Math 425
Introduction to Probability
Lecture 19

Kenneth Harris
kaharri@umich.edu

Department of Mathematics
University of Michigan

March 2, 2009

Discrete Random Variable

☞ A real-valued function $p : \mathbb{R} \to \mathbb{R}$ is a probability mass function if there is an enumeration of real numbers $A_p = \{a_1, a_2, a_3, \ldots\}$ satisfying the following

$$\begin{array}{lll} (a) & p(a) \geq 0 & \text{for all } a, \\ (b) & p(a) = 0 & \text{if } a \notin A_p, \\ (c) & \sum_{a \in A_p} p(a) = 1 & \end{array}$$

☞ Intuitively, $p(a)$ gives the “probability" of a occurring.

Probability density functions

☞ A probability mass function “defines" a probability on $\mathbb{R}$.

Theorem

Let $p : \mathbb{R} \to \mathbb{R}$ be a probability mass function which assigns nonzero values only to the reals in the set A_p.
Then the function $\mathbf{P}$ on $\mathbb{R}$ defined on each subset E by

$$\mathbf{P}(E) = \sum_{a \in E \cap A_p} p(a)$$

is a probability function: for each subset E and F

$$\begin{array}{lll} (a) & \mathbf{P}(E) \geq 0 & \\ (b) & \mathbf{P}(\mathbb{R}) = 1 & \\ (c) & \mathbf{P}(E \cup F) = \mathbf{P}(E) + \mathbf{P}(F) & \text{when } E \cap F = \emptyset. \end{array}$$

Note: $\mathbf{P}$ is well-defined for any subset E of $\mathbb{R}$.

Discrete Random Variable

☞ Let S be a sample space with probability function $\mathbf{P}$.
A random variable $X : S \to \mathbb{R}$ is said to be discrete if there is a probability mass function p_X with

$$p_X(a) = \mathbf{P}\{X = a\}.$$

☞ For any subset E of $\mathbb{R}$,

$$\mathbf{P}\{X \in E\} = \sum_{a \in E,\ p_X(a) > 0} p_X(a)$$

Example

☞ There are problems for which discrete random variables are not appropriate.

Example. Consider the experiment of choosing a number in the interval $[0,1]$ “at random".
Let $S = [0,1]$ (any real number in $[0,1]$ can be an outcome).
Let X be the value selected (a continuous random variable).

Problem. Each number $a \in [0,1]$ is equally likely to be selected, so we must have

$$\mathbf{P}\{X = a\} = 0 \quad \text{but} \quad \mathbf{P}\{0 \le X \le 1\} = 1.$$

Basic Events

☞ In the discrete case, the probabilities of the basic events, $\mathbf{P}\{X = a\}$, were the building blocks for determining $\mathbf{P}\{X \in E\}$ for all subsets E of S.

☞ When $S \subseteq E$ we will take intervals $[a,b]$ to be our basic events. We will use the probabilities

$$\mathbf{P}\{a \le X \le b\} \qquad \text{where } a, b \in \mathbb{R} \cup \{-\infty, +\infty\}.$$

to determine the probability $\mathbf{P}\{X \in E\}$ of all other events E.

Density function

Definition

A function $f : \mathbb{R} \to \mathbb{R}$ is integrable if $\int_a^b f(x)\,dx$ exists for all $a, b \in \mathbb{R}$.

An integrable function $f : \mathbb{R} \to \mathbb{R}$ is called a probability density function if it satisfies the following two conditions:

$$(a) \qquad f(t) \ge 0 \qquad \text{for all } t \in \mathbb{R}$$

$$(b) \qquad \int_{-\infty}^{\infty} f(t)\,dt = 1.$$

Probability density functions

☞ A probability density function “defines" a probability on $\mathbb{R}$.

