



НАЦИОНАЛЬНЫЙ БАНК КАЗАХСТАНА

ДИСКРИМИНАНТНЫЙ АНАЛИЗ В ОЦЕНКЕ КРЕДИТНЫХ РИСКОВ

**Департамент исследований и статистики
Экономическое исследование №2018-2**

Керимхан Жанар

Экономические исследования и аналитические записки Национального Банка Республики Казахстан (далее – НБРК) предназначены для распространения результатов исследований НБРК, а также других научно-исследовательских работ сотрудников НБРК. Экономические исследования распространяются для стимулирования дискуссий. Мнения, высказанные в документе, выражают личную позицию автора и могут не совпадать с официальной позицией НБРК.

Автор выражает благодарность Константину Орлову и Ернуру Даней за помощь в проведении исследования и полезные комментарии.

Дискриминантный анализ в оценке кредитных рисков.

Август 2018 года

NBRK – WP – 2018 – 2

Дискриминантный анализ в оценке кредитных рисков

Аннотация

Дискриминантный анализ осуществляется путем построения прогностической модели для классификации объектов на определенные группы, в зависимости от характеристик, предсказывающих их будущее поведение. Разделение на группы происходит с помощью дискриминантной функции, которая строится по набору наблюдений, для которых принадлежность к группе уже известна, то есть по обучающей выборке¹. Иными словами, задачей дискриминантного анализа является определение правила, по которому объекты будут отнесены к какой-либо группе.

В рамках данного исследования построена модель, которая позволила выявить наиболее информативные индикаторы финансового состояния предприятия и определить приближение кризисной ситуации на предприятии (вероятность банкротства), и на этой основе классифицировать их по степени надежности. Предприятия в данном случае рассматриваются в качестве потенциальных заемщиков, от результатов производственно-хозяйственной деятельности которых зависят стабильность и устойчивость банковского сектора.

Для эмпирических оценок были использованы данные ежеквартального мониторинга предприятий Национального Банка. На основе финансовых показателей предприятий из обучающей выборки построены функции классификации. Полученные функции были проверены на данных тестовой выборки, где заранее известна принадлежность к группе. В результате, правильность распределения составила 93,3%.

Ключевые слова: Дискриминантная функция, классификационная функция, модели прогнозирования банкротства, мониторинг предприятий, обучающая выборка.

Классификация JEL: C38; G33.

¹ Обучающая выборка – это выборка, где заранее известна принадлежность объектов к той или иной группе.

Содержание

1. Введение	3
2. Обзор литературы	5
3. Методология исследования и исходные данные	8
4. Обсуждение результатов	13
5. Заключение	19
6. Список литературы	21
7. Приложения	22

1. Введение

Одним из важных условий для устойчивого роста экономики является стабильность функционирования финансовой системы, которая, в свою очередь, зависит от эффективности банковских вложений в реальный сектор экономики. Согласно статистическим данным, по итогам 2017 года большую часть ссудного портфеля банков (70%) занимают займы корпоративному сектору, которые имеют более низкое качество, чем займы физическим лицам². Доля займов с просроченной задолженностью свыше 90 дней в ссудном портфеле банков составляет около 10%, а доля сомнительных займов еще больше, которые к тому же не покрываются провизиями. Высокая доля таких кредитов может неблагоприятно отразиться на ликвидности банков и даже снизить устойчивость банковского сектора.

Основной причиной возникновения проблемных кредитов является неадекватная оценка кредитного риска или ухудшение финансового положения предприятий-заемщиков.

Целью данного исследования является определение переменных, дающих наилучшее предсказание банкротства и позволяющих наилучшим образом дискриминировать предприятия по степени надежности. В этих целях использованы данные предприятий-участников мониторинга реального сектора.

Дискриминантный анализ на основе использования прошлого опыта позволяет предсказать поведение новых наблюдений (предприятий) и определить их принадлежность к определенной группе по принципу максимального сходства характеристик новых наблюдений с признаками обучающей выборки.

Преимущество дискриминантного анализа по сравнению с другими методами классификации предприятий (методом скоринга, рейтинговой оценки или методом, основанным на оценке соответствия нормативным значениям финансовых показателей) заключается в том, что при построении дискриминантной функции автоматически определяются переменные, различающие группы наилучшим способом. В отличие от рейтинговых оценок, которые являются сравнительной оценкой предприятий, дискриминантный анализ является методом прогнозирования банкротства, основанным на прошлом опыте. Этот метод позволяет предсказывать поведение наблюдаемых объектов на основе имеющихся стереотипов поведения аналогичных объектов.

Данное исследование состоит из трех частей: в первом разделе приводится обзор литературы, где описываются сферы применения дискриминантного анализа, а также существующие модели прогнозирования банкротства. Во второй части поясняются теоретические аспекты проведения дискриминантного анализа, описываются переменные, используемые для

² Отчет о финансовой стабильности за 2015-2017 годы, стр. 94
<http://www.nationalbank.kz/?docid=619&switch=russian>

построения модели, процесс построения дискриминантной функции и классификации предприятий. В третьей части представлены полученные результаты, рассматривается возможность применения дискриминантного анализа в аналитической работе Национального Банка. В заключении делаются выводы, а также рекомендации для дальнейшего исследования.

2. Обзор литературы

Обзор литературы показал, что дискриминантный анализ широко используется не только при построении кризис-прогнозных моделей в экономике, но и в медицине (для прогнозирования состояния пациента в зависимости от переменных состояния больного), маркетинге (для определения поведения покупателей в зависимости от привычек, интересов), в области образования (для определения выбора студентов колледжа о дальнейшем обучении), в психологии, социологии и т.д.

Методы дискриминантного анализа разрабатывались, начиная с 1950-х годов, такими учеными как Прасанта Чандра Махаланобис (индийский экономист и статистик, 1893-1972), Гарольд Хоттеллинг (американский экономист и статистик, 1895-1973), Рональд Фишер (английский статистик, биолог-эволюционист, генетик, 1890-1962), и другими [1].

Впервые метод дискриминантного анализа был применен в банковской сфере для определения характеристик заемщика, погасившего кредит в срок, и разделения в отдельную группу от тех, кто не погасил или погасил с задержкой. Результаты исследования использовались для принятия решения относительно новых заемщиков по имеющейся характеристике заемщиков, погасивших или не погасивших кредит.

