

**ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ ОБРАЗОВАТЕЛЬНОЕ  
УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ  
«ОРЕНБУРГСКИЙ ГОСУДАРСТВЕННЫЙ АГРАРНЫЙ УНИВЕРСИТЕТ»**

**МЕТОДИЧЕСКИЕ МАТЕРИАЛЫ ДЛЯ ОБУЧАЮЩИХСЯ ПО ОСВО-  
ЕНИЮ ДИСЦИПЛИНЫ**

**Б1.О.04  
МАТЕМАТИЧЕСКОЕ МОДЕЛИРОВАНИЕ БИОЛОГИЧЕСКИХ ПРОЦЕССОВ.  
КОМПЬЮТЕРНЫЕ ТЕХНОЛОГИИ В БИОЛОГИИ**

**Направление подготовки (специальность) 06.04.01 Биология**

**Профиль образовательной программы Микробиология**

**Форма обучения очно-заочная**

## СОДЕРЖАНИЕ

1. Тематическое содержание дисциплины .....	3
2. Методические рекомендации по выполнению курсовой работы (проекта) .....	85
3. Методические рекомендации по выполнению индивидуальных домашних заданий (контрольных работ) .....	85

## **1. Тематическое содержание дисциплины**

### **1.1. Тема 1: «Математическое моделирование биологических процессов.**

#### **1.1.1. Перечень и краткое содержание рассматриваемых вопросов:**

##### **1. Наименование вопроса № 1.**

Понятие о математическом моделировании. Компьютерные и математические модели. Современная классификация моделей биологических процессов.

Краткое содержание вопроса № 1. Понятие о математическом моделировании. Компьютерные и математические модели. Современная классификация моделей биологических процессов.

Задание для работы: ознакомиться с понятием о математическом моделировании, компьютерных и математических моделей, современной классификацией моделей биологических процессов.

##### **Понятие о математическом моделировании**

**Модель** - материальный объект, система математических зависимостей или программа, имитирующая структуру или функционирование исследуемого объекта.

**Моделирование** - представление различных характеристик поведения физической или абстрактной системы с помощью другой системы.

**Математическое моделирование** - метод исследования процессов и явлений на их математических моделях.

Изучение компьютерного математического моделирования открывает широкие возможности для осознания связи информатики с математикой и другими науками - естественными и социальными. Компьютерное математическое моделирование в разных своих проявлениях использует практически весь аппарат современной математики.

Математическое моделирование не всегда требует компьютерной поддержки. Каждый специалист, профессионально занимающийся математическим моделированием, делает все возможное для аналитического исследования модели. Аналитические решения (т. е. представленные формулами, выражающими результаты исследования через исходные данные) обычно удобнее и информативнее численных. Возможности аналитических методов решения сложных математических задач, однако, очень ограничены и, как правило, эти методы гораздо сложнее численных.

В компьютерном моделировании доминируют численные методы, реализуемые на компьютерах. Однако понятия "аналитическое решение" и "компьютерное решение" отнюдь не противостоят друг другу, так как: а) все чаще компьютеры при математическом моделировании используются не только для численных расчетов, но и для аналитических преобразований: б) результат аналитического исследования математической модели часто выражен столь сложной формулой, что при взгляде на нее не складывается восприятия описываемого ей процесса. Этую формулу нужно представить графически, проиллюстрировать в динамике, иногда даже озвучить, т. е. проделать то, что называется "визуализацией абстракций". При этом компьютер - незаменимое техническое средство.

**Классификация математических моделей.** К классификации математических моделей можно подходить по-разному, положив в основу классификации различные прин-

ципы.

1) Классификация моделей по отраслям наук (математические модели в физике, биологии, социологии и т. д.);

2) Классификация моделей по применяемому математическому аппарату (модели, основанные на применении обыкновенных дифференциальных уравнений, дифференциальных уравнений в частных производных, стохастических методов, дискретных алгебраических преобразований и т. д.);

3) Классификация моделей с точки зрения целей моделирования.

- **дескриптивные (описательные) модели;**
- **оптимизационные модели;**
- **многокритериальные модели;**
- **игровые модели;**
- **имитационные модели.**

**Имитационная модель** - описание системы и ее поведения, которое может быть реализовано и исследовано в ходе операций на компьютере.

**Имитационное моделирование** - исследование поведения сложной системы на ее модели. Можно сказать, что чаще всего имитационное моделирование применяется для того, чтобы описать свойства большой системы при условии, что поведение составляющих ее объектов очень просто и четко сформулировано. Математическое описание тогда сводится к уровню статистической обработки результатов моделирования при нахождении макроскопических характеристик системы. Такой компьютерный эксперимент фактически претендует на воспроизведение натурного эксперимента. Имитационное моделирование позволяет осуществить проверку гипотез, исследовать влияние различных факторов и параметров.

**Математической моделью биологического процесса** называется система соотношений, уравнений, неравенств, геометрических объектов и т. д., которые в математической форме отображают, воспроизводят или имитируют наиболее важные особенности и свойства реального биологического процесса с целью его анализа, прогнозирования, получения научных и практических результатов.

Замена реального процесса математической моделью даёт возможность безболезненно, достаточно быстро и с минимальными затратами исследовать все мыслимые сценарии развития биологического процесса, спрогнозировать результаты и последствия.

Особенности математической модели во многом определяются типом моделируемого процесса. Но любая математическая модель в процессе создания проходит ряд схожих этапов, проходит проверку на адекватность и затем используется для анализа реальных биологических процессов.

При моделировании, выборе и формулировке модели, определяющими обстоятельствами являются объект, цель и метод (средства) моделирования. В данной науке объектами моделирования являются биологические процессы разного уровня организации. Методами моделирования служат методы динамической теории систем. Средства – дифференциальные и разностные уравнения, методы качественной теории дифференциальных уравнений, компьютерная симуляция. Цели моделирования: выяснение механизмов взаимодействия элементов системы, идентификация и верификация параметров модели по экспериментальным данным, оценка устойчивости системы (модели), прогноз поведения системы при различных внешних воздействиях,

различных способах управления и прочая, оптимальное управление системой в соответствии с выбранным критерием оптимальности.

Условно все математические модели биологических систем можно разделить на регрессионные, качественные и имитационные. *Регрессионные зависимости* – это формулы, описывающие связь различных характеристик системы, не претендую на физический или биологический смысл этих зависимостей. Для построения регрессионной модели достаточно статистически достоверных наблюденных корреляций между переменными или параметрами системы.

В результате рассмотрения вопроса № 1 студенты должны:

- освоить понятие о математическом моделировании, компьютерных и математических моделях, современной классификации моделей биологических процессов, динамических дифференциальных моделях биологических процессов.

## 2. Наименование вопроса № 2. Динамические модели биологических процессов.

Краткое содержание вопроса № 2. Динамические модели биологических процессов.

Задание для работы:

2.1 Рассмотреть динамические модели биологических процессов: модель динамики численности изолированной популяции, логистическую модель А. Лотки.

2.2, а). Сформулировать задачу *Коии математической модели естественного роста изолированной популяции*

$$\frac{dx}{dt} = (a - b)x, x_0 = x(0)$$

при конкретных значениях  $a, b, x_0$ .

б) Решив задачу методом разделения переменных, найти экспоненциальную динамику численности микроорганизмов *модели естественного роста изолированной популяции*

$$x(t) = x_0 e^{(a-b)t}, x_0 = x(0).$$

в) Иллюстрировать графически найденную динамику и выяснить тенденцию (размножение или вымирание).

2.3 Иметь представление о моделях биохимической регуляции белкового синтеза, моделях внутриклеточных колебаний кальция, моделях Вольтерра.

Предлагаются для ознакомления следующие динамические модели биологических процессов:

- 1) Модели динамики численности изолированной популяции,
- 2) Логистическая модель А. Лотки.
- 3) Модель биохимической регуляции белкового синтеза.
- 4) Модель внутриклеточных колебаний кальция.
- 5) Модели Вольтерра.

Рассмотрим динамические модели биологических процессов на примере модели динамики численности изолированной популяции, логистической модели А. Лотки.

Пусть  $t$  – время,  $x(t)$  – количество микроорганизмов в популяции к моменту вре-

мени  $t$ ,  $R(x)$  – количество родившихся и  $S(x)$  – количество умерших особей популяции микроорганизмов. Дифференциальное уравнение динамики численности популяции имеет вид

$$\frac{dx}{dt} = R(x) - S(x). \quad (1)$$

В простейшем случае, когда рождаемость и смертность пропорциональны численности, т. е.

$$R(x) = a \Psi_k, S(x) = b \Psi_k$$

уравнение (1) принимает вид

$$\frac{dx}{dt} = (a - b) \Psi_k. \quad (2)$$

так называемого уравнения естественного роста замкнутой популяции.

*Математическая модель естественного роста замкнутой популяции.* Найти решение уравнения (2), удовлетворяющее заданному начальному условию

$$x_0 = x(0). \quad (3)$$

Такая математическая задача в теории дифференциальных уравнений называется *задачей Коши*.

Решая задачу (2), (3) методом разделения переменных найдём экспоненциальную динамику численности микроорганизмов *модели естественного роста замкнутой популяции*

$$x(t) = x_0 \Psi^{(a-b)t}, x_0 = x(0). \quad (4)$$

График функции (2.8) при положительных (размножение) и отрицательных (вымирание) значениях константы  $r = a - b$  скорости роста представлен на рисунке 1.

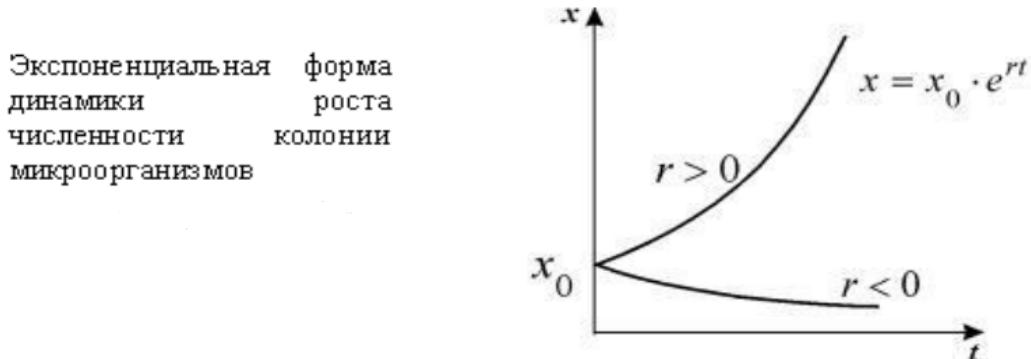


Рис. 1

Эта модель сыграла важную роль в развитии математической биологии и экологии.

Альфред Лотка (Alfred Lotka), 1880–1949, один из создателей математической биологии) пришел к уравнению динамики замкнутой популяции, которое стало известно как *логистическое уравнение*, которое оно наиболее часто встречается в литературе в виде

$$\frac{dx}{dt} = (a - b) \Psi_k - \frac{x}{K}$$

Здесь  $a - b$  и  $K$  — положительные параметры. Когда  $x(t)$  мало, то  $x(t)$  растет экспоненциально. Параметр  $K$  интерпретируется как *потенциальная емкость экологической системы*, которая определяется доступным наличным количеством ресурсов. Величина  $K$  определяет предельное значение численности популяции.

В результате рассмотрения вопроса № 2 студенты должны:

- освоить понятие о динамических дифференциальных моделях биологических процессов: модели динамики численности изолированной популяции, логистической модели А. Лотки, иметь представление о моделях биохимической регуляции белкового синтеза, модели внутриклеточных колебаний кальция, моделях Вольтерра.

- приобрести умения и навыки создания и аналитического исследования динамических моделей биологических процессов: модели динамики численности изолированной популяции и логистической модели А. Лотки.

### 3. Наименование вопроса № 3.

Вероятностные и статистические модели в биологии.

3.1 Наименование вопроса № 3.1. Вероятностные и статистические модели в биологии: оценивание характеристик генеральной совокупности по выборке – методы группировки, таблицы частот и гистограммы.

3.2. Наименование вопроса № 3.2. Вероятностный и статистический методы изучения моделей в биологии: оценивание характеристик генеральной совокупности по выборке – числовые характеристики статистического распределения.

3.3. Наименование вопроса № 3.3. Вероятностный и статистический методы изучения моделей в биологии: проверка гипотез о параметрах и виде распределения, доверительные интервалы.

3.1 Краткое содержание вопроса № 3.1. Вероятностные и статистические модели в биологии: оценивание характеристик генеральной совокупности по выборке – методы группировки, таблицы частот и гистограммы.

Задание для работы:

1. Краткие теоретические сведения.

2. Основное задание. По данной выборке объема  $n = 100$  (Таблица 1) следует 1) Составить вариационный ряд (эмпирический ряд распределения) **сгруппированный** (Таблица 2).

2) Составить интервальный эмпирический ряд распределения (Таблица 3).

3) Объединив крайние интервалы с малыми значениями частот составить интервальный вариационный ряд (эмпирический ряд распределения) - Таблица 4.

4) Заменив каждый интервал его серединой составить точечный ряд распределения (Таблица 5).

- 5) Найти эмпирическую функцию распределения  $F^*(x)$  по точечному ряду распределения в Таблице 5, нарисовать график  $F^*(x)$  (Рис. 1)
- 6) Выполнить графическое представление статистических данных: построить полигон относительных частот по точечному ряду распределения фрагмента Таблицы 5 (Рис. 2).
- 7) Составить таблицу плотности относительных частот (таблицу 6) с шагом  $h=2$  (длина интервала).
- 8) Составить гистограмму плотности относительных частот по таблице 6, взяв из неё первую, вторую и последнюю строки (Рис. 3).
- 9) Соединив соседние вершины прямоугольников гистограммы на Рис.3, получить линию эмпирической плотности (Рис. 4).

### **1. Генеральная и выборочная совокупности. Повторная и бесповторная выборки. Репрезентативная выборка. Способы отбора.**

Под **генеральной совокупностью** подразумевают всю совокупность изучаемых объектов, или все множество возможных значений изучаемой случайной величины (СВ)  $X$ . Количество элементов генеральной совокупности называют **объемом  $N$**  этой совокупности.

**Случайной выборкой** или просто **выборкой объема  $n$**  значений СВ  $X$  называется набор значений этой случайной величины

$$x_1, x_2, x_3, \dots, x_n,$$

полученных в ходе  **$n$**  соответствующих независимых наблюдений, измерений или экспериментов.

Иногда выборку также называют **статистическими данными**, или **статистикой реализаций (СВ)  $X$** , или просто **статистикой**.

Различают повторные и бесповторные выборки. Говорят, о собственно случайных, механических, серийных и типических выборках.

### **2. Вариационный ряд (эмпирический ряд распределения). Табличное представление статистических данных**

Пусть  $X$  – случайная величина, множество всевозможных значений которой образует генеральную совокупность,  $x_i$  – значения случайной величины  $X$ , полученные в ходе  **$n$**  соответствующих независимых наблюдений, измерений;  $n_i$  – выборочные (эмпирические) частоты наблюдаемых значений  $x_i$ ;  $w_i = \frac{n_i}{n}$  – выборочные (эмпирические) относительные частоты наблюдаемых значений  $x_i$ .

**Определение.** **Вариационный ряд** – ранжированный ряд наблюдаемых значений случайной величины  $X$  (вариантов)  $x_1, x_2, x_3, \dots, x_n$ , с указанием соответствующих весов, т.е. частот  $n_i$  и относительных частот  $\frac{n_i}{n}$ .

Дана выборка значений случайной величины  $X$

**Таблица 1**

3	-5	5	-6	4	-6	4	-6	-5	6
1	-4	2	-5	1	-5	1	-4	-4	2
0	-3	0	-4	0	-4	0	-4	-3	7
-1	-3	0	-3	-1	-3	-1	-3	-2	0
-2	-2	-1	-2	-1	-2	-1	-2	-2	-1
-2	-1	-2	-2	-2	-2	-2	-1	-1	-2
-3	0	-3	-1	-3	-1	-3	-1	0	-3
-4	1	-3	0	-4	0	-4	0	1	-3
-5	2	-4	1	-4	1	-5	2	2	-4
-6	1	-5	3	-6	3	-6	5	10	-5

а) Вариационный ряд (эмпирический ряд распределения) **простой**

$x_i$	-6	-6	-6	-5	-5	...	-2	-2	...	0	0	0	...	6	7	8	9	10
$n_i$	1	1	1	1	1	...	1	1	...	1	1	1	...	1	1	1	1	1

б) Вариационный ряд (эмпирический ряд распределения) **сгруппированный**

**Таблица 2**

$x_i$	-6	-5	-4	-3	-2	-1	0	1	2	3	4	5	6	7	8	9	10	$\Sigma$
$n_i$	6	8	12	13	14	13	11	8	5	3	2	2	1	1	0	0	1	<b>100</b>
$\frac{n_i}{n}$	0,06	0,08	0,12	0,13	0,14	0,13	0,11	0,08	0,05	0,03	0,02	0,02	0,01	0,01	0	0	0,01	<b>1</b>

При большом объёме выборки и большом числе различных по величине элементов выборки пользоваться простым и сгруппированным вариационными рядами неудобно. В таком случае пользуются интервальным вариационным рядом, который строится следующим образом.

Под группировкой данных понимают также разбиение вариационного ряда на интервалы.

Интервальный вариационный ряд  
(интервальный эмпирический ряд распределения)

**Таблица 3**

$[x_i, x_{i+1})$	$[-7; -5)$	$[-5; -3)$	$[-3; -1)$	$[-1; 1)$	$[1; 3)$	$[3; 5)$	$[5; 7)$	$[7; 9)$	$[9; 11]$
$n_i$	6	20	27	24	13	5	3	1	1
$\frac{n_i}{n}$	0,06	0,2	0,27	0,24	0,13	0,05	0,03	0,01	0,01
$\frac{n_i}{hn}$	0,03	0,1	0,135	0,12	0,065	0,025	0,015	0,005	0,005

Объединим крайние интервалы с малыми значениями частот. Получим

Интервальный вариационный ряд (эмпирический ряд распределения)

**Таблица 4**

$[x_i, x_{i+1})$	$[-7; -5)$	$[-5; -3)$	$[-3; -1)$	$[-1; 1)$	$[1; 3)$	$[3; 5)$	$[5; 7)$	$[7; 9)$
$n_i$	6	20	27	24	13	5	3	2
$\frac{n_i}{n}$	0,06	0,2	0,27	0,24	0,13	0,05	0,03	0,02
$\frac{n_i}{hn}$	0,03	0,1	0,135	0,12	0,065	0,025	0,015	0,01

В Таблице 4 величины  $n_i$  и  $w_i = \frac{n_i}{n}$  называются **интервальной частотой** и **интервальной относительной частотой**.

Заменяя каждый интервал его серединой и получаем новый ряд распределения – точечный

**Таблица 5**

$[x_i, x_{i+1})$	-6	-4	-2	0	2	4	6	8
$n_i$	6	20	27	24	13	5	3	2
$\frac{n_i}{n}$	0,06	0,2	0,27	0,24	0,13	0,05	0,03	0,02
$\frac{n_i}{hn}$	0,03	0,1	0,135	0,12	0,065	0,025	0,015	0,01

### 3. Эмпирическая функция распределения

Пусть  $X$  – случайная величина, множество всевозможных значений которой образует генеральную совокупность. Ни закон распределения этой случайной величины (ряд распределения или функция распределения, плотность распределения), ни её численные характеристики и параметры (математическое ожидание, дисперсия, мода, медиана и т. д.), вообще говоря, не известны. По выборке

$$\{x_1, x_2, x_3, \dots, x_n\}$$

надо составить представление о законе распределения СВ  $X$  и параметрах её распределения в генеральной совокупности (выдвинуть достоверную статистическую гипотезу).

Неизвестный истинный закон распределения СВ  $X$  в генеральной совокупности называется **теоретическим**.

Истинное значение того или иного параметра, численной характеристики распределения изучаемой СВ  $X$  (в генеральной совокупности) называют его **теоретическим, или генеральным значением**.

Закон распределения в выборке, который является приближением к неизвестному истинному закону распределения СВ  $X$  в генеральной совокупности, называется **выборочным или эмпирическим**.

Вместо неизвестных значений параметров, численных характеристик распределения СВ  $X$  приходится находить их приближения на данной конкретной выборке, называемые **статистическими оценками, выборочными значениями**, или точечными значениями, или **выборочной оценкой** (выборочное математическое ожидание, выборочная дисперсия). **Выборочные характеристики** иногда называют также **эмпирическими**. Например, функция распределения в выборке называется **эмпирической (выборочной) функцией распределения**. Если  $F(x)$  – истинная функция распределения СВ  $X$  генеральной совокупности, то через  $F^*(x)$  обозначается эмпирическая функция распределения выборки, через  $\varphi^*(x)$  – эмпирическая (выборочная) функция плотности вероятности.

Важно четко понимать разницу между теоретическим значением параметра и его выборочными оценками. Теоретическое значение того или иного параметра распределения СВ в общем случае определяется только его плотностью распределения, т. е. бесконечной информацией. Данные же выборки всегда конечны. Поэтому никогда нет возможности найти точное теоретическое значение параметра.

Очень важно также, что выборочное значение параметра само является случайной величиной, поскольку рассчитывается по данным случайной выборки.

Напомним, что закон распределения ДСВ  $X$  сопоставляет значениям СВ  $x_i$  их вероятности  $p_i$ . Потому можно трактовать эмпирический ряд распределения (Таблица 2) как эмпирический (выборочный) закон распределения, сопоставляющий значениям  $x_i$  их эмпирические вероятности

$$p_i^* = w_i = \frac{n_i}{n}$$

т. е. выборочные (эмпирические) относительные частоты наблюдаемых значений  $x_i$ .

**Эмпирической функцией распределения  $F^*(x)$  (функцией распределения выборки)** называется относительная частота появления события, состоящего в том, что случайная величина  $X$  примет значение, меньшее заданного  $x$ , т. е.

$$F^*(x) = w(X < x) = \frac{n_x}{n},$$

где  $n_x$  – число вариантов, меньших  $x$ ,  $n$  – объем выборки. Другими словами, для данного  $x$  эмпирическая функция распределения представляет накопленную частоту.

Эмпирическую функцию распределения  $F^*(x)$  составляем по точечному ряду распределения Таблицы 5.

**Таблица 5**

$[x_i, x_{i+1})$	-6	-4	-2	0	2	4	6	8
$n_i$	6	20	27	24	13	5	3	2
$\frac{n_i}{n}$	0,06	0,2	0,27	0,24	0,13	0,05	0,03	0,02
$\frac{n_i}{hn}$	0,03	0,1	0,135	0,12	0,065	0,025	0,015	0,01

$$F^*(x) = \begin{cases} 0, & x \leq -6, \\ 0,06, & -6 < x \leq -4, \\ 0,26, & -4 < x \leq -2, \\ 0,53, & -2 < x \leq 0, \\ 0,77, & 0 < x \leq 2, \\ 0,90, & 2 < x \leq 4, \\ 0,95, & 4 < x \leq 6, \\ 0,98, & 6 < x \leq 8, \\ 1, & x > 8. \end{cases}$$

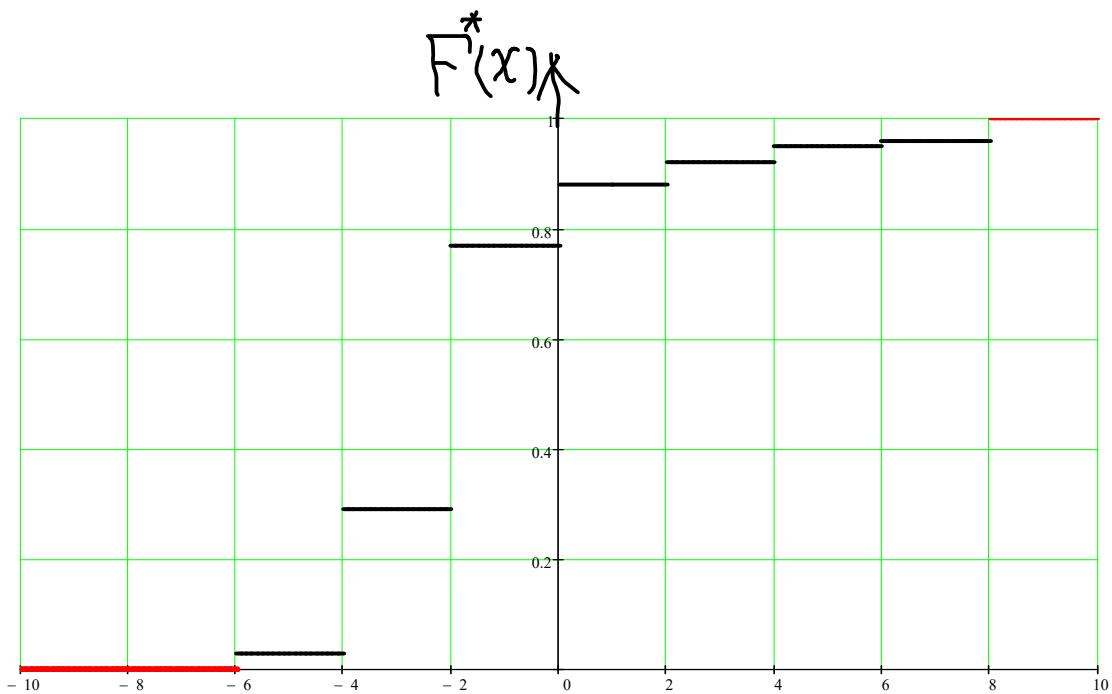


Рис. 1

Эмпирическая функция распределения  $F^*(x)$  позволяет определить эмпирическую вероятность (относительную частоту)  $P^*(\alpha < X < \beta)$  того, что значения СВ  $X$  попадут в интервал  $(\alpha, \beta)$ :  $P^*(\alpha < X < \beta) = F^*(\beta) - F^*(\alpha)$ . Например,

$$P^*(-3,5 < X < 4,8) = F^*(4,8) - F^*(-3,5) = 0,95 - 0,26 = 0,69 (69\%).$$

#### 4. Графическое представление статистических данных

Для наглядного представления о поведении исследуемой случайной величины в выборке можно строить различные графики. Наиболее часто используют следующие виды графического представления характеристик выборки: **полигон (многоугольник)** частот или относительных частот, **гистограмма** относительных частот или плотности относительных частот. Гистограмма и полигон позволяют выявить преобладающие значения СВ  $X$  и характер распределения частот и относительных частот.

