Тестирование распределений*

Игорь Хейфец†

Российская экономическая школа, Москва, Россия

В эссе рассмотрены методы тестирования предположений о форме функций распределений при помощи теории эмпирических процессов. На примерах показаны трудности при применении таких методов — эффект от оценивания параметра, зависимость асимптотики от распределения — и способы их преодоления — мартингальное преобразование, бутстрап. Тесты применимы, в частности, к GARCH- и диффузионным моделям.

1 Введение

Методы оценивания и прогнозирования эконометрических моделей основываются на некоторых предположениях о моделях. Если эти предположения неверны, мы не можем быть уверены в полученных результатах. Оценки могут уже не быть эффективными или даже стать несостоятельными. Например, оценка методом максимального правдоподобия (ММП) при неверном предположении, что распределение случайных величин (СВ) является нормальным, перестает быть эффективной, а при неверном предположении, что распределение является t_5 (т.е. распределением Стьюдента с 5 степенями свободы), перестает быть даже состоятельной. Другие примеры последствий плохой спецификации см. в Крил (2008). Поэтому важно научиться тестировать предпосылки, на которые опираются модели.

В данном эссе мы рассмотрим некоторые методы тестирования предположений о функциональной форме функций распределений (ФР), называемыми для краткости критериями согласия (англ. goodness-of-fit). Проверку другого типа предпосылок, таких как о функциональной форме среднего и дисперсии, форме связей между переменными (например, линейности), симметричности плотностей, независимости или некоррелированности ошибок с регрессорами и т.п., мы напрямую здесь рассматривать не будем. На практике знать функциональную форму ФР важно не только для применения различных методов оценки и прогнозирования динамики курсов акций, индексов, обменных курсов, процентных ставок, но также и при непосредственном моделировании распределения, например, распределения неравенства доходов среди населения или распределения доходностей портфелей.

2 Безусловные критерии согласия

Пусть у нас имеется выборка независимых СВ $X_1, X_2, ..., X_n$ длины n. Мы хотим протестировать простую гипотезу, что X_i имеют распределение $F(x)$:

$$H_0: \quad X_i \sim F \quad \forall\, i = 1, ..., n, \tag{1}$$

против любой альтернативы. Определим *эмпирическую* (или выборочную) ФР X_i как

$$F_n(x) = \frac{1}{n}\sum_{i=1}^{n} \mathbb{I}\{X_i \leq x\},$$

где $\mathbb{I}\{\cdot\}$ — индикатор-функция. Если гипотеза верна, то $F_n(x)$ должна быть близка к предполагаемой $F(x)$.

*Цитировать как: Хейфец, Игорь (2011) «Тестирование распределений», Квантиль, №9, стр. 25–34. Citation: Kheifets, Igor (2011) “Goodness-of-fit testing,” Quantile, No.9, pp. 25–34.

†Адрес: 117418, г. Москва, Нахимовский проспект, 47, офис 2114. Электронная почта: *ikheifets@nes.ru*

Можно определить «близость» двух функций как

$$D_n = \sup_x |F_n(x) - F(x)|, \tag{2}$$

т. е. взять sup-метрику (не вдаваясь в подробности, как определить пространства функций, чтобы это была метрика в строгом смысле). По теореме Гливенко–Кантелли $D_n \xrightarrow{\text{п.н.}} 0$ (почти наверное) при $n \to \infty$. Далее,

$$\sqrt{n} D_n \xrightarrow{d} \sup_x |B(F(x))|$$

при $n \to \infty$, где $B(\cdot)$ — Броуновский мост. Если $F(x)$ непрерывна, то предельное распределение $\sqrt{n}D_n$ не зависит от $F(x)$, совпадает с распределением $\sup_{x\in[0,1]} |B(x)|$ и равно

$$K(x) = 1 - 2\sum_{i=1}^{\infty}(-1)^{i-1}e^{-2i^2x^2} = \frac{\sqrt{2\pi}}{x}\sum_{i=1}^{\infty} e^{-(2i-1)^2\pi^2/(8x^2)} \tag{3}$$

(Kolmogoroff, 1933), что называется *распределением Колмогорова. Свобода от распределения* (англ. distribution free) — важное свойство статистики, позволяющее получить критические значения для любого n вне зависимости от $F(x)$. Этот тест называют *критерием Колмогорова* или *критерием Колмогорова–Смирнова.* Если взять ω^2-метрику

$$\omega_n^2 = \int_{-\infty}^{\infty} [F_n(x) - F(x)]^2 \,\mathrm{d}F(x),$$

то получится *тест Крамера–фон Мизеса* (также известного как *критерий* ω^2 или *критерий Смирнова*). В зависимости от метрики свойства критериев могут меняться. Мощность будет зависеть также от типа альтернативы, и сказать, какой критерий оптимален в общем случае, нельзя. Однако, такие критерии *состоятельны против всех альтернатив* (от лат. *omnibus* — «для всех»), т. е. любое отклонение от истинного распределения будет замечено при достаточно большой выборке, в противоположность критериям, которые хорошо работают против одних альтернатив и не работают против остальных.