Один из примеров применения дискриминантного анализа в медицине – исследование группы российских ученых для моделирования обобщенных показателей уровня здоровья медицинских работников онкологического диспансера. В качестве признаков для проведения дифференцирования использовались такие показатели, как работа в отделении лучевой диагностики, работа в хирургических отделениях, работа с излучающей аппаратурой, наличие у работника среднего или высшего профессионального образования, возраст сотрудника, пол и стаж работы в отделении [2].

В оценке экономического положения предприятий с применением дискриминантной функции в зарубежных странах широкое распространение получили Z-анализ Альтмана, Z-модель Лиса, Z-модель Таффлера и др.

Фундаментальные исследования в этой области принадлежат профессору Нью-Йоркского университета Э. Альтману, который разработал модель определения кредитоспособности субъектов хозяйствования на основе исследования финансовых показателей 66 предприятий, из которых 33 потерпели банкротство (Z-score). В общем случае модель Альтмана имеет вид:

$$Z=A1*X1+A2*X2+...An*Xn \quad (1)$$

Где, $A1...An$ – коэффициенты или веса соответствующих переменных;
 $X1...Xn$ – финансовые коэффициенты предприятий;
 Z – значение оценки.

На сегодняшний день в экономической литературе известны 7 моделей Альтмана (Двухфакторная модель Альтмана; Пятифакторная модель Альтмана для компаний, чьи акции котируются на бирже; Модель Альтмана для компаний, чьи акции не торгуются на биржевом рынке; Z-модель Альтмана для непроизводственных компаний; Модель Альтмана для развивающихся рынков; Модель Альтмана-Сабато; Семифакторная модель), которые отличаются друг от друга значимостью весовых коэффициентов и факторами риска [3].

Самая популярная из них – пятифакторная модель, опубликованная в 1968 году, в которую вошли 5 из 22 финансовых показателей:

$$Z = 1,2X_1 + 1,4X_2 + 3,3X_3 + 0,6X_4 + X_5 \quad (2)$$

Где, X_1 - оборотный капитал к сумме активов предприятия.

X_2 - нераспределенная прибыль к сумме активов предприятия

X_3 - прибыль до налогообложения к общей стоимости активов.

X_4 - рыночная стоимость собственного капитала / бухгалтерская (балансовая) стоимость всех обязательств.

X_5 - объем продаж к общей величине активов предприятия.

В результате подсчета Z-показателя для конкретного предприятия делается заключение:

- Если $Z < 1,81$ – вероятность банкротства составляет от 80 до 100%;
- Если $2,77 \leq Z < 1,81$ – средняя вероятность краха компании от 35 до 50%;
- Если $2,99 < Z < 2,77$ – вероятность банкротства не велика от 15 до 20%;
- Если $Z \leq 2,99$ – ситуация на предприятии стабильна, риск неплатежеспособности в течение ближайших двух лет крайне мал.

Точность прогноза в этой модели на горизонте одного года составляет 95%, на два года – 83%, что является ее достоинством [3].

Далее на основе работ Альтмана были проведены многочисленные исследования с целью оценки финансового состояния предприятий, например модель Спрингейта, разработанная Гордоном Л. В. в университете Симона Фрейзера (Канада) в 1978 году, который, следуя предложенному Альтманом подходу, исследовал 19 наиболее часто используемых финансовых коэффициентов. В окончательном варианте модели осталось только четыре показателя.

$$Z = 1.03 * K_1 + 3.07 * K_2 + 0.66 * K_3 + 0.4 * K_4 \quad (3)$$

Где, K_1 - Оборотные средства/Валюта баланса;

K_2 - Прибыль до н/о/Валюта баланса;

K_3 - Прибыль до налогообложения/Краткосрочные обязательства;

K_4 - Выручка/Валюта баланса.

Если Z-значение предприятия меньше 0,862, то предприятие получает оценку «крах». При создании данной модели автор использовал данные 40 предприятий и достиг 92,5% точности предсказания банкротства на год вперед [4].

Также следует отметить разработку российских ученых Иркутской государственной экономической академии, которые разработали R-модель прогноза риска банкротства.

$$R = 8,38 * K1 + K2 + 0,054 * K3 + 0,63 * K4 \quad (4)$$

Где, K1 – коэффициент обеспеченности собственными оборотными средствами;

K2 – рентабельность собственного капитала;

K3 – оборачиваемость активов;

K4 – отношение чистой прибыли к сумме себестоимости продукции, коммерческих и управленческих расходов.

Результаты интерпретируются в соответствии со следующей шкалой:

- меньше 0 - вероятность банкротства: Максимальная (90%-100%).
- 0-0,18 - вероятность банкротства: Высокая (60%-80%).
- 0,18-0,32 - вероятность банкротства: Средняя (35%-50%).
- 0,32-0,42 - вероятность банкротства: Низкая (15%-20%).
- больше 0,42 - вероятность банкротства: Минимальная (до 10%) [5].

Сайфуллин Р.С. и Кадыков Г.Г. предложили использовать для оценки финансового состояния предприятий рейтинговое число:

$$R = 2K_o + 0,1K_{тл} + 0,08K_{и} + 0,45K_{м} + K_{пр} \quad (5)$$

Где, K_о – коэффициент обеспеченности собственными средствами;

K_{тл} – коэффициент текущей ликвидности;

K_и – коэффициент оборачиваемости активов;

K_м – коммерческая маржа (рентабельность реализации продукции);

K_{пр} – рентабельность собственного капитала.

При полном соответствии финансовых коэффициентов их минимальным нормативным уровням рейтинговое число будет равно единице, то есть предприятие имеет удовлетворительное финансовое состояние. Финансовое состояние предприятий с рейтинговым числом менее единицы характеризуется как неудовлетворительное [6].

Таким образом, использование дискриминантного анализа позволит отобрать наиболее информативные финансовые коэффициенты из обучающей выборки и на их основе отнести предприятия к той или иной группе риска.

