Stat 330
HK 10

9.8, 9.9, 9.10a, *9.11a, 9.13 , z extra problems

1. (Ex 9.8)

Sol:

(a) \[ \bar{x} \pm Z_{\alpha/2} \cdot \frac{6}{\sqrt{n}} = 37.7 \pm (1.645) \times \frac{9.2}{\sqrt{100}} = 37.7 \pm 1.5134 = [36.1866, 39.2134] \]

(b) Test H0: \( \mu \leq 35 \) vs. H1: \( \mu > 35 \).

Reject H0 if the test statistic \( Z > Z_{0.01} = 2.326 \).

The observed test statistic is

\[ Z = \frac{\bar{x} - \mu_0}{\frac{6}{\sqrt{n}}} = \frac{37.7 - 35}{\frac{9.2}{\sqrt{100}}} = 2.9348 \]

belongs to the rejection region. Therefore, reject H0 in favor of H1.

Yes, these data provide significant evidence that the mean number of current users is greater than 35.

*=*
2. (Ex 9.9)

Sol:
(a) \( \bar{x} \pm Z_{0.025} \cdot \frac{6}{\sqrt{n}} = 42 \pm (1.96) \times \frac{5}{\sqrt{64}} = 42 \pm 1.225 \)
\[ \approx [40.775, \ 43.225] \]

(b) \( P \left[ 40.775 \leq x \leq 43.225 \right] = P \left[ \frac{40.775 - 40}{6} \leq \frac{x - 40}{6} \leq \frac{43.225 - 40}{6} \right] \)
\[ = P \left[ \frac{40.775 - 40}{5} \leq Z \leq \frac{43.225 - 40}{5} \right] = \Phi(0.845) - \Phi(0.155) \]
\[ = 0.7405 - 0.5616 = 0.1789 \quad \text{(using R)} \]

Note: The individual value, the time of one installation, is not very likely to belong to the computed 95% confidence interval (C.I.). The C.I. is computed for the population mean \( \mu \), not for the individual value.

3. (Ex 9.10a)

Sol:
The standard deviation is unknown. Therefore, the interval is
\[ \bar{x} \pm t_{\frac{a}{2}} \cdot \frac{s}{\sqrt{n}}, \]
where \( n = 1 - 0.90 = 0.10 \), \( n = 3 \), \( t_{\frac{a}{2}} = t_{0.05} = 2.9200 \quad \text{(with 2 df.)} \),
\( \bar{x} = \frac{30 + 50 + 70}{3} = 50 \)
\[ s = \sqrt{\frac{(30 - 50)^2 + (50 - 50)^2 + (70 - 50)^2}{3 - 1}} = \sqrt{\frac{800}{2}} = 20 \]
Then, C.I. is

\[ 50 \pm 2.7200 \times \frac{20}{\sqrt{3}} = 50 \pm 33.7173 = [16.2827, 83.7173] \]

\[ \star \star \]

4. (Ex 9.11a)

**Sol:**

\[ \hat{p} = \frac{24}{200} = 0.12 \]

For \( a = 1 - 0.96 = 0.04 \),

\[ Z_{\frac{a}{2}} = Z_{0.02} = 2.0537 \hspace{1cm} (By \ R) \]

\[ \hat{p} \pm Z_{\frac{a}{2}} \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = 0.12 \pm 2.0537 \times \sqrt{\frac{0.12 \times 0.88}{200}} = 0.12 \pm 0.0472 \]

\[ = [0.0728, 0.1672] \hspace{1cm} \star \star \]

5. (Ex 9.13)

**Sol:**

(a)

\[ \bar{x} \pm \frac{z \times 0.05}{\sqrt{n}} = 0.62 \pm (1.96) \times \frac{0.2}{\sqrt{52}} = 0.62 \pm 0.0344 \]

\[ = [0.5656, 0.6744] \hspace{1cm} \star \star \]

(b) If \( \mu = 0.6 \)

\[ P [ \bar{x} \geq 0.62 ] = P \left( \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}} \geq \frac{0.62 - \mu}{\frac{0.2}{\sqrt{52}}} \right) = P \left( Z \geq \frac{0.62 - 0.6}{\frac{0.2}{\sqrt{52}}} \right) \]

\[ = P ( Z \geq 0.7211 ) = 1 - \Phi (0.7211) = 1 - 0.7646 = 0.2354 \hspace{1cm} \star \star \]
Extra 1:
Sol:
(a) - (c): \( \mathbb{R} \in \mathbb{R} \) output.

