

# Микроэконометрика

## Модели с неслучайным отбором

Потанин Богдан Станиславович

старший преподаватель, кандидат экономических наук

2021-2022

- Иногда значение зависимой переменной наблюдается лишь при соблюдении некоторого условия (правила).

# Неслучайный отбор

## Мотивация

- Иногда значение зависимой переменной наблюдается лишь при соблюдении некоторого условия (правила).
- Например, зарплата наблюдается лишь для работающих индивидов, а затраты на покупку в приложении лишь для тех, кто им пользуется.

- Иногда значение зависимой переменной наблюдается лишь при соблюдении некоторого условия (правила).
- Например, зарплата наблюдается лишь для работающих индивидов, а затраты на покупку в приложении лишь для тех, кто им пользуется.
- В отличие от модели Тобина модели с неслучайным отбором предполагают, что правило, определяющее попадание наблюдений в выборку, моделируется отдельно.

- Иногда значение зависимой переменной наблюдается лишь при соблюдении некоторого условия (правила).
- Например, зарплата наблюдается лишь для работающих индивидов, а затраты на покупку в приложении лишь для тех, кто им пользуется.
- В отличие от модели Тобина модели с неслучайным отбором предполагают, что правило, определяющее попадание наблюдений в выборку, моделируется отдельно.
- Например, в моделях с неслучайным отбором одновременно моделируется как зарплата, так и занятость индивида. Причем различные факторы могут по разному (в том числе с разным знаком) влиять на ожидаемую зарплату и вероятность занятости.

# Усеченное двумерное нормальное распределение

Частный случай с односторонним усечением одной компоненты

- Случайные величины  $X_1$  и  $X_2$  имеют совместное нормальное распределение:

$$(X_1, X_2) \sim \mathcal{N} \left( \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}, \begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix} \right)$$

# Усеченное двумерное нормальное распределение

Частный случай с односторонним усечением одной компоненты

- Случайные величины  $X_1$  и  $X_2$  имеют совместное нормальное распределение:

$$(X_1, X_2) \sim \mathcal{N} \left( \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}, \begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix} \right)$$

- Напомним, что:

$$E(X_1 | x_2 = t) = \mu_1 + \rho \frac{\sigma_1}{\sigma_2} (t - \mu_2),$$

# Усеченное двумерное нормальное распределение

Частный случай с односторонним усечением одной компоненты

- Случайные величины  $X_1$  и  $X_2$  имеют совместное нормальное распределение:

$$(X_1, X_2) \sim \mathcal{N} \left( \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}, \begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix} \right)$$

- Напомним, что:

$$E(X_1|x_2 = t) = \mu_1 + \rho \frac{\sigma_1}{\sigma_2} (t - \mu_2), \quad \text{Var}(X_1|x_2 = t) = (1 - \rho^2)\sigma_1^2$$



# Усеченное двумерное нормальное распределение

Частный случай с односторонним усечением одной компоненты

- Случайные величины  $X_1$  и  $X_2$  имеют совместное нормальное распределение:

$$(X_1, X_2) \sim \mathcal{N} \left( \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}, \begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix} \right)$$

- Напомним, что:

$$E(X_1|x_2 = t) = \mu_1 + \rho \frac{\sigma_1}{\sigma_2} (t - \mu_2), \quad \text{Var}(X_1|x_2 = t) = (1 - \rho^2)\sigma_1^2$$

- При усечении  $X_2$  получаем:

# Усеченное двумерное нормальное распределение

Частный случай с односторонним усечением одной компоненты

- Случайные величины  $X_1$  и  $X_2$  имеют совместное нормальное распределение:

$$(X_1, X_2) \sim \mathcal{N} \left( \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}, \begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix} \right)$$

- Напомним, что:

$$E(X_1|X_2 = t) = \mu_1 + \rho \frac{\sigma_1}{\sigma_2} (t - \mu_2), \quad \text{Var}(X_1|X_2 = t) = (1 - \rho^2)\sigma_1^2$$

- При усечении  $X_2$  получаем:

$$E(X_1|X_2 > t) = \mu_2 + \rho\sigma_2\lambda(t^*),$$

$$\lambda(t^*) = \frac{\phi(t^*)}{\Phi(t^*)},$$

$$t^* = \frac{t - \mu_1}{\sigma_1}$$

# Усеченное двумерное нормальное распределение

Частный случай с односторонним усечением одной компоненты

- Случайные величины  $X_1$  и  $X_2$  имеют совместное нормальное распределение:

$$(X_1, X_2) \sim \mathcal{N} \left( \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}, \begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix} \right)$$

- Напомним, что:

$$E(X_1|X_2 = t) = \mu_1 + \rho \frac{\sigma_1}{\sigma_2} (t - \mu_2), \quad \text{Var}(X_1|X_2 = t) = (1 - \rho^2)\sigma_1^2$$

- При усечении  $X_2$  получаем:

$$E(X_1|X_2 > t) = \mu_2 + \rho\sigma_2\lambda(t^*), \quad E(X_1|X_2 < t) = \mu_2 - \rho\sigma_2\lambda(-t^*)$$

$$\lambda(t^*) = \frac{\phi(t^*)}{\Phi(t^*)},$$

$$t^* = \frac{t - \mu_1}{\sigma_1}$$

# Усеченное двумерное нормальное распределение

Частный случай с односторонним усечением одной компоненты

- Случайные величины  $X_1$  и  $X_2$  имеют совместное нормальное распределение:

$$(X_1, X_2) \sim \mathcal{N} \left( \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}, \begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix} \right)$$

- Напомним, что:

$$E(X_1|X_2 = t) = \mu_1 + \rho \frac{\sigma_1}{\sigma_2} (t - \mu_2), \quad \text{Var}(X_1|X_2 = t) = (1 - \rho^2)\sigma_1^2$$

- При усечении  $X_2$  получаем:

$$\begin{aligned} E(X_1|X_2 > t) &= \mu_2 + \rho\sigma_2\lambda(t^*), & E(X_1|X_2 < t) &= \mu_2 - \rho\sigma_2\lambda(-t^*) \\ \text{Var}(X_1|X_2 > t) &= \sigma_2^2 (1 - \rho^2\delta(t^*)), \\ \lambda(t^*) &= \frac{\phi(t^*)}{\Phi(t^*)}, & \delta(t^*) &= \lambda(t^*)(\lambda(t^*) - t^*), & t^* &= \frac{t - \mu_1}{\sigma_1} \end{aligned}$$

