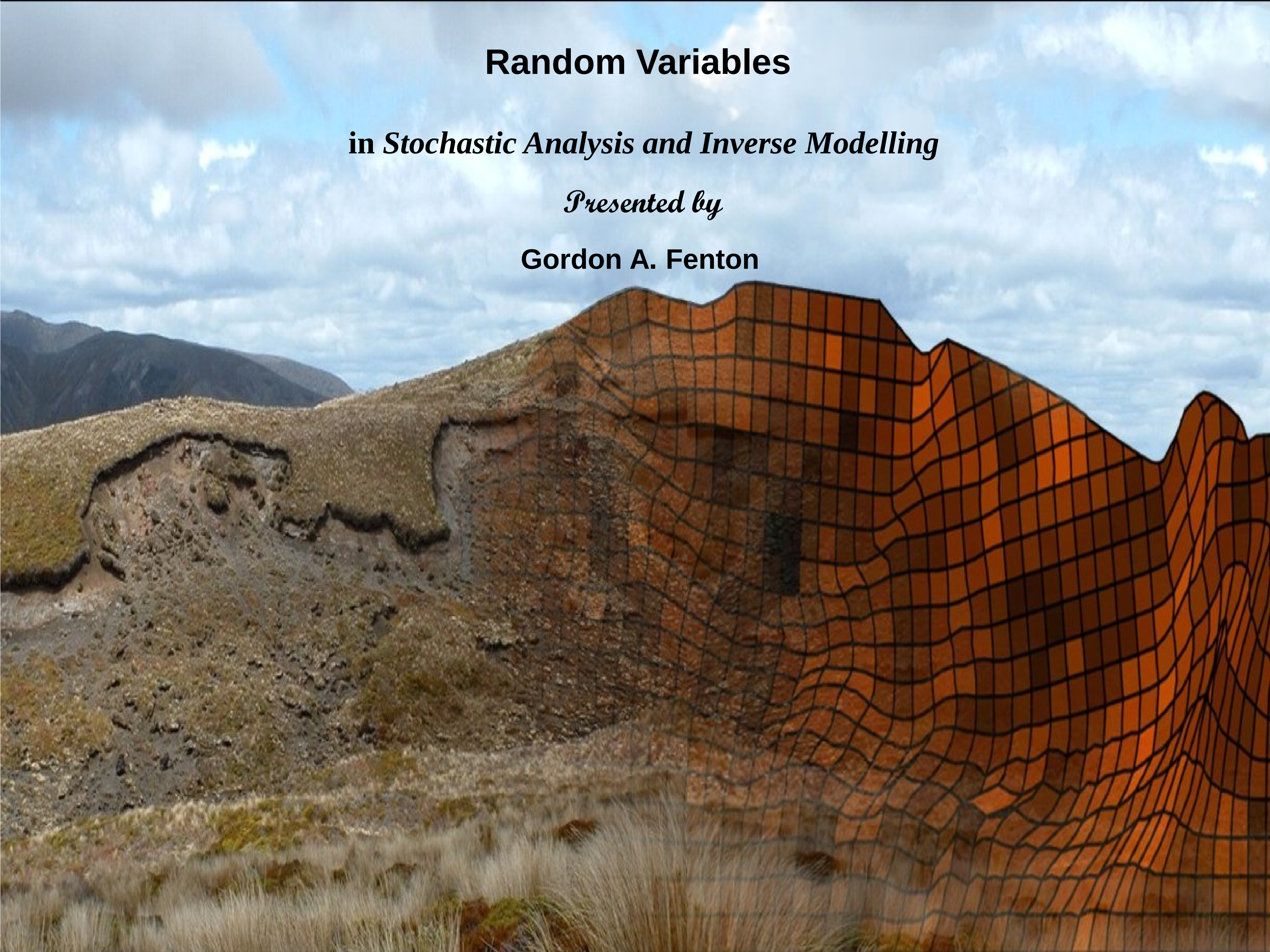
Random Variables

in *Stochastic Analysis and Inverse Modelling*

Presented by

Gordon A. Fenton



Random Variables

A random variable is a mapping from a real event to a number, i.e.,

Let $X_i = 1$ if the i'th pile adequately supports its load

$= 0$ if not

Now we can use math. For example

$$\sum_{i=1}^{n} X_i$$

gives the number of piles which adequately support their loads, out of n piles.

