# Applied seminar “Quantitative Analysis”

# Colored Noise in Financial Time Series

**Pavlova Anna**

Проектно-учебная лаборатория «Искусственный интеллект в математических финансах»

https://cs.hse.ru/iai/aimf/

Броуновская частица — это объект, подверженный случайным флуктуациям за счет столкновений с молекулами среды. Броуновское движение описывает траекторию движения этой частицы. Классическая модель с белым шумом:

$$\frac{dv}{dt} = -\gamma v + \xi(t),$$

где $v(t)$ − скорость частицы, $\gamma$ − коэффициент вязкости, $\xi(t)$ − **белый шум**.

В идеализированном броуновском движении предполагается, что случайные удары молекул жидкости (или газа) по частице независимы во времени. Это означает, что сила случайных флуктуаций в момент времени $t$ **никак не зависит от того, какие силы действовали на частицу в предыдущие моменты времени**. Однако в реальных физических системах, особенно в вязких средах, ситуация сложнее. Действительно, молекулы среды имеют собственную структуру и динамику, и их влияние на броуновскую частицу не является мгновенным и независимым. Если молекулы толкнули частицу в одном направлении, то следующие удары с большей вероятностью будут направлены в том же направлении в течение некоторого времени. Это приводит к **коррелированным шумам**, которые создают эффект памяти.

В финансовых рынках аналогом вязкой среды можно считать рыночную ликвидность и устойчивые тренды, которые создают эффект памяти в динамике цен. В идеализированной модели случайного блуждания предполагается, что изменения цен активов независимы во времени, аналогично независимым ударам молекул в классическом броуновском движении. Однако в реальности рыночные процессы обладают инерцией и временной корреляцией, что делает их похожими на движение частицы в вязкой среде. Таким образом, концепция вязкости в финансах проявляется через запаздывания в реакции рынка, долгосрочные корреляции и влияние крупных участников, создающих эффект памяти в финансовых временных рядах.

От статистических характеристик шума, действующего на конкретно рассматриваемую систему, в значительной мере соответствует ее стохастическая динамика в целом. Известно, что **случайный процесс** задан, если известен полный набор многомерных функций распределения вероятностей. Далее рассматриваются только **стационарные в широком смысле** внешние шумы, т.е. используется корреляционная теория, когда принимаются во внимание лишь одно- и двумерные распределения этих шумов. В этом случае на первый план выходят среднее и корреляционная функция шума. В соответствии с **теоремой Винера – Хинчина**, корреляционная функция однозначно может быть задана посредством спектральной плотности шума, вид которой может кардинально повлиять на качественные свойства стохастической системы в целом. Это можно показать на простом, но важном примере.

Пусть дан **стохастический риманов интеграл**

$$\rho(t) = \int_0^t X(s)ds$$

от некоторого стационарного в широком смысле процесса $X(t)$ с нулевым средним. Предположим, что $t$ достаточно большое. Известно, что полученный таким образом процесс не является стационарным. Введем автокорреляционную функцию

$$R_X(\tau) = E[X(t)X(t + \tau)]$$

и продемонстрируем это свойство, посчитав дисперсию этого процесса.

$$\sigma_\rho^2 = \langle\rho^2(t)\rangle = E[(\int_0^t X(s)ds)^2] = E[\int_0^t X(s)ds\int_0^t X(u)du] = \int_0^t\int_0^t E[X(s)X(u)]dsdu.$$

Математическое ожидание интеграла совпадает с интегралом от математического ожидания благодаря линейности ожидания.

Поскольку процесс X(t) стационарный, его автокорреляционная функция

$$R_X(\tau) = E[X(t)X(t + \tau)]$$

зависит только от разности времени $\tau = s - u$. Тогда

$$E[X(s)X(u)] = R_X(s - u).$$

Введем новые переменные: $\tau = s - u$, $v = u$. Тогда $s = \tau + v$, $u = v$. Поскольку изменились переменные, нужно преобразовать интеграл с помощью Якобиана. Якобиан используется при замене переменных в интеграле, чтобы правильно учитывать изменение объема области интегрирования:

$$|J| = |\frac{\partial(s,u)}{\partial(\tau, v)}| = 1.$$

Таким образом, $dsdu = |J|d\tau dv = d\tau dv$.

