Эконометрический ликбез: ошибки спецификации

Основы теории квази- и псевдо-правдоподобия*

Станислав Анатольев†

Российская экономическая школа, Москва, Россия
CERGE-EI, Прага, Чехия

Данное эссе содержит краткий обзор концепций и методов, связанных с неверно специфицированным распределением при применении принципа максимального правдоподобия: квази-плотность, псевдо-плотность, квази-правдоподобие, псевдо-правдоподобие и т.д. Обзор сопровождается примерами и задачами.

1 Методы квази-правдоподобия

Методом максимального квази-правдоподобия (ММКП) называется метод максимального правдоподобия, применяемый в отсутствие в модели спецификации (условной) плотности. В качестве плотности применяется так называемая квази-плотность, с осознанием, что она не является истинной плотностью. Неверная спецификация, вроде бы непозволительная в расчёте на состоятельность оценок, не всегда приводит к катастрофическим результатам. Данный анализ была впервые проделан Халбертом Уайтом (White 1982). Мы предполагаем, что используется случайная выборка.

Рассмотрим в качестве примера случай $y \sim \mathcal{D}(\mu_0, \sigma_0^2)$, где μ_0 и σ_0^2 — истинные математическое ожидание и дисперсия y, а $\mathcal{D}$ — её истинная неизвестная плотность. Предположим, что исследователь применяет ММКП на основе нормальной плотности, т.е. спецификации $y \sim \mathcal{N}(\mu, \sigma^2)$. Конечно же, $\mathcal{D}$ не обязана быть $\mathcal{N}$. Логарифмическая квази-плотность равна

$$\log f(y|\mu, \sigma^2) = \text{const} - 0.5\log(\sigma^2) - \frac{(y-\mu)^2}{2\sigma^2}.$$

Посмотрим, какие параметры решают задачу максимизации ожидаемой лог-квази-плотности в популяции:

$$\begin{aligned}
\arg\max{}_{\mu,\sigma^2} \mathbb{E}\left[\log f(y|\mu,\sigma^2)\right] &= \arg\max{}_{\mu,\sigma^2} \mathbb{E}\left[\text{const} - 0.5\log(\sigma^2) + \frac{(y-\mu)^2}{2\sigma^2}\right] \\
&= \arg\min{}_{\mu,\sigma^2} \left\{\log(\sigma^2) + \frac{(\mu-\mu_0)^2 + \sigma_0^2}{\sigma^2}\right\}.
\end{aligned}$$

Приравняв производные нулю, легко убедиться, что решением (или, как говорят, *псевдо-истинными параметрами*) является пара μ_0, σ_0^2, то есть истинные параметры решают задачу максимизации ожидаемой лог-квази-плотности в популяции, а значит, построенные методом аналогий оценки ММКП будут состоятельными, несмотря на то, что использовалась не истинная, а квази-плотность. Критическим фактором данного феномена является, конечно же, выбор нормальной квази-плотности. Обобщением данного примера является линейная (или

*Цитировать как: Анатольев, Станислав (2019) «Основы теории квази- и псевдо-правдоподобия», Квантиль, №14, стр. 45–52. Citation: Anatolyev, Stanislav (2019) "Basics of quasi- and pseudo-likelihood theories," Quantile, No.14, pp. 45–52.

†Адрес: 121353, г. Москва, Сколковское шоссе, дом 45. Электронная почта: *sanatoly@nes.ru*

нелинейная) регрессия среднего, где ММКП-оценки регрессионной функции, рассчитанные на основе нормальной условно гомоскедастичной плотности, состоятельны хотя бы потому, что совпадают с МНК-оценками (или НЛМНК-оценками).

Рассмотрим, однако, следующий похожий пример. Пусть $y \sim \mathcal{N}(\mu_0, \sigma_0^2)$, где μ_0 и σ_0^2 — истинные математическое ожидание и дисперсия y. Предположим, что исследователь применяет ММКП на основе нормальной плотности, но ошибается в спецификации среднего: $y \sim \mathcal{N}(0, \sigma^2)$. Логарифмическая квази-плотность равна

$$\log f(y|\sigma^2) = \text{const} - 0.5\log(\sigma^2) - \frac{y^2}{2\sigma^2}.$$