**Полигон** частот – ломаная линия с координатами  $(x_i, n_i)$ , где  $x_i$  откладываются на оси абсцисс, а  $n_i$  – на оси ординат. Если на оси ординат откладывать не абсолютные, а относительные частоты  $w_i = n_i/n$ , то получим полигон относительных частот.

Полигон служит обычно для представления дискретного вариационного ряда. В системе координат  $(x, w)$  строятся точки, соответствующие значениям частот или относительных частот ряда, а затем эти точки соединяются прямыми линиями. Для группированного ряда полигон относительных частот строится по точечному ряду Таблицы 5.

#### Фрагмент Таблицы 5

$[x_i, x_{i+1})$	-6	-4	-2	0	2	4	6	8
$n_i/n$	0,06	0,2	0,27	0,24	0,13	0,05	0,03	0,02



Рис.2

Для непрерывного признака графической иллюстрацией служит гистограмма, то есть ступенчатая фигура, состоящая из прямоугольников, основаниями которых служат частичные интервалы длиной  $h$ , а высотами — отрезки длиной  $n_i$  (гистограмма частот) или  $\frac{n_i}{n}$  (гистограмма относительных частот) или  $\frac{n_i}{hn}$  — (гистограмма плотности относительных частот).

В первом случае площадь гистограммы равна объему выборки, во втором — единице, в третьем — тоже 1.

Гистограмма — это прямоугольники, с основаниями, равными интервалам значений признака и высотами, равными относительным частотам или их плотностям.

Полигон (гистограмма) аналогичны кривой распределения (плотности распределения), эмпирическая функция распределения — это функции распределения случайной величины.

Гистограмма — это диаграмма, используемая, как правило, для представления интервального вариационного ряда. Наиболее существенное отличие от полигона в том, что частота и относительная частота отображаются не точкой, а прямой, параллельной оси абсцисс на всем интервале. Это объясняется тем, что данная частота (относительная частота) относится не к дискретному значению признака, а ко всему интервалу.

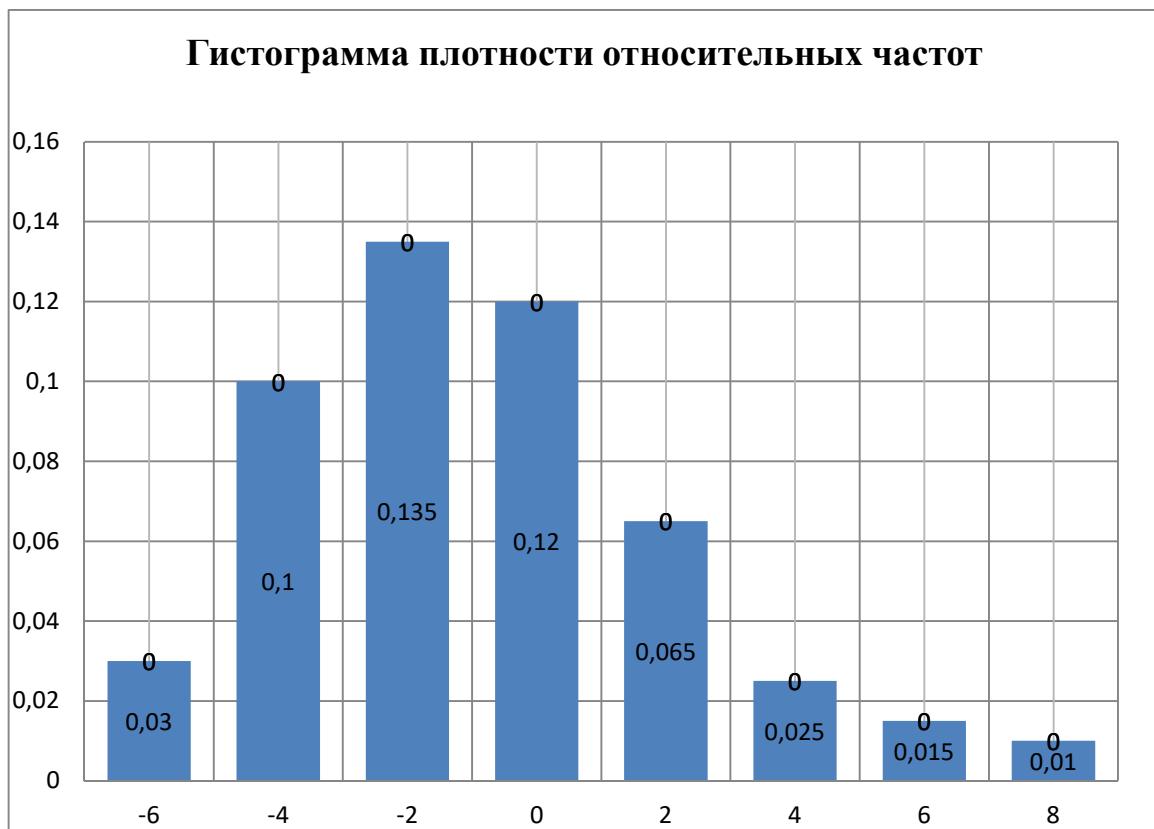
Составим гистограмму плотности относительных частот. Для того составим таблицу 6.

**Таблица 6**

$[x_i, x_{i+1})$	$[-7; -5)$	$[-5; -3)$	$[-3; -1)$	$[-1; 1)$	$[1; 3)$	$[3; 5)$	$[5; 7)$	$[7; 9)$	$\Sigma$
$\bar{x}_i$	-6	-4	-2	0	2	4	6	8	
$n_i$	6	20	27	24	13	5	3	2	100
$\frac{n_i}{n}$	0,06	0,2	0,27	0,24	0,13	0,05	0,03	0,02	1
$\frac{n_i}{hn}$	0,03	0,1	0,135	0,12	0,065	0,025	0,015	0,01	0,5

**Фрагмент Таблицы 6**

$[x_i, x_{i+1})$	$[-7; -5)$	$[-5; -3)$	$[-3; -1)$	$[-1; 1)$	$[1; 3)$	$[3; 5)$	$[5; 7)$	$[7; 9)$	$\Sigma$
$\bar{x}_i$	-6	-4	-2	0	2	4	6	8	
$\frac{n_i}{hn}$	0,03	0,1	0,135	0,12	0,065	0,025	0,015	0,01	0,5



**Рис. 3**

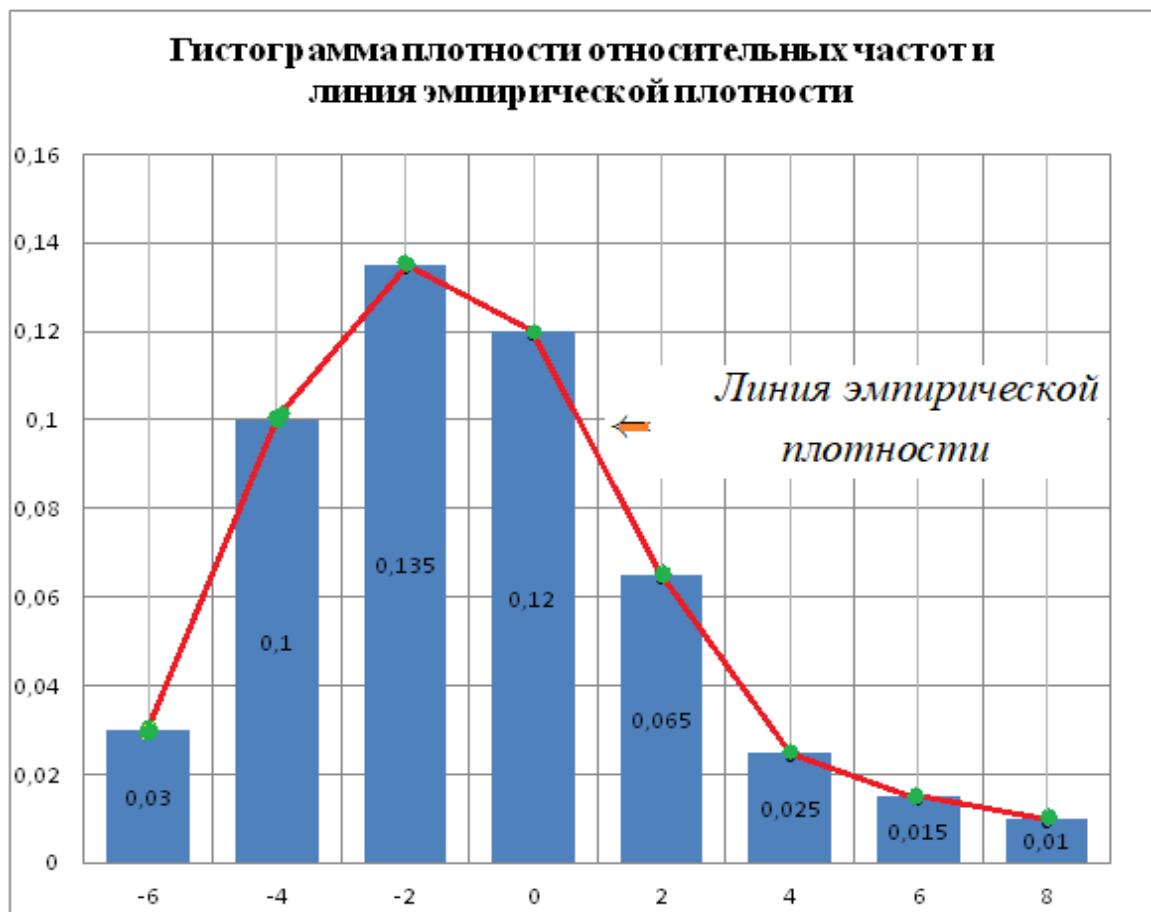


Рис. 4

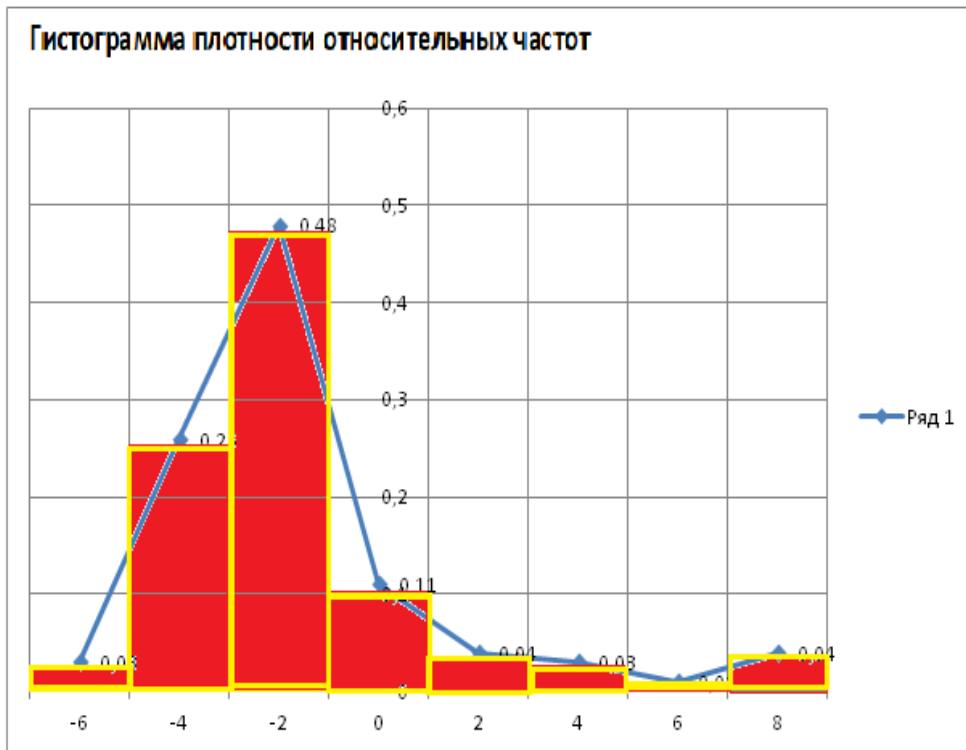


Рисунок с MathCAD

Найденная ломаная является графиком эмпирического (опытного) закона распределения плотности вероятностей признака  $X$  (случайной величины) в генеральной совокупности (в отличие от неизвестного истинного, теоретического закона распределения).

В результате рассмотрения вопроса № 3.1 студенты должны:

- ознакомиться с вероятностными и статистическими методами изучения моделей в биологии: оцениванием характеристик генеральной совокупности по выборке - методами группировки, построения таблиц частот и гистограмм;

- приобрести умения и навыки оценивания характеристик генеральной совокупности по выборке: методов группировки, построения таблиц частот и гистограмм.

3.2 Краткое содержание вопроса 3.2. Вероятностный и статистический методы изучения моделей в биологии: оценивание характеристик генеральной совокупности по выборке – числовые характеристики статистического распределения.

Задание для работы:

1. Краткие теоретические сведения.

2. Основное задание. Используя результаты выполнения задания вопроса 3.1 для той же выборки объёма  $n = 100$  (Таблица 1) следует найти

1) *Размах* выборки по формуле

$$R = x_{\max} - x_{\min}.$$

сгруппированного вариационного ряда (эмпирического ряда распределения) (Таблица 2).

2) Выборочную **медиану** по формуле (4) или (5) сгруппированного вариационного ряда (эмпирического ряда распределения) (Таблица 2) и **моду**.

3) **Выборочное среднее** по формуле

$$\bar{x}_e = \bar{x} = \frac{\sum_{i=1}^m x_i^* \cdot n_i}{n} = \sum_{i=1}^m x_i^* \cdot \omega_i,$$

где -  $m$  количество интервалов,  $\omega_i = \frac{n_i}{n}$  – частность (относительная частота), соответствующая  $i$ -й варианте или  $i$ -му частичному интервалу;  $x_i^*$  – середина  $i$ -го частичного интервала группировки для интервальный точечного вариационного ряда (эмпирического ряда распределения) - Таблица 5.

4) **Выборочную дисперсию** по формуле

$$D_e = \frac{\sum_{i=1}^m (x_i - \bar{x}) \cdot n_i}{n},$$

для интервальный точечного вариационного ряда (эмпирического ряда распределения) - Таблица 5.

### 5) среднее выборочное квадратичное отклонение

$$\sigma_e = \sqrt{D_e}.$$

### Числовые характеристики вариационных рядов

К числовым характеристикам обычно относят так называемые средние величины и величины, характеризующие разброс данных вокруг средних величин, а также некоторые другие дополнительные параметры, описывающие характер распределения опытных данных. Средние величины – это характеристики, обобщенно представляющие одним числом всю выборку.

Существует несколько видов средних величин: средняя арифметическая, средняя геометрическая, средняя гармоническая, средняя квадратическая и т. д. Какой из них пользоваться в каждом конкретном случае определяется тем, какое свойство ряда желательно описать данной величиной.

Первый шаг к осмыслению скрытых в выборке закономерностей – это ее графическое представление, то есть построение гистограммы, полигона частот и эмпирической функции распределения. Однако выборки, имеющие похожие графические изображения, могут различаться своими числовыми характеристиками. Числовые характеристики вариационных рядов вычисляют по данным, полученным в результате наблюдений (статистическим данным), поэтому их называют также статистическими характеристиками или оценками. Выборочные характеристики являются оценками соответствующих характеристик генеральной совокупности. Эти оценки должны удовлетворять определенным требованиям. В соответствии с важнейшими требованиями, оценки должны быть:

- *несмешенными*, то есть стремиться к истинному значению характеристики генеральной совокупности при неограниченном увеличении количества испытаний;

- *состоятельными*, то есть с ростом размера выборки оценка должна стремиться к значению соответствующего параметра генеральной совокупности с вероятностью, приближающейся к 1;

- *эффективными*, то есть для выборок равного объема используемая оценка должна иметь минимальную дисперсию. Выборка может характеризоваться следующими числовыми характеристиками.

#### 1. Характеристики положения.

Самой известной и наиболее употребляемой характеристикой любого вариационного ряда является его средняя арифметическая, называемая также *выборочным средним*. Средняя арифметическая характеризует значения признака, вокруг которого концентрируются наблюдения, т. е. центральную тенденцию распределения. При статистическом анализе выборки, кроме средней арифметической, широко применяют структурные, или порядковые, средние, к которым относятся *медиана* и *мода*. Выборочное среднее рассчитывается по формуле

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

(1)

где  $x_i$  – варианты **простого** вариационного ряда,  $n$  – объём выборки.

Если же анализируемые данные представлены в виде сгруппированного вариационного ряда, то для вычисления выборочного среднего применяется одно из следующих соотношений:

- для дискретного вариационного ряда

$$\bar{x}_e = \bar{x} = \frac{\sum_{i=1}^m x_i \cdot n_i}{n}, \quad (2)$$

где –  $m$  количество интервалов;

- для интервального вариационного ряда

$$\bar{x}_e = \bar{x} = \frac{\sum_{i=1}^m x_i^* \cdot n_i}{n} = \sum_{i=1}^m x_i^* \cdot \omega_i, \quad (3)$$

где  $\omega_i = \frac{n_i}{n}$  – частность (относительная частота), соответствующая  $i$ -й варианте или  $i$ -му

частичному интервалу;  $x_i^*$  – середина  $i$ -го частичного интервала группировки.

Выборочная **медиана** разбивает выборку пополам: слева и справа от нее оказывается одинаковое число элементов выборки. Если число элементов выборки четно,  $n = 2k$ , то выборочную медиану определяют по формуле

$$MeX = \frac{x_k + x_{k+1}}{2}, \quad (4)$$

При нечетном  $n = 2k + 1$  объеме выборки медиану находят по формуле

$$MeX = x_{k+1}. \quad (5)$$

**Мода**  $MoX$  используется для нахождения наиболее часто встречающегося в выборке значения.

## 2. Характеристики рассеяния.

Для получения полного представления о вариационном ряде (определен центральную тенденцию распределения с помощью характеристик положения) далее оценивают

рассеяние (вариацию, изменчивость) исследуемого признака вокруг этих величин. Простейшим и, весьма приближенным показателем вариации (изменчивости), является вариационный *размах*. Размах вариации наиболее полезен, если нужен быстрый и общий взгляд на изменчивость при сравнении большого количества выборок. *Размах* выборки вычисляется по формуле

$$R = x_{\max} - x_{\min}. \quad (1)$$

Но наибольший интерес представляют меры вариации (рассеяния) наблюдений вокруг средних величин, в частности, вокруг средней арифметической. К таким оценкам относятся *выборочная дисперсия* и *среднее квадратичное отклонение*. Дисперсия выборки – это параметр, характеризующий степень разброса элементов выборки относительно среднего значения  $x$ . Чем больше дисперсия, тем дальше отклоняются значения элементов выборки от среднего значения.

**Выборочная дисперсия** находится по формуле

$$D_e = \frac{\sum_{i=1}^m (x_i - \bar{x})^2 \cdot n_i}{n}, \quad (7)$$

если данные представлены в виде сгруппированного вариационного ряда

Выборочная дисперсия обладает одним существенным недостатком: если среднее арифметическое выражается в тех же единицах, что и значения случайной величины, то, согласно определению, дисперсия выражается уже в квадратных единицах. Этого недостатка можно избежать, если использовать в качестве меры вариации признака *среднее выборочное квадратичное отклонение*

$$\sigma_e = \sqrt{D_e}.$$

Существуют ещё и *исправленная дисперсия*, и *исправленное среднее квадратичное отклонение*.

Для выборки из предыдущего занятия находим

$$R = 10 - (-6) = 16,$$

$$MeX = 2, MoX = -2,$$

$$\bar{x} = -2,26, D_e(X) = 8,6124, \sigma_e(X) = 2,93.$$

В результате рассмотрения вопроса № 3.2 студенты должны

- получить представление о вероятностных и статистических методах изучения моделей в биологии; нахождение численных характеристик вариационных рядов;

- приобрести умения и навыки нахождение численных характеристик вариационных рядов.

3.3 Краткое описание вопроса № 3.3. Вероятностный и статистический методы изучения моделей в биологии: проверка гипотез о параметрах и виде распределения, доверительные интервалы.

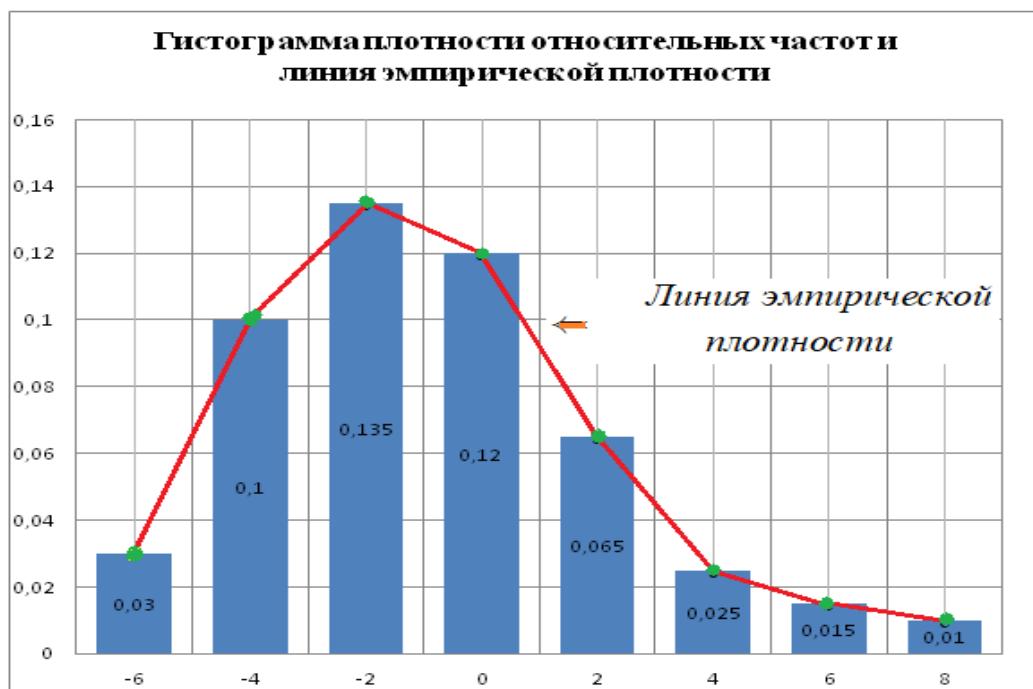
Задание для работы:

1. Краткие теоретические сведения.

2. Основное задание. Используя результаты выполнения **задания** в вопросе № 3.2 для той же выборки объёма  $n = 100$  (Таблица 1) следует
- 6) Выдвинуть **статистическую гипотезу о виде теоретического закона** (нормального) **распределения** в генеральной совокупности по результатам пункта 9) задания вопроса 3.1
  - 7) Вычислить теоретические частоты и сравнить их с эмпирическими.
  - 8) Проверить согласие гипотезы о нормальном законе распределения изучаемой СВ  $X$  по критерию Пирсона.
  - 9) Найти доверительный интервал для оценки неизвестного математического ожидания нормального распределения.

**Предположение о виде теоретического закона распределения случайной величины в генеральной совокупности (статистическая гипотеза о виде теоретического закона распределения)**

В пункте 9) задания вопроса 3.1 на Рис.4, построена линия эмпирической плотности



По виду линии эмпирической плотности делаем предположение о теоретическом законе распределения. В данном примере линия эмпирической плотности напоминает кривую Гаусса. Следовательно, можно предположить, что закон распределения СВ  $X$  нормальный. Плотность нормально распределенной СВ определяется формулой

$$\varphi(x) = \frac{1}{\sigma\sqrt{2\pi}} \cdot e^{-\frac{(x-a)^2}{2\sigma^2}}.$$

Параметры этого распределения:  $a$  – математическое ожидание СВ  $X$ ,  $\sigma$  – среднее квадратическое отклонение СВ  $X$ . Положим  $a \approx \bar{x} \Rightarrow a \approx -2$ ;  $\sigma \approx \bar{\sigma}_e \Rightarrow \sigma \approx 3$ .

