

# Динамика условных моментов высоких порядков и прогнозирование стоимостной меры риска\*

Григорий Франгуриди†

*Российская экономическая школа, Москва, Россия*

В настоящей работе исследуется возможность улучшения качества прогнозов стоимостной меры риска VaR с помощью моделирования динамики скошенности и эксцесса условных распределений финансовых доходностей. Для однодневных прогнозов VaR, вычисленных для пяти ликвидных акций индекса S&P500 из различных секторов, мы сравниваем стандартную модель GARCH со скошенным студентовским распределением ошибок с набором моделей с динамическими моментами высоких порядков, таких как модели типа ARCD с нормальным обратным гауссовским и обобщённым скошенным студентовским распределением ошибок. В качестве тестов на предсказательную способность используются как скоринговые тесты для левых хвостов предсказанных плотностей, так и LR-тесты корректности прогнозов VaR. Мы предлагаем модель со скошенным обобщённым распределением ошибок с динамическими скошенностью и куртозисом, которая даёт по крайней мере не менее качественные прогнозы VaR, чем аналогичные модели, и в то же время является вычислительно менее сложной.

*Ключевые слова: стоимостная мера риска, условное распределение, скошенность, куртозис, финансовые доходности*

*Классификация JEL: C22, C53, C58*

## 1 Введение

Стоимостная мера риска value-at-risk (VaR) является в настоящее время одним из самых распространённых показателей риска в финансовой индустрии. По определению, VaR доверительной вероятности  $p$  (близкой к 1) портфеля ценных бумаг  $\Pi_{t+1}$ , оцениваемый на день  $t+1$ , есть квантиль уровня  $1-p$  распределения логарифмической доходности  $r_{t+1}^{\Pi} = \log(\Pi_{t+1}/\Pi_t)$ . Таким образом, VaR можно понимать как значение, меньше которого завтрашняя доходность не упадёт с вероятностью  $p$ . VaR имеет несколько важных преимуществ над другими мерами риска. Во-первых, совокупные риски портфеля представляются всего одним числом с прозрачной интерпретацией. Во-вторых, VaR можно применять к портфелям, состоящим из ценных бумаг любого типа, будь то обыкновенные акции или экзотические опционы, что позволяет сравнивать риски портфелей разной структуры. В-третьих, расчёт VaR реализован почти во всех пакетах для финансовых вычислений, что позволяет использовать его финансистам, не владеющим программированием. Наконец, расчёт VaR является обязательным требованием Базельского комитета по банковскому надзору в рамках вычислений показателей достаточности капитала коммерческих банков. Так как во многих развитых странах банковские регуляторы требуют выполнения Базельских соглашений, подотчётные банки не могут обойтись без расчёта VaR. Более того, многие банки ежедневно рассчитывают VaR портфелей кредитов и ценных бумаг для внутренней отчётности по контролю за рисками.

---

\*Цитировать как: Франгуриди, Григорий (2014) «Динамика условных моментов высоких порядков и прогнозирование стоимостной меры риска», Квантиль, №12, стр. 69–82. Citation: Franguridi, Grigory (2014) “Higher order conditional moment dynamics and forecasting value-at-risk,” Quantile, No.12, pp. 69–82.

†Адрес: 117418, г. Москва, Нахимовский проспект, 47, офис 1721(5). Электронная почта: [gfranguridi@nes.ru](mailto:gfranguridi@nes.ru)

Однако нельзя не отметить и некоторые недостатки VaR. Очевидно, что сведение рисков портфеля к одному числу сильно *упрощает* истинную структуру рисков. VaR может также сильно *искажать* представление о «настоящих» рисках: например, два портфеля могут иметь одинаковый VaR, даже если при условии пробы VaR по одному из этих портфелей потери будут заведомо больше, чем по другому. Иными словами, VaR не может отразить разницу в левых хвостах распределений доходностей, находящихся за  $(1 - p)$ -квантилем. Более того, VaR не всегда поощряет диверсификацию портфеля, т.е. не является *субаддитивным*: не для всех случайных величин (доходностей)  $r_1$  и  $r_2$  верно  $\text{VaR}(r_1 + r_2) \leq \text{VaR}(r_1) + \text{VaR}(r_2)$  (см. контрпример в книге Föllmer & Schied, 2004).<sup>1</sup>

В банковской индустрии обычно интересуются *условным* дневным VaR, т.е. квантилем завтрашнего распределения финансовой доходности с учётом наблюдаемой истории доходностей. Как правило, для прогнозирования условного VaR специфицируется и оценивается параметрическая модель доходностей, после чего прогноз условного VaR вычисляется как квантиль предсказанной условной плотности. Одними из наиболее распространённых моделей здесь являются разновидности GARCH-моделей с распределениями ошибок, характеризующимися тяжёлыми хвостами и скошенностью. В настоящей работе мы показываем, что учёт *динамики* условных моментов высоких порядков позволяет улучшить качество прогнозирования VaR (вне обучающей выборки) по сравнению с моделями с фиксированными скошенностью и куртозисом как в смысле поведения процесса пробоев VaR, так и в смысле корректности предсказания левых хвостов условных распределений доходностей. Мы также строим новую модель SGE-ACD со скошенным обобщённым распределением ошибок Theodossiou (2000) и авторегрессионной динамикой на (параметры, отвечающие за) условные моменты второго, третьего и четвёртого порядков, которая по качеству прогнозов VaR и левых хвостов условных распределений неотличима от более вычислительно сложных моделей с динамикой высоких моментов, таких как SGT-GARCH с обобщённым скошенным распределением Стюдента из работы Bali, Mo & Tang (2008) и NIG-ACD с нормальным обратным гауссовским распределением из работы Wilhelmsson (2009).

## 2 Моделирование стоимостной меры риска

### 2.1 Модели с постоянными условными моментами высоких порядков

Пусть дана выборка доходностей  $\{r_t\}_{t=1}^T$ , а задача состоит в оценивании однопериодного  $\text{VaR}_{T+1|T}^p$  — истинного  $(1 - p)$ -квантиля распределения доходности  $r_{T+1}$ , условно на предыстории  $r_T, r_{T-1}, \dots, r_1$ .

