Larson and Marx Section 3.1

Gene Quinn

Random Variables

Given a sample space $\mathcal{S}$, we have defined probabilities in terms of a probability function that maps the power set $\mathcal{P}(S)$ into the real numbers

$$P : \mathcal{P}(S) \mapsto \mathcal{R}$$

in a way that is consistent with the Kolmogorov axioms.

Random Variables

Given a sample space $\mathcal{S}$, we have defined probabilities in terms of a probability function that maps the power set $\mathcal{P}(S)$ into the real numbers

$$P : \mathcal{P}(S) \mapsto \mathcal{R}$$

in a way that is consistent with the Kolmogorov axioms.

It is often the case that the power set $\mathcal{P}(S)$ has far more granularity than applications require, and in fact things can be greatly simplified by "redefining" the sample space to elimate some of the detail.

Random Variables

Example:

Six dice are thrown. By the multiplication rule, there are

$$6 \cdot 6 \cdot 6 \cdot 6 \cdot 6 \cdot 6 \quad = \quad 46,656$$

distinct outcomes in the sample space.

Random Variables

Example:

Six dice are thrown. By the multiplication rule, there are

$$6 \cdot 6 \cdot 6 \cdot 6 \cdot 6 \cdot 6 \quad = \quad 46,656$$

distinct outcomes in the sample space.

However, chances are we are only interested in the **sum** of the six faces that come up.

Consequently, we really only care which of the 31 possible totals

$$\{6, 7, 8, \ldots, 34, 35, 36\}$$

has occurred.

Random Variables

One way we can "simplify" a sample space with $46,656$ elements down to one with 31 mutually exclusive outcomes is to define a *function* from the original sample space into the numbers from 6 to 36.

Every outcome from the original sample space is mapped into exactly one number in the set $\{6, 7, \ldots, 35, 36\}$.

Random Variables

One way we can "simplify" a sample space with $46,656$ elements down to one with 31 mutually exclusive outcomes is to define a *function* from the original sample space into the numbers from 6 to 36.

Every outcome from the original sample space is mapped into exactly one number in the set $\{6, 7, \ldots, 35, 36\}$.

In fact, this is standard practice in the study of probability, leading to the following

Definition:

A real-valued function defined on a sample space is called a **random variable**

Random Variables

When we say that the function is defined on a sample space, what we mean is that the independent variable is a point in a sample space, i.e., the outcome of an experiment.

Random Variables

When we say that the function is defined on a sample space, what we mean is that the independent variable is a point in a sample space, i.e., the outcome of an experiment.

The choice of the term **random variable** is somewhat misleading since it is, in fact, a *function*.

Some authors consider it unfortunate that the more appropriate term **random function** is not used.

Random Variables

When we say that the function is defined on a sample space, what we mean is that the independent variable is a point in a sample space, i.e., the outcome of an experiment.

The choice of the term **random variable** is somewhat misleading since it is, in fact, a *function*.

Some authors consider it unfortunate that the more appropriate term **random function** is not used.

At least in the beginning, it is somewhat helpful to think

"random function"

when you see the term **"random variable"**.

Random Variables

Compounding the confusion resulting from calling a random function a **random variable** is the standard notation for random variables.

Random Variables

Compounding the confusion resulting from calling a random function a **random variable** is the standard notation for random variables.

A random variable is usually designated by an upper case alphabetic character, often in bold type:

$$X = 17$$

Random Variables

Compounding the confusion resulting from calling a random function a **random variable** is the standard notation for random variables.

A random variable is usually designated by an upper case alphabetic character, often in bold type:

$$X = 17$$

In the context of our tossing six dice experiment, if X was a random variable defined to be the sum of the six faces, this is how we would denote the outcome that the sum was 17.

Random Variables

If we think of X as a function defined on S (which it is),

$$X : S \mapsto \{6, 7, \ldots, 35, 36\}$$

we could write

$$X(s) = 17$$

Random Variables

If we think of X as a function defined on S (which it is),

$$X : S \mapsto \{6, 7, \ldots, 35, 36\}$$

we could write

$$X(s) = 17$$

This notation would indicate more clearly the nature of the situation, namely that an event $s \in S$ occurred and the function defining the random variable X maps this event to the number 17.

