

**Эконометрика, 2017-2018, 3 модуль**  
**Семинар 7**  
**26.02.18**

**Для Группы Э\_Б2015\_Э\_3**  
**Семинарист О.А.Демидова**

**Модели бинарного выбора**

1) Демидова, Малахов, 11.5

**Задание 11.5.**

**База данных rlms14 (файл rlms14.dta)**

В файле rlms14.dta приведены данные о заработной плате, возрасте, образовании, профессиональной принадлежности россиян из базы данных RLMS, 14 раунд опроса, проводившийся в 2005 г.

**Переменные:**

income – ответ на вопрос: «Сколько денег в течение последних 30 дней Вы получили по основному месту работы после вычета налогов и отчислений?»

wage - ответ на вопрос: «Ваша среднемесячная зарплата за последние 12 месяцев?»

rsu – первичная единица отбора (1 – Санкт – Петербург, ..., 38 – Амурская обл.),

marst – семейное положение (1 – никогда в браке не состояли, 2- состоите в зарегистрированном браке, 3 – живете вместе, но не зарегистрированы, 4 – разведены и в браке не состоите, 5 – вдовец/вдова),

occup – профессиональная группа (1- военнослужащие, 2 – законодатели, крупные чиновники, управленцы, 3 – профессионалы с высшим образованием, 4 – конторские служащие, 5 – занятые в сфере обслуживания, 6 – квалифицированные с/х работники, 7 – ремесленники, 8 – промышленные рабочие),

educ – образование (1 – 0 классов школы, 23 – аспирантура и т.п. с дипломом),

diplom – законченное образование (1 – 0-6 классов, 1-2 – незаконченное среднее образование, 3 - законченное среднее образование, 4 - законченное среднее образование, 5 - законченное среднее специальное образование, 6 - законченное высшее образование и выше),

gender – пол респондента (1 – мужской, 2 – женский),

birth\_year – год рождения,

industry – ответ на вопрос: «в какой отрасли Вы работаете?» (1 – легкая и пищевая промышленность, ...),

beginning – ответ на вопрос: «с какого года Вы работаете на этом предприятии?»

boss – ответ на вопрос: «У Вас есть подчиненные на этой работе?» (1 – да, 2 – нет),

subordinates – ответ на вопрос: «Сколько у Вас подчиненных?»

duration\_weekh - ответ на вопрос: «Сколько часов в среднем продолжается Ваша обычная рабочая неделя?»

life\_sat - ответ на вопрос: «Насколько Вы удовлетворены или не удовлетворены Вашей жизнью в целом?» (1 - «Полностью удовлетворены», ..., 5 – «Совсем не удовлетворены»),

job\_sat - ответ на вопрос: «Насколько Вы удовлетворены или не удовлетворены Вашей работой в целом?» (1 - «Полностью удовлетворены», 2 - «Скорее удовлетворены», 3 – «И да, и нет», 4 – «Не очень удовлетворены», 5 – «Совсем не удовлетворены»),

job\_career - ответ на вопрос: «Насколько Вы удовлетворены или не удовлетворены возможностями Вашего профессионального роста?» (1 - «Полностью удовлетворены», ..., 5 - «Совсем не удовлетворены»),  
children - ответ на вопрос: «Сколько у Вас детей?»,  
child\_less18 - ответ на вопрос: «Сколько у Вас детей моложе 18 лет?».

По данным файла rlms14.dta с помощью статистического пакета STATA попытайтесь выявить факторы, влияющие на удовлетворенность россиян жизнью в целом, работой, возможностями профессионального роста.

Например, для выявления факторов, влияющих на удовлетворенность россиян жизнью в целом

- 1) Создайте зависимую переменную life, равную 1 для тех, кто на вопрос «Насколько Вы удовлетворены или не удовлетворены Вашей жизнью в целом?» ответили «Полностью удовлетворены» или «Скорее удовлетворены» (т.е. для кого значения переменной life\_sat равно 1 или 2).
- 2) Оцените логит и пробит модели с зависимой переменной life, объясняющие переменные выберите самостоятельно (при необходимости создайте новые, например возраст age = 2005 - birth\_year и т.п.), Вычислите предельные эффекты объясняющих факторов в выбранной Вами точке. Дайте интерпретацию полученным результатам. Используйте статистический пакет STATA.

