

МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ
РЕСПУБЛИКИ КАЗАХСТАН

ИННОВАЦИОННЫЙ ЕВРАЗИЙСКИЙ УНИВЕРСИТЕТ
МАГИСТРАТУРА

Кафедра «Информатики и вычислительной техники»

Магистерская диссертация

**ИССЛЕДОВАНИЕ И РАЗРАБОТКА МОДЕЛЕЙ
ПРОГНОЗИРОВАНИЯ ПОСТУПЛЕНИЯ ВЫПУСКНИКОВ ШКОЛ
ПАВЛОДАРСКОЙ ОБЛАСТИ В ВЫСШИЕ И СРЕДНИЕ УЧЕБНЫЕ
ЗАВЕДЕНИЯ**

6N0111 «Информатика»

Исполнитель _____ Хаирова С. А.
(подпись, дата)

Научный руководитель

к.т.н., профессор _____ Зайцева Н. М.
(подпись, дата)

Допущена к защите:

Зав. кафедрой «ИВТ»
к.т.н., профессор _____ Деревягин С.И.
(подпись, дата)

Павлодар, 2008

АНДАТПА

Бұл диссертациялық жұмыс Павлодар облысының мектеп түлектерінің жоғарғы және орта білім мекемелеріне түсуін болжам моделін жасауына және талдауына арналған.

Жұмыс барысында жүргізілген талдаулар көмегімен түлектердің оқу орны таңдауына әсер ететін басты факторлар айқындалып, Павлодар облысының түлектерінің оқу орнына түсуін болжамдайтын компьютерлік моделі жасалған.

Жұмыстың тәжірибелік маңызы диссертациялық талдауды теориялық және ғылым методикалық қорытындылары аймақтағы жоғарғы және орта білімнің оқу стратегиясына әсер етуіне арналған.

АННОТАЦИЯ

Данная диссертационная работа посвящена исследованию и разработки моделей прогнозирования поступления выпускников школ Павлодарской области в высшие и средние учебные заведения

В работе на основе проведенных исследований выявлены основные факторы, влияющие на выбор выпускниками учебного заведения и разработана компьютерная модель прогнозирования поступления выпускников Павлодарской области.

Практическая значимость работы состоит в том, что полученные теоретические и научно-методические результаты диссертационного исследования предназначены способствовать стратегии развития высшего и среднего образования в регионе.

ANNOTATION

The given dissertational work is devoted to research and development of models of forecasting of receipt of graduates of schools of the Pavlodar area in the maximum and average educational institutions.

In work are revealed the major factors influencing a choice by graduates of an educational institution and the computer model of forecasting of receipt of graduates of the Pavlodar area is developed which made on the basis of the lead researches.

The practical importance of work consists of received theoretical and scientifically-methodical.

РЕФЕРАТ

Магистерская диссертация выполнена в объеме 125 страницы и содержит 35 иллюстраций, 47 использованных источников.

Ключевые слова: система образования, высшие учебные заведения, средние учебные заведения, прогнозирование, корреляционный и дисперсионный анализы, однофакторные и многофакторные модели прогнозирования, математическое моделирование, информационные технологии в образовании, компьютерная модель.

Цель работы: построение математической модели прогнозирования поступления выпускников школ в высшие и средние учебные заведения.

Объект исследования: выпускники школ в системе образования Республики Казахстан.

Предмет исследования: факторы, влияющие на поступление выпускников школ в высшие и средние учебные заведения.

Методы исследования: корреляционный и дисперсионный анализы, системный анализ, статистические исследования, анализ статистического прогнозирования и имитационное моделирование.

Новизна: математическая модель, которая позволяет прогнозировать поступление выпускников школ в высшие и средние учебные заведения.

Полученные результаты: разработана компьютерная модель прогнозирования поступления выпускников школ региона в высшие и средние учебные заведения, предназначенная для планирования образовательного процесса в высших и средних учебных заведениях Павлодарской области.

СОДЕРЖАНИЕ

	ВВЕДЕНИЕ	8
1	ИССЛЕДОВАНИЕ ФАКТОРОВ, ВЛИЯЮЩИХ НА ПОСТУПЛЕНИЕ ВЫПУСКНИКОВ ШКОЛ В СРЕДНИЕ И ВЫСШИЕ УЧЕБНЫЕ ЗАВЕДЕНИЯ	11
1.1	Анализ состояния системы образования Республики Казахстан	11
1.2	Основные социальные факторы, влияющие на выбор выпускниками учебного заведения	22
1.3	Анализ данных социального опроса	33
1.4	Анализ социально-экономических данных региона Павлодарской области	34
1.5	Выводы по первой главе	53
2	ПОСТРОЕНИЕ МОДЕЛИ ПРОГНОЗИРОВАНИЯ ПОСТУПЛЕНИЯ ВЫПУСКНИКОВ В ВЫСШИЕ И СРЕДНИЕ УЧЕБНЫЕ ЗАВЕДЕНИЯ	54
2.1	Анализ методов построения моделей прогнозирования	54
2.1.1	Однофакторная модель	58
2.1.2	Многофакторная модель	67
2.1.3	Критерии адекватности модели	73
2.2	Основные этапы процесса моделирования	76
2.3	Постановка задачи моделирования поступления выпускников школ в высшие и средние учебные заведения	83
2.4	Выводы по второй главе	87
3	РЕАЛИЗАЦИЯ МОДЕЛИ НА ЭВМ И ЕЕ ИССЛЕДОВАНИЕ	88
3.1	Проектирование таблиц базы данных	88
3.2	Выбор инструментария разработки программного продукта	91
3.3	Разработка интерфейса	93
3.5	Выводы по третьей главе	99
4	ПРАКТИЧЕСКАЯ АПРОБАЦИЯ	100
4.1	Апробация и тестирование компьютерной модели прогнозирования	100
4.2	Выводы по четвертой главе	107
	ЗАКЛЮЧЕНИЕ	108
	СПИСОК ИСПОЛЬЗУЕМЫХ ИСТОЧНИКОВ	109
	ПРИЛОЖЕНИЕ А «Анкета социального опроса»	114
	ПРИЛОЖЕНИЕ Б «Листинг программы»	116

2 ПОСТРОЕНИЕ МАТЕМАТИЧЕСКОЙ МОДЕЛИ ПРОГНОЗИРОВАНИЯ ПОСТУПЛЕНИЯ ВЫПУСКНИКОВ В ВЫСШИЕ И СРЕДНИЕ УЧЕБНЫЕ ЗАВЕДЕНИЯ

2.1 Анализ методов построения моделей прогнозирования

В современной практической деятельности, в различных областях науки, техники, технологий и экономики очень часто приходится сталкиваться с необходимостью выявления и оценки влияния отдельных факторов на изменчивость качественного признака. Если анализ применяется для изучения влияния нескольких факторов на некоторый признак, то его называют многофакторным. Среди многообразия методов многофакторного анализа можно выделить корреляционный, регрессионный и дисперсионный анализ.

Корреляционный анализ – это выявление статистической зависимости между случайными величинами, не имеющими строго функционального характера, при которой изменение одной из случайных величин приводит к изменению математического ожидания другой. Его задача – это количественное определение тесноты связей между множеством факторных и результативных признаков. Первоначально исследования корреляции проводились в биологии, а позднее распространились и на другие области, в том числе на социально-экономическую.

Регрессионный анализ называют основным методом современной математической статистики для выявления неявных и завуалированных связей между данными наблюдений. Корреляция и регрессия тесно связаны между собой: первая оценивает силу (тесноту) статистической связи, вторая исследует ее форму. И корреляция, и регрессия служат для установления соотношений между явлениями и для определения наличия или отсутствия связи между ними.[10, с.11-14]

Дисперсионный анализ (от латинского *Dispersio* – рассеивание) – статистический метод, позволяющий анализировать влияние различных факторов на исследуемую переменную. Целью дисперсионного анализа

является проверка значимости различия между средними с помощью сравнения дисперсий. Дисперсию измеряемого признака разлагают на независимые слагаемые, каждое из которых характеризует влияние того или иного фактора или их взаимодействия. Последующее сравнение таких слагаемых позволяет оценить значимость каждого изучаемого фактора, а также их комбинации.

При истинности нулевой гипотезы (о равенстве средних в нескольких группах наблюдений, выбранных из генеральной совокупности), оценка дисперсии, связанной с внутригрупповой изменчивостью, должна быть близкой к оценке межгрупповой дисперсии.

При проведении исследования рынка часто встает вопрос о сопоставимости результатов. Например, проводя опросы по поводу потребления какого-либо товара в различных регионах страны, необходимо сделать выводы, на сколько данные опроса отличаются или не отличаются друг от друга. Сопоставлять отдельные показатели не имеет смысла и поэтому процедура сравнения и последующей оценки производится по некоторым усредненным значениям и отклонениям от этой усредненной оценки. Изучается вариация признака. За меру вариации может быть принята дисперсия.

На практике часто возникают задачи более общего характера – задачи проверки существенности различий средних выборочных нескольких совокупностей. Например, требуется оценить влияние различного сырья на качество производимой продукции, решить задачу о влиянии количества удобрений на урожайность сельхоз продукции.

Иногда дисперсионный анализ применяется, чтобы установить однородность нескольких совокупностей (дисперсии этих совокупностей одинаковы по предположению; если дисперсионный анализ покажет, что и математические ожидания одинаковы, то в этом смысле совокупности однородны). Однородные же совокупности можно объединить в одну и тем самым получить о ней более полную информацию, следовательно, и более надежные выводы.[11, с.14-16]

Основные понятия дисперсионного анализа

В процессе наблюдения за исследуемым объектом качественные факторы произвольно или заданным образом изменяются. Конкретная реализация фактора (например, определенный температурный режим, выбранное оборудование или материал) называется уровнем фактора или способом обработки. Модель дисперсионного анализа с фиксированными уровнями факторов называют моделью первого порядка, модель со случайными факторами - моделью второго порядка. Существуют модели дисперсионного анализа в которых часть факторов имеют фиксированное количество уровней, а часть случайное. Благодаря варьированию фактора можно исследовать его влияние на величину отклика. В настоящее время общая теория и методика проведения дисперсионного анализа разработана для моделей первого порядка, то есть для моделей с фиксированным уровнем факторов.

В зависимости от количества факторов, определяющих вариацию результативного признака, дисперсионный анализ подразделяют на однофакторный и многофакторный. Основными схемами организации исходных данных с двумя и более факторами являются перекрестная и иерархическая классификация. Перекрестная классификация, характерная для моделей первого порядка, в которых каждый уровень одного фактора сочетается с каждой градацией другого фактора. Иерархическая или иначе ее называют гнездовая классификация, характерная для модели второго порядка, в которой каждому случайному, наудачу выбранному значению одного фактора соответствует свое подмножество значений второго фактора.

Если одновременно исследуется зависимость отклика от качественных и количественных факторов, то есть факторов смешанной природы, то используется ковариационный анализ.

При обработке данных наиболее теоретически разработанной и поэтому распространенной считается модель первого порядка с фиксированным уровнем факторов. В модели дисперсионного анализа с фиксированными эффектами исследователь намеренно устанавливает строго определенные уровни изучаемого фактора. Термин фиксированный эффект в данном контексте имеет

тот смысл, что самим исследователем фиксируется количество уровней фактора и различия между ними. При повторении эксперимента он или другой исследователь выберет те же самые уровни фактора.

При проведении дисперсионного анализа должны выполняться ряд статистических допущений. Независимо от уровня фактора, величины отклика имеют нормальный, Гауссовский закон распределения и одинаковую дисперсию. Такое равенство дисперсий называется гомогенностью. Таким образом, изменение способа обработки сказывается лишь на положении случайной величины отклика, которое характеризуется средним значением или медианой. Поэтому все наблюдения отклика принадлежат сдвиговому семейству нормальных распределений.

Считается, что методика проведения дисперсионного анализа является «робастной». Этот термин, используемый статистиками, означает, что данные допущения могут быть в некоторой степени нарушены, но несмотря на это, методику проведения дисперсионного анализа можно использовать.

В основе дисперсионного анализа лежит разделение дисперсии на части или компоненты. Вариацию, обусловленную влиянием фактора, положенного в основу группировки, характеризует межгрупповая дисперсия, обозначаемая через σ^2 . Она является мерой вариации частных средних по группам, которые обозначаются как $\bar{x}_j$ вокруг общей средней и определяется по формуле:

$$\sigma^2 = \frac{\sum_{j=1}^k (\bar{x}_j - \bar{x})^2 \times n_j}{\sum_{j=1}^k n_j}, \quad (1)$$

где,

σ^2 - межгрупповая дисперсия

$\bar{x}_j$ - вариации частных средних по группам

$\bar{x}$ - общая средняя

k - число групп;

n_j - число единиц в j -ой группе;

$\bar{x}_j$ - частная средняя по j -ой группе;

$\bar{x}$ - общая средняя по совокупности единиц.

Вариацию, обусловленную влиянием прочих факторов, характеризует в каждой группе внутригрупповая дисперсия, обозначаемая как σ_j^2 :

$$\sigma_j^2 = \frac{\sum_{i=1}^n (x_{ij} - \bar{x}_j)^2}{n} \quad (2)$$

Между общей дисперсией, которая обозначается через σ_0^2 , внутригрупповой дисперсией σ^2 и межгрупповой дисперсией $\bar{\sigma}^2$ существует соотношение:

$$\sigma_0^2 = \bar{\sigma}^2 + \sigma^2. \quad (3)$$

Внутригрупповая дисперсия объясняет влияние неучтенных при группировке факторов, а межгрупповая дисперсия объясняет влияние факторов группировки на среднее значение по группе.

Вышеизложенный материал описывает общую схему дисперсионного анализа. Для раскрытия методики проведения дисперсионного анализа, необходимо более подробно остановиться на описании случаев многофакторного дисперсионного анализа. Для начала рассмотрим частный случай однофакторного дисперсионного анализа.[13, с.45-46]

2.1.1 Однофакторная модель. Под однофакторным анализом принимают анализ влияния одного фактора на изменчивость одного признака. При этом отдельные интервалы значений исследуемого фактора называются его уровнями.

Однофакторная дисперсионная модель имеет вид:

$$x_{ij} = F_j + \varepsilon_{ij}, \quad (4)$$

где,

x_{ij} – значение исследуемой переменной, полученной на i -м уровне

фактора ($i=1,2,\dots,t$) с j -м порядковым номером ($j=1,2,\dots,n$);

F_i – эффект, обусловленный влиянием i -го уровня фактора;

ε_{ij} – случайная компонента, или возмущение, вызванное влиянием

неконтролируемых факторов, т.е. вариацией переменной внутри отдельного уровня.

Пусть, например, необходимо выяснить, имеются ли существенные различия между партиями изделий по некоторому показателю качества, т.е. проверить влияние на качество одного фактора - партии изделий. Если включить в исследование все партии сырья, то влияние уровня такого фактора систематическое, то есть является моделью с фиксированным уровнем факторов, а полученные выводы применимы только к тем отдельным партиям, которые привлекались при исследовании.

Пусть имеется m партий изделий. Из каждой партии отобрано соответственно $n_1, n_2, \dots, n_m$ изделий (для простоты полагается, что $n_1=n_2=\dots=n_m=n$). Значения показателя качества этих изделий представлены в матрице наблюдений:

$$\begin{pmatrix} X_{11} & X_{12} & \dots & X_{1n} \\ X_{21} & X_{22} & \dots & X_{2n} \\ \dots & \dots & \dots & \dots \\ X_{m1} & X_{m2} & \dots & X_{mn} \end{pmatrix} = (x_{ij}), (i = 1, 2, \dots, m; j = 1, 2, \dots, n). \quad (5)$$

Необходимо проверить существенность влияния партий изделий на их качество.

Если полагать, что элементы строк матрицы наблюдений – это численные

значения случайных величин $X_1, X_2, \dots, X_m$, выражающих качество изделий и имеющих нормальный закон распределения с математическими ожиданиями соответственно $a_1, a_2, \dots, a_m$ и одинаковыми дисперсиями σ^2 , то данная задача сводится к проверке нулевой гипотезы $H_0: a_1 = a_2 = \dots = a_m$, осуществляемой в дисперсионном анализе.

Усреднение по какому-либо индексу обозначено звездочкой (или точкой) вместо индекса, тогда средний показатель качества изделий i -й партии, или групповая средняя для i -го уровня фактора, примет вид:

$$\bar{X}_{i*} = \frac{\sum_{j=1}^n X_{ij}}{n}, \quad (6)$$

где,

$\bar{X}_{i*}$ – среднее значение по столбцам;

$\bar{X}_{ij}$ – элемент матрицы наблюдений;

n – объем выборки.

А общая средняя:

$$\bar{X}_{..} = \frac{\sum_{i=1}^m \sum_{j=1}^n X_{ij}}{mn} = \frac{\sum_{i=1}^m \bar{X}_{i*}}{m}. \quad (7)$$

Сумма квадратов отклонений наблюдений x_{ij} от общей средней $\bar{X}_{..}$ равна:

$$\sum_{i=1}^m \sum_{j=1}^n (x_{ij} - \bar{x}_{**})^2 = \sum_{i=1}^m \sum_{j=1}^n \left(\bar{x}_{i\cdot} - \bar{x}_{**} \right)^2 + \sum_{i=1}^m \sum_{j=1}^n \left(x_{ij} - \bar{x}_{i\cdot} \right)^2 +$$

(8)

$$+ 2 \sum_{i=1}^m \sum_{j=1}^n \left(x_{ij} - \bar{x}_{i\cdot} \right) \left(\bar{x}_{i\cdot} - \bar{x}_{**} \right)$$

или рассчитывается по формуле:

$$Q = Q_1 + Q_2 + Q_3, \quad (9)$$

где,

Q – общая сумма квадратов отклонений

Q_1 – сумма квадратов отклонений групповых средних от общей средней

Q_2 – сумма квадратов отклонений наблюдений групповых средних или внутригрупповая (остаточная) сумма квадратов отклонений

Последнее слагаемое равно нулю

$$\sum_{i=1}^n (X_{ij} - \bar{X}_{i*}) = 0. \quad (10)$$

так как сумма отклонений значений переменной от ее средней равна нулю, т.е.

$$\sum_{i=1}^n (X_{ij} - \bar{X}_{i*})^2 = 0. \quad (11)$$

Первое слагаемое можно записать в виде:

$$Q_1 = \sum_{i=1}^m \sum_{j=1}^n (\bar{X}_{i*} - \bar{X}_{**})^2 = n \sum_{i=1}^m (\bar{X}_{i*} - \bar{X}_{**})^2. \quad (12)$$

В результате получается тождество:

$$Q = Q_1 + Q_2, \quad (13)$$

где,

$Q = \sum_{i=1}^m \sum_{j=1}^n (x_{ij} - \bar{x}_{**})^2$ - общая, или полная, сумма квадратов отклонений;

$Q_1 = n \sum_{i=1}^m (\bar{x}_{i*} - \bar{x}_{**})^2$ - сумма квадратов отклонений групповых средних от общей средней, или межгрупповая сумма квадратов отклонений;

$Q_2 = \sum_{i=1}^m \sum_{j=1}^n (x_{ij} - \bar{x}_{i*})^2$ - сумма квадратов отклонений наблюдений от групповых средних, или внутригрупповая сумма квадратов отклонений.

