

Lec 1 Probability Models And Axioms

Mohammed Asir Shahid

2021-05-15

Contents

1	1. Motivation	6
1.1	We need to be able to make decisions under uncertainty . . .	6
2	2. Lecture 1 overview and slides	6
2.1	Downloaded PDF for lecture slides	6
3	3. Sample space	6
3.1	Two Steps:	6
3.1.1	Describe possible outcomes	6
3.1.2	Describe beliefs about likelihood	6
3.2	List set of possible outcomes Ω	6
3.3	The elements out our set should have some properties	6
3.3.1	Mutually exclusive	6
3.3.2	Collectively exhaustive	6
3.3.3	At the right granularity	7
3.4	Ex.	7
3.4.1	Flipping a coin. Our outcomes are heads or tails. . . .	7
4	4. Exercise: Sample space	7
4.1	For the experiment of flipping a coin, and for each one of the following choices, determine whether we have a legitimate sample space:	7
4.2	$\Omega = \{\text{Heads and it is raining, Heads and it is not raining, Tails}\}$	7
4.2.1	Yes	7
4.3	$P(A \cup B) = \{ \text{Heads and it is raining, Tails and it is not raining, Tails} \}$	7
4.3.1	No	7

5	5. Sample space examples	7
5.1	Discrete/finite example	7
5.1.1	Two rolls of a tetrahedral die	7
5.2	Continuous example	8
5.2.1	If we have a dart board, where we are throwing darts with X and Y axis, we can record the coordinates as real numbers with infinite precision. This gives us an infinite amount of possibilities for our sample space, thus we have a continuous sample space.	8
6	6. Exercise: Tree representations	8
6.1	Paul checks the weather forecast. If the forecast is good, Paul will go out for a walk. If the forecast is bad, then Paul will either stay home or go out. If he goes out, he might either remember or forget his umbrella. In the tree diagram below, identify the leaf that corresponds to the event that the forecast is bad and Paul stays home.	8
6.1.1	1 Represents the forecast being good, as the decision making ends there. Paul will definitely go for a walk. .	9
6.1.2	If the forecast is bad, we have choices to make. 2 represents the weather being good and Paul deciding to stay home.	9
6.1.3	If the forecast is bad, Paul could still go out. In which case he could either remember or forget his umbrella which would be 3 and 4.	9
7	7. Probability axioms	9
7.1	Now we can see how to figure out which occurrences are more/less likely to occur.	9
7.2	However, we have an issue when it comes to continuous sample spaces.	9
7.2.1	The probability of hitting any single point on a continuous space is essentially 0.	9
7.3	Event: A subset of the sample space.	9
7.3.1	We will instead assign probability to events.	9
7.4	Axioms:	9
7.4.1	Nonnegativity: $P(A) \geq 0$	9
7.4.2	Normalization: $P(\Omega) = 1$	9
7.4.3	Finite Additivity:	9

7.4.4	These axioms are the only ones that need to be stated. Anything else, such as $P(A) \leq 1$ is already implied from our axioms.	10
8	8. Exercise: Axioms	10
8.1	Let A and B be events on the same sample space, with $P(A)=0.6$ and $P(B)=0.7$. Can these two events be disjoint?	10
8.1.1	No because $P(A) + P(B) > 1$. If they were disjoint, then $P(A) + P(B) = P(A \cup B) = 1.3$ which is greater than 1. This would contradict the normalization axiom	10
9	9. Simple properties of probabilities	10
9.1	Consequences of axioms:	10
9.1.1	(A) $P(A) \geq 0$ implies that $P(A) \leq 1$	10
9.1.2	(B) $P(\Omega) = 1$ implies that $P(\emptyset) = 0$	10
9.1.3	(C) $P(A \cup B) = P(A) + P(B)$ implies that $P(A) + P(A^c) = 1$ and that $P(A \cup B \cup C) = P(A) + P(B) + P(C)$ for k disjoint sets.	10
10	10. Exercise: Simple properties	11
10.1	Let A, B, and C be disjoint subsets of the sample space. For each one of the following statements, determine whether it is true or false. Note: “False” means “not guaranteed to be true.	11
10.1.1	$P(A) + P(A^c) + P(B) = P(A \cup A^c \cup B)$	11
10.1.2	$P(A) + P(B) \leq 1$	11
10.1.3	$P(A^c) + P(B) \leq 1$	11
10.1.4	$P(A \cup B \cup C) \geq P(A \cup B)$	11
11	11. More properties of probabilities	12
11.1	More consequences of axioms:	12
11.1.1	If $A \subset B$, then $P(A) \leq P(B)$	12
11.1.2	$P(A \cup B) = P(A) + P(B) - P(A \cap B)$ even if these sets are not disjoint.	12
11.1.3	$P(A \cup B) \leq P(A) + P(B)$	12
12	12. Exercise: More properties	13
12.1	Let A, B, and C be subsets of the sample space, not necessarily disjoint. For each one of the following statements, determine whether it is true or false. Note: “False” means “not guaranteed to be true.”	13

