Микроэконометрика

Спецификация моделей бинарного выбора

Потанин Богдан Станиславович

старший преподаватель, кандидат экономических наук

2021-2022

Проблемы спецификации

Список основных проблем

- Гетероскедастичность – дисперсии случайных ошибок могут зависеть от регрессоров.

Проблемы спецификации

Список основных проблем

- Гетероскедастичность – дисперсии случайных ошибок могут зависеть от регрессоров.
- Распределение случайной ошибки – может существенно отличаться от нормального или логистического.

Проблемы спецификации

Список основных проблем

- Гетероскедастичность – дисперсии случайных ошибок могут зависеть от регрессоров.
- Распределение случайной ошибки – может существенно отличаться от нормального или логистического.
- Панельные эффекты – при использовании панельных данных следует учитывать фиксированные или случайные эффекты.

Проблемы спецификации

Список основных проблем

- Гетероскедастичность – дисперсии случайных ошибок могут зависеть от регрессоров.
- Распределение случайной ошибки – может существенно отличаться от нормального или логистического.
- Панельные эффекты – при использовании панельных данных следует учитывать фиксированные или случайные эффекты.
- Случайные коэффициенты - влияние различных независимых переменных на вероятность может различаться для каждого наблюдения.

- Предположим, что условная дисперсия случайных ошибок может зависеть от регрессоров, вследствие чего уравнение условной дисперсии принимает вид:
$$Var\left(\varepsilon_i | w_i\right) = \left[g\left(w_i\gamma\right)\right]^2$$
Где $w_i\gamma$ – это линейный индекс уравнения дисперсии, в котором w_i является вектором регрессоров, а γ – вектором коэффициентов. При этом w_i не обязательно совпадает с регрессорами x_i уравнения латентной переменной y_i^*.

- Предположим, что условная дисперсия случайных ошибок может зависеть от регрессоров, вследствие чего уравнение условной дисперсии принимает вид:

$$Var\left(\varepsilon_i | w_i\right) = \left[g\left(w_i\gamma\right)\right]^2$$

Где $w_i\gamma$ – это линейный индекс уравнения дисперсии, в котором w_i является вектором регрессоров, а γ – вектором коэффициентов. При этом w_i не обязательно совпадает с регрессорами x_i уравнения латентной переменной y_i^*.

- В качестве $g(.)$ обычно полагают линейную функцию или экспоненту.

- Предположим, что условная дисперсия случайных ошибок может зависеть от регрессоров, вследствие чего уравнение условной дисперсии принимает вид:
$$Var\left(\varepsilon_i|w_i\right) = \left[g(w_i\gamma)\right]^2$$
Где $w_i\gamma$ – это линейный индекс уравнения дисперсии, в котором w_i является вектором регрессоров, а γ – вектором коэффициентов. При этом w_i не обязательно совпадает с регрессорами x_i уравнения латентной переменной y_i^*.
- В качестве $g(.)$ обычно полагают линейную функцию или экспоненту.
- Из соображений стандартизации обычно также полагают $g(0) = 1$.

- Предположим, что условная дисперсия случайных ошибок может зависеть от регрессоров, вследствие чего уравнение условной дисперсии принимает вид:

$$Var\left(\varepsilon_i|w_i\right) = \left[g(w_i\gamma)\right]^2$$

  Где $w_i\gamma$ – это линейный индекс уравнения дисперсии, в котором w_i является вектором регрессоров, а γ – вектором коэффициентов. При этом w_i не обязательно совпадает с регрессорами x_i уравнения латентной переменной y_i^*.

- В качестве $g(.)$ обычно полагают линейную функцию или экспоненту.
- Из соображений стандартизации обычно также полагают $g(0) = 1$.
- При $\gamma = (0, ..., 0)^T$ получаем $w_i\gamma = 0$, откуда $g(w_i\gamma) = g(0) = 1$, а значит дисперсия не зависит от w_i. Поэтому тестирование гетероскедастичности эквивалентно тестированию гипотезы $H_0 : \gamma = (0, ..., 0)^T$. Ее можно протестировать при помощи LR-теста, LM-теста или теста Вальда.

- Предположим, что условная дисперсия случайных ошибок может зависеть от регрессоров, вследствие чего уравнение условной дисперсии принимает вид:

$$Var\left(\varepsilon_i|w_i\right) = \left[g(w_i\gamma)\right]^2$$

Где $w_i\gamma$ – это линейный индекс уравнения дисперсии, в котором w_i является вектором регрессоров, а γ – вектором коэффициентов. При этом w_i не обязательно совпадает с регрессорами x_i уравнения латентной переменной y_i^*.

- В качестве $g(.)$ обычно полагают линейную функцию или экспоненту.
- Из соображений стандартизации обычно также полагают $g(0) = 1$.
- При $\gamma = (0, ..., 0)^T$ получаем $w_i\gamma = 0$, откуда $g(w_i\gamma) = g(0) = 1$, а значит дисперсия не зависит от w_i. Поэтому тестирование гетероскедастичности эквивалентно тестированию гипотезы $H_0 : \gamma = (0, ..., 0)^T$. Ее можно протестировать при помощи LR-теста, LM-теста или теста Вальда.
- Отсутствие учета гетероскедастичности может привести к несостоятельности оценок вследствие неверной спецификации функции правдоподобия.

- Для простоты далее будем рассматривать гетероскедастичную пробит модель и предположим $g(t) > 0$ при любом $t \in R$.

- Для простоты далее будем рассматривать гетероскедастичную пробит модель и предположим $g(t) > 0$ при любом $t \in R$.
- С учетом гетероскедастичности получаем выражение для условной вероятности:

$$P(y_i = 1|x_i, w_i) = P(\varepsilon_i \geq -x_i\beta) = 1 - \Phi\left(\frac{-x_i\beta}{g(w_i\gamma)}\right) = \Phi\left(\frac{x_i\beta}{g(w_i\gamma)}\right)$$

- Для простоты далее будем рассматривать гетероскедастичную пробит модель и предположим $g(t) > 0$ при любом $t \in R$.
- С учетом гетероскедастичности получаем выражение для условной вероятности:

$$P(y_i = 1|x_i, w_i) = P(\varepsilon_i \geq -x_i\beta) = 1 - \Phi\left(\frac{-x_i\beta}{g(w_i\gamma)}\right) = \Phi\left(\frac{x_i\beta}{g(w_i\gamma)}\right)$$

- Рассмотрим непрерывную переменную v_i, которая входит и в x_i и в w_i с коэффициентами β_v и γ_v соответственно, тогда:

$$\frac{\partial P(y_i = 1|x_i, w_i)}{\partial v_i} = \phi\left(\frac{x_i\beta}{g(w_i\gamma)}\right)\frac{\beta_v g(w_i\gamma) - \gamma_v g'(w_i\gamma)x_i\beta}{g(w_i\gamma)^2}$$

Где $g'(w_i\gamma)$ является производной $g(.)$ в точке $w_i\gamma$.

- Для простоты далее будем рассматривать гетероскедастичную пробит модель и предположим $g(t) > 0$ при любом $t \in R$.
- С учетом гетероскедастичности получаем выражение для условной вероятности:

$$P(y_i = 1|x_i, w_i) = P(\varepsilon_i \geq -x_i\beta) = 1 - \Phi\left(\frac{-x_i\beta}{g(w_i\gamma)}\right) = \Phi\left(\frac{x_i\beta}{g(w_i\gamma)}\right)$$

- Рассмотрим непрерывную переменную v_i, которая входит и в x_i и в w_i с коэффициентами β_v и γ_v соответственно, тогда:

$$\frac{\partial P(y_i = 1|x_i, w_i)}{\partial v_i} = \phi\left(\frac{x_i\beta}{g(w_i\gamma)}\right)\frac{\beta_v g(w_i\gamma) - \gamma_v g'(w_i\gamma)x_i\beta}{g(w_i\gamma)^2}$$

  Где $g'(w_i\gamma)$ является производной $g(.)$ в точке $w_i\gamma$.
- Для того, чтобы получить формулу предельного эффекта для переменной, входящей лишь в x_i или w_i, достаточно предположить равенство нулю β_v или γ_v соответственно.

