Part IA: Mathematics for Natural Sciences B
Examples Sheet 8: Probability spaces, conditional probability, and combinatorics

Model Solutions

Please send all comments and corrections to jmm232@cam.ac.uk.

Sample spaces and events

1. In an experiment, two fair four-sided dice are rolled. We define:

$$S_1 = \{(i,j) : i \text{ is the result of the first die, } j \text{ is the result of the second die}\},$$
$$S_2 = \{\text{the sum of the results is odd, the sum of the results is even}\},$$
$$S_3 = \{\text{the sum of the results is prime, the first die shows } 1, \text{ the first die shows } 2\}.$$

Which of S_1, S_2, S_3 are valid sample spaces for the experiment?

Solution: Recall that a *sample space* for an experiment is a set that includes all possible outcomes of an experiment. In the above cases:

- S_1 covers all possible pairs of dice rolls, hence covers all outcomes of the experiment. Thus it is a valid sample space.
- S_2 also covers all possible outcomes of the experiment, since the sum of the dice must be odd *or* even. Thus it is also a valid sample space.
- S_3 does not cover all possible outcomes of the experiment, since if the first dice shows 2 and the second dice shows 4, then the sum of the results is 6. This is not prime, so this outcome is not included in the proposed sample space. Hence S_3 is not a valid sample space.

---

2. Given a (discrete) sample space S, define an *event*. Write down in set notation:

(a) a sample space for the result of a 12-sided die roll, the event corresponding to getting a three, the event corresponding to getting an even result, and the event corresponding to getting a prime result;

(b) a sample space for the result of flipping three coins, the event corresponding to getting all tails, the event corresponding to getting an even number of tails, and the event corresponding to getting more heads than tails.

Solution: An *event* is a subset $E \subseteq S$ of the sample space. For the given experiments, we have:

(a) A sample space for the result of a 12-sided die roll is just $S = \{1, 2, 3, 4, 5, 6, ..., 12\}$ corresponding to all possible numbers shown by the dice. The event corresponding to getting a three is $E_3 = \{3\}$. The event corresponding to getting an even result if $E_{\text{even}} = \{2, 4, 6, 8, 10, 12\}$. The event corresponding to getting a prime result is $E_{\text{prime}} = \{2, 3, 5, 7, 11\}$.

(b) A sample space for the result of flipping three coins is just:

$$S = \{(i, j, k) : i, j, k \in \{\text{heads}, \text{tails}\}\}.$$

The event corresponding to getting all tails is $E_{\text{all tails}} = \{(\text{tails}, \text{tails}, \text{tails})\}$. The even corresponding to getting an even number of tails is:

$$E_{\text{even tails}} = \{(\text{tails}, \text{tails}, \text{heads}), (\text{tails}, \text{heads}, \text{tails}), (\text{heads}, \text{tails}, \text{tails}), (\text{heads}, \text{heads}, \text{heads})\}.$$

Finally, the event corresponding to getting more heads than tails is:

$$E_{\text{heads}>\text{tails}} = \{(\text{heads}, \text{heads}, \text{tails}), (\text{heads}, \text{tails}, \text{heads}), (\text{tails}, \text{heads}, \text{heads}), (\text{heads}, \text{heads}, \text{heads})\}.$$

---

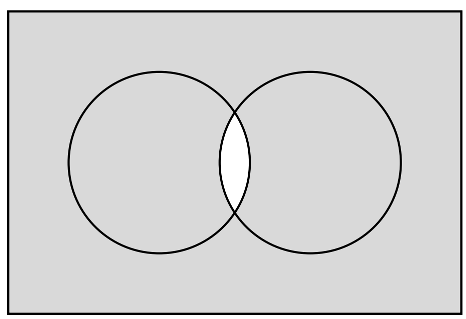
3. Suppose that S is a sample space and A, B, C are events. By drawing appropriate diagrams, show that:

(a) $\overline{(A \cap B)} = \overline{A} \cup \overline{B}$, (b) $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$,

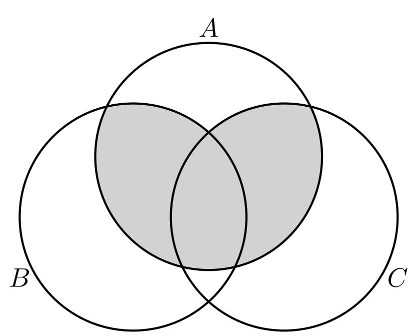
---

➾ Solution:

(a) The diagram below shows the relevant region in grey. The notation $\overline{A \cap B}$ means 'not in both A and B'. The notation $\overline{A} \cup \overline{B}$ means 'not in A or not in B'. These sets evidently coincide:



(b) Again, the diagram below shows the relevant region in grey. The notation $A \cap (B \cup C)$ means 'in A, and in B or C'. The notation $(A \cap B) \cup (A \cap C)$ means 'in A and B, or in A and C'. These evidently coincide.



4. Suppose that S is a sample space and A is an event. Simplify the expressions:

(a) $A \cap S$, (b) $(A \cap \overline{A}) \cup (A \cup \overline{A}) \cup \overline{A}$.

Solution: We have:

(a) $A \cap S$ is the set of all outcomes which are in both the set A and the set S. But all outcomes are in S by definition, hence $A \cap S = A$.

(b) $A \cap \overline{A}$ is the set of all outcomes which are in the set A, and are also outside of the set A. Hence $A \cap \overline{A} = \emptyset$ is the empty set. $A \cup \overline{A}$ is the set of all outcomes which are in the set A or are out of the set A. Hence $A \cup \overline{A} = S$ is the entire sample space. This leaves:

$$\emptyset \cup A \cup \overline{A},$$

which is the set of all outcomes which are either in the empty set (there are none there though!), or in the set A, or out of the set A. This gives the entire sample space as a result, S.

Probability measures

5. Let S be a (discrete) sample space, and let $\mathcal{F}$ be the set of all associated events. What are the three basic *Kolmogorov axioms* that a probability measure $\mathbb{P} : \mathcal{F} \to \mathbb{R}$ must satisfy?

Now suppose that $S = \{\omega_1, \omega_2, \omega_3\}$ is a sample space containing three outcomes.

(a) Write down the set of all possible events associated with this sample space.

(b) Show that if we are given the probabilities $\mathbb{P}(\{\omega_1\}), \mathbb{P}(\{\omega_2\})$, then we may deduce the probabilities of all other events using the basic axioms.

(c) Similarly, show that if we are instead given the probabilities $\mathbb{P}(\{\omega_1, \omega_2\}), \mathbb{P}(\{\omega_1, \omega_3\})$, then we may deduce the probabilities of all other events using the basic axioms.

