STAT 201 - Week-In-Review 8

Dr. Prasenjit Ghosh

Problem Solutions

1. Consider a non-negative continuous random variable X with $\mathrm{P}(X > 6.45) = 0.17$.

   (a) Find $\mathrm{P}(X \geq 6.45)$.

   Solution: $\mathrm{P}(X \geq 6.45) = \mathrm{P}(X > 6.45) = 0.17$.

   (b) Find $\mathrm{P}(X \leq 6.45)$.

   Solution: $\mathrm{P}(X \leq 6.45) = 1 - \mathrm{P}(X > 6.45) = 1 - 0.17 = 0.83$.

   (c) Find $\mathrm{P}(X < 6.45)$.

   Solution: $\mathrm{P}(X < 6.45) = \mathrm{P}(X \leq 6.45) = 0.83$.

   (d) Find $\mathrm{P}(X = 6.45)$.

   Solution: $\mathrm{P}(X = 6.45) = 0$.

2. Suppose a STAT 201 student just found out that she scored 75 on Midterm 1, with a z-score of -0.5. Her friend, who is in a different section with the same professor, also scored 75 but had a z-score of 0.5. What can you conclude about the average Midterm 1 scores in the classes?

   (a) The Midterm 1 class averages must both be equal to 75.

   (b) The student's class had a lower Midterm 1 average than her friend's class average.

   (c) The student's class had a higher Midterm 1 average than her friend's class average.

   (d) The Midterm 1 scores for the two classes must have the same standard deviations.

   (e) The class averages cannot be compared unless we know the corresponding standard deviations.

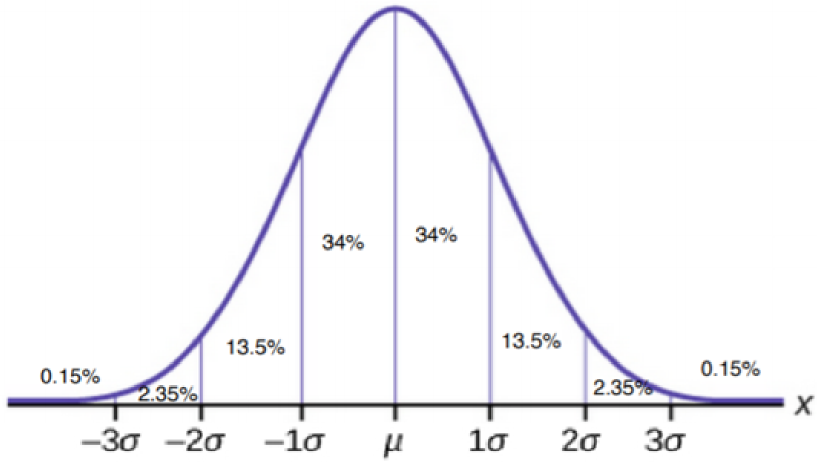
3. It is known that when a specific type of radish is grown in a certain manner without fertilizer the weights (X) of the radishes produced are normally distributed with a mean of 40 gram and a standard deviation of 10 gram.

   (A) Use the empirical rule to determine the percentage of radishes grown without fertilizer with weights less than 50 grams.

   Solution: According to the problem, $X \sim N(\mu = 40, \sigma = 10)$. Then, using the empirical rule, we obtain

$$\mathrm{P}(X < 50) = \mathrm{P}(X \leq 50) = \mathrm{P}(Z \leq 1) = 0.0015 + 0.0135 + 0.34 + 0.34 = 0.84.$$

   Hence, the required percentage is 84%.



(B) Use the empirical rule to determine the percentage of radishes grown without fertilizer with weights between 20 grams and 60 grams.

Solution: Using the empirical rule, we obtain

$$\mathrm{P}(20 < X < 60) = \mathrm{P}(20 \leq X \leq 60) = \mathrm{P}(-2 \leq Z \leq 2) = 0.95.$$

Hence, the required percentage is 95%.

(C) Use the empirical rule to determine the percentage of radishes grown without fertilizer with weights more than 60 grams.

Solution: Using the empirical rule, we obtain

$$\mathrm{P}(X > 60) = \mathrm{P}(Z > 2) = 0.0235 + 0.0015 = 0.025$$

Hence, the required percentage is 2.5%.

