

Глава 7, итерация 1

http://www.youtube.com/watch?v=0BJTraRatFg&list=PLCFTZGvx_s3f_saaG5Hl-wC6uSJb0WFM

7-1-1 Суть метода максимального правдоподобия

<http://www.youtube.com/watch?v=ahu8Xt7LEP0>

1:53 Изменяем заголовок на «Метод максимального правдоподобия»

1:53 под заголовком появляется «ML — Maximum likelihood»

1:55 добавляем надпись ниже:

* Есть неизвестный параметр θ (!!!монтажёр: θ и $\hat{\theta}$ всегда синим цветом)

1:58 добавляем надпись ниже:

* Хотим построить оценку $\hat{\theta}$

2:00 добавляем надпись ниже:

* В качестве оценки неизвестного параметра θ возьмём такое число $\hat{\theta}$, при котором вероятность имеющихся данных максимальна.

2:16 под заголовком сразу появляется пункта:

* Наблюдения, количества звонков: $y_1 = 0, y_2 = 1, y_3 = 2, y_4 = 0$.

* Модель для наблюдений:

y_i	0	1	2
$P(.)$	p	$2p$	$1 - 3p$

2:37 ниже таблицы появляется вопрос задачи:

Оцените \hat{p} методом максимального правдоподобия

7-1-2

<http://www.youtube.com/watch?v=c-zRZ29bz1U>

0:23 переставить слова в появляющейся фразе, должно быть:

Для непрерывных случайных величин максимизируется плотность вероятности

1:47 исправить формулу (сделать предлог «при» чёрным цветом, заменить x на y), должно получиться:

Модель: наблюдения независимы, $f(y) = \lambda e^{-\lambda y}$ при $y > 0$.

1:55 добавляем надпись (сейчас она появляется, но позже)

* Найдите $\hat{\lambda}$

1:58 пропадает строка с игреками, надо её вернуть обратно

7-1-3 Построение доверительных интервалов

http://www.youtube.com/watch?v=_K00JaZ4G1w

0:28 сделать «при» черным цветом:

* Состоятельны: $\hat{\theta}_{ML} \rightarrow \theta$ при $n \rightarrow \infty$

0:38 сделать «при» черным цветом:

* Асимптотически несмещены: $E(\hat{\theta}_{ML}) \rightarrow \theta$ при $n \rightarrow \infty$

0:56 исправить появляющийся пункт на:

* Асимптотически эффективны:

Дисперсия $Var(\hat{\theta}_{ML})$ наименьшая среди асимптотически несмещенных оценок

1:11 дополнительной надписи не появляется (старые пока остаются)

1:38 исправляем первый пункт (под заголовком слайда) на:

* Оценки $\hat{\theta}_{ML}$ асимптотически нормальны

1:40 в появляющейся строке делаем «при» черным цветом

1:50 исправить в появляющейся строке букву l на более каллиграфическую ℓ (а то больно она на I похожа), и слово «информация» с маленькой буквы, должно быть:

I — информация Фишера, $I = -E(\ell''(\theta))$

2:36 снова используем более каллиграфическую ℓ , должно быть:

Наблюдаемая информация Фишера: $\hat{I} = -\ell''(\hat{\theta})$

2:59 во второй формуле сделать более каллиграфическую ℓ , должно быть:

$$se(\hat{\theta}) = \sqrt{Var(\hat{\theta}_{ML})} = \sqrt{-(\ell''(\hat{\theta}))^{-1}}$$

7-1-4 Проверка гипотез. LR тест.

<http://www.youtube.com/watch?v=0BJTraRatFg>

0:24 добавить пропущенные двоеточия, отделить «хотя» от «бы», должно быть:

H_0 : Система из q уравнений на неизвестные параметры

H_a : Хотя бы одно из q условий не выполнено

0:47 в формуле сделать 0 нижним индексом:

$$LR = 2(l(\hat{\theta}) - l(\hat{\theta}_{H_0})) \sim \chi_q^2$$

7-1-5 Логит-модель

<http://www.youtube.com/watch?v=3RULyVUXV74>

0:45 зелёная линия, образующая угол слева — убрать

1:10 под заголовком слайда появляется надпись

$$y_i = \begin{cases} 1, y_i^* \geq 0 \\ 0, y_i^* < 0 \end{cases}$$

* Скрытая переменная: $y_i^* = \beta_1 + \beta_2 x_i + \varepsilon_i$

1:18 ниже появляется надпись:

* Пробит-модель: $\varepsilon_i \sim N(0, 1)$

1:27 ниже надписи «Пробит-модель...» появляется надпись (в формуле исправлено x на t):

* Логит-модель: $\varepsilon_i \sim logistic$, $f(t) = e^{-t}/(1 + e^{-t})^2$

1:35 ниже надписи «Логит-модель...» появляется надпись

* Логистическое распределение похоже на $N(0, 1.6^2)$

1:55 старые надписи стираем, заменяем заголовок слайда на

Вероятность $P(y_i = 1)$

1:56 под заголовком появляется начало формулы

$$P(y_i = 1) = P(y_i^* \geq 0) = P(\beta_1 + \beta_2 x_i + \varepsilon_i \geq 0) =$$

2:10 формула продолжается ниже

$$= P(-\varepsilon_i \leq \beta_1 + \beta_2 x_i) = P(\varepsilon_i \leq \beta_1 + \beta_2 x_i) = F(\beta_1 + \beta_2 x_i)$$

2:40 – 8:58 – вместо старого куска с доской поставить отснятый 1 апреля :)

