

Statistical Simulations Examples

Abhi Pendela

Rahif Mansoor

Rohan Girish

2024-12-08

Contents

Theoretical Sampling Distribution	1
Confidence Level Coverage	2
Complete Analysis Example	5
Additional Examples	8

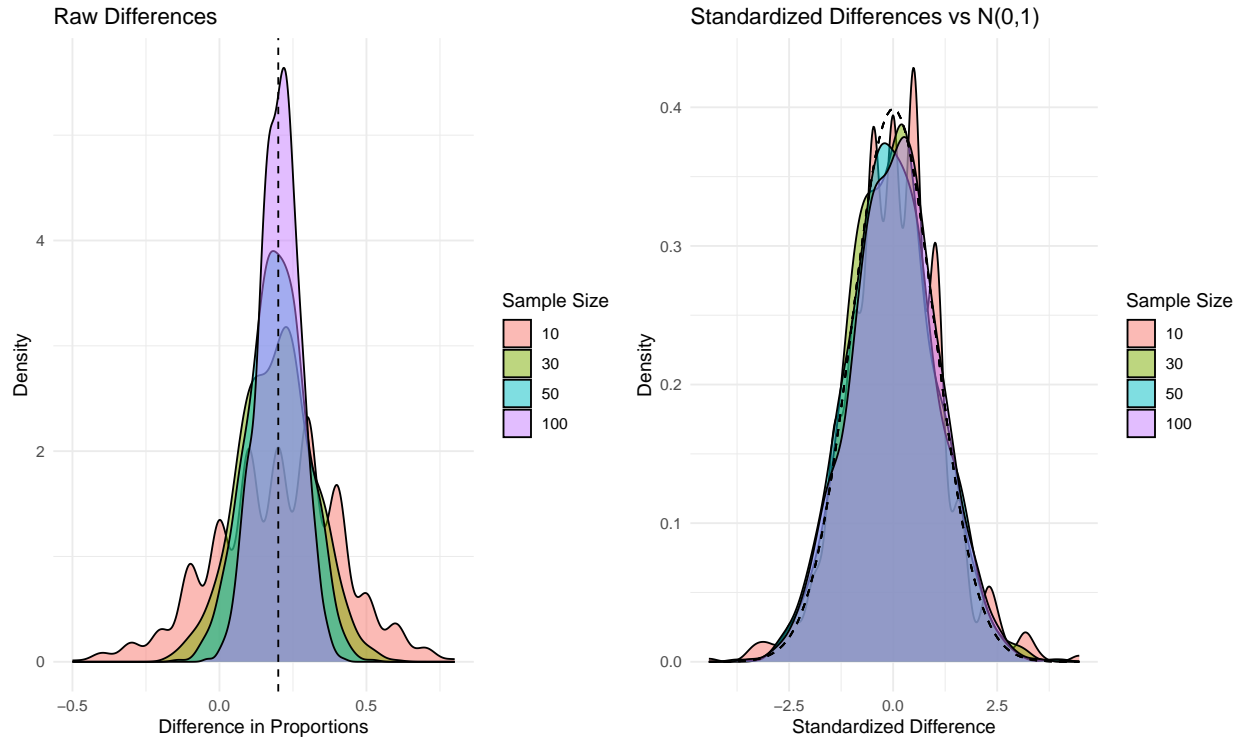
```
# Load required packages
library(ggplot2)
library(gridExtra)

# Source our functions
source("R/core_functions.R")
source("R/visualization_functions.R")
```

Theoretical Sampling Distribution

We'll demonstrate how the sampling distribution of the difference in proportions approaches normality as sample size increases.

```
# Compare different sample sizes
plot_sampling_dist(
  n_values = c(10, 30, 50, 100), # varying sample sizes
  p1 = 0.7,                       # probability for group 1
  p2 = 0.5,                       # probability for group 2
  reps = 1000                     # number of simulations
)
```

