

1. When conducting the work of Lab 11, we conducted the test that uses the Central Limit Theorem even though the sample size was “small” (i.e.,  $n < 30$ ). It turns out, that how “far off” the  $t$ -test is can be computed using a first-order Edgeworth approximation for the error. Below, we will do this for the the further observations.

- (a) Boos and Hughes-Oliver (2000) note that

$$P(T \leq t) \approx F_Z(t) + \underbrace{\frac{\text{skew}}{\sqrt{n}} \frac{(2t^2 + 1)}{6} f_Z(t)}_{\text{error}},$$

where  $f_Z(\cdot)$  and  $F_Z(\cdot)$  are the Gaussian PDF and CDF and skew is the skewness of the data. What is the potential error in the computation of the  $p$ -value when testing  $H_0 : \mu_X = 0; H_a : \mu_X < 0$  using the zebra finch further data?

```
library(tidyverse)
finches.dat <- read_csv("zebrafinches.csv")

## Rows: 25 Columns: 3
## -- Column specification -----
## Delimiter: ","
## dbl (3): closer, further, diff
##
## i Use 'spec()' to retrieve the full column specification for this data.
## i Specify the column types or set 'show_col_types = FALSE' to quiet this message.

alpha <- 0.05
t <- t.test(x = finches.dat$further,
            alternative = "two.sided",
            mu = 0,
            conf.level = 0.95)

t.score.further <- -7.778

n <- length(finches.dat$further)

library(e1071)
further.skew <- skewness(finches.dat$further)

f.of.z <- dnorm(t.score.further, 0, 1)

F.of.z <- pnorm(t.score.further, 0, 1)

numerator.a <- further.skew / sqrt(n)
numerator.b <- (2*(t.score.further*t.score.further) + 1)/6

ans <- F.of.z + numerator.a*numerator.b*f.of.z
print(ans)

## [1] -1.189085e-13
```

