаком минус, то ее отнимают, если плюс, то прибавляют к условной средней. За условную среднюю обычно принимают среднее значение класса с наибольшей частотой вариант. В нашем примере взято среднее значение 4-го класса (223,5). Для наглядности его нужно выделить двумя жирными линиями.

Далее определяем условное отклонение для каждого класса. В четвертой строке вариационного ряда в шестой колонке ставят нуль:

$$a = \frac{w_i - A}{k} = \frac{223,5 - 223,5}{8} = 0.$$

Далее в сторону уменьшения средних значений классов ставим  $-1$ ,  $-2$ ,  $-3$ , а в сторону увеличения соответственно  $+1$ ,  $+2$ ,  $+3$  ...  $+6$ . Заполняем последующие колонки таблицы.

Вычислим поправку к условной средней:

$$b = \pm \frac{119}{100} = +1,19.$$

Средняя арифметическая величина именованная, т.е. выраженная в тех единицах измерения, что и признак, для которого она вычислена:

$$\bar{X} = 223,5 + 8 \cdot 1,19 = 234,2 \approx 234 \text{ кг.}$$

### 3.3. Среднее квадратическое (стандартное) отклонение $\sigma$

Среднее квадратическое (стандартное) отклонение  $\sigma$  — основной показатель, характеризующий степень изменчивости изучаемого признака, вычисляют по формуле:

$$\sigma = \pm k \sqrt{\frac{\sum fa^2}{n} - b^2} = \pm 9 \sqrt{\frac{517}{100} - 1,19^2} = \pm 9 \cdot 1,94 = 17,4 \approx 17 \text{ кг.}$$

Он применяется для построения кривых распределения признака в совокупности, определения интенсивности отбора, селекционного дифференциала и эффекта селекции. С его помощью можно выяснить границы вариационного ряда, если не известны лимиты. Установлено, что разнообразие признака располагается в следующих интервалах от средней арифметической (рис. 4):

$\bar{X} \pm 1\sigma$  — 68,3% всех вариант;

$\bar{X} \pm 2\sigma$  — 95,5%;

$\bar{X} \pm 3\sigma$  — 99,7%.

Следует учитывать, что в пределах  $\bar{X} \pm 0,67\sigma$  находится 50% особей всей совокупности. Это так называемые типичные особи, которые отличаются высокой приспособленностью к конкретным средовым факторам, что выражается в их более высоких воспроизводительных качествах, сохранности, долголетию и т.д.

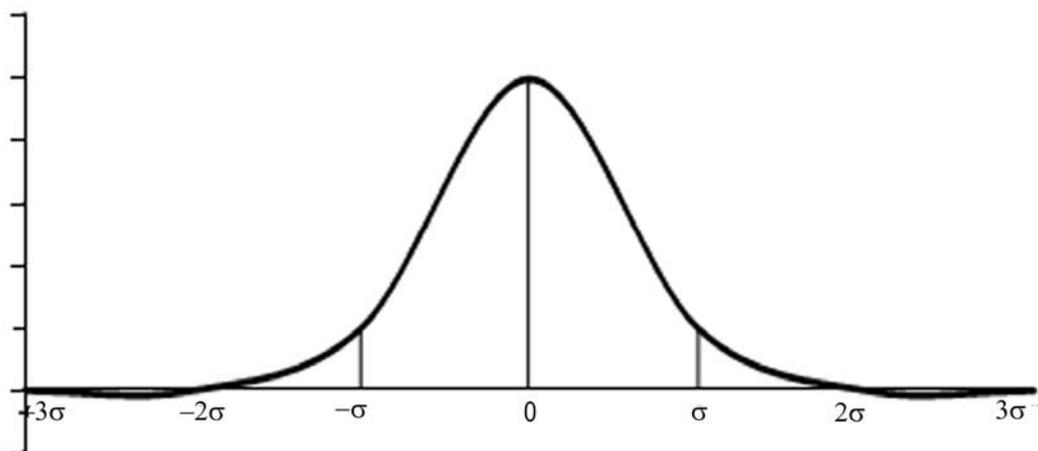


Рис. 4. Плотность нормального распределения

Если же отдельные варианты отклоняются от  $\bar{X}$  за  $\pm 3\sigma$ , то они не принадлежат данному вариационному ряду, а являются членами другой биологической совокупности и из анализа их необходимо исключить.

### 3.4. Коэффициент вариации $C_v$

Среднее квадратическое отклонение — величина именованная и выражается в определенных единицах измерения (л, кг, %, см, шт. и т.д.). Однако, чтобы имелась возможность сравнения степени изменчивости разных признаков, этот показатель нужно выразить в относительных единицах — в процентах к своей средней арифметической величине.

Такой параметр называется коэффициентом вариации или коэффициентом изменчивости. Формула для его вычисления имеет следующий вид:

$$C_v = \frac{\sigma}{\bar{X}} 100\%.$$

Коэффициентом изменчивости необходимо пользоваться и тогда, когда сравнивается степень варьирования одинаковых признаков разных совокупностей при наличии значительной разницы между средними этих групп. Так, если в первом стаде средний удой 2 000 кг,  $\sigma_1 = \pm 500$  кг, во втором, соответственно, 1 000 и 800 кг. Если судить по величине среднего квадратического отклонения, то большая изменчивость должна быть во втором стаде, т.к.  $\sigma_2 = 800$  кг против  $\sigma_1 = 500$  кг. Однако, вычислив  $C_v$ , мы увидим, что большая изменчивость будет в первом стаде:

$$C_{v_1} = \frac{500}{2000} 100\% = 25\%, \quad C_{v_2} = \frac{800}{4000} 100\% = 20\%.$$

Принято считать, что признаки со слабой изменчивостью характеризуются значением  $C_v$  не менее 10%, средней — 11-20%, высокой — более 21%. Коэффициентом изменчивости можно воспользоваться для определения объема выборки с целью получения достоверных результатов (приложение 1).

### 3.5. Взаимосвязь между признаками

Признаки и свойства животных находятся в определенной взаимосвязи. Наблюдениями установлено, что увеличение удоя сопровождается уменьшением жира в молоке; чем больше масса яйца, тем больше масса цыпленка; существует связь между устойчивостью и восприимчивостью матерей и их дочерей по ряду различных заболеваний и т.д. Особенностью такой связи является то, что каждому значению признака соответствует несколько значений другого. Из приведенных примеров видно, что корреляция может быть различной. Она бывает положительной и отрицательной. При положительной увеличение (уменьшение) одного признака сопровождается увеличением (уменьшением) второго, а при отрицательной — увеличение одного ведет к уменьшению другого.

Различным может быть не только характер, но и степень связи. Для оценки силы и направления взаимосвязи между признаками вычисляют коэффициент корреляции  $r$ . Он колеблется от 0 до  $\pm 1$ . При этом знак, стоящий перед цифровым значением  $r$ , свидетельствует о характере корреляции. Если  $r$  ниже 0,5, то связь считается слабой, при величине равной 0,51-0,7 — средней, если коэффициент более 0,71 — связь высокая. Следовательно, чем ближе значение  $r$  к единице, тем сильнее взаимосвязь; чем ближе к нулю — тем слабее связь между признаками.

#### *Вычисление коэффициента корреляции*

Вычислим величину корреляции между живой массой телок (кг) костромской породы в возрасте 12 мес. и величиной их удоя за 1 лактацию (кг). Воспользуемся данными, представленными в таблице 4.

Работа с выборкой идет в том же порядке, как это было при составлении вариационных рядов и вычисления его основных параметров. Но так как признака два, то удои обозначим буквой  $x$ , а живую массу —  $y$ . Определяем границы классов для каждого вариационного ряда. Потом составляем корреляционную решетку (табл. 5). Записываем вверху по горизонтали границы классов по живой массе  $y$  от минимального до максимального значения, а по вертикали аналогично располагаем границы по величине удоя  $x$ , т.е. расположим их взаимно перпендикулярно. Затем разносим варианты по классам, при этом учитываем одновременно сразу значения двух признаков. Чтобы показатели первой первотелки поместить в необходимую клеточку, нужно найти, в какой класс она попадет по величине удоя и в какой по живой массе, и на их пересечении ставим точку. Подсчитываем частоту вариант  $f_x$  и  $f_y$ , выделяем жирными линиями классы с условной средней  $A_x$  (4352-4591) и  $A_y$  (220-230).

