MAE301 Homework 2

February 2, 2026

Please study the following topics:

- Central Limit Theorem (CLT)
- Z-test (large-sample hypothesis testing, known variance)
- T-test (small-sample or unknown variance)
- F-test (comparing two variances)
- Chi-square tests (goodness-of-fit, independence)

Problem 1: CLT for a Production Line

Statement:
A production line makes metal rods with a length X_i (in cm). Each rod length has mean $\mu = 100$ and standard deviation $\sigma = 2$. Assume X_i are i.i.d. random variables.

1. If you sample $n = 40$ rods at random, what is the approximate distribution of the sample average $\overline{X}$ for large n?
2. Using the CLT, find $P(99.5 \leq \overline{X} \leq 100.5)$ when $n = 40$.

Solution Sketch:

- (a) By the Central Limit Theorem, $\overline{X} \approx \mathcal{N}\big(\mu, \frac{\sigma^2}{n}\big)$ for large n. Hence $\overline{X} \sim \mathcal{N}\big(100, \frac{2^2}{40}\big)$ approximately.
- (b) $\overline{X}$ has mean 100 and std. dev. $\frac{2}{\sqrt{40}} = \frac{2}{6.3246} \approx 0.316$. Then:

  $$P(99.5 \leq \overline{X} \leq 100.5) \approx P\Big(\frac{99.5 - 100}{0.316} \leq Z \leq \frac{100.5 - 100}{0.316}\Big) = P(-1.58 \leq Z \leq 1.58).$$

  So $P(99.5 \leq \overline{X} \leq 100.5) = 2\Phi(1.58) - 1$.

Problem 2: CLT for Component Weights

Statement:
An engineering firm packs 25 electronic components per box. Let W_i be the weight (grams) of the ith component. Suppose W_i have mean $\mu = 50$ grams and std. dev. $\sigma = 4$ grams, i.i.d.

1. Let $S_{25} = \sum_{i=1}^{25} W_i$ be the total weight in the box. Approximate the distribution of S_{25} for large samples.

2. Find approximately $P(S_{25} > 1300)$ using the CLT.

Solution Sketch:

- (a) By the CLT, $S_{25} \approx \mathcal{N}(n\mu,\, n\sigma^2) = \mathcal{N}(25 \times 50,\ 25 \times 4^2) = \mathcal{N}(1250,\, 400)$.
- (b) Std. dev. is $\sqrt{400} = 20$. Then

$$P(S_{25} > 1300) \approx P\Big(Z > \frac{1300 - 1250}{20}\Big) = P(Z > 2.5).$$

  For a standard normal Z, $P(Z > 2.5) \approx 0.0062$.

Problem 3: Z-Test for Mean Diameter (Known Variance)

Statement:
A factory produces ball bearings with a supposed mean diameter of 10 mm. The population variance $\sigma^2 = 0.04\,\text{mm}^2$ is known. A random sample of $n = 50$ ball bearings yields sample mean $\overline{X} = 10.03\,\text{mm}$. We want to test:

$$H_0 : \mu = 10 \quad \text{vs} \quad H_1 : \mu \neq 10,$$

at the 5% significance level (two-sided).

1. Compute the test statistic.

2. State whether H_0 is rejected.

Solution Sketch:

- *Z-test statistic*:

$$Z = \frac{\overline{X} - \mu_0}{\sigma/\sqrt{n}} = \frac{10.03 - 10}{0.2/\sqrt{50}} = \frac{0.03}{0.2/7.071} = \frac{0.03}{0.0283} \approx 1.06.$$

- For a two-sided test at $\alpha = 0.05$, critical region is $|Z| > z_{0.025} \approx 1.96$. Since $|1.06| < 1.96$, we *fail to reject* H_0 at 5% significance.

Problem 4: One-Sample T-Test for Durability

Statement:
An R&D lab tests the durability (in hours) of a new composite material. The design specification claims the mean lifetime is 1200 hours. Because the population variance is *unknown*, a sample of $n = 10$ specimens is tested, yielding:

$$\overline{X} = 1180, \quad s = 50 \, \text{(sample std. dev.)}.$$

We want to test:

$$H_0 : \mu = 1200 \quad \text{vs} \quad H_1 : \mu < 1200,$$

at $\alpha = 0.05$. (Assume normality for lifetimes.)
Solution Sketch:

- *T statistic*:
$$T = \frac{\overline{X} - \mu_0}{s/\sqrt{n}} = \frac{1180 - 1200}{50/\sqrt{10}} = \frac{-20}{15.811} \approx -1.265.$$
- With *df* $= n - 1 = 9$, a one-tailed test at 5% significance has critical value $t_{0.05,9} \approx -1.833$ (for the left tail). Since $-1.265 > -1.833$, we *fail to reject* H_0 at $\alpha = 0.05$.

