Day_01

July 11, 2023

1 Statistics Interview Questions And Answers

2 A. What is Z-test?

A z-test is a statistical test that is used to compare the mean of a sample to the mean of a population. It is a parametric test, which means that it assumes that the data follows a normal distribution. The z-test is most commonly used when the sample size is large (greater than 30) and the population variance is known.

The z-test statistic is calculated as follows:

`z = (sample mean - population mean) / (standard deviation / sqrt(sample size))`

where:

- z is the z-test statistic
- sample mean is the mean of the sample data
- population mean is the mean of the population
- standard deviation is the standard deviation of the population
- sqrt(sample size) is the square root of the sample size

The z-test statistic is then compared to a critical value to determine whether the null hypothesis should be rejected. The critical value is the value of z that has a probability of 0.50 (or 50%) in the tail of the normal distribution. For example, if the significance level is 0.05, then the critical value would be 1.96. This means that there is a 5% chance of obtaining a z-value greater than 1.96 if the null hypothesis is true.

If the z-test statistic is greater than the critical value, then the null hypothesis is rejected. This means that there is enough evidence to conclude that the mean of the sample is different from the mean of the population.

Here is an example of a z-test:

A teacher claims that the mean score of students in his class is greater than 82. The standard deviation of the scores is 20, and a sample of 81 students was selected. The mean score of the sample is 90.

To test the teacher's claim, we can use a z-test. The z-test statistic is calculated as follows:

`z = (90 - 82) / (20 / sqrt(81)) = 2.23`

The critical value for a two-tailed z-test with a significance level of 0.05 is 1.96. Since the z-test statistic is greater than the critical value, we can reject the null hypothesis. This means that there is enough evidence to conclude that the mean score of the students in the teacher's class is greater than 82.

I hope this helps! Let me know if you have any other questions.

2.1 Practical Implemetation Of Z-test

[ ]:
```
import math

def z_test(mean, standard_deviation, sample_mean, significance_level):
  """Performs a z-test.

  Args:
    mean: The population mean.
    standard_deviation: The population standard deviation.
    sample_mean: The sample mean.
    significance_level: The significance level.

  Returns:
    The z-statistic.
  """

  z_statistic = (sample_mean - mean) / standard_deviation
  return z_statistic

def main():
  """Runs the z-test program.

  Prints the z-statistic.
  """

  mean = 100
  standard_deviation = 10
  sample_mean = 105
  significance_level = 0.05
  z_statistic = z_test(mean, standard_deviation, sample_mean,␣
 ↪significance_level)
  print(f"The z-statistic is: {z_statistic}")

if __name__ == "__main__":
  main()
```

```
The z-statistic is: 0.5
```

[ ]:
```
import scipy.stats as stats

# Statistical data
sample_mean = 70.5
```

```
sample_std = 5.1
population_mean = 65.0
sample_size = 50

# Perform a z-test
z_statistic = (sample_mean - population_mean) / (sample_std / (sample_size ** 0.
 ↪5))
p_value = 1 - stats.norm.cdf(z_statistic)

# Output the results
print("Z-statistic:", z_statistic)
print("P-value:", p_value)
```

```
Z-statistic: 7.625661365737279
P-value: 1.2101430968414206e-14
```

3 B. What is T-test?

A t-test is a statistical test that is used to compare the means of two groups. It is a parametric test, which means that it assumes that the data follows a normal distribution. The t-test is most commonly used when the sample size is small (less than 30) or when the population variance is unknown.

There are three main types of t-tests:

- **One-sample t-test:** This test is used to compare the mean of a sample to a hypothesized value. For example, you could use a one-sample t-test to test the hypothesis that the average height of American adults is 5'10".
- **Two-sample t-test:** This test is used to compare the means of two independent groups. For example, you could use a two-sample t-test to test the hypothesis that the average IQ of men is different from the average IQ of women.
- **Paired t-test:** This test is used to compare the means of two dependent groups. For example, you could use a paired t-test to test the hypothesis that the average weight of people before and after a diet is different.

The t-test statistic is calculated as follows:

`t = (mean1 - mean2) / (spooled) / sqrt(n)`

where:

- t is the t-test statistic
- mean1 is the mean of the first sample
- mean2 is the mean of the second sample
- spooled is the pooled standard deviation
- n is the sample size

The t-test statistic is then compared to a critical value to determine whether the null hypothesis should be rejected. The critical value is the value of t that has a probability of 0.50 (or 50%) in the tail of the t-distribution. For example, if the significance level is 0.05, then the critical value would be 1.96. This means that there is a 5% chance of obtaining a t-value greater than 1.96 if the null hypothesis is true.

If the t-test statistic is greater than the critical value, then the null hypothesis is rejected. This means that there is enough evidence to conclude that the means of the two groups are different.

I hope this helps! Let me know if you have any other questions.

3.1 Practical Implementation Of t-test...

[ ]:
```
import math

def t_test(mean1, standard_deviation1, sample_size1, mean2,␣
↪standard_deviation2, sample_size2):
  """Performs a t-test.

  Args:
    mean1: The mean of the first sample.
    standard_deviation1: The standard deviation of the first sample.
    sample_size1: The size of the first sample.
    mean2: The mean of the second sample.
    standard_deviation2: The standard deviation of the second sample.
    sample_size2: The size of the second sample.

  Returns:
    The t-statistic.
  """

  pooled_standard_deviation = math.sqrt(
      (standard_deviation1 ** 2 / sample_size1) + (standard_deviation2 ** 2 /␣
↪sample_size2))
  t_statistic = (mean1 - mean2) / pooled_standard_deviation
  return t_statistic

def main():
  """Runs the t-test program.

  Prints the t-statistic.
  """

  mean1 = 100
  standard_deviation1 = 10
  sample_size1 = 100
  mean2 = 105
  standard_deviation2 = 10
  sample_size2 = 100
  t_statistic = t_test(mean1, standard_deviation1, sample_size1, mean2,␣
↪standard_deviation2, sample_size2)
```

```
  print(f"The t-statistic is: {t_statistic}")

if __name__ == "__main__":
  main()
```

```
The t-statistic is: -3.5355339059327373
```

[ ]:
```
import scipy.stats as stats

# Sample data
group1 = [1.2, 1.5, 2.1, 1.8, 1.6]
group2 = [1.3, 1.7, 1.9, 2.2, 2.4]

# Perform a t-test (e.g., independent two-sample t-test)
t_statistic, p_value = stats.ttest_ind(group1, group2)

# Output the results
print("T-statistic:", t_statistic)
print("P-value:", p_value)
```

```
T-statistic: -1.0650014966747516
P-value: 0.31796027009264544
```

4 1. What is a z-test and when is it used?

