А.Л. Васнев

НОВЫЙ НЕПАРАМЕТРИЧЕСКИЙ МЕТОД ПОСТРОЕНИЯ ПСЕВДОВЫБОРКИ ПРИ БУТСТРАПЕ АВТОРЕГРЕССИИ

Препринт #BSP/2001/0.. R

Эта статья основана на дипломной работе, выполненной в РЭШ в 2001 году в рамках исследовательской программы "Преобразование государственного сектора в экономиках переходного периода", финансируемой Фондом Форда, проект "Финансирование капитальных вложений в России". Я благодарен руководителю проекта профессору Станиславу Анатольеву (РЭШ) за неоценимую помощь в работе.

Москва

2001

Васнев А.Л. Новый непараметрический метод построения псевдовыборки при бутстрапе авторегрессии. / Препринт #BSP/2001/0.. R. - М.: Российская экономическая школа, 2001. - 29с.(Англ)

В случае независимых одинаково распределенных данных применение бутстрапа обеспечивает более точные результаты, чем при использовании асимптотического распределения, поскольку бутстрап лучше приближает конечномерное распределение для пивотальных статистик. В контексте авторегрессии наилучшие свойства в большинстве случаев показал бутстрап с перекрывающимися блоками.

В работе предлагается новый непраметрический метод построения бутстраповской выборки при бутстрапе авторегрессии, основанный на приближении истинного процесса генерации данных дискретной марковской цепью. Новый метод сравнивался в симуляциях с асимптотическим методом построения критических областей и блочным бутстрапом с перекрывающимися блоками. Предложенный метод наиболее точно приближает конечномерное распределение, что подтверждается наилучшим совпадением реальных размеров тестов с номинальными.

При помощи нового метода была проверена гипотеза о непредсказуемости реального временного ряда для ВНП США в классе расширения AR моделей. При помощи бутстрапа гипотеза не отвергается, использование асимптотического распределения дает противоположный результат.

Vasnev A.L. A Nonparametric Method for Bootstrap Resampling in Autoregressions. / Working paper #BSP/2001/0.. R. - Moscow, New Economic School, 2001. - 29p.(Engl)

The bootstrap is an alternative to the asymptotic approach and allows to improve statistical inferences in finite samples, since the bootstrap distribution gives more exact approximation of finite sample distribution for the pivotal statistics. In the time series context the moving block bootstrap has one of the best performance.

We propose the non-parametric method for the construction of the bootstrap pseudosample. This method is based on the approximation of the true data generation process by the discrete Markov chain. The method is compared with the asymptotic and the moving block bootstrap in simulation. From the good size properties it may be concluded that the new method gives the best approximation of the finite sample distribution.

The hypothesis of unpredictability of GNP USA in the extension of AR models is tested. The bootstrap does not reject the hypothesis. The using of the asymptotic distribution leads to opposite result.

ISBN 5-8211-0136-0

Содержание

Введение

Бутстрап является альтернативой асимптотическому подходу при построении статистических выводов и имеет лучшие свойства в конечных выборках. Основной идеей бутстрапа является замена истинного механизма генерации данных выборочным распределением. В случае независимых одинаково распределенных данных процедура построения бутстраповских выборок сводится к вытягиванию с возвращением данных из исходной выборки. Такая процедура не проходит в контексте временных рядов, так как разрушает зависимость между наблюдениями. Одно из решений - наложить ограничения на модель генерации данных, чтобы выделить независимые одинаково распределенные инновации, тогда применима упомянутая выше процедура. Другой способ - строить бутстраповские выборки исходя из исходной, без значительного разрушения временной структуры, как это делается в блочном бутстрапе.

Существующие непараметрические способы построения бутстраповских выборок можно рассматривать в контексте построения переходных вероятностей из одной точки исходной выборки в другую, которые зависят от истории бутстраповской выборки. С такой точки зрения отчетливо видны недостатки существующих методов. Вычисление переходных вероятностей очень резко меняется от точки к точке либо очень слабо зависит от предыстории бутстраповской выборки. Например в блочном бутстрапе механизм сильно зависит от того где находится предыдущая выбранная точка, в конце блока или нет. В стационарном бутстрапе вероятности зависят только от одной точки в предыстории.

В работе предлагается новый непраметрический метод построения бутстраповской выборки, который использует бóльшую информацию об истории и который создает выборку без искусственных разрывов между наблюдениями. Первоначальная идея взята из приближения исходного процесса марковской цепью, которая имеет точки исходной выборки в качестве состояний. Вычисление переходных вероятностей включает информацию из исходной выборки и зависит от истории бутстраповской выборки. Важным преимуществом предложенного метода является более полное использование информации, заложенной в выборке, без наложения сильных ограничений на модель. В частности, метод использует информацию об автокорреляции и условной гетероскедастичности, что делает его привлекательным для приложений. Монте Карло симуляции показали, что предложенный метод дает лучшее приближение конечномерного распределения статистики, в смысле близости реальных размеров тестов к номинальным, по сравнению с блочным бутстрапом и асимптотикой.

Обзор литературы

Первоначально бутстрап был введен Эфроном (1979) для независимых одинаково распределенных данных. Данный метод позволяет получить более точное приближение распределения интересующей нас статистики чем при использовании асимптотического распределения.

В случае временных рядов обычный бутстрап непременим, так как он нарушает структуру временного ряда. Тем не менее существуют модификации, которые позволяют применить первоначальную идею в контексте временных рядов. Различные вариации бутстрапа для временных рядов могут быть представлены следующей таблицей.

Бутстрап остатков		
	параметрический	
	непараметрический	
		Решеточный бутстрап
		Холески-фактор бутстрап
Блочный бутстрап		
	неперекрывающиеся блоки(Künsch)	
	перекрывающиеся блоки	
		Скользящие блоки (Carlstein)
		Бутстрап блоков из блоков
		Обесшумленный бутстрап
	стационарный	
непараметрический бутстрап (Hansen)		

В случае независимых одинаково распределенных данных как способ генерации псевдовыборки, так и выбор статистики для бутстрапа не представляют больших проблем. Псевдовыборка строится при помощи вытягивания с возвращением данных из исходной выборки, в качестве статистики для бутстрапа нужно выбирать асимптотически пивотальную статистику, то есть предельное распределение которой не зависит от неизвестных параметров. В случае временных рядов статистику по прежнему нужно выбирать пивотальной, но генерация псевдовыборок становится достаточно сложной задачей. Если есть наблюдения $\{x_t\}_{t=1}^{T}$, то, фактически, имеется лишь одно наблюдение о траектории процесса, по которому нужно приблизить истинный процесс генерации данных.

Одна из возможностей – сделать предположение о модели генерации данных, например AR модель, и свести задачу к бутстрапу независимых одинаково распределенных инноваций. Другая возможность, сохраняющая структуру зависимости между наблюдениями, – построение псевдовыборки при помощи вытаскивния наблюдений из исходной выборки блоками. Первая называется бутстрапом остатков, вторая – блочный бутстрап.

В случае параметрического бутстрапа остатков на примере AR (p) модели $A(L)y_t = \varepsilon_t$ применяется следующая процедура (Berkowitz, Kilian 1996).

