Topic 20: Hypothesis Testing

02-680: Essentials of Mathematics and Statistics

November 25, 2024

Once we have taken our data and tried to fit a model to it, we often want to know if that fit model matches our original assumptions about the data. We call this ***hypothesis testing***, and we use these techniques to put an actual value on this match.

We add some terminology on what we've been talking about with respect to statistics and frame things as follows: we first define our ***alternate hypothesis***, which we will denote H_1, this will be the set of events that define what we want to ask about the confidence in happening; we then define the ***null hypothesis***, which we denote H_0, this is the set of events that are all outcomes other than than our alternate. So lets say were asking if a drug has a measurable impact on cholesterol, the null hypothesis would be that cholesterol stayed the same and the alternate hypothesis would be that it changed.

We usually refer to a hypothesis test telling us if we should ***reject*** or ***retain*** the null hypothesis.

1 Defining Errors

Errors occur when the hypothesis test tells us something thats wrong. So in the example above if the test tells us to reject the null (that is, its confident that the cholesterol changed) but in reality it didn't change we call this a ***Type I*** error. On the other hand if our test tells us to retain the null but in reality the value *did* change we call that a ***Type II*** error.

		Hypothesis Test Result	
		Retain H_0	Reject H_0
Truth	H_0		Type I Error
	H_1	Type II Error	

We say the

Type I Error Rate is $p(\text{reject } H_0, H_0 \text{ is true})$,

Type II Error Rate is $p(\text{retain } H_0, H_1 \text{ is true})$, and

Statistical ***Power*** is 1 - Type II Error Rate.

The last point means that the higher power tests have a stronger ability to detect signals for H_1.

Let's look at it visually, first for what we call a ***two-sided*** test, that is $H_0 : \mu = x$ and $H_1 : \mu \neq x$.