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?
   The critical t value is 1.729 after 20 months.

   (b) What values of t_{30} provide statistically discernible support for the alternative hypothesis?
   The critical t value is 1.699 after 30 months.

   (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).
   The type 1 error for this approach is 0.0735.

   (d) **Optional Challenge:** Can you find a value of $\alpha < 0.05$ that yields a Type I error rate of 0.05?
   If a desired Type 1 error is 0.05, the alpha that should be used is 0.034.

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?

```
##  Beta_10_2  Beta_2_10 Beta_10_10
##     0.0270     0.0781     0.0512
```

   The type 1 errors for each of the following distributions for the left tailed test are listed above.

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

```
##  Beta_10_2  Beta_2_10 Beta_10_10
##     0.0772     0.0325     0.0494
```

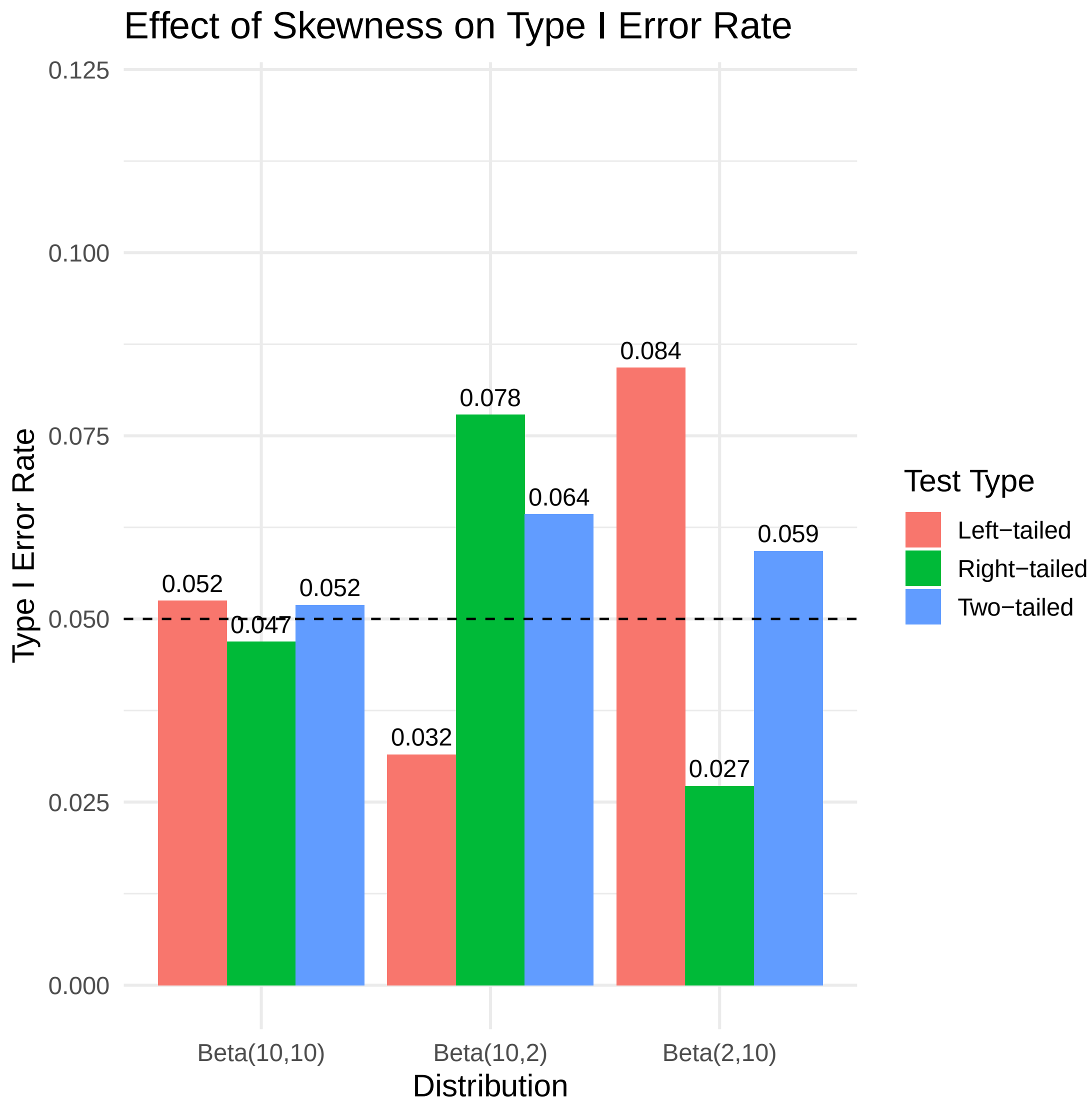
   The type 1 errors for each of the following distributions for the right tailed test are listed above.

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

```
##  Beta_10_2  Beta_2_10 Beta_10_10
##     0.0571     0.0603     0.0545
```

   The type 1 errors for each of the following distributions for the two tailed test are listed above.

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



The t-test assumes that the data is symmetric, so when the data is skewed, the data becomes biased in the direction of the skew, leading to an incorrect Type 1 error.

References

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