Stochastic Processes

Summary

Table 1 Definitions

Outcome	Each thing that can occur in an expiriment is called an outcome. In the example of tossing a coin we have two outcomes 'heads' or 'tails' which we can denote by the letters H and T respectively.
Sample Space	The set of all possible outcomes of an experiment is known as the sample space. By convention we label it Ω.. In our simple coin tossing scenario we would have $\Omega = \{H, T\}$ If we toss two coins our sample space would become $\Omega = \{HH, HT, TH, TT\}$
Event	A subset of the probability space is ` an event. We define an event using the following notation. $A = \{\varpi \in \Omega; \varpi = H\}$ "This means the set of all outcomes ϖsuch that $\varpi$is a head".
Probability Measure	A probability measure P is a function that assigns to each element ϖ in Ω a probability such that $$\sum_{\omega \in \Omega} P(\varpi) = 1$$ Since an event A is a subset of Ω then the probability of an event is given by $$P(A) = \sum_{\omega \in A} P(\varpi)$$
Probability Space	A probability space (Ω, P) consists of a sample space and a probability measure. The sample space is the set of outcomes and the probability measure is a function that assigns to each element ϖin Ω a value in $[0,1]$ such that
Random Variable	A random variable x is a real valued function defined on Ω . Put another way a random variable maps each outcome from the sample space Ω to a real number.
Proabability Distribution	We now introduce one more important concept, that of a probability distribution. A random variable is a function defined on Ω whereas its distribution is a tabulation of the probabilities that the random variable takes its various values. A random variable is **not** a distribution.

TABLE 2 PROPERTIES OF EXPECTATION

1. Definition of Expectation	$E[X] = \sum_{\omega \epsilon \Omega} X(\omega)P(\omega)$
2. Expectation of constant multiple	$E[cX] = cE[X]$ for any constant c
3. Exectation of constant addition	$\mathrm{E[X + b] = E[X] + b}$ for any constant b
4. Linearity of expectation	$\mathrm{E[aX + b] = aE[X] + b}$
5. Expecation of a sum of random variables	$E[X_1 + \cdots X_n] = E[X_1] + \cdots + E[X_n]$
6. Expectation of a function of a random variable	$E[g(X)] = \sum_{\omega \epsilon \Omega} g(X(\omega))P(\omega)$

TABLE 3 PROPERTIES OF VARIANCE

1. Definition of Variance	$Var[X] = \sum_{\omega \in \Omega} [X(\omega) - E(X)]^2 P(\omega)$ $Var[X] = E[X^2] - (E[X])^2$
2. Variance of constant	$Var[a] = 0$
3. Variance of addition	$Var[\mathrm{X} + a] = Var[\mathrm{X}]$
4. Variance of a constant multiple	$Var[aX] = a^2 Var[X]$
5. Sum of two independent random variables	$Var[X + Y] = Var[X] + Var[Y]$
6. Sum of n IIR variable	$n.Var[X]$

Worked example

A Single experiment

Imagine a random event that involves the tossing of a single coin. We have two *outcomes*, heads or tails giving us *a sample space* of

$$\Omega = \{H, T\}$$

Furthermore, let us define two *random variables*. The first x_1 takes the value of plus one if we obtain a head and minus one if we obtain a tail.

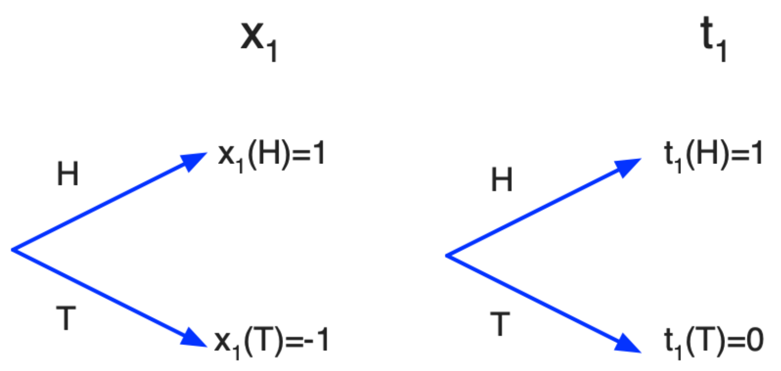
$$x_1(H) = 1, x_1(T) = -1$$

The second takes a value of plus one if we obtain a head and zero if we get a tail

$$t_1(H) = 1, t_1(T) = 0$$

Notice that our random variables do not say anything about the probability of a head or tail. They just tell us what value we assign to the outcomes of the sample space.

