MAT 2377 3X (Spring 2011)

Assignment 4 - Solutions

[4] 1.

(a) We compute the expectations of the estimators :

$$E[\hat{\mu}_1] = 2\,E[X_1] - E[X_4] = 2\mu - \mu = \mu$$

and

$$\begin{aligned} E[\hat{\mu}_2] &= (5/8)\,E[X_1] + (1/8)E[X_2] + (1/8)E[X_3] + (1/8)E[X_4] \\ &= (5/8)\,\mu + (1/8)\mu + (1/8)\mu + (1/8)\mu \\ &= \mu. \end{aligned}$$

Since $E[\hat{\mu}_1] = \mu$ and $E[\hat{\mu}_2] = \mu$, then both estimators are **unbiased** for estimating μ.

(b) We compute the variance for each estimator :

$$V[\hat{\mu}_1] = 2^2\,V[X_1] + (-1)^2\,E[X_4] = 4\sigma^2 + \sigma^2 = 5\,\sigma^2$$

and

$$\begin{aligned} V[\hat{\mu}_2] &= (5/8)^2\,V[X_1] + (1/8)^2 V[X_2] + (1/8)^2 V[X_3] + (1/8)^2 V[X_4] \\ &= (5/8)^2\,\sigma^2 + (1/8)^2\sigma^2 + (1/8)^2\sigma^2 + (1/8)^2\sigma^2 \\ &= (28/64)\sigma^2. \end{aligned}$$

Since $V[\hat{\mu}_2] < V[\hat{\mu}_1]$, then the second estimator is better for estimating μ.

[3] 2. Let $\overline{X}$ be the sample of $n = 45$ observations from a population with a mean $\mu = 75$ and a standard deviation $\sigma = 1.6$. By the Central Limit Theorem,

$$\overline{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right) = N\left(75, \frac{1.6^2}{45}\right)$$

approximately, since $n = 50$ is a large sample size. The probability that the mean pull-off force of the 45 observations will be between 74.25 and 75.75 pounds is

$$\begin{aligned} P(74.25 < \overline{X} < 75.75) &\approx \Phi\left(\frac{75.75 - 75}{1.6/\sqrt{45}}\right) - \Phi\left(\frac{74.25 - 75}{1.6/\sqrt{45}}\right) \\ &= \Phi(3.14) - \Phi(-3.14) \\ &= 0.9992 - 0.0008 = 0.9984. \end{aligned}$$

[4] 3.

(a) The sample mean is

$$\overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i = \frac{19.22}{8} = 2.4025$$

and the sample standard deviation is

$$s = \sqrt{\frac{(\sum_{i=1}^{n} x_i^2) - (\sum_{i=1}^{n} x_i)^2/n}{n-1}} = \sqrt{\frac{(48.114) - (19.22)^2/8}{8-1}} = 0.52617.$$

(b) Using $\sigma = 0.52617$, and solving

$$n \geq \left(\frac{z_{.025}\,\sigma}{E}\right)^2 = \left(\frac{(1.96)(0.52617)}{0.1}\right)^2 = 106.4.$$

We will require $n = 107$ observations.

[1] 4. (a) The sample mean is 9.325 nonconforming coils and the sample median 8 nonconforming coils.

[1] (b) The range of the sample is max − min = $19-3 = 16$, the interquartile range is $q_3 - q_1 = 12 - 6 = 6$ and the standard deviation is $s = 4.485804$.

[1] (c) The outer fence for the boxplot is at $q_3 + 1.5\text{IQR} = 12 + 1.5(6) = 21$. Since the largest value in the sample is 19, then there are no outliers on the right side of the range of the sample.

The inner fence for the boxplot is at $q_1 - 1.5\text{IQR} = 6 - 1.5(6) = -3$. Since the smallest value in the sample is 3, then there are no outliers on the left side of the range of the sample.

Therefore there are no outliers.

[1] (d) The distribution is skewed to the right.

[2] (e)

i. The sample mean $\overline{x} = 9.325$ is a point estimate for μ and the sample standard deviation $s = 4.485804$ is a point estimate for σ.

ii. The estimated standard error for the estimate of the mean is

$$\hat{\sigma}_{\overline{X}} = \frac{s}{\sqrt{n}} = \frac{4.485804}{\sqrt{40}} = 0.70927.$$

iii. **Conditions :** the sample size $n = 40$ is large and σ is unknown.

