



Правительство Российской Федерации

Федеральное государственное автономное образовательное учреждение
высшего профессионального образования
«Национальный исследовательский университет
«Высшая школа экономики»

Факультет прикладной политологии

Программа дисциплины

Анализ политологических данных в пакете Stata

для направления 030200.62 «Политология» подготовки бакалавра

Автор программы:

Камалова Рита Ульфатовна, преподаватель (rkamalova@hse.ru)

Одобрена на заседании кафедры высшей математики «__» _____ 2012 г.
Зав. кафедрой А.А. Макаров

Рекомендована секцией УМС «_____» «__» _____ 2012 г.
Председатель _____

Утверждена УС факультета прикладной политологии «__» _____ 2012 г.
Ученый секретарь _____

Москва, 2012

Настоящая программа не может быть использована другими подразделениями университета и другими вузами без разрешения кафедры-разработчика программы.



1 Область применения и нормативные ссылки

Настоящая программа учебной дисциплины устанавливает минимальные требования к знаниям и умениям студента и определяет содержание и виды учебных занятий и отчетности.

Программа предназначена для преподавателей, ведущих данную дисциплину, учебных ассистентов и студентов направления 030200.62 «Политология» подготовки бакалавра, изучающих дисциплину «Анализ политологических данных в пакете Stata».

Программа разработана в соответствии с рабочим учебным планом университета по направлению 030200.62 «Политология» подготовки бакалавра, утвержденным в 2012 г.

2 Цели освоения дисциплины

Целью освоения дисциплины «Анализ политологических данных в пакете Stata» является выработка компетенций по анализу данных политологических и социально-экономических данных (преимущественно методами регрессионного анализа) в пакете Stata.

3 Компетенции обучающегося, формируемые в результате освоения дисциплины

В результате освоения дисциплины студент должен:

- Знать базовые методы анализа данных, в том числе регрессии, позволяющие решать некоторые задачи социальных наук;
- Иметь навыки их реализации в пакете Stata;
- Уметь интерпретировать и представлять полученные результаты;
- Знать ряд известных источников и баз политологических и социально-экономических данных.

4 Место дисциплины в структуре образовательной программы

Настоящая дисциплина относится к факультативам направления 030200.62 «Политология» подготовки бакалавра. Для указанного направления дисциплина является дисциплиной по выбору.

Изучение данной дисциплины базируется на следующих дисциплинах:

- «Алгебра и анализ» (1 курс),
- «Начала теории вероятностей и математической статистики» (1 курс).

Основные положения дисциплины должны быть использованы в дальнейшем при изучении следующих дисциплин:

- «Многомерный статистический анализ в политологии»,
- «Сравнительная политика»,
- «Политический анализ»,
- «Политическое поведение»,
- «Анализ электоральных данных»,
- «Прогнозирование социально-политических процессов»,
- «Многомерные типологии политических систем современных государств».



5 Тематический план учебной дисциплины

№	Название раздела	Всего часов	Аудиторные часы		Самостоятельная работа
			Лекции	Практические занятия	
1	Обзор источников политологических и социально-экономических данных и средств их анализа. Пакет Stata.	4	1	1	2
2	Описательные статистики, графики и рисунки в Stata.	8	2	2	4
3	Связь качественных признаков в политологии. Связь признаков в количественных и порядковых шкалах	6	1	1	4
4	Регрессионный анализ. Модель парной линейной регрессии	20	6	6	8
5	Регрессионный анализ. Модель множественной линейной регрессии. Разложение вариации	18	4	4	10
6	Мультиколлинеарность. Автокорреляция	14	4	4	6
7	Гетероскедастичность. Влиятельные наблюдения	14	4	4	6
8	Эффекты взаимодействия переменных и фиктивные переменные	12	4	4	4
9	Нелинейные взаимосвязи в регрессии.	12	4	4	4
	Итого	108	30	30	48

6 Формы контроля знаний студентов

Тип контроля	Форма контроля	1 год				Параметры
		1	2	3	4	
Текущий (неделя)	Контрольная работа			*		Письменная работа, 80 минут.
	Домашнее задание				*	Самостоятельная письменная работа по итогам анализа массива политологических и социально-экономических данных с использованием изученных методов. Объем: 8-15 тыс. символов с пробелами.
Итоговый	Зачет				*	Письменный зачет, 100 минут.

