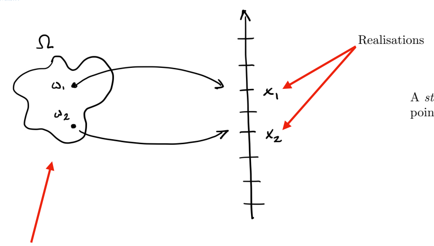
Stochastic variables

Andreas Jakobsson

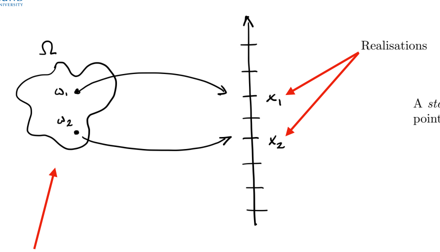
Stochastic variables



The sample space, Ω, contains all possible outcomes

A *stochastic variable* is a mapping from an event in the sample space, Ω, to a point on the real line. This outcome is the *realisation*.

Stochastic variables



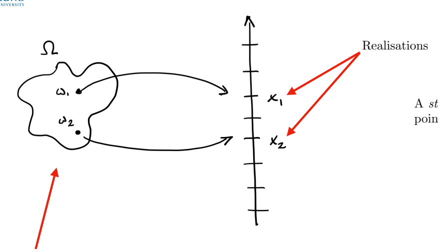
The sample space, Ω, contains all possible outcomes

A *stochastic variable* is a mapping from an event in the sample space, Ω, to a point on the real line. This outcome is the *realisation*.

Sample space	Realisations
 →	{ 1, 3, 5, 1, 6, 2 }

This would be an example of a one-dimensional *discrete* stochastic variable.

Stochastic variables



The sample space, Ω, contains all possible outcomes

A *stochastic variable* is a mapping from an event in the sample space, Ω, to a point on the real line. This outcome is the *realisation*.

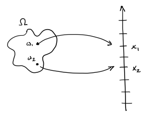
Sample space	Realisations
 →	{ 1, 3, 5, 1, 6, 2 }

This would be an example of a one-dimensional *discrete* stochastic variable.

Sample space	Realisations
 →	{ 31.451, 42.327, π }

This would be an example of a one-dimensional *continuous* stochastic variable.

Stochastic variables

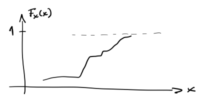


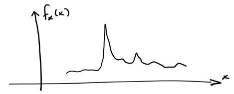
The mapping is characterised by the *probability distribution function*

$$F_X(x) = P(X \leq x) = \int_{-\infty}^{x} f_X(x)dx$$

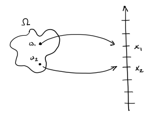
where $f_X(x)$ is the *probability density function* (pdf)

$$f_X(x) = \frac{d}{dx} F_X(x)$$





Stochastic variables

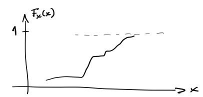


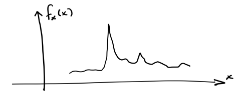
The mapping is characterised by the *probability distribution function*

$$F_X(x) = P(X \leq x) = \int_{-\infty}^{x} f_X(x)dx$$

where $f_X(x)$ is the *probability density function* (pdf)

$$f_X(x) = \frac{d}{dx} F_X(x)$$



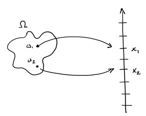


The *expected value* is defined as

$$E\{X\} = \int_{-\infty}^{\infty} x f_X(x)dx = m_X$$

This is the value one can expect the variable to take, on average.

Stochastic variables

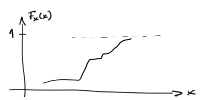


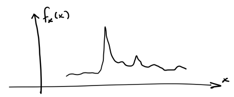
The mapping is characterised by the *probability distribution function*

$$F_X(x) = P(X \leq x) = \int_{-\infty}^{x} f_X(x)dx$$

where $f_X(x)$ is the *probability density function* (pdf)

$$f_X(x) = \frac{d}{dx} F_X(x)$$





The *expected value* is defined as

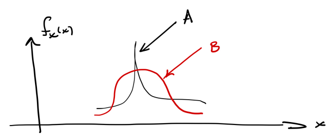
$$E\{X\} = \int_{-\infty}^{\infty} x f_X(x)dx = m_X$$

This is the value one can expect the variable to take, on average.