В формуле 13 заключена основная идея дисперсионного анализа. Применительно к рассматриваемой задаче формула 13 показывает, что общая вариация показателя качества, измеренная суммой Q , складывается из двух компонент – Q_1 и Q_2 , характеризующих изменчивость этого показателя между партиями (Q_1) и изменчивость внутри партий (Q_2), характеризующих одинаковую для всех партий вариацию под воздействием неучтенных факторов.

В дисперсионном анализе анализируются не сами суммы квадратов отклонений, а так называемые средние квадраты, являющиеся несмещенными оценками соответствующих дисперсий, которые получаются делением сумм квадратов отклонений на соответствующее число степеней свободы.

Число степеней свободы определяется как общее число наблюдений минус число связывающих их уравнений. Поэтому для среднего квадрата s_1^2 , являющегося несмещенной оценкой межгрупповой дисперсии, число степеней свободы $k_1=m-1$, так как при его расчете используются m групповых средних, связанных между собой одним уравнением, показанном в формуле 7. А для среднего квадрата s_2^2 , являющегося несмещенной оценкой внутригрупповой дисперсии, число степеней свободы $k_2=mn-m$, так как при ее расчете используются все количество наблюдений.

Таким образом средние квадраты рассчитывается по формулам:

$$S_1^2 = Q_1/(m-1), \quad (14)$$

$$S_2^2 = Q_2/(mn-m). \quad (15)$$

Если найти математические ожидания средних квадратов S_1^2 и S_2^2 , подставить в их формулы выражения один, то через параметры модели получится:

$$MS_1 = \frac{n}{m-1} M \left[\sum_{i=1}^m (\mu + F_i + \epsilon_{ij} - \mu - F_* - \epsilon_{*j})^2 \right] =$$

$$= \frac{n}{m-1} M \left[\sum_{i=1}^m (F_i - F_*)^2 + (\epsilon_{ij} - \epsilon_{*j})^2 \right] = \frac{n}{m-1} M \left[\sum_{i=1}^m (F_i - F_*)^2 \right] +$$

$$+ \frac{n}{m-1} M \left[\sum_{i=1}^m \sum_{j=1}^n (\epsilon_{ij} - \epsilon_{*j})^2 \right] =$$

$$= \frac{n}{m-1} M \left[\sum_{i=1}^m (F_i - F_*)^2 \right] + \sigma^2. \quad (16)$$

А с учетом свойств математического ожидания формулы примут вид:

$$M \left[2 \sum_{i=1}^m (F_i - F_*) (\varepsilon_{i*} - \varepsilon_{**}) \right] = 0, \text{ а}$$

$$- \frac{2}{m} \sum_{i=1}^m (F_i - F_*) (\varepsilon_{i*} - \varepsilon_{**}) = \frac{1}{m} M \left[\frac{\sum_{i=1}^m \sum_{j=1}^n (\varepsilon_{ij} - \varepsilon_{**})^2}{n-1} \right] = \frac{1}{m} \sum_{i=1}^m \sigma_{\varepsilon_{ij}}^2 = \frac{1}{m} \sum_{i=1}^m \sigma^2 = \frac{1}{m} m \sigma^2 = \sigma^2. \quad (17)$$

Для модели первого порядка с фиксированными уровнями фактора $F_i (i=1,2,...,m)$ – величины неслучайные, поэтому:

$$M(S_1^2) = n \sum_{i=1}^m (F_i - F_*)^2 / (m-1) + \sigma^2. \quad (18)$$

Гипотеза H_0 примет вид $F_i = F_* (i = 1,2,...,m)$, т.е. влияние всех уровней фактора одно и то же. В случае справедливости этой гипотезы:

$$M(S_1^2) = M(S_2^2) = \sigma^2. \quad (19)$$

В таблице 26 представлена обобщенная методика вычисления значений, с помощью дисперсионного анализа.

Таблица 26 – Базовая таблица дисперсионного анализа

Компоненты дисперсии	Сумма квадратов	Число степеней свободы	Средний квадрат	Математическое ожидание среднего квадрата
Межгрупповая	$Q_1 = n \sum_{i=1}^m (\bar{x}_{i*} - \bar{x}_{**})^2$	$m-1$	$S_1^2 = Q_1/(m-1)$	$M(S_1^2) = \begin{cases} \frac{n}{m-1} \sum_{i=1}^m (F_i - F_*)^2 + \sigma^2 (\text{модель I}) \\ n\sigma_F^2 + \sigma^2 (\text{модель II}) \end{cases}$
Внутригрупповая	$Q_2 = \sum_{i=1}^m \sum_{j=1}^n (x_{ij} - \bar{x}_{i*})^2$	$mn-m$	$S_2^2 = Q_2/(mn-m)$	$M(S_2^2) = \sigma^2$
Общая	$Q = \sum_{i=1}^m \sum_{j=1}^n (x_{ij} - \bar{x}_{**})^2$	$mn-1$		

Гипотеза H_0 примет вид $\sigma_F^2 = 0$, а дисперсия будет рассчитана по формуле:

$$M(S_1^2) = M(S_2^2) = \sigma^2. \quad (20)$$

В однофакторном комплексе средние квадраты S_1^2 и S_2^2 , являются несмещенными и независимыми оценками одной и той же дисперсии σ^2 , поэтому проверка нулевой гипотезы H_0 свелась к проверке существенности различия несмещенных выборочных оценок S_1^2 и S_2^2 дисперсии σ^2 .

Гипотеза H_0 отвергается, если фактически вычисленное значение статистики $F = S_1^2 / S_2^2$ больше критического $F_{\alpha; K_1; K_2}$, определенного на уровне значимости α при числе степеней свободы $k_1 = m-1$ и $k_2 = mn-m$, и принимается, если $F < F_{\alpha; K_1; K_2}$.

F- распределение Фишера (для $x > 0$) имеет следующую функцию плотности (для $\nu = 1, 2, \dots$; $\omega = 1, 2, \dots$):

$$f(x) = \frac{\tilde{A}[(\nu + \omega)/2]}{\tilde{A}(\nu) \tilde{A}(\omega)} * (\nu \nu / \omega^{\nu/2} * x^{(\nu/2-1)} * (1 + [\nu \nu / \omega * x])^{-(\nu+\omega)/2}), \quad (21)$$

где,

ν, ω - степени свободы;

Γ - гамма-функция.

Применительно к данной задаче опровержение гипотезы H_0 означает наличие существенных различий в качестве изделий различных партий на рассматриваемом уровне значимости.

Для вычисления сумм квадратов Q_1, Q_2, Q часто бывает удобно использовать следующие формулы, при которых не обязательно находить средние квадраты отклонений от группового среднего:

$$Q_1 = \frac{\sum_{i=1}^m \left[\sum_{j=1}^n x_{ij} \right]^2}{n} - \frac{\left[\sum_{i=1}^m \sum_{j=1}^n x_{ij} \right]^2}{mn}, \quad (22)$$

$$Q_2 = \sum_{i=1}^m \sum_{j=1}^n x_{ij}^2 - \frac{\sum_{i=1}^m \left[\sum_{j=1}^n x_{ij} \right]^2}{n}, \quad (23)$$

$$Q = \sum_{i=1}^m \sum_{j=1}^n x_{ij}^2 - \frac{\left[\sum_{i=1}^m \sum_{j=1}^n x_{ij} \right]^2}{mn}. \quad (24)$$

Таким образом, процедура однофакторного дисперсионного анализа состоит в проверке гипотезы H_0 о том, что имеется одна группа однородных экспериментальных данных против альтернативы о том, что таких групп больше, чем одна. Под однородностью понимается одинаковость средних значений и дисперсий в любом подмножестве данных. При этом дисперсии могут быть как известны, так и неизвестны заранее. Если имеются основания полагать, что известная или неизвестная дисперсия измерений одинакова по всей совокупности данных, то задача однофакторного дисперсионного анализа сводится к исследованию значимости различия средних в группах данных.[14, с.23]

2.1.1 Многофакторная модель. Следует сразу же отметить, что принципиальной разницы между многофакторным и однофакторным анализом нет. Многофакторный анализ не меняет общую логику дисперсионного анализа, а лишь несколько усложняет ее, поскольку, кроме учета влияния на зависимую переменную каждого из факторов по отдельности, следует оценивать и их совместное действие. Таким образом проведение многофакторного дисперсионного анализа, касается в основном возможности оценить межфакторное взаимодействие, но также остается возможность оценивать влияние каждого фактора в отдельности. Процедура многофакторного дисперсионного анализа в варианте ее компьютерного использования несомненно более экономична, поскольку всего за один запуск решает сразу две задачи: оценивается влияние каждого из факторов и их взаимодействие.

Общая схема двухфакторного анализа, данные которого обрабатываются методом дисперсионного анализа с двумя влияющими на факторами – А и В, представлена на рисунке 21.

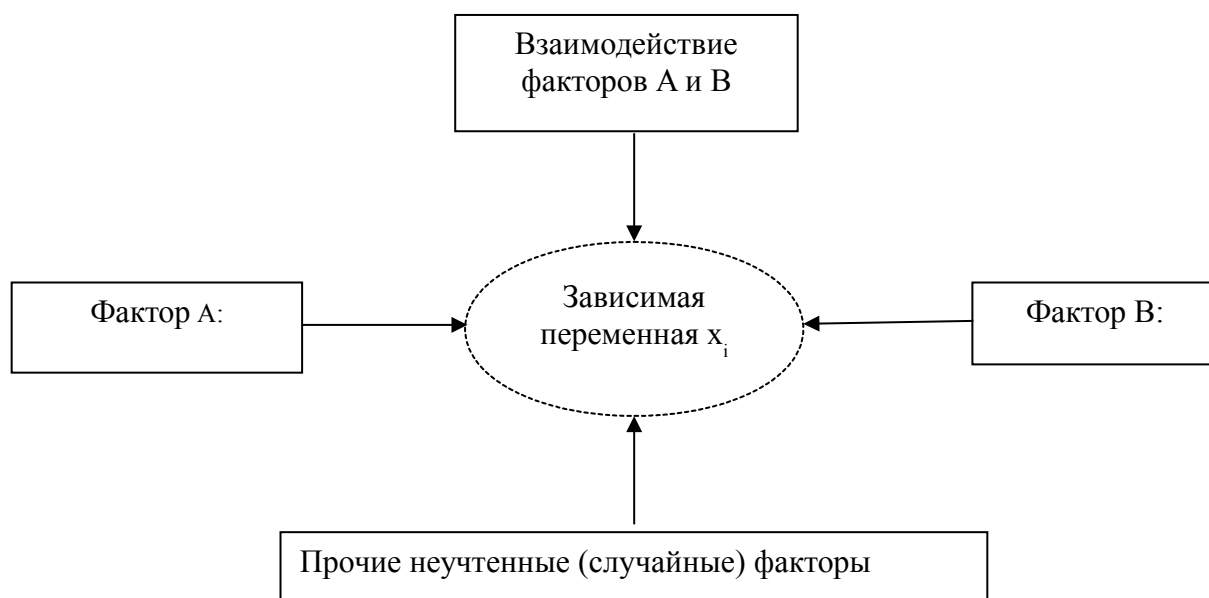


Рисунок 21 – Схема двухфакторного эксперимента

Данные, подвергаемые многофакторному дисперсионному анализу, часто обозначают в соответствии с количеством факторов и их уровней.

Предположив, что в рассматриваемой задаче о качестве различных m

партий изделия изготавливались на разных t станках и требуется выяснить, имеются ли существенные различия в качестве изделий по каждому фактору - A - партия изделий; B - станок.

В результате получается переход к задаче двухфакторного дисперсионного анализа.

Все данные представлены в таблице 2, в которой по строкам - уровни A_i фактора A , по столбцам — уровни B_j фактора B , а в соответствующих ячейках, таблицы находятся значения показателя качества изделий x_{ijk} ($i=1,2,\dots,m$; $j=1,2,\dots,l$; $k=1,2,\dots,n$).

Таблица 27 – Показатели качества изделий

	B_1	B_2	...	B_j	...	B_l
A 1	$x_{111},$..., x_{11k}	$x_{121},$..., x_{12k}	...	$x_{1j1}, \dots, x_{1jk}$	...	$x_{1l1},$..., x_{1lk}
A 2	$x_{211},$..., x_{21k}	$x_{221},$..., x_{22k}	...	$x_{2j1}, \dots, x_{2jk}$	...	$x_{2l1},$..., x_{2lk}
...	...	...	...	...	...	...
A_i	$x_{i11},$..., x_{i1k}	$x_{i21},$..., x_{i2k}	...	$x_{ij1}, \dots, x_{ijk}$	...	$x_{jl1},$..., x_{jlk}
...	...	...	...	...	...	...
A m	$x_{m11},$..., x_{m1k}	$x_{m21},$..., x_{m2k}	...	$x_{mj1}, \dots, x_{mjk}$	...	$x_{ml1},$..., x_{mlk}

Двухфакторная дисперсионная модель примет вид:

$$x_{ijk} = \mu + F_i + G_j + I_{ij} + \varepsilon_{ijk}, \quad (25)$$

где,

x_{ijk} - значение наблюдения в ячейке ij с номером k ;

μ - общая средняя;

F_i - эффект, обусловленный влиянием i -го уровня фактора A ;

G_j - эффект, обусловленный влиянием j -го уровня фактора B ;

I_{ij} - эффект, обусловленный взаимодействием двух факторов, т.е.

отклонение от средней по наблюдениям в ячейке ij от суммы
 первых трех слагаемых в модели (25);
 ε_{ijk} - возмущение, обусловленное вариацией переменной внутри
 отдельной ячейки.

Предполагается, что ε_{ijk} имеет нормальный закон распределения $N(0; \sigma^2)$, а
 все математические ожидания F^* , G^* , I_i^* , I_j^* равны нулю.

Групповые средние находятся по формулам:

в ячейке:

$$\bar{X}_{ij*} = \frac{\sum_{k=1}^n X_{ijk}}{n}, \quad (26)$$

по строке:

$$\bar{X}_{i**} = \frac{\sum_{j=1}^l \bar{X}_{ij*}}{l}, \quad (27)$$

по столбцу:

$$\bar{X}_{*j*} = \frac{\sum_{i=1}^m \bar{X}_{ij*}}{m}, \quad (28)$$

общая средняя:

$$\bar{X}_{***} = \frac{\sum_{i=1}^m \sum_{j=1}^l \bar{X}_{ij*}}{ml}. \quad (29)$$

В таблице 28 представлен общий вид вычисления значений, с помощью
 дисперсионного анализа.[14, с.45-47]

Таблица 28 – Базовая таблица дисперсионного анализа

Компоненты дисперсии	Сумма квадратов	Число степеней свободы	Средние квадраты
Межгрупповая (фактор А)	$Q_1 = n \sum_{i=1}^m (\bar{x}_{i**} - \bar{x}_{***})^2$	$m-1$	$S_1^2 = \frac{Q_1}{m-1}$
Межгрупповая (фактор В)	$Q_2 = mn \sum_{j=1}^l (\bar{x}_{*j*} - \bar{x}_{***})^2$	$l-1$	$S_2^2 = \frac{Q_2}{l-1}$
Взаимодействие	$Q_3 = n \sum_{i=1}^m \sum_{j=1}^l (\bar{x}_{ij*} - \bar{x}_{i**} - \bar{x}_{*j*} + \bar{x}_{***})^2$	$(m-1)(l-1)$	$S_3^2 = \frac{Q_3}{(m-1)(l-1)}$
Остаточная	$Q_4 = \sum_{i=1}^m \sum_{j=1}^l \sum_{k=1}^n (x_{ijk} - \bar{x}_{ij*})^2$	$mln - ml$	$S_4^2 = \frac{Q_4}{mln - ml}$
Общая	$Q = \sum_{i=1}^m \sum_{j=1}^l \sum_{k=1}^n (x_{ijk} - \bar{x}_{***})^2$	$mln - 1$	

Проверка нулевых гипотез H_A , H_B , H_{AB} об отсутствии влияния на рассматриваемую переменную факторов А, В и их взаимодействия АВ осуществляется сравнением отношений $\frac{S_1^2}{S_4^2}$, $\frac{S_2^2}{S_4^2}$, $\frac{S_3^2}{S_4^2}$ с соответствующими табличными значениями F – критерия Фишера – Снедекора.

Если $n=1$, т.е. при одном наблюдении в ячейке, то не все нулевые гипотезы могут быть проверены так как выпадает компонента Q_3 из общей суммы квадратов отклонений, а с ней и средний квадрат S_3^2 , так как в этом случае не может быть речи о взаимодействии факторов.

С точки зрения техники вычислений для нахождения сумм квадратов Q_1 , Q_2 , Q_3 , Q_4 , Q целесообразнее использовать следующие формулы:

$$Q_1 = \frac{\sum_{i=1}^m \left[\sum_{j=1}^l \sum_{k=1}^n x_{ijk} \right]^2}{\ln} - \frac{\left[\sum_{i=1}^m \sum_{j=1}^l \sum_{k=1}^n x_{ijk} \right]^2}{m \ln}, \quad (30)$$

$$Q_2 = \frac{\sum_{j=1}^l \left[\sum_{i=1}^m \sum_{k=1}^n x_{ijk} \right]^2}{mn} - \frac{\left[\sum_{i=1}^m \sum_{j=1}^l \sum_{k=1}^n x_{ijk} \right]^2}{m \ln}, \quad (31)$$

$$Q_4 = \sum_{i=1}^m \sum_{j=1}^l \sum_{k=1}^n x_{ijk}^2 - \frac{\sum_{i=1}^m \sum_{j=1}^l \left[\sum_{k=1}^n x_{ijk} \right]^2}{n}, \quad (32)$$

$$Q = \sum_{i=1}^m \sum_{j=1}^l \sum_{k=1}^n x_{ijk}^2 - \frac{\left[\sum_{i=1}^m \sum_{j=1}^l \sum_{k=1}^n x_{ijk} \right]^2}{m \ln}, \quad (33)$$

$$Q_3 = Q - Q_1 - Q_2 - Q_4. \quad (34)$$

Отклонение от основных предпосылок дисперсионного анализа - нормальности распределения исследуемой переменной и равенства дисперсий в ячейках не сказывается существенно на результатах дисперсионного анализа при равном числе наблюдений в ячейках, но чувствительно при неравном их числе. Кроме того, при неравном числе наблюдений в ячейках резко возрастает сложность аппарата дисперсионного анализа. Поэтому рекомендуется планировать схему с равным числом наблюдений в ячейках, а если встречаются недостающие данные, то возмещать их средними значениями других наблюдений в ячейках. При этом, однако, искусственно введенные недостающие данные не следует учитывать при подсчете числа степеней свободы.[14, с.47]

Описанная выше методика проведения многофакторного анализа должна быть основой для построения модели прогнозирования поступления выпускников. Следует отметить, что методика проведения дисперсионного анализа оперирует известными математическими функциями и требует проведения определенной последовательности вычислений. Используемые в

дисперсионном анализе функции математической статистики и проводимые вычисления в полной мере могут быть описаны с использованием компьютерных алгоритмов, поэтому можно перейти к следующему этапу – постановке задачи моделирования поступления выпускников школ в высшие и средние учебные заведения и непосредственной разработке математической модели.

