Департамент образования Вологодской области
ФГБОУ ВПО «Вологодская государственная молочнохозяйственная академия имени Н.В. Верещагина»
БОУ ДОД ВО «Областная станция юных натуралистов»

МАТЕМАТИЧЕСКОЕ МОДЕЛИРОВАНИЕ ЛЕСНЫХ ЭКОСИСТЕМ

Учебное пособие

К144303 2

Вологда–Молочное
2012

ВВЕДЕНИЕ

Лес как объект хозяйственной деятельности человека одновременно является и объектом его научных исследований. Поскольку лесные насаждения – это сложные природные образования, то и научное их изучение требует от исследователя знания разнообразных методов анализа. Лес – составная часть географического ландшафта и состоит из совокупности разных видов деревьев, кустарников, трав, мхов, животных и микроорганизмов, разнообразно взаимосвязанных и влияющих друг на друга и на окружающую среду. Однако, лес следует рассматривать не только в пространстве, но и во времени, учитывая его собственное развитие. Эти динамические процессы являются результатом сложных взаимоотношений между составляющими его организмами (борьба за существование и естественный отбор, постоянное обновление и развитие, процессы обмена веществ и энергии).

Рис. 1. Искусственно созданные насаждения сосны обыкновенной (фото Ю.М. Авдеева)

Лес создает свою внутреннюю структуру и особую биологическую обстановку для животных и микроорганизмов. Здесь не только растения приспосабливаются друг к другу, но и животные к растениям. На все это оказывает влияние и условия внешней среды. Совокупность растений, животных и микроорганизмов, характеризующихся определенными отношениями между собой, приспособленностью к окружающей среде, называется ***биоценозом.*** Биоценоз, формируясь под влиянием взаимодействующих живых существ и условий среды, сам оказывает преобразующее действие на условия среды.

Вместе с условиями среды он образует единство, представляющее качественно новую составную часть природы. Такое единство между живыми организмами и средой их существования и есть ***биогеоценоз.*** По определению В.Н. Сукачева, «биогеоценоз – это совокупность на известном протяжении земной поверхности однородных природных условий (атмосферы, горной породы, растительности, животного мира и мира микроорганизмов, почвы и гидрологических условий), имеющая свою особую специфику взаимодействий этих слагающих ее компонентов и определенный тип обмена веществом и энергией их между собой и с другими явлениями природы и представляющая собой внутреннее противоречивое единство, находящееся в постоянном движении, развитии».

Рис. 2. Женские генеративные органы – макростробилы сосны обыкновенной (фото Ю.М. Авдеева)

В лесу, с учетом не только растительных, но и других компонентов, складывается ***лесной биогеоценоз.*** Под лесным биогеоценозом В.Н. Сукачев понимает «...всякий участок леса, однородный на известном протяжении по составу, структуре и свойствам слагаемых его компонентов и по взаимоотношениям между ними».

Близкой к понятию биогеоценоза является *экосистема* – термин введеный для обозначения динамической открытой системы, где организмы, почва и климат являются составными частями. Вообще, термин «экосистема» часто применяется в широком и очень узком смысле. Экосистемой называют и всю экосистему земного шара, и пруд со всеми населяющими его организмами или участок леса и, даже, небольшую группу деревьев.

Рассматривая лес как биогеоценоз, необходимо помнить, что *древостой является* основным связующим звеном в биологической системе леса, который оказывает сильное влияние на среду, в том числе и среду обитания большей части компонентов самого лесного биогеоценоза, их состав, размещение и жизнедеятельность.

Рис. 3. Шмель или земляная пчела
(фото Ю.М. Авдеева)

Изучение лесных биогеоценозов должно проводиться в комплексе, во всем многообразии связей между его частями и процессами, протекающими внутри него. Вместе с тем, необходим более углубленный анализ отдельно взятых компонентов системы и происходящих внутри нее процессов. Это делает необходимым использование системного подхода, т.е. принципа системности в познании биологических явлений природы, как в частностях, так и в целом. Отсюда следует вывод, что основной задачей при изучении биогеоценоза – выявление всех разнообразных связей между отдельными компонентами и окружающей средой.

Следует отметить, что лесные биогеоценозы характеризуются высокой изменчивостью в пространстве, поскольку условия микросреды в их границах всегда несколько отличаются друг от друга. Большая вариабельность (изменчивость) признаков обусловлена и хозяйственной деятельностью человека. Изменение климатических и погодных условий, периодические их колебания могут еще больше увеличить вариабельность лесных биогеоценозов. Все это существенно затрудняет процесс исследований.

Большую помощь в этом оказывают статистические методы исследования, которые позволяют получить средние данные с определенной вероятностью и степенью достоверности.

Моделирование лесных фитоценозов позволяет количественные изменения массовых явлений представить в виде конкретных математических моделей, тем самым повысить эффективность массовых наблюдений биологических явлений природы при научных исследованиях. Системный подход для моделирования сложных процессов в природе – одна из ведущих идей современного естествознания. Моделирование однородных статистических совокупностей раскрывает перед исследователями динамику причинно-следственных взаимосвязей между составными элементами и показывает, что за случайными явлениями стоят закономерности, которые доступны описанию точными математическими моделями. При этом особое место занимают методы многомерной статистики, которые позволяют количественные изменения массовых явлений представить математическими моделями, дополнив их корреляционным, регрессионным, дисперсным и другими видами анализа.

При научных исследованиях биологических явлений природы наиболее эффективным является ***метод массовых наблюдений***. Для этого сначала производят большое количество наблюдений в натуре, характеризующих те или иные явления. Собранный материал обрабатывают, анализируют, делают соответствующие выводы и устанавливают те или иные закономерности. Рассмотренный путь установления закономерностей называется ***прямым индук-***

тивным методом, когда от отдельных фактов переходят к общим положениям.

Статистический анализ массовых наблюдений дополняет и углубляет познания биологических явлений природы, позволяет объективно оценить полученные результаты и сделать обоснованные выводы с определенной степенью достоверности. Обработка материалов наблюдения без применения методов математической статистики является неполной и нередко приводит к неправильным выводам в отношении изучаемого явления.

1 ГРУППИРОВКА И ОБРАБОТКА ДАННЫХ КОЛИЧЕСТВЕННОЙ ИЗМЕНЧИВОСТИ

1.1 Вычисление статистических характеристик при малой выборке

Малой выборкой принято считать вариационный ряд с небольшим количеством единиц наблюдения (менее 30).

Среднее значение – это обобщающая характеристика изучаемого признака в исследуемой совокупности, отражающая его типичный уровень в расчете на единицу совокупности в конкретных условиях пространства и времени.

Среднее значение при малой выборке вычисляется по формуле:

$$\mathrm{M} = \frac{x_1 + x_2 + x_3 ... + x_k}{n} = \frac{\sum x_i}{n}, \tag{1.1}$$

где $x_1, x_2, x_3, ..., x_k$ – варианты;

n – количество единиц наблюдения.

Среднее квадратичное отклонение (стандартное отклонение) – это степень рассеяния ряда распределения, показывая отклонение 68% единиц наблюдения от среднего значения, т.е. среднее отклонение отдельных объектов выборки от среднего значения в 68 случаях из 100.

При малой выборке этот показатель определяется по формуле:

$$\sigma = \sqrt{\frac{\sum (x_i - M)^2}{n-1}}. \tag{1.2}$$

Среднеквадратическое отклонение измеряется в тех же единицах, что и среднее значение.

Основная ошибка среднего значения выражает величину, на которую отличается среднее значение выборки от среднего значения генеральной совокупности и определяется по формуле:

$$m_M = \frac{\sigma}{\sqrt{n}}. \tag{1.3}$$

Основная ошибка, как правило, записывается вместе со средним значением через знак ±, и измеряется в тех же единицах.

Коэффициент изменчивости (вариации) – основное отклонение, выраженное в процентах от среднего:

$$C = \frac{\sigma}{M} 100\%. \tag{1.4}$$

Изменчивость ряда может быть малой ($C \leq 10\%$), средней ($C = 10{,}1 \ldots 30\%$) и большой ($C \geq 30{,}1\%$).

Точность опыта характеризует процент расхождения между выборочной и генеральной средними, являясь ошибкой наблюдения.

$$P = \frac{C}{\sqrt{n}}, \tag{1.5}$$

или

$$P = \frac{m_M}{M} 100\%. \tag{1.6}$$

Точность опыта считается высокой, если она менее 5%, или удовлетворительной (6…10%). В остальных случаях результаты считаются не точными.

Достоверность среднего значения является показателем его надежности и числено, равен отношению среднего значения к его основной ошибке:

$$t_1 = \frac{M}{m_M}. \tag{1.7}$$

Среднее значение достоверно если t_1 больше четырех. Иногда результаты исследований считают достоверными если t_1 больше или равен трем. В случае если показатель достоверности менее трех, то среднее значение нельзя использовать при формулировании выводов.

Пример:

В насаждении выборочно определили диаметр у 16 учетных деревьев:

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

Необходимо установить основные статистические показатели: M, σ, m_M, C, P, t_1.

Для начала следует заполнить вспомогательную таблицу 1.

Таблица 1 – Вспомогательная таблица для расчета основных статистических показателей непосредственным способом

Номер п.п.	Варианты (диаметры, см) x_i	Центральные отклонения $x_i - M$	Квадраты центрального отклонения $(x_i - M)^2$
1	16,6	-12,3	151,29
2	19,5	-9,4	88,36
3	21,7	-7,2	51,84
4	22,6	-6,3	39,69
5	26,4	-2,5	6,25
6	26,2	-2,7	7,29
7	28	-0,9	0,81
8	30,1	+1,2	1,44
9	30,3	+1,4	1,96
10	31,5	+2,6	6,76
11	33,7	+4,8	23,04
12	37,3	+8,4	70,56
13	30,4	+1,5	2,25
14	33,2	+4,3	18,49
15	33,9	+5	25
16	41,1	+12,2	148,84
n=16	Σ 462,5	+41,4 - 41,3 Σ ±82,7	Σ 643,87

Сначала следует рассчитать средний диаметр:

$$\text{M} = \frac{\sum x_j}{n} = \frac{462{,}5}{16} = 28{,}9 \text{ см.}$$

Затем нужно вычислить среднеквадратическое отклонение:

$$\sigma = \sqrt{\frac{\Sigma (x_i - M)^2}{n-1}} = \frac{643{,}87}{15} = 6{,}55 \text{ см.}$$

Тогда основная ошибка среднего значения составит:

$$m_M = \frac{\sigma}{\sqrt{n}} = \frac{6,55}{4} = 1,64 \approx 1,6\text{ см}.$$

Коэффициент изменчивости:

$$C = \frac{\sigma}{M}100\% = \frac{6,55}{28,9}100 = 22,66\%.$$

Поскольку значение коэффициента вариации находится в пределах 10…30%, то изменчивость признака средняя.

Рассчитаем точность опыта:

$$P = \frac{C}{\sqrt{n}} = \frac{22,66}{\sqrt{16}} = 5,7\%,$$

или по другой формуле:

$$P = \frac{m_M}{M}100\% = \frac{1,64}{28,9} = 5,7\%.$$

Точность опыта удовлетворительная.

Для определения надежности суждения следует найти показатель достоверности t_1:

$$t_1 = \frac{M}{m_M} = \frac{28,9}{1,64} = 17,6.$$

Средний диаметр определен достоверно, т.к. t_1 равен 17,6, что больше четырех ($t_1>4$).

Вывод: *Средний диаметр насаждения составляет 28,9±1,6 см. Изменчивость признака средняя. Точность опыта удовлетворительная, а полученные результаты достоверны.*

1.2 Группировка и обработка данных количественной изменчивости при большой выборке

Для наглядности научного исследования и при большой выборке (25-30 и более единиц наблюдения) исходные опытные данные группируются в вариационный ряд.

Вариационный ряд показывает число повторений значений признака, по которому изучается статистическая совокупность. Вариационный ряд имеет варианты, интервалы, численности и частности.

Варианта – это признак, по которому изучается статистическая совокупность (диаметр, высота, протяженность безсучковой зоны и т.д.).

Интервал – величина (шаг) ступени (класса) или какой-либо градации.

Численность – это число единиц наблюдения.

Частность – это процентное отношение численности интервала от общего количества единиц наблюдения.

При составлении вариационного ряда важно правильно установить число интервалов, зависимое от количества единиц наблюдения. Здесь можно воспользоваться формулой:

$$k = 1 + 3{,}322 \lg n, \tag{1.8}$$

где k – число интервалов;

n – численность выборки.

Для удобства можно воспользоваться следующими продержками:

Численность:	25 – 40	41 – 60	61 – 100	101 – 200	≥201
Количество интервалов:	5 – 6	6 – 8	7 – 10	8 – 12	9 – 15

После определения числа интервалов рассчитывают их величину, т.е. шаг ступени. Для этого разность между максимальным и минимальным значением вариационного ряда делят на число интервалов:

$$i = \frac{x_{max} - x_{min}}{k}, \tag{1.9}$$

где x_{max} – максимальная варианта,

x_{min} – минимальная варианта.

Величину интервала следует округлять до целых единиц. Округление шага ступени приводит к увеличению или уменьшению количества интервалов при разноске по ним вариант. Для правильного построения вариационного ряда необходимо чтобы фактическое количество ступеней не отличалось более чем на две от расчетного.

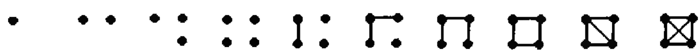
Для установления пределов класса, необходимо начиная от минимальной варианты последовательно прибавлять шаг ступени. При этом значение нижнего предела первой ступени можно принимать меньше минимальной варианты, так, чтобы среднее значение класса было целым числом.

Среднее значение интервала иначе называют ступенью вариационного ряда – это полусумма предельных значений интервала.

Значение вариант разносятся по ступеням, т.е. группируются в классы. Рабочая запись осуществляется, как правило, способом конвертов (так «называемая точковка»). Сначала ставятся точки по углам будущего квадрата (цифры 1-4), затем точки поочередно соединяют линиями по сторонам квадрата (цифры 5-8). Цифру 9 и 10 записывают, соединяя противоположные углы:

Цифра *1* *2* *3* *4* *5* *6* *7* *8* *9* *10*

Рабочая запись



При количестве вариант более десяти в пределах класса эту операцию повторяют снова.

После группировки вариант по классам рассчитывают основные статистические показатели.

1.2.1 Вычисление статистических показателей вариационного ряда непосредственным способом

Среднее значение здесь определяется по формуле:

$$\mathrm{M} = \frac{x_1 n_1 + x_2 n_2 + x_3 n_3 ... + x_k n_k}{n_1 + n_2 + n_3 + ... + n_k} = \frac{\sum x_j n_j}{\sum n}, \qquad (1.10)$$

Среднее квадратичное (основное) отклонение вычисляется как формуле:

$$\sigma = \sqrt{\frac{\sum (x_j - M)^2 n_j}{\sum n}}. \qquad (1.11)$$

Основную ошибку, коэффициент изменчивости, точность и достоверность среднего значения определяют также как и при малой выборке.