Тогда теоретический закон распределения (статистическая гипотеза) будет задан формулой

$$\varphi(x) = \frac{1}{\sigma_e \cdot \sqrt{2\pi}} \cdot e^{-\frac{(x-\bar{x})^2}{2\cdot\sigma_e^2}} = \frac{1}{3 \cdot \sqrt{2\pi}} \cdot e^{-\frac{(x-2)^2}{18}}. \quad (8)$$

Для вариационного ряда относительная частота  $\frac{n_i}{n}$  может рассматриваться как эмпирическая вероятность распределения, т. е.

$$\frac{n_i}{n} = p_i^*$$

может рассматриваться как эмпирическая вероятность попадания СВ  $X$  в интервал  $(x_i; x_{i+1})$

$$\frac{n_i}{n} = p^*(x_i < X < x_{i+1}), \quad i = 1, 2, \dots, m.$$

Полагая  $p(x_i < X < x_{i+1}) \approx p^*(x_i < X < x_{i+1})$ ,  $i = 1, 2, \dots, m$ , получим

$$\frac{n_i}{n} \approx p(x_i < X < x_{i+1}), \quad i = 1, 2, \dots, m, \quad (*)$$

где  $p(x_i < X < x_{i+1})$  уже теоретическая (в генеральной совокупности) вероятность попадания СВ  $X$  в интервал  $(x_i; x_{i+1})$ .

Для определения вероятности попадания СВ  $X$  в интервал  $(x_i; x_{i+1})$  в ТВ существует формула

$$p(x_i < X < x_{i+1}) = \int_{x_i}^{x_{i+1}} \varphi(x) dx = F(x_{i+1}) - F(x_i) \quad (**)$$

Функция нормального распределения имеет вид

$$F(x) = 0,5 + \Phi\left(\frac{x - a}{\sigma}\right).$$

Поэтому, для нормально распределённой СВ  $X$

$$p(x_i < X < x_{i+1}) = \Phi\left(\frac{x_{i+1} - \bar{x}_e}{\bar{\sigma}_e}\right) - \Phi\left(\frac{x_i - \bar{x}_e}{\bar{\sigma}_e}\right).$$

Таким образом,

$$\frac{n_i}{n} \approx p(x_i < X < x_{i+1}) = \Phi\left(\frac{x_{i+1} - \bar{x}_\epsilon}{\bar{\sigma}_\epsilon}\right) - \Phi\left(\frac{x_i - \bar{x}_\epsilon}{\bar{\sigma}_\epsilon}\right).$$

$$\frac{n_i}{n} \approx \Phi\left(\frac{x_{i+1} - \bar{x}_\epsilon}{\bar{\sigma}_\epsilon}\right) - \Phi\left(\frac{x_i - \bar{x}_\epsilon}{\bar{\sigma}_\epsilon}\right).$$

$$n_i^0 \approx n \cdot \left[ \Phi\left(\frac{x_{i+1} - \bar{x}_\epsilon}{\bar{\sigma}_\epsilon}\right) - \Phi\left(\frac{x_i - \bar{x}_\epsilon}{\bar{\sigma}_\epsilon}\right) \right], \quad i = 1, 2, \dots, m. \quad (9)$$

Найденные по формуле (9) частоты называются теоретическими и обозначаются  $n_i^0$ . Для их вычисления берутся концы  $x_i$  и  $x_{i+1}$  интервалов из интервального ряда таблицы 6.

**Таблица 6**

$[x_i, x_{i+1})$	$[-7; -5)$	$[-5; -3)$	$[-3; -1)$	$[-1; 1)$	$[1; 3)$	$[3; 5)$	$[5; 7)$	$[7; 9)$	$\Sigma$
$\bar{x}_i$	-6	-4	-2	0	2	4	6	8	
$n_i$	6	20	27	24	13	5	3	2	100

В нашем примере  $n = 100$ ,  $\bar{x}_\epsilon = -2$ ,  $\bar{\sigma}_\epsilon = 3$ , поэтому

$$n_i^0 \approx 100 \cdot \left[ \Phi\left(\frac{x_{i+1} + 2}{3}\right) - \Phi\left(\frac{x_i + 2}{3}\right) \right], \quad i = 1, 2, \dots, 8. \quad (10)$$

Это формула для вычисления теоретических частот в нашем примере.

Последовательно вычисляем теоретические частоты, считая

$$x_1 = -7, x_2 = -5, x_3 = -3, x_4 = -1, x_5 = 1, x_6 = 3, x_7 = 5, x_8 = 7, x_9 = 9.$$

\*\*\*\*\*

$$i = 1$$

$$\begin{aligned} n_1^0 &\approx 100 \cdot \left[ \Phi\left(\frac{x_2 + 2}{3}\right) - \Phi\left(\frac{x_1 + 2}{3}\right) \right] = 100 \cdot \left[ \Phi\left(\frac{-5 + 2}{3}\right) - \Phi\left(\frac{-7 + 2}{3}\right) \right] = \\ &= 100 \cdot [\Phi(-1) - \Phi(-1,667)] = 100 \cdot [\Phi(1,667) - \Phi(1)] = \\ &= 100 \cdot [0,4525 - 0,3413] = 100 \cdot 0,1112 = 11,12. \end{aligned}$$

\*\*\*\*\*

$$i = 2$$

$$n_2^0 \approx 100 \cdot \left[ \Phi\left(\frac{x_3+2}{3}\right) - \Phi\left(\frac{x_2+2}{3}\right) \right] = 100 \cdot \left[ \Phi\left(\frac{-3+2}{3}\right) - \Phi\left(\frac{-5+2}{3}\right) \right] =$$

$$= 100 \cdot [\Phi(-0,333) - \Phi(-1)] = 100 \cdot [\Phi(1) - \Phi(0,333)] =$$

$$= 100 \cdot [0,3413 - 0,1293] = 100 \cdot 0,212 = 21,2 .$$


---

$$i = 3$$

$$n_3^0 \approx 100 \cdot \left[ \Phi\left(\frac{x_4+2}{3}\right) - \Phi\left(\frac{x_3+2}{3}\right) \right] = 100 \cdot \left[ \Phi\left(\frac{-1+2}{3}\right) - \Phi\left(\frac{-3+2}{3}\right) \right] =$$

$$= 100 \cdot [\Phi(0,333) - \Phi(-0,333)] = 100 \cdot [2 \cdot \Phi(0,333)] =$$

$$= 100 \cdot [2 \cdot 0,1293] = 100 \cdot 0,2586 = 25,86 .$$

\*\*\*\*\*

$$i = 4$$

$$n_4^0 \approx 100 \cdot \left[ \Phi\left(\frac{x_5+2}{3}\right) - \Phi\left(\frac{x_4+2}{3}\right) \right] = 100 \cdot \left[ \Phi\left(\frac{1+2}{3}\right) - \Phi\left(\frac{-1+2}{3}\right) \right] =$$

$$= 100 \cdot [\Phi(1) - \Phi(0,333)] = 100 \cdot [0,3413 - 0,1293] = 100 \cdot 0,212 = 21,2 .$$

\*\*\*\*\*

$$i = 5$$

$$n_5^0 \approx 100 \cdot \left[ \Phi\left(\frac{x_6+2}{3}\right) - \Phi\left(\frac{x_5+2}{3}\right) \right] = 100 \cdot \left[ \Phi\left(\frac{3+2}{3}\right) - \Phi\left(\frac{1+2}{3}\right) \right] =$$

$$= 100 \cdot [\Phi(1,667) - \Phi(1)] = 100 \cdot [0,4525 - 0,3413] = 100 \cdot 0,1112 = 11,12 .$$

\*\*\*\*\*

$$i = 6$$

$$n_6^0 \approx 100 \cdot \left[ \Phi\left(\frac{x_7+2}{3}\right) - \Phi\left(\frac{x_6+2}{3}\right) \right] = 100 \cdot \left[ \Phi\left(\frac{5+2}{3}\right) - \Phi\left(\frac{3+2}{3}\right) \right] =$$

$$= 100 \cdot [\Phi(2,333) - \Phi(1,667)] = 100 \cdot [0,49 - 0,4525] = 100 \cdot 0,0375 = 3,75 .$$

\*\*\*\*\*

$$i = 7$$

$$n_7^0 \approx 100 \cdot \left[ \Phi\left(\frac{x_8+1,6}{2,86}\right) - \Phi\left(\frac{x_7+1,6}{2,86}\right) \right] = 100 \cdot \left[ \Phi\left(\frac{7+2}{3}\right) - \Phi\left(\frac{5+2}{3}\right) \right] =$$

$$=100 \cdot [\Phi(3)-\Phi(2,333)]=100 \cdot [0,4987-0,4898]=100 \cdot 0,00865=0,865.$$

\*\*\*\*\*

$$i=8$$

$$n_8^0 \approx 100 \cdot \left[ \Phi\left(\frac{x_9+2}{3}\right) - \Phi\left(\frac{x_8+2}{3}\right) \right] = 100 \cdot \left[ \Phi\left(\frac{9+2}{3}\right) - \Phi\left(\frac{7+2}{3}\right) \right] =$$

$$=100 \cdot [\Phi(3,667)-\Phi(3)]=100 \cdot [0,5-0,4986]=100 \cdot 0,0014=0,14.$$

\*\*\*\*\*

В результате получим сравнительную таблицу эмпирических и теоретических частот

$[x_i, x_{i+1})$	$[-7; -5)$	$[-5; -3)$	$[-3; -1)$	$[-1; 1)$	$[1; 3)$	$[3; 5)$	$[5; 7)$	$[7; 9)$	$\Sigma$
$n_i$	6	20	27	24	13	5	3	2	100
$n_i^0$	11,12	21,2	25,86	21,2	11,12	3,75	0,87	0,14	

Проверяем согласие гипотезы о нормальном законе распределения изучаемой СВ  $X$  по критерию Пирсона, вычислив значение

$$\chi_{\text{набл}}^2 = \sum \frac{(n_i - n_i^0)^2}{n_i^0}$$

и сравнив его со значением  $\chi_{kp}^2$ , найденным по таблице критических точек распределения Пирсона  $\chi^2$  при уровне значимости  $\alpha = 0,01$  и числе степеней свободы

$$k = l - 1 - 2 = l - 3,$$

где  $l$  – количество интервалов. Если  $\chi_{\text{набл}}^2 < \chi_{kp}^2$ , то гипотеза о нормальном законе распределения принимается.

Находим доверительный интервал для оценки неизвестного математического ожидания

$$P(|\bar{x}_\epsilon - a| < \delta) = 2\Phi(t) \Rightarrow P(\bar{x}_\epsilon - \delta < a < \bar{x}_\epsilon + \delta) = 2\Phi(t),$$

$$\text{где } t = \frac{\delta \sqrt{n}}{\bar{\sigma}_\epsilon} \Rightarrow \delta = \frac{t \bar{\sigma}_\epsilon}{\sqrt{n}}.$$

$$P(\bar{x}_\epsilon - \delta < a < \bar{x}_\epsilon + \delta) = \gamma,$$

$$P\left(\bar{x}_\epsilon - \frac{t \bar{\sigma}_\epsilon}{\sqrt{n}} < a < \bar{x}_\epsilon + \frac{t \bar{\sigma}_\epsilon}{\sqrt{n}}\right) = 2\Phi(t) = \gamma.$$

$$2\Phi(t) = \gamma, \quad \Phi(t) = \frac{\gamma}{2},$$

$$\bar{x}_e - \frac{t\bar{\sigma}_e}{\sqrt{n}} < a < \bar{x}_e + \frac{t\bar{\sigma}_e}{\sqrt{n}} \Rightarrow a \in \left( \bar{x}_e - \frac{t\bar{\sigma}_e}{\sqrt{n}}; \bar{x}_e + \frac{t\bar{\sigma}_e}{\sqrt{n}} \right).$$

$$\gamma = 0,95, \Phi(t) = \frac{\gamma}{2} = 0,475 \Rightarrow t = 1,96,$$

Так как  $\bar{\sigma}_e = 3$ ,  $\bar{x}_e = -2$ , то  $\delta = \frac{t\bar{\sigma}_e}{\sqrt{n}} = \frac{1,96 \cdot 3}{\sqrt{10}} = 0,588 \approx 0,57$ . Потому доверительный

интервал для оценки неизвестного математического ожидания нормального распределения при известной точности оценки  $\delta \approx 0,57$  с надёжностью (доверительной вероятностью)  $\gamma = 0,95$  будет

$$\left( \bar{x}_e - \frac{t\bar{\sigma}_e}{\sqrt{n}}; \bar{x}_e + \frac{t\bar{\sigma}_e}{\sqrt{n}} \right) = (-2,57; -2 + 0,57) = (-2,16; -1,43).$$

В результате рассмотрения вопроса № 3.3 студенты должны

- получить представление о вероятностных и статистических методах изучения моделей в биологии: проверке гипотезы о параметрах и виде распределения, нахождении доверительных интервалов для оценки неизвестного математического ожидания нормального распределения;

- приобрести умения и навыки проверки гипотез о параметрах и виде распределения, нахождения доверительных интервалов для оценки неизвестного математического ожидания нормального распределения.

#### 4. Наименование вопроса № 4.

Оптимационные модели биологических процессов и систем. Оптимационные модели на графах и сетях.

4.1 Наименование вопроса № 4.1. Математические модели оптимационных задач линейного программирования в биологии.

4.1.1 Наименование вопроса № 4.1.1. Математическая модель задачи об оптимизации распределения ресурсов.

4.1.2 Наименование вопроса № 4.1.2. Математическая модель калькуляционной задачи линейного программирования (задачи о выборе оптимального рациона кормления животных).

4.1.3 Наименование вопроса № 4.1.3. Математическая модель задачи об оптимизации транспортных расходов.

4.1.4 Наименование вопроса № 4.1.4 Математическая модель задачи о назначениях, задачи целочисленного программирования.

4.2 Наименование вопроса № 4.2. Математические модели задач оптимизации на графах и сетях в биологии: задача об остеове минимального веса сети и её решение по алгоритму Краскала.

4.1.1. Краткое содержание вопроса № 4.1.1: Математическая модель задачи об оптимизации распределения ресурсов.

1. Формулировка задачи о распределении ресурсов.

**Задача.** Предприятие с помощью биотехнологий производит и продаёт продукцию двух видов: «1 Продукт» и «2 Продукт». Для производства продукции используются ресурсы двух категорий: А и В. Расходы ресурсов А и В на производство единицы продукции каждого вида, запасы ресурсов и цены продукции приведены в таблице I.

Таблица I

Ресурсы	Расход ресурсов на ед. продукции		Запасы ресурсов
	1 Продукт	2 Продукт	
А	1	2	3
В	3	1	3
Количество продукции	$x_1$	$x_2$	
Цены	2(ден. ед.)	1(ден. ед.)	

Выяснить, какое количество продукции каждого вида надо производить предприятию (составить план производства), чтобы получить максимум прибыли.

2. Создание математической модели задачи, понятие о задаче линейного программирования, математическом программировании.

Для составления математической модели задачи прежде всего **введём переменные (неизвестные) задачи**:  $x_1$  - количество продукции 1-го вида, а  $x_2$  - количество продукции 2-го вида, производимые предприятием.

Ограниченнность запасов ресурсов приводит к **ограничениям на  $x_1$  и  $x_2$**

- ограничения на расход ресурса А:  $x_1 + 2 \cdot x_2 \leq 3$ ,
- ограничения на расход ресурса В:  $3 \cdot x_1 + x_2 \leq 3$ .

Кроме того,  $x_1, x_2 \geq 0$ .

Качество решения задачи определяется с помощью **целевой функции**  $Z(x_1, x_2)$  **данной задачи** – функции, определяющей доход предприятия от продажи продукции:

$$Z(x_1, x_2) = 2 \cdot x_1 + x_2.$$

Задача об определении плана производства продукции свелась к следующей математической задаче: *найти вектор  $(x_1, x_2)$  (план производства), координаты которого удовлетворяют системе ограничений*

$$\begin{cases} x_1 + 2 \cdot x_2 \leq 3, \\ 3 \cdot x_1 + x_2 \leq 3 \end{cases}$$

и условиям неотрицательности  $x_1, x_2 \geq 0$ , который доставляет максимум целевой функции  $Z(x_1, x_2) = 2 \cdot x_1 + x_2$ .

Эту математическую задачу принято записывать в виде

$$Z(x_1, x_2) = 2 \cdot x_1 + x_2 \rightarrow \max, \quad (1)$$

$$\begin{cases} x_1 + 2 \cdot x_2 \leq 3, \\ 3 \cdot x_1 + x_2 \leq 3, \end{cases} \quad (2)$$

$$x_1, x_2 \geq 0$$

(3)

и называть **математической моделью** данной производственной задачи.

Подобные задачи называются **задачами линейного программирования**. Они изучаются в разделе математики, называемом **математическим программированием**. Так как переменные  $x_1$  и  $x_2$  входят в систему ограничений (2) и целевую функцию  $Z(x_1, x_2)$  линейно, то эту задачу математического программирования называют **задачей линейного программирования**.

Множество точек декартовой плоскости  $(x_1, x_2)$ , координаты которых удовлетворяют системе ограничений (2) и условиям неотрицательности (3), называется областью допустимых решений задачи линейного программирования (областью допустимых планов). В данной задаче она представляет собой выпуклый четырёхугольник. Значения  $x_1^*$  и  $x_2^*$  из области допустимых планов, при которых  $Z(x_1, x_2)$  принимает наибольшее значение в этой области, называются **оптимальными (оптимальный план)**, а соответствующее наибольшее значение  $Z^* = 2 \cdot x_1^* + x_2^*$  является **оптимальным значением прибыли**. Таким образом, задача о распределении ресурсов является задачей оптимизации и её математической моделью служит задача линейного программирования, заключающаяся в поиске оптимального плана и оптимального значения целевой функции.

Задачей оптимизации может быть поиск наименьшего значения.

В результате рассмотрения вопроса № 4.1.1 студенты должны:

- освоить понятие о задачах (моделях) оптимизации распределения ресурсов, моделях (задачах) линейного программирования в биологии;

- приобрести умения и навыки создания моделей линейного программирования в биологии.

4.1.2 Краткое содержание вопроса № 4.1.2. Математическая модель калькуляционной задачи линейного программирования (задачи о выборе оптимального рациона кормления животных).

Задание для работы:

1. Формулировка калькуляционной задачи линейного программирования (задачи о выборе оптимального рациона кормления животных).

2. Создание математической модели задачи.

1. Формулировка калькуляционной задачи линейного программирования (задачи о

выборе оптимального рациона кормления животных).

**Задача.** Составляется комбинированный корм из трёх злаков: кукурузы, овса и ржи. Калорийность и содержание витамина С в одном килограмме каждого злака, а также цена одного кг каждого злака указаны в таблице:

	Кукуруза	Овёс	Рожь
Калорийность(ккал)	200	175	100
Содержание С (в Гр)	5	1	3
Цена (ден. ед.)	6	4	1

Составить наиболее дешёвый комбинированный корм, 1кг которого содержал бы не менее 150 ккал и не менее 3 г витамина С.

## 2. Создание математической модели задачи.

Обозначим содержание кукурузы, овса и ржи в 1кг комбикорма через  $x_1$ ,  $x_2$ ,  $x_3$  соответственно, а стоимость злаков в комбикорме  $Z$ . Тогда математическая модель задачи об оптимизации производства комбикорма формулируется следующим образом: найти вектор  $x = (x_1; x_2; x_3)$  (называемый планом), доставляющий минимум целевой функции задачи (функции затрат)

$$Z = 6 \cdot x_1 + 4 \cdot x_2 + x_3 \rightarrow \min ,$$

координаты которого  $x_1$ ,  $x_2$ ,  $x_3$  удовлетворяют системе ограничений

$$\begin{cases} 200 \cdot x_1 + 175 \cdot x_2 + 100 \cdot x_3 \geq 150, \\ 5 \cdot x_1 + x_2 + 3 \cdot x_3 \geq 3, \\ x_1 + x_2 + x_3 = 1 \end{cases}$$

и условиям неотрицательности

$$x_1 \geq 0. x_2 \geq 0. x_3 \geq 0 .$$

В результате рассмотрения вопроса № 4.1.2 студенты должны:

- освоить понятие о калькуляционных моделях (задачах) линейного программирования в биологии;
- приобрести умения и навыки создания калькуляционных моделей линейного программирования в биологии.

## 4.1.3 Краткое содержание вопроса № 4.1.3. Математическая модель задачи об оптимизации транспортных расходов.

Задание для работы:

1. Формулировка специальной задачи линейного программирования: транспортной задачи.
  2. Создание математической модели задачи.
1. Формулировка специальной задачи линейного программирования: транспортной задачи.

**Задача об оптимизации перевозок.** Четыре отделения сельхозпредприятия  $B_1, B_2, B_3, B_4$  закупают корма у трёх поставщиков  $A_1, A_2, A_3$ . Запасы кормов у поставщиков, потребности сельхозпредприятия в кормах и стоимость перевозки единицы продукта от поставщика к потребителю даны в таблице.

Потребители		Потребители				Запасы кормов у поставщиков
		$B_1$	$B_2$	$B_3$	$B_4$	
Поставщики	$A_1$	10 $x_{11}$	0 $x_{12}$	20 $x_{13}$	11 $x_{14}$	$a_1 = 15$
	$A_2$	12 $x_{21}$	7 $x_{22}$	9 $x_{23}$	20 $x_{24}$	$a_2 = 25$
$A_3$	0 $x_{31}$	14 $x_{32}$	16 $x_{33}$	18 $x_{34}$	$a_3 = 5$	
	$b_1 = 5$	$b_2 = 15$	$b_3 = 15$	$b_4 = 10$	45	
Потребность в кормах						

Значительную часть расходов с /х предприятия составляют именно транспортные расходы. Минимизировать расходы предприятия: составить такой план перевозок, при котором суммарные транспортные расходы будут минимальными, все запасы поставщиков будут вывезены, все потребности отделений с.\ х. предприятия будут удовлетворены.

## 2. Создание математической модели задачи.

Обозначим через  $x_{ij}$  количество кормов, перевозимое от поставщика  $A_i$  к потребителю  $B_j$ ,  $c_{ij}$ - стоимость перевозок по этому маршруту,  $i = 1, 2, 3; j = 1, 2, 3, 4$ .

**Целевая функция:** транспортные расходы на перевозку кормов вычисляются по формуле

$$\begin{aligned}
 Z = \sum_{i=1}^3 \sum_{j=1}^4 c_{ij} \cdot x_{ij} = & 10 \cdot x_{11} + 0 \cdot x_{12} + 20 \cdot x_{13} + 11 \cdot x_{14} + \\
 & + 12 \cdot x_{21} + 7 \cdot x_{22} + 9 \cdot x_{23} + 20 \cdot x_{24} + \\
 & + 0 \cdot x_{31} + 14 \cdot x_{32} + 16 \cdot x_{33} + 18 \cdot x_{34} \rightarrow \min.
 \end{aligned}$$

**Ограничения на переменные задачи.**

**Ограничения вывоза:** из  $A_1 \quad x_{11} + x_{12} + x_{13} + x_{14} = 15$ ,

из  $A_2 \quad x_{21} + x_{22} + x_{23} + x_{24} = 25$ ,

из  $A_3 \quad x_{31} + x_{32} + x_{33} + x_{34} = 5$ .

**Ограничения ввоза:** ввоз в  $B_1 \quad x_{11} + x_{21} + x_{31} = 5$ ,

$$\text{ввоз в } B_2 \quad x_{12} + x_{22} + x_{32} = 15,$$

$$\text{ввоз в } B_3 \quad x_{13} + x_{23} + x_{33} = 15,$$

$$\text{ввоз в } B_4 \quad x_{14} + x_{24} + x_{34} = 10.$$

Ограничения на знаки (значения) переменных:  $x_{ij} \geq 0$ .

Задача закрытого типа:  $a_1 + a_2 + a_3 = b_1 + b_2 + b_3 + b_4 = 45$ .

В результате рассмотрения вопроса № 4.1.3 студенты должны:

- освоить понятие о специальных задачах (транспортной задаче) линейного программирования;
- приобрести умения и навыки создания моделей специальных задач линейного программирования (транспортной задачи).

4.1.4 Краткое содержание вопроса № 4.1.4. Математическая модель задачи о назначениях, задачи целочисленного программирования.

Задание для работы:

1. Формулировка специальной задачи линейного программирования: задача о назначениях, задачи целочисленного программирования. Понятие о динамическом программировании.
2. Создание математической модели задачи.

1. Формулировка специальной задачи линейного программирования: задача о назначениях, задачи целочисленного программирования. Понятие о динамическом программировании.

**Задача о назначениях.** Производится классификация 7 болезней по 5 категориям. Результаты тестирования каждого заболевания (уровень опасности заболевания) по каждой категории выражены в баллах по 10-балльной шкале и представлены матрицей

*	$P_1$	$P_2$	$P_3$	$P_4$	$P_5$
$S_1$	7	5	7	6	7
$S_2$	6	4	8	4	9
$S_3$	8	6	4	3	8
$S_4$	7	7	8	5	7
$S_5$	5	9	7	9	5
$S_6$	6	8	6	4	7
$S_7$	7	7	8	6	4

$C =$

Определить самые опасные заболевания в каждой из 5 категорий так, чтобы сумма баллов выбранных заболеваний была наибольшей (суммарный уровень опасности вы-

бранных заболеваний был наибольшим). Каждое заболевание может быть самым опасным только в одной категории и все категории должны быть заняты.

