**P-value: Explained**

[[NitinKumar Sharma](https://medium.com/@nitin.data1997?source=post_page-----bc9649c32dd2--------------------------------)](https://medium.com/@nitin.data1997?source=post_page-----bc9649c32dd2--------------------------------)

[NitinKumar Sharma](https://medium.com/@nitin.data1997?source=post_page-----bc9649c32dd2--------------------------------)

·

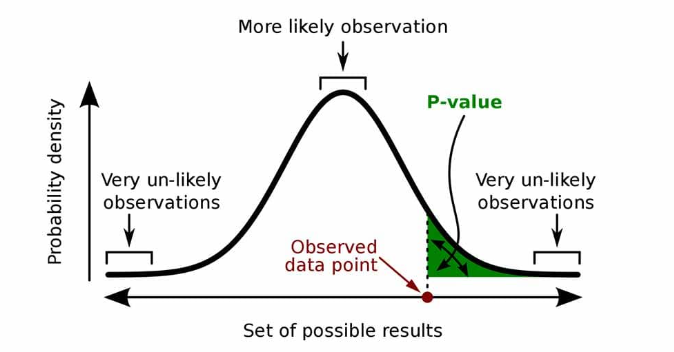
[Follow](https://medium.com/m/signin?actionUrl=https%3A%2F%2Fmedium.com%2F_%2Fsubscribe%2Fuser%2Fce6169a8b73e&operation=register&redirect=https%3A%2F%2Fmedium.com%2F%40nitin.data1997%2Fp-value-explained-to-a-10-year-old-kid-bc9649c32dd2&user=NitinKumar+Sharma&userId=ce6169a8b73e&source=post_page-ce6169a8b73e----bc9649c32dd2---------------------post_header-----------)

6 min read

·

Sep 13, 2023

131



Imagine you have a big jar of colorful marbles. You want to know if the blue marbles are heavier than the red ones. So, you randomly pick a few marbles of each color and weigh them.

***Example 1:***

* If the blue marbles you picked are just a little bit heavier than the red ones, the p-value will be high. It means the weight difference could be due to chance, like some blue marbles being naturally heavier.

***Example 2:***

* But, if the blue marbles you picked are much, much heavier than the red ones, the p-value will be very low. This suggests that the weight difference is likely not due to chance but because blue marbles are generally heavier.

So, the p-value tells us how likely it is that what we observed (the weight difference) happened by chance. A low p-value suggests that what we observed is more likely to be true and not just a coincidence.

In simpler terms, think of the p-value as a detective that helps us decide if something we found is a real clue or just random noise. If the detective (p-value) says it’s a strong clue (low p-value), we take it seriously; if not, we might think it’s just a lucky coincidence (high p-value).

Let’s look at more examples:

***Example 3:***

Coin Toss Imagine you and your friend are flipping a fair coin (which should land heads or tails 50% of the time). If you flip it 10 times and it lands heads 7 times and tails 3 times, you might wonder if the coin is biased.

* A high p-value in this case would mean that the results you got (7 heads, 3 tails) are quite common with a fair coin.
* A low p-value would suggest that these results are unusual and the coin might not be fair.

***Example 4:***

Exam Scores Let’s say you and your classmates take a math test, and you get a score of 90 while your friend scores 60. You might want to know if your higher score is just due to luck or if you’re actually better at math.

* A high p-value here would indicate that the score difference could happen by chance.
* A low p-value would suggest that your higher score is likely because you’re better at math.

**Example 5:**Soccer Game In a soccer game, your team wins 10 out of 20 matches. You wonder if your team is better than the other teams in your league.

* A high p-value would imply that winning half of the games is common and might not mean your team is superior.
* A low p-value would indicate that winning half the games is unusual, and your team might indeed be better.

In all these examples, the p-value helps us decide whether the results we’ve observed are likely due to random chance or if they indicate something significant. It’s like having a detective for your data to separate coincidences from real discoveries.

***P-value***

What is this P-value? The P-value is the probability value that the correlation between these two variables is statistically significant. Normally, we choose a significance level of 0.05, which means that we are 95% confident that the correlation between the variables is significant.

By convention, when the

* p-value is << 0.001: we say there is strong evidence that the correlation is significant.
* the p-value is << 0.05: there is moderate evidence that the correlation is significant.
* the p-value is << 0.1: there is weak evidence that the correlation is significant.
* the p-value is >> 0.1: there is no evidence that the correlation is significant.

