

БЕЛОРУССКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ

УДК 519.2

Бодягин  
Игорь Александрович

**Статистическое прогнозирование  
авторегрессионных временных рядов  
при наличии цензурирования**

Автореферат диссертации на соискание ученой степени  
кандидата физико-математических наук

по специальности 01.01.05 — теория вероятностей и  
математическая статистика

Минск, 2012

Работа выполнена в Белорусском государственном университете.

Научный руководитель —

**Харин Юрий Семенович**

доктор физико-математических наук, профессор, член-корреспондент Национальной академии наук Беларуси, заведующий кафедрой математического моделирования и анализа данных Белорусского государственного университета.

Официальные оппоненты:

**Янович Леонид Александрович**

доктор физико-математических наук, профессор, член-корреспондент Национальной академии наук Беларуси, главный научный сотрудник отдела нелинейного и стохастического анализа ГНУ “Институт математики Национальной академии наук Беларуси”;

**Муха Владимир Степанович**

доктор технических наук, профессор, заведующий кафедрой информационных технологий автоматизированных систем УО “Белорусский государственный университет информатики и радиоэлектроники”.

Опонирующая организация—

Учреждение образования “Гродненский государственный университет имени Янки Купалы”.

Защита состоится 21 декабря 2012 г. в 12.00 на заседании совета по защите диссертаций Д 02.01.07 при Белорусском государственном университете по адресу: 220030, Республика Беларусь, г. Минск, ул. Ленинградская, 8 (корпус юридического факультета), ауд. 407, тел. ученого секретаря: (017) 209-57-09.

С диссертацией можно ознакомиться в Фундаментальной библиотеке Белорусского государственного университета.

Автореферат разослан “20” ноября 2012 года.

Ученый секретарь совета

по защите диссертаций Д 02.01.07,  
доктор физико-математических наук,  
профессор



Н. В. Лазакович

## КРАТКОЕ ВВЕДЕНИЕ

В настоящее время задачи статистического прогнозирования будущих значений временного ряда (случайного процесса с дискретным временем) по имеющимся настоящим и прошлым наблюдениям возникают во многих приложениях: в экономике, метеорологии, технике, медицине<sup>1</sup>. Для описания временных рядов с зависимыми наблюдениями и прогнозирования будущих значений широко используется модель гауссовского временного ряда и модель авторегрессии<sup>2</sup>. Одним из важных свойств авторегрессии является то, что любой стационарный в широком смысле временной ряд может быть представлен данной моделью конечного или бесконечного порядка. Это позволяет для любого стационарного в широком смысле временного ряда строить авторегрессионную модель конечного порядка с заданной точностью и затем использовать ее для статистического прогнозирования. При статистическом прогнозировании временных рядов возникают следующие задачи: 1) построение оптимальной прогнозирующей статистики в смысле минимума некоторого риска прогнозирования в случае известных параметров модели; 2) статистическое оценивание параметров модели в случае, когда они не известны, и построение подстановочной (“plug-in”) прогнозирующей статистики, которая строится путем подстановки в оптимальную прогнозирующую статистику вместо истинных значений параметров модели их статистических оценок<sup>3</sup>.

Статистические выводы (прогнозы, оценки) зависят не только от наблюдений  $X_0$ , но и от модельных априорных предположений<sup>4</sup>. На практике наблюдения часто регистрируются с различными искажениями их гипотетической модели: выбросами, пропусками, цензурированием<sup>5</sup>. При наличии искажений классические статистические выводы часто либо не применимы вовсе, либо становятся смещенными и несостоятельными. В связи с этим в математической статистике развивается теория робастности (от англ. robust), в рамках которой решаются задачи построения статистических оценок и прогнозирующих статистик, устойчивых к конкретным видам искажений гипотетической модели наблюдений.

В настоящей диссертации рассматривается такой вид искажений, как цен-

<sup>1</sup> Бокс, Дж. Анализ временных рядов. Прогноз и управление: в 2 т. / Дж. Бокс, Г. Дженкинс; пер. с англ. А. Л. Левшина; под ред. В. Ф. Писаренко. — М.: Мир, 1974. — 1 т. — 406 с.

<sup>2</sup> Андерсон, Т. Статистический анализ временных рядов / Т. Андерсон; пер. с англ. И. Г. Журбенко, В. П. Носко; под ред. Ю. К. Беляева. — М.: Мир, 1976. — 755 с.

