Part 1: Continuous Random Variables

Section 4.1 Continuous Random Variables
Section 4.2 Probability Distributions & Probability Density Functions
Section 4.3 Cumulative Distribution Functions
Section 4.4 Mean and Variance
Moving on to continuous random variables...

For a continuous random variable $X$, the range of $X$ includes all values in an interval of real numbers. This could be an infinite interval such as $(-\infty, \infty)$.

For example, let $X$ be a random variable such that $X$ can take on any value in the interval $[0, 1]$. Then, $X$ is a continuous random variable.
Continuous Random Variables

- Since there is an infinite number of possible values for $X$, we describe the probability distribution with a smooth curve.

Example (Continuous random variable)

Suppose $X$ is a random variable such that $X \in [0, 1]$, and there is a high probability that $X$ is near 0.15 and a small probability that $X$ is 0 or 1. Specifically, the distribution of the random variable is shown as:

This curve is called a probability density function.
Continuous Random Variables

- We can contrast this probability distribution with that of a discrete random variable which has mass at only ‘distinct’ $x$-values.

- **The area under a probability density function is 1.** This compares to the sum of the masses for a discrete random variable being equal to 1.

![Graph of a probability density function]

- The probability density function is denoted as $f(x)$, same notation is the probability mass function, as $f(x)$ describes the distribution of a random variable.
How likely is it that $X$ falls between 0.18 and 0.22?

We can answer this question by considering the area under the curve between 0.18 and 0.22.

This area is approximately 0.1181 for the given $f(x)$ above.
If \( f(x) \) is a known function, such as, \( f(x) = C \cdot x^{\alpha-1}(1 - x)^{\beta-1} \),

which is the case for the above where \( C, \alpha, \beta \) are known constants, then we could answer this question through integration:

\[
P(0.18 \leq x \leq 0.22) = \int_{0.18}^{0.22} f(x) \, dx
\]
Probability Density Function

Definition (Probability Density Function)

For a continuous random variable $X$, a **probability density function** is a function such that

1. $f(x) \geq 0$ (above the horizontal axis)
2. $\int_{-\infty}^{\infty} f(x) \, dx = 1$ (area under curve is equal to 1)
3. $P(a \leq x \leq b) = \int_{a}^{b} f(x) \, dx$ = area under $f(x)$ from $a$ to $b$ for any $a$ and $b$

- A probability density function is zero for $x$ values that cannot occur, and it is assumed to be zero wherever it is not specifically defined.

- We use $f(x)$ to calculate a probability through integration.
In the continuous case, every distinct \( x \)-value has zero width (there’s infinitely many of them), and the probability for a single specific \( x \)-value is zero...

\[
P(X = x) = 0.
\]

Instead, we find probabilities for intervals of the random variable, not singular specific values, like \( P(0.18 \leq X \leq 0.20) \).

For example, consider weight.

Without the rounding, we see weight as a continuous variable, and there are an infinite number of possible weights in \((0, \infty)\). On this perception, we could show the distribution of weights using a probability density function.
With rounding to the nearest pound, when we see a weight of 45 on the scale, there are an infinite number of possible weights that end up with this scale weight... anything between 44.5 and 45.5 pounds.

The probability of observing 45 on the scale can be found by integrating the probability density function $f(x)$ of the continuous variable weight $X$ over the range $44.5 \leq x \leq 45.5$.

Because no singular $x$-value has positive mass, the difference between $\leq$ and $<$ doesn’t matter for the continuous variable...

If $X$ is a continuous random variable, for $x_1$ and $x_2$,

\[
P(x_1 \leq X \leq x_2) = P(x_1 < X \leq x_2) = P(x_1 \leq X < x_2) = P(x_1 < X < x_2)
\]
There are many common probability density functions with known $f(x)$.

- Normal Density
- Uniform Density
- Exponential Density
- Gamma Density
- Log-normal Density
- Weibull Density
But they can also take on *unusual* shapes too...

**Example (Proportion of people returning a survey)**

The proportion of people who respond to a certain mail-order solicitation is a continuous random variable $X$ that has density function

$$f(x) = \begin{cases} \frac{2(x+2)}{5} & 0 < x < 1 \\ 0 & \text{otherwise} \end{cases}$$

1. Plot $f(x)$
2. Show that $P(0 < x < 1) = 1$
3. Find the probability that more than 1/4 but fewer than 1/2 of the people contacted will respond to this type of solicitation.
Example (Proportion of people returning a survey, cont.)

1) Plot $f(x)$

2) Show that $P(0 < x < 1) = 1$
### Example (Proportion of people returning a survey, cont.)

3) Find the probability that more than $\frac{1}{4}$ but fewer than $\frac{1}{2}$ of the people contacted will respond to this type of solicitation.
Example (Time to failure)

The probability density function of the time to failure of an electronic component in a copier (in hours) is

\[
f(x) = \begin{cases} 
\frac{1}{3000} e^{-x/3000} & x > 0 \\
0 & \text{otherwise}
\end{cases}
\]

(a) What does \( f(x) \) look like?

(b) Determine the probability that a component fails in the interval from 1000 to 2000 hours.

(c) At what time do we expect 10% of the components to have failed
Example (Time to failure)

\[ f(x) = \begin{cases} \frac{1}{3000} e^{-x/3000} & x > 0 \\ 0 & \text{otherwise} \end{cases} \]

(a)

(b) Determine the probability that a component fails in the interval from 1000 to 2000 hours.
Example (Time to failure, cont.)

\[ f(x) = \begin{cases} 
\frac{1}{3000}e^{-x/3000} & x > 0 \\
0 & \text{otherwise}
\end{cases} \]

(c) At what time do we expect 10% of the components to have failed
Example (Fantasy football point distributions)

Alvin Karama (RB)

- **BOOM CHANCE**: 19.4% (≥23.8 pts)
- **BUST CHANCE**: 20.0% (<13.6 pts)

Jake Elliott (K)

- **BOOM CHANCE**: 25.1% (≥10.9 pts)
- **BUST CHANCE**: 9.7% (<4.7 pts)

NOTE: Don’t look at the y-axis stated ‘probability %’, not relevant as probabilities are areas under the curve.
As we did with discrete random variables, we often want to compute the cumulative probabilities, or \( P(X \leq x) \).

