Self-Check Quiz - Results

Attempt 1 of 2

Written Jun 16, 2025 3:06 PM - Jun 16, 2025 4:30 PM

Attempt Score	**33 / 35 - A**
Overall Grade (Highest Attempt)	**33 / 35 - A**

Foundations of Statistics

Question 1 1 / 1 point

Which statement correctly describes the relationship between a population and a sample?

- ✓ ● A sample is used to make inferences about the population.
- ○ A population consists only of the subjects that were actually measured.
- ○ A sample includes every member of the population.
- ○ A sample and a population are always the same size.

Question 2 1 / 1 point

Which of the following statements correctly distinguishes a parameter from a statistic?

- ○ A parameter can change depending on the sample, while a statistic remains constant.
- ○ A parameter is always known, while a statistic is always unknown.
- ○ A parameter is an estimate, while a statistic is an exact value.
- ✓ ● A parameter summarizes data for an entire population, while a statistic summarizes data for a subset of the population.

Question 3 1 / 1 point

Match each of the following descriptions for datasets with the data visualization tool that would be the **most** appropriate for displaying the data for that dataset.

✓ __3__ This data visualization tool would be the **most** appropriate for visualizing the relationship between income (in US dollars) and size of residence (in square feet) for 200 individuals

✓ __2__ This data visualization tool would be the **most** appropriate tool for displaying information about the favorite color for each individual in a class of 100

✓ __1__ This data visualization tool would be the **most** appropriate for visualizing the heights (in centimeters) of 500 college athletes?

1. Histogram
2. Bar Chart (or Bar Plot)
3. Scatterplot

Question 4 1 / 1 point

Suppose $X_1, X_2, \ldots, X_n$ are independent and identically distributed (iid) random variables with mean μ. Which of the following statements is true?

✓ ● The expectation of the sample mean $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ is μ

○ The expectation of the sample mean $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ is μ/n

○ The expectation of the sum $\sum_{i=1}^{n} X_i$ is μ

○ The expectation of the sum $\sum_{i=1}^{n} X_i$ is μ/n

Question 5

1 / 1 point

Suppose $X_1, X_2, \ldots, X_n$ are independent and identicaly distributed (iid) random variables with mean μ and variance σ^2. Which of the following statements is true?

○ The variance of the sample mean $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ is σ^2

○ The variance of the sum $\sum_{i=1}^{n} X_i$ is $n^2\sigma^2$

✓ ● The variance of the sample mean $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ is σ^2/n

○ The variance of the sum $\sum_{i=1}^{n} X_i$ is σ^2/n^2

Question 6

0 / 1 point

Assume that $X_1, X_2, \ldots, X_n$ are (iid) normally distributed random variables with mean μ and variance σ^2. That is $X_i \overset{iid}{\sim} \mathcal{N}(\mu, \sigma^2)$. Which of the following statements is true?

➡ ○ The distribution of $\sum_{i=1}^{n} X_i$ is $\mathcal{N}(n\mu, n\sigma^2)$

✗ ● The distribution of $\sum_{i=1}^{n} X_i$ is $\mathcal{N}(\mu, \sigma^2/n)$

○ The distribution of the sample mean $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ is $\mathcal{N}(n\mu, n\sigma^2)$

○ The distribution of the sample mean $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ is $\mathcal{N}(n\mu, \sigma^2/n)$

Question 7

1 / 1 point

Assume that $X_1, X_2, \ldots X_n$ are (iid) normally distributed random variables with mean μ and variance σ^2. That is $X_i \overset{iid}{\sim} \mathcal{N}(\mu, \sigma^2)$. Which of the following statements is true?

○ The distribution $\frac{\sqrt{n}(\bar{X}-\mu)}{\sigma}$ is $\mathcal{N}(0, \sigma^2/n)$

○ The distribution $\frac{\sqrt{n}(\bar{X}-\mu)}{\sigma}$ is t_n (t-distribution with n degrees of freedom)

○ The distribution $\frac{\sqrt{n}(\bar{X}-\mu)}{\sigma}$ is t_{n-1} (t-distribution with $n-1$ degrees of freedom)

✓ ● The distribution $\frac{\sqrt{n}(\bar{X}-\mu)}{\sigma}$ is $\mathcal{N}(0, 1)$

Question 8 1 / 1 point

Assume that $X_1, X_2, \ldots, X_n$ are (iid) normally distributed random variables with mean μ and variance σ^2. That is $X_i \overset{iid}{\sim} N\left(\mu, \sigma^2\right)$.

What is the distribution of the sample variance $S^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \mu)^2$ when μ is known?

○ The distribution of S^2 when μ is known is $S^2 \sim \chi^2_{n-1}$ (chi-square distribution with $n-1$ degrees of freedom).

○ The distribution of S^2 when μ is known is $S^2 \sim \chi^2_n$ (chi-square distribution with n degrees of freedom).

✓ ● The distribution of S^2 when μ is known is $nS^2/\sigma^2 \sim \chi^2_n$ (chi-square distribution with n degrees of freedom).

○ The distribution of S^2 when μ is known is $(n-1)S^2/\sigma^2 \sim \chi^2_{n-1}$ (chi-square distribution with n degrees of freedom).

Question 9 1 / 1 point

Assume that $X_1, X_2, \ldots, X_n$ are (iid) normally distributed random variables with mean μ and variance σ^2. That is $X_i \overset{iid}{\sim} N\left(\mu, \sigma^2\right)$.

What is the distribution of the sample variance $S^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \mu)^2$ when μ is **unknown** and must be estimated with $\bar{x}$?

○ The distribution of S^2 when μ is **not** known is $S^2 \sim \chi^2_n$ (chi-square distribution with n degrees of freedom).

