**Контрольные вопросы**

1. **Что такое временной ряд? Каковы его составляющие?**

*Временной ряд* – последовательность измерений, упорядоченных в неслучайные моменты времени. Анализ временных рядов основывается на предположении, что последовательные значения в файле данных наблюдаются через равные промежутки времени.

1. **Модель временного ряда? Назовите типы моделей временного ряда.**

Каждый уровень ряда *xt* таким образом можно представить в виде некоторой функции *F*(*ft*, *vt*, *εt*), вид этой функции определяет взаимосвязь компонент уровня ряда и указывает на модель ряда:

- Аддитивная *xt* = *ft* + *vt* + *εt*

- Мультипликативная *xt* = *ft* \* *vt* \* *εt*

- Смешанная *xt* = *ft* \* *vt* + *εt*

1. **В каком случае применяют трансформацию данных ряда? Какие виды трансформации данных вам известны?**

*Трансформация данных* — это преобразование данных к определенному представлению, формату или виду, оптимальному с точки зрения *конкретного* метода анализа Для разных задач анализа могут потребоваться разные методы трансформации. *Типичными средствами трансформации данных* являются следующие.

1. **Что такое стационарный временной ряд и коэффициент автокорреляции? Что является его статистической оценкой?**

### Стационарные временные ряды

Большое значение в анализе временных рядов имеют ***стационарные*** временные ряды, вероятностные свойства которых не изменяются во времени.

Простейшей моделью стационарного временного ряда является Гауссовский белый шум, имеющей нормальное распределение *N*(0,1). Временной ряд, основанный на модели белого шума, можно отнести к классу стохастических временных рядов.

1. **Что такое коррелограмма? Какие выводы можно сделать после ее анализа?**

.

Функцию *rτ* называют ***выборочной автокорреляционной функцией***, а ее график – ***коррелограммой***.

1. **Что такое частичная автокорреляция?**

В таком случае используются интегрированные временные ряды – это нестационарный временной ряд, который может быть приведен к стационарному путем взятия последовательной разности ряда *k*-го порядка (такой ряд называется интегрированным *k*-го порядка). Интегрированные временные ряды еще называют разностно-стационарными (DS-ряд).

1. **Какие тесты вы знаете для определения стационарности временного ряда?**

***Расширенный тест Дики-Фуллера*** используется для определения наличия единичного корня. Временной ряд имеет ***единичный корень***, или порядок интеграции один, если его первые разности образуют стационарный ряд. Нулевая гипотеза говорит о наличии единичного корня.

***KPSS тест*** используется для определения стационарности уровня или тренда ряда. Нулевая гипотеза – стационарный тренд (уровень). Значение *p* ниже 0.05 говорит о необходимости взятия разницы (в случае приведения ряда к стационарному).

1. **Что такое единичный корень и как проверить его отсутствие?**

Нестационарный процесс, первые разности которого стационарны, называется интегрированным процессом первого порядка и обозначается *I*(1). Стационарный процесс в этих обозначениях является *I*(0)-рядом. Как известно, при моделировании зависимости между временными рядами на основе регрессионного анализа важно знать, являются ли анализируемые ряды стационарными или нет. В случае стационарности оценки, полученные на основе метода наименьших квадратов, будут «хорошими», т.е. несмещёнными, эффективными и состоятельными. Но если моделируемые временные ряды не стационарны, велика вероятность получить ложную регрессию, особенно, если эти ряды имеют тренды. Отличить ложную регрессию от истинной можно, опираясь на теорию коинтеграции.

*yt = α + βxt + εt*. (2.12)

1. **Какие наиболее часто используемые на практике критерии проверки «наличия-отсутствия» тренда?**

### Проверка наличия тренда

Рассмотрим наиболее часто используемые на практике критерии проверки «наличия-отсутствия» тренда: критерий серий, основанный на медиане выборки, метод Фостера-Стюарта и метод сравнения средних.

1. **В чём заключается сглаживание ряда? Какие методы сглаживания вам известны?**

Сглаживание (выравнивание) ряда - это **построение нового ряда данных, значения которого максимально приближены к значениям последнего ряда, но на графике изображаются более гладкой, ровной линией**. Эта процедура обычно выполняется для того, чтобы выявить основную тенденцию развития явления и продолжить ее на будущее, то есть сделать прогноз.

