



Правительство Российской Федерации

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

Общеуниверситетская кафедра высшей математики

Программа дисциплины

Анализ категориальных данных в статистических пакетах

для направления 41.03.04. «Политология» подготовки бакалавра

Автор программы:

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

Одобрена на заседании кафедры высшей математики «__» _____ 2016 г.
Зав. кафедрой А.А. Макаров

Рекомендована секцией УМС «_____» «__» _____ 2016 г.
Председатель _____

Утверждена УС факультета прикладной политологии «__» _____ 2016 г.
Ученый секретарь _____

Москва, 2017

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 41.03.04. «Политология» подготовки бакалавра, изучающих дисциплину «Анализ категориальных данных в статистических пакетах».

Программа разработана в соответствии с рабочим учебным планом университета по направлению 030200.62 «Политология» подготовки бакалавра, утвержденным в 2014 г.

2 Цели освоения дисциплины

Целью освоения дисциплины «Анализ категориальных данных в статистических пакетах» является выработка компетенций по решению политологических и социально-экономических задач, связанных с анализом данных в том случае, когда изучаемые зависимые признаки имеют дискретную природу.

3 Компетенции обучающегося, формируемые в результате освоения дисциплины

В результате освоения дисциплины студент должен:

- Знать основные модели бинарного и множественного выбора, позволяющие решать задачи анализа политических и социально-экономических процессов;
- Иметь навыки их реализации в пакетах Stata и R;
- Уметь интерпретировать и представлять полученные результаты;
- Знать в каких областях применяются изученные модели.

В результате освоения дисциплины студент осваивает следующие компетенции:

<i>Компетенция</i>	<i>Код по ФГОС/ НИУ</i>	<i>Дескрипторы — основные признаки освоения (показатели достижения результата)</i>	<i>Формы и методы обучения, способствующие формированию и развитию компетенции</i>
Умение логически верно, аргументировано и ясно строить устную и письменную речь	ОНК2	Демонстрирует способность написания профессиональных текстов в виде эссе/ статьи	Семинары: обсуждение статей ведущих зарубежных реферируемых журналов. Зачет: самостоятельная подготовка эссе/ статьи в соответствии со стандартами ведущих зарубежных реферируемых журналов.
Использование основных положений и методов социальных, гуманитарных и экономических наук при решении профессиональных	ОНК4	Демонстрирует способность отбирать положения политической науки, а также смежных наук, позволяющие интерпретировать результаты анализа	Семинары: обсуждение статей ведущих зарубежных реферируемых журналов. Зачет: самостоятельная подготовка эссе/ статьи в соответствии со стандартами ведущих зарубежных реферируемых журналов.



задач, способность анализировать социальнозначимые проблемы и процессы		данных	
Способность применять математические методы анализа к социальным, политическим и экономическим явлениям	ОНК7	Оценивает и интерпретирует регрессионные модели, графики, корректно представляет результаты	Семинары: реализация методов эконометрики «руками», а также на компьютерах; обсуждение статей ведущих зарубежных реферируемых журналов.
Способность работать с информацией в глобальных компьютерных сетях	ИК2	Демонстрирует способность к нахождению статистической информации в сети Интернет	Зачет: самостоятельная подготовка эссе/ статьи в соответствии со стандартами ведущих зарубежных реферируемых журналов.
Владение иностранным языком на уровне, достаточном для разговорного общения, а также для поиска и анализа иностранных источников информации	ИК3	Демонстрирует способность к чтению профессиональной литературы на английском языке.	Семинары: обсуждение статей ведущих зарубежных реферируемых журналов. Зачет: самостоятельная подготовка эссе/ статьи в соответствии со стандартами ведущих зарубежных реферируемых журналов.

4 Место дисциплины в структуре образовательной программы

Настоящая дисциплина относится к курсам направления 41.03.04. «Политология» подготовки бакалавра. Для указанного направления дисциплина является дисциплиной по выбору.

Изучение данной дисциплины базируется на следующих дисциплинах:

- «Математика и статистика» (1 курс),
- «Вероятностно-статистические модели в политологии» (2 курс),

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

- «Многомерный статистический анализ в политологии»,
- «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	Модели бинарного выбора	40	10	10	20
2	Модели множественного неупорядоченного выбора	32	8	8	16
3	Модели множественного упорядоченного выбора	32	8	8	16
4	Теория статистического оценивания и методы оптимизации	24	6	6	12
Итого		128	32	32	64

6 Формы контроля знаний студентов

Тип контроля	Форма контроля	1 год				Параметры
		1	2	3	4	
Текущий	Контрольная работа	*				Письменная работа, 80 минут.
Текущий	Эссе		*			Самостоятельная письменная работа по итогам анализа массива политологических и социально-экономических данных с использованием изученных методов. Объем: 25-40 тыс. символов с пробелами.
Итоговый	Экзамен		*			Письменная работа, 80 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.

