Basic Statistics

Normal Distributions

Learning, Teaching
and Student Engagement

Normal Distributions

Learning Intentions

Today we will understand:

- The basic properties of probability
- How frequency distributions are used to calculate probability
- Properties of a normal probability distribution

What is Probability?

- **Experiment –** the process of measuring or observing an activity for the purpose of collecting data
- **Outcome –** a particular result of an experiment
- **Sample space –** all possible outcomes of the experiment
- **Event –** one or more outcomes that are of interest in the experiment and which is a subset of the sample space

Image accessed: http://www.relevantclassroom.com/blog/post/roll-the-dice

What is Probability?

- **Experiment –** rolling a pair of dice
- **Outcome –** rolling a pairs of fours with the dice
- **Sample space –** {1,1 2,2 3,3 4,4 5,5 6,6 1,2 1,3 1,4 etc}
- **Event –** rolling a total of 2, 3, 4 or 5 with two dice

Image accessed: http://www.relevantclassroom.com/blog/post/roll-the-dice

What is Probability?

JAMES COOK UNIVERSITY AUSTRALIA

- **Probability –** the likelihood of a particular outcome/event
- Represented as a number between 0 and 1

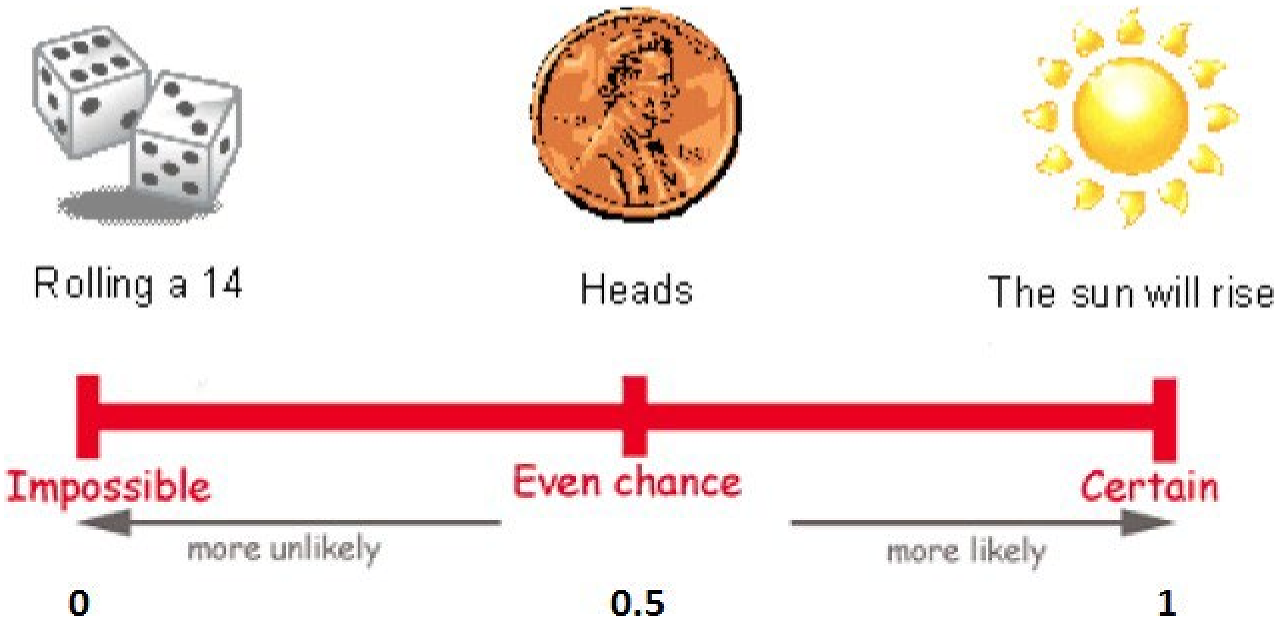


Image accessed: http://www.webquest.hawaii.edu/kahihi/mathdictionary/P/probability.php

What is Probability?

