Random Variables

Andreas Klappenecker

Texas A&M University

What is a Random Variable?

Random variables are **functions** that associate a numerical value to each outcome of an experiment.

For instance, if we roll a pair of dice, then the sum of the two face values is a random variable.

Similarly, if we toss a coin three times, then the observed number of heads is a random variable.

Definition of a Random Variable

Definition

Let $\mathcal{F}$ be a σ-algebra over the sample space Ω. A **random variable** X is a function $X : \Omega \to \mathbf{R}$ such that the set

$$\{z \in \Omega \mid X(z) \leqslant x\}$$

is an event contained in $\mathcal{F}$ for all $x \in \mathbf{R}$.

For brevity, we will say that X is defined on the σ-algebra $(\Omega, \mathcal{F})$.

It should be clear from this definition that there is nothing **random** about a random variable, it is simply a (measurable) function.

Example

The definition ensures that a random variable can be used to specify events in a convenient way. There are a number of notational conventions which help to express events in an even more compact way. For instance, the event $\{z \in \Omega \,|\, X(z) \leqslant x\}$ is denoted shortly by $X \leqslant x$, an idiosyncratic but standard notation.

Example

If X is the random variable denoting the sum of the face values of a pair of dice, then $X \leqslant 3$ denotes the event $\{(1,1),(1,2),(2,1)\}$.

Example

Example

If Y is the random variable counting the number of heads in three subsequent coin tosses, then $Y \leqslant 0$ is the event $\{(\text{tail},\text{tail},\text{tail})\}$, and $Y \leqslant 1$ is the event $\{(\text{tail},\text{tail},\text{tail}), (\text{head},\text{tail},\text{tail}), (\text{tail},\text{head},\text{tail}), (\text{tail},\text{tail},\text{head})\}$.

Digression on Measurable Functions and Borel σ-Algebras

Measurable Functions

Let $(\Omega_1, \mathcal{F}_1)$ and $(\Omega_2, \mathcal{F}_2)$ be two measurable spaces, that is Ω_i is a sample space and $\mathcal{F}_i$ is a σ-algebra over Ω.

Definition

A function $f : \Omega_1 \to \Omega_2$ is called **measurable** if and only if $f^{-1}(E) \in \mathcal{F}_1$ for all events E in $\mathcal{F}_2$.

Examples of Measurable Functions

Let $(\Omega_1, \mathcal{F}_1)$ and $(\Omega_2, \mathcal{F}_2)$ be two measurable spaces.

Example

The identity function $\iota : (\Omega_1, \mathcal{F}_1) \to (\Omega_1, \mathcal{F}_1)$ with $\iota(x) = x$ is a measurable function.

Example

Any map $f : (\Omega, 2^{\Omega}) \to (\Omega_2, \mathcal{F}_2)$ is measurable.

Example

Any map $f : (\Omega_1, \mathcal{F}_1) \to (\Omega_2, \{\emptyset, \Omega_2\})$ is measurable.

Generated σ-Algebra

Proposition

Let $(\Omega_2, \mathcal{F}_2)$ *be a measurable space.*
Suppose that $X\colon \Omega_1 \to \Omega_2$ *is a map. The preimage*

$$X^{-1}(\mathcal{F}_2) = \{X^{-1}(E) \mid E \in \mathcal{F}_2\}$$

is the smallest σ*-algebra such that* X *is measurable.*

Notation

We write $\sigma(X) = \{X^{-1}(E) \mid E \in \mathcal{F}_2\}$ and call $\sigma(X)$ the σ-algebra generated by X.

Proof

We will now show that $\sigma(X) = \{X^{-1}(E) \mid E \in \mathcal{F}_2\}$ is indeed a σ-algebra.

