Our third test covers the material from 3/31 to 4/26 on the syllabus—significance tests.

A basic overview of the material

The major idea of this material is a test of hypotheses or significance tests. You should consult sections 10-2 and 3 in the book to make sure you are familiar with the vocabulary terms.

The basic idea is that we want to understand something about an **unkown** population parameter. For us this is either μ, the population mean, or p the population proportion. We use the sample mean, $\bar{X}$, or the sample proportion, $\hat{p}$ to make the inference. How this is done involves a test of two competing hypotheses: the null hypotheses, H_0 and the alternative hypotheses, H_A. To help decide which hypothesis is preferable a **test statistic** is used whose distribution under the null hypothesis is known.

An observed value of the test statistic can be used two ways to see if the results are statistically significant. The p-value can be computed, as is done on the computer. This is compared to the significance level. Alternatively, in class we used the significance level to find the acceptance and rejection regions. We can then decide to accept or reject depending on the observed value of the test statistic.

The basic skills involve 3 steps:

1. Specify the two hypotheses
2. Decide on a test statistic and compute its observed value
3. Compute a p-value or the rejection region and compare.

We have a few different tests to keep straight:

Test for mean when variance is known or n is large This test has hypotheses

$$H_0 : \mu = a, \quad H_A : \mu < a \text{ or } \mu \neq a \text{ or } \mu > a$$

where a is some number. The alternative is one of 3 possibilities: less, two-sided or greater.

For these assumptions the test statistic

$$Z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} \text{ (when } \sigma \text{ is known) or } \frac{\bar{x} - \mu}{s/\sqrt{n}}$$

has a standard normal distribution.

Test for mean when n is small This test has the same hypotheses as the previous, only when n is small the distribution of the test statistic is different. We use

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

which will have the t-distribution with $n - 1$ degrees of freedom under the null hypothesis.

Two sample test of difference of means When testing the difference of means the hypotheses are

$$H_0 : \mu_1 = \mu_2, \quad H_A : \mu_1 <, \neq, > \mu_2$$

(one of the three signs). The test statistic used is

$$T = \frac{\bar{x} - \bar{y}}{s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}},$$

where s_p is the pooled standard deviation and is defined by

$$s_p^2 = \frac{(n_1 - 1)s_x^2 + (n_2^2 - 1)s_y^2}{n_1 + n_2 - 2}.$$

Under H_0 and an assumption that the unknown variances are equal, the distribution of T is the t-distribution with $n_1 + n_2 - 2$ degrees of freedom.

Test for a proportion The test for a sample proportion involves the hypotheses

$$H_0 : p = a, \quad p <, \neq, > a,$$

where a is some number between 0 and 1. The test statistic used is

$$Z = \frac{\hat{p} - p}{\sqrt{p(1 - p)/n}}.$$

Under H_o this is a standard normal so this is used to compute p-values.

Some sample problems

The problems should seem similar after awhile as the procedure is the same: formulate the hypotheses, find the observed value, compute the p-value.

For each of these questions, write down your two hypotheses, your test statistic and how you found the p-value. The first one is done for you,

1. An informal poll of 100 found that 66 thought Mick Jagger becoming a Knight was silly (unless he brought back the head of a dragon). Do a test of significance to see if the true population percentage is more than 60 percent?
   An answer to this would look like this:
   This is a test of proportion with the hypotheses

   $$H_0 : p = .60, \quad H_A : p > .60$$

   The test statistic is

   $$Z = \frac{\hat{p} - p}{\sqrt{p(1 - p)/n}}$$

   with observed value

   $$\frac{\frac{66}{100} - .60}{\sqrt{.6(.4)/100}} = 1.2$$

The critical value is found from the normal table or the t-table using infinite degrees of freedom. For a significance level of 0.05 the critical value is 1.645. We then see that 1.2 is in the acceptance region.

To do this with p-values we compute

$$P(Z > 1.2) = 1/2 - .3849 = .12$$

The p-value fo 0.12 is more than 0.05 so again the same result: we would accept the null hypothesis.

2. A poll of 2003 poll of 2018 Canadians found that 1070 supported gay marriage. Do a test of significance to see if this indicates that the true percentage for Canadians is more than 50%.

3. A 2003 European opinion poll of George Bush found only 16 percent of 200 people supported George Bush. Does this indicate that the population proportion of those who support George Bush is different than the 0.5 proportion in the US?

4. Last month the average time to park on campus was 8 minutes. This month (as more students have dropped out) it seems to take less time. Suppose, the sample average for 10 trips is 7 minutes with a sample standard deviation of 2 minutes. Does this indicate that the average time is less or is the difference explainable by sampling variation?

5. A test to determine if echinacea is beneficial in treating the common cold was setup as follows. If a child reported cold symptoms then they were randomly assigned to be given a treatment of echinacea or a placebo treatment. The time to recover was measured and is summarized in the table below

```
group       n   sample mean  sample sd
--------------------------------------
echinacea   200     5.3         2.5
placebo     207     5.4         2.5
```

Is this evidence that the echinacea group had a quicker recovery?

6. Is the line at the DMV improved? Historically, it took 65 minutes to do something at the DMV. After changes were made a sample of 21 people found their time was only 59 minutes. Is this evidence that the mean time has decreased?

7. A machine sometimes needs recalibration. It should have a mean of 16 and variance of 4, but sometimes the mean gets out of calibration. In a sample of 10 runs the sample mean was 15.8. Is this evidence that the population mean is different from 16?