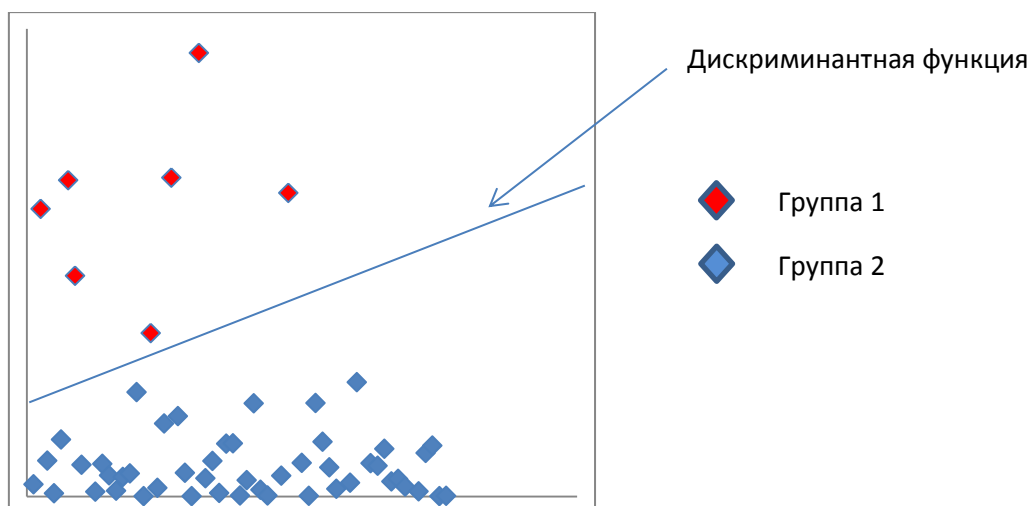
3. Методология исследования и исходные данные

В данном разделе поясняются теоретические аспекты проведения дискриминантного анализа, описываются переменные, используемые для построения модели, процесс построения дискриминантной функции и классификации предприятий.

Наиболее наглядно теория дискриминантного анализа может быть объяснена для случая двух классов. На рисунке 1 представлено разделение наблюдений на две категории с помощью одной дискриминантной функции. Для разделения на 3 категории нужны две дискриминантные функции, на 4 категории – три и т.д.

Рисунок 1

Визуализация разделения наблюдений с помощью дискриминантной функции в двухмерном пространстве



В n -мерном пространстве речь идет о поиске линейной комбинации вида

$$F(x) = \sum_{n=1}^n a_n X_n \quad (6)$$

Где, $x_1 \dots x_n$ – дискриминантные переменные
 $a_0 \dots a_n$ – веса дискриминантных переменных
 n – число переменных

Основная идея дискриминантного анализа заключается в том, чтобы определить, отличаются ли совокупности по среднему значению какой-либо переменной (или линейной комбинации переменных), и затем использовать эту переменную, чтобы предсказать для новых членов их принадлежность к той или иной группе. Если среднее значение определенной переменной заметно различается для двух совокупностей, то можно сказать, что переменная разделяет данные совокупности[7].

Процедура дискриминантного анализа состоит из двух основных частей:

- определение наиболее информативных характеристик (дискриминантных переменных) групп по обучающей выборке и построение на этой основе дискриминантной функции вида:

$$F(x) = \sum_{n=1}^n a_n X_n \quad (6)$$

- классификация новых объектов (наблюдений) с помощью полученной модели, подставляя значения переменных каждого объекта в дискриминантную функцию.

Непосредственное выполнение анализа можно разделить на несколько этапов:

1. Подготовка данных;
2. Выбор лучших переменных из подготовленных данных;
3. Построение модели на основе дискриминантных уравнений;
4. Анализ полученной модели и ее проверка на тестовых данных.

Подготовка данных

Дискриминантные переменные, которые используются в качестве признаков, отличающих один класс от другого, должны отвечать некоторым требованиям:

- 1) дискриминантные переменные должны быть количественными и линейно независимыми;
- 2) каждая из дискриминантных переменных должна быть подчинена нормальному закону распределения;
- 3) количество объектов каждого класса должно превышать число предикторов (эмпирическое правило – не менее чем в 5 раз) [8];
- 4) наблюдения в обучающей классификации должны быть представлены в равных пропорциях. Это связано с тем, если в обучающей выборке большинство наблюдений относится к одному классу, то вероятность отнесения нового наблюдения в тот же класс выше.

Учитывая вышеуказанные рекомендации, необходимые для снижения ошибки классификации, для данного исследования отобраны 300 предприятий-участников мониторинга с разным положением финансовой устойчивости. В качестве дискриминирующих переменных исследуются 11 финансовых коэффициентов, отвечающих всем требованиям, предъявляемым к дискриминантным переменным (Таблица 1).

Исходные данные

	Показатель	Формула расчета
X ₁	Темп роста собственного капитала (CAGR)	$(СК_к/СК_н)^{1/3}-1$
X ₂	Темп роста ДРП на одного работника (CAGR)	$(ДРПР_к/ДРПР_н)^{1/3}-1$ Где, ДРПР = ДРП за год / Среднегодовая численность занятых
X ₃	Темп роста ДРП (CAGR)	$(ДРП_к/ДРП_н)^{1/3}-1$
X ₄	Коэффициент оборачиваемости основных средств (КООС _н)	ДРП за год /Среднегодовая стоимость долгосрочных активов
X ₅	Коэффициент финансовой независимости (КФН _н)	Среднегодовая стоимость СК/ Среднегодовая стоимость активов
X ₆	Коэффициент оборачиваемости активов (КОА _н)	ДРП за год/Среднегодовая стоимость активов
X ₇	Период погашения дебиторской задолженности (ППДЗ _н)	360/ОДЗ
X ₈	Оборачиваемость дебиторской задолженности (ОДЗ _н)	ДРП за год/Среднегодовое значение краткосрочной дебиторской задолженности
X ₉	Коэффициент текущей ликвидности (КТЛ _н)	Среднегодовая стоимость краткосрочных активов/Среднегодовая стоимость краткосрочных обязательств
X ₁₀	Коэффициент быстрой ликвидности (КБЛ _н)	(Среднегодовая стоимость краткосрочных активов – среднегодовая стоимость запасов)/ Среднегодовая стоимость краткосрочных обязательств
X ₁₁	Валовый коэффициент прибыли (ВКП _н)	$(ДРП-СРП)/ДРП$

ДРП – доход от реализации продукции

ДРПР - доход от реализации продукции на работника

СРП – себестоимость реализованной продукции

СК – собственный капитал (СК_к –на конец периода, СК_н –на начало)

ОДЗ - Оборачиваемость дебиторской задолженности

CAGR - Compound Annual Growth Rate. Характеризует среднегодовую скорость роста показателя в течение определенного времени.