Таблица 4. Исходные данные для вычисления коэффициента корреляции

| Удой, кг —<br>ж. масса, кг |
|----------------------------|----------------------------|----------------------------|----------------------------|----------------------------|
| 4 177 — 220                | 3 903 — 255                | 3 361 — 197                | 5 074 — 270                | 3 703 — 240                |
| 4 188 — 227                | 4 557 — 222                | 4 345 — 262                | 4 110 — 269                | 4 548 — 244                |
| 4 238 — 254                | 3 865 — 200                | 4 512 — 258                | 4 433 — 276                | 4 917 — 240                |
| 4 179 — 260                | 4 427 — 226                | 3 971 — 259                | 4 360 — 224                | 3 746 — 230                |
| 4 616 — 260                | 4 778 — 250                | 4 504 — 278                | 4 651 — 253                | 4 286 — 230                |
| 3 819 — 280                | 5 253 — 243                | 5 548 — 263                | 4 008 — 272                | 5 038 — 240                |
| 4 465 — 223                | 4 427 — 226                | 3 982 — 225                | 4 215 — 264                | 5 490 — 248                |
| 4 374 — 225                | 4 673 — 255                | 5 451 — 278                | 4 609 — 210                | 4 628 — 290                |
| 4 451 — 223                | 3 865 — 200                | 4 355 — 232                | 4 231 — 235                | 3 151 — 280                |
| 4 394 — 224                | 5 379 — 226                | 4 765 — 255                | 4 600 — 290                | 4 493 — 265                |
| 4 082 — 262                | 5 180 — 260                | 4 155 — 221                | 4 101 — 233                | 3 694 — 266                |
| 5 145 — 260                | 3 652 — 276                | 5 253 — 243                | 3 588 — 230                | 3 572 — 240                |
| 4 360 — 280                | 4 682 — 261                | 3 877 — 274                | 4 623 — 269                | 3 743 — 243                |
| 4 511 — 245                | 3 973 — 279                | 4 131 — 250                | 5 074 — 270                | 4 896 — 264                |
| 5 348 — 265                | 5 213 — 224                | 3 879 — 238                | 4 816 — 250                | 4 899 — 230                |
| 5 124 — 253                | 4 613 — 252                | 5 419 — 271                | 3 720 — 261                | 4 731 — 231                |
| 4 691 — 270                | 4 955 — 225                | 5 294 — 240                | 3 946 — 238                | 5 238 — 270                |
| 3 199 — 265                | 5 027 — 230                | 3 601 — 212                | 4 177 — 220                | 4 460 — 270                |
| 3 592 — 239                | 4 815 — 299                | 4 778 — 250                | 4 900 — 268                | 4 532 — 270                |
| 4 815 — 248                | 4 196 — 256                | 3 168 — 242                | 4 455 — 283                | 4 262 — 257                |

Таблица 5. Корреляционная решетка

$y \backslash x$	197	209	220	231	242	253	264	275	286	297	$f_x$	$a_x$	$f_x a_x$	$f_x a_x^2$
	208	219	230	241	252	263	274	285	296	307				
3151-3391	• 1				• 1		• 1	• 1			4	-5	-20	100
3392-3631		1	• 1	• 2				<b>II</b>			4	-4	-16	64
3632-3871	• 2		• 1	• 1	• 1	• 1	• 1	• 2			9	-3	-27	81
3872-4111	<b>I</b>		• 2	• 3		• 2	• 3	• 1			11	-2	-22	44
4112-4351			• 5	• 1	• 1	• 5	• 2				14	-1	-14	14
4352-4591			• 9	• 1	• 2	• 1	• 2	• 4			19	0	0	0
4592-4831		• 1		• 2	• 5	• 5	• 1		• 2	• 1	17	+1	17	17
4832-5071	<b>III</b>		• 3	• 1			• 2		<b>IV</b>		6	+2	12	24
5072-5311			• 1	• 1	• 2	• 3	• 3				10	+3	30	90
5312-5551			• 1		• 1	• 2	• 1	• 1			6	+4	24	96
$f_y$	3	2	23	12	13	19	16	9	2	1	100		$\sum f_x a_x = +16$	$\sum f_x a_x^2 = 530$
$a_y$	-2	-1	0	+1	+2	+3	+4	+5	+6	+7				
$f_y a_y$	-6	-2	0	12	26	57	64	45	12	7			$\sum f_y a_y = 215$	
$f_y a_y^2$	12	2	0	12	52	171	256	225	72	49			$\sum f_y a_y^2 = 851$	

$$b_x = \frac{\sum f_x a_x}{n} = -\frac{16}{100} = -0,16$$

$$b_y = \frac{\sum f_y a_y}{n} = \frac{215}{100} = +2,15.$$

В решетке образовались четыре квадранта. Далее определяем ряд условных отклонений  $a_x$  и  $a_y$ . Вычисляем произведение  $f_x a_x$  и  $f_y a_y$ ;  $f_x a_x^2$  и  $f_y a_y^2$  и их суммы.

Коэффициент корреляции  $r$  рассчитывается по формуле

$$r = \frac{\sum f a_x a_y - n b_x b_y}{n \sigma_x \sigma_y},$$

где  $\sum f_x a_x$  — сумма произведений числа вариантов на отклонение по ряду  $x$  и ряду  $y$  одновременно.

Остальные символы в формуле известны. Однако среднее квадратическое отклонение берется не в именованных величинах, а в количестве классовых промежутков, т.е. цифра, полученная после извлечения квадратного корня на  $k$  (величина классового промежутка) не умножается и формула этого параметра имеет следующий вид:

$$\sigma = \pm \sqrt{\frac{\sum f a^2}{n} - b^2}.$$

Чтобы вычислить величину, нужно обратиться к корреляционной решетке, в которой как мы выяснили образовались четыре квадрата, их необходимо пронумеровать римскими цифрами. В каждом из них необходимо найти этот показатель. При этом частоту  $f$  в каждой клеточке умножаем на  $a_x$  и  $a_y$  одновременно.

I квадрант	II квадрант	III квадрант	IV квадрант
$1(-5)(-2) = 10$	$1(-5)2 = -10$	$1(-1)1 = -1$	$2 \cdot 1 \cdot 1 = 2$
$1(-4)(-1) = 4$	$1(-5)4 = -20$	$\Sigma = -1$	$5 \cdot 1 \cdot 2 = 10$
$2(-3)(-2) = 12$	$1(-5)5 = -25$		$2 \cdot 1 \cdot 3 = 6$
$\Sigma = +26$	$2(-4)1 = -8$		$1 \cdot 1 \cdot 4 = 4$
	$1(-3)1 = -3$		$2 \cdot 1 \cdot 6 = 12$
	$1(-3)2 = -6$		$1 \cdot 1 \cdot 7 = 7$
	$1(-3)3 = -9$		$1 \cdot 2 \cdot 1 = 2$
	$1(-3)4 = -12$		$2 \cdot 2 \cdot 4 = 16$
	$2(-3)5 = -30$		$1 \cdot 3 \cdot 1 = 3$
	$3(-2)1 = -6$		$2 \cdot 3 \cdot 2 = 12$
	$2(-2)3 = -12$		$3 \cdot 3 \cdot 3 = 27$
	$3(-2)4 = -24$		$3 \cdot 3 \cdot 4 = 36$
	$1(-2)5 = -10$		$1 \cdot 2 \cdot 3 = 6$
	$1(-1)1 = -1$		$1 \cdot 4 \cdot 4 = 16$
	$1(-1)2 = -2$		$1 \cdot 4 \cdot 5 = 20$
	$2(-1)4 = -8$		$\Sigma = +213$
	$\Sigma = -201$		

Сумма произведений отклонений числа вариантов по ряду  $x$  и ряду  $y$  в первом и четвертом квадратах положительная, во втором и третьем — отрицательная. Таким образом, мы получили

$$\sum f a_x a_y = (+26) + (-201) + (-1) + (+213) = +37.$$

Все необходимое для вычисления  $r$  известно, находим его значение:

$$r = \frac{+37 - 100 \cdot (-0,16) \cdot 2 \cdot 15}{100 \cdot 2,29 \cdot 1,97} = \frac{+71,4}{451,13} = +0,16.$$

Следовательно, между удоем первотелок и их живой массой в возрасте 12 мес. существует слабая положительная связь. Так как коэффициент корреляции вычислен не по генеральной, а выборочной совокупности, он имеет свою статистическую ошибку:

$$S_r = \pm \frac{1 - r^2}{\sqrt{n}} = \pm \frac{1 - 0,16^2}{\sqrt{100}} = \pm 0,09.$$

Далее определяем степень достоверности  $r$ :

$$t_r = \frac{r}{S_r} = \frac{0,16}{0,09} = 1,78.$$

При числе степеней свободы  $\nu = n - 2 = 100 - 2 = 98$  находим  $t_{st}$  (приложение 2), они равны  $1,98 \div 2,63 \div 3,39$ . Так как найденное значение  $t_r$  меньше  $t_{st}$ , то не представляется возможным утверждать, что у исследованных коров костромской породы существует слабая положительная связь между их удоем за 1-ю лактацию и живой массой в возрасте одного года, хотя тенденция к этому имеется.