2.1.2 Критерии адекватности модели. Для практического использования моделей корреляционного анализа большое значение имеет их адекватность, т.е. соответствие фактическим статистическим данным. Корреляционный и регрессионный анализ обычно проводится для ограниченной по объёму совокупности. Поэтому показатели регрессии и корреляции – параметры уравнения регрессии, коэффициенты корреляции и детерминации могут быть искажены действием случайных факторов. Чтобы проверить, насколько эти показатели характерны для всей генеральной совокупности, не являются ли они результатом стечения случайных обстоятельств, необходимо проверить адекватность построенных статистических моделей.

Критерий Стьюдента

При численности объектов анализа до 30 единиц возникает необходимость проверки значимости (существенности) каждого коэффициента регрессии. При этом выясняют насколько вычисленные параметры характерны для отображения комплекса условий: не являются ли полученные значения параметров результатами действия случайных причин.

Значимость коэффициентов простой линейной регрессии (применительно к совокупностям, у которых $n < 30$) осуществляют с помощью t-критерия Стьюдента. При этом вычисляют расчетные (фактические) значения t-критерия:

$$t_{a_0} = |a_0| \frac{\sqrt{n-2}}{\sigma_{ocm}} \quad t_{a_1} = |a_1| \frac{\sqrt{n-2}}{\sigma_{ocm}} \sigma_x \quad (35)$$

- для параметра a_0 ;
- для параметра a_1 ;

$$\sigma_{ocm} = \sqrt{\sum (y - \hat{y})^2 / n} \quad (36)$$

- где n - объём выборки;
- среднее квадратическое отклонение результативного признака от выравненных значений $\hat{y}$:

$$\sigma_x = \sqrt{\sum (x - \bar{x})^2 / n} \quad \text{или} \quad \sigma_x = \sqrt{\frac{\sum x^2}{n} - \left(\frac{\sum x}{n} \right)^2} \quad (37)$$

- среднее квадратическое отклонение факторного признака x от общей средней $\bar{x}$.

Вычисленные по вышеприведенным формулам значения сравнивают с критическими t , которые определяют по таблице Стьюдента с учетом принятого уровня значимости β и числом степеней свободы вариации $\nu = n - 2$. В социально-экономических исследованиях уровень значимости β обычно принимают равным 0,05. Параметр признаётся значимым (существенным) при условии, если $t_{расч} > t_{табл}$. В таком случае практически невероятно, что найденные значения параметров обусловлены только случайными совпадениями. **Критерий Фишера**

Критерий Фишера позволяет сравнивать величины выборочных дисперсий двух независимых выборок. Для вычисления $F_{эмп}$ нужно найти отношение дисперсий двух выборок, причем так, чтобы большая по величине дисперсия находилась бы в числителе, а меньшая – в знаменателе. Формула вычисления критерия Фишера такова:

$$F_{\text{эмп}} = \frac{\sigma_x^2}{\sigma_y^2}, \quad (38)$$

где σ_x^2 , σ_y^2 - дисперсии первой и второй выборки соответственно.

Так как, согласно условию критерия, величина числителя должна быть больше или равна величине знаменателя, то значение $F_{\text{эмп}}$ всегда будет больше или равно единице.

Число степеней свободы определяется также просто:

$k_1 = n_1 - 1$ для первой выборки (т.е. для той выборки, величина дисперсии которой больше) и $k_2 = n_2 - 1$ для второй выборки.

Критические значения критерия Фишера находятся по величинам k_1 и k_2 по таблице адекватности модели по критерию Фишера.

Если $t_{\text{эмп}} > t_{\text{крит}}$, то нулевая гипотеза принимается, в противном случае принимается альтернативная.

Критерий Пирсона

Наиболее широко известен коэффициент корреляции Пирсона, характеризующий степень линейной зависимости между переменными. Он определяется, как:

$$r = \frac{\sum_i (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_i (x_i - \bar{x})^2} \sqrt{\sum_i (y_i - \bar{y})^2}} \quad (39)$$

Используя этот коэффициент, следует учитывать, что лучше всего он подходит для оценки взаимосвязи между двумя нормальными переменными. Если распределение переменных отличается от нормального, то он по-прежнему продолжает характеризовать степень взаимосвязи между ними, но к нему уже нельзя применять методы проверки на значимость. Также коэффициент корреляции Пирсона не очень устойчив к выбросам - при их наличии можно ошибочно сделать вывод о наличии корреляции между

переменными. Достоинством критерия Пирсона является его универсальность: с его помощью можно проверять гипотезы о различных законах распределения.

Таким образом, для проверки адекватности данной математической модели были использованы критерии Пирсона и Фишера.[15, с.78-80]

2.2 Основные этапы процесса моделирования

В общетеоретическом плане моделирование означает осуществление каким-либо способом отображения или воспроизведения действительности для изучения имеющихся в ней объективных закономерностей. Обобщенно моделирование можно определить как метод опосредствованного познания, при котором изучаемый объект (оригинал) находится в некотором соответствии с другим объектом (моделью), причем объект-модель способен в том или ином отношении замещать оригинал на некоторых стадиях познавательного процесса. Стадии познания, на которых может происходить такая замена, а также формы соответствия модели и оригинала могут быть различны. В процессе изучения модель служит самостоятельным объектом, позволяющим получить при его исследовании некоторые знания об объекте-оригинале. Но с другой стороны, модель используется для получения такой информации об оригинале, которую затруднительно или невозможно получить путем непосредственного исследования оригинала. Ясно, что для того, чтобы модель могла выполнить эту задачу, она должна быть не только сходной с оригиналом, но и отличаться от него. Отличие от оригинала — обязательный признак модели. Лишь в том случае, когда созданная модель отличается от оригинала как раз в тех отношениях, которые препятствуют непосредственному познанию оригинала, появляется возможность обойти эти препятствия.[20, с.10]

Таким образом, отношение модели к оригиналу противоречиво: она обязательно должна быть сходна с ним и отлична от него. Благодаря этому в процессе моделирования от установления отношений сходства между одними элементами модели и оригинала можно осуществить переход к установлению отношений сходства между другими элементами оригинала и модели. Именно

наличие такого перехода и дает возможность получить новые данные об оригинале, о тех его свойствах, связях, и соотношениях, которые на начальной ступени моделирования были еще не известны.

В настоящее время существуют и широко применяются различные виды моделей. В работе рассматривается и используется только один вид, а именно — математические модели, реализуемые на ЭВМ. В отличие от физического моделирования математическое моделирование позволяет изучать только те параметры оригинала, которые имеют математическое описание и связаны математическими соотношениями в уравнениях, относящихся как к модели, так и к оригиналу.

Математическая модель реального объекта есть некоторый математический объект, поставленный в соответствие данному физическому объекту. Естественно, что должны быть известны соотношения, которые выражают в виде математических зависимостей реальные физические связи.

Процесс математического моделирования включает в себя ряд этапов. Начинается он с постановки задачи, решение которой должно быть получено посредством использования модели. В соответствии с конкретной задачей необходимо выделить некоторые свойства и связи (отношения), исследование которых может привести к достижению цели. Следовательно, на данном этапе рассматриваемый объект изучается лишь под определенным углом зрения, т. е. более или менее односторонне. Второй этап — разработка математической модели. Создание математической модели предполагает предварительное осознание ее общей схемы, после чего можно переходить непосредственно к использованию метода моделирования. При разработке математической модели создается объект, в котором интересующие свойства и отношения оригинала могут быть изучены проще, чем при непосредственном его изучении. Объект как бы освобождается от связей и отношений, затрудняющих его познание. Необходимо подчеркнуть, что математическая модель может появиться только как следствие четкого формального описания рассматриваемого объекта с требуемой степенью приближения к действительности. На третьем этапе

математическая модель, о которой ранее говорилось как о средстве познания, становится объектом исследования, проводится вычислительный эксперимент. При этом все действия производятся над моделью и направлены непосредственно на получение знаний об этом объекте, на установление законов его развития, его свойств и отношений. Все эксперименты производятся только с моделью. Предметом теоретического анализа также является математическая модель. Важным преимуществом исследования модели является наличие возможности повторять многие явления для различных исходных условий и с различным характером их изменения во времени. При этом объекты, являвшиеся ранее лишь объектами наблюдения, благодаря математическому моделированию становятся объектами экспериментирования. На четвертом, заключительном этапе процесса математического моделирования необходимо рассматривать вопрос о переносе значений, полученных на математической модели, на реальный объект изучения. Знания, получаемые при исследовании математической модели, относятся прямо и непосредственно лишь к объекту изучения, т. е. к самой модели. Исследователя, однако, интересует не модель как таковая и не ее свойства, а свойства другого объекта (оригинала), который замещается моделью в процессе исследования. Возможность такого перевода знаний существует благодаря наличию определенного соответствия элементов и отношений модели элементам и отношениям оригинала. Связи соответствующих элементов и отношений одного множества (математической модели) с элементами и отношениями другого множества (оригинала) устанавливаются в процессе моделирования. Возможность установления этих связей в свою очередь позволяет определить правила переноса знаний, полученных с помощью математической модели, на моделируемый объект.

Вместе с тем при использовании метода математического моделирования необходимо иметь в виду качественно-количественную специфику моделируемого объекта и учитывать ее при переносе знаний с математической модели на оригинал. Здесь имеется в виду, прежде всего вопрос о точности

получаемых данных. Среди главных факторов, от которых зависит точность результатов, получаемых в процессе математического моделирования, следует отметить степень адекватности описания объекта.

Методические принципы, принятые при исследовании, опыт и теоретические знания исследователя, а также уровень развития науки вообще определяют, насколько существенные связи и отношения объекта были выделены для их описания и последующего изучения посредством математического моделирования и с какой точностью они были описаны.

Таким образом, успешность применения математического моделирования решающим образом зависит от того, насколько удачно была построена модель. При этом необходимо иметь в виду и характер соответствия модели оригиналу, и степень изученности модели, и удобство оперирования с ней.[24, с.21]

Вычислительный эксперимент — метод изучения устройств или физических процессов с помощью математического моделирования. Он предполагает, что вслед за построением математической модели проводится ее численное исследование, позволяющее изучить поведение исследуемого объекта в различных условиях или в различных модификациях. Численное исследование модели дает возможность определять разнообразные характеристики процессов, оптимизировать конструкции или режимы функционирования проектируемых устройств. Более того, случается, что в ходе вычислительного эксперимента исследователь неожиданно открывает новые процессы и свойства, о которых ему ранее ничего не было известно.

Вычислительный эксперимент занимает промежуточное положение между натурным экспериментом и аналитическим исследованием.

Натурный (физический) эксперимент при надлежащей постановке может, вообще говоря, дать исчерпывающие и надежные результаты. И все же во многих случаях предпочтение отдается вычислительному эксперименту.

Дело в том, что в вычислительном эксперименте в роли опытной установки выступает не конкретное физическое устройство, а программа. Ее построение и

последующие модификации, как правило, требуют существенно меньших затрат, чем подобные манипуляции над реальным объектом.

Кроме того, в опытной установке нередко просто невозможно бывает воссоздать некоторые критические режимы или экстремальные условия. Поэтому математическое моделирование может оказаться практически единственно возможным способом исследования.

При аналитическом подходе так же, как и в вычислительном эксперименте, строится математическая модель. Но исследуется эта модель исключительно посредством аналитических выкладок, без привлечения каких-либо численных методов. Если аналитических выкладок оказывается достаточно, то данный подход приводит к строгому точному решению.

Однако на практике, как это ни парадоксально, аналитическому подходу обычно отводится роль инструмента для (сравнительно быстрого) получения грубых оценок. Объясняется это тем, что аналитическими выкладками удастся ограничиться только для несложных, сильно упрощенных моделей реальных процессов. Получаемое тут строгое аналитическое решение на самом деле в силу исходного огрубления модели оказывается весьма далеким от совершенства. Напротив, численные методы, применяемые в вычислительном эксперименте, дают возможность изучать более сложные модели, достаточно полно и точно отражающие исследуемые процессы.

Отмеченные достоинства вычислительного эксперимента вывели его в число основных методов исследования таких крупных физических и инженерно-технических проблем, как задачи ядерной энергетики, освоения космического пространства и др.

Программные комплексы, обслуживающие вычислительный эксперимент, объемны и сложны, в их создание вовлечен многочисленный отряд программистов. Поэтому особую актуальность приобретает изучение возникающих здесь конфигурационных построений, которые, как будет видно из дальнейшего изложения, постоянно находятся в центре внимания участников такого рода разработок.[25, с.79]

В цикле вычислительного эксперимента можно выделить следующие этапы (см. рисунок 22):

Этап первый. Построение математической модели (составление уравнений, описывающих исследуемое явление).

Этап второй. Выбор численных методов расчета (построение дискретной модели, аппроксимирующей исходную математическую задачу, построение разностной схемы, разработка вычислительного алгоритма и т. д.).

Этап третий. Создание программы, реализующей вычислительный алгоритм.

Этап четвертый. Проведение расчетов и обработка полученной информации.

Этап пятый. Анализ результатов расчетов, сравнение (если это возможно) с натурным экспериментом.



Рисунок 22 – Цикл вычислительного эксперимента

Обычно на последнем (пятом) этапе исследователь приходит к заключению о том, что необходимо внести определенные изменения в решения, принятые на этапах первом, втором или третьем.

Так, может выясниться, что построенная модель недостаточно хорошо отражает особенности исследуемого явления. В этом случае модель корректируется, вносятся соответствующие поправки в численные методы и реализующие их программы, и выполняется новый расчет. Тем самым цикл вычислительного эксперимента воспроизводится в полном объеме.

При анализе результатов могут быть выявлены какие-либо недостатки используемых численных методов, связанные, в частности, с соображениями точности или эффективности. Изменение методов влечет за собой изменение соответствующих программ и т.д. Иначе говоря, цикл повторяется в несколько сокращенном виде (этапы со второго по пятый).

Наконец, может оказаться неудачным некоторое программное решение, например выбранный способ работы с внешней памятью. Пересмотр таких решений приводит к повторению этапов с третьего по пятый.

Вместе с тем при использовании метода математического моделирования необходимо иметь в виду качественно-количественную специфику моделируемого объекта и учитывать ее при переносе знаний с математической модели на оригинал. Здесь имеется в виду, прежде всего вопрос о точности получаемых данных. Среди главных факторов, от которых зависит точность результатов, получаемых в процессе математического моделирования, следует отметить степень адекватности описания объекта.

Методические принципы, принятые при исследовании, опыт и теоретические знания исследователя, а также уровень развития науки вообще определяют, насколько существенные связи и отношения объекта были выделены для их описания и последующего изучения посредством математического моделирования и с какой точностью они были описаны.

Таким образом, успешность применения математического моделирования решающим образом зависит от того, насколько удачно была построена модель. При этом необходимо иметь в виду и характер соответствия модели оригиналу, и степень изученности модели, и удобство оперирования с ней.[26, с.21-26]

2.3 Постановка задачи моделирования поступления выпускников школ в высшие и средние учебные заведения

Целью диссертационной работы является построение математической модели прогнозирования поступления выпускников школ в высшие и средние учебные заведения.

Постановка цели исследования обусловила необходимость решения следующих задач:

- Провести мониторинг динамических рядов данных, характеризующих социально-экономическое развитие региона;
- На основе методологической базы проанализировать методы построения однофакторной и многофакторной модели прогнозирования;
- Разработать математическую модель прогнозирования;
- Разработать рекомендации к использованию математической модели.

Для построения математической модели был выявлен ряд факторов, влияющих на поступление выпускников школ в высшие и средние учебные заведения. К социально-экономическим факторам региона, влияющим на поступление выпускников в высшие и средние учебные заведения, относятся:

1. Уровень жизни населения:

- денежные доходы населения;
- среднемесячная заработная плата;
- минимальная заработная плата;
- величина прожиточного минимума в среднем на душу населения;
- стоимость продовольственной корзины.

2. Образование:

- число дошкольных учреждений, численность детей;
- число государственных общеобразовательных школ, численность учащихся;
- число негосударственных общеобразовательных школ, численность учащихся;

- число государственных средних учебных заведений, численность учащихся;
- число негосударственных средних учебных заведений, численность учащихся;
- прием и выпуск учащихся средних учебных заведений;
- число государственных высших учебных заведений, численность студентов;
- число негосударственных высших учебных заведений, численность студентов;
- прием студентов и выпуск специалистов;
- стоимость обучения в высших и средних учебных заведениях;
- инвестиции в сферу образования.

3. Территория, плотность и численность населения:

- административно-территориальное деление Павлодарской области;
- плотность населения (городское и сельское);
- численность населения (городское и сельское);
- средний возраст мужчин и женщин.

4. Общие показатели воспроизводства населения:

- рождаемость (число родившихся);
- смертность (число умерших);
- младенческая смертность;
- естественный прирост.

5. Браки и разводы:

- количество браков;
- количество разводов.

6. Миграция:

- распределение мигрантов, в разрезе городов и районов;
- миграция населения по уровню образования;
- распределение мигрантов по возрасту

7. Безработица:

- численность безработных, состоящих на учете в органах занятости;
- численность безработных, не состоящих на учете в органах занятости;
- численность экономически активного населения.

Построение математической модели основывается на методе математического прогнозирования, где:

- известные значения параметров – это социально-экономические факторы региона павлодарской области;
- число наблюдений – это значение известного параметра в данный интервал времени;
- прогнозируемое значение – это число выпускников школ, которые будут поступать в высшие и средние учебные заведения.