- 1 means certain
- For a single event, the probability of all outcomes must equal 1
- For example, if the probability of the home football team winning the match is 0.7, the probability that they lose is 0.3
- Pr(win) = 0.7
- Pr(lose) = 1 – 0.7

  = 0.3

Image accessed: http://www.shmoop.com/basic-statistics-probability/probability.html

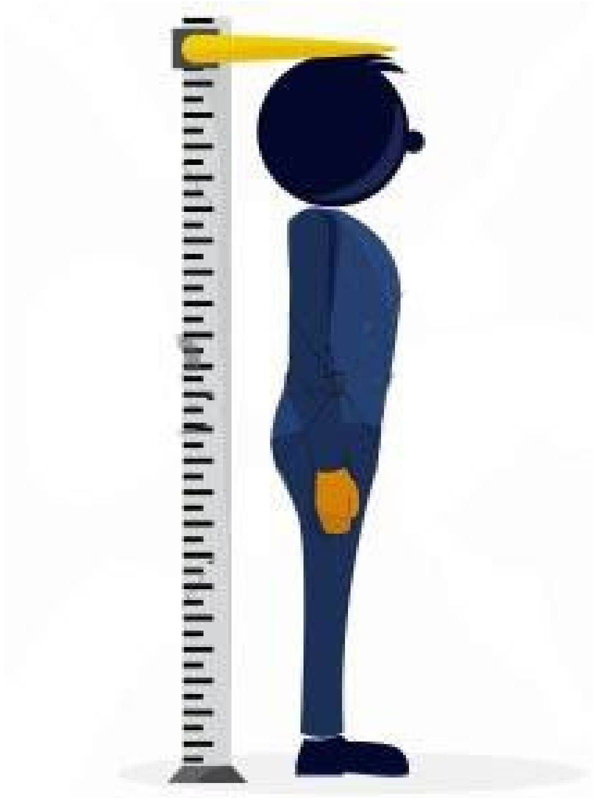
Probability of Outcomes

- **Categorical (discrete) outcomes**
- Heads/tails, win/lose
- Probability of specific outcomes

Image accessed: http://resultsci.com/risking-business-success-coin-toss/

Probability of Outcomes

- Continuous measurements
- Weight, height
- Probability of outcome within a specific range



Relative Frequency

- Probability of an outcome is its relative frequency

- The proportion of times the event would occur if the experiment was repeated over and over again

  Pr(event) = (# times event occurs) / (# trials)

- There are 395 passengers on a plane, 195 females and 200 males. If we choose an adult at random from the group the probability that our choice is female:

  195/395 = 0.49

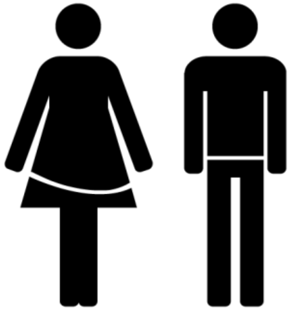


Image accessed: http://michaeljholley.com/2013/03/28/is-escape-vs-reality-a-malefemale-thing/

Probability of Outcomes

- Accuracy of estimate improves with sample size

- Toss a coin 10 times – 6 heads, 4 tails

- Pr(heads) = 6/10 = 0.6
- Pr(tails) = 1 – 0.6 = 0.4

If we run more trials, the relative frequency will approach the actual probability (0.5)

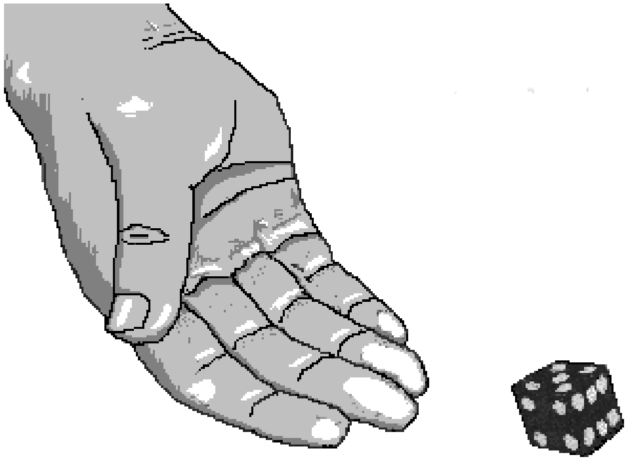
Image accessed: http://ict.channelsteve.com/wp-content/uploads/2012/10/