2. Создание математической модели задачи.

Обозначим через  $x_{ij}$  переменные задачи:

$x_{ij} = 1$ , если заболевание  $S_i$  выбирается самым опасным в категории  $P_j$ ;

$x_{ij} = 0$ , если заболевание  $S_i$  не является самым опасным в категории  $P_j$ ;

$c_{ij}$ - уровень опасности заболевания  $S_i$  (количество баллов по результатам тестирования) в категории  $P_j$ ,  $i = 1, 2, 3, 4, 5, 6, 7; j = 1, 2, 3, 4, 5$ .

*Целевая функция:* суммарный уровень опасности всех заболеваний в баллах вычисляется по формуле

$$Z = \sum_{i=1}^7 \sum_{j=1}^5 c_{ij} \cdot x_{ij} = 7 \cdot x_{11} + 5 \cdot x_{12} + 7 \cdot x_{13} + 6 \cdot x_{14} + 7 \cdot x_{15} + \\ + 6 \cdot x_{21} + 4 \cdot x_{22} + 8 \cdot x_{23} + 4 \cdot x_{24} + 9 \cdot x_{25} + \\ + 8 \cdot x_{31} + 6 \cdot x_{32} + 4 \cdot x_{33} + 3 \cdot x_{34} + 8 \cdot x_{35} + \\ + 7 \cdot x_{41} + 7 \cdot x_{42} + 8 \cdot x_{43} + 5 \cdot x_{44} + 7 \cdot x_{45} + \\ + 5 \cdot x_{51} + 9 \cdot x_{52} + 7 \cdot x_{53} + 9 \cdot x_{54} + 5 \cdot x_{55} + \\ + 6 \cdot x_{61} + 8 \cdot x_{62} + 6 \cdot x_{63} + 4 \cdot x_{64} + 7 \cdot x_{65} + \\ + 7 \cdot x_{71} + 7 \cdot x_{72} + 8 \cdot x_{73} + 6 \cdot x_{74} + 4 \cdot x_{75} \rightarrow \max.$$

*Ограничения на переменные задачи.*

*Ограничения на лидерство одного заболевания:* каждое заболевание может быть самым опасным только в одной категории

$$\text{для } S_1 \quad x_{11} + x_{12} + x_{13} + x_{14} + x_{15} = 1,$$

$$\text{для } S_2 \quad x_{21} + x_{22} + x_{23} + x_{24} + x_{25} = 1,$$

$$\text{для } S_3 \quad x_{31} + x_{32} + x_{33} + x_{34} + x_{35} = 1,$$

$$\text{для } S_4 \quad x_{41} + x_{42} + x_{43} + x_{44} + x_{45} = 1,$$

$$\text{для } S_5 \quad x_{51} + x_{52} + x_{53} + x_{54} + x_{55} = 1,$$

$$\text{для } S_6 \quad x_{61} + x_{62} + x_{63} + x_{64} + x_{65} = 1,$$

$$\text{для } S_7 \quad x_{71} + x_{72} + x_{73} + x_{74} + x_{75} = 1.$$

*Ограничения по занятости категорий:* в каждой категории может быть только один лидер

для  $P_1 \quad x_{11} + x_{21} + x_{31} + x_{41} + x_{51} + x_{61} + x_{71} = 1$ ,

для  $P_2 \quad x_{12} + x_{22} + x_{32} + x_{42} + x_{52} + x_{62} + x_{72} = 1$ ,

для  $P_3 \quad x_{13} + x_{23} + x_{33} + x_{43} + x_{53} + x_{63} + x_{73} = 1$ ,

для  $P_4 \quad x_{14} + x_{24} + x_{34} + x_{44} + x_{54} + x_{64} + x_{74} = 1$ ,

для  $P_5 \quad x_{15} + x_{25} + x_{35} + x_{45} + x_{55} + x_{65} + x_{75} = 1$ ,

*Ограничения на знаки (значения) переменных:  $x_{ij} \geq 0$ ,  $x_{ij}$  - двоичные числа.*

Задача открытого типа: вводятся две фиктивные категории с нулевыми столбцами баллов. Матрица  $C^*$  становится квадратной.

*	$P_1$	$P_2$	$P_3$	$P_4$	$P_5$	$P_6$	$P_7$
$S_1$	7	5	7	6	7	0	0
$S_2$	6	4	8	4	9	0	0
$S_3$	8	6	4	3	8	0	0
$S_4$	7	7	8	5	7	0	0
$S_5$	5	9	7	9	5	0	0
$S_6$	6	8	6	4	7	0	0
$S_7$	7	7	8	6	4	0	0

В результате рассмотрения вопроса № 4.1.4 студенты должны:

- освоить понятие о специальных задачах линейного программирования – задаче о назначениях, задачах целочисленного программирования;
- приобрести умения и навыки создания моделей специальных задач линейного программирования – задаче о назначениях, задачах целочисленного программирования.

4.2 краткое содержание вопроса № 4.2. Математические модели задач оптимизации на графах и сетях в биологии: задача об остове минимального веса сети и её решение по алгоритму Краскала.

Задание для работы:

1. Формулировка задачи оптимизации на графах и сетях в биологии: задача об остове минимального веса сети.
2. Алгоритм Краскала решения задачи об остове минимального веса сети.
3. Создание математической модели задачи.

1. Формулировка задачи оптимизации на графах и сетях в биологии: задача об остове минимального веса сети.

Пусть  $G$  – связный нагруженный граф (сеть). Задача построения *минимального оствового дерева* заключается в том, чтобы из множества оствовых деревьев найти такое, у которого сумма длин ребер минимальна.

Приведем типичный случай, когда возникает необходимость построения минимального оствового дерева графа.

Биологическая система представляет собой сеть узлов, соединённых каналами связи (дугами), имеющими веса (уровень взаимодействия). Необходимо выделить в этой системе минимальную устойчивую подсистему, являющуюся оствовым деревом минимального веса.

## 2. Алгоритм Краскала решения задачи об оствове минимального веса сети.

Задачу построения минимального дерева-остова можно решить с помощью алгоритма Краскала. Приведем описание алгоритма по шагам.

*Шаг 1.* Отсортируем ребра графа по не убыванию весов.

*Шаг 2.* Полагаем, что каждая вершина относится к своей компоненте связности.

*Шаг 3.* Проходим ребра в «отсортированном» порядке. Для каждого ребра выполняем следующую проверку:

а) если вершины, соединяемые данным ребром, лежат в разных компонентах связности, то объединяем эти компоненты в одну, а рассматриваемое ребро добавляем к минимальному дереву-остову;

б) если вершины, соединяемые данным ребром, лежат в одной компоненте связности, то исключаем ребро из рассмотрения, так как при включении данного ребра образуется цикл.

*Шаг 4.* Если есть еще нерассмотренные ребра и не все компоненты связности объединены в одну, то переходим к шагу 3, иначе алгоритм завершает работу:

а) если при этом просмотрены все ребра, но не все компоненты связности объединены в одну, то для исходного графа невозможно построить покрывающее дерево;

б) если просмотрены все ребра, и все компоненты связности объединены в одну, то для исходного графа построено минимальное покрывающее дерево.

Задача. Граф  $G$  содержит 5 вершин. Расстояния между вершинами заданы таблицей

1. Найти его минимальное дерево-остов (минимальное покрывающее дерево).

Таблица 1

	1	2	3	4	5
1	0	5	8	2	7
2	5	0	9	2	5
3	8	9	0	10	10
4	2	2	10	0	7
5	7	5	10	7	0

*Решение.* Так как матрица симметрична, можно рассматривать, например, только элементы, расположенные выше или ниже главной диагонали. Упорядочим рёбра по длине.

Исходные  
данные

Упорядоченные по длине ребра

ребро	длина
(1,2)	5
(1,3)	8
(1,4)	2
(1,5)	7
(2,3)	9
(2,4)	2
(2,5)	5
(3,4)	10
(3,5)	10
(4,5)	7

ребро	длина	Включить в дерево-остов
(1,4)	2	Да
(2,4)	2	Да
(1,2)	5	Нет
(2,5)	5	Да
(1,5)	7	Нет
(4,5)	7	Нет
(1,3)	8	Да
(2,3)	9	Нет
(3,4)	10	Нет
(3,5)	10	Нет

1. Включаем в дерево-остов ребро (1, 4). Множество вершин, включенных в дерево-остов  $V = \{1, 4\}$  (рис. 21).
2. Следующим кандидатом на включение в дерево-остов является ребро (2, 4). Добавление вершины 2 к множеству  $V$  и ребра (2, 4) к дереву не создает цикла, так как вершина 2 не входит в множество  $V$ . После добавления ребра (2, 4) к дереву множество вершин, включенных в дерево-остов  $V = \{1, 2, 4\}$  (рис. 22).
3. Добавление ребра (1, 2) приведет к образованию цикла (рис. 23). Поэтому не включаем это ребро в дерево-остов.
4. Следующим кандидатом на включение в дерево-остов является ребро (2, 5). Его включение не создает цикла, поэтому  $V = \{1, 2, 4, 5\}$  (рис. 24).
5. Добавление ребра (1, 5) приведет к образованию цикла (рис. 25). Поэтому не включаем это ребро в дерево-остов.
6. Следующим кандидатом на включение в дерево-остов является ребро (4, 5). Однако его добавление приведет к образованию цикла. Поэтому не включаем это ребро в дерево-остов.

7. Включение ребра  $(1, 3)$  не создает цикла, поэтому  $V = \{1, 2, 3, 4, 5\}$  (рис. 26).
8. Так как все вершины графа вошли в дерево, то получено покрывающее дерево с минимальным весом, равным 17.

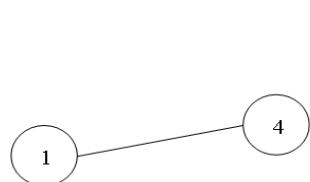


Рисунок 1

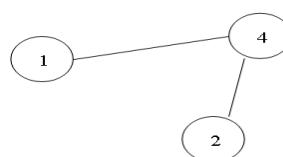


Рисунок 2

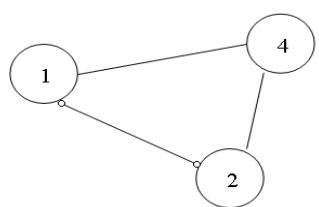


Рисунок 3

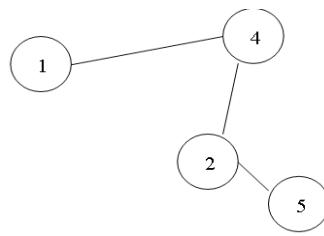


Рисунок 4

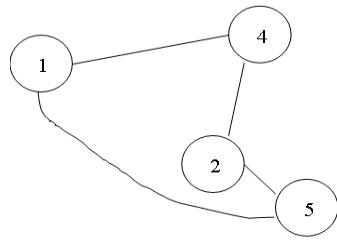


Рисунок 5

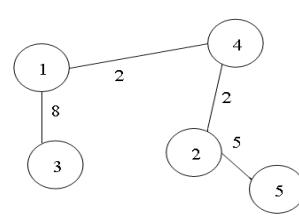


Рисунок 6

3. Создание математической модели задачи.

*Задача.* Найти минимальное дерево-остов для графа, приведенного на рис. 7.

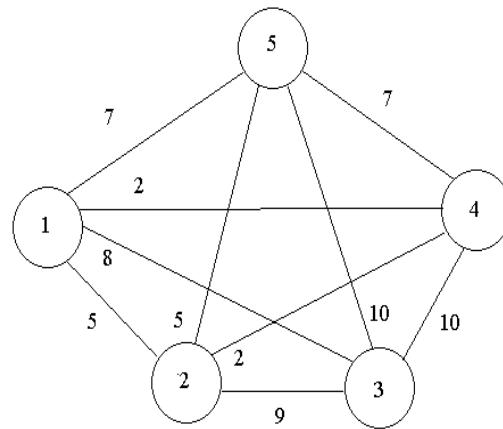


Рисунок 7

Пусть переменная  $X_{ij}$  означает наличие ребра вида  $(i, j)$  в покрывающем дереве, а  $a_{ij}$  интерпретируется как длина указанного ребра.

*Решение.* В заданном графе 5 вершин и 10 ребер. Следовательно, переменными математической модели этой индивидуальной задачи о построении минимального дерева-остова являются 10 переменных

$$x_{12}, x_{13}, x_{14}, x_{15}, x_{23}, x_{24}, x_{25}, x_{34}, x_{35}, x_{45}.$$

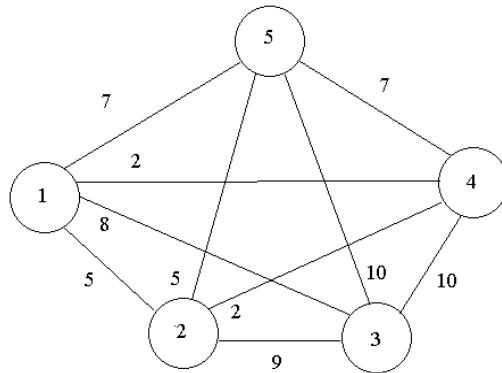


Рисунок 8

Математическая постановка (математическая модель) такой индивидуальной задачи приведена ниже.

$$5x_{12} + 8x_{13} + 2x_{14} + 7x_{15} + 9x_{23} + 2x_{24} + 5x_{25} + 10x_{34} + \\ + 10x_{35} + 7x_{45} \rightarrow \min_{x \in \Delta_\beta}$$

Множество ограничений выглядит так:

$$\left\{ \begin{array}{l} x_{12} + x_{13} + x_{14} + x_{15} \geq 1; \\ x_{12} + x_{23} + x_{24} + x_{25} \geq 1; \\ x_{13} + x_{23} + x_{34} + x_{35} \geq 1; \\ x_{14} + x_{24} + x_{34} + x_{45} \geq 1; \\ x_{15} + x_{25} + x_{35} + x_{45} \geq 1; \\ x_{12} + x_{13} + x_{14} + x_{15} + x_{23} + x_{24} + x_{25} + x_{34} + x_{35} + x_{45} = 4; \\ x_{12}, x_{13}, x_{14}, x_{15}, x_{23}, x_{24}, x_{25}, x_{34}, x_{35}, x_{45} \in \{0, 1\}. \end{array} \right.$$

В результате рассмотрения вопроса № 4.2 студенты должны:

- освоить понятие о задачах (моделях) оптимизации на графах и сетях в биологии, алгоритмы их решения;
- приобрести умения и навыки создания моделей оптимизации на графах и сетях в биологии, использования алгоритмов их решения.

## 1.2. Тема 2: «Компьютерные технологии в биологии. Базы и банки данных.

### 1.2.1 Перечень и краткое содержание рассматриваемых вопросов:

#### 1. Наименование вопроса № 1.

Понятие о биоинформатике: цели и задачи биоинформатики, основные направления биоинформатики. История становления биоинформатики. Особенности биоинформа-

ционных данных, генетическая информация. Информационные и компьютерные технологии в биоинформатике. Интернет-технологии в биоинформатике. Биоинформационные данные, сети и базы Применение биоинформатики. Вклад советских и российских учёных в биоинформатику, А. А. Ляпунов.

Краткое содержание вопроса №1.

Задание для работы:

1. Понятие о биоинформатике: цели и задачи биоинформатики, основные направления биоинформатики, история становления биоинформатики:

- понятие о биоинформатике,
- цели и задачи биоинформатики,
- основные направления биоинформатики,
- история становления биоинформатики.

2. Особенности биоинформационных данных, генетическая информация:

- выявить особенности биоинформационных данных;
- составить на примерах представление о генетической информации;
- привести примеры и отсортировать области применение биоинформатики.

3. Информационные и компьютерные технологии в биоинформатике. Интернет-технологии в биоинформатике. Биоинформационные данные, сети и базы. Вклад советских и российских учёных в биоинформатику, А. А. Ляпунов.

1). Информационные и компьютерные технологии в биоинформатике:

- информационные и компьютерные технологии в биоинформатике:
- аппаратные средства;
- программное обеспечение, системное и прикладное;
- языки программирования, сценарии и разметки;
- операционные системы:

2). Информационные и компьютерные технологии в биоинформатике:

- выявить роль единых протоколов связи;
- выявить способы доступа к базам данных через интернет;
- выявить назначение IP-адресации;
- иерархическая система имён доменов;
- поиск веб-страниц биоинформатики и молекулярной биологии;
- веб-узлы биоинформатики и молекулярной биологии,
- веб-серверы биоинформатики и молекулярной биологии,

3). Биоинформационные данные, сети и базы:

- способ кодирования (представления, хранения, обработки) информации о последовательностях ДНК последовательностями 8 битовых слов в двоичном коде;
- кодирование информации о белковых последовательностях в двоичном коде;
- стандартные форматы для кодирования и записи в текстовые файлы информации о последовательностях ДНК или белка: ASCII (Американский стандартный код для обмена информацией) и FASTA (быстрое выравнивание);
- EMBnet, европейская сеть предоставления биоинформационных и образовательных услуг сотрудникам лабораторий разных государств Европы, через специально выделенные узлы, работающие на местных языках;
- SRS, сетевой браузер для выборки данных молекулярной биологии (система выборки последовательностей SRS) из баз данных европейской сети EMBnet;
- NCBI, Национальный центр биотехнологической информации США;
- Entrez, сетевой браузер для выборки данных молекулярной биологии из баз данных, объединённых в сети NCBI.

4). Примеры конкретных программных комплексов биоинформатики.

5). Вклад советских и российских учёных в биоинформатику, А. А. Ляпунов. Состояние биоинформатики в России. Центр Коллективного Пользования по биоинформатике в СО РАН.

## **1. Понятие о биоинформатике: цели и задачи биоинформатики, основные направления биоинформатики, история становления биоинформатики.**

Биоинформатика — это область науки, разрабатывающая и применяющая вычислительные алгоритмы для систематизации и анализа генетической информации с целью определения молекулярных основ биологических процессов с последующим использованием этих знаний на практике. Ее основная задача — разработка вычислительных алгоритмов для анализа и систематизации данных о структуре и функциях биологических молекул, прежде всего нуклеиновых кислот и белков. Объем генетической информации, накапливаемой в банках данных, начал увеличиваться с возрастающей скоростью после того, как были разработаны быстрые методы секвенирования (расшифровки нуклеотидных последовательностей ДНК). Биоинформационные методы позволяют не просто обрабатывать этот огромный массив данных, но и выявлять закономерности, которые не всегда можно заметить при обычном эксперименте, предсказывать функции генов и зашифрованных в них белков, строить модели взаимодействия генов в клетке, конструировать лекарства.

Биоинформатика как наука появилась на стыке молекулярной биологии, генетики,

математики и компьютерных технологий и обязана своим появлением накоплению обширных экспериментальных данных, особенно заметно это накопление стало проходить начиная с 70-х гг. прошлого века. Главной проблемой, потребовавшей нового подхода, стал анализ последовательностей генов и белков, которые стало возможно «прочитать» в эксперименте. Появление данных о пространственной структуре макромолекул также требовало новых методов работы с ними. Информации, получаемой в эксперименте, было больше, чем возможностей человека к запоминанию фактов и их анализу.

### *Цели и задачи биоинформатики*

В настоящее время задачи биоинформатики — это почти исключительно задачи молекулярной биологии. Причина этого в том, что за последние 20–25 лет накоплен поистине колоссальный экспериментальный материал именно о строении и функционировании биологических молекул (белков и нуклеиновых кислот), в качестве примера достаточно привести геном человека. Этот материал требует развитых компьютерных методов для своего анализа.

Основными разделами биоинформатики являются компьютерная геномика, решающая проблемы расшифровки генетических «текстов», хранящихся в последовательностях нуклеотидов ДНК (РНК), и метабономика, исследующая организацию метаболизма клетки и его управления со стороны генома. Важное значение для развития основных разделов биоинформатики имеет создание компьютерных баз данных по молекулярной биологии, обеспечивающих геномику и метабономику необходимыми для их развития экспериментальными данными в достаточном количестве и удобной для использования форме. Среди основных задач биоинформатики — описание генных сетей, изобретение новых лекарств с заданными свойствами, разработка компьютерных моделей процессов, происходящих в организме.

К биоинформатике часто относят также информационные службы, обеспечивающую накопление, хранение и использование приобретаемых наукой знаний о биологических системах. Целью биоинформатики является как накопление биологических знаний в форме, обеспечивающей их наиболее эффективное использование, так и построение и анализ математических моделей биологических систем и их элементов. Информация о строении материальных элементов, обеспечивающих функционирование организма, хранится в последовательности нуклеотидов ДНК (или РНК), образующей его геном. Установление нуклеотидных последовательностей ДНК геномов организмов (секвенирование) стало к началу XXI в. хорошо освоенной и, видимо, достаточно рентабельной технологией. Количество секвенированных геномов быстро увеличивается и определяется в основном только объемом средств, которые можно затратить на эти цели. По сути биоинформа-

тика включает в себя три компонента: создание баз данных, позволяющих осуществлять хранение крупных наборов биологически данных и управлять ими; разработку алгоритмов и методов статистического анализа для определения отношений между элементами крупных наборов данных; использование этих средств для анализа и интерпретации биологических данных различного типа — в частности, последовательностей ДНК, РНК и белков, белковых структур, профилей экспрессии генов, биохимических путей.

Цели биоинформатики можно свести к следующим:

1. Организовать данные таким образом, чтобы исследователи имели доступ к текущей информации, хранящейся в базах данных, и смогли вносить в нее новые записи по мере получения новых сведений.
2. Развивать программные средства и информационные ресурсы, которые помогают в управлении данными и в их анализе.
3. Применять эти средства для анализа данных и интерпретации полученных результатов таким образом, чтобы они имели биологический смысл.

В целом задачи биоинформатики состоят в анализе информации, закодированной в биологических последовательностях, что предполагает:

- обнаружение генов в последовательности ДНК различных организмов;
- развитие методов изучения структуры и (или) функции новых расшифрованных последовательностей и соответствующих структурных областей ДНК;
- определение семейств родственных последовательностей и построение моделей;
- выравнивание последовательностей и восстановление филогенетических деревьев с целью выявления эволюционных связей;
- обнаружение мишеней для медикаментозного воздействия и отыскание перспективных опытных соединений.

#### *Основные направления биоинформатики*

Есть несколько основных направлений этого раздела науки, в зависимости от исследуемых объектов:

- биоинформатика последовательностей;
- структурная биоинформатика;
- компьютерная геномика.

Хотя биоинформатике как науке всего лишь около 30 лет, в ней уже существуют свои традиционные направления: компьютерный анализ ДНК, РНК и белковых последовательностей, реконструкция пространственных структур биополимеров, теоретический и компьютерный анализ структурно-функциональной организации геномов и белков. Сюда

же относится и развитие баз данных по молекулярно-генетической тематике, структурированию экспериментальных данных. Биоинформатика по своей сути – это интегративная наука, являющаяся инструментом в руках тех, кто занимается биологией, молекулярной биологией, молекулярной генетикой, медицинской генетикой, фармакологией, а также биохимией, биофизикой и т. д.

**2. Выявить особенности биоинформационных данных; составить на примерах представление о генетической информации; привести примеры и отсортировать области применения биоинформатики.**

Данные в биологии, биоинформатике

- стали не описательными, а аналитическими (количественными), невероятно точными, дискретными;
- имеют вторую особенность – их огромное количество (банки данных нуклеотидных последовательностей содержат 16 млрд нуклеиновых пар оснований);
- позволяют увидеть картину мира живых существ чётко и целиком;
- позволяют изучать организмы во времени: назад в прошлое, вперёд в будущее;
- применяются в других областях (медицина, с/х и т. д.).

*Генетическая информация*

Носителями генетической информации являются молекулы ДНК. Генетическая информация – совокупность генов, регулирующая целенаправленную деятельность любой живой клетки, определяется не самими основаниями ДНК, а последовательностью их расположения. Изучение невозможно без компьютерного анализа.

**3. Информационные и компьютерные технологии в биоинформатике. Интернет-технологии в биоинформатике. Биоинформационные данные, сети и базы. Вклад советских и российских учёных в биоинформатику, А. А. Ляпунов.**

В системной биологии важнейшую роль играют методы биоинформатики. Они дают возможность с помощью компьютеров

- накапливать и интегрировать в базы данных экспериментальную информацию;
- осуществлять ее компьютерный анализ;
- проводить математическое моделирование структурно-функциональной организации живых систем;
- предсказывать новые свойства живых систем;
- на этой основе планировать новые этапы экспериментальных исследований.

Приведём конкретные примеры баз и банков данных, базовых пакетов, программных средств для полного анализа макромолекул, а также других программных продуктов

по биоинформатике с их адресами в Интернете, разработанных как отечественными, так и зарубежными учеными.