Solution: The *Kolmogorov axioms* state that a probability measure $\mathbb{P} : \mathcal{F} \to \mathbb{R}$, a function on the space of events which tells us the probability of that event, must satisfy the following:

(K1) For any event $E \subseteq S$, we have $\mathbb{P}(E) \geq 0$.

(K2) The entire sample space has probability 1, $\mathbb{P}(S) = 1$.

(K3) If $E_1, E_2, \ldots$ are a series of disjoint events (that is, they all have zero intersection with one another), we have:

$$\mathbb{P}\left(E_1 \cup E_2 \cup \cdots\right) = \mathbb{P}(E_1) + \mathbb{P}(E_2) + \cdots .$$

From these basic axioms, we can deduce other results. For example:

Complement rule: For any event E, we have $\mathbb{P}(\overline{E}) = 1 - \mathbb{P}(E)$.

Proof: The events $E, \overline{E}$ are disjoint. Hence $\mathbb{P}(E \cup \overline{E}) = \mathbb{P}(E) + \mathbb{P}(\overline{E})$. However, $E \cup \overline{E} = S$, the entire sample space, so $\mathbb{P}(E \cup \overline{E}) = \mathbb{P}(S) = 1$. Thus we have:

$$\mathbb{P}(E) + \mathbb{P}(\overline{E}) = 1 \qquad \Leftrightarrow \qquad \mathbb{P}(\overline{E}) = 1 - \mathbb{P}(E),$$

as required. □

We now consider the sample space $S = \{\omega_1, \omega_2, \omega_3\}$ given in the question.

(a) The set of all events associated with this sample space is:

$$\mathcal{F} = \{\emptyset, \{\omega_1\}, \{\omega_2\}, \{\omega_3\}, \{\omega_1, \omega_2\}, \{\omega_1, \omega_3\}, \{\omega_2, \omega_3\}, \{\omega_1, \omega_2, \omega_3\}\}.$$

(b) Given the probabilities $\mathbb{P}(\{\omega_1\}), \mathbb{P}(\{\omega_2\})$, we can deduce that probability of the event $\{\omega_3\}$:

$$1 = \mathbb{P}(S) = \mathbb{P}(\{\omega_1\} \cup \{\omega_2\} \cup \{\omega_3\}) = \mathbb{P}(\{\omega_1\}) + \mathbb{P}(\{\omega_2\}) + \mathbb{P}(\{\omega_3\}),$$

so we have:

$$\mathbb{P}(\{\omega_3\}) = 1 - \mathbb{P}(\{\omega_1\}) - \mathbb{P}(\{\omega_2\}).$$

This allows us to deduce the probabilities of all other events, since we can write $\{\omega_1, \omega_2\} = \{\omega_1\} \cup \{\omega_2\}$, etc.

(c) Given the probabilities $\mathbb{P}(\{\omega_1, \omega_2\}), \mathbb{P}(\{\omega_1, \omega_3\})$, by the complement rule we can obtain:

$$\mathbb{P}(\{\omega_3\}) = 1 - \mathbb{P}(\{\omega_1, \omega_2\}), \qquad \mathbb{P}(\{\omega_2\}) = 1 - \mathbb{P}(\{\omega_1, \omega_3\}).$$

We can then obtain $\mathbb{P}(\{\omega_1\})$ by using the fact that $1 = \mathbb{P}(\{\omega_1\}) + \mathbb{P}(\{\omega_2\}) + \mathbb{P}(\{\omega_3\})$. We are now back in case (b), and hence can obtain all probabilities, as required.

6.

(a) From the axioms for a probability measure, prove that for any two events A, B (not necessarily exclusive), we have $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$. Generalise this formula to three events A, B, C.

(b) A card is drawn randomly from a standard pack. Using the generalised formula in part (a), determine the probability that the card either shows a prime number, is a spade, or is red.

Solution: (a) Observe that we can split the event $A \cup B$ into three disjoint sets as:

$$A \cup B = (A \cap \overline{B}) \cup (A \cap B) \cup (B \cap \overline{A})$$

Hence we have:

$$\mathbb{P}(A \cup B) = \mathbb{P}(A \cap \overline{B}) + \mathbb{P}(A \cap B) + \mathbb{P}(B \cap \overline{A}). \tag{†}$$

Next, observe that we can split the events A, B into two disjoint sets respectively as:

$$A = (A \cap \overline{B}) \cup (A \cap B), \qquad B = (B \cap \overline{A}) \cup (A \cap B).$$

Hence we have:

$$\mathbb{P}(A) = \mathbb{P}(A \cap \overline{B}) + \mathbb{P}(A \cap B), \qquad \mathbb{P}(B) = \mathbb{P}(B \cap \overline{A}) + \mathbb{P}(A \cap B).$$

Substituting for $\mathbb{P}(A \cap \overline{B})$ and $\mathbb{P}(B \cap \overline{A})$ in equation (†), we get the result:

$$\mathbb{P}(A \cup B) = \mathbb{P}(A) - \mathbb{P}(A \cap B) + \mathbb{P}(A \cap B) + \mathbb{P}(B) - \mathbb{P}(A \cap B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B).$$

The obvious generalisation to three events is:

$$\mathbb{P}(A \cup B \cup C) = \mathbb{P}(A) + \mathbb{P}(B) + \mathbb{P}(C) - \mathbb{P}(A \cap B) - \mathbb{P}(A \cap C) - \mathbb{P}(B \cap C) + \mathbb{P}(A \cap B \cap C).$$

(b) Let S be the sample space comprising the possible cards drawn from the pack, so that the size of the sample space is 52. Let E_{prime} be the event corresponding to drawing a prime number; the prime cards in each suit are $2, 3, 5, 7$, so we have $\mathbb{P}(E_{\text{prime}}) = 4/13$. Let E_{spade} be the event corresponding to drawing a spade; the spades are one of four suits, hence $\mathbb{P}(E_{\text{spade}}) = 1/4$. Let E_{red} be the event corresponding to drawing a red card; there are equal numbers of red and black cards in the deck, hence $E_{\text{red}} = 1/2$.

Next, there are four prime cards which are also spades, which gives $\mathbb{P}(E_{\text{prime}} \cap E_{\text{spade}}) = 4/52 = 1/13$. There are eight prime cards which are also red, which gives $\mathbb{P}(E_{\text{prime}} \cap E_{\text{red}}) = 8/52 = 2/13$. There are no spades which are red, because spades are a black suit, hence $\mathbb{P}(E_{\text{spade}} \cap E_{\text{red}}) = 0$. Finally, there are no cards which are all three of spades, red and prime, because spades are a black suit. Hence $\mathbb{P}(E_{\text{spade}} \cap E_{\text{red}} \cap E_{\text{prime}}) = 0$.