4. The impurity level in a batch of chemicals is approximately normally distributed with a population mean of $\mu = 4\%$ and a population standard deviation of $\sigma = 1.5\%$.

   For a randomly selected batch of chemicals, find the approximate probability that the impurity level is between 3.4% and 4.3%.

   Solution: Let the random variable X denote the impurity level in a randomly selected batch of chemicals. Then, according to the problem,

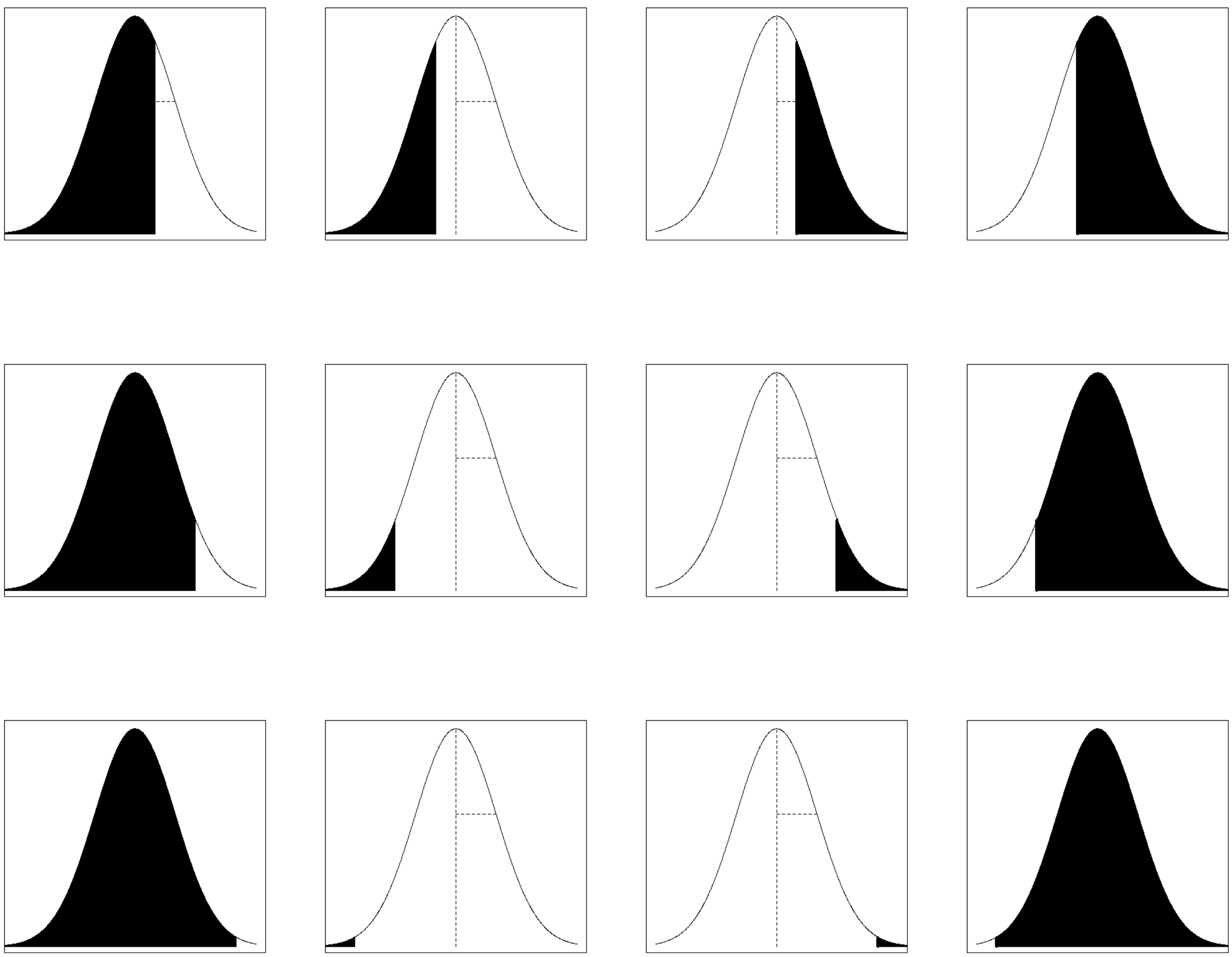
$$X \overset{a}{\sim} \mathrm{N}(\mu = 4,\ \sigma = 1.5).$$