✓ ● The distribution of S^2 when μ is **not** known is $(n-1)S^2/\sigma^2 \sim \chi^2_{n-1}$ (chi-square distribution with n degrees of freedom).

○ The distribution of S^2 when μ is **not** known is $nS^2/\sigma^2 \sim \chi^2_n$ (chi-square distribution with n degrees of freedom).

○ The distribution of S^2 when μ is **not** known is $S^2 \sim \chi^2_{n-1}$ (chi-square distribution with $n-1$ degrees of freedom).

Question 10 1 / 1 point

Match each of the following definitions with the corresponding statistical distribution.

✓ __2__ Standard normal random variable divided by the square root of an independent chi-square random variable divided by its degrees of freedom

✓ __3__ Ratio of two independent chi-squared random variables, each of which is divided by their respective degrees of freedom

✓ __1__ Square of a standard normal random variable

1. χ^2 distribution
2. t-distribution
3. F-distribution

Question 11 1 / 1 point

Assume that $X_1, X_2, \ldots X_n$ are iid random variables with mean μ and variance $\sigma^2 < \infty$, but are not necessarily normally distributed. If the sample size n is sufficiently large, what does the Central Limit Theorem (CLT) state?

○ That the quantity $\frac{\sqrt{n}(\bar{X}-\mu)}{\sigma}$ is approximately normal with mean μ and variance σ^2

○ That the quantity $\frac{n(\bar{X}-\mu)}{\sigma}$ is approximately normal with mean 0 and variance 1

○ That the quantity $\frac{\sqrt{n}(\bar{X}-\mu)}{\sigma}$ is approximately normal with mean 0 and variance σ^2

✓ ● That the quantity $\frac{\sqrt{n}(\bar{X}-\mu)}{\sigma}$ is approximately normal with mean 0 and variance 1

Question 12	1 / 1 point

Assume that $X_1, X_2, \ldots, X_n$ form a random sample of iid random variables with mean μ and variance σ^2. What is the sampling distribution of the sample mean $\bar{X}$

✓ ● The sampling distribution of the sample mean $\bar{X}$ is approximately normal with mean μ and variance σ^2/n

○ The sampling distribution of the sample mean $\bar{X}$ is approximately normal with mean μ and variance 1

○ The sampling distribution of the sample mean $\bar{X}$ is approximately normal with mean μ and variance $\sqrt{\sigma^2/n}$

○ The sampling distribution of the sample mean $\bar{X}$ is approximately normal with mean 0 and variance 1

Question 13	1 / 1 point

Assume that $X_1, X_2, \ldots, X_n$ are normally distributed iid random variables with mean μ and variance σ^2. When σ is known, what is the distribution for $\frac{\sqrt{n}(\bar{X}-\mu)}{\sigma}$?

○ $\mathcal{N}(0, \sigma^2)$

○ t_{n-1} (t-distribution with $n-1$ degrees of freedom)

✓● $\mathcal{N}(0, 1)$

○ t_n (t-distribution with n degrees of freedom)

Question 14 1 / 1 point

Assume that $X_1, X_2, \ldots, X_n$ are normally distributed iid random variables with mean μ and variance σ^2. When σ **is unknown** and must be estimated using the sample standard deviation $S = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2}$, what is the distribution for $\frac{\sqrt{n}(\bar{X}-\mu)}{S}$?

○ t_n (t-distribution with n degrees of freedom)

✓● t_{n-1} (t-distribution with $n-1$ degrees of freedom)

○ $\mathcal{N}(0, \sigma^2)$

○ $\mathcal{N}(0, 1)$

Question 15 1 / 1 point

Assume that newborn dolphins weights (in lbs) are normally distributed with unknown mean μ lbs. Previously, the average weight for newborn dolphins was μ_0 lbs, but it is believed that the average weight has decreased since then. You wish to test this hypothesis.

What are the correct null and alternative hypotheses for this test?

○ $H_0 : \bar{X} = \mu_0$ vs $H_a : \bar{X} < \mu_0$

○ $H_0 : \mu = \mu_0$ vs $H_a : \mu > \mu_0$

○ $H_0 : \bar{X} = \mu_0$ vs $H_a : \bar{X} > \mu_0$

✓● $H_0 : \mu = \mu_0$ vs $H_a : \mu < \mu_0$

Question 16 1 / 1 point

Now assume that weights (in lbs) for newborn dolphins are still normally distributed, but it is believed that the average weight for males, denoted as μ_{male}, is different than that for females, denoted as μ_{female}. You wish to test this hypothesis.

What are the correct null and alternative hypotheses for this test?

○ $H_0 : \bar{X}_{male} \neq \bar{X}_{female}$ vs $H_a : \bar{X}_{male} = \bar{X}_{female}$

✓● $H_0 : \mu_{male} = \mu_{female}$ vs $H_a : \mu_{male} \neq \mu_{female}$

○ $H_0 : \mu_{male} \neq \mu_{female}$ vs $H_a : \mu_{male} = \mu_{female}$

○ $H_0 : \bar{X}_{male} = \bar{X}_{female}$ vs $H_a : \bar{X}_{male} \neq \bar{X}_{female}$

Question 17 1 / 1 point

Assume that newborn dolphins weights (in lbs) are normally distributed with unknown mean μ lbs and **known standard deviation** σ lbs. You collect a random sample of size n and compute the sample mean $\bar{X}$ and sample standard deviation S for this sample. You wish to test the hypothesis that the true average weight of newborn dolphins is different than the hypothesized value of μ_0 lbs.

Which is the **most** appropriate test statistic for testing this hypothesis?