12.1.1	$P((A \cap B) \cup (C \cap A^c)) \leq P(A \cup B \cup C)$	13
12.1.2	$P(A \cup B \cup C) = P(A \cap C^c) + P(C) + P(B \cap A^c \cap C^c)$	13
13	13. A discrete example	14
13.1	We can move from the abstract to the concrete.	14
13.2	When we have two rolls of a tetrahedral dice, we have 16 possible outcomes, all of which are equally likely.	14
13.3	$P(X = 1) = \frac{1}{4}$	14
13.3.1	Where X is the value of the first roll.	14
13.4	Let $Z = \min(X, Y)$	14
13.4.1	$P(Z = 4) = \frac{1}{16}$	14
13.4.2	$P(Z = 2) = \frac{1}{16} + \frac{2}{16} + \frac{2}{16} = \frac{5}{16}$	14
13.5	This example is something called a Discrete Uniform Law. . .	14
13.5.1	Assume Ω consists of n equally likely elements	14
13.5.2	Assume A consists of k elements	14
13.5.3	$P(A) = k * \frac{1}{n}$	14
14	14. Exercise: Discrete probability calculations	14
14.1	Consider the same model of two rolls of a tetrahedral die, with all 16 outcomes equally likely. Find the probability of the following events:	14
14.1.1	The value in the first roll is strictly larger than the value in the second roll.	14
14.1.2	The sum of the values obtained in the two rolls is an even number.	15
15	15. A continuous example	15
15.1	Probability calculation: Continuous example	15
15.1.1	(x, y) such that $0 \leq x, y \leq 1$	15
15.2	Probability calculation steps	16
15.2.1	Specify the sample space	16
15.2.2	Specify a probability law	16
15.2.3	Identify an event of interest	16
15.2.4	Calculate	16
16	16. Exercise: Continuous probability calculations	16

16.1	Consider a sample space that is the rectangular region $[0,1] \times [0,2]$, i.e., the set of all pairs (x,y) that satisfy $0 \leq x \leq 1$ and $0 \leq y \leq 2$. Consider a “uniform” probability law, under which the probability of an event is half of the area of the event. Find the probability of the following events:	16
16.1.1	The two components x and y have the same values . .	16
16.1.2	The value, x , of the first component is larger than or equal to the value, y , of the second component.	16
16.1.3	The value of x^2 is larger than or equal to the value of y	16
17	17. Countable additivity	17
17.1	Probability calculation: discrete but infinite sample space. . .	17
17.1.1	Sample space: $\{1, 2, 3, \dots\}$	17
18	18. Exercise: Using countable additivity	18
18.1	Let the sample space be the set of positive integers and suppose that $P(n) = 1/2^n$, for $n = 1, 2, \dots$. Find the probability of the set $\{3, 6, 9, \dots\}$, that is, of the set of positive integers that are multiples of 3.	18
18.1.1	For this, we can use countable additivity	18
18.1.2	$\frac{1}{2^3} + \frac{1}{2^6} + \frac{1}{2^9} + \dots = \frac{\frac{1}{2^3}}{1 - (\frac{1}{2^3})} = 1/7$	18
19	19. Exercise: Uniform probabilities on the integers	18
19.1	Let the sample space be the set of all positive integers. Is it possible to have a “uniform” probability law, that is, a probability law that assigns the same probability c to each positive integer?	18
19.1.1	This cannot be the true. We have the following two cases:	18
20	20. Exercise: On countable additivity	19
20.1	Let the sample space be the two-dimensional plane. For any real number x , let A_x be the subset of the plane that consists of all points of the vertical line through the point $(x, 0)$, i.e. $A_x = \{(x, y) : y \in \mathbb{R}\}$	19
20.1.1	Do the axioms of probability theory imply that the probability of the union of the sets A_x (which is the whole plane) is equal to the sum of the probabilities $P(A_x)$	19
20.1.2	Do the axioms of probability theory imply that	19

