Mathematics in the New Zealand Curriculum
Second Tier

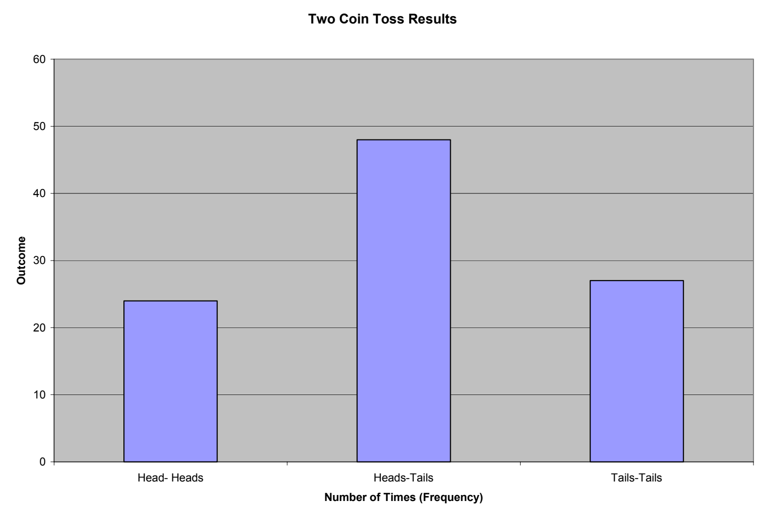
Strand: Statistics   Thread: Probability   Level: Four

Achievement Objective:

- Investigate situations that involve elements of chance by comparing experimental distributions with expectations from models of the possible outcomes, acknowledging variation and independence.
- Use simple fractions and percentages to describe probabilities

Exemplars of student performance:

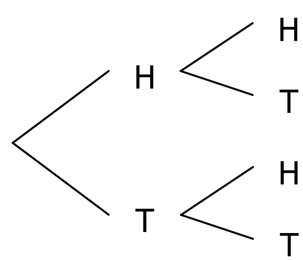
Exemplar One: Given a situation in which two coins are tossed, Centaine carries out 100 two-coin tosses
and records the distribution (pattern of occurrence), using a bar chart.

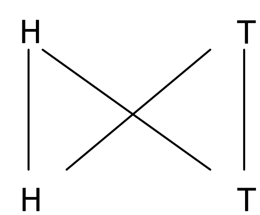


She finds the frequency (total occurrence) of each outcome; 48 Heads-Tails, 24 Heads-Heads, and 28 Tails-Tails.
Centaine notes that about half of the time the result is a Heads-Tails mixture, while two heads and two tails occurs approximately one-quarter of the time. She accepts the variation of her results from these exact proportions.

Centaine uses several different recording methods to find all of the possible outcomes for a two coin toss, including a tree diagram, table, and network.

Coin 1/ Coin 2	Heads	Tails
Heads	H-H	H-T
Tails	H-T	T-T





She recognizes that the theoretical probability of a Heads-Tails mix is one half (50%) and of two heads or two tails is one quarter (25%). Centaine links these theoretical proportions to the results of the trialing, acknowledging that the sample varies from the predicted proportions but the variation is reasonable.

Centaine's thinking exemplifies level four because she makes predictions for the likely outcome of an event based on analysis of all the possible outcomes. She describes the probabilities using simple fractions and accepts reasonable variation in a sample.

Exemplar Two: Each student in the class rolls a standard dice (1-6) 24 times. Here are the results from the whole class. Rua is able to predict expected variation of frequency (total occurrence) for any number in a future trial of 24 dice rolls. He does so using the distribution of frequencies shown in the class table of results.

Dice Number	1	2	3	4	5	6
Allan	4	4	5	6	3	3
Betty	5	6	4	1	4	4
Cath	6	3	3	3	2	7
Dee	3	5	4	3	6	3
Fred	4	3	5	3	5	4
Gil	4	7	3	2	4	4
Hene	5	3	3	3	6	4
Iria	2	7	3	4	4	4
Jobe	2	2	1	10	5	4

etc.