1. выбрать p
2. по исходным данным оценить параметры $\hat{A}(L)$
3. построить бутстраповские инновации ε_t^*, выбирая с возращением из эмпирических остатков $\hat{\varepsilon}_t = \hat{A}(L)y_t$
4. выбрать вектор начальных данных Y_0^*
5. сгенерировать псевдовыборку $\hat{A}(L)y_t^* = \varepsilon_T^*$ исходя из Y_0^*
6. вычислить интересуемые статистики по псевдовыбоке
7. повторить 3-6 чтобы построить эмпирическое распределение статистики

Если в модель не включена константа, то остатки нужно рецентрировать, чтобы среднее равнялось 0 (Berkowits, Kilian 1996, Horowitz 1999). Также, чтобы сделать поправку на количество оцениваемых параметров, обычно остатки домножаются на $[T - p/(T - p - d)]^{1/2}$, где d - количество оцениваемых коэффициентов (Berkowits, Kilian 1996). Вектор первоначальных наблюдений вычисляется либо из остатков ε_t^*, либо берется блок из исходных данных.

Похожие процедуры применяются при бутсрапе ARMA и VAR моделей.

Главный недостаток данного подхода – необходимость выбора модели, что является достаточно сложной задачей. Даже если известен класс моделей, которому принадлежит истинный процесс генерации данных, например AR модели, то качество бутстраповских выводов сильно зависит от выбора правильного порядка p (Berkowits, Kilian 1996).

Решеточный бутстрап приближает исходный процесс генерации данных при помощи авторегрессии порядка $p(T)$, где $p(T) \to \infty$ и $p(T) = o(T)$ при $T \to \infty$. Предполагается, что при помощи такого 'просеивания' мы получаем остатки, которые гораздо ближе к iid данным. Тогда приведенная выше процедура для бутстрапа AR модели верна и для решеточного бутстрапа.

Холески-фактор бутстрап использует точное до второго порядка представление стационарного временного ряда при помощи разложения Холески серийной матрицы ковариаций и набора нескоррелированных гомоскедастичных остатков. То есть если $Y = (y_1, \ldots, y_T)' \sim (\mu, \Sigma)$, разложение Холески $PP' = \Sigma$, то представление $\mu + P\varepsilon$, где $\varepsilon = (\varepsilon_1, \ldots, \varepsilon_T)'$ и

ε_t нескоррелированны и одинаково распределены, можно использовать для генерации псевдовыборок. Для этого необходимо оценить μ, Σ и вектор ошибок. После чего можно при помощи бутстрапа для независимых одинаково распределенных данных строить псвдовыборку для ошибок, которую затем преобразовывать в псевдовыборку для исходного ряда.

Чтобы сохранить зависимость между наблюдениями в исходной выборке, блочный бутстрап выбирает не по одному наблюдению, а целыми блоками. При построении блоков существует ряд альтернатив: блоки фиксированной длины перекрывающиеся (Künsch), неперекрывающиеся (Carlstein), блоки случайной, распределенной по геометрическому закону длины (стационарный бутстрап, Politis&Romano).

Главная сложность в блочном бутстрапе – это выбор длины блока. Маленькие блоки разрушают зависимость между наблюдениями, при больших блоках псевдовыборки будут похожими. В литературе есть несколько способов выбора длины блоков. Berkowitz, Kilian (1996) предлагают в случае, когда модель известна, например AR(2), или состоятельно оценена, например выбран порядок авторегрессионной модели при помощи информационного критерия, выбирать блоки, которые дают наилучшие результаты при Монте Карло симуляциях в заданной модели. Hall, Horowitz and Jing(1995) предлагают разбить выборку на несколько подвыборок, при поиощи перебора различных вариантов выбрать блоки, которые дают хорошие результате при бутстрапе в подвыборках, откуда получить длину блока для исходной выборки.

В случае когда интересуемая нас статистика существенно зависит от порядка наблюдений, например при бутстрапе AR моделей, то следует применять бутстрап блоков из блоков. Основная суть данного метода – вытаскивать блоки уравнений. То есть если у нас есть AR(1) модель и при блочном бутстрапе получилась некая псевдовыборка, то в качестве регрессоров нужно брать не предыдущее наблюдение в псевдовыборке, а предыдущее наблюдение из исходной выборки. Тем самым отличие от обычного блочного бутстрапа будет только в местах склейки блоков.

Обесшумленный бутстрап позволят частично разрешить проблему выбора длины блока. Поскольку в исходных данных зависимость между наблюдениями существенна, что не позволяет выбирать маленькие блоки и тем самым ограничивает разнообразие псевдовыборок, то предлагается применить к исходным данным фильтр, например AR(1), получить остатки и применить блочный бутстрап к остаткам, а затем из псевдоостатков построить псевдовыборку. Предполагается, что применение фильтра схватывает существенную долю зависимости между данными, поэтому к остатком можно применять блочный бутстрап с меньшей длиной блока. Тем самым получается достаточно разнообразное семейство псевдовыборок, в которых сохранена зависимость между наблюдениями.

При непараметрическом бутстрапе (Hansen 1999) для построения псевдовыборки наблюдения вытаскиваются по одному, что напоминает бутстрап для *iid* случая, но наблюдения из исходной выборки вытаскиваются с разными вероятностями. Эти вероятности вычисляются исходя из оптимизационной задачи и зависят от исходной выборки и предыстории псевдовыборки. В оригинальной процедуре считаются известными ограничения на моменты, которым удовлетворяет генеральная совокупность, из которой пришла исходная выборка, поэтому вероятности для вытаскивания нового наблюдения для псевдовыборки вычисляются так, чтобы были выполнены аналоги этих ограничений на моменты в псевдовыборке. Запишем задачу формально.

Пусть y_t - анализируемый временной ряд. Обозначим $x_t = (y_{t-1}, \ldots, y_{t-m})$ предысторию. Пусть известно, что временной ряд удовлетворяет условию

$$E(g(y_t, x_t; \theta_0)|\mathcal{F}_{t-1}) = 0$$

и мы имеем исходную выборку $y_1, \ldots, y_T$. Пусть построена часть псевдовыборки до момента t, то есть известна предыстория псевдовыборки x_t^*. Тогда y_t вытаскивается из исходной выборки с вероятностями $p_1, \ldots, p_T$, которые являются решением задачи

$$\hat{p} = (\hat{p}_1, \ldots, \hat{p}_T) = \arg\max \sum_{i=1}^{T} w_i \log(p_i)$$

с ограничениями

$$p_i \geq 0, \quad \sum_{i=1}^{T} p_i = 1, \quad \sum_{i=1}^{T} g(y_t, x_t^*; \hat{\theta}) p_i = 0$$

Веса w_i вычисляются при помощи ядерной функции

$$w_i = \frac{K_h(x_i - x_t^*)}{\sum_{j=1}^{T} K_h(x_j - x_t^*)}$$

Если не накладывать ограничения на моменты, то есть считать, что модель неизвестна, то решением задачи будет

$$p_i = w_i,$$

то есть y_i вытаскивается с большей вероятностью, если его предыстория x_i близка к предыстории, получившейся к данному моменту в псевдовыборке.

Статистика для бутстрапа. Как и в случае *iid* данных необходимо выбирать асимптотически пивотальные статистики, например в случае AR моделей для OLS оценок коэффициентов используются t-статистики. Здесь возникает проблема оценивания стандартных ошибок, а также смещение (Horowitz 1999) из-за того, что процесс построения псевдовыборки при блочном бутстрапе не воспроизводит структуру зависимости истинного DGP.