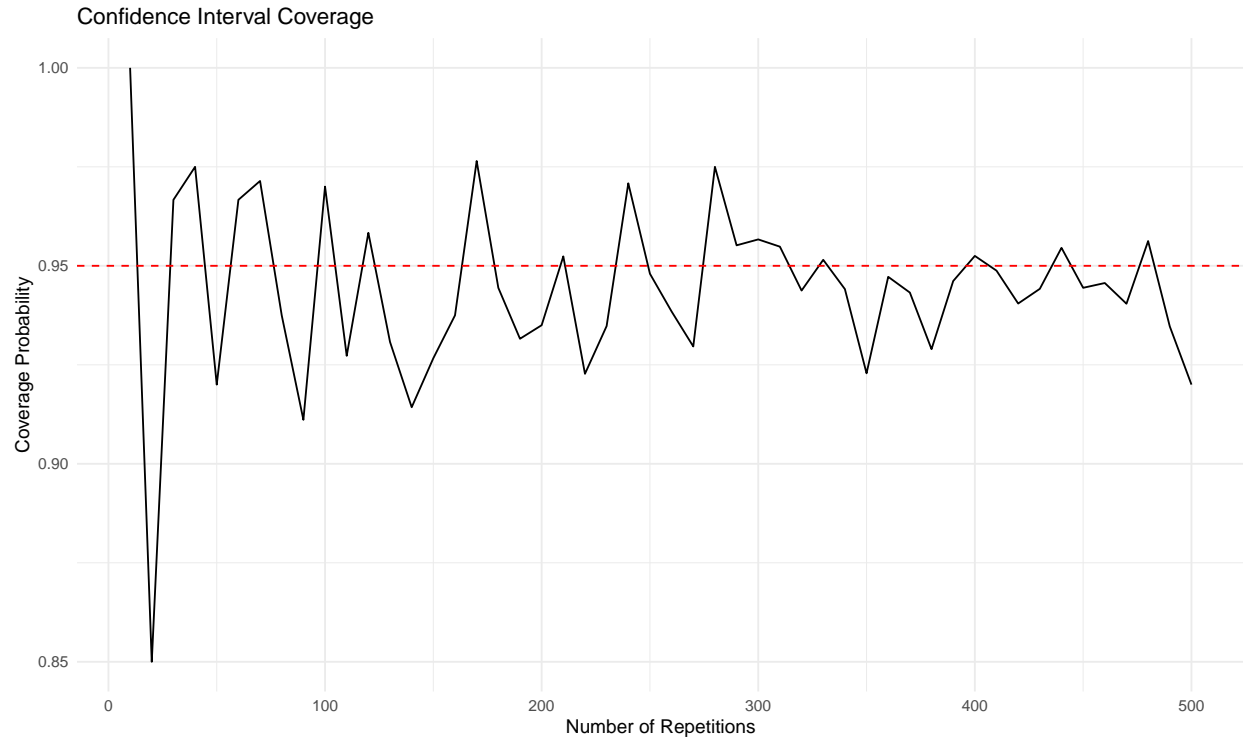


The left plot shows differences in proportions centered around the true difference of 0.2 (dashed line), while the right plot displays standardized differences compared to $N(0,1)$. As sample size increases from 10 to 100, both distributions become notably narrower and more peaked, with the smaller sample sizes showing wider disparity. The standardized differences align with the standard normal distribution at larger sample sizes, clearly showing the Central Limit Theorem

Confidence Level Coverage

Let's examine how the empirical coverage of confidence intervals approaches the nominal level as we increase the number of simulations.

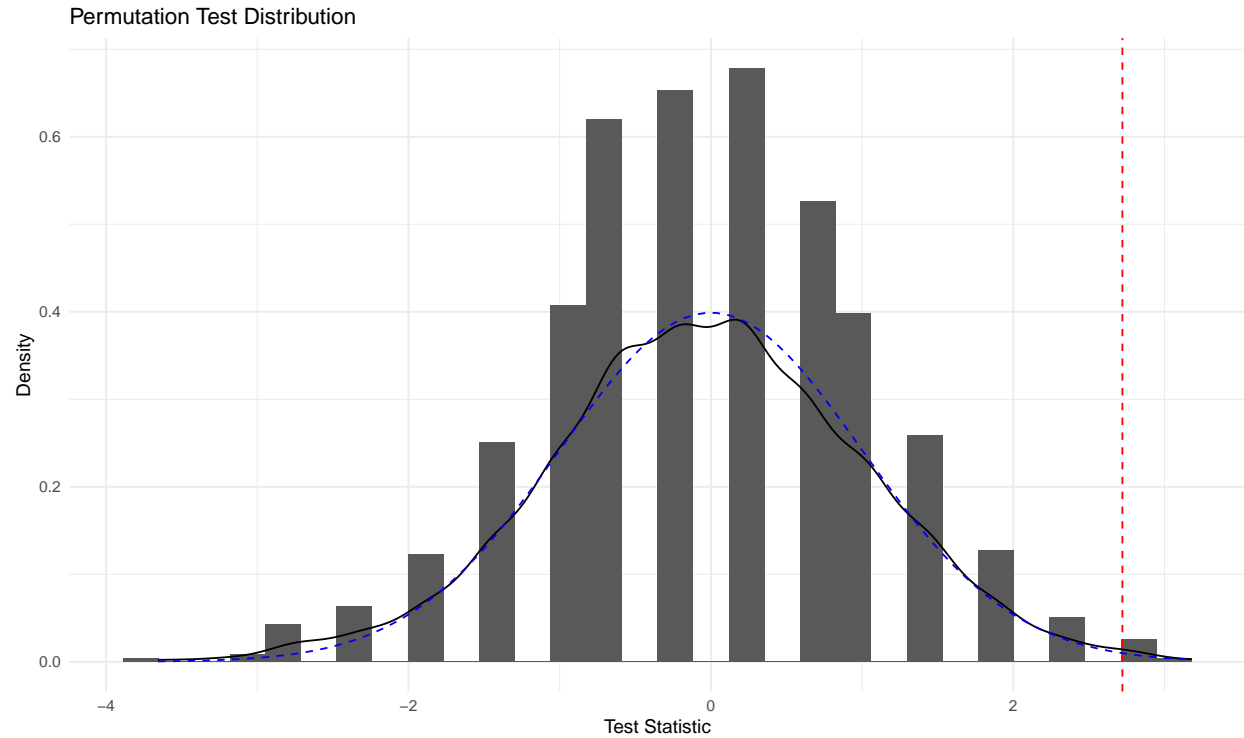
```
# Look at coverage probability for 95% confidence intervals
plot_confidence_coverage(
  p1 = 0.7,          # probability for group 1
  p2 = 0.5,          # probability for group 2
  n1 = 50,           # sample size group 1
  n2 = 50,           # sample size group 2
  alpha = 0.05,      # 95% confidence level
  max_reps = 500     # maximum number of repetitions
)
```



The plot shows the coverage probability of confidence intervals over increasing numbers of repetitions, showing initial high variability with coverage ranging from 0.9 to 1.0 for smaller numbers of repetitions. As the number of repetitions increases beyond 200, the coverage probability shows less extreme fluctuations and tends to stabilize closer to the 95% level (indicated by the dashed line), though some natural random variation still happens ## 3. Permutation Test

We'll perform a permutation test to assess whether the difference in proportions is statistically significant.

