

# Setting Up a Test of Difference in Population Proportions

Julie Deeke Statistics Course Developer





#### C.S. Mott Children's Hospital Poll

C.S. Mott Children's Hospital conducted a national poll on an issue in children's health, water safety. We will be looking at an example about swimming lessons.





#### Research Question

Is there a significant difference between the population proportions of parents of black children and parents of Hispanic children who report that their child has had some swimming lessons?

Populations - All parents of black children age 6-18 and all parents of Hispanic children age 6-18

Parameter of Interest - p<sub>1</sub> - p<sub>2</sub>

We'll let 1 = black and 2 = Hispanic



#### Research Question

Is there a significant difference between the population proportions of parents of black children and parents of Hispanic children who report that their child has had some swimming lessons?

Populations - All parents of black children age 6-18 and all parents of Hispanic children age 6-18

Parameter of Interest - p<sub>1</sub> - p<sub>2</sub>

Test for a significant difference in the population proportions of parents reporting that their child has had swimming lessons at the 10% significance level.



#### Hypotheses

$$H_0: p_1 - p_2 = 0$$

$$H_a: p_1 - p_2 \neq 0$$

$$\alpha = 0.10$$



## **Survey Results**

- A sample of 247 parents of black children age 6 -18 was taken with 91 saying that their child has had some swimming lessons.
- A sample of 308 parents of Hispanic children age 6
  -18 was taken with 120 saying that their child has had some swimming lessons.



## Assumptions

We need to assume that we have two independent random samples.

We also need <u>large enough sample sizes</u> to assume that the distribution of our estimate is normal. That is, we need  $n_1p$ ,  $n_1(1-p)$ ,  $n_2p$ , and  $n_2(1-p)$  to all be at least 10.



#### Assumptions

We need to assume that we have two independent random samples.

We also need <u>large enough</u> that the distribution of consists, we need  $n_1p$ ,  $n_1(1-p)$ , least 10.

That is, we need to estimate the common proportion, and then make sure that we would expect at least 10 yes's and 10 no's in each sample.



#### **Checking Assumptions**

$$\hat{p} = (91+120)/(247+308) = 211/555 = 0.38$$

$$247(0.38) = 94; 247(1-0.38) = 153;$$

$$308(0.38) = 117; 308(1-0.38) = 191$$



## **Checking Assumptions**

$$\hat{p} = (91+120)/(247+308) = 211/555 = 0.38$$

$$247(0.38) = 94; 247(1-0.38) = 153;$$

$$308(0.38) = 117; 308(1-0.38) = 191$$

If this assumption is not met, we can perform different tests that bypass this assumption.



#### **Best Estimate of the Parameter**

$$\hat{p_1} = 91/247 = 0.37$$

$$\hat{p}_2 = 120/308 = 0.39$$

$$\hat{p}_1 - \hat{p}_2 = 0.37 - 0.39 = -0.02$$