21	21. Interpretations and uses of probabilities	19
21.1	Narrow view:	19
21.1.1	A branch of Mathematics	19
21.2	Are probabilities frequency?	19
21.2.1	$P(\text{Coin toss yields heads})=1/2$	19
21.2.2	$P(\text{The president of ... will be reelected})=0.7$	19
21.3	Probabilities are often interpreted as:	20
21.3.1	Description of beliefs	20
21.3.2	Betting preferences	20
21.4	The role of probability theory	20
21.4.1	A framework for analyzing phenomena with uncertain outcomes	20
21.5	Real world generates data which can be used by statistics to create models which can then be used by probability theory analysis that can be used to make real world predictions. . . .	20

1 1. Motivation

1.1 We need to be able to make decisions under uncertainty

2 2. Lecture 1 overview and slides

2.1 Downloaded PDF for lecture slides

3 3. Sample space

3.1 Two Steps:

3.1.1 Describe possible outcomes

3.1.2 Describe beliefs about likelihood

3.2 List set of possible outcomes Ω

3.3 The elements out our set should have some properties

3.3.1 Mutually exclusive

1. If one outcome happened, others can not happen

3.3.2 Collectively exhaustive

1. All the elements of the set make up all possibilities

3.3.3 At the right granularity

3.4 Ex.

3.4.1 Flipping a coin. Our outcomes are heads or tails.

1. Let's say we also looked outside while flipping the coin.
 - (a) Either it's heads and rain, heads and no rain, tails and rain, or tails and no rain.
 - i. This is also a valid sample space.
 - ii. However, since we are just curious about whether its heads or tails, this second sample space isn't as relevant.

4 4. Exercise: Sample space

4.1 For the experiment of flipping a coin, and for each one of the following choices, determine whether we have a legitimate sample space:

4.2 $\Omega = \{\text{Heads and it is raining, Heads and it is not raining, Tails}\}$

4.2.1 Yes

4.3 $P(A \cup B) = \{\text{Heads and it is raining, Tails and it is not raining, Tails}\}$

4.3.1 No

5 5. Sample space examples

5.1 Discrete/finite example

5.1.1 Two rolls of a tetrahedral die

1. We have pairs of numbers. One containing the result of the first die and another containing the result of the second die.
 - (a) Result could be written as (1,1) if first die was 1 and second die was 1, (1,2) if first die was 1 and second die was 2, etc.
2. Sequential description-

- (a) An experiment with several stages such as this can be described as a tree <https://i.imgur.com/k6H14NP.png>
- (b) This shows us a distinction in the stages, with 16 different outcomes.

5.2 Continuous example

- 5.2.1 If we have a dart board, where we are throwing darts with X and Y axis, we can record the coordinates as real numbers with infinite precision. This gives us an infinite amount of possibilities for our sample space, thus we have a continuous sample space.