# Усеченное двумерное нормальное распределение

Частный случай с односторонним усечением одной компоненты

- Случайные величины  $X_1$  и  $X_2$  имеют совместное нормальное распределение:

$$(X_1, X_2) \sim \mathcal{N} \left( \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}, \begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix} \right)$$

- Напомним, что:

$$E(X_1|X_2 = t) = \mu_1 + \rho \frac{\sigma_1}{\sigma_2} (t - \mu_2), \quad \text{Var}(X_1|X_2 = t) = (1 - \rho^2)\sigma_1^2$$

- При усечении  $X_2$  получаем:

$$\begin{aligned} E(X_1|X_2 > t) &= \mu_2 + \rho\sigma_2\lambda(t^*), & E(X_1|X_2 < t) &= \mu_2 - \rho\sigma_2\lambda(-t^*) \\ \text{Var}(X_1|X_2 > t) &= \sigma_2^2 (1 - \rho^2\delta(t^*)), & \text{Var}(X_1|X_2 < t) &= \sigma_2^2 (1 - \rho^2\delta(-t^*)) \\ \lambda(t^*) &= \frac{\phi(t^*)}{\Phi(t^*)}, & \delta(t^*) &= \lambda(t^*)(\lambda(t^*) - t^*), & t^* &= \frac{t - \mu_1}{\sigma_1} \end{aligned}$$

# Неслучайный отбор

## Формулировка

- Имеется два уравнения:

Целевое уравнение:  $y_i^* = x_i\beta + \varepsilon_i$

Уравнение отбора:  $z_i^* = w_i\gamma + u_i$

# Неслучайный отбор

## Формулировка

- Имеется два уравнения:

Целевое уравнение:  $y_i^* = x_i\beta + \varepsilon_i$

Уравнение отбора:  $z_i^* = w_i\gamma + u_i$

- Значение  $y_i^*$  наблюдается лишь при соблюдении определенного условия (правила):

$$z_i = \begin{cases} 1, & \text{если } z_i^* \geq 0 \\ 0, & \text{в противном случае} \end{cases} \quad y_i = \begin{cases} y_i^*, & \text{если } z_i = 1 \\ \text{не наблюдаем}, & \text{в противном случае} \end{cases}$$

# Неслучайный отбор

## Формулировка

- Имеется два уравнения:

Целевое уравнение:  $y_i^* = x_i\beta + \varepsilon_i$

Уравнение отбора:  $z_i^* = w_i\gamma + u_i$

- Значение  $y_i^*$  наблюдается лишь при соблюдении определенного условия (правила):

$$z_i = \begin{cases} 1, & \text{если } z_i^* \geq 0 \\ 0, & \text{в противном случае} \end{cases} \quad y_i = \begin{cases} y_i^*, & \text{если } z_i = 1 \\ \text{не наблюдаем}, & \text{в противном случае} \end{cases}$$

- Например,  $y_i^*$  может отражать потенциальную заработную плату индивида, которая наблюдается лишь в случае, когда индивид работает, то есть  $z_i = 1$ .



# Неслучайный отбор

## Формулировка

- Имеется два уравнения:

Целевое уравнение:  $y_i^* = x_i\beta + \varepsilon_i$

Уравнение отбора:  $z_i^* = w_i\gamma + u_i$

- Значение  $y_i^*$  наблюдается лишь при соблюдении определенного условия (правила):

$$z_i = \begin{cases} 1, & \text{если } z_i^* \geq 0 \\ 0, & \text{в противном случае} \end{cases} \quad y_i = \begin{cases} y_i^*, & \text{если } z_i = 1 \\ \text{не наблюдаем}, & \text{в противном случае} \end{cases}$$

- Например,  $y_i^*$  может отражать потенциальную заработную плату индивида, которая наблюдается лишь в случае, когда индивид работает, то есть  $z_i = 1$ .
- Для простоты предположим, что случайные ошибки имеют совместное нормальное распределение:

$$(u_i, \varepsilon_i) \sim \mathcal{N} \left( \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & \rho\sigma \\ \rho\sigma & \sigma^2 \end{bmatrix} \right), \text{i.i.d.}$$

- Поскольку  $y_i$  наблюдается лишь при  $z_i = 1$ , то:

$$E(y_i | w_i, x_i) = E(y_i^* | z_i = 1, w_i, x_i) = x_i \beta + E(\varepsilon_i | u_i \geq -w_i \gamma, w_i, x_i),$$

- Поскольку  $y_i$  наблюдается лишь при  $z_i = 1$ , то:

$$E(y_i | w_i, x_i) = E(y_i^* | z_i = 1, w_i, x_i) = x_i \beta + E(\varepsilon_i | u_i \geq -w_i \gamma, w_i, x_i),$$

где по свойствам усеченного двумерного нормального распределения:

$$E(\varepsilon_i | u_i \geq -w_i \gamma, w_i, x_i) = \frac{\phi(w_i \gamma)}{\Phi(w_i \gamma)} = \lambda_i(w_i \gamma) = \lambda_i,$$

- Поскольку  $y_i$  наблюдается лишь при  $z_i = 1$ , то:

$$E(y_i | w_i, x_i) = E(y_i^* | z_i = 1, w_i, x_i) = x_i \beta + E(\varepsilon_i | u_i \geq -w_i \gamma, w_i, x_i),$$

где по свойствам усеченного двумерного нормального распределения:

$$E(\varepsilon_i | u_i \geq -w_i \gamma, w_i, x_i) = \frac{\phi(w_i \gamma)}{\Phi(w_i \gamma)} = \lambda_i(w_i \gamma) = \lambda_i,$$

- Отметим, что  $\lambda_i$  часто именуют лямбдой Хекмана.