### *Вычисление коэффициента регрессии*

При установлении связи между признаками часто возникает необходимость выяснить насколько увеличится один из признаков, если второй, связанный с ним, изменится на единицу измерения. К чему может привести односторонний отбор только по одному признаку? Это можно выяснить путем вычисления коэффициента регрессии  $R$  по следующим формулам:

$$R_{x/y} = r \frac{\sigma_x}{\sigma_y} = 0,16 \frac{2,26 \cdot 240}{1,97 \cdot 11} = 4,00 \text{ кг},$$

$$R_{y/x} = r \frac{\sigma_y}{\sigma_x} = 0,16 \frac{1,97 \cdot 11}{2,26 \cdot 240} = 0,01 \text{ кг}.$$

При этом  $\sigma$  вычисляется с поправкой на величину классового промежутка. Следовательно, при увеличении живой массы первотелок в возрасте 12 мес. на 1 кг, их удои за 1-ю лактацию увеличатся на 4 кг, а при увеличении удоя на 1 кг, живая масса в анализируемом возрастном периоде увеличится на 0,01 кг (100 грамм).

Коэффициенты корреляции и регрессии вычисляются не только для установления связи между двумя варьирующими признаками одного и того же организма. Они используются для вычисления коэффициентов наследуемости  $h^2$  (доля генетической изменчивости в общей фенотипической) и повторяемости  $r_w$  (степень соответствия оценок признака за разные промежутки времени). При этом  $r$  и  $R$  вычисляются между одноименными признаками в двух смежных поколениях (например, удой матерей — удой дочерей и т.д.). Техника их вычисления такая же как и при расчете коэффициента корреляции. При этом можно воспользоваться простейшими формулами:

$$h^2 = 2r_{M-D}; \quad h^2 = 2R_{M-D}; \quad r_M = r.$$

### 3.6. Оценка достоверности статистических показателей

#### *Определение необходимого объема совокупности*

В основу расчетов необходимого объема выборки положены такие параметры, как  $\bar{X}$ ,  $C_v$ ,  $\sigma$ , при этом можно пользоваться как научными литературными данными, так и личными наблюдениями. Наиболее простой способ — это использование коэффициента изменчивости, тогда достоверный объем выборки можно рассчитать по формуле

$$n = \frac{C_v^2 t^2}{\varepsilon^2},$$

где  $C_v$  — коэффициент вариации;

$t$  — критерий достоверности (см. приложение 1);

$\varepsilon$  — допустимая погрешность (1, 3 или 5%).

#### *Уровни вероятности и значимости*

Всего выделяют три уровня вероятности безошибочного прогноза, обозначенные  $p_1 \geq 0,95$ ;  $p_2 \geq 0,99$  и  $p_3 \geq 0,999$ , которым соответствуют три уровня значимости  $\alpha \leq 0,05$ ;  $\alpha \leq 0,01$  и  $\alpha \leq 0,001$ . Уровень значимости обозначает вероятность получения ошибки от установленных с определенной вероятностью результатов. С его помощью можно установить, в каком проценте случаев возможна ошибка в результатах. Если уровень значимости  $p \leq 0,05$  (5%-ный), то это значит, что в силу случайности ошибка будет в 5% случаев. Вероятность  $p \geq 0,95$  показывает, что из 100 повторений в 95 будут получены ожидаемые результаты.

### Статистические ошибки (S)

В практической работе основные параметры совокупности вычисляют не по генеральной совокупности, а по части ее — выборке. Вследствие этого возникают ошибки, называемые статистическими, которые представляют собой величину, на которую параметр выборочной совокупности отличается от одноименного в генеральной совокупности. Для каждого параметра вычисляется своя статистическая ошибка. Для нашего примера они равны:

– статистическая ошибка для средней арифметической:

$$S_{\bar{x}} = \pm \frac{\sigma}{\sqrt{n}} = \pm \frac{17}{\sqrt{100}} = 1,7 \approx 2 \text{ кг};$$

– статистическая ошибка для среднего квадратического отклонения:

$$S_{\sigma} = \pm \frac{\sigma}{\sqrt{2n}} = \pm \frac{17}{\sqrt{200}} = 1,2 \approx 1 \text{ кг};$$

– статистическая ошибка для коэффициента вариации:

$$S_{C_v} = \pm \frac{C_v}{\sqrt{2n}} = \pm \frac{7}{\sqrt{200}} = 0,49 \approx 0,5\%.$$

Имея значения параметра и получив его статистическую ошибку, мы можем определить границы, в которых будет находиться данный параметр в генеральной совокупности. Статистики подсчитали, что он может находиться в пределах  $\pm 3m_x$  от выборочного. В данном примере это представлено следующим образом:

$$\bar{X} = 234 - 3 \cdot 2 = 228,$$

т.е. генеральная средняя находится в  $234 + 3 \cdot 2 = 240$  пределах от 228 до 240 кг.

### Ошибки выборочной разности

В практической работе и исследованиях нередко приходится встречаться с необходимостью сравнения двух групп (совокупностей) по ряду показателей. Для оценки достоверности разности между средними арифметическими двух выборок применяется критерий достоверности  $t_d$ , который вычисляется по формуле

$$t_d = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{S_{x_1}^2 + S_{x_2}^2}} \geq t_{st},$$

где  $\bar{X}_1$  и  $\bar{X}_2$  — средние арифметические;

$S_{x_1}$  и  $S_{x_2}$  — их статистические ошибки;

$t_{st}$  — стандартное значение критерия, определяемого по таблице Стьюдента (см. приложение 2) с учетом числа степеней свободы ( $v = n_1 + n_2 - 2$ ) для трех уровней вероятности безошибочного прогноза;  $n_1$  и  $n_2$  — численность сравниваемых групп.

*Например.* Вычислить критерий достоверности разности по живой массе баранов породы прекос разных конституциональных типов ( $n = 200$  гол. в каждой группе), шерстно-мясной  $77 \pm 2,2$  кг и мясной тип  $85 \pm 2,6$  кг:

$$t_{d_{AC}} = \frac{85 - 77}{\sqrt{2,2^2 + 2,6^2}} = 2,23,$$

$$v = 200 + 200 - 2 = 398, \quad P \geq 0,95.$$

Следовательно, бараны мясного типа достоверно превосходят по живой массе своих сверстников шерстно-мясного типа.

### *Критерий ХИ-КВАДРАТ ( $\chi^2$ )*

Этот показатель показывает степень отличия наблюдаемых показателей от тех, которые должны были бы получаться в соответствии с существующими закономерностями. Он используется при сравнении достоверности разности между двумя и более группами, для изучения влияния определенных факторов в изменчивости признаков. Критерий соответствия  $\chi^2$  вычисляется по формуле

$$\chi^2 = \sum \frac{(O - E)^2}{E},$$

где  $O$  — фактически наблюдаемое число;

$E$  — теоретически ожидаемое число.

*1-й метод.* Допустим от 150 овцематок, при условии получения одинцов, получено 150 ягнят, в т.ч. 80 баранчиков и 70 ярочек. Соответствует ли это теоретически ожидаемому 1:1 (75 баранчиков и 75 ярочек). Составим для удобства таблицу 6.

*Таблица 6. Исходные данные*

Классы	Распределение	
	фактическое	теоретическое
Баранчики	80	75
Ярочки	70	75

Разница между фактическим и теоретически ожидаемым количеством баранчиков  $O - E = 80 - 75 = 5$ ,  $(O - E)^2 = 5^2 = 25$ ; ярочек соответственно  $70 - 75 = -5^2 = 25$ .

Тогда

$$\chi^2 = \frac{25}{150} + \frac{25}{150} = \frac{50}{150} = 0,33.$$

Обратимся к таблице стандартных значений  $\chi^2$  (приложение 3), учитывая при этом число степеней свободы  $v = 2 - 1 = 1$ . Полученное значение  $\chi^2 = 0,33$  значительно ниже стандартного — 3,84, что свидетельствует о том, что наблюдаемое соотношение количества баранчиков и ярочек соответствует теоретическому ожиданию.

