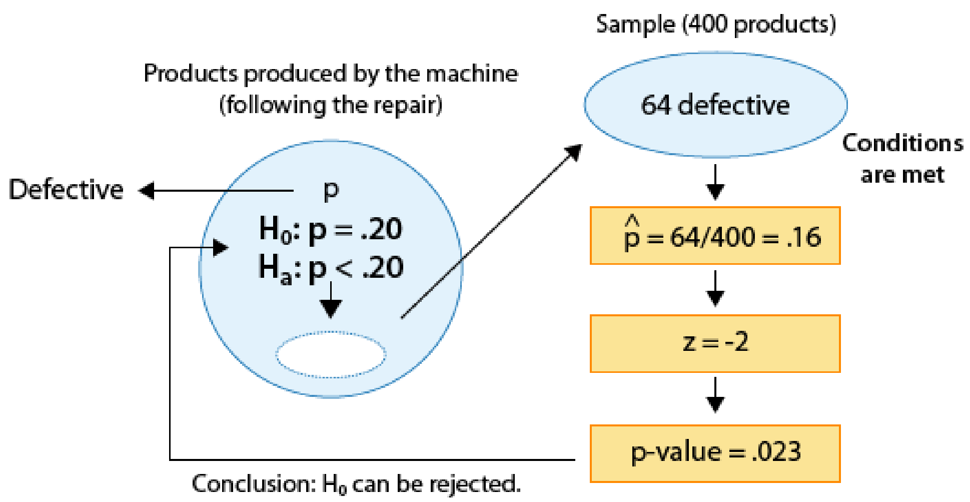
Statistics Package Exercise: Exploring the Effect of Sample Size on the Significance of Sample Results

Learning Objective: Apply the concepts of: sample size, statistical significance vs. practical importance, and the relationship between hypothesis testing and confidence intervals.

The purpose of this activity is to give you guided practice exploring the effect of sample size on the significance of sample results, and help you get a better sense of this effect. Another important goal of this activity is to help you understand the distinction between statistical significance and practical importance.

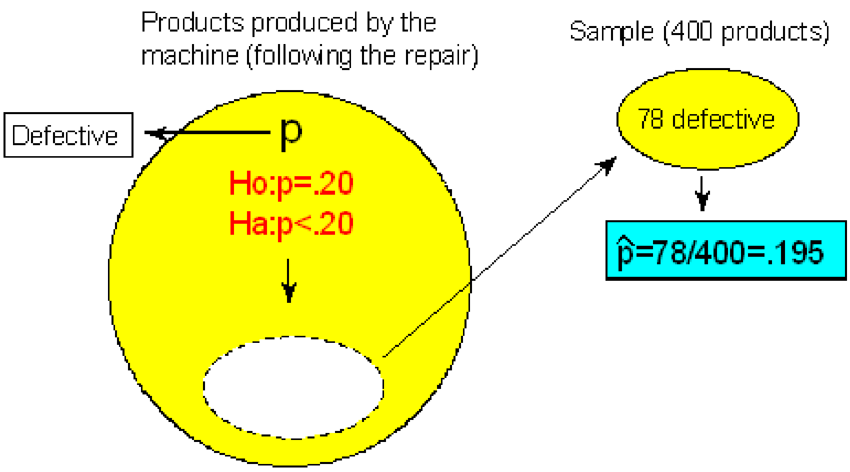
Background:

For this activity, we will use example 1. Here is a summary of what we have found:



The results of this study—64 defective products out of 400—were statistically significant in the sense that they provided enough evidence to conclude that the repair indeed reduced the proportion of defective products from 0.20 (the proportion prior to the repair). Even though the results—a sample proportion of defective products of 0.16—are statistically significant, it is not clear whether the results indicate that the repair was effective enough to meet the company's needs, or, in other words, whether these results have a practical importance. If the company expected the repair to eliminate defective products almost entirely, then even though statistically, the results indicate a significant reduction in the proportion of defective products, this reduction has very little practical importance, because the repair was not effective in achieving what it was supposed to. To make sure you understand this important distinction between statistical significance and practical importance, we will push this a bit further.

