Larson and Marx Section 2.3

Gene Quinn

The Probability Function

Recall the following definitions:

An **experiment** is a procedure that:

- in theory, can be repeated an infinite number of times
- has a well-defined set of possible outcomes

The Probability Function

Recall the following definitions:

An **experiment** is a procedure that:

- in theory, can be repeated an infinite number of times
- has a well-defined set of possible outcomes

A **sample outcome** is one of the possible outcomes of an experiment.

The Probability Function

Recall the following definitions:

An **experiment** is a procedure that:

- in theory, can be repeated an infinite number of times
- has a well-defined set of possible outcomes

A **sample outcome** is one of the possible outcomes of an experiment.

The **sample space** is the set containing all possible sample outcomes of an experiment.

The Probability Function

Recall the following definitions:

An **experiment** is a procedure that:

- in theory, can be repeated an infinite number of times
- has a well-defined set of possible outcomes

A **sample outcome** is one of the possible outcomes of an experiment.

The **sample space** is the set containing all possible sample outcomes of an experiment.

An **event** is a subset of the sample space, including the subsets represented by individual outcomes.

The Probability Function

We are now in a position to define the **probability function**

The **probability function** associates a real number with each event defined on the sample space.

The Probability Function

We are now in a position to define the **probability function**

The **probability function** associates a real number with each event defined on the sample space.

The domain of the probability function consists of the events defined on the sample space, which include the individual outcomes.

The Probability Function

Suppose S is the sample space of an experiment and $\mathcal{Q}$ is the collection of subsets of S that contains all events defined on the sample space.

The **probability function**

$$P : \mathcal{Q} \mapsto \mathcal{R}$$

maps **subsets** of the sample space representing defined events into the real numbers.

The Probability Function

Suppose S is the sample space of an experiment and $\mathcal{Q}$ is the collection of subsets of S that contains all events defined on the sample space.

The **probability function**

$$P : \mathcal{Q} \longmapsto \mathcal{R}$$

maps **subsets** of the sample space representing defined events into the real numbers.

The Kolmogorov axioms are necessary and sufficient conditions for characterizing the probability function P.

The Probability Function

Suppose S is the sample space of an experiment and $\mathcal{Q}$ is the collection of subsets of S that contains all events defined on the sample space.

The **probability function**

$$P : \mathcal{Q} \mapsto \mathcal{R}$$

maps **subsets** of the sample space representing defined events into the real numbers.

The Kolmogorov axioms are necessary and sufficient conditions for characterizing the probability function P.

If the sample space S is finite, three axioms are necessary and sufficient for characterizing the probability function P.

The Probability Function

Suppose S is the sample space of an experiment and $\mathcal{Q}$ is the collection of subsets of S that contains all events defined on the sample space.

The **probability function**

$$P : \mathcal{Q} \mapsto \mathcal{R}$$

maps **subsets** of the sample space representing defined events into the real numbers.

The Kolmogorov axioms are necessary and sufficient conditions for characterizing the probability function P.

If the sample space S is finite, three axioms are necessary and sufficient for characterizing the probability function P.

If S has an infinite number of elements, a fourth axiom is needed.

The Kolmogorov Axioms

An **axiom** is a statement that is assumed to be true, without requiring proof.

The Kolmogorov Axioms

An **axiom** is a statement that is assumed to be true, without requiring proof.

Axioms are used to prove other statements, called **theorems**.

The Kolmogorov Axioms

Axiom 1 Let A be any event defined over a sample space S. Then

$$P(A) \geq 0$$

The Kolmogorov Axioms

Axiom 1 Let A be any event defined over a sample space S. Then

$$P(A) \geq 0$$

Axiom 2 If S is the entire sample space, then

$$P(S) = 1$$

The Kolmogorov Axioms

Two events A and B are called **mutually exclusive** if $A \cap B = \emptyset$.

The Kolmogorov Axioms

Two events A and B are called **mutually exclusive** if $A \cap B = \emptyset$.

Axiom 3:

Let A and B be any two mutually exclusive events defined over S. Then

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

The Kolmogorov Axioms

Two events A and B are called **mutually exclusive** if $A \cap B = \emptyset$.

Axiom 3:

Let A and B be any two mutually exclusive events defined over S. Then

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

When the sample space S is finite, these three axioms are necessary and sufficient for characterizing the probability function P.

The Kolmogorov Axioms

When the sample space S is not finite, the following additional axiom is required.

The Kolmogorov Axioms

When the sample space S is not finite, the following additional axiom is required.

Axiom 4:

Let $A_1, A_2, \ldots$ be events over S.