- (b) Compute the error for  $t$  statistics from -10 to 10 and plot a line that shows the error across  $t$ . Continue to use the skewness and the sample size for the zebra finch further data.

```

t.vals <- seq(-10, 10, length.out = 1000)

calculate.error <- function(t.val){
  skew <- skewness(finches.dat$further)
  n <- length(finches.dat$further)

  f.of.z <- dnorm(t.val,0,1)
  F.of.z <- pnorm(t.val,0,1)

  numerator.a <- skew / sqrt(n)
  numerator.b <- (2*(t.val*t.val) + 1)/6

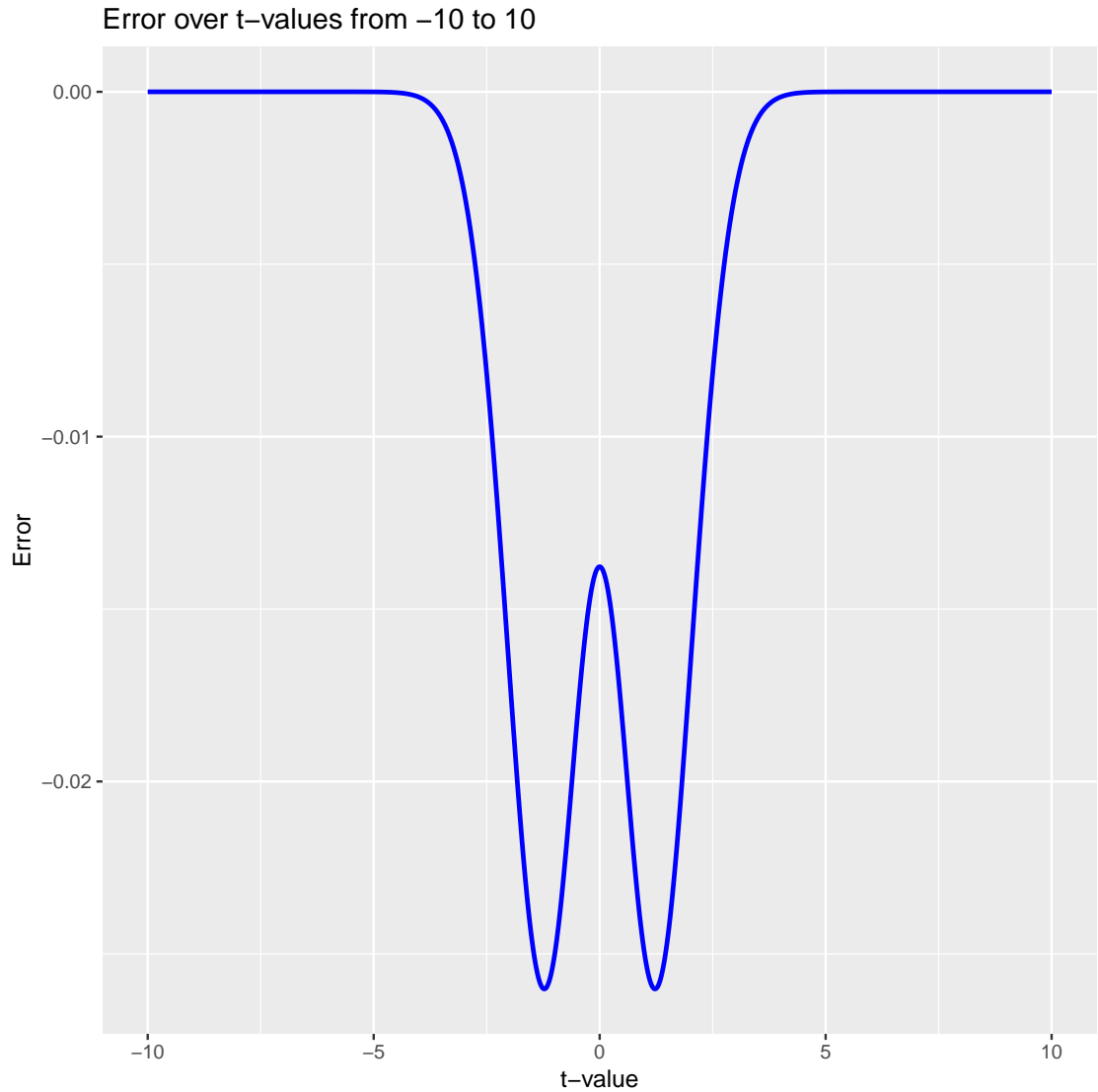
  ans <- numerator.a*numerator.b*f.of.z
  return(ans)
}

error.vals <- numeric(1000)
for (i in 0:length(t.vals)){
  t.val <- t.vals[i]
  error <- calculate.error(t.val)
  error.vals[i] = error
}

#plot the results
plot.data <- tibble(
  t = t.vals,
  metric = error.vals
)

ggplot(plot.data, aes(x = t, y = metric)) +
  geom_line(color = "blue", linewidth = 1.0) +
  labs(title = "Error over t-values from -10 to 10", x = "t-value", y = "Error")

```



- (c) Suppose we wanted to have a tail probability within 10% of the desired  $\alpha = 0.05$ . Recall we did a left-tailed test using the further data. How large of a sample size would we need? That is, we need to solve the error formula equal to 10% of the desired left-tail probability:

$$0.10\alpha \stackrel{\text{set}}{=} \underbrace{\frac{\text{skew}}{\sqrt{n}} \frac{(2t^2 + 1)}{6} f_Z(t)}_{\text{error}},$$

which yields

$$n = \left( \frac{\text{skew}}{6(0.10\alpha)} (2t^2 + 1) f_Z(t) \right)^2.$$

```
skew <- skewness(finches.dat$further)

alpha <- .05

t <- qnorm(alpha, 0, 1)
```

```
f.of.z <- dnorm(t,0,1)

part.a <- skew/ (6*(.10*alpha))
part.b <- ((2*t*t) + 1) * f.of.z

n <- (part.a*part.b)*(part.a*part.b)
print(n)

## [1] 520.8876
```