He notes that the frequency variation over thirty samples of 24 dice tosses is between zero and eleven. He concludes that zero occurrence of a number is very unlikely and an occurrence of over twelve is even more unlikely.

He is given this hypothetical record of two classmates' 24 dice throws.

Dice Number	1	2	3	4	5	6
Zoe	4	4	4	4	4	4

Dice Number	1	2	3	4	5	6
Yves	1	12	0	10	1	1

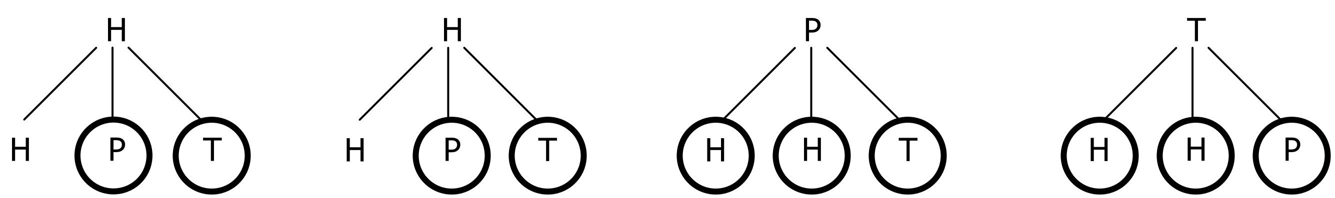
Rua concludes that both records are very unlikely, as Zoe's sample lacks any variation while Yves sample has frequencies that contain too much variation.

Rua's thinking exemplifies Level Four because he accepts frequency variation for 24 coin tosses and considers whether this variation is reasonable or unreasonable. He is able to recognise when the results of others are unlikely as they contain too much or too little variation based on the distribution of frequencies from his own large sample. Rua does so without using any formal measures of sampling variation.

Exemplar Three:

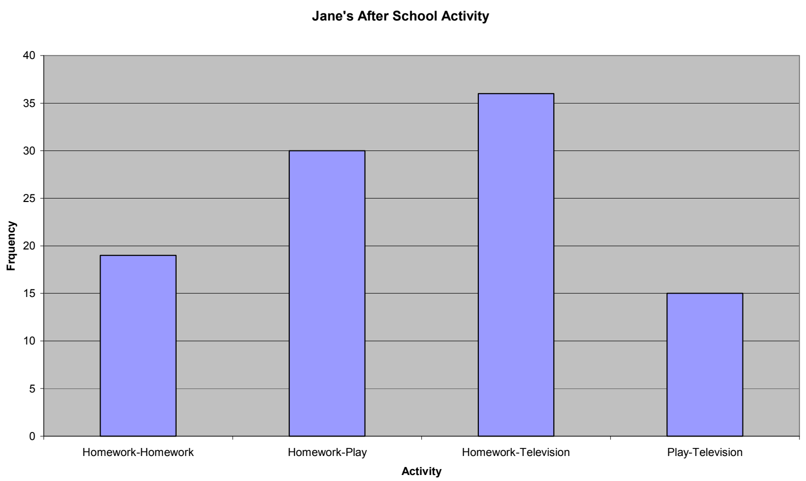
Jane has this scenario. To decide which two things she will do after school she shakes a plastic jar containing four ping-pong balls. On the balls are written homework, homework, play, television. The first two balls to come out tell the things to do. If homework comes out twice she has an extra long session. "What are Jane's chances of doing at least one fun thing each day after school?"

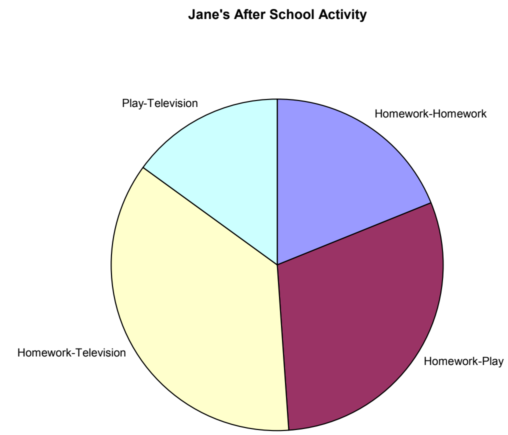
Jane draws a tree diagram to model the situation and circles the outcomes that include at least one fun thing.



