

MDM 4U

NORMAL DISTRIBUTIONS

Distributions like the binomial probability distribution deals with DISCRETE DATA. The possible values of the random variable are natural numbers, because they arise from counting processes (usually numbers of successes vs. failures). Many characteristics of a population such as the wing span of a bird population are continuous in nature, and have decimal values. These continuous variables have statistical distributions as well.

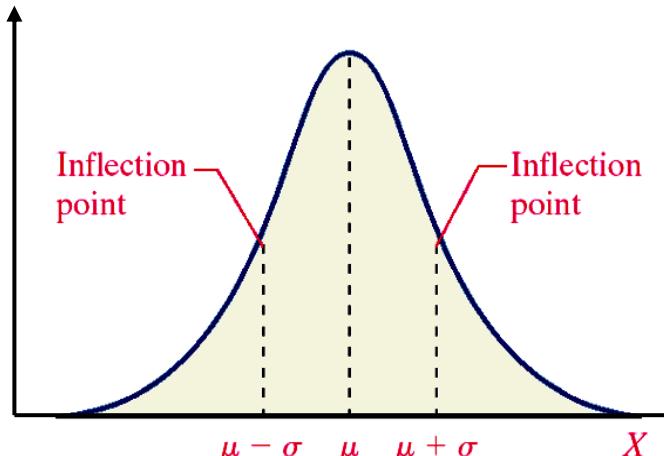
Continuous probability distributions allow fractional values and can be graphed as smooth curves, unlike discrete distributions which are based on bar graphs. The quantities that are measured (and expressed as relative frequencies) tend to be distributed symmetrically and unimodally about the mean. In the distribution of a set of data, statisticians observe a "bell curve" shape so often that its model is referred to as the NORMAL DISTRIBUTION.

COMMON CHARACTERISTICS OF UNIMODAL DATA:

1. Data are symmetrically distributed about the mean.
2. Probability of getting a result in a certain interval decreases as the distance the interval is from the mean increases.
3. With very small intervals for the data, the relative frequency of the distribution would be a smooth curve called a NORMAL CURVE.

NORMAL CURVE PROPERTIES:

- ① frequencies are **symmetric about the mean, μ**
- ② The **standard deviation, σ** , is the distance from the mean to the points of inflection. [points on the curve where the curve changes from a "smile" to a "frown" and vice versa.]
- ③ **Area under the curve equals one.** (sum of all relative frequencies.)



A population that follows a normal distribution can be completely described by its mean, μ , and standard deviation, σ , so the notation used to describe a normal curve is...

$$X \sim N(\mu, \sigma^2)$$

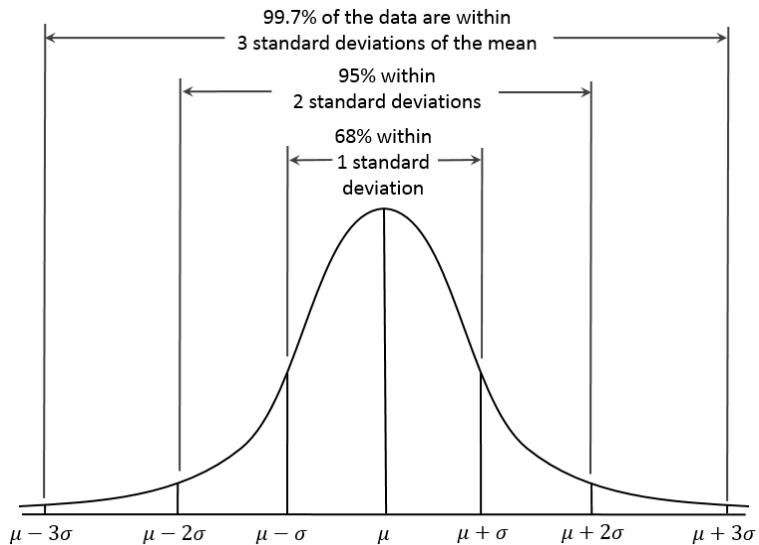
... which is read as "X is distributed normally with mean μ and standard deviation σ ."

NOTE: The term σ^2 is known as the **variance** and we determine the standard deviation as σ by taking the square root of the variance.

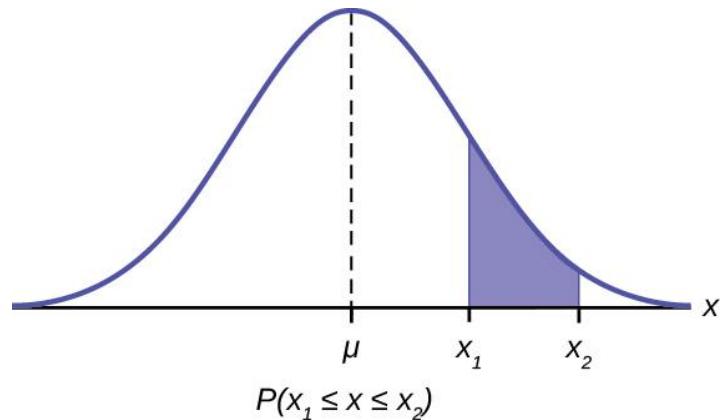
The area under the curve represents all of the data; that is, 100% of the data.

