Module 6

Alex Schlesener

2023-09-24

Contents

Objective

- Conduct analyses on two different independent sample means:
  - Special forces soldiers who completed Standard training
  - Special forces soldiers who completed new training
- Examine scores of various criteria:
  - Maneuvering Efficacy
  - Sensor Operation Proficiency
  - Declarative Knowledge
- Summarize findings of the statistical analyses.

Description of data

A sample of 11 special forces soldiers were evaluated regarding scores on various criteria including Maneuvering Efficacy, Sensor Operation Proficiency, and Declarative Knowledge. (Higher scores are better.) Of the 11 soldiers, 6 completed Standard training, while 5 completed New training. The data is in the file `uav.csv`.

Using variables *Maneuver*, *Sensor*, and *Knowledge*, we compared the results of soldiers who completed Standard training and soldiers who completed New training.

Import data

```r
dat <- read.csv("Data/uav.csv")
str(dat)
```

```
## 'data.frame':    11 obs. of  4 variables:
##  $ Training : int  1 0 0 1 1 0 0 0 1 0 ...
##  $ Maneuver : int  29 22 19 28 20 20 18 17 25 14 ...
##  $ Sensor   : int  52 30 42 54 39 51 25 25 49 31 ...
##  $ Knowledge: int  20 14 16 24 18 19 20 24 28 17 ...
```

When using the `str()` function, we can see that there are 11 observations and 4 variables in the data frame. We will focus on *Maneuver*, *Sensor* and *Knowledge*, relative to *Training*.

Convert categorical variables to factors

In the present data, *Training* is a categorical variable, and should be converted into a factor. This way, we can examine what 0 and 1 represent.

```r
dat$Training <- factor(dat$Training)
levels(dat$Training)        # Confirm the levels of the diabetes factor
```

```
## [1] "0" "1"
```

```r
summary(dat)                # 355 No and 177 Yes
```

```
##  Training    Maneuver         Sensor        Knowledge
##  0:6      Min.   :14.00   Min.   :25.00   Min.   :14.00
##  1:5      1st Qu.:18.50   1st Qu.:30.50   1st Qu.:17.50
##           Median :20.00   Median :41.00   Median :20.00
##           Mean   :21.64   Mean   :39.91   Mean   :20.82
##           3rd Qu.:25.50   3rd Qu.:50.00   3rd Qu.:24.00
##           Max.   :29.00   Max.   :54.00   Max.   :29.00
```

Based on this information, we can see that 6 soldiers completed Standard training, whereas 5 soldier completed New training.

Examine the data

Maneuver Scores Calculate descriptive statistics and examine a boxplot of the outcome (Maneuver) by the categorical predictor (Training).

```r
aggregate(Maneuver ~ Training, dat, mean)       # mean in each group
```

```
##   Training Maneuver
## 1        0 18.33333
## 2        1 25.60000
```

```r
aggregate(Maneuver ~ Training, dat, median)     # median in each group
```

```
##   Training Maneuver
## 1        0     18.5
## 2        1     26.0
```

```r
aggregate(Maneuver ~ Training, dat, sd)          # standard deviation in each group
```

