Assignment 2 Solutions

1. The first property of density curves is that all density curves must lie above the x-axis, which is true in this case.

   The second property of density curves is that the area underneath the density curve must be 1. We can see that the area under the density curve can be split up into a rectangle and two triangles. So the area underneath the density curve is simply the sum of the areas of the rectangle and the two triangles.

   The rectangle has a width of 0.8 - 0.2 = 0.6 and the height of the rectangle is 1.25. So the area of the rectangle is $(0.6)(1.25) = 0.75$.

   The two triangles have the same area as each other since the length of their bases and their heights are the same. The area of one of these triangles is $(1/2)(0.2)(1.25) = 0.125$. So the area of the two triangles combined is $2(0.125) = 0.25$.

   Finally, the total area underneath the density curve is $0.75 + 0.25 = 1$.

2. The proportion of values of X that are between 0.2 and 0.8 is the area of the region (the rectangle) between 0.2 and 0.8 divided by the area underneath the entire density curve.

   The area of the region between 0.2 and 0.8 is $(0.6)(1.25) = 0.75$. The area underneath the entire density curve is 1. So the proportion of values of X that are between 0.2 and 0.8 is $\frac{0.75}{1} = 0.75$.

3. The proportion of values of X that are less than 0.6 is the area of the region under the density curve to the left of 0.6 divided by the area under the entire density curve.

   The region under the density curve to the left of 0.6 consists of the left triangle in the diagram plus the rectangle that extends from 0.2 to 0.6. We have found above that the left triangle has an area of 0.125. The area of the rectangle that extends from 0.2 to 0.6 is $(0.6 - 0.2)(1.25) = 0.5$. So the total area under the density curve to the left of 0.6 is $0.125 + 0.5 = 0.625$. Finally, the proportion of values of X that are less than 0.6 is $\frac{0.625}{1} = 0.625$.

4. The area under the entire density curve is 1. If the proportion to the left of 0.6 is 0.625, then the proportion to the right of 0.6 is $1 - 0.625 = 0.375$.

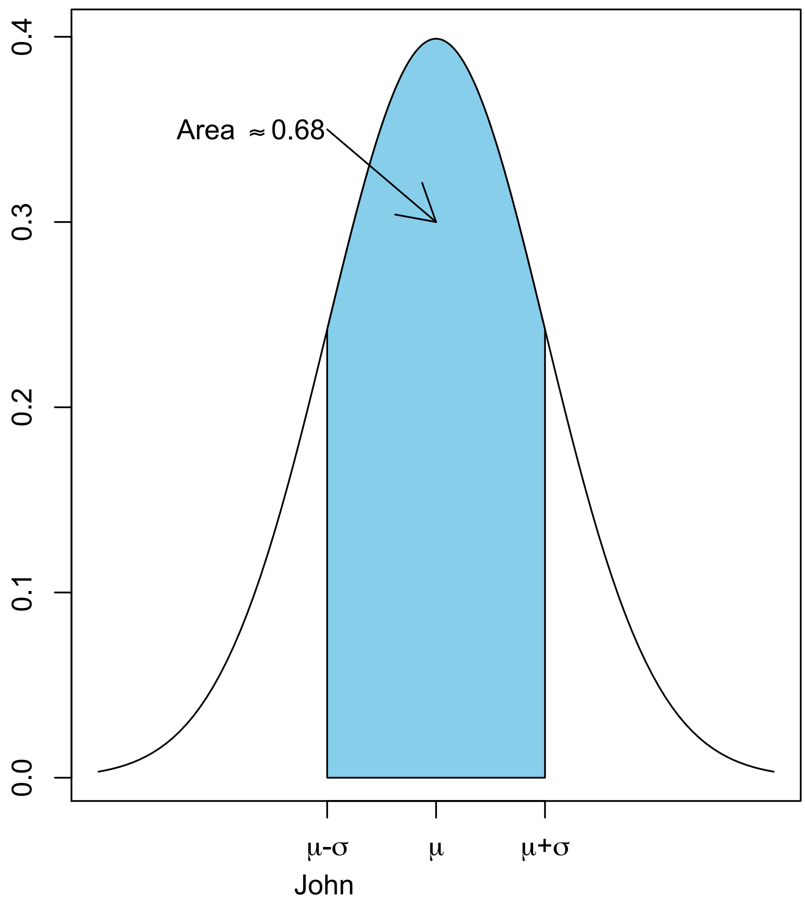
5. We know that the area of the middle rectangle (between 0.2 and 0.8) is 0.75, as found above. So, if we want an area of 0.5 in the region between 0.2 and some value x, the value of x that we're looking for must be to the left of 0.8.

