# CHAPTER 3
# PROBABILITY DISTRIBUTIONS

---

## Learning Objectives:

1. *Define probability.*
2. *Determine rare events.*
3. *Distinguish between classical and empirical approaches of computing probability.*
4. *Use basic probability rules.*
5. *Define random variable.*
6. *Distinguish between discrete and continuous random variables.*
7. *Identify the Binomial random variable.*
8. *Apply the Binomial probability model.*
9. *Identify the Normal curve and its characteristics.*
10. *Find probabilities under the standard Normal curve by reading Z-table.*
11. *Convert any Normal curve to the standard Normal curve, and find the corresponding probability.*
12. *Convert any probability given from the Normal curve to find the corresponding random variable X value.*
13. *Apply the Normal distribution in application problems.*
14. *Compute the probability of the Normal distribution using Minitab Express.*
15. *Interpret Minitab Express outputs of Binomial and Normal distributions.*

---

## Content

# 1. Introduction to Probability

## 1.1 What is Probability?

**Probability** is the mathematical way of quantifying _______________, in order to make predictions in the real world. Some examples of chance or probabilistic statements we encounter in our daily lives are:

- Weather report says "there is a 70% chance of rain today."
- Doctor says "there is a 20% chance of complications from the surgery."
- Singapore Pools says "there is a 1 in 175 million chance of winning the lottery."
- NASA says "the probability of a giant asteroid slamming Earth is very low."

Probabilities take values between ___ and ___, both inclusive.
Probability close to 1 means the event will ___________________, and probability close to 0 means the event is _________________ to happen.

## 1.2 Calculating Probability

There are two ways to calculate probability, ______________ and ______________.
To illustrate both further, let us first define some terms.

- **Experiment** or **trial** is an occurrence that has an _____________ outcome.
- **Sample space** is the set of ______________________________. The probabilities of all possible outcomes will add up to ___.
- **Event** is ___________________________ of the outcomes.
- **Rare event** is an event with ______________ probability, i.e. close to _______.

**Classical probability** is used when each outcome in a sample space is equally likely to occur.
The classic probability of an event A is given by:

$$P(A) = \frac{\text{number of outcomes in A}}{\text{total number of outcomes in sample space}}$$

**Empirical probability** is based on observations obtained from probability experiments.
The empirical probability of an event A is simply the **relative frequency** of event A:

$$P(A) = \frac{\text{number of times A occurs}}{\text{number of times experiment is conducted}} = \frac{\text{frequency of A}}{\text{total frequency}}$$

***Example 1:*** In each of the probability statements below, decide between classical and empirical. Also, identify the rare event.

| | |
|---|---|
| P(getting a head in a coin toss) = 50% | Classical / Empirical |
| P(rain today) = 70% | Classical / Empirical |
| P(winning the lottery) = 0.00000001 | Classical / Empirical |
| P(complications from surgery) = $\frac{1}{5}$ | Classical / Empirical |

## Case Study 1: Ceramic Insulators

By Jarek Tuszyński / CC-BY-SA-3.0, CC BY-SA 3.0, https://commons.wikimedia.org/w/index.php?curid=17724597

In Singapore, high-voltage electric power transmission cables are installed underground. However, in many neighbouring countries and further, these cables are above ground and overhead.

Although these cables are usually left bare and uninsulated, insulators are required at the points where they are supported by utility poles or transmission towers. Such insulators are often of ceramic material due to its non-conductivity and heat-withstanding property.

**Formulating questions**

As the ceramic insulators are exposed to weather elements, they are expected to be able to withstand thermal shock, which is the sudden change in temperature. **What is the likelihood of a randomly selected ceramic insulator shattering due to thermal shock?**

**Collecting data**

300 random ceramic insulators from the same manufacturer is tested, of which, 4 shattered under thermal shock.

**Analyzing data**

P(ceramic insulator shatter) =

P(ceramic insulator withstand thermal shock) =

**Interpreting results**

What is the likelihood of a randomly selected ceramic insulator shattering?

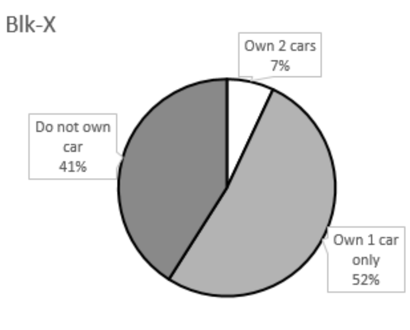
Is it rare for a randomly selected ceramic insulator to shatter? Should it be?

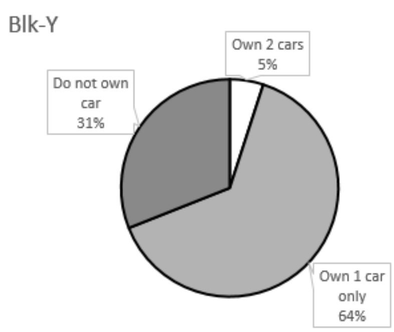
## 1.3 Probability Rules

To compute the probability of multiple events, we can use the following probability rules:

| **Addition Rule for Disjointed Events** | **Multiplication Rule for Independent Events** |
|---|---|
| If event A and event B are <u>disjointed</u>, they do not ____________. | If event A and event B are <u>independent</u>, they do not ____________ each other. |
| P(A **or** B) = P(A) + P(B) | P(A **and** B) = P(A) × P(B) |

***Example 2:*** In a private estate, there are two apartment blocks, Blk-X and Blk-Y.




(a) What is the probability that a random household from Blk-X owns 1 **or** 2 cars?

(b) What is the probability that both a random household from Blk-X **and** a random household from Blk-Y do not own car?

# 2. Random Variables

## 2.1 Discrete and Continuous Random Variables

A random variable is a variable whose numeric value is based on the outcome of a random event. A random variable can be classified as discrete or continuous.

A **discrete** random variable has a _________________________ number of possible outcomes that can be ________. On the contrary, a **continuous** random variable has an _____________ number of possible outcomes.

***Example 3:*** Are the following random variables discrete or continuous?

| | |
|---|---|
| Number of stocks in the Straits Times Index that have share prices increase on a given day. | Discrete / Continuous |
| Volume of water in a 500-ml bottle. | Discrete / Continuous |
| Number of highway fatalities in a country. | Discrete / Continuous |
| Weight of a chemical compound. | Discrete / Continuous |
| Room temperature at 12pm on a particular day in Singapore. | Discrete / Continuous |
| Number of heads that comes up when a coin is tossed four times. | Discrete / Continuous |

## 2.2 Probability of Random Variables

Consider an experiment where a fair coin is tossed four times. Here is the sample space:

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| HHHH | HHHT | HHTH | HTHH | THHH | HHTT | HTHT | THHT |
| HTTH | TTHH | THTH | HTTT | THTT | TTHT | TTTH | TTTT |

Remember that each of the above outcomes is ________________________.

Suppose that we are interested in the number of heads that comes up, so we define:

X = number of _________________ in four coin tosses

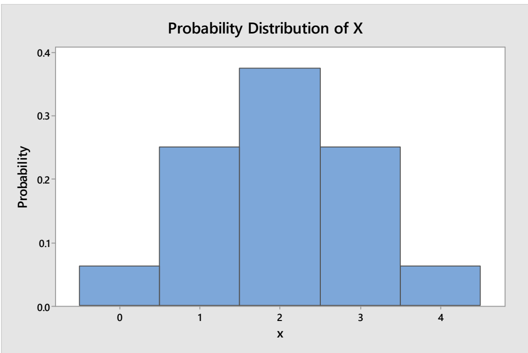
Then X is called a random variable which is the numerical outcome of a random phenomenon. Note that we do not care when in the sequence of tosses we get heads or tails, just the overall number of heads that comes up.

The **probability distribution** of a random variable X tells us the values that the random variable can take on and the probabilities of each value. Here is the probability distribution of X in our four coin tosses, presented in a table form:

| Values of X | | | | | |
|---|---|---|---|---|---|
| Probability, P(X = $x$) | | | | | |

Notice that:

- X is a ___________ variable since it has a finite number of possible values.
- Each of the outcomes listed in the table is possible, but not equally likely.
- The sum of all the probabilities is ___.
- P(X = $x$) or p($x$) denotes the probability associated with a particular value $x$.
- The probability distribution of four coin tosses can also be represented using a graph. The horizontal axis shows us the possible values of $x$, and the height of each bar represents the probability for that value.




***Example 4:*** Use the probability distribution table of four coin tosses to answer the following questions.

(a) What is the most likely number of heads from four coin tosses?

(b) What is the probability of obtaining no heads in four coin tosses?

(c) Find P(X = 2) and P(X < 2).

(d) What is the probability of obtaining at least one head in four coin tosses?

(e) Is it rare to obtain all tails in four coin tosses?

## Case Study 2: Space Shuttle Challenger

*Reference: "Random Variables: Against All Odds—Inside Statistics." Films Media Group, 2013,* [http://fod.infobase.com.ezp1.lib.sp.edu.sg/portalplaylists.aspx?wid=151497&xtid=111539](http://fod.infobase.com.ezp1.lib.sp.edu.sg/portalplaylists.aspx?wid=151497&xtid=111539).