A z-test is a statistical test used to determine whether the means of two populations are significantly different when the population standard deviations are known and the sample sizes are large. It is based on the standard normal distribution, which has a mean of 0 and a standard deviation of 1.

The z-test is used in the following scenarios:

1. Hypothesis Testing: The z-test is used to test hypotheses about the population mean when the population standard deviation is known. It helps determine if the observed difference between sample means is statistically significant or simply due to chance.
2. Comparing Means: The z-test is used to compare the means of two independent groups or samples. It determines if there is a significant difference between the means of the two populations being studied.
3. Quality Control: The z-test is used in quality control to assess whether a production process is operating within acceptable limits. It helps determine if the measured sample mean falls within the acceptable range defined by the population mean.
4. A/B Testing: The z-test can be used in A/B testing to compare the performance of two different versions of a website, application, or marketing campaign. It helps determine if the observed difference in outcomes between the two versions is statistically significant.

To perform a z-test, the following steps are typically followed:

1. Formulate Hypotheses: Define the null hypothesis (H0) and alternative hypothesis (Ha) based on the research question.

2. Set Significance Level: Determine the desired level of significance ( ) to control the probability of Type I error.
3. Calculate Test Statistic: Compute the z-statistic using the formula z = (x -) / (/ √n), where x is the sample mean, is the population mean, is the population standard deviation, and n is the sample size.
4. Determine Critical Value: Find the critical value corresponding to the desired level of significance ( ) and the chosen test (one-tailed or two-tailed).
5. Compare Test Statistic and Critical Value: Compare the test statistic with the critical value. If the test statistic falls within the critical region, reject the null hypothesis; otherwise, fail to reject the null hypothesis.
6. Draw Conclusion: Based on the comparison, draw a conclusion regarding the statistical significance of the observed difference between sample means.

The z-test is widely used when sample sizes are large and the population standard deviation is known. However, when the population standard deviation is unknown or the sample size is small, the t-test is more appropriate.

5 2. How is the z-score calculated?

The z-score (also known as the standard score) is a measure that indicates how many standard deviations an individual data point is away from the mean of a distribution. It helps to standardize values and allows for meaningful comparisons across different distributions.

The formula to calculate the z-score is as follows: z = (x -) / Where: - x represents the individual data point - represents the mean of the distribution - represents the standard deviation of the distribution

The z-score formula calculates the difference between the individual data point and the mean, and then divides it by the standard deviation. This normalization process allows us to express the value in terms of standard deviations from the mean.

A positive z-score indicates that the data point is above the mean, while a negative z-score indicates that the data point is below the mean. The magnitude of the z-score indicates how far the data point deviates from the mean in terms of standard deviations.

For example, let's say we have a dataset with a mean of 50 and a standard deviation of 10. If we want to calculate the z-score for a data point of 65, we can use the formula:

z = (65 - 50) / 10 z = 1.5

The z-score of 1.5 indicates that the data point is 1.5 standard deviations above the mean.

Similarly, if we had a data point of 40, the calculation would be: z = (40 - 50) / 10 z = -1

In this case, the z-score of -1 indicates that the data point is 1 standard deviation below the mean.

The z-score is a powerful tool in statistics as it allows us to compare data points from different distributions, identify outliers, and make meaningful interpretations based on their deviation from the mean.

6 3. What is the central limit theorem and how is it related to the z-test?

The Central Limit Theorem (CLT) is a fundamental concept in statistics that states that the sampling distribution of the means (or sums) of a large number of independent and identically distributed random variables approaches a normal distribution, regardless of the shape of the population distribution. The CLT is related to the z-test in the following way:

1. Assumption of Normality: The CLT allows us to assume that the sampling distribution of the mean (or sum) of a sufficiently large sample from any population will be approximately normally distributed, even if the population distribution is not normal. This assumption is crucial for the z-test, as it relies on the assumption of normality.
2. Calculation of Test Statistic: The z-test uses the standard normal distribution as the reference distribution. The test statistic in a z-test is calculated by standardizing the sample mean using the population standard deviation (or an estimated standard deviation). This standardization is achieved by subtracting the population mean from the sample mean and dividing it by the standard deviation. By doing this, the test statistic follows a standard normal distribution.
3. Hypothesis Testing: In the z-test, the z-statistic is compared to critical values from the standard normal distribution to make inferences about the population parameter of interest. The CLT ensures that, for a sufficiently large sample size, the sampling distribution of the mean is approximately normal, allowing for the use of z-scores and critical values.

In summary, the CLT provides the theoretical basis for the z-test, enabling the use of the standard normal distribution to make statistical inferences about population parameters based on sample means. It allows us to make assumptions about the behavior of sample means, regardless of the underlying population distribution, as long as certain conditions are met (e.g., sufficiently large sample size, independence of observations)

The central limit theorem (CLT) states that the sampling distribution of the mean will always be normally distributed, as long as the sample size is large enough. This means that the z-test can be used to compare the mean of a sample to the mean of a population, even if the population is not normally distributed.

The CLT is based on the following assumptions:

- The data is sampled from a population with a finite mean and standard deviation.
- The samples are independent of each other.
- The sample size is large enough.

The CLT states that as the sample size increases, the sampling distribution of the mean will approach a normal distribution, regardless of the shape of the population distribution. This is because the central limit theorem is a law of large numbers, which states that the average of a large number of random variables will approach the expected value of the random variable.

The z-test is a parametric test, which means that it assumes that the data follows a normal distribution. However, the CLT allows us to use the z-test even if the population is not normally distributed, as long as the sample size is large enough.

The z-test statistic is calculated as follows:

```
z = (sample mean - population mean) / (standard deviation / sqrt(sample size))
```

where:

- z is the z-test statistic
- sample mean is the mean of the sample data
- population mean is the mean of the population
- standard deviation is the standard deviation of the population
- sqrt(sample size) is the square root of the sample size

The z-test statistic is then compared to a critical value to determine whether the null hypothesis should be rejected. The critical value is the value of z that has a probability of 0.50 (or 50%) in the tail of the normal distribution. For example, if the significance level is 0.05, then the critical value would be 1.96. This means that there is a 5% chance of obtaining a z-value greater than 1.96 if the null hypothesis is true.

If the z-test statistic is greater than the critical value, then the null hypothesis is rejected. This means that there is enough evidence to conclude that the mean of the sample is different from the mean of the population.

