

Пластичность волатильности

СЕРГЕЙ С. СТЕПАНОВ ¹

АННОТАЦИЯ

Обсуждаются проблемы нестационарности финансовых рынков, связанные с динамическим характером волатильности цен. Предложена новая методика вычисления текущей волатильности рынка. Дано простое и наглядное объяснение причин появления заметных корреляций между последовательными значениями волатильностей или квадратов доходности. Показано, что после устранения нестационарности автокорреляции существенно снижаются, становясь статистически незначимыми. Рассмотрены причины негауссова характера вероятности распределения доходностей. На данных фондового и валютного рынков продемонстрировано, что устранение нестационарности существенно снижает экспесс распределения, приближая его к нормальной функции Гаусса. Предложены статистические критерии контроля степени сглаживания эмпирических значений волатильности. Вводится гипотеза о гладкой, нестохастической природе волатильности и причинах её изменения.

1 Введение

По-видимому, наиболее характерным свойством финансовых рынков является их нестационарность. Статистические параметры ценовой динамики изменяются со временем. Этот факт неприятен как для исследователей, так и для участников рынка, так как обнаруженные закономерности или построенные торговые системы со временем утрачивают свою актуальность. Наилучшим решением проблемы нестационарности стало бы включение её в вероятностную модель функционирования рынка.

Одной из важнейших характеристик доходности финансового инструмента является его волатильность. Не вызывает сомнения, что волатильность изменяется со временем [1]. Существуют периоды "спокойного" поведения рынка и периоды с повышенной волатильностью. Волатильность характеризует "температуру" рынка, величину его эмоционального напряжения. Предсказание её будущего значения является исключительно важным, например, для определения цен опционов или управления рисками портфельных инвесторов. Понимание причин и характера нестационарности волатильности приводит также к более глубокому проникновению в сущность функционирования финансовых рынков.

¹Research Center of "Altus Assets Activities", www.Altus.ua,
. Department of Theoretical Physics, Dnepropetrovsk National University, Ukraine,
. E-mail: steps@Altus.ua

В термин *волатильность* вкладывают, как минимум, четыре различных смысла - 1) эмоциональная характеристика рынка, 2) выборочное средне - квадратичное отклонение логарифмических доходностей, 3) его "истинное" значение по генеральной совокупности, и 4) подразумеваемая волатильность в опционных контрактах [2]. Мы используем волатильность во втором и третьем смыслах. Нас будет интересовать, какое определение выборочной волатильности приводит к минимальной ошибке в той или иной *модели* случайного процесса. "Истинная" волатильность, естественно, является величиной ненаблюдаемой, и вопрос её природы и оценки особенно усложняется нестационарным характером рынков.

Общепринятым подходом является рассмотрение волатильности как стохастической величины [3], [4], [5]. Одна из причин этого - высокие автокорреляционные коэффициенты между волатильностями или квадратами доходности, посчитанными за различные последовательные периоды времени [6]. После практических нулевых корреляций между логарифмическими доходностями обнаружение подобной закономерности оказывает неизгладимое впечатление.

В вероятностных моделях с переменной волатильностью блуждание цены $x(t)$ описывается дискретными или непрерывными уравнениями, параметры которых, в свою очередь, являются случайными. Широкую популярность получила модель GARCH(p, q) и её всевозможные обобщения [7], [8]. В этом случае временная шкала разбивается на конечные интервалы времени (лаги) длительностью τ , так что рассматриваются только цены "закрытия" этих интервалов $x_k = x(k \cdot \tau)$, где $k = 0, 1, \dots$. Логарифмические доходности $r_k = \ln x_k / x_{k-1}$ являются независимыми случайными величинами, с переменной волатильностью σ_k , квадрат которой линейно зависит от предыдущих квадратов волатильностей и доходностей:

$$r_k = \sigma_k \varepsilon_k, \quad \sigma_k^2 = \alpha_0 + \sum_{i=1}^p \alpha_i \sigma_{k-i}^2 + \sum_{i=1}^q \beta_i r_{k-i}^2. \quad (1)$$

Здесь и далее ε_k - это нескоррелированные нормированные случайные числа с нулевым средним и единичной дисперсией: $\overline{\varepsilon_i} = 0$, $\overline{\varepsilon_i^2} = 1$, $\overline{\varepsilon_i \cdot \varepsilon_j} = 0$. Черта сверху, как обычно, обозначает среднее по всем возможным реализациям ε_i .

В непрерывном классе моделей как для ценовой динамики, так и для динамики волатильности выбирается то или иное стохастическое уравнение Ито. Например [9], цена может совершать обычное логарифмическое блуждание, а волатильность - подчиняться уравнению Орнштейна - Уленбека:

$$\frac{dx}{x} = \mu dt + \sigma(t) \delta W_1, \quad d \ln \sigma = \beta \cdot (\alpha - \ln \sigma) dt + \gamma \delta W_2, \quad (2)$$

где $\delta W_1, \delta W_2$ - нескоррелированные винеровские переменные $\delta W = \varepsilon \sqrt{dt}$. Собственно, именно для этого класса моделей принято использование термина "стохастическая волатильность", хотя мы его будем использовать в несколько более широком смысле.

И в дискретных, и в непрерывных моделях иногда вводится одна или несколько "скрытых" стохастических переменных, функцией которых является волатильность. Возможны и другие, достаточно сложные подходы [29]. Общее, что их объединяет, - это вероятностное описание локальной динамики значений волатильности (дискретное или непрерывное) [10].

Существуют обширные эмпирические исследования, посвящённые тестированию предсказательной силы стохастических моделей [11], и критика их методики [12]. В целом, в последние годы, начал накапливаться определённый скепсис касательно прогностических возможностей подобных моделей. Недавно были высказаны соображения о том, что автокорреляционные коэффициенты положительно определённых величин могут возникать по причине их нестационарности [30].

С эффектом нестационарности непосредственно связана также проблема поиска распределения вероятностей доходности финансового инструмента. Хорошо известно, что это распределение не описывается функцией Гаусса. Оно имеет толстые хвосты и, как следствие, высокий эксцесс и большие вероятности заметных выбросов доходностей. Начиная с революционных работ Мандельброта [13], этот факт постепенно стал стандартом в финансовом инженеринге. Однако, при построении распределения вероятности случайных чисел неявно подразумевается их стационарность, которой мы не имеем на финансовых рынках.

Идея о том, что нестационарность может быть причиной негауссовой распределения доходности восходит ещё к классической работе [31]. Там же она подвергается критике. Тем не менее, вопрос о характере распределения и влиянии на него нестационарности требует более аккуратного рассмотрения.

Ниже мы приводим аргументы в пользу того, что волатильность $\sigma(t)$ является *гладкой*, а не стохастической функцией времени. Будет дано объяснение происхождения высоких автокорреляционных коэффициентов и негауссовой распределения доходностей. Сформулирована гипотеза о том, что волатильность обладает свойством пластиичности. Под воздействием нерегулярных, относительно редких и принципиально не предсказываемых шоковых воздействий на рынок она постепенно деформируется. После прекращения подобных воздействий происходит процесс релаксации, в результате которого волатильность плавно снижается.

Последовательность материала статьи следующая. Сначала мы обсудим новую меру измерения волатильности и продемонстрируем её эффективность. Затем будут перечислены основные эмпирические свойства автокорреляционных коэффициентов, связанных с волатильностью, и предложена простая нестационарная модель, в которой подобная автокорреляция возникает. Механизму появления автокорреляций с характерными свойствами будет дана очень наглядная графическая форма и предложен простой математический аппарат для проведения необходимых вычислений.

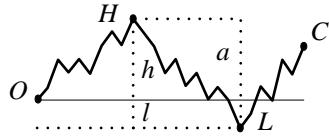
Свидетельством того, что подобный механизм реализуется на финансовых рынках, служат нулевые автокорреляционные коэффициенты для разностей волатильности соседних дней или её остатков после выделения гладкой составляющей $\sigma(t)$. Будут приведены эмпирические доказательства этих фактов на примерах фондового и валютного рынков.

Далее мы покажем, что нормирование доходностей на $\sigma(t)$ приводит к существенному снижению эксцесса распределения, приводя в ряде случаев его к нормальному виду. Будут разработаны определённые статистические критерии контроля степени сглаживания данных. Рассмотрены аргументы касательно локального постоянства "истинной" волатильности. В заключении сформулирован ряд предположений о возможных свойствах динамики волатильности.

В приложения, представляющие собой замкнутые информационные блоки, вынесены различные технические детали.

2 Измерение волатильности

Исторические данные ценовой динамики различных финансовых инструментов обычно агрегированы в точки, между которыми существует определённый период времени (лаг). Наиболее доступными являются дневные лаги, реже встречаются часовые, и ещё реже - минутные. Кроме цены закрытия C_t (последнее значение лага), общепринятыми являются: котировка открытия O_t (первая цена лага), максимальное H_t и минимальное L_t значения цены. При помощи этих четырех цен можно образовать три независимые *относительные* величины, которые мы будем называть *базисом* лага:



$$\begin{aligned} h_t &= H_t - O_t \\ l_t &= O_t - L_t \\ r_t &= C_t - O_t \end{aligned} \quad (3)$$

Рис.1: Характеристики волатильности

Высота подъема h и *глубина* опускания цены l – величины положительные. С их помощью определяется *амплитуда размаха* цены $a = h + l$. Доходность r может быть как положительной, так и отрицательной.

В моделях аддитивного винеровского блуждания $dx = \mu dt + \sigma dW$ четвёрка $\{O_t, H_t, L_t, C_t\}$ непосредственно являются ценами. Для логарифмического блуждания $dx/x = \mu dt + \sigma dW$ это логарифмы ценовых значений $\ln x$. Поэтому в этом случае, например, размах a_t будет равен логарифму отношения максимума к минимуму цены $a_t = \ln H_t/L_t$, доходность – логарифмической доходности $r_t = \ln C_t/O_t$, и т.д.

Под волатильностью σ лага длительностью T будем подразумевать усреднение по достаточно большому числу лагов отклонения доходности r от среднего: $\sigma^2 = \langle (r - \bar{r})^2 \rangle$. Если волатильность σ постоянна, то значения положительных величин $\{h, l, |r|, a\}$ в той или иной мере характеризуют её значение. Чем волатильность рынка выше, тем вероятнее будут их большие значения. В частности, если сноса нет ($\mu = 0$), средние значения пропорциональны волатильности: $\bar{a} = 1.596 \cdot \sigma$, $\bar{h} = \bar{l} = |r| = 0.798 \cdot \sigma$ (см. приложение А).

Однако, информационная значимость каждого из параметров и их возможных комбинаций различна. Распределения плотности вероятности для $P(a)$ и $P(h)$, $P(l)$, $P(|r|)$, в случае винеровского процесса, имеют следующий вид (пунктирные линии – это средние значения при $\sigma = 1$):

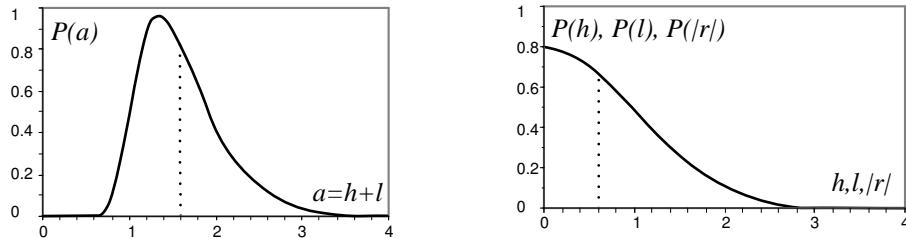


Рис.2: Плотности вероятности базиса

Из четвёрки $\{a, h, l, |r|\}$ только амплитуда размаха имеет достаточно узкий максимум в окрестности среднего значения. Плотности вероятности остальных величин монотонно снижаются с их ростом. Заметим также, что h , l и $|r|$ с высокой вероятностью могут иметь значения близкие к нулю. Амплитуда размаха " a ", наоборот, избегает быть нулевой и вероятность того, что $a < 0.75\sigma$, равна 0.002. Так, часто рынок закрывается с близким к нулю приростом цен $|r| \sim 0$, в то время как его волатильность в течении дня была существенной.

Чем уже распределение вероятности для меры измерения волатильности, тем она лучше. Для положительно определённой величины v относительную степень узости распределения можно характеризовать отношением σ_v/\bar{v} , где $\sigma_v^2 = (v - \bar{v})^2$ - усреднение квадратов отклонения от среднего \bar{v} . Для амплитуды размаха $\sigma_a/\bar{a} = 0.30$, что в более чем в два раза меньше аналогичного отношения, например, для высоты $\sigma_h/\bar{h} = 0.76$.

Возникает естественный вопрос - не существует ли комбинации базисных величин $f(h, l, r)$, которая имела бы более узкое распределение, чем амплитуда ценового размаха "a"? Этой теме посвящена обширная литература (см., например [21] - [25]).

В этой статье мы будем использовать простую модификацию амплитуды размаха. Если динамика цены внутри лага сопровождается существенным трендом $|r| \neq 0$ (не важно, вверх или вниз), то волатильность может быть ниже, чем при том же размахе, но в отсутствие тренда ($|r| = 0$). Поэтому целесообразно понижать значение амплитуды размаха в случае больших $|r|$. Будем называть *модифицированной амплитудой размаха* следующую величину:

$$v = a - \frac{|r|}{2}.$$

Ниже в таблице приведены статистические параметры (среднее значение (av), среднеквадратичное отклонение от него (si), асимметрия распределения (as) и его эксцесс (ex) при $\sigma = 1$:

Таблица 1.: Статистические параметры распределений для $|r|$, a и v .

	av	si	as	ex	si/av
$ r $	0.798	0.603	1.00	0.87	0.76
a	1.596	0.476	0.97	1.24	0.30
v	1.197	0.300	0.53	0.26	0.25

Видно, что относительная ширина распределения модифицированной амплитуды размаха $\sigma_v/\bar{v} = 0.25$, что лучше, чем у простого размаха a . Статистические параметры показывают также, что распределение для "v" более симметрично вокруг максимума и имеет меньший эксцесс, чем "a". Приведём вид распределения для $P(v)$ вместе с $P(a)$ (пунктир), и выражения для среднего v и его квадрата в случае броуновского блуждания:

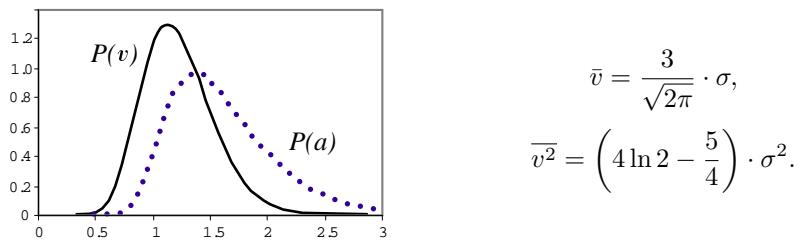


Рис.3: Плотности вероятности $P(v)$, $P(a)$ и средние для v и v^2

Таким образом, модифицированная амплитуда размаха оказывается лучшей мерой волатильности, чем простой размах, и существенно лучше, чем модуль логарифмической доходности. В приложении B. мы сравниваем модифицированную амплитуду размаха с некоторыми другими известными способами измерения волатильности. При аналогичной или меньшей ошибке определения волатильности мера "v" имеет существенно более простое определение и несомненность для малого числа лагов, поэтому далее в статье мы будем широко её использовать.