Наиболее примитивный способ оценивания (*historically simulated, HS-VaR*) состоит в вычислении эмпирического  $(1 - p)$ -квантиля по подвыборке  $\{r_t\}_{t=T-L+1}^T$  длины  $L \leq T$ . Как отмечают Manganelli & Engle (2001), такой подход является логически противоречивым. Если мы предполагаем, что процесс  $\{r_t\}$  эргодичен и стационарен, то  $\text{HS-VaR}_{t+1|t}^p$  сходится к  $(1 - p)$ -квантилю *безусловного* распределения доходностей при неограниченном увеличении длины окна  $L$ , и тогда разумно положить  $L = T$ . Однако, на практике обычно требуется оценка *условного* квантиля, и значит длина окна не может быть слишком большой: в противном случае оценка будет испорчена слишком давними наблюдениями. Даже если требуется оценить безусловный квантиль, качество оценки  $\text{HS-VaR}_{t+1|t}^p$  в конечных выборках оставляет желать лучшего, особенно если для процесса  $\{r_t\}$  характерна сильная кластеризация волатильности, что для финансовых доходностей верно почти всегда.

Для оценивания условного квантиля существует множество подходов, которые можно грубо разделить на две категории: моделирование условного распределения доходностей (в том

<sup>1</sup>В работе Danielsson, de Haan, Peng & Vries (2001) показано, впрочем, что в большинстве практически релевантных случаев VaR является субаддитивной мерой риска.

числе полу- и непараметрические методы) и непосредственное моделирование условного квантиля. В настоящей работе мы сконцентрируем внимание только на моделях первого типа.

Положим  $\mathcal{F}_{t-1} = \{r_{t-1}, r_{t-2}, \dots\}$ ,  $\mu_t = \mathbb{E}[r_t | \mathcal{F}_{t-1}]$ ,  $\sigma_t^2 = \mathbb{V}[r_t | \mathcal{F}_{t-1}]$  и представим динамику  $r_t$  в стандартном виде

$$\begin{aligned} r_t &= \mu_t + \eta_t, \\ \eta_t &= \sigma_t \varepsilon_t, \quad \{\varepsilon_t\} \sim iid(0, 1), \\ \varepsilon_t | \mathcal{F}_{t-1} &\sim F(\cdot | \theta). \end{aligned}$$

Описываемые ниже модели отличаются только в динамике условной дисперсии  $\sigma_t^2$  и предположениях об условном распределении  $F$  стандартизированной ошибки  $\varepsilon_t$ . Условное среднее  $\mu_t$  как правило незначимо отличается от нуля, поэтому моделируется редко, так что будем считать, что  $\mu_t \equiv 0$ . После получения оценок условной дисперсии  $\hat{\sigma}_{t+1}^2$  и условной функции распределения  $F(\cdot | \hat{\theta})$  оценка VaR вычисляется просто:

$$\widehat{\text{VaR}}_{t+1|t}^p = \hat{\sigma}_{t+1} F^{-1}(1 - p | \hat{\theta}).$$

Простейшая модель условной дисперсии — так называемая методология RiskMetrics, разработанная инвестиционным банком JPMorgan, см. Zumbach (2007). Условная дисперсия имеет динамику

$$\sigma_t^2 = \lambda \sigma_{t-1}^2 + (1 - \lambda) r_{t-1}^2.$$

Легко видеть, что условная дисперсия представляется в виде «экспоненциально сглаженной» суммы квадратов доходностей прошедших периодов:

$$\sigma_t^2 = (1 - \lambda) \sum_{i=1}^{\infty} \lambda^i r_{t-i}^2,$$

а значит не обладает желательным свойством возвращения к среднему. Более того, структура волатильности в модели RiskMetrics линейная: условная дисперсия  $\sigma_{t+k|t}^2$  доходности за  $k$  дней равна  $k \sigma_{t+1|t}^2$ .

Модель RiskMetrics является ограничением более общей модели обобщённой авторегрессионной условной гетероскедастичности GARCH( $p, q$ ), предложенной в работе Bollerslev (1986):

$$\sigma_t^2 = \omega + \alpha_1 \eta_{t-1}^2 + \dots + \alpha_q \eta_{t-q}^2 + \beta_1 \sigma_{t-1}^2 + \dots + \beta_p \sigma_{t-p}^2,$$

где  $p$  и  $q$  — натуральные числа. В отличие от RiskMetrics, модель GARCH обладает свойством возвращения к среднему:

$$\sigma_t^2 - \sigma^2 = \alpha(r_{t-1}^2 - \sigma^2) + \beta(\sigma_{t-1}^2 - \sigma^2),$$

где величина  $\sigma = \omega / (1 - \alpha - \beta)$  интерпретируется как *долгосрочная волатильность*.

При моделировании распределения стандартизированной ошибки  $\varepsilon_t$  необходимо найти баланс между простотой вида распределения (а значит, и соответствующей функции правдоподобия) и гибкостью спецификации. Например, для нормального распределения  $\mathcal{N}(0, 1)$  известен явный вид функции правдоподобия; более того, теория квазиравдоподобия для GARCH-моделей, развитая в работе Bollerslev & Wooldridge (1992), показывает, что даже в случае неверной спецификации нормального распределения оценивание может давать состоятельные оценки. С другой стороны, очевидным недостатком (стандартного) гауссовского

распределения является отсутствие параметров и неспособность уловить такой важный стилизованный факты о распределении доходностей, как скошенность<sup>2</sup>.

В следующем параграфе будут описаны модели с динамическими моментами высоких порядков; параметрические распределения стандартизованных ошибок в этих моделях находят применение и в более простом случае постоянных моментов.

## 2.2 Модели с динамическими условными моментами высоких порядков

Многие работы, например, Harvey & Siddique (1999), Jondeau & Rockinger (2003) и Bali, Mo & Tang (2008) показывают, что условные скошенности и эксцессы финансовых доходностей не являются постоянными во времени. В настоящем параграфе делается обзор некоторых моделей динамики моментов высоких порядков.