GenBank – <http://www.ncbi.nlm.nih.gov/>

GenBank Банк данных по нуклеотидным последовательностям (3400000000 пар оснований в 461000 последовательностей).

SWISS-PROT – <http://www.expasy.ch/sprot/sprot-top.html> Аннотированный банк данных по аминокислотным последовательностям белков.

PIR – <http://www.ncbi.nlm.nih.gov/PIR/PIRSearch.html> Аннотированный банк данных по аминокислотным последовательностям белков, организованных в соответствии с гомологией и таксономией.

PBD – <http://www.rcsb.org/pdb/> Банк данных по 3D структуре биологических макромолекул.

NDB – <http://ndbserver.rutgers.edu> Банк данных по нуклеиновым кислотам. Включает структуры ДНК и РНК вместе с их 3-хмерными изображениями.

ProDom – <http://protein.toulouse.inra.fr/prodom.html> Банк данных по доменам белков. Web-серверы, предоставляющие пользователю генетическую информацию, оснащены комплексом программных средств для поиска информации в банках данных и анализа нуклеотидных и аминокислотных последовательностей. В качестве запросов при поиске последовательностей в банках данных могут использоваться номенклатурные названия генов, организмов, ключевые слова и др.

В качестве примера предложим программу Auto Dok, которая является программой для автоматического докинга. С ее помощью можно посмотреть, как молекулы лекарств или кандидатов на роль лекарств взаимодействуют в известной 3D-структуре. В частности, программа применяется для разработки лекарств, специфически связывающихся с тем или иным белком.

Здесь же приведем примеры основных программ сравнения аминокислотных и нуклеотидных последовательностей.

ACT – (Artemis Comparison Tool) – геномный анализ.

Arlequin – анализ популяционно-генетических данных.

Bio Edit – редактор множественного выравнивания аминокислотных и нуклеотидных последовательностей.

Numerics – коммерческий универсальный пакет программ по биоинформатике; BLAST – поиск родственных последовательностей в базе данных аминокислотных и нуклеотидных последовательностей.

ClustalW – множественное выравнивание аминокислотных и нуклеотидных после-

довательностей;

FASTA – набор алгоритмов определения схожести аминокислотных и нуклеотидных последовательностей;

Mesquite – программа для сравнительной биологии на языке Java.

Muscle – множественное сравнение аминокислотных и нуклеотидных последовательностей. Более быстрая и точная программа в сравнении ClustalW;

Pop Gene – анализ генетического разнообразия популяций; Populations – популяционно-генетический анализ.

Примером интегрированного инструмента биолога является также Unipro UGENE. Это свободно распространяемое программное обеспечение для работы молекулярного биолога. Пользовательский интерфейс этого продукта обеспечивает:

- простую и удобную работу с последовательностями; – визуализацию хроматограмм;
- использование редактора множественного выравнивания последовательностей;
- просмотр трехмерных моделей PDB и MMDB с поддержкой стереорежима;
- просмотр филогенетических деревьев;
- применение конструктора вычислительных схем, автоматизирующего процесс анализа;
- поддержку сохранения изображений в векторные форматы для удобства публикаций.

Ряд оригинальных компьютерных программ, баз и банков данных, созданных российскими учеными, можно также найти по разным поисковикам на других многочисленных сайтах по биоинформатике.

Отметим ряд институтов, где в России ведутся работы по биоинформатике. В Москве организован Институт математических проблем биологии РАН. Там занимаются решением разных биологических проблем при поддержке информационных технологий, таких как структурная и сравнительная геномика, основные молекулярно-биологические механизмы, молекулярно-генетические системы управления, протомиля, метаболита, базы биологических данных, математическое обеспечение биологических экспериментов.

При Институте проблем передачи информации РАН (ИППИ РАН) в 2003 г. основан учебно-научный центр «Биоинформатика». Также к научным институтам такой тематики относятся следующие учреждения: ВНИИ «Генетика» (Москва); Институт белка РАН (Пущино); Институт биоорганической химии РАН (Москва); Институт молекулярной биологии РАН (Москва); Институт биомедицинской химии РАМН (Москва);

Институт цитологии и генетики СО РАН (Новосибирск) и ряд других.

Web-сайты имеются у следующих институтов: Институт биомедицинской химии РАМН (Москва) – <http://www.ibmh.msk.su/bioinform>

- Институт цитологии и генетики СО РАН (Новосибирск) – <http://www.mgs.bionet.nsc.ru> и другие.

В СО РАН (Новосибирский Академгородок) существует центр для разработки экспериментальных технологий в области биоинформатики. Для биологов НГУ и СО РАН создали единую информационную базу по биоинформатике. Совместными усилиями нескольких институтов СО РАН и Новосибирского государственного университета открыт Центр Коллективного Пользования (ЦКП) «БИОИНФОРМАТИКА» Суперкомпьютер центра один из самых быстрых в мире. Его производительность 16,5 терафлопс (16,5 триллионов операций в секунду), по мощности он не уступают аналогичным, установленным в центре Vital-IT в Лозанне (Швейцария), система хранения данных рассчитана на 48 терабайт. Центр «Биоинформатика», по сути, – это огромная вычислительная система, суперкомпьютер, с помощью которого можно моделировать лекарства от рака и СПИДа, расшифровывать геномы растений и живых организмов. Прочитать миллиарды и миллионы букв в геномах человека, пшеницы, вируса гепатита или энцефалита сегодня позволяет ДНК-секвенатор. В будущем ЦКП сможет помочь и специалистам смежных областей науки. Например, ученые планируют сравнить геном древнего гоминида из алтайской пещеры с геномом современного человека В ЦКП более 400 пользователей, это около 20 академических институтов СО РАН. Есть и учебные институты, и фирмы. и др.

Цель данного центра – создать сетевую инфраструктуру, которая объединит в технологический комплекс экспериментальные установки с вычислительными комплексами и хранилищами данных. Задачами самой структуры являются биоинформационная поддержка научно-исследовательских работ, подготовка специалистов в области компьютерной системной биологии и биоинформатики, разработка новых экспериментально-теоретических технологий. Активное использование вычислительных мощностей ЦКП предполагается в рамках проекта СО РАН «Геномика, протеомика, биоинформатика». Организациями-учредителями проекта стали Институт цитологии и генетики СО РАН, Институт вычислительной математики и математической геофизики СО РАН, Международный Томографический центр СО РАН, Институт химической биологии и фундаментальной медицины СО РАН, Институт математики СО РАН, Институт вычислительных технологий СО РАН и Новосибирский государственный университет. С 1998 г. в Институте цитологии и генетики СО РАН регулярно проводятся Международные конференции по биоинформатике регуляции и структуры геномов и системной биологии (Bioinformatics of

Genome Regulation and Structure System Biology – BGRS/SB).

Следует также отметить и заслуги сибирских ученых в создании новой науки. К ним относится А. А. Ляпунов, выдающийся русский ученыи математик. В 2011 г. отмечалось 100 лет со дня его рождения. Последние годы своей жизни он жил и работал в Академгородке. Надо сказать, что, будучи «отцом отечественной кибернетики», он заинтересовался биологией еще в 30-е годы прошлого века. Позднее, в 50-е годы, он снова обратился к биологии. Сложность биологических систем и процессов их эволюции привлекала его как объект приложения методов исследований, характерных для дескриптивной теории множеств. Таким образом, он стоял у истоков создания новой науки на стыке математики и биологии.

В результате рассмотрения вопроса № 1 студенты должны

- освоить понятие о биоинформатике,
- составить представление о целях и задачах биоинформатики; основных направлениях биоинформатики, истории становления биоинформатики, особенностях биоинформационных данных, генетической информации и областях применение биоинформатики; о вкладе советских и российских учёных в биоинформатику (А. А. Ляпунов);
- приобрести умения и навыки работы с литературой и интернет-ресурсами по темам, связанным с основными направлениями биоинформатики, особенностями биоинформационных данных, генетической информацией и областях применение биоинформатики.

2. Наименование вопроса № 2. Углубленное изучение возможностей современных офисных технологий для использования в профессиональной деятельности, научных исследованиях в биологии.

2.1 Наименование вопроса № 2.1 Углубленное изучение возможностей современных офисных технологий для использования в профессиональной деятельности, научных исследованиях в биологии. Работа с пакетом Microsoft Office: таблицы в MS Word, форматирование таблиц, использование формул в таблицах, стандартные функции.

2.2 Наименование вопроса № 2.2. Углубленное изучение возможностей современных офисных технологий для использования в профессиональной деятельности, научных исследованиях в биологии. Работа с пакетом Microsoft Office: таблицы MS Excel, ввод и редактирование данных, вставка формул, стандартные функции.

2.3 Наименование вопроса № 2.3. Углубленное изучение возможностей современных офисных технологий для использования в профессиональной деятельности, научных исследованиях в биологии. Работа с пакетом Microsoft Office: MS Access, создание баз данных, создание межтабличных связей, сортировка и фильтрация данных; запросы и отчёты.

2.1 Краткое содержание вопроса № 2.1 Углубленное изучение возможностей со-

временных офисных технологий для использования в профессиональной деятельности, научных исследованиях в биологии. Работа с пакетом Microsoft Office: таблицы в MS Word, форматирование таблиц, использование формул в таблицах, стандартные функции.

Задание для работы:

1. Ознакомиться с краткими теоретическими сведениями (даны в описании работы).
2. Выполнить практическую работу:
  - работа с таблицами в MS Word,
  - форматирование таблиц
  - использование формул в таблицах,
  - стандартные функции.
3. Ответить на контрольные вопросы.

### **Краткие теоретические сведения**

Для создания таблицы можно использовать кнопку **Таблица–Добавить таблицу** на панели инструментов или команду **Таблица + Вставить + Таблица**.

Для выделения таблицы используется команда **Таблица + Выделить + Таблица** или кнопка в левом верхнем углу таблицы.

Для вставки элементов таблицы их необходимо выделить и воспользоваться командой **Таблица + Добавить + ...**, а для удаления элементов - **Таблица + Удалить + ....**

Для форматирования таблицы можно:

- Объединять ячейки - **Таблица + объединить ячейки**, разбивать ячейки - **Таблица + разбить ячейки**;
- Изменение ширины столбцов и высоты строк - **Таблица + Свойства таблицы**;
- Преобразование текста в таблицу - **Таблица + Преобразовать + Текст в таблицу**;
- Преобразование таблицы в текст - **Таблица + Преобразовать + Таблицу в текст**;
- Наложение (выбор) автоформата на таблицу - **Таблица + Автоформат**;
- Выравнивание высоты строк и ширины столбцов - **Таблица + Автоподбор+ +Выровнять ширину столбцов или выровнять высоту строк**;
- Сортировка элементов таблицы - **Таблица + Сортировка**;
- Вставка формулы в таблицу для подсчета значений - **Таблица + Формула**.

Для пересчета значений формул необходимо выделить таблицу и нажать функциональную клавишу F9.

Кроме стандартных элементов диалога в окне содержатся еще два компонента: **Диаграмма** и **Таблица данных**. Состав команд главного меню, естественно, ориентирован на работу по редактированию и настройке диаграмм.

Панель инструментов **Форматирование** также появляется при запуске MS Word. Панель содержит кнопки, позволяющие форматировать графические объекты, содержимое ячеек данных и объекты диаграммы.

Окно с диаграммой позволяет оперативно оценивать тот вид, который диаграмма будет иметь после закрытия диалога. Вид диаграммы зависит от содержимого таблицы данных и выполненных настроек формата диаграммы. В составе диаграммы наиболее важными являются следующие элементы – **Ось значений, Ось категорий, Ось ряда данных**.

**ных, Основные линии сетки оси значений, Легенда.** Кроме того, в составе диаграммы имеются элементы, которые позволяют улучшить ее внешний вид, например, стены и основание.

Данные на диаграмме отображаются с помощью рядов данных, отображаемых в виде наборов линий, столбцов, точек (в плоских диаграммах) или секторов (в круговых диаграммах).

Для изменения отображаемых в диаграмме данных соответствующие изменения нужно ввести в таблицу данных. При изменении формата чисел в таблице данных меняется формат соответствующих подписей данных на диаграмме.

Управление составом элементов диаграммы выполняется с помощью вкладок диалогового окна **Параметры диаграммы**, вызываемого одноименной командой меню **Диаграмма**. С помощью вкладок названного окна можно задать и отменить отображение легенды, сетки диаграммы, названий, осей диаграммы, таблицы данных.

**Задание:**

1. Запустите программу Microsoft Word.
2. Добавьте таблицу из 4 строк и 6 столбцов., щелкнув кнопку **Добавить таблицу** на панели инструментов.
3. Введите в таблицу текст:

Специальности	1-й курс	2-й курс	3-й курс	4-й курс.	Всего
Биология	200	180	160	140	
ВСЭ	170	150	130	110	
Иняз	80	70	60	50	

4. Установите курсор в столбец «Всего» по строке Биология и введите команду **Таблица + Формула**. В диалоговом окне **Формула** введите формулу **=SUM(LEFT)**. Повторите эти действия для 3 и 4 строки.
5. Установите для названий специальностей полужирный шрифт, а для названия таблицы цвет текста – синий. Выделите все столбцы, начиная со второго и установите выравнивание по центру.
6. Выделите всю таблицу и выполните команду **Формат + Границы и заливка**. В диалоговом окне установите тип обрамления **Сетка**, толщину линии - **1,5 пт**.
7. Выделите первую строку текста выполните команду **Формат + Границы и заливка**. Во вкладке **Заливка** установите тип узора заливки ячейки - **25%**.
8. Выделите названия специальностей и выполните команду **Таблица + Сортировка**. Установите сортировку для 1 столбца по возрастанию.
9. Создайте следующую таблицу в виде табулированного текста (переход между ячейками таблицы осуществляется клавишей **Tab**, а переход на новую строку таблицы – клавишей **Enter**):

	Дисциплина	Группа	Сред. балл	Всего сдавало	Отл	Хор	Удовл	Неуд	Неяв
1.	Биология	11а	3,88	32	12	10	6	3	1
2.	Биология	11б	3,52	27	7	9	6	3	2
3.	ВСЭ	21а	3,43	28	9	8	3	5	3

4.	ВСЭ	21б	3,52	29	8	8	8	3	2
	ИТОГО		3,59	116	36	35	23	14	8
1.	Иняз	11а	3,75	32	8	12	10	1	1
2.	Иняз	11б	3,75	32	12	9	6	3	2
3.	Иняз	21а	3,58	31	12	8	3	5	3
4.	Иняз	21б	3,46	28	7	8	8	3	2
	ИТОГО		3,64	123	39	37	27	12	8
	Средние показатели		3,56	239	75	72	50	26	16

- Установите курсор в столбец **Средний балл** по строке **Итого** и введите команду **Таблица + Формула**. В диалоговом окне **Формула** введите формулу **=Average()**. Повторите эти действия по второй строке **Итого** и по строке **Средние показатели**.
- С помощью команды **Таблица + Формула** просчитайте **Итого** и **Средние показатели** для столбца **Всего сдавало**. В диалоговом окне **Формула** введите формулу **=SUM(Above)**.
- Повторите эти действия для столбцов **Отлично, Хорошо, Удовл., Неуд., Неявки**.
- Преобразуйте набранный текст в таблицу.
- Добавьте в таблицу верхнюю строку для заголовка.
- Объедините ячейки в верхней строке и напишите «Сведения об успеваемости студентов факультета биотехнологий за 2019/2020 учебный год»
- Шрифт в таблице установите 14 пт.
- Для шапки таблицы установите полужирный шрифт и выравнивание по центру.
- Задайте ширину второго столбца – 3,5 см.
- Задайте ширину столбцов 6-10 — 1,8 см.
- Выделите более жирным границы строк ИТОГО как показано в задании.
- Задайте заливку красным цветом для строки "Средние показатели".
- Скопируйте данную таблицу еще раз в конец документа, для первой таблицы задайте Автоформат – Цветной 1, а для второй – Изысканная таблица.
- Вторую таблицу преобразуйте в текст с разделителем \_!
- Создайте следующую таблицу:

#### Отчет по микробиологическим анализам

	январь	февраль	март	апрель	май	июнь	Всего
Объект 1	10	11	9	7	5	14	
Объект 2	18	10	15	17	21	15	
Объект 3	12	18	14	16	17	21	
Объект 4	30	44	26	28	20	19	
Объект 5	25	21	19	12	14	9	
Итого							
Максимальное число микробиологических анализов объекта 1							
Минимальное число микробиологических анализов объекта 4							

- Просчитайте значения в пустых ячейках таблицы используя вставку формул (**Всего, Максимальное, Минимальное**).
- Измените несколько значений в таблице и пересчитайте значение формул.

27. Сохраните файл в папке с именем вашей группы под названием **Таблицы ФИ**. (Вместо ФИ укажите полностью свою фамилию и имя).

### **Контрольные вопросы:**

1. Как вставить таблицу в текст?
2. Какие существуют действия для преобразования таблицы?
3. Как выделяются элементы в таблице?
4. Какие свойства таблицы можно изменить?
5. Как необходимо набрать текст для дальнейшего преобразования его в таблицу?
6. Как преобразовать таблицу в текст?
7. Какие типы автоформатов таблиц Вы знаете?
8. Как произвести сортировку элементов таблицы?
9. Как производится ссылка на ячейки таблицы при вводе формулы?

В результате рассмотрения вопроса № 2.1 студенты должны:

- освоить использование формул в таблицах, стандартные функции при создании и форматировании таблиц и оформлении документов Microsoft Word;
- приобрести умения и навыки использования формул в таблицах, стандартные функции при создании и форматировании таблиц и оформлении документов Microsoft Word.

2.2 Краткое содержание вопроса № 2.2. Углубленное изучение возможностей современных офисных технологий для использования в профессиональной деятельности, научных исследованиях в биологии. Работа с пакетом Microsoft Office: таблицы MS Excel, ввод и редактирование данных, вставка формул, стандартные функции.

Задание для работы:

1. Ознакомиться с краткими теоретическими сведениями (даны в описании работы).
2. Выполнить практическую работу:
  - работа с таблицами MS Excel,
  - ввод и редактирование данных,
  - вставка формул в таблицы,
  - стандартные функции.
3. Ответить на контрольные вопросы.

### **Краткие теоретические сведения**

Для вставки нового листа необходимо выделить лист, перед которым надо вставить новый лист и Выбрать Вставка + Лист.

Для удаления листа необходимо его выделить и выбрать Правка + Удалить либо щелкнуть правой кнопкой мыши на ярлыке листа и выбрать команду «Удалить».

**Маркер заполнения** — это небольшой черный квадрат в углу выделенного диапазона. Попав на маркер заполнения, указатель мыши принимает вид черного креста.

**Автозаполнение** — это функция, которая помогает быстрее вводить данные. При перетаскивании маркера заполнения ячейки может происходить не только копирование одних и тех же значений. Например, введите в какую-либо ячейку число 1, а в соседнюю

справа от нее - число 2. Затем отметьте обе ячейки и протяните маркер заполнения вдоль строки. Появится ряд значений: 1,2,3,4...

Диапазонам и ячейкам можно присваивать имена. Необходимо выделить какой-либо диапазон или ячейку, а затем выбрать Вставка, Имя, Присвоить и в появившемся диалоговом окне ввести имя. После этого можно выделять необходимый диапазон по его имени и использовать это имя в ссылках и формулах.

Чтобы выделить диапазон по его имени, надо щелкнуть стрелку вниз в поле «Имя» в строке формул и выбрать имя диапазона из списка.

Обратите внимание, что в диалоговом окне «Имя» адреса ячеек или диапазонов появляются со знаком \$. Это так называемые абсолютные ссылки, которые не изменяются в процессе различных операций Excel.

**Вставка строк и столбцов** выбрать Вставка + Строки или Вставка + Столбцы. Вставленные строки или столбцы будут иметь то же форматирование, что и выделенные до этого ячейки.

### **Автосуммирование и Автovычисления**

Так как одной из наиболее часто употребляющихся функций является СУММ, в Excel предусмотрен быстрый способ ее ввода:

1. Выделить ту ячейку, в которую необходимо вставить сумму. Лучше, если она расположена в конце строки или столбца данных — это поможет Excel «догадаться», какие ячейки необходимо просуммировать.
2. Щелкнуть кнопку «Автосумма» на стандартной панели инструментов. В выделенную ячейку будет введена функция СУММ и адрес диапазона левее или выше ячейки.
3. Если выбранный Excel диапазон будет неверным, то необходимо исправить формулу «вручную» в строке формул, либо перетащить курсор мыши через необходимый диапазон.
4. Нажать клавишу Enter либо щелкнуть кнопку Enter в строке формул.

Можно просто дважды щелкнуть кнопку «Автосумма» и функция СУММ будет сразу вставлена в выделенную ячейку.

### **Работа с именами ячеек и диапазонов**

Формуле или константе можно присвоить имя, если выбрать Вставка, Имя, Присвоить. Величины и функции, имеющие имена (например, Реализация или Прибыль), удобно использовать в разных местах книги. Кроме этого, ввод сложных функций значительно сокращается, если их части имеют краткие имена.

### **Присваивание имени диапазону ячеек.**

1. Указать ячейку, либо выделить диапазон, либо выделить область несмежных ячеек.
2. Установить указатель на поле имен в левой части строки формул и нажать кнопку мыши.
3. Ввести имя.

Создание имен из заголовков строк и столбцов.

1. Выделить область, в которой следует присвоить имена строкам или столбцам. Выделенная область должна содержать строку или столбец заголовков.
2. Выбрать Вставка, Имя, Создать.

3. В появившемся окне в группе флажков «По тексту» указать расположение заголовков, из которых следует создать имена.

### **Вставка имени в формулу.**

Чтобы вставить имя в формулу, его предварительно следует присвоить. Если имя присвоено, то при вводе или исправлении формулы нужно выбрать Вставка, Имя, Вставить, а затем выбрать необходимое имя из списка.

### **Правка имен ячеек, диапазонов, формул и ссылок.**

1. Необходимо выбрать Вставка + Имя + Присвоить
2. Выбрать необходимый элемент из списка
3. Для правки ссылки исправить ее в поле «Формула»
4. Для правки имени надо ввести новое имя и нажать кнопку Добавить, после этого старое имя можно удалить.

При удалении имен следует иметь в виду, что листах книги могут появиться ошибки, если в ячейках используются удаленные имена.

Существует возможность создания собственных форматов чисел. Пользовательский числовой формат создается путем описания **шаблона** отображаемых данных, который включает 4 секции для описания форматов чисел, даты, времени и текста. Секции отделяются друг от друга при помощи символа - разделителя списка Windows. (Для русского языка это обычно точка с запятой. Проверить или изменить этот символ можно в панели управления - «Язык и стандарты»).

Для установки в ячейках необходимого обрамления (границ ячеек) и заливки цветом (затенения) можно использовать кнопки «Внешние границы» и «Выделение цветом» на панели форматирования (При этом необходимо предварительно выделить ячейки).

Для установки любых параметров границ ячеек и необходимо выбрать Формат, Ячейки и щелкнуть вкладку «Граница». Затем с помощью соответствующих кнопок установить необходимые параметры границ, в том числе тип линии, цвет и с каких сторон ячеек будут установлены границы. При этом в средней части окна можно увидеть, как будут выглядеть ячейки после изменения их границ.

Для установки любых параметров заливки надо выделить необходимые ячейки, а затем выбрать Формат, Ячейки и щелкнуть вкладку «Вид». На этой вкладке можно установить цвет фона и узор для выделенных ячеек. При этом в поле «Образец» можно увидеть, каким будет фон после изменения параметров.

### **Автоформат и кнопка «Формат по образцу»**

Автоформат предоставляет 16 форматов таблиц, которые можно применить к диапазону ячеек. Для применения автоформата необходимо:

1. Выделить диапазон ячеек.
2. Выбрать Формат, Автоформат. На экране появится окно «Автоформат», в левой части которого расположен список форматов, а в поле «Образец» виден внешний вид будущего формата.
3. Необходимо отметить нужный формат в списке.
4. Чтобы исключить из выбранного формата некоторые элементы, можно щелкнуть кнопку «Параметры».
5. Нажать Ok.

Для копирования форматов можно, вначале выделив необходимые ячейки, скопировать их в буфер обмена, а затем использовать команду Правка, Специальная вставка и в появившемся окне отметить флажок «Форматы».

Очень удобно для копирования форматов использовать кнопку «Формат по образцу» (в виде «кисточки») на панели стандартной инструментов:

1. Выделить ячейку (или ячейки) с форматом, который необходимо скопировать и вставить на новом месте.
2. Щелкнуть кнопку «Формат по образцу». Указатель мыши примет вид «кисточки» с расположенным рядом с ней знаком плюс.
3. Перетащить курсор мыши через ячейки, к которым необходимо применить скопированный формат.

Можно копировать формат одновременно в несколько мест. Для этого надо щелкнуть кнопку «Формат по образцу» дважды. После этого курсор мыши будет иметь вид кисточки до тех пор, пока не будет нажата клавиша ESC.

