- **Significance Testing with one proportion**

  Visit the Hyperstat site to learn about what significance testing looks like when dealing with one proportion. Be sure to only work through the 4 pages on this topic.

  A manufacturer is interested in whether people can tell the difference between a new formulation of a soft drink and the original formulation. The new formulation is cheaper to produce so if people cannot tell the difference, the new formulation will be manufactured. A sample of 100 people is taken. Each person is given a taste of both formulations and asked to identify the original. Sixty-two percent of the subjects correctly identified the new formulation. Is this proportion significantly different from 50%?

  1. The first step in hypothesis testing is to specify the null hypothesis and an alternative hypothesis. In testing proportions, the null hypothesis is that π, the proportion in the population, is equal to some specific value. In this example, the null hypothesis is that $\pi = 0.5$. The alternate hypothesis is $\pi \neq 0.5$.

  2. The second step is to choose a significance level. Assume the 0.05 level is chosen.

  3. The third step is to compute the difference between the sample proportion (p) and the value of π specified in the null hypothesis. In this example, $p - \pi = 0.62 - 0.5 = 0.12$.

  4. The fourth step is to compute p, the probability (or probability value). It is the probability of obtaining a difference between the proportion and the value specified by the null hypothesis as large or larger than the difference observed in the experiment. The general formula for significance testing as applied to this problem If p is greater than π then the formula is:

     $$z = \frac{p - \pi - \frac{0.5}{N}}{\sigma_p}.$$

     If p is less than π then the formula is:

     $$z = \frac{p - \pi + \frac{0.5}{N}}{\sigma_p}$$

     Note that the correction always makes z smaller.

     N is the sample size and σ_p is the standard error of a proportion. The formula for a standard error of a proportion is:

$$\sigma_p = \sqrt{\frac{\pi(1-\pi)}{N}}$$.

The term $\frac{0.5}{N}$ is the correction for continuity. For the present problem,

$$\sigma_p = \sqrt{\frac{(0.5)(0.5)}{100}} = 0.05.$$

Therefore,

$$z = \frac{0.62 - 0.50 - \frac{0.5}{100}}{0.05} = 2.3.$$

A z table can be used to determine that the two-tailedprobability value for a z of 2.3 is 0.0214.

5. The probability computed in Step 4 is compared to the significance level stated in Step 2. Since the probability value (0.0214) is less than the significance level of 0.05, the effect is statistically significant.

6. Since the effect is significant, the null hypothesis is rejected. It is concluded that the proportion of people choosing the original formulation is greater than 0.50.

7. This result might be described in a report as follows:

   The proportion of subjects choosing the original formulation (0.62) was significantly greater than 0.50, z = 2.3, p = 0.021. Apparently at least some people are able to distinguish between the original formulation and the new formulation.

Summary of Computations

1. Specify the null hypothesis and an alternative hypothesis.

2. Compute the proportion in the sample.

3. Compute

$$\sigma_p = \sqrt{\frac{\pi(1-\pi)}{N}}.$$

4. If p > π then compute

$$z = \frac{p - \pi - \frac{0.5}{N}}{\sigma_p}$$

otherwise, compute

$$z = \frac{p - \pi + \frac{0.5}{N}}{\sigma_p}$$

5. Use a z table to compute the probability value from z.

Assumptions

1. Observations are sampled randomly and independently.

2. The adequacy of the normal approximation depends on the sample size (N) and π. Although there are no hard and fast rules, the following is a guide to needed sample size: If π is between 0.4 and 0.6 then an N of 10 is adequate. If π is as low as 0.2 or as high as 0.8 then N should be at least 25. For π as low as 0.1 or as high as 0.9, N should be at least 30. A conservative rule of thumb is that both Nπ and N(1 - π) should be greater than 10.