On the morning of January 28, 1986, the space shuttle Challenger 7 broke apart shortly after lift-off.

After thorough investigation, a commission of experts found that the accident was caused by failure in at least one of the O-rings. The O-rings were supposed to seal field joints on the rocket boosters to contain hot, pressurized gases within the boosters.

| | |
|---|---|
| **Formulating questions** | Has the risk of this failure been adequately evaluated? Could the disaster have been predicted? |
| **Collecting data** | The first step in a probability analysis of field joint failure is to calculate the probability of failure in one of them. Under the Challenger flight conditions, the probability of failure of a particular field joint is 0.023, which means that each individual field joint has a probability of success of 0.977. But a space shuttle has six field joints. So for the entire system to succeed, all six field joints have to succeed, i.e. no failures. |
| **Analysing data** | Let X be the number of failures in the six field joints.<br>(a) What is the probability that none of the field joints fail?<br>$P(X = 0) =$<br>(b) What is the probability that at least one field joint fail?<br>$P(X \geq 1) =$ |
| **Interpreting results** | (a) Is the failure of a single field joint considered a rare event?<br>(b) Is the failure of at least one field joints considered a rare event? What is the implication of this on the safety of a space shuttle mission? |

# 3. Discrete Random Variable: Binomial

## 3.1 The Binomial Distribution

Probability models provide us with a list of all possible outcomes and proportions for how often they would each occur in the long run. We can use a probability model to find the following:

| Scenario | Possible Outcomes |
|---|---|
| How many times can we expect to get heads on coin tosses? | head vs. tail |
| How many daffodil blossoms can we expect to see in spring, based on the number of bulbs planted in the previous autumn? | bloom vs. none |
| How many children in a family is expected to inherit a genetic disease? | sick vs. healthy |

There is a commonality among these scenarios; they are all concerned with things that have only two possible outcomes. Traditionally, we think of one possible outcome as a **success** and the other as a **failure**. What we are interested in is the overall count of successes. The count forms a particular kind of discrete probability model, the **Binomial distribution**.

There are four conditions to identify in the Binomial distribution:

#1. There is a repeated fixed number of trials or observations, ***n***.

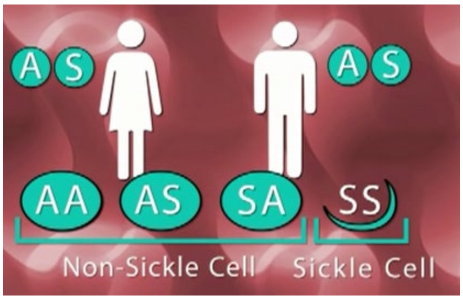
#2. All these trials are independent. That is, the outcome of one trial does not change the probabilities of other trials.

#3. Each trial should end in one of two outcomes: success or failure.

#4. The probability of success, ***p***, must be the same for all trials.

## Case Study 3: Sickle Cell Disease

*Reference: Binomial Distribution: Against All Odds—Inside Statistics. (2013). Films Media Group. Available at:* http://fod.infobase.com/PortalPlaylists.aspx?wID=151497&xtid=111540




In people with sickle cell disease, the sickle hemoglobin molecules cause the normally round red blood cells to destore into a sickle shape, which causes blockages in the blood vessels. Tissues downstream are starved of oxygen, causing damage and much pain.

The genes that determine an individual's hemoglobin type are inherited, one version from each parent. Since it is a recessive disease, the child needs to receive two bad versions of the gene, one from each carrier parent, to have the disease.

| | |
|---|---|
| **Formulating questions** | Public health officials want to know the mean number of children with sickle cell disease in a family where the parents are carriers. |
| **Collecting data** | Inheritance of the sickle cell disease, if both parents are carriers, fits the Binomial distribution.<br>• There are 2 possible outcomes in each child conceived: sick or healthy<br>• The outcome for each child is independent.<br>• The number of children in a particular family and the parents' genetic makeup do not change.<br>• The probability of a child having sickle cell disease ('success') is 0.25, and is the same for each pregnancy. |
| **Analysing data** | Let X be the number of children with sickle cell disease in a family with 6 children.<br>Possible values of $x$ are: 0, 1, 2, 3, 4, 5, 6<br>Number of trials, $n = 6$<br>Probability of success, $p = 0.25$<br>Mean, $\mu = np = 6 \times 0.25 = 1.5$ |
| **Interpreting results** | So the mean number of children with sickle cell disease, in families of six children where both parents are carriers, is ______.<br>If their first child turns out to have sickle cell disease, does it mean that their next 3 children will not have sickle cell disease? Explain. |

## 3.2 The Binomial Probability Model

To denote a Binomial random variable X, with number of trials $n$ and probability of success in each trial $p$, we write:

$$\boxed{X \sim B(n, p)}$$

The possible values that X can take are 0, 1, 2, 3, …, $n$.
Then, the probability of getting exactly $x$ successes out of $n$ trials is given by the formula:

$$\boxed{P(X = x) = {}_nC_x \; p^x q^{n-x}}$$

where $q =$ ______ , is the probability of ____________ in each trial, and

${}_nC_x$ is called the ________________________.

Furthermore, $\boxed{\text{expected number of success} = \mu = np}$

***Example 5:*** Suppose that in a space shuttle, there are 6 field joints working independently. The probability of each field joint failing is 0.023.

(a) Explain why this scenario can fit the Binomial distribution.

(b) Let X be the number of field joints that fail out of 6. Derive the probability distribution of X.

***Example 6:*** If two parents are both carriers of the sickle cell disease, the chance of their child having the disease is 1 out of 4.

(a) If a pair of parents with these conditions have 6 children, what is the probability that:

(i) half of them will have sickle cell disease?

(ii) at most one child will have sickle cell disease?

(b) Another pair of parents with the same conditions have 5 children.

(i) How many children do they expect to have sickle cell disease?

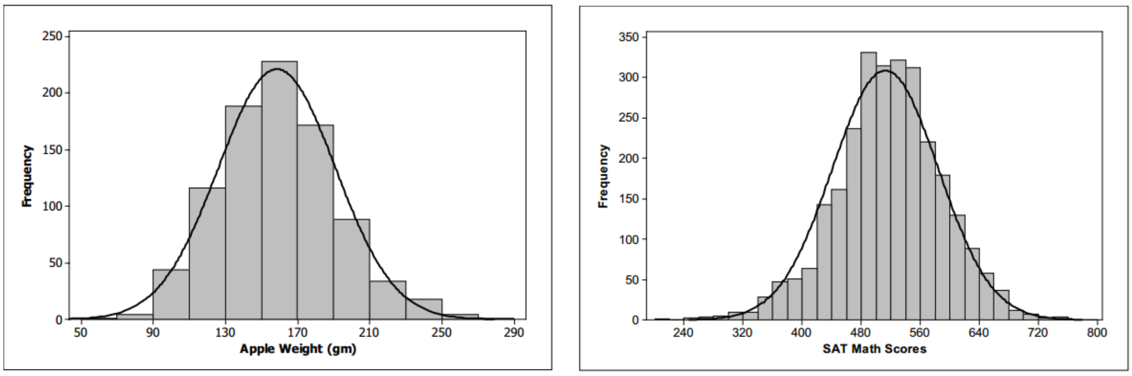
(ii) Is it rare to have exactly 4 children with sickle cell disease?

# 4. Continuous Random Variable: Normal

## 4.1 The Normal Curve

Histograms are often used to graph the distribution of the sample values for one particular quantitative variable. To make it even easier to focus on general shapes, sometimes statisticians draw a smooth curve through a histogram. These curves, drawn over histograms, summarize the overall patterns in data sets.

The curves can also be compared to spot similarities in shapes, even if the data sets that are being compared might not be of similar source or scale. For example, histograms of weights of Gala apples from an orchard and SAT Math scores from entering students at a US state university, are shown here:



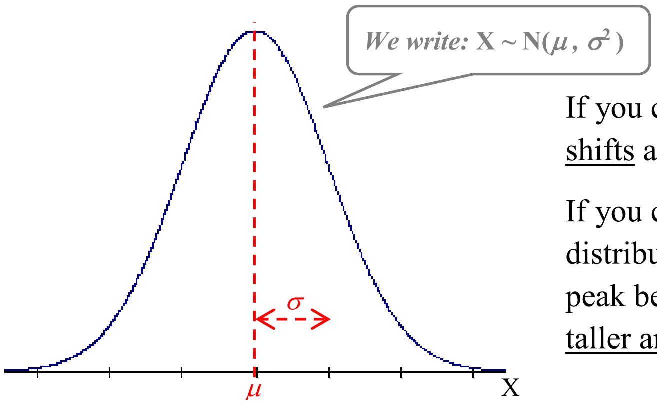

*Reference: "Normal Curves: Against All Odds—Inside Statistics," director. Films Media Group, 2013,* *http://fod.infobase.com/portalplaylists.aspx?wid=151497&xtid=111526.*

Notice that the curves on both histograms roughly have the same shape, even though the data sets are unrelated.