Probability Distribution

- If we roll a 6-sided dice, then any of the six possible outcomes are equally likely

- Probability of each outcome is 1/6

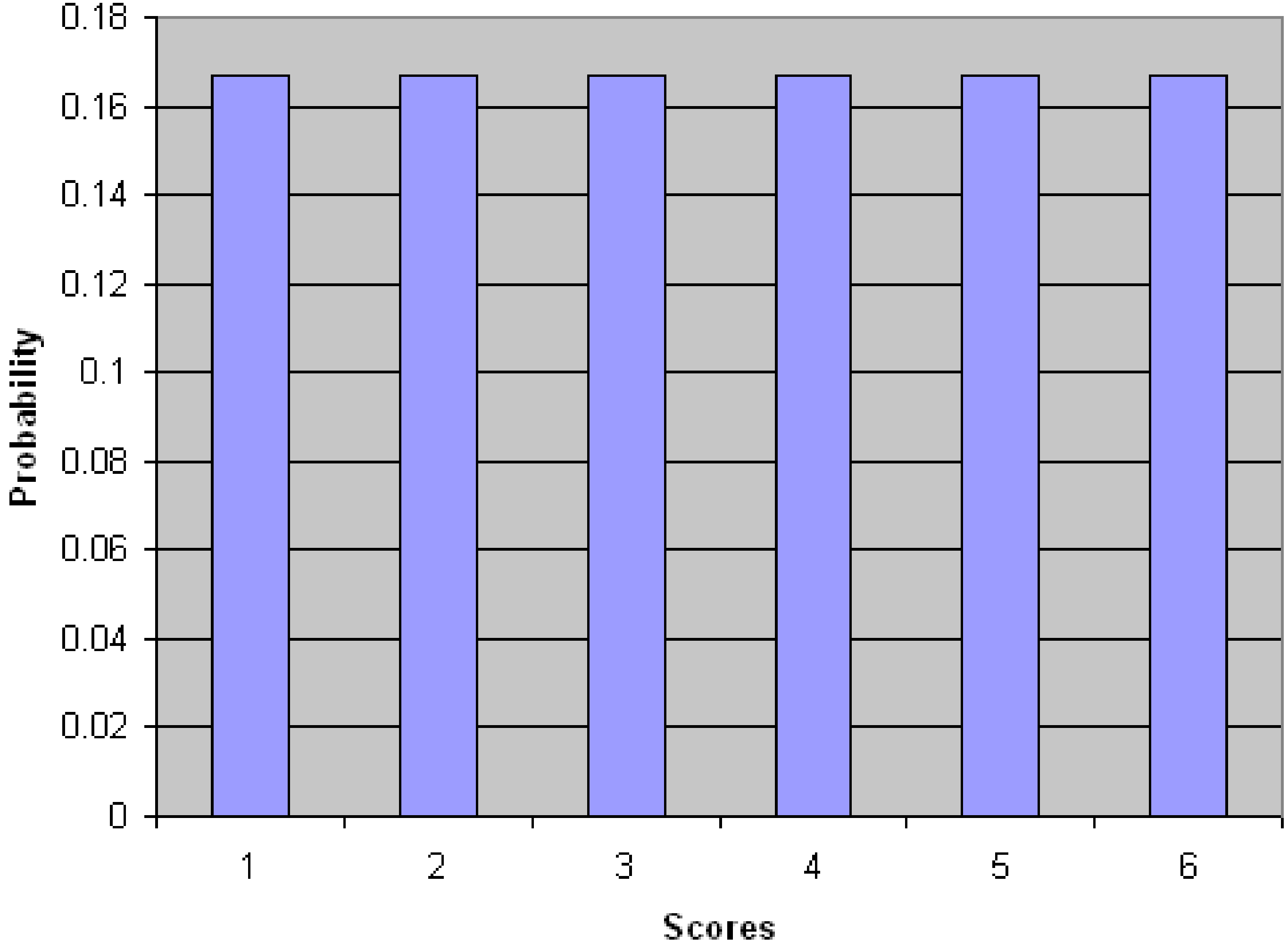
- The probabilities of all outcomes must equal 1



