ISE 315: Engineering Statistics

Homework 4 Solution

Class Section: F-04, Instructor: Mansur M. Arief
Department of Industrial and Systems Engineering, KFUPM

Set A

Problem 1 [18 pt] — Z-Test for the Mean (Known σ, Two-Sided)

Given: $\mu_0 = 50$ ms, $\sigma = 12$ ms, $n = 36$, $\bar{x} = 46.4$ ms, $\alpha = 0.05$.

Solution:

Step 1: Hypotheses:

$$H_0 : \mu = 50 \qquad \text{vs.} \qquad H_1 : \mu \neq 50$$

Step 2: Test statistic:

$$z_0 = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} = \frac{46.4 - 50}{12/\sqrt{36}} = \frac{-3.60}{2.00} = -1.8000$$

Step 3: P-value: Two-sided test, so P-value $= 2\,P(Z > |z_0|) = 2\,P(Z > 1.8000)$.

From the standard normal table, $\Phi(1.80) = 0.9641$, so $P(Z > 1.80) = 0.0359$.

Therefore, P-value $= 2(0.0359) = 0.0719$.

Step 4: Decision: Since P-value $= 0.0719 > 0.05 = \alpha$, we **fail to reject** H_0.

Conclusion: At $\alpha = 0.05$, there is insufficient evidence to conclude that the mean flagging time differs from 50 ms.

Scoring Rubric

Error	Deduction
Wrong or missing hypotheses	-3
Uses t instead of z (σ known)	-4
Wrong test statistic computation	-4
Wrong P-value computation or bounding	-3
Wrong decision/conclusion or missing justification	-4
Missing or no work shown	-15

Problem 2 [10 pt] — Type II Error and Power

Given: From Problem 1: $H_0 : \mu = 50$ vs. $H_1 : \mu \neq 50$, $n = 36$, $\sigma = 12$ ms, $\alpha = 0.05$. True mean $\mu = 47$ ms.

Solution:

(a) **Type II error (β): [5 pt]**

We fail to reject H_0 when $-z_{0.025} \leq Z_0 \leq z_{0.025}$, i.e., when $\bar{X}$ falls in the interval

$$50 - 1.96(2.00) \leq \bar{X} \leq 50 + 1.96(2.00) \quad \Longrightarrow \quad 46.0800 \leq \bar{X} \leq 53.9200$$

Under $\mu = 47$:

$$\beta = P\left(46.0800 \leq \bar{X} \leq 53.9200 \,\middle|\, \mu = 47\right) = \Phi\left(\frac{53.9200 - 47}{2.00}\right) - \Phi\left(\frac{46.0800 - 47}{2.00}\right)$$

$$= \Phi(3.46) - \Phi(-0.46) = 0.9997 - 0.3228 = \boxed{0.6770}$$

(b) **Power: [5 pt]**

$$\text{Power} = 1 - \beta = 1 - 0.6770 = \boxed{0.3230}$$

Scoring Rubric

Error	Deduction
Wrong fail-to-reject region or critical boundaries	-3
Wrong standardization under true μ	-3
Wrong Φ lookup or arithmetic	-2
Part (b): wrong power formula	-2
Missing or no work shown	-8

Problem 3 [18 pt] — T-Test for the Mean (One-Sided Upper)

Given: $\mu_0 = 85\%$, $n = 16$, $\bar{x} = 88.0\%$, $s = 5.5\%$, $\alpha = 0.05$.

Solution:

Step 1: Hypotheses: (Claim: accuracy exceeds 85%)

$$H_0 : \mu = 85 \qquad \text{vs.} \qquad H_1 : \mu > 85$$

Step 2: Test statistic: $\nu = 15$.

$$t_0 = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} = \frac{88.0 - 85}{5.5/\sqrt{16}} = \frac{3.00}{1.3750} = 2.1818$$

Step 3: P-value: Upper-tail test. From the t-table with $\nu = 15$: $0.01 < \text{P-value} < 0.025$. (Exact P-value $= 0.0227$.)