3 Внутридневная волатильность

Продемонстрируем эффективность модифицированной амплитуды размаха на реальных данных. Рассмотрим 15-минутные котировки на рынке Форекс за период с 2004 по 2008 год для пары EURUSD. Произведём их агрегирование в дневные точки, вычисляя, кроме минимального и максимального значения, также внутридневную волатильность на основе логарифмических доходностей 15-минутных лагов:

$$\sigma^2 = \frac{n}{n-1} \sum_{i=1}^n (r_i - \bar{r})^2. \quad (4)$$

В течение дня мы имеем $n = 96 = 4 \cdot 24$ пятнадцатиминутных лагов. Множитель n в (4) приводит 15-минутную волатильность к дневному значению. Временная динамика *внутридневной волатильности* имела следующий вид (1250 торговых дней, без учёта выходных и основных праздников):

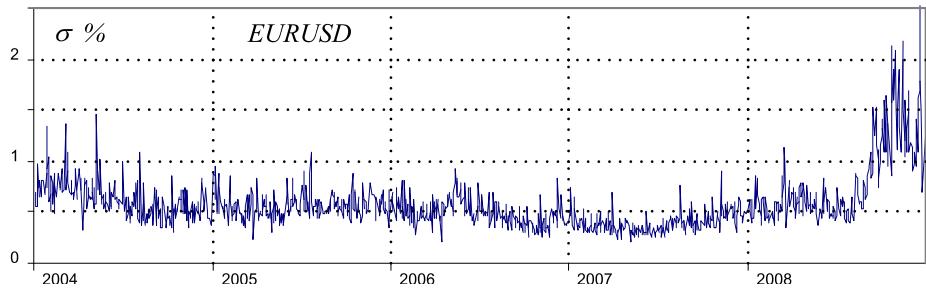


Рис.4: Внутридневная волатильность EURUSD

Видно, что, начиная с осени 2008-го года, волатильность валютного рынка, как, впрочем, и других финансовых рынков, существенно выросла, в связи с обострением финансового кризиса. Однако, даже в докризисный период волатильность имеет ярко выраженную нестационарную составляющую.

Предположим, что полученная по выборке из $n = 96$ чисел выборочная волатильность лучше характеризует "истинную" волатильность, чем только дневной базис из трёх чисел $\{h, l, r\}$ [14], [15], [16], [17]. Более качественная мера волатильности, основанная на базисе, должна быть лучше скоррелирована с внутридневной волатильностью. Построим точечные диаграммы зависимости ежедневных значений v_t , a_t и $|r_t|$ от внутридневной волатильности σ_t (EURUSD 2004-2008):

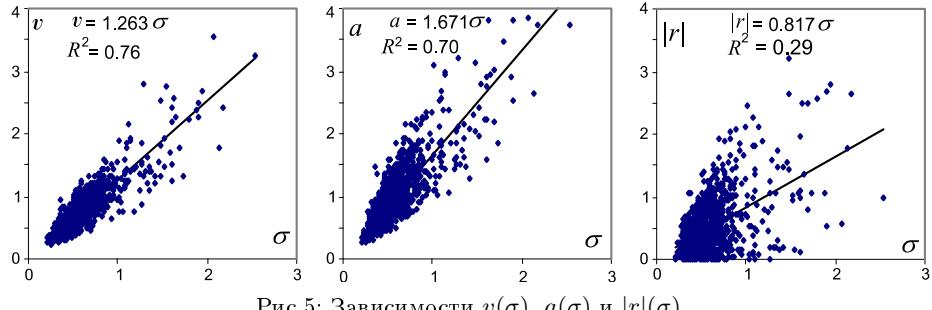


Рис.5: Зависимости $v(\sigma)$, $a(\sigma)$ и $|r|(\sigma)$

Несложно видеть, что v_t и a_t существенно лучше связаны с σ_t , чем $|r_t|$. Переход от амплитуд "a" к модифицированным амплитудам "v" даёт определённый выигрыш, однако он, конечно, не столь значителен.

Похожие результаты получаются и для других валют. Наклоны регрессионных прямых v_t/σ_t и a_t/σ_t для шести валютных пар имеют следующие значения:

	eurusd	gbpusd	usdchf	usdjpy	usdcad	audusd	average
$\langle v/\sigma \rangle$	1.263	1.260	1.289	1.251	1.241	1.243	1.258
$\langle a/\sigma \rangle$	1.671	1.665	1.692	1.640	1.621	1.660	1.658
$\langle r /\sigma \rangle$	0.817	0.809	0.807	0.776	0.761	0.834	0.801

В каждом случае ошибка линейной аппроксимации для v была меньше, чем для a , и существенно меньше, чем для $|r|$.

Несмотря на заметный разброс, значения v/σ , a/σ и $|r|/\sigma$ близки к своим теоретическим величинам 1.197, 1.596 и 0.798, возникающим при случайному винеровском блуждании. Тем не менее, необходимо помнить, что, например, отношение $v/\sigma = 3/\sqrt{2\pi}$ справедливо только для броуновского блуждания с *нормальным* распределением доходности лага. На практике это не совсем так, и отношение v/σ может быть равно некоторой константе, отличной от $3/\sqrt{2\pi}$, определение значения которой мы обсудим ниже.

Ещё одним признаком значимости модифицированной амплитуды размаха цены являются автокорреляционные коэффициенты, которые будут объектом нашего интереса:

$$\rho_s(v) = \text{cor}(v_t, v_{t-s}) = \frac{\langle (v_t - \bar{v})(v_{t-s} - \bar{v}) \rangle}{\sigma_v^2}, \quad (5)$$

где усреднение проводится по всем наблюдаемым значениям $v_t = v_1, \dots, v_n$. Для дневных курсов EURUSD (2004-2008) получаются следующие диаграммы автокорреляции как функции параметра сдвига s :

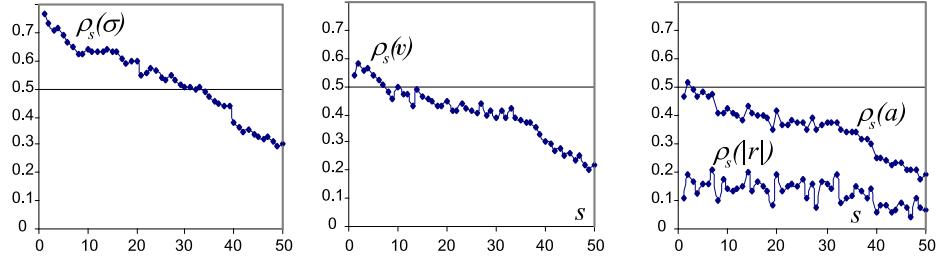


Рис.6: Коррелограммы волатильности для EURUSD

Видно, что наиболее высокими являются автокорреляции внутридневной волатильности $\rho_1(\sigma) = 0.77$, затем идут модифицированные амплитуды размаха цены $\rho_1(v) = 0.54$, простой амплитуды $\rho_1(a) = 0.47$, и самые маленькие значения коэффициентов у модуля логарифмической доходности $\rho_1(|r|) = 0.11$.

Высокие автокорреляции появляются при рассмотрении самых разнообразных финансовых инструментов и являются довольно интригующим фактом [27]. Для сравнения, 1-й автокорреляционный коэффициент доходности курса EURUSD равен $\rho_1(r) = -0.02$, что с учётом двойной статистической ошибки 0.06 (1250 торговых дней) соответствует отсутствию корреляции. Эта непредсказуемость рынка является проявлением его эффективности.

Однако для модулей доходности, и тем более волатильностей, это не так. На основании этого факта строится огромное количество стохастических моделей, претендующих на предсказание будущих значений волатильности. В большинстве своём эти модели носят *эмпирический* характер, не объясняя *причин* появления автокорреляций. Одной из наших задач будет предложить подобное объяснение.

4 Эмпирические особенности автокорреляций

В целях дальнейшего анализа отметим ряд особенностей поведения автокорреляционных коэффициентов, связанных с волатильностью.

1. Автокорреляции убывают монотонно и очень медленно.

Этот результат широко известен. Существует ряд исследований по определению функциональной зависимости автокорреляционного коэффициента от параметра сдвига s . Обычно автокорреляции аппроксимируют степенной функцией $s^{-\mu}$. При этом параметр μ оказывается достаточно маленьким.

2. Автокорреляции тем выше, чем шире временной интервал.

Рассмотрим поведение автокорреляционных коэффициентов ежедневной модифицированной амплитуды размаха $v = a - |r|/2$ для фондового индекса S&P500 за период с 2001 по 2006 год. Разобьём этот интервал на два трёхлетних периода 2001-2003 и 2004-2006. В первом случае было $n = 752$ торговых дня, а во втором - $n = 755$. Вычислим автокорреляционные коэффициенты каждого периода по отдельности и автокорреляцию по объединённым данным.

Результирующие автокоррелограммы представлены на рисунке ниже (объединённая автокорреляция повторена на каждом из них):

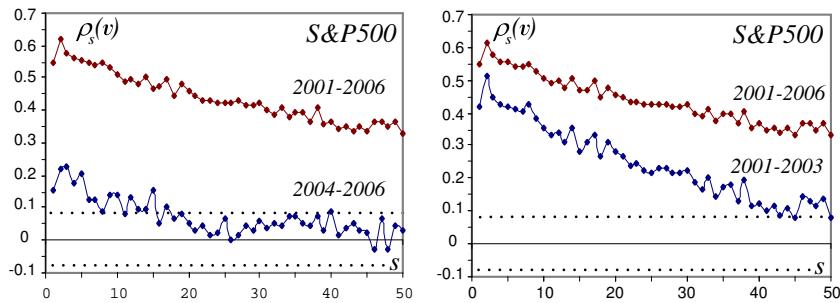


Рис.7: Коррелограммы S&P500 за различные интервалы времени

Видно, что суммарная коррелограмма лежит выше коррелограмм каждого из периодов. Эта закономерность, вообще говоря, выполняется не всегда, и условия, при которых она возникает, будут понятны из дальнейших рассуждений.

Здесь и далее пунктирные горизонтальные линии на коррелограммах образуют коридор с двойной ошибкой $\pm 2/\sqrt{n}$, где n - количество чисел, участвующих в вычислении автокорреляционных коэффициентов. В таблице приведены основные статистические параметры ежедневной логарифмической доходности индекса S&P500 на различных этапах:

Период	n	\bar{r}	σ	as	ex	p_0	p_1	$\rho_1(v)$
2004-2006	755	0.032	0.659	-0.02	0.25	55.9	69.4	0.16
2001-2003	752	-0.023	1.376	0.20	1.27	48.9	71.4	0.42
2001-2006	1507	0.005	1.078	0.15	2.84	52.4	75.7	0.55

Кроме среднего (\bar{r}), дневной волатильности σ , асимметрии (as) и эксцесса (ex) вычислены доли в процентах положительных доходностей $p_0 = p(r > 0)$ и доли их попаданий в одну сигму: $p_1 = p(|r - \bar{r}| < \sigma)$. Отметим, что p_1 существенно устойчивее по отношению к большим выбросам, чем эксцесс. Так, за период 1950-2008 ($n = 14844$) мы получим $ex = 32.2$, $p_1 = 78.9$. Замена только трех последовательных дней, сопровождающих биржевой крах 1987 года с понедельника, 19 октября, $r = -22.9\%$, $+5.2\%$, $+8.7\%$, на одно суммарное падение $r = -9\%$ в три раза уменьшает эксцесс $ex = 10.5$, и лишь незначительно - вероятность $p_1 = 78.3$.

Из приведенной выше таблицы видно, что когда рынок спокоен (2004-2006: $\sigma = 0.659\%$), он достаточно близок к нормальному ($ex = 0.25$). После расширения интервалов времени нормальность существенно ухудшается. Одновременно с этим начинает возрастать автокорреляция волатильностей $\rho_1(v) = \text{cor}(v_t, v_{t-1})$.

Аналогично обстоит ситуация на валютном рынке. Отбрасывание кризисного 4-го квартала 2008 года существенно снижает автокорреляционные коэффициенты статистик, связанных с волатильностью для EURUSD:

Период	n	$\rho_1(\sigma)$	$\rho_1(v)$	$\rho_1(a)$	$\rho_1(r)$
1Q-2004 .. 4Q-2008	1302	0.80	0.54	0.47	0.11
1Q-2004 .. 3Q-2008	1215	0.51	0.25	0.16	0.01

Заметим, что при этом происходит уменьшение количества дней, по которым вычисляются автокорреляционные коэффициенты, всего на 7%.

3. Точечная диаграмма волатильности имеет форму "венника".

Построим точечные графики модифицированных амплитуд размаха $\{v_{t-1}, v_t\}$, иллюстрирующие "наличие временной памяти" волатильности для трёх рассмотренных выше периодов индекса S&P500:

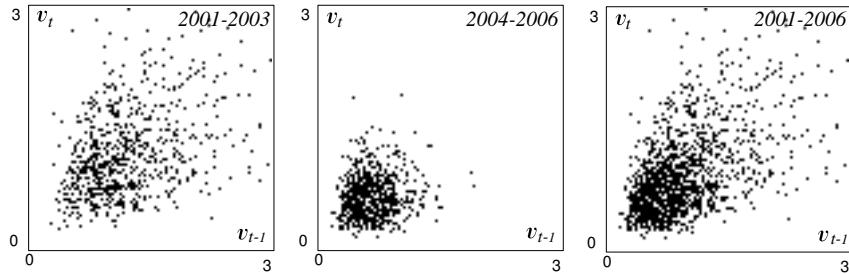


Рис.8: Зависимости $a_t(a_{t-1})$, S&P500

Как видно из диаграммы, точки заполняют область с характерной формой "венника", расширяющегося в область положительных значений. Естественно, он тем более ярко выражен, чем выше автокорреляционные коэффициенты.

Форма области $\sigma_t = f(\sigma_{t-s})$ практически не зависит от величины сдвига и способа измерения волатильности. Для валютной пары EURUSD на интервале 2004-2008 мы имеем следующие точечные диаграммы внутридневных волатильностей, полученных по 15-минутным лагам:

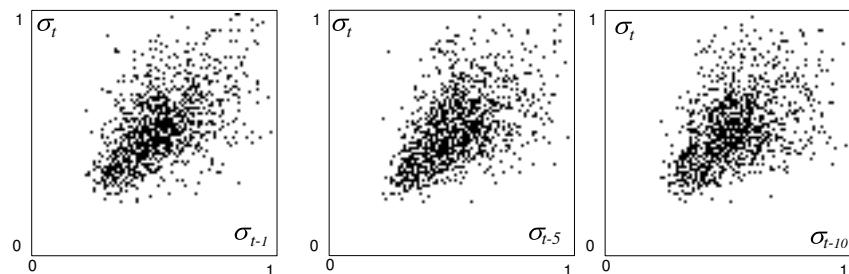


Рис.9: Зависимости $\sigma_t(\sigma_{t-s})$, $s = 1, 5, 10$, EURUSD

На этих рисунках в качестве сдвига взят один день ($s = 1$), неделя ($s = 5$), и две недели ($s = 10$). Видно, что форма "венника" заметным образом не изменяется, постепенно расплываясь с понижением автокорреляционного коэффициента.

5 Когда автокорреляции не затухают

Медленное снижение значений автокорреляционного коэффициента с увеличением параметра сдвига, на самом деле, должно настораживать. Существуют очень простые модели, не связанные со стохастичностью волатильности, в которых возникает подобный эффект.

