

sample statistics problems with solutions

sample statistics problems with solutions are a cornerstone for anyone looking to grasp statistical concepts and apply them in real-world scenarios. This comprehensive guide delves into a variety of statistical problems, offering clear, step-by-step solutions to illuminate fundamental principles. From basic probability and descriptive statistics to inferential techniques like hypothesis testing and regression analysis, we'll cover essential areas. Whether you're a student preparing for exams, a researcher analyzing data, or a professional seeking to enhance your analytical skills, these sample statistics problems with solutions will serve as an invaluable resource for learning and practice. We aim to demystify complex statistical methods through practical examples, making statistics more accessible and understandable.

Table of Contents

- Descriptive Statistics Problems and Solutions
- Probability and Distributions Problems and Solutions
- Inferential Statistics Problems and Solutions
- Correlation and Regression Problems and Solutions

Descriptive Statistics Problems and Solutions

Descriptive statistics is the foundation of data analysis, focusing on summarizing and organizing data to make it understandable. It involves measures of central tendency, variability, and distribution. Understanding these concepts is crucial before moving on to more complex statistical methods. These sample statistics problems with solutions will help solidify your understanding of how to describe and interpret datasets.

Calculating Mean, Median, and Mode

The mean, median, and mode are key measures of central tendency. The mean is the average of all data points, calculated by summing all values and dividing by the number of values. The median is the middle value in a dataset that has been ordered from least to greatest. If there's an even number of data points, the median is the average of the two middle values. The mode is the value that appears most frequently in the dataset. Identifying these measures provides a snapshot of the typical value within a set of data, helping to identify patterns and outliers in your sample statistics.

Problem: A small business owner recorded the number of sales each day for a week: 5, 7, 8, 5, 10, 9, 6. Calculate the mean, median, and mode for these sales figures.

Solution:

- **Mean:** $(5 + 7 + 8 + 5 + 10 + 9 + 6) / 7 = 50 / 7 \approx 7.14$. The average number of sales per day is approximately 7.14.
- **Median:** First, order the data: 5, 5, 6, 7, 8, 9, 10. The middle value is 7. The median number of sales is 7.
- **Mode:** The number 5 appears twice, more than any other number. The mode is 5.

Determining Standard Deviation and Variance

Variance and standard deviation measure the spread or dispersion of data points around the mean. Variance is the average of the squared differences from the mean. Standard deviation is the square root of the variance, providing a measure of dispersion in the same units as the original data. These metrics are vital for understanding the variability within a sample statistics set and are foundational for inferential statistics.

Problem: Using the sales data from the previous problem (5, 7, 8, 5, 10, 9, 6), calculate the variance and standard deviation. (Mean ≈ 7.14).

Solution:

1. Calculate the difference of each data point from the mean and square it:
 - $(5 - 7.14)^2 \approx 4.58$
 - $(7 - 7.14)^2 \approx 0.02$
 - $(8 - 7.14)^2 \approx 0.74$
 - $(5 - 7.14)^2 \approx 4.58$
 - $(10 - 7.14)^2 \approx 8.18$
 - $(9 - 7.14)^2 \approx 3.46$
 - $(6 - 7.14)^2 \approx 1.30$
2. Sum the squared differences: $4.58 + 0.02 + 0.74 + 4.58 + 8.18 + 3.46 + 1.30 \approx 22.86$.
3. Calculate the variance (for a sample, divide by $n-1$): $22.86 / (7 - 1) = 22.86 / 6 \approx 3.81$.
4. Calculate the standard deviation (square root of variance): $\sqrt{3.81} \approx 1.95$.

The variance is approximately 3.81, and the standard deviation is approximately 1.95 sales.

Probability and Distributions Problems and Solutions

Probability deals with the likelihood of events occurring, forming the basis for statistical inference. Understanding different probability distributions allows us to model and predict phenomena. These sample statistics problems with solutions cover fundamental probability concepts and common distributions.