2-й метод. Имеются две линии мышей: высоколейкозная (АКР) и низколейкозная (СС57). В обычных условиях из 44 мышей АКР выжило 22 и пало 22, а в линии СС57 выжило 29, а пало 6. Необходимо оценить методом  $\chi^2$  различие между линиями. Составим таблицу 7.

Таблица 7. Исходные данные

Линия	Выжило	Пало	Всего
АКР	$22 - a$	$22 - b$	$44 - a - b$
СС57	$29 - c$	$6 - d$	$35 - c - d$
Всего	$51 - a - c$	$28 - b - d$	$n = 79$

Рассчитаем критерий  $\chi^2$ :

$$\begin{aligned}\chi^2 &= \frac{(ad - bc)^2 n}{(a + b)(c + d)(a + c)(b + d)} = \\ &= \frac{(22 \cdot 6 - 22 \cdot 29)^2 \cdot 79}{(22 + 29)(29 + 6)(22 + 29)(22 + 6)} = 9,2.\end{aligned}$$

Число степеней свободы  $\nu = (2 - 1)(2 - 1) = 1$ . Табличное значение  $\chi^2$  при  $\nu = 1$  составляет  $3,8 \div 6,6 \div 10,8$  (приложение 3), т.к. фактическая величина (9,2) больше табличного значения 6,6, но меньше, чем 10,8, нулевая гипотеза отвергается с высоким уровнем вероятности  $P \geq 0,99$ . Значит, между линиями существуют достоверные различия.

#### 4. СТАТИСТИЧЕСКОЙ АНАЛИЗ В МАЛОЧИСЛЕННЫХ ВЫБОРКАХ ( $n = 3 \div 30$ )

В таких совокупностях значения вариант не группируются в классы, а вычисление статистических параметров удобнее вести с использованием метода условной средней. За ее величину может быть взято значение любой или наиболее часто встречающейся варианты.

Например: имеются показатели живой массы ягнят при рождении: 3,0; 4,5; 3,0; 3,5; 4,2; 4,6; 5,0; 3,0; 3,8; 3,2 кг. Значение вариант удобнее располагать в столбик, во втором следует показать отклонение ( $\Delta$ ), в третьем квадрат отклонения значения каждой варианты от условной средней ( $\Delta^2$ ), которая выделяется рамкой (табл. 8).

*Таблица 8. Данные для вычисления основных статистических показателей*

№	Значение вариант $v$	Отклонение $\Delta = v - A$	$\Delta^2$
1	3,0	0	0
2	4,5	+1,5	2,25
3	<b>3,0</b>	0	0
4	3,5	+0,5	1,25
5	4,2	+1,2	1,44
6	4,6	+1,6	2,56
7	5,0	+2,0	4,0
8	3,0	0	0
9	3,8	+0,8	0,64
10	3,2	+0,2	0,04
	$\Sigma v = 37,8$	$\Sigma \Delta = +7,8$	$\Sigma \Delta^2 = 11,18$

Средняя арифметическая:

$$\bar{x} = A \pm \frac{\Sigma \Delta}{n},$$

где  $A$  — условная средняя;

$\Sigma \Delta$  — поправка к условной средней, если она со знаком (+), то прибавляется, если (-), то вычитается.

Среднее квадратическое отклонение:

$$\sigma = \pm \sqrt{\frac{\Sigma \Delta^2}{n-1}} = \pm \sqrt{\frac{11,18}{9}} = \pm 1,1 \text{ кг.}$$

Коэффициент вариации:

$$C_v = \frac{\sigma}{\bar{x}} 100\% = \frac{1,1}{3,8} 100\% = 28,9\%.$$

Статистическая ошибка средней арифметической:

$$S_{\bar{x}} = \pm \frac{\sigma}{\sqrt{n-1}} = \pm \frac{1,1}{\sqrt{9}} = 0,37.$$

*Коэффициент корреляции*

Он вычисляется по формуле

$$r = \frac{\alpha_1 + \alpha_2}{2\sqrt{\alpha_1\alpha_2}},$$

где  $\alpha$  — сумма квадратов центральных отклонений ( $\alpha_1$  — для первого признака,  $\alpha_2$  — для второго):

$$\alpha = \Sigma \Delta^2 = \frac{(\Sigma \Delta)^2}{n},$$

где  $\Delta$  — отклонение вариант от условной средней ( $\Delta_1$  — для первого признака,  $\Delta_2$  — для второго).

Находим разницу:

$$d = \Delta_1 - \Delta_2 \text{ или } d = v_1 - v_2.$$

*Пример.* Определить направление и величину связи между содержанием жира  $v_1$  и белка  $v_2$  в молоке коров костромской породы (табл. 9).

*Таблица 9. Данные для вычисления фенотипической корреляции*

№	Содержание		$\Delta_1$	$\Delta_2$	$\Delta_1^2$	$\Delta_2^2$	$d$	$d^2$
	жира $v_1$	белка $v_2$						
1	<u>4,5</u>	3,6	0	+0,1	0	0,01	+0,9	0,81
2	4,0	3,2	-0,5	-0,3	0,25	0,09	+0,8	0,64
3	4,6	<u>3,5</u>	+0,1	0	0,01	0	+1,1	1,21
4	5,0	3,6	+0,5	+0,1	0,25	0,01	+1,4	1,96
5	4,0	3,3	-0,5	-0,2	0,25	0,04	+0,7	0,49
6	4,1	3,3	-0,4	-0,2	0,16	0,04	+0,8	0,64
7	4,5	3,5	0	0	0	0	+1,0	1,00
8	4,8	3,6	+0,3	+0,1	0,09	0,01	+1,2	1,44
9	4,9	3,7	+0,4	+0,2	0,16	0,04	+1,2	1,44
10	5,2	3,8	+0,7	+0,3	0,49	0,04	+1,4	1,96
			$\Sigma \Delta_1 =$ +0,6	$\Sigma \Delta_2 =$ +0,1	$\Sigma \Delta_1^2 =$ 1,66	$\Sigma \Delta_2^2 =$ 0,33	$\Sigma d =$ 10,5	$\Sigma d^2 =$ 11,59

Отсюда условная средняя по каждому признаку при  $n = 10$  гол.:

$$A_1 = 4,5\%; \quad A_2 = 3,5\%.$$

Полученные расчетные данные занесем в таблицу 10.

Таблица 10. Расчетные данные

Показатель	Признак		<i>d</i>
	1	2	
$\Sigma\Delta$	0,6	0,1	10,5
$(\Sigma\Delta)^2$	0,36	0,01	110,25
$(\Sigma\Delta)^2/n$	0,036	0,001	11,025
$\Sigma\Delta^2$	1,66	0,33	11,59
$\alpha$	1,624	0,329	0,565

Вычисляем коэффициент корреляции  $r$ :

$$r = \frac{1,624 + 0,329 - 0,565}{2\sqrt{1,164 \cdot 0,329}} = \frac{1,36}{1,75} = 0,937 \approx +0,94.$$

Статистическая ошибка коэффициента корреляции  $S_r$ :

$$S_r = \pm \sqrt{\frac{1-r^2}{n-2}} = \pm \sqrt{\frac{1-(0,94)^2}{8}} = 0,12.$$

Критерий достоверности коэффициента корреляции:

$$t_r = \frac{r}{S_r} = \frac{0,94}{0,12} = 0,83.$$

По таблице Стьюдента находим, что  $t_r = 7,83$  при  $\nu = 8$  превышает стандартное значение при вероятности  $P \geq 0,999$ , что свидетельствует о статистической достоверности коэффициента корреляции. Значит между содержанием жира и белка в молоке у коров костромской породы существует высокая, устойчивая, положительная связь. Следовательно, с увеличением содержания жира в молоке увеличивается и содержание белка.

## 5. СТАТИСТИЧЕСКИЙ АНАЛИЗ ПРИЗНАКОВ С АЛЬТЕРНАТИВНОЙ ИЗМЕНЧИВОСТЬЮ

Признаки, имеющие только два возможных взаимоисключающих состояния, называются альтернативными (пол мужской – женский, здоровые – больные животные и др.). Средняя арифметическая для альтернативных признаков отражает долю или процент особей, имеющих  $p$  и не имеющих данный признак  $q$  в совокупности.

*Пример.* В стаде из 1 000 коров заболело маститом 50 голов, а 950 остались здоровыми. Совокупность состоит из двух групп: первая — больные, вторая — здоровые животные. Обозначим численность I группы как  $P_1$ , соответственно второй —  $P_0$ , а общее число —  $n$ .