Основополагающей работой, посвящённой моделированию динамики моментов высоких порядков, является статья Hansen (1994). В ней предложена общая методология моделирования условных плотностей, имеющих авторегрессионную структуру (*autoregressive conditional density, A/R/CD*), и получено обобщение распределение Стьюдента на случай нетривиальных скошенности и куртозиса. Плотность SGT-распределения (*skewed generalized t*) имеет вид

$$f(z|\lambda, \eta, \kappa) = C \left( 1 + \frac{|z + \delta|^\kappa}{(\eta + 1)/\kappa (1 + \operatorname{sgn}(z + \delta)\lambda)^\kappa \theta^\kappa} \right)^{\frac{\eta+1}{\kappa}},$$

где

$$\begin{aligned} C &= 0.5\kappa \left( \frac{\eta + 1}{\kappa} \right)^{-\frac{1}{\kappa}} B \left( \frac{\eta}{\kappa}, \frac{1}{\kappa} \right)^{-1} \theta^{-1}, \\ \theta &= \frac{1}{\sqrt{g - \rho^2}}, \\ \rho &= 2\lambda B \left( \frac{\eta}{\kappa}, \frac{1}{\kappa} \right)^{-1} \left( \frac{\eta + 1}{\kappa} \right)^{\frac{1}{\kappa}} B \left( \frac{\eta - 1}{\kappa}, \frac{2}{\kappa} \right), \\ g &= (1 + 3\lambda^2) B \left( \frac{\eta}{\kappa}, \frac{1}{\kappa} \right)^{-1} \left( \frac{\eta + 1}{\kappa} \right)^{\frac{2}{\kappa}} B \left( \frac{\eta - 2}{\kappa}, \frac{3}{\kappa} \right), \\ \delta &= \rho\theta. \end{aligned}$$

Здесь параметр  $\lambda \in (-1, 1)$  отвечает за скошенность,  $\eta > 2$  — за толщину хвостов,  $\kappa > 0$  — за куртозис, а за  $B(\cdot)$  обозначена бета-функция. Чтобы избежать необходимости решать задачу максимального правдоподобия с ограничениями, Хансен предлагает использовать подходящие логистические преобразования параметров. Автор накладывает простую ARCH-динамику на параметры и показывает, что полученная модель показывает отличные результаты для дневных данных по обменному курсу USD/CHF. Jondeau & Rockinger (2003) приводят список возможных спецификаций динамики параметров в таких моделях, в том числе спецификации типа асимметричной GARCH-модели и спецификации для динамики третьего и четвёртого начальных моментов непосредственно.

Аналогичная методология применяется в работе Bali, Mo & Tang (2008). Авторы используют SGT-распределение и результаты Theodossiou (1998) для моделирования стандартизованных доходностей с динамикой скошенности, куртозиса и толщины хвостов. Условное среднее моделируется простой авторегрессией первого порядка AR(1), а условная дисперсия — с помощью различных GARCH-моделей. Параметры  $\lambda_t, \eta_t, \kappa_t$  записываются в виде

<sup>2</sup>Здесь необходимо отметить, что речь идёт про *безусловную* скошенность, которая вряд ли возможна без *условной* скошенности. Что касается куртозиса, то можно показать, что даже в гауссовской модели GARCH безусловное распределение доходностей имеет тяжёлые хвосты; в этом отношении предположение о нормальности не является столь ограничительным.

$\lambda_t = -1 + 2/(1 + \exp(\tilde{\lambda}_t))$ ,  $\eta_t = 2 + \exp(\tilde{\eta}_t)$  и  $\kappa_t = \exp(\tilde{\kappa}_t)$ , после чего на неограниченные вспомогательные параметры  $\tilde{\lambda}_t$ ,  $\tilde{\eta}_t$  и  $\tilde{\kappa}_t$  накладывается обычная ARCH-динамика

$$\tilde{\lambda}_t = \lambda^0 + \lambda^1 z_{t-1} + \lambda^2 \tilde{\lambda}_{t-1}, \quad (1)$$

$$\tilde{\eta}_t = \eta^0 + \eta^1 z_{t-1} + \eta^2 \tilde{\eta}_{t-1}, \quad (2)$$

$$\tilde{\kappa}_t = \kappa^0 + \kappa^1 z_{t-1} + \kappa^2 \tilde{\kappa}_{t-1}, \quad (3)$$

Показано, что такая SGT-GARCH модель генерирует очень точные и робастные оценки VaR, причём гипотеза о постоянстве параметров во времени уверенно отвергается.

В работе Wilhelmsson (2009) предлагается использовать нормальное обратное гауссовское (*Normal Inverse Gaussian, NIG*) распределение для моделирования стандартизованных доходностей. NIG-распределение обладает двумя основными преимуществами. Во-первых, как показано в работе Jondeau & Rockinger (2003), по сравнению с SGT-распределением Hansen (1994) NIG-распределение позволяет моделировать более широкий класс комбинаций скошенности и куртозиса. Во-вторых, использование NIG-распределения имеет теоретическое обоснование: в модели Barndorff-Nielsen (1997) со скрытым процессом появления информации доходности имеют условное NIG-распределение. Плотность NIG-распределения имеет вид

$$f(x; \alpha, \beta, \mu, \delta) = \frac{\alpha}{\pi} \exp\left(\delta\sqrt{\alpha^2 - \beta^2} - \beta\gamma\right) q\left(\frac{x - \mu}{\delta}\right)^{-1} K_1\left(\delta\alpha q\left(\frac{x - \mu}{\delta}\right)\right) \exp(\beta x),$$

где  $q(z) = \sqrt{1 + z^2}$ , а  $K_1(\cdot)$  — модифицированная функция Бесселя порядка 3 и индекса 1. Параметр  $\alpha$  отвечает за куртозис,  $\beta$  — за скошенность,  $\mu$  — за центр плотности на действительной оси, а  $\sigma$  — масштабирующий параметр. NIG-распределение замкнуто относительно взятия свёртки, а значит, если однодневные доходности NIG-распределены, то многодневные доходности также получаются NIG-распределёнными. Однако дисперсия NIG-распределения зависит от нескольких параметров, что неудобно для моделирования динамики волатильности. Чтобы избавиться от этой зависимости, Jensen & Lunde (2001) предлагают перепараметризацию  $\bar{\alpha} = \alpha\delta$ ,  $\bar{\beta} = \beta\delta$ .