Basic Probability Calculation

Basic probability is calculated as the number of favorable outcomes divided by the total number of possible outcomes. This concept is applied in numerous scenarios, from coin flips to complex event analysis.

Problem: A bag contains 5 red marbles and 3 blue marbles. If you randomly draw one marble, what is the probability of drawing a red marble?

Solution:

- Total number of marbles = 5 (red) + 3 (blue) = 8.
- Number of favorable outcomes (drawing a red marble) = 5.
- Probability of drawing a red marble = (Number of red marbles) / (Total number of marbles) = $5 / 8$.

Understanding the Normal Distribution

The normal distribution, often called the bell curve, is a continuous probability distribution characterized by its symmetrical shape. Many natural phenomena, like heights and IQ scores, follow this distribution. Key to working with it is the z-score, which indicates how many standard deviations a data point is from the mean.

Problem: The average height of adult males in a certain country is 175 cm with a standard deviation of 7 cm. If a male is randomly selected, what is the probability that his height is between 170 cm and 180 cm? Assume heights are normally distributed.

Solution:

1. Calculate the z-scores for the two heights:
 - For 170 cm: $z = (170 - 175) / 7 = -5 / 7 \approx -0.71$.
 - For 180 cm: $z = (180 - 175) / 7 = 5 / 7 \approx 0.71$.
2. Look up the probabilities associated with these z-scores in a standard normal distribution table (or use a calculator). The probability of a z-score being less than 0.71 is approximately 0.7611. The probability of a z-score being less than -0.71 is approximately 0.2389.

3. The probability of the height being between 170 cm and 180 cm is the difference between these two probabilities: $0.7611 - 0.2389 = 0.5222$.

Therefore, there is approximately a 52.22% chance that a randomly selected adult male from this country will have a height between 170 cm and 180 cm.

Inferential Statistics Problems and Solutions

Inferential statistics involves using sample data to make generalizations or predictions about a larger population. This often includes hypothesis testing, confidence intervals, and estimation. These sample statistics problems with solutions illustrate common inferential techniques.

Conducting a One-Sample Z-Test

A one-sample z-test is used to determine whether a sample mean is significantly different from a known or hypothesized population mean, assuming the population standard deviation is known and the sample is large or the population is normally distributed.

Problem: A manufacturer claims that the average lifespan of their light bulbs is 1000 hours. A random sample of 50 bulbs showed an average lifespan of 980 hours with a known population standard deviation of 80 hours. At a significance level of 0.05, is there enough evidence to reject the manufacturer's claim?

Solution:

1. **State hypotheses:**

- Null Hypothesis (H_0): $\mu = 1000$ (The average lifespan is 1000 hours).
- Alternative Hypothesis (H_1): $\mu \neq 1000$ (The average lifespan is not 1000 hours).

2. **Calculate the test statistic (z-score):**

$$\begin{aligned} z &= (\bar{x} - \mu) / (\sigma / \sqrt{n}) \\ z &= (980 - 1000) / (80 / \sqrt{50}) \\ z &= -20 / (80 / 7.07) \\ z &= -20 / 11.31 \approx -1.77 \end{aligned}$$

3. **Determine the critical value or p-value:** For a two-tailed test at $\alpha = 0.05$, the critical z-values are ± 1.96 . Alternatively, the p-value for $z = -1.77$ is approximately 0.076.

4. **Make a decision:** Since the calculated z-score (-1.77) is between -1.96 and +1.96 (or p-value $0.076 > 0.05$), we fail to reject the null hypothesis.

5. **Conclusion:** At the 0.05 significance level, there is not enough evidence to reject the manufacturer's claim that the average lifespan of their light bulbs is 1000 hours.

Constructing a Confidence Interval for a Mean

A confidence interval provides a range of values within which the true population parameter is likely to lie, with a certain level of confidence. For a mean, it's calculated using the sample mean, sample standard deviation (or population standard deviation), and a critical value from a distribution.