**Definition (Cumulative Distribution Function)**

The cumulative distribution function (CDF) of a continuous random variable \( X \) is

\[
F(x) = P(X \leq x) = \int_{-\infty}^{x} f(u)du
\]

for \(-\infty < x < \infty\).

To find \( F(x) \), find the anti-derivative of \( f(x) \).

And vice versa...

If you’re given \( F(x) \), you can find \( f(x) \) through differentiation as \( f(x) = \frac{dF(x)}{dx} \).
### Example (Uniform prob. density function)

Let the continuous random variable $X$ denote the current measured in a thin copper wire in milliamperes. Assume that the range of $X$ is $[0, 20\ mA]$, and assume that the probability density function of $X$ is $f(x) = 0.05$ for $x \in [0, 20]$, and $f(x) = 0$ o.w.

(a) Use the cumulative distribution function to determine what proportion of the current measurements is less than 10 mA, i.e. find $P(X \leq 10) = F(10)$.

**ANS:** First, find $F(x) = P(X \leq x)$.

For this problem, the CDF consists of three expressions related to the three pieces of the domain as $(-\infty, 0), [0, 20], (20, \infty)$. This segmentation occurs because $f(x) = 0$ when $x < 0$ and $x > 20$. 
Cumulative Distribution Functions

Example (Uniform prob. density function, cont.)

We know $F(x)=0$ for $x<0$ (no probability accumulated left of $x = 0$) and $F(x)=1$ for $x \geq 20$ (all probability accumulated by $x = 20$).

Now we just need to define $F(x)$ in $[0, 20]$...

$$[0, 20] : F(x) = P(X \leq x) = \int_0^x f(u)du$$

$$= \int_0^x 0.05du$$

$$= 0.05u\bigg|_0^x = 0.05x$$

Thus,

$$F(x) = \begin{cases} 
0 & x < 0 \\
0.05x & 0 \leq x \leq 20 \\
1 & x > 20 
\end{cases}$$
Example (Uniform prob. density function, cont.)

Now use the function $F(x)$ to calculate $P(X \leq 10) = F(10)$

$F(10) = 0.05 \times 10 = 0.50$

Or, 50% of the probability is accumulated at the midway point of $x = 10$ for this uniform distribution on the interval $[0, 20]$.

(b) Graph the cumulative distribution function (CDF) for this uniform probability density.
A cumulative distribution function (CDF) describes how probability is accumulated as $x$ goes from left to right or from $-\infty$ to $+\infty$.

In the uniform distribution case above, we accumulate probability at a constant rate (thus, the straight line).
Cumulative Distribution Functions

Example (Recall the Time to Failure example...)

The probability density function of the time to failure of an electronic component in a copier (in hours) is

\[
f(x) = \begin{cases} 
\frac{1}{3000}e^{-x/3000} & x > 0 \\
0 & \text{otherwise}
\end{cases}
\]

Determine the cumulative distribution function (CDF) from \( f(x) \) and use it to calculate when 50% of the components have failed.
Cumulative Distribution Functions

Example (Recall the Time to Failure example…)

**ANS:** Determine the CDF or \( F(x) = P(X \leq x) \)

\((-\infty, 0) : F(X) = 0 \)

\([0, \infty) : F(X) = \)

Find \( x \) such that \( F(x) = P(X \leq x) = 0.50 \).
The mean and variance of continuous random variables can be computed similar to those for discrete random variables, but for continuous random variables, we will be integrating over the domain of $X$ rather than summing over the possible values of $X$.

**Definition (Mean and Variance of Continuous Random Variable)**

Suppose $X$ is a continuous random variable with probability density function $f(x)$. The mean or expected value of $X$, denoted as $\mu$ or $E(X)$, is

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

Definition (Mean and Variance of a Continuous Random Variable)

The variance of $X$, denoted as $V(X)$ or $\sigma^2$, is

$$\sigma^2 = V(X) = E(X - \mu)^2 = E(X^2) - (E X)^2 = \int_{-\infty}^{\infty} (x - \mu)^2 f(x) dx = \int_{-\infty}^{\infty} x^2 f(x) dx - \mu^2$$

The standard deviation of $X$ is $\sigma = \sqrt{\sigma^2}$
Example (Mean and Variance of a Continuous Random Variable)

For the uniform probability density function described earlier, \( f(x) = 0.05 \) for \( 0 \leq x \leq 20 \), compute \( E(X) \) and \( V(X) \).

ANS:
Continuous Random Variables

Definition (Expected Value of a Function)

If \( X \) is a continuous variable with probability density function \( f(x) \),

\[
E[h(x)] = \int_{-\infty}^{\infty} h(x) f(x) \, dx
\]

Example (Weight of delivered packages (p. 115 problem 4-36))

The probability density function of the weight of packages delivered by a post office is

\[
f(x) = \frac{70}{(69x^2)} \text{ for } 1 < x < 70 \text{ pounds.}
\]

If the cost is $2.50 per pound, what is the mean shipping cost of a package?
Example (Weight of delivered packages (p. 115 problem 4-36), cont.)

\[ f(x) = \frac{70}{69x^2} \text{ for } 1 < x < 70 \text{ pounds.} \]

If the cost is $2.50 per pound, what is the mean shipping cost of a package?

ANS: