

Chapter Goals

To introduce you to the process of Statistical Decision Making.

You will

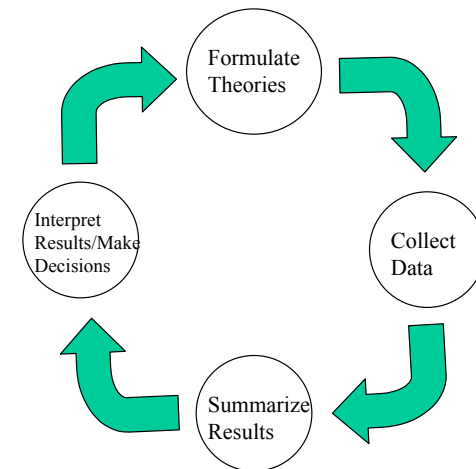
- learn to think critically about information that is presented
- learn to examine assumptions
- learn to evaluate evidence

What is Statistics?

Statistics is the science of data

The Scientific Method

1. Formulate a theory
2. Collect data to test the theory
3. Analyze the results
4. Interpret the results, and make decisions



Example:

Exercise *Does the data always conclusively prove or disprove the theory?*

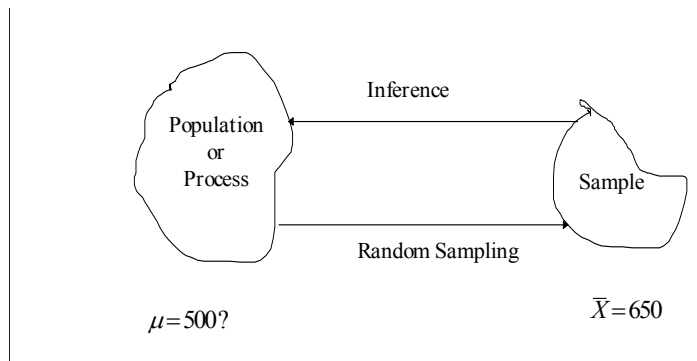
Scientific Method (continued)

The scientific method is an iterative process. In general, we **reject a theory** if the data were *unlikely* to occur if the theory were in fact *true*.

Statistical Decision Making

Statistical Inference

To use sample data to make generalizations about a larger data set (population)



Statistical Inference: What is it?

Definition A *population* is the total set of elements of interest for a given problem.

Definition A *sample* is a part of the population under study so that inferences can be drawn from it about the population.

Definition *Statistical inference* is the process of drawing conclusions about the population based on information from a sample

Testing Theories

Hypotheses Competing theories that we want to test about a population are called *Hypotheses* in statistics. Specifically, we label these competing theories as *Null Hypothesis* (H_0) and *Alternative Hypothesis* (H_1 or H_a).

H_0 : The null hypothesis is the status quo or the prevailing viewpoint.

H_1 : The alternative hypothesis is the competing belief. It is the statement that the researcher is hoping to prove.

Example: Taking an aspirin every other day for 20 years can cut your risk of colon cancer nearly in half, a study suggests. According to the American Cancer Society, the lifetime risk of developing colon cancer is 1 in 16.

H_0 :

H_1 :

Let's do it! 1.2

(*New York Times*, 1/21/1997) Winter can give you a cold because it forces you indoors with coughers, sneezers, and wheezers.

Toddlers can give you a cold because they are the original Germs “R” Us. But, can going postal with the boss or fretting about marriage give a person a post-nasal drip?

Yes, say a growing number of researchers. A psychology professor at Carnegie Mellon University, Dr. Sheldon Cohen, said his most recent studies suggest that stress doubles a person's risk of getting a cold.

The percentage of people exposed to a cold virus who actually get a cold is 40%. The researcher would like to assess if stress increases this percentage. So, the population of interest is people who are under stress. State the appropriate hypothesis for assessing the researcher's theory regarding the population.

H_0 :

H_1 :

Deciding Which Theory to Support

Decision making is based on the “rare event” concept. Since the null hypothesis is the status quo, we assume that it is true unless the observed result is extremely unlikely (rare) under the null hypothesis.

Definition *If the data were indeed unlikely to be observed under the assumption that H_0 is true, and therefore we reject H_0 in favor of H_1 , then we say that the data are **statistically significant**.*

Examples:

Let's do it! 1.3

Last month a large supermarket chain received many customer complaints about the quantity of chips in a 16-ounce bag of a particular brand of potato chips. Wanting to assure its customers that they were getting their money's worth, the chain decided to test the following hypothesis concerning the true average weight (in ounces) of a bag of such potato chips in the next shipment received from the supplier:

H_0 :

H_1 :

Suppose you concluded H_1 . Could you be wrong in your decision? What if you did not reject H_0 ? Could you be wrong in your decision?

Errors in Decision Making

In our current justice system, the defendant is presumed innocent until proven guilty. The null and alternative hypothesis that represents this is:

H_0 :

H_1 :

		The Truth	
		The defendant is:	
		Innocent	Guilty
Your decision based on the data	Innocent		
	Guilty		

Definition *Rejecting* the null hypothesis H_0 when in fact it *is true* is called a **Type I** error. *Accepting* the null hypothesis H_0 when in fact it *is not true* is called a **Type II** error.

Note: Rejecting the null hypothesis is usually considered the more serious error than accepting it.

		The Truth	
		H_0 is true	H_1 is true
Your decision based on the data	H_0		
	H_1		

- α = Type I error
- = The chance of rejecting H_0 when in fact H_0 is true
- = $P(H_1|H_0)$
- β = Type II error
- = The chance of accepting H_0 when in fact H_1 is true
- = $P(H_0|H_1)$

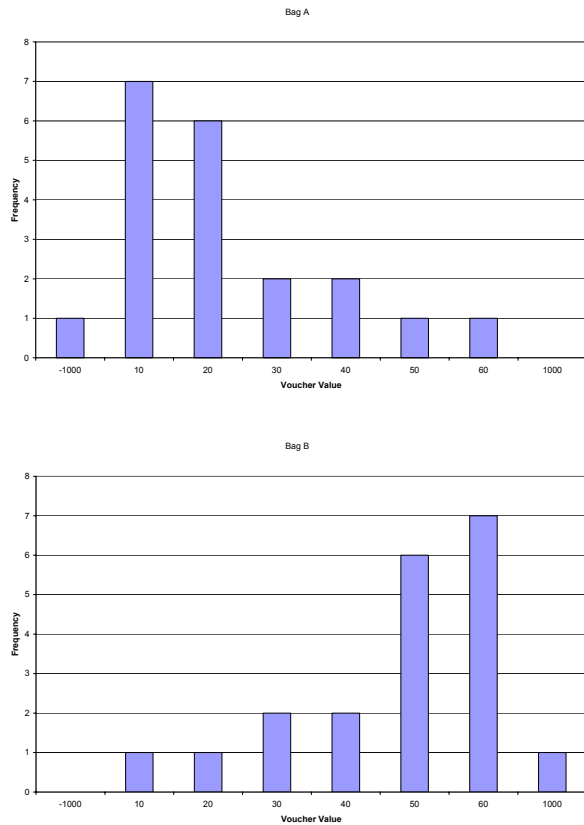
What's in the Bag?

Objective To explore the various aspects of decision making

Problem statement There are two identical looking bags, Bag A and Bag B. Each bag contains 20 vouchers. The contents of the bag, i.e., the face value and the frequency of voucher values, are as follows:

Face Value (\$)	Bag A	Bag B
	Frequency	Frequency
-1000	1	0
10	7	1
20	6	1
30	2	2
40	2	2
50	1	6
60	1	7
1000	0	1
Total	20	20

Frequency Plots



If you could see both bags, which one would you choose?

Game Rules

- The objective is to pick Bag B.
- You will be shown only **one** of the bags.
- You will be allowed to gather some data from the bag, and based on that information, you must decide whether to take the shown bag (because you think that it is Bag B), or the other bag (because you think that the shown bag is Bag A).
- Initially, the data will consist of selecting just one voucher from the shown bag (without looking into it). In this case, we say that we are taking a sample of size $n = 1$.