6 6. Exercise: Tree representations

- 6.1 Paul checks the weather forecast. If the forecast is good, Paul will go out for a walk. If the forecast is bad, then Paul will either stay home or go out. If he goes out, he might either remember or forget his umbrella. In the tree diagram below, identify the leaf that corresponds to the event that the forecast is bad and Paul stays home.

https://courses.edx.org/assets/courseware/v1/641ba94a18d06ac5698cf9a413c42bd2/asset-v1:MITx+6.431x+2T2021+type@asset+block/images_texshop_image.jpg

- 6.1.1 1 Represents the forecast being good, as the decision making ends there. Paul will definitely go for a walk.
- 6.1.2 If the forecast is bad, we have choices to make. 2 represents the weather being good and Paul deciding to stay home.
- 6.1.3 If the forecast is bad, Paul could still go out. In which case he could either remember or forget his umbrella which would be 3 and 4.

7 7. Probability axioms

- 7.1 Now we can see how to figure out which occurrences are more/less likely to occur.
- 7.2 However, we have an issue when it comes to continuous sample spaces.
 - 7.2.1 The probability of hitting any single point on a continuous space is essentially 0.
- 7.3 Event: A subset of the sample space.
 - 7.3.1 We will instead assign probability to events.
 - 1. $P(A)$
 - 2. So even though individual events have 0 probability, these events have some probability.

7.4 Axioms:

7.4.1 Nonnegativity: $P(A) \geq 0$

7.4.2 Normalization: $P(\Omega) = 1$

- 1. The probability of our sample space occurring is 1.

7.4.3 Finite Additivity:

If $A \cap B = \emptyset$, then $P(A \cup B) = P(A) + P(B)$

- 1. When the intersection between A and B is the empty set, then they are considered disjoint. Then $P(A \cup B) = P(A) + P(B)$
 - (a) This is due to the fact that there is no overlap in $P(A)$ or $P(B)$.

2. This axiom needs to be refined and strengthened later.

7.4.4 These axioms are the only ones that need to be stated. Anything else, such as $P(A) \leq 1$ is already implied from our axioms.

8 8. Exercise: Axioms

8.1 Let A and B be events on the same sample space, with $P(A)=0.6$ and $P(B)=0.7$. Can these two events be disjoint?

8.1.1 No because $P(A) + P(B) > 1$. If they were disjoint, then $P(A) + P(B) = P(A \cup B) = 1.3$ which is greater than 1. This would contradict the normalization axiom

9 9. Simple properties of probabilities

9.1 Consequences of axioms:

9.1.1 (A) $P(A) \geq 0$ implies that $P(A) \leq 1$

9.1.2 (B) $P(\Omega) = 1$ implies that $P(\emptyset) = 0$

9.1.3 (C) $P(A \cup B) = P(A) + P(B)$ implies that $P(A) + P(A^c) = 1$ and that $P(A \cup B \cup C) = P(A) + P(B) + P(C)$ for k disjoint sets.

1. An element and its complement makes up Ω and the intersection of a set and its complement is \emptyset .

2. From (B), $P(\Omega) = 1$ which is equal to $P(A \cup A^c)$. This is equal to $P(A) + P(A^c)$.

(a) Based on this, we can write that $P(A) = 1 - P(A^c) \leq 1$.

3. $1 = P(\Omega) + P(\Omega^c)$

(a) $1 = 1 + P(\Omega^c) \Rightarrow 1 = 1 + P(\emptyset) \Rightarrow P(\emptyset) = 0$

4. When A, B, C are disjoint, $P(A \cup B \cup C) = P(A) + P(B) + P(C)$.

(a) We can think of the $P(A \cup B \cup C)$ as $P((A \cup B) \cup C) = P(A \cup B) + P(C) = P(A) + P(B) + P(C)$

- i. This logic can be continued for k disjoint sets. Can prove by induction.
 - ii. If (A_1, \dots, A_k) then $P(A_1 \cup \dots \cup A_k) = \sum_{i=1}^k P(A_i)$
5. $P(\{s_1, s_2, \dots, s_k\}) = P(\{s_1\} \cup \dots \cup \{s_k\}) = P(\{s_1\}) + \dots + P(\{s_k\}) = P(s_1) + \dots + P(s_k)$

10 10. Exercise: Simple properties

10.1 Let A, B, and C be disjoint subsets of the sample space. For each one of the following statements, determine whether it is true or false. Note: “False” means “not guaranteed to be true.”

