УДК 519.234.3

Проверка статистической значимости результатов механических торговых систем

Аннотация. *Статья посвящена статистическим методам проверки результатов механических торговых систем на наличие их предсказательной способности. В качестве торговой системы использовалась стратегия, основанная на пересечении двух скользящих средних. Для статистических проверок были использованы следующие методы: статический бутстреп и метод перестановки Монте-Карло. Были рассмотрены теоретические аспекты методов и их практическое использование. В результате статистических тестов было установлено, что скользящие средние не обладают предсказательной способностью.*

Ключевые слова: *торговые системы; статическая проверка; бутстреп; скользящие средние; WRC.*

Abstract. *The article is devoted to statistical methods for testing the results of automated trading systems for existence of predictive power. As for trading system used a strategy based on the intersection of two moving averages. For statistical tests the following methods used: bootstrap and permutation method of Monte Carlo. The theoretical aspects of the methods and their practical use were considered. Results of statistical tests found that moving averages are do not have predictive power.*

Keywords: *trading systems; statistical tests; bootstrap; moving averages; WRC.*

Дэмбэрэлсурэн Нямхуу,
студент Финансового университета
✉ *Nyamhyy14@yahoo.com*

Рассмотрим последовательность ежедневных значений цен акций ОАО «Сбербанк России» с 20 июня 2007 г. до 29 мая 2015 г. (1960 дней) [1]. Составим простую торговую стратегию на основе пересечения двух скользящих средних с разными периодами. При пересечении скользящей средней с меньшим периодом (быстрой) снизу вверх со скользящей средней с большим периодом (медленной) открываем позицию на покупку ценной бумаги. При обратном пересечении открываем позицию на продажу ценной бумаги.

Результаты торговой стратегии

Торговая стратегия имеет два параметра: n_1 и n_2, периоды скользящих средних ($n_1 < n_2$). Прибыльность торговой стратегии в основном зависит от оптимальности подобранных параметров. С помощью статистических программ было проведено тестирование стратегии при различных параметрах $n_1 \in [1; 90]$ и $n_2 \in [2; 100]$ при условии $n_1 < n_2$ (4905 комбинаций). При параметрах $n_1 = 13$ и $n_1 = 15$ была получена максимальная доходность — 853% (серая линия). Для сравнения: стратегия «Покупай и держи» (черная линия) имеет доходность 33% за весь период тестирования. Результаты сравнения представлены на *рис. 1*.

Как видим, наибольшую доходность стратегия пересечения скользящих средних принесла в период кризиса (2008–2009), когда цены акции падали высокими темпами. С одной стороны, получена значительная прибыль, что может говорить об опровержении гипотезы об отсутствии предсказательной способности, но, с другой стороны,

Научный руководитель: ***Зададаев С.А.,*** *кандидат физико-математических наук, доцент кафедры «Теория вероятностей и математическая статистика».*

