## I. Random Variables as Functions

Recall that random variables are functions from the sample space $\Omega$ to the set of all real numbers $\mathbb{R}$ that satisfies certain properties. Thus a random variable X is a function of the form $\mathrm{X}: \Omega \rightarrow \mathbb{R}$, and for each $\omega \in \Omega$, the random variable takes the value $X(\omega)$. Until now, we have only discussed discrete random variables. Discrete random variables are the random variables that can be expressed as $X: \Omega \rightarrow S$ for a countable (i.e., finite or countably infinite) subset $S$ of $\mathbb{R}$. In the following, we first define the continuous random variables relying on your knowledge of calculus. Then we present the general definition of a random variable, which subsumes both discrete and continuous random variables, as well as the random variables that are neither discrete nor continuous.

## A. Continuous Random Variables and PDF

A function $\mathrm{X}: \Omega \rightarrow \mathbb{R}$ defines a continuous random variable for a probability model iff there exists a non-negative function $f_{\mathrm{X}}$ that satisfies the following identity for every $x \in \mathbb{R}$

$$
\begin{equation*}
\mathbf{P}(\mathrm{X} \leq x)=\int_{-\infty}^{x} f_{\mathrm{X}}(\tau) \mathrm{d} \tau \tag{1}
\end{equation*}
$$

The non-negative function $f_{\mathrm{X}}$ is called the Probability Density Function (PDF) of X . The integral used to state the the condition given in (1), is the Riemann integral, which you have learned in calculus.

Remark 1. There are not just one but multiple ways to define the integral formally; the main difference between them is not in terms of the values we get as a result but in terms of their applicability. Certain integrals are defined for some definitions of integral and not for other definitions. Throughout our discussion of continuous random variables, we will be dealing with piece-wise continuous PDFs for which these subtleties do not arise.

Remark 2. The condition given in (1) for being a continuous random variable is not just about the function $X: \Omega \rightarrow \mathbb{R}$; it is also about the probability law and hence the probability model. The dependence on the probability model was not explicit in the discrete random variables, but it was present nonetheless. Recall that we have defined the probability mass function $p_{Y}$ of a discrete random variable Y as $p_{\mathrm{Y}}(y):=\mathbf{P}(\mathrm{Y}=y)$ for all $y \in \mathbb{R}$. Thus the definition of PMF implicitly assumes that the set $\{\omega \mid \mathrm{Y}(\omega)=y\}$ is an event in the probability model for all $y \in \mathbb{R}$. The same assumption needs to be spelled out in the definition of discrete random variables in order to be rigorous.

Remark 3. The condition given in (1) implies together with the axioms of probability that

$$
\begin{equation*}
\mathbf{P}(\mathrm{X} \in \mathrm{~B})=\int_{\mathrm{B}} f_{\mathrm{X}}(\tau) \mathrm{d} \tau \tag{2}
\end{equation*}
$$

for any B that can be obtained by countable unions, intersections, and complements of open intervals. Evidently, if (2) holds for all such B then (1) holds for all $x \in \mathbb{R}$ as well. Thus the condition given in (2) can be used in the definition of continuous random variables. Our textbook uses this alternative -but equivalent- definition, without explicitly referring to the condition of being expressible via countable unions, intersections, and complements of open intervals.

The PDF plays a role for continuous random variables that is analogous to the role played by the PMF for the discrete random variables.

- For any continuous random variable X , its $\mathrm{PDF} f_{\mathrm{X}}$ is non-negative, i.e. $f_{\mathrm{X}}(\tau) \geq 0$ for all $\tau \in \mathbb{R}$. This property is similar to the non-negativity of PMF for discrete random variables.
- However, the value of the PDF $f_{\mathrm{X}}(\tau)$ at a particular $\tau$ is NOT equal to $\mathbf{P}(\mathrm{X}=\tau)$, in fact $f_{\mathrm{X}}(\tau)$ is not the probability of any particular event. Hence, it can take values greater than one as well.
- The PDF can be interpreted as the probability per unit length, hence the use of the term density. Furthermore, at the points where $f_{\mathrm{X}}$ is continuous we have $\mathbf{P}(\mathrm{X} \in[x, x+\delta]) \approx f_{\mathrm{X}}(x) \delta$ for small $\delta$. In fact at the points where $f_{\mathrm{X}}$ is continuous this relation can be stated formally as

$$
\begin{equation*}
\lim _{\delta \downarrow 0} \frac{\mathbf{P}(\mathrm{X} \in[x, x+\delta])}{\delta}=f_{\mathbf{X}}(x) \tag{3}
\end{equation*}
$$

However, at the points where $f_{\mathrm{X}}$ is not continuous we can not make either of the assertions.
Example 1. Let $f_{\mathrm{X}}$ be the uniform PDF on $[0,1]$, i.e.

$$
f_{\mathrm{X}}(x)= \begin{cases}1 & x \in[0,1] \\ 0 & x \notin[0,1]\end{cases}
$$

Note that $\mathbf{P}(\mathrm{X} \in[1,1+\delta])=0$ for all $\delta>0$, but $f_{\mathrm{X}}(1)=1$.

