UQ, STAT2201, 2017,
Lectures 3 and 4
Unit 3 – Probability Distributions.
Random Variables
A **random variable** $X$ is a numerical (integer, real, complex, vector etc.) summary of the outcome of the random experiment.

The **range** or **support** of the random variable is the set of possible values that it may take. Random variables are usually denoted by capital letters.
A **discrete random variable** is an integer/real-valued random variable with a finite (or countably infinite) range.

A **continuous random variable** is a real valued random variable with an interval (either finite or infinite) of real numbers for its range.
Experiment $\Rightarrow$ Outcome, $\omega$, from the sample space.

$X(\omega) \equiv \text{Random Variable (function of the outcome)}.$

$P(X \in \mathcal{U}) = P\left(\{\omega \mid X(\omega) \in \mathcal{U}\}\right).$
Example: Dig a hole searching for gold.

\[ \Omega \equiv \text{all possible outcomes (many ways to define this)} \]

\[ X \equiv \text{Weight of gold found in grams} \]

\[ P(X > 20) = P\left( \{ \omega \mid X(\omega) \in \mathcal{U} \} \right) \]

with \( \mathcal{U} = \{ x \mid x > 20 \} \).
Probability Distributions
The **probability distribution** of a random variable $X$ is a description of the probabilities associated with the possible values of $X$.

There are several common alternative ways to describe the probability distribution, with some differences between discrete and continuous random variables.
While not the most popular in practice, a unified way to describe the distribution of any scalar valued random variable $X$ (real or integer) is the **cumulative distribution function**, 

$$F(x) = P(X \leq x).$$

It holds that

1. $0 \leq F(x) \leq 1$.
2. $\lim_{x \to -\infty} F(x) = 0$.
3. $\lim_{x \to \infty} F(x) = 1$.
4. If $x \leq y$, then $F(x) \leq F(y)$. That is, $F(\cdot)$ is non-decreasing.
Examples to understand:

\[ F(x) = \begin{cases} 
0, & x < -1, \\
0.3, & -1 \leq x < 1, \\
1, & 1 \leq x. 
\end{cases} \]
Distributions are often summarised by numbers such as the **mean**, $\mu$, **variance**, $\sigma^2$, or **moments**. These numbers, in general do not identify the distribution, but hint at the general location, spread and shape.

The **standard deviation** of $X$ is $\sigma = \sqrt{\sigma^2}$ and is particularly useful when working with the Normal distribution.

More on these soon.
Discrete Random Variables
Given a discrete random variable $X$ with possible values $x_1, x_2, \ldots, x_n$, the **probability mass function** of $X$ is,

$$p(x) = P(X = x).$$

Note: In [MonRun2014] and many other sources, the notation used is $f(x)$ (as a pdf of a continuous random variable).
A probability mass function, $p(x)$ satisfies:

1. $p(x_i) \geq 0$.
2. $\sum_{i=1}^{n} p(x_i) = 1$.

The **cumulative distribution function** of a discrete random variable $X$, denoted as $F(x)$, is

$$F(x) = \sum_{x_i \leq x} p(x_i).$$
\( P(X = x_i) \) can be determined from the \textit{jump} at the value of \( x \). More specifically

\[
p(x_i) = P(X = x_i) = F(x_i) - \lim_{x \uparrow x_i} F(x_i).
\]
Back to the example:

\[
F(x) = \begin{cases} 
0, & x < -1, \\
0.3, & -1 \leq x < 1, \\
1, & 1 \leq x. 
\end{cases}
\]

What is the pmf?
The **mean** or **expected value** of a discrete random variable $X$, is

$$\mu = E(X) = \sum_{x} x \, p(x).$$
The expected value of $h(X)$ for some function $h(\cdot)$ is:

$$E \left[ h(X) \right] = \sum_x h(x) p(x).$$
The $k$’th **moment** of $X$ is,

$$E(X^k) = \sum_x x^k p(x).$$
The **variance** of $X$, is

$$\sigma^2 = V(X) = E(X - \mu)^2 = \sum_x (x - \mu)^2 p(x) = \sum_x x^2 p(x) - \mu^2.$$
The Discrete Uniform Distribution
A random variable $X$ has a **discrete uniform distribution** if each of the $n$ values in its range, $x_1, x_2, \ldots, x_n$, has equal probability. I.e.

$$p(x_i) = 1/n.$$
Suppose that $X$ is a discrete uniform random variable on the consecutive integers $a, a + 1, a + 2, \ldots, b$, for $a \leq b$. The mean and variance of $X$ are

$$E(X) = \frac{b + a}{2}$$
and
$$V(X) = \frac{(b - a + 1)^2 - 1}{12}.$$
To compute the mean and variance of the discrete uniform, use:

\[
\sum_{k=1}^{n} k = \frac{n(n + 1)}{2}, \quad \sum_{k=1}^{n} k^2 = \frac{n(n + 1)(2n + 1)}{6}
\]
\[ E(X) = \sum_{k=a}^{b} k \frac{1}{b-a+1} = \]
\[ E(X^2) = \sum_{k=a}^{b} k^2 \frac{1}{b-a+1} = \]
The Binomial Distribution
The setting of $n$ independent and identical Bernoulli trials is as follows:

