

HANDOUT Two-Way ANOVA with Interaction

P.B. Matheson adopted from S.M.Liao

STAT 230

Runners Data

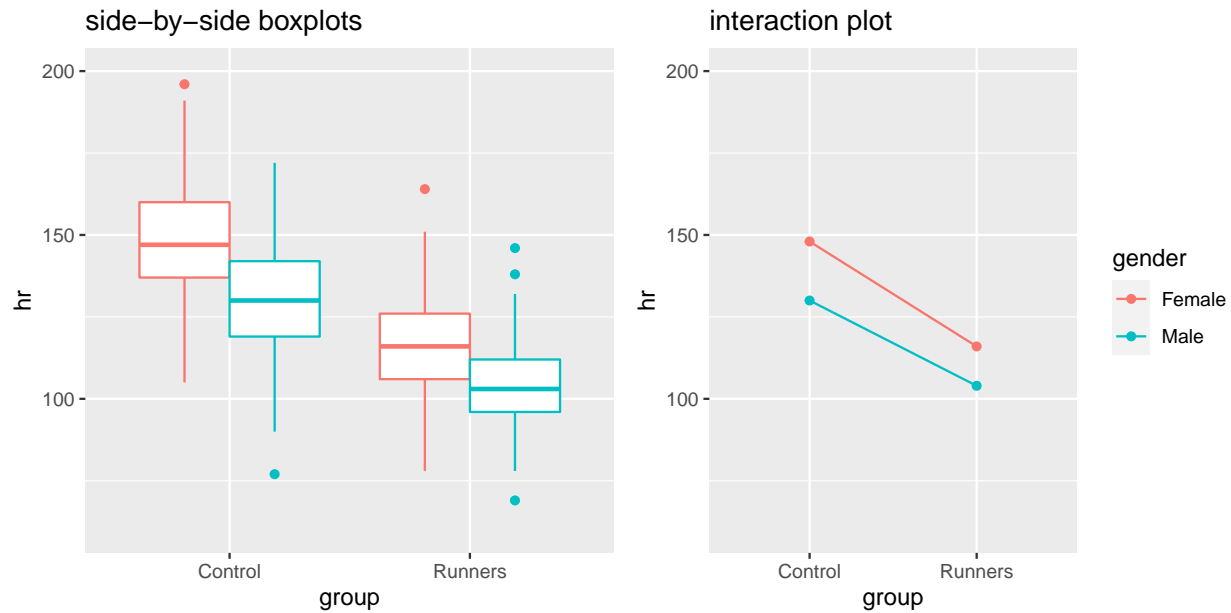
Recall that we used this dataset on the heart rates `hr` after six minutes of exercise on a treadmill. The two factors are: `group` (Control vs. Runners) and `gender` (Female vs. Male), and we want to see whether or not heart rates differed on average across either groups or genders.

| ## | group | gender | min | Q1 | median | Q3 | max | mean | sd | n | missing |
|------|---------|--------|-----|-----|--------|-----|-----|--------|--------|-----|---------|
| ## 1 | Control | Female | 105 | 137 | 147 | 160 | 196 | 148.00 | 16.271 | 200 | 0 |
| ## 2 | Runners | Female | 78 | 106 | 116 | 126 | 164 | 115.98 | 15.972 | 200 | 0 |
| ## 3 | Control | Male | 77 | 119 | 130 | 142 | 172 | 130.00 | 17.100 | 200 | 0 |
| ## 4 | Runners | Male | 69 | 96 | 103 | 112 | 146 | 103.97 | 12.499 | 200 | 0 |

There are many different ways to get interaction plots. The lab for this week has another method of plotting and later in this handout is a simpler way as well.

Note that the side by side graphs is obtained with `grid.arrange`.

```
## Chunk options: fig.width=8, fig.height=4, message=FALSE ##
bxplot <- gf_boxplot(hr ~ group, color = ~ gender, data = Runners,
                    title = "side-by-side boxplots", show.legend = FALSE) %>%
  gf_lims(y = c(60, 200))
interplot <- gf_point(hr ~ group, color = ~ gender, data = Runners,
                     group = ~ gender, stat = "summary", title = "interaction plot") %>%
  gf_line(hr ~ group, color = ~ gender, data = Runners,
          group = ~ gender, stat = "summary") %>%
  gf_lims(y = c(60, 200))
grid.arrange(bxplot, interplot, ncol = 2) ## Need to load the package `gridExtra`
```