9:18 исправить вторую формулу на:

$$\ln \frac{P(y_i = 1)}{P(y_i = 0)} = \beta_1 + \beta_2 x_i$$

7-1-6 Интерпретация коэффициентов в логит-модели. Отношение шансов и вероятности. (доска),
6:19

http://www.youtube.com/watch?v=mSV_h2sigS8 ок

7-1-7 Предельные эффекты и прогнозирование

<http://www.youtube.com/watch?v=cjCls8z6gok>

1:04 опечатка в заголовке слайда, должно быть «Интерпретация»

3:35 появляется дополнительный второй пункт:

* Точечный прогноз вероятности: $\hat{P}(y_f = 1) = F(\hat{y}_f^*)$

3:50 появляется два дополнительных пункта:

* Доверительный интервал для $E(\hat{y}_f^*)$:

$$[\hat{y}_f^* - z_{cr} se(\hat{y}_f^*); \hat{y}_f^* + z_{cr} se(\hat{y}_f^*)]$$

* Доверительный интервал для вероятности $P(y_f = 1)$:

$$[F(\hat{y}_f^* - z_{cr} se(\hat{y}_f^*)); F(\hat{y}_f^* + z_{cr} se(\hat{y}_f^*))]$$

!!!монтажеру: эти четыре пункта висят с 3:50 до 4:07 (больше ничего в этот период не добавляется)

4:20 исправляем появляющийся пункт на:

* Логит-модель: $y_i^* = \beta_1 + \beta_2 x_i + u_i$, где u_i примерно $N(0, 1.6^2)$

4:40 появляется дополнительный пункт:

* Логит-модель: $\frac{y_i^*}{1.6} = \frac{\beta_1}{1.6} + \frac{\beta_2}{1.6} x_i + \frac{u_i}{1.6}$, где $\frac{u_i}{1.6}$ примерно $N(0, 1)$

4:43 увеличение не нужно, по смыслу лучше все формулы обозревать

4:57 появляется дополнительный пункт:
 Пробит-модель: $y_i^* = \beta_1 + \beta_2 x_i + \varepsilon_i$, где $\varepsilon_i \sim N(0, 1)$
 5:05 появляется дополнительный пункт ниже всех:
 * $\{y_i = 1\} \Leftrightarrow \{y_i^* > 0\} \Leftrightarrow \{y_i^*/1.6 > 0\}$
 5:26 и далее — отрезать — чёрный экран!!!
 7-1-8 Несуществование ML оценок. Заключение
<http://www.youtube.com/watch?v=08980dxrXW4>
 10:15 добавляем пункт
 * Метод максимального правдоподобия. Позволяет получать оценки неизвестных параметров.
 10:20 добавляем ниже пункт:
 * Логит и пробит модели. Модели для зависимой переменной, принимающей значения 0 и 1.
 10:30 добавляем ниже пункт:
 * МНК не подходит для моделирования бинарной зависимой переменной
 7-2-1 Графический анализ качественных переменных
http://www.youtube.com/watch?v=_7cfL8C98tc
 3:38-3:49 удалить неудачный фрагмент (далее всё идет ок и я говорю те же слова, только без оговорок)
 7-2-2 Оценка коэффициентов и прогнозирование скрытой переменной
<http://www.youtube.com/watch?v=MdwF-uRru9Q>
 3:08 (примерно, где я в камеру смотрю) сюда вставить кусок этого фрагмента с доской (7:40 и до конца)
 7:40 и до конца (весь кусок с доской) нужно вырезать и вставить в точку 3:08
 7-2-3 Доверительный интервал для вероятности и LR-тест
 0:16 исправить название фрагмента (на синей полосе внизу) на «Доверительный интервал для вероятности и LR-тест»
http://www.youtube.com/watch?v=w_20Cd0tC5k
 2:48 — 4:08 вырезать полностью
 в конец вставить начало из 7-2-4 (от начала и до 2:34)

7-2-4 Пределные эффекты

<http://www.youtube.com/watch?v=4hl0bAdp45E>

0:16 исправить название фрагмента (на синей полосе внизу) на «Пределные эффекты» от начала и до 2:34 вырезать и вставить в конец 7-2-3

4:21–4:29 вырезать (там я говорю, что надо разместить слева скриншот из программы)

4:29 – 9:20 на прозрачной доске слева надо разместить выдачу из софта (я рассказываю справа, а слева висит выдача):

```
maBina(m_logit)

##           effect error t.value p.value
## (Intercept)  0.821 0.091   8.979  0.000
## sexmale      -0.551 0.029 -18.673  0.000
## age          -0.008 0.001  -5.391  0.000
## pclass2nd    -0.264 0.046  -5.747  0.000
## pclass3rd    -0.485 0.047 -10.271  0.000
## fare         0.000 0.000   0.191  0.849
```

```
maBina(m_logit,x.mean = FALSE)

##           effect error t.value p.value
## (Intercept)  0.528 0.059   8.979  0.000
## sexmale      -0.551 0.029 -18.673  0.000
## age          -0.005 0.001  -5.391  0.000
## pclass2nd    -0.264 0.046  -5.747  0.000
## pclass3rd    -0.485 0.047 -10.271  0.000
## fare         0.000 0.000   0.191  0.849
```

9:21-9:22 вырезать «угу»

7-2-5 ROC кривая

<http://www.youtube.com/watch?v=G2iKEg0QJ0c>

0:16 исправить название фрагмента (на синей полосе внизу) на «ROC кривая»