Приведём далее полную спецификацию модели NIG-ACD. Положим  $\bar{\gamma}_t = \sqrt{\bar{\alpha}_t^2 + \bar{\beta}_t^2}$  и  $\rho_t = \beta_t/\alpha_t \in [0, 1)$ . Тогда доходности  $\{r_t\}$  имеют вид

$$r_t = \mu + \bar{\gamma}_t^{1/2} \delta_t \rho_t + \varepsilon_t, \quad (4)$$

$$\varepsilon_t = \delta_t \eta_t, \quad \eta_t \sim NIG\left(\bar{\alpha}_t, \bar{\beta}_t, -\bar{\gamma}_t^{1/2} \rho_t, \bar{\gamma}_t^{3/2} / \bar{\alpha}_t\right), \quad (5)$$

$$\delta_t^v = c + b\delta_{t-1}^v + a(|\varepsilon_{t-1}| - \tau\varepsilon_{t-1})^v, \quad (6)$$

$$\bar{\gamma}_t = \exp\left(\lambda_0 + \lambda_1 \varepsilon_{t-1}^2 + \lambda_2 \log(\bar{\gamma}_{t-1})\right), \quad (7)$$

$$\tilde{\rho}_t = \theta_0 + \theta_1 \varepsilon_{t-1}. \quad (8)$$

где  $\tilde{\rho}_t = \log\left(\frac{1+\rho_t}{1-\rho_t}\right) \in \mathbb{R}$  — удобная перепараметризация параметра скошенности  $\rho$ . Условное стандартное отклонение  $\delta_t$  имеет динамику типа асимметричной степенной ARCH-модели (APARCH), условная «крутизна»  $\bar{\gamma}_t$  эволюционирует согласно экспоненциальной ARCH-модели, а модифицированный параметр скошенности  $\tilde{\rho}_t$  имеет простую MA(1)-динамику. Сложный вид спецификации обусловлен тем, что необходимо гарантировать нулевое среднее и единичную дисперсию ошибок  $\eta_t$ .

Условное распределение  $r_{t+1}$  в такой модели имеет вид

$$r_{t+1} | \mathcal{F}_t \sim NIG\left(\alpha_t, \beta_t, \mu, \bar{\gamma}_t^{3/2} / \alpha_t\right).$$

Таким образом, модель легко оценивается условным методом максимального правдоподобия.

Wilhelmsson (2009) показывает, что для дневных доходностей индекса S&P500 за период с 1962 по 2005 годы модель NIG-ACD с динамическими моментами высоких порядков показывает лучшие результаты, чем стандартные GARCH-модели с нормальным и NIG распределением ошибок, модели ARCD из работы Hansen (1994), а также NIG-ACD модели с постоянными моментами высоких порядков.

Мы предлагаем использовать спецификацию для динамики параметров  $\lambda_t$  и  $\kappa_t$  из модели SGT-GARCH и спецификацию условной волатильности из модели NIG-ACD для скошенного обобщённого распределения ошибок (*skewed generalized error, SGE*), предложенного в Theodossiou (2000), и показываем далее, что новая модель по крайней мере не менее успешна в смысле качества прогнозирования VaR, одновременно являясь вычислительно более простой. Плотность SGE-распределения имеет вид:

$$f(x|\mu, \sigma, \kappa, \lambda) = \frac{C}{\sigma} \exp\left(-\frac{|x - \mu + \delta\sigma|^\kappa}{[1 - \operatorname{sgn}(x - \mu - \delta\sigma)\lambda]^\kappa \theta^\kappa \sigma^\kappa}\right),$$

где

$$\begin{aligned} C &= \frac{\kappa}{2\theta} \Gamma(1/\kappa)^{-1}, \\ \theta &= \Gamma(1/\kappa)^{1/2} \Gamma(3/\kappa)^{-1/2} S(\lambda)^{-1}, \\ \delta &= 2\lambda A S(\lambda)^{-1}, \\ S(\lambda) &= \sqrt{1 + 3\lambda^2 - 4A^2\lambda^2}, \\ A &= \Gamma(2/\kappa) \Gamma(1/\kappa)^{-1/2} \Gamma(3/\kappa)^{-1/2}. \end{aligned}$$

Здесь параметр  $\lambda \in (-1, 1)$  отвечает за скошенность и контролирует скорость убывания плотности в окрестности моды распределения, параметр  $\kappa > 0$  контролирует тяжесть хвостов и высоту пика распределения, а параметры  $\mu = \mathbb{E}[x]$  и  $\sigma = \sqrt{\mathbb{V}[x]}$ . Частными случаями SGE-распределения являются распределение Субботина ( $\lambda = 0$ ), распределение Лапласа ( $\lambda = 0$ ,  $\kappa = 1$ ), нормальное распределение ( $\lambda = 0$ ,  $\kappa = 2$ ); при  $\kappa \rightarrow \infty$  и фиксированном  $\lambda = 0$  распределение становится всё более «равномерным»<sup>3</sup>.

Таким образом, в новой SGE-ACD модели параметры распределения ошибок  $\mu = 0$ ,  $\sigma = 1$ , параметры  $\lambda_t$  и  $\kappa_t$  имеют динамику (1) и (3) соответственно, а в качестве модели волатильности  $\delta_t = \sqrt{\mathbb{V}[r_t | \mathcal{F}_{t-1}]}$  используется APARCH (6).

### 3 Тестирование предсказательной способности моделей VaR

После оценивания набора моделей и вычисления прогнозов VaR по каждой модели необходимо ответить на два вопроса: (1) можно ли для конкретной модели считать, что она «успешна» в прогнозировании VaR? (2) какая из моделей «лучше» с точки зрения прогнозирования VaR?

Ответ на первый вопрос зависит от определения «успешности» модели, под которой наиболее естественно понимать корректное поведение процесса пробоев прогнозов VaR. В этом случае можно воспользоваться классическими результатами Christoffersen (1998), который предлагает трёхшаговую процедуру тестирования качества построенных доверительных интервалов<sup>4</sup>. Сначала тестируется корректность *безусловного накрытия*, т.е. гипотеза о том, что вычисленные прогнозы VaR действительно пробиваются с вероятностью  $1 - p$ . Затем тестируется отсутствие кластеризации пробоев VaR, т.е. корректность *условного накрытия*. На последнем шаге тестируется *совокупная* корректность накрытия.

<sup>3</sup>Однако предел не определён, так как равномерного распределения на  $\mathbb{R}$  не существует.

<sup>4</sup>Заметим, что VaR можно рассматривать как односторонний интервал доверительной вероятности  $p$  вида  $(\text{VaR}, +\infty)$ .