1. Since $X^{-1}(\varnothing) = \varnothing$, we have $\varnothing \in \sigma(X)$.
2. For $E \in \mathcal{F}_2$, we have $X^{-1}(E^c) = X^{-1}(E)^c$. Thus, $\sigma(X)$ is closed under complements.
3. If $E_1, E_2, \ldots$ are events in $\mathcal{F}_2$, then

$$X^{-1}\left(\bigcup_{k=1}^{\infty} E_k\right) = \bigcup_{k=1}^{\infty} X^{-1}(E_k).$$

Thus, $\sigma(X)$ is closed under countable unions. $\square$

Examples

Example

If $X \equiv c$ is a constant function, then the preimage of a Borel set B satisfies

$$X^{-1}(B) = \begin{cases} \emptyset & \text{if } c \notin B, \\ \Omega & \text{if } c \in B. \end{cases}$$

Thus, $\sigma(X) = \{\emptyset, \Omega\}$.

Example

If the range of X is a set $\{a, b\}$ with two elements, then

$$\sigma(X) = \{\emptyset, X^{-1}(\{a\}), X^{-1}(\{b\}), \Omega\}.$$

Remark

A function X from a measurable space $(\Omega, \mathcal{F})$ to the set of real numbers $\mathbf{R}$ is called **measurable** if and only if $X^{-1}(E) \in \mathcal{F}$ for all so-called Borel sets E of $\mathbf{R}$.

We will now describe how to define the Borel σ-algebra $\mathcal{B}(\mathbf{R})$ that is comprised of the Borel sets of $\mathbf{R}$.

Open Sets of the Real Numbers

Definition

A subset S of the set $\mathbf{R}$ of real numbers is called **open** if and only if for every element x in S there exists an $\epsilon > 0$ such that

$$(x - \epsilon, x + \epsilon) \subseteq S.$$

In particular, any open interval is an open subset of $\mathbf{R}$.

Borel σ-Algebra

Let $\mathcal{O}$ denote the set of all open subsets of **R**. Thus,

$$\mathcal{O} = \{S \in P(\mathbf{R}) \mid S \text{ open}\}.$$

Definition

The **Borel σ-algebra** $\mathcal{B} = \mathcal{B}(\mathbf{R})$ of the set of real numbers is given by the smallest σ-algebra of **R** containing all open sets

$$\mathcal{B} = \sigma(\mathcal{O}).$$

The elements of $\mathcal{B}$ are called the **Borel sets** of **R**.

Open Sets

Proposition

Let S be an open subset of the real numbers. Then S is the finite or countable union of open intervals.

Proof

Proof.

For each element x in S, we define the set

$$I_x = \bigcup \{(a,b) \mid a < x < b, (a,b) \subseteq S\}.$$

Then I_x is the largest open interval in S containing x.

If $x \neq y$, then either $I_x = I_y$ or $I_x \cap I_y = \emptyset$. Indeed, if $z \in I_x \cap I_y$, then $I_x \cup I_y$ is an open interval, hence $I_x \cup I_y = I_x = I_y$.

Since each interval (a, b) contains a rational number, we have

$$S = \bigcup_{x \in S} \{I_x \mid x \in S\} = \bigcup_{x \in S \cap \mathbf{Q}} \{I_x \mid x \in S\}. \quad \square$$

Open Intervals

Let $\mathcal{O}_0$ denote the set of open subintervals of $\mathbf{R}$,

$$\mathcal{O}_0 = \{(a, b) \mid a, b \in \mathbf{R}, a < b\}.$$

Corollary

Then the Borel σ-algebra is given by

$$\sigma(\mathcal{O}_0) = \sigma(\mathcal{O}).$$

Let $\mathcal{O}_h$ denote the set of half-closed subintervals $(-\infty, a]$ given by

$$\mathcal{O}_h = \{(-\infty, a] \mid a \in \mathbf{R}\}.$$

Corollary

The Borel σ-algebra is given by

$$\sigma(\mathcal{O}_h) = \sigma(\mathcal{O}_o) = \sigma(\mathcal{O}).$$

The complement of $(-\infty, a]$ is given by (a, ∞).