Пример:

Имеются измерения диаметров 32 деревьев:

14,5 *16,6 18,1 19,5 20,1 21,7 22,2 22,6 23,8 26,4 26,8 26,2 27,1 28*
28,5 30,1 29,7 30,3 30,9 31,5 32,1 33,7 34,4 37,3 29,8 30,4 31,7 33,2
34,7 33,9 39 ***41,1***

Необходимо рассчитать основные статистические показатели методом сгруппированных данных.

$k = 1 + 3{,}322 \lg n = 1 + 3{,}322 \cdot 1{,}505 = 6$

$i = \frac{x_{max} - x_{min}}{k} = \frac{41{,}1 - 14{,}5}{6} = 4{,}4 \approx 4$

14,5 + 4= 18,5 = 16,5

Таблица 2 – Распределение диаметров по интервалам

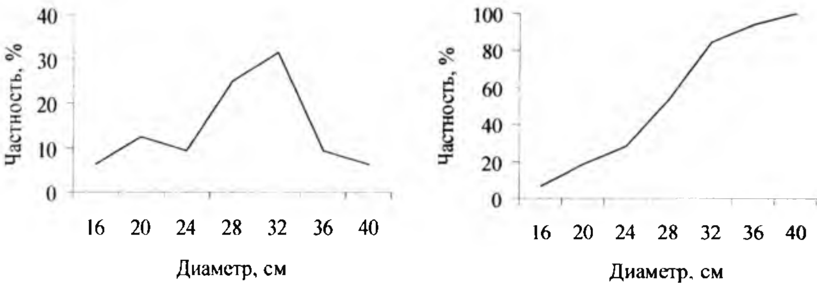
Границы классов, см	Среднее значение класса, см	Численность, шт.	
		в рабочей записи	в цифрах
14-17,9	16		2
18-21,9	20		4
22-25,9	24		3
26-29,9	28		8
30-33,9	32		10
34-37,9	36		3
38-41,9	40		2
Итого	-	32	32

Рассчитывается частность (процент) по ступеням толщины (табл. 3).

Таблица 3 – Вариационный ряд по ступеням толщины

Среднее значение интервала	16	20	24	28	32	36	40	Всего
Численность, шт.	2	4	3	8	10	3	2	32
Частность, %	6,3	12,5	9,4	25,0	31,3	9,4	6,3	100

Графическое отображение вариационного ряда приведено на рисунке 1, а. Суммируя частность по интервалам, получаем: 16 – 6,3%; 20 – 6,3+12,5=18,8%; 24 – 18,8+9,4=28,2% и т.д. (рис. 4, б)



а) *б)*

Рис.4. Графическое выражение вариационного ряда:
а) Распределение частностей; б) Кумулята частностей

Для вычисления основных статистических показателей непосредственным способом строится вспомогательная таблица (табл.4).

Таблица 4 – Вычисление статистических показателей непосредственным способом

Диаметр x_i, см	Численность n_i, шт.	$x_i n_i$	x_i-M	$(x_i\text{-}M)^2$	$(x_i\text{-}M)^2 n_i$
16	2	32	-12,6	158,76	317,52
20	4	80	-8,6	73,96	295,84
24	3	72	-4,6	21,16	63,48
28	8	224	-0,6	0,36	2,88
32	10	320	3,4	11,56	115,6
36	3	108	7,4	54,76	164,28
40	2	80	11,4	129,96	259,92
Всего	32	916	-	-	1219,52

Сначала рассчитается средний диаметр:

$$\mathrm{M} = \frac{x_1 n_1 + x_2 n_2 + x_3 n_3 ... + x_k n_k}{n_1 + n_2 + n_3 + ... + n_k} = \frac{\sum x_j n_j}{\sum n} = \frac{916}{32} = 28{,}6\,\text{см};$$

Далее необходимо определить среднеквадратичное отклонение:

$$\sigma = \sqrt{\frac{\sum (x_j - M)^2 n_j}{\sum n}} = \sqrt{\frac{1219{,}52}{32}} = \sqrt{38{,}11} = 6{,}17\,\text{см};$$

В таком случае основная ошибка среднего значения составит:

$$m_M = \frac{\sigma}{\sqrt{\sum n}} = \frac{6{,}17}{5{,}66} = 1{,}09 \approx 1{,}1\,\text{см}.$$

Коэффициент изменчивости:

$$C = \frac{\sigma}{M} 100\% = \frac{6{,}17}{28{,}6} 100 = 21{,}57\,\%.$$

Поскольку значение коэффициента вариации находится в пределах 10…30%, то изменчивость признака средняя.

Точность опыта:

$$P = \frac{C}{\sqrt{\sum n}} = \frac{21{,}57}{\sqrt{32}} = 3{,}8\,\%,$$

или по другой формуле:

$$P = \frac{m_M}{M} 100\% = \frac{1{,}09}{28{,}6} = 3{,}8\,\%.$$

Точность опыта высокая (т.к. менее 5%).

Для определения надежности суждения следует найти показатель достоверности t_1:

$$t_1 = \frac{M}{m_M} = \frac{28{,}6}{1{,}09} = 26{,}2\,.$$

Средний диаметр определен достоверно, т.к. t_1 равен 26,2, что больше четырех ($t_1 > 4$).

Вывод: *Средний диаметр насаждения составляет 28,6±1,1 см. Изменчивость признака средняя. Точность опыта высокая, а полученные результаты достоверны.*

1.2.2 Вычисление статистических показателей вариационного ряда с использованием начальных моментов способом произведений

Тот же пример можно решить по способу моментов. Этот прием облегчает целый ряд сложных статистических расчетов.

Среднее значение по данному способу вычисляется по формуле:

$$M = x_0 \pm i m_1, \tag{1.12}$$

где x_0 – начальная варианта (условно-принятая для вычисления начальных моментов);

i – величина интервала;

m_1 – первый начальный момент.

Начальным моментом статистической величины называется сумма произведений тех или иных степеней вспомогательных отклонений (от x_0) на соответствующую численность, деленная на сумму всех численностей.

Начальные моменты вычисляются по формуле:

$$m_a = \frac{\sum (x_i - x_0)^a n_i}{\sum n}, \tag{1.13}$$

где m_a – начальный момент некоторой степени a (обычно до четвертой степени);

x_i – варианты ряда распределения;

x_0 – начальное значение;

n_i – численности интервала;

Σn – сумма численностей.

Среднее квадратичное отклонение рассчитывается по формуле:

$$\sigma = i\sqrt{m_2 - m_1^2}\ . \tag{1.14}$$

Для расчетов следует оформить вспомогательную таблицу.

Таблица 5 – Вычисление начальных моментов по способу произведений

Диаметр x_i, см	Численность n_i, шт.	k_i	n_ik_i	$n_ik_i^2$	$n_ik_i^3$	$n_ik_i^4$
16	2	-3	-6	18	-54	162
20	4	-2	-8	16	-32	64
24	3	-1	-3	3	-3	3
28	8	0	0	0	0	0
32	10	1	10	10	10	10
36	3	2	6	12	24	48
40	2	3	6	18	54	162
Всего	32	-	5	77	-1	449

За начальное значение лучше выбрать варианту на середине ряда распределения. Правильность выбора начального значения определяется величиной первого начального момента (чем меньше его значение, тем вернее выбрано начальное значение).

Заполнив графы таблицы, рассчитывая значения по формулам, приведенным для каждого столбца, следует рассчитать начальные моменты:

$$m_1 = \frac{\sum n_i k_i}{\sum n} = \frac{5}{32} = +0{,}156;$$

$$m_2 = \frac{\sum n_i k_i^2}{\sum n} = \frac{77}{32} = 2{,}406;$$

$$m_3 = \frac{\sum n_i k_i^3}{\sum n} = \frac{677}{32} = -0{,}031;$$

$$m_4 = \frac{\sum n_i k_i^4}{\sum n} = \frac{18065}{32} = 14{,}031;$$

Первый и третий начальные моменты следует обязательно приводить со знаком «плюс» или «минус».

Установив первый начальный момент можно рассчитать среднее значение:

$$M = x_0 \pm i m_1 = 28 + 4 \cdot 0{,}156 = 28{,}624 \approx 28{,}6 \text{ см}.$$

Основное отклонение:

$$\sigma = i\sqrt{m_2 - m_1^2} = 4\sqrt{2{,}406 - 0{,}156^2} = 6{,}17 \text{ см}.$$

Таким образом, рассчитанные среднее значение и основное отклонение аналогичны результатам, полученным ранее непосредственным способом. Оставшиеся статистические показатели находятся по формулам, описанным в разделе 1.2.1.

1.2.3 Центральные моменты

Для характеристики вариационного ряда используют такие показатели как мера косости и мера крутости. Для их определения необходимо произвести расчет центральных моментов.

Центральным моментом статистической величины называется сумма произведений тех или иных степеней центральных отклонений (от М) на соответствующую численность, деленная на сумму всех численностей.

Для расчета центральных моментов используют начальные моменты, оперируя следующими соотношениями:

$$\mu_1 = 0; \tag{1.15}$$

$$\mu_2 = m_2 - m_1^2; \tag{1.16}$$

$$\mu_3 = m_3 - 3m_2m_1 + 2m_1^3; \tag{1.17}$$

$$\mu_4 = m_4 - 4m_3m_2 + 6m_2m_1^2 - 3m_1^4, \tag{1.18}$$

где μ_1, μ_2, μ_3, μ_4 – соответственно первый, второй, третий и четвертый центральные моменты,

m_1, m_2, m_3, m_4 – соответственно первый, второй, третий и четвертый начальные моменты.

Таким образом, подставив значения начальных моментов рассчитанных ранее (m_1 = + 0,156; m_2 =2,406; m_3 = – 0,031; m_4 = 14,031) в формулы, получим центральные моменты:

$$\mu_2 = m_2 - m_1^2 = 2{,}406 - 0{,}156^2 = 2{,}382;$$

$$\mu_3 = m_3 - 3m_2m_1 + 2m_1^3 = -0{,}031 - 3 \cdot 2{,}406 \cdot 0{,}156 + 2 \cdot 0{,}156^3 = -1{,}149;$$

$$\mu_4 = m_4 - 4m_3m_2 + 6m_2m_1^2 - 3m_1^4 =$$

$$= 14{,}031 - 4 \cdot (-0{,}031) \cdot 2{,}406 + 6 \cdot 2{,}406 \cdot 0{,}156^2 - 3 \cdot 0{,}156^4 = 14{,}678$$

2 КОРРЕЛЯЦИЯ

Коэффициент корреляции и корреляционное отношение определяют для оценки тесноты связи между двумя или несколькими статистическими величинами. Коэффициент корреляции используют, если связь между этими величинами прямолинейна. При криволинейной зависимости определяют корреляционное отношение.

Коэффициент корреляции вычисляют по формуле:

при малой выборке:

$$r = \frac{\sum (x_i - M_x)(y_i - M_y)}{\sqrt{\sum (x_i - M_x)^2 (y_i - M_y)^2}}, \qquad (2.1)$$

при большой выборке:

$$r = \frac{\sum x_i y_i - nM_x M_y}{\sqrt{\left(\sum x_i^2 - nM_x^2\right)\left(\sum y_i^2 - nM_y^2\right)}}. \qquad (2.2)$$

Коэффициент корреляции колеблется в пределах от 0 до +1 при прямой зависимости и от 0 до -1 при обратной связи. Корреляционное отношение всегда положительно от 0 до +1.

О тесноте связи судят по следующим придержкам:

Коэффициент корреляции, или корреляционное отношение	Теснота связи
Менее 0,30	Слабая
0,31 – 0,50	Умеренная
0,51 – 0,70	Значительная
0,71 – 0,90	Высокая
0,91 и более	Очень высокая

Если теснота связи равняется единице то, ее называют функциональной.

2.1 Вычисление коэффициента корреляции непосредственным способом

Пример:

В насаждении выборочно определили высоту и диаметр у 16 учетных деревьев:

H	*17,5*	*18,5*	*19,5*	*20*	*20,5*	*21*	*21,5*	*22,5*	*23*	*24*	*24*	*24,5*	*25,5*	*26,5*	*27,5*	*28*
D	*16,6*	*19,5*	*21,7*	*22,6*	*26,4*	*26,2*	*28*	*30,1*	*30,3*	*31,5*	*33,7*	*37,3*	*30,4*	*33,2*	*33,9*	*41,1*

Необходимо установить связь между этими показателями.

Для решения задачи необходимо построить вспомогательную таблицу.

Таблица 6 – Вычисление коэффициента корреляции (при малой выборке)

Признаки		Центральные отклонения				
Диаметр, x_i, см	Высота, y_i, м	$x_i - M_x$	$y_i - M_y$	$(x_i - M_x)$ $*(y_i - M_y)$	$(x_i - M_x)^2$	$(y_i - M_y)^2$
16,6	17,5	-12,3	-5,3	65,2	151,3	28,1
19,5	18,5	-9,4	-4,3	40,4	88,4	18,5
21,7	19,5	-7,2	-3,3	23,8	51,8	10,9
22,6	20	-6,3	-2,8	17,6	39,7	7,8
26,4	20,5	-2,5	-2,3	5,8	6,3	5,3
26,2	21	-2,7	-1,8	4,9	7,3	3,2
28	21,5	-0,9	-1,3	1,2	0,8	1,7
30,1	22,5	+1,2	-0,3	-0,4	1,4	0,1
30,3	23	+1,4	+0,2	0,3	2,0	0,0
31,5	24	+2,6	+1,2	3,1	6,8	1,4
33,7	24	+4,8	+1,2	5,8	23,0	1,4
37,3	24,5	+8,4	+1,7	14,3	70,6	2,9
30,4	25,5	+1,5	+2,7	4,1	2,3	7,3
33,2	26,5	+4,3	+3,7	15,9	18,5	13,7
33,9	27,5	+5	+4,7	23,5	25,0	22,1
41,1	28	+12,2	+5,2	63,4	148,8	27,0
Σ 462,5	364,0	+41,4 -41,3 +0,2	-21 +21	Σ 288,8	Σ 643,9	Σ 151,5

$$M_x = \frac{\sum x_i}{n} = \frac{362{,}5}{16} = 28{,}9 \text{ см};$$

$$M_y = \frac{\sum y_i}{n} = \frac{364{,}0}{16} = 22{,}8 \text{ м},$$

где n – количество единиц наблюдения в нашем случае равное 16.

