Basic Probability

Library, Teaching and Learning
2014

Lincoln University
Te Whare Wānaka o Aoraki
CHRISTCHURCH•NEW ZEALAND

New Zealand's specialist land-based university

Basic rules for calculating simple probability

Basic definition	Formula, symbols
Probability of an event, A, occurring:	$= \frac{\boldsymbol{number\ of\ successful\ outcomes}}{\boldsymbol{total\ number\ of\ possible\ outcomes}}$ $= \boldsymbol{P(A)}$
Complementary Events Events in the whole sample space but not one of the outcomes included in A are complementary.	$\boldsymbol{P(\tilde{A}) = P(not\ A) = 1 - P(A)}$
Limits of *P* P(any event occurring) lies between 0 and 1 If event A is certain **not** to happen: If event A is **certain** to happen:	$0 \le P(A) \le 1$ $\boldsymbol{P(A) = 0}$ $\boldsymbol{P(A) = 1}$
Union or General Addition Rule Probability that either one or other event occurs	$\boldsymbol{P(A\ or\ B) = P(A + B)}$
Mutually Exclusive Events Events that cannot both occur at the same time have no intersection. For events that are mutually exclusive:	$P(A\ or\ B) = P(A + B) = P(A) + P(B)$ $P(A\ and\ B) = 0$
Non- Mutually Exclusive Events Events that can both occur at the same time have some intersection. *Note that this rule applies regardless, as if there is no intersection, zero will be subtracted.*	$\boldsymbol{P(A\ or\ B) = P(A) + P(B) - P(A\ and\ B)}$
Statistically Independent Events Events where the occurrence of one event does not influence the likelihood of the other occurring **Note the reverse of this is also true:** **If $\boldsymbol{P(A\ and\ B) = P(A) \times P(B)}$, then A and B are independent.** **This is a specifically mathematical definition. Do not rely on "gut feeling" or instinct to tell you whether two events are statistically independent or not.**	If A and B are independent, then $\boldsymbol{P(A\ and\ B) = P(A) \times P(B)}$

Conditional Probability

This arises when we are calculating the probabilities of a particular event, A, given that we know the condition of another event, B. It is the probability that an event occurs <u>given that another event has occurred</u>.

$$P(A|B) = \frac{P(A \text{ and } B)}{P(B)} = \frac{n(A \text{ and } B)}{n(B)}$$

P(A|B) means : "The probability that A will occur given that B has already occurred."

Also, *Note:* If the events are **independent**, then

$$P(A|B) = \frac{P(A \text{ and } B)}{P(B)}$$

$$= \frac{P(A) \times P(B)}{P(B)} = P(A)$$

i.e., if events A and B are independent then the conditional probability that A occurs, given that event B has occurred, is simply the probability that event A occurs.

Expected Value

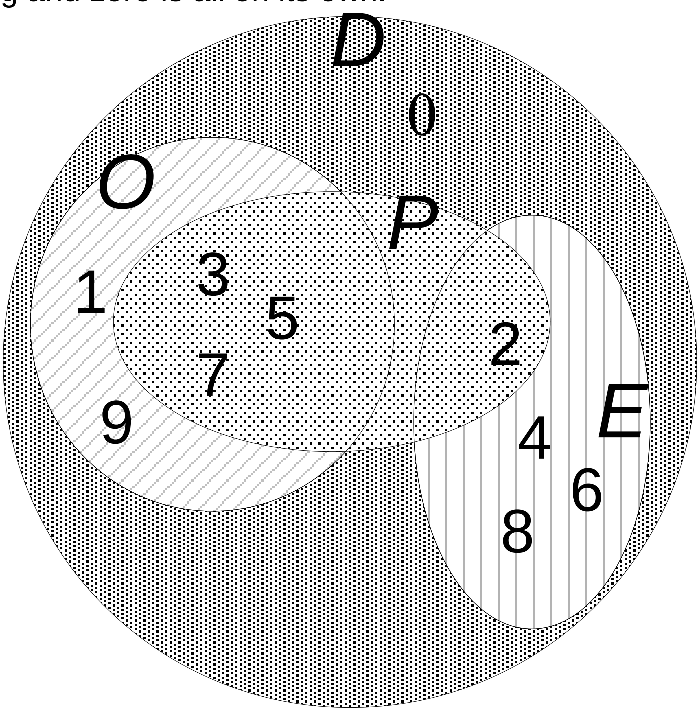
The ***expected value*** of a random variable is the ***mean*** of the random variable

$$E(X) = \mu = x_1 p(x_1) + x_2 p(x_2) + x_3 p(x_3) + \ldots x_n p(x_n)$$

That is, to work out the expected value of a random variable, multiply each possible value of *X* by its probability and add these products.