Discrete and Continuous Random Variables

Discrete : A random variable is called a discrete random variable if its set of possible outcomes is countable. This usually occurs for any random variable which is a count of occurrences or of items.

Continuous : A random variable is called a continuous random variable if it can take on values on a continuous scale. This is usually the case with measured data, such as cohesion.

Examples:

1) Let X be the number of blows in an Standard Penetration Test

-- X is *discrete.*

2) Let Y be the number of piles driven in one day

-- Y is discrete.

3) Let Z be the time till consolidation settlement exceeds some threshold

-- Z is continuous.

4) Let W be the undrained shear strength of a soil deposit.

--W is continuous.

Random variables are *unknown*.

The *most* we can say about a random variable is what its probability is of taking on each of its possible values.

We call this a *probability distribution*. For example if the random variable X can take on possible values 0, 1, 2, and 3, then the complete description of X would be the four numbers;

$$\mathrm{P}[X = 0] = 0.22$$

$$\mathrm{P}[X = 1] = 0.47$$

$$\mathrm{P}[X = 2] = 0.18$$

$$\mathrm{P}[X = 3] = 0.13$$

(for example) which must add up to 1.0.

Probability Distributions

Discrete Random Variables: The set of probabilities

$$P[X = x_1] = p_1$$

$$P[X = x_2] = p_2$$

$$\vdots$$

is called a Probability Mass Function (PMF)

Continuous Random Variables:

$$P[X = 12.000000000.....000] = 0$$

so we must define probabilities over a range

$$P[x < X \le x + dx] = f_X(x)dx$$

where $f_x(x)$ is called a Probability Density Function (PDF)

Probability Density Function

Definition :

The function $f(x)$ is a probability density function (PDF) for the continuous random variable X, defined over the set of real numbers, if

1) $0 \leq f(x) < \infty, \quad \text{for all} \quad -\infty < x < \infty,$

2) $\int_{-\infty}^{\infty} f(x)\,dx = 1$ (the area under the PDF is unity)

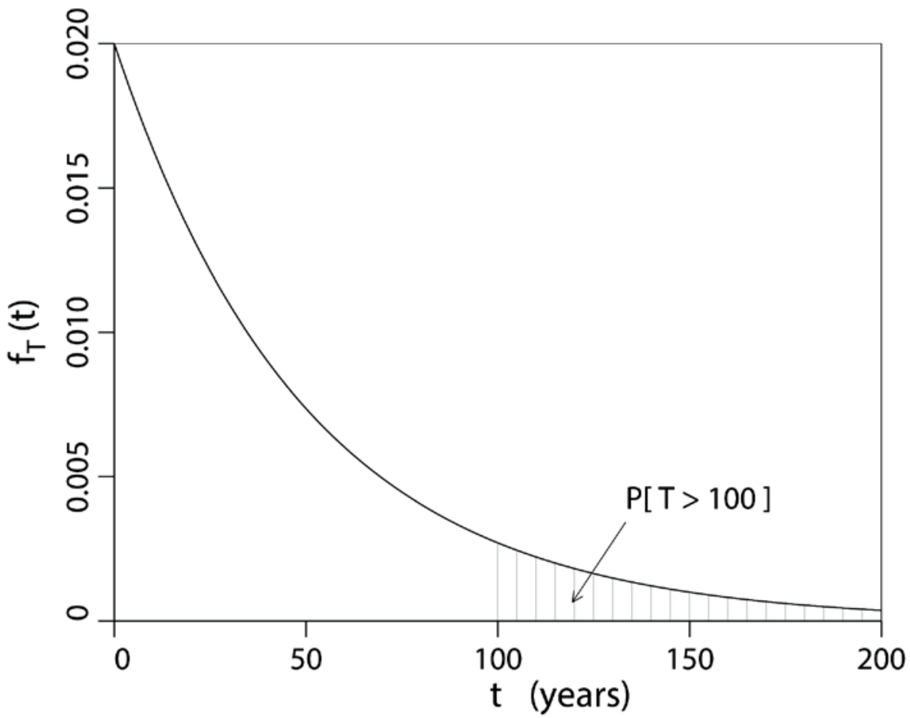
3) $P[a < X < b] = \int_{a}^{b} f(x)\,dx$

NOTE : it is important to recognize that, in the continuous case, $f(x)$ is not a probability. It has units of probability per unit length. In order to get probabilities, we have to find *areas* under the pdf, ie. sum up values of $f(x)\,dx$.

