

- 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.

- 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)
library(xtable)
library(VGAM)
library(e1071)
library(pwr)
library(effectsize)
### Part A ### -> using further observations
dat.finch< read.csv("zebrafinches.csv")

further <- dat.finch$further
mu0 <- 0
t.further <- t.test(further,
                    mu = mu0,
                    alternative = "less")
t.stat.further <- t.further$statistic
n <- length(further)
skewness <- skewness(further)
fz <- dnorm(t.stat.further)
Fz <- pnorm(t.stat.further)

(edgeworth.approx.error <- (skewness/sqrt(n)) *
  (((2*(t.stat.further)^2 + 1)/6)*(fz)))

##          t
## -1.226006e-13

# Potential Error in the computation of the p-value is -1.226006e-13.
probability <- Fz + edgeworth.approx.error
```

- 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.

```
### Part B ###
t.values <- seq(-10,10, length = 1000)
fz2 <- dnorm(t.values)

error.vals <- (skewness/sqrt(n)) * (((2*(t.values)^2 + 1)/6)*(fz2))

error.tvals <- tibble(
  t = t.values,
  error = error.vals)
```

```

)
(summary(error.tvals))

##           t           error
##  Min.    :-10    Min.    :-2.602e-02
## 1st Qu.: -5     1st Qu.: -8.061e-03
## Median : 0      Median : -2.560e-06
## Mean   : 0      Mean   : -5.172e-03
## 3rd Qu.: 5      3rd Qu.: 0.000e+00
## Max.   : 10     Max.   : 0.000e+00

error.t.plot <- ggplot()+
  geom_line(data = error.tvals,
    aes(x = t, y = error),
    color = "lightblue")+
  theme_bw()+
  labs(title = "Edgeworth Approximation Error",
    x = "T Values (-10,10)",
    y = "Error")

```

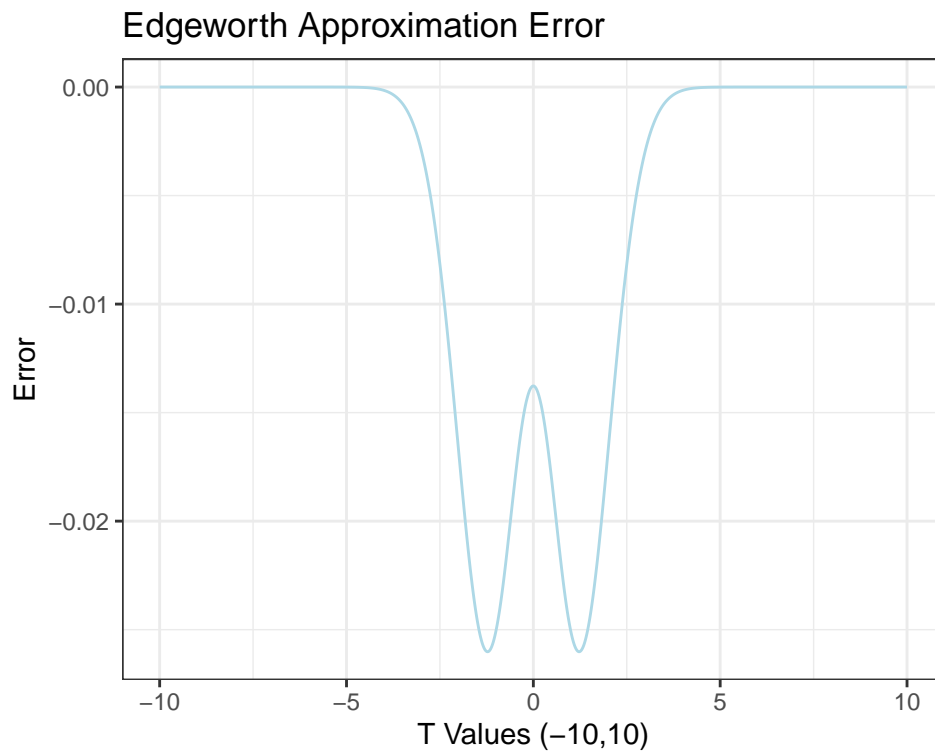


Figure 1: Edgeworth Approximation Error Across t values

- (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:

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

which yields

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

```
### Part C ### - error in the rejection region
skewness <- skewness(further)
tval <- qnorm(0.05)
fz3 <- dnorm(tval)

(min.n <- ((skewness/(6*(0.10*0.05))) * (2*tval^2 + 1) *fz3)^2)

## [1] 520.8876
```

We would need a sample size of $n = 520.8876$.

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).

```
library(boot)
### Part A ###

# Closer Resamples
R <- 10000
s.closer <- sd(dat.finchess$closer)
n.closer <- length(dat.finchess$closer)
resamples.closer <- tibble(tstat=rep(NA, R),
                           xbar=rep(NA, R))

for(i in 1:R){
  curr.resample <- sample(x = dat.finchess$closer,
                         size = n.closer,
                         replace = T)
  resamples.closer$tstat[i] <- (mean(curr.resample)-0)/(s.closer/sqrt(n.closer))
  resamples.closer$xbar[i] <- mean(curr.resample)
}

# shift so H0 is true
delta.closer <- mean(resamples.closer$tstat) - 0 # null mu0 = 0

resamples.null.closer <- resamples.closer |>
  mutate(tstat.shifted = tstat - delta.closer)

(mean(resamples.null.closer$tstat.shifted))

## [1] 8.914758e-16
```

```

# Further Resamples
R <- 10000
s.further <- sd(dat.finches$further)
n.further <- length(dat.finches$further)
resamples.further <- tibble(tstat=rep(NA, R),
                           xbar=rep(NA, R))

for(i in 1:R){
  curr.resample <- sample(x = dat.finches$further,
                        size = n.further,
                        replace = T)
  resamples.further$tstat[i] <- (mean(curr.resample)-0)/(s.further/sqrt(n.further))
  resamples.further$xbar[i] <- mean(curr.resample)
}

# shift so H0 is true
delta.further <- mean(resamples.further$tstat) - 0 # null mu0 = 0
resamples.null.further <- resamples.further |>
  mutate(tstat.shifted = tstat - delta.further)

(mean(resamples.null.further$tstat.shifted))

## [1] -1.207923e-17

# Difference Resamples
R <- 10000
s.diff <- sd(dat.finches$diff)
n.diff <- length(dat.finches$diff)
resamples.diff <- tibble(tstat=rep(NA, R),
                       xbar=rep(NA, R))

for(i in 1:R){
  curr.resample <- sample(x = dat.finches$diff,
                        size = n.diff,
                        replace = T)
  resamples.diff$tstat[i] <- (mean(curr.resample)-0)/(s.diff/sqrt(n.diff))
  resamples.diff$xbar[i] <- mean(curr.resample)
}

# shift so H0 is true
delta.diff <- mean(resamples.diff$tstat) - 0 # null mu0 = 0
resamples.null.diff <- resamples.diff |>
  mutate(tstat.shifted = tstat - delta.diff)

(mean(resamples.null.diff$tstat.shifted))

## [1] -8.820833e-16

```

- (b) Compute the bootstrap p -value for each test using the shifted resamples. How do these compare to the t -test p -values?
 - (c) What is the 5th 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.
- (a) 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

(b) Compute the randomization test p -value for each test.

(c) Compute the randomization confidence interval by iterating over values of μ_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)

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

- (a) Perform a simulation study to evaluate the Type I error for conducting this hypothesis test both ways.
- (b) 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.