Cancer Clusters

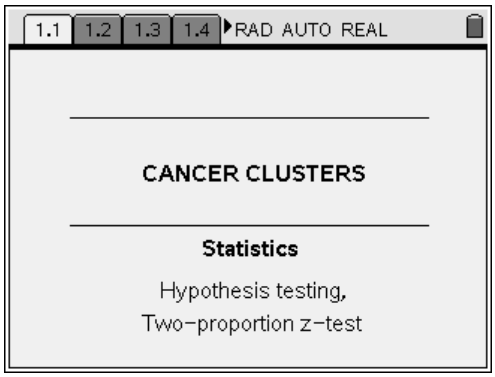
ID: 9996

Name _____

Class _____

In this activity, you will explore:

- *Hypothesis testing*
- *Two-proportion z-tests*



Open the file *StatAct06_CancerClusters_EN.tns* on your handheld and follow along with your teacher to work through the activity. Use this document as a reference and to record your answers.

Cancer Clusters

If individuals in an area seem to have an unusually high occurrence of cancer, many people believe that this might represent a "cancer cluster." A statistical technique called a two-proportion *z*-test can be used to decide whether the rate of cancer is unusually high in this location.

Consider the data on cases of cancer, based on a hypothetical random sample of 50,000 people from each of six states in New England. A large sample is used here because the rate of incidence of cancer is quite small.

State	Cancer Cases
Connecticut	268
Maine	291
Massachusetts	268
New Hampshire	251
Rhode Island	281
Vermont	249

Based on these data, does there seem to be an association between location and incidence of cancer?

Maine and New Hampshire are neighboring states. What percent of the Maine sample had cancer? What percent of the New Hampshire sample had cancer? Do you think that the difference between these two percentages is considerable?

Exercises

1. Connecticut and Rhode Island are also neighboring states. Compare their cancer data using a two-proportion *z*-test. Choose an appropriate alternate hypothesis. What is the *p*-value for this test? Explain the meaning of this number in the context of this example. Do you have enough information to conclude that one rate is larger than the other?

2. The sample data in this activity indicate that 268 out of 50,000 people in both the Connecticut and Massachusetts samples were diagnosed with cancer. Does this mean that the *actual* rates of incidence of cancer in the total populations of the two states were the same? Compare their cancer data using a two-proportion *z*-test. Choose an appropriate alternate hypothesis. What is the *p*-value for this test? Explain the meaning of this number in the context of this example.

3. Determine the relationship between the p-value obtained in a two-proportion *z*-test with the alternate hypothesis $p_1 < p_2$ or $p_1 > p_2$ versus the *p*-value for the alternate hypothesis $p_1 \neq p_2$.