Homework 2 Solution

July 9, 2014

1 Ex. 7. Ch. 2

$$\bar{x} = \frac{67}{91}$$
$$\hat{se}(\bar{x}) = \frac{\sqrt{\bar{x}(1-\bar{x})}}{\sqrt{n-1}}\sqrt{\frac{N-n}{N}} = 0.039$$
$$CI = [\bar{x} - 1.96 \times se(\bar{x}), \bar{x} + 1.96 \times se(\bar{x})] = [0.660, 0.813]$$

2 Ex. 10. Ch. 2

2.1 a.

By CLT, for large n, the probability distribution of

$$Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}}$$

is approximately standard normal. So the 95% CI for μ is $[\bar{x} - 1.96 \times \sigma/\sqrt{n}, \bar{x} + 1.96 \times \sigma/\sqrt{n}]$ whose width is $4\sigma/\sqrt{n}$. If the width is less than 4%, then

$$n = 10^4\sigma^2$$

Ignore FPC. Use the sample percentage $\bar{x}$ to estimate the population variance: $\hat{\sigma}^2 = \bar{x}(1-\bar{x})$. Plug in, then we get

$$n = 10^4\bar{x}(1-\bar{x})$$

By the sample data, $\bar{x} = \frac{67}{91}$, then $n = 1941$.

2.2 b.

$\pi = 0.5$ makes the population variance be the largest, which is $0.5 \times (1 - 0.5) = 0.25$.

$$n = 10^4 \times 0.25 = 2500$$

3 Ex. 11. Ch. 2

Use the sample mean to estimate the population mean. Ignore FPC. The estimate of SE equals to

$$\hat{se}(\bar{x}) = \sqrt{\frac{\bar{x}(1-\bar{x})}{n-1}}$$

.

3.1 a.

Let $\hat{se}(\bar{x})$ be less than 1%. If $\bar{x}$ is roughly 50%, then $n \approx 2500$ and if $\bar{x}$ is roughly 10%, then $n \approx 900$.

3.2 b.

Let $\hat{se}(\bar{x})$ be less roughly than 10% of the population parameter. For $\bar{x}$ equals roughly to 50%, $n = 400\bar{x}(1-\bar{x}) \approx 100$. For $\bar{x}$ equals roughly to 10%, $n = 10^4\bar{x}(1-\bar{x}) \approx 900$.

4 Ex. 13. Ch. 2

4.1 a.

Let $N = 314$ be the number of whole population and $N_w = 131$. The population fraction is

$$\begin{aligned} \pi &= \frac{\sum_{i=1}^{N} \mathbb{1}\{x_i \text{ is woman and played video games}\}}{N_w} \\ &= \frac{\sum_{i=1}^{N} \mathbb{1}\{x_i \text{ is woman and played video games}\}}{N} \frac{N}{N_w} \end{aligned}$$

The expected value of $\hat{\pi}$ is

$$\mathbb{E}[\hat{\pi}] = \mathbb{E}[\bar{v}]\frac{N}{N_w}$$

where,

$$\begin{aligned} \mathbb{E}[\bar{v}] &= \mathbb{E}\left[\frac{\sum_{i=1}^{n} \mathbb{1}\{x_i \text{ is woman and played video games}\}}{n}\right] \\ &= P(x_i \text{ is woman and played video games}) \\ &= \frac{\sum_{i=1}^{N} \mathbb{1}\{x_i \text{ is woman and played video games}\}}{N} \end{aligned}$$

Then, we have $\mathbb{E}[\hat{\pi}] = \pi$, which implies that $\hat{\pi}$ is an unbiased estimator.

4.2 b

$$\hat{se}(\hat{\pi}) = \sqrt{\frac{\hat{\pi}(1-\hat{\pi})}{n-1}}\sqrt{\frac{N-n}{N}}$$

where $N = 314$, $n = 91$. Plug in, then we get $\hat{se}(\hat{\pi}) = \sqrt{0.0189\bar{v}(1-2.396\bar{v})}$
Note that it's unable to specify both $\hat{\pi}$ and $\bar{v}$ unless the number of woman played video games in the sample is available.

5 Ex. 15. Ch. 2

Let a random variable Y to be the number of female students in the sample and random variable $X_i = \mathbb{1}(x_i$ is a woman). Then we could write Y as

$$Y = \frac{\sum_{i=1}^{n} X_i}{n}$$

As $\mathbb{E}[X_i] = 131/314 = 0.4172$ and $Var(X_i) = \frac{131}{314}\left(1 - \frac{131}{314}\right)\frac{223}{314} = 0.1732$. By CLT, the distribution of $\frac{Y-\mathbb{E}[X_i]}{\sigma/\sqrt{n}}$ goes to the standard normal distribution as $n \longrightarrow \infty$. Using the normal approximation to compute the probability of having 38 female,

$$P(37.5 \leq Y < 38.5) = \Phi(0.1347) - \Phi(-0.117) = 0.10014$$

whereas the exact calculation in the text is 0.1003. The two answers are nearly the same and it indicates that CLT is a good approximation when sample size $n = 91$.

6 Ex. 17. Ch. 2

6.1 a.

False. A confidence interval does not predict that the true value of the parameter has a particular probability of being in the confidence interval given the data actually obtained.

6.2 b.

False. What we could say is that the sample percentage has 0.95 probability within two standard error of μ.

6.3 c.

False. The object we consider for the interval is the population percentage not the students.

6.4 d.

True.

7 Ex 18. Ch. 2

7.1 a.

$$\mathbb{E}[\hat{x}] = \mathbb{E}[x_{I(1)}] = \mu$$
$$Var(\hat{x}) = Var(x_{I(1)}) = \sigma^2$$

7.2 b.

$$\mathbb{E}[\tilde{x}] = \mathbb{E}[2x_{I(1)} - x_{I(2)}] = 2\mathbb{E}[x_{I(1)}] - \mathbb{E}[x_{I(2)}] = 2\mu - \mu = \mu$$
$$Var(\tilde{x}) = Var(2x_{I(1)} - x_{I(2)}) = 2^2 Var(x_{I(1)}) + (-1)^2 Var(x_{I(1)}) = 4\sigma^2 + \sigma^2 = 5\sigma^2$$

7.3 c.

$$\mathbb{E}[x^*] = \mathbb{E}[2\bar{x}] = 2\mathbb{E}[\bar{x}] = 2\mu$$
$$Var(x^*) = Var(2\bar{x}) = 2^2 \frac{Var(x)}{n} = \frac{4}{n}\sigma^2$$

8 Ex. 19. Ch. 2

$\hat{x}$ has the weights $(w_1, w_2, \cdots, w_n) = (1, 0, \cdots, 0)$, $\tilde{x}$ has the weights $(w_1, w_2, \cdots, w_n) = (2, -1, 0, \cdots, 0)$ and x^* has the weights $(w_1, w_2, \cdots, w_n) = (\frac{2}{n}, \frac{2}{n}, \cdots, \frac{2}{n})$. So all of them are special forms of $\bar{x}_w$.
We know that

$$\begin{aligned} \mathbb{E}[\bar{x}_w] &= \mathbb{E}\Big[\sum_{i=1}^{n} w_i x_{I(i)}\Big] \\ &= \sum_{i=1}^{n} w_i \mathbb{E}[x_{I(i)}] \\ &= \mu \sum_{i=1}^{n} w_i \end{aligned}$$

So the expected value of $\bar{x}_w$ is the same as μ only if $\sum_{i=1}^{n} w_i = 1$.