TRANSFORMATIONS OF RANDOM VARIABLES

Consider a discrete/continuous random variable X with range $\mathbb{X}$ and probability distribution described by mass/pdf f_X, or cdf F_X. Suppose g is a real-valued function whose domain includes $\mathbb{X}$, and suppose that

$$g: \quad \begin{array}{l} \mathbb{X} \longrightarrow \mathbb{Y} \\ x \longmapsto y \end{array}$$

Then $Y = g(X)$ is also a random variable as Y is a function from Ω to $\mathbb{R}$.

For $A \subseteq \mathbb{R}$, the event $[\ Y \in A\ ]$ is an event in terms of the transformed variable Y. If f_Y is the mass/density function for Y, then

$$\mathrm{P}[\ Y \in A\ ] = \begin{cases} \displaystyle\sum_{y \in A} f_Y(y) & Y \text{ discrete} \\ \\ \displaystyle\int_A f_Y(y)\ dy & Y \text{ continuous} \end{cases}$$

We wish to derive the probability distribution of random variable Y; in order to do this, we first consdier the inverse transformation g^{-1} from $\mathbb{Y}$ to $\mathbb{X}$ defined for set $A \subseteq \mathbb{Y}$ (and for $y \in \mathbb{Y}$) by

$$g^{-1}(A) = \{\ x \in \mathbb{X}\ :\ g(x) \in A\} \qquad g^{-1}(y) = \{\ x \in \mathbb{X}\ :\ g(x) = y\}$$

that is, $g^{-1}(A)$ is the set of points in $\mathbb{X}$ that map into A, and $g^{-1}(y)$ is the set of points in $\mathbb{X}$ that map to y, under transformation g. By construction, we have

$$\mathrm{P}[\ Y \in A\ ] = \mathrm{P}\left[\ X \in g^{-1}(A)\ \right]$$

and hence $[\ Y \in A\ ]$ and $[\ X \in g^{-1}(A)\ ]$ are **equivalent** events.

Consider first the cdf of Y, F_Y, evaluated at a point $y \in \mathbb{R}$. We have

$$F_Y(y) = \mathrm{P}[\ Y \leq y\ ] = \mathrm{P}[\ g(X) \leq y\ ] = \begin{cases} \displaystyle\sum_{x \in A_y} f_X(x) & \text{if } X \text{ is discrete} \\ \\ \displaystyle\int_{A_y} f_X(x)\ dx & \text{if } X \text{ is continuous} \end{cases}$$

where $A_y = \{\ x \in \mathbb{X}\ :\ g(x) \leq y\}$. Often, the set is A_y is easy to identify for a given y, that is, $F_Y(y) = \mathrm{P}[\ g(X) \leq y\ ] = \mathrm{P}[\ x_1 \leq X \leq\ x_2]$, where x_1 and x_2 depend on y and g or g^{-1}. The main objective is therefore to identify the set A_y.

1-1 Transformations

The mapping $g(X)$ is a function of X from $\mathbb{X}$ which is *1-1 and onto* $\mathbb{Y}$ if,

(i) for each $x \in \mathbb{X}$, there exists one and only one y such that $y = g(x)$, and

(ii) for each $y \in \mathbb{Y}$, there exists an $x \in \mathbb{X}$ such that $g(x) = y$.

The following theorem gives the distribution for random variable $Y = g(X)$ when g is 1-1.

THEOREM

Let X be a random variable with mass/density function f_X and support $\mathbb{X}$. Let g be a 1-1 function from $\mathbb{X}$ onto $\mathbb{Y}$ with inverse g^{-1}. Then $Y = g(X)$ is a random variable with support $\mathbb{Y}$ and

Discrete Case : The mass function of random variable Y is given by

$$f_Y(y) = f_X(g^{-1}(y)) \qquad y \in \mathbb{Y} = \{\ y \mid f_Y(y) > 0\ \}.$$

where x is the unique solution of $y = g(x)$ (so that $x = g^{-1}(y)$).

Continuous Case : The pdf of random variable Y is given by

$$f_Y(y) = f_X(g^{-1}(y)) \left| \frac{d}{dt}\left\{g^{-1}(t)\right\}_{t=y} \right| \qquad y \in \mathbb{Y} = \{\ y \mid f_Y(y) > 0\ \}$$

where $y = g(x)$, provided that the derivative $\dfrac{d}{dt}\left\{g^{-1}(t)\right\}$ is *continuous* and *non-zero* on $\mathbb{Y}$.

PROOF

Discrete case: by direct calculation, $f_Y(y) = \mathrm{P}[\ Y = y\ ] = \mathrm{P}[\ g(X) = y\ ] = \mathrm{P}[\ X = g^{-1}(y)\ ] = f_X(x)$. where $x = g^{-1}(y)$, and hence $f_Y(y) > 0 \Longleftrightarrow f_X(x) > 0$.

Continuous case: function g is either (I) a monotonic increasing, or (II) a monotonic decreasing function.

Case (I): If g is *increasing*, then for $x \in \mathbb{X}$ and $y \in \mathbb{Y}$, we have that $g(x) \leq y \Longleftrightarrow x \leq g^{-1}(y)$. Therefore, for $y \in \mathbb{Y}$,

$$F_Y(y) = \mathrm{P}[\ Y \leq y\ ] = \mathrm{P}[\ g(X) \leq y\ ] = \mathrm{P}[\ X \leq g^{-1}(y)\ ] = F_X(g^{-1}(y))$$

and, by differentiation, because g is monotonic increasing,

$$f_Y(y) = f_X(g^{-1}(y)) \frac{d}{dt}\left\{g^{-1}(t)\right\}_{t=y} = f_X(g^{-1}(y)) \left| \frac{d}{dy}\left\{g^{-1}(y)\right\}_{t=y} \right| \qquad \text{as } \frac{d}{dt}\left\{g^{-1}(t)\right\} > 0.$$

Case (II): If g is *decreasing*, then for $x \in \mathbb{X}$ and $y \in \mathbb{Y}$ we have $g(x) \leq y \Longleftrightarrow x \geq g^{-1}(y)$. Therefore, for $y \in \mathbb{Y}$,

$$F_Y(y) = \mathrm{P}[\ Y \leq y\ ] = \mathrm{P}[\ g(X) \leq y\ ] = \mathrm{P}[\ X \geq g^{-1}(y)\ ] = 1 - F_X(g^{-1}(y))$$

so

$$f_Y(y) = -f_X(g^{-1}(y)) \frac{d}{dt}\left\{g^{-1}(y)\right\} = f_X(g^{-1}(y)) \left| \frac{d}{dt}\left\{g^{-1}(t)\right\}_{t=y} \right| \qquad \text{as } \frac{d}{dt}\left\{g^{-1}(t)\right\} < 0.$$

DEFINITION

Suppose transformation $g : \mathbb{X} \longrightarrow \mathbb{Y}$ is 1-1, and is defined by $g(x) = y$ for $x \in \mathbb{X}$. Then the **Jacobian** of the transformation, denoted $J(y)$, is given by

$$J(y) = \left| \frac{d}{dt}\left\{g^{-1}(t)\right\}_{t=y} \right|$$

that is, the absolute value of first derivative of g^{-1} evaluated at $y = g(x)$. Note that the inverse transformation $g^{-1} : \mathbb{Y} \longrightarrow \mathbb{X}$ has Jacobian $1/J(x)$.