Вычисляем коэффициент корреляции:

$$r = \frac{\sum (x_i - M_x)(y_i - M_y)}{\sqrt{\sum (x_i - M_x)^2 (y_i - M_y)^2}} = \frac{288{,}8}{\sqrt{643{,}8 \cdot 151{,}5}} = 0{,}92.$$

Ошибка коэффициента корреляции рассчитывается по формуле:

$$m_r = \frac{1 - r^2}{\sqrt{n}}, \qquad (2.3)$$

тогда:

$$m_r = \frac{1 - (0{,}92)^2}{\sqrt{16}} = \frac{1 - 0{,}84}{4} = 0{,}04\,.$$

Расчет показателя достоверности коэффициента корреляции осуществляется по формуле

$$t_{\text{r}} = \frac{r}{m_r}, \qquad (2.4)$$

получаем:

$$t_{\text{r}} = \frac{r}{m_r} = \frac{0{,}92}{0{,}04} = 23{,}0,$$

коэффициент корреляции достоверный, так как больше 3.

Оценка значимости r проводится по t-распределению Стьюдента $t_{\text{ф}}$ большего t_{st}. Определяется расчетный критерий:

$$t_{\text{ф}} = \frac{\text{r}}{\sqrt{1 - \text{r}^2}} \cdot \sqrt{\text{n} - 2} = \frac{0{,}94}{\sqrt{1 - (0{,}92)^2}} \cdot \sqrt{16 - 2} = 8{,}9$$

Табличный критерий Стьюдента (t_{st}) находится в приложении 2 (при числе степеней свободы: df = n – 2 = 16 – 2 = 14) и равен 2,1; 3,0; 4,1 соответственно при 5; 1; 0,1% уровне значимости , таким образом сравнивая t_{st} равный 8,9 с табличным следует заключить, что $t_ф$ больше t_{st}. Это значит, что наш вывод о тесноте связи будет существенен на 0,1% уровне значимости.

Вывод: *Между диаметром и высотой существует прямая очень высокая (r = 0,92±0,04) связь, близкая к функциональной.*

2.2 Вычисление меры связи для сгруппированных данных

2.2.1 Вычисление коэффициента корреляции

Пример:

У каждого из 32 деревьев измерены диаметр и высота:

h	*17,5*	*18,5*	*18,5*	*19*	*20*	*20*	*20,5*	*21*	*21,5*	*21,5*	*22*	*22,5*	*22*	*23,5*	*23*	*24*
d	*17,4*	*17,6*	*18,3*	*20,2*	*21,3*	*22,5*	*25,5*	*25,1*	*26,1*	*26,4*	*26,9*	*30*	*30,5*	*27,4*	*27,9*	*28*

h	*23*	*23,5*	*23,5*	*23,5*	*24*	*24*	*24,5*	*24,5*	*24,5*	*25*	*26*	*26*	*27*	*28*	***28,5***	***28,5***
d	*28,5*	*29,6*	*30,1*	*30,9*	*31,5*	*30,4*	*33,2*	*33,7*	*34,3*	*33,5*	*35,8*	*36*	*36,7*	*37,9*	*38,6*	***39,3***

Определить коэффициент корреляции

Для расчета коэффициента корреляции предварительно следует подготовить рабочую таблицу (табл. 6).

Таблица 6 – Расчет вспомогательных величин для вычисления коэффициента корреляции (r) и корреляционного отношения (η)

Диаметр x_i, см	Высота y_i, м: 18	20	22	24*	26	28	n_x	k_x	$n_x k_x$	$n_x k^2_x$	$n_{xy}k_y$: 18	20	22	24	26	28	$\Sigma n_{xy}k_y$	$\Sigma n_{xy}k_y k_x$	$(\Sigma n_{xy}k_y)^2$	$\frac{(\sum n_{xy}k_y)^2}{n_x}$
16	2						2	-12	-24	288	-12						-12	144	144	72,0
20	1	1					2	-8	-16	128	-6	-4					-10	80	100	50,0
24		3	1				4	-4	-16	64		-12	-2				-14	56	196	49,0
28*			3	5			8	0	0	0			-6	0			-6	0	36	4,5
32			2	6	1		9	4	36	144			-4	0	2		-2	-8	4	0,4
36				1	2	2	5	8	40	320				0	4	8	12	96	144	28,8
40						2	2	12	24	288						8	8	96	64	32,0
n_y	3	4	6	12	3	4	32		44	1232							-24	464		236,7
k_y	-6	-4	-2	0	2	4														
$n_y k_y$	-18	-16	-12	0	6	16	-24													
$n_y k^2_y$	108	64	24	0	12	64	272													
$n_{xy}k_x$ 16	-24																			
$n_{xy}k_x$ 20	-8	-8																		
$n_{xy}k_x$ 24		-12	-4																	
$n_{xy}k_x$ 28			0	0																
$n_{xy}k_x$ 32			8	24	4															
$n_{xy}k_x$ 36				8	16	16														
$n_{xy}k_x$ 40						24														
$\Sigma n_{xy}k_x$	-32	-20	4	32	20	40	44													
$\Sigma n_{xy}k_y k_x$	192	80	-8	0	40	160	464													
$(\Sigma n_{xy}k_x)^2$	1024	400	16	1024	400	1600														
$\frac{(\sum n_{xy}k_x)^2}{n_y}$	341	100	2,67	85,3	133	400	1063													

Примечание: * - x_0 и y_0.

В соответствии с предварительно рассчитанными интервалами по каждому из признаков производят разноску частот (в данном случае количество деревьев соответствующих классу диаметра и высоты) в таблицу распределения. Далее производят расчет вспомогательных величин по формулам, приведенным в таблице.

Коэффициент корреляции находят по формуле:

$$r = \frac{\mu_{1/1}}{\sigma'_x \sigma'_y}, \qquad (2.5)$$

где $\mu_{1/1}$ – первый центральный момент произведения двух статистических величин;

σ'_x σ'_y – неименованные средние квадратичные отклонения статистических величин x, y.

Первый центральный момент определяется по формуле:

$$\mu_{1/1} = m_{1/1} - m_{1x} m_{1y}, \qquad (2.6)$$

где $m_{1/1}$ – первый начальный момент произведения двух статистических величин;

m_{1x}, m_{1y} – первые начальные моменты статистических величин.

Первый начальный момент произведения двух статистических величин находят по формуле:

$$m_{1/1} = \frac{\sum n_{xy} k_y k_x}{\sum n}, \qquad (2.7)$$

где $\sum n_{xy} k_y k_x$ - сумма произведений отклонений по численности;

$\sum n$ - общая численность выборки.

Таким образом, в соответствии с формулами расчета начальных моментов приведенных в п.1.2.2 производим следующие расчеты.

Начальные моменты:

$$m_{1(x)} = \frac{\sum n_x k_x}{\sum n} = \frac{44}{32} = +1{,}375;$$

$$m_{2(x)} = \frac{\sum n_x k_{xi}^2}{\sum n} = \frac{1232}{32} = 38{,}500;$$

$$m_{1(y)} = \frac{\sum n_y k_y}{\sum n} = \frac{-24}{32} = -0{,}750;$$

$$m_{2(y)} = \frac{\sum n_y k_y^2}{\sum n} = \frac{272}{32} = 8{,}500$$

Вторые центральные моменты:

$$\mu_{2(x)} = m_{2(x)} - m_{1(x)}^2 = 38{,}500 - 1{,}375^2 = 36{,}609$$

$$\mu_{2(y)} = m_{2(y)} - m_{1(y)}^2 = 8{,}500 - 0{,}750^2 = 7{,}937$$

Первый начальный момент двух статистических величин:

$$m_{1/1} = \frac{\sum n_{xy} k_y k_x}{\sum n} = \frac{464}{32} = 14{,}500$$

Первый центральный момент произведения двух статистических величин:

$$\mu_{1/1} = m_{1/1} - m_{1x} m_{1y} = 14{,}500 - 1{,}375 \cdot (-0{,}750) = 15{,}531$$

Среднее квадратичное отклонение по x:

$$\sigma'_x = \sqrt{\mu_{2(x)}} = \sqrt{36{,}609} = 6{,}051$$

Среднее квадратичное отклонение по y:

$$\sigma'_y = \sqrt{\mu_{2(y)}} = \sqrt{7{,}937} = 2{,}817$$

Коэффициент корреляции параметров:

$$r = \frac{\mu_{1/1}}{\sigma'_x \sigma'_y} = \frac{15{,}531}{6{,}051 \cdot 2{,}817} = 0{,}91$$

Основная ошибка коэффициента корреляции:

$$m_r = \pm \frac{1 - r^2}{\sqrt{\sum n}} = \frac{1 - 0{,}91^2}{\sqrt{32}} = 0{,}03$$

Достоверность коэффициента корреляции:

$$t_1 = \frac{r}{m_r} = \frac{0{,}91}{0{,}03} = 30{,}33$$

Поскольку $t_l>4$ коэффициент корреляции достоверен.

Вывод: *Между диаметром и высотой наблюдается статистически достоверная очень высокая связь (r=0,91±0,03).*

2.2.2 Вычисление корреляционного отношения

Многие признаки в биометрии имеют криволинейный характер связи. Поэтому расчет коэффициента корреляции в данном случае не отражает в полной мере тесноту зависимости параметров. При нелинейной связи принято рассчитывать *корреляционное отношение* η по формуле:

$$\eta^2_{y/x} = \frac{1}{\mu_{2(y)}}\left[\frac{1}{\sum n}\sum\frac{\left(\sum n_{xy}k_y\right)^2}{n_x} - m^2_{1(y)}\right], \qquad (2.8)$$

где $\mu_{2(y)}$ – второй центральный момент по параметру у (высоте);

$\sum n_{xy}k_y k_x$ – сумма произведений отклонений по численности;

$\sum n$ – общая численность выборки.

$m^2_{1(y)}$ – первый начальный момент по высоте.

Пример:

Необходимо определить корреляционное отношение по исходным данным примера п.2.2.1.

Для расчета корреляционного отношения пользуются рассчитанными в предыдущем пункте параметрами уравнения.

$$\eta^2_{y/x} = \frac{1}{\mu_{2(y)}}\left[\frac{1}{\sum n}\sum\frac{\left(\sum n_{xy}k_y\right)^2}{n_x} - m^2_{1(y)}\right] = \frac{1}{7{,}937}\left[\frac{1}{32}\cdot 236{,}7-(-0{,}750)^2\right] = 0{,}85$$

Полученный результат (η^2) называют показателем силы влияния или индексом детерминации. Этот статистический показатель характеризует долю влияния факториального признака (в данном случае диаметр – x) на результативный (в данном случае высота – y). Таким образом, в нашем случае высота деревьев на 85% зависит от их диаметров.

Само корреляционное отношение получают путем извлечения корня из показателя силы влияния.

$\eta = \sqrt{0,85} = 0,92$.

Основная ошибка корреляционного отношения определяется по стандартной формуле:

$$m_\eta = \pm\frac{1-\eta^2}{\sqrt{\sum n}} = \frac{1-0,92^2}{\sqrt{32}} = 0,03.$$

Достоверность показателя:

$$t_1 = \frac{\eta}{m_\eta} = \frac{0,92}{0,03} = 30,67.$$

Вывод: *Корреляционное отношение диаметров деревьев с их высотами составляет 0,92±0,03, связь достоверная (t>4) очень высокая (η>0,91).*

2.2.3 Мера линейности и показатель криволинейности

Для того чтобы судить о характере взаимосвязи (прямолинейности или криволинейности) необходимо рассчитать следующие показатели:

Меру линейности находят по формуле:

$$\xi = \eta^2 - r^2. \qquad (2.9)$$

В нашем примере (п.2.2.1 и п.2.2.2):

$\xi = \eta^2 - r^2 = 0,92^2 - 0,91^2 = 0,018$.

Основная ошибка меры линейности определяется:

$$m_\xi \approx \pm\sqrt{\frac{\xi}{\sum n}}, \qquad (2.10)$$

где Σn – сумма численностей.

В нашем примере:

$$m_\xi \approx \pm\sqrt{\frac{\xi}{\sum n}} = \pm\sqrt{\frac{0,018}{32}} = 0,024.$$

О линейности связи судят по достоверности статистического показателя:

$$t_1 = \frac{\xi}{m_\xi} = \frac{0{,}018}{0{,}024} = 0{,}75.$$

Поскольку достоверность меры линейности меньше трех следует сделать заключение об отсутствии криволинейной зависимости. Корреляция линейная.

Показатель криволинейности:

$$Kp = \frac{\eta^2 - r^2}{1 - r^2}. \qquad (2.11)$$

В нашем случае:

$$Kp = \frac{\eta^2 - r^2}{1 - r^2} = \frac{0{,}92^2 - 0{,}91^2}{1 - 0{,}91^2} = 0{,}10.$$

Таким образом, показатель криволинейности незначителен.

2.3 Непараметрические методы корреляционного анализа

Непараметрические методы позволяют определить связь не только количественных признаков, но и качественных. В основу этих методов положен принцип ранжирования значений статистического ряда.

Для оценки тесноты связи между несколькими признаками определяют *коэффициент конкордации* ω, вычисляемый по формуле:

$$\omega = \frac{12S}{m^2(n^3 - n)}, \qquad (2.12)$$

где m – число факторов;

n – число ранжируемых единиц;

S – сумма квадратов отклонений рангов.

Сумму квадратов отклонений рангов определяют следующим образом:

$$S=\sum_{1}^{n}\left(\sum_{1}^{m}r_{ij}\right)^{2}-\frac{\left(\sum_{1}^{n}\sum_{1}^{m}r_{ij}\right)^{2}}{n}. \quad (2.13)$$

Пример:

Студентом пятого курса Михаилом Казаковым под руководством Р.С. Хамитова были отобраны образцы шишек ели с клонов плюсовых деревьев на Диковской ЛСП. В качестве показателей семенной продуктивности были проанализированы масса семян содержащихся в одной шишке, масса 1000 шт. семян, урожайность семян с одного дерева.

Таблица 7 – Показатели семенной продуктивности клонов плюсовых деревьев

Номер клона	Показатели семенной продуктивности		
	масса семян в шишках, г	масса 1000 шт. семян, г	урожайность семян, г/дер.
55	0,34	5,70	11,60
139	1,02	7,48	41,99
141	0,81	7,18	34,10
143	1,03	7,65	16,50
181	0,55	5,62	15.23
232	0,59	6,54	5,94
234	1,33	7,65	29,12
236	1,32	7,34	20,60
239	1,11	7,42	22,64
250	0,98	7,91	34,90
252	0,74	5,83	9,68
253	0,68	6,72	17,89
254	1,25	5,88	14,71
270	1,08	8,49	20,80
271	0,37	4,63	8,79
274	0,84	6,66	33,30
275	0,93	7,79	14,73
442	0,95	7,00	4,70

Для расчета коэффициента конкордации проранжируем показатели семенной продуктивности, заполним таблицу 8, и рассчитаем вспомогательные коэффициенты:

$\sum_{1}^{m} r_{ij}$ и $\left(\sum_{1}^{m} r_{ij}\right)^2$.

