* 1. **Tests About a Population Proportion**

**The One-Sample z Test for a Proportion**

Choose an SRS of size n from a large population that contains an unknown proportion p of successes. To test the hypothesis H0 : p = p0, compute the z statistic

 

Find the P-value by calculating the probability of getting a z statistic this large or larger in the direction specified by the alternative hypothesis Ha:

 

The same three conditions we used for confidence intervals must also be satisfied:

* **Random**
* **Normal**: np0 > 10 and n(1 - p0) > 10
* **Independent**: population is at least 10 times as large as the sample.

**Carrying Out a Significance Test (4-step process)**

Recall our basketball player who claimed to be an 80% free-throw shooter. A teammate doubted him and told him to prove it. In an SRS of 50 free-throws, he made 32. His sample proportion of made shots, 32/50 = 0.64, is much lower than what he claimed.

Does it provide convincing evidence against his claim at a 5% significance level?

**Step 1: State the hypotheses and parameter of interest**

**Step 2: Check the conditions**

**Step 3: Calculate the test statistic (z-score) and find the p-value**

**Step 4: Interpret the result in context of the problem**

**Example**: Potato Chips

A potato-chip producer has just received a truckload of potatoes from its main supplier. If the producer determines that more than 8% of the potatoes in the shipment have blemishes, the truck will be sent away to get another load from the supplier. A supervisor selects a random sample of 500 potatoes from the truck. An inspection reveals that 47 of the potatoes have blemishes. Carry out a significance test at the α= 0.10 significance level. What should the producer conclude? (Use the 4-step process)

**Two –sided Tests**

 

According to the Centers for Disease Control and Prevention (CDC) Web site, 50% of high school students have never smoked a cigarette. Taeyeon wonders whether this national result holds true in his large, urban high school. For his AP Statistics class project, Taeyeon surveys an SRS of 150 students from his school. He gets responses from all 150 students, and 90 say that they have never smoked a cigarette. What should Taeyeon conclude? Give appropriate evidence to support your answer.

**Step 1: State the hypotheses and parameter of interest**

**Step 2: Check the conditions**

**Step 3: Calculate the test statistic (z-score) and find the p-value**

**Step 4: Interpret the result in context of the problem**

**Example**: Coin Toss

Suppose a coin toss turns up 12 heads out of 20 trials. At .05 significance level, can one reject the null hypothesis that the coin toss is fair?

**Why Confidence Intervals Give More Information**

The result of a significance test is basically a decision to reject *H0* or fail to reject *H0*. When we reject *H0*, we’re left wondering what the actual proportion *p* might be. A confidence interval might shed some light on this issue.

Taeyeon found that 90 of an SRS of 150 students said that they had never smoked a cigarette. Find a 95% confidence interval for the actual proportion of students at his high school that have never smoked cigarettes.

**Confidence Intervals and Two-Sided Tests**

There is a link between confidence intervals and two-sided tests. The 95% confidence interval gives an approximate range of p0’s that would not be rejected by a two-sided test at the α = 0.05 significance level.

* A two-sided test at significance level α (say, α = 0.05) and a 100(1 –α)% confidence interval (a 95% confidence interval if α = 0.05) give similar information about the population parameter.
* If the sample proportion falls in the “fail to reject *H0*” region, like the green value in the figure, the resulting 95% confidence interval would include *p0*. In that case, both the significance test and the confidence interval would be unable to rule out *p0* as a plausible parameter value.
* However, if the sample proportion falls in the “reject *H0*” region, the resulting 95% confidence interval would not include *p0*. In that case, both the significance test and the confidence interval would provide evidence that *p0* is not the par

**Example #1**: Rafael Nadal

Does Rafael Nadal have different results when playing on clay? In a 2-year period, overall Rafa won 136 out of 162 matches. On clay, he won 35 out of 39 matches. Conduct a 5% significance test to determine if his results are different on clay. Then construct a 95% confidence interval to confirm your results.

**Example #2**: Financial Audit

When an accounting firm audits a company’s financial records for fraud, there is no fraud when about 30.1% of the numbers in financial records begin with the digit 1. However, if the proportion of first digits that are 1 is significantly different from 0.301 in a random sample of records, a much more thorough investigation of the company is done. Suppose that a random sample of 300 expenses from a company’s financial records results in only 68 expenses that begin with the digit 1. Should a more thorough investigation of this company? Conduct a 5% significance test. Then construct a 95% confidence interval to confirm your results.