

Секция «Вычислительная математика, математическое моделирование и численные методы»

## РЕКОНСТРУКЦИЯ ДЛИННЫХ РЯДОВ ДЕМОГРАФИЧЕСКИХ ДАННЫХ

*Темрук Александр Дмитриевич*

*Студент (магистр)*

Московский физико-технический институт, Москва, Россия

*E-mail: atem@mail.ru*

1. Постановка проблемы. Предметом рассмотрения являются показатели, представленные в виде временных рядов различной длины. Решается задача по реконструкции значений некоторых статистических характеристик полной выборки на основе заданных частичных выборок. В качестве таких характеристик были выбраны *квантили* рассматриваемой величины, т.е. такие ее значения, которые эта величина превышает на определенной доле рассматриваемых объектов. Работа выполнялась при поддержке РФФИ (проект 15-01-07944-а).

2. Актуальность проблемы. С подобными задачами часто сталкиваются при анализе демографических данных, регистрация которых начиналась в разных странах в различное время. Причем чем более развитой была страна, тем, как правило, раньше в ней начинался сбор демографической статистики. Поэтому выборка стран, для которых доступны данные, не может служить неизменным эталоном при описании ситуации в разные годы. Действительно, чем дальше мы будем уходить в прошлое, тем сильнее будем ограничивать анализ лишь сравнительно благополучными странами.

Реконструкция квантилей какого-либо демографического коэффициента на полной выборке стран позволяет определить место любой страны в мире относительно широкого и неизменного набора стран на всем промежутке времени, за который для этой страны представлены значения коэффициента. Это позволяет оценить динамику выбранного демографического коэффициента в этой стране относительно других стран, вошедших в выборку, без учета общего тренда.

Очевидно, что проблема реконструкции временных рядов может стоять и при анализе социально-экономических показателей стран и групп стран.

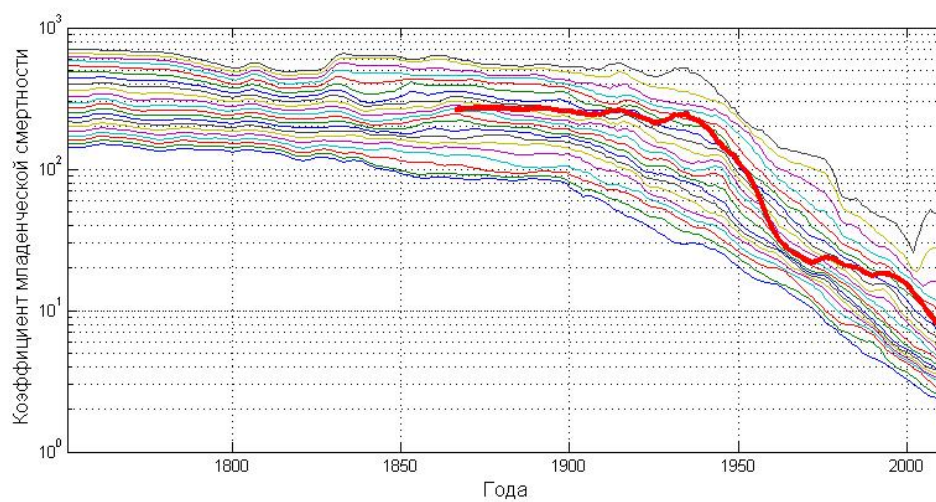
3. Математические методы и инструмент. Основная часть данных была взята из баз данных ресурса [1]. Материалы для России были дополнены значениями из работ [2], [3]. Исходные данные проходили стадию пре-процессинга: восстановление пропущенных значений при помощи интерполяции. Проводилась сортировка стран в порядке уменьшения года начала сбора данных. Содержание первого этапа: построение квантильных уровней для всех стран на всем временном промежутке при помощи регрессионной модели. Содержание второго этапа: представление значений показателя каждой страны в пространстве этих квантилей. Разработанный алгоритм был реализован на платформе MatLab.

4. Результаты экспериментов. В качестве первоначальных данных были использованы значения следующих демографических показателей по 49 странам: коэффициент младенческой смертности, общий коэффициент смертности, общий коэффициент рождаемости, коэффициент суммарной рождаемости. Реконструированная динамика процентилей позволила описать общие закономерности изменения рассматриваемых величин и эволюцию относительного положения избранной страны относительно неизменной выборки других стран.

**Источники и литература**

- 1) Демоскоп Weekly. Приложения. <http://demoscope.ru/weekly/pril.php>
- 2) Андреев Е.М., Дарский Л.Е., Харьковская Т.Л. Демографическая история России: 1927–1957. – М.: Информатика, 1998. – 187 с.
- 3) Демографическая модернизация России. 1900–2000/ Под ред. А.Г.Вишневого. – М.: Новое издательство, 2006. – 608 с.

### Иллюстрации



**Рис. 1.** Графики реконструированных перцентилей коэффициента младенческой смертности и график его значений для России (красная линия).