Семинар 4.
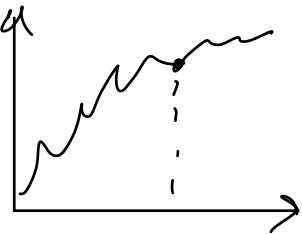## Общий алгоритм.

0) Разбить данные на трейн и тест.
Тест строго после трейна

1) Визуальный анализ
Паттерны (тренд, сезонность, ...)
Понять, какие преобразования

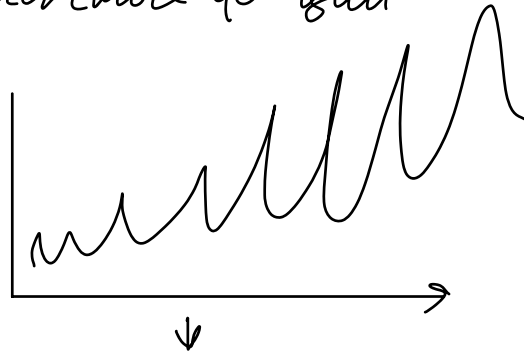2) Преобразовать данные
1) ln
2) любые монотонные фо-ции
3) Box - Cox

$$\hat{y}_t = \begin{cases} \dfrac{y_t^{\lambda} - 1}{\lambda}, & \lambda \neq 0 \\ \ln(y_t), & \lambda = 0 \end{cases}$$

3) Выбор модели

Ручная

Отбираем модели-кандидаты

Автоматическая

Auto-ETS, Auto-ARIMA

Кросс-валидация на трейне

In-sample оценка

1) Качество на трейне

2) Информационные критерии

Выбор лучших моделей

Пост-тестирование

(a) Тестирование автокорреляций в остатках

б) Тесты остатков на нормальность

Остатки похожи на белый шум?

$u_t \sim WN$, $E(u_t) = 0$, $Var(u_t) = \sigma^2$
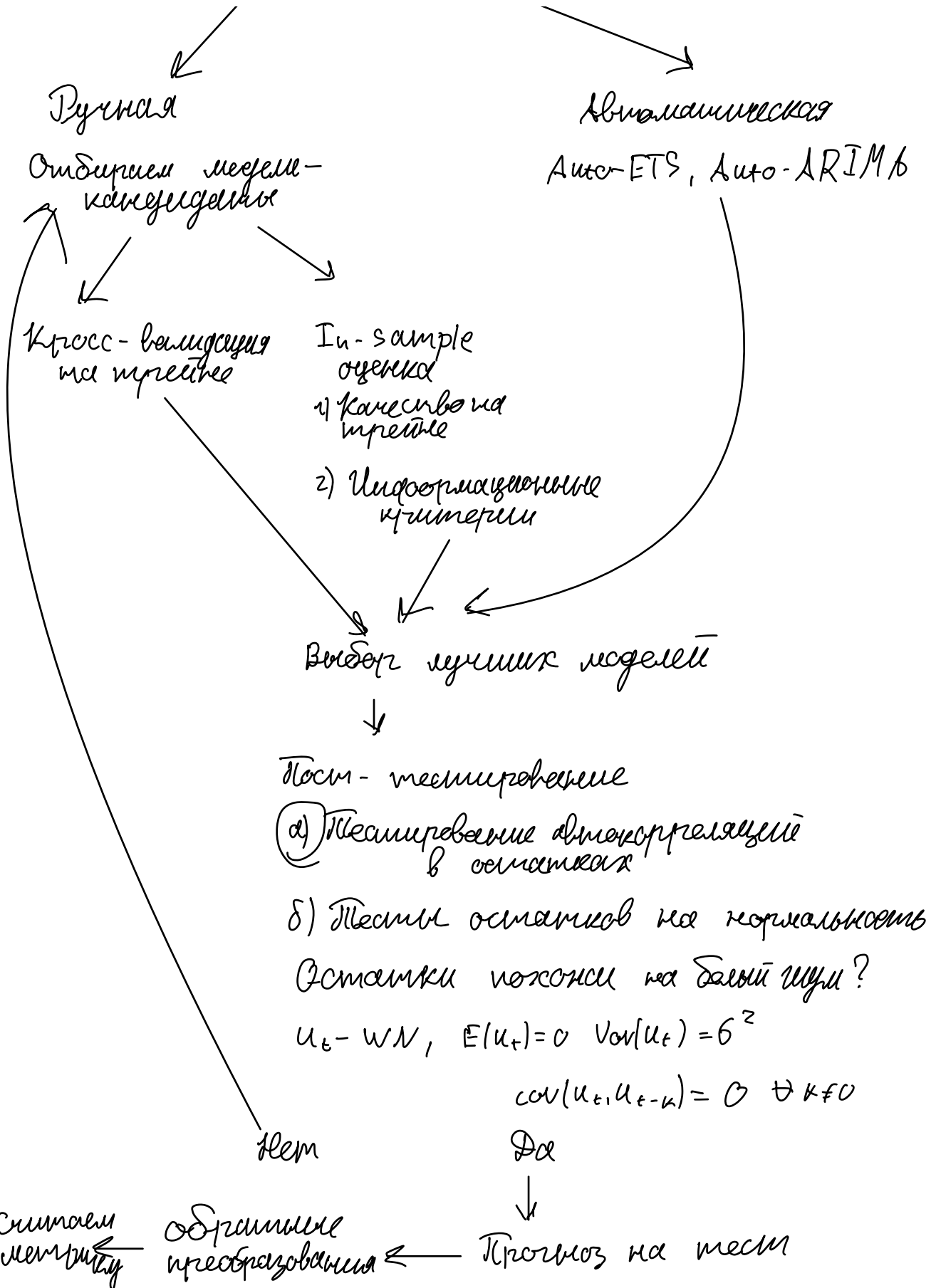
$cov(u_t, u_{t-k}) = 0 \quad \forall k \neq 0$