- По аналогии с классическими моделями бинарного выбора оценивание осуществляется с помощью метода максимального правдоподобия. При этом функция правдоподобия имеет вид:

$$L\left(\beta,\gamma;y|X,W\right)=\prod_{i:y_i=1}\Phi\left(\frac{x_i\beta}{g(w_i\gamma)}\right)\prod_{i:y_i=0}1-\Phi\left(\frac{x_i\beta}{g(w_i\gamma)}\right)$$

- По аналогии с классическими моделями бинарного выбора оценивание осуществляется с помощью метода максимального правдоподобия. При этом функция правдоподобия имеет вид:

$$L\left(\beta,\gamma;y|X,W\right)=\prod_{i:y_i=1}\Phi\left(\frac{x_i\beta}{g(w_i\gamma)}\right)\prod_{i:y_i=0}1-\Phi\left(\frac{x_i\beta}{g(w_i\gamma)}\right)$$

- При некоторых формах $w_i\gamma$ регрессоры w_i не должны включать константу. Например, при использовании наиболее популярной спецификации $g(w_i\gamma)=e^{\gamma_0+w_i\gamma}$ имеем:

$$P(y_i=1|x_i,w_i)=\Phi\left(\frac{x_i\beta}{e^{\gamma_0+w_i\gamma}}\right)=\Phi\left(\frac{x_i\left(\beta/e^{\gamma_0}\right)}{e^{w_i\gamma}}\right)=\Phi\left(\frac{x_i\beta^*}{e^{w_i\gamma}}\right),\text{ где }\beta^*=\beta/e^{\gamma_0}$$

- По аналогии с классическими моделями бинарного выбора оценивание осуществляется с помощью метода максимального правдоподобия. При этом функция правдоподобия имеет вид:

$$L\left(\beta,\gamma;y|X,W\right)=\prod_{i:y_i=1}\Phi\left(\frac{x_i\beta}{g(w_i\gamma)}\right)\prod_{i:y_i=0}1-\Phi\left(\frac{x_i\beta}{g(w_i\gamma)}\right)$$

- При некоторых формах $w_i\gamma$ регрессоры w_i не должны включать константу. Например, при использовании наиболее популярной спецификации $g(w_i\gamma)=e^{\gamma_0+w_i\gamma}$ имеем:

$$P(y_i=1|x_i,w_i)=\Phi\left(\frac{x_i\beta}{e^{\gamma_0+w_i\gamma}}\right)=\Phi\left(\frac{x_i\left(\beta/e^{\gamma_0}\right)}{e^{w_i\gamma}}\right)=\Phi\left(\frac{x_i\beta^*}{e^{w_i\gamma}}\right)\text{, где }\beta^*=\beta/e^{\gamma_0}$$

  В результате мы можем заменить β/e^{γ_0} на β^* при максимизации функции правдоподобия. Но каждому значению β^* соответстует бесконечное число комбинаций параметров β и γ_0. В результате β и γ_0 оказываются неидентифицируемы, а функция правдоподобия – имеет бесконечное число максимумов при наличии γ_0. Во избежание данной проблемы без потери общности полагают $\gamma_0=0$.

Модели бинарного выбора с гетероскедастичностью

Тестирование гетероскедастичности при помощи LM-теста, часть 1

- Рассмотрим тестирование гипотезы об отсутствии гетероскедастичности $H_0 : \gamma = (0, ..., 0)^T$ с помощью LM-теста.

- Рассмотрим тестирование гипотезы об отсутствии гетероскедастичности $H_0 : \gamma = (0, ..., 0)^T$ с помощью LM-теста.
- Производная логарифма правдоподобия по γ_k имеет вид:

$$\frac{\partial \ln L}{\partial \gamma_k} = \sum_{i:y_i=1} \frac{\phi\left(\frac{x_i\beta}{g(w_i\gamma)}\right)}{\Phi\left(\frac{x_i\beta}{g(w_i\gamma)}\right)} \frac{-w_{ki}g'(w_i\gamma)x_i\beta}{g(w_i\gamma)^2} + \sum_{i:y_i=0} \frac{\phi\left(\frac{x_i\beta}{g(w_i\gamma)}\right)}{1-\Phi\left(\frac{x_i\beta}{g(w_i\gamma)}\right)} \frac{x_{wi}g'(w_i\gamma)x_i\beta}{g(w_i\gamma)^2}$$

- Рассмотрим тестирование гипотезы об отсутствии гетероскедастичности $H_0 : \gamma = (0, ..., 0)^T$ с помощью LM-теста.
- Производная логарифма правдоподобия по γ_k имеет вид:

$$\frac{\partial \ln L}{\partial \gamma_k} = \sum_{i:y_i=1} \frac{\phi\left(\frac{x_i\beta}{g(w_i\gamma)}\right)}{\Phi\left(\frac{x_i\beta}{g(w_i\gamma)}\right)} \frac{-w_{ki}g'(w_i\gamma)x_i\beta}{g(w_i\gamma)^2} + \sum_{i:y_i=0} \frac{\phi\left(\frac{x_i\beta}{g(w_i\gamma)}\right)}{1-\Phi\left(\frac{x_i\beta}{g(w_i\gamma)}\right)} \frac{x_{wi}g'(w_i\gamma)x_i\beta}{g(w_i\gamma)^2}$$

- Для ограниченной модели $\gamma = (0, ..., 0)^T$ получаем:

$$\frac{\partial \ln L\left(\beta, 0....0; y|X, W\right)}{\partial \gamma_k} = \sum_{i:y_i=1} \frac{\phi\left(x_i\beta\right)}{\Phi\left(x_i\beta\right)}\left(-w_{ki}g'(0)x_i\beta\right) + \sum_{i:y_i=0} \frac{\phi\left(x_i\beta\right)}{1-\Phi\left(x_i\beta\right)} w_{ki}g'(0)x_i\beta =$$

- Рассмотрим тестирование гипотезы об отсутствии гетероскедастичности $H_0 : \gamma = (0, ..., 0)^T$ с помощью LM-теста.
- Производная логарифма правдоподобия по γ_k имеет вид:

$$\frac{\partial \ln L}{\partial \gamma_k} = \sum_{i:y_i=1} \frac{\phi\left(\frac{x_i\beta}{g(w_i\gamma)}\right)}{\Phi\left(\frac{x_i\beta}{g(w_i\gamma)}\right)} \frac{-w_{ki}g'(w_i\gamma)x_i\beta}{g(w_i\gamma)^2} + \sum_{i:y_i=0} \frac{\phi\left(\frac{x_i\beta}{g(w_i\gamma)}\right)}{1-\Phi\left(\frac{x_i\beta}{g(w_i\gamma)}\right)} \frac{x_{wi}g'(w_i\gamma)x_i\beta}{g(w_i\gamma)^2}$$