- Поскольку  $y_i$  наблюдается лишь при  $z_i = 1$ , то:

$$E(y_i | w_i, x_i) = E(y_i^* | z_i = 1, w_i, x_i) = x_i \beta + E(\varepsilon_i | u_i \geq -w_i \gamma, w_i, x_i),$$

где по свойствам усеченного двумерного нормального распределения:

$$E(\varepsilon_i | u_i \geq -w_i \gamma, w_i, x_i) = \frac{\phi(w_i \gamma)}{\Phi(w_i \gamma)} = \lambda_i(w_i \gamma) = \lambda_i,$$

- Отметим, что  $\lambda_i$  часто именуют лямбдой Хекмана.
- В результате регрессионное уравнение может быть записано как:

$$y_i = x_i \beta + \lambda_i + v_i,$$

- Поскольку  $y_i$  наблюдается лишь при  $z_i = 1$ , то:

$$E(y_i | w_i, x_i) = E(y_i^* | z_i = 1, w_i, x_i) = x_i \beta + E(\varepsilon_i | u_i \geq -w_i \gamma, w_i, x_i),$$

где по свойствам усеченного двумерного нормального распределения:

$$E(\varepsilon_i | u_i \geq -w_i \gamma, w_i, x_i) = \frac{\phi(w_i \gamma)}{\Phi(w_i \gamma)} = \lambda_i(w_i \gamma) = \lambda_i,$$

- Отметим, что  $\lambda_i$  часто именуют лямбдой Хекмана.
- В результате регрессионное уравнение может быть записано как:

$$y_i = x_i \beta + \lambda_i + v_i, \quad v_i = \varepsilon_i - \lambda_i \implies E(v_i | x_i, w_i) = 0$$

- Поскольку  $y_i$  наблюдается лишь при  $z_i = 1$ , то:

$$E(y_i | w_i, x_i) = E(y_i^* | z_i = 1, w_i, x_i) = x_i \beta + E(\varepsilon_i | u_i \geq -w_i \gamma, w_i, x_i),$$

где по свойствам усеченного двумерного нормального распределения:

$$E(\varepsilon_i | u_i \geq -w_i \gamma, w_i, x_i) = \frac{\phi(w_i \gamma)}{\Phi(w_i \gamma)} = \lambda_i(w_i \gamma) = \lambda_i,$$

- Отметим, что  $\lambda_i$  часто именуют лямбдой Хекмана.
- В результате регрессионное уравнение может быть записано как:

$$y_i = x_i \beta + \lambda_i + v_i, \quad v_i = \varepsilon_i - \lambda_i \implies E(v_i | x_i, w_i) = 0$$

- Без учета  $\lambda_i$  при  $\rho \neq 0$  и наличии корреляции между  $\lambda_i$  и  $x_i$  МНК оценки коэффициентов  $\beta$  окажутся несостоятельными вследствие проблемы пропущенной переменной.

# Метод Хекмана: двухшаговая процедура

## Оценивание

- **Идея метода:** истинное значение  $\lambda_i$  исследователю неизвестно, поскольку зависит от неизвестных параметров  $\gamma$ . Однако, оценив параметры  $\gamma$  можно получить состоятельную оценку  $\hat{\lambda}_i$  и использовать ее вместо  $\lambda_i$  для того, чтобы избежать смещения в оценках вследствие пропущенной переменной.



# Метод Хекмана: двухшаговая процедура

## Оценивание

- **Идея метода:** истинное значение  $\lambda_i$  исследователю неизвестно, поскольку зависит от неизвестных параметров  $\gamma$ . Однако, оценив параметры  $\gamma$  можно получить состоятельную оценку  $\hat{\lambda}_i$  и использовать ее вместо  $\lambda_i$  для того, чтобы избежать смещения в оценках вследствие пропущенной переменной.
- Двухшаговая процедура оценивания:
  - **Первый шаг:** при помощи пробит модели оцениваются параметры  $\gamma$ . В силу инвариантности ММП оценок состоятельная оценка  $\lambda_i$  рассчитывается как  $\hat{\lambda}_i = \lambda_i(w_i\hat{\gamma})$ .

# Метод Хекмана: двухшаговая процедура

## Оценивание

- **Идея метода:** истинное значение  $\lambda_i$  исследователю неизвестно, поскольку зависит от неизвестных параметров  $\gamma$ . Однако, оценив параметры  $\gamma$  можно получить состоятельную оценку  $\hat{\lambda}_i$  и использовать ее вместо  $\lambda_i$  для того, чтобы избежать смещения в оценках вследствие пропущенной переменной.
- **Двухшаговая процедура оценивания:**
  - **Первый шаг:** при помощи пробит модели оцениваются параметры  $\gamma$ . В силу инвариантности ММП оценок состоятельная оценка  $\lambda_i$  рассчитывается как  $\hat{\lambda}_i = \lambda_i(w_i\hat{\gamma})$ .
  - **Второй шаг:** В регрессионное уравнение для  $y_i$  подставляется  $\hat{\lambda}_i$  в качестве дополнительного регрессора с коэффициентом  $\rho\sigma$ . Затем  $\beta$  и  $\rho\sigma$  оцениваются при помощи МНК.

# Метод Хекмана: двухшаговая процедура

## Оценивание

- **Идея метода:** истинное значение  $\lambda_i$  исследователю неизвестно, поскольку зависит от неизвестных параметров  $\gamma$ . Однако, оценив параметры  $\gamma$  можно получить состоятельную оценку  $\hat{\lambda}_i$  и использовать ее вместо  $\lambda_i$  для того, чтобы избежать смещения в оценках вследствие пропущенной переменной.
- **Двухшаговая процедура оценивания:**
  - **Первый шаг:** при помощи пробит модели оцениваются параметры  $\gamma$ . В силу инвариантности ММП оценок состоятельная оценка  $\lambda_i$  рассчитывается как  $\hat{\lambda}_i = \lambda_i(w_i\hat{\gamma})$ .
  - **Второй шаг:** В регрессионное уравнение для  $y_i$  подставляется  $\hat{\lambda}_i$  в качестве дополнительного регрессора с коэффициентом  $\rho\sigma$ . Затем  $\beta$  и  $\rho\sigma$  оцениваются при помощи МНК.
- Состоятельные оценки  $\hat{\sigma}^2$  и  $\hat{\rho}$  можно получить как:

$$\hat{\sigma}^2 = \sum_{i=1}^n (y_i - \hat{y}_i)^2 + (\hat{\rho}\sigma)^2 \hat{\lambda}_i (\lambda_i - w_i\gamma),$$

