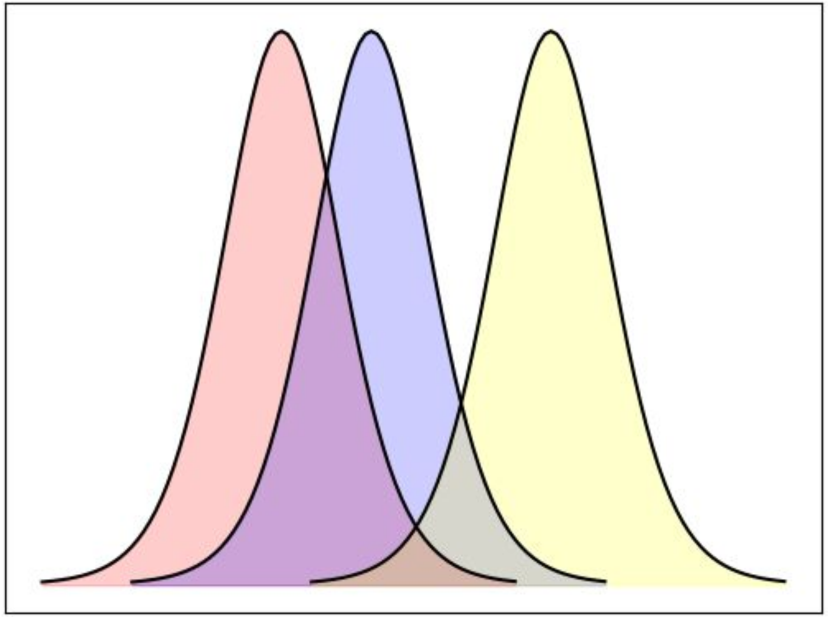
# Analysis of Variance (ANOVA)

Jolla Kullgren
Department of Chemistry - Ångström

# one-way ANOVA F-test



- The F-test can be used to assess whether the expected values of a quantitative variable within several pre-defined groups differ from each other.

- It can be seen as a generalization of the t-test for more than two data-sets.

UPPSALA UNIVERSITET

# F-test main formula

$$F = \frac{\text{"beetween-group variablity"}}{\text{"within-group variablity"}} = \frac{\sum_{i=1}^{K} n_i(\bar{x}_i - \bar{x})^2/(K-1)}{\sum_{i=1}^{K}\sum_{j=1}^{n_i}(x_{ij} - \bar{x}_i)^2/(N-K)}$$

$K$ : Number of groups

$n_i$ : Samples in group $i$

$\bar{x}_i$ : Mean of group $i$

$\bar{x}$ : Overall mean

$x_{ij}$ : $j^{\text{th}}$observation in the $i^{\text{th}}$ group.

