

```

import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
import statsmodels.api as sm
import statsmodels.formula.api as smf
import math

pos = [0.0, 1.0, 2.0, 3.0, 4.0, 5.0, 6.0, 7.0, 8.0, 9.0, 10.0]
neg = [10.0, 9.0, 8.0, 7.0, 6.0, 5.0, 4.0, 3.0, 2.0, 1.0, 0.0]

df_pos = pd.DataFrame(data=pos)
df_neg = pd.DataFrame(data=neg)

# вероятность успеха при каждом исходе
count = 10
poss_pos=[pos[i]/(pos[i]+neg[i]) for i in range(1,len(pos)-1)]
poss_neg=[neg[i]/(pos[i]+neg[i]) for i in range(1,len(pos)-1)]

oddspos=[poss_pos[i]/(1.0-poss_pos[i]) for i in range(len(poss_pos))]
oddsneg=[poss_neg[i]/(1.0-poss_neg[i]) for i in range(len(poss_neg))]

oddspos

[0.11111111111111112,
 0.25,
 0.4285714285714286,
 0.6666666666666667,
 1.0,
 1.4999999999999998,
 2.333333333333333,
 4.000000000000001,
 9.000000000000002]

oddsneg

[9.000000000000002,
 4.000000000000001,
 2.333333333333333,
 1.4999999999999998,
 1.0,
 0.6666666666666667,
 0.4285714285714286,
 0.25,
 0.11111111111111112]

ln_oddspos=np.log(oddspos)
ln_oddsneg=np.log(oddsneg)

ln_oddspos

```

```
array([-2.19722458, -1.38629436, -0.84729786, -0.40546511,
0.40546511, 0.84729786, 1.38629436, 2.19722458])
```

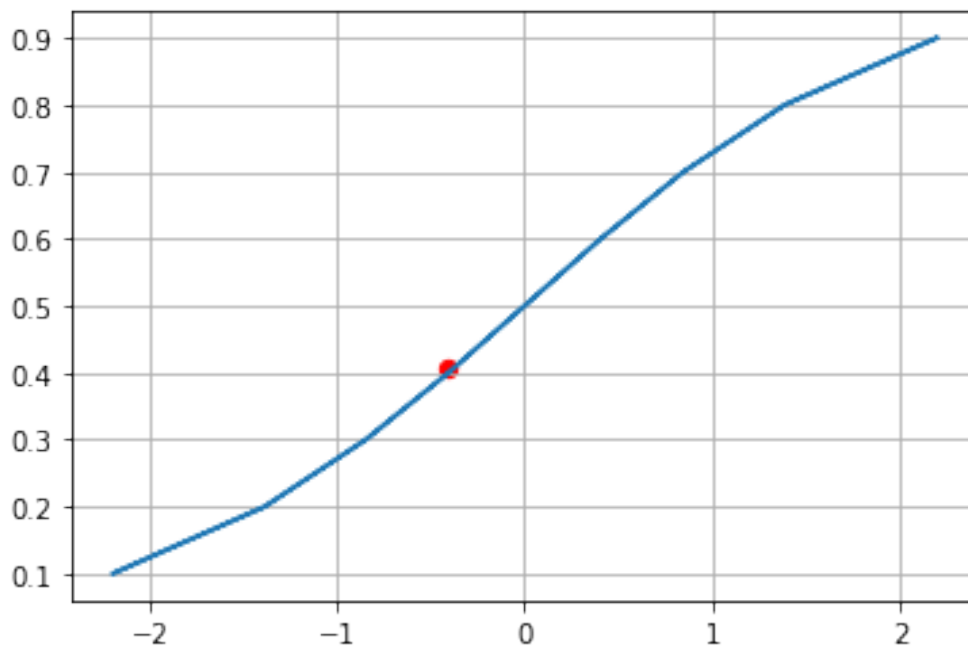
```
ln_oddsneg
```

```
array([ 2.19722458, 1.38629436, 0.84729786, 0.40546511,
0.40546511, -0.84729786, -1.38629436, -2.19722458])
```

```
p=0.4
```

```
odds=p/(1-p)
```

```
plt.plot(ln_oddsneg.tolist()+ln_oddsposs.tolist(), poss_neg+poss_pos)
plt.scatter(math.log(odds), p, color='red', s=40)
plt.grid()
plt.show()
```



```
df = pd.read_csv('Titanic.csv')
df = df[df.Age.notnull()]
```

```
df
```

	PassengerId	Survived	Pclass	\
0	1	0	3	
1	2	1	1	
2	3	1	3	
3	4	1	1	
4	5	0	3	
...	
885	886	0	3	

886	887	0	2
887	888	1	1
889	890	1	1
890	891	0	3

SibSp \	Name	Sex	Age
0	Braund, Mr. Owen Harris	male	22.0
1			
1	Cumings, Mrs. John Bradley (Florence Briggs Th...	female	38.0
1			
2	Heikkinen, Miss. Laina	female	26.0
0			
3	Futrelle, Mrs. Jacques Heath (Lily May Peel)	female	35.0
1			
4	Allen, Mr. William Henry	male	35.0
0			
..
...			
885	Rice, Mrs. William (Margaret Norton)	female	39.0
0			
886	Montvila, Rev. Juozas	male	27.0
0			
887	Graham, Miss. Margaret Edith	female	19.0
0			
889	Behr, Mr. Karl Howell	male	26.0
0			
890	Dooley, Mr. Patrick	male	32.0
0			

	Parch	Ticket	Fare	Cabin	Embarked
0	0	A/5 21171	7.2500	NaN	S
1	0	PC 17599	71.2833	C85	C
2	0	STON/O2. 3101282	7.9250	NaN	S
3	0	113803	53.1000	C123	S
4	0	373450	8.0500	NaN	S
..
885	5	382652	29.1250	NaN	Q
886	0	211536	13.0000	NaN	S
887	0	112053	30.0000	B42	S
889	0	111369	30.0000	C148	C
890	0	370376	7.7500	NaN	Q

[714 rows x 12 columns]

```
binom = smf.glm(formula='Survived ~ 1', data=df,
family=sm.families.Binomial()).fit()
binom.summary()
```

```
<class 'statsmodels.iolib.summary.Summary'>
"""
```

Generalized Linear Model Regression Results

```
=====
Dep. Variable:          Survived    No. Observations:
714
Model:                  GLM         Df Residuals:
713
Model Family:          Binomial     Df Model:
0
Link Function:         logit        Scale:
1.0000
Method:                 IRLS        Log-Likelihood:
-482.26
Date:                  Mon, 12 Dec 2022    Deviance:
964.52
Time:                  11:12:20    Pearson chi2:
714.
No. Iterations:         4

Covariance Type:        nonrobust
```

```
=====
=====
              coef      std err          z      P>|z|      [0.025
0.975]
-----
-----
Intercept    -0.3799      0.076     -4.985      0.000     -0.529
-0.230
=====
=====
"""
```

```
p = df.Survived.value_counts()[1] / (df.Survived.value_counts()[0]+
df.Survived.value_counts()[1])
```

```
p
```

```
0.4061624649859944
```

```
binom.aic
```

```
966.5159648555248
```

```
sns.countplot(x='Survived', data=df, palette='hls')
```

```
<matplotlib.axes._subplots.AxesSubplot at 0x7f7f463cce50>
```