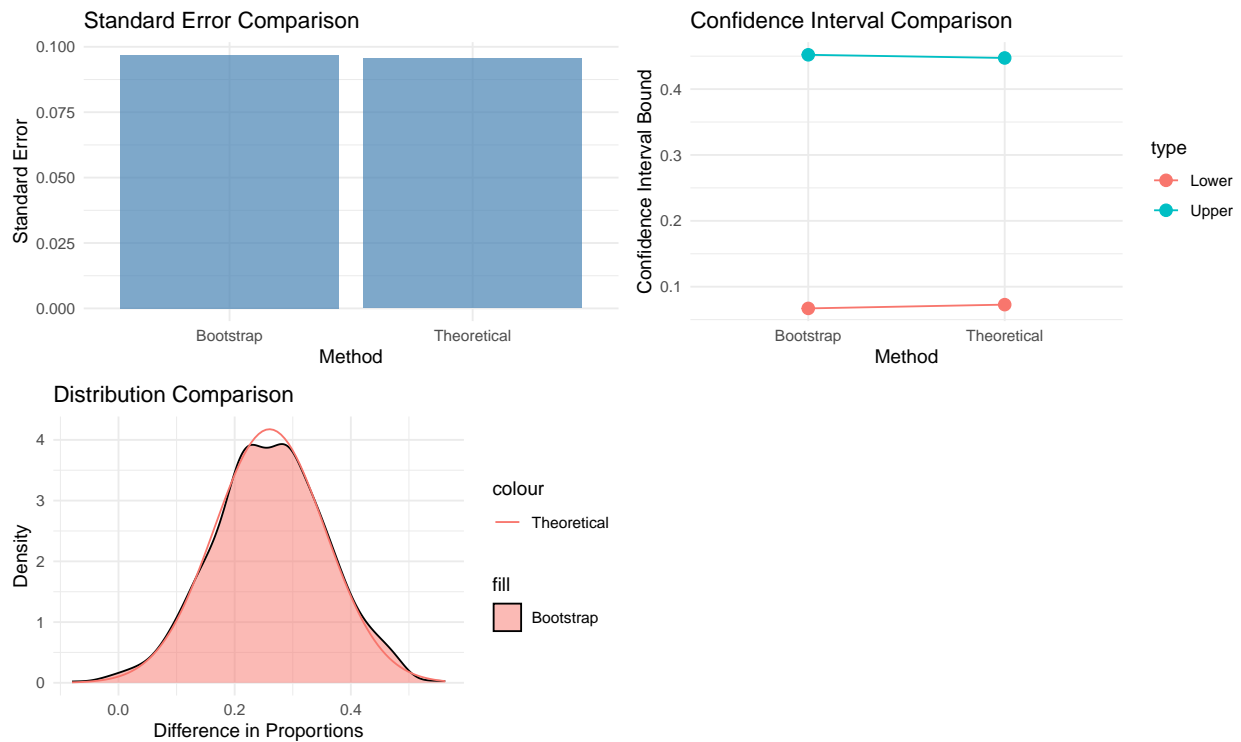
```
# Generate data with a known difference  
set.seed(123) # for reproducibility  
data <- sim_binary_data(0.7, 0.5, 50, 50)  
  
# Perform and visualize permutation test  
plot_permutation_test(data, reps = 1000)
```



The histogram and overlaid normal curve (blue dashed line) show the null distribution of test statistics obtained through permutation, which closely approximates a standard normal distribution centered at zero. The observed test statistic (red dashed line) falls in the extreme right tail of the distribution, providing strong evidence against the null hypothesis of equal proportions between the two groups. ## 4. Bootstrap Analysis

Compare bootstrap and theoretical methods for estimating uncertainty.

```
# Compare bootstrap and theoretical approaches
plot_bootstrap_comparison(data, reps = 1000, conf_level = 0.95)
```



The three plots compare bootstrap and theoretical methods, with the top panels showing nearly identical standard errors and confidence interval bounds between the two approaches. The bottom distribution plot demonstrates excellent agreement between the bootstrap sampling distribution (pink fill) and the theoretical normal distribution (black line), showing both methods provide estimates of the difference in proportions.

Complete Analysis Example

```
# Set parameters
p1 <- 0.7 # probability for group 1
p2 <- 0.5 # probability for group 2
n <- 50 # sample size per group
reps <- 1000 # number of repetitions

# Generate data
set.seed(456) # for reproducibility
data <- sim_binary_data(p1, p2, n, n)

# Calculate observed difference and SE
results <- calc_prop_diff(data)
cat("Observed difference:", round(results$diff, 3), "\n")
```

```
## Observed difference: 0.08
```

```
cat("Standard error:", round(results$se, 3), "\n")
```

```
## Standard error: 0.098
```

```
# Perform permutation test
perm_results <- permutation_test(data, reps)
p_value <- mean(abs(perm_results$null_dist) >= abs(perm_results$obs_stat))
cat("Permutation test p-value:", round(p_value, 3), "\n")
```