Доля больных животных (т.е. имеющих данный признак) определяется по формуле

$$p = \frac{P_1}{n} = \frac{50}{1000} = 0,05, \text{ или } 5\%.$$

Доля здоровых животных (т.е. без данного признака) составляет

$$p = \frac{P_0}{n} = \frac{950}{1000} = 0,95, \text{ или } 95\%.$$

Среднее квадратическое отклонение вычисляют по формуле

$$\sigma = \sqrt{pq} = \sqrt{0,95 \cdot 0,05} = 0,218, \text{ или } 21,8\%.$$

Величина статистической ошибки доли величин одинакова как для доли больных, так и для доли здоровых животных:

$$S_p = S_q = \sqrt{\frac{pq}{n}} = \sqrt{\frac{0,05 \cdot 0,95}{1000}} = 0,007 \text{ или } 0,7\%.$$

Вычисление достоверности разности между выборочными долями или процентами:

$$t_d = \frac{P_x - P_y}{\sqrt{S_x^2 + S_y^2}} \geq t_{st}.$$

*Пример.* В семействе коровы Красы 998 учтено в трех поколениях 40 потомков, и у 20 из них выявлен лейкоз, а из 36 потомков семейства коровы Симпатии 200 — 9 голов. Необходимо установить, различаются ли эти два семейства по восприимчивости к лейкозу.

Определим долю больных потомков:

$$P_x = 20 / 40 = 0,5 \text{ и } P_y = 9 / 36 = 0,25.$$

Величины статистических ошибок соответственно будут равны:

$$S_x = \sqrt{\frac{0,5(1-0,5)}{40}} = \sqrt{\frac{0,5 \cdot 0,5}{40}} = \sqrt{0,00625} = 0,08,$$
$$S_y = \sqrt{\frac{0,25(1-0,25)}{36}} = \sqrt{\frac{0,25 \cdot 0,75}{36}} = \sqrt{0,00521} = 0,07.$$

Полученные данные подставляем в формулу:

$$t_d = \frac{0,5 - 0,25}{\sqrt{0,08^2 + 0,07^2}} = \frac{0,25}{\sqrt{0,0113}} = \frac{0,25}{0,106} = 2,36 \quad (p \geq 0,95).$$

По таблице Стьюдента с учетом числа степеней свободы

$$\nu = n_x + n_y - 2 = 40 + 36 - 2 = 74$$

находим  $t_{st} = 1,98 \div 2,63 \div 3,39$ . Так как  $t_d = 2,36$ , это соответствует первому уровню вероятности безошибочного прогноза. Следовательно, разница заболеваемости лейкозом между потомками в анализируемых семействах достоверна. Это значит, что потомки в семействе Красы 200 отличаются большей восприимчивостью к лейкозу

*Коэффициент корреляции  
для альтернативных признаков*

Для его определения первичные данные размещают в четырехпольной корреляционной решетке. Коэффициент корреляции  $r_a$  вычисляют по формуле

$$r_a = \frac{(P_1 P_2 - P_2 P_3) - \frac{n}{2}}{\sqrt{(P_1 + P_2)(P_3 + P_4)(P_1 + P_3)(P_2 + P_4)}}$$

где  $P_1, P_2, P_3, P_4$  — частоты, распределившиеся в четырех клетках корреляционной решетки.

*Пример.* Определить величину связи между резистентностью цыплят к пуллорозу (белому поносу) и степенью их отселекционированности по данному признаку (табл. 11).

*Таблица 11. Исходные данные*

Группа цыплят	После заражения живой культурой		Итого
	выжило	пало	
Исходная отселекционированная	115 ( $P_1$ ) 560 ( $P_3$ )	105 ( $P_2$ ) 58 ( $P_4$ )	$P_1 + P_2 = 220$ $P_3 + P_4 = 618$
Итого	$P_1 + P_3 = 675$	$P_2 + P_4 = 163$	$n = 838$

Отсюда коэффициент корреляции равен

$$r_a = \frac{(115 \cdot 58 - 560 \cdot 105) - \frac{838}{2}}{\sqrt{220 \cdot 618 \cdot 675 \cdot 163}} = \frac{-52549}{122306} = -0,43.$$

Статистическая ошибка коэффициента корреляции:

$$S_r = \frac{1 - r_a^2}{\sqrt{n}} = \frac{1 - 0,43^2}{\sqrt{838}} = 0,006.$$

Достоверность коэффициента корреляции:

$$t_r = \frac{r}{m_r} = \frac{0,43}{0,006} = 71,67.$$

Так как  $t_r$  значительно выше  $t_{st} = 3,29$ , то коэффициент корреляции высокодостоверен ( $p \geq 0,999$ ).

Следовательно, с увеличением интенсивности селекции снижается заболевание цыплят пуллорозом.

## 6. ОСНОВЫ ДИСПЕРСИОННОГО АНАЛИЗА

На формирование признаков у сельскохозяйственных животных влияют генетические факторы, факторы внешней среды, внутренние факторы организма, методы селекции. Они вызывают варьирование изучаемого признака как в сторону увеличения, так и в сторону уменьшения. Степень и направленность воздействия различных факторов неодинакова, поэтому исключительно важно определить долю влияния каждого из них на изменчивость признака. Это можно выяснить с помощью дисперсионного анализа, в основу которого положен математический метод вычисления дисперсии или изменчивости, выраженной через  $\sigma^2$ , при этом в анализе используют общую дисперсию по всей выборке  $\sigma_y^2$ , факториальную  $\sigma_x^2$  и остаточную  $\sigma_z^2$  (табл. 12, 13).

Таблица 12. Техника расчета при дисперсионном анализе однофакторного комплекса для малых групп

Показатели варьирующего признака	Градация			Свободные показатели. Число градации $q = 3$
	Салат	Каро	Ладок	
Значение признака $v$	36, 39, 34	40, 38, 28	40, 40, 36	$\sum v = 331$ (суммируем по градациям)
$v^2$	$36^2 + 39^2 + 34^2 =$ $= 3\ 973$	$40^2 + 38^2 + 28^2 =$ $= 3\ 828$	$40^2 + 40^2 + 36^2 =$ $= 4\ 496$	$\sum v^2 = 12\ 297$ (Возводят в квадрат каждое значение $v$ и суммируют по строчке)
Число голов $n_x$	3	3	3	$n = 9$ (Суммируют по градациям)
Объем градации $\sum v_x$	$36 + 39 + 34 =$ $= 109$	$40 + 38 + 28 =$ $= 106$	$40 + 40 + 36 =$ $= 116$	
$(\sum v_x)^2$	$(109)^2 = 11\ 881$	$(106)^2 = 11\ 235$	$(116)^2 = 13\ 445$	
$h = (\sum v_x)^2/n_x$	$\frac{11\ 881}{3} = 3\ 960$	$\frac{11\ 235}{3} = 3\ 745$	$\frac{13\ 445}{3} = 4\ 485$	$\sum h = 12\ 190$ (суммируют по строчке)
Частное среднее града- ции $M_x = \sum v_x/n$	$\frac{109}{3} = 36$	$\frac{106}{3} = 35$	$\frac{116}{3} = 39$	

Определяем вспомогательную величину:

$$H = (\sum v)^2/n = (331)^2/9 = 12\,173,$$

$$C_y = \sum v^2 - H = 12\,297 - 12\,173 = 124,$$

$$C_z = \sum v^2 - \sum h = 12\,297 - 12\,190 = 107,$$

$$C_x = \sum h - H = 12\,190 - 12\,173 = 17,$$

$$C_y = C_z + C_x = 107 + 17 = 124.$$

Таблица 13. Свободная таблица дисперсионного комплекса

Показатель	Разнообразие		
	фактическое x	остаточное y	общее z
Дисперсии C	17	107	127
Число степеней свободы v	$q - 1 = 3 - 1 = 2$	$n - q = 9 - 3 = 6$	$n - 1 = 9 - 1 = 8$
Варианса $\sigma^2$	$\frac{C_x}{V_x} = \frac{17}{2} = 8,5$	$\frac{C_z}{V_z} = \frac{107}{6} = 15,8$	$\frac{C_y}{V_y} = \frac{124}{8} = 15,5$
Степень влияния	$\frac{C_x}{C_y} = \frac{17}{124} = 0,12$	$\frac{C_z}{C_y} = \frac{107}{124} = 0,9$	

Показатель достоверности влияния:

$$F_{эмн} = \frac{\sigma_x^2}{\sigma_z^2} = \frac{8,5}{15,8} = 0,5,$$

$$F_{ст} = \frac{v_1}{v_2} = \frac{2}{6} = 5,1 \div 27,0 \text{ (см. приложение 3).}$$

**Вывод:** влияние генетического фактора (быков-производителей) равно 12%, а влияние паратипических факторов 88%.

Дисперсионный анализ позволяет проводить изменение степени наследственности признака, сравнение наследственной и комбинативной способности лидерных животных, решает задачи по определению величины влияния генетических и паратипических факторов на признак, прогнозировать развитие признака при заданном или имеющемся комплексе условий или симптомов.

