ЛЕКЦИИ ПО ДИСЦИПЛИНЕ «ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ В АПК

Лекция 1.
СТАТИСТИЧЕСКИЕ ХАРАКТЕРИСТИКИ КОЛИЧЕСТВЕННОЙ И КАЧЕСТВЕННОЙ ИЗМЕНЧИВОСТИ

Количественная изменчивость. Основными статистическими характеристиками количественной изменчивости являются средняя арифметическая $(\bar{x})$, дисперсия (s^2), стандартное отклонение (s), ошибка средней арифметической $(s_{\bar{x}})$, коэффициент вариации *(V)* и относительная ошибка выборочной средней $(s_{\bar{x}}\%)$.

Средняя арифметическая $\bar{x}$ представляет собой обобщённую, абстрактную характеристику всей совокупности в целом. Если сумму всех вариант $(X_1 + X_2 + \cdots + X_n)$ обозначить через $\sum X$, а число всех вариант через n, то формула для определения простой средней арифметической примет следующий вид:

$$\bar{x} = \sum \frac{X}{n},$$

Взвешенную среднюю арифметическую вычисляют по формуле:

$$\bar{x} = \frac{f_1X_1 + f_2X_2 + \cdots + f_nX_n}{f_1 + f_2 + \cdots + f_n} = \frac{\sum fX}{n},$$

где X — значение признака, варианты; f – частота встречаемости каждой варианты, признака; n – общее число измеренных значений, сумма всех частот, $(n = \sum f)$

Основное свойство средней арифметической заключается в равенстве суммы всех положительных и всех отрицательных отклонений от нее, т. е. сумма центральных отклонений всех отдельных вариант от X равна нулю: $\sum(X - \bar{x}) = (X_1 - \bar{x}) + (X_2 - \bar{x}) + \cdots + (X_n - \bar{x}) = 0$. Если $\sum(X - \bar{x})$ оказалась неравной нулю, значит, допущена ошибка в вычислениях.

Дисперсия s^2 и стандартное отклонение s служат основными мерами вариации, рассеяния изучаемого признака. Дисперсия представляет собой частное от деления суммы квадратов отклонений $\sum(X - \bar{x})^2$ на число всех измерений без единицы $(n-1)$:

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

Размерность дисперсии равна квадрату размерности изучаемого признака, что неудобно и заставляет ввести для измерения рассеяния другую характеристику, имеющую размерность варьирующей величины и называемую **стандартным или средним квадратическим отклонением.** Его получают извлечением квадратного корня из дисперсии:

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

Если исходные наблюдения сгруппированы, и частоты групп обозначены через f, то дисперсию и стандартное отклонение вычисляют по формулам:

$$s^2 = \frac{\sum f(X - \bar{x})^2}{n-1} \quad \text{и} \quad s = \sqrt{s^2} = \sqrt{\frac{\sum f(X - \bar{x})^2}{n-1}}.$$

Для вычисления дисперсии s^2 следует определить отклонения всех вариант X от среднего арифметического $(X - \bar{x})$, возвести каждое такое отклонение в квадрат $(X - \bar{x})^2$ и сумму этих квадратов $\sum(X - \bar{x})^2$ разделить на число всех измерений без единицы $(n-1)$. Для вычисления стандартного отклонения необходимо извлечь квадратный корень из дисперсии.

Из математической статистики известно, что при определении любых средних величин сумму всех показателей необходимо делить на число независимых друг от друга

величин. В связи с этим в формулах сумму квадратов отклонений $\sum(X-\bar{x})^2$ делят не на общее число наблюдений, а на число без единицы, так как одно любое отклонение зависимое и может быть найдено из равенства $\sum(X-\bar{x})=0$. Остальные отклонения могут свободно варьировать, принимать любые значения. Число свободно варьирующих величин называется числом степеней свободы или числом степеней свободы вариации. Оно обозначается *v* и в простейшем случае равно $n-1$.

При вычислении средней арифметической $\bar{x}$ все величины независимы друг от друга, поэтому сумма их делится на общее число вариант n. Но когда уже известен ряд наблюдений от X_1 до X_n, каждое значение ряда, так же как и каждое отклонение $(X-\bar{x})$, можно легко определить по значению $\bar{x}$ и значениям остальных $n-1$ вариант ряда. Действительно, любое отклонение зависит от величины всех остальных и численно равно сумме их, взятых с обратным знаком, так как сумма всех отклонений $\sum(X-\bar{x})=0$. Поэтому неизвестное нам отклонение должно свести эту сумму к нулю. Следовательно, отклонение одной любой варианты от $\bar{x}$ как бы лишено свободы вариации и точно определяется варьированием всех остальных вариант, т. е. $n-1$. В связи с этим число независимых величин при определении s^2 и s равно не *n*, а $n-1$.

При вычислении дисперсии и стандартного отклонения по основным формулам нередко возникают технические неудобства. Средняя арифметическая обычно получается в виде числа с дробью, поэтому центральные отклонения $(X-\bar{x})$ и особенно квадраты их $(X-\bar{x})^2$ получаются многозначными, что затрудняет счетную работу и ведет к ошибкам. Поэтому разработано несколько способов вычисления s^2 и s, которые значительно упрощают арифметические расчеты. Они основаны на том, что для получения суммы квадратов центральных отклонений $\sum(X-\bar{x})^2$ достаточно взять отклонения от любого произвольного числа А (условной средней, произвольного начала) и произвести с ними действия по формуле:

$$\sum(X-\bar{x})^2=\sum(X-A)^2-\frac{[\sum(X-A)]^2}{n}=\sum X_1{}^2-\frac{\sum X_1{}^2}{n},$$

Если за условную среднюю (произвольное начало) принять нуль, то формула примет следующий вид:

$$\sum(X-\bar{x})^2=\sum X^2-\frac{(\sum X)^2}{n},$$

Формула $\sum X_1{}^2-\frac{\sum X_1{}^2}{n}$ сильно облегчает работу по вычислению дисперсии и стандартного отклонения для больших групп с многозначными числами. Условную среднюю А выбирают с таким расчетом, чтобы разности (X-A) были возможно меньше. Часто в качестве А берут целое число, близкое к предполагаемой средней. Среднюю арифметическую в этом случае вычисляют по формуле:

$$\bar{x}=A+\frac{\sum X_1}{n}$$

Для малых групп с малозначными числами сумму квадратов отклонений легче получить по формуле:

$$\sum(X-\bar{x})^2=\sum X^2-\frac{(\sum X)^2}{n}.$$

Стандартное отклонение служит показателем, который дает представление о наиболее вероятной средней ошибке отдельного, единичного наблюдения, взятого из данной совокупности. В пределах одного значения $(\pm 1s)$ *укладывается примерно* $^2/_3$ *всех наблюдений, или, точнее, 68,3% всех вариант, т. е. основное ядро изучаемого ряда величин. Поэтому стандартное отклонение называют также основным отклонением вариационного ряда.* Следовательно, возможны отклонения от $\bar{x}$, превосходящие $\pm 1s$, но вероятность их по мере удаления отклонений от $\pm 1s$ всё время уменьшается. Так,

вероятность встретить варианту,| отклоняющуюся от $\bar{x}$ на величину больше $\pm 3s$, составляет всего около 0,3%. Поэтому *утроенное значение стандартного отклонения принято считать предельной ошибкой отдельного наблюдения,* и, следовательно, почти все значения вариант в вариационном ряду укладываются в пределах $\pm 3s$. Шестикратное значение среднего квадратического отклонения (от $+3s$,до $-3s$,) дает ясное представление о ширине ряда наблюдений, о его рассеянии.

Коэффициент вариации V – стандартное отклонение, выраженное в процентах к средней арифметической данной совокупности:

$$V = \frac{s}{\bar{x}} 100.$$

Коэффициент вариации является относительным показателем изменчивости. Использование коэффициента вариации имеет смысл при изучении вариации признака, принимающего только положительные значения. Не имеет смысла, например, коэффициент вариации, вычисленный для характеристики колеблемости среднегодовой температуры, близкой к *0°*, когда варьирующий признак принимает как положительные, так и отрицательные значения.

Изменчивость принято считать незначительной, если коэффициент вариации не превышает 10%; средней, если V выше 10%, но менее 20%, и значительной, если коэффициент вариации более 20%.

Для характеристики степени выравненности материала иногда целесообразно использовать величину, дополняющую значение коэффициента вариации до 100. Этот показатель называют коэффициентом выравненности и определяют по равенству *B=100—V.*

Коэффициенты изменчивости и выравненности, будучи отвлеченными числами, выраженными в процентах, дают возможность сравнивать варьирование признаков разной размеренности, например высоты и массы, содержания азота и площади листьев, а также при сравнении изменчивости величин, уровень которых резко различен (например, урожай льноволокна и корнеплодов). При изучении вариабельности признаков одинаковой размеренности необходима известная осторожность – коэффициент вариации может дать искаженное представление об изменчивости, например, при разных значениях $\bar{x}$ и одинаковых *s*. В этих случаях степень вариации необходимо оценивать величиной s^2 или *s*.

Ошибка выборочной средней или ошибка выборки $s_{\bar{x}}$ является мерой отклонения выборочной средней $\bar{x}$ от средней всей (генеральной) совокупности μ. Ошибки выборки возникают вследствие неполной репрезентативности (представительности) выборочной совокупности и свойственны только выборочному методу исследования. Они связаны с перенесением результатов, полученных при изучении выборки, на всю генеральную совокупность. Величина ошибок зависит от степени изменчивости изучаемого признака и от объема выборки.

Ошибка выборочной средней прямо пропорциональна выборочному стандартному отклонению *s* и обратно пропорциональна корню квадратному из числа измерений *n*, т. е.

$$s_{\bar{x}} = \frac{s}{\sqrt{n}} = \sqrt{\frac{s^2}{n}},$$

Ошибки выборки выражают в тех же единицах измерения, что и варьирующий признак, и приписывают к соответствующим средним со знаками ±, т.е $\bar{x} \pm s_{\bar{x}}$.

Ошибка средней арифметической тем меньше, чем меньше варьирует опытный материал и чем из большего числа измерений вычислено среднее арифметическое. Ошибка выборки, выраженная в процентах от соответствующей средней, называется относительной ошибкой выборочной средней:

$$s_{\bar{x}}\% = \frac{s_{\bar{x}}}{\bar{x}} 100.$$

Относительную ошибку средней иногда обозначают буквой ***P*** *и* называют «точностью опыта», «точностью анализа». Следует признать крайне неудачным это укоренившееся понятие. При одних и тех же значениях выборочных средних возрастание величины ***P***

свидетельствует о том, что опыт становится менее точным, так как чем больше абсолютная ошибка эксперимента, тем выше и относительная ошибка, т. е. ***P.*** Кроме того, указанное обстоятельство вносит в понятие «точность» элемент двойственности, величина ***P*** часто необоснованно используется для оценки качества опытной работы и браковки полевых опытов. Так, если ***P*** превышает 5%, то рекомендуется совершенствовать методику, а опыты с *P>7*-8% браковать.

Такой подход очень условен, так как значение ***P*** зависит не только от методического уровня эксперимента, но и от урожайности возделываемой культуры. Без учета уровня урожайности часто опыты, имеющие практически равные абсолютные ошибки и, следовательно, равноценные по точности, могут по величине ***P*** классифицироваться по-разному.

Например, в опытах с зерновыми, проведенных на участках с низким, средним и высоким уровнем плодородия при средней урожайности, равной соответственно 14; 25 и 45 ц с 1 га, могут быть получены близкие значения ошибок $s_{\bar{x}} \approx 1{,}5$ ц с 1 га. По фактической точности, мерой которой и является абсолютная ошибка $s_{\bar{x}}$, эти опыты равноценны. Однако по величине ***P*** первый опыт относят к «недостоверным» и бракуют ***(P***=10,7%*)*, для второго надо выяснить причины низкой «точности» ***(P****=6,0%)*, а третий опыт проведен достаточно «точно» ***(P****=3,3%)*. Понятно, что в данном случае величина ***P*** вводит экспериментатора в заблуждение относительно фактической точности опыта.