Problem: A researcher wants to estimate the average study time of college students. A random sample of 100 students reported an average study time of 15 hours per week with a sample standard deviation of 4 hours. Construct a 95% confidence interval for the true average study time.

Solution:

1. Identify values:

- Sample mean ($\bar{x}$) = 15 hours
- Sample standard deviation (s) = 4 hours
- Sample size (n) = 100
- Confidence level = 95%

2. Find the critical value (z): For a 95% confidence level, the z value is 1.96 (from standard normal distribution tables).

3. Calculate the margin of error (ME):

$$ME = z (s / \sqrt{n})$$

$$ME = 1.96 (4 / \sqrt{100})$$

$$ME = 1.96 (4 / 10)$$

$$ME = 1.96 \cdot 0.4 = 0.784$$

4. Construct the confidence interval:

$$\text{Confidence Interval} = \bar{x} \pm ME$$

$$\text{Confidence Interval} = 15 \pm 0.784$$

$$\text{Confidence Interval} = (14.216, 15.784)$$

We are 95% confident that the true average study time for college students is between 14.216 and 15.784 hours per week.

Correlation and Regression Problems and Solutions

Correlation measures the strength and direction of a linear relationship between two variables, while regression analysis allows us to model and predict the relationship. These sample statistics problems with solutions cover the basics of understanding and applying these techniques.

Calculating the Correlation Coefficient

The Pearson correlation coefficient (r) ranges from -1 to +1. A value of +1 indicates a perfect positive linear relationship, -1 indicates a perfect negative linear relationship, and 0 indicates no linear relationship.

Problem: Consider the following pairs of data points for hours studied and exam scores: (5, 60), (7, 75), (4, 50), (8, 85), (6, 70). Calculate the Pearson correlation coefficient.

Solution:

Calculating the Pearson correlation coefficient (r) typically involves several steps using formulas for means, standard deviations, and covariance. For a small dataset like this, it's often done using statistical software or calculators. The formula is:

$$r = \frac{\sum[(x_i - \bar{x})(y_i - \bar{y})]}{\sqrt{[\sum(x_i - \bar{x})^2 \sum(y_i - \bar{y})^2]}}$$

After performing the calculations (which are extensive to show step-by-step here but involve finding the mean for x and y , then the deviations, squared deviations, and products of deviations):

The calculated Pearson correlation coefficient (r) for this dataset is approximately 1.0.

This indicates a very strong positive linear relationship between hours studied and exam scores.

Performing Simple Linear Regression

Simple linear regression models the relationship between an independent variable (x) and a dependent variable (y) using a linear equation: $y = a + bx$, where ' a ' is the y -intercept and ' b ' is the slope.

Problem: Using the same data as the correlation problem (hours studied and exam scores), find the equation of the simple linear regression line. Predict the exam score for a student who studies for 9 hours.

Solution:

The coefficients for the regression line (slope ' b ' and intercept ' a ') are calculated using formulas involving sums of x , y , x^2 , y^2 , and xy . For this dataset:

- **Slope (b):** $b = [n\sum(xy) - \sum x \sum y] / [n\sum(x^2) - (\sum x)^2]$

- **Intercept (a):** $a = (\sum y - b \sum x) / n$

Performing these calculations yields:

- Slope (b) ≈ 10.71
- Intercept (a) ≈ -5.71

The equation of the regression line is approximately: **Exam Score = -5.71 + 10.71 (Hours Studied)**

Prediction for 9 hours of study:

$$\text{Exam Score} = -5.71 + 10.71 \times 9$$

$$\text{Exam Score} = -5.71 + 96.39$$

$$\text{Exam Score} \approx 90.68$$

A student who studies for 9 hours is predicted to score approximately 90.68 on the exam.

Frequently Asked Questions

A researcher wants to estimate the average height of adult women in a city. They collect a random sample of 100 women and find the sample mean height to be 165 cm with a sample standard deviation of 8 cm. What is the 95% confidence interval for the average height of adult women in the city?

