ACCESS TO SCIENCE, ENGINEERING AND AGRICULTURE: MATHEMATICS 2

MATH00040

SEMESTER 2 2018/2019

DR. ANTHONY BROWN

5. Probability

5.1. Introduction to Probability.

You are probably familiar with the elementary concepts concepts of probability. For example, if I were to ask you what is the probability of getting a head when tossing a coin, you would say it is one in two and if I were to ask you what is the probability of getting an even number when rolling a die, then you would say that it is one in two. What we will do in the first section is to introduce some formal definitions to put these sorts of ideas on a firm footing before we more on to more advanced topics.

In probability we consider the results of various types of 'experiment', such as tossing a coin, rolling a die or drawing a card from a deck of cards. Before we do anything concrete, we need to give some definitions.

Definition 5.1.1 (Sample Space)**.** The set of all possible outcomes of an experiment is called the *sample space*, which we will usually denote by S. A particular outcome, that is, an element of S is called a *sample point*.

As usual a couple of examples will make things clearer.

Example 5.1.2.
(1) Suppose that we throw an ordinary six sided die, then the sample space is $S = \{1, 2, 3, 4, 5, 6\}$ and $1, 2, 3, 4, 5$ and 6 are the sample points.
(2) Suppose that we toss two coins, then the sample space is $\{HH, HT, TH, TT\}$ and HH, HT, TH and TT are the sample points. Note that HT and TH are two different sample points.

We also have a name that denotes a subset of the sample space.

Definition 5.1.3 (Event)**.** An event A in the sample space S is a subset of S and so is a set containing possible outcomes. Note that potentially A can be the empty set or can be equal to S.

Here are a couple of examples.

Example 5.1.4. (1) Suppose that we throw an ordinary six sided die, then 'throwing an even number' is the event $A = \{2, 4, 6\}$.
(2) Suppose that we toss two coins, then 'getting at least one head' is the event $\{HH, HT, TH\}$.

Note that if A and B are events then $A \cup B$ is the event that occurs if A occurs or if B occurs or if both A and B occur. Similarly $A \cap B$ is the event that occurs if both A and B occur.

Definition 5.1.5 (Mutually Exclusive)**.** If A and B are events with $A \cap B = \emptyset$, then we say that A and B are *mutually exclusive.*

Again here are a couple of examples.

Example 5.1.6. (1) Suppose that we throw an ordinary six sided die, suppose that A is the event $A = \{1, 3, 5\}$ and suppose that B is the event $B = \{4, 5, 6\}$. Then $A \cup B$ is the event $\{1, 3, 4, 5, 6\}$ and $A \cap B$ is the event $\{5\}$. Since $A \cap B \neq \emptyset$, the events A and B are not mutually exclusive.
(2) Suppose that we toss two coins, suppose that A is the event $\{HH, HT\}$ and suppose that B is the event $\{TH, TT\}$. Then $A \cup B$ is the event $\{HH, HT, TH, TT\}$ (it must happen) and $A \cap B$ is the event $\emptyset$ (it can't happen). Since $A \cap B = \emptyset$, the events A and B are mutually exclusive.

Definition 5.1.7 (Complement)**.** If A is an event in a sample space S, then A^c, the *complement* of A, is the event that consists of all the outcomes not in A.

Note that A and A^c are mutually exclusive and that $S = A \cup A^c$. Here is an example.

Example 5.1.8. Suppose that we throw an ordinary six sided die and suppose that A is the event $A = \{1, 2\}$. Then A^c is the event $A^c = \{3, 4, 5, 6\}$.

In this section we will only be looking at sample spaces containing a finite number of sample points. In such spaces we have the following definition.

Definition 5.1.9 (Finite Probability Space)**.** A sample space S is made into a *finite probability space* by assigning to each sample point a number called its *probability.* If a_i is the ith sample point then $P(a_i)$ denotes the probability of a_i occurring.

Of course we don't have complete freedom in assigning probabilities, they have to agree with the following lemma.

Lemma 5.1.10. *Let $S = \{a_1, a_2, \ldots, a_n\}$ be a finite probability space. Then*

(1) $P(a_i) > 0$ *for all* i.
(2) $P(a_1) + P(a_2) + \cdots + P(a_n) = 1$.

Remark 5.1.11. Note that these conditions imply that $0 < P(a_i) \leqslant 1$ for all sample points a_i.

Here are a couple of examples.

Example 5.1.12. (1) Suppose that we throw an ordinary six sided die, then $P(1) = P(2) = P(3) = P(4) = P(5) = P(6) = \dfrac{1}{6}$.

(2) Suppose that we toss two coins, then $P(HH) = P(HT) = P(TH) = P(TT) = \dfrac{1}{4}$.

We also need to know how to calculate the probability of an event.

Definition 5.1.13 (Probability of an Event)**.** If A is any non-empty event in a finite probability space, then the probability of A is the sum of the probabilities of all the sample points in A. We also let $P(\emptyset) = 0$.

Remark 5.1.14. Note that Lemma 5.1.10 and Definition 5.1.13 together imply that if A is any event in a finite probability space, then $0 \leqslant P(A) \leqslant 1$.

Warning 5.1.15. So if you are ever asked to calculate a probability and you end up with an answer less than 0 or greater than 1, then you must have gone wrong somewhere.

Here are a couple of examples of probabilities of events.

Example 5.1.16. (1) Suppose that we throw an ordinary six sided die and suppose that A is the event $A = \{1, 3, 5\}$. Then

$$P(A) = P(1) + P(3) + P(5) = \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = \frac{1}{2}.$$

(2) Suppose that we toss two coins and suppose that A is the event $\{HH, HT, TT\}$. Then

$$P(A) = P(HH) + P(HT) + P(TT) = \frac{1}{4} + \frac{1}{4} + \frac{1}{4} = \frac{3}{4}.$$

5.2. **Addition and Multiplication Rules.**

In this section, we will look at how to combine probabilities that are known to give other probabilities. The simplest case is adding the probabilities of two events A and B where A and B are mutually exclusive.