Заметим, что если гипотеза (1) верна, и $F(x)$ непрерывна, то СВ

$$U_i = F(X_i) \sim U[0,1] \tag{4}$$

равномерно распределены на $[0,1]$. На этом факте основан альтернативный способ вывода статистики Колмогорова. Рассмотрим *эмпирический процесс*

$$V_n(u) = \sqrt{n}\left[F_n^U(u) - u\right], \quad u \in [0,1], \tag{5}$$

где $F_n^U(u)$ — эмпирическая ФР U_i. Этот процесс сравнивает эмпирическую и предполагаемую ФР для U_i. При нулевой гипотезе эта разность должна быть маленькой для всех $u \in [0,1]$. Если взять sup-метрику, то мы получим статистику $\sqrt{n}D_n$ для равномерных распределений.

Для многомерных случайных величин ситуация более сложная. Пусть у нас имеется выборка независимых случайных векторов размерности K

$$X_1, X_2, ..., X_n,$$

где $X_i = (X_{i1}, X_{i2}, ..., X_{iK})$, $i = 1, ..., n$, и мы хотим протестировать гипотезу (1), что X_i имеют совместное распределение $F(x)$, $x = (x_1, x_2, ..., x_K) \in R^K$. Пусть $F_n(x)$ обозначает совместную эмпирическую ФР

$$F_n(x) = \frac{1}{n}\sum_{i=1}^{n}\prod_{k=1}^{K}\mathbb{I}\{X_{ik} \le x_k\}.$$

Если гипотеза верна, то непараметрическая оценка ФР $F_n(x)$ должна быть близка к $F(x)$. Для $K > 1$ распределение sup- и ω^2-статистик зависит от $F(x)$, то есть нарушена свобода от распределения.

Rosenblatt (1952) предложил следующее преобразование (*преобразование Розенблатта*):

$$U_{ik} = \Pr\{X_{ik}|X_{ik-1}, ..., X_{i1}\} \quad \forall\, i = 1, ..., n,\ k = 1, ..., K. \tag{6}$$

Если $X_i \sim F$, то $U_i = (U_{i1}, U_{i2}, ..., U_{iK}) \sim U[0,1]^K$ равномерно распределены на единичном кубе размерности K. Этим можно воспользоваться для построения теста. Иногда условные вероятности в (6) известны в явном виде. Например, их легко получить для многомерного нормального распределения.

На практике вместо проверки конкретной ФР необходимо научиться проверять принадлежность к параметрическому семейству ФР $\{F(x, \theta),\ \ \theta \in \Theta\}$. Например, ММП дает состоятельные и асимптотически эффективные оценки для всего семейства нормальных данных. Рассмотрим *сложную гипотезу*

$$H_0 : \ \ \exists\, \theta_0\ \ X_i \sim F(x, \theta_0) \quad \forall\, i = 1, ..., n. \tag{7}$$

Если гипотеза (7) верна, то непараметрическая оценка $F_n(x)$ должна быть близка к параметрической оценке $F(x, \hat\theta)$, где $\hat\theta$ — $\sqrt{n}$-состоятельная оценка истинного параметра θ_0.

Вернемся к одномерному случаю. Так как $F(x, \hat\theta)$ не совпадает с $F(x, \theta_0)$, $\hat U_i = F(X_i, \hat\theta)$ не распределены равномерно, в отличие от U_i в (4). Эмпирический процесс

$$\hat V_n(u) = \sqrt{n}\left[F_n^{\hat U_i}(u) - u\right], \quad u \in [0,1], \tag{8}$$