Putting all this information together, we have the required probability:

$$\mathbb{P}(E_{\text{prime}} \cup E_{\text{spade}} \cup E_{\text{red}}) = \frac{4}{13} + \frac{1}{4} + \frac{1}{2} - \frac{1}{13} - \frac{2}{13} = \frac{43}{52}.$$

Conditional probability

7. Suppose that S is a (discrete) sample space, $\mathcal{F}$ is the set of all events, and $\mathbb{P} : \mathcal{F} \to \mathbb{R}$ is a probability measure.

   (a) Define the *conditional probability* $\mathbb{P}(B|A)$ of an event B given an event A.

   (b) Show that, for a fixed A, the conditional probability function $\mathbb{P}(\cdot|A) : \mathcal{F} \to \mathbb{R}$ satisfies the three basic axioms for a probability measure.

   (c) State the definition for two events A, B being *independent* under a probability measure, and explain why this definition makes sense using the definition of conditional probability.

➻ **Solution:** (a) The conditional probability is defined by:

$$\mathbb{P}(B|A) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(A)}.$$

This tells us the probability of B, *given that A has already occurred*. This makes sense, because the probability of B should now be shrunk down to the probability of $B \cap A$ (since we know that A has already happened, so any outcomes in the original B that were not contained in A are no longer allowed). We also need to normalise to the size of the new effective sample space, which becomes A itself, hence the division by $\mathbb{P}(A)$.

(b) If we fix A, and consider the probability measure $\mathbb{P}(B|A)$ for events B, we can prove it satisfies each of Kolmogorov's axioms. We have:

(K1) $\mathbb{P}(B|A) = \mathbb{P}(A \cap B)/\mathbb{P}(A) \geq 0$ for all events B, since $\mathbb{P}(A \cap B), \mathbb{P}(A)$ are both non-negative (and we assume that $\mathbb{P}(A) > 0$).

(K2) $\mathbb{P}(S|A) = \mathbb{P}(S \cap A)/\mathbb{P}(A) = \mathbb{P}(A)/\mathbb{P}(A) = 1$, as required.

(K3) Finally, for an events $B_1, B_2, \ldots$ which are all mutually disjoint, we have:

$$\begin{aligned}
\mathbb{P}(B_1 \cup B_2 \cdots |A) &= \frac{\mathbb{P}((B_1 \cup B_2 \cdots) \cap A)}{\mathbb{P}(A)} \\
&= \frac{\mathbb{P}((B_1 \cap A) \cup (B_2 \cap A) \cup \cdots)}{\mathbb{P}(A)} \\
&= \frac{\mathbb{P}(B_1 \cap A)}{\mathbb{P}(A)} + \frac{\mathbb{P}(B_2 \cap A)}{\mathbb{P}(A)} + \cdots \\
&= \mathbb{P}(B_1|A) + \mathbb{P}(B_2|A) + \cdots,
\end{aligned}$$

using the distributive law for intersection over union, which we proved in Question 3(b).

Hence, conditional probability indeed satisfies all the axioms of a usual probability measure.

(c) A, B are said to be *independent events* if $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$. This makes sense, because if A, B are independent, we have:

$$\mathbb{P}(B|A) = \frac{\mathbb{P}(B \cap A)}{\mathbb{P}(A)} = \frac{\mathbb{P}(B)\mathbb{P}(A)}{\mathbb{P}(A)} = \mathbb{P}(B).$$

That is, the event A occurring does not affect the probability of the event B occurring.

8. A box of 100 gaskets contains ten gaskets with type-A defects only, five with type-B defects only, and two with both types of defect. Given that a gasket drawn at random has a type-A defect, what is the probability that it also has a type-B defect?

➾ **Solution:** Let S be the sample space, which is $S = \{\text{no defects, type A only, type B only, both defects}\}$. We are given the probabilities:

$$\mathbb{P}(\{\text{type A only}\}) = \frac{10}{100} = \frac{1}{10}, \qquad \mathbb{P}(\{\text{type B only}\}) = \frac{5}{100} = \frac{1}{20}, \qquad \mathbb{P}(\{\text{both defects}\}) = \frac{2}{100} = \frac{1}{50}.$$

All of these events are mutually disjoint. The probability we wish to calculate is:

$$\begin{aligned}
\mathbb{P}(\{\text{both defects}\}|\{\text{type A only, both defects}\}) &= \frac{\mathbb{P}(\{\text{both defects}\} \cap \{\text{type A only, both defects}\})}{\mathbb{P}(\{\text{type A only, both defects}\})} \\
&= \frac{\mathbb{P}(\{\text{both defects}\})}{1/10 + 1/50} \\
&= \frac{1/50}{6/50} = \frac{1}{6}.
\end{aligned}$$

---

9. Your supervisor has two children, who are either boys or girls. Assuming equal probability of either gender, determine: (a) the probability that at least one child is a boy, given that at least one is a girl; (b) the probability that at least one child is a boy, given that the *younger* child is a girl.

➾ **Solution:** Let B represent a boy and G a girl. Let the pair (i, j) represent a pair of children, with the ith child older than the jth child. Then the sample space for this experiment is:

$$S = \{(B, B), \qquad (B, G), \qquad (G, B), \qquad (G, G)\}.$$

All of these outcomes are equally likely, so:

$$\mathbb{P}(\{(B, B)\}) = \mathbb{P}(\{(B, G)\}) = \mathbb{P}(\{(G, B)\}) = \mathbb{P}(\{(G, G)\}) = \frac{1}{4}.$$

Notice all these events are also mutually disjoint. Beginning the question proper, we have:

(a) We wish to calculate the probability:

$$\mathbb{P}(\{(B, G), (B, B), (G, B)\}|\{(B, G), (G, G), (G, B)\}) = \frac{\mathbb{P}(\{(B, G), (G, B)\})}{\mathbb{P}(\{(B, G), (G, G), (G, B)\})} = \frac{2/4}{3/4} = \frac{2}{3}.$$

(b) Now, we instead wish to calculate the probability:

$$\mathbb{P}(\{(B, G), (B, B), (G, B)\}|\{(B, G), (G, G)\}) = \frac{\mathbb{P}(\{(B, G)\})}{\mathbb{P}(\{(B, G), (G, G)\})} = \frac{1/4}{2/4} = \frac{1}{2}.$$

