Simple A/B Test with Visualization

2023-11-15

### *I use this project to demonstrate my ability to conduct A/B analyses. For our purposes, I simulate data that represents collected watch time from two different advertisements, shown to 100 viewers each. I conduct an analysis using a Welch’s T-test to determine whether one ad may lead to increased watch time than the other.*

# Simulating data

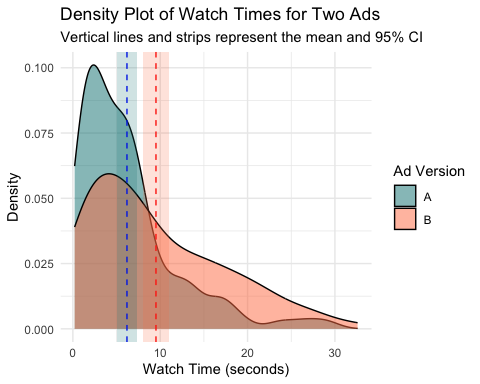
### In simulating the data, I create two vectors with 100 data points each, generated from different normal distributions. I chose to simulate with an exponential distribution because I imagine the dropoff of viewers at the beginning of an ad would be greater than towards the end, so a normal distribution might not be the best fit.

# Set seed for reproducibility  
set.seed(0)  
  
# Simulate watch times using an exponential distribution  
rate\_A <- 1/6 # Rate parameter for Ad A  
rate\_B <- 1/10 # Rate parameter for Ad B  
  
watch\_times\_A <- rexp(100, rate\_A)   
watch\_times\_B <- rexp(100, rate\_B)

# Visualizing

### I start by visualizing the data (using a density plot) to make sure it looks as expected, and plot the means and their 95% confidence intervals. A histogram could have also been used. For our purposes I think a density plot is prettier haha.

# Load necessary library  
library(ggplot2)  
  
# Combine the data into a data frame and label them  
watch\_times\_df <- data.frame(  
 Time = c(watch\_times\_A, watch\_times\_B),  
 Ad = factor(c(rep("A", length(watch\_times\_A)), rep("B", length(watch\_times\_B))))  
)  
  
# Calculate the means and standard errors for each ad  
mean\_watch\_time\_A <- mean(watch\_times\_A)  
mean\_watch\_time\_B <- mean(watch\_times\_B)  
se\_watch\_time\_A <- sd(watch\_times\_A) / sqrt(length(watch\_times\_A))  
se\_watch\_time\_B <- sd(watch\_times\_B) / sqrt(length(watch\_times\_B))  
  
# Calculate the 95% confidence intervals  
ci\_watch\_time\_A <- c(mean\_watch\_time\_A - 1.96 \* se\_watch\_time\_A, mean\_watch\_time\_A + 1.96 \* se\_watch\_time\_A)  
ci\_watch\_time\_B <- c(mean\_watch\_time\_B - 1.96 \* se\_watch\_time\_B, mean\_watch\_time\_B + 1.96 \* se\_watch\_time\_B)  
  
# Set colors  
color\_A <- "#008080"  
color\_B <- "#FF7F50"  
  
# Plot  
ggplot(watch\_times\_df, aes(x = Time, fill = Ad)) +  
 geom\_density(alpha = 0.5) +  
 geom\_vline(xintercept = mean\_watch\_time\_A, color = "blue", linetype = "dashed", linewidth = 0.5) +  
 geom\_vline(xintercept = mean\_watch\_time\_B, color = "red", linetype = "dashed", linewidth = 0.5) +  
 annotate("rect", xmin = ci\_watch\_time\_A[1], xmax = ci\_watch\_time\_A[2], ymin = -Inf, ymax = Inf, alpha = 0.2, fill = color\_A) +  
 annotate("rect", xmin = ci\_watch\_time\_B[1], xmax = ci\_watch\_time\_B[2], ymin = -Inf, ymax = Inf, alpha = 0.2, fill = color\_B) +  
 scale\_fill\_manual(name = "Ad Version", values = c(color\_A, color\_B)) +  
 labs(title = "Density Plot of Watch Times for Two Ads",  
 subtitle = "Vertical lines and strips represent the mean and 95% CI",  
 x = "Watch Time (seconds)",  
 y = "Density") +  
 theme\_minimal()

 ### On first impression alone, the means look to be different, given how their standard deviations do not overlap. The distributions are clearly not normal, so some caution should be taken when conducting the analysis.

# T-Test

### I chose to conduct a t-test to determine which mean is higher. A typical t-test (Student’s) assumes the data are normally distributed, independent, and of equal variance. The data is independent, but not normally distributed and have difference variances. Since there are many data points for each ad, (n=100>>30), because of the Central Limit Theorem, the results of this t-test should still be reliable. Since the variance looks different as well, I elect to use a Welch’s T-test rather than a standard T-test, because it is designed to handle that difference (and the test does not lose much power in doing so!).

### I am testing whether the ads have the same or different watch times. The null hypothesis is that they are the same, and I will be testing with alpha=0.05.

# Perform a t-test  
t\_test\_result <- t.test(watch\_times\_A, watch\_times\_B, var.equal = FALSE)  
  
# Output the result  
t\_test\_result

##   
## Welch Two Sample t-test  
##   
## data: watch\_times\_A and watch\_times\_B  
## t = -3.482, df = 182.16, p-value = 0.0006231  
## alternative hypothesis: true difference in means is not equal to 0  
## 95 percent confidence interval:  
## -5.202786 -1.439112  
## sample estimates:  
## mean of x mean of y   
## 6.192882 9.513831

# Conclusion

### The p-value being much less than 0.05 (0.000623) confirms our visual suspicion that the means are different. We have significant evidence to claim that ad B is associated with higher watch time and should be the advertisement chosen.