Example: Suppose that the time to failure, T in years, of a clay barrier has the following probability density function:

$$f_T(t) = \lambda e^{-\lambda t} \quad \text{for } t \geq 0 \quad \text{where } \lambda = 0.02$$

What is the probability that T will exceed 100 years?



$$\text{P}\left[T > 100\right] = \int_{100}^{\infty} \lambda e^{-\lambda t}\, dt$$

$$= e^{-100\lambda} = e^{-100(0.02)}$$

$$= e^{-2} = 0.1353$$

Cumulative Distribution Function

An "equivalent" description of a random variable is the cumulative distribution function, $F_X(x)$, defined as

$$F_X(x) = \mathrm{P}[X \le x] = \sum_{x_i \le x} \mathrm{P}[X = x_i] \quad \text{(discrete case)}$$

$$= \int_{-\infty}^{x} f_X(s)\,ds \qquad \text{(continuous case)}$$

In the continuous case, knowledge of $F_X(x)$ allows us to avoid integration. For example, if $f_T(t) = \lambda e^{-\lambda t}$, for $t \ge 0$ then

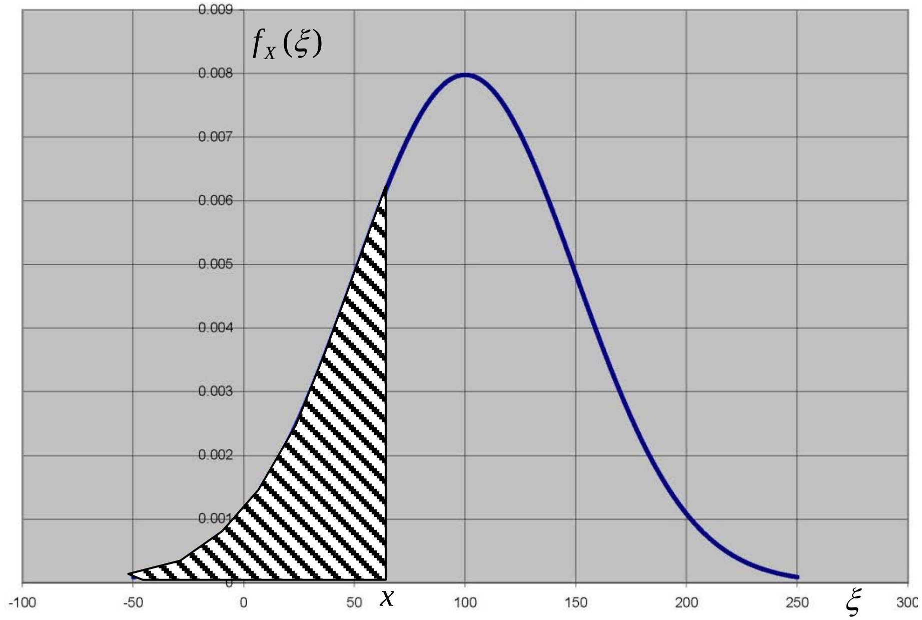
$$F_T(t) = \mathrm{P}[T \le t] = \int_0^t f_T(s)\,ds = 1 - e^{-\lambda t}$$

and $\mathrm{P}[T > 100] = 1 - F_T(100) = 1 - \left(1 - e^{-100\lambda}\right) = e^{-100\lambda}$

Cumulative Distribution Function (CDF)

The Cumulative Distribution Function for a random variable gives the probability that the random variable will take a value less than or equal to a given value *x*. Hence,

$$F_X(x) = P[X \le x] = \int_{-\infty}^{x} f_X(\xi)\,d\xi$$



Expectation Operator

Let a random variable X be described by a probability density function (PDF), $f_X(x)$

If $g(X)$ is a function of the random variable X, then the expected value of $g(X)$ is its average value after it has been weighted by the PDF of X, thus

$$\mathrm{E}\left[g(X)\right] = \int_{-\infty}^{\infty} g(x)\, f_X(x)\, dx$$

Examples: $$\mathrm{E}\left[X^2\right] = \int_{-\infty}^{\infty} x^2\, f_X(x)\, dx$$

$$\mathrm{E}\left[\cos(X - \mu_X)\right] = \int_{-\infty}^{\infty} \cos(x - \mu_X)\, f_X(x)\, dx$$

