

Unit 9: Probability and Statistics

Grade 9 Math

Theoretical Probability – $\frac{\text{number of favorable outcomes to the event}}{\text{number of possible outcomes}}$

When we refer to probability we are usually referring to theoretical probability. It is what we EXPECT will happen.

Example: What is the probability of rolling a number greater than 4 on a standard die?

$$\frac{2}{6} = \frac{1}{3} \quad \begin{array}{l} \text{The 2 favorable outcomes are 5, 6} \\ \text{All the possible outcomes are 1, 2, 3, 4, 5, and 6.} \end{array}$$

Experimental Probability – the probability of an event calculated from experimental results.

Example: If you rolled a die 60 times, based on the theoretical probability you expect to land on 5 or 6, 20 times (one third). However, you may get different results when you complete the actual experiment. This is called experimental probability.

If you toss a coin ten times, what is the theoretical probability that you will land on heads? $\frac{5}{10} = \frac{1}{2}$

Student's Experiment: 7 heads, 3 tails so $\frac{7}{10}$

In theoretical probability we expected it to be 5 out of 10, since a head or tail has a 50% chance of happening. However, our experimental probability was different – it occurred 7 out of 10 times.

Sometimes when we make a decision, we take certain things for granted. When we do this, we make **assumptions**.

Ex: A health magazine claims that people who exercise 3 times per week for an hour each time reduce their risk of heart disease by 40%. Monique plays soccer 3 times a week for an hour each time. So, Monique thinks she has reduced her risk of heart disease by 40%.

- A). What assumptions is Monique making?
- B). How might the outcome change if the assumptions change?

ANSWERS:

A). Monique assumes that no other factors increase the risk of developing heart disease.

Other factors that should be considered are:

- Eating habits
- Blood pressure
- Family history
- Exposure to smoking
- Age

B). If Monique eats a lot of junk food, or she lives with people who smoke, Monique's risk of developing heart disease might increase.

If Monique's family has no history of heart disease, she lives with non-smokers and she eats a healthy diet, her risk of developing heart disease might decrease.

Probability - the likelihood that an event will occur. Can be written as **P(event)**

Ex #1: Hannah made 44 out of 50 basketball shots attempted. What is the probability she will sink the next shot? Write the answer as a fraction, decimal and percent.

$$P(\text{sinking next shot}) = \frac{\text{\# of favourable outcomes}}{\text{total \# outcomes}}$$

$$\begin{aligned} P(\text{sinking next shot}) &= \frac{44}{50} \text{ as a fraction} &= \frac{22}{25} \text{ as a fraction in lowest terms} \\ &= 0.88 \text{ as a decimal} &= 88\% \text{ as a percent} \end{aligned}$$

Ex #2: A bag contains 20 marbles. There are 15 red marbles and 5 blue. What is the theoretical probability that you will pick red?

$$P(\text{red marble}) = \frac{15}{20} = \frac{3}{4} = 0.75 \times 100 = 75\%$$

Ex #3A: What is the theoretical probability of rolling a 5 on a die?

$$P(5) = \frac{1}{6} = 0.1666$$

Ex#3B: In an experiment, you rolled a dice 25 times and rolled a 5 five times. What was your experimental probability?

$$EP = \frac{5}{25} = \frac{1}{5} = 0.20 \times 100 = 20\%$$

In general: The higher the number of trials you do, the closer the experimental probability will be to the theoretical probability.

Subjective Judgement

"Buying a lottery ticket today because you feel lucky", would be an example of using **subjective judgement**. This means the probability is based on emotional views and a strong personal opinion. Like basing probability on a hunch or a feeling.

Ex #4: Explain whether each decision is based on theoretical probability, experimental probability or subjective judgement.

- A. Listening to a song you dislike on the radio because you feel the next song must be one of your favourites.
- ↳ this is based on subjective judgement because you have a hunch that the next song to be played on the radio station will be one that you like.
- B. It is Isabella's experience that 4 out of 5 times, the prize in the cereal box is found at the bottom of the box. So she opens the cereal from the bottom.
- ↳ Isabella's decision to open the box from the bottom to find her prize is based on past experience. This is an example of experimental probability.
- C. Kayden is rolling a die 8 times. So far he has rolled a "4" , 7 times in a row. He predicts the next roll will likely not be a 4 since each number has an equal chance of being rolled.
- ↳ Kayden is using theoretical probability because he knows the probability of rolling any number is 1 out of 6. Therefore, the probability of him not rolling a 4 is greater and is 5 out of 6.