Стандартные ошибки вычисляются при помощи HAC формулы, с использованием различных весов, а также оцениванием ковариаций только внутри блоков. Также стандартные ошибки можно вычислять при помощи VARHAC оценки (den Haan, Levin 1996).

Интерпретация существующих схем в терминах переходных вероятностей

Существующие схемы построения псевдовыборки можно рассмотреть в следующем контексте. Пусть имеется исходная выборка $y_1, \ldots, y_T$ и построена часть псевдовыборки $y_1^*, \ldots, y_s^*$, тогда следующее наблюдение для псевдовыборки y_{s+1}^* выбирается из исходной выборки в соответствии с вероятностями p_t (то есть $y_{s+1}^* = y_t$ с вероятностью p_t и $\sum p_t = 1$). В общем виде эти вероятности зависят от истории псевдовыборки и вычисляются исходя из информации содержащейся в данных в соответствии с некоторым правилом, то есть $p_t = p_t(y_1^*, \ldots, y_s^*; y_1, \ldots, y_T; s)$.

Так вероятности $p_t = \frac{1}{T}$ соответствует схеме для независимых одинаково распределенных данных и не зависят от предыстории псевдовыборки и самой выборки.

В случае стационарного бутстрапа, если в уже построенной части псевдовыборки $y_s^* = y_k$, то следующее псевдонаблюдение y_{s+1}^* с вероятностью p будет равно y_{k+1} и с вероятностью $1-p$ будет произвольным наблюдением из исходной выборки. Более формально, $p_t = p + (1-p)\frac{1}{T}$ при $t = k+1$ и $p_t = (1-p)\frac{1}{T}$ при $t \neq k+1$. Здесь переходные вероятности зависят от последнего наблюдения в истории псевдовыборки и от исходной выборки.

Рассмотрим теперь блочный бутстрап с неперекрывающимися блоками длины l (для простоты предположим, что $T = nl$). Здесь процедура выбора y_{s+1}^* зависит существенным образом от того в какой части блока находилось наблюдение y_s^*. Если это было последнее наблюдение в блоке, то y_{s+1}^* выбирается с равными вероятностями из $y_1, y_{1+l}, y_{1+2l} \ldots, y_{1+(n-1)l}$, если оно находилась внутри блока, то вероятности вырождаются и с вероятностью 1 берется следующая точка внутри блока. Для определения местонахождения точки в данном случае по-прежнему достаточно последней точки из имеющейся части псевдовыборки, но здесь механизм определения резко меняется в зависимости от этой информации. Он либо детерменистичен, либо случаен.

В блочном бутстрапе с перекрывающимися блоками ситуация похожая. То есть y_{s+1}^* выбирается с равными вероятностями из $y_1, y_2, y_3 \ldots, y_{T-l+1}$, если y_s^* - последняя точка блока и с вероятностью 1 берется следующая точка внутри блока, если y_s^* внутри блока. Но теперь недостаточно одного наблюдения для определения местонахождения точки y_s^* по отношению к блоку, в некоторых случаях необходима вся история псевдовыборки. Но по прежнему механизм определения вероятностей очень резко меняется от этого местонахождения.

Непараметрический бутстрап Хансена в терминах переходных вероятностей выглядит наиболее совершенным. При вычислении вероятностей используется информация из выборки, информация о предыстории в псевдовыборке, а также механизм вычисления вероятностей не меняется резко от точки к точке. Но этот метод требует знания модели, то есть ограничений на моменты, а также требует много времени для вычислений.

Итак, эти схемы обладают рядом недостатков. Поэтому необходима некоторая модификация этих схем с целью более полного использования информации, содержащейся в исходной выборке.

Марковский бутстрап

Итак общая постановка задачи. Имеется исходная выборка $y_1, \dots, y_T$ и построена часть псевдовыборки $y_1^*, \dots, y_s^*$. Тогда $y_{s+1}^* = y_t$ с вероятностью p_t.

$$p_t = p_t(y_1^*, \dots, y_s^*; y_1, \dots, y_T; s)$$

Найдем эти вероятности исходя из приближения исходного процесса дискретной марковской цепью, которая имеет в качестве состояний наблюдения исходной выборки, а переходные вероятности находятся по некоторому правилу и зависят от исходной выборки.

Простейший случай. Рассмотрим пары $\{(y_k, y_{k+1})\}_{k=1}^{T-1}$. Если рассмотреть эти точки в двумерной плоскости, то эти точки находятся в квадрате $[y_{(1)}, y_{(T)}] \times [y_{(1)}, y_{(T)}]$. Разделим этот квадрат на части, для этого поделим каждую сторону на I частей $\Delta_1, \Delta_2, \dots, \Delta_I$, возможные варианты такого разбиения будут даны ниже. После разбиения найдем отрезок Δ_k, в который попала y_s^*, после чего определим вероятности

$$p_t = \begin{cases} 0, & y_{t-1} \notin \Delta_k \\ 1/\#\{y_l \in \Delta_k\}, & y_{t-1} \in \Delta_k \end{cases}$$

То есть если y_s^* попала в Δ_k, то мы с равной вероятностью переходим в точки, у которых предыстория также попадает в этот отрезок.

Выбор разбиения $[y_{(1)}, y_{(T)}]$ на отрезки Δ_i является одним из параметров модели. В симуляциях были рассмотрены равномерное эмпирическое разбиения и разбиение на равные отрезки. Обозначим $\Delta_i = [a_i, a_{i+1}]$.
Равномерное эмпирическое разбиение (с одинаковым количеством

точек предыстории в каждом отрезке):

$$a_1 = y_{min}, \quad \#\{a_i \le y_s \le a_{i+1}\} = m, \quad i = 1, \ldots, I, \quad I = \frac{T}{m}, \quad (m = pT)$$

Разбиение на равные отрезки

$$a_1 = y_{min}, \quad a_{i+1} = a_i + \frac{y_{max} - y_{min}}{I}, \quad i = 1, \ldots, I, \quad I = \frac{T}{m}, \quad (m = pT)$$

Даже в простейшем случае видны преимущества МС бутстрапа. Во-первых, переходные вероятности зависят от предыстории бутстраповской выборки. Во-вторых, при определении вероятностей используется сама выборка. Наконец, вероятности считаются по единой схеме, которая не меняется от точки к точке.

Сглаженный марковский бутстрап

В простейшем случае, если история псевдовыборки пришла в отрезок Δ_k, то вероятности перехода в точки исходной выборки, у которых предыстория также попала в этот отрезок, были одинаковыми. Естественным обобщением является учет того, насколько далеко легла предыстория в выборке от предыстории точки, в которую мы хотим перейти. Эту информацию можно учесть при помощи ядерной функции, тогда получаем следующую формулу для переходных вероятностей:

$$p_t = p_t(y_1^*, \ldots, y_s^*; y_1, \ldots, y_T; s) = K_{h(y_s^*)}(y_s^* - y_t),$$

где в качестве ядерной функции K возможны различные варианты, а также $h(y_s^*)$ является отдельным параметром выбора. Эта формула напоминает непараметрический бутстрап Хансена в простейшем случае (без использования ограничений на моменты), но здесь h зависит от y_s^*, что дает дополнительные преимущества. И как показали расчеты, при h, не зависящем от предыстории, метод дает худшие результаты.