Presentation of information

As well as just being written out, information can be presented in a table or as a diagram.
Examine the following information.
The set of digits, D, contains the numbers {0, 1, 2, 3, 4, 5, 6, 7, 8, 9}
The set of even numbers, E, is {2, 4, 6, 8}
The set of odd numbers, O, is {1, 3, 5, 7, 9}
The set of prime numbers, P, is {2, 3, 5, 7}
This Venn diagram shows the relationships between the sets. Note there are some numbers in more than one grouping and zero is all on its own.



D = digits E = even numbers O = odd numbers P = prime numbers

A summary of this information could have been written in table form, showing the ***number*** of digits in each category:

	Odd numbers	**Even numbers**	**Neither**	**Total**
Prime	3	1	0	4
Not Prime	2	3	1	6
Total	5	4	1	10

That is, there are '3' digits that are both odd and prime, '2' digits that are both odd and not prime and '5' odd digits in total etc.

Study how the following probabilities are calculated, using the rules given above.

$P(even) = \frac{4}{10} = 0.4$ $\quad P(odd) = \frac{5}{10} = 0.5$ $\quad P(neither\ even\ or\ odd) = \frac{1}{10} = 0.1$

$P(prime\ and\ even) = \frac{1}{10} = 0.1$(from Venn Diagram)

or using formula: $\boldsymbol{P(A\ or\ B) = P(A) + P(B) - P(A\ and\ B)}$

$$P(prime\ or\ even) = \frac{4}{10} + \frac{4}{10} - \frac{1}{10} = \frac{7}{10}$$

$P(prime\ and\ odd) = \frac{3}{10}$ $\quad P(prime\ or\ odd) = \frac{5}{10} + \frac{4}{10} - \frac{3}{10} = \frac{6}{10}$

$P(even, odd\ or\ neither) = 1$ $\quad P(odd\ and\ even) = 0$

Probabilities are very easy to calculate if data is given in table form. If you are given information not in table form, try to tabulate it before you start your calculations.
The table is sometimes referred to as a *contingency table*.

In some texts you will see:

$\cap$ used for "intersection" and defined by the word "and".
$\cup$ used for "union" and defined by the word "or".

PROBABILITY

Addition law (events not mutually exclusive):

P(A or B) = P (A) + P(B) – P(A and B)

For mutually exclusive events: **P(A or B) = P(A) + P(B)**
P(A and B) = 0

Multiplication law: **P(A and B) = P(A)P(B|A)=P(B)P(A|B)**

If statistically independent:

P(A|B) = P(A) and P(B|A) = P(B)
P(A and B) = P(A)P(B)

DO NOT ABANDON YOUR OWN LOGIC – think about the questions and the likely answer.

PROBABILITY - PRACTICE QUESTIONS

1. If two events are mutually exclusive, the probability that they both occur is:

   A 0.00 B 0.50 C 1.00. D Cannot be determined from the information given

2. When using the general multiplication rule, $P(A \text{ and } B)$ is equal to:

   A $P(A|B).P(B)$ B $P(A).P(B)$ C $P(B)/P(A)$ D $P(A)/P(B)$

3. A recent survey of banks revealed the following distribution for the interest rate being charged on a home loan (based on a 15-year mortgage with a 20% deposit):

Interest rate	7.0%	7.5%	8.0%	8.5%	> 8.5%
Probability	0.12	0.23	0.24	0.35	0.06

   If a bank is selected at random from this distribution, what is the chance that the interest rate charged on a home loan will exceed 8.0%?

   A 0.06 B 0.41 C 0.59 D 1.00

*Use the following information for the next **two** questions.*

Mothers Against Drunk Driving is a very visible group whose main focus is to educate the public about the harm caused by drunk drivers. A study was recently done that emphasised the problem we all face with drinking and driving. Four hundred accidents that occurred on a Saturday night were analysed. Two items noted were the number of vehicles involved and whether alcohol played a role in the accident. The numbers are shown below:

	Number of Vehicles Involved Totals			**Totals**
Did alcohol play a role?	1	2	3	
Yes	50	100	20	170
No	25	175	30	230
Totals	75	275	50	400

4. What proportion of accidents involved alcohol and a single vehicle?

   A 25/400 B 50/400 C 195/400 D 245/400

5. Given that alcohol was not involved, what proportion of the accidents were multiple vehicle?

   A 50/170 B 120/170 C 205/230 D 25/230

6. The connotation ‘expected value’ or ‘expected gain’ from playing Roulette at a casino means:

   A the amount you expect to ‘gain’ on a single play
   B the amount you expect to ‘gain’ in the long run over many plays
   C the amount you need to ‘break even’ over many plays
   D the amount you should expect to ‘gain’ if you are lucky

7. If two events are collectively exhaustive, the probability that one or the other occurs is

A 0 B 0.50 C 1.00 D Cannot be determined from the information given

8. There are 100 female students and 230 male students in a class. The probability that a randomly picked student is a female is:

A 0 B 0.50 C 0.30 D Cannot be determined from the information given

9. According to a survey of American households, the probability that the residents own two cars IF annual household income is over $25,000 is 80%. Of the households surveyed, 60% had incomes over $25,000 and 70% had two cars. The probability that the residents of a household own two cars AND have an income less than or equal to $25,000 a year is:

A 0.12 B 0.18 C 0.22 D 0.48

10. A company has two machines that produce widgets. An older machine produces 23% defective widgets, while the new machine produces only 8% defective widgets. In addition, the new machine produces three times as many widgets as the older machine does. Given a randomly chosen widget was tested and found to be defective, what is the probability it was produced by the new machine?