○ $T = \frac{(\bar{X}-\mu)}{S/n}$

○ $Z = \frac{(\bar{X}-\mu)}{\sigma/n}$

✓ ● $Z = \frac{(\bar{X}-\mu)}{\sigma/\sqrt{n}}$

○ $T = \frac{(\bar{X}-\mu)}{S/\sqrt{n}}$

Question 18 1 / 1 point

Again assume that newborn dolphins weights (in lbs) are normally distributed, but now both the mean μ and standard deviation σ are **unknown**. You collect a random sample of size n and compute the sample mean $\bar{X}$ and sample standard deviation S for this sample. You wish to test the hypothesis that the true average weight of newborn dolphins is different than the hypothesized value of μ_0 lbs.

Which is the **most** appropriate test statistic for testing this hypothesis?

✓ ● $T = \frac{(\bar{X}-\mu)}{S/\sqrt{n}}$

○ $T = \frac{(\bar{X}-\mu)}{S/n}$

○ $Z = \frac{(\bar{X}-\mu)}{\sigma/\sqrt{n}}$

○ $Z = \frac{(\bar{X}-\mu)}{\sigma/n}$

Question 19 1 / 1 point

Now assume that weights (in lbs) for newborn dolphins are still normally distributed, but it is believed that the average weight for males, denoted as μ_{male}, is different than that for females, denoted as μ_{female}. Furthermore, assume that the standard deviation

σ for both males and females **is the same and is known**. You collect a random sample of n male newborn dolphins and n female newborn dolphins and compute the sample means $\bar{X}_{male}$, $\bar{X}_{female}$ and sample standard deviation S. You wish to use these to test the hypothesis that the true average weights for male and female newborn dolphins are different.

Which is the **most** appropriate test statistic for testing this hypothesis?

✓ ● $Z = \frac{\bar{X}_{male} - \bar{X}_{female}}{\sigma\sqrt{2/n}}$

○ $Z = \frac{\bar{X}_{male} - \bar{X}_{female}}{\sigma\sqrt{1/n}}$

○ $T = \frac{\bar{X}_{male} - \bar{X}_{female}}{S\sqrt{1/n}}$

○ $T = \frac{\bar{X}_{male} - \bar{X}_{female}}{S\sqrt{2/n}}$

Question 20 1 / 1 point

Now assume that weights (in lbs) for newborn dolphins are still normally distributed, but it is believed that the average weight for males, denoted as μ_{male}, is different than that for females, denoted as μ_{female}. Furthermore, assume that the standard deviation σ for both males and females **is the same and is unknown**. You collect a random sample of n male newborn dolphins and n female newborn dolphins and compute the sample means $\bar{X}_{male}$, $\bar{X}_{female}$ and sample standard deviation S. You wish to use these to test the hypothesis that the true average weights for male and female newborn dolphins are different.

Which is the **most** appropriate test statistic for testing this hypothesis?

✓ ● $T = \frac{\bar{X}_{male} - \bar{X}_{female}}{S\sqrt{2/n}}$

○ $Z = \frac{\bar{X}_{male} - \bar{X}_{female}}{\sigma\sqrt{2/n}}$

○ $T = \frac{\bar{X}_{male} - \bar{X}_{female}}{S\sqrt{1/n}}$

○ $Z = \frac{\bar{X}_{male} - \bar{X}_{female}}{\sigma\sqrt{1/n}}$

Question 21 1 / 1 point

You are interested in testing the following hypothesis:

$H_0 : \mu = \mu_0$ vs $H_a : \mu \neq \mu_0$

You are provided with a test statistic Z, which follows a normal distribution with mean 0 and variance 1 when the H_0 is true (i.e., $\mu = \mu_0$). You collect a random sample of observations and compute a value for the test statistic of $z = 2.10$. Compute a P-value for this hypothesis test. Round your to **four** decimal places.

○ 0.0714

○ Cannot be computed

✓ ● 0.0357

○ 0.0179

Question 22 1 / 1 point

You are interested in testing the following hypothesis:

$H_0 : \mu = \mu_0$

$H_a : \mu \neq \mu_0$

You are now provided with a test statistic T, which follows a t-distribution with 15 degrees of freedom when H_0 is true (i.e., $\mu = \mu_0$). You collect a random sample of observations and compute a value for the test statistic of $t = 1.8$. Based on this information and using a significance level of $\alpha = 0.05$, what decision would you make regarding H_0?

✓ ● The P-value corresponding to $t = 1.8$ is above 0.05, therefore we fail to reject H_0

○ The P-value corresponding to $t = 1.8$ is above 0.05, therefore we reject H_0

○ The P-value corresponding to $t = 1.8$ is below 0.05, therefore we reject H_0

○ The P-value corresponding to $t = 1.8$ is below 0.05, therefore we fail to reject H_0

Question 23 1 / 1 point

Assume that $X_1, X_2, \ldots, X_n$ are normally distributed iid random variables with mean μ and variance σ^2, both of which are **unknown**. You collect a random sample of $n = 21$ observations and compute the sample mean $\bar{x}$ and sample standard deviation s for this sample and wish to use these to compute a 99% confidence interval for μ.

Which of the following answers gives the correct confidence interval for μ in this scenario?