Теперь определим новые пределы интегрирования. Исходные пределы для переменных: $0 < s < t$, $0 < u < t$. После замены $\tau = s - u$, $v = u$. Максимальное значение для $\tau$ при $s = t$, $u = 0 \to \tau = t$. Минимальное значение для $\tau$ при $s = 0$, $u = t \to \tau = - t$. Тогда $\tau \in [- t, t]$. Из $s = \tau + v$ и $s < t \to v \le t - \tau$. Из $v = u \ge 0 \to v \ge 0$. Из $\tau = s - u \ge - v \to v \ge - \tau$ (актуально при $\tau < 0$). Итоговые пределы: $v \in [max(0, - \tau), min(t, t - \tau)]$.

Исходный интеграл преобразуется в

$$\int_{-t}^{t}\int_{max(0,-\tau)}^{min(t,t-\tau)} R_X(\tau)dvd\tau.$$

Интегрируем по $v$. Для фиксированного предела $\tau$, пределы $v$ дают:

$$\int_{max(0,-\tau)}^{min(t,t-\tau)} dv = min(t, t - \tau) - max(0, - \tau).$$

Упростим пределы.

При $\tau \ge 0$: $max(- \tau, 0) = 0$; $min(t, t - \tau) = t - \tau$. Имеем $min(t, t - \tau) - max(0, - \tau) = t - \tau$.

При $\tau < 0$: $max(-\tau, 0) = -\tau$; $min(t, t - \tau) = t$.

Имеем $min(t, t - \tau) - max(0, -\tau) = t + \tau$.

Объединим два случая и получим $min(t, t - \tau) - max(0, -\tau) = t - |\tau|$.

Интеграл принимает вид

$$\int_{-t}^{t} (t - |\tau|) R_X(\tau) d\tau.$$

Для больших $t$, множитель $t - |\tau| \approx t$ при $|\tau| \ll t$. Рассмотрение больших значений времени t в анализе дисперсии стохастического интеграла связано с асимптотическим поведением системы, то есть с её долговременными свойствами. Это упрощает анализ и характерно для многих физических систем в долговременной перспективе. Тогда имеем

$$\sigma_\rho^2 = t \int_{-\infty}^{+\infty} R_X(\tau) d\tau + A(t).$$

Обратите внимание на переход пределов интегрирования с $-t, t$ к $-\infty, +\infty$. Мы это делаем, так как t достаточно большое.

Имеем:

$$\sigma_\rho^2 = tD_\rho + A(t),$$

где

$$D_\rho = \int_{-\infty}^{+\infty} R_X(\tau),$$

а $A(t)$ - это остаточный член, который состоит из константы и интеграла, затухающего при больших $t$.

Таким образом, дисперсия интеграла $\rho(t)$ при больших t начинает возрастать пропорционально времени! Пропорциональная зависимость $\sigma_\rho^2$ от времени обычно называется **законом диффузии**, а коэффициент пропорциональности $D_\rho$ — **коэффициентом диффузии.**

Коэффициент диффузии — это ключевая характеристика случайного процесса, которая определяет, с какой скоростью растет дисперсия (вариабельность) процесса со временем. Другими словами, он показывает, насколько быстро "разлетаются" траектории процесса от своего среднего значения. Это важно, потому что именно через дисперсию можно оценить риск, нестабильность и степень неопределенности в моделируемых системах. В **финансовых моделях** коэффициент диффузии используется для описания волатильности активов. Например, в модели Блэка-Шоулза цена акции описывается стохастическим уравнением $dS(t) = \mu S(t)dt + \sigma S(t)dW(t)$. Здесь $\sigma$ - волатильность, играющая роль коэффициента диффузии, который определяет степень случайных колебаний цен.

В соответствии с **теоремой Винера Хинчина**, спектральная плотность процесса представляется интегралом Фурье:

$$S_X(\omega) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} R_X(\tau) e^{-i\omega\tau} d\tau.$$

Эта теорема связывает корреляционную функцию случайного процесса с его спектральной плотностью. Грубо говоря, она говорит: если у вас есть случайный сигнал (процесс), то его частотное содержание (спектральная плотность) можно получить через корреляцию значений сигнала в разные моменты времени. **Корреляционная функция** случайного процесса показывает, насколько похожи значения сигнала в разные моменты времени. **Спектральная плотность** показывает, какие частоты содержатся в сигнале. Теорема Винера–Хинчина утверждает, что спектральная плотность – это просто **преобразование Фурье** корреляционной функции.
Из соотношений полученной дисперсии и интеграла Фурье следует, что $D_\rho = 2\pi S_X(0)$, т.е. коэффициент диффузии пропорционален спектральной плотности взятой на нулевой частоте. Таким образом, такая важная математическая характеристика, как коэффициент диффузии $D_\rho$, зависит лишь от значения спектральной плотности $S_X(0)$ источников флуктуаций, обуславливающих диффузионный процесс. Это сразу наталкивает на мысль, что величина $S_X(0)$ может служить одним из параметров, характеризующим определенные классы шумов. Остановимся на этом более подробно.

