# To Do

**Read Section 5.1-5.3**

**Do End-of-Chapter Problems 1-15.**

**Assignment 3 due Friday Nov. 11.**

# Today's Lecture

(1) Introduction to Tests of Hypothesis
(2) Null hypothesis, Alternative Hypothesis
(3) Test statistic or Discrepancy Measure
(4) Steps of a Test of Hypothesis
(5) Interpretation of a p-value

# Tests of Hypotheses – Chapter 5

**Statistical tests of hypotheses are conducted in a very similar manner to a North American criminal trial.**

**We begin by specifying a single "default" hypothesis ("the defendant is innocent").**

**This hypothesis is called the null hypothesis, and is denoted by $H_0$.**

# Null Hypothesis

**The name “null hypothesis” originates from the fact that in experiments which are conducted to determine if a new treatment has an effect the default hypothesis is usually that the treatment has no effect, that is, the treatment has a “null” effect.**

# Alternative Hypothesis

**We collect some data.**

**On the basis of the observed data we decide how plausible the null hypothesis is.**

**Of course, there is also an alternative hypothesis, denoted by, $H_A$.**

**In many cases $H_A$ is simply that $H_0$ is not true ("the defendant is guilty").**

# Experiment to Test ESP Claim

**To prove or disprove the claim that Student has ESP we conduct an experiment.**

**Let $Y$ = the number of correct outcomes in 25 trials.**

**We assume $Y$ has a Binomial distribution with $n = 25$. (Is this a valid assumption?)**

**Let $\theta$ be the probability of a correct outcome.**

**If Student has no ESP and they are just guessing, then $\theta = 0.5$ whereas if Student has ESP then $\theta \neq 0.5$. (What does $\theta < 0.5$ mean?)**

## Questions

**(1) What should the null hypothesis be?**

**$H_0$: $\theta = 0.5$ (Student does not have ESP, they are just guessing) or $H_0$: $\theta \neq 0.5$ (Student does have ESP)?**

**(2) The possible observed values of $Y$ are 0,1,...,25. What observed values of $Y$ are consistent with $H_0$?**

**(3) What observed values of $Y$ provide evidence against $H_0$?**

**(4) How do we measure the strength of the evidence against $H_0$?**

# Answers

**(1) We choose $H_0$: $\theta = 0.5$ (Student does not have ESP) to be the null hypothesis.**

**(2) Observed values of $Y$ close to $E(Y) = 12.5$ are consistent with $H_0$.**

**(3) Values of $Y$ close to 0 or 25 provide evidence against $H_0$.**

# More Questions

**(4) Suppose Student identifies the correct outcome 15 times.**

**Is this "close enough" to 12.5 to decide they do not have ESP?**

**That is, if the observed value of $Y$ is 15 is there evidence against $H_0$?**

**What if we observe $Y$ = 18 or $Y$ = 20?**

# Test Statistics and Discrepancy Measures

**To answer the question of "how close is close" we use a test statistic or discrepancy measure.**

**Definition:**

**A test statistic or discrepancy measure is a function of the data $D = g(Y)$ that is constructed to measure the degree of "agreement" between the data $Y$ and the null hypothesis $H_0$.**

# Test Statistics and Discrepancy Measures

Usually we define $D$ so that $D = 0$ represents the best possible agreement between the data and $H_0$.

"Large" values of $D$ indicate poor agreement between the data and $H_0$.

# Discrepancy Measure for ESP Example

**In the ESP example, it seems reasonable to use the test statistic or discrepancy measure**

$$D(Y) = |Y - 12.5|$$

**since E($Y$) = 12.5 if Student is guessing.**

**The possible values for $D(Y)$ are 0.5, 1.5, ..., 12.5.**

# Discrepancy Measure for ESP Example

**Small observed values of *D* (close to 0.5) would lead us to believe that Student does not have ESP, that is, there is no evidence against $H_0$: $\theta = 0.5$.**

**Large observed values of *D* (close to 12.5) would lead us to believe that Student does have ESP, that is, there is evidence against $H_0$: $\theta = 0.5$.**

# Discrepancy Measure for ESP Example

**We still have the problem of deciding how large the observed of *D* should be before we would conclude significant evidence against the null hypothesis $H_0$: $\theta = 0.5$.**

**Suppose Student correctly identified 15 outcomes.**

**The observed value of $D = |Y - 12.5|$ is**
**$d = |15 - 12.5| = 2.5$.**

**What is the probability of observing a value of *D* greater than or equal to $d = 2.5$ if $H_0$: $\theta = 0.5$ is true?**

