Errors. Power. More power.

AP Statistics
Chapter 21

When we perform a hypothesis test…

	H_o True	H_o False
Reject H_o	**Type I error** α	✓
Fail to reject H_o	✓	**Type II error** β

Consider a test for a serious disease…

What are the hypothesis?

Ho: the person is HEALTHY

Ha: the person is NOT healthy

Type I error: **Decide that the person has the disease…**
…but in fact they are healthy!

Consequence:

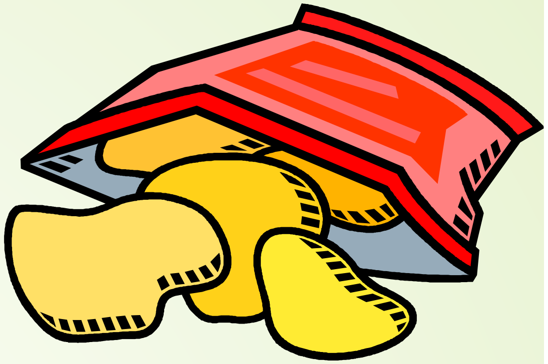
Person undergoes further testing…

Type II error: **Decide that the person is healthy…**
…but in fact they have the disease…

Consequence:

Person fails to get the treatment they need…

Lay's chip company tests a sample of potatoes from a truckload for E-coli to determine whether or not to accept the truckload.



What are the hypothesis?

Ho: the potatoes are good

Ha: the potatoes are bad

Type I error: **Decide the potatoes are BAD when they really are good.**

Type II error: **Decide the potatoes are good when they really are BAD.**

From the **potato supplier's** viewpoint, which is more serious? **A type I error**

From **Lay's chip company's** viewpoint, which is more serious? **A type II error**

A school district is considering purchasing laptops for all of its 10th and 11th grade students, in hopes that using the devices will improve student achievement on end-of-year exams.

Is this a **one-tailed** or a **two-tailed** test?

One-tailed. The district won't spend $$$ on laptops if exam performance gets WORSE

What are the hypothesis?

Ho: student achievement will not improve
Ha: student achievement improves (due to laptops)

Type I error: Decide the laptops make a positive impact... when really they don't.

Type II error: Decide the laptops don't make a difference... when really they would.

(let's think about a criminal trial...)

Let's pretend we have a defendant that we **KNOW** is **GUILTY** of a crime.

Ho: the defendant is innocent
Ha: the defendant is guilty

What do we need to convict the criminal?

(i.e., reject the null hypothesis)

STRONG EVIDENCE.

Without it, we risk committing a type II error.

(let's talk about treatments for terrible diseases...)

Let's pretend we have a new treatment that we KNOW is better than the old…

Ho: the new treatment **is the same as** the old
Ha: the new treatment **is better than** the old

What do we need to prove that the new treatment is better? (again – to reject the null hypothesis?)

again – STRONG EVIDENCE!

In hypothesis testing, and we call this…

Power

the probability of correctly rejecting a false Ho.

$$\text{Power} = 1 - \beta$$

3 ways to increase power:

(a look at the consequences of each)

- **increase α** lowers P(type II) and increases power… but increases P(type I)
- **increase n** lowers P(type I) and P(type II) AND increases power (but requires a larger sample size)
- **increase effect size**

  ALSO lowers P(type I) and P(type II) AND increases power

AP Statistics – Errors & Power with Hypothesis Testing

1. In preparation for the upcoming flu season, a pharmaceutical company has spent **MILLIONS OF DOLLARS** and **MONTHS OF RESEARCH** and development preparing a flu vaccine. The current vaccine is believed to be effective in preventing the flu on 47% of adult patients, and the company is certain that the new version of the flu vaccine will be even more effective.

   The researchers for the company are preparing a clinical trial with 100 patients, at the end of which they will perform a hypothesis test at the $\alpha = 0.01$ level of significance. The results of this test will be used to determine whether the company will proceed with mass production of the new vaccine.

   a) State the hypotheses for this company's test (both with symbols and with words).

   p = true proportion of all adult patients on whom the new vaccine is effective

   H_0: $p = 0.47$ (the new vaccine is no more effective than the old)
   H_A: $p > 0.47$ (the new vaccine is MORE effective than the old)

   b) Identify what a type I and type II error would be, in the context of this problem. Also identify a practical consequence of each.

   Type I Error (H_0 is true, but we reject it)
   In truth, the new vaccine is NOT any more effective than the current, but the company decides that it is. As a result, they go on producing a vaccine that is no better than the current vaccine (which may cause people to waste money on it, etc).

   Type II Error (H_0 is false, but we fail to reject it)
   In truth, the new vaccine is better than the current, however the test fails to show this, so we decide that it is not more effective. As a result, the company does NOT proceed with production of the new vaccine, and patients who might benefit from it will not have a chance to receive it.

   c) The biomedical engineers who developed the new vaccine believe that it should be effective in preventing the flu on 60% of patients. Based on this alternative, along with the sample size of 100 and a $\alpha = 1\%$ level of of significance, the power of the test is calculated to be 0.612*.
   Interpret the meaning of this value (power = 0.612) in the context of this scenario.

   (Power is the probability of correctly rejecting a false H_0)
   There is roughly a 0.612 probability (61.2%) that the test shows that the new vaccine is more effective than the current vaccine, assuming that this is in fact correct.

   d) What are **2 ways** the company could increase the power of the hypothesis test?
   Use a larger sample size (increase n) or test with a higher alpha level (increase α).

   e) Let's suppose that the company sends the biomedical team back to the drawing board to improve the effectiveness of the new vaccine. If "p" represents the **true** effectiveness of the new vaccine, under which of the following scenarios would the power of the test be the greatest? (HINT: THINK ABOUT THE HA)

   p = 0.1   p = 0.2   p = 0.45   p = 0.5   [✓ p = 0.6]   (this is the value that MOST agrees with the Ha)

If you're curious about how to actually CALCULATE power with this information, don't. Calculating power is beyond the scope of the AP Statistics curriculum.