Tutorial 5: Solutions

February 14, 2018

Question 7.1.6, Page 285

On the basis of extensive tests, the yield point of a particular type of mild steel-reinforcing bar is known to be normally distributed with $\sigma = 100$. The composition of bars has been slightly modified, but the modification is not believed to have affected either the normality or the value of σ.

(a) Assuming this to be the case, if a sample of 25 modified bars resulted in a sample average yield point of 8439 lb, compute a 90% CI for the true average yield point of the modified bar.

Our sample is normal with known sigma, $\sigma = 100$.
$n = 25$, $\alpha = 0.10 \Rightarrow 1 - \frac{\alpha}{2} = 0.95$. A 90% CI is:

$$
\begin{aligned}
&\overline{x} \pm (z_{0.95}) \left(\frac{\sigma}{\sqrt{n}}\right) \\
&= 8439 \pm (1.6449) \left(\frac{100}{5}\right) \\
&= 8439 \pm 32.898 \\
&= (8406.102\ ,\ 8471.898)
\end{aligned}
$$

(b) How would you modify the interval in part (a) to obtain a confidence level of 92%?

$\alpha = 0.08 \Rightarrow 1 - \frac{\alpha}{2} = 0.96$
Therefore all we need to do is change $z_{0.95}$ to $z_{0.96}$. What do we notice? If $\overline{x}$, σ, and n stay constant, since $z_{0.96} > z_{0.95}$, increasing the confidence level of our CI increases the width of the CI!

Question 7.1.10, Page 285

A random sample of $n = 15$ heat pumps of a certain type yielded the following observations on lifetime (in years):

2.0	1.3	6.0	1.9	5.1	0.4	1.0	5.3
15.7	0.7	4.8	0.9	12.2	5.3	0.6	

(a) Assume that the lifetime distribution is exponential and use an argument parallel to that of Example 7.5 to obtain a 95% CI for expected (true average) lifetime.

We will use an argument somewhat parallel to Example 7.5. The textbook uses exponential pdf of form:

$$f(x;\lambda) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0, \lambda > 0 \\ 0 & \text{otherwise} \end{cases}$$

Where $\mathbf{E}(X) = \frac{1}{\lambda}$.

However, Assignment 3 Question 2a implies a pdf of form:

$$f(x;\lambda) = \begin{cases} \dfrac{1}{\lambda} e^{-x/\lambda} & x \geq 0, \lambda > 0 \\ 0 & \text{otherwise} \end{cases}$$

Where $\mathbf{E}(X) = \lambda$.

For this question I will use the pdf similar to the one required for your assignment. It is still similar to Example 7.5 with a small switch of λ for $1/\lambda$. We are given from the example (and from the assignment) that $\dfrac{2n\overline{X}}{\lambda} \sim {\chi^2}_{2n}$. We derive a **95**% CI for a sample of $n = 15$:

$$\mathbf{P}\left({\chi^2}_{2n,1-\frac{\alpha}{2}} < \frac{2n\overline{X}}{\lambda} < {\chi^2}_{2n,\frac{\alpha}{2}}\right) = 1-\alpha$$

$$\mathbf{P}\left({\chi^2}_{30,0.975} < \frac{2\cdot 15\cdot \overline{X}}{\lambda} < {\chi^2}_{30,0.025}\right) = 0.95$$

$$\mathbf{P}\left(\frac{16.791}{2\cdot 15\cdot \overline{X}} < \frac{1}{\lambda} < \frac{46.976}{2\cdot 15\cdot \overline{X}}\right) = 0.95$$

$$\mathbf{P}\left(\frac{30\cdot \overline{X}}{16.791} > \lambda > \frac{30\cdot \overline{X}}{46.976}\right) = 0.95$$

$$\mathbf{P}\left(\frac{30\cdot \overline{X}}{46.976} < \lambda < \frac{30\cdot \overline{X}}{16.791}\right) = 0.95$$

Therefore a 95% CI for λ with a sample size of $n = 15$ is:

$$\left(\frac{30\overline{X}}{46.976}, \frac{30\overline{X}}{16.791}\right).$$

Using the given data, we compute $\overline{x} = 63.2/15$. Thus our 95% CI is (2.691, 7.528).

(b) How should the interval of part (a) be altered to achieve a confidence level of 99%?

The general form of the 100(1 - α)% CI for $n = 15$ is:

$$\left(\frac{30\overline{X}}{{\chi^2}_{30,\frac{\alpha}{2}}} , \frac{30\overline{X}}{{\chi^2}_{30,1-\frac{\alpha}{2}}} \right)$$

If we now seek a 99% CI, then $\alpha = 0.01$ so $\frac{\alpha}{2} = 0.005$ and $1 - \frac{\alpha}{2} = 0.995$.
${\chi^2}_{30,0.005} = 53.672$, ${\chi^2}_{30,0.995} = 13.787$.