This special shape is called **normal curve**. It is symmetric with one peak, or simply called, **bell-shaped**. The mean $\mu$ and the median are at the same point right in the middle. In fact, many distributions in the natural world exhibit this normal curve shape.

The bars in the histogram represent the actual sample data collected. The curve represents our idealized assumption of what the whole population would look like, based on the actual data.

An important feature of any normal curve is that it can be completely defined by its **mean** $\mu$, and its **standard deviation** $\sigma$.



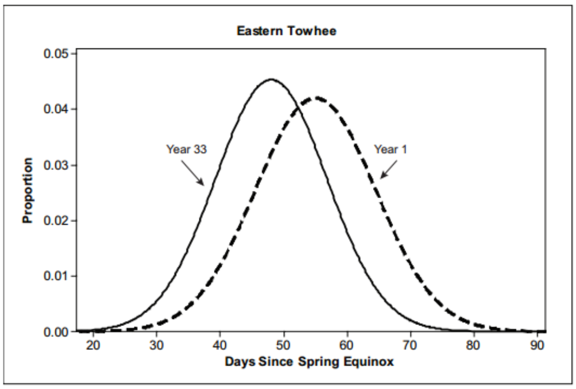

If you change the mean $\mu$, the whole curve just shifts along the $x$-axis.

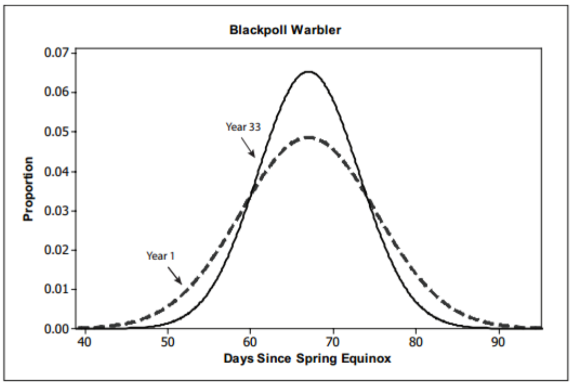
If you change the standard deviation $\sigma$, the distribution spread will change, such that the peak becomes flatter and wider, or becomes taller and narrower.

It is easier to make comparisons if we convert each bell-shaped smooth histogram into a **normal density curve**. To do this we change the scaling on the $y$-axis from a simple count or frequency, to a relative frequency or proportion.

With this new scale, the total area under the density curve is 1, and represents 100% of the data. Thus, 50% of the data falls below mean $\mu$, and the other 50% of the data falls above.

***Example 7:*** The normal density curves, in the graphs below, represent migration pattern of the Eastern Towhee and the Blackpoll Warbler birds in Manomet since 1970, at year 1 and year 33.




*Reference: "Normal Curves: Against All Odds—Inside Statistics," director. Films Media Group, 2013, http://fod.infobase.com/portalplaylists.aspx?wid=151497&xtid=111526.*

The following observations can be made from the Eastern Towhee graphs:

- The mean days of arrival for year 1 is later, because its curve is to the ________ of year 33's curve.
- The standard deviation of days of arrival for year ____ is smaller, because its curve is taller and pointier, and the data less spread out.
- The first arrivals for both years are happening about the same time.
- By day 48, about 50% of the birds had arrived for year 33, but only about ____% had arrived for year 1.
- About half of the birds had arrived for year 1 by day ____.

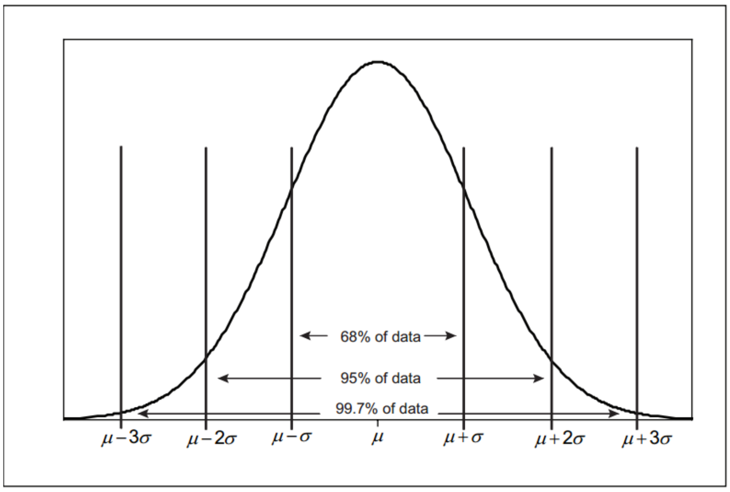
The following observations can be made from the Blackpoll Warbler graphs:

- The mean days of arrival for both years are ____________________.
- The first arrivals for year ____ is later.
- By day 56, about 10% of the birds had arrived for year 1, but only about ____% had arrived for year 33.
- About half of the birds had arrived for both year 1 and year 33 by day ____.

## 4.2 Empirical Rule

Recall that a normal curve is symmetric, single-peaked and bell-shaped, and it is completely described by its mean $\mu$ and standard deviation $\sigma$.

Normal curves have a unique feature that can be summed up by the **empirical rule**. It is also known as the ________________ rule.




- Approximately _____% of the data falls within _____standard deviation of the mean.
- Approximately _____% of the data falls within _____standard deviations of the mean.
- Approximately _____% of the data falls within _____standard deviations of the mean.
- The __________________________ is a natural yardstick for any measurements that follow a normal distribution.

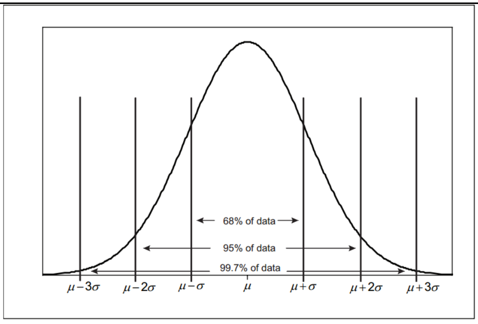
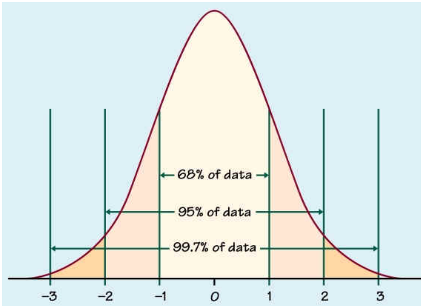
***Example 8:*** The birth weight of (full-term) babies in the United States (U.S.) is normally distributed with mean $\mu = 3.4$ kg and standard deviation $\sigma = 0.5$ kg.
Sketch the normal curve for the birth weight of U.S. babies, marking out the standard deviation as the "yardsticks".

Hence,

- approximately _____% of U.S. babies weigh between _____ kg and _____ kg;
- approximately _____% of U.S. babies weigh between _____ kg and _____ kg;
- approximately _____% of U.S. babies weigh between _____ kg and _____ kg;
- approximately _____% of U.S. babies weigh below _____ kg or above _____ kg.

## 4.3 Z-Score

We can figure out what is called the **standardized** value of any observation. This unitless value, often called a **z-score**, tells us how many ________________________ our observation falls from the mean and in which direction. It is a way to convert data from a normal distribution into a **standard normal distribution**. Let's see the similarities and differences:

| **normal distribution** | **standard normal distribution** |
|---|---|
|   |   |
| **Centre is at mean $\mu$** | **Centre is at mean $\mu = 0$** |
| **Spread is standard deviation $\sigma$** | **Spread is standard deviation $\sigma = 1$** |
| **We write: X ~ N($\mu$, $\sigma^2$)** | **We write: Z ~ N(0, $1^2$)** |
| Empirical rule:<br>• about 68% of data between $\mu-\sigma$ and $\mu+\sigma$<br>• about 95% of data between $\mu-2\sigma$ and $\mu+2\sigma$<br>• about 99.7% of data between $\mu-3\sigma$ and $\mu+3\sigma$ | Empirical rule:<br>• about 68% of data between −1 and 1<br>• about 95% of data between −2 and 2<br>• about 99.7% of data between −3 and 3 |

Mathematically, to convert $x$-value to z-score: $\boxed{Z = \dfrac{X-\mu}{\sigma}}$

Z-scores allow us to compare observations from two different normal distributions by standardizing them with a common scale.

***Example 9:*** The birth weight of (full-term) babies in the United States (U.S.) is normally distributed with mean $\mu = 3.40$ kg and standard deviation $\sigma = 0.50$ kg. Comparatively, the birth weight of (full-term) babies in Germany is normally distributed with mean $\mu = 3.56$ kg and standard deviation $\sigma = 0.45$ kg. A guideline is to be set such that babies who weigh 4.5 kg and above at birth is considered "overweight" and will require closer monitoring by pediatricians.

(a) Convert the guideline to z-scores for both countries.

(b) Sketch the standard normal curve and locate the z-scores on the curve.

(c) Which of the two countries will have more "overweight" babies?

## 4.4 Reading Probability from Z-Table

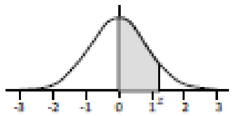
How do we find the proportion of data if the z-score does not fall exactly at 1, 2 or 3 standard deviations away from the mean?

