The Statistical Sleuth in R: Chapter 6

Linda Loi Ruobing Zhang Kate Aloisio Nicholas J. Horton*

July 25, 2024

Contents

1 Introduction

This document is intended to help describe how to undertake analyses introduced as examples in the Third Edition of the *Statistical Sleuth* (2013) by Fred Ramsey and Dan Schafer. More information about the book can be found at `http://www.proaxis.com/~panorama/home.htm`. This file as well as the associated `knitr` reproducible analysis source file can be found at `http://www.math.smith.edu/~nhorton/sleuth3`.

This work leverages initiatives undertaken by Project MOSAIC (`http://www.mosaic-web.org`), an NSF-funded effort to improve the teaching of statistics, calculus, science and computing in the undergraduate curriculum. In particular, we utilize the mosaic package, which was written to simplify the use of R for introductory statistics courses. A short summary of the R needed to teach introductory statistics can be found in the mosaic package vignette (`http://cran.r-project.org/web/packages/mosaic/vignettes/MinimalR.pdf`).

To use a package within R, it must be installed (one time), and loaded (each session). The package can be installed using the following command:

*Department of Mathematics and Statistics, Smith College, nhorton@smith.edu

```
> install.packages('mosaic')                # note the quotation marks
```

Once this is installed, it can be loaded by running the command:

```
> require(mosaic)
```

This needs to be done once per session.

In addition the data files for the *Sleuth* case studies can be accessed by installing the `Sleuth3` package.

```
> install.packages('Sleuth3')               # note the quotation marks
```

```
> require(Sleuth3)
```

We also set some options to improve legibility of graphs and output.

```
> trellis.par.set(theme=col.mosaic())  # get a better color scheme for lattice
> options(digits=3)
```

The specific goal of this document is to demonstrate how to calculate the quantities described in Chapter 6: Linear Combinations and Multiple Comparisons of Means using R.

2 Discrimination Against the Handicapped

Do equivalent candidates with the same qualifications but different disabilities get treated differentially? This is the question addressed in case study 6.1 in the *Sleuth*.

2.1 Summary statistics and graphical display

We begin by reading the data and summarizing the variables.

```
> case0601$Handicap = relevel(case0601$Handicap, ref="Amputee")
> summary(case0601)

     Score            Handicap
 Min.   :1.40   Amputee   :14
 1st Qu.:3.70   Crutches  :14
 Median :5.05   Hearing   :14
 Mean   :4.93   None      :14
 3rd Qu.:6.10   Wheelchair:14
 Max.   :8.50

> favstats(Score ~ Handicap, data=case0601)
```

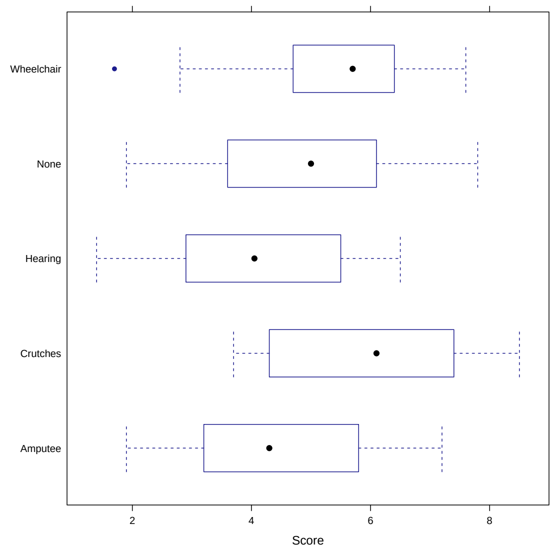
```
    Handicap min   Q1 median   Q3 max mean   sd  n missing
1    Amputee 1.9 3.30   4.30 5.72 7.2 4.43 1.59 14       0
2   Crutches 3.7 4.50   6.10 7.15 8.5 5.92 1.48 14       0
3    Hearing 1.4 3.02   4.05 5.30 6.5 4.05 1.53 14       0
4       None 1.9 3.73   5.00 6.05 7.8 4.90 1.79 14       0
5 Wheelchair 1.7 4.73   5.70 6.35 7.6 5.34 1.75 14       0
```