Jane says that in only two out of twelve possible outcomes she does only homework. This is one-sixth of the time. Asked to predict how many times this will happen on the 100 school days left this year, she estimates that between 10 and 20 after school sessions will be spent doing only homework.

Jane carries out an experiment to verify her expectations. She trials 100 events of rolling two ping-pong balls out of the plastic bottle and uses both a bar graph and a pie chart to display the distribution of results.



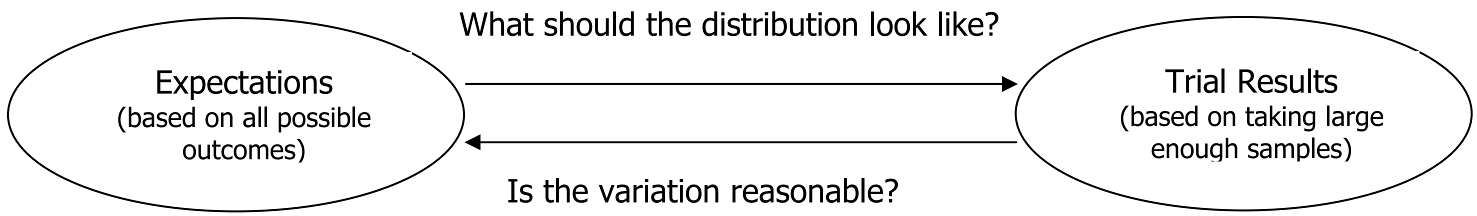


Jane accepts that the results are in line with her expectations but accepts that sampling variation might gives results outside the 10-20 range on occasions. She accepts this as reasonable because this is a situation involving chance. The pie chart confirms her prediction that the occurrence of Homework-Play and Homework-Television is about one-third each.

Jane’s thinking exemplifies level four because she is able to construct a model for all the possible outcomes of a simple two-stage event and describe her expectations for a given outcome using fractions. She accepts reasonable variation from her predicted outcome as she compares her trial data to her expectation.