<sup>3</sup> Chatfield, C. Time-series forecasting / C. Chatfield. — London: Chapman & Hall/CRC, 2000. — 265 p.

<sup>4</sup> Хьюбер, П. Робастность в статистике / П. Хьюбер; пер. с англ. И. А. Маховой, В. И. Хохловой; под ред. И. Г. Журбенко. — М.: Мир, 1984. — 304 с.

<sup>5</sup> Харин, Ю. С. Оптимальность и робастность в статистическом прогнозировании / Ю. С. Харин. — Мн.: БГУ, 2008. — 263 с.

зурирование, которое заключается в том, что лишь часть значений временного ряда наблюдается полностью, а об остальных значениях известно лишь, что они принадлежат некоторым числовым множествам, как правило, интервалам<sup>6</sup> конечной или бесконечной длины. Цензурированные данные встречаются при статистическом анализе в медицине, экономике, метеорологии, технике и многих других приложениях. Причинами возникновения цензурированных данных могут быть следующие<sup>7</sup>: наличие у измерительных приборов пределов измерений, специфика хранения данных, особенности проведения экспериментов и др.

Отметим, что цензурирование активно изучается в математической теории надежности для независимых выборок<sup>7</sup>, однако случай зависимых наблюдений — временных рядов — при наличии цензурирования остается малоизученным и, следовательно, актуальным направлением исследований, которому и посвящена настоящая диссертационная работа. При этом в литературе<sup>8</sup> отмечается, что цензурирование “сильно искажает” статистические свойства исходного временного ряда и значительно усложняет статистический анализ и прогнозирование.

Статистическое оценивание параметров и прогнозирование цензурированных временных рядов относится к робастной статистике, точнее к ее разделу: статистический анализ “неполных данных”. За рубежом исследованиями в данном направлении занимаются М. Genton, D. F. Heitjan, S. Zeger, P. K. Sen, N. Balakrishnan и др. В Беларуси исследованиями в области робастной статистики и, в частности, статистического анализа неполных данных занимаются Ю. С. Харин, Е. Е. Жук, В. С. Муха, М. С. Абрамович, В. И. Малюгин.

Указанные актуальные проблемы и определяют направление исследований настоящей диссертационной работы.

## ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

### Связь работы с крупными научными программами (проектами) и темами

Тема диссертации соответствует направлению фундаментальных и прикладных научных исследований “12.1. Физические и математические методы

<sup>6</sup> Klein, J. P. Survival analysis. Techniques for censored and truncated data / J. P. Klein, M. L. Moeschberger. — New York :Springer, 2003. — 536 p.

<sup>7</sup> Sen, P. K. Censoring in theory and practice: statistical perspectives and controversies / P. K. Sen // IMS lectures Notes. — 1995. — N 27. — P. 177–192.

<sup>8</sup> Zeger, S. L. Regression analysis with censored autocorrelated data / S. L. Zeger, R. Brookmeyer // Journal of the American statistical association. — 1986. — Vol. 81, No. 395. — P. 722–728.

и их применение для решения актуальных проблем естествознания, техники, новых технологий, экономики и социальных наук”, определенному Перечнем приоритетных направлений фундаментальных и прикладных научных исследований Республики Беларусь в области естественных, технических, гуманитарных и социальных наук на 2011–2015 годы. Результаты диссертационного исследования использованы при выполнении в Белорусском государственном университете следующих научно-исследовательских работ:

1) НИР “Разработка методов, алгоритмов и программных средств статистического прогнозирования процессов в сложных стохастических системах” в рамках ГКПНИ “Инфотех” (2006–2010 гг.), номер госрегистрации 20062250;

2) НИР “Разработка вероятностных моделей, методов и алгоритмов статистического анализа временных рядов с конечным пространством состояний и “длинной памятью” в рамках ГПФИ “Матмодели” (2006–2010 гг.), номер госрегистрации 20062206; на отдельные подзадания этой НИР диссертанту по конкурсу выделены гранты БГУ “Статистический анализ и прогнозирование марковских случайных последовательностей” (2009 г.), “Вероятностно-статистический анализ марковских случайных последовательностей” (2010 г.);

3) НИР “Статистический анализ данных сложной структуры и его применения в компьютерных системах моделирования и обработки информации” (2006–2010 гг.), номер госрегистрации 20062641;