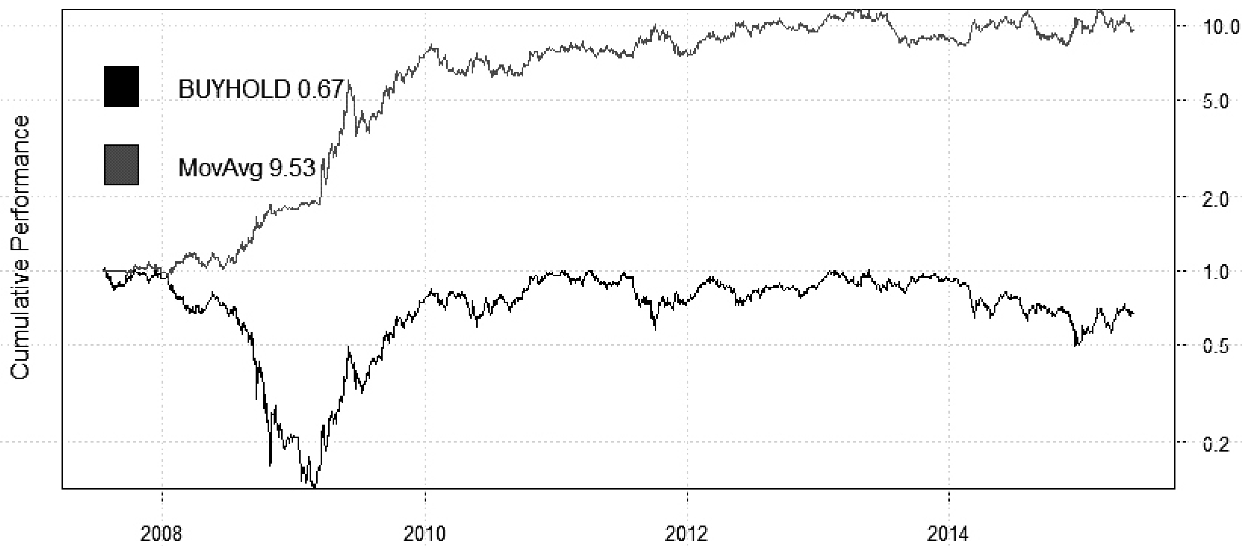

Рис. 1

возможна ситуация, когда такой результат был получен по чистой случайности. Поэтому необходимо провести тесты на проверку значимости полученных результатов.

Статистический бутстреп

Бутстреп-процедура (*bootstrap*) была предложена (*B. Efron*, 1979) как некоторое обобщение алгоритма «складного ножа», чтобы не уменьшать каждый раз число элементов по сравнению с исходной совокупностью. Основная идея бутстрепа по Б. Эфрону (1988) состоит в том, чтобы методом статистических испытаний многократно извлекать повторные выборки из эмпирического распределения. А именно: берется конечная совокупность из n членов исходной выборки x_1, x_2, …, x_{n-1}, x_n, откуда на каждом шаге из n последовательных итераций с помощью датчика случайных чисел, равномерно распределенных на интервале $[1, n]$, «вытягивается» произвольный элемент x_k, который снова «возвращается» в исходную выборку (т. е. может быть извлечен повторно).

Например, при $n = 8$ одна из таких комбинаций имеет вид $x_4, x_2, x_8, x_2, x_1, x_2, x_4, x_5$, т.е. отдельные элементы могут повторяться. Этим способом можно сформировать любое, сколь угодно большое число бутстреп-выборок. В результате легкой модификации частотного распределения реализаций исходных данных можно ожидать, что каждая следующая генерируемая псевдовыборка будет возвращать значение параметра, немного отличающееся от вычисленного для первоначальной совокупности. Образующийся разброс значений показателя дает возможность построения доверительных интервалов и других полезных выборочных параметров анализируемой величины.

Фундаментальной проблемой статистики является получение наиболее корректной оценки параметров выборочного распределения. Пусть дана выборка $x_1, x_2, \ldots, x_n$ и предполагается, что это набор независимых и одинаково распределенных реализаций случайной величины, извлеченных из генеральной совокупности X. Задача заключается в изучении свойств некоторой статистики $f(x_1, x_2, \ldots, x_n)$, которую мы трактуем как выборочную оценку. Обычно мы имеем некоторый сдвиг вычисленного значения параметра относительно его истинной величины, который вызывается многими причинами. Во-первых, выборочные значения имеют погрешность измерений, во-вторых, нет особенных гарантий, что выборка состоит из независимых и случайных значений, и, наконец, при оценке параметра мы обычно задаемся какими-то предположениями о законе распределения X. Бутстреп предоставляет более экономный способ, позволяющий обойтись без дополнительных измерений, и построить, например, доверительные интервалы анализируемой статистики на основе разброса значений анализируемого показателя, полученного в процессе имитации.

В контексте проверки гипотез нулевая гипотеза будет состоять в том, что торговая стратегия не имеет предсказательной способности (ожидаемая дневная доходность равна нулю). Альтернативная гипотеза: стратегия имеет предсказательную способность (ожидаемая дневная доходность не равна нулю). Для бутстрепа необходимо воспользоваться дневной доходностью стратегии и сформировать псевдовыборки с повторением. Но также необходимо убрать тренд из данных. Для обоснования данной процедуры приведем пример. Пусть цена за период тестирования выросла на 50%. И возьмем в качестве стратегии генератор случайных чисел на отрезке $[0; 1]$. Каждый день получаем случайное значение генератора, и если значение находится на полуинтервале $[0; 0{,}9]$, покупаем, а если значение находится на отрезке $[0{,}9; 1]$, продаем. Логично предположить, что доходность стратегии будет больше 0, и можно утверждать о наличии предсказательной способности. Хотя это было лишь случайностью. Полная детрендированная доходность тестируемой стратегии рассчитывается по следующей формуле для k торговли:

$$\ln\left(\text{Доходность}\right)=\mathop{e}_{\text{i=1}}^{\text{k}}\ln\frac{\text{Цена продажи}_i}{\text{Цена покупки}_i}-\frac{\text{Длительность торговли}_i}{\text{Длительность тестирования}}\times\ln\frac{\text{Цена}_{\text{конец}}}{\text{Цена}_{\text{начало}}}\times\left(-1\right)^{\text{D}},$$

где $\text{D}=0$ для позиций на покупку и $\text{D}=1$ для позиций на продажу.

$\text{Цена}_{\text{начало}}$ и $\text{Цена}_{\text{конец}}$ — цены на начало и конец тестирования.

В общем случае детрендированная доходность[1] стратегий на каждый i день рассчитывается по представленной ниже формуле при условии, что в этот день была открытая позиция на покупку или на продажу, иначе $\text{Доходность}_i=0$:

$$\text{Доходность}_i=\ln\frac{\text{Цена закрытия}_i}{\text{Цена открытия}_i}-\ln\frac{\text{Цена}_{\text{конец}}}{\text{Цена}_{\text{начало}}}\times\left(-1\right)^{\text{D}}.$$

Рассмотрим последовательность шагов для бутстрепа:

1. Средняя дневная доходность вычитается из доходности каждого дня (центрирование) — исходная выборка.
2. Для каждой псевдовыборки выбираем n (объем исходной выборки) дневных центрированных доходностей в случайном порядке с повторением и рассчитываем среднюю дневную доходность (бутстреп средняя).
3. Повторяем достаточно большое количество раз второй шаг.
4. Формируем распределение из бутстреп средних.
5. Получаем — значение (вероятность ошибки при отклонении нулевой гипотезы).

При объеме **1960** доходностей в каждой псевдовыборке было сформировано **30 000** псевдовыборок. Гистограмма распределения представлена на *рис. 2*.

Традиционно в качестве наблюдаемого значения используется ожидаемая дневная доходность исходной выборки без центрирования [2], что фактически эквивалентно стандартной проверке гипотезы о нулевом математическом ожидании нормально распределенной доходности при неизвестной дисперсии:

$$t_{набл}=\frac{\overline{x_в}}{\sigma_в}\sqrt{n}\sim t(n-1),$$

где $t\,(n-1)$ — распределение Стьюдента с $n-1$ степенью свободы. При этом критическая область с уровнем значимости α для, например, двусторонней альтернативы будет иметь вид:

$$|\,t_{набл}\,|\geq t_{1-/2}(n-1),$$

где $t_{1-\alpha/2}\,(n-1)$ — соответствующая квантиль распределения.

---

[1] Для всех тестов используются только детрендированные дневные доходности.

