556: MATHEMATICAL STATISTICS I

NON 1-1 TRANSFORMATIONS

Suppose that X is a continuous r.v. with range $\mathbb{X} \equiv (0, 2\pi)$ whose pdf f_X is constant

$$f_X(x) = \frac{1}{2\pi} \qquad 0 < x < 2\pi$$

and zero otherwise. This pdf has corresponding continuous cdf

$$F_X(x) = \frac{x}{2\pi} \qquad 0 < x < 2\pi$$

Example 1 Consider transformed r.v. $Y = \sin X$. Then the range of Y, $\mathbb{Y}$ is $[-1, 1]$, but the transformation is not 1-1. However, from first principles, we have

$$F_Y(y) = P_Y[Y \leq y] = P_X[\sin X \leq y]$$

Now, by inspection of Figure 1, for $y > 0$, we can easily identify the required set A_y : it is the union of **two** disjoint intervals $A_y = [0, x_1] \cup [x_2, 2\pi] = [0, \arcsin y] \cup [\pi - \arcsin y, 2\pi]$

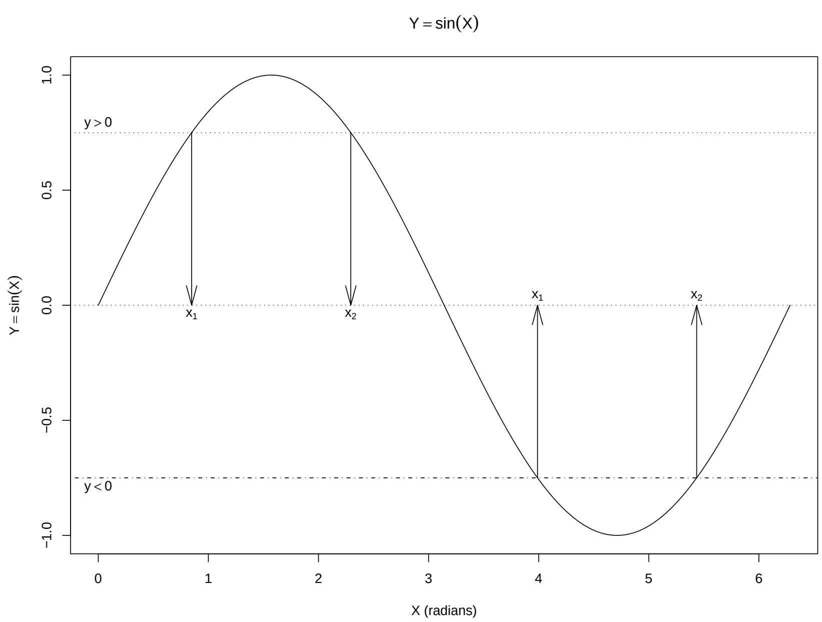

Figure 1: Computation of A_y for $Y = \sin X$

$$F_Y(y) = P_X[\sin X \leq y] = P_X[X \leq x_1] + P_X[X \geq x_2] = \{P_X[X \leq x_1]\} + \{1 - P_X[X < x_2]\}$$

$$= \left\{\frac{1}{2\pi}\arcsin y\right\} + \left\{1 - \frac{1}{2\pi}(\pi - \arcsin y)\right\} = \frac{1}{2} + \frac{1}{\pi}\arcsin y$$

For $y < 0$, the calculation is slightly different, with a horizontal line at a given y cutting through the part of the sine curve in the interval $(\pi, 2\pi)$. In this case, remembering that the arcsine function takes values on $[-\pi/2, \pi/2]$, we have that $x_1 = \pi - \arcsin y$ and $x_2 = 2\pi + \arcsin y$. Thus

$$\begin{aligned} F_Y(y) &= P_X[\sin X \leq y] = P_X[x_1 \leq X \leq x_2] = P_X[X \leq x_2] - P_X[X \leq x_1] \\ &= \frac{1}{2\pi}(2\pi + \arcsin y) - \frac{1}{2\pi}(\pi - \arcsin y) = \frac{1}{2} + \frac{1}{\pi}\arcsin y \end{aligned}$$

so, in fact, the answer is unchanged. Hence, by differentiation

$$f_Y(y) = \frac{1}{\pi}\frac{1}{\sqrt{1-y^2}} \qquad -1 \le y \le 1$$

and zero otherwise.

Example 2 Consider transformed r.v. $Y = \sin^2 X$. Then the range of Y, $\mathbb{Y}$, is $[0, 1]$, but the transformation is not 1-1. However, from first principles, we have

$$F_Y(y) = P_Y[Y \le y] = P_X[\sin X \le y]$$

In Figure 2, we identify the required set A_y : it is the union of **three** disjoint intervals

$$A_y = [0, x_1] \cup [x_2, x_3] \cup [x_4, 2\pi]$$

where

$$x_1 = \arcsin(\sqrt{y}) \qquad x_2 = \pi - \arcsin(\sqrt{y}) \qquad x_3 = \pi + \arcsin(\sqrt{y}) \qquad x_4 = 2\pi - \arcsin(\sqrt{y})$$