## 7. ЗАДАНИЯ ДЛЯ САМОСТОЯТЕЛЬНОЙ РАБОТЫ

*Задача 1.* По имеющимся данным (табл. 14) составьте вариационный ряд и вычислите основные статистические константы.

*Таблица. 14. Удой первотелок костромской породы, кг*

4376	3979	5516	4526	4566	5143	4555	5226	4958	6214
4153	6490	5936	5563	4795	4132	5530	4359	4781	5557
4612	6390	6269	6389	4674	5485	5973	5240	5800	6402
4332	4543	5092	5322	4338	5362	6746	3884	5656	5903
4279	6627	4623	5037	4320	4061	3813	5684	3350	6330
5013	4537	4793	5390	5346	5312	5654	6435	6091	5024
3694	4812	4417	4455	3027	5020	4993	3849	4008	4434
3735	5485	5020	3708	5869	5646	5135	4041	5223	5668
5704	5574	4010	4250	4948	5220	5291	3590	3645	4668
067	5740	5230	3605	3325	5846	4557	4510	4409	4543
4047	6336	4478	4631	5433	5331	4507	4896	5328	5668
4011	5105	5161	5151	4543	6103	5113	4516	3813	5483
4025									

*Задача 2.* По имеющимся данным (табл. 15) составьте вариационный ряд и вычислите основные статистические константы.

*Таблица. 15. Живая масса кур в возрасте 18 месяцев, кг*

2,35	1,85	1,80	1,75	1,80	2,30	1,70	1,90	1,80	1,80	1,45
1,95	1,95	2,10	1,75	1,75	2,20	1,75	1,65	1,95	2,00	1,85
1,70	1,65	2,0	1,85	1,80	2,35	1,55	2,20	1,75	1,20	1,60
1,70	1,90	1,70	1,95	1,75	1,95	1,80	1,65	1,80	2,00	1,80
2,00	1,75	1,90	1,95	1,70	1,65	1,70	1,70	1,85	1,35	1,55
2,05	1,90	1,55	1,60	2,00	1,70	1,85	1,85	1,85	230	1,80
1,70	1,60	1,90	2,60	2,00	1,25	1,75	2,10	1,40	1,65	2,15
2,10	1,85	1,95	1,75	2,05	2,05	1,80	1,95	1,65	2,25	1,40
1,90	1,50	1,85	2,05	1,40	2,10	2,10	2,00	1,80	1,95	2,05
1,45	1,90									

Задача 3. По имеющимся данным (табл. 16) вычислить взаимосвязь между удоем  $x$  и МДЖ ( $y$ ) в молоке коров костромской породы.

Таблица. 16. Исходные данные (выборка)

№	$x$	$y$	$x$	$y$	$x$	$y$	$x$	$y$	$x$	$y$
1	6 492	3,7	3 440	3,8	5 102	3,8	4 317	3,8	3 608	3,8
2	4 830	3,7	4 758	3,6	5 186	3,7	3 705	3,7	5 018	3,7
3	4 098	3,9	3 925	3,6	3 309	3,7	3 845	3,8	4 043	4,0
4	2 434	3,95	2 825	3,7	3 118	3,84	6 303	4,1	1 938	3,75
5	3 230	4,0	5 679	3,8	6 346	3,8	3 234	4,0	4 573	3,9
6	5 211	3,9	4 867	3,7	4 217	3,9	4 928	3,7	3 804	3,98
7	3 108	4,0	4 485	4,0	3 655	3,9	4 662	4,2	4 672	3,7
8	2 769	4,0	3 017	4,0	4 520	3,9	5 620	4,0	3 178	3,9
9	3 976	4,1	3 855	3,9	3 409	3,8	3 190	4,0	3 289	3,9
10	4 019	3,8	3 000	4,0	3 215	4,0	2 829	4,0	4 335	3,9
11	3 077	4,1	4 120	4,0	4 605	4,1	2 254	4,4	3 764	4,0
12	4 911	4,0	3 226	4,3	2 642	4,2	3 567	3,9	3 270	3,8
13	3 283	3,7	2 547	4,0	3 270	3,8	2 974	3,8	3 180	4,0
14	3 180	4,0	4 108	4,2	3 413	3,7	3 801	4,0	3 784	3,9
15	3 780	4,0	2 707	4,0	3 687	3,9	3 775	4,0	3 750	4,0
16	2 652	3,8	3 093	4,0	2 863	3,9	2 563	3,9	3 325	3,8
17	2 526	4,0	3 473	3,9	2 195	3,9	4 679	4,0	3 090	4,1
18	3 100	3,7	2 605	3,9	4 438	4,0	2 995	3,9	3 976	3,9
19	3 009	3,7	3 757	3,9	3 926	3,9	3 079	3,9	2 767	4,0
20	3 465	3,9	3 254	3,9	2 096	4,0	3 353	3,8	3 260	3,9

Задача 4. По имеющимся данным (табл. 17) вычислить взаимосвязь между удоем  $x$  и живой массой  $y$  коров костромской породы.

Таблица. 17. Исходные данные (выборка)

№	$x$	$y$								
1	2 201	590	2 720	531	2 510	686	2 752	550	4 030	653
2	3 107	570	4 123	556	4 226	620	2 753	572	3 245	550
3	2 730	542	3 447	576	3 885	542	3 783	481	4 238	600
4	2 147	670	3 270	445	3 848	478	4 127	620	1 785	472
5	3 237	510	2 344	442	3 492	535	2 583	522	3 086	510
6	4 106	526	3 255	530	3 550	506	3 860	446	4 149	535
7	3 374	450	3 271	520	1 012	430	2 537	650	1 303	540
8	2 419	595	3 756	522	2 742	528	2 227	482	3 914	423
9	2 563	483	2 853	534	3 603	423	4 656	583	4 280	506
10	3 670	550	3 842	447	2 810	620	1 998	487	2 560	478
11	3 299	570	3 560	465	4 650	500	3 005	600	2 699	560
12	1 983	580	4 113	520	2 706	540	3 647	500	3 450	600
13	2 381	500	2 742	500	4 637	552	3 374	586	3 130	570
14	3 112	510	3 195	620	2 519	540	4 438	600	2 418	500
15	2 340	540	3 490	550	3 463	510	4 177	510	1 507	426
16	2 818	540	2 720	606	3 084	530	2 965	500	3 000	580
17	3 100	560	2 726	520	3 510	540	2 505	560	2 913	600
18	1 815	550	1 901	564	2 923	520	2 224	500	2 100	450
19	2 150	480	3 120	470	2 358	570	4 094	520	1 569	500
20	1 761	540	1 794	400	1 700	420	1 569	560	1 900	540

**Задача 5.** По имеющимся данным (табл. 18) вычислить взаимосвязь между живой массой  $x$  и высотой в холке  $y$  коров костромской породы

*Таблица. 18. Исходные данные (выборка)*

№	$x$	$y$								
1	380	114	440	116	430	115	450	117	460	118
2	473	119	468	118	480	120	475	119	485	120
3	492	120	495	120	500	121	508	122	516	123
4	539	123	520	122	524	123	532	122	538	123
5	527	122	505	121	514	121	540	124	542	124
6	553	125	564	125	572	126	578	126	557	124
7	579	125	567	124	580	127	588	128	530	124
8	592	129	603	129	610	129	692	128	697	128
9	503	115	510	116	520	118	525	120	530	124
10	532	126	537	127	540	118	539	128	603	129
11	548	118	550	121	557	123	560	124	562	126
12	570	127	572	130	578	131	580	123	600	125
13	615	127	619	129	519	131	660	123	690	129
14	693	131	670	129	380	115	400	120	419	122
15	419	124	420	116	430	118	432	119	440	121
16	450	124	459	128	460	113	470	116	475	120
17	480	121	485	121	490	122	492	124	499	125
18	490	127	700	129	690	130	699	131	698	133
19	632	127	620	120	650	124	641	128	537	127
20	603	125	538	124	671	129	480	129	480	121

**Задача 6.** Вычислите основные статистические константы (среднюю арифметическую, среднее квадратическое отклонение и коэффициент вариации с их ошибками) по данным выборки при характеристике свиноматок уржумской породы по плодовитости

9 11 15 12 11 14 11 10 11 13 12 12 11 12

**Задача 7.** Вычислите основные статистические константы (среднюю арифметическую, среднее квадратическое отклонение и коэффициент вариации с их ошибками) по данным выборки живой массы коров костромской породы в возрасте первого отела

529 497 530 545 436 565 515 495 481 500

**Задача 8.** Вычислите основные статистические константы (среднюю арифметическую, среднее квадратическое отклонение и коэффициент вариации с их ошибками) по данным выборки сбора меда в ульях

40 32 22 42 15 35 33 28 30 45 42 40 40 33

*Задача 9.* Вычислите основные статистические константы (среднюю арифметическую, среднее квадратическое отклонение и коэффициент вариации с их ошибками) по данным выборки по данным (дней) свиней крупной белой породы при достижении живой массы 100 кг.

