# МА процессы

# Стационарные процессы

# Стационарные процессы: план

- Определение стационарного процесса.
- Автоковариационная функция.
- Случайное блуждание и независимые величины.

# Стационарный процесс

Случайный процесс с постоянными характеристиками.

**Стационарность в широком смысле**

Процесс $(y_t)$ стационарен в широком смысле, если для любых $t$ и $k$:

$$\begin{cases} \mathbb{E}(y_t) = \mu \\ \mathrm{Cov}(y_t, y_{t+k}) = \gamma_k \end{cases}$$

**Стационарность в узком смысле**

Процесс $(y_t)$ стационарен в узком смысле, если для любого $k$ закон распределения вектора $(y_t, y_{t+1}, y_{t+2}, \ldots, y_{t+k})$ не зависит от $t$.

# Стационарность: много уравнений

У стационарного процесса $(y_t)$:

$$\mathbb{E}(y_5) = \mathbb{E}(y_7) = \mathbb{E}(y_{100}) = \mathbb{E}(y_{135}) = \ldots = \mu$$

$$\mathrm{Var}(y_5) = \mathrm{Var}(y_7) = \mathrm{Var}(y_{100}) = \mathrm{Var}(y_{135}) = \ldots = \gamma_0$$

$$\mathrm{Cov}(y_5, y_7) = \mathrm{Cov}(y_8, y_{10}) = \mathrm{Cov}(y_8, y_6) = \ldots = \gamma_2$$

$$\mathrm{Cov}(y_1, y_5) = \mathrm{Cov}(y_8, y_{12}) = \mathrm{Cov}(y_8, y_4) = \ldots = \gamma_4$$

# Стационарный процесс: пример

**Независимые наблюдения**

Величины $(y_t)$ независимы и одинаково распределены с конечным ожиданием $\mu_y$ и конечной дисперсией $\sigma_y^2$.

$$\mu_y = \mathbb{E}(y_t)$$

$$\gamma_0 = \mathrm{Cov}(y_t, y_t) = \mathrm{Var}(y_t) = \sigma_y^2.$$

$$\gamma_k = \mathrm{Cov}(y_t, y_{t+k}) = 0, \text{ при } k \geq 1.$$

# Нестационарный процесс: пример

**Случайное блуждание**

$$\begin{cases} y_0 = \mu \\ y_t = y_{t-1} + u_t, \text{ при } t \geq 1 \end{cases},$$

где $u_t$ — белый шум.

В явном виде: $y_t = \mu + u_1 + u_2 + \ldots + u_t$.

$$\mu_y = \mathbb{E}(y_t)$$

$$\gamma_0 = \mathrm{Cov}(y_t, y_t) = \mathrm{Var}(y_t) = \mathrm{Var}(\mu + u_1 + \ldots + u_t) = t\sigma_u^2.$$

$$\gamma_k = \mathrm{Cov}(y_t, y_{t+k}) = \mathrm{Cov}(y_t, y_t + u_{t+1} + \ldots + u_{t+k}) = \mathrm{Var}(y_t).$$

# Случайное блуждание и случайная выборка

тут график!

# Автоковариационная функция

**Определение**

Для стационарного процесса $(y_t)$ функцию $\gamma_k = \mathrm{Cov}(y_t, y_{t+k})$ называют автоковариационной.

**Определение**

Для стационарного процесса $(y_t)$ функцию $\rho_k = \mathrm{Corr}(y_t, y_{t+k})$ называют автокорреляционной.

# Связь функций

$$\rho_k = \mathrm{Corr}(y_t, y_{y+j}) = \frac{\mathrm{Cov}(y_t, y_{y+j})}{\sqrt{\mathrm{Var}(y_t)\,\mathrm{Var}(y_{t+k})}} = \frac{\gamma_k}{\sqrt{\gamma_0 \gamma_0}} = \frac{\gamma_k}{\gamma_0}$$

# Автоковариационная функция — наше всё!

**Теоремка**

Если вектор $(y_t, y_{t+1}, \ldots, y_{t+k})$ имеет многомерное нормальное распределение при любом количестве компонент, то константа $\mu = \mathbb{E}(y_t)$ и функция $\gamma_k = \operatorname{Cov}(y_t, y_{t+k})$ полностью определяют конечномерные распределения случайного процесса $(y_t)$.

# Стационарность: итоги

- Постоянные $\mathbb{E}(y_t)$, $\gamma_k = \mathrm{Cov}(y_t, y_{t+k})$.
- Автоковариационная функция.
- Случайное блуждание нестационарно.
- Случайная выборка стационарна.