Figure 1 Random variables are real valued functions on the sample space

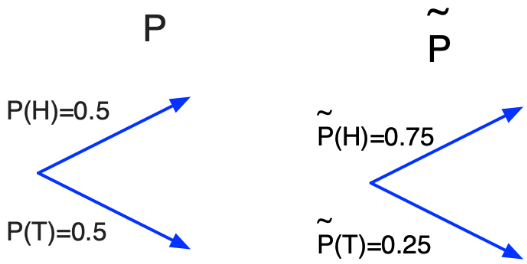


A probability measure is a real valued function that maps each outcome of the sample space to a probability. If our coin is fair we could have a measure P such that

$$P(H) = 0.5, P(T) = 0.5$$

We might however have a different measure for a loaded coin

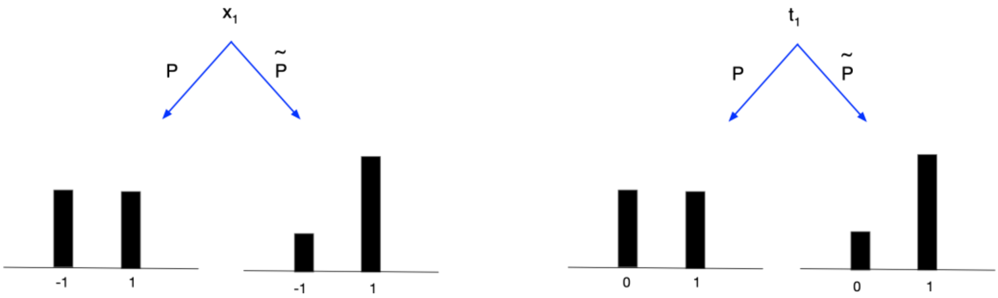
$$\tilde{P}(H) = 0.75, \tilde{P}(T) = 0.25$$



Distribution

Applying a probability measure to a random variable gives us a distribution. The distribution shows the probability of each value of the random variable. Since we have two random variable and two measures we have four probability distributions

Figure 2Applying measures to random variables give us distributions



Mean and variance

We can define the expectation or expected value of any random variable X under a probability measure P as

$$E(X) = \sum_{\omega \in \Omega} X(\omega) P(\omega)$$

For any actual value of a random variable X we can calculate the difference between that value and the expectation $X(\omega) - E(X)$ .We might ask the question "on average how much does a given value differ from the expected value?" We could calculate the average difference as $\sum_{\omega \in \Omega} [X(\omega) - E(X)] P(\omega)$ however where the distribution is symmetric around the mean this value will be zero. A more instructive measure is given by calculating the average of the difference squared

$$\sum_{\omega \in \Omega} [X(\omega) - E(X)]^2 P(\omega)$$

We now add the mean and variance values to the four distributions ontained by applying the two measures to our two random variables