Описание задачи математического моделирования:

Составим матрицу, состоящую из известных параметров и прогнозируемого значения:

$$\begin{bmatrix} x_1 & a_1 & b_1 & \dots & c_1 \\ x_2 & a_2 & b_2 & \dots & c_2 \\ x_3 & a_3 & b_3 & \dots & c_3 \\ \dots & \dots & \dots & \dots & \dots \\ x_n & a_n & b_n & \dots & c_n \end{bmatrix} \quad (40)$$

где: x_n – прогнозируемое значение;

$a_i, b_i, \dots, c_i; i = \overline{1, n}$ - известные значения параметров;

n - число наблюдений.

Необходимо спрогнозировать значение x_n , в зависимости от $a_i, b_i, \dots, c_i; i = \overline{1, n}$, используя метод корреляционного многофакторного прогнозирования.

Для решения поставленной задачи необходимо:

1. Выявить наличия взаимосвязей между: x и a , x и b , x и $\dots$, x и c .

Для этого необходимо вычислить:

$$r_{xa}, r_{xb}, r_{x\dots}, r_{xc}; i = \overline{1, n-1} \quad (41)$$

Выбираем явно-зависимые параметры, т.е. $r > 0.5$, где:

$$r_{xy} = \frac{E(x * y) - E(x) * E(y)}{\sqrt{(E(x^2) - (E(x))^2) * (E(y^2) - (E(y))^2)}} \quad (42)$$

E – математическое ожидание.

$$E(z) = \frac{1}{m} \sum_{i=1}^m z_i \quad (43)$$

2. Общий коэффициент корреляции вычисляется по следующей формуле:

$$R = r_{xa} * r_{xb} * r_{x...} * r_{xc} \quad (44)$$

3. На промежутке $[\min x_i; 2 * \max x_i]$ с шагом Δx , вычисляем $x_{nj} = x_{nj-1} + \Delta x$

$$\Delta x = \frac{\max x_i - \min x_i}{\vartheta} \quad (45)$$

ϑ - точность вычислений $R' = r'_{xa} * r'_{xb} * r'_{x...} * r'_{xc}$.

4. Выбираем x_{nj} , соответствующий значению $\max R'$.

5. Погрешность прогнозирования в данном случае вычисляется как:

$$\varepsilon = |R' - R| \quad (46)$$

В результате решения поставленной задачи получаем:

- x_{nj} - искомое прогнозируемое значение;
- ε - погрешность.[41, с.23]

2.4 Выводы по второй главе

Таким образом, во второй главе диссертационного исследования рассмотрены основные этапы процесса моделирования, проведен анализ

методов построения однофакторной и многофакторной моделей прогнозирования и обоснован выбор критерия адекватности математической модели. Предлагается математическая модель для прогнозирования поступления выпускников школ в высшие и средние учебные заведения. Данная модель может быть применена в системе образования Павлодарской области для наиболее эффективного планирования учебного процесса.

3 РЕАЛИЗАЦИЯ МОДЕЛИ НА ЭВМ И ЕЕ ИССЛЕДОВАНИЕ

3.1 Проектирование таблиц базы данных

Вся информация, необходимая для построения модели прогнозирования поступления выпускников школ в высшие и средние учебные заведения, хранится в базе данных Access.

Для моделирования прогнозирования поступления выпускников сформирована база данных Baza_kriteriev.mdb, которая состоит из двух таблиц: krits, main.

Таблица Krits предназначена для хранения наименований всех критериев и представлена в таблице 29.

Таблица 29 - Структура таблицы Krits

Наименование поля	Тип поля	Описание
ID_Krit	Числовой	Индексированное поле, используемое для связи с интервалом времени
Krit_Name	Текстовый	Название критерия

Таблица Main предназначена для хранения временного интервала времени, значения соответствующего данному интервалу времени и поля необходимого для связывания ее с главной таблицей (см. таблица 30).

Таблица 30 - Структура таблицы Main

Наименование поля	Тип поля	Описание
ID_krit	Числовой	Индексированное поле, используемое для связи с критериями
Data_krit	Числовой	Временной интервал
Znachenie	Числовой	Значение критерия

Многие СУБД при связывании таблиц автоматически выполняют контроль целостности вводимых в базу данных в соответствии с установленными

связями. В конечном итоге это повышает достоверность хранимой в БД информации.

Кроме того, установление связи между таблицами облегчает доступ к данным. Связывание таблиц при выполнении таких операций как поиск, просмотр, редактирование, выборка и подготовка отчетов обычно обеспечивает возможность обращения к произвольным полям связанных записей. Это уменьшает количество явных обращений к таблицам данных и число манипуляций в каждой из них.

Между таблицами krita и main установлена связь «один ко многим».

Главной при связи таблиц krita и main является таблица - krita, а подчиненной – main. Структура связи представлена на рисунке 23.

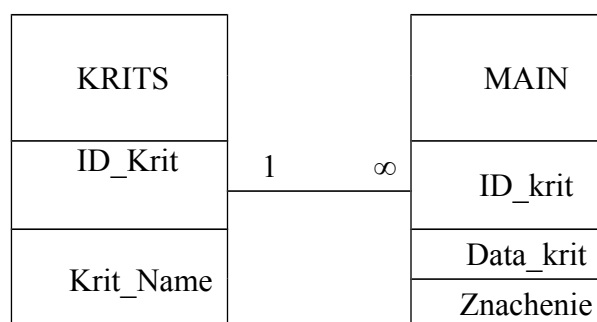


Рисунок 23 – Межтабличные связи

Из вышеперечисленных двух таблиц ключевой является krita, так как именно в ней храниться вся необходимая информация для прогнозирования поступления выпускников школ в высшие и средние учебные заведения.

Структурная схема алгоритма работы модели схематично представлена на рисунке 24.

К входной информации относится перечень критериев, временной интервал, значение соответствующее данному интервалу времени.

К выходной информации относится значение прогнозируемого параметра.

Таким образом, на основе предлагаемой математической модели прогнозирования можно проанализировать степень влияния социально-экономических факторов региона Павлодарской области на поступление

выпускников школ в учебные заведения и на основе полученных результатов проводить планирование учебного процесса на следующий учебный год.

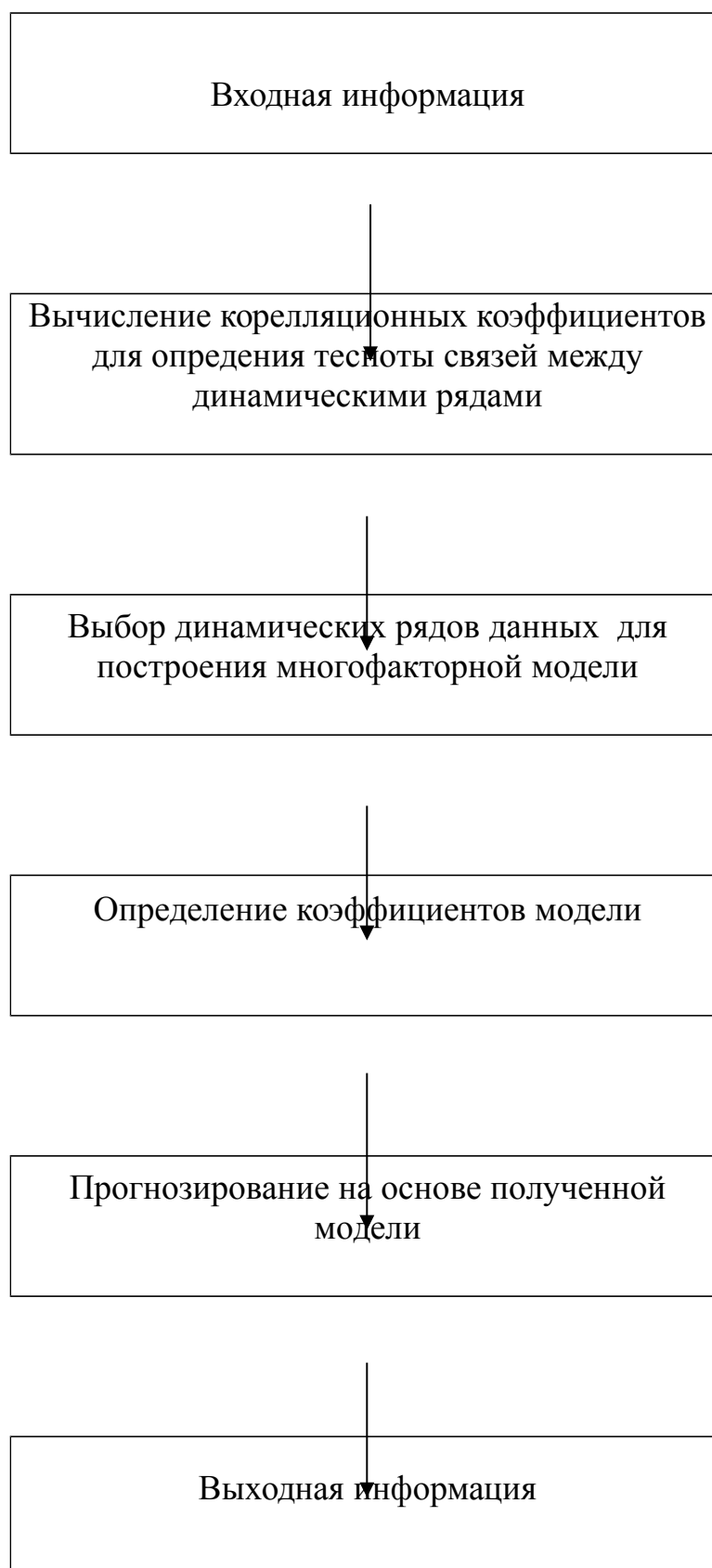


Рисунок 24 – Структурная схема алгоритма работы модели прогнозирования поступления выпускников школ в высшие и средние учебные заведения.

3.2 Выбор инструментария разработки программного продукта

Выбор технических средств реализации методов обучения на ЭВМ остается одной из главных задач разработчика. На сегодняшний день большинство разработок представляют собой полностью законченное программное обеспечение, реализованное в некоторой среде программирования. Это объясняется, с одной стороны, доступностью данных языков программирования и их большими возможностями в области объектного проектирования, а с другой стороны, это достаточно высокая компьютерная грамотность разработчиков именно в среде программирования высокого уровня.

Популярность языков программирования зависит от многих факторов, среди которых исторически сложившиеся традиции, цены на программное обеспечение и применимость языков для конкретных практических задач в данном регионе.

Мировая индустрия средств разработки приложений движется в направлении максимального упрощения процесса создания программ, переводя его на визуальный уровень. Это позволяет разработчику сосредоточиться только на логике решаемой задачи. Поэтому необходимо использовать проверенные средства программного обеспечения, сочетая разумную достаточность с простыми технологиями быстрой разработки. С точки зрения удобства применения для реализации прикладных разработок система объектного программирования Delphi представляется достаточно удобной средой. Большой объем (более 200) визуальных компонентов, необходимых для формирования пользовательского интерфейса, возможность создания многоуровневых приложений, способных работать с разными СУБД на разных компьютерах в локальной и глобальной сети, создание Web – приложений, основанных на TCP/IP протоколах, и другие возможности среды программирования Delphi позволяют создавать проекты максимально просто и качественно.

Модель прогнозирования поступления выпускников в высшие и средние учебные заведения реализована в объектно-ориентированной среде Delphi, так как именно здесь можно без особого труда реализовать следующие задачи:

- степень защиты от несанкционированного изменения данных в системе;
- составления необходимого перечня выходных форм;
- организовать дружественный интерфейс;
- возможность ввода информации и вывода отчетов средствами пакета MsOffice.

База данных представляет собой совокупность специальным образом организованных данных, хранимых в памяти вычислительной системы и отображающих состояние объектов и их взаимосвязей.

Технология Microsoft ActiveX Data Objects (ADO) – это пользовательский интерфейс к любым типам данных, обеспечивающий универсальный доступ к источникам данных из приложений БД. Связь с данными осуществляется посредством так называемой технологии OLE DB.

Технология ADO и интерфейсы OLE DB обеспечивают для приложений единый способ доступа к источникам данных различных типов. Например, приложение, использующее ADO, может применять одинаково сложные операции и к данным, хранящимся на корпоративном сервере SQL, и к электронным таблицам, и локальным СУБД.

Так как технология ADO основана на стандартных интерфейсах COM, которые являются системным механизмом Windows, это сокращает общий объем работающего программного кода и позволяет распространять приложения БД без вспомогательных программ и библиотек.

Microsoft Access имеет возможности, которые позволяют произвести переход на программирование доступа к данным (таблицам, хранимым процедурам, представлениям) через объекты данных ActiveX (ADO) средства доступа OLE DB.

3.3 Разработка интерфейса

Компьютерная модель разработана в интегрированной среде Delphi. Программа состоит из 5 модулей, обеспечивающих работу со всеми таблицами базы данных. Все модули можно условно разделить на следующие группы: модуль данных, модули для работы с фреймами, модуль главной формы программы, модуль под справочную информацию о программе.

Все основные компоненты приложения расположены на фреймах, каждый из которых выполняет определенное действие.

Многофакторная модель прогнозирования основана на общекорреляционном анализе известных данных. Для прогнозирования последующего значения используются значения корреляционно зависимых факторов.

Главное окно приложения содержит главное меню, которое состоит из 3 пунктов, как показано на рисунке 25.

Пример работы с программой

1. Для начала расчетов необходимо выбрать в главном меню Данные ->Критерии.

2. В поле «Критерии», можно указать известные статистические параметры.

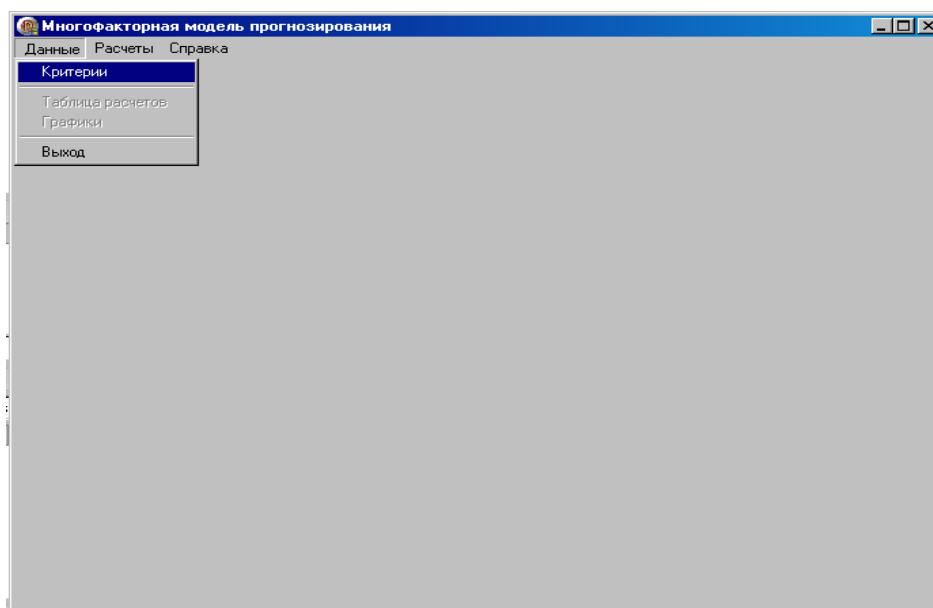


Рисунок 25 - Главное меню программы

3. В поле «Значение» необходимо указать значение данного параметра в наблюдаемый период, как показано на рисунке 26.

4. После Ввода данных, необходимо построить «Таблицу расчета» (см. рисунок 26).

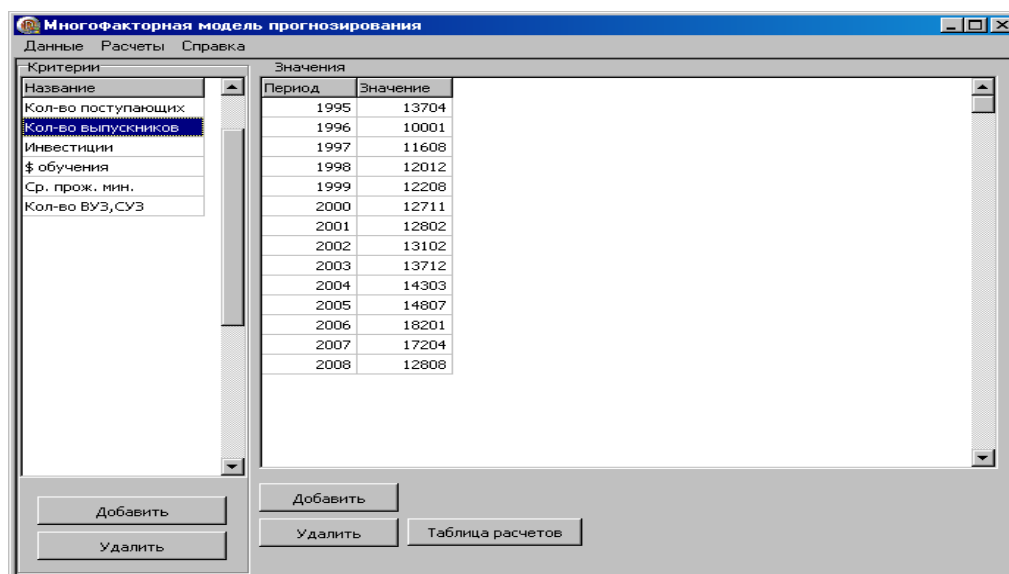
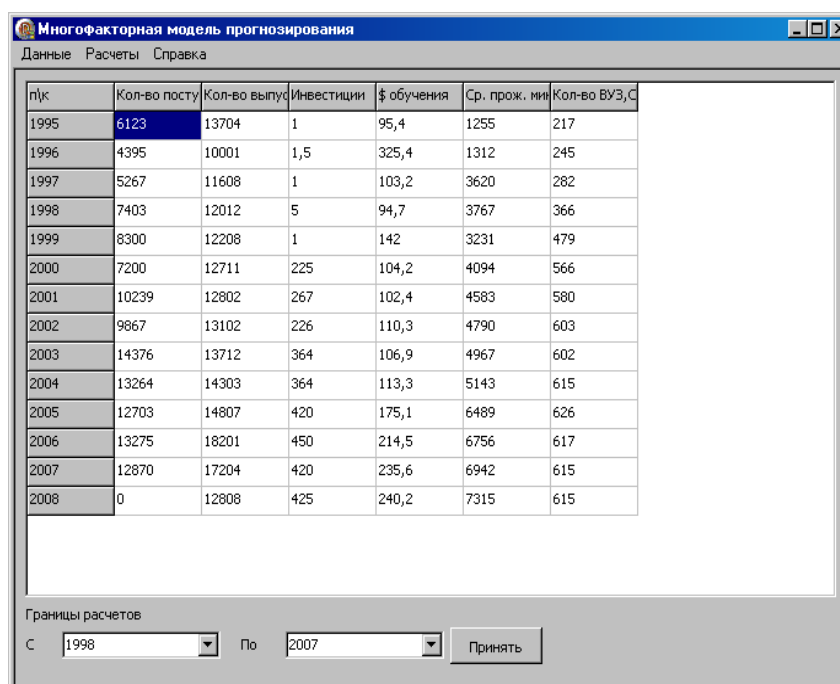


Рисунок 26 - Ввод статистических данных

5. Для определения границы расчетов прогнозирования необходимо ввести значения «С», «По» после чего нажать кнопку «Принять» (см. рисунок 27).



