MATH 556: MATHEMATICAL STATISTICS I

PROBABILITY PRIMER

8. RANDOM VARIABLES & PROBABILITY MODELS

A **random variable (r.v.)** X is a function defined on a space $(\Omega, \mathcal{F})$ that associates a *real number* $X(\omega) = x$ with each "event" $A \in \mathcal{F}$. We regard X as a (possibly many-to-one) mapping from Ω to $\mathbb{R}$

$$X: \quad \begin{array}{c} \mathcal{F} \longrightarrow \mathbb{R} \\ A \longmapsto x \end{array}$$

We note the equivalence of events in $\mathcal{F}$ and their images under X. Consider the space $(\mathbb{R}, \mathcal{B})$, that is the real line $\mathbb{R}$, and the sigma-algebra $\mathcal{B}$ of sets that are open, half-open or closed subsets of the reals, that is, sets of the form

$$(a, b) \qquad (a, b] \qquad [a, b) \qquad [a, b]$$

and countable unions and intersections of these sets. Consider $B \in \mathcal{B}$, and consider the pre-image

$$B^{-1} = \{\omega \in \Omega : X(\omega) \in B\}$$

For X to be a valid random variable, we require that $B^{-1} \in \mathcal{F}$. Strictly, when referring to random variables, we should make explicit the connection to original sample space Ω and write $X(\omega)$ for individual sample outcomes, and

$$P_X[X \in B] = \mathrm{P}(\{\omega : X(\omega) \in B\})$$

for events. However, generally, we will suppress this and merely refer to X.

In deducing the probability measure for X, P_X, we must have

$$\mathrm{P}(B^{-1}) = P_X[X \in B].$$

for all $B \in \mathcal{B}$. The notation $\mathrm{P}(\cdot)$ will be retained to refer to probabilities associated with events in Ω, whereas $P_X[\,\cdot\,]$ will be used when referring to events in $\mathbb{R}$.

Probability Functions: Consider the real function of a real argument, F_X, defined by

$$F_X(x) = P_X((-\infty, x]) = \int_{(-\infty, x]} P_X(dx)$$

for real values of x. Note that $X \in (-\infty, x]$ is equivalent to $X \leq x$, and, by definition,

$$F_X(x) = \int_{-\infty}^{x} dF_X(t).$$

F_X defines the **probability distribution** of X. The nature of F_X determines how we can manipulate this function. There are (essentially) three cases to consider:

1. F_X is a *step-function*, that is, F_X changes only at a certain (countable) set of x values.

2. F_X is a *continuous* function.

3. F_X is a mixture of 1. and 2.

8.1. Discrete Random Variables

A random variable X is **discrete** if F_X is a step-function, that is, if the set of all values at which F_X changes, denoted $\mathbb{X}$, is **countable**

- $\mathbb{X} \equiv \{x_1, x_2, \ldots, x_n\}$ (that is, a **finite** list)
- $\mathbb{X} \equiv \{x_1, x_2, \ldots\}$ (that is, a countably **infinite** list)

where, without loss of generality it is assumed that

$$x_1 < x_2 < \cdots$$

If X is discrete, then it follows the probability of event $(X \in B)$ can be decomposed

$$\mathrm{P}(\{\omega : X(\omega) \in B\}) = P_X[X \in \{x_i : x_i \in B\}] = \sum_{i:x_i \in B} P_X(X = x_i)$$

so that F_X can be represented as a sum of probabilities

$$F_X(x) = \sum_{x_i \leq x} P_X(X \in \{x_i\}) = \sum_{x_i \leq x} P_X[X = x_i].$$

8.1.1. Probability Mass Function

The function f_X, defined on $\mathbb{X}$ by

$$f_X(x) = P_X[X = x] \qquad x \in \mathbb{X}$$

that assigns probability to each $x \in \mathbb{X}$ is the (discrete) **probability mass function**, or **pmf**. For completeness, we define

$$f_X(x) = 0 \qquad x \notin \mathbb{X}$$

so that f_X is defined for all $x \in \mathbb{R}$. Thus $\mathbb{X}$ is the *support* of random variable X (or the pmf f_X): $\mathbb{X}$ is the set of $x \in \mathbb{R}$ such that $f_X(x) > 0$.