# Метод Хекмана: двухшаговая процедура

## Оценивание

- **Идея метода:** истинное значение  $\lambda_i$  исследователю неизвестно, поскольку зависит от неизвестных параметров  $\gamma$ . Однако, оценив параметры  $\gamma$  можно получить состоятельную оценку  $\hat{\lambda}_i$  и использовать ее вместо  $\lambda_i$  для того, чтобы избежать смещения в оценках вследствие пропущенной переменной.
- **Двухшаговая процедура оценивания:**
  - **Первый шаг:** при помощи пробит модели оцениваются параметры  $\gamma$ . В силу инвариантности ММП оценок состоятельная оценка  $\lambda_i$  рассчитывается как  $\hat{\lambda}_i = \lambda_i(w_i\hat{\gamma})$ .
  - **Второй шаг:** В регрессионное уравнение для  $y_i$  подставляется  $\hat{\lambda}_i$  в качестве дополнительного регрессора с коэффициентом  $\rho\sigma$ . Затем  $\beta$  и  $\rho\sigma$  оцениваются при помощи МНК.
- Состоятельные оценки  $\hat{\sigma}^2$  и  $\hat{\rho}$  можно получить как:

$$\hat{\sigma}^2 = \sum_{i=1}^n (y_i - \hat{y}_i)^2 + (\hat{\rho}\hat{\sigma})^2 \hat{\lambda}_i (\lambda_i - w_i\gamma), \quad \hat{\rho} = \frac{\widehat{\rho\sigma}}{\sqrt{\hat{\sigma}^2}}$$

# Метод Хекмана: двухшаговая процедура

## Тестирование гипотез

- Случайные ошибки гетероскедастичны, поскольку по свойствам усеченного двумерного нормального распределения:

$$E(v_i^2 | z_i = 1, w_i, x_i) = \sigma^2 (1 - \rho^2 \delta_i), \quad \delta_i = \lambda_i (\lambda_i - w_i \gamma)$$

# Метод Хекмана: двухшаговая процедура

## Тестирование гипотез

- Случайные ошибки гетероскедастичны, поскольку по свойствам усеченного двумерного нормального распределения:

$$E(v_i^2 | z_i = 1, w_i, x_i) = \sigma^2 (1 - \rho^2 \delta_i), \quad \delta_i = \lambda_i (\lambda_i - w_i \gamma)$$

- Для коррекции ковариационной матрицы оценок регрессионных коэффициентов необходимо учесть как гетероскедастичность, так и то, что вместо истинного значения  $\lambda_i$  используется его оценка, зависящая от  $\hat{\gamma}$ , откуда:

$$\widehat{\text{As.Cov}}(\hat{\beta}^*) = \hat{\sigma}^2 (X_*^T X_*)^{-1} (\hat{A}_1 + \hat{A}_2) (X_*^T X_*)^{-1}$$

# Метод Хекмана: двухшаговая процедура

## Тестирование гипотез

- Случайные ошибки гетероскедастичны, поскольку по свойствам усеченного двумерного нормального распределения:

$$E(v_i^2 | z_i = 1, w_i, x_i) = \sigma^2 (1 - \rho^2 \delta_i), \quad \delta_i = \lambda_i (\lambda_i - w_i \gamma)$$

- Для коррекции ковариационной матрицы оценок регрессионных коэффициентов необходимо учесть как гетероскедастичность, так и то, что вместо истинного значения  $\lambda_i$  используется его оценка, зависящая от  $\hat{\gamma}$ , откуда:

$$\widehat{\text{As.Cov}}(\hat{\beta}^*) = \hat{\sigma}^2 (X_*^T X_*)^{-1} (\hat{A}_1 + \hat{A}_2) (X_*^T X_*)^{-1}$$

$$\hat{A}_1 = X_*^T (I - \hat{\rho}^2 \hat{\Delta}) X_*,$$

# Метод Хекмана: двухшаговая процедура

## Тестирование гипотез

- Случайные ошибки гетероскедастичны, поскольку по свойствам усеченного двумерного нормального распределения:

$$E(v_i^2 | z_i = 1, w_i, x_i) = \sigma^2 (1 - \rho^2 \delta_i), \quad \delta_i = \lambda_i (\lambda_i - w_i \gamma)$$

- Для коррекции ковариационной матрицы оценок регрессионных коэффициентов необходимо учесть как гетероскедастичность, так и то, что вместо истинного значения  $\lambda_i$  используется его оценка, зависящая от  $\hat{\gamma}$ , откуда:

$$\widehat{\text{As.Cov}}(\hat{\beta}^*) = \hat{\sigma}^2 (X_*^T X_*)^{-1} (\hat{A}_1 + \hat{A}_2) (X_*^T X_*)^{-1}$$
$$\hat{A}_1 = X_*^T (I - \hat{\rho}^2 \hat{\Delta}) X_*, \quad \hat{A}_2 = \hat{\rho}^2 (X_* \hat{\Delta} W) \widehat{\text{As.Cov}}(\hat{\gamma}) (X_* \hat{\Delta} W)^T,$$



# Метод Хекмана: двухшаговая процедура

## Тестирование гипотез

- Случайные ошибки гетероскедастичны, поскольку по свойствам усеченного двумерного нормального распределения:

$$E(v_i^2 | z_i = 1, w_i, x_i) = \sigma^2 (1 - \rho^2 \delta_i), \quad \delta_i = \lambda_i (\lambda_i - w_i \gamma)$$

- Для коррекции ковариационной матрицы оценок регрессионных коэффициентов необходимо учесть как гетероскедастичность, так и то, что вместо истинного значения  $\lambda_i$  используется его оценка, зависящая от  $\hat{\gamma}$ , откуда:

$$\widehat{\text{As.Cov}}(\hat{\beta}^*) = \hat{\sigma}^2 (X_*^T X_*)^{-1} (\hat{A}_1 + \hat{A}_2) (X_*^T X_*)^{-1}$$

$$\hat{A}_1 = X_*^T (I - \hat{\rho}^2 \hat{\Delta}) X_*, \quad \hat{A}_2 = \hat{\rho}^2 (X_* \hat{\Delta} W) \widehat{\text{As.Cov}}(\hat{\gamma}) (X_* \hat{\Delta} W)^T,$$

где  $\beta_* = (\beta, \rho\sigma)^T$  и  $X_*$  является матрицей регрессоров, полученной за счет присоединения столбца  $\hat{\lambda}$  к матрице  $X$  справа.

