Math 146 8.1 - Basics of Hypothesis Tests

A *hypothesis* is a **claim** about a property or characteristic of a population.
Examples:

"Normal body temperature is 98.6°F", "Mean income is \$33000/yr", "The drug is effective for 60% of patients"

A ***hypothesis test*** is a procedure to test the validity of that statement or claim.

Question? Does a sample of 40 people with a mean temperature of 98.2 °F provide enough evidence to support the claim that normal body temperature is less than 98.6°F? (We'll do this one in section 8.3.)

COMPONENTS OF A HYPOTHESIS TEST

1) The Null Hypothesis is statement of the population parameter *equaling* a value indicating no change or no effect. From the statements above:

$H_0: \mu = 98.6\,°\text{F}$ $\quad$ $H_0: \mu = \$33000/\text{yr}$ $\quad$ $H_0: p = 0.60$

2) The Alternative Hypothesis is the statement or claim if the **null hypothesis is false**:

$H_1: \mu < 98.6\,°\text{F}$ $\quad$ $H_1: \mu \neq \$33000/\text{yr}$ $\quad$ $H_1: p > 0.60$

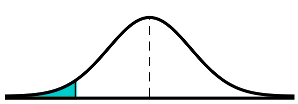
(Sometimes H_a: is used for the alternative hypothesis.)

Important!!!

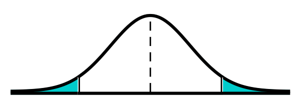
- If you are trying to ***support*** a claim, make it the ***alternative*** hypothesis.
- If you are trying to ***reject*** a claim, make it the ***null*** hypothesis.

Possible types of hypothesis set-ups:

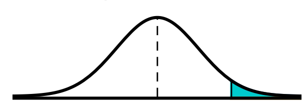
$H_0: \mu = 98.6\,°\text{F}$
$H_1: \mu < 98.6\,°\text{F}$
Left-tail test



$H_0: \mu = \$33000/\text{yr}$
$H_1: \mu \neq \$33000/\text{yr}$
Two-tailed test



$H_0: p = 0.42$
$H_1: p > 0.42$
Right-tail test



The "tail", or type, is determined by the alternative hypothesis enequality.

Question? In a hypothesis test we either reject the null or fail to reject the null. So, how do we determine if a null hypothesis should be rejected (or not rejected)?

Answer: Design an experiment, survey, etc., to collect good data. Using the Central Limit Theorem and probabilities, we can determine if the result of the experiment, *assuming the null is true*, is a rare "impossible" event, or a typically expected event. If it's "rare" we reject the null, if it's not "rare" we fail to reject the null.

Method I (**The Traditional Method**): compare the test statistic against a critical value z^* (or t^*) to determine how "rare" a sample statistic is, i.e., if the test statistic falls in a shaded region in the normal curves above, it's a rare event and we reject the null. If the statistic falls in the unshaded region, the event is not rare and we fail to reject the null.

The test statistics are:

proportions: $z = \frac{\hat{p}-p}{\sqrt{\frac{pq}{n}}}$ $\qquad$ **means:** $z = \frac{\bar{x}-\mu}{\sigma/\sqrt{n}}$ or $t = \frac{\bar{x}-\mu}{s/\sqrt{n}}$

Method II (**The P-value Method**) We compare the probability of obtaining the test statistic to a predetermined α, say, $\alpha = 0.05$, or 5%. If the probability of getting our test statistic result, ***assuming the null hypothesis is true***, is less than 5%, then we conclude that that is highly unlikely, and therefore conclude the null is false and **reject** the null hypothesis.

EXAMPLE 1 (hypothesis test for a population proportion)

Suppose a researcher claims the drug is actually effective for fewer than 60% of patients. Here's the setup:

null-hypothesis: $H_0 : p = 0.60$

alternative hypothesis: $H_1 : p < 0.60$ **Claim**

alpha level $\alpha = 0.05$

Notice that we made the "claim" the alternative in order to attempt to *support* the claim.

Now, suppose we conduct an experiment on 500 people and find that 267 reacted favorably to the treatment of the drug. The probability of this result, assuming the null is true, is only $p = 0.0013$. Since this is less than $\alpha = 0.05$, we conclude that it would be such a rare event that the null statement of 60% **must** be false. We reject the null hypothesis and support the claim that the true percentage is actually less than 60%.

Since most current technology can calculate p-values so easily, that is now typically the preferred method.

Errors from Hypothesis Tests

Consider the null hypothesis: H_0 : the patient is NOT sick.

Two "good" outcomes of a hypothesis test are:

(1) rejecting a false null hypothesis (the patient is sick, we reject the null, and conclude he is sick)

(2) failing to reject a true null hypothesis (the patient is not sick, we fail to reject the null, and conclude he's not sick.)

But, there are also two possible "bad" conclusions or errors:

(1) If a *true* null hypothesis is rejected (the patient is not sick, but we conclude he is sick), this is called a **Type I** error (also,a false positive).

(2) If a *false* null hypothesis is not rejected (the patient IS sick and we conclude he is not), this is called a **Type II** error (a false negative).

The probability of a Type I error is α (this is the α we choose, in that we are willing to make a Type I error 5% of the time if $\alpha = 0.05$), and the probability of a Type II error is β, which is difficult to calculate. Also, α and β can be controlled depending on the consequences of the error. Note: $1 - \beta$ is called *the power of the test*, and is used to gauge the effectiveness of the test.

Example 2*:* Write down the null and alternative hypotheses given the claim "*The parachute is packed correctly*". Identify the Type I error and the Type II error (and the correct outcomes). Which is better $\alpha < \beta$ or $\alpha > \beta$?

Example 3*:* You are the first EMT on the scene of an accident, and discover the first victim of the crash. Write down the null and alternative hypotheses given the claim "*The victim is alive*". Identify the Type I error and the Type II error (and the correct outcomes). Which is better $\alpha < \beta$ or $\alpha > \beta$?