Когда мы анализируем временной ряд, например, цену акции, мы рассматриваем последовательность значений во времени. **Частота** в этом контексте означает скорость изменений во времени. Если цена меняется медленно, это низкочастотный процесс; если быстро — высокочастотный. На фондовом рынке цена меняется медленно в течение дней, недель, месяцев. Это **низкочастотный** процесс: большие движения происходят редко. Спектр такого процесса концентрируется около малых частот. На криптовалютном рынке цены скачут быстрее. Это **высокочастотный** процесс: в единицу времени много изменений. Спектр такого процесса смещен в сторону высоких частот. **Спектральная плотность мощности (СПМ)** случайного процесса — это мера того, какие частоты вносят наибольший вклад в дисперсию временного ряда. Другими словами, СПМ показывает, на каких временных масштабах происходят основные колебания в данных. В финансах спектральный анализ помогает понимать природу рыночных движений, например: есть ли долгосрочные тренды? (**низкие** частоты) и преобладают ли краткосрочные шумовые колебания? (**высокие** частоты). Если спектр содержит низкие частоты, то цены изменяются медленно (типично для долгосрочных инвестиций). Если спектр доминирует на высоких частотах, то рынок волатильный.

Наиболее часто используемой является математическая модель источников в виде гауссовского белого шума $X(t) = w(t)$ с постоянной спектральной плотностью $S_w(0) = S_w(\omega) = const$, с т.е. такого стационарного случайного процесса, у которого $w(t_1)$ и $w(t_2)$ некоррелированы, если $t_1 \neq t_2$ и $\langle w(t) \rangle = 0$. $w(t)$ называют белым шумом по аналогии с белым светом, имеющим в видимой области спектра почти равномерное распределение. **Белый шум** — это случайный процесс, у которого **все частоты имеют одинаковую мощность**. СПМ белого шума постоянна - это означает, что колебания вносят одинаковый вклад на **всех временных масштабах**. В высокочастотной торговле (HFT) шумовые движения цены могут быть похожи на белый шум (случайные скачки без структуры). В моделировании финансовых временных рядов ошибки в регрессиях часто моделируют белым шумом, если нет корреляции между наблюдениями.

Строго говоря, $w(t)$ можно представить только в виде некоторого **обобщенного случайного процесса**. Белый шум нельзя представить как обычный случайный процесс, так как его дисперсия формально бесконечна, а его корреляционная функция содержит дельта-функцию. Однако его можно корректно определить в пространстве **обобщённых функций** (распределений), где он действует не на отдельные точки, а на целые функции. Остановимся на этом более подробно.

Основные свойства белого шума - нулевое среднее и дельта-коррелированность (отсутствие корреляции во времени). **Его АКФ:**

$$R_w(t' - t'') = E[X(t')X(t'')] = \sigma^2\delta(t' - t'').$$

При этом $\sigma^2$определяется как **интенсивность белого шума**. **Дельта-функция** - обобщенная функция, которая формально не принимает конечных значений в традиционном смысле. Однако говорят, что она бесконечно велика в точке t' = t'' и равна нулю везде, кроме этой точки. Она обладает следующими свойствами: Все энергия сосредоточена в одной точке (t' = t''), а для всех других t', t'' корреляция исчезает. Это идеализированное представление: на практике шум может иметь небольшую корреляцию на очень коротких временных масштабах. Ее интеграл по всей оси дает 1. Так, если бы белый шум был обычной случайной функцией, то у него должна была бы существовать дисперсия $E[X(t)^2]$. Но из свойства

$$E[X(t')X(t'')] = \sigma^2\delta(t' - t'')$$

следует, что при $t' = t''$: $E[X(t)^2] = \sigma^2\delta(0)$. Однако формально $\delta(0)$ **бесконечно велика**! Это значит, что белый шум не может существовать как обычный процесс: его значения в конкретных точках времени просто **не определены в стандартном смысле**.

