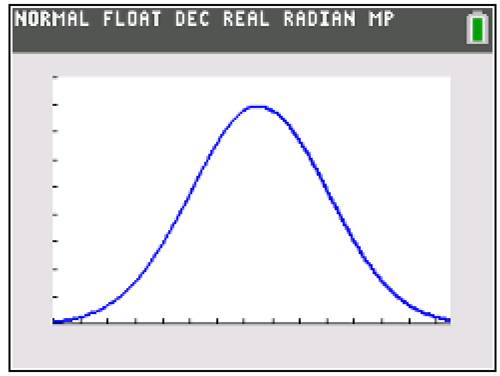
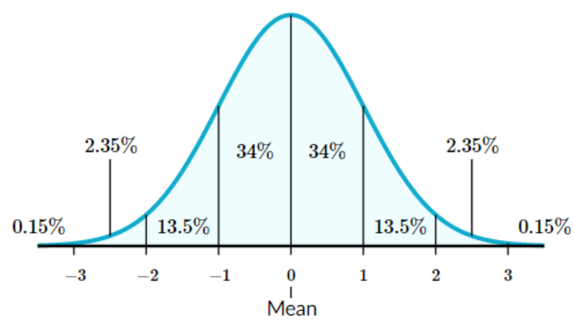
Name ______________________
Class ______________________

In this activity, you will use the idea of a normal distribution to pull together multiple areas of probability and statistics. You will start with basic ideas using means, standard deviations, quartiles, interquartile ranges, and z-scores, then progressing to conditional probabilities.



A normal distribution is a probability distribution that is symmetric about the mean, showing that data near the mean are more frequent in occurrence than data far from the mean. In graphical form, the normal distribution appears as a “bell curve”.



Data is symmetrically distributed with no skew, see the graph to the right. There are three rules to remember about a normal distribution:

1. Symmetrical Bell Shape
2. Mean = Median, both are located at the center of the distribution
3. 68% of the data falls within one standard deviation of the mean.

Problem 1 - Basics

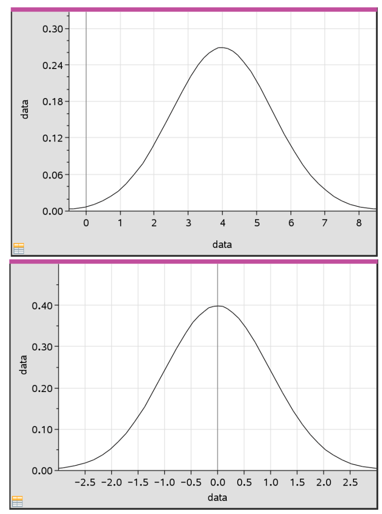
Before we apply the idea of normal distribution to real world scenarios, let us recall what we have learned with some practice.

Using the data: 1, 2, 2, 3, 3, 3, 3, 4, 4, 4, 4, 4, 5, 5, 5, 5, 6, 6, 7

1. Find its mean and standard deviation.

2. Find Q_1, Median, Q_3, and the Interquartile Range.

3. Discuss with a classmate if you think this is a normal distribution. Explain. If you think it is, graph this normal distribution on your handheld.



When a normal distribution occurs, you can center your data around the actual mean, but if the full data is not given, and only its statistics, you can standardize your data to z-scores and center the data around the mean of 0 and standard deviation of 1.

To find the z-score, you will use the formula:

$$z = \frac{x - \mu}{\sigma}$$

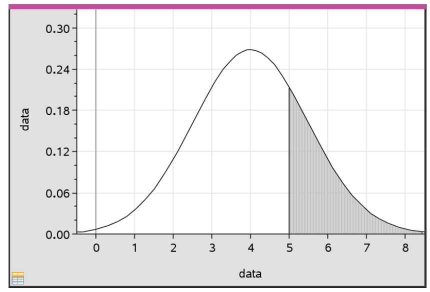
Where x is the given data value, μ is the mean, and σ is the standard deviation.

4. Convert $x = 7$ to a z-score.

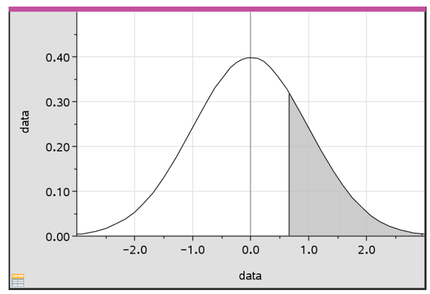
5. Given the z-score of 1, find its corresponding data value x.

These normal distributions can also be used to find probabilities. The cumulative probability of an event occurring is 100% or 1. The total area or shading under a normal distribution curve is also represented by 100% or 1. The shading can be broken down into individual probabilities. Using your handheld, you can find these probabilities by pressing **2nd vars (distr)**, **2 normalcdf** and fill in each line with the appropriate information for your cumulative probabilities.