Theorem 5.2.1. *Let A and B be mutually exclusive events in a finite probability space. Then*

$$P(A \cup B) = P(A) + P(B).$$

Remark 5.2.2. If we let $B = A^c$ in Theorem 5.2.1, then since $P(A \cup A^c) = 1$, we obtain $P(A^c) = 1 - P(A)$.

Here is an example of the use of Theorem 5.2.1.

Example 5.2.3. Suppose that we select a card from a deck of cards. Calculate the probability of selecting a red card given that the probability of selecting a heart is $\frac{1}{4}$ and that the probability of selecting a diamond is $\frac{1}{4}$.
Let A be the event 'we select a heart' and let B be the event 'we select a diamond'. Then since A and B are mutually exclusive, we can say that the probability of selecting a red card is $P(A \cup B) = P(A) + P(B) = \frac{1}{4} + \frac{1}{4} = \frac{1}{2}$.

The situation is slightly more complicated when the events A and B are not mutually exclusive. However if we look at Figure 1, we see that if we just calculate $P(A) + P(B)$ then we will be counting $P(A \cap B)$ twice, so we have to subtract it at the end. This gives us Theorem 5.2.4.

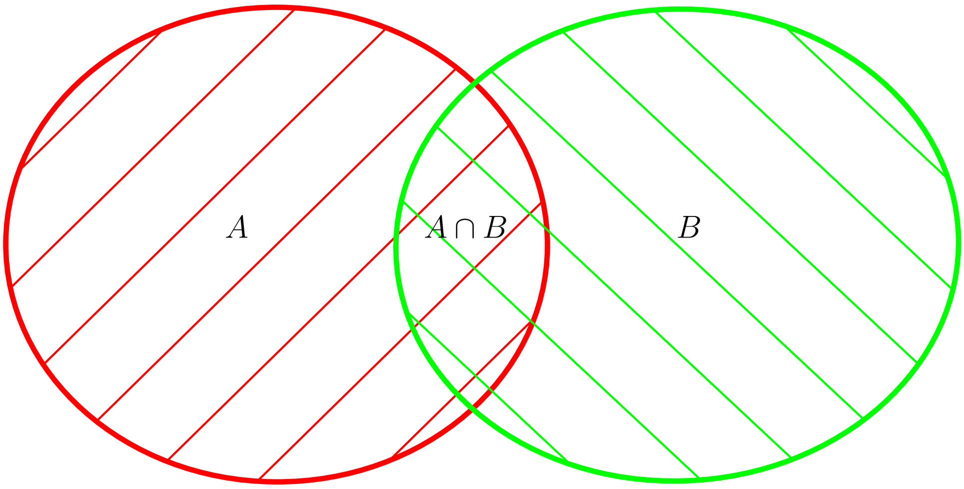


Figure 1. Non Mutually Exclusive Events.

Theorem 5.2.4 (The Addition Rule)**.** *Let A and B be events in a finite probability space. Then*

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

Remark 5.2.5. Although we have only looked at finite probability spaces in this section, Theorem 5.2.4 does in fact hold for all probability spaces.

Theorem 5.2.4 can be used in many more types of situation than Theorem 5.2.1 and can be a very useful tool when trying to calculate probabilities. Here are a couple of examples.

Example 5.2.6. (1) If we pick a card from a deck of cards, what is the probability that it is either a club or a nine?
Let A be the event 'we select a club' and let B be the event 'we select a nine', so that $P(A \cup B)$ is the probability that it is either a club or a nine. Since $P(A \cap B) \neq 0$, the events A and B are not mutually exclusive and so

we have to use Theorem 5.2.4 rather than Theorem 5.2.1. It is easy to see that $P(A) = \frac{1}{4}$, $P(B) = \frac{1}{13}$ and $P(A \cap B) = \frac{1}{52}$. Hence, by Theorem 5.2.4, we have that the probability that it is either a club or a nine is

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) = \frac{1}{4} + \frac{1}{13} - \frac{1}{52} = \frac{4}{13}.$$

(2) In a group of 102 undergraduate students at UCD, there are 54 first years, 49 females and 21 female first years. If one of these students is selected at random, what is the probability that they will either be a first year or female?
Let A be the event 'the student is a first year' and let B be the event 'the student is female', so that $P(A \cup B)$ is the probability that the student is either a first year or female. Again, since $P(A \cap B) \neq 0$, the events A and B are not mutually exclusive and so we have to use Theorem 5.2.4 rather than Theorem 5.2.1. We are given that $P(A) = \frac{54}{102}$, $P(B) = \frac{49}{102}$ and $P(A \cap B) = \frac{21}{102}$. Hence, by Theorem 5.2.4, we have that the probability that the student selected is either a first year or female is

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) = \frac{54}{102} + \frac{49}{102} - \frac{21}{102} = \frac{82}{102}.$$

Now that we have had a look at how to add probabilities, we will turn our attention to how we multiply them. First we need the following definition.

Definition 5.2.7 (Independent Events)**.** Let A and B be events in a finite probability space. If $P(A \cap B) = P(A)P(B)$, then A and B are said to be *independent*.

In one sense this doesn't really help us calculate probabilities since in order to use $P(A \cap B) = P(A)P(B)$, we need to know that the events are independent, but in order to know that two events are independent, then we need to know that $P(A \cap B) = P(A)P(B)$ holds! However, in practice, if it is clear that two events are independent (i.e., the probability of one occurring does not depend on whether the other has occurred) then we can use $P(A \cap B) = P(A)P(B)$. Here is an example.

Example 5.2.8. Suppose that we toss a coin and throw a die. What is the probability of getting a head and a 6?
Let A be the event that the coin comes down a head and let B be the event that we throw a 6, so that the probability of getting a head and a 6 is $P(A \cap B)$. In this case it is intuitively clear that the events A and B are independent, so we can use $P(A \cap B) = P(A)P(B)$. Hence the probability of getting a head and a 6 is $P(A \cap B) = P(A)P(B) = \frac{1}{2} \cdot \frac{1}{6} = \frac{1}{12}$.