Each outcome is mutually exclusive – only one of the possible outcomes can occur

Image accessed: http://www.instructables.com/id/Single-Die-Gambling-Game/step3/The-Rules/

Probability Distribution



Discrete Probability Distribution

- This probability distribution is **uniform –** all outcomes are equally likely
- **Discrete** – it is not continuous. You cannot get an outcome of 5.45
- A plot of a probability distribution must have a total area of 1 (in this case each bar has an area of 1/6)

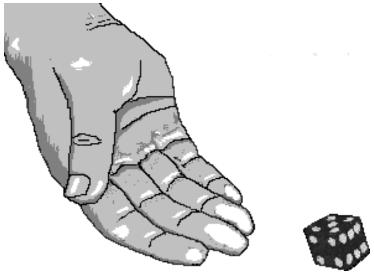


Image accessed: http://www.instructables.com/id/Single-Die-Gambling-Game/step3/The-Rules/

Example: Bag with 5000 green and 5000 yellow balls

- Population =
- Take ball out at random
- Probability of Green ball =
- Put the ball back in the bag and mix
- Probability of picking Yellow ball =
- Total probability =

Example: Bag with 5000 green and 5000 yellow balls

- Population = 10,000
- Take ball out at random
- Probability of Green ball = ½
- Put the ball back in the bag and mix
- Probability of picking Yellow ball = ½
- Total probability = ½ + ½ = 1

Example 2

- You take out 6 balls in sequence (take out ball record colour, then place it back in the bag. Then take out second ball record color, place it back in the bag. Do this 6 times
- What is the probability that the first 6 balls you retrieved were all Green?

- Probability of the first ball being Green =
- Probability of the 2nd ball being Green =
- Probability of the 3rd ball being Green =
- Probability of the 4th ball being Green =
- Probability of the 5th ball being Green =
- Probability of the 6th ball being Green =
- What is the probability that the first 6 balls you retrieved were all Green?

- Probability of the first ball being Green = ½
- Probability of the 2^{nd} ball being Green = ½
- Probability of the 3^{rd} ball being Green = ½
- Probability of the 4^{th} ball being Green = ½
- Probability of the 5^{th} ball being Green = ½
- Probability of the 6^{th} ball being Green = ½
- What is the probability that the first 6 balls you retrieved were all Green?
- ½ x ½ x ½ x ½ x ½ x ½ = 1/64

Rule of multiplication:- Probability of 2 independent events occurring simultaneously is the product of their individual probabilities

Example 2

- What is the probability of getting 5 green and 1 yellow ball?

