Phil/LPS 31 Introduction to Inductive Logic
Lecture 11

David Mwakima
dmwakima@uci.edu
Department of Logic and Philosophy of Science
University of California, Irvine

May 10th 2023

Topics

- Recap of Kolmogorov Axioms
- Probability Spaces
- Calculating Probabilities Using the Axioms (or Rules)
- Deriving Other Rules of Probability Theory

Recap of Kolmogorov Axioms

- We defined a field $\mathcal{F}$ as a non-empty collection of sets Ω that is closed under finite applications of set-theoretic operations of taking unions and complements.

Recap of Kolmogorov Axioms

- We defined a field $\mathcal{F}$ as a non-empty collection of sets Ω that is closed under finite applications of set-theoretic operations of taking unions and complements.
- To say that a field $\mathcal{F}$ is closed under finite applications of set theoretic operations means:

Recap of Kolmogorov Axioms

- We defined a field $\mathcal{F}$ as a non-empty collection of sets Ω that is closed under finite applications of set-theoretic operations of taking unions and complements.
- To say that a field $\mathcal{F}$ is closed under finite applications of set theoretic operations means:
  (1) $\Omega \in \mathcal{F}$

Recap of Kolmogorov Axioms

- We defined a field $\mathcal{F}$ as a non-empty collection of sets Ω that is closed under finite applications of set-theoretic operations of taking unions and complements.

- To say that a field $\mathcal{F}$ is closed under finite applications of set theoretic operations means:
  (1) $\Omega \in \mathcal{F}$
  (2) If $A \in \mathcal{F}$, then $A^c \in \mathcal{F}$

Recap of Kolmogorov Axioms

- We defined a field $\mathcal{F}$ as a non-empty collection of sets Ω that is closed under finite applications of set-theoretic operations of taking unions and complements.

- To say that a field $\mathcal{F}$ is closed under finite applications of set theoretic operations means:
  (1) $\Omega \in \mathcal{F}$
  (2) If $A \in \mathcal{F}$, then $A^c \in \mathcal{F}$
  (3) If $A, B \in \mathcal{F}$, then $A \cup B \in \mathcal{F}$

Recap of Kolmogorov Axioms

- We defined a field $\mathcal{F}$ as a non-empty collection of sets Ω that is closed under finite applications of set-theoretic operations of taking unions and complements.
- To say that a field $\mathcal{F}$ is closed under finite applications of set theoretic operations means:
  (1) $\Omega \in \mathcal{F}$
  (2) If $A \in \mathcal{F}$, then $A^c \in \mathcal{F}$
  (3) If $A, B \in \mathcal{F}$, then $A \cup B \in \mathcal{F}$
- The Kolmogorov Axioms characterize what a probability function P defined on a field is as follows:

Recap of Kolmogorov Axioms

- We defined a field $\mathcal{F}$ as a non-empty collection of sets Ω that is closed under finite applications of set-theoretic operations of taking unions and complements.
- To say that a field $\mathcal{F}$ is closed under finite applications of set theoretic operations means:
  (1) $\Omega \in \mathcal{F}$
  (2) If $A \in \mathcal{F}$, then $A^c \in \mathcal{F}$
  (3) If $A, B \in \mathcal{F}$, then $A \cup B \in \mathcal{F}$
- The Kolmogorov Axioms characterize what a probability function P defined on a field is as follows:
  - $P(\Omega) = 1$ (Normalized)

Recap of Kolmogorov Axioms

- We defined a field $\mathcal{F}$ as a non-empty collection of sets Ω that is closed under finite applications of set-theoretic operations of taking unions and complements.
- To say that a field $\mathcal{F}$ is closed under finite applications of set theoretic operations means:
  (1) $\Omega \in \mathcal{F}$
  (2) If $A \in \mathcal{F}$, then $A^c \in \mathcal{F}$
  (3) If $A, B \in \mathcal{F}$, then $A \cup B \in \mathcal{F}$
- The Kolmogorov Axioms characterize what a probability function P defined on a field is as follows:
  - $P(\Omega) = 1$ (Normalized)
  - $0 \leq P(A) \leq 1$ for $A \in \mathcal{F}$ (Non-negative)