Таблица 8 – Расчет вспомогательных коэффициентов для определения суммы квадратов отклонения рангов

Номер клона	Ранг по показателям			$\sum_{1}^{m} r_{ij}$	$\left(\sum_{1}^{m} r_{ij}\right)^2$
	масса семян в шишках	массы 1000 шт. семян	урожайности семян		
55	18	16	14	48	2304
139	7	6	1	14	196
141	12	9	3	24	576
143	6	4	10	20	400
181	16	17	11	44	1936
232	15	13	17	45	2025
234	1	4	5	10	100
236	2	8	8	18	324
239	4	7	6	17	289
250	8	2	2	12	144
252	13	15	15	43	1849
253	14	11	9	34	1156
254	3	14	13	30	900
270	5	1	7	13	169
271	17	18	16	51	2601
274	11	12	4	27	729
275	10	3	12	25	625
442	9	10	18	37	1369
Итого				512	17692

Сумма квадратов рангов:

$$S = 17692 - \frac{512^2}{18} = 3128.$$

Величина коэффициента конкордации:

$$\omega = \frac{12 \cdot 3128}{3^2\left(18^3 - 18\right)} = 0{,}72.$$

Вывод: *Масса семян содержащихся в одной шишке, масса 1000 шт. семян, урожайность семян с одного дерева в высокой степени взаимосвязаны между собой ($\omega = 0{,}72$).*

Коэффициент контингенции вычисляется в случаях, если наблюдается только наличие или отсутствие изучаемого признака или признак может принимать лишь два значения.

Пример:

По результатам исследования формового разнообразия сосны обыкновенной в 17-летних культурах Псковского лесхоза установлено, что процент левых изомеров у узкокронных форм составляет 68, правых – 32, а у ширококронных: левых – 32, правых – 68 (Маслаков и др., 1978)[1].

Для установления тесноты связи сведем в табл.9.

Таблица 9 – Вычисление коэффициента контингенции

Формы по кроне	Изомеры		Всего
	левые	правые	
Узкокронные	68 (a)	32 (b)	100
Ширококронные	32 (c)	68 (d)	100
Всего	100	100	200

Коэффициент контингенции (А) вычисляется по формуле:

$$A = \frac{ad - bc}{\sqrt{(a+b)(b+c)(a+c)(c+d)}}, \qquad (2.14)$$

где a,b,c,d - численности противоположных признаков.

Тогда коэффициент контингенции для частностей:

$$A = \frac{ad - bc}{10000}. \qquad (2.15)$$

Подставляя цифровые значения из таблицы 9 в формулу 2.15 получаем:

[1] Маслаков Е.Л. Формы сосны обыкновенной в культурах / Е.Л. Маслаков, А.М. Голиков, А.И. Толстопятенко // Восстановление леса на Северо-западе РСФСР. – Л.: ЛенНИИЛХ, 1978. – С.120-123.

$$A=\frac{ad-bc}{10000}=\frac{68\cdot 68-32\cdot 32}{10000}=0{,}36.$$

Основная ошибка коэффициента А вычисляется по формуле:

$$m_A=\frac{1-A^2}{\sqrt{N}}, \tag{2.16}$$

где N – общая численность выборочной совокупности.

Если данные в таблице приведены в виде процентов (частности) то ошибку определяют по формуле:

$$m_A=\frac{1-A^2}{\sqrt{200}}. \tag{2.17}$$

Для нашего примера:

$$m_A=\frac{1-0.36^2}{\sqrt{200}}=\frac{0{,}87}{14{,}1}=0{,}062.$$

Достоверность коэффициента сходства рассчитывается по формуле

$$t=\frac{A}{m_A}. \tag{2.18}$$

Следовательно

$$t=\frac{0{,}36}{0{,}062}=5{,}8.$$

Коэффициент сходства достоверен, поскольку $t = 5{,}8 > 4$.

Вывод: *Представленность левых и правых изомеров связана с формой кроны. Связь прямая, умеренная.*

В случае если в основу корреляционной решетки положены три, и более признаков для определения наличия связи рассчитывают критерий χ^2 («хи-квадрат»).

Пример:

В Чагринской роще С.М. Хамитовой у 42 плодоносящих кедров отобраны образцы шишек. Для каждого дерева была определены типичная форма развитой шишки и тип апофиза семенной чешуи. Результаты учета приведены в таблице.

Таблица 10 – Распределение деревьев по морфологическим формам шишек*

Форма апофиза	Форма шишки				Всего
	округлая (p_a)	яйцевидная (p_b)	коническая (p_c)	цилиндрическая (p_d)	
плоский (p_1)	- (0,2)	- (0,5)	2 (2,4)	8 (6,9)	$N_1=10$
бугорчатый (p_2)	1 (0,5)	1 (1,0)	3 (5,0)	16 (14,5)	$N_2=21$
крючковатый (p_3)	- (0,3)	1 (0,5)	5 (2,6)	5 (7,6)	$N_3=11$
Всего	$N_a=1$	$N_b=2$	$N_c=10$	$N_d=29$	$N=42$

Примечание: * - В скобках указаны теоретические частоты

Теоретические частоты рассчитывают по формуле:

$$p_{a1} = \frac{N_a \cdot N_1}{N}, \quad (2.19)$$

Расчетное значение χ^2 определяют по уравнению:

$$\chi^2 = \sum \frac{(n_{ij} - p_{ij})^2}{p_{ij}}. \quad (2.20)$$

$$\chi^2 = \frac{(0-0,2)^2}{0,2} + \ldots + \frac{(5-7,6)^2}{7,6} = 6,150.$$

Расчетное значение χ^2 сравнивают с табличным (приложение 4), при разных уровнях значимости. Для этого необходимо найти число степеней свободы. В нашем случае $k = (4-1)(3-1) = 6$. При уровне вероятности 0,99% табличный критерий χ^2 для данного числа степеней составляет 8,6. Таким образом, $\chi^2_{\text{расч}} < \chi^2_{\text{табл}}$. Следовательно, гипотеза о наличии связи не опровергается.

Далее можно рассчитать показатели тесноты связи: коэффициенты сопряженности К.Пирсона и А.А. Чупрова.

Коэффициент сопряженности К.Пирсона рассчитывается по формуле:

$$P = \sqrt{\frac{\chi^2}{N + \chi^2}}, \qquad (2.21)$$

а *коэффициент взаимной сопряженности А.А. Чупрова* определяют следующим образом:

$$C = \sqrt{\frac{\chi^2}{N\sqrt{(k_1 - 1)(k_2 - 1)}}}. \qquad (2.22)$$

Таким образом, получаем:

$$P = \sqrt{\frac{6,150}{42 + 6,150}} = \sqrt{0,128} = 0,36,$$

$$C = \sqrt{\frac{6,150}{42\sqrt{3 \cdot 2}}} = \sqrt{\frac{6,150}{102,9}} = 0,24.$$

Вывод: *Между формой шишки и типом апофиза наблюдается некоторая связь. Коэффициент сопряженности К. Пирсона указывает на наличие умеренной связи.*

3 РЕГРЕССИОННЫЙ АНАЛИЗ

Изучение связи величин двух или нескольких признаков при корреляционном анализе основано на установлении их тесноты и формы. Собственно математическое моделирование изучаемого явления связи разрабатывается на основе *регрессионного анализа*.

Полученные опытные данные величины результативного признака (y) по градациям факториальной оси графика (x) выстраиваются в виде ломаной линии. Задачей регрессионного анализа является выравнивание опытных данных при помощи аналитических уравнений. Выравнивание опытных данных в виде функции осуществляется по методу наименьших квадратов. Тренд располагается таким образом, чтобы сумма квадратов отклонений опытных данных от выровненных значений была наименьшей по сравнению с суммой квадратов отклонений, которые получаются при любом другом проведении выравнивания.

Характер влияния факториального признака изменение опытных данных, т.е. на предполагаемый тренд, учитывается при выборе аналитического уравнения.

Наиболее простой моделью является уравнение прямой линии. Такое уравнение используют, если с увеличением одного показателя наблюдается пропорциональное увеличение или уменьшение другого:

$$y = a + bx. \tag{3.1}$$

Парабола второго порядка используется, если при возрастании одного признака другой возрастает, при этом, с каждой градацией по факториальной оси возрастание увеличивается, т.е. тренд имеет один плавный изгиб:

$$y = a + bx + cx^2. \tag{3.2}$$

При S-образном характере изгиба можно использовать функцию:

$$y = ax^{b + c\lg x}. \tag{3.3}$$

Если при увеличении факториального признака результативный увеличивается замедленно, применяют логарифмические кривые:

$$y = a + b\lg x\,; \tag{3.4}$$

$$y = a + bx + c\lg x\,. \tag{3.5}$$

В случае, когда при возрастании одного признака другой интенсивно возрастает, для выравнивания берется показательная кривая:

$$y = ab^x\,. \tag{3.6}$$

В обратном случае выбирается уравнение гиперболы:

$$y = a + \frac{b}{x}\,. \tag{3.7}$$

При изначально постепенном увеличении результативного признака, но постепенном переходе в последующем к пропорциональное увеличение может быть адекватна степенная функция:

$$y = ax^b\,, \tag{3.8}$$

где у – зависимая переменная;

x – независимая переменная (факториального показателя);

a, b, c, d – постоянные неизвестные подлежащие расчету в уравнениях.

Пример:

В насаждении выборочно определили высоту и диаметр у 16 учетных деревьев:

H	*17.5*	*18,5*	*19,5*	*20*	*20,5*	*21*	*21,5*	*22.5*	*23*	*24*	*24*	*24,5*	*25,5*	*26.5*	*27,5*	*28*
D	*16,6*	*19,5*	*21,7*	*22,6*	*26,4*	*26,2*	*28*	*30,1*	*30,3*	*31,5*	*33,7*	*37,3*	*30,4*	*33,2*	*33,9*	*41,1*

Необходимо аппроксимировать зависимость высоты деревьев от их таксационного диаметра.

Вычисление уравнения прямой $y = a + bx$

Предварительно составляется два нормальных уравнения. Для первого нормального уравнения все члены уравнения умножаются на коэффициент *a*, т.е. на единицу, и суммируют. Для нахождения второго нормального уравнения все члены уравнения прямой умножаются на коэффициент *b*, т.е. на *x*, и суммируются.

Таким образом:

$$\left.\begin{aligned}\sum y &= an + b\sum x\\ \sum xy &= a\sum x + b\sum x^2\end{aligned}\right\}, \tag{3.9}$$

В соответствии с нормальными уравнениями составляется вспомогательная таблица.

Таблица 11 – Расчет уравнения прямой $y = a + bx$

Диаметр	Высота	x^2	xy	Высота по уравнению $\tilde{y}$	$(y-\tilde{y})^2$
18	18,1	324	325,8	20,3	5,38
20	24,2	400	484,0	23,9	15,04
22	29,5	484	649,0	27,4	29,55
24	34,2	576	820,8	31,0	48,92
26	31,8	676	826,8	34,6	73,14
28	37,5	784	1050,0	38,1	102,21
Всего: 138	175,3	3244	4156,4	175,3	274,24

Данные таблицы подставляются в систему уравнений:

$$\begin{cases}175{,}3 = 6a + 138b\\ 4156{,}4 = 138a + 3244b\end{cases},$$

Каждая часть уравнения разделяется на свой коэффициент a, тогда:

$$\begin{cases} 29{,}217 = a + 23b \\ 30{,}119 = a + 23{,}507b \end{cases}$$

Для нахождения коэффициента b (коэффициента регрессии) из второго уравнения вычитается первое:

$$\begin{array}{l} -\begin{array}{l} 30{,}119 = a + 23{,}507b \\ \underline{29{,}217 = a + 23b} \end{array} \\ 0{,}902 = 0{,}507b \end{array}$$

$$b = \frac{0{,}902}{0{,}507} = 1{,}779.$$

Тогда:

$30{,}119 = a + 23{,}507b;$

$30{,}119 = a + 23{,}507 \cdot 1{,}779;$

$a = 30{,}119 - 41{,}821 = -11{,}702.$

Следовательно, уравнение регрессии примет вид:

$$\tilde{y} = 1{,}779x - 11{,}702.$$

Для получения высоты по уравнению в него подставляются диаметры по ступеням.

Основная ошибка уравнения вычисляется:

$$m_y = \pm\sqrt{\frac{\sum(y-\tilde{y})^2}{n-e}}, \qquad (3.10)$$

где $\sum(y-\tilde{y})^2$ – сумма квадратов отклонений между опытными и расчетными высотами;

n – количество точек по которым вычислялось уравнение;

e – количество коэффициентов в уравнении.

В нашем примере:

$$m_y = \pm\sqrt{\frac{274{,}24}{6-2}} = \pm 8{,}28 \text{ м}$$

Вычисление уравнения параболы второго порядка $y = a + bx + cx^2$

Для вычисления уравнения параболы второго порядка решается система уравнений:

$$\begin{cases} \sum y = an + b\sum x + c\sum x^2 \\ \sum xy = a\sum x + b\sum x^2 + c\sum x^2 \\ \sum x^2 y = a\sum x^2 + b\sum x^3 + c\sum x^4 \end{cases} \qquad (3.11)$$

Для расчета заполняется вспомогательная таблица

Таблица 12 – Вычисление уравнения параболы второго порядка

Диаметр x, см	Высота y, м	x^2	x^3	x^4	xy	x^2y	Высота по уравнению $\tilde{y}$	$(y-\tilde{y})^2$
18	18,1	324	5832	104976	325,8	5864,4	18,0	0,01
20	24,2	400	8000	160000	484,0	9680,0	24,3	0,01
22	29,5	484	10648	234256	649,0	14278,0	29,3	0,02
24	34,2	576	13824	331776	820,8	19699,2	32,9	1,69
26	31,8	676	17576	456976	826,8	21496,8	35,0	10,24
28	37,5	784	21952	614656	1050,0	29400,0	35,9	2,56
Всего: 138	175,3	3244	77832	1902640	4156,4	100418,4	175,4	14,53

В нашем примере:

$$\begin{cases} 175{,}3 = 6a + 138b + 3244c \\ 4156{,}4 = 138a + 3244b + 77832c \\ 100418{,}4 = 3244a + 77832b + 1902640c \end{cases}$$

Части уравнения делятся на свой коэффициент a:

$$\begin{cases} 29{,}217 = a + 23b + 540{,}667c \\ 30{,}119 = a + 23{,}507b + 564c \\ 30{,}955 = a + 23{,}992b + 586{,}510c \end{cases}$$

Из третьего уравнения системы последовательно вычитается первое и второе:

$$\begin{array}{r} -\begin{array}{l} 30{,}955 = a + 23{,}992b + 586{,}510c \\ \underline{29{,}217 = a + 23b + 540{,}667c} \end{array}, \\ 1{,}738 = 0{,}992b + 45{,}843c \end{array}$$

$$\begin{array}{r} -\begin{array}{l} 30{,}955 = a + 23{,}992b + 586{,}510c \\ \underline{30{,}119 = a + 23{,}507b + 564c} \end{array}. \\ 0{,}836 = 0{,}485b + 22{,}510c \end{array}$$

Решается система уравнений для коэффициентов b и c:

$$\begin{cases} 1{,}738 = 0{,}992b + 45{,}843c \\ 0{,}836 = 0{,}485b + 22{,}510c \end{cases}.$$

Обе части уравнения делятся на свой коэффициент b, и решается система уравнений:

$$\begin{cases} 1{,}752 = b + 46{,}213c \\ 1{,}724 = b + 46{,}412c \end{cases}$$

Находится разность уравнений:

$$\begin{array}{r} -\begin{array}{l} 1{,}724 = b + 46{,}412c \\ \underline{1{,}752 = b + 46{,}213c} \end{array}, \\ -0{,}028 = 0{,}199c \end{array}$$

$$c = \frac{-0{,}028}{0{,}199} = -0{,}171,$$

тогда

$$1{,}752 = b + 46{,}213 \cdot (-0{,}171),$$

$$b = 1{,}752 + 7{,}902 = 9{,}654 .$$

Находится коэффициент a:

$$29{,}217 = a + 23 \cdot 9{,}654 + 540{,}667 \cdot (-0{,}171)$$

$$a = 29{,}217 - 222{,}042 + 92{,}454 = -100{,}371$$

Таким образом:

$$\tilde{y} = -100{,}371 + 9{,}654x - 0{,}171x^2 .$$

Основная ошибка уравнения:

$$m_y = \pm\sqrt{\frac{14{,}53}{6-2}} = \pm 1{,}91\text{ м}.$$

Выбор оптимального уравнения:

Оптимальная модель связи двух статистических величин выбирается на основании ошибки уравнений. В нашем примере наименьшая ошибка наблюдается у уравнения параболы второго порядка (m_y=±1,91 м).