(c) What is a 95% CI for the standard deviation of the lifetime distribution? [Hint: What is the standard deviation of an exponential random variable?]

Recall that the variance of an exponential random variable is λ^2. Hence the standard deviation is λ. A CI for the standard deviation will yield the same interval as (a)!!

Question 7.2.14, Page 293

Some super long story... Among the children in the study, 514 came from households that used coal for cooking or heating or both. Their FEV1 mean was 1427 with a standard deviation of 325.

(a) Calculate and interpret a 95% (two-sided) CI for true average FEV1 level in the population of all children from which the sample was selected. Does it appear that the parameter of interest has been accurately estimated?

$n = 514 > \mathbf{40}$ so by CLT, the sample mean is approximately normally distributed. It was pointed out by your classmate that we actually need $n > 40$ due to the added variability of using s instead of σ. In addition, s will approximate σ well. We proceed with a large sample CI:
$z_{1-\frac{\alpha}{2}} = z_{0.975} = 1.96$

$$\begin{aligned} &\overline{x} \pm (z_{1-\frac{\alpha}{2}}) \left(\frac{s}{\sqrt{n}} \right) \\ &= 1427 \pm 1.96 \frac{325}{\sqrt{514}} \\ &= 1427 \pm 28.097 \\ &= (1398.90 \text{ , } 1455.10) \end{aligned}$$

The parameter of interest has been accurately estimated as we have a very small width. We conclude with **approximately** 95% confidence that the true FEV1 mean among these children lies in the above interval.

(b) Suppose the investigators had made a rough guess of 320 for the value of s before collecting data. What sample size would be necessary to obtain an interval width of 50 ml for a confidence level of 95%?

Suppose that $s = 320$. Then

$$w = \frac{2 \cdot z_{0.975} \cdot s}{\sqrt{n}}$$
$$n = \left(\frac{2 \cdot 1.96 \cdot 320}{50} \right)^2 = 629.407 \rightarrow 630$$

The investigators will need to sample 630 children.

Question 7.2.20, Page 294

TV advertising agencies face increasing challenges in reaching audience members because viewing TV programs via digital streaming is gaining in popularity. The Harris poll reported on November 13, 2012, that 53% of 2343 American adults surveyed said they have watched digitally streamed TV programming on some type of device.

(a) Calculate and interpret a CI at the 99% confidence level for the proportion of all adult Americans who watched streamed programming up to that point in time.

We are given that $\widehat{p} = 0.53$. $n = 2343 > 40$, $n\widehat{p} = 1241.79 > 10$, $n\widehat{q} = 1101.21 > 10$. Thus we can apply the large sample CI for proportions.
$\alpha = 0.01 \Rightarrow z_{1-\frac{\alpha}{2}} = z_{0.995} = 2.576$.

Using the super long formula given by (7.10) on page 289, we obtain:

$$\frac{0.5314 \pm 0.0266}{1.0028} = (0.5034\ ,\ 0.5564)$$

Thus we conclude with **approximately** 99% confidence that the true value of p will lie between 0.5034 and 0.5564.

(b) What sample size would be required for the width of a 99% CI to be at most 0.05 irrespective of the value of $\widehat{p}$?

We consider that $\widehat{p}\widehat{q} = \widehat{p}(1-\widehat{p})$ is a quadratic function of $\widehat{p}$ which is maximized at $\widehat{p} = 0.5$. So we take $\widehat{p}$ and $\widehat{q}$ to both be 0.5 and use formula (7.12). This gives us:

$$n = 2647.675 \to 2648$$

So we will need to sample 2648 people to obtain a 99% CI with the specified width that is irrespective of $\widehat{p}$.

Question 7.3.32, Page 302

Investigators developed a new test that adds cyclic strain to a level well below breakage and determines the number of cycles to break a condom. A sample of 20 condoms of one particular type resulted in a sample mean number of 1584 and a sample standard deviation of 607. Calculate and interpret a CI at the 99% confidence level for the true average number of cycles to break. [Note: The article presented the results of hypothesis tests based on the t distribution; the validity of these depends on assuming normal population distributions.]

We assume normality as recommended in the question.
$\alpha = 0.01 \Rightarrow \frac{\alpha}{2} = 0.005$. $t_{0.005,19} = 2.861$

$$\begin{aligned} &\overline{x} \pm t_{\frac{\alpha}{2},n-1}\frac{s}{\sqrt{n}} \\ &= 1584 \pm 2.861\frac{607}{\sqrt{20}} \\ &= (1195.68, 1972.32) \end{aligned}$$

We conclude with 99% confidence that the true average number of cycles to break lies within the above interval.