Chapter 4
Continuous Random Variables and Probability Distributions

Part 2: More on Continuous Random Variables

Section 4.4 Continuous Uniform Distribution
Section 4.5 Normal Distribution
Continuous and Discrete Random Variables

**Continuous Random Variable**

- $X$ can take on all possible values in an interval of real numbers.
  
e.g. $X \in [0, 1]$

- Probability density function, $f(x)$

- Cumulative distribution function, $F(x) = P(X \leq x) = \int_{-\infty}^{x} f(u) \, du$

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

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

**Discrete Random Variable**

- $X$ can take on only distinct ‘discrete’ values in a set.
  
e.g. $X \in \{0, 1, 2, 3, \ldots, \infty\}$

- Probability mass function, $f(x)$

- Cumulative distribution function, $F(x) = P(X \leq x) = \sum_{x_i \leq x} f(x_i)$

- $\mu = E(X) = \sum_{x} x \, f(x)$

- $\sigma^2 = V(X) = E(X - \mu)^2$
  
  \[ = \sum_{x} (x - \mu)^2 \, f(x) \]
  
  \[ = E(X^2) - [E(X)]^2 \]
  
  \[ = \sum_{x} x^2 \, f(x) - \mu^2 \]
Continuous Uniform Distribution

- The simplest continuous distribution
- $X$ falls between $a$ and $b$.
- It’s uniformly distributed over the interval $[a, b]$.
- $f(x)$ has a constant value, and $f(x) = \frac{1}{b-a}$
- This coincides with the area under the curve being 1.

Example (Uniform(2,4))
Continuous Uniform Distribution

**Definition (Continuous Uniform Distribution)**

A continuous random variable $X$ with probability density function

$$f(x) = \frac{1}{b-a}, \quad a \leq x \leq b$$

is a continuous uniform random variable.

**Definition (Mean and Variance for Continuous Uniform Dist’n)**

If $X$ is a continuous uniform random variable over $a \leq x \leq b$

$$\mu = E(X) = \frac{(a+b)}{2},$$

and

$$\sigma^2 = V(X) = \frac{(b-a)^2}{12}$$
Example (Uniform(0,20))

For the uniform probability density function described earlier with $a = 0$ and $b = 20$, $f(x) = \frac{1}{20} = 0.05$ for $0 \leq x \leq 20$.

Find $E(X)$ and $V(X)$ using the formulas.

**ANS:** $a = 0$, $b = 20$

\[
\mu = E(X) = \frac{(0+20)}{2} = 10
\]

\[
\sigma^2 = V(X) = \frac{(20-0)^2}{12} = 33.33
\]
Continuous Uniform Distribution

Example (Shampoo bottle volume)

The volume, $X$, of shampoo filled into a container is uniformly distributed between 374 and 380 milliliters.

1) Find the cumulative distribution function (CDF) for $X$.

2) Use the CDF to find the volume of shampoo that is exceeded by 95% of all the volumes (i.e., the threshold for the lowest 5%).

3) Graph $F(x)$.

ANS: 1) $F(x) =$?
Example (Shampoo bottle volume, cont.)

**ANS:** 2) Use the CDF to find the volume of shampoo that is exceeded by 95% of all the volumes (i.e. the threshold for the lowest 5%).

**ANS:** 3) Graph $F(x)$. 
Normal Distribution

- Perhaps the most widely used distribution of a random variable.
- Arises naturally in physical phenomena.
- Two parameters completely define a normal probability density function, $\mu$ and $\sigma^2$.
- $\mu$ is the expected value, or center of the distribution.
- $\sigma^2$ is the variance of the distribution, and quantifies spread.
- Symmetrical distribution.
A normal distribution can occur anywhere along the real number line.

It always has a bell-shape.

The parameter $\mu$ tells us where it is centered, and where there’s a high probability of $X$ occurring.

$\sigma^2$ tells us how spread-out the distribution is.

Recall that the area under the curve must be a 1.

**Figure 4-10** Normal probability density functions for selected values of the parameters $\mu$ and $\sigma^2$. 
Normal Distribution: 68-95-99.7 Rule

- Special result of normal distribution:
  - Recall: $\sigma$ is the standard deviation of $X$, $\sigma = \sqrt{V(X)}$
  - 68% of the observations lie within 1 std. deviation of the mean.
  - 95% of the observations lie within 2 std. deviation of the mean.
  - 99.7% of the observations lie within 3 std. deviation of the mean.
  - Very little area under the curve lies beyond $3\sigma$ away from the mean.
Example (Weight of contents in cereal box)

A box of Quazar cereal states there are 15 oz. of cereal in a box.