As was the case with mutually exclusive events when we were adding probabilities, it is not generally the case that events are independent. Luckily we can also generalise the formula for multiplying probabilities. First we need some notation.

Definition 5.2.9 (Conditional Probability)**.** Let A and B be events in a finite probability space. Then the *conditional probability* that A occurs given that B has occurred is denoted $P(A|B)$.

Here is an example of what this actually means.

Example 5.2.10. Suppose that we draw two cards from a deck of cards (without replacing the first before drawing the second). Let B be the event that we draw an ace on the first draw and let A be the event that we draw an ace on the second draw. The probability $P(A|B)$ is then the probability of drawing an ace on the second draw given that an ace has already been drawn. Since an ace has already been drawn, there are 3 aces left out of 51 cards. Hence $P(A|B) = \frac{3}{51}$.

Now that we have our definition of conditional probability, we can generalise our formula for multiplying probabilities.

Theorem 5.2.11 (Multiplication Rule)**.** *Let A and B be events in a finite probability space. Then*

$$P(A \cap B) = P(A|B)P(B). \tag{1}$$

Remark 5.2.12. If the events A and B are independent, then we also have $P(A \cap B) = P(A)P(B)$. If we combine this equation with (1), it follows that $P(A|B)P(B) = P(A)P(B)$, so that if $P(B) \neq 0$, we have $P(A|B) = P(A)$. Of course $P(A|B) = P(A)$ also holds if $P(B) = 0$, so we always have $P(A|B) = P(A)$ if the events A and B are independent. This corresponds to our intuitive idea of independent events: the fact that B has occurred doesn't affect the probability of A occurring.

Theorem 5.2.11 can be used in two very different ways. Firstly it can be used to calculate probabilities and secondly it can be used to decide if events are independent if this is not intuitively obvious. In this course we will only use it to calculate probabilities: here are a couple of examples.

Example 5.2.13.

(1) A maths lecturer gave his class two tests. Eighty percent of the class passed the first test and sixty percent of the class passed both tests. If we pick a student who passed the first test at random, then what is the probability that they also passed the second test?
Let A be the event that a student passes the second test and let B be the event that a student passes the first test. Note there is nothing special about this order, I have just done it this way so we can use Theorem 5.2.11 directly without swapping the A and the B. Now, the probability we want is $P(A|B)$ and we are given in the question that $P(B) = 0.8$ and that $P(A \cap B) = 0.6$. Hence if we pick a student who passed the first test at random, then the probability that they also passed the second test is

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{0.6}{0.8} = \frac{3}{4}.$$

(2) A bag contains black and white marbles. Two marbles are chosen with the second being chosen without the first being replaced. The probability of selecting a black marble followed by a white marble is 0.36, and the probability of selecting a black marble first is 0.49. What is the probability of selecting a white marble second, given that the first marble drawn was black?
Let A be the event that a white marble is selected second and let B be the event that a black marble is selected first. Then the probability we want is $P(A|B)$ and we are given in the question that $P(A \cap B) = 0.36$ and that $P(B) = 0.49$. Hence the probability of selecting a white marble second, given that the first marble drawn was black is

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{0.36}{0.49} = \frac{36}{49}.$$

5.3. The Binomial Distribution.

For the rest of this chapter, we will look at three different 'probability distributions' which can be used to model various different processes. The first one we will look at is the binomial distribution which allows us to calculate probabilities in a binomial experiment.

Definition 5.3.1 (Binomial Experiment)**.** A *binomial experiment* has the following properties.

(1) The experiment consists of a fixed number of trials.
(2) Each trial has precisely two possible outcomes, which are some times called success and failure.
(3) The probability of success is the same for each trial.

Example 5.3.2. An example of a binomial experiment would be to toss a coin ten times. In this case we could call getting a head success and getting a tail failure. Note there are a fixed number of trials (ten) and the probability of success at each trial is the same (one half).

It is easy to calculate the probabilities associated with a binomial experiment.

Theorem 5.3.3 (Binomial Distribution)**.** *Suppose that we have performed a binomial experiment with n trials, where the probability of each trial yielding a success is p. Then the probability of obtaining k successes (where $0 \leqslant k \leqslant n$) is* $\binom{n}{k} p^k (1-p)^{n-k}$.

Remark 5.3.4. (1) Sometimes we let X denote the number of successes. In this case, we can restate Theorem 5.3.3 as $P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$ and we also write $X \sim B(n, p)$, where the B indicates we are dealing with a binomial distribution. When using this notation, we say X is a random variable.

(2) Another variation is that we sometimes call the probability of a failure q (so that $q = 1 - p$), and then Theorem 5.3.3 becomes $P(X = k) = \binom{n}{k} p^k q^{n-k}$.

Here are a couple examples of using Theorem 5.3.3 to calculate probabilities.

Example 5.3.5. (1) Suppose that we toss a coin ten times. What is the probability of getting seven heads?
Let us call tossing a head a success and let X denote the number of successes we get in ten tosses, so that we want to find $P(X = 7)$. (Note we could also call a success getting a tail and in this case we would have to find $P(X = 3)$, where X is still the number of successes). Since $n = 10$, $k = 7$, $p = \frac{1}{2}$ and $q = 1 - p = \frac{1}{2}$, Theorem 5.3.3 yields

$$P(X = 7) = \binom{10}{7} \left(\frac{1}{2}\right)^7 \left(\frac{1}{2}\right)^{10-7} = \frac{15}{128}.$$

(2) Suppose that we are given a bag containing six white balls and seven black balls and suppose that we draw nine balls from the bag, replacing each ball before drawing the next. What is the probability of drawing five black balls?
Let us call drawing a black ball a success and let X denote the number of successes we get in nine draws, so that we want to find $P(X = 5)$. In this case $n = 9$, $k = 5$, $p = \frac{7}{13}$ and $q = 1 - p = \frac{6}{13}$. Hence Theorem 5.3.3 yields

$$P(X = 5) = \binom{9}{5} \left(\frac{7}{13}\right)^5 \left(\frac{6}{13}\right)^{9-5} \simeq 0.259.$$

It is also easy to calculate the expected value and variance of a binomially distributed random variable.