1. **В чём суть метода экспоненциального скользящего среднего?**

Сущность метода экспоненциального сглаживания. Сущность метода состоит в том, что **динамический ряд сглаживается с помощью взвешенной «скользящей средней», в которой веса подчиняются экспоненциальному закону**. Другими словами, чем дальше от конца временного ряда отстоит точка, для которой вычисляется взвешенная скользящая средняя, тем меньше «участия она принимает» в разработке прогноза.

1. **Как по значениям автокорреляционной функции можно судить о наличии таких компонент ряда как тренд или сезонная компонента? Каким способом ее можно выделить в случае ее наличия?**

Так как коэффициент *ρ*(*τ*) измеряет корреляцию между членами одного и того же ряда, его называют *коэффициентом автокорреляции*, а зависимость *ρ*(*τ*) *автокорреляционной функцией* (ACF).

1. **Приведите линейную авторегрессионную модель p-го порядка? Как определить порядок авторегрессионной модели?**

Если исследуемый процесс *xt* в момент *t* определяется лишь его значениями в предшествующий период *t*-1, то рассматривают ***авторегрессионную модель* 1-*го порядка***:

*xt* = *b*0 + *b*1 *xt*-1 + *εt* (*t*=1, 2, …, *n*).

1. **Приведите модель скользящего среднего q-го порядка? Как определить порядок модели скользящего среднего?**

Параметрическая ***модель скользящего среднего*** (MA) представляется формулой:

,

где 𝜃𝑖 – параметр, 𝜀𝑖 – белый шум, *q* – порядок скользящего среднего.

1. **Что из себя представляют модель ARMA и ARIMA?**

**ARMA** - **модель** временных рядов использующая одновременно и авторегрессию и скользящие среднее. **ARIMA** - нестационарная\* **модель** времяных рядов выраженная в авторегрессии и скользящем среднем. SARIMA - тоже самое что **ARIMA**, только учитывает сезонность. \* Если спросят, то дать определение стационарности.

1. **Какие критерии используются для выбора наиболее информативной модели?**

Алгоритмы отбора информативных признаков могут быть представлены следующими группами: Wrappers (оберточные), Filters (фильтровочные) и Embedded (встроенные в машины). (Я оставлю без точного перевода эти термины в виду размытости их звучания для русскоязычного сообщества — прим. мое.)  
  
Оберточные алгоритмы создают поднаборы, используя поиск в пространстве возможных входных переменных и оценивают полученные поднаборы входов путем обучения полной модели на имеющихся данных. Оберточные алгоритмы могут быть очень дорогими и рискуют переобучать модель. (Если не используется валидационная выборка — прим. мое.)  
  
Фильтровочные алгоритмы похожи на оберточные в том, что они также выполняют поиск поднаборов входных данных, но, вместо запуска полной модели, важность поднабора для выходной переменной оценивается с помощью более простого (фильтровочного) алгоритма.  
  
Встроенные в машины алгоритмы оценивают важность входных признаков с помощью эвристики, заранее заложенной в обучение.

1. **Зачем проверяют остатки ряда на белый шум?**

Это важно по двум основным причинам:

1. **Предсказуемость**: Если ваш временной ряд - белый шум, то по определению он случайный. Вы не можете разумно моделировать это и делать прогнозы.
2. **Модельная диагностика**В идеале серия ошибок из модели прогноза временного ряда должна быть белым шумом.

Диагностика модели является важной областью прогнозирования временных рядов.

1. **Опишите алгоритм исследования временного ряда.**

## Компоненты и временного ряда и аналитические коэффициенты

Основные понятия в статистическом анализе временных рядов:

**Тренд** - компонента, описывающая долгосрочное изменение уровня ряда.

**Сезонность**- компонента, обозначаемая как Q, описывает циклические изменения уровня ряда.

**Ошибка (random noise)** - непрогнозируемая случайная компонента, описывает нерегулярные изменения в данных, необъяснимые другими компонентами.

**Автокорреляция** — статистическая взаимосвязь между последовательностями величин одного ряда. Это один из самых важных коэффициентов в анализе временного ряда. Чтобы посчитать автокорреляцию, используется корреляция между временным рядом и её сдвинутой копией от величины временного сдвига. Сдвиг ряда называется лагом.