The area under the standard normal density curve represents the proportion of data or probability. This can be found using a standard normal table or statistical software.

There are different formats of the **standard normal table** or simply, **z-table**. We shall use the format that shows areas (probabilities) that are measured from the centre of the standard normal curve, that is, from zero to the desired z-score.

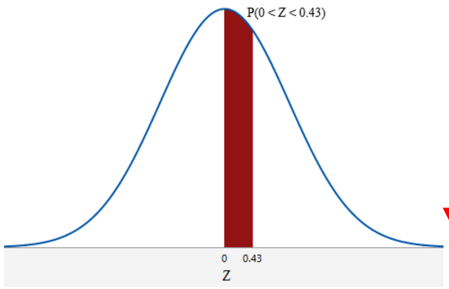
Area under the Standard Normal Curve from 0 to $z$

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



| z | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|---|---|---|---|---|---|---|---|---|---|---|
| 0.0 | .0000 | .0040 | .0080 | .0120 | .0160 | .0199 | .0239 | .0279 | .0319 | .0359 |
| 0.1 | .0398 | .0438 | .0478 | .0517 | .0557 | .0596 | .0636 | .0675 | .0714 | .0754 |
| 0.2 | .0793 | .0832 | .0871 | .0910 | .0948 | .0987 | .1026 | .1064 | .1103 | .1141 |
| 0.3 | .1179 | .1217 | .1255 | .1293 | .1331 | .1368 | .1406 | .1443 | .1480 | .1517 |
| 0.4 | .1554 | .1591 | .1628 | .1664 | .1700 | .1736 | .1772 | .1808 | .1844 | .1879 |
| 0.5 | .1915 | .1950 | .1985 | .2019 | .2054 | .2088 | .2123 | .2157 | .2190 | .2224 |
| 0.6 | .2258 | .2291 | .2324 | .2357 | .2389 | .2422 | .2454 | .2486 | .2518 | .2549 |
| 0.7 | .2580 | .2612 | .2642 | .2673 | .2704 | .2734 | .2764 | .2794 | .2823 | .2852 |
| 0.8 | .2881 | .2910 | .2939 | .2967 | .2996 | .3023 | .3051 | .3078 | .3106 | .3133 |
| 0.9 | .3159 | .3186 | .3212 | .3238 | .3264 | .3289 | .3315 | .3340 | .3365 | .3389 |
| 1.0 | .3413 | .3438 | .3461 | .3485 | .3508 | .3531 | .3554 | .3577 | .3599 | .3621 |
| 1.1 | .3643 | .3665 | .3686 | .3708 | .3729 | .3749 | .3770 | .3790 | .3810 | .3830 |
| 1.2 | .3849 | .3869 | .3888 | .3907 | .3925 | .3944 | .3962 | .3980 | .3997 | .4015 |
| 1.3 | .4032 | .4049 | .4066 | .4082 | .4099 | .4115 | .4131 | .4147 | .4162 | .4177 |
| 1.4 | .4192 | .4207 | .4222 | .4236 | .4251 | .4265 | .4279 | .4292 | .4306 | .4319 |
| 1.5 | .4332 | .4345 | .4357 | .4370 | .4382 | .4394 | .4406 | .4418 | .4429 | .4441 |
| 1.6 | .4452 | .4463 | .4474 | .4484 | .4495 | .4505 | .4515 | .4525 | .4535 | .4545 |
| 1.7 | .4554 | .4564 | .4573 | .4582 | .4591 | .4599 | .4608 | .4616 | .4625 | .4633 |
| 1.8 | .4641 | .4649 | .4656 | .4664 | .4671 | .4678 | .4686 | .4693 | .4699 | .4706 |
| 1.9 | .4713 | .4719 | .4726 | .4732 | .4738 | .4744 | .4750 | .4756 | .4761 | .4767 |
| 2.0 | .4772 | .4778 | .4783 | .4788 | .4793 | .4798 | .4803 | .4808 | .4812 | .4817 |
| 2.1 | .4821 | .4826 | .4830 | .4834 | .4838 | .4842 | .4846 | .4850 | .4854 | .4857 |
| 2.2 | .4861 | .4864 | .4868 | .4871 | .4875 | .4878 | .4881 | .4884 | .4887 | .4890 |
| 2.3 | .4893 | .4896 | .4898 | .4901 | .4904 | .4906 | .4909 | .4911 | .4913 | .4916 |
| 2.4 | .4918 | .4920 | .4922 | .4925 | .4927 | .4929 | .4931 | .4932 | .4934 | .4936 |
| 2.5 | .4938 | .4940 | .4941 | .4943 | .4945 | .4946 | .4948 | .4949 | .4951 | .4952 |
| 2.6 | .4953 | .4955 | .4956 | .4957 | .4959 | .4960 | .4961 | .4962 | .4963 | .4964 |
| 2.7 | .4965 | .4966 | .4967 | .4968 | .4969 | .4970 | .4971 | .4972 | .4973 | .4974 |
| 2.8 | .4974 | .4975 | .4976 | .4977 | .4977 | .4978 | .4979 | .4979 | .4980 | .4981 |
| 2.9 | .4981 | .4982 | .4982 | .4983 | .4984 | .4984 | .4985 | .4985 | .4986 | .4986 |
| 3.0 | .4987 | .4987 | .4987 | .4988 | .4988 | .4989 | .4989 | .4989 | .4990 | .4990 |
| 3.1 | .4990 | .4991 | .4991 | .4991 | .4992 | .4992 | .4992 | .4992 | .4993 | .4993 |
| 3.2 | .4993 | .4993 | .4994 | .4994 | .4994 | .4994 | .4994 | .4995 | .4995 | .4995 |
| 3.3 | .4995 | .4995 | .4995 | .4996 | .4996 | .4996 | .4996 | .4996 | .4996 | .4997 |
| 3.4 | .4997 | .4997 | .4997 | .4997 | .4997 | .4997 | .4997 | .4997 | .4997 | .4998 |
| 3.5 | .4998 | .4998 | .4998 | .4998 | .4998 | .4998 | .4998 | .4998 | .4998 | .4998 |
| 3.6 | .4998 | .4998 | .4999 | .4999 | .4999 | .4999 | .4999 | .4999 | .4999 | .4999 |
| 3.7 | .4999 | .4999 | .4999 | .4999 | .4999 | .4999 | .4999 | .4999 | .4999 | .4999 |
| 3.8 | .4999 | .4999 | .4999 | .4999 | .4999 | .4999 | .4999 | .4999 | .4999 | .4999 |
| 3.9 | .5000 | .5000 | .5000 | .5000 | .5000 | .5000 | .5000 | .5000 | .5000 | .5000 |

So, how do we read this z-table? Here is an illustration:




Given z-score = 0.43, then
$P(0 < Z < 0.43) = 0.1664$

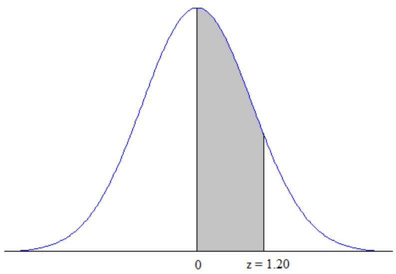
**Read down the first column for 1st decimal place of z-score, then the first row for the 2nd decimal place of z-score.**

| z | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|---|---|---|---|---|---|---|---|---|---|---|
| 0.0 | .0000 | .0040 | .0080 | .0120 | .0160 | .0199 | .0239 | .0279 | .0319 | .0359 |
| 0.1 | .0398 | .0438 | .0478 | .0517 | .0557 | .0596 | .0636 | .0675 | .0714 | .0754 |
| 0.2 | .0793 | .0832 | .0871 | .0910 | .0948 | .0987 | .1026 | .1064 | .1103 | .1141 |
| 0.3 | .1179 | .1217 | .1255 | .1293 | .1331 | .1368 | .1406 | .1443 | .1480 | .1517 |
| 0.4 | .1554 | .1591 | .1628 | .1664 | .1700 | .1736 | .1772 | .1808 | .1844 | .1879 |

Reading from the z-table, the area between 0 and desired z-score 0.43 is 0.1664, or 16.64%.

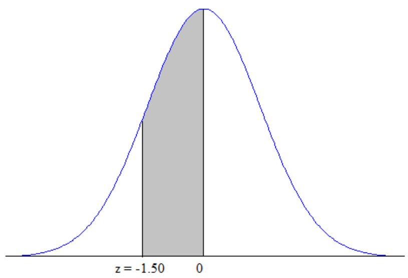
***Example 10:*** Find the probability in each of the following:

(a)




$P(0 < Z < 1.20) =$

(b)

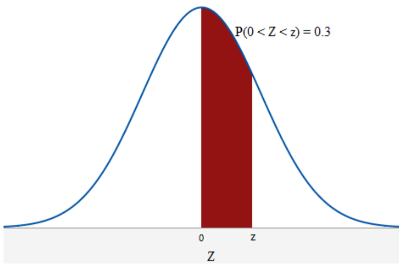



$P(-1.50 < Z < 0) =$

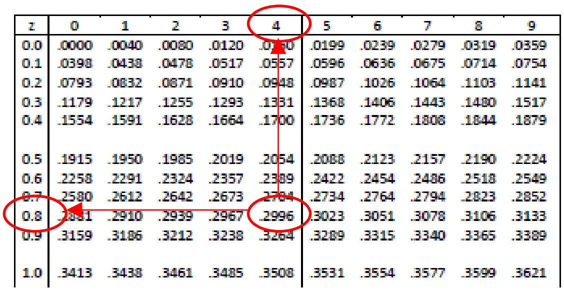
(c) $P(Z < 0.37) =$

(d) $P(-0.72 < Z < 1.5) =$

Conversely, if we are given the probability (area), how do we get the corresponding z-score from the z-table? Here are two illustrations:




Given area between 0 and z-score is 0.3, then z = 0.84.