This is a confidence interval problem for a population mean when the population standard deviation is unknown. We use the t-distribution.

1. Identify the values: Sample mean ($\bar{x}$) = 165 cm, sample standard deviation (s) = 8 cm, sample size (n) = 100, confidence level = 95%.
2. Determine the degrees of freedom: $df = n - 1 = 100 - 1 = 99$.
3. Find the critical t-value: For a 95% confidence level and $df=99$, the critical t-value ($t^{0.025}$) is approximately 1.984 (from t-tables or statistical software).
4. Calculate the standard error: $SE = s / \sqrt{n} = 8 / \sqrt{100} = 8 / 10 = 0.8$ cm.
5. Calculate the margin of error: $ME = t^{0.025} SE = 1.984 \cdot 0.8 = 1.5872$ cm.
6. Construct the confidence interval: $CI = \bar{x} \pm ME = 165 \pm 1.5872$.

Therefore, the 95% confidence interval for the average height of adult women in the city is (163.41 cm, 166.59 cm) approximately.

A company claims that their new light bulbs last an average of 15,000 hours. A sample of 50 light bulbs from a recent production run had a mean lifespan of 14,500 hours with a standard deviation of 500 hours. Perform a hypothesis test to determine if the company's claim is statistically significant at the $\alpha = 0.05$ level.

This is a one-sample z-test for a population mean, assuming the sample size is large enough or the population is normally distributed.

1. State the hypotheses:
 - Null Hypothesis (H_0): $\mu = 15,000$ hours (The company's claim is true).
 - Alternative Hypothesis (H_1): $\mu \neq 15,000$ hours (The company's claim is not true - this is a two-tailed test).
2. Set the significance level: $\alpha = 0.05$.
3. Calculate the test statistic (z-score):

- $z = (\bar{x} - \mu_0) / (\sigma / \sqrt{n})$
- Since the population standard deviation (σ) is unknown, we use the sample standard deviation (s) as an estimate, especially with a large sample size ($n=50$).
- $z = (14,500 - 15,000) / (500 / \sqrt{50}) = -500 / (500 / 7.071) \approx -500 / 70.71 \approx -7.07$.
- 4. Determine the critical values or p-value:
 - For a two-tailed test with $\alpha = 0.05$, the critical z-values are ± 1.96 .
 - Alternatively, find the p-value associated with $z = -7.07$. This p-value will be extremely small (much less than 0.0001).
- 5. Make a decision:
 - Since the calculated z-score (-7.07) is far beyond the critical values of ± 1.96 (i.e., $|-7.07| > 1.96$), or since the p-value is much less than α ($p < 0.05$), we reject the null hypothesis.

Conclusion: At the 0.05 significance level, there is sufficient evidence to conclude that the average lifespan of the company's light bulbs is significantly different from their claim of 15,000 hours.

A survey of 500 randomly selected voters found that 280 support Candidate A. What is the sample proportion of voters who support Candidate A, and what is the 99% confidence interval for the true proportion of voters who support Candidate A?

This problem involves calculating a sample proportion and its confidence interval.

Part 1: Sample Proportion

1. Identify values: Number of successes (x) = 280, sample size (n) = 500.
2. Calculate sample proportion ($\bar{p}$): $\bar{p} = x / n = 280 / 500 = 0.56$.

The sample proportion of voters who support Candidate A is 0.56 or 56%.