6.2 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 заданиях 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. Накопленная оценка учитывает результаты студента по текущему контролю следующим образом:

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

где $k_{\text{к/р}} = 0.2$, $k_{\text{эссе}} = 0.5$, $k_{\text{ауд.}} = 0.3$.



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:

$$O_{результ} = O_{накопл.} + O_{зачет},$$

где $k_{накопл.} = 0.6$, $k_{зачет} = 0.4$.

Округление результирующей оценки также производится в соответствии с правилами математики.

Если в зачетной работе допущены грубые ошибки (например, интерпретация регрессионного коэффициента без учета нелинейности модели регрессии. Примерный список таких ошибок будет сообщен студентам), то работа оценивается не выше 4 баллов.

В диплом выставляется результирующая оценка по учебной дисциплине.

7 Содержание дисциплины

Раздел 1. Модели бинарного выбора

- Линейная модель регрессии. Дискретные зависимые переменные. Латентный признак и наблюдаемые исходы. Линейная вероятностная модель.
- Оценивание моделей с дискретными зависимыми переменными. Метод максимального правдоподобия (ММП). Модели бинарного выбора: логистическая регрессия (логит-модель) и пробит-модель. Условия на ошибки моделей бинарного выбора.
- Интерпретация коэффициентов логистической регрессии при непрерывных и категориальных переменных. Шансы и отношения шансов. Предсказанные вероятности. Доверительные интервалы. Графическое представление результатов. Сравнение коэффициентов логит- и пробит-моделей.
- Проверка гипотез о линейных ограничениях на коэффициенты: тест Вальда, тест отношения правдоподобия.
- Проверка гипотез об адекватности модели (критерии качества модели), сравнение моделей. (pseudo-R², AIC, BIC). Анализ остатков модели. Выявление нетипичных наблюдений.

Основная литература

1. Long, J. Scott. 1997. *Regression Models for Categorical and Limited Dependent Variables*. Thousand Oaks, CA: Sage. Ch. 2, 3, 4.
2. Long, J. Scott and Jeremy Freese. 2005. *Regression Models for Categorical Dependent Variables Using Stata*. 2nd Edition. College Station, TX: Stata Press. Ch. 3, 4
3. Stata Topics. Logistic (and Categorical) Regression UCLA: Academic Technology Services, Statistical Consulting Group. (http://www.ats.ucla.edu/stat/stata/topics/logistic_regression.htm).

Дополнительная литература

1. 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. (Гл. 7 с. 296-316).
2. Agresti, A. (2002). *Categorical Data Analysis*. John Wiley & Sons, Inc., Hoboken, New Jersey. Pp. 120-125, 165-196, 211-230.
3. Allison, P. D. (1999). Comparing logit and probit coefficients across groups. *Sociological Methods and Research* 28: 186–208.
4. Burnham, Kenneth P. and David R. Anderson (2004). Multimodel Inference: Understanding AIC and BIC in Model Selection. *Sociological Methods & Research*, 33(261), pp. 261-304.
5. Cook, R. D. and Weisberg, S. (1999). *Applied Regression Including Computing and Graphics*. New York: Wiley. Ch. 21, 22.



6. Gujarati, D.N. Basic econometrics. New York McGraw-Hill, 2003. Ch. 15.1-15.10.
7. Peng, C.-Y. J., Lee, K. L., & Ingersoll, G. M. (2002). An introduction to logistic regression analysis and reporting. *The Journal of Educational Research*, 96(1):3-14. (<http://sta559s11.pbworks.com/w/file/fetch/37766848/IntroLogisticRegressionPengEducResearch.pdf>).
8. Peng, C.-Y. J., & So, T.-S. H. (2002). Logistic regression analysis and reporting: A primer. *Understanding Statistics*, 1(1), 31-70. (<http://www.indiana.edu/~jopeng51/teaching-logistic.pdf>).
9. Bo Hu, Jun Shao and Mari Palta (2006). Pseudo-R² in Logistic Regression Model, *Statistica Sinica* 16:847-860. (<http://www3.stat.sinica.edu.tw/statistica/oldpdf/A16n39.pdf>).
10. Logistic Regression with Stata. UCLA: Academic Technology Services, Statistical Consulting Group. (http://www.ats.ucla.edu/stat/stata/seminars/stata_logistic/default.htm).
11. Logistic Regression with Stata by Xiao Chen, Phil Ender, Michael Mitchell & Christine Wells. Stata Web Books. UCLA: Academic Technology Services, Statistical Consulting Group. (<http://www.ats.ucla.edu/stat/stata/webbooks/logistic/>).
12. Stock, J., Watson, M. Introduction to Econometrics, Pearson Addison Wesley, 2008. Pp. 383-400.
13. What are pseudo R-squareds? UCLA: Academic Technology Services, Statistical Consulting Group. (http://www.ats.ucla.edu/stat/mult_pkg/faq/general/pseudo_rsquareds.htm).