Theorem 5.3.6 (Expected value and Variance)**.** *Let X be a binomially distributed random variable with parameters n and p. That is let $X \sim B(n, p)$. Then the expected value of X is $E[X] = np$ and the variance of X is $Var[X] = np(1 - p)$.*

Here are a couple of examples.

Example 5.3.7. (1) Suppose that we toss a coin eight times. What is the expected number of heads obtained? What is the variance of the number of heads obtained?
In this case $n = 8$ and $p = \frac{1}{2}$, so the expected number of heads obtained is $np = 4$ (as we would expect intuitively) and the variance of the number of heads obtained is $np(1 - p) = 2$.

(2) Suppose that $X \sim B(25, 0.6)$. What are the expected value and variance of X?
Here $n = 25$ and $p = 0.6$, so $\mathrm{E}[X] = np = 25 \times 0.6 = 15$ and
$\mathrm{Var}[X] = np(1 - p) = 25 \times 0.6(1 - 0.6) = 6$.

Sometimes we might want to know what is called a cumulative probability. For example, we might want to know what is the probability of getting at most seven heads when we toss a coin twelve times, or we might want to know the probability of getting at least 7 heads when tossing a coin eleven times. While we could find these probabilities by summing up all the individual probabilities, this is a lot of work and we would probably make a mistake somewhere. Luckily there are tables of these probabilities available and we will now look at a couple of examples of using these.

Example 5.3.8. Suppose that we toss a coin twelve times. What is the probability of getting at most seven heads?
Let us call tossing a head a success and let X denote the number of successes we get in twelve tosses, so that we want to find $P(X \leqslant 7)$. In this case the probability of a success is $p = 0.5$ and we are tossing a coin twelve times, so $n = 12$. Since we want the probability of getting at most seven heads, we look at the $c = 7$ row and the $p = 0.5$ column in the $n = 12$ block - see Figure 2.

Table: Cumulative Binomial probabilities (*continued*)

		p										
	c	0.05	0.10	0.20	0.30	0.40	0.50	0.60	0.70	0.80	0.90	0.95
n = 12	0	0.540	0.282	0.069	0.014	0.002	0.000	0.000	0.000	0.000	0.000	0.000
	1	0.882	0.659	0.275	0.085	0.020	0.003	0.000	0.000	0.000	0.000	0.000
	2	0.980	0.889	0.558	0.253	0.083	0.019	0.003	0.000	0.000	0.000	0.000
	3	0.998	0.974	0.795	0.493	0.225	0.073	0.015	0.002	0.000	0.000	0.000
	4	1.000	0.996	0.927	0.724	0.438	0.194	0.057	0.009	0.001	0.000	0.000
	5	1.000	0.999	0.981	0.882	0.665	0.387	0.158	0.039	0.004	0.000	0.000
	6	1.000	1.000	0.996	0.961	0.842	0.613	0.335	0.118	0.019	0.001	0.000
	7	1.000	1.000	0.999	0.991	0.943	0.806	0.562	0.276	0.073	0.004	0.000
	8	1.000	1.000	1.000	0.998	0.985	0.927	0.775	0.507	0.205	0.026	0.002
	9	1.000	1.000	1.000	1.000	0.997	0.981	0.917	0.747	0.442	0.111	0.020
	10	1.000	1.000	1.000	1.000	1.000	0.997	0.980	0.915	0.725	0.341	0.118
	11	1.000	1.000	1.000	1.000	1.000	1.000	0.998	0.986	0.931	0.718	0.460
	12	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

FIGURE 2. The probability of getting at most seven heads when tossing a coin twelve times.

Using the table, we see that the required probability is $P(X \leqslant 7) \simeq 0.806$.

There is a slight complication in that the tables only show the probabilities of getting at most c successes, so if we want to get at least d successes, then we first have to transform the question into an appropriate form. The following is an example of this.

Example 5.3.9. Suppose that we are given a bag containing six white balls and nine black balls and suppose that we draw seventeen balls from the bag, replacing each ball before drawing the next. What is the probability of drawing at least ten black balls?

Let us call drawing a black ball a success and let X denote the number of successes we get in seventeen draws, so that we want to find $P(X \geqslant 10)$. In this case we have that the probability of a success is $p = \dfrac{9}{15} = 0.6$ and we are drawing seventeen balls, so $n = 17$. Since the table doesn't directly give us $P(X \geqslant 10)$, we have to use the fact that $P(X \geqslant 10) = 1 - P(X \leqslant 9)$. So we look at the $c = 9$ row and the $p = 0.6$ column in the $n = 17$ block - see Figure 3.

Table: Cumulative Binomial probabilities (*continued*)

		p										
	c	0.05	0.10	0.20	0.30	0.40	0.50	0.60	0.70	0.80	0.90	0.95
n = 17	0	0.418	0.167	0.023	0.002	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	1	0.792	0.482	0.118	0.019	0.002	0.000	0.000	0.000	0.000	0.000	0.000
	2	0.950	0.762	0.310	0.077	0.012	0.001	0.000	0.000	0.000	0.000	0.000
	3	0.991	0.917	0.549	0.202	0.046	0.006	0.000	0.000	0.000	0.000	0.000
	4	0.999	0.978	0.758	0.389	0.126	0.025	0.003	0.000	0.000	0.000	0.000
	5	1.000	0.995	0.894	0.597	0.264	0.072	0.011	0.001	0.000	0.000	0.000
	6	1.000	0.999	0.962	0.775	0.448	0.166	0.035	0.003	0.000	0.000	0.000
	7	1.000	1.000	0.989	0.895	0.641	0.315	0.092	0.013	0.000	0.000	0.000
	8	1.000	1.000	0.997	0.960	0.801	0.500	0.199	0.040	0.003	0.000	0.000
	9	1.000	1.000	1.000	0.987	0.908	0.685	0.359	0.105	0.011	0.000	0.000
	10	1.000	1.000	1.000	0.997	0.965	0.834	0.552	0.225	0.038	0.001	0.000
	11	1.000	1.000	1.000	0.999	0.989	0.928	0.736	0.403	0.106	0.005	0.000
	12	1.000	1.000	1.000	1.000	0.997	0.975	0.874	0.611	0.242	0.022	0.001
	13	1.000	1.000	1.000	1.000	1.000	0.994	0.954	0.798	0.451	0.083	0.009
	14	1.000	1.000	1.000	1.000	1.000	0.999	0.988	0.923	0.690	0.238	0.050
	15	1.000	1.000	1.000	1.000	1.000	1.000	0.998	0.981	0.882	0.518	0.208
	16	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.998	0.977	0.833	0.582
	17	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

FIGURE 3. The probability of getting at most nine black balls when drawing a ball seventeen times.

