NORMAL DISTRIBUTIONS

The normal distribution refers to a family of continuous probability distributions defined by the normal equation:

$$Y = \left(\frac{1}{\sigma\sqrt{2\pi}}\right)e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

Where *X* is a normal random variable, μ is the mean, and σ^2 is the standard deviation.

The normal equation is the probability density function for the normal distribution. The graph of the normal distribution depends on two factors - the mean and the standard deviation. The mean of the distribution determines the location of the center of the graph, and the standard deviation determines the height and width of the graph. When the standard deviation is large, the curve is short and wide; when the standard deviation is small, the curve is tall and narrow. All normal distributions look like a symmetric, bell-shaped curve, as shown below.

There are few key points to remember when dealing with continuous probability distribution:

- The total area under the normal curve is equal to 1.
- The probability that a normal random variable *X* equals any particular value is 0.
- The probability that *X* is greater than a certain value *a* equals the area under the normal curve bounded by *a* and plus infinity.
- The probability that *X* is less than a certain value *a* equals the area under the normal curve bounded by *a* and minus infinity.

Every normal curve (regardless of its mean or standard deviation) conforms to the 68-95-99.7 rule, that is:

- 68% of the area under the curve falls within 1 standard deviation of the mean.
- 95% of the area under the curve falls within 2 standard deviations of the mean.

psst…don't forget to take notes

- 99.7% of the area under the curve falls within 3 standard deviations of the mean.

To find the probability associated with a normal random variable of a standard normal distribution we use the z-score and the normal distribution table.

Every normal random variable *X* can be transformed into a *z* score using the following equation:

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

Where *X* is a normal random variable, μ is the mean, and σ is the standard deviation of *X*.

The normal distribution table shows a cumulative probability associated with a particular z-score. Table rows show the whole number and tenths place of the z-score. Table columns show the hundredths place. The cumulative probability appears in the cell of the table.

Often times we may not be interested in the probability that a normal random variable falls between minus infinity and a given value. We may want to know the probability that it lies between a given value and plus infinity, or lies between two given values. These probabilities are easy to compute from a normal distribution table. Consider the following:

- The probability that a standard normal random variable (z) is greater than a given value (a), P(Z < a), is determined by 1 - P(Z < a).
- The probability that a normal random variables lies between two values, P(a < Z < b). is P(Z < b) - P(Z < a).

Often, events in the real world follow a normal distribution. This allows us to use the normal distribution as a model for assessing probabilities associated with these events.

The analysis follows two steps:

- Transform raw data in to the form of z-scores using the equation:

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

- Once the data has been transformed into z-scores, use the normal distribution tables, to find probabilities associated with the z-scores.

Concept Example:

The following problem introduces the concept reviewed within this module. Use this content as a primer for the subsequent material.

A company pays its employees an average wage of $3.25 an hour with a standard deviation of 60 cents. If the wages are approximately normally distributed, determine the proportion of the workers getting wages between $2.75 and $3.69 an hour.

Solution:

To analyze this problem, knowing it is normally distributed, we complete the following process:

- Transform raw data in to the form of z-scores using the equation:

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

- Once the data has been transformed into z-scores, use the normal distribution tables, to find probabilities associated with the z-scores.

So we know that:

$$X_1 = 2.75$$
$$X_2 = 3.69$$
$$\mu = 3.25$$
$$\sigma = .60$$

So:

$$z_1 = \frac{(2.75 - 3.25)}{.60} = .833$$
$$z_2 = \frac{(3.69 - 3.25)}{.60} = .733$$

Which means that:

$$P\left(2.75 < X < 3.69\right) = P(-.833 < z < .733)$$

Using the normal distribution tables we find that:

$$P\left(-.833\right) = .298$$
$$P\left(.733\right) = .268$$

So:

$$.298 + .268 = .566$$

About 56.6% of the workers have wages between \$2.75 and \$3.69 an hour.