

```
In [22]: 1 import scipy
2 from scipy import stats
3 import pandas as pd
4 import math
```

- V prvom prípade se jedna o pravy test (pred/po -- parova data)
- V druhom prípade se jedna o dvouvýberový t-test (dva výbery)

RUCE

```
In [12]: 1 df = pd.read_csv("data-ruce.txt", delimiter=" ")
2
3 # two-sided test
4 # nevim jestli pocet zachvevu ocekavat mensi nebo vetsi
5 # zajima me, jestli se to proste lisi
6 stats.ttest_rel(df["ruka_pred"], df["ruka_po"])
```

Out[12]: Ttest_relResult(statistic=-3.1814457556571267, pvalue=0.00873863596242495)

- H_0 : same $E(X)$
- H_1 : different $E(X)$
- $\alpha = 0.01$
- $pvalue = 0.008 < 0.01$... **zamítame H_0** - zamítame stejné střední hodnoty

ZUBY

```
In [20]: 1 data = pd.read_csv("data-zuby.txt", delimiter=" ")
2 stats.ttest_ind(data["zuby_pomlau"], data["zuby_sok"])
```

Out[20]: Ttest_indResult(statistic=2.019354722091951, pvalue=0.06302124008545723)

- H_0 : same $E(X)$
- H_1 : different $E(X)$
- $\alpha = 0.05$
- $pvalue = 0.06 > 0.05$... **nezamítame H_0** , ve prospěch toho že dva výbery mají různé střední hodnoty

Konfidenční interval:

```
In [39]: 1 X = data["zuby_pomlau"]
2 Y = data["zuby_sok"]
3 n = X.size
4 df = ( 2 * n - 2 )
5 std_x = X.std()
6 std_y = Y.std()
7 std_x_y = math.sqrt(((n - 1) * (std_x)**2 + (n - 1) * (std_y)**2) / df)
8 mu = X.mean() - Y.mean()
9 t = scipy.stats.t.ppf(0.95, df) * std_x_y * math.sqrt(1/n + 1/n)
10 print ('95% konfidenční interval pro rozdíl středních hodnot síly potřebné k
```

95% konfidenční interval pro rozdíl středních hodnot síly potřebné k drcení je:
[0.149, 2.176]