Bayes' theorem

10. State and prove *Bayes' theorem*. Give an interpretation of each of the terms that arise.

➾ **Solution:** Bayes' theorem states:

$$\mathbb{P}(B|A) = \frac{\mathbb{P}(A|B)\mathbb{P}(B)}{\mathbb{P}(A)}.$$

To prove this, simply use the definition of conditional probability:

$$\mathbb{P}(B|A) = \frac{\mathbb{P}(B \cap A)}{\mathbb{P}(A)} = \frac{(\mathbb{P}(B \cap A)/\mathbb{P}(B)) \cdot \mathbb{P}(B)}{\mathbb{P}(A)} = \frac{\mathbb{P}(A|B)\mathbb{P}(B)}{\mathbb{P}(A)}.$$

To interpret Bayes' theorem, we think about 'updating' our knowledge according to experimental results. On the right hand side, $\mathbb{P}(B)$ represents our 'prior' probability of the event B being observed. If we observe A however, the probability gets updated by: (i) multiplying $\mathbb{P}(B)$ by $\mathbb{P}(A|B)$, the probability of A being observed under the assumption that B occurs (called the *likelihood*); (ii) normalising by the overall probability of A happening (called the *evidence*). Overall, $\mathbb{P}(B|A)$ is the *posterior* probability of B occurring given that A has already occurred.

---

11. You randomly choose a biscuit from one of two seemingly identical jars. Jar A has 10 chocolate biscuits and 30 plain; jar B has 20 chocolate and 20 plain biscuits. Unfortunately, you choose a plain biscuit. What is the probability that you chose from jar A?

➾ **Solution:** The sample space is $S = \{(A, \text{plain}), (A, \text{choc}), (B, \text{plain}), (B, \text{choc})\}$, where the first half of the pair represents the jar we have selected from. Let $A = \{(A, \text{plain}), (A, \text{choc})\}$ be the event of choosing from jar A, and similarly let B be the event of choosing from jar B. Let $P = \{(A, \text{plain}), (B, \text{plain})\}$ be the event of choosing a plain biscuit, and similarly let C be the event of choosing a chocolate biscuit. Observe that:

- The probability of selecting from jar A and jar B is equal, so that:

  $$\mathbb{P}(A) = \mathbb{P}(B) = \frac{1}{2}$$

- Since A, B are mutually disjoint and $A \cup B = S$, the probability of choosing a plain biscuit may be decomposed as:

  $$\mathbb{P}(P) = \mathbb{P}((P \cap A) \cup (P \cap B)) = \mathbb{P}(P \cap A) + \mathbb{P}(P \cap B) = \mathbb{P}(P|A)\mathbb{P}(A) + \mathbb{P}(P|B)\mathbb{P}(B).$$

  The decomposition of this event in this way is extremely useful, and very often used alongside Bayes' theorem (it is sometimes called the *law of total probability*). The probability of choosing a plain biscuit given we have selected A is $3/4$, and the probability of choosing a plain biscuit given we have selected B is $1/2$. Hence we can simplify this probability to:

  $$\mathbb{P}(P) = \frac{3}{4} \cdot \frac{1}{2} + \frac{1}{2} \cdot \frac{1}{2} = \frac{5}{8}.$$

Now by Bayes' theorem, we have:

$$\mathbb{P}(A|P) = \frac{\mathbb{P}(P|A)\mathbb{P}(A)}{\mathbb{P}(P)} = \frac{(3/4) \cdot (1/2)}{5/8} = \frac{3}{5}.$$

12. (**The base rate fallacy**) Suppose that a disease affects one person in a thousand, and that a medical test for the disease accurately classifies 99% of all cases. What is the probability that, in a random screening exercise, a person who tests positively for the disease actually has the disease?

Solution: The sample space is $S = \{(\text{positive, disease}), (\text{positive, no disease}), (\text{negative, disease}), (\text{negative, no disease})\}$, with each outcome representing the outcome of the test and whether the person actually has the disease. Define:

- $D = \{(\text{positive, disease}), (\text{negative, disease})\}$ to be the event of having the disease. We are given that $\mathbb{P}(D) = 1/1000$, and hence $\mathbb{P}(\overline{D}) = 999/1000$.
- $P = \{(\text{positive, disease}), (\text{positive, no disease})\}$ to be the event of testing positive. We are given that the test accurately classifies 99% of all cases, which tells us that if you have the disease, then the probability of testing positive is $99/100$. That is, $\mathbb{P}(P|D) = 99/100$. It also tells us that if you don't have the disease, then the probability of testing negative is $1/100$. That is, $\mathbb{P}(P|\overline{D}) = 1/100$.

Observe also that $D, \overline{D}$ are disjoint, and $D \cup \overline{D} = S$, so we can write:

$$\mathbb{P}(P) = \mathbb{P}((P \cap D) \cup (P \cap \overline{D})) = \mathbb{P}(P \cap D) + \mathbb{P}(P \cap \overline{D}) = \mathbb{P}(P|D)\mathbb{P}(D) + \mathbb{P}(P|\overline{D})\mathbb{P}(\overline{D}).$$

Substituting the known values from above, we have that the overall probability of testing positive is:

$$\mathbb{P}(P) = \frac{99}{100} \cdot \frac{1}{1000} + \frac{1}{100} \cdot \frac{999}{1000} = \frac{1098}{100000}.$$

We wish to calculate $\mathbb{P}(D|P)$, the probability of having the disease if you tested positive. We apply Bayes' theorem:

$$\mathbb{P}(D|P) = \frac{\mathbb{P}(P|D)\mathbb{P}(D)}{\mathbb{P}(P)} = \frac{(99/100) \cdot (1/1000)}{1098/100000} = \frac{99}{1098} = \frac{11}{122}.$$

Combinatorics

13.

(a) How many ways are there to order n distinct objects?

(b) How many ways are there to order r objects from a set of n distinct objects?

(c) How many ways are there to choose a subset of r objects from a set of n distinct objects?

(d) How many ways are there to arrange n identical objects into r groups?

Solution: Most of these were covered in the lectures.

(a) There are $n!$ ways of ordering n distinct objects. This is because there are n ways to pick the first object, $n-1$ ways to pick the second, $n-2$ ways to pick the third, etc. This gives a total of $n \cdot (n-1) \cdot (n-2) \cdot \ldots = n!$ ways.

(b) There are n ways of choosing the first object in our list, then $n-1$ of choosing the second, etc, down to $n-(r-1)$ ways of choosing the rth object in our list. In total, that gives $n(n-1)\ldots(n-(r-1))$ ways of choosing an ordering of the objects. Hence there are a total of:

$$n(n-1)\ldots(n-(r-1)) = \frac{n!}{(n-r)!}.$$

(c) We now want to take the answer in (b), and ignore ordering. There are $r!$ possible orderings of our list, each of which we must consider equivalent, hence we divide by this factor in (b). The result is:

$$\frac{n!}{(n-r)!r!} = \binom{n}{r},$$

the binomial coefficient.

