

В помощь изучающим эконометрику

Мини-словарь англоязычных эконометрических терминов, часть 3*

Александр Цыплаков[†]

Новосибирский государственный университет, Новосибирск, Россия

В этой части словаря комментируются англоязычные эконометрические термины *estimator*, *quantile*, *nested*, *marginal* и др. Как и в предыдущих частях, акцент делается на уточнении значения терминов с целью избежать возможной путаницы и некорректностей при интерпретации.

Estimator

Основная проблема с переводом терминов, связанных с получением статистических оценок и изучением их свойств, состоит в том, что русское слово «оценка» может означать по крайней мере три разные вещи. Во-первых, это собственно оценка, т. е. значение, полученное в конкретной выборке — по-английски *estimate*. Во-вторых, это метод (или функция) в соответствии с которым получают оценки в первом смысле — по-английски *estimator*. В-третьих, это процесс получения оценок в первом смысле — по-английски *estimation*. Во избежание путаницы *estimator* лучше переводить как «метод оценивания» или «оцениватель», а *estimation* — оценивание. К сожалению, использование термина «оценка» во втором смысле прочно укоренилось в русскоязычной статистической литературе. Если говорят о свойствах оценки, таких как несмещенность или эффективность, то на самом деле имеют в виду свойства соответствующего метода, поскольку бессмысленно, например, говорить об эффективности конкретного числа. Особенно следует следить за адекватностью перевода слова «оценка» с русского на английский, потому что если значение слова «оценка» в русском тексте обычно можно понять из контекста, то «consistent estimate», употребленное в неудачном месте, может вызвать у читателя недоумение.

Estimation: оценивание, оценка (как процесс).

Estimator: метод оценивания, способ оценивания, оценка, оцениватель.

Estimate (гл.): оценивать.

Estimate (сущ.): оценка.

Estimated: оцененный.

Estimating: оценивающий, оценочный.

Estimating equations: оценивающие уравнения, уравнения, задающие оценку.

Estimand: оцениваемая величина.

*Цитировать как: Цыплаков, Александр (2014) «Мини-словарь англоязычных эконометрических терминов, часть 3», Квантиль, №12, стр. 45–51. Citation: Tsyplakov, Alexander (2014) “A mini-dictionary of English econometric terminology III,” *Quantile*, No.12, pp. 45–51.

[†] Адрес: 630090, г. Новосибирск, Весенний проезд, 6–44. Электронная почта: tsy@academ.org

Quantile

Пусть случайная величина с носителем $[a, b]$ имеет функцию распределения $F(x)$, которая непрерывна и возрастает на $[a, b]$ (здесь a и b могут быть бесконечными). Тогда квантиль (quantile) случайной величины для вероятности p определяется как $F^{-1}(p)$. В частности, при $p = 1/2$ получаем медиану распределения.

Как аббревиатура слова **quantile** может использоваться буква **Q**. Например, **Q-Q plot** — это сокращение от **quantile-quantile plot**.

Если отрезок $[0, 1]$ разделить на n равных промежутков, то для соответствующих вероятностей $1/n, \dots, (n-1)/n$ имеем семейство «равномерных» квантилей. Второе значение слова **quantile** в английском языке — квантиль из такого семейства. В этом смысле i -я квантиль равна $F^{-1}(i/n)$, $i = 1, \dots, n-1$.

Для некоторых n у квантилей, соответствующих разбиению на n частей, есть специальные названия, происходящие от соответствующих латинских числительных. В частности, децили соответствуют разбиению на 10 частей, квартили — на 4 части, (более редко встречающиеся) квинтили — на 5 частей. Слова **quantile**, **quartile** и **quintile** легко спутать, но надо помнить, что первый из этих терминов общий, в отличие от остальных двух. (А еще есть английское слово **quantal** — «квантовый».)

Из квантилей во втором смысле особо выделяют нижнюю (**lower**), то есть первую, и верхнюю (**upper**), то есть $(n-1)$ -ю. Так нижняя квартиль соответствует вероятности $1/4$, а верхняя — вероятности $3/4$ (еще одна квартиль — вторая — это, конечно, просто медиана). Разность между верхней и нижней квартилями называют интерквартильным размахом (**interquartile range**). Соответственно, разность между девятой и первой децилями называют интердецильным размахом.