If

$$A_i \cap A_j = \emptyset \quad \text{for each} \quad i \neq j$$

then

$$P\left(\bigcup_{i=1}^{\infty}\right) = \sum_{i=1}^{\infty} P(A_i)$$

Properties of the Probability Function

Any function on the sample space that satisfies the Kolmogorov axioms must have certain basic properties.

These properties follow as theorems from the axioms, and require proof.

Properties of the Probability Function

Any function on the sample space that satisfies the Kolmogorov axioms must have certain basic properties.

These properties follow as theorems from the axioms, and require proof.

Theorem:

Let A be any event defined on the sample space. Then

$$P(A^c) = 1 - P(A)$$

i.e., for any event A, $P(A)$ and $P(A^c)$ must add to 1.

Properties of the Probability Function

Proof:

By Axiom 2 and the definition of the sample space,

$$P(S) = 1 \quad \text{and} \quad S = A \cup A^c$$

Properties of the Probability Function

Proof:

By Axiom 2 and the definition of the sample space,

$$P(S) = 1 \quad \text{and} \quad S = A \cup A^c$$

By Axiom 3, since A and A^c are mutually exclusive,

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

and the result follows from the fact that $A \cup A^c = S$.

Properties of the Probability Function

Theorem:

If $A \subset B$, then

$$P(A) \leq P(B)$$

Properties of the Probability Function

Theorem:

If $A \subset B$, then

$$P(A) \leq P(B)$$

Proof: Write the event B in the equivalent form

$$B = A \cup (B \cap A^c)$$

and note that

$$A \cap (B \cap A^c) = \emptyset$$

By Axiom 3, since A and $B \cap A^c$ are mutually exclusive,

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

and by Axiom 1, $P(B \cap A^c) \geq 0$, so $P(A) \leq P(B)$.

Properties of the Probability Function

Theorem:

For any events A and B in the sample space S,

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

Properties of the Probability Function

Theorem:

For any events A and B in the sample space S,

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

Proof: Write A as the union of two disjoint sets,

$$A = (A \cap B^c) \cup (A \cap B)$$

Then by Axiom 3,

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

Properties of the Probability Function

Theorem:

For any events A and B in the sample space S,

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

Proof: Write A as the union of two disjoint sets,

$$A = (A \cap B^c) \cup (A \cap B)$$

Then by Axiom 3,

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

By a similar argument,

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

Properties of the Probability Function

Adding the two equations

$$
\begin{aligned}
P(A) &= P(A \cap B^c) + P(A \cap B) \\
P(B) &= P(B \cap A^c) + P(A \cap B)
\end{aligned}
$$

gives

$$P(A) + P(B) = [P(A \cap B^c) + P(B \cap A^c) + P(A \cap B)] + P(A \cap B)$$

Properties of the Probability Function

Adding the two equations

$$\begin{aligned} P(A) &= P(A \cap B^c) + P(A \cap B) \\ P(B) &= P(B \cap A^c) + P(A \cap B) \end{aligned}$$

gives

$$P(A) + P(B) = [P(A \cap B^c) + P(B \cap A^c) + P(A \cap B)] + P(A \cap B)$$

Recognizing that

$$(A \cap B^c) \cup (B \cap A^c) \cup (A \cap B) = A \cup B,$$

and the three sets on the left are mutually exclusive, so by Theroem 2.3.5,

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

Properties of the Probability Function

Substituting $P(A \cup B)$ for

$$P(A \cap B^c) + P(B \cap A^c) + P(A \cap B)$$

gives us

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

and if follows that

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

Properties of the Probability Function

Substituting $P(A \cup B)$ for

$$P(A \cap B^c) + P(B \cap A^c) + P(A \cap B)$$

gives us

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

and if follows that

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

This theorem is used very frequently.

Conditional Probability

If we know for a fact that some event has occurred, the probabilities of other events may need to be revised.

This revised probability of an event is known as the **conditional probability** that the event occurs, given the knowledge that some other event has definitely occurred.

Conditional Probability

If we know for a fact that some event has occurred, the probabilities of other events may need to be revised.

This revised probability of an event is known as the **conditional probability** that the event occurs, given the knowledge that some other event has definitely occurred.

Definition:

Let A and B be any two events defined on S such that $P(B) > 0$.

The **conditional probability** of A given that B has already occurred, denoted by

$$P(A|B)$$

is given by

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

Conditional Probability

The following rearrangement of the formula defining conditional probability is often useful:

$$P(A \cap B) = P(A|B) \cdot P(B)$$