Properties: A function f_X is a probability mass function for discrete random variable X with support $\mathbb{X}$ of the form $\{x_1, x_2, \ldots\}$ if and only if

$$\text{(i) } f_X(x_i) \geq 0 \qquad\qquad \text{(ii) } \sum_i f_X(x_i) = 1$$

Clearly as $f_X(x) = P_X[X = x]$, we must have $0 \leq f_X(x) \leq 1$ for all $x \in \mathbb{R}$.

8.1.2. Discrete Cumulative Distribution Function

The **cumulative distribution function**, or **cdf**, F_X of a discrete r.v. X is defined by

$$F_X(x) = P_X[X \leq x] \qquad x \in \mathbb{R}.$$

Connection between F_X and f_X: Let X be a discrete random variable with support $\mathbb{X} \equiv \{x_1, x_2, \ldots\}$, pmf f_X and cdf F_X. For any real value x, if $x < x_1$, then $F_X(x) = 0$, and for $x \geq x_1$,

$$F_X(x) = \sum_{x_i \leq x} f_X(x_i)$$

so that, for $i = 2, 3, \ldots$,

$$f_X(x_i) = F_X(x_i) - F_X(x_{i-1})$$

with, for completeness, $f_X(x_1) = F_X(x_1)$.

Properties:

(i) In the limiting cases,

$$\lim_{x\to-\infty} F_X(x) = 0 \qquad\qquad \lim_{x\to\infty} F_X(x) = 1.$$

(ii) F_X is **continuous from the right** (but not continuous) on $\mathbb{R}$ that is, for $x \in \mathbb{R}$,

$$\lim_{h\to 0^+} F_X(x+h) = F_X(x)$$

but, if $x \in \mathbb{X}$,

$$\lim_{h\to 0^-} F_X(x+h) \neq F_X(x)$$

that is, the "left limit" is not equal to the "right limit" at x values in $\mathbb{X}$.

(iii) F_X is **non-decreasing**, that is

$$a < b \Longrightarrow F_X(a) \leq F_X(b)$$

(iv) For $a < b$,

$$P_X[a < X \leq b] = F_X(b) - F_X(a)$$

Notes:

- The functions f_X and/or F_X can both be used to describe the probability distribution of random variable X.
- The function f_X is non-zero only at the elements of $\mathbb{X}$.
- The function F_X is a **step-function**, which takes the value zero at minus infinity, the value one at infinity, and is non-decreasing with points of discontinuity at the elements of $\mathbb{X}$.
- The right-continuity of F_X is denoted in plots by the use of a filled circle, $\bullet$, as in the example below.
- In the discrete case, F_X is **not differentiable** for all $x \in \mathbb{R}$; at points of continuity (that is, for $x \notin \mathbb{X}$), it is differentiable, and the derivative is zero.

Example 1 Consider a coin tossing experiment where a fair coin is tossed repeatedly under identical experimental conditions, with the sequence of tosses independent, until a Head is obtained. For this experiment, the sample space, Ω is then the set of sequences

$$(\{H\}, \{TH\}, \{TTH\}, \{TTTH\} \ldots)$$

with associated probabilities 1/2, 1/4, 1/8, 1/16, $\ldots$.

Define discrete random variable X by $X(\omega) = x \Longleftrightarrow$ first H on toss x. Then

$$f_X(x) = P_X[X = x] = \left(\frac{1}{2}\right)^x \qquad x = 1, 2, 3, \ldots$$

and zero otherwise. For $x \geq 1$, let $\lfloor x \rfloor$ be the largest integer not greater than x. Then

$$F_X(x) = \sum_{x_i \leq x} f_X(x_i) = \sum_{i=1}^{\lfloor x \rfloor} f_X(i) = 1 - \left(\frac{1}{2}\right)^{\lfloor x \rfloor}$$

and $F_X(x) = 0$ for $x < 1$.