# Метод Хекмана: двухшаговая процедура

## Тестирование гипотез

- Случайные ошибки гетероскедастичны, поскольку по свойствам усеченного двумерного нормального распределения:

$$E(v_i^2 | z_i = 1, w_i, x_i) = \sigma^2 (1 - \rho^2 \delta_i), \quad \delta_i = \lambda_i (\lambda_i - w_i \gamma)$$

- Для коррекции ковариационной матрицы оценок регрессионных коэффициентов необходимо учесть как гетероскедастичность, так и то, что вместо истинного значения  $\lambda_i$  используется его оценка, зависящая от  $\hat{\gamma}$ , откуда:

$$\widehat{\text{As.Cov}}(\hat{\beta}^*) = \hat{\sigma}^2 (X_*^T X_*)^{-1} (\hat{A}_1 + \hat{A}_2) (X_*^T X_*)^{-1}$$

$$\hat{A}_1 = X_*^T (I - \hat{\rho}^2 \hat{\Delta}) X_*, \quad \hat{A}_2 = \hat{\rho}^2 (X_* \hat{\Delta} W) \widehat{\text{As.Cov}}(\hat{\gamma}) (X_* \hat{\Delta} W)^T,$$

где  $\beta_* = (\beta, \rho\sigma)^T$  и  $X_*$  является матрицей регрессоров, полученной за счет присоединения столбца  $\hat{\lambda}$  к матрице  $X$  справа.

- Элементы  $\hat{A}_1$  и  $\hat{A}_2$  позволяют учесть гетероскедастичность и использование оценок  $\lambda_i$  вместо истинных значений соответственно.

# Метод Хекмана: двухшаговая процедура

## Ограничения исключения (exclusion restrictions)

- Лямбда Хекмана  $\lambda(w_i\gamma)$  крайне близка к линейной функции при  $w_i\gamma \in (-\infty, 2)$ , то есть в данном диапазоне  $\lambda(w_i\gamma) \approx cw_i\gamma$ , где  $c \in R_{++}$ .

# Метод Хекмана: двухшаговая процедура

## Ограничения исключения (exclusion restrictions)

- Лямбда Хекмана  $\lambda(w_i\gamma)$  крайне близка к линейной функции при  $w_i\gamma \in (-\infty, 2)$ , то есть в данном диапазоне  $\lambda(w_i\gamma) \approx c w_i\gamma$ , где  $c \in R_{++}$ .
- Из-за этого при сильном сходстве между  $x_i$  и  $w_i$  может возникнуть сильная коллинеарность между  $\lambda(w_i\gamma)$  и  $x_i$ .

# Метод Хекмана: двухшаговая процедура

## Ограничения исключения (exclusion restrictions)

- Лямбда Хекмана  $\lambda(w_i\gamma)$  крайне близка к линейной функции при  $w_i\gamma \in (-\infty, 2)$ , то есть в данном диапазоне  $\lambda(w_i\gamma) \approx c w_i\gamma$ , где  $c \in R_{++}$ .
- Из-за этого при сильном сходстве между  $x_i$  и  $w_i$  может возникнуть сильная коллинеарность между  $\lambda(w_i\gamma)$  и  $x_i$ .
- Эта коллинеарность часто приводит к существенному снижению в эффективности оценок.

# Метод Хекмана: двухшаговая процедура

## Ограничения исключения (exclusion restrictions)

- Лямбда Хекмана  $\lambda(w_i\gamma)$  крайне близка к линейной функции при  $w_i\gamma \in (-\infty, 2)$ , то есть в данном диапазоне  $\lambda(w_i\gamma) \approx cw_i\gamma$ , где  $c \in R_{++}$ .
- Из-за этого при сильном сходстве между  $x_i$  и  $w_i$  может возникнуть сильная коллинеарность между  $\lambda(w_i\gamma)$  и  $x_i$ .
- Эта коллинеарность часто приводит к существенному снижению в эффективности оценок.
- Для смягчения проблемы коллинеарности, как правило, исследователи пытаются обеспечить наличие **ограничений исключения**: регрессоров, входящих в  $w_i$  и не входящих в  $x_i$ .

# Метод Хекмана: двухшаговая процедура

## Ограничения исключения (exclusion restrictions)

- Лямбда Хекмана  $\lambda(w_i\gamma)$  крайне близка к линейной функции при  $w_i\gamma \in (-\infty, 2)$ , то есть в данном диапазоне  $\lambda(w_i\gamma) \approx c w_i\gamma$ , где  $c \in R_{++}$ .
- Из-за этого при сильном сходстве между  $x_i$  и  $w_i$  может возникнуть сильная коллинеарность между  $\lambda(w_i\gamma)$  и  $x_i$ .
- Эта коллинеарность часто приводит к существенному снижению в эффективности оценок.
- Для смягчения проблемы коллинеарности, как правило, исследователи пытаются обеспечить наличие **ограничений исключения**: регрессоров, входящих в  $w_i$  и не входящих в  $x_i$ .
- Например, исследователь может предположить, что количество детей влияет на вероятность занятости (входит в  $w_i$ ), но не влияет на зарплату индивида (не входит в  $x_i$ ).

# Метод Хекмана: двухшаговая процедура

## Ограничения исключения (exclusion restrictions)

- Лямбда Хекмана  $\lambda(w_i\gamma)$  крайне близка к линейной функции при  $w_i\gamma \in (-\infty, 2)$ , то есть в данном диапазоне  $\lambda(w_i\gamma) \approx c w_i\gamma$ , где  $c \in R_{++}$ .
- Из-за этого при сильном сходстве между  $x_i$  и  $w_i$  может возникнуть сильная коллинеарность между  $\lambda(w_i\gamma)$  и  $x_i$ .
- Эта коллинеарность часто приводит к существенному снижению в эффективности оценок.
- Для смягчения проблемы коллинеарности, как правило, исследователи пытаются обеспечить наличие **ограничений исключения**: регрессоров, входящих в  $w_i$  и не входящих в  $x_i$ .
- Например, исследователь может предположить, что количество детей влияет на вероятность занятости (входит в  $w_i$ ), но не влияет на зарплату индивида (не входит в  $x_i$ ).
- В качестве более устойчивой к отсутствию ограничений исключения альтернативы вместо двухшаговой процедуры можно воспользоваться методом максимального правдоподобия.