Раздел 2. Модели множественного неупорядоченного выбора

- Номинальные переменные. Мультиномиальная логистическая регрессия. Интерпретация коэффициентов мультиномиальной логистической регрессии.
- Условная логистическая регрессия (Conditional logit).
- Предположение о независимости от посторонних альтернатив.

Основная литература

1. Long, J. Scott. 1997. *Regression Models for Categorical and Limited Dependent Variables*. Thousand Oaks, CA: Sage. Ch. 6. (Полочный индекс в библиотеке НИУ ВШЭ – 519 L87).
2. Long, J. Scott and Jeremy Freese. 2005. *Regression Models for Categorical Dependent Variables Using Stata*. 2nd Edition. College Station, TX: Stata Press. Ch 3, 6. (Полочный индекс в библиотеке НИУ ВШЭ – 519 L87).
3. Stata Topics. Logistic (and Categorical) Regression UCLA: Academic Technology Services, Statistical Consulting Group. (http://www.ats.ucla.edu/stat/stata/topics/logistic_regression.htm).

Дополнительная литература

1. 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. (Гл. 7 с. 324-329).
2. Agresti, A. (2002). *Categorical Data Analysis*. John Wiley & Sons, Inc., Hoboken, New Jersey. Pp. 267-302.
3. McFadden, D., 'Conditional Logit Analysis of Qualitative Choice Behaviour', in Zarembka (ed.), *Frontiers in Econometrics*, Academic Press, 1973. (<http://elsa.berkeley.edu/reprints/mcfadden/zarembka.pdf>)
4. R Data Analysis Examples. Multinomial Logistic Regression. UCLA: Academic Technology Services, Statistical Consulting Group. (<http://www.ats.ucla.edu/stat/r/dae/mlogit.htm>).
5. Stata Data Analysis Examples. Multinomial Logistic Regression. UCLA: Academic Technology Services, Statistical Consulting Group. (<http://www.ats.ucla.edu/stat/stata/dae/mlogit.htm>).
6. Stata Library. Understanding RR Ratios in Multinomial Logistic Regression. UCLA: Academic Technology Services, Statistical Consulting Group. (http://www.ats.ucla.edu/stat/stata/library/odds_ratio_mlogit.htm).



7. Steckel J, Vanhonacker W. 1988. A heterogeneous conditional logit model of choice. *Journal of Business and Economic Statistics* 6: 391-398.
8. Thurner P. (2000) The empirical application of the spatial theory of voting in multiparty systems with random utility models. *Electoral Studies*, No. 19, pp. 493–517.

Раздел 3. Модели множественного упорядоченного выбора

- Порядковые переменные. Порядковая логистическая регрессия. Оценивание порядковой логистической регрессии. Интерпретация коэффициентов порядковой логистической регрессии. Предположение о параллельности регрессий: тест Бранта.
- Обобщенная модель множественного упорядоченного выбора (Generalized ordered logit) / частичная модель пропорциональных шансов (Partial proportional odds model).

Основная литература

1. Long, J. Scott. 1997. *Regression Models for Categorical and Limited Dependent Variables*. Thousand Oaks, CA: Sage. Ch. 5. (Полочный индекс в библиотеке НИУ ВШЭ – 519 L87).
2. Long, J. Scott and Jeremy Freese. 2005. *Regression Models for Categorical Dependent Variables Using Stata*. 2nd Edition. College Station, TX: Stata Press. Ch. 3, 5. (Полочный индекс в библиотеке НИУ ВШЭ – 519 L87).
3. Stata Topics. Logistic (and Categorical) Regression UCLA: Academic Technology Services, Statistical Consulting Group. (http://www.ats.ucla.edu/stat/stata/topics/logistic_regression.htm).