Вследствие недостаточной обоснованности и двойственности понятия «точность опыта» в дальнейшем мы не будем им пользоваться. Вызывает возражение и дальнейшее использование буквы ***P*** для обозначения относительной ошибки средней. Известно, что этим символом во всех руководствах и учебниках по математической статистике обозначается вероятность. В данной работе относительную ошибку будем обозначать символом $s_{\bar{x}}\%$.

Качественная изменчивость. В биологических и агрономических исследованиях часто приходится иметь дело с качественной изменчивостью признаков: разная форма и окраска семян и плодов, расщепление гибридов и т. д. Частным случаем качественной изменчивости является альтернативная, при которой варьирующие признаки представляют собой одну из двух возможностей (альтернатив) – наличие или отсутствие признака, например мужские или женские экземпляры, растения больные и здоровые, колосья остистые и безостые и т. п. Группировка результатов наблюдений при качественном варьировании сводится к распределению совокупности объектов на группы (классы) с разными качественными признаками.

Основными статистическими показателями (параметрами) качественной изменчивости являются доля признака, показатель изменчивости, коэффициент вариации и ошибка выборочной доли.

Доля признака, или относительная численность (частота) отдельной варианты в данной совокупности. Доля признака обозначается через p_1, p_2, p_3 и т. д. и может быть выражена в частях единицы или в процентах. В первом случае сумма всех долей в пределах данной совокупности или ряда распределения равна единице, а во втором – 100%.

Доля признака – это отношение численности каждого из членов ряда n_1, n_2, n_3 и т. д. к численности совокупности ***N,*** т. е. вероятность появления данного признака в изучаемой совокупности:

$$p_1 = \frac{n_1}{N}; \quad p_2 = \frac{n_2}{N}; \quad p_3 = \frac{n_3}{N} \quad \text{и т.д.}$$

При альтернативной (двояковозможной) изменчивости доля одного признака обозначается через *p*, а второго через *q.* На основании очевидного равенства $p + q = 1{,}0$ (или 100%), так как вероятность двух противоположных событий всегда равна единице (100%), значение *q=1-p.*

Показатель изменчивости качественного признака *s* характеризует варьирование величин ряда относительно друг друга. Значение показателя изменчивости определяется по формуле:

$$s = \sqrt[k]{p_1 \cdot p_2 \cdot p_3 \cdot \dots \cdot p_k},$$

где p_1, p_2 и т. д – доли признака (или процентные значения их) в общей совокупности, k – число градаций признака.

Когда $k>2$, то формулу для вычисления показателя изменчивости удобнее прологарифмировать:

$$lg\ s = \frac{lg p_1 + lg p_2 + \cdots + lg p_k}{k}.$$

Если изучаемая совокупность представлена объектами с двумя градациями признака (альтернативная изменчивость), то

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

где p и q – доли признака, выраженные в частях единицы или процентах.

Например, показатель изменчивости при $p = 0{,}10$ и $q = 0{,}90$ будет равен:

$$s = \sqrt{pq} = \sqrt{0{,}10 \cdot 0{,}90} = 0{,}30 \text{ (или 30\%)}.$$

В зависимости от соотношения p и q значение s изменяется от 0 до 0,5. Максимальная изменчивость качественного признака $s_{макс}$ будет наблюдаться тогда, когда $p=q=0{,}5$ и $s_{макс} = \sqrt{0{,}5 \cdot 0{,}5} = 0{,}50$ (или 50%). Значения максимальной (наибольшей) изменчивости для распределений с разным числом градаций качественных признаков даны ниже:

Число градаций признака	$s_{макс}$	Число градаций признака	$s_{макс}$
2	0,500 (50,0%)	5	0,200 (20,0%)
3	0,333 (33,3%)	6	0,167 (16,7%)
4	0,250 (25,0%)	7	0,143 (14,3%)

Пользуясь величинами максимальных значений $s_{макс}$, можно вычислить **коэффициент вариации качественных признаков** – фактический показатель изменчивости, выраженный в процентах к максимально возможной изменчивости:

$$V_p = \frac{s}{s_{макс}} \cdot 100.$$

Коэффициент вариации характеризует относительную степень изменчивости изучаемых признаков и широко используется для сравнительной оценки выравненности различных совокупностей. Максимальное значение $V_p = 100\%$ наблюдается при $s = s_{макс}$.

Ошибка выборочной доли s_p – мера отклонения доли признака выборочной совокупности p от доли его по всей генеральной совокупности P вследствие неполной представительности (репрезентативности) выборки. Ошибку доли вычисляют по формуле:

$$s_p = \frac{s}{\sqrt{N}},$$

где s – показатель изменчивости качественного признака; N – объём выборки.

Для альтернативного варьирования, когда значение $s = \sqrt{pq}$, формула ошибки выборочной доли примет вид:

$$s = \sqrt{pq/N}.$$

Здесь p и q могут быть выражены в долях единицы или процентах.

Вероятность встретить p (или q) в интервале $p \pm s_p$ составляет около 68%, в интервале $p \pm 2s_p$ – 95% и в интервале $p \pm 3s_p$ – около 99%. Следовательно, подобно количественной изменчивости, все значения p с вероятностью 99% укладываются в пределах тройной ошибки выборочной доли.

Лекция 2.
ВЫЧИСЛЕНИЕ СТАТИСТИЧЕСКИХ ХАРАКТЕРИСТИК ВЫБОРКИ ПРИ КОЛИЧЕСТВЕННОЙ ИЗМЕНЧИВОСТИ ПРИЗНАКА

К количественным относят признаки, которые могут быть охарактеризованы количественно, – урожай с делянки, число, высота и масса растений, содержание белка и клейковины в зерне и т. д. Различают два вида количественной изменчивости: непрерывную и прерывистую или дискретную. В первом случае значения признака выражены мерами объема, длины, массы и т. д., во втором различия между единицами наблюдения выражаются целыми числами, между которыми нет, и не может быть переходов, например число зерен в колосе и т. д.

Выборки, состоящие из 20–30 единиц наблюдения, называют малыми, а выборки большего объема – большими.

После изучения выборочная совокупность представляет собой ряд варьирующих значений признака, записанных в той последовательности, в какой они были получены. Статистические характеристики вычисляются по формулам таблицы 1.

В таблице 1 через *X* обозначены отдельные значения признака в малых выборках и групповые средние в больших выборках; X_1 – преобразованные значения исходных дат; *A* – произвольное начало, условная средняя; *f* – частота, численность группы; *n* – объём выборки; *t* – теоретическое значение критерия Стьюдента.

Для вычисления средней арифметической и суммы квадратов (числитель дисперсии) в таблице дано несколько формул. Все они дают практически одинаковые результаты.

При вычислениях исходные даты целесообразно преобразовать так, чтобы отбросить лишние цифры и опустить запятые. Последние потом вновь восстанавливают. Преобразование (кодирование) может осуществляться вычитанием от результатов измерении одного и того же числа *A,* умножением или делением исходных дат на одно и то же число *K,* а также одновременным проведением двух действий.

При работе с преобразованными (закодированными) датами необходимо иметь в виду, что вычитание или прибавление условной средней *A*, т.е. изменение начала отсчета, не оказывает влияния на сумму квадратов и поправка необходима лишь при определении средней арифметической. Если преобразование осуществляется путём умножения или деления, то для получения окончательных результатов среднее арифметическое и сумму квадратов надо скорректировать: среднее – в первом случае надо разделить, во втором – умножить на число кода *K,* а сумму квадратов соответственно разделить или умножить на K^2.

Таблица 1 – Формулы для вычисления статистических характеристик выборки при количественной изменчивости

Показатель	Малая выборка (несгруппированные данные)	Большая выборка (сгруппированные данные)
Средняя арифметическая	$\bar{x} = \frac{\sum X}{n} = A + \frac{\sum X_1}{n}$	$\bar{x} = \frac{\sum fX}{n} = A + \frac{\sum fX_1}{n}$
Дисперсия	$s^2 = \frac{\sum(X-\bar{x}^2)^2}{n-1} = \frac{\sum X^2-(\sum X)^2 \div n}{n-1} = \frac{\sum X_1{}^2-(\sum X_1)^2 \div n}{n-1}$	$s^2 = \frac{\sum f(X-\bar{x}^2)^2}{n-1} = \frac{\sum fX^2-(\sum fX)^2 \div n}{n-1} = \frac{\sum fX_1{}^2-(\sum fX_1)^2 \div n}{n-1}$
Стандартное отклонение	$s = \sqrt{s^2}$	
Коэффициент вариации	$V = \frac{s}{\bar{x}} 100$	
Ошибка средней	$s_{\bar{x}} = \frac{s}{\sqrt{n}} = \sqrt{\frac{s^2}{n}}$	
Относительная ошибка средней	$s_{\bar{x}}\% = \frac{s_{\bar{x}}}{\bar{x}} 100$	
Доверительный интервал для среднего значения	$\bar{x} \pm t s_{\bar{x}}$	
Степень свободы	n-1	

Лекция 3.
СТАТИСТИЧЕСКИЕ МЕТОДЫ ПРОВЕРКИ ГИПОТЕЗ

Вопрос о статистической проверке гипотез – один из основных при применении математической статистики в научных исследованиях. Статистические методы или критерии проверки гипотез – надежная основа принятия тех или иных решений при некоторой неопределенности, обусловленной случайной вариацией изучаемых явлений. Они применяются всегда, когда необходимо использовать выборочное наблюдение для суждения о законе распределения совокупности, для решения вопроса о существенности разности между выборочными средними, для установления принадлежности варианты к данной совокупности и соответствия между фактическими и теоретическими распределениями частот.

Практически проверка гипотез часто сводится к сравнению статистических характеристик, оценивающих параметры законов распределения, т. е. к проверке определенных статистических гипотез. Вообще статистической гипотезой называют научное предположение о тех или иных статистических законах распределения рассматриваемых случайных величин, которое может быть проверено на основе выборки. В большинстве случаев задача сводится к проверке гипотезы об отсутствии реального различия между фактическими и теоретически ожидаемыми наблюдениями. Эту гипотезу называют нулевой гипотезой и обозначают символом H_0.

Если в результате проверки H_0 различия между фактическими и гипотетическими показателями близки к нулю или находятся в области допустимых значений, то нулевая гипотеза не отвергается, а если различия оказываются в критической для данного статистического критерия области, которые при нашей гипотезе невозможны, а потому несовместимы с ней, H_0 опровергается. Принятие нулевой гипотезы означает, что данные наблюдении не противоречат предположению об отсутствии различий между фактическими и гипотетическими (теоретическими) или между двумя рядами фактических распределений, но не доказывают отсутствия такого различия. Отбрасывание гипотезы означает, что эмпирические данные несовместимы с H_0, а верна другая, альтернативная гипотеза.

Справедливость нулевой гипотезы проверяется вычислением статистических критериев проверки для определенного уровня значимости.

Уровень значимости определяется конкретными задачами исследования; он характеризует, в какой мере мы рискуем ошибиться, отвергая нулевую гипотезу. Чем меньше уровень значимости, тем меньше вероятность отвергнуть H_0, когда она верна, или, как говорят, совершить ошибку I рода, но тем больше вероятность совершить ошибку II рода, когда не отвергают H_0, в действительности неверную. Уровень значимости не измеряет степень риска, связанный с принятием неверной гипотезы (ошибка II рода), он контролирует лишь ошибку I рода.

Для проверки статистической гипотезы H_0 используют критерии двух видов: параметрические и непараметрические.

Параметрическими называют критерии, которые основаны на предположении, что распределение признака в совокупности подчиняется некоторому известному закону, например закону нормального распределения. К таким критериям относятся, в частности, критерии *t* и *F*, применение которых требует вычисления оценок параметров распределения.

Непараметрическими называют критерии, использование которых не требует предварительного вычисления оценок неизвестных параметров распределения и даже приближенного значения закона распределения признака. Они могут применяться и тогда, когда распределение сильно отклоняется от нормального. С другой стороны, непараметрические критерии менее эффективны по сравнению с параметрическими, и поэтому их целесообразно использовать только в предварительных исследованиях.

Лекция 4.