Исследование метода в стилизованных моделях

Метод исследовался на модели AR(1) когда истинный процесс генерации данных один из следующих: AR(1), AR(2) или нелинейный. Выбор таких вариантов позволяет определить поведение методов в случае когда истинный процесс генерации данных известен, когда в силу каких-либо причин произошла ошибка в определении порядка авторегрессии (например в силу откланений состоятельной оценки для порядка авторегрессии при помощи информационного критерия в конечных выборках), или когда истинный процесс генерации данных нелинейный, но мы его анализируем в терминах проекции на предысторию процесса. Итак, формально были выбраны следующие модели:

$$y_t = \mu + \alpha y_{t-1} + e_t, \tag{1}$$

$$(1-\alpha L)(1-\rho_2 L)y_t = e_t \quad \sim \quad y_t = \mu + (\alpha+\rho_2)y_{t-1} - \alpha\rho_2 y_{t-2} + e_t,^{1} \tag{2}$$

$$y_t = \mu + \alpha y_{t-2}e_{t-1} + e_t,^{2} \tag{3}$$

где

$$e_t = \eta_t\sqrt{\omega + \gamma e_{t-1}^2}.$$

[1] $\rho_2 = .5$ При оценивании этой модели как AR(1) $y_t = \tilde{\mu} + \tilde{\alpha} y_{t-1} + v_t$ нужно использовать псевдоистинный параметр $\tilde{\alpha} = \frac{Cov(y_t, y_{t-1})}{Var(y_t)} = \frac{\gamma_1}{\gamma_0} = \frac{\alpha+\rho_2}{1+\alpha\rho_2}$, тогда $Cov(v_t, y_{t-1})$=0

[2] псевдоистинный параметр $\tilde{\alpha} = 0$

Распределение η_t было выбранно стандартным нормальным $\mathcal{N}(0,1)$.
Параметры: $\omega = 1$, $\mu = 0$, $\gamma \in \{0, .5, \}$, $\alpha = 0$ для AR(1) модели, $\alpha = -0.5$ для AR(2) модели и $\alpha = 0.5$ для нелинейной модели. Такие параметры для α были выбраны, чтобы проанализировать метод в случае несвязных процессов (not persistent).
Размер выборки $T = 30, 60, 120$ и $R = 5000$ [3] количество симуляций.

В результате симуляций вычислялись истинные ошибки первого рода при проверке гипотез при помощи симметричных двусторонних и несимметричных односторонних t-процентных критических значений, посчитанных с использованием асимптотического распределения, распределения полученного в симуляциях при помощи бутстрапа с перекрывающимися блоками и при помощи марковского бутстрапа.

Для вычисления стандартных ошибок использовались часто используемая в приложениях формула Вайта, формула Нюи-Веста для учета автокорреляции в ошибках, но поскольку она на учитывает степень корреляции ошибок в каждом конкретном случае, то была использована также и VARHAC формула.

[3] для асимптотики и прекрывающихся блоков $R = 10000$

Описание экспериментов

Временной ряд $y_1, \ldots, y_T$ оценивался как AR(1) модель.

$$y_t = \mu + \alpha y_{t-1} + e_t$$

В дальнейшем будут использоваться следующие обозначения:

$$Y = (y_2, \ldots, y_T)', \quad X = (x_t') = [(1, y_1)', \ldots, (1, y_{T-1})']'$$

OLS оценка $\hat{\beta} = (\hat{\mu}, \hat{\alpha})' = (X'X)^{-1}X'Y$, остатки $\hat{e} = Y - X\beta$, $Z = (z_t') = (x_t'\hat{e}_t)$,

t-статистика для коэффициента α расчитывалась по формуле

$$t = \frac{\hat{\alpha} - \alpha}{se(\hat{\alpha})},$$

где стандартная ошибка расчитывалась по Вайту, Нюи-Весту и VARHAC.

Формула Нюи-Веста

$$\hat{V} = \hat{Q}_{xx}^{-1}\hat{Q}_{xe}\hat{Q}_{xx}^{-1} =$$

$$= \frac{T-1}{T-3}(X'X)^{-1}\left(Z'Z + \sum_{j=1}^{m}\omega(\frac{j}{m+1})\sum_{l=1}^{T-1-j}(z_{l+j}z_l' + z_l z_{l+j}')\right)(X'X)^{-1}$$

где $m = \left[4(\frac{T-1}{100})^{1/3}\right]$, в качестве весовых функций использовались

$\omega(x) = (1-x)$	стандартная
$\omega(x) = 1$	равномерная
$\omega(x) = \begin{cases} 1, & \lvert x\rvert \le \alpha \\ 1 - \frac{\lvert x\rvert-\alpha}{1-\alpha}, & \alpha < \lvert x\rvert \end{cases}$	трапециидальная
$\omega(x) = 1 - \lvert x\rvert^q$	Парзеновская

Стандартная ошибка по Вайту соответствует $m = 0$ в приведенной выше формуле. Стандартная ошибка VARHAC строится следующим образом. Для каждого столбца $Z1$, $Z2$ матрицы Z прогонялись регрессии на $Z_{-1}, \ldots, Z_{-k}$, где $k = 1, \ldots, 4$, выбиралась модель с минимальным AIC[4] критерием (также рассматривалась модель с $k = 0$, то есть когда в качестве e берется исходный вектор). Итак, для каждого столбца выбиралась своя модель, из них составлялась общая VAR модель для Z

$$Z_t = \sum_{k=1}^{\widehat{k}} \hat{A}_k Z_{t-k} + \hat{e}_t,$$

после чего строились $\widetilde{V}$ - оценка дисперсии ошибок VAR модели, и итоговая оценка для $\hat{Q}_{xe}$ вычислялась по формуле $\hat{Q}_{xe} = (I - \sum_{k=1}^{\widehat{k}} \hat{A}_k)^{-1} \widetilde{V} (I - \sum_{k=1}^{\widehat{k}} \hat{A}'_k)^{-1}$.

Поведение нового метода сравнивалось с результатами асимптотического подхода и блочного бутстрапа.

Асимптотика. В 10000 симуляциях рассматривалось действительный процент попадания t-статистики в 1%, 5% и 10% критические области, построенные по стандартному нормальному распределению [5]. То есть считалось частота событий $\{|t| > 2.57\}$, $\{|t| > 1.96\}$, $\{|t| > 1.64\}$ для симметричных двухсторонних интервалов и $\{t > 2.33\}$, $\{t > 1.64\}$, $\{t > 1.28\}$ и $\{t < -2.33\}$, $\{t < -1.64\}$, $\{t < -1.28\}$ для сравнения с несимметричными односторонними интервалами.