○ $\left(\bar{x} - 2.576 \times \frac{s}{\sqrt{21}}, \bar{x} + 2.576 \times \frac{s}{\sqrt{21}}\right)$

✓ ● $\left(\bar{x} - 2.845 \times \frac{s}{\sqrt{21}}, \bar{x} + 2.845 \times \frac{s}{\sqrt{21}}\right)$

○ $\left(\bar{x} - 1.960 \times \frac{s}{\sqrt{21}}, \bar{x} + 1.960 \times \frac{s}{\sqrt{21}}\right)$

○ $\left(\bar{x} - 2.086 \times \frac{s}{\sqrt{21}}, \bar{x} + 2.086 \times \frac{s}{\sqrt{21}}\right)$

Question 24 0 / 1 point

Assume that $X_1, X_2, \ldots, X_n$ are normally distributed iid random variables with mean μ_X and that $Y_1, Y_2, \ldots, Y_n$ are normally distributed iid random variables with mean μ_Y. Assume that the variance for all observations is σ^2 and that it is **unknown**. Let

$n = 12$, meaning that there are 24 observations in total, 12 for X and 12 for Y. Using these, you compute the sample means $\bar{x}$, $\bar{y}$ and a single estimate of the sample standard deviation s. You wish to construct a 95% confidence interval for the difference of means $(\mu_X - \mu_Y)$.

Which of the following answers gives the correct confidence interval for $\mu_X - \mu_Y$ in this scenario?

✗ ● $\left((\bar{x}-\bar{y}) - 2.074 \times s\sqrt{\frac{1}{12}}, (\bar{x}-\bar{y}) + 2.074 \times s\sqrt{\frac{1}{12}}\right)$

○ $\left((\bar{x}-\bar{y}) - 1.960 \times s\sqrt{\frac{2}{12}}, (\bar{x}-\bar{y}) + 1.960 \times s\sqrt{\frac{2}{12}}\right)$

○ $\left((\bar{x}-\bar{y}) - 1.960 \times s\sqrt{\frac{1}{12}}, (\bar{x}-\bar{y}) + 1.960 \times s\sqrt{\frac{1}{12}}\right)$

➡ ○ $\left((\bar{x}-\bar{y}) - 2.074 \times s\sqrt{\frac{2}{12}}, (\bar{x}-\bar{y}) + 2.074 \times s\sqrt{\frac{2}{12}}\right)$

Question 25

1 / 1 point

Which of the following options gives the best descriptions of Type I Error and Power in the context of hypothesis testing?

✓ ● Type I Error: Rejecting the null hypothesis when it is true.

Power: Rejecting the null hypothesis when it is not true.

○ Type I Error: Failing to reject the null hypothesis when it is not true.

Power: Rejecting the null hypothesis when it is not true.

○ Type I Error: Rejecting the null hypothesis when it is not true.

Power: Rejecting the null hypothesis when it is true.

○ Type I Error: Rejecting the null hypothesis when it is true.

Power: Failing to reject the null hypothesis when it is not true.

Linear Regression Analysis

Question 26 1 / 1 point

Let $Y_1, Y_2, \ldots, Y_n$ represent the weights of n individuals and let $X_1, X_2, \ldots, X_n$ represent the corresponding heights for each individual. You would like to understand the relationship between height and weight using a simple linear regression model. Which of the following is the **most appropriate** way to express this model?

○ $Y_i = \beta_1 X_i + \epsilon_i$ where $\epsilon_i \overset{iid}{\sim} N(0, \sigma^2)$

○ $Y_i = (\beta_0 + \beta_1) X_i + \epsilon_i$ where $\epsilon_i \overset{iid}{\sim} N(0, \sigma^2)$

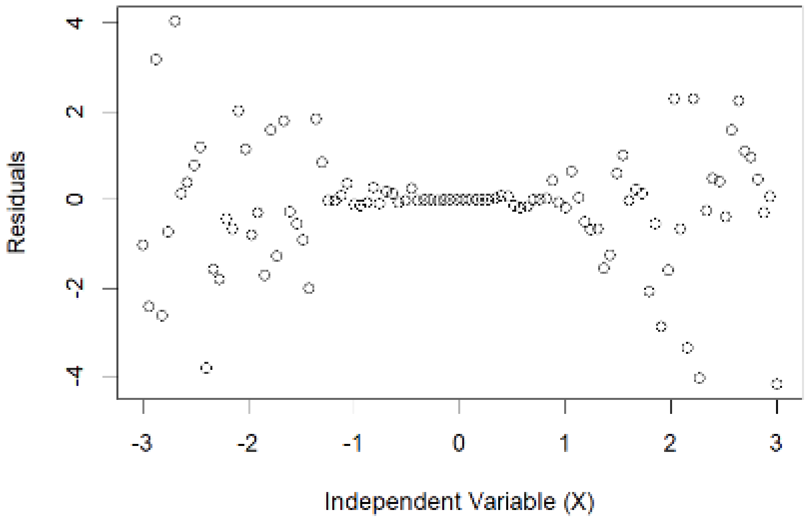
✓ ● $Y_i = \beta_0 + \beta_1 X_i + \epsilon_i$ where $\epsilon_i \overset{iid}{\sim} N(0, \sigma^2)$

○ $Y_i = \beta_0 + \beta_1 X_i + \epsilon_i$ where $\epsilon_i \sim N(0, \sigma_i^2)$

