AP Statistics
Mr Murphy
10.1 Hypothesis Tests and Errors in Hypothesis Testing

HW Pg. 581 #10.2, 10.4, 10.11, 10.13, 10.14, 10.16, 10.18

Goals: 1. Determine the null and alternate hypothesis in context.
2. Determine Type I and Type II errors in context.
3. Recommendation of appropriate levels of significance.
4. Be kind.

- A **test hypotheses or test procedure** is a method for using sample data to decide between two competing claims (hypothesis) about a population characteristic.

- The **null hypothesis**, denoted by H_0, is a claim about a population characteristic that is initially *assumed to be true*.

- The **alternate hypothesis**, denoted by H_a, is the competing claim.

- The two possible conclusions are then,
  - *reject* H_0 (conclude H_a is true) or
  - *fail to reject* H_0 (conclude H_0 is true).

- The form of a null hypothesis is

H_0: population characteristic = hypothesized value

The alternate hypothesis has one of the following three forms:

H_a: population characteristic > hypothesized value

H_a: population characteristic < hypothesized value

H_a: population characteristic $\neq$ hypothesized value

- Hypotheses are **always** based on **parameters**, NEVER statistics.

Ex1 Costco brand 13 – W light bulbs state on the package "Avg. Life 1000 Hr." Let μ denote the true mean life of Costco 13 – W light bulbs. People who purchased this brand would be unhappy if μ is actually less than the advertised value. How would we write our null and alternative hypotheses?

Ex2 Because of the variation in the manufacturing process, tennis balls produced by a particular machine do not have identical diameters. Let μ denote the true average diameter for tennis balls currently being produced. Suppose that the machine was initially calibrated to achieve the design specification μ = 3in. However, the manufacturer is now concerned that the diameters no longer conform to this specification. Determine the null and alternative hypotheses.

Ex3 A certain university has decided to introduce the use of plus and minus with letter grades, as long as there is evidence that more than 60% of the faculty favor the change. A random sample of faculty will be selected, and the resulting data will be used to test the relevant hypotheses. If *p* represents the true proportion of all faculty that favor a change to plus-minus grading, which of the following pair of hypotheses should the administration test:

$$H_0: p = 0.6 \qquad H_a: p < 0.6$$

or

$$H_0: p = 0.6 \qquad H_a: p > 0.6$$

Explain your choice.

Ex4 The manufacturer of a refrigerator system for beer kegs produces refrigerators that are supposed to maintain a true mean temperature, μ, of $40^\circ F$, ideal for a certain type of German pilsner. The owner of the brewery does not agree with the refrigerator manufacturer, and claims he can prove that the true mean temperature is incorrect. What are the appropriate null and alternative hypothesis?

(a) $H_0: \ \mu \neq 40^\circ$; $H_a: \ \mu = 40^\circ$

(b) $H_0: \ \mu \geq 40^\circ$; $H_a: \ \mu < 40^\circ$

(c) $H_0: \ \mu \leq 40^\circ$; $H_a: \ \mu > 40^\circ$

(d) $H_0: \ \mu = 40^\circ$; $H_a: \ \mu \neq 40^\circ$

(e) $H_0: \ \mu = 40^\circ$; $H_a: \ \mu > 40^\circ$

Type I vs. Type II Errors

Errors can be made when testing hypotheses. In any hypothesis test we have 4 possibilities: two are correct and two are errors.

	Fail to reject H_0	**Reject H_0**
H_0 **true**	*Hooray!*	**Type I error**
H_a **true**	**Type II error**	*Hooray!*

- **Type I Error** : the error of rejecting H_0 when H_0 is true
- **Type II Error**: the error of failing to reject H_0 when H_0 is false

Ex5 The U.S. Department of Transportation reported that during a recent period, 77% of all domestic passenger flights arrived on time (meaning within 15 minutes of the scheduled arrival). Suppose that an airline with a poor on-time record decides to offer its employees a bonus if, in an upcoming month, the airline's proportion of on-time flights exceeds the overall industry rate of 0.77. Let *p* be the true proportion of the airline's flights that are on time during the month of interest.