Therefore, we can associate any area under the curve with a specific percentage of the data. For instance,

- $\pm\sigma = 68\% \text{ of data}$
- $\pm 2\sigma = 95\% \text{ of data}$
- $\pm 3\sigma = 99.5\% \text{ of data}$



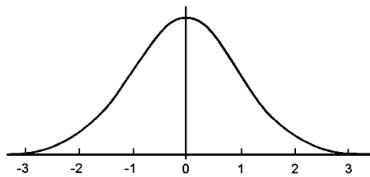
Probabilities for normally distributed populations are equal to areas under the distribution curve. The area under the curve from x_1 to x_2 gives the probability $P(x_1 < X < x_2)$ that data values for X will lie between the values x_1 and x_2 .



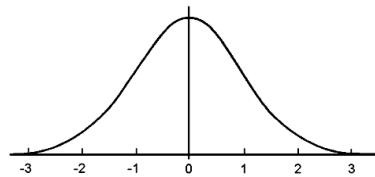
Areas under a curve can be determined to a very high degree of accuracy using a **standard normal distribution table of values** (found on pages 606-607). The table provides probabilities for a normally distributed variable with mean 0 and standard deviation 1.

EXERCISE ①: Given $X \sim N(0, 1)$, illustrate each probability question in the diagram, then determine the probability using the tables on Pages 606-607.

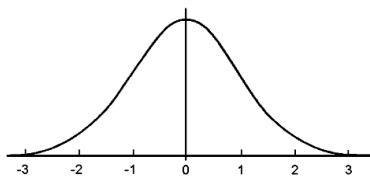
A) $P(x < 1)$



B) $P(x > 1.26)$



C) $P(-1.1 < x < 0.17)$



Most continuous random variables do not have data distributed normally with $\mu = 0$ and $\sigma = 1$.

The **z-score** formula (developed in chapter 2) is used to standardize the given random variable and its distribution.

z-score for a value of the random variable: $z = \frac{x - \mu}{\sigma}$

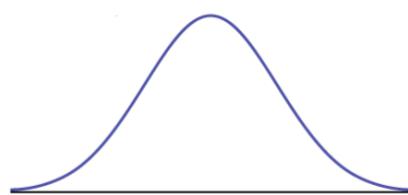
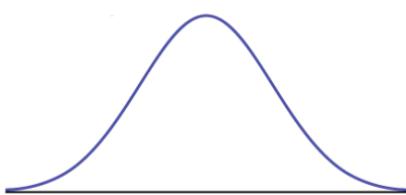
If $X \sim N(\mu, \sigma^2)$, then $z = \frac{x - \mu}{\sigma} \sim N(0, 1)$

EXERCISES: [Assume all problems to be normally distributed.]

② Given $X \sim N(75, 25)$, determine each probability and illustrate in a diagram.

A) $P(X > 82)$

B) $P(60 < X < 70)$



③ Find the probability. Draw a diagram to illustrate the probability.

A) $X \sim N(50, 16)$; $P(X < 45)$ B) $X \sim N(31, 3^2)$; $P(34 < X < 38)$

④ Ben's commuting time to school has a mean of 52.8 min and a standard deviation of 15.6 minutes. Assuming that distribution of times to be normal, determine the percentage that...

A) Ben's commuting times are less than 40 min?

B) commuting times are between 30 and 40 min?

⑤ A university Chemistry professor was not pleased with his class' final exam marks. He decided to "bell curve" the marks. If the exam results had an average of 55% and standard deviation of 20%, and he prefers a mean of 70% and standard deviation of 15%, how would a student's mark of 40% change?

⑥ The following data represent the marks on a science test:
[Assume the distribution of the marks to be normal.]

35, 67, 89, 14, 90, 66, 100, 99, 76, 93, 77, 15, 73, 90, 50, 62

A) Determine the mean and standard deviation.

B) Determine the probability that a student's mark is a passing mark.

C) Determine the probability that a mark is between 60% and 70%.

D) If the mean and standard deviation were adjusted to 75% and 20% respectively, what would happen to a mark of 40%?

⑦ The heights of basketball players on the Toronto Raptors team appear to be much lower than the average of the NBA teams. If 20% of the heights of the players in the NBA are below 6'3" and 30% of the players are below 6'6",

A) Determine the mean and standard deviation of the NBA players' heights.

B) What is the probability that a randomly chosen NBA player has a height...

- i) less than Kyle Lowry?
- ii) greater than the tallest player of this year's league?
- ii) greater than 7'6"?

⑧ **Field goal percentage** in basketball is the *ratio* of field goals *made* to field goals *attempted*. Overall, the league average field goal percentage is estimated to be 0.451, and the standard deviation to be 0.0525.

- A) What is the probability a randomly chosen NBA player has a FGP greater than Stephen Curry : 0.4910108?
- B) The NBA career record for field goal percentage is held by **DeAndre Jordan** at 0.677. What is the probability a future player will score above Jordan's FGP?
- C) A player earns a bonus in a season if their FGP is in the top 15% of the league's average FGP. To what FGP would players have to reach in the season if they are expecting to receive a bonus?
- D) What is the range of FGP between which 75% of the players will score?

⑨ The coach of a track team can send only the top 5% of her runners to a regional track meet. For the members of her team, times for a 1-km run are normally distributed with a mean of 5.6 min and a standard deviation of 0.76 min.

- A) What is the cut-off time to determine which members of the team qualify for the regional meet?
- B) In order to remain a member of the track team, a runner must not score a time in the bottom 10%. What running time must a member make?

⑩ On average 18.5% of drivers on a highway travel at speeds below the speed limit of 100 km/h, while 10% of vehicles travel above 130 km/h.

- A) Determine the mean and standard deviation of travel speeds.
- B) What is the probability that a random driver travels at a speed
 - i) greater than 100 km/h?
 - ii) between 60 and 80 km/h?
 - iii) in excess of 150 km/h?