MH1820 Introduction to Probability and Statistical Methods

Tutorial 11 (Week 12) $\boxed{\text{Solution}}$

Problem 1 (Confidence Intervals)

A random sample of 100 automobile owners in Tekong state shows that an automobile is driven on average $23,500$ km per year with a standard deviation of 3900 km. Assume the population distribution is normal.
Construct a 99% confidence interval for the average number of km per year an automobile is driven in this state.

$\boxed{\textbf{Solution}}$ The population variance is unknown. Since the sample size $n = 100 \geq 30$, we may approximate σ by the sample standard deviation s. We have $n = 100$, $\overline{x} = 23,500$ and $s = 3900$. For a 99% confidence interval, we have $\alpha = 0.01$ and $z_{\alpha/2} = z_{0.005} = 2.575$ ($z_{0.005} = 0.257$ or 0.258 are acceptable). Thus, a 99% confidence interval is

$$\overline{x} - z_{0.005}\frac{s}{\sqrt{n}} < \mu < \overline{x} + z_{0.005}\frac{s}{\sqrt{n}}$$

$$23500 - 2.575 \cdot \frac{3900}{\sqrt{100}} < \mu < 23500 + 2.575 \cdot \frac{3900}{\sqrt{100}}$$

$$22495.75 < \mu < 24504.25$$

$\square$

Problem 2 (Confidence Intervals)

A random sample of 10 chocolate energy bar of a certain brand has, on average, 230 calories per bar, with a stantard deviation of 15 calories. Construct a 99% confidence interval for the true mean calorie content of this brand of energy bar. Assume that the population distribution of the calorie content is normal.

$\boxed{\textbf{Solution}}$ We are given $n = 10$, $\overline{x} = 230$, $s = 15$. Since $n < 30$, we have the t-distribution for $T = \frac{\overline{X}-\mu}{s/\sqrt{n}}$ with degree of freedom $n - 1 = 9$. For a 99% confidence interval, we have $\alpha = 0.01$ and $t_{\alpha/2}(n-1) = t_{0.005}(9)$. From the table, we have $t_{0.005}(9) = 3.250$. Thus, a 99% confidence interval for the mean μ is

$$\overline{x} - t_{0.005}(9)\frac{s}{\sqrt{n}} < \mu < \overline{x} + t_{0.005}(9)\frac{s}{\sqrt{n}}$$

$$230 - 3.25 \cdot \frac{15}{\sqrt{10}} < \mu < 230 + 3.25 \cdot \frac{15}{\sqrt{10}}$$

$$214.58 < \mu < 245.42.$$

$\square$

Problem 3 (Confidence Intervals)

Suppose the fat content of certain steaks follows a $N(\mu, \sigma^2)$ distribution. The following observations $x_1, \dots, x_{16}$ for the fat content are given.

$$5.33, 4.25, 3.15, 3.70, 1.61, 6.39, 3.12, 6.59, 3.53, 4.74, 0.11, 1.60, 5.49, 1.72, 4.15, 2.28$$

Suppose that both μ and σ^2 are *unknown*.

(i) Find 90%, 95%, and 99% confidence intervals for μ.

(ii) Find 90%, and 95% confidence intervals for σ^2.

$\boxed{\textbf{Solution}}$

(i) We have $n = 16 < 30$, $\overline{x} = 3.61$ and $S = 1.847$. For small sample size, we have a t-distribution for $T = \frac{\overline{X}-\mu}{S/\sqrt{n}}$ with degree of freedom $n-1$. For a $100(1-\alpha)\%$ confidence interval, let $t_{\alpha/2}(n-1)$ be the upper percentage point. A $100(1-\alpha)\%$ confidence interval for μ is given by

$$\overline{X} - t_{\alpha/2}(n-1)\frac{S}{\sqrt{n}} < \mu < \overline{X} + t_{\alpha/2}(n-1)\frac{S}{\sqrt{n}} \tag{1}$$