(d)
\[
\mu_i = E(x) \\
\mu = \bar{x} = 7.9 \\
\hat{\mu}_i = 8.9
\]
\[
\sigma^2_i = \text{var}(x) \\
\sigma^2 = \frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{x})^2 = 162.2233
\]
\[
\hat{\sigma}_i^2 = 142.2233
\]
(Note the variance of sample reported by \( \bar{R} \) is \( \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})^2 \);
(\( n-1 \)) instead of \( n \) is the denominator.)

Extra 2:
Sol:
(a)
In 9.3 (e) of Homework 9, we derived MLE for \( (\mu, \sigma^2) \).
\[
\hat{\mu} = \bar{x}, \quad \hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{x})^2
\]
So using \( \bar{R} \), we find
\[
\bar{x} = 131.16, \\
\frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{x})^2 = 365.2544
\]
\[
\hat{\mu} = 131.16, \quad \hat{\sigma}^2 = 365.2544 \text{ are MLE for } \mu \text{ and } \sigma^2.
\]
(b) C.I. of \( \mu \) is:
\[
\bar{x} \pm t_{0.05} \cdot \frac{s}{\sqrt{n}} = 131.16 \pm 1.96 \times \frac{15}{\sqrt{25}} = 131.16 \pm 5.88
\]
\[
= [125.28, 137.04]
\]

(c) C.I. of \( \mu \) is:
\[
\bar{x} \pm t_{0.025} \cdot \frac{s}{\sqrt{n}}
\]
where \( n = 25 \), \( t_{0.025} = 2.0639 \) \((\text{with } 25-1 = 24 \text{ df.})\),
\[
s = 19.5057
\]
\[
\bar{x} \pm t_{0.025} \cdot \frac{s}{\sqrt{n}} = 131.16 \pm 2.0639 \times \frac{19.5057}{\sqrt{25}} = 131.16 \pm 8.0516
\]
\[
= [123.1084, 139.2116]
\]

(d), (e) see R output.
Extra Problem

1.

(a)

> x1 <- c(6, -4, -3, 10, 1, 3, 6, 11, 13, 7, -3, 7, 12, 2, -7, 8, -6, -7, 14, 13, 21, 13, 31, 14, 24, 7, 7, 21, 57, 19)

> hist(x1, breaks=14)

The distribution is skewed. It is right skewed.

(b)

> summary(x1)

Min. 1st Qu.  Median    Mean  3rd Qu.    Max.
Yes, a boxplot reports five point summary. (See (c).)

(c)

> boxplot(x1)

Yes, there is an outlier.

\[ IQR = Q3 - Q1 = 13.75 - 2.25 = 11.5 \]

\[ Q1 - 1.5 \times IQR = -15 \]

\[ Q3 + 1.5 \times IQR = 13.75 + 1.5 \times 11.5 = 31 \]

But there is one point 57, which is larger than 31. So it is an outlier.
(d) (e)

\(x_2 \leftarrow c(125, 121, 125, 137, 151, 108, 132, 134, 120, 118, 139, 124, 129, 177, 115, 106, 153, 125, 91, 126, 152, 154, 127, 122, 168)\)

\(x_3 \leftarrow \text{rnorm}(25*100, 130, 15)\)

\(m_1 \leftarrow \text{matrix}(x_3, nrow = 100)\)

\(ci \leftarrow \text{function}(v1, \text{confLevel})\{\)

\(n \leftarrow \text{length}(v1);\)

\(\text{meanV1} \leftarrow \text{mean}(v1);\)

\(\text{sdV1} \leftarrow \text{sd}(v1);\)

\(t \leftarrow \text{qt}(1 - (1 - \text{confLevel})/2, n-1);\)

\(\text{lb} \leftarrow \text{meanV1} - t * \text{sdV1}/\sqrt{n};\)

\(\text{ub} \leftarrow \text{meanV1} + t * \text{sdV1}/\sqrt{n};\)

\(\text{return}(\text{c(lb, ub)});\)\}

\(ci\text{Matrix} \leftarrow \text{apply}(m_1, 1, ci, \text{confLevel} = 0.95)\)

\(\text{sum}(ci\text{Matrix}[1,] <= 130 \& 130 <= ci\text{Matrix}[2,])\)

[1] 92

(d)

Answer: There are 92 out of 100 confidence intervals include true population mean 130.

(e)

\(\text{plot}(y=c(1,100), x=c(\text{min}(ci\text{Matrix}[1,]), \text{max}(ci\text{Matrix}[2,])), \text{type="n"}, \text{xlab="The Range of C.I.'s"}, \text{ylab=""}, \text{main="Plot of 100 C.I.'s and the true population mean 130"});\)
for(i in 1:100){
    lines(y=c(i,i), x=ciMatrix[,i]);
}

abline(v=130)

Plot of 100 C.I.'s and the true population mean 130