2. Complete the following steps to revisit the analyses from lab 11 using the bootstrap procedure.

- (a) Now, consider the zebra finch data. We do not know the generating distributions for the closer, further, and difference data, so perform resampling to approximate the sampling distribution of the  $T$  statistic:

$$T = \frac{\bar{x}_r - 0}{s/\sqrt{n}},$$

where  $\bar{x}_r$  is the mean computed on the  $r^{th}$  resample and  $s$  is the sample standard deviation from the original samples. At the end, create an object called `resamples.null.closer`, for example, and store the resamples shifted to ensure they are consistent with the null hypotheses at the average (i.e., here ensure the shifted resamples are 0 on average, corresponding to  $t = 0$ , for each case).

```
R <- 1000
mu0 <- 0

resamples.closer <- tibble(t = numeric(R))
for(i in 1:R){
  curr.sample <- sample(x= finches.dat$closer, size = length(finches.dat$closer), replace = T)
  resamples.closer$t[i] <- (mean(curr.sample) - mu0)/(sd(finches.dat$closer)/sqrt(length(finches
)})

resamples.further <- tibble(t = numeric(R))
for(i in 1:R){
  curr.sample <- sample(x= finches.dat$further, size = length(finches.dat$further), replace = T)
  resamples.further$t[i] <- (mean(curr.sample) - mu0)/(sd(finches.dat$further)/sqrt(length(finches
)})

resamples.difference <- tibble(t = numeric(R))
for(i in 1:R){
  curr.sample <- sample(x= finches.dat$diff, size = length(finches.dat$diff), replace = T)
  resamples.difference$t[i] <- (mean(curr.sample) - mu0)/(sd(finches.dat$diff)/sqrt(length(finches
)})

mean.t.closer <- mean(resamples.closer$t)
print(mean.t.closer)

## [1] 8.403617

mean.t.further <- mean(resamples.further$t)
print(mean.t.further)

## [1] -7.775494

mean.t.difference <- mean(resamples.difference$t)
print(mean.t.difference)
```

```
## [1] 8.54412
```

- (b) Compute the bootstrap  $p$ -value for each test using the shifted resamples. How do these compare to the  $t$ -test  $p$ -values?

```
t.obs.closer <- (mean(finches.dat$closer) - 0) / (sd(finches.dat$closer)/sqrt(length(finches.d
t.obs.further <- (mean(finches.dat$further) - 0) / (sd(finches.dat$further)/sqrt(length(finches
t.obs.difference <- (mean(finches.dat$diff) - 0) / (sd(finches.dat$diff)/sqrt(length(finches.d

resamples.null.closer <- resamples.closer$t - mean(resamples.closer$t)
resamples.null.further <- resamples.further$t - mean(resamples.further$t)
resamples.null.difference <- resamples.difference$t - mean(resamples.difference$t)

pval.closer <- mean(resamples.null.closer <= t.obs.closer)
pval.further <- mean(resamples.null.further <= t.obs.further)
pval.difference <- mean(resamples.null.difference <= t.obs.difference)

print(pval.closer)
## [1] 1
print(pval.further)
## [1] 0
print(pval.difference)
## [1] 1
```

- (c) What is the 5<sup>th</sup> percentile of the shifted resamples under the null hypothesis? Note this value approximates  $t_{0.05, n-1}$ . Compare these values in each case.
- (d) Compute the bootstrap confidence intervals using the resamples. How do these compare to the  $t$ -test confidence intervals?
3. Complete the following steps to revisit the analyses from lab 11 using the randomization procedure.
- Now, consider the zebra finch data. We do not know the generating distributions for the closer, further, and difference data, so perform the randomization procedure
  - Compute the randomization test  $p$ -value for each test.
  - Compute the randomization confidence interval by iterating over values of  $\mu_0$ .  
**Hint:** You can “search” for the lower bound from  $Q_1$  and subtracting by 0.0001, and the upper bound using  $Q_3$  and increasing by 0.0001. You will continue until you find the first value for which the two-sided  $p$ -value is greater than or equal to 0.05.
4. **Optional Challenge:** In this lab, you performed resampling to approximate the sampling distribution of the  $T$  statistic using

$$T = \frac{\bar{x}_r - 0}{s/\sqrt{n}}.$$

I’m curious whether it is better/worse/similar if we computed the statistics using the sample standard deviation of the resamples ( $s_r$ ), instead of the original sample ( $s$ )

$$T = \frac{\bar{x}_r - 0}{s_r/\sqrt{n}}.$$

- Perform a simulation study to evaluate the Type I error for conducting this hypothesis test both ways.
- Using the same test case(s) as part (a), compute bootstrap confidence intervals and assess their coverage – how often do we ‘capture’ the parameter of interest?

## References

Boos, D. D. and Hughes-Oliver, J. M. (2000). How large does  $n$  have to be for  $z$  and  $t$  intervals? *The American Statistician*, 54(2):121–128.