

# Микроэконометрика

## Модели бинарного выбора

Потанин Богдан Станиславович

доцент, научный сотрудник, кандидат экономических наук

2023-2024

# Модели порядкового выбора

## Мотивация

- Часто исследователи сталкиваются с порядковыми зависимыми переменными:

Зависимая переменная	Возможные значения
Здоровье	плохое (0), среднее (1), хорошее (2)
Удовлетворенностью жизнью	очень низкая (0), низкая (1), средняя (2), высокая (3), очень высокая (4)
Уровень образования	отсутствует (0), незаконченное среднее (1), среднее (2), среднее специальное (3), высшее (4), кандидат наук (5) доктор наук (6)
Уровень дохода (в тысячах рублей)	менее 50 (0), 50-100 (1), более 100 (3)
Степень согласия с некоторым утверждением	полностью не согласен (0), скорее не согласен (1), скорее согласен (2), полностью согласен (3)

# Модели порядкового выбора

## Мотивация

- Часто исследователи сталкиваются с порядковыми зависимыми переменными:

Зависимая переменная	Возможные значения
Здоровье	плохое (0), среднее (1), хорошее (2)
Удовлетворенностью жизнью	очень низкая (0), низкая (1), средняя (2), высокая (3), очень высокая (4)
Уровень образования	отсутствует (0), незаконченное среднее (1), среднее (2), среднее специальное (3), высшее (4), кандидат наук (5) доктор наук (6)
Уровень дохода (в тысячах рублей)	менее 50 (0), 50-100 (1), более 100 (3)
Степень согласия с некоторым утверждением	полностью не согласен (0), скорее не согласен (1), скорее согласен (2), полностью согласен (3)

- Для построения регрессий с порядковыми зависимыми переменными используются модели порядкового выбора.

# Модели порядкового выбора

## Формализация

- Латентная переменная определяется по аналогии с моделями бинарного выбора:

$$y_i^* = x_i\beta + \varepsilon_i$$

# Модели порядкового выбора

## Формализация

- Латентная переменная определяется по аналогии с моделями бинарного выбора:

$$y_i^* = x_i\beta + \varepsilon_i$$

- Наблюдаемое значение порядковой переменной определяется интервалом, в который попала латентная переменная:

$$y_i = \begin{cases} 0, & \text{если } y_i^* \leq c_1 \\ 1, & \text{если } c_1 < y_i^* \leq c_2 \\ 2, & \text{если } c_2 < y_i^* \leq c_3 \\ \vdots & \\ J, & \text{если } y_i^* > c_J \end{cases}, c = (c_1, \dots, c_J)$$

где  $c_j$  являются пороговыми значениями, при пересечении которых порядковая переменная принимает очередное значение.

# Модели порядкового выбора

## Формализация

- Латентная переменная определяется по аналогии с моделями бинарного выбора:

$$y_i^* = x_i\beta + \varepsilon_i$$

- Наблюдаемое значение порядковой переменной определяется интервалом, в который попала латентная переменная:

$$y_i = \begin{cases} 0, & \text{если } y_i^* \leq c_1 \\ 1, & \text{если } c_1 < y_i^* \leq c_2 \\ 2, & \text{если } c_2 < y_i^* \leq c_3 \\ \vdots & \\ J, & \text{если } y_i^* > c_J \end{cases}, c = (c_1, \dots, c_J)$$

где  $c_j$  являются пороговыми значениями, при пересечении которых порядковая переменная принимает очередное значение.

- При  $J = 1$  модели порядкового выбора превращаются в модели бинарного выбора.

# Модели порядкового выбора

Оценивание с помощью метода максимального правдоподобия

- Для удобства в дальнейшем будем рассматривать порядковую пробит модель, то есть предполагая  $\varepsilon_i \sim \mathcal{N}(0, 1)$ .

# Модели порядкового выбора

Оценивание с помощью метода максимального правдоподобия

- Для удобства в дальнейшем будем рассматривать порядковую пробит модель, то есть предполагая  $\varepsilon_i \sim \mathcal{N}(0, 1)$ .
- Вероятность попадания в  $j$ -ю категорию рассчитывается как:

$$\begin{aligned} P(y_i = j | x_i) &= P(c_j < y_i^* \leq c_{j+1} | x_i) = P(c_j < x_i\beta + \varepsilon_i \leq c_{j+1} | x_i) = \\ &= P(c_j - x_i\beta < \varepsilon_i \leq c_{j+1} - x_i\beta | x_i) = \Phi(c_{j+1} - x_i\beta) - \Phi(c_j - x_i\beta), \end{aligned}$$

где для удобства расчета вероятностей попадания в наибольшую и наименьшую категории полагаем  $c_0 = -\infty$  и  $c_{J+1} = \infty$ .