4) НИР “Методы статистического анализа цепей Маркова с частичными связями при наличии искажений и их применение в системах защиты информации” договор с БРФФИ № Ф09М-042 (2009–2011 гг.), номер госрегистрации 20091347;

5) НИР “Статистическое прогнозирование авторегрессионных временных рядов при наличии цензурирования” в рамках гранта Министерства образования Республики Беларусь на 2011 г., номер госрегистрации 20112196;

6) НИР “Разработка методов робастного статистического анализа случайных последовательностей и полей при наличии неоднородностей и цензурирования” в рамках ГПНИ “Конвергенция” (2011–2015 гг.), номер госрегистрации 20111047; на отдельное подзадание этой НИР диссертанту по конкурсу выделен грант БГУ “Робастный статистический анализ марковских случайных последовательностей” (2011 г.);

7) НИР “Робастные статистические выводы и их применение в компьютерных системах моделирования и анализа данных” (2011–2015 гг.), номер госрегистрации 20120355; на отдельные подзадания этой НИР диссертанту по конкурсу выделен грант БГУ “Статистический анализ данных сложной структуры” (2012 г.).

## **Цель и задачи исследования**

Цель диссертационной работы: построить состоятельные статистические оценки параметров модели и прогнозирующие статистики для авторегрессионных временных рядов при наличии цензурирования и исследовать их вероятностные свойства. Для достижения поставленной цели требуется решить следующие основные задачи:

1) для стационарных авторегрессионных временных рядов при наличии цензурирования в случае известных параметров модели (полной априорной информации) построить оптимальную прогнозирующую статистику, минимизирующую условный среднеквадратический риск (среднеквадратическую ошибку) прогнозирования, и исследовать свойства риска оптимальной прогнозирующей статистики;

2) построить статистические оценки максимального правдоподобия параметров гауссовского временного ряда, и, как частный случай, гауссовского авторегрессионного временного ряда при наличии интервального цензурирования;

3) построить статистические оценки параметров авторегрессионного временного ряда при наличии цензурирования справа по методу моментов и исследовать их вероятностные свойства.

Объектом исследования являются гауссовские и авторегрессионные временные ряды при наличии цензурирования, часто используемые на практике при решении задач статистического прогнозирования. Предмет исследования — оптимальные прогнозирующие статистики в смысле минимума условного среднеквадратического риска (среднеквадратической ошибки) прогнозирования и статистические оценки параметров модели на основе методов моментов и максимального правдоподобия; на основе этих статистик и оценок удастся построить алгоритмы прогнозирования временных рядов, имеющие практическую ценность.

## **Положения, выносимые на защиту**

1. Построенная оптимальная прогнозирующая статистика, минимизирующая условный среднеквадратический риск прогнозирования для авторегрессионного временного ряда при наличии цензурирования, и асимптотическое разложение условного риска в случае одного интервально цензурированного значения при стремлении к нулю длины интервала цензурирования.

2. Найденное представление функции правдоподобия для гауссовского временного ряда при наличии конечного интервального цензурирования в виде сходящегося функционального ряда и основанный на этом представлении ал-

горитм построения приближений для оценки максимального правдоподобия.

3. Состоятельная статистическая оценка параметров модели авторегрессии при наличии цензурирования справа, основанная на методе моментов для вспомогательного временного ряда.

### **Личный вклад соискателя**

Основные результаты, приведенные в диссертации, получены автором самостоятельно. Научному руководителю в совместных работах принадлежит выбор направлений исследования, предметные постановки задач и формулировки общих подходов к их решению, обсуждение результатов.

### **Апробация результатов диссертации**

Основные результаты диссертации были представлены и обсуждались на заседаниях Республиканского научного семинара кафедры математического моделирования и анализа данных БГУ и Научно-исследовательского института прикладных проблем математики и информатики “Математическое моделирование сложных систем, анализ данных и защита информации”, научных конференциях студентов и аспирантов Белорусского государственного университета, на крупных Республиканских и Международных научных конференциях, в том числе на Международной конференции “International conference on robust statistics” (14–19 июня 2009 года, г. Парма, Италия), 6-ой Международной конференции “Математические методы в теории надежности. Теория. Методы. Приложения” (22–29 июня 2009 года, г. Москва), Международной конференции “International conference on robust statistics” (28 июня–2 июля 2010 года, г. Прага, Чехия), 9-ой Международной научной конференции “Computer data analysis and modeling: complex stochastic data and systems” (7–11 сентября 2010 года, г. Минск), Международной научной конференции студентов и молодых ученых “Theoretical and applied aspects of cybernetics” (21–25 февраля 2011 года, г. Киев, Украина), Международной конференции “International conference on robust statistics” (27 июня–1 июля 2011 года, г. Вальядолид, Испания), Международном научном семинаре “Applied methods of statistical analysis. Simulations and statistical inference” (20–22 сентября 2011 года, г. Новосибирск, Россия).

