P-values! What do they say about the data?

Part 1: Effects of difference between sample proportion and null proportion on the $P$-value.

For each $P$-value: after obtaining it, determine its reciprocal, and use this to get a rough idea of how small it is. For example: $P$-value = 0.0647; reciprocal = 15.456; $P$-value $\approx 1 / 15$. Whatever probability the $P$-value represents, the “event” occurs approximately once in every 15 repeats on average: Such an event is not very common (“usual”), but also not extremely uncommon (it is somewhat “unusual”).

A national survey of restaurant employees found that 75% said that work stress had a negative impact on their personal lives. An executive of a restaurant chain wants to know whether his employees differ from all restaurant employees, and will test

$$H_0: p = 0.75 \quad H_1: p \neq 0.75.$$ 

A random sample of 100 employees of the chain finds that 82 answer “Yes” when asked “Does work stress have a negative impact on your personal life.

1. Compute the test statistic $Z$. Then obtain the $P$-value for this test.

$$Z = \quad P\text{-value} = \quad \approx 1 \text{ in } ______$$

2. The significance levels that are usually proposed in textbook exercises are 0.10, 0.05 and 0.01. What is the decision for each of these levels?

<table>
<thead>
<tr>
<th>Level</th>
<th>0.01</th>
<th>0.05</th>
<th>0.10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Decision</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

3. How different is the sample result from 0.75? Subtract to find out.

4. At another chain a similar study is conducted – again with 100 employees sampled – and it turns out that 68 of the 100 people answer “Yes.” How different is this result from 0.75?

What do you think the $P$-value will be for this chain’s data / study – higher than, lower than, or the same as 0.1056 (the solution to 1)?

5. Obtain the test statistic and $P$-value for a result of 68 / 100. Compare to your answer to 1, and then check back on your thoughts from 4.

$$Z = \quad P\text{-value} = \quad \approx 1 \text{ in } ______$$

6. Suppose instead the survey result were 78 / 100. A $P$-value is a quantitative assessment of the evidence in the data in favor of the alternative hypothesis. Which data, 82 / 100 or 78 / 100, is most favorable to the alternative hypothesis? Why?
7. Determine the test statistic and P-value for a result of 78 / 100. Is the P-value larger or smaller than the P-value from 1?

\[ Z = \]  
\[ P\text{-value} = \approx 1 \text{ in } \]______

8. Suppose instead the survey result were 83 / 100. Which data, 83 / 100 or 82 / 100 is most favorable to the alternative hypothesis?

9. Determine the test statistic and P-value for a result of 83 / 100.

\[ Z = \]  
\[ P\text{-value} = \approx 1 \text{ in } \]______

How does this P-value compare to the 0.1056 that accompanies a result of 82 / 100 (see 6 and 7 above).

10. (See the pattern? See how it works?) What do you think will happen to the P-value (higher than/lower than/the same as 0.0647) if the result is instead 85 / 100?

11. Determine the test statistic and P-value for a result of 85 / 100.

\[ Z = \]  
\[ P\text{-value} = \approx 1 \text{ in } \]______

12. For each result considered in the rows of the table below, determine the P-value (you’ve done some already) and appropriate decision as in the columns of the table.

<table>
<thead>
<tr>
<th>Result</th>
<th>P-value</th>
<th>( \alpha = 0.10 )</th>
<th>( \alpha = 0.05 )</th>
<th>( \alpha = 0.01 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \hat{p} = 78 / 100 )</td>
<td>0.4884</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \hat{p} = 82 / 100 )</td>
<td>0.1060</td>
<td>Do not reject ( H_0 )</td>
<td>Do not reject ( H_0 )</td>
<td>Do not reject ( H_0 )</td>
</tr>
<tr>
<td>( \hat{p} = 83 / 100 )</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \hat{p} = 85 / 100 )</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \hat{p} = 87 / 100 )</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

13. If you know that the decision is to reject \( H_0 \) at the 10% level, can you be certain that \( H_0 \) is rejected at a lower level (say 5% or 1%)?

14. If you know that the decision is to reject \( H_0 \) at the 1% level, can you be certain that \( H_0 \) is rejected at a higher level (say 5% or 10%)?
What follows is a plot of the $P$-value for all of the possible outcomes in a sample of size $n = 100$. (There’s no reason you couldn’t plot this fairly completely, but it does take some time.)