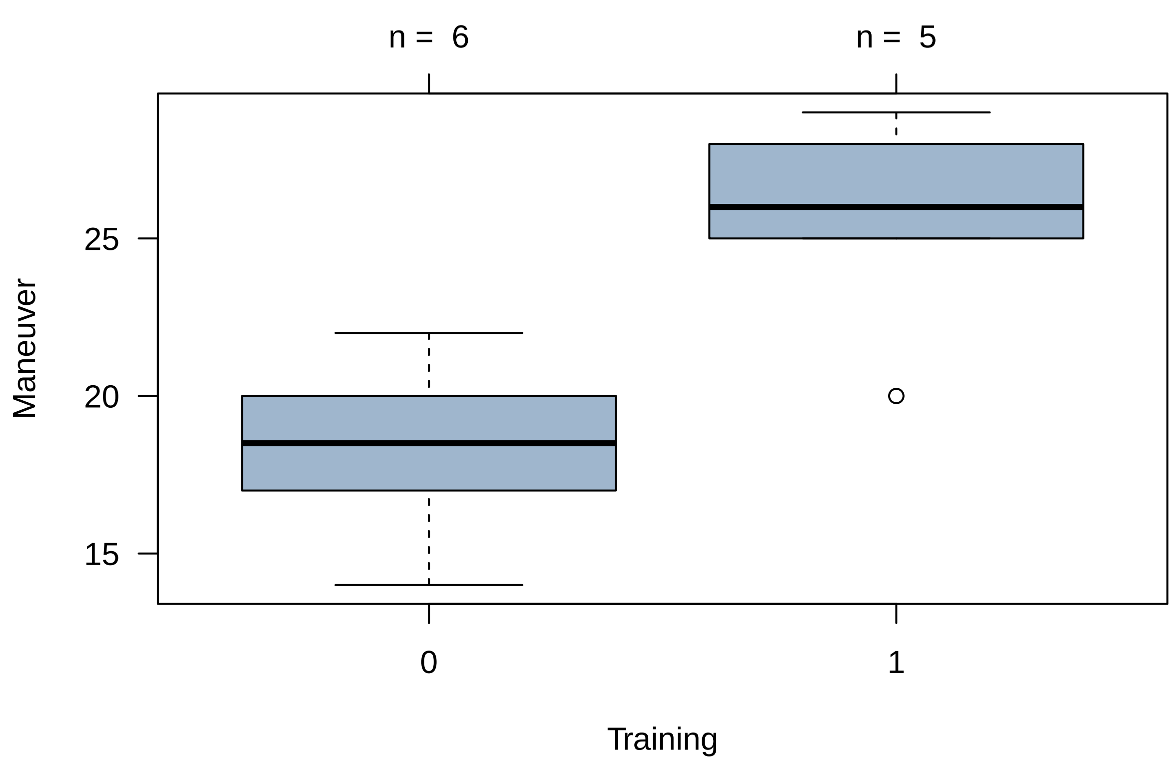
```
##   Training Maneuver
## 1        0 2.732520
## 2        1 3.507136
```

```r
aggregate(Maneuver ~ Training, dat, length)      # sample size in each group
```

```
##   Training Maneuver
## 1        0        6
## 2        1        5
```

```r
boxplot(Maneuver ~ Training, data = dat,
        col = "slategray3",
        las = 1,
        xlab = "Training",
        ylab = "Maneuver")

axis(3, at = 1:2, labels = paste("n = ", table(dat$Training)))
```



Sensor Scores Calculate descriptive statistics and examine a boxplot of the outcome (Sensor) by the categorical predictor (Training).

```
aggregate(Sensor ~ Training, dat, mean)        # mean in each group
```

```
##   Training Sensor
## 1        0     34
## 2        1     47
```

```
aggregate(Sensor ~ Training, dat, median)      # median in each group
```

```
##   Training Sensor
## 1        0   30.5
## 2        1   49.0
```

```
aggregate(Sensor ~ Training, dat, sd)          # standard deviation in each group
```

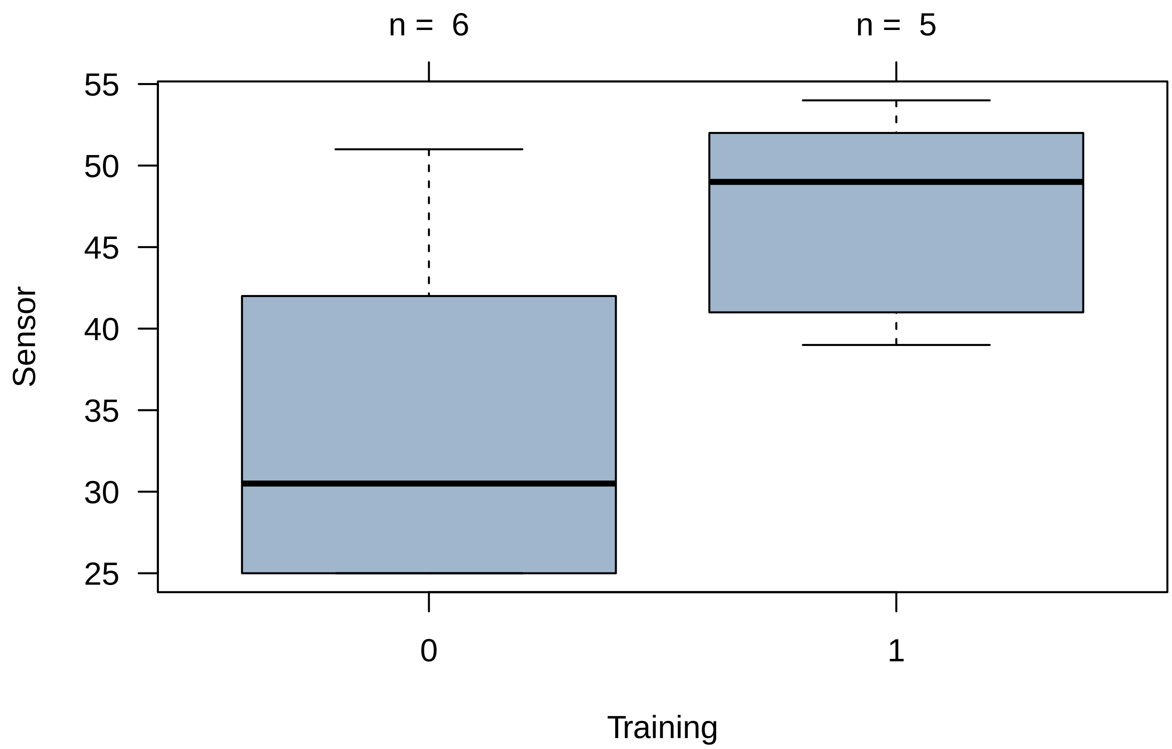
```
##   Training    Sensor
## 1        0 10.392305
## 2        1  6.670832
```

```
aggregate(Sensor ~ Training, dat, length)      # sample size in each group
```

```
##   Training Sensor
## 1        0      6
## 2        1      5
```

```
boxplot(Sensor ~ Training, data = dat,
        col = "slategray3",
        las = 1,
        xlab = "Training",
        ylab = "Sensor")

axis(3, at = 1:2, labels = paste("n = ", table(dat$Training)))
```