Рассмотрим, например, обычное логарифмическое блуждание:

$$\frac{dx}{x} = \mu dt + \sigma \delta W. \quad (6)$$

Смоделируем "двадцатилетнюю" (5000=20·250 торговых дней) эволюцию цены, при которой в первые 10 лет волатильность была постоянной и равной $\sigma_1 = 1\%$, а во второе десятилетие она скачком повысилась до $\sigma_2 = 2\%$. При этом винеровская переменная представляется в виде $\delta W = \varepsilon \sqrt{dt}$, где ε - нормально распределенное случайное число с нулевым средним и единичной дисперсией. В качестве малого интервала времени dt выбрана одна секунда $dt = 1/(24 \cdot 60 \cdot 60)$, см. приложение C.

Динамика ежедневных значений модифицированной амплитуды размаха $v_t = a_t - |r_t|/2$ на "переломных" 10-м и 11-ом годах имела вид (время в "днях"):

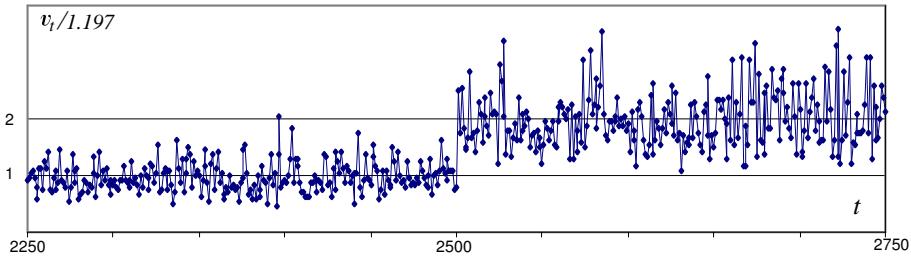


Рис.10: Два года блуждания на "перескоке" волатильности

Подобный ряд с нестационарностью в виде ступеньки обладает заметными автокорреляционными коэффициентами для модуля доходности (второй рисунок) и ещё большими - для амплитуды размаха (третий рисунок):

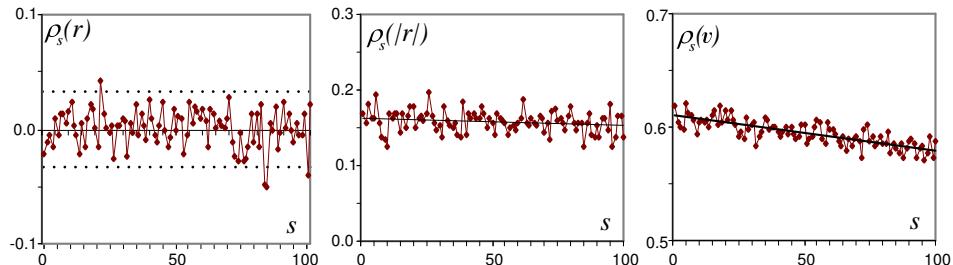


Рис.11: Автокорреляции доходности, её модуля и модифицированной амплитуды.

Характерно, что они очень медленно убывают с ростом параметра сдвига s . В противоположность $|r_t|$ и v_t , корреляции доходностей цены r_t (первый рисунок ниже), в пределах двух стандартных ошибок, равны нулю.

Таким образом, несмотря на статистическую независимость двух соседних дней, возникают корреляционные закономерности. Обратим внимание, что независимыми являются не только доходности r , но и их модули $|r|$, или амплитуды размаха v . Если бы волатильность была постоянной двадцать лет, то все коррелограммы $\rho_s(|r|)$ и $\rho_s(v)$ были бы равны нулю. При появлении нестационарности ситуация изменяется.

Причину подобного эффекта несложно понять. Ниже на трёх точечных диаграммах представлены значения логарифмических доходностей, их модулей и модифицированных амплитуд двух соседних дней в течение первого десятилетия эволюции с постоянной волатильностью $\sigma = 1\%$.

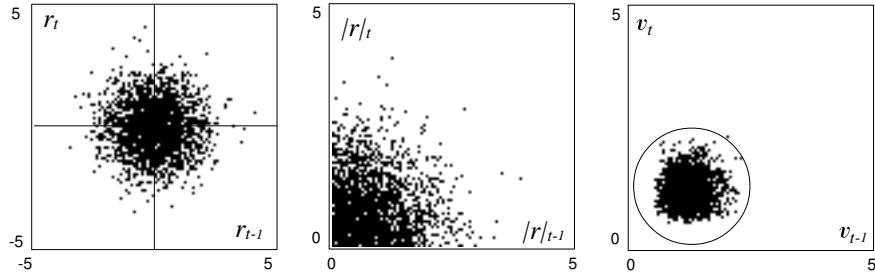


Рис.12: Первое десятилетие

В первом случае точки образуют практически симметричное облако, и, естественно, корреляция равна нулю. Во втором и третьем - полной симметрии нет, так как этим свойством не обладают плотности вероятности $P(|r|)$ и $P(v)$. Однако, в силу независимости последовательных дней, корреляционный коэффициент равен нулю. Если, например, $x = v_t$, а $y = v_{t-1}$, то независимость означает, что совместная плотность вероятности равна произведению плотностей вероятности каждой величины $P(x, y) = P(x) \cdot P(y)$. Поэтому при любом распределении ковариация будет равна нулю: $(x - \bar{x})(y - \bar{y}) = 0$.

Важным является тот факт, что для доходностей r_t центр облака расположен в начале координат, тогда как для положительно определённых величин $|r_t|$ и v_t он сдвинут вправо и вверх в область положительных значений.

Добавим теперь на диаграммы точки второго десятилетия:

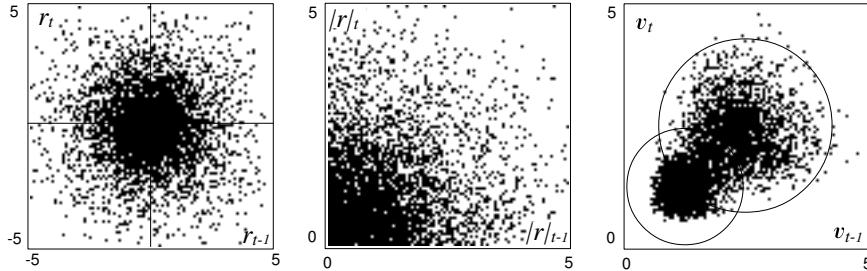


Рис.13: Оба десятилетия

Для логарифмических доходностей (первая диаграмма) происходит наложение двух облаков с одинаковым центром $\bar{r} = 0$. Результирующее облако остаётся симметричным, поэтому автокорреляция по-прежнему равна нулю. В случае амплитуд размаха (третья диаграмма) возникают два несовпадающих облака, одно из которых соответствует $\sigma_1 = 1\%$, а второе $\sigma_2 = 2\%$ (напомним, что $\bar{v} = 1.197\sigma$). Перемычка между облаками размывается, и в результате получается фигура с характерной формой веника (верхнее облако имеет больший размер). Через неё можно, методом наименьших квадратов, провести прямую, наклон которой и будет пропорционален корреляционному коэффициенту.

Диаграмма не изменится, если мы возьмём сдвиг в два дня $\{v_t, v_{t-2}\}$, так как, за исключением переходных точек в момент скачка волатильности, значения каждого десятилетия будут кластеризоваться в своём облаке.

Несколько сложнее ситуация обстоит со второй точечной диаграммой для модулей доходности $\{|r_{t-1}|, |r_t|\}$. Визуально она не отличается от аналогичной для первого десятилетия. Корреляция, тем не менее, возникает. Чтобы понять её происхождение, необходимо расширить стандартные статистические соотношения на случай нестационарных данных.

6 Нестационарная статистика

Пусть параметры распределения случайной величины x плавно изменяются со временем. Если мы, *не учитывая этого*, производим усреднение наблюдений за некоторый интервал времени T , то среднее значение \bar{x} по всем данным фактически будет равно:

$$\bar{x} = \langle \bar{x}(t) \rangle = \frac{1}{T} \int_0^T \bar{x}(t) dt, \quad \text{где} \quad \bar{x}(t) = \int_{-\infty}^{\infty} x \cdot P(x, t) dx. \quad (7)$$

Другими словами, мы в каждый фиксированный момент времени вычисляем *локальное среднее* $\bar{x}(t)$, а затем равномерно усредняем все такие средние по интервалу времени T (угловые скобки). Аналогично определим *локальную дисперсию*:

$$\sigma^2(t) = \int_{-\infty}^{\infty} (x - \bar{x}(t))^2 \cdot P(x, t) dx = \overline{x^2}(t) - \bar{x}(t)^2. \quad (8)$$

Дисперсия, вычисленная по всем данным, будет равна:

$$\sigma^2 = \overline{(x - \bar{x})^2} = \langle \overline{x^2}(t) \rangle - \langle \bar{x}(t) \rangle^2 = \langle \sigma^2(t) \rangle + \langle \overline{x^2}(t) \rangle - \langle \bar{x}(t) \rangle^2, \quad (9)$$

где угловые скобки, как и выше, обозначают усреднение по интервалу времени T . Таким образом, σ^2 определяется суммой взвешенной локальной дисперсии $\langle \sigma^2(t) \rangle$ и временной дисперсии среднего (второе и третье слагаемые).

В частном случае *параметрической нестационарности* величину x_t можно представить в виде $x_t = \mu(t) + \sigma(t) \cdot \eta_t$, где η_t - стационарные независимые случайные числа с нулевым средним и единичной дисперсией ($\bar{\eta} = 0$, $\eta^2 = 1$). Среднее этих величин равно $\langle \mu(t) \rangle$, а дисперсия: $\langle \sigma(t)^2 \rangle + \langle \mu(t)^2 \rangle - \langle \mu(t) \rangle^2$.

Рассмотрим теперь две *локально* независимые величины x и y . Их независимость означает, что плотность совместной вероятности в фиксированный момент времени t расщепляется $P(x, y, t) = P(x, t) \cdot P(y, t)$, и

$$\overline{x \cdot y}(t) = \bar{x}(t) \cdot \bar{y}(t). \quad (10)$$

Однако при усреднении по всем данным эти величины уже перестают быть независимыми. Действительно, временное среднее произведения:

$$\overline{x \cdot y} = \frac{1}{T} \int_0^T \int_{-\infty}^{\infty} x \cdot y P(x, y, t) dx dy dt = \frac{1}{T} \int_0^T \bar{x}(t) \bar{y}(t) dt = \langle \bar{x}(t) \cdot \bar{y}(t) \rangle, \quad (11)$$

вообще говоря, не равно произведению временных средних: $\bar{x} \cdot \bar{y} = \langle \bar{x}(t) \rangle \cdot \langle \bar{y}(t) \rangle$. Если локальные средние $\bar{x}(t)$, $\bar{y}(t)$ отличны от нуля, то, в общем случае, будет не нулевым и корреляционный коэффициент. В примере предыдущего раздела среднее значение доходностей в каждом десятилетии было равно нулю, поэтому их автокорреляция не возникала. Для положительных величин $|r|$ и a средние значения не нулевые, и, несмотря на независимость двух последовательных дней, возникает автокорреляция.

Таким образом, локально независимые величины, обладая однотипной долгосрочной нестационарностью, неизбежно становятся зависимыми. Однако эта зависимость имеет не стохастическую, а "детерминированную", гладкую природу, связанную времененной синхронизацией.

Например, если нестационарность волатильности имеет форму ступеньки с одинаковой длительностью, среднее значение и дисперсия по всем данным равны:

$$\bar{x} = \frac{\bar{x}_1 + \bar{x}_2}{2}, \quad \sigma^2 = \frac{\sigma_{x1}^2 + \sigma_{x2}^2}{2} + \frac{(\bar{x}_1 - \bar{x}_2)^2}{4}, \quad (12)$$

где x обозначает " $|r|$ " или " v ", статистические параметры первого десятилетия равны \bar{x}_1, σ_{x1} , а второго \bar{x}_2, σ_{x2} . Если сдвиг при вычислении автокорреляционного коэффициента мал по сравнению с длительностью T , в первом приближении краевыми эффектами можно пренебречь и считать, что $x = v_t$ и $y = v_{t-1}$ - две независимые величины. Их ковариация равна:

$$\bar{x} \cdot \bar{y} - \bar{x} \cdot \bar{y} = \frac{\bar{x}_1 \bar{y}_1 + \bar{x}_2 \bar{y}_2}{2} - \frac{\bar{x}_1 + \bar{x}_2}{2} \cdot \frac{\bar{y}_1 + \bar{y}_2}{2} = \frac{(\bar{x}_1 - \bar{x}_2)(\bar{y}_1 - \bar{y}_2)}{4}. \quad (13)$$

Так как $\bar{x}_i = \bar{y}_i$, то для автокорреляционного коэффициента получаем:

$$cor(x, y) = \frac{(\bar{x}_1 - \bar{x}_2)^2}{(\bar{x}_1 - \bar{x}_2)^2 + 2(\sigma_{x1}^2 + \sigma_{x2}^2)}. \quad (14)$$

Явно видно, что такая "корреляция" для нестационарных данных возникает только у величин с неравными средними. Для v и $|r|$ их средние и сигмы пропорциональны волатильности логарифмической доходности $\bar{x} = \alpha\sigma, \sigma_x = \beta\sigma$. Для модуля доходности $\alpha = 0.795, \beta = 0.605$, а для модифицированной амплитуды размаха цены $\alpha = 1.197, \beta = 0.300$. Если волатильность рынка изменяется, то изменяется и средние значения положительно определённых величин v и $|r|$. При удвоении волатильности логарифмической доходности возникает корреляция $1/(1 + 10(\beta/\alpha)^2)$. Для модуля доходности она равна 0.15, а для модифицированной амплитуды размаха 0.61. Именно это мы и наблюдали в численном эксперименте.

В общем случае для получения автокорреляционных коэффициентов, зависящих от сдвига s , необходимо воспользоваться их определением в виде сумм. Однако более компактным будет непрерывный предел. Будем считать T и сдвиг s величинами непрерывными. В случае n лагов длительностью τ каждый $T = n\tau$, а $s = k\tau$, причём $k \ll n$. Пусть некоторая положительно определённая величина σ_t , связанная с волатильностью, промодулирована нестационарной составляющей $\sigma_t = \sigma(t) \cdot \theta_t$, где θ_t - стационарное случайное число с единичным средним. Например, для модифицированной амплитуды $\theta_t = v_t \sqrt{2\pi}/3$. Так как случайные числа θ_t различных моментов времени нескоррелированы, но положительно определены, то $\theta_t \cdot \bar{\theta}_{t-s}$ равно единице при $s \neq 0$, и $\bar{\theta}^2$ при $s = 0$. Определим ковариационный коэффициент для случая $s \neq 0$ следующим образом:

$$\gamma_s(\sigma) = \langle \sigma_t \cdot \sigma_{t-s} \rangle - \langle \sigma_t \rangle^2 = \frac{1}{T-s} \int_s^T \sigma(t) \sigma(t-s) dt - \left[\frac{1}{T} \int_0^T \sigma(t) dt \right]^2. \quad (15)$$

Заметим, что это не единственная возможность в случае конечной выборки длительностью T . В любом случае мы требуем, чтобы ковариация (15) была равна нулю, если $\sigma(t) = const$. Дисперсия положительной величины σ_t равна $\gamma_0 = \bar{\theta}^2 \langle \sigma^2(t) \rangle - \langle \sigma(t) \rangle^2$. Соответственно, автокорреляционный коэффициент $\rho_s = \gamma_s/\gamma_0$ позволяет найти зависимость от параметра сдвига в различных ситуациях нестационарности (см. приложение D).