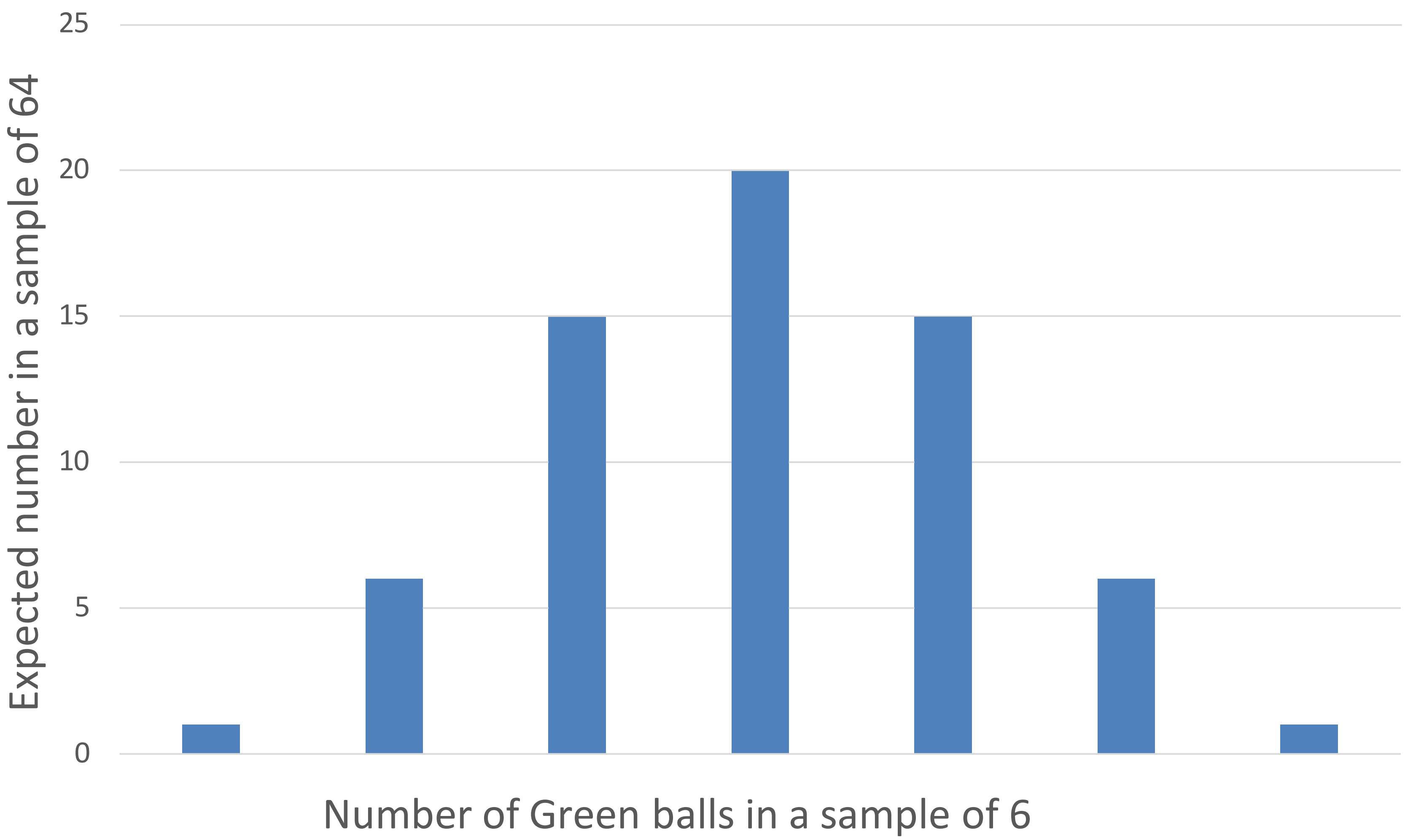
Example 2

- What is the probability of getting 5 green and 1 yellow ball?
- 6 combination in which we can get 5 green and 1 yellow ball
  - YGGGGG → 1/64
  - GYGGGG → 1/64
  - GGYGGG → 1/64
  - GGGYGG → 1/64
  - GGGGYG → 1/64
  - GGGGGY → 1/64

Rule of addition:- Probability of 2 mutually exclusive events occurring simultaneously is the sum of their individual probabilities

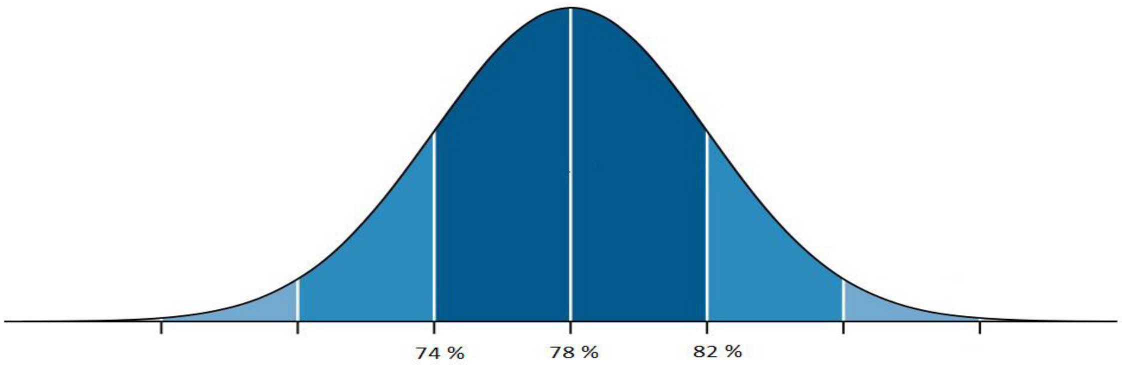
- Probability = 1/64 + 1/64 +1/64 + 1/64 +1/64 + 1/64 = 6/64