A total of 70 undergraduate students from a U.S. university were randomly assigned to view the tapes, 14 to each tape. The five kinds of tapes are: *None*, *Amputee*, *Crutches*, *Hearing* and *Wheelchair*. After reviewing the tape, each subject rated the qualifications of the application on 0-10 scale. Among the five handicap conditions, the *Crutches* group gave the highest mean score, while the *Hearing* group gave the lowest mean score. This is summarized on page 150 and in Display 6.1 of the *Sleuth*.

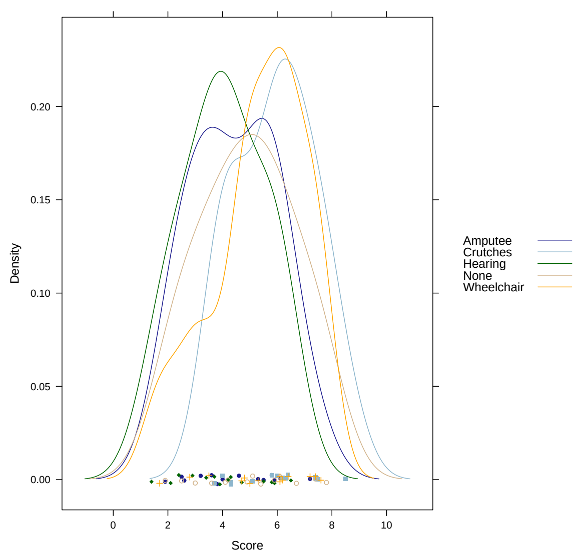
```
> with(subset(case0601, Handicap=="None"), stem(Score, scale=2))


  The decimal point is at the |

  1 | 9
  2 | 5
  3 | 06
  4 | 129
  5 | 149
  6 | 17
  7 | 48

> with(subset(case0601, Handicap=="Amputee"), stem(Score, scale=2))


  The decimal point is at the |

  1 | 9
  2 | 56
  3 | 268
  4 | 06
  5 | 3589
  6 | 1
  7 | 2

> with(subset(case0601, Handicap=="Crutches"), stem(Score, scale=1))


  The decimal point is at the |

  3 | 7
```

```
  4 | 033
  5 | 18
  6 | 0234
  7 | 445
  8 | 5

> with(subset(case0601, Handicap=="Hearing"), stem(Score, scale=2))


  The decimal point is at the |

  1 | 4
  2 | 149
  3 | 479
  4 | 237
  5 | 589
  6 | 5

> with(subset(case0601, Handicap=="Wheelchair"), stem(Score, scale=2))


  The decimal point is at the |

  1 | 7
  2 | 8
  3 | 5
  4 | 78
  5 | 03
  6 | 1124
  7 | 246
```

```
> bwplot(Handicap ~ Score, data=case0601)
```




```
> densityplot(~ Score, groups=Handicap, auto.key=TRUE, data=case0601)
```



The stem plots show the applicant qualification scores given by objectives. The boxplots and the density plots show that all the distributions are approximately normally distributed.

2.2 One-way ANOVA

First we fit the one way analysis of variance (ANOVA) model, using all of the groups. This corresponds to the interpretations on page 151.

```
> anova(lm(Score ~ Handicap, data=case0601))
```

```
Analysis of Variance Table

Response: Score
          Df Sum Sq Mean Sq F value Pr(>F)
Handicap   4   30.5    7.63    2.86   0.03
Residuals 65  173.3    2.67
```

The p-value provides some evidence that subjects rate qualifications differently according to handicap status.