Посмотрим, какой параметр дисперсии решают задачу максимизации ожидаемой лог-квази-плотности в популяции:

$$\begin{aligned}\arg\max\nolimits_{\sigma^2} \mathbb{E}\left[\log f(y|\sigma^2)\right] &= \arg\max\nolimits_{\sigma^2} \mathbb{E}\left[\text{const} - 0.5\log(\sigma^2) - \frac{y^2}{2\sigma^2}\right] \\ &= \arg\min\nolimits_{\sigma^2} \left\{\log(\sigma^2) + \frac{\mu_0^2 + \sigma_0^2}{\sigma^2}\right\}.\end{aligned}$$

Приравняв производные нулю, легко убедиться, что решением (псевдо-истинным параметром) является значение $\mu_0^2 + \sigma_0^2$, то есть ММКП-оценка дисперсии несостоятельна, хотя использовалась квази-плотность с истинной формой. Здесь негативный результат является следствием неверной спецификации среднего.

Мораль данных двух примеров в том, что при удачном выборе неверной квази-плотности всё ещё остаётся надежда на состоятельное оценивание (по крайней мере, некоторых) параметров, в то время как в случае неверной спецификации условных моментов (в данном случае – среднего) рассчитывать на состоятельность ММКП-оценок вряд ли приходится, даже если с формой плотности «угадали».

Если ММКП-оценка оказались состоятельна, неверная спецификация плотности всё же имеет некоторые последствия — хотя оценка по-прежнему асимптотически нормальна, она асимптотически неэффективна. Кроме того, равенство информационной матрицы нарушено, и асимптотическая дисперсия принимает форму «сэндвича». Соответственно, состоятельно оценивать нужно целый «сэндвич», или, как говорят, строить *робастные оценки* (по отношению к спецификации плотности).

Задача 1: Пусть спецификация эконометрической модели следующая:

$$y|x \sim \mathcal{N}(x^\delta, 1),$$

где x положительно почти наверное. Однако, истинное условное распределение не является нормальным, а на самом деле даже не является симметричным; условная дисперсия не равна 1, а на самом деле равна $\frac{1}{2}$. Каковы последствия применения ММП в расчёте на то, что спецификация верна? Выборка $\{(x_i, y_i)\}_{i=1}^n$ случайная.

Решение задачи 1: Легко видеть, что ММКП в качестве решения даёт оценку НЛМНК

$$\hat{\delta} = \arg\min_d \sum_{i=1}^n \left(y_i - x_i^d\right)^2,$$

которая является состоятельной назависимо от истинной и предполагаемой дисперсий ошибки. Правильная асимптотика для этой оценки следующая:

$$\sqrt{n}\left(\hat{\delta} - \delta\right) \xrightarrow{d} \mathcal{N}\left(0, V_\delta\right),$$

где асимптотическая дисперсия равна $V_\delta = \frac{1}{2}E\left[x^{2\delta}\log(x)^2\right]^{-1}$. Однако, в расчёте на неверное значение дисперсии ошибки исследователь завысит V_δ в два раза (если не воспользуется «сэндвичным» представлением при оценивании асимптотической дисперсии). □

Задача 2: Положим, данные $\{(x_i, y_i)\}_{i=1}^n$ — случайная выборка из популяции, где среднее y условно на x равно $g(x,\beta)$, а дисперсия y условно на x равна $\sigma^2(x)$, т.е. зависит от x. Мы оцениваем вектор параметров β, полагаясь на нормальное распределение с верно специфицированным средним $g(x,b)$, но неверно специфицированной дисперсионной функцией σ^2, то есть неверно налагая условную гомоскедастичность. Будет ли вектор β состоятельно оцениваться? Какую величину будет состоятельно оценивать оценка дисперсии?

Решение задачи 2: Условная квази-плотность одного наблюдения равна

$$\log f\left(y|x,b,s^2\right) = \text{const} - \frac{1}{2}\log s^2 - \frac{1}{2s^2}\left(y - g\left(x,b\right)\right)^2.$$

Популяционные величины, которые будут состоятельно оцениваться, следующие:

$$\begin{aligned}\begin{pmatrix}\beta_*\\ \sigma_*^2\end{pmatrix} &= \arg\max_{b,s^2}\mathbb{E}\left[-\frac{1}{2}\log s^2 - \frac{1}{2s^2}\left(y-g\left(x,b\right)\right)^2\right]\\ &= \arg\max_{b,s^2}\left\{-\frac{1}{2}\log s^2 - \frac{1}{2s^2}\left(\mathbb{E}\left[\sigma^2\left(x\right)\right] + \mathbb{E}\left[\left(g\left(x,\beta\right)-g\left(x,b\right)\right)^2\right]\right)\right\}.\end{aligned}$$