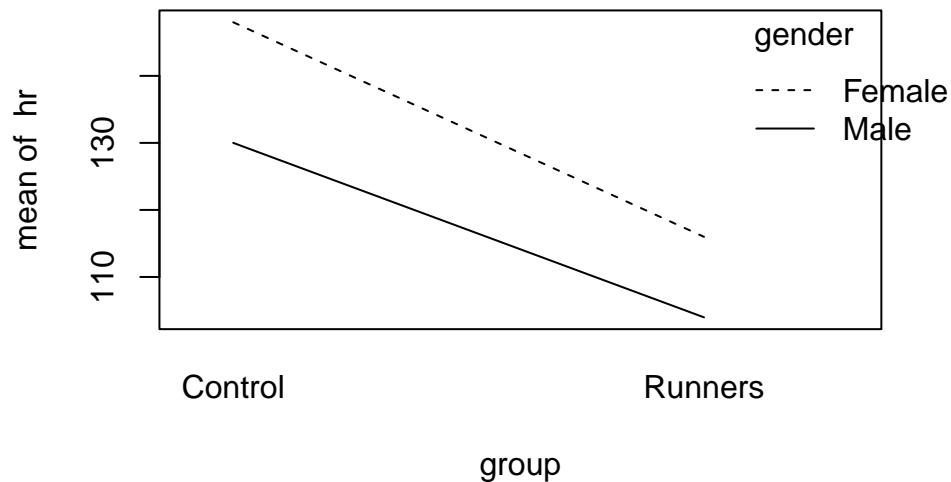
From the introductory lecture on two-way ANOVA we learned that the difference in average heart rates between genders was significant at $\alpha = 0.05$ (gender main effect) as well as the difference between runners and non-runners (group main effect).

QUESTION: Using the *left* plot (side-by-side boxplot), do you think there is an **INTERACTION** effect?

QUESTION: Using the *right* plot (**interaction plot**), do you think there is an **INTERACTION** effect?

Here is another way (easier way?) to create an interaction plot in R via the function `interaction.plot()`.

```
## Syntax: with(Dataset, interaction.plot(factorA, factorB, Response)) ##
with(Runners, interaction.plot(group, gender, hr))
```



Fit a two-way ANOVA model WITH interaction

Now let's include the interaction term using *. Again, we can fit an ANOVA model via either `lm()` or `aov()`.

```
mod2 <- lm(hr ~ group*gender, data = Runners)      #added interaction term by using *
anova(mod2)
```

```
aovmod2 <- aov(hr ~ group*gender, data = Runners)
summary(aovmod2)
```

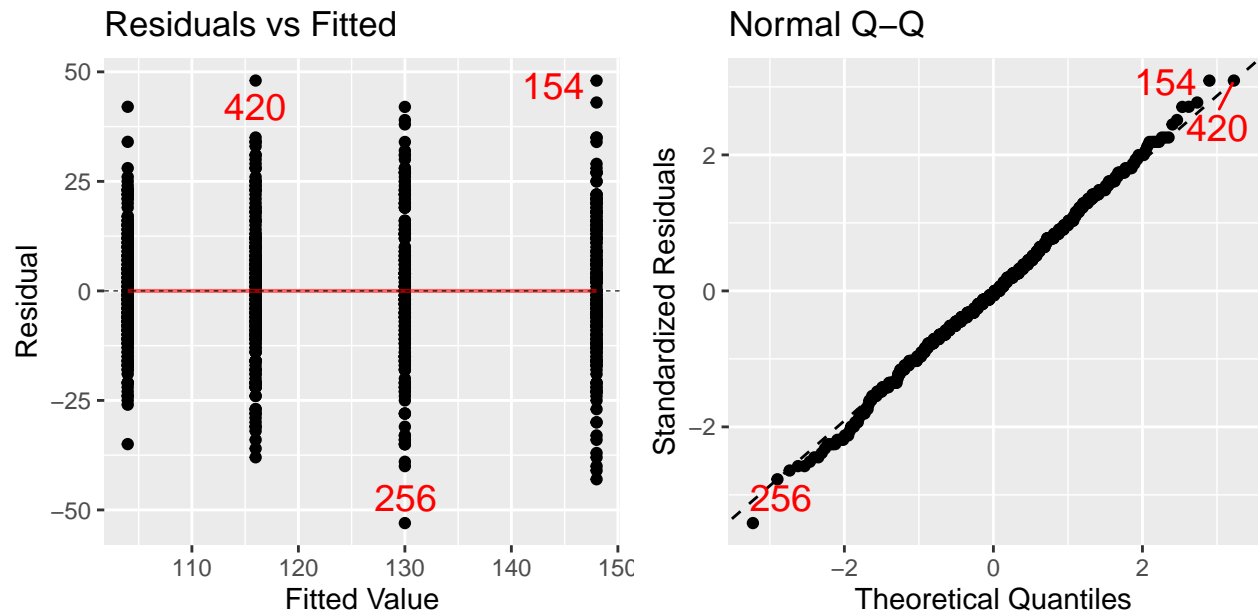
```
##           Df Sum Sq Mean Sq F value Pr(>F)
## group      1 168432  168432   695.65 <2e-16 ***
## gender     1  45030   45030   185.98 <2e-16 ***
## group:gender 1   1794    1794     7.41 0.0066 **
## Residuals 796 192730     242
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
model.tables(aovmod2)      ## estimated parameters
```

```
## Tables of effects
##
## group
## group
## Control Runners
##   14.51 -14.51
##
## gender
## gender
```

```
## Female    Male
## 7.502 -7.502
##
## group:gender
##      gender
## group   Female   Male
## Control 1.4975 -1.4975
## Runners -1.4975 1.4975
```