Figure 3 Mean and variance

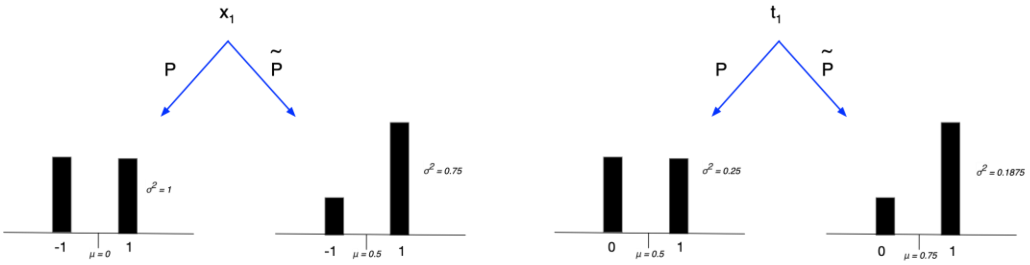


TABLE 4 SUMMARY

Outcome	H	Each outcome is a thing that can occur in an experiment
Sample Space	$\Omega = \{H, T\}$	The set of all possible outcomes that can occur in an experiment is called the sample space
Event	$A = \{\varpi \in \Omega; \varpi = H\}$	A subset of the sample space is called an event
Probability Measure	$P(H) = P(T) = 0.5$	A probability measure P is a function that assigns to each element ϖ in Ω a probability such that $\sum_{\omega \in \Box} P(\varpi) = 1$
Probability Space	$(\{H, T\}, P(H) = P(T)) = 0.5)$	A probability space consists of a sample space and a probability measure
Random Variable	$x_1(H) = 1, x_1(T) = -1$	A random variable is a real valued function defined on the sample space.
Probability Distribution	$P(x_1 = 1) = 0.5, P(x_1 = -1) = 0.5$	Tabulation of the probabilities that the random variable takes its various values.
Expectation	$(X) = x_1(H)P(H) + x_1(T)P(T) = 0$	We define the expected value of our random variable under the probability measure P
Variation	$[x_1(H) - E(x_1)]^2\tilde{P}(H) + [x_1(T) - E(x_1)]^2\tilde{P}(H) = 1.5$	

Performing twice

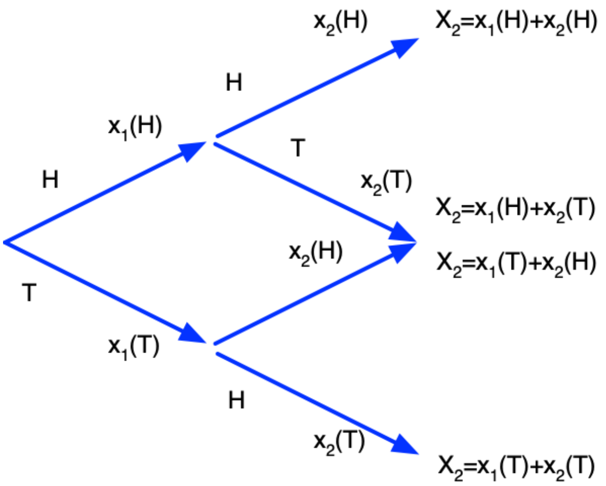
We can create a new game by playing the original games multiple times. If we play the original game twice then our new game effectively involves tossing the coin two times and our sample space becomes $\Omega = \{HH, HT, TH, TT\}$. We can define two new random variables by summing the original variables

$X_2 = x_1 + x_2$.

$T_2 = t_1 + t$.

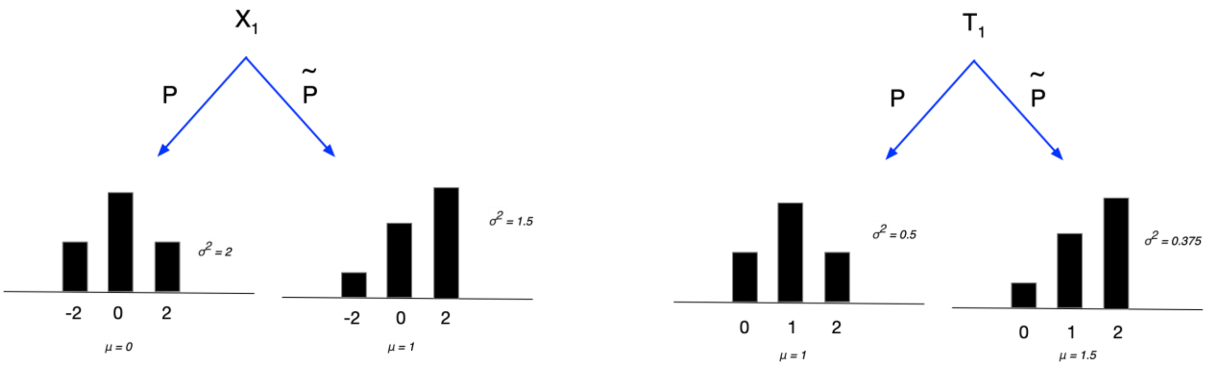
The following diagram shows how summing two independent random variables works for $X_2 = x_1 + x_2$. Notice how there are two ways of achieving the outcome that a head and a tail occur $\{HT, TH\}$

Figure 4 Summing two instances of the same random variable



Under our two measures our distributions of the two variables are as follows

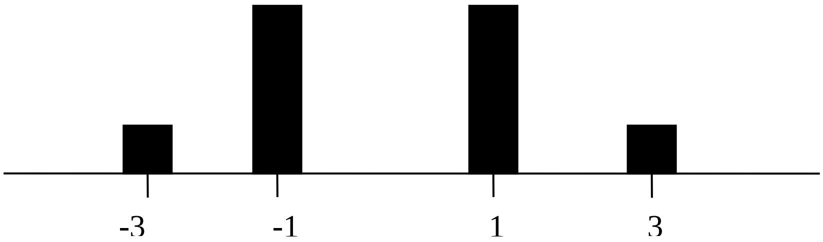
Figure 5 Mean and variance



Performing three times

Let us go one-step further and look at the event obtained by summing three of the original events. $X_3 = x_1 + x_2 + x_3$ Under the probability measure P we get the following distribution, whose mean is zero and whose variance is three.