Нет

Да

Считаем метрику

Обратные преобразования

Прогноз на тест

Тесты на автокорреляцию

$y_1 \dots y_T$

$\hat{y}_1 \dots \hat{y}_T$

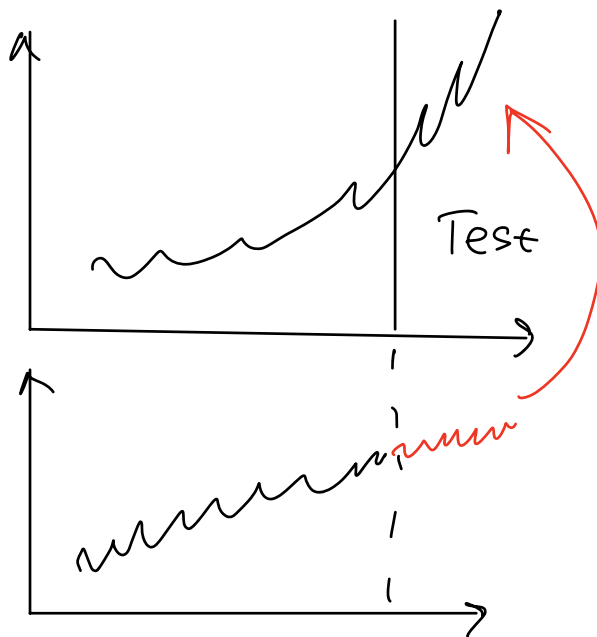$\boxed{e_1 \mid \dots \mid e_T}$

1) Визуальный анализ
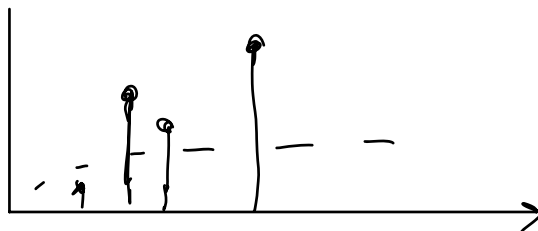
2) portmanteau
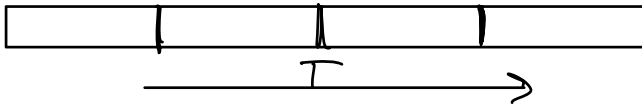
a) Box–Pierce $corr(e_t, e_{t-k})$

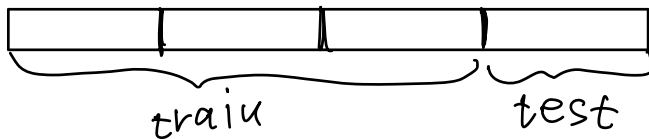$$Q_{BP} = T \sum_{k=1}^{h} \hat{\rho}_k^2 \sim \chi_h^2$$

б) Ljung–Box

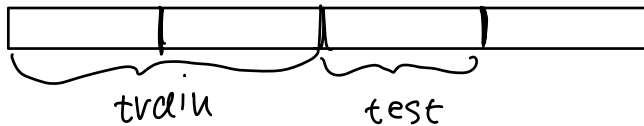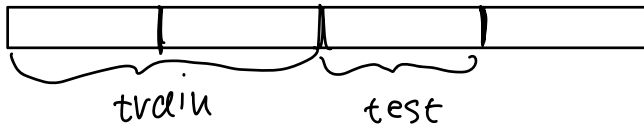$$T(T+2) \sum_{k=1}^{h} \frac{\hat{\rho}_k^2}{T-k} \sim \chi_h^2$$

Test

Кросс-валидация.

1) Expanding window

train   test

train   test

2) Sliding window

train   test

train   test

Информационные критерии

AIC (AICc)   BIC

Akaike IC   Bayesian IC

$$AIC = 2K - 2\ln(L) \qquad BIC = K\ln T - 2\ln(L)$$

кол-во параметров
модели