- Check the conditions:



QUESTION: Do you think the conditions check out based on the plots?

Return to the ANOVA table. There are now three F-statistics and the third one is 7.41.

QUESTION: What set of hypotheses is this used to test? What would your conclusion be here?

Note that having a **significant interaction term** means **interpretations of main effects cannot be done without including interaction**; in other words, when interaction is present, it might not be meaningful to interpret the main effects alone. Instead, we focus on inference about *conditional effects* - see more details below.

QUESTION: Do you need to use multiple comparisons procedures here to find where the differences are?

Let's see what happens if we use Tukey's HSD in this setting to identify differences when we have an *interaction* term in the model.

```
TukeyHSD(mod2)
```

```
##    Tukey multiple comparisons of means
##      95% family-wise confidence level
##
## Fit: aov(formula = x)
##
## $group
##              diff      lwr      upr p adj
## Runners-Control -29.02 -31.18 -26.86    0
##
## $gender
##              diff      lwr      upr p adj
## Male-Female -15.005 -17.165 -12.845    0
##
## $'group:gender'
##              diff      lwr      upr p adj
## Runners:Female-Control:Female -32.015 -36.021 -28.0091    0
## Control:Male-Control:Female   -18.000 -22.006 -13.9941    0
## Runners:Male-Control:Female   -44.025 -48.031 -40.0191    0
## Control:Male-Runners:Female    14.015  10.009  18.0209    0
## Runners:Male-Runners:Female   -12.010 -16.016  -8.0041    0
## Runners:Male-Control:Male     -26.025 -30.031 -22.0191    0
```

QUESTION: How do these comparisons differ from the comparisons we did before? (What is added?) What do these results show you?

To make inference about the *Runners* effect (compared to *non-runners*), we would need to look at that *conditioning on* (separately for) different 'genders'.

- For Females: We are 95% confident that the average heart rate of runners is between 28 and 36 bpm *lower* than that of non-runners; the estimated difference between their mean heart rate is 32 bpm, which is quite large as it's 205.70379% as big as the typical size of the difference between two females.
- For Males: We are 95% confident that the average heart rate of runners is between 22 and 30 bpm *lower* than that of non-runners; the estimated difference between their mean heart rate is 26 bpm, which is also large as it's 167.13433% as big as the typical size of the difference between two males.

You can make inference about the *gender* effect using the same way.

Note that this example was a 2x2 ANOVA with interaction with **balanced** cells (each cell had 200 observations) - which is a relatively simple case in two-way ANOVA.

Recall that an ANOVA model is equivalent to a regression model with indicator variables - this doesn't only apply to one-way ANOVA model; it works for multi-way ANOVA models too!

```
msummary(mod2)
```

```
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      148.00      1.10  134.51  <2e-16 ***
## groupRunners     -32.01      1.56  -20.57  <2e-16 ***
## genderMale       -18.00      1.56  -11.57  <2e-16 ***
## groupRunners:genderMale  5.99      2.20   2.72  0.0066 **
##
## Residual standard error: 15.6 on 796 degrees of freedom
## Multiple R-squared:  0.528, Adjusted R-squared:  0.526
## F-statistic: 296 on 3 and 796 DF, p-value: <2e-16
```

Fitted Model:

As for the ANOVA output, it's difficult to interpret either coefficients for the main effects alone, WITHOUT interpreting the coefficient of the interaction term when the interaction term is significant. Again, you would want to interpret those main effects *conditionally*.