### **Условное форматирование**

Если необходимо выделить на рабочем листе какие-либо данные, имеющие определенные значения, то можно использовать условное форматирование:

1. Выделить ячейки, которые необходимо отформатировать.
2. Выбрать Формат, Условное форматирование.
3. В появившемся окне необходимо сформировать условие, согласно которому будут отбираться ячейки для применения условного формата и, нажав кнопку «Формат», установить сам формат, которым будут отмечены ячейки, удовлетворяющие условию.
4. Нажимая кнопку «А также», можно установить до 3-х условий и соответствующих им форматов. В условиях кроме значений можно указывать формулы, возвращающие значение, ИСТИНА либо ЛОЖЬ. Для удаления условий следует использовать кнопку «Удалить».
5. После сформирования всех необходимых условий и форматов к ним следует нажать кнопку Ok.

Условные форматы можно копировать при помощи кнопки «Формат по образцу».

Если из нескольких указанных условий более одного принимают истинное значение, то применяется только тот формат, который соответствует первому истинному условию.

Если ни одно из заданных условий не принимает истинного значения, то формат ячеек остается прежним.

### **Изменение ширины столбцов и высоты строк.**

Изменять ширину столбцов и высоту строк проще всего при помощи мыши, перетащив границу заголовка (прямоугольника, в котором находится номер строки или названия столбца) при помощи мыши. Для автоматической подгонки высоты строки или ширины столбца необходимо передвинуть курсор мыши на правую границу заголовка столбца или нижнюю границу заголовка строки и дважды щелкнуть левой кнопкой мыши. Можно также выделить сразу несколько строк или столбцов и установить высоту одной из выделенных строк либо ширину одного из выделенных столбцов – тогда автоматически установится высота всех выделенных строк либо ширина выделенных столбцов.

Для точной установки высоты строк следует выбрать Формат, Стока, Высота. При этом следует иметь в виду, что высота строки измеряется в пунктах (1/72 дюйма) в диапазоне от 0 до 409. (Если установить высоту строки равной 0, то она будет скрыта.)

Для точной установки ширины столбцов следует выбрать Формат, Столбец, Ширина и в появившемся окне ввести число в диапазоне от 0 до 255 (Это число приблизительно равно количеству символов стандартного шрифта, которое поместиться в ячейке указанной ширины). Если ввести ширину столбца равной 0, то столбец будет скрыт.

### **Ввод формул**

Чтобы ввести формулу с клавиатуры, надо выполнить следующие действия:

1. Щелкнуть ячейку, в которую необходимо ввести формулу
2. Набрать знак равенства (=)
3. Набрать формулу. Она появится в строке формул.
4. Нажать Enter или щелкнуть «галочку» в строке формул. Excel вычислит результат.

### **Задание:**

1. Запустите Microsoft Excel.
2. В ячейку **C10** введите цифру «1», а в ячейку **D10** цифру «2» Выделите диапазон **C10:D10** и протяните маркер заполнения до ячейки **H10**.
3. В ячейку **C10** введите слово «Январь» и протяните маркер заполнения до ячейки **H10**.
4. Используя команду **Файл + Параметры + Дополнительно + Общие + Создать списки для сортировки и заполнения (изменить списки)** создайте новый список со следующими элементами:  
Объект 1  
Объект 2  
Объект 3  
Объект 4  
Объект 5  
Объект 6  
Объект 7  
Объект 8
5. В ячейку **B11** введите слово «Объект 1» и протяните маркер заполнения вниз до ячейки **B18**.
6. В ячейку **B19** введите «Итого», в ячейку **B20** – «Из них в статусе – готов».
7. В ячейку **B21** скопируйте текст из ячейки **B11** и вновь протяните маркер заполнения вниз до ячейки **B28**.
8. В ячейку **B29** введите «Итого».
9. Выделите диапазон **B6:B8** и импортируйте его данные в список.
10. В ячейку **G6** введите «Темпы роста» и перетащите маркер заполнения в ячейку **G8**.
11. В диапазоне **C11:H15** введите следующие данные:

10	11	9	7	5	14
18	10	15	17	21	15
12	18	14	16	17	21
30	44	26	28	20	19
25	21	19	12	14	9

12. В ячейке **C16** введите число «5», а в ячейке **D16** число «7», выделите диапазон **C16:D16** и протяните маркер заполнения до ячейки **H16**.
13. В ячейке **C17** введите число «8», а в ячейке **D17** число «10», выделите диапазон **C17:D17** и протяните маркер заполнения до ячейки **H17**.
14. В ячейке **C18** введите число «12», а в ячейке **D18** число «14», выделите диапазон **C18:D18** и протяните маркер заполнения до ячейки **H18**.
15. В ячейку **I10** введите слово «Всего».
16. Используя кнопку **Автосумма** на панели инструментов, заполните столбец **Всего** для заполненных строк таблицы.
17. В строке «Итого» просчитайте сумму по столбцам.
18. Перед столбцом **A** вставьте дополнительный столбец. Переставьте столбцы **B** и **A**. В результате на экране должна получиться следующая картина (см рисунок 3)
19. Диапазону **C11:C29** присвойте имя **Объекты**.
20. В диапазоне **C10:I19** каждой строке и столбцу присвойте имена по заголовкам верхней строки и левого столбца.
21. Выделите блоки с помощью поля имени и с помощью функциональной клавиши **F5**.

Данные по отчётом о микробиологических анализах									
Составил									
Дата		06.05.2020							
Полученные данные		Темпы роста							
		Коэффициент н/ёл							
10 Отчёт		Коэффициент наукоёмкости							
11		Январь	Февраль	Март	Апрель	Май	Июнь	Всего	
12      Объект 1		10	11	9	7	5	14	56	
13      Объект 2		18	10	15	17	21	15	96	
14      Объект 3		12	18	14	16	17	21	98	
15      Объект 4		30	44	26	28	20	19	167	
16      Объект 5		25	21	19	12	14	9	100	
17      Объект 6		5	7	9	11	13	15	60	
18      Объект 7		8	10	12	14	16	18	78	
19      Объект 8		12	14	16	18	20	22	102	
20      Итого		120	135	120	123	123	133	757	

Рисунок 3 Внешний вид окна документа

22. После **Лист3** добавьте новый лист.
23. Переименуйте **Лист1** и присвойте ему имя «**Микробиологические анализы 2020**».
24. Переместите **Лист2** после **Лист4**.
25. Удалите **Лист3**.
26. Сохраните файл под именем **ПЗ-10 Excel \_ФИ**.
27. В ячейку **B4** введите функцию **СЕГОДНЯ**.

28. В ячейку **K10** введите слово «**Среднее**», в ячейку **L10** - слово «**Максимальное**», в ячейку **M10** введите слово «**Минимальное**».
29. В столбце **K** рассчитайте среднее значение цифр соответствующей строки с января по июнь (функция **СРЗНАЧ**).
30. В столбцах **L** и **M** рассчитайте максимальное (функция **МАКС**), минимальное (функция **МИН**) значение цифр соответствующей строки с января по июнь, используя имена диапазона ячеек.
31. В ячейку **D21** введите формулу  $= D11*D7$  и не закрывая ячейки измените тип ссылки в ячейке **D7** на абсолютную.
32. Скопируйте содержимое данной ячейки по вертикали до ячейки **D28**.
33. Каждую ячейку в полученном вертикальном столбце скопируйте по горизонтали до столбца **I**.
34. Рассмотрите, как в формулах изменяются относительные и абсолютные ссылки на ячейки при копировании.
35. Для полученных данных рассчитайте строку **Итого**, столбцы **Всего, Среднее, Максимальное, Минимальное**.
36. В ячейку **C30** введите функцию расчета количества непустых ячеек в диапазоне **C1:C29** (функция **СЧЕТЗ**).
37. В ячейку **D30** введите функцию расчета количества ячеек с числами в диапазоне **D1:D20** (функция **СЧЕТ**).
38. На **Листе 2** выделите диапазон **I7:K7** и поверните текст на  $90^0$  и выровняйте его по центру по горизонтали и по вертикали.
39. На листе «**Микробиологические анализы 2020**» выделите диапазон **D11:J18** и установите для чисел пользовательский формат **###0"штук"**. Если в ячейках появятся значки **####**, то измените ширину столбцов.
40. Выделите диапазон **D10:J10** и поверните текст в этих ячейках на  $45^0$ .
41. Для диапазонов **C6:D8** и **C10:M29** установите вокруг них толстые рамки.
42. Для диапазона **J11:J29** установите условное форматирование; выделите красным цветом и полужирным шрифтом те ячейки, значения которых больше или равно 90.

#### **Контрольные вопросы:**

- 1.Какие существуют способы заполнения и редактирования ячеек?
- 2.Как выделить смежные и несмежные диапазоны?
- 3.Как создать и импортировать списки?
- 4.Как производиться автосуммирование строк и столбцов?
- 5.как перейти между листами в одной рабочей книге?
- 6.Какие существуют способы ввода формул в ячейку?
- 7.Как в ячейку ввести функцию?
- 8.Как задать автоформат выделенному диапазону ячеек?
- 9.Какие существуют способы форматирования ячеек?
10. Как наложить пользовательский формат на ячейку?
11. Как задать условное форматирование для выделенного диапазона ячеек?
12. Как в формуле сделать ссылку на другой лист?
13. Как в формулах обозначаются абсолютные и относительные ссылки?

В результате рассмотрения вопроса № 2.2 студенты должны:

- освоить ввод и редактирование ячеек Microsoft Excel, ввод формул и функций, а также форматирование данных в листах Microsoft Excel;
- приобрести умения и навыки ввода и редактирования ячеек Microsoft Excel, ввода формул и функций, а также форматирования данных в листах Microsoft Excel.

2.3 Краткое содержание вопроса № 2.3. Углубленное изучение возможностей современных офисных технологий для использования в профессиональной деятельности, научных исследованиях в биологии. Работа с пакетом Microsoft Office: MS Access, создание баз данных, создание межтабличных связей, сортировка и фильтрация данных; запросы и отчёты.

Задание для работы:

1. Ознакомиться с краткими теоретическими сведениями (даны в описании работы).
2. Выполнить практическую работу:
  - MS Access, создание баз данных,
  - создание межтабличных связей,
  - сортировка и фильтрация данных,
  - создание запросов и отчётов.
3. Ответить на контрольные вопросы.

### **Краткие теоретические сведения**

Сразу после запуска Access на экране появляется окно, в котором можно установить флажок *Новая база данных* и затем нажать кнопку ОК. После этого на экране появится окно *Файл новой базы данных*, в котором надо ввести название новой базы данных, выбрать папку, где она будет храниться и затем нажать кнопку Создать. После этого на экране появится окно базы данных, в котором перечислены объекты базы данных: *Таблицы, Запросы, Формы, Отчеты, Макросы, Модули*.

#### **Планирование баз данных.**

Перед созданием новой базы данных обязательно необходимо хорошо продумать следующие вопросы:

- 1 Какие данные будут храниться в БД и как их организовать наилучшим образом? Это позволит определить, какие потребуются таблицы и какие связи необходимо организовать между ними.
- 2 Какие действия с данными необходимо будет производить в процессе эксплуатации БД? Это позволит определить, какие потребуются формы.
- 3 Какие документы необходимо будет выводить на печать? Это позволит определить, какие потребуются отчеты.

Таблицы и связи между ними представляют собой так называемую модель данных, которая является основой любой БД. В теории БД существуют так называемые правила нормализации данных, которые позволяют устранить избыточность, противоречивость и непоследовательность модели данных.

При проектировании таблиц базы данных рекомендуется:

#### *1 Избегать повторения информации*

Если информация повторяется, то логичнее разбить информацию в две таблицы, задать ключевые поля и связать таблицы по ключевым полям.

#### *2 Избегать повторяющихся групп*

Повторяющаяся группа — это столбцы, которые повторяются в пределах одной и той же строки для хранения нескольких значений одного вида.

Например, мы храним таблицу со списком зарегистрированных фирм, и в ней 10 столбцов с наименованиями фирм-учредителей. Но количество учредителей чаще всего меньше 10-ти и поэтому многие поля в этой таблице будут пустыми.

С другой стороны, если когда-нибудь придется зарегистрировать фирму с количеством учредителей больше 10, то придется заводить новый столбец. Если наша таблица связана с другими, и мы уже разработали много форм и отчетов, то такая ситуация повлечет за собой переделку всей базы данных.

### *3 Каждая таблица должна описывать одну сущность реального мира*

Не следует смешивать, например, в одной таблице сведения о сотрудниках фирмы и заключенных ими договорах.

### *4 Там, где это возможно, следует использовать коды (первичные ключи)*

Первичные ключи помогают связывать таблицы.

### *5 Справочную информацию следует помещать в отдельные таблицы.*

#### **Создание таблиц в Access.**

Для создания новой таблицы в базе данных необходимо в окне базы данных выбрать объект *Таблица*, а затем нажать кнопку *Создать*. В появившемся окне будет предложено выбрать один из следующих способов создания:

- 1 Режим таблицы
- 2 Конструктор
- 3 Мастер таблиц
- 4 Импорт таблиц
- 5 Связь с таблицами

#### **Конструктор таблиц.**

Чтобы добавить поле, в верхней части окна таблицы в режиме конструктора следует ввести имя поля и определить его тип.

*Имя поля* должно содержать не более 64 символов и может включать любые комбинации букв, цифр и пробелов, а также специальных символов, за исключением точки, восклицательного знака, надстрочного символа и прямых скобок. Имя не должно начинаться с пробела и содержать управляемые символы.

*Тип данных* определяет, какого вида данные допускается вводить в поле.

Уникальная метка, называемая *ключом*, используется для определения каждой записи таблицы. Подобно тому, как номерной знак однозначно определяет автомобиль, ключ определяет запись.

Ключевые поля в таблицах используются для создания межтабличных связей. Чтобы определить ключ, необходимо выделить строку с описанием нужного поля и нажать пиктограмму Ключ.

Для задания свойства поля надо выбрать его в верхней части окна конструктора таблиц и в нижней части окна ввести значения этого свойства или выбрать его из списка.

По окончании описания полей таблицы, необходимо закрыть окно конструктора.

Для наполнения таблицы данными откройте ее в режиме таблицы (двойной щелчок по значку таблицы в окне база данных) и внесите информацию, соответственно типам данных каждого поля.

От того, как новая таблица связана с остальными, зависит, какое из ее полей следует назначить первичным ключом. В теории баз данных известны 4 варианта *связей между двумя таблицами*, называемых обычно *отношениями*.

**Связь Один-к-одному.** Каждой записи *первой* таблицы соответствует не больше одной записи *второй* таблицы и *наоборот*. Соответствие записей устанавливается в результате поиска в поле, являющемся *первичным ключом* одной из таблиц, значения поля, называемого *внешним ключом* второй таблицы.

**Связь Многие-к-одному.** Любой записи *второй* таблицы может соответствовать любое количество записей *первой* таблицы, но не *наоборот*. В этом случае ключевое поле *первой* таблицы будет *внешним ключом*, и повторяющиеся значения в нем допускаются.

**Связь Один-ко-многим.** Первичный ключ *первой* таблицы (поле, содержащее уникальные значения), связывается с внешним ключом *второй* таблицы (значения поля могут повторяться). При этом каждой записи *первой* таблицы может соответствовать несколько записей *второй*. Можно сказать, что «один-ко-многим» — это «многие-к-одному» *наоборот*.

**Связь Многие-ко-многим.** Каждой записи одной таблицы может соответствовать любое количество записей другой таблицы и *наоборот*. Соответственно, поля *обеих* таблиц, по которым осуществляется связь, являются *внешними ключами* и могут содержать повторяющиеся значения.

### **Создание межтабличных связей.**

Для создания межтабличных связей в СУБД MS Access существует команда *Файл – Схема данных*, а также на панели инструментов расположена одноименная пиктограмма *Схема данных*.

В окне схемы данных связи можно создавать путем перетаскивания полей из одной таблицы в другую.

При создании межтабличной связи очень важно, из какой таблицы в какую перетаскивается поле. Поля надо перетаскивать из таблицы со стороны «один» в таблицу со стороны «многие». При этом в окне *Изменение связей* главная таблица (со стороны «один») окажется слева под заголовком *Таблица/запрос*, а таблица со стороны «многие» — справа под заголовком *Связанная таблица/запрос*. При попытке сделать *наоборот* мы получим сообщение об ошибке.

На созданной в схеме связи следует щелкнуть правой кнопкой мыши, чтобы установить дополнительные параметры:

– в окне *Изменение связей* щелкнуть на кнопке *Объединение*. Появится окно *Параметры объединения*. Если необходимо, например, чтобы в дальнейшем при объединении данных таблиц отображались все записи таблицы «Пациенты», независимо от того, имеется ли соответствующая запись в таблице «Посещения», надо выбрать 2-й тип объединения.

В диалоговом окне *Изменение связей* есть очень полезный флажок – *Обеспечение целостности данных*. Следует установить его, чтобы Access отказывалась сохранять в подчиненной таблице записи, относящиеся к несуществующей записи в главной таблице.

Щелкните на кнопке *Создать*, чтобы новая связь появилась в окне *Схема данных*.

Связи между таблицами в Access можно создавать разными способами. На первых порах удобнее всего использовать команду *Вставка - Поле подстановки* в режиме конструктора для вызова мастера создания связей.

## ЗАДАНИЕ

### I Создание таблиц базы данных

Создадим базу данных, содержащую сведения о студентах потока. Для этого выполним следующие действия:

1. Запустим программу **MS Access**.
2. При запуске появится диалоговое окно, в котором надо выбрать строку *Новая база данных*.
3. В окне *Файл новой базы данных* указать имя новой БД- **ДЕКАНАТ** и сохранить в папке ПР-10.
4. Нажмите кнопу **Создать**. В появившемся окне *База данных* активизировать вкладку *Таблицы* и щелкнуть на кнопке **Создать**.
5. Создать таблицу, воспользовавшись *Конструктором*. В окне *Новая таблица* выбрать пункт *Конструктор* и подтвердить выбор.
6. Определить поля таблицы. В появившемся окне создать поля базы данных, согласно следующей таблице.

Поле	Тип поля	Размер поля
Номер	Счетчик	
Фамилия	Текстовое	15
Имя	Текстовое	10
Отчество	Текстовое	15
Дата рождения	Дата	Краткий формат
Группа	Текстовое	7
Адрес	Текстовое	20

7. Для ввода типа поля использовать значок контекстного меню, который появляется при установке курсора в столбец *Тип данных*.
8. Определить первичный ключ для таблицы. В данной таблице ключевым является поле *Номер*. Чтобы сделать поле ключевым, нужно выделить его и выбрать из меню *Правка* команду *Ключевое поле* или нажать кнопку **Ключевое поле** на панели инструментов.
9. Закрыть заполненную таблицу. При закрытии сохраните ее под именем **Студенты**.

### II Ввод и редактирование данных

10. В окне *База данных* появилось имя сохраненной таблицы. Для того, чтобы вводить данные, надо открывать ее в режиме таблицы. Щелкнуть на кнопку **Открыть**.
11. Занести в таблицу 6–7 записей. Для поля *Группа* использовать номера 56, 57, 58. Отредактировать введенные данные в таблицу: заменить во второй записи фамилию.
12. В поле *Дата рождения* изменить в первой записи год рождения.
13. Удалить последнюю запись в таблице. Для этого нужно выделить ее: установить курсор мыши к левой границе таблицы до изменения его в виде стрелки, направленной вправо, щелкнуть мышью и нажать клавишу **Delete**.
14. Добавить еще две записи.
15. Сохранить таблицу и закрыть ее.

### III Создание многотабличной БД

1. Создать таблицы **СЕССИЯ** и **СТИПЕНДИЯ**, используя ту же технологию, что и при создании таблицы **СТУДЕНТЫ** в пункте 1. Атрибуты поля *Номер* таблицы **СЕССИЯ**

должны быть такими же, как атрибуты этого же поля таблицы СТУДЕНТЫ. Состав полей и их свойства следующие:

### СЕССИЯ

Признак ключа	Поле	Тип поля	Размер поля
Ключ	Номер	Числовое	
	Оценка 1	Числовое	Фиксированный
	Оценка 2	Числовое	Фиксированный
	Оценка 3	Числовое	Фиксированный
	Оценка 4	Числовое	Фиксированный
	Результат	Текстовое	5

### СТИПЕНДИЯ

Признак ключа	Поле	Тип поля	Размер поля
Ключ	Результат	Текстовое	5
	Процент	Числовое	Процентный

2. Заполнить таблицы данными; оценки в записи ввести на свое усмотрение так, чтобы в записях присутствовали разные комбинации оценок из четырех групп:

Неуд.	Хор.	Хор1.	Отл
За Удовл, и неудов.	За две 4 и более	5 5 5 4	5 5 5 5

3. В поле *Результат* данные заносить в соответствии с представленной таблицей, например, если в записи три оценки 5 и одна оценка 4, то в результат занести хор1.

4. Поле *Процент* заполнить в соответствии со следующей таблицей:

Результат	Процент
Неуд.	0,00%
Хор.	100,00%
Хор1.	200,00%
Отл.	300,00%

5. Сохранить обе таблицы и закрыть их.

### IV Установление связей между таблицами

6. В окне *База данных Деканат* должны быть имена трех таблиц: СТУДЕНТЫ, СЕССИЯ, СТИПЕНДИЯ. Для установления связей выполнить команду *Сервис – Схема данных*.

7. В появившемся окне выполнить добавление всех трех таблиц в схему.

8. Установить связи между таблицами СТУДЕНТЫ и СЕССИЯ. Для этого протащить указатель мыши от поля *Номер* таблицы СТУДЕНТЫ к полю *Номер* таблицы СЕССИЯ при нажатой клавише мыши.

9. В появившемся диалоговом окне *Связи* активизировать значок *Обеспечение целостности данных*, отношение «Один к одному», активизировать значки *Каскадное обновление связанных полей* и *Каскадное удаление связанных полей*. Прочитать встроенную справку об этих значках. Нажать кнопку **Создать**.

10. Установить связь между таблицами СТИПЕНДИЯ и СЕССИЯ. Для этого протащить указатель мыши от поля *Результат* таблицы СТИПЕНДИЯ к полю *Результат* таблицы СЕССИЯ. Здесь отношение «Один ко многим».

11. Закрыть окно *Схема данных*, при выходе сохранить связи.

### **Задание для самостоятельной работы**

1. Добавьте в базу данных **Деканат** таблицу **ФАКУЛЬТЕТЫ**, содержащую следующую информацию:

Шифр специальности	Название специальности	Название факультета
06.04.01	«Биология»	Биотехнологий
36.04.02	«Зоотехния»	Биотехнологий
06.04.01	«ИБАС»	ИУР и КБ
09.03.01	«ИВТ»	ИУР и КБ
27.03.04	«УТС»	ИУР и КБ

2. Добавьте в таблицу **СТУДЕНТЫ** поле шифр специальности и заполните его соответствующими данными из таблицы **ФАКУЛЬТЕТЫ**. Свяжите таблицы **СТУДЕНТЫ** и **ФАКУЛЬТЕТЫ**.

Сохраните базу данных **Деканат**.

### **V Поиск, сортировка и отбор данных**

1. Откройте таблицу **СТУДЕНТЫ** в режиме формы.

2. Осуществите поиск студента по полю *Фамилия*. Для этого:

- установите курсор в строку поля, по которому будет осуществляться поиск, то есть *Фамилия*
- выполните команду *Правка – Найти* или нажмите пиктограмму *Найти*.

3. Закрыть окно формы.

4. Откройте таблицу **СТУДЕНТЫ**.

5. Отсортируйте записи таблицы в алфавитном порядке по полю *Фамилия*. Для этого:

- установите курсор в поле *Фамилия*;
- выполните команду *Записи – Сортировка – По возрастанию* или воспользоваться соответствующей пиктограммой

6. Используя фильтр, вывести на экран список студентов группы 21 ИБАС. Для этого:

- Выполнить команду *Записи – Изменить фильтр* или использовать пиктограмму
- Установить курсор в поле *Группа*, и из контекстного меню выбрать номер нужной группы
- Выполнить команду *Фильтр – Применить фильтр* или использовать соответствующую пиктограмму
- Чтобы убрать фильтр, воспользуйтесь командой *Записи – Удалить фильтр* или применить ту же пиктограмму.
- Закрыть таблицу.

7. Открыть таблицу **СТУДЕНТЫ** в режиме формы и выполнить тот же фильтр: отобрать студентов группы БУА-12. Для этого выполните действия:

- В поле *Формы* найти запись, которая содержит индекс нужной группы – 21 ИБАС
- Выполнить команду *Записи – Фильтр – Фильтр по выделенному* или воспользоваться соответствующей пиктограммой на панели инструментов

8. Удалить фильтр

### **Контрольные вопросы**

1. Назовите основные элементы окна Access.
2. Перечислите основные объекты окна базы данных.
3. Какие режимы работы используются для работы с таблицей?

4. Что такое *Конструктор* в СУБД Access?
5. Для чего служит ключевое поле?
6. Что такое *счетчик*?
7. Какой тип данных следует использовать для создания поля, содержащего рисунки?
8. Что такое маска ввода? Какие знаки используются для работы с маской?
9. Объясните для чего необходимо связывать таблицы при работе с базами данных?
10. Перечислите виды связей между таблицами базы данных.
11. Назовите обязательные условия при создании связей между главной и подчиненной таблицами?
12. Как вы понимаете связь «Один к одному»?
13. Что означает связь «Один ко многим»?
14. Что такое связь «Многие ко многим»?
15. Что такое целостность данных?

