

import wooldridge as woo

import numpy as np

import pandas as pd

import seaborn as sns

import matplotlib.pyplot as plt

from stargazer.stargazer import Stargazer

from IPython.core.display import HTML

import statsmodels.formula.api as smf

from **statsmodels.stats.outliers\_influence** import **variance\_inflation\_factor**

Wage1 = woo.dataWoo('wage1')

# OLS regression: the target multiple regression model

lmres = smf.ols('np.log(wage) ~ educ+exper+tenure', data=Wage1).fit()

st=Stargazer([lmres])

HTML(st.render\_html())

* calculate VIF :

X = Wage1[["educ","exper","tenure"]].copy()

X['Intercept'] = 1

# Compute and view VIF

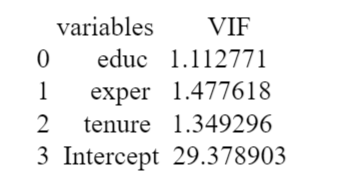
vif = pd.DataFrame()

vif["variables"] = X.columns

vif["VIF"] = [**variance\_inflation\_factor(X.values, i)** for i in range(X.shape[1])]

# View results using print

print(vif)



* Plotting the residuals in a histogram

from **scipy.stats** import **norm**

x\_axis = np.arange(-2, 2, 0.001) #a vector: from -2 to 2

res=pd.DataFrame(lmres.resid)

sns.distplot(res, fit=norm);