ДИСПЕРСИОННЫЙ АНАЛИЗ

Дисперсионный анализ разработан и введён в практику сельскохозяйственных и биологических исследований английским учёным Р.А. Фишером, который открыл закон распределения отношений средних квадратов (дисперсий):

$$\frac{\text{средний квадрат выборочных средних}}{\text{средний квадрат объектов}} = \frac{s_1^2}{s_2^2} = F.$$

Дисперсионный анализ широко используется для планирования эксперимента и статистической обработки его данных. Если в недалеком прошлом считали, что роль математика состоит лишь в анализе экспериментальных данных, то работы Р. Л. Фишера коренным образом изменили эту точку зрения, и в настоящее время статистическое планирование опыта в соответствии с требованиями дисперсионного анализа и математическая интерпретация результатов – непременные условия успешного получения ответов на вопросы, интересующие экспериментатора. Статистически обоснованный план эксперимента определяет и метод математического анализа результатов. Поэтому современный эксперимент нельзя правильно спланировать, не зная основ дисперсионного анализа.

При дисперсионном анализе одновременно обрабатывают данные нескольких выборок (вариантов), составляющих единый статистический комплекс, оформленный в виде специальной рабочей таблицы. Структура статистического комплекса и его последующий анализ определяются схемой и методикой эксперимента.

Сущностью дисперсионного анализа является расчленение общей суммы квадратов отклонений и общего числа степеней свободы на части – компоненты, соответствующие структуре эксперимента, и оценка значимости действия и взаимодействии изучаемых факторов по F-критерию.

Если обрабатывают однофакторныс статистические комплексы, состоящие из нескольких независимых выборок, например *l*-вариантов в вегетационном опыте, то общая изменчивость результативного признака, измеряемая общей суммой квадрант C_Y расчленяется на два компонента: варьирование между выборками (вариантами) C_V и внутри выборок C_Z. Следовательно, в общей форме изменчивость признака может быть представлю на выражением:

$$C_Y = C_V + C_Z$$

Здесь вариация между выборками (вариантами) представляет ту часть общей дисперсии, которая обусловлена действием изучаемых факторов, а дисперсия внутри выборок характеризует случайное варьирование изучаемого признака, т. е. ошибку эксперимента.

Общее число степеней свободы *(N—1)* также расчленяется на две части – степени свободы для вариантов *(l-1)* и для случайного варьирования *(N-l):*

$$N-1=(l-1)+(N-1)$$

Если обрабатывают однофакторные сопряженные статистические комплексы, когда выборки (варианты) связаны каким-то общим контролируемым условием, например наличием *n* организованных повторений в полевом опыте, общая сумма квадратов разлагается на три части: варьирование повторений C_p, вариантов C_V и случайное C_Z. В подобных случаях общая изменчивость и общее число степеней свободы могут быть представлены выражениями:

$$C_Y = C_p + C_V + C_Z;$$

$$(N-1)=(n-1)+(l-1)+(n-1)(l-1)$$

Суммы квадратов отклонений по данным полевого опыта – статистического комплекса с *l*-вариантами и *n*-повторениями – находят обычно в такой последовательности. В исходной таблице определяют суммы по повторениям *P,* вариантам *V* и общую сумму всех наблюдений $\sum X$. Затем вычисляют:

1) общее число наблюдений $N=ln$;

2) корректирующий фактор (поправку) $C = (\sum X)^2/N$;

3) общую сумму квадратов $C_Y = \sum X^2 - C$;

4) сумму квадратов для повторений $C_p = \sum P^2/l - C$;

5) сумму квадратов для вариантов $C_V = \sum V^2/n - C$;

6) сумму квадратов для ошибки (остаток) $C_Z = C_Y - C_p - C_V$.

Две последние суммы квадратов C_V и C_Z делят на соответствующие им степени свободы, т. е. приводят к сравниваемому виду – одной степени свободы вариации. В результате получают два средних квадрата (дисперсии):

$$\text{вариантов } s_V^2 = \frac{C_V}{l-1} \quad \text{и} \quad \text{ошибки } s^2 = \frac{C_V}{(n-1)(l-1)}.$$

Эти средние квадраты и используют в дисперсионном анализе для оценки значимости действия изучаемых факторов. Оценка проводится путем сравнения дисперсии вариантов s_V^2 с дисперсией ошибки s^2 по критерию $F = s_v^2/s^2$. Таким образом, за базу – единицу сравнения принимают средний квадрат случайной дисперсии, которая определяет случайную ошибку эксперимента. При этом проверяемой нулевой гипезой служит предположение: все выборочные средние являются оценками одной генеральной средней и, следовательно, различия между ними несущественны. Если $F_{\text{факт}} = s_v^2/s^2 < F_{\text{теор}}$, то нулевая гипотеза H_0:$d=0$ не отвергается; между всеми выборочными средними нет существенных различий, и на этом проверка заканчивается. Нулевая гипотеза отвергается, когда $F_{\text{факт}} = s_v^2/s^2 \geq F_{\text{теор}}$. В этом случае дополнительно оценивают существенность частных различий по НСР и определяют, между какими средними имеются значимые разности.

Теоретическое значение критерия *F* для принятого в исследовании уровня значимости равного 5% находят по таблицам 1-2 В приложения В с учетом числа степеней свободы для дисперсии вариантов и случайной дисперсии.

Лекция 5.
КОРРЕЛЯЦИЯ И РЕГРЕССИЯ

Вопросы:
1. Корреляция
2. Регрессия

1. Корреляция

При проведении математической обработки опытных данных редко приходится иметь дело с точными и определёнными функциональными связями, когда каждому значению одной величины соответствует строго определённое значение другой величинью. Здесь чаще встречаются такие соотношения между переменными, когда каждому значению признак *X* соответствует не одно, а множество возможных значений признака *Y*, т.е. их распределение. Такие связи, обнаруживаемые лишь при массовом изучении признаков, в отличие от функциональных называют стохастическими (вероятностными) или корреляционными.

При изучении корреляционных связей возникают два основных вопроса – о тесноте связи и о форме связи. Для измерения тесноты и формы связи используют специальные статистические методы, называемые корреляцией и регрессией.

По форме корреляция может быть линейной и криволинейной, по направлению прямой и обратной. Корреляцию и регрессию называют простой, если исследуется связь между двумя признаками, и множественной, когда изучается зависимость между тремя и более признаками.

В качестве числового показателя простой линейной корреляции, указывающей на тесноту (силу) и направление связи *X* на *Y* используют коэффициент корреляции, обозначаемый буквой *r*.

Линейный коэффициент корреляции является безразмерной величиной, изменяющейся в области $-1<r<+1$.

Коэффициент корреляции рассчитывают по формуле:

$$r = \frac{\sum(X-\bar{x})(Y-\bar{y})}{\sqrt{\sum(X-\bar{x})^2\sum(Y-\bar{y})^2}}$$

Связи между признаками могут быть слабыми и сильными (тесными). Их критерии оцениваются по шкале Чеддока:

$0{,}1 < r_{xy} < 0{,}3$: слабая;
$0{,}3 < r_{xy} < 0{,}5$: умеренная;
$0{,}5 < r_{xy} < 0{,}7$: заметная;
$0{,}7 < r_{xy} < 0{,}9$: высокая;
$0{,}9 < r_{xy} < 1$: весьма высокая;

Квадрат коэффициента корреляции (r^2) называется коэффициентом детерминации и обозначается d_{xy}. Он показывает долю или % (при умножении доли на 100%) тех изменений, которые в данном явлении зависят от изучаемого фактора.

Регрессия

Под регрессией понимается изменение результативного признака *Y* (функции) при определённом изменении одного или нескольких факториальных (аргументов).

Связь между функцией и аргументом выражается уравнением регрессии или производственной функцией. При простой регрессии уравнение кратко обозначается *Y=f(X)* и при множественной *Y=f(X, Z, V, …)*. Если степень связи между признаками велика, то по уравнению регрессии можно предсказать значение результативного признака для определённых значений факториального признака.

Уравнение парной линейной корреляционной связи называется *уравнением парной регрессии* и имеет вид:

$$y = a + b_j x,$$

где b_{yx} – коэффициент регрессии *Y* по *X*.

Коэффициент регрессии вычисляются по формулам:

$$b_j = \frac{\sum(x-\bar{x})(y-\bar{y})}{\sum(x-\bar{x})^2}$$

где $\bar{x}$ и $\bar{y}$ – средние арифметические для ряда *X* и *Y*.

Коэффициент регрессии b_j показывает, как изменяется *Y* при изменении *X* на единицу измерения, и выражается в единицах *Y*.

Таким образом, коэффициентом линейной регрессии называется число, показывающее, в каком направлении и на какую величину изменяется в среднем признак *Y* (функция) при изменении признака *X* (аргумента) на единицу измерения. Коэффициент регрессии имею знак коэффициента корреляции.

Коэффициенты регрессии b_j являются именованными числами, выраженными в разных единицах измерения, и поэтому несравнимы друг с другом.

Для преобразования их в сравнимые относительные показатели рассчитываются показатели, которые называют *стандартизованным коэффициентом регрессии или β-коэффициентом.*

$$\beta_i = b_j \frac{\sigma_{x_j}}{\sigma_y}$$

β-коэффициент при факторе x_j, определяет меру влияния вариации фактора x_j на вариацию результативного признака *y* при отвлечении от сопутствующей вариации других факторов, входящих в уравнение регрессии.

Бета-коэффициент показывает, на какую часть величины своего среднего квадратичного отклонения изменится в среднем значение результативного признака при изменении факторного признака на величину его среднеквадратического отклонения при фиксированном на постоянном уровне значении остальных независимых переменных.

Общий вид многофакторного линейного уравнения регрессии имеет вид:

$$Y = a + b_1 x_1 + b_2 x_2 + \cdots + b_n x_n = a + \sum_{j=1}^{n} b_j x_j$$

Если же связь отдельного фактора с результативным признаком не является линейной, то производят линеаризацию уравнения путем замены или преобразования величины факторного признака и аргументов.

Для исключения из уравнения регрессии аргументов не существенно влияющих на результативный фактор применяется так называемый пошаговый итеративный процесс, в результате которого в уравнении регрессии остаются только те аргументы, которые существенно влияют на результативный фактор.

Лекция 6.

КЛАСТЕРНЫЙ АНАЛИЗ

Задача разделения исходных данных на однородные совокупности решается с помощью кластерного анализа. При этом не требуется априорной информации о распределении генеральной совокупности. Иерархическая агломеративная процедура последовательно объединяет элементы кластеров (групп) сначала самых близких, а затем всё более отдалённых друг от друга. В начальном состоянии каждый кластер (группа) содержит наблюдение, а расстояние между кластерами определяется обычной евклидовой метрикой. Следовательно, расстояние между *l*-м и *r*-м кластерами равно:

$$\rho_{rl} = \sqrt{\sum_{k=1}^{n} (x_{rk} - x_{lk})^2}.$$

Расстояние между кластерами после объединения определяется по принципу «ближайшего соседа», т.е. расстояние между кластером и результатом объединения кластеров $S_{(rl)} = S_r \cup S_l$ равно:

$$\rho(S_k, S_{(r,l)}) = \frac{1}{2}\rho_{rk} - \frac{1}{2}\rho_{kl} - \frac{1}{2}|\rho_{rk} - \rho_{kl}|.$$

Лекция 7.

ПРИМЕНЕНИЕ ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ В ЖИВОТНОВОДСТВЕ

Вопросы:

1. Уравнение роста.
2. Молочное хозяйство
3. Производство говядины
4. Выбраковка
5. Рационы кормления

1. УРАВНЕНИЯ РОСТА

В этом разделе дается обзор наиболее часто встречающихся уравнений роста. Там, где это возможно, эти уравнения будут выводиться на основе анализа простых моделей – обычно путем интегрирования дифференциального уравнения типа (1) или (2).

$$\frac{dW}{dt} = h(W) \qquad (1)$$

$$\frac{dW}{dt} = u(W,t) \text{ – уравнение Гомпертца,} \qquad (2)$$

где: W – сухая масса вещества, накопленная растением в процессе роста, t – время роста растения, h – некоторая функция роста растения, u – есть функция от W и t.