By default, the use of the linear model (regression) function displays the pairwise differences between the first group and each of the other groups. Note that the overall test of the model is the same.

```
> summary(lm(Score ~ Handicap, data=case0601))


Call:
lm(formula = Score ~ Handicap, data = case0601)

Residuals:
   Min     1Q Median     3Q    Max
-3.643 -1.209  0.114  1.329  2.900

Coefficients:
                   Estimate Std. Error t value Pr(>|t|)
(Intercept)           4.429      0.436   10.15    5e-15
HandicapCrutches      1.493      0.617    2.42    0.018
HandicapHearing      -0.379      0.617   -0.61    0.542
HandicapNone          0.471      0.617    0.76    0.448
HandicapWheelchair    0.914      0.617    1.48    0.143

Residual standard error: 1.63 on 65 degrees of freedom
Multiple R-squared:  0.15, Adjusted R-squared:  0.0974
F-statistic: 2.86 on 4 and 65 DF,  p-value: 0.0301
```

The reference group here is *Amputee*, followed by *Crutches*, *Hearing*, *None* and *Wheelchair*.

Another way of viewing these results is through a model table, which displays the differences between the grand mean and the group means.

```
> model.tables(aov(Score ~ Handicap, data=case0601))

Tables of effects

 Handicap
Handicap
   Amputee   Crutches    Hearing       None Wheelchair
   -0.5000     0.9929    -0.8786    -0.0286     0.4143
```

Or by:

```
> mean(Score ~ Handicap, data=case0601)-mean(~ Score, data=case0601)

   Amputee   Crutches    Hearing       None Wheelchair
   -0.5000     0.9929    -0.8786    -0.0286     0.4143
```

2.3 Contrasts and linear combination

The Tukey-Kramer test is a reasonable method for these data. We can use this to verify the calculation on page 151.

```
> TukeyHSD(aov(lm(Score ~ Handicap, data=case0601)), "Handicap", ordered=TRUE, c(0,1,-1,0,0), 

  Tukey multiple comparisons of means
    95% family-wise confidence level
    factor levels have been ordered

Fit: aov(formula = lm(Score ~ Handicap, data = case0601))

$Handicap
                     diff    lwr  upr p adj
Amputee-Hearing     0.379 -1.353 2.11 0.972
None-Hearing        0.850 -0.882 2.58 0.644
Wheelchair-Hearing  1.293 -0.439 3.02 0.235
Crutches-Hearing    1.871  0.140 3.60 0.028
None-Amputee        0.471 -1.260 2.20 0.940
Wheelchair-Amputee  0.914 -0.817 2.65 0.578
Crutches-Amputee    1.493 -0.239 3.22 0.123
Wheelchair-None     0.443 -1.289 2.17 0.952
Crutches-None       1.021 -0.710 2.75 0.469
Crutches-Wheelchair 0.579 -1.153 2.31 0.881
```

Based on the Tukey-Kramer procedure, the difference is estimated to be higher for the *Crutches* tapes.

Next, we calculate the comparison of *Amputee/Hearing* to *Crutches/Wheelchair*.

```
> require(gmodels)
> fit.contrast(lm(Score ~ Handicap, data=case0601), "Handicap", c(-1, 1, -1, 0, 1), conf.int=0

                           Estimate Std. Error t value Pr(>|t|) lower CI
Handicap c=( -1 1 -1 0 1 )     2.79      0.873    3.19  0.00218     1.04
                           upper CI
Handicap c=( -1 1 -1 0 1 )     4.53
```

The results indicate a statistically significant difference between the average scores given to the *Wheelchair* and *Crutches* handicaps and the average scores given to the *Amputee* and *Hearing* handicaps.