6.1 Критерии оценки знаний, навыков

Оценки за домашние задания и зачет выставляются, исходя из следующих критериев:

- правильность решения задачи,
- полнота решения задачи,
- наличие и корректность интерпретации полученных результатов.

Оценки по всем формам текущего и итогового контроля выставляются по 10-ти балльной шкале.

Письменный зачет состоит из 2 частей: **А** и **В**. В части **А** будут представлены **двенадцать** задач разного уровня сложности. Вам нужно решить **по одной задаче в каждом тематическом блоке**. Если будут сделаны попытки решить больше необходимого числа задач из этой части, *проверены будут те, которые следуют в работе раньше*.



В части **В** будут представлены **три** задачи. Задачи части **В** нужно решать, только если в работе представлены решения **любых семи задач** части **А**.

7 Содержание дисциплины

Раздел 1. Обзор источников политологических и социально-экономических данных и средств их анализа. Пакет Stata.

Основная литература

1. Long, J. Scott and Jeremy Freese. 2005. *Regression Models for Categorical Dependent Variables Using Stata*. 2nd Edition. College Station, TX: Stata Press. Ch. 1, 2.
2. A brief description of Stata. (http://www.stata-press.com/manuals/stata10/guide_brief.pdf).
3. Basics of Stata. (2008). (<http://www.ics.uci.edu/~dvd/Courses/Stata/stata.pdf>).
4. Учебно-методические материалы для преподавания курсов по применению количественного инструментария к решению социально-экономических задач (<http://www.hse.ru/jesda/mathbase/>).

Дополнительная литература

5. Единый архив социологических данных (рук. — Л. Б. Косова) // Экономическая социология (электронный журнал), 2010, т. 11, № 1, с. 133–140. (http://www.hse.ru/data/2011/12/08/1208204985/ecsoc_t11_n1.pdf#page=133)
6. Вакуленко Е.С., Камалова Р.У., Кисельгоф С.Г., Смирнова Ж.И., Стукал Д.К., Хавенсон Т.Е. Разработка учебно-методических материалов для преподавания курсов по применению количественного инструментария к решению социально-экономических задач. // Экономическая социология (электронный журнал), 2010. Т. 11. № 4. С. 190—194. (http://ecsoc.hse.ru/data/2010/09/29/1234590525/ecsoc_t11_n4.pdf#page=190)
7. Long, J. Scott. 2009. The workflow of data analysis using Stata. TX: Stata Press. (Полочный индекс в библиотеке НИУ ВШЭ – 004 L87).
8. Resources to help you learn and use Stata. UCLA: Academic Technology Services, Statistical Consulting Group. (<http://www.ats.ucla.edu/stat/stata/>).
9. Rodríguez, Germán. (<http://data.princeton.edu/stata/>).

Раздел 2. Описательные статистики в Stata. Графики и рисунки в Stata.

Основная литература

1. Тюрин Ю.Н., Макаров А.А. Анализ данных на компьютере: учебное пособие. – М.: ИД «ФОРУМ», 2008. – С. 24 – 49, 131 – 138.
2. Descriptive information and statistics. UCLA: Academic Technology Services, Statistical Consulting Group. (<http://www.ats.ucla.edu/stat/stata/modules/descript.htm>).
3. Graphics. UCLA: Academic Technology Services, Statistical Consulting Group. (<http://www.ats.ucla.edu/stat/stata/topics/graphics.htm>).
4. Long, J. Scott and Jeremy Freese. 2005. *Regression Models for Categorical Dependent Variables Using Stata*. 2nd Edition. College Station, TX: Stata Press. Ch. 1, 2.

Дополнительная литература

5. Data Management. UCLA: Academic Technology Services, Statistical Consulting Group. (http://www.ats.ucla.edu/stat/stata/topics/data_management.htm).
6. Long, J. Scott. 2009. The workflow of data analysis using Stata. TX: Stata Press. (Полочный индекс в библиотеке НИУ ВШЭ – 004 L87).
7. Крыштановский А.О. Анализ социологических данных с помощью пакета SPSS : учеб. пособие для вузов. – М.: Изд. дом ГУ ВШЭ, 2006. – С. 10-39.
8. Тюрин Ю.Н., Макаров А.А., Высоцкий И.Р., Яценко И.В. Теория вероятностей и статистика. — М.: МЦНМО, 2004. – С. 27-44, 44-63, 189-205.