Процентиль (**percentile**, иногда просто **centile**) соответствует разбиению на 100 промежутков. Термин **percentile** может использоваться почти как синоним слова **quantile** в первом смысле. По-видимому, это связано с тем, что 100 промежутков — это уже «много» и более мелкое деление с практической точки зрения не всегда нужно. Квантили в привычном нам первом смысле по-английски называют также **fractile**. То есть “0.35-fractile” — это “35th percentile”. (**Fractile** желательно не путать со словом **fractal** — «фрактал».)

Support: носитель (распределения).

Quantile: квантиль.

Median: медиана.

Q-Q plot (quantile-quantile plot): диаграмма квантиль-квантиль, Q-Q диаграмма.

Quartile/quintile/decile: квартиль/квинтиль/дециль.

Upper/lower quartile: нижняя/верхняя квартиль.

Interquartile/interdecile: интерквартильный/интердецильный, межквартильный/междецильный.

Interquartile range: интерквартильный размах, межквартильный размах, интерквартильная широта.

Percentile (centile): процентиль, перцентиль, квантиль.

Fractile: фрактиль, квантиль.

Quantile: журнал *Квантиль*.

Nested, encompassing

Когда модель A является частным случаем модели B , т. е. A можно получить из B , наложив на параметры некоторые ограничения (и, возможно, сделав замену переменных), то говорят, что модель A вложена (*nested*) в модель B . Если же A и B не сводятся друг к другу с помощью наложения ограничений на параметры, то говорят, что A и B являются невложенными (*non-nested*); примером могут служить модели логит и пробит. Если правильность спецификации одной модели проверяют на основе сопоставления с невложенной альтернативной моделью, то говорят о невложенных гипотезах. Термин связан не с прямым значением глагола *to nest* — «гнездиться», а с переносным: «помещать один объект внутрь другого» (как одну матрешку внутрь другой) или «содержать другой объект». Соответственно *to be nested* (редко также *to nest*) «находиться внутри другого объекта».

В том же контексте выбора модели есть похожий термин *encompassing* (охват). *Принцип охвата* заключается в том, что если модель правильно специфицирована, то она должна быть способна объяснять результаты использования альтернативной модели. О выборе модели, невложенных и охватывающих моделях см. Gougiou & Monfort (1994), Pesaran & Weeks (2001), Greene (2012).

Кроме вложенных моделей и гипотез в эконометрике известна также модель, называемая *nested logit*, в основе которой лежит иерархия вложенных друг в друга структур — множеств альтернатив. См. Greene (2012).

Nested/non-nested: вложенный/невложенный.

Nested model/hypothesis: вложенная модель/гипотеза.

Nested logit: вложенный логит.

Nest (гл.): вкладывать, включать (как частный случай).

Nest (сущ.): «гнездо», подмножество альтернатив (во вложенном логите).

Be nested: вкладываться, быть частным случаем.

Nesting: вложение (одной гипотезы/модели/структуры в другую).

Encompass: охватывать, заключать в себе.

Encompassing: охват; охватывающий, заключающий в себе.

Encompassing principle: принцип охвата.

Eigenvalue

Приставка *eigen-*, используемая в английских словах *eigenvalue* и *eigenvector* имеет немецкое происхождение и означает «собственный». Эти термины — частичный перевод с немецкого, где они имеют вид *Eigenwert* и *Eigenvektor*.

Eigenvalue: собственное значение, собственное (характеристическое) число.

Eigenvector: собственный вектор.

Marginal

В эконометрику термин **marginal** попал разными путями. С одной стороны, это общий статистический термин, относящийся к распределениям нескольких случайных величин. Например, если есть две случайные величины X и Y , то можно рассматривать их совместное распределение, условное распределение одной переменной относительно другой, а также безусловное распределение одной переменной. Последнее распределение называется в таком контексте маргинальным¹ или частным. Кроме распределений, маргинальными могут быть плотности, вероятности и т. п. С этим значением связаны производные термины **marginalize** и **marginalization**.