# Частные корреляции

# Частные корреляции: план

- Проекция для случайных величин.
- Общее определение.
- Частная автокорреляционная функция.

# Геометрия случайных величин

### Длина и угол

Дисперсия $\mathrm{Var}(R)$ — квадрат длины случайной величины.

Корреляция $\mathrm{Corr}(L, R)$ — косинус угла между случайными величинами.

### Ортогональность

Величины $L$ и $R$ ортогональны, если $\mathrm{Cov}(L, R) = 0$.

# Проекция

**Обозначение**

$Best(L; R_1, R_2, \ldots R_n)$ — линейная комбинация 1 и $R_1$, …, $R_n$, наиболее похожая на $L$.

$\hat{L} = Best(L; R_1, R_2, \ldots R_n)$ если:

- $\hat{L} = \alpha_0 \cdot 1 + \alpha_1 R_1 + \ldots + \alpha_n R_n$;
- Ожидание $\mathbb{E}((L - \hat{L})^2)$ минимально.

# Как найти проекцию?

Хотим найти $\hat{L} = Best(L; R_1, R_2, \ldots R_n)$:

$$\hat{L} = \alpha_0 \cdot 1 + \alpha_1 R_1 + \ldots + \alpha_n R_n.$$

Как найти коэффициенты?

- Минимизация:

$$\mathbb{E}((L - \hat{L})^2) \to \min$$

- Решение системы:

$$\begin{cases} \mathbb{E}(L) = \mathbb{E}(\hat{L}); \\ \operatorname{Cov}(L, R_i) = \operatorname{Cov}(\hat{L}, R_i) \text{ при всех } i; \end{cases}$$

# Частная корреляция

**Определение**

$$\mathrm{pCorr}(U, D; R_1, R_2, \dots, R_n) = \mathrm{Corr}(U^*, D^*),$$

где

$$U^* = U - Best(U; R_1, R_2, \dots, R_n),$$

$$D^* = D - Best(D; R_1, R_2, \dots, R_n).$$

Величины $U^*$ и $D^*$ — это очищенные версии $U$ и $D$.

$$\mathrm{Cov}(U^*, R_i) = 0, \quad \mathrm{Cov}(D^*, R_i) = 0.$$

# Два угла на графике

обычная и частная корреляции

# PACF

**Определение**

Для стационарного процесса $(y_t)$ функцию

$$\varphi_{kk} = \mathrm{pCorr}(y_t, y_{t+k}; y_{t+1}, \ldots, y_{t+k-1}).$$

называют частной автокорреляционной.

# ACF и PACF: интуиция

Для стационарного процесса!

- ACF:

$$\rho_k = \mathrm{Corr}(y_t, y_{t+k}).$$

Общая сила связи $y_t$ и $y_{t+k}$.

- PACF:

$$\varphi_{kk} = \mathrm{pCorr}(y_t, y_{t+k}; y_{t+1}, \ldots, y_{t+k-1}).$$

Сила связи $y_t$ и $y_{t+k}$ при разорванных связях через промежуточные наблюдения.

# Почему двойной индекс?

$$\varphi_{33} = \mathrm{pCorr}(y_t, y_{t+3}; y_{t+1}, y_{t+2}).$$

$$\varphi_{23} = \mathrm{pCorr}(y_t, y_{t+2}; y_{t+1}, y_{t+3}).$$

$$\varphi_{13} = \mathrm{pCorr}(y_t, y_{t+1}; y_{t+2}, y_{t+3}).$$

# Выборочная PACF через остатки

**Корреляция остатков**

$PACF_4$ — выборочная корреляция между остатками $a_t$ и остатками $b_t$.

$a_t$ — остатки из регрессии

$$y_t \text{ на } 1, y_{t-1}, y_{t-2}, y_{t-3}.$$

$b_t$ — остатки из регрессии

$$y_{t-4} \text{ на } 1, y_{t-1}, y_{t-2}, y_{t-3}.$$

# Выборочная PACF через коэффициент

**Оценка коэффициента**

$PACF_4$ — оценка последнего коэффициента в множественной регрессии:

$$\hat{y}_t = \hat{\beta} + \hat{\beta}_1 y_{t-1} + \ldots + \hat{\beta}_4 y_{t-4}, \quad PACF_4 = \hat{\beta}_4.$$

# Выборочная и истинная PACF

- Истинная PACF есть только у стационарного процесса.
- Выборочную PACF можно посчитать у любого процесса.
- По выборочной PACF иногда можно судить о стационарности.
- Оба способа дают состоятельные оценки для стационарного процесса.
- Способ с выборочной корреляцией остатков гарантирует числа из отрезка $[-1; 1]$.

