# Regression Analysis

## Analysis of Variance

**Nicoleta Serban, Ph.D.**
*Professor*
School of Industrial and Systems Engineering

Mean Pairwise Comparison


1

# About This Lesson


2

# Pairwise Comparison of Means

One primary goal of ANOVA might be to determine which treatment means are bigger or smaller. One way to do this is to compare all *k(k-1)/2* pairs of treatments. For a (1 - $\alpha$) confidence interval for the mean difference $\mu_i - \mu_j$:

$$\left(\hat{\mu}_i - \hat{\mu}_j\right) \pm q_{\alpha,k,N-k}\sqrt{\frac{MSE}{2}\left(\frac{1}{n_i}+\frac{1}{n_j}\right)}$$

- Estimate of difference in means
- $\alpha$ percentile of "studentized range" distribution
- Standard deviation/error of estimator



# Difference Between $t_\alpha$ and $q_\alpha$

**Correct for simultaneous inference**:

- $q > t$ (at any fixed $\alpha$ and df)
- Intervals are wider to compensate for the fact that we are making simultaneous comparisons (multiplicity correction)

**Why?**

**95%** CIs for two populations $\Rightarrow$ (.95)(.95) ≈ .90 $\Rightarrow$ The simultaneous or joint confidence level for the two parameters is roughly **90%**.

**95%** CIs for three populations $\Rightarrow$ (.95)(.95)(.95) ≈ .86 $\Rightarrow$ The simultaneous or joint confidence level for the three parameters is roughly **86%**.

# Example1: Global Suicide by Region

Which country regions have different suicide rates?

Georgia Tech

5

# Pairwise Comparison

*TukeyHSD(aov(suicidesper100k ~ region, data=suicide_data))*

Tukey multiple comparisons of means

95% family-wise confidence level

$region

| | diff | lwr | upr | p adj |
|---|---|---|---|---|
| EASTERN EUROPE-ASIA | 7.1256986 | -0.8654681 | 15.1168654 | 0.1218931 |
| GLOBAL WEST-ASIA | 1.3948384 | -6.3253621 | 9.1150390 | 0.9998655 |
| LATIN AMER. & CARIB-ASIA | -2.4242761 | -9.7079484 | 4.8593961 | 0.9848625 |
| MIDDLE EAST-ASIA | -7.8183246 | -17.4646356 | 1.8279865 | 0.2171605 |
| NORTHERN AMERICA-ASIA | 1.8826591 | -18.6470201 | 22.4123382 | 0.9999996 |
| OCEANIA-ASIA | -0.6423728 | -13.5277421 | 12.2429965 | 1.0000000 |
| SUB-SAHARAN AFRICA-ASIA | -4.2457218 | -17.1310911 | 8.6396474 | 0.9858800 |
| WESTERN ASIA-ASIA | -9.6996143 | -30.2292935 | 10.8300649 | 0.8717761 |
| WESTERN EUROPE-ASIA | 2.4643324 | -10.4210369 | 15.3497016 | 0.9997844 |
| GLOBAL WEST-EASTERN EUROPE | -5.7308602 | -12.5740866 | 1.1123662 | 0.1809537 |
| LATIN AMER. & CARIB-EASTERN EUROPE | -9.5499748 | -15.8966379 | -3.2033117 | 0.0002123 |
| MIDDLE EAST-EASTERN EUROPE | -14.9440232 | -23.9039098 | -5.9841367 | 0.000026 |

……..

Georgia Tech

6

# Pairwise Comparison

*TukeyHSD(aov(suicidesper100k ~ region, data=suicide_data))*

Tukey multiple comparisons of means

95% family-wise confidence level

$region

| | diff | lwr | upr | p adj |
|---|---|---|---|---|
| EASTERN EUROPE-ASIA | 7.1256986 | -0.8654681 | 15.1168654 | 0.1218931 |
| GLOBAL WEST-ASIA | 1.3948384 | -6.3253621 | 9.1150390 | 0.9998655 |
| LATIN AMER. & CARIB-ASIA | -2.4242761 | -9.7079484 | 4.8593961 | 0.9848625 |
| MIDDLE EAST-ASIA | -7.8183246 | -17.4646356 | 1.8279865 | 0.2171605 |
| NORTHERN AMERICA-ASIA | 1.8826591 | -18.6470201 | 22.4123382 | 0.9999996 |
| OCEANIA-ASIA | -0.6423728 | -13.5277421 | 12.2429965 | 1.0000000 |
| SUB-SAHARAN AFRICA-ASIA | -4.2457218 | -17.1310911 | 8.6396474 | 0.9858800 |
| WESTERN ASIA-ASIA | -9.6996143 | -30.2292935 | 10.8300649 | 0.8717761 |
| WESTERN EUROPE-ASIA | 2.4643324 | -10.4210369 | 15.3497016 | 0.9997844 |
| GLOBAL WEST-EASTERN EUROPE | -5.7308602 | -12.5740866 | 1.1123662 | 0.1809537 |
| LATIN AMER. & CARIB-EASTERN EUROPE | -9.5499748 | -15.8966379 | -3.2033117 | 0.0002123 |
| MIDDLE EAST-EASTERN EUROPE | -14.9440232 | -23.9039098 | -5.9841367 | 0.000026 |

- 10 different categories, total of 45 different pairwise comparisons
- Two groups with only one observation and three groups with three observations – not sufficient data for comparison
- Only three pairs have an adjusted p-value smaller than 0.05: Latin America vs Eastern Europe, Middle East vs Eastern Europe and Middle East vs Global West

……..

Georgia Tech

7

# ANOVA Example 2: Keyboard Layout

Three different keyboard layouts are being compared in terms of typing speed.

Which mean typing times for the three keyboard layouts are different?

| Layout 1 | Layout 2 | Layout 3 |
|---|---|---|
| 23.8 | 30.2 | 27.0 |
| 25.6 | 29.9 | 25.4 |
| 24.0 | 29.1 | 25.6 |
| 25.1 | 28.8 | 24.2 |
| 25.5 | 29.1 | 24.8 |
| 26.1 | 28.6 | 24.0 |
| 23.8 | 28.3 | 25.5 |
| 25.7 | 28.7 | 23.9 |
| 24.3 | 27.9 | 22.6 |
| 26.0 | 30.5 | 26.0 |
| 24.6 | * | 23.4 |
| 27.0 | * | * |

Georgia Tech

8

# Pairwise Comparison

*TukeyHSD(aov(speed ~ layout))*
Tukey multiple comparisons of means
95% family-wise confidence level

Fit: aov(formula = speed ~ layout)

$layout

| | diff | lwr | upr | p adj |
|---|---|---|---|---|
| 2-1 | 3.9850000 | 2.854395 | 5.1156053 | 0.0000000 |
| 3-1 | -0.3613636 | -1.463581 | 0.7408538 | 0.7008915 |
| 3-2 | -4.3463636 | -5.500092 | -3.1926352 | 0.0000000 |

- Keyboard layout 2 has a statistically significantly higher typing time than keyboard layouts 1 and 3, on average.
- It is plausible that keyboard layouts 1 and 3 have similar typing time, on average.

Georgia Tech

9

# Summary

Georgia Tech

10