

# HW05

107070008

**Question 1)** Use the slider bars of the simulation to the values your colleague found, and confirm from the visualization that we cannot reject the null hypothesis. Consider the scenarios (a – d) independently using the simulation tool. For each scenario, start with the initial parameters above, then adjust them to answer the following questions: i. Would this scenario create systematic or random error (or both or neither)? ii. Which part of the t-statistic or significance (diff, sd, n, alpha) would be affected? iii. Will it increase or decrease our power to reject the null hypothesis? iv. Which kind of error (Type I or Type II) becomes more likely because of this scenario?

a. You discover that your colleague wanted to target the general population of Taiwanese users of the product. However, he only collected data from a pool of young consumers, and missed many older customers who you suspect might use the product much less every day.

- i. Systematic error
- ii. Diff will be effected, because we should gain more samples to make our measurement representative.
- iii. Decrease the power to reject null hypothesis.
- iv. TypeII error

b. You find that 20 of the respondents are reporting data from the wrong wearable device, so they should be removed from the data. These 20 people are just like the others in every other respect.

- i. Random error
- ii. N will be effected. We will get less data to make our result correct.
- iii. Decrease the power to reject null hypothesis.
- iv. TypeII error

c. A very annoying professor visiting your company has criticized your colleague's "95% confidence" criteria, and has suggested relaxing it to just 90%.

- i. systematic error
- ii. It will affect the significant level of our measurement.
- iii. Increase the power to reject null hypothesis.
- iv. TypeI error

d. Your colleague has measured usage times on five weekdays and taken a daily average. But you feel this will underreport usage for younger people who are very active on weekends, whereas it over-reports usage of older users.

- i. systematic error
- ii. It will increase the sd of smartphone usage.
- iii. Increase the power to reject null hypothesis.
- iv. TypeII error

## Question 2)

```
verizon <- read.csv("verizon.csv")
time <- verizon$Time
```

a. Recreate the traditional hypothesis test of last week using high-level built-in functions of R: (you may have to see the R help documentation, google how to use them, or ask for help on Teams)

- (i) Use the `t.test()` function to conduct a one-sample, one-tailed t-test: report 99% confidence interval of the mean, t-value, and p-value

```
t.test(time, mu=7.6, alternative="less", conf.level=0.99)
```

```
##
## One Sample t-test
##
## data: time
## t = 2.5608, df = 1686, p-value = 0.9947
## alternative hypothesis: true mean is less than 7.6
## 99 percent confidence interval:
##      -Inf 9.360414
## sample estimates:
## mean of x
## 8.522009
```

- (ii) Use the `power.t.test()` function to tell us the power of the test.

```
hyp <- 7.6
power.t.test(n = length(time), delta = mean(time) - hyp,
             sd = sd(time), alternative = "one.sided")
```

```
##
## Two-sample t test power calculation
##
##          n = 1687
##      delta = 0.9220095
##        sd = 14.78848
##   sig.level = 0.05
##      power = 0.5657309
## alternative = one.sided
##
## NOTE: n is number in *each* group
```

b. Let's use bootstrapped hypothesis testing to re-examine this problem:

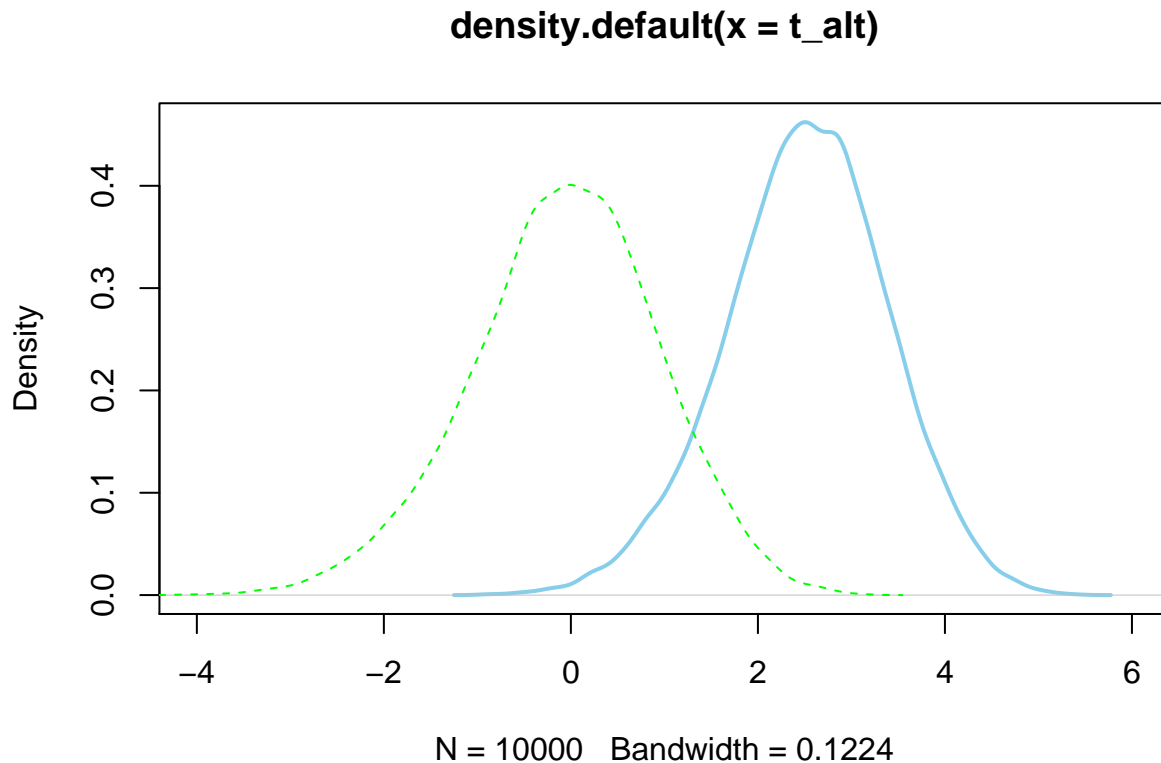
- (i) Retrieve the original t-value from traditional methods (above).

```
se <- sd(time) / sqrt(length(time))
t_value <- (mean(time) - hyp)/se
t_value
```

```
## [1] 2.560762
```

