Multiple Comparisons

Prof. Lauren Perry

After the ANOVA

- $H_0 : \mu_1 = \mu_2 = \cdots = \mu_k$
- $H_A : \mu_i \neq \mu_j$ for at least one pair (i, j)

We've rejected the null hypothesis... now what?

Multiple Comparisons

Why don't we start with pairwise comparisons?

- ▶ Type I Error Rate

Test Corrections

- Bonferroni correction
- Tukey's HSD
- Benjamini-Hochberg
- . . . and many more!

Tukey's HSD (honest significant difference)

▶ This method is a good balance of dealing with Type I error without being too aggressive.

▶ Essentially a t-test, but one that corrects for family-wise error rate (multiple comparisons).

▶ Compare absolute t values, $|t| = \left| \frac{\bar{x}_i - \bar{x}_j}{SE} \right|$ to the *studentized range distribution*

  ▶ Adjusted p-value
  ▶ Family-wise confidence intervals

studentized range distribution

Let $X_1, \ldots, X_n$ be iid $N(\mu, \sigma^2)$ and let R be the range of X.

Then $q_{n,\nu} = R/\hat{\sigma}$ has the studentized range distribution where ν is the degrees of freedom used in estimating $\hat{\sigma}$.

Example: Chick Weights

```
anova(aov(weight ~ feed, chickwts))
```

```
## Analysis of Variance Table
##
## Response: weight
##           Df Sum Sq Mean Sq F value    Pr(>F)
## feed       5 231129   46226  15.365 5.936e-10 ***
## Residuals 65 195556    3009
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Example: Chick Weights

```
TukeyHSD(aov(weight ~ feed, chickwts))
```

```
##   Tukey multiple comparisons of means
##     95% family-wise confidence level
##
## Fit: aov(formula = weight ~ feed, data = chickwts)
##
## $feed
##                         diff         lwr        upr      p adj
## horsebean-casein  -163.383333 -232.346876  -94.41979 0.0000000
## linseed-casein    -104.833333 -170.587491  -39.07918 0.0002100
## meatmeal-casein    -46.674242 -113.906207   20.55772 0.3324584
## soybean-casein     -77.154762 -140.517054  -13.79247 0.0083653
## sunflower-casein     5.333333  -60.420825   71.08749 0.9998902
## linseed-horsebean   58.550000  -10.413543  127.51354 0.1413329
## meatmeal-horsebean 116.709091   46.335105  187.08308 0.0001062
## soybean-horsebean   86.228571   19.541684  152.91546 0.0042167
```

Example: Chick Weights

```
plot(TukeyHSD(aov(weight ~ feed, chickwts)))
```

95% family–wise confidence level