Consider the same example, but suppose that when the company examined the 400 randomly selected products, they found that 78 of them were defective (instead of 64 in the original problem):



- R
- StatCrunch
- TI Calculator
- Minitab
- Excel

R Instructions

From the background we know that there are

```
n=400, x=78
```

, and the null value is

```
p=0.20
```

. Here are the basic commands:

- ```
  p = prop.test(x=78,n=400,p=0.20,alternative="less",conf.level=0.95, correct=FALSE);p
  ```

To calculate

```
z
```

, enter the following command. (Since the sample proportion is less than

```
p = 0.2
```

, we know that z will be negative, so we take the negative square root.)

- ```
  z = -sqrt(p$statistic);z
  ```

The provided p-value is equivalent to the p-value we might find from the z-test we hand calculate for proportions.

Learn By Doing (1/1 point)

Based on the output, comment on the (statistical) significance of the results, and state your conclusions in context.

Your Answer:

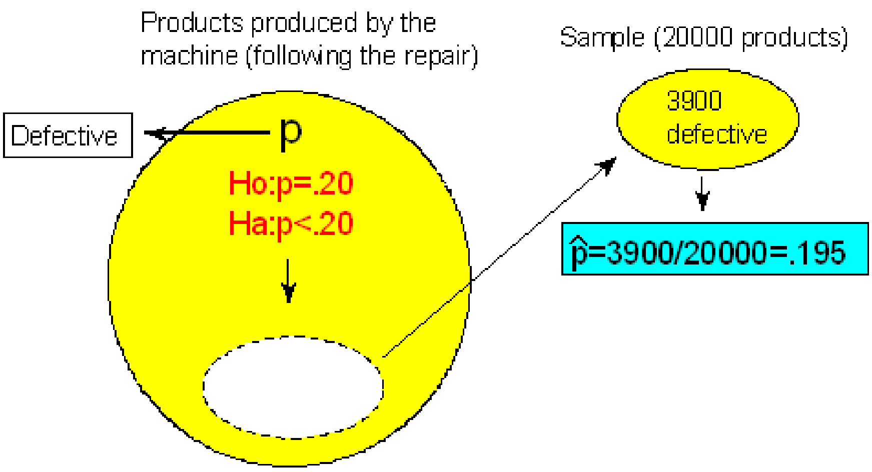
p-value is 0.40 and larger than 0.05

Our Answer:

RStatCrunch TI Calculator MinitabExcel R Here is the R output: Based on the large p-value (0.401) we conclude that the results are not statistically significant. In other words, the data do not provide evidence to conclude that the proportion of defective products has been reduced. StatCrunch Here is the StatCrunch output: Based on the large p-value (.401) we conclude that the results are not statistically significant. In other words, the data do not provide evidence to conclude that the proportion of defective products has been reduced. TI Calculator Here is the output: Based on the large p-value (.401) we conclude that the results are not statistically significant. In other words, the data do not provide evidence to conclude that the proportion of defective products has been reduced. Minitab Here is the Minitab output: Based on the large p-value (.401) we conclude that the results are not statistically significant. In other words, the data do not provide evidence to conclude that the proportion of defective products has been reduced. Excel Here are the results of our test: X: 78 n: 400 p-hat: 0.195 z: -0.25 p-value: 0.401 Based on the large p-value (.401) we conclude that the results are not statistically significant. In other words, the data do not provide evidence to conclude that the proportion of defective products has been reduced.

Resubmit Reset

Consider now another variation on the same problem. Assume now that over a period of a month following the repair, the company randomly selected 20,000 products, and found that 3,900 of them were defective.



Note that the sample proportion of defective products is the same as before, 0.195, which as we established before, does not indicate any practically important reduction in the proportion of defective products.

Learn By Doing (1/1 point)

Carry out the test. Based on the output, comment on the (statistical) significance of the results and state your conclusions in context.

Your Answer:

p-value is now 0.04 and so enough evidence to reject Ho.