Using the table, we see that the required probability is

$$P(X \geqslant 10) = 1 - P(X \leqslant 9) \simeq 1 - 0.359 = 0.641.$$

5.4. **The Poisson Distribution.**

The Poisson distribution is used for discrete processes that occur at a known average rate, either in time or in space. It can be used to find the probability of a given number of events occurring during a given length of time or in a particular volume of space. Usually Poisson processes are most useful when the average number of occurrences is 'low'. Here are some examples of the problems that we will be able to solve.

Example 5.4.1. (1) It has been observed that the average number of traffic accidents in Dublin on a weekday morning between 7am and 9am is one. What is the probability that there will be exactly one accident in Dublin during those times next Tuesday?

(2) A particular type of bacterium is randomly distributed in a certain river at an average concentration of one per 50 cm^3 of water. If we draw from the river a test tube containing 25 cm^3 of water, what is the chance that the sample contains exactly two of these bacteria?

(3) The number of births per week in a particular town is five. What is the probability that there will be exactly three births in the next four weeks?

(4) It has been observed that the average rate of customers arriving at a supermarket checkout between 5pm and 6pm on a Friday evening is one per minute. What is the probability of less than four customers arriving between 5.30pm and 5.32pm on Friday?

(5) It has been observed that the average rate of cars arriving at a petrol station between 5pm and 6pm on a Sunday evening is one per ten minutes. What is the probability of at least five cars arriving between 5.30pm and 6pm on Sunday?

We will say that a distribution is a Poisson distribution if it satisfies a certain equation.

Definition 5.4.2 (Poisson Distribution)**.** A random variable X is said to have a *Poisson distribution* with *parameter* λ if

$$P(X = k) = \frac{\lambda^k e^{-\lambda}}{k!}, \text{ where } k = 0, 1, 2, \ldots \tag{2}$$

In practice however, we will proceed by assuming a process is a Poisson process and use (2) to calculate probabilities. So let us answer the first three questions in Example 5.4.1 by assuming that each of the processes are Poisson processes.

Example 5.4.3. (1) It has been observed that the average number of traffic accidents in Dublin on a weekday morning between 7am and 9am is one. What is the probability that there will be exactly one accident in Dublin during those times next Tuesday?
Here we will assume that this is a Poisson process with parameter $\lambda = 1$, where in this case we have taken λ to be the average number of accidents in the two hour period. If we let X be the number of accidents in Dublin between 7am and 9pm next Tuesday, then we have to calculate $P(X = 1)$. In general it should be clear what you should take for λ but this is something you will have to decide for each question. Hence, using (2), we have that the required probability is

$$P(X = 1) = \frac{1^1 e^{-1}}{1!} = e^{-1} \simeq 0.368.$$

Perhaps this is a slightly lower probability than our intuition would suggest.

(2) A particular type of bacterium is randomly distributed in a certain river at an average concentration of one per 50 cm^3 of water. If we draw from the river a test tube containing 25 cm^3 of water, what is the chance that the sample contains exactly two of these bacteria?

Here we will assume that this is a Poisson process with parameter $\lambda = \dfrac{1}{2}$, where in this case we have taken λ to be the average number of bacteria per 25 cm^3 of water. Thus we have to calculate $P(X = 2)$, where X is the number of these bacteria in the test tube containing 25 cm^3 of water. Hence, using (2), we have that the required probability is

$$P(X = 2) = \frac{\left(\frac{1}{2}\right)^2 e^{-\frac{1}{2}}}{2!} = \frac{e^{-\frac{1}{2}}}{8} \simeq 0.076.$$

(3) The number of births per week in a particular town is five. What is the probability that there will be exactly three births in the next four weeks?
Here we will assume that this is a Poisson process with parameter $\lambda = 20$, where in this case we have taken λ to be the average number of births per four weeks. Thus we have to calculate $P(X = 3)$, where X is the number of births in the town in the next four weeks. Hence, using (2), the required probability is

$$P(X = 3) = \frac{(20)^3 e^{-20}}{3!} = \frac{4000e^{-20}}{3} \simeq 0.00000275.$$

This is an extremely low probability as one would expect but perhaps it is even lower than we might expect.

The expected value and variance of X takes a particularly simple form for a Poisson process.

Theorem 5.4.4 (Expected value and Variance)**.** *Let X be a random variable that has a Poisson distribution with parameter λ. Then the expected value of X is $E[X] = \lambda$ and the variance of X is $Var[X] = \lambda$.*

The final two questions in Example 5.4.1 are about cumulative probabilities and, as was the case with the binomial distribution, these are most easily answered using tables (although we could use (2) if we wanted). We will now answer these questions by assuming the processes are Poisson processes.

