Computational Foundations of Cognitive Science

Lecture 20: Discrete and Contiuous Random Variables; Distributions and Densities

Frank Keller

School of Informatics
University of Edinburgh
keller@inf.ed.ac.uk

February 23, 2010

Reading: Freund, Chs. 3.1–3.4.

Discrete Random Variables

Definition: Random Variable

If S is a sample space with a probability measure and X is a real-valued function defined over the elements of S, then X is called a random variable.

We will denote random variable by capital letters (e.g., X), and their values by lower-case letters (e.g., x).

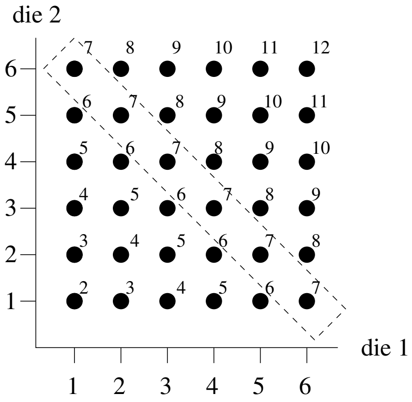
Example

Given an experiment in which we roll a pair of dice, let the random variable X be the total number of points rolled with the two dice.

For example $X = 7$ picks out the set $\{(1,6),(2,5),(3,4),(4,3),(5,2),(6,1)\}$.

Discrete Random Variables

This can be illustrated graphically:



For each outcome, this graph lists the value of X, and the dotted area corresponds to $X = 7$.

Discrete Random Variables

Example

Assume a balanced coin is flipped three times. Let X be the random variable denoting the total number of heads obtained.

Outcome	Probability	x
HHH	$\frac{1}{8}$	3
HHT	$\frac{1}{8}$	2
HTH	$\frac{1}{8}$	2
THH	$\frac{1}{8}$	2

Outcome	Probability	x
TTH	$\frac{1}{8}$	1
THT	$\frac{1}{8}$	1
HTT	$\frac{1}{8}$	1
TTT	$\frac{1}{8}$	0

Hence, $P(X = 0) = \frac{1}{8}$, $P(X = 1) = P(X = 2) = \frac{3}{8}$, $P(X = 3) = \frac{1}{8}$.

Probability Distributions

Definition: Probability Distribution

If X is a discrete random variable, the function given by $f(x) = P(X = x)$ for each x within the range of X is called the probability distribution of X.

Theorem: Probability Distribution

A function can serve as the probability distribution of a discrete random variable X if and only if its values, $f(x)$, satisfy the conditions:

1. $f(x) \geq 0$ for each value within its domain;
2. $\sum_x f(x) = 1$, where x over all the values within its domain.

Probability Distributions

Example

For the probability function defined in the previous example:

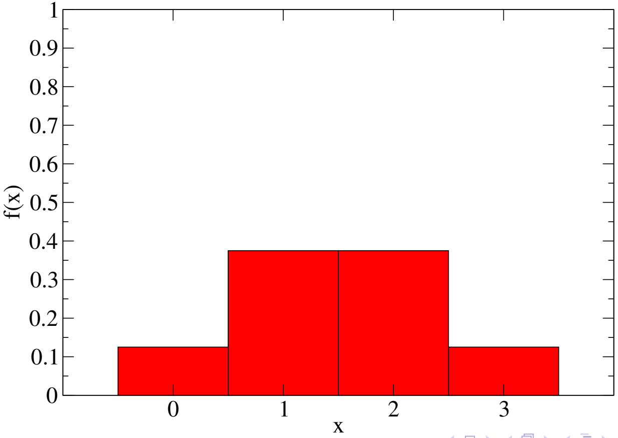
x	$f(x) = P(X = x)$
0	$\frac{1}{8}$
1	$\frac{3}{8}$
2	$\frac{3}{8}$
3	$\frac{1}{8}$

This function can be written more concisely as:

$$f(x) = \frac{4 - |3 - 2x|}{8}$$

Probability Distributions

A probability distribution is often represented as a *probability histogram*. For the previous example:



Mid-lecture Problem

A balanced coin is tossed four times. Define a random variable X that indicates the number of heads obtained. Find a formula for the probability distribution of X.

Now generalize this formula to n coin tosses.

Cumulative Distribution

In many cases, we're interested in the probability for values $X \leq x$, rather than for $X = x$.

Definition: Cumulative Distribution

If X is a discrete random variable, the function given by:

$$F(x) = P(X \leq x) = \sum_{t \leq x} f(t) \text{ for } -\infty < x < \infty$$

where $f(t)$ is the value of the probability distribution of X at t, is the cumulative distribution of X.