- Для ограниченной модели $\gamma = (0, ..., 0)^T$ получаем:

$$\frac{\partial \ln L\left(\beta, 0....0; y|X, W\right)}{\partial \gamma_k} = \sum_{i:y_i=1} \frac{\phi\left(x_i\beta\right)}{\Phi\left(x_i\beta\right)} \left(-w_{ki}g'(0)x_i\beta\right) + \sum_{i:y_i=0} \frac{\phi\left(x_i\beta\right)}{1-\Phi\left(x_i\beta\right)} w_{ki}g'(0)x_i\beta =$$

$$= g'(0) \sum_{i=1}^{n} \frac{\phi\left(x_i\beta\right)}{\Phi\left(\left(2y_i-1\right)x_i\beta\right)} w_{ki}\left(1-2y_i\right)x_i\beta$$

- Рассмотрим тестирование гипотезы об отсутствии гетероскедастичности $H_0 : \gamma = (0, ..., 0)^T$ с помощью LM-теста.
- Производная логарифма правдоподобия по γ_k имеет вид:

$$\frac{\partial \ln L}{\partial \gamma_k} = \sum_{i:y_i=1} \frac{\phi\left(\frac{x_i\beta}{g(w_i\gamma)}\right)}{\Phi\left(\frac{x_i\beta}{g(w_i\gamma)}\right)} \frac{-w_{ki}g'(w_i\gamma)x_i\beta}{g(w_i\gamma)^2} + \sum_{i:y_i=0} \frac{\phi\left(\frac{x_i\beta}{g(w_i\gamma)}\right)}{1-\Phi\left(\frac{x_i\beta}{g(w_i\gamma)}\right)} \frac{x_{wi}g'(w_i\gamma)x_i\beta}{g(w_i\gamma)^2}$$

- Для ограниченной модели $\gamma = (0, ..., 0)^T$ получаем:

$$\frac{\partial \ln L(\beta, 0....0; y|X, W)}{\partial \gamma_k} = \sum_{i:y_i=1} \frac{\phi(x_i\beta)}{\Phi(x_i\beta)}(-w_{ki}g'(0)x_i\beta) + \sum_{i:y_i=0} \frac{\phi(x_i\beta)}{1-\Phi(x_i\beta)} w_{ki}g'(0)x_i\beta =$$

$$= g'(0)\sum_{i=1}^{n} \frac{\phi(x_i\beta)}{\Phi((2y_i-1)x_i\beta)} w_{ki}(1-2y_i)x_i\beta$$

- Поскольку частные производные отличаются лишь множителями x_{wi}, то градиент принимает вид:

$$\nabla_\gamma \ln L(\beta, 0....0; y|X, W) = g'(0)\sum_{i=1}^{n} \frac{\phi(x_i\beta)}{\Phi((2y_i-1)x_i\beta)} w_i^T(1-2y_i)x_i\beta$$

- По аналогии нетрудно показать, что:

$$\nabla_{\beta} \ln L\left(\beta, 0....0; y|X, W\right) = g'(0)\sum_{i=1}^{n}\frac{\phi\left(x_i\beta\right)}{\Phi\left(\left(2y_i-1\right)x_i\beta\right)}x_i^T\left(2y_i-1\right)$$

- По аналогии нетрудно показать, что:

$$\nabla_\beta \ln L\left(\beta, 0\ldots0; y|X, W\right) = g'(0)\sum_{i=1}^{n}\frac{\phi\left(x_i\beta\right)}{\Phi\left(\left(2y_i-1\right)x_i\beta\right)}x_i^T\left(2y_i-1\right)$$

- В результате получаем:

$$\nabla \ln L_F = \begin{bmatrix}\nabla_\beta L_F\\ \nabla_\gamma L_F\end{bmatrix},\text{ где } \ln L_F = \ln L\left(\hat{\beta}, 0\ldots0; y|X, W\right),$$

где $\hat{\beta}$ является оценкой обычной (ограниченной) пробит модели, откуда $\nabla_\beta L_F = (0, ..., 0)$.

- По аналогии нетрудно показать, что:

$$\nabla_\beta \ln L\left(\beta, 0....0; y|X, W\right) = g'(0)\sum_{i=1}^{n}\frac{\phi\left(x_i\beta\right)}{\Phi\left(\left(2y_i-1\right)x_i\beta\right)}x_i^T\left(2y_i-1\right)$$

- В результате получаем:

$$\nabla \ln L_F = \begin{bmatrix}\nabla_\beta L_F\\ \nabla_\gamma L_F\end{bmatrix}, \text{ где } \ln L_F = \ln L\left(\hat{\beta}, 0....0; y|X, W\right),$$

где $\hat{\beta}$ является оценкой обычной (ограниченной) пробит модели, откуда $\nabla_\beta L_F = (0, ..., 0)$.

- В итоге предполагая $g'(0)\ne 0$ нетрудно записать тестовую статистику LM-теста в виде (для расчета $\widehat{As.Cov}_F$ воспользуемся произведением якобианов), при котором она не зависит от $g'(0)$:

$$\left(\nabla \ln L_F\right)^T\left(J\left(L_F\right)^T J\left(L_F\right)\right)^{-1}\ln L_F =$$

- По аналогии нетрудно показать, что:

$$\nabla_\beta \ln L\left(\beta, 0....0; y|X, W\right) = g'(0)\sum_{i=1}^{n}\frac{\phi\left(x_i\beta\right)}{\Phi\left(\left(2y_i-1\right)x_i\beta\right)}x_i^T\left(2y_i-1\right)$$

- В результате получаем:

$$\nabla \ln L_F = \begin{bmatrix}\nabla_\beta L_F\\ \nabla_\gamma L_F\end{bmatrix},\text{ где } \ln L_F = \ln L\left(\hat{\beta}, 0....0; y|X, W\right),$$

где $\hat{\beta}$ является оценкой обычной (ограниченной) пробит модели, откуда $\nabla_\beta L_F = (0, ..., 0)$.

- В итоге предполагая $g'(0)\neq 0$ нетрудно записать тестовую статистику LM-теста в виде (для расчета $\widehat{As.Cov}_F$ воспользуемся произведением якобианов), при котором она не зависит от $g'(0)$:

$$\left(\nabla \ln L_F\right)^T\left(J\left(L_F\right)^T J\left(L_F\right)\right)^{-1}\ln L_F = \left(\ln L_F/g'(0)\right)^T\left(\left[J\left(L_F\right)^T/g'(0)\right]\left[J\left(L_F\right)/g'(0)\right]\right)^{-1}\ln L_F/g'(0)$$

- По аналогии нетрудно показать, что:

$$\nabla_{\beta} \ln L\left(\beta, 0....0; y|X, W\right) = g'(0) \sum_{i=1}^{n} \frac{\phi\left(x_i\beta\right)}{\Phi\left(\left(2y_i - 1\right) x_i\beta\right)} x_i^T \left(2y_i - 1\right)$$

- В результате получаем:

$$\nabla \ln L_F = \begin{bmatrix} \nabla_{\beta} L_F \\ \nabla_{\gamma} L_F \end{bmatrix}, \text{ где } \ln L_F = \ln L\left(\hat{\beta}, 0....0; y|X, W\right),$$

где $\hat{\beta}$ является оценкой обычной (ограниченной) пробит модели, откуда $\nabla_{\beta} L_F = (0, ..., 0)$.

- В итоге предполагая $g'(0) \neq 0$ нетрудно записать тестовую статистику LM-теста в виде (для расчета $\widehat{As.Cov}_F$ воспользуемся произведением якобианов), при котором она не зависит от $g'(0)$:

$$\left(\nabla \ln L_F\right)^T \left(J\left(L_F\right)^T J\left(L_F\right)\right)^{-1} \ln L_F = \left(\ln L_F / g'(0)\right)^T \left(\left[J\left(L_F\right)^T / g'(0)\right] \left[J\left(L_F\right) / g'(0)\right]\right)^{-1} \ln L_F / g'(0)$$

- Таким образом, значение тестовой статистики не зависит от формы функции $g(.)$. Следовательно, благодаря LM-тесту мы тестируем гипотезу об отсутствии гетероскедастичности против альтернативы о том, что она определяется широким классом функций вида $g(.)$.

Тестирование гипотезы о распределении случайной ошибки

Общая идея

- Некорректная спецификация распределении случайной ошибки ε_i в модели бинарного выбора может привести к несостоятельности оценок.

Тестирование гипотезы о распределении случайной ошибки

Общая идея

- Некорректная спецификация распределении случайной ошибки ε_i в модели бинарного выбора может привести к несостоятельности оценок.
- На практике распределение случайной ошибки ε_i может отличаться от нормального или логистического. Для простоты рассмотрим подходы к тестированию гипотезы о стандартном нормальном распределении случайных ошибок $\varepsilon_i \sim \mathcal{N}(0, 1)$.

Тестирование гипотезы о распределении случайной ошибки

Общая идея

- Некорректная спецификация распределении случайной ошибки ε_i в модели бинарного выбора может привести к несостоятельности оценок.
- На практике распределение случайной ошибки ε_i может отличаться от нормального или логистического. Для простоты рассмотрим подходы к тестированию гипотезы о стандартном нормальном распределении случайных ошибок $\varepsilon_i \sim \mathcal{N}(0,1)$.
- Предполагается, что случайные ошибки в полной модели распределены в соответствии с некоторым распределением Θ с вектором параметров θ, то есть в полной модели $\varepsilon_i \sim \Theta(\theta)$. Кроме того, при некотором $\theta = \theta^0$ распределение Θ совпадает со стандартным нормальным $\mathcal{N}(0,1)$, то есть $\Theta(\theta^0) = \mathcal{N}(0,1)$. В результате гипотеза о нормальности принимает параметрический вид $H_0 : \theta = \theta_0$.

- Некорректная спецификация распределении случайной ошибки ε_i в модели бинарного выбора может привести к несостоятельности оценок.
- На практике распределение случайной ошибки ε_i может отличаться от нормального или логистического. Для простоты рассмотрим подходы к тестированию гипотезы о стандартном нормальном распределении случайных ошибок $\varepsilon_i \sim \mathcal{N}(0,1)$.
- Предполагается, что случайные ошибки в полной модели распределены в соответствии с некоторым распределением Θ с вектором параметров θ, то есть в полной модели $\varepsilon_i \sim \Theta(\theta)$. Кроме того, при некотором $\theta = \theta^0$ распределение Θ совпадает со стандартным нормальным $\mathcal{N}(0,1)$, то есть $\Theta(\theta^0) = \mathcal{N}(0,1)$. В результате гипотеза о нормальности принимает параметрический вид $H_0 : \theta = \theta_0$.
- Далее полная модель может быть оценена при помощи ММП, где θ выступает в качестве дополнительного вектора оцениваемых (вместе с β) параметров. Для тестирования гипотезы можно воспользоваться LR-тестом, LM-тестом или тестом Вальда. Обычно наиболее удобным является LM-тест, поскольку позволяет не оценивать полную модель, что при сложных формах Θ бывает технически затруднительным.

- В качестве альтернативы стандартному нормальному распределению рассмотрим распрделение из семейства кривых Пирсона с параметрами θ_1 и θ_2, функция распределения которого имеет вид:

$$\Phi(t + \theta_1 t^2 + \theta_2 t^3)$$

- В качестве альтернативы стандартному нормальному распределению рассмотрим распрделение из семейства кривых Пирсона с параметрами θ_1 и θ_2, функция распределения которого имеет вид:

$$\Phi(t + \theta_1 t^2 + \theta_2 t^3)$$

- При $\theta_1 = \theta_2 = 0$ данное распределение совпадает со стандартным нормальным, поэтому тестирование гипотезы о корректности применения пробит модели эквивалентно проверке параметрической гипотезы $H_0 : \theta_1 = \theta_2 = 0$.

- В качестве альтернативы стандартному нормальному распределению рассмотрим распрделение из семейства кривых Пирсона с параметрами θ_1 и θ_2, функция распределения которого имеет вид:

$$\Phi(t + \theta_1 t^2 + \theta_2 t^3)$$

- При $\theta_1 = \theta_2 = 0$ данное распределение совпадает со стандартным нормальным, поэтому тестирование гипотезы о корректности применения пробит модели эквивалентно проверке параметрической гипотезы $H_0 : \theta_1 = \theta_2 = 0$.
- Оценивание полной модели предполагает максимизацию функции правдоподобия по параметрам β, θ_1 и θ_2 с использованием $\Phi(x_i\beta + \theta_1 x_i\beta^2 + \theta_2 x_i\beta^3)$ вместо $\Phi(x_i\beta)$.

- В качестве альтернативы стандартному нормальному распределению рассмотрим распрделение из семейства кривых Пирсона с параметрами θ_1 и θ_2, функция распределения которого имеет вид:

$$\Phi(t + \theta_1 t^2 + \theta_2 t^3)$$

- При $\theta_1 = \theta_2 = 0$ данное распределение совпадает со стандартным нормальным, поэтому тестирование гипотезы о корректности применения пробит модели эквивалентно проверке параметрической гипотезы $H_0 : \theta_1 = \theta_2 = 0$.
- Оценивание полной модели предполагает максимизацию функции правдоподобия по параметрам β, θ_1 и θ_2 с использованием $\Phi(x_i\beta + \theta_1 x_i\beta^2 + \theta_2 x_i\beta^3)$ вместо $\Phi(x_i\beta)$.
- Поскольку оценивание полной модели является достаточно затруднительной технической задачей, то для проверки гипотезы о стандартном нормальном распределении случайной ошибки предпочтительно воспользоваться LM-тестом.

- В качестве альтернативы стандартному нормальному распределению рассмотрим распрделение из семейства кривых Пирсона с параметрами θ_1 и θ_2, функция распределения которого имеет вид:

$$\Phi(t+\theta_1 t^2+\theta_2 t^3)$$

- При $\theta_1=\theta_2=0$ данное распределение совпадает со стандартным нормальным, поэтому тестирование гипотезы о корректности применения пробит модели эквивалентно проверке параметрической гипотезы $H_0:\theta_1=\theta_2=0$.
- Оценивание полной модели предполагает максимизацию функции правдоподобия по параметрам β, θ_1 и θ_2 с использованием $\Phi(x_i\beta+\theta_1 x_i\beta^2+\theta_2 x_i\beta^3)$ вместо $\Phi(x_i\beta)$.
- Поскольку оценивание полной модели является достаточно затруднительной технической задачей, то для проверки гипотезы о стандартном нормальном распределении случайной ошибки предпочтительно воспользоваться LM-тестом.
- Недостаток данного подхода заключается в рассмотрении достаточно узкого класса распределений в качестве альтернативы стандартному нормальному.

- Галлант и Ничка предложили распределение со следующими функцией плотности:

$$f_{GN}(t;\theta) = \frac{(1+\theta_1 t + ... + \theta_K t^K)^2}{\sum\limits_{i=0}^{K}\sum\limits_{j=0}^{K}\theta_i\theta_j M(i+j)}\phi(t),$$

где через $M(k)$ обозначается k-й начальный момент стандартного нормального распределения.

- Галлант и Ничка предложили распределение со следующими функцией плотности:

$$f_{GN}(t;\theta)=\frac{(1+\theta_1 t+...+\theta_K t^K)^2}{\sum\limits_{i=0}^{K}\sum\limits_{j=0}^{K}\theta_i\theta_j M(i+j)}\phi(t),$$

  где через $M(k)$ обозначается k-й начальный момент стандартного нормального распределения.
- При $\theta_1=...=\theta_K=0$ данное распределение совпадает со стандартным нормальным, поэтому тестирование гипотезы о нормальности эквивалентно тестированию соблюдения данного ограничения на парамеры.