Part 2: Effects of sample size on the $P$-value.

15. An outcome of $\hat{p} = \frac{78}{100}$ yields a test statistic of $Z = 0.693$ and corresponding $P$-value of 0.4884. Suppose we had the same proportion, but based on a larger sample size. If the sample size is instead 400, and $\hat{p} = 0.78$, then how many of the 400 employees answer “Yes?”

16. Is a sample of size $n = 400$ more, or less, informative than a sample of size 100?

17. Find the value of the test statistic and the $P$-value for the test when $\frac{312}{400}$ people answer “Yes.” How do these compare to the results stated in 15?

$$Z = \quad P\text{-value} = \quad \approx 1 \text{ in } \square$$

From before: A $P$-value is a quantitative assessment of the evidence in the data in favor of the alternative hypothesis. You also previously worked out this: (For samples of the same size) lower $P$-values are indicative of more evidence in favor of the alternative. Applying this to your $P$-value from 17, we conclude that a larger sample with the same difference of 0.07 from 0.75 as for a smaller sample gives *more evidence* that $p \neq 0.75$ (for the population).
18. Complete this table (use D N R for “Do not reject” and R for “Reject”)

<table>
<thead>
<tr>
<th>Result</th>
<th>Z =</th>
<th>P-value</th>
<th>( \alpha = 0.10 )</th>
<th>( \alpha = 0.05 )</th>
<th>( \alpha = 0.01 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \hat{p} = 78 / 100 = 0.78 )</td>
<td>0.693</td>
<td>0.4884</td>
<td>D N R H(_0)</td>
<td>D N R H(_0)</td>
<td>D N R H(_0)</td>
</tr>
<tr>
<td>( \hat{p} = 312 / 400 = 0.78 )</td>
<td>1.386</td>
<td>0.1689</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \hat{p} = 780 / 1000 = 0.78 )</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \hat{p} = 3120 / 4000 = 0.78 )</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Do you see the pattern with the Z values? If \( n \) is increased (decreased) by a multiplicative factor of \( k^2 \), the test statistic \( Z \) increases (decreases) by a factor of \( k \).

19. (Review 1, 4 and 5 above.) Without doing any calculations: What are the values of the test statistic \( Z \) and \( P \)-value for a result of \( \hat{p} = 720 / 1000 = 0.72 \)?

\[ Z = \quad P\text{-value} = \quad \approx 1 \text{ in } \_ \_ \_ \_ \_ \]

20. Consider a result of \( \hat{p} = 751117 / 1000000 = 0.751117 \). Notice that this result is almost indistinguishable from the null value of 0.75. Yet the sample size is huge. What are the test statistic \( Z \) and \( P \)-value for such a result? Is the null hypothesis rejected at \( \alpha = 0.01 \)?

\[ Z = \quad P\text{-value} = \quad \approx 1 \text{ in } \_ \_ \_ \_ \_ \]

21. Now consider a result of \( \hat{p} = 22 / 25 = 0.88 \). This is quite different from the null value of 0.75. But the sample size is quite small (it’s particularly small for Success / Failure data). What is the \( P \)-value for such a result? Is the null hypothesis rejected at \( \alpha = 0.10 \)?

22. Next consider a result of \( \hat{p} = 8 / 20 = 0.40 \). The sample size is even smaller now. What is the \( P \)-value for such a result? Is the null hypothesis rejected at \( \alpha = 0.01 \)? How about \( \alpha = 0.001 \)?

23. Finally: Is it possible to make strong conclusions based on very small samples? If so – what about the data allows this to happen?
### Solutions

1. \( Z = 1.62; \ P\text{-value} = 0.1060 \approx 1 \text{ in } 9 \)
2. Do not reject for all three.
3. It’s 0.07 above 0.75.
4. It’s 0.07 below 0.75.
5. \( Z = -1.62; \ P\text{-value} = 0.1060 \approx 1 \text{ in } 9 \) (the same)
6. 82 is more favorable – 82/100 is farther from 0.75 than 78/100.
7. \( Z = 0.69; \ P\text{-value} = 0.4884 \approx 1 \text{ in } 2 \)
8. 83/100 is more favorable.
9. \( Z = 1.85; \ P\text{-value} = 0.0647 \approx 1 \text{ in } 15 \). The \( P\text{-value} \) is smaller.
10/11. \( Z = 2.31; \ P\text{-value} = 0.0209 \approx 1 \text{ in } 48 \). The \( P\text{-value} \) is smaller.

