Transformations of Random Variables

September, 2009

We begin with a random variable X and we want to start looking at the random variable $Y = g(X) = g \circ X$ where the function

$$g : \mathbb{R} \to \mathbb{R}.$$

The **inverse image** of a set A,

$$g^{-1}(A) = \{x \in \mathbb{R}; g(x) \in A\}.$$

In other words,

$$x \in g^{-1}(A) \text{ if and only if } g(x) \in A.$$

For example, if $g(x) = x^3$, then $g^{-1}([1, 8]) = [1, 2]$

For the singleton set $A = \{y\}$, we sometimes write $g^{-1}(\{y\}) = g^{-1}(y)$. For $y = 0$ and $g(x) = \sin x$, $g^{-1}(0) = \{k\pi; k \in \mathbb{Z}\}$.

If g is a one-to-one function, then the inverse image of a singleton set is itself a singleton set. In this case, the inverse image naturally defines an inverse function. For $g(x) = x^3$, this inverse function is the cube root. For $g(x) = \sin x$ or $g(x) = x^2$ we must limit the domain to obtain an inverse function.

Exercise 1. *The inverse image has the following properties:*

- $g^{-1}(\mathbb{R}) = \mathbb{R}$
- *For any set A, $g^{-1}(A^c) = g^{-1}(A)^c$*
- *For any collection of sets $\{A_\lambda; \lambda \in \Lambda\}$,*

$$g^{-1}\left(\bigcup_\lambda A_\lambda\right) = \bigcup_\lambda g^{-1}(A).$$

As a consequence the mapping

$$A \mapsto P\{g(X) \in A\} = P\{X \in g^{-1}(A)\}$$

satisfies the axioms of a probability. The associated probability $\mu_{g(X)}$ is called the **distribution** of $g(X)$.

1 Discrete Random Variables

For X a discrete random variable with probabiliity mass function f_X, then the probability mass function f_Y for $Y = g(X)$ is easy to write.

$$f_Y(y) = \sum_{x \in g^{-1}(y)} f_X(x).$$

Example 2. *Let X be a uniform random variable on $\{1, 2, \ldots n\}$, i. e., $f_X(x) = 1/n$ for each x in the state space. Then $Y = X + a$ is a uniform random variable on $\{a + 1, 2, \ldots a + n\}$*

Example 3. *Let X be a uniform random variable on $\{-n, -n + 1, \ldots, n - 1, n\}$. Then $Y = |X|$ has mass function*

$$f_Y(y) = \begin{cases} \frac{1}{2n+1} & \text{if } x = 0, \\ \frac{2}{2n+1} & \text{if } x \neq 0. \end{cases}$$

2 Continuous Random Variable

The easiest case for transformations of continuous random variables is the case of g one-to-one. We first consider the case of g increasing on the range of the random variable X. In this case, g^{-1} is also an increasing function.

To compute the cumulative distribution of $Y = g(X)$ in terms of the cumulative distribution of X, note that

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

Now use the chain rule to compute the density of Y

$$f_Y(y) = F_Y'(y) = \frac{d}{dy} F_X(g^{-1}(y)) = f_X(g^{-1}(y)) \frac{d}{dy} g^{-1}(y).$$

For g decreasing on the range of X,

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

and the density

$$f_Y(y) = F_Y'(y) = -\frac{d}{dy} F_X(g^{-1}(y)) = -f_X(g^{-1}(y)) \frac{d}{dy} g^{-1}(y).$$

For g decreasing, we also have g^{-1} decreasing and consequently the density of Y is indeed positive,

We can combine these two cases to obtain

$$f_Y(y) = f_X(g^{-1}(y)) \left| \frac{d}{dy} g^{-1}(y) \right|.$$

Example 4. *Let U be a uniform random variable on $[0, 1]$ and let $g(u) = 1 - u$. Then $g^{-1}(v) = 1 - v$, and $V = 1 - U$ has density*

$$f_V(v) = f_U(1 - v)|-1| = 1$$

on the interval $[0, 1]$ and 0 otherwise.

Example 5. *Let X be a random variable that has a uniform density on $[0,1]$. Its density*

$$f_X(x) = \begin{cases} 0 & \text{if } x < 0, \\ 1 & \text{if } 0 \le x \le 1, \\ 0 & \text{if } x > 1. \end{cases}$$

Let $g(x) = x^p,\ p \neq 0$. Then, the range of g is $[0,1]$ and $g^{-1}(y) = y^{1/p}$. If $p > 0$, then g is increasing and

$$\frac{d}{dy} g^{-1}(y) = \begin{cases} 0 & \text{if } y < 0, \\ \frac{1}{p} y^{1/p-1} & \text{if } 0 \le y \le 1, \\ 0 & \text{if } y > 1. \end{cases}$$

This density is unbounded near zero whenever $p > 1$.

