0. python imports

In [36]:
```
import pandas as pd
from scipy.stats import ttest_rel, ttest_1samp, ttest_ind
```

1. data loading

In [4]:
```
blood_pressure = pd.read_csv('./data/blood_pressure.csv')
blood_pressure.head()
```

Out[4]:

	before	after
0	136.713072	92.432965
1	134.735618	105.022643
2	127.529115	82.242766
3	144.527126	93.607172
4	124.214720	103.212223

In [35]:
```
ab_test = pd.read_csv('./data/ab_test.csv')
ab_test.head()
```

Out[35]:

	a	b
0	0.27	13.61
1	6.08	21.53
2	13.74	9.23
3	9.70	5.36
4	7.00	12.90

2. hypothesis test example (related samples)

test related distributions, is the differences between them due to chance?

In [8]:
```
ttest_rel(blood_pressure['after'], blood_pressure['before'])
```

Out[8]:
```
Ttest_relResult(statistic=-27.291841767560236, pvalue=7.303035069608042e-48)
```

test related distributions as mean difference is zero

In [26]:
```
blood_pressure['diff'] = blood_pressure['after'] - blood_pressure['before']
blood_pressure.head()
```

Out[26]:

	before	after	diff
0	136.713072	92.432965	-44.280107
1	134.735618	105.022643	-29.712975
2	127.529115	82.242766	-45.286349
3	144.527126	93.607172	-50.919953
4	124.214720	103.212223	-21.002497

In [33]:
```
ttest_1samp(blood_pressure['diff'], 0)
```

Out[33]:
```
Ttest_1sampResult(statistic=-27.291841767560236, pvalue=7.303035069608042e-48)
```

ojo: https://stackoverflow.com/questions/15984221/how-to-perform-two-sample-one-tailed-t-test-with-numpy-scipy (https://stackoverflow.com/questions/15984221/how-to-perform-two-sample-one-tailed-t-test-with-numpy-scipy)

3. hypothesis test example (independent samples)

assuming equal variances

In [39]:
```
ttest_ind(ab_test['b'], ab_test['a'], equal_var=True)
```

Out[39]:
```
Ttest_indResult(statistic=2.637533181209767, pvalue=0.009713140852447347)
```

assuming unequal variances (Welch's)

In [41]:
```
ttest_ind(ab_test['a'], ab_test['b'], equal_var=False)
```

Out[41]:
```
Ttest_indResult(statistic=-2.637533181209767, pvalue=0.009776243024828825)
```