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?

```
## [1] 1.729133
```

(b) What values of t_{30} provide statistically discernible support for the alternative hypothesis?

```
## [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).

```
## [1] 0.0708
```

(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? (Proportions are given in order of $Beta$(10,2), $Beta$(2,10), $Beta$(10,10))

```
## [1] 0.0298 0.0816 0.0469
```

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

```
## [1] 0.0788 0.0289 0.0507
```

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

```
## [1] 0.0592 0.0617 0.0472
```

(d) How does skewness of the underlying population distribution effect Type I error across the test types? (Answer: The t-test assumes a relatively symmetrical distribution for data, so skewness could inflate or deflate a testing error. As we can see from this example, the direction of the skewness would decrease the Type I error in that direction and increase the error in the other direction. Using the two tailed test as the normal, we can compare it against the skewness and Type I errors of the other graphs. For example, the $Beta$(10,2) distribution is skewed left, and it produced a lower Type I error to the left but a higher proportion of it to the right. These error proportions are flipped for the $Beta$(2,10) distribution which is skewed right.)

References

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