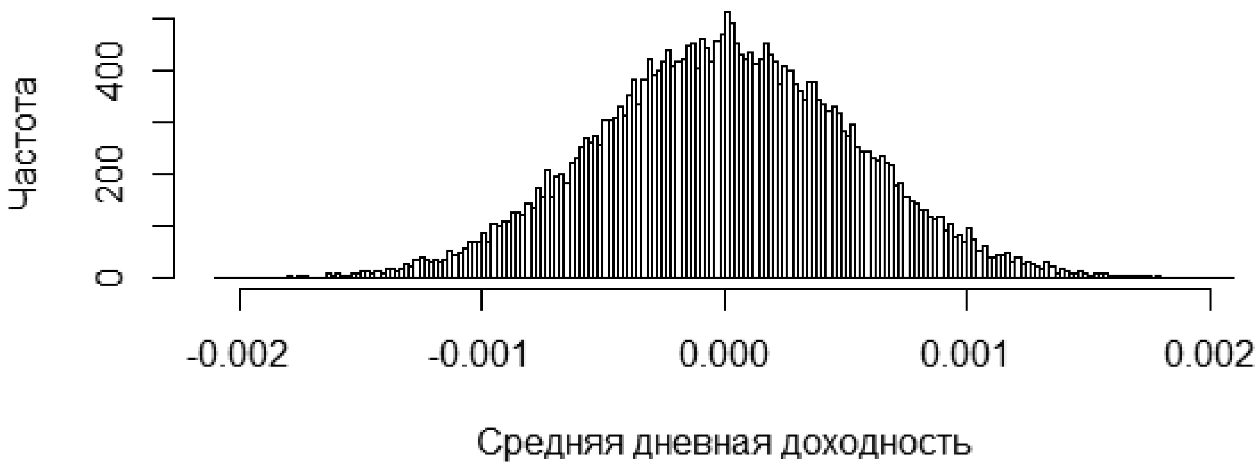

Рис. 2

В нашем случае наблюдаемое значение средней дневной доходности оказывается на уровне **0,1152% (0,001152)**. Количество значений из бутстрепа, которые превышают это уровень, равно **441**. При объеме 30 000 $\boldsymbol{p}$–значение равно $\frac{441}{30000} = \mathbf{0{,}0147}$. На уровне значимости $\alpha = \mathbf{0{,}05}$ можно отвергнуть нулевую гипотезу о нулевой или отрицательной средней дневной доходности.

Однако следует учесть тот факт, что параметры скользящих средних у исходной были оптимизированы. Следовательно, возможна ситуация, когда параметры были подогнаны под историю и имеется большое количество прибыльных дней. Поэтому проведем проверку Уайта.

White Reality Check (WRC) использует бутстреп для извлечения выборочного распределения подходящего для тестирования нулевой гипотезы, что все тестированные правила имеют ожидаемую доходность, равную нулю [3].

Для наглядности возьмем упрощенный пример *WRC*. Пусть исследователь протестировал два правила ($\boldsymbol{N} = \mathbf{2}$) в течение 5 дней ($\boldsymbol{T} = \mathbf{5}$). Первое правило имеет ожидаемую доходность 1%, а второе — 2%.

1. Средняя дневная доходность вычитается из доходности каждого дня (центрирование) для каждого правила.

2. Составляем выборку с повторением, размера $\boldsymbol{T}$ из $\mathbf{1, 2, \ldots, T}$. Предположим, что получилась следующая выборка: 2, 3, 4, 1, 1.

3. Для первого правила берем доходности 2, 3, 4, 1, 1 дней и находим ожидаемую доходность. И то же самое проделываем для второго правила. Максимальная ожидаемая доходность, полученная среди правил, является значением для выборки.

4. Повторяем шаги со 2-го до 3-го достаточно большое количество раз (например, $\boldsymbol{L} = \mathbf{15000}$ раз).

5. Выборочное распределение максимальных ожидаемых доходностей из совокупности $\boldsymbol{N} = \mathbf{2}$ правил, у которых ожидаемая доходность равна нулю, формируется из $\boldsymbol{L} = \mathbf{15000}$ значений.

6. $\boldsymbol{p}$–значение может быть рассчитано как доля из $\boldsymbol{L} = \mathbf{15000}$ значений, превышающих максимальную ожидаемую дневную доходность лучшего правила (**2%)**.

Было тестировано $\boldsymbol{N} = \mathbf{4905}$ правил (комбинаций параметров). Период тестирования составляло $\boldsymbol{T} = \mathbf{1960}$ дней, и выборка получена из $\boldsymbol{L} = \mathbf{15000}$ значений. Гистограмма распределения представлена на *рис. 3*. Максимальная ожидаемая доходность равна **0,1152% (0,001152),** $\boldsymbol{p}$–значение равно $\frac{3561}{15000} = \mathbf{0{,}2374}$ и уровень значимости $\alpha = \mathbf{0{,}05}$. Следовательно, нулевая гипотеза принимается на уровне значимости $\alpha = \mathbf{0{,}05}$.