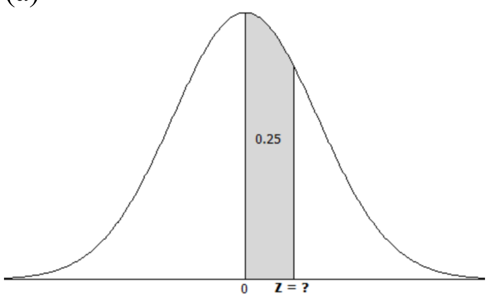


| z | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|---|---|---|---|---|---|---|---|---|---|---|
| 0.0 | .0000 | .0040 | .0080 | .0120 | .0160 | .0199 | .0239 | .0279 | .0319 | .0359 |
| 0.1 | .0398 | .0438 | .0478 | .0517 | .0557 | .0596 | .0636 | .0675 | .0714 | .0754 |
| 0.2 | .0793 | .0832 | .0871 | .0910 | .0948 | .0987 | .1026 | .1064 | .1103 | .1141 |
| 0.3 | .1179 | .1217 | .1255 | .1293 | .1331 | .1368 | .1406 | .1443 | .1480 | .1517 |
| 0.4 | .1554 | .1591 | .1628 | .1664 | .1700 | .1736 | .1772 | .1808 | .1844 | .1879 |
| 0.5 | .1915 | .1950 | .1985 | .2019 | .2054 | .2088 | .2123 | .2157 | .2190 | .2224 |
| 0.6 | .2258 | .2291 | .2324 | .2357 | .2389 | .2422 | .2454 | .2486 | .2518 | .2549 |
| 0.7 | .2580 | .2612 | .2642 | .2673 | .2704 | .2734 | .2764 | .2794 | .2823 | .2852 |
| 0.8 | .2881 | .2910 | .2939 | .2967 | .2996 | .3023 | .3051 | .3078 | .3106 | .3133 |
| 0.9 | .3159 | .3186 | .3212 | .3238 | .3264 | .3289 | .3315 | .3340 | .3365 | .3389 |
| 1.0 | .3413 | .3438 | .3461 | .3485 | .3508 | .3531 | .3554 | .3577 | .3599 | .3621 |

Reading from z-table, area of 0.3 (or 30%) is closest to area of 0.2996. Tracing back to the 1st column and 1st row, gives desired z-score of 0.84.

***Example 11:*** Find the unknown values in each of the following:

(a)




(b)


(c) Given $P(Z > a) = 0.2676$, find $a$.

(d) Given $P(Z > c) = 0.6845$, find $c$.

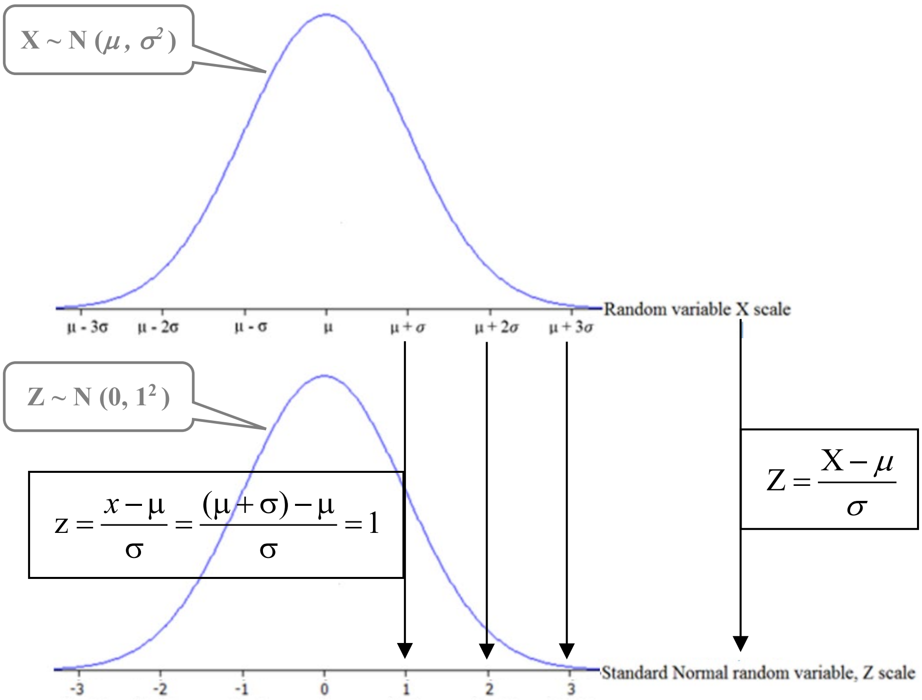
(e) If $P(-1.7 < Z < m) = 0.0666$, find $m$.

(f) If $P(-n < Z < n) = 0.251$, find $n$.

## 4.5 Normal Random Variable X

We now know how to obtain probabilities related to z-scores, not restricted by the Empirical Rule. What if we want to find probabilities concerned with any normal random variable X?

Any normal curve can be changed into the **standard** normal curve by transforming the scale on the horizontal axis into a z-score scale.




***Example 12:*** Given that a random variable X has a normal distribution with $\mu = 50$ and $\sigma = 10$, find the following probabilities:

(a) $P(X \geq 45)$

(b) $P(X > 62)$

(c) $P(45 < X < 62)$

***Example 13:*** Back to "overweight" babies. Let random variable X be the birth weight (in kg) of a U.S. baby, so $X \sim N\,(3.40, 0.50^2)$. Let random variable Y be the birth weight (in kg) of a German baby, so $Y \sim N\,(3.56, 0.45^2)$. Recall that the guideline of classifying "overweight" babies is 4.5 kg and above.

(a) Find the proportion of U.S. babies classified as "overweight".
(b) Find the proportion of German babies classified as "overweight".
(c) Find the percentage of German babies who weigh above 4 kg at birth, but below guideline of "overweight".
(d) What is the birth weight of a U.S. baby who is in the $90^{th}$ percentile?

***Example 14:*** The measurements inside a diameter of a cast-iron pipe is normally distributed with mean 5.01 cm and standard deviation 0.025 cm. The specification limits are set at $5.00 \pm 0.05$ cm. What percentage of the pipes are <u>not</u> acceptable?

***Example 15:*** The dimension of a circular mechanical part is normally distributed with mean 2 cm. The specification limits are set at $2.00 \pm 0.05$ cm.

(a) If the standard deviation is 0.03 cm, what is the percentage of mechanical parts that are within the specification limits?

(b) How small must the standard deviation be if 95% of the mechanical parts must be within specification limits?

# TUTORIAL 3

1. A box contains 100 items, of which, 27 are oversized and 16 are undersized. The items which are not of the right size will be rejected and the rest will be accepted.
   (a) What is the probability that a randomly selected item from this box is undersized?
   (b) What is the probability that a randomly selected item from this box is accepted?

2. In a class of 25 students who took a Mathematics test, the grade distribution is as follows:

| Grade | A | B | C | D | Fail |
|---|---|---|---|---|---|
| Count | 8 | 10 | 4 | 2 | 1 |

   (a) What is the probability that a randomly selected student from this class scored 'A'?
   (b) What is the failure rate (i.e. percentage of students who failed) of this class?
   (c) What is the probability that a random student from this class scored 'C' or 'D'?

3. In a sample of 446 cars stopped at a roadblock, 34 of the drivers did not have their seatbelts fastened. Among these 34 drivers, 21 of them were first-time offenders and let off with a warning; the rest were given demerit points and fined.
   (a) What is the sample space in this scenario?
   (b) What is the probability that a random driver stopped at that roadblock will have his/her seatbelt fastened?
   (c) What is the probability that a random driver stopped at that roadblock will be given demerit points and fined for not fastening seatbelt? Is this a rare event?

