Analysis of postural sway data

AUTHOR
Steve Simon and Leroy Wheeler

PUBLISHED
October 9, 2024

This program reads data and runs a two-sample t-test. Consult the data dictionary for information about the data itself.

Libraries

```
library(broom)
library(tidyverse)
```

Read data

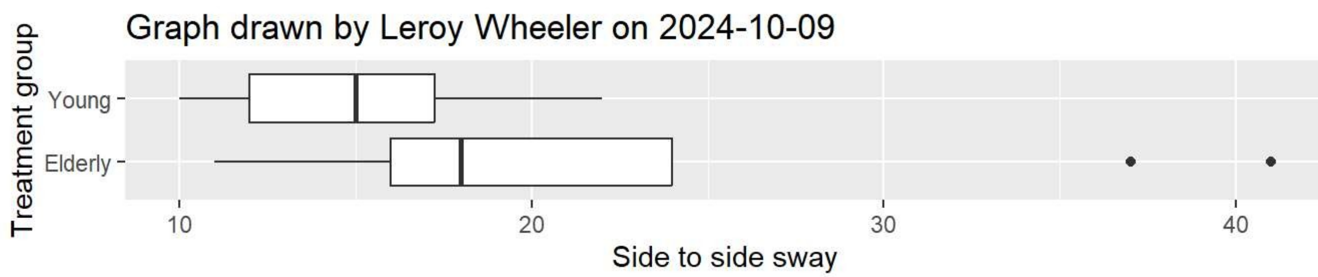
```
sway <- read_tsv(
  file="../data/postural-sway.txt",
  col_types="cnn")
names(sway) <- tolower(names(sway))
glimpse(sway)
```

```
Rows: 17
Columns: 3
$ age      <chr> "Elderly", "Elderly", "Elderly", "Elderly", "Elderly", "Elder…
$ fbsway   <dbl> 19, 30, 20, 19, 29, 25, 21, 24, 50, 25, 21, 17, 15, 14, 14, 2…
$ sidesway <dbl> 14, 41, 18, 11, 16, 24, 18, 21, 37, 17, 10, 16, 22, 12, 14, 1…
```

Boxplot of fbsway by age

```
sway |>
  ggplot(aes(age, sidesway)) +
    geom_boxplot() +
```

```
  ggtitle("Graph drawn by Leroy Wheeler on 2024-10-09") +
  xlab("Treatment group") +
  ylab("Side to side sway") +
  coord_flip()
```



The elderly patients have generally higher side to side sway values. The elderly group also appear to have more variation which include two extreme outliers.

Descriptive statistics for memory by treatment

```
sway |>
  group_by(age) |>
  summarize(
    sidesway_mn=mean(sidesway),
    sidesway_sd=sd(sidesway),
    n=n())
```

```
# A tibble: 2 × 4
  age     sidesway_mn sidesway_sd     n
  <chr>         <dbl>       <dbl> <int>
1 Elderly        22.2       10.3      9
2 Young          15.1        3.91     8
```

The average side-to-side sway is higher in the elderly patients. There is more variation in the elderly group, possibly caused by the two outliers.

Two-sample t-test

```r
m1 <- t.test(
  sidesway ~ age,
  data=sway,
  alternative="two.sided",
  var.equal=TRUE)
m1
```

```

	Two Sample t-test

data:  sidesway by age
t = 1.8349, df = 15, p-value = 0.08643
alternative hypothesis: true difference in means between group Elderly and group Young is not equal to 0
95 percent confidence interval:
 -1.146965 15.341409
sample estimates:
mean in group Elderly   mean in group Young 
             22.22222              15.12500 
```

The 95% confidence interval includes the value of zero for the difference in the sample means of sidesway values. Further, the calculated p value is greater than our alpha which is set at 0.05. Therefore we conclude that there is not enough difference in the sample means to support rejecting the null hypothesis.

Equivalent analysis using linear regression

```r
m2 <- lm(sidesway ~ age, data=sway)
tidy(m2)
```

```
# A tibble: 2 × 5
  term        estimate std.error statistic     p.value
  <chr>          <dbl>     <dbl>     <dbl>       <dbl>
```

```
1 (Intercept)    22.2       2.65      8.38 0.000000487
2 ageYoung       -7.10      3.87     -1.83 0.0864
```

```r
confint(m2)
```

```
                2.5 %    97.5 %
(Intercept)  16.56676 27.877688
ageYoung    -15.34141  1.146965
```

Using a linear model, R program has given a value of zero to the elderly group and a value of one to the young group. The y-intercept is equivalent to the mean sidesway of the elderly group and has a value of 22.2. The young group mean value is determined by adding -7.09 to the elderly group mean and is equal to 15.1. The 95% confidence interval includes the possibility that the difference between the two means is zero so there is not enough difference between the two sample means to support rejecting the null hypothesis.

Sample size justification for postural sway study

This program evaluates various sample size calculations for a proposed replication of the postural-sway study in a different population. It was written by Steve Simon and Leroy Wheeler on 2024-10-09 and is placed in the public domain.