Part 2: 99% Confidence Interval for Population Proportion

1. Calculate the complement of the sample proportion: $\bar{q} = 1 - \bar{p} = 1 - 0.56 = 0.44$.
2. Check conditions for z-interval for proportions: $n\bar{p} \geq 10$ and $n\bar{q} \geq 10$. ($500 \cdot 0.56 = 280 \geq 10$, and $500 \cdot 0.44 = 220 \geq 10$. Conditions met).
3. Find the critical z-value: For a 99% confidence level, the critical z-value ($z^{0.005}$) is approximately 2.576.
4. Calculate the standard error of the proportion: $SE = \sqrt{(\bar{p} \bar{q}) / n} = \sqrt{(0.56 \cdot 0.44) / 500} = \sqrt{0.2464 / 500} = \sqrt{0.0004928} \approx 0.022199$.
5. Calculate the margin of error: $ME = z SE = 2.576 \cdot 0.022199 \approx 0.0572$.
6. Construct the confidence interval: $CI = \bar{p} \pm ME = 0.56 \pm 0.0572$.

Therefore, the 99% confidence interval for the true proportion of voters who support Candidate A is (0.5028, 0.6172) or (50.28%, 61.72%).

A manufacturing plant produces widgets. The quality control

department wants to know if the mean diameter of the widgets is within the specified tolerance of 5 cm. They take a sample of 20 widgets and find a sample mean diameter of 5.05 cm and a sample standard deviation of 0.1 cm. What is the appropriate statistical test to use and why?

The appropriate statistical test to use is a one-sample t-test for the mean. Here's why:

1. Comparing a sample mean to a hypothesized population mean: The goal is to determine if the sample mean diameter (5.05 cm) is significantly different from the target population mean diameter (5 cm).
2. Population standard deviation is unknown: The problem provides the sample standard deviation (0.1 cm), not the population standard deviation. When the population standard deviation is unknown, the t-distribution is used instead of the z-distribution.
3. Small sample size (potentially): While $n=20$ is not extremely small, it's often considered small enough where the t-distribution is preferred over the z-distribution if the population is assumed to be approximately normally distributed.

Therefore, a one-sample t-test is the most suitable method to assess if the observed sample mean deviates significantly from the claimed population mean.

Two independent samples of students were taken from two different schools to compare their average scores on a standardized math test. School A had 30 students with a mean score of 78 and a standard deviation of 6. School B had 35 students with a mean score of 75 and a standard deviation of 7. Perform an independent samples t-test to determine if there is a significant difference in the average scores between the two schools at the $\alpha = 0.01$ significance level.

This is an independent samples t-test to compare the means of two independent groups.

1. State the hypotheses:

Null Hypothesis (H_0): $\mu_a = \mu_e$ (The mean scores are equal).

Alternative Hypothesis (H_1): $\mu_a \neq \mu_e$ (The mean scores are different - two-tailed test).

2. Set the significance level: $\alpha = 0.01$.

3. Identify values:

School A: $n_a = 30$, $\bar{x}_a = 78$, $s_a = 6$.

School B: $n_e = 35$, $\bar{x}_e = 75$, $s_e = 7$.

4. Calculate the degrees of freedom: Since the sample sizes and standard deviations are different, we use the Welch-Satterthwaite equation for degrees of freedom, which is complex. A common approximation for equal variances (if assumed) is $df = n_a + n_e - 2 = 30 + 35 - 2 = 63$. However, assuming unequal variances is safer and will result in a more conservative test (lower df).

Using statistical software or a calculator for Welch's t-test, the degrees of freedom are approximately 61.3.

5. Calculate the test statistic (t-score): Assuming unequal variances (more robust):

$$t = [(\bar{x}_a - \bar{x}_e) - 0] / \sqrt{[(s_a^2 / n_a) + (s_e^2 / n_e)]}$$

$$t = [(78 - 75) - 0] / \sqrt{[(6^2 / 30) + (7^2 / 35)]}$$

$$t = 3 / \sqrt{[(36 / 30) + (49 / 35)]}$$

$$t = 3 / \sqrt{1.2 + 1.4}$$

$$t = 3 / \sqrt{2.6} \approx 3 / 1.612 \approx 1.861.$$

6. Determine the critical t-value or p-value:

For a two-tailed test with $\alpha = 0.01$ and $df \approx 61.3$, the critical t-values are approximately ± 2.665 . Alternatively, the p-value for $t=1.861$ with $df=61.3$ is approximately 0.067.