Выбор лучших переменных из подготовленных данных

Для выбора лучших переменных необходимо сделать группировку объектов обучающей выборки. Данный набор информации позволит выявить, какие характеристики отличают объекты, относящиеся к одной группе, от объектов, относящихся к другой группе. Иными словами, какие характеристики наиболее отличают наблюдения одной группы от другой. В этих целях предприятия, участвующие в мониторинге Национального Банка, в зависимости от динамики показателей производственно-хозяйственной деятельности были заранее распределены на 3 группы по типу финансовой устойчивости:

- предприятия с абсолютной устойчивостью («устойчивые») – группа А³;
- предприятия с нормальной устойчивостью («нормальные») – В;
- предприятия с критическим финансовым состоянием («критические») – С.

Распределение предприятий для обучающей выборки было осуществлено экспертным путем. Предприятия с положительным темпом роста дохода от реализации продукции за последние 3 года и положительным коэффициентом финансовой независимости, значениями коэффициентов быстрой и текущей ликвидности выше среднего значения вошли в группу А; в группу С вошли убыточные предприятия с отрицательным темпом роста ДРП и собственного капитала, имеющие низкие значения коэффициентов быстрой и текущей ликвидности; остальные предприятия вошли в группу В.

Исходные данные для обучающей выборки представлены в виде матрицы для каждой группы:

$$X_i = \begin{pmatrix} x_{11} & \cdots & x_{n1} \\ \vdots & \ddots & \vdots \\ x_{1m} & \cdots & x_{nm} \end{pmatrix} \quad (7)$$

Где, i – обозначение группы, к которой относятся наблюдения (А,В,С)

n – количество переменных

m – количество наблюдений

x_{nm} – значение n -й переменной m -го предприятия

Построение модели на основе дискриминантных уравнений

Далее строится дискриминантная функция, представляющая собой линейную функцию и разделяющая наблюдения на группы по нескольким признакам:

$$F(x) = a_1 * x_1 + a_2 * x_2 + \dots + a_n * x_n \quad (8)$$

где $a_1 \dots a_n$ – коэффициенты дискриминантной функции или веса соответствующих переменных

³ Обозначения классов для формул

$x_1 \dots x_n$ – финансовые коэффициенты предприятий (признаки, независимые переменные, или предикторы)

Функция строится так, чтобы ошибка классификации была минимальной. Здесь необходимо оценить значения коэффициентов дискриминантной функции, которые рассчитываются исходя из того, что внутригрупповая вариация рассматриваемых объектов должна быть минимальной, а межгрупповая вариация – максимальной. То есть $a_1 \dots a_n$ определяются таким образом, чтобы средние значения функций классификации в группах (\bar{f}_1 и \bar{f}_2) как можно больше различались между собой, то есть их разница должна быть максимальной. Таким образом, используя транспонированную матрицу отклонений наблюдаемых значений исходных переменных от их средних величин в группе и объединенную ковариационную матрицу, определяется вектор коэффициентов дискриминантной функции ($a_1 \dots a_n$) по следующей формуле:

$$A = S^{-1}(\bar{X}_1 - \bar{X}_2) \quad (9)$$

Где, A – вектор коэффициентов дискриминантной функции
 \bar{X}_1 и \bar{X}_2 – векторы средних значений в первой и второй группах
 S – матрица, обратная совместной ковариационной матрице

Далее значения коэффициентов переменных каждого наблюдения подставляются в формулу (8), полученные значения функций сравниваются с константой (K), наблюдения выше которой относятся к одной группе, ниже – к другой группе.

Константа, равноудаленная от среднего значения дискриминантной функции, рассчитывается по следующей формуле:

$$K = \frac{1}{2}(\bar{f}_1 + \bar{f}_2) \quad (10)$$

где \bar{f}_1, \bar{f}_2 – среднее значение функции для группы 1 и группы 2, соответственно.

Средние значения функции для каждой группы рассчитываются по формуле:

$$\bar{f}_i(x) = a_1 * \bar{x}_{1i} + a_2 * \bar{x}_{2i} + \dots + a_n * \bar{x}_{ni} \quad (11)$$

где $\bar{x}_1 \dots \bar{x}_n$ – средние значения финансовых коэффициентов в группе i ;
 i – группа, для которой рассчитывается среднее значение.

Когда групп больше двух, строятся не дискриминантные функции, а *классификационные функции* для каждой группы. Классификационная функция также представляет собой линейную комбинацию признаков,

которые предназначены для определения, к какой группе с наибольшей вероятностью относятся наблюдения.

$$F_A(x) = c_A + b_{A1} * x_1 + b_{A2} * x_2 + \dots + b_{An} * x_n \quad (12)$$

$$F_B(x) = c_B + b_{B1} * x_1 + b_{B2} * x_2 + \dots + b_{Bn} * x_n \quad (13)$$

$$F_C(x) = c_C + b_{C1} * x_1 + b_{C2} * x_2 + \dots + b_{Cn} * x_n \quad (14)$$

Для определения принадлежности к группе значения признаков новых наблюдений подставляются в классификационные функции ((12)-(14)) и рассчитываются значения классификационных функций для каждого наблюдения. Наблюдения относят к той группе, для которой получен максимальный показатель классификации $F(x)$.

Далее для построения модели на основе эмпирических данных был использован программный пакет «Statistica», который не только строит дискриминантную функцию, автоматически отбрасывая незначимые переменные, но и проводит классификацию новых наблюдений по группам, результаты которой более подробно описаны в следующем разделе.

4. Обсуждение результатов

Для проведения дискриминантного анализа подготовлена обучающая выборка, состоящая из 300 предприятий, которая представлена предприятиями из каждой группы в равных пропорциях: 100 устойчивых, 100 нормальных и 100 критических. Исходные данные представляют собой динамику показателей финансового состояния предприятий за 3 года до момента построения модели в соответствии с таблицей 1.

Обучающая выборка была разделена на тренировочную и тестовую. Тренировочная выборка, состоящая из 270 предприятий (по 90 предприятий из каждой группы), используется для построения дискриминантной и классификационных функций. А остальные 30 предприятий из тестовой выборки (по 10 предприятий из каждой группы) – для оценки качества классификации, то есть для определения степени правильности распределения предприятий по построенной модели.

При применении дискриминантного анализа обычно имеется несколько переменных, и задача состоит в том, чтобы установить, какие из переменных вносят значимый вклад в дискриминацию между совокупностями [7]. Поиск наиболее дискриминирующих переменных осуществляется на основе среднего значения переменных. Если среднее значение определенной переменной заметно различается для двух совокупностей, то можно сказать, что переменная разделяет данные совокупности. Из таблицы 2 видно, что все переменные, кроме X11, имеют значимо различные средние для разных классов.

Средние значения переменных по классам

Class	Means (Лист1 in DB 1st)											Valid N
	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11	
Устойчивые (A)	0,32	0,39	0,33	32,31	0,58	1,16	101,93	13,74	9,15	5,02	0,23	90
Нормальные (B)	0,58	0,09	0,10	11,92	0,53	1,78	72,26	24,90	3,23	2,17	0,24	90
Критические (C)	-0,20	-0,19	-0,19	136,23	0,14	2,84	94,55	26,44	1,19	0,73	0,26	90
All Grps	0,23	0,09	0,08	60,15	0,42	1,93	89,58	21,69	4,52	2,64	0,24	270

Если рассмотреть диаграмму размаха независимых переменных (рисунок 2), можно увидеть что диапазон изменений значений переменной X4 заметно шире. А диаграмма размаха переменной X4 внутри группы показывает, что такая изменчивость вокруг центрального положения в большей степени характерна для класса «Критические» (рисунок 3).

Рисунок 2

Диаграмма размаха независимых переменных

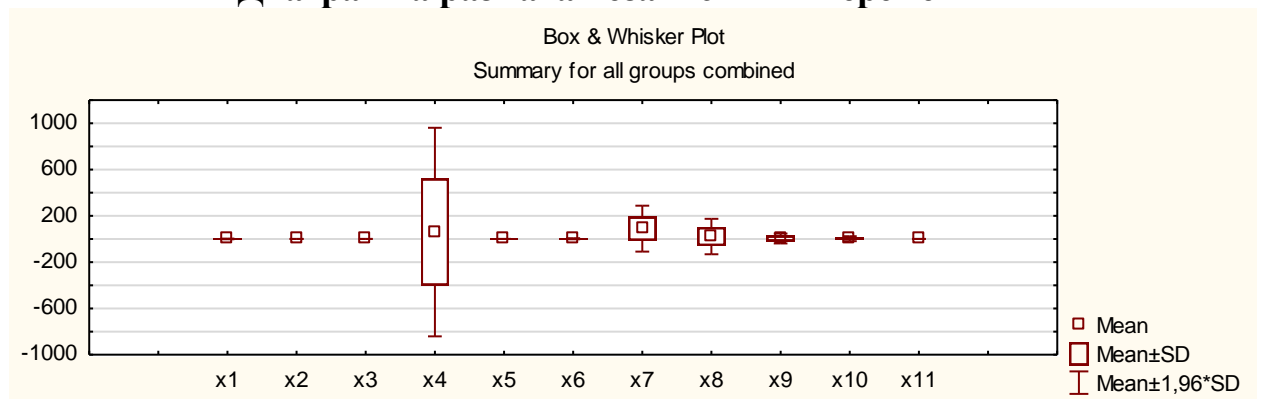
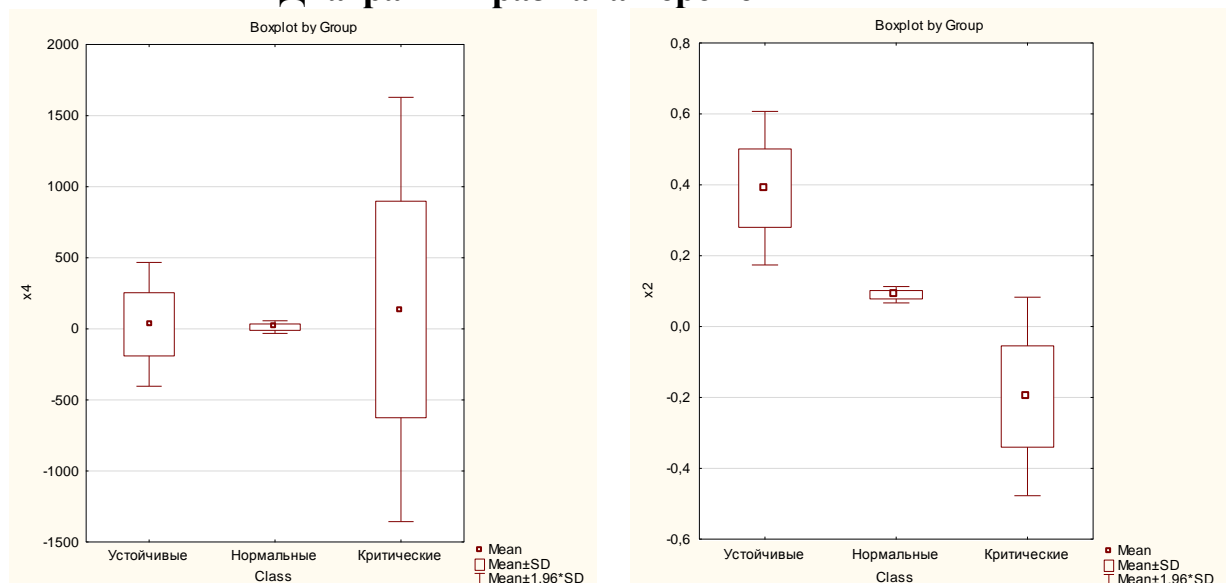


Рисунок 3

Диаграммы размаха переменных X4 и X2



В отличие от переменной X2, центральные точки переменной X4 в разных группах сильно не разнятся, что говорит о том, что переменная X4 не позволяет производить дискриминацию между группами. Такие переменные автоматически исключаются из модели при проведении пошагового дискриминантного анализа, так как они не вносят значимого вклада в дискриминацию наблюдений.

Пошаговый дискриминантный анализ проводится двумя способами: пошаговый с включением, пошаговый с исключением. При первом методе на каждом шаге просматриваются все переменные и выбираются те, которые дают наибольший вклад в различие между совокупностями, а при втором методе – все переменные будут сначала включены в модель, а затем на каждом шаге будут исключаться переменные, вносящие малый вклад в предсказание[7].

В результате дискриминантного анализа, проведенного методом пошагового исключения, из 11 переменных в модели остались только 3 наиболее «важные» (таблица 3).

