

# Практическое задание 2

*Практическое задание не сдается на проверку и не оценивается.*

## Построение доверительных интервалов в Python

В Python механизм вычисления доверительного интервала немного отличается от R. Поэтому в рамках этого задания вам предлагается работать не со сводной информацией (сколько людей участвовало в эксперименте и сколько выполнили определенные действия), а с исходным массивом данных.

Дан массив `users`, состоящий из 0 и 1, где 0 соответствует людям, не кликнувшим на кнопку «Оформить заказ», а 1 — людям, кликнувшим на кнопку «Оформить заказ».

```
import numpy as np
users = np.array([1, 0, 0, 0, 1, 1, 1, 0, 1])
```

1. Используя Python, посчитайте и сохраните в переменную `n` общее число людей в группе — число элементов массива `users`.
2. Используя Python, посчитайте и сохраните в переменную `p` долю единиц в массиве `users` (эта та доля, на основе которой мы строим доверительный интервал).
3. Посчитайте значение *стандартной ошибки доли* — показателя, который измеряет изменчивость значений в выборке, и сохраните результат в переменную `se`.

Стандартная ошибка показывает, насколько сильно значения в выборке отличаются от среднего значения (более формально, это оценка стандартного отклонения доли).

Значение стандартной ошибки зависит от доли единиц в выборке и от общего числа элементов в выборке. Считается она просто:

$$se = \sqrt{\frac{p(1-p)}{n}}.$$

4. Импортируйте модуль `stats` из библиотеки `scipy`:

```
import scipy.stats as st
```

Используйте функцию `interval()` из подмодуля `norm` для построения 95%-ного доверительного интервала для доли людей, кликнувших на кнопку «Оформить заказ».

**Подсказка:** `st.norm.interval(confidence, loc, scale)`, где `confidence` — уровень доверия, `loc` — середина распределения, то есть интересующая нас доля, `scale` — показатель разброса значений, то есть стандартная ошибка доли.

5. Проинтерпретируйте полученный доверительный интервал.

### Дополнение для желающих: доверительный интервал для среднего

1. Создайте `numpy`-массив `ages` со значениями возраста посетителей сайта, если известно, что значения такие: 23, 24, 38, 34, 36, 37, 39.
2. Постройте 95%-ный доверительный интервал для среднего возраста посетителей сайта, используя массив `ages`. Проинтерпретируйте полученный доверительный интервал.

**Подсказка:** импортируйте необходимый модуль `statsmodels.stats.api`, извлеките из него функцию `DescrStatsW`, примените ее к массиву значений, а потом примените метод `.tconfint_mean()`, который выведет на экран границы доверительного интервала.

```
import statsmodels.stats.api as stm
stm.DescrStatsW(ages).tconfint_mean()
```