1. Should researchers rely on the accuracy of self-reported measurements? To investigate this question, a study was conducted. Fifteen randomly-selected people were asked to report their weights on a questionnaire, and then were weighed on an accurate scale. For each subject, the difference between the measured weight and the self-reported weight was computed. These differences were found to have a mean of 3.2 and a standard deviation of 3.6 pounds. The differences showed no evidence of outliers or nonnormality.

   (a) (5 points) Compute a level 0.90 confidence interval for the population mean difference.

   ANS: For these data, \( n = 15 \) and \( s = 3.6 \), which means that \( \hat{\sigma}(\bar{Y}) = \frac{3.6}{\sqrt{15}} = 0.9295 \). In addition, \( t_{n-1, 0.05} = t_{14, 0.05} = 1.7613 \), so a level 0.90 confidence interval for \( \mu \) is

   \[ 3.2 \pm (0.9295)(1.7613) = (1.563, 4.837) \]

   (b) (5 points) What does the interval you computed in (a) suggest about the mean difference between measured and reported weights? Justify your response.

   ANS: Since the interval contains all positive values, it suggests that on average measured weight is higher than reported weight. Specifically, it suggests that measured weight is between 1.505 and 4.895 pounds greater than reported weight, on average.

   (c) (5 points) A new, randomly-selected individual is going to join a second phase of the study. Predict the difference between her actual and reported weight with 90% confidence.

   ANS: A 90% prediction interval is

   \[ \hat{Y}_{\text{new}} \pm \hat{\sigma}(Y_{\text{new}} - \hat{Y}_{\text{new}})t_{n-1, 0.05} \]

   where \( \hat{Y}_{\text{new}} = \bar{Y} = 3.2 \), \( t_{n-1, 0.05} = t_{14, 0.05} = 1.7613 \), and

   \[ \hat{\sigma}(Y_{\text{new}} - \hat{Y}_{\text{new}}) = S\sqrt{1 + \frac{1}{n}} = 3.6\sqrt{1 + \frac{1}{15}} = 3.718. \]

   The resulting interval is

   \[ 3.2 \pm (3.718)(1.7613) = (-3.349, 9.749) \]

   (d) (5 points) Based on the interval from (c), do you predict that the new individual will report a lower weight than her actual weight? Justify your response.

   ANS: No, since the prediction interval contains 0, we cannot predict which weight will be higher.

   (e) (5 points) Obtain an interval that with 90% confidence will contain at least 99% of all population differences between measured and reported weights.

   ANS: The appropriate tolerance interval is computed as

   \[ 3.2 \pm (3.562)(3.6) = (-9.623, 16.023) \]

2. A computer manufacturer randomly samples computer mice from the production line and subjects them to wear tests. In one such test, the left mouse button is pressed repeatedly by a testing machine until it fails. Data on the number of press-release cycles until failure in 12 test mice has mean 6709.6 and standard deviation 117.2. The process is stationary, and the data show no evidence of nonnormality.

   (a) (10 points) Construct a level 0.99 confidence interval for the population mean cycles to failure.

   ANS: The interval is

   \[ \bar{Y} \pm \frac{s}{\sqrt{n}}t_{11, 0.995} = 6709.6 \pm \frac{117.2}{\sqrt{12}} = 3.1058 = (6604.5, 6814.7) \]
(b) **(20 points)** Construct an interval based on the 12 values in the data set which you are 99 percent confident will contain the number of cycles to failure of the next mouse tested. If the next mouse fails after 6208 cycles, what will you conclude based on the interval you calculated?

**ANS:** We construct a level 0.99 prediction interval. The interval is

$$
\bar{y} \pm s \sqrt{1 + \frac{1}{n} t_{11, 0.995}} = 6709.6 \pm 117.2 \sqrt{1 + \frac{1}{12} 3.1058} = (6330.7, 7088.5)
$$

Since 6208 is below the interval, we conclude that the performance of this mouse is inferior to that of the population as represented by the 12 mice in the sample.

(c) **(20 points)** Construct an interval you are 99% confident will contain at least 95% of all cycles to failure values in the population. Interpret this interval in terms of “99% confidence”.

**ANS:** We construct a level 0.99 tolerance interval for proportion 0.95 of the population values. The interval is

$$
\bar{y} \pm K \cdot s = 6709.6 \pm 3.870 \cdot 117.2 = (6256.0, 7163.2)
$$

**Interpretation:** In repeated sampling, approximately 99% of all intervals constructed will contain at least 95% of all population values.

3. **(10 points)** Consider a level $L$ confidence interval for the mean $\mu$ in the C+E model where $\sigma$ is known. If I want to double the precision of the interval (i.e., make it half as wide), how much must I increase the sample size? How about if I want to triple the precision? Give a general formula for how much I must increase the sample size $n$ to make the interval $k$ times as precise.

**ANS:** We know from formula (5.20), p. 257 of the text, that the sample size is inversely proportional to the square of the desired precision. Therefore, to double the precision, we need to quadruple the sample size. To triple the precision, we need 9 times the sample size. To increase precision by a factor of $k$, we need to obtain $k^2$ times the number of observations.