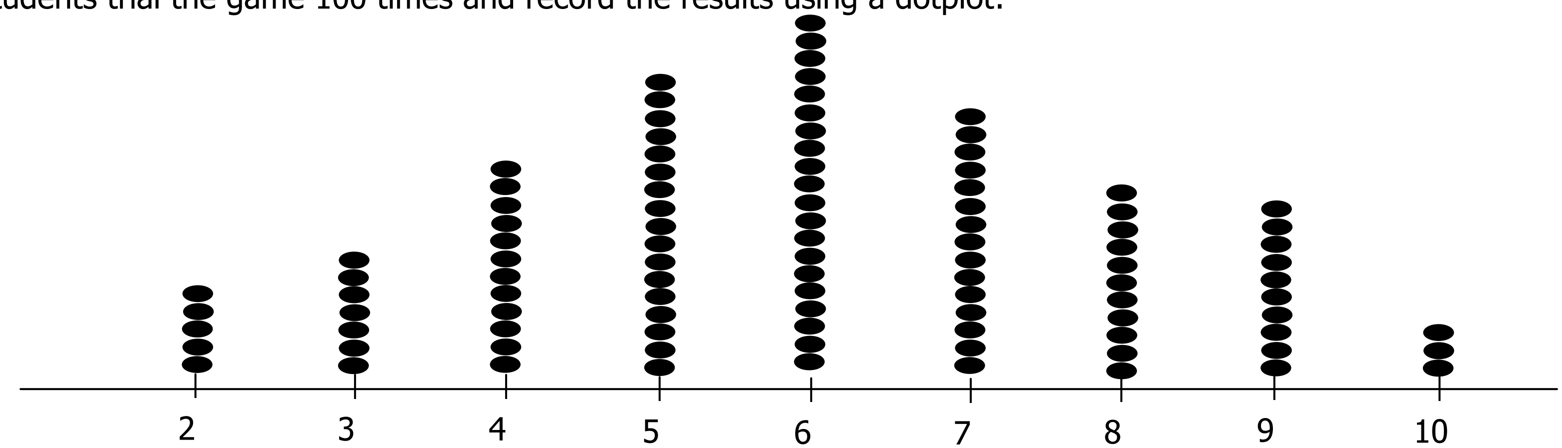
Important teaching ideas

Students at level four are learning to recognise that probabilities are estimated from trialing, and /or, in some cases, predicted by finding all the possible outcomes. They are also learning to detect when there is synergy or lack of synergy between their theoretical predictions and the results of trialing, given the natural variation that occurs with sampling.



To demonstrate this, consider a game played in pairs. Each player independently enters a secret single digit number, 1-5, on his or her calculator. They add the numbers to produce an odd or even answer. Is it better to be the person getting a point when the answer is even or better to be the odd winner? Students might consider the chances equal since there are the same number of odd numbers and even numbers. This is an expectation that is intuitive and not based on a model of all the possible outcomes. Extensive trialing might contradict their beliefs.

So the students trial the game 100 times and record the results using a dotplot:



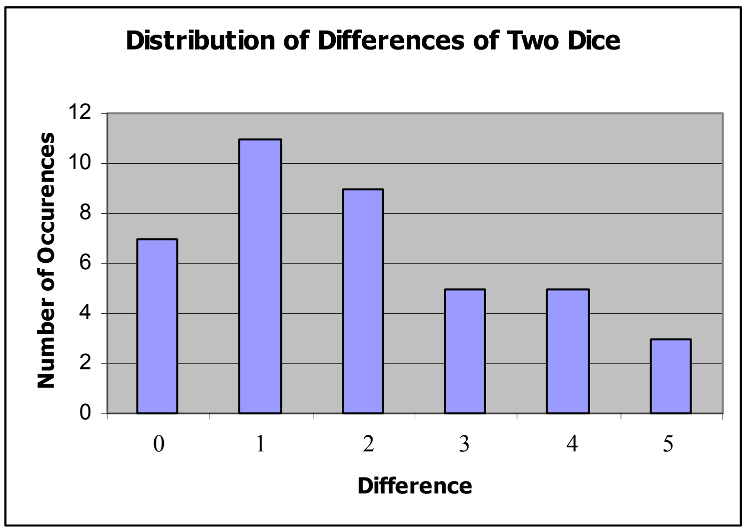
They note that 51 of the 100 games results in a win to even numbers. This is little variation from their original expectation but they seek a model of all the outcomes. They choose a table as the best recording method.

+	1	2	3	4	5
1	2	3	4	5	6
2	3	4	5	6	7
3	4	5	6	7	8
4	5	6	7	8	9
5	6	7	8	9	10

The table gives the students a model of all the outcomes. They assign probabilities to all the outcomes, e.g. five out of the 25 outcomes are sums of six so six has a $\frac{5}{25}$ or $\frac{1}{5}$ chance of occurring. The number of outcomes that are even is 13 out of 25 which is slightly more than one-half. The trial is supportive since the results are close to these expected values. The students conclude that the even answer person has a marginally better chance of winning.

Provide experiences in which the theoretical probability can be established easily using students' own strategies, or by using more formalized techniques like drawing tables, and diagrams. Drawing marbles from a can, using spinners, rolling dice and tossing coins are good examples. Balance this with situations in which the student can only estimate the probability by trialing, for example, "What proportion, of all people your age, is left-handed?" This reinforces the reality of statistical investigation where the actual probabilities are estimated through sampling.