и статистика $\hat D_n = \sup_{u\in[0,1]} |\hat V_n(u)|$ имеет при каждом новом значении параметра распределение, отличное от D_n, т.е. мы наблюдаем *эффект от оценивания параметров*, и нельзя использовать критические значения (3) (Durbin, 1973). Используя теорию слабой сходимости эмпирических процессов (см. Billingsley, 1968), обычно можно вывести, как изменится процесс, построенный на оценках на месте истинных параметров. Например, для (8) можно показать, что равномерно по $u \in [0, 1]$

$$\hat V_n(u) = V_n(u) - g(u)\sqrt{n}(\hat\theta - \theta_0) + o_p(1), \tag{9}$$

где $g(u)$ — некоторая функция, зависящая от производной $F(x, \theta)$ по параметру (Bai, 2003). Обычно используют один из двух способов получения критических значений. Первый способ заключается в преобразовании $\hat V_n(u)$, убирающем дополнительный член в разложении (9). См. например, мартингальное преобразование, описанное в Khmaladze (1981) и примененное в Bai (2003), Delgado & Stute (2008) и др. Распределение преобразованного процесса и статистики можно затабулировать, но сами преобразования зависят от исходных моделей. Второй способ заключается в бутстрапировании $\hat D_n$, см. Andrews (1997) и Corradi & Swanson (2005). Далее мы рассмотрим подробнее оба этих способа.

3 Тестирование условных распределений

3.1 Обобщенный критерий Колмогорова и параметрический бутстрап

Andrews (1997) обобщил критерий Колмогорова для тестирования условных распределений. Пусть дана выборка независимых одинаково распределенных (НОР) случайных векторов $Z_i = (Y_i', X_i')' \in R^{V+K}$. Рассмотрим гипотезу

$$H_0 : \ \ \exists\, \theta_0\ \ Y_i|X_i \sim F(y|x, \theta_0) \quad \forall\, i = 1, ..., n. \tag{10}$$

Пусть $H_n(z)$ и $G_n(x)$ — эмпирические ФР Z_i и X_i соответственно. Рассмотрим полупараметрическую оценку совместной ФР Z_i, основанную на гипотетической условной $F(y|x,\theta_0)$ и эмпирической маржинальной $G_n(x)$:

$$F_n(z,\theta) = \frac{1}{n}\sum_{i=1}^{n} F(y|X_i,\theta)\mathbb{I}\{X_i \leq x\}, \quad z = (y,x)' \in R^{V+K}.$$

Сравнивая полупараметрическую (опирающуюся на $\sqrt{n}$-состоятельную оценку $\hat{\theta}$) и эмпирическую ФР, получаем статистику

$$CK_n = \sqrt{n}\max_{i\leq n}\left|H_n\left(Z_i\right) - F_n(Z_i,\hat{\theta})\right|.$$

В отличие от статистики Колмогорова (2), здесь достаточно брать максимум по конечному числу элементов выборки, что значительно сокращает время вычислений. Распределение статистики подвержено эффекту от оценивания параметров. Критические значения можно получить с помощью параметрического бутстрапа.

На примере статистики CK_n опишем подробнее алгоритм получения критических значений с помощью бутстрапа.[1] Задача бутстрапа состоит в приближении распределения статистики для СВ предполагаемым распределением вне зависимости, совпадает ли оно с истинным или нет. Эмпирическое распределение H_n для этих целей не подходит, так как не генерирует предполагаемое распределение, если гипотеза (10) не верна. Естественным кандидатом для генерирования Y_i является гипотетическое условное распределение $F(y|x,\theta_0)$, или $F(y|x,\hat{\theta})$, если параметр неизвестен. Регрессоры X_i можно сгенерировать из G_n или брать одни и те же исходные данные для каждой симуляции (*бутстрап с фиксированными регрессорами*). Основываясь на симуляциях Монте Карло, Andrews (1997) сделал вывод, что бутстрап с фиксированными регрессорами работает лучше. Распределение СВ, получаемых таким бутстрапом, известно с точностью до конечномерного параметра θ, поэтому процедура называется параметрическим бутстрапом.

В итоге параметрический бутстрап выглядит следующим образом:

1. По данным $Z_i = (Y_i', X_i')'$ получаем оценку параметра $\hat{\theta}$ и статистику CK_n.
2. С помощью $F(y|x,\hat{\theta})$, условно на X_i, генерируем Y_{bi}.
3. По $Z_{bi} = (Y_{bi}', X_i')'$ получаем оценку параметра $\hat{\theta}^b$ и статистику CK_n^b.
4. Повторяем шаги 2–3 для $b = 1,...,B$. Считываем p-значение статистики CK_n, равное пропорции CK_n^b, больших CK_n.

