**Section 10.2 - Comparing Two Means** (pp. 634-653)

**Background** - Just as we considered comparing the proportions from two separate populations, we also may want to compare the *means* of two different populations. Again this might be looking at the results of random sampling or experimentation.

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

From Chapter 7, we saw that the sampling distribution of $\overbar{x}$ has the following properties:

* Shape:
* Center:
* Spread:

Again, we can use the formulas for combining two independent random variables to describe the distribution of $\overbar{x}\_{1}-\overbar{x}\_{2}$ :

* Mean:
* Standard Deviation:

|  |
| --- |
| **The Sampling Distribution of** $\overbar{x}\_{1}-\overbar{x}\_{2}$Choose an SRS of size *n1* from Population 1 with mean *μ1* and standard deviation *σ1* and an independent SRS of size *n2* from Population 2 with mean *μ2* and standard deviation *σ2*.* Shape:
* Center:
* Spread:
 |

**Example** - The Hyena Potato Chip Company buys potatoes from two different suppliers, Riderwood Farms and Camberley, Inc. The weights of the potatoes from Riderwood are approximately Normally distributed with a mean of 175 grams and a standard deviation of 25 grams. The weights of the potatoes from Camberley are approximately Normally distributed with a mean of 180 grams and a standard deviation of 30 grams. When the shipments arrive at the factory, inspectors randomly select a sample of 20 potatoes from each shipment and weigh them. They are surprised when the average weight of potatoes from Riderwood $\overbar{x}\_{r}$ is higher than the average weight of the potatoes from Camberley $\overbar{x}\_{c}$ .

a. Describe the shape, center and spread of the sampling distribution of $\overbar{x}\_{c}-\overbar{x}\_{r}$ .

b. Find the probability that the mean weight of the Riderwood sample is larger than the mean weight of the Camberley sample. Should the inspectors have been surprised?

HW: 25-28, 31, 33, 35, 51

**Confidence Intervals for** $\overbar{x}\_{1}-\overbar{x}\_{2}$

|  |  |  |
| --- | --- | --- |
| **Estimate** | **Two –sample t interval for μ1-μ2 (2-SampTInt)**$$\left(\overbar{x}\_{1}-\overbar{x}\_{2}\right)\pm t^{\*}\sqrt{\frac{s\_{1}^{2}}{n\_{1}}+\frac{s\_{2}^{2}}{n\_{1}}}$$df = min(n1 - 1, n2 - 1) | **Random**: Data from random samples or randomized experiment**Normal**: Population distributions Normal or large samples (n1≥30, n2≥30)**Independent**: Observations and independent samples or groups; *10% condition* if sampling without replacement |

**Example** - Do plastic bags from Target or plastic bags from Walmart hold more weight? A group of AP Statistics students decided to investigate by filling a random sample of 5 bags from each store with common grocery items until the bags ripped. Then they weighed the contents of items in each bag to determine its capacity. Here are the results in grams:

**Target** 12572 13999 11215 15447 10896

**Walmart** 9552 10896 6983 8767 9972

a. Construct and interpret a 99% confidence interval for the difference in the mean capacity of plastic grocery bags from Target and Walmart.

b. Does the interval provide convincing evidence that there is a difference in the mean capacity between the stores?

**Technology**

**Significance Tests for** $\overbar{x}\_{1}-\overbar{x}\_{2}$

|  |  |  |
| --- | --- | --- |
| **Test** | **Two-sample t test for μ1-μ2 (2-SampTTest)**$$t=\frac{(\overbar{x}\_{1}-\overbar{x}\_{2})-(μ\_{1}-μ\_{2})}{\sqrt{\frac{s\_{1}^{2}}{n\_{1}}+\frac{s\_{2}^{2}}{n\_{2}}}}$$df = min(n1 - 1, n2 - 1) | **Random**: Data from random samples or randomized experiment**Normal**: Population distributions Normal or large samples (n1≥30, n2≥30)**Independent**: Observations and independent samples or groups; *10% condition* if sampling without replacement |

**Example** - In commercials for Bounty paper towels, the manufacturer claims that they are the “quicker picker-upper.” But are they also the stronger picker-upper? Two AP Statistics students selected a random sample of 30 Bounty paper towels and 30 generic paper towels and measured their strength when wet. To do this, they uniformly soaked each paper towel with 4 ounces of water, held two opposite edges of the paper towel, and counted how many quarters each paper towel could hold until ripping, alternating brands. Here are the results:

**Bounty:**
106 111 106 120 103 112 115 125 116 120 126 125 116 117 114 118 126 120 126 125 116 117 114 118 126 120 115 116 121 113 111 128 124 125 127 123 115 114
**Generic:**
77 103 89 79 88 86 100 90 81 84 84 96 87 79 90 86 88 81 91 94 90 89 85 83 89 84 90 100 94 87

a. Display these distributions using parallel boxplots and briefly compare the distributions. Based only on the boxplots, discuss whether or not you think that the mean of Bounty is significantly higher than the mean of the generic.

b. Use a significance test to determine whether there is convincing evidence that wet Bounty paper towels can hold more weight, on average, than wet generic paper towels can.

c. Interpret the *P*-value from part b in the context of the question.

**Technology**

**Using the Two-Sample *t* Procedures**

**1. The Normal Condition**

* *Sample size less than 15*:
* *Sample size at least 15*:
* *Large samples*:

**2. The Pooled Two-Sample *t* Procedures**

**3. Inference for Experiments**

**4. Using Two-Sample *t* Procedures versus Using a Paired *t*-Test**

**Application** - Suppose you are designing an experiment to determine if students perform better on tests when there are no distractions, such as a teacher talking on a phone. You have access to two classrooms and 30 volunteers who are willing to participate in your experiment.

a. Design an experiment so that a two-sample *t* test would be the appropriate inference method.

b. Design an experiment so that a paired *t* test would be the appropriate inference method.

c. Which experiment is better? Why?

HW: 41, 43, 45, 53, 57-60, 65