Таким образом, автокорреляции между мерами волатильности могут возникать по причине достаточно гладкой нестационарности, а не в силу стохастической природы волатильности. Возникает естественный вопрос: не обязаны ли наблюдаемые на различных финансовых рынках заметные автокорреляции волатильности подобному механизму?

7 Корреляция разностей

Простейший способ устранения относительно гладких нестационарностей во временных рядах - это переход к разностям величин. Если v_t испытывает локально постоянный снос, то это будет приводить к появлению автокорреляций. Разность двух последовательных величин подобный снос устраниет. Даже, если тренд данных v_t медленно изменяет своё направление, то в рамках восходящих и нисходящих участков значения разностей меняются незначительно и становятся локально квазистационарными.

Рассмотрим изменение модифицированной амплитуды размаха цены:

$$\delta v_t = v_t - v_{t-1}. \quad (16)$$

В качестве данных возьмём ежедневную статистику по фондовому индексу S&P500 за период 1990-2008 (4791 торговый день) и курса EURUSD (1999-2008, 2495 дня, исключая праздники). Построим сначала автокорреляционные коэффициенты амплитуд ежедневного размаха цены $\rho_s(v) = \text{cor}(v_t, v_{t-s})$:

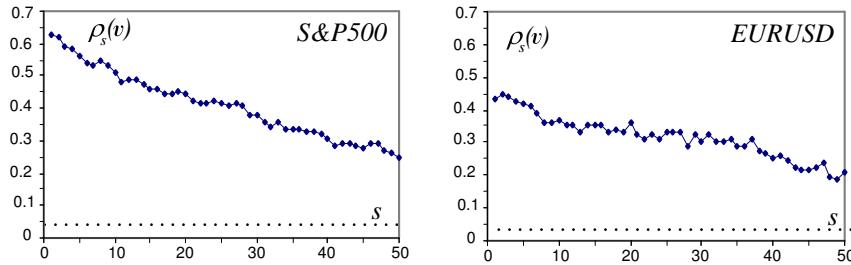


Рис.14: S&P500 и EURUSD коррелограммы амплитуд размаха цены

Как обычно, коэффициенты ρ_s достаточно высокие, однако автокорреляции для индекса S&P500 более значительны, чем для курса EURUSD, и имеют меньшие колебания.

Перейдём теперь к разностям амплитуд двух соседних дней. Их автокорреляция $\rho_s(\delta v)$ тут же резко падает:

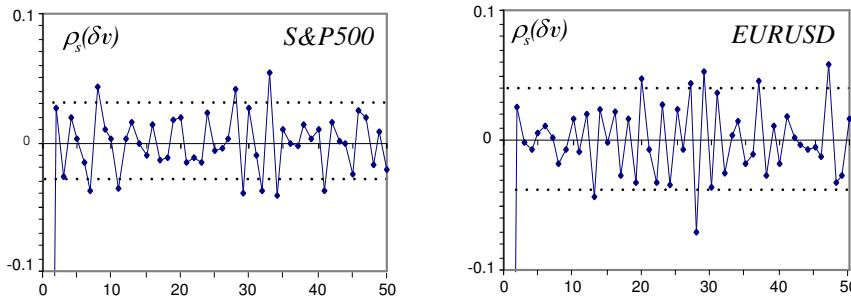


Рис.15: S&P500 и EURUSD коррелограммы после перехода к разностям

Отличия разительны. Второй автокорреляционный коэффициент в случае индекса S&P500 падает в 24 раза, со значения 0.618 до величины 0.026. Для курса EURUSD снижение составляет 17 раз - с 0.449 до 0.027

Пунктирные линии на всех рисунках означают двойную стандартную ошибку, равную $0.03 = 2/\sqrt{4791}$ для индекса S&P500 и $0.04 = 2/\sqrt{2495}$ для EURUSD.

Наглядно исчезновение корреляционной зависимости можно продемонстрировать на точечных диаграммах связи последовательных значений v_t и v_{t-s} .

Ниже на точечных диаграммах явно видны "корреляции" между $\{v_t, v_{t-1}\}$ индекса S&P500 и их отсутствие для $\{\delta v_t, \delta v_{t-2}\}$:

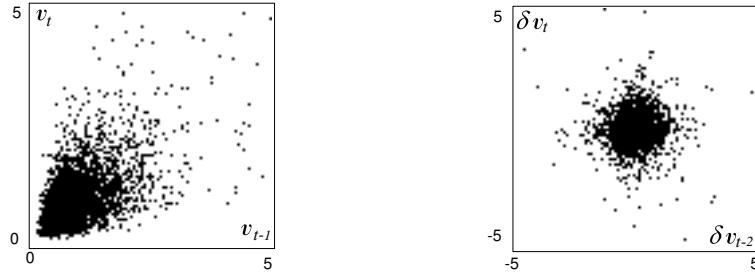


Рис.16: Точечные диаграммы для S&P500 до и после перехода к разностям

На рисунке слева точки заполняют область с характерной формой веника, тогда как справа мы имеем симметричное облако с нулевой корреляцией. Похожие результаты обнуления автокорреляционных коэффициентов получаются также для модулей логарифмической доходности $|r_t|$, и других финансовых инструментов.

Заметим, правда, что для разностей δv_t существует высокая отрицательная автокорреляция со сдвигом в один день $\rho_1(\delta v) = \text{cor}(\delta v_t, \delta v_{t-1})$. В примере выше она равна -0.49 для S&P500 и -0.53 для EURUSD. Однако её происхождение связано не со стохастической динамикой волатильности, а с эффектом перекрытия. Поясним это на следующем примере. Предположим, что справедлива простейшая модель:

$$v_t = \sigma \cdot \theta_t, \quad (17)$$

где $\sigma = \text{const}$, а θ_t - независимые стационарные положительные случайные числа, возникающие по причине ошибок конечности выборки по которой измеряется волатильность. В этом случае изменения $\delta v_t = \sigma \cdot (\theta_t - \theta_{t-1})$ имеют нулевое среднее $\bar{\delta v}_t = 0$. Первый автоковариационный коэффициент равен:

$$\langle \delta v_t \cdot \delta v_{t-1} \rangle = \sigma^2 \langle (\theta_t - \theta_{t-1}) \cdot (\theta_{t-1} - \theta_{t-2}) \rangle = -\sigma^2 \cdot [\bar{\theta^2} - \bar{\theta}^2] = -\sigma^2 \cdot \sigma_\theta^2, \quad (18)$$

где σ_θ^2 - дисперсия случайных величин θ . Среднее квадрата возникает в слагаемом $-\langle \theta_{t-1} \cdot \theta_{t-1} \rangle = \bar{\theta^2}$, которое и ответственно за эффект перекрытия. Аналогичным образом получается дисперсия разности $\langle \delta v_t^2 \rangle = 2\sigma^2 \sigma_\theta^2$. Поэтому первый автокорреляционный коэффициент в точности равен $\rho_1(\delta v) = -0.5$, что и наблюдается выше. Корреляции со сдвигами $s > 1$ будут нулевыми, так как перекрытия уже не возникает.

Тот факт, что для автокорреляций разностей с хорошей степенью точности выполняются соотношения $\rho_1(\delta v) = -0.5$ и $\rho_s(\delta v) = 0$ при $s > 1$ свидетельствует в пользу модели (17). Однако, если бы параметр σ был константой, то не возникло бы корреляций между последовательными значениями волатильности $\rho_s(v) = 0$ (в силу независимости θ_t). Она может возникать, как мы показали выше, в результате *плавного* изменения величины σ со временем. Таким образом, фактически $\sigma = \sigma(t)$, и является гладкой функцией времени.

Как для окончательного прояснения ситуации с $\rho_1(\delta v)$, так и для целей дальнейших исследований нам необходима методология выделения гладкой нестационарной составляющей волатильности.

8 Выделение гладкой нестационарности

Для выделения медленно меняющейся составляющей во временном ряду $x_k = x(t_k)$ мы будем использовать фильтр Ходрика-Прескотта (далее НР-фильтр). Гладкая составляющая s_k ряда находится в результате минимизации квадратов её отклонений от эмпирических данных x_k , одновременно с требованием минимальности кривизны s_k :

$$\sum_{k=1}^n (x_k - s_k)^2 + \lambda \cdot \sum_{k=2}^{n-1} (\nabla^2 s_k)^2 = \min, \quad (19)$$

где вторая производная в разностях равна $\nabla^2 s_k = (s_{k+1} - s_k) - (s_k - s_{k-1})$. Степень гладкости s_k будет тем выше, чем больше параметр λ . Значения λ варьируются в очень широком диапазоне, поэтому мы будем приводить его десятичный логарифм ν , представляя $\lambda = 10^\nu$.

При сглаживании сильно зашумлённых данных всегда присутствует прозвол в выборе параметра λ . Если λ мал, существует опасность обнаружить нестационарность там, где её нет. При слабом сглаживании гладкая составляющая будет повторять любые локальные флуктуации, не имеющие к нестационарности никакого отношения. С другой стороны, при сильном сглаживании мы рискуем упустить важные детали интересующей нас динамики.

Поэтому нам необходим некоторый статистический критерий степени сглаживания для уменьшения возможного произвола. Как обычно, будем в качестве эталона использовать модель случайного блуждания.

Среднее значение логарифмической доходности равно относительному изменению цены внутри временного лага $r_t = \ln C_t / O_t$. Волатильность будем восстанавливать по слаженному среднему модифицированной амплитуды размаха цены внутри лага $\sigma(t) = (a - |r|/2) \cdot \sqrt{2\pi}/3$. Говоря о волатильности, всегда подразумеваем волатильность лага (минутную, часовую, дневную и т.д.).

Если число дискретных блужданий цены внутри лага достаточно велико, то, независимо от их распределения, логарифмические доходности r_t будут нескоррелированными *гауссовыми* случайными числами. Сгладим их среднее значение $\bar{r}(t)$ при помощи НР-фильтра с различными параметрами λ и вычислим типичную величину $Err[\bar{r}(t)]$ колебаний $\bar{r}(t)$ вокруг среднего \bar{r} по всем эмпирическим точкам:

$$Err[\bar{r}(t)] = \sqrt{\langle (\bar{r}(t) - \bar{r})^2 \rangle}. \quad (20)$$

Аналогично определяется ошибка вычисления слаженной волатильности лага. Численные эксперименты показывают, что эти ошибки, с хорошей степенью точности, убывают с ростом параметра λ следующим образом:

$$Err[\bar{r}(t)] \approx \frac{0.50 \sigma}{\lambda^{1/8}}, \quad Err[\sigma(t)] \approx \frac{0.15 \sigma}{\lambda^{1/8}}, \quad (21)$$

и практически не зависят от числа эмпирических точек n . Более того, ошибки не зависят и от типа распределения (в случае дискретной модели лагового блуждания). Малая степень $1/8$ объясняет причину необходимости использования широкого диапазона для изменения параметра λ .

Соотношения (21) задают типичный коридор колебаний слаженных величин $\bar{r}(t)$ и $\sigma(t)$, которые являются флуктуациями и статистически не значимы в случае постоянства волатильности. Поэтому они будут нашими ориентирами, по крайней мере, на горизонтальных участках $\sigma(t)$.

Приведём типичный пример численного моделирования ($\sigma = 1$, $n = 1000$) для трёх значений λ ($\nu = \log_{10} \lambda$):

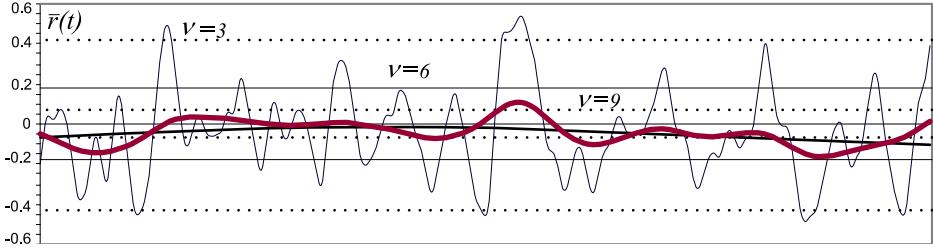


Рис.17: Сглаженное среднее гауссова шума

Более жирная линия соответствует $\lambda = 1000000$ ($\nu = 6$), а тонкая - $\lambda = 1000$ ($\nu = 3$). Сплошные горизонтальные "уровни значимости" определяют двойную ошибку $\pm 2\text{Err}[\bar{r}(t)]$ в случае $\nu = 6$, а пунктирные - для $\nu = 3$ и $\nu = 9$. В отличие от уровней значимости корреляционных коэффициентов, мы имеем гладкую величину $\bar{r}(t)$, которая может некоторое время "живь" вне заданного ошибкой диапазона. Тем не менее, соотношения (21) характеризуют значение типичных колебаний сглаженной величины для случайных данных.

Однако в ситуации нестационарности, которая нас, собственно, и интересует, необходимо выдерживать баланс между гладкостью и отсутствием излишнего сглаживания. Так, если $\sigma(t) = 1 + 0.5 \cdot \sin(2\pi t/T)$, где T - общая длительность эксперимента, получаем следующие варианты сглаживания волатильности, оцененной по модифицированной амплитуде размаха:

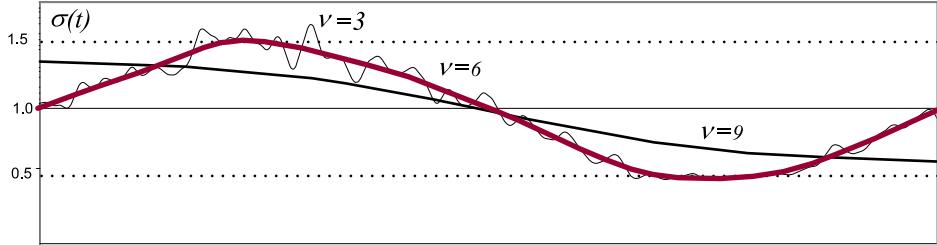


Рис.18: Сглаженная волатильность блуждания с $\sigma(t) = 1 + 0.5 \cdot \sin(2\pi t/T)$

В данном случае оптимальным значением была $\nu = 6$, так как $\nu = 3$ испытывает шумящие колебания вокруг истинной волатильности, а $\nu = 9$ - фактически не "ловит" синусоиду. Однако ситуация сильно ухудшается, если волатильность испытывает скачок. Так, пусть половину из $n = 1000$ "торговых дней" волатильность была $\sigma = 1\%$, а вторую половину $\sigma = 2\%$. Тогда сглаживания с различными λ дают такие результаты:

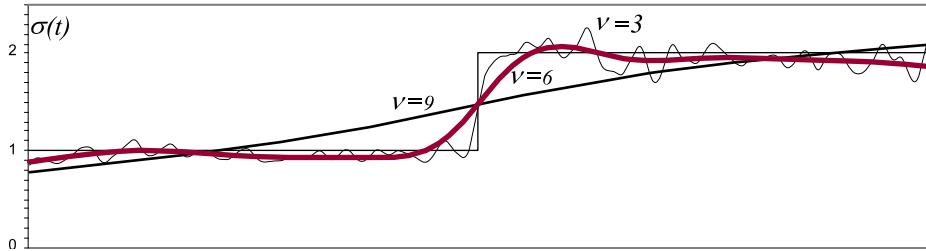


Рис.19: Сглаженная волатильность блуждания $\sigma(t)$ ступеньки

Видно, что в этом случае $\nu = 6$ существенно размывает ступеньку. Сглаживание с $\nu = 3$ размывает скачок волатильности существенно меньше, но зато даёт шумящие и незначимые колебания при постоянстве σ .

