Comparing Means for Two Independent Samples: An Example

Brady T. West

Example: Comparing Means in Two Groups

Research Question:
Considering African-American adults living in the U.S. in 2015-2016, did **males and females have significantly different mean** systolic blood pressure?

Inference Approaches:

- Form a confidence interval for the difference in the two means
- Perform a two-sample t-test for the difference in the two means
- Be sure to check assumptions!

Approach 1: Form a Confidence Interval

Males: Mean = 131.01, standard deviation = 20.59, n = 536
Females: Mean = 125.79, standard deviation = 19.06, n = 599

- **Best Point Estimate:** Difference in sample means is

  131.01 – 125.79 = 5.22 mmHg

- **Interpretation:** In 2015-2016, we estimate the mean systolic blood pressure for all male black adults was 5.22 mmHg ***higher*** than that for all female black adults.

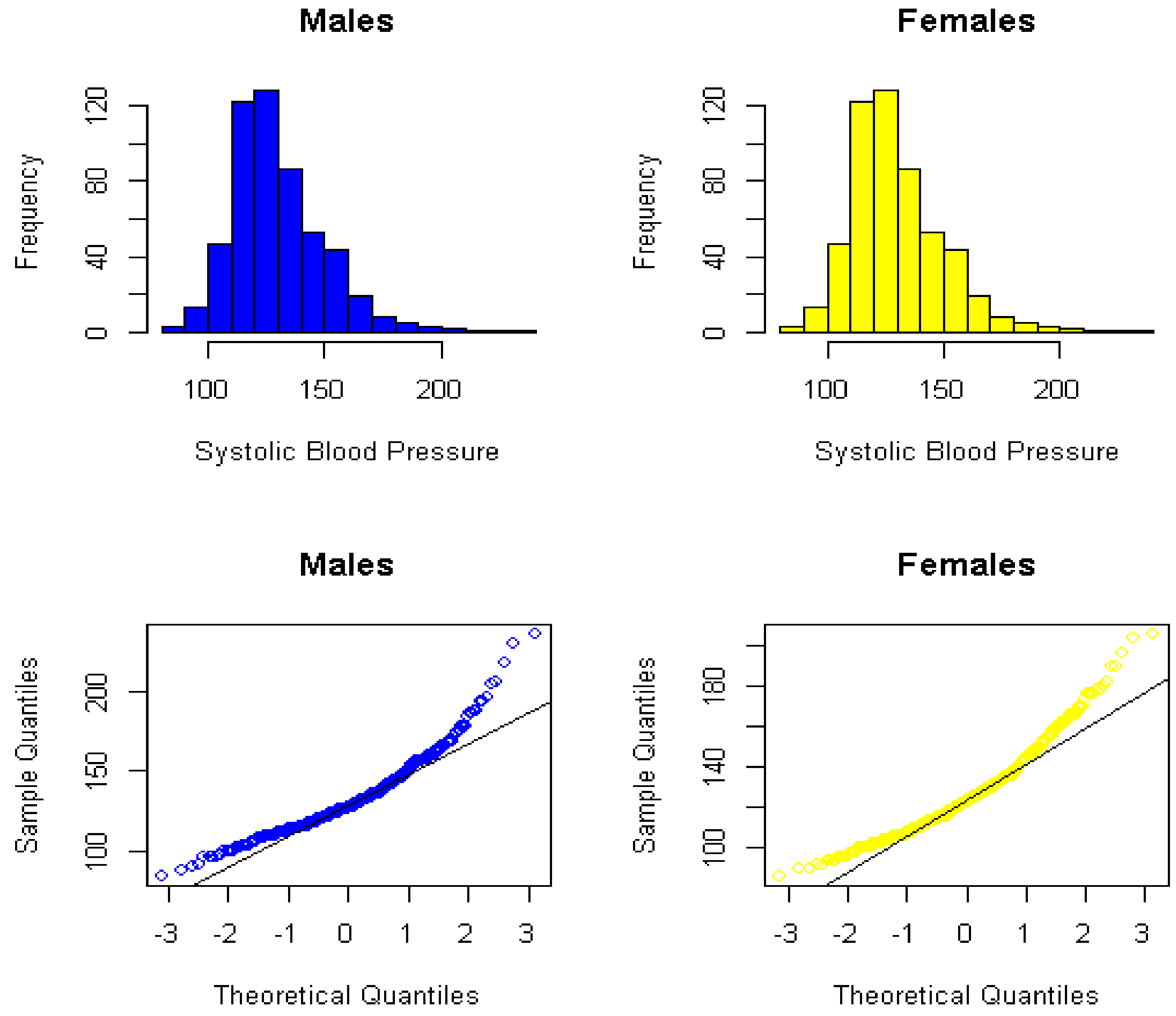
Approach 1: Form a Confidence Interval

Males: Mean = 131.01, standard deviation = 20.59, n = 536
Females: Mean = 125.79, standard deviation = 19.06, n = 599

Note: sample standard deviations are similar.

Let's **check some assumptions** and
decide which confidence interval approach is reasonable
(pooled or unpooled?)

