Probability Overview

Gene Quinn

Experiments and Outcomes

An **experiment** is a procedure which:

1) Has a well-defined set of possible results known as **outcomes**

and

2) Can be repeated an infinite number of times (at least in theory).

Experiments and Outcomes

An **experiment** is a procedure which:

1) Has a well-defined set of possible results known as **outcomes**

and

2) Can be repeated an infinite number of times (at least in theory).

Each of the possible results is called a **sample outcome**

Sample Outcomes and Sample Spaces

The set of possible results or outcomes of an experiment taken as a whole is called the **sample space**.

Sample Outcomes and Sample Spaces

The set of possible results or outcomes of an experiment taken as a whole is called the **sample space**.

A subset of the sample space is called an **event**.

Events include:

- an individual outcome
- the entire sample space
- the empty set $\emptyset$

Probability Functions

A **probability function** is a real-valued function that maps *events* into real numbers in the closed interval $[0, 1]$ in a manner that is consistent with the Kolmogorov axioms.

Probability Functions

A **probability function** is a real-valued function that maps *events* into real numbers in the closed interval $[0, 1]$ in a manner that is consistent with the Kolmogorov axioms.

The domain of a probability function is a collection of sets, usually the power set of a sample space.

A probability function always takes values in $[0, 1]$.

(You should convince yourself that choosing any values outside this interval would inevitably cause the Kolmogorov axioms to be violated)

Discrete Random Variables

Definition

A **discrete random variable** is a *real-valued function* whose domain is a sample space S having finite or countably infinite cardinality.

A random variable maps an outcome of an experiment into a real number.

Discrete Random Variables

Definition

A **discrete random variable** is a *real-valued function* whose domain is a sample space S having finite or countably infinite cardinality.

A random variable maps an outcome of an experiment into a real number.

Random variables are denoted by upper case letters: X, Y

Discrete Random Variables

Definition

A **discrete random variable** is a *real-valued function* whose domain is a sample space S having finite or countably infinite cardinality.

A random variable maps an outcome of an experiment into a real number.

Random variables are denoted by upper case letters: X, Y

The range of a random variable (i.e., the set of values it can assume) often has far fewer elements than the underlying sample space, because many outcomes often map to one real number.

Probability Overview

Keep in mind the following similarities and differences between a *probability function* and a *random variable*:

Probability Function	**Random Variable**
maps sets into real numbers	maps sets into real numbers
domain is the power set of a sample space	domain is a sample space
only takes values in $[0, 1]$	can take any real value, positive or negative
must be consistent with Kolmogorov axioms	assignment of real numbers is arbitrary

Probability Density Function

A **probability density function** maps a random variable into a real number in the interval $[0, 1]$.

Probability Density Function

A **probability density function** maps a random variable into a real number in the interval $[0, 1]$.

Definition:

The **probability density function** or pdf of a random variable X, $p_X(k)$ is defined by

$$p_X(k) = P(s \in S \mid X(s) = k)$$

Probability Density Function

A **probability density function** maps a random variable into a real number in the interval $[0, 1]$.

Definition:

The **probability density function** or pdf of a random variable X, $p_X(k)$ is defined by

$$p_X(k) = P(s \in S \mid X(s) = k)$$

Usually the simpler notation which omits explicit reference to s and S is used,

$$p_X(k) = P(X = k)$$

Probability Density Function

A **probability density function** maps a random variable into a real number in the interval $[0, 1]$.

Definition:

The **probability density function** or pdf of a random variable X, $p_X(k)$ is defined by

$$p_X(k) = P(s \in S \mid X(s) = k)$$

Usually the simpler notation which omits explicit reference to s and S is used,

$$p_X(k) = P(X = k)$$

In other words, the probability density function for a discrete random variable maps each value the random variable can assume into the probability that it assumes that value.

Probability Density Function

Every discrete random variable X has an associated probability density function (pdf).

Probability Density Function

Every discrete random variable X has an associated probability density function (pdf).

The value of the pdf is defined to be zero for any value of k that is not in the range of X.

Examples

Two fair dice are rolled. The sample space (the set of possible outcomes) is a set of ordered pairs:

$$S = \{(x, y) \mid x, y \in \{1, 2, 3, 4, 5, 6\}\}$$

Examples

Two fair dice are rolled. The sample space (the set of possible outcomes) is a set of ordered pairs:

$$S = \{(x, y) \,|\, x, y \in \{1, 2, 3, 4, 5, 6\}\}$$

The 36 elements of the sample space S are:

(1,1)	(1,2)	(1,3)	(1,4)	(1,5)	(1,6)
(2,1)	(2,2)	(2,3)	(2,4)	(2,5)	(2,6)
(3,1)	(3,2)	(3,3)	(3,4)	(3,5)	(3,6)
(4,1)	(4,2)	(4,3)	(4,4)	(4,5)	(4,6)
(5,1)	(5,2)	(5,3)	(5,4)	(5,5)	(5,6)
(6,1)	(6,2)	(6,3)	(6,4)	(6,5)	(6,6)

Examples

Two fair dice are rolled. The sample space (the set of possible outcomes) is a set of ordered pairs:

$$S = \{(x,y) \,|\, x,y \in \{1,2,3,4,5,6\}\}$$

The 36 elements of the sample space S are:

(1,1)	(1,2)	(1,3)	(1,4)	(1,5)	(1,6)
(2,1)	(2,2)	(2,3)	(2,4)	(2,5)	(2,6)
(3,1)	(3,2)	(3,3)	(3,4)	(3,5)	(3,6)
(4,1)	(4,2)	(4,3)	(4,4)	(4,5)	(4,6)
(5,1)	(5,2)	(5,3)	(5,4)	(5,5)	(5,6)
(6,1)	(6,2)	(6,3)	(6,4)	(6,5)	(6,6)

There are $2^{36} = 68,719,476,736$ possible events.