1. For 90% confidence interval, $\alpha = 0.1$ and $t_{\alpha/2}(n-1) = t_{0.05}(15) \approx 1.753$. By (1), a 90% confidence interval for μ is
$$[2.800, 4.419].$$

2. For 95% confidence interval, $\alpha = 0.05$ and $t_{\alpha/2}(n-1) = t_{0.025}(15) \approx 2.131$. By (1), a 95% confidence interval for μ is
$$[2.626, 4.594].$$

3. For 99% confidence interval, $\alpha = 0.01$ and $t_{\alpha/2}(n-1) = t_{0.005}(15) \approx 2.947$. By (1), a 99% confidence interval for μ is
$$[2.249, 4.971].$$

(ii) We have

$$X = \frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1).$$

For each $\alpha \in (0,1)$, let $\chi^2_\alpha(n-1)$ denote the upper percentage point. Recall that the $100(1-\alpha)$ confidence interval is given by

$$\frac{(n-1)S^2}{\chi^2_{\alpha/2}(n-1)} \leq \sigma^2 \leq \frac{(n-1)S^2}{\chi^2_{1-\alpha/2}(n-1)}. \tag{2}$$

Note that $n = 16$ and $S^2 \approx 3.412$.

1. For 90% confidence interval, we have $\alpha = 0.1$ and

$$\chi^2_{1-\alpha/2}(n-1) = \chi^2_{0.95}(15) \approx 7.261 \quad \text{and} \quad \chi^2_{\alpha/2}(n-1) = \chi^2_{0.05}(15) \approx 25.$$

   By (2), a 90% confidence interval for σ^2 is

$$[2.047, 7.050].$$

2. For 95% confidence interval, we have $\alpha = 0.05$ and

$$\chi^2_{1-\alpha/2}(n-1) = \chi^2_{0.975}(15) \approx 6.262 \quad \text{and} \quad \chi^2_{\alpha/2}(n-1) = \chi^2_{0.025}(15) \approx 27.49.$$

   By (2), a 95% confidence interval for σ^2 is

$$[1.861, 8.172].$$

Problem 4 (Hypothesis Testing)

In the journal Hypertension, researchers report that individuals who practice Transcendental Meditation (TM) lower their blood pressure significantly. If a random sample of 225 male TM practitioners meditate for 8.5 hours per week with a standard deviation of 2.5 hours, does that suggest that, on average, men who use TM meditate more than 8 hours per week? Use a significance level of $\alpha = 0.05$. State the null hypothesis, alternative hypothesis, test statistic and the conclusion.

$\boxed{\textbf{Solution}}$ Let X be the number hours per week a male TM practitioner meditate.
Sample: $n = 225$, $\overline{x} = 8.5$, $s = 2.5$.

- Null hypothesis H_0: $\mu = 8$
- Alternative hypothesis H_1: $\mu > 8$
- Level of significance: $\alpha = 0.05$, $z_\alpha = z_{0.05} = 1.645$
- Statistic: $T = \frac{\overline{X}-\mu}{S/\sqrt{n}} \sim N(0,1)$ (for large sample with unknown population variance). Based on the sample, we have (assuming H_0 is true)

$$t = \frac{8.5 - 8}{2.5/\sqrt{225}} = 3.00.$$

- p-value: (assuming H_0 is true)

$$\begin{aligned} \mathbb{P}(T \geq t) &= \mathbb{P}(T \geq 3.0) \\ &= 1 - \Phi(3) = 0.0013 \end{aligned}$$

- Conclusion: Since the p-value is less than $\alpha = 0.05$, we reject the null hypothesis.

$\square$

Problem 5 (Hypothesis Testing)

An electrical firm manufactures light bulbs that have a lifetime that is approximately normally distributed with a mean of 800 hours and a standard deviation of 40 hours. A random sample of 30 bulbs has an average life of 788 hours. Use a 0.01 level of significance to test the hypothesis that $\mu = 800$ hours against the alternative hypothesis, $\mu \neq 800$ hours.

