Statistics 3.2.2 Exercise 2

Waist-to-Hip Ratios Men/Women

Lucas Liona

Table of contents

Problem Description

Waist-to-hip ratios of men and women sampled from the workforce-study database are provided. The data includes different age groups (45 and >45) for both males and females.

Data Loading and Preparation

```
# Load required libraries
library(ggplot2)
library(dplyr)
```

```
Attaching package: 'dplyr'
```

```
The following objects are masked from 'package:stats':

    filter, lag

The following objects are masked from 'package:base':

    intersect, setdiff, setequal, union
```

```r
# Read the data
waist_data <- read.table("waists.txt", header = TRUE, sep = "\t")

# Check the structure of our data
str(waist_data)
```

```
'data.frame':   80 obs. of  3 variables:
 $ sex   : chr  "male" "male" "male" "male" ...
 $ ratios: num  0.84 0.99 0.78 0.91 1.06 0.98 0.86 0.89 0.87 0.92 ...
 $ agegp : chr  "<=45" "<=45" "<=45" "<=45" ...
```

```r
# View summary statistics
summary(waist_data)
```

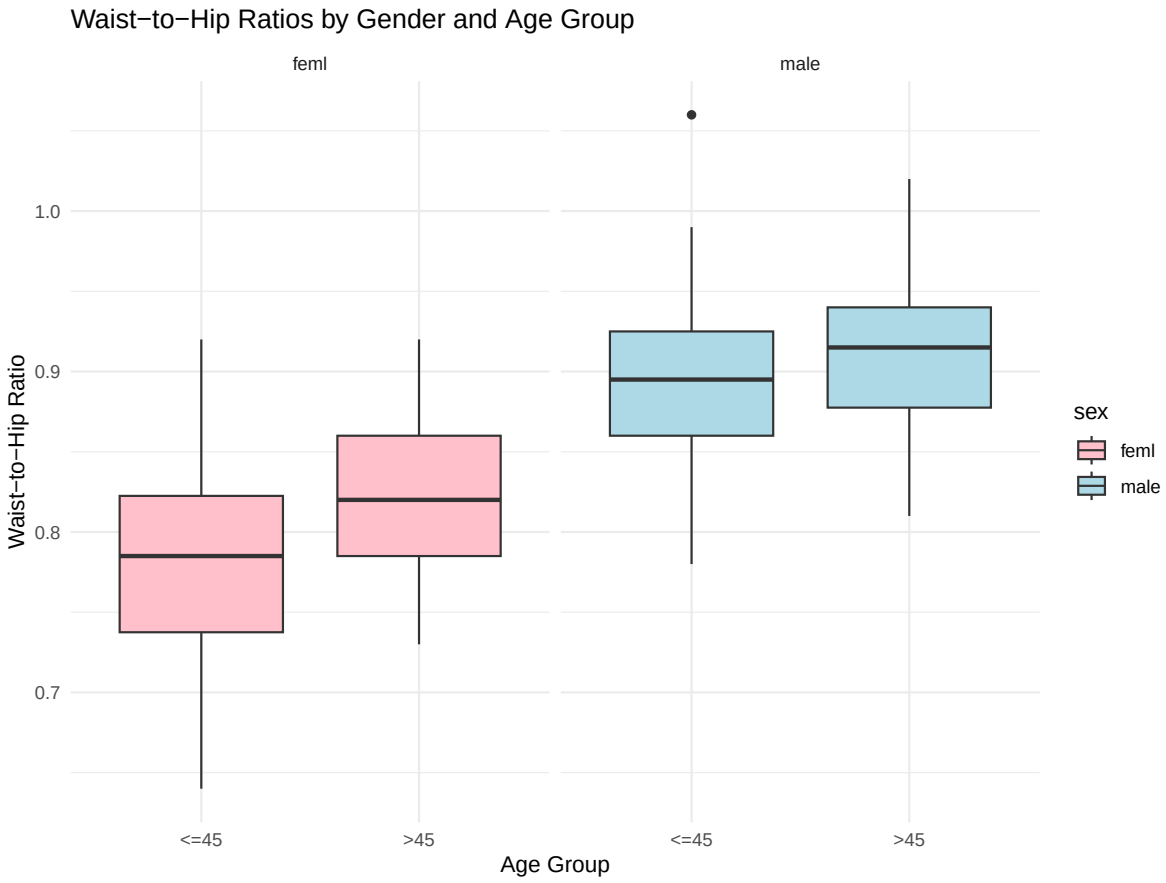
```
     sex                ratios          agegp
 Length:80          Min.   :0.6400   Length:80
 Class :character   1st Qu.:0.8000   Class :character
 Mode  :character   Median :0.8550   Mode  :character
                    Mean   :0.8535
                    3rd Qu.:0.9100
                    Max.   :1.0600
```

(a) Box Plots

Let's create box plots to compare the four groups by gender and age group:

```r
# Create box plots for each group
ggplot(waist_data, aes(x = agegp, y = ratios, fill = sex)) +
  geom_boxplot() +
  facet_wrap(~ sex) +
  labs(title = "Waist-to-Hip Ratios by Gender and Age Group",
       x = "Age Group",
```

```
       y = "Waist-to-Hip Ratio") +
  theme_minimal() +
  scale_fill_manual(values = c("male" = "lightblue", "feml" = "pink"))
```