Checking Assumptions

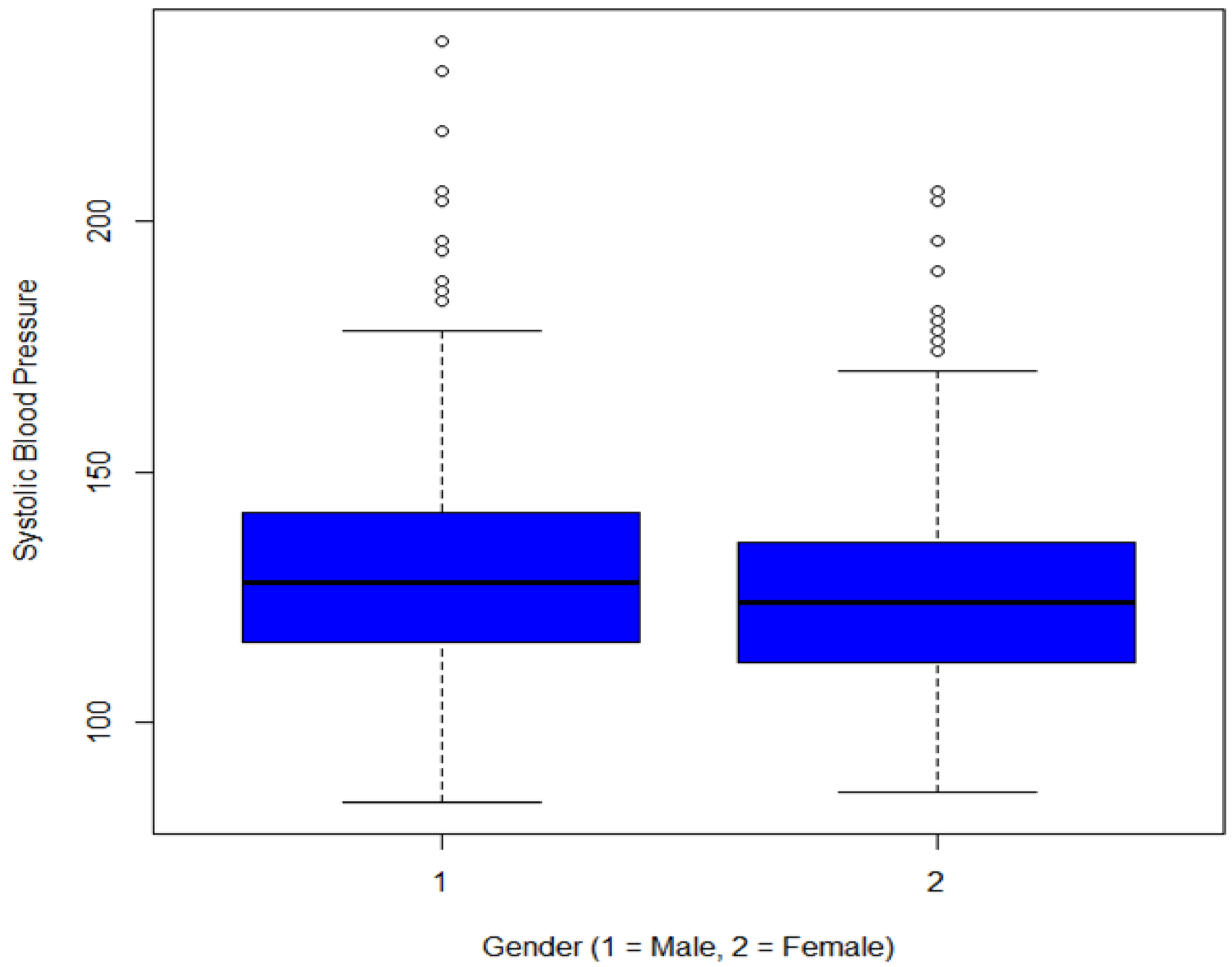


The distributions have a moderate right skew for both males and females...

The normality assumption is important for the two-sample t-test!

(large sample sizes, so can also rely on CLT)

Checking Assumptions



Some evidence of slightly higher variance for males

→ consider robustness of results to different assumptions about variances in two groups!

Approach 1: Form a Confidence Interval

- **Assuming equal variance (pooling):**

  95% CI for difference in means = (2.91 mmHg, 7.53 mmHg)

  Interval doesn't include 0 → **Significant difference!**

- **Assuming unequal variance (no pooling):**

  95% CI for difference in means = (2.90 mmHg, 7.54 mmHg)

Same conclusion!

Result robust to possible violation of assumption.

Approach 2: Two-Sample t-test

- **Null :** Males and Females have equal population means
- **Alternative :** Males and Females have **different** means

Alternative allows male mean to be *either* greater or less than female mean
→ **two-tailed test** need more evidence against null hypothesis to reject it!

Significance Level = 5%

Approach 2: Two-Sample t-test

Assumptions:

- **Normal** distribution of blood pressure in each population
  - May not be reasonable, based on previous histograms and QQ plots for each sample
- **Same standard deviation** in each population
  - Somewhat reasonable and techniques robust,
    but we can examine both pooled and unpooled test results

Approach 2: Two-Sample t-test

- **Assuming equal variance (pooling):**

  t = 4.436, df = 1133, p-value < 0.001

We **reject the null hypothesis**; means are different!

- **Assuming unequal variance (no pooling):**

  t = 4.417, df = 1094.3, p-value < 0.001

Same conclusion!

Result robust to possible violation of assumption.

What if Normality Doesn't Hold?

- Not convinced the variable of interest follows a normal distribution in each population?
  → **non-parametric test** that does not assume normality
- Non-parametric analog of two-sample t-test
  = **Mann-Whitney test**
  ~ compares locations of distributions using medians

What if Normality Doesn't Hold?

- Mann-Whitney Test Result: p-value < 0.001
  - We reject null that both distributions have identical "locations"
- Conclusion is robust to potential violations of normality!

Consistent evidence of robust difference
in central tendencies of the two distributions,
regardless of assumptions made
and approach to inference used

What's Next?

How to compare two means based on ***paired data***

Examples:

1. blood pressure measurements
   from right and left arms of same subjects
2. measures of a continuous outcome
   before and after an intervention