1. Continuous random variables
2. Normal distribution
3. Transformation of random variables
Continuous Random Variables

A continuous random variable $X$:

- Is a function that maps the sample space to real numbers
- Is defined on intervals; not on distinct points (discrete r.v.’s)
  - There are infinitely many points within any interval
  - We cannot assign probability mass to each point!

Examples:

- Annual income for a family
- A company’s stock price for a share
- A person’s height or weight
Recall that we can discretize a continuous variable into categories/groups.

Here we use intervals of length 1 to create a discrete r.v. with a probability mass function.
Recall that because every value within an interval is possible, the *height* $h_i$ of each bin is determined by:

1. $f_i = \text{the relative frequency associated to a given class } C_i$
2. $\Delta_i = \text{the length of class } C_i$

Specifically,

$$h_i = \frac{f_i}{\Delta_i}$$

Note that $f_i$ can be interpreted as the area of each bin.

Also, the probability of all groups must add up to 1

$$\sum_i h_i \times \Delta_i = \sum_i f_i = 1$$
We can make the interval length as small as possible to discretize a continuous distribution.
What happens if the interval $\Delta_i$ gets extremely small?

- The height of the bins becomes smoother (almost like a continuous function)!
- If $\Delta_i \to 0$ then the relative frequency (the area of the rectangle) gets very small

$$\lim_{\Delta_i \to 0} \Delta_i \times h_i = f_i = 0$$

The probability of a continuous random variable grouped on an infinitely small interval is equal to zero!

- Loosely, we call $h_i$ the **density** of the random variable
Probability Density Function

Suppose we want to know the following probability of an interval \([a, b]\)

\[ P(a < X < b) \]

- We can approximate it by using a large number of small bins \(C_i\) with length \(\Delta_i\) and add up all the rectangles between \(a\) and \(b\)

\[
P(a < X < b) \approx \sum_{C_i \in [a,b]} \Delta_i \times h_i = \sum_{C_i \in [a,b]} f_i
\]

- If the interval shrinks to zero then

\[
\lim_{a \to b} P(a < X < b) = P(a) = P(b) = 0!!!
\]
Probability Density Function

For continuous random variable

- \( p(x) \) is called **probability density function (pdf)**
  - It represents how dense the probability is at point \( x \)
  - It does not indicate the probability of \( x \)

- Intuitively, every single exact value, \( X = x \) has an infinitesimally small probability (therefore zero)

- Think about the histogram! If the number of breaks increases (e.g. \( \infty \)), the frequencies go to zero
Let $A$ be some interval on the real line; the probability of $P(A)$ is defined by the area under the curve traced by $p(x)$. 
$F(x_0)$ is called cumulative distribution function (cdf).

$$F(x_0) = P(X \leq x_0) = \int_{-\infty}^{x_0} p(x) \, dx$$

$F(x_0)$ is simply the area under the density $p(x)$ up to $x_0$. 
Cumulative Distribution Function

$F(x_0)$ is very useful for calculating probabilities in regions. Examples:

$$P(a \leq X \leq b) = P(X \leq b) - P(X \leq a) = F(b) - F(a)$$
This expression for $P(a \leq X \leq b)$ can also be found by:

- Start with $P(X \leq b) = F(b)$. By total probability,

$$P(X \leq b) = P((X \leq b) \cap (X \leq a)) + P((X \leq b) \cap (X > a))$$

- Since $a < b$, $(X \leq b) \cap (X \leq a) = (X \leq a)$

- Also, $(X \leq b) \cap (X > a) = (a \leq X \leq b)$

$$P(X \leq b) = P(X \leq a) + P(a \leq X \leq b)$$

$$F(b) = F(a) + P(a \leq X \leq b)$$

$$P(a \leq X \leq b) = F(b) - F(a)$$
Each continuous random variable $X$ is associated with a density function $p(x)$ and a cumulative distribution function $F(x)$.

Because the total probability must be 1:

$$\int_{-\infty}^{\infty} p(x) \, dx = 1$$

$$\lim_{a \to \infty} F(a) = 1 \quad \text{and} \quad \lim_{a \to -\infty} F(a) = 0$$

$$P(X \geq a) = 1 - F(a)$$
Consider the following density of $X$:

$$p(x) = x + \frac{1}{2} \quad 0 \leq x \leq 1$$

What is the probability for:

1. $P(0 \leq X \leq 1)$
2. $P(X \leq 0.2)$
3. $P(X \geq 0.5)$
4. $P(0.4 \leq X \leq 0.6)$
The CDF of the previous example is

\[ F(x_0) = \frac{1}{2} (x_0^2 + x_0) \]

Note that \( F(0) = 0 \) and \( F(1) = 1 \).

Calculate the previous probabilities using \( F(x) \).