More About Expectations

Expectation of a sum = sum of expectations

$$\mathrm{E}\left[a_1 X_1 + a_2 X_2\right] = a_1 \mathrm{E}\left[X_1\right] + a_2 \mathrm{E}\left[X_2\right]$$

Linear function of a single random variable (SRV)

$$\mathrm{E}\left[a + bX\right] = a + b\mathrm{E}\left[X\right]$$

Linear function of a two random variables

$$\mathrm{E}\left[a + bX + cY\right] = a + b\mathrm{E}\left[X\right] + c\mathrm{E}\left[Y\right]$$

Sum of multiple random variables

$$\mathrm{E}\left[X_1 + X_2 + \ldots + X_n\right] = \mathrm{E}\left[X_1\right] + \mathrm{E}\left[X_2\right] + \ldots + \mathrm{E}\left[X_n\right]$$

Sum of functions of random variables

$$\mathrm{E}\left[f\left(X\right) + g\left(Y\right)\right] = \mathrm{E}\left[f\left(X\right)\right] + \mathrm{E}\left[g\left(Y\right)\right]$$

Main Descriptors of a Random Variable

Mean: $\mu_X = \mathrm{E}[X] = \sum_i x_i \mathrm{P}[X = x_i]$ (discrete case)

$= \int_{-\infty}^{\infty} x\, f_X(x)\,dx$ (continuous case)

Variance: $\sigma_X^2 = \mathrm{Var}[X] = \mathrm{E}\left[(X - \mu_X)^2\right] = \mathrm{E}\left[X^2\right] - \mathrm{E}^2[X]$

$= \sum_i (x_i - \mu_X)^2 \mathrm{P}[X = x_i]$ (discrete case)

$= \int_{-\infty}^{\infty} (x - \mu_X)^2 f_X(x)\,dx$ (continuous case)

Skewness: $\psi_X = \dfrac{\mathrm{E}\left[(X - \mu_X)^3\right]}{\sigma_X^3} = \dfrac{1}{\sigma_X^3} \int_{-\infty}^{\infty} (x - \mu_X)^3 f_X(x)\,dx$

IDENTITIES RELATING TO VARIANCE

Single random variable (SRV)

$$\begin{aligned}\mathrm{Var}[X] &= \mathrm{E}\left[(X-\mu_X)^2\right] \\ &= \mathrm{E}\left[X^2 - 2X\mu_X + \mu_X^2\right] \\ &= \mathrm{E}\left[X^2\right] - 2\mathrm{E}^2[X] + \mathrm{E}^2[X] \\ &= \mathrm{E}\left[X^2\right] - \mathrm{E}^2[X]\end{aligned}$$

Linear function of a single random variable (SRV)

$$\begin{aligned}\mathrm{Var}[a+bX] &= \mathrm{E}\left[\left(\left(a+bX\right)-\left(a+b\mu_X\right)\right)^2\right] \\ &= b^2\mathrm{E}\left[(X-\mu_X)^2\right] \\ &= b^2\mathrm{Var}[X]\end{aligned}$$