С другой стороны, в экономической теории есть понятие предельного влияния одной переменной на другую (предельная полезность, предельная производительность). С этим связано понятие предельного влияния (эффекта) объясняющей переменной на ожидаемое значение или другую характеристику объясняемой переменной в эконометрических моделях. Для модели линейной регрессии линейной по регрессорам предельный эффект регрессора — это его коэффициент. Но в более общем случае коэффициенты эконометрической модели могут не иметь такой очевидной интерпретации и тогда для анализа результатов оценивания имеет смысл использовать подходящим образом определенное предельное значение (маргинальное значение). Например в модели логит это может быть производная расчетной вероятности по регрессору.

Есть также использование термина **marginal** в общем значении «граничный», например **marginal significance level**.

Margin: край, граница, грань, предел, разница, преимущество.

Marginal: маргинальный, маржинальный, частный (в статистике), предельный (в экономической теории), граничный.

Marginal distribution: маргинальное (маржинальное) распределение, частное распределение, безусловное распределение (в отличие от условного).

Marginal (сущ.): маргинальное (маржинальное) распределение.

Joint distribution: совместное распределение.

Conditional distribution: условное распределение.

Unconditional distribution: безусловное распределение.

Marginalize: маргинализировать, получать маргинальное распределение из совместного распределения.

Marginalization: маргинализация, получение маргинального распределения из совместного.

Marginal effect: предельное (маргинальное) влияние, предельный (маргинальный) эффект.

Marginal value: предельное (маргинальное) значение.

Marginal significance level: граничный уровень значимости.

On the margin: на грани, с небольшой разницей.

¹Относительно термина «маргинальный» (буквально «лежащий на краю/границе») в русском языке есть неоднозначность. Вариант «маргинальный» отражает заимствование из средневековой латыни (наряду с «маргиналиями» — заметками на полях). Несколько менее распространен вариант «маржинальный», соответствующий терминам «маржа», «маржинализм», заимствованным из французского языка и отражающим французское произношение.

Smoothing, bandwidth, kernel

Непараметрические методы оценивания, дающие сглаженную оценку некоторой функции (плотности, регрессии или спектральной плотности), обычно имеют параметр, от которого зависит степень гладкости оценки. Для так называемых ядерных оценок естественным параметром сглаживания является **bandwidth** — ширина полосы или ширина окна, отвечающая за масштабирование ядерной функции. Ядерная оценка в данной точке — это взвешенное среднее несглаженных данных, причем вес, задаваемый масштабированной ядерной функцией, убывает до нуля при удалении от точки. «Далекие» данные входят с нулевым или незначительным весом. (Встречаются и волнообразные ядерные функции, и функции с областями отрицательных значений, но и они в пределе стремятся к нулю.) Когда ширина полосы большая, в оценке используется широкая окрестность данной точки и оценка получается гладкой, т. е. ее дисперсия низкой.

Название по-видимому связано с обработкой сигналов и спектральным анализом. **Band** в этом контексте — это полоса (т. е. интервал) частот, а **bandwidth** — это ширина такой полосы. При сглаживании в частотной области — получении ядерной оценки спектральной плотности на основе периодограммы — ядерную функцию принято называть (спектральным) окном. Отсюда термин «ширина окна» в русском языке. В английском языке **window width** тоже встречается, но не так часто. Хотя ядерные оценки плотности и ядерные регрессии не имеют отношения к полосе частот, но для соответствующего параметра сглаживания используют тот же термин **bandwidth**.

О непараметрической регрессии можно прочесть в Härdle (1990) (русский перевод Хардле 1993) и Расин (2008), о сглаженных оценках спектральной плотности — в Jenkins & Watts (1998) (русский перевод Дженкинс & Ваттс 1971) и Hamilton (1994).

Smoothing: сглаживание.

Undersmoothing/oversmoothing: недостаточное/избыточное сглаживание.

Smoothing parameter: параметр сглаживания.

Kernel: ядро.

Kernel estimation: ядерное оценивание.

Kernel function: ядерная функция.

Spectral window: спектральное окно.

Bandwidth: ширина полосы, ширина окна.