The CLT is an important theorem in statistics because it allows us to use parametric tests, such as the z-test, even when the population is not normally distributed. This is a valuable tool for researchers because it allows them to make inferences about populations based on samples.

7 4. What is a t-test and when is it used?

A t-test is a statistical test used to determine whether the means of two groups are significantly different from each other. It is commonly used when the sample sizes are small and the population standard deviation is unknown. The t-test assesses the likelihood that the observed difference between the sample means is due to chance or represents a true difference in the population means. The t-test is used in the following scenarios: 1. Comparing Means: The t-test is used to compare the means of two independent groups or samples. It helps determine if there is a significant difference between the means of the two populations being studied.

2. Paired Samples: The t-test can be used to compare the means of two related or paired samples. This is often done when the same group of subjects is measured before and after a treatment or intervention.

3. One-Sample Test: The t-test can also be used to compare the mean of a single sample to a known or hypothesized value. This is called a one-sample t-test and helps determine if the sample mean significantly differs from the population mean.

4. Assumptions Testing: The t-test is used to test assumptions in statistical analyses, such as normality assumptions or assumptions of equal variances in different groups.

The t-test is based on the t-distribution, which is similar to the normal distribution but with fatter tails. The test calculates a t-statistic, which measures the difference between the sample means relative to the variability within the samples. The calculated t-statistic is then compared to critical values from the t-distribution to determine statistical significance.

To perform a t-test, the following steps are typically followed:

1. Formulate Hypotheses: Define the null hypothesis (H0) and alternative hypothesis (Ha) based on the research question.

2. Set Significance Level: Determine the desired level of significance ( ) to control the probability of Type I error.

3. Choose the Appropriate Test: Select the appropriate type of t-test based on the study design (independent samples, paired samples, or one-sample).

4. Calculate Test Statistic: Compute the t-statistic using the appropriate formula for the chosen test.

5. Determine Degrees of Freedom: Calculate the degrees of freedom, which depend on the sample sizes and study design.

6. Determine Critical Value: Find the critical value corresponding to the desired level of significance ( ) and the degrees of freedom.

7. Compare Test Statistic and Critical Value: Compare the test statistic with the critical value. If the test statistic falls w ithin the critical region, reject the null hypothesis; otherwise, fail to reject the null hypothesis.

8. Draw Conclusion: Based on the comparison, draw a conclusion regarding the statistical significance of the observed difference between sample means.

The t-test is a widely used statistical test when sample sizes are small and the population standard deviation is unknown. It allows researchers to make inferences about population means based on sample means while accounting for variability within the samples.

8 5. What is the difference between a one-sample t-test and a two-sample t-test?

The difference between a one-sample t-test and a two-sample t-test lies in the comparison being made and the study design involved. Here's a breakdown of each test: One-Sample t-test: - Comparison: A one-sample t-test is used to determine if the mean of a single sample significantly differs from a known or hypothesized value. It assesses whether the sample provides sufficient evidence to reject the null hypothesis that the population mean is equal to the hypothesized value.

- Study Design: In a one-sample t-test, data is collected from a single group or sample, and the mean of that sample is compared to a specified value.
- Assumptions: The assumptions of a one-sample t-test include random sampling, independence of observations, normality of the population distribution (or a sufficiently large sample size for the Central Limit Theorem to apply), and homogeneity of variances (if applicable).

Two-Sample t-test: - Comparison: A two-sample t-test is used to compare the means of two independent groups or samples. It determines if there is a significant difference between the means of the two populations being studied.

- Study Design: In a two-sample t-test, data is collected from two separate and independent groups. The mean of one group is compared to the mean of the other group to evaluate if there is a statistically significant difference between them.
- Assumptions: The assumptions of a two-sample t-test include random and independent sampling from each group, normality of the population distributions (or sufficiently large sample

sizes), and equal variances between the groups (unless using a modified version of the t-test that assumes unequal variances).

In summary, a one-sample t-test compares the mean of a single sample to a known or hypothesized value, while a two-sample t-test compares the means of two independent samples. The choice between the two tests depends on the research question and the design of the study.

The main difference between a one-sample t-test and a two-sample t-test is that a one-sample t-test compares the mean of a sample to a hypothesized value, while a two-sample t-test compares the means of two independent samples.

Here is a table that summarizes the key differences between the two tests:

Feature	One-sample t-test	Two-sample t-test
Number of samples	1	2
Type of comparison	Compares the mean of a sample to a hypothesized value	Compares the means of two independent samples
Assumptions	The population is normally distributed	The populations are normally distributed and the samples are independent
t-test statistic	`t = (sample mean - hypothesized value) / (standard deviation / sqrt(sample size))`	`t = (mean1 - mean2) / (spooled) / sqrt(n)`

Here are some examples of when you might use each type of t-test:

- **One-sample t-test:** You could use a one-sample t-test to test the hypothesis that the average height of American adults is 5'10".
- **Two-sample t-test:** You could use a two-sample t-test to test the hypothesis that the average IQ of men is different from the average IQ of women.

9 6. How is the t-statistic calculated?

The t-statistic is calculated in different ways depending on the type of t-test being performed. Here are the formulas for calculating the t-statistic for different t-tests:

1. One-Sample t-test: The one-sample t-test compares the mean of a single sample to a known or hypothesized value. Formula: t = (x -) / (s / √n) Where:

- t is the t-statistic
- x is the sample mean
- is the hypothesized population mean
- s is the sample standard deviation
- n is the sample size

2. Independent Samples t-test (Equal Variances): The independent samples t-test compares the means of two independent groups or samples, assuming equal variances. Formula: t = (x1 - x2) / √((s1^2 / n1) + (s2^2 / n2)) Where:

- t is the t-statistic
- x1 and x2 are the means of the two samples
- s1 and s2 are the standard deviations of the two samples
- n1 and n2 are the sample sizes of the two samples

3. Independent Samples t-test (Unequal Variances): The independent samples t-test compares the means of two independent groups or samples, allowing for unequal variances. Formula: $t = (x1 - x2) / \sqrt{((s1\hat{}2 / n1) + (s2\hat{}2 / n2))}$ Where:

- t is the t-statistic
- x1 and x2 are the means of the two samples
- s1 and s2 are the standard deviations of the two samples
- n1 and n2 are the sample sizes of the two samples Note: In the case of unequal variances, a modified version of the formula known as Welch's t-test is used, which accounts for the difference in variances.