Based on the above information, we will be testing the following hypothesis:

H_0 : The shown bag is Bag A

H_1 : The shown bag is Bag B

Type I error $\alpha =$

Type II error $\beta =$

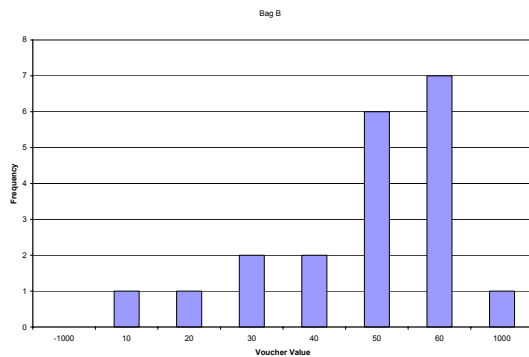
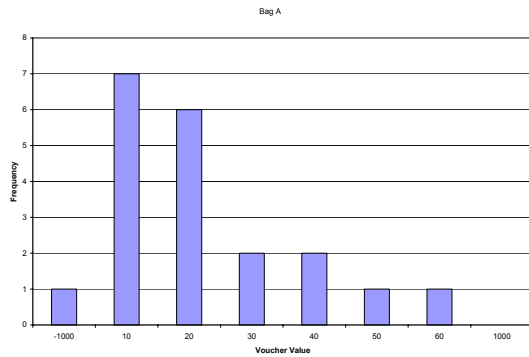
Exercise *If the voucher you selected was \$60, what would you decide? What if the voucher was \$10 instead*

Forming a Decision Rule

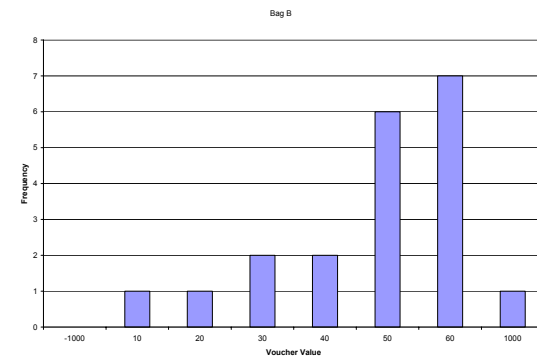
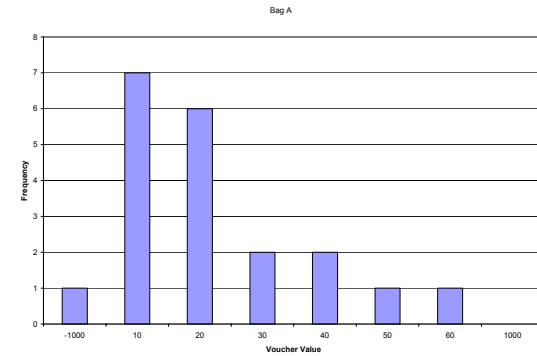
Face Value (\$)	Chance if	Chance if
	Bag A	Bag B
-1000	1/20	0/20
10	7/20	1/20
20	6/20	1/20
30	2/20	2/20
40	2/20	2/20
50	1/20	6/20
60	1/20	7/20
1000	0/20	1/20

What values of the voucher (or in what direction of voucher values) support the alternative hypothesis H_1 ? That is, what is the direction of extreme?

Decision Rule 1 Reject the null hypothesis H_0 in favor of the alternative hypothesis H_1 if the voucher value is $\geq \$50$. Otherwise, do not reject the null hypothesis H_0 .