   Then the required probability is:

$$\begin{aligned} \mathrm{P}(3.4 < \mathrm{X} < 4.3) &= \mathrm{P}\left(\mathrm{Z} < \frac{4.3 - \mu}{\sigma}\right) - \mathrm{P}\left(\mathrm{Z} < \frac{3.4 - \mu}{\sigma}\right) \\ &= \mathrm{P}(\mathrm{Z} < 0.2) - \mathrm{P}(\mathrm{Z} < -0.4) \\ &= 0.57926 - 0.34458 \\ &= 0.23468. \end{aligned}$$

5. The distribution of salt per cubic meter of seawater is N(μ = 9.29, σ = 2.2) grams. We want to know the probability that a random meter of seawater would have more than 12.59 grams of salt per cubic meter.

   Which picture below shows that probability?



   Solution: The 3rd image on the 2nd row or, equivalently, the 2nd image on the 3rd column.

6. The mean June midday temperature in Desertville is 36°C and the standard deviation is 3°C. Assuming this data to be normally distributed, how many days in June would you expect the midday temperature to be between 39°C and 42°C? Round up your answer to the next positive integer.

   Solution: Let the random variable X denote the midday temperature in Desertville on a randomly selected day in June. Then, according to the problem,

$$X \overset{a}{\sim} \mathrm{N}(\mu = 36,\ \sigma = 3).$$

First we find the probability that the the June midday temperature in Desertville on a randomly selected day lies between 39°C and 42°C:

$$\begin{aligned} p &= P(39 < X < 42) \\ &= P\left(Z < \frac{42-\mu}{\sigma}\right) - P\left(Z < \frac{39-\mu}{\sigma}\right) \\ &= P(Z < 2) - P(Z < 1) \\ &= 0.97725 - 0.84134 \\ &= 0.13591. \end{aligned}$$

For i = 1, ..., 30, define a random variable Y_i that takes the value 1 if the midday temperature in Desertville on the i-th day of June lies between 39°C and 42°C, and Y_i takes the value 0 otherwise.

Then, for each i = 1, ..., 30, $Y_i \sim \mathrm{Bernoulli}(p)$, whence $E(Y_i) = 0 \cdot (1-p) + 1 \cdot p = p$.

Define a new random variable Y as

$$Y = Y_1 + Y_2 + \cdots + Y_{30},$$

which denotes the number of days in the midday temperature in Desertville lies between 39°C and 42°C.

Then, using the linearity property of expectation, it follows:

$$E(Y) = E(Y_1 + \cdots + Y_{30}) = E(Y_1) + \cdots + E(Y_{30}) = p + \cdots + p = 30p = 30 \times 0.13591 = 4.0773.$$

Since the number of days cannot contain a decimal part, the desired expected number of days in June the midday temperature in Desertville lies between 39°C and 42°C would be 5.

Remember that you need to round up $E(Y)$ if it is not a whole number.

7. When it rains in College Station the average amount of rain is $N(\mu = 7.92, \sigma = 2.2)$ cm. If it rains too much the City Council worries about flooding. Fortunately, it only happens 3.92% of the time. How much rain is enough to worry the City Council?

Solution: Let the random variable X denote the amount of rain in College Station on a random day. Then $X \sim N(\mu = 7.92, \sigma = 2.2)$ cm. We need to find an x such that $P(X > x) = 0.0392$, that is, $P(X \leq x) = 1 - 0.0392 = 0.9608$.

This means that x is the 96.08-th percentile of the distribution of X.

Then,

$$\begin{aligned} x &= \mu + \sigma \cdot z_{0.0392} \\ &= 7.92 + 2.2 \times 1.76 \quad [\text{since } z_{0.0392} = 1.76 \text{ (follows from z-table)}] \\ &= 11.792. \end{aligned}$$

Hence, the required amount of rain in College Station must be 11.792 cm which is enough to make the City Council worried.

