

- 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?

```
finch.data <- read.csv("zebrafinches.csv")
n = nrow(finch.data)
further.sum <- finch.data |>
  summarize(
    mean = mean(further),
    sd = sd(further),
    skew = skewness(further)
  )

close.t.stat <- t.test(x=finch.data$further, mu = 0, alternative = "less")

t.stat <- close.t.stat$statistic

fZ.at.t <- dnorm(t.stat, mean = 0, sd = 1)

error.val <- ((further.sum$skew)/(sqrt(n)))*(((2*t.stat^2)+1)/6)*fZ.at.t
```

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

fZ <- dnorm(t.seq, mean = 0, sd = 1)

skew.further <- further.sum$skew

error.seq <- (skew.further / sqrt(n)) *
  ((2 * t.seq^2 + 1) / 6) *
  fZ

error.df <- tibble(
  t = t.seq,
  error = error.seq
)
error.df

## # A tibble: 1,000 x 2
##       t      error
##   <dbl>   <dbl>
## 1 -10    -5.34e-22
## 2 -9.98   -6.49e-22
## 3 -9.96   -7.90e-22
## 4 -9.94   -9.60e-22
## 5 -9.92  -1.17e-21
## 6 -9.90  -1.42e-21
## 7 -9.88  -1.72e-21
## 8 -9.86  -2.09e-21
## 9 -9.84  -2.53e-21
## 10 -9.82 -3.07e-21
## # i 990 more rows
```

- (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

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

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  $r^{\text{th}}$  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).

- (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 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.

- (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  $\mu_0$ .  
**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 ( $s$ )

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

- (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.