<table>
<thead>
<tr>
<th>Result</th>
<th>( P\text{-value} )</th>
<th>( \alpha = 0.10 )</th>
<th>( \alpha = 0.05 )</th>
<th>( \alpha = 0.01 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \hat{p} = 78/100 )</td>
<td>0.4884</td>
<td>Do not reject ( H_0 )</td>
<td>Do not reject ( H_0 )</td>
<td>Do not reject ( H_0 )</td>
</tr>
<tr>
<td>( \hat{p} = 82/100 )</td>
<td>0.1060</td>
<td>Do not reject ( H_0 )</td>
<td>Do not reject ( H_0 )</td>
<td>Do not reject ( H_0 )</td>
</tr>
<tr>
<td>( \hat{p} = 83/100 )</td>
<td>0.0647</td>
<td>Reject ( H_0 )</td>
<td>Do not reject ( H_0 )</td>
<td>Do not reject ( H_0 )</td>
</tr>
<tr>
<td>( \hat{p} = 85/100 )</td>
<td>0.0209</td>
<td>Reject ( H_0 )</td>
<td>Reject ( H_0 )</td>
<td>Do not reject ( H_0 )</td>
</tr>
<tr>
<td>( \hat{p} = 87/100 )</td>
<td>0.0056</td>
<td>Reject ( H_0 )</td>
<td>Reject ( H_0 )</td>
<td>Reject ( H_0 )</td>
</tr>
</tbody>
</table>

13. No.
14. Yes.
15. 0.78*400 = 312.
17. \( Z = 1.39; \ P\text{-value} = 0.1659 \approx 1 \text{ in } 6 \). The \( P\text{-value} \) is smaller.

<table>
<thead>
<tr>
<th>Result</th>
<th>( Z = )</th>
<th>( P\text{-value} )</th>
<th>( \alpha = 0.10 )</th>
<th>( \alpha = 0.05 )</th>
<th>( \alpha = 0.01 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \hat{p} = 78/100 = 0.78 )</td>
<td>0.693</td>
<td>0.4884</td>
<td>D N R ( H_0 )</td>
<td>D N R ( H_0 )</td>
<td>D N R ( H_0 )</td>
</tr>
<tr>
<td>( \hat{p} = 312/400 = 0.78 )</td>
<td>1.386</td>
<td>0.1689</td>
<td>D N R ( H_0 )</td>
<td>D N R ( H_0 )</td>
<td>D N R ( H_0 )</td>
</tr>
<tr>
<td>( \hat{p} = 780/1000 = 0.78 )</td>
<td>2.191</td>
<td>0.0285</td>
<td>R ( H_0 )</td>
<td>R ( H_0 )</td>
<td>D N R ( H_0 )</td>
</tr>
<tr>
<td>( \hat{p} = 3120/4000 = 0.78 )</td>
<td>4.382</td>
<td>0.0000</td>
<td>R ( H_0 )</td>
<td>R ( H_0 )</td>
<td>R ( H_0 )</td>
</tr>
</tbody>
</table>

19. \( 720/1000 = 0.72 \), which is 0.03 below 0.75. The \( P\text{-value} \) is the same as for \( 780/1000 = 0.78 \) which is 0.03 above 0.75. The \( P\text{-value} \) is 0.0285, about 1 in 35.
20. \( Z = 2.580; \ P\text{-value} = 0.0099 \approx 1 \text{ in } 101 \). The null is rejected.
21. \( Z = 1.501; \ P\text{-value} = 0.1333 \approx 1 \text{ in } 8 \). The null is not rejected.
22. \( Z = -3.615; \ P\text{-value} = 0.0003 \approx 1 \text{ in } 3327 \). The null is rejected at the 0.001 level.
23. Yes it is (see 22). However, when this happens, the difference is quite large. In 22 the difference is \( 0.40 - 0.75 = 0.35 \) (35%).
Meaning from P-values