Для наглядности опытные и выровненные данные изображаются графически.

Рис. 5. Зависимость высоты дерева от таксационного диаметра ствола

4 ДИСПЕРСИОННЫЙ АНАЛИЗ

Задачей *дисперсионного анализа*, предложенного Р.Э. Фишером, является выявление статистического влияния одного или нескольких факторов на результативный признак. Влияние различных факторов по-разному отражается на варьировании признака. Вариация признака, выраженная суммой квадратов отклонений, называется дисперсией. Дисперсионный анализ по количеству изучаемых факторов, положенных в основание комплекса, подразделяется на однофакторный, двухфакторный и многофакторный. Общая дисперсия признака складывается из одной или нескольких дисперсий обусловленной действием факторов и случайной дисперсии. Метод анализа основан на разложении общей дисперсии статистического комплекса на составляющие компоненты:

$$SS_O = SS_\Phi + SS_C, \tag{4.1}$$

где SS_O – общая дисперсия;

SS_Φ – факториальная дисперсия;

SS_C – случайная дисперсия.

Сопоставление факториальной дисперсии с общей показывает *долю влияния фактора*:

$$\eta^2 = \frac{SS_\Phi}{SS_O}, \tag{4.2}$$

где η^2 – показатель силы влияния (индекс детерминации).

Показатель достоверности влияния определяется по *критерию Фишера*:

$$F = \frac{\sigma_\Phi^2}{\sigma_c^2}, \tag{4.3}$$

где F – эмпирический критерий достоверности силы влияния;

σ_Φ^2 – факториальная варианса;

σ_c^2 – случайная варианса.

Факториальная варианса рассчитывается по формуле:

$$\sigma_{\phi}^{2} = \frac{SS_{\phi}}{g-1}, \tag{4.4}$$

где g – число градаций изучаемого фактора;

$g - 1$ – число степеней свободы.

Случайная варианса:

$$\sigma_{c}^{2} = \frac{SS_{c}}{N-g}, \tag{4.5}$$

где N – численность всего комплекса;

$N - g$ – число степеней свободы.

Ошибка показателя силы влияния:

$$m_{\eta} = \pm(1-\eta^{2})\frac{g-1}{N-g}. \tag{4.6}$$

4.1 Однофакторный дисперсионный комплекс

Пример:

Имеются результаты определения средней высоты сосны на пробных площадях в разных условиях местопроизрастания (табл. 13):

Таблица 13 – Средняя высота сосны по типам леса

Вариант опыта (ТУМ)	Средняя высота на пробных площадях, м
Лишайниковый	18,5; 17,5; 18; 18
Брусничный	18,5; 18; 18; 20
Черничный	19,5; 20; 20,5; 19,5
Кисличный	20; 20,5; 22; 21

Пробные площади были заложены таким образом, чтобы исключить влияние прочих факторов на результативный признак (одинаковый средний возраст, подзона тайги и т.д.).

В ходе исследования предстоит установить, влияет ли тип условий местопроизрастания (ТУМ) на рост насаждений сосны.

Исходные данные по вариантам опыта (типам леса) заносятся в таблицу 14.

Таблица 14 – Однофакторный дисперсионный комплекс

Вариант опыта (ТУМ)	Значение признака (средняя высота, м)	Число повторностей по вариантам	Сумма высот Σy, м	Среднее значение признака M, м
Лишайниковый	18,5; 17,5; 18; 18	4	72,0	18
Брусничный	18,5; 18; 18; 20	4	74,5	18,6
Черничный	19,5; 20; 20,5; 19,5	4	79,5	19,9
Кисличный	20; 20,5; 22; 21	4	83,5	20,9
Всего	–	16	309,5	–

Затем рассчитываются средние значения по вариантам (М) и по всему комплексу (M_0).

Средние значения по вариантам:

$$M_k = \frac{\sum y}{n_k}, \tag{4.7}$$

где Σy – сумма значений изучаемого признака внутри варианта;

n_k – число повторностей (единиц наблюдения) в варианте.

$$M_1 = \frac{\sum y}{n_1} = \frac{72}{4} = 18;$$

$$M_2 = \frac{\sum y}{n_2} = \frac{74{,}5}{4} = 18{,}6;$$

$$M_3 = \frac{\sum y}{n_3} = \frac{79{,}5}{4} = 19{,}9;$$

$$M_4 = \frac{\sum y}{n_4} = \frac{83{,}5}{4} = 20{,}9.$$

Среднее значение по комплексу:

$$M_o = \frac{\sum\sum y}{N}, \qquad (4.8)$$

где $\Sigma\Sigma y$ – сумма всех вариант комплекса;

N – общее количество единиц наблюдений в комплексе.

$$M_o = \frac{309{,}5}{16} = 19{,}34\,.$$

Рассчитывается сумма квадратов по вариантам:

$$\sum S = \sum S_1 + \sum S_2 + \sum S_3 + \ldots + \sum S_k\,, \qquad (4.9)$$

где ΣS_1, ΣS_2, ΣS_3, …, ΣS_k – средние квадраты вариант комплекса рассчитываемые по формуле:

$$\sum S_k = \frac{\left(\sum y_k\right)^2}{n_k}\,. \qquad (4.10)$$

В нашем случае:

$$\sum S_1 = \frac{72{,}0^2}{4} = 1296{,}0\,;$$

$$\sum S_2 = \frac{74{,}5^2}{4} = 1387{,}6\,;$$

$$\sum S_3 = \frac{79{,}5^2}{4} = 1580{,}1\,;$$

$$\sum S_4 = \frac{83{,}5^2}{4} = 1743{,}1\,.$$

$$\sum S = 1296{,}0 + 1387{,}6 + 1580{,}1 + 1780{,}1 = 6006{,}8\,.$$

Сумма квадратов всех высот по комплексу:

$$\sum (y)^2 = y_1^2 + y_2^2 + y_3^2 + \ldots + y_k^2\,. \qquad (4.11)$$

В нашем примере:

$\Sigma(y)^2 = 18{,}5^2 + 17{,}5^2 + 18{,}0^2 + 18{,}0^2 + 18{,}5^2 + 18{,}0^2 + 18{,}0^2 + 20{,}0^2 +$

$$+ 19{,}5^2 + 20{,}0^2 + 20{,}5^2 + 19{,}5^2 + 20{,}0^2 + 20{,}5^2 + 22{,}0^2 + 21{,}0^2 = 6012{,}8$$

Рассчитанные вспомогательные величины будут полезны для нахождения дисперсий:

$$SS_{\Phi} = \sum S - \frac{\left(\sum\sum y\right)^2}{N} = 6006{,}8 - \frac{309{,}5^2}{16} = 19{,}9\,;$$

$$SS_C = \sum (y)^2 - \sum S = 6012{,}8 - 6006{,}8 = 6{,}0\,;$$

$$SS_O = \sum (y)^2 - \frac{\left(\sum\sum y\right)^2}{N} = 6012{,}8 - \frac{309{,}5^2}{16} = 25{,}9\text{, или}$$

$$SS_O = SS_{\Phi} + SS_C = 19{,}9 + 6{,}0 = 25{,}9\,.$$

Теперь можно рассчитать показатель силы влияния:

$$\eta^2 = \frac{SS_{\Phi}}{SS_O} = \frac{19{,}9}{25{,}9} = 0{,}77\,.$$

Корреляционное отношение, показывающее тесноту связи изучаемого признака:

$$\eta = \sqrt{0{,}77} = 0{,}88\,.$$

Поскольку корреляционное отношение составляет 0,88, теснота связи высокая (более 0,71, но менее 0,91).

Далее необходимо рассчитать критерий Фишера, служащий показателем достоверности силы влияния. Для этого следует установить межгрупповую (факториальную) вариансу σ_{ϕ}^2 ($MS_{\text{межгр}}$):

$$\sigma_{\Phi}^2 = \frac{SS_{\Phi}}{g-1} = \frac{19{,}9}{4-1} = 6{,}63\,,$$

случайную вариансу σ_c^2 ($MS_{\text{внутригр}}$):

$$\sigma_c^2 = \frac{SS_C}{N-g} = \frac{6{,}0}{16-4} = 0{,}50\,.$$

Критерий Фишера:

$$F = \frac{\sigma_{\Phi}^2}{\sigma_c^2} = \frac{6{,}63}{0{,}50} = 13{,}3\,.$$

Рассчитанный показатель достоверности силы влияния сравнивают с табличным (стандартным) критерием Фишера при разном уровне значимости. Табличное значение определяется по приложению 3. Число степеней свободы в нашем случае $df_1 = g - 1 = 4 - 1 = 3$; $df_2 = N - g = 16 - 4 = 12$. Стандартное значение при 5% уровне значимости – 8,7, при 1% уровне значимости – 27,3. Таким образом, $F_ф > F_{05}$ (13,3 > 8,7), но $F_ф < F_{01}$ (13,3 < 27,3). Следовательно влияние условий местопроизрастания на высоту насаждений доказано на 5% уровне значимости.

Ошибка показателя силы влияния:

$$m_\eta = \pm\left(1-\eta^2\right)\frac{g-1}{N-g} = \pm(1-0{,}77)\frac{3}{12} = \pm 0{,}08.$$

Таким образом, сила влияния типа условий местопроизрастания на высоту формируемых насаждений составляет 0,77±0,8. Следовательно, высота насаждений на 77% предопределена типом условий местопроизрастания.

Результаты исследований заносят в итоговую таблицу 15.

Таблица 15 – Результаты дисперсионного анализа однофакторного комплекса

Источник вариации	Дисперсия SS	Степень свободы df	Варианса MS	$F_ф$	F_{05}	$\eta \pm m_\eta$
Межгрупповая (факториальная)	19,9	3	6,63	13,3	8,7	0,77±0,8
Внутригрупповая (случайная)	6,0	12	0,50	–	–	
Общая	25,9	15	–	–	–	

Дисперсионный анализ выявляет лишь наличие, или отсутствие факта различия между вариантами. Однако, различия между конкретными вариантами может и не быть. Поэтому в случае, если достоверность влияния фактора на результативный признак установлена, необходимо оценить различия между самими вариантами.

Наиболее распространенным методом оценки этого различия является выявление наименьшей существенной разности (НСР).

Для ее расчета сначала определяют разность между средними величинами признака вариантов:

$$d = M_1 - M_2 . \tag{4.12}$$

Затем находят ошибку разности:

$$m_d = \sqrt{\sigma_c^2 \left(\frac{n_1 + n_2}{n_1 n_2} \right)}, \tag{4.13}$$

где σ_c^2 – внутригрупповая варианса ($MS_{внутригр}$);

n_1, n_2 – численности сравниваемых групп.

Если число сравниваемых повторностей равно (при равномерном поиске) формула примет вид:

$$m_d = \sqrt{\frac{2\sigma_c^2}{n}} . \tag{4.14}$$

Вычисленные разности заносят в вспомогательную таблицу-решетку 16.

Таблица 16 – Разность средних значений по вариантам

Тип условий местопроизра- стания	Разность средних значений по ТУМ			
	лишайнико-вый	брусничный	черничный	кисличный
Лишайниковый	0	0,6	1,9	2,9
Брусничный		0	1,3	2,3
Черничный			0	1,0
Кисличный				0

Ошибка разности средних в данном случае для всех вариантов:

$$m_d = \sqrt{\sigma_c^2 \left(\frac{n_1 + n_2}{n_1 n_2} \right)} = \sqrt{0{,}5 \left(\frac{4+4}{4 \cdot 4} \right)} = 0{,}7 ,$$

или

$$m_d = \sqrt{\frac{2\sigma_c^2}{n}} = \sqrt{\frac{2 \cdot 0{,}5}{4}} = 0{,}7.$$

Далее рассчитывают НСР (*наименьшую существенную разность*) между средними данными вариантов по формуле:

$$HCP = t \cdot m_d, \tag{4.15}$$

где t – табличное значение критерия Стьюдента при заданном уровне вероятности.

При вероятности 0,95 и числе степеней свободы случайной вариансы df = 12 стандартное значение критерия Стьюдента – t_{st} = 2,2 (приложение 2).

В нашем случае:

$$HCP_{05} = t \cdot m_d = 2{,}2 \cdot 0{,}7 = 1{,}5.$$

Следовательно, если между вариантами наблюдается разница 1,5 м, то можно говорить о наличии существенного различия между ними. Исходя из данных о различии высоты исследуемых насаждений приведенных в таблице 14, очевидно, что достоверного существенного различия не наблюдается между лишайниковыми и брусничными, а также между брусничными и черничными, кисличными и черничными условиями местопроизрастания. В остальных случаях различия достоверны.

Для наглядности полученные результаты приводят в итоговой таблице, проранжировав их по возрастанию результативного признака (табл. 17).

Таблица 17 – Средняя высота насаждений по вариантам опыта

Вариант опыта (ТУМ)	Среднее значение признака М, м	Процент к лучшему варианту	Ранг
Кисличный	20,9	100	I
Черничный	19,9	95	I
Брусничный	18,6	89	II
Лишайниковый	18,0	86	III
HCP_{05}	1,5	7	–

Вывод: *Высота насаждений при 95% вероятности безошибочного заключения на 77% предопределена типом условий местопроизрастания. Достоверно наибольшей высоты к моменту исследований достигли насаждения в кисличном типе условий местопроизрастания.*

4.2 Двухфакторный дисперсионный комплекс без повторений

Общая дисперсия при *двухфакторном дисперсионном анализе* складывается из двух факториальных дисперсий и дисперсии ошибки (случайной):

$$SS_o = SS_A + SS_B + SS_c, \tag{4.16}$$

где SS_0 – общая дисперсия;

SS_A – факториальная дисперсия первого фактора;

SS_B – факториальная дисперсия второго фактора;

SS_c – случайная дисперсия.