# Частная корреляция: итоги

- Ковариация задаёт геометрию.
- Частная корреляция — корреляция очищенных величин.
- Во временных рядах очищаем два наблюдения от промежуточных.
- Оцениваем частную корреляцию.

# MA процессы

# MA процессы: план

- Определение и запись с лагами.
- Стационарность.
- ACF и PACF.
- Неединственность записи.

# Лаговый оператор

**Определение**

Для процесса $(y_t)$, определённого при $t \in \mathbb{Z}$, лагированным процессом $Ly_t$ называют ту же последовательность величин со сдвинутым индексом,

$$Ly_t = y_{t-1}.$$

$$L^2 y_t = L \cdot L \cdot y_t = L \cdot y_{t-1} = y_{t-2}.$$

$$\Delta y_t = y_t - y_{t-1} = (1 - L) y_t.$$

$$\Delta_{12} y_t = y_t - y_{t-12} = (1 - L^{12}) y_t.$$

# MA процесс

**Определение**

Процесс $(y_t)$, который можно представить в виде

$$y_t = \mu + u_t + \alpha_1 u_{t-1} + \ldots + \alpha_q u_{t-q},$$

где $\alpha_q \neq 0$ и $(u_t)$ — белый шум, называют $MA(q)$ процессом.

MA — Moving Average — скользящее среднее.

Пример $MA(1)$ процесса:

$$y_t = 5 + u_t + 0.3 u_{t-1},$$

где $(u_t)$ — некоторый белый шум.

Нормировка коэффициента при $u_t$ к единице.

# Запись с лагами

**MA с лаговым полиномом**

Процесс $(y_t)$, который можно представить в виде

$$y_t = \mu + P(L)u_t,$$

где $P(L)$ — многочлен степени $q$ от лага $L$ с $P(0) = 1$, а $(u_t)$ — белый шум, называют $MA(q)$ процессом.

Пример $MA(2)$ процесса:

$$y_t = 5 + (1 - 0.2L + 0.3L^2)u_t,$$

где $(u_t)$ — белый шум.

# Стационарность MA

**Теорема**

Любой $MA(q)$ процесс стационарен.

**Доказательство на примере**

$$\mathbb{E}(5 + u_t + 0.6u_{t-1} + 0.2u_{t-2}) = 5$$

$$\begin{multline}\mathrm{Cov}(5 + u_t + 0.6u_{t-1} + 0.2u_{t-2}, \\ 5 + u_{t+k} + 0.6u_{t+k-1} + 0.2u_{t+k-2}) = \gamma_k\end{multline}$$

Ковариация для $(u_t)$ определяется совпадающими индексами.

При изменении $t$ совпадающие индексами пары те же.

# ACF

**Теорема**

У $MA(q)$ процесса теоретическая автокорреляция $\rho_k$ равна нулю при $k > q$

**Доказательство**

Считаем $\gamma_3 = \mathrm{Cov}(y_t, y_{t+3})$ для $MA(2)$:

$$\gamma_3 = \mathrm{Cov}(5+u_t+0.6u_{t-1}+0.2u_{t-2}, 5+u_{t+3}+0.6u_{t+3-1}+0.2u_{t+3-2})$$

Нет совпадающих индексов у белого шума!

Побочный результат: для $MA(q)$ процесса $\rho_q \neq 0$.

# PACF

**Теорема**

У $MA(q)$ процесса теоретическая частная автокорреляция $\varphi_{kk}$ экспоненциально быстро сходится к нулю.

$$|\varphi_{kk}| < b_0 \cdot r^k, \text{ где } r \in (0;1).$$

# ACF и прогнозы

Традиционно $MA(q)$ процесс оценивают предполагая совместную нормальность $(y_t)$.

Из нулевой $\rho_k = 0$ при $k > q$ следует независимость $y_t$ и $y_{t+k}$.

Прогнозы больше, чем на $q$ шагов вперёд, совершенно одинаковые.

$$(y_{T+q+1} \mid \mathcal{F}_T) \sim (y_{T+q+2} \mid \mathcal{F}_T) \sim (y_{T+q+3} \mid \mathcal{F}_T) \sim \ldots$$

# Прогнозы для $MA(2)$

картинка

# А корректно ли определение?

Нюанс: $(y_t)$ — наблюдаемый ряд, $(u_t)$ — единорог.

Машенька: этот $y_t$ — $MA(1)$ процесс.

Вовочка: этот $y_t$ — $MA(2)$ процесс.