Recap of Kolmogorov Axioms

- We defined a field $\mathcal{F}$ as a non-empty collection of sets Ω that is closed under finite applications of set-theoretic operations of taking unions and complements.
- To say that a field $\mathcal{F}$ is closed under finite applications of set theoretic operations means:
  (1) $\Omega \in \mathcal{F}$
  (2) If $A \in \mathcal{F}$, then $A^c \in \mathcal{F}$
  (3) If $A, B \in \mathcal{F}$, then $A \cup B \in \mathcal{F}$
- The Kolmogorov Axioms characterize what a probability function P defined on a field is as follows:
  - $P(\Omega) = 1$ (Normalized)
  - $0 \leq P(A) \leq 1$ for $A \in \mathcal{F}$ (Non-negative)
  - If $A \cap B = \emptyset$, then $P(A \cup B) = P(A) + P(B)$ (Additive)

Recap of Kolmogorov Axioms

- We defined a field $\mathcal{F}$ as a non-empty collection of sets Ω that is closed under finite applications of set-theoretic operations of taking unions and complements.
- To say that a field $\mathcal{F}$ is closed under finite applications of set theoretic operations means:
  (1) $\Omega \in \mathcal{F}$
  (2) If $A \in \mathcal{F}$, then $A^c \in \mathcal{F}$
  (3) If $A, B \in \mathcal{F}$, then $A \cup B \in \mathcal{F}$
- The Kolmogorov Axioms characterize what a probability function P defined on a field is as follows:
  - $P(\Omega) = 1$ (Normalized)
  - $0 \leq P(A) \leq 1$ for $A \in \mathcal{F}$ (Non-negative)
  - If $A \cap B = \emptyset$, then $P(A \cup B) = P(A) + P(B)$ (Additive)
- A probability function is a normalized, non-negative and additive real-valued set function defined on a field.

Probability Spaces

- Kolmogorov Axioms tell you what the rules for a probability function should be.

Probability Spaces

- Kolmogorov Axioms tell you what the rules for a probability function should be.
- This means what many many functions in mathematics can be probability functions so long as they satisfy these axioms. Because of this we say that there are different interpretations of the sort of things we call probability functions. (We shall come back to this later in the course!)

Probability Spaces

- Kolmogorov Axioms tell you what the rules for a probability function should be.
- This means what many many functions in mathematics can be probability functions so long as they satisfy these axioms. Because of this we say that there are different interpretations of the sort of things we call probability functions. (We shall come back to this later in the course!)
- Because of how often probability functions defined on a field show up, and also due to historical reasons, mathematicians have a special name for the field $\mathcal{F}$ together with the probability function defined on it. It is called a probability space.

Probability Spaces

- Recall the definition of a field. A field $\mathcal{F}$ is closed under finite applications of set theoretic operations means:

Probability Spaces

- Recall the definition of a field. A field $\mathcal{F}$ is closed under finite applications of set theoretic operations means:
  (1) $\Omega \in \mathcal{F}$

Probability Spaces

- Recall the definition of a field. A field $\mathcal{F}$ is closed under finite applications of set theoretic operations means:
  (1) $\Omega \in \mathcal{F}$
  (2) If $A \in \mathcal{F}$, then $A^c \in \mathcal{F}$

Probability Spaces

- Recall the definition of a field. A field $\mathcal{F}$ is closed under finite applications of set theoretic operations means:
  (1) $\Omega \in \mathcal{F}$
  (2) If $A \in \mathcal{F}$, then $A^c \in \mathcal{F}$
  (3) If $A, B \in \mathcal{F}$, then $A \cup B \in \mathcal{F}$

Probability Spaces

- Recall the definition of a field. A field $\mathcal{F}$ is closed under finite applications of set theoretic operations means:
  (1) $\Omega \in \mathcal{F}$
  (2) If $A \in \mathcal{F}$, then $A^c \in \mathcal{F}$
  (3) If $A, B \in \mathcal{F}$, then $A \cup B \in \mathcal{F}$
- The set Ω in a probability space is called the sample space. We shall denote the sample space with S. The sample space S is the set of all possible outcomes of an experiment whose outcome is governed by chance.