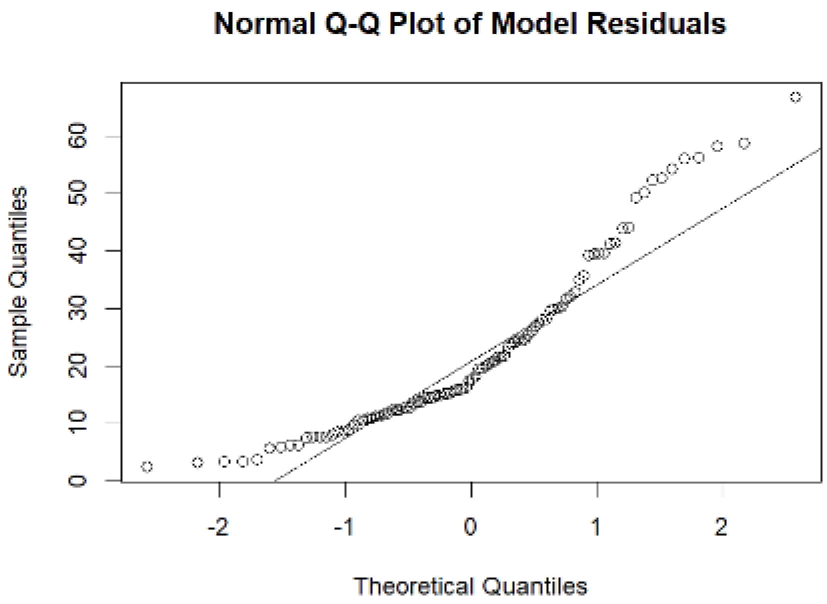
Question 27 1 / 1 point

Select the plot that most clearly indicates there is a violation of the assumption of linearity for the simple linear regression model.

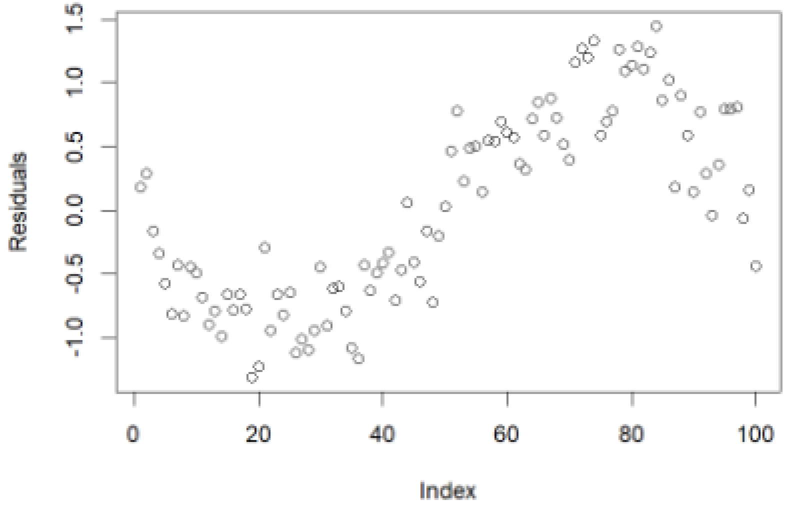
○



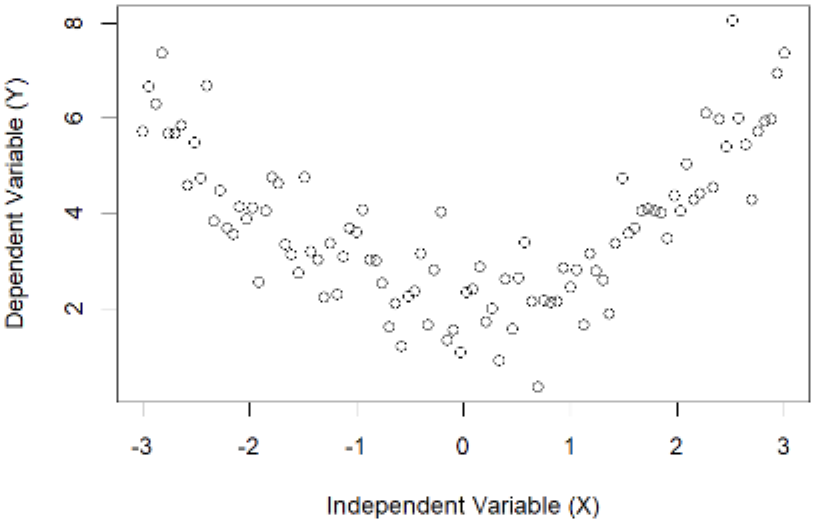
○


Normal Q-Q Plot of Model Residuals
Sample Quantiles
0
10
20
30
40
50
60
-2
-1
0
1
2
Theoretical Quantiles