### **Опубликованность результатов диссертации**

Результаты диссертационной работы опубликованы в 23 научных работах. Из них 5 статей в научных журналах в соответствии с пунктом 18 Положения



о присуждении ученых степеней и присвоении ученых званий в Республике Беларусь (общим объемом 4 авторских листа), 12 статей в сборниках материалов научных конференций, 6 тезисов.

## Структура и объем диссертации

Диссертационная работа состоит из перечня условных обозначений, введения, общей характеристики работы, четырех глав, заключения, библиографического списка. Полный объем диссертации составляет 115 страниц, в том числе 7 рисунков на 4 страницах, 1 таблица на 1 странице. Библиографический список содержит 111 наименований, включая 23 собственные публикации соискателя, на 11 страницах.

## ОСНОВНОЕ СОДЕРЖАНИЕ

В главе 1 определены модели гауссовских и авторегрессионных временных рядов при наличии цензурирования. Рассматриваются их основные свойства и доказываются вспомогательные результаты, используемые в главах 2-4.

В разделе 1.2 вводятся в рассмотрение гауссовские и авторегрессионные временные ряды. Временной ряд  $x_t \in \mathbb{R}$ ,  $t \in \mathbb{Z}$ , заданный на вероятностном пространстве  $(\Omega, \mathcal{F}, \mathbb{P})$ , называется авторегрессионным порядка  $p \in \mathbb{N}$  (AR( $p$ )), если для него справедливо стохастическое разностное уравнение:

$$x_t = \sum_{i=1}^p \theta_i x_{t-i} + u_t, \quad t \in \mathbb{Z}, \quad (1)$$

где последовательность независимых в совокупности случайных величин  $\{u_t\}_{t=-\infty}^{+\infty}$  такова, что  $\mathbf{E}\{u_t\} = 0$ ,  $\mathbf{D}\{u_t\} = \sigma^2 < +\infty$ ;  $\theta_1, \dots, \theta_p \in \mathbb{R}$  — коэффициенты авторегрессии.

Определим дополнительные условия:

**У1)** Коэффициенты авторегрессии  $\theta_1, \dots, \theta_p$  таковы, что все корни  $z_1, \dots, z_p \in \mathbb{C}$  характеристического многочлена  $P(z) = z^p - \sum_{i=1}^p \theta_i z^{p-i}$  лежат внутри единичного круга ( $|z_i| < 1$ ,  $i = 1, \dots, p$ );

**У2)**  $\{u_t\}_{t=-\infty}^{+\infty}$  — независимые в совокупности, одинаково распределенные случайные величины, имеющие гауссовский (нормальный) закон распределения:

$$\mathcal{L}(u_t) = \mathcal{N}_1(0, \sigma^2), \quad \sigma^2 < +\infty.$$



В разделе 1.3 рассматривается общая постановка задачи статистического прогнозирования и подходы к ее решению, вводится понятие робастности.

В разделе 1.4 представлена модель цензурированных временных рядов. Пусть на вероятностном пространстве  $(\Omega, \mathcal{F}, \mathbf{P})$  определен некоторый временной ряд  $\mathbf{x}_t \in \mathbb{R}$ , наблюдаемый в дискретные моменты времени  $t = 1, 2, \dots, T$ ,  $T \in \mathbb{N}$  — количество наблюдений. Будем говорить, что в моменты времени  $1 \leq t_1 < t_2 < \dots < t_M \leq T$  временной ряд  $\mathbf{x}_t$  подвержен цензурированию (censoring), если в эти моменты времени вместо истинных значений  $\mathbf{x}_t$  наблюдаются лишь случайные события:

$$A_t^* = \{\mathbf{x}_t \in A_t\} \in \mathcal{F}, t \in T_C, \quad (2)$$

где  $A_t \in \mathcal{B}(\mathbb{R})$  — заданное борелевское множество,  $T_C = \{t_1, \dots, t_M\}$  — заданное множество моментов цензурирования,  $M \in \{0, 1, \dots, T\}$  — число цензурированных наблюдений. Случай  $M = 0$  соответствует классическому случаю отсутствия цензурирования. Обозначим:  $\Pi_C = A_{t_1} \times \dots \times A_{t_M}$ .