[4] $AIC = \ln \frac{e'e}{T} + 2 * k * 2/T$

[5] Для $\xi \sim N(0,1)$ $P\{|\xi| > 2.57\} = 1\%$, $P\{|\xi| > 1.96\} = 5\%$, $P\{|\xi| > 1.64\} = 10\%$, $P\{\xi > 2.33\} = P\{\xi < -2.33\} = 1\%$, $P\{\xi > 1.64\} = P\{\xi < -1.64\} = 5\%$, $P\{\xi > 1.28\} = P\{\xi < -1.28\} = 10\%$,

Блочный бутстрап. В 10000 симуляциях расчитывася действительный процент попадания t-статистики в 1%, 5% и 10% критические области, построенные исходя из бутстраповского распределения t-статистики, вычисленного при помощи B бутстраповских повторений. При построении псевдовыборки использовались перекрывающиеся блоки длины l. То есть если обозначить $s_1, \ldots, s_{[T/l]+1}$ случайные целые числа выбранные из $[1, T-l+1]$, то псевдовыборка строилась как

$$y_{s_1}, y_{s_1+1}, \ldots, y_{s_1+l-1}, y_{s_2}, \ldots, y_{s_2+l-1}, \ldots, y_{s_{[T/l]+1}}, \ldots, y_{s_{[T/l]+1}+T-[T/l]l-1}$$

Для каждой псевдовыборки проводилась процедура оценки коэффициентов как описано выше и вычислялась t-статистика

$$t(\hat{\alpha}^*) = \frac{\hat{\alpha}^* - \hat{\alpha}}{se(\hat{\alpha}^*)},$$

после чего по набору этих статистик находились критические значения для симметричных двусторонних интервалов $|t|_{[B(1-0.01)]}$, $|t|_{[B(1-0.05)]}$, $|t|_{[B(1-0.10)]}$ и для несимметричных односторонних интервалов $t_{[B*0.01]}$, $t_{[B*0.05]}$, $t_{[B*0.10]}$, $t_{[B(1-0.01)]}$, $t_{[B(1-0.05)]}$, $t_{[B(1-0.10)]}$. Далее находился процент попадания t-статистики в критические области.

Марковский бутстрап. Алгоритм отличается от предыдущего только в части построения псевдовыборки. Для построения псевдовыборки предварительно вычислялось разбиение $\{a_i\}_{i=1}^{I+1}$ отрезка $[y_{(1)}, y_{(T)}]$ по одному из предложенных ранее способов. Далее по этому разбиению вычислялась матрица переходов $P = (p_{ij})$, где p_{ij} - вероятность прейти в точку y_j если предыстория бутстраповской псевдовыборки попала в y_i. Если обозначить I_k отрезок куда попадала точка y_i, а n - общее количество точек попавших в этот отрезок, то i-я строка матрицы P вычисляется по формуле

$$p_{ij} = \begin{cases} 0, & y_{j-1} \notin I_k \\ 1/n, & y_{j-1} \in I_k \end{cases}$$

Отдельной обработки требует точка y_T. Если эта точка попала в отрезок I_k вместе с другими точками, то применима предыдущая формула. Но возможна ситуация, когда в I_k кроме y_T ничего не попало (такая ситуация не возможна при равномерном эмпирическом разбиении и редка для разбиения на равные отрезки). Для любой другой точки это не играет большой роли, так как всегда имеется точка для перехода в исходной выборке, но для последней точки не понятно в какую точку переходить. Тогда вычисляется прогноз для $\widehat{y}_{T+1}$, вычисленный по МНК оценкам исходной выборки, и переходим в точку исходной выборки, ближайшую к $\widehat{y}_{T+1}$. Но возможно еще более редкая ситуация, когда ближайшей к $\widehat{y}_{T+1}$ оказывается y_T, тогда с некоторого момента псевдовыборка может вырождиться и принимать значение y_T. В таких случаях полагалось, что из y_T мы можем перейти в любую другую точку исходной выборки с равной вероятностью.

Псевдовыборка строится по вычисленной матрице переходов следующим образом. Первая точка выбирается случайно. Далее если есть предыстория $y_l^* = y_i$, то следующая точка выбирается исходя из распределения, соответствующего i-ой строке матрицы P.

Сглаженный марковский бутстрап. Матрица переходных вероятностей составляется по формуле

$$p_{ij} = K_{h(y_i)}(y_i - y_{j-1}).$$

В качестве ядерной функции использовалось ядро Епанечникова, $h(y_i)$ вычесляется по правилу ближайших соседей, то есть таким, чтобы в отрезок $[y_i - h, y_i + h]$ попало заданное количество точек из исходной выборки.

Результаты экспериментов

Таблицы с результатами представлены в приложениях. В таблицах показаны результаты полученные при симуляциях Монте Карло. Асимптотика (ASY), блочный бутстрап (MBB), марковский бутстрап с равномерным эмпирическим разбиением (MCB1), марковский бутстрап с разбиением на равные отрезки (MCB2) и сглаженный марковский бутстрап (SMCB) сравниваются при помощи реальных размеров тестов, построенных по 1%, 5% и 10% квантилям. Приведены результаты для симметричных двусторонних квантилей, несимметричных односторонних левых и правых квантилей. Для каждой модели расматривались гомоскедастичные и гетероскедастичные варианты, а также варианты с расчетом стандартных ошибок по Вайту (W), Нюи-Весту (NW) и VARHAC формулам. В таблицах приведены данные для $T = 30$, для MBB размер блока равен 4, для MCB количество отрезков равно 8, для SMCB количество ближайших соседей равно 5.

В таблице 1a) даны результаты для идеального случая, когда истинный процесс генерации данных совпадает с оцениваемой моделью, ошибки нескоррелированы и гомоскедастичны. Даже в идеальном случае результаты асимптотических выводов оставляют желать лучшего.

Хотя асимптотические односторонние реальные размеры более близки к реальным, чем бутстраповские, но имеющие большую скорость сходимости симметричные тесты хуже. Причем асимптотические результаты сильно зависят от способа расчета стандартных ошибок, более сложные формулы дают худшие результаты. Бутстрап улучшает результаты для симметричных интервалов, блочный бутстрап и марковский бутстрап с равномерным эмпирическим распределением ведут себя примерно одинаково, сглаженный бутстрап не дает улучшения, равномерное эмпирическое разбиение предпочтительнее разбиения на равные отрезки. Если смотреть на результаты односторонних тестов, то марковский бутстрап дает лучшие результаты, и, что не маловажно, применение левосторонних и правосторонних квантилей дает симметричные результаты, что свидетельствует о симметричном приближении. Важной особенностью является то, что результаты на ухудшаются при расчете стандартных ошибок другими способами, а в некоторых случаях, например VARHAC для MCB2, даже улучшаются.

В случае когда истинный процесс совпадает с оцениваемым, а ошибки нескоррелированы и гетероскедастичны (такая ситуация отображена в таблице 1б), все предыдущие выводы справедливы. Необходимо сделать лишь одно замечание об ухудшении асимптотических выводов.

Если истинный процесс генерации данных, AR(2) модель, не совпадает с оцениваемой, AR(1) моделью, то ошибки в AR(1) модели будут скоррелированы и применение формулы Вайта для стандартных ошибок дает неверный результат, что отражается в результатах симуляций в таблицах 2а) и 2б). При состоятельном расчете стандартных ошибок реальные асимтотические размеры тестов не имеют ничего общего с номинальными. Блочный и марковский бутстрап дают примерно одинаковые результаты для симметричных тестов и, по прежнему, марковский бутстрап лидирует в приближении конечномерного распределения, если рассматривать односторонние тесты. Использование VARHAC формулы позволяет достичь практически номинальных размеров для MBB и SMCB. Марковский бутстрап с разбиением на равные отрезки уступает альтернативным методам построения псевдовыборки. Все сказанное выше справедливо и для гетероскедастичного варианта AR(2) модели, с небольшим ухудшением конкретных чисел.