If $p < 0$, then g is decreasing. Its range is $[1,\infty)$, and

$$\frac{d}{dy} g^{-1}(y) = \begin{cases} 0 & \text{if } y < 1, \\ -\frac{1}{p} y^{1/p-1} & \text{if } y \ge 1, \end{cases}$$

In this case, Y is a Pareto distribution with $\alpha = 1$ and $\beta = -1/p$. We can obtain a Pareto distribution with arbitrary α and β by taking

$$g(x) = \left(\frac{x}{\alpha}\right)^{1/\beta}.$$

If the transform g is not one-to-one then special care is necessary to find the density of $Y = g(X)$. For example if we take $g(x) = x^2$, then $g^{-1}(y) = \sqrt{y}$.

$$F_y(y) = P\{Y \le y\} = P\{X^2 \le y\} = P\{-\sqrt{y} \le X \le \sqrt{y}\} = F_X(\sqrt{y}) - F_X(-\sqrt{y}).$$

Thus,

$$\begin{aligned} f_Y(y) &= f_X(\sqrt{y}) \frac{d}{dy}(\sqrt{y}) - f_X(-\sqrt{y}) \frac{d}{dy}(-\sqrt{y}) \\ &= \frac{1}{2\sqrt{y}} (f_X(\sqrt{y}) + f_X(-\sqrt{y})) \end{aligned}$$

If the density f_X is symmetric about the origin, then

$$f_y(y) = \frac{1}{\sqrt{y}} f_X(\sqrt{y}).$$

Example 6. *A random variable Z is called a* **standard normal** *if its density is*

$$\phi(z) = \frac{1}{\sqrt{2\pi}} \exp(-\frac{z^2}{2}).$$

A calculus exercise yields

$$\phi'(z) = -\frac{1}{\sqrt{2\pi}} z \exp(-\frac{z^2}{2}) = -z\phi(z), \qquad \phi''(z) = \frac{1}{\sqrt{2\pi}} (z^2 - 1) \exp(-\frac{z^2}{2}) = (z^2 - 1)\phi(z).$$

Thus, ϕ has a global maximum at $z = 0$, it is concave down if $|z| < 1$ and concave up for $|z| > 1$. This show that the graph of ϕ has a bell shape.

*$Y = Z^2$ is called a χ^2 (***chi-square***) random variable with one degree of freedom. Its density is*

$$f_Y(y) = \frac{1}{\sqrt{2\pi y}} \exp(-\frac{y}{2}).$$

3 The Probability Transform

Let X a continuous random variable whose distribution function F_X is strictly increasing on the possible values of X. Then F_X has an inverse function.

Let $U = F_X(X)$, then for $u \in [0,1]$,

$$P\{U \le u\} = P\{F_X(X) \le u\} = P\{U \le F_X^{-1}(u)\} = F_X(F_X^{-1}(u)) = u.$$

In other words, U is a uniform random variable on $[0,1]$. Most random number generators simulate independent copies of this random variable. Consequently, we can simulate independent random variables having distribution function F_X by simulating U, a uniform random variable on $[0,1]$, and then taking

$$X = F_X^{-1}(U).$$

Example 7. *Let X be uniform on the interval $[a,b]$, then*

$$F_X(x) = \begin{cases} 0 & \text{if } x < a, \\ \frac{x-a}{b-a} & \text{if } a \le x \le b, \\ 1 & \text{if } x > b. \end{cases}$$

Then

$$u = \frac{x-a}{b-a}, \quad (b-a)u + a = x = F_X^{-1}(u).$$

Example 8. *Let T be an exponential random variable. Thus,*

$$F_T(t) = \begin{cases} 0 & \text{if } t < 0, \\ 1 - \exp(-t/\beta) & \text{if } t \ge 0. \end{cases}$$

Then,

$$u = 1 - \exp(-t/\beta), \quad \exp(-t/\beta) = 1 - u, \quad t = -\frac{1}{\beta}\log(1-u).$$

Recall that if U is a uniform random variable on $[0,1]$, then so is $V = 1-U$. Thus if V is a uniform random variable on $[0,1]$, then

$$T = -\frac{1}{\beta}\log V$$

is a random variable with distribution function F_T.

Example 9. *Because*

$$\int_\alpha^x \frac{\beta\alpha^\beta}{t^{\beta+1}}\, dt = -\alpha^\beta t^{-\beta}\Big|_\alpha^x = 1 - \left(\frac{\alpha}{x}\right)^\beta.$$

A Pareto random variable X has distribution function

$$F_X(x) = \begin{cases} 0 & \text{if } x < \alpha, \\ 1 - \left(\frac{\alpha}{x}\right)^\beta & \text{if } x \ge \alpha. \end{cases}$$

Now,

$$u = 1 - \left(\frac{\alpha}{x}\right)^\beta \quad 1 - u = \left(\frac{\alpha}{x}\right)^\beta, \quad x = \frac{\alpha}{(1-u)^{1/\beta}}.$$

As before if $V = 1 - U$ is a uniform random variable on $[0, 1]$, then

$$X = \frac{\alpha}{V^{1/\beta}}$$

is a Pareto random variable with distribution function F_X.