- Галлант и Ничка предложили распределение со следующими функцией плотности:

$$f_{GN}(t;\theta) = \frac{(1+\theta_1 t + ... + \theta_K t^K)^2}{\sum_{i=0}^{K}\sum_{j=0}^{K}\theta_i\theta_j M(i+j)}\phi(t),$$

  где через $M(k)$ обозначается k-й начальный момент стандартного нормального распределения.
- При $\theta_1 = ... = \theta_K = 0$ данное распределение совпадает со стандартным нормальным, поэтому тестирование гипотезы о нормальности эквивалентно тестированию соблюдения данного ограничения на парамеры.
- За счет полинома в числителе данное распределение может аппроксимировать широкий класс распределений. Чем больше K, тем более широкий класс распределений можно достаточно точно аппроксимировать.

- Галлант и Ничка предложили распределение со следующими функцией плотности:

$$f_{GN}(t;\theta)=\frac{(1+\theta_1 t+...+\theta_K t^K)^2}{\sum\limits_{i=0}^{K}\sum\limits_{j=0}^{K}\theta_i\theta_j M(i+j)}\phi(t),$$

где через $M(k)$ обозначается k-й начальный момент стандартного нормального распределения.

- При $\theta_1=...=\theta_K=0$ данное распределение совпадает со стандартным нормальным, поэтому тестирование гипотезы о нормальности эквивалентно тестированию соблюдения данного ограничения на парамеры.
- За счет полинома в числителе данное распределение может аппроксимировать широкий класс распределений. Чем больше K, тем более широкий класс распределений можно достаточно точно аппроксимировать.
- Функция распределения и моменты распределения Галланта и Нички считаются как:

$$F_{GN}(t;\theta)=\frac{\sum\limits_{i=0}^{K}\sum\limits_{j=0}^{K}\theta_i\theta_j M(i+j;t)}{\sum\limits_{i=0}^{K}\sum\limits_{j=0}^{K}\theta_i\theta_j M(i+j)}\Phi(t),\qquad M_{GN}(k;\theta)=\frac{\sum\limits_{i=0}^{K}\sum\limits_{j=0}^{K}\theta_i\theta_j M(i+j+k)}{\sum\limits_{i=0}^{K}\sum\limits_{j=0}^{K}\theta_i\theta_j M(i+j)},$$

где через $M(k;t)$ обозначается k-й начальный момент стандартного нормального распределения, усеченного сверху в точке t.

- Распределение Галланта и Нички часто используют для того, что аппроксимировать неизвестное распределение случайной ошибки.

- Распределение Галланта и Нички часто используют для того, что аппроксимировать неизвестное распределение случайной ошибки.
- При различных параметрах θ данное распределение имеет различные математическое ожидание и дисперсию, что может вызвать проблемы с идентифицируемостью β в моделях бинарного выбора.

- Распределение Галланта и Нички часто используют для того, что аппроксимировать неизвестное распределение случайной ошибки.
- При различных параметрах θ данное распределение имеет различные математическое ожидание и дисперсию, что может вызвать проблемы с идентифицируемостью β в моделях бинарного выбора.
- Во избежание данных проблем распределение галланта и Нички удобно стандартизировать к нулевому математическому ожиданию и единичной дисперсии.

- Распределение Галланта и Ничкi часто используют для того, что аппроксимировать неизвестное распределение случайной ошибки.
- При различных параметрах θ данное распределение имеет различные математическое ожидание и дисперсию, что может вызвать проблемы с идентифицируемостью β в моделях бинарного выбора.
- Во избежание данных проблем распределение галланта и Ничкi удобно стандартизировать к нулевому математическому ожиданию и единичной дисперсии.
- Обозначим через Z случайную величину, имеющую распределение Галланта и Ничкi с вектором параметров θ. Тогда функция распределения случайной величины $Z^* = \frac{Z-E(Z)}{\sqrt{Var(Z)}}$, имеющей стандартизированое распределение Галланта и Ничкi GN^*, будет иметь вид:

$$F_{GN^*}(t;\theta) = F_{Z^*}(t;\theta) = P(Z^* \leq t) = P\left(\frac{Z-E(Z)}{\sqrt{Var(Z)}} \leq t\right) =$$

Полупараметрические модели бинарного выбора

Стандартизация распределения Галланта и Нички

- Распределение Галланта и Нички часто используют для того, что аппроксимировать неизвестное распределение случайной ошибки.
- При различных параметрах θ данное распределение имеет различные математическое ожидание и дисперсию, что может вызвать проблемы с идентифицируемостью β в моделях бинарного выбора.
- Во избежание данных проблем распределение галланта и Нички удобно стандартизировать к нулевому математическому ожиданию и единичной дисперсии.
- Обозначим через Z случайную величину, имеющую распределение Галланта и Нички с вектором параметров θ. Тогда функция распределения случайной величины $Z^* = \frac{Z-E(Z)}{\sqrt{Var(Z)}}$, имеющей стандартизированое распределение Галланта и Нички GN^*, будет иметь вид:

$$F_{GN^*}(t;\theta) = F_{Z^*}(t;\theta) = P(Z^* \leq t) = P\left(\frac{Z-E(Z)}{\sqrt{Var(Z)}} \leq t\right) =$$

$$= P\left(Z < \sqrt{Var(Z)}t + E(Z)\right) = P\left(Z < \sqrt{M_{GN}(2;\theta) - M_{GN}(1;\theta)^2}t + M_{GN}(1;\theta)\right) =$$

- Распределение Галланта и Нички часто используют для того, что аппроксимировать неизвестное распределение случайной ошибки.
- При различных параметрах θ данное распределение имеет различные математическое ожидание и дисперсию, что может вызвать проблемы с идентифицируемостью β в моделях бинарного выбора.
- Во избежание данных проблем распределение галланта и Нички удобно стандартизировать к нулевому математическому ожиданию и единичной дисперсии.
- Обозначим через Z случайную величину, имеющую распределение Галланта и Нички с вектором параметров θ. Тогда функция распределения случайной величины $Z^* = \frac{Z-E(Z)}{\sqrt{Var(Z)}}$, имеющей стандартизированое распределение Галланта и Нички GN^*, будет иметь вид:

$$F_{GN^*}(t;\theta) = F_{Z^*}(t;\theta) = P(Z^* \leq t) = P\left(\frac{Z-E(Z)}{\sqrt{Var(Z)}} \leq t\right) =$$

$$= P\left(Z < \sqrt{Var(Z)}t + E(Z)\right) = P\left(Z < \sqrt{M_{GN}(2;\theta) - M_{GN}(1;\theta)^2}t + M_{GN}(1;\theta)\right) =$$

$$= F_Z\left(\sqrt{M_{GN}(2;\theta) - M_{GN}(1;\theta)^2}t + M_{GN}(1;\theta);\theta\right) = F_{GN}\left(\sqrt{M_{GN}(2;\theta) - M_{GN}(1;\theta)^2}t + M(1;\theta);\theta\right)$$

- Предположим, что ε_i имеет стандартизированное распределение Галланта и Нички GN^*, откуда:

$$P(y_i = 1|x_i) = 1 - F_{GN^*}(-x_i\beta;\theta)$$

- Предположим, что ε_i имеет стандартизированное распределение Галланта и Нички GN^*, откуда:

$$P(y_i = 1|x_i) = 1 - F_{GN^*}(-x_i\beta;\theta)$$

- Предполагая, что распределение ε_i может быть достаточно точно приближено стандартизированным распределением Галланта и Нички можно оценить β и θ за счет максимизации функции квази (псевдо) правдоподобия:

$$L(\beta,\theta;y|X) = \prod_{i:y_i=1} 1 - F_{GN^*}(-x_i\beta;\theta) \prod_{i:y_i=0} F_{GN^*}(-x_i\beta;\theta)$$

При этом степень полинома K может быть подобрана исходя из минимизации информационного критерия Акаике.