7. Make a decision:

Since the calculated t-value (1.861) is less extreme than the critical t-values (± 2.665), or since the p-value (0.067) is greater than α (0.01), we fail to reject the null hypothesis.

Conclusion: At the 0.01 significance level, there is not enough evidence to conclude that there is a significant difference in the average scores on the standardized math test between the two schools.

A marketing team wants to test the effectiveness of two different advertisement campaigns (Campaign A and Campaign B) on sales. They select 10 pairs of similar stores, expose one store in each pair to Campaign A and the other to Campaign B, and record the sales for a month. The sales differences (Campaign A - Campaign B) are: 100, 150, -50, 200, 75, 120, 0, 180, -20, 90. Is there a significant difference in sales between the two campaigns at the $\alpha = 0.05$ significance level?

This is a paired samples t-test because the data comes from matched pairs of stores (one exposed to Campaign A, the other to Campaign B), allowing us to control for store-specific variations.

1. State the hypotheses:

Null Hypothesis (H_0): $\mu_e = 0$ (There is no significant difference in sales between the two campaigns).

Alternative Hypothesis (H_1): $\mu_e \neq 0$ (There is a significant difference in sales between the two campaigns - two-tailed test).

2. Set the significance level: $\alpha = 0.05$.

3. Calculate the differences (d) and their mean ($\bar{d}$) and standard deviation (s_e):

Differences: 100, 150, -50, 200, 75, 120, 0, 180, -20, 90

Sum of differences = $100 + 150 - 50 + 200 + 75 + 120 + 0 + 180 - 20 + 90 = 845$

Sample size (n) = 10

Mean difference ($\bar{d}$) = Sum of differences / n = $845 / 10 = 84.5$

Calculate the sample standard deviation of the differences (s_e). This involves calculating the variance of the differences and then taking the square root.

Variance of differences (s_e^2) ≈ 7263.89

Standard deviation of differences (s_e) = $\sqrt{7263.89} \approx 85.23$

4. Calculate the test statistic (t-score):

$t = (\bar{d} - 0) / (s_e / \sqrt{n}) = 84.5 / (85.23 / \sqrt{10}) \approx 84.5 / (85.23 / 3.162) \approx 84.5 / 26.95 \approx 3.135$.

5. Determine the critical t-value or p-value:

Degrees of freedom (df) = n - 1 = 10 - 1 = 9.

For a two-tailed test with $\alpha = 0.05$ and $df = 9$, the critical t-values are ± 2.262 .

Alternatively, the p-value for $t=3.135$ with $df=9$ is approximately 0.013.

6. Make a decision:

Since the calculated t-value (3.135) is greater than the critical t-value (2.262) (i.e., $|3.135| > 2.262$), or since the p-value (0.013) is less than α (0.05), we reject the null hypothesis.

Conclusion: At the 0.05 significance level, there is sufficient evidence to conclude that there is a significant difference in sales between Campaign A and Campaign B.

A quality inspector is checking for defects in a batch of electronic components. They sampled 200 components and found 12 defective ones. Calculate the sample proportion of defective components and determine the 90% confidence interval for the true proportion of defective components.

This problem involves calculating a sample proportion and its confidence interval for a proportion.

Part 1: Sample Proportion

1. Identify values: Number of defective components (x) = 12, sample size (n) = 200.
2. Calculate sample proportion ($\bar{p}$): $\bar{p} = x / n = 12 / 200 = 0.06$.

The sample proportion of defective components is 0.06 or 6%.