Кодирование такого алгоритма не составляет труда. Однако, время исполнения программы может оказаться ощутимым. Для теоретического обоснования применимости бутсрапа Andrews (1997) вывел для своей статистики разложение типа (9) и показал, что оно асимптотически одинаково как при нулевой гипотезе, так и при использовании сгенерированных бутстрапом данных.

3.2 Динамические модели и мартингальное преобразование

Ослабим теперь предположение, что Z_i являются НОР. А именно, предположим, что их распределение может меняться с i, и Y_i могут быть зависимы. Для простоты пусть $V = 1$. Маржинальные распределения не учитывают ни кросс-, ни временных зависимостей, поэтому при работе с кросс-секциями и временными рядами моделируют условные распределения.

[1] Обзоры бутстрапа см. в журнале *Квантиль* №3 (2007).

Пусть Ω_i — вся информация, доступная нам «сегодня», т.е. минимальная σ-алгебра, содержащая $Z_1, Z_2, ..., Z_{i-1}, X_i$. Рассмотрим гипотезу

$$H_0: \quad \exists\, \theta_0 \quad Y_i|\Omega_i \sim F_i(y|\cdot, \theta_0) \quad \forall\, i = 1, ..., n. \tag{11}$$

Эту гипотезу можно переформулировать в терминах оценок прогноза ФР, см. Diebold, Gunther & Tay (1998), в противоположность оценкам точечного и интервального прогнозов.

С помощью гипотезы (11) можно протестировать большое количество динамических моделей. Приведем пример одного класса таких моделей. Пусть

$$Y_i = \mu(i, \theta) + \sigma(i, \theta)\varepsilon_i, \tag{12}$$

где $\mu(i, \theta)$ и $\sigma(i, \theta)$ Ω_i-измеримы, и ε_i — НОР с ФР $F_{\varepsilon i}(\varepsilon, \theta)$. Этот класс включает в себя популярные модели гетероскедастичности. Скажем, модель AR(1)-GARCH(1,1)-N может быть представлена в виде

$$\begin{aligned} \mu(i, \theta) &= \omega_0 + \omega_1 Y_{i-1}, \\ \sigma^2(i, \theta) &= \alpha_0 + \beta_1 \sigma^2(i-1, \theta) + \alpha_1 \left(Y_{i-1} - \omega_0 - \omega_1 Y_{i-2}\right)^2, \end{aligned}$$

$F_{\varepsilon i}(\varepsilon, \theta) \sim N\left(\mu_\varepsilon, \sigma_\varepsilon^2\right)$, и $\theta = (\omega_0, \omega_1, \alpha_0, \alpha_1, \beta_1, \mu_\varepsilon, \sigma_\varepsilon)$. В то же время, не все параметрические модели можно записать в форме условного распределения. Например, GARCH-модели с не специфицированными ФР инноваций, GARCH-модели с некоррелированными (но, возможно, зависимыми) инновациями, SV-модели.[2] Для модели (12) легко выписать соответствующее условное распределение

$$F_i(y|\Omega_i, \theta) = F_{\varepsilon i}\left(\frac{y - \mu(i, \theta)}{\sigma(i, \theta)}, \theta\right).$$

Если гипотеза (11) верна, а $F(y|\cdot, \theta_0)$ непрерывна по y, то

$$U_i = F_i(Y_i|\Omega_i, \theta_0) \quad \sim \quad \text{НОР} \quad U[0, 1], \tag{13}$$

т.е. независимы и равномерно распределены на $[0, 1]$. Diebold, Gunther & Tay (1998) предложили визуальные методы проверки (13), а именно, с помощью гистограмм и коррелограмм первых четырех центральных моментов. Эти методы дают хорошее наглядное представление о спецификации, однако не могут учесть эффект от оценивания параметров. Действительно, при проверке сложной гипотезы мы не знаем истинное значение параметра θ_0. Пусть у нас есть $\sqrt{n}$-состоятельная оценка $\hat{\theta}$. Тогда $\hat{U}_i = F_i(Y_i|\Omega_i, \hat{\theta})$, так же как и в безусловном случае, не обязаны быть равномерными и независимыми даже при правильной спецификации.