○


Residuals
-1.0
-0.5
0.0
0.5
1.0
1.5
0
20
40
60
80
100
Index

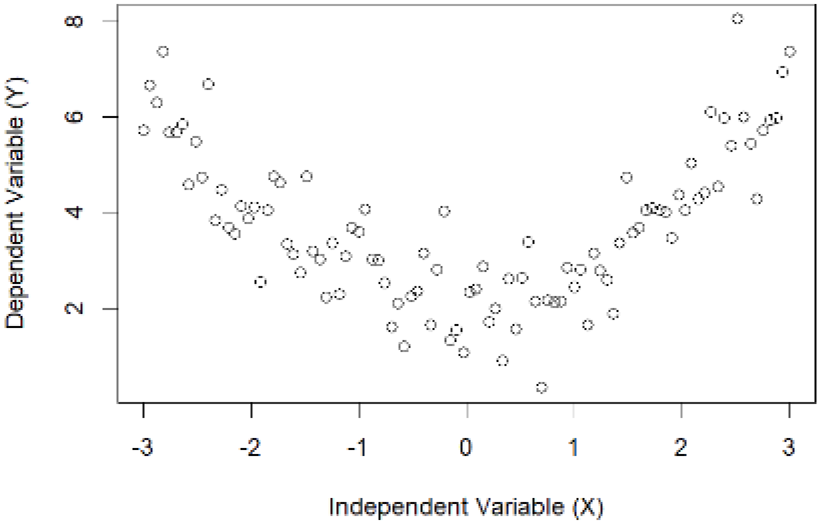
✓ ●



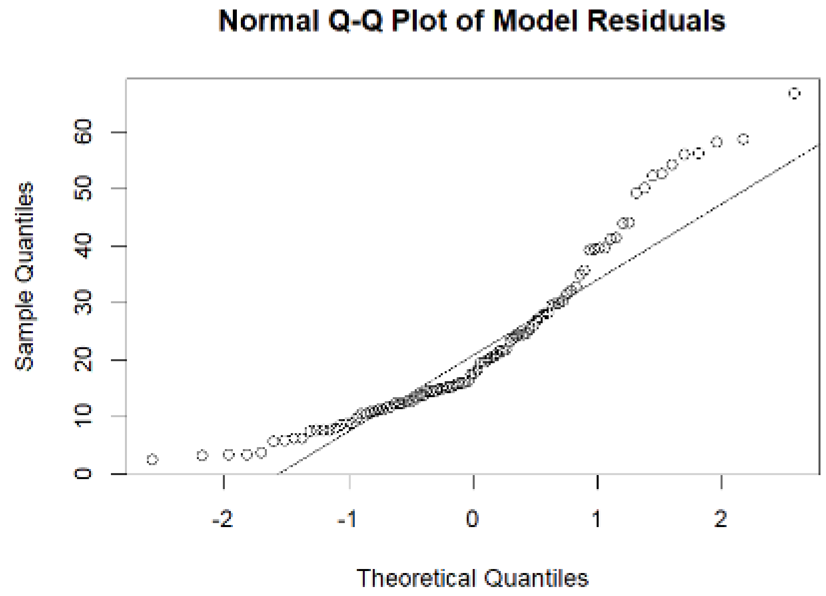
Question 28 1 / 1 point

Select the plot that most clearly indicates there is a violation of the assumption of normality for the simple linear regression model.

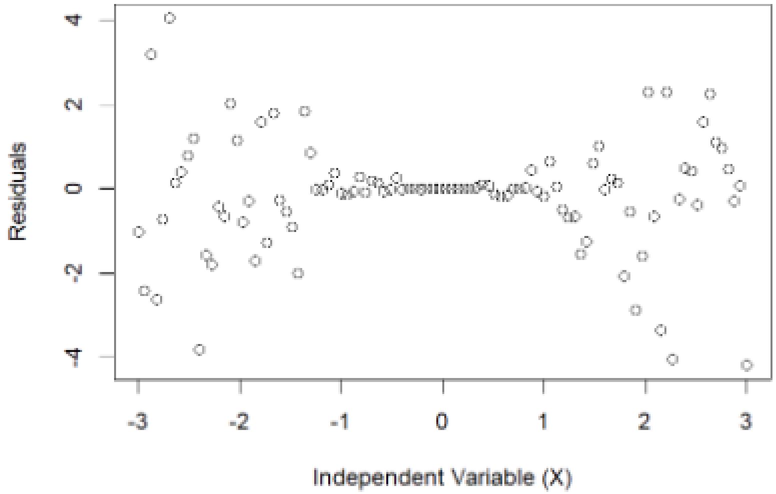
○



✓ ●



○



Question 29 1 / 1 point

Let b_0 and b_1 represent estimates of the model coefficients, β_0 and β_1, for a simple linear regression model. Let y_i represent an observed response (dependent variable) for an individual and let x_i represent the corresponding value for an observed predictor (independent variable). Which of the following equations can be used to compute the residual e_i?

○ $e_i = b_0 + b_1 x_i$

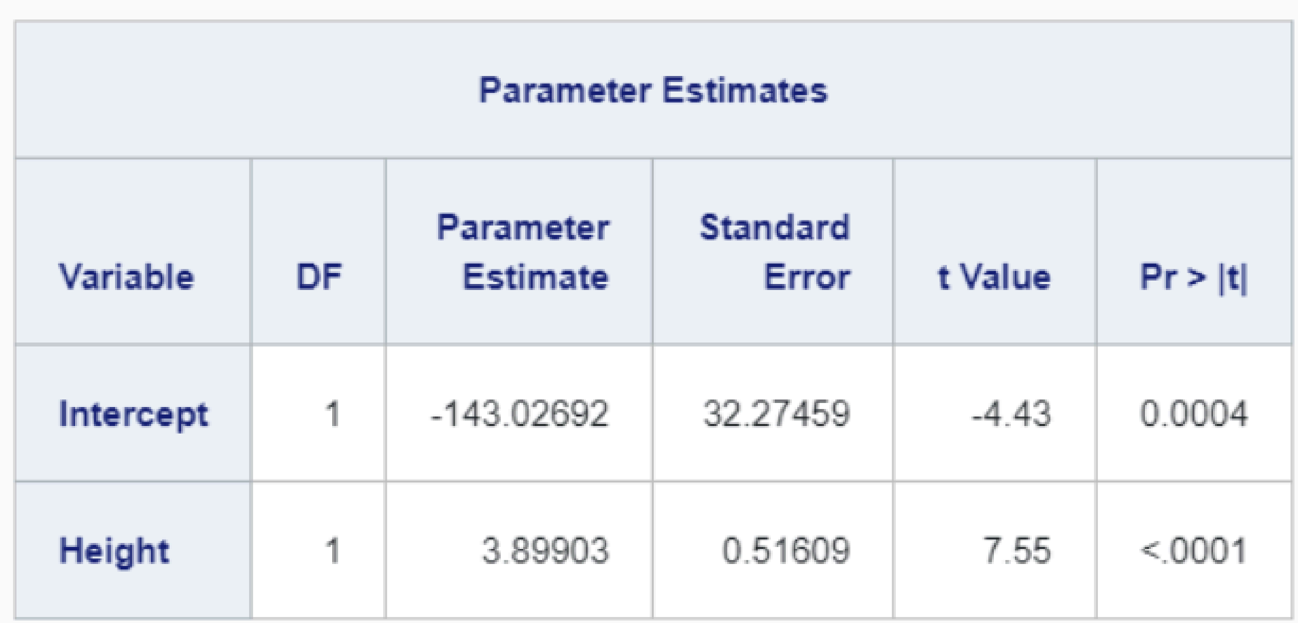
○ $e_i = \beta_0 + \beta_1 x_i$

✓ ● $e_i = y_i - b_0 - b_1 x_i$

○ $e_i = y_i - \beta_0 - \beta_1 x_i$

Question 30 1 / 1 point

Below is the SAS output for a simple linear regression model that uses an individual's height to predict their weight. Use this output to construct a 95% confidence interval for the model parameter for height. Assume that $n = 22$ observations were used to construct this model.



Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|
Intercept	1	-143.02692	32.27459	-4.43	0.0004
Height	1	3.89903	0.51609	7.55	<.0001

✓ ● (2.82, 4.98)

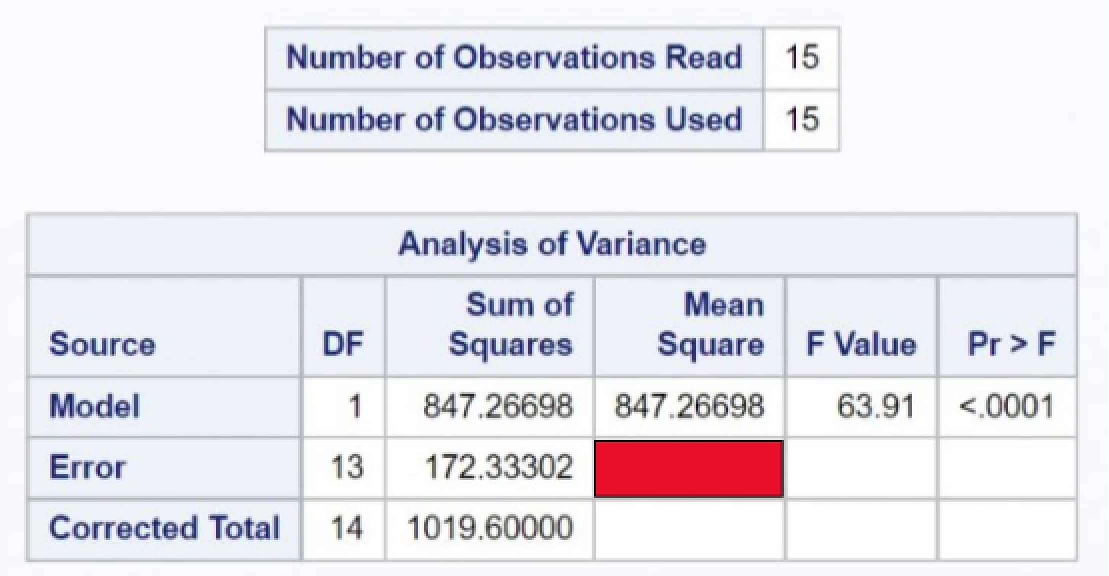
○ (3.67, 4.13)

○ (3.34, 4.45)

○ (2.89, 4.91)

Question 31 1 / 1 point

Below is the SAS ANOVA output for a simple linear regression model that uses the number of hours studied for an exam to predict the exam score. Let y represent an observed exam score, $\hat{y}$ represent a predicted exam score using the model, and $\bar{y}$ represent the mean exam score for all students.



Number of Observations Read	15
Number of Observations Used	15

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	1	847.26698	847.26698	63.91	<.0001
Error	13	172.33302	🟥		
Corrected Total	14	1019.60000			

Which of the following formulas was used to calculate the number for the Sum of Squares for Error, 172.33302?

○ $\sum_{i=1}^{15} (y_i - \bar{y})^2$

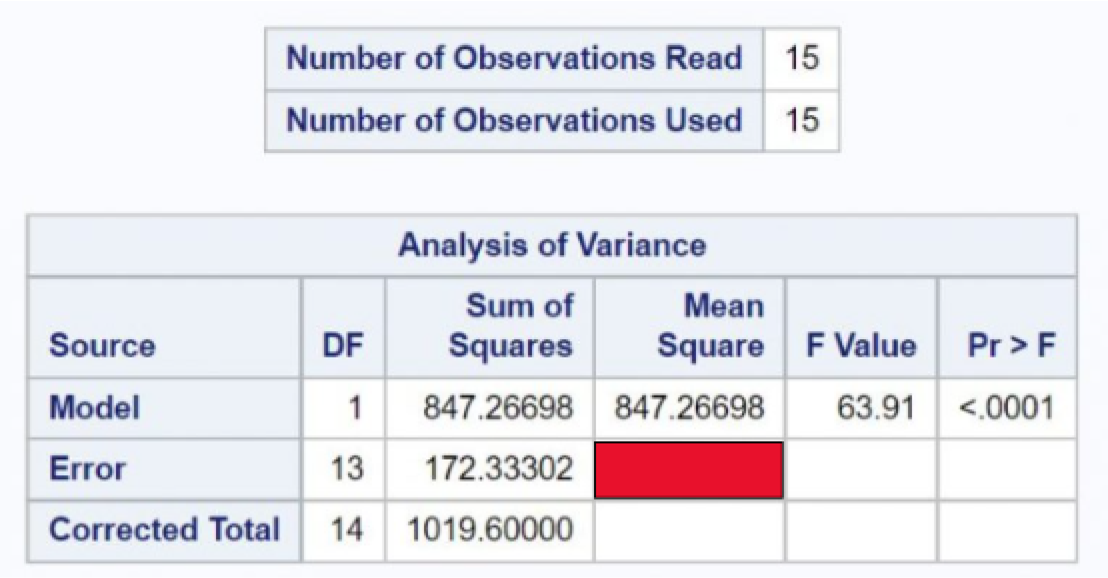
✓ ● $\sum_{i=1}^{15} \left(y_i - \hat{y}_i\right)^2$

○ $\sum_{i=1}^{15} \left(\hat{y}_i - \bar{y}\right)^2$

○ $\sum_{i=1}^{15} (y_i - \bar{y})^2$

Question 32 1 / 1 point

Below is the SAS ANOVA output for a simple linear regression model that uses the number of hours studied for an exam to predict the exam score.