Этот метод анализа полезен и при изучении рассеяния под действием одного фактора, но с учетом повторений:

$$SS_o = SS_Ф + SS_П + SS_c, \tag{4.17}$$

где $SS_Ф$ – факториальная дисперсия первого фактора;

$SS_П$ –дисперсия повторений;

SS_c – случайная дисперсия.

Пример:

Имеются результаты определения средней высоты сосны на пробных площадях в разных условиях местопроизрастания и с градациями по вариантам внесения удобрений (табл. 18):

Таблица 18 – Средняя высота сосны по типам леса и вариантам внесения удобрений

Варианты по ТУМ	Варианты по внесению удобрений			
	Контроль	N	NP	NPK
Лишайниковый	18,5	19,5	20,0	20,5
Брусничный	19,5	20,0	20,5	21,0
Черничный	20,5	21,0	21,5	22,0
Кисличный	21,5	22,0	22,5	23,0

Следует установить доли влияния факторов на высоту насаждений.

Сначала рассчитывают вспомогательные данные: сумму высот, средние значения признака и дисперсии по вариантам, аналогично п.4.1, раздельно по строкам и столбцам. Результаты заносят в таблицу 19.

Таблица 19 – Двухфакторный дисперсионный комплекс без повторений

Вариант опыта (ТУМ)	Число повторностей по вариантам n	Сумма высот Σy, м	Среднее значение признака M,	$\Sigma S=(\Sigma y)^2/N$
Лишайниковый	4	78,5	19,6	1540,56
Брусничный	4	81,0	20,3	1640,25
Черничный	4	85,0	21,3	1806,25
Кисличный	4	89,0	22,3	1980,25
Всего	N=16	$\Sigma\Sigma y$=333,5	–	6967,31
Контроль	4	80,0	20,0	1600,00
N	4	82,5	20,6	1701,56
NP	4	84,5	21,1	1785,06
NPK	4	86,5	21,6	1870,56
Всего	N=16	$\Sigma\Sigma y$=333,5	–	6957,18

Сумма квадратов всех высот по комплексу:

$$\sum (y)^2 = y_1^2 + y_2^2 + y_3^2 + \ldots + y_k^2 = 6973{,}25.$$

Дисперсии:

$$SS_A = \sum S_A - \frac{\left(\sum\sum y\right)^2}{N} = 6967{,}31 - \frac{333{,}5^2}{16} = 15{,}92;$$

$$SS_B = \sum S_B - \frac{\left(\sum\sum y\right)^2}{N} = 6957{,}18 - \frac{333{,}5^2}{16} = 5{,}79;$$

$$SS_O = \sum (y)^2 - \frac{\left(\sum\sum y\right)^2}{N} = 6967{,}31 - \frac{333{,}5^2}{16} = 21{,}86;$$

$$SS_C = SS_O - \left(SS_A + SS_B\right) = 21{,}86 - (15{,}92 + 5{,}79) = 0{,}15.$$

Вариансы (MS):

$$\sigma_A^2 = \frac{SS_A}{g_A - 1} = \frac{15{,}92}{3} = 5{,}31;$$

$$\sigma_B^2 = \frac{SS_B}{g_B - 1} = \frac{5{,}79}{3} = 1{,}93;$$

$$\sigma_C^2 = \frac{SS_C}{(g_A - 1)\cdot(g_B - 1)} = \frac{0{,}15}{3\cdot 3} = 0{,}02.$$

Критерий Фишера:

$$F_A = \frac{\sigma_A^2}{\sigma_c^2} = \frac{5{,}31}{0{,}02} = 265{,}5;$$

$$F_B = \frac{\sigma_A^2}{\sigma_c^2} = \frac{1{,}93}{0{,}02} = 96{,}5.$$

Результаты исследований заносят в итоговую таблицу 20.

Влияние обоих факторов достоверно при 99% вероятности безошибочного наблюдения, поскольку $F_A > F_{01}$ и $F_B > F_{01}$.

Таблица 20 – Результаты дисперсионного анализа двухфакторного комплекса без повторений

Источник вариации	Дисперсия SS	Степень свободы df	Варианса MS	$F_ф$	F_{05}	F_{01}
Межгрупповая А (строки)	15,92	3	5,31	265,5	8,8	27,3
Межгрупповая В (столбцы)	5,79	3	1,93	96,5	8,8	27,3
Остаточная (погрешность)	0,15	9	0,02	–	–	–
Общая	21,86	15	–	–	–	–

Рассчитываются показатели силы (доли) влияния факторов.

Фактора А (ТУМ):

$$\eta_A^2 = \frac{SS_A}{SS_O} = \frac{15{,}92}{21{,}86} = 0{,}73;$$

фактора В (внесение удобрений):

$$\eta_B^2 = \frac{SS_B}{SS_O} = \frac{5{,}79}{21{,}86} = 0{,}26;$$

прочих факторов

$$\eta_C^2 = \frac{SS_C}{SS_O} = \frac{0{,}15}{21{,}86} = 0{,}01.$$

На 73% высота насаждений обусловлена типом условий местопроизрастания, а на 26% – внесением удобрений.

Корреляционное отношение:

$$\eta_A = \sqrt{0{,}73} = 0{,}85;$$

$$\eta_B = \sqrt{0{,}26} = 0{,}51.$$

Зависимость высоты насаждений от типа условий местопроизрастания высокая, а от варианта внесения удобрений – значительная.

Для определения достоверного различия между вариантами рассчитывают НСР отдельно по результатам влияния факторов А и В. Предварительно рассчитывают решетки разностей по отдельным факторам.

В нашем примере предварительно определяют разность значений по вариантам типа условий местопроизрастания (табл. 21).

Таблица 21 – Разность средних значений по вариантам типа условий местопроизрастания

Тип условий местопроизрастания	Разность средних значений по ТУМ			
	лишайниковый	брусничный	черничный	кисличный
Лишайниковый	0,0	0,7	1,7	2,7
Брусничный		0,0	1,0	2,0
Черничный			0,0	1,0
Кисличный				0,0

Определяют ошибку разности:

$$m_d = \sqrt{\frac{2\sigma_c^2}{n}} = \sqrt{\frac{2 \cdot 0{,}15}{4}} = 0{,}3\,.$$

Наименьшее существенное различие при $df = 9$ и 1% уровне значимости:

$$HCP_{01} = t \cdot m_d = 3{,}3 \cdot 0{,}3 = 1{,}0\,.$$

Результаты приводят в итоговой таблице 22.

Таблица 22 – Средняя высота насаждений по вариантам опыта

Вариант опыта (ТУМ)	Среднее значение признака M, м	Процент к лучшему варианту	Ранг
Кисличный	22,3	100	I
Черничный	21,3	96	II
Брусничный	20,3	91	III
Лишайниковый	19,6	88	III
$НСР_{01}$	1,0	4	–

Аналогичным образом определяют существенность различий по фактору В (табл.23).

Таблица 23 – Разность средних значений по вариантам типа условий местопроизрастания

Тип условий местопроизрастания	Разность средних значений по ТУМ			
	Контроль	N	NP	NPK
Контроль	0,0	0,6	1,1	1,6
N		0,0	0,5	1,0
NP			0,0	0,5
NPK				0,0

НСР по фактору В здесь определять не следует, т.к. он будет равен предыдущему. Результаты приводят в итоговой таблице 24. Ранжировку в случаях, если имеется контрольный вариант опыта, производят от него.

Таблица 24 – Средняя высота насаждений по вариантам опыта

Вариант опыта (ТУМ)	Среднее значение признака *M*, м	Процент к контролю	Ранг по отношению к контролю
NPK	21,6	108	I
NP	21,1	106	I
N	20,6	103	II
Контроль (st)	20,0	100	–
$НСР_{01}$	1,0	5	–

Вывод: *Влияние типа условий местопроизрастания сказываются более существенно, чем внесение минеральных удобрений. Доля влияния первого фактора – 0,73, второго – 0,26. Высота насаждений увеличивается с улучшением типа условий местопроизрастания и при внесении комплекса макроэлементов.*

4.2 Двухфакторный дисперсионный комплекс с повторениями

Пример:

Заложен опыт по выявлению влияния рубок ухода в разрезе типов условий местопроизрастания (табл. 25). В каждом из четырех типах условий местопроизрастания заложили по две пробные площади, где осуществили уход и по две, где его не проводили (контроль).

Таблица 25 – Результаты определения средней высоты насаждений по вариантам опыта

ТУМ	Средняя высота насаждений, м	
	контроль	рубки ухода
Лишайниковый	18,5; 19,5	20,0; 20,5
Брусничный	19,5; 20,0	20,5; 21,0
Черничный	20,5; 21,0	21,5; 22,0
Кисличный	21,5; 22,0	22,5; 23,0

Рассчитываются вспомогательные величины:

Средний квадрат суммы всех вариант комплекса:

$$S_{\Sigma} = \frac{\left(\sum\sum y\right)^2}{N} = \frac{333{,}5^2}{16} = 6951{,}39.$$

Сумма средних квадратов суммы по всем градациям комплекса:

$$\sum S = \sum \frac{\left(\sum y_k\right)^2}{n_k}. \qquad (4.18)$$

В нашем примере:

$\Sigma S = (18{,}5 + 19{,}5)^2 : 2 + (20{,}0 + 20{,}5)^2 : 2 + (19{,}5 + 20{,}0)^2 : 2 + \ldots +$
$+ (22{,}5 + 23{,}0)^2 : 2 = 6972{,}88.$

Сумма квадратов всех вариант:

$$\sum (y)^2 = y_1^2 + y_2^2 + y_3^2 + \ldots + y_k^2 = 6973{,}25.$$

Дисперсия по первому фактору:

$$SS_A = \sum S_A - S_{\Sigma}, \qquad (4.19)$$

где ΣS_A – сумма средних квадратов сумм каждой градации первого фактора,

S_A – сумма средних квадратов по каждой градации первого фактора определяется:

$$S_A = \frac{\left(\sum y_A\right)^2}{n_A}. \qquad (4.20)$$

Результаты заносят в таблицу 26:

Таблица 26 – Вспомогательная таблица для расчета дисперсии по первому фактору

Градация фактора	n	Σy	S_A	SS_A
Лишайниковый	4	78,5	1540,56	SS_A=6967,31-6951,39=15,92
Брусничный	4	81,0	1640,25	
Черничный	4	85,0	1806,25	
Кисличный	4	89,0	1980,25	
Итого	16	333,5	6967,31	

Дисперсия по второму фактору:

$$SS_B = \sum S_B - S_\Sigma \cdot \quad (4.21)$$

$$S_\Sigma = \frac{\left(\sum\sum y\right)^2}{N} = \frac{333{,}5^2}{16} = 6951{,}39.$$

Таблица 27 – Вспомогательная таблица для расчета дисперсии по второму фактору

Градация фактора	n	Σy	S_A	SS_B
Контроль	8	162,5	3300,78	SS_B=6955,91-6951,39=4,52
Рубки ухода	8	171,0	3655,13	
Итого	16	333,5	6955,91	

Дисперсия по сочетанию факторов АВ:

$$SS_{AB} = SS_X - SS_A - SS_B, \quad (4.22)$$

где SS_X – дисперсия по суммарному взаимодействию,

$$SS_X = \sum S - S_\Sigma, \quad (4.23)$$

SS_A, SS_B – дисперсии по первому и второму фактору.

Составляется вспомогательная таблица для расчета сочетания факторов А и В (табл. 28).

Таблица 28 – Вспомогательная таблица для расчета дисперсии по первому фактору

Фактор А	Фактор В	n	Σy	Σy^2	S	$\Sigma(y)^2$
1	1	2	38,0	1444,0	722,00	722,50
	2	2	40,5	1640,3	820,13	820,25
2	1	2	39,5	1560,3	780,13	780,25
	2	2	41,5	1722,3	861,13	861,25
3	1	2	41,5	1722,3	861,13	861,25
	2	2	43,5	1892,3	946,13	946,25
4	1	2	43,5	1892,3	946,13	946,25
	2	2	45,5	2070,3	1035,13	1035,30
Итого		16	333,5	13944,1	6971,91	6973,25

Сред квадрат суммы всех вариант комплекса:

$S_{\Sigma} = 6951{,}39.$

Дисперсия по суммарному действию обоих факторов:

$$SS_X = \sum S - S_{\Sigma} = 6971{,}91 - 6951{,}39 = 20{,}52\,.$$

Дисперсия сочетания факторов:

$$SS_{AB} = SS_X - SS_A - SS_B = 20{,}52 - 15{,}92 - 4{,}52 = 0{,}08\,.$$

Случайная дисперсия по суммарному действию неорганизованных факторов:

$$SS_C = \sum (v)^2 - \sum S = 6973{,}25 - 6971{,}91 = 1{,}34\,.$$

Общая дисперсия всего комплекса:

$$SS_O = \sum (v)^2 - S_{\Sigma} = 6973{,}25 - 6951{,}39 = 21{,}86\,.$$

Показатель силы влияния:

1. Первого фактора (типа условий местопроизрастания):

$$\eta_A^2 = \frac{SS_A}{SS_O} = \frac{15{,}92}{21{,}86} = 0{,}73\,.$$

2. Второго фактора (рубок ухода):

$$\eta_B^2 = \frac{SS_B}{SS_O} = \frac{4{,}52}{21{,}86} = 0{,}21\,.$$

3. Влияние сочетаний обоих факторов:

$$\eta_{AB}^2 = \frac{SS_{AB}}{SS_O} = \frac{0{,}08}{21{,}86} = 0{,}004\,.$$

4. Суммарное действие организованных факторов:

$$\eta_X^2 = \frac{SS_X}{SS_O} = \frac{20{,}52}{21{,}86} = 0{,}94\,.$$

5. Суммарное действие неорганизованных факторов:

$$\eta_X^2 = \frac{SS_C}{SS_O} = \frac{1{,}34}{21{,}86} = 0{,}06\,.$$

Для вычисления достоверности каждого вида факториальных влияний составляется таблица 29.