Theorem

If $f : \mathbb{R} \to \mathbb{R}$ is a probability density function, then

$$(a) \qquad \int_a^b f(t)\,dt \ge 0 \qquad \text{for all } a, b \in \mathbb{R} \cup \{-\infty, +\infty\}$$

$$(b) \qquad \int_{-\infty}^{\infty} f(t)\,dt = 1$$

$$(c) \qquad \int_{E \cup F} f(t)\,dt = \int_E f(t)\,dt + \int_F f(t)\,dt \qquad \text{when } E \cap F = \emptyset.$$

☠ *Warning. (c) holds provided the integrals have a value.*

Continuous Probability Space

Definition

A sample space $S \subseteq \mathbb{R}$ with probability function $\mathbf{P}$ is called a continuous sample space if there is a probability density function f such that

$$\mathbf{P}\{[a,b] \cap S\} = \int_a^b f(t)\,dt \qquad \text{for all } a, b \in \mathbb{R} \cup \{-\infty, +\infty\}.$$

An event is any subset $E \subset S$ such that $\int_E f(t)\,dt$ has a value.

Warning. It is NOT true that $\int_E f(x)\,dx$ has a value for every subset E of $\mathbb{R}$. We will only consider subsets which do.

Density function

Definition

Let S be a continuous sample space and $X : S \to \mathbb{R}$.

We say X is a continuous random variable if there is a probability density function f_X such that

$$\mathbf{P}\{a \le X \le b\} = \int_a^b f_X(t)\,dt \qquad \text{for all } a, b \in \mathbb{R} \cup \{-\infty, +\infty\}.$$

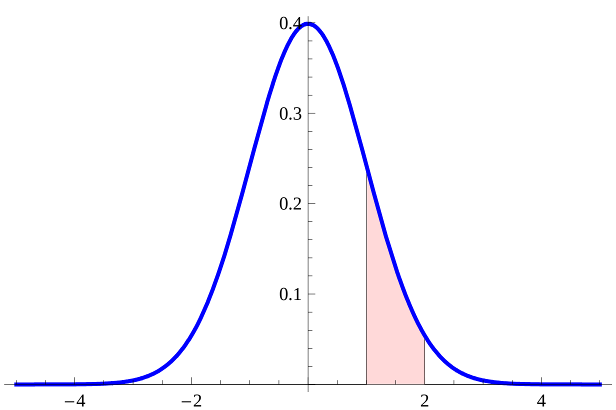
If $E \subseteq \mathbb{R}$ is an event, then

$$\mathbf{P}\{X \in E\} = \int_E f_X(t)\,dt.$$

Probability density function

☞ The probability of $\mathbf{P}\{X \in E\} = \int_E f(t)\,dt$.

The probability density function is $f(t) = \frac{1}{\sqrt{2\pi}} e^{-t^2}$, a normal density, the event is $E = [1, 2]$, and $\mathbf{P}\{1 \le X \le 2\} \approx 0.1359$.



Probability of basic events

☞ Let X be a continuous random variable. Then,

$$\begin{aligned}\int_a^b f(t)\,dt &= \mathbf{P}\{a \le X \le b\} = \mathbf{P}\{a < X \le b\} \\ &= \mathbf{P}\{a \le X < b\} = \mathbf{P}\{a < X < b\}\end{aligned}$$

☞ More generally, if E and F differ by a finite number of elements, then

$$\mathbf{P}\{X \in F\} = \int_E f(t)\,dt = \int_F f(t)\,dt = \mathbf{P}\{X \in F\}.$$

This is still true if E and F differ by a countable number of elements.

Discrete vs. Continuous random variables

☞ The probability function of a discrete random variable X on sample space S is determined by a probability mass function p_X.

- $p_X(a)$ is a probability (the likelihood of t).
- The probability of basic events $\mathbf{P}\{X = a\}$ determine all probabilities.
- All subsets E of S have a well determined probability $\mathbf{P}\{X \in E\}$.

☞ The probability function of a continuous random variable X on sample space S is determined by a probability density function f_X.

- $f_X(t)$ is NOT a probability, only the integral $\int_E f_X(t)\,dt$ is.
- The probability of basic events $\mathbf{P}\{a \le X \le b\}$ determine all probabilities. In fact, $\mathbf{P}\{X = a\} = 0 = \int_a^a f(t)\,dt = 0$.
- Not all subsets E of S have a well determined probability $\mathbf{P}\{X \in E\}$. Fortunately, all "reasonable sets" do.