Cumulative Distributions

Example

Consider the probability distribution $f(x) = \frac{4-|3-2x|}{8}$ from the previous example. The values of the cumulative distribution are:

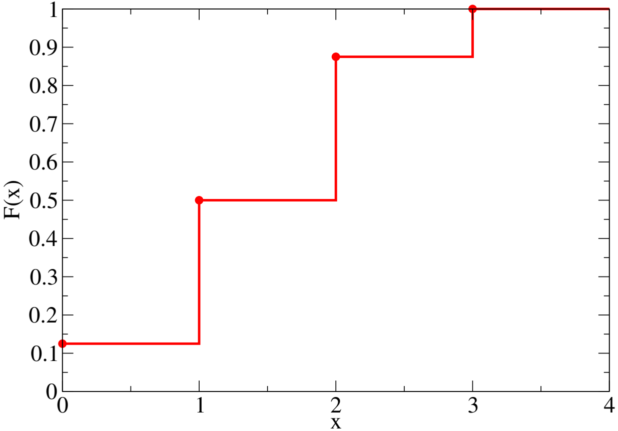
x	$f(x)$	$F(x)$
0	$\frac{1}{8}$	$\frac{1}{8}$
1	$\frac{3}{8}$	$\frac{4}{8}$
2	$\frac{3}{8}$	$\frac{7}{8}$
3	$\frac{1}{8}$	$\frac{8}{8}$

Note that $F(x)$ is defined for all real values of x:

$$F(x) = \begin{cases} 0 & \text{for } x < 0 \\ \frac{1}{8} & \text{for } 0 \leq x < 1 \\ \frac{4}{8} & \text{for } 1 \leq x < 2 \\ \frac{7}{8} & \text{for } 2 \leq x < 3 \\ \frac{8}{8} & \text{for } x \geq 3 \end{cases}$$

Cumulative Distributions

The cumulative distribution can be graphed; for the previous example:



Cumulative Distributions

Theorem: Cumulative Distributions

The values $F(x)$ of the cumulative distribution of a discrete random variable X satisfies the conditions:

1. $F(-\infty) = 0$ and $F(\infty) = 1$;
2. if $a < b$, then $F(a) \leq F(b)$ for any real numbers a and b.

Example

Consider the example of $F(x)$ on the previous slide:

1. $F(-\infty) = 0$ as $F(0) < 0$ by definition; $F(\infty) = 1$ as $F(x) = 1$ for $x \geq 3$ by definition;
2. $F(a) < F(b)$ holds for $(0, 1)$, $(1, 2)$, $(2, 3)$ by definition; $F(a) = F(b)$ holds for all other values of a and b.

Continuous Random Variables

We only dealt with discrete (integer-valued) random variables. In many situations, continuous (real-valued) random variables occur.

Examples

The outcomes of real-life experiments are often continuous:

- An experimental subject reacts to a picture by pressing a button (e.g., to indicate if the picture is familiar): the reaction time (in ms) is a continuous random variable.
- An EEG machine measures the electrical brain activity when a subjects reads a word: the current (in μV) is a continuous random variable.

Definition of probability distribution and cumulative distribution can be extended to the continuous case.

Probability Density Functions

Extend definitions from discrete to continuous random variables:

- use intervals $a \leq X \leq b$ instead of discrete values $X = x$;
- use integration over intervals instead of summation over discrete values.

Definition: Probability Density Function

A function with values $f(x)$, defined over the set of all real numbers, is called a probability density function (pdf) of the continuous random variable X if and only if:

$$P(a \leq X \leq b) = \int_a^b f(x)dx$$

for any real constants a and b with $a \leq b$.

Probability Density Functions

Example

Assume a continuous random variable X with the pdf:

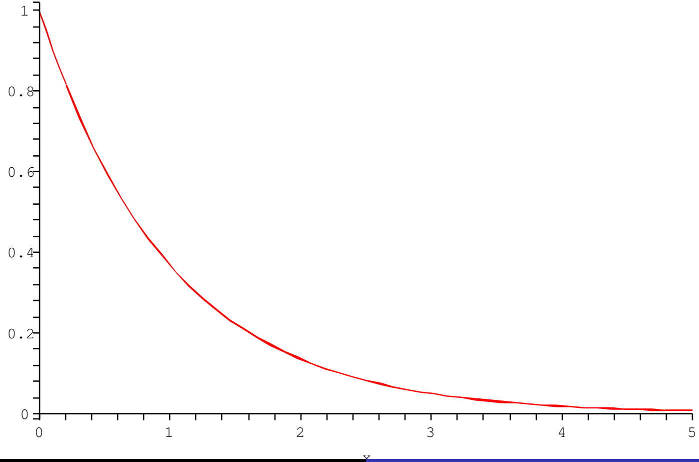
$$f(x) = \begin{cases} e^{-x} & \text{for } x > 0 \\ 0 & \text{elsewhere} \end{cases}$$

Compute the probability for the interval $0 \leq X \leq 1$:

$$\begin{aligned} P(a \leq X \leq b) &= \int_a^b f(x)dx = \int_0^1 e^{-x}dx = -e^{-x}\Big|_0^1 \\ &= (-e^{-1}) - (-e^0) = -\frac{1}{e} + 1 = 0.63 \end{aligned}$$

Probability Density Functions

Plot the function on the previous slide:



Probability Density Functions

Theorem: Intervals of pdfs

If X is a continuous random variable and a and b are real constants with $a \leq b$, then:

$$P(a \leq X \leq b) = P(a \leq X < b) = P(a < X \leq b) = P(a < X < b)$$

Theorem: Valid pdfs

A function can serve as the pdf of a continuous random variable X if its values, $f(x)$, satisfy the conditions:

1. $f(x) \geq 0$ for each value within its domain;
2. $\int_{-\infty}^{\infty} f(x)dx = 1$.