4. A random sample of 250 youths between 18 and 25 years-old was selected and asked about the number of email accounts they signed up for and their primary email account which they use most often. The data collected is tallied into the following table:

<table>
<tr><td colspan="2" rowspan="2"></td><td colspan="4">Number of email accounts</td></tr>
<tr><td>1</td><td>2</td><td>3</td><td>4 or more</td></tr>
<tr><td rowspan="3">Primary email account</td><td>Gmail</td><td>30</td><td>28</td><td>17</td><td>7</td></tr>
<tr><td>Outlook</td><td>25</td><td>31</td><td>26</td><td>10</td></tr>
<tr><td>Yahoo</td><td>20</td><td>26</td><td>19</td><td>11</td></tr>
</table>

   If a random youth is selected, what is the probability that the youth:
   (a) has 2 email accounts and uses Gmail primarily
   (b) has and uses Outlook only (i.e. has only 1 email account)
   (c) has 1 email account only
   (d) uses Yahoo primarily
   (e) does not use Gmail primarily
   (f) has at least 2 email accounts
   (g) has at least 2 email accounts and uses Outlook primarily
   (h) has at most 2 email accounts and does not use Yahoo primarily

5. Decide whether each of the scenarios described below fits the Binomial distribution. If it does, identify the values of $n$, $p$ and $q$, and list all the possible values of X. If it does not, explain why.

   (a) Cyanosis is the condition of having bluish skin due to insufficient oxygen in the blood. About 80% of the babies born with cyanosis recover fully. A hospital is caring for five babies born with cyanosis. The random variable X represents the number of babies that fully recover from cyanosis.

   (b) A survey company called 1000 people to ask whether they "agree, disagree or have no opinion" about the latest suggestion to abolish national service in a certain country. The random variable X represents the number of people in the survey who agree to the suggestion.

   (c) An inventory study determines that, on average, demand for a certain type of item is made 5 times a day. The random variable X represents the number of demands for that item per day.

   (d) It is conjectured that an impurity exists in three out of ten drinking wells in a rural community. To gain some insights to the extent of the problem, it is determined that some testing is necessary. Since testing all the wells is too resource-intensive, 15 wells were randomly selected for testing. The random variable X is the number of wells that contain impurity.

6. Refer to the coin toss example in Section 2.2 on page 5, where X is the number of heads that comes up in four coin tosses.

   (a) Explain why X can be a Binomial random variable.

   (b) What is the mean number of heads tossed in four coin tosses?

   (c) Use the Binomial probability model to derive the probability distribution of X.

7. Given $X \sim B(12, 0.4)$, find the following probabilities:

   (a) $P(X = 2)$

   (b) $P(X < 3)$

   (c) $P(X \geq 4)$

   (d) $P(2 \leq X < 5)$

8. Of all the students in a school, 30% travel to school by school bus. In a random sample of 10 students selected, find the following:

   (a) What is the mean number of students who travel to school by bus?

   (b) What is the probability that exactly 3 students travel to school by bus?

   (c) What is the probability that half of the students travel to school by bus?

   (d) Is it rare that more than 8 students travel to school by bus?

9. [*Refer to Case Study 2 on page 6*]
A space shuttle has two boosters, with each booster having three field joints. The field joint design has been improved such that the success launch rate of each field joint is now 0.985 (instead of 0.977).
Let X be the number of failures in the six field joints, and assume that the six field joints are independent.
(a) What is the probability that a single field joint will fail?
(b) What is the probability that exactly one field joint in the shuttle will fail?
(c) What is the probability that at least one field joint in the shuttle will fail?
(d) Has the safety of a space shuttle mission improved?
(e) Does X still fit a Binomial distribution if the field joints are not independent of each other?

10. An inspection plan for a microchip factory operates as follows:
A sample of ten microchips are randomly selected from a large batch.
If none from the sample is defective, then accept the batch.
If more than one microchip from the sample is defective, then reject the batch.
If exactly one microchip from the sample is defective, take another sample of ten microchips from the batch, and accept the batch only if none from this second sample is defective.
If a batch of microchips with 5% defectives is inspected, find the following:
(a) What is the mean number of defectives in the sample?
(b) What is the probability that the batch is accepted after the first sampling?
(c) What is the probability that the batch is accepted only after the second sampling?
(d) What is the probability that the batch is accepted?

11. Answer the following questions based on the Minitab Express output given for a Binomial random variable:
(a) How many trials were conducted?
(b) What is the probability of a success?
Is this a rare event?
(c) What is the probability of four successes?
Is this a rare event?

Probability Density Function: 4

Summary

| Input | |
|---|---|
| Distribution | Binomial |
| Number of trials | 35 |
| Event probability | 0.5 |
| Input value | 4 |

Probability Density

| x | P(X = x) |
|---|---|
| 4 | 0.000002 |

12. Answer the following questions based on the Minitab Express output given for a Binomial random variable:
    (a) How many trials were conducted?
    (b) What is the probability of a success?
    (c) What is the probability of getting more than 20 successes?
    (d) What is the probability of getting at least 20 successes?

**Cumulative Distribution Function: 20**

Summary

| Input | |
|---|---|
| Distribution | Binomial |
| Number of trials | 40 |
| Event probability | 0.35 |
| Input value | 20 |

Cumulative Probability

| x | P(X ≤ x) |
|---|---|
| 20 | 0.982719 |

13. Suppose that the random variable X is normally distributed with mean $\mu = 86$ and standard deviation $\sigma = 5$.
    (a) Sketch the normal curve and apply the Empirical Rule.
    (b) Hence, use the Empirical Rule to estimate the following probabilities:
        (i) $P(X < 96)$ (ii) $P(X \le 81)$ (iii) $P(76 \le X < 91)$
    (c) Convert the indicated $x$-value to z-score, hence find the corresponding probabilities.
        (i) $x = 80$ ; $P(X < 80)$ (ii) $x = 92$ ; $P(X > 92)$
        (iii) $x = 100$ ; $P(X < 100)$ (iv) $x = 72$ ; $P(X > 72)$
        (v) $x = 70$ ; $P(70 < X < 80)$ (vi) $x_1 = 85$, $x_2 = 95$ ; $P(85 < X < 95)$

14. Find the desired z-scores given the following probabilities:
    (a) $P(0 < Z < a) = 0.4753$ (b) $P(b < Z < 0) = 0.129$
    (c) $P(Z < c) = 0.97$ (d) $P(Z > d) = 0.864$
    (e) $P(Z > k) = 0.0217$ (f) $P(Z < l) = 0.271$
    (g) $P(-1 < Z < m) = 0.5$ (h) $P(1.5 < Z < w) = 0.0018$

15. Telephone calls from a call centre are monitored and found to have a mean duration of 452 seconds and a standard deviation of 123 seconds. Suppose that the distribution of call durations is approximately normal, determine the following:
    (a) What is the percentage of calls that last more than 10 minutes (i.e. 600 seconds)?
    (b) What is the percentage of calls that last more than 5 minutes?
    (c) What is the percentage of calls with duration between 300 seconds and 480 seconds?
    (d) If 250 calls are made from the call centre on a particular day, what is the mean number of calls that last more than 5 minutes? Round off your answer to the nearest whole number.

16. Components made by a certain process have a thickness which is normally distributed about a mean of 3.00 cm and a standard deviation of 0.03 cm. A component is classified as defective if its thickness lies outside the limits of 2.95 cm to 3.05 cm. Find the proportion of defective components. Hence, in a batch of 500 components, how many will be defective on average?

17. The mass of a bag of cookies is normally distributed with mean 450 g and standard deviation 15 g. Bags of cookies that have mass in the upper 7.5% are too heavy and must be repackaged. What is the maximum allowable mass a bag of cookies can be such that repackaging is not required?

18. You sell a brand of automobile tyre which has a life expectancy that is normally distributed with a mean of 30,000 km and a standard deviation of 2,500 km. You want to give a guarantee for free replacement of tyres that wear out too quickly. How should you word your guarantee if you are willing to replace approximately 10% of the tyres you sell?

19. A normal distribution has mean $\mu = 62.4$. Find its standard deviation if 20% of the area under the curve lies to the right of 79.2.

20. A vending machine is calibrated to dispense coffee into a 250ml paper cup. The amount of coffee dispensed into the cup is normally distributed with a standard deviation of 10ml. If the machine is allowed to overfill the cup 1% of the time, what should be set as the mean amount of coffee to be dispensed?

21. Suppose the life in hours of a certain electronic tube is normally distributed with mean $\mu = 160$ hours. The specification limits call for the product to last between 120 hours and 200 hours with probability 0.95. What is the maximum allowable standard deviation that the process can have and still maintain its quality?

22. Suppose a chemical manufacturer produces a product that is marketed in plastic bottles. The material is toxic, so the bottles must be tightly sealed. The manufacturer of the bottles must produce the bottles and caps within very tight specification limits. Suppose the caps will be acceptable to the chemical manufacturer only if their diameters are between 0.497 and 0.503 inch. When the manufacturing process for the caps is in control, cap diameter can be described by a normal distribution with $\mu = 0.5$ inch and $\sigma = 0.0015$ inch. If the process is *in control* (i.e. within specification limits), what percentage of the bottle caps would have diameters outside the chemical manufacturer's specification limits? Use empirical rule to approximate the answer.