Видно, что в оптимуме псевдо-истинное значение $\beta_* = \beta$, то есть β будет состоятельно оцениваться (если условие идентификации выполнено). Оценка дисперсии будет состоятельно оценивать псевдо-истинный параметр $\sigma_*^2 = \mathbb{E}\left[\sigma^2\left(x\right)\right]$. □

Добавим, что всё вышесказанное также выполняется для моделей на стационарных эргодичных временных рядах, если регрессионная ошибка имеет структуру мартингальных приращений. «Сэндвичные» робастные оценки дисперсии в таким случаях называются оценками Боллерслева–Вулдриджа, по именам авторов статьи Bollerslev & Wooldridge (1992).

2 Методы псевдо-правдоподобия

Повествование в данном разделе опирается на работу Gourieroux, Monfort & Trognon (1984).

2.1 Оценка ММПП первого порядка

Рассмотрим следующую *Модель 1*:

$$\mathbb{E}\left[y|x\right] = m\left(x,\theta_0\right),$$

где $m(\cdot,\cdot)$ — известная функция, а $\theta_0 \in \Theta$ — истинное значение неизвестного параметра θ. Рассмотрим семейство плотностей $f(u,\mu)$, где μ — среднее распределения. Видим, что данная задача — обычная регрессионная постановка, в рамках которой исследователь намерен применять ММП без, возможно, достаточных на то оснований.

Определим оценку *максимального псевдо-правдоподобия первого порядка* (ММПП1):

$$\hat{\theta}_1 = \arg\max_{\theta\in\Theta}\frac{1}{n}\sum_{i=1}^n \log f\left(y_i, m\left(x_i,\theta\right)\right).$$

Определим также *линейно-экспоненциальное семейство плотностей* как семейство, содержащее в себе плотности, записываемые в следующей форме:

$$f\left(u,\mu\right) = \exp\left(A(\mu) + B\left(u\right) + C\left(\mu\right)u\right),$$

где $A(\mu), B(u), C(\mu)$ — скалярные функции, удовлетворяющие определённым ограничениям, среди которых, в частности, следующее: $\partial A(\mu)/\partial\mu + \mu\partial C(\mu)/\partial\mu = 0$, и где μ — среднее f. Данное семейство включает, например, следующие плотности:

- Нормальная с единичной дисперсией: $f(u,\mu) = (2\pi)^{-1/2}\exp\left(-\frac{1}{2}(u-\mu)^2\right)$. Для неё $C(\mu) = \mu$.
- Пуассонова: $f(u,\mu) = \exp(-\mu)\mu^u/u!$. Для неё $C(\mu) = \log\mu$.
- Гамма с фиксированным параметром α: $f(u,\mu) = u^{\alpha-1}(\mu/\alpha)^{-\alpha}\exp(-\alpha u/\mu)/\Gamma(\alpha)$. Для неё $C(\mu) = -\alpha/\mu$.
- Бернулли: $f(u,\mu) = \mu^u(1-\mu)^{1-u}$. Для неё $C(\mu) = \log(\mu/(1-\mu))$.

Заметим, что такие распространённые плотности, как хи-квадрат, бета, Стьюдента, Вейбулла и многие другие не принадлежат рассматриваемому классу.

Мы, конечно, не случайно рассматриваем данное семейство плотностей. Именно для таких плотностей ММПП1 даёт состоятельные оценки регрессионных параметров, даже если истинная плотность отличается от выбранной.

Теорема 1: Если псевдо-плотность $f(u,\mu)$ принадлежит линейно-экспоненциальному семейству, то

$$\hat\theta_1 \xrightarrow{p} \theta_0$$

и

$$\sqrt{n}\left(\hat\theta_1 - \theta_0\right) \xrightarrow{d} \mathcal{N}\left(0, \mathcal{J}_1^{-1}\mathcal{I}_1\mathcal{J}_1^{-1}\right),$$

где

$$\mathcal{J}_1 = \mathbb{E}\left[\left.\frac{\partial C(\mu)}{\partial\mu}\right|_{m(x,\theta_0)} \frac{\partial m(x,\theta_0)}{\partial\theta}\frac{\partial m(x,\theta_0)}{\partial\theta'}\right]$$