Раздел 3. Связь качественных признаков в политологии. Связь признаков в количественных и порядковых шкалах.

Основная литература

1. Доугерти К. Введение в эконометрику: Учебник. М.: ИНФРА-М, 1999. Гл. 1. С. 34–52.
2. Тюрин Ю.Н., Макаров А.А. Анализ данных на компьютере: учебное пособие. – М.: ИД «ФОРУМ», 2008. – С. 244-263.
3. Stata Annotated Output. Correlation. UCLA: Academic Technology Services, Statistical Consulting Group. (http://www.ats.ucla.edu/stat/stata/output/stata_corr_output.htm).

Дополнительная литература

4. Крыштановский А.О. Анализ социологических данных с помощью пакета SPSS : учеб. пособие для вузов. — М.: Изд. дом ГУ ВШЭ, 2006. – С. 47-81.
5. Айвазян С.А., Мхитарян В.С. Прикладная статистика и основы эконометрики. Т.2: Теория вероятностей и прикладная статистика. — М.: Юнити, 2001. Гл. 11. с. 396-457.

Раздел 4. Регрессионный анализ. Модель парной линейной регрессии. Парная линейная регрессия как линейная аппроксимация условного математического ожидания. Оценивание коэффициентов парной линейной регрессии: метод наименьших квадратов. Интерпретация коэффициентов регрессии при непрерывных и категориальных переменных.

Основная литература

1. Доугерти К. Введение в эконометрику: Учебник. М.: ИНФРА-М, 1999. Гл. 2. С. 53–72.
2. Gujarati, D.N. Basic econometrics. New York McGraw-Hill, 2003. Ch. 1, 2.
3. Regression Analysis. UCLA: Academic Technology Services, Statistical Consulting Group. (<http://www.ats.ucla.edu/stat/stata/topics/regression.htm>).
4. Stata Annotated Output. Regression Analysis. UCLA: Academic Technology Services, Statistical Consulting Group. (http://www.ats.ucla.edu/stat/stata/output/reg_output.htm).

Дополнительная литература

5. Тюрин Ю.Н., Макаров А.А. Анализ данных на компьютере: учебное пособие. – М.: ИД «ФОРУМ», 2008. – С. 208-219.
6. Айвазян С.А., Мхитарян В.С. Прикладная статистика и основы эконометрики. Т.2: Основы эконометрики. - М.: ЮНИТИ, 2001. – 656 с.
7. Крыштановский А.О. Анализ социологических данных с помощью пакета SPSS : учеб. пособие для вузов. – М.: Изд. дом ГУ ВШЭ, 2006. – С. 115-146.
8. Cook, R. D. and Weisberg, S. (1999). Applied Regression Including Computing and Graphics. New York: Wiley.
9. Kritzer, H.M., 1996. The Data Puzzle: The Nature of Interpretation in Quantitative Research, *American Journal of Political Science*, 40(1): 1-32.

Раздел 5. Модель множественной линейной регрессии. Теорема Гаусса-Маркова и теорема Рао о свойствах МНК-оценок. Статистический вывод в регрессии: статистическая значимость коэффициентов. Критерии качества моделей, сравнение моделей. Разложение вариации.

Основная литература

1. Доугерти К. Введение в эконометрику: Учебник. М.: ИНФРА-М, 1999. Гл. 3. С. 73–114, Гл. 5. С. 134–140, 146–155, Обзор. С. 15-20, 26-28.
2. Gujarati, D.N. Basic econometrics. New York McGraw-Hill, 2003. Ch. 3, 4.

Дополнительная литература

3. Доугерти К. Введение в эконометрику: Учебник. М.: ИНФРА-М, 1999. Гл. 6. 165-199.