9 Автокорреляция остатков

Воспользуемся НР-фильтром для выделения гладкой нестационарной составляющей волатильности и устранения её из данных. Нас будут интересовать остаточные после такого выделения значения волатильности и их автокорреляционные коэффициенты.

Рассмотрим ежедневные модифицированные амплитуды размаха цены $v_t = a_t - |r_t|/2$ для курса EURUSD за период 1999-2008. С их помощью оценим ежедневную волатильность $\sigma_t = v_t \sqrt{2\pi}/3$.

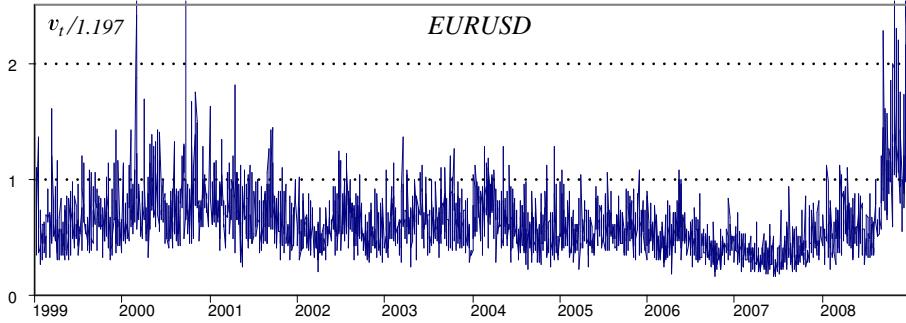


Рис.20: Волатильность EURUSD по амплитудам размаха

Выделим нестационарность при помощи НР-фильтра. Жирная линия на рисунке ниже представляет собой сглаженную волатильность с $\lambda = 1000000$ ($\nu = 6$). Двойная ошибка, в соответствии с формулой (21), при значении волатильности 0.5 (среднее за 2004-2007 г.г.) равна ± 0.026 . Фактически, это лишь немногим более толщины линии. Поэтому изгибы нестационарной волатильности $\sigma(t)$ при $\nu = 6$ можно считать статистически значимыми:

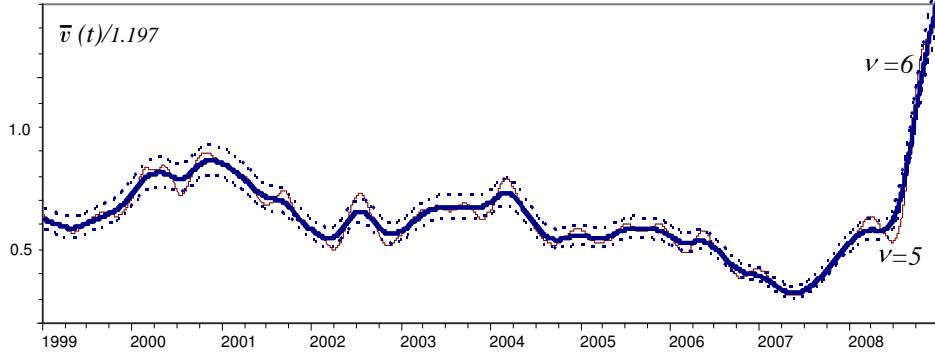


Рис.21: Сглаживание волатильности НР-фильтром

Иная ситуация при сглаживании с параметром $\nu = 5$. Будем отталкиваться от кривой $\sigma(t)$ для $\nu = 6$. Построим вокруг неё коридор двойной ошибки $(1 \pm 0.036) \cdot \sigma(t)$ (пунктирные линии) соответствующий значимости для сглаживания параметром $\nu = 5$. Видно, что сглаженная волатильность при $\nu = 5$ (тонкая линия) изгибается внутри этого коридора. Поэтому эти изгибы, по всей видимости, не являются статистически значимыми. Однако начало осеннего "перелома" 2008-го года $\nu = 5$, по-видимому, ухватывает заметно лучше.

Как понятно из предыдущего раздела, НР-фильтр выдерживает минимальную и примерно постоянную кривизну всей кривой. Поэтому он даёт хорошие результаты на относительно гладких участках и вносит определённые искажения на переломах.

Устраним теперь из данных гладкий тренд $\sigma(t)$. Для этого мы будем не вычитать его, как принято при обработке временных рядов, а делить на него:

$$\sigma_t \rightarrow \frac{\sigma_t}{\sigma(t)}. \quad (22)$$

Смысл подобной процедуры понятен. Происходит переход к нормированной волатильности для всего ряда. В результате волатильности выравнивают не только своё среднее положение равное 1, но и дисперсию:

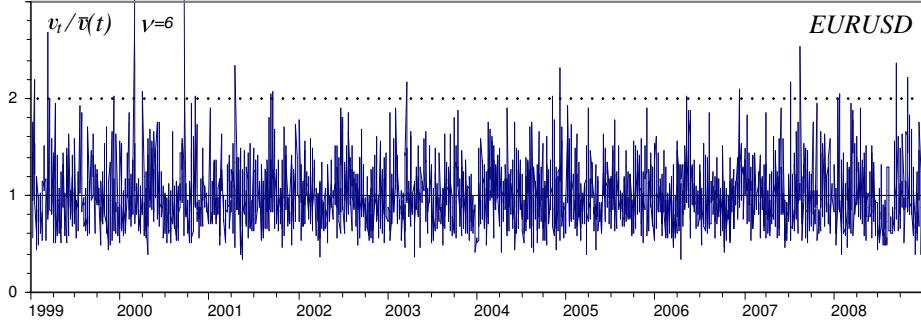


Рис.22: Амплитуды после нормировки

Сравним автокорреляционные коэффициенты до процедуры нормализации (22) (первый рисунок), и после неё (второй и третий):

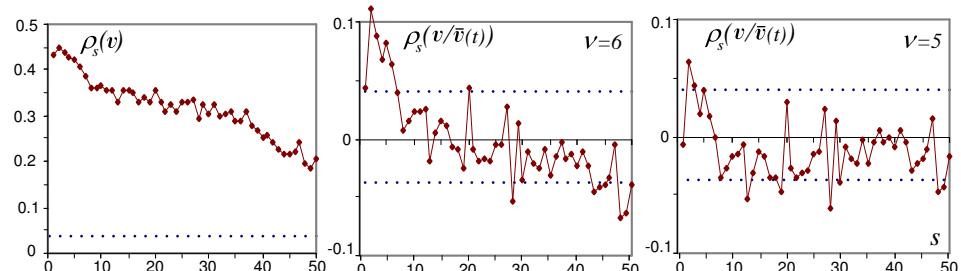


Рис.23: Автокорреляции до и после нормировки

Как мы видим, автокорреляции падают почти в 10 раз. Это же относится и к первому корреляционному коэффициенту, который в случае разностей амплитуд имел значение -0.50. Следовательно, его происхождение действительно было связано с эффектом перекрытия.

Заметим, что при процедуре нормализации мы делим все дневные амплитуды на сглаженную величину $\sigma(t)$. Однако при её вычислении используется множество значений σ_t в окрестности текущего времени t . В результате соседние значения $\sigma(t)$ будут существенно скоррелированы. Это может приводить к возникновению небольших автокорреляций после нормирования. Тем не менее, $\rho_s(v/v(t))$ очень малы.

Таким образом, как простой переход к разностям, так и устранение гладкой составляющей волатильности при помощи НР-фильтра делает корреляционные коэффициенты, фактически, статистически незначимыми. Простое объяснение причин появления автокорреляции в условиях нестационарности, в совокупности с этим, заставляет усомниться в стохастической природе волатильности. Однако причину её шумящей составляющей необходимо дополнительно исследовать. Мы вернёмся к этому вопросу в последнем разделе.

10 Назад к нормальному распределению

Как уже упоминалось во введении, существует большое число исследований, которые изучают распределение вероятности логарифмической доходности. Общепринятым стал факт его ненормальности. Однако, когда мы говорим о плотности вероятности как функции только доходности $P(r)$, мы, очевидно, подразумеваем стационарность случайных чисел r , так как не вводим зависимости от времени. Чтобы получить более или менее надёжные статистические результаты, при построении $P(r)$ выбирается как можно более широкий интервал времени, содержащий большое число данных n .

В условиях нестационарности подобные действия заметно искажают "истинный" вид распределения. Если статистические параметры зависят от времени, то и плотность распределения также не будет стационарной $P(r, t)$. Будем считать, что нестационарность является *параметрической* и сосредоточена только в волатильности $\sigma(t)$. Пусть для $P(r, t) = P(r, \sigma(t))$ справедливо распределение Гаусса ($r_t = \sigma(t) \cdot \varepsilon_t$):

$$P(r, t) = \frac{1}{\sigma(t)\sqrt{2\pi}} e^{-\frac{1}{2} \frac{r^2}{\sigma^2(t)}}. \quad (23)$$

Второй и четвёртый моменты равны: $\overline{r^2} = \langle \sigma^2(t) \rangle$, $\overline{r^4} = 3 \langle \sigma^4(t) \rangle$, и в общем случае, несмотря на гауссовость распределения, его "интегральный" эксцесс без учёта нестационарности становится отличным от нуля:

$$ex = 3 \cdot \left[\frac{\langle \sigma^4(t) \rangle}{\langle \sigma^2(t) \rangle^2} - 1 \right]. \quad (24)$$

В модельном примере 20-летнего блуждания с удвоением волатильности экспесс объединённых данных равен $27/25=1.08$. В более общем случае на негауссовость могут влиять и другие виды нестационарности, например, снос доходности: $r_t = \mu(t) + \sigma(t) \cdot \varepsilon_t$.

Посмотрим, что произойдёт с эмпирическими данными после устранения нестационарности. Для этого разделим все r_t на величину волатильности в данный момент времени. Её текущее значение будем получать при помощи сглаживания НР-фильтром дневных модифицированных амплитуд размаха $\sigma_t = (a_t - |r_t|/2)\sqrt{2\pi}/3$. Таким образом, мы проводим следующее преобразование исходных логарифмических доходностей:

$$r_t \rightarrow r'_t = \frac{r_t}{\sigma_t}. \quad (25)$$

Подобная нормализация делает промодулированные функцией $\sigma(t)$ случайные числа r'_t стационарными.

Ниже в таблице приведены статистические параметры логарифмической доходности индекса S&P500 за период 1990-2008. Общее число торговых дней равно $n = 4791$, доля положительных значений во всех случаях 52.8%.

	aver	sigma	asym	excess	p_1
r	0.020	1.137	-0.23	10.18	78.9
$\nu = 6$	0.051	1.199	-0.15	1.19	71.1
$\nu = 5$	0.055	1.187	-0.12	0.82	70.4
$\nu = 4$	0.059	1.177	-0.08	0.51	69.5

Первая строка соответствует статистикам до преобразования нормирования (25). Затем идут статистики после преобразования, в котором использовалось сглаживание с различным параметром $\nu = \log_{10} \lambda$.

Стоит обратить внимание на колонки *excess* и p_1 . Видно, что сглаживание радикально уменьшает значение этих параметров. Это происходит даже для достаточно плавной функции $\sigma(t)$, соответствующей $\nu = 6$. Ниже на рисунке она представлена более толстой линией:

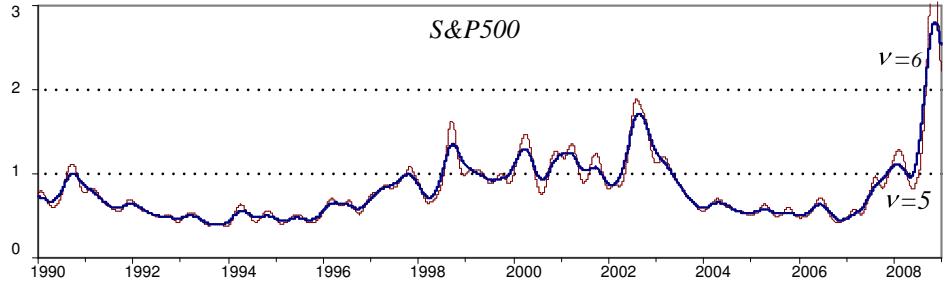


Рис.24: Гладкая волатильность S&P500 для различных параметров ν

Отметим также тот факт, что волатильность после нормирования несколько больше единицы. Это свидетельствует о том, что для фондового рынка необходимо использовать поправочный множитель в отношении ν/σ и фактически $\sigma \approx a - |r|/2$ без коэффициента $\sqrt{2\pi}/3$.

Чем меньше параметр ν , тем сильнее изгибы волатильности $\sigma(t)$, так как на усреднении начинают сказываться флуктуации доходности. Понятно, что в этом случае происходит уменьшение эксцесса даже для стационарных негауссовых случайных чисел. Чтобы контролировать этот эффект, проведём следующий эксперимент. Исходные пары ежедневной доходности и волатильности $\{r_t, \sigma_t\}$ произвольным образом перемешаем для устранения нестационарности. После этого будем применять сглаживание НР-фильтром и нормализацию (25) как к исходным данным (*original*), так и к перемешанным (*mixed*). Ниже на рисунках представлена зависимость эксцесса (слева) и вероятности попадания доходности в одну сигму p_1 (справа), как функций параметра сглаживания $\nu = \log_{10} \lambda$.

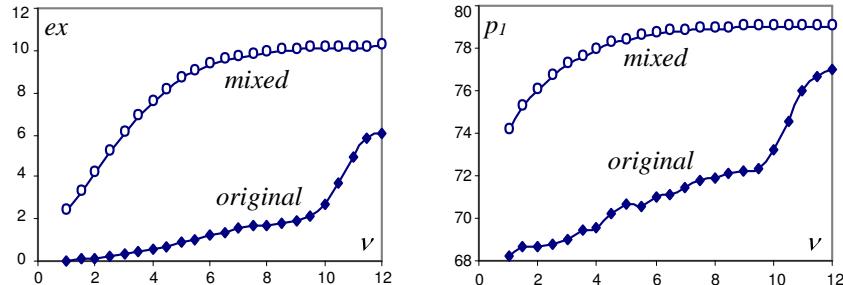


Рис.25: Эксцесс *ex* и вероятность p_1 для разных ν , 1990-2008

Хорошо видно, что выше значений $\nu \sim 6$ эксцесс и вероятность p_1 для перемешанных данных убывают очень слабо. В то же время статистические параметры негауссности исходных данных стремительно падают.

В качестве некоторого критерия оптимального значения ν можно выбрать точку, в которой разница между статистикой по перемешанным и исходным данным достигает своего максимального значения.

Ещё одним аргументом существенного вклада эффектов нестационарности в негауссовость распределения является разворачивающийся в 2008-м году финансовый кризис. Как видно из таблицы, эксцесс за период 1990-2008 составил *ex* = 10.2. Однако, достаточно отбросить волатильный 2008-й год, и эксцесс уменьшится в три раза до значения *ex* = 3.8. Количество торговых дней, по которым он вычислялся, уменьшается при этом всего на 5% до $n = 4528$.