Examples

The usual choice for a probability function on S is:

$$P(s) = \frac{1}{36}, \quad s \in S$$

This is equivalent to the statement the 36 possible outcomes are equally likely.

Examples

The usual choice for a probability function on S is:

$$P(s) = \frac{1}{36}, \quad s \in S$$

This is equivalent to the statement the 36 possible outcomes are equally likely.

For event E (E being an arbitrary subset of S) that contains more than one outcome, simply assign

$$P(E) = \frac{n(E)}{36}$$

where $n(E)$ is the cardinality of E.

You should convince yourself that this probability function satisfies the first three Kolmogorov axioms.

Examples

We may proceed to define a random variable X on S, the set of outcomes, by the formula

$$X \;=\; u+v \quad \text{for} \quad (u,v)\in S=\{(u,v)\,|\,u,v\in\{1,2,3,4,5,6\}\}$$

Examples

We may proceed to define a random variable X on S, the set of outcomes, by the formula

$$X \;=\; u + v \quad \text{for} \quad (u, v) \in S = \{(u, v) \,|\, u, v \in \{1, 2, 3, 4, 5, 6\}\}$$

The range of the random variable X contains 11 elements:

$$X \in \{2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\}$$

which is a considerably simpler set than either the sample space or its power set.

Examples

A probability density function $p_X(k)$ can be defined for X, with

$$p_X(k) \;=\; P(X=k), \quad k \in \{2,3,4,5,6,7,8,9,10,11,12\}$$

Examples

A probability density function $p_X(k)$ can be defined for X, with

$$p_X(k) \;=\; P(X = k), \quad k \in \{2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\}$$

By counting the number of outcomes that produce each value of k for X, we can assign specific values to $p_X(k)$:

$$p_X(k) \;=\; P(X = k) = \begin{cases} 1/36 & \text{if} \quad k = 2 \text{ or } k = 12 \\ 2/36 & \text{if} \quad k = 3 \text{ or } k = 11 \\ 3/36 & \text{if} \quad k = 4 \text{ or } k = 10 \\ 4/36 & \text{if} \quad k = 5 \text{ or } k = 9 \\ 5/36 & \text{if} \quad k = 6 \text{ or } k = 8 \\ 6/36 & \text{if} \quad k = 7 \end{cases}$$

Examples

Suppose an experiment consists of tossing a fair coin three times. The sample space contains 8 possible outcomes:

$$S = \{HHH, HHT, HTH, THH, HTT, THT, TTH, TTT\}$$

Examples

Suppose an experiment consists of tossing a fair coin three times.

The sample space contains 8 possible outcomes:

$$S = \{HHH, HHT, HTH, THH, HTT, THT, TTH, TTT\}$$

As in the previous example, it is usual to define the probability function P on S so that the outcomes are equally likely.

Then $P(s) = 1/8$ for any element s of the sample space S, and for an arbitrary subset E of the power set of S (that is, an arbitrary event E), define $P(E)$ to be the cardinality of E divided by 8:

$$P(E) = \frac{n(E)}{8}$$

Examples

Suppose an experiment consists of tossing a fair coin three times. The sample space contains 8 possible outcomes:

$$S = \{HHH, HHT, HTH, THH, HTT, THT, TTH, TTT\}$$

As in the previous example, it is usual to define the probability function P on S so that the outcomes are equally likely.

Then $P(s) = 1/8$ for any element s of the sample space S, and for an arbitrary subset E of the power set of S (that is, an arbitrary event E), define $P(E)$ to be the cardinality of E divided by 8:

$$P(E) = \frac{n(E)}{8}$$

Once again, you can verify that P satisfies the first 3 Kolmogorov axioms.

Examples

Define a random variable X on S as follows:

$$X = \text{Number of heads obtained in three tosses}$$

Examples

Define a random variable X on S as follows:

$$X = \text{Number of heads obtained in three tosses}$$

The range of X is then $\{0, 1, 2, 3\}$. By counting the outcomes with 0, 1, 2, and 3 heads and using P as defined above, we can construct a probability density function $p_X(k)$ for X:

$$p_X(k) = P(X = k) = \begin{cases} 1/8 & \text{if} \quad X = 0 \\ 3/8 & \text{if} \quad X = 1 \\ 3/8 & \text{if} \quad X = 2 \\ 1/8 & \text{if} \quad X = 3 \end{cases}$$

Examples

We will show that in general if an experiment consists of n independent Bernoulli trials with probability of success p and X is the number of successes, the probability density of X can be written as:

$$p_X(k) \;=\; P(X=k) \;=\; \binom{n}{k} p^k (1-p)^{1-k}, \quad k = 0, 1, \ldots, n$$

Examples

We will show that in general if an experiment consists of n independent Bernoulli trials with probability of success p and X is the number of successes, the probability density of X can be written as:

$$p_X(k) = P(X = k) = \binom{n}{k} p^k (1-p)^{1-k}, \quad k = 0, 1, \ldots, n$$

It is easy to verify that these probabilities add to 1 because

$$\sum_{k=0}^{n} p_X(k) = \sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{1-k}$$

is just the expansion of $(p + (1-p))^n$, and $(p + (1-p)) = 1$ so this is 1^n.