Homework5

PS 2010

Notice: The grading will mainly focus on the steps not the final results. This means even your final results are wrong but you follow the correct steps, you will be given the full score.

Random Sample and Sampling Distribution

1. Sample mean: $\bar{x} = \frac{7+6+8+4+2+7+6+7+6+5}{10} = 5.8$. Sample variance: $s^2 = \frac{1}{9}\left[(7-5.8)^2 + (6-5.8)^2 + \cdots + (5-5.8)^2\right] = \frac{1}{9}(27.6) = 3.067$. Sample sd: $s = \sqrt{3.067} = 1.75$

2. In this question, because we know the population variance, and the population is normally distributed, we can use z statistic. We need to transform to standard normal and then calculate the probability. $P(-3 \leq X - \mu \leq 3) = P(-3\sqrt{9}/4 \leq \sqrt{n}(\bar{X} - \mu)/4 \leq 3\sqrt{9}/4) = P(-2.25 \leq \sqrt{n}(\bar{X} - \mu)/\sigma \leq 2.25) = 0.988 - 0.012 = 0.976$.

3.

$$
\begin{aligned}
P\left(-0.5 \leq \bar{X}_n - \mu \leq 0.5\right) &= P\left(-\frac{0.5\sqrt{n}}{\sigma} \leq \frac{\sqrt{n}\left(\bar{X}_n - \mu\right)}{\sigma} \leq \frac{0.5\sqrt{n}}{\sigma}\right) \\
&= P\left(-2 \leq \frac{\sqrt{n}\left(\bar{X}_n - \mu\right)}{\sigma} \leq 2\right) \\
&\approx P(-2 \leq Z \leq 2\ ) \text{ where } Z \sim \mathcal{N}(0,1) \\
&\approx 0.9544
\end{aligned}
$$

4.

1. $\mu_{\bar{x}} = \mu = 112$ and $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{40}{\sqrt{50}} = 5.65685$
2. $$
\begin{aligned}
P(110 < \bar{X} < 114) &= P\left(\frac{110 - \mu_{\bar{x}}}{\sigma_{\bar{x}}} < Z < \frac{114 - \mu_{\bar{x}}}{\sigma_{\bar{x}}}\right) \\
&= P\left(\frac{110 - 112}{5.65685} < Z < \frac{114 - 112}{5.65685}\right) \\
&= P(-0.35 < Z < 0.35) = 0.6368 - 0.3632 = 0.2736
\end{aligned}
$$

3.

$$\begin{aligned}
P(\bar{X} > 113) &= P\left(Z > \frac{113 - \mu_{\bar{x}}}{\sigma_{\bar{x}}}\right) \\
&= P\left(Z > \frac{113 - 112}{5.65685}\right) \\
&= P(Z > 0.18) \\
&= 1 - P(Z < 0.18) = 1 - 0.5714 = 0.4286
\end{aligned}$$

5.

We first calculate the expected value of each estimator.

$$\begin{aligned}
&\mathrm{E}\left[\hat{\mu}_1\right] = \mathrm{E}\left[\frac{X_1 + X_2 + X_3}{3}\right] = \frac{1}{3}(\mu + \mu + \mu) = \mu \\
&\mathrm{E}\left[\hat{\mu}_2\right] = \mathrm{E}\left[\frac{X_1}{4} + \frac{X_2 + \cdots + X_{n-1}}{2(n-2)} + \frac{X_n}{4}\right] = \frac{\mu}{4} + \frac{(n-2)\mu}{2(n-2)} + \frac{\mu}{4} = \frac{\mu}{4} + \frac{\mu}{2} + \frac{\mu}{4} = \mu \\
&\mathrm{E}\left[\hat{\mu}_3\right] = \mathrm{E}[\bar{X}] = \mu
\end{aligned}$$

Thus we see that each estimator is unbiased. Next we calculate the variance of each estimator.