# of Green balls	# of Yellow balls	Outcome probability	Probability %
6	0	1/64	1.56
5	1	6/64	9.38
4	2	15/64	23.44
3	3	20/64	31.25
2	4	15/64	23.44
1	5	6/64	9.38
0	6	1/64	1.26
	Total	64/64	100


25
20
15
10
5
0
Expected number in a sample of 64
Number of Green balls in a sample of 6

Normal Probability Distribution

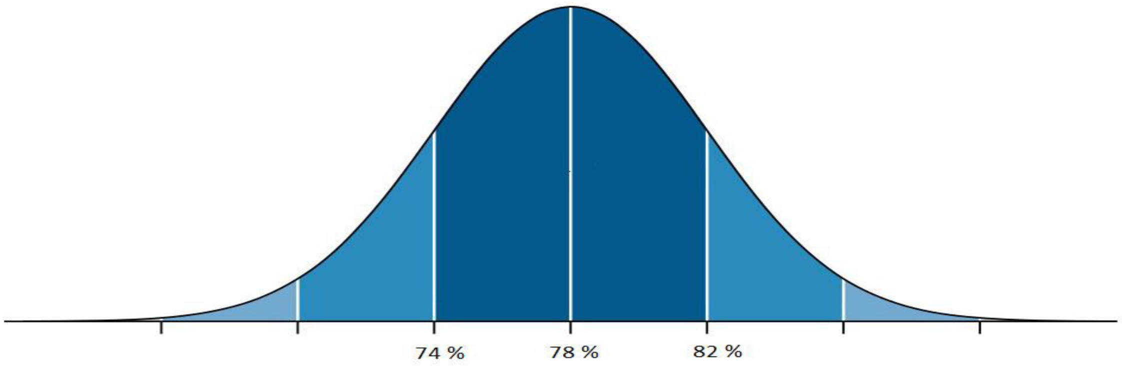
- Continuous variables that follow normal probability distribution have several distinct features



- The mean, mode and median are the same value
- The distribution is bell shaped and symmetrical around the mean
- The total area under the curve is equal to 1

