

Запуск новой функциональности

Нахождение доверительного интервала для среднего

```
In from scipy import stats as st

confidence_interval = st.t.interval(alpha = alpha, df=df,
loc=sample.mean(), scale=sample.sem())

# alpha – уровень значимости;
# df – количество степеней свободы, равное n – 1;
# loc – среднее распределение, равное оценке среднего.
# Для выборки sample = sample.mean();
# scale – стандартное отклонение распределения, равное оценке стандартной ошибки.
# Вычисляется так: sample.sem()
```

Извлечение подвыборки для бутстрепа

```
In from numpy.random import RandomState
state = RandomState(12345)

# без возвращения
print(example_data.sample(frac=1, replace=False, random_state=state))
# С возвращением
print(example_data.sample(frac=1, replace=True, random_state=state))
```

Анализ A/B теста с помощью бутстрепа

```
In import pandas as pd
import numpy as np

# фактическая разность средних значений в группах
AB_difference = samples_B.mean() - samples_A.mean()

alpha = 0.05
state = np.random.RandomState(12345)

bootstrap_samples = 1000
count = 0
for i in range(bootstrap_samples):
    # подсчитываем, сколько раз разница целевых показателей
    # фактическое значение при условии верности нулевой гипотезы
    united_samples = pd.concat([samples_A, samples_B])
    subsample = united_samples.sample(frac=1, replace=True, random_state=state)

    subsample_A = subsample[:len(samples_A)]
    subsample_B = subsample[len(samples_A):]
    bootstrap_difference = subsample_B.mean() - subsample_A.mean()

    if bootstrap_difference >= AB_difference:
        count += 1

pvalue = 1. * count / bootstrap_samples
print('p-value =', pvalue)

if pvalue < alpha:
    print("Отвергаем нулевую гипотезу: скорее всего, целевой показатель увеличился")
else:
    print("Не отвергаем нулевую гипотезу: скорее всего, целевой показатель не увеличился")
```

Словарь

A/B-тестирование или сплит-тестирование

техника проверки гипотез, которая позволяет оценить, как изменение сервиса или продукта повлияет на пользователей. Техника заключается в разделении всех пользователей на две группы, первая группа видит сервис или продукт без изменений, а вторая видит обновлённый интерфейс

Проблема подглядывания

общий результат искажается, если новые данные поступают в начале эксперимента

Ошибка первого рода

когда нулевая гипотеза верна, но она отклоняется (ложно-положительный результат)

Ошибка второго рода

когда нулевая гипотеза не верна, но принимается (ложно-отрицательный результат)

Доверительный интервал

отрезок числовой оси, в который с заданной вероятностью попадает нужный нам параметр генеральной совокупности