4. Paired Samples t-test: The paired samples t-test compares the means of two related or paired samples. Formula: $t = (\bar{x}d - \ d) / (sd / \sqrt{n})$ Where:

- t is the t-statistic
- x̄d is the mean of the differences between the paired observations
- d is the hypothesized mean difference
- sd is the standard deviation of the differences
- n is the number of pairs

The t-statistic measures the difference between sample means relative to the variability within the samples. It quantifies how large the observed difference between sample means is compared to the expected variability based on the sample sizes and standard deviations. The t-statistic is then compared to critical values from the t-distribution to determine statistical significance.

10 7. What is the p-value in hypothesis testing?

In hypothesis testing, the p-value is a probability value that measures the strength of evidence against the null hypothesis. It quantifies the likelihood of obtaining the observed data or more extreme data if the null hypothesis is true.

The p-value is used to make decisions in hypothesis testing based on a predetermined significance level (alpha), which represents the threshold for rejecting the null hypothesis. Here's how the p-value is interpreted:

1. If the p-value is less than the significance level (p-value < alpha), the result is statistically significant. It indicates that the observed data is unlikely to occur by chance alone if the null hypothesis is true. In such cases, the null hypothesis is rejected in favor of the alternative hypothesis.

2. If the p-value is greater than or equal to the significance level (p-value >= alpha), the result is not statistically significant. It suggests that the observed data is likely to occur by chance even if the null hypothesis is true. In these cases, there is insufficient evidence to reject the null hypothesis.

It's important to note that the p-value does not directly measure the probability of the null hypothesis being true or false. Instead, it quantifies the evidence against the null hypothesis based on the observed data. A smaller p-value suggests stronger evidence against the null hypothesis.

The choice of significance level (alpha) is subjective and depends on the desired balance between Type I and Type II errors. Commonly used significance levels are 0.05 (5%) and 0.01 (1%).

It's crucial to interpret the p-value within the context of the specific hypothesis test and research question. It helps researchers make informed decisions about whether to reject or fail to reject the null hypothesis based on the strength of evidence provided by the data.

11 8. How do you interpret the p-value in hypothesis testing?

The interpretation of the p-value in hypothesis testing depends on the predetermined significance level (alpha) and the result of the hypothesis test. Here are three common interpretations: 1. If the p-value is less than the significance level (p-value < alpha): - Interpretation: The result is statistically significant. - Explanation: The p-value indicates that the observed data is unlikely to occur by chance alone if the null hypothesis is true. There is sufficient evidence to reject the null hypothesis in favor of the alternative hypothesis.

2. If the p-value is greater than or equal to the significance level (p-value >= alpha):

- Interpretation: The result is not statistically significant.
- Explanation: The p-value suggests that the observed data is likely to occur by chance even if the null hypothesis is true. There is insufficient evidence to reject the null hypothesis.

3. If the p-value is very small (e.g., p-value < 0.001):

- Interpretation: The result is highly statistically significant.
- Explanation: A very small p-value indicates strong evidence against the null hypothesis. The observed data is highly unlikely to occur by chance if the null hypothesis is true. It's important to note that the interpretation of the p-value should always be considered in the context of the specific hypothesis test and research question. It's not enough to solely rely on the p-value; other factors such as effect size, practical significance, and study design should also be taken into account.

Additionally, the interpretation of the p-value does not directly indicate the probability of the null hypothesis being true or false. It quantifies the strength of evidence against the null hypothesis based on the observed data. Therefore, a small p-value does not guarantee that the alternative hypothesis is true, and a large p-value does not prove that the null hypothesis is true. Overall, the interpretation of the p-value helps researchers make informed decisions about whether to reject or fail to reject the null hypothesis based on the strength of evidence provided by the data.

11.1 Practical Implemetation Of P-Value...

[ ]:
```
import scipy.stats as stats

# Statistical data
sample_data = [1.2, 1.5, 2.1, 1.8, 1.6, 1.3, 1.7, 1.9, 2.2, 2.4]
population_mean = 2.0

# Perform a hypothesis test (e.g., one-sample t-test)
t_statistic, p_value = stats.ttest_1samp(sample_data, population_mean)
```

```python
# Output the results
print("T-statistic:", t_statistic)
print("P-value:", p_value)
```

```
T-statistic: -1.8703377008862208
P-value: 0.09424784185146451
```

12 9. What is the significance level (alpha) in hypothesis testing?

The significance level (alpha) in hypothesis testing is the predetermined threshold used to make decisions about rejecting or failing to reject the null hypothesis. It represents the maximum probability of committing a Type I error, which is the error of rejecting the null hypothesis when it is actually true.

Commonly used significance levels are 0.05 (5%) and 0.01 (1%), but the choice of alpha depends on the specific context, research field, and the consequences of Type I and Type II errors.

Here's how the significance level is used in hypothesis testing: 1. Hypotheses formulation: The null hypothesis (H0) and alternative hypothesis (Ha) are defined based on the research question. The null hypothesis typically represents the absence of an effect or relationship, while the alternative hypothesis represents the presence of an effect or relationship.

2. Test statistic calculation: The test statistic, such as a t-statistic or z-statistic, is calculated based on the sample data and the chosen hypothesis test.

3. Comparison with critical value: The critical value(s) corresponding to the chosen significance level (alpha) and the specific test are determined from the reference distribution (e.g., t-distribution, z-distribution). These critical values determine the boundary for rejecting the null hypothesis.

4. Decision making: The test statistic is compared to the critical value(s) to make a decision:

- If the test statistic falls in the rejection region (beyond the critical value), the null hypothesis is rejected, indicating that the result is statistically significant at the chosen significance level.
- If the test statistic falls in the non-rejection region (within the critical value), the null hypothesis is not rejected, suggesting that the result is not statistically significant at the chosen significance level.

Choosing the appropriate significance level involves considering the trade-off between Type I and Type II errors. A smaller alpha (e.g., 0.01) reduces the probability of Type I error but increases the likelihood of Type II error (failing to reject the null hypothesis when it is false). A larger alpha (e.g., 0.05) increases the probability of Type I error but decreases the likelihood of Type II error.

The significance level is a critical aspect of hypothesis testing as it determines the stringency of evidence required to reject the null hypothesis. It helps researchers establish a clear threshold for statistical significance and make informed decisions based on the strength of evidence provided by the data.

13 10. What is the difference between a one-tailed and a two-tailed test?

