



Курс «Основы статистического анализа в демографии»

4–15 февраля 2013г.

Российская экономическая школа

Преподаватель

[Трифон И. Мисов](#) (Max Planck Institute for Demographic Research, Germany; [Universität Rostock, Germany](#))

Краткая информация о курсе

Демографические данные позволяют определять вероятность наступления и интенсивность демографических событий, а также их распределение во времени. Курс посвящен методам их статистического анализа. В нем будут представлены методы как классической, так и современной математической статистики, которые будут реализованы на стандартном для демографии языке программирования R, к программной среде которого имеется свободный доступ. Теоретические результаты будут проиллюстрированы на основе многочисленных демографических примеров.

Курс дает возможность

- ознакомиться с классическими и современными методами статистического анализа демографических данных
- получить практический опыт применения этих методов на широко используемом в демографических исследованиях языке программирования R для решения конкретных научно-практических задач
- научиться интерпретировать полученные количественные результаты

Курс рассчитан на преподавателей, научных сотрудников, аспирантов, студентов старших курсов и магистратуры по демографии и смежным дисциплинам. Основные знания математики, математической статистики и демографии желательны, но не обязательны, поскольку в рамках курса все необходимые понятия будут четко введены. Знакомство с языком программирования R, на котором будут реализованы алгоритмы анализа данных, тоже желательно, но не обязательно.

Содержание курса

Курс имеет целью представить наиболее корректные и практически полезные методы для статистического анализа демографических данных на доступном количественном и качественном уровне и освоить основы решения статистических задач на языке R.

Как статистически осмысленно сравнить режимы смертности разных лет? Как исследовать динамику бездетности? Как оценить вероятность зачатия для курящих и некурящих женщин? Как оценить интенсивность заболевания раком в зависимости от ряда факторов? Насколько вероятно заключить брак в зависимости от возраста? Ответы на эти и другие важные вопросы могут быть получены с помощью корректного применения статистических методов, которым посвящен курс „Основы статистического анализа в демографии“.

Курс состоит из трех частей. Первая из них служит введением в тематику: каковы бывают демографические данные, как проводить их предварительный (описательный) анализ, что необходимо оценивать.

Вторая часть посвящена методам классической математической статистики, с помощью которой оцениваются вероятность наступления и интенсивность демографических событий. Основной акцент ставится на разъяснение применения метода максимального правдоподобия (ММП).

Третья часть курса является введением в анализ выживаемости (Survival Analysis), с помощью которого оценивается наступление демографических событий во времени. Сначала будет пояснена особенность данных типа времени жизни и то, почему для их анализа нужно применять методы неклассической статистики. Затем будут представлены методы параметрического и непараметрического оценивания дожития и их практическое применение. В конце будет изучено статистическое оценивание с учетом ненаблюдаемой неоднородности и его применение в области анализа смертности.

Основная литература

1. Agresti A., and B. Finlay (2009). *Statistical Methods for the Social Sciences*. 4th edition. Pearson Prentice Hall.
2. Rodriguez, G.. *Lecture Notes*: <http://data.princeton.edu/wws509/notes/>
3. Klabfleisch, J.D., and R.L. Prentice (1980). *The Statistical Analysis of Failure Time Data*. New York: Wiley.

Дополнительные материалы и текст используемого компьютерного кода будут выданы участникам во время проведения курса.

Программа курса «Основы статистического анализа в демографии»

I. Введение: демографические данные и общая постановка их статистического анализа

II. Статистический анализ вероятности наступления и интенсивности демографических событий

1. Дескриптивная статистика в R: предварительный анализ данных
2. Статистические модели как механизмы генерирования данных. Применение теории вероятностей в задачах демографии и эпидемиологии
3. Параметрические статистические модели: оценивание методом максимального правдоподобия
4. Примеры применения метода максимального правдоподобия в демографических моделях:
 - a) возрастные коэффициенты смертности: биномиальная модель
 - b) плодовитость: геометрическая модель
 - c) число смертей: пуассоновская модель
5. Регрессионные параметрические статистические модели :
 - a) использование противозачаточных средств: логистическая регрессия
 - b) заболеваемость раком: пуассоновская регрессия

III. Введение в статистический анализ данных типа времени жизни (анализ выживаемости)

1. Особенности данных типа времени жизни: цензурирование
2. Примеры параметрических моделей:
 - a) модель Вейбулля
 - b) человеческая смертность: модель Гомперца
3. Непараметрические модели: оценка Каплана-Майера и регрессия Кокса.
4. Ненаблюдаемая неоднородность в данных по смертности: модели уязвимости

Практическая информация

Курс проводится на базе РЭШ по адресу г. Москва, Нахимовский проспект, д. 47. Занятия будут проходить в течение 10 рабочих дней с **4 по 15 февраля 2013 г.** с 10.00 до 13.30. С 15:00 до 16:00 – консультации по сложным вопросам курса (при необходимости). Занятия будут проводиться на русском языке, но литература для самостоятельного изучения преимущественно на английском. Поэтому знание английского языка на базовом уровне является обязательным, хороший уровень английского будет давать преимущества при зачислении на курс.

Теоретические лекции будут сопровождаться практическими занятиями. Кроме того, программа курса подразумевает много самостоятельной работы. Для успешного окончания курса участникам курса будет предложено либо представить решения задач, которые будут выдаваться в конце каждой темы, либо подготовить минипроект в том случае, когда имеется конкретный набор данных для анализа, на которых можно продемонстрировать изучаемые методы.

По окончании курса будет выдан сертификат международного образца.

Для работы на практических занятиях участники курса должны иметь персональные компьютеры (ноутбуки).

Заявки на участие в курсе принимаются на сайте [РЭШ](#) до **5 декабря 2012 года** включительно. Заявки на участие, присланные позже 1 декабря, не рассматриваются. На курс будут отобраны 15 человек. О зачислении на курс участники будут извещены в срок до **14 декабря 2012 года**.

Общежитие

Места в общежитии иногородним слушателям предоставляются. Иногородние участники самостоятельно оплачивают проживание в общежитии и проезд.

По всем вопросам просим обращаться к организаторам:

Елена Чурилова echurilova@nes.ru

Дмитрий Жданов Jdanov@demogr.mpg.de