(d) This is a fun part, based on a method called 'stars and bars'. Suppose that we represent the n identical objects by a list of n *stars*:

$$\underbrace{\star\star\star\ldots\star}_{n \text{ times}}.$$

To partition n up into r sets, we insert $r-1$ *bars*, for example for $n=4, r=3$ we could have:

$$\star\star\,|\,\star\,|\star, \qquad \text{or} \qquad \star\star\star\star||$$

In general, we will have n stars and $r-1$ bars, and any arrangement of these will give rise to a different grouping. Hence there are a total of:

$$\binom{n+r-1}{n}$$

ways.

14. Prove the following property of the binomial coefficients:

$$\binom{n+1}{r+1} = \binom{n}{r+1} + \binom{n}{r},$$

using: (a) the expression for the binomial coefficients in terms of factorials; (b) the combinatorial interpretation of the binomial coefficients in terms of combinations. Explain how this property of binomial coefficients relates to *Pascal's triangle*.

➾ **Solution:** We perform the proof in both of the required ways:

(a) In terms of factorials, we have:

$$\begin{aligned}\binom{n}{r+1} + \binom{n}{r} &= \frac{n!}{(r+1)!(n-r-1)!} + \frac{n!}{r!(n-r)!} \\ &= \frac{n!}{r!(n-r-1)!}\left(\frac{1}{r+1} + \frac{1}{n-r}\right) \\ &= \frac{n!}{r!(n-r-1)!}\left(\frac{n+1}{(r+1)(n-r)}\right) \\ &= \frac{(n+1)!}{(r+1)!((n+1)-(r+1))!} \\ &= \binom{n+1}{r+1},\end{aligned}$$

as required.

(b) We know that $\binom{n+1}{r+1}$ is the number of ways to choose a subset of $r+1$ objects from a set of $n+1$ distinct objects. We can perform this choice by splitting the set in a different way though: imagine we choose one special element from the set of $n+1$ objects. Then, we can *either* choose all of our $r+1$ objects from the remaining n; there are $\binom{n}{r+1}$ ways of doing this. *Or, alternatively,* we can choose our special element and then choose the remaining r objects from the remaining n objects; there are $\binom{n}{r}$ ways of doing this. This establishes the identity:

$$\binom{n+1}{r+1} = \binom{n}{r+1} + \binom{n}{r}$$

in a purely combinatorial (i.e. counting) way!

This relates to Pascal's triangle, because it tells us that the $(r+1)$th entry of the $(n+1)$th row is equal to the sum of the $(r+1)$th entry and the rth entry in the nth row - exactly the normal way that we construct the binomial coefficients.

15. In one of the National Lottery games, six balls are drawn at random from 49 balls, numbered from one to 49. You pick six different numbers.

(a) What is the probability that your six numbers match those drawn?

(b) What is the probability that exactly r of the numbers you choose match those drawn?

(c) What is the probability that five numbers of those you choose match those drawn and that your sixth number matches a 'bonus ball' drawn from those remaining after the first six balls are drawn?

Solution:

(a) The number of ways to choose 6 objects from 49 is:

$$\binom{49}{6}.$$

All of these ways are equally likely, which implies that the probability is the reciprocal of this number.

(b) Of the six balls that are drawn, r of our balls must match - there are $\binom{6}{r}$ ways of doing this. The number of ways of filling up the remaining numbers drawn is the number of ways of choosing $6-r$ objects from $49-r$, which is:

$$\binom{6}{r}\cdot\binom{49-r}{6-r}.$$

Hence, the probability is:

$$\binom{6}{r}\cdot\binom{49-r}{6-r}\Bigg/\binom{49}{6}.$$

(c) First of all, from above there are $6\cdot(49-5)=6\cdot 43$ ways in which the first five balls drawn match ours. There is a $1/43$ chance that a further ball drawn from the remaining ones will match our sixth. Hence the total probability is:

$$6\Bigg/\binom{49}{6}.$$

---

16. Suppose that n distinguishable particles are placed randomly into N different states. A particular configuration of this system is such that there are n_s particles in state s, where $1\le s\le N$. If the ordering of particles in any particular state does not matter, show that the number of ways of realising a particular configurations is:

$$\frac{n!}{n_1!n_2!...n_N!}.$$

Solution: There are $n!$ arrangements of the particles in total. Amongst the first state, there are $n_1!$ orderings, so ignoring possible reorderings, there are now $n!/n_1!$ arrangements. Continuing in this fashion gives this result.

17. Letters $A, B, C, D, E,$ and F are written in a random order, but without repetition, into places $1, 2, 3, 4, 5$ and 6. Explaining your reasoning in each case, how many distinct orderings:

(a) exist in total?

(b) have F in the sixth place?

(c) have E or F in the sixth place?

(d) have E in the fifth place, and F in the sixth place?

(e) have E in the fifth place, or F in the sixth place?

(f) have E in the fifth place, or F in the sixth place, but *not both*?

Now, the letters A, B, C, D, E and F are instead partitioned into two bins, where order does not matter in a given bin. We say that a partition is of *type* $[a, b]$, if a letters are placed into the first bin, and b letters are placed into the second bin.

(g) How many partitions of type $[4, 2]$ are there?

(h) Assuming that from all the possible orderings given enumerated in part (a), the letters in the first four places are placed into the first bin, and the letters in the final two places are placed into the second bin, how many times do A, B, C, D end up in the first bin overall?

(i) Calculate the product of your answers to the two previous parts, and explain the value you obtain.

(j) Repeat the calculation of parts (g)-(i) for each of the possible types of partitions.

Solution:

(a) There are 6, letters which gives $6! = 720$ different permutations.

(b) Fixing F in the sixth place, there are five remaining letters, which gives $5! = 120$ permutations.

(c) There are $5!$ permutations with F in the sixth place, and $5! = 120$ permutations with E in the sixth place, hence there are $2 \cdot 5! = 240$ in total.

(d) If we demand both E in the fifth place, and F in the sixth place, there are four letters remaining with $4! = 24$ possible permutations.

(e) This is given by the number of ways of having E in the fifth place, plus the number of ways of having F in the sixth place, minus the number of ways of having E in the fifth place *and* F in the sixth place, to avoid double counting (this reminds us of the identity $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$). We have $2 \cdot 5! - 4! = 240 - 24 = 216$.

(f) Assuming that we don't have both in the fifth and sixth place, if E is in the fifth place, then there are 4 remaining choices for the sixth slot, and then $4!$ arrangements of the other four slots. This gives $4 \cdot 4! = 48$. Similarly for F, giving $48 + 48 = 96$ in total.

(g) We choose 4 letters from 6, which gives $\binom{6}{4} = 15$.