Step 4: Decision: Since P-value $< 0.05 = \alpha$, we **reject** H_0.

Equivalently, $t_0 = 2.1818 > t_{0.05,15} = 1.753$, so we reject H_0.

Conclusion: At $\alpha = 0.05$, there is sufficient evidence to support the team's claim that the accuracy exceeds 85%.

Scoring Rubric

Error	Deduction
Wrong or missing hypotheses (e.g., two-sided instead of one-sided)	-3
Uses z instead of t (σ unknown, small sample)	-4
Wrong test statistic computation	-4
Wrong P-value bounding or critical value comparison	-3
Wrong decision/conclusion or missing justification	-4
Missing or no work shown	-15

Problem 4 [18 pt] — Chi-Square Test for Variance (One-Sided Upper)

Given: $\sigma_0^2 = 12\ \text{m}^2$, $n = 20$, $s^2 = 17.0\ \text{m}^2$, $\alpha = 0.01$.

Solution:

Step 1: Hypotheses:

$$H_0 : \sigma^2 = 12 \qquad \text{vs.} \qquad H_1 : \sigma^2 > 12$$

Step 2: Test statistic: $\nu = 19$.

$$\chi_0^2 = \frac{(n-1)\,s^2}{\sigma_0^2} = \frac{19 \times 17.0}{12} = 26.92$$

Step 3: P-value: Upper-tail test. From the χ^2 table with $\nu = 19$: P-value > 0.05.

(Exact P-value $= 0.1066$.)

Step 4: Decision: Since P-value $> 0.01 = \alpha$, we **fail to reject** H_0.

Equivalently, $\chi_0^2 = 26.92 < \chi^2_{0.01,19} = 36.191$, so we fail to reject H_0.

Conclusion: At $\alpha = 0.01$, there is insufficient evidence that the variance exceeds the safety specification of $\sigma^2 = 12\ \text{m}^2$.

Scoring Rubric

Error	Deduction
Wrong or missing hypotheses	-3
Wrong test statistic formula (e.g., uses z or t)	-4
Wrong computation of χ_0^2	-4
Wrong P-value bounding or critical value comparison	-3
Wrong decision/conclusion or missing justification	-4
Missing or no work shown	-15

Problem 5 [18 pt] — Z-Test for a Proportion (One-Sided Lower)

Given: $p_0 = 0.88$, $n = 200$, $x = 170$, $\alpha = 0.01$.

Solution:

Step 1: Hypotheses: (Test whether compliance has fallen below 88%)

$$H_0 : p = 0.88 \qquad \text{vs.} \qquad H_1 : p < 0.88$$

Step 2: Sample proportion: $\hat{p} = 170/200 = 0.85$.

Step 3: Test statistic:

$$z_0 = \frac{\hat{p} - p_0}{\sqrt{p_0(1-p_0)/n}} = \frac{0.85 - 0.88}{\sqrt{0.88 \times 0.12/200}} = \frac{-0.0300}{0.0230} = -1.3056$$

Step 4: P-value: Lower-tail test, so P-value $= P(Z < -1.3056) = \Phi(-1.3056) = 0.0958$.

Step 5: Decision: Since P-value $= 0.0958 > 0.01 = \alpha$, we **fail to reject** H_0.

Conclusion: At $\alpha = 0.01$, there is insufficient evidence that the tool's fairness compliance rate has fallen below 88%.

Scoring Rubric

Error	Deduction
Wrong or missing hypotheses (e.g., upper-tail or two-sided)	-3
Wrong $\hat{p}$ computation	-3
Wrong SE formula (e.g., uses $\hat{p}$ instead of p_0)	-3
Wrong test statistic computation	-3
Wrong P-value or wrong tail	-3
Wrong decision/conclusion	-3
Missing or no work shown	-15

Problem 6 [18 pt] — T-Test for the Mean (One-Sided Lower)

Given: $\mu_0 = 150$ ms, $n = 12$, $\bar{x} = 138$ ms, $s = 16$ ms, $\alpha = 0.1$.

