Hw_2: Ch 6: Ex 32, 34, Ch 7: Ex 2, 6, 16, and 20

Julia Andrus

Chp 6: 32, 34

32a) Let $X_1, ...X_n$ be a random sample from a uniform distribution on $[0, \theta]$. Then the mle of θ is $\hat{\theta} = Y = \max(X_i)$. Use the fact that $Y \leq y$ iff each $X_i \leq y$ to the derivative the cdf of Y. Then show that the pdf of $Y = \max(X_i)$ is:

$$f_y(y) = \begin{cases} \frac{ny^{n-1}}{\theta^n} & 0 \leq y \leq \theta \\ 0 & \text{otherwise} \end{cases}$$

cdf:

$$F_Y(y) = P(Y \leq y) = P(max(X_i) \leq y$$

all the X_i are small, if max is small:

$$= P(X_1 \leq y * ....X_n \leq y)$$

they are independent:

$$= P(X_1 \leq y) * ....(X_n \leq y)$$

cdf of uniform distribution on $[0, \theta]$:

$$= (\frac{y}{0})^n, 0 \leq y \leq \theta$$

Now taking the derivative:

$$\frac{d}{df_Y} = \begin{cases} \frac{ny^{n-1}}{\theta^n} & 0 \leq y \leq \theta \\ 0 & \text{otherwise} \end{cases}$$

32b) Use the result of part (a) to show that mle is biased but that (n + 1)max(X_i)/n is unbiased.

For the estimator ($E(y) = \theta$) to be unbiased, then:

$$E(Y) = \int_0^\theta y * \frac{ny^n - 1}{\theta^n} dy$$

$$= \frac{n}{\theta^n} \frac{y^n + 1}{n+1} |_0^\theta$$

$$= \frac{n}{n+1} \theta$$

As you can see that $E(Y) \neq \theta$, and $\therefore$ the estimator is biased.

To show that (n + 1)max(X_i)/n is unbiased:

Let $\tilde{Y} = \frac{n+1}{n}$, then

$$E(\tilde{Y}) = E\frac{(n+1)}{n}$$

Using result from (a):

$$= E(Y)\frac{n+1}{n}$$

$$= \theta \frac{n}{n+1} \frac{n+1}{n}$$

Cancel:

$$= \theta$$

This shows that $\tilde{Y}$ is unbiased.

34) The mean squared error of an estimator $\hat{\theta}$ is MSE $(\hat{\theta}) = E(\hat{\theta} - \hat{\theta})^2$. Consider $\hat{\sigma}^2 = KS^2$, where S^2= sample variance. If $\hat{\theta}$ is unbiased, then MSE $(\hat{\theta}) = V(\hat{\theta}) + (bias)^2$. What value of K minimizes the mean squared error of this estimator when the population distribution is normal?

First we have to find the mean squared error of $\hat{\sigma}^2 = KS^2$. Since S^2 is our estimator, then:

$$MSE(\hat{\sigma}^2) = V(\hat{\sigma}^2) + (bias)(\sigma^2)$$

$Bias(\sigma^2) = E(\hat{\sigma}^2 - \hat{\sigma}^2)^2 = E(KS^2) - \sigma^2 = K\sigma^2 - \sigma^2 = \sigma^2(k-1)$.

Leta find the variance given that S^2 is unbiased of σ^2:

$V(\hat{\sigma 2}) = V(KS^2)$

$= K^2V(S^2) = K^2(E((S^2)^2 - (E(S^2))^2)$

$= K^2(\frac{n+1}{n-1}\sigma^4 - (\sigma^2)^2)$

$= K^2\sigma^4(\frac{n+1}{n-1} - 1)$

Now to find the value of K that minimize the MSE, we need to take the derivative of the MSE:

$MSE(\sigma^2) = K^2\sigma^4(\frac{n+1}{n-1} - 1) - (\sigma^2(k-1))^2$

$= K^2\sigma^4(\frac{n+1}{n-1} - 1) - \sigma^4(k-1)^2$

Now doing the derivative:

$\frac{d}{dK}MSE(\sigma^2) = \frac{d}{dK}(K^2\sigma^4(\frac{n+1}{n-1} - 1) - \sigma^4(k-1)^2)$

$= 2K\sigma^4(\frac{n+1}{n-1} - 1) - 2\sigma^4(K-1)$

Now equal it to 0 so we can find the minimizer of the MSE:

$2K\sigma^4(\frac{n+1}{n-1} - 1) - 2\sigma^4(K-1) = 0$

$2K\sigma^4(\frac{n+1}{n-1} - 1) = 2\sigma^4(K-1)$

$K(\frac{n+1}{n-1} - 1) = (K-1)$

.

.

.

$K = \frac{n-1}{n-3}$

Answer: $K = \frac{n-1}{n-3}$

Ch 7: Ex 2, 6, 16, and 20

2a) Each of the following is a confidence interval for μ=true average (i.e., population mean) resonance frequency (Hz) for all tennis rackets of a certain type. (114.4 , 115.6) and (114.1 , 115.9). What is he value of the sample mean resonance frequency?

To find the sample mean, we just

$$\overline{X} = \frac{x_1 + x_2 ... x_n}{n}$$

So for (114.4 , 115.6):

$$\overline{X} = \frac{114.4 + 115.6)}{2} = 115$$

For (114.1 , 115.9):

$$\overline{X} = \frac{114.1 + 115.9)}{2} = 115$$

Answer: Mean for both intervals is 115

2b) Both intervals were calculated from the same sample data. The Confidence level for one of these intervals of 90% and the other has 99%. Which has 90%?