191 186 191 188 194 199 200 199 194 201 193

*Задача 10.* Вычислите основные статистические константы (среднюю арифметическую, среднее квадратическое отклонение и коэффициент вариации с их ошибками) по данным выборки среднесуточного удоя коров швицкой породы

21,8 22,6 19,1 26,9 21,1 24,3 24,0 19,0

*Задача 11.* По данным о живой массе и удое коров черно-пестрой породы (табл. 19, 20) выясните достоверность разности в разном возрасте.

*Таблица. 19. Исходные данные*

Возраст по числу отелов	$n$	$\bar{X} \pm m_{\bar{X}}$
1-й	190	$415 \pm 3,7$
2-й	165	$538 \pm 4,1$
3-й	133	$559 \pm 5,1$
4-й	128	$565 \pm 5,2$

*Таблица. 20. Исходные данные*

Лактация по счету	$n$	$\bar{X} \pm m_{\bar{X}}$
1-я	103	$3415 \pm 173$
2-я	111	$4066 \pm 175$
3-я	102	$4621 \pm 105$
4-я	100	$4780 \pm 102$

*Задача 12.* Живая масса баранов породы прекос (по 200 голов в каждой группе):

– тип С (мясошерстный) —  $\bar{\chi} \pm m = 81 \pm 1,7$ ;

– тип С<sup>+</sup> (шерстно-мясной) —  $\bar{\chi} \pm m = 77 \pm 2,2$ ;

– тип С<sup>-</sup> (мясной) —  $\bar{\chi} \pm m = 85 \pm 2,6$ .

Вычислите критерий достоверности разности этих групп по живой массе.

*Задача 13.* Морфологические признаки пчел.

*Таблица. 21. Исходные данные*

Признак	Среднерусская порода $\bar{x} \pm t$	Башкирская порода $\bar{x} \pm t$
Длина хоботка	$6,12 \pm 0,010$	$6,26 \pm 0,009$
Длина крыла	$9,19 \pm 0,026$	$9,60 \pm 0,004$
Ширина крыла	$3,15 \pm 0,019$	$3,17 \pm 0,004$
Ширина воскового зеркала	$2,15 \pm 0,022$	$2,56 \pm 0,020$
Число зацепок, шт.	$20,42 \pm 0,598$	$19,84 \pm 0,048$

Вычислите критерий достоверности разности между указанными признаками у пчел разных пород.

*Задача 14.* Имеются следующие данные (табл. 22), характеризующие кур Y, Z и X. Выясните, по каким показателям имеется статистически достоверная разница.

*Таблица. 22. Исходные данные*

Показатели	Y		Z		X	
	n	$\bar{x} \pm t$	n	$\bar{x} \pm t$	n	$\bar{x} \pm t$
Половая зрелость, дн.	3 740	$168,65 \pm 0,20$	4 119	$167,85 \pm 0,26$	1 011	$106 \pm 0,4$
Масса яиц в 273 дн., г	3 621	$57,25 \pm 0,06$	3 806	$55,90 \pm 0,05$	940	$53,68 \pm 0,103$
Масса яиц в 476 дн., г	2 753	$60,69 \pm 0,08$	2 699	$60,58 \pm 0,08$	875	$56,98 \pm 0,12$
Яйценоскость за 273 дн., шт.	3 706	$81,52 \pm 0,19$	3 943	$82,40 \pm 0,19$	948	$69,55 \pm 0,41$
Яйценоскость за 476 дн., шт.	3 143	$240,49 \pm 0,36$	3 203	$242,19 \pm 0,39$	786	$210,15 \pm 0,79$

*Задача 15.* По данным о живой массе коров черно-пестрой породы (табл. 23) выясните достоверность разности.

*Таблица. 23. Исходные данные*

Возраст по числу отелов	n	$\bar{x} \pm t$
1-й	190	$415 + 3,7$
2-й	165	$538 \pm 4,1$
3-й	133	$559 \pm 5,1$
4-й	128	$565 + 5,2$

*Задача 16.* При изучении влияния инбридинга (родственного спаривания) на мясные качества маток различных семейств получены следующие показатели среднесуточного прироста (табл. 24).

*Таблица 24. Показатели среднесуточного прироста у инбредного и аутбредного потомства свиней*

Семейство	<i>n</i>	$\bar{X} \pm t$
«Волшебница»:		
аутбредные	108	716 ± 6,5
инбредные	76	718 ± 16,2
«Черная птичка»:		
аутбредные	44	667 ± 8,6
инбредные	8	829 ± 8,5

Одинаково ли влияние инбридинга в обоих семействах? Какова достоверность полученных показателей среднесуточного прироста между аутбредными и инбредными животными в обоих семействах?

## 8. ВОПРОСЫ ДЛЯ САМОКОНТРОЛЯ

1. Что такое генеральная совокупность?
2. Что такое выборки, как они составляются?
3. Как составляется вариационный ряд?
4. Какие бывают типы распределения и вариационных кривых?
5. Перечислите средние величины и их использование.
6. Как вычисляется средняя арифметическая величина в малых » больших выборках?
7. Какими свойствами обладают средние величины?
8. Какие показатели характеризуют разнообразие признаков?
9. Как вычисляется среднее квадратическое отклонение в малых » больших выборках?
10. Как вычисляется среднее квадратическое отклонение для альтернативных признаков?
11. Как вычисляется коэффициент фенотипической корреляции в малых и больших выборках?
12. В чем заключается различие связи между признаками при положительных и отрицательных значениях коэффициента корреляции?
13. Как вычисляется коэффициент корреляции для альтернативных признаков?
14. В каких случаях используется коэффициент ранговой корреляции?
15. Приведите формулы вычисления коэффициентов генетической корреляции.
16. Что характеризуют коэффициенты регрессии? В чем различие между коэффициентами  $R_{x/y}$  и  $R_{y/x}$ ?
17. В чем различие между коэффициентами  $r$  и  $R$ ?
18. Что такое ошибки репрезентативности? Чем отличаются они от ошибок измерения и вычисления?
19. Как вычисляют ошибку средней арифметической величины?
20. Приведите формулы вычисления ошибок  $\sigma$ ,  $C_v$ ,  $r$ ,  $R$ .
21. Как изменяется величина  $m_x$  при изменении объема выборки и величины сигмы?
22. Что такое доверительные вероятности?
23. Какие доверительные вероятности можно использовать в биологических, зоотехнических и ветеринарных исследованиях?
24. Как определяется достоверность выборочных показателей?
25. Как оценивается достоверность разности между средними величинами двух выборок?
26. Что такое критерий соответствия (хи-квадрат) и как он используется в генетических исследованиях?

27. В чем заключается цель дисперсионного анализа? Что называется общей, факториальной и остаточной дисперсией?

28. Какие бывают дисперсионные комплексы? Чем они характеризуются?

29. Как составляют однофакторный дисперсионный комплекс и вычисляют вспомогательные величины?

30. Какие показатели используются для оценки достоверности влияния изучаемого фактора?

## 9. КРАТКИЙ СЛОВАРЬ ТЕРМИНОВ, ПРИМЕНЯЕМЫХ В БИОМЕТРИИ

**Альтернативные признаки** — взаимоисключающие друг друга (больной – здоровый, высокий – низкий и т.п.).

**Аномалия** — отклонение от нормы.

**Асимметрия** — отсутствие или нарушение симметрии при распределении вариант по классам вариационного ряда.

**Варианта** — отдельное значение варьирующего признака.

**Варьирование, вариабельность** — особая форма биологической изменчивости, выражающаяся в виде внутригруппового индивидуального разнообразия.

**Вариационный ряд** — двойной ряд чисел, в котором произведена группировка их в классы в зависимости от величины изучаемого признака.