Формальное правило проверки (13) можно осуществить с помощью теории эмпирических процессов. Используя $\hat{U}_i$, построим эмпирическую ФР $F_n^{\hat{U}}(u)$ и процесс $\hat{V}_n(u)$, см. (8). Bai (2003) вывел для этого случая аналог разложения (9) и предложил использовать мартингальное преобразование (Khmaladze, 1981) для получения критических значений с учетом эффекта от оценивания параметра. А именно, при дополнительных предположении о непрерывной дифференцируемости ФР по параметру, разложение (9) имеет место с

$$g(u) = \operatorname*{plim}_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} \left.\frac{\partial F_i(x|\Omega_i, \theta_0)}{\partial \theta}\right|_{x = F_i^{-1}(u|\Omega_i, \theta_0)}, \tag{14}$$

что позволяет увидеть в явном виде и обработать эффект от оценивания параметров. Например, мартингальное преобразование выглядит следующим образом (Bai, 2003):

$$\hat{W}_n(u) = \hat{V}_n(u) - \int_0^u \left(\bar{g}(s)' C^{-1}(s) \int_s^1 \bar{g}(\tau) d\hat{V}_n(\tau)\right) ds,$$

[2]См. обзор по GARCH моделям в Росси (2010), а по SV моделям — в Цыплаков (2010).

где $\bar{g}(u) = (1, \dot{g}(u)')'$, $\dot{g}(u)$ — производная $g(u)$ и $C(u) = \int_u^1 \bar{g}(\tau)\bar{g}(\tau)'d\tau$. Процесс $\hat{W}_n(u)$ слабо сходится к Броуновскому процессу, поэтому, опираясь на теорему о непрерывном преобразовании, распределение статистик на основе преобразованного процесса легко просимулировать и затабулировать. Сложность применения данного метода заключается в том, что выражение для $g(u)$ в (14) надо вычислять явно для каждой ФР. Примеры таких аналитических формул для GARCH(1,1), ARCH(p,q) и нелинейных динамических регрессий приведены в Bai (2003).

Теперь покажем, что тесты на базе V_n обнаруживают *локальные $\sqrt{n}$-альтернативы*, то есть распределения, приближающиеся к истинным со скоростью $\sqrt{n}$:

$$G_{ni}\left(y|\Omega_i,\theta_0\right) = \left(1 - \frac{\delta}{\sqrt{n}}\right) F_i\left(y|\Omega_i,\theta_0\right) + \frac{\delta}{\sqrt{n}} H_i\left(y|\Omega_i,\theta_0\right),$$

где $\delta > 0$, $1 < \delta/\sqrt{n}$, а H_i — условное распределение отличное от F_i, так что

$$k(u) = \operatorname*{plim}_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} H_i\left(F_i^{-1}\left(u|\Omega_i,\theta_0\right)|\Omega_i,\theta_0\right) - u \neq 0.$$

При локальной альтернативе новые СВ $U_i^* = G_{ni}(Y_i|\Omega_i,\hat{\theta})$ равномерно распределены и независимы, поэтому теперь нам понадобится разложение $\hat{V}_n(u)$ вокруг $V_n^*(u)$, построенному по U_i^*

$$V_n^*(u) = \sqrt{n}\left[F_n^{U_i^*}(u) - u\right], \quad u \in [0,1],$$

и слабо сходящемуся к Броуновскому процессу. Bai (2003) установил аналог (9) для локальной альтернативы:

$$\hat{V}_n(u) = V_n^*(u) - g(u)\sqrt{n}(\hat{\theta} - \theta_0) + \delta k(u) + o_p(1)$$

равномерно по $u \in [0,1]$. Если теперь применить мартингальное преобразование, то получим, что $\hat{W}_n(u)$ слабо сходится к Броуновскому процессу $W(u)$ плюс добавка:

$$\hat{W}_n(u) \Rightarrow W(u) + \delta k(u) - \delta\phi_g(k)(u),$$

где $\phi_g(k)(u) = \int_0^u (\dot{g}(s)'C^{-1}(s)\int_s^1 \dot{g}dk)ds$. Таким образом, если удается показать, что добавка ненулевая, то критерий обнаруживает локальные альтернативы.

Описанный тест имеет существенный недостаток. Легко видеть, что процесс (8) проверяет в явном виде равномерность, но не зависимость U_i. Поэтому такой тест плохо отлавливает динамику и зависимости в моделях. Например, если истинный процесс является авторегрессионным второго порядка, AR(2) с нормальными ошибками, а предполагаемый — AR(1), то такой тест имеет лишь номинальную мощность. Тест не работает и в общем случае при нарушении динамики и эллиптических ошибках (Corradi & Swanson, 2005). Более формально, (8) проверяет (11) в среднем, т. е. правильнее говорить о

$$H_0: \quad \exists\,\theta_0 \quad E\left(F_i\left(y|\Omega_i,\theta_0\right)\right) = F_i(y) \quad \forall\, i = 1,...,n.$$

Однако при такой нулевой гипотезе не гарантируется независимость U_i, меняются предельные распределения статистик, а это значит, что ни параметрический бутстрап, ни мартингальное преобразование использовать нельзя. Для проверки такой гипотезы Corradi & Swanson (2005) предложили приблизить предельное распределение статистики D_n, построенной по (8) с помощью блочного бутстрапа.