Example 5.4.5. (1) It has been observed that the average rate of customers arriving at a supermarket checkout between 5pm and 6pm on a Friday evening is one per minute. What is the probability of less than four customers arriving in four minutes beginning at 5.30pm on Friday?
Here we take $\lambda = 4$, where λ is the average number of customers arriving per four minutes between 5pm and 6pm on a Friday evening. Thus we have to calculate $P(X \leqslant 3)$, where X is the number of customers arriving in the given two minutes. Note that we asked for probability of **LESS** than four customers, so this means three or less customers. So we look at the $x = 3$ row in the $\lambda = 4$ column - see Figure 4.
Using the table, we see that the required probability is

$$P(X < 4) = P(X \leqslant 3) \simeq 0.4335.$$

	λ												
x	3.4	3.6	3.8	4	4.2	4.4	4.6	4.8	5	5.2	5.4	5.6	5.8
0	0.0334	0.0273	0.0224	0.0183	0.0150	0.0123	0.0101	0.0082	0.0067	0.0055	0.0045	0.0037	0.0030
1	0.1468	0.1257	0.1074	0.0916	0.0780	0.0663	0.0563	0.0477	0.0404	0.0342	0.0289	0.0244	0.0206
2	0.3397	0.3027	0.2689	0.2381	0.2102	0.1851	0.1626	0.1425	0.1247	0.1088	0.0948	0.0824	0.0715
3	0.5584	0.5152	0.4735	0.4335	0.3954	0.3594	0.3257	0.2942	0.2650	0.2381	0.2133	0.1906	0.1700
4	0.7442	0.7064	0.6678	0.6288	0.5898	0.5512	0.5132	0.4763	0.4405	0.4061	0.3733	0.3422	0.3127
5	0.8705	0.8441	0.8156	0.7851	0.7531	0.7199	0.6858	0.6510	0.6160	0.5809	0.5461	0.5119	0.4783
6	0.9421	0.9267	0.9091	0.8893	0.8675	0.8436	0.8180	0.7908	0.7622	0.7324	0.7017	0.6703	0.6384
7	0.9769	0.9692	0.9599	0.9489	0.9361	0.9214	0.9049	0.8867	0.8666	0.8449	0.8217	0.7970	0.7710

FIGURE 4. The probability of less than or equal to three customers arriving at the checkout during a period of two minutes.

(2) It has been observed that the average rate of cars arriving at a petrol station between 5pm and 6pm on a Sunday evening is one per five minutes. What is the probability of at least five cars arriving between 5.30pm and 6pm on Sunday?

Here we take $\lambda = 6$, where λ is the average number of cars arriving in the thirty minutes between 5.30pm and 6pm on Sunday. Thus we have to calculate $P(X \geqslant 5)$, where X is the number of customers arriving in the given thirty minutes. As was the case with the binomial distribution, there is a slight complication since the tables only show $P(X \leqslant x)$. However, this is not a great problem since $P(X \geqslant 5) = 1 - P(X \leqslant 4)$ (note that we have $P(X \geqslant 5) = 1 - P(X \leqslant 4)$ **NOT** $P(X \geqslant 5) = 1 - P(X \leqslant 5)$). To find $P(X \leqslant 4)$, we look at the $x = 4$ row in the $\lambda = 6$ column - see Figure 5.

	λ												
x	6	6.2	6.4	6.6	6.8	7	7.2	7.4	7.6	7.8	8	8.5	9
0	0.0025	0.0020	0.0017	0.0014	0.0011	0.0009	0.0007	0.0006	0.0005	0.0004	0.0003	0.0002	0.0001
1	0.0174	0.0146	0.0123	0.0103	0.0087	0.0073	0.0061	0.0051	0.0043	0.0036	0.0030	0.0019	0.0012
2	0.0620	0.0536	0.0463	0.0400	0.0344	0.0296	0.0255	0.0219	0.0188	0.0161	0.0138	0.0093	0.0062
3	0.1512	0.1342	0.1189	0.1052	0.0928	0.0818	0.0719	0.0632	0.0554	0.0485	0.0424	0.0301	0.0212
4	0.2851	0.2592	0.2351	0.2127	0.1920	0.1730	0.1555	0.1395	0.1249	0.1117	0.0996	0.0744	0.0550
5	0.4457	0.4141	0.3837	0.3547	0.3270	0.3007	0.2759	0.2526	0.2307	0.2103	0.1912	0.1496	0.1157
6	0.6063	0.5742	0.5423	0.5108	0.4799	0.4497	0.4204	0.3920	0.3646	0.3384	0.3134	0.2562	0.2068
7	0.7440	0.7160	0.6873	0.6581	0.6285	0.5987	0.5689	0.5393	0.5100	0.4812	0.4530	0.3856	0.3239
8	0.8472	0.8259	0.8033	0.7796	0.7548	0.7291	0.7027	0.6757	0.6482	0.6204	0.5925	0.5231	0.4557
9	0.9161	0.9016	0.8858	0.8686	0.8502	0.8305	0.8096	0.7877	0.7649	0.7411	0.7166	0.6530	0.5874

FIGURE 5. The probability of less than or equal to four cars arriving at the petrol station during a period of thirty minutes.

Using the table, we see that the required probability is

$$P(X \geqslant 5) = 1 - P(X \leqslant 4) \simeq 1 - 0.2851 = 0.7149.$$

5.5. **The Normal Distribution.**

In the final section we will look at the normal distribution (sometimes called the Gaussian distribution). This distribution is extremely important since it crops up in numerous different areas. It is different to the binomial and Poisson distributions since it is a continuous distribution. We won't worry too much about exactly what this means in this course but the main thing to remember is that we can only use this distribution to calculate probabilities of the form $P(a \leqslant X \leqslant b)$ (where we accept a

may be minus infinity and b may be plus infinity) rather than $P(X = c)$. It might seem counter intuitive but all probabilities of the form $P(X = c)$ are zero - this is really just equivalent to the fact that all definite integrals of the form $\int_c^c f(x)\,dx$ are also zero.

In fact there are many different normal distributions, one corresponding to each given mean and standard deviation.

Definition 5.5.1 (Normal Distribution)**.** A random variable X is said to have a *normal distribution* with mean μ and standard deviation σ if we have

$$f(X, \mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(X-\mu)^2}{2\sigma^2}}. \tag{3}$$

I have shown a selection of these in Figure 6. Note that as with all probability distributions, the area between each of the them and the x-axis is one.