$\boxed{\textbf{Solution}}$ Let X be the lifetime of the light bulbs. $X \sim N(\mu = 800, \sigma^2 = 40^2)$. Sample: $n = 30$, $\overline{x} = 788$

- Null hypothesis H_0: $\mu = 800$
- Alternative hypothesis H_1: $\mu \neq 800$
- Level of significance: $\alpha = 0.01$, $z_\alpha = z_{0.005} = 2.575$
- Statistic: $T = \frac{\overline{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$ (for sample with known population variance). We have $\mathbb{E}[T] = 0$. Based on the sample, we have (assuming H_0 is true)
$$t = \frac{788 - 800}{40/\sqrt{30}} = -1.64.$$
- p-value: (assuming H_0 is true)
$$\begin{aligned} \mathbb{P}(|T - \mathbb{E}[T]| \geq |t - \mathbb{E}[T]|) &= \mathbb{P}(|T - 0| \geq |-1.64 - 0|) \\ &= \mathbb{P}(|T| \geq 1.64) = \mathbb{P}(T \geq 1.64) + \mathbb{P}(T \leq -1.64) \\ &= 2 \times \mathbb{P}(T \leq -1.64) = 2\Phi(-1.64) = 2(0.0505) = 0.101 > \alpha = 0.05 \end{aligned}$$
- Conclusion: Since the p-value is more than $\alpha = 0.05$, we **do not** reject the null hypothesis. There is no strong evidence to conclude that the mean lifetime is not 800.

Problem 6 (Hypothesis Testing)

Past experience indicates that the time required for high school seniors to complete a standardized test is a normal random variable with a mean of 35 minutes. A random sample of 20 high school seniors took an average of 33.1 minutes to complete this test with a standard deviation of 4.3 minutes. Test the hypothesis that $\mu = 35$ against the alternative hypothesis that $\mu < 35$, at the 0.05 level of significance.

$\boxed{\textbf{Solution}}$ Let X be the time required for a high school senior to complete a standardized test. Given that $X \sim N(\mu = 35, \sigma^2)$, where σ is unknown.
Sample: $n = 20$, $\overline{x} = 33.1$, $s = 4.3$. Small sample size.

- Null hypothesis H_0: $\mu = 35$

- Alternative hypothesis H_1: $\mu < 800$
- Level of significance: $\alpha = 0.05$, $t_\alpha(n-1) = t_{0.05}(19) = 1.729$.
- Statistic: $T = \frac{\overline{X}-\mu}{S/\sqrt{n}}$ has t-distribution with degree of freedom $n-1$ (for samll sample size with unknown population variance). Based on the sample, we have (assuming H_0 is true)

$$t = \frac{33.1 - 35}{4.3/\sqrt{20}} = -1.976.$$

- p-value: (assuming H_0 is true)

$$\begin{aligned} \mathbb{P}(T \le t) &= \mathbb{P}(T \le -1.976) \\ &< \mathbb{P}(T \le -1.729) = \mathbb{P}(T \ge 1.729) = 0.05 = \alpha \end{aligned}$$

- Conclusion: Since the p-value is less than $\alpha = 0.05$, we **reject** the null hypothesis. There is strong evidence to conclude that the mean time required by high school senior to complete the standardized test is less than 35 minutes.

$\square$

Answer Keys. **Q1.** $22495.75 < \mu < 24504.25$ **Q2.** $214.58 < \mu < 245.42$ **Q3(i).**$[2.800, 4.419]$, $[2.626, 4.594]$, $[2.249, 4.971]$ **Q3(ii).** $[2.047, 7.050]$, $[1.861, 8.172]$ **Q4.** H_0: $\mu = 8$, H_1: $\mu > 8$,$T = \frac{\overline{X}-\mu}{S/\sqrt{n}}$, p-value is 0.0013, RejectH_0 **Q5.** p-value is 0.101, Do not reject H_0 **Q6.** p-value is less than $\alpha = 0.05$, Reject H_0.