It is important that students develop thinking about distributions. This involves consideration of the centre, spread and shape of the distribution. For example, tossing two dice and finding the difference between the numbers, e.g. 5 and 1 gives a difference of 4. After many trials the bar graph of the distribution might look like this:

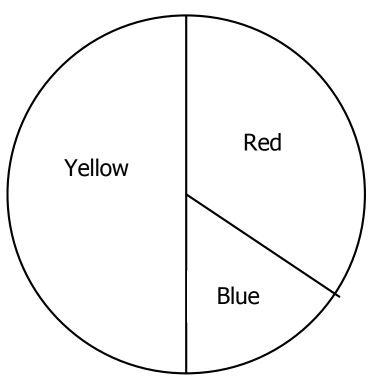


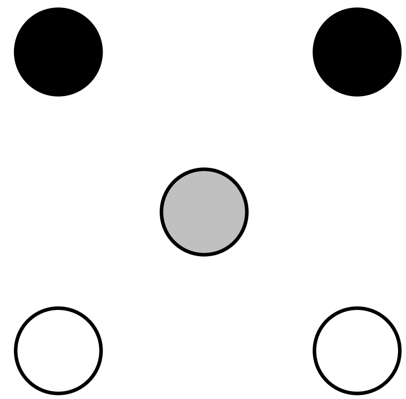
Focus on why the distribution is not symmetrical, by comparing the trial results to all the possible outcomes. The shape of the data suggests that differences of 0, 1, and 2 are more likely to occur that differences of 3, 4, and 5. A difference of 1 is the most common. In this case graphing the sampling distribution provokes the creation of a model to explain what happens. Many models will work including a table, list, tree diagram or network.

-	1	2	3	4	5	6
1	0	1	2	3	4	5
2	1	0	1	2	3	4
3	2	1	0	1	2	3
4	3	2	1	0	1	2
5	4	3	2	1	0	1
6	5	4	3	2	1	0

The table shows that 10 out of 36 outcomes give differences of 1 while 2 out of 36 outcomes give differences of 5. The theoretical probabilities explain the shape of the distribution.

Students need to connect their usual representations of fractions, like part-whole diagrams and ratios of discrete objects, with theoretical probabilities. Use situations in which the possibilities are easy to model physically or diagrammatically. Situations like using spinners, link nicely to students part whole models of fractions. Sampling discrete items from a container links nicely to ratio models:





The probability of getting blue on this spinner is $\frac{1}{6}$.

The probability of getting a black marble is $\frac{2}{5}$.

Connect other ways to represent probabilities, including:

- Percentages, e.g. $\frac{2}{5}$ chance can be expressed as 40%
- Decimals, e.g. $\frac{2}{5}$ chance can be expressed as 0.4
- Ratios, e.g. odds of 1:5 means there is one desired outcome and five undesired outcomes for an event (spinner above showing blue with one spin)

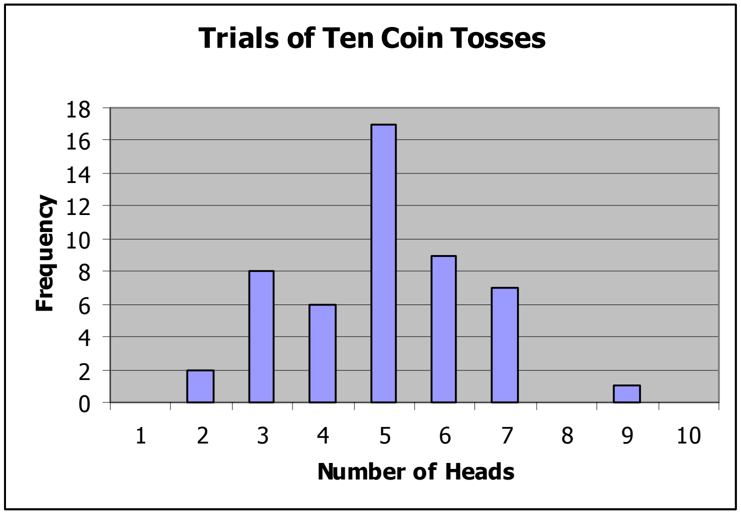
Students also need to have a repertoire of recording methods that help them find all the possible outcomes of increasingly complex events. These events should involve one or two stages. The outcomes of the events might be:

- Conditional, i.e. the outcome of one stage depends on the outcome that of the stage that precedes it. For example, five marbles in a can, two blue and three yellow. Take two marbles out of the can, with an aim of getting one of each colour. The chance of the second marble selected being of a certain colour is affected by the selection of the first marble.
- Independent, i.e. the outcome of the first stage had no impact on the possible outcomes of the second stage, e.g. two coins are tossed. The outcome of one coin toss has no impact on the possible outcomes of the second.

The repertoire of recording strategies should include organised lists, tables, tree diagrams of counts, and network diagrams (see level three).

Develop students' ideas about what variation seems reasonable for simple events involving chance while avoiding formal ideas about confidence intervals. Computer technology allows students to conduct many trials of an event very quickly. This requires an acceptance on their part that the software is actually enacting the event. Students should be encouraged to look at the effect of the number of trials has on the match between the trial results and a theoretical model based on all of the possible outcomes. The key idea is that larger samples usually give a closer match. This requires students to focus on proportions rather than absolute variation.

For example, students might use an applet that simulates trials of coin tossing. In this case, the theoretical model predicts that there is a half chance of heads and a half chance of tails. Each student uses the applet to produce a sample of ten tosses, then 100 tosses, then 1000 tosses, etc. The students note that in ten tosses the results are 6 heads and 4 tails. In absolute terms, the results are one off the predicted value and in proportional terms, that is 10%. For 100 tosses, the results are 52 heads and 48 tails. While the absolute difference of two from the theoretical prediction is greater, proportionally it is a tighter fit to the predicted 50%.Using the same software, create distributions from the collected results of all students' trials. For example, a class set of ten coin toss trials might be graphed like this:



Based on 50 trials students might conclude that $\frac{47}{50}$ (94%) of the trials resulted in the number of heads being in the range three to seven. Students could explore what happens to their conclusion as more trial samples are considered, and what the distribution for trials of 100 coin tosses might look like.
Note that we delay formal understanding of long-run frequency and confidence intervals until higher levels.

Look for situations in students' everyday lives that involve probability. Contexts such as weather, health risks, traffic lights, sports results and dice/card games can offer opportunities for probability investigations. Gambling scenarios such as Lotto and horse-racing are highly motivational and often make students aware of the poor chances of success. However, treat these contexts with sensitivity and with awareness of peoples' beliefs concerning gambling and the health risks of problem gambling. In some cases the theoretical probability can be determined by counting the number of possibilities while in most situations probabilities are estimated from a large scale sampling, e.g. weather forecasts are based on a long history of conditions.

Possible Resources
Figure It Out, Statistics Levels 3-4, Pages 17 – 24.
Figure It Out Years 7-8, Statistics Book One (Level 4), Pages 1, 19 – 24.
Figure It Out Years 7-8, Statistics Book Two (Level 4+), Pages 17 – 24.
nzmaths website units:
http://www.nzmaths.co.nz/node/139 (Probability: Beat it)
http://www.nzmaths.co.nz/node/137 (Probability: Top Drop)
http://www.nzmaths.co.nz/node/142 (Probability: Murphy's Law)
http://www.nzmaths.co.nz/node/143 (Probability: Gambling - who really wins?)
http://www.nzmaths.co.nz/node/119 (Probability: Greedy Pig)
Learning Objects
http://www.nzmaths.co.nz/learningobjects/317/4
http://nlvm.usu.edu/en/nav/frames_asid_305_g_3_t_5.html
http://nlvm.usu.edu/en/nav/frames_asid_310_g_3_t_5.html
http://nlvm.usu.edu/en/nav/frames_asid_186_g_3_t_5.html?open=activities
http://illuminations.nctm.org/ActivityDetail.aspx?ID=143