Многофакторная модель прогнозирования

Данные Расчеты Справка

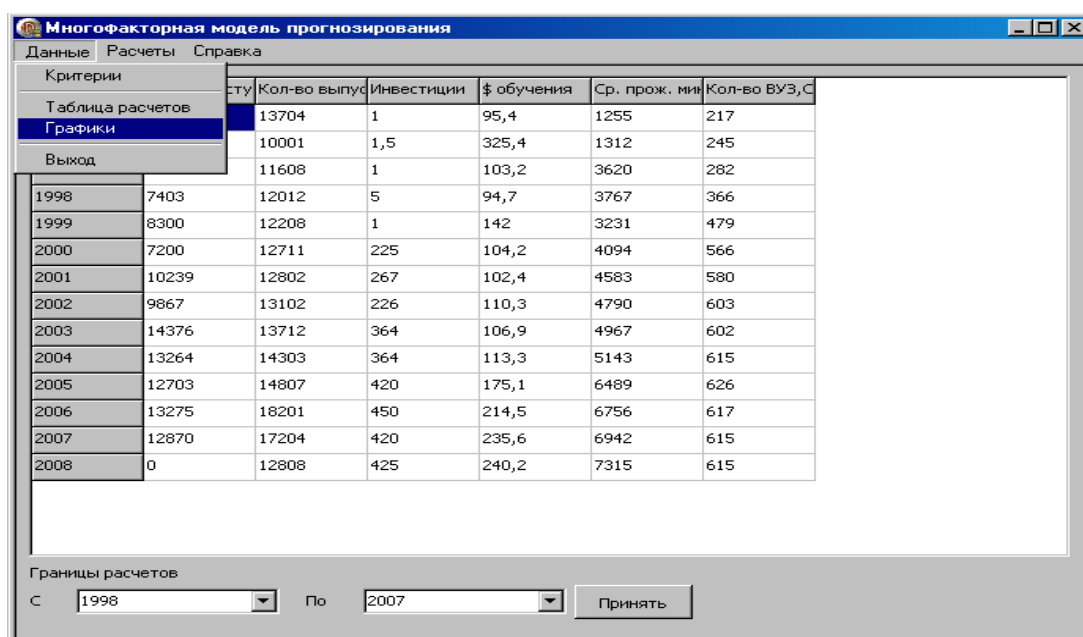
п/к	Кол-во посту	Кол-во выпуск	Инвестиции	\$ обучения	Ср. прож. ми	Кол-во ВУЗ, С
1995	6123	13704	1	95,4	1255	217
1996	4395	10001	1,5	325,4	1312	245
1997	5267	11608	1	103,2	3620	282
1998	7403	12012	5	94,7	3767	366
1999	8300	12208	1	142	3231	479
2000	7200	12711	225	104,2	4094	566
2001	10239	12802	267	102,4	4583	580
2002	9867	13102	226	110,3	4790	603
2003	14376	13712	364	106,9	4967	602
2004	13264	14303	364	113,3	5143	615
2005	12703	14807	420	175,1	6489	626
2006	13275	18201	450	214,5	6756	617
2007	12870	17204	420	235,6	6942	615
2008	0	12808	425	240,2	7315	615

Границы расчетов

с 1998 По 2007 Принять

Рисунок 27 - Расчетная таблица

6. Для более наглядного отображения зависимости факторов, строятся «Графики» распределения величин исследуемых критериев.



Многофакторная модель прогнозирования

Данные Расчеты Справка

Критерии

Таблица расчетов

Графики

Выход

п/к	Кол-во посту	Кол-во выпуск	Инвестиции	\$ обучения	Ср. прож. ми	Кол-во ВУЗ, С
1995	6123	13704	1	95,4	1255	217
1996	4395	10001	1,5	325,4	1312	245
1997	5267	11608	1	103,2	3620	282
1998	7403	12012	5	94,7	3767	366
1999	8300	12208	1	142	3231	479
2000	7200	12711	225	104,2	4094	566
2001	10239	12802	267	102,4	4583	580
2002	9867	13102	226	110,3	4790	603
2003	14376	13712	364	106,9	4967	602
2004	13264	14303	364	113,3	5143	615
2005	12703	14807	420	175,1	6489	626
2006	13275	18201	450	214,5	6756	617
2007	12870	17204	420	235,6	6942	615
2008	0	12808	425	240,2	7315	615

Границы расчетов

с 1998 По 2007 Принять

Рисунок 28 - Пункт главного меню «Графики»

7. Произвести построение графиков можно, после выбора необходимых критериев, нажатием кнопки «Построить», как показано на рисунке 29.

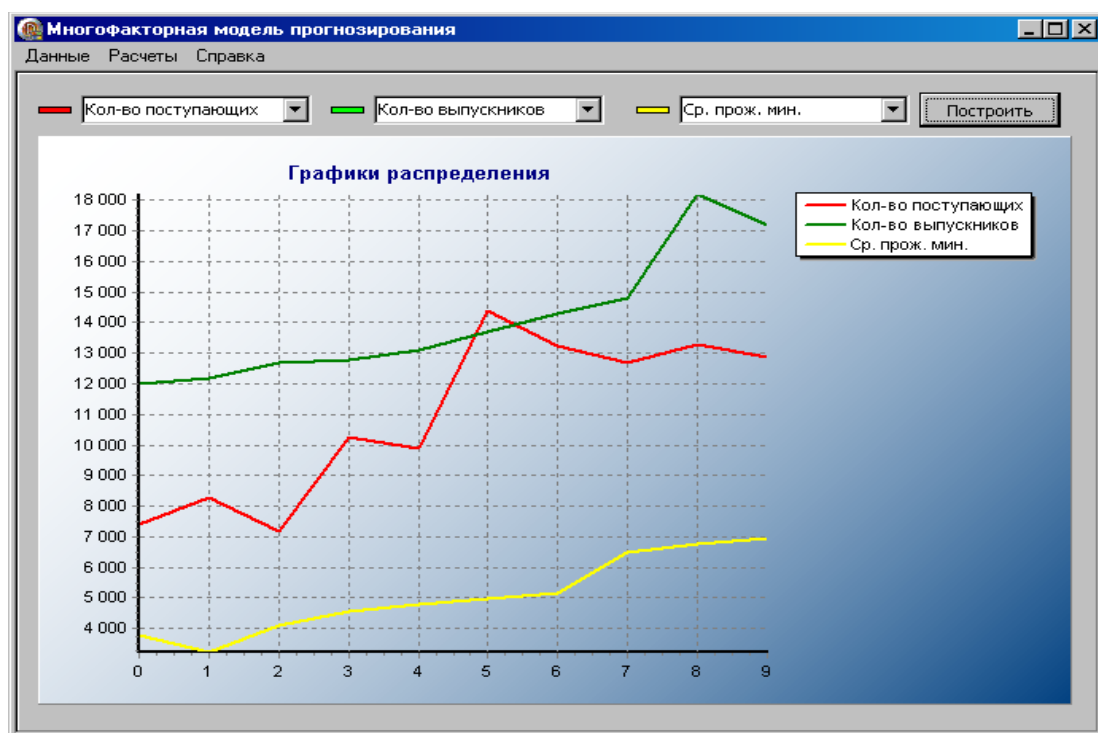


Рисунок 29 - Графики распределения факторов

8. Произвести вычисление Корреляции, между введенными критериями можно в меню Расчеты -> Корреляция (см. рисунок 30).

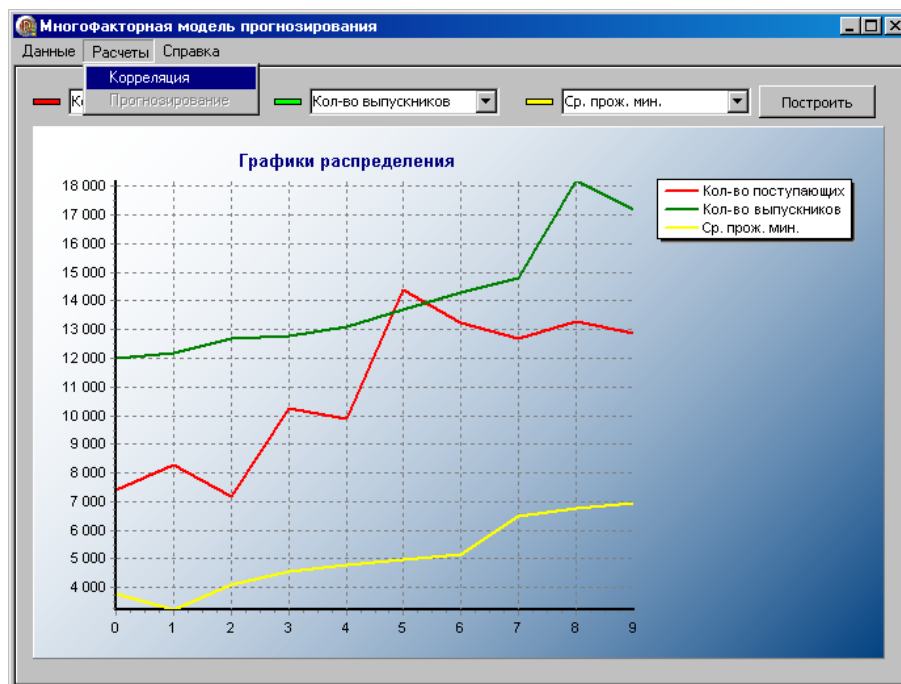
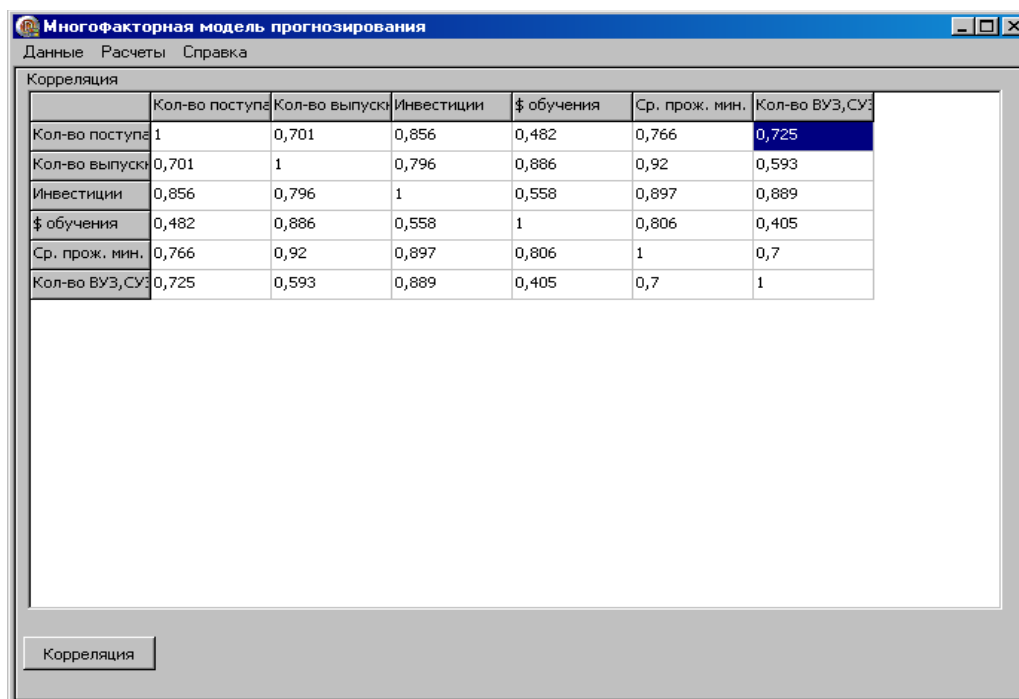


Рисунок 30 - Пункт меню «Корреляция»

9. По нажатию кнопки «Корреляция» произойдет анализ зависимости критериев (см. рисунок 31).



Многофакторная модель прогнозирования

Данные Расчеты Справка

Корреляция

	Кол-во поступающих	Кол-во выпускников	Инвестиции	\$ обучения	Ср. прож. мин.	Кол-во ВУЗ, СУЗ
Кол-во поступающих	1	0,701	0,856	0,482	0,766	0,725
Кол-во выпускников	0,701	1	0,796	0,886	0,92	0,593
Инвестиции	0,856	0,796	1	0,558	0,897	0,889
\$ обучения	0,482	0,886	0,558	1	0,806	0,405
Ср. прож. мин.	0,766	0,92	0,897	0,806	1	0,7
Кол-во ВУЗ, СУЗ	0,725	0,593	0,889	0,405	0,7	1

Корреляция

Рисунок 31 - Прогнозирование зависимой величины

10. Для перехода к этапу прогнозирования необходимо выбрать пункт меню Расчеты -> Прогнозирование (см. рисунок 32).

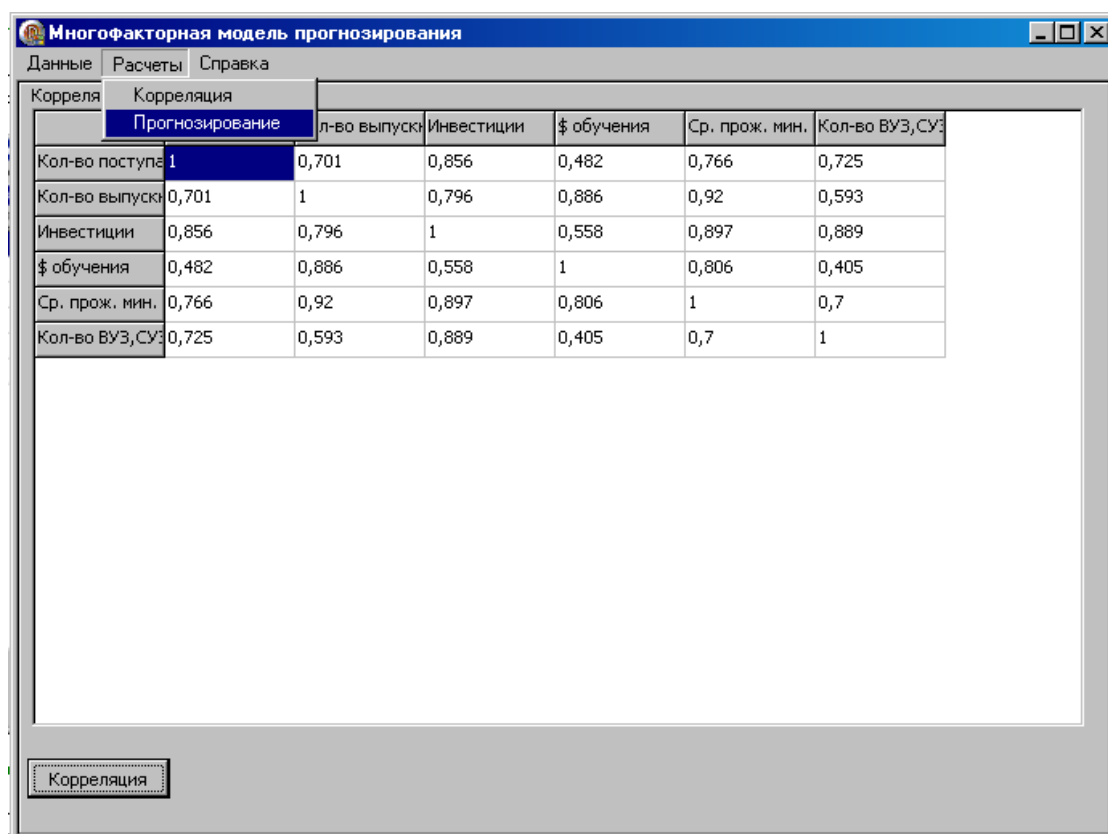


Рисунок 32 - Пункт меню «Прогнозирование»

11. В этом меню можно выбрать прогнозируемый параметр и подтвердить свои действия нажатием на кнопку «Принять».

12. Для прогнозирования необходимо указать Значения известных факторов, и коэффициент «Сглаживания».

13. Запустить расчет прогнозирования можно нажатием кнопки «Прогноз» (см. рисунок 33)

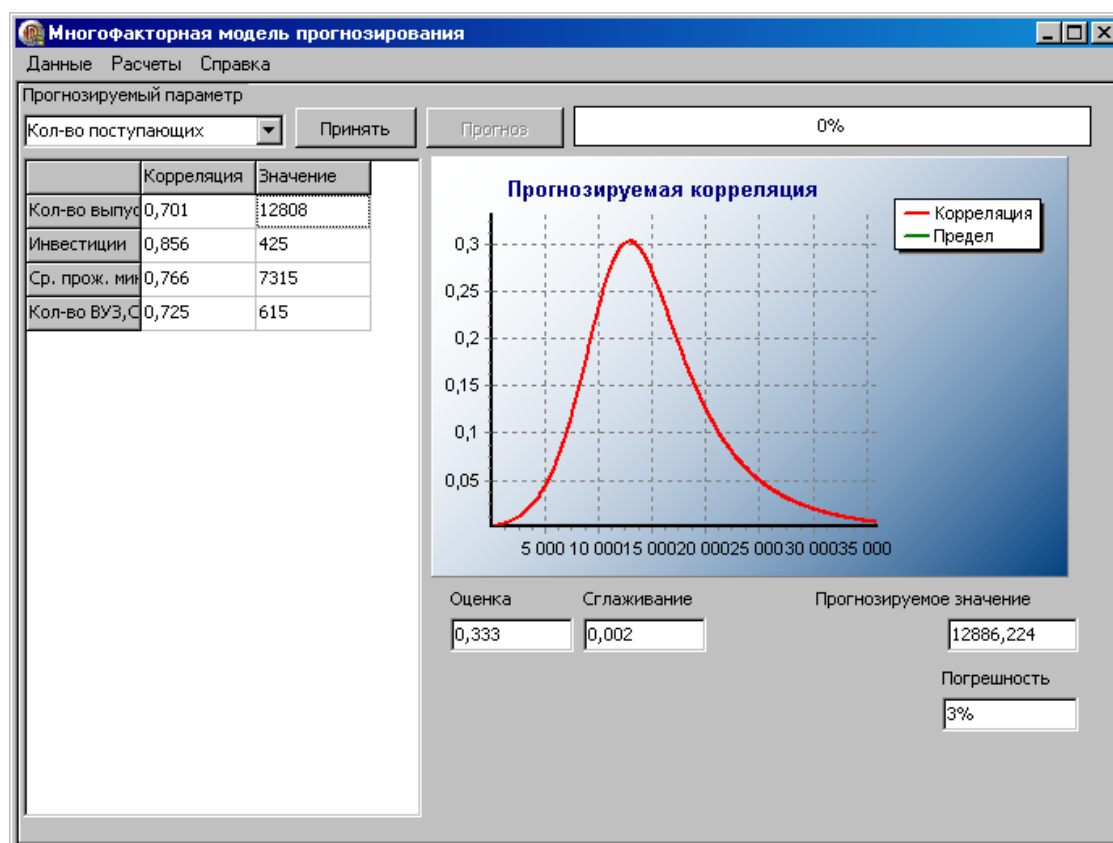


Рисунок 33 - Пункт меню «Прогнозирование»

В случае если значение критерия в какой-либо период неизвестно, программа выбирает нулевое значение.