Ниже приведен график зависимости эксцесса и вероятности p_1 от параметра сглаживания ν для перемешанных и исходных данных ежедневных доходностей индекса S&P500 за период 1990-2007:

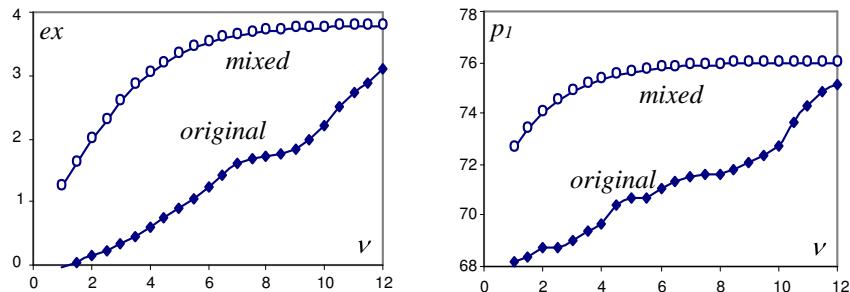


Рис.26: Эксцесс ex и вероятность p_1 для разных ν , S&P500, 1990-2007

Видно, что, хотя начальное значение эксцесса достаточно невелико, оно, тем не менее, статистически значимо падает при устранении в данных нестационарности. Можно считать значимым для нормализованных данных эксцесс порядка $ex = 1$, т.е. в 4-е раза меньше, чем у исходных.

Приведём гистограммы распределения плотности вероятности и граф нормальной вероятности [31], формально построенные по исходным нестационарным данным:

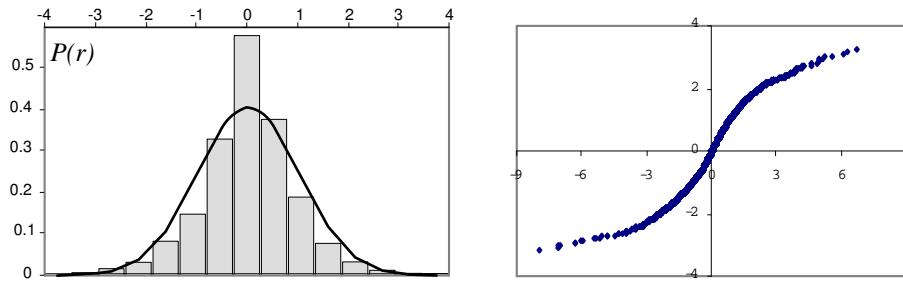


Рис.27: Распределение доходностей S&P500 1990-2008

и данным после процедуры их нормализации (25):

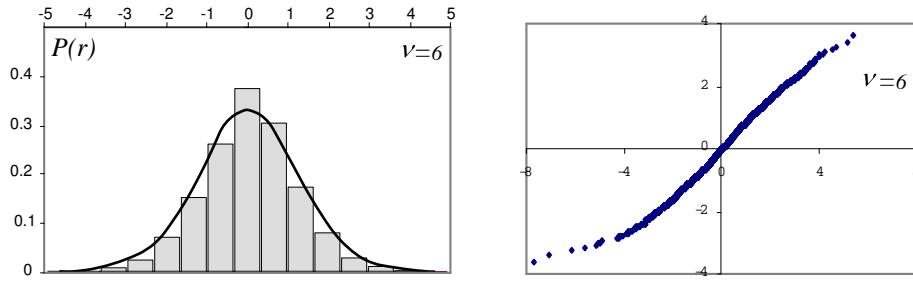


Рис.28: Распределение доходностей S&P500 1990-2008 после нормализации

Линия без маркеров на рисунках соответствует распределению Гаусса. Граф нормальной вероятности является зависимостью $y = f(r)$ связи $F_N(y) = F(r)$, где $F_N(y)$ - интегральное нормальное распределение, а $F(r)$ - эмпирическое интегральное распределение для доходностей. Если эмпирическое распределение $F(r)$ является гауссовым, то этот график должен быть прямой линией. Видно, что после нормализации плотность вероятности становится существенно более нормальной. Отклонения от прямой на графике особенно заметны для больших отрицательных выбросов, что связано с редкими негативными шоковыми воздействиями на рынок.

Рассмотрим, для сравнения, распределение вероятности ежедневных доходностей валютного рынка на примере курса EURUSD за период 1999-2008. Основные статистические параметры до нормирования (первая строка) и после сглаживания с различными параметрами ν приведены в таблице ниже:

	<i>aver</i>	<i>sigma</i>	<i>asym</i>	<i>excess</i>	p_1
r	0.008	0.652	0.05	1.3	72.7
$\nu = 10$	0.017	1.022	0.03	0.8	71.5
$\nu = 6$	0.022	0.995	0.00	0.1	69.3
$\nu = 4$	0.022	0.993	0.01	0.1	69.0

Видно, что у исходных данных сравнительно небольшой эксцесс, однако после сглаживания он уменьшается ещё сильнее. Среднее значение волатильности после нормировки близко к единице. Это говорит о том, что $\sigma = \nu\sqrt{2\pi}/3$ является хорошей несмешённой оценкой ежедневной волатильности доходности курса.

Тест на статистическую значимость падения эксцесса и вероятности p_1 свидетельствует о близком к нулю эксцессе нормированных доходностей:

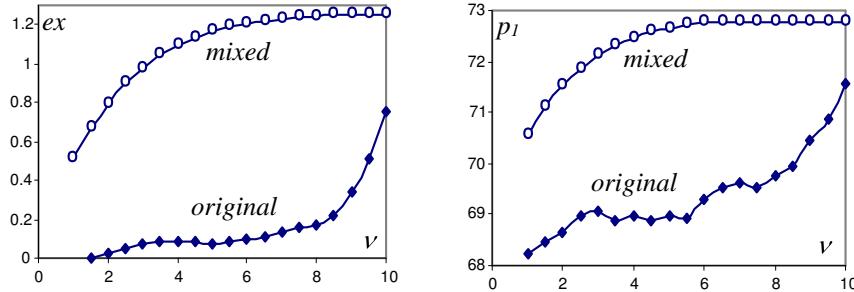


Рис.29: Эксцесс ex и вероятность p_1 для разных ν , EURUSD, 1999-2008

Соответствующая гистограмма и график нормальной вероятности имеют вид:

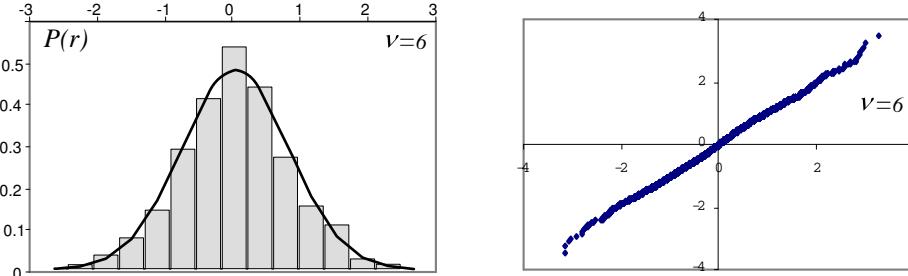


Рис.30: Распределение доходностей EURUSD, 1999-2008 после нормализации

В результате получается практически каноническое нормальное распределение с отклонениями, достаточно типичными для сравнительно небольшой выборки ($n = 2495$).

Мы не будем проводить более детальный статистический анализ формы распределения, ограничившись этими иллюстративными примерами. Предполагается (см. Заключение), что наблюдаемые данные являются смесью нормально распределённых флюктуаций рынка, промодулированных нестационарной волатильностью, и редких шоковых воздействий. Поэтому даже после устранения нестационарности данных в них могут оставаться шоковые выбросы, которые делают суммарное распределение слабо-негауссовым.

11 Квазистационарность волатильности

Нулевые автокорреляционные коэффициенты между последовательными значениями волатильности, вообще говоря, не исключают возможности её стохастического описания. В частности, мы можем записать следующую простейшую дискретную модель:

$$r_t = \sigma_t \cdot \nu_t, \quad \sigma_t = \sigma \cdot (1 + \beta \cdot \mu_t), \quad (26)$$

где ν_i и μ_i - независимые случайные числа, а σ, β - константы. Однако в таком виде интерпретация волатильности σ_t как случайной величины достаточно бессмысленна. Фактически, мы возвращаемся к обычной стационарной модели $r_t = \sigma \varepsilon_t$, где $\varepsilon_t = \nu_t + \beta \cdot \mu_t \nu_t$. В частности, если ν_i и μ_i имеют нормальное распределение, то распределение для ε_i будет уже не нормальным с эксцессом, равным $6\beta^2(2 + \beta^2)/(1 + \beta^2)^2$. Тем не менее, вопрос о *локальном* постоянстве "истинной" волатильности остаётся открытым.

Проведём некоторые статистические оценки. Рассмотрим сперва модифицированную амплитуду вероятности. Разброс её значений при постоянной волатильности σ происходит по причине конечной ширины плотности распределения $P(v)$. Её аналитический вид можно получить из уравнения (31) приложения А в виде следующего ряда:

$$P(v) = (32v^4 - 9)\mathbf{N}(2v) + \sum_{k=2}^{\infty} \left\{ \frac{4(2k-1)^2}{k^2(k-1)^2} \mathbf{N}_1 - \frac{8k^2(1+k^2 - 4(k^4 - k^2)v^2)}{(k^2-1)^2} \mathbf{N}_2 \right\},$$

где $\mathbf{N}_1 = \mathbf{N}(2(2k-1)v)$, $\mathbf{N}_2 = \mathbf{N}(2kv)$ - нормированные гауссовые функции (см. приложение А). Ниже приведены интегральные вероятности попадания величины $\sigma = v\sqrt{2\pi}/3$ в интервал $[0..v_0]$ (первая строка - v_0 , вторая - вероятность в процентах):

0.5	0.6	0.7	0.8	0.9	1.0	1.1	1.2	1.3	1.4	1.5	1.6
0.6	3.3	10.4	22.5	37.8	53.7	68.1	79.5	87.7	93.1	96.4	98.2

В интервал $[0 .. 1.5]$ модифицированная амплитуда размаха $v\sqrt{2\pi}/3$ должна попадать около 96.4% дней. Ниже 0.5 она может опускаться крайне редко.

Если мы устраним (25) при помощи сглаживания ($\nu = 4$) нестационарность в ежедневных модифицированных амплитудах размаха v для EURUSD за 2007-2008 год, получится следующая динамика:

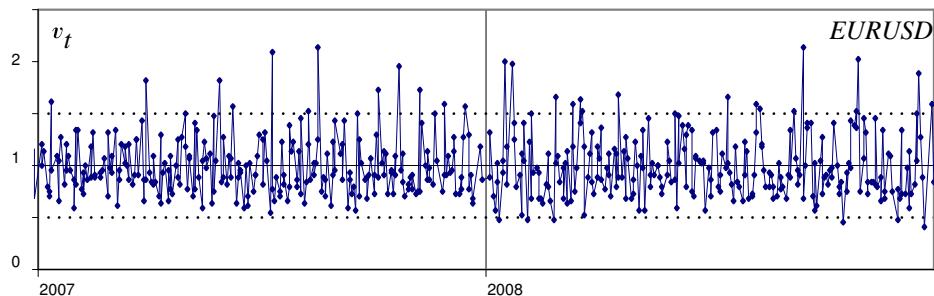


Рис.31: Стационарная амплитуда EURUSD в 2007-2008

Пунктирные линии в случае броуновского блуждания соответствуют вероятности 96% попадания в интервал $0.5 < v\sqrt{2\pi}/3 < 1.5$. Видно, что за исключением достаточно редких выбросов большинство ежедневных волатильностей, оцененных по модифицированной амплитуде вероятности, попадают в пунктирный коридор. Выбросов из него несколько больше, чем 4% (в году около 250 торговых дней, $250*4\% = 10$). Незначительное превышение можно интерпретировать, как редкие шоковые воздействия на рынок, не связанные с "типичной" для него динамикой, особенно в кризисный 2008-й год.

Как мы уже обсуждали выше, "ежедневную" волатильность можно получить, не только при помощи модифицированных амплитуд размаха, но и вычисляя её значения на основе внутридневных лагов для, например, пятнадцати - минутных точек. Ниже приведена динамика волатильности, полученная таким способом за период 2007-2008 для курса EURUSD:

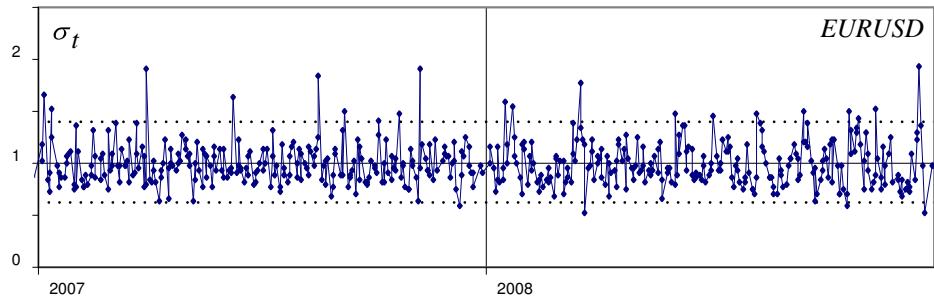


Рис.32: Внутридневная волатильность EURUSD в 2007-2008

В данном случае разброс значений связан с конечностью выборки, по которой вычисляется волатильность ($n = 96 = 4 \cdot 24$). Чтобы провести уровни значимости, необходимо знание соответствующего распределения вероятности. Как известно, ошибка вычисления стационарной волатильности определяется моментами четвёртого порядка и, в случае большого эксцесса, она будет достаточно велика.

Внутридневные 15-минутные данные обладают заметным эксцессом. При формальном его вычислении для EURUSD 2004-2008 получается значение порядка 20. Такой высокий эксцесс обусловлен долгосрочной нестационарностью, существенными циклическими эффектами активности внутри дня, а также рядом других специфических причин, на которых мы сейчас останавливаться не будем.

Для получения оценки уровней значимости проведём следующий простой эксперимент с данными. Вычислим логарифмические доходности на основе 15 - минутных лагов курса EURUSD. Затем, для сохранения внутридневной периодичности, перемешаем доходности с одинаковыми значениями времени суток. Другим словами, случайным образом переставляем все лаги в 00:00, отдельно от них перемешиваем лаги в 00:15, и т.д. По этим синтетическим данным, потерявшим любую память, кроме внутридневных циклов, вычислим значения внутридневной волатильности. Отнормируем их таким образом, чтобы среднее было равно единице, и построим соответствующее распределение вероятности. Оказывается, что в коридор 1 ± 0.4 попадет около 96% данных. Именно эти уровни, характеризующие "типичный" разброс волатильности за счёт конечности данных, проведены на рисунке выше пунктирными линиями. Данные достаточно хорошо укладываются в этот коридор.

Подчеркнем, что проделанные выше вычисления являются скорее оценкой, чем строгим статистическим анализом, который, возможно, и неуместен без построения полной модели нестационарности данных на различных временных масштабах. Однако выглядит достаточно правдоподобным предположение о локальном постоянстве волатильности для доходности дневных лагов. Другими словами, дневная волатильность рынка, по всей видимости, является гладкой, достаточно медленно изменяющейся функцией времени. В данный момент её значение можно считать постоянным и определяющим стохастическую динамику доходности цен финансового инструмента.

12 Заключение

Основной тезис, который мы аргументировали в настоящей работе, состоит в следующем:

Волатильность и другие статистические параметры стоит рассматривать как плавно изменяющиеся функции времени. Они определяют локально - квазистационарную стохастическую динамику цен финансовых инструментов. Относительно *редко и нерегулярно* возникают точечные шоковые воздействия на рынки, которые изменяют ежедневные значения доходностей, и, накапливаясь, могут воз действовать на долгосрочные значения волатильности.

Ситуация напоминает деформацию пластичного материала после серии ударов по нему и постепенное восстановление формы при их отсутствии. Исключительный интерес представляет исследование свойств подобной пластичности волатильности, особенно в целях предсказания времени её возврата на долгосрочные уровни.