Таблица 3

Переменные, вошедшие в модель

N=258	Discriminant Function Analysis Summary (Лист1 in DB 20180719 4) Step 8, N of vars in model: 3; Grouping: Class (3 grps) Wilks' Lambda: ,09816 approx. F (6,506)=184,84 p<0,0000					
	Wilks' Lambda	Partial Lambda	F-remove (2,253)	p-value	Toler.	1-Toler. (R-Sqr.)
Темп роста СК	0,118225	0,830256	25,8627	0,000000	0,899199	0,100801
Темп роста ДРП на одного работника	0,435574	0,225351	434,8456	0,000000	0,997768	0,002232
Коэффициент финансовой независимости на начало периода	0,133040	0,737802	44,9552	0,000000	0,898647	0,101353

Лямбда Уилкса – это отношение меры внутригрупповой вариации к общей мере вариации. Значение лямбды Уилкса меняется от 0 до 1 и показывает степень дискриминации в модели. Чем ближе ее значение к 0, тем лучше статистическая значимость мощности дискриминации. В нашем случае Лямбда Уилкса = 0,09816, то есть можно говорить, что есть и достоверные различия между группами, и качественная классификация.

Значения Лямбды Уилкса в первом столбце таблицы 3 показывают значения Лямбды Уилкса для модели, где соответствующая переменная не используется. Из таблицы 3 видно, что без переменной «Темп роста ДРП на одного работника», качество дискриминации было бы значительно хуже.

Partial Lambda – отношение Лямбды Уилкса после добавления данной переменной к Лямбде Уилкса до ее добавления. То есть чем меньше значение данного показателя, тем ценнее данная переменная.

Толерантность (Toler.) является мерой избыточности переменной в модели. Чем меньше ее значение, тем меньшую дополнительную информацию несет переменная.

Низкое значение P-value говорит о статистической значимости данных переменных. Таким образом, из таблицы 3 следует, что данные переменные могут быть успешно включены в модель.

Здесь следует отметить, что по модели, построенной при первоначальной попытке, когда выборка состояла из 270 предприятий, процент правильно классифицированных наблюдений составил 97,7% (264 объекта), то есть не все объекты попали в исходные группы.

Процедура построения модели, позволяющей получить 100%-ю правильность распределения, состоит из нескольких итерации, после каждой итерации из выборки исключаются конкретные объекты с несовпадением (в данном случае 6 предприятий, попавших в другие группы).

В результате 100%-я правильность распределения была достигнута после нескольких итераций по модели, построенной на основе данных 258 предприятий, из которых «устойчивых» – 89, «нормальных» – 89, «критических» – 80 (таблица 4). Разделение данных наблюдений с помощью дискриминантных функций представлено на рисунке 4.

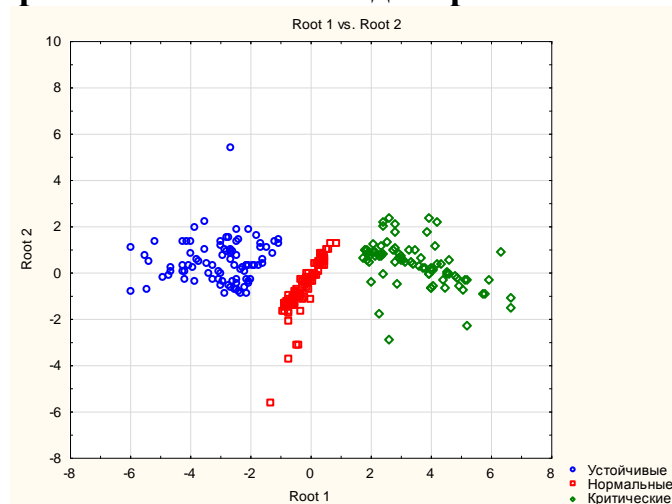
Таблица 4

Классификационная матрица

Group	Classification Matrix (Лист1 in DB 20180719 1st) Rows: Observed classifications Columns: Predicted classifications			
	Percent Correct	Устойчивые p=,33333	Нормальные p=,33333	Критические p=,33333
Устойчивые	100	89	0	0
Нормальные	100	0	89	0
Критические	100	0	0	80
Total	100	89	89	80

Рисунок 4

Диаграмма рассеяния на основе дискриминантных функций



Теперь можно строить классификационные функции, которые позволяют распределить новые наблюдения по группам. В таблице 5 представлены коэффициенты для каждой переменной, вошедшей в классификационные функции.

Таблица 5

Классификационные функции

Variable	Classification Functions; grouping: Class (Лист1 in DB 20180719 4)		
	Устойчивые p=,33333	Нормальные p=,33333	Критические p=,33333
Темпы роста СК	0,9631	1,2636	-0,3090
Темпы роста ДРП на одного работника	38,1442	9,2163	-20,1732
КФН _n	11,8331	10,9004	1,1091
Константа	-12,1982	-4,8299	-3,4049

Таким образом, классификационные функций для каждой группы выглядят следующим образом:

$$\text{Устойчивые: } F_A(x) = -12,1982 + 0,9631 * x_1 + 38,1442 * x_2 + 11,8331 * x_5 \quad (15)$$

$$\text{Нормальные: } F_B(x) = -4,8299 + 1,2636 * x_1 + 9,2163 * x_2 + 10,9004 * x_5 \quad (16)$$

$$\text{Критические: } F_C(x) = -3,4049 - 0,3090 * x_1 - 20,1732 * x_2 + 1,1091 * x_5 \quad (17)$$

Где, x_1 – Темпы роста собственного капитала за последние 3 года;
 x_2 – Темпы роста ДРП на одного работника за последние 3 года;
 x_5 – Коэффициент финансовой независимости на начало периода.

Далее для распределения новых наблюдений по группам необходимо подставить значения показателей каждого предприятия в каждую классификационную функцию ((15)-(17)). Новые объекты будут относиться к той группе, для которой получен максимальный показатель классификации $F(x)$.

Полученные функции были проверены на тестовых данных из 30 наблюдений, которые не участвовали в построении модели. В результате достигнута 93,3%-я правильность распределения (из 30 предприятий 28 были распределены правильно, 2 предприятия из группы «Критические» попали в группу «Нормальные»).

Таким образом, учитывая значение Лямбды Уилкса (0,09816) полученной модели, построенной на основе данных за последние 3 года, можно утверждать, что с вероятностью 90,2% (=1-0,09816), новые наблюдения будут относиться к той или иной категории на горизонте 3 лет, если предприятие не изменит стратегию или не изменится состояние экономики.