и

$$\mathcal{I}_1 = \mathbb{E}\left[\left(\left.\frac{\partial C(\mu)}{\partial\mu}\right|_{m(x,\theta_0)}\right)^2 \mathbb{V}[y|x]\,\frac{\partial m(x,\theta_0)}{\partial\theta}\frac{\partial m(x,\theta_0)}{\partial\theta'}\right].$$

Задача 3: Установите, приводит ли использование отрицательной биномиальной плотности

$$f(u,\mu) = \frac{\Gamma(a+u)}{\Gamma(a)\Gamma(1+u)}\left(\frac{\mu}{a}\right)^u\left(1+\frac{\mu}{a}\right)^{-(a+u)},$$

где μ — его среднее, а a — произвольная известная константа, к состоятельному оцениванию θ_0 в регрессии среднего

$$\mathbb{E}[y|x] = m(x,\theta_0),$$

когда истинное условное распределение является условно гетероскедастичным и нормальным, со скедастичной функцией $\mathbb{V}[y|x] = 2m(x,\theta_0)$.

Решение задачи 3: Да. Данная плотность принадлежит линейно-экспоненциальному семейству, причём $C(\mu) = \log(\mu) - \log(a+\mu)$. Форма скедастичной функции не влияет на данное свойство. □

Задача 4: Рассмотрим регрессионную модель $\mathbb{E}[y|x] = m(x, \theta_0)$. Предположим, что ММПП1-оценка на основе псевдо-плотности $f(z, \mu)$, параметризованной средним μ, состоятельно оценивает истинный параметр θ_0. Рассмотрим другую псевдо-плотность $h(z, \mu, \varsigma)$, параметризованную двумя параметрами, средним μ и неким параметром ς, которая включает $f(z, \mu)$ как особый случай. Используя пример распределения Вейбулла с плотностью

$$h(z, \mu, \varsigma) = \varsigma \left(\frac{\Gamma(1 + \varsigma^{-1})}{\mu} \right)^{\varsigma} z^{\varsigma - 1} \exp \left(- \left(\frac{\Gamma(1 + \varsigma^{-1})}{\mu} z \right)^{\varsigma} \right) \cdot \mathbb{I}[z \geq 0], \quad \varsigma > 0,$$

покажите, что ММПП1-оценка на основе $h(z, \mu, \varsigma)$ необязательно состоятельно оценивает θ_0. Каково эконометрическое объяснение этого, возможно контринтуитивного, результата?

Решение задачи 4: Действительно, распределение Вейбулла сводится к экспоненциальному, когда $\varsigma = 1$, а ММПП1 с экспоненциальной псевдо-плотностью состоятельно оценивает θ_0. Когда ς также оценивается, вектором параметров является $(\mu, \varsigma)'$, а псевдо-скор имеет вид

$$\begin{aligned} \frac{\partial \log h(z, \mu, \varsigma)}{\partial (\mu, \varsigma)'} &= \frac{\partial \left(\log \varsigma + \varsigma \log \Gamma\left(1 + \varsigma^{-1}\right) - \varsigma \log \mu + (\varsigma - 1) \log z - \left(\Gamma\left(1 + \varsigma^{-1}\right) z/\mu\right)^{\varsigma} \right)}{\partial (\mu, \varsigma)'} \\ &= \begin{pmatrix} \left(\left(\Gamma\left(1 + \varsigma^{-1}\right) z/\mu\right)^{\varsigma} - 1 \right) \varsigma/\mu \\ \dots \end{pmatrix}. \end{aligned}$$

Заметим, что псевдо-скор для μ имеет нулевое матожидание тогда и только тогда, когда $\mathbb{E}\left[\left(\Gamma\left(1 + \varsigma^{-1}\right) z/\mu\right)^{\varsigma}\right] = 1$. Очевидно, это условие выполнено при $\varsigma = 1$, но может не выполняться для других ς. Например, если $\varsigma = 2$, мы имеем $\mathbb{E}\left[z^2\right] = \mu^2 \Gamma(2.5) / \Gamma(1.5)^2 = 1.5\mu^2/\Gamma(1.5)$, что противоречит нулевому ожидаемому преводо-скор. Псевдо-истинное значение ς_* для ς можно получить решением системы, приравняв преводо-скор к нулю, но маловероятно, что решением будет 1. Данное явление можно объяснить наличием постороннего параметра, псевдо-истинное значение которого не имеет отношения к задаче и оценивание которого пагубно влияет на оценивание интересующего нас параметра. □

Следующий естественный вопрос — раз состоятельных асимптотически нормальных оценок ММПП1 целое множество, какая (если такая есть) из них наиболее асимптотически эффективная? Ответ на этот вопрос содержится в следующем предложении.