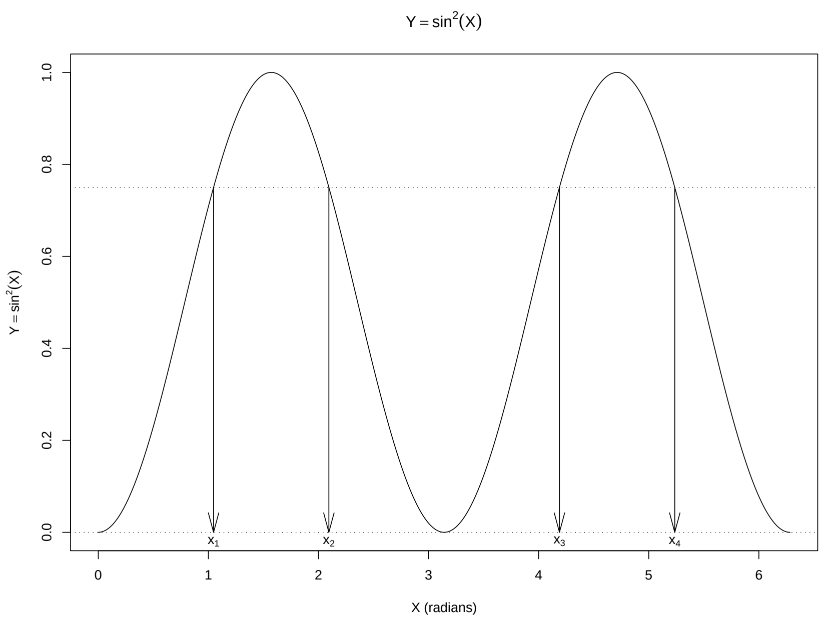

Figure 2: Computation of A_y for $Y = \sin^2 X$

$$F_Y(y) = P_X\left[\sin^2 X \le y\right] = P_X[X \le x_1] + P_X[x_2 < X \le x_3] + P_X[x_4 < X \le 2\pi]$$

$$= F_X(x_1) + \{F_X(x_3) - F_X(x_2)\} + \{1 - F_X(x_4)\}$$

$$= \frac{x_1}{2\pi} + \left\{\frac{x_3}{2\pi} - \frac{x_2}{2\pi}\right\} + \left\{1 - \frac{x_4}{2\pi}\right\} = \frac{2}{\pi}\arcsin(\sqrt{y})$$

and hence, by differentiation

$$f_Y(y) = \frac{1}{\pi}\frac{1}{\sqrt{(1-y)y}} \qquad 0 \le y \le 1$$

and zero otherwise.

Example 3 Consider transformed r.v. $T = \tan X$. Then the range of T, $\mathbb{T}$ is $\mathbb{R}$, but the transformation is not 1-1. However, from first principles, we have, for $t > 0$

$$F_T(t) = P_T\left[T \leq t\right] = P_X\left[\tan X \leq t\right]$$

Figure 3 helps identify the required set A_t: in this case, it is the union of three disjoint intervals

$$A_t = [0, x_1] \cup \left[\frac{\pi}{2}, x_2\right] \cup \left[\frac{3\pi}{2}, 2\pi\right] = \left[0, \tan^{-1} t\right] \cup \left[\frac{\pi}{2}, \pi + \tan^{-1} t\right] \cup \left[\frac{3\pi}{2}, 2\pi\right]$$

(note, for values of $t < 0$, the union will be of only two intervals, but the calculation proceeds identically) so that

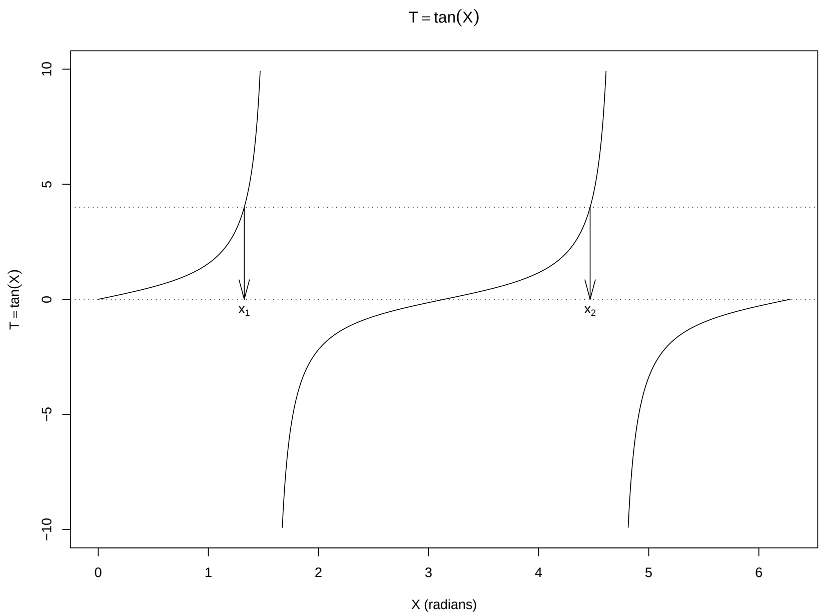

Figure 3: Computation of A_t for $T = \tan X$

$$F_T(t) = P_X\left[\tan X \leq t\right] = P_X\left[X \leq x_1\right] + P_X\left[\frac{\pi}{2} \leq X \leq x_2\right] + P_X\left[\frac{3\pi}{2} \leq X \leq 2\pi\right]$$

$$= \left\{\frac{1}{2\pi}\tan^{-1} t\right\} + \frac{1}{2\pi}\left\{\pi + \tan^{-1} t - \frac{\pi}{2}\right\} + \frac{1}{2\pi}\left\{2\pi - \frac{3\pi}{2}\right\} = \frac{1}{\pi}\tan^{-1} t + \frac{1}{2}$$

and hence, by differentiation

$$f_T(t) = \frac{1}{\pi}\frac{1}{1+t^2} \qquad t \in \mathbb{R}$$

The case for $t < 0$ yields the same result.