Такой подход облегчит интерпретацию параметров зависимостей типа «сухая масса – время». Хотя чаще всего на практике для аппроксимации временных рядов роста применяется полулогарифмическая шкала, она таит в себе для неопытных исследователей определенные опасности: например, кривая сложной формы может на таком графике полностью утратить свою специфику. Чтобы этого избежать, решения уравнений, наиболее характерных для каждого из рассматриваемых здесь типов, изображаются как в линейном, так и в полулогарифмическом масштабе.

На рисунке 1 схематически изображена простая двухкомпонентная модель.

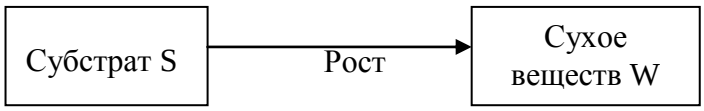

Рис. 1 – Замкнутая двухкомпонентная модель роста. Система замкнута, поскольку не имеет ни входов, 'ни выходов: предполагается, что в процессе роста преобразование материала первого компонента (субстрата) в материал второго компонента (сухое вещество) происходит без потерь Различные предположения относительно зависимости скорости процесса (темпа роста) от W и S приводят к различным формализациям.

Если допустить, что на рассматриваемом отрезке времени система не получает из внешней среды и не теряет никакого материала, то справедливо

$$\frac{dW}{dt} = -\frac{dS}{dt}; \qquad (3)$$

$$\frac{dW}{dt} + \frac{dS}{dt} = \frac{d}{dt}(W+S) = 0,$$

так что

$$W + S = const = W_0 + S_0 = W_f + S_f = C, \qquad (4)$$

где W_0 и S_0 – исходные значения W и S в момент t=0, W_f и S_f – значения, к которым приближаются эти параметры при $t \to \infty$ (в допущении, что система со временем приходит в стабильное состояние), C – постоянная величина.

Далее темп роста можно представить в виде некоторой функции v, такой, что

$$\frac{dW}{dt} = v(W,S). \qquad (5)$$

Поскольку из уравнения (4) следует, что $S=C-W$, путем подстановки S в уравнение (5) можно получить

$$\frac{dW}{dS} = v(W, C-W) = h(W), \qquad (6),$$

где h – функция только одной переменной W, как и в уравнении (5).

Таким образом, мы пришли к задаче с одной переменной состояния. Теперь предстоит решить кардинальный вопрос – какую функцию следует использовать в уравнении (5). Ответ на него определит вид уравнения (6) и после интегрирования – вид самого уравнения роста, то есть зависимости.

$$W = f(t), \qquad (7)$$

Где f – определённая функциональная связь.

Дальнейшие наши выводы будут опираться на различные предположения относительно функции v в выражении (5), то есть на несколько вполне определенных уравнений роста, каждое из которых можно интерпретировать в терминах модели, изображенной на рисунке 1.

Простой экспоненциальный рост с резкой отсечкой

Принятые допущения: количество энергии роста пропорционально количеству сухой массе W; механизм роста «работает» с максимальным темпом на протяжении всего времени, пока существует питательная среда; процесс роста необратим и прекращается, как только истощается питательная среда. Уравнение (5) приобретает вид

$$\frac{dW}{dt} = \mu W \qquad (7)$$

где μ — параметр, уже упоминавшийся под названием удельного или относительного темпа роста.

Параметр μ зависит, во-первых, от вида сухой массы W, соответствующей (в заданной пропорции) ресурсу питательной среды, и, во-вторых, от производительности или скорости, с которой «работает» механизм роста. Интегрирование уравнения (7) дает

$$W = W_0 e^{\mu t} \quad \text{при} \quad 0 \le t \le t_f \qquad (8a)$$

$$W = W_f \quad \text{при} \quad t > t_f \qquad (8b)$$

Когда $W = W_f, \quad S = 0$, то из уравнения (6) следует

$$W_f = W_0 + S_0, \qquad (9)$$

и рост внезапно прекращается, когда

$$t_f = \{ln[(W_0 + S_0)/W_0]\}/\mu \qquad (10)$$

(здесь приводится результат подстановки $t = t_f$ и $W = W_0 + S_0$ в выражение (8a).

Простой экспоненциальный рост, ограниченный реальным ресурсом питательной среды, графически представлен на рисунке 2*а*.

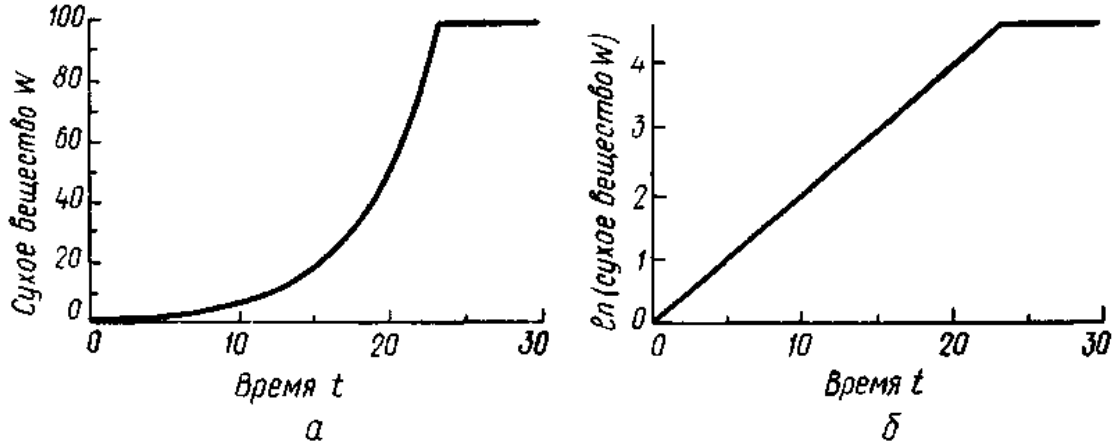

Рис. 2. Простой экспоненциальный рост с резкой отсечкой. Сухое вещество W и время t – в произвольных единицах. Кривые заданы выражением (8) при W_0=1, W_f=100, μ=0,2; t_f=23,0

При использовании полулогарифмического масштаба кривая роста преобразуется в прямую линию (рис. 2*б*), так как из выражения (8) следует, что

$$lnW = lnW_0 + \mu t \qquad (11)$$

Мономолекулярное уравнение

Это уравнение описывает ход простой необратимой химической реакции первого порядка. Принятые допущения: количество энергии роста неизменно и не зависит от количества сухой массы *W*; механизм роста «работает» со скоростью, пропорциональной ресурсу питательной среды *S*; рост необратим. В данном случае вместо уравнения (7) имеем

$$\frac{dW}{dt} = kS \qquad (12)$$

где *k* – постоянная величина

Пользуясь соотношением (4), произведем подстановку в уравнении (12), что при значениях $S(f)=0$ и $S = W_f - W$ дает

$$\frac{dW}{dt} = k(W_f - W) \qquad (13)$$

Интегрируя правую и левую часть

$$\int_{W_0}^{W} \frac{dW}{W_f - W} = \int_{0}^{t} kdt,$$

получаем

$$ln\left(\frac{W_f - W_0}{W_f - W}\right) = kt.$$

Последнее выражение можно записать в виде соотношения

$$W = W_f - (W_f - W_0)e^{-kt}. \qquad (14)$$

Если задаться исходной массой, $W_0 = 0$, то выражение (14) приобретает еще более простую форму

$$W = W_f(1 - e^{-kt}).$$

Зависимость (14) графически представлена на рисунке 3. Темп роста непрерывно падает, и кривая не имеет точки перегиба. В этом можно убедиться, исследуя вторую производную:

$$\frac{d^2W}{dt^2} = -k^2(W_f - W) = -(W_f - W_0)k^2e^{-kt},$$

которая обращается в нуль только при $t \to \infty$ либо при $W \to W_f$.

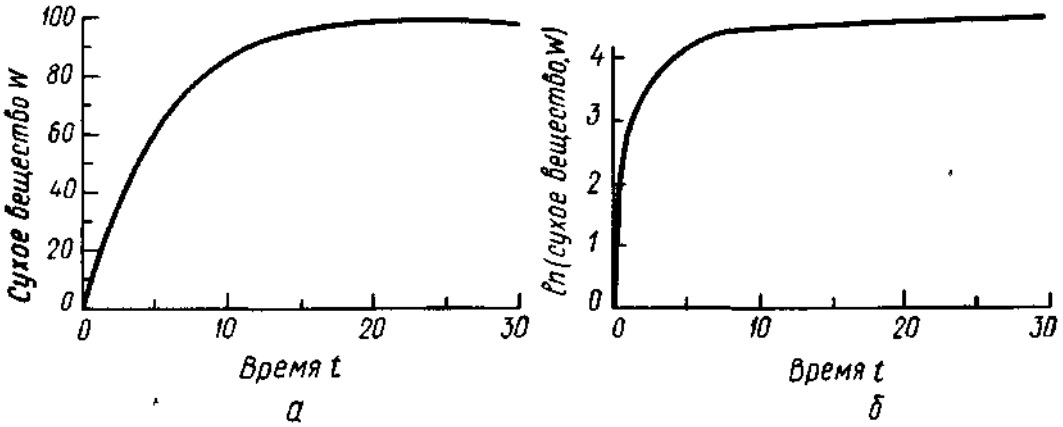

Рис. 3. Мономолекулярный рост. Сухое вещество *W* и время *t* – в произвольных единицах. Кривые заданы выражением (14) при $W_0 = 1, \quad W_f = 100, \quad k = 0{,}2.$

Уравнение логистического роста

В предыдущих разделах были рассмотрены две крайние ситуации: в уравнении (7) темп роста зависит только от энергии роста (которая пропорциональна количеству сухой массы *W)* и не зависит от ресурса питательной среды; в уравнении (13) темп роста зависит только от ресурса питательной среды и не зависит от сухой массы. При выводе уравнения логистического роста делается двоякое допущение: энергия роста пропорциональна сухой массе *W*; механизм роста «работает» со скоростью, пропорциональной ресурсу питательной среды *S*, процесс роста необратим. В соответствии с (7) и (12) при этих условиях справедливо уравнение

$$\frac{dW}{dt} = k^1 WS \qquad (15)$$

где k^1 – постоянная величина.

Подставляя далее из (4) $S = W_f - W$ и принимая $S_f = 0$, можно получить

$$\frac{dW}{dt} = k^1 W(W_f - W). \qquad (16)$$

Воспользовавшись соотношением

$$k^1 = \mu / W_f \qquad (17)$$

можно перейти от константы k^1 к уже известному нам параметру μ, то есть

$$\frac{dW}{dt} = \mu W\left(1 - \frac{W}{W_f}\right) \qquad (18)$$

Полученное выражение с привлечением метода элементарных дробей можно переписать в виде

$$\int_{W_0}^{W} \left(\frac{1}{W_f - W} + \frac{1}{W}\right) dW = \int_0^t \mu dt$$

и затем после интегрирований и элементарных преобразований получить

$$W = \frac{W_0 W_f e^{\mu t}}{W_0 + (W_f - W_0) e^{-\mu t}} \qquad (19)$$

Чаще логистическое уравнение записывают в форме

$$W = \frac{W_0 W_f}{W_0 + (W_f - W_0) e^{-\mu t}} \qquad (20)$$

Анализ любого из двух последних выражений показывает, что при $W_0 \ll W_f$ для малых значений t (подстановка $W_0 = 0$ в знаменатель) справедливо приближенное равенство

$$W \approx W_0 e^{-\mu t} \qquad (21)$$

из которого следует, что при указанных условиях имеет место экспоненциальный рост с начальным темпом, равным μ. Так как $t \to \infty$ и $W \to W_f$, процесс роста носит асимптотический характер.

Дифференцирование (18) дает

$$\frac{1}{\mu}\frac{d^2W}{dt^2} = \frac{dW}{dt}\left(1 - \frac{2W}{W_f}\right)$$

Приравнивая правую часть нулю, убеждаемся, что существует точка перегиба при

$$W = \frac{1}{2} W_f \qquad (22)$$

Подстановкой последнего выражения в уравнение (20) можно установить, что точка эта соответствует моменту времени $t = t^*$, где

$$t^* = \frac{1}{\mu} ln\left(\frac{W_f - W_0}{W_0}\right) \qquad (23)$$

Соответствующая кривая роста изображена на рисунке 4. Логистическое уравнение – первое среди рассмотренных здесь, дающее кривую с точкой перегиба.