- Предположим, что ε_i имеет стандартизированное распределение Галланта и Нички GN^*, откуда:

$$P(y_i = 1|x_i) = 1 - F_{GN^*}(-x_i\beta;\theta)$$

- Предполагая, что распределение ε_i может быть достаточно точно приближено стандартизированным распределением Галланта и Нички можно оценить β и θ за счет максимизации функции квази (псевдо) правдоподобия:

$$L(\beta,\theta;y|X) = \prod_{i:y_i=1} 1 - F_{GN^*}(-x_i\beta;\theta) \prod_{i:y_i=0} F_{GN^*}(-x_i\beta;\theta)$$

  При этом степень полинома K может быть подобрана исходя из минимизации информационного критерия Акаике.

- Приставка квази (псевдо) означает, что $F_{GN^*}(-x_i\beta;\theta)$ отражает не истинное распределение случайной ошибки, а функцию, при помощи которой оно аппроксимируется. Поэтому для оценивания асимптотической ковариационной матрицы необходимо использовать сэндвич.

- Предположим, что ε_i имеет стандартизированное распределение Галланта и Нички GN^*, откуда:

$$P(y_i = 1|x_i) = 1 - F_{GN^*}(-x_i\beta; \theta)$$

- Предполагая, что распределение ε_i может быть достаточно точно приближено стандартизированным распределением Галланта и Нички можно оценить β и θ за счет максимизации функции квази (псевдо) правдоподобия:

$$L(\beta, \theta; y|X) = \prod_{i:y_i=1} 1 - F_{GN^*}(-x_i\beta; \theta) \prod_{i:y_i=0} F_{GN^*}(-x_i\beta; \theta)$$

  При этом степень полинома K может быть подобрана исходя из минимизации информационного критерия Акаике.

- Приставка квази (псевдо) означает, что $F_{GN^*}(-x_i\beta; \theta)$ отражает не истинное распределение случайной ошибки, а функцию, при помощи которой оно аппроксимируется. Поэтому для оценивания асимптотической ковариационной матрицы необходимо использовать сэндвич.
- Данный подход позволяет ослабить допущение о конкретной форме распределения случайной ошибки, однако требует существенных вычислительных мощностей для поиска глобального максимума функции квази правдоподобия.

- Через j обозначим индекс группы, к которой принадлежит наблюдение в панельной выборке, а через $i \in \{1, ..., T\}$ – индекс данного наблюдения в группе. Например, индекс j может отражать название фирмы, а индекс i – год наблюдения по этой фирме.

- Через j обозначим индекс группы, к которой принадлежит наблюдение в панельной выборке, а через $i \in \{1, ..., T\}$ – индекс данного наблюдения в группе. Например, индекс j может отражать название фирмы, а индекс i – год наблюдения по этой фирме.
- При наличии случайных эффектов случайная ошибка специфицируется как:

$$\varepsilon_{ij} = \varepsilon_{ij}^{(1)} + \varepsilon_{j}^{(2)}$$

Где $\varepsilon_{j}^{(2)}$ именуется случайным эффектом. Предполагается, что $\varepsilon_{j}^{(2)}$ и $\varepsilon_{ij}^{(1)}$ независимы и не зависят от x_{ij}. Допущение о независимости $\varepsilon_{j}^{(2)}$ и x_{ij} часто называется слабым место модели.

- Через j обозначим индекс группы, к которой принадлежит наблюдение в панельной выборке, а через $i \in \{1, ..., T\}$ – индекс данного наблюдения в группе. Например, индекс j может отражать название фирмы, а индекс i – год наблюдения по этой фирме.
- При наличии случайных эффектов случайная ошибка специфицируется как:

$$\varepsilon_{ij} = \varepsilon_{ij}^{(1)} + \varepsilon_{j}^{(2)}$$

Где $\varepsilon_{j}^{(2)}$ именуется случайным эффектом. Предполагается, что $\varepsilon_{j}^{(2)}$ и $\varepsilon_{ij}^{(1)}$ независимы и не зависят от x_{ij}. Допущение о независимости $\varepsilon_{j}^{(2)}$ и x_{ij} часто называется слабым место модели.

- Без потери общности предполагается, что $E(\varepsilon_{ij}^{(1)}|x_{ij}) = E(\varepsilon_{j}^{(2)}|x_{ij}) = 0$, а также $Var(\varepsilon_{ij}^{(1)}|x_{ij}) = 1$ и $Var(\varepsilon_{j}^{(2)}|x_{ij}) = \sigma^2$.

- Через j обозначим индекс группы, к которой принадлежит наблюдение в панельной выборке, а через $i \in \{1, ..., T\}$ – индекс данного наблюдения в группе. Например, индекс j может отражать название фирмы, а индекс i – год наблюдения по этой фирме.
- При наличии случайных эффектов случайная ошибка специфицируется как:

$$\varepsilon_{ij} = \varepsilon_{ij}^{(1)} + \varepsilon_{j}^{(2)}$$

Где $\varepsilon_{j}^{(2)}$ именуется случайным эффектом. Предполагается, что $\varepsilon_{j}^{(2)}$ и $\varepsilon_{ij}^{(1)}$ независимы и не зависят от x_{ij}. Допущение о независимости $\varepsilon_{j}^{(2)}$ и x_{ij} часто называется слабым место модели.

- Без потери общности предполагается, что $E(\varepsilon_{ij}^{(1)}|x_{ij}) = E(\varepsilon_{j}^{(2)}|x_{ij}) = 0$, а также $Var(\varepsilon_{ij}^{(1)}|x_{ij}) = 1$ и $Var(\varepsilon_{j}^{(2)}|x_{ij}) = \sigma^2$.
- Наличие случайных эффектов в модели можно протестировать как $H_0 : \sigma^2 = 0$.

Модели бинарного выбора для панельных данных

Расчет вероятностей

- Обратим внимание, что (используя law of total expectation):

$$P(y_{ij} = 1|x_{ij}) = P(\varepsilon_{ij}^{(1)} + \varepsilon_{j}^{(2)} > -x_{ij}\beta|x_{ij}) = P(\varepsilon_{ij}^{(1)} > -\varepsilon_{j}^{(2)} - x_{ij}\beta|x_{ij}) =$$

Модели бинарного выбора для панельных данных

Расчет вероятностей

- Обратим внимание, что (используя law of total expectation):

$$P(y_{ij}=1|x_{ij})=P(\varepsilon_{ij}^{(1)}+\varepsilon_{j}^{(2)}>-x_{ij}\beta|x_{ij})=P(\varepsilon_{ij}^{(1)}>-\varepsilon_{j}^{(2)}-x_{ij}\beta|x_{ij})=$$
$$=E\left(P(\varepsilon_{ij}^{(1)}>-\varepsilon_{j}^{(2)}-x_{ij}\beta|x_{ij})\right)=E_{\varepsilon_j^{(2)}}\left(E\left[P\left(\varepsilon_{ij}^{(1)}>-\varepsilon_{j}^{(2)}-x_{ij}\beta|x_{ij},\varepsilon_{j}^{(2)}\right)\right]\right)=$$

