Indian Institute of Technology, Kanpur
Probability Theory, 18–19

SCRIBE

4

Instructors: Neeraj Misra
Authors: Gurpreet Singh
Date: Decemenber 1, 2017

Transformation of Random Variables

1. Discrete Random Variables

Suppose we have a discrete random variable X, we need to find a random variable Y such that $Y = h(X)$, where h is a function $h : \mathbb{R} \to \mathbb{R}$.

Theorem 4.1. Consider two random variables X and Y such that $Y = h(X)$. We can say that the set $h(S_X) = \{h(x) \mid x \in S_X\}$ is the support for the random variable Y *i.e.* $S_Y = h(S_X)$.

Proof. Let $y \in h(S_X)$ *i.e.* there is some x^** such that $h(x^*) = y$, then

$$\begin{aligned} \mathbb{P}[Y = y] &= \mathbb{P}[h(X) = y] \\ &= \mathbb{P}\left[X = h^{-1}(y)\right] \\ &> 0 \end{aligned}$$

Hence, for every $y \in h(S_X)$, the probability $\mathbb{P}[Y = y]$ is non-zero.

$$\begin{aligned} \mathbb{P}[Y \in h(S_X)] &= \mathbb{P}[h(X) \in h(S_X)] \\ &= \mathbb{P}[X \in S_X] \quad = \quad 1 \end{aligned}$$

Since for every $y \in h(S_X)$, the probability $\mathbb{P}[Y = y]$ is non-zero and the probability $\mathbb{P}[Y \in h(S_X)] = 1$, we can say that $h(S_X)$ is the support for the random variable Y. $\square$

For $y \in S_Y = h(S_X)$,

$$\begin{aligned} \mathbb{P}[Y = y] &= \mathbb{P}[h(X) = y] \\ &= \mathbb{P}\left[X \in h^{-1}(y)\right] \\ &= \sum_{x \in h^{-1}(y)} f_X(x) \end{aligned}$$

Hence, we can finally write the pmf of the random variable Y where $Y = h(X)$ with support $S_Y = h(S_X) = \{h(x) \mid x \in S_X\}$

$$f_Y(y) \quad = \quad \begin{cases} \sum_{x \in h^{-1}(y)} f_X(x) & \text{if } y \in S_Y \\ 0 & \text{else} \end{cases}$$

2. Absolutely Continuous Variable

Consider a random variable X and a random variable $Y = h(X)$. As discussed in the previous section, if X is discrete, Y is also discrete. However, this is not true for an absolutely continuous variable. Consider the following example.

Example 4.1. Let X be an absolutely continuous r.v. with pdf

$$f_X(x) = \begin{cases} \frac{1}{2} & \text{if } -1 < x < 1 \\ 0 & \text{else} \end{cases}$$

Let $Y = \lfloor X \rfloor$. Then, we can write Y as

$$Y = \begin{cases} -1 & \text{if } -1 < X < 0 \\ 0 & \text{if } 0 \leq X < 1 \end{cases}$$

Therefore, we can say

$$\begin{aligned} \mathbb{P}[Y = -1] &= \mathbb{P}[-1 < X < 0] \\ &= \int_{-1}^{0} f_x(x)dx \\ &= \frac{1}{2} \end{aligned}$$

$$\begin{aligned} \mathbb{P}[Y = 0] &= \mathbb{P}[0 \leq X < 1] \\ &= \int_{0}^{1} f_x(x)dx \\ &= \frac{1}{2} \end{aligned}$$

Clearly, Y is discrete with support $S_Y = \{-1, 0\}$, and pmf

$$f_Y(y) = \begin{cases} \frac{1}{2} & \text{if } y \in -1, 0 \\ 0 & \text{else} \end{cases}$$

We now give a formula to transform Absolutely Continuous Random variables (say X and Y) such that $Y = h(X)$, for which the support is a collection of disjoint intervals $S_X = \bigcup_n S_{n,X}$, such that for each set $S_{n,X}$, $h_n : S_{n,x} \longrightarrow \mathbb{R}$ is strictly monotone with inverse function $h^{-1}(y)$ with $\frac{d}{dy} h^{-1}(y)$ is continuous.

Let $h(S_{n,X} = \{h(x) \mid x \in S_{n,X}\}$, then the pdf for Y (f_Y) is written as

$$f_Y(y) = \sum_n f_Y\left(h_n{}^{-1}(y)\right) \left| \frac{d}{dy} h_n{}^{-1}(y) \right| \mathbb{I}[y \in h(S_{n,X})]$$

Exercise 4.1. Let X be an absolutely continuous r.v. with pdf

$$f_X(x) = \begin{cases} \frac{\lfloor x \rfloor}{2} & \text{if } -1 < x < 1 \\ \frac{x}{3} & \text{if } 1 < x < 2 \\ 0 & \text{else} \end{cases}$$

Let $Y = X^2$, then

(i) find the pdf of Y and hence the CDF of Y

(ii) find the CDF of Y and hence the pdf of Y