Statistical Inference and Hypothesis Testing in Data Science Applications

DTSA 5003 offered on Coursera

by the University of Colorado, Boulder

Instructor: J.N. Corcoran

Week One
Lesson Four
Errors in Hypothesis Testing

A "Rejection Rule"

As the operators of the salmon hatchery, we are going to test the hypotheses

$$H_0 : \mu \leq 28 \qquad \text{versus} \qquad H_1 : \mu > 28.$$

To do this, we will take a random sample of size n:

$$X_1, X_2, \ldots, X_n \overset{iid}{\sim} N(\mu, \sigma^2)$$

and look at the sample mean $\overline{X}$. We assume that H_0 is true but are looking for evidence that H_1 might be true. H_1 is saying that the true mean for the entire population is larger than we are assuming. How would this be reflected in $\overline{X}$? If H_1 were true, we would expect $\overline{X}$ to be larger than 28. Per our previous discussion though, it is entirely possible for $\overline{X}$ to be larger than 28 even if the true mean is not. We actually won't be comfortable saying that that μ is larger than 28 until $\overline{X}$ is "significantly" larger.

In summary, we assume that H_0 is true but decide to reject it, in favor of H_1 if

$$\overline{X} > c$$

for some "large" constant c that we deem to be "significant".

Types of Errors in Hypothesis Testing

The null hypothesis H_0 is either true or it is false. We are assuming it is true and will reject it if our rejection criterion is met. We will never "accept" H_0 as true. The two possibilities are that we "reject H_0" or "fail to reject H_0". It is kind of like the American justice system where someone is pronounced "guilty" or "not guilty" but no one is ever pronounced "innocent".

Let's go back to the fish length example and our decision rule for rejection of H_0 when $\overline{X}$ is "large". Suppose that we have determined that $c = 29.5$ is our definition of "large". We will reject H_0, in favor of H_1 if $\overline{X} > 29.5$. (This cutoff has been completely made up at this point. We will learn how to find a reasonable value of c soon!)

It is possible that H_0 is true, that $\mu \leq 28$ and yet we observe $\overline{X} > 29.5$ and decide to reject H_0. Similarly, is possible that H_0 is false, that $\mu > 28$ and yet we observe an $\overline{X}$ that is below 29.5 and decide not to reject H_0. In both cases, we have made an error. The two types of error, called Type I error and Type II error, are summarized in the following table.

		Your Decision	
		Fail to Reject H_0	Reject H_0
Reality	H_0 True	✓	**Type I Error**
	H_0 False	**Type II Error**	✓

Neither type of error (Type I or Type II) is inherently more serious. The errors must be interpreted within the context of a problem.

In the next Lesson, we will discuss how to find the cutoff for rejection of H_0 . It will be based on a pre-specified tolerance for error.