Предложение: Минимум асимптотической дисперсии $\mathcal{J}_1^{-1} \mathcal{I}_1 \mathcal{J}_1^{-1}$ равен

$$\left(\mathbb{E} \left[\frac{1}{\mathbb{V}[y|x]} \frac{\partial m(x, \theta_0)}{\partial \theta} \frac{\partial m(x, \theta_0)}{\partial \theta'} \right] \right)^{-1}$$

и достигается при

$$\left. \frac{\partial C(\mu)}{\partial \mu} \right|_{m(x, \theta_0)} \propto \frac{1}{\mathbb{V}[y|x]}.$$

Легко вывести, что нормальное условное распределение оптимально использовать при условной гомоскедастичности ($\partial C(\mu) / \partial \mu = 1$), условное распределение Пуассона — при условной дисперсии, пропорциональной скедастичной функции ($\partial C(\mu) / \partial \mu = \mu^{-1}$), условное Гамма-распределения — при условной дисперсии, пропорциональной квадрату скедастичной функции ($\partial C(\mu) / \partial \mu = -\alpha \mu^{-2}$) и т.д. Неудивительно, что оптимальность распределения имеет место тогда, когда условное среднее и условная дисперсия находятся в соотношении, характерном для данного распределения.

2.2 Оценка ММПП второго порядка

Рассмотрим теперь *Модель 2*:

$$\begin{aligned} \mathbb{E}\left[y|x\right] &= m\left(x,\theta_0\right), \\ \mathbb{V}\left[y|x\right] &= v\left(x,\theta_0\right), \end{aligned}$$

где $m\left(\cdot,\cdot\right)$ и $v\left(\cdot,\cdot\right)$ — известные функции, а $\theta_0\in\Theta$ — истинное значение неизвестного параметра θ.

Рассмотрим семейство плотностей $f\left(u,\mu,v\right)$, где μ — среднее распределения, а v — его дисперсия. Видим, что в данной задаче задана не только форма регрессии среднего, но и скедастичная функция, и опять же исследователь намерен применить ММП. Самым популярным контекстом такой постановки является класс GARCH-моделей для волатильности.

Определим оценку *максимального псевдо-правдоподобия второго порядка* (ММПП2):

$$\hat{\theta}_2=\arg\max_{\theta\in\Theta}\frac{1}{n}\sum_{i=1}^{n}\log f\left(y_i,m\left(x_i,\theta\right),v\left(x_i,\theta\right)\right)$$

Определим также *квадратично-экспоненциальное семейство плотностей* как семейство, содержащее в себе плотности, записываемые в следующей форме:

$$f\left(u,\mu,v\right)=\exp\left(A(\mu,v)+B\left(u\right)+C\left(\mu,v\right)u+D\left(\mu,v\right)u^2\right),$$

где $A(\mu,v),B\left(u\right),C\left(\mu,v\right),D\left(\mu,v\right)$ — скалярные функции, удовлетворяющие определённым ограничениям, и где μ — среднее f, а v — дисперсия f. Данное семейство включает, в частности, следующие плотности:

- Нормальное: $f\left(u,\mu,v\right)=\left(2\pi v\right)^{-1/2}\exp\Big(-\left(u-\mu\right)^2/\left(2v\right)\Big)$.
- Трёхточечное: принимающее значения из множества $\{-1,0,+1\}$ с вероятностями p_1, p_2, $1-p_1-p_2$, соответственно.

И вновь, именно для плотностей из рассматриваемого класса ММПП2 даёт состоятельные оценки регрессионных и скедастичных параметров, даже если истинная плотность отличается от выбранной.

