# Math 3215: Lecture 8

Will Perkins

February 2, 2012

## 1 The Statistical Method

Today is a bit of change of gears: we've been doing basic probability, but now we'll take a detour and talk about statistics. The moral of today's story is that nothing really is new: the statistical method is just a framework for using probability models to test hypotheses. It's important to understand the precise framework well, however. Today we will learn the statistical method in its most basic form.

## 2 The Null Hypothesis

A statistical experiment begins with an assumption about the world that you want to disprove. This assumption must be specific enough that you can build a probability model around it and calculate probabilites explicitly.

**Examples of null hypotheses**

- The coin is fair (probability of heads $= \frac{1}{2}$
- The average height of a Georgia Tech student is 6 feet (and the heights are distributed like a bell curve)
- The machine makes an average of one bad nail every $10,000$ nails
- 48% of voters will vote for Abe Lincoln
- Shoppers choice of milk (full vs skim) is independent of how much they spend
- SAT prep classes do not affect SAT scores

**Bad examples (these cannot be null hypotheses!)**

- The coin is not fair
- More voters prefer Lincoln than prefer Douglas
- The average height of Georgia Tech students is above 5 feet
- Shoppers choice of milk has some dependence on how much they spend
- SAT prep classes improve SAT scores

Often the null hypothesis is denoted $H_0$.

## 3 The Alternate Hypothesis

The *alternate hypothesis*, denoted $H_1$, is the opposite of the null hypothesis. Unlike the null hypothesis, the alternate hypothesis need not exactly specify a probablility distribution.

### 3.1 One-sided vs. Two-sided alternatives

Consider the null hypothesis "The coin is fair." The probability distribution under the null hypothesis has $\Pr[Heads] = \frac{1}{2}$. What is the alternate hypothesis? There are three options:

1. The coin is not fair ($p \neq .5$)
2. The coin is biased in favor of heads ($p > .5$)
3. The coin is biased in favor of tails ($p < .5$)

The first alternate hypothesis is a two-sided alternative, the second two are both one-sided alternatives.

- For the null hypotheses listed in section 1, come up with a two-sided alternate hypothesis and one or two one-sided alternate hypotheses.

## 4 Possible Conclusions

Before we describe a statistical test of the null hypothesis, let's describe the possible conclusions we can draw with the statistical method. *There are only two possible conclusions*:

1. We reject the null hypothesis with high confidence (and specify our level of confidence)
2. We do not have enough evidence to reject

*Important note*: we never confirm the null hypothesis with the statistical method, we simply declare that we have not seen enough evidence to reject it.

## 5 p-values and Confidence

### 5.1 Confidence Level

How do we come to such a conclusion? The first step, before we even observe data, is to set a confidence level, often denoted by $\alpha$. $\alpha$ is a meaure of how confident we must be to reject the null hypothesis; the smaller the $\alpha$, the more confident we must be. A common choice of $\alpha$ in social science is .05, and in other sciences .01 or even .001.

### 5.2 p-value

The *p-value* for an experiment is the probability that we would see what we observed or something more extreme under the null hypothesis.

- At its heart, the p-value is simply a probably calculation under the distribution described by $H_0$.
- 'What we saw or something more extreme' depends on whether $H_1$ is one-sided or two-sided.

Example:

- $H_0 = \{p = .5\}$. We flip the coin 10 times and get 7 heads. Under each of the three possible alternate hypotheses, describe the event 'What we saw or something more extreme'
- Calculate the p-value in each case.

## 6 Conclusion

We finish our test as follows:

- if $p \leq \alpha$, then we reject the null hypothesis. We can report $\alpha$ and $p$
- If $p > \alpha$, then we report that we do not have enough evidence to reject. We can report $p$.