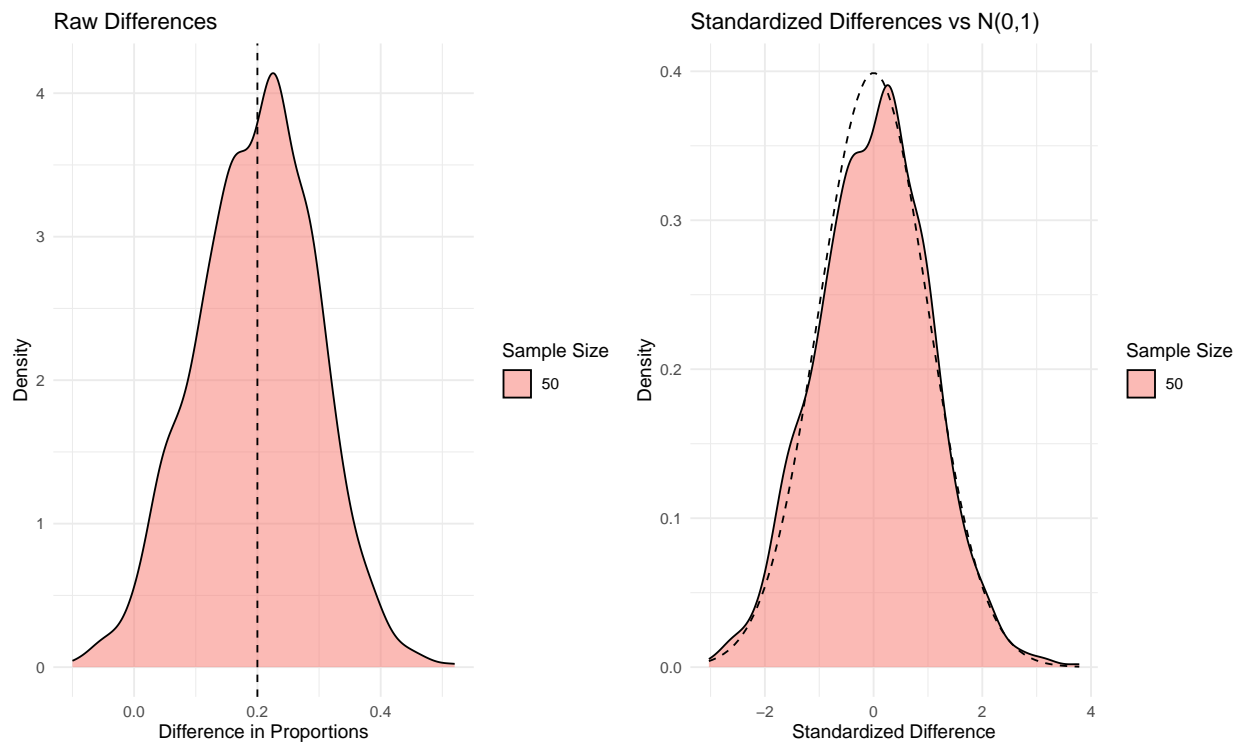
```
## Permutation test p-value: 0.559
```

```
# Calculate bootstrap confidence interval
boot_samples <- bootstrap_samples(data, reps)
boot_ci <- quantile(boot_samples, c(0.025, 0.975))
cat("Bootstrap 95% CI: (", round(boot_ci[1], 3), ",", round(boot_ci[2], 3), ")\n")
```