Figure 6 Distribution of X_3 under P



- $\mu = 0$
- $\sigma^2 = 0.125(3-0)^2 + 0.125(-3-0)^2 + 0.375(1-0)^2 + 0.375(-1-0)^2 = 3.0$

Figure 7Distribution of T2 under P

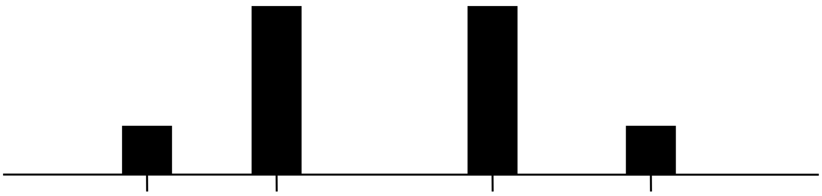
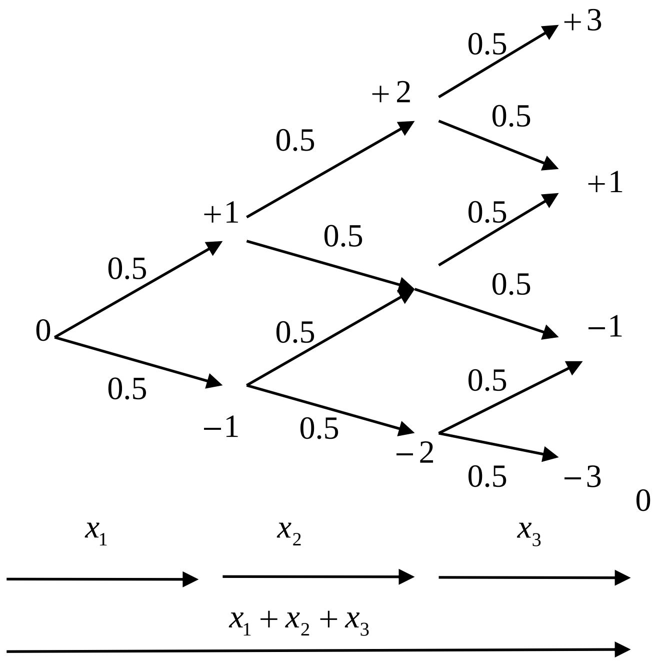


Figure 8 Tree for X_3



Generalizing

If we then perform n identical tosses of the coin and define n identical random variables $x_1, x_2, x_3, \ldots, x_n$ we can define a new random variable as the sum of the individual random variables

$$X_n = x_1 + \cdots + x_n.$$

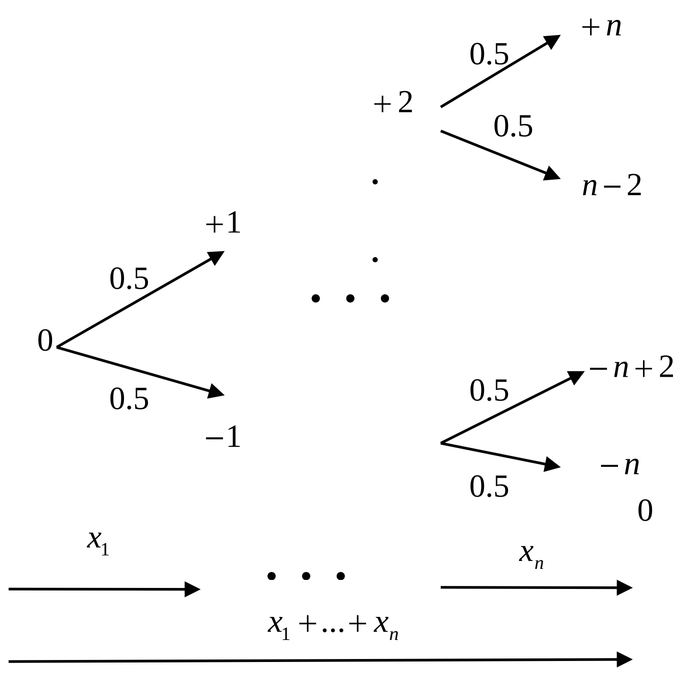
And similarly for $t_1, t_2, t_3, \ldots, t_n$ we can define a new random variable