Дополнительная литература

1. 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. (Гл. 7 с. 316-324).
2. Beyond Binary: Ordinal Logistic Regression in Stata. UCLA: Academic Technology Services, Statistical Consulting Group. (http://www.ats.ucla.edu/stat/stata/seminars/stata_BeyondBinaryLogistic/Stata_ordinal_seminar.htm).
3. Peterson, Bercedis and Frank E. Harrell. (1990). Partial Proportional Odds Models for Ordinal Response Variables, *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, Vol. 39, No. 2, pp. 205-217.
4. Fu, V. 1998. sg88: Estimating generalized ordered logit models. *Stata Technical Bulletin* 44: 27–30.

Раздел 4. Теория статистического оценивания и методы оптимизации

- Теория статистического оценивания. Метод максимального правдоподобия (ММП) и стандартные ошибки (дальта метод).
- Методы оптимизации.

Основная литература

1. Agresti, A. (2002). *Categorical Data Analysis*. John Wiley & Sons, Inc., Hoboken, New Jersey. Pp. 132-153, 576-594.

Дополнительная литература

1. Agresti, A. (1996). *An Introduction to Categorical Data Analysis*. New York: John Wiley & Sons, Inc. Pp. 600-615.
2. Cameron, A. Colin and Pravin K. *Microeconometrics: Methods and Applications*. Cambridge University Press, 2005. Pp. 116-163, 166-220, 336-352, 943-955.

8 Образовательные технологии

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



9 Оценочные средства для текущего контроля и аттестации студента

9.1 Тематика заданий текущего контроля

Эссе (пример)

Эссе сдается поэтапно.

- обзор литературы и гипотезы (~ 10 окт.) [20% оценки за эссе];
- таблицы выдачи и результаты (~ 20 ноября) [30% оценки за эссе];
- итоговый текст эссе (~ конец курса) [50% оценки за эссе].

Итоговый текст сдается в электронном и печатном виде. Отсутствие печатного текста автоматически влечет неудовлетворительный балл за домашнее задание.

Эссе --- самостоятельная (индивидуальная) письменная работа по результатам анализа массива опросных данных с использованием изученных методов на тему <<Социально-экономические и психологические факторы электорального поведения россиян>>.

Вам даны три массива данных исследования <<Курьер>> Аналитического Центра Юрия Левады, относящиеся к трем разным периодам времени. Задача состоит в том, чтобы построить три одинаковые и разумные модели для трех <<временных точек>> и сравнить полученные результаты. Сравнить в смысле неизменности порождающего данные процесса: выяснить, насколько существенно меняются коэффициенты и содержательные результаты, получаемые по регрессионным моделям.

Для анализа предлагаются следующие зависимые переменные:

1. Электоральный выбор.
2. Поддержка В. Путина и Д. Медведева.
3. Готовность участвовать в протестах.

(Уточнение зависимых переменных, а также их индивидуальное распределение см. в Приложении. Желание сменить назначенную тему на другую (из предложенного списка) может быть обосновано в письме на адрес rkamalova@hse.ru при наличии договоренности об обмене с однокурсником/-цей.)

Предикторы должны быть выбраны самостоятельно в соответствии с теоретическими соображениями (научной литературой).

Требования к структуре эссе:

1. Титульный лист;
2. Аннотация;
3. Введение (постановка задачи, краткий обзор литературы, описывающий исследовательскую проблему и показывающий, куда Вы вносите вклад);
4. Гипотезы (формулируются и обосновываются содержательные гипотезы);
5. Данные и используемые методы (предполагается описание используемых данных и методов; выбор методов требует обоснования; не обязательно использовать все методы, которые вам знакомы);
6. Результаты (представляются в табличной форме, соответствующей зарубежным публикационным стандартам и сопровождаются обязательной интерпретацией);
7. Обсуждение полученных результатов (описывается, в какой мере полученные результаты соответствуют существующей литературе; обсуждаются возможные причины расхождений; обсуждается устойчивость



полученных результатов к смене массива данных, изменению методов анализа);

8. Заключение (формулируются выводы работы, обсуждаются теоретические следствия из полученных выводов, указываются возможные направления дальнейшего уточнения результатов);
9. Список литературы.

Оценка выставляется по 10-балльной шкале. Оцениваются адекватность формулировки задачи/проблемы, соответствие заявленных для проверки гипотез сформулированной задаче, корректность подбора методов и интерпретации результатов, обоснованность выводов, проверка устойчивости результатов. Обращайте внимание на грамотность русского языка и корректность использования терминов, аккуратность оформления, в т.ч. библиографии.