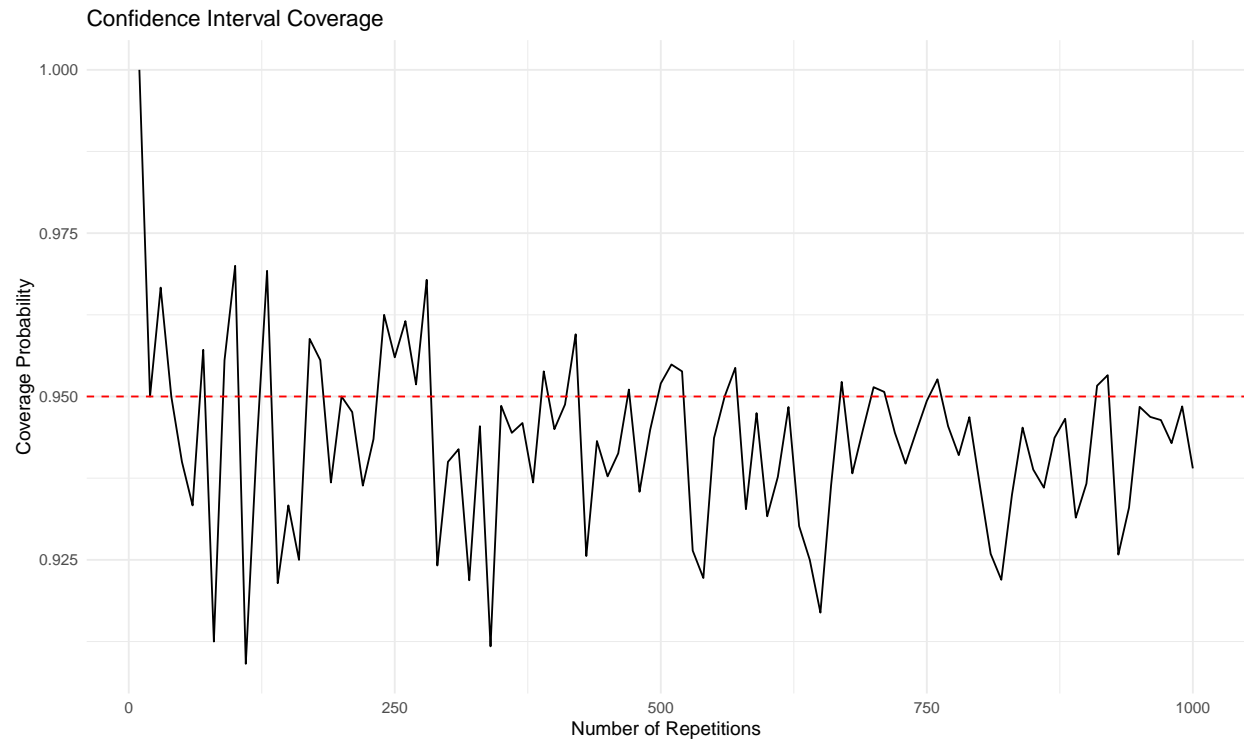
```
## Bootstrap 95% CI: ( -0.117 , 0.265 )
```

```
# Visualize all aspects
par(mfrow = c(2,2))

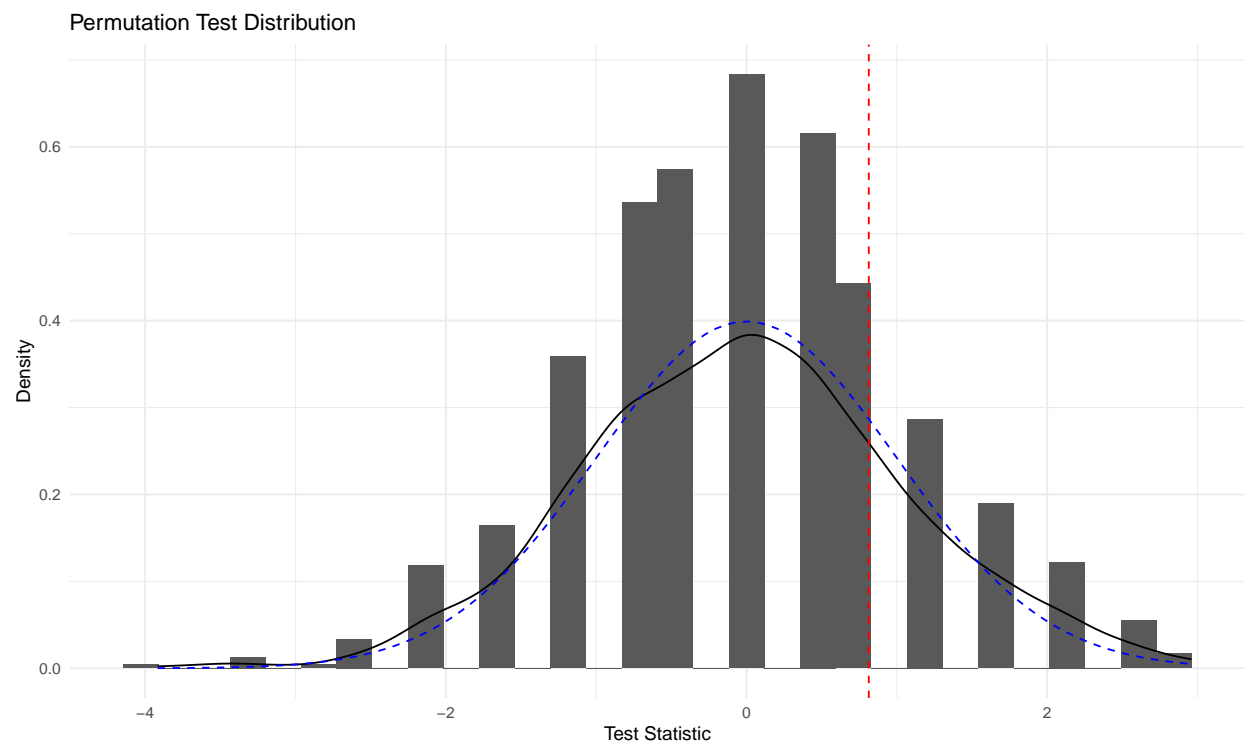
# Sampling distribution
plot_sampling_dist(c(n), p1, p2, reps)
```



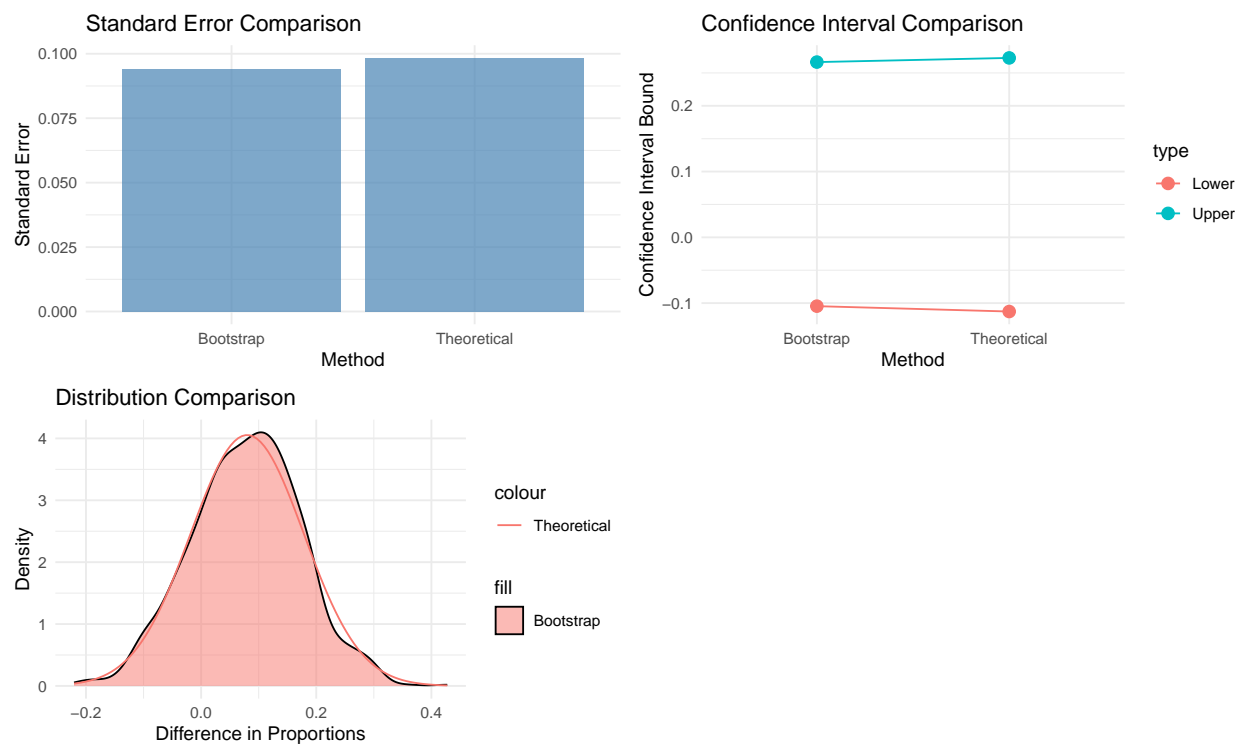
```
# Confidence coverage
plot_confidence_coverage(p1, p2, n, n, max_reps = reps)
```



```
# Permutation test  
plot_permutation_test(data, reps)
```