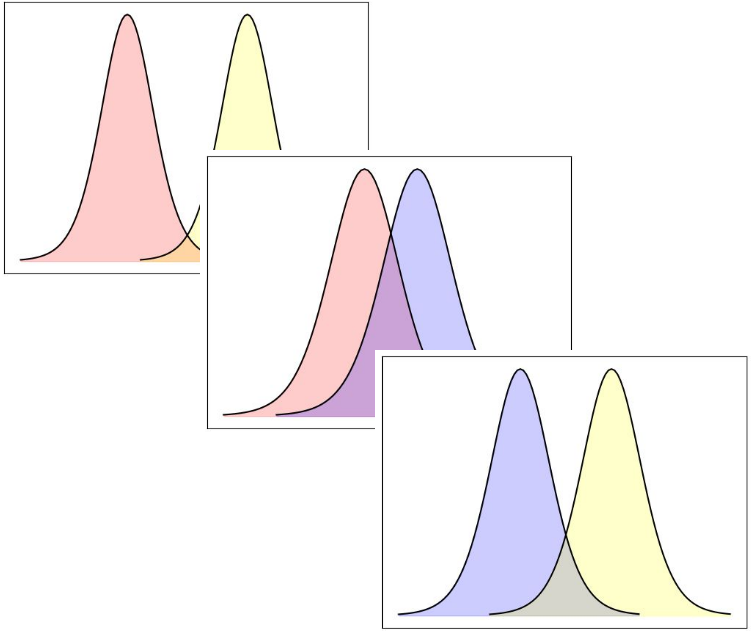
$N$ : Overall sample size

# Example: ANOVA in python:

```
from scipy import stats
import numpy as np

N1=5 ; off1 = 0.1 ; w1=0.4
N2=4 ; off2 = 0.6 ; w2=0.3
N3=6 ; off3 = 0.3 ; w3=0.2

data_1 = w1*np.random.rand(N1)+off1
data_2 = w2*np.random.rand(N2)+off2
data_3 = w3*np.random.rand(N3)+off3

F_test, p_value =stats.f_oneway(data_1,data_2,data_3)
print("t-test: ",F_test," -> P-value: ",p_value)

================================================================
t-test:  15.011495298649004  -> P-value:  0.0005422077738381697
```

P-value suggest we can reject null hypothesis. The mean value are not significantly the same. However, we do not know if they are all different or if one is different from the other two...

# Pair-wise independent two-sample t-test



Using multiple ***independent two-sample t-test*** we can try to locate the “outlier”.

- Using multiple independent two-sample t-test we can identify “outlier”
- With 95 % confidence “mislabel” 5 % of the times. With many comparisons, this become a substantial problem.
- The *Bonferroni correction*: Scale the significance level with the number of comparisons. With a 5 % significance level, p-values should be $0.05/N_{pairs}$ or less.

UPPSALA UNIVERSITET

# Example: pair-wise t-tests in python:

```
from scipy import stats
import numpy as np

print("Pair 1-2")
t_test, p_value = stats.ttest_ind(data_1, data_2)
print(t_test,p_value)

print("Pair 1-3")
t_test, p_value = stats.ttest_ind(data_1, data_3)
print(t_test,p_value)

print("Pair 2-3")
t_test, p_value = stats.ttest_ind(data_2, data_3)
print(t_test,p_value)
```

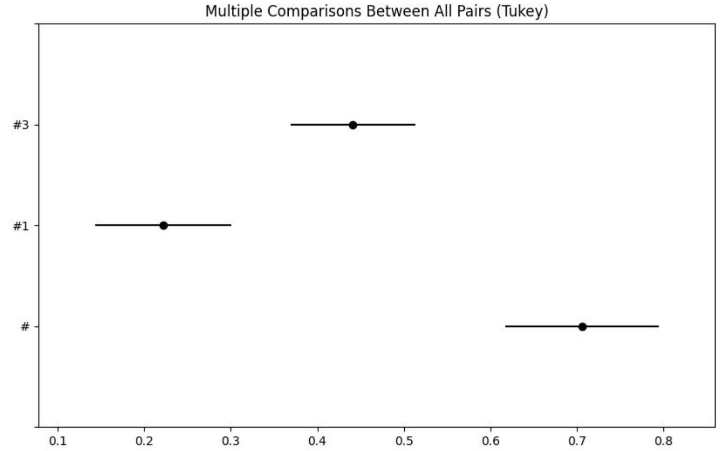
The output is shown on the next slide.

# Example: pair-wise t-tests in python (continued):

```
=================================================================
=

Pair 1-2
-7.294292120993236 0.00016356196701124712
Pair 1-3
-3.6621612014725207 0.00521797192193281
Pair 2-3
5.023586272797825 0.0010222980431426679
```

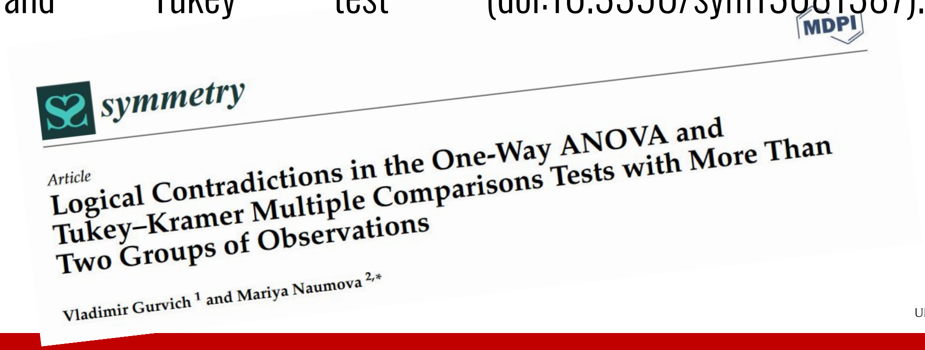
P-values suggest we can reject all  null hypothesis. This is true also when applying the Bonferroni correction, in which case our threshold should be 0.05/3.

