Lecture 11: Sept 15

Last time

- Discrete Random Variables

Today

- Continuous Random Variables
- Transformations of Random Variables

Continuous Random Variables

Definition A random variable X is *continuous* if $F_X(x)$ is a continuous function of x.

Definition A random variable X is *absolutely continuous* if $F_X(x)$ is an absolutely continuous function of x.

Definition A function $F(x)$ is *absolutely continuous* if it can be written

$$F(x) = \int_{-\infty}^{x} f(x)dx.$$

Absolute continuity is stronger than continuity but weaker than differentiability. An example of an absolutely continuous function is one that is:

- continuous everywhere
- differentiable everywhere, except possibly for a countable number of points.

Definition The *probability density function* or pdf, $f_X(x)$, of a continuous random variable X is the function that satisfies

$$F_X(x) = \int_{-\infty}^{x} f_X(t)dt \quad \text{for all } x.$$

Notation: We write $X \sim F_X(x)$ for the expression "X has a distribution given by $F_X(x)$" where we read the symbol "$\sim$" as "is distributed as". Similarly, we can write $X \sim f_X(x)$ or , if X and Y have the same distribution, $X \sim Y$.

Theorem A function $f_X(x)$ is a pdf (or pmf) of a random variable X if and only if

1. $f_X(x) \geqslant 0$ for all x.
2. $\int_{-\infty}^{\infty} f_X(x)dx = 1$ (pdf) or $\sum_x f_X(x) = 1$ (pmf).

Example Suppose $F(x) = 1 - e^{-\lambda x}$ for $x > 0$ and $F(x) = 0$ otherwise. Is $F(x)$ a cdf? What is the associated pdf?
solution:

Notes

- If X is a continuous random variable, then $f(x)$ is not the probability that $X = x$. In fact, if X is an absolutely continuous random variable with density function $f(x)$, then $\Pr(X = x) = 0$. (Why?)
  proof

- Because $\Pr(X = a) = 0$, all the following are equivalent:

$$\Pr(a \leqslant X \leqslant b), \quad \Pr(a \leqslant X < b) \quad , \quad \Pr(a < X \leqslant b) \quad \text{and} \quad \Pr(a < X < b)$$

- $f(x)$ can exceed one!

Transformations of Random Variables

Theorem If X is a r.v. with sample space $\mathcal{X} \subset \mathbb{R}$ and cdf $F_X(x)$, then any function of X, say $Y = g(X)$ is also a random variable. The new random variable Y has a new sample space $\mathcal{Y} = g(X) \subset \mathbb{R}$. The objective is to find the cdf $F_Y(y)$ of Y.

Probability mapping: For any set $A \subset \mathcal{Y}$:

$$\begin{aligned} \Pr(Y \in A) &= \Pr(g(X) \in A) \\ &= \Pr(\{x \in \mathcal{X} : g(x) \in A\}) \\ &= \Pr(X \in g^{-1}(A)), \end{aligned}$$

where we have defined

$$g^{-1}(A) = \{x \in \mathcal{X} : g(x) \in A\}.$$

Notice that $g^{-1}(A)$ is well defined even if $g(\cdot)$ is not necessarily bijective.

Example (Binomial transformation) A discrete random variable X has a *binomial distribution* if its pmf is of the form

$$f_X(x) = \Pr(X = x) = \binom{n}{x} p^x (1-p)^{n-x}, \quad x = 0, 1, \ldots, n,$$

where n is a positive integer and $0 \leqslant p \leqslant 1$. Values such as n and p that can be set to different values, producing different probability distributions, are called *parameters*. Consider a random variable $Y = g(X)$, where $g(x) = n - x$; that is, $Y = n - X$. Here $\mathcal{X} = \{0, 1, \ldots, n\}$

and $\mathcal{Y} = \{y : y = g(x), x \in \mathcal{X}\} = \{0, 1, \ldots, n\}$. For any $y \in \mathcal{Y}$, $n - x = g(x) = y$ if and only if $x = n - y$. Therefore, $g^{-1}(y) = n - y$ and

$$\begin{aligned} f_Y(y) &= \sum_{x \in g^{-1}(y)} f_X(x) \\ &= f_X(n-y) \\ &= \binom{n}{n-y} p^{n-y}(1-p)^{n-(n-y)} \\ &= \binom{n}{y}(1-p)^y p^{n-y}. \end{aligned}$$

Therefore, Y also has a binomial distribution, but with parameters n and $1-p$.

Example (exercise 2.3) Suppose X has the geometric pmf $f_X(x) = \frac{1}{3}(\frac{2}{3})^x, x = 0, 1, 2, \ldots$. Determine the probability distribution of $Y = X/(X+1)$. Note that here both X and Y are discrete random variables. To specify the probability distribution of Y, specify its pmf.
Solution: