Example 1:

A 1996 report from the US Consumer Product Safety Commission claimed that 90% of all American homes have at least one smoke detector. A city's fire department has been running a public safety campaign about smoke detectors with posters, billboards, and ads on radio and TV and in the newspaper. The city wonders if the effort has raised the local level about the 90% national rate. Building inspectors visit 400 randomly selected homes and find that 376 have detectors. Is this strong evidence that the local rate is higher than the national rate?

So…in the example above we made a decision from the z test statistic. Let’s answer the same question using a P-value. **What is a P-Value?**

The key outcome of an inferential statistical test is a *P* value, which is the probability of obtaining a result—such as a difference in blood pressure change between a treatment and control group—as extreme or more extreme than the one observed if no true effect existed in the population.

Consider a study on the effects of a new drug for treating hypertension. Over the study period, the experimental and control groups experienced a decrease in systolic blood pressure of 6.2 mm Hg and 1.2 mm Hg, respectively. Is the 5.0 mm Hg difference truly due to the effects of the drug, or is the difference simply due to chance factors? The statistical test would yield a *P* value, which is the probability of observing a 5.0 mm Hg or greater difference in systolic blood pressure in repeated experiments if the null hypothesis were true. Let's say that the statistical test revealed a *P* value of .99. Here's the interpretation: If it's really true that the drug has no effect on systolic blood pressure, we would observe a 5.0 mm Hg difference between experimental and control subjects in 99 out of 100 repeated experiments. If we obtained such a result so frequently, we would be confident that the null hypothesis is tenable and that the drug doesn't really reduce blood pressure.

What if the test revealed a *P* value of .01? Here's the interpretation: If it's really true that the drug has no effect on systolic blood pressure, we would observe a 5.0 mm Hg difference between experimental and control subjects in 1 out of 100 repeated experiments. If we obtained such a result so rarely under the assumption that the null hypothesis is true, we would have to doubt that the null hypothesis is tenable. We would thus accept the research hypothesis that the drug does reduce blood pressure. We would say that the 5.0 mm Hg difference is statistically *significant*.

As the *P* value gets lower (i.e., closer to 0% and farther away from 100%), researchers are more inclined to accept the research hypothesis and to reject the null hypothesis. As you know, before beginning studies scientists conventionally establish cutoff points, which are called *alpha values*, to indicate what *P* values they will accept as significant. In many physiological studies, alpha values are set at .05 or .01. As you know, if the *P* values that are calculated in statistical tests are less than alpha—for example, if $P < .05$— the researchers would conclude that their study results are statistically significant.

A relatively simple way to interpret *P* values is to think of them as representing how likely a result would occur by chance. For a calculated *P* value of .001, we can say that the observed outcome would be expected to occur by chance only 1 in 1,000 times in repeated tests on different samples of the population.

How do we find the P-Value?

Conclusion from the P-Value

Example 2:

There are supposed to be 20% purple M&M's. Suppose a bag of 122 has only 21 of the purple ones. Does this contradict the company's 20% claim? This a two-tail test since we are equally interested to discover that there are too many.

Example 3: A union spokesperson claims that 75% of union members will support a strike if their basic demands are not met. A company negotiator believes that the true percentage is lower and runs a hypothesis test at the 10% significance level. What is the conclusion if 87 out of an SRS of 125 union members say they will strike?

Try!

A cancer research group surveys 500 women more than 40 years old to test the hypothesis that 28% of women in this age group have regularly scheduled mammograms. Should the hypothesis be rejected at the 5% level of significance if 151 of the women responded affirmatively?

Homework #10

Test all at a 5% level of significance.

1. In a group of 371 Pitt students, 42 were left-handed. Is this significantly lower than the proportion of all Americans who are left-handed, which is .12?
2. A university has found over the years that out of all the students who are offered admission, the proportion who accept is 0.70. After a new director of admissions is hired, the university wants to check if the proportion of students accepting has changed significantly. Suppose they offer admission to 1200 students and 888 accept. Is this evidence of a change from status quo?