3.1 Continuous Random Variables

Today's Objectives

- What is the difference between a discrete and continuous distributions?
- What is a pdf?
- What is the cdf in the continuous case? How can we calculate the formula for a given distribution?
- How do we calculate probabilities for a continuous random variable?
- How do we calculate E[X], Var[X], mgf, percentiles for continuous random variables?

Fundamental Theorem of Calculus (review)

Fundamental Theorem of Calculus, Part 1

If $f(x)$ is continuous over an interval $[a, b]$, and the function $F(x)$ is defined by

$$F(x) = \int_a^x f(t)\, dt,$$

then $F'(x) = f(x)$ over $[a, b]$.

Types of Random Variables

Sample spaces belong to one of two types:

- Discrete
- Continuous

Discrete: Rolling die and counting the total number of spots

- Countable (can be countably infinite)

Continuous: Choosing a random number from the interval [0,1]

- Uncountable

How we calculate probabilities

Discrete: Each outcome in the sample space is assigned a probability by the pmf, $f(x)$. $P(X = x) = f(x)$

Continuous: This no longer works because a continuous space has an uncountably infinite number of outcomes.

We need to talk about the probability that X falls within some interval or range. For all x, $P(X = x) = 0$

Probability density function (pdf)

A **continuous random variable**, X, can be described with a **probability density function**, $f(x)$.

$f(x)$ is an integrable function that satisfies the following three conditions:

(a) $f(x) \geq 0, \qquad x \in S.$

(b) $\int_S f(x)\,dx = 1.$

(c) If $(a, b) \subseteq S$, then the probability of the event $\{a < X < b\}$ is

$$P(a < X < b) = \int_a^b f(x)\,dx.$$

Continuous distribution example

$$f(x) = \begin{cases} k(x^2 + x), & 0 \le x \le 1, \\ 0 \quad\quad\quad\quad , & elsewhere \end{cases}$$

Find the value *k* that makes *f*(*x*) a probability density function (pdf).

$$\begin{aligned} 1 &= \int_{-\infty}^{\infty} f(x)dx \\ &= k\int_0^1 (x^2 + x)dx \\ &= k\left(\frac{x^3}{3} + \frac{x^2}{2}\right)\Big|_0^1 \\ &= k\left(\frac{5}{6}\right) \end{aligned}$$

Cumulative Distribution Function (cdf)

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(t)d(t), \qquad -\infty \leq x \leq \infty$$

Based on the fundamental theorem of calculus,

$$F'(x) = f(x)$$

cdf example

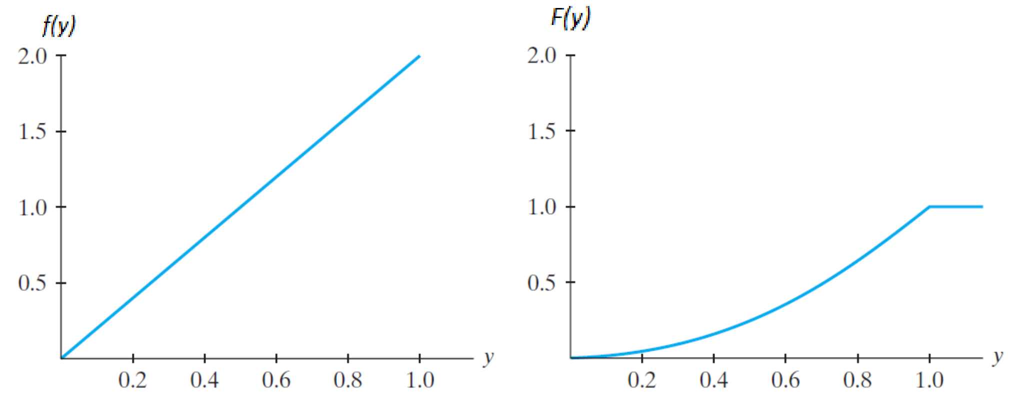

Figure 3.1-2 Continuous distribution pdf and cdf

Example 3.1-1 Let Y be a continuous random variable with pdf $f(y) = 2y$, $0 < y < 1$. The cdf of Y is defined by

$$F(y) = \begin{cases} 0, & y < 0, \\ \displaystyle\int_0^y 2t\,dt = y^2, & 0 \le y < 1, \\ 1, & 1 \le y. \end{cases}$$