Для случая НОР Z_i и $V = 1$, Delgado & Stute (2008) предложили использовать двумерное преобразование Розенблатта (6). Если гипотеза (11) верна, преобразованные $U_i = F(X_i,\theta_0)$ и $V_i = F(Y_i|X_i,\theta_0)$ независимы и равномерно распределены на единичном квадрате $(U_i,V_i) \sim$

$U[0,1]^2$. Это можно проверить сравнивая ФР, оцененную непараметрически и обозначенную как $F_n^{U,V}(u,v)$, и предполагаемую совместную ФР (U_i, V_i):

$$V_{2n}(u,v) = \sqrt{n}\left[F_n^{U,V}(u,v) - uv\right], \quad u,\ v \in [0,1]. \tag{15}$$

Этот процесс слабо сходится к двумерному Броуновскому мосту. Однако, при тестировании сложной гипотезы наблюдается эффект от оценивания параметра. Для борьбы с этим Delgado & Stute (2008) используют мартингальное преобразование к Броуновскому мосту.

Для тестирования в моделях временных рядов без экзогенных регрессоров, т. е. для проверки (11) при зависимых Y_i и $V = 0$, Kheifets (2010) предложил использовать процесс

$$V_{2n}(u_1,u_2) = \sqrt{n}\left[F_n^{U_{-1},U}(u_1,u_2) - u_1u_2\right], \quad u_1,\ u_2 \in [0,1], \tag{16}$$

где $F_n^{U_{-1},U}$ обозначает совместную эмпирическую ФР (U_i, U_{i-1}), U_i получены в (13). Можно рассматривать также и другие лаги (U_i, U_{i-j}), $j = 1, ..., i-1$. В отличие от (15), (16) не является суммой независимых слагаемых, что не позволяет применять стандартные предельные теоремы. Kheifets (2010) изучил слабую сходимость процессов (16) и доказал законность использования параметрического бутстрапа для приближения критических значений sup-статистики

$$D_{2n} = \sup_{(u_1,u_2)\in[0,1]^2} |V_{2n}(u_1,u_2)|$$

и ω^2-статистики

$$\omega_{2n}^2 = \int_0^1\int_0^1 V_{2n}^2(u_1,u_2)du_1du_2.$$

4 Тестирование плотностей и диффузий

Для тестирования плотностей можно использовать подходы, заключающиеся в сравнении параметрических, зависящих от модели, и непараметрических оценок плотностей. В качестве непараметрических оценок плотностей берут, например, ядерные или серийные оценки, см. обзор Расин (2008). В отличие от оценок ФР, такие приближения сходятся медленнее, и свойства тестов зависят от настроек непараметрической оценки (например, выбор ядерной функции и ширины окна). Такие критерии не могут отличить $\sqrt{n}$-локальные альтернативы. В то же время, благодаря меньшей, чем $\sqrt{n}$ скорости сходимости, нет эффекта от оценивания параметра при тестировании сложных гипотез.

Для оценки производных ценных бумаг (деривативов) используются модели непрерывного времени, в которых величины задаются как решения стохастических дифференциальных уравнений. Как проверять спецификацию таких моделей? Рассмотрим, как применять описанные методы для тестирования диффузионных процессов

$$dX_t = \mu_0(X_t,t)dt + \sigma_0(X_t,t)dW_t, \tag{17}$$

где W_t — стандартный Броуновский процесс. Тестирование параметрической диффузии заключается в проверке того, что дрейф $\mu_0(X_t,t)$ и диффузия $\sigma_0(X_t,t)$ принадлежат параметрическим семействам $\mu(X_t,t,\theta)$ и $\sigma(X_t,t,\theta)$, $\theta \in \Theta$ с вероятностью 1. Такая спецификация включает в себя большое количество моделей. Например, модель Vasicek (1977) задается как

$$\begin{aligned}\mu(x,t,\theta) &= \beta(\alpha - x),\\ \sigma(x,t,\theta) &= \gamma^2,\end{aligned}$$

и $\theta = (\alpha, \beta, \gamma)$, а модель Cox, Ingersoll & Ross (1985) задается как

$$\begin{aligned} \mu(x,t,\theta) &= \beta(\alpha - x), \\ \sigma(x,t,\theta) &= \gamma^2 x, \end{aligned}$$

и $\theta = (\alpha, \beta, \gamma)$.