Выходные данные: прогнозируемое значение, погрешность вычислений, оценка общей зависимости факторов.

Таким образом, с помощью компьютерной модели прогнозирования поступления выпускников школ в высшие и средние учебные заведения можно сделать прогноз на будущее, который необходим для эффективного планирования учебного процесса.

В заключении следует отметить, что полноценное развитие и использование новых информационных технологий и, в частности, программных разработок в системе образования, может привести к оптимизации процесса обучения. Учебный процесс, в свою очередь, будет органично связан с новыми достижениями в области информационно – коммуникационных технологий, что, в конечном счете, повлечет за собой повышение качества профессиональной подготовки будущих специалистов.

3.4 Выводы по третьей главе

При реализации предложенной модели прогнозирования поступления выпускников выполняется постановка задач, устанавливается состав входной и выходной информации, проектируется алгоритм работы, разрабатывается пользовательский интерфейс.

При внедрении разработанной модели прогнозирования поступления выпускников, она приспособливается к организационной структуре предприятия и происходит лишь модернизация методов работы.

Степень риска от внедрения модели минимальна, так как затраты незначительны и организационная структура предприятия не меняется.

С помощью данной модели представиться возможность, опираясь на результаты прошлых лет, прогнозировать значение поступления выпускников и принимать нужные решения в планировании учебного процесса учебными заведениями.

4 ПРАКТИЧЕСКАЯ АПРОБАЦИЯ КОМПЬЮТЕРНОЙ МОДЕЛИ

4.1 Апробация и тестирование компьютерной модели прогнозирования

Одним из важнейших этапов жизненного цикла ИС является этап тестирования и проверки. На этом этапе выявляются все существующие недостатки ИС, не замеченные в процессе разработки.

Качество программного продукта характеризуется набором свойств, определяющих, насколько продукт «хорош» с точки зрения заинтересованных сторон, таких как заказчик продукта, спонсор, конечный пользователь, разработчики и специалисты по тестированию и программному обеспечению, инженеры поддержки, сотрудники отделов маркетинга, обучения и продаж.

Каждый из участников может иметь различное представление о продукте и о том, насколько он хорош или плох, то есть о том, насколько высоко качество продукта. Таким образом, постановка задачи обеспечения качества продукта выливается в задачу определения заинтересованных лиц, их критериев качества и затем нахождения оптимального решения, удовлетворяющего этим критериям.

Тестирование является одним из наиболее устоявшихся способов обеспечения качества разработки программного обеспечения и входит в набор эффективных средств проверки качества программного продукта.

С технической точки зрения тестирование заключается в выполнении приложения на некотором множестве исходных данных и сверке получаемых результатов с заранее известными (эталонными) с целью установить соответствие различных свойств и характеристик приложения заказанным свойствам.

Как одна из основных фаз процесса разработки программного продукта (Дизайн приложения - Разработка кода - Тестирование), тестирование характеризуется достаточно большим вкладом в суммарную трудоемкость разработки продукта.

Широко известна оценка распределения трудоемкости между фазами создания программного продукта: 40% - 20% - 40%, из чего следует, что наибольший эффект в снижении трудоемкости может быть получен прежде всего на фазах Design и Testing. Поэтому основные вложения в автоматизацию или генерацию кода следует осуществлять, прежде всего, на этих фазах.

Хотя в современном индустриальном программировании автоматизация тестирования является широко распространенной практикой, в то же время технология верификации требований и спецификаций пока делает только первые шаги. Задачей ближайшего будущего является движение в сторону такого распределения трудоемкости, чтобы суммарная цена обнаружения большинства дефектов стремилась к минимуму за счет обнаружения преимущественного числа на наиболее ранних фазах разработки программного продукта.

Требования к идеальному критерию тестирования:

- Критерий должен быть достаточным, т.е. показывать, когда некоторое конечное множество тестов достаточно для тестирования данной программы;
- Критерий должен быть полным, т.е. в случае ошибки должен существовать тест из множества тестов, удовлетворяющих критерию, который раскрывает ошибку;
- Критерий должен быть надежным, т.е. любые два множества тестов, удовлетворяющих ему, одновременно должны раскрывать или не раскрывать ошибки программы;
- Критерий должен быть легко проверяемым, например вычисляемым на тестах.

Для нетривиальных классов программ в общем случае не существует полного и надежного критерия, зависящего от программ или спецификаций. Поэтому мы стремимся к идеальному общему критерию через реальные частные.

Классы критериев:

- Структурные критерии используют информацию о структуре

программы (критерии «белого ящика»);

- Функциональные критерии формулируются в описании требований к программному изделию (критерии «черного ящика»);

- Критерии стохастического тестирования формулируются в терминах проверки наличия заданных свойств у тестируемого приложения средствами проверки некоторой статистической гипотезы;

- Мутационные критерии ориентированы на проверку свойств программного изделия на основе подхода Монте-Карло.

Структурные критерии. Структурные критерии используют модель программы в виде «белого ящика», что предполагает знание исходного текста программы или спецификации программы в виде потокового графа управления. Структурная информация понятна и доступна разработчикам подсистем и модулей приложения, поэтому данный класс критериев часто используется на этапах модульного и интеграционного тестирования (Unit testing, Integration testing).

Структурные критерии базируются на основных элементах УГП, операторах, ветвях и путях.

Условие критерия тестирования команд - набор тестов в совокупности должен обеспечить прохождение каждой команды не менее одного раза. Это слабый критерий, он, как правило, используется в больших программных системах, где другие критерии применить невозможно.

Условие критерия тестирования ветвей - набор тестов в совокупности должен обеспечить прохождение каждой ветви не менее одного раза. Это достаточно сильный и при этом экономичный критерий, поскольку множество ветвей в тестируемом приложении конечно и не так уж велико. Данный критерий часто используется в системах автоматизации тестирования.

Условие критерия тестирования путей - набор тестов в совокупности должен обеспечить прохождение каждого пути не менее одного раза. Если программа содержит цикл (в особенности с неявно заданным числом итераций), то число итераций ограничивается константой (часто - 2, или числом классов

выходных путей).

Автоматизация тестирования. Использование различных подходов к тестированию определяется их эффективностью применительно к условиям, определяемым промышленным проектом.

В реальных случаях работа группы тестирования планируется так, чтобы разработка тестов начиналась с момента согласования требований к программному продукту (выпуск Requirement Book, содержащей высокоуровневые требования к продукту) и продолжалась параллельно с разработкой дизайна и кода продукта. В результате, к началу системного тестирования создаются тестовые наборы, содержащие тысячи тестов. Большой набор тестов обеспечивает всестороннюю проверку функциональности продукта и гарантирует качество продукта, но пропуск такого количества тестов на этапе системного тестирования представляет проблему. Ее решение лежит в области автоматизации тестирования, т.е. в автоматизации разработки.

Качество программного продукта и тестирование.

Качество программного продукта можно оценить некоторым набором характеристик, определяющих, насколько продукт «хорош» с точки зрения всех потенциально заинтересованных в нем сторон. Такими сторонами являются: заказчик продукта, спонсор, конечный пользователь, разработчики продукта, специалисты по тестированию программного продукта, инженеры поддержки, отдел обучения, отдел продаж.

Каждый из участников может иметь различное представление о продукте и по-разному судить о том, насколько он хорош или плох, то есть насколько высоко качество продукта. С точки зрения разработчика, продукт может быть настолько хорош, насколько хороши заложенные в нем алгоритмы и технологии.

Пользователю продукта, скорее всего, безразличны детали внутренней реализации, его в первую очередь волнуют вопросы функциональности и надежности. Спонсора интересует цена и совместимость с будущими технологиями. Таким образом, задача обеспечения качества продукта

выливается в задачу определения заинтересованных лиц, согласования их критериев качества и нахождения оптимального решения, удовлетворяющего этим критериям.

В рамках подобной задачи группа тестирования рассматривается не просто как еще одна заинтересованная сторона, но и как сторона, способная оценить удовлетворение выбранных критериев и сделать вывод о качестве продукта с точки зрения других участников. К сожалению, далеко не все критерии могут быть оценены группой тестирования. Поэтому ее внимание в основном сосредоточено на критериях, определяющих качество программного продукта с точки зрения конечного пользователя.

Тестирование как способ обеспечения качества. Тестирование, с технической точки зрения, есть процесс выполнения приложения на некоторых входных данных и проверка получаемых результатов с целью подтвердить их корректность по отношению к результату.

Тестирование не позиционируется в качестве единственного способа обеспечения качества. Оно является частью общей системы обеспечения качества продукта, элементы которой выбираются по критерию наибольшей эффективности применения в конкретном проекте.

В каждом конкретном проекте элементы системы должны быть выбраны так, чтобы обеспечить приемлемое качество, исходя из приоритетов и имеющихся ресурсов. Выбирая элементы для системы обеспечения качества конкретного продукта, можно применить комбинированное тестирование, обзоры кода, аудит.

При подобном выборе некоторые качества, например легкость модификации и исправления дефектов, не будут оценены и, возможно, выполнены. Задачей тестирования в рассматриваемом случае будет обнаружение дефектов и оценка удобства использования продукта, включая полноту функциональности.

Исходя из задач, поставленных перед группой тестирования в конкретном проекте, выбирается соответствующая стратегия тестирования. Так, в данном

примере, ввиду необходимости оценить удобство использования и полноту функциональности, преимущественный подход к разработке тестов следует планировать на основе использования сценариев.

Итак, основная последовательность действий при выборе и оценке критериев качества программного продукта включает:

- Определение всех лиц, так или иначе заинтересованных в исполнении и результатах данного проекта;
- Определение критериев, формирующих представление о качестве для каждого из участников;
- Приоритетность критериев, с учетом важности конкретного участника для компании, выполняющей проект, и важности каждого из критериев для данного участника;
- Определение набора критериев, которые будут отслежены и выполнены в рамках проекта, исходя из приоритетов и возможностей проектной команды. Постановка целей по каждому из критериев;
- Определение способов и механизмов достижения каждого критерия;
- Определение стратегии тестирования исходя из набора критериев, попадающих под ответственность группы тестирования, выбранных приоритетов и целей.

В магистерской диссертации проведено исследование и разработана компьютерная модель прогнозирования поступления выпускников школ в высшие и средние учебные заведения.

Целью данной компьютерной модели является прогноз поступления выпускников школ в высшие и средние учебные заведения при влиянии социально-экономических факторов региона.

В процессе тестирования в разработанную компьютерную модель в качестве входных данных заносились фактические данные с отдела статистики Павлодарской области с 2000 по 2006 гг. для получения статистических данных на 2007 год.

В результате расчета с помощью программы была выявлена небольшая

разница между теоретическим и фактическим значениями прогнозируемого параметра [см. рисунок 34, рисунок 35].

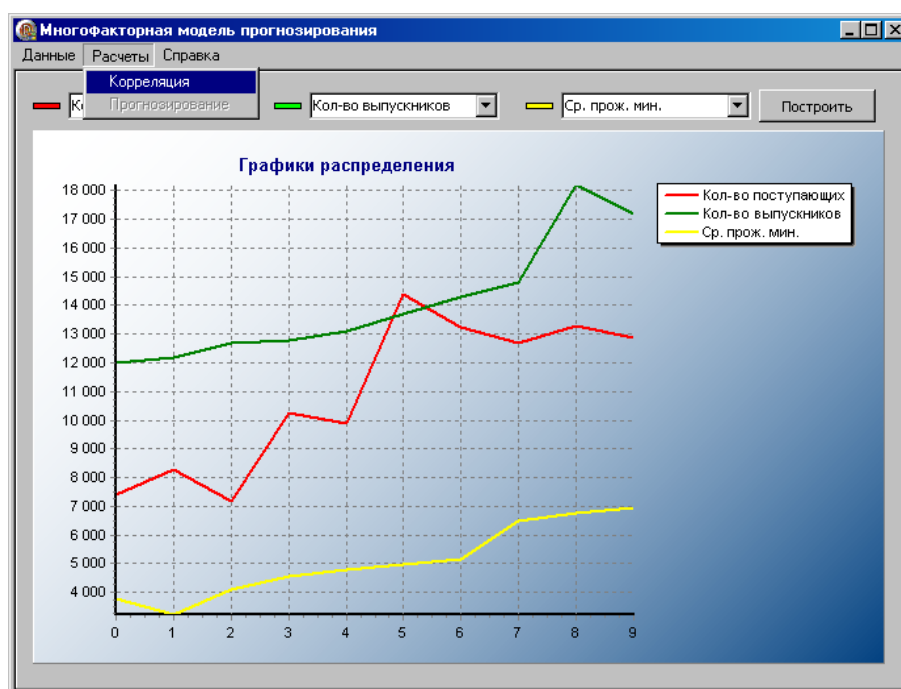


Рисунок 34 - Пункт меню «Корреляция»

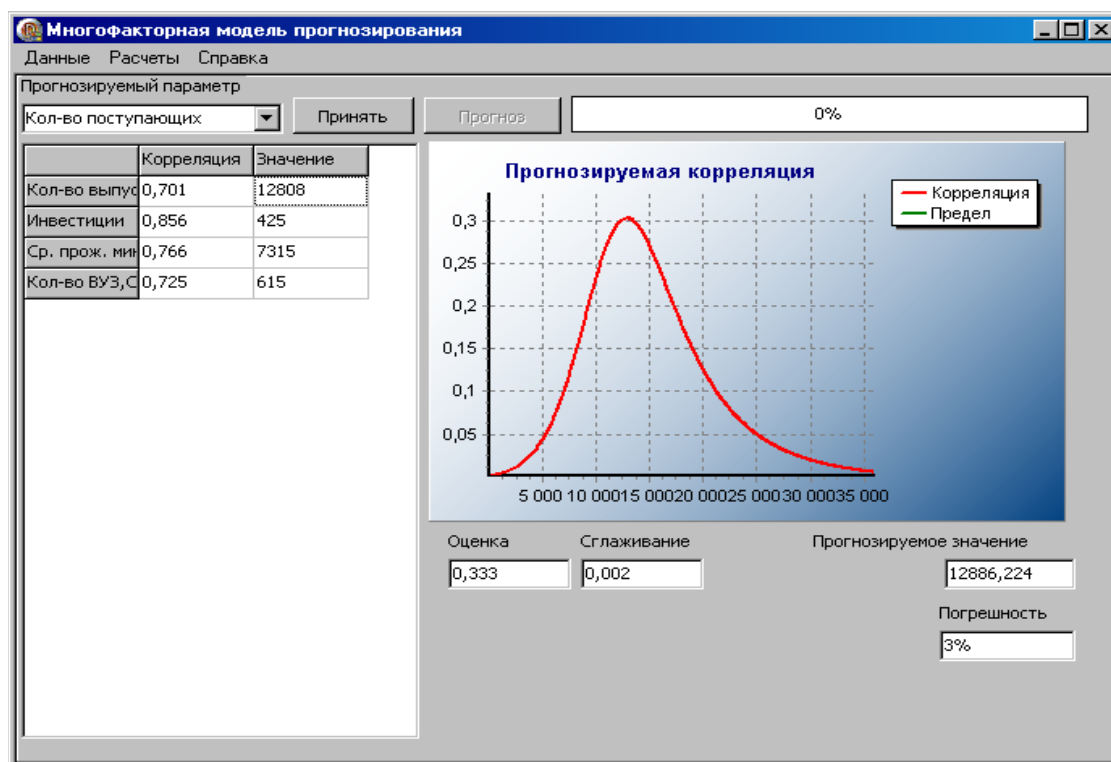


Рисунок 35 - Пункт меню «Прогнозирование»

Таким образом, в процессе внедрения компьютерной модели предполагается решение следующих задач:

- определение социально-экономических факторов, оказывающих влияние на поступление выпускников в учебные заведения;
- формирование динамического (за ряд лет) анализа развития региона в целом;
- определение зависимостей прогнозируемого параметра от известных статистических величин;
- подготовка предложений по планированию учебного процесса в заведениях.

Использование разработанной модели позволит:

- провести анализ зависимости поступления выпускников школ от социально-экономических факторов региона;
- сделать прогноз на поступление выпускников для эффективного планирования учебного процесса.

4.2 Выводы по четвертой главе

Таким образом, при внедрении модели прогнозирования поступления выпускников школ появляется возможность проведения анализа зависимости поступления выпускников школ от социально-экономических факторов региона, прогнозирование поступления выпускников школ, что, в свою очередь, является необходимым фактором для эффективного планирования учебного процесса.

ЗАКЛЮЧЕНИЕ

Эволюция развития человеческого общества давно определила роль образования как наиболее значимого ресурса в формировании перспективного экономического потенциала для отдельных государств и мирового сообщества в целом.

В Послании народу Казахстана в марте 2004 г. Президент страны Н.А. Назарбаев отметил, что: «Конкурентоспособность нации, в первую очередь, определяется уровнем ее образованности. Безусловно, что полная интеграция в мировое образовательное пространство требует поднятия системы казахстанского образования на международный уровень».

Прогнозирование поступления выпускников школ является серьезным фактором для эффективного планирования образовательного процесса в высших и средних учебных заведениях Павлодарской области и, как следствие, поднятия системы образования региона на высокий уровень.

В диссертационном исследовании был определен ряд факторов, влияющих на поступление выпускников в высшие и средние учебные заведения и построена математическая модель.

