David Youngberg
BSAD 210—Montgomery College

LECTURE 19: HYPOTHESES AND TYPES OF ERROR

I. Hypotheses
   a. *Null Hypothesis*—assertion corresponding to the default position, where there is no significant different, or where nothing is happening
      i. Example: Income ***doesn't*** predict how much you spend
   b. *Alternative Hypothesis*—assertion that claims there is a significant difference.
      i. Example: Income ***does*** predict how much you spend
   c. Alternative hypothesis can be either one-tailed or two-tailed.
      i. Example: More income predicts that you spend less (one-tailed), that you spend more (one-tailed), or that you spend either more or less (two-tailed).
      ii. We usually focus on two-tailed tests.
   d. *Level of significance*—determines a cut-off point when the null hypothesis is rejected or failed to be rejected. The standard is 95% (or 1.96).
   e. There are two types of mistakes you can make when working with a null hypothesis.
      i. *Type I error*—false positive; a non-match is declared a match—or when you reject the null hypothesis and you should fail to reject it
      ii. *Type II error*—false negative; a good match is not detected—or when you fail to reject the null hypothesis and you should reject it

Examples

Type I	*Type II*
Convicting an innocent person	Letting the guilty go free
Approving a damaging drug	Rejecting a beneficial drug
Befriending a jerk	Ignoring a nice person
Funding a poor investment	Passing on a good investment

   f. Type I and Type II errors are equally undesirable, but Type II errors are insidious because they are harder to notice when they happen.
      i. *In general*, Type I errors are self correcting; Type II errors are not. But precisely because Type I errors are self correcting, the

fact that one made an error at all is evident thus there is a tendency for people to commit Type II errors.

II. Simple tests of hypothesis

a. Suppose you take a sample and you're interested if its difference from the mean (higher or lower) is because of random chance or because there's something special about the sample you took.

b. You know the population standard deviation:

$$z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}}$$

c. You don't know the population standard deviation:

$$t = \frac{\bar{x} - \mu}{s/\sqrt{n}}$$

d. You are using a proportion:

$$z = \frac{p - \pi}{\sqrt{\frac{\pi(1-\pi)}{n}}}$$

i. Where π is the estimate of the population's proportion and p is the sample estimate.