Random Variables

If we think of X as a function defined on S (which it is),

$$X : S \mapsto \{6, 7, \ldots, 35, 36\}$$

we could write

$$X(s) = 17$$

This notation would indicate more clearly the nature of the situation, namely that an event $s \in S$ occurred and the function defining the random variable X maps this event to the number 17.

But this notation is never used, so we have to live with $X = 17$.

Random Variables - Notation

Suppose X is a random variable.

Let

$$R = \{x_1, x_2, x_3, \ldots\}$$

be the set of values which X can assume.

Random Variables - Notation

Suppose X is a random variable.

Let

$$R = \{x_1, x_2, x_3, \ldots\}$$

be the set of values which X can assume.

The fact that we wrote R in the form of a list implies that R is either finite or countably infinite.

Random Variables - Notation

Suppose X is a random variable.

Let

$$R = \{x_1, x_2, x_3, \ldots\}$$

be the set of values which X can assume.

The fact that we wrote R in the form of a list implies that R is either finite or countably infinite.

Random variables which assume only a finite or countable number of values are called **discrete random variables**

Random Variables - Notation

For any element of R, say

$$x_i \in R\,,$$

the collection of all sample points on which X assumes the value x_i constitutes the **event** that

$$X = x_i$$

Random Variables - Notation

For any element of R, say

$$x_i \in R\,,$$

the collection of all sample points on which X assumes the value x_i constitutes the **event** that

$$X = x_i$$

The probability of this event is denoted by

$$P(X = x_i)$$

The Probability Distribution

Definition:

The function

$$f(x_i) \quad = \quad P(X = x_i), \qquad i = 1, 2, \ldots$$

which is defined on the set of values R assumed by the random variable X is called the

(probability) **distribution** of the random variable X.

The Probability Distribution

Definition:

The function

$$f(x_i) \quad = \quad P(X = x_i), \qquad i = 1, 2, \ldots$$

which is defined on the set of values R assumed by the random variable X is called the

(probability) **distribution** of the random variable X.

The Probability Distribution

If a function f is a probability distribution, the values it assumes are /textitprobabilities.

Consequently, a probability distribution function can assume only nonnegative values:

$$f(x_i) \quad \geq \quad 0 \qquad \forall i$$

The Probability Distribution

If a function f is a probability distribution, the values it assumes are /textitprobabilities.

Consequently, a probability distribution function can assume only nonnegative values:

$$f(x_i) \quad \geq \quad 0 \qquad \forall i$$

Furthermore, the sum of the values of the probability distribution function over all values assumed by the random variable must be one:

$$\sum_i f(x_i) \quad = \quad 1$$

Example: The Binomial Distribution

An experiment consists of tossing 3 coins, or, equivalently, tossing the same coin 3 times.

We are interested in the number of heads obtained.

Example: The Binomial Distribution

An experiment consists of tossing 3 coins, or, equivalently, tossing the same coin 3 times.

We are interested in the number of heads obtained.

We define a random variable X representing the number of heads. Actually, X is a function from the sample space S into the set $A = \{0, 1, 2, 3\}$, defined by the following table:

$$X : S \mapsto A \quad \text{such that} \begin{cases} \{t,t,t\} & \to & 0 \\ \{t,t,h\}, \{t,h,t\}, \{h,t,t\} & \to & 1 \\ \{t,h,h\}, \{h,t,h\}, \{h,h,t\} & \to & 2 \\ \{h,h,h\} & \to & 3 \end{cases}$$

Example: The Binomial Distribution

The event

$$X = 2$$

will occur if the experiment produces one of the outcomes

$$\{t, h, h\}, \{h, t, h\}, \{h, h, t\}$$

Example: The Binomial Distribution

The event

$$X = 2$$

will occur if the experiment produces one of the outcomes

$$\{t, h, h\}, \{h, t, h\}, \{h, h, t\}$$

We can now define a **probability distribution** f on A, the set of possible values of the random variable X.