Таким образом, модель цензурирования задается множеством моментов цензурирования  $T_C$  и борелевскими множествами  $\{A_t : t \in T_C\} \subset \mathcal{B}(\mathbb{R})$ .

В разделе 1.5 проведен аналитический обзор литературы, посвященной задачам статистического анализа и прогнозирования в случаях неполных или искаженных данных: основное внимание уделяется анализу цензурированных выборок и временных рядов. Проведенный обзор обосновывает актуальность цели и задач диссертационного исследования.

**Глава 2** посвящена задаче построения оптимальной прогнозирующей статистики, минимизирующей среднеквадратический риск (среднеквадратическую ошибку) прогнозирования, в случае известных параметров модели авторегрессии (полной априорной информации) при наличии цензурирования и исследованию ее некоторых вероятностных свойств.

В разделе 2.1 определяется одношаговая прогнозирующая статистика, условный и усредненный риски прогнозирования. Рассматривается авторегрессионный временной ряд  $\mathbf{x}_t$ , задаваемый (1), при наличии цензурирования (2). Параметры модели авторегрессии  $(p, \theta_1, \dots, \theta_p, \sigma^2)$  предполагаются известными. Результатом наблюдения является случайный вектор  $T - M$  зарегистрированных значений временного ряда  $\mathbf{x}_t$ :

$$X_O = (\mathbf{x}_t)_{t \in T_O}' \in \mathbb{R}^{T-M}, T_O = \{1, \dots, T\} \setminus T_C, \quad (3)$$

и случайное событие:

$$A_C^* = \bigcap_{t \in T_C} A_t^*, \quad (4)$$

означающее, что остальные  $M$  значений временного ряда цензурированы.

# СПИСОК ПУБЛИКАЦИЙ СОИСКАТЕЛЯ ПО ТЕМЕ ДИССЕРТАЦИИ

## Статьи в научных журналах

1. *Гурип, А. С.* Статистический анализ векторных авторегрессионных временных рядов с гетероскедастичностью / А. С. Гурип, И. А. Бодягин, А. В. Федорук // Вестн. Белорус. гос. ун-та. Сер. 1, Физ., матем., информат. — 2009. — № 3. — С. 94–99.

2. *Бодягин, И. А.* Прогнозирование авторегрессионных временных рядов при наличии цензурирования / И. А. Бодягин, Ю. С. Харин // Вестн. сибирского гос. аэрокосмического ун-та. — 2010. — № 5(31). — С. 22–25.

3. *Харин, Ю. С.* Об оптимальном прогнозировании авторегрессионных временных рядов при наличии интервального цензурирования / Ю. С. Харин, И. А. Бодягин // Вес. нац. акад. навук. Беларусі. Сер. фіз.-мат. навук. — 2011. — № 1. — С. 60–69.

4. *Бодягин, И. А.* Функция правдоподобия для цензурированных гауссовских временных рядов / И. А. Бодягин, Ю. С. Харин // Весці нац. акад. навук. Беларусі. Сер. фіз.-мат. навук. — 2012. — № 2. — С. 4–14.

5. *Бодягин, И. А.* Оценивание параметров цензурированного авторегрессионного временного ряда по методу моментов / И. А. Бодягин, Ю. С. Харин // Вестн. Белорус. гос. ун-та. Сер. 1, Физ., матем., информат. — 2012. — № 3. — С. 60–66.

## Статьи в сборниках научных трудов

6. *Бодягин, И. А.* Статистический анализ авторегрессионных временных рядов при наличии цензурирования / И. А. Бодягин, Ю. С. Харин // НИРС-2009: сб. науч. тр. студентов высших учебных заведений Республики Беларусь / Белорус. гос. ун-т; редкол.: А. И. Жук [и др.]. — Минск, 2010. — С. 14–15.