10.1.1 $P(A) + P(A^c) + P(B) = P(A \cup A^c \cup B)$

- 1. This statement is false. While A and the compliment of A are known to be disjoint, we do not know if this also applies to A and B or A^c and B.

10.1.2 $P(A) + P(B) \leq 1$

- 1. This statement is true since A and B are disjoint, $P(A) + P(B) = P(A \cup B)$ which can not be greater than 1.

10.1.3 $P(A^c) + P(B) \leq 1$

- 1. This is false. For example, consider A situation where $A = \emptyset$. Then $A^c = \Omega$. Now given that A and B are disjoint, no matter what B is, $P(\Omega) + P(B) > 1$.

10.1.4 $P(A \cup B \cup C) \geq P(A \cup B)$.

- 1. This must be true. Given that these are disjoint, we know that $P(A \cup B \cup C) = P(A) + P(B) + P(C)$. And we also know that $P(A \cup B) = P(A) + P(B)$. So given that the probability of an event must be nonnegative, we know $P(A \cup B \cup C)$ must be greater than or equal to $P(A \cup B)$.

11 11. More properties of probabilities

11.1 More consequences of axioms:

11.1.1 If $A \subset B$, then $P(A) \leq P(B)$

1. This seems obvious, but how can we prove this?
 - (a) B can be written as $A \cup (B \cap A^c)$.
 - (b) Thus $P(B) = P(A) + P(B \cap A^c) \geq P(A)$
 - (c) So since probabilities are nonnegative, we know that $P(A) \leq P(B)$

11.1.2 $P(A \cup B) = P(A) + P(B) - P(A \cap B)$ even if these sets are not disjoint.

1. The union of A and B can be broken down into 3 distinct parts.
 - (a) $a = P(A \cap B^c)$
 - (b) $b = P(A \cap B)$
 - (c) $c = P(B \cap A^c)$
2. Then $P(A \cup B) = a + b + c$
 - (a) $P(A) + P(B) - P(A \cap B) = (a + b) + (b + c) - b = a + b + c$
 - (b) Thus we can see that this is true.
3. This leads us to the following union bound:

11.1.3 $P(A \cup B) \leq P(A) + P(B)$

1. This is called the union bound and can be used when proving the probability of something is less than the probability of something else.
2. $P(A \cup B \cup C) = P(A) + P(A^c \cap B) + P(A^c \cap B^c \cap C)$
 - (a) We have several pieces:
 - i. $A \cup B \cup C = A \cup (B \cap A^c) \cup (C \cap A^c \cap B^c)$
 - (b) These three pieces are disjoint, thus the additivity axiom gives us:
 - i. $P(A \cup B \cup C) = P(A) + P(B \cap A^c) + P(C \cap A^c \cap B^c)$

12 12. Exercise: More properties

12.1 Let A , B , and C be subsets of the sample space, not necessarily disjoint. For each one of the following statements, determine whether it is true or false. Note: “False” means “not guaranteed to be true.”

12.1.1 $P((A \cap B) \cup (C \cap A^c)) \leq P(A \cup B \cup C)$

1. This statement must be true. This is due to the fact that $A \cap B$ cannot be any bigger than $A \cup B$ and $C \cap A^c$ cannot be any bigger than C . Thus this statement is true.

12.1.2 $P(A \cup B \cup C) = P(A \cap C^c) + P(C) + P(B \cap A^c \cap C^c)$

1. This statement must be true. It is the same property we saw earlier in 11.

13 13. A discrete example

13.1 We can move from the abstract to the concrete.

13.2 When we have two rolls of a tetrahedral dice, we have 16 possible outcomes, all of which are equally likely.

13.3 $P(X = 1) = \frac{1}{4}$

13.3.1 Where X is the value of the first roll.

13.4 Let $Z = \min(X, Y)$

13.4.1 $P(Z = 4) = \frac{1}{16}$

13.4.2 $P(Z = 2) = \frac{1}{16} + \frac{2}{16} + \frac{2}{16} = \frac{5}{16}$

13.5 This example is something called a Discrete Uniform Law.

13.5.1 Assume Ω consists of n equally likely elements

13.5.2 Assume A consists of k elements

13.5.3 $P(A) = k * \frac{1}{n}$

14 14. Exercise: Discrete probability calculations

14.1 Consider the same model of two rolls of a tetrahedral die, with all 16 outcomes equally likely. Find the probability of the following events:

14.1.1 The value in the first roll is strictly larger than the value in the second roll.

1. This event contains the times the first roll is strictly larger than the second. The first will be equal to the second a fourth of the time. The two rolls are not equal $\frac{3}{4}$ of the time, which in this case means 12 of the 16 rolls will not be equal. In half of these occurrences, the first roll will be greater than the second roll, and vice versa in the other half. Thus, we have a $\frac{6}{16}$ probability that the first roll will be strictly greater than the second.

14.1.2 The sum of the values obtained in the two rolls is an even number.

1. This event will happen half the time, since half our numbers are even numbers and half are odd. So the probability will be $\frac{8}{16}$

15 15. A continuous example

15.1 Probability calculation: Continuous example

15.1.1 (x, y) such that $0 \leq x, y \leq 1$

1. We need to specify a uniform probability law
 - (a) Meaning, the probability is equal to the area.
 - (b) $P(\{(x, y) | x + y \leq \frac{1}{2}\})$
 - i. These points form a triangle between the line $x + y = \frac{1}{2}$
 - ii. So the probability is equal to the area of the triangle which is $\frac{1}{2} * \frac{1}{2} * \frac{1}{2} = \frac{1}{8}$
 - (c) $P(\{(0.5, 0.3)\})$
 - i. This event consists of only one element, a single point.
 - ii. The area of a single point is 0, so the probability of having this single point, or any other single point, is simply 0.

15.2 Probability calculation steps

15.2.1 Specify the sample space

15.2.2 Specify a probability law

15.2.3 Identify an event of interest

15.2.4 Calculate

16 16. Exercise: Continuous probability calculations

16.1 Consider a sample space that is the rectangular region $[0,1] \times [0,2]$, i.e., the set of all pairs (x,y) that satisfy $0 \leq x \leq 1$ and $0 \leq y \leq 2$. Consider a “uniform” probability law, under which the probability of an event is half of the area of the event. Find the probability of the following events:

16.1.1 The two components x and y have the same values

1. This event is simply a line, which has no area. Thus the probability of this will be 0.

16.1.2 The value, x , of the first component is larger than or equal to the value, y , of the second component.

1. This would make up a triangle with points at $(0,1)$, $(1,0)$, $(1,1)$. The area of such a triangle is 0.5 and the total area of the region is 2. Thus the probability of the x value being larger than the y value is $\frac{1}{4}$.

16.1.3 The value of x^2 is larger than or equal to the value of y

1. This corresponds to the region below the curve $y = x^2$, with x from 0 to 1. This can be seen as $\int_0^1 x^2 dx = \frac{x^3}{3} \Big|_0^1 = \frac{1}{3}$. Given that the total area is 2, the probability of x^2 being greater than or equal to y is $\frac{1}{6}$.