Таблицы 3а) и 3б) показывают случай нелинейного истинного процесса генерации данных. Здесь однозначно лидирующую роль занимает марковский бутстрап, причем как при односороннем тестированнии, так и при двустороннем симметричном. Использование VARHAC формулы для стандартных ошибок по прежнему улучшает результаты. От использования асимптотики лучше воздерживаться. В данном случае для результатов бутстрапа нет существенного различия при гомоскедастичных или гетероскедастичных ошибках истинного процесса генерации данных, потому что при оценки нелинейного процесса при помощи AR(1) модели ошибки будут скоррелированы и гетероскедастичны в любом случае. Поэтому цифры получаются похожими, хотя заметны некоторые улучшения в симметричных двусторонних тестах и ухудшения в односторонних, но эти изменения могут отчасти объясняться погрешностью симуляций.

На основе полученных результатов можно сделать выыод, что построение псевдовыборки при помощи марковского бутстрапа обеспечивает более точное приближение конечномерноого распределения t-статистики. На практике следует применять марковский бутстрап с равномерным эмпирическим разбиением или сглаженный марковский бутстрап. При расчете стандартных ошибок лучше пользоваться HAC формулами, поскольку они не дают проигрыша в идеальных случаях (когда нет серийной корреляции) и всегда состоятельны.

Приложения к реальным данным

Для анализа был выбран ряд первой разности ежегодных данных по ВНП США за период 1909-1988 [6] (первая разность взята поскольку ВНП США является интегрируемым порядка 1, см.например Hansen 1999a). Ряд моделировался в классе расширения AR моделей при помощи квадратичных членов и на основаниии BIC-критерия была выбрана AR(1) модель.

$$y_t = c_1 + c_2 y_{t-1} + \varepsilon_t$$

Оценки (в скобках указаны стандартные ошибки, посчитанные по формуле Нюи-Веста (NW) / VARHAC) $c_1 = 0.020(0.007/0.008)$, $c_2 = 0.334(0.119/0.135)$. t-статистика для второго коэффициента 2.81/2.47 превосходит критические значения посчитанные по асимптотике, то есть при использовании асимптотических статистических выводов коэффициент значим на 5% уровне. Применение блочного бутстрапа с длиной блока 4, 6 и 8 дает следующие 5% критические значения: 2.88/3.44, 2.70/3.34, 2.75/3.56 соот-

[6] Данные взяты с сайта Б.Хансена (http://www.ssc.wisc.edu/ bhansen/) - содержат расширенные данные Нельсона-Плоссера в натуральных логарифмах.

ветственно. Применение равномерного эмпирического марковского бутстрапа с параметрами $I = 7, 8, 9$ дает следущие результаты: 2.56/2.81, 2.44/3.00, 2.65/3.03 соответственно. Такие результаты ставят вопрос о непредсказуемости исследуемого ряда в классе линейных моделей.

На основании AIC-критерия была выбрана модель

$$y_t = c_0 + c_1 y_{t-1} + c_3 y_{t-3} + c_{23} y_{t-2} y_{t-3} + \varepsilon_t$$

Оценки (в скобках указаны стандартные ошибки, посчитанные по формуле Нюи-Веста (NW)) $c_0 = 0.022(0.009)$, $c_1 = 0.325(0.135)$, $c_3 = -0.278(0.109)$, $c_{23} = 2.685(1.689)$. Значения для t-статистики c_1, c_3 и c_{23} равны 2.41, -2.55 и 1.59 соответственно и превосходят критические значения посчитанные по асимптотике, то есть при использовании асимптотических статистических выводов коэффициенты (кроме c_{23}) значимы на 5% уровне. Проверка гипотезы на совместную значимость c_1, c_3 и c_{23} дает значение статистики Вальда равным 14.48 и использование асимптотики отвергает нулевую гипотезу $H_0 : c_1 = c_3 = c_{23} = 0$ на 5% уровне значимости. Применение блочного бутстрапа с длиной блока 4, 6 и 8 дает следующие 5% критические значения: 2.92, 2.73, 2.88 для c_1, 3.95, 3.33, 3.17 для c_3, 3.67, 3.27, 3.24 для c_{23}, 32.83, 28.08, 30.33 для статистики Вальда. Применение равномерного эмпирического марковского бутстрапа с параметрами $I = 7, 8, 9$ дает следущие результаты: 2.44, 2.50, 2.71 для c_1, 4.45, 4.50, 4.44 для c_3, 4.11, 4.20, 4.15 для c_{23}, 40.47, 41.19, 40.49 для статистики Вальда.

Как видно из симуляций, использование асимптотики слишком часто отвергает гипотезу о незначимости коэффициентов, то есть в нашем случае гипотезу о непредсказуемости процесса в расширении AR процессов. Применение марковского бутстрапа позволяет скорректировать выводы и поскольку этот метод показал хорошую точность в симуляциях, то следует сделать вывод о непредсказуемости процесса в расширениии AR моделей и исследовать ряд в еще более широком классе.

Заключение

Марковский бутстрап прост в реализации, не требует много времени для вычислений в реальных приложениях, показал наилучшее приближение конечномерного распределения t-статистики по сравнению с асимптотикой и блочным бутстрапом. Все это делает его привлекательным для использования в приложениях и построения корректных статистических выводов.

Список литературы

Berkiwitz, J. and Kilian, L. (2000). Recent Developments in Bootstrapping Time Series, *Econometric Reviews*, 19(1), February 2000, 1-54.

Bergstrom, P. (1999). Bootstrap Methods and Applications in Ecometrics - a Brief Survey, *working paper* Uppsala University.

Carlstein, E. (1986). The use of subseries methods for estimating the variance of a general statistic from a stationary time series, *Annals of Statistics*, 14, 1171-1179.

Efron, B. (1979). Bootstrap methods: another look at the jackknife, *Annals of Statistics*, 7, 1-26.

den Haan, W.J., Lewin, A. (1996) A practitioner's guide to robust covariance matrix estimation, *discussion paper*, University of California, San Diego.

Hall, P., J.L. Horowitz, and B.-Y. Jing (1995). On blocking rules for the bootstrap with dependent data, *Biometrika*, 82, 561-574.

Hall, P., J.L. Horowitz, (1996). Bootstrap Critical Values for Tests Based on Generalized-Method-of-Moments Estimators, *Econometrica*, Vol. 64, No. 4, 891-916

Hansen, B.E. (1999a). The Grid Bootstrap and Autoregressive model, *Review of Economics and Statistics*, 81 (4), 594-607.

Hansen, B.E. (1999b). Non-Parametric Dependent Data Bootstrap for Conditional Moment Models, (working paper, University of Wisconsin-Madison) Presented at the 2000 World Congress Meetings in Seattle, Washington.

Hardle, W. and Linton O. (1994). Applied Nonparametric Methods, *Handbook of econometrics*, volume IV, pp. 2297-2341, Elsevier science B.V.

Horowitz, J.L. (1997). Bootstrap methods in econometrics: theory and numerical perfomance, Econometric society monograph "Advances in economics and econometrics: theory and application"edited D.M.Kreps and K.F.Willis, *Cambrige University Press*, 1997, 188-222.

Horowitz, J.L. (1999). The bootstrap, Handbook of econometrics, forthcoming.

Horowitz, J.L. and Savin, N.E. (2000). Empirically relevant critical values for hypothesis tests: A bootstrap approach, *Journal of econometrics* 95(2000) 375-389.

Kilian, L. (1998). Small-Sample Confidence Intervals for Impulse Response Function, *Review of Economics and Statistics*, 80 (2), 218-230.