## ANSWERS

1. (a) 0.16 (b) 0.57
2. (a) 0.32 (b) 0.04 (c) 0.24
3. (a) "fastened seatbelt", "did not fasten seatbelt and let off with warning", "did not fasten seatbelt and given demerit points and fined" (b) 0.924 (c) 0.0291, yes
4. (a) 0.112 (b) 0.1 (c) 0.3 (d) 0.304 (e) 0.672 (f) 0.7 (g) 0.268 (h) 0.456
5. (a) Yes. $X \sim B(n = 5, p = 0.8)$, $q = 0.2$
   (b) No, because there are 3 possible outcomes in each trial – agree, disagree, neutral. However, if we group the outcomes such that 'success' is agree and 'failure' is either disagree or neutral, then we can fit a Binomial distribution $X \sim B(n = 1000, p = 1/3)$.
   (c) No, because the number of trials is not fixed, but possibly infinite.
   (d) Yes. $X \sim B(n = 15, p = 0.3)$, $q = 0.7$
6. (a) Fixed number of trials (four tosses => $n = 4$); only 2 possible outcomes (H or T); probability of success is constant from trial to trial ($p = 0.5$), this means the trials are independent. (b) 2 (c) Refer to probability distribution table on page 5.
7. (a) 0.0639 (b) 0.0834 (c) 0.775 (d) 0.419
8. (a) 3 (b) 0.267 (c) 0.103 (d) $P(X > 8) = 0.000144$, rare
9. (a) 0.015 (b) 0.0834 (c) 0.0867 (d) Improved (was 13%), but still not safe because failed mission is not a rare event. (e) No, because it will violate one of the conditions of a Binomial distribution.
10. Let X be the number of defective microchips in a sample, then $X \sim B(10, 0.05)$
    (a) 0.5 (b) 0.5987 (c) 0.1887 (d) 0.7874
11. (a) 35 (b) 0.5, no (c) 0.000002, yes
12. (a) 40 (b) 0.35 (c) 0.0173 (d) 0.0363
13. (a) About 68% between 81 and 91, about 95% between 76 and 96, and about 99.7% between 71 and 101. (b) (i) 97.5% (ii) 16% (iii) 81.5%
    (c) (i) −1.2, 0.1151 (ii) 1.2, 0.1151 (iii) 2.8, 0.9974
    (iv) −2.8, 0.9974 (v) −3.2, 0.1144 (vi) −0.2, 1.8, 0.5434
14. (a) 1.965 (b) −0.33 (c) 1.88 (d) −1.10 (e) 2.02 (f) −0.61 (g) 0.41 (h) 1.51
15. (a) 11.51% (b) 89.25% (c) 48.35% (d) 223
16. 0.095, 47.5
17. 471.6 g
18. "Tyres that wear out before 26,800 km will be replaced free of charge!"
19. 20
20. 226.7 ml
21. 20.41 hours
22. The specified control limits are 2 S.D. away from the mean, hence by Empirical rule, about 5% of bottle caps would have diameters outside of the specification limits.

# LAB 3A : Binomial Distribution

**Learning Objectives:**

1. *Find probability of Binomial distribution using Minitab Express.*
2. *Find cumulative probability of Binomial distribution using Minitab Express.*

## Task 1

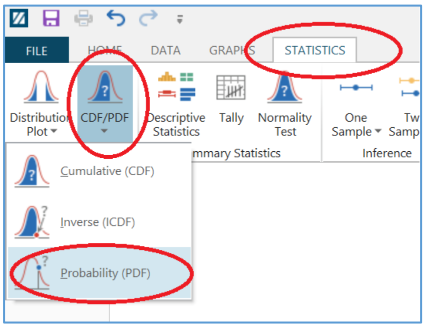
**Consider an experiment of tossing a dice 30 times. Assuming that the dice is fair, what is the probability of not getting any '6' in 30 tosses of a dice?**

Let X represents the Binomial random variable, number of '6' in 30 tosses of a dice.

Using the Binomial formula, $P(X=0) = {}_{30}C_0 \left(\frac{1}{6}\right)^0 \left(\frac{5}{6}\right)^{30} = 0.0042$

Now let us try to obtain this probability using Minitab Express.

Step 1: In Minitab Express, select **STATISTICS > CDF/PDF > Probability (PDF)**.



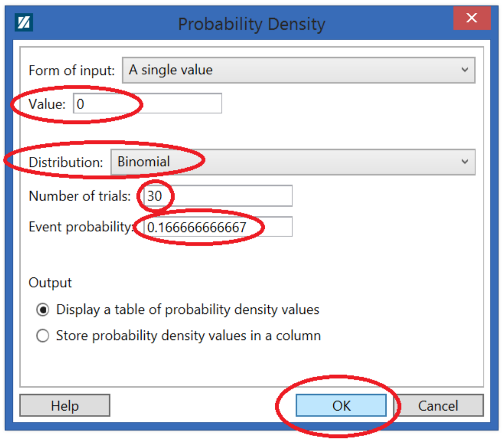

Step 2: For **Value**, key in *0*. (This means you want the number of successes to be zero.)

For **Distribution**, select *Binomial*.

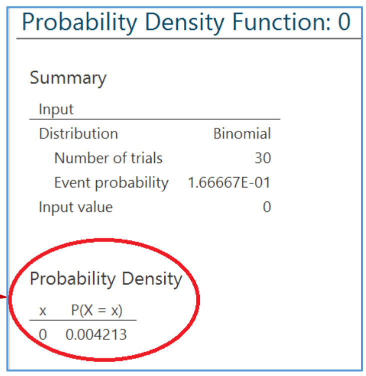
For **Number of trials**, key in *30*.

For **Event probability**, key in *0.166666666667*. (This is the probability of success, which in our case, is $\frac{1}{6}$ converted into decimal number.)

Click **OK**.




Step 3: The results will be displayed in the output window.




*Interpreting output*:
The probability of no success in 30 trials is 0.004213.
Notice that this is the same answer obtained using Binomial formula.

## Task 2

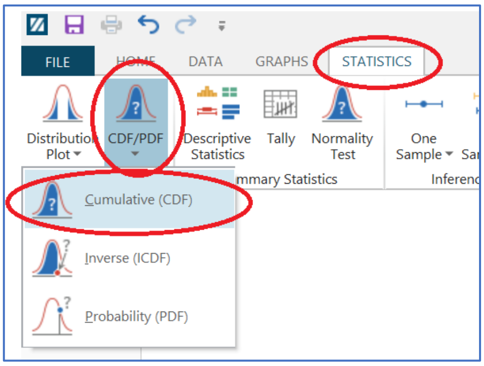
**Assuming that the dice is fair, what is the probability of getting <u>at most</u> five '6' in 30 tosses of a dice?**

For this example, calculation using Binomial formula is possible, but tedious!

$$P(X \le 5) = P(X=5) + P(X=4) + P(X=3) + P(X=2) + P(X=1) + P(X=0)$$
$$= 0.6165$$

Using Minitab Express to obtain this will be much easier!

<u>Step 1</u>: In Minitab Express, select **STATISTICS > CDF/PDF > Cumulative (CDF)**.




<u>Step 2</u>: Note that you are calculating <u>cumulative probability</u>.

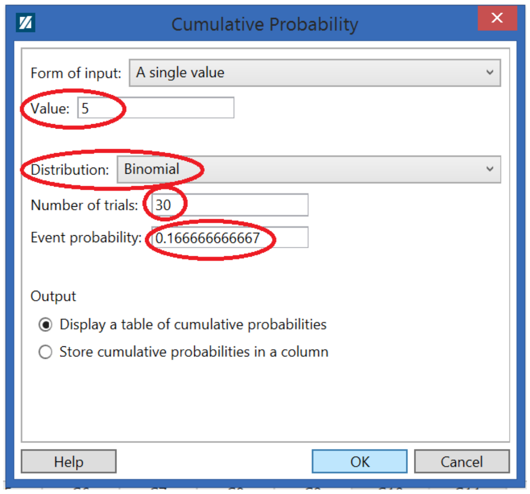
For **Value**, key in *5*. (This means you want the number of success to be <u>at most</u> five.)

For **Distribution**, select *Binomial*.

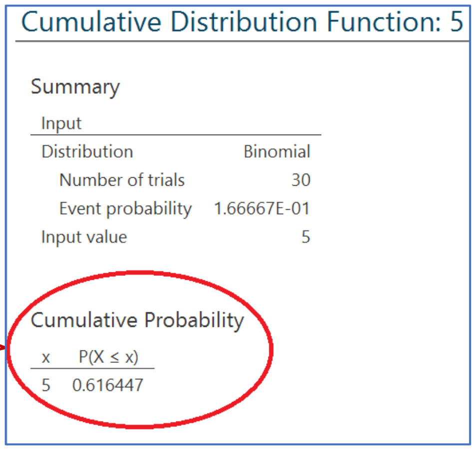
For **Number of trials**, key in *30*.

For **Event probability**, key in *0.166666666667*. (This is the probability of success, which in our case, is $\frac{1}{6}$ converted into decimal number.)

Click **OK**.




<u>Step 3</u>: The results will be displayed in the output window.




*Interpreting output*:
The probability of at most five successes in 30 trials is 0.616447. Notice that this is the same answer obtained using Binomial formula.

