

## 10.4 Inference as Decision



### The 1995 O.J. Simpson trial: the situation



- Nicole Brown Simpson and Ronald Goldman were brutally murdered sometime after 10:00 pm on June 12, 1994. Nicole was the wife of O.J. Simpson, former NFL player (winner of 1968 Heisman Trophy, Hall of Famer).
- O.J., like any American accused of a crime, is assumed innocent.
- Evidence collected: bloody size 12 Bruno Magli shoeprint, bloody glove, blood spots on white Ford Bronco, neighbor heard a thump on the condominium wall, ...
- Criminal trial: O.J. is accused of the murders
- Civil trial: O.J. is accused of wrongfully causing the deaths

### O.J. Simpson trial: jury decisions

- **Criminal trial:** did not reject assumption of innocence.  
Behave as if O.J. is innocent.
- **Civil trial:** reject assumption of innocence.
- Behave as if O.J. is guilty.
- *Was an error made in either trial?*

### Errors in Trials

		Truth (what is actually the case)	
		Innocent	Guilty
Jury decision	Not Guilty	OK	<b>ERROR</b>
	Guilty	<b>ERROR</b>	OK

O.J. innocent  $\Rightarrow$  error made in civil trial.

O.J. guilty  $\Rightarrow$  error made in criminal trial.

### Errors in Hypothesis Testing

		$H_0$ really true	$H_0$ really false
Decision	Do not reject $H_0$	OK	? $\beta$ TYPE II ERROR
	Reject $H_0$	TYPE I ERROR	$\alpha$ power OK

Type I  $\approx$  false alarm.

Type II  $\approx$  a miss.

**Example: Grade inflation?**

Has the average college gpa increased beyond 2.7?

$$H_0: \mu = 2.7$$

$$H_a: \mu > 2.7$$

Random sample  
of students

Statistics

$$n = 36$$

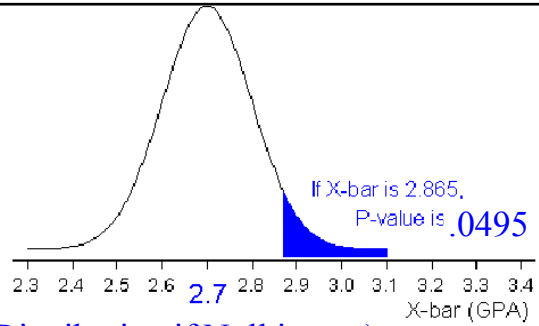
$$\sigma = 0.6$$

and  $\bar{x}$  is also found

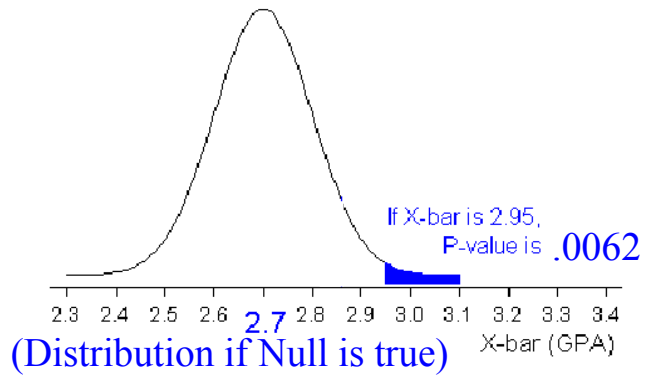
**Decision Rule**

We'll choose  $\alpha = 0.05$ .

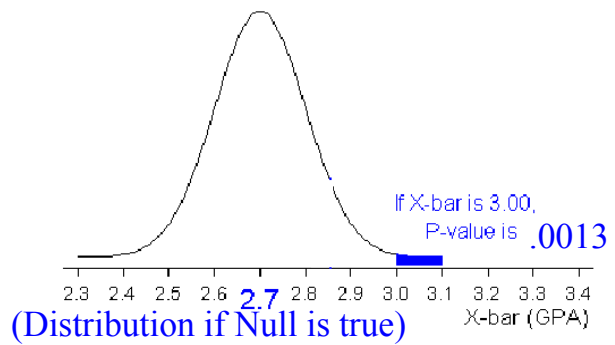
If  $\bar{x} = 2.865 \dots$



If  $\bar{x} = 2.95 \dots$

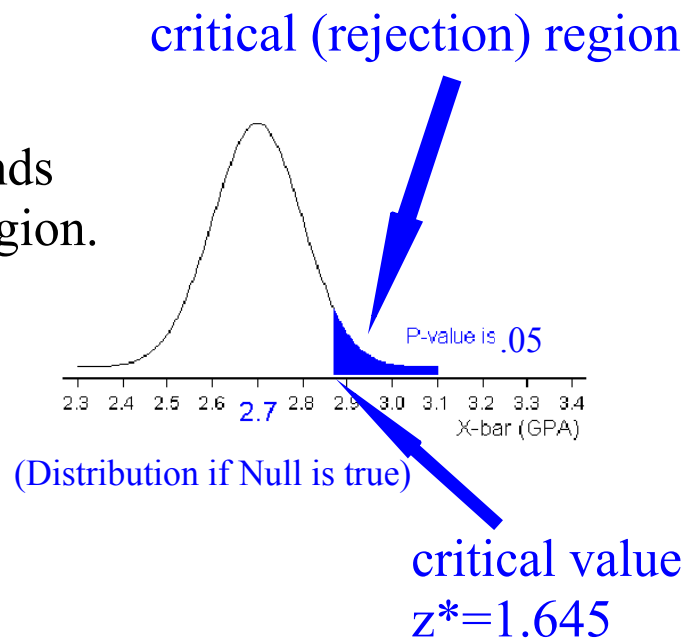


If  $\bar{x} = 3.0 \dots$



Each time,  
 $p\text{-value} \leq 0.05$ ,  
 so we reject  $H_0$ .

