FinalStudyGuide2

Joel Boynton

December 4, 2025

Contents

[11pt, letterpaper]article [utf8]inputenc [margin=1in]geometry amsmath, amssymb booktabs enumitem fancyhdr titlesec

% Header/Footer Setup STAT 3128 - Practice Exam Final Exam Practice (Remake) 1

% Title Formatting {**STAT 3128 Final Exam Practice (Remake)**}
Instructions:

- Please show all your work in the blank area provided below each question.
- Round your answers to **four decimal places**.
- Illegible handwriting will not be awarded any credit.

Question 1

Researchers want to see if a **High-Protein (HP)** diet is associated with higher muscle mass volume compared with a **Standard-Protein (SP)** diet. The following table summarizes collected muscle mass index data.

Diet	**Sample Size (n)**	**Sample Mean ($\bar{x}$)**	**Sample Std Dev (s)**
High-Protein (HP)	14	28	5.2
Standard-Protein (SP)	13	24	3.8

Assuming the populations are normal, test the following hypothesis at the significance level $\alpha = 0.05$:

$$H_0 : \mu_{hp} = \mu_{sp} \quad \text{VS} \quad H_a : \mu_{hp} > \mu_{sp}$$

where μ_{hp} and μ_{sp} represent the population average muscle mass for High-Protein and Standard-Protein diets, respectively.

Question 2

The reaction time of a chemical process under standard conditions is known to be normally distributed with mean value $\mu = 45$ seconds and standard deviation $\sigma = 4$ seconds. Engineers have proposed a new catalyst designed to **decrease** average reaction time. Because of the cost, evidence should strongly suggest an improvement before the catalyst is adopted. Experimental data consists of reaction times from $n = 36$ test runs.

[label=()]From the sample, it is given that $\sum_{i=1}^{n} x = 1566$. Using the significance level $\alpha = 0.01$, test the following hypothesis against its alternative:

$$H_0 : \mu = 45 \quad \text{VS} \quad H_a : \mu < 45$$

Find the **Type II error rate (β)** of the test in (a) when the true

alternative mean is $\mu_a = 43$.

Question 3

A journal article reports that a sample of size $n = 16$ was used as a basis for calculating a 90% Confidence Interval (CI) for the true average elasticity of a polymer. We know that the population is normally distributed, but σ is unknown. The reported 90% confidence interval was $(145.2500, 154.7500)$.

You think that the reported confidence interval is too narrow and a **99% confidence level** is more appropriate. What are the lower and upper limits of the 99% interval?

Question 4

A real estate agency wants to estimate the selling price of houses based on their size. They collected data on house size (X in 100 sq ft) and selling price (Y in \$10,000s) for $n = 20$ houses. A summary of the data is provided below:

$$\bar{x} = 20.0, \quad \bar{y} = 30.0$$
$$\sum x_i^2 = 8,500, \quad \sum y_i^2 = 18,600$$
$$\sum x_i y_i = 12,400$$
$$\sum (y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i)^2 = 144.0$$

[label=()]Construct the simple linear regression model that predicts Price (y) based on Size (x). The agency wants to test whether the estimated regression line is reliable. Test the following hypothesis at the significance level of $\alpha = 0.05$:

$$H_0 : \beta_1 = 0 \quad \text{VS} \quad H_a : \beta_1 \neq 0$$

A new house has a size of 2,400 sq ft ($x = 24$). Using the linear

regression model created in (a), predict the selling price.

Question 5

An engineer is studying the relationship between the pressure setting (x) on a machine and the material density (y). The following summary statistics were calculated for $n = 10$ observations:

$$\sum x_i = 50, \quad \sum y_i = 200$$
$$\sum x_i^2 = 290, \quad \sum y_i^2 = 4100$$
$$\sum x_i y_i = 1080$$
$$\sum (y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i)^2 = 32.0$$
$$\text{Minimum } x = 2, \quad \text{Maximum } x = 8$$

[label=()]Estimate the true average change in density (y) associated with a 1-unit increase in pressure (x) using a **95% confidence interval**. Construct a **95% Confidence Interval for the mean response**

at $x = 5.0$. Suppose we have a new predictor value $x_{new} = 12$. Would

you recommend calculating a 95% CI for the mean response of 12? Explain why or why not.

Answer Key

Question 1: Two-Sample t-Test

- **Degrees of Freedom (Satterthwaite):**

$$\nu \approx \frac{(5.2^2/14 + 3.8^2/13)^2}{\frac{(5.2^2/14)^2}{13} + \frac{(3.8^2/13)^2}{12}} = \frac{9.2550}{0.2869 + 0.1028} \approx 23.7 \to \mathbf{23}$$

- **Test Statistic:**

$$t = \frac{28 - 24}{\sqrt{1.9314 + 1.1108}} = \frac{4}{1.7442} = \mathbf{2.2933}$$

- **Conclusion:** $t_{crit}(df = 23, \alpha = 0.05) = 1.714$. Since $2.2933 > 1.714$, **Reject** H_0.

Question 2: Z-Test & Type II Error

- **(a) Hypothesis Test:** $\bar{x} = 1566/36 = 43.5$.

$$Z = \frac{43.5 - 45}{4/\sqrt{36}} = -2.2500$$

  Critical value ($Z_{0.01}$) = -2.33. Since $-2.25 > -2.33$, **Fail to Reject** H_0.

- **(b) Type II Error (β):** Rejection region boundary (in terms of $\bar{x}$):

$$\bar{x}_{crit} = 45 - 2.33(0.6667) = 43.4467$$

  Calculate Z under $H_a(\mu = 43)$:

$$Z = \frac{43.4467 - 43.0}{0.6667} = 0.6700$$

$$\beta = P(Z < 0.6700) = \mathbf{0.7486}$$

Question 3: Confidence Interval Manipulation

- $\bar{x} = 150.0$, Margin (E_{90}) = 4.75.
- Find s: $E_{90} = t_{0.05,15}(s/\sqrt{n}) \to 4.75 = 1.753(s/4) \to s = 10.8386$.
- New E_{99}: $t_{0.005,15} = 2.947$. $E_{99} = 2.947(10.8386/4) = 7.9856$.
- **99% CI:** $150 \pm 7.9856 \to (\mathbf{142.0144}, \mathbf{157.9856})$.

Question 4: Simple Linear Regression

- $S_{xx} = 500, S_{xy} = 400$. $\hat{\beta}_1 = 0.8, \hat{\beta}_0 = 14.0$.
- **(a)** $\hat{y} = 14.0 + 0.8x$.
- **(b)** $MSE = 144/18 = 8$. $s_{\beta_1} = \sqrt{8/500} = 0.1265$.

$$t = 0.8/0.1265 = \mathbf{6.3241} > 2.101 \to \textbf{Reject } H_0.$$

- **(c)** $\hat{y} = 14 + 0.8(24) = \mathbf{33.2}$.

Question 5: SLR Intervals

- $S_{xx} = 40, S_{xy} = 80, \hat{\beta}_1 = 2, MSE = 4$.
- **(a) CI for Slope:** $2 \pm 2.306\sqrt{4/40} \to (\mathbf{1.2708}, \mathbf{2.7292})$.
- **(b) CI Mean Response ($x = 5$):** $20 \pm 2.306\sqrt{4(0.1)} \to (\mathbf{18.5415}, \mathbf{21.4585})$.
- **(c) No.** $x = 12$ is outside the data range (Extrapolation).