$$T_n = t_1 + \cdots + t_n$$

The sample space of T_n and X_n is then $\Omega = \{\omega_1\omega_2 \ldots \omega_1\}, \omega_i \in \{H, T\}$

In the general case to calculate the probability of obtaining k heads in n tosses we need to take into account the probability of a head on a single toss which we call p and the number of paths through the decision tree that come to that number of heads. The paths are given by the binomial co-efficients $\binom{n}{k} = \frac{n!}{(n-k)!k!}$ and the probability becomes $\binom{n}{k} p^k (1-p)^{n-k}$ The

Figure 9 Paths through the tree for X_n under measure P



Our distribution depends on both the random variable and the probability measure. Under our measure P representing a fairly weighted coin the expectations of our two random variables are given by $E(X_n) = 0$ and $E(T_n) = \frac{n}{2}$

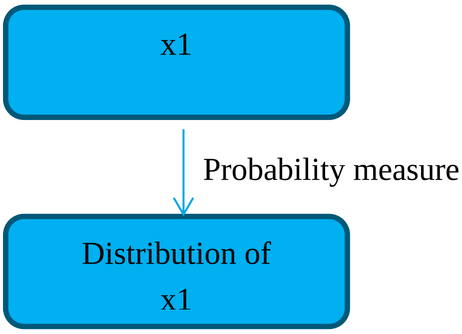
We can intepret X as the distance from the origin if we move one unit in a positive direction whenever we obtain a head and one unit in a negative direction whenever we obtain a tail. This is the ‘random walk’ interpretation.

In general the sum of n independent, identically distributed random variables with mean μ and variance σ^2 is a random variable with mean $n\mu$ and variance $n\sigma^2$ For a proof of why this is the case see the proofs section below.

Details

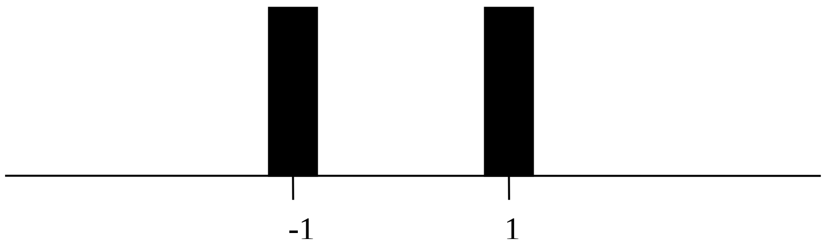
Probability Distribution

We now introduce one more important concept, that of a probability distribution. A random variable is a function defined on Ω whereas its distribution is a tabulation of the probabilities that the random variable takes its various values. A random variable is **not** a distribution.



Under the probability measure P defined on Ω either a head or tail are equally likely so our distribution becomes

$$P(x_1 = 1) = 0.5, P(x_1 = -1) = 0.5$$



Risk and Pricing Solutions

Expectation

We can define the expectation or expected value of any random variable X under a probability measure P as.

$$E(X) = \sum_{\omega \epsilon} X(\omega)P(\omega)$$

- Weighted average of the values the random variable X can take
- Weighting by the probability of each value
- Measure of centrality

Expectation of Variable Squared

We are often interested in expectation of the square of the variable which we call the mean squared.

$$E({x_1}^2) = [x_1(H)]^2P(H) + [x_1(T)]^2P(T) = 0.5 + 0.5 = 1.0$$

$$\tilde{E}({x_1}^2) = [x_1(H)]^2\tilde{P}(H) + [x_1(T)]^2\tilde{P}(T) = 0.75 + 0.25 = 1.0$$