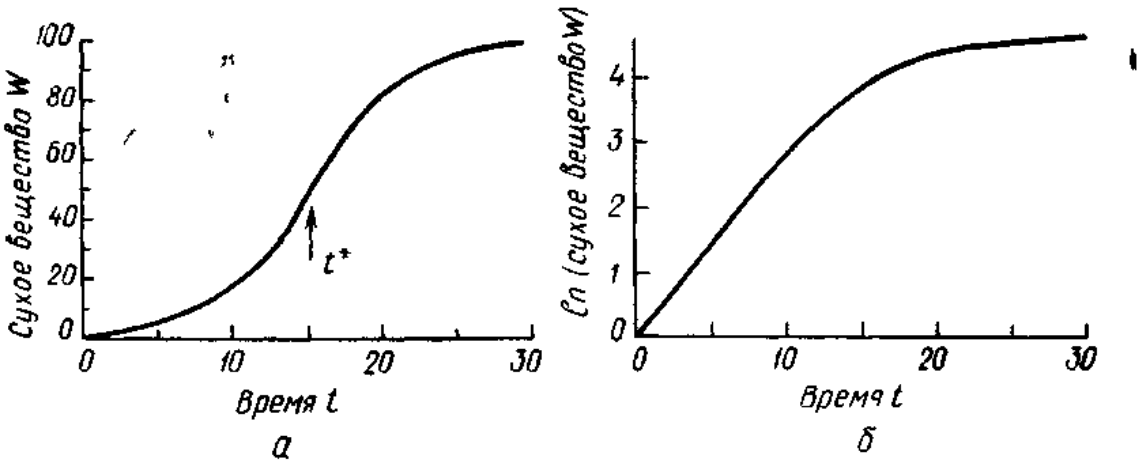

Рис 4. Логистический рост. Сухое вещество W и время t – в произвольных единицах Кривые заданы выражением (20) при $W_0 = 1,\ W_f = 100,\ \mu = 0{,}3$. Через t^* обозначена точка перегиба согласно выражению (23) $t^* = 15{,}32$.

2. МОЛОЧНОЕ ХОЗЯЙСТВО

Молочной ферме, при всей ее значимости в сельском хозяйстве, разработчики моделей уделили относительно мало внимания. Это едва ли покажется удивительным, если принять в расчет, что молочное производство – самый многофункциональный из всех видов животноводческой деятельности, в рамках которого необходимо заниматься и управлять такими сложными физиологическими процессами, как беременность, лактация, формирование тканей тела.

Одна из первых публикаций по указанной тематике принадлежит Крэбтри [3] и содержит описание имитационной модели (запрограммированной на языке высокого уровня DYNAMO) интенсивного молочного хозяйства. В ходе имитации реализуются подмодели выращивания трав, энергетических и кормовых потребностей коровы, структуры стада и движения денежной наличности. Описания этих компонентов модели строятся с помощью эмпирических соотношений либо заимствуются из литературы, либо конструируются непосредственно из экспериментальных данных.

Байуотер и Дент [4] разработали методическое руководство по конструированию детализированной модели, предназначенной для имитации процессов усвоения и перераспределения питательных веществ в организме дойной коровы. Цель моделирования – помочь в управлении молочным стадом. Упомянутая работа расценивалась авторами как стартовая в серии статей, посвященных подробному описанию различных компонентов модели.

Байуотер [5] во второй статье этой серии предложил схему для моделирования надоев и состава молока, основанную на анализе литературных источников. Никаких материалов по апробации этой подмодели не имеется. Насколько нам известно, последующие работы объявленной серии пока не появились.

Чудлей [6] описывает простую экономическую модель небольшой молочной фермы в Кении. Он полагает, что эта модель может быть использована как составной элемент в анализе затрат и выпуска, предпринимаемом с целью оптимизации распределения ресурсов на исследования и проектирование по программе повышения эффективности мелких ферм.

Модели, призванные сыграть роль помощников в управлении молочными хозяйствами, оказываются особенно эффективными, когда их можно реализовать прямо на ферме, пользуясь программируемым калькулятором. Среди публикаций, посвященных этой проблеме, следует отметить работы Вуда [7], а также Франса и др. [8].

Вуд [7] предлагает модель для расчета кормовой базы, исходя из типа молочного производства на ферме и планируемых приростов живой массы.

Франс и др. [8] описали модель, которая позволяет планировать объем производства молока, доход от его продажи, потребность в концентрированных кормах, сравнивать экономический эффект от применения различных схем отела. В качестве лактационной кривой выбран график неполной гамма-функции; эта функция используется также для расчета производства молока в течение всего лактационного периода.

Линн и Спайк [9] выпустили обзор по использованию программируемых калькуляторов при решении задач кормления и содержания молочного скота, а Кирк [10] опубликовал работу, где дана оценка эффективности использования этих устройств в рамках программы борьбы с маститом в США.

Важным фактором в производстве молока является сезонность. Регулярность снабжения молоком зависит главным образом от распределения дат отелов, которые, в свою очередь, подвержены влиянию сезонных факторов, в особенности такого, как доступность качественной травы. Вуд [11, 12, 13, 14, 15] с помощью регрессионного анализа исследовал сезонные вариации надоев по большой выборке данных лактации, применив при этом в качестве аппроксимирующей кривой кривую грамма-функции. Результаты Вуда свидетельствуют о том, что коровы зимнего отела в среднем вносят в суммарный надой больший вклад, чем коровы весеннего отела, что суточный надой падает в зимние месяцы и

возрастает весной, что для концентраций жира и белка также характерны сезонные изменения.

Поставщики молока и агенты по его сбыту сталкиваются со стабильным спросом на питьевое молоко и скоропортящиеся молочные продукты и, для того чтобы обеспечить приемлемую стабильность поступлений молока на свои перерабатывающие предприятия, закупают его у фермеров по сезонным ценам. Киллен и Кин [16] построили модель линейного программирования, предназначенную для оптимизации распределения дат отелов с тем, чтобы удовлетворить спрос на продукты и установить твердые сезонные цены на молоко, которые были бы приемлемы для фермеров. Постановка задачи выглядит следующим образом:

Минимизировать

$$\sum_{j=1}^{12} c_j x_j + d \sum_{j=1}^{12} \left(\sum_{j=1}^{12} a_{ij} x_j - m_i \right)$$

с учётом

$$\sum_{j=1}^{12} a_{ij} x_j \geq m_i, \quad i = 1, 2, \cdots, 12.$$

где x_j – общий удой в *j*-й системе, включающей в себя коров, отелившихся в *j*-м месяце, л/год, c_j – переменные издержки производства, связанные с отелом коров в *j*-м месяце, фунт стерлингов/л, *d* – затраты на реализацию избыточной продукции, ф. ст./л; a_{ij} – доля молока, надоенного в *i*-и месяц от коровы, отелившейся в *j*-м месяце, m_i– спрос в *i-м* месяце, л.

Киллен и Кин значения a_{ij} заимствовали у Вуда [13]; естественно, что

$$\sum_{j=1}^{12} a_{ij} = 1, \qquad j = 1, 2, \cdots, 12.$$

Решение этой задачи линейного программирования дает такие значения $\mathbf{x}_j$, которые соответствуют оптимальному объему потребления от каждой из 12 систем. Для получения рациональной системы сезонных цен Киллен и Кин решали двойственную задачу:

максимизировать

$$\sum_{j=1}^{12} m_i y_i - d \sum_{j=1}^{12} m_i$$

с учетом

$$\sum_{j=1}^{12} a_{ij} y_i \leq \mathrm{c}_j + d, \qquad j = 1, 2, \cdots, 12,$$

где $y_i \geq 0$, фунт стерлингов/л – двойственные переменные.

Величины y_i^* могут интерпретировать такие оценки месячного производства, которые максимизируют стоимость полного потребления молока агентством сбыта (с более общей интерпретацией двойственных переменных можно познакомиться в работе Таха [17]). Рассчитанные значения y_i^* затем используют в качестве базиса для выбора системы сезонных цен: цены, назначаемые ежемесячно для закупок у фермера, p_i ф. ст/л, при этом определяются как

$$p_i = y_i^* + k, i = 1, 2, \cdots, 12,$$

где *k* — фиксированный элемент стоимости каждого литра молока, а y_i^* – надбавка, выплачиваемая фермеру в месяце *i*.

Другие примеры использования линейного программирования в молочном производстве приводятся в работе Рейеса и др. [18], где описана многопериодная модель линейного программирования, используемая для оценки альтернативных режимов

кормления и содержания животных на протяжении всего производственного цикла (интервала между отелами) на молочных фермах северо-восточного Техаса.

3. ПРОИЗВОДСТВО ГОВЯДИНЫ

Опубликованные к настоящему моменту модели по системам производства говядины можно разбить на две категории – те, которые опираются на технику компьютерной имитации, и те, в основе которых лежат оптимизационные методы, в частности линейное программирование. Степень общности таких моделей колеблется от уровня отдельного предприятия до отрасли на национальном уровне. В данном разделе будут рассмотрены работы по моделированию систем производства говядины. Обзор более ранних работ содержится в статье Жоандета и Картрайта [19].

Большинство из описанных в литературе моделей — имитационного типа. Райен [20] описывает имитационную модель группы животных, проходящих через откормочную площадку. Программирование модели осуществлялось на языке FORTRAN, а в основу ее структуры положены суточные изменения живой массы группы животных. Модель предназначена для оценки альтернативных стратегий, в том числе критериев принятия решения о продаже продукции, цен на выбракованных животных, а также эффективность альтернативных форм хозяйствования, отражающейся в соответствующих показателях роста и смертности животных. Кроме того, рассматривались последствия изменений рыночных цен на говядину, стоимость кормов.

Холтер и др. [21] также предложили имитационную модель, предназначенную для сопоставления различных стратегий стимулирования фермеров Венесуэлы к преобразованию их хозяйств из традиционных в современные высокоприбыльные предприятия.

Морли [22] описывает модель откорма бычков-кастратов в возрасте 2-4 лет в условиях, специфичных для срединных, то есть удаленных от моря земель Аргентины. Для аналогичных природных условий в Австралии построена модель [23] для исследования влияния суперфосфата на рост травы при различных комбинациях концентрации скота на пастбище, цен на говядину и времени реализации, а также с учетом погодных условий (средние, неблагоприятные). А модель Уайта и др. [24] опирается на технику линейного программирования для анализа альтернатив развития животноводства на промышленной основе в Гвиане.

Сопоставление альтернативных технологий производства говядины и стратегий их реализации в регионах выгульных пастбищ Южной Африки осуществляется с помощью модели Лау, Гросскопфа и Грёневальда [25], которая имитирует формирование прибыли в четырех системах – от прямой продажи телят-отъемышей до организации сбыта бычков-трехлеток. Уравнения регрессии, используемые в модели, соотносят норму тела (в процентах) и массу телят при отъеме с концентрацией скота на пастбище и количеством осадков.

Модели Сандерса и Картрайта [26, 27] имитируют функционирование предприятия по производству говядины в широком диапазоне условий внешней среды и форм хозяйствования для различных генотипов крупного рогатого скота; модель нашла приложение в управлении предприятиями США, Южной Америки и Африки.

Конглетоном и Гудвиллом [28, 29, 30] описана имитационная система, моделирующая производство говядины и запрограммированная на алгоритмическом языке DYNAMO. Компоненты модели имитируют динамику поголовья в молочном стаде и продуктивном стаде, составляемом из телят молочного стада. С её помощью можно разрабатывать планы случек скота, оценивать стратегии выбраковки, осуществлять обоснованный выбор, например, содержать молочных телок или продавать их.

Примером стохастической имитационной модели, используемой при принятии решений, может служить система Бека и Диллона [31], предназначавшаяся для выбора

стратегии при ликвидации бруцеллеза крупного рогатого скота. В качестве исходных использовались данные о поголовье скота, распространенности заболевания и эффективности принимаемых мер в данном районе, а оценке подлежали требуемый объем работ, нормативы выбраковки и продолжительность противобруцеллезной кампании при различных стратегиях ее проведения.