Для достижения цели исследования были решены следующие задачи:

- проведение мониторинга динамических рядов данных, характеризующих социально-экономическое развитие региона;
- методологической базы проанализированы методы построения однофакторной и многофакторной модели прогнозирования;
- разработана математическую модель прогнозирования.

Исходя из поставленных задач, в работе разработана и предложена компьютерная модель прогнозирования поступления выпускников, позволяющая определить зависимость прогнозируемого параметра от социально-экономических факторов региона. Т.е. модель позволяет не только выявить зависимость поступления выпускников в учебные заведения, но и дает возможность осуществлять анализ социально-экономических факторов прошлых лет, оценить нынешнее положение и производить долгосрочное

планирование.

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

- 1 «Приоритетные направления развития образования и науки на ближайший период» К.Е.Кушербаев, министр Образования и науки РК. Из выступления на расширенном заседании Коллегии Министерства.// «Актуальные проблемы высшей школы» 2000 год.
- 2 Бизнес-образование в Республике Казахстан: «Проблемы и стратегические ориентиры на XXI век». С.Сатубалдин, исполнительный директор КИМЭП при Президенте Республики Казахстан, академик, доктор экономических наук, профессор.// «Актуальные проблемы высшей школы» 2000 год.
- 3 Кушербаев К.Е., Ахметов А.К., Абылкасымова А.Е. Стратегия развития высшего образования в Республике Казахстан. – Алматы: Изд.-во «Білім», 1998. – . 232с.
- 4 Приказ Министерства образования и науки РК «Об утверждении Правил о порядке планирования расходов на подготовку кадров с высшим профессиональным образованием и расходов на обучение одного студента». №681 от 18 сентября 2002. – Астана.
- 5 Приказ Министерства образования и науки РК «Об утверждении финансовых нормативов по высшему образованию». №467 от 13 июня 2002. – Астана.
- 6 Законодательство об образовании в Республике Казахстан. – Алматы: ЮРИСТ, 2005. – 212с.
- 7 Аксельрод В.И. Моделирование и реализация системы дополнительного краеведческого образования учащихся// Дополнительное образование.-1999.-№2.-С.9-12.
- 8 Аксенова Л.И. Модель комплексного образовательного учреждения//Дефектология.-2002.-№4.-С.15-23.
- 9 Бабошина Е.Б. Культурологическая модель гуманистической образовательной среды//Школа.-2002.-№6.-С.22-26.

- 10 Кундышева Е.С. Математическое моделирование в экономике: Учебное пособие для вузов / Под ред. Б.А. Сулакова. – 2-е изд., перераб. и доп. – М.: Дашков и К, 2005. – 352 с.
- 11 Бондаревская Е.В. Ценностно-смысловые ориентиры и стратегические направления развития сельской школы [Опытно-экспериментальная работа по моделированию образоват. пространства в сел. школе] //Педагогика.-2002.-№5.-С.52-64.
- 12 Башмаков М.И. Индивидуальная образовательная программа средней школы(Концепции и модели) //Школьные технологии.-2000.-№4.-С.17-23.
- 13 Математические методы в экономике: Учебник для вузов / О.О. Замков, А.В. Толтопятенко: Под ред. А.В. Сидоровича. – 2-е изд. – М.: Дело и сервис, 1999. – 368 с.
- 14 Экономико-математические методы и прикладные модели: Учебное пособие для вузов / В.В. Федосеев, А.Н. Гармаш и др.; под ред. В.В. Федосеева. – М: - ЮНИТИ, 1999. – 391 с.
- 15 Хачатрян С.Р. и др. Методы и модели решения экономических задач. Учебное пособие / С.Р. Хачатрян, М.В. Пинегина, В.П. Буянов. – М.: Экзамен, 2005. – 384 с.
- 16 Яндхиала Б. Г. Тилак Глобальные тенденции и финансирование высшего образования (перевод с англ. О. Попова). Алма Матер №2, 2005, С. 5 - 7.
- 17 Горячев А.В., Шафрин Ю.А. Практикум по информационным технологиям. – М.: БИНОМ, 2003. – 272с.
- 18 Тарасюк М.В. Защищенные информационные технологии. Проектирование и применение. – М.: СОЛОН - Пресс, 2004. – 192с.
- 19 Видт И.Е. Культурологическая интерпретация эволюции образовательных моделей//Педагогика.-2003.-№3.-С.32-38.
- 20 Гафурова Н.В. Интеллектуально-личностная модель образования как условие развития одаренности учащихся на этапе «школа-

- вуз»//Инновации в образовании.-2004.-№4.-С.119-129.
- 21 Гришин В.Н., Панфилова Е.Е. Информационные технологии в профессиональной деятельности: Учебник.- М.: ФОРУМ ИНФРА-М, 2005.-416 с.
 - 22 Наука и образование в стратегии регионального развития: Материалы Республиканской научно-практической конференции, посвященной 100-летию К.И.Сатпаева (25.10.1999 г.). – Павлодар: ПГУ. Ч.II. – 1999. – 358 с.
 - 23 Головизнина Н.Л. Областная очно-заочная историко-краеведческая школа как инновационная модель системы дополнительного образования//Внешкольник.-2003.-№6.-С.14-15.
 - 24 Гончарова Е.Ю. Модель школы индивидуального выбора//Завуч.-2000.-№4.-С.21-48.
 - 25 Гребнева С.В. Моделирование воспитательной системы в школе-гимназии с углубленным изучением//Классный руководитель.-2000.-№2.-С.79-82.
 - 26 Дахин А.Н. Педагогическое моделирование: сущность, эффективность и неопределенность//Педагогика.-2003.-№4.-С.21-26.
 - 27 Дахин А.Н. Педагогическое моделирование как средство модернизации образования в открытом информационном сообществе // Стандарты и мониторинг в образовании.-2004.-№4.-С.46-60.
 - 28 Дахин А.Н. Содержание образования как культуросообразная модель жизнедеятельности // Стандарты и мониторинг в образовании.-2003.-№6.-С.51-58.
 - 29 Бор М. 3. Основы экономических исследований. Логика, методология, организация, методика. - М.: Издательство «ДИС», 1998.-144с.
 - 30 Вовненко О. В. Как выбрать профессию // Вечерний Мурманск. - 2001.- 9 авг. С. 4

- 31 Гинзбург М.Р. Психологическое содержание личностного самоопределения // Вопр. Психол. 1994. №3. - 255с.
- 32 Голубков Е.П. Маркетинговые исследования: теория, методология и практика. -2-е изд., доп. и перераб. - М.: Издательство «Дело и Сервис», 2001. - 112 с.
- 33 Клименко А.И. Практика тестирования. - М.: Профиздат, 1991. - 80 с.
- 34 Климов Е.А. Как выбрать профессию: Кн. Для учащихся ст. классов сред. Шк. - 2-е изд., доп. и дораб. -М.: Просвещение, 1990. - 159с.
- 35 Кожухарь Л.И. Основы общей теории статистики. — М.: Финансы и статистика, 1999. -144с
- 36 Королёва Т.В. Подготовка молодёжи к выходу на рынок труда // Служба кадров.- 2000.-С.34 - 36.
- 37 Фофенова Н. А. Молодёжный клуб - путь к трудоустройству // Полярная правда 2000.-18 февр. Стр. 3
- 38 Аверкин А.Н., Батыршин И.З., Блишун А.Ф. и др. Нечеткие множества в моделях управления и искусственного интеллекта // Под ред. Д.А. Поспелова. – М.: Наука, 1986. – 312 с.
- 39 Аветисян Д.О. Проблемы информационного поиска: (Эффективность, автоматическое кодирование, поисковые стратегии) - М.: Финансы и статистика, 1981. - 207 с.
- 40 Айвазян С.А., Бежаева З.И., Староверов О.В. Классификация многомерных наблюдений. – М.: Статистика, 1974. – 240 с.
- 41 Айвазян С.А., Енюков И.С., Мешалкин Л.Д. Прикладная статистика. Основы моделирования и первичная обработка данных. Справочное издание. – М.: Финансы и статистика, 1983. – 472 с.
- 42 Айвазян С.А., Енюков И.С., Мешалкин Л.Д. Прикладная статистика: Исследование зависимостей: Справочник. – М.: Финансы и статистика, 1985. – 182с.

- 43 Айвазян С.А. , Мхитарян В.С. Прикладная статистика и основы эконометрики. – М. Юнити, 1998. – 1024 с.
- 44 З. ВФерстер Э., Ренц Б. Методы корреляционного и регрессионного анализа. М.: Финансы и статистика, 1983. – 879 с.
- 45 Архангельский А.Я. Программирование в Delphi 5. – М.:ЗАО “Издательство БИНОМ”, 2000. – 1070с.
- 46 Фаронов В.В. Программирование баз данных в Delphi 6. Учебный курс. – СПб.:Питер, 2002.— 352с.
- 47 Гофман В.Э., Хомоненко А.Д. Delphi. Быстрый старт. – СПб.: БХВ Санкт-Петербург, 2002.- 208с.

ПРИЛОЖЕНИЕ А

(обязательное)

Анкета для социального опроса

1. Анкета (на каждого студента).
2. Инструкция к заполнению анкеты.

Инструкция к заполнению анкеты.

Уважаемые студенты! В данной анкете Вам предлагается перечень социальных факторов, которые влияют на выбор учебного заведения выпускниками школ. Отметьте, пожалуйста, по 5-ти балльной шкале каждый из факторов, который повлиял на Ваш выбор учебного заведения. В конце анкеты Вы можете вписать свой вариант ответа. Если Вы затрудняетесь ответить на какой-нибудь пункт данной анкеты, обращайтесь, пожалуйста, к консультанту, давшему Вам анкету. Так как, Вы уже являетесь студентом и Вам предложили заполнить эту анкету, то ответьте на вопросы так, как Вы думали до поступления в ВУЗ, либо можете передать анкету знакомому абитуриенту.

СПАСИБО.

Вид анкеты

1. Ф. И. О. студента: _____

2. Образование: _____

3. Какие социальные факторы, на Ваш взгляд наиболее сильно влияют на выбор выпускниками учебного заведения (отметьте галочкой, по 5-ти балльной системе):

ПРИЛОЖЕНИЕ А

(продолжение)

Таблица А.1 – Социальные факторы

Социальные факторы	1 - балл	2 - балла	3 - балла	4 - балла	5 - баллов
Жилищные условия					
Здоровье (внутренние возможности и ограничения)					
Информированность					
Квалификация профессорско-преподавательского состава заведения					
Позиция друзей					
Позиция учителей и школьных педагогов					
Позиция членов семьи					
Престиж					
Профессиональный личный план					
Самооценка					
Склонности (интересы, мотивы, труды)					
Соотношение количества студентов к количеству профессорско-преподавательского состава в заведении					
Способности					
Среднее количество студентов в группах					
Стоимость обучения					
Уровень притязаний на общественное признание					
Учебные заведения, представленные на территории проживания					
Ваш фактор: _____					

ПРИЛОЖЕНИЕ В

(обязательное)

ЛИСТИНГ ПРОГРАММЫ

```

unit Unit1;

interface

uses
  Windows, Messages, SysUtils, Variants, Classes, Graphics, Controls, Forms,
  Dialogs, ExtCtrls, ActnList, DB, DBClient, StdCtrls, Grids, DBGrids, ComCtrls,
  AppEvnts, Buttons, OleCtrls, SHDocVw, TeEngine, TeeFunci, Series, TeeProcs,
  Chart, Menus, dblookup, StdActns, DBCtrls, Gauges;
Type TAR = array of real;
type
  TForm1 = class(TForm)
    PageControl1: TPageControl;
    TabSheet1: TTabSheet;
    TabSheet2: TTabSheet;
    TabSheet3: TTabSheet;
    ApplicationEvents1: TApplicationEvents;
    TabSheet6: TTabSheet;
    MainMenu1: TMainMenu;
    N1: TMenuItem;
    N3: TMenuItem;
    N4: TMenuItem;
    N5: TMenuItem;
    N6: TMenuItem;
    N9: TMenuItem;
    N10: TMenuItem;
    N12: TMenuItem;
    N14: TMenuItem;
    Panel3: TPanel;
    GroupBox2: TGroupBox;
    DBGrid4: TDBGrid;
    Label1: TLabel;
    SpeedButton1: TSpeedButton;
    SpeedButton2: TSpeedButton;
    SpeedButton3: TSpeedButton;
    SpeedButton4: TSpeedButton;
    DBGrid5: TDBGrid;
    BitBtn1: TBitBtn;
    Panel1: TPanel;
    TRAS: TStringGrid;
    N2: TMenuItem;
    N11: TMenuItem;
    Label9: TLabel;
    ComboBox1: TComboBox;
    ComboBox2: TComboBox;
    Label10: TLabel;
    Label11: TLabel;
    BitBtn2: TBitBtn;
    Panel4: TPanel;
    Chart1: TChart;
    Series1: TLineSeries;
    TeeFunction1: TAddTeeFunction;
    Series2: TLineSeries;
    TeeFunction2: TAddTeeFunction;
    Series3: TLineSeries;
    ComboBox3: TComboBox;
    ComboBox4: TComboBox;
    ComboBox5: TComboBox;
    Shape1: TShape;
    Shape2: TShape;
    Shape3: TShape;
    BitBtn3: TBitBtn;
    CDSRAS: TClientDataSet;
    CDSRASKritName: TStringField;
    CDSRASZnach: TFloatField;
    TabSheet5: TTabSheet;
    Panel5: TPanel;
  end;

```

```

TCor: TStringGrid;
Label6: TLabel;
BitBtn4: TBitBtn;
DBGrid1: TDBGrid;
DataSource4: TDataSource;
TabSheet7: TTabSheet;
TabSheet4: TTabSheet;
Memo1: TMemo;
Memo2: TMemo;
N7: TMenuItem;
N8: TMenuItem;
Panel6: TPanel;
BitBtn5: TBitBtn;
BitBtn6: TBitBtn;
Chart2: TChart;
LineSeries1: TLineSeries;
LineSeries2: TLineSeries;
ComboBox6: TComboBox;
Edit1: TEdit;
Edit2: TEdit;
Edit3: TEdit;
Edit4: TEdit;
Label2: TLabel;
Label3: TLabel;
Label4: TLabel;
Label5: TLabel;
Label7: TLabel;
Tpro: TStringGrid;
Panel7: TPanel;
WebBrowser1: TWebBrowser;
Gauge1: TGauge;
procedure ApplicationEvents1Exception(Sender: TObject; E: Exception);
procedure Button5Click(Sender: TObject);
procedure SpeedButton2Click(Sender: TObject);
function MIR(X: array of real):double;
function MAR(X: array of real):double;
function MO(X:array of real):double;
Function MO2(X,Y:array of real):double;
Function Correl(X,Y:array of real):Double;
function Sred(X: array of real):double;
Function S1(X: array of real):double;
Function S2(X,Y:array of real):Double;
Function Correl2(X,Y:array of real):Double;
Function Mark1(ms:array of TAR):real;
procedure Edit1KeyPress(Sender: TObject; var Key: Char);
procedure Button6Click(Sender: TObject);
procedure FormCreate(Sender: TObject);
Function FtoA(DSet:Tdataset;FN:String):TAR;
procedure Button4Click(Sender: TObject);
procedure N12Click(Sender: TObject);
procedure SpeedButton1Click(Sender: TObject);
procedure SpeedButton3Click(Sender: TObject);
procedure SpeedButton4Click(Sender: TObject);
procedure DBGrid5CellClick(Column: TColumn);
procedure BitBtn1Click(Sender: TObject);
procedure DBGrid5MouseWheelDown(Sender: TObject; Shift: TShiftState;
  MousePos: TPoint; var Handled: Boolean);
procedure ComboBox1Change(Sender: TObject);
procedure BitBtn2Click(Sender: TObject);
procedure BitBtn3Click(Sender: TObject);
procedure BitBtn4Click(Sender: TObject);
procedure BitBtn5Click(Sender: TObject);
procedure BitBtn6Click(Sender: TObject);
procedure ComboBox6KeyPress(Sender: TObject; var Key: Char);
procedure N9Click(Sender: TObject);
procedure N5Click(Sender: TObject);
procedure N6Click(Sender: TObject);
procedure N11Click(Sender: TObject);
procedure N10Click(Sender: TObject);
procedure N7Click(Sender: TObject);
procedure N8Click(Sender: TObject);

private
  ID_Krit:integer;
  { Private declarations }
public
  Wtar:Tar;
  Xp:array of integer;
  Rs:array of real;

```

```

Puper: array of TAR;
M1,M2:real;
  { Public declarations }
end;

var
  Form1: TForm1;
implementation

uses Unit2, Unit4, Unit3, Unit5;

{$R *.dfm}
function TForm1.MAR(X: array of real): double;
var i:integer;
begin
  Result:=X[0];
  for I := 0 to Length(X) - 1 do
    begin
      if Result>x[i] then Result:=X[i];
    end;
  end;
end;

function TForm1.Mark1(ms: array of TAR): real;
var I:integer;
begin
  Result:=1;
  for I:= 1 to length(ms)-1 do
    begin
      Result:=Result*Correl(ms[0],ms[i]);
    end;
  end;
end;

function TForm1.MIR(X: array of real):Double;
var i:integer;
begin
  Result:=X[0];
  for I := 0 to Length(X) - 1 do
    begin
      if Result<X[i] then Result:=X[i];
    end;
  end;
end;

procedure TForm1.ApplicationEvents1Exception(Sender: TObject; E: Exception);
begin
  Showmessage('Произошла непредвиденная ошибка.');
```

```

end;

procedure TForm1.BitBtn1Click(Sender: TObject);
var i,j,k:integer; ravno:boolean; val:real;
begin
  ComboBox3.Items.Clear;
  ComboBox4.Items.Clear;
  ComboBox5.Items.Clear;
  ComboBox6.Items.Clear;

  DataModule4.Main.Filtered:=false;
  DataModule4.Main.First;
  setlength(Xp,1);
  Xp[0]:=DataModule4.MainData_krit.Value;
  for I := 0 to DataModule4.Main.RecordCount - 1 do
    begin
      ravno:=false;
      for j := 0 to Length(Xp) do
        if Xp[j]=DataModule4.MainData_krit.Value then ravno:=true;
        if ravno=false then
          begin
            SetLength(Xp,length(Xp)+1);
            Xp[length(Xp)-1]:=DataModule4.MainData_krit.Value;
          end;
      DataModule4.Main.Next;
    end;
  TRAS.RowCount:=length(Xp)+1;
  TRAS.ColCount:=DataModule4.Krits.RecordCount+1;
  TRAS.Cells[0,0]:='n\к';
  DataModule4.Krits.First;
  for j := 0 to DataModule4.Krits.RecordCount - 1 do
    begin
      For I := 0 to Length(xp) - 1 do
        begin