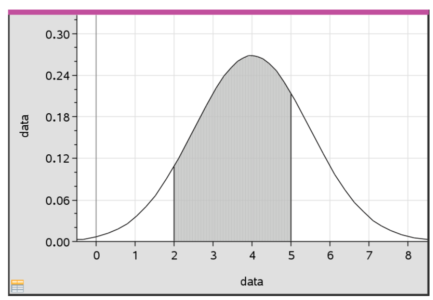
6. Find the probability of selecting a piece of data that is greater than 5. This can also be written as $P(X > 5)$.



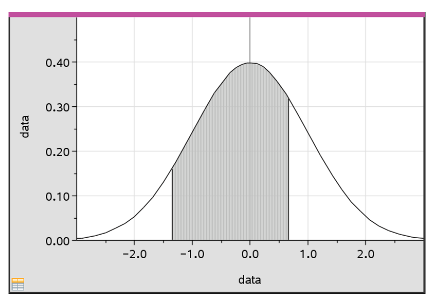
7. With a classmate, write the problem in question number 6 using z-scores, then find this probability.



8. Find the probability of selecting a piece of data that is between the values of 2 and 5. This can also be written as $P(2 < X < 5)$.



9. With a classmate, write the problem in question number 8 using z-scores, then find its probability.



10. (i) Find $P(X < 2.5)$. Sketch the normal distribution curve and shade the region under the curve that represents this probability.

(ii) With a classmate, write the problem in part (i) using z-scores, find its probability and sketch the normal distribution curve and shade the region under the curve that represents this probability.

Problem 2 – Using Statistics and Probability to Find the Data

What if you were given the probability of certain data being selected, could you find individual pieces of this data?

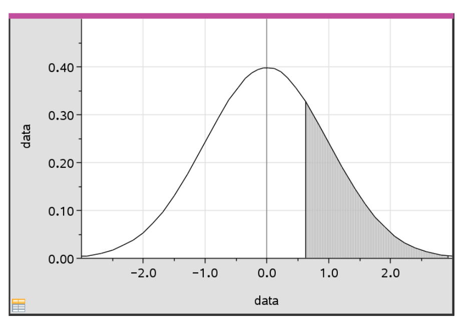
Notation to be familiar with:

Normal Distribution centered around the mean of the given data: $X \sim N(\mu, \sigma^2)$
where μ = mean, σ^2 = variance

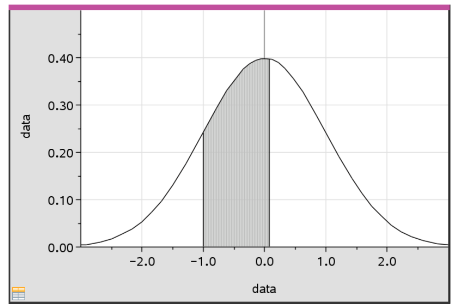
Normal Distribution standardized around the mean score of 0: $Z \sim N(0, 1)$
where 0 = mean and 1 = variance

1. Given $Z \sim N(0, 1)$, find a when $P(Z < a) = 0.576$.
   ****Handheld tip:** You will use the inverse Normal function by pressing **2nd, vars (distr), 3 invNorm(**, then fill in the **Area**, **mean** and **standard deviation** (if not standardized), and choose the placement of the area (tail).

2. Find a such that:
   (i) $P(Z > a) = 0.261$



(ii) $P(-1 < Z < a) = 0.372$



What if the raw data is not given, but the statistics of the data are, could you find individual pieces of data?

3. Given a set of data with a mean of 29 and a standard deviation 2.32, $X \sim N(29, 2.32^2)$, find the Interquartile Range of the data.

Problem 3 – Real World Scenarios

1. The weights of oranges sold at a grocers are normally distributed with a mean weight of 175 g and a standard deviation of 25 g.
   (a) If an orange is chosen at random, find the probability that its weight lies between 160 g and 190 g.

   (b) Find the weight exceeded by 15% of the oranges.

2. The grades of 400 students in an examination are normally distributed with a mean of 60 and a standard deviation of 10.
   (a) If 7% of the students obtain a grade of w marks or more, find the value of w.

   (b) If 15% of the students fail by getting f or less, find the value of f.

Further IB Application

1. A company manufactures fan blades for ceiling fans. The lengths of the blades, L cm, are normally distributed with a mean 65 and a standard deviation of σ. The interquartile range is 7. Find the value of σ.

2. A bakery makes two flavors of cupcakes: Red Velvet and Vanilla.

(a) The weights, R grams, of the red velvet cupcakes are normally distributed with a mean of 24 g and standard deviation of 1.6 g. Find the probability that a randomly selected red velvet cupcake weighs less than 21 g.

(b) The weights, V grams, of the vanilla cupcakes are normally distributed with a mean of 22 g and standard deviation of 1.4 g.
Each day 65% of the cupcakes made are vanilla.
On a particular day, a cupcake is randomly selected from all those made at the baker.

(i) Find the probability that the randomly selected cupcake weighs less than 21 g.

(ii) Given that a randomly selected cupcake weighs less than 21 g, find the probability that it is red velvet.