Furthermore, $\bigcup_{n=1}^{\infty}(-\infty, b - \frac{1}{n}] = (-\infty, b)$.

Thus, $(a, b) = (a, \infty) \cap (-\infty, b)$ when $a < b$.

Random Variables

Definition (Alternate Version)

A random variable $X : \Omega \to \mathbf{R}$ is a measurable function from a sample space Ω to the set of real numbers $\mathbf{R}$

This definition is equivalent to the previous one.

Proposition

Let $(\Omega, \mathcal{F})$ be a measurable space. The following are equivalent:

1. $X : \Omega \to \mathbf{R}$ *is* $\mathcal{F}$*-measurable.*
2. *For all* $c \in \mathbf{R}$, $X^{-1}((c, \infty))$ *is measurable.*
3. *For all* $c \in \mathbf{R}$, $X^{-1}([c, \infty))$ *is measurable.*
4. *For all* $c \in \mathbf{R}$, $X^{-1}((-\infty, c))$ *is measurable.*
5. *For all* $c \in \mathbf{R}$, $X^{-1}((-\infty, c])$ *is measurable.*

We write these preimages in the form

$$X > c,\ X \geqslant c,\ X < c, \text{ and } X \leqslant c.$$

Proposition

Let $(\Omega, \mathcal{F})$ *be a measurable space. The following are equivalent:*

1. $X : \Omega \to \mathbf{R}$ *is* $\mathcal{F}$*-measurable.*
2. *For all* $c \in \mathbf{R}$, $X^{-1}((c, \infty))$ *is measurable.*
3. *For all* $c \in \mathbf{R}$, $X^{-1}([c, \infty))$ *is measurable.*
4. *For all* $c \in \mathbf{R}$, $X^{-1}((-\infty, c))$ *is measurable.*
5. *For all* $c \in \mathbf{R}$, $X^{-1}((-\infty, c])$ *is measurable.*

We write these preimages in the form

$$X > c,\ X \geqslant c,\ X < c, \text{ and } X \leqslant c.$$

Proposition

Let $(\Omega, \mathcal{F})$ *be a measurable space. The following are equivalent:*

1. $X : \Omega \to \mathbf{R}$ *is* $\mathcal{F}$*-measurable.*
2. *For all* $c \in \mathbf{R}$, $X^{-1}((c, \infty))$ *is measurable.*
3. *For all* $c \in \mathbf{R}$, $X^{-1}([c, \infty))$ *is measurable.*
4. *For all* $c \in \mathbf{R}$, $X^{-1}((-\infty, c))$ *is measurable.*
5. *For all* $c \in \mathbf{R}$, $X^{-1}((-\infty, c])$ *is measurable.*

We write these preimages in the form

$$X > c,\ X \geqslant c,\ X < c, \text{ and } X \leqslant c.$$

Proposition

Let $(\Omega, \mathcal{F})$ be a measurable space. The following are equivalent:

1. $X : \Omega \to \mathbf{R}$ *is* $\mathcal{F}$*-measurable.*
2. *For all* $c \in \mathbf{R}$, $X^{-1}((c, \infty))$ *is measurable.*
3. *For all* $c \in \mathbf{R}$, $X^{-1}([c, \infty))$ *is measurable.*
4. *For all* $c \in \mathbf{R}$, $X^{-1}((-\infty, c))$ *is measurable.*
5. *For all* $c \in \mathbf{R}$, $X^{-1}((-\infty, c])$ *is measurable.*

We write these preimages in the form

$$X > c,\ X \geqslant c,\ X < c,\ \text{and } X \leqslant c.$$

Proposition

Let $(\Omega, \mathcal{F})$ be a measurable space. The following are equivalent:

1. $X : \Omega \to \mathbf{R}$ *is* $\mathcal{F}$*-measurable.*
2. *For all* $c \in \mathbf{R}$, $X^{-1}((c, \infty))$ *is measurable.*
3. *For all* $c \in \mathbf{R}$, $X^{-1}([c, \infty))$ *is measurable.*
4. *For all* $c \in \mathbf{R}$, $X^{-1}((-\infty, c))$ *is measurable.*
5. *For all* $c \in \mathbf{R}$, $X^{-1}((-\infty, c])$ *is measurable.*

We write these preimages in the form

$$X > c,\ X \geqslant c,\ X < c, \text{ and } X \leqslant c.$$

Types of Random Variables

Indicator Random Variables

Let $(\Omega, \mathcal{F})$ be a measurable space.