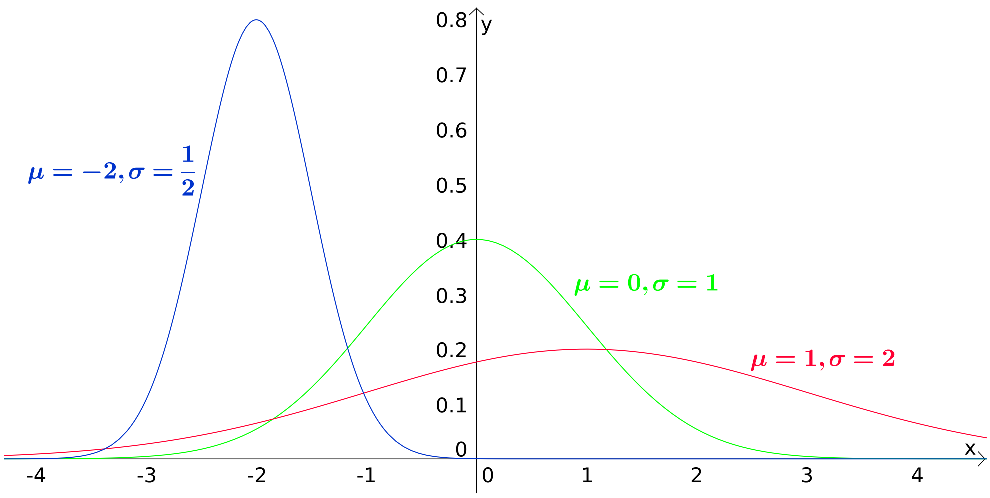


FIGURE 6. Normal distributions.

We can immediately write down the probability of X being between a and b. It is

$$P(a \leqslant X \leqslant b) = \int_a^b \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}.$$

That is it is just the area under the curve between a and b. However this does not get us very far. The problem is that the function f in (3) is not integrable algebraically. All is not lost however. There are tables that we can use, but this causes another problem: we clearly couldn't have a different table for each mean and each standard deviation. The way we get around this is to just have one table for what we call the standard normal distribution.

Definition 5.5.2 (Standard Normal Distribution)**.** The *standard* normal distribution is the one with mean $\mu = 0$ and standard deviation $\sigma = 1$. That is, it is the green curve in Figure 6. So the standard normal distribution is given by

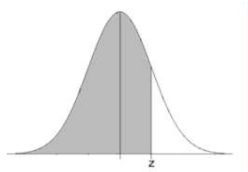
$$f(Z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}.$$

Note that when dealing with the standard normal distribution, we generally use Z rather than X.

The general strategy used to solve problems is to first convert the problem into one involving the standard normal distribution and then solve it using the tables. We will first look at some problems where we don't have to convert the problem and then look at some where we do.

Example 5.5.3. (1) If Z is a random variable with the standard normal distribution, find $P(Z \leqslant 1)$.
First note that since the probability of $Z = 1$ is zero, it doesn't matter if we write $P(Z \leqslant 1)$ or $P(Z < 1)$. This is in complete contrast to the binomial and Poisson distributions, where the difference between $\leqslant 1$ and < 1 is crucial. To find $P(Z \leqslant 1)$ we can just go straight to the tables - see Figure 7.

Standard Normal Cumulative Probability Table



Cumulative probabilities for POSITIVE z-values are shown in the following table:

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830

FIGURE 7. The probability $P(Z \leqslant 1)$.

From the tables, we see that $P(Z \leqslant 1) \simeq 0.8413$.

(2) If Z is a random variable with the standard normal distribution, find $P(Z > -2.14)$.
Since the tables only show $P(Z \leqslant -2.14)$ we have to first find that and then use $P(Z > -2.14) = 1 - P(Z \leqslant -2.14)$.
Figure 8 shows that $P(Z \leqslant -2.14) \simeq 0.0162$. Hence

$$P(Z > -2.14) = 1 - P(Z \leqslant -2.14) \simeq 1 - 0.0162 = 0.9838.$$

Cumulative probabilities for NEGATIVE z-values are shown in the following table:

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
-3.4	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0002
-3.3	0.0005	0.0005	0.0005	0.0004	0.0004	0.0004	0.0004	0.0004	0.0004	0.0003
-3.2	0.0007	0.0007	0.0006	0.0006	0.0006	0.0006	0.0006	0.0005	0.0005	0.0005
-3.1	0.0010	0.0009	0.0009	0.0009	0.0008	0.0008	0.0008	0.0008	0.0007	0.0007
-3.0	0.0013	0.0013	0.0013	0.0012	0.0012	0.0011	0.0011	0.0011	0.0010	0.0010
-2.9	0.0019	0.0018	0.0018	0.0017	0.0016	0.0016	0.0015	0.0015	0.0014	0.0014
-2.8	0.0026	0.0025	0.0024	0.0023	0.0023	0.0022	0.0021	0.0021	0.0020	0.0019
-2.7	0.0035	0.0034	0.0033	0.0032	0.0031	0.0030	0.0029	0.0028	0.0027	0.0026
-2.6	0.0047	0.0045	0.0044	0.0043	0.0041	0.0040	0.0039	0.0038	0.0037	0.0036
-2.5	0.0062	0.0060	0.0059	0.0057	0.0055	0.0054	0.0052	0.0051	0.0049	0.0048
-2.4	0.0082	0.0080	0.0078	0.0075	0.0073	0.0071	0.0069	0.0068	0.0066	0.0064
-2.3	0.0107	0.0104	0.0102	0.0099	0.0096	0.0094	0.0091	0.0089	0.0087	0.0084
-2.2	0.0139	0.0136	0.0132	0.0129	0.0125	0.0122	0.0119	0.0116	0.0113	0.0110
-2.1	0.0179	0.0174	0.0170	0.0166	0.0162	0.0158	0.0154	0.0150	0.0146	0.0143
-2.0	0.0228	0.0222	0.0217	0.0212	0.0207	0.0202	0.0197	0.0192	0.0188	0.0183

FIGURE 8. The probability $P(Z \leqslant -2.14)$.

(3) If Z is a random variable with the standard normal distribution, find $P(-1.51 \leqslant Z \leqslant 1.37)$.
Here we will use the fact that

$$P(-1.51 \leqslant Z \leqslant 1.37) = P(Z \leqslant 1.37) - P(Z \leqslant -1.51).$$

This can be seen from Figure 9, where the area $P(Z \leqslant 1.37)$ has been shaded red and the area $P(Z \leqslant -1.51)$ has been shaded green, so that $P(-1.51 \leqslant Z \leqslant 1.37)$ is the area that is just shaded red.