The difference between a one-tailed test and a two-tailed test lies in the directionality of the hypothesis and the way statistical significance is assessed. Here's a breakdown of each test: One-Tailed Test:

- Hypothesis Direction: In a one-tailed test, the alternative hypothesis (Ha) is formulated to specifically test for a difference or relationship in one direction. It states that there is an effect or difference in a specific direction.
- Statistical Significance: The p-value in a one-tailed test is calculated by considering the probability of observing a test statistic as extreme as the one obtained, only in the specified direction of the alternative hypothesis.
- Research Question Example: "The new treatment is expected to improve the mean score significantly."

Two-Tailed Test: - Hypothesis Direction: In a two-tailed test, the alternative hypothesis (Ha) is formulated to test for a difference or relationship in both directions. It states that there is a general effect or difference, without specifying a particular direction.

- Statistical Significance: The p-value in a two-tailed test is calculated by considering the probability of observing a test statistic as extreme as the one obtained, in either direction (both tails) of the null hypothesis.
- Research Question Example: "The new treatment is expected to have a different mean score compared to the control group."

The choice between a one-tailed and a two-tailed test depends on the research question and prior knowledge or expectations about the direction of the effect. Consider the following factors when deciding which test to use: 1. Directionality: If there is a specific reason or expectation to test for an effect in one direction, a one-tailed test is appropriate. For example, if previous research suggests that a new treatment would only increase scores, a one-tailed test can focus on that specific direction.

2. Non-Directionality: If there is no prior expectation or if the research question is more general, a two-tailed test is suitable. This allows for the possibility of an effect in either direction.
3. Sample Size: A one-tailed test has more power to detect an effect in a specific direction compared to a two-tailed test, given the same sample size. However, it may be less appropriate if there is a possibility of an effect in the opposite direction.

It's important to note that the choice between one-tailed and two-tailed tests should be made before analyzing the data to avoid biasing the results. It is also essential to justify the choice based on prior knowledge, theory, or logical reasoning.

14 C. A/B Testing:

A/B testing, also known as split testing, is a method of comparing two versions of a webpage or app against each other to determine which one performs better. In A/B testing, two versions of a page are shown to different segments of website visitors at the same time. The goal is to determine which version of the page results in the desired outcome, such as increased conversions or decreased bounce rates.

A/B testing is a type of experimental design that can be used to test different hypotheses about how users interact with a website or app. For example, you could use A/B testing to test the hypothesis that a new design for your homepage will increase conversions.

To run an A/B test, you will need to create two versions of the page that you want to test. These versions should be as similar as possible, with the only difference being the element that you are testing. For example, if you are testing the hypothesis that a new design for your homepage will increase conversions, you would create two versions of your homepage, one with the old design and one with the new design.

Once you have created the two versions of the page, you will need to create a traffic split. This means that you will need to decide how many visitors will see each version of the page. For example, you could decide to split the traffic 50/50, so that half of the visitors see the old design and half of the visitors see the new design.

Once you have created the traffic split, you will need to track the results of the test. This means that you will need to track the number of conversions for each version of the page. You can use Google Analytics or another analytics tool to track the results of the test.

Once you have tracked the results of the test, you will need to analyze the data to determine which version of the page performed better. The version of the page that performed better is the one that you should implement on your website or app.

A/B testing is a powerful tool that can be used to improve the performance of your website or app. By testing different versions of your pages, you can determine which elements are most effective and make changes to improve your results.

Here are some of the benefits of A/B testing:

- It can help you to improve the performance of your website or app.
- It can help you to identify the elements that are most effective.
- It can help you to make changes to improve your results.
- It can help you to save time and money.

If you are looking to improve the performance of your website or app, then A/B testing is a tool that you should consider using.

15 11. What is A/B testing and why is it important?

A/B testing, also known as split testing or bucket testing, is a controlled experiment method used to compare two versions of a webpage, application, marketing campaign, or any other product or feature. It helps determine which version performs better in terms of user behavior, conversion rates, click-through rates, or other key performance indicators (KPIs).

Here's why A/B testing is important: 1. Data-Driven Decision Making: A/B testing allows companies to make data-driven decisions rather than relying on assumptions or intuition. By testing different versions of a product or feature, organizations can gather empirical evidence on which design, content, or functionality resonates best with their target audience.

2. Optimization and Improvement: A/B testing helps optimize and improve products, user experiences, and marketing strategies. By systematically testing different variations, companies can identify and implement changes that lead to better performance, increased conversions, higher engagement, or other desired outcomes.
3. User Experience Enhancement: A/B testing enables organizations to enhance the user experience by identifying and implementing changes that resonate with users. It helps understand user preferences, behavior patterns, and pain points, leading to iterative improvements that drive user satisfaction and loyalty.
4. Mitigating Risks: A/B testing minimizes the risks associated with implementing major changes or new features. By testing variations with a subset of users, organizations can assess the impact before rolling out changes to the entire user base. This helps identify potential issues, validate hypotheses, and avoid costly mistakes.
5. Continuous Improvement: A/B testing fosters a culture of continuous improvement within organizations. It encourages teams to test hypotheses, learn from results, and iterate on designs and strategies. This iterative process helps companies stay agile, adapt to changing market conditions, and maintain a competitive edge.
6. Cost-Effective Solution: A/B testing offers a cost-effective approach to optimize and validate changes. It allows companies to allocate resources efficiently by focusing on variations that demonstrate a statistically significant improvement, rather than investing in changes that may not yield desired results.

Overall, A/B testing plays a vital role in driving data-driven decision making, optimizing user experiences, and continuous improvement. It allows organizations to test, measure, and refine their products and strategies based on empirical evidence, leading to improved performance, user satisfaction, and business success.

16 12. How do you design an A/B test?

Designing an A/B test involves several key steps to ensure a well-structured and reliable experiment. Here's an example of how to design an A/B test:

1. Identify the Objective: Clearly define the goal of your A/B test. For example, let's say you run an e-commerce website, and your objective is to increase the conversion rate for a specific product page.
2. Choose the Variable to Test: Determine the specific element or variation you want to test. In this case, you might choose to test different variations of the product page's call-to-action (CTA) button.
3. Define Hypotheses: Formulate clear and specific hypotheses for each variation. For example, "Hypothesis A: Changing the CTA button color from blue to green will increase the conversion rate," and "Hypothesis B: Changing the CTA button text from 'Buy Now' to 'Shop Now' will increase the conversion rate."
4. Define Sample Size: Determine the required sample size to achieve statistical power and significance. This involves considering factors such as desired effect size, statistical confidence level, and expected variability. Use sample size calculators or statistical tools to determine the appropriate sample size for your test.