Part 2: 90% Confidence Interval for Population Proportion

1. Calculate the complement of the sample proportion: $\bar{q} = 1 - \bar{p} = 1 - 0.06 = 0.94$.
2. Check conditions for z-interval for proportions: $n\bar{p} \geq 10$ and $n\bar{q} \geq 10$. ($200 \cdot 0.06 = 12 \geq 10$, and $200 \cdot 0.94 = 188 \geq 10$. Conditions met).
3. Find the critical z-value: For a 90% confidence level, the critical z-value ($z^{0.05}$) is approximately 1.645.
4. Calculate the standard error of the proportion: $SE = \sqrt{(\bar{p} \bar{q}) / n} = \sqrt{(0.06 \cdot 0.94) / 200} = \sqrt{0.0564 / 200} = \sqrt{0.000282} \approx 0.01679$.
5. Calculate the margin of error: $ME = z SE = 1.645 \cdot 0.01679 \approx 0.02761$.
6. Construct the confidence interval: $CI = \bar{p} \pm ME = 0.06 \pm 0.02761$.

Therefore, the 90% confidence interval for the true proportion of defective components is (0.03239, 0.08761) or (3.24%, 8.76%).

A study investigates the relationship between hours spent studying per week and final exam scores for a group of 50 university students. The Pearson correlation coefficient (r) between study hours and exam scores is calculated to be 0.75. Is this correlation statistically significant at the $\alpha = 0.01$ significance level?

This problem involves testing the statistical significance of a Pearson correlation coefficient.

1. State the hypotheses:

Null Hypothesis (H_0): $\rho = 0$ (There is no linear correlation between study hours and exam scores).

Alternative Hypothesis (H_1): $\rho \neq 0$ (There is a significant linear correlation between study hours and exam scores - two-tailed test).

2. Set the significance level: $\alpha = 0.01$.

3. Identify values: Sample correlation coefficient (r) = 0.75, sample size (n) = 50.

4. Calculate the test statistic (t-score for correlation):

$$t = r \sqrt{(n - 2) / (1 - r^2)}$$

$$t = 0.75 \sqrt{(50 - 2) / (1 - 0.75^2)}$$

$$t = 0.75 \sqrt{48} / \sqrt{1 - 0.5625}$$

$$t = 0.75 \cdot 6.928 / \sqrt{0.4375}$$

$$t = 5.196 / 0.6614 \approx 7.858.$$

5. Determine the critical t-value or p-value:

Degrees of freedom (df) = $n - 2 = 50 - 2 = 48$.

For a two-tailed test with $\alpha = 0.01$ and $df = 48$, the critical t-values are approximately ± 2.682 .

Alternatively, the p-value for $t = 7.858$ with $df = 48$ is extremely small (much less than 0.0001).

6. Make a decision:

Since the calculated t-value (7.858) is greater than the critical t-value (2.682) (i.e., $|7.858| > 2.682$), or since the p-value is much less than α ($p < 0.01$), we reject the null hypothesis.

Conclusion: At the 0.01 significance level, there is sufficient evidence to conclude that a statistically significant linear correlation exists between hours spent studying per week and final exam scores.

A company is launching a new product and wants to estimate the proportion of consumers who are likely to purchase it. They conduct a survey of 800 potential consumers and find that 240 express an intention to buy. What is the margin of error for a 95% confidence interval for the proportion of consumers who intend to buy?

This problem requires calculating the margin of error for a confidence interval for a population proportion.

1. Identify values: Sample size (n) = 800, number of successes (consumers intending to buy) (x) = 240, confidence level = 95%.

2. Calculate the sample proportion ($\bar{p}$): $\bar{p} = x / n = 240 / 800 = 0.30$.

3. Calculate the complement of the sample proportion ($\bar{q}$): $\bar{q} = 1 - \bar{p} = 1 - 0.30 = 0.70$.

4. Find the critical z-value: For a 95% confidence level, the critical z-value ($z^{0.025}$) is 1.96.

5. Calculate the standard error of the proportion: $SE = \sqrt{(\bar{p} \bar{q}) / n} = \sqrt{(0.30 \cdot 0.70) / 800} = \sqrt{0.21 / 800} = \sqrt{0.0002625} \approx 0.01620$.