Поскольку в наших задачах процесс w(t) будет содержаться в правой части стохастических дифференциальных уравнений, будем использовать следующую (традиционную) интерпретацию. **Любое дифференциальное уравнение есть предельная запись некоего допредельного (дискретного) уравнения.** Если w(t) содержится в правой части дифференциального уравнения, то в его дискретном эквиваленте будет содержаться выражение $w(t_k)\Delta t_k$, где k - целое, а $\Delta t_k = t_{k+1} - t_k$.

Стохастические дифференциальные уравнения (**СДУ**) — это уравнения, которые включают случайные компоненты, такие как белый шум. Здесь речь идет о том, что белый шум w(t) будет присутствовать в правой части этих уравнений, обычно как источник случайных колебаний или помех. Дифференциальное уравнение — это **непрерывное приближение к дискретной модели**. То есть, любое дифференциальное уравнение можно рассматривать как предел, к которому стремится некоторое дискретное уравнение, когда шаг дискретизации стремится к нулю. Это означает, что дифференциальное уравнение описывает поведение системы в непрерывном времени, но может быть представлено как результат перехода от дискретной модели (с конечным шагом времени). Выражение $w(t_k)\Delta t_k$ — это дискретизация белого шума. В

дискретном случае белый шум умножается на шаг времени, что позволяет моделировать его воздействие на систему на каждом шаге.
Обычно выражение $w(t_k)\Delta t_k$ заменяется на $\Delta W_k = W(t_{k+1}) - W(t_k)$, где $W(t)$ - некоторый гауссовский случайный процесс, если положить $\rho(t) = W(t)$ и принять условия $D_w = 2\pi S_W(0) = 2\pi S_W(\omega) = const$. Этот процесс принято называть винеровским процессом. Иногда его называют процессом броуновского движения.

Внимание на предыдущую задачу. Все, что было сказано до этого, может применяться к $X(t) = w(t)$, так как это так же стационарный процесс с нулевым средним Т.е. мы можем сказать, что ранее мы рассмотрели, в частности, именно белый шум. Интеграл от белого шума и является винеровским процессом. Из предыдущей задачи его среднее равно нулю. Его дисперсия, как отмечалось, $\sigma_W^2 \approx tD_W$. При этом

$$D_W = \int_{-\infty}^{+\infty} R_w(\tau) = \int_{-\infty}^{+\infty} \sigma^2 \delta(t' - t'') = \sigma^2 \int_{-\infty}^{+\infty} \delta(t' - t'').$$

Мы знаем, что интеграл дельта-функции по всей оси равен 1! Примем в качестве нормирования $\sigma^2 = 1$. Тогда дисперсия винеровского процесса равна просто $t$! Таким образом, благодаря свойствам белого шума **мы доказали, что дисперсия винеровского процесса растет пропорционально времени.**

Винеровский процесс играет основополагающую роль в современной математической теории случайных процессов. В частности, на его основе построен класс марковских диффузионных процессов, когда решение практических задач сводится к решению диффузионного уравнения в частных производных, известного как **уравнение Фоккера – Планка** (Фоккера – Планка – Колмогорова, Эйнштейна – Фоккера) для условной плотности вероятности перехода в некоторое состояние, при условии, что в начальный момент времени процесс находился в известном состоянии. Уравнение Фоккера-Планка (УФП) играет ключевую роль в описании эволюции вероятностного распределения случайного процесса. Оно особенно важно при моделировании динамики систем, подверженных стохастическим флуктуациям, например, в физике, биологии и финансах.

Вообще говоря, понятие белого шума вызывает **серьезные математические проблемы**, поскольку реализации (выборочные функции) белого шума не являются дифференцируемыми. Белый шум не является обычной функцией времени, а представляет собой обобщенный случайный процесс (в терминах теории обобщённых функций). Его проблема в том, что его реализации не дифференцируемы и даже не являются функциями в привычном смысле. Его автокорреляционная функция представляет собой дельта-функцию, что означает бесконечную мощность на любой частоте, что невозможно в реальных системах. В реальной практике дельта-функция, используемая в определении белого шума, заменяется на некоторую приближенную симметричную функцию $\Delta(\tau)$ с конечной шириной $2\tau_\Delta$, обладающую свойством нормировки $\int_{-\infty}^{+\infty} \Delta(\tau)d\tau = 1$. Величина $\tau_\Delta$ называется временем корреляции случайного процесса с АКФ $R_\Delta(\tau) = \sigma^2\Delta(\tau)$. Оно определяет характерный масштаб затухания корреляционной функции $R_\Delta(\tau)$. Если $\tau_\Delta \to 0$, то случайный процесс превращается в белый шум, поскольку его автокорреляционная функция приближается к дельта-функции.