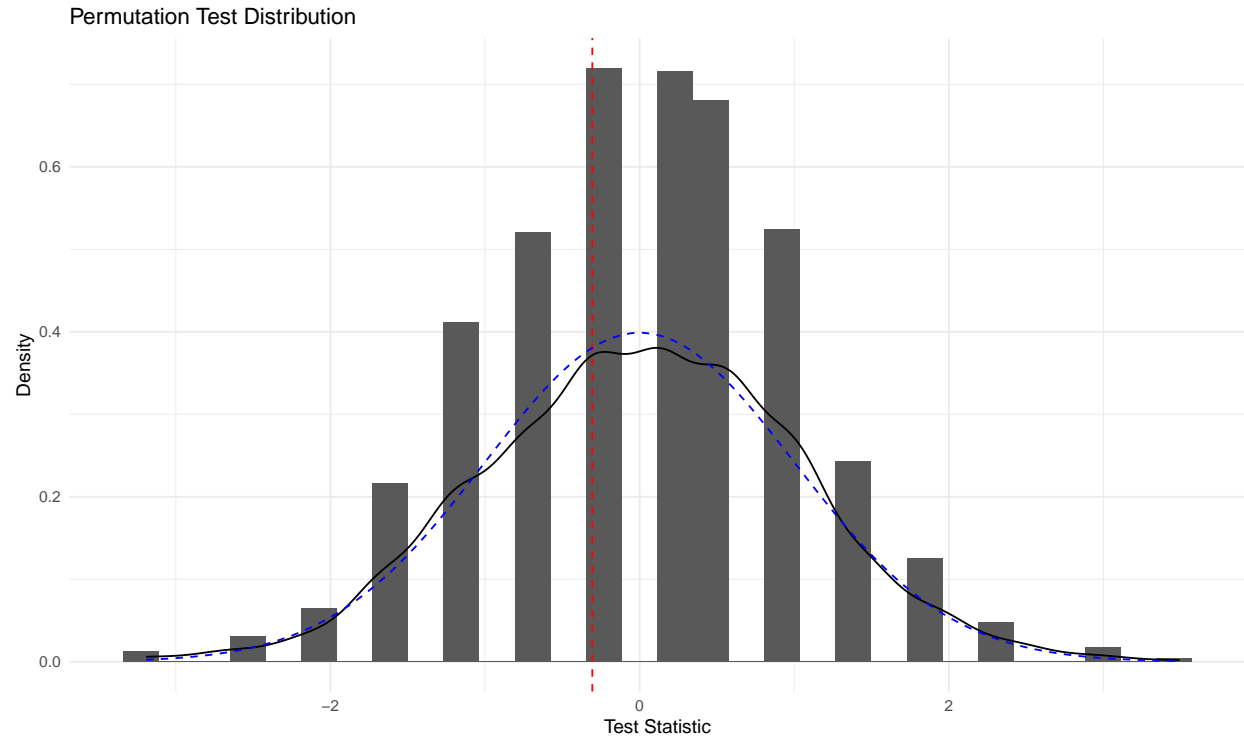
```
# Bootstrap comparison
plot_bootstrap_comparison(data, reps)
```