## Статьи в сборниках материалов научных конференций

7. *Бодягин, И. А.* О статистическом прогнозировании авторегрессионных временных рядов при наличии неполных наблюдений / И. А. Бодягин, Ю. С. Харин // Информационные системы и технологии: материалы Междунар. конф., Минск, 4–6 ноября 2008 г. / Белорус. гос. ун-т; редкол.: А. Н. Курбацкий [и др.]. — Минск, 2008. — С. 24–29.

8. *Kharin, Yu. S.* On robust forecasting of autoregressive time series under censoring / Yu. S. Kharin, I. A. Badziahin // Pattern recognition and information processing: Proc. of the 10th Intern. conf., Minsk, 19–21 May, 2009. / Belorussian State Univ.; editors: V. Krasnoproshin, S. Ablameyko, R. Sadykhov. — Minsk, 2009. — P. 73–77.

9. *Бодягин, И. А.* Об оптимальном прогнозировании авторегрессионных временных рядов при наличии интервального цензурирования / И. А. Бодягин // Информационные системы и технологии: материалы V Междунар. конф.-форума, Минск, 16–17 ноября 2009 г.: в 2 ч. / Белорус. гос. ун-т; редкол.: Н. И. Листопад [и др.] — Минск, 2009. — Ч. 1. — С. 106–109.

10. *Бодягин, И. А.* Об оценках максимального правдоподобия параметров  $AR(1)$ -временных рядов при наличии цензурирования / И. А. Бодягин // 67-я научно-практическая конференция студентов и аспирантов Белорус. гос. ун-та: сб. работ, Минск, 17–20 мая 2010 г.: в 3 ч. / Белорус. гос. ун-т; редкол.: С. А. Мазаник [и др.] — Минск, 2011 — Ч. 1 — С. 7–10.

11. *Петлицкий, А. И.* Искаженные цепи маркова с частичными связями / А. И. Петлицкий, И. А. Бодягин // Теория вероятностей, математическая статистика и их приложения: Материалы Междунар. науч. конф., Минск, 22–25 февраля 2010 г. / РИВШ; редкол.: Н. Н. Труш [и др.] — Минск, 2010. — С. 278–283.

12. *Бодягин, И. А.* Об оценках максимального правдоподобия параметров авторегрессии при наличии цензурирования справа / И. А. Бодягин // Белорусская статистика: вчера, сегодня, завтра: Материалы науч.-практич. конф., Минск, 23–24 августа 2010 г. / Нац. статистический комитет Республики Беларусь; редкол.: И. А. Костевич [и др.]. — Минск, 2010. — С. 308–309.

13. *Badziahin, I. A.* On maximum likelihood estimation of parameters for censored autoregressive time series / I. A. Badziahin, Yu. S. Kharin // Computer data analysis and modeling: complex stochastic data and systems: proc. of the ninth Intern. conf., Minsk, Sept. 7–11, 2010.: in 2 vol. / Belorussian State University; editors: S. Aivazian [et. al]. — Minsk, 2010. — Vol. 1 — P. 122–125.

14. *Бодягин, И. А.* Об оценках максимального правдоподобия параметров авторегрессии при наличии интервального цензурирования / И. А. Бодягин // Информационные системы и технологии: материалы VI Междунар. науч. конф., Минск, 24–25 ноября 2010 г. / Белорус. гос. ун-т; редкол.: А. Н. Курбацкий [и др.] — Минск, 2010. — С. 125–128.

15. *Бодягин, И. А.* Об оценках максимального правдоподобия параметров авторегрессии в случае интервального цензурирования / И. А. Бодягин // Theoretical and applied aspects of cybernetics: proc. of the Intern. scient. conf. of students and young scientists, Kyiv, Febr. 21–25, 2011. / National University

of Kyiv; editors: I. O. Lytvinenko, D. O. Terletskyi. — Kyiv, 2011 — P. 263–267.

16. *Kharin, Yu. S.* Statistical forecasting for censored autoregressive time series / Yu. S. Kharin, I. A. Badziahin // Applied methods of statistical analysis. Simulations and statistical inference: proc. of the Intern. Workshop, Novosibirsk, September 20–22, 2011. / Nation. state technical university; editors.: B. Lemeshko, M. Nikulin, N. Balakrishnan. — Novosibirsk, 2011. — P. 310–317.