Модель диффузии характеризует маржинальную плотность X_t, $\pi_0(x)$, и транзитивные плотности $p_0(x,t|y,s)$, $\;s < t$ (условная плотность $X_t = x$ при $X_s = y$), поэтому ее тестирование можно свести к тестированию плотностей. В качестве данных используют $\{X_{\tau\Delta}\}_{\tau=1}^n$, полученные дискретизацией временного периода $[0,T]$ интервалами длины Δ из непрерывной модели для X_t. Ait-Sahalia (1996) предложил тестировать диффузии, проверяя маржинальные плотности

$$H_0: \;\; \exists\, \theta_0 \;\; \pi(x,\theta_0) = \pi_0(x) \quad \text{почти всюду.}$$

Маржинальная плотность выражается в явном виде через дрейф и диффузию

$$\pi(x,\theta) = \frac{\xi(\theta)}{\sigma(x,\theta)} \exp\left[\int_{x_0}^{x} \frac{2\mu(u,\theta)}{\sigma(u,\theta)} du\right],$$

где выбор x_0 и $\xi(\theta)$ обеспечивает интегрирование плотности к 1. Непараметрическая ядерная оценка этой же плотности дает

$$\hat{\pi}_0(x) = \frac{1}{n} \sum_{\tau=1}^{n} \frac{1}{b_n} K\left(\frac{u - X_{\tau\Delta}}{b_n}\right),$$

где $K(\cdot)$ и b_n — ядерная функция и ширина окна. Ait-Sahalia (1996) установил асимптотическую нормальность статистики

$$M = n b_n \min_{\theta} \frac{1}{n} \sum_{\tau=1}^{n} \left(\pi(X_{\tau\Delta,\theta}) - \hat{\pi}_0(X_{\tau\Delta})\right)^2 .$$

Как и критерий, предложенный в Bai (2003), такой тест плохо контролирует динамику.

Вместо маржинальных плотностей можно рассматривать переходные плотности, которые полностью определяют динамику. Рассмотрим

$$H_0: \;\; \exists\, \theta_0 \;\; p(x,t|y,s,\theta_0) = p_0(x,t|y,s) \quad \text{п.н.} \tag{18}$$

Транзитивные плотности плохо оцениваются непараметрически из-за большой размерности. Однако, при нулевой гипотезе (18)

$$U_\tau = \int_{-\infty}^{\tau\Delta} p\left(x, \tau\Delta | X_{(\tau-1)\Delta}, (\tau-1)\Delta, \theta_0\right) dx \quad \sim \quad \text{НОР} \;\; U[0,1]. \tag{19}$$

Hong & Li (2005) использовали для проверки (19) разность ядерной оценки совместной плотности $(U_\tau, U_{\tau-1})$, $\hat{g}(u_1, u_2)$ и единицы — плотности равномерного распределения

$$Q = \int_0^1 \int_0^1 \left[\hat{g}(u_1, u_2) - 1\right]^2 du_1 du_2.$$

Можно рассматривать также и другие лаги $(U_\tau, U_{\tau-j})$, $j = 1, ..., \tau - 1$. После должной нормировки предельное распределение оказывается стандартным нормальным, но скорость сходимости ниже, чем $\sqrt{n}$.

5 Заключение

В данном эссе мы рассмотрели различные методы тестирования предположений о функциональной форме ФР и плотностей. Часто такие критерии основываются на эмпирическом процессе, сравнивающем две оценки подходящей ФР: непараметрическую, сходящуюся к истинной ФР всегда, и параметрическую, сходящуюся к истинной ФР только если предположение верно.

Не всегда возможно затабулировать критические значения таких тестов. Распределение может зависеть от распределения и от факта оценивания параметров. В этом случае используют либо преобразование процесса (например, мартингальное преобразование к Броуновскому процессу), либо бутстрапируют критические значения. Мы видели примеры применения параметрического бутстрапа для обобщенного критерия Колмогорова и мартингального преобразования для табулирования критических значений теста в динамических моделях. В обоих случаях важную роль играет разложение оцененного процесса вокруг истинного. Это разложение также важно для вывода распределения процесса при локальной альтернативе.