Заметим, что во многих случаях для других стационарных процессов $X(t)$ также можно ввести конечное время корреляции $\tau_X$, которое является удобным качественным параметром. Это время характеризует интервал, на котором корреляционную функцию $R_X(\tau)$ еще можно считать отличной от нуля и значения случайного процесса на этом интервале остаются все еще статистически зависимыми. Если за интервал времени, равный времени корреляции, изменением состояния динамической системы или статистического ансамбля можно пренебречь, то модель белого шума, когда $\tau_X = 0$, является удобным приближением для решения конкретных задач. Однако, например, в статистической физике и особенно в статистической радиофизике, где почти всегда необходимо учитывать инерционность динамических систем, приходится рассматривать коррелированные процессы, для которых $\tau_X \neq 0$. В этом случае в силу соотношения неопределенности, свойственного преобразованию Фурье, спектр становится неравномерным. В современной физической литературе подобные шумы обычно называют **окрашенными**.

Простейшим и одним из наиболее популярных примеров окрашенного шума является известный **процесс Орнштейна – Уленбека**. Таким процессом описывается, например, скорость свободной броуновской частицы. Процесс Орнштейна – Уленбека $X(t)$ задается стохастическим дифференциальным уравнением:

$$dX_t = \theta(\mu - X_t)dt + \sigma dW_t,$$

где: $X_t$ - случайный процесс, $\theta > 0$ - коэффициент возврата к среднему (частота затухания), $\mu -$ среднее значение, $\sigma$ - интенсивность шума, $W_t$ - стандартный винеровский процесс (случайное броуновское движение). Это уравнение описывает процесс, который в среднем стремится к среднему, но подвержен случайным флуктуациям.

В отличие от белого шума, процесс Орнштейна – Уленбека обладает **конечным временем корреляции**. Его **автокорреляционная функция** (АКФ) имеет вид $R_X(\tau) = \frac{\sigma^2}{2\theta}e^{-\theta|\tau|}$. АКФ **затухает экспоненциально** с характерным временем $\tau = \frac{1}{\theta}$. Это означает, что значения $X(t)$ и $X(t + \tau)$ остаются взаимосвязанными в течение времени $\tau$. Так, процесс Орнштейна – Уленбека имеет экспоненциально спадающую корреляционную функцию. Это значит, что за время $\tau$ корреляция падает в $e$ раз. Чем больше $\theta$, тем быстрее процесс теряет память. Для белого шума АКФ была бы дельта-функцией, но здесь мы видим, что шум имеет структуру и "окрашен".

Как процесс Орнштейна – Уленбека связан с **броуновским движением**? Скорость броуновской частицы в жидкости подчиняется уравнению Ленарда – Больцмана. Если разделить на массу, это уравнение принимает вид, который полностью совпадает с процессом Орнштейна – Уленбека. ОУ-процесс можно понимать как броуновское движение **с трением**. Без возвратного члена $\theta(\mu - X_t)dt$ процесс был бы обычным броуновским движением.

ОУ-процесс используется для моделирования процентных ставок (модель Васичека):

$$dr_t = a(b - r_t)dt + \sigma dW_t,$$

где $r_t$ - процентная ставка, $a$ - скорость возврата к среднему уровню $b$, а σ - волатильность.

Мы знаем, что дисперсия случайного процесса равна его автокорреляционной функции при лаге τ=0: $R_X(0) = \frac{\sigma^2}{2\theta}$, так как это значение показывает, насколько процесс коррелирует сам с собой в тот же момент времени. Дисперсия ОУ-процесса **обратно пропорциональна** коэффициенту возврата θ. Чем **больше** θ, тем **меньше** разброс значений процесса. Чем **выше** интенсивность шума σ, тем **больше** дисперсия.

Отметим, что процесс Орнштейна – Уленбека является единственным случайным процессом, который является одновременно гауссовским, марковским и стационарным (теорема Дуба).