Example

☞ Choose a number "at random" in the interval $[0, 1]$.
What are the probabilities of the basic events $\mathbf{P}\{a \le X \le b\}$?

☞ Any real number is "equally likely" to be chosen.
So, any subinterval of $[0, 1]$ will have the same probability as any other subinterval of $[0, 1]$ of the same length.

$$\mathbf{P}\{0 \le X \le \frac{1}{2}\} = \mathbf{P}\{\frac{1}{2} \le X \le 1\}$$
$$\mathbf{P}\{0 \le X \le \frac{1}{3}\} = \mathbf{P}\{\frac{2}{3} \le X \le 1\}$$

Thus, the length of the subinterval determines its probability.

Probabilities of Basic Events

☞ Since $\mathbf{P}$ is a probability function, we must have

$$\mathbf{P}\{a \le X \le b\} = b - a \qquad \text{when } 0 \le a \le b \le 1$$

More generally, for any a and b (real or infinite)

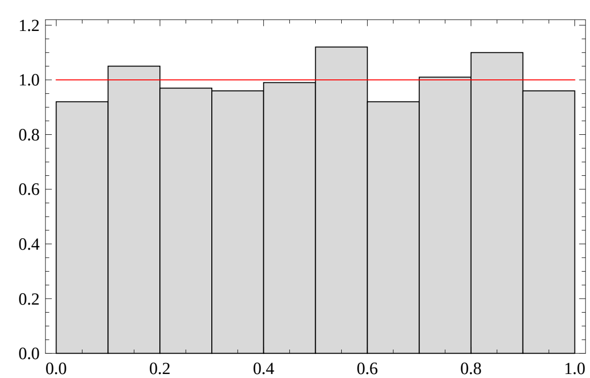
$$\mathbf{P}\{a \le X \le b\} = \mathbf{P}\left(\{a \le X \le b\} \cap [0, 1]\right).$$

Examples.

$$\mathbf{P}\{-\infty < X < \infty\} = \mathbf{P}\{0 \le X \le 1\} = 1$$
$$\mathbf{P}\{\frac{1}{2} \le X \le 27\} = \mathbf{P}\{\frac{1}{2} \le X \le 1\} = \frac{1}{2}$$
$$\mathbf{P}\{-1 \le X \le \frac{1}{3}\} = \mathbf{P}\{0 \le X \le \frac{1}{3}\} = \frac{1}{3}$$

Probabilities of Basic Events

☞ A plot of 1000 randomly chosen real numbers in $[0, 1]$ broken into intervals of length 0.1. The area of each rectangle gives the proportion of numbers in the interval. The red line is $f(t) = 1$, the area under f approximates the probability.



Probabilities of Arbitrary Events

☞ The probability of basic events can be computed using the integral with probability density function:

$$\mathbf{P}\{a \le X \le b\} = \int_a^b f(t)\,dt \qquad f(t) = \begin{cases} 1 & \text{if } 0 \le t \le 1 \\ 0 & \text{otherwise} \end{cases}$$

☞ The probability of an arbitrary event $E \subseteq \mathbb{R}$ is

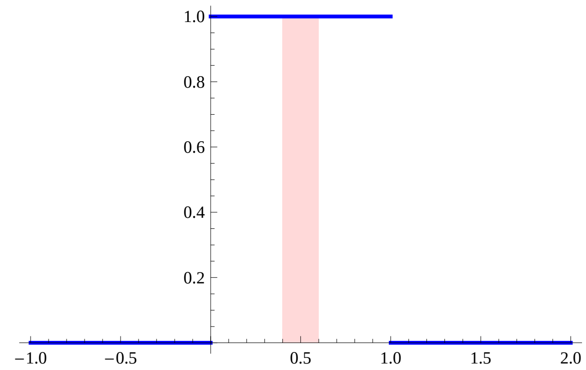
$$\mathbf{P}\{X \in E\} = \int_E f(t)\,dt$$

And when $E \subseteq [0, 1]$,

$$\mathbf{P}\{X \in E\} = \int_E dt$$

Probabilities of Basic Events

☞ The probability of $\mathbf{P}\{X \in E\}$ is $\int_E f(t)\,dt$, the area under f along E. A graph of $f(t)$ with $E = [0.4, 0.6]$.