- For any continuous random variable X , integral of its PDF over the real line is equal to one, i.e. $\int_{-\infty}^{\infty} f_{\mathrm{X}}(\tau) \mathrm{d} \tau=1$. This property is similar to the sum of the PMF of a discrete random variable over all possible values of that discrete random varible being equal to 1 .


## B. General Random Variables and CDF

For both continuous and discrete random variables $X$ the subset $\{\omega: X(\omega) \leq \tau\}$ of the sample space $\Omega$ is an event for every $\tau \in \mathbb{R}$. Thus we can calculate the probability $\mathbf{P}(\mathrm{X} \leq \tau)$ for every $\tau \in \mathbb{R}$ both when X is a continuous random variable and and when $X$ is a discrete random variable. In particular,

$$
\mathbf{P}(\mathrm{X} \leq \tau)= \begin{cases}\sum_{k \leq x} p_{\mathrm{X}}(k) & \text { if } \mathrm{X} \text { is a discrete random variable } \\ \int_{-\infty}^{x} f_{\mathrm{X}}(\tau) \mathrm{d} \tau . & \text { if } \mathrm{X} \text { is a continuous random variable }\end{cases}
$$

in fact this property can be taken as the definition of a random variable. A function $\mathrm{X}: \Omega \rightarrow \mathbb{R}$ defines a random variable for a probability model iff there exists a function $F_{\mathrm{X}}$ that satisfies the following identity for every $x \in \mathbb{R}$

$$
\begin{equation*}
F_{\mathbf{X}}(x)=\mathbf{P}(\mathbf{X} \leq x) \tag{4}
\end{equation*}
$$

The function $F_{\mathrm{X}}$ is called the Cumulative Distribution Function (PDF) of X .
The CDF $F_{X}$ of any random variable $X$ satisfies the following properties as a result of the the definition of CDF and the properties of probability laws.

- $F_{\mathrm{X}}$ is monotonically nondecreasing, i.e., if $x \leq y$ then $F_{\mathrm{X}}(x) \leq F_{\mathrm{X}}(y)$.
- $F_{\mathrm{X}}(x)$ converges to zero as $x$ converges to $-\infty$, i.e., $\lim _{x \downarrow-\infty}=0$.
- $F_{\mathrm{X}}(x)$ converges to one as $x$ converges to $\infty$, i.e., $\lim _{x \uparrow \infty}=1$.

If $X$ is a continuous random variable then as a result of (1) and (4) we have

$$
\begin{equation*}
F_{\mathrm{X}}(x)=\int_{-\infty}^{x} f_{\mathrm{X}}(\tau) \mathrm{d} \tau \tag{5}
\end{equation*}
$$

Thus using the fundamental theorem of calculus we can deduce that

$$
\begin{equation*}
f_{\mathrm{X}}(x)=\frac{\mathrm{d} F_{\mathrm{X}}}{\mathrm{~d} x}(x) \quad \text { at } x \text { values at which } f_{\mathrm{X}} \text { is continuous. } \tag{6}
\end{equation*}
$$

## II. Recommended Questions and Solutions of Select Questions

- Recommended Questions For Section 3.1: 1, 2, 3
- Recommended Questions For Section 3.2: 5, 6, 7, 8, 9

Solutions of the Problems 5 and 7 are presented in the following pages. If you try to solve these problems yourself first and read the solution afterward, you would make most out of these solutions.

Solution of Problem 5.


Figure 1.