8. Suppose the distribution of scores of students in an exam follows approximately a normal distribution with a mean of 74, and a standard deviation of 10. Find the middle 95% scores of students in the exam.

Solution: Let the random variable X denote the score obtained by a randomly selected student in the exam. Then, $X \overset{a}{\sim} \mathrm{N}(\mu = 74,\ \sigma = 10)$.

By middle 95% scores, we mean that there exist scores x and y, with $x < y$, which are symmetrically located around the mean $\mu = 74$, such that the area under the curve in between them is 0.95, and the areas on both the tails are 2.5% each:

$$\mathrm{P}(x \le X \le y) = 0.95,\ \mathrm{P}(X \le x) = 0.025,\ \text{and}\ \mathrm{P}(X > y) = 0.025.$$

This means, that x and y would be the 2.5−th and the 97.5−th percentiles of the distribution of X, respectively.

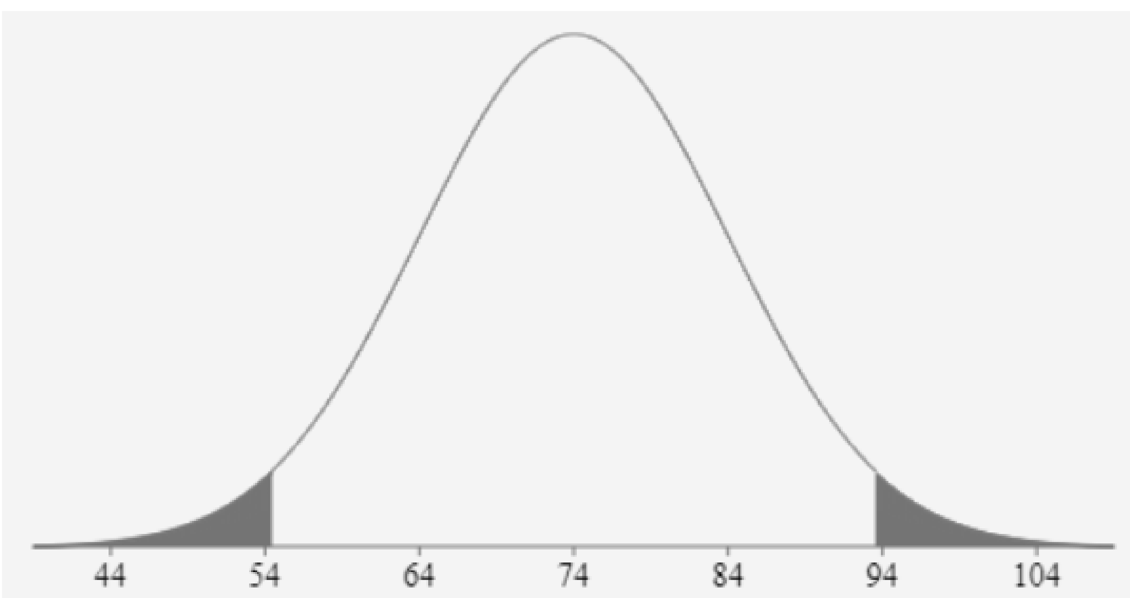
Now, the 97.5−th percentile of X is given by

$$\begin{aligned} x_{0.025} &= \mu + \sigma \cdot z_{0.025} \\ &= 74 + 10 \times 1.96 \quad [\text{ since } z_{0.025} = 1.96] \\ &= 93.6, \end{aligned}$$

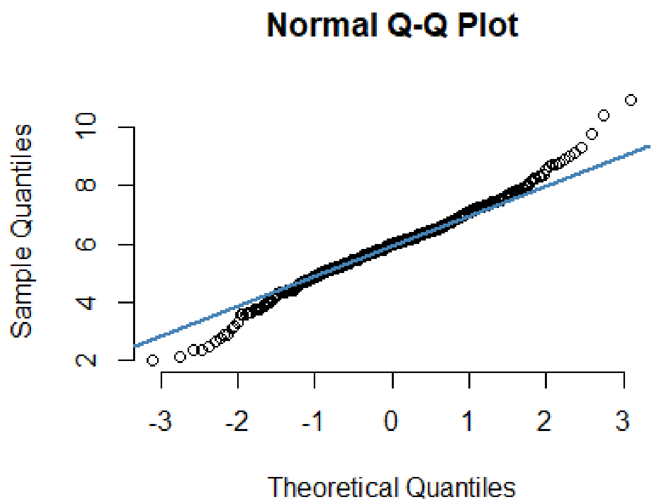
and the 2.5−th percentile of X is given by