A 95% confidence interval for μ is

$$\overline{x} \pm z_{.025}\frac{s}{\sqrt{n}} = 9.325 \pm 1.96\,(0.70927) = [7.93, 10.72]$$

iv. The length of the 90% CI will be smaller in comparison to the length of the 95% CI. The precision of the estimate increases as we decrease the level of confidence, since we require a smaller proportion of the confidence intervals to contain the true value of the parameter.

[4] 5.

(a) **Conditions :** normal population with σ known. We will use the following z-test statistic

$$Z_0 = \frac{\overline{X} - 12}{0.5/\sqrt{15}}.$$

Since it is a left-sided alternative, then we will reject H_0 if $z_0 < -z_{0.05} = -1.645$, where z_0 is the observed value of the test statistic. Since the critical region is

$$\frac{\overline{x} - 12}{0.5/\sqrt{15}} = z_0 < -1.645,$$

we can rewrite it in terms of $\overline{x}$. We will reject H_0 if

$$\overline{x} < -1.645\left(\frac{0.5}{\sqrt{15}}\right) + 12 = 11.78763.$$

(b) The probability of committing an error of type II if in reality $\mu = 11.5$ is

$$\begin{aligned}\beta(11.5) &= P\left(\text{not falling in the C.R.}|\mu = 11.5\right)\\ &= P\left(\overline{X} \geq 11.78763|\mu = 11.5\right)\\ &= 1 - \Phi\left(\frac{11.78763 - 11.5}{0.5/\sqrt{25}}\right)\\ &= 1 - \Phi(2.23)\\ &= 1 - 0.9871 = 0.0129.\end{aligned}$$

Alternative Solution : If $\mu = 11.5$, then the sampling distribution of the z-test statistic is

$$Z_0 \sim N\left(\frac{\sqrt{n}(\mu_1 - \mu_0)}{\sigma}, 1\right) = N\left(\frac{\sqrt{15}(11.5 - 12)}{0.5}, 1\right) = N(-3.87298, 1).$$

Hence, The probability of committing an error of type II if in reality $\mu = 11.5$ is

$$\begin{aligned}\beta(11.5) &= P\left(\text{not falling in the C.R.}|\mu = 11.5\right)\\ &= P\left(Z_0 \geq -1.645|\mu = 11.5\right)\\ &= 1 - \Phi\left(\frac{-1.645 - (-3.87298)}{\sqrt{1}}\right)\\ &= 1 - \Phi(2.23)\\ &= 1 - 0.9871 = 0.0129.\end{aligned}$$

[1] 6. (a) A point estimate for the mean syrup content in a beverage is

$$\hat{\mu} = \overline{x} = 1.05 \ \text{ fluid ounces,}$$

and its estimated standard error is

$$\hat{\sigma}_{\overline{X}} = \frac{s}{\sqrt{n}} = \frac{0.15}{\sqrt{25}} = 0.03 \ \text{ fluid ounces.}$$

[2] (b) **Conditions :** normal population with σ unknown.

A 95% confidence interval for the mean syrup content is

$$\overline{x} \pm t_{.025,24} \frac{s}{\sqrt{n}} = 1.05 \pm 2.064\,(0.03) = [0.99, 1.11].$$

[2] (c) **Conditions :** normal population with σ unknown. So we are performing a t-test. The test statistic is

$$T_0 = \frac{\overline{X} - 1}{S/\sqrt{25}}.$$

Since it is a two-sided alternative, then we will reject H_0 if

$$|t_0| > t_{0.025,24} = 2.064,$$

where t_0 is the observed value of the test statistic. The observed value of the test statistic is

$$t_0 = \frac{1.05 - 1}{0.15/\sqrt{25}} = 1.6667.$$

Conclusion : Since $|t_0| = 1.667 < 2.064$, then we cannot reject H_0. At a level of significance of $\alpha = 5\%$, the data do not suggest that $\mu \neq 1$.

[1] (d) Since it is a two-sided test, then the p-value is

$$P = 2\,P(T_0 > |1.6667|) = 2\,P(T_0 > 1.6667),$$

where T_0 has a t distribution with $\nu = n - 1 = 24$ degrees of freedom. From Table V, we get

$$0.05 < P(T_0 > 1.6667) < 0.10.$$

Thus, the p-value falls in the following interval :

$$0.10 < P < 0.20.$$

[/**27**]