To verify the calculations on page 155 we used the following contrast:

```
> fit.contrast(lm(Score ~ Handicap, data=case0601), "Handicap", c(-0.5, 0.5, -0.5, 0, 0.5), co

                                   Estimate Std. Error t value Pr(>|t|)
Handicap c=( -0.5 0.5 -0.5 0 0.5 )     1.39      0.436    3.19  0.00218
                                   lower CI upper CI
Handicap c=( -0.5 0.5 -0.5 0 0.5 )    0.521     2.26
```

Other multiple comparison procedures could also be implemented. The following shows the calculation on page 164.

```
> require(agricolae)

Loading required package:  agricolae

> LSD.test(aov(lm(Score ~ Handicap, data=case0601)), "Handicap")   # LSD
> HSD.test(aov(lm(Score ~ Handicap, data=case0601)), "Handicap")   # Tukey-Kramer
> LSD.test(aov(lm(Score ~ Handicap, data=case0601)), "Handicap", p.adj=c("bonferroni"))  # Bon
> scheffe.test(aov(lm(Score ~ Handicap, data=case0601)), "Handicap")    # Scheffe
```

The "Significant Difference" in each test result is the "95% interval half-width" described in the book.

3 Pre-existing Preference of Fish

Was Charles Darwin right that sexual selection is driven by females? This is the question addressed in case study 6.2 in the *Sleuth*.

3.1 Summary statistics and graphical display

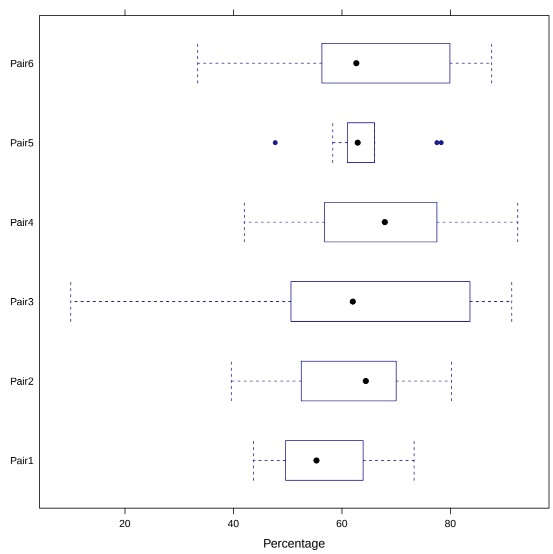
We begin by reading the data and summarizing the variables.

```
> summary(case0602)

   Percentage      Pair        Length
 Min.   :10.0   Pair1:16   Min.   :28.0
 1st Qu.:53.1   Pair2:14   1st Qu.:31.0
 Median :61.5   Pair3:17   Median :34.0
 Mean   :62.1   Pair4:14   Mean   :32.8
 3rd Qu.:71.8   Pair5: 9   3rd Qu.:34.0
 Max.   :92.4   Pair6:14   Max.   :35.0

> favstats(Percentage ~ Pair, data=case0602)
```

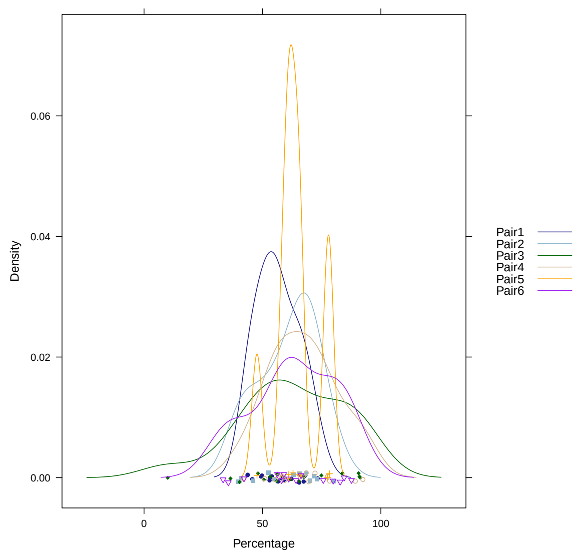
```
   Pair  min   Q1 median   Q3  max mean    sd  n missing
1 Pair1 43.7 49.7   55.3 63.1 73.3 56.4  9.02 16       0
2 Pair2 39.6 53.1   64.4 69.6 80.2 60.9 12.48 14       0
3 Pair3 10.0 50.6   62.0 83.6 91.3 62.4 22.29 17       0
4 Pair4 42.0 57.2   67.9 76.2 92.4 67.0 14.33 14       0
5 Pair5 47.7 61.0   62.9 66.0 78.3 64.2  9.41  9       0
6 Pair6 33.4 56.7   62.7 78.9 87.6 63.3 17.68 14       0
```

A total of 84 female fish were involved in this experiment, which is shown on page 153.

```
> bwplot(Pair ~ Percentage, data=case0602)
```




```
> densityplot(~ Percentage, groups=Pair, auto.key=TRUE, data=case0602)
```



Besides the distribution of pair 5, all distributions of other pairs are approximately normally distributed.