1. \( P(0 \leq X \leq 1) = F(1) - F(0) = 1 \)
2. \( P(X \leq 0.2) = F(0.2) = 0.12 \)
3. \( P(X \geq 0.5) = 1 - F(0.5) = 0.625 \)
4. \( P(0.4 \leq X \leq 0.6) = F(0.6) - F(0.4) = 0.2 \)
The Gaussian (Normal) Distribution

- The famous bell-curve
- First introduced by Abraham de Moivre in 1733
- Pierre-Simon Laplace used the normal distribution in the analysis of errors of experiments
- Approximately normal distributions occur in many situations
  - Measurement errors (repeated measurements of the same quantity)
  - Financial variables (stock returns, with some modifications)
  - Distribution in testing and intelligence (IQ)
- Galton Board [http://www.youtube.com/watch?v=6YDHBFVIvIs](http://www.youtube.com/watch?v=6YDHBFVIvIs)
Normal distribution is the bell-shaped curve. It is defined over all possible real numbers. The probability density function is:

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(x-\mu)^2}$$

with the constants

- $\mu = \text{mean}$
- $\sigma^2 = \text{variance}$
- $\pi \approx 3.14159$ and $e \approx 2.71828$. 
Normal random variables with variance 1 and different means
Normal Distribution: SDs

Normal random variables with mean 0 and different variances

SD
1
2
4

Density

Thais Paiva  STA 111 - Summer 2013 Term II  Lecture 6, 07/09/2013
*The orange line also has SD = 2. The legend is incorrect.
In the previous graph, note that

- The CDF is non-decreasing
- Smaller variance → more steep
- Larger mean → CDF shifts horizontally to the right
- Mean/Mode/Median of $X = F^{-1}(0.5)$
- Normal distribution is symmetric around $F^{-1}(0.5)$
A person is considered to have mental retardation when:

1. IQ is below 70
2. Significant limitations exist in two or more adaptive skill areas
3. Condition is present from childhood

What percentage of people have IQ that meet the first criterion of mental retardation?

Let’s assume IQ is normally distributed with mean $\mu = 100$ and standard deviation $\sigma = 16$. 
Example: $P(IQ < 70)$ Area under the density

IQ histogram: $P(IQ<70)$

![IQ histogram image]

Thais Paiva  STA 111 - Summer 2013 Term II  Lecture 6, 07/09/2013
Example: $P(IQ < 70)$ on CDF
Reggie Jackson, one of the greatest baseball players ever, has an IQ of 140. What percentage of people have higher IQ than Reggie?

Marilyn vos Savant, self-proclaimed smartest person in the world, has a reported IQ of 205. What percentage of people have IQ scores smaller than Marilyn’s score?

Mensa is a society for “intelligent people”. To qualify for Mensa, one needs to be in at least the upper 2% of the population in IQ score. What is the score needed to qualify for Mensa?
In order to find numerical values for the previous questions we need the cumulative distribution function. The easiest way to evaluate this is to utilize the standard Normal distribution.

A **standard normal distribution** is defined as follow:

\[ p(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} \]

The standard Normal distribution is often denoted as \( Z \). It has \( \mu = 0 \) and \( \sigma^2 = 1 \).

\[ Z \sim N(0, 1) \]

(This notation says \( Z \) follows a normal distribution with mean 0 and variance 1)
Remember that the CDF of the normal distribution depends on $\mu$ and $\sigma^2$.

However, we can transform any normal random variable to one with mean 0 and variance 1 by

1. Subtracting the mean $\mu \rightarrow$ center at zero
2. Dividing by the SD $\sigma \rightarrow$ reduce the spread (variance) to 1

Let $X \sim N (\mu, \sigma^2)$. Then

$$\frac{X - \mu}{\sigma} = Z \sim N(0, 1)$$
Let $X \sim N(\mu, \sigma^2)$. We are interested in finding the probability of $X$ between the interval $[a, b]$. Note the equality:

$$P(a \leq X \leq b) = P\left(\frac{a - \mu}{\sigma} \leq \frac{X - \mu}{\sigma} \leq \frac{b - \mu}{\sigma}\right)$$

$$= P\left(Z_1 \leq \frac{X - \mu}{\sigma} \leq Z_2\right)$$

where

$$\frac{X - \mu}{\sigma} = Z \sim N(0, 1)$$

and $Z_1$ and $Z_2$ are some transformed constants. These are known as $Z$-scores.
Reggie Jackson’s IQ is $x = 140$

By knowing that the IQ is normally distributed with mean $\mu = 100$ and standard deviation $\sigma = 16$, we have

$$P(X \geq 140) = P \left( \frac{X - \mu}{\sigma} \geq \frac{140 - \mu}{\sigma} \right)$$

$$= P \left( \frac{X - 100}{16} \geq \frac{140 - 100}{16} \right)$$

$$= P(Z \geq 2.5) = 0.006$$

0.006 comes from a table

The book’s table reports $P(y > 2.5)$
Marilyn vos Savant’s IQ is $x = 205$. We want $P(X < 205)$

We have

$$\frac{x - \mu}{\sigma} = \frac{205 - 100}{16} = 6.5$$

In the book’s table $P(y > 6.5)$ is not reported because it is very small. Therefore we can say $P(y > 6.5) \approx 0$

$P(X < 205) = P(Z < 6.5) = 1 - 0 \approx 100\%$

Remember symmetry.
To qualify for Mensa, one needs to be in at least the upper 2% of the population in IQ score.