Probability Spaces

- Recall the definition of a field. A field $\mathcal{F}$ is closed under finite applications of set theoretic operations means:
  (1) $\Omega \in \mathcal{F}$
  (2) If $A \in \mathcal{F}$, then $A^c \in \mathcal{F}$
  (3) If $A, B \in \mathcal{F}$, then $A \cup B \in \mathcal{F}$
- The set Ω in a probability space is called the sample space. We shall denote the sample space with S. The sample space S is the set of all possible outcomes of an experiment whose outcome is governed by chance.
- The sets $A \in \mathcal{F}$ in probability spaces are called, you guessed it, chance events or simply just events. We shall denote events with the letters E or F, with or without numerical subscripts.

Probability Spaces: Sample Spaces and Events

- An example of an experiment whose outcome is governed by chance is flipping a slice of toasted bread buttered on one side and seeing whether the buttered side shows “up” or “down”.

Probability Spaces: Sample Spaces and Events

- An example of an experiment whose outcome is governed by chance is flipping a slice of toasted bread buttered on one side and seeing whether the buttered side shows "up" or "down".
- Here the sample space can be written down by listing its members as a set:

$$S = \{\{\text{buttered side up}\}, \{\text{buttered side down}\}\}$$

Probability Spaces: Sample Spaces and Events

- An example of an experiment whose outcome is governed by chance is flipping a slice of toasted bread buttered on one side and seeing whether the buttered side shows "up" or "down".
- Here the sample space can be written down by listing its members as a set:

$$S = \{\{\text{buttered side up}\}, \{\text{buttered side down}\}\}$$

- The events here are: $E = \{\text{buttered side up}\}$ and $F = \{\text{buttered side down}\}$

Probability Spaces: Sample Spaces and Events

- Exercise. Specify the sample space S and list all the events of the following experiments:

Probability Spaces: Sample Spaces and Events

- Exercise. Specify the sample space S and list all the events of the following experiments:
  - The experiment of flipping a fair coin once.

Probability Spaces: Sample Spaces and Events

- Exercise. Specify the sample space S and list all the events of the following experiments:
  - The experiment of flipping a fair coin once.
  - The experiment of flipping a fair coin twice.

Probability Spaces: Sample Spaces and Events

- Exercise. Specify the sample space S and list all the events of the following experiments:
  - The experiment of flipping a fair coin once.
  - The experiment of flipping a fair coin twice.
  - The experiment of rolling a fair die once.

Probability Spaces: Sample Spaces and Events

- Exercise. Specify the sample space S and list all the events of the following experiments:
  - The experiment of flipping a fair coin once.
  - The experiment of flipping a fair coin twice.
  - The experiment of rolling a fair die once.
- See Homework 6 for more exercises.

Probability Spaces: Null Events

- We have already said that the $S \in \mathcal{F}$ and the sets $E \in \mathcal{F}$ have special names: the sample space and events, respectively.

Probability Spaces: Null Events

- We have already said that the $S \in \mathcal{F}$ and the sets $E \in \mathcal{F}$ have special names: the sample space and events, respectively.
- Since $\emptyset \in \mathcal{F}$ (Why?), it also has a special name: the null event.

Probability Spaces: Null Events

- We have already said that the $S \in \mathcal{F}$ and the sets $E \in \mathcal{F}$ have special names: the sample space and events, respectively.
- Since $\emptyset \in \mathcal{F}$ (Why?), it also has a special name: the null event.
- We shall say what two events E and F are mutually exclusive if $(E \cap F) = \emptyset$, i.e., the intersection of E and F is the null event.

Probability Spaces: Null Events

- We have already said that the $S \in \mathcal{F}$ and the sets $E \in \mathcal{F}$ have special names: the sample space and events, respectively.
- Since $\emptyset \in \mathcal{F}$ (Why?), it also has a special name: the null event.
- We shall say what two events E and F are mutually exclusive if $(E \cap F) = \emptyset$, i.e., the intersection of E and F is the null event.
- Exercise. Specify the null events in the following experiments:

Probability Spaces: Null Events

- We have already said that the $S \in \mathcal{F}$ and the sets $E \in \mathcal{F}$ have special names: the sample space and events, respectively.
- Since $\emptyset \in \mathcal{F}$ (Why?), it also has a special name: the null event.
- We shall say what two events E and F are mutually exclusive if $(E \cap F) = \emptyset$, i.e., the intersection of E and F is the null event.
- Exercise. Specify the null events in the following experiments:
  - The experiment of flipping a fair coin once.