Таблица 29 – Результаты дисперсионного анализа двухфакторного комплекса с повторениями

Источник вариации	Дисперсия SS	Степень свободы df	Варианса MS	$F_{ф}$	F_{05}	F_{01}
Факториальная А	15,92	4 – 1 = 3	5,31	31,2	4,1	7,6
Факториальная В	4,52	2 – 1 = 1	4,52	26,6	5,3	11,3
Сочетания факторов А и В	0,08	3 x 1 = 3	0,03	0,2	4,1	–
Суммарная организованных факторов X	20,52	4 x 2 – 1 = 7	2,93	17,2	3,5	6,1
Остаток (случайная)	1,34	16 – 4 x 2 = 8	0,17	–	–	–
Общая	21,86	15	1,46	–		–

Вывод: *Наибольший вклад в высоту формируемых насаждений имеет тип условий местопроизрастания (73 %). Проведение рубок ухода в меньшей степени влияет на высоту насаждений – 21%.*

СПИСОК ЛИТЕРАТУРЫ

1. Гусев, И.И. Моделирование экосистем [Текст]: учебное пособие / И.И. Гусев. – Архангельск: Изд-во Арханг. гос. техн. ун-та, 2002. – 112 с.

2. Дворецкий, М.Л. Практическое пособие по вариационной статистике [Текст] / М.Л. Дворецкий. – М.: Лесная пром-сть, 1971. – 102 с.

ПРИЛОЖЕНИЯ

Приложение 1

Исходные данные для выполнения задания

1		2		3		4		5	
h, м	*d, см*	*h, м*	*d, см*	*h, м*	*d, см*	*h, м*	*d, см*	*h, м*	*d, см*
18,5	17,9	18,5	17,9	18,5	17,9	18,5	17,8	17,0	15,8
17,5	16,2	18,0	16,2	17,0	16,3	17,0	16,4	17,5	17,4
18,0	21,6	17,5	14,1	17,5	18,6	17,5	18,3	17,5	18,1
18,0	20,3	18,0	20,3	18,0	20,3	18,0	20,6	18,5	20,3
18,5	18,3	18,0	18,2	18,0	18,4	19,0	18,6	18,0	20,6
18,0	20,0	18,5	20,0	18,5	20,2	19,5	20,9	19,0	20,5
18,0	19,2	20,5	21,2	20,5	21,4	20,5	21,2	20,5	21,4
20,0	18,0	19,5	18,6	19,5	18,8	19,5	22,0	19,5	21,9
19,5	21,9	20,0	21,9	20,0	21,9	20,0	23,0	20,5	23,2
20,0	20,2	20,0	20,2	20,0	22,2	20,0	22,8	20,0	22,5
20,5	19,5	20,5	22,1	20,5	23,5	20,5	24,4	20,5	24,6
19,5	19,1	19,5	23,0	19,5	23,0	21,0	25,7	20,0	25,5
20,0	18,6	22,0	25,8	21,5	25,3	21,5	26,0	21,0	25,8
20,5	21,5	20,5	23,7	22,0	28,7	22,0	26,8	21,5	26,6
22,0	22,0	22,0	27,2	22,0	27,2	22,0	28,7	22,5	28,2
21,0	25,6	22,0	28,4	22,0	28,4	22,5	30,3	22,0	29,4
22,0	23,0	21,0	30,2	21,0	30,2	22,5	32,1	22,5	31,4
22,5	31,5	22,0	31,0	22,0	31,0	23,0	25,9	22,0	32,7
22,0	32,3	24,5	27,5	24,5	27,5	23,5	28,4	23,5	28,0
24,0	28,0	24,0	28,0	24,0	32,0	23,5	28,0	23,5	29,6
23,0	29,6	23,0	30,3	23,0	30,3	24,0	30,2	24,0	30,0
24,0	33,7	24,0	33,7	24,0	33,6	24,0	31,8	24,0	31,4
23,5	30,8	23,5	34,7	23,5	36,1	24,5	32,6	24,5	33,6
24,0	33,2	25,5	33,1	25,5	33,2	24,5	34,2	24,5	34,5
26,0	34,0	26,0	33,8	26,0	32,4	25,0	29,8	24,0	36,0
25,0	34,5	25,0	34,5	25,0	34,3	25,5	31,5	25,0	29,7
26,0	35,7	26,0	35,7	26,0	35,7	26,0	32,7	26,0	32,3
26,5	36,8	26,5	36,8	26,5	37,8	26,5	35,8	26,0	31,0
28,0	28,4	28,0	33,5	28,0	32,5	26,5	36,5	26,5	32,4
27,0	29,3	27,0	36,9	27,0	32,9	27,0	35,9	26,5	35,3
28,0	38,7	28,0	37,7	28,0	37,7	28,0	39,3	27,0	36,2
28,5	41,3	28,5	41,7	28,5	41,7	28,5	41,7	28,5	41,1

6		7		8		9		10	
h, м	*d, см*	*h, м*	*d, см*	*h, м*	*d, см*	*h, м*	*d, см*	*h, м*	*d, см*
17,0	15,2	17,5	14,3	17,0	14,5	17,5	16,5	17,0	15,5
17,5	17,6	17,5	15,6	17,5	16,6	17,0	17,2	17,5	17,1
17,5	18,3	18,0	18,2	18,0	18,1	18,0	18,3	18,5	18,4
18,0	20,1	18,5	19,1	18,5	19,5	18,5	19,4	18,0	19,5
18,0	20,7	18,0	20,6	18,5	20,1	18,5	20,5	18,0	20,9
19,0	20,4	19,0	20,7	19,5	21,7	19,0	21,7	19,0	21,2
20,5	21,7	19,5	21,8	19,5	22,2	19,5	22,2	20,0	21,8
19,5	22,1	20,0	22,4	20,0	22,6	20,0	22,6	19,5	22,3
20,0	22,6	20,0	23,6	20,0	23,8	20,0	23,8	19,5	24,7
20,5	24,1	20,5	25,8	20,5	26,4	20,5	26,4	20,0	25,6
20,5	24,8	20,5	26,5	20,5	26,8	20,5	26,8	20,5	26,8
21,0	25,3	21,0	26,0	21,0	26,2	21,0	28,2	21,0	25,6
22,0	26,5	22,0	27,6	22,0	27,1	22,0	29,5	21,5	27,5
21,5	26,9	21,5	26,7	21,5	28,0	21,5	30,0	21,5	28,3
22,5	29,2	22,0	31,0	22,0	28,5	22,0	32,4	22,0	30,1
22,0	30,4	22,5	32,4	22,5	30,1	22,5	33,1	22,5	30,8
22,0	31,9	23,0	28,5	23,0	29,7	23,0	29,7	23,0	29,5
23,5	28,0	23,0	29,3	23,0	30,3	23,0	30,3	23,5	29,9
24,0	33,3	23,5	30,3	23,5	30,9	24,0	30,9	23,5	31,3
23,5	33,6	24,0	31,6	24,0	31,5	23,5	31,0	24,0	32,5
24,0	34,1	24,5	32,8	24,5	32,1	24,5	32,0	24,0	33,1
24,0	35,0	24,0	33,5	24,0	33,7	24,0	34,7	24,5	33,8
24,5	36,5	24,0	34,2	24,0	34,4	24,5	35,4	24,0	34,4
24,0	37,2	24,5	37,5	24,5	37,3	24,5	36,3	24,5	34,7
25,0	30,4	25,0	29,7	25,0	29,8	25,0	32,8	25,0	32,8
25,5	31,7	25,5	30,1	25,5	30,4	26,0	33,4	25,5	33,2
25,5	33,9	26,0	31,0	26,0	31,4	26,5	33,6	26,0	33,7
26,0	35,3	26,5	33,5	26,5	33,2	26,5	35,2	26,0	35,2
26,5	37,4	26,0	34,5	26,0	34,7	26,5	35,7	26,5	36,7
27,0	32,3	27,5	33,6	27,5	33,9	27,0	33,9	27,5	33,9
28,0	38,2	27,5	38,0	27,5	39,0	27,5	38,7	28,0	37,5
28,5	41,2	28,0	41,7	28,0	41,1	27,5	40,3	28,5	40,1

11		12		13		14		15	
h, м	*d, см*	*h, м*	*d, см*	*h, м*	*d, см*	*h, м*	*d, см*	*h, м*	*d, см*
18,0	16,2	17,5	17,4	17,0	16,5	17,0	17,0	17,5	17,2
17,5	17,9	18,0	18,3	18,0	17,6	17,5	17,8	18,0	17,6
17,0	18,2	17,5	18,8	17,5	18,2	18,0	18,4	18,5	18,7
18,5	18,5	18,0	19,5	18,0	20,3	18,5	20,4	19,0	20,2
19,0	19,8	18,0	20,2	18,5	21,5	19,5	21,1	19,5	21,4
20,0	20,6	18,5	21,3	19,5	21,5	20,0	22,3	20,5	22,6
20,0	20,7	20,5	22,4	19,0	23,2	19,5	24,0	20,5	26,5
20,0	21,6	19,5	22,8	19,5	24,8	20,5	26,1	21,0	25,1
20,5	22,3	20,0	23,9	20,5	26,2	21,0	24,9	21,5	26,0
20,5	23,4	21,0	24,3	20,5	26,6	22,0	26,1	21,5	26,4
20,5	24,7	21,5	25,6	21,0	25,5	21,0	26,7	22,0	26,9
21,0	25,4	22,5	26,1	21,0	27,3	22,0	30,1	22,5	30,2
21,0	26,3	22,0	27,0	21,5	28,0	22,5	30,5	22,0	30,7
21,5	28,2	22,5	27,5	21,5	28,3	23,0	27,0	23,5	27,4
21,5	28,7	22,0	27,8	22,5	29,0	23,5	27,8	23,0	27,9
22,0	30,2	22,5	30,1	22,0	28,8	23,5	28,1	24,0	28,0
22,5	31,0	23,0	28,5	22,5	30,2	24,0	28,7	23,0	28,5
23,0	28,9	24,5	29,6	22,0	31,0	24,0	29,4	23,5	29,6
23,0	29,7	24,0	30,4	23,0	29,6	24,0	30,3	23,5	30,1
23,5	30,3	24,5	31,5	23,0	29,8	23,5	31,2	24,0	30,9
23,5	31,8	24,0	32,0	23,5	30,2	24,5	33,5	24,0	31,5
23,5	32,4	24,5	34,6	23,5	30,9	24,0	31,6	23,5	30,4
24,0	33,3	24,5	35,1	24,0	32,1	24,5	34,1	24,5	33,1
24,5	34,2	25,5	33,2	24,5	34,3	25,0	29,6	24,5	33,7
25,0	33,9	25,0	33,7	24,5	34,7	25,5	33,4	24,5	34,3
25,5	34,8	26,0	35,8	25,0	33,3	26,0	34,3	25,0	33,5
26,0	35,5	26,5	36,7	25,5	33,8	25,5	35,4	26,0	35,8
26,5	36,1	26,0	37,8	26,0	35,9	26,5	36,8	26,0	36,0
27,0	36,4	27,0	33,9	26,5	36,5	27,0	32,7	27,0	36,7
27,5	36,9	28,5	36,9	27,0	36,9	27,5	36,9	28,0	37,9
27,5	37,6	28,0	37,7	27,0	37,5	28,0	37,6	28,5	38,6
28,0	38,4	28,0	38,7	28,0	38,4	28,0	40,5	28,5	39,4

16		17		18		19		20	
h, м	*d, см*	*h, м*	*d, см*	*h, м*	*d, см*	*h, м*	*d, см*	*h, м*	*d, см*
18,5	17,9	18,5	17,9	18,5	17,9	18,5	17,8	17,0	15,8
17,5	16,2	18,0	16,2	17,0	16,3	17,0	16,4	17,5	17,4
18,0	21,6	17,5	14,1	17,5	18,6	17,5	18,3	17,5	18,1
18,0	20,3	18,0	20,3	18,0	20,3	18,0	20,6	18,5	20,3
18,5	18,3	18,0	18,2	18,0	18,4	19,0	18,6	18,0	20,6
18,0	20,0	18,5	20,0	18,5	20,2	19,5	20,9	19,0	20,5
18,0	19,2	20,5	21,2	20,5	21,4	20,5	21,2	20,5	21,4
20,0	18,0	19,5	18,6	19,5	18,8	19,5	22,0	19,5	21,9
19,5	21,9	20,0	21,9	20,0	21,9	20,0	22,6	20,5	22,2
20,0	20,3	20,0	20,2	20,0	22,2	20,0	23,0	20,0	23,4
20,5	19,5	20,5	22,1	20,5	23,5	20,5	24,4	20,5	24,6
19,5	19,1	19,5	23,0	19,5	23,0	21,0	25,7	20,0	25,5
20,0	18,6	22,0	25,8	21,5	25,3	21,5	26,0	21,0	25,8
20,5	21,5	20,5	23,7	22,0	28,7	22,0	26,8	21,5	26,6
22,0	22,0	22,0	27,2	22,0	27,2	22,0	28,7	22,5	28,2
21,0	25,6	22,0	28,4	22,0	28,4	22,5	30,3	22,0	29,4
22,0	23,0	21,0	30,2	21,0	30,2	22,5	32,1	22,5	31,4
22,5	31,5	22,0	31,0	22,0	31,0	23,0	25,9	22,0	32,7
22,0	32,3	24,5	27,5	24,5	27,5	23,5	28,4	23,5	28,0
24,0	28,0	24,0	28,0	24,0	32,0	23,5	28,0	23,5	29,6
23,0	29,6	23,0	30,3	23,0	30,3	24,0	30,2	24,0	30,0
24,0	33,7	24,0	33,7	24,0	33,6	24,0	31,8	24,0	31,4
23,5	30,8	23,5	34,7	23,5	36,1	24,5	32,6	24,5	33,6
24,0	33,2	25,5	33,1	25,5	33,2	24,5	34,2	24,5	34,5
26,0	34,0	26,0	33,8	26,0	32,4	25,0	29,8	24,0	36,0
25,0	34,5	25,0	34,5	25,0	34,3	25,5	31,5	25,0	29,7
26,0	35,7	26,0	35,7	26,0	35,7	26,0	32,7	26,0	32,3
26,5	36,8	26,5	36,8	26,5	37,8	26,5	35,8	26,0	31,0
28,0	28,4	28,0	33,5	28,0	32,5	26,5	36,5	26,5	32,4
27,0	29,3	27,0	36,9	27,0	32,9	27,0	35,9	26,5	35,3
28,5	38,9	28,0	39,8	28,0	37,7	28,0	39,3	27,0	36,2
28,0	40,3	28,5	40,4	28,5	40,6	28,5	40,2	28,5	40,3