### **Методические указания.**

Полезные команды:

```
set more off
gen life = 0
replace life = 1 if job_sat < 3
gen age = 2005 - birth_year
tab educ
tab diplom
logit life gender income children age i.diplom
label list psu
logit life gender income children age i.diplom if psu == 2
gen higheduc = 0
replace higheduc = 1 if diplom == 6
probit life gender income children age higheduc if psu == 2
mfx
sum gender income children age higheduc
mfx, at (1 5000 0 25 1)
probit life income children age higheduc if psu == 2 & gender == 2
```

### **Дополнительные задания.**

- 3) Оцените качество оценки моделей

Полезные команды:

```
estat classification
estat gof
estat classification, cutoff(0.4)
```

- 4) Постройте график кривых чувствительности и специфичности

Полезная команда:

```
lsens
```

- 5) Постройте график ROC – кривой и вычислите площадь под ней

Полезная команда:

```
lroc
```

**Эконометрика, 2017-2018, 3 модуль**  
**Семинары 7-8**  
**26.02.18, 05.03.18**  
**Для Группы Э\_Б2015\_Э\_3**  
**Семинарист О.А.Демидова**

**Модели бинарного выбора**

1) Демидова, Малахов, 11.5

**Задание 11.5.**

**База данных rlms14 (файл rlms14.dta)**

В файле rlms14.dta приведены данные о заработной плате, возрасте, образовании, профессиональной принадлежности россиян из базы данных RLMS, 14 раунд опроса, проводившийся в 2005 г.

**Переменные:**

income – ответ на вопрос: «Сколько денег в течение последних 30 дней Вы получили по основному месту работы после вычета налогов и отчислений?»

wage - ответ на вопрос: «Ваша среднемесячная зарплата за последние 12 месяцев?»

rsu – первичная единица отбора (1 – Санкт – Петербург, ..., 38 – Амурская обл.),

marst – семейное положение (1 – никогда в браке не состояли, 2- состоите в зарегистрированном браке, 3 – живете вместе, но не зарегистрированы, 4 – разведены и в браке не состоите, 5 – вдовец/вдова),

occup – профессиональная группа (1- военнослужащие, 2 – законодатели, крупные чиновники, управленцы, 3 – профессионалы с высшим образованием, 4 – конторские служащие, 5 – занятые в сфере обслуживания, 6 – квалифицированные с/х работники, 7 – ремесленники, 8 – промышленные рабочие),

educ – образование (1 – 0 классов школы, 23 – аспирантура и т.п. с дипломом),

diplom – законченное образование (1 – 0-6 классов, 1-2 – незаконченное среднее образование, 3 - законченное среднее образование, 4 - законченное среднее образование, 5 - законченное среднее специальное образование, 6 - законченное высшее образование и выше),

gender – пол респондента (1 – мужской, 2 – женский),

birth\_year – год рождения,

industry – ответ на вопрос: «в какой отрасли Вы работаете?» (1 – легкая и пищевая промышленность, ...),

beginning – ответ на вопрос: «с какого года Вы работаете на этом предприятии?»

boss – ответ на вопрос: «У Вас есть подчиненные на этой работе?» (1 – да, 2 – нет),

subordinates – ответ на вопрос: «Сколько у Вас подчиненных?»

duration\_weekh - ответ на вопрос: «Сколько часов в среднем продолжается Ваша обычная рабочая неделя?»