# Метод Хекмана: метод максимального правдоподобия

## Оценивание

- Оценки параметров  $\beta$ ,  $\rho$  и  $\sigma$  можно также получить за счет максимизации функции правдоподобия:

$$L(\beta, \rho, \sigma; y, z|X, W) = \prod_{i: z_i=1} f_{y_i|x_i, w_i}(y_i) P(z_i = 1|y_i, x_i, w_i) \prod_{i: z_i=0} P(z_i = 0|x_i, w_i) =$$

# Метод Хекмана: метод максимального правдоподобия

## Оценивание

- Оценки параметров  $\beta$ ,  $\rho$  и  $\sigma$  можно также получить за счет максимизации функции правдоподобия:

$$\begin{aligned} L(\beta, \rho, \sigma; y, z|X, W) &= \prod_{i:z_i=1} f_{y_i|x_i, w_i}(y_i) P(z_i = 1|y_i, x_i, w_i) \prod_{i:z_i=0} P(z_i = 0|x_i, w_i) = \\ &= \prod_{i:z_i=1} \frac{1}{\sigma} \phi\left(\frac{y_i - x_i\beta}{\sigma}\right) \Phi\left(\frac{\rho(y_i - x_i\beta)/\sigma + w_i\gamma}{\sqrt{1 - \rho^2}}\right) \prod_{i:z_i=0} (1 - \Phi(w_i\gamma)) \end{aligned}$$

# Метод Хекмана: метод максимального правдоподобия

## Оценивание

- Оценки параметров  $\beta$ ,  $\rho$  и  $\sigma$  можно также получить за счет максимизации функции правдоподобия:

$$\begin{aligned} L(\beta, \rho, \sigma; y, z|X, W) &= \prod_{i:z_i=1} f_{y_i|x_i, w_i}(y_i) P(z_i = 1|y_i, x_i, w_i) \prod_{i:z_i=0} P(z_i = 0|x_i, w_i) = \\ &= \prod_{i:z_i=1} \frac{1}{\sigma} \phi\left(\frac{y_i - x_i\beta}{\sigma}\right) \Phi\left(\frac{\rho(y_i - x_i\beta)/\sigma + w_i\gamma}{\sqrt{1 - \rho^2}}\right) \prod_{i:z_i=0} (1 - \Phi(w_i\gamma)) \end{aligned}$$

- Оценки ММП метода более эффективны, чем оценки двухшаговой процедуры.

# Метод Хекмана: метод максимального правдоподобия

## Оценивание

- Оценки параметров  $\beta$ ,  $\rho$  и  $\sigma$  можно также получить за счет максимизации функции правдоподобия:

$$\begin{aligned} L(\beta, \rho, \sigma; y, z|X, W) &= \prod_{i:z_i=1} f_{y_i|x_i, w_i}(y_i) P(z_i = 1|y_i, x_i, w_i) \prod_{i:z_i=0} P(z_i = 0|x_i, w_i) = \\ &= \prod_{i:z_i=1} \frac{1}{\sigma} \phi\left(\frac{y_i - x_i\beta}{\sigma}\right) \Phi\left(\frac{\rho(y_i - x_i\beta)/\sigma + w_i\gamma}{\sqrt{1 - \rho^2}}\right) \prod_{i:z_i=0} (1 - \Phi(w_i\gamma)) \end{aligned}$$

- Оценки ММП метода более эффективны, чем оценки двухшаговой процедуры.
- Недостаток ММП заключается в сложности технической реализации, связанной с возможностью наличия несколько локальных максимумов функции правдоподобия.

# Метод Хекмана: метод максимального правдоподобия

## Оценивание

- Оценки параметров  $\beta$ ,  $\rho$  и  $\sigma$  можно также получить за счет максимизации функции правдоподобия:

$$\begin{aligned} L(\beta, \rho, \sigma; y, z|X, W) &= \prod_{i:z_i=1} f_{y_i|x_i, w_i}(y_i) P(z_i = 1|y_i, x_i, w_i) \prod_{i:z_i=0} P(z_i = 0|x_i, w_i) = \\ &= \prod_{i:z_i=1} \frac{1}{\sigma} \phi\left(\frac{y_i - x_i\beta}{\sigma}\right) \Phi\left(\frac{\rho(y_i - x_i\beta)/\sigma + w_i\gamma}{\sqrt{1 - \rho^2}}\right) \prod_{i:z_i=0} (1 - \Phi(w_i\gamma)) \end{aligned}$$

- Оценки ММП метода более эффективны, чем оценки двухшаговой процедуры.
- Недостаток ММП заключается в сложности технической реализации, связанной с возможностью наличия несколько локальных максимумов функции правдоподобия.
- Для тестирования гипотезы о наличии неслучайного отбора достаточно проверить  $H_0 : \rho = 0$  для ММП или  $H_0 : \rho\sigma = 0$  для двухшаговой процедуры. Если нулевая гипотеза не отвергается, то МНК оценки несостоятельны, что мотивирует применение метода Хекмана.

- Предельный эффект переменной  $x_{ik}$  на обычное математическое ожидание имеет такой же вид, как в случае с обычной линейной регрессией:

$$\frac{\partial E(y_i^* | x_i)}{\partial x_{ik}} = \beta_k$$

- Предельный эффект переменной  $x_{ik}$  на обычное математическое ожидание имеет такой же вид, как в случае с обычной линейной регрессией:

$$\frac{\partial E(y_i^* | x_i)}{\partial x_{ik}} = \beta_k$$

- Предельный эффект на условное математическое ожидание рассчитывается как:

$$\frac{\partial E(y_i^* | a < y_i^* < b)}{\partial x_{ik}} = \beta_k - \gamma_* \rho \sigma \delta_i,$$

где  $\gamma_*$  является коэффициентом при  $x_{ki}$  в уравнении отбора, если  $x_{ki}$  входит в  $w_i$ . В противном случае  $\gamma_* = 0$ .

