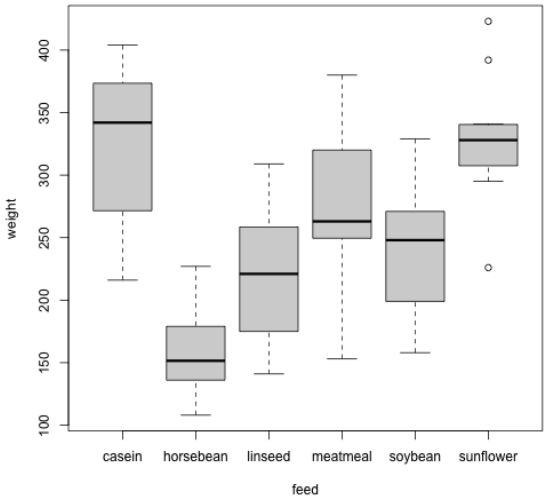
Chapter 10

Example: This example uses the **chickwts** dataset again. This dataset include 71 observations of the growth rate of chickens using a variety of feed supplements. The boxplot is produced again below.

```
data(chickwts)
boxplot(weight~feed,data=chickwts)
```



Using the theory from Chapter 10, we are ready to test if there is evidence of a difference in mean weights of chicks for the six types of feed. Our ANOVA F-test hypotheses will be: $H_0 : \mu_1 = \mu_2 = \mu_3 = \mu_4 = \mu_5 = \mu_6$ vs. H_a : at least one μ_i differs $(i = 1, 2, 3, 4, 5, 6)$.

Let's run this F-test at $\alpha = 0.01$.

```
fit = aov(weight~feed,data=chickwts)
summary(fit)

##             Df Sum Sq Mean Sq F value   Pr(>F)
## feed         5 231129   46226   15.37 5.94e-10 ***
## Residuals   65 195556    3009
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

The resulting F-test statistic is $F_{test} = 15.37$ $(df_1 = 5, df_2 = 65)$ and a p-value of $5.94 \cdot 10^{-10}$. At a significance level of $\alpha = 0.01$, we would reject H_0 and conclude that at least one of the true mean weights differs by diet.

Which μ_i differ? We can answer this by running all pairwise comparisons of μ_i and μ_j $(i \neq j)$ and making a multiple comparisons adjustment. Let's use a family-wise error rate of $\alpha = 0.01$.

```
TukeyHSD(fit,conf.level = 0.99)

##   Tukey multiple comparisons of means
##     99% family-wise confidence level
##
## Fit: aov(formula = weight ~ feed, data = chickwts)
##
## $feed
##                             diff         lwr        upr     p adj
## horsebean-casein    -163.383333 -245.964363 -80.802304 0.0000000
## linseed-casein      -104.833333 -183.571256 -26.095411 0.0002100
## meatmeal-casein      -46.674242 -127.181777  33.833292 0.3324584
## soybean-casein       -77.154762 -153.028522  -1.281001 0.0083653
## sunflower-casein       5.333333  -73.404589  84.071256 0.9998902
## linseed-horsebean     58.550000  -24.031030 141.131030 0.1413329
## meatmeal-horsebean   116.709091   32.439113 200.979069 0.0001062
## soybean-horsebean     86.228571    6.373743 166.083400 0.0042167
## sunflower-horsebean  168.716667   86.135637 251.297696 0.0000000
## meatmeal-linseed      58.159091  -22.348444 138.666626 0.1276965
## soybean-linseed       27.678571  -48.195189 103.552332 0.7932853
## sunflower-linseed    110.166667   31.428744 188.904589 0.0000884
## soybean-meatmeal     -30.480519 -108.189144  47.228105 0.7391356
## sunflower-meatmeal    52.007576  -28.499959 132.515111 0.2206962
## sunflower-soybean     82.488095    6.614335 158.361856 0.0038845
```

There are several statistically significant differences in mean chick weights at $\alpha = 0.01$.

Casein differs from horsebean, linseed, and soybean.
Horsebean differs from casein, meatmeal, soybean, and sunflower.
Linseed differs from casein and sunflower.
Meatmeal differs from horsebean.
Sunflower differs from horsebean, linseed, and soybean.
Soybean differs from sunflower, horsebean, and casein.

```
plot(TukeyHSD(fit,conf.level = 0.99))
```