life\_sat - ответ на вопрос: «Насколько Вы удовлетворены или не удовлетворены Вашей жизнью в целом?» (1 - «Полностью удовлетворены», ..., 5 – «Совсем не удовлетворены»),

job\_sat - ответ на вопрос: «Насколько Вы удовлетворены или не удовлетворены Вашей работой в целом?» (1 - «Полностью удовлетворены», 2 - «Скорее удовлетворены», 3 – «И да, и нет», 4 – «Не очень удовлетворены», 5 – «Совсем не удовлетворены»),

job\_career - ответ на вопрос: «Насколько Вы удовлетворены или не удовлетворены возможностями Вашего профессионального роста?» (1 - «Полностью удовлетворены», ..., 5 - «Совсем не удовлетворены»),  
children - ответ на вопрос: «Сколько у Вас детей?»,  
child\_less18 - ответ на вопрос: «Сколько у Вас детей моложе 18 лет?».

По данным файла rlms14.dta с помощью статистического пакета STATA попытайтесь выявить факторы, влияющие на удовлетворенность россиян жизнью в целом, работой, возможностями профессионального роста.

Например, для выявления факторов, влияющих на удовлетворенность россиян жизнью в целом

- 1) Создайте зависимую переменную life, равную 1 для тех, кто на вопрос «Насколько Вы удовлетворены или не удовлетворены Вашей жизнью в целом?» ответили «Полностью удовлетворены» или «Скорее удовлетворены» (т.е. для кого значения переменной life\_sat равно 1 или 2).
- 2) Оцените логит и пробит модели с зависимой переменной life, объясняющие переменные выберите самостоятельно (при необходимости создайте новые, например возраст age = 2005 - birth\_year и т.п.), Вычислите предельные эффекты объясняющих факторов в выбранной Вами точке. Дайте интерпретацию полученным результатам. Используйте статистический пакет STATA.

### **Методические указания.**

Полезные команды:

```
set more off
gen life = 0
replace life = 1 if job_sat < 3
gen age = 2005 - birth_year
tab educ
tab diplom
logit life gender income children age i.diplom
label list psu
logit life gender income children age i.diplom if psu == 2
gen higheduc = 0
replace higheduc = 1 if diplom == 6
probit life gender income children age higheduc if psu == 2
mfx
sum gender income children age higheduc
mfx, at (1 5000 0 25 1)
probit life income children age higheduc if psu == 2 & gender == 2
```

### **Дополнительные задания.**

- 3) Оцените качество оценки моделей

Полезные команды:

```
estat classification
estat gof
estat classification, cutoff(0.4)
```

- 4) Постройте график кривых чувствительности и специфичности

Полезная команда:

```
lsens
```

- 5) Постройте график ROC – кривой и вычислите площадь под ней

Полезная команда:

```
lroc
```

**Эконометрика, 2017-2018, 3 модуль**  
**Семинары 7-8**  
**26.02.18, 05.03.18**  
**Для Группы Э\_Б2015\_Э\_3**  
**Семинарист О.А.Демидова**

**Модели бинарного выбора**

1) Демидова, Малахов, 11.5

**Задание 11.5.**

**База данных rlms14 (файл rlms14.dta)**

В файле rlms14.dta приведены данные о заработной плате, возрасте, образовании, профессиональной принадлежности россиян из базы данных RLMS, 14 раунд опроса, проводившийся в 2005 г.

**Переменные:**

income – ответ на вопрос: «Сколько денег в течение последних 30 дней Вы получили по основному месту работы после вычета налогов и отчислений?»

wage - ответ на вопрос: «Ваша среднемесячная зарплата за последние 12 месяцев?»