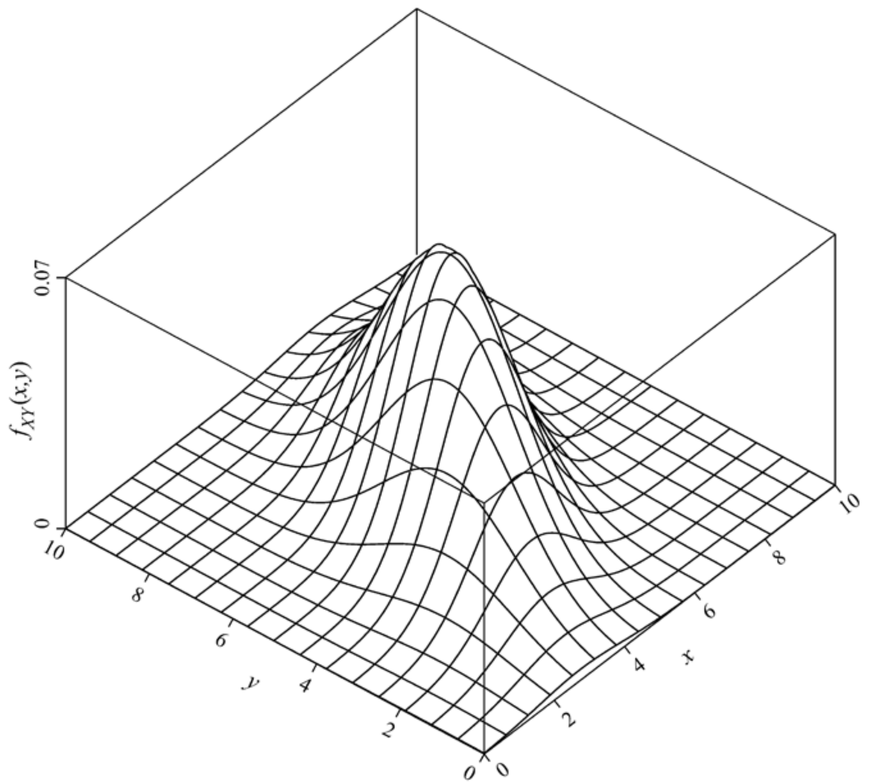
hence

$$\sigma_{a+bX} = b\sigma_X$$

Pairs of Random Variables

If we consider two random variables, X and Y, their relationship is fully prescribed by a bivariate probability distribution;

$$\mathrm{P}\left[X = x \cap Y = y\right] = f_{XY}(x, y)\,dx\,dy$$



Covariance

The most important piece of information about the relationship between two random variables is their *degree of linear dependence*. This is determined by their *covariance*;

$$\mathrm{Cov}[X,Y] = \mathrm{E}[(X-\mu_X)(Y-\mu_Y)] = \mathrm{E}[XY] - \mathrm{E}[X]\mathrm{E}[Y]$$

where $\mathrm{E}[XY] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xy\, f_{XY}(x,y)\,dx\,dy$

If X and Y are *independent*, then

$$\mathrm{P}[X = x \cap Y = y] = \mathrm{P}[X = x]\mathrm{P}[Y = y]$$

so that $f_{XY}(x,y) = f_X(x)\, f_Y(y)$ and $\mathrm{E}[XY] = \mathrm{E}[X]\mathrm{E}[Y]$
which means that $\mathrm{Cov}[X,Y] = 0$
(NOTE: the converse is not generally true)

Linear function of two random variables

$$
\begin{aligned}
&\mathrm{Var}\left[a+bX+cY\right] \\
&=\mathrm{E}\left[\left(\left(a+bX+cY\right)-\left(a+b\mu_X+c\mu_Y\right)\right)^2\right] \\
&=\mathrm{E}\left[\left(b\left(X-\mu_X\right)+c\left(Y-\mu_Y\right)\right)^2\right] \\
&=\mathrm{E}\left[\left(b^2\left(X-\mu_X\right)^2+2bc\left(X-\mu_X\right)\left(Y-\mu_Y\right)+c^2\left(Y-\mu_Y\right)^2\right)\right] \\
&=b^2\mathrm{Var}\left[X\right]+2bc\mathrm{Cov}\left[X,Y\right]+c^2\mathrm{Var}\left[Y\right]
\end{aligned}
$$

where $\mathrm{Cov}\left[X,Y\right]$ is the "Covariance" of X and Y, defined by
$\mathrm{Cov}\left[X,Y\right]=\mathrm{E}\left[\left(X-\mu_X\right)\left(Y-\mu_Y\right)\right]$

Covariance is a measure of the linear dependence between random variables, e.g. when X is greater than its mean, Y is also greater than its mean, and so on.

Correlation Coefficient

Since the covariance has dimensions of *XY*, its magnitude is not intuitively meaningful.
The correlation coefficient directly reflects the degree of linear dependence between *X* and *Y*

$$\rho_{XY} = \frac{\mathrm{Cov}[X,Y]}{\sigma_X \sigma_Y}$$

It can be shown that $-1 \le \rho_{XY} \le 1$ since $\mathrm{Var}\left[\frac{X}{\sigma_X} \pm \frac{Y}{\sigma_Y}\right] \ge 0$

If $\rho_{XY} = \pm 1$, then *X* and *Y* are completely linearly dependent, i.e.