Рассмотрим модель CIR - модификацию модели Васичека:

$$dr_t = a(b - r_t)dt + \sigma\sqrt{r_t}dW_t.$$

Если **заменить** $dW_t$ на приращения процесса, генерирующего красный шум (назовём его $dZ_t$), уравнение модели CIR примет вид:

$$dr_t = a(b - r_t)dt + \sigma\sqrt{r_t}dW_t$$

Приращения $dZ_t$ будут коррелированы во времени, что означает, что случайные шоки не являются независимыми, а имеют тенденцию «сохраниться» на некоторый промежуток времени. Модель перестанет быть марковской, поскольку будущее распределение $r_{t+1}$ теперь зависит не только от текущего значения $r_t$, но и от прошлых значений шума $dZ_t$. Это усложняет теоретический анализ, но может сделать модель более реалистичной для описания финансовых процессов, где наблюдаются эффекты памяти. Шоковые воздействия будут иметь длительное влияние — после удара значение $r_t$ может дольше отклоняться от среднего уровня $b$, что приводит к эффекту кластеризации волатильности или более «задержанным» реакциям.

В чем состоят **преимущества** замены $dW_t$ в CIR на красный шум? Во многих финансовых временных рядах наблюдаются эффекты долгосрочной корреляции (например, кластеризация волатильности). Модель с красным шумом может лучше описывать такие эффекты, поскольку шок имеет память и его влияние затухает постепенно. В реальных экономических системах воздействие внешних факторов (новостей, шоков) часто не является мгновенным и независимым, а имеет тенденцию к затуханию с течением времени. Модель с красным шумом учитывает этот факт, что может повысить точность прогнозов и оценок риска. Использование коррелированного шума позволяет варьировать скорость затухания корреляции (например, параметр, аналогичный θ в ОУ-процессе, может задавать характерное время памяти шума). Это дает возможность адаптировать модель под особенности конкретного финансового временного ряда. Так же в некоторых ситуациях, например, на рынках с высокой ликвидностью или в периоды кризисов, шоки могут оказывать длительное влияние. Модель с красным шумом может отразить такое поведение, в отличие от классической модели CIR с белым шумом.

Подходы с глубоким обучением и использование нейронных сетей получают все большее внимание, в том числе и при решении финансовых и экономических проблем. В этом исследовании мы использовали многослойную модель LSTM для прогнозирования точек бифуркации (точки смена дисперсии, среднего, тренда в данных). По метрикам Accuracy, Balanced Accuracy, Recall и Precision мы видим, что модель с цветным шумом демонстрирует большую предиктивность данных. Эти метрики используются для оценки качества моделей машинного обучения, особенно в задачах классификации. Давайте разберём каждую из них.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN}$$

**Accuracy** показывает, какую долю предсказаний модель сделала правильно. Допустим, у нас есть модель для предсказания дефолта заемщиков. Если Accuracy = 95%, это означает, что в 95% случаев модель верно определила, выплатит ли клиент кредит или нет. Однако если классы сильно несбалансированы (например, 98% клиентов исправно платят, а 2% — нет), высокая точность может быть обманчивой, так как модель может просто предсказывать "не дефолт" в 100% случаев и всё равно иметь высокое значение Accuracy.

$$Balanced\ Accuracy = \frac{1}{2}\left(\frac{TP}{TP+FN} + \frac{TN}{TN+FP}\right)$$

**Balanced Accuracy** полезна при дисбалансе классов. Она усредняет долю правильных предсказаний для каждого класса, предотвращая ситуацию, когда модель слишком ориентируется на доминирующий класс. Если мы разрабатываем модель обнаружения мошеннических транзакций, мошеннические операции (Fraud) составляют всего 1% данных, а обычные транзакции (Non-Fraud) — 99%. Простая accuracy может быть высокой, даже если модель почти никогда не предсказывает "Fraud". Balanced Accuracy исправляет этот дисбаланс, давая равный вес каждому классу.

$$Precision = \frac{TP}{TP+FP}$$

**Precision** показывает, какая доля объектов, предсказанных как положительные, действительно являются положительными. Полезно в задачах, где ложные срабатывания (FP) критичны. Допустим, модель предсказывает, что акция компании X скоро сильно вырастет, и инвестор принимает решение о покупке. Высокий precision означает, что большинство таких прогнозов окажутся верными. Если Precision низкий, модель часто ошибается, предлагая купить акции, которые не растут.

$$Recall = \frac{TP}{TP+FN}$$

**Recall** показывает, насколько хорошо модель находит все истинно положительные объекты. Высокий Recall важен в ситуациях, где пропуск (FN) критичен. В случае обнаружения мошенничества Recall важен, так как мы хотим находить как можно больше мошеннических операций. Если у модели Recall = 80%, это значит, что 80% мошеннических транзакций были корректно идентифицированы, а 20% — пропущены.

