Section 8.1: Hypotheses and Testing Procedure

Concepts and Formulae:

- Suppose there is a claim or a statement.
- The null hypothesis is an answer in which the claim or the statement is true, and denoted by H_0.
- The alternative hypothesis is an answer in which the claim or the statement is NOT true, and denoted by H_a (H_A or H_1).
- If we conclude H_0, we say "we accept H_0 or fail to reject H_0"; if we conclude H_a, we say "we reject H_0 and conclude H_a".

A testing procedure is specified by the following steps:

- Choose a test statistic. A test statistic is a function of data which makes the decision about the rejection or acceptance of H_0.

- A rejection range is a set in which we reject H_0 if the test statistic is in the set and accept H_0 if the test statistic is outside of the set.

- Let T be the test statistic and R be the rejection range. Then, we conclude

$$\begin{cases} H_0 & \text{if } T \notin R \\ H_a & \text{if } T \in R \end{cases}$$

In general, a test can be summarized into the following table.

Conclude	Truth H_0	H_a
H_0	Correct	Type II Error
H_a	Type I Error	Correct

- Type I error rate (probability) is defined by

$$P(\text{Conclude } H_a | \text{Truth is } H_0)$$

  and type II error rate (probability) is defined by

$$P(\text{Conclude } H_0 | \text{Truth is } H_a).$$

- We want to make both probabilities small. However, this is usually impossible in real applications. Thus classically, we only control type I error probabilities.

- The significance level denoted by α is the maximum of type I error probabilities. Therefore, if we choose $\alpha = 0.05$, we guarantee the type I error probability is less than or equal to 0.05.

First example of Section 8.1: Examples 8.1 and 8.3 on textbook. Suppose old rate of "no visible damage" is 25%. An experiment took 20 samples and want to know whether the rate increases with a new method.

- Null hypothesis is

  $$H_0 : \text{rate does not increase}$$

  versus the alternative

  $$H_a : \text{rate increases.}$$

- Under H_0, the count X of visible damage follows $Bin(20, 0.25)$. Under H_a, $X \sim Bin(20, p)$ with $p > 0.25$. Then, we can write

  $$H_0 : p = 0.25 \leftrightarrow H_a : p > 0.25.$$

- Assume the rejection range is

  $$R = \{8, 9, 10, 11, \cdots, 20\}.$$

- Then, the type I error probability is

$$P(X \geq 8|p = 0.25) = 1 - B(7; 20, 0.25) \\ = 0.102.$$

- The significance level $\alpha = 0.102$.

- The type II error probability is

$$P(X < 8|p > 0.25) = B(7; 20, p)$$

which is a function of p.

- Suppose we change the testing problem as

$$H_0 : p \leq 0.25 \leftrightarrow H_a : p > 0.25.$$

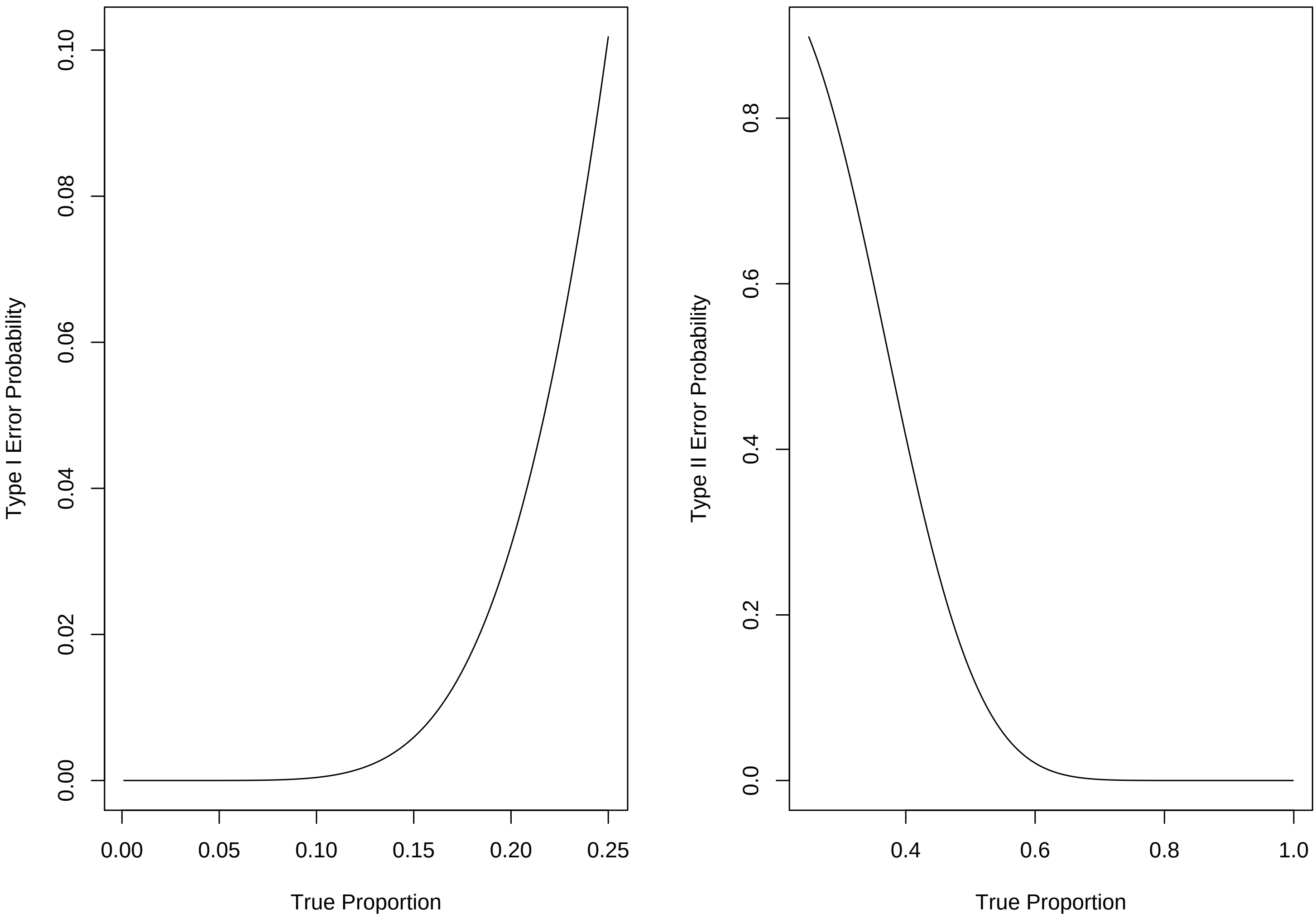
Then, the type I error probability is

$$1 - B(7; 20, p)$$

for $p \leq 0.25$ and the type II error probability is

$$B(7; 20, p)$$

for $p > 0.25$. In this case, we still have the significance level $\alpha = 0.102$.



Graph of Type I and Type II error probabilities

The significance level α is the maximum of type I error probabilities, which gives

$$\alpha = 0.102.$$

Second example of Section 8.1: Examples 8.2 and 8.4 on textbook. Assume we choose 25 samples from $N(\mu, 81)$ and suppose we test

$$H_0 : \mu = 75 \leftrightarrow H_a : \mu < 75.$$

Suppose we use $R = \bar{X} < 70.8$, i.e., we conclude

$$\begin{cases} H_0 & \text{if } \bar{X} \geq 70.8 \\ H_a & \text{if } \bar{X} < 70.8 \end{cases}$$