Calculating continuous probabilities

$f(y) = 2y,\ 0 < y < 1.$ *Calculate* P[½ < Y < ¾]

Using the CDF:

F(¾) – F(½) = (¾)^2 – (½)^2 = **5/16**

Calculating continuous probabilities

$f(y) = 2y,\ 0 < y < 1.$ $\quad Calculate$ P[¼ < Y < 2].

Using the cdf:

F(2) – F(¼) = 1 – (¼)^2 = **15/16**

cdf example

$$f(x) = \begin{cases} k(x^2 + x), & 0 \le x \le 1, \\ 0 \qquad\qquad , & elsewhere \end{cases}$$

Find an expression for the cdf, $F(x)$.

When $0 \le x \le 1$,

$$\begin{aligned} F(x) &= \int_{-\infty}^{x} f(t)dt \\ &= \int_{-\infty}^{0} f(t)dt + \int_{0}^{x} f(t)dt \\ &= 0 + \frac{6}{5}\left(\frac{x^3}{3} + \frac{x^2}{2}\right) \\ &= \frac{6}{5}\left(\frac{x^3}{3} + \frac{x^2}{2}\right) \end{aligned}$$

$$F(x) = \begin{cases} 0 & , x < 0 \\ \frac{6}{5}\left(\frac{x^3}{3} + \frac{x^2}{2}\right) & , 0 \le x \le 1 \\ 1 & , x > 1 \end{cases}$$

E[X] and Var[X] for Continuous RV

$$\mu = E(X) = \int_{-\infty}^{\infty} x f(x)\, dx$$

$$\sigma^2 = \mathrm{Var}(X) = E[(X - \mu)^2] = \int_{-\infty}^{\infty} (x - \mu)^2 f(x)\, dx$$

MGF for Continuous RV

The mgf is still $E[e^{tX}]$:

$$M(t) = \int_{-\infty}^{\infty} e^{tx} f(x)\, dx$$

Percentiles

The **(100p)th percentile** is a number π_p such that the area under $f(x)$ to the left of π_p is p. That is,

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

- For example, the 60th percentile is the number, π_p $(or\ x)$, such that $F(\pi_p)$ = 0.6

Continuous Example

Suppose f_X(x) $= 4x^3, 0 \leq x \leq 1$.

Find $P(0 \leq X \leq \frac{1}{2})$ ans: (1/16)

Find E[X]. ans: (4/5)

Find the 20th percentile of X. ans: (0.6687)

Properties of continuous pdf/cdf

The pdf for a continuous random variable does **not** need to be a continuous function. For example:

$$f(x) = \begin{cases} \dfrac{1}{2}, & 0 < x < 1 \quad \text{or} \quad 2 < x < 3, \\ 0, & \text{elsewhere,} \end{cases}$$

The cdf will always be a continuous function.

Cdf of previous distribution

$$f(x) = \begin{cases} \frac{1}{2}, & 0 < x < 1 \quad \text{or} \quad 2 < x < 3, \\ 0, & \text{elsewhere}, \end{cases}$$

$$F(x) = \begin{cases} 0, & x < 0 \\ \frac{1}{2}x, & 0 < x < 1 \\ \frac{1}{2}, & 1 \leq x \leq 1 \\ \frac{1}{2} + \frac{1}{2}(x-2), & 2 < x < 3 \\ 1, & x \geq 3 \end{cases}$$

Note: the use of "$\leq$" and "$<$" are interchangeable for continuous distributions. Why?

A quick note about continuous pdfs

A continuous pdf does not need to be bounded above

- (that means $f(x)$ for a continuous RV can be larger than 1)
- (Remember for discrete RVs, the pmf $f(x)$ is bounded by 1.)

The **area** between a continuous pdf and the x-axis must still equal 1. (total probability)