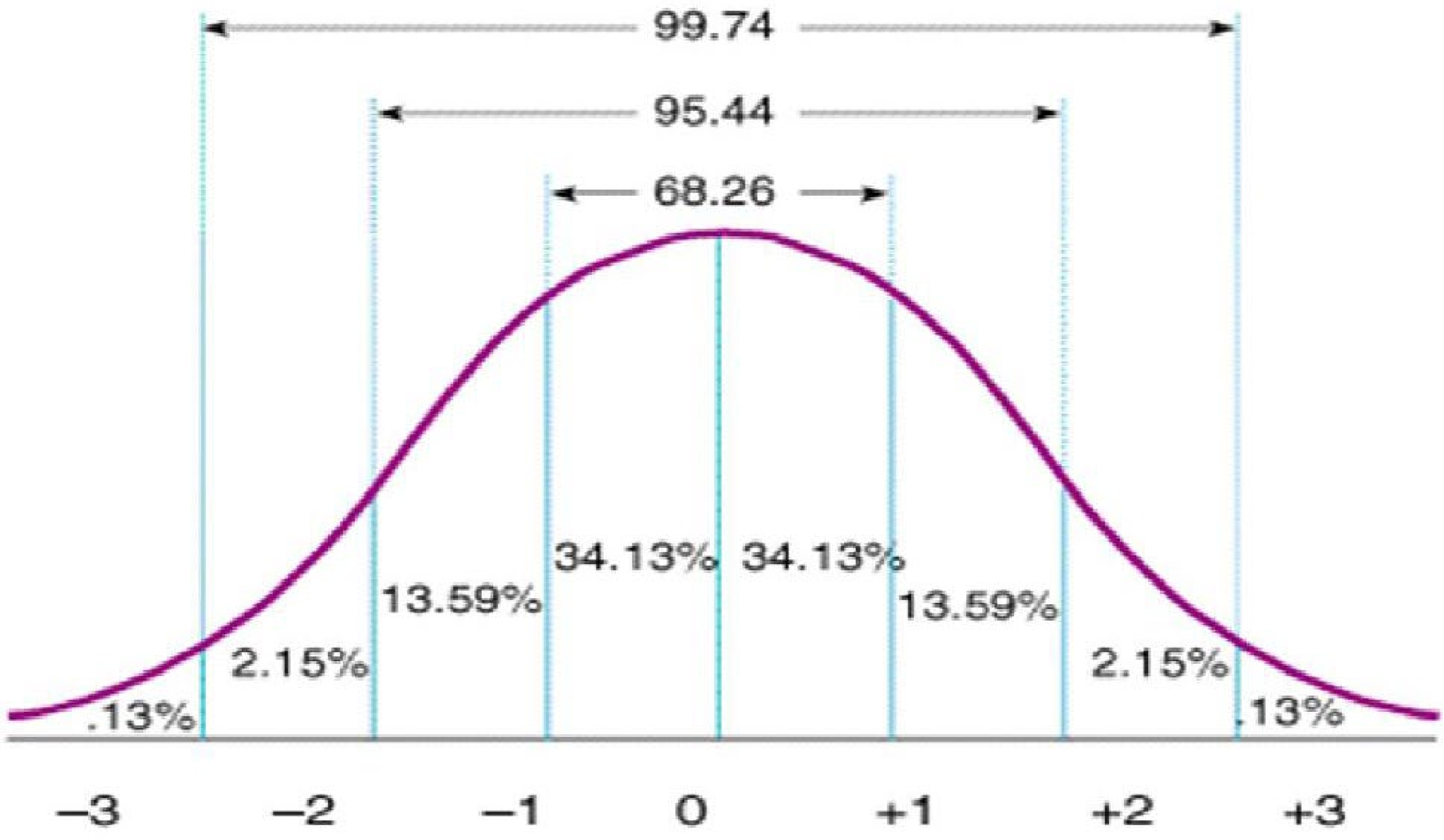
Normal Probability Distribution

- Because the area under the curve = 1 and the curve is symmetrical, we can say the probability of getting more than 78 % is 0.5, as is the probability of getting less than 78 %



- To define other probabilities (ie. The probability of getting 81 % or less) we need to define the **standard normal distribution**

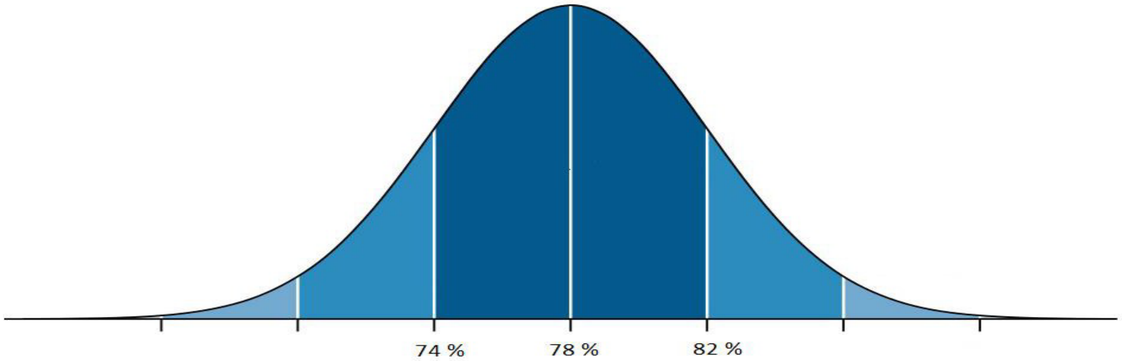
Standard Normal Distribution



- Normal distribution with μ = 0 and SD = 1

Normal Probability Distribution

- To determine the probability of getting 81 % or less



- Determine how many standard deviations the value of 81 % is from the mean of 78 %