   So, the region of interest that has an area of 0.5 is simply a rectangle. We want the area to be 0.5, we know the height of the rectangle is 1.25, so the base of the rectangle must have length $\frac{0.5}{1.25} = 0.4$. In order for the base to have a length of 0.4, the value of x must be $0.2 + 0.4 = 0.6$.

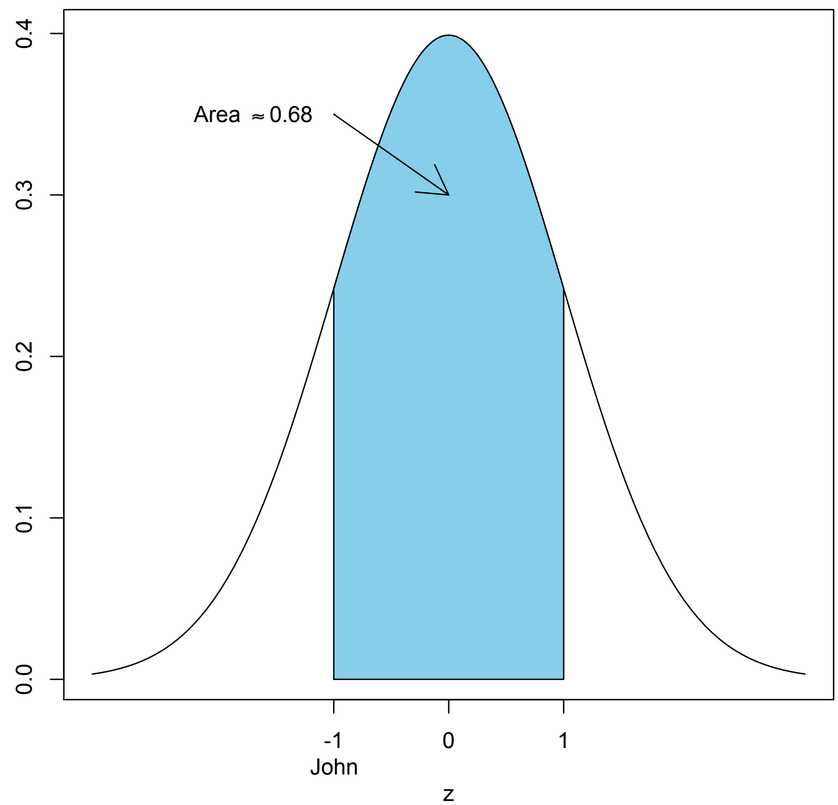
6. If John completes the exam in a time that is one standard deviation below the mean completion time, that means John's completion time x corresponds to a z-score of $z = -1$. Another way of stating this is that John's completion time x is equal to $\mu - \sigma$ so that the z-score is:

$$z = \frac{(\mu - \sigma) - \mu}{\sigma} = -1$$

Yet another way of seeing this is as follows:



which corresponds to the following picture with the standard normal distribution:



Since $P(-1 < Z < 1) \approx 0.68$, that means $P(-1 < Z < 0) \approx 0.34$. Furthermore, $P(Z < 0) = 0.5$, so $P(Z < -1) \approx 0.5 - 0.34 = 0.16$. That means approximately 16% of students complete the exam in less time than John (or approximately 84% of students take longer to complete the exam than John). Since there are 100 students writing the exam, approximately 84 students take longer to complete the exam than John.