Салливен, Картрайт и Фаррис [32] описали способ интеграции двух моделей – кормопроизводства и скотоводства – с целью имитации процесса получения мясной и молочной продукции в тропических районах Восточной Африки. Задача состояла в изучении влияния прогрессивных способов содержания скота на трансформацию традиционной деревенской системы выпаса.

Бек, Харрисон и Джонстон [33] разработали модель, позволяющую прогнозировать прибыли и оценивать вероятности риска для предприятия, специализирующегося по выгульному мясному скотоводству. Особое внимание уделено моделированию влияния на прибыли (убытки), которое оказывает концентрация скота на пастбище. Работоспособность модели иллюстрируется примерами её приложения к управлению хозяйствами Нового Южного Уэльса (Австралия).

Большинство опубликованных в литературе оптимизационных моделей систем производства говядины – это большие задачи линейного программирования (ЛП). Уилтон и др. [34] приводят описание одной из таких ЛП-моделей, предназначенных для оптимизации интегрированной системы производства говядины. Модель содержит блоки растениеводства, откорма и селекции, где формализуются соответствующие требования к площадям, трудовым ресурсам, помещениям для содержания животных и хранения кормов. С помощью упомянутой ЛП-модели был проведен анализ трех различных по размеру предприятий по выращиванию чистопородных коров, которые содержались в условиях, типичных для провинции Онтарио. Моррис, Паркинс и Уилтон [35] использовали эту же модель для оценки влияния, которое оказывают на экономическую эффективность предприятия различия в живой массе коров, уровнях потенциальных надоев и режимах кормления.

Миллер, Бринкс и Сазерленд [36] предложили ЛП-модель для выработки стратегии управления типичным для Южного Колорадо горным ранчо на предстоящий год. Модель рассчитывает значения показателей по каждому виду деятельности (например, размеры молочного стада и стада бычков-годовиков, площадь лугов под сенокосами и т. д.), которые максимизируют чистую прибыль хозяйства с учетом ограничений на земельные, трудовые и финансовые ресурсы. Целесообразность расширения предприятия оценивается путем исследования оптимального решения методами параметрического программирования. Модель содержит 127 ограничений и 274 переменных. Еще более высокий уровень общности подхода при большей продолжительности планируемого периода продемонстрирован в предложенной Миллером и др. [37] ЛП-модели системы производства говядины в США, задача которой – определение такой кормовой базы и таких режимов кормления мясного скота, которые максимизировали бы объем первичного производства говядины в течение планируемого пятилетия. США представлены в модели пятью производящими регионами с межрегиональной транспортировкой телят, когда этого требуют оптимальные планы кормления. Йорке и др. [38] использовали эту модель для сравнения результатов оптимизации производства говядины в США по двум разным критериям — минимуму затрат на откорм и минимуму затрат энергии, получаемой за счет полезных ископаемых.

Примеров использования методов оптимизации, которые опирались бы не на линейное программирование, в литературе мало. Мейер и Ньюэт [39], а также Кеннеди [40] предлагают находить оптимальные значения живой массы закупаемого и продаваемого скота при содержании его на откормочных площадках, а также оптимальные режимы кормления с помощью аппарата динамического программирования. Конвей [41, 42] предложил модель выбора экономического решения, максимизации превышения дохода над переменными затратами при содержании бычков-кастратов на подножном корме; модель использует аппарат производственных функций.

4. ВЫБРАКОВКА

Моделирование применяется также для выработки стратегий ремонта в непрерывно функционирующих животноводческих предприятиях, как, например, в случае замены коров в молочном стаде. Данная проблема аналогична той, которая возникает в промышленности при выработке оптимальной стратегии обновления станочного парка на заводе. Гартнер и Герберт [43], а также Уолсингам, Эдельстен и Брокингтон [44] применили имитационные модели для исследования, выбраковки и ремонта молочного стада и популяции кроликов. Проблема ремонта хорошо поддается решению с помощью метода динамического программирования, и этим подходом воспользовались Уайт [45] для описания ремонта стада несушек на предприятии по производству яиц, а также Мак-Артур [46], Смит [47], Стьюарт и др. [48] для исследования стратегий ремонта молочных стад. Этот подход ниже иллюстрируется сильно упрощенным, но показательным примером, заимствованным в работе Тросби [49].

Пусть g_i и h_i обозначают соответственно годовой доход и затраты, связанные с содержанием молочной коровы в возрасте *i (i=0, 1, 2, ...)*, причем возраст отсчитывается от начала продуктивной жизни (*i=0*). Пусть s_i, есть доход от выбраковки коровы в возрасте *i* и пусть ремонт осуществляется только в момент, близкий к началу ее продуктивной жизни (то есть в возрасте *i=0*). Каждый год возможно два исхода: *K – сохранение* в стаде, а *R – выбраковка и ремонт*. Определим f_i, как оптимальную величину до хода от содержания коровы в возрасте i. Примем, что $g_i = h_i$ и $f_i = 0$ для $i > 5$. Тогда $f_i \geq 0$ для всех *i*. Рекуррентное соотношение, определяющее f_i имеет вид

$$f_i = max\ K, R \begin{bmatrix} K: g_i - h_i + f_{i+1} \\ R: s_i + g_0 - h_0 + f_1 \end{bmatrix}.$$

Допустим, что g_i, h_i и s_i, приняли следующие значения:

i	g_i	h_i	s_i
0	4	2	2
1	7	3	3
2	10	3	5
3	9	4	7
4	6	4	9
5	5	4	6

Приложение рекуррентных соотношений дает

$$f_5 = max \begin{bmatrix} K: 5 - 4 + 0 \\ R: 6 + 4 - 2 + f_1 \end{bmatrix} = max \begin{bmatrix} 1 \\ 8 + f_1 \end{bmatrix} = 8 + f_1 \rightarrow \text{выбраковка и ремонт}$$

$$f_4 = max \begin{bmatrix} K: 6 - 4 + f_5 = 8 + f_1 \\ R: 9 + 4 - 2 + f_1 \end{bmatrix} = max \begin{bmatrix} 10 + f_1 \\ 11 + f_1 \end{bmatrix} = 11 + f_1 \rightarrow \text{выбраковка и ремонт}$$

$$f_3 = max \begin{bmatrix} K: 9 - 4 + f_4 = 11 + f_1 \\ R: 7 + 4 - 2 + f_1 \end{bmatrix} = max \begin{bmatrix} 16 + f_1 \\ 8 + f_1 \end{bmatrix} = 16 + f_1 \rightarrow \text{сохранение}$$

Таким образом, решение заключается в выбраковке и замене коровы после того, как она прожила четыре продуктивных года (то есть когда i=4). Заметим, что рекуррентное соотношение вычисляет f_i из f_{i+1}, и поэтому порядок вычислений $f_5 \rightarrow f_4 \rightarrow f_3$.

5. РАЦИОНЫ КОРМЛЕНИЯ

Составление рациона – важный аспект управления многими животноводческими предприятиями, в особенности, если там приняты системы интенсивного откорма, когда животным скармливают в больших количествах покупные продукты. Проблема очень напоминает составление диеты минимальной стоимости в госпиталях и других учреждениях и составление кормовых смесей для сельскохозяйственных животных [50, 51]), где цель заключается в нахождении самой дешевой комбинации продуктов, которая вместе с тем содержала бы необходимые питательные вещества в нужных количествах.

Составляя рацион, фермер обычно старается так определить количество каждого продукта, подлежащего ежедневной раздаче животным, чтобы требования, предъявляемые к его энергетической и питательной ценности, удовлетворялись с минимальными издержками. Что касается жвачных, то для них упомянутая проблема формализуется следующим образом:

Минимизировать целевую функцию:

$$z \sum_{j=1}^{n} c_j x_j$$

с учетом ограничений:

1. По энергии:

$$\sum_{j=1}^{n} e_j x_j \geq E_{req}$$

2. По белку:

$$\sum_{j=1}^{n} pe_j x_j \geq P_{req}$$

3. По усвоению:

$$\sum_{j=1}^{n} x_j \leq F_{max}$$

4. По условию не отрицательности:

$$x_j \geq 0 \text{ для всех } j.$$

где n – число продуктов, доступных фермеру; x_j – количество продукта j, скармливаемого животному, кг сухого вещества (СВ) в сутки; c_j – стоимость продукта j, руб./кг СВ; e_j – концентрация энергии в j-м продукте, МДж/кг СВ; p_j – концентрация белка в j-м продукте, г/белка/кг СВ; E_{req} – энергетическая потребность животного, МДж/сут.; P_{req} – потребность животного в белке, г/сут; F_{max} – максимальное потребление корма животным, кг/сут. СВ.

Исследуемая проблема может выдвинуть и другие ограничения, такие как потребность жвачных в определенных минеральных веществах и минимально допустимое содержание в рационе сырой клетчатки.

В Великобритании при составлении рационов потребность в энергии и ее концентрацию в настоящее время принято выражать в единицах метаболической энергии (МЭ), то есть той доли энергии, содержащейся в корме, которая может быть утилизована жвачным животным (с учетом потерь с калом, мочой и метаном, рассчитываемых по методике Министерства сельского хозяйства, рыболовства и продовольствия Великобритании [53]), а потребность в белках и их концентрацию – в единицах сырой клетчатки (сырая клетчатка = азот·6,25). Поэтому главное внимание здесь будет сосредоточено на методиках составления рационов, опирающихся на эту общепринятую в Великобритании систему.

Задача составления рационов, если она формализована с помощью уравнений приведённых выше, лучше всего решается методами линейного программирования, определяющими значения переменных x_j, когда величины правой части – константы. К сожалению, так бывает редко; как правило, эти величины представляют собой функции от живой массы животных и от объема производства. Упомянутую трудность обычно преодолевают, рассчитывая группы рационов в соответствии с единичными приростами живой массы и объемов производства и полагая величины правой части константами для каждой такой «ступеньки».

Очередная проблема возникает в связи с тем, что и требования к метаболической энергии (МЭ) и уровень максимального потребления корма могут изменяться в зависимости от концентрации МЭ в рационе r, МДж/кг сухого вещества. Эффективность расходования МЭ на рост животного претерпевает ощутимые изменения при переходе от одного рациона к

другому (смена *r),* когда животное находится в фазе роста или на откорме. Преодолеть эту трудность можно, выражая соответствующие ограничения в единицах не метаболической, а чистой энергии по методике Харкинса, Эвардса и Мак-Дональда [54]. Чистая энергия определяется как МЭ, из которой вычтен прирост тепла, то есть это та часть энергии, содержащейся в корме, которая расходуется животным на поддержание жизнедеятельности и рост. Система с переменной чистой энергии широко используется при составлении рационов; читателю, желающему глубже ознакомиться с системой, следует обратиться к отчету Министерства сельского хозяйства, рыболовства и продовольствия Великобритании [52].

Было бы неразумно пренебречь при анализе проблемой максимального потребления. Максимальный для жвачного животного уровень потребления корма *I*, кг сухого вещества на 1 кг живой массы, как функция концентрации энергии в рационе *r* достаточно точно формализуется с помощью уравнения равносторонней гиперболы

$$I(r) = k_1 \frac{r}{r + k_2} - k_3,$$

где k_1, k_2 и k_3 – положительные числа.

Поскольку

$$r = \sum_{j=1}^{n} e_j x_j \Big/ \sum_{j=1}^{n} x_j,$$

проблема составления рациона теоретически поддается решению методами сепарабельного программирования. Однако на практике этот подход применяется редко: как правило, в рассматриваемой ситуации пользуются методом проб и ошибок. Максимальное потребление корма F_{max} рассчитывают по произведению *wI(r),* где *w* – живая масса, кг, животного. Начальное значение F_{max} соответствует, к примеру, *r*=10 МДж/кг сухого вещества, то есть $F_{max} = wI(r = 10)$, и дальнейшее решение разворачивается по схеме линейного программирования. Пусть концентрация МЭ окончательного рациона есть r_1. Если потребление этого рациона $\sum x_j$ больше либо равно $wI(r = r_1)$, то тогда F_{max} устанавливают равным $wI(r = r_1)$ и производят повторный цикл решения по схеме линейного программирования. Процедура повторяется до тех пор, пока не будет найден оптимальный рацион, удовлетворяющий критерию

$$\sum_{j=1}^{n} x_j \leq wI\left(r = \sum_{j=1}^{n} e_j x_j \Big/ \sum x_j\right).$$

Чтобы повысить вычислительную эффективность этого подхода, можно прибегнуть к параметрическому программированию.