6. Calculate the margin of error (ME): $ME = z \cdot SE = 1.96 \cdot 0.01620 \approx 0.03175$.

Therefore, the margin of error for a 95% confidence interval for the proportion of consumers who intend to buy is approximately 0.0318 or 3.18%.

Additional Resources

Here are 9 book titles related to sample statistics problems with solutions, featuring italicized titles:

1. *A First Course in Statistics: Understanding Sample Data with Worked Examples*

This introductory textbook focuses on the fundamental concepts of statistics, with a particular emphasis on how to interpret and analyze sample data. It provides clear explanations of common statistical procedures and includes a wealth of step-by-step examples that demonstrate how to solve typical problems encountered in the field. Students will find this book invaluable for building a strong foundation in statistical thinking and application.

2. *Applied Statistics for the Social Sciences: Problem Sets and Solutions*

Designed for students in social science disciplines, this book bridges the gap between statistical theory and real-world research. It presents statistical methods relevant to social research, such as regression analysis and hypothesis testing, and dedicates significant attention to practice problems. Each problem is accompanied by a detailed solution, allowing readers to check their understanding and learn from their mistakes.

3. *Statistics for Engineers: Problem-Solving with Real-World Data*

This practical guide caters to engineering students and professionals who need to apply statistical techniques to solve engineering challenges. It covers essential topics like experimental design, quality control, and data analysis using statistical software. The book is rich with case studies and practice problems derived from actual engineering scenarios, complete with comprehensive solutions.

4. *The Practice of Statistics in the Life Sciences: Exercises and Answers*

Tailored for students in biology, health sciences, and related fields, this resource offers a hands-on approach to learning statistical concepts. It highlights the importance of statistics in scientific discovery and data interpretation within the life sciences. The book features a broad range of exercises that mirror the types of statistical problems encountered in biological research, with all solutions provided.

5. *Business Statistics: From Concepts to Solutions*

This comprehensive textbook equips business students and professionals with the statistical tools necessary for informed decision-making. It covers descriptive statistics, inferential statistics, and forecasting methods relevant to the business world. The book's strength lies in its abundant application-oriented problems, each followed by a thorough solution that explains the reasoning process.

6. *Introductory Statistics: Bridging Theory and Practice Through Problems*

This accessible text aims to make statistics less intimidating for beginners by directly connecting theoretical concepts to practical problem-solving. It covers foundational statistical ideas, including probability, distributions, and hypothesis testing, in a clear and concise manner. The book's core strength is its extensive collection of solved problems that illustrate how to apply statistical knowledge effectively.

7. *Statistics for Psychologists: A Problem-Based Approach*

For psychology students, this book offers a focused approach to understanding and applying statistical methods in psychological research. It explores statistical techniques commonly used in psychology, such as t-tests, ANOVA, and chi-square tests, with an emphasis on interpretation. Each chapter includes numerous practice problems designed to solidify understanding, each with a

detailed answer and explanation.

8. Quantitative Methods for Economists: Solved Problems in Statistics

This book is designed for economics students who need to master statistical techniques for economic analysis. It covers essential quantitative methods, including regression models and time series analysis, with a focus on their application in economic contexts. The text provides a wealth of meticulously worked-out problems that demonstrate the application of statistical theory to economic data.

9. Data Analysis with Confidence: Statistical Problems and Solutions for Every Learner

This user-friendly guide is perfect for anyone looking to build confidence in their statistical analysis skills. It breaks down complex statistical concepts into manageable steps and offers a variety of problems that cover a wide range of applications. The book's strength lies in its detailed solutions, which not only provide the correct answer but also explain the underlying logic and methods used.

Sample Statistics Problems With Solutions

Sample Statistics Problems With Solutions

Related Articles

- [scale factor word problems worksheet](#)
- [san juan puerto rico travel guide](#)
- [samsung convection oven manual](#)

[Back to Home](#)