Alternative (older) method:  
Find critical value  $z^*$  that  
gives  $p\text{-value} = 0.05$   
&  
reject  $H_0$  if your  $z$  score lands  
in the critical (rejection) region.



Significance level,  $\alpha = P(\text{Type I error})$ .

Common values:  $\alpha = 0.01, 0.05, \text{ or } 0.10$ .

Choose value based on seriousness of Type I error.

- In criminal trials:

Type I error: innocent sentenced to imprisonment or death.

"Beyond a reasonable doubt."  $\Rightarrow \alpha$  really small, such as 0.001.

- In civil trials:

Type I error: innocent fined.

"Preponderance of evidence."  $\Rightarrow \alpha$  not so small, such as 0.10.



Example: Serious Type I Error

A new drug is supposed to reduce blood pressure, but can have serious side effects.

Null Hypothesis: new drug has no impact on blood pressure		
	$H_0$ is true	$H_0$ is false
Decision	Do not reject $H_0$	OK
	Reject $H_0$	Approve drug, though <u>not</u> effective.
		Don't approve drug, which <u>is</u> effective.
		$\beta$ power
		$\alpha$ OK

Result: Expose patients to serious side effects without real benefit.

Therefore, we make  $\alpha$  small, perhaps 0.01 or less.

Example: Serious Type II Error

Snow is forecast.

$H_0$ : No big changes in road conditions due to snow

Null: snow won't be bad enough to cancel school.  
(no effect or no difference from a normal day)

Decision

	$H_0$ is true	$H_0$ is false
Do not reject $H_0$	OK have school, roads good ?	$\beta$ bad roads, & in school
Reject $H_0$	$\alpha$ no snow & no school	power OK school off, bad roads

Result: People in danger trying to get to school.

Example: Not so serious Type I Error

A university suspects grade inflation is occurring.

Null: (No grade inflation-- Avg. GPA still the same.)		
	$H_0$ is true	$H_0$ is false
Decision	Do not reject $H_0$	OK  ? $\beta$ Decide GPAs are not inflated, school becomes a joke
	Reject $H_0$	Mistakenly conclude avg. GPA has increased.  $\alpha$ power OK

Result: Instructors grade harder, students become unhappy, but diplomas may be considered impressive.

Therefore, we make  $\alpha$  bigger, perhaps up to 0.10.

Type II Error: Fail to reject false  $H_0$ .

If  $H_0$  is false, ( $H_a$  is true)

$$\begin{array}{rcccl} P(\text{correctly reject } H_0 | H_0 \text{ is false}) & + & P(\text{incorrectly retain } H_0 | H_0 \text{ is false}) & = & 1 \\ \text{Power} & & + & P(\text{Type II error}) & = & 1 \end{array}$$

**So, Power = 1 - P(Type II error)**

The **Power** of a test is the probability of correctly rejecting  $H_0$  when it's false.

## Power

- is probability of correctly rejecting false  $H_0$ .
- 0 is bad!
- 1 is good!
- **Calculating is not** on the AP Exam.
- **Knowing about is** on the AP Exam.

<http://wise.cgu.edu/power/powerapplet1.html>



Change: Move  $H_a$  distribution,  $\alpha$ ,  $\sigma$ , or  $n$ .

<http://www.stat.uiowa.edu/~rlenth/Power/>



<http://www.stat.sc.edu/~ogden/javahtml/power/power.html>



### These affect power...

- Difference between value under  $H_0$  and the actual value
- $P(\text{Type I error}) = \alpha$
- Standard deviation
- Sample size

### Power is higher when...

- the true mean is far from the mean specified in  $H_0$ .
- $\alpha$  is higher (though this increases the chance of a Type I error).
- $\sigma$  is smaller.
- $n$  is larger.

Men's Health magazine reported that "research has shown that blood pressure may increase by as much as 10 points in nonhabitual coffee drinkers." To follow-up on this research, a researcher wants to test the hypotheses  $H_0: \mu = 10$ ,  $H_a: \mu > 10$ . Identify the type I and type II errors.

	$H_0$ really true	$H_0$ really false
Decision	Do not reject $H_0$	OK
	Reject $H_0$	OK

Men's Health magazine also reported that "Drinks that contain caffeine increase blood pressure by nearly 4 points, on top of speeding up your heart rate by an average of 2 beats per minute. It's enough to push a borderline heart problem into the danger zone." In doing the research on heart rate, suppose the scientists tested the hypotheses  $H_0: \mu = 2$ ,  $H_a: \mu > 2$ . Identify the type I and type II errors.

	$H_0$ really true	$H_0$ really false
Decision	Do not reject $H_0$	OK
	Reject $H_0$	OK



The Eastern Mediterranean Health Journal reported that “The Islamic Republic of Iran has been recognized as an area of iodine deficiency since 1969. Extensive studies in the 1980s found goiter to be hyperendemic in many areas. Complications of severe iodine deficiency, such as retardation in physical and mental development, severe neurological, psychomotor and auditory deficits, and hypothyroidism have also been observed. Though there is essentially the same mean daily salt intake in the two Iranian cities of Rasht and Sari, but the mean concentration of iodine in urine was significantly [different] in the two cities.” Suppose that the scientists used the hypotheses  $H_0: \mu_1 = \mu_2$ ,  $H_a: \mu_1 \neq \mu_2$ . Identify the type I and type II errors.

	$H_0$ really true	$H_0$ really false
Decision	Do not reject $H_0$	OK
	Reject $H_0$	OK