Используя построенные функции классификации, предприятия, участвующие в мониторинге⁴, были распределены на группы, и по каждой группе были рассчитаны средние значения коэффициента оборачиваемости активов (таблица 6).

Таблица 6

Среднее значение коэффициента оборачиваемости активов (КОА)⁵

	Устойчивые	Нормальные	Критические	Средневзвешенное
Крупные	1.360	1.126	1.117	1.169
Средние	1.883	1.785	0.845	1.500
Малые	2.771	2.910	1.094	2.275
По всем	2.104	2.020	1.002	1.727

Из таблицы 6 следует, что чем хуже финансовая устойчивость предприятий, тем ниже коэффициент оборачиваемости активов, который характеризует эффективность использования активов.

Результаты распределения предприятий по степени надежности в разрезе отраслей, регионов, а также по размерности предприятий по состоянию на конец 2017 года представлены в приложениях 7-9, а прогнозы распределения через 3 года – в приложениях 10-12.

Относительно кредитного риска со стороны реального сектора по результатам распределения можно сделать следующие выводы.

Около 30% предприятий реального сектора, участвующих в мониторинге Национального Банка, имеют критическое финансовое состояние. Доля таких предприятий в строительстве заметно больше, чем в других отраслях (41,2%), наименьшая доля предприятий с критическим положением приходится на предприятия электроснабжения и водоснабжения (таблица 7). В данной выборке доля предприятий с устойчивым финансовым состоянием по экономике составила 20%.

В региональном разрезе доля предприятий с критическим состоянием в г.Алматы выше, чем в других регионах (38,5%), наименьшая доля таких предприятий в Восточно-Казахстанской области (22,9%), а наибольшая доля устойчивых предприятий приходится на Мангистаускую область – 27% (таблица 8). Большинство крупных предприятий имеют нормальное финансовое состояние (60,2%). Доля ненадежных предприятий среди малых предприятий заметно выше, чем среди крупных (таблица 9).

⁴ В выборку вошли 2130 предприятий, стабильно участвовавшие в опросах Национального Банка в течение последних 3 лет.

⁵ КОА рассчитывается как отношение дохода от реализации продукции к среднегодовой стоимости активов. Показатель характеризует эффективность использования активов с точки зрения объема продаж. При избыточных вложениях в активы сокращается свободный денежный поток. В случае недостаточных инвестиций в активы, как правило, снижаются объемы продаж [6]

Как показывают результаты прогноза по построенной модели, в целом по республике ситуация на горизонте 3 лет значительно не изменится, хотя в Актыобинской области заметно вырастет доля нестабильных предприятий (с 34,3% до 38,1%).

5. Заключение

Полученные результаты позволяют сделать следующие выводы:

1. Для предсказания финансового состояния предприятия более важное значение имеют не статические показатели производственно-хозяйственной деятельности, а динамические. Во многих исследованиях, посвященных дискриминантному анализу, для оценки вероятности банкротства предприятий используются финансовые коэффициенты по состоянию на определенный период. А результаты данного исследования показывают, что стратегия управления предприятием может значимо изменить положение предприятия. К примеру, предприятия, где рост дохода сопровождается еще большим ростом численности занятых, могут оказаться в критическом положении. Данный показатель (X2) оказался самым значимым в построенной модели, это означает, что показатель, характеризующий рост прибыльности на одного работника, наилучшим образом распределяет предприятия по степени надежности.

2. Для более достоверной оценки кредитного риска со стороны реального сектора целесообразно применение полученной модели для классификации предприятий генеральной совокупности. В этих целях необходимо использование данных официальной статистики Комитета по статистике, как для малых, так и для крупных и средних предприятий. В дальнейшем такой анализ может быть проведен при наличии необходимых данных для расчета переменных, вносящих значимый вклад в дискриминацию между группами.

3. Финансовые коэффициенты, хотя и являются ключевыми, не единственные факторы прогнозирования банкротства предприятий. Стабильность предприятий зависит не только от внутреннего менеджмента в компании и показателей, которые вошли в модель, но также значимо могут зависеть от состояния экономики. Поэтому построение модели для прогнозирования банкротства с учетом будущего состояния экономики может стать объектом дальнейших исследований по данной теме. В качестве показателя, отражающего будущее состояние экономики, могут быть использованы прогнозы ВВП и Композитный опережающий индикатор⁶,

⁶ Композитный опережающий индикатор (КОИ) используется для определения поворотных точек делового цикла и дает качественную информацию о состоянии и направлении динамики экономической активности. КОИ отражает обобщенную оценку мнений руководителей опрошенных предприятий относительно сложившейся и ожидаемой ситуации на предприятиях и обладает свойством опережения динамики реального ВВП на 1-2 квартала. Построение КОИ основано на методологии ОЭСР («OECD System of Composite Leading Indicators, *Methodology Guideline*», OECD 2012).

построенный на данных мониторинга предприятий Национального Банка. Для достижения более достоверного распределения целесообразно рассмотреть возможность построения квадратических дискриминантных функций.

4. Результаты данного исследования, кроме как для отражения оценки кредитного риска со стороны реального сектора в ежеквартальных обзорах по мониторингу предприятий, могут быть использованы:

- банками для выявления важных характеристик, определяющих вероятность банкротства заемщиков;
- предприятиями для корректировки внутреннего менеджмента (для предупреждения банкротства) при необходимости.