Если главная цель модели – не пропускать бифуркации (например, чтобы заранее подготовиться к изменениям ставок), то важнее **Recall**. Если важно минимизировать ложные срабатывания (например, чтобы не принимать ошибочные инвестиционные решения), то приоритетной метрикой будет **Precision**. Если данные несбалансированы (например, бифуркации встречаются редко), лучше смотреть на **Balanced Accuracy**.

Результаты показывают, что модели цветного шума, в частности розового и красного, **превосходят эталонную модель с белым шумом по всем ключевым показателям, демонстрируя высокую эффективность при обнаружении точек перехода.** Эти результаты свидетельствуют о том, что использование цветных шумов может улучшить возможности прогнозирования при работе со стохастическими моделями, аналогичных модели Васичека и CIR, способствуя обнаружению значительных изменений в финансовых временных рядах.

## Вопросы и ответы

**Почему интеграл от дельта-функции равен 1?**

- Дельта-функция Дирака была введена П. Дираком для описания сосредоточенных воздействий. В нестрогой форме она определялась как «странная» функция, равная 0 всюду, кроме одной точки, где она равна бесконечности, причем интеграл от этой функции равен 1. В этом смысле понятие дельта-функции аналогично физическим понятиям точечной массы или точечного заряда. Дельта-функция не может быть определена как классическая функция, но является примером обобщенных функций.
  Для понимания интеграла полезно представить себе некую фигуру на плоскости с единичной площадью, например, треугольник. Если уменьшать основание данного треугольника и увеличивать высоту так, чтобы площадь была неизменной, то в предельном случае мы получим треугольник с малым основанием и очень большой высотой. По предположению его площадь равна

единице, что и показывает интеграл. Вместо треугольника можно без ограничения общности использовать любую фигуру.

**Данные по какой процентной ставке были использованы в исследовании?**

- Была использована US Federal Funds Rate. Это краткосрочная процентная ставка, по которой банки предоставляют ночные (overnight) кредиты друг другу для поддержания резервных требований.

**Чтобы использовать белый шум, нужно взять интеграл от него же, что будет являться винеровским процессом. Значит, чтобы использовать цветной шум, нужно брать интеграл так же от какого-то случайного процесса?**

- Для генерации цветных шумов в Python используется алгоритм Тиммера-Кёнинга (Timmer, J. and Koenig, M.: On generating power law noise. Astron. Astrophys. 300, 707-710 (1995)) и основанная на нем библиотека colorednoise. Этот алгоритм использует быстрое преобразование Фурье (FFT). Сначала задаётся спектральная плотность $S(f) \sim 1/f^{\alpha}$, затем случайно генерируются амплитуды и фазы для каждой частоты. После этого создается комплексный спектр с масштабированием по $S(f)$ и симметрией для получения реального сигнала. Наконец, обратное FFT преобразует спектр в временной ряд с заданными шумовыми характеристиками. Этот метод позволяет быстро и точно моделировать временные ряды с нужной корреляцией.

**На каких данных обучалась модель?**

- В исследовании использованы две модели - многослойная нейросеть LSTM и классификатор Catboost на градиентном бустинге - для прогнозирования точек перехода путем присвоения значений 0 или 1 в зависимости от наличия в них аномалии, выявленной алгоритмом PELT. Далее для каждой точки $y_i$ выбираются предыдущие 100 значений, т.е. $X = \{y_{i-100}, y_{i-99}, ..., y_{i-1}\}$ и считается их дельта. Таким образом, в качестве параметров в модель передается:

$$X = \{y_{i-100} - y_{i-101}, y_{i-99} - y_{i-100}, ..., y_{i-1} - y_{i-2}\}.$$

Получаем набор данных для обучения:

$$Y = \{y_{100}, y_{101}, ..., y_n\}, X = \{X_{100}, X_{101}, ..., X_n\},$$

где

$$X_j = \{\Delta y_{j-100}, \Delta y_{j-99}, ..., \Delta y_{j-1}\}.$$

Для каждого коррелированного шума модель обучается на 80%

сгенерированных данных, а оставшиеся 20% данных используются для теста модели.