Scenario 1

- Replicate postural sway study
  - Different populations
  - Same outcome measure
- Research hypothesis, $H_0\ \mu_1 - \mu_2 = 0$
- Standard deviations: 9.77, 4.09
- MCID = 4

```
power.t.test(
  n=NULL,
  delta=4,
  sd=9.8,
  sig.level=0.05,
  power=0.9,
  type="two.sample",
  alternative="two.sided")
```

```

     Two-sample t test power calculation 

              n = 127.1097
          delta = 4
             sd = 9.8
      sig.level = 0.05
          power = 0.9
    alternative = two.sided

NOTE: n is number in *each* group
```

With a sample of 128 patients, we would have 90% power for detecting a 4 unit difference in postural sway, using a two-sided test at an alpha level of 0.05.

Scenario 2, MCID = 2

```
power.t.test(
  n=NULL,
  delta=2,
  sd=9.8,
  sig.level=0.05,
  power=0.9,
  type="two.sample",
  alternative="two.sided")
```

```

     Two-sample t test power calculation 

              n = 505.5288
          delta = 2
             sd = 9.8
      sig.level = 0.05
          power = 0.9
    alternative = two.sided

NOTE: n is number in *each* group
```

If we wanted to be able to detect a 2 unit difference, we would need a sample size that is about four times as large, assuming standard deviation, our alpha and power were unchanged.

Scenario 3, MCID=8

```
power.t.test(
  n=NULL,
  delta=8,
```

```
  sd=9.8,
  sig.level=0.05,
  power=0.9,
  type="two.sample",
  alternative="two.sided")
```

```
     Two-sample t test power calculation 

              n = 32.52648
          delta = 8
             sd = 9.8
      sig.level = 0.05
          power = 0.9
    alternative = two.sided

NOTE: n is number in *each* group
```

With a sample of 33 patients, we would have 90% power for detecting an 8 unit difference in postural sway, using a two-sided test at an alpha level of 0.05.

Scenario 4, sd=4.9

```
power.t.test(
  n=NULL,
  delta=4,
  sd=4.9,
  sig.level=0.05,
  power=0.9,
  type="two.sample",
  alternative="two.sided")
```

```
     Two-sample t test power calculation 

              n = 32.52648
```

```
          delta = 4
             sd = 4.9
      sig.level = 0.05
          power = 0.9
    alternative = two.sided

NOTE: n is number in *each* group
```

With a sample of 33 patients, we would have 90% power for detecting a 4 unit difference in postural sway with a standard deviation of 4.9, using a two-sided test at an alpha level of 0.05.

Scenario 5, sd=19.6

```
power.t.test(
  n=NULL,
  delta=4,
  sd=19.6,
  sig.level=0.05,
  power=0.9,
  type="two.sample",
  alternative="two.sided")
```

```

     Two-sample t test power calculation

              n = 505.5288
          delta = 4
             sd = 19.6
      sig.level = 0.05
          power = 0.9
    alternative = two.sided

NOTE: n is number in *each* group
```

If we increased our pooled standard deviation from 9.8 to 19.6, we would need 506 patients to have 90% power for detecting a 4 unit difference in postural sway, using a two-sided test at an alpha level of 0.05.

Scenario 6, alpha=0.01

```
power.t.test(
  n=NULL,
  delta=4,
  sd=9.8,
  sig.level=0.01,
  power=0.9,
  type="two.sample",
  alternative="two.sided")
```

```

     Two-sample t test power calculation 

              n = 180.2936
          delta = 4
             sd = 9.8
      sig.level = 0.01
          power = 0.9
    alternative = two.sided

NOTE: n is number in *each* group
```

With a sample of 181 patients, we would have 90% power for detecting a 4 unit difference in postural sway with a standard deviation of 9.8, using a two-sided test at an alpha level of 0.01.

Scenario 7, alpha=0.10

```
power.t.test(
  n=NULL,
  delta=4,
  sd=9.8,
  sig.level=0.1,
  power=0.9,
```

```
  type="two.sample",
  alternative="two.sided")
```

```

     Two-sample t test power calculation 

              n = 103.4925
          delta = 4
             sd = 9.8
      sig.level = 0.1
          power = 0.9
    alternative = two.sided

NOTE: n is number in *each* group
```

With a sample of 104 patients, we would have 90% power for detecting a 4 unit difference in postural sway with a standard deviation of 9.8, using a two-sided test at an alpha level of 0.1.

Scenario 8, power=0.8

```
power.t.test(
  n=NULL,
  delta=4,
  sd=9.8,
  sig.level=0.05,
  power=0.8,
  type="two.sample",
  alternative="two.sided")
```

```

     Two-sample t test power calculation 

              n = 95.19575
          delta = 4
             sd = 9.8
      sig.level = 0.05
```

```
      power = 0.8
alternative = two.sided

NOTE: n is number in *each* group
```

With a sample of 96 patients, we would have 80% power for detecting a 4 unit difference in postural sway with a standard deviation of 9.8, using a two-sided test at an alpha level of 0.05.

Scenario 9, power=0.95

```
power.t.test(
  n=NULL,
  delta=4,
  sd=9.8,
  sig.level=0.05,
  power=0.95,
  type="two.sample",
  alternative="two.sided")
```

```

     Two-sample t test power calculation 

              n = 156.9685
          delta = 4
             sd = 9.8
      sig.level = 0.05
          power = 0.95
    alternative = two.sided

NOTE: n is number in *each* group
```

With a sample of 157 patients, we would have 95% power for detecting a 4 unit difference in postural sway with a standard deviation of 9.8, using a two-sided test at an alpha level of 0.05.