Формально, обозначим за  $I_t = \mathbb{I}\left(r_t > \widehat{\text{VaR}}_{t|t-1}^p\right)$ ,  $t = 1, \dots, T$  индикатор отсутствия пробоя прогноза VaR в момент  $t$ . Тогда при нулевой гипотезе о корректном *безусловном* накрытии  $I_t \sim \mathcal{B}(p)$ , а значит статистика отношения правдоподобия

$$\text{LR}_{UC} = 2 \log \left[ \frac{\hat{\pi}_1^{T_1} (1 - \hat{\pi}_1)^{T-T_1}}{(1 - \alpha)^{T_1} \alpha^{T-T_1}} \right] \xrightarrow{d} \chi^2(1),$$

где  $T_1$  — количество единиц в последовательности индикаторов и  $\hat{\pi}_1 = T_1/T$ .

Для тестирования отсутствия кластеризации Christoffersen (1998) предлагает тестировать гипотезу о независимости индикаторов против альтернативной гипотезы о том, что последовательность индикаторов образует марковскую цепь первого порядка. В такой постановке статистика отношения правдоподобия

$$\text{LR}_{IND} = 2 \log \left[ \frac{(1 - \hat{\pi}_{01})^{T_0 - T_{01}} \hat{\pi}_{01}^{T_{01}} (1 - \hat{\pi}_{11})^{T_1 - T_{11}} \hat{\pi}_{11}^{T_{11}}}{\hat{\pi}_1^{T_1} (1 - \hat{\pi}_1)^{T - T_1}} \right] \xrightarrow{d} \chi^2(1),$$

где  $T_{ij} = \#\{I_t = i, I_{t+1} = j\}$  и  $\hat{\pi}_{ij} = T_{ij}/T_i$ .

Наконец, в статье показывается, что при корректном *совокупном* накрытии (эквивалентно, верности обеих нулевых гипотез выше или  $I_t \sim iid \mathcal{B}(p)$ ) статистика

$$\text{LR}_{CC} = \text{LR}_{UC} + \text{LR}_{IND} \xrightarrow{d} \chi^2(2).$$

В работе Christoffersen & Pelletier (2004) отмечается, что во многих случаях асимптотическое приближение работает плохо, и предлагается бутстрапирование статистики  $\text{LR}_{CC}$ .

Что касается ответа на второй вопрос, то в качестве критерия для сравнения (двух) моделей VaR можно использовать некоторую меру качества предсказания левых хвостов условных распределений доходностей. Ясно, что хороший прогноз левого хвоста будет давать качественную оценку VaR, однако обратное неверно, а значит, тестирование на одинаковое качество прогноза левых хвостов — несколько избыточная задача с точки зрения меры VaR. Тем не менее, для параметрических моделей такой подход является естественным и позволяет косвенно судить о том, насколько одна из моделей более подходит для прогнозирования мер риска, более чувствительных к распределению потерь, например, Expected Shortfall.

Одну из методологий сравнения качества out-of-sample прогнозов распределений предлагают Diks, Panchenko & van Dijk (2011). Авторы модифицируют существующие скоринговые методы, основанные на взвешенном правдоподобии, чтобы обойти проблему смещения стандартных тестов в сторону прогнозов, которые приписывают интересующему исследователя хвосту большую совокупную вероятность. А именно, они определяют *цензурированный скор*  $S^{csl}(\hat{f}_t(\cdot); r_{t+1})$  для прогноза  $\hat{f}_t(\cdot)$  истинной условной плотности  $f_{t+1|t}(\cdot)$  доходности  $r_{t+1}$ :

$$S^{csl}(\hat{f}_t(\cdot); r_{t+1}) = \mathbb{I}_{A_t}(r_{t+1}) \log \hat{f}_t(r_{t+1}) + \mathbb{I}_{\mathbb{R} \setminus A_t}(r_{t+1}) \log \int_{\mathbb{R} \setminus A_t} \hat{f}_t(s) ds,$$

где  $\mathbb{I}_A(\cdot)$  — индикатор множества  $A$ . Нас интересуют левые хвосты, поэтому положим  $A_t := (-\infty, q)$  для некоторого малого квантиля  $q$  распределения  $r$ . Авторы показывают, что при некоторых технических предположениях цензурированный скор  $S^{csl}$  является *корректно построенным* в том смысле, что для него выполняется аналог информационного неравенства: средний цензурированный скор любого прогноза  $\hat{f}_t(\cdot)$  не может быть больше среднего цензурированного скор истинной плотности  $f_t(\cdot)$ , т.е.

$$\mathbb{E}_t S^{csl}(\hat{f}_t(\cdot); r_{t+1}) \leq \mathbb{E}_t S^{csl}(f_t(\cdot); r_{t+1}).$$

Таблица 1: Описательные статистики для дневных доходностей с 01.01.1996 до 31.12.2010

	ВАС	F	GE	JPM	PFE
Mean	0,000613	0,000651	0,000428	0,000673	0,000421
Std	0,0317	0,0291	0,0208	0,0275	0,0191
Min	-0,2897	-0,2500	-0,1279	-0,2072	-0,1115
Max	0,3527	0,2951	0,1970	0,2610	0,1017
Skew	0,910	0,655	0,310	0,828	-0,035
Kurt	27,18	16,24	10,65	15,23	6,13
JB	92584	27848	9274	23972	1547
ACF(1)	0,00208	0,00165	-0,00930	-0,07739	0,00276
ACF(2)	0,014	0,049	-0,016	-0,030	-0,0929
ACF(3)	-0,037	0,037	-0,008	-0,020	-0,022
Q(5)	28,70	31,91	11,65	31,14	36,02

Mean, Std, Min, Max, Skew и Kurt означают среднее, стандартное отклонение, минимум, максимум, эмпирическую скошенность и куртозис соответственно. JB означает статистику Харке–Бера (5%-ое критическое значение 5,9773). ACF(1), ACF(2), ACF(3) — первые три лага эмпирической автокорреляционной функции. Q(5) — статистика Льюнга–Бокса (5%-ое критическое значение 11,0705).