# Probability that *D* is greater or equal to *d*

**If $H_0$: $\theta = 0.5$ is true then**

**$P(D \geq 2.5; H_0) = P(|Y - 12.5| \geq 2.5)$**

**where $Y \sim \text{Binomial}(25, 0.5)$**

**$= P(Y \leq 10) + P(Y \geq 15)$**

**$= 1 - P(11 \leq Y \leq 14)$**

$$= 1 - \sum_{y=11}^{14} \binom{25}{y} (0.5)^{25}$$

**$= 0.4244$**

**How do we interpret this number?**

# How do we interpret this probability?

**Suppose we tested a large number of students to see if they have ESP using the same experiment.**

**Suppose also that these students did not have any special ESP ability and they were only guessing.**

**Then approximately 42% of these students would observe a value of $D = |Y - 12.5|$ greater than or equal to Student's value of $d = |15 - 12.5| = 2.5$.**

# How do we interpret this probability?

**This does not prove Student does not have ESP but it does indicate that there is little evidence based on the observed data to support the hypothesis that Student does have ESP.**

**The probability P($D \geq 2.5$) = 0.4244, in this example, is called the *p*-value of the test of hypothesis.**

# Definition of a *p*-value

**Suppose we use the test statistic $D = D(Y)$ to test the hypothesis $H_0$. Suppose also that $d = D(y)$ is the observed value of $D$.**

**Definition:**

**The *p*-value of the test of hypothesis $H_0$ using test statistic $D$ is *p*-value = $P(D \geq d;\ H_0)$.**

**In other words, the *p*-value is the probability of observing a value of the test statistic greater than or equal to the observed value of the test statistic assuming $H_0$ is true.**

# Steps of a Statistical Test of Hypothesis

**(1) Assume that the null hypothesis $H_0$ will be tested using data $Y$.**

**(2) Adopt a test statistic or discrepancy measure $D(Y)$ for which, large values of $D$ are less consistent with $H_0$. Let $d = D(y)$ be the corresponding observed value of $D$.**

**(3) Calculate**

$$p\text{-value} = \mathrm{P}(D \geq d \text{ assuming } H_0 \text{ is true})$$
$$= \mathrm{P}(D \geq d;\, H_0)$$

**(4) Draw a conclusion based on the p-value. (This is the tricky part!)**

# Test of Hypothesis for ESP Experiment

**(1) In the ESP experiment $Y$ = number of correct answers out of 25, $Y \backsim$ Binomial(25,θ) and the null hypothesis was $H_0$: θ = 0.5 (Student is guessing.)**

**(2) We used the test statistic $D(Y) = |Y - 12.5|$ and the observed value was $d = |15 - 12.5| = 2.5$.**

**(3) $p$-value = P ($D \geq 2.5$ ; assuming $H_0$ is true)**

$$= P(D \geq 2.5\ ;\ H_0)$$

$$= P(|Y - 12.5| \geq 2.5)$$

**where $Y \backsim$ Binomial(25,0.5)**

$$= 0.4244$$

# Interpreting the p-value

**(4) If we did this experiment over and over again then about 42% of the time we would observe a result as or more unusual than *d* = 2.5 if the person is just guessing.**

**There is no evidence based on the data to suggest that the hypothesis "Student is guessing" is false.**

# Interpreting the *p*-value

**If $d$, the observed value of $D$ is large, and consequently the *p*-value = $P(D \geq d; H_0)$ is small then one of the following two statements is true:**

**(1) $H_0$ is true but by chance we have observed an outcome that does not happen very often when $H_0$ is true**

**or**

**2) $H_0$ is false.**

**The problem is we don't know which statement is true.**

# Interpreting the *p*-value

**Suppose *p*-value = 0.001, then the event "observe a *D* value as unusual or more unusual as we have observed" happens only about 1 time out of 1000, that is, not very often.**

**We interpret small *p*-values (< 0.01) as indicating that the observed data are providing strong evidence against the null hypothesis $H_0$.**

# Interpreting the *p*-value

**If the *p*-value is large ( > 0.1), then we have not observed anything unusual when $H_0$ is true so there is no evidence based on the observed data to suggest that the null hypothesis $H_0$ is false.**

# Interpreting the *p*-value

**Suppose *p*-value = 0.05, then the event "observe a *D* value as unusual or more unusual as we have observed" happens only about 5 times out of 100.**

**We interpret a *p*-value close to 0.05 as indicating that the observed data are providing evidence against the null hypothesis $H_0$.**

# Guidelines for Interpreting the *p*-value

**These are only guidelines for this course.**

**The interpretation of a *p*-value must always be made in the context of a given study.**

| *p*-value | Interpretation |
|---|---|
| $p > 0.1$ | There is no evidence against $H_0$ based on the data. |
| $0.05 < p \leq 0.1$ | There is some evidence against $H_0$ based on the data. |
| $0.01 < p \leq 0.05$ | There is evidence against $H_0$ based on the data. |
| $0.001 < p \leq 0.01$ | There is strong evidence against $H_0$ based on the data. |
| $p \leq 0.001$ | There is very strong evidence against $H_0$ based on the data. |