Variation from expected value

For any actual value of a random variable X we can calculate the difference between that value and the expectation $X(\omega) - E(X)$ .We might ask the question “on average how much does a given value differ from the expected value?” We could calculate the average difference as $\sum_{\omega \in \square}[X(\omega) - E(X)]P(\omega)$ however where the distribution is symmetric around the mean this value will be zero. A more instructive measure is given by calculating the average of the difference squared

$$\sum_{\omega \in \Omega} [X(\omega) - E(X)]^2P(\omega)$$

Under our two probability measures we get

$$[x_1(H) - E(x_1)]^2P(H) + [x_1(T) - E(x_1)]^2P(H) = 0.5 + 0.5 = 1.0$$

$$[x_1(H) - E(x_1)]^2\tilde{P}(H) + [x_1(T) - E(x_1)]^2\tilde{P}(H) = 0.5 + 0.5 = 1.5$$

Expectation of a function of random variable

$$E[g(X)] = \sum_{\omega \epsilon \Omega} g(X(\omega))P(\omega)$$

- The expectation of a function of a random variable is **not equal** to the function of the expectation $E[g(X)] \neq g[E(X)]$

Manipulations

Scaling and Shifting

We now know that if we have a process x with mean zero and variance one we can scale and shift it to a process with mean μ and variance σ^2 by adding μ and multiplying by $\sqrt{\sigma^2} = \sigma$. Our new random variable $(x + \mu)\sigma$is now distributed with mean μ and variance σ^2

Even more useful is the fact that if we know that a random process x is distributed with mean μ and variance σ^2 then we also know that the random variable $\frac{(x-\mu)}{\sigma}$ is distributed with mean zero and variance one

Increasing the number of steps

From our previous sections we can see that if we sum n identical independent random variables with mean zero and unit variance we obtain a new random variable

$$X_n = \sum_{i=1}^{n} x_n$$

with mean zero and variance n. But what happens as we increase the number of steps? In the limit as$n \to \infty$ $X_n$is normally distributed with mean $n\mu$ and variance $n\sigma^2$.

Using the standard normal distribution

The standard normal distribution with mean zero and standard deviation one has been studied extensively and its properties are well known. So if we have a random variable x which we know is normally distributed with mean μ and variance σ^2 then we can represent it via a scale and shift of the standard normal distribution $\varphi(0,1)$as follows

$$x = \mu + \sigma\varphi(0,1)$$

Stochastic Time Series

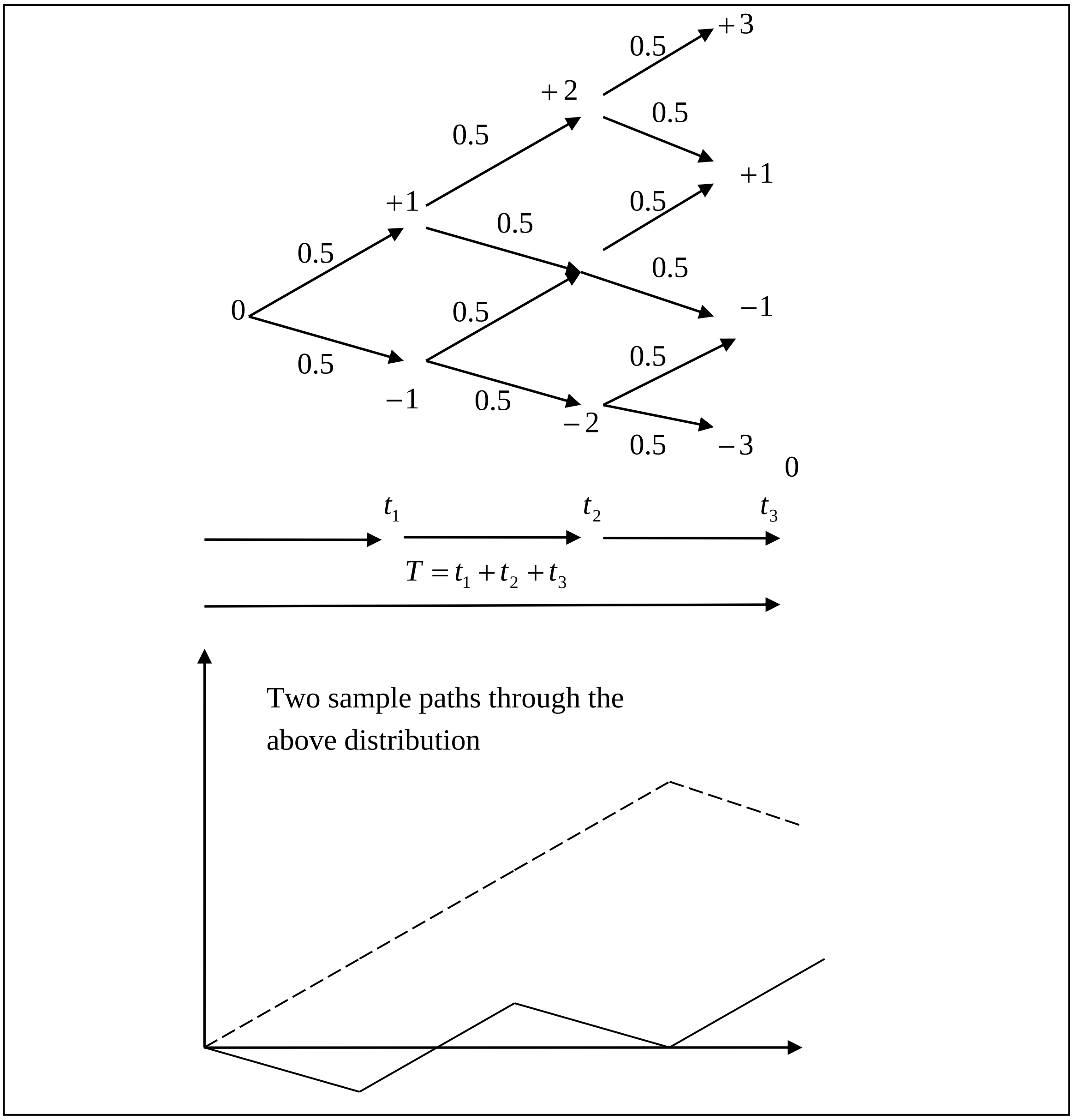
We know the sum of n identical, independently distributed random variables with mean μ and finite non-zero variances σ^2 is a random variable with mean $n\mu$and variance $n\sigma^2$.