The point chosen uniformly from the points of the triangle can be interpreted as the outcome of the experiment $\omega$. The distance of $\omega$ to the base of the triangle is the random variable $X$. One such triangle whose height is $h$ and a particular outcome $\omega$ and the resultant $\mathrm{X}(\omega)$ is given in Figure 1
Since the line segments $D E$ and $A B$ parallel, triangles $A B C$ and $D E C$ are similar triangles. Thus the length of the line segments $D E$ and $A B$ are related to the height of the triangle and the value of random variable $X$ as follows:

$$
\frac{|D E|}{|A B|}=\frac{h-\mathrm{X}}{h} .
$$

Thus $\mathrm{X} \leq \tau$ if and only if $|D E| \geq|A B|\left[1-\frac{\tau}{h}\right]$ for all $\tau \in[0, h]$. Hence

$$
\mathbf{P}(\mathrm{X} \leq \tau)=\mathbf{P}\left(\frac{|D E|}{|A B|} \geq 1-\frac{\tau}{h}\right) \quad \forall \tau \in[0, h]
$$

Since we have a uniform probability law, for any $\tau \in[0, h]$ the probability of the event $\left\{\frac{|D E|}{|A B|} \geq 1-\frac{\tau}{h}\right\}$ is equal to the ratio of the area of the trapezoid $A B \widetilde{D} \widetilde{E}$ to the area of the triangle $A B C$, where $|\widetilde{D} \widetilde{E}|=|A B|\left(1-\frac{\tau}{h}\right)$. Thus

$$
\begin{array}{rlr}
\mathbf{P}(\mathrm{X} \leq \tau) & =\frac{\text { area of } A B \widetilde{D} \widetilde{E}}{\text { area of } A B C} & \\
& =\frac{\frac{|A B|+|\widetilde{D} \widetilde{E}|}{2} \tau}{\frac{|A B|}{2} h} & \forall \tau \in[0, h]
\end{array}
$$

Invoking the identity $|\widetilde{D} \widetilde{E}|=|A B|\left(1-\frac{\tau}{h}\right)$ and using the definition of CDF we get

$$
F_{\mathrm{X}}(\tau)= \begin{cases}0 & \tau<0 \\ \left(2-\frac{\tau}{h}\right) \frac{\tau}{h} & \tau \in[0, h] \\ 1 & \tau>h\end{cases}
$$

Using (6), we can determine the value of $f_{\mathrm{X}}(\tau)$ for all $\tau \in \mathbb{R} \backslash\{0\}$, i.e. for real values of $\tau$ other than 0 . Consequently we get

$$
f_{\mathrm{X}}(\tau)= \begin{cases}\frac{2}{h}\left(1-\frac{\tau}{h}\right) & \tau \in[0, h] \\ 0 & \tau \notin[0, h]\end{cases}
$$

Solution of Problem 7. (a) Since we have a uniform law on the disk we can find the CDF of the random variable X at $\tau$ using the ratio of the disk whose radius is $\tau$ to the area of the disk whose radius is $r$. Thus


Figure 2.

$$
F_{\mathrm{X}}(\tau)= \begin{cases}0 & \tau<0 \\ \frac{\tau^{2}}{r^{2}} & \tau \in[0, r] \\ 1 & \tau>r\end{cases}
$$

Using (6), we can determine the value of PDF of X at $\tau$, i.e. $f_{\mathrm{X}}(\tau)$, for all $\tau \in \mathbb{R} \backslash\{r\}$, i.e. for real values of $\tau$ other than $r$. Consequently we get

$$
f_{\mathrm{X}}(\tau)= \begin{cases}\frac{2 \tau}{r^{2}} & \tau \in[0, r]  \tag{7}\\ 0 & \tau \notin[0, r]\end{cases}
$$

Thus the expected value of X is given by

$$
\begin{aligned}
\mathbf{E}[\mathrm{X}] & =\int_{0}^{r} x \frac{2 x}{r^{2}} \mathrm{~d} x \\
& =\frac{2 r}{3} .
\end{aligned}
$$

The second moment of $X$ can be found in a similar way,

$$
\begin{aligned}
\mathbf{E}\left[\mathrm{X}^{2}\right] & =\int_{0}^{r} x^{2} \frac{2 x}{r^{2}} \mathrm{~d} x \\
& =\frac{r^{2}}{2} .
\end{aligned}
$$

Then the variance of X can be found using the identity $\operatorname{var}(\mathrm{X})=\mathbf{E}\left[\mathrm{X}^{2}\right]-\mathbf{E}[\mathrm{X}]^{2}$ :

$$
\operatorname{var}(\mathrm{X})=\frac{r^{2}}{18}
$$

(b) We will use the following general principle in order to solve this problem: If the random variable Y can be expressed as a function of the random variable X , i.e. if $\mathrm{Y}=f(\mathrm{X})$ for some function $f(\cdot)$, then for any $\mathrm{B} \subset \mathbb{R}$ the event $\{\mathrm{Y} \in \mathrm{B}\}$ can be expressed as the event $\left\{\mathrm{X} \in f^{-1}(\mathrm{~B})\right\}$, i.e.