В результате рассмотрения вопроса № 2.3 студенты должны:

- освоить технологию создания базы данных, создания межтабличных связей, сортировки и фильтрации данных, создания запросов и отчётов;
- приобрести умения и навыки создания базы данных, создания межтабличных связей, сортировки и фильтрации данных, создания запросов и отчётов.

### 3. Наименование вопроса № 3.

Компьютерные технологии изучения вероятностных и статистических моделей в биологии и обработки результатов научных исследований в биологии, статистической обработки биологических данных.

Краткое содержание вопроса № 3. Компьютерные технологии изучения вероятностных и статистических моделей в биологии и обработки результатов научных исследований в биологии, статистической обработки биологических данных.

Задание для работы:

1. Формулировка статистической модели (задачи): дана выборка объёма  $n = 30$  (Таблица 1). Выполнить описательную статистику выборки с помощью Надстройки «Пакет Анализа» MS Excel.

2. Выполнение описательной статистики выборки с помощью Надстройки «Пакет Анализа» MS Excel.

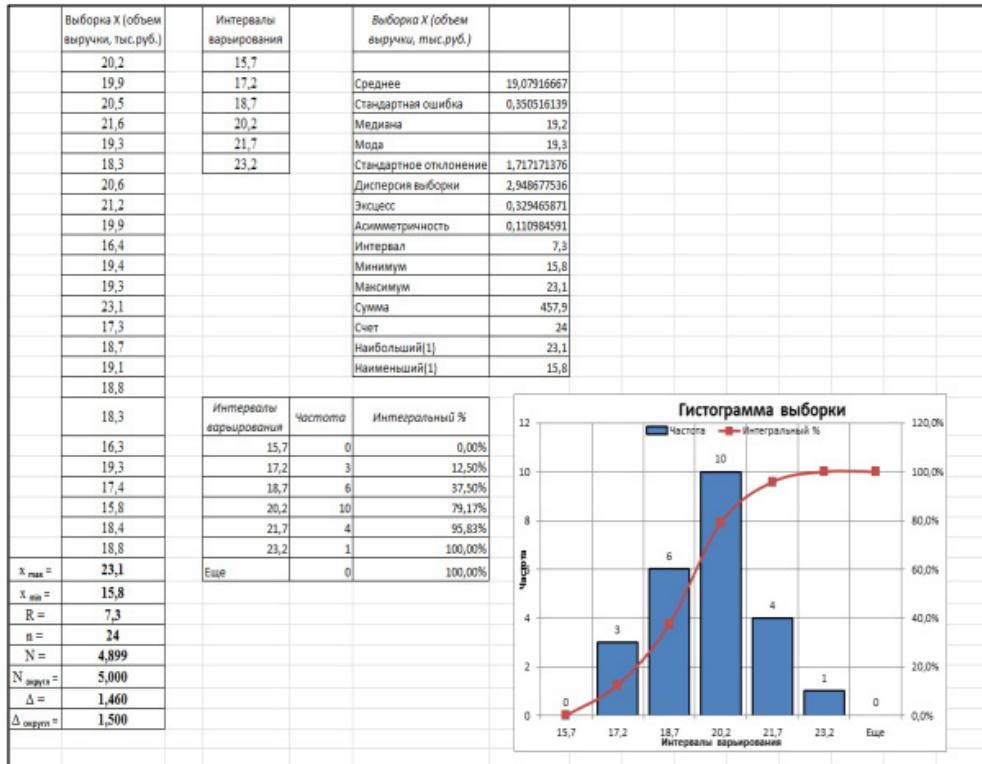
1. Формулировка статистической модели (задачи): дана выборка объёма  $n = 30$  (Таблица 1); выполнить описательную статистику выборки с помощью Надстройки «Пакет Анализа» MS Excel.

**Таблица 1**

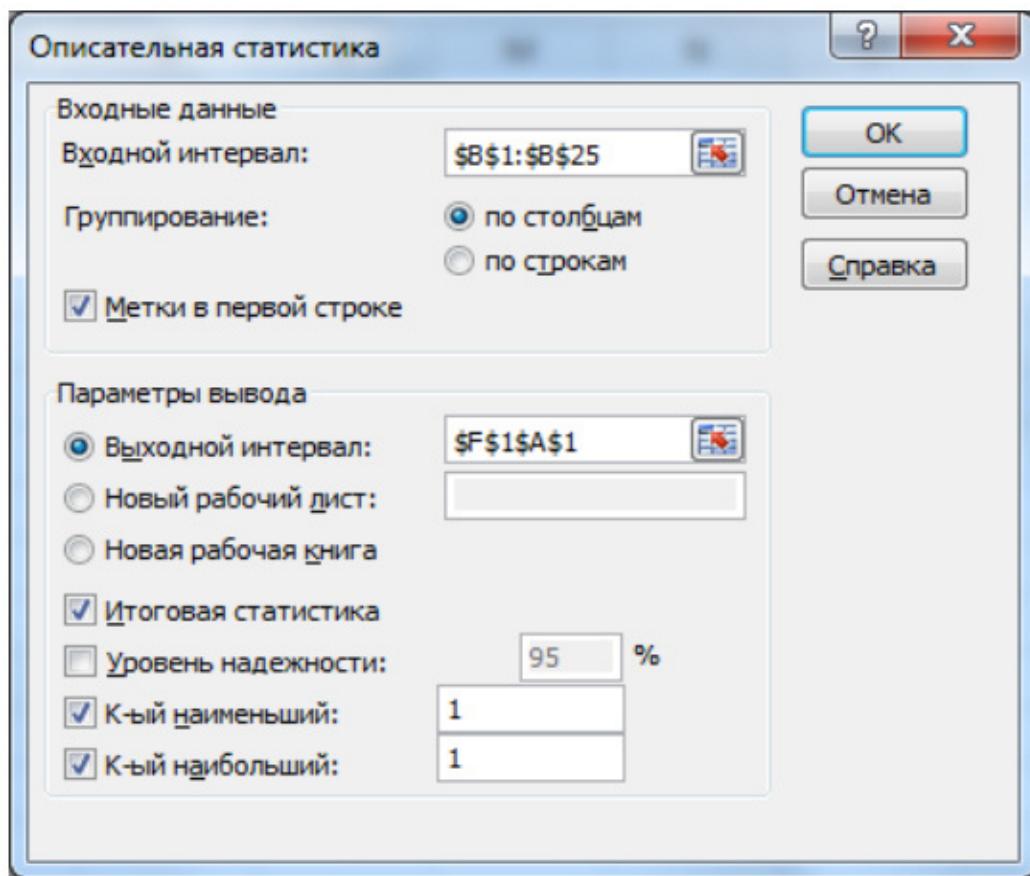
3	0	3	-6	4	-2	4	-1	-1	-2
1	-2	-2	-5	1	-5	1	-2	-1	-1
0	-3	0	-3	0	-2	0	-1	-3	-3

Для решения задачи в MS Excel необходимо:

1. Идентифицировать свою работу, переименовав Лист1 в Титульный лист и записав номер практической работы, ее название, кто выполнил и проверил.
2. Переименовать Лист 2 в Исходные данные и набрать столбец исходных данных.
3. Вычислить величины  $x_{max}$ ,  $x_{min}$ ,  $R$ ,  $n$ , используя встроенные функции Excel *МАКС*, *МИН*, *СЧЕТ*, *КОРЕНЬ*.



4. Сформировать столбец интервалов группировки. Наберите команду Данные → Анализ данных → Гистограмма и в появившемся диалоговом окне выполните нужные установки. Отформатируйте полученную таблицу и построенную гистограмму выборки.
5. Наберите команду Данные → Анализ данных → Описательная статистика и в появившемся диалоговом окне выполните нужные установки.



В результате рассмотрения вопроса № 3 студенты должны:

- освоить понятие о компьютерных технологиях изучения статистических моделей, статистической обработке биологических данных с Excel: оценивании характеристик генеральной совокупности по выборке, создании таблиц частот и гистограмм
- приобрести умения и навыки использования компьютерных технологий изучения статистических моделей, статистической обработки биологических данных с Excel: оценивании характеристик генеральной совокупности по выборке, создании таблиц частот и гистограмм.

#### 4. Наименование вопроса № 4.

Компьютерные технологии исследования оптимизационных моделей в биологии с Excel.

4.1 Наименование вопроса № 4.1. Компьютерные технологии решения оптимизационных задач линейного программирования в биологии с Excel.

4.1.1 Наименование вопроса № 4.1.1. Решение задачи об оптимизации распределения ресурсов.

4.1.2 Наименование вопроса № 4.1.2. Решение калькуляционной задачи линейного программирования (задачи о выборе оптимального рациона кормления животных).

4.1.3 Наименование вопроса № 4.1.3. Решение задачи об оптимизации транспортных расходов.

4.1.4 Наименование вопроса № 4.1.4 Решение задачи о назначениях, задачи целочисленного программирования.

4.2 Наименование вопроса № 4.2. Компьютерные технологии решения задач (моделей)

лей) оптимизации на графах и сетях в биологии: задача об остеове минимального веса сети и её решение по алгоритму Краскала с Excel.

4.1.1. Краткое содержание вопроса № 4.1.1: Решение задачи об оптимизации распределения ресурсов.

1. Формулировка задачи о распределении ресурсов.

**Задача.** Предприятие с помощью биотехнологий производит и продаёт продукцию двух видов: «1 Продукт» и «2 Продукт». Для производства продукции используются ресурсы двух категорий: A и B. Расходы ресурсов A и B на производство единицы продукции каждого вида, запасы ресурсов и цены продукции приведены в таблице I.

Таблица 1

Ресурсы	Расход ресурсов на ед. продукции		Запасы ресурсов
	1 Продукт	2 Продукт	
A	1	2	3
B	3	1	3
Количество продукции	$x_1$	$x_2$	
Цены	2(ден. ед.)	1(ден. ед.)	

Выяснить, какое количество продукции каждого вида надо производить предприятию (составить план производства), чтобы получить максимум прибыли.

2. Математическая модель задачи.

Ранее, в вопросе № 4 Темы 1, была построена математическая модель этой задачи

$$Z(x_1, x_2) = 2 \cdot x_1 + x_2 \rightarrow \max, \quad (1)$$

$$\begin{cases} x_1 + 2 \cdot x_2 \leq 3, \\ 3 \cdot x_1 + x_2 \leq 3, \end{cases} \quad (2)$$

$$x_1, x_2 \geq 0$$

(3)

3. Компьютерные технологии исследования модели о распределении ресурсов с надстройкой Excel «Поиск решения».

3.1. Ввод данных и формул в таблицу Excel. Открыть Книгу **Excel**, Лист1.

-Объединим ячейки B1 и C1. Для этого выделить ячейки, нажать правую кнопку мыши. В появившемся окне вызвать «Формат ячеек», затем «Выравнивание» и поставить галочку против опции «объединение ячеек», нажать ОК. В объединённые ячейки впишем заголовок «Переменные».

-В ячейку A2 вписать «Имя», в A3- «План», в ячейку A4 «Цена», в B2- «1 Продукт», в C2- «2 Продукт», в D2 «Прибыль».

-В ячейки B4 и C4 заносятся значения цен на продукцию.

-Для переменных  $x_1$  и  $x_2$  отводятся ячейки B3 и C3. Это изменяемые(рабочие) ячейки, в них исходные данные не заносятся и в результате решения задачи в эти ячейки будут вписаны оптимальные значения. Таблица данных будет иметь вид

-В ячейке D4 после окончания решения задачи будет указана оптимальное значение прибыли (целевая ячейка). С этой целью в ячейку D4 вводится формула для вычисления значений целевой функции  $Z = 2 \cdot x_1 + x_2$ . Для этого надо выполнить следующие операции:

- 1) курсор в D4, выделить эту ячейку,
- 2) щёлкнув по кнопке  $f_x$  вызвать Мастера функций, в открывшемся окне в категории «10 недавно использовавшихся» выбрать «Математические», а затем «СУММПРОИЗВ», ОК.

В появившемся окне «Аргументы функции» в поле «Массив 1» ввести адреса изменяемых ячеек B3:C3(протаскивая курсор мыши по ячейкам), в поле «Массив 2» вводятся адреса ячеек с ценами на продукцию B4:C4, «Массив 3» игнорируется. Нажать OK. В ячейке D4 появится число 0.

-Объединить ячейки B5 и C5 и вписать «Ограничения», в A6- «Ресурсы», в B6 и C6  $x_1$  и  $x_2$ , в D6 «Расход», в E6 «Запасы», A7 и A8 значки ресурсов, в поле B7:C8- нормы расхода ресурсов.

-В ячейку D7 вводится формула вычисления израсходованного ресурса A  $x_1 + 2 \cdot x_2$ , в ячейку D8- формула израсходованного ресурса B  $3 \cdot x_1 + x_2$  (также, как и формула целевой функции).

- В ячейки E7 и E8 вносим размеры запасов ресурсов.

Данные и формулы введены. Интерфейс задачи будет иметь вид

### 3.2. Использование надстройки Excel «Поиск решения».

Надстройка Excel «Поиск решения» при первом использовании должна быть предварительно активирована. Открыв Excel, нажать кнопки «Office» → «Параметры Excel» → «Надстройки» → «Неактивные надстройки приложений» → выделить строку «Поиск решения» → «Управление: надстройки Excel» → «перейти» → OK.

Щёлкнув на ленте кнопку «Данные», затем «Поиск решений» откроем окно «Поиск решений».

-В поле «Установить целевую ячейку» ввести адрес целевой ячейки D4, щёлкнув по ней курсором мыши.

-Выбрать «равной максимальному значению».

-В поле «изменяя ячейки» указать адреса B3:C3.

-В поле «Ограничения» щёлкнуть «Добавить». После появления поля «Добавление ограничения» в поле «Ссылка на ячейку:» сделать ссылку на ячейку D7, выбрать знак  $\leq$ , в поле «Ограничение:» ввести адрес ячейки с запасом ресурса A- E7. Вновь выбрать «Добавить» провести ввод ограничения по ресурсу B, затем по ограничению  $x_1, x_2 \geq 0$ . После этого нажать OK.

### 3.3. Настройка параметров решения задачи.

Выбрав в окне «Поиск решений» опцию «Параметры» в появившемся окне «Параметры поиска решения» установить флажок в поле «Линейная модель». При таком выборе при решении задачи будет использоваться симплекс-метод. Остальные значения можно оставить без изменения. Нажать OK.

	A	B	C	D	E	F	G	H	I	J	K
1		<b>Переменные</b>									
2	Имя	1 Продукт	2 Продукт	Прибыль							
3	План	0,6	1,2								
4	Цена	2	1	2,4							
5		<b>Ограничения</b>									
6	Ресурсы	X <sub>1</sub>	X <sub>2</sub>	Расход	Запасы						
7	A	1	2	3	3						
8	B	3	1	3	3						
9											
10											
11											
12											
13											

### 2.3. Завершение решения задачи и просмотр результатов.

В окне «Поиск решений» нажимаем кнопку «Выполнить». Появляется окно «Результаты поиска решения». Можно выбрать тип отчёта, сохранить найденное решение или восстановить исходные значения, OK.

Лин\_прогр-1.xlsx - Microsoft Excel

Поиск решения

Установить целевую ячейку: \$D\$4

Равно:  максимальному значению  значению: 0

Изменяя ячейки: \$B\$3:\$C\$3

Ограничения:

- \$B\$3:\$C\$3 >= 0
- \$D\$7 <= \$E\$7
- \$D\$8 <= \$E\$8

Параметры

Добавить Изменить Удалить Восстановить Справка

Лин\_прогр-1.xlsx - Microsoft Excel

Результаты поиска решения

Решение найдено. Все ограничения и условия оптимальности выполнены.

Тип отчета

Сохранить найденное решение  Восстановить исходные значения

OK Отмена Сохранить сценарий... Справка

В ячейках B3 и C3 появятся оптимальные значения плана 0,6 и 1,2, а в ячейке D4 оптимальное значение прибыли 2,4. Задача решена.

Лин_прогр-1.xlsx -									
Главная Вставка Разметка страницы Формулы Данные Рецензирование Вид Надстройки									
Шрифт Выравнивание									
A1									
1	A	В	С	D	E	F	G	H	I
2	Имя	1 Продукт	2 Продукт	Прибыль					
3	План	0,6	1,2						
4	Цена	2	1	2,4					
5	Ограничения								
6	Ресурсы	X <sub>1</sub>	X <sub>2</sub>	Расход	Запасы				
7	A	1	2	3	3				
8	B	3	1	3	3				
9									

В результате рассмотрения вопроса № 4.1.1 студенты должны:

- освоить понятие о задачах (моделях) оптимизации распределения ресурсов, моделях (задачах) линейного программирования в биологии и технологии их решение с Excel;
- приобрести умения и навыки создания моделей линейного программирования в биологии и их решения с Excel.

4.1.2. Краткое содержание вопроса № 4.1.2. Решение калькуляционной задачи линейного программирования (задачи о выборе оптимального рациона кормления животных).

1. Формулировка калькуляционной задачи.

**Задача.** Составляется комбинированный корм из трёх злаков: кукурузы, овса и ржи. Калорийность и содержание витамина С в одном килограмме каждого злака, а также цена одного кг каждого злака указаны в таблице:

	Кукуруза	Овёс	Рожь
Калорийность(ккал)	200	175	100
Содержание С (в Гр)	5	1	3
Цена (ден. ед.)	6	4	1

Составить наиболее дешёвый комбинированный корм, 1кг которого содержал бы не менее 150 ккал и не менее 3 г витамина С.

2. Математическая модель задачи.

Ранее, в вопросе № 4 Темы 1, была построена математическая модель этой задачи

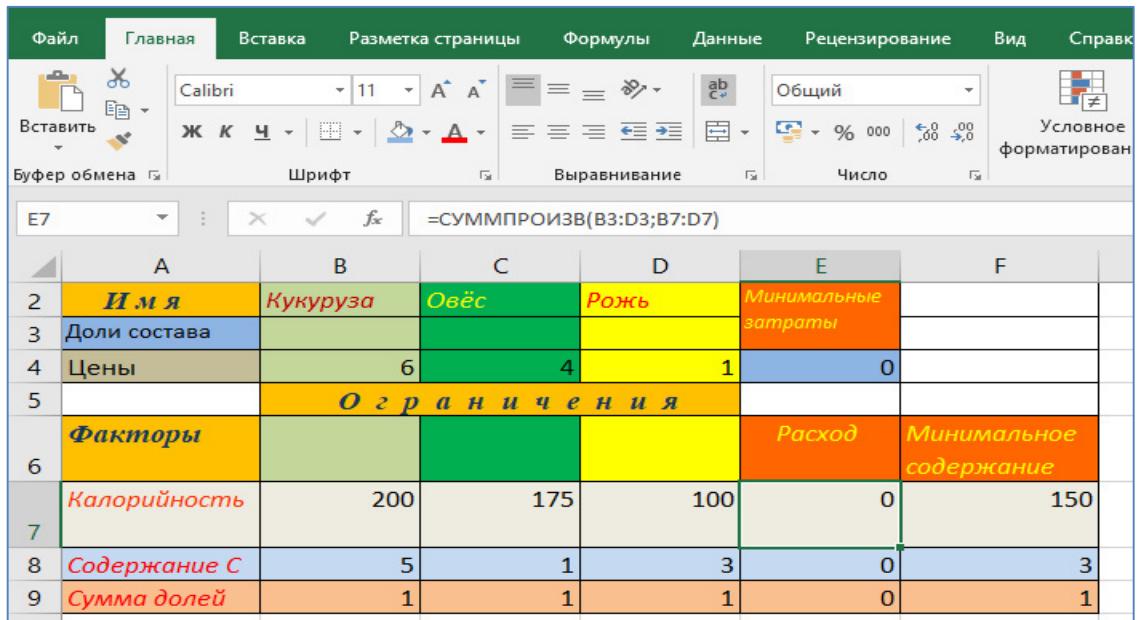
$$Z = 6 \cdot x_1 + 4 \cdot x_2 + x_3 \rightarrow \min ,$$

$$\begin{cases} 200 \cdot x_1 + 175 \cdot x_2 + 100 \cdot x_3 \geq 150, \\ 5 \cdot x_1 + x_2 + 3 \cdot x_3 \geq 3, \\ x_1 + x_2 + x_3 = 1, \end{cases}$$

$$x_1 \geq 0. x_2 \geq 0. x_3 \geq 0.$$

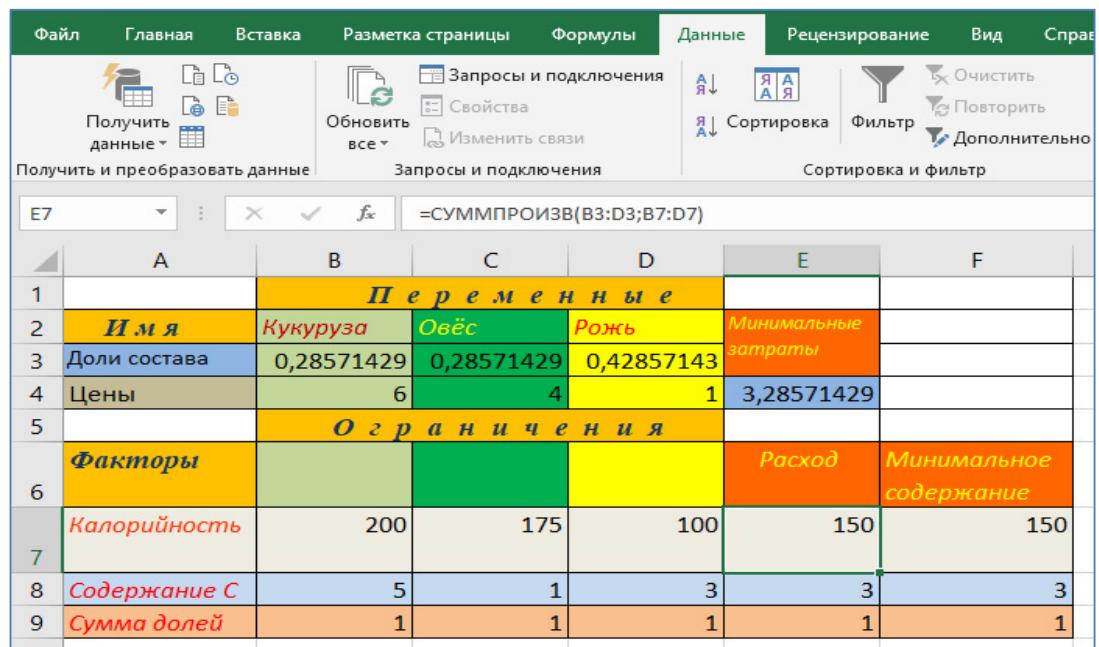
3. Компьютерные технологии исследования калькуляционной модели (задачи) линейного программирования с надстройкой Excel «Поиск решения».

1) Интерфейс задачи после ввода данных и формул



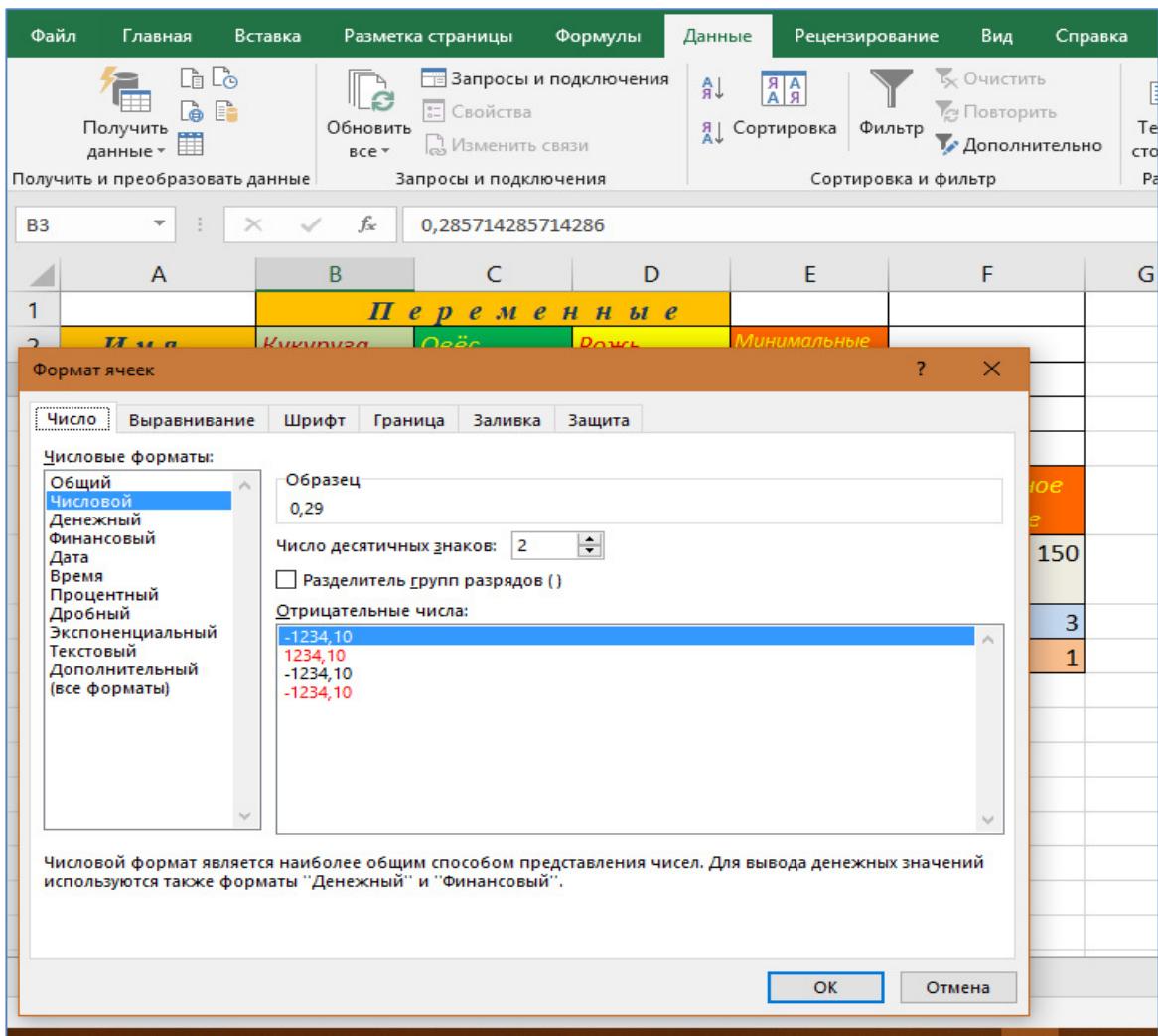
A	В	С	D	E	F
<b>Имя</b>	<i>Кукуруза</i>	<i>Овёс</i>	<i>Рожь</i>	<i>Минимальные затраты</i>	
<b>Доли состава</b>					
<b>Цены</b>	6	4	1	0	
<b>Ограничения</b>					
<b>Факторы</b>				<i>Расход</i>	<i>Минимальное содержание</i>
<i>Калорийность</i>	200	175	100	0	150
<i>Содержание С</i>	5	1	3	0	3
<i>Сумма долей</i>	1	1	1	0	1

2) Интерфейс задачи после применения «Поиска решения»



A	В	С	D	E	F
<b>Переменные</b>					
<b>Имя</b>	<i>Кукуруза</i>	<i>Овёс</i>	<i>Рожь</i>	<i>Минимальные затраты</i>	
<b>Доли состава</b>	0,28571429	0,28571429	0,42857143		
<b>Цены</b>	6	4	1	3,28571429	
<b>Ограничения</b>					
<b>Факторы</b>				<i>Расход</i>	<i>Минимальное содержание</i>
<i>Калорийность</i>	200	175	100	150	150
<i>Содержание С</i>	5	1	3	3	3
<i>Сумма долей</i>	1	1	1	1	1

3) Открыв формат ячеек (число) B3, C3, D3, E4 взять значения с двумя десятичными знаками.



Лин_прогр-1 - Excel						
Переменные						
Имя	Кукуруза	Овёс	Рожь	Минимальные затраты		
Доли состава	0,29	0,29	0,43			
Цены	6	4	1	3,29		
Ограничения						
Факторы				Расход	Минимальное содержание	
Калорийность	200	175	100	150	150	
Содержание С	5	1	3	3	3	
Сумма долей	1	1	1	1	1	1

В результате рассмотрения вопроса № 4.1.2 студенты должны:

- освоить понятие о калькуляционных моделях (задачах) линейного программирования в биологии и технологии их решение с Excel;
- приобрести умения и навыки создания калькуляционных моделей линейного программирования в биологии и их решения с Excel.

4.1.3 Краткое содержание вопроса № 4.1.3. Решение задачи об оптимизации транспортных расходов.

1. Формулировка задачи об оптимизации транспортных расходов.

**Задача об оптимизации перевозок.** Четыре отделения сельхозпредприятия  $B_1, B_2, B_3, B_4$  закупают корма у трёх поставщиков  $A_1, A_2, A_3$ . Запасы кормов у поставщиков, потребности сельхозпредприятия в кормах и стоимость перевозки единицы продукта от поставщика к потребителю даны в таблице.

Потребители		Потребители				Запасы кормов у поставщиков
		$B_1$	$B_2$	$B_3$	$B_4$	
Поставщики	$A_1$	10 $x_{11}$	0 $x_{12}$	20 $x_{13}$	11 $x_{14}$	$a_1 = 15$
	$A_2$	12 $x_{21}$	7 $x_{22}$	9 $x_{23}$	20 $x_{24}$	$a_2 = 25$
$A_3$	0 $x_{31}$	14 $x_{32}$	16 $x_{33}$	18 $x_{34}$	$a_3 = 5$	
	$b_1 = 5$	$b_2 = 15$	$b_3 = 15$	$b_4 = 10$		<b>45</b>
Потребность в кормах						

Значительную часть расходов с /х предприятия составляют именно транспортные расходы. Минимизировать расходы предприятия: составить такой план перевозок, при котором суммарные транспортные расходы будут минимальными, все запасы поставщиков будут вывезены, все потребности отделений с.\ х. предприятия будут удовлетворены.

2. Математическая модель задачи.

Ранее, в вопросе № 4 Темы 1, была построена математическая модель этой задачи

*Целевая функция:* транспортные расходы вычисляются по формуле

$$\begin{aligned}
 Z = \sum_{i=1}^3 \sum_{j=1}^4 c_{ij} \cdot x_{ij} = & 10 \cdot x_{11} + 0 \cdot x_{12} + 20 \cdot x_{13} + 11 \cdot x_{14} + \\
 & + 12 \cdot x_{21} + 7 \cdot x_{22} + 9 \cdot x_{23} + 20 \cdot x_{24} + \\
 & + 0 \cdot x_{31} + 14 \cdot x_{32} + 16 \cdot x_{33} + 18 \cdot x_{34} \rightarrow \min.
 \end{aligned}$$

*Ограничения на переменные задачи.*

Ограничения вывоза: из  $A_1$   $x_{11} + x_{12} + x_{13} + x_{14} = 15$ ,

из  $A_2$   $x_{21} + x_{22} + x_{13} + x_{14} = 25$ ,

из  $A_3$   $x_{31} + x_{32} + x_{33} + x_{34} = 5$ .

Ограничения ввоза: ввоз в  $B_1$   $x_{11} + x_{21} + x_{31} = 5$ ,

ввоз в  $B_2$   $x_{12} + x_{22} + x_{32} = 15$ ,

ввоз в  $B_3$   $x_{13} + x_{23} + x_{33} = 15$ ,

ввоз в  $B_4$   $x_{14} + x_{24} + x_{34} = 10$ .

Ограничения на знаки (значения) переменных:  $x_{ij} \geq 0$ .

Задача закрытого типа:  $a_1 + a_2 + a_3 = b_1 + b_2 + b_3 + b_4 = 45$ .

3. Компьютерные технологии исследования модели (специальной задачи) линейного программирования – транспортной задачи – с надстройкой Excel «Поиск решения».

**Решение задачи в Excel.**

Книга1 - Microsoft Excel											
Главная Вставка Разметка страницы Формулы Данные Рецензирование Вид Надстройки											
Вставка											
Стиль											
Шрифт											
Выравнивание											
	J23										
1	Поставщики	Потребители				Запасы					
2	A1	B1	B2	B3	B4						
3	10	0	20	11	15						
4	12	7	9	20	25						
5	0	14	16	18	5						
6	Потребности	5	15	15	10						
7	Поставщики										
8	A1										
9	A2										
10	A3										
11											
12											
13											

	A	B	C	D	E	F	G	H
1	Постав-		Потребители					
2	щики	$B_1$	$B_2$	$B_3$	$B_4$	Запасы		
3	$A_1$	10	0	20	11	15		
4	$A_2$	12	7	9	20	25		
5	$A_3$	0	14	16	18	5		
6	Потреб-	5	15	15	10			
7	ности							
8	$A_4$							
9	$A_2$							
10	$A_3$							
11								
12								

	A	B	C	D	E	F	G	H
1	Постав-		Потребители					
2	щики	$B_1$	$B_2$	$B_3$	$B_4$	Запасы		
3	$A_1$	10	0	20	11	15	315	
4	$A_2$	12	7	9	20	25		
5	$A_3$	0	14	16	18	5		
6	Потребн	5	15	15	10			
7		Переменные						
8	$A_1$	$x_{11}$	0	$x_{12}$	5	$x_{13}$	0	$x_{14}$
9	$A_2$	$x_{21}$	0	$x_{22}$	10	$x_{23}$	15	$x_{24}$
10	$A_3$	$x_{31}$	5	$x_{32}$	0	$x_{33}$	0	$x_{34}$
11		5	15		15	10		
12								
13								

	A1	B	C	D	E	F	G	H	I
1	Поставщики								
2		<i>Потребители</i>							
3	<i>A<sub>1</sub></i>	10	0	20	11	15	315		
4	<i>A<sub>2</sub></i>	12	7	9	20	25			
5	<i>A<sub>3</sub></i>	0	14	16	18	5			
6	<i>Потребн</i>	5	15	15	10				
7		<i>Переменные</i>							
8	<i>A<sub>1</sub></i>	<i>x<sub>11</sub></i>	0	<i>x<sub>12</sub></i>	5	<i>x<sub>13</sub></i>	0	<i>x<sub>14</sub></i>	10
9	<i>A<sub>2</sub></i>	<i>x<sub>21</sub></i>	0	<i>x<sub>22</sub></i>	10	<i>x<sub>23</sub></i>	15	<i>x<sub>24</sub></i>	0
10	<i>A<sub>3</sub></i>	<i>x<sub>31</sub></i>	5	<i>x<sub>32</sub></i>	0	<i>x<sub>33</sub></i>	0	<i>x<sub>34</sub></i>	0
11			5	15	15	10			
12									

В результате рассмотрения вопроса № 4.1. 3 студенты должны:

- освоить понятие о специальных задачах (транспортной задаче) линейного программирования и технологию их решения с Excel;
- приобрести умения и навыки создания моделей специальных задач линейного программирования (транспортной задачи) и технологии их решения с Excel.

4.1.4 Краткое содержание вопроса № 4.1.4 Решение задачи о назначениях, задачи целочисленного программирования.

1. Формулировка специальной задачи линейного программирования: задача о назначениях, задачи целочисленного программирования.

**Задача о назначениях.** Производится классификация 7 болезней по 5 категориям. Результаты тестирования каждого заболевания (уровень опасности заболевания) по каждой категории выражены в баллах по 10-балльной шкале и представлены матрицей

*	$P_1$	$P_2$	$P_3$	$P_4$	$P_5$
$S_1$	7	5	7	6	7
$S_2$	6	4	8	4	9
$S_3$	8	6	4	3	8
$S_4$	7	7	8	5	7
$S_5$	5	9	7	9	5
$S_6$	6	8	6	4	7
$S_7$	7	7	8	6	4

Определить самые опасные заболевания в каждой из 5 категорий так, чтобы сумма баллов выбранных заболеваний была наибольшей (суммарный уровень опасности выбранных заболеваний был наибольшим). Каждое заболевание может быть самым опасным только в одной категории и все категории должны быть заняты.

## 2. Математическая модель задачи.

Ранее, в вопросе № 4 Темы 1, была построена математическая модель этой задачи

*Целевая функция:* суммарный уровень опасности всех заболеваний в баллах вычисляется по формуле

$$\begin{aligned}
 Z = \sum_{i=1}^7 \sum_{j=1}^5 c_{ij} \cdot x_{ij} = & 7 \cdot x_{11} + 5 \cdot x_{12} + 7 \cdot x_{13} + 6 \cdot x_{14} + 7 \cdot x_{15} + \\
 & + 6 \cdot x_{21} + 4 \cdot x_{22} + 8 \cdot x_{23} + 4 \cdot x_{24} + 9 \cdot x_{25} + \\
 & + 8 \cdot x_{31} + 6 \cdot x_{32} + 4 \cdot x_{33} + 3 \cdot x_{34} + 8 \cdot x_{35} + \\
 & + 7 \cdot x_{41} + 7 \cdot x_{42} + 8 \cdot x_{43} + 5 \cdot x_{44} + 7 \cdot x_{45} + \\
 & + 5 \cdot x_{51} + 9 \cdot x_{52} + 7 \cdot x_{53} + 9 \cdot x_{54} + 5 \cdot x_{55} + \\
 & + 6 \cdot x_{61} + 8 \cdot x_{62} + 6 \cdot x_{63} + 4 \cdot x_{64} + 7 \cdot x_{65} + \\
 & + 7 \cdot x_{71} + 7 \cdot x_{72} + 8 \cdot x_{73} + 6 \cdot x_{74} + 4 \cdot x_{75} \rightarrow \max.
 \end{aligned}$$

*Ограничения на переменные задачи.*

*Ограничения на лидерство одного заболевания:* каждое заболевание может быть самым опасным только в одной категории

$$\text{для } S_1 \quad x_{11} + x_{12} + x_{13} + x_{14} + x_{15} = 1,$$

$$\text{для } S_2 \quad x_{21} + x_{22} + x_{23} + x_{24} + x_{25} = 1,$$

$$\text{для } S_3 \quad x_{31} + x_{32} + x_{33} + x_{34} + x_{35} = 1,$$

$$\text{для } S_4 \quad x_{41} + x_{42} + x_{43} + x_{44} + x_{45} = 1,$$

$$\text{для } S_5 \quad x_{51} + x_{52} + x_{53} + x_{54} + x_{55} = 1,$$

$$\text{для } S_6 \quad x_{61} + x_{62} + x_{63} + x_{64} + x_{65} = 1,$$

$$\text{для } S_7 \quad x_{71} + x_{72} + x_{73} + x_{74} + x_{75} = 1.$$

*Ограничения по занятости категорий:* в каждой категории может быть только один лидер

$$\text{для } P_1 \quad x_{11} + x_{21} + x_{31} + x_{41} + x_{51} + x_{61} + x_{71} = 1,$$

$$\text{для } P_2 \quad x_{12} + x_{22} + x_{32} + x_{42} + x_{52} + x_{62} + x_{72} = 1,$$

$$\text{для } P_3 \quad x_{13} + x_{23} + x_{33} + x_{43} + x_{53} + x_{63} + x_{73} = 1,$$

для  $P_4 \quad x_{14} + x_{24} + x_{34} + x_{44} + x_{54} + x_{64} + x_{74} = 1$ ,

для  $P_5 \quad x_{15} + x_{25} + x_{35} + x_{45} + x_{55} + x_{65} + x_{75} = 1$ ,

*Ограничения на знаки (значения) переменных:  $x_{ij} \geq 0$ ,  $x_{ij}$  - двоичные числа.*

Задача открытого типа: вводятся две фиктивные категории с нулевыми столбцами баллов. Матрица  $C^*$  становится квадратной.

*	$P_1$	$P_2$	$P_3$	$P_4$	$P_5$	$P_6$	$P_7$
$S_1$	7	5	7	6	7	0	0
$S_2$	6	4	8	4	9	0	0
$S_3$	8	6	4	3	8	0	0
$S_4$	7	7	8	5	7	0	0
$S_5$	5	9	7	9	5	0	0
$S_6$	6	8	6	4	7	0	0
$S_7$	7	7	8	6	4	0	0

3. Компьютерные технологии исследования модели (специальной задачи) линейного программирования – задачи о назначениях – с надстройкой Excel «Поиск решения».

БАЛЛЫ								Сумма баллов
Болезни	P1	P2	P3	P4	P5	P6	P7	42
S1	7	5	7	6	7	0	0	
S2	6	4	8	4	9	0	0	
S3	8	6	4	3	8	0	0	
S4	7	7	8	5	7	0	0	
S5	5	9	7	9	5	0	0	
S6	6	8	6	4	7	0	0	
S7	7	7	8	6	4	0	0	
ПЕРЕМЕННЫЕ								
	P1	P2	P3	P4	P5	P6	P7	
S1	0	0	0	0	0	0	1	1
S2	0	0	0	0	1	0	0	1
S3	1	0	0	0	0	0	0	1
S4	0	0	1	0	0	0	0	1
S5	0	0	0	1	0	0	0	1
S6	0	1	0	0	0	0	0	1
S7	0	0	0	0	0	1	0	1
	1	1	1	1	1	1	1	

В результате рассмотрения вопроса № 4.1.4 студенты должны:

- освоить понятие о специальных задачах линейного программирования – задаче о назначениях, задачах целочисленного программирования – и технологию их решения с Excel;
- приобрести умения и навыки создания моделей специальных задач линейного программирования – задаче о назначениях, задачах целочисленного программирования – и технологии их решения с Excel.

4.2 Краткое содержание вопроса № 4.2. Компьютерные технологии решения задач (моделей) оптимизации на графах и сетях в биологии: задача об остове минимального веса сети и её решение по алгоритму Краскала с Excel.

1. Формулировка задачи оптимизации на графах и сетях в биологии: задача об остове минимального веса сети.

Задача. Сеть G содержит 5 вершин. Расстояния между вершинами заданы таблицей

- Найти его минимальное дерево-остов (минимальное покрывающее дерево).

Таблица 2

	1	2	3	4	5
1	0	5	8	2	7
2	5	0	9	2	5
3	8	9	0	10	10
4	2	2	10	0	7
5	7	5	10	7	0

## 2. Математическая модель задачи.

Ранее, в вопросе № 4 Темы 1, была построена математическая модель этой задачи

$$5x_{12} + 8x_{13} + 2x_{14} + 7x_{15} + 9x_{23} + 2x_{24} + 5x_{25} + 10x_{34} + \\ + 10x_{35} + 7x_{45} \rightarrow \min_{x \in \Delta_\beta}$$

Множество ограничений выглядит так:

$$\left\{ \begin{array}{l} x_{12} + x_{13} + x_{14} + x_{15} \geq 1; \\ x_{12} + x_{23} + x_{24} + x_{25} \geq 1; \\ x_{13} + x_{23} + x_{34} + x_{35} \geq 1; \\ x_{14} + x_{24} + x_{34} + x_{45} \geq 1; \\ x_{15} + x_{25} + x_{35} + x_{45} \geq 1; \\ x_{12} + x_{13} + x_{14} + x_{15} + x_{23} + x_{24} + x_{25} + x_{34} + x_{35} + x_{45} = 4; \\ x_{12}, x_{13}, x_{14}, x_{15}, x_{23}, x_{24}, x_{25}, x_{34}, x_{35}, x_{45} \in \{0, 1\}. \end{array} \right.$$

## 3. Компьютерная реализация алгоритму Краскала решения задачи об остове минимального веса с Excel.

Воспользуемся надстройкой «Поиск решения», входящей в состав MS EXCEL. Расположим исходные данные на рабочем листе так, как на рис. 29. На рис. 30–31 приведены окна надстройки и ее параметров.

На рис. 32 показано состояние рабочего листа после выполнения решения с использованием надстройки. На рис. 33 изображено минимальное дерево-остов для данного графа. Заметим, что вес этого дерева в точности соответствует полученному по алгоритму Краскала весу (задача 3. 1).

	A	В	C	D	E
1	начальная вершина ребра	конечная вершина ребра	вес ребра	Переменные	Ограничения
2	1	2	5	0	=СУММ(D2:D5)
3	1	3	8	0	=СУММ(D2:D6:D8)
4	1	4	2	0	=СУММ(D3:D6:D9:D10)
5	1	5	7	0	=СУММ(D4:D7:D9:D11)
6	2	3	9	0	=СУММ(D5:D8:D10:D11)
7	2	4	2	0	=СУММ(D2:D11)
8	2	5	5	0	
9	3	4	10	0	
10	3	5	10	0	
11	4	5	7	0	
15	Целевая функция				
16	=СУММПРОИЗВ(D2:D11;C2:C11)				

Рисунок 9

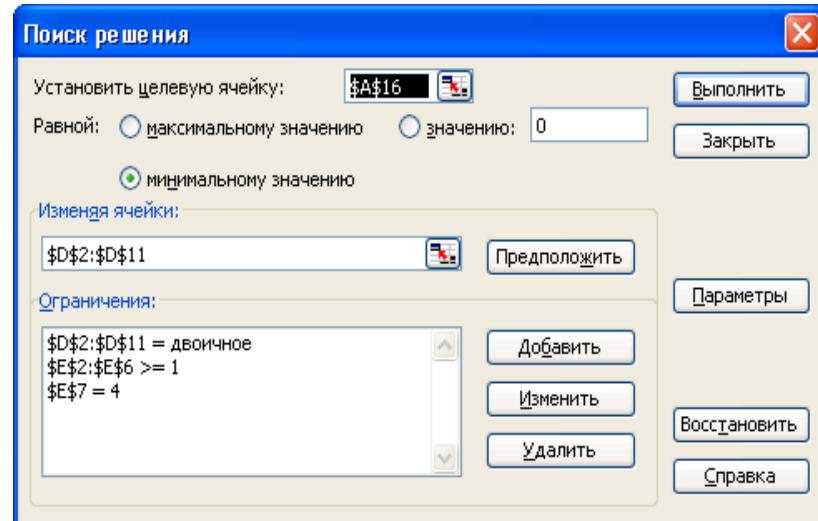


Рисунок 10

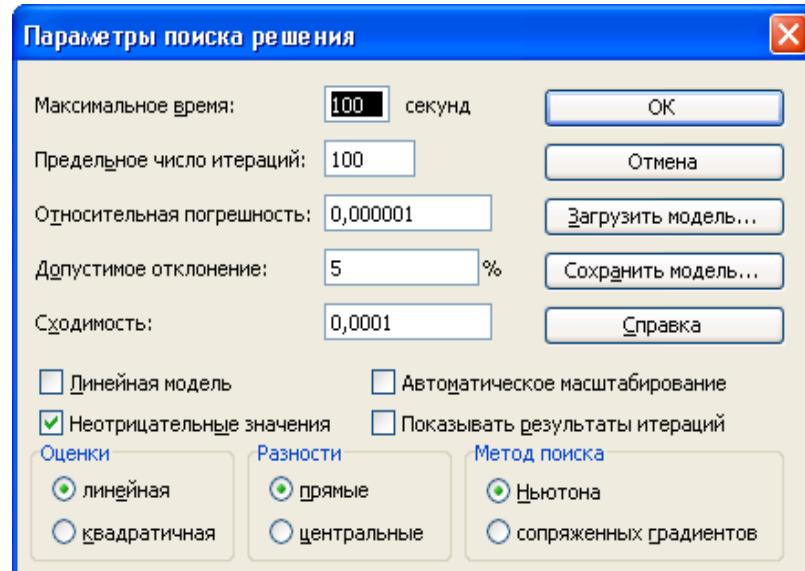


Рисунок 11

	A	B	C	D	E
1	начальная вершина ребра	конечная вершина ребра	вес ребра	Переменные	Ограничения
2		1	2	5	0
3		1	3	8	1
4		1	4	2	1
5		1	5	7	0
6		2	3	9	0
7		2	4	2	1
8		2	5	5	1
9		3	4	10	0
10		3	5	10	0
11		4	5	7	0
15	Целевая функция				
16			17		

Рисунок 12

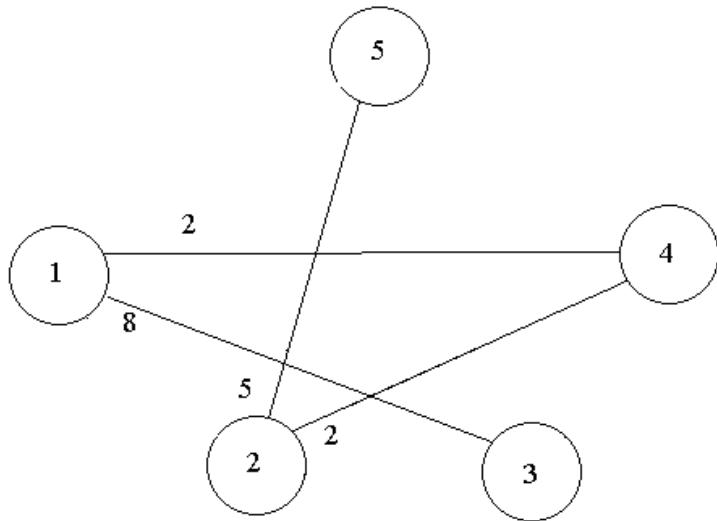


Рисунок 13

В результате рассмотрения вопроса № 4.2 студенты должны:

- освоить понятие о задачах (моделях) оптимизации на графах и сетях в биологии, алгоритмы их решения и компьютерную реализацию на примере задачи об остове минимального веса сети с Excel;
- приобрести умения и навыки создания моделей оптимизации на графах и сетях в биологии, использования алгоритмов их решения и компьютерной реализации на примере задачи об остове минимального веса сети с Excel.

## 2. Методические рекомендации по выполнению курсовой работы (проекта)

Курсовые работы (проекты) по дисциплине не предусмотрены учебным планом.

## 3. Методические рекомендации по выполнению индивидуальных домашних заданий (контрольных работ)

Индивидуальные домашние задания по дисциплине не предусмотрены рабочей программой дисциплины.