**Вариация** — разнообразие, видоизменение одного и того же признака, наблюдаемое в совокупности.

**Вероятность** — объективная возможность осуществления чего-либо, степень надежности или достоверности чего-либо.

**Выборка** — часть генеральной совокупности, предназначенная для ее характеристики.

**Гистограмма** — изображение вариационного ряда в виде столбиковой диаграммы, в которой высоты столбиков соответствуют частотам классов.

**Группировка** — объединение особей, вариант для совместного изучения, расположение вариант в таблицы или ряды.

**Дисперсия** — характеристика вариации признака, выражаемая средним квадратом отклонений вариант от средней арифметической данной совокупности.

**Доверительные вероятности** — степени надежности безошибочной оценки параметров генеральной совокупности по данным выборки.

**Изменчивость** — одно из общих свойств жизни, выражающееся в форме внутригруппового разнообразия.

**Корреляция** — связь между двумя варьирующими величинами, не имеющая функционального характера.

**Кривая распределения** — линейный график вариационного ряда.

**Критерий** — мерило надежности, достоверности суждений, выводов, их соответствие истине, реальной действительности.

**Лимиты** — границы вариации признака, максимальная и минимальная варианты совокупности или ряда.

**Медиана** — непараметрическая средняя, которая находится точно в середине вариационного ряда, делит его на две равные части.

**Мода** — наиболее часто встречающаяся варианта.

**Параметр** — величина, входящая в математическую формулу и сохраняющая свое постоянное значение в условиях данной совокупности.

**Популяция** — внутригрупповая группировка или некоторая совокупность одного и того же вида.

**Признак** — любая особенность в строении, функциях, по которой можно отличить один объект наблюдения от другого.

**Разнообразие** — неоднородность особей по каким-либо признакам.

**Ранжирование** — размещение чего-либо в определенном порядке в направлении возрастания или убывания значений изучаемого признака.

**Ранг** — порядковый номер ранжированных значений признака.

**Регрессия** — математическое выражение линии связи между коррелируемыми признаками.

**Репрезентативность** — степень соответствия выборочных параметров параметрам генеральной совокупности.

**Рендомичность** — случайность, неотобранность, при которой любой член генеральной совокупности имеет одинаковую вероятность войти в выборочную совокупность.

**Совокупность (генеральная)** — это теоретически бесконечно большая группа особей.

**Совокупность (выборочная)** — часть генеральной совокупности, предназначенная для ее характеристики.

**Степени свободы** — числа независимых величин, участвующих в образовании обобщенных статистических показателей.

**Уровни значимости** — показатели, используемые при проверке статистических гипотез, связанные с такими значениями вероятности, при которых появление ожидаемых событий в данных условиях считается практически невозможным. Чем меньше уровень значимости, тем меньше вероятность отвергнуть гипотезу.

**Частота** — абсолютная численность вариантов, указывающая на то, как часто они встречаются в совокупности.

**Явление** — отдельный факт, единичное событие, их множество составляет массовое явление.

## СПИСОК РЕКОМЕНДУЕМЫХ ИСТОЧНИКОВ

1. Бакай, А.В. Генетика / А.В. Бакай, И.И. Кочиш, Г.Г. Скрипниченко. — М. : КолосС, 2006. — 448 с.
2. Бакай, А.В. Практикум по генетике / А.В. Бакай, И.И. Кочиш, Г.Г. Скрипниченко и др. — М. : КолосС, 2010. — 301 с.
3. Лакин, Г.Ф. Биометрия. — М. : Высшая школа, 1980. — 292 с.
4. Петухов, В.Л. Ветеринарная генетика / В.Л. Петухов, А.И. Жигачев, Г.А. Назарова. — М. : Агропромиздат, 1996. — 384 с.
5. Плохинский, Н.А. Руководство по биометрии для зоотехников. — М. : Колос, 1969. — 256 с.

## ПРИЛОЖЕНИЯ

### Приложение 1

#### Определение объема выборки для получения достоверных результатов

Коэффициент изменчивости $C_v, \%$	Значение уровня вероятности $P \geq$		
	0,95	0,99	0,999
Объем выборки, гол			
Допустимая погрешность опыта 5%			
8	10	17	28
10	16	27	44
12	23	39	63
14	31	53	85
16	41	69	112
18	52	88	141
20	64	108	174
Допустимая погрешность опыта 3%			
8	28	48	77
10	44	75	121
12	64	108	174
14	87	147	237
16	114	192	310
18	144	243	392
20	178	300	480
Допустимая погрешность опыта 1%			
8	2566	433	697
10	400	676	1089
12	576	973	1568
14	784	1325	2134
16	1024	1730	2708
18	1296	2190	3528
20	1600	2704	4356

Значение критерия  $t_d$   
 для трех степеней вероятности,  
 $\nu$  — число степеней свободы

Число степеней свободы, $\nu$	Уровень вероятности			Число степе- ней свободы, $\nu$	Уровень вероятности		
	0,95	0,99	0,999		0,95	0,99	0,999
1	12,71	63,66	637,00	19	2,09	2,86	3,88
2	4,30	9,92	31,60	20	2,09	2,85	3,85
3	3,18	5,84	12,94	21	2,08	2,83	3,82
4	2,78	4,60	8,61	22	2,07	2,82	3,79
5	2,57	4,03	6,86	23	2,07	2,81	3,77
6	2,45	3,71	5,96	24	2,06	2,80	3,75
7	2,37	3,50	5,41	25	2,06	2,79	3,73
8	2,31	3,36	5,04	26	2,06	2,78	3,71
9	2,26	3,25	4,78	27	2,05	2,77	3,69
10	2,23	3,17	4,59	28	2,05	2,76	3,67
11	2,20	3,11	4,44	29	2,05	2,76	3,66
12	2,18	3,06	4,32	30	2,04	2,75	3,65
13	2,16	3,01	4,22	35-39	2,03	2,72	3,59
14	2,15	2,98	4,14	40-44	2,02	2,70	3,55
15	2,13	2,95	4,07	45-60	2,01	2,66	3,50
16	2,12	2,92	4,02	70-100	1,98	2,63	3,39
17	2,11	2,90	3,97	> 120	1,96	2,58	3,29
18	2,10	2,88	3,92				

Приложение 3

Стандартные значения  $\chi^2$  для трех уровней значимости  $\alpha$   
и числа степеней свободы  $\nu$

Число степеней свободы, $\nu$	Уровень значимости			Число степеней свободы, $\nu$	Уровень значимости		
	0,05	0,01	0,001		0,05	0,01	0,001
1	3,84	6,64	10,83	11	19,68	24,72	31,26
2	5,99	9,21	13,82	12	21,03	26,22	32,91
3	7,81	11,34	16,27	13	22,36	27,69	34,53
4	9,49	13,28	18,46	14	23,68	29,14	36,12
5	11,07	15,09	20,52	15	25,00	30,58	37,70
6	12,59	16,81	22,46	16	26,30	32,00	39,25
7	14,07	18,48	24,32	17	27,59	33,41	40,79
8	15,51	20,09	26,12	18	28,87	34,81	42,31
9	16,92	21,67	27,88	19	30,14	36,19	43,82
10	18,31	23,21	29,59	20	31,41	37,57	45,32

*Учебно-теоретическое издание*

**Генетика и биометрия. Часть 1. Биометрические методы анализа качественных и количественных признаков** : учебное пособие для студентов направления подготовки 36.03.02 «Зоотехния» очной и очно-заочной форм обучения / сост. С.Г. Белокуров. — Караваяево : Костромская ГСХА, 2015. — 46 с.

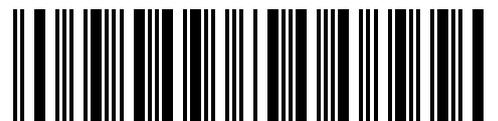
Гл. редактор Н.В. Киселева  
Редактор выпуска Т.В. Тарбеева  
Корректор Т.В. Кулинич

© Федеральное государственное бюджетное образовательное учреждение высшего образования "Костромская государственная сельскохозяйственная академия" 156530, Костромская обл., Костромской район, пос. Караваяево, уч. городок, д. 34, КГСХА

Компьютерный набор. Подписано в печать 09/11/2015.  
Заказ №1122. Формат 84x60/16. Тираж 100 экз.  
Усл. печ. л. 2,88. Бумага офсетная. Отпечатано  
24/12/2015. Цена 22,00 руб.

Отпечатано с готовых оригинал-макетов в академической типографии на цифровом дубликаторе.  
Качество соответствует предоставленным оригиналам.  
вид издания: первичное (редакция от 18.09.2015 № 131)

Цена 22,00 руб.



2015\*1122