(h) This question asks for the number of arrangements of A, B, C, D amongst the first four slots, and the number of arrangements of E, F amongst the final two slots. This gives $4! \cdot 2! = 48$.

(i) The product is $\binom{6}{4} \cdot 4! \cdot 2! = 6!$, which is the total number of arrangements. This must be the case because one way of counting the arrangements is to say: 'count the number of ways to split the letters into a group of four and a group of two, then count the number of ways that the group of four can be permuted and the group of two can be permuted'.

(j) The possible types of partition are $[6,0]$, $[5,1]$, $[4,2]$, $[3,3]$, and the remaining cases are symmetric with these. We have:

- The number of arrangements of type $[6,0]$ is $\binom{6}{0} = 1$. The number of arrangements that result in A, B, C, D, E, F falling into the first bin is $6!$. The product is evidently $6!$.
- The number of arrangements of type $[5,1]$ is $\binom{6}{1} = 6$. The number of arrangements that result in A, B, C, D, E falling into the first bin is $5!$. The product is $6!$.
- We have already done $[4,2]$ above.
- The number of arrangements of type $[3,3]$ is $\binom{6}{3} = 20$. The number of arrangements that result in A, B, C falling into the first bin is $3! \cdot 3! = 36$. The product is $6!$.

Miscellaneous probability space problems

18. A box contains $N_B \geq 2$ blue balls and $N - N_B \geq 2$ non-blue balls. An experiment consists of three consecutive stages: drawing a ball from a box, returning it or not returning it, then drawing a second ball from the box. The event B_i represents a blue ball being drawn on the ith draw, for $i = 1, 2$. The event R represents returning a ball on the second stage of the experiment. The probability of event R is $\mathbb{P}(R) = r$.

(a) Write down the sample space of the experiment, and find the probabilities of all of the possible outcomes.

(b) Hence, find:

  (i) $\mathbb{P}(B_2)$;
  (ii) $\mathbb{P}(B_1 \cap B_2)$;
  (iii) $\mathbb{P}(R|B_1 \cap B_2)$.

(c) By sketching the graph of $\mathbb{P}(R|\overline{B}_1 \cap B_2)$ as a function of r, show that $\mathbb{P}(R|\overline{B}_1 \cap B_2) \leq r$.

Solution: (a) A valid sample space is:

$$S = \{(c_1, \text{return}, c_2), (c_1, \text{not return}, c_2) : c_1, c_2 = \text{blue, not blue}\}.$$

The required probabilities are:

$$\begin{aligned}
\mathbb{P}(\{(\text{blue, return, blue})\}) &= \frac{N_B}{N} \cdot r \cdot \frac{N_B}{N} = \frac{N_B^2 r}{N^2}, \\
\mathbb{P}(\{(\text{blue, return, not blue})\}) &= \frac{N_B}{N} \cdot r \cdot \frac{N - N_B}{N} = \frac{N_B(N - N_B)r}{N^2}, \\
\mathbb{P}(\{(\text{not blue, return, blue})\}) &= \frac{N - N_B}{N} \cdot r \cdot \frac{N_B}{N} = \frac{N_B(N - N_B)r}{N^2}, \\
\mathbb{P}(\{(\text{not blue, return, not blue})\}) &= \frac{N - N_B}{N} \cdot r \cdot \frac{N - N_B}{N} = \frac{(N - N_B)^2 r}{N^2}, \\
\mathbb{P}(\{(\text{blue, not return, blue})\}) &= \frac{N_B}{N} \cdot (1 - r) \cdot \frac{N_B}{N - 1} = \frac{N_B^2(1 - r)}{N(N - 1)}, \\
\mathbb{P}(\{(\text{blue, not return, not blue})\}) &= \frac{N_B}{N} \cdot (1 - r) \cdot \frac{N - N_B}{N - 1} = \frac{N_B(N - N_B)(1 - r)}{N(N - 1)}, \\
\mathbb{P}(\{(\text{not blue, not return, blue})\}) &= \frac{N - N_B}{N} \cdot (1 - r) \cdot \frac{N_B}{N - 1} = \frac{N_B(N - N_B)(1 - r)}{N(N - 1)}, \\
\mathbb{P}(\{(\text{not blue, not return, not blue})\}) &= \frac{N - N_B}{N} \cdot (1 - r) \cdot \frac{N - N_B}{N - 1} = \frac{(N - N_B)^2(1 - r)}{N(N - 1)}.
\end{aligned}$$

(b) Using the result from part (a), we have:

(i) Since B_2 contains all outcomes which have 'blue' in the third stage of the experiment, we have:

$$\begin{aligned}
\mathbb{P}(B_2) &= \frac{N_B^2 r}{N^2} + \frac{N_B(N - N_B)r}{N^2} + \frac{N_B^2(1 - r)}{N(N - 1)} + \frac{N_B(N - N_B)(1 - r)}{N(N - 1)} \\
&= \frac{N_B r}{N} + \frac{N_B(1 - r)}{N - 1}.
\end{aligned}$$

(ii) The event $B_1 \cap B_2$ corresponds to both the first stage and the third stage of the experiment showing 'blue'. This gives the probability:

$$\mathbb{P}(B_1 \cap B_2) = \frac{N_B^2 r}{N^2} + \frac{N_B^2(1-r)}{N(N-1)}.$$

(iii) The conditional probability $\mathbb{P}(R|B_1 \cap B_2)$ can be written out as:

$$\mathbb{P}(R|B_1 \cap B_2) = \frac{\mathbb{P}(R \cap B_1 \cap B_2)}{\mathbb{P}(B_1 \cap B_2)}.$$

The event $R \cap B_1 \cap B_2$ corresponds to exactly one outcome, when the first stage is 'blue', the second stage is 'return', and the third stage is 'blue'. which has probability $\mathbb{P}(R \cap B_1 \cap B_2) = N_B^2 r/N^2$. Hence the required probability is:

$$\begin{aligned}\mathbb{P}(R|B_1 \cap B_2) &= \frac{N_B^2 r/N^2}{N_B^2 r/N^2 + N_B^2(1-r)/N(N-1)} \\ &= \frac{rN - r}{N - r}.\end{aligned}$$

(c) The conditional probability $\mathbb{P}(R|\overline{B}_1 \cap B_2)$ can be written out as:

$$\mathbb{P}(R|\overline{B}_1 \cap B_2) = \frac{\mathbb{P}(R \cap \overline{B}_1 \cap B_2)}{\mathbb{P}(\overline{B}_1 \cap B_2)},$$

and hence we can compute it similarly to part (b)(iii). We have:

$$\mathbb{P}(\overline{B}_1 \cap B_2) = \frac{N_B(N-N_B)r}{N^2} + \frac{N_B(N-N_B)(1-r)}{N(N-1)}$$

Since $\mathbb{P}(R \cap \overline{B}_1 \cap B_2) = N_B(N-N_B)r/N^2$, we have:

$$\begin{aligned}\mathbb{P}(R|\overline{B}_1 \cap B_2) &= \frac{N_B(N-N_B)r/N^2}{N_B(N-N_B)r/N^2 + N_B(N-N_B)(1-r)/N(N-1)} \\ &= \frac{r/N}{r/N + (1-r)/(N-1)} \\ &= \frac{rN - r}{N - r},\end{aligned}$$

similarly to part (b)(iii).