5. Split Test Groups: Randomly assign users to two groups: the control group and the variation group. The control group experiences the existing version (baseline), while the variation group experiences the modified version (variation). Ensure that the groups are comparable and representative of your target audience.
6. Implement Tracking: Set up appropriate tracking mechanisms to collect relevant data and metrics. This may involve using tools like Google Analytics or other analytics platforms to capture user interactions, conversion events, or other key metrics related to your objective.
7. Run the Experiment: Launch the A/B test and expose each group to their respective variations. Ensure that the test runs for a sufficient duration to gather an adequate sample size and to account for potential temporal effects.
8. Analyze Results: Analyze the collected data to compare the performance of the control and variation groups. Use statistical analysis techniques, such as hypothesis testing or confidence intervals, to assess the statistical significance of the observed differences.
9. Draw Conclusions: Based on the analysis, evaluate whether the observed differences in performance are statistically significant and align with the hypotheses. Determine the winning variation, if any, and draw conclusions regarding the impact on the conversion rate.
10. Implement and Monitor: If the A/B test results indicate a significant improvement in the variation group, implement the winning variation as the new default. Continuously monitor the performance and gather user feedback to inform further optimizations.

Remember, it is crucial to adhere to ethical guidelines, ensure proper statistical methods, and avoid biases during the A/B testing process. Additionally, documenting the entire process, including the rationale behind design decisions and outcomes, is valuable for future reference and knowledge sharing.

17 13. What is the null hypothesis in A/B testing?

In A/B testing, the null hypothesis (H0) represents the assumption that there is no significant difference or effect between the control group (A) and the variation group (B). It assumes that any observed differences in the test results are due to random chance or sampling variability. In the context of A/B testing, the null hypothesis typically states that the two variations being compared have the same conversion rate, click-through rate, user engagement, or any other metric of interest. The null hypothesis denies the presence of any meaningful difference between the groups.

For example, let's consider an A/B test where you are comparing two versions of a website's landing page: the control version and a variation with a modified headline. The null hypothesis in this case could be: "The modification to the headline has no significant impact on the click-through rate."

By assuming the null hypothesis, you are setting up a scenario where any observed differences between the control and variation groups are considered to be solely due to random chance. The purpose of the A/B test is to gather evidence to either support or reject the null hypothesis based on the observed data.

In hypothesis testing, the aim is to statistically test the null hypothesis against the alternative hypothesis (Ha), which represents the claim or hypothesis that there is a significant difference or effect between the groups. The goal is to gather enough evidence to reject the null hypothesis in favor of the alternative hypothesis, indicating that the observed differences are not due to chance but are meaningful and significant.

The null hypothesis serves as the default assumption in A/B testing and provides a benchmark against which the alternative hypothesis is compared. The test results are interpreted by evaluating the statistical evidence to either support or reject the null hypothesis, helping make informed decisions about which variation performs better or whether further optimization is needed.

18 14. How do you calculate statistical power in A/B testing?

Statistical power in A/B testing is the probability of correctly rejecting the null hypothesis (H0) when it is false. In other words, it measures the sensitivity of a statistical test to detect a true effect or difference between the control and variation groups. A higher statistical power indicates a greater likelihood of detecting a significant difference if it truly exists. To calculate the statistical power in A/B testing, you need to consider several factors: 1. Effect Size: The effect size represents the magnitude of the difference or effect you expect to observe between the control and variation groups. It is typically expressed as the standardized difference, such as Cohen's d or the standardized mean difference. The effect size should be estimated based on prior knowledge, pilot studies, or practical significance.

2. Sample Size: The sample size refers to the number of observations or participants in each group (control and variation). A larger sample size generally leads to greater statistical power. The sample size affects both the variability (standard deviation) and the precision of the estimate.
3. Significance Level: The significance level (alpha) is the threshold used to determine statistical significance. It is typically set to 0.05 (5%) or 0.01 (1%). The significance level affects the critical values used in hypothesis testing.
4. Variability: The variability or standard deviation of the outcome variable in the population impacts the statistical power. A smaller standard deviation increases the power of the test, making it easier to detect a significant difference.
5. Type I and Type II Errors: The statistical power is inversely related to the Type II error rate (beta), which is the probability of failing to reject the null hypothesis when it is false. The power is equal to 1 minus the Type II error rate (1 - beta). It is also related to the Type I error rate (alpha), as lowering the alpha level increases the power but also increases the likelihood of committing a Type I error.

To calculate the statistical power, you can use statistical power analysis methods or power calculators. These methods typically require information about the effect size, sample size, significance level, and variability. By adjusting the input values, such as effect size and sample size, you can determine the required sample size to achieve a desired power level or assess the power given a specific sample size.

It's important to note that achieving high statistical power is desirable, as it increases the chances of detecting meaningful effects. However, it often requires larger sample sizes. Balancing the power, sample size, and other practical considerations is crucial in designing an A/B test to ensure reliable results.

19 15. How do you calculate sample size for an A/B test?

Calculating the sample size for an A/B test involves considering several factors, such as the desired level of statistical power, significance level, expected effect size, and variability. There are various methods and formulas available to estimate the sample size. Here's an overview of one commonly used approach: 1. Determine the Statistical Power: Decide on the desired level of statistical power, which represents the probability of correctly detecting a true effect or difference if it exists. Commonly used power levels are 0.80 (80%) or 0.90 (90%). Higher power levels require larger sample sizes.

2. Choose the Significance Level: Set the desired significance level (alpha), which is the threshold for rejecting the null hypothesis. The most common values are 0.05 (5%) and 0.01 (1%).

3. Estimate the Expected Effect Size: Based on prior knowledge, pilot studies, or assumptions, estimate the effect size you expect to observe between the control and variation groups. This can be expressed as a standardized difference, such as Cohen's d or the standardized mean difference.

4. Assess Variability: Estimate the expected variability or standard deviation of the outcome variable in the population. This can be based on previous data or expert knowledge.

5. Select a Sample Size Calculation Method: Choose an appropriate sample size calculation method based on the study design and assumptions. Commonly used methods include the t-test for means, chi-square test for proportions, or regression-based approaches.

6. Perform Sample Size Calculation: Use a sample size calculator or statistical software that supports sample size calculations. Input the desired power level, significance level, expected effect size, and variability to calculate the required sample size.

7. Adjust for Practical Considerations: Consider any practical constraints, such as budget, time, or feasibility, that may impact the ability to achieve the calculated sample size. Balance the desired power level with practical limitations.