В основном, мы рассматривали одномерные модели. Для тестирования многомерных динамических моделей можно использовать аналог преобразования Розенблатта, см., например, Diebold, Hahn & Tay (1999), Bai & Chen (2008), Kheifets (2011). Мы также предполагали непрерывность ФР. Velasco & Kheifets (2010) исследовали аналог (13) без предположения о непрерывности ФР и описали критерии для тестирования динамических моделей выбора. Анализ и сравнение других методов тестирования моделей выбора см. в Mora & Moro-Egido (2007).

Рассмотренные тесты должны работать против общих отклонений в противовес параметрическим тестам, заточенным на определенные альтернативы.

Список литературы

Крил, М. (2008). Некоторые ловушки параметрической инференции. *Квантиль* 4, 1–6.

Расин, Дж. (2008). Непараметрическая эконометрика: вводный курс. *Квантиль* 4, 7–56.

Росси, Э. (2010). Одномерные GARCH-модели: обзор. *Квантиль* 8, 1–67.

Цыплаков, А. (2010). Сделать тайное явным: искусство моделирования с помощью стохастической волатильности. *Квантиль* 8, 69–122.

Ait-Sahalia, Y. (1996). Testing Continuous-Time Models of the Spot Interest Rate. *Review of Financial Studies* 9, 385–426.

Andrews, D.W.K. (1997). A conditional Kolmogorov test. *Econometrica* 65, 1097–1128.

Bai, J. (2003). Testing Parametric Conditional Distributions of Dynamic Models. *The Review of Economics and Statistics* 85, 531–549.

Bai, J. & Z. Chen (2008). Testing multivariate distributions in GARCH models. *Journal of Econometrics* 143, 19–36.

Billingley, P. (1968) Convergence of Probability Measures. New York: John Wiley & Sons.

Corradi, V. & R. Swanson (2006). Bootstrap conditional distribution test in the presence of dynamic misspecification. *Journal of Econometrics* 133, 779–806.

Cox, J.C., Ingersoll, J.E. & S.A. Ross (1985). A theory of the term structure of interest rates. *Econometrica* 53, 385–407.

Delgado, M. & W. Stute (2008). Distribution-free specification tests of conditional models. *Journal of Econometrics* 143, 37–55.

Diebold, F.X., Gunther, T. & A.S. Tay (1998). Evaluating density forecasts with applications to finance and management. *International Economic Review* 39, 863–883.

Diebold, F.X., Hahn, J. & A.S. Tay (1999). Multivariate density forecast evaluation and calibration in financial risk management: high frequency returns on foreign exchange. *Review of Economics and Statistics* 81, 661–673

Durbin, J. (1973). Weak convergence of sample distribution functions when parameters are estimated. *Annals of Statistics* 1, 279–290.

Hong, Y. & H. Li (2005). Nonparametric Specification Testing for continuous-Time Models with Applications to Term Structure of Interest Rates. *Review of Financial Studies* 72, 499–541.

Kheifets, I. (2010). Specification tests for nonlinear time series models. Mimeo, Universidad Carlos III de Madrid.

Kheifets, I. (2011). Specification tests for multivariate time series models. Mimeo, Universidad Carlos III de Madrid.

Khmaladze, E.V. (1981). Martingale approach in the theory of goodness-of-tests. *Theory of Probability and its Applications* 26, 240–257.

Kolmogorov, A.N. (1933). Sulla determinazione empirica di una legge di distribuzione. *Giornale dell'Istituto Italiano degli Attuari* 4, 83.

Mora, J. & A.I. Moro-Egido (2007). On specification testing of ordered discrete choice models. *Journal of Econometrics* 143, 191–205.

Rosenblatt, M. (1952). Remarks on a Multivariate Transformation. *Annals of Mathematical Statistics* 23, 470–72.

Vasicek, O. (1977). An equilibrium characterization of the term structure. *Journal of Financial Economics* 5, 177–188.

Velasco, C. & I. Kheifets (2010) Model evaluation of the Fed monetary rules. Mimeo, Universidad Carlos III de Madrid.

Goodness-of-fit testing

Igor Kheifets

New Economic School, Moscow, Russia

We consider goodness-of-fit tests based on the empirical process theory. There are two main obstacles in obtaining critical values for such tests: the parameter estimation effect and distribution dependence. We discuss solutions to these problems: martingale transformation and bootstrap. As an illustration we show how to test GARCH and diffusion models.