$$
\{\omega \mid \mathrm{Y}(\omega) \in \mathrm{B}\}=\left\{\omega \mid \mathrm{X}(\omega) \in f^{-1}(\mathrm{~B})\right\} \quad \forall \mathrm{B}
$$

where

$$
f^{-1}(\mathrm{~B}):=\{\tau \in \mathbb{R} \mid f(\tau) \in \mathrm{B}\}
$$

Consequently

$$
\mathbf{P}(\mathrm{Y} \in \mathrm{~B})=\mathbf{P}\left(\mathrm{X} \in f^{-1}(\mathrm{~B})\right)
$$

In order to determine $f^{-1}(\mathrm{~B})$ in terms of smaller $f^{-1}(\cdot)$ for smaller sets we use the following two properties:

- $f^{-1}\left(\mathrm{~B}_{1} \cup \mathrm{~B}_{2}\right)=f^{-1}\left(\mathrm{~B}_{1}\right) \cup f^{-1}\left(\mathrm{~B}_{2}\right)$ for any $\mathrm{B}_{1}$ and $\mathrm{B}_{2}$.
- $f^{-1}\left(\mathrm{~B}_{1} \cap \mathrm{~B}_{2}\right)=\emptyset$ provided that $\mathrm{B}_{1} \cap \mathrm{~B}_{2}=\emptyset$.

Note that the random variable $S$ can be expressed as a function of the random variable X as $\mathrm{S}=g(\mathrm{X})$ where $g: \mathbb{R} \geq 0 \rightarrow \mathbb{R}$

$$
g(x):= \begin{cases}\frac{1}{x} & x \in(0, t]  \tag{8}\\ 0 & x \in(t, \infty) \text { or } x=0\end{cases}
$$

Thus

$$
\begin{array}{rlrlrl}
\{\omega \mid \mathrm{S}(\omega)<\tau\} & =\emptyset & \forall \tau \leq 0 & & \text { because } g^{-1}((-\infty, \tau))=\emptyset \text { for all } \tau \leq 0, & \text { (9a) } \\
\{\omega \mid \mathrm{S}(\omega)=0\} & =\{\omega \mid \mathrm{X}(\omega)=0\} \cup\{\omega \mid \mathrm{X}(\omega) \in(t, \infty)\} & & & \text { because } g^{-1}(\{0\})=\{0\} \cup(t, \infty), \\
\{\omega \mid \mathrm{S}(\omega) \in(0, \tau)\} & =\emptyset & & \forall \tau \in\left(0, \frac{1}{t}\right) & & \text { because } g^{-1}((0, \tau))=\emptyset \text { for all } \tau \in\left(0, \frac{1}{t}\right),
\end{array}
$$

On the other hand we can calculate the probabilities of the events of the form $\{\mathrm{X} \in \mathrm{B}\}$ using $\mathbf{P}(\mathrm{X} \in \mathrm{B})=\int_{\mathrm{B}} f_{\mathrm{X}}(x) \mathrm{d} x$ and the expression for the CDF of $X$ given in (7):

$$
\begin{align*}
\mathbf{P}(\mathrm{X}=0) & =0  \tag{10a}\\
\mathbf{P}(\mathrm{X} \in(t, \infty)) & =1-\frac{t^{2}}{r^{2}}  \tag{10b}\\
\mathbf{P}(\mathbf{X}=t) & =0  \tag{10c}\\
\mathbf{P}\left(\mathrm{X} \in\left(0, \frac{1}{\tau}\right)\right) & =\frac{1}{r^{2} \tau^{2}} \tag{10d}
\end{align*}
$$

Using (9) and (10) together with $F_{\mathrm{S}}(\tau)=\mathbf{P}(\mathrm{S} \leq \tau)$ we get

$$
F_{\mathrm{S}}(\tau)= \begin{cases}0 & \tau<0 \\ 1-\frac{t^{2}}{r^{2}} & \tau \in\left[0, \frac{1}{t}\right] \\ 1-\frac{1}{r^{2} \tau^{2}} & \tau>\frac{1}{t}\end{cases}
$$

Note that $S$ is not a continuous random variable because the condition given in (1) can not be met because of the discontinuity of its CDF $F_{\mathrm{S}}$ at $0 . F_{\mathrm{S}}$ is plotted in Figure 3.


Remark 4. Note that the event $\{X=0\}$ has zero probability, i.e., $\mathbf{P}(X=0)=0$, thus the value of the function $g(\cdot)$ defined in (8) at zero is immaterial for the cumulative distribution function $F_{\mathrm{S}}$, i.e. we get the same cumulative distribution function $F_{\mathrm{S}}$ no matter how we define $g(0)$.

Figure 3.