$$\begin{aligned}
&\mathrm{Var}\left[\hat{\mu}_1\right] = \mathrm{Var}\left[\frac{X_1 + X_2 + X_3}{3}\right] = \frac{1}{9}\left(\sigma^2 + \sigma^2 + \sigma^2\right) = \frac{\sigma^2}{3} \\
&\mathrm{Var}\left[\hat{\mu}_2\right] = \mathrm{Var}\left[\frac{X_1}{4} + \frac{X_2 + \cdots + X_{n-1}}{2(n-2)} + \frac{X_n}{4}\right] = \frac{\sigma^2}{16} + \frac{(n-2)\sigma^2}{4(n-2)^2} + \frac{\sigma^2}{16} = \frac{n\sigma^2}{8(n-2)} \\
&\mathrm{Var}\left[\hat{\mu}_3\right] = \mathrm{Var}[\bar{X}] = \frac{\sigma^2}{n}
\end{aligned}$$

Since each estimator is unbiased, and for any estimator

$$\mathrm{MSE}[\hat{\theta}] = \mathrm{E}\left[(\hat{\theta} - \theta)^2\right] = (\mathrm{bias}[\hat{\theta}])^2 + \mathrm{var}[\hat{\theta}]$$

we see that the resulting mean squared errors are simply the variances.

$$\begin{aligned}
\mathrm{MSE}\left[\hat{\mu}_1\right] &= \frac{\sigma^2}{3} \\
\mathrm{MSE}\left[\hat{\mu}_2\right] &= \frac{n\sigma^2}{8(n-2)} \\
\mathrm{MSE}\left[\hat{\mu}_3\right] &= \frac{\sigma^2}{n}
\end{aligned}$$

Bias and Consistency

6.

First, note that $\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$ is an unbiased estimator as we have seen many times before.

$$\mathrm{E}\left[\bar{X}_n\right] = \mathrm{E}\left[\sum_{i=1}^{n} X_i\right] = \frac{1}{n} \cdot n\theta = \theta$$

Also, as we have seen before, the variance of $\bar{X}_n$ is given by

$$\mathrm{Var}\left[\bar{X}_n\right] = \frac{\sigma^2}{n} = \frac{1}{n}$$

Since we have an unbiased estimator, we simply need for the variance to vanish as n goes to infinity. We see that

$$\lim_{n\to\infty} \operatorname{Var}\left[\bar{X}_n\right] = \lim_{n\to\infty} \frac{1}{n} = 0$$

Thus $\bar{X}_n$ is a consistent estimator for θ.

Confidence Interval in Large and Small Sample

7.

1) Here, we have a "small" n, σ unkown, and a sample from a normal distribution, so we use the confidence interval

$$\bar{x} \pm t_{\alpha/2,n-1}\frac{s}{\sqrt{n}}$$

We have,

- $n = 6$
- $\bar{x} = 21.4$
- $s = 0.8$
- $1 - \alpha = 0.95$, so $\alpha/2 = 0.025$
- $t_{\alpha/2,n-1} = t_{0.025,5} = 2.571$

Therefore,

$$21.4 \pm 2.571\frac{0.8}{\sqrt{6}}$$

2) Here, we have a "large" n, σ unkown, and a sample from a normal distribution, so we use the confidence interval

$$\bar{x} \pm z_{\alpha/2}\frac{s}{\sqrt{n}}$$

since for "large" n normal is a good approximation of t.

We have,

- $n = 42$
- $\bar{x} = 17.2$,
- $s = 8$,
- $1 - \alpha = 0.90$, so $\alpha/2 = 0.05$,
- $z_{\alpha/2} = z_{0.05} = 1.645$.

Therefore,

$$17.2 \pm 1.645\frac{8}{\sqrt{42}}$$

8.

Here, we have a "small" n, σ unkown, and a sample from a normal distribution, so we use the confidence interval

$$\bar{x} \pm t_{\alpha/2,n-1}\frac{s}{\sqrt{n}}$$

We have,

- $n = 25$,
- $\bar{x} = 56$,
- $s = 8$,
- $1 - \alpha = 0.99$, so $\alpha/2 = 0.005$,
- $t_{\alpha/2,n-1} = t_{0.005,24} = 2.797$.

$$56 \pm 2.797 \frac{8}{\sqrt{25}}$$

9.

$$\hat{p} = \frac{x}{n} = \frac{397}{749} = 0.53$$
$$z_{\alpha/2} = z_{0.05} = 1.645$$
$$\hat{p} \pm z_{\alpha/2} \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$
$$0.53 \pm 1.645 \sqrt{\frac{0.53(1-0.53)}{749}}$$