Analysis of postural sway data

AUTHOR
Michael Dang

PUBLISHED
October 15, 2024

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

This program was written by Steve Simon on 2024-10-07 with a minor fix to the documentation on 2024-10-13. It is placed in the public domain.

Libraries

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

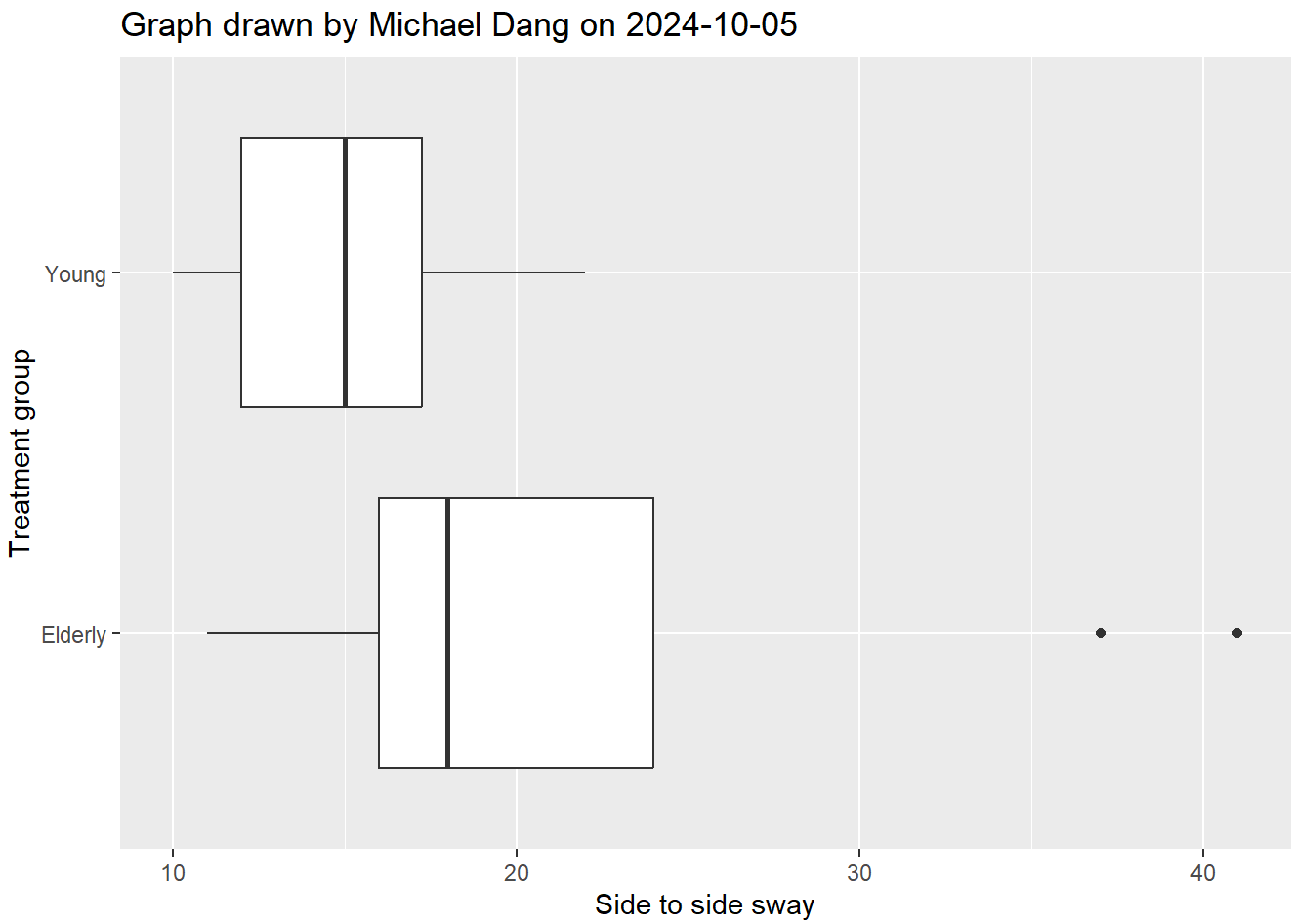
Read data

```
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 side-to-side sway by age

```
sway |>
  ggplot(aes(age, sidesway)) +
    geom_boxplot() +
    ggtitle("Graph drawn by Michael Dang on 2024-10-05") +
    xlab("Treatment group") +
    ylab("Side to side sway") +
    coord_flip()
```



Elderly patients tend to have generally higher sway values compared to the younger group, indicating more instability in their sway. The variation (interquartile range) between the two groups appears similar, suggesting comparable variability in their sway behavior.

Descriptive statistics for side-to-side sway by age

```
sway |>
  group_by(age) |>
  summarize(
    sts_mn=mean(sidesway),
    sts_sd=sd(sidesway),
    n=n())
```

```
# A tibble: 2 × 4
  age     sts_mn sts_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 extreme outlier.

Two-sample t-test

```
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 two-sample t-test result indicates that the t-statistic is 1.8349 with 15 degrees of freedom, and the p-value is 0.08643.
- Since the p-value is greater than the conventional significance level (0.05), we fail to reject the null hypothesis. This suggests that there is no statistically significant difference between the mean side-to-side sway values for elderly and young patients at the 5% significance level.
- However, the confidence interval (-1.15 to 15.34) indicates the range within which the true difference in means likely falls, implying that while a difference exists, the evidence isn't strong enough to conclude it definitively under these test conditions.

Equivalent analysis using linear regression

```
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
```

```
confint(m2)
```

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

The p-value in the linear regression analysis is the same as in two-sample t-test and same as the confidence interval.