HYPOTHESIS TESTS I

Video 1. Watch the introduction to hypothesis tests.

We start with a **null hypothesis** H_0, which we'll assume to be true until we have evidence to the contrary. Exactly what constitutes contrary evidence is determined by our choice of **alternative hypothesis** H_1. The null hypothesis should specify the distribution of a **test statistic**. We collect data, calculate the observed value of the test statistic, then calculate the **P-value** (or observed significance level) of our data. We **reject** H_0 in favor of H_1 if the p-value is small (usually less than 0.05). Otherwise we **fail to reject** H_0.

Method. For tests about a proportion, the null hypothesis should be $H_0 : \theta = \theta_0$ and our test statistic is the sample total T. Under H_0, $T \sim \text{binom}(n, \theta_0)$.

1. A July, 2018 NPR/IPSOS poll asked respondents if they support or oppose "building a wall or fence along the entire U.S./Mexico border." Let θ be the proportion of the population in question (e.g. Midwesterners) that **oppose** building a wall or fence. Test $H_0 : \theta = 0.5$ against $H_1 : \theta > 0.5$ for the following populations. State your p-values and conclusions clearly.

a) 115 of 217 people in the Midwest oppose the wall

b) 150 of 264 people in the West oppose the wall

c) 198 of 401 people in the South oppose the wall

Date: March 25, 2020.

2. The article "Analysis of Reserve and Regular Bottlings: Why Pay for a Difference Only the Critics Claim to Notice?" reported on an experiment to determine if wine tasters could correctly distinguish between reserve and regular versions of a wine. In each trial tasters were given 4 indistinguishable containers of wine, two of which contained the regular version and two of which contained the reserve version of the wine. The taster then selected 3 of the containers, tasted them, and was asked to identify which one of the 3 was different from the other 2. In 855 trials, 346 resulted in correct distinctions. Does this provide compelling evidence that wine tasters can distinguish between regular and reserve wines?

a) Start with a null hypothesis that the tasters can't distinguish between the wines. State this as a hypothesis about the proportion of times the tasters correctly identify the odd wine out.

b) State an alternative hypothesis.

c) Calculate the observed significance level of the experimental results.

d) What is your conclusion? Any additional comments/thoughts?

Method. For tests about the mean of a normally distributed population (with unknown variance), the null hypothesis should be $H_0 : \mu = \mu_0$ and our test statistic is $T = \dfrac{\overline{X} - \mu_0}{S/\sqrt{n}}$. Under H_0, T has a t-distribution with $n - 1$ degrees of freedom.

3. The mean unrestrained compressive strength for a random sample of 45 specimens of a particular type of brick was 3107 psi, and the sample standard deviation was 188 psi. Does the data indicate that the true average unrestrained compressive strength is less than the design value of 3200? Assume the population is normally distributed and test $H_0 : \mu = 3200$ against $H_1 : \mu < 3200$ at a significance level of 0.05.

Definition. When conducting a hypothesis test there are two types of error:

(1) **Type I error** is rejecting H_0 when H_0 is actually true.
(2) **Type II error** is failing to reject H_0 when H_0 is actually false.

The **significance level** of a test is the probability of a type I error.

Ideally, the probability of both types of error would be small. Unfortunately, decreasing the probability of type I error (by lowering the threshold for significance from 0.05 to 0.01, say) increases the probability of type II error (and vice versa).

4. The EPA has determined that the Maximum Contaminant Level Goal (MCLG, "The level of a contaminant in drinking water below which there is no known or expected risk to health") for nitrates in drinking water is 10mg/L. Imagine you have been hired by the City of Spokane to monitor drinking water safety. Your plan is to collect a random sample of water from 25 different sources and test the samples for nitrate levels, then use the sample mean as an estimate of mean level in all of Spokane's water.

a) State your null and alternative hypotheses in both non-mathematical language and as $H_0 : \mu =$ ___ with a corresponding H_1.

b) State in plain language what a Type I error would be for your hypotheses (e.g. if you were explaining your results to the Mayor).

c) Repeat for a Type II error.

d) What significance level do you think makes sense for the test?

e) The EPA statement on the effect of nitrate contamination: "Infants below the age of six months who drink water containing nitrate in excess of the MCL could become seriously ill and, if untreated, may die. Symptoms include shortness of breath and blue-baby syndrome." Does this change the significance level you want to use for your test? Does this change how you want to set up the null and alternative hypotheses?