Uniform Random Variable

Definition

We say X is a uniform random variable on an interval I with real-valued endpoints $\alpha < \beta$, when its probability density function is given by

$$f(t) = \begin{cases} \frac{1}{\beta-\alpha} & \text{if } \alpha \le t \le \beta \\ 0 & \text{otherwise} \end{cases}$$

☞ The random variable giving the outcome of the experiment which selects a real number in the interval $I = [\alpha, \beta]$ "at random" is a uniform random variable. It does not matter if we exclude one or both endpoints.

Uniform Random Variable

☞ Verify.

$$\mathbf{P}\{\alpha \le X \le \beta\} = \int_\alpha^\beta \frac{1}{\beta - \alpha}\,dt = 1.$$

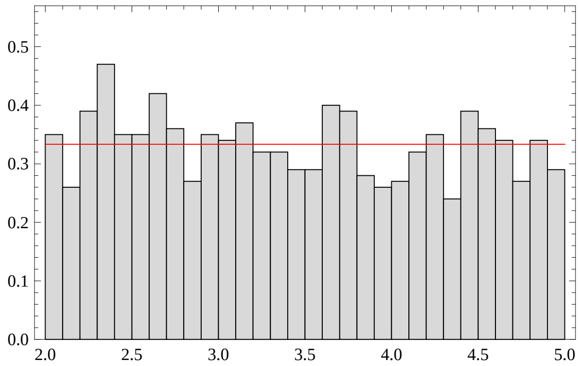
☞ If $I_1, I_2 \subseteq [\alpha, \beta]$ are intervals of the same length ℓ, then

$$\int_{I_1} \frac{1}{\beta-\alpha}\,dt = \frac{\ell}{\beta-\alpha} = \int_{I_2} \frac{1}{\beta-\alpha}\,dt$$

Computer Verification

☞ X picks a number at random in the interval $[2, 5]$. 1000 points were chosen; the area of each rectangle is the proportion of points in the corresponding interval of length 0.1.

The red line is the p.d.f for X: $f(t) = \frac{1}{3}$ for $t \in [2, 5]$.



Example

Example. A real number is chosen at random from $[0, 1]$ with uniform probability, and then this number is squared. X is the r.v. which represents this is result.

What is the density of X?

☞ Let U be the uniform random variable for this experiment. So,

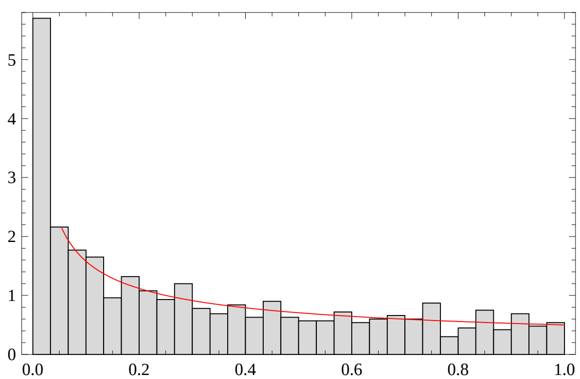
$$f_U(t) = \begin{cases} 1 & \text{if } 0 \leq t \leq 1 \\ 0 & \text{otherwise.} \end{cases}$$

Furthermore, $X = U^2$. Can we determine f_X from f_U?

Example – continued

☞ A plot of 1000 randomly chosen and squared numbers in the interval $[0, 1]$ and partitioned into 30 equal subintervals.

The red line is $f(t) = \frac{1}{2\sqrt{t}}$, and is very close to giving the area of the rectangles in each subinterval. Note that $f(t)$ blows-up at 0.



Cumulative Distribution Function

☞ There is another kind of function, closely related to the density function, which is of great importance when considering continuous random variables.

Definition

Let X be a continuous real-valued random variable. The cumulative distribution function of X is defined by the equation

$$F_X(a) = \mathbf{P}\{X \leq a\} \qquad \text{for all } a \in \mathbb{R}.$$

Example

Example. Let X be a uniform random variable on the interval $[0, 2\pi)$. What is the cumulative distribution for X?