Метод перестановки Монте-Карло

Метод перестановки Монте-Карло (МК) также может быть использован для генерирования выборочного распределения, необходимого для оценки статистической значимости найденных правил при тестировании. Оба *WRC* и МК проверяют нулевую гипотезу, что вся совокупность

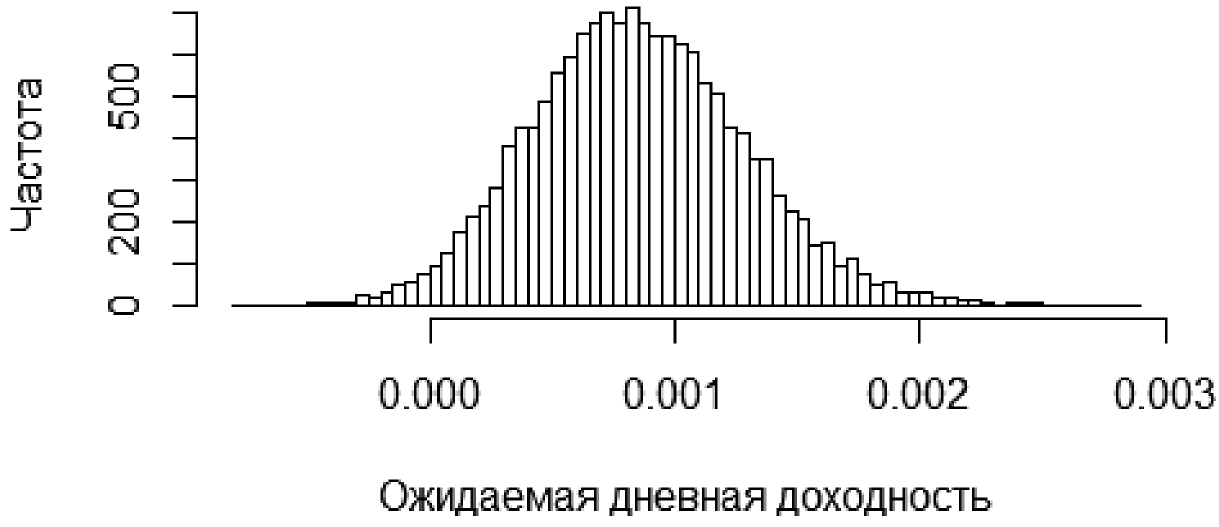

Рис. 3

тестированных правил не имеет предсказательной способности. Но МК формулирует гипотезу немного по-другому. МК сочетает выходные значения правила случайно с дневным изменением цены из $\boldsymbol{T}$. По сравнению с *WRC* МК случайно сочетает без повторения. Доходность, полученная в результате случайного сочетания, становится критерием сравнения с исходной доходностью. Если правила обладают предсказательной способностью, то результат сочетания должен выдавать существенно лучшие результаты.

Доходность случайного сочетания рассчитывается умножением выходного значения правила (+1 и −1 для покупки и продажи соответственно) на случайно перестановленные изменения цены. Если выходное значение правила +1 и случайно выбранное изменение цены дня положительное, то, следовательно, и доходность будет положительной. Убытки получаются в те дни, когда знаки противоположные. После расчета доходности для правила рассчитывается ожидаемая доходность. Это проделывается для каждого правила из $\boldsymbol{N}$, и будет получено $\boldsymbol{N}$ ожидаемых доходностей. Максимальная ожидаемая доходность среди $\boldsymbol{N}$ правил выбирается в качестве значения для выборочного распределения. Весь процесс повторяется достаточно большое количество раз ($\boldsymbol{L > 500}$).

После получения выборочного распределения рассчитывается -значение, равное количест-

