Applet #3 Activities

To provide an estimate of the mean μ of a population based on a sample we can use the sample mean. This is intuitive. However, it is clear that the sample mean will be a better estimator in some cases than others. For instance, it would be hoped that an estimate based on a hundred observations provides a better estimate than one based on ten. Moreover, the sample mean is a random variable. Its value will vary from sample to sample. With these points in mind, it seems desirable to supplement the estimate of μ with something that indicates how *confident* we are that the particular sample mean we obtained is a "good guess" at μ. Let us focus on the special case of sampling from a Normal distribution.

1. Consider a production line producing bolts whose lengths are independent and from the Normal distribution with mean $\mu = 300$ mm and standard deviation 2 mm. On the first day, a quality control engineer takes a sample of size $n = 20$, finds the mean length of the sample and records it. On the second day, the engineer takes another sample of size 20, finds the mean and records it. On the third day the engineer does the same, and repeats this procedure every day for a month. At the end of this time, the engineer plots a histogram of the means.

   (a) Broadly describe the shape and location of the histogram you would expect to see.

   (b) Provide a range within which about 95% of the means recorded should fall.

2. The above leads us to the idea of *confidence intervals* for μ. An applet to assist your understanding of confidence intervals can be accessed at

   `http://www.zoology.ubc.ca/~whitlock/kingfisher/CIMean.htm`

   Play with this applet for a while, and complete the tutorial. Note that although the population parameter σ is given in the applet, it is not used in computing the intervals. Why is this?

3. What statistic is used to estimate σ?

4. In the tutorial you have obtained many 95% confidence intervals, each based on a given sample size n, starting with $n = 10$. Keeping the values of the mean $\mu = 105$ and standard deviation $\sigma = 30$ fixed, create 95% confidence for each of twenty samples of size $n = 10$ by clicking "Repeated Samples" and stopping when the number of "Successes" and "Failures" sum to twenty. Are all the confidence intervals the same width?

5. How many of the twenty 95% confidence intervals you saw in the previous part are "Successes" in that they contain μ?

6. Construct a table such as the one below:

		Proportion of intervals containing μ
	10	
	15	
	20	
n	25	
	30	
	40	

   Each row in the table corresponds to a set of 500 simulations you will perform from the $N(105, 30)$ distribution. Enter in each row the observed "success rate", the proportion of 95% confidence intervals in the 500 created that contain the population mean. Do not worry too much if your numbers of intervals differ a little from 500; you can enter proportions based on numbers of intervals close to 500.

7. How if at all does the width of a 95% confidence interval change as n increases? The width appears to

   Increase   Stay the same   Decrease

8. How if at all does the probability that a 95% confidence interval contains the population mean μ change as n increases? The probability

will

Increase   Stay the same   Decrease

Explain your reasoning.

9. The applet allows sampling from Normal distributions with different means and variances. Construct a table such as the one below:

		Proportion of intervals containing μ
	10	
	15	
	20	
n	25	
	30	
	40	

   Each row in the table corresponds to a set of 500 simulations you will perform from the $N(105, \sigma)$ distribution, for your own choice of $\sigma > 30$. Enter in each row the observed "success rate", the proportion of 95% confidence intervals in the 500 created that contain the population mean. Do not worry too much if your numbers of intervals differ a little from 500; you can enter proportions based on numbers of intervals close to 500.

10. Explore how the confidence intervals change when σ is altered, say by repeating the previous part for $\sigma < 30$. How if at all does the width of a 95% confidence interval change as σ increases? The width will

Increase   Stay the same   Decrease