A 0.08 B 0.15 C 0.489 D 0.511

*Use the following information for the next **two** questions.*

A certain sales company has both male and female employees. These employees either worked overtime (extra hours) or did not. The probability that an employee chosen at random was male was 0.60. The probability that a randomly chosen employee worked overtime was 0.45.

11. What is the probability that an employee chosen at random will be female?

12. The probability that an employee chosen at random is both male AND works overtime is 0.25. What is the probability that a randomly chosen employee is male OR works overtime? Hint: to answer this question it could help to construct a 2x2 contingency table.

*Use the following information for the next **three** questions.*

The marks (pass or fail) of 100 QMET103 students were summarised according to student gender:

	Passed	**Failed**
Male	**20**	**20**
Female	**45**	**15**

13. If a student is selected at random, what is the probability that the student passed QMET103?

14. If a student is selected at random, what is the probability that the student failed QMET103 AND is male?

15. Given that the selected student had passed, what is the probability that the student was male?

16. A local retail store surveyed 1000 people and asked whether they intended to purchase a large television over the next 12 months. Twelve months later, the same respondents were contacted and asked whether they actually purchased the television.

Their responses are summarized in the following table:

Planned to purchase	Actually Purchased	
	Yes	No
Yes	200	50
No	100	650

a) What is the probability that a randomly selected person planned to purchase a large television?

b) What is the probability that a randomly selected person planned to purchase a television **AND** actually purchased a television?

c) What is the probability that a randomly selected person planned to purchase a television **OR** actually purchased a television?

d) Given that a randomly selected person planned to purchase a television, what is the probability that he/she actually purchased a television?

e) Are the two events, planning to purchase a television and actually purchasing a television, statistically independent? (Show working).

17. 300 students were sampled to determine attitudes to internal assessment workloads. Students from both Commerce and Science Divisions were sampled and the following table produced:

	Workload too light	**Workload about right**	**Workload too much**
Science	20	30	50
Commerce	100	20	80

a) What is the probability that a randomly selected person in the sample considers the workload too light?

b) What is the probability that a randomly selected person in the sample considers the workload about right AND too light?

c) What is the probability that a randomly selected person in the sample is a commerce student OR considers the workload too much?

d) Given that a randomly selected student is from the Commerce Division, what is the probability that the student considers the workload about right?

e) What is the probability that a randomly selected student is not a science student AND they think the workload is too light?

18. There are 50 students in the Lincoln University Rugby Club and 20 of them take vitamin C daily. 30% Rugby Club students catch a cold each year. 20% of students who take Vitamin C every day caught a cold last year.

   a) Prepare a contingency table for the above information.

   b) What is the probability that a randomly selected student who does not take Vitamin C every day caught a cold last year?

   c) Given that the randomly selected student caught a cold last year, what is the probability that he takes Vitamin C?

   d) Are taking Vitamin C and catching a cold independent events? Support your answer with appropriate mathematical calculations.

19. A soft drink company is interested in introducing a new Cola brand to the market. Initially they developed three different flavours and want to select the flavour which would be the most popular one. Their research department randomly selected 100 males and 100 females and asked them to choose the best flavour between the three flavours (say A, B and C). The results are summarised in the following table:

Flavour	***Male***	***Female***
A	25	30
B	35	50
C	40	20

   a) What is the probability that a person likes flavour A?

   b) What is the probability that a randomly selected person is a female and likes flavour C?

   c) Given that the randomly selected person is male, what is the probability that he likes flavour C?

   d) What is the probability that a randomly selected person is female or likes the flavour A?

   e) If two persons are randomly selected without replacement, what is the probability that both persons selected will like flavour C?

SOLUTIONS

Questions 1 - 6

1	A	2	A	3	B	4	B
5	C	6	B				

Questions 7-10

1	C	2	C	3	C	4	D

Questions 11-15

1	0.4	2	0.8	3	0.65	4	0.2	5	0.31

Question 16

A	0.25	B	0.2	C	0.35	D	0.8	E	NO

Question 17

a	0.4	b	0	c	0.83	d	0.1	e	0.33

Question 18

	Took Vit C	NO Vit C	Total		
Caught cold	4	11	15	b	0.3667
NO Cold	16	19	35	c	0.2667
Total	20	30	50		

d

$$P(Cold)\times P(vitC) = 0.4\times 0.3 = 0.12;\quad P(cold\ \ and\ \ vit\,C) = 0.08 \neq 0.12$$

$$Hence\ not\ Statistically\ Independent$$

Question 19

a	0.275	b	0.1	c	0.4	d	0.625	e	0.0889