

Введение в статистику

Проект 7. Обязательный

Проходной порог

Кейс, проблема, идея



Кейс

- Компания выбирает стажёра в отдел финансового мониторинга
- Главное требование — повышенная внимательность
- Первый этап отбора на стажировку — тест из нескольких вопросов с вариантами ответа
- В каждом вопросе **только 1 верный ответ**

Компании нужно выбрать:

- Количество вопросов в тесте
- Уровень сложности вопросов

У компании есть 2 образа кандидата:

- Подходящий — тот, кто подходит по уровню внимательности
- Неподходящий — тот, кто склонен допускать слишком много ошибок



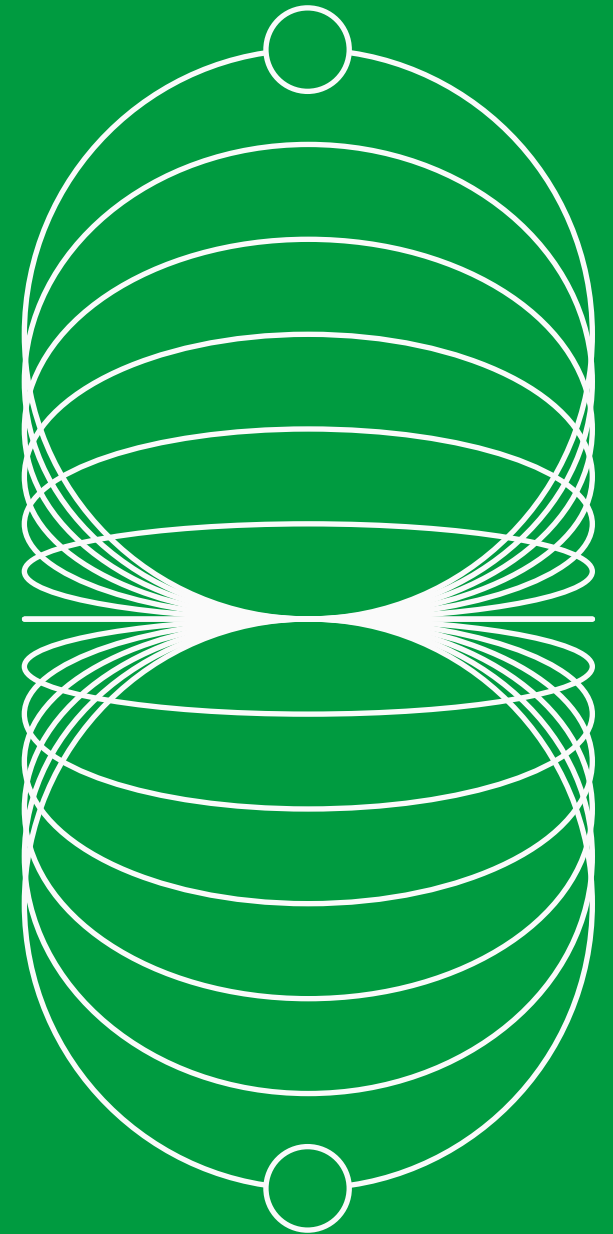
Проблема

- Долгий тест может отпугнуть подходящих кандидатов
- Короткий тест не позволит отличить их от неподходящих
- Простой тест пропустит на второй этап отбора слишком много кандидатов
- Сложный тест пропустит на второй этап отбора слишком мало кандидатов



Идея решения

Принимать решение о количестве вопросов и сложности на основе расчётов, а не просто интуитивно



Роль, задача, результат



Роль

Аналитик отдела продаж, которого HR-отдел попросил помочь



Задача

- 01 Сформулировать предположения модели — то, на что будут опираться выводы
- 02 С помощью модели спрогнозировать итоги первого этапа отбора согласно предложениям HR-отдела, предложить своё количество вопросов в тесте и уровень сложности вопросов
- 03 Подготовить слайды с презентацией результатов