Так не бывает!

# Однако!

Машенька: этот $y_t$ — $MA(1)$ процесс с уравнением

$$y_t = 5 + u_t + 0.5u_{t-1}, \quad \sigma_u^2 = 4.$$

Вовочка: этот $y_t$ — $MA(1)$ процесс с уравнением

$$y_t = 5 + u_t + 2u_{t-1}, \quad \sigma_u^2 = 1.$$

Здесь нет противоречия: $(u_t)$ — единорог!

# Противоречия нет

Машенька: этот $y_t$ — $MA(1)$ процесс с уравнением

$$y_t = 5 + \nu_t + 0.5\nu_{t-1}, \quad \sigma^2_\nu = 4.$$

Вовочка: этот $y_t$ — $MA(1)$ процесс с уравнением

$$y_t = 5 + u_t + 2u_{t-1}, \quad \sigma^2_u = 1.$$

Связь $(u_t)$ и $(\nu_t)$:

$$(1 + 0.5L)\nu_t = (1 + 2L)u_t.$$

# MA процессы: итоги

- $MA(q)$ — взвешивание нескольких белых шумов.
- $MA(q)$ — стационарный процесс.
- $ACF$ резко зануляется, $PACF$ стремится к нулю.
- Неединственность записи.

$MA(\infty)$

# $MA(\infty)$: план

- Определение.
- Существование бесконечных сумм.
- Стационарность.

# $MA(\infty)$

**Определение**

Процесс $(y_t)$, который можно представить в виде

$$y_t = \mu + u_t + \alpha_1 u_{t-1} + \alpha_2 u_{t-2} + \ldots,$$

где $(u_t)$ — белый шум, бесконечное количество $\alpha_i \neq 0$ и $\sum_{i=1}^{\infty} \alpha_i^2 < \infty$, называется $MA(\infty)$ процессом.

$MA(\infty)$:

$$y_t = 5 + u_t + 0.5u_{t-1} + 0.5^2 u_{t-2} + 0.5^3 u_{t-3} + \ldots$$

А так нельзя:

$$y_t = 5 + u_t + \frac{1}{\sqrt{2}} u_{t-1} + \frac{1}{\sqrt{3}} u_{t-2} + \frac{1}{\sqrt{4}} u_{t-3} + \ldots$$

# Сходимости

**Теорема**

Если $\sum_{i=0}^{\infty} \alpha_i^2 < \infty$ и $(u_t)$ — стационарный процесс с нулевым ожиданием, то последовательность частичных сумм $y_t^q$ вида

$$y_t^q = \mu + \sum_{i=0}^{q} \alpha_i u_{t-i}$$

сходится при $q \to \infty$ в среднеквадратичном, по вероятности и по распределению.

Нюанс: сходимость взвешенной суммы гарантирована для стационарного $(u_t)$.

**Бонус**

…и получающийся процесс $(y_t)$ стационарен.

# Виды сходимости: $q \to \infty$

**$y_t^q \to y_t$ в среднеквадратичном**

$$\mathbb{E}((y_t - y_t^q)^2) \to 0.$$

**$y_t^q \to y_t$ по вероятности**

$$\mathbb{P}\left(\left|y_t - y_t^q\right| > \varepsilon\right) \to 0 \text{ для любого числа } \varepsilon > 0.$$

**$y_t^q \to y_t$ по распределению**

$$\mathbb{P}(y_t^q \le c) \to \mathbb{P}(y_t \le c)$$

в точках непрерывности $F(c) = \mathbb{P}(y_t \le c)$.

# Теорема Вольда

**Теорема**

Если $(y_t)$ — стационарный процесс, то он представим в виде:

$$y_t = \sum_{i=0}^{\infty} \alpha_i u_{t-i} + r_t,$$

где

- $(u_t)$ — белый шум,
- $\sum \alpha_i^2 < \infty$,
- $r_t$ — линейно предсказуемый случайный процесс,
- $\mathrm{Cov}(u_t, r_t) = 0$.

Ахтунг: deterministic часто ошибочно переводят как последовательность констант.

# Предсказуемый процесс

**Правильное определение**

Процесс $(r_t)$ называется линейно предсказуемым, если

- $(r_t)$ стационарен,
- $r_t = \beta_0 + \beta_1 r_{t-1} + \beta_2 r_{t-2} + \ldots + \beta_p r_{t-p}$.

# $MA(\infty)$: плюсы

- Стационарный процесс.
- Богатая структура корреляций $\rho_k$.
- Практически любой стационарный процесс.

