$$P(A \cup B) = P(A) + P(B) - P(A \cap B).$$
If $P(A \cap B) = 0$, A & B are mutually exclusive and $P(A \cup B) = P(A) + P(B)$.

$$P(A \cap B) = P(A)P(B|A).$$
If $P(B|A) = P(B)$, A & B are independent and $P(A \cap B) = P(A)P(B)$.

Tutorial discussion: Q2, Q3, Q4, Q7 and Q8 marked with *

1. (**Multiple choice**) From a set of 8 observations on weights of baby cows, Bigjoh obtained the following: $\sum_{i=1}^{8} x_i = 427$ and $\sum_{i=1}^{8} x_i^2 = 22805$. The mean and variance of this sample (respectively) are (2d.p.)

   (a) 53.38 and 1.19 (b) 53.38 and 1.41 (c) 53.40 and 1.41 (d) 53.38 and 1.98 (e) none of the above.

2. *(**Multiple choice**) A set of 17 observations is given in the following frequency table, where x_i denotes the ith observation and f_i denotes its frequency:

x_i	1	6	8	3	5
f_i	3	1	7	4	2

   The mean of this sample is (2dp):

   (a) 87   (b) 23   (c) 5.12   (d) 2.26   (e) cannot calculate.

3. *Find the variance of the data in Q2.

4. *(**Multiple choice**) A fair six-sided die is rolled twice. Let A be the event such that a total of strictly less than 4 occurs. The probability, $P(A)$ is:

   (a) 1/6   (b) 3/6   (c) 9/36   (d) 1/36   (e) 3/36

5. The letters A, B and C were drawn on cards and placed in a hat. The cards were then selected one at a time without replacement. Find the probability that

   (a) the letters spell CAB   (b) the first letter is B (c) the letter C is last

   (d) the letter A is *not* last.

6. Suppose that the smoking habits of the parents are independent. Given that P(mother is a current smoker)=0.40 and P(father is a current smoker)=0.50, draw a tree diagram for this information. What is the probability that both the father and mother are current smokers?

7. *(**Ref. Q4**) Let B be the event of observing a total divisible by 3.
   (a) Write down the set B.

   (b) Find the conditional probability, $P(A|B)$ without using the formula.

   (c) Find the conditional probability, $P(A|B)$ using $P(A|B) = \dfrac{P(A \cap B)}{P(B)}$.

   Use R to answer these questions.

8. *Each student in a class of 65 was given a packet of 100 seeds and a mildy saline solution. They were required to place the seeds on blotting paper, keep the seeds moistened, and report the number of seeds that had germinated by the end of 14 days. These results are given on the web in germination:

84	88	81	91	84	88	85	90	88	85	87	93	90	88	85	87	90	85	85	89	89	91
85	79	90	90	86	89	92	80	88	84	89	90	86	92	87	89	85	84	82	89	89	89
85	86	86	88	87	83	89	87	87	91	86	91	93	94	88	88	93	87	87	91	89	

(a) Read the data using

    x = scan(file=url("http://www.maths.usyd.edu.au/math1015/r/germination.txt"))

(b) Give the five number summary of these data.

(c) Draw a boxplot for these data. Do you notice of any outliers? Comment on the shape.

(d) Draw a stem-and-leaf plot for this set. Is this shape consistent with the shape in (iii)?

(e) We do not have any prior information about the probability, p, of a seed of this type germination within two weeks. Using the available data, find an estimate $\hat{p}$ of p. **Hint:** mean/100 is a reasonable estimate of p.

(f) Using the available data, find an estimate of the probability that 90 out of 100 seeds germinate for a student. **Hint:** to count the number of students with 90 seeds germinated, the command is `length(x[x==90])`.

(g) Using the available data, find an estimate of the probability that 79 or fewer (out of 100) seeds germinating. **Hint:** to count the number of students with 79 or fewer seeds germinated, the command is `length(x[x<=79])`.

9. The `hospital` data of discharged patients contains the following columns:

```
Column  Label
1       ID no.
2       Duration of hospital stay
3       Age
4       Sex  1=male  2=female
5       First temperature following admission
6       First WBC(x1000) following admission
7       Received antibiotic 1=yes  2=no
8       Received bacterial culture 1=yes  2=no
9       Service 1=med 2=surg.
```

Read the data using

`dat=read.table(file=url("http://www.maths.usyd.edu.au/math1015/r/hospital.txt"),skip=1)`

(a) Find the mean of hospital `stay` among female discharged patients. **Hint:** the command `sex=dat[,4]` set a vector of `sex`. Then the subset of `stay` for female patients is `stay.f=stay[sex==2]`

(b) Use the data to estimate the probability that a discharged patient is a female. **Hint:** use `length(stay.f)` to count the number of female patients.

(c) Condition on female patients, use the data to estimate the probability that a discharged patient stays longer than a week.

---

1. (**Multiple choice**) Consider the following frequency table for x_i values:

x_i	2	4	7	10	12
f_i	1	3	6	8	11

The sample mean is (2dp):

(a) 9.24   (b) 9.28   (c) 53.60   (d) 0.11   (e) 1.99.

2. A six-sided die is loaded in such a way that an even number is twice as likely to occur as an odd number. Find the probability that a number (strictly) less than 4 occurs in a single toss.

3. In Q2, let A be the event that an even number turns up and let B be the event that a number is divisible by 3 occurs. Find $A \cap B$ and $A \cup B$ and deduce that A and B are not mutually exclusive events. Evaluate $P(A \text{ and } B)$ (ie. $P(A \cap B)$) and $P(A \text{ or } B)$ (ie. $P(A \cup B)$) and verify that $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.

4. A "fair" die has two of its faces white, one red and three green. It is thrown three times. Use probability rules to find the probability that

(a) a white face is uppermost at each throw (Hint: The throws are independent)

(b) the same colour is uppermost at each throw (Hint: Express this outcome as the *union* of three mutually exclusive events).