Доступные данные

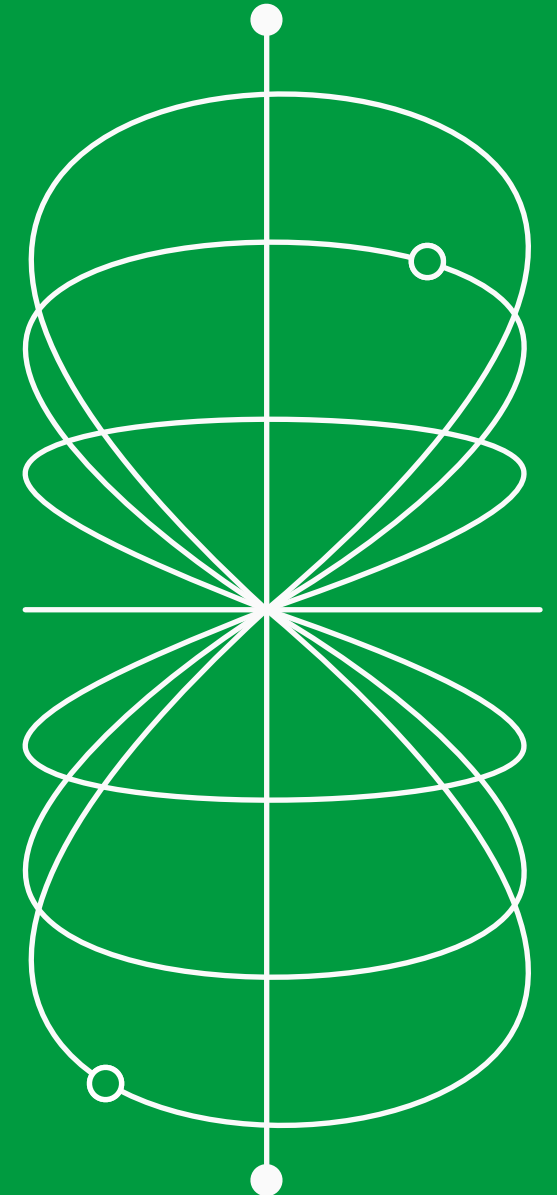
Шанс правильно ответить на вопрос

Сложность заданий	Сильный кандидат	Слабый кандидат
Проще, чем в работе	99%	95%
Сложность заданий	Сильный кандидат	Слабый кандидат
Как в работе	98%	90%
Сложность заданий	Сильный кандидат	Слабый кандидат
Сложнее, чем в работе	70%	20%



Ожидаемый результат

- Презентация на слайдах в формате pdf
- Jupyter Notebook в Google Colab с расчётами



Требования к слайдам

Если слайды или Jupyter Notebook не приложен, решение кейса **оценивается в 0 баллов**

Понятность и внешний вид

- Внешний вид презентации не мешает воспринимать информацию
- Понятно, на какие вопросы отвечает каждый слайд
- Содержимое таблиц, графиков понятно из слайда без необходимости открывать исходный датасет
- Выводы явно сформулированы

МАКСИМУМ 5 БАЛЛОВ

Обоснованность

- Выводы основаны на таблицах, графиках, показателях, полученных из данных
- Таблицы и графики получены скриншотом или картинкой из Jupyter Notebook, поэтому их можно перепроверить
- Выводы явно сформулированы

МАКСИМУМ 5 БАЛЛОВ

Реакция заказчика

- **0 баллов**
Не принимает, ищет другого исполнителя
- **1 балл**
Частично принимает, считает необходимым отдать на доработку текущему исполнителю
- **2 балла**
Принимает, готов пересылать слайды от своего имени, под свою ответственность

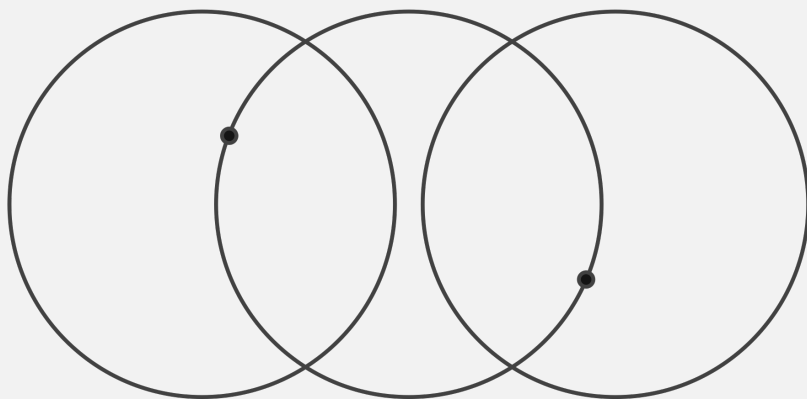
МАКСИМУМ 2 БАЛЛА

Максимальная оценка — **10 баллов**

Бонусные баллы могут поднять основную оценку, но не выше **10 баллов**, даже если в сумме получится 11 или 12 баллов

Синий уровень

Подход к решению



01

Подготовить слайд с предпосылками и постановкой задачи:

- Сложность вопросов «как в работе» — описать, что это значит, про вероятности
- Вероятность прохождения теста сильным кандидатом должна быть не меньше 80%

02

Подготовить слайд, который отвечает на поставленный вопрос в этих предпосылках:

- На одном графике показать, как от порога зависят 2 вероятности — пройти для сильного и пройти для слабого
- Выбрать оптимальный порог и обосновать выбор; если у выбора есть дополнительные предпосылки, изложить их на слайде