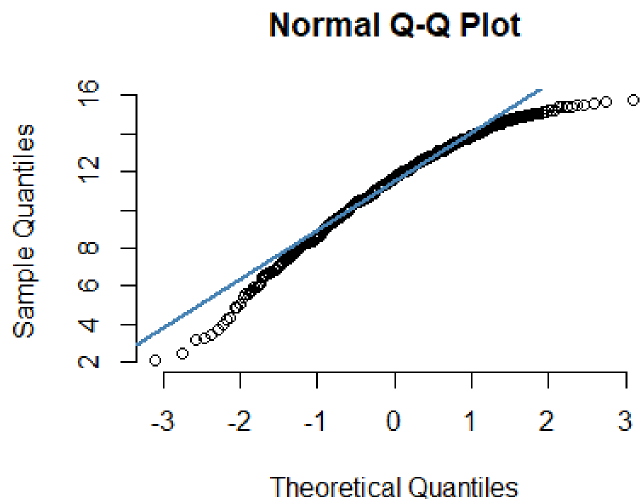
$$\begin{aligned} x_{0.975} &= \mu + \sigma \cdot z_{0.975} \\ &= 74 + 10 \times (-1.96) \quad [\text{ since } z_{0.975} = -z_{0.025} = 1.96] \\ &= 54.4. \end{aligned}$$

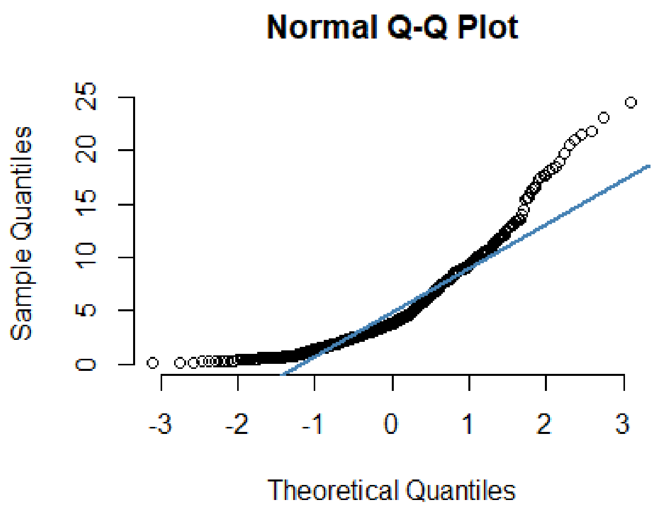
Thus, the required middle 95% scores are x = 54.4, and y = 93.6.

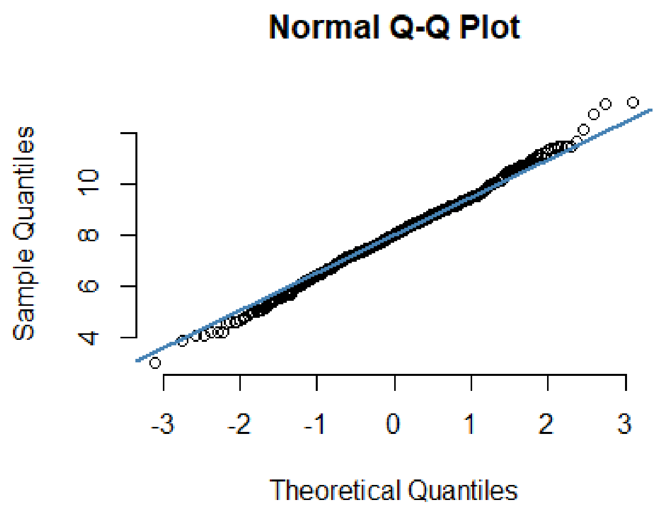


9. Which of the following QQ-plots should correspond to a sample of size n = 500 drawn from a normal distribution?









Solution: The fourth one.