21		22		23		24		25	
h, м	*d, см*	*h, м*	*d, см*	*h, м*	*d, см*	*h, м*	*d, см*	*h, м*	*d, см*
18,0	16,2	17,5	17,4	17,0	16,5	17,0	17,0	17,5	17,2
17,5	17,9	18,0	18,3	18,0	17,6	17,5	17,8	18,0	17,6
17,0	18,2	17,5	18,8	17,5	18,2	18,0	18,4	18,5	18,7
18,5	18,5	18,0	19,5	18,0	20,3	18,5	20,4	19,0	20,2
19,0	19,8	18,0	20,2	18,5	21,5	19,5	21,1	19,5	21,4
20,0	20,6	18,5	21,3	19,5	21,5	20,0	22,3	20,5	22,6
20,0	20,7	20,5	22,4	19,0	23,2	19,5	24,0	20,5	26,5
20,0	21,6	19,5	22,8	19,5	24,8	20,5	26,1	21,0	25,1
20,5	22,3	20,0	23,9	20,5	26,2	21,0	24,9	21,5	26,0
20,5	23,5	21,0	24,2	20,5	26,7	22,0	26,2	21,5	26,5
20,5	24,7	21,5	25,6	21,0	25,5	21,0	26,7	22,0	26,9
21,0	25,4	22,5	26,1	21,0	27,3	22,0	30,1	22,5	30,2
21,0	26,3	22,0	27,0	21,5	28,0	22,5	30,5	22,0	30,7
21,5	28,2	22,5	27,5	21,5	28,3	23,0	27,0	23,5	27,4
21,5	28,7	22,0	27,8	22,5	29,0	23,5	27,8	23,0	27,9
22,0	30,2	22,5	30,1	22,0	28,8	23,5	28,1	24,0	28,0
22,5	31,0	23,0	28,5	22,5	30,2	24,0	28,7	23,0	28,5
23,0	28,9	24,5	29,6	22,0	31,0	24,0	29,4	23,5	29,6
23,0	29,7	24,0	30,4	23,0	29,6	24,0	30,3	23,5	30,1
23,5	30,3	24,5	31,5	23,0	29,8	23,5	30,2	24,0	30,9
23,5	31,8	24,0	32,0	23,5	30,2	24,5	32,5	24,0	32,5
23,5	32,4	24,5	34,6	23,5	31,6	24,0	31,6	23,5	30,4
24,0	33,3	24,5	35,1	24,0	32,8	24,5	34,1	24,5	33,1
24,5	34,2	25,5	33,2	24,5	34,3	25,0	29,6	24,5	33,7
25,0	33,9	25,0	33,7	24,5	34,7	25,5	33,4	24,5	34,3
25,5	34,8	26,0	35,8	25,0	33,3	26,0	34,3	25,0	33,5
26,0	35,5	26,5	36,7	25,5	33,8	25,5	35,4	26,0	35,8
26,5	36,1	26,0	37,8	26,0	35,9	26,5	36,8	26,0	36,0
27,0	36,4	27,0	33,9	26,5	36,5	27,0	32,7	27,0	36,7
27,5	36,9	28,5	36,9	27,0	36,9	27,5	36,9	28,0	37,9
27,5	37,6	28,5	37,7	27,5	37,5	28,5	37,6	28,0	38,6
28,0	38,4	28,0	38,7	28,0	38,4	28,0	40,5	28,5	39,4

26		27		28		29		30	
h, м	*d, см*	*h, м*	*d, см*	*h, м*	*d, см*	*h, м*	*d, см*	*h, м*	*d, см*
17,0	15,2	17,5	14,3	17,0	14,5	17,5	16,5	17,0	15,5
17,5	17,6	17,5	15,6	17,5	16,6	17,0	17,2	17,5	17,1
17,5	18,3	18,0	18,2	18,0	18,1	18,0	18,3	18,5	18,4
18,0	20,1	18,5	19,1	18,5	19,5	18,5	19,4	18,0	19,5
18,0	20,7	18,0	20,6	18,5	20,1	18,5	20,5	18,0	20,9
19,0	20,4	19,0	20,7	19,5	21,7	19,0	21,7	19,0	21,2
20,5	21,7	19,5	21,8	19,5	22,2	19,5	22,2	20,0	21,8
19,5	22,1	20,0	22,4	20,0	23,6	20,0	23,2	19,5	22,3
20,0	22,6	20,0	24,3	20,0	24,8	20,0	23,8	19,5	24,1
20,5	23,5	20,5	25,8	20,5	26,4	20,5	26,4	20,0	25,4
20,5	24,8	20,5	26,5	20,5	26,8	20,5	26,8	20,5	26,8
21,0	25,3	21,0	26,0	21,0	26,2	21,0	28,2	21,0	25,6
22,0	26,5	22,0	27,6	22,0	27,1	22,0	29,5	21,5	27,5
21,5	26,9	21,5	26,7	21,5	28,0	21,5	30,0	21,5	28,3
22,5	29,2	22,0	31,0	22,0	28,5	22,0	32,4	22,0	30,1
22,0	30,4	22,5	32,4	22,5	30,1	22,5	33,1	22,5	30,8
22,0	31,9	23,0	28,5	23,0	29,7	23,0	29,7	23,0	29,5
23,5	28,0	23,0	29,3	23,0	30,3	23,0	30,3	23,5	29,9
24,0	33,3	23,5	30,3	23,5	30,9	24,0	30,9	23,5	31,3
23,5	33,6	24,0	31,6	24,0	31,5	23,5	31,5	24,0	32,5
24,0	34,1	24,5	32,1	24,5	32,0	24,5	32,0	24,0	33,1
24,0	35,0	24,0	33,5	24,0	33,7	24,0	34,7	24,5	33,8
24,5	36,5	24,0	34,2	24,0	34,4	24,5	35,4	24,0	34,4
24,0	37,2	24,5	37,5	24,5	37,3	24,5	36,3	24,5	34,7
25,0	30,4	25,0	29,7	25,0	29,8	25,0	32,8	25,0	32,8
25,5	31,7	25,5	30,1	25,5	30,4	26,0	33,4	25,5	33,2
25,5	33,9	26,0	31,0	26,0	31,4	26,5	33,6	26,0	33,7
26,0	35,3	26,5	33,5	26,5	33,2	26,5	35,2	26,0	35,2
26,5	37,4	26,0	34,5	26,0	34,7	26,5	35,7	26,5	36,7
27,0	32,3	27,5	33,6	27,5	33,9	27,0	33,9	27,5	33,9
28,0	38,2	27,5	38,0	27,5	39,0	27,5	38,7	28,0	37,5
28,5	40,5	28,0	41,1	28,0	41,2	27,5	40,3	28,5	40,1

Приложение 2

Стандартные значения критерия Стьюдента

Число степеней свободы	Критерий Стьюдента t_{st} при вероятности безошибочного заключения p		
	0,95	**0,99**	**0,999**
1	12,7	63,7	637,0
2	4,3	9,9	31,6
3	3,2	5,8	12,9
4	2,8	4,6	8,6
5	2,6	4,0	6,9
6	2,4	3,7	6,0
7	2,4	3,5	5,3
8	2,3	3,4	5,0
9	2,3	3,3	4,8
10	2,2	3,2	4,6
11	2,2	3,1	4,4
12	2,2	3,1	4,3
13	2,2	3,0	4,2
14 – 15	2,1	3,0	4,1
16 – 17	2,1	2,9	4,0
18 – 20	2,1	2,9	3,9
21 – 24	2,1	2,8	3,8
25 – 28	2,1	2,8	3,7
29 – 30	2,0	2,8	3,7
31 – 34	2,0	2,7	3,7
35 – 42	2,0	2,7	3,6
43 – 62	2,0	2,7	3,5
63 – 175	2,0	2,6	3,4
≥ 176	2,0	2,6	3,3

Приложение 3

Значения верхних пределов критерия Фишера F
(верхняя строка 5%-ный уровень значимости, нижняя – 1-ный уровень значимости)

df_2	Степень свободы для большей дисперсии (df_1)																
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**16**	**24**	**30**	**50**	**∞**
1	161	200	216	225	230	234	237	239	241	242	243	244	246	249	250	252	254
	4052	**4999**	**5403**	**5625**	**5764**	**5889**	**5928**	**5981**	**6022**	**6056**	**6082**	**6106**	**6169**	**6234**	**6258**	**6302**	**6366**
2	18,5	19,0	19,2	19,3	19,3	19,3	19,4	19,4	19,4	19,4	19,4	19,4	19,4	19,4	19,5	19,5	19,5
	98,5	**99,0**	**99,2**	**99,3**	**99,3**	**99,3**	**99,3**	**99,4**	**99,4**	**99,4**	**99,4**	**99,4**	**99,4**	**99,4**	**99,5**	**99,5**	**99,5**
3	10,1	9,6	9,3	9,1	9,0	8,9	8,9	8,8	8,8	8,8	8,8	8,7	8,7	8,6	8,6	8,6	8,5
	34,1	**30,8**	**29,5**	**28,7**	**28,2**	**27,9**	**27,7**	**27,5**	**27,3**	**27,2**	**27,1**	**27,1**	**26,8**	**26,6**	**26,5**	**26,4**	**26,1**
4	7,7	6,9	6,6	6,4	6,3	6,2	6,1	6,0	6,0	6,0	5,9	5,9	5,8	5,8	5,7	5,7	5,7
	21,2	**18,0**	**16,7**	**16,0**	**15,5**	**15,2**	**15,0**	**14,8**	**14,7**	**14,6**	**14,5**	**14,4**	**14,2**	**13,9**	**13,8**	**13,7**	**13,5**
5	6,6	5,8	5,4	5,2	5,1	5,0	4,9	4,8	4,8	4,8	4,7	4,7	4,6	4,5	4,5	4,4	4,4
	16,3	**13,3**	**12,1**	**11,4**	**11,0**	**10,7**	**10,5**	**10,3**	**10,2**	**10,1**	**10,0**	**9,9**	**9,7**	**9,5**	**9,4**	**9,2**	**9,0**
6	6,0	5,1	4,8	4,5	4,4	4,3	4,2	4,2	4,1	4,0	4,0	4,0	3,9	3,8	3,8	3,8	3,7
	13,7	**10,9**	**9,8**	**9,2**	**8,8**	**8,5**	**8,3**	**8,1**	**8,0**	**7,9**	**7,8**	**7,7**	**7,5**	**7,3**	**7,2**	**7,1**	**6,9**
7	5,6	4,7	4,4	4,1	4,0	3,9	3,8	3,7	3,7	3,6	3,6	3,6	3,5	3,4	3,4	3,3	3,2
	12,3	**9,6**	**8,5**	**7,9**	**7,5**	**7,2**	**7,0**	**6,8**	**6,7**	**6,6**	**6,5**	**6,5**	**6,3**	**6,1**	**6,0**	**5,9**	**5,6**
8	5,3	4,5	4,1	3,8	3,7	3,6	3,5	3,4	3,4	3,3	3,3	3,3	3,2	3,1	3,1	3,0	2,9
	11,3	**8,7**	**7,6**	**7,0**	**6,6**	**6,4**	**6,2**	**6,0**	**5,9**	**5,8**	**5,7**	**5,7**	**5,5**	**5,3**	**5,2**	**5,1**	**4,9**
9	5,1	4,3	3,9	3,6	3,5	3,4	3,3	3,3	3,2	3,2	3,1	3,1	3,0	2,9	2,9	2,8	2,7
	10,6	**8,0**	**7,0**	**6,4**	**6,1**	**5,8**	**5,6**	**5,5**	**5,4**	**5,3**	**5,2**	**5,1**	**5,0**	**4,7**	**4,6**	**4,5**	**4,3**
10	5,0	4,1	3,7	3,5	3,3	3,2	3,1	3,1	3,0	3,0	3,0	2,9	2,8	2,7	2,7	2,6	2,5
	10,0	**7,6**	**6,6**	**6,0**	**5,6**	**5,4**	**5,2**	**5,1**	**5,0**	**4,9**	**4,8**	**4,7**	**4,5**	**4,3**	**4,3**	**4,1**	**3,9**
12	4,8	3,9	3,5	3,3	3,1	3,0	2,9	2,9	2,8	2,8	2,7	2,7	2,6	2,5	2,5	2,4	2,3
	9,3	**6,9**	**6,0**	**5,4**	**5,1**	**4,8**	**4,7**	**4,5**	**4,4**	**4,3**	**4,2**	**4,2**	**4,0**	**3,8**	**3,7**	**3,6**	**3,4**
16	4,5	3,6	3,2	3,0	2,9	2,7	2,7	2,6	2,5	2,5	2,5	2,4	2,3	2,2	2,2	2,1	2,0
	8,5	**6,2**	**5,3**	**4,8**	**4,4**	**4,2**	**4,0**	**3,9**	**3,8**	**3,7**	**3,6**	**3,6**	**3,4**	**3,2**	**3,1**	**3,0**	**2,8**

Приложение 4

Значения верхнего предела χ^2

Число степеней свободы	Критерий χ^2 при уровнях значимости							
	0,99	0,98	0,95	0,90	0,80	0,70	0,50	0,30
1	1,6	2,7	3,8	5,4	6,6	7.9	9,5	10,8
2	3,2	4,6	6,0	7,8	9,2	11,6	12,3	13,8
3	4,6	6,3	7,8	9,8	11,3	12,8	14,8	16,3
4	6,0	7,8	9,5	11,7	13,3	13,9	16,9	18,5
5	7,3	9,2	11,1	13,4	15,1	16,3	18,9	20,5
6	8,6	10,6	12,6	15,0	16,8	18,6	20,7	22,5
7	9,8	12,0	14,1	16,6	18,5	20,3	22,6	24,3
8	11,0	13,4	15,5	18,2	20,1	21,9	24,3	26,1
9	12,2	14,7	16,9	19,7	21,7	23,6	26,1	27,9
10	13,4	16,0	18,3	21,2	23,2	25,2	27,7	29,6
11	14,6	17,3	19,7	22,6	24,7	26,8	29,4	31,3
12	15,8	18,5	21,0	24,1	26,2	28,3	31,0	32,9
13	17,0	19,8	22,4	25,5	27,7	29,8	32,5	34,5
14	18,2	21,1	23,7	26,9	29,1	31,0	34,0	36,1
15	19,3	22,3	25,0	28,3	30,6	32,5	35,5	37,7
16	20,5	23,5	26,3	29,6	32,0	34,0	37,0	39,2
17	21,6	24,8	27,6	31,0	33,4	35,5	38,5	40,8
18	22,8	26,0	28,9	32,3	34,8	37,0	40,0	42,3
19	23,9	27,2	30,1	33,7	36,2	38,5	41,5	43,8
20	25,0	28,4	31,4	35,0	37,6	40,0	43,0	45,3
21	26,2	29,6	32,7	36,3	38,9	41,5	44,5	46,8
22	27,3	30,8	33,9	37,7	40,3	42,5	46,0	48,3
23	28,4	32,0	35,2	39,0	41,6	44,0	47,5	49,7
24	29,6	33,2	26,4	40,3	43,0	45,5	48,5	51,2
25	30,7	34,4	37,7	41,6	44,3	47,0	50,0	52,6
26	31,8	35,6	38,9	42,9	45,6	48,0	51,5	54,1
27	32,9	36,7	40,1	44,1	47,0	49,5	53,0	55,5
28	34,0	37,9	41,3	45,4	48,3	51,0	54,5	56,9
29	35,1	39,1	42,6	46,7	49,9	52,5	56,0	58,3
30	36,3	40,3	43,8	48,0	50,9	54,0	57,5	59,7

Содержание