What are the null and alternate hypotheses?

What are the Type I and Type II errors in this context?

***AP Note**: Always write errors in terms of H_a*

What are the consequences of these errors?

- The probability of a Type I error is denoted by α (alpha) and is called the **level of significance** of the test. Thus, a test with $\alpha = 0.01$ is said to have a level of significance of 0.01 or to be a level 0.01 test.

- The probability of a Type II error is denoted by β (beta).

- An ideal test procedure would result in both $\alpha = 0$ and $\beta = 0$ -- but in order for this to happen we would need to conduct a census.

- After assessing the consequences of Type I and Type II errors, identify which error to control - *using a smaller α increases β, and vice versa.*

Ex6 Suppose that you are an inspector for the Fish and Game Department and that you are given the task of determining whether to prohibit fishing along part of the Oregon coast. You will close an area to fishing if it is determined that fish in that region have unacceptably high mercury content.

(a) Assuming that a mercury concentration of 5 ppm is considered the maximum safe concentration, which of the following pairs of hypotheses would you test:

$$H_0 : \mu = 5 \qquad\qquad H_0 : \mu = 5$$
$$H_a : \mu > 5 \qquad \text{or} \qquad H_a : \mu < 5$$

(b) Would you prefer a significance level of 0.1 or 0.01 for your test? Explain.

Checkpoint

1. What type of error occurs if you reject H_0, when, in fact, it is true?

(a) Type 1 error
(b) Type 2 error
(c) Type 3 error
(d) either a Type 1 or Type 2 error, depending on the level of significance.
(e) either a Type 1 or Type 2 error, depending on whether the test is one tail or two tail.

2. A psychologist claims that more than 6.1 percent of the population suffers from professional problems due to extreme shyness. Determine the null and alternate hypotheses.

(a) $H_0: \quad p < 6.1\%$
$H_a: \quad p \geq 6.1\%$

(b) $H_0: \quad p = 6.1\%$
$H_a: \quad p < 6.1\%$

(c) $H_0: \quad p > 6.1\%$
$H_a: \quad p \leq 6.1\%$

(d) $H_0: \quad p = 6.1\%$
$H_a: \quad p > 6.1\%$

(e) $H_0: \quad p = 6.1\%$
$H_a: \quad p \neq 6.1\%$

3. In hypothesis testing,

(a) the less the likelihood of a Type I error, the less the likelihood of Type II error
(b) the less the likelihood of a Type I error, the more the likelihood of Type II error
(c) the likelihood Type II errors will not be affected by the likelihood Type I errors
(d) the sum of the probabilities of Type I and Type II errors must equal 1
(e) the probability of committing a Type I error is β

4. In the past, the mean running time for a certain type of flashlight battery has been 9.6 hours. The manufacturer has introduced a change in the production method and wants to perform a significance test to determine whether the mean running time has increased as a result. The hypotheses are:

$$H_0: \quad \mu = 9.6 \text{ hours}$$
$$H_a: \quad \mu > 9.6 \text{ hours}$$

If the hypothesis test concludes the production change increases battery life, the manufacturer will change production methods and pass the increased cost onto the consumer. From the standpoint of the consumer, what α and β levels should be chosen?

(a) $\alpha = 0.05,\ \beta = 0.05$
(b) $\alpha = 0.07,\ \beta = 0.05$
(c) $\alpha = 0.10,\ \beta = 0.01$
(d) $\alpha = 0.05,\ \beta = 0.07$
(e) $\alpha = 0.01,\ \beta = 0.10$

5. Rejecting a true alternate hypothesis

(a) is a Type I error.
(b) has the probability of $1-\beta$ of occurring.
(c) has the probability of α of occurring.
(d) is a Type II error.
(e) is a correct decision.