# Модели порядкового выбора

## Оценивание с помощью метода максимального правдоподобия

- Для удобства в дальнейшем будем рассматривать порядковую пробит модель, то есть предполагая  $\varepsilon_i \sim \mathcal{N}(0, 1)$ .
- Вероятность попадания в  $j$ -ю категорию рассчитывается как:

$$\begin{aligned} P(y_i = j | x_i) &= P(c_j < y_i^* \leq c_{j+1} | x_i) = P(c_j < x_i \beta + \varepsilon_i \leq c_{j+1} | x_i) = \\ &= P(c_j - x_i \beta < \varepsilon_i \leq c_{j+1} - x_i \beta | x_i) = \Phi(c_{j+1} - x_i \beta) - \Phi(c_j - x_i \beta), \end{aligned}$$

где для удобства расчета вероятностей попадания в наибольшую и наименьшую категории полагаем  $c_0 = -\infty$  и  $c_{J+1} = \infty$ .

- Пользуясь полученными вероятностями запишем функцию правдоподобия:

$$L(\beta, c_1, \dots, c_J; y | X) = \prod_{i=1}^n \Phi(c_{y_i+1} - x_i \beta) - \Phi(c_{y_i} - x_i \beta)$$

# Модели порядкового выбора

## Оценивание с помощью метода максимального правдоподобия

- Для удобства в дальнейшем будем рассматривать порядковую пробит модель, то есть предполагая  $\varepsilon_i \sim \mathcal{N}(0, 1)$ .
- Вероятность попадания в  $j$ -ю категорию рассчитывается как:

$$\begin{aligned} P(y_i = j | x_i) &= P(c_j < y_i^* \leq c_{j+1} | x_i) = P(c_j < x_i \beta + \varepsilon_i \leq c_{j+1} | x_i) = \\ &= P(c_j - x_i \beta < \varepsilon_i \leq c_{j+1} - x_i \beta | x_i) = \Phi(c_{j+1} - x_i \beta) - \Phi(c_j - x_i \beta), \end{aligned}$$

где для удобства расчета вероятностей попадания в наибольшую и наименьшую категории полагаем  $c_0 = -\infty$  и  $c_{J+1} = \infty$ .

- Пользуясь полученными вероятностями запишем функцию правдоподобия:

$$L(\beta, c_1, \dots, c_J; y | X) = \prod_{i=1}^n \Phi(c_{y_i+1} - x_i \beta) - \Phi(c_{y_i} - x_i \beta)$$

- Особенности распределения случайных ошибок, включая гетероскедастичность, случайные эффекты и отклонение от нормальности, учитываются точно так же, как в моделях бинарного выбора.

# Модели порядкового выбора

## Предельные эффекты непрерывных переменных

- Через  $x_{ki}$  обозначим  $k$ -й элемент (регрессор) вектора независимых переменных  $x_i$ . Предположим, что он измерен в непрерывной шкале.

# Модели порядкового выбора

## Предельные эффекты непрерывных переменных

- Через  $x_{ki}$  обозначим  $k$ -й элемент (регрессор) вектора независимых переменных  $x_i$ . Предположим, что он измерен в непрерывной шкале.
- Предельный эффект переменной  $x_{ki}$  на вероятность попадания зависимой порядковой переменной в  $j$ -ю категорию рассчитывается как:

$$\frac{\partial P(y_i = j | x_i)}{\partial x_{ki}} = \beta_k (\phi(c_j - x_i \beta) - \phi(c_{j+1} - x_i \beta))$$

# Модели порядкового выбора

## Предельные эффекты непрерывных переменных

- Через  $x_{ki}$  обозначим  $k$ -й элемент (регрессор) вектора независимых переменных  $x_i$ . Предположим, что он измерен в непрерывной шкале.
- Предельный эффект переменной  $x_{ki}$  на вероятность попадания зависимой порядковой переменной в  $j$ -ю категорию рассчитывается как:

$$\frac{\partial P(y_i = j | x_i)}{\partial x_{ki}} = \beta_k (\phi(c_j - x_i \beta) - \phi(c_{j+1} - x_i \beta))$$

- Знак предельного эффекта однозначно определяется лишь для крайних категорий: совпадает со знаком  $\beta_k$  при  $j = J$  и имеет противоположный знак при  $j = 0$ . Для средних категорий  $0 < j < J$  знак предельного эффекта зависит от разницы функций плотности, а значит и от  $x_i$ .

# Модели порядкового выбора

## Предельные эффекты непрерывных переменных

- Через  $x_{ki}$  обозначим  $k$ -й элемент (регрессор) вектора независимых переменных  $x_i$ . Предположим, что он измерен в непрерывной шкале.
- Предельный эффект переменной  $x_{ki}$  на вероятность попадания зависимой порядковой переменной в  $j$ -ю категорию рассчитывается как:

$$\frac{\partial P(y_i = j | x_i)}{\partial x_{ki}} = \beta_k (\phi(c_j - x_i \beta) - \phi(c_{j+1} - x_i \beta))$$