# $MA(\infty)$: проблема

Оценить невозможно: бесконечное число параметров $\alpha_i$.

Решение: введём ограничения на $\alpha_i$!

# $MA(\infty)$: итоги

- Быстро стремящиеся к нулю коэффициенты.
- Стационарный процесс.
- Пока не ясно как оценивать.

## Условие обратимости

# Условие обратимости: план

- Два варианта условия.
- Единственность записи.
- Возможность поймать единорога.

# Помним о проблеме!

Машенька: этот $y_t$ — $MA(1)$ процесс с уравнением

$$y_t = 5 + u_t + 0.5u_{t-1}, \quad \sigma_u^2 = 4.$$

Вовочка: этот $y_t$ — $MA(1)$ процесс с уравнением

$$y_t = 5 + u_t + 2u_{t-1}, \quad \sigma_u^2 = 1.$$

Хотим единственной формы записи для одного $MA(q)$ процесса.

# Лаговый полином $MA(q)$ части

Рассмотрим запись

$$y_t = 5 + u_t + 0.6u_{t-1} + 0.2u_{t-2},$$

представим в виде

$$y_t = 5 + (1 + 0.6L + 0.2L^2)u_t.$$

Лаговый многочлен:

$$P(L) = 1 + 0.6L + 0.2L^2.$$

# Характеристический полином $MA(q)$ части

Рассмотрим запись

$$y_t = 5 + u_t + 0.6u_{t-1} + 0.2u_{t-2},$$

оставим только белый шум

$$u_t + 0.6u_{t-1} + 0.2u_{t-2},$$

подставим геометрическую прогрессию $u_t = \lambda^t$ и сократим

$$\lambda^2 + 0.6\lambda + 0.2.$$

Характеристический многочлен:

$$\phi(\lambda) = \lambda^2 + 0.6\lambda + 0.2.$$

# Связь многочленов

$$y_t = 5 + u_t + 0.6u_{t-1} + 0.2u_{t-2}$$

**Теоремка**

$$P(x) = x^q \cdot \phi(1/x)$$

Ахтунг: путаница в названиях!

$P(L) = 1 + 0.6L + 0.2L^2$ — лаговый многочлен,

$\phi(\lambda) = \lambda^2 + 0.6\lambda + 0.2$ — характеристический.

# Условие обратимости

**Характеристический вариант**

Уравнение $MA(q)$ процесса удовлетворяет условию обратимости, если у характеристического многочлена $\phi(\lambda)$ все корни $|\lambda_i| < 1$.

**Лаговый вариант**

Уравнение $MA(q)$ процесса удовлетворяет условию обратимости, если у лагового многочлена $P(L)$ все корни $|\ell_i| > 1$.

# Пример обратимой записи $MA(1)$

$$y_t = 5 + u_t + 0.5u_{t-1}, \quad \sigma_u^2 = 4.$$

$$\lambda^1 + 0.5 \cdot \lambda^0$$

$$\phi(\lambda) = \lambda + 0.5$$

$$\lambda_1 = -0.5.$$

# Пример необратимой записи $MA(1)$

$$y_t = 5 + u_t + 2u_{t-1}, \quad \sigma_u^2 = 1.$$

$$\lambda^1 + 2 \cdot \lambda^0$$

$$\phi(\lambda) = \lambda + 2$$

$$\lambda_1 = -2.$$

# Нюанс

**Разница**

Стационарность — это свойство самого процесса $(y_t)$.

Обратимость — это свойство записи процесса (уравнения) для $(y_t)$.

Один и тот же процесс $(y_t)$ можно записать с помощью обратимого $MA(q)$ уравнения и с помощью необратимого $MA(q)$ уравнения.

В этих уравнениях будут фигурировать разные единороги $(u_t)$.

# Единственность записи

**Теорема**

$MA(q)$ процесс допускает единственную обратимую запись.

**Теорема**

$MA(\infty)$ процесс допускает единственную запись.

# Попутный бонус

**Теорема**

Если запись $MA(q)$ процесса обратима, то $u_t$ можно представить в

$$u_t = c + \sum_{i=0}^{\infty} \pi_i y_{t-i},$$

где $\sum |\pi_i| < \infty$.

Возможность примерно найти $u_t$ иногда важна для интерпретации.

# Условие обратимости: итоги

- Корни характеристического многочлена $|\lambda_i| < 1$.
- Корни лагового многочлена $|\ell_i| > 1$.
- $MA(\infty)$ имеет единственную запись.
- $MA(q)$ имеет единственную запись при обратимости.
- Возможность восстановить $u_t$.