Knowledge Scores Calculate descriptive statistics and examine a boxplot of the outcome (Knowledge) by the categorical predictor (Training).

```
aggregate(Knowledge ~ Training, dat, mean)        # mean in each group
```

```
##   Training Knowledge
## 1        0  18.33333
## 2        1  23.80000
```

```
aggregate(Knowledge ~ Training, dat, median)      # median in each group
```

```
##   Training Knowledge
## 1        0        18
## 2        1        24
```

```
aggregate(Knowledge ~ Training, dat, sd)          # standard deviation in each group
```

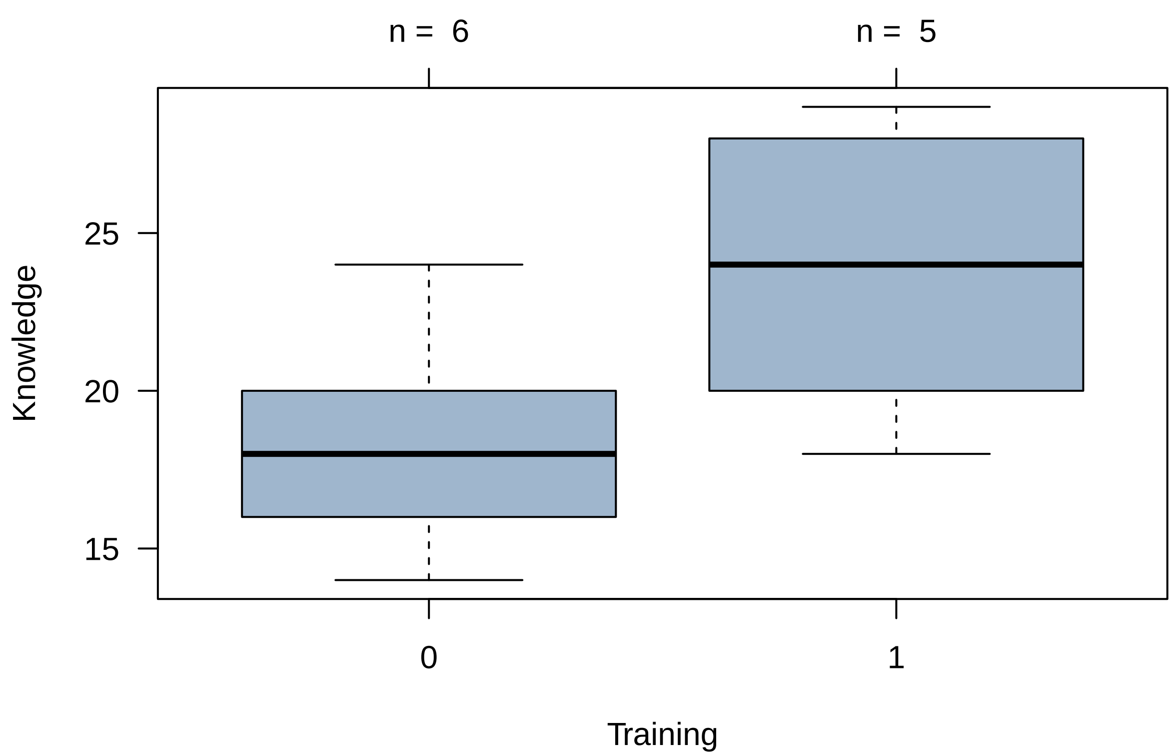
```
##   Training Knowledge
## 1        0  3.502380
## 2        1  4.816638
```

```r
aggregate(Knowledge ~ Training, dat, length)      # sample size in each group
```

```
##   Training Knowledge
## 1        0         6
## 2        1         5
```

```r
boxplot(Knowledge ~ Training, data = dat,
        col = "slategray3",
        las = 1,
        xlab = "Training",
        ylab = "Knowledge")

axis(3, at = 1:2, labels = paste("n = ", table(dat$Training)))
```