Solution:

Step 1: Hypotheses: (Claim: latency is less than 150 ms)

$$H_0 : \mu = 150 \qquad \text{vs.} \qquad H_1 : \mu < 150$$

Step 2: Test statistic: $\nu = 11$.

$$t_0 = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} = \frac{138 - 150}{16/\sqrt{12}} = \frac{-12.00}{4.6188} = -2.5981$$

Step 3: P-value: Lower-tail test. Since $|t_0| = 2.5981$, from the t-table with $\nu = 11$: $0.01 <$ P-value < 0.025.

(Exact P-value $= 0.0124$.)

Step 4: Decision: Since P-value $< 0.1 = \alpha$, we **reject** H_0.

Equivalently, $t_0 = -2.5981 < -t_{0.1,11} = -1.363$, so we reject H_0.

Conclusion: At $\alpha = 0.1$, there is sufficient evidence to support the startup's claim that the average inference latency is less than 150 ms.

Scoring Rubric

Error	Deduction
Wrong or missing hypotheses (e.g., upper-tail or two-sided)	-3
Uses z instead of t (σ unknown, small sample)	-4
Wrong test statistic computation	-4
Wrong P-value bounding or critical value comparison	-3
Wrong decision/conclusion or missing justification	-4
Missing or no work shown	-15

Set A Total: $18 + 10 + 18 + 18 + 18 + 18 =$ **100 points.**

Set B

Problem 1 [18 pt] — Z-Test for the Mean (Known σ, Two-Sided)

Given: $\mu_0 = 50$ ms, $\sigma = 14$ ms, $n = 49$, $\bar{x} = 53.6$ ms, $\alpha = 0.05$.

Solution:

Step 1: Hypotheses:

$$H_0 : \mu = 50 \qquad \text{vs.} \qquad H_1 : \mu \neq 50$$

Step 2: Test statistic:

$$z_0 = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} = \frac{53.6 - 50}{14/\sqrt{49}} = \frac{3.60}{2.00} = 1.8000$$

Step 3: P-value: Two-sided test, so P-value $= 2\,P(Z > |z_0|) = 2\,P(Z > 1.8000)$.

From the standard normal table, $\Phi(1.80) = 0.9641$, so $P(Z > 1.80) = 0.0359$.

Therefore, P-value $= 2(0.0359) = 0.0719$.

Step 4: Decision: Since P-value $= 0.0719 > 0.05 = \alpha$, we **fail to reject** H_0.

Conclusion: At $\alpha = 0.05$, there is insufficient evidence to conclude that the mean flagging time differs from 50 ms.

Scoring Rubric

Error	Deduction
Wrong or missing hypotheses	-3
Uses t instead of z (σ known)	-4
Wrong test statistic computation	-4
Wrong P-value computation or bounding	-3
Wrong decision/conclusion or missing justification	-4
Missing or no work shown	-15

Problem 2 [10 pt] — Type II Error and Power

Given: From Problem 1: $H_0 : \mu = 50$ vs. $H_1 : \mu \neq 50$, $n = 49$, $\sigma = 14$ ms, $\alpha = 0.05$. True mean $\mu = 53$ ms.

Solution:

(a) **Type II error (β): [5 pt]**

We fail to reject H_0 when $-z_{0.025} \leq Z_0 \leq z_{0.025}$, i.e., when $\bar{X}$ falls in the interval

$$50 - 1.96(2.00) \leq \bar{X} \leq 50 + 1.96(2.00) \quad \Longrightarrow \quad 46.0800 \leq \bar{X} \leq 53.9200$$