Graphs of the probability mass function (top) and cumulative distribution function (bottom) are shown in Figure 1. Note that the mass function is only non-zero at points that are elements of X, and that the cdf is defined for all real values of x, but is only continuous from the right. F_X is therefore a step-function.

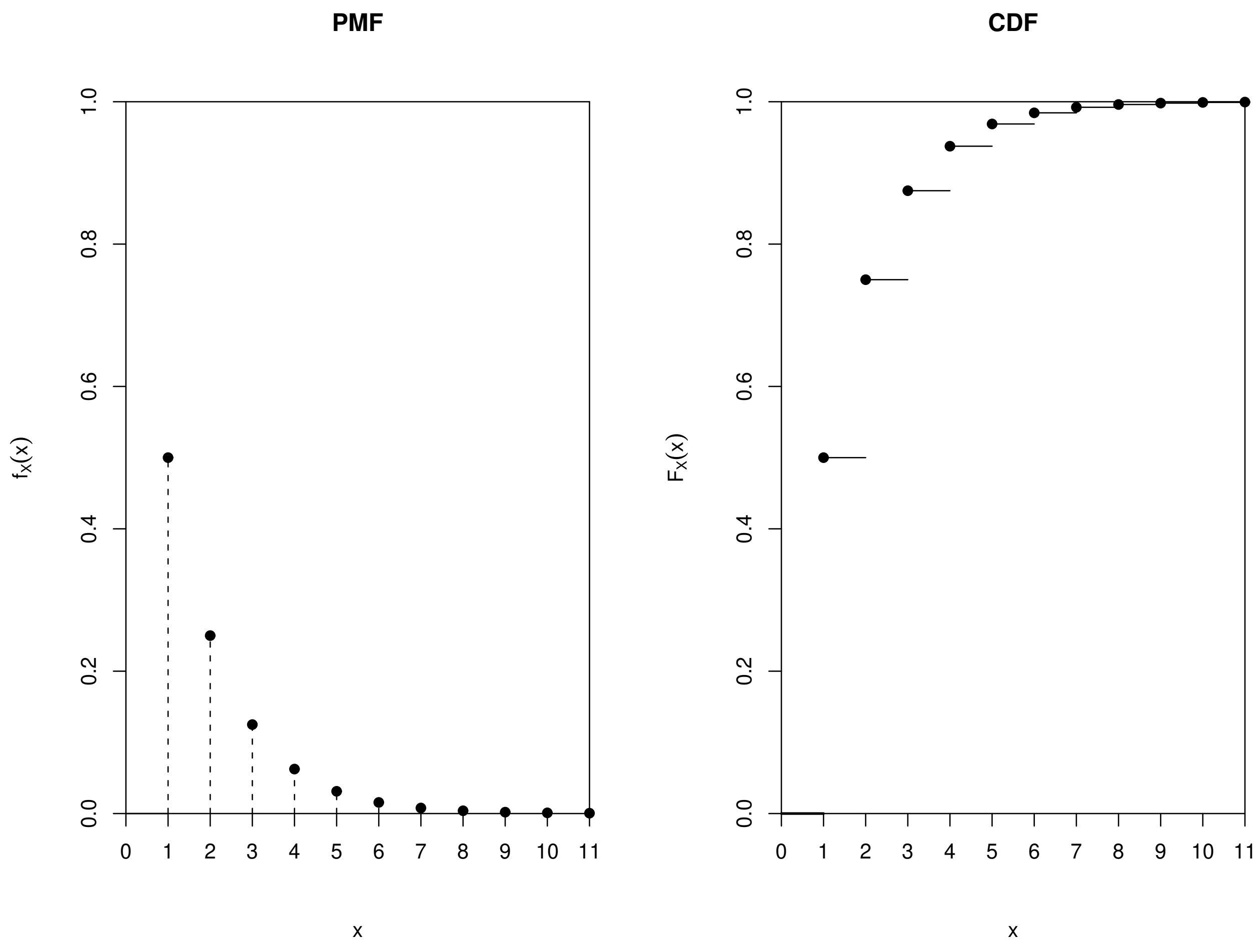

Figure 1: PMF $f_X(x) = \left(\frac{1}{2}\right)^x, x = 1, 2, 3, \ldots$ and CDF $F_X(x) = 1 - \left(\frac{1}{2}\right)^{\lfloor x \rfloor}$