Number of Observations Read	15
Number of Observations Used	15

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	1	847.26698	847.26698	63.91	<.0001
Error	13	172.33302	🟥		
Corrected Total	14	1019.60000			

What number belongs in the red box corresponding to the Mean Square of Error (MSE)?

✓ ● 13.25639

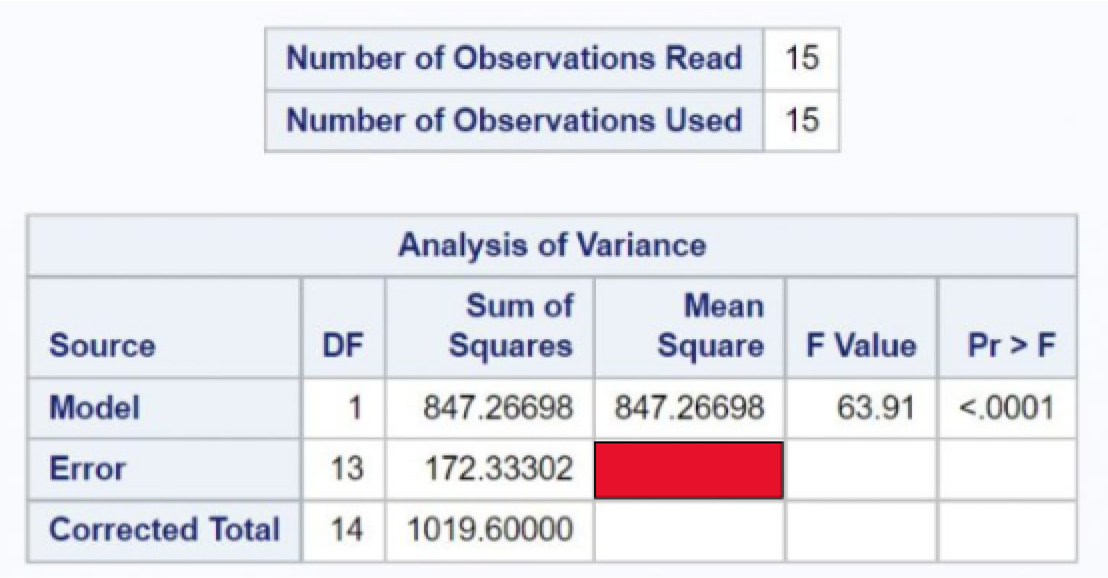
○ 4.91645

○ 172.33302

○ 674.93400

Question 33 1 / 1 point

Below is the SAS ANOVA output for a simple linear regression model that uses the number of hours studied for an exam to predict the exam score.



Number of Observations Read	15
Number of Observations Used	15

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	1	847.26698	847.26698	63.91	<.0001
Error	13	172.33302	🟥		
Corrected Total	14	1019.60000			

What are the degrees of freedom for the F-statistic that is used to test the hypothesis that the number of hours studied has no effect on the resulting exam score?

✓ ● 1 and 13

○ 13 and 14

○ 1 and 14

○ 1 and 15

Question 34 1 / 1 point

Suppose you wish to use information about the number of years of education for an individual, X_1, and an individual's age (in years), X_2, to predict their annual salary (in dollars), Y. You decide to use a multiple linear regression model to do this. Which of the following is the **most appropriate** way to express this model?

○ $Y_i = \beta_1 X_{i1} + \beta_2 X_{i2} + \epsilon_i$ where $\epsilon_i \overset{iid}{\sim} N(0, \sigma^2)$

○ $Y_i = Y_{i1} + Y_{i2}$

$Y_{i1} = \beta_0 + \beta_1 X_{i1} + \epsilon_{i1}$ where $\epsilon_{i1} \overset{iid}{\sim} N(0, \sigma^2)$

$Y_{i2} = \beta_0 + \beta_1 X_{i2} + \epsilon_{i2}$ where $\epsilon_{i2} \overset{iid}{\sim} N(0, \sigma^2)$

✓ ● $Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \epsilon_i$ where $\epsilon_i \overset{iid}{\sim} N(0, \sigma^2)$

○ $Y_i = \beta_0 + \beta_1 (X_{i1} + X_{i2}) + \epsilon_i$ where $\epsilon_i \overset{iid}{\sim} N(0, \sigma^2)$

Question 35	1 / 1 point

You have collected a random sample of 50 individuals and have recorded three pieces of information about each individual: The number of years of education they possess, X_1, their age (in years), X_2, and their annual income (in dollars), Y. You fit two different models:

Model 1: A model using both X_1 and X_2 to predict Y

Model 2: A model using only X_1 to predict Y

For each model, you have the ANOVA output which includes the Sum of Squares for Error (SSE) for each model, denoted as $SSE_1 = 350$ and $SSE_2 = 400$ respectively. You wish to use this information to test whether it is acceptable to exclude X_2 from the model.

What is the value of the F-statistic that would be calculated for this test?

○ 5.88

○ 6.85

○ 6.00

✓ ● 6.71

Done