rsu – первичная единица отбора (1 – Санкт – Петербург, ..., 38 – Амурская обл.),

marst – семейное положение (1 – никогда в браке не состояли, 2- состоите в зарегистрированном браке, 3 – живете вместе, но не зарегистрированы, 4 – разведены и в браке не состоите, 5 – вдовец/вдова),

occup – профессиональная группа (1- военнослужащие, 2 – законодатели, крупные чиновники, управленцы, 3 – профессионалы с высшим образованием, 4 – конторские служащие, 5 – занятые в сфере обслуживания, 6 – квалифицированные с/х работники, 7 – ремесленники, 8 – промышленные рабочие),

educ – образование (1 – 0 классов школы, 23 – аспирантура и т.п. с дипломом),

diplom – законченное образование (1 – 0-6 классов, 1-2 – незаконченное среднее образование, 3 - законченное среднее образование, 4 - законченное среднее образование, 5 - законченное среднее специальное образование, 6 - законченное высшее образование и выше),

gender – пол респондента (1 – мужской, 2 – женский),

birth\_year – год рождения,

industry – ответ на вопрос: «в какой отрасли Вы работаете?» (1 – легкая и пищевая промышленность, ...),

beginning – ответ на вопрос: «с какого года Вы работаете на этом предприятии?»

boss – ответ на вопрос: «У Вас есть подчиненные на этой работе?» (1 – да, 2 – нет),

subordinates – ответ на вопрос: «Сколько у Вас подчиненных?»

duration\_weekh - ответ на вопрос: «Сколько часов в среднем продолжается Ваша обычная рабочая неделя?»

life\_sat - ответ на вопрос: «Насколько Вы удовлетворены или не удовлетворены Вашей жизнью в целом?» (1 - «Полностью удовлетворены», ..., 5 – «Совсем не удовлетворены»),

job\_sat - ответ на вопрос: «Насколько Вы удовлетворены или не удовлетворены Вашей работой в целом?» (1 - «Полностью удовлетворены», 2 - «Скорее удовлетворены», 3 – «И да, и нет», 4 – «Не очень удовлетворены», 5 – «Совсем не удовлетворены»),

job\_career - ответ на вопрос: «Насколько Вы удовлетворены или не удовлетворены возможностями Вашего профессионального роста?» (1 - «Полностью удовлетворены», ..., 5 - «Совсем не удовлетворены»),  
children - ответ на вопрос: «Сколько у Вас детей?»,  
child\_less18 - ответ на вопрос: «Сколько у Вас детей моложе 18 лет?».

По данным файла rlms14.dta с помощью статистического пакета STATA попытайтесь выявить факторы, влияющие на удовлетворенность россиян жизнью в целом, работой, возможностями профессионального роста.

Например, для выявления факторов, влияющих на удовлетворенность россиян жизнью в целом

- 1) Создайте зависимую переменную life, равную 1 для тех, кто на вопрос «Насколько Вы удовлетворены или не удовлетворены Вашей жизнью в целом?» ответили «Полностью удовлетворены» или «Скорее удовлетворены» (т.е. для кого значения переменной life\_sat равно 1 или 2).
- 2) Оцените логит и пробит модели с зависимой переменной life, объясняющие переменные выберите самостоятельно (при необходимости создайте новые, например возраст age = 2005 - birth\_year и т.п.), Вычислите предельные эффекты объясняющих факторов в выбранной Вами точке. Дайте интерпретацию полученным результатам. Используйте статистический пакет STATA.

### **Методические указания.**

Полезные команды:

```
set more off
gen life = 0
replace life = 1 if job_sat < 3
gen age = 2005 - birth_year
tab educ
tab diplom
logit life gender income children age i.diplom
label list psu
logit life gender income children age i.diplom if psu == 2
gen higheduc = 0
replace higheduc = 1 if diplom == 6
probit life gender income children age higheduc if psu == 2
mfx
sum gender income children age higheduc
mfx, at (1 5000 0 25 1)
probit life income children age higheduc if psu == 2 & gender == 2
```

### **Дополнительные задания.**

- 3) Оцените качество оценки моделей

Полезные команды:

```
estat classification
estat gof
estat classification, cutoff(0.4)
```

- 4) Постройте график кривых чувствительности и специфичности

Полезная команда:

```
lsens
```

- 5) Постройте график ROC – кривой и вычислите площадь под ней

Полезная команда:

```
Lroc
```

- 2) Демидова, Малахов, задачи 11.1-11.4
- 3) Борzych, Демешев, задача 6.12.