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? **Solution:** We can calculate the value of t_{20} that provide discernible support for the alternative by finding the t-value that is where the $\alpha = 0.05$ cutoff is. For t_{20}, that value is 1.7291, so any values greater than or equal to that provide support for the alternative.

```
#values of t20 that provide support
alpha = 0.05
n = 20  # sample size
df <- n - 1  # degrees of freedom
t_crit_20 <- qt(1 - alpha, df = df)
```

(b) What values of t_{30} provide statistically discernible support for the alternative hypothesis? **Solution:** Using the same approach as for t_{20}, but replacing n=20 with n=30, we find that the $\alpha = 0.05$ t-value is 1.6991, and consequently any values greater than this provide statistically discernible support for the alternative.

(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).

Solution: To conduct the simulation study, we take n number of observations (X) from the Laplace distribution $a = 0$ and $b = 4.0$. For this set of X, we then calculate the t-statistic using the formula provided above, as we can calculate the mean and standard deviation for the simulated observations and we also know n (either 20 or 30). Then, we compare this t-statistic to the critical points we calculated in a) and b), where if the simulated t-statistic is greater than the critical point, we reject the null hypothesis. By counting the number of rejections in the 10000 samples, and then dividing it by the total number of samples (10000), we can calculate the Type I error. We find that for t_{20}, the $TypeIerror = 0.0510$ and for t_{20} $TypeIerror = 0.0494$. These values are close to the expected $\alpha = 0.05$.

```
#Simulation Study for Type I error
R = 10000 #number of sims
alpha = 0.05
set.seed = 151
a = 0
b = 4
n = 20 #t_20
rejects = 0 #will add to this for each time we reject H0

for (i in 1:R) {
x <- rlaplace(n, location = a, scale = b)
#take observation (x) from the Laplace dist. n times
t = mean(x) / (sd(x) / sqrt(n)) #T-stat formula
```

```r
#Compare to critical point calculated in a) and b)
if (t > t_crit_20){
  rejects = rejects + 1
}
}
TypeI_20 = rejects/R

n = 30 #t_30
rejects = 0 #will add to this for each time we reject H0

for (i in 1:R) {
x <- rlaplace(n, location = a, scale = b)
#take observation (x) from the Laplace dist. n times
t = mean(x) / (sd(x) / sqrt(n)) #T-stat formula

#Compare to critical point calculated in a) and b)
if (t > t_crit_30){
  rejects = rejects + 1
}
}
TypeI_30 = rejects/R
```

(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?

Solution: We conducted the simulation in the same way as for calculating the Type I error for the Laplace distribution. The code was repeated the same for parts b) and c) as in part a), where just α and β for the Beta distribution were changed. All the results can be seen in the table after part c.

```r
##################################
# Question 2
##################################
R = 10000
n = 15
#critical t-point which is a = 0.05
alpha <- 0.05
df <- n - 1  # degrees of freedom

t15_rtail = qt(1 - alpha, df = df)  #right-tailed text
t15_ltail = qt(alpha, df = df) #left-tailed test
t15_2tail_low = qt(alpha/2, df = df) # left side of two tailed
t15_2tail_high = qt(1 - alpha/2, df = df) #right side of two tailed

##########
#Beta(10,2)
##########
a = 10
b = 2
mu = a/(a+b)

left_reject = 0
right_reject = 0
twotail_reject = 0

for (i in 1:R){
  x = rbeta(n = n, shape1 = a, shape2 = b)
  #left-tailed test
  t_left = t.test(x, alternative = "less", mu = mu)$statistic
  if (t_left < t15_ltail){
    left_reject = left_reject + 1
  }
  #right-tailed test
  t_right = t.test(x, alternative = "greater", mu = mu)$statistic
  if (t_right > t15_rtail){
    right_reject = right_reject + 1
  }
  #two-tail test
  t_twotail = t.test(x, alternative = "two.sided", mu = mu)$statistic
  if (t_twotail < t15_2tail_low | t_twotail > t15_2tail_high){
    twotail_reject = twotail_reject + 1
```

```
  }
}


#Calculating the Type I error
TypeI_left_102 = left_reject/R
TypeI_right_102 = right_reject/R
TypeI_2tail_102 = twotail_reject/R
```

(b) What proportion of the time do we make an error of Type I for a right-tailed test?

(c) What proportion of the time do we make an error of Type I for a two-tailed test?

	Beta	left_tailed	right_tailed	two_tailed
1	Beta(10,2)	0.03	0.08	0.06
2	Beta(2,10)	0.08	0.03	0.06
3	Beta(10,10)	0.05	0.05	0.05

(d) How does skewness of the underlying population distribution effect Type I error across the test types?

Solution: The Beta(10,2) distribution is left skewed. As such, it has a lower Type I error for the left-tailed distribution and a higher error for the right-tailed. The inverse is true for the Beta(2,10) as it is right skewed. Finally, as Beta(10,10) is symmetric, it makes sense that the left and right-tailed tests result in the same Type I error, which would also match the two-tailed test.

References

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