☞ The density f_X is given by

$$f_X(t) = \begin{cases} \frac{1}{2\pi} & \text{if } 0 \le t < 2\pi \\ 0 & \text{otherwise} \end{cases}$$

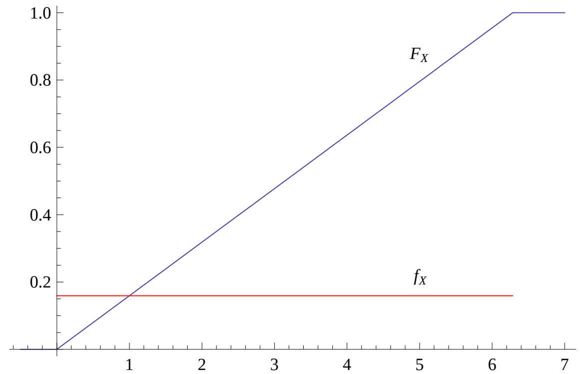
The cumulative distribution F_X is given by

$$F_X(a) = \int_{-\infty}^{a} \frac{1}{2\pi}\, dt = \begin{cases} 0 & \text{if } a < 0 \\ \frac{a}{2\pi} & \text{if } 0 \le a \le 2\pi \\ 1 & \text{otherwise} \end{cases}$$

Example – continued

☞ The cumulative distribution and density for a uniform random variable on $[0, 2\pi)$.

Note that F_X is continuous, although f_X is not.



Theorem

☞ The density function f_X and cumulative distribution function F_X of a continuous random variable X are related in a very nice way. The following is just the statement of the Second Fundamental Theorem of Calculus.

Theorem

Suppose X is a continuous random variable with density $f_X(t)$ which is continuous, except perhaps finitely many points.

The cumulative distribution function F_X is given by

$$F_X(a) = \int_{-\infty}^{a} f(t)\, dt \qquad \text{for each } a \in \mathbb{R}$$

It is continuous and is related to f_X by

$$\frac{d}{dt} F_X(t) = f_X(t).$$

Theorem

☞ This gives a converse of the previous theorem. Given the cumulative distribution function F_X, you can recover the density function f_X. It is really just a restatement of the First Fundamental Theorem of Calculus.

Theorem

Suppose X is a continuous random variable with cumulative distribution function F_X.

The density function f_X is given by

$$\frac{d}{dt} F_X(t) = f_X(t).$$

and is related to F_X by

$$F_X(a) = \int_{-\infty}^{a} f(t)\, dt \qquad \text{for each } a \in \mathbb{R}$$

Example – return

☞ Sometimes it is easier to specify the cumulative distribution function, then it is the density function.

Example – return. A real number is chosen at random from $[0, 1]$ with uniform probability, and then this number is squared. X is the r.v. which represents this is result.

What is the cumulative distribution of X? What is the density of X?

Example – continued

☞ Let U represent the chosen real number; so, $X = U^2$.
For $0 \leq t \leq 1$,

$$\begin{aligned} F_X(a) &= \mathbf{P}\{X \leq a\} \\ &= \mathbf{P}\{U^2 \leq a\} \\ &= \mathbf{P}\{U \leq \sqrt{a}\} \\ &= \int_0^{\sqrt{a}} dt = \sqrt{a}. \end{aligned}$$

Since $0 \leq X \leq 1$, the cumulative distribution function for X is

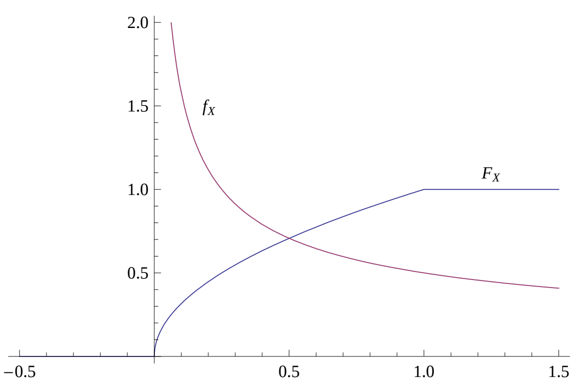
$$F_X(t) = \begin{cases} 0 & \text{if } t \leq 0, \\ \sqrt{t} & \text{if } 0 \leq t \leq 1, \\ 1 & \text{if } t \geq 1. \end{cases}$$

The density function of X is

$$f_X(t) = \begin{cases} \frac{1}{2\sqrt{t}} & \text{if } 0 \leq t \leq 1, \\ 0 & \text{otherwise.} \end{cases}$$

Example – continued

☞ The cumulative distribution $F_X(t) = \sqrt{t}$ and density function $f_X(t) = \frac{1}{2\sqrt{t}}$ on $[0, 1]$. Note that F_X is continuous, although f_X is not.