6. Список литературы

1. Классификация с обучением. Дискриминантный анализ. Основные понятия. <https://studme.org>
2. Ермолина Т.А., Шиловская Н.А., Мартынова Н.А., Калинин А.Г., Красильников С.В., Оценка состояния медицинских работников с применением многомерного статистического анализа\ Казаньский медицинский журнал №2, 2012
3. Модель Альтмана (Z-score). Анализ финансового состояния предприятия.
http://afdanalyse.ru/publ/finansovyj_analiz/1/bankrot_1/13-1-0-10
4. Сайт Финансовый менеджер (FM)
http://financem.info/bankruptcy_model_springate.html
5. Прогноз ИГЭА риска банкротства (иркутская модель)
http://afdanalyse.ru/publ/finansovyj_analiz/1/prognoz_igeha_riska_bankrotstva/13-1-0-40
6. Инна Сорокина. (2012) Оценка вероятности банкротства предприятия-заемщика. www.bankir.ru
7. Дискриминантный анализ. Электронный учебник по статистике Statsoft. <http://statsoft.ru/home/textbook/modules/stdiscan.html>
8. Виктор Гречков Дискриминантный анализ. Статья в электронном ресурсе «Лаборатория» www.advlab.ru
9. А.А. Макарова. Использование средств многомерной классификации в банковской сфере. Вестник СамГУ. 2014. № 8 (119)
10. Финансы организаций (предприятий) учебно-методический материал
https://www.audit-it.ru/finanaliz/terms/analysis/altman_z_model.html
11. www.accountingverse.com / Financial ratio analysis – раздел сайта по анализу показателей финансовых отчетов
12. Бригхэм Ю., Эрхардт М. Анализ финансовой отчетности // Финансовый менеджмент = Financial management. Theory and Practice. — 10-е изд./Пер. с англ. под. ред. к.э.н. Е. А. Дорофеева.. — СПб.: Питер, 2007.
13. А.А. Халафян, Е.Ю. Пелипенко. Дискриминантный анализ в определении кредитоспособности предприятий.
14. Дискриминантный анализ в Statistica
15. Г.П. Бессокирная. (2003), Статистические методы и анализ данных. Дискриминантный анализ для отбора информативных переменных.
16. Н.Н. Буреева. (2007), Многомерный статистический анализ с использованием ППП Statistica, учебно –методический материал
17. АНАЛИЗ ДАННЫХ. Учебник для академического бакалавриата.pdf
18. А.А. Халафян, Е.Ю. Пелипенко. (2010) Оценка платежеспособности российских предприятий на основе современных технологий статистического моделирования. Научный журнал КУБ ГАУ №62(08)

7. Приложения

Оценка текущего периода

Таблица 7

Распределение предприятий по степени устойчивости в разрезе отраслей

	Устойчивые	Нормальные	Критические
Сельское хозяйство	23.8%	49.2%	26.9%
Горнодобывающая пром-сть	23.3%	49.6%	27.1%
Обрабатывающая пром-сть	26.3%	49.3%	24.5%
Электроснабжение	9.3%	70.7%	20.0%
Водоснабжение	12.0%	88.0%	0.0%
Строительство	20.6%	38.2%	41.2%
Торговля	20.6%	44.4%	34.9%
Транспорт и складирование	15.4%	53.8%	30.8%
Услуги по проживанию и питанию	15.4%	61.5%	23.1%
Информация и связь	6.1%	69.7%	24.2%
Операция с недвижимым имуществом	15.2%	50.0%	34.8%
Профессиональная и научно-техническая деятельность	22.0%	39.0%	39.0%
Прочие отрасли	14.3%	60.7%	25.0%
По экономике в целом	20.9%	48.9%	30.3%

Таблица 8

Региональное распределение предприятий по степени устойчивости

	Устойчивые	Нормальные	Критические
Акмолинская область	18.0%	46.8%	35.3%
Актюбинская область	20.0%	45.7%	34.3%
Алматинская область	26.4%	42.7%	30.9%
Атырауская область	20.6%	50.0%	29.4%
Восточно-Казахстанская область	19.9%	57.2%	22.9%
г. Алматы	17.0%	44.5%	38.5%
г. Астана	23.6%	46.5%	29.9%
Джамбульская область	24.1%	38.9%	37.0%
Западно-Казахстанская область	25.8%	45.2%	29.0%
Карагандинская область	23.0%	50.8%	26.2%
Костанайская область	26.0%	46.0%	28.0%
Кызылординская область	12.9%	58.3%	28.8%
Мангистауская область	27.0%	45.9%	27.0%
Павлодарская область	21.4%	44.1%	34.5%
Северо-Казахстанская область	17.4%	55.6%	27.1%
Южно-Казахстанская область	17.8%	56.8%	25.4%
По республике в целом	20.9%	48.9%	30.3%

Таблица 9

Распределение предприятий по размерностям

	Устойчивые	Нормальные	Критические
Крупные	19.3%	60.2%	20.5%
Средние	21.3%	47.3%	31.4%
Малые	21.8%	44.5%	35.6%
По республике в целом	20.9%	48.9%	30.3%

Прогноз на следующие 3 года

Таблица 10

Прогноз распределения предприятий по степени устойчивости в разрезе отраслей

	Устойчивые	Нормальные	Критические
Сельское хозяйство	23.1%	47.7%	29.2%
Горнодобывающая пром-сть	21.1%	49.6%	29.3%
Обрабатывающая пром-сть	25.5%	49.8%	24.6%
Электроснабжение	9.3%	68.0%	22.7%
Водоснабжение	12.0%	88.0%	0.0%
Строительство	20.2%	38.7%	41.2%
Торговля	20.8%	44.8%	34.3%
Транспорт и складирование	15.4%	52.2%	32.4%
Услуги по проживанию и питанию	15.4%	59.6%	25.0%
Информация и связь	6.1%	72.7%	21.2%
Операция с недвижимым имуществом	15.2%	51.1%	33.7%
Профессиональная и научно-техническая деятельность	20.0%	38.0%	42.0%
Прочие отрасли	14.3%	67.9%	17.9%
По экономике в целом	20.4%	48.9%	30.7%

Таблица 11

Прогноз распределения предприятий по размерностям

	Устойчивые	Нормальные	Критические
Крупные	19.1%	59.2%	21.7%
Средние	20.2%	47.3%	32.5%
Малые	21.3%	44.3%	34.4%
По республике в целом	20.4%	48.9%	30.7%

Прогноз регионального распределения предприятий по степени устойчивости

	Устойчивые	Нормальные	Критические
Акмолинская область	18.0%	47.5%	34.5%
Актюбинская область	18.1%	43.8%	38.1%
Алматинская область	24.5%	41.8%	33.6%
Атырауская область	20.6%	50.6%	28.8%
Восточно-Казахстанская область	18.7%	54.8%	26.5%
г. Алматы	15.9%	46.7%	37.4%
г. Астана	22.8%	47.2%	29.9%
Джамбульская область	24.1%	38.0%	38.0%
Западно-Казахстанская область	25.8%	46.8%	27.4%
Карагандинская область	23.0%	49.7%	27.3%
Костанайская область	25.0%	45.0%	30.0%
Кызылординская область	12.9%	55.3%	31.8%
Мангистауская область	24.3%	43.2%	32.4%
Павлодарская область	21.4%	49.7%	29.0%
Северо-Казахстанская область	18.1%	59.0%	22.9%
Южно-Казахстанская область	18.6%	53.4%	28.0%
По республике в целом	20.4%	48.9%	30.7%