4. Тюрин Ю.Н., Макаров А.А. Анализ данных на компьютере: учебное пособие. – М.: ИД «ФОРУМ», 2008. – С. 208-219.
5. Крыштановский А.О. Анализ социологических данных с помощью пакета SPSS : учеб. пособие для вузов. – М.: Изд. дом ГУ ВШЭ, 2006. – С. 146-166.
6. Cook, R. D. and Weisberg, S. (1999). *Applied Regression Including Computing and Graphics*. New York: Wiley.
7. Lewis-Beck, Michael S. and Andrew Skalaban. (1990). “When to Use R-Squared”. *The Political Methodologist* 3(2): 11–12.
8. King G. (1990). “When Not to Use R-Squared.” *The Political Methodologist*, 3(2): 9–11.
9. Goertzel, T., 2002. “Myths of Murder and Multiple Regression”, *The Sceptical Inquirer* 26, 1: 19-23 (<http://crab.rutgers.edu/%7Egoertzel/mythsofmurder.htm>).

Раздел 6. Мультиколлинеарность. Источники мультиколлинеарности. Последствия мультиколлинеарности для статистического вывода. Способы диагностики. Меры борьбы с мультиколлинеарностью. Автокорреляция. Источники автокорреляции. Последствия автокорреляции для статистического вывода. Способы диагностики. Меры борьбы с автокорреляцией.

Основная литература

1. Доугерти К. Введение в эконометрику: Учебник. М.: ИНФРА-М, 1999. С. Гл. 5. 155–164, Гл. 7. 217–234.
2. Gujarati, D.N. *Basic econometrics*. New York McGraw-Hill, 2003. Ch. 10, 12.

Дополнительная литература

3. Cook, R. D. and Weisberg, S. (1999). *Applied Regression Including Computing and Graphics*. New York: Wiley.

Раздел 7. Гетероскедастичность. Источники гетероскедастичности. Последствия гетероскедастичности для статистического вывода. Способы диагностики. Меры борьбы с гетероскедастичностью. Влиятельные наблюдения.

Основная литература

4. Доугерти К. Введение в эконометрику: Учебник. М.: ИНФРА-М, 1999. Гл. 7. С. 200–217.
5. Gujarati, D.N. *Basic econometrics*. New York McGraw-Hill, 2003. Ch. 11.

Дополнительная литература

6. Braumoeller B. (2006). “Explaining Variance.” *Political Analysis*, Vol.14, No.3, pp. 268–290.
7. Cook, R. D. and Weisberg, S. (1999). *Applied Regression Including Computing and Graphics*. New York: Wiley.

Раздел 8. Эффекты взаимодействия переменных и фиктивные переменные. Фиктивные переменные как способ включения в анализ категориальных переменных. Фиктивные переменные для константы и для коэффициента наклона. Взаимодействия категориальных переменных. Взаимодействия непрерывных переменных.

Основная литература

1. Доугерти К. Введение в эконометрику: Учебник. М.: ИНФРА-М, 1999. Гл. 9. С. 262–285.
2. Gujarati, D.N. *Basic econometrics*. New York McGraw-Hill, 2003. Ch. 9.
3. Brambor Th., Clark W., Golder M. (2006). “Understanding Interaction Models: Improving Empirical Analyses.” *Political Analysis*, Vol. 14, pp. 63–82.

Дополнительная литература



4. Крыштановский А.О. Анализ социологических данных с помощью пакета SPSS : учеб. пособие для вузов. – М.: Изд. дом ГУ ВШЭ, 2006. – С. 166-182.
5. Cook, R. D. and Weisberg, S. (1999). Applied Regression Including Computing and Graphics. New York: Wiley.

Раздел 9. Нелинейные взаимосвязи в регрессии. Линеаризация модели, нелинейной по регрессорам.

Основная литература

1. Доугерти К. Введение в эконометрику: Учебник. М.: ИНФРА-М, 1999. Гл. 4. С. 115–132.
2. Gujarati, D.N. “Basic econometrics”. New York McGraw-Hill, 2003. Ch. 6.

Дополнительная литература

3. Тюрин Ю.Н., Макаров А.А. Анализ данных на компьютере: учебное пособие. – М.: ИД «ФОРУМ», 2008. – С. 219-225.
4. Cook, R. D. and Weisberg, S. (1999). Applied Regression Including Computing and Graphics. New York: Wiley.

