

```
In [41]: import numpy as np
import matplotlib.pyplot as plt
import scipy.stats as sps
import pandas as pd

from statsmodels.distributions.empirical_distribution import ECDF
from sklearn.neighbors import KernelDensity
import statsmodels.api as sm
from statsmodels.sandbox.stats.multicomp import multipletests
from skidmarks import wald_wolfowitz

%matplotlib inline
```

```
In [9]: wine_data = pd.read_csv('wine.data', header = None, sep=',')

sample = []
sample.append(wine_data.values[:,1])
sample.append(wine_data.values[:,4])
sample.append(wine_data.values[:,3])
sample.append(wine_data.values[:,8])
sample.append(wine_data.values[:,9])

slump_test = pd.read_csv('slump_test.data', header = 0, sep=',')
sample.append(slump_test.values[:, -1])

vowel_context = pd.read_csv('vowel-context.data', header = None, sep='\s+')
sample.append(vowel_context.values[:,7])
```

1: Alcohol

3: Ash

4: Alcalinity of ash

8: Nonflavanoid phenols

9: Proanthocyanins

```

In [47]: num_tests = 4
def normility(samples, alpha = 0.05, method='bonferroni'):
    N = num_tests * len(samples)
    p_val = np.zeros(N)
    for i in range(len(samples)):
        p_val[0 + num_tests * i] = sps.kstest(samples[i], sps.norm(np.mean(samples[i]), np.std(samples[i])))[1]
        # st, p_val[1 + 2 * i] = sps.jarque_bera(samples[i])
        # st, p_val[1 + 2 * i] = sps.skewtest(samples[i])
        p_val[1 + num_tests * i] = sps.shapiro(samples[i])[1]
        p_val[2 + num_tests * i] = sps.normaltest(samples[i])[1]
    return multipletests(p_val, alpha=alpha, method=method)

for method in ['fdr_by']:
    reject, pvals_corrected = normility(np.array(sample), alpha=0.05, method=method)
    print("method %s" % (method))
    for i in range(len(sample)):
        print("%d гипотеза : " % (i) + str(reject[ num_tests * i: num_tests * (i+1)]))

method fdr_by
0 гипотеза : [False False  True]
1 гипотеза : [ True False False]
2 гипотеза : [ True False False]
3 гипотеза : [ True  True  True]
4 гипотеза : [ True False False]
5 гипотеза : [ True False False]
6 гипотеза : [ True False False]

```

Принимаем гипотезу о нормальности распределения четвёртой выборки

Критерий Бенджамина Хохтберга в данном случае нельзя использовать из-за явной зависимости части гипотез