All-Star Baseball

Example

When I was a kid I used to play *All-Star Baseball*. Each player was a disc-shaped card that was placed in a spinner. Along the circumference were regions denoting baseball events (see next slide). If the pointer of the spinner stopped in a region, you read-off the relevant event.

☞ The area of arc of a region corresponds to the percent of time the player performed that event in their at-bats. For example Babe Ruth hit a homerun about 7% of the time, so his homerun arc (region **1** on the disc) was about 7% arc, or 0.44 radians (25 degrees).

All-Star Baseball

Sample Space

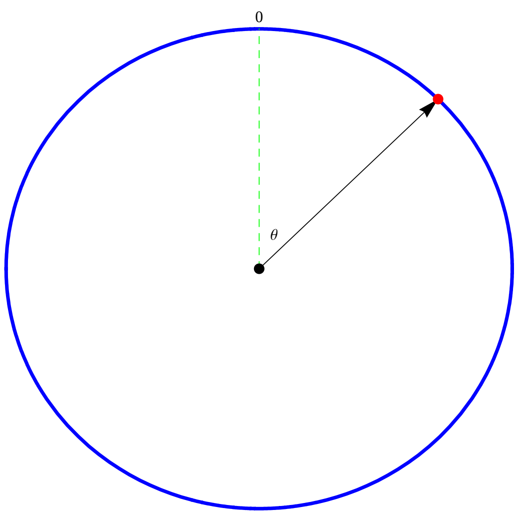
☞ Assume the spinner randomly points to a point on the circumference. Fix twelve noon to be 0, and measure the angle clockwise from noon in radians. This angle uniquely determines the point picked-out by the spinner.

☞ The sample space S is the interval $[0, 2\pi)$. Let X be the uniform random variable providing the outcome of the spinner. Then, the p.d.f for X is

$$f(t) = \begin{cases} \frac{1}{2\pi} & \text{if } 0 \leq t < 2\pi \\ 0 & \text{otherwise} \end{cases}$$

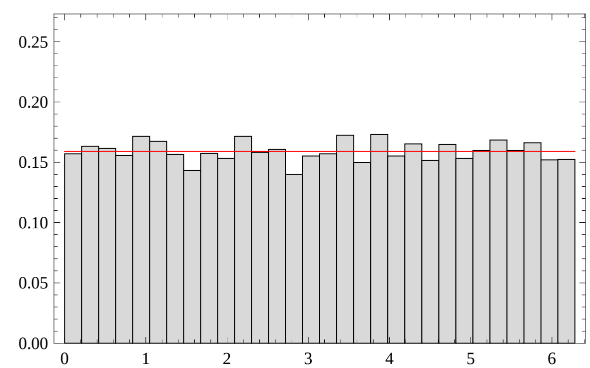
Spinner Sample Space

☞ The angle θ radians is measured from high noon to pointer, clockwise.



Computer Verification

☞ Distribution of values in the interval $[0, 2\pi)$ in a simulation of 10,460 “Ruthian at bats". There are 30 intervals of equal length. The red line is at $f(t) = \frac{1}{2\pi} \approx 0.159$.



Computer Verification

☞ I simulated Babe Ruth's career 10,460 at bats using the measured arcs on the Babe Ruth All-Star Card. (The amount of arc determined the probability of each event: homerun, double, strikeout, etc.)

	HR	Doubles	Triples	Strikeouts	Walks	Batting Average
Babe Ruth	714	506	136	1330	2062	.342
Simulation	734	493	127	1308	2028	.343

☞ Pretty darn accurate game. However, when I was younger I was a master spinner, and could make the spinner stop nearly at will.