- Обратим внимание, что (используя law of total expectation):

$$P(y_{ij}=1|x_{ij})=P(\varepsilon_{ij}^{(1)}+\varepsilon_{j}^{(2)}>-x_{ij}\beta|x_{ij})=P(\varepsilon_{ij}^{(1)}>-\varepsilon_{j}^{(2)}-x_{ij}\beta|x_{ij})=$$
$$=E\left(P(\varepsilon_{ij}^{(1)}>-\varepsilon_{j}^{(2)}-x_{ij}\beta|x_{ij})\right)=E_{\varepsilon_{j}^{(2)}}\left(E\left[P\left(\varepsilon_{ij}^{(1)}>-\varepsilon_{j}^{(2)}-x_{ij}\beta|x_{ij},\varepsilon_{j}^{(2)}\right)\right]\right)=$$
$$=E_{\varepsilon_{j}^{(2)}}\left(P\left(\varepsilon_{ij}^{(1)}>-\varepsilon_{j}^{(2)}-x_{ij}\beta|x_{ij}\right)\right)=E_{\varepsilon_{j}^{(2)}}\left(1-F_{\varepsilon_{ij}^{(1)}}\left(-x_{ij}\beta-\varepsilon_{j}^{(2)}\right)\right)$$

- Обратим внимание, что (используя law of total expectation):

$$P(y_{ij}=1|x_{ij})=P(\varepsilon_{ij}^{(1)}+\varepsilon_{j}^{(2)}>-x_{ij}\beta|x_{ij})=P(\varepsilon_{ij}^{(1)}>-\varepsilon_{j}^{(2)}-x_{ij}\beta|x_{ij})=$$
$$=E\left(P(\varepsilon_{ij}^{(1)}>-\varepsilon_{j}^{(2)}-x_{ij}\beta|x_{ij})\right)=E_{\varepsilon_{j}^{(2)}}\left(E\left[P\left(\varepsilon_{ij}^{(1)}>-\varepsilon_{j}^{(2)}-x_{ij}\beta|x_{ij},\varepsilon_{j}^{(2)}\right)\right]\right)=$$
$$=E_{\varepsilon_{j}^{(2)}}\left(P\left(\varepsilon_{ij}^{(1)}>-\varepsilon_{j}^{(2)}-x_{ij}\beta|x_{ij}\right)\right)=E_{\varepsilon_{j}^{(2)}}\left(1-F_{\varepsilon_{ij}^{(1)}}\left(-x_{ij}\beta-\varepsilon_{j}^{(2)}\right)\right)$$

- Применяя закон больших чисел (ЗБЧ), соответствующее математическое ожидание можно приблизительно вычислить, насимулировав большое количество q независимых случайных величин $Z_1,...,Z_q$ из распределения такого же, как у $\varepsilon_{j}^{(2)}$:

- Обратим внимание, что (используя law of total expectation):

$$P(y_{ij}=1|x_{ij})=P(\varepsilon_{ij}^{(1)}+\varepsilon_{j}^{(2)}>-x_{ij}\beta|x_{ij})=P(\varepsilon_{ij}^{(1)}>-\varepsilon_{j}^{(2)}-x_{ij}\beta|x_{ij})=$$
$$=E\left(P(\varepsilon_{ij}^{(1)}>-\varepsilon_{j}^{(2)}-x_{ij}\beta|x_{ij})\right)=E_{\varepsilon_{j}^{(2)}}\left(E\left[P\left(\varepsilon_{ij}^{(1)}>-\varepsilon_{j}^{(2)}-x_{ij}\beta|x_{ij},\varepsilon_{j}^{(2)}\right)\right]\right)=$$
$$=E_{\varepsilon_{j}^{(2)}}\left(P\left(\varepsilon_{ij}^{(1)}>-\varepsilon_{j}^{(2)}-x_{ij}\beta|x_{ij}\right)\right)=E_{\varepsilon_{j}^{(2)}}\left(1-F_{\varepsilon_{ij}^{(1)}}\left(-x_{ij}\beta-\varepsilon_{j}^{(2)}\right)\right)$$

- Применяя закон больших чисел (ЗБЧ), соответствующее математическое ожидание можно приблизительно вычислить, насимулировав большое количество q независимых случайных величин $Z_1,...,Z_q$ из распределения такого же, как у $\varepsilon_{j}^{(2)}$:

$$P(y_{ij}=1|x_i)\approx\frac{1}{q}\sum_{i=1}^{q}1-F_{\varepsilon_{ij}^{(1)}}(-x_{ij}\beta-Z_q)$$

- По аналогии можно рассчитывать и совместные вероятности, например, обозначая через $Z_1^{(1)},...,Z_q^{(1)}$ и $Z_1^{(2)},...,Z_q^{(2)}$ две независимые выборки из того же распределения, что и $\varepsilon_j^{(2)}$, получаем:

$$P(y_{1j}=1,y_{2j}=1|x_{1j},x_{2j})=E_{\varepsilon_j^{(2)}}\left(E\left[P\left(\varepsilon_{1j}^{(1)}>-\varepsilon_j^{(2)}-x_{1j}\beta,\varepsilon_{2j}^{(1)}>-\varepsilon_j^{(2)}-x_{2j}\beta|x_{1j},x_{2j},\varepsilon_j^{(2)}\right)\right]\right)=$$

- По аналогии можно рассчитывать и совместные вероятности, например, обозначая через $Z_1^{(1)},...,Z_q^{(1)}$ и $Z_1^{(2)},...,Z_q^{(2)}$ две независимые выборки из того же распределения, что и $\varepsilon_j^{(2)}$, получаем:

$$P(y_{1j}=1,y_{2j}=1|x_{1j},x_{2j})=E_{\varepsilon_j^{(2)}}\left(E\left[P\left(\varepsilon_{1j}^{(1)}>-\varepsilon_j^{(2)}-x_{1j}\beta,\varepsilon_{2j}^{(1)}>-\varepsilon_j^{(2)}-x_{2j}\beta|x_{1j},x_{2j},\varepsilon_j^{(2)}\right)\right]\right)=$$

$$=E_{\varepsilon_j^{(2)}}\left(E\left[P\left(\varepsilon_{1j}^{(1)}>-\varepsilon_j^{(2)}-x_{1j}\beta|x_{1j},\varepsilon_j^{(2)}\right)P\left(\varepsilon_{2j}^{(1)}>-\varepsilon_j^{(2)}-x_{2j}\beta|x_{2j},\varepsilon_j^{(2)}\right)\right]\right)\approx$$

- По аналогии можно рассчитывать и совместные вероятности, например, обозначая через $Z_1^{(1)},...,Z_q^{(1)}$ и $Z_1^{(2)},...,Z_q^{(2)}$ две независимые выборки из того же распределения, что и $\varepsilon_j^{(2)}$, получаем:

$$P(y_{1j}=1, y_{2j}=1|x_{1j}, x_{2j}) = E_{\varepsilon_j^{(2)}}\left(E\left[P\left(\varepsilon_{1j}^{(1)} > -\varepsilon_j^{(2)} - x_{1j}\beta, \varepsilon_{2j}^{(1)} > -\varepsilon_j^{(2)} - x_{2j}\beta|x_{1j}, x_{2j}, \varepsilon_j^{(2)}\right)\right]\right) =$$

$$= E_{\varepsilon_j^{(2)}}\left(E\left[P\left(\varepsilon_{1j}^{(1)} > -\varepsilon_j^{(2)} - x_{1j}\beta|x_{1j}, \varepsilon_j^{(2)}\right)P\left(\varepsilon_{2j}^{(1)} > -\varepsilon_j^{(2)} - x_{2j}\beta|x_{2j}, \varepsilon_j^{(2)}\right)\right]\right) \approx$$

$$\approx \frac{1}{q}\sum_{i=1}^{q}\left(1 - F_{\varepsilon_{ij}^{(1)}}\left(-Z_q^{(1)} - x_{1j}\beta\right)\right)\left(1 - F_{\varepsilon_{ij}^{(1)}}\left(-Z_q^{(2)} - x_{2j}\beta\right)\right)$$