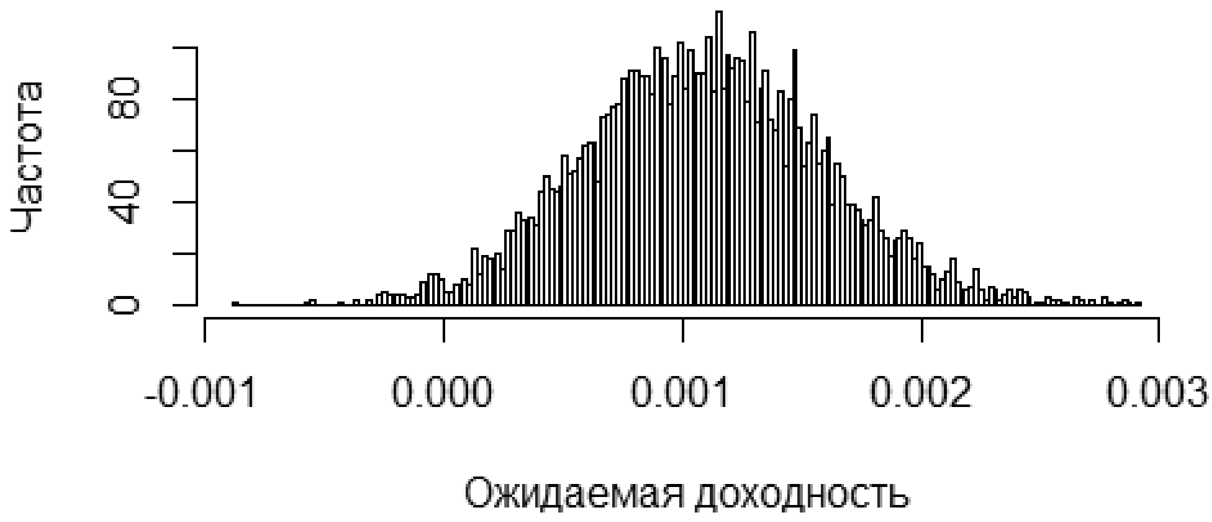

Рис. 4

День	1	2	3	4	5	6	7	8
Позиция	+1	+1	+1	0	0	−1	−1	−1
Дневное изменение цены	0.54%	−0.32%	1.54%	0.69%	−1.02%	−0.68%	1.20%	−2.50%
Результат	0.54%	−0.32%	1.54%	0.00%	0.00%	0.68%	−1.20%	2.50%
Ожидаемая доходность	**0.47%**							

День	1	2	3	4	5	6	7	8
Позиция	+1	+1	+1	0	0	−1	−1	−1
Дневное изменение цены	0.54%	0.69%	1.54%	−0.66%	1.20%	−0.32%	−2.50%	−1.02%
Результат	0.54%	0.69%	1.54%	0.00%	0.00%	0.32%	2.50%	1.02%
Ожидаемая доходность	**0.83%**							

ву значений выборки, превышающих ожидаемую доходность лучшего правила. Ограничением МК является отсутствие возможности построения доверительного интервала, так как МК не проверяет гипотезу об ожидаемой доходности правила. Нулевая гипотеза МК состоит в том, что тестированные правила имеют случайную корреляцию с будущим рыночным изменением. Значение может быть рассчитано как доля из $\boldsymbol{L}$ значений, превышающих максимальную ожидаемую дневную доходность лучшего правила.

Для иллюстрации представлены некоторые возможные комбинации для одного правила (*табл.*).

При $\boldsymbol{N = 4905}$ тестированных правилах, $\boldsymbol{T = 1959}$ данных о изменений цены за период тестирования и максимальной ожидаемой доходности **0,1152% (0,001152)**, было проведено $\boldsymbol{L = 6000}$ перестановок.

В результате МК-значение равно $\frac{2663}{6000} = \mathbf{0{,}4438}$.

На уровне значимости $\alpha = \mathbf{0{,}05}$ принимается нулевая гипотеза о случайной корреляции выходного значения правила с будущим изменением цены. Гистограмма распределения представлена на *рис. 4*.

Заключение

Статистические тесты подтвердили гипотезу об отсутствии предсказательной способности у стратегии скользящих средних. Было показано, что большая доходность в период тестирования стратегии не может свидетельствовать о наличии предсказательной способности. Также следует учесть тот факт, что тестирование проводилось без управления капиталом. Каждый раз покупались или продавалась ценные бумаги на одинаковую сумму. В реальной жизни происходит реинвестирование прибыли, и можно проводить различные по объему сделки. Используя бутстреп при наличии репрезентативной выборки генеральной совокупности, можно находить распределение любого представляющего интерес параметра.

Литература

1. Финам. Электронный ресурс: http://bit.ly/1L1rBcJ (дата обращения: 06.06.2015).
2. Aronson D. Evidence-based technical analysis: applying the scientific method and statistical inference to trading signals. John Wiley & Sons, 2011. T. 274.
3. White H. A reality check for data snooping // Econometrica, 2000. C. 1097–1126.