

MANOVA is a technique which determines the **effects of independent categorical variables on multiple continuous dependent variables**. It is usually used to compare several groups with respect to multiple continuous variables.

```
In [ ]: # import libraries
import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt
from statsmodels.multivariate.manova import MANOVA
```

```
In [ ]: # dataset load
df = sns.load_dataset('iris')
df.head()
```

```
Out[ ]:      sepal_length  sepal_width  petal_length  petal_width  species
0           5.1           3.5           1.4           0.2     setosa
1           4.9           3.0           1.4           0.2     setosa
2           4.7           3.2           1.3           0.2     setosa
3           4.6           3.1           1.5           0.2     setosa
4           5.0           3.6           1.4           0.2     setosa
```

```
In [ ]: df.columns
```

```
Out[ ]: Index(['sepal_length', 'sepal_width', 'petal_length', 'petal_width',
              'species'],
              dtype='object')
```

```
In [ ]: # MANOVA by the statsmodel library
mov = MANOVA.from_formula('sepal_length + sepal_width + petal_length + petal_width ~
# here all 4 columns length and width are dependent on species while species are ind
print(mov.mv_test())
```

Multivariate linear model

```
=====
-----
Intercept      Value  Num DF  Den DF    F Value  Pr > F
-----
Wilks' lambda  0.0170  4.0000  144.0000  2086.7720  0.0000
Pillai's trace  0.9830  4.0000  144.0000  2086.7720  0.0000
Hotelling-Lawley trace  57.9659  4.0000  144.0000  2086.7720  0.0000
Roy's greatest root  57.9659  4.0000  144.0000  2086.7720  0.0000
-----

-----
species        Value  Num DF  Den DF    F Value  Pr > F
-----
Wilks' lambda  0.0234  8.0000  288.0000  199.1453  0.0000
Pillai's trace  1.1919  8.0000  290.0000   53.4665  0.0000
Hotelling-Lawley trace  32.4773  8.0000  203.4024  582.1970  0.0000
Roy's greatest root  32.1919  4.0000  145.0000 1166.9574  0.0000
=====
```

In []:

```
import pandas as pd
x = pd.DataFrame((mova.mv_test().results['species']['stat']))
x
```

Out[]:

	Value	Num DF	Den DF	F Value	Pr > F
Wilks' lambda	0.023439	8	288.0	199.145344	0.0
Pillai's trace	1.191899	8.0	290.0	53.466489	0.0
Hotelling-Lawley trace	32.47732	8	203.40239	582.197018	0.0
Roy's greatest root	32.191929	4	145	1166.957433	0.0

- The output shows the analysis using different test statistics. Wilks' lambda ,Pillai's trace, Hotelling-Lawley trace and Roy's greatest root are used but there is no absolute consensus in the statistical literature as to which test statistic should be preferred.
 - The p-values are shown in the right column and are all inferior to 0.05, which confirms that features has an impact on dependant variable.it is significant.
-