Таким образом, стохастичность рынков состоит из двух составляющих - 1) естественной гауссово распределённой ежедневной доходности с медленно изменяющей волатильностью и 2) шоковых воздействий, происходящих достаточно редко. Шоки предполагаются принципиально непредсказуемыми, однако их воздействие на волатильность и её последующая эволюция должны быть предметом исследований.

Вообще говоря, шоки, до некоторой степени, являются "вещью в себе". В реальности достаточно сложно отделить "неестественное" поведение рынка в результате шокового воздействия от его "обычной" волатильности. Аналитики и экономические обозреватели ежедневно находят новостное объяснение любым скачкам цен. С другой стороны, такие события, как банкротство Lehman Brothers, вряд ли являются рядовыми новостями.

Волатильность может постепенно повышаться и в результате относительно незначимого негативного новостного фона, если он длится достаточно долго. Так, постепенный рост волатильности с начала 2007 года был результатом именно такого "мягкого" давления на рынки со стороны сектора недвижимости. С осени 2008 года этот рост стал взрывным и беспрецедентным для новейшей истории финансовых рынков. Произошло это, как известно в финансовом секторе, что спровоцировало лавинообразный эффект кризиса доверия и панических настроений. Это, в свою очередь, немедленно сказалось на реальном сегменте экономики.

Наконец, рост волатильности обычно сопровождает "немотивированный" ажиотаж на рынке в процессе возникновения того или иного финансового пузыря. Естественно, высокая волатильность сохраняется и в период его сущности. В фазе "спокойного" роста волатильность обычно медленно снижается.

Пики на диаграммах динамики нестационарной волатильности наводят на аналогию с резонансными явлениями в физических системах. Это, в свою очередь, подразумевает наличие некоторых уравнений, описывающих подобную динамику. Не вызывает также сомнения факт существования механизма релаксации и определённого времени жизни резонанса, после которого происходит затухание возбуждения системы.

Говоря о плавном ходе изменения волатильности надо иметь ввиду, что это относится к "типовым" долгосрочным ситуациям на рынке. Иногда, тем не менее, могут происходить бифуркационные, скачкообразные перестройки статистических параметров определяющих стохастическую динамику цен. По-видимому с такой бифуркационной перестройкой мы имели дело в сентябре 2008 года. Выход из этой бифуркации скорее всего будет достаточно плавный и затяжной.

Негауссовый характер рынков определяется двумя причинами. Во-первых, он является артефактом некритичного постулирования стационарности в условиях, когда её на самом деле нет. Эта составляющая, в принципе, устранима. После приведения данных в стационарный вид они существенно снижают свою негауссовость. Тем не менее, редкие шоковые воздействия, смешанные со стационарными гауссовыми доходностями, делают распределение слабо-негауссовым. Особенно это проявляется в случае фондового рынка, который обладает внебиржевым периодом накопления негативных или позитивных новостей. Когда рынки открываются, появляется возможность "разового выплескивания" накопившихся эмоций. Круглосуточно работающие валютные рынки могут "плавнее" реагировать на появление подобных шоковых воздействий.

Автокорреляционные коэффициенты между различными мерами, характеризующими волатильность, также возникают по причине нестационарности данных и исчезают после её устранения. В этом смысле они, естественно, являются свидетельством долгосрочной памяти, но не имеют ни чего общего с краткосрочными стохастическими свойствами волатильности, которые предполагаются в соответствующих моделях. По всей видимости, дальнейшие исследования должны сконцентрироваться на прогнозировании гладкой динамики волатильности, а не на её стохастических теориях.

A Приложение: Броуновское блуждание.

Приведём основные соотношения теории броуновского движения, описываемого стохастическим уравнением $dx = \mu dt + \sigma dW$. Рассмотрим сначала случай отсутствия сноса ($\mu = 0$). Без потери общности можно считать, что в начальный момент времени $x(0) = 0$. Максимальное и минимальное значения x за период $0 \leq t \leq T$ равны H и L , и $r = x(T)$. Высота подъёма $h = H$ и глубина опускания $l = -L$ всегда положительны, и $-l \leq r \leq h$. Амплитуда размаха равна $a = h + l$. Ниже рассматривается случай единичной волатильности $\sigma = 1$ и единичного интервала времени $T = 1$. Для их восстановления необходимо во всех "размерных" величинах r , h , l , a проделать замену $r \rightarrow r/\sigma\sqrt{T}$. Это же необходимо сделать и в дифференциалах dr и т.п. при интегрировании плотностей вероятностей. Для сокращения используется функция нормального распределения: $\mathbf{N}(x) = e^{-x^2/2}/\sqrt{2\pi}$.

- Исходным соотношением является вероятность того, что x не поднимется выше h и не опустится ниже $-l$, закрывшись на доходности r :

$$p(-l < L, H < h, r) = \sum_{k=-\infty}^{\infty} \{\mathbf{N}(r + 2ka) - \mathbf{N}(r + 2l + 2ka)\}. \quad (27)$$

Эту формулу получил Феллер в 1951 [20]. Отметим также исключительно полезный справочник [26]. Из вероятности Феллера (27) выводятся другие распределения. Так, для доходности, высоты и глубины имеем:

$$P(r) = \mathbf{N}(r), \quad P(h) = 2\mathbf{N}(h), \quad P(l) = 2\mathbf{N}(l). \quad (28)$$

Плотность вероятности размаха a выражается в виде бесконечного ряда по гауссовому базису:

$$P(a) = 8 \sum_{k=1}^{\infty} (-1)^{k+1} \cdot k^2 \cdot \mathbf{N}(ka). \quad (29)$$

Ряд достаточно быстро сходится для всех $a \neq 0$. Характерным свойством распределения Феллера $P(a)$ является экстремально быстрое снижение плотности вероятности при уменьшении a . Приведём некоторые значения интегральных вероятностей $F(a) = p(H - L < a)$:

a	0.750	1.000	1.500	2.000	2.500	3.000	3.500	4.000
$F(a)$	0.002	0.063	0.487	0.819	0.950	0.989	0.998	1.000

Ниже значения 0.75 ($\sigma = 1$) параметр a опускается только в 2-х случаях из 1000. Среднее значение $\bar{a} = 1.5958$, сигма $\sigma_a = 0.29798 \cdot \bar{a}$. В интервал одной сигмы $\bar{a} \pm \sigma_a = [1.120 .. 2.071]$ попадает 71.6% значений a . В двойную сигму $\bar{a} \pm 2\sigma_a = [0.645 .. 2.547]$ попадет 95.6% значений, причём выпадания из этого интервала, практически, должны встречаться только сверху.

Совместные плотности вероятности для высоты ($r \leq h$) глубины ($-l \leq r$) и амплитуды ($|r| \leq a$) имеют вид:

$$P(h, r) = 2(2h - r) \cdot \mathbf{N}(2h - r), \quad P(l, r) = 2(2l + r) \cdot \mathbf{N}(2l + r). \quad (30)$$

$$P(a, r) = 4 \sum_{k=-\infty}^{\infty} k \cdot \left\{ -|r| - k(2k+3)a + k \cdot (a - |r|)(2ka + |r|)^2 \right\} \cdot \mathbf{N}(|r| + 2ka). \quad (31)$$

Заметим, что $P(a, -r) = P(a, r)$, и $P(a, |r|) = 2P(a, r)$.

- Приведём таблицу средних значений различных величин:

$$\begin{aligned}\bar{r} &= 0, & \bar{r^2} &= 1, & \bar{r^3} &= 0, & \bar{r^4} &= 3, \\ \bar{h} &= \sqrt{\frac{2}{\pi}}, & \bar{h^2} &= 1, & \bar{h^3} &= \sqrt{\frac{8}{\pi}}, & \bar{h^4} &= 3, \\ \bar{a} &= \sqrt{\frac{8}{\pi}}, & \bar{a^2} &= 4 \ln 2, & \bar{a^3} &= \frac{(2\pi)^{3/2}}{3}, & \bar{a^4} &= 9 \cdot \zeta[3], \\ \bar{v} &= \frac{3}{\sqrt{2\pi}}, & \bar{v^2} &= 4 \ln 2 - \frac{5}{4}, & \bar{v^3} &= \frac{21 + \pi^2}{6\sqrt{2\pi}}, & \bar{v^4} &= 6 \ln 2 - \frac{27}{16} + \frac{3}{8} \cdot \zeta[3],\end{aligned}$$

где $\zeta[n] = \sum_{k=1}^{\infty} k^{-n}$ - функция Римана. Средние для l и $|r|$ эквивалентны h . Ниже даны значения некоторых смешанных произведений:

$$\begin{aligned}\bar{h} \bar{r} &= \frac{1}{2}, & \bar{h} \bar{r^2} &= \frac{4}{3} \sqrt{\frac{2}{\pi}}, & \bar{h} \bar{r^3} &= \frac{3}{2}, & \bar{h} \bar{r^4} &= \frac{24}{5} \sqrt{\frac{2}{\pi}}, \\ \bar{l} \bar{r^n} &= (-1)^n \cdot \bar{h} \bar{r^n}, & \bar{a} \bar{r^{2n+1}} &= 0, & \bar{a} \bar{r^{2n}} &= 2 \cdot \bar{h} \bar{r^{2n}}, & \bar{a} \bar{|r|} &= \frac{3}{2}.\end{aligned}$$

Другие средние, а также их производящую функцию можно найти в [22].

Для блуждания со сносом $dx = \mu dt + \sigma dW$ будем использовать уже определённые ранее плотности без сноса. Для того, чтобы восстановить время и волатильность, необходимо дополнительно сделать замену сноса $\mu \rightarrow \mu T/\sigma\sqrt{T}$. Плотность вероятности для доходности равна:

$$P_\mu(r) = \mathbf{N}(r - \mu) = e^{\mu r - \mu^2/2} P(r).$$

Выражения для совместных плотностей [26]:

$$\begin{aligned}P_\mu(h, r) &= e^{\mu r - \mu^2/2} P(h, r), & P_\mu(l, r) &= e^{\mu r - \mu^2/2} P(l, r), \\ P_\mu(a, r) &= e^{\mu r - \mu^2/2} P(a, r), & P_\mu(h, l, r) &= e^{\mu r - \mu^2/2} P(h, l, r).\end{aligned}$$

Таким образом, во всех случаях плотности соответствующие $\mu = 0$ умножаются на фактор $e^{\mu r - \mu^2/2}$. При наличии сноса:

$$\bar{r} = \mu, \quad \bar{r^2} = 1 + \mu^2, \quad \bar{r^3} = 3\mu + \mu^3, \quad \bar{r^4} = 3 + 6\mu^2 + \mu^4.$$

Выражения для средних значений других величин достаточно громоздки. Однако, так как для финансовых данных $\mu \ll \sigma = 1$, уместно разложить в ряд фактор $e^{\mu r - \mu^2/2}$ и использовать средние для случая $\mu = 0$. В результате:

$$\bar{h} = \sqrt{\frac{2}{\pi}} + \frac{\mu}{2} + \frac{\mu^2}{3\sqrt{2\pi}} - \frac{\mu^4}{60\sqrt{2\pi}} + \dots, \quad \bar{|r|} = \sqrt{\frac{2}{\pi}} + \frac{\mu^2}{\sqrt{2\pi}} - \frac{\mu^4}{12\sqrt{2\pi}} + \dots, \quad (32)$$

$$\bar{l} = \sqrt{\frac{2}{\pi}} - \frac{\mu}{2} + \frac{\mu^2}{3\sqrt{2\pi}} - \frac{\mu^4}{60\sqrt{2\pi}} + \dots, \quad \bar{a} = \sqrt{\frac{8}{\pi}} + \frac{2\mu^2}{3\sqrt{2\pi}} - \frac{\mu^4}{30\sqrt{2\pi}} + \dots \quad (33)$$

Средние значения высоты и глубины линейны по μ , и далее в разложении идут только чётные степени. Амплитуды размаха и модуля доходности зависят только от чётных степеней μ . Отметим также простые конечные соотношения: $\bar{h} - \bar{l} = \bar{r} = \mu$, $\bar{h^2} + \bar{l^2} = 2 + \mu^2$, $\bar{h} \bar{r} = \bar{h^2} - 1/2$, $\bar{l} \bar{r} = 1/2 - \bar{l^2}$.

В Приложение: Меры волатильности.

Для положительной случайной величины $z > 0$ ширину её распределения можно характеризовать относительной ошибкой σ_z/\bar{z} , где σ_z , как обычно, обозначает среднеквадратичное отклонение $\sigma_z^2 = \frac{1}{(z - \bar{z})^2}$.

Заметим, что относительные ширины распределений для z и z^2 различны, поэтому, вообще говоря, существуют различные критерии оптимальности мер измерения волатильности. Например, для вычисления стационарной волатильности обычно используется, усреднение либо квадратов доходностей, либо квадратов амплитуд размаха [21]:

$$\sigma_R^2 = \frac{1}{n-1} \sum_{t=1}^n (r_t - \bar{r})^2, \quad \sigma_P^2 = \frac{1}{n} \sum_{t=1}^n \frac{a_t^2}{4 \ln 2}. \quad (34)$$

Так как в этой статье мы изучаем нестационарный характер волатильности, и для её сглаживания используем нелинейный НР-фильтр, то удобнее усреднять непосредственно волатильности σ , а не их квадраты, которые, как мы увидим ниже, при малых n дают, к тому же, смещённое значение σ . Тем не менее, говоря о различных мерах волатильности, мы будем вычислять относительную ширину как самой величины, так и её квадрата.

Приведём некоторые известные меры. Базовой будет мера Parkinson (1980) [21], равная амплитуде размаха $v_P = a$. Garman and Klass (1980) [22], в классе аналитических функций по h, l, r , предложили следующую оптимальную комбинацию, которая лучше меры Parkinson:

$$v_{GK}^2 = 0.511 \cdot a^2 - 0.019(r \cdot (h-l) + 2h \cdot l) - 0.383 \cdot r^2. \quad (35)$$

Отметим также более простую и не зависящую от сноса μ меру Rogers and Satchell (1991) [23]:

$$v_{RS}^2 = h \cdot (h - r) + l \cdot (l + r). \quad (36)$$

Покажем, что простейшая линейная модификация меры Parkinson

$$v_\beta = a - \beta \cdot |r| \quad (37)$$

с некоторой константой $\beta > 0$ приводит к более узкому распределению, чем амплитуда размаха. Если в качестве критерия узости использовать относительную волатильность σ_v , то при помощи средних из приложения А несложно найти оптимальное значение коэффициента β :

$$\frac{(a - \beta \cdot |r|)^2}{(\bar{a} - \beta \cdot |r|)^2} = \min \quad \Rightarrow \quad \beta = 6 - 8 \ln 2 \approx 0.455. \quad (38)$$

Однако, так как критерий σ_v/\bar{v} не является единственным, и в силу слабой чувствительности относительной волатильности от β , мы в статье использовали значение $\beta = 1/2$ и обозначение $v = a - |r|/2$. Далее $v_\beta = a - 0.455 \cdot |r|$.

Заметим, что существует ещё одна простая мера волатильности, сравнимая по эффективности с (37), следующего вида:

$$v_F = \frac{a}{1 + r^2/a^2}. \quad (39)$$

Хотя вероятность нулевого значения a для конечной длительности лага T исчезающе мала, необходимо, тем не менее, доопределить $v_F = 0$ при $a = 0$. Вообще, (37) и (39) не являются аналитическими функциями по a и r и выпадают из действия леммы приложения В. работы [22].

