**PENJELASAN TABEL ANOVA**

Tabel ANOVA (Analysis of Variance) menguji penerimaan (acceptability) model dari perspektif statistik dalam bentuk analisis sumber keragaman. ANOVA ini sering juga diterjemahkan sebagai analisis ragam. Output Anova Penjelasan Uji Regresi Excel.

Kodingan :

import numpy as np

import pandas as pd

import scipy

from scipy import stats

def ANOVATAB(y,yhat,n,m):

dfn=n

dfd=m-n-1

ybar=np.average(y)

SSR=sum((yhat-ybar)\*\*2)

SSE=sum((y-yhat)\*\*2)

print (SSE)

print (dfd)

SST=sum((y-ybar)\*\*2)

MSR=SSR/dfn

MSE=SSE/dfd

Fs=MSR/MSE

ks =1-scipy.stats.f.cdf(Fs,dfn,dfd)

data\_table={

'SS' : [SSR,SSE,SST],

'df' : [dfn,dfd,m-1],

'MS' : [MSR,MSE,'-'],

'Fs' : [Fs,'-','-'],

'pval' : [ks,'-','-']

}

return pd.DataFrame(data\_table)

n=1

m=len(y\_pred)

print('Tabel Anovanya')

result=ANOVATAB(y\_test,y\_pred,n,m)

result

Output :

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **SS** | | **df** | **MS** | **Fs** | **Pval** |
| 0 | 2.087941e+09 | 1 | 2087941040.83831 | 26.165926 | 0.002188 |
| 1 | 4.787771e+08 | 6 | 79796185.052866 | - | - |
| 2 | 3.071911e+09 | 7 | - | - | - |

1. **SS (Sum of Square)**

**SSR = Sum of Square Regression (2.087941e+09)**

**SSE = Sum of Square Residual (4.787771e+08)**

**SST = Sum of Square Total (3.0719113+09)**

Kolom SS (Sum of Square) atau jumlah kuadrat untuk regression diperoleh dari penjumlahan kuadrat dari prediksi variabel terikat (permintaan) dikurangi dengan nilai rata-rata permintaan dari data sebenarnya. Jadi secara manual kita cari dulu rata-rata permintaan dari data asli kita. Kemudian masing-masing prediksi permintaan (lihat tabel residual output di bawah) dikurangi dengan rata-rata tersebut kemudian dikuadratkan. Selanjutnya, seluruh hasil perhitungan tersebut dijumlahkan.

Contohnya, rata-rata permintaan dari data kita = 820. Berdasarkan tabel residual output dibawah, untuk observasi pertama prediksi permintaan = 498.2362193. Selanjutnya kita hitung (498.24 – 820 )2 = 103531.93. Untuk observasi kedua dihitung (262.98 – 820)2 = 310271.8. Demikian seterusnya sampai data terakhir. Selanjutnya, hasil-hasil perhitungan tersebut dijumlahkan dan hasilnya = 769993.78.

Kolom SS untuk residual diperoleh dari jumlah pengkuadratan dari residual. Lihat cara menghitung residual pada tabel residual output dibawah. Nilai-nilai residual tersebut dikuadratkan, kemudian hasilnya dijumlahkan dan hasilnya adalah46006.22.

Kolom SS untuk total adalah penjumlahan dari SS untuk regresi dengan dengan SS untuk residual. Sebenarnya SS total ini adalah variasi (besar-kecil,naik-turun) dari permintaan. Ini diukur dengan mengurangi nilai masing-masing permintaan aktual dengan rata-ratanya, kemudian dikuadratkan. Hasil perhitungan tersebut kemudian dijumlahkan.

Lalu, apa artinya dari angka-angka tersebut? Sekarang perhatikan ketiga hasil kita, SS regresi, SS residual dan SS total.

SS total kita adalah 816000. Artinya, variasi dari pemintaan yang dikuadratkan adalah sebesar nilai tersebut. Lalu apa yang menyebabkan permintaan tersebut bervariasi ? Sebagian berasal dari variabel bebas (harga dan pendapatan) yaitu sebesar 769993.78 (regresi). Lalu sisanya, yang sebesar 46006.22 disebabkan oleh variabel lain yang juga mempengaruhi pendapatan, tetapi tidak dimasukkan dalam model (residual).

Kalau kita bandingkan (bagi) antara SS regresi dengan SS total, maka akan kita dapatkan proporsi dari total variasi permintaan yang disebabkan oleh variasi harga dan pendapatan. Coba kita bagi: 769993.78/816000=0.9436. Nilai 0,9436 adalah R2 atau koefisien determinasi yang telah kita bahas diatas.

1. **Df (Degree of Freedom)**

Degree of Freedom (df) atau derajat bebas dari total adalah n-1, dimana n adalah banyaknya observasi. Karena observasi kita ada 10, maka derajat bebas total adalah 9. Derajat bebas dari model regresi adalah 2, karena ada dua variabel bebas dalam model kita (harga dan pendapatan). Derajat bebas untuk residual adalah sisanya yaitu derajat bebas total – derajat bebas regresi = 9 – 2 = 7.

1. **Mean of Square**

Selanjutnya kolom berikutnya dari ANOVA adalah kolom MS (Mean of Square) atau rata-rata jumlah kuadrat. Ini adalah hasil bagi antara kolom SS dengan kolom df. Dari perhitungan MS ini, selanjutnya dengan membagi antara MS Regresi dengan MS Residual didapatkan nilai F. Nilai F ini yang dikenal dengan F hitung dalam pengujian hipotesa dibandingkan dengan nilai F tabel. Jika F hitung > F tabel, maka dapat dinyatakan bahwa secara simultan (bersama-sama) harga dan pendapatan berpengaruh signifikan terhadap permintaan.

Selain itu, kita juga bisa membandingkan antara taraf nyata dengan p-value (dalam istilah Excel adalahSignificance F). Jika taraf nyata > dari p-value maka kesimpulannya sama dengan di atas. Misalnya kita menetapkan taraf nyata 5%. Karena p-value (Significance F) = 0.000, maka dapat disimpulkan bahwa harga dan pendapatan secara bersama-bersama berpengaruh signifikan terhadap permintaan.

1. Koefisien Regresi