1. We first find \( z \) such that \( \Pr\{Z < z\} = 0.98 \), i.e. \( \Pr\{Z > z\} = 0.02 \). We get \( z = 2.05 \). It gives \( x - 25 \over \sqrt{4} = 2.05 \) i.e. \( x = 29.10 \) cm.

2. (a) Either method was acceptable, as long as it was reasonably defended. Histogram could be preferred because the many decimal places make a stem and leaf cumbersome; stem and leaf may have been better if there is a desire to retain the precision of the points. Constructed with some artistry, both will show the general shape of the distribution, location, spread, and possible outliers, so these were not criterion that would justify one or the other.

(b) \( 0.18 \times 15 = 2.7 \), so the 0.18 quantile is the 3rd smallest data point, i.e. 8.012 cm.

(c) \( s^2 = \frac{1043.218 - 124.965^2/15}{14} = 0.1525 \)

(d) i. The sample mean of the original 16 observations is equal to the sample mean of the available 15 observations. Intuitively, if you think of the mean as the balance point of the data, adding another point exactly on the fulcrum will maintain the balance. Algebraically, if we let \( \frac{x}{n} = \bar{x} \), where \( x = \) sum of 15 remaining points, and \( n=15 \), we see that the mean of the original 16 observations would be \( \frac{x + \bar{x}}{n+1} \). A few algebraic manipulations show this to be equal to \( \frac{x}{n} \), which proves the statement.

ii. The sample variance of the original 16 observations is smaller than the sample variance of the available 15 observations. We established above that the mean was unchanged. Adding another observation at the mean will therefore leave the numerator in the formula for the sample variance unchanged, since the deviation of the mean from the mean is zero. However, the denominator will increase by one, causing the sample variance of the original 16 observations to be slightly smaller.

3. (a) i. \( \Pr\{B\} = 0.5 \), \( \Pr\{I\} = \Pr\{\text{boy} \& \text{immunized}\} + \Pr\{\text{girl} \& \text{immunized}\} = 0.5 \times 0.95 + 0.5 \times 0.40 = 0.66 \).

ii. \( \Pr\{I|B\} = 0.40 \) is not the same as \( \Pr\{I\} = 0.66 \), so \( B \) and \( I \) are not independent.

(b) i. \( \Pr\{Y = 2\} = 1 - 0.1832 - 0.3136 = 0.5032 \)

ii. \( \mathbb{E}(Y) = 1 \times 0.3136 + 2 \times 0.5032 = 1.32 \)
\( \text{var}(Y) = (0 - 1.32)^2 \times 0.1832 + (1 - 1.32)^2 \times 0.3136 + (2 - 1.32)^2 \times 0.5032 = 0.584 \)

iii. \( Y \) does not have a binomial distribution because the immunization status of the two children are not independent: these two children are not sampled independently of each other. If I know the first child is immunized, then it is more likely that this child is a girl, and so it becomes more likely that the second child is also immunized.

4. (a) \( \Pr\{Y \geq 2\} = 1 - \Pr\{Y = 0\} - \Pr\{Y = 1\} \). Using the binomial formula, \( \Pr\{Y = 0\} = (.75)^{20} = 0.00317 \) and \( \Pr\{Y = 1\} = 20 \times 0.25 \times (.75)^{19} = 0.0211 \). We get \( \Pr\{Y \geq 2\} = 0.9757 \).

(b) i. \( Y \) has a binomial distribution, so \( \mathbb{E}(Y) = np = 140 \times 0.25 = 35 \) and \( \text{var}(Y) = npq = 35 \times 0.75 = 26.25 \).

ii. Since \( np = 35 \geq 5 \) and \( nq = 105 \geq 5 \), we can approximate the binomial distribution by the normal \( N(35, 26.25 = (5.12)^2) \). We get
\[ \Pr\{Y \leq 28\} \approx \Pr\left\{Z \leq \frac{28 - 35}{5.12}\right\} = \Pr\{Z \leq -1.366\} \]

By symmetry, this is also \( \Pr\{Z \geq 1.366\} = 0.086 \).

5. (a) We need to calculate \( \Pr\{\bar{Y} \geq 40m\} \) for each species, and see which is the largest. We get
\[ \Pr\left\{Z \geq \frac{40 - 37}{\sqrt{70/50}}\right\} = \Pr\{Z \geq 2.535\} = 0.0055 \]
for species A, while for species B we get
\[ \Pr\left\{Z \geq \frac{40 - 38}{\sqrt{25/50}}\right\} = \Pr\{Z \geq 2.83\} = 0.0023 \]

So the arborist should prefer trees A.

(b) We used \( \bar{Y} \sim N(\mu, \sigma^2/n) \). This is true when
• we assume that trees form an random sample, with trees attaining independent heights, and

• the distribution of tree heights is normal, or the sample is large enough.

There are 50 trees here, so this is large enough, and the second condition is met. However, the first condition is probably not met: If one tree gets a disease, the trees next to it will probably get the disease as well, and a whole range of trees will not grow well. Also, if one tree grows high, the trees next to it are probably in the shade and will not grow as high as they would have grown otherwise.