Kilian, L. (1999). Finite-Sample Properties of Percentile and Percentile-t Bootstrap Confidence Intervals for Impulse Responses, *Review of Economics and Statistics*, 81 (4), 652-660.

Kunsch, H.R. (1989). The jackknife and the bootstrap for general stationary observations, *Annals of Statistics*, 17, 1217-1241.

Lahiri, S.N. (1997). Theoretical comparisons of block bootstrap methods, *working paper*, Department of Statistics, Iowa State University, Ames, IA.

Maddala, G.S. and In-Moo Kim (1998). Unit roots, cointegration, and structural Change, 10. Small Sample inference: bootstrap methods, pp. 309-336, *Cambridge University Press* 1998.

Politis, D.N. and J.P. Romano (1994). The Stationary Bootstrap, *Journal of the American Statistical Association*, 89, 1303-1313.

Silverman, B.W. (1986). Density Estimation for Statistics and Data Analysis. *London: Chapman and Hall*, pp. 45, 86-87.

Приложения

Таблица 1a: AR(1) модель, гомоскедастичный вариант ($\gamma = 0.0$)

	$P\{\lvert t\rvert \geq q_\alpha^{sym}\}$			$P\left\{t \leq q_\alpha^{left}\right\}$			$P\left\{t \geq q_\alpha^{right}\right\}$		
	α=1%	α=5%	α=10%	α=1%	α=5%	α=10%	α=1%	α=5%	α=10%
W									
ASY	2,5%	7,8%	13,3%	2,8%	9,0%	15,4%	1,2%	4,2%	7,7%
MBB	1,8%	6,5%	11,5%	6,3%	13,6%	20,2%	6,1%	12,5%	18,4%
MCB1	1,8%	5,6%	10,9%	3,0%	8,3%	14,1%	2,9%	8,8%	14,3%
MCB2	2,0%	6,2%	12,1%	3,1%	8,6%	14,7%	3,8%	9,0%	14,1%
SMCB	1,6%	5,7%	11,1%	3,7%	9,3%	15,3%	3,7%	9,6%	15,7%
NW									
ASY	5,1%	12,7%	18,9%	5,1%	12,4%	19,2%	2,3%	6,6%	10,4%
MBB	1,5%	5,7%	10,4%	5,5%	12,6%	18,8%	5,6%	11,8%	17,5%
MCB1	2,1%	5,6%	10,3%	2,9%	8,2%	13,1%	2,5%	8,3%	13,6%
MCB2	1,8%	6,5%	11,8%	2,7%	7,9%	13,6%	3,2%	7,8%	13,4%
SMCB	1,7%	5,5%	10,3%	2,9%	8,9%	14,3%	3,1%	9,6%	15,8%
VARHAC									
ASY	13,3%	21,6%	28,3%	10,7%	18,2%	24,2%	5,4%	10,1%	14,1%
MBB	1,2%	4,6%	9,7%	3,7%	11,6%	18,2%	4,1%	11,0%	16,8%
MCB1	1,3%	5,1%	10,1%	1,8%	6,9%	12,8%	2,3%	8,2%	14,4%
MCB2	1,3%	5,3%	10,9%	1,7%	6,9%	12,3%	2,5%	8,4%	14,4%
SMCB	1,4%	5,4%	9,8%	2,3%	8,4%	14,2%	3,0%	8,3%	14,7%

Таблица 1б: AR(1) модель, гетероскедастичный вариант ($\gamma = 0.5$)

	$P\{\lvert t\rvert \geq q_\alpha^{sym}\}$			$P\{t \leq q_\alpha^{left}\}$			$P\{t \geq q_\alpha^{right}\}$		
	α=1%	α=5%	α=10%	α=1%	α=5%	α=10%	α=1%	α=5%	α=10%
W									
ASY	3,7%	9,6%	15,7%	3,5%	10,0%	15,9%	1,9%	5,7%	9,7%
MBB	1,5%	5,8%	10,5%	8,4%	16,0%	21,6%	8,3%	15,2%	20,3%
MCB1	1,7%	6,2%	11,3%	4,5%	11,1%	16,3%	3,6%	9,3%	14,9%
MCB2	1,7%	6,0%	11,8%	4,4%	10,3%	15,6%	5,7%	11,7%	17,3%
SMCB	1,4%	5,0%	9,2%	5,9%	13,1%	19,1%	3,8%	10,8%	16,7%
NW									
ASY	6,8%	14,9%	22,2%	5,6%	13,4%	19,8%	3,7%	8,7%	13,2%
MBB	1,3%	5,6%	10,5%	7,8%	15,1%	21,3%	8,7%	15,7%	21,3%
MCB1	1,6%	5,7%	10,6%	4,8%	11,6%	17,1%	3,6%	9,1%	14,8%
MCB2	1,9%	6,5%	11,8%	5,4%	11,1%	17,2%	4,7%	10,9%	16,5%
SMCB	1,2%	4,5%	8,8%	5,5%	13,0%	19,3%	4,3%	11,0%	17,4%
VARHAC									
ASY	19,7%	29,6%	36,7%	14,1%	21,8%	27,1%	9,2%	14,9%	19,2%
MBB	1,2%	5,6%	11,1%	5,9%	14,3%	21,3%	7,2%	15,4%	21,4%
MCB1	1,2%	5,8%	10,6%	3,5%	10,3%	16,5%	3,2%	9,7%	15,6%
MCB2	1,9%	6,3%	12,0%	3,6%	10,2%	16,3%	4,0%	10,9%	16,4%
SMCB	1,4%	5,8%	10,6%	4,4%	12,2%	19,3%	4,1%	11,5%	18,1%

Таблица 2а: AR(2) модель, гомоскедастичный вариант ($\gamma = 0.0$)

	$P\{\lvert t\rvert \geq q_\alpha^{sym}\}$			$P\{t \leq q_\alpha^{left}\}$			$P\{t \geq q_\alpha^{right}\}$		
	α=1%	α=5%	α=10%	α=1%	α=5%	α=10%	α=1%	α=5%	α=10%
W									
ASY	8,2%	17,4%	25,2%	7,8%	17,2%	24,2%	3,3%	8,0%	12,6%
MBB	4,1%	11,2%	18,1%	10,0%	19,1%	25,5%	8,5%	14,8%	19,8%
MCB1	4,5%	11,9%	19,3%	6,4%	15,0%	21,0%	4,8%	10,6%	15,6%
MCB2	5,1%	12,9%	20,4%	6,8%	15,1%	21,0%	5,6%	11,3%	16,1%
SMCB	4,4%	11,2%	18,7%	6,8%	14,8%	21,4%	5,7%	12,7%	18,2%
NW									
ASY	9,3%	18,6%	25,7%	8,7%	17,5%	24,6%	3,7%	8,2%	12,4%
MBB	2,8%	8,5%	14,5%	8,3%	16,7%	22,6%	6,7%	13,3%	18,6%
MCB1	3,5%	9,6%	14,9%	4,9%	11,8%	17,8%	3,4%	8,5%	14,0%
MCB2	4,1%	10,1%	16,5%	5,3%	11,8%	18,3%	4,3%	9,7%	14,4%
SMCB	3,2%	9,2%	15,8%	6,0%	13,8%	20,6%	3,9%	10,2%	15,5%
VARHAC									
ASY	16,4%	25,8%	32,8%	13,9%	22,2%	28,1%	6,0%	10,6%	14,5%
MBB	1,6%	6,1%	11,7%	5,9%	14,8%	21,6%	5,2%	12,0%	17,8%
MCB1	1,7%	6,7%	12,8%	2,7%	9,3%	16,5%	2,8%	8,7%	14,4%
MCB2	2,0%	6,8%	13,1%	2,5%	9,3%	16,3%	3,0%	9,2%	14,1%
SMCB	1,7%	6,1%	11,6%	3,0%	10,0%	16,9%	3,0%	9,5%	15,0%