Then, the type I error is

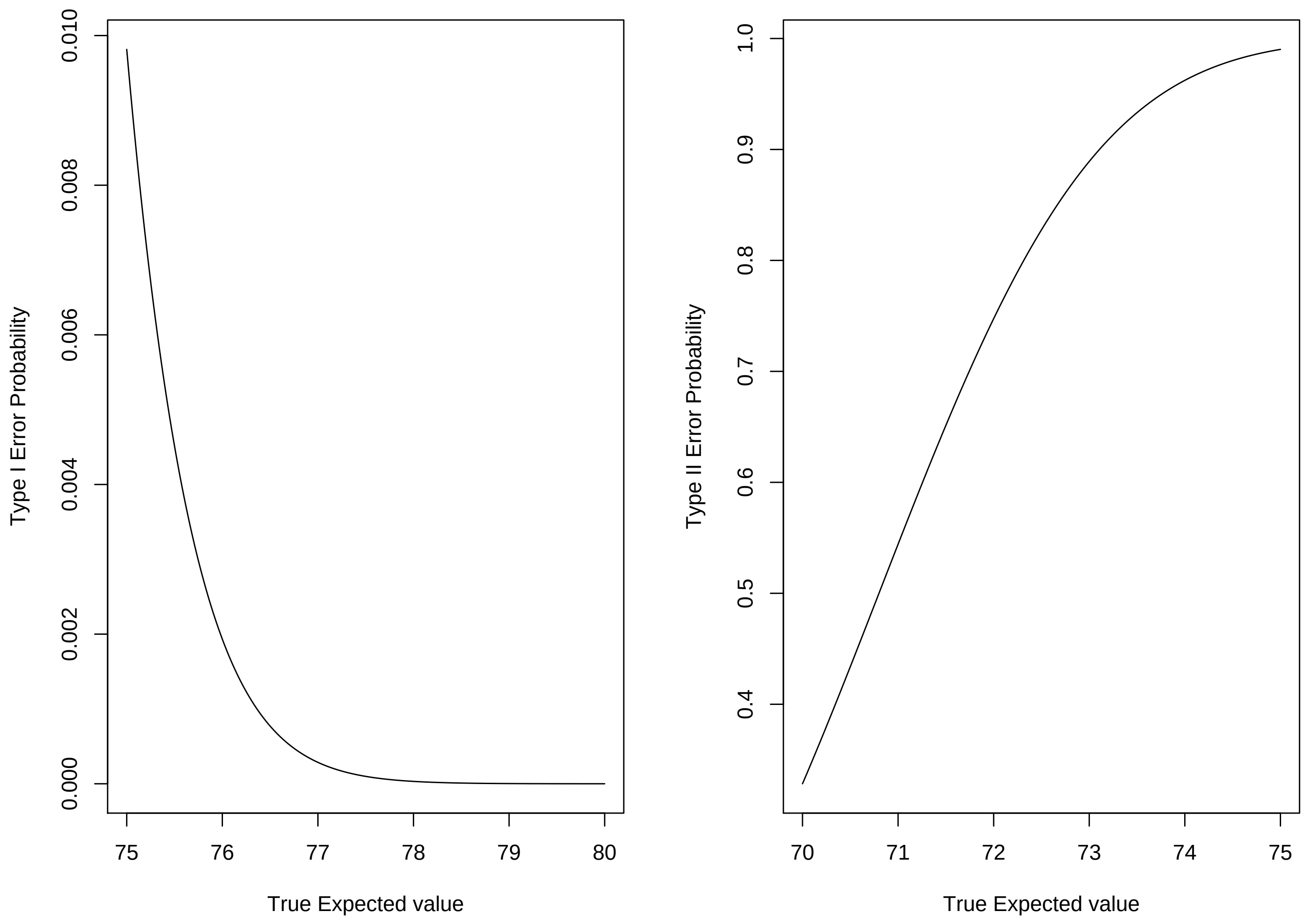
$$\begin{aligned} & P(\bar{X} \leq 70.8 \text{ if } \mu = 75) \\ = & \Phi(\frac{70.8 - 75}{\sqrt{81/25}}) \\ = & \Phi(-2.33) = 0.01. \end{aligned}$$

Then, the significance level is $\alpha = 0.01$ and type II error probability is

$$\beta(\mu) = P(\bar{X} \geq 70.8) = 1 - \Phi(\frac{70.8 - \mu}{1.8}).$$

for $\mu < 75$, which is a function of μ for $\mu < 75$. For example, we have

$$\beta(72) = 1 - \Phi(-0.67) = 0.7486.$$



Graph of Type I and Type II error probabilities

The significance level α is the maximum of type I error probabilities, which gives

$$\alpha = 0.01.$$