**Section 10.1 - Comparing Two Proportions** (pp. 609-629)

**Background** - Often times we want to compare proportions of individuals with a certain characteristic from two different populations. This might be the result of sampling two distinct populations or comparing the effectiveness of two distinct treatments in an experiment.

**The Sampling Distribution of a Difference between Two Proportions**

From Chapter 7, we saw that the sampling distribution of has the following properties:

* Shape:
* Center:
* Spread:

We can use the formulas for combining two independent random variables to describe the sampling distribution of :

* Mean:
* Standard Deviation:

|  |
| --- |
| **The Sampling Distribution of**  Choose an SRS of size *n1*  from Population 1 with proportion of successes and an independent SRS of size *n2* from Population 2 with proportion of successes *p2*.   * Shape: * Center:      * Spread: |

**Example** – Nathan and Kyle both work for the Department of Motor Vehicles (DMV), but in different states. In Nathan’s state, 80% of the registered cars are made by American manufacturers. In Kyle’s state, only 60% of the registered cars are made by American manufacturers. Nathan selects a random sample of 100 cars from his state and Kyle selects a random sample of 70 cars from his state. Let be the difference in the sample proportion of cars made by American manufacturers.

1. What is the shape of the sampling distribution of ? Why?
2. Find the mean of the sampling distribution. Show work.
3. Find the standard deviation of the sampling distribution. Show work.

**Confidence Intervals for *p1-p2***

|  |  |  |
| --- | --- | --- |
| **Estimate** | **Two-sample z interval for p1-p2 (2-PropZInt)**  An approximate level C confidence interval for *p1-p2* is  where z\* is the standard Normal critical value. | **Random**: The data are producd by a random sample of size n1 from population 1 and a random sample of size n2 from population 2 or by two groups of size n1 and n2 in a randomized experiment.  **Normal**: The counts of “successes” and “failures” in each sample or group -- n1p1, n1(1-p1), n2p2, n2(1-p2) – are at least 10.  **Independent**: Both the samples or groups themselves and the individual observations in each sample or group are independent. When sampling without replacement, check that the two populations are at least 10 times as large as the corresponding samples (the *10% condition*). |

**Example** - Many news organizations conduct polls asking adults in the US if they approve of the job the President is doing. How did President Obama’s approval rating change from August 2009 to September 2010? According to a CNN poll of 1024 randomly selected US adults on September 1-2, 2010, 50% approved of Mr. Obama’s performance. A CNN poll of 1010 randomly selected US adults on August 28-30, 2009, showed that 53% approved of Mr. Obama’s performance.

(1) Use the results of these polls to construct and interpret a 90% confidence interval for the change in Mr. Obama’s approval rating among all US adults.

(2) Based on your interval, is there convincing evidence that Mr. Obama’s job approval rating changed between August 2009 and September 2010?

HW: 5, 7, 9, 11

**Significance Tests for *p1-p2***

|  |  |  |
| --- | --- | --- |
| **Test** | **Two-sample z test for p1-p2 (2-PropZTest)**  **Significance tests** of *H0: p1 – p2 = 0* use the **pooled (combined) sample proportion**  The **two-sample z test for *p1-p2*** uses the test statistic    with P-values calculated from the standard Normal distribution. | **Random**: The data are producd by a random sample of size n1 from population 1 and a random sample of size n2 from population 2 or by two groups of size n1 and n2 in a randomized experiment.  **Normal**: The counts of “successes” and “failures” in each sample or group -- n1p1, n1(1-p1), n2p2, n2(1-p2) – are at least 10.  **Independent**: Observations and independent samples or groups; *10% condition* if sampling without replacement |

**Example** - Are teenagers going deaf? (“Say what?”) In a study of 3000 randomly selected teenagers in 1988-1994, 15% showed some hearing loss. In a similar study of 1800 teenagers in 2005-2006, 19.5% showed some hearing loss. (*Arizona Daily Star*, August 18, 2010)

Do these data give convincing evidence that the proportion of all teens with hearing loss has increased? Between the two studies, Apple introduced the iPod. If the results of the test are statistically significant, can we blame iPods for the increased hearing loss in teenagers?

**Inference for Experiments**

**Example** - In an effort to reduce health care costs, General Motors sponsored a study to help employees to stop smoking. In the study, half of the subjects were randomly assigned to receive up to $750 for quitting smoking for a year while the other half were simply encouraged to use traditional methods to stop smoking. None of the 878 volunteers knew that there was a financial incentive when they signed up. At the end of one year, 15% of those in the financial rewards group had quit smoking while only 5% in the traditional group had quit smoking. Do the results of this study give convincing evidence that a financial incentive helps people quit smoking? (*Arizona Daily Star*, February 11, 2009.) What was the *scope* of this experiment?

HW: 13, 15, 17, 21, 23, 25-28.