Probability Spaces: Null Events

- We have already said that the $S \in \mathcal{F}$ and the sets $E \in \mathcal{F}$ have special names: the sample space and events, respectively.
- Since $\emptyset \in \mathcal{F}$ (Why?), it also has a special name: the null event.
- We shall say what two events E and F are mutually exclusive if $(E \cap F) = \emptyset$, i.e., the intersection of E and F is the null event.
- Exercise. Specify the null events in the following experiments:
  - The experiment of flipping a fair coin once.
  - The experiment of flipping a fair coin twice.

Probability Spaces: Null Events

- We have already said that the $S \in \mathcal{F}$ and the sets $E \in \mathcal{F}$ have special names: the sample space and events, respectively.
- Since $\emptyset \in \mathcal{F}$ (Why?), it also has a special name: the null event.
- We shall say what two events E and F are mutually exclusive if $(E \cap F) = \emptyset$, i.e., the intersection of E and F is the null event.
- Exercise. Specify the null events in the following experiments:
  - The experiment of flipping a fair coin once.
  - The experiment of flipping a fair coin twice.
  - The experiment of rolling a fair die once.

Probability Spaces: Null Events

- We have already said that the $S \in \mathcal{F}$ and the sets $E \in \mathcal{F}$ have special names: the sample space and events, respectively.
- Since $\emptyset \in \mathcal{F}$ (Why?), it also has a special name: the null event.
- We shall say what two events E and F are mutually exclusive if $(E \cap F) = \emptyset$, i.e., the intersection of E and F is the null event.
- Exercise. Specify the null events in the following experiments:
  - The experiment of flipping a fair coin once.
  - The experiment of flipping a fair coin twice.
  - The experiment of rolling a fair die once.
- See Homework 6 for more exercises.

Using the axioms to calculate $P(E)$

- One way of reading the third Kolmogorov Axiom of probability functions is that if $A, B \in \mathcal{F}$ are mutually exclusive, then $P(A \cup B) = P(A) + P(B)$.

Using the axioms to calculate $P(E)$

- One way of reading the third Kolmogorov Axiom of probability functions is that if $A, B \in \mathcal{F}$ are mutually exclusive, then $P(A \cup B) = P(A) + P(B)$.
- We shall call this rule of probability theory, the special disjunction rule.

Using the axioms to calculate $P(E)$

- One way of reading the third Kolmogorov Axiom of probability functions is that if $A, B \in \mathcal{F}$ are mutually exclusive, then $P(A \cup B) = P(A) + P(B)$.
- We shall call this rule of probability theory, the special disjunction rule.
- Exercise. Use the special disjunction rule and Kolmogorov axioms to find the probability of the following events.

Using the axioms to calculate $P(E)$

- One way of reading the third Kolmogorov Axiom of probability functions is that if $A, B \in \mathcal{F}$ are mutually exclusive, then $P(A \cup B) = P(A) + P(B)$.
- We shall call this rule of probability theory, the special disjunction rule.
- Exercise. Use the special disjunction rule and Kolmogorov axioms to find the probability of the following events.
  - E = Two heads in two tosses of a coin

Using the axioms to calculate $P(E)$

- One way of reading the third Kolmogorov Axiom of probability functions is that if $A, B \in \mathcal{F}$ are mutually exclusive, then $P(A \cup B) = P(A) + P(B)$.
- We shall call this rule of probability theory, the special disjunction rule.
- Exercise. Use the special disjunction rule and Kolmogorov axioms to find the probability of the following events.
  - E = Two heads in two tosses of a coin
  - E = Rolling a six

Using the axioms to calculate $P(E)$

- One way of reading the third Kolmogorov Axiom of probability functions is that if $A, B \in \mathcal{F}$ are mutually exclusive, then $P(A \cup B) = P(A) + P(B)$.
- We shall call this rule of probability theory, the special disjunction rule.
- Exercise. Use the special disjunction rule and Kolmogorov axioms to find the probability of the following events.
  - E = Two heads in two tosses of a coin
  - E = Rolling a six
  - E = At least one head in two tosses of a coin

Using the axioms to calculate $P(E)$

- One way of reading the third Kolmogorov Axiom of probability functions is that if $A, B \in \mathcal{F}$ are mutually exclusive, then $P(A \cup B) = P(A) + P(B)$.
- We shall call this rule of probability theory, the special disjunction rule.
- Exercise. Use the special disjunction rule and Kolmogorov axioms to find the probability of the following events.
  - E = Two heads in two tosses of a coin
  - E = Rolling a six
  - E = At least one head in two tosses of a coin
  - E = Rolling a one or a six.