Our Answer:

RStatCrunch TI CalculatorMinitabExcel R Here is the R output: Even though the sample results are similar to what we got before (sample proportion of 0.195), since they are based on a much larger sample (20,000 compared to 400) now they are statistically significant (at the .05 level, since 0.039 is less than 0.05). In this case, we can therefore reject Ho and conclude that the repair reduced the proportion of defective products to below 0.20. Summary: This is perhaps an "extreme" example, yet it is effective in illustrating the important distinction between practical importance and statistical significance. A reduction of 0.005 (or .5%) in the proportion of defective products probably does not carry any practical importance, however, because of the large sample size, this reduction is statistically significant. In general, with a sufficiently large sample size you can make any result that has very little practical importance statistically significant. This suggests that when interpreting the results of a test, you should always think not only about the statistical significance of the results but also about their practical importance. StatCrunch Here is the StatCrunch output: Even though the sample results are similar to what we got before (sample proportion of .195), since they are based on a much larger sample (20,000 compared to 400) now they are statistically significant (at the .05 level, since .039 is less than .05). In this case, we can therefore reject H0 and conclude that the repair reduced the proportion of defective products below .20. Summary: This is perhaps an “extreme” example, yet it is effective in illustrating the important distinction between practical importance and statistical significance. A reduction of .005 (or 0.5%) in the proportion of defective products probably does not carry any practical importance, however, because of the large sample size, this reduction is statistically significant. In general, with a sufficiently large sample size you can make any result that has very little practical importance statistically significant. This suggests that when interpreting the results of a test, you should always think not only about the statistical significance of the results but also about their practical importance. TI Calculator If you enter: and choose CALCULATE, then press ENTER, you should see: Even though the sample results are similar to what we got before (sample proportion of .195), since they are based on a much larger sample (20,000 compared to 400) now they are statistically significant (at the .05 level, since .039 is less than .05). In this case, we can therefore reject Ho and conclude that the repair reduced the proportion of defective products to below .20. Summary: This is perhaps an "extreme" example, yet it is effective in illustrating the important distinction between practical importance and statistical significance. A reduction of .005 (or .5%) in the proportion of defective products probably does not carry any practical importance, however, because of the large sample size, this reduction is statistically significant. In general, with a sufficiently large sample size you can make any result that has very little practical importance statistically significant. This suggests that when interpreting the results of a test, you should always think not only about the statistical significance of the results but also about their practical importance. Minitab Here is the Minitab output: Even though the sample results are similar to what we got before (sample proportion of .195), since they are based on a much larger sample (20,000 compared to 400) now they are statistically significant (at the .05 level, since .039 is less than .05). In this case, we can therefore reject Ho and conclude that the repair reduced the proportion of defective products to below .20. Summary: This is perhaps an "extreme" example, yet it is effective in illustrating the important distinction between practical importance and statistical significance. A reduction of .005 (or .5%) in the proportion of defective products probably does not carry any practical importance, however, because of the large sample size, this reduction is statistically significant. In general, with a sufficiently large sample size you can make any result that has very little practical importance statistically significant. This suggests that when interpreting the results of a test, you should always think not only about the statistical significance of the results but also about their practical importance. Excel Here are the results of our test: X: 3,900 n: 20,000 p-hat: 0.195 z: -1.77 p-value: 0.039 Even though the sample results are similar to what we got before (sample proportion of .195), since they are based on a much larger sample (20,000 compared to 400) now they are statistically significant (at the .05 level, since .039 is less than .05). In this case, we can therefore reject Ho and conclude that the repair reduced the proportion of defective products to below .20. Summary: This is perhaps an "extreme" example, yet it is effective in illustrating the important distinction between practical importance and statistical significance. A reduction of .005 (or .5%) in the proportion of defective products probably does not carry any practical importance, however, because of the large sample size, this reduction is statistically significant. In general, with a sufficiently large sample size you can make any result that has very little practical importance statistically significant. This suggests that when interpreting the results of a test, you should always think not only about the statistical significance of the results but also about their practical importance.

Resubmit Reset