```

```

Tras.Cells[0,i+1]:=inttostr(xp[i]);
DataModule4.Main.filtered:=false;
DataModule4.Main.filter:='(ID_krit='+DataModule4.KritsID_Krit.AsString+') AND (Data_krit='+inttostr(xp[i])+')' ;
//ShowMessage(DataModule4.Main.filter);
DataModule4.Main.filtered:=True;
if DataModule4.Main.RecordCount<1 then val:=0 else val:=(round(DataModule4.MainZnachenie.Value*1000)/1000);
Tras.Cells[j+1,i+1]:=floattostr(val);
end;
Tras.Cells[j+1,0]:=DataModule4.KritsKrit_Name.AsString;
ComboBox3.Items.Add(DataModule4.KritsKrit_Name.AsString);
ComboBox4.Items.Add(DataModule4.KritsKrit_Name.AsString);
ComboBox5.Items.Add(DataModule4.KritsKrit_Name.AsString);
ComboBox6.Items.Add(DataModule4.KritsKrit_Name.AsString);
DataModule4.Krits.Next;
end;
DataModule4.Krits.First;
DataModule4.Main.First;
For I := 0 to Length(xp) - 2 do
begin
    ComboBox1.Items.Add(inttostr(xp[i]));
    ComboBox1.ItemIndex:=0;
    ComboBox2.Items.Add(inttostr(xp[i]));
    ComboBox2.ItemIndex:=ComboBox2.Items.Count-1;
end;
N6.Enabled:=true;
N11.Enabled:=true;
ComboBox3.ItemIndex:=0;
ComboBox4.ItemIndex:=0;
ComboBox5.ItemIndex:=0;
ComboBox6.ItemIndex:=0;
panel1.visible:=true;
panel3.visible:=false;
panel4.visible:=false;
panel5.visible:=false;
panel6.visible:=false;
panel7.visible:=false;
end;

procedure TForm1.BitBtn2Click(Sender: TObject);
var i,j:integer;
begin
    CDSRAS.Filtered:=false;
    for I := 0 to CDSRAS.RecordCount - 1 do
        CDSRAS.Delete;
    SetLength(Xp,0);
    i:=ComboBox1.ItemIndex-1;
    while ComboBox1.Items.Strings[i]<>ComboBox2.Items.Strings[ComboBox2.ItemIndex] do
        begin
            i:=i+1;
            SetLength(Xp,Length(Xp)+1);
            Xp[Length(Xp)-1]:=strtoint(ComboBox1.Items.Strings[i]);
            //memo2.Lines.Add(ComboBox1.Items.Strings[i]);
        end;
    for I := 1 to TRAS.ColCount - 1 do
    for J := 1 to TRAS.RowCount - 1 do
    if (StrToInt(TRAS.Cells[0,j])>=StrToInt(ComboBox1.Text)) and (StrToInt(TRAS.Cells[0,j])<=StrToInt(ComboBox2.Text)) then
        begin
            CDSRAS.Append;
            CDSRASKritName.Value:=TRAS.Cells[i,0];
            CDSRASZnach.Value:=strtofloat(TRAS.Cells[i,j]);
            CDSRAS.Post;
        end;
    N6.Enabled:=true;
    N9.Enabled:=true;
end;

procedure TForm1.BitBtn3Click(Sender: TObject);
var i:integer;
begin
    Series1.Clear;
    Series2.Clear;
    Series3.Clear;
    Series1.Title:=ComboBox3.Text;
    Series2.Title:=ComboBox4.Text;
    Series3.Title:=ComboBox5.Text;
    CDSRAS.Filtered:=false;
    CDSRAS.Filter:='KritName='+QuotedStr(ComboBox3.Text);
    CDSRAS.Filtered:=true;
    CDSRAS.First;

```

```

for I := 0 to CDSRAS.RecordCount - 1 do
begin
Series1.Add(CDSRAS.Znach.Value, ",Shape1.Brush.Color);
CDSRAS.Next;
end;

CDSRAS.Filtered:=false;
CDSRAS.Filter:='KritName='+QuotedStr(ComboBox4.Text);
CDSRAS.Filtered:=true;
CDSRAS.First;
for I := 0 to CDSRAS.RecordCount - 1 do
begin
Series2.Add(CDSRAS.Znach.Value, ",Shape2.Brush.Color);
CDSRAS.Next;
end;

CDSRAS.Filtered:=false;
CDSRAS.Filter:='KritName='+QuotedStr(ComboBox5.Text);
CDSRAS.Filtered:=true;
CDSRAS.First;
for I := 0 to CDSRAS.RecordCount - 1 do
begin
Series3.Add(CDSRAS.Znach.Value, ",Shape3.Brush.Color);
CDSRAS.Next;
end;
CDSRAS.First;
end;

procedure TForm1.BitBtn4Click(Sender: TObject);
var i,j:integer; Xs,Ys:string; X,Y:Tar;
begin
for I := 1 to Tras.ColCount - 1 do
begin
Tcor.Cells[i,0]:=Tras.Cells[i,0];
Tcor.Cells[0,i]:=Tras.Cells[i,0];
end;
Tcor.RowCount:=Tras.ColCount;
Tcor.ColCount:=Tras.ColCount;
for I := 1 to Tras.ColCount - 1 do
begin
Tcor.Cells[i,0]:=Tras.Cells[i,0];
Tcor.Cells[0,i]:=Tras.Cells[i,0];
Xs:=Tcor.Cells[i,0];
CDSRAS.Filtered:=false;
CDSRAS.Filter:='KritName='+QuotedStr(Xs);
CDSRAS.Filtered:=true;
CDSRAS.First;
X:=FtoA(CDSRAS,'Znach');
for j := 1 to Tras.ColCount - 1 do
begin
Ys:=Tcor.Cells[0,j];
CDSRAS.Filtered:=false;
CDSRAS.Filter:='KritName='+QuotedStr(Ys);
CDSRAS.Filtered:=true;
CDSRAS.First;
Y:=FtoA(CDSRAS,'Znach');
Tcor.Cells[i,j]:=FloatToStr(round(Correl(X,Y)*1000)/1000);
end;
end;
CDSRAS.Filtered:=false;
N10.Enabled:=true;
end;

procedure TForm1.BitBtn5Click(Sender: TObject);
var i,j,k:integer; Mark:real;
begin
k:=1; Mark:=1;
Tpro.RowCount:=2;
Tpro.FixedRows:=1;
Tpro.Cells[1,0]:='Корреляция';
for I := 1 to tcor.RowCount - 1 do
if ComboBox6.Text=tCor.Cells[i,0] then k:=i;
Tpro.Cells[2,0]:='Значение';
for j := 1 to tcor.RowCount - 1 do
begin
if (j<>k)and(strtfloat(tCor.Cells[k,j])>0.5) then
begin
Tpro.Cells[0,Tpro.RowCount-1]:=tCor.Cells[0,j];
Tpro.Cells[1,Tpro.RowCount-1]:=tCor.Cells[k,j];

```

```

    Mark:=Mark*StrToFloat(tCor.Cells[k,j]);
    Tpro.Cells[2,Tpro.RowCount-1]:=0';
    Tpro.RowCount:=Tpro.RowCount+1;
end;
end;
Tpro.RowCount:=Tpro.RowCount-1;
if Tpro.RowCount<2 then showmessage('Прогнозируемый параметр не имеет явно зависимых критериев');
edit1.Text:=floattostr(Round(Mark*1000)/1000);
M1:=Round(Mark*1000)/1000;
BitBtn6.Enabled:=true;
end;

```

```

procedure TForm1.BitBtn6Click(Sender: TObject);
var i,j,k:integer; Xs,Ys:string; mkr,mrk,p,a,b,min,max,Mmrk,mp1,mp2:double; X,Y:tar;
begin
LineSeries1.Clear;
LineSeries2.Clear;
Xs:=ComboBox6.Text;
CDSRAS.Filtered:=false;
CDSRAS.Filter:='KritName='+Quotedstr(Xs);
CDSRAS.Filtered:=true;
X:=FtoA(CDSRAS,'Znach');
k:=length(X);
a:=MAR(X);
b:=MIR(X);
min:=2*a-b;
max:=2*b+a;
p:=min; mp1:=0;mp2:=0; Mmrk:=0;
while p<max do
begin
SetLength(X,k+1);
x[k]:=p; mrk:=1;
for I := 1 to Tpro.RowCount - 1 do
begin
Ys:=Tpro.Cells[0,i];
CDSRAS.Filtered:=false;
CDSRAS.Filter:='KritName='+Quotedstr(Ys);
CDSRAS.Filtered:=true;
Y:=FtoA(CDSRAS,'Znach');
SetLength(Y,length(Y)+1);
Y[length(Y)-1]:=StrToFloat(Tpro.Cells[2,i]);
//sleep(1);
//memo1.Lines.add(floattostr(Correl(X,Y)));
mrk:=mrk*Correl(X,Y);
end;
if mrk>Mmrk then begin Mmrk:=mrk; mp1:=p; end;
memo1.Lines.Add('X =' +floattostr(p)+' Оценка'+floattostr(mrk));
mkr:=strtofloat(edit1.Text);
LineSeries1.AddXY(p,mrk,"");
LineSeries2.AddXY(p,mkr,"");
mrk:=1;
Gauge1.Progress:=Round(p/Max*100);
p:=p+strtofloat(edit4.Text)*(MAX-MIN);
end;
mp2:=Round(Abs(Mkr-Mmrk)*1000)/10;
edit2.Text:=floattostr(Mp1);
edit3.Text:=floattostr(Mp2)+'%';
BitBtn6.Enabled:=false;
Gauge1.Progress:=0;
end;

```

```

procedure TForm1.Button4Click(Sender: TObject);
begin
Setlength(Puper,2);
SetLength(Puper[0],2);
SetLength(Puper[1],2);
Puper[0][0]:=1;
Puper[0][1]:=2;
Puper[1][0]:=3;
Puper[1][1]:=4;
//Memo2.Lines.Add(FloatToStr(Puper[0][0]));
//Memo2.Lines.Add(FloatToStr(Puper[0][1]));
//Memo2.Lines.Add(FloatToStr(Puper[1][0]));
//Memo2.Lines.Add(FloatToStr(Puper[1][1]));
end;

```

```

procedure TForm1.Button5Click(Sender: TObject);
begin
form1.Close;

```



```

end;

procedure TForm1.Button6Click(Sender: TObject);
var Res1,Res2,s1,s2:real; i:integer;
begin
  //for I := 0 to SG.RowCount - 1 do
  begin
    // Res1:=Res1+strtofloat(Sg.Cells[2,i])/SG.RowCount*Sqr(strtofloat(Sg.Cells[1,i])+strtofloat(Edit1.Text));
    // Res2:=Res2+strtofloat(Sg.Cells[2,i])/SG.RowCount*Sqr(strtofloat(Sg.Cells[1,i])-strtofloat(Edit1.Text));
    // s1:=s1*strtofloat(Sg.Cells[1,i]);
  end;
  edit2.Text:=floattostr((Res1-s1)*(1-strtofloat(Edit1.Text)));
  edit3.Text:=floattostr((Res2+s1)*(1-strtofloat(Edit1.Text)));
end;

procedure TForm1.ComboBox1Change(Sender: TObject);
var i:integer;
begin
  ComboBox2.Items.Clear;
  for I := ComboBox1.ItemIndex to ComboBox1.Items.Count - 1 do
    ComboBox2.Items.Add(ComboBox1.Items.Strings[i]);
  ComboBox2.ItemIndex:=ComboBox2.Items.Count-1;
end;

procedure TForm1.ComboBox6KeyPress(Sender: TObject; var Key: Char);
begin
  key:=chr(0);
end;

function TForm1.Correl(X, Y: array of real): Double;
var Exy,Exx,Eyy,Ex,Ey:double;
begin
  Result:=0;
  Exy:=Round(Mo2(X,Y)*1000)/1000;
  //Memo2.Lines.add('Exy='+floattostr(Exy));
  Exx:=Round(Mo2(X,X)*1000)/1000;
  //Memo2.Lines.add('Exx='+floattostr(Exx));
  Eyy:=Round(Mo2(Y,Y)*1000)/1000;
  //Memo2.Lines.add('Eyy='+floattostr(Eyy));
  Ex:=Round(Mo(X)*1000)/1000;
  //Memo2.Lines.add('Ex='+floattostr(Ex));
  Ey:=Round(Mo(Y)*1000)/1000;
  //Memo2.Lines.add('Ey='+floattostr(Ey));
  // if (Eyy-Ey*Ey=0)or(Exx-Ex*Ex=0)or((Exx-Ex*Ex)*(Eyy-Ey*Ey)<0) then Result:=13 else
  Result:=(Exy-Ex*Ey)/Sqrt((Exx-Ex*Ex)*(Eyy-Ey*Ey));
  memo1.Lines.Add('Корреляция =' + floattostr(Result));
  // if (ABS(result)>1)or((ABS(result)<0.0001)) then Result:=0;
  if Result>1 then result:=1;
end;

function TForm1.Correl2(X, Y: array of real): Double;
var Exy,Ex,Ey:double;
begin
  Exy:=S2(X,Y);
  Ex:=S1(X);
  Ey:=S1(Y);
  Result:=Exy/(Ex*Ey);
end;

procedure TForm1.DBGrid5CellClick(Column: TColumn);
begin
  ID_Krit:=DataModule4.KritID_Krit.Value;
  DataModule4.Main.Filtered:=false;
  DataModule4.Main.Filter:='ID_krit = '+inttostr(ID_Krit);
  DataModule4.Main.Filtered:=True;
end;

procedure TForm1.DBGrid5MouseWheelDown(Sender: TObject; Shift: TShiftState;
  MousePos: TPoint; var Handled: Boolean);
begin
  Form1.DBGrid5CellClick(nil);
end;

procedure TForm1.Edit1KeyPress(Sender: TObject; var Key: Char);
begin
  case key of
    '0'..'9',',',chr(8):key:=key;
  else key:=chr(0);
  end;
end;

```

```

end;
end;

procedure TForm1.FormCreate(Sender: TObject);
begin
WebBrowser1.Navigate('c:\Program files\Model\Help.mht');
end;

function TForm1.FtoA(DSet: Tdataset; FN: String): TAR;
var i:integer;
begin
SetLength(Result,0);
DSet.First;
for I := 0 to DSet.RecordCount - 1 do
begin
SetLength(Result,Length(Result)+1);
if DSet.FieldByName(FN).Value=NULL then
Result[Length(Result)-1]:=0 else Result[Length(Result)-1]:=DSet.FieldByName(FN).AsFloat;
Memo2.Lines.Add(floattostr(Result[Length(Result)-1]));
DSet.Next;
end;
DSet.First;
end;

function TForm1.MO(X: array of real): Double;
var i,j:integer; s:string;
begin
s:='Математическое ожидание (';
Result:=0;
j:=Length(x);
for I := 0 to j-1 do
begin
Result:=Result+X[i];
s:=s+floattostr(X[i])+' ';
end;
Result:=result/j;
s:=s+') = '+floattostr(result);
Memo1.Lines.Add(S);
end;

function TForm1.MO2(X, Y: array of real): Double;
var i,j:integer; s:string;
begin
Result:=0;
s:='Дисперсия (';
j:=length(X);
for I := 0 to Length(x)-1 do
begin
Result:=Result+(X[i]*Y[i]);
s:=s+floattostr(X[i])+' ';
end;
Result:=Result/j;
s:=s+') = '+floattostr(result);
Memo1.Lines.Add(S);

end;

procedure TForm1.N10Click(Sender: TObject);
begin
panel1.visible:=false;
panel3.visible:=false;
panel4.visible:=false;
panel5.visible:=false;
panel6.visible:=true;
panel7.visible:=false;
end;

procedure TForm1.N11Click(Sender: TObject);
begin
panel1.visible:=true;
panel3.visible:=false;
panel4.visible:=false;
panel5.visible:=false;
panel6.visible:=false;
panel7.visible:=false;

end;

procedure TForm1.N12Click(Sender: TObject);

```

```

begin
Form1.Close;
end;

procedure TForm1.N5Click(Sender: TObject);
begin
panel1.visible:=false;
panel3.visible:=true;
panel4.visible:=false;
panel5.visible:=false;
panel6.visible:=false;
panel7.visible:=false;
end;

procedure TForm1.N6Click(Sender: TObject);
begin
panel1.visible:=false;
panel3.visible:=false;
panel4.visible:=true;
panel5.visible:=false;
panel6.visible:=false;
panel7.visible:=false;
end;

procedure TForm1.N7Click(Sender: TObject);
begin
panel1.visible:=false;
panel3.visible:=false;
panel4.visible:=false;
panel5.visible:=false;
panel6.visible:=false;
panel7.visible:=true;
end;

procedure TForm1.N8Click(Sender: TObject);
begin
Form5:=Tform5.Create(Self);
Form5.ShowModal;
Form5.free;
end;

procedure TForm1.N9Click(Sender: TObject);
begin
panel1.visible:=false;
panel3.visible:=false;
panel4.visible:=false;
panel5.visible:=true;
panel6.visible:=false;
panel7.visible:=false;
end;

function TForm1.S1(X: array of real): double;
var i:integer; Xr:double;
begin
Xr:=Sred(X);
result:=0;
for I := 0 to Length(x) - 1 do
  result:=result+Sqr(x[i]-Xr);
Result:=Sqrt(result);
end;

function TForm1.S2(X, Y: array of real): Double;
var i:integer; xr,yr:double;
begin
Xr:=Sred(X);
Yr:=Sred(Y);
for I := 0 to length(X) - 1 do
  result:=Result+(x[i]-xr)*(y[i]-yr);
end;

procedure TForm1.SpeedButton1Click(Sender: TObject);
begin
Form2:=TForm2.Create(Self);
Form2.ShowModal;
Form2.Free;
end;

procedure TForm1.SpeedButton2Click(Sender: TObject);
begin

```

```

DataModule4.Krits.Delete;
DataModule4.Main.Close;
DataModule4.Main.Open;
end;

procedure TForm1.SpeedButton3Click(Sender: TObject);
begin
Form3:=TForm3.Create(Self);
Form3.Id_Krit:=ID_Krit;
Form3.ShowModal;
Form3.Free;
end;

procedure TForm1.SpeedButton4Click(Sender: TObject);
begin
DataModule4.Main.Delete;
end;

function TForm1.Sred(X: array of real): double;
var i:integer;
begin
result:=0;
for I := 0 to Length(x) - 1 do
  Result:=Result+x[i];
  Result:=Result/Length(x);
end;

end.

```