8.2. Continuous Random Variables

A random variable X is termed **continuous** if the function F_X defined on $\mathbb{R}$ by

$$F_X(x) = P_X[X \leq x]$$

for $x \in \mathbb{R}$ is a **continuous** function on $\mathbb{R}$, that is, for $x \in \mathbb{R}$,

$$\lim_{h \to 0} F_X(x + h) = F_X(x).$$

8.2.1. Continuous Cumulative Distribution Function

The **cumulative distribution function**, or **cdf**, F_X of a continuous r.v. X is defined by

$$F_X(x) = P_X[X \leq x] \qquad x \in \mathbb{R}.$$

8.2.2. Probability Density Function

A cumulative distribution function F_X is **absolutely continuous** if it can be written

$$F_X(x) = \int_{-\infty}^{x} f_X(t)\, dt$$

for some function f_X, termed the **probability density function**, or **pdf**, of X. For any suitable set B,

$$P_X[X \in B] = \int_B f_X(x)\, dx.$$

Directly from the definition, at values of x where F_X is differentiable x,

$$f_X(x) = \frac{d}{dt}\left\{F_X(t)\right\}_{t=x}.$$

It follows that a function f_X is a pdf for a continuous random variable X **if and only if**

$$\text{(i) } f_X(x) \geq 0 \qquad\qquad \text{(ii) } \int_{-\infty}^{\infty} f_X(x)\, dx = 1$$

This result follows direct from definitions and properties of F_X. Note that in the continuous case, there is no requirement that f_X is bounded above; it can be defined to be zero at an arbitrary set of points, or defined piecewise on intervals of $\mathbb{R}$.

Properties:

(i) If X is continuous, $f_X(x) \neq P_X[X = x]$, as

$$P_X[X = x] = \lim_{h \to 0} [F_X(x+h) - F_X(x)] = 0$$

as F_X is continuous.

(ii) For the cdf of a continuous r.v.,

$$\lim_{x \to -\infty} F_X(x) = 0 \qquad \lim_{x \to \infty} F_X(x) = 1$$

(iii) For $a < b$,

$$P_X[a < X \leq b] = P_X[a \leq X < b] = P_X[a \leq X \leq b] = P_X[a < X < b] = F_X(b) - F_X(a)$$

Example 2 Consider an experiment to measure the length of time that an electrical component functions before failure. The sample space of outcomes of the experiment, Ω is $^+$, and if A_x is the event that the component functions for longer than $x > 0$ time units, suppose that

$$P(A_x) = \exp\left\{-x^2\right\}.$$

Define random variable X by $X(\omega) = x \iff$ component fails at time $x \in \mathbb{R}$. Then, $f_X(x) = 0$ for $x \leq 0$, and for $x > 0$,