Лекция 8.
СЕПАРАБЕЛЬНОЕ ПРОГРАММИРОВАНИЕ

Сепарабельное* программирование, разработанное Миллером [55] , представляет собой метод нелинейного программирования, который может быть применен для нахождения глобального или локального оптимума в целом ряде нелинейных задач. Он допускает использование в целевой функции и в ограничениях линейной в остальном задачи нелинейных функций одной переменной (например, $x_1, \ln x_2, \ \exp x_3$).

(* Происхождение термина связано с понятием сепарабельной функции, то есть нелинейного выражения, которое можно представить в виде суммы *n* функций,
по одной для каждой из переменных x_j:

$$c_x = \sum_{j=1}^{n} c_j (x_j).$$

Сепарабельное программирование приложимо к задачам, в которых все нелинейные функции сепарабельны.)

Требование о том, чтобы нелинейные члены были функциями одной переменной, не слишком обременительно, поскольку, например, нелинейности типа произведений можно с помощью простых преобразований заменять суммами и разностями функций одной переменной. Например, член $x_1 x_2$ можно заменить разностью $u_1^2 - u_2^2$, где

$$u_1 = (x_1 + x_2)/2 \text{ и } u_2 = (x_1 + x_2)/2$$

а $x_1^{a_1} x_2^{a_2} x_3^{a_3} \cdots x_n^{a_n}$ – заменяется на e^u,

где

$u = a_1 lnx_1 + a_2 lnx_2 + a_3 lnx_3 + \cdots + a_n lnx_n$.

После того как все нелинейности преобразованы в функции одной переменной, каждую из этих функций можно заменить ее кусочно-линейной аппроксимацией. Рассмотрим нелинейную функцию *f(v)*, изображенную на рисунке 5.

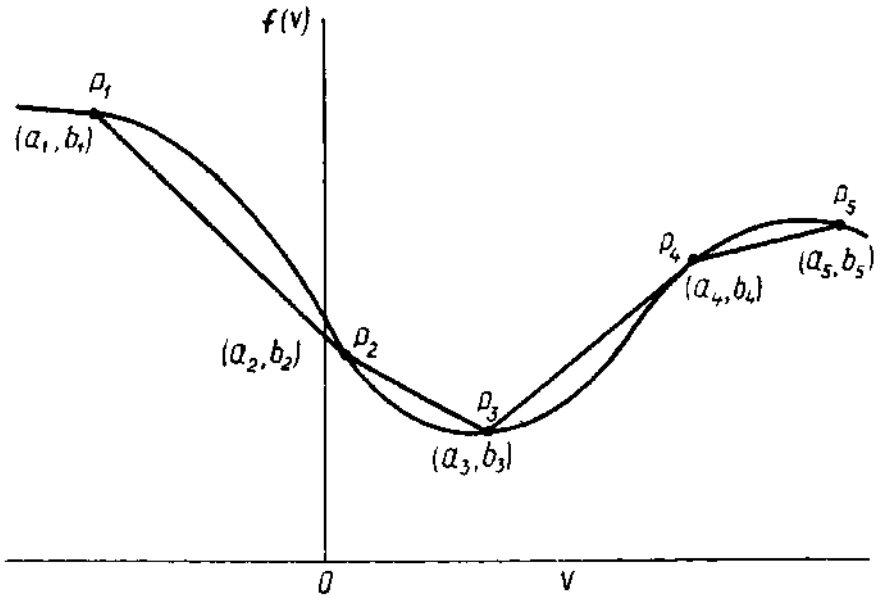

Рис. 5 Кусочно-линейная аппроксимация нелинейной функции одной переменной.

Там же показана кусочно-линейная аппроксимация функции *f(v)* в виде ломаной $P_1P_2P_3P_4P_5$, где координаты точек $P_i (i = 1,2, \dots, 5)$ обозначены (a_i, b_i). Формально она описывается так:

$$\sum_{i=1}^{5} \lambda_i = 1 \qquad (c.1)$$

$$\sum_{i=1}^{5} a_i \lambda_i = v \qquad (c.2)$$

$$\sum_{i=1}^{5} b_i \lambda_i = f_A(v) \qquad (c.3)$$

где $f_A(v)$ – аппроксимация для $f(v)$.

$$\lambda_i \geq 0 \text{ для } i = 1,2,\dots,5 \qquad (c.4)$$

Вершины кусочно-линейной аппроксимации для $f(v)$ получаются, когда в соотношениях (с.1–с.3) все λ_i кроме одной, равны нулю. Любая другая точка аппроксимации определяется, когда две смежные переменные λ принимают положительные значения, в то время как остальные приравниваются нулю. Не требуется, чтобы *v* и *f(v)* были неотрицательными, точки P_i не обязательно должны быть одинаково расположены, а интервал $(a_1\ a_5)$ может быть и больше исследуемого.

В сепарабельном программировании λ_i называют *специальными переменными,* а связи (с.1) и (с.2) – *условиями выпуклости и привязки* для них. Итак, нелинейная задача может быть линеаризована путем подстановки вместо каждой нелинейности *f(v)* её кусочно-линейной аппроксимации. При этом множество специальных переменных включается в общий список переменных, а условия выпуклости и привязки для каждой нелинейности – в число ограничений задачи.

Линеаризованную задачу теперь можно решать по схеме линейного программирования, дополнив ее специальными проверками, к которым следует прибегать, когда специальные переменные рассматриваются как возможные «кандидатуры» для введения в базис. Эти проверки должны гарантировать, что из каждого набора λ_i в базис войдёт не более двух специальных переменных и что члены этой пары будут смежными.

ЛИТЕРАТУРА

1. Доспехов Б. А. Методика полевого опыта (с основами статистической обработки результатов исследований) – 5-е изд., доп. и перераб. – М.: Агропромиздат, 1985. – 351 с.

2. Урбах В.Ю. Биометрические методы (статистическая обработка опытных данных в биологии, сельском хозяйстве и медицине). – М.: «Наука». – 1964. – 416 с.

3. Crabtree J.R. Towards a dairy enterprise model. In: The Use of Models in Agricultural and Biological Research (ed. J.G. W. Jones). – 1970. – pp. 35-42. – Hurley, Maidenhead: Grassalnd Research Institute.

4. Bywater A.C. and Dent J.B. Simulation of the intake and partition of nutrients by the diary cow: part I – management control in the diary enterprise: philosophy and general model construction. – Agricultural Systems, 1. – 1976. – pp. 245-260.

5. Bywater A.C. Simulation of the intake and partition of nutrients by the dairy cow: part II – the yield and composition of milk. – Agricultural Systems, 1. – 1976. – pp. 261-279.

6. Chudleigh P. A model of the small scale enterprise: an aid to resource allocation in agricultural development. – Agricultural Systems, 2. – 1977. – pp. 67-82.

7. Wood P.D.P. A simple model of lactation curves for milk yield, food requirement and body weight. – Animal Production, 28. – 1979. – pp. 55-63.

8. France J., Heal H.D. St. C., Marsden S. and Frost B. A dairy herdcash flow model. – Agricultural Systems, 8. – 1982. – pp. 129-142.

9. Linn J.G. and Spike P.L. Programmable calculators and their application to feeding and management of dairy cattle. – Journal of Dairy Science, 63. – 1980. – p. 1390.

10. Kirk J.H. Application of programmable calculators to mastitis control programs. – Journal of Dairy Science, 64. – 1981. – pp. 2048-2058.

11. Wood P.D.P. Algebraic models of the lactation curves for milk, fat and protein production, with estimates of seasonal variation. – Animal Production, 22. – 1976. – pp. 35-40.

12. Wood P.D.P. Breed variations in the shape of the lactation curve of cattle and their implications for efficiency. – Animal Production, 31. – 1980. – pp. 133-141.

13. Wood P.D.P. Factors affecting the shape of the lactation curve in cattle. – Animal Production, 11. – 1969. – pp. 307-316.

14. Wood P.D.P. A note on regional variations in the seasonality of milk production in dairy cattle. – Animal Production, 32. – 1981. – pp. 105-108.

15. Wood P.D.P. The relationship between the month of calving and milk production. – Animal Production, 12. – 1970. – pp. 235-239.

16. Killen L. and Keane M. A linear programming model of seasonality in milk production. – Journal of the Operational Research Society, 29. – 1978. – pp. 625-631.

17. Taha H.A. Operations Research – an Introduction. New York: Macmillan. – 1971.

18. Reyes A.A., Blake R.W., Shumway C.R. and Long J.T. Multistage optimization model for dairy production. – Journal of Dairy Science, 64. – 1981. – pp. 2003-2016.

19. Joandet G.E. and Cartwright T.C. Modelling beef production systems. –Journal of Animal Science, 41. – 1975. – pp. 1238-1246.

20. Ryan T.J. A beef feedlot simulation model. – Journal of Agricultural Economics, 25. – 1974. – pp. 265-276.

21. Halter A.N., Miller S.F., De Jongh R. and Estrada H.R. Application of systems-simulation in the Venezuelan cattle industry. – Agricultural Systems, 1. – 1976. pp. 139-162.

22. Morley F.N.W. In: Applications in Agricultural Modelling (ed. A.J. de Boer and C.W. Rose). – 1977. – pp. 131-146. – Brisbane: Australian Institute of Agricultural Science.

23. Curll M.L. Simulation: an aid to decisions on superphosphate use for beef production. – Agricultural Systems, 3. – 1977. – pp. 195-204.

24. Kelley White T., McCarl B.A., May R.D. and Spreen T.H. A systems analysis of the Guvanese livestock industry. – Agricultural Systems, 3. – 1978. – pp. 47-66.

25. Louw A., Grosskopf and Groenewald J.A. Beef production systems and sales strategies in an extensive ranching in South Africa. – Agricultural Systems, 4. – 1979. – pp. 101-114.

26. Sanders J.O. and Cartwright T.C. A general cattle production systems model. I: Structure of the model. – Agricultural Systems, 4. – 1979. – pp. 217-227.

27. Sanders J.O. and Cartwright T.C. A general cattle production systems model. Part 2 – procedures user for simulating animal performance. – Agricultural Systems, 4. – pp. 289-309.

28. Congleton Jr.W.R. and Goodwill R.E. Simulated comparisons of breeding plans for beef production – part 1: a dynamic model to evaluate the effects of maung plan on herd age structure and productivity. – Agricultural Systems, 5. – 1980. – pp. 207-219.

29. Congleton Jr.W.R. and Goodwill R.E. Simulated comparisons of breeding plans for beef production – part 2: Hereford, Angus and Charolais sites bred to Hereford – Angus dams to produce feeder calves. – Agricultural Systems, 5. – 1980. –pp.225-232.

30. Congleton Jr.W.R. and Goodwill R.E. Simulated comparisons of breeding plans for beef production – part 3: systems for producing feeder-calves involving intensive culling and additional breeds of sire. – Agricultural Systems, 5. – 1980. –pp. 309-318.

31. Beck A.C. and Dilon M. Brucellosis cradication planning an application of simulation modeling. – Agricultural Systems, 3. – 165-180.

32. Sultivan G.M., Cartwrite T.C., Farris D.E. Simulation of production systems in East Africa by use of interfaced farrago and cattle models. – Agricultural Systems, 7. – 1981. –pp. 245-265.

33. Beck A.C., Harrison J. and Johnston J.H. Using simulation to assess the risks and returns from pasture improvement for beef production in agriculturally underdeveloped regions. – Agricultural Systems, 8. – 1982. – pp.55-71.

34. Wilton J.W., Morris C.A., Jensen E.O. and Pfeiffer W.C. A linear programming model for beef cattle production. – Canadian Journal of Animal Science, 54. – 1974. – pp. 693-707.

