

banknote authentication

October 25, 2018

```
In [1]: import numpy as np
data = np.genfromtxt("https://archive.ics.uci.edu/ml/machine-learning-databases/00267/da
,delimiter=",")
X = data[:,0:3]
y = data[:,4]
```

```
In [2]: import statsmodels.api as sm
model = sm.Logit(y, sm.add_constant(X))
result = model.fit(method="cg",maxiter=1000)
```

```
/home/nbuser/anaconda3_501/lib/python3.6/site-packages/statsmodels/compat/pandas.py:56: FutureWarning
from pandas.core import datetools
```

```
Optimization terminated successfully.
Current function value: 0.019424
Iterations: 46
Function evaluations: 131
Gradient evaluations: 131
```

```
In [3]: result.summary()
```

```
Out[3]: <class 'statsmodels.iolib.summary.Summary'>
"""
```

```

                        Logit Regression Results
=====
Dep. Variable:          y      No. Observations:          1372
Model:                Logit   Df Residuals:          1368
Method:               MLE     Df Model:              3
Date:                 Thu, 25 Oct 2018   Pseudo R-squ.:        0.9717
Time:                 03:37:06   Log-Likelihood:       -26.649
converged:            True     LL-Null:              -942.56
                                LLR p-value:              0.000
=====
                coef    std err          z      P>|z|      [0.025    0.975]
-----
const          6.8818    1.383      4.977    0.000     4.172    9.592
```

x1	-6.7797	1.394	-4.864	0.000	-9.511	-4.048
x2	-3.5044	0.693	-5.060	0.000	-4.862	-2.147
x3	-4.4614	0.900	-4.958	0.000	-6.225	-2.698

=====

Possibly complete quasi-separation: A fraction 0.86 of observations can be perfectly predicted. This might indicate that there is complete quasi-separation. In this case some parameters will not be identified.
 ""

In [4]: result.conf_int()

Out[4]: array([[4.17171943, 9.59186301],
 [-9.51141385, -4.04792873],
 [-4.86180827, -2.14700806],
 [-6.22497278, -2.69788499]])