- Предельный эффект переменной  $x_{ik}$  на обычное математическое ожидание имеет такой же вид, как в случае с обычной линейной регрессией:

$$\frac{\partial E(y_i^* | x_i)}{\partial x_{ik}} = \beta_k$$

- Предельный эффект на условное математическое ожидание рассчитывается как:

$$\frac{\partial E(y_i^* | a < y_i^* < b)}{\partial x_{ik}} = \beta_k - \gamma_* \rho \sigma \delta_i,$$

где  $\gamma_*$  является коэффициентом при  $x_{ki}$  в уравнении отбора, если  $x_{ki}$  входит в  $w_i$ . В противном случае  $\gamma_* = 0$ .

- Предельный эффект на условное математическое ожидание целевой переменной складывается из предельного эффекта на безусловное математическое ожидание  $\beta_k$  и части, обусловленной наличием неслучайного отбора  $\gamma_* \rho \sigma \delta_i$ .



- При нарушении допущения о совместном нормальном распределении случайных ошибок оценки метода Хекмана могут оказаться несостоятельными.

- При нарушении допущения о совместном нормальном распределении случайных ошибок оценки метода Хекмана могут оказаться несостоятельными.
- В качестве альтернативы допущению о конкретной форме совместного распределения случайных ошибок условное математическое ожидание случайной ошибки основного уравнения можно аппроксимировать при помощи полинома  $k$ -й степени:

$$E(\varepsilon_i | z_i = 1, w_i, x_i) \approx \sum_{t=0}^k \tau_t g(w_i \gamma)^t, \quad \tau = (\tau_1, \dots, \tau_k),$$

- При нарушении допущения о совместном нормальном распределении случайных ошибок оценки метода Хекмана могут оказаться несостоятельными.
- В качестве альтернативы допущению о конкретной форме совместного распределения случайных ошибок условное математическое ожидание случайной ошибки основного уравнения можно аппроксимировать при помощи полинома  $k$ -й степени:

$$E(\varepsilon_i | z_i = 1, w_i, x_i) \approx \sum_{t=0}^k \tau_t g(w_i \gamma)^t, \quad \tau = (\tau_1, \dots, \tau_k),$$

где  $g(w_i \gamma)$  является произвольно выбираемой сглаживающей функцией, в качестве которой, как правило, рассматривают  $g(w_i \gamma) = w_i \gamma$  или  $g(w_i \gamma) = \lambda(w_i \gamma)$ .

- При нарушении допущения о совместном нормальном распределении случайных ошибок оценки метода Хекмана могут оказаться несостоятельными.
- В качестве альтернативы допущению о конкретной форме совместного распределения случайных ошибок условное математическое ожидание случайной ошибки основного уравнения можно аппроксимировать при помощи полинома  $k$ -й степени:

$$E(\varepsilon_i | z_i = 1, w_i, x_i) \approx \sum_{t=0}^k \tau_t g(w_i \gamma)^t, \quad \tau = (\tau_1, \dots, \tau_k),$$

где  $g(w_i \gamma)$  является произвольно выбираемой сглаживающей функцией, в качестве которой, как правило, рассматривают  $g(w_i \gamma) = w_i \gamma$  или  $g(w_i \gamma) = \lambda(w_i \gamma)$ .

- На **первом шаге** параметры  $\gamma$  оцениваются при помощи полупараметрической модели бинарного выбора (например, метода Галланта и Нички), а на **втором шаге** все  $k$  переменных  $g(w_i \hat{\gamma})^t$  подставляются в целевое уравнение в качестве регрессоров (в дополнении к  $x_i$ ), в котором параметры  $\beta$  и  $\tau$  оцениваются с помощью МНК.

- При нарушении допущения о совместном нормальном распределении случайных ошибок оценки метода Хекмана могут оказаться несостоятельными.
- В качестве альтернативы допущению о конкретной форме совместного распределения случайных ошибок условное математическое ожидание случайной ошибки основного уравнения можно аппроксимировать при помощи полинома  $k$ -й степени:

$$E(\varepsilon_i | z_i = 1, w_i, x_i) \approx \sum_{t=0}^k \tau_t g(w_i \gamma)^t, \quad \tau = (\tau_1, \dots, \tau_k),$$

где  $g(w_i \gamma)$  является произвольно выбираемой сглаживающей функцией, в качестве которой, как правило, рассматривают  $g(w_i \gamma) = w_i \gamma$  или  $g(w_i \gamma) = \lambda(w_i \gamma)$ .

- На **первом шаге** параметры  $\gamma$  оцениваются при помощи полупараметрической модели бинарного выбора (например, метода Галланта и Нички), а на **втором шаге** все  $k$  переменных  $g(w_i \hat{\gamma})^t$  подставляются в целевое уравнение в качестве регрессоров (в дополнении к  $x_i$ ), в котором параметры  $\beta$  и  $\tau$  оцениваются с помощью МНК.
- Оценки данного метода состоятельные и асимптотически нормальные. Для тестирования гипотез и оценивания асимптотической ковариационной матрицы оценок регрессионных коэффициентов, как правило, применяют бутстрап.

- По мере увеличения степени полинома  $k$  растет точность аппроксимации, что позволяет снизить смещение оценок. Однако, вместе с ростом  $k$  увеличивается и число оцениваемых параметров, а также усугубляется проблема коллинеарности, что приводит к росту дисперсии оценок.

- По мере увеличения степени полинома  $k$  растет точность аппроксимации, что позволяет снизить смещение оценок. Однако, вместе с ростом  $k$  увеличивается и число оцениваемых параметров, а также усугубляется проблема коллинеарности, что приводит к росту дисперсии оценок.
- Оптимальная степень полинома  $k$ , как правило, подбирается с помощью leave-one-out кросс-валидации.

- По мере увеличения степени полинома  $k$  растет точность аппроксимации, что позволяет снизить смещение оценок. Однако, вместе с ростом  $k$  увеличивается и число оцениваемых параметров, а также усугубляется проблема коллинеарности, что приводит к росту дисперсии оценок.
- Оптимальная степень полинома  $k$ , как правило, подбирается с помощью leave-one-out кросс-валидации.
- Создается  $n$  (объем исходной выборки) выборок объема  $n - 1$ , каждая из которых формируется за счет исключения из исходной выборки одного (каждый раз разного) наблюдения.