Probability Density Functions

Example

Assume a random variable X with the pdf $f(x)$ as follows. Is this a valid pdf?

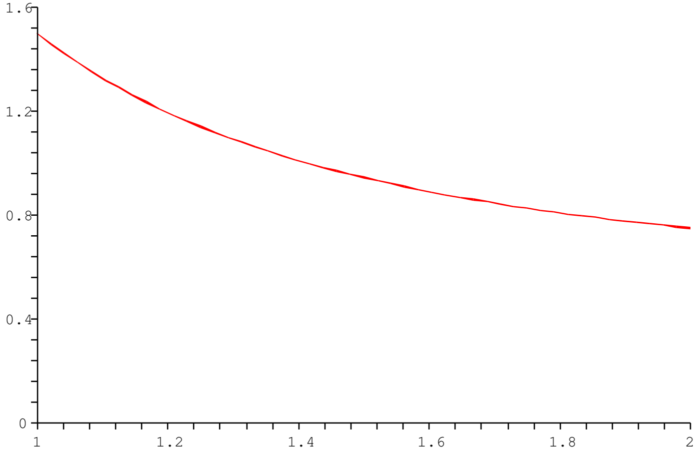
$$f(x) = \begin{cases} \frac{1}{x^2} + \frac{1}{2} & \text{for } 1 < x \leq 2 \\ 0 & \text{elsewhere} \end{cases}$$

$f(x) \geq 0$ is true by definition. To show $\int_{-\infty}^{\infty} f(x)dx = 1$, integrate:

$$\begin{aligned} \int_{-\infty}^{\infty} f(x)dx &= \int_{1}^{2} \frac{1}{x^2} + \frac{1}{2}dx = -\frac{1}{x} + \frac{1}{2}x \Big|_{1}^{2} \\ &= (-\frac{1}{2} + \frac{1}{2} \cdot 2) - (-\frac{1}{1} + \frac{1}{2} \cdot 1) = 1 \end{aligned}$$

Probability Density Functions

Plot the function on the previous slide:



Cumulative Distributions

In analogy with the discrete case, we can define:

Definition: Cumulative Distribution

If X is a continuous random variable and the value of its probability density function at t is $f(t)$, then the function given by:

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(t)dt \text{ for } -\infty < x < \infty$$

is the cumulative distribution of X.

Intuitively, the cumulative distribution captures the area under the curve defined by $f(t)$ from $-\infty$ to x.

Cumulative Distributions

Example

Assume a continuous random variable X with the pdf:

$$f(t) = \begin{cases} e^{-t} & \text{for } t > 0 \\ 0 & \text{elsewhere} \end{cases}$$

Integrate for $t > 0$:

$$\begin{aligned} F(x) = P(X \leq x) &= \int_{-\infty}^{x} f(t)dt = \int_{0}^{x} e^{-t}dt = -e^{-t}\Big|_0^x \\ &= (-e^{-x}) - (-e^{0}) = -e^{-x} + 1 \end{aligned}$$

Therefore the cumulative distribution of X is:

$$F(x) = \begin{cases} -e^{-x} + 1 & \text{for } x > 0 \\ 0 & \text{elsewhere} \end{cases}$$

Cumulative Distributions

Theorem: Value of Cumulative Distribution

If $f(x)$ and $F(x)$ are the values of the pdf and the cumulative distribution function of X at x, then:

$$P(a \leq X \leq b) = F(b) - F(a)$$

for any real constants a and b with $a \leq b$ and:

$$f(x) = \frac{dF(x)}{dx}$$

where the derivative exists.

Cumulative Distributions

Example

Use the theorem on the previous slide to compute the probability $P(0 \leq X \leq 1)$ for $f(t)$:

$$P(0 \leq X \leq 1) = F(1) - F(0) = (-e^{-1}) - (-e^{-0}) = -\frac{1}{e} + 1 = 0.63$$

Also, verify the derivative of $F(x)$:

$$\frac{dF(x)}{dx} = \frac{d(-e^{-x})}{dx} = e^{-x}$$

Summary

- A random variable picks out a subset of the sample space;
- a probability distribution returns a probability for each value of a random variable;
- a cumulative distribution sums all the values of a probability up to a threshold;
- probability density functions are the probability distributions for continuous random variables;
- cumulative distributions can also be defined for continuous random variables.