Using the axioms to calculate $P(E)$

- One way of reading the third Kolmogorov Axiom of probability functions is that if $A, B \in \mathcal{F}$ are mutually exclusive, then $P(A \cup B) = P(A) + P(B)$.
- We shall call this rule of probability theory, the special disjunction rule.
- Exercise. Use the special disjunction rule and Kolmogorov axioms to find the probability of the following events.
  - E = Two heads in two tosses of a coin
  - E = Rolling a six
  - E = At least one head in two tosses of a coin
  - E = Rolling a one or a six.
- See Homework 6 for more exercises.

Using the axioms to calculate $P(E)$

- What the exercises have shown us is that so long as events are: (1) mutually exclusive and; (2) we have a finite sample space, we can use Axiom 3 and Axiom 1 to define:

$$P(E) = \frac{\text{Number of elements in } E}{\text{Number of all possible outcomes in } S}$$

Using the axioms to calculate $P(E)$

- What the exercises have shown us is that so long as events are: (1) mutually exclusive and; (2) we have a finite sample space, we can use Axiom 3 and Axiom 1 to define:

$$P(E) = \frac{\text{Number of elements in } E}{\text{Number of all possible outcomes in } S}$$

- We shall take this as the "definition" of the probability of an event in this class, which some of you may be familiar with from high school or elementary school.

Using the axioms to calculate $P(E)$

- What the exercises have shown us is that so long as events are: (1) mutually exclusive and; (2) we have a finite sample space, we can use Axiom 3 and Axiom 1 to define:

$$P(E) = \frac{\text{Number of elements in } E}{\text{Number of all possible outcomes in } S}$$

- We shall take this as the "definition" of the probability of an event in this class, which some of you may be familiar with from high school or elementary school.
- In more advanced classes, you will learn that this definition of $P(E)$ is too simple! What if the size of Ω is infinite? Obviousness is ....

Using the axioms to calculate $P(E)$

- What the exercises have shown us is that so long as events are: (1) mutually exclusive and; (2) we have a finite sample space, we can use Axiom 3 and Axiom 1 to define:

$$P(E) = \frac{\text{Number of elements in } E}{\text{Number of all possible outcomes in } S}$$

- We shall take this as the "definition" of the probability of an event in this class, which some of you may be familiar with from high school or elementary school.
- In more advanced classes, you will learn that this definition of $P(E)$ is too simple! What if the size of Ω is infinite? Obviousness is ....
- But, even in the infinite case, we will still take the additive sum (by Axiom 3) and normalize (by Axiom 1) using integration.

Using the axioms to calculate $P(E)$

- We are now in the position to “derive” other rules of probability functions that we can use in actual calculations.

Using the axioms to calculate $P(E)$

- We are now in the position to “derive” other rules of probability functions that we can use in actual calculations.
- Let us label, for future reference, the three Kolmogorov axioms for probability spaces as:

Using the axioms to calculate $P(E)$

- We are now in the position to "derive" other rules of probability functions that we can use in actual calculations.
- Let us label, for future reference, the three Kolmogorov axioms for probability spaces as:

Rule 1 $P(S) = 1$

Using the axioms to calculate $P(E)$

- We are now in the position to "derive" other rules of probability functions that we can use in actual calculations.
- Let us label, for future reference, the three Kolmogorov axioms for probability spaces as:

Rule 1 $P(S) = 1$
Rule 2 $0 \leq P(E) \leq 1$

Using the axioms to calculate $P(E)$

- We are now in the position to "derive" other rules of probability functions that we can use in actual calculations.
- Let us label, for future reference, the three Kolmogorov axioms for probability spaces as:

Rule 1 $P(S) = 1$

Rule 2 $0 \leq P(E) \leq 1$

Rule 3 If $E \cap F = \emptyset$, then $P(E \cup F) = P(E) + P(F)$ (Special Disjunction Rule)

Using the axioms to calculate $P(E)$

- We are now in the position to "derive" other rules of probability functions that we can use in actual calculations.
- Let us label, for future reference, the three Kolmogorov axioms for probability spaces as:

  Rule 1 $P(S) = 1$

  Rule 2 $0 \leq P(E) \leq 1$

  Rule 3 If $E \cap F = \emptyset$, then $P(E \cup F) = P(E) + P(F)$ (Special Disjunction Rule)
- We now add the following "derived" rules of probability theory. (Note that in the following rules whenever we write E^c we shall mean $(E^c \cap S)$.)