Для конкурирующих прогнозов  $\hat{f}_t(\cdot)$  и  $\hat{g}_t(\cdot)$  необходимо протестировать гипотезу о равенстве средних цензурированных скоров

$$H_0 : \mathbb{E}_t S^{csl}(\hat{f}_t(\cdot); r_{t+1}) = \mathbb{E}_t S^{csl}(\hat{g}_t(\cdot); r_{t+1}) \quad \forall t = 0, \dots, n-1.$$

Тестовая статистика имеет вид стандартизованной средней разности эмпирических цензурированных скоров двух прогнозов:

$$DM_n = \frac{\bar{d}_n^*}{\sqrt{\hat{\sigma}_n^2/n}},$$

где

$$\bar{d}_n^* = \frac{1}{n} \sum_{t=0}^{n-1} \left( S^{csl}(\hat{f}_t(\cdot); r_{t+1}) - S^{csl}(\hat{g}_t(\cdot); r_{t+1}) \right),$$

а  $\hat{\sigma}_n^2$  — состоятельная оценка асимптотической дисперсии величины  $\bar{d}_n^*$ . Применяя классический результат работы Diebold & Mariano (1995), авторы показывают, что при нулевой гипотезе статистика  $DM_n$  имеет стандартное нормальное асимптотическое распределение.

Таблица 2:  $P$ -значения тестов на предсказательную способность VaR для различных моделей

	BAC	F	GE	JPM	PFE
HS-VaR					
Безусловный	0,01	0,07*	0,04	0,00	0,22*
Условный	0,00	0,03	0,00	0,00	0,01
Совместный	0,00	0,01	0,00	0,00	0,02
N-GARCH(1,1)-VaR					
Безусловный	0,56*	0,49*	0,35*	0,44*	0,02
Условный	0,34*	0,35*	0,29*	0,06*	0,39*
Совместный	0,59*	0,61*	0,15*	0,13*	0,07*
SGT-GARCH(1,1)-VaR с постоянными моментами высоких порядков					
Безусловный	0,93*	0,89*	0,13*	0,64*	0,06*
Условный	0,95*	0,39*	0,42*	0,04	0,84*
Совместный	0,99*	0,68*	0,24*	0,13*	0,16*
NIG-GARCH(1,1)-VaR с постоянными моментами высоких порядков					
Безусловный	0,88*	0,81*	0,43*	0,53*	0,19*
Условный	0,94*	0,28*	0,43*	0,08*	0,80*
Совместный	0,99*	0,65*	0,36*	0,18*	0,20*
SGT-GARCH(1,1)-VaR с динамическими моментами высоких порядков					
Безусловный	0,93*	0,89*	0,13*	0,64*	0,06*
Условный	0,95*	0,39*	0,42*	0,04	0,84*
Совместный	0,99*	0,68*	0,24*	0,13*	0,16*
NIG-ACD-VaR с динамическими моментами высоких порядков					
Безусловный	0,88*	0,81*	0,43*	0,53*	0,19*
Условный	0,94*	0,28*	0,43*	0,08*	0,80*
Совместный	0,99*	0,65*	0,36*	0,18*	0,20*
SGE-ACD-VaR с динамическими моментами высоких порядков					
Безусловный	0,83*	0,71*	0,67*	0,62*	0,49*
Условный	0,88*	0,31*	0,29*	0,14*	0,23*
Совместный	0,89*	0,59*	0,43*	0,46*	0,25*

Нулевые гипотезы тестов на корректное безусловное, условное и совокупное покрытие, соответственно:  $\mathbb{E}(I_t) = 1-p$ ,  $I_t \sim$  независимы и  $I_t \sim \text{iid } \mathcal{B}(1-p)$ , где  $I_t$  — индикатор пробоя VaR. Звёздочками обозначены случаи неотвержения нулевой гипотезы на 5%-ом уровне значимости.

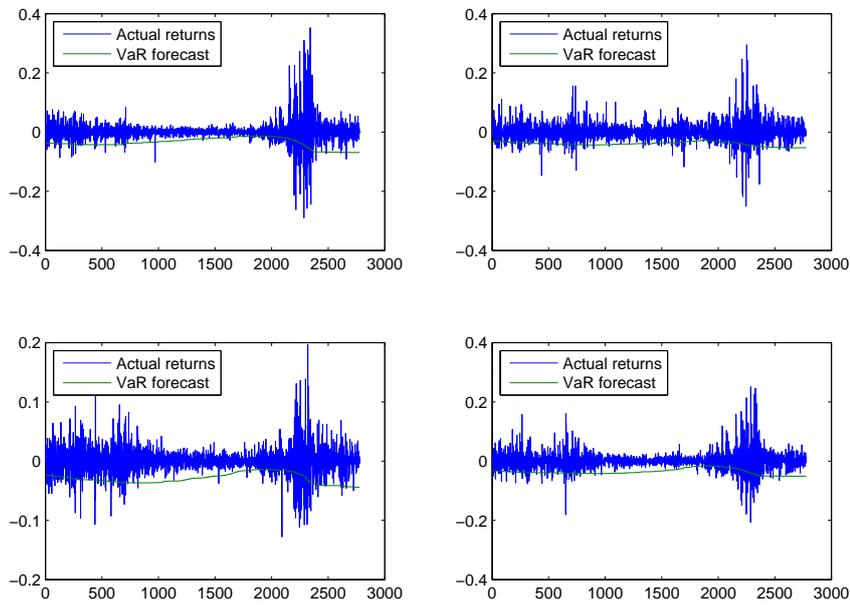


Рис. 1: HS-VaR для акций BAC, F, GE, JPM

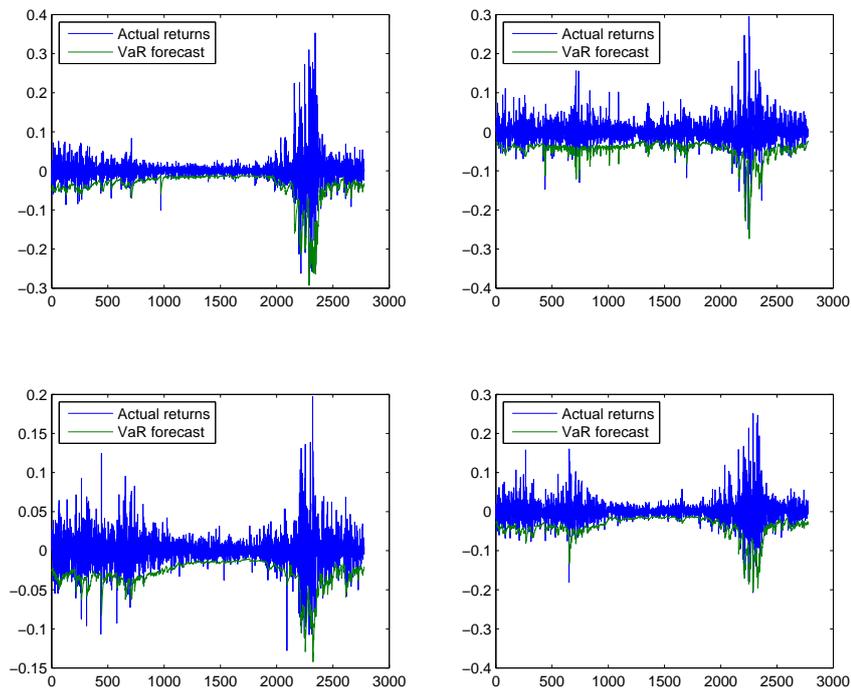


Рис. 2: SGE-ACD-VaR для акций BAC, F, GE, JPM

Таблица 3:  $P$ -значения тестов на одинаковую предсказательную способность для левых хвостов распределений для пар моделей с динамическими моментами высоких порядков