It's important to note that sample size calculations are based on assumptions and estimations, and the actual results may vary. It's always a good practice to monitor the progress of the A/B test and make adjustments if necessary.

Additionally, it's crucial to consider the ethical implications of sample size determination, ensuring that the sample size is appropriate for obtaining reliable and meaningful results without unnecessarily exposing individuals to experimental conditions.

20 16. What is the difference between conversion rate and click-through rate?

The difference between conversion rate and click-through rate lies in the specific actions they measure and the context in which they are commonly used. Here's a breakdown of each metric: Conversion Rate:

- Definition: Conversion rate is a metric that measures the percentage of users or visitors who take a desired action or complete a predefined goal. It quantifies the effectiveness of a specific conversion event, such as making a purchase, signing up for a service, filling out a form, or any other action that aligns with the objective of a website, landing page, or marketing campaign.

- Calculation: The conversion rate is calculated by dividing the number of conversions by the total number of visitors or users, and then multiplying by 100 to express it as a percentage.
- Importance: Conversion rate is crucial for businesses and marketers as it indicates the effectiveness of their efforts in converting visitors into customers or achieving desired outcomes.

It helps evaluate the performance of marketing campaigns, optimize user experiences, and identify areas for improvement. Click-Through Rate (CTR): - Definition: Click-through rate measures the percentage of users or visitors who click on a specific link, ad, or call-to-action compared to the total number of impressions or views. It primarily focuses on the rate of engagement or interaction with a particular element, such as an email, search result, banner ad, or social media post.

- Calculation: CTR is calculated by dividing the number of clicks by the number of impressions or views, and then multiplying by 100 to express it as a percentage.
- Importance: CTR is commonly used in online advertising, email marketing, search engine optimization (SEO), and other digital marketing contexts. It helps assess the performance and relevance of ad campaigns, determine the effectiveness of headlines or creative elements, and make informed decisions regarding ad placements and targeting strategies.

In summary, conversion rate measures the percentage of users who complete a specific goal or action, such as making a purchase, signing up, or completing a form. Click-through rate, on the other hand, measures the percentage of users who click on a particular link or element compared to the total number of impressions or views. Conversion rate focuses on overall effectiveness and goal achievement, while click-through rate assesses engagement and interaction with specific elements or content. Both metrics are important in evaluating the performance of marketing efforts and optimizing user experiences.

21 17. How do you analyze the results of an A/B test?

Analyzing the results of an A/B test involves several steps to assess the performance and statistical significance of the variations being compared. Here's an overview of the analysis process: 1. Data Preparation: Gather the relevant data from the control and variation groups, including the metrics or KPIs measured during the test period. Ensure the data is clean, properly formatted, and ready for analysis. 2. Descriptive Statistics: Calculate descriptive statistics for each variation, such as means, medians, standard deviations, or other appropriate measures. This helps understand the central tendency and variability of the data in each group. 3. Statistical Testing: Conduct appropriate statistical tests to compare the performance of the control and variation groups. The choice of test depends on the nature of the data and the hypothesis being tested. Commonly used tests include t-tests, chi-square tests, or regression analysis. 4. Statistical Significance: Assess the statistical significance of the observed differences between the groups. Compare the p-value (probability value) obtained from the statistical test to the predetermined significance level (alpha). If the p-value is less than alpha, the results are considered statistically significant, indicating that the observed differences are unlikely to be due to chance. 5. Effect Size Calculation: Calculate the effect size to quantify the magnitude of the observed differences between the groups. This helps evaluate the practical significance or meaningfulness of the differences. Common effect size measures include Cohen's d, relative risk, or odds ratio, depending on the type of data and the analysis conducted. 6. Confidence Intervals: Calculate confidence intervals around the estimated effect sizes. This provides a range of plausible values for the true effect in the population. 7. Interpretation: Interpret the results in the context of the hypotheses tested and the research question. Consider the statistical significance, effect size, and practical implications of the observed differences. Determine whether the results support the alternative hypothesis or fail to reject the null hypothesis. 8. Reporting: Document the results, including the analysis methods, key findings, statistical significance, effect sizes, and any limitations or assumptions made during the analysis. Present the results in a clear and concise manner, with appropriate visualizations or summary statistics to facilitate understanding and communication.

It's important to note that the analysis process should be conducted rigorously, following appropriate statistical techniques and considering any assumptions or limitations. Additionally, consult with domain experts or statisticians when analyzing complex or specialized data to ensure accurate interpretation and conclusions.

22 18. What is a confidence interval in A/B testing?

In A/B testing, a confidence interval is a range of values that provides an estimate of the true effect or difference between the control and variation groups. It quantifies the uncertainty associated with the point estimate and provides a measure of the precision of the observed effect. The confidence interval is calculated based on the data collected during the A/B test and is commonly expressed with a specified level of confidence, such as 95% or 99%. The confidence level represents the proportion of intervals, constructed from repeated sampling, that are expected to contain the true population parameter.

For example, if a 95% confidence interval is calculated for the difference in conversion rates between the control and variation groups, it means that in 95% of repeated sampling, the confidence intervals would capture the true difference.

The interpretation of a confidence interval is as follows: 1. Confidence Level: The specified confidence level represents the level of confidence in which the interval is constructed. A higher confidence level (e.g., 95%) results in a wider interval, providing greater certainty but sacrificing precision. 2. Range of Plausible Values: The confidence interval provides a range of plausible values for the true effect or difference. It includes both positive and negative values, reflecting the possibility of a beneficial or detrimental effect. 3. Null Hypothesis: If the confidence interval includes zero or overlaps with zero, it suggests that the observed difference is not statistically significant. In such cases, there is insufficient evidence to reject the null hypothesis. 4. Practical Significance: The width of the confidence interval reflects the precision of the estimate. A narrower interval indicates greater precision and provides more confidence in the estimate's practical significance.

It's important to note that a confidence interval is an estimation of the true effect based on the observed data, and it does not guarantee that the true parameter falls within the interval. It provides a range of plausible values, allowing for uncertainty in the estimation process.

In A/B testing, confidence intervals are commonly used alongside point estimates, such as mean differences or conversion rate differences, to provide a more comprehensive understanding of the observed effect and its precision. They help researchers and decision-makers make informed judgments about the practical significance and reliability of the observed differences.

23 19. How do you handle multiple comparisons in A/B testing?