Type I error $\alpha =$
 Type II error $\beta =$



Type I error $\alpha =$
 Type II error $\beta =$

Summary

1. **Decision Rule** Reject H_0 if voucher $\geq \$50$

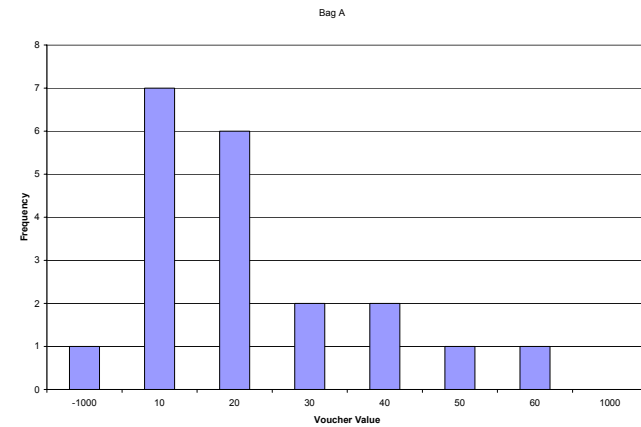
Rejection Region \$50 or more

We say ... the cutoff is \$50, and larger values are more extreme

P-Values

Suppose we select a voucher. Assuming that H_0 is true, how likely is it that we would get the observed voucher value, or something more extreme?

Example:



Question: What kind of p-values values support H_1 ?

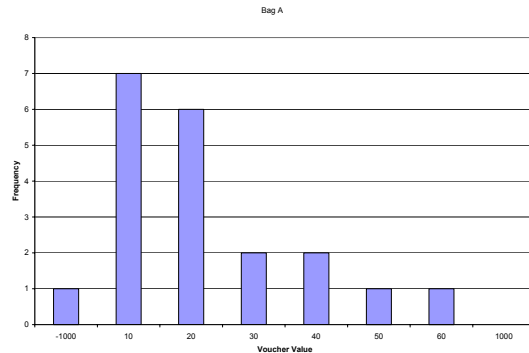
Decision Making and P-Values

Consider the hypothesis that we considered earlier for the two-bag experiment:

H_0 : The shown bag is Bag A

H_1 : The shown bag is Bag B

Using $\alpha = 0.10$, what is the decision rule?



Now, if you draw a voucher with a face value of \$30, which hypothesis would you conclude? For this voucher value, can you calculate the *p-value*?

Relationship between α and *p-values*

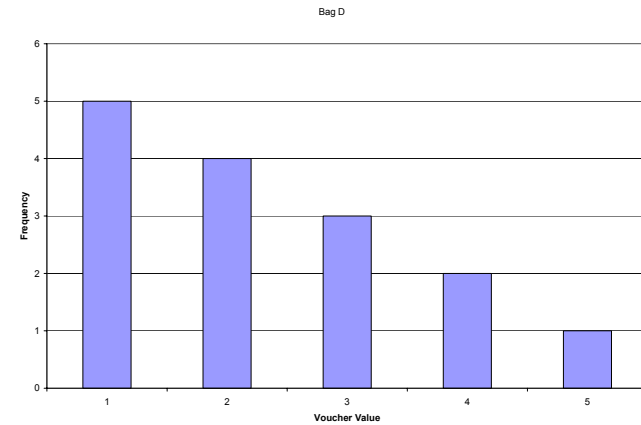
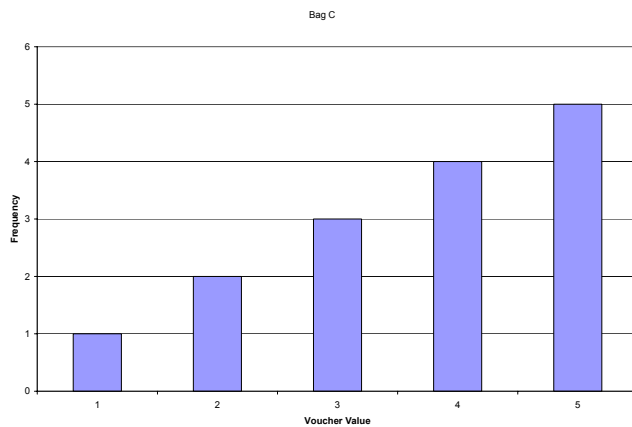
If $p\text{-value} > \alpha$, Do Not Reject the null hypothesis H_0 , otherwise

if $p\text{-value} \leq \alpha$, Reject the null hypothesis H_0 in favor of the alternative hypothesis H_1 .