Если объем и характер заимствования, оформленного в виде ссылки, ставят под сомнение самостоятельность выполнения работы, преподаватель действует в соответствии с п. 2.5 Положения о плагиате НИУ ВШЭ.

Задание выполняется в пакете для анализа данных Stata. Для выполнения домашней работы студентам рекомендуется пользоваться руководствами к использованию Stata от UCLA: Academic Technology Services.

9.2 Вопросы для оценки качества освоения дисциплины

Проверочная работа (пример)

1. Покажите, что $P(y_i = 1|X) = F(\beta_0 + \beta_1 x_i)$, приняв, что $y_i = 1$, если латентный $y_i^* > 0$, и $y_i = 0$ в противном случае.
2. Вы оценили модель логистической регрессии на данных GSS (*General Social Survey*) по США

$$P(y_i = 1|X_i) = 2.5 - 0.11education_i + 0.53male_i + 0.70bush92_i - 0.07children_i$$

```
. logit cappun educ male bush92 chldrn, nolog
```

```
Logistic regression                Number of obs   =       5570
                                   LR chi2(4)       =       248.17
                                   Prob > chi2       =       0.0000
Log likelihood = -2831.6609        Pseudo R2      =       0.0420
```

	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
cappun						
educ	-.1133811	.0127096	-8.92	0.000	-.1382916	-.0884707
male	.5295012	.0667557	7.93	0.000	.3986624	.66034
bush92	.7005197	.0736915	9.51	0.000	.556087	.8449524
chldrn	-.0748111	.0221608	-3.38	0.001	-.1182455	-.0313767
_cons	2.500112	.1964794	12.72	0.000	2.115019	2.885204

Зависимая переменная --- отношение к смертной казни (0 --- отрицательное), *education* --- количество лет обучения, *male* --- мужчина (дамми), *bush92* --- голосовал за Буша в 1992 г. (дамми), *children* --- число детей у респондента. Проинтерпретируйте коэффициенты при количестве лет обучения и половой принадлежности в терминах отношения шансов.



3. Во сколько раз шансы мужчины без детей на то, чтобы поддерживать смертную казнь, больше/меньше, чем у женщины с 2 детьми (при прочих равных (зафиксированных на среднем значении)).
4. Какими критериями можно проверить значимость коэффициентов при предикторах в модели бинарного выбора:
 1. Критерий Фишера
 2. Критерий Вальда
 3. Критерий отношения правдоподобия
 4. Критерий Хаусмана
5. Рассчитайте для модели поддержки смертной казни AIC и BIC.
6. У альтернативной модели без переменной *bush92* AIC = 8174.8, BIC = 8202.7. Какую Вы предпочтете? Обоснуйте свой выбор.
7. Рассчитайте скорректированный R^2_{count} , специфичность и чувствительность модели.

```
. estat classification, cutoff(.6)
```

```
Logistic model for cappun
```

Classified	True		Total
	D	~D	
+	4224	1163	5387
-	104	79	183
Total	4328	1242	5570

8. Нарисуйте ROC-кривую. Подпишите оси.

10 Учебно-методическое и информационное обеспечение дисциплины

10.1 Базовые учебники

1. Long, J. Scott. 1997. Regression Models for Categorical and Limited Dependent Variables. Thousand Oaks, CA: Sage. (Полочный индекс в библиотеке НИУ ВШЭ – 519 L87).
2. 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).

10.2 Дополнительная литература

3. Айвазян С.А. Методы эконометрики. Учебник. – М.: Инфра-М, 2010. 512 с.
4. Айвазян С. А., Мхитарян В. С. Теория вероятностей и прикладная статистика, в 2 ч. М.: Юнити, 2001.
5. 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.
6. Крыштановский А.О. Анализ социологических данных с помощью пакета SPSS. М.: ГУ–ВШЭ, 2006.
7. Шандор З. Мультиномиальные модели дискретного выбора // Квантиль, №7, с. 9-19.
8. Agresti, A. (2002). Categorical Data Analysis. John Wiley & Sons, Inc., Hoboken, New Jersey.