(ii) Bootstrap the null and alternative t-distributions.

```
bootstrap_null_alt <- function(sample0, hyp_mean) {  
  resample <- sample(sample0, length(sample0), replace=TRUE)  
  resample_se <- sd(resample) / sqrt(length(resample))  
  t_stat_alt <- (mean(resample) - hyp_mean) / resample_se  
  t_stat_null <- (mean(resample) - mean(sample0)) / resample_se  
  c(t_stat_alt, t_stat_null)  
}  
boot_t_stats <- replicate(10000, bootstrap_null_alt(time, hyp))  
t_alt <- boot_t_stats[1,]  
t_null <- boot_t_stats[2,]  
plot(density(t_alt), xlim=c(-4,6), col="skyblue", lwd = 2)  
lines(density(t_null), lty = "dashed", col = "green")
```

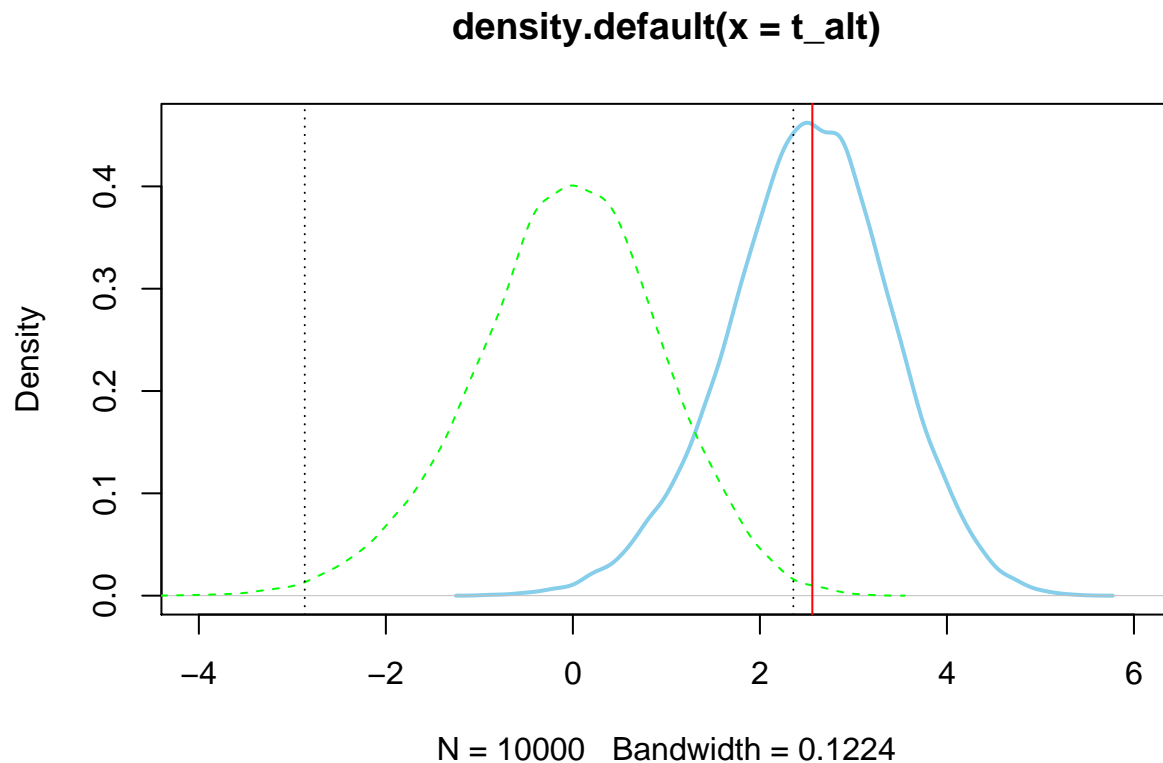


(iii) Find the 99% cutoff value for critical null values of  $t$  (from the bootstrapped null); What should our test conclude when comparing the original  $t$ -value to the 99% cutoff value?

```
cutoff_99 <- quantile(t_null, probs=c(0.005, 0.995))  
cutoff_99
```

```
##      0.5%      99.5%  
## -2.867418  2.357681
```

```
plot(density(t_alt), xlim=c(-4,6), col="skyblue", lwd = 2)
lines(density(t_null), lty = "dashed", col = "green")
abline(v=t_value, col = "red")
abline(v=cutoff_99, lty = "dotted")
```



(iv) Compute the p-value and power of our bootstrapped test.

```
null_probs <- ecdf(t_null)
one_tailed_pvalue <- 1 - null_probs(t_value)
one_tailed_pvalue
```

```
## [1] 0.0025
```

```
cutoff_99 <- quantile(t_null, probs=c(0.005, 0.995))
alt_probs <- ecdf(t_alt)
alt_probs(cutoff_99[1]) + (1 - alt_probs(cutoff_99[2]))
```

```
## [1] 0.5852
```