35. Morris C.A., Parkins J.J. and Wilton J.W. Effects of greep feeding, nature cow weight and milk yield on farm gross margins in an integrated beef production model. – Canadian Journal of Animal Science, 56. – 1976. – pp. 87-97.

36. Miller W.C., Brinks J.S. and Sutherland T. M. Computer assisted management decisions for beef production systems. – Agricultural Systems, 3. – 1978. –pp. 147-158.

37. Miller W.C., Ward G.M., Yorks T.P., Rossiter D.L. and Combs J.J. A mathematical model of the United States beef production systems. – Agricultural Systems, 5. – 1980. – pp. 295-307.

38. Yorks T.P., Miller W.C., Combs J.J. and Ward G.M. Energy minimized v. cost minimized alternatives for the US beef production system. – Agricultural Systems, 6. – 1980. – pp. 121-129.

39. Meyer C.F. and Newett R.J. Dynamic programming for feedlot optimization. – Management Science, 16. – 1970. –pp. 410-426.

40. Farazdaghi H. and Harris P.M. Plant competition and crop yield. – Nature, London, 217. – 1968. – pp. 289-290.

41. Conway A.G. A production Function for grazing cattle. 1. Model of the relationship between liveweight gain and stocking rate. – Irish Journal of Agricultural economics and Rural Sociology, 5. – 1974. – pp. 1-13.

42. Conway A.G. A production Function for grazing cattle. 4. An economic decision model for grazing steers. – Irish Journal Agricultural Economics and Rural Sociology, 5. – 1974. – pp. 57-108.

43. Gartner J.A. and Herbert W.A. A preliminary model to investigate culling and replacement policy in dairy herds. – Agricultural Systems, 4. – 1979. – pp. 189-215.

44. Walsingham J.M., Edelsten P.R. and Brockington N.R. Simulation of the management of a rabbit population for meat production. – Agricultural Systems, 2. – 1977. – pp. 85-98.

45. White W.C. The determination of an optimal policy for a continually operating egg production enterprise. – Journal of Farm Economics, 41. – 1959. – pp. 1535-1542.

46. McArthur A.T.G. Application of dynamic programming to the culling decision in dairy cattle. – Proceedings of the New Zealand Society of Animal Production, 33. – 1973. – pp. 141-147.

47. Smith B.J. Dynamic programming of the dairy cow replacement problem. – American Journal of Agricultural Economics, 55. – 1973. – 100-104.

48. Stewart H.M., Burnside E.B., Wilton J.W. and Pfeiffer W.C. A dynamic programming approach to cutting decisions in commercial dairy herds. – Journal of Dairy Science, 60. – 1977. – pp. 601-617.

49. Throsdy C.D. Some dynamic programming model for farm management research. – Journal of Agricultural Economics, 16. – 1964. – pp. 98-110.

50. Dent J.B. and Casey H. Linear Programming and Animal Nutrition. – London: Crosby Lockwood. – 1967.

51. Beneke R.R. and Winterboer R. Linear Programming Applications to Agriculture. – Ames: Iowa State University Press. – 1973.

52. Ministry of Agriculture, Fisheries and Food. Energy Allowances and Feeding Systems for Ruminants. – Technical Bulletin No 33. – London: HMSO. – 1975.

53. Caltagirone I. E. Landmark examples in classical biological control. – Annual Review of Entomology, 26. – 1981. – pp. 213-232.

54. Harkins J., Edwards R.A. and Mc Donald P. A new net energy system for ruminants. – Animal Production, 19. – 1974. – pp. 141-148.

55. Miller C.E. The simplex methods for local separable programming . In: Recent Advances in Mathematical Programming (ed. R.L. Graves and P. Wolfe) – New York: Mc Graw-Hill. – 1963. – pp. 311-317.

Приложение А

Таблица А1 – Значения критерия t на 5 и 1%-ном уровне значимости

Число степеней свободы	Уровни значимости	
	0,05	0,01
1	12,71	63,66
2	4,30	9,93
3	3,18	5,84
4	2,78	4,60
5	2,57	4,03
6	2,45	3,71
7	2,37	3,50
8	2,31	3,36
9	2,26	3,25
10	2,23	3,17
11	2,20	3,11
12	2,18	3,06
13	2,16	3,01
14	2,15	2,98
15	2,13	2,95
16	2,12	2,92
17	2,11	2,90
18	2,10	2,88
19	2,09	2,86
20	2,09	2,85
21	2,08	2,83
22	2,07	2,82
23	2,07	2,81
24	2,06	2,80
25	2,06	2,79
26	2,06	2,78
27	2,05	2,77
28	2,05	2,76
29	2,05	2,76
30	2,04	2,75
50	2,01	2,68
100	1,98	2,63
∞	1,96	2,58

Приложение В

Таблица В1 – Значения критерия F на 5% уровне значимости v_1 – степень свободы для большей дисперсии, которая берётся числителем; v_2 – степень свободы знаменателя

v_2	v_1																				
	1	2	3	4	5	6	7	8	9	10	11	12	14	16	20	24	30	40	50	100	∞
1	161	200	216	225	230	234	237	239	241	242	243	244	245	246	248	249	250	251	252	253	254
2	18,51	19,00	19,16	19,25	19,30	19,33	19,36	19,37	19,38	19,39	19,40	19,41	19,42	19,43	19,44	19,45	19,46	19,47	19,47	19,49	19,50
3	10,13	9,55	9,28	9,12	9,01	8,94	8,88	8,84	881	8,78	8,76	8,74	8,71	8,69	8,66	8,64	8,62	8,6	8,58	8,56	8,53
4	7,71	6,94	6,59	6,39	6,26	6,16	6,09	6,04	6,00	5,96	5,93	5,91	5,87	5,84	5,80	5,77	5,74	5,71	5,70	5,66	5,63
5	6,61	5,79	5,41	5,19	5,05	4,95	4,88	4,82	4,78	4,74	4,70	4,68	4,64	4,60	4,56	4,53	4,50	4,46	4,44	4,40	4,36
6	5,99	5,14	4,76	4,53	4,39	4,28	4,21	4,15	4,10	4,06	4,03	4,00	3,96	3,92	3,87	3,84	3,81	3,77	3,75	3,71	3,67
7	5,59	4,74	4,35	4,12	3,97	3,87	3,79	3,73	3,68	3,63	3,60	3,57	3,52	3,49	3,44	3,41	3,38	3,34	3,32	3,28	3,23
8	5,32	4,46	4,07	3,84	3,69	3,58	3,50	3,44	3,39	3,34	3,31	3,28	3,23	3,20	3,15	3,12	3,08	3,05	3,03	2,98	2,93
9	5,12	4,26	3,86	3,63	3,48	3,37	3,29	3,23	3,18	3,13	3,10	3,07	3,02	2,98	2,93	2,90	2,86	2,82	2,80	2,76	2,71
10	4,96	4,10	3,71	3,48	3,33	3,22	3,14	3,07	3,02	2,97	2,94	2,91	2,86	2,82	2,77	2,74	2,70	2,67	2,64	2,59	2,54
11	4,84	3,98	3,59	3,36	3,20	3,09	3,01	2,95	2,90	2,86	2,82	2,79	2,74	2,70	2,65	2,61	2,57	2,53	2,50	2,45	2,40
12	4,75	3,88	3,49	3,26	3,11	3,00	2,92	2,85	2,80	2,76	2,72	2,69	2,64	2,60	2,54	2,50	2,46	2,42	2,40	2,35	2,30
13	4,64	3,80	3,41	3,18	3,02	2,92	2,84	2,77	2,72	2,67	2,63	2,60	2,55	2,51	2,46	2,42	2,38	2,34	2,32	2,26	2,21
14	4,60	3,74	3,34	3,11	2,96	2,85	2,77	2,70	2,65	2,60	2,56	2,53	2,48	2,44	2,39	2,35	2,31	2,27	2,24	2,19	2,13
15	4,54	3,60	3,29	3,06	2,90	2,79	2,70	2,64	2,59	2,55	2,51	2,48	2,43	2,39	2,33	2,29	2,25	2,21	2,18	2,12	2,07
16	4,49	3,63	3,24	3,01	2,85	2,74	2,66	2,59	2,54	2,49	2,45	2,42	2,37	2,33	2,28	2,24	2,20	2,16	2,13	2,07	2,01
17	4,45	3,59	3,20	2,96	2,81	2,70	2,62	2,55	2,50	2,45	2,41	2,38	2,33	2,29	2,23	2,19	2,15	2,11	2,08	2,02	1,96
18	4,41	3,55	3,16	2,93	2,77	2,66	2,58	2,51	2,46	2,41	2,37	2,34	2,29	2,25	2,19	2,15	2,11	2,07	2,04	1,98	1,92
19	4,38	3,52	3,13	2,90	2,74	2,63	2,55	2,48	2,43	2,38	2,34	2,31	2,26	2,21	2,15	2,11	2,07	2,02	2,00	1,94	1,88
20	4,35	3,49	3,10	2,87	2,71	2,60	2,52	2,45	2,40	2,35	2,31	2,28	2,23	2,18	2,12	2,08	2,04	1,99	1,96	1,90	1,84
21	4,32	3,47	3,07	2,84	2,68	2,57	2,49	2,42	2,37	2,32	2,28	2,25	2,20	2,15	2,09	2,05	2,00	1,96	1,93	1,87	1,81
22	4,30	3,44	3,05	2,82	2,66	2,55	2,47	2,40	2,35	2,30	2,26	2,23	2,18	2,13	2,07	2,03	1,98	1,93	1,91	1,84	1,78
23	4,28	3,42	3,03	2,80	2,64	2,53	2,45	2,38	2,32	2,28	2,24	2,20	2,14	2,10	2,04	2,00	1,96	1,91	1,88	1,82	1,76
24	4,26	3,40	3,01	2,78	2,62	2,51	2,43	2,36	2,30	2,26	2,22	2,18	2,13	2,09	2,02	1,98	1,94	1,89	1,86	1,80	1,73
25	4,24	3,38	2,99	2,76	2,60	2,49	2,41	2,34	2,25	2,24	2,20	2,16	2,11	2,06	2,00	1,96	1,92	1,87	1,84	1,77	1,71
26	4,22	3,37	2,98	2,74	2,59	2,47	2,39	2,32	2,27	2,22	2,18	2,15	2,10	2,05	1,99	1,95	1,90	1,85	1,82	1,76	1,69
28	4,20	3,34	2,92	2,71	2,56	2,44	2,36	2,29	2,24	2,19	2,15	2,12	2,07	2,02	1,96	1,91	1,87	1,81	1,78	1,72	1,65
30	4,17	3,32	2,92	2,69	2,53	2,42	2,34	2,27	2,21	2,16	2,12	2,09	2,04	1,99	1,93	1,89	1,84	1,79	1,76	1,69	1,62
40	4,08	3,23	2,84	2,61	2,45	2,34	2,25	2,18	2,12	2,07	2,04	2,00	1,95	1,90	1,84	1,79	1,74	1,69	1,66	1,59	1,51
50	4,03	3,18	2,79	2,56	2,40	2,29	2,20	2,13	2,07	2,02	1,98	1,95	1,90	1,85	1,78	1,74	1,69	1,63	1,60	1,52	1,44
60	4,00	3,15	2,76	2,52	2,37	2,25	2,17	2,10	2,04	1,99	1,95	1,93	1,86	1,81	1,75	1,70	1,65	1,59	1,56	1,48	1,39
100	3,94	3,09	2,70	2,46	2,30	2,19	2,10	2,03	1,97	1,92	1,88	1,85	1,79	1,75	1,68	1,63	1,57	1,51	1,48	1,39	1,28
200	3,89	3,04	2,65	2,41	2,26	2,14	2,05	1,98	1,92	1,87	1,83	1,80	1,74	1,69	1,62	1,57	1,52	1,45	1,42	1,32	1,19
∞	3,84	2,99	2,60	2,37	2,21	2,09	2,01	1,94	1,88	1,83	1,79	1,75	1,69	1,64	1,57	1,52	1,46	1,40	1,35	1,24	1,00