# Tukey's-test




Using multiple ***independent two-sample t-test*** we can try to locate the "outlier".

- Another way of performing a post hoc-test is the so-called Tukey's test.
- The result from the test can be summarized in a plot a shown to the left.
- Note: There can be "logical contradictions" between ANOVA and Tukey test (doi:10.3390/sym13081387).


MDPI

symmetry

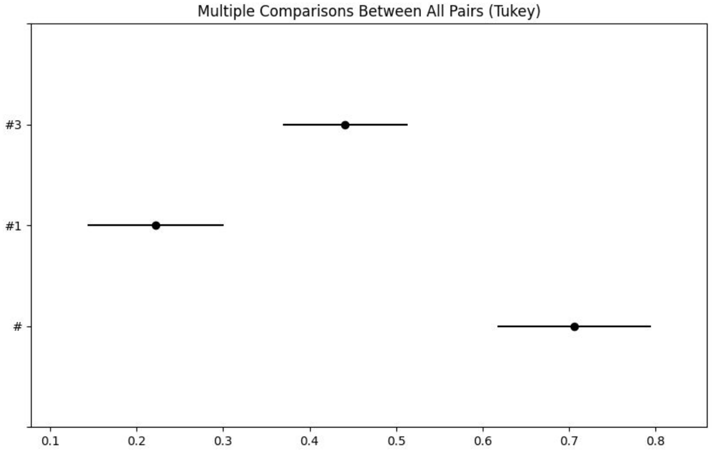
*Article*

**Logical Contradictions in the One-Way ANOVA and Tukey–Kramer Multiple Comparisons Tests with More Than Two Groups of Observations**

Vladimir Gurvich [1] and Mariya Naumova [2,*]

UPPSALA UNIVERSITET

# Tukey's-test




Using multiple ***independent two-sample t-test*** we can try to locate the "outlier".

- We use the following test statistics for a pair A-B:

$$q^{test} = \frac{\bar{x}_A - \bar{x}_B}{s_{A+B}/\sqrt{n_{A+B}}}$$

$s_{A+B}$ and $n_{A+B}$ refer to the variance and number of samples in the combined "A+B" group.

- The comparison is made to the studentized range distribution from the entire set of groups.

$$q = \frac{\bar{x}_{max} - \bar{x}_{min}}{\sqrt{2} s_{Tot}/\sqrt{n_{Tot}}}$$

Where max, min and Tot, refer to minimum, maximum, and the total numbers of samples.

# Example: Tukey's test in python:

```
from statsmodels.stats.multicomp import pairwise_tukeyhsd
Import pandas as pd

data_1=pd.DataFrame({"data-set": len(data_1)*["#1"], "data":data_1 })
data_2=pd.DataFrame({"data-set": len(data_2)*["#2"], "data":data_2 })
data_3=pd.DataFrame({"data-set": len(data_3)*["#3"], "data":data_3 })

data=pd.concat([data_1,data_2,data_3],ignore_index=True)
tukey = pairwise_tukeyhsd(endog=data['data'],  # Data
                          groups=data['data-set'],   # Groups
                          alpha=0.05)          # Significance level

tukey.summary()
=====================================================================
Multiple Comparison of Means - Tukey HSD, FWER=0.05
group1     group2     meandiff   p-adj      lower     upper    reject
#2         #1         -0.4845    0.0        -0.652    -0.317   True
#2         #3         -0.2649    0.0024     -0.4261   -0.1037  True
#1         #3          0.2196    0.0058      0.0684    0.3708  True
```

The test is consistent with the pair-wise t-tests.

# Summary

- ANOVA can be used to assess whether the expected values of a quantitative variable within several pre-defined groups differ from each other.
- ANOVA cannot identify which(s) group(s) that differ from the others.
- Outliers can be identified by performing pair-wise t-tests, or with Tukey's test.
- One should note that the results from ANOVA and Tukey's test need not be consistent.