Additional Examples

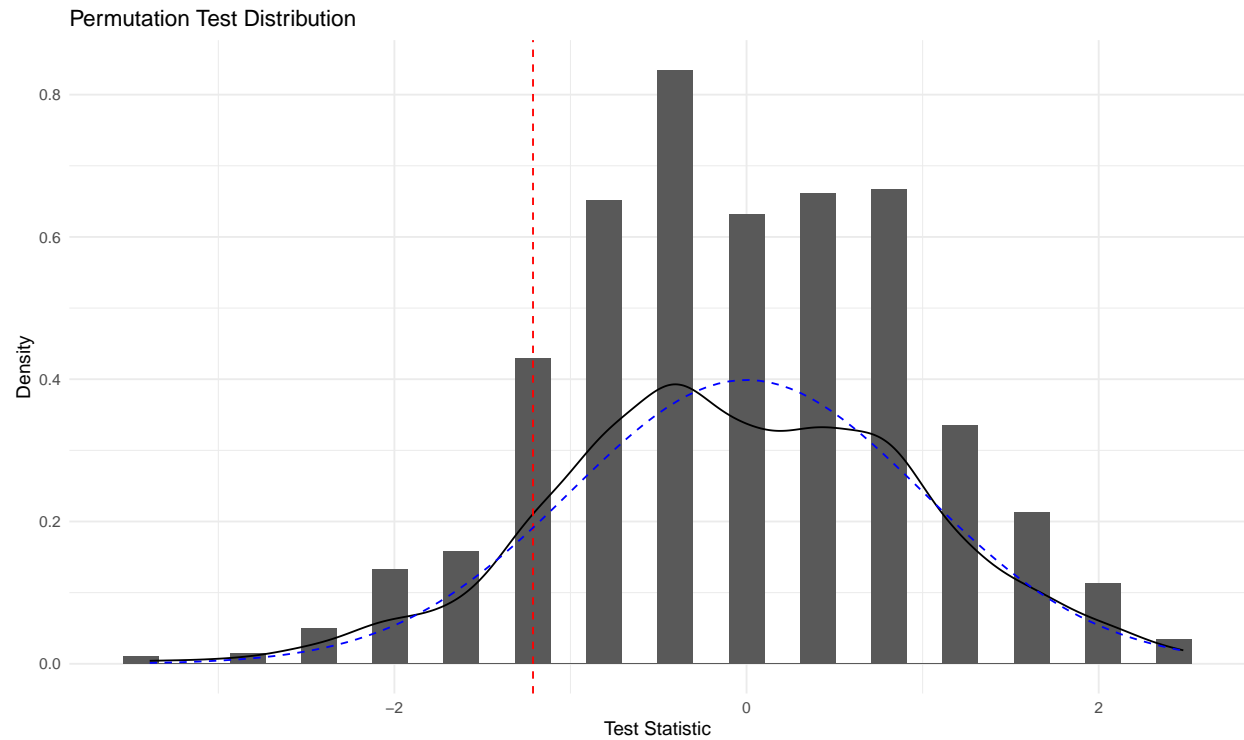
Testing Different Sample Size Ratios

```
# Compare cases with unequal sample sizes
data_unequal <- sim_binary_data(0.7, 0.5, 30, 70)
plot_permutation_test(data_unequal, reps = 1000)
```

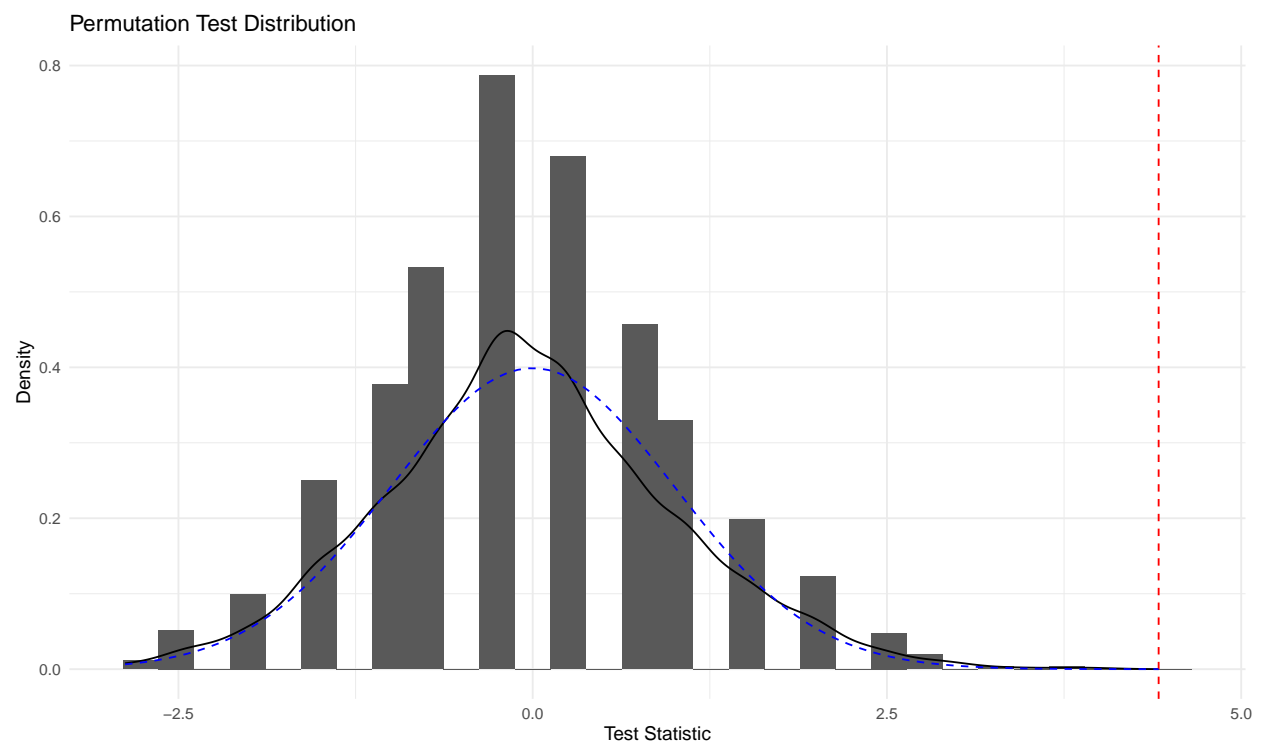



The permutation distribution (gray histogram) closely follows the theoretical normal curve (blue dashed line), and the observed test statistic (vertical dashed line) lies near zero, showing that despite the theoretical difference in proportions, this particular sample with unequal group sizes doesn't provide strong evidence against the null hypothesis of equal proportions. ### Testing Different Effect Sizes

```
# Small effect  
data_small <- sim_binary_data(0.52, 0.50, 50, 50)  
plot_permutation_test(data_small, reps = 1000)
```



```
# Large effect
data_large <- sim_binary_data(0.8, 0.4, 50, 50)
plot_permutation_test(data_large, reps = 1000)
```