At $r = 0$, the probability is zero. At $r = 1$, the probability is 1. Computing the derivative of the probability with respect to r, we have:

$$\frac{d}{dr}\left(\frac{rN - r}{N - r}\right) = \frac{(N-1)(N-r) + (rN - r)}{N - r)^2} = \frac{N(N-1)}{(N-r)^2}.$$

This is everywhere positive, so the function is strictly increasing on the interval $[0, 1]$. Note that at $r = 0$, the gradient is $N(N-1)/N^2 < 1$. Note also that:

$$\frac{d^2}{dr^2}\left(\frac{rN - r}{N - r}\right) = \frac{2N(N-1)}{(N-r)^3},$$

so that the second derivative is also positive. This implies that the gradient of the function is also strictly increasing on the interval $[0, 1]$. In particular, the function must start with a gradient shallower than r, then approach the value 1 from below as $r \to 1$, because if it crosses the line r at any point, its gradient would have to decrease to arrive at the value 1.

19. A factory produces good bananas with probability p and bad bananas with probability $1-p$. The bananas are placed on a conveyor belt and inspected by n different workers sequentially. Worker k notices a good banana with probability g_k, and removes it from the conveyor belt in this instance. Worker k notices a bad banana with probability b_k, and also removes it from the conveyor belt in this instance. Assume that $0 < p < 1$ and that $g_k > 0, b_k > 0$ for all $k = 1, ..., n$.

An experiment is conducted where a banana, which may be good or bad, is placed on the conveyor belt for inspection.

(a) Write down the sample space for the experiment.

(b) Let G be the event that a good banana is produced, and let X_k be the event that the banana is removed from the conveyor belt by worker k. Find:

(i) $\mathbb{P}(G \cap X_1)$;

(ii) $\mathbb{P}(X_1)$;

(iii) $\mathbb{P}(\overline{G}|X_2)$;

(iv) $\mathbb{P}(G|X_2 \cup X_3 \cup ... \cup X_n)$;

(v) $\mathbb{P}(X_{k+1} \cup X_{k+2} \cup ... \cup X_n)$;

(vi) $\mathbb{P}(G|X_1 \cup X_2 \cup ... \cup X_k)$.

(c) If $b_1 = b_2 = ... = b_n$, and $n = 98$, find the minimal value of p for which $\mathbb{P}(G|X_1) = 0.99$.

➡ **Solution:** (a) A valid sample space is:

$$S = \{(g,1), (g,2), ..., (g,n), (g, \text{not removed}), (b,1), (b,2)..., (b,n), (b, \text{not removed})\},$$

where the first entry in each pair is whether the banana is good (g) or bad (b), and the second entry is the index of the worker who removes the banana from the conveyor belt (including the case where the banana is not removed at all).

We can write down the probabilities of the outcomes as follows:

- (g,k) corresponds to the first $k-1$ workers not noticing a good banana, but the kth worker correctly identifying it. Thus the probability is:
$$\mathbb{P}(\{(g,k)\}) = p(1-g_1)(1-g_2)...(1-g_{k-1})g_k.$$

- $(g, \text{not removed})$ corresponds to all n workers not noticing a good banana. Thus the probability is:
$$\mathbb{P}(\{(g, \text{not removed})\}) = p(1-g_1)(1-g_2)...(1-g_n).$$

- (b,k) corresponds to the first $k-1$ workers not noticing a bad banana, but the kth worker correctly identifying it. Thus the probability is:
$$\mathbb{P}(\{(g,k)\}) = (1-p)(1-b_1)(1-b_2)...(1-b_{k-1})b_k.$$

- $(b, \text{not removed})$ corresponds to all n workers not noticing a bad banana. Thus the probability is:
$$\mathbb{P}(\{(b, \text{not removed})\}) = (1-p)(1-b_1)(1-b_2)...(1-b_n).$$

(b) For each of the events, we have:

(i) $G \cap X_1$ is the event where the banana is good, and the first worker notices that it is good, and hence removes it from the production line. This is precisely the outcome $(g, 1)$, which we showed above has probability pg_1.

(ii) X_1 is the event where any banana is removed from the production line by the first worker. If it is good, this happens with probability pg_1, but if it is bad, this happens with probability $(1-p)b_1$. The total probability is:

$$\mathbb{P}(X_1) = pg_1 + (1-p)b_1.$$

(iii) Using the definition of conditional probability, we can rewrite the given probability as:

$$\mathbb{P}(\overline{G}|X_2) = \frac{\mathbb{P}(\overline{G} \cap X_2)}{\mathbb{P}(X_2)} = \frac{(1-p)(1-b_1)b_2}{p(1-g_1)g_2 + (1-p)(1-b_1)b_2},$$

since the numerator is the probability of a bad banana being removed by the second worker, and the denominator is the probability of any banana being removed by the second worker.

(iv) Again, using the definition of conditional probability (and the identity from Question 3(b)), we have:

$$\mathbb{P}(G|X_2 \cup X_3 \cup ... \cup X_n) = \frac{\mathbb{P}((G \cap X_2) \cup (G \cap X_3) \cup ... \cup (G \cap X_n))}{\mathbb{P}(X_2 \cup X_3 \cup ... \cup X_n)}$$

$$= \frac{p(1-g_1)\left[g_2 + (1-g_2)g_3 + ... + (1-g_2)...(1-g_{n-1})g_n\right]}{p(1-g_1)\left[g_2 + ... + (1-g_2)...(1-g_{n-1})g_n\right] + (1-p)(1-b_1)\left[b_2 + ... + (1-b_2)...(1-b_{n-1})b_n\right]}.$$

(v) This probability is simply given by:

$$\mathbb{P}(X_{k+1} \cup ... \cup X_n) = p(1-g_1)...(1-g_k)\left[g_{k+1} + (1-g_{k+1})g_{k+2} + ... + (1-g_{k+1})...(1-g_{n-1})g_n\right]$$
$$+ (1-p)(1-b_1)...(1-b_k)\left[b_{k+1} + (1-b_{k+1})b_{k+2} + ... + (1-b_{k+1})...(1-b_{n-1})b_n\right].$$