- Знак предельного эффекта однозначно определяется лишь для крайних категорий: совпадает со знаком  $\beta_k$  при  $j = J$  и имеет противоположный знак при  $j = 0$ . Для средних категорий  $0 < j < J$  знак предельного эффекта зависит от разницы функций плотности, а значит и от  $x_i$ .
- Также, знак предельного эффекта определен для вероятности попадания не менее (по аналогии для не более) чем в определенную категорию:

$$\begin{aligned} P(y_i \geq j | x_i) &= P(y_i^* > c_j) = P(\varepsilon_i > c_j - x_i \beta) = \Phi(x_i \beta - c_j) \implies \\ &\implies \partial P(y_i \geq j | x_i) / \partial x_{ki} = \beta_k \phi(x_i \beta - c_j) \end{aligned}$$

# Модели порядкового выбора

## Допущение о параллельности регрессии

- Вероятность попадания в ту или иную категорию зависит от линейного индекса  $x_i\beta$  с постоянным вектором коэффициентов  $\beta$  – **допущение о параллельности регрессии** (parallel regression assumption).

# Модели порядкового выбора

## Допущение о параллельности регрессии

- Вероятность попадания в ту или иную категорию зависит от линейного индекса  $x_i\beta$  с постоянным вектором коэффициентов  $\beta$  – **допущение о параллельности регрессии** (parallel regression assumption).
- Предполагается, что данное допущение является нереалистичным, поскольку регрессоры могут по-разному (в том числе в разном направлении) влиять на вероятность перехода из одной категории в другую.



# Модели порядкового выбора

## Допущение о параллельности регрессии

- Вероятность попадания в ту или иную категорию зависит от линейного индекса  $x_i\beta$  с постоянным вектором коэффициентов  $\beta$  – **допущение о параллельности регрессии** (parallel regression assumption).
- Предполагается, что данное допущение является нереалистичным, поскольку регрессоры могут по-разному (в том числе в разном направлении) влиять на вероятность перехода из одной категории в другую.
- В силу этого допущения коэффициентов  $\beta$  (кроме константы) можно оценить за счет бинарной регрессии с заменой порядковой переменной  $y_i$  на бинарную переменную, принимающую значение 1 при  $y_i \geq j$  и 0 – в противном случае, где  $0 < j < J$ . Обозначим полученную таким образом оценку (без константы) как  $\hat{\beta}^{(j)}$ .

# Модели порядкового выбора

## Допущение о параллельности регрессии

- Вероятность попадания в ту или иную категорию зависит от линейного индекса  $x_i\beta$  с постоянным вектором коэффициентов  $\beta$  – **допущение о параллельности регрессии** (parallel regression assumption).
- Предполагается, что данное допущение является нереалистичным, поскольку регрессоры могут по-разному (в том числе в разном направлении) влиять на вероятность перехода из одной категории в другую.
- В силу этого допущения коэффициентов  $\beta$  (кроме константы) можно оценить за счет бинарной регрессии с заменой порядковой переменной  $y_i$  на бинарную переменную, принимающую значение 1 при  $y_i \geq j$  и 0 – в противном случае, где  $0 < j < J$ . Обозначим полученную таким образом оценку (без константы) как  $\hat{\beta}^{(j)}$ .
- Если допущение о параллельности регрессии соблюдается, то коэффициенты  $\beta^{(j)}$  не зависят от  $j$ , что позволяет сформулировать нулевую гипотезу  $H_0 : \beta_1 = \dots = \beta_{J-1}$ . Проверка этой гипотезы осуществляется с помощью **теста Бранта**, основанного на тесте Вальда.

# Модели порядкового выбора

## Допущение о параллельности регрессии

- Вероятность попадания в ту или иную категорию зависит от линейного индекса  $x_i\beta$  с постоянным вектором коэффициентов  $\beta$  – **допущение о параллельности регрессии** (parallel regression assumption).
- Предполагается, что данное допущение является нереалистичным, поскольку регрессоры могут по-разному (в том числе в разном направлении) влиять на вероятность перехода из одной категории в другую.
- В силу этого допущения коэффициентов  $\beta$  (кроме константы) можно оценить за счет бинарной регрессии с заменой порядковой переменной  $y_i$  на бинарную переменную, принимающую значение 1 при  $y_i \geq j$  и 0 – в противном случае, где  $0 < j < J$ . Обозначим полученную таким образом оценку (без константы) как  $\hat{\beta}^{(j)}$ .
- Если допущение о параллельности регрессии соблюдается, то коэффициенты  $\beta^{(j)}$  не зависят от  $j$ , что позволяет сформулировать нулевую гипотезу  $H_0 : \beta_1 = \dots = \beta_{J-1}$ . Проверка этой гипотезы осуществляется с помощью **теста Бранта**, основанного на тесте Вальда.
- Поскольку не существует процесса генерации данных, способного учесть описанное различие в регрессионных коэффициентах, то отклонение нулевой гипотезы принято связывать не непосредственно с нарушением описанного допущения, а с иными проблемами спецификации модели, такими как спецификация линейного индекса и распределения случайной ошибки.