$$Y = a + bX$$

with b negative if $\rho_{XY} = -1$ and positive if $\rho_{XY} = +1$

Correlation Coefficient

If the random variables are perfectly correlated then $Y = a + bX$, so that $\sigma_Y = \sqrt{b^2\sigma_X^2} = |b|\sigma_X$, and we can write

$$\begin{aligned}
\mathrm{Cov}[X,Y] &= \mathrm{E}\left[X(a+bX)\right] - \mathrm{E}[X]\mathrm{E}[a+bX] \\
&= a\mathrm{E}[X] + b\mathrm{E}\left[X^2\right] - a\mathrm{E}[X] - b\mathrm{E}^2[X] \\
&= b\left(\mathrm{E}\left[X^2\right] - \mathrm{E}^2[X]\right) \\
&= b\mathrm{Var}[X] \\
&= b\sigma_X^2 \\
&= \mathrm{sign}(b)\sigma_X\sigma_Y \qquad \text{(i.e. having the same sign as } b\text{)}
\end{aligned}$$

The correlation coefficient between X and Y is thus

$$\rho_{XY} = \frac{\mathrm{Cov}[X,Y]}{\sigma_X\sigma_Y} = \frac{\mathrm{sign}(b)\sigma_X\sigma_Y}{\sigma_X\sigma_Y} = \mathrm{sign}(b) = \pm 1$$

Variance of a Sum

Variance of a sum = sum of all possible covariances

$$\text{Var}\left[a_1X_1 + a_2X_2\right] = \text{Cov}\left[a_1X_1, a_1X_1\right] + \text{Cov}\left[a_1X_1, a_2X_2\right]$$
$$+ \text{Cov}\left[a_2X_2, a_1X_1\right] + \text{Cov}\left[a_2X_2, a_2X_2\right]$$
$$= a_1^2\text{Cov}\left[X_1, X_1\right] + 2a_1a_2\text{Cov}\left[X_1, X_2\right] + a_2^2\text{Cov}\left[X_2, X_2\right]$$

where $\text{Cov}\left[X_i, X_i\right] = \text{E}\left[\left(X_i - \mu_{X_i}\right)\left(X_i - \mu_{X_i}\right)\right]$

$$= \text{E}\left[\left(X_i - \mu_{X_i}\right)^2\right] = \text{Var}\left[X_i\right]$$

so that $\text{Var}\left[a_1X_1 + a_2X_2\right] = a_1^2\sigma_{X_1}^2 + a_2^2\sigma_{X_2}^2 + 2a_1a_2\text{Cov}\left[X_1, X_2\right]$

If X_1 and X_2 are uncorrelated, then $\text{Cov}\left[X_1, X_2\right] = 0$ and

Variance of a sum = sum of the variances

Variance of a Sum

In general if $Y = \sum_{i=1}^{n} a_i X_i$ then

$$\mu_Y = \mathrm{E}[Y] = \mathrm{E}\left[\sum_{i=1}^{n} a_i X_i\right] = \sum_{i=1}^{n} \mathrm{a}_i \mathrm{E}[X_i]$$

$$\sigma_Y^2 = \mathrm{Var}[Y] = \mathrm{Var}\left[\sum_{i=1}^{n} a_i X_i\right] = \sum_{i=1}^{n}\sum_{j=1}^{n} a_i a_j \mathrm{Cov}\left[X_i, X_j\right]$$

Note: when $j = i$ the 'diagonal' terms are

$$a_i^2 \mathrm{Cov}[X_i, X_i] = a_i^2 \mathrm{Var}[X_i]$$

When the X's are uncorrelated, then $\mathrm{Cov}\left[X_i, X_j\right] = 0$ for $i \neq j$ and we are left with

$$\sigma_Y^2 = \mathrm{Var}\left[\sum_{i=1}^{n} a_i X_i\right] = \sum_{i=1}^{n} a_i^2 \,\mathrm{Var}[X_i]$$

Variance of an Average

Note that if $Y = \frac{1}{n}\sum_{i=1}^{n} X_i$, i.e. $a_i = 1/n$,

and $\mathrm{Cov}\left[X_i, X_j\right] = \sigma_X^2 \rho_{ij}$ then $\sigma_Y^2 = \sigma_X^2 \left(\frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n} \rho_{ij} \right)$

The bracketed quantity is the average correlation coefficient between all possible pairs. We can write this as

$$\sigma_Y^2 = \sigma_X^2 \gamma(n)$$

where $\gamma(n)$ is the variance reduction function

$$\gamma(n) = \left(\frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n} \rho_{ij} \right)$$

If the X's are uncorrelated, then $\gamma(n) = \frac{1}{n}$ as expected