In class we talked about how the length of confidence interval depends on couple of points. One of the points mentioned about the sample size (n): the larger the value, the shorter the length of confidence interval. So for 90% C.I, the interval would have to be narrower. If we have a wider interval, then our C.I. would be higher. With that being said, I believe that (114.4 , 115.6) has 90% C.I.

6a) Mild-reinforcing bar is known to be normally distributed with $\sigma = 100$. The composition of bars has been slightly modified, but does not affect the normality or value of σ. If n=25, $\overline{X} = 8439$lb, compute a 90% CI for the true average yield point of the modified bar.

First we have to find what the area for a 90% interval:

$$A = \frac{1 + .90}{2} = .95$$

Now using the appendix in the back of the book to find z-score:

$$.95 = 1.64$$

$$z = 1.64$$

Now we have to use this formula to compute the intervals: $\overline{X} + \pm z(\frac{\sigma}{n})$

n=25
$\overline{X} = 8439$lb
$\sigma = 100$
z=1.64
So,

$$\overline{X} + \pm z(\frac{\sigma}{n})$$
$$8439 + \pm 1.64(\frac{100}{25})$$
$$(8439 - 1.64(\frac{100}{25}), 8439 + 1.64(\frac{100}{25}))$$
$$(8432.44, 8445.56)$$

Answer: (8432.44, 8445.56)

6b) How would you modify the interval in part (a) to obtain a confidence level of 92%?

The interval would have to a a bit wider to get a higher CI. So, same process as we did in part (a):

First we have to find what the area for a 90% interval:

$$A = \frac{1 + .92}{2} = .96$$

Now using the appendix in the back of the book to find z-score:

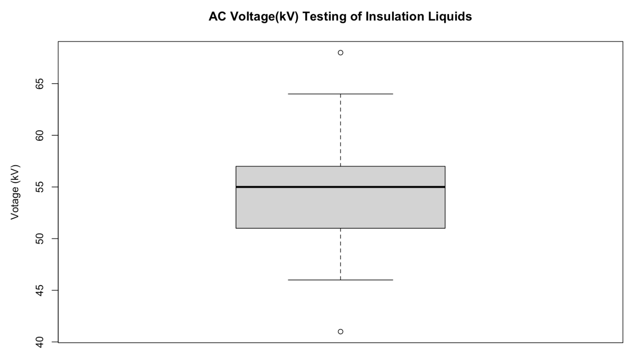
$$.96 = 1.75$$

$$z = 1.75$$

Once changing z that correspond with the percentage CI, then I can find out what my interval is by going through the same process I did in part (a). Note that the interval will be a little wider because our higher CI.

16a) Construct a boxplot of the data and comment on interesting features.

I have constructed a boxplot in R:



This boxplot is not symmetric so that means that is it not normally distributed.

16b) Calculate and interpret a 95% CI for true average breakdown voltage μ. Does it appear that μ has been precisely estimated? Explain.

We find some information to use the margin of error formula: margin of error = $E = z_{\frac{\alpha}{2}} \times \frac{s}{\sqrt{(n)}}$:

$\overline{X} = \frac{x_1+x_2+.....x_n}{n} = 54.70833$

$s^2 = \sigma^2 = \sum(x^2) - (\overline{X})^2 = 27.35993$

$s = \sigma = \sqrt{(\sigma^2)} = 5.230672$

n=48

z=1.96

So now that we have all of our info, we can solve:

$$z_{\frac{\alpha}{2}} \times \frac{s}{\sqrt{(n)}}$$

$$1.96 \times \frac{5.230672}{\sqrt{(48)}}$$

$$E = 1.4798$$

Now we can use this formula to get our intervals: $\overline{X} \pm E$

$$\overline{X} \pm E$$

$$(\overline{X} - E, \overline{X} + E)$$

$$(54.70833 - 1.4798, 54.70833 + 1.4798)$$

$$(53.2285, 56.1881)$$

Answer: (53.2285, 56.1881)

16c) all values of breakdown voltage are between 40 and 70. What sample size would be appropriate for the 95% CI to have a width of 2 kV (so that m is estimated to within 1 kV with 95% confidence)?

For this problem, we have to us this formula: $n = (\frac{z_{\frac{\alpha}{2}}\sigma}{E})^2$

We first need to find the population deviation of 40 to 70, which can be estimated by 1/4: $\sigma = \frac{70-40}{4} = 7.5$

Now we can use the formula to find our n:

$$n = (\frac{z_{\frac{\alpha}{2}}\sigma}{E})^2$$

$$n = (\frac{1.96 \times 7.5}{1.4798})^2$$

$$n = 217$$

Answer: n=217

20a) Calculate and interpret a confidence interval at the 99% confidence level

We are given:

$$n = 2343$$

$$\hat{p} = 53\% = .53$$

$$z = 2.575$$

The margin error formula is: $E = z_{\frac{\alpha}{2}}(\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$

$E = 2.575(\sqrt{\frac{.53(1-.53)}{2343}}$

$E = 0.0266$

Now to ge the intervals: $\hat{p} \pm E$

$.53 - .0266 = .5034$

$.53 + .0266 = 0.5566$

Answer: (0.5034 , 0.5566)

20b) What sample size would be required for the width of a 99% CI to be at most .05 irrespective of the value of $\hat{p}$?

Since we are interested in finding the sample size of 99% CI to be at most .05 irrespective of the value of $\hat{p}$, the we use this formula: $n = \frac{(z_{\frac{\alpha}{2}})^2 \times \hat{p}(1-\hat{p})}{E^2}$

z=2.575

E=0.05

$\hat{p} = .53$

So,

$$n = \frac{2.575^2 \times .53(1 - .53)}{.053^2}$$

$$n = 661$$

Answer: 661 (rounded to the nearest whole number)