8 Образовательные технологии

Занятия проводятся в форме лекций и практических занятий в компьютерном классе.

9 Порядок формирования оценок по дисциплине

Преподаватель оценивает работу студентов на практических занятиях: активность в дискуссиях, правильность решения задач на семинаре, правильность и своевременность решения задач в текущих домашних заданиях¹ и прочих заданиях, которые выдаются на семинарских занятиях. Оценки за работу на практических занятиях преподаватель выставляет в рабочую ведомость. Накопленная оценка по 10-ти балльной шкале за работу на семинарских и практических занятиях определяется перед промежуточным или итоговым контролем – $O_{\text{ауд}}$.

Накопленная оценка учитывает результаты студента по текущему контролю следующим образом:

$$O_{\text{накопл.}} = k_{\text{к/р}} \cdot O_{\text{к/р}} + k_{\text{д/з}} \cdot O_{\text{д/з}} + k_{\text{ауд}} \cdot O_{\text{ауд}},$$

где $k_{\text{к/р}} = 0.4$, $k_{\text{д/з}} = 0.3$, $k_{\text{ауд}} = 0.3$.

Округление каждого компонента накопленной оценки производится в соответствии с правилами математики и происходит до расчета накопленной оценки. Округление накопленной оценки также производится в соответствии с правилами математики.

Результирующая оценка за дисциплину рассчитывается следующим образом:

$$O_{\text{результ}} = k_{\text{накопл}} \cdot O_{\text{накопл}} + k_{\text{зачет}} \cdot O_{\text{зачет}},$$

где $k_{\text{накопл}} = 0.5$, $k_{\text{зачет}} = 0.5$.

Округление результирующей оценки также производится в соответствии с правилами математики.

В случае если $O_{\text{д/з}} < 4$, то до зачета студент может переписать домашнее задание в соответствии с полученными замечаниями. Тогда выполнение домашнего задания в соответствии с требованиями оценивается в 4 балла.

В диплом выставляется результирующая оценка по учебной дисциплине.

¹ Имеются ввиду текущие домашние работы, которые не включаются в РУП, это не форма текущего контроля «Домашнее задание».



10 Учебно-методическое и информационное обеспечение дисциплины

10.1 Базовые учебники

1. Доугерти К. Введение в эконометрику: Учебник. 2-е изд. – М.: ИНФРА-М, 1999. – 402 с. (Полочный индекс в библиотеке НИУ ВШЭ – 330.4 Д71).
2. Gujarati, D.N. “Basic econometrics”. New York McGraw-Hill, 2003. - 1002 с. (Полочный индекс в библиотеке НИУ ВШЭ – 330.4 G94).

10.2 Основная литература

3. Крыштановский А.О. Анализ социологических данных с помощью пакета SPSS: учеб. пособие для вузов. — М.: ГУ–ВШЭ, 2006. – 281 с.
4. Тюрин Ю.Н., Макаров А.А. Анализ данных на компьютере: учебное пособие. – М.: ИД «ФОРУМ», 2008. – 368 с.
5. Cook, R. D. and Weisberg, S. (1999). Applied Regression Including Computing and Graphics. New York: Wiley.
6. Brambor, T., W.R. Clark, M. Golder, 2006. Understanding Interaction Models: Improving Empirical Analysis, *Political Analysis* 14: 63-82.
7. A brief description of Stata. (http://www.stata-press.com/manuals/stata10/guide_brief.pdf).
8. Basics of Stata. (2008). (<http://www.ics.uci.edu/~dvd/Courses/Stata/stata.pdf>).
9. Учебно-методические материалы для преподавания курсов по применению количественного инструментария к решению социально-экономических задач (<http://www.hse.ru/jesda/mathbase/>).
10. Long, J. Scott and Jeremy Freese. 2005. *Regression Models for Categorical Dependent Variables Using Stata*. 2nd Edition. College Station, TX: Stata Press. (Полочный индекс в библиотеке НИУ ВШЭ – 519 L87).