Теорема 2: Если псевдо-плотность $f\left(u,\mu,v\right)$ принадлежит квадратично-экспоненциальному семейству, то

$$\hat{\theta}_2\xrightarrow{p}\theta_0$$

и

$$\sqrt{n}\left(\hat{\theta}_2-\theta_0\right)\xrightarrow{d}\mathcal{N}\left(0,\mathcal{J}_2^{-1}\mathcal{I}_2\mathcal{J}_2^{-1}\right),$$

где

$$\mathcal{J}_2=\mathbb{E}\left[\left.\frac{\partial^2\log f\left(u,m\left(x,\theta\right),v\left(x,\theta\right)\right)}{\partial\theta\partial\theta'}\right|_{\theta=\theta_0}\right]$$

и

$$\mathcal{I}_2=\mathbb{E}\left[\left.\frac{\partial\log f\left(u,m\left(x,\theta\right),v\left(x,\theta\right)\right)}{\partial\theta}\right|_{\theta=\theta_0}\left.\frac{\partial\log f\left(u,m\left(x,\theta\right),v\left(x,\theta\right)\right)}{\partial\theta'}\right|_{\theta=\theta_0}\right].$$

Задача 5: Рассмотрим следующую модель:

$$y_i = \alpha + e_i, \quad i = 1, \ldots, n,$$

где ненаблюдаемые e_i, условно на x_i, распределены симметрично вокруг нуля с дисперсией $x_i^2\sigma^2$, где σ^2 неизвестна. Выборка (y_i, x_i) случайная. Выведите ММПП2-оценки параметров α и σ^2 на основе нормального распределения и их асимптотические свойства.

Решение задачи 5: ММПП2-оценка на основе нормального распределения определяется как решение следующей задачи:

$$\begin{pmatrix} \hat{\alpha} \\ \hat{\sigma}^2 \end{pmatrix} = \arg\max_{\alpha,\sigma^2} \left\{ \text{const} - \frac{n}{2}\log\sigma^2 - \frac{1}{\sigma^2}\sum_{i=1}^{n}\frac{(y_i-\alpha)^2}{2x_i^2} \right\}.$$

Решая её, получаем

$$\hat{\alpha} = \left(\sum_{i=1}^{n}\frac{1}{x_i^2}\right)^{-1}\sum_{i=1}^{n}\frac{y_i}{x_i^2}, \quad \hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}\frac{(y_i-\hat{\alpha})^2}{x_i^2}.$$

Используя теорию ММПП2-оценивания или же напрямую, получаем следующие асимптотические дисперсии:

$$V_{\hat{\alpha}} = \sigma^2\left(\mathbb{E}\left[x^{-2}\right]\right)^{-1}$$

и

$$V_{\hat{\sigma}^2} = \mathbb{E}\left[\frac{\mathbb{E}\left[(y-\alpha)^4|x\right]}{x^4}\right] - \sigma^4.$$

2.3 Дополнительный результат

Оказывается, что специфицированные выше семейства плотностей — единственные, при которых гарантируется состоятельность соответствующих МПП-оценок.

Предложение:

1. Необходимым и достаточным условием состоятельности ММПП1-оценки является принадлежность $f(u, \mu)$ линейно-экспоненциальному семейству плотностей.

2. Необходимым и достаточным условием состоятельности ММПП2-оценки является принадлежность $f(u, \mu, \upsilon)$ квадратично-экспоненциальному семейству плотностей.

Данный результат означает, что при использовании распределений, не входящих в классы экспоненциальных семейств, исследователь, скорей всего, получает неосостоятельные оценки параметров регрессии среднего и скедастичной функции. Из этого следует, в частности, что при необходимости применения ММКП- и ММПП-оценивания (например, в динамических моделях, описывающих распределения финансовых доходностей) нужно быть предельно осторожным и перестраховываться, выбирая максимально возможную гибкость формы условной плотности (скажем, использовать скошенное распределения Стьюдента вместо обычного распределения Стьюдента).

Список литературы

Bollerslev, T. & J.M. Wooldridge (1992). Quasi-maximum likelihood estimation and inference in dynamic models with time-varying covariances. *Econometric Reviews* 11, 143–172.

Gourieroux, C., A. Monfort & A. Trognon (1984). Pseudo maximum likelihood methods: Theory. *Econometrica* 52, 681–700.

White, H. (1982). Maximum likelihood estimation of misspecified models. *Econometrica* 50, 1–25.

Basics of quasi- and pseudo-likelihood theories

Stanislav Anatolyev

CERGE-EI, Prague, Czech Republic

New Economic School, Moscow, Russia

This essay contains a brief review of concepts and methods related to the principle of maximum likelihood based on misspecified distributions: quasi-density, pseudo-density, quasi-likelihood, pseudo-likelihood, etc. The review is accompanied with examples and problems.