Now it seems a little bit confusing looking at the definition. Let’s correlate this with above examples.

**Example 3:** Coin Toss with p-values Suppose you’re flipping a coin 10 times, and it lands heads 7 times and tails 3 times. You want to know if the coin is biased.

1. You would calculate the p-value by performing a statistical test, like a chi-square test. This test would compare your observed results (7 heads, 3 tails) to what you would expect from a fair coin (5 heads, 5 tails).
2. If the p-value is high (let’s say above 0.05), it means your observed results are not significantly different from what you’d expect from a fair coin. In this case, you might conclude that the coin is likely fair.
3. If the p-value is low (below 0.05), it suggests that your results are significantly different from what you’d expect by chance. You might conclude that the coin could be biased.

**Example 4:** Exam Scores with p-values When comparing your score of 90 to your friend’s score of 60, you might want to know if the difference is significant.

1. Calculate the p-value using a statistical test, like a t-test. This test compares the means (average scores) of both groups.
2. A high p-value indicates that the score difference could occur by chance, suggesting no significant difference between your math abilities.
3. A low p-value suggests that the score difference is unlikely due to chance, indicating a significant difference in math abilities.

**Example 5:**Soccer Game with p-values To determine if your team is better than other teams in the league, you’d involve p-values.

1. Perform a statistical test, such as a chi-square test or a binomial test, to compare your team’s win rate to the expected win rate (50% if all teams are equally matched).
2. A high p-value suggests that your team’s win rate is not significantly different from what you’d expect by chance, indicating no clear superiority.
3. A low p-value implies that your team’s win rate significantly differs from what you’d expect by chance, suggesting that your team might be better or worse than average.

***Question:  
What will happen to the p-value in the second case if you observe the same values for the sample mean and the sample standard deviation for both the cases?***

Options:  
A) It will increase.  
B) It will decrease.  
C) It will stay the same.  
D) Cannot be determined.

**Answer:  
B) It will decrease.**

**Explanation:**  
With an increase in the sample size, the denominator of the Z-score decreases, which makes the absolute value of Z-score increase. This causes the p-value to decrease. Increasing the sample size narrows the distribution of sample means, making it more likely for the sample mean to fall in the critical region, resulting in a smaller p-value.

***Question:  
Consider the null hypothesis that a process produces no more than the maximum permissible rate of defective items. In this situation, a type-II error would be \_\_.***

Options:  
A) To conclude that the process does not produce more than the maximum permissible rate of defective items when it actually does not.  
B) To conclude that the process produces more than the maximum permissible rate of defective items when it actually does not.  
C) To conclude that the process produces more than the maximum permissible rate of defective items when it actually does.  
D) To conclude that the process does not produce more than the maximum permissible rate of defective items when it actually does.

**Answer:  
D) To conclude that the process does not produce more than the maximum permissible rate of defective items when it actually does.**

**Explanation:**  
A type-II error in hypothesis testing means failing to reject a false null hypothesis. In this context, it means you fail to detect that the process produces more defective items than the maximum permissible rate when, in fact, it does not produce more defective items.

***Question:  
A screening test for a serious but curable disease is similar to hypothesis testing. In this instance, it is better to increase the probability of \_\_\_.***

Options:  
A) Making a type-I error, i.e., not provide treatment when it is needed.  
B) Making a type-I error, i.e., provide treatment when it is not needed.  
C) Making a type-II error, i.e., not provide treatment when it is needed.  
D) Making a type-II error, i.e., provide treatment when it is not needed.

**Answer:  
B) Making a type-I error, i.e., provide treatment when it is not needed.**

**Explanation:**  
In this context, making a type-I error means providing treatment when it is not actually needed. Since the treatment has no serious side effects, it is better to err on the side of caution and provide treatment to avoid missing the disease.

**Thank me Later & Don’t forget to click on clap button!!!**

[P Value](https://medium.com/tag/p-value?source=post_page-----bc9649c32dd2---------------p_value-----------------)

[Statistical Analysis](https://medium.com/tag/statistical-analysis?source=post_page-----bc9649c32dd2---------------statistical_analysis-----------------)

[Chi Square](https://medium.com/tag/chi-square?source=post_page-----bc9649c32dd2---------------chi_square-----------------)

[Understand P Value](https://medium.com/tag/understand-p-value?source=post_page-----bc9649c32dd2---------------understand_p_value-----------------)