- По мере увеличения степени полинома  $k$  растет точность аппроксимации, что позволяет снизить смещение оценок. Однако, вместе с ростом  $k$  увеличивается и число оцениваемых параметров, а также усугубляется проблема коллинеарности, что приводит к росту дисперсии оценок.
- Оптимальная степень полинома  $k$ , как правило, подбирается с помощью leave-one-out кросс-валидации.
- Создается  $n$  (объем исходной выборки) выборок объема  $n - 1$ , каждая из которых формируется за счет исключения из исходной выборки одного (каждый раз разного) наблюдения.
- На каждой из этих выборок при заданном  $k$  методом Ньюи оцениваются параметры модели, а затем с их помощью предсказывается  $\hat{y}_i$  – значение исключенного из выборки наблюдения  $y_i$ .

- По мере увеличения степени полинома  $k$  растет точность аппроксимации, что позволяет снизить смещение оценок. Однако, вместе с ростом  $k$  увеличивается и число оцениваемых параметров, а также усугубляется проблема коллинеарности, что приводит к росту дисперсии оценок.
- Оптимальная степень полинома  $k$ , как правило, подбирается с помощью leave-one-out кросс-валидации.
- Создается  $n$  (объем исходной выборки) выборок объема  $n - 1$ , каждая из которых формируется за счет исключения из исходной выборки одного (каждый раз разного) наблюдения.
- На каждой из этих выборок при заданном  $k$  методом Ньюи оцениваются параметры модели, а затем с их помощью предсказывается  $\hat{y}_i$  – значение исключенного из выборки наблюдения  $y_i$ .
- Рассчитывается  $RMSE_k = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2}$  мера качества модели при данной степени полинома.

- По мере увеличения степени полинома  $k$  растет точность аппроксимации, что позволяет снизить смещение оценок. Однако, вместе с ростом  $k$  увеличивается и число оцениваемых параметров, а также усугубляется проблема коллинеарности, что приводит к росту дисперсии оценок.
- Оптимальная степень полинома  $k$ , как правило, подбирается с помощью leave-one-out кросс-валидации.
- Создается  $n$  (объем исходной выборки) выборок объема  $n - 1$ , каждая из которых формируется за счет исключения из исходной выборки одного (каждый раз разного) наблюдения.
- На каждой из этих выборок при заданном  $k$  методом Ньюи оцениваются параметры модели, а затем с их помощью предсказывается  $\hat{y}_i$  – значение исключенного из выборки наблюдения  $y_i$ .
- Рассчитывается  $RMSE_k = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2}$  мера качества модели при данной степени полинома.
- К счастью существует аналитическая формула, позволяющая рассчитать  $RMSE_k$  без необходимости  $n$  раз оценивать параметры модели.

- По мере увеличения степени полинома  $k$  растет точность аппроксимации, что позволяет снизить смещение оценок. Однако, вместе с ростом  $k$  увеличивается и число оцениваемых параметров, а также усугубляется проблема коллинеарности, что приводит к росту дисперсии оценок.
- Оптимальная степень полинома  $k$ , как правило, подбирается с помощью leave-one-out кросс-валидации.
- Создается  $n$  (объем исходной выборки) выборок объема  $n - 1$ , каждая из которых формируется за счет исключения из исходной выборки одного (каждый раз разного) наблюдения.
- На каждой из этих выборок при заданном  $k$  методом Ньюи оцениваются параметры модели, а затем с их помощью предсказывается  $\hat{y}_i$  – значение исключенного из выборки наблюдения  $y_i$ .
- Рассчитывается  $RMSE_k = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2}$  мера качества модели при данной степени полинома.
- К счастью существует аналитическая формула, позволяющая рассчитать  $RMSE_k$  без необходимости  $n$  раз оценивать параметры модели.
- Выбирается степень  $k$  с наименьшим  $RMSE_k$ .

- По мере увеличения степени полинома  $k$  растет точность аппроксимации, что позволяет снизить смещение оценок. Однако, вместе с ростом  $k$  увеличивается и число оцениваемых параметров, а также усугубляется проблема коллинеарности, что приводит к росту дисперсии оценок.
- Оптимальная степень полинома  $k$ , как правило, подбирается с помощью leave-one-out кросс-валидации.
- Создается  $n$  (объем исходной выборки) выборок объема  $n - 1$ , каждая из которых формируется за счет исключения из исходной выборки одного (каждый раз разного) наблюдения.
- На каждой из этих выборок при заданном  $k$  методом Ньюи оцениваются параметры модели, а затем с их помощью предсказывается  $\hat{y}_i$  – значение исключенного из выборки наблюдения  $y_i$ .
- Рассчитывается  $RMSE_k = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2}$  мера качества модели при данной степени полинома.
- К счастью существует аналитическая формула, позволяющая рассчитать  $RMSE_k$  без необходимости  $n$  раз оценивать параметры модели.
- Выбирается степень  $k$  с наименьшим  $RMSE_k$ .
- При использовании бутстрапа кросс-валидацию необходимо проводить каждую итерацию.

# Модели с неслучайным отбором

## Краткие дополнительные комментарии

- Помимо метода Ньюи существуют и иные подходы к ослаблению допущения о совместном нормальном распределении случайных ошибок в моделях с неслучайным отбором. Например, можно воспользоваться методом Галланта и Нички для аппроксимации соответствующего совместного распределения и получить оценки за счет максимизации функции квази-правдоподобия.

- Помимо метода Ньюи существуют и иные подходы к ослаблению допущения о совместном нормальном распределении случайных ошибок в моделях с неслучайным отбором. Например, можно воспользоваться методом Галланта и Нички для аппроксимации соответствующего совместного распределения и получить оценки за счет максимизации функции квази-правдоподобия.
- Во многих исследованиях рассматриваются альтернативные механизмы неслучайного отбора наблюдений. Например, в качестве уравнения отбора можно использовать мультиномиальную логит модель или порядковую пробит модель. Также, рассматриваются модели с несколькими правилами отбора, когда, например, наблюдения по зарплате доступны лишь для работающих индивидов (первое правило), согласившихся ответить на вопрос о зарплате (второе правило).