Using the axioms to calculate $P(E)$

- We are now in the position to "derive" other rules of probability functions that we can use in actual calculations.
- Let us label, for future reference, the three Kolmogorov axioms for probability spaces as:

Rule 1 $P(S) = 1$

Rule 2 $0 \leq P(E) \leq 1$

Rule 3 If $E \cap F = \emptyset$, then $P(E \cup F) = P(E) + P(F)$ (Special Disjunction Rule)

- We now add the following "derived" rules of probability theory. (Note that in the following rules whenever we write E^c we shall mean $(E^c \cap S)$.)

Rule 4 $P(E \cup E^c) = 1$

Using the axioms to calculate $P(E)$

- We are now in the position to "derive" other rules of probability functions that we can use in actual calculations.
- Let us label, for future reference, the three Kolmogorov axioms for probability spaces as:

Rule 1 $P(S) = 1$

Rule 2 $0 \leq P(E) \leq 1$

Rule 3 If $E \cap F = \emptyset$, then $P(E \cup F) = P(E) + P(F)$ (Special Disjunction Rule)

- We now add the following "derived" rules of probability theory. (Note that in the following rules whenever we write E^c we shall mean $(E^c \cap S)$.)

Rule 4 $P(E \cup E^c) = 1$

Rule 5 $P(E^c) = 1 - P(E)$

Using the axioms to calculate $P(E)$

- We are now in the position to "derive" other rules of probability functions that we can use in actual calculations.
- Let us label, for future reference, the three Kolmogorov axioms for probability spaces as:

Rule 1 $P(S) = 1$

Rule 2 $0 \leq P(E) \leq 1$

Rule 3 If $E \cap F = \emptyset$, then $P(E \cup F) = P(E) + P(F)$ (Special Disjunction Rule)

- We now add the following "derived" rules of probability theory. (Note that in the following rules whenever we write E^c we shall mean $(E^c \cap S)$.)

Rule 4 $P(E \cup E^c) = 1$

Rule 5 $P(E^c) = 1 - P(E)$

Rule 6 $P(\emptyset) = 0$

Using the axioms to calculate $P(E)$

- We are now in the position to "derive" other rules of probability functions that we can use in actual calculations.
- Let us label, for future reference, the three Kolmogorov axioms for probability spaces as:

Rule 1 $P(S) = 1$

Rule 2 $0 \leq P(E) \leq 1$

Rule 3 If $E \cap F = \emptyset$, then $P(E \cup F) = P(E) + P(F)$ (Special Disjunction Rule)

- We now add the following "derived" rules of probability theory. (Note that in the following rules whenever we write E^c we shall mean $(E^c \cap S)$.)

Rule 4 $P(E \cup E^c) = 1$

Rule 5 $P(E^c) = 1 - P(E)$

Rule 6 $P(\emptyset) = 0$

Rule 7 $P(E) = P(E \cap F) + P(E \cap F^c)$ (Partition of an Event Rule)

Using the axioms to calculate $P(E)$

- We are now in the position to "derive" other rules of probability functions that we can use in actual calculations.

- Let us label, for future reference, the three Kolmogorov axioms for probability spaces as:

  Rule 1 $P(S) = 1$
  Rule 2 $0 \leq P(E) \leq 1$
  Rule 3 If $E \cap F = \emptyset$, then $P(E \cup F) = P(E) + P(F)$ (Special Disjunction Rule)

- We now add the following "derived" rules of probability theory. (Note that in the following rules whenever we write E^c we shall mean $(E^c \cap S)$.)

  Rule 4 $P(E \cup E^c) = 1$
  Rule 5 $P(E^c) = 1 - P(E)$
  Rule 6 $P(\emptyset) = 0$
  Rule 7 $P(E) = P(E \cap F) + P(E \cap F^c)$ (Partition of an Event Rule)

- See Homework 6 for the derivation of more rules.