The caliper of a tree is the diameter of the tree measured six inches from the ground. A researcher measures the caliper of sixteen trees sampled at random from all trees between 12 and 14 feet tall in a nursery. The measurements in inches are displayed here.

1.18 1.19 1.21 1.33 1.44 1.49 1.50 1.51 1.51 1.70 1.21 2.05 1.65 1.76 1.18 1.19 1.80

The mean and standard deviation of these values are 1.555 and 0.255 inches, respectively.

(a) Find the lower and upper quartiles and the median of the data and display the data with a boxplot.

Solution:

Here is the data sorted.

[1] 1.18 1.19 1.21 1.33 1.44 1.49 1.50 1.51 1.51 1.64 1.65 1.70 1.76 1.80 1.90
[16] 2.05

As there are \( n = 16 \) observations and 16 is even, the median is the average of the two middle observations after sorting. As 16 is also divisible by four, the lower and upper quartiles also fall between observations, counting up by 4 from the bottom and 4 from the top to find the most extreme end of each pair. The median is 1.51 and the quartiles are 1.385 and 1.73.

Here is a boxplot of the data.

![Boxplot](image)

(b) Find a 99% confidence interval for the mean caliper of 12 to 14 foot tall trees in the nursery. Interpret the confidence interval in the context of the problem.

Solution:

(You may have noticed that the mean and sd I report in the problem description are not quite right. This is because I calculated the mean and sd on raw data data before I rounded the values to two decimal places for display in the exam. The mean and sd of the printed data are mean = 1.554 and sd = 0.254.)

There are 15 degrees of freedom, so the correct \( t \) multiplier is 2.947. The margin of error is \( 2.947 \times 0.255/\sqrt{16} = 0.19 \). It is good practice to round the margin of error to two significant digits and then round the estimate to the same level of precision. A 99% confidence interval is as follows.

\[
1.37 < \mu < 1.75
\]

The population is trees of a certain size in the nursery.

In the context of the problem,

We are 99% confident that the mean caliper of trees in the nursery with heights between 12 and 14 feet is between 1.37 and 1.75 inches.
(c) Is your confidence interval valid? Comment on the features of the data apparent in your numerical and graphical summaries that might affect the validity of the confidence interval in the previous part.

Solution: The procedure for constructing the confidence interval is predicated on the assumptions that the underlying population is normal and that the sample was selected at random. We are told in the problem description to treat the data as a random sample. The boxplot is fairly symmetric (there is a very small skew to the right) indicating that the confidence interval is valid.

Substantial outliers or very strong skewness would be reason to be concerned with a small sample of size 16, but this data looks fine. Notice in particular that the small sample size by itself is not a reason to invalidate a confidence interval. The small sample size simply results in a larger margin of error.

(d) Find a 95% confidence interval for the proportion of trees in the nursery with heights between 12 and 14 feet whose calipers are greater than 2 inches.

Solution: Only one of 16 trees have calipers larger than two inches. Plugging in to the textbook formula, we have $\hat{p} = (1 + 2)/(16 + 4) = 0.15$. The standard error is $\sqrt{(0.15)(0.85)/20} = 0.0798$. We use z multipliers for proportions, so the margin of error is $1.96 \times 0.0798 = 0.16$. The 95% confidence interval for $p$ by the formula is $-0.01 < p < 0.31$. Many of you noticed the impossibility of having a negative probability and adjusted the interval to be $0 < p < 0.31$ which is good.

The real issue is that a confidence interval for $p$ when the sample proportion is so small really ought not be symmetric and based on a normal distribution, but this is the best you could do within the context of this course.

(e) Answer TRUE or FALSE. For a different species of tree, if a 90% confidence interval for the mean caliper $\mu$ of trees with heights between 12 and 14 feet tall is computed to be $2.1 < \mu < 2.6$, then 90% of the individual trees in this new population have calipers between 2.1 and 2.6 inches.

Solution: The statement is false. Confidence intervals are statements about the location of a parameter (such as $\mu$) and not about the locations of individuals.