(1) There are $n$ trials.
(1) The trials are independent.
(2) Each trial results in only two possible outcomes, labelled as “success” and “failure”.
(3) The probability of a success in each trial denoted as $p$ is the same for all trials.
Binomial Example: Number of digs finding gold.

\[ n = 5 \text{ digs in different spots.} \]

\[ p = 0.1 \text{ chance of finding gold in each spot.} \]
The random variable $X$ that equals the number of trials that result in a success is a **binomial random variable** with parameters $0 \leq p \leq 1$ and $n = 1, 2, \ldots$. The probability mass function of $X$ is

$$p(x) = \binom{n}{x} p^x (1 - p)^{n-x}, \quad x = 0, 1, \ldots, n.$$
Useful to remember from algebra: the binomial expansion for constants $a$ and $b$ is

$$(a + b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k}.$$
If $X$ is a binominal random variable with parameters $p$ and $n$, then,

\[ E(X) = np \quad \text{and} \quad V(X) = np(1 - p). \]
Example (cont.): Number of digs finding gold \((n = 5, p = 0.1)\):
Continuous Random Variables
Given a continuous random variable \( X \), the **probability density function** (pdf) is a function, \( f(x) \) such that,

1. \( f(x) \geq 0. \)
2. \( f(x) = 0 \) for \( x \) not in the range.
3. \( \int_{-\infty}^{\infty} f(x) \, dx = 1. \)
4. For small \( \Delta x \), \( f(x) \Delta x \approx P(X \in [x, x + \Delta x]) \).
5. \( P(a \leq X \leq b) = \int_{a}^{b} f(x) \, dx = \text{area under } f(x) \text{ from } a \text{ to } b. \)
Given the pdf, \( f(x) \) we can get the cdf as follows:

\[
F(x) = P(X \leq x) = \int_{-\infty}^{x} f(u)du \quad \text{for} \quad -\infty < x < \infty.
\]
Given the cdf, $F(x)$ we can get the pdf:

$$f(x) = \frac{d}{dx} F(x).$$
The **mean** or **expected value** of a continuous random variable \( X \), is

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

The **expected value** of \( h(X) \) for some function \( h(\cdot) \) is:

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

The \( k'\)th **moment** of \( X \) is,

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

The **variance** of \( X \), is

\[
\sigma^2 = V(X) = \int_{-\infty}^{\infty} (x - \mu)^2 \, f(x) \, dx = \int_{-\infty}^{\infty} x^2 \, f(x) \, dx - \mu^2.
\]
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** or “uniform random variable” for short.
If $X$ is a continuous uniform random variable over $a \leq x \leq b$, the **mean** and **variance** are:

$$
\mu = E(X) = \frac{a + b}{2} \quad \text{and} \quad \sigma^2 = V(X) = \frac{(b - a)^2}{12}.
$$
The Normal Distribution
A random variable $X$ with probability density function

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

is a **normal random variable** with parameters $\mu$ where $-\infty < \mu < \infty$, and $\sigma > 0$. For this distribution, the parameters map directly to the mean and variance,

$$E(X) = \mu \quad \text{and} \quad V(X) = \sigma^2.$$

The notation $N(\mu, \sigma^2)$ is used to denote the distribution. Note that some authors and software packages use $\sigma$ for the second parameter and not $\sigma^2$. 
A normal random variable with a mean and variance of:

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

is called a **standard normal random variable** and is denoted as \( Z \). The cumulative distribution function of a standard normal random variable is denoted as

\[ \Phi(z) = F_Z(z) = P(Z \leq z), \]

and is tabulated in a table.
It is very common to compute $P(a < X < b)$ for $X \sim N(\mu, \sigma^2)$. This is the typical way:

$$P(a < X < b) = P(a - \mu < X - \mu < b - \mu)$$

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

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

$$= \Phi\left(\frac{b - \mu}{\sigma}\right) - \Phi\left(\frac{a - \mu}{\sigma}\right).$$

We get:

$$F_X(b) - F_X(a) = F_Z\left(\frac{b - \mu}{\sigma}\right) - F_Z\left(\frac{a - \mu}{\sigma}\right).$$
The Exponential Distribution
The **exponential distribution** with parameter $\lambda > 0$ is given by the **survival function**,

$$
\bar{F}(x) = 1 - F(x) = P(X > x) = e^{-\lambda x}.
$$

The random variable $X$ represents the distance between successive events from a Poisson process with mean number of events per unit interval $\lambda > 0$. 
The probability density function of $X$ is

$$f(x) = \lambda e^{-\lambda x} \quad \text{for} \quad 0 \leq x < \infty.$$ 

Note that sometimes a different parameterisation, $\theta = 1/\lambda$ is used (e.g. in the Julia Distributions package).
The **mean** and **variance** are:

\[
\mu = E(X) = \frac{1}{\lambda} \quad \text{and} \quad \sigma^2 = V(X) = \frac{1}{\lambda^2}
\]
The exponential distribution is the only continuous distribution with range \([0, \infty)\) exhibiting the **lack of memory property**. For an exponential random variable \(X\),

\[
P(X > t + s \mid X > t) = P(X > s).
\]
Monte Carlo Random Variable Generation
Monte Carlo simulation makes use of methods to transform a uniform random variable in a manner where it follows an arbitrary given distribution. One example of this is if $U \sim \text{Uniform}(0, 1)$ then $X = -\frac{1}{\lambda} \log(U)$ is exponentially distributed with parameter $\lambda$. 