Conduct test on two independent sample means

We conducted a test on the differences between two independent means. This test was performed on three different variables: *Maneuver*, *Sensor*, and *Knowledge*.

```r
t.test(Maneuver ~ Training, data = dat, var.equal = T)
```

```
##
##  Two Sample t-test
```

```
##
## data:  Maneuver by Training
## t = -3.8702, df = 9, p-value = 0.003788
## alternative hypothesis: true difference in means between group 0 and group 1 is not equal to 0
## 95 percent confidence interval:
##  -11.51412  -3.01921
## sample estimates:
## mean in group 0 mean in group 1
##        18.33333        25.60000
```

```
t.test(Sensor ~ Training, data = dat, var.equal = T)
```

```
##
##  Two Sample t-test
##
## data:  Sensor by Training
## t = -2.4036, df = 9, p-value = 0.03966
## alternative hypothesis: true difference in means between group 0 and group 1 is not equal to 0
## 95 percent confidence interval:
##  -25.2348727  -0.7651273
## sample estimates:
## mean in group 0 mean in group 1
##              34              47
```

```
t.test(Knowledge ~ Training, data = dat, var.equal = T)
```

```
##
##  Two Sample t-test
##
## data:  Knowledge by Training
## t = -2.1815, df = 9, p-value = 0.05703
## alternative hypothesis: true difference in means between group 0 and group 1 is not equal to 0
## 95 percent confidence interval:
##  -11.13539   0.20206
## sample estimates:
## mean in group 0 mean in group 1
##        18.33333        23.80000
```

Check homogeneity of variance assumption

Use Levene's Test to ensure the validity of the test statistic via determining the homoscedasticity assumption.

```
library(car)
```

```
## Loading required package: carData
```

```
leveneTest(Maneuver ~ Training, dat, center = mean)
```

```
## Levene's Test for Homogeneity of Variance (center = mean)
##       Df F value Pr(>F)
## group  1  0.1779  0.683
##        9
```

```r
leveneTest(Sensor ~ Training, dat, center = mean)
```

```
## Levene's Test for Homogeneity of Variance (center = mean)
##       Df F value Pr(>F)
## group  1  1.2688 0.2891
##        9
```

```r
leveneTest(Knowledge ~ Training, dat, center = mean)
```

```
## Levene's Test for Homogeneity of Variance (center = mean)
##       Df F value Pr(>F)
## group  1  0.8954 0.3687
##        9
```

Summary

1A) To assess the effectiveness of the two types of special forces soldier Raven B operation training, an independent samples t test was conducted using maneuver scores as the dependent variable. The average impact for New Training (M = 25.60, SD = 3.51) was greater than that of Standard Training (M = 18.33, SD = 2.73) and this difference was statistically significant, $t(9)$ = -3.87, p = .004, Cohen's d = -2.34. The 95% confidence interval for the mean difference was [-11.51, -3.019]. Overall, New Training appears to be more effective than Standard Training at maneuver scores.

To assess the effectiveness of the two types of special forces soldier Raven B operation training, an independent samples t test was conducted using sensor scores as the dependent variable. The average impact for New Training (M = 47.00, SD = 6.67) was greater than that of Standard Training (M = 34.00, SD = 10.39) and this difference was statistically significant, $t(9)$ = -2.40, p = .04, Cohen's d = -1.46. The 95% confidence interval for the mean difference was [-25.24, -.765]. Overall, New Training appears to be more effective than Standard Training at sensor scores.

To assess the effectiveness of the two types of special forces soldier Raven B operation training, an independent samples t test was conducted using knowledge scores as the dependent variable. The average impact for New Training (M = 23.80, SD = 4.82) was greater than that of Standard Training (M = 18.33, SD = 3.50) and this difference was not statistically significant, $t(9)$ = -2.18, p = .057, Cohen's d = -1.32. The 95% confidence interval for the mean difference was [-11.14, .202]. Overall, New Training does not appear to be more effective than Standard Training at knowledge scores.

1B) Overall, New Training appears to be more effective than Standard Training, specifically regarding maneuver efficacy and sensor operation proficiency. Average scores in declarative knowledge were higher in New training, but this did not reach a level of significance.

2A) The homogeneity of variance indicates that the population variance between each group does not differ from one another. In other words, homoscedasticity indicates that the population variances are the same.

2B) This assumption was never mentioned in a one sample test because only one population variance was being evaluated, and homoscedasticity is measured when comparing two or more group sizes.