9. Agresti, A. and I. Liu. (2005). The analysis of ordered categorical data: An overview and a survey of recent developments, invited discussion paper for the Spanish Statistical Journal, TEST.
10. Allison, P. D. 1999. Comparing logit and probit coefficients across groups. *Sociological Methods and Research* 28: 186–208.
11. Beyond Binary: Ordinal Logistic Regression in Stata. UCLA: Academic Technology Services, Statistical Consulting Group. (http://www.ats.ucla.edu/stat/stata/seminars/stata_BeyondBinaryLogistic/Stata_ordinal_seminar.htm).
12. Bo Hu, Jun Shao and Mari Palta (2006). Pseudo-R² in Logistic Regression Model, *Statistica Sinica* 16:847-860. (<http://www3.stat.sinica.edu.tw/statistica/oldpdf/A16n39.pdf>).
13. Burnham, Kenneth P. and David R. Anderson (2004). Multimodel Inference: Understanding AIC and BIC in Model Selection. *Sociological Methods & Research*, 33(261), pp. 261-304.
14. Cameron, A. Colin and Pravin K. Trivedi. 2005. *Microeconometrics: Methods and Applications*, Cambridge University Press, New York.
15. Cameron, A. Colin and Pravin K. Trivedi. 2009. *Microeconometrics Using Stata*, Cambridge University Press, New York.
16. Fu, V. 1998. sg88: Estimating generalized ordered logit models. *Stata Technical Bulletin* 44: 27–30.
17. Gujarati, D.N. *Basic econometrics*. New York McGraw-Hill, 2003.
18. Hosmer, D. and Lemeshow, S. (2000) *Applied Logistic Regression* (Second Edition). New York: John Wiley & Sons, Inc..
19. Jann, B. 2005. Making regression tables from stored estimates. *Stata Journal* 5: 288–308.
20. *Logistic Regression with Stata*. UCLA: Academic Technology Services, Statistical Consulting Group. (http://www.ats.ucla.edu/stat/stata/seminars/stata_logistic/default.htm).
21. McFadden, D., 'Conditional Logit Analysis of Qualitative Choice Behaviour', in Zarembka (ed.), *Frontiers in Econometrics*, Academic Press, 1973. (<http://elsa.berkeley.edu/reprints/mcfadden/zarembka.pdf>)
22. Menard, S. (1995). *Applied logistic regression analysis*. Thousand Oaks, CA: Sage.
23. Peng, C.-Y. J., & So, T.-S. H. (2002). Logistic regression analysis and reporting: A primer. *Understanding Statistics*, 1(1), 31-70. (<http://www.indiana.edu/~jopeng51/teaching-logistic.pdf>).
24. Peng, C.-Y. J., Lee, K. L., & Ingersoll, G. M. (2002). An introduction to logistic regression analysis and reporting. *The Journal of Educational Research*, 96(1):3-14. (<http://sta559s11.pbworks.com/w/file/attach/37766848/IntroLogisticRegressionPengEducResearch.pdf>).
25. Peterson, Bercedis and Frank E. Harrell. (1990). Partial Proportional Odds Models for Ordinal Response Variables, *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, Vol. 39, No. 2, pp. 205-217.
26. *Stata Data Analysis Examples. Multinomial Logistic Regression*. UCLA: Academic Technology Services, Statistical Consulting Group. (<http://www.ats.ucla.edu/stat/stata/dae/mlogit.htm>).
27. *Stata Topics. Logistic (and Categorical) Regression* UCLA: Academic Technology Services, Statistical Consulting Group. (http://www.ats.ucla.edu/stat/stata/topics/logistic_regression.htm).
28. Steckel J, Vanhonacker W. 1988. A heterogeneous conditional logit model of choice. *Journal of Business and Economic Statistics* 6: 391-398.
29. Stock, J., Watson, M. *Introduction to Econometrics*, Pearson Addison Wesley, 2008.
30. Thurner P. (2000) The empirical application of the spatial theory of voting in multiparty systems with random utility models. *Electoral Studies*, No. 19, pp. 493–517



31. Williams, R. 2006. Generalized ordered logit/partial proportional odds models for ordinal dependent variables. *Stata Journal* 6: 58–82.

10.3 Программные средства

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

10.4 Дистанционная поддержка дисциплины

При выполнении домашних работ студентам рекомендуется пользоваться материалами

- Единого архива экономических и социологических данных НИУ ВШЭ (<http://sophist.hse.ru/>) (преимущественно результаты опросов населения Аналитического Центра Юрия Левады, Фонда Общественное Мнение, Всероссийского центра изучения общественного мнения),
- Межуниверситетского консорциума по политическим и социальным исследованиям (ICPSR) (<http://www.icpsr.umich.edu/icpsrweb/ICPSR/index.jsp>),
- 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

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