- По аналогии можно рассчитывать и совместные вероятности, например, обозначая через $Z_1^{(1)},...,Z_q^{(1)}$ и $Z_1^{(2)},...,Z_q^{(2)}$ две независимые выборки из того же распределения, что и $\varepsilon_j^{(2)}$, получаем:

$$P(y_{1j}=1,y_{2j}=1|x_{1j},x_{2j})=E_{\varepsilon_j^{(2)}}\left(E\left[P\left(\varepsilon_{1j}^{(1)}>-\varepsilon_j^{(2)}-x_{1j}\beta,\varepsilon_{2j}^{(1)}>-\varepsilon_j^{(2)}-x_{2j}\beta|x_{1j},x_{2j},\varepsilon_j^{(2)}\right)\right]\right)=$$

$$=E_{\varepsilon_j^{(2)}}\left(E\left[P\left(\varepsilon_{1j}^{(1)}>-\varepsilon_j^{(2)}-x_{1j}\beta|x_{1j},\varepsilon_j^{(2)}\right)P\left(\varepsilon_{2j}^{(1)}>-\varepsilon_j^{(2)}-x_{2j}\beta|x_{2j},\varepsilon_j^{(2)}\right)\right]\right)\approx$$

$$\approx\frac{1}{q}\sum_{i=1}^{q}\left(1-F_{\varepsilon_{ij}^{(1)}}\left(-Z_q^{(1)}-x_{1j}\beta\right)\right)\left(1-F_{\varepsilon_{ij}^{(1)}}\left(-Z_q^{(2)}-x_{2j}\beta\right)\right)$$

- Функция правдоподобия модели бинарного выбора со случайными эффектами будет иметь вид:

$$L\left(\beta,\sigma^2;y|X\right)=\prod_{j:y_{1j}=0,\ldots,y_{Tj}=0}P(y_{1j}=0,...,y_{Tj}=0)\times...\times\prod_{j:y_{1j}=1,\ldots,y_{Tj}=1}P(y_{1j}=1,...,y_{Tj}=1)$$

- По аналогии можно рассчитывать и совместные вероятности, например, обозначая через $Z_1^{(1)}, ..., Z_q^{(1)}$ и $Z_1^{(2)}, ..., Z_q^{(2)}$ две независимые выборки из того же распределения, что и $\varepsilon_j^{(2)}$, получаем:

$$P(y_{1j} = 1, y_{2j} = 1|x_{1j}, x_{2j}) = E_{\varepsilon_j^{(2)}}\left(E\left[P\left(\varepsilon_{1j}^{(1)} > -\varepsilon_j^{(2)} - x_{1j}\beta, \varepsilon_{2j}^{(1)} > -\varepsilon_j^{(2)} - x_{2j}\beta|x_{1j}, x_{2j}, \varepsilon_j^{(2)}\right)\right]\right) =$$

$$= E_{\varepsilon_j^{(2)}}\left(E\left[P\left(\varepsilon_{1j}^{(1)} > -\varepsilon_j^{(2)} - x_{1j}\beta|x_{1j}, \varepsilon_j^{(2)}\right) P\left(\varepsilon_{2j}^{(1)} > -\varepsilon_j^{(2)} - x_{2j}\beta|x_{2j}, \varepsilon_j^{(2)}\right)\right]\right) \approx$$

$$\approx \frac{1}{q}\sum_{i=1}^{q}\left(1 - F_{\varepsilon_{ij}^{(1)}}\left(-Z_q^{(1)} - x_{1j}\beta\right)\right)\left(1 - F_{\varepsilon_{ij}^{(1)}}\left(-Z_q^{(2)} - x_{2j}\beta\right)\right)$$

- Функция правдоподобия модели бинарного выбора со случайными эффектами будет иметь вид:

$$L\left(\beta, \sigma^2; y|X\right) = \prod_{j:y_{1j}=0,...,y_{Tj}=0} P(y_{1j} = 0, ..., y_{Tj} = 0) \times ... \times \prod_{j:y_{1j}=1,...,y_{Tj}=1} P(y_{1j} = 1, ..., y_{Tj} = 1)$$

Максимизация соответствующей функции с использование аппроксимаций вероятностей, указанных выше, именуется симуляционным методом максимального правдоподобия.

- Рассмотрим модель, в которой у каждого индивида различаются коэффициенты, отражающие влияние регрессоров на латентную переменную:

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

Модели бинарного выбора со случайными коэффициентами

Формулировка и оценивание

- Рассмотрим модель, в которой у каждого индивида различаются коэффициенты, отражающие влияние регрессоров на латентную переменную:

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

- Для того, чтобы учесть возможность различия эффектов различных переменных на зависимую бинарную переменную можно предположить, что все регрессионные коэффициенты являются случайными величинами. Для простоты предположим, что они имеют нормальное распределение $\beta_{ik} \sim \mathcal{N}\left(\beta_k, \sigma_k^2\right)$.

- Рассмотрим модель, в которой у каждого индивида различаются коэффициенты, отражающие влияние регрессоров на латентную переменную:

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

- Для того, чтобы учесть возможность различия эффектов различных переменных на зависимую бинарную переменную можно предположить, что все регрессионные коэффициенты являются случайными величинами. Для простоты предположим, что они имеют нормальное распределение $\beta_{ik} \sim \mathcal{N}\left(\beta_k, \sigma_k^2\right)$.
- По аналогии с моделями со случайными эффектами для оценивания параметров распределения случайных коэффициентов, то есть β_k и σ_k^2, можно воспользоваться методом симуляционного правдоподобия.

- Рассмотрим модель, в которой у каждого индивида различаются коэффициенты, отражающие влияние регрессоров на латентную переменную:

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

- Для того, чтобы учесть возможность различия эффектов различных переменных на зависимую бинарную переменную можно предположить, что все регрессионные коэффициенты являются случайными величинами. Для простоты предположим, что они имеют нормальное распределение $\beta_{ik} \sim \mathcal{N}\left(\beta_k, \sigma_k^2\right)$.
- По аналогии с моделями со случайными эффектами для оценивания параметров распределения случайных коэффициентов, то есть β_k и σ_k^2, можно воспользоваться методом симуляционного правдоподобия.
- При $\sigma_k^2 = 0$ коэффициент является не случайным, а детерминированным, что можно использовать для тестирования соответствующей гипотезы.

- Рассмотрим модель, в которой у каждого индивида различаются коэффициенты, отражающие влияние регрессоров на латентную переменную:

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

- Для того, чтобы учесть возможность различия эффектов различных переменных на зависимую бинарную переменную можно предположить, что все регрессионные коэффициенты являются случайными величинами. Для простоты предположим, что они имеют нормальное распределение $\beta_{ik} \sim \mathcal{N}\left(\beta_k, \sigma_k^2\right)$.
- По аналогии с моделями со случайными эффектами для оценивания параметров распределения случайных коэффициентов, то есть β_k и σ_k^2, можно воспользоваться методом симуляционного правдоподобия.
- При $\sigma_k^2 = 0$ коэффициент является не случайным, а детерминированным, что можно использовать для тестирования соответствующей гипотезы.
- Существенный недостаток данной модели заключается в нереалистичном допущении о независимости β_{ik} и x_i. В таком случае предположение о гетерогенности эффекта регрессоров в популяции вступает, в содержательном смысле, в противоречие с допущением о том, что эта гетерогенность не связана с наблюдаемыми характеристиками исследуемых объектов x_i.