

# Домашнее задание

## Задача 1

Для маркетингового исследования случайным образом необходимо отобрать  $N$  людей. Этим людям далее демонстрируется кофе марки Nescafe в упаковке с новым дизайном: вместо классической этикетки в коричнево-золотых тонах – новогодняя этикетка в красно-зеленых тонах. После этого фиксируется, сколько людей захотели купить этот кофе, а точнее, доля таких людей.

Рассчитайте необходимый объем выборки для тестовой группы, если известно, что

- выбран 95%-ный уровень доверия;
- погрешность не должна превышать 3%;
- нет определенных ожиданий относительно доли людей, которые захотят кофе с измененным дизайном этикетки (считается, что человек с равной вероятностью выберет кофе или не выберет).

Для вычислений можно использовать базовые средства R или Python, но в таком случае необходимо привести код в тексте с решением.

## Задача 2

Решите задачу 1, учитывая, что теперь ожидаемая доля людей, которые захотят купить кофе с измененным дизайном этикетки, известна и равна 0.67 (группа аналитиков нашла старый отчет с результатами аналогичного эксперимента).

Какой из необходимых объемов выборки больше: в случае, когда мы считаем, что шансы выбрать и не выбрать кофе одинаковы или в случае, когда ожидаемая доля выбравших кофе нам известна и равна 0.67? Сделайте вывод о связи между ожидаемой долей и необходимым объемом выборки при прочих равных условиях (тот же уровень доверия).

## Задача 3

В файле `experiment.csv` сохранены результаты эксперимента, описанного в задаче 1. Данные представляют собой последовательности из 0 и 1, где 1 соответствуют людям, которые захотели выбрать кофе с измененным дизайном этикетки, а 0 - людям, которые не захотели.

### Описание данных:

`test`: результаты по тестовой группе (людям предлагали кофе с новым дизайном этикетки); `control`: контрольная группа (людям предлагали кофе со старым дизайном этикетки).

1. Загрузите в R данные из файла `experiment.csv` с помощью функции `read.csv()`.
2. Постройте в R 90%-ный доверительный интервал для доли людей, которые захотели купить кофе, в тестовой группе. Проинтерпретируйте полученный доверительный интервал.

В решении приведите используемый код, границы доверительного интервала и текст с интерпретацией.

3. Постройте в R 90%-ный доверительный интервал для доли людей, которые захотели купить кофе, в контрольной группе. Проинтерпретируйте полученный доверительный интервал.

В решении приведите используемый код, границы доверительного интервала и текст с интерпретацией.

4. Сравните полученные доверительные интервалы для тестовой и контрольной группы. Можно ли на основании полученных результатов сделать определенный вывод о том, что новый дизайн этикетки лучше (привлекает больше потенциальных покупателей)? Приведите ваши выводы и объяснения.
5. Выполните пункты 1-3 в Python. Изменились ли ваши выводы относительно полезности нового дизайна этикетки? Приведите код, результаты и текст с выводами.