Apr. 18 Statistic for the day: 
In New York, 25% of the 293 deaths among pregnant women between 1987 and 1991 were homicides.

Assignment: 
Read Chapter 23 
Do exercises 1, 2, 3, 8, 16

The reasoning:
The p-value is the probability that, if the skeptic is correct, we would have observed the result we did (p. 415).

Thus, if the p-value is small, the assumption “skeptic is correct” looks suspicious, so we support the research advocate.

If the p-value is not small, supporting the research advocate might be riskier (because we might be making a type I error), so we do not support the research advocate.

(Note: “Support the research advocate” is the same as “reject the null hypothesis” in this scenario.)

Hypotheses for calcium study?

Alternative (Research) Hypothesis: Calcium will reduce the severity of PMS, as compared with placebo.
Null Hypothesis (Skeptic): There is no effect due to calcium except the placebo effect.

The alternative in this case is one-sided.

Calcium and PMS
(Case study 21.1, pp. 400-402, and exercise 3, p. 428.)
A study was conducted to see if a calcium supplement relieves the symptoms of premenstrual syndrome (PMS). Women were randomly assigned a placebo or a calcium supplement and a measure of severity of PMS was recorded.

<table>
<thead>
<tr>
<th></th>
<th>Placebo</th>
<th>Calcium</th>
</tr>
</thead>
<tbody>
<tr>
<td>mean</td>
<td>.60</td>
<td>.43</td>
</tr>
<tr>
<td>SD</td>
<td>.52</td>
<td>.40</td>
</tr>
<tr>
<td>sample size</td>
<td>228</td>
<td>212</td>
</tr>
</tbody>
</table>

The p-value is less than .0001
What is your decision?

P-value < .0001.
The researchers concluded that calcium is effective. Which type of error could they have committed?

Type 1 error: claim calcium is effective (Research Advocate) when it is not (Skeptic is correct).

Type 2 error: claim calcium is not effective (Skeptic) when it is effective (Research Advocate is correct).

Since the researchers supported the Research Advocate, the only possible error is Type 1. But with small prob.
What would be the consequences of making a Type 1 error in this experiment? a Type 2 error?

Type 1 error: claim calcium is effective (Research Advocate) when it is not (Skeptic is correct).
Consequence: Perhaps market a product that is not helpful. Put a bogus drug on the market.

Type 2 error: claim calcium is not effective (Skeptic) when it is effective (Research Advocate is correct).
Consequence: Fail to market an effective drug.

Note: we could not have committed a Type 2 error since the researchers supported the Research Advocate.

The computation of the p-value:

<table>
<thead>
<tr>
<th></th>
<th>Placebo</th>
<th>Calcium</th>
</tr>
</thead>
<tbody>
<tr>
<td>mean</td>
<td>.60</td>
<td>.43</td>
</tr>
<tr>
<td>SD</td>
<td>.52</td>
<td>.40</td>
</tr>
<tr>
<td>sample size</td>
<td>228</td>
<td>212</td>
</tr>
<tr>
<td>SEM</td>
<td>.034</td>
<td>.027</td>
</tr>
<tr>
<td>SD of the difference</td>
<td>sqrt(.034^2 + .027^2) = .043</td>
<td></td>
</tr>
</tbody>
</table>

Test Statistic = (.60 - .43)/.043 = 3.95

From p. 157, the proportion BEYOND 3.78 is less than .0001. So p-value less than .0001.

Hiring a bartender

You have just become the new owner of Zeno’s.

Your first task is to hire a new bartender.

You would like to hire someone with experience, someone who can distinguish among beers.

You decide to hire on the basis of the ability to discriminate between Bud and Yuengling.

For example:
A test consists in picking Yuengling from 3 Buds and 1 Yuengling which have been randomly ordered.

You give the prospective bartender 15 tests and record the number of successes and also the proportion of successes.

The rule for hiring the person will entail that the person get at least a certain number of successes.

Suppose the person gets 6 out of 15 correct.
Proportion = 6/15 = .4

What sort of a proportion would you expect if the person was guessing? 1/4 or .25

We want the probability of hiring a bogus bartender to be small.
Recall our test:
4 beers: 3 Buds and 1 Yuengling
Person must pick the Yuengling.

Gave the test 15 times and the person got 40% correct.
Let’s find a test statistic and then a p-value.

Cast the problem of hiring a bartender as a hypothesis testing problem.

Null Hypothesis (Skeptic): The person is guessing

Alternative Hypothesis (Research Advocate):
The person can discriminate between beers.

Type 1 error: Hire a bartender who can not discriminate.
p-value = Probability of committing a Type 1 error.
= Probability of hiring a bogus bartender
How to compute the test statistic.

Assuming the Skeptic (null) is correct, the standard deviation of the sample proportion is:

\[ \sqrt{\frac{0.25 \times 0.75}{15}} = .112 \]

Note: We used .25×.75, not .40×.60, in the st. dev.

Test statistic or standardized score:

\[ \frac{\text{Estimate} - \text{Null value}}{\text{Standard dev. of estimate}} = \frac{.40 - .25}{.112} = 1.34 \]

How to compute the p-value.

The test statistic was 1.34. Looking it up on p. 157, we see that the percent above 1.34 is 9% (because the percent below is 91%).

Therefore, the p-value is .09.

In this case, we would probably support the Skeptic and not hire the person; the probability of a Type 1 error is too big.

If the person is guessing, there is a 1% chance they will get at least 6 out of 15 correct.

Suppose the person gets 9 out of 15 correct.
Proportion = 0.6

The test statistic is 3.125; the p-value is less than .0013.

Decision?
Hire the person.
p-value = probability of hiring a person when the person is guessing.
p-value is small so it is unlikely that we are hiring a bogus bartender.

How to compute the test statistic.

Assuming the Skeptic (null) is correct, the standard deviation of the sample proportion is:

\[ \sqrt{\frac{0.25 \times 0.75}{15}} = .112 \]

Note: We used .25×.75, not .40×.60, in the st. dev.

Test statistic or standardized score:

\[ \frac{\text{Estimate} - \text{Null value}}{\text{Standard dev. of estimate}} = \frac{.60 - .25}{.112} = 3.125 \]

How to compute the p-value.

The test statistic was 3.125. Looking it up on p. 157, we see that the percent above 3.00 is 0.13% (because the percent below is 99.87%).

Therefore, the p-value is .0013.

In the case of the person who got 9 out of 15 correct with a p-value of .09, we decided not to hire the person.

In the case of the person who got 6 out of 15 correct with a p-value less than .0013, we decided to hire her.

In the first case is it possible to commit a Type 1 error?
No. Type 1 error = hire a bogus bartender.

In the first case is it possible to commit a Type 2 error?
Yes. Type 2 error = do not hire a good bartender.

In the second case is it possible to commit a Type 1 error?
Yes.
You have a suspicious coin.

You flip this coin 100 times and observe 60 heads.
Set up a statistical hypothesis test to determine whether this is convincing evidence that the coin is weighted unfairly to come up heads.