In reality, the amount of cereal in a box varies from box to box. Suppose the amount has a normal distribution with $\mu = 15$ and $\sigma^2 = 0.04$.

What percentage of boxes have between 14.6 oz and 15.4 oz. of cereal?

ANS:
What about computing probabilities for values other than 
\( \mu \pm 1\sigma, \mu \pm 2\sigma, \mu \pm 3\sigma \)

Definition (Normal distribution)

A random variable \( X \) with probability density function

\[
f(x) = \frac{1}{\sqrt{2\pi}\cdot\sigma} \cdot e^{-\frac{1}{2\sigma^2}(x-\mu)^2} \quad -\infty < x < \infty
\]

is a normal random variable with parameters \( \mu \) and \( \sigma \), where \( -\infty < \mu < \infty \), and \( \sigma > 0 \), and \( \pi = 3.14159 \ldots \) and \( e = 2.71828 \ldots \)

Also, \( E(X) = \mu \) and \( V(X) = \sigma^2 \)

The notation \( \mathcal{N}(\mu, \sigma^2) \) will be used to denote the distribution.
Example (Weight of contents in cereal box)

What percent of boxes contain less than 14.5 oz. of cereal?

\[ P(X \leq 14.5) = F(14.5) = \int_{-\infty}^{14.5} f(x) \, dx \]

\[ = \int_{-\infty}^{14.5} \frac{1}{\sqrt{2\pi} \cdot \sigma} e^{-\frac{1}{2\sigma^2}(x-\mu)^2} \, dx \]

\[ = \frac{1}{\sqrt{2\pi} \cdot (0.2)} \int_{-\infty}^{14.5} e^{-\frac{1}{2(0.04)}(x-15)^2} \, dx \]

This cannot be done in closed form, instead we'll use statistical tables (Appendix A: A-8 and A-9) to calculate.
There is an infinite number of distinct normal distributions (any $\mu$ and $\sigma^2$ define one).

But, we only need one statistical table to compute probabilities for EVERY normal.

This is because every normal distribution can be shifted and scaled (i.e. stretched or shrunk) to look like the Standard Normal Distribution (shown below).
Normal Distribution: Standard Normal $N(0, 1)$

**Definition (Standard Normal Distribution)**

A normally distributed random variable with

$$
\mu = 0 \quad \text{and} \quad \sigma^2 = 1
$$

is a standard normal random variable and is denoted as $Z$.

- We say $Z$ is distributed $N(0, 1)$, or $Z \sim N(0, 1)$.

- The cumulative distribution function, $F(z) = P(Z \leq z)$, of a standard normal random variable is denoted as

$$
\Phi(z) = P(Z \leq z)
$$
The Standard Normal Distribution

\[ \Phi(1.5) = P(Z \leq 1.5) = 0.93319 \]

Appendix A: Cumulative Standard Normal Distribution

(A-8 and A-9 in Appendix of book).

Table for determining cumulative probabilities for \( Z \):

<table>
<thead>
<tr>
<th>( z )</th>
<th>0.00</th>
<th>0.01</th>
<th>0.02</th>
<th>0.03</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.50000</td>
<td>0.50399</td>
<td>0.50398</td>
<td>0.51197</td>
</tr>
<tr>
<td>\vdots</td>
<td>\vdots</td>
<td>\vdots</td>
<td>\vdots</td>
<td>\vdots</td>
</tr>
<tr>
<td>1.5</td>
<td>0.93319</td>
<td>0.93448</td>
<td>0.93574</td>
<td>0.93699</td>
</tr>
</tbody>
</table>
Example (Standard Normal Distribution)

Find \( P(Z \leq 1.52) = \Phi(1.52) \).

**ANS:** See Appendix A: A-8 and A-9.

Find row and column for \( z \)-value of \( z=1.52 \).

\[
P(Z \leq 1.52) = 0.93574
\]

The table provides cumulative distributions. These are areas under the normal curve of \( f(x) \) to the left of a given \( z \)-value.
Example (Standard Normal Distribution)

Find \( P(Z \leq -1.25) = \Phi(-1.25) \).

**ANS:** From Table III, \( P(Z \leq -1.25) = 0.10565 \)
Example (Standard Normal Distribution)

Find \( P(Z > 1.26) \).