7. $P(-0.88 < Z < z) = P(Z < z) - P(Z \leq -0.88) = P(Z < z) - 0.1894 = 0.7977$

   $$\Rightarrow P(Z < z) = 0.7977 + 0.1894 = 0.9871 \Rightarrow z = 2.23$$

8. $P(Z > z) = 0.0537 \Rightarrow P(Z \leq z) = 1 - 0.0537 = 0.9463 \Rightarrow z = 1.61$

9. $P(90 < X < 110) = P\left(\frac{90-105}{10} < Z < \frac{110-105}{10}\right) = P(-1.50 < Z < 0.50)$

   $= P(Z < 0.50) - P(Z < -1.50) = 0.6915 - 0.0668 = 0.6247$

10. We want to find the value x such that $P(X < x) = 0.10$. We first find the value z such that $P(Z < z) = 0.10$ to be $z = -1.28$. We solve for x as follows:

$$z = \frac{x - \mu}{\sigma} \Rightarrow x = \mu + z\sigma = 105 + (-1.28)(10) = 92.2$$

11. The 65th percentile of completion times is the completion time x such that 65% of all completion times are less than x. Therefore, we want to find the value x such that $P(X < x) = 0.65$. We first find the value z such that $P(Z < z) = 0.65$ to be $z = 0.39$. We solve for x as follows:

$$z = \frac{x - \mu}{\sigma} \Rightarrow x = \mu + z\sigma = 105 + (0.39)(10) = 108.9$$

12. We know that $P(X > 200) = 0.2266$. We first find the value z such that $P(Z > z) = 0.2266 \Rightarrow P(Z < z) = 1 - 0.2266 = 0.7734$. From Table 2, we find this value to be $z = 0.75$. We now solve for σ as follows:

$$z = \frac{x - \mu}{\sigma} \Rightarrow \sigma = \frac{x - \mu}{z} = \frac{200 - 194}{0.75} = 8$$

13. We know that $P(X > 85) = 0.0375$. We first find the value z such that $P(Z > z) = 0.0375 \Rightarrow P(Z < z) = 1 - 0.0375 = 0.9625$. From Table 2, we find this value to be $z = 1.78$. We now solve for μ as follows:

$$z = \frac{x - \mu}{\sigma} \Rightarrow \mu = x - z\sigma = 85 - (1.78)(8) = 70.76$$

14. (a) There are two factors in this experiment – frequency and dose.
    Frequency has three factor levels – once per day, twice per day, four times per day.
    Dose has three factor levels – 325 mg, 500 mg, 650 mg.

    (b) There are $3 \times 3 = 9$ treatments:
    once per day/325 mg, once per day/500 mg, once per day/650 mg,
    twice per day/325 mg, twice per day/500 mg, twice per day/650 mg,
    four times per day/325 mg, four times per day/500 mg, four times per day/650 mg.

    (c) We allocate an equal number of experimental units to each treatment. That is, we allocate $45/9 = 5$ experimental units to each treatment.

    (d) No, there is no blocking variable in this experiment. This is a completely randomized design.

15. (a) There is only factor in this experiment – type of engine analyzer.
    Type of engine analyzer has two factor levels – computerized, electronic.
    Note that the blocking variable is not a factor.

    (b) Since there is only one factor, the treatments are the same as the factor levels – computerized engine analyzer, electronic engine analyzer.

    (c) Yes, this is a randomized block design, so there is a blocking variable. It is car size. (There are three blocks – compact cars, intermediate size cars, and full-sized cars.)
    Note that a blocking variable is a lurking variable (it may have an influence on the response variable, but it is not the variable that we are studying for its effect on the response variable) that is brought into the design of the experiment to minimize its effect on the response variable.