Example: The Binomial Distribution

The event

$$X = 2$$

will occur if the experiment produces one of the outcomes

$$\{t, h, h\}, \{h, t, h\}, \{h, h, t\}$$

We can now define a **probability distribution** f on A, the set of possible values of the random variable X.

f will be defined as:

$$f(x_i) \quad = \quad P(X = x_i) \qquad x_i \in \{0, 1, 2, 3\}$$

Example: The Binomial Distribution

We will define f by the following table:

$$f : X \mapsto [0,1] \quad \text{such that} \begin{cases} 0 & \rightarrow & P(X=0) \\ 1 & \rightarrow & P(X=1) \\ 2 & \rightarrow & P(x=2) \\ 3 & \rightarrow & P(x=3) \end{cases}$$

Example: The Binomial Distribution

We will define f by the following table:

$$f : X \mapsto [0,1] \quad \text{such that} \begin{cases} 0 & \rightarrow & P(X=0) \\ 1 & \rightarrow & P(X=1) \\ 2 & \rightarrow & P(x=2) \\ 3 & \rightarrow & P(x=3) \end{cases}$$

Defined in this fashion, f is called a **probability distribution** for the random variable X.

In this situation, we also say that the random variable X has distribution f.

Example: The Binomial Distribution

If the coin is fair, each of the eight outcomes in the sample space S has probability $1/8$, so we expect the following probabilities for the random variable X:

$$f : X \mapsto [0,1] \quad \text{such that} \begin{cases} 0 & \rightarrow & P(X=0) & = & 1/8 \\ 1 & \rightarrow & P(X=1) & = & 3/8 \\ 2 & \rightarrow & P(x=2) & = & 3/8 \\ 3 & \rightarrow & P(x=3) & = & 1/8 \end{cases}$$

Example: The Binomial Distribution

If the coin is fair, each of the eight outcomes in the sample space S has probability $1/8$, so we expect the following probabilities for the random variable X:

$$f : X \mapsto [0,1] \quad \text{such that} \begin{cases} 0 & \rightarrow & P(X=0) & = & 1/8 \\ 1 & \rightarrow & P(X=1) & = & 3/8 \\ 2 & \rightarrow & P(x=2) & = & 3/8 \\ 3 & \rightarrow & P(x=3) & = & 1/8 \end{cases}$$

If we add up the values of $f(X)$ over the domain of f, which is $\{0,1,2,3\}$, the set of values that the random variable X can assume, we have

$$\sum f(x_i) \quad = \quad \sum P(X=x_i), \quad x_i \in \{0,1,2,3\} \quad = 1$$

as required.

Example: The Binomial Distribution

Suppose instead of a coin toss, we considered an experiment with two possible outcomes that are not necessarily equally likely.

Example: The Binomial Distribution

Suppose instead of a coin toss, we considered an experiment with two possible outcomes that are not necessarily equally likely.

We will call the two outcomes "success" and "failure", and denote them by "s" and "f".

Example: The Binomial Distribution

Suppose instead of a coin toss, we considered an experiment with two possible outcomes that are not necessarily equally likely.

We will call the two outcomes "success" and "failure", and denote them by "s" and "f".

Since these are the only possible outcomes of the experiment, our probability theorems require that

$$P(success) = 1 - P(failure)$$

Example: The Binomial Distribution

Suppose instead of a coin toss, we considered an experiment with two possible outcomes that are not necessarily equally likely.

We will call the two outcomes "success" and "failure", and denote them by "s" and "f".

Since these are the only possible outcomes of the experiment, our probability theorems require that

$$P(success) = 1 - P(failure)$$

We will denote the probability of success by p.

Example: The Binomial Distribution

Our experiment now had two outcomes, "success" and "failure".

Example: The Binomial Distribution

Our experiment now had two outcomes, "success" and "failure".

We define a random variable X representing the number of successes. Actually, X is a function from the sample space S into the set $A = \{0, 1, 2, 3\}$, defined by the following table:

$$X : S \mapsto A \quad \text{such that} \begin{cases} \{f, f, f\} & \rightarrow \quad 0 \\ \{f, f, s\}, \{f, s, f\}, \{s, f, f\} & \rightarrow \quad 1 \\ \{f, s, s\}, \{s, f, s\}, \{s, s, f\} & \rightarrow \quad 2 \\ \{h, h, h\} & \rightarrow \quad 3 \end{cases}$$