	VAC	F	GE	JPM	PFE
NIG-ACD и SGT-GARCH(1,1)	0,12*	0,04	0,07*	0,08*	0,12*
SGE-ACD и SGT-GARCH(1,1)	0,25*	0,63*	0,12*	0,32*	0,19*
SGE-ACD и NIG-ACD	0,86*	0,78*	0,56*	0,42*	0,53*

Нулевая гипотеза  $H_0 : \mathbb{E}_t S(\hat{f}_t; r_{t+1}) = \mathbb{E}_t S(\hat{g}_t; r_{t+1})$  для всех  $t$  (одинаковая предсказательная способность). Звёздочками обозначены случаи неотвержения  $H_0$  на 5%-ом уровне значимости.

## 4 Эмпирическое сравнение моделей VaR с динамическими моментами высоких порядков

### 4.1 Данные

Мы используем дневные доходности по ценам закрытия нескольких ликвидных<sup>5</sup> акций, входящих в индекс широкого рынка S&P500, — Bank of America, Ford Motor Co., General Electric, J.P. Morgan и Pfizer, — с 1996 по 2010 годы включительно, что даёт 3778 наблюдений каждого ряда доходностей. Рассматриваемый период включает турбулентность финансового кризиса, что позволяет отследить реакцию различных способов оценивания VaR на отрицательные рыночные шоки. Описательные статистики представлены в таблице 1. Видно, что безусловное распределение доходностей далеко от гауссовского. Тест Льюнга–Бокса отвергает гипотезу об отсутствии автокорреляции на 5%-м уровне значимости, что может служить первым признаком наличия динамических моментов. Эмпирическая скошенность положительна для всех бумаг, кроме одной, что необычно, т.к. считается, что скошенность безусловного распределения доходностей должна быть отрицательной, см., например, Manganelli & Engle (2001). Эмпирический куртозис значительно больше 3 (значения, соответствующего гауссовскому распределению), что может свидетельствовать о наличии тяжёлых хвостов у распределения доходностей.

### 4.2 Оценивание и тестирование моделей

Для каждой модели прогноз 5%-го однодневного VaR вычисляется в скользящем окне длины<sup>6</sup>  $L = 1000$  с 10.03.1997 до конца выборки 31.12.2010.

Оценим для начала VaR простым выборочным квантилем (HS-VaR), в целях иллюстрации неадекватности метода. В первой секции таблицы 2 приведены результаты тестов на корректность накрытия. Видно, что HS-VaR плох не только как оценка условного VaR, но и для оценивания безусловного VaR. Причиной этого может быть недостаточность длины окна или нестационарность доходностей. Неудовлетворительное качество оценивания условного VaR видно на Рис. 1: HS-VaR не способен быстро отразить недавние шоки, и поэтому характеризуется большими запаздываниями.

Посмотрим теперь на результаты прогнозирования VaR с помощью моделей GARCH(1,1) с нормальным и скошенным студентовским распределением ошибок. Результаты тестирования моделей представлены во второй и третьей секциях таблицы 2. Видно, что GARCH(1,1) даёт значительно более точные прогнозы VaR, чем выборочный квантиль. Как условно, так

<sup>5</sup>См. веб-страницу [www.marketvolume.com/stocks/mosttraded.asp?s=nyse](http://www.marketvolume.com/stocks/mosttraded.asp?s=nyse)

<sup>6</sup>Такая длина окна стандартна для вычисления 5%-го однодневного VaR, см., например, Kuuster, Mittnik & Paolella (2006).

и безусловно модель показывает себя на удивление точной: только для акций PFE N-GARCH имеет неточное безусловное покрытие (однако для всех рядов нельзя отвергнуть гипотезу о точном совместном покрытии). Ответы VaR на отрицательные шоки для GARCH моделей по сравнению с наивным историческим квантилем характеризуются значительно меньшими задержками.

Теперь перейдём к моделям к динамическими условными моментами высоких порядков. Прогнозы VaR в модели SGT-GARCH хороши как в смысле условного, так и безусловного покрытия, и доминируют  $p$ -значения модели NIG-ACD, см. таблицу 2. Последняя показывает равномерно лучшие результаты, чем простая модель NIG-GARCH(1,1) с постоянными моментами, что согласуется с выводами работы Wilhelmsson (2009). То же верно и для предложенной нами модели SGE-ACD (см. последнюю секцию таблицы 2). Рис. 2 показывает, что модель SGE-ACD позволяет получить прогнозы VaR с меньшей кластеризацией и более быстрой реакцией на неожиданные скачки волатильности.

Так как  $p$ -значения тестирования на корректность покрытия для модели SGE-ACD близки к  $p$ -значениям для моделей SGT-GARCH и NIG-ACD, встаёт естественный вопрос: имеет ли модель SGE-ACD значимо лучшую (худшую) предсказательную способность для VaR, чем модели SGT-GARCH и NIG-ACD? Скоринговый метод Diks, Panchenko & van Dijk (2011), описанный в предыдущем разделе, идеально подходит для разрешения этого вопроса. Результаты в таблице 3 показывают, что доминирование модели SGT-GARCH над моделью NIG-ACD может быть ложным. В самом деле, только для акций Ford нулевая гипотеза об одинаковой предсказательной способности в левом хвосте распределения может быть отвергнута на 5% уровне значимости, причём соответствующее  $p$ -значение всего лишь немногим меньше 5%. Сравнивая аналогичным образом модель SGE-ACD попарно с моделями SGT-GARCH(1,1) и NIG-ACD, из таблицы 3 видим, что тест не может различить предсказательную способность трёх моделей в хвостах распределения, и значит, скорее всего, не сможет различить качество предсказаний условных квантилей VaR.

