

# SPPH567 TUTORIAL IV

# TODAY'S TASKS

- Do one-way ANOVA
- Run a Tukey-Kramer test
- Do simple linear regression with a categorical independent variable

**GOOD NEWS!**  
**THE THREE TASKS CAN ALL BE DONE**  
**AT THE SAME TIME IN DEDUCER!**

# STEP 1

1. Go to Analysis  
> Linear Model

Console

File Edit Workspace Data Analysis Plots Extras Packages & Data Window Help

Linear Regression Model

dat

Filter:

UID  
Community  
MainRadon  
Location  
Foundation  
Soil  
LowerWindow  
Separation  
Heating  
AC  
HomeAge1990  
RockType  
MainRadon.tr  
Radon\_subLOD  
Radon\_subLODhalf  
Radon\_subLODhalf.tr  
Radon\_subLODroot2

Outcome  
Radon\_subLODroot2.tr

As Numeric

As Factor  
Potential

Weights

Subset

Continue Reset Cancel

```
> load("G:/cou  
>
```

2. Put the outcome  
variable here

3. Put the  
categorical  
variable here

4. Click Continue

# STEP 2

Linear Regression Model Builder

Specify

Outcomes  
Radon\_subLODroot2.tr

Variables  
Potential

2-way  
3-way

+

⋮

\*

—

IN

poly

Model  
Potential

Remove

Continue Reset Cancel

I. Click Continue

# STEP 3

The screenshot shows the 'Linear Regression Model Explorer' window. The main window title is 'Radon\_subLODroot2.tr ~ Potential'. The 'General' tab is selected. A 'Post-Hoc Comparisons' dialog box is open, showing the 'Factors' list with 'Potential' selected. The 'Post-Hoc' list is empty. The 'Type' is set to 'Tukey' and the 'Correction' is set to 'No Correction'. The 'Estimate confidence intervals' checkbox is unchecked. The 'OK' and 'Cancel' buttons are visible at the bottom of the dialog box. In the background, the 'Options' menu is open, with 'Post Hoc' highlighted. The 'Update Model' button is also visible.

2. You should see this window pop up

3. Select the variable in "Factors" and click the arrow

1. Click Post Hoc

# STEP 4

Linear Regression Model Explorer

Radon\_subLODroot2.tr ~ Potential

General Diagnostics Terms Added Variable

Preview

```
>.gui.work  
tential, da  
  
>Anova (.gu  
  
Anova Table  
  
Response: I  
  
Potential  
Residuals  
----  
Signif. co
```

Post-Hoc Comparisons

Factors

Post-Hoc  
Potential

Type: Tukey

Estimate confidence intervals

Correction: single-step

OK Cancel

Update Model

Run Reset Cancel

1. Now you should see the variable here

2. Select Tukey

3. Check Estimate CI

4. Select Correction as single-step

5. Click OK

# STEP 5

Linear Regression Model Explorer

Radon\_subLODroot2.tr ~ Potential

General Diagnostics Terms Added Variable

Preview

```
>.gui.working.env$model.lm <- lm(formula=Radon_subLODroot2.tr ~
potential, data=.gui.working.env$dat, na.action=na.omit)

>Anova(.gui.working.env$model.lm, type='II')

Anova Table (Type II tests)

Response: Radon_subLODroot2.tr
      Sum Sq  Df F value    Pr(>F)
Potential  37.47   2  21.812 5.073e-10 ***
Residuals 971.38 1131
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

>summarylm(.gui.working.env$model.lm)

Call:
lm(formula = Radon_subLODroot2.tr ~ Potential, data = .gui.worki
.env$dat,
    na.action = na.omit)
```

Options

Post Hoc

Tests

Plots

Means

Export

Update Model

1. Click Run

Run Reset Cancel



## STEP 6

Now you should see a lot of output in the Console, with codes in red that separate the results into four parts:

```
> Anova(model.lm,type='II')
```

1. Results of the one-way ANOVA

```
> summarylm(model.lm)
```

2. Results of the linear regression

```
> summary(glht(model.lm,linfct=mcp('Potential'='Tukey')),test=adjusted("single-step"))
```

3. Results of the Tukey-Kramer test

```
> confint(glht(model.lm,linfct=mcp('Potential'='Tukey')))
```

4. Confidence intervals of the group differences from the Tukey-Kramer test

# OUTPUT BREAK-DOWN - I

```
> Anova(model.lm, type='II')
Anova Table (Type II tests)

Response: Radon subLODroot2.tr

      Sum Sq   Df F value    Pr(>F)
Potential  37.47    2  21.812 5.073e-10 ***
Residuals 971.38 1131
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

The first part of the output is the one-way ANOVA result

You can get the Mean Squared Error by dividing Sum Sq by Df for Potential and Residuals respectively

# OUTPUT BREAK-DOWN - 2

```
> summarylm(model.lm)
```

Call:

```
lm(formula = Radon_subLODroot2.tr ~ Potential, data = dat, na.action = na.omit)
```

Residuals:

Min	1Q	Median	3Q	Max
-1.6826	-0.6864	-0.1111	0.5757	3.6442

Coefficients:

	Estimate	Std. Error	t value	Pr(> t )	
(Intercept)	3.98540	0.04511	88.342	< 2e-16	***
PotentialMOD	-0.38363	0.07069	-5.427	7.01e-08	***
PotentialHIGH	0.05807	0.06380	0.910	0.363	

---

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.9268 on 1131 degrees of freedom

Multiple R-squared: 0.03714, Adjusted R-squared: 0.03544

F-statistic: 21.81 on 2 and 1131 DF, p-value: 5.073e-10

The second part of the output is the simple linear regression

MOD compared with LOW

HIGH compared with LOW

How is the F-stats and P-value compared with results from the last page?

# OUTPUT BREAK-DOWN - 3

The third part of the output is the Tukey-Kramer test for the pair-wise comparison

```
> summary(glht(model.lm, linfct=mcp('Potential'="Tukey")), test=adjusted("single-step"))
```

Simultaneous Tests for General Linear Hypotheses

Multiple Comparisons of Means: Tukey Contrasts

Fit: lm(formula = Radon\_subLODroot2.tr ~ Potential, data = dat, na.action = na.omit)

Linear Hypotheses:

	Estimate	Std. Error	t value	Pr(> t )	
MOD - LOW == 0	-0.38363	0.07069	-5.427	<1e-04	***
HIGH - LOW == 0	0.05807	0.06380	0.910	0.633	
HIGH - MOD == 0	0.44170	0.07069	6.249	<1e-04	***

---

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

(Adjusted p values reported -- single-step method)

How is this compared with results in the linear regression?

# OUTPUT BREAK-DOWN - 4

The fourth part of the output is also from the Tukey-Kramer test

```
> confint(glht(model.lm, linfct=mcp('Potential'="Tukey")))
```

Simultaneous Confidence Intervals

Multiple Comparisons of Means: Tukey Contrasts

Fit: lm(formula = Radon\_subLODroot2.tr ~ Potential, data = dat, na.action = na.omit)

Quantile = 2.3455

95% family-wise confidence level

Linear Hypotheses:

	Estimate	lwr	upr
MOD - LOW == 0	-0.38363	-0.54943	-0.21783
HIGH - LOW == 0	0.05807	-0.09157	0.20772
HIGH - MOD == 0	0.44170	0.27590	0.60751

Instead of standard errors and p-values, it gives you the 95% confidence intervals