**Как отправить задание**

* Создайте гитхаб репозиторий  test\_intro
* Используя git, залейте результаты в ветку test\_alfa
* Пошлите ссылку на свой репозиторий по завершении

**Условия**

* Используйте jupyter notebook и python 3
* Если вы делаете какие-то предположения, подробно изложите их в ноутбуке
* Рекомендуется использовать набор библиотек (matplotlib, numpy, pandas, scipy, statsmodel, seaborn)
* Пользуйтесь Википедией, чтобы познакомиться с незнакомыми терминами
* Допускается свобода действий в каждом задании
* Пожалуйста ограничьтесь 8-10 часами

**Задания**

**1**

* Загрузите дневные котировки 3 акций (Apple, Google, Facebook) из открытых источников за последние 10 лет:
  + <https://www.alphavantage.co/documentation/>
  + <https://blog.quandl.com/api-for-stock-data>
  + <https://www.quora.com/What-are-some-good-APIs-to-get-real-time-stock-quotes>
* Для однозначности используйте цены закрытия

**2**

* Посчитайте daily returns , где i -день, и i + 1 –след. день
* Посчитайте статистики распределения return-ов.
* Для каждой акции:
  + - Оцените, пожалуйста, схожесть полученного распределения дневных returns и одного из известных распределений – например, Gaussian, Student-t с n степенями свободы и т.д. Используйте визуальные и статистические критерии.
    - Распределение доходностей стационарно по времени?
    - Есть ли взаимосвязь между прошлыми и будущими значениями временного ряда для отдельно выбранной акции? Если да, то значима ли она (для 1 акции)?
* Есть ли интересные взаимосвязи между return-ами различных инструментов в один и тот же период по времени? (Предположение: поиск корреляции, Granger causality, PCA decomposition of returns)

**3**

Предположим, каждый день вы получаете предсказание величины

Введем понятие **Profit-and-loss (pnl)** для каждого дня

Пусть имеет распределение standard Gaussian с mean = 0 и std = 1.

**Задача**: Сгенерируйте несколько случайных реализаций , постройте графики и

**4**

Для временного ряда (например, доходностей) коэффициент **Sharpe Ratio** определяется как

Представьте, что каждый день вы делаете случайное предсказание согласно стандартному распределению для конкретной акции. Качество ваших предсказаний можно охарактеризовать с помощью

*Hit rate - это мера точности: то, как часто наш прогноз совпадает по направлению движения акции .*

это случайная величина, имеющая множество реализаций

Постройте scatter plot от hit rate в диапазоне от 0 до 1

**\*5**

Скачайте таблицу данных по адресу:  
<https://drive.google.com/file/d/1B6fA8FAz874hl2vdpkGPND0IU9gd0AAO/view?usp=sharing>  
Каждая строка – экземпляр, имеющий признаки a, b, c, d и класс. Часть таблицы не размечена. Постройте модель классификации на представленных данных и заполните пустые значения (тестовая выборка), сохранив метки тестовой выборки в виде простого списка чисел в ‘Y\_test.csv’ файл. Файл прикрепите в репозиторий. Все рассуждения и процесс построения модели изложите в ноутбуке.