17. *Бодягин, И. А.* О прогнозировании цензурированных временных рядов / И. А. Бодягин // Международный конгресс по информатике: информационные системы и технологии: материалы Междунар. науч. конгресса, Минск, 31 октября – 3 ноября 2011 г.: в 2 ч. / Белорус. гос. ун-т; редкол.: С. В. Абламейко [и др.]. — Минск, 2011. — Ч. 1 — С. 34–39.

### Тезисы докладов

18. *Гурин, А. С.* Статистический анализ векторных авторегрессионных временных рядов с гетероскедастичностью / А. С. Гурин, И. А. Бодягин, А. В. Федорук // X Белорусская математическая конференция: тезисы докл. Междунар. науч. конф., Минск, 3–7 ноября 2008 г. / Белорус. гос. ун-т; редкол.: С. Г. Красовский, А. А. Ленин. — Минск, 2008. — Ч. 5. — С. 40–41.

19. *Badziahin, I. A.* On robustness of forecasting for AR(p) time series under interval censoring / I. A. Badziahin, Yu. S. Kharin // International conference on robust statistics: abstr. of the Intern. conf., Parma, Italy, 14–19 June, 2009. / Universita degli Studi di Parma; editors: C. Agostinelli [et. al.]. — Parma, 2009. — P. 9.

20. *Харин, Ю. С.* Об оптимальном прогнозировании авторегрессионных временных рядов с интервальным цензурированием / Ю. С. Харин, И. А. Бодягин // Математические методы в теории надежности: Расширенные тезисы докл. VI Междунар. конф., Москва, 22–29 июня, 2009 г. / Российский ун-т нефти и газа; редкол.: Н. А. Кузнецов, В. В. Рыков. — Москва, 2009. — С. 415–419.

21. *Badziahin, I. A.* On estimation of parameters for autoregressive time series under left censoring / I. A. Badziahin, Yu. S. Kharin // International conference on robust statistics: abstr. of the Intern. conf. Prague, Czech Republic, 28 June – 2 July 2010. — Prague, 2010. — P. 9–10.

22. *Бодягин, И. А.* О функции правдоподобия для гауссовских временных рядов при наличии интервального цензурирования / И. А. Бодягин // НИРС-2011: сб. тезисов докладов Респ. науч. конф. студентов и аспирантов высших учебных заведений Республики Беларусь, Минск, 18 октября, 2011 г. / Белорус. гос. ун-т; редкол.: С. В. Абламейко [и др.] — Минск, 2011. — С. 72.

23. *Badziahin, I. A.* Likelihood function for Gaussian time series under interval censoring / I. A. Badziahin, Yu. S. Kharin // International conference on robust statistics: book of abstr. of the Intern. conf., Valladolid, Spain, June 27 – July 1, 2011. / Universidad de Valladolid. — Valladolid, 2011 — P. 35.

# РЭЗІЮМЭ

Бадзягін Ігар Аляксандравіч

Статыстычнае прагназаванне аўтарэгрэсійных часовых шэрагаў  
пры наяўнасці цэнзурыравання

*Ключавыя словы:* аўтарэгрэсійны часовы шэраг, цэнзурыраванне, сярэднеквадратычная рызыка прагназавання, функцыя праўдападобнасці, метады момантаў.

Дыссертацыйная праца мае сваей мэтай пабудаванне слухных статыстычных ацэнак параметраў мадэлі і прагназуючых статыстык для аўтарэгрэсійных часовых шэрагаў пры наяўнасці цэнзурыравання і даследаванні іх імавернасных ўласцівасцяў. Пры даследаванні выкарыстоўваліся метады матэматычнага аналізу, тэорыі імавернасцей, матэматычнай статыстыкі, камп'ютэрнага мадэлявання.