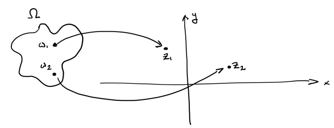
The *variance* is defined as

$$\begin{aligned} V\{X\} &= E\{(X - m_X)^2\} \\ &= E\{X^2 - 2m_X X + m_X^2\} = E\{X^2\} - 2m_X E\{X\} + m_X^2 \\ &= E\{X^2\} - m_X^2 \\ &= \int_{-\infty}^{\infty} (x - m_X)^2 f_X(x)dx \end{aligned}$$

This is a measure how much different realisations can be expected to differ from each other.



Stochastic variables



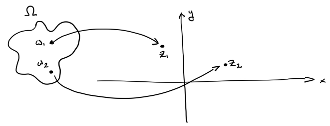
A *stochastic variable* may have more than one dimension. For example, the mapping from the sample space can be to a 2-D variable $z = [\ x \quad y\ ]$, making the realisation a point in a 2-D space.

Sample space

Realisations

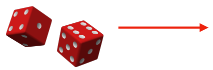
$z_1 = [\ 2 \quad 4\ ]$, $z_2 = [\ 5 \quad 1\ ]$

Stochastic variables



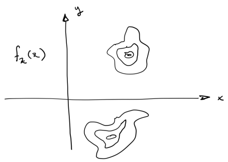
A *stochastic variable* may have more than one dimension. For example, the mapping from the sample space can be to a 2-D variable $z = \begin{bmatrix} x & y \end{bmatrix}$, making the realisation a point in a 2-D space.

Sample space



Realisations

$z_1 = \begin{bmatrix} 2 & 4 \end{bmatrix}$, $z_2 = \begin{bmatrix} 5 & 1 \end{bmatrix}$



The pdf $f_Z(z)$ is a 3-D function

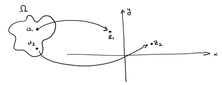
In this case, the mapping is characterised by the probability distribution function

$$F_Z(z) = F_{X,Y}(x,y) = P(X \leq x, Y \leq y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f_{X,Y}(x,y)dxdy$$

where the pdf $f_{X,Y}(x,y)$ is

$$f_{X,Y}(x,y) = \frac{\partial}{\partial x \partial y} F_{X,Y}(x,y) = f_Z(z)$$

Stochastic variables



If X and Y are *statistically independent*, the pdf is separable, i.e.,

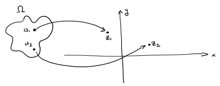
$$f_{X,Y}(x,y) = f_X(x)f_Y(y)$$

This is a very strong assumption. A weaker assumption is that of the variables being *uncorrelated*, implying that

$$E\{XY\} = E\{X\}E\{Y\}$$

It the variables are independent, they are also uncorrelated (show this!), but not necessarily the other way around.

Stochastic variables



If X and Y are *statistically independent*, the pdf is separable, i.e.,

$$f_{X,Y}(x,y) = f_X(x)f_Y(y)$$

This is a very strong assumption. A weaker assumption is that of the variables being *uncorrelated*, implying that

$$E\{XY\} = E\{X\}E\{Y\}$$

It the variables are independent, they are also uncorrelated (show this!), but not necessarily the other way around.

If the variables depend on each other, this dependence can be measured via the cross-covariance function

$$r_{XY} = C\{X,Y\} = E\{(X - m_x)(Y - m_Y)^*\} = E\{XY^*\} - m_X m_Y^*$$

where $(\cdot)^*$ denotes the conjugate. Clearly, $r_{XY} = 0$ if X and Y are uncorrelated.

As the variance and cross-covariance scale depend on the scale of the stochastic variables, one instead often use the correlation coefficient

$$\rho_{XY} = \frac{C\{X,Y\}}{\sqrt{V\{X\}}\sqrt{V\{Y\}}}$$

which is bounded $0 \leq \rho_{XY} \leq 1$.