17 17. Countable additivity

17.1 Probability calculation: discrete but infinite sample space.

17.1.1 Sample space: $\{1, 2, 3, \dots\}$

1. For example, an experiment with a coin tossed and we are looking for the first time a heads occur. The first heads could be the first toss, second, third, etc so the sample space is discrete and infinite.
2. Let us specify a probability law.
 - (a) We have $P(n) = \frac{1}{2^n}, n = 1, 2, 3, \dots$
 - (b) Do these add to 1?
 - i. $\sum_{n=1}^{\infty} \frac{1}{2^n} = \frac{1}{2} \sum_{n=0}^{\infty} \frac{1}{2^n} = \frac{1}{2} * \frac{1}{1-(\frac{1}{2})} = 1$
 - (c) How can we calculate the probability of a general event?
 - i. Probability that the outcome is even is equal to
 - ii. $P(\{2, 4, 6, \dots\}) = P(\{2\} \cup \{4\} \cup \{6\} \cup \dots) = P(2) + P(4) + P(6) + \dots = \frac{1}{2^2} + \frac{1}{2^4} + \frac{1}{2^6} = \frac{1}{4}(1 + \frac{1}{4} + \frac{1}{4^2} + \dots) = \frac{1}{4} \cdot \frac{1}{1-\frac{1}{4}} = \frac{1}{3}$
 - iii. This leads to the Countable Additivity Axiom
 - A. If A_1, A_2, \dots is an infinite sequence of disjoint events, then $P(A_1 \cup A_2 \cup \dots) = P(A_1) + P(A_2) + \dots$
 - B. Having a sequence here is important.
 - C. Otherwise, we could say that a unit square with a union of disjoint single elements, each with a probability of 0, lead to $P(\Omega) = 0$, which cannot be the case.
 - D. Since the disjoint elements in the unit square are not countable and cannot be arranged in a sequence, the Countable Additivity Axiom does not apply.

18 18. Exercise: Using countable additivity

18.1 Let the sample space be the set of positive integers and suppose that $P(n)=1/2^n$, for $n=1,2,\dots$. Find the probability of the set $\{3,6,9,\dots\}$, that is, of the set of positive integers that are multiples of 3.

18.1.1 For this, we can use countable additivity

18.1.2 $\frac{1}{2^3} + \frac{1}{2^6} + \frac{1}{2^9} + \dots = \frac{\frac{1}{2^3}}{1-(\frac{1}{2^3})} = 1/7$

19 19. Exercise: Uniform probabilities on the integers

19.1 Let the sample space be the set of all positive integers. Is it possible to have a “uniform” probability law, that is, a probability law that assigns the same probability c to each positive integer?

19.1.1 This cannot be the true. We have the following two cases:

1. $c=0$

(a) If $c=0$, then we have the following by countable additivity

i. $1 = P(\Omega) = P(\{1\} \cup \{2\} \cup \{3\} \cup \dots) = P(\{1\}) + P(\{2\}) + \dots = 0 + 0 + \dots = 0$

A. Thus we have $1 = 0$ which is a contradiction.

2. $c>0$

(a) If $c>0$, then there is some integer k such that $kc>1$. k could simply be $\frac{1}{c} + 1$. This would lead to the total probability being greater than 1, which cannot be the case.

20 20. Exercise: On countable additivity

20.1 Let the sample space be the two-dimensional plane. For any real number x , let A_x be the subset of the plane that consists of all points of the vertical line through the point $(x, 0)$, i.e. $A_x = \{(x, y) : y \in \mathbb{R}\}$

20.1.1 Do the axioms of probability theory imply that the probability of the union of the sets A_x (which is the whole plane) is equal to the sum of the probabilities $P(A_x)$.

1. No, they do not imply this. This is due to the fact that this collection of sets is not countable, as the set of real numbers is not countable. Thus the additivity axiom does not apply.

20.1.2 Do the axioms of probability theory imply that

$$P(A \cup A_2 \cup \dots) = \sum_{x=1}^{\infty} P(A_x)$$

1. Yes, they do imply this. This is due to the fact that we are dealing with a sequence of disjoint events.

21 21. Interpretations and uses of probabilities

21.1 Narrow view:

21.1.1 A branch of Mathematics

1. Axioms \implies theorems.
2. This can be looked at without looking at what probability really means
3. "Theorem": Frequency of event A is $P(A)$

21.2 Are probabilities frequency?

21.2.1 $P(\text{Coin toss yields heads})=1/2$

1. In this case, it makes sense to think of probability as a frequency

21.2.2 $P(\text{The president of ... will be reelected})=0.7$

1. Thinking of this as a frequency doesn't make as much sense, because there is only a single election.

21.3 Probabilities are often interpreted as:

21.3.1 Description of beliefs

21.3.2 Betting preferences

21.4 The role of probability theory

21.4.1 A framework for analyzing phenomena with uncertain outcomes

1. Rules for consistent reasoning
2. Used for predictions and decisions

21.5 Real world generates data which can be used by statistics to create models which can then be used by probability theory analysis that can be used to make real world predictions.