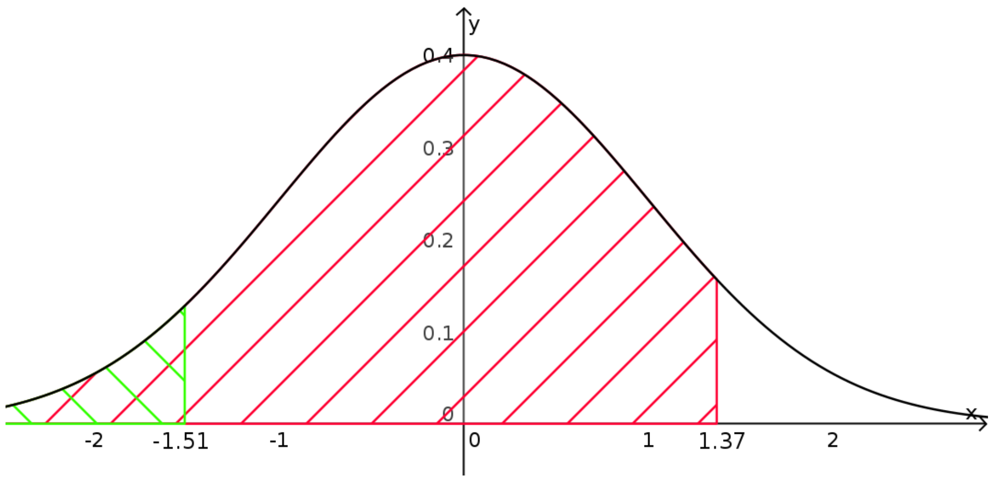


FIGURE 9. The probability $P(-1.51 \leqslant Z \leqslant 1.37)$.

Now, using the tables, we see that $P(Z \leqslant 1.37) \simeq 0.9147$ and $P(Z \leqslant -1.51) \simeq 0.0655$. Hence

$$\begin{aligned} P(-1.51 \leqslant Z \leqslant 1.37) &= P(Z \leqslant 1.37) - P(Z \leqslant -1.51) \\ &\simeq 0.9147 - 0.0655 \\ &= 0.8492. \end{aligned}$$

To finish we will look at some examples where we have to first convert the problem into one using the standard normal distribution. The method is really quite simple and we will state it as a theorem.

Theorem 5.5.4. *If X is a random variable which is normally distributed with mean μ and standard deviation σ then*

$$P(a \leqslant X \leqslant b) = P\left(\frac{a-\mu}{\sigma} \leqslant Z \leqslant \frac{b-\mu}{\sigma}\right),$$

where Z is a random variable which has the standard normal distribution.

Here are some examples to show how this works in practice.

Example 5.5.5. (1) Suppose that the marks in a particular exam are normally distributed with mean 60 and standard deviation 20. What is the probability of a student scoring less than 50?
Suppose that X is a normally distributed random variable with mean 60 and standard deviation 20. Then we want to find $P(X < 50) = P(X \leqslant 50)$. Note that in an exam where the marks are only whole numbers then this will not be true, but we are allowed to assume it since we are told in the question that the marks are normally distributed. Now $\mu = 60$ and $\sigma = 20$, so that $\frac{50-\mu}{\sigma} = \frac{50-60}{20} = -0.5$. Hence, using Theorem 5.5.4, we have that $P(X \leqslant 50) = P(Z \leqslant -0.5)$ and using the tables, we find that the probability of a student scoring less than 50 in the exam is approximately 0.3085.

(2) Suppose that the heights of adult males in Ireland are normally distributed with mean 176cm and standard deviation 5cm. What is the probability of a man chosen at random in Ireland being taller than 183cm?
Suppose that X is a normally distributed random variable with mean 176 and standard deviation 5. Then we want to find $P(X \geqslant 183)$. Now $\mu = 176$ and $\sigma = 5$, so that $\frac{183-\mu}{\sigma} = \frac{183-176}{5} = 1.4$. Hence, using Theorem 5.5.4, we have that $P(X \geqslant 183) = P(Z \geqslant 1.4)$. However $P(Z \geqslant 1.4) = 1 - P(Z \leqslant 1.4)$ and using the tables we have that $1 - P(Z \leqslant 1.4) = 1 - 0.9192 = 0.0808$. Thus the probability of a man chosen at random in Ireland being taller than 183cm is approximately 0.0808.

(3) A particular brand of light bulb is known to have a life normally distributed with mean 1250 hours and standard deviation 150 hours. What is the probability of a randomly selected bulb of this brand lasting between 1300 hours

and 1500 hours?

Suppose that X is a normally distributed random variable with mean 1250 and standard deviation 150. Then we want to find $P(1300 \leqslant X \leqslant 1500)$. Now $\mu = 1250$ and $\sigma = 150$, so that $\dfrac{1300-\mu}{\sigma} = \dfrac{1300-1250}{150} = \dfrac{1}{3}$ and $\dfrac{1500-\mu}{\sigma} = \dfrac{1500-1250}{150} = \dfrac{5}{3}$. Hence, using Theorem 5.5.4, we have that $P(1300 \leqslant X \leqslant 1500) = P\left(\dfrac{1}{3} \leqslant Z \leqslant \dfrac{5}{3}\right)$. But, as in Example 5.5.3 (3),

$$P\left(\frac{1}{3} \leqslant Z \leqslant \frac{5}{3}\right) = P\left(Z \leqslant \frac{5}{3}\right) - P\left(Z \leqslant \frac{1}{3}\right).$$

Using the tables, $P\left(Z \leqslant \dfrac{5}{3}\right) \simeq 0.9522$ and $P\left(Z \leqslant \dfrac{1}{3}\right) \simeq 0.6306$. Hence the probability of a randomly selected bulb of this brand lasting between 1300 hours and 1500 hours is approximately $0.9522 - 0.6306 = 0.3216$.