Under $\mu = 53$:

$$\beta = P\left(46.0800 \leq \bar{X} \leq 53.9200 \,\middle|\, \mu = 53\right) = \Phi\left(\frac{53.9200 - 53}{2.00}\right) - \Phi\left(\frac{46.0800 - 53}{2.00}\right)$$

$$= \Phi(0.46) - \Phi(-3.46) = 0.6772 - 0.0003 = \boxed{0.6770}$$

(b) **Power: [5 pt]**

$$\text{Power} = 1 - \beta = 1 - 0.6770 = \boxed{0.3230}$$

Scoring Rubric

Error	**Deduction**
Wrong fail-to-reject region or critical boundaries	-3
Wrong standardization under true μ	-3
Wrong Φ lookup or arithmetic	-2
Part (b): wrong power formula	-2
Missing or no work shown	-8

Problem 3 [18 pt] — T-Test for the Mean (One-Sided Upper)

Given: $\mu_0 = 85\%$, $n = 20$, $\bar{x} = 87.5\%$, $s = 6.0\%$, $\alpha = 0.05$.

Solution:

Step 1: Hypotheses: (Claim: accuracy exceeds 85%)

$$H_0 : \mu = 85 \qquad \text{vs.} \qquad H_1 : \mu > 85$$

Step 2: Test statistic: $\nu = 19$.

$$t_0 = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} = \frac{87.5 - 85}{6.0/\sqrt{20}} = \frac{2.50}{1.3416} = 1.8634$$

Step 3: P-value: Upper-tail test. From the t-table with $\nu = 19$: $0.025 < \text{P-value} < 0.05$. (Exact P-value $= 0.0390$.)

Step 4: Decision: Since P-value $< 0.05 = \alpha$, we **reject** H_0.

Equivalently, $t_0 = 1.8634 > t_{0.05,19} = 1.729$, so we reject H_0.

Conclusion: At $\alpha = 0.05$, there is sufficient evidence to support the team's claim that the accuracy exceeds 85%.

Scoring Rubric

Error	**Deduction**
Wrong or missing hypotheses (e.g., two-sided instead of one-sided)	-3
Uses z instead of t (σ unknown, small sample)	-4
Wrong test statistic computation	-4
Wrong P-value bounding or critical value comparison	-3
Wrong decision/conclusion or missing justification	-4
Missing or no work shown	-15

Problem 4 [18 pt] — Chi-Square Test for Variance (One-Sided Upper)

Given: $\sigma_0^2 = 12\ \text{m}^2$, $n = 25$, $s^2 = 18.0\ \text{m}^2$, $\alpha = 0.01$.

Solution:

Step 1: Hypotheses:

$$H_0 : \sigma^2 = 12 \qquad \text{vs.} \qquad H_1 : \sigma^2 > 12$$

Step 2: Test statistic: $\nu = 24$.

$$\chi_0^2 = \frac{(n-1)\,s^2}{\sigma_0^2} = \frac{24 \times 18.0}{12} = 36.00$$

Step 3: P-value: Upper-tail test. From the χ^2 table with $\nu = 24$: P-value > 0.05.

(Exact P-value $= 0.0549$.)

Step 4: Decision: Since P-value $> 0.01 = \alpha$, we **fail to reject** H_0.

Equivalently, $\chi_0^2 = 36.00 < \chi^2_{0.01,24} = 42.980$, so we fail to reject H_0.

Conclusion: At $\alpha = 0.01$, there is insufficient evidence that the variance exceeds the safety specification of $\sigma^2 = 12\ \text{m}^2$.

Scoring Rubric

Error	Deduction
Wrong or missing hypotheses	-3
Wrong test statistic formula (e.g., uses z or t)	-4
Wrong computation of χ_0^2	-4
Wrong P-value bounding or critical value comparison	-3
Wrong decision/conclusion or missing justification	-4
Missing or no work shown	-15

Problem 5 [18 pt] — Z-Test for a Proportion (One-Sided Lower)

Given: $p_0 = 0.88$, $n = 250$, $x = 210$, $\alpha = 0.01$.

Solution:

Step 1: Hypotheses: (Test whether compliance has fallen below 88%)

$$H_0 : p = 0.88 \qquad \text{vs.} \qquad H_1 : p < 0.88$$

Step 2: Sample proportion: $\hat{p} = 210/250 = 0.84$.

Step 3: Test statistic:

$$z_0 = \frac{\hat{p} - p_0}{\sqrt{p_0(1-p_0)/n}} = \frac{0.84 - 0.88}{\sqrt{0.88 \times 0.12/250}} = \frac{-0.0400}{0.0206} = -1.9462$$

Step 4: P-value: Lower-tail test, so P-value $= P(Z < -1.9462) = \Phi(-1.9462) = 0.0258$.

Step 5: Decision: Since P-value $= 0.0258 > 0.01 = \alpha$, we **fail to reject** H_0.

Conclusion: At $\alpha = 0.01$, there is insufficient evidence that the tool's fairness compliance rate has fallen below 88%.

Scoring Rubric

Error	Deduction
Wrong or missing hypotheses (e.g., upper-tail or two-sided)	-3
Wrong $\hat{p}$ computation	-3
Wrong SE formula (e.g., uses $\hat{p}$ instead of p_0)	-3
Wrong test statistic computation	-3
Wrong P-value or wrong tail	-3
Wrong decision/conclusion	-3
Missing or no work shown	-15

Problem 6 [18 pt] — T-Test for the Mean (One-Sided Lower)

Given: $\mu_0 = 150$ ms, $n = 15$, $\bar{x} = 141$ ms, $s = 18$ ms, $\alpha = 0.1$.

Solution:

Step 1: Hypotheses: (Claim: latency is less than 150 ms)

$$H_0 : \mu = 150 \qquad \text{vs.} \qquad H_1 : \mu < 150$$

Step 2: Test statistic: $\nu = 14$.

$$t_0 = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} = \frac{141 - 150}{18/\sqrt{15}} = \frac{-9.00}{4.6476} = -1.9365$$

Step 3: P-value: Lower-tail test. Since $|t_0| = 1.9365$, from the t-table with $\nu = 14$: $0.025 <$ P-value < 0.05.

(Exact P-value $= 0.0366$.)

Step 4: Decision: Since P-value $< 0.1 = \alpha$, we **reject** H_0.

Equivalently, $t_0 = -1.9365 < -t_{0.1,14} = -1.345$, so we reject H_0.

Conclusion: At $\alpha = 0.1$, there is sufficient evidence to support the startup's claim that the average inference latency is less than 150 ms.

Scoring Rubric

Error	Deduction
Wrong or missing hypotheses (e.g., upper-tail or two-sided)	-3
Uses z instead of t (σ unknown, small sample)	-4
Wrong test statistic computation	-4
Wrong P-value bounding or critical value comparison	-3
Wrong decision/conclusion or missing justification	-4
Missing or no work shown	-15

Set B Total: $18 + 10 + 18 + 18 + 18 + 18 =$ **100 points.**

Set C

Problem 1 [18 pt] — Z-Test for the Mean (Known σ, Two-Sided)

Given: $\mu_0 = 50$ ms, $\sigma = 16$ ms, $n = 64$, $\bar{x} = 46.0$ ms, $\alpha = 0.05$.

Solution:

Step 1: Hypotheses:

$$H_0 : \mu = 50 \qquad \text{vs.} \qquad H_1 : \mu \neq 50$$

Step 2: Test statistic:

$$z_0 = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} = \frac{46.0 - 50}{16/\sqrt{64}} = \frac{-4.00}{2.00} = -2.0000$$

Step 3: P-value: Two-sided test, so P-value $= 2\,P(Z > |z_0|) = 2\,P(Z > 2.0000)$.

From the standard normal table, $\Phi(2.00) = 0.9772$, so $P(Z > 2.00) = 0.0228$.

Therefore, P-value $= 2(0.0228) = 0.0455$.

Step 4: Decision: Since P-value $= 0.0455 < 0.05 = \alpha$, we **reject** H_0.

Conclusion: At $\alpha = 0.05$, there is sufficient evidence that the mean flagging time differs from 50 ms.

Scoring Rubric

Error	**Deduction**
Wrong or missing hypotheses	-3
Uses t instead of z (σ known)	-4
Wrong test statistic computation	-4
Wrong P-value computation or bounding	-3
Wrong decision/conclusion or missing justification	-4
Missing or no work shown	-15

Problem 2 [10 pt] — Type II Error and Power

Given: From Problem 1: $H_0 : \mu = 50$ vs. $H_1 : \mu \neq 50$, $n = 64$, $\sigma = 16$ ms, $\alpha = 0.05$. True mean $\mu = 46$ ms.

Solution:

(a) **Type II error (β): [5 pt]**

We fail to reject H_0 when $-z_{0.025} \leq Z_0 \leq z_{0.025}$, i.e., when $\bar{X}$ falls in the interval

$$50 - 1.96(2.00) \leq \bar{X} \leq 50 + 1.96(2.00) \quad \Longrightarrow \quad 46.0800 \leq \bar{X} \leq 53.9200$$

Under $\mu = 46$:

$$\beta = P\Big(46.0800 \leq \bar{X} \leq 53.9200 \;\Big|\; \mu = 46\Big) = \Phi\left(\frac{53.9200 - 46}{2.00}\right) - \Phi\left(\frac{46.0800 - 46}{2.00}\right)$$

$$= \Phi(3.96) - \Phi(0.04) = 1.0000 - 0.5160 = \boxed{0.4840}$$

(b) **Power: [5 pt]**

$$\text{Power} = 1 - \beta = 1 - 0.4840 = \boxed{0.5160}$$

Scoring Rubric

Error	Deduction
Wrong fail-to-reject region or critical boundaries	-3
Wrong standardization under true μ	-3
Wrong Φ lookup or arithmetic	-2
Part (b): wrong power formula	-2
Missing or no work shown	-8

Problem 3 [18 pt] — T-Test for the Mean (One-Sided Upper)

Given: $\mu_0 = 85\%$, $n = 25$, $\bar{x} = 87.2\%$, $s = 5.5\%$, $\alpha = 0.05$.

Solution:

Step 1: Hypotheses: (Claim: accuracy exceeds 85%)

$$H_0 : \mu = 85 \qquad \text{vs.} \qquad H_1 : \mu > 85$$

Step 2: Test statistic: $\nu = 24$.

$$t_0 = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} = \frac{87.2 - 85}{5.5/\sqrt{25}} = \frac{2.20}{1.1000} = 2.0000$$

Step 3: P-value: Upper-tail test. From the t-table with $\nu = 24$: $0.025 < \text{P-value} < 0.05$.

(Exact P-value = 0.0285.)

Step 4: Decision: Since P-value $< 0.05 = \alpha$, we **reject** H_0.

Equivalently, $t_0 = 2.0000 > t_{0.05,24} = 1.711$, so we reject H_0.

Conclusion: At $\alpha = 0.05$, there is sufficient evidence to support the team's claim that the accuracy exceeds 85%.

Scoring Rubric

Error	Deduction
Wrong or missing hypotheses (e.g., two-sided instead of one-sided)	−3
Uses z instead of t (σ unknown, small sample)	−4
Wrong test statistic computation	−4
Wrong P-value bounding or critical value comparison	−3
Wrong decision/conclusion or missing justification	−4
Missing or no work shown	−15

Problem 4 [18 pt] — Chi-Square Test for Variance (One-Sided Upper)

Given: $\sigma_0^2 = 12\ \text{m}^2$, $n = 16$, $s^2 = 19.5\ \text{m}^2$, $\alpha = 0.01$.

Solution:

Step 1: Hypotheses:

$$H_0 : \sigma^2 = 12 \qquad \text{vs.} \qquad H_1 : \sigma^2 > 12$$

Step 2: Test statistic: $\nu = 15$.

$$\chi_0^2 = \frac{(n-1)\,s^2}{\sigma_0^2} = \frac{15 \times 19.5}{12} = 24.38$$

Step 3: P-value: Upper-tail test. From the χ^2 table with $\nu = 15$: P-value > 0.05.

(Exact P-value $= 0.0590$.)

Step 4: Decision: Since P-value $> 0.01 = \alpha$, we **fail to reject** H_0.

Equivalently, $\chi_0^2 = 24.38 < \chi^2_{0.01,15} = 30.578$, so we fail to reject H_0.

Conclusion: At $\alpha = 0.01$, there is insufficient evidence that the variance exceeds the safety specification of $\sigma^2 = 12\ \text{m}^2$.

Scoring Rubric

Error	Deduction
Wrong or missing hypotheses	-3
Wrong test statistic formula (e.g., uses z or t)	-4
Wrong computation of χ_0^2	-4
Wrong P-value bounding or critical value comparison	-3
Wrong decision/conclusion or missing justification	-4
Missing or no work shown	-15

Problem 5 [18 pt] — Z-Test for a Proportion (One-Sided Lower)

Given: $p_0 = 0.88$, $n = 300$, $x = 258$, $\alpha = 0.01$.

Solution:

Step 1: Hypotheses: (Test whether compliance has fallen below 88%)

$$H_0 : p = 0.88 \qquad \text{vs.} \qquad H_1 : p < 0.88$$

Step 2: Sample proportion: $\hat{p} = 258/300 = 0.86$.

Step 3: Test statistic:

$$z_0 = \frac{\hat{p} - p_0}{\sqrt{p_0(1-p_0)/n}} = \frac{0.86 - 0.88}{\sqrt{0.88 \times 0.12/300}} = \frac{-0.0200}{0.0188} = -1.0660$$

Step 4: P-value: Lower-tail test, so P-value $= P(Z < -1.0660) = \Phi(-1.0660) = 0.1432$.

Step 5: Decision: Since P-value $= 0.1432 > 0.01 = \alpha$, we **fail to reject** H_0.

Conclusion: At $\alpha = 0.01$, there is insufficient evidence that the tool's fairness compliance rate has fallen below 88%.

Scoring Rubric

Error	Deduction
Wrong or missing hypotheses (e.g., upper-tail or two-sided)	−3
Wrong $\hat{p}$ computation	−3
Wrong SE formula (e.g., uses $\hat{p}$ instead of p_0)	−3
Wrong test statistic computation	−3
Wrong P-value or wrong tail	−3
Wrong decision/conclusion	−3
Missing or no work shown	−15

Problem 6 [18 pt] — T-Test for the Mean (One-Sided Lower)

Given: $\mu_0 = 150$ ms, $n = 10$, $\bar{x} = 139$ ms, $s = 14$ ms, $\alpha = 0.1$.

Solution:

Step 1: Hypotheses: (Claim: latency is less than 150 ms)

$$H_0 : \mu = 150 \qquad \text{vs.} \qquad H_1 : \mu < 150$$

Step 2: Test statistic: $\nu = 9$.

$$t_0 = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} = \frac{139 - 150}{14/\sqrt{10}} = \frac{-11.00}{4.4272} = -2.4846$$

Step 3: P-value: Lower-tail test. Since $|t_0| = 2.4846$, from the t-table with $\nu = 9$: $0.01 <$ P-value < 0.025.

(Exact P-value $= 0.0174$.)

Step 4: Decision: Since P-value $< 0.1 = \alpha$, we **reject** H_0.

Equivalently, $t_0 = -2.4846 < -t_{0.1,9} = -1.383$, so we reject H_0.

Conclusion: At $\alpha = 0.1$, there is sufficient evidence to support the startup's claim that the average inference latency is less than 150 ms.

Scoring Rubric

Error	Deduction
Wrong or missing hypotheses (e.g., upper-tail or two-sided)	-3
Uses z instead of t (σ unknown, small sample)	-4
Wrong test statistic computation	-4
Wrong P-value bounding or critical value comparison	-3
Wrong decision/conclusion or missing justification	-4
Missing or no work shown	-15

Set C Total: $18 + 10 + 18 + 18 + 18 + 18 =$ **100 points.**