## 5 Заключение

Разрушительные последствия глобальной рецессии не оставляют сомнений в том, что эффективный риск-менеджмент должен быть важнейшим приоритетом мировых финансовых институтов. Сложность динамики финансовых рынков делает необходимым использование продвинутых моделей для расчёта различных мер риска, таких как VaR, который получил в последнее время широкое распространение в бизнесе из-за концептуальной простоты и регуляторных требований к его вычислению.

В настоящей работе производится сравнение некоторых моделей финансовых доходностей с динамическими условными моментами высоких порядков, таких как скошенность и куртозис. Так как считается, что непостоянство условных моментов критически сказывается на предсказании условных распределений доходностей, в качестве меры качества моделей выступают прогнозы VaR. Мы действительно находим, что по качеству условных прогнозов VaR модели с динамическими моментами, такие как GARCH(1,1) со скошенным студентовским (SGT) распределением ошибок и модель авторегрессионной условной нормальной обратной гауссовской плотности NIG-ACD превосходят свои более традиционные аналоги с постоянными условными моментами. Предложенная нами комбинация данных моделей (SGE-ACD) со скошенным обобщённым распределением ошибки позволяет получить столь же качественные прогнозы VaR, в то же время являясь вычислительно менее сложной, особенно в сравнении с моделью NIG-ACD.

В заключение необходимо отметить, что в последнее время набирают популярность полупараметрические методы оценивания VaR, а также использование теории экстремальных значений для более точного оценивания хвостов условных распределений, см., например,

Kuester, Mittnik & Paoletta (2006). Насколько нам известно, в настоящее время не существует достаточно полного сравнения таких методов с параметрическим подходом, а значит, разработка методологии для такого сравнения может являться перспективным направлением дальнейших исследований.

## Благодарности

Автор искренне благодарит своих научных руководителей в Российской экономической школе Станислава Анатольева и Станислава Храпова, а также участников исследовательского проекта РЭШ по анализу временных рядов 2012–2013 гг. за ценные замечания и плодотворные обсуждения.

## Список литературы

- Bali, T. G., H. Mo & Y. Tang (2008). The role of autoregressive conditional skewness and kurtosis in the estimation of conditional VaR. *Journal of Banking & Finance* 32, 269–282.
- Barndorff-Nielsen, O. E. (1997). Normal inverse gaussian distributions and stochastic volatility modelling. *Scandinavian Journal of Statistics* 24, 1–13.
- Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics* 31, 307–327.
- Bollerslev, T. & J. M. Wooldridge (1992). Quasi-maximum likelihood estimation and inference in dynamic models with time-varying covariances. *Econometric Reviews* 11, 143–172.
- Christoffersen P. & D. Pelletier (2004). Backtesting value-at-risk: a duration based approach. *Journal of Financial Econometrics* 2, 84–108.
- Christoffersen, P. F. (1998). Evaluating interval forecasts. *International Economic Review* 39, 841–862.
- Danielsson, J., D. Jorgensen, S. Mandira, G. Samorodnitsky & C. Vries (2005). Subadditivity re-examined: the case for value-at-risk. Working paper. Financial Markets Group, London School of Economics and Political Science.
- Diebold, F. & R. Mariano (1995). Comparing predictive accuracy. *Journal of Business & Economic Statistics* 13, 253–263.
- Diks, C., V. Panchenko & D. van Dijk (2011). Likelihood-based scoring rules for comparing density forecasts in tails. *Journal of Econometrics* 163, 215–230.
- Föllmer, H. & A. Schied (2004). *Stochastic finance. An introduction in discrete time*. Walter de Gruyter, Berlin.
- Hansen, B. (1994). Autoregressive conditional density estimation. *International Economic Review* 35, 705–730.
- Harvey, C. & A. Siddique (1999). Autoregressive conditional skewness. *Journal of Financial & Quantitative Analysis* 34, 465–487.
- Jensen, M. B. & A. Lunde (2001). The NIG-S&ARCH model: a fat-tailed, stochastic and autoregressive conditional heteroskedastic volatility model. *Econometrics Journal* 4, 319–342.
- Jondeau, E. & M. Rockinger (2003). Conditional volatility, skewness and kurtosis: existence, persistence and comovements. *Journal of Economic Dynamics and Control* 27, 1699–1737.
- Kuester, K., S. Mittnik & M. Paoletta (2006). Value-at-risk prediction: a comparison of alternative strategies. *Journal of Financial Econometrics* 4, 53–89.
- Manganelli, S. & R. Engle (2001). Value-at-risk models in finance. Working paper, European Central Bank.
- Theodossiou, P. (1998). Financial data and the skewed generalized t distribution. *Management Science* 44, 1650–1661.
- Theodossiou, P. (2000). Skewed generalized error distribution of financial assets and option pricing. SSRN working paper 219679.
- Wilhelmsson, A. (2009). Value-at-risk with time-varying variance, skewness and kurtosis – the NIG-ACD model. *Econometrics Journal* 12, 82–104.
- Zumbach, G. (2007). A gentle introduction to the Risk Metrics 2006 methodology. New York, RiskMetrics group.

# Higher order conditional moment dynamics and forecasting value-at-risk

Grigory Franguridi

*New Economic School, Moscow, Russia*

We empirically investigate the possibilities for enhancing value-at-risk predictions by explicit modelling conditional higher order moment dynamics of financial returns. Using one-day-ahead VaR forecasts for 5 highly liquid constituents of the S&P500 index from different industrial sectors, we compare performances of the benchmark GARCH model with skewed generalized Student's innovations with a set of models allowing for time-varying asymmetry and kurtosis such as ARCD-type models with normal inverse gaussian and skewed generalized Student's errors. As predictive accuracy tests we exploit both the scoring rules for left tail forecasts and likelihood-ratio tests for correct (un)conditional quantile forecasts. We also propose a parsimonious ARCD model with the skewed generalized error distribution for innovations, asymmetric power ARCH for volatility and autoregressive dynamics for skewness and kurtosis related parameters which is shown to perform not worse than the aforementioned models in terms of VaR prediction accuracy, while being computationally less demanding.

*Keywords: value-at-risk, conditional distribution, skewness, kurtosis, financial returns*

*JEL Classification: C22, C53, C58*