Window width: ширина окна.

Back-testing

Термин **back-testing** пришел из финансов. Если, например, трейдер хочет использовать новую торговую стратегию, то прежде, чем работать с реальными денежными средствами, ему имеет смысл опробовать ее на исторических данных. В эконометрике этот термин применяют к прогнозным моделям, когда проверяется, насколько качественные прогнозы некоторого (обычно финансового) показателя данная модель давала бы, если бы была использована в прошлом. В терминологии Дж. Стока и М. Уотсона (Stock & Watson 2011) такая имитация работы прогнозной модели в реальном времени — это псевдо-вневыборочное прогнозирование (**pseudo out-of-sample forecasting**).

Back-testing: обратное тестирование, историческое тестирование, бэктестинг.

Out-of-sample forecasting: вневыборочное прогнозирование.

Time domain, frequency domain

Само слово **domain** общенаучное и может переводиться как «область» или «сфера». В частности, в математике это область определения функции.

Разделение **time domain** и **frequency domain** (временная область и частотная область) пришло из инженерных приложений, в которых сигнал (например, звуковой) можно представить как функцию времени, т. е. временной ряд, а можно как функцию от частоты. Такие понятия как «лаг», «автоковариация», «автокорреляция» и т. п. относятся к временной области, а «периодограмма», «спектр» — к частотной. Слабо стационарный случайный процесс можно представить как во временной, так и в частотной области и описание процесса в одной области имеет аналоги в другой области. Связь между двумя областями определяется преобразованием Фурье. Во многом выбор области — это вопрос простоты представления; см. Hamilton (1994, p. 152). Например, определение причинности по Грейнджеру (Granger 1969) было изначально дано в частотной области, но сейчас практически всегда с ней работают во временной области.

Time domain: временная область.

Frequency domain: частотная область.

Domain: область, сфера.

Domain (of function): область определения функции.

Spectrum: спектр.

Spectral: спектральный.

Spectral density: спектральная плотность.

Power spectrum: спектр мощности.

Periodogram: периодограмма.

Fourier transform (transformation): преобразование Фурье.

Cross-spectrum: кросс-спектр.

Bispectrum: биспектр.

Granger causality: причинность/обусловленность по Грейнджеру.

Список литературы

Дженкинс, Г. & Д. Ваттс (1971). *Спектральный анализ и его приложения*. Москва: Мир.

Расин, Дж. (2008). Непараметрическая эконометрика: вводный курс. *Квантиль* №4, 7–56.

Хардле, В. (1993). *Прикладная непараметрическая регрессия*. Москва: Мир.

Gourieroux, C. & A. Monfort (1994). Testing non-nested hypotheses. Глава 44 в *Handbook of Econometrics* под редакцией R. F. Engle & D. L. McFadden, Elsevier, том 4, стр. 2583–2637.

- Granger, C. W. J. (1969). Investigating causal relations by econometric models and cross-spectral methods. *Econometrica* 37, 424–438.
- Greene, W.H. (2012). *Econometric Analysis* (7th edition). Boston: Prentice Hall.
- Hamilton, J. D. (1994). *Time Series Analysis*. Princeton: Princeton University Press.
- Härdle, W. (1990). *Applied Nonparametric Regression*. Cambridge: Cambridge University Press.
- Jenkins, G. M. & D. Watts (1998). *Spectral Analysis and Its Applications*. Boca Raton: Emerson Adams.
- Pesaran, M. H. & M. Weeks (2001). Nonnested hypothesis testing: An overview. Глава в *A Companion to Theoretical Econometrics* под редакцией В. Н. Baltagi, Blackwell Publishing, 279–309.
- Stock, J. H. & M. W. Watson (2011). *Introduction to Econometrics* (3rd edition). Amsterdam: Addison-Wesley Longman.

A mini-dictionary of English econometric terminology III

Alexander Tsyplakov

Novosibirsk State University, Novosibirsk, Russia

This part of the dictionary comments on English econometric terms estimator, quantile, nested, marginal, and some others. Emphasis is again placed on accurate definitions of their meaning to avoid possible confusion and incorrect interpretation.

