Solutions to Worktext Exercises

Chapter 8

Visualizing Properties of Estimators

Basic Learning Exercises

1. Population being Sampled: Mean ____0____ Variance ____57.07___

   The population mean μ is being estimated with the sample mean, $\overline{X} = \frac{1}{n}\sum_{i=1}^{n} x_i$.

2. a) The bottom graph shows a dot plot of one sample superimposed upon the population's distribution. b) The triangle is the sample mean of the sample being displayed in the dot plot. c) The top graph is a histogram of the sample means.

3. Average of Sample Means _–0.9 to 0.9_ Standard Error of Sample Means _1.56 to 2.55_
   CLT Predicted Mean __0__ CLT Predicted Standard Error ___1.95___
   The experimental results should be very close to the predicted results. The red triangle shows the population mean μ, and the blue triangle shows the average of the sample means. If the average of the sample means (blue triangle) is close to the population mean (red triangle), it demonstrates that the estimator may be an unbiased estimator for the population parameter (when sampling from the distribution shown).

4. If one selects n independent observations from any population with finite mean μ and finite standard deviation σ, then the sampling distribution has a mean μ and standard error $\sigma/\sqrt{n}$. If n is large enough, the sampling distribution is normally distributed.

5. Standard Deviation ___7.55___ CLT Predicted Standard Error __0.755__
   The standard error equals $\sigma/\sqrt{n}$. The standard deviation measures the dispersion of the population being sampled, while the standard error measures the dispersion of the sampling distribution of the sample mean.

6. Average of Sample Means _–0.1 to 0.1_ Standard Error of Sample Means _0.72 to 0.82_
   In both exercises the average of the sample means is close to the population mean and the standard error of the sample means is close to what the CLT predicted it would be. This is especially true in this exercise because of the larger number of replications. The standard error decreased as n increased. Both sampling distributions were also normally distributed.

7. The sampling distribution is not normally distributed.

8. The sampling distribution is now normally distributed. This illustrates that the sampling distribution will be normally distributed if the sample size is large enough.

Intermediate Learning Exercises

9. The expected value of the estimator equals the parameter you are estimating.

10. The CLT states that the mean of the sampling distribution of the sample mean is the population mean μ. In other words, the expected value of the sample mean is μ. This means that the sample mean is an unbiased estimator for μ.

11. The sample median is the 50th percentile. If n is odd, it is the middle observation, and if n is even, it is the average of the middle two observations. The sample median is based upon the ranking of the observations and not the values themselves. Therefore, it is not a sufficient estimator because it did not use *all* the information in the sample.

12. No, it does not provide an unbiased estimator in this case. The histogram shows that the average of the sample means is not close to the population mean.

13. Yes, it does produce an unbiased estimator in this case because the histogram is centered at μ and the red and blue triangles are close together.

14. In exercise 13 a symmetric distribution was sampled, while in exercise 12 a very skewed distribution was sampled.

15.

n	Average of Sample Means	Standard Error of Sample Means
10	0.98 to 1.02	0.118 to 0.135
30	0.99 to 1.01	0.069 to 0.078
90	0.99 to 1.01	0.039 to 0.045

16. a) A consistent estimator is one whose probability of being close to the parameter being estimated increases as the sample size increases. b) One can demonstrate that an estimator is consistent by showing that it is unbiased (or asymptotically unbiased) and that its standard error decreases as the sample size increases. c) Exercise 15 shows that the sample mean is a consistent estimator because it is unbiased and because, as the sample size increases, the standard error decreases.

17. An estimator is relatively efficient compared with another estimator if it has a smaller variance and if *both* are unbiased. One can show that an estimator is relatively efficient compared with another estimator by showing that both are unbiased but that one has a smaller variance than the other.

18.

n	Average of Sample Medians	Standard Error of Sample Medians
10	0.98 to 1.02	0.150 to 0.175
30	0.99 to 1.01	0.087 to 0.100

19. Both estimators are unbiased, but clearly the sample mean has a smaller standard error.

Advanced Learning Exercises

20. a) The sample variance estimator is $\sum_{i=1}^{n} \frac{(x_i - \overline{x})^2}{n-1}$. b) The population variance, σ^2, is 100. c) The sample variance is generally between 95 and 112 with 500 replications. d) It is an unbiased estimator for σ^2. e) You can see this by comparing the location of the red and blue triangles on the histogram. If an estimator is unbiased, they will be close together.

21. The population variance is 25. The sample variance is generally between 23.8 and 26 for 500 replications. The sample variance is *always* an unbiased estimator for σ^2.

22. The histogram does not have the scaled χ^2 distribution but is still an unbiased estimator for σ^2.

23. No, the histogram does not have the scaled χ^2 distribution. Since a sample size of 100 is not large enough when you sample from a distribution that is almost normal, a theorem similar to the CLT cannot be proven about the sample variance.

24. a) The sample MSD estimator is $\sum_{i=1}^{n} \frac{\left(x_i - \overline{x}\right)^2}{n}$. b) The population variance is 81. c) The average sample MSD is generally between 56 and 68 using 500 replications. d) The average sample MSD is a biased estimator for σ^2 since its mean differs from 81.

25. The bias has been substantially reduced with a sample size of 100. The average sample MSD is generally between 80 and 82 using 500 replications. This demonstrates that the average sample MSD is an asymptotically unbiased estimator for σ^2.

26. a) The bootstrap estimator of the variance is the square root of the variance of the K resampled means multiplied by $\sqrt{n}$ and squared. b) The 50 is the number of times each sample is resampled. c) Because you must resample each sample 50 times, it takes longer. d) The bootstrap estimator is relatively efficient compared with the sample variance because both are unbiased estimators for σ^2 and the bootstrap has a smaller variance.