$$F_X(x) = P_X[X \leq x] = 1 - P_X(A_x) = 1 - \exp\left\{-x^2\right\}$$

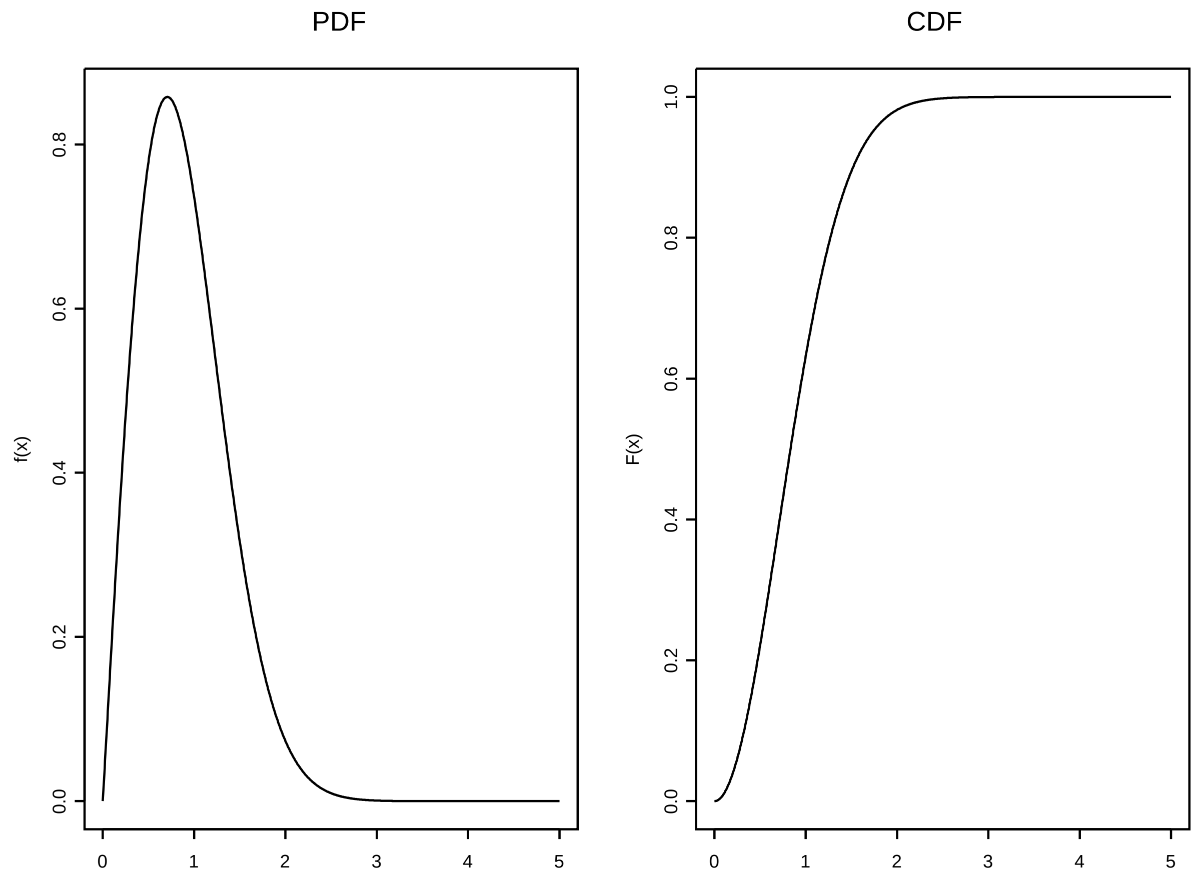

Figure 2: PDF $f_X(x) = 2x \exp\left\{-x^2\right\},\ x > 0$, and CDF $F_X(x) = 1 - \exp\left\{-x^2\right\},\ x > 0$

and $F_X(x) = 0$ if $x \leq 0$. Hence if $x > 0$,

$$f_X(x) = \frac{d}{dt}\left\{F_X(t)\right\}_{t=x} = 2x \exp\left\{-x^2\right\}.$$

Graphs of the probability density function (top) and cumulative distribution function (bottom) are shown in Figure 2. Note that both the pdf and cdf are defined for all real values of x, and that both are continuous functions.

$$F_X(x) = \int_{-\infty}^{x} f_X(t)dt = \int_{0}^{x} f_X(t)\, dt$$

as $f_X(x) = 0$ for $x \leq 0$, and also that

$$\int_{-\infty}^{\infty} f_X(x)\, dx = \int_{0}^{\infty} f_X(x)\, dx = 1.$$