**ANS:**

\[
P(Z > 1.26) = 1 - P(Z \leq 1.26)
\]
\[
= 1 - 0.89616
\]
\[
= 0.10384
\]
Example (Standard Normal Distribution)

Find $P(-1.25 \leq Z \leq 0.37)$.

**ANS:** $P(-1.25 \leq Z \leq 0.37) = P(Z \leq 0.37) - P(Z \leq -1.25)

= 0.64431 - 0.10565

= 0.53866$
How do we compute probabilities for our cereal example?

For \( X \sim N(15, 0.2^2) \), how do we use the table to find \( P(X \leq 14.5) \)?

We first **shift** the random variable to be centered at 0 (i.e. subtract the mean).

\[
X^* = X - \mu = X - 15
\]

Then, we **scale** it to have a standard deviation of 1 (i.e. divide by the standard deviation).

\[
Z = X^{**} = \frac{X^*}{\sigma} = \frac{X-\mu}{\sigma} = \frac{X-15}{0.2}
\]

After this shift and scale phrased as “subtract the mean, divide by the standard deviation”, then this new variable \( X^{**} = \frac{X-\mu}{\sigma} \) is a \( Z \) random variable, or a standard normal random variable, or a \( N(0, 1) \) random variable.
Definition (Standardizing a Normal Random Variable)

If $X$ is a normal random variable with $E(X) = \mu$ and $V(X) = \sigma^2$, then the random variable

$$Z = \frac{X - \mu}{\sigma}$$

is a normal random variable with $E(Z) = 0$ and $V(Z) = 1$. That is, $Z$ is a standard normal random variable.

$Z$ represents “the distance $X$ is from its mean” in terms of the number of standard deviations.
Normal Distribution: Standardizing

- Let \( X \sim N(10, 2^2) \) \{i.e. \( X \) is not a std. normal r.v. \}

- “Subtract the mean, divide by the standard deviation”, \( Z = \frac{X - \mu}{\sigma} \)

\[
P(X \leq 13) = P \left( \frac{X - \mu}{\sigma} \leq \frac{13 - 10}{2} \right)
= P(Z \leq 1.5)
= 0.9332 \quad \{\text{from Table III}\}
\]

*Figure 4-15  Standardizing a normal random variable.*
Normal Distribution: Standardizing

Standardizing to Calculate a Probability

- Suppose $X$ is a normal random variable with mean $\mu$ and variance $\sigma^2$ or $X \sim N(\mu, \sigma^2)$, then

$$P(X \leq x) = P\left(\frac{X-\mu}{\sigma} \leq \frac{x-\mu}{\sigma}\right) = P(Z \leq z)$$

where $Z$ is a standard normal random variable, and $z = \frac{(x-\mu)}{\sigma}$ is the z-value obtained by standardizing $X$.

Then, we obtain probabilities from $Z$–table or Table III.

- Again, there are an infinite number of normal distributions, but we only need one table since any $N(\mu, \sigma^2)$ can be related to the $N(0, 1)$. 
Back to the cereal example...

What percent of boxes contain less than 14.5 oz. of cereal?

Recall that the amount in a cereal box is normally distributed with mean 15 oz. and standard deviation of 0.2 oz.

ANS:
Example (Finding a percentile of the cereal box distribution)

Find the cereal box amount (in oz.) at which 10% of the cereal boxes have less than this much cereal (i.e. find the threshold at which 10% of the boxes fall below this amount).

**ANS:** Recall, \( X \sim N(15, 0.2^2) \) <\textit{sketch here}>

First, on the \( Z \)-table, find the probability 0.10 in the \textit{middle} of the table.

Then, find the \( z \)-value (specific row and column) that coincides with this probability. You have found ‘\( z \)’ such that \( P(Z < z) = 0.10 \). This can be denoted using inverse function notation for the CDF as \( \Phi^{-1}(0.10) = z \).

You should find \( \Phi^{-1}(0.10) = -1.28 \).
Example (Finding a percentile of the cereal box distribution)

The $z$-value sought after goes where the $?$ is at:

$$P(Z \leq ?) = 0.10$$

$$P(Z \leq -1.28) = 0.10 \quad \text{or} \quad \Phi^{-1}(0.10) = -1.28$$

relevant $z$-value$= -1.28$

“Unstandardize” the $z$-value to get the $x$-value:

$$z = \frac{x-\mu}{\sigma} \quad \Rightarrow \quad x = \mu + z\sigma$$

$$x = 15 + (-1.28)(0.2)$$

$$x = 14.744$$