У дыссертацыйнай працы атрыманы наступныя новыя навуковыя рэзультаты. Для аўтарэгрэсійных часовых шэрагаў пры наяўнасці цэнзурыравання пабудаваны аптымальныя прагназуючыя статыстыкі, мінімізуючыя ўмоўную сярэднеквадратычную рызыку прагназавання, і знойдзены ўмоўныя рызыкі прагназуючых статыстык. Пабудаваны асімптатычны расклад ўмоўнай рызыкі аптымальнай прагназуючай статыстыкі для мадэлі аўтарэгрэсіі першага парадку з адзіным інтэрвальна цэнзурыраваным значэннем пры імгненні да нуля даўжыні інтэрвалу цэнзурыравання. Пабудаваны расклад у шэраг функцыі праўдападобнасці для гаўссаўскіх часовых шэрагаў пры наяўнасці фінітнага інтэрвальнага цэнзурыравання. Распрацаваны алгарытм пабудавання прыблізных ацэнак максімальнай праўдападобнасці на аснове знойдзенага раскладу функцыі праўдападобнасці. Пабудаваны статыстычныя ацэнкі параметраў мадэлі аўтарэгрэсіі пры наяўнасці цэнзурыравання справа, заснаваныя на метады момантаў для асобага дапаможнага часовага шэрагу. Даказана слухнасць пабудаваных ацэнак пры выкананні некаторых умоў на чацвёртых моманты дапаможная часовага шэрагу.

Атрыманыя тэарытычныя рэзультаты могуць быць выкарыстаны пры статыстычным аналізу і прагназаванні рэальных дадзенных пры наяўнасці цэнзурыравання ў эканоміцы, фінансах, тэхніке, медыцыне, а таксама ў навуковым працэсе пры чытанні спецыяльных курсаў па матэматычнай статыстыцы на матэматычных факультэтах універсітэтаў.

# РЕЗЮМЕ

Бодягин Игорь Александрович

Статистическое прогнозирование авторегрессионных временных рядов при наличии цензурирования

*Ключевые слова:* авторегрессионный временной ряд, цензурирование, среднеквадратический риск прогнозирования, функция правдоподобия, метод моментов

Диссертационная работа имеет целью построение состоятельных статистических оценок параметров модели и прогнозирующих статистик для авторегрессионных временных рядов при наличии цензурирования и исследование их вероятностных свойств. При исследовании использовались методы математического анализа, теории вероятностей, математической статистики, компьютерного моделирования.

В диссертационной работе получены следующие новые научные результаты. Для авторегрессионных временных рядов при наличии цензурирования построены оптимальные прогнозирующие статистики, минимизирующие условный среднеквадратический риск прогнозирования, и найдены условные риски оптимальных прогнозирующих статистик. Построено асимптотическое разложение условного риска оптимальной прогнозирующей статистики для модели авторегрессии первого порядка с единственным интервально цензурированным значением при стремлении к нулю длины интервала цензурирования. Построено разложение в ряд функции правдоподобия для гауссовского временного ряда при наличии конечного интервального цензурирования. Разработан алгоритм построения приближенной оценки максимального правдоподобия на основе построенного разложения функции правдоподобия. Построены статистические оценки параметров модели авторегрессии при наличии цензурирования справа, основанные на методе моментов для специального вспомогательного временного ряда. Доказана состоятельность построенных оценок при выполнении некоторых дополнительных условий на четвертые моменты вспомогательного временного ряда.

Полученные теоретические результаты могут быть использованы при статистическом анализе и прогнозировании реальных данных при наличии цензурирования в экономике, финансах, технике, медицине, а также в учебном процессе при чтении специальных курсов по математической статистике на математических факультетах университетов.



# SUMMARY

Badziahin Ihar Alexandrovich

Forecasting of autoregressive time series observed under censoring

*Key words:* autoregressive time series, censoring, mean-square forecasting risk, likelihood function, method of moments.

The dissertation has the following objectives: to construct new statistical estimators of model parameters and forecasting statistics for autoregressive time series observed under censoring; to analyze their probabilistic properties. *The techniques used include* mathematical analysis, probability theory, methods of mathematical statistics and computer simulation.

In this dissertation the following new scientific results are obtained. The optimal forecasting statistics that minimize the conditional mean-square risk for autoregressive time series observed under censoring are constructed. Values of conditional risks for optimal forecasting statistics are found. Asymptotic expansion of optimal forecasting statistic conditional risk for first order autoregression is found. Representation of the likelihood function for gaussian time series observed under interval censoring in the form of functional expansion is found. Algorithm of approximate maximum likelihood estimators construction is proposed. Statistical estimators of autoregression model parameters are constructed by using the method of moments for special auxiliary time series. Consistency of constructed estimators are proved under some additional general conditions.

The obtained theoretical results can be used in statistical analysis and forecasting of real data observed under censoring in economics, finance, technique, medicine. Also they can be used in teaching process by reading special courses of statistical analysis in universities.