Normal Probability Distribution

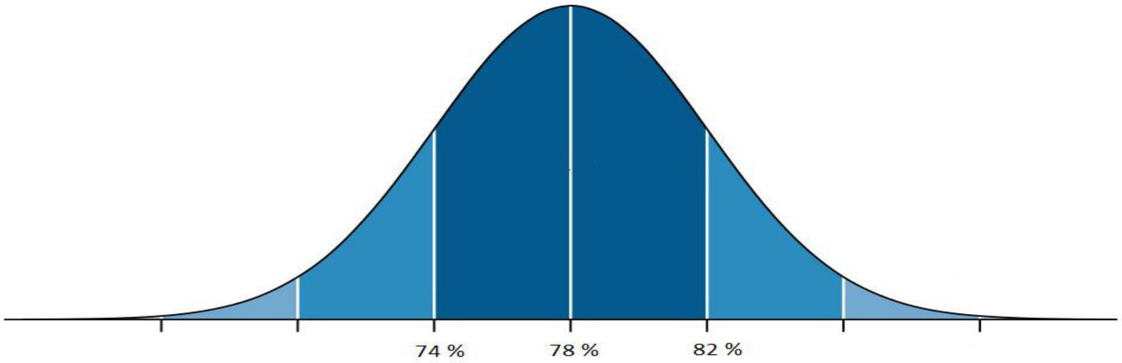
- We do this using the following formula

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

x = the normally distributed random variable of interest
μ = the mean for the normal distribution
σ = the standard deviation of the normal distribution
z = the z-score (the number of standard deviations between x and μ)

Normal Probability Distribution

- To determine the probability of getting 81 % or less



$$z = \frac{x - \mu}{\sigma} = \frac{81 - 78}{4} = 0.75$$

Normal Probability Distribution

- Now that you have the standard z-score (0.75), use a z-score table to determine the probability

x	0.0	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621

Normal Probability Distribution

JAMES COOK UNIVERSITY AUSTRALIA

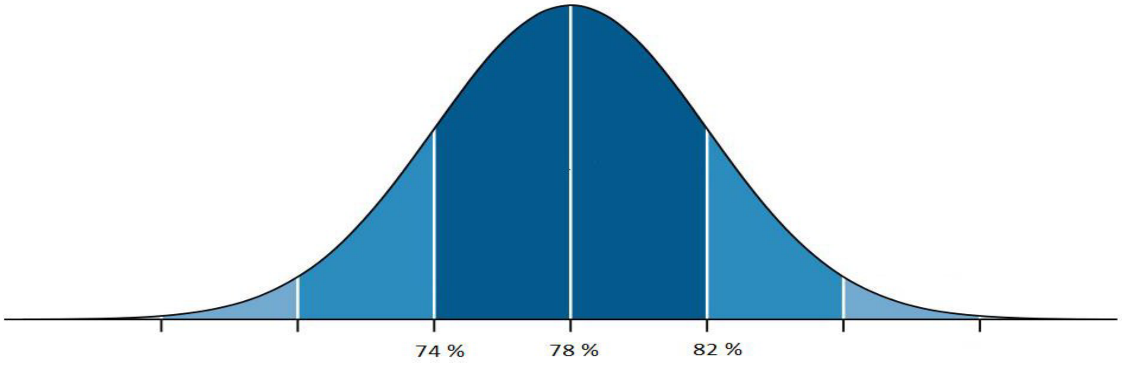
- Z = 0.75, in this example, so we go to the 0.7 row and the 0.05 column

x	0.0	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621

Normal Probability Distribution

- The probability that the z-score will be equal to or less than 0.75 is 0.7734

- Therefore, the probability that the score will be equal to or less than 81 % is 0.7734



- There is a 77.34 % chance I will get 81 % or less on my test