Lecture 13: Sept 21

Last time

- Counting Techniques
- Transformations of Random Variables

Today

- Transformations of continuous random variables

Theorem Suppose a continuous random variable X has cdf $F_X(x)$, let $Y = g(X)$, and let $\mathcal{X}$ and $\mathcal{Y}$ be defined as

$$\mathcal{X} = \{x : f(x) > 0\} \quad \text{and} \quad \mathcal{Y} = \{y : y = g(x) \text{ for some } x \in \mathcal{X}\}.$$

Then,

1. If g is an increasing function on $\mathcal{X}$, $F_Y(y) = F_X(g^{-1}(y))$ for $y \in \mathcal{Y}$.
2. If g is a decreasing function on $\mathcal{X}$, $F_Y(y) = 1 - F_X(g^{-1}(y))$ for $y \in \mathcal{Y}$.

Proof: We start with

$$\begin{aligned} F_Y(y) &= \Pr(Y \leqslant y) \\ &= \Pr(g(X) \leqslant y) \end{aligned}$$

Theorem Let X have pdf $f_X(x)$ and let $Y = g(X)$, where g is a monotone function. Let $\mathcal{X}$ and $\mathcal{Y}$ be defined as

$$\mathcal{X} = \{x : f(x) > 0\} \quad \text{and} \quad \mathcal{Y} = \{y : y = g(x) \text{ for some } x \in \mathcal{X}\}.$$

Suppose that $f_X(x)$ is continuous on $\mathcal{X}$ and that $g^{-1}(y)$ has a continuous derivative on $\mathcal{Y}$. Then the pdf of Y is given by

$$f_Y(y) = \begin{cases} f_X(g^{-1}(y))\left|\frac{d}{dy}g^{-1}(y)\right| & y \in \mathcal{Y} \\ 0 & \textit{otherwise.} \end{cases}$$

Proof:
From last theorem, we have the cdf forms $F_Y(y)$. Then $f_Y(y) = \frac{d}{dy}F_Y(y)$. (finish the proof)

Example (Square transformation) Suppose X is a continuous random variable. For $y > 0$, the cdf of $Y = X^2$ is

$$F_Y(y) = \Pr(Y \leqslant y) = \Pr(X^2 \leqslant y) = \Pr(-\sqrt{y} \leqslant X \leqslant \sqrt{y}).$$

Because x is continuous, we can drop the equality from the left endpoint and obtain

$$\begin{aligned} F_Y(y) &= \Pr(-\sqrt{y} < X \leqslant \sqrt{y}) \\ &= \Pr(X \leqslant \sqrt{y}) - \Pr(X \leqslant -\sqrt{y}) = F_X(\sqrt{y}) - F_X(-\sqrt{y}). \end{aligned}$$

The pdf of Y can now be obtained from the cdf by differentiation: where we use the chain rule to differentiate $F_X(\sqrt{y})$ and $F_X(-\sqrt{y})$.

Example (Linear transformation) Suppose X is a continuous random variable with pdf $f_X(x)$. Let

$$Y = a + bX, \quad \frac{dy}{dx} = b.$$

Then

$$f_Y(y) = f_X\left[g^{-1}(y)\right]|\frac{dx}{dy}| = f_X(\frac{y-a}{b})\frac{1}{|b|}.$$

This transformation is often used when X has mean 0 and standard deviation 1. The linear transformation above creates a random variable Y with a distribution that has the same shape as that of X but has mean a and variance b^2.

Conversely, if Y has mean a and standard deviation b, then $X = (Y - a)/b$ has mean 0 and standard deviation 1. This is called sometimes the "Studentized" transformation.

Example (Normal distribution) Let $X \sim N(0, 1)$:

$$f_X(x) = \frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}, \quad -\infty < x < \infty.$$

The transformation

$$Y = \mu + \sigma X, \quad X = \frac{Y-\mu}{\sigma}$$

yields

$$f_Y(y) = f_X(\frac{y-\mu}{\sigma})\frac{1}{\sigma} = \frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(y-\mu)^2}{2\sigma^2}}.$$

More generally, a distribution is a member of the class of *location-scale* distributions if the distribution of a linear transformation of a random variable with that distribution has the same distribution, but with different parameters.