Таблица 2б: AR(2) модель, гетероскедастичный вариант ($\gamma = 0.5$)

	$P\{\lvert t\rvert \geq q_\alpha^{sym}\}$			$P\{t \leq q_\alpha^{left}\}$			$P\{t \geq q_\alpha^{right}\}$		
	α=1%	α=5%	α=10%	α=1%	α=5%	α=10%	α=1%	α=5%	α=10%
W									
ASY	9,6%	18,4%	25,8%	8,5%	17,1%	23,2%	3,9%	8,7%	13,5%
MBB	3,7%	10,6%	17,1%	12,3%	20,4%	25,9%	11,0%	17,5%	22,6%
MCB1	4,1%	11,6%	18,3%	8,9%	17,4%	23,5%	5,7%	12,0%	17,6%
MCB2	4,4%	11,9%	19,3%	8,3%	15,9%	21,7%	7,4%	13,5%	18,3%
SMCB	3,6%	9,9%	15,9%	10,1%	17,9%	24,0%	6,5%	13,8%	19,9%
NW									
ASY	11,5%	21,6%	29,7%	9,6%	18,9%	25,6%	5,2%	10,8%	15,5%
MBB	3,1%	8,8%	15,3%	11,3%	19,0%	24,9%	9,8%	16,5%	22,0%
MCB1	3,4%	9,2%	15,6%	8,2%	15,6%	22,0%	5,3%	10,7%	16,0%
MCB2	3,5%	10,0%	16,2%	7,0%	14,1%	20,4%	6,9%	13,2%	18,2%
SMCB	2,7%	8,2%	13,2%	8,6%	17,1%	23,4%	6,2%	12,5%	18,2%
VARHAC									
ASY	23,3%	32,9%	39,5%	17,2%	24,8%	30,2%	9,6%	14,7%	18,7%
MBB	2,3%	7,4%	14,1%	8,6%	17,8%	24,1%	8,3%	16,3%	22,4%
MCB1	2,6%	8,6%	14,2%	5,9%	14,0%	19,6%	4,0%	11,0%	16,9%
MCB2	2,4%	8,2%	14,4%	5,1%	12,0%	18,5%	5,4%	12,1%	17,7%
SMCB	2,0%	7,2%	13,0%	6,3%	15,7%	22,3%	4,9%	12,2%	18,0%

Таблица 3а: Нелинейная модель, гомоскедастичный вариант ($\gamma = 0.0$)

	$P\{\lvert t\rvert \geq q_\alpha^{sym}\}$			$P\{t \leq q_\alpha^{left}\}$			$P\{t \geq q_\alpha^{right}\}$		
	α=1%	α=5%	α=10%	α=1%	α=5%	α=10%	α=1%	α=5%	α=10%
W									
ASY	4,7%	11,6%	18,1%	3,9%	10,7%	17,0%	2,7%	7,3%	11,3%
MBB	2,8%	8,8%	15,2%	7,9%	16,3%	22,7%	9,0%	15,4%	20,6%
MCB1	2,1%	7,6%	13,5%	3,8%	10,9%	17,0%	4,2%	10,3%	15,7%
MCB2	2,8%	9,5%	15,9%	4,5%	11,8%	17,8%	5,3%	10,7%	16,3%
SMCB	2,3%	7,4%	12,8%	4,3%	11,2%	16,9%	5,4%	12,3%	18,1%
NW									
ASY	8,6%	18,1%	25,6%	7,7%	16,3%	23,1%	4,2%	9,4%	13,8%
MBB	3,2%	8,7%	15,1%	8,2%	16,3%	22,5%	8,5%	15,7%	21,2%
MCB1	2,4%	7,7%	13,5%	3,7%	9,9%	16,7%	4,7%	10,8%	16,9%
MCB2	3,1%	9,4%	15,5%	3,9%	10,7%	16,4%	5,7%	11,7%	17,1%
SMCB	2,0%	7,2%	12,7%	4,1%	12,0%	17,9%	5,1%	11,7%	18,4%
VARHAC									
ASY	20,3%	30,5%	37,0%	15,5%	23,3%	28,9%	8,5%	13,7%	18,1%
MBB	2,2%	7,5%	13,7%	6,1%	15,3%	22,0%	6,9%	14,7%	20,1%
MCB1	2,1%	7,9%	13,9%	3,4%	10,2%	17,6%	3,2%	10,1%	15,9%
MCB2	2,4%	8,1%	14,6%	2,9%	10,0%	16,6%	4,1%	10,1%	15,6%
SMCB	2,5%	7,5%	13,5%	3,8%	11,7%	18,8%	4,2%	12,0%	17,6%

Таблица 3б: Нелинейная модель, гетероскедастичный вариант ($\gamma = 0.5$)

	$P\{\|t\| \geq q_\alpha^{sym}\}$			$P\{t \leq q_\alpha^{left}\}$			$P\{t \geq q_\alpha^{right}\}$		
	α=1%	α=5%	α=10%	α=1%	α=5%	α=10%	α=1%	α=5%	α=10%
W									
ASY	5,5%	12,3%	18,7%	3,9%	10,6%	17,2%	3,7%	8,1%	12,8%
MBB	1,9%	7,2%	13,1%	9,1%	17,2%	23,2%	11,4%	18,4%	23,6%
MCB1	2,1%	6,9%	11,9%	6,1%	13,2%	18,6%	5,6%	12,2%	18,1%
MCB2	2,2%	7,2%	12,4%	6,3%	13,1%	19,0%	7,7%	14,2%	19,5%
SMCB	1,9%	6,2%	10,8%	7,0%	14,4%	20,5%	5,5%	12,6%	18,3%
NW									
ASY	9,9%	19,6%	27,6%	7,9%	16,3%	22,6%	5,5%	11,3%	16,3%
MBB	2,0%	7,2%	13,2%	10,4%	18,9%	25,0%	10,8%	17,5%	22,5%
MCB1	2,4%	6,8%	12,2%	6,4%	13,3%	19,4%	5,3%	11,8%	17,8%
MCB2	1,9%	7,0%	13,2%	5,8%	13,0%	18,9%	7,5%	15,2%	20,8%
SMCB	1,9%	6,4%	11,4%	7,0%	14,6%	20,7%	5,7%	13,8%	19,6%
VARHAC									
ASY	28,9%	38,8%	45,3%	19,9%	27,5%	32,3%	12,1%	17,8%	21,9%
MBB	2,7%	9,5%	15,8%	9,0%	19,0%	25,6%	9,9%	18,7%	24,1%
MCB1	2,4%	8,7%	15,5%	5,5%	13,9%	20,6%	5,6%	13,5%	19,5%
MCB2	3,2%	10,2%	17,2%	5,2%	12,9%	19,2%	7,3%	15,7%	22,1%
SMCB	2,7%	8,7%	14,8%	6,6%	15,4%	22,5%	6,2%	14,3%	20,1%