Let A be a subset of Ω. Then the indicator function $I_A : (\Omega, \mathcal{F}) \to \mathbf{R}$ given by

$$I_A(x) = \begin{cases} 1 & \text{if } x \in A \\ 0 & \text{otherwise.} \end{cases}$$

is a random variable if and only if $A \in \mathcal{F}$. We call I_A the **indicator random variable** of the event A.

Simple Random Variable

A random variable is called **simple** if and only if it is a linear combination of a finite number of indicator random variables with disjoint support.

In other words, if X is a simple random variable, then there exist pairwise disjoint events $A_1, \ldots, A_n$ and real numbers $s_1, \ldots, s_n$ such that

$$X = \sum_{k=1}^{n} s_k I_{A_k}.$$

Any nonnegative random variable can be approximated by a sequence of simple random variables.

Discrete Random Variable

A **discrete random variable** is a random variable with countable range, which means that the set $\{X(z) \mid z \in \Omega\}$ is countable.

The convenience of a discrete random variable X is that one can define events in terms of values of X, for instance in the form $X \in A$ which is short for

$$\{z \in \Omega \mid X(z) \in A\}.$$

If the set A is a singleton, $A = \{x\}$, then we write $X = x$.

Probability Distributions

Probability Distribution

Let $(\Omega, \mathcal{F}, \Pr)$ denote a probability space. Suppose that $X\colon \Omega \to \mathbf{R}$ is a random variable on the probability space.

For each Borel set B of $\mathcal{B}(\mathbf{R})$, we define

$$\Pr_X(B) = \Pr[X^{-1}(B)].$$

Then $\Pr_X$ is a probability measure on the Borel σ-algebra $\mathcal{B}(\mathbf{R})$. We call $\Pr_X$ the **probability distribution** of the random variable X.

We note that $(\mathbf{R}, \mathcal{B}(\mathbf{R}), \Pr_X)$ is a probability space.

Probability Distribution

The probability distribution of a random variable X is determined by the values

$$\Pr[X \leqslant x]$$

for all $x \in \mathbf{R}$. This is simply another way of writing

$$\Pr[X \leqslant x] = \Pr[X^{-1}((-\infty, x])].$$

Densities

Let X be a discrete random variable defined on a σ-algebra $(\Omega, \mathcal{F})$. Let Pr be a probability measure on $\mathcal{F}$. The **density function** p_X of a discrete random variable X is defined by

$$p_X(x) = \Pr[X = x].$$

The density function describes the probabilities of the events $X = x$.

Note that the density function is sometimes also called the **probability mass function**.

Example

Let $(\Omega, 2^{\Omega}, \Pr)$ be the probability space of a pair of fair dice, that is, the sample space $\Omega = \{1,2,3,4,5,6\} \times \{1,2,3,4,5,6\}$, and Pr is the uniform probability measure, $\Pr[A] = |A|/36$ for any subset A of Ω. Let X denote the random variable denoting the sum of the face values of the two dice. The density function and the distribution function of X are tabulated below:

x	2	3	4	5	6	7	8	9	10	11	12
$\Pr[X = x]$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$
$\Pr[X \leqslant x]$	$\frac{1}{36}$	$\frac{3}{36}$	$\frac{6}{36}$	$\frac{10}{36}$	$\frac{15}{36}$	$\frac{21}{36}$	$\frac{26}{36}$	$\frac{30}{36}$	$\frac{33}{36}$	$\frac{35}{36}$	$\frac{36}{36}$