

# 19-703 Homework 1

Samuel Jones

January 28, 2018

## Simmons Situation B Example Replication

```
# Homework 1 write a simulation in R that replicates Situation B from
# Simmons et al. (2011). "Results for Situation B were obtained by conducting
# one t test after collecting 20 observations per cell and another after
# collecting an additional 10 observations per cell." (Simmons, et al 2011)

#This code is based on the code used for Situation A (Davis, 2018)
SituationB <- function(pval){
  # Group 1 generate 30
  Likers1 <- rnorm(30, 0, 1)
  # Group 2 generate 30
  Likers2 <- rnorm(30, 0, 1)
  # For the first t-test, use the first 20 observations
  t.test1 <- t.test(Likers1[1:20], Likers2[1:20], var.equal = TRUE)
  # For the second t-test, use all 30
  t.test2 <- t.test(Likers1[1:30], Likers2[1:30], var.equal = TRUE)
  # significant = 1 if at least one test gives p < pval, 0 otherwise
  significant <- ifelse(t.test1$p.value < pval | t.test2$p.value < pval, 1, 0)
  return(significant) #value returned by the function
}

replicates <- replicate(15000, SituationB(.05))
SitBMean <- mean(replicates)
SitBMean

## [1] 0.0758
```

## Do you agree with Feynman's Rule? Is it practical?

I do agree with Feynman's Rule. As scientists, especially those of us who work in the public policy arena, I think we have a responsibility to accurately convey what we know, what we don't know and what we think we know. We're all familiar with a quote made popular by Mark Twain, but of unknown origin, "There are three kinds of lies: lies, damned lies, and statistics." Unfortunately, it only takes a few irresponsible scientists (or politicians) to take liberties to forward their own agenda to erode public trust and reinforce the aforementioned quote. Practicality is another matter altogether. I'm sure it's difficult to get works published that undermine themselves by explaining all the ways the analysis could be wrong. If the work is published, will those who read it judge the work to be of less work because the scientist attempted to follow Feynmans' Rule. Since we are particularly concerned with policy makers implementing our findings, we also have a responsibility to ensure that the significance of our findings is properly communicated. Perhaps the best we can do is to ensure our analysis is as "bulletproof" as possible and to not oversell our findings.

**Which of the Neuroskeptic's 9 circles are most relevant for your field? Why, and what corrective policies do you think are necessary (if any)?**

I'm not an expert in my field yet, but from what I've seen so far, the fourth circle seems to be a danger. Perhaps it's only because my research is still in the exploratory stages that I feel that way. As to policy fixes, I'm not convinced policy can fix them. I think it's more of a culture change that needs to take place. Policies may be able to help nudge the culture changes, but if the culture change were to lead, the change could take place much quicker. If journals weren't so concerned with p-values, and looked at an article based on the strength of its argument and analysis regardless of if the p-value was "significant" or not. Perhaps there are better measures than the p-value, but if there are, I don't know what they are.

**When are the results of a t-test meaningless? In other words, what additional information do you need to make sense of the results of a t-test?**

T-tests are meaningless when the following 3 pieces of information are absent: "1) the estimated size of the population mean difference ( $\bar{Y}_1 - \bar{Y}_2$ ), 2) the sample variances  $\sigma_1^2$  and  $\sigma_2^2$ , and 3) the sample sizes n and m. With this information, you can calculate an estimate of the effect size". (Davis, 2018)

**Explain the two-sample t-test qualitatively and mathematically. What is in the denominator of a two-sample t-test? What would violations of independence within each group do to the denominator? What about violations of independence between groups?**

The two-sample t-test helps us to measure the difference between the population means.

$$t - test = \frac{EstimateOfAThing - TestValueOfThingUnderH_0}{StandardErrorOfEstimateOfAThing} \text{ or } t - test = \frac{(\bar{x} - \bar{y}) - (\mu_x - \mu_y)_{H_0}}{\sqrt{Var(\bar{x} - \bar{y})}} \text{ or } t = \frac{\bar{x} - \bar{y}}{\sqrt{\frac{\sigma_x^2}{n} + \frac{\sigma_y^2}{m}}}$$

Violation of independence within the sample groups increases the covariance between the groups, which increases the variance of the sample groups' means and decreases the t-value. Violation of independence between groups decreases covariance between the two groups, which will decrease the variance sample groups' means and increases the t-value. (Davis, 2018)

**Without doing any math, which would yield more bias in Simmons et al. (2011) Situation A: Two dependent variables that are highly correlated with each other or weakly correlated with each other?**

More bias would come from two highly correlated variables than two weakly correlated variables. This is most easily illustrated by thinking about two perfectly correlated variables. Effectively we'd only be testing one variable and thus only have one opportunity to reject the null hypothesis.

**Create an account at OSF or Harvard's Dataverse or GitHub.**

Please visit my GitHub repository for a copy of this pdf and the source code referenced therein.