# LAB 3B : Normal Distribution

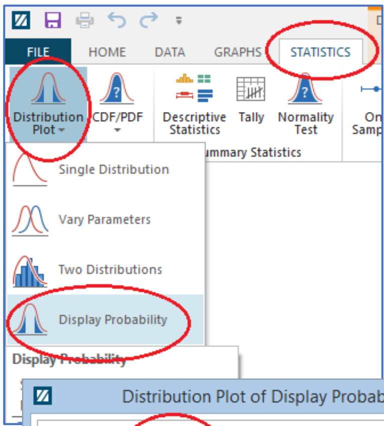
**Learning Objectives:**

1. *Find probability of standard normal distribution using Minitab Express.*
2. *Find probability of any normal distribution using Minitab Express.*
3. *Find the z-score given the probability in the standard normal distribution using Minitab Express.*

## Task 1

**On page 15 of this chapter, we found that P(0 < Z < 0.43) = 0.1664 by using the z-table. Let us obtain this number using Minitab Express.**

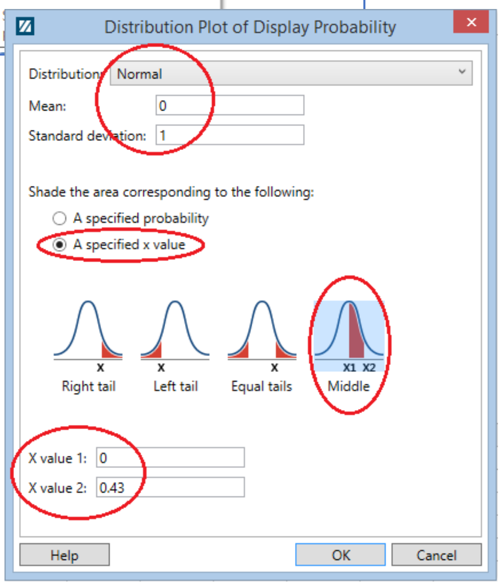
Step 1: In Minitab Express, select **STATISTICS > Probability Distribution > Display Probability**.



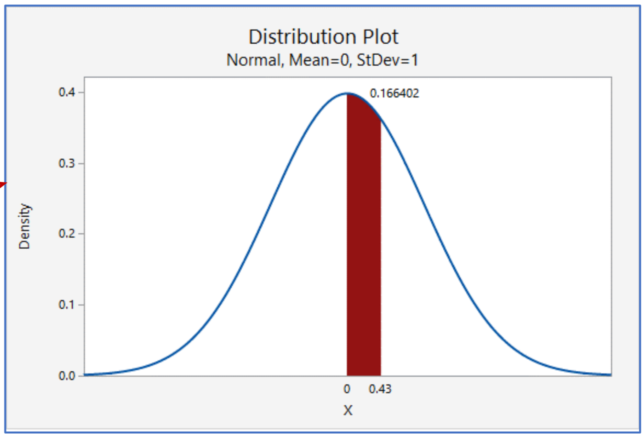

Step 2: For **Distribution**, select *Normal*.
For **Mean**, key in *0*.
For **Standard deviation**, key in *1*.
For **Shade the area corresponding to the following:**

- select the option *A specified x value*
- and then select the curve labelled *Middle*

For **X value 1**, key in *0*.
For **X value 2**, key in *0.43*.
Click **OK**.




Step 3: The results will be displayed in the output window.




*Interpreting output*:
The probability (area) of Z between 0 and 0.43 is 0.166402. Notice that this is the same answer obtained using z-table.

**Task 2**

**Suppose X is normally distributed with $\mu = 100$ and $\sigma = 20$, let us obtain $P(90 < X < 140)$ using Minitab Express.**

Step 1: Same as in Task 1, select **STATISTICS** > **Probability Distribution** > **Display Probability**.

Step 2: For **Distribution**, select *Normal*.

For **Mean**, key in *100*.

For **Standard deviation**, key in *20*.

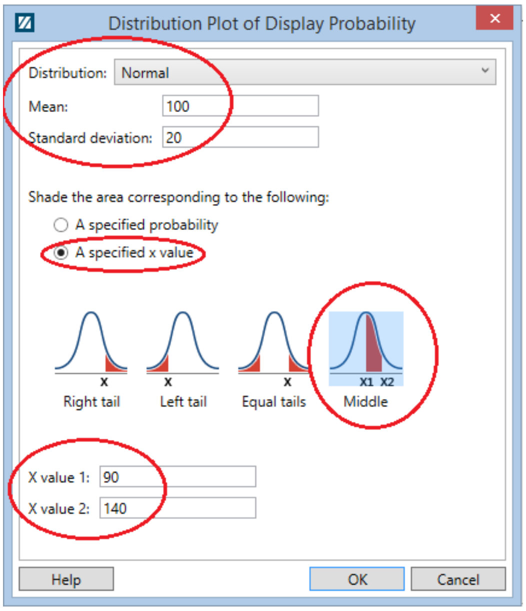
For **Shade the area corresponding to the following:**

- select the option *A specified x value*
- and then select the curve labelled *Middle*.

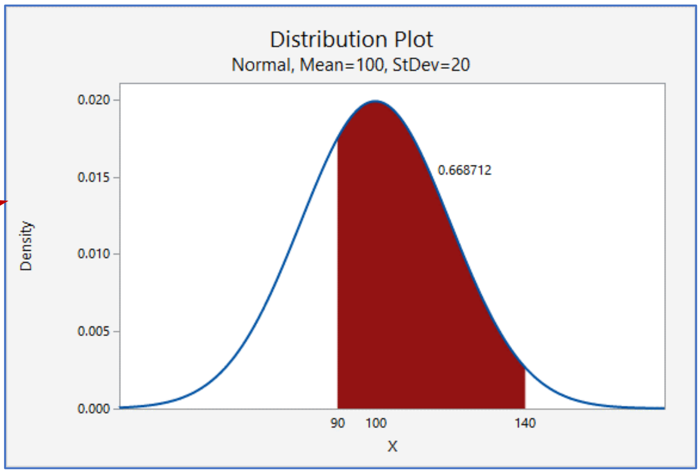
For **X value 1**, key in *90*.

For **X value 2**, key in *140*.

Click **OK**.




Step 3: The results will be displayed in the output window.




*Interpreting output*:
The probability (area) of X between 90 and 140 is 0.668712.

**Task 3**

**Suppose we know that P(Z > z) = 0.1314, that is, the area under the standard normal curve to the right of z is 0.1314, let us use Minitab Express to find the value of z.**

Step 1: Same as in Task 1 and Task 2, select **STATISTICS** > **Probability Distribution** > **Display Probability**.

Step 2: For **Distribution**, select *Normal*.

For **Mean**, key in *0*.

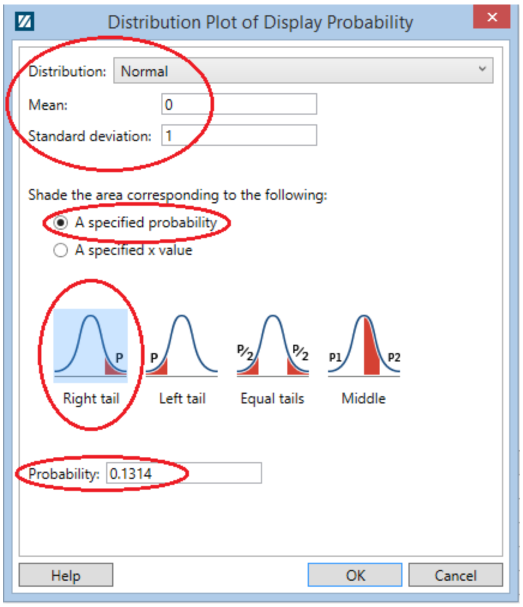
For **Standard deviation**, key in *1*.

For **Shade the area corresponding to the following:**

- select the option *A specified probability*
- and then select the curve labelled *Right tail*

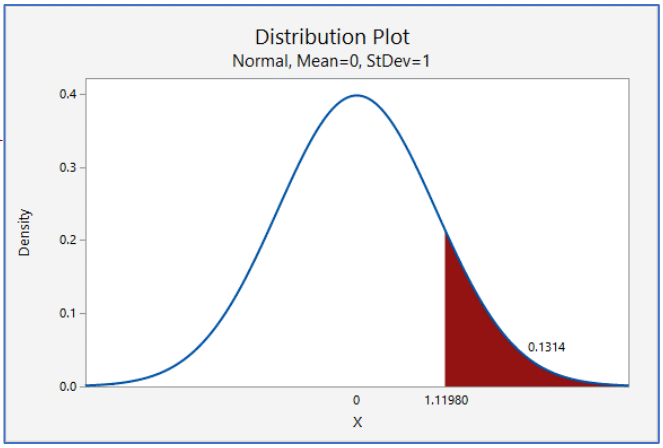
For **Probability**, key in *0.1314*.

Click **OK**.




Step 3: The results will be displayed in the output window.

*Interpreting output*:
If P(Z > z) = 0.1314, then z is 1.11980.




**Investigative Task**

**In Minitab Express, plot the Binomial distribution for *p* = 0.3, with different values of *n* from small (say, 10) to very large (say, 1000). What do you observe about the shape of the distribution as *n* gets larger? (You may repeat this for another value of *p*.)**