

Measurements cheat sheet

Этот листочек для измерения времени работы одной и той же программы на одном и том же датасете. Примеры будут на языке Python, чтобы инструкцию смогло осилить как можно больше народу. Некоторые моменты сильно упрощены, но для студенческих работ это ОК. В курсе матстатистики вам расскажут, как оно на самом деле работает.

1. Подготовьте тестовый стенд: отключите обновления, сбросьте все кэши, отключите фоновую музыку, закройте браузеры, стабилизируйте частоту CPU (например, выставьте на минимальную; если выставить на максимальную, CPU начнет перегреваться) etc. Проверьте алгоритм, входные данные, средство измерения на адекватность. Проконсультируйтесь с научным руководителем.
2. Проведите серию измерений. Желательно, чтобы количество замеров было больше 20 (ну, например, 40), но и не слишком большое: десятки тысяч измерений тоже довольно бестолково. Получите список значений:

$$t = (X_1, X_2, X_3, \dots, X_n)$$

3. Подготовьтесь к анализу:

```
from scipy import stats
import numpy as np
import matplotlib.pyplot as plt
```

4. Изучите, как распределены данные. Постройте гистограмму, посмотрите на нее. На гистограмме не должно быть сильных выбросов. Если выброс одиночный, то его можно отбросить¹. Но только если он один! Если их несколько, то см пункт 1.

```
plt.hist(t)
```

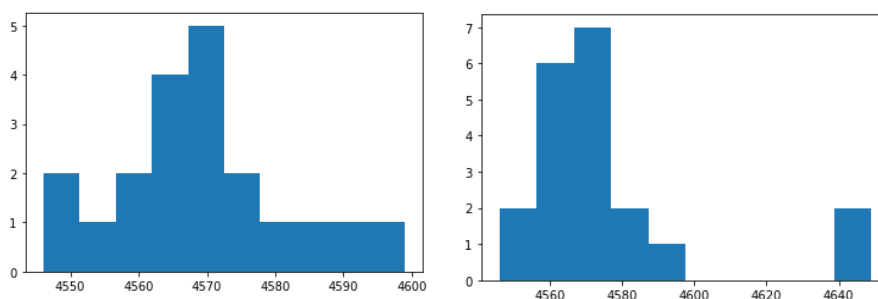


Рис. 1: Пример «хорошей» (слева) и «плохой» (справа) гистограмм

5. Проверьте, что данные проходят некоторые тесты на нормальность. В нашем случае признаком успешного прохождения теста является $p\text{-value} > 0.05$ хотя бы на одном тесте. Если тесты не проходятся, см пункт 1.

```
stats.normaltest(t)
stats.shapiro(t)
```

6. Вычислите среднее и стандартное отклонение. Стандартное отклонение не должно быть слишком большим (5 – 10% от среднего). Если оно большое, см пункт 1.

```
np.mean(t)
np.std(t, ddof=1)
```

7. Определите, нужен ли вам доверительный интервал. Если да, то вычислите его. В примере используется уровень доверия 95%:

```
stats.t.ppf(0.975, df=len(t)-1)*stats.sem(t)
```

Примечание: 0.975 — это не опечатка.

8. Определите, какая именно случайная погрешность вам нужна, затем сложите полученную случайную погрешность с погрешностью инструмента измерения.
9. Проведите округление в соответствии с принятыми правилами. Не нужно выносить на слайды те числа, которые напечатал скрипт:

Погрешность : 12.879372652772872 \rightarrow 13

Среднее : 4568.8 \rightarrow 4569

Интервал записывается так : 4569 \pm 13

- В погрешности оставьте одну значащую цифру (или две, если первая значащая цифра единица)
- Среднее округлите до той точности, с которой приведена погрешность.

10. Если у вас много разных величин разных порядков, имеет смысл вычислять относительные погрешности вместо абсолютных. Так нагляднее.

¹Методологически надо применить формальные критерии (напр., Шовене), но мы упростим себе жизнь и просто отбросим.