MAT 2377 – Assignment 3 – Solutions

Deadline: Thursday Apr 08, 2021 at 3:00 pm

Total = 100 marks

1. [not marked] Let $\bar{X}$ be the mean resistance of the 50 selected resistors. Since $n = 50$ is large, then CLT tells us

$$\frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \sim N(0,1) \text{ approximately.}$$

Thus, we have

$$P(\bar{X} \geq 1005) = 1 - \Phi\left(\frac{1005 - 1000}{200/\sqrt{50}}\right) = 1 - \Phi(0.35) = 0.3632.$$

2. [30 marks]

(a) [15 marks] Since σ is unknown, we should use T-CI. A 95% confidence interval for the mean lifetime of a 75 watts lightbulb is given by

$$\bar{x} \pm b\,\frac{s}{\sqrt{n}} = [1002.3; 1025.7],$$

where $\bar{x} = 1014$, $s = 25$, $n = 20$, and $b = 2.093$ based on T_{19} distribution.

(b) [15 marks] Since $\sigma = 25$ is now known, we should use Z-CI. We want

$$2a\frac{\sigma}{\sqrt{n}} \leq 9,$$

where $a = 1.96$ such that $P(-a < Z < a) = 0.95$. Solving the above equation, we get $n = 118.5679 \approx 119$.

3. [40 marks]

(a) [10 marks] Let μ pressure require for breakage of the fiber in psi. We want to test

$$H_0 : \mu = 200 \quad \text{against} \quad H_1 : \mu > 200.$$

(b) [15 marks] The sample mean is $\bar{x} = \sum_{i=1}^{n} x_i/n = 1619/8 = 202.375$. The observed value of the z-test statistic is

$$z_0 = \frac{\bar{x} - 200}{\sigma/\sqrt{n}} = \frac{202.375 - 200}{4.5/\sqrt{8}} = 1.4928.$$

(c) [15 marks] Since it is a right-sided alternative, the p-value is give by

$$\text{p-value} = P(Z > 1.4928) = P(Z < -1.4928) \approx 0.0681.$$

(i) Since $0.0681 > 0.05$, we fail to reject H_0 at level $\alpha = 0.05$ and do not have evidence to justify H_1.
(ii) Since $0.0681 < 0.1$, we reject H_0 at level $\alpha = 0.1$ and conclude H_1 is true.

(d) [not marked] The observed value of the test statistic is

$$z_0 = \frac{\bar{x} - 200}{\sigma/\sqrt{n}} = \frac{202.375 - 200}{4.5/\sqrt{30}} = 2.8908.$$

Since it is a right-sided alternative,

$$\text{p-value} = P(Z > 2.8908) = P(Z < -2.8908) \approx 0.0019.$$

Since $0.0019 < 0.05$, we reject H_0 at level $\alpha = 0.05$ and conclude H_1 is true.

4. [40 marks]

(a) [10 marks] We want

$$P(X < 200) = \int_0^{200} \frac{3\,x^2}{(400)^3}\,dx = \left.\frac{x^3}{400^3}\right|_0^{200} = \frac{1}{8} = 0.125.$$

(b) [15 marks] The mean lifetime is

$$\mu_X = \mathrm{E}[X] = \int_0^{400} x\,\frac{3\,x^2}{(400)^3}\,dx = \left.\frac{3\,x^4}{4(400^3)}\right|_0^{400} = 300.$$

We have

$$\mathrm{E}[X^2] = \int_0^{400} x^2\,\frac{3\,x^2}{(400)^3}\,dx = \left.\frac{3\,x^5}{5(400^3)}\right|_0^{200} = 96000,$$

thus the standard deviation is

$$\sigma_X = \sqrt{\mathrm{E}[X^2] - \mu_X^2} = \sqrt{96000 - (300)^2} = 77.45967.$$

(c) [15 marks] Let $\bar{X}$ be the average lifetime of the 50 selected parts. Since $n = 50$ is large, then the CLT tells us

$$\frac{\bar{X} - \mu_X}{\sigma_X/\sqrt{n}} \sim N(0,1) \text{ approximately.}$$

We want

$$P(\bar{X} < 275) = \Phi\left(\frac{275 - 300}{77.45967/\sqrt{50}}\right) = \Phi(-2.28) = 0.0113.$$

5. [not marked]

(a) The sample mean is $\bar{x} = \sum_{i=1}^n x_i/n = 520.7/20 = 26.035$, and the sample standard deviation is

$$s = \sqrt{\frac{(\sum_{i=1}^n x_i^2) - (\sum_{i=1}^n x_i)^2/n}{n-1}} = \sqrt{\frac{(13991.41) - (520.7)^2/20}{20-1}} = 4.7848.$$

(b) The rank for the median is $(n+1)50\% = 10.5$. Thus, the median is

$$\tilde{x} = y_{10} + 0.5\,(y_{11} - y_{10}) = 26.7 + 0.5\,(26.9 - 26.7) = 26.8.$$

The rank for q_1 is $(n+1)\,25\% = 5.25$. Thus, the first quartile is

$$q_1 = y_5 + 0.25\,(y_6 - y_5) = 21.8 + 0.25\,(22.3 - 21.8) = 21.925.$$

The rank for q_3 is $(n+1)\,75\% = 15.75$. Thus, the third quartile is

$$q_3 = y_{15} + 0.75\,(y_{16} - y_{15}) = 29.2 + 0.75\,(30.7 - 29.2) = 30.325.$$

(c) The lower and upper fences are

$$\text{lower fence} = q_1 - 1.5\,\text{IQR} = 21.925 - 1.5\,(30.325 - 21.925) = 9.325$$

and

$$\text{upper fence} = q_3 + 1.5\,\text{IQR} = 30.325 + 1.5\,(30.325 - 21.925) = 42.925.$$

No values are outside the fences. So it is unlikely that there are outliers.

6. [not marked]

(a) The distribution of the calcium concentration is unimodal and skewed to the left. The distribution of mortality is approximately symmetric and unimodal.

(b) Let us consider the calcium concentration. The lower and upper fences are

$$\text{lower fence} = q_1 - 1.5\,\text{IQR} = 14 - 1.5\,(75 - 14) = -77.5$$

and

$$\text{upper fence} = q_3 + 1.5\,\text{IQR} = 75 + 1.5\,(75 - 14) = 166.5.$$

No values are outside the fences. So there are no outliers in the sample of calcium concentration.

Let us consider the mortality. The lower and upper fences are

$$\text{lower fence} = q_1 - 1.5\,\text{IQR} = 1379 - 1.5\,(1668 - 1379) = 945.5$$

and

$$\text{upper fence} = q_3 + 1.5\,\text{IQR} = 1668 + 1.5\,(1668 - 1379) = 2101.5.$$

No values are outside the fences. So it is unlikely that there are outliers in the sample.

(c) **(i)** North

(ii) South

(iii) North

(iv) The larger IQR is for the South, so the mortality is more dispersed in the South.