Problem 5: Paired T-Test for Two Sensor Types

Statement:
An engineer compares two types of sensors (A and B) for measuring the same voltage source. For each of 8 trials, they record the difference $d_i = (\text{Sensor A reading}) - (\text{Sensor B reading})$. The sample differences have mean $\overline{d} = 0.05\,\text{V}$ and sample std. dev. $s_d = 0.02\,\text{V}$. We test:

$$H_0 : \mu_d = 0 \quad \text{vs} \quad H_1 : \mu_d \neq 0,$$

assuming normal differences.
Solution Sketch:

- *Paired T-test statistic*:
$$T = \frac{\overline{d} - 0}{s_d/\sqrt{n}} = \frac{0.05}{0.02/\sqrt{8}} = \frac{0.05}{0.00707} \approx 7.07.$$
- *df* $= 7$. For a two-sided test at 5% significance, $t_{0.025,7} \approx 2.365$. Since $7.07 > 2.365$, we *reject* H_0 and conclude a significant difference.

Problem 6: F-Test for Variances of Two Manufacturing Processes

Statement:
Two different manufacturing processes produce resistors. We suspect different variance in resistor resistance. Process 1 sample: $n_1 = 12$, sample variance $s_1^2 = 0.0012\,\Omega^2$. Process 2 sample: $n_2 = 10$, sample variance $s_2^2 = 0.0008\,\Omega^2$. Test at 5% significance:

$$H_0 : \sigma_1^2 = \sigma_2^2 \quad \text{vs.} \quad H_1 : \sigma_1^2 \neq \sigma_2^2.$$

Solution Sketch:

- Compute $F = \frac{s_1^2}{s_2^2} = \frac{0.0012}{0.0008} = 1.5$. Let $df_1 = n_1 - 1 = 11$, $df_2 = n_2 - 1 = 9$.
- Two-sided F-test: We compare F to the critical values $F_{\alpha/2}(11, 9)$ and $F_{1-\alpha/2}(11, 9)$. Typically, we ensure $F \geq 1$ by assigning the larger sample variance to the numerator.
- Numerically, if $F_{\text{lower}} = F_{0.025}(11, 9) \approx 0.316$ and $F_{\text{upper}} = F_{0.975}(11, 9) \approx 3.07$ (approximate values), then 1.5 lies between 0.316 and 3.07. So we *fail to reject* H_0 at 5% level, i.e. no evidence of different variances.

Problem 7: Chi-Square Goodness-of-Fit for Failure Modes

Statement:
An engineer classifies $n = 200$ observed failures into 4 *failure modes*:

$$A,\ B,\ C,\ D.$$

The theoretical model predicts probabilities $(p_A, p_B, p_C, p_D) = (0.4, 0.3, 0.2, 0.1)$. Actual observed counts were (80, 55, 50, 15). Use a *chi-square goodness-of-fit* test at 5% significance to see if the data deviate from the predicted proportions.
Solution Sketch:

- Expected counts: $E_A = 200 \times 0.4 = 80$, $E_B = 60$, $E_C = 40$, $E_D = 20$.
- Compute $\chi^2 = \sum_i \frac{(O_i - E_i)^2}{E_i} = \frac{(80-80)^2}{80} + \frac{(55-60)^2}{60} + \frac{(50-40)^2}{40} + \frac{(15-20)^2}{20}$. So

$$\chi^2 = 0 + \frac{25}{60} + \frac{100}{40} + \frac{25}{20} = 0 + 0.4167 + 2.5 + 1.25 = 4.1667.$$

- Degrees of freedom $df = k - 1 = 4 - 1 = 3$. Critical $\chi^2_{0.95,3} \approx 7.815$. Since $4.1667 < 7.815$, we *fail to reject* the model at 5% significance.

Problem 8: Chi-Square Test for Independence in a Contingency Table

Statement:
A reliability engineer records whether a car battery fails *early* or *late* for two brands (Brand X and Brand Y). The data from 300 batteries is:

	Early Fail	Late Fail	Row Totals
Brand X	40	90	130
Brand Y	30	140	170
Column Totals	70	230	300

Use a chi-square test to check if *failure time* is independent of *brand* at $\alpha = 0.05$.
Solution Sketch:

- *Expected counts* if independent:

$$E(\text{X, Early}) = \frac{130 \times 70}{300} = 30.33, \quad E(\text{X, Late}) = \frac{130 \times 230}{300} = 99.67,$$
$$E(\text{Y, Early}) = \frac{170 \times 70}{300} = 39.67, \quad E(\text{Y, Late}) = \frac{170 \times 230}{300} = 130.33.$$

- Compute

$$\chi^2 = \sum_{\text{cells}} \frac{(O - E)^2}{E}.$$

  So

$$\chi^2 = \frac{(40 - 30.33)^2}{30.33} + \frac{(90 - 99.67)^2}{99.67} + \frac{(30 - 39.67)^2}{39.67} + \frac{(140 - 130.33)^2}{130.33}.$$

  Evaluate each term:

$$\frac{(9.67)^2}{30.33} + \frac{(-9.67)^2}{99.67} + \frac{(-9.67)^2}{39.67} + \frac{(9.67)^2}{130.33}.$$

  Numerically, it comes out around $\chi^2 \approx 7.52$ (approx).

- *df* $= (\text{rows} - 1) \times (\text{cols} - 1) = 1 \times 1 = 1$. The critical value $\chi^2_{0.95,1} \approx 3.84$. Since $7.52 > 3.84$, *reject* independence at 5% level. Failure time is not independent of brand.

Problem 9: Z-Interval for Mean Voltage (Large Sample)

Statement:
A company wants a 95% *confidence interval* for the mean voltage of a newly designed battery cell. They know from past data that the standard deviation is $\sigma = 0.5\,\text{V}$. A random sample of $n = 64$ cells gives sample mean $\overline{X} = 4.8\,\text{V}$. Compute the 95% *Z-interval* for μ.
Solution Sketch:

- Standard error of the mean: $\sigma_{\overline{X}} = \frac{0.5}{\sqrt{64}} = \frac{0.5}{8} = 0.0625$.
- 95% Z-critical value is $z_{0.025} = 1.96$.
- Confidence interval:
$$\overline{X} \pm z_{0.025}\, \sigma_{\overline{X}} = 4.8 \pm 1.96 \times 0.0625 = 4.8 \pm 0.1225.$$
So the 95% CI is $(4.6775,\ 4.9225)$.

Problem 10: Two-Sample T-Interval for Difference of Means

Statement:
Two engineering teams measure the tensile strength (in MPa) of a new alloy under different heat treatments, labeled A and B. Assume each sample is from normal populations with unknown (but possibly unequal) variances.

- Treatment A: $n_A = 14,\ \overline{X}_A = 520,\ s_A^2 = 40$.
- Treatment B: $n_B = 10,\ \overline{X}_B = 505,\ s_B^2 = 45$.

Construct a 90% *two-sample T-interval* for the difference $\mu_A - \mu_B$ assuming Welch's formula for degrees of freedom.
Solution Sketch:

- The difference $\overline{X}_A - \overline{X}_B = 520 - 505 = 15$.
- Standard error for Welch's formula:
$$SE = \sqrt{\frac{s_A^2}{n_A} + \frac{s_B^2}{n_B}} = \sqrt{\frac{40}{14} + \frac{45}{10}} = \sqrt{2.857 + 4.5} = \sqrt{7.357} \approx 2.71.$$
- Degrees of freedom (Welch's approximation):
$$df \approx \frac{\left(\frac{40}{14} + \frac{45}{10}\right)^2}{\frac{(\frac{40}{14})^2}{n_A - 1} + \frac{(\frac{45}{10})^2}{n_B - 1}}.$$
Numerically, $df \approx$ (somewhere around 15-18). One can compute precisely or estimate. Suppose it is about $df \approx 16$ for illustration.
- For a 90% two-sided interval, $t_{0.05, df\approx 16} \approx 1.746$.
- Then the interval is
$$(\overline{X}_A - \overline{X}_B) \pm t_{\alpha/2,\, df}\, SE = 15 \pm 1.746 \times 2.71 \approx 15 \pm 4.73.$$
So we get $(10.27,\ 19.73)$ as the approximate 90% CI for $\mu_A - \mu_B$.