In the first plot (small effect: $p_1=0.52$ vs $p_2=0.50$), the observed test statistic (dashed vertical line) falls

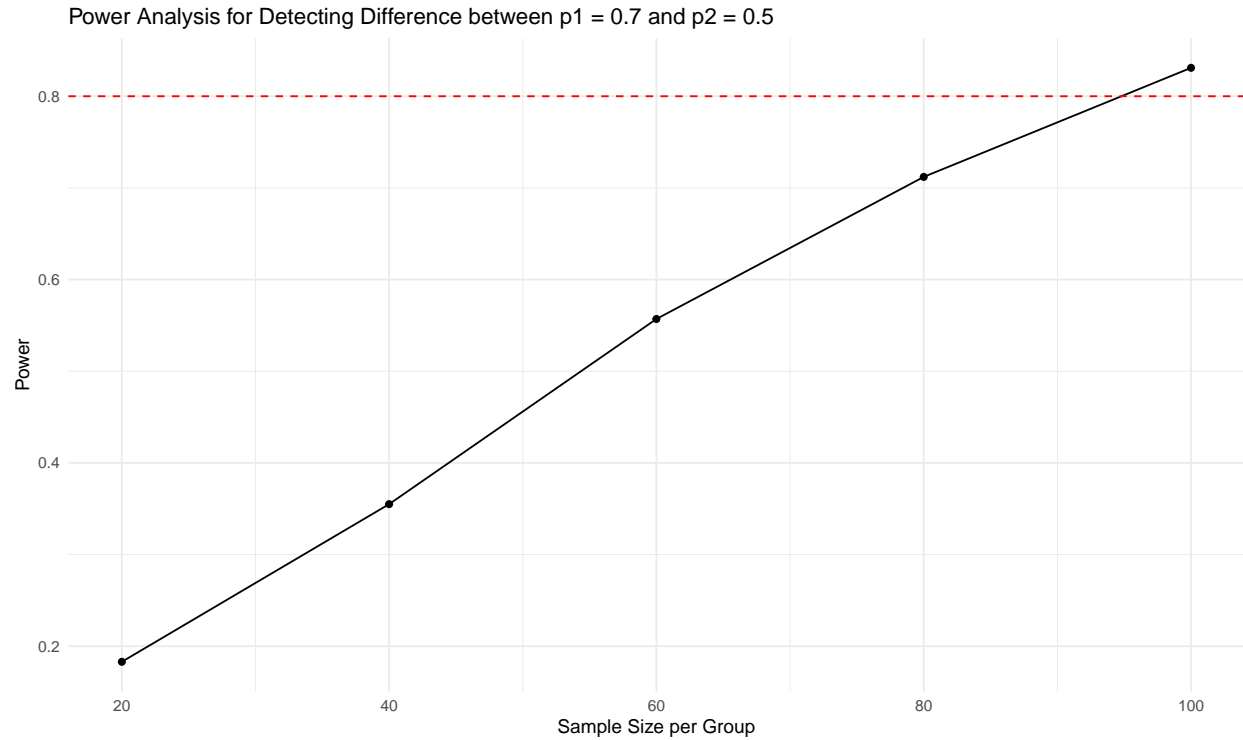
close to the center of the null distribution, showing no significant difference between groups. The second plot (large effect: $p_1=0.8$ vs $p_2=0.4$) shows the observed test statistic far in the right tail of the null distribution, providing strong evidence of a significant difference between the groups. Both plots show that the permutation distribution closely shows the normal distribution (blue dashed line).

Power Analysis Example

```
# Function to calculate power
calc_power <- function(p1, p2, n, reps = 1000, alpha = 0.05) {
  results <- replicate(reps, {
    data <- sim_binary_data(p1, p2, n, n)
    perm_results <- permutation_test(data, 100)
    p_value <- mean(abs(perm_results$null_dist) >= abs(perm_results$obs_stat))
    p_value <= alpha
  })
  mean(results)
}

# Calculate power for different sample sizes
n_seq <- seq(20, 100, by = 20)
power_results <- sapply(n_seq, function(n) calc_power(0.7, 0.5, n))

# Plot power curve
ggplot(data.frame(n = n_seq, power = power_results), aes(x = n, y = power)) +
  geom_line() +
  geom_point() +
  geom_hline(yintercept = 0.8, linetype = "dashed", color = "red") +
  labs(x = "Sample Size per Group",
       y = "Power",
       title = "Power Analysis for Detecting Difference between  $p_1 = 0.7$  and  $p_2 = 0.5$ ") +
  theme_minimal()
```



The plot shows how the ability to detect a difference between $p_1=0.7$ and $p_2=0.5$ increases with sample size, starting at around 20% power with 20 samples per group and reaching 85% power with 100 samples per group. The red dashed line at 0.8 indicates the conventional target power level, suggesting that a sample size of slightly under 100 per group would be needed to achieve adequate statistical power for detecting this specific difference in proportions.