- \( P(X > x_0) = 0.02. \) We need to find \( x_0 \)

- By looking at the table we see that the 2% upper tail probability is associated with \( Z_0 = 2 \)

\[
P(X \geq x_0) = P \left( \frac{X - \mu}{\sigma} \geq \frac{X_0 - \mu}{\sigma} \right) = P \left( Z \geq \frac{X_0 - 100}{16} \right)
\]

\[
P(Z \geq 2) = 0.02
\]

- Therefore

\[
\frac{X_0 - 100}{16} = 2 \rightarrow X_0 = 2 \times 16 + 100 = 132
\]
Let $X$ be distributed according to $p_X(x)$ with $\text{mean}(X) = \mu$, $\text{var}(X) = \sigma^2$, and $\text{sd}(X) = \sigma$. We apply the transformation $Y = aX + b$.

\[
\text{Mean}(Y) := E(Y) = \int_D y p_Y(y)dy = \int_D f(x)p_X(x)dx
\]

\[
= \int_D (ax + b) p_X(x)dx
\]

\[
= \int_D a x p_X(x)dx + \int_D b p_X(x)dx
\]

\[
= a \int_D x p_X(x)dx + b \int_D p_X(x)dx = a\mu + b
\]

Therefore $E(Y) = aE(X) + b$
What about the variance of $Y = aX + b$?

$$
Var(Y) = \int_D [y - E(Y)]^2 p_Y(y) dy = \int_D [ax + b - E(ax + b)]^2 p_X(x) dx
$$

$$
= \int_D [ax + b - (a\mu + b)]^2 p_X(x) dx = a^2 \int_D (x - \mu)^2 p_X(x) dx
$$

$$
= a^2 \sigma^2
$$

Therefore

- $Var(Y) = a^2 Var(X)$
What about the standard deviation of $Y = aX + b$?

- $SD(Y) = \sqrt{Var(Y)} = \sqrt{a^2\sigma^2} = a\sigma$

To summarize, if we have $Y = aX + b$ then

$$E(Y) = a \ E(X) + b$$

$$Var(Y) = a^2 \ Var(X)$$

$$SD(Y) = a \ SD(X)$$

- We have seen this before!
For the normal distribution, we said that if $X \sim N(\mu, \sigma^2)$ then

$$Y = \frac{X - \mu}{\sigma} \sim N(0, 1)$$

It is easy to see that in this case

$$Y = \left(\frac{1}{\sigma}\right) X + \left(-\frac{\mu}{\sigma}\right) = aX + b$$

Therefore

$$E(Y) = \left(\frac{1}{\sigma}\right) E(X) + \left(-\frac{\mu}{\sigma}\right) = \left(\frac{1}{\sigma}\right) \mu + \left(-\frac{\mu}{\sigma}\right) = 0$$

$$Var(Y) = \left(\frac{1}{\sigma}\right)^2 Var(X) = \frac{\sigma^2}{\sigma^2} = 1$$
Suppose that \( X \) is income in Dollars and we want to transform to Euros.

- \( E(\text{income}) = \$30,000 \)
- \( \text{Var}(\text{income}) = (\$15,000)^2 \)
- \( SD(\text{income}) = \$15,000 \)

Then given that $1.5 = 1 \€$

- \( E(\text{income}) = \$30,000 \times \frac{1\€}{\$1.5} = 20,000\€ \)
- \( \text{Var}(\text{income}) = (\$15,000)^2 \left( \frac{1\€}{\$1.5} \right)^2 = (10,000\€)^2 \)
- \( SD(\text{income}) = \$15,000 \cdot \frac{1\€}{\$1.5} = 10,000\€ \)
Phil and Kim do not know whether to buy a house now or wait a year, in which case a price increase may put a house beyond their reach. Their best guess is that, if they wait a year, the price increase ($X$) will be approximately normal, with mean of 8% and, reflecting the uncertainty of the market, a standard deviation of 10%.

- If the price increase exceeds 25% they feel they will be unable to afford a house. What is the chance of this?

$$P(X > 25) = P(Z > \frac{25-8}{10}) = P(Z > 1.7) = 0.045$$

- On the other hand, if the price drops, they will have won their gamble handsomely. What is the chance of this?

$$P(X < 0) = P(Z < \frac{0-8}{10}) = P(Z < -0.8) = P(Z > 0.8) = 0.212$$