(vi) Finally, using the definition of conditional probability, we have:

$$\mathbb{P}(G|X_1 \cup X_2 ... \cup X_k) = \frac{\mathbb{P}((G \cap X_1) \cup (G \cap X_2) \cup ... \cup (G \cap X_k))}{\mathbb{P}(X_1 \cup X_2 \cup ... \cup X_k)}$$

$$= \frac{p\left[g_1 + (1-g_1)g_2 + ... + (1-g_1)...(1-g_{k+1})g_k\right]}{p\left[g_1 + (1-g_1)g_2 + ... + (1-g_1)...(1-g_{k+1})g_k\right] + (1-p)\left[b_1 + (1-b_1)b_2 + ... + (1-b_1)...(1-b_{k+1})b_k\right]}.$$

(c) Let $b = b_1 = ... = b_n$. Then the given probability is:

$$\mathbb{P}(G|X_1) = \frac{\mathbb{P}(G \cap X_1)}{\mathbb{P}(X_1)} = \frac{pg_1}{pg_1 + (1-p)b} = 0.99.$$

Rearranging, we have:

$$pg_1 = 99(1-p)b \qquad \Leftrightarrow \qquad p = \frac{99b}{g_1 + 99b} \qquad \Leftrightarrow \qquad p = \frac{99}{g_1/b + 99}.$$

The only thing that constrains these probabilities is that they must sum to one, giving:

$$p\left[g_1 + (1-g_1)g_2 + ... + (1-g_1)(1-g_2)...(1-g_n)\right] + (1-p)\left[b + (1-b)b + ... + (1-b)^{n-1}b + (1-b)^n\right] = 1.$$

We can recognise part of the second term as the sum of a geometric progression, hence reducing this to:

$$p\left[g_1 + (1-g_1)g_2 + ... + (1-g_1)(1-g_2)...(1-g_n)\right] + (1-p) = 1 \qquad \Leftrightarrow \qquad g_1 + (1-g_1)g_2 + ... + (1-g_1)(1-g_2)...(1-g_n) = 1.$$

So it is clear that if we pick b, g_1, then p is still unconstrained. Therefore to minimise p, we must minimise $99/(g_1/b + 99)$ as a function of g_1/b. Taking b arbitrarily small, and g_1 arbitrarily close to 1, we may make p as close to zero as we wish. Thus, there is no minimal value of p.

20. (**Sampling with replacement**) A bag is filled with N white balls and M black balls. Balls are drawn from the bag sequentially without replacement. Let W_i denote the event that the ith ball drawn is white, and let B_i denote the event that the ith ball drawn is black.

(a) Find $\mathbb{P}(W_1), \mathbb{P}(W_2)$, and $\mathbb{P}(W_3)$. Hence, conjecture and prove a general formula for $\mathbb{P}(W_i)$.

(b) What does the event $W_i \cap W_j$ represent? Find $\mathbb{P}(W_1 \cap W_2), \mathbb{P}(W_1 \cap W_3)$ and $\mathbb{P}(W_2 \cap W_3)$. Hence, conjecture and prove a general formula for $\mathbb{P}(W_i \cap W_j)$.

(c) What does the event $W_i \cup W_j$ represent? Find $\mathbb{P}(W_1 \cup W_2), \mathbb{P}(W_1 \cup W_3)$ and $\mathbb{P}(W_2 \cup W_3)$. Hence, conjecture and prove a general formula for $\mathbb{P}(W_i \cup W_j)$.

(d) Using Bayes' theorem, compute $\mathbb{P}(W_2|W_1)$ and $\mathbb{P}(W_3|W_2)$. Hence, conjecture and prove a general formula for $\mathbb{P}(W_{i+1}|W_i)$.

(e) Show that the probability of obtaining exactly $n \leq N$ white balls in a total of x draws is given by:

$$\frac{\binom{N}{n}\binom{M}{x-n}}{\binom{N+M}{x}}.$$

Solution: For this question, it is useful to imagine that all of the balls have been drawn out of the bag. They are drawn in some order, such as $WBBWBW$... where W denotes a white ball and B denotes a black ball. There are N white balls on the list and M black balls on the list, with total length $N+M$. There are a total of:

$$\frac{(N+M)!}{N!M!}$$

different lists, each of which is equally probably.

(a) The probability of $\mathbb{P}(W_i)$ is the probability that the ith position on our list is a white ball. If we fix this, there are:

$$\frac{(N+M-1)!}{(N-1)!M!}$$

ways of arranging the remaining balls of the list. This gives a probability of:

$$\frac{(N+M-1)!}{(N-1)!M!} \cdot \frac{N!M!}{(N+M)!} = \frac{N}{N+M},$$

regardless of the index i.

(b) The event $W_i \cap W_j$ represents the event that both the i, jth balls are white. This corresponds to both the ith and jth entries of our list being white. If we fix those, there are:

$$\frac{(N+M-2)!}{(N-2)!M!}$$

ways of arranging the remaining balls of the list. This gives a probability of:

$$\frac{(N+M-2)!}{(N-2)!M!} \cdot \frac{N!M!}{(N+M)!} = \frac{N(N-1)}{(N+M)(N+M-1)},$$

regardless of the indices i, j (assuming they are distinct, else we are in case (a) again).

(c) The event $W_i \cup W_j$ represents the even that either the ith ball is white or the jth ball is white. We can obtain this probability using the formula $\mathbb{P}(W_i \cup W_j) = \mathbb{P}(W_i) + \mathbb{P}(W_j) - \mathbb{P}(W_i \cap W_j)$, which gives:

$$\frac{2N}{N+M} - \frac{N(N-1)}{(N+M)(N+M-1)} = \frac{2N(N+M-1) - N(N-1)}{(N+M)(N+M-1)} = \frac{N(N+2M-1)}{(N+M)(N+M-1)}.$$

(d) Using the definition of conditional probability, we have:

$$\mathbb{P}(W_i|W_j) = \frac{\mathbb{P}(W_i \cap W_j)}{\mathbb{P}(W_j)} = \frac{N(N-1)}{(N+M)(N+M-1)} \cdot \frac{N+M}{N} = \frac{N-1}{N+M-1}.$$

(e) Out of the first x balls in our list, we need n of them to be white and $x-n$ of them to be black. There are $\binom{N}{x}$ ways of choosing the white balls, and $\binom{M}{n-x}$ ways of choosing the black balls, for a total of $\binom{N}{x} \cdot \binom{M}{n-x}$ ways. There are a total of $\binom{N+M}{x}$ ways of choosing the first x balls. This gives the required probability.