10.3 Дополнительная литература

11. Айвазян С. А., Мхитарян В. С. Прикладная статистика и основы эконометрики. В 2 частях. — М.: Юнити, 2001.
12. Айвазян С.А. Методы эконометрики. Учебник. — М.: Инфра-М, 2010. — 512 с.
13. Вакуленко Е.С., Камалова Р.У., Кисельгоф С.Г., Смирнова Ж.И., Стукал Д.К., Хавенсон Т.Е. Разработка учебно-методических материалов для преподавания курсов по применению количественного инструментария к решению социально-экономических задач. // Экономическая социология (электронный журнал), 2010. Т. 11. № 4. С. 190—194. (http://ecsoc.hse.ru/data/2010/09/29/1234590525/ecsoc_t11_n4.pdf#page=190)
14. Вербик М. Путеводитель по современной эконометрике. Пер. с англ. В. А. Банникова. Под науч. ред. и предисл. С.А. Айвазяна. — М.: Научная книга, 2008. — 616 с.
15. Единый архив социологических данных (рук. — Л. Б. Косова) // Экономическая социология (электронный журнал), 2010, т. 11, № 1, с. 133–140. (http://www.hse.ru/data/2011/12/08/1208204985/ecsoc_t11_n1.pdf#page=133)
16. Магнус Я.Р., Катышев П.К., Пересецкий А.А. Эконометрика. Начальный курс. 6-е изд., перераб. и доп. — М.: Дело, 2004. — 576 с.
17. Тюрин Ю.Н., Макаров А.А., Высоцкий И.Р., Яценко И.В. Теория вероятностей и статистика. — М.: МЦНМО, 2004.
18. Braumoeller B. (2006). “Explaining Variance.” *Political Analysis*, Vol.14, No.3, pp. 268–290.
19. Data Management. UCLA: Academic Technology Services, Statistical Consulting Group. (http://www.ats.ucla.edu/stat/stata/topics/data_management.htm).



20. Descriptive information and statistics. UCLA: Academic Technology Services, Statistical Consulting Group. (<http://www.ats.ucla.edu/stat/stata/modules/descript.htm>).
21. Fox, J., 2008. Applied Regression Analysis and Generalized Linear Models, London: Sage.
22. Goertzel, T., 2002. “Myths of Murder and Multiple Regression”, *The Sceptical Inquirer* 26, 1: 19-23 (<http://crab.rutgers.edu/%7Egoertzel/mythsofmurder.htm>).
23. Graphics. UCLA: Academic Technology Services, Statistical Consulting Group. (<http://www.ats.ucla.edu/stat/stata/topics/graphics.htm>).
24. King G. (1990). “When Not to Use R-Squared.” *The Political Methodologist*, 3(2): 9–11.
25. Kritzer, H.M., 1996. The Data Puzzle: The Nature of Interpretation in Quantitative Research, *American Journal of Political Science*, 40(1): 1-32.
26. Lewis-Beck, Michael S. and Andrew Skalaban. (1990). “When to Use R-Squared”. *The Political Methodologist* 3(2): 11–12.
27. Long, J. Scott. 2009. The workflow of data analysis using Stata. TX: Stata Press. (Полочный индекс в библиотеке НИУ ВШЭ – 004 L87).
28. Regression Analysis. UCLA: Academic Technology Services, Statistical Consulting Group. (<http://www.ats.ucla.edu/stat/stata/topics/regression.htm>).
29. Resources to help you learn and use Stata. UCLA: Academic Technology Services, Statistical Consulting Group. (<http://www.ats.ucla.edu/stat/stata/>).
30. Stata Annotated Output. Regression Analysis. UCLA: Academic Technology Services, Statistical Consulting Group. (http://www.ats.ucla.edu/stat/stata/output/reg_output.htm).

10.4 Программные средства

Для успешного освоения дисциплины, студент использует пакет для анализа данных Stata.

10.5 Дистанционная поддержка дисциплины

При выполнении домашних работ студентам рекомендуется пользоваться материалами

- Единого архива экономических и социологических данных НИУ ВШЭ (<http://sophist.hse.ru/>),
- Проекта «Разработка учебно-методических материалов для преподавания курсов по применению количественного инструментария к решению социально-экономических задач» (<http://www.hse.ru/jesda/mathbase/>).

11 Материально-техническое обеспечение дисциплины

Пакет для анализа данных Stata. Возможно самостоятельное использование других программ: Excel, SPSS, R.