P-Values (Continued)

Consider two identical bags C and D with the following distribution of voucher values:

Face Value (\$)	Bag C (H_0)		Bag D (H_1)	
	Frequency	Chance	Frequency	Chance
1	1	1/15	5	5/15
2	2	2/15	4	4/15
3	3	3/15	3	3/15
4	4	4/15	2	2/15
5	5	5/15	1	1/15



Let's do it! 1.6

Now, consider the following hypothesis:

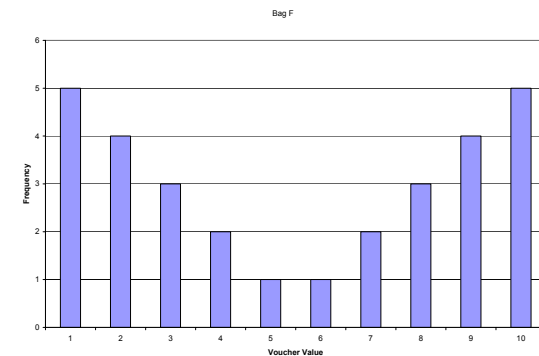
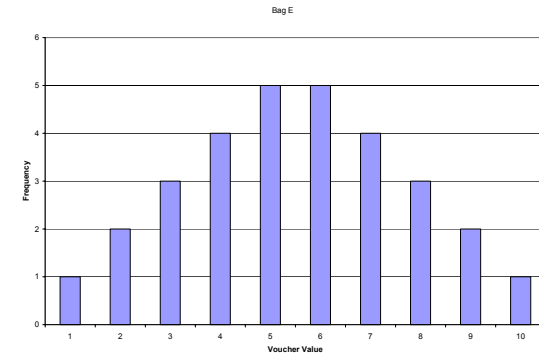
H_0 : The shown bag is Bag C

H_1 : The shown bag is Bag D

1. Suppose that the observed voucher (sample size of $n = 1$) is \$2. What is the p -value?
2. Would you accept or reject the null hypothesis for the following levels of $\alpha = 0.10, 0.05, 0.01$.

P-Values (Continued)

Consider two identical bags E and F with the following distribution of voucher values:



Let's do it! 1.7

Consider the hypothesis

H_0 : The shown bag is Bag E

H_1 : The shown bag is Bag F

1. If the decision rule is Reject H_0 if the selected voucher value is ≤ 1 or ≥ 10 , then what are α and β ?
2. Suppose the observed voucher value is \$2. What is the p -value?
3. Would you accept or reject the null hypothesis for the following levels of $\alpha = 0.10, 0.05, 0.01$.

Let's do it! 1.8

The following table summarizes the results of three studies:

Study A

H_0 : The true average lifetime ≥ 54

H_1 : The true average lifetime < 54

P -value = 0.0251

Study B

H_0 : The average time to relief for Treatment I is equal to the average time to relief for Treatment II

H_1 : The average time to relief for Treatment I is not equal to the average time to relief for Treatment II

P -value = 0.0018

Study C

H_0 : The true proportion of adults who work 2 jobs is ≤ 0.33

H_1 : The true proportion of adults who work 2 jobs is > 0.33

P -value = 0.3590

1. For which study do the results show the most support for the null hypothesis?

2. Suppose Study A concluded that the data supported the alternative hypothesis that the true average lifetime is less than 54 months, but in fact the true average lifetime is greater than or equal to 54 months. Is this a Type I (α) or Type II (β) error?

3. For each of the three above studies, determine if the rejection region would be on the one-sided left tailed, one-sided right tailed, or two-sided.
 - Study A
 - Study B
 - Study C

Significant versus Important

- With a large enough sample size, even a small difference can be found statistically significant – that is, the difference is hard to explain by chance alone. This does not necessarily make the difference important.
- On the other hand, an important difference may not be statistically significant if the sample size is too small.