3.2 One-way ANOVA

First we fit the one way analysis of variance (ANOVA) model, using all of the groups:

```
> anova(lm(Percentage ~ Pair, data=case0602))

Analysis of Variance Table

Response: Percentage
          Df Sum Sq Mean Sq F value Pr(>F)
Pair       5    939     188    0.79   0.56
Residuals 78  18637     239
```

The p-value is not small, and does not provide much evidence that the mean percentage of time with the yellow-sword male significantly differed from one male pair to another back in the population.

By default, the use of the linear model (regression) function displays the pairwise differences between the first group and each of the other groups. Note that the overall test of the model is the same.

```
> summary(lm(Percentage ~ Pair, data=case0602))


Call:
lm(formula = Percentage ~ Pair, data = case0602)

Residuals:
```

```
   Min     1Q Median     3Q    Max
-52.43  -8.41   0.25  10.86  28.87

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)    56.41       3.86   14.60   <2e-16
PairPair2       4.48       5.66    0.79    0.431
PairPair3       6.02       5.38    1.12    0.267
PairPair4      10.59       5.66    1.87    0.065
PairPair5       7.80       6.44    1.21    0.229
PairPair6       6.93       5.66    1.22    0.224

Residual standard error: 15.5 on 78 degrees of freedom
Multiple R-squared:  0.048, Adjusted R-squared:  -0.0131
F-statistic: 0.786 on 5 and 78 DF,  p-value: 0.563
```

The reference group here is pair 1, followed by pairs 2-6. Another way of viewing these results is through a model table, which displays the differences between the grand mean and the group means.

```
> model.tables(aov(Percentage ~ Pair, data=case0602))

Tables of effects

 Pair
     Pair1  Pair2   Pair3  Pair4 Pair5  Pair6
    -5.722 -1.243  0.3008  4.871 2.083  1.207
rep 16.000 14.000 17.0000 14.000 9.000 14.000
```

Or by:

```
> mean(Percentage ~ Pair, data=case0602)-mean(~ Percentage, data=case0602)

 Pair1  Pair2  Pair3  Pair4  Pair5  Pair6
-5.722 -1.243  0.301  4.871  2.083  1.207
```

3.3 Contrasts and linear combination

We can calculate the values on page 152 and Display 6.5 on page 158 using contrasts.

```
> require(gmodels)
> lc = fit.contrast(lm(Percentage ~ Pair, data=case0602), "Pair", c(5, -3, 1, 3, -9, 3), conf.i

                         Estimate Std. Error t value Pr(>|t|) lower CI upper CI
Pair c=( 5 -3 1 3 -9 3 )    -25.1       54.8  -0.458    0.648     -134     83.9
```

```
> t=round(lc[, "t value"], 2); t

[1] -0.46

> pt(t, 78, lower.tail=TRUE)

[1] 0.323
```

The t-value is -0.46 and the one-sided p-value is 0.32.

```
> mean(mean(Percentage ~ Pair, data=case0602))

[1] 62.4

> t.test(mean(Percentage ~ Pair, data=case0602))


 One Sample t-test

data:  mean(Percentage ~ Pair, data = case0602)
t = 43, df = 5, p-value = 1e-07
alternative hypothesis: true mean is not equal to 0
95 percent confidence interval:
 58.6 66.1
sample estimates:
mean of x
     62.4
```

The estimated mean percentage of time spent with the yellow-sword male is 62.378%. The one-sided p-value< 0.0001, and the 95% confidence interval is (58.637%, 66.119%).