- Кроме ширины распределения, в качестве критерия иногда используется отсутствие или слабая зависимость от сноса μ . Заметим, что для дневных или более коротких лагов $\mu \ll \sigma$. Поэтому этот критерий не является столь значимым. Предложенная выше мера модифицированной амплитуды размаха, как и сама амплитуда, зависит от μ . Однако эта зависимость существенно слабее, чем у амплитуды. Если воспользоваться разложениями (32), (33), для v_β можно записать:

$$\bar{v}_\beta = (2 - \beta) \cdot \sqrt{\frac{2}{\pi}} + \frac{(2 - 3\beta)\mu^2}{3\sqrt{2\pi}} - \frac{(2 - 5\beta)\mu^4}{60\sqrt{2\pi}} + \dots$$

Видно, что коэффициент при μ^2 в случае $\beta = 1/2$ в четыре раза меньше, чем в случае $\beta = 0$ ($v_P = a$). Соответственно, в четыре раза меньше и зависимость от μ . При $\beta = 2/3$ (далее $v_{2/3}$) коэффициент при μ^2 становится равным нулю и зависимость от μ ослабевает ещё сильнее, хотя полностью она исчезает только для меры Rogers and Satchell.

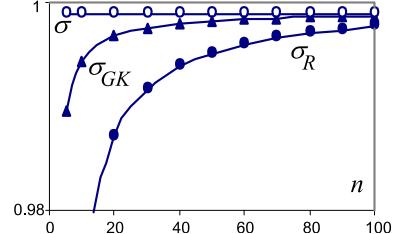
- Сравним теперь статистические параметры различных мер волатильности:

Мера	\bar{v}	\bar{v}^2	σ_v	as	ex	p_1	σ_v/\bar{v}	σ_{v^2}/\bar{v}^2
v_P	1.596	2.773	0.476	0.97	1.24	70.6	0.298	0.638
v_{RS}	0.960	0.998	0.275	0.46	0.42	69.5	0.287	0.576
$v_{2/3}$	1.064	1.217	0.292	0.52	0.29	68.4	0.275	0.557
v_{GK}	0.968	0.998	0.245	0.60	0.39	68.6	0.253	0.519
v_F	1.254	1.673	0.316	0.53	0.28	68.4	0.252	0.513
v	1.197	1.523	0.300	0.53	0.26	68.2	0.251	0.511
v_β	1.233	1.615	0.308	0.55	0.29	68.3	0.250	0.510

Курсивным шрифтом приведены значения Монте-Карло моделирования по 3.5 миллионам лагов, каждый из которых являлся блужданием из миллиона тиков. В этом случае для средних и волатильности в последнем знаке возможна ошибка порядка ± 0.002 . Для определения остальных значений (не курсивных) использовались аналитические выражения.

- В условиях нестационарности данных часто необходимо проводить усреднение по достаточно малому числу наблюдений n . В этом случае начинает проявляться смещённость квадратичных мер для оценки волатильности σ . Даже, если мы вычисляем классический квадрат волатильности σ_R^2 по несмещённой формуле (34), значение σ_R будет смещено, так как при усреднении по большому числу выборок, каждая из которых имеет размер n , мы имеем: $\langle \sigma_R^2 \rangle = \sigma^2$, но $\langle \sqrt{\sigma_R^2} \rangle \neq \sigma$. Если нас интересует значение именно волатильности, а не её квадрата, лучше для нестационарных данных использовать линейные, а не квадратичные меры.

Для иллюстрации эффекта смещения справа приведены графики средних значений волатильности, полученных усреднением по большому числу выборок по n значений в каждой для стандартного определения σ_R и меры $\sigma_{RG} = \sqrt{v_{RG}^2}$ (35) по сравнению с линейной мерой $\sigma = (a - |r|/2)\sqrt{2\pi}/3$.



Таким образом, мера $v = a - |r|/2$ имеет достаточно узкое распределение, и, следовательно, ошибку измерения волатильности. При этом её очевидным преимуществом является простота, по сравнению с мерами v_{RS} и v_{GK} . Кроме этого она не смещена при малых размерах выборки, что существенно при исследовании эффектов нестационарности.

C Приложение: Моделирование блуждания

При моделировании логарифмического блуждания $dx/x = \sigma\delta W$ минимальный временной шаг выбирался равным одной секунде $dt = 1/(60 \cdot 60 \cdot 24)$ (время измеряется в днях), а в качестве случайной переменной ε в винеровском слагаемом $\delta W = \varepsilon\sqrt{dt}$ взяты нормально распределённые случайные числа с единичной волатильностью и нулевым средним (функция RndG()):

```

int lag=60*60*24;      // один день, секундные тики
int num = 250*20;      // двадцать лет

Float dt=1.0/lag, sqrt_dt=sqrt(dt);

Float x=100;
for(int t=0; t<num; t++) {
    Float min=x, max=x, x0=x;
    for(int i=0; i<lag; i++) {                                // агрегирование
        Float sigma = (t<num/2) ? 0.01: 0.02;
        x += ( sigma*RndG() * sqrt_dt ) * x;
        if(x>max) max=x;
        if(x<min) min=x;
    }
    r = 100*log(x/x0);
    a = 100*log(max/min);
    printf("%d\t%g\t%g\t%g\n", t, r, a, fabs(r));
}

```

Понятно, что все рассмотренные в приложении А. соотношения справедливы в непрерывном пределе винеровского блуждания. Если внутри лага (как это обычно и бывает на практике) конечное число тиков, то результаты могут иногда заметно отличаться от асимптотических.

Приведём зависимость значений средних и их волатильностей различных величин в зависимости от количества тиков внутри лага (первая колонка lag):

lag	\bar{a}	σ_a	\bar{h}	σ_h	$ \bar{r} $	$\sigma_{ r }$
10	1.267	0.489	0.633	0.599	0.633	0.599
100	1.483	0.477	0.741	0.602	0.742	0.603
1000	1.560	0.476	0.779	0.603	0.780	0.603
10000	1.585	0.475	0.793	0.602	0.792	0.602
100000	1.591	0.476	0.795	0.603	0.796	0.603
1000000	1.595	0.475	0.797	0.603	0.798	0.603

Таким образом, например, среднее значение амплитуды размаха \bar{a} достаточно медленно сходится к значению непрерывного предела $\bar{a} = 1.696$.

Вообще, соотношению непрерывного и дискретного в теории и практике финансов уделяется недостаточно внимания. Ценовая динамика финансовых инструментов обладает, до определённого уровня, свойствами фрактальности. К тому же в непрерывном пределе, благодаря дифференциальному исчислению, заметно упрощается описание стохастических процессов. Однако на ультракоротких временах мы имеем существенно дискретные процессы взаимодействия экономических агентов и эффекты конечности изменения цены. Кроме этого, наличие спредов и транзакционных комиссий реально не позволяет делать займы и производить непрерывное перестраивание портфеля, как это обычно предполагается, например, в теории опционов. Иногда подобные ограничения дискретности могут оказывать существенное влияние на справедливость конечных выводов.

D Приложение: Автокорреляции

Рассмотрим примеры автокорреляционных коэффициентов в условиях нестационарности (15) в некоторых частных случаях.

- Для волатильности в виде ступеньки со значением σ_1 длительностью $f \cdot T$ и σ_2 в течение $(1-f) \cdot T$ для дисперсии имеем:

$$\gamma_0(\sigma) = \sigma_\theta^2 \cdot [f \sigma_1^2 + (1-f) \sigma_2^2] + f(1-f) \cdot (\sigma_2 - \sigma_1)^2, \quad (40)$$

где $\sigma_\theta^2 = \overline{(\theta - 1)^2}$ - дисперсия нормированной положительно определённой случайной величины с единичным средним. Ковариационный коэффициент $\gamma_s(\sigma) = \langle \sigma_t \cdot \sigma_{t-s} \rangle - \langle \sigma_t \rangle^2$ равен:

$$\gamma_s(\sigma) = f(1-f) \cdot (\sigma_1 - \sigma_2)^2 + (\sigma_1 - \sigma_2)(\sigma_2 f - \sigma_1(1-f)) \cdot \frac{s}{T-s} \quad (41)$$

При большом периоде усреднения T по сравнению со сдвигом s остаётся только первое слагаемое, которое мы получили в шестом разделе.

- В случае линейного роста волатильности $\sigma(t) = \sigma_0 \cdot (1 + \beta t/T)$, с постоянной скоростью β дисперсия γ_0 равна:

$$\frac{\gamma_0(\sigma)}{\sigma_0^2} = \sigma_\theta^2 \cdot \left(1 + \beta + \frac{\beta^2}{3}\right) + \frac{\beta^2}{12} \quad (42)$$

Автоковариационный коэффициент:

$$\frac{\gamma_s(\sigma)}{\sigma_0^2} = \frac{\beta^2}{12} \cdot \left(1 - 2 \frac{s}{T} - 2 \frac{s^2}{T^2}\right) \quad (43)$$

- Если волатильность испытывает периодические колебания $\sigma(t) = \sigma_0(1 + \beta \sin(2\pi m t/T))$, где $m = 1, 2, \dots$ - целые числа, то волатильность x в условиях такой нестационарности равна:

$$\frac{\gamma_0(\sigma)}{\sigma_0^2} = \sigma_\theta^2 \cdot \left(1 + \frac{\beta^2}{2}\right) + \frac{\beta^2}{2} \quad (44)$$

Ковариационные коэффициенты становятся периодическими функциями сдвига s :

$$\frac{\gamma_s(\sigma)}{\sigma_0^2} = \beta^2 \cdot \left\{ \frac{1}{2} \cos(2m\pi \frac{s}{T}) + \frac{T/4m\pi}{(T-s)} \sin(2m\pi \frac{s}{T}) \right\} \quad (45)$$

Ниже приведены эмпирические графики автокорреляционных коэффициентов всех этих трёх случаев вместе с теоретической кривой (тонкая линия):

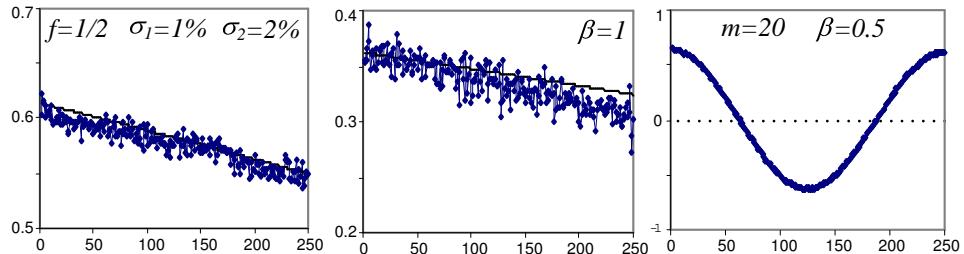


Рис.33: Коррелограммы трёх видов нестационарностей

Список литературы

- [1] G. W. Schwert, 1989, Why Does Stock Market Volatility Change Over Time? *The Journal of Finance*, Vol.44, No.5, pp. 1115-1153.
- [2] M. McAleer, M.C. Medeiros, 2008 Realized volatility: A review, *Econometric Reviews*, 27(1-3) pp.10-45.
- [3] T.G. Andersen, T. Bollerslev, 1998, Answering the skeptics: Yes, standard volatility models do provide accurate forecasts, *International Economic Review*, Vol. 39, No. 4, pp.885-905.
- [4] Stochastic volatility. Selected Readings, 2005, Edited by N. Shephard, Oxford.
- [5] Forecasting Volatility in the Financial Markets, 2007, Third edition, Edited by J.Knight, S. Satchell,
- [6] R.F. Engle, A.J. Patton, 2001, What good is volatility model?.
- [7] R.F. Engle, 1982, Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation, *Econometrica*, Vol.50, No.4, pp.987-1008.
- [8] R. Engle, 2002, New frontiers for ARCH models *Journal of Applied Econometrics*, Vol.15, pp.425-446.
- [9] S. Alizadeh, M.W. Brandt, F.X. Diebold, 2002, Range - based estimation of stochastic volatility models, *The Journal of Finance*, Vol.LVII, No.3. p.1047
- [10] Ser-Huang Poon, C.W.J. Granger, 2003, Forecasting Volatility in Financial Markets: A Review, *Journal of Economic Literature*, Vol.XLI, pp. 478-539.
- [11] T.A. Silvey, 2007 An investigation of the relative performance of GARCH models versus simple rules in forecasting volatility, in "Forecasting Volatility in the Financial Markets", Third edition.
- [12] George A. Christodoulakis and Stephen E. Satchell, 2007, 10 Hashing GARCH: a reassessment of volatility forecasting performance in "Forecasting Volatility in the Financial Markets", Third edition.
- [13] B. Mandelbrot, 1963 The variation of certain speculative prices, *The Journal of Business*, Vol.36, No.4, pp.394-419.
- [14] O.E. Barndorff-Nielsen, N.Shephard, 2000, Econometric analysis of realised volatility and its use in estimating stochastic volatility models,
- [15] B. Biais, L. Glosten, C. Spatt, 2005, Market microstructure: A survey of microfoundations, empirical results, and policy implications *Journal of Financial Markets*, No.8, pp.217-264.
- [16] A.Madhavan, 2000 Market microstructure: A survey *Journal of Financial Markets*, 3, pp.205-258.
- [17] F.M. Bandi, J.R. Russell, 2003, Separating microstructure noise from volatility
- [18] S. Figlewski, 1994, Forecasting volatility using historical data Working Paper Series.
- [19] T.G. Andersen, T. Bollersley, F.X. Diebold, P. Labys, 2000, The Distribution of Realized Exchange Rate Volatility *Penn Institute for Economic Research*.

- [20] W. Feller, 1950, The asymptotic distribution of the range of sums of independent random variables, *The Annals of Mathematical Statistics*, pp.427-432.
- [21] M. Parkinson, 1980 The extreme value method for estimating the variance of the rate of return, *The Journal of Business*, Vol.53, No.1.
- [22] M.B. Garman, M.J.Klass, 1980, On the estimation of security price volatilities from historical data, *The Journal of Business*, Vol.53, No.1.
- [23] L.C.G. Rogers, S.E.Satchell, 1991, Estimating variance from high, low and closing prices, *The Annals of Applied Probability*, Vol.1. No.4, pp.504-512.
- [24] L.C.G. Rogers, F.Zhou, 2008 Estimating correlation from high, low, opening and closing prices, *The Annals of Applied Probability*, Vol.18. No.2, pp.813-823.
- [25] D. Yang, Q. Zhang, 2000, Drift-independent volatility estimation based on high, low, open, and close prices, *The Journal of Business*, Vol.73, No.3, pp.477-491.
- [26] A.N. Borodin, P. Salminen, 2000 Handbook of Brownian Motion - Facts and Formulae, Basel: Birkhauser.
- [27] R. Cont, 2001, Empirical properties of asset returns: stylized facts and statistical issues, *Quantitative Finance* Vol.1, pp.223-236.
- [28] P.Boyle, A.Ananthanarayan, 1997, The impact of variance estimation in option valuation models,
- [29] T. Eraker, 2004, Do stock prices and volatility jump? Reconciling evidence from spot and option prices. *The Journal of Finance*, Vol.LIX, No.3 pp.1367-1402.
- [30] T. Mikosch, C. Starica, 2004, Nonstationarities in financial time series, the long-range dependence, and the IGARCH effects. *The Review of Economics and Statistics*, 86(1) pp.378-390.
- [31] E.F. Fama, 1965, The behavior of stock-market prices, *The Journal of Business*, Vol.38, No.1, pp.34-105

(c) 2009, Altus.ua