Example (Square root of an exponential RV) Suppose $X \sim exp(\lambda)$, so that

$$f_X(x) = \begin{cases} \lambda e^{-\lambda x} & x \geqslant 0 \\ 0 & otherwise \end{cases}$$

and consider the distribution of $Y = \sqrt{X}$. The transformation

$$y = g(x) = \sqrt{x}, \quad x \geqslant 0$$

is one-to-one and has an inverse $x = y^2$ with $dx/dy = 2y$. Thus
This distribution is a particular form of the Rayleigh distribution and is a special case of the Weibull distribution.

Theorem (Probability integral transformation) Let X have continuous cdf $F_X(x)$ and define the random variable Y as $Y = F_X(X)$. Then Y is uniformly distributed on $(0, 1)$, that is, $\Pr(Y \leqslant y) = y, 0 < y < 1$.

Before we prove this theorem, we will digress for a moment and look at F_X^{-1}, the inverse of the cdf F_X, in some detail. If F_X is strictly increasing, then F_X^{-1} is well defined by

$$F_X^{-1}(y) = x \iff F_X(x) = y.$$

However, if F_X is constant on some interval, then F_X^{-1} is not well defined as Figure 13.1 illustrates. Any $x_1 \leqslant x \leqslant x_2$ satisfies $F_X(x) = y$

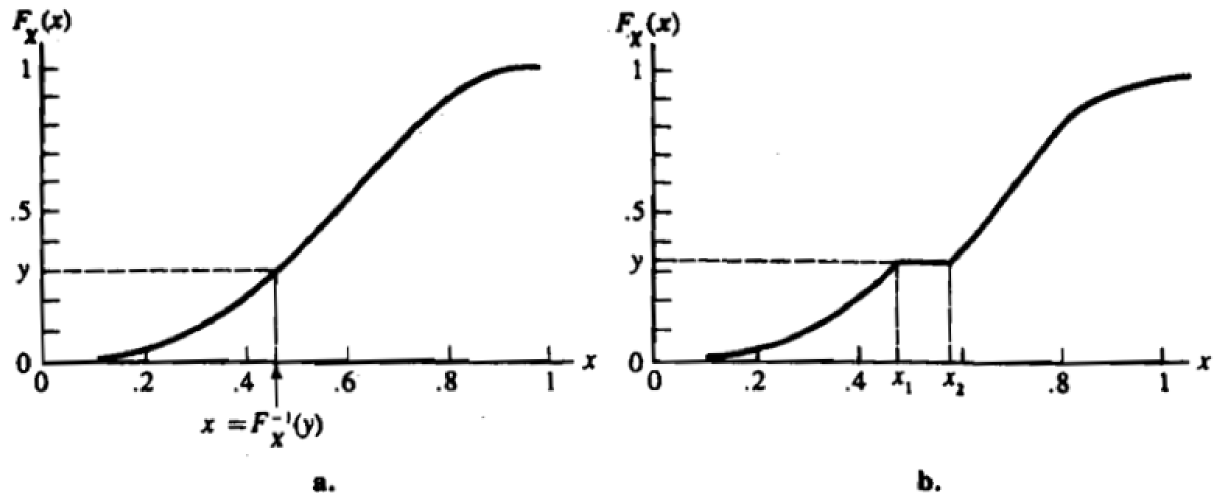

Figure 13.1: Figure 2.1.2. (a) $F_X(x)$ strictly increasing; (b) $F_X(x)$ nondecreasing

This problem is avoided by defining F_X^{-1} for $0 < y < 1$ by

$$F_X^{-1}(y) = \inf\{x : F_X(x) \geqslant y\}.$$

With this definition, for Figure 13.1(b), we have $F_X^{-1}(y) = x_1$.
Proof:

One application of the probability integral transformation is in the generation of random samples from a particular distribution. If it is required to generate an observation X from a population with cdf F_X, we need only generate a uniform random number V, between 0 and 1, and solve for x in the equation $F_X(x) = u$.