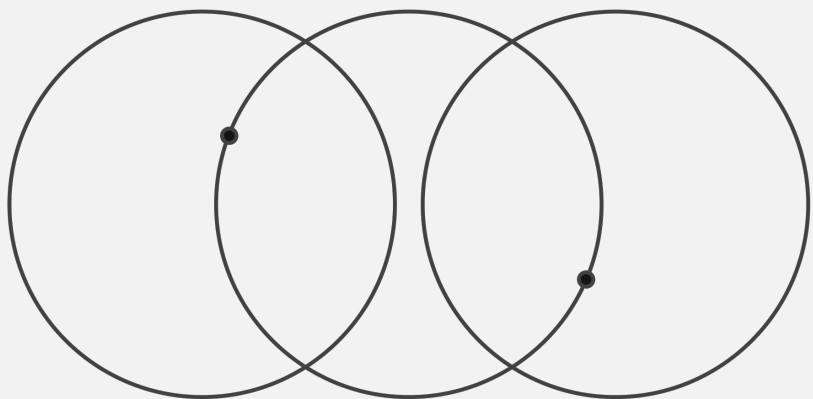
03

Тест состоит из 20 вопросов. Задачи теста:

- Выяснить, с какой вероятностью сильный и слабый кандидат пройдут на второй этап в зависимости от проходного порога
- Определить оптимальный порог

Красный уровень

Подход к решению



01

Подготовить слайды синего уровня

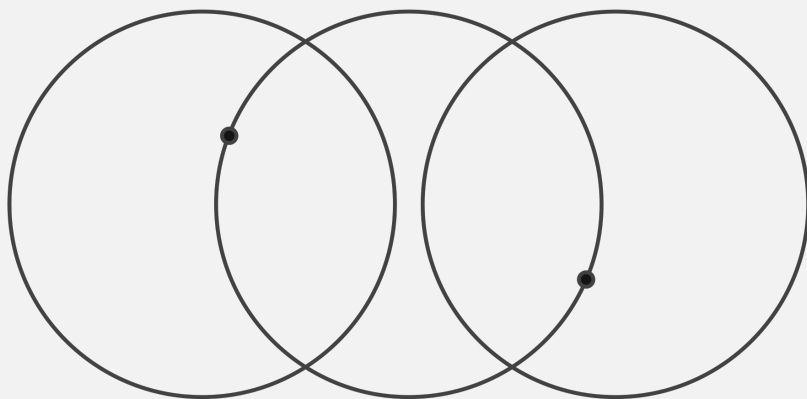
02

Подготовить слайд с предпосылками красного уровня:

- Сложность вопросов «как в работе» — описать, что это значит, про вероятности
- Вероятность прохождения теста сильным кандидатом должна быть не меньше 80%
- Порог определяется так, чтобы отсеять как можно больше слабых кандидатов при выполнении предыдущего пункта

Красный уровень

Подход к решению



01

Задача 1: выяснить, как вероятность отсева слабого кандидата зависит от количества вопросов теста

Подготовить слайд с результатами:

- график необходимой зависимости
- вывод

02

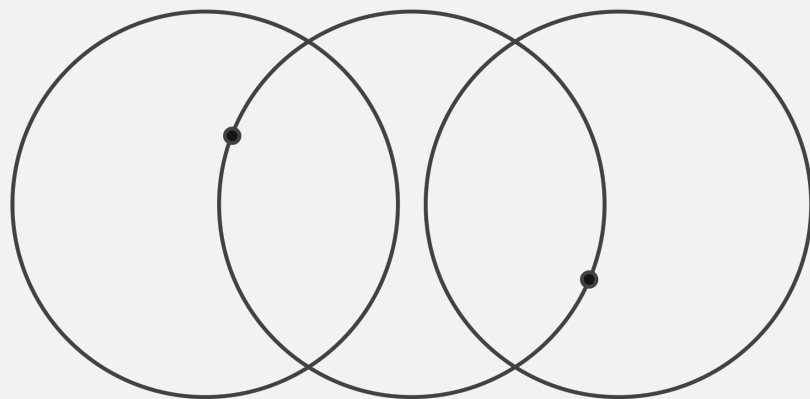
Задача 2: выяснить, сколько вопросов минимум должно быть в тесте, который удовлетворяет условиям выше. Вероятность отсева слабого кандидата должна быть хотя бы 70%

Подготовить слайд с результатами:

- график необходимой зависимости
- вывод

Чёрный уровень

Подход к решению



01

Подготовить слайды синего и красного уровня

02

Подготовить слайд с решением предыдущей задачи для каждого уровня сложности вопросов

Требования: вероятность прохождения сильным — 80%, отсева слабого — 70%



Сложность 1: «проще, чем в работе»

Вероятности правильно ответить такие-то; количество вопросов и проходной балл



Сложность 2: «как в работе»

Предпосылки, оптимальное количество вопросов и проходной балл



Сложность 3: «сложнее, чем в работе»

Предпосылки, оптимальное количество вопросов и проходной балл

Нужно подготовить слайд, сравнивающий все 3 сценария, и обосновать выбор одного из них

Что и когда нужно сдать



Что сдавать?

- Презентация в слайдах в формате pdf
- Jupyter Notebook в Google Colab с расчётами



Когда сдавать?

Сроки сдачи указаны в информационной системе

Сдача проекта – **необходимое условие** прохождения курса