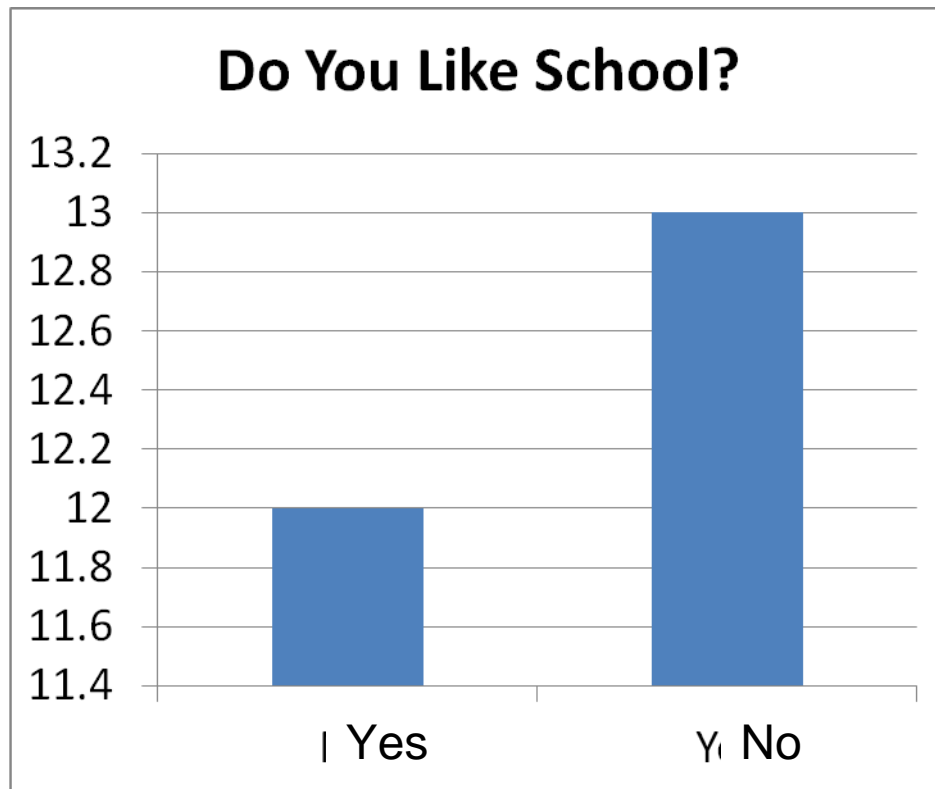
Manipulating Data

Statistics can be used to manipulate data in a way that best suits our own needs. It can be manipulated to give different results, and support opposing views. Therefore, be careful when collecting and analyzing data. We do not want to portray false information.

Ex #5: Jon wants to learn how to snowboard but does not want to take lessons. His mother insists that Jon take lessons. They find an article that claims: **"68% of snowboarding injuries occur during beginner lessons."** Explain how both Jon and his mother can use this statistic to support their own opinion.

- ↳ Jon's Mother can argue that the statistic proves lessons are important since beginners are prone to accidents.
- ↳ Jon can argue the statistic is a good reason not to take lessons because the likelihood of getting injured during a lesson is high.

Ex #6: How is this graph misleading?



- ↳ the way the bars are drawn makes it look like way more people dislike school than like school.
- ↳ however, if you notice the number, 25 people were asked 12 said yes and 13 said no. Therefore, it was only one more person who said they disliked school....not way more.
- ↳ It's almost 50/50. The bars should be almost the same height.

Potential Problems with Collecting Data

1. Timing – *when* data are collected could affect the results.

Ex: The results of a survey on ski-jackets may differ if the survey is conducted in the summer instead of the winter.

2. Privacy – people may not want to share personal information.

Ex: Students may not want to share their school grades.

3. Cultural Sensitivity – the survey must not offend other cultures.

Ex: A question about Christmas shopping may offend those who do not celebrate Christmas.

4. Cost – the cost of collecting data should be considered.

Ex: the cost of mailing a survey to a large number of people may be too expensive.

5. Time – the amount of time needed to collect the data should be considered.

Ex: A 30 minute survey conducted over the lunch period may not interest students. Most would not want to give up 30 minutes of their lunch!

6. Use of Language – the way a question is worded could lead people to answer in a certain way.

Ex: This question may lead students to answer “Yes”
Don’t you think the cafeteria food is too expensive?

A better wording might be....

Do you think food in the cafeteria is priced

a). too low ? b). just right ? c). too high ?

7. Bias - The question influences responses in favour of, or against the topic of data collection.

Ex: Don't you think movie ticket prices are too high? There is already a bias against the price in the question.

8. Ethics - The data collected must not be used for other purposes, otherwise your actions are considered unethical.

Ex: Surveying your class to find out their favourite snacks so you can have them available at your birthday party. But then you use the information to sell your classmates snacks between classes.

Sample vs Population

When a census is conducted, data are collected from all people or items in the population.

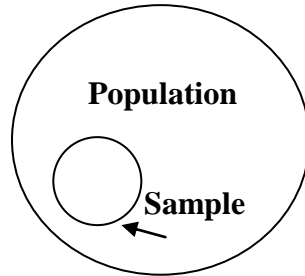
For example, if you survey all the grade 9 students at your school then a **census** was conducted.

Is it always reasonable to survey the entire population?

Why not?

- Cost
- Time
- Difficult
- Almost impossible to complete

When a population is too large, it is often necessary to collect a representative portion of the population – called a **sample**.



For example, surveying 50 out of the 300 grade 9 students in a large junior high school is a sample.

The **sample should represent the population**.

For example, the sample should contain a proportional number of girls and boys.

If the Grade 9 population of 300 students has 150 girls and 150 boys, then the sample of 50 students should have 25 girls and 25 boys.

The **size of the sample** should also be considered.

For example, if the sample size is too small the results may not represent the population. Ten grade 9 students may not provide **valid conclusions** (represent what is typical of the population).

Selecting a Sample

Simple Random Sampling

Each member of the population has an equal chance of being chosen.

Example: The name of each student in the class is put into a hat and 5 names are drawn.

Systematic Sampling

Every n^{th} member of the population is chosen.

Example: Every 10th name in the telephone directory is chosen.

Cluster Sampling

The population is divided into groups and every member of one group is chosen.

Example: A school has five grade 9 classes. All students in one class are chosen.

Self-Selected Sampling

Members of the population volunteer to be chosen.

Examples: 40 students volunteer to take a survey about homework habits.
NTV news has a question of the week. Viewers volunteer to phone in their answer.

Convenience Sampling

Members of the population who are convenient are chosen.

Example: The principal chooses the first 3 boys she sees to help her move chairs into the gym.

Stratified Random Sampling

Some members of each group of the population are randomly chosen.

Example: The school's population is divided into grades and 20 students from each grade are chosen.