# Understanding Distributions of Random Variables

Ravi Saripalli

29th Sep. 2021

## 1 Expressing Discrete PDF as a continuous function

Consider a discrete distribution function

$$f(x_i) = p_k \quad x_i \in \{x_1, x_2, ...x_n\}$$

Its cumulative distribution

$$F(x < x_k) = \sum_{i=1}^{k} p_k$$

This function can be expressed as continuous CDF in x with the help of step functions

$$F(x) = \sum_{i=1}^{k} f(x_i) S(x - x_i) \quad x_1 < x < x_n$$

where

$$S(x) = \begin{cases} 0, & \text{if } x < 0 \\ 1, & \text{otherwise} \end{cases}$$

If we differentiate the above function we should get continuous equivalent of discrete distribution $P(x_i)$

$$f(x) = \frac{dF}{dx} = \sum_{i=1}^{k} f(x_i) \delta(x - x_i) \quad x_1 < x < x_n$$

*Usefulness of this is yet to be explored by me.*

## 2 PDF of transformed random variable

Given a random continuous variable $x$ with known PDF $P_X(x)$ determine the PDF of a transformed variable $P_Y(y)$ with transformation $Y = g(X)$.

This is achieved expressing the CDF of transformed variable $Y$ in terms of CDF of original variable $X$. with appropriate reverse transformation.

$$\begin{aligned}
F_Y(y) &= P(Y < y) \\
&= P(g(X) < y) \\
&= P(X < g^{-1}(y)) \;\; \text{provided g is monotonically increasing} \\
&= F_X(g^{-1}(y))
\end{aligned}$$

Note that $g^{-1}$ function refers to reverse mapping from Y to X. Also, the requirement for the function $g(X)$ to be a monotonically increasing function in the interval of interest is to ensure the inequality is preserved during reverse mapping.

The PDF of transformed variable can now be determined with differentiation of above CDF with respect to y.

$$\begin{aligned}
P_Y(y) &= \frac{d}{dy} F_Y(y) \\
&= \frac{d}{dy} F_X(g^{-1}(y)) \\
&= \frac{dx}{dy} \frac{d}{dx} F_X(g^{-1}(y)) \\
&= \frac{dx}{dy} P_X(g^{-1}(y)) \\
&= \frac{d}{dy}(g^{-1}(y)) \; P_X(g^{-1}(y))
\end{aligned}$$

Although above result is derived for monotonically increasing $g(X)$, for non-monotonic function case, one can segment the functions into monotonically increasing and decreasing segements and a generic result can be obtained as follows.

Assuming the $g(X)$ has $k$ segments of montonically increasing or decreasing regions,

$$P_Y(y) = \sum_{i=1}^{k} \left| \frac{d}{dy}(g_k^{-1}(y)) \right| \; P_X(g_k^{-1}(y)) \quad y \in Y$$

Note the absolute sign on the derivative term is to account for effect of sign change between monotonically increasing and decreasing function gradients. *How one can do it for very complicated functions is not clear to me.*

## 3 Transformed PDF with multiple random variables

But my interest is to find PDF of a function $g(x1, x2, x3)$ given $x1, x2, x3$ are members of a random variables with their respective PDF functions are known.