1. A group of researchers is running an experiment over the course of 30 months, with a single observation collected at the end of each month. Let $X_1, ..., X_{30}$ denote the observations for each month. From prior studies, the researchers know that

$$X_i \sim f_X(x),$$

but the mean μ_X is unknown, and they wish to conduct the following test

$$H_0 : \mu_X = 0$$
$$H_a : \mu_X > 0.$$

At month k, they have accumulated data $X_1, ..., X_k$ and they have the t-statistic

$$T_k = \frac{\bar{X} - 0}{S_k/\sqrt{n}}.$$

The initial plan was to test the hypotheses after all data was collected (at the end of month 30), at level $\alpha = 0.05$. However, conducting the experiment is expensive, so the researchers want to "peek" at the data at the end of month 20 to see if they can stop it early. That is, the researchers propose to check whether t_{20} provides statistically discernible support for the alternative. If it does, they will stop the experiment early and report support for the researcher's alternative hypothesis. If it does not, they will continue to month 30 and test whether t_{30} provides statistically discernible support for the alternative.

(a) What values of t_{20} provide statistically discernible support for the alternative hypothesis? - **At month 20, because we obtained a critical t-value using the right-tailed test, we can reject the null hypothesis if T20 $>$ 1.729.**

```
alpha = 0.05
t20.crit <- qt(1-alpha, df = 20 - 1)
t20.crit

## [1] 1.729133
```

(b) What values of t_{30} provide statistically discernible support for the alternative hypothesis? - **At month 30, using the same right tailed test to obtain a critical t-value, we can reject the null hypothesis if T30 $>$ 1.699.**

```
alpha = 0.05
t30.crit <- qt(1-alpha, df = 30 - 1)
t30.crit

## [1] 1.699127
```

(c) Suppose $f_X(x)$ is a Laplace distribution with $a = 0$ and $b = 4.0$. Conduct a simulation study to assess the Type I error rate of this approach.
Note: You can use the `rlaplace()` function from the `VGAM` package for `R` (Yee, 2010). - **By conducting a simulation study under the assumtion $f_X(x)$ is a Laplace distribution with $a = 0$ and $b = 4.0$, we were able to calculate an overall rate of type 1 errors of 7.19 Percent.**

```
n.sim = 10000
#1c

sim.tbl <- tibble(
early_reject = logical(n.sim),
overall      = logical(n.sim)
)


for (i in 1:n.sim) {
x <- rlaplace(30, location = 0, scale = 4)

# interim look
t20 <- mean(x[1:20]) / (sd(x[1:20]) / sqrt(20))
sim.tbl$early_reject[i] <- (t20 > t20.crit)

if (sim.tbl$early_reject[i]) {
```

```
    # if rejected early, overall is also TRUE
    sim.tbl$overall[i] <- TRUE
  } else {
    # final look
    t30 <- mean(x) / (sd(x) / sqrt(30))
    sim.tbl$overall[i] <- (t30 > t30.crit)
  }
}


sim.tbl.rates <- sim.tbl |>
mutate(late_reject = !early_reject & overall)|>
summarize(
  early_rate   = mean(early_reject),
  late_rate    = mean(late_reject),
  overall_rate = mean(overall)
)
sim.tbl.rates

## # A tibble: 1 x 3
##   early_rate late_rate overall_rate
##        <dbl>     <dbl>        <dbl>
## 1     0.0491    0.0273       0.0764
```

(d) **Optional Challenge:** Can you find a value of $\alpha < 0.05$ that yields a Type I error rate of 0.05?

2. Perform a simulation study to assess the robustness of the T test. Specifically, generate samples of size $n = 15$ from the Beta(10,2), Beta(2,10), and Beta(10,10) distributions and conduct the following hypothesis tests against the actual mean for each case (e.g., $\frac{10}{10+2}$, $\frac{2}{10+2}$, and $\frac{10}{10+10}$).

   (a) What proportion of the time do we make an error of Type I for a left-tailed test? **- The simulation test we have performed below suggests that we make a type I error while using a left-tailed test 3.02% of the time for the Beta(10,2) distribution, 7.81% of the time for the Beta(2,10) distribution, and 5.18% of the time for the Beta(10,10) distribution.**

```
n_sim      <- 10000
n          <- 15
df         <- n - 1

tcrit_left <- qt(alpha, df=df)

tcrit_right <- qt(1 - alpha, df = df)
# two-sided critical value
tcrit_2s   <- qt(1 - alpha/2, df = df)

#List of parameters for each distribution
params <- list(
Beta_10_2  = list(a=10, b=2,  mu=10/12),
Beta_2_10  = list(a=2,  b=10, mu=2/12),
Beta_10_10 = list(a=10, b=10,mu=0.5)
)

####Rejection rate when using a left tailed test
type1_rates_left <- sapply(params, function(p) {
rejections <- replicate(n_sim, {
  x     <- rbeta(n, p$a, p$b)
  tstat <- (mean(x) - p$mu) / (sd(x) / sqrt(n))
  tstat < tcrit_left
})
mean(rejections)
})
type1_rates_left

##  Beta_10_2  Beta_2_10 Beta_10_10
##     0.0291     0.0766     0.0498
```

   (b) What proportion of the time do we make an error of Type I for a right-tailed test? **- The simulation test we have performed below suggests that we make a type I error while using a right-tailed test 7.74% of the time for the Beta(10,2) distribution, 2.75% of the time for the Beta(2,10) distribution, and 4.85% of the time for the Beta(10,10) distribution.**

```
type1_rates_right <- sapply(params, function(p) {
rejections <- replicate(n_sim, {
  x     <- rbeta(n, p$a, p$b)
```

```
  tstat <- (mean(x) - p$mu) / (sd(x) / sqrt(n))
  # reject if t-statistic is too large
  tstat > tcrit_right
})
mean(rejections)
})
type1_rates_right

##  Beta_10_2  Beta_2_10 Beta_10_10
##     0.0776     0.0273     0.0469
```

(c) What proportion of the time do we make an error of Type I for a two-tailed test? - **The simulation test we have performed below suggests that we make a type I error while using a two-tailed test** 5.75% **of the time for the Beta(10,2) distribution,** 6.28% **of the time for the Beta(2,10) distribution, and** 5.08% **of the time for the Beta(10,10) distribution.**

```
type1_rates_two <- sapply(params, function(p) {
  mean(replicate(n_sim, {
    x     <- rbeta(n, p$a, p$b)
    tstat <- (mean(x) - p$mu) / (sd(x) / sqrt(n))
    (tstat < -tcrit_2s) | (tstat > +tcrit_2s)
  }))
})
type1_rates_two

##  Beta_10_2  Beta_2_10 Beta_10_10
##     0.0603     0.0602     0.0551
```

(d) How does skewness of the underlying population distribution effect Type I error across the test types? - **The direction skewness of each distribution appears to have a positive cooreation with the direction of the tail of each test, as the rate of type I errors occurring is the lowest for each distribution in the same direction that each distribution is skewed.**

Distribution	Left_Tailed	Right_Tailed	Two_Tailed
Beta(10,2)	0.0302	0.0781	0.0518
Beta(2,10)	0.0774	0.0275	0.0485
Beta(10,10)	0.0575	0.0628	0.0508

Table 1: Empirical Type I Error Rates by Test Type and Distribution

References

Yee, T. W. (2010). The VGAM package for categorical data analysis. *Journal of Statistical Software*, 32(10):1–34.