Handling multiple comparisons in A/B testing is important to maintain the overall statistical validity and control the risk of Type I errors. When conducting multiple comparisons, such as testing multiple variations against a control group, the likelihood of obtaining false-positive results (rejecting the null hypothesis incorrectly) increases.

Here are a few approaches to address the issue of multiple comparisons: 1. Bonferroni Correction: The Bonferroni correction is a widely used method to adjust the significance level (alpha) for multiple comparisons. It divides the desired alpha level by the number of comparisons being made. For example, if you are testing three variations against a control group and desire a significance level of 0.05, the Bonferroni-corrected alpha level would be 0.05 / 3 = 0.0167. This adjustment reduces the likelihood of making a Type I error but may increase the likelihood of Type II errors. 2. Holm-Bonferroni Method: The Holm-Bonferroni method is another approach to address multiple comparisons. It applies a step-down procedure where the p-values are sorted in ascending order, and the significance level is adjusted sequentially. The most significant p-value is compared to an adjusted alpha, and if it exceeds the threshold, the remaining p-values are not tested. If the first p-value is significant, subsequent p-values are tested against adjusted thresholds. 3. False Discovery Rate (FDR) Control: The False Discovery Rate control is a method that controls the expected proportion of false positives among all significant results. It allows for a higher number of false positives while still maintaining a controlled overall error rate. The Benjamini-Hochberg procedure is a commonly used approach to control the FDR. 4. Prioritization and Pre-registration: To avoid the issue of multiple comparisons altogether, it is advisable to plan and pre-register your hypotheses and comparisons before conducting the A/B test. Clearly define your primary hypothesis and focus on testing specific variations. This helps reduce the temptation to cherry-pick significant results from multiple comparisons.

It's important to note that adjusting for multiple comparisons reduces the chances of false positives but may increase the chances of false negatives (Type II errors). The choice of the correction method depends on the specific context, research goals, and acceptable trade-offs between Type I and Type II errors.

Additionally, transparent reporting of all comparisons made, including significant and non-significant results, is crucial to maintain scientific integrity and provide a comprehensive understanding of the A/B test outcomes.

24 20. What are some common challenges in A/B testing and how do you overcome them?

A/B testing, while a valuable tool for data-driven decision making, comes with its own set of challenges. Here are some common challenges in A/B testing and strategies to overcome them:

1. Insufficient Sample Size: Having an inadequate sample size can result in low statistical power and unreliable results. To overcome this challenge, ensure that the sample size is calculated properly before running the test. Consider factors such as effect size, desired statistical power, and significance level. If the sample size is limited, consider running the test for a longer duration or focusing on high-impact variations.

2. Selection Bias: Selection bias occurs when the assignment of users to the control and variation groups is not random or representative. To mitigate this challenge, use random assignment

methods to ensure that participants have an equal chance of being in either group. Randomization helps minimize the influence of confounding factors and makes the groups comparable.

3. Novelty Effect and Seasonality: Users may react differently to changes due to the novelty effect, where they may exhibit different behavior simply because they are experiencing something new. Additionally, external factors like seasonality can impact the test results. To address these challenges, consider running the test for a sufficiently long duration to account for these effects and ensure that the results are stable and representative.

4. Multiple Comparisons: Conducting multiple comparisons without appropriate adjustments can lead to an increased risk of false positives. Implement strategies like Bonferroni correction, Holm-Bonferroni method, or controlling the False Discovery Rate (FDR) to address this challenge. Plan and pre-register your hypotheses and comparisons to reduce the temptation to cherry-pick significant results.

5. Interpretation of Results: Interpreting the results of an A/B test requires careful consideration of statistical significance, effect size, and practical significance. Ensure that you have a clear understanding of the metrics being measured and their relevance to the business objective. Consider the context, the goals of the test, and the impact of the observed differences in making informed decisions.

6. External Factors and Confounding Variables: External factors or confounding variables can influence the test results and make it challenging to attribute the observed differences solely to the variations being tested. To address this, carefully design the test, control for known confounders, and collect additional data or consider segmentation to analyze results based on relevant factors.

7. Practical Constraints: A/B testing may face practical constraints such as limited resources, time, or technical limitations. Prioritize your tests based on impact and feasibility. Continuously iterate and learn from each test to optimize resource allocation and overcome practical challenges. By addressing these challenges proactively, A/B testing can yield reliable and actionable insights, enabling data-driven decision making and iterative improvement of products, experiences, and strategies.

24.1 Practical Implemetation Of A/B testing..

Certainly! Here's an example of a Python program that performs an A/B test:

```
import numpy as np
import scipy.stats as stats

# Control group data
control_group = [3, 5, 7, 8, 9, 5, 6, 7, 4, 5]

# Experimental group data
experimental_group = [6, 8, 9, 10, 12, 7, 8, 9, 7, 8]

# Perform the A/B test (e.g., independent two-sample t-test)
t_statistic, p_value = stats.ttest_ind(control_group, experimental_group)

# Output the results
```

```python
print("T-statistic:", t_statistic)
print("P-value:", p_value)
```

In this example, we have two groups: the control group and the experimental group. You can replace the `control_group` and `experimental_group` lists with your own sets of data.

The `ttest_ind()` function from the `scipy.stats` module is used to perform an independent two-sample t-test. It calculates the t-statistic and the p-value for the two groups. The t-statistic measures the difference in means between the control and experimental groups, while the p-value tells us the probability of obtaining a t-statistic as extreme as the observed one, assuming the null hypothesis is true (i.e., there is no difference between the means of the two groups).

Note that you may need to install the `scipy` library if it's not already installed on your system. You can install it using `pip` by running the command `pip install scipy`.

Remember to interpret the results of the A/B test carefully and consider other factors such as sample size, statistical power, and practical significance when making decisions based on the test.

[ ]:
```python
import random

def a_b_testing(n):
  """Performs A/B testing on a sample of size n.

  Args:
    n: The size of the sample.

  Returns:
    The version of the page that performed better.
  """

  version_a_conversions = 0
  version_b_conversions = 0

  for i in range(n):
    version = random.choice(["version_a", "version_b"])
    if version == "version_a":
      version_a_conversions += 1
    else:
      version_b_conversions += 1

  if version_a_conversions > version_b_conversions:
    return "version_a"
  else:
    return "version_b"

def main():
  """Runs the A/B testing program.

  Prints the version of the page that performed better.
```

```
    """

    n = 1000
    version = a_b_testing(n)
    print(f"The version that performed better is: {version}")

if __name__ == "__main__":
    main()
```

```
The version that performed better is: version_a
```

25 Thank You!