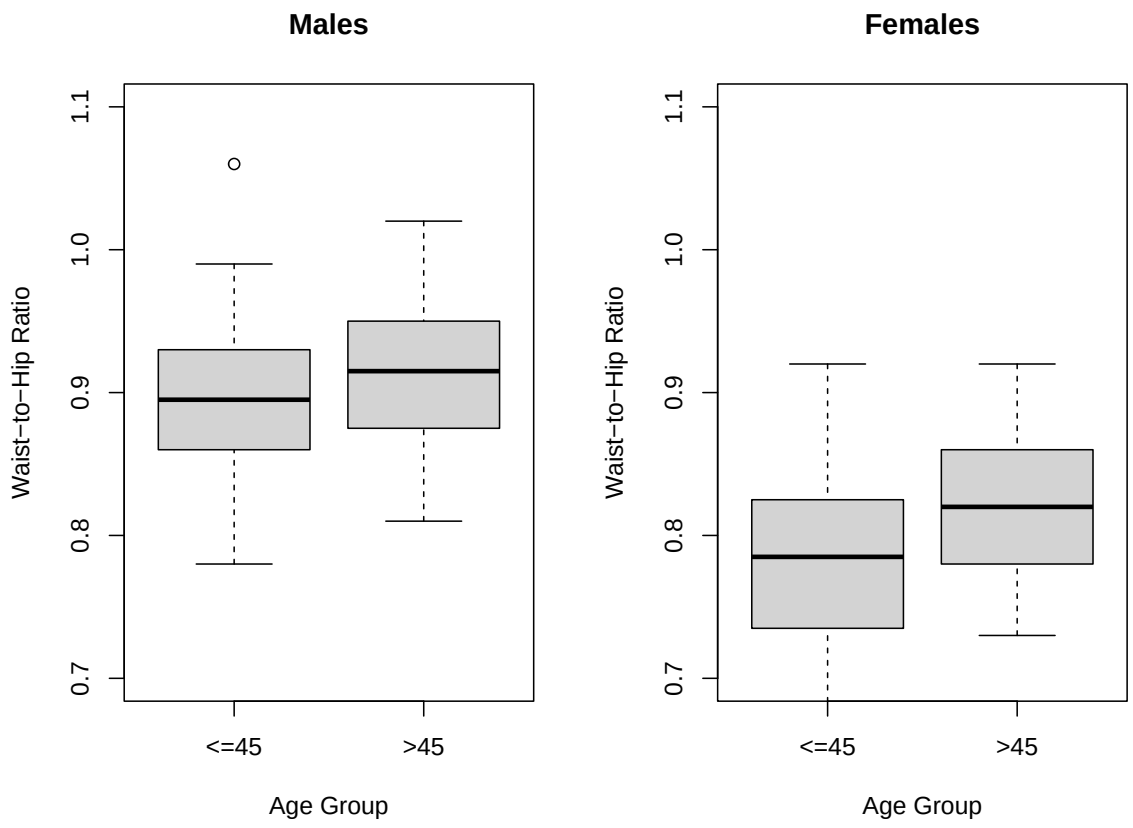



```
# Alternative: Side-by-side box plots
par(mfrow = c(1, 2))

# Plot for males
male_data <- waist_data[waist_data$sex == "male", ]
boxplot(ratios ~ agegp, data = male_data,
        main = "Males",
        xlab = "Age Group",
        ylab = "Waist-to-Hip Ratio",
        ylim = c(0.7, 1.1))

# Plot for females
```

```r
female_data <- waist_data[waist_data$sex == "feml", ]
boxplot(ratios ~ agegp, data = female_data,
        main = "Females",
        xlab = "Age Group",
        ylab = "Waist-to-Hip Ratio",
        ylim = c(0.7, 1.1))
```



(b) Analysis of Results

Examining the box plots reveals several important patterns:

Gender Differences:

- Women consistently have lower waist-to-hip ratios than men across both age groups.

- Male ratios typically range from about 0.85 to 1.0, while female ratios range from about 0.7 to 0.9.

Age Effects:

- For both genders, the older age group (>45) shows slightly higher waist-to-hip ratios on average.
- The age effect appears more pronounced in women, with a clearer separation between the age groups.

Variability:

- All groups show relatively similar spreads (interquartile ranges).
- The data appears approximately symmetric for all groups with a few outliers in each.

The clear separation between male and female distributions suggests that gender is a strong predictor of waist-to-hip ratio. Age also appears to be a factor, though less pronounced than gender.

Statistical Summary

Let's calculate summary statistics for each group:

```
# Group summary statistics
waist_data %>%
  group_by(sex, agegp) %>%
  summarise(
    n = n(),
    mean = mean(ratios),
    median = median(ratios),
    sd = sd(ratios),
    min = min(ratios),
    max = max(ratios)
  ) %>%
  knitr::kable(digits = 3, caption = "Summary Statistics by Group")
```

```
`summarise()` has grouped output by 'sex'. You can override using the `.groups`
argument.
```

Table 1: Summary Statistics by Group

sex	agegp	n	mean	median	sd	min	max
feml	<=45	20	0.779	0.785	0.067	0.64	0.92
feml	>45	20	0.821	0.820	0.051	0.73	0.92
male	<=45	20	0.900	0.895	0.067	0.78	1.06
male	>45	20	0.914	0.915	0.058	0.81	1.02

Conclusion

The analysis confirms that both gender and age influence waist-to-hip ratios:

1. **Gender is the primary factor**: Men consistently have higher ratios than women, likely reflecting biological differences in fat distribution patterns.
2. **Age has a secondary effect**: Both genders show increased ratios with age, possibly due to changes in body composition over time.

These differences are important for health assessments, as waist-to-hip ratio is often used as an indicator of health risks associated with body fat distribution.