$X_1 = x_1$	Mean μ and standard deviation σ^2
$X_2 = x_1 + x_2$	Mean 2μ and standard deviation $2\sigma^2$
$X_3 = x_1 + x_2 + x_3$	Mean 3μ and standard deviation $3\sigma^2$
$X_4 = x_1 + x_2 + x_3 + x_4$	Mean 4μ and standard deviation$4\sigma^2$
$X_n = x_1 + x_2 + x_3 + \ldots \rightleftarrows + x_n$	Mean $n\mu$ and standard deviation$n\sigma^2$

What if we were to replace the subscript n with the subscript t as follows?

$X_1 = x_1$	Mean μ and standard deviation σ^2
$X_2 = x_1 + x_2$	Mean 2μ and standard deviation $2\sigma^2$
$X_3 = x_1 + x_2 + x_3$	Mean 3μ and standard deviation $3\sigma^2$
$X_4 = x_1 + x_2 + x_3 + x_4$	Mean 4μ and standard deviation $4\sigma^2$
$X_t = x_1 + x_2 + x_3 + \ldots + x_t$	Mean $t\mu$ and standard deviation $t\sigma^2$

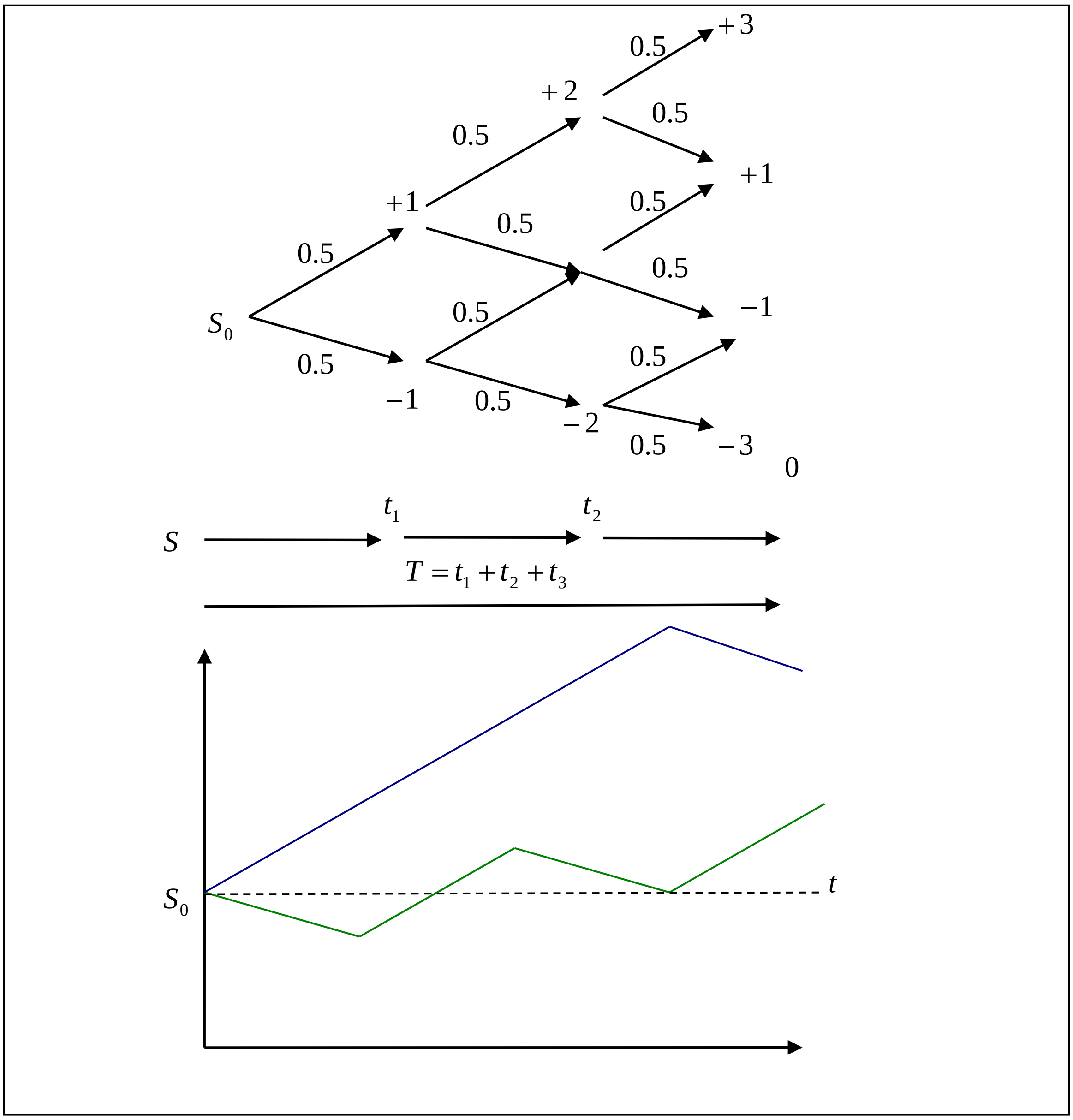
We can view X_t as a random variable which describes the displacement of a particle that starts at the origin and is displaced the distance X_t at time t.



Once again we note that it is the variance that scales linearly. The total variance in this example is three so each individual part has variance one.

Stock Price Process

Consider a stock price that starts at some level S_0. On each tick one of two equally probable outcomes can occur; either the stock moves up by one dollar or it moves down by one dollar. Over three ticks we could model our stock using the process in the previous section with one small modification; we adjust the mean to be the current stock level.



We have created a simple stochastic process to model the change in value of our stock over time. We know from the introductory sections that if over one unit of time our stochastic process has variance $\sigma^2$and mean zero then over a period of time T it will have variance $\sigma^2 T$

Risk and Pricing Solutions

Normal distribution notation

$\phi(0,1)$ Normal distribution with mean zero and variance of one

$\phi(0,2)$ Normal distribution with mean zero and variance of two

$\phi(0,2) = \sqrt{2}\phi(0,1)$

$\phi(0,t) = \sqrt{t}\phi(0,1)$

If a process has a variance of σ^2per unit of time then we get

$\phi(0,\sigma^2 t) = \sigma^2\sqrt{t}\phi(0,1)$

We also know that if we break down a stochastic process with variance $\sigma^2 T$ into n equal steps represented by the random variables x_i then at each step we will need the variance of x_i to equal $\frac{\sigma^2 T}{n}$. The change in our asset price can then be modelled as $X = \sum_{i=1}^{n} x_i$

In the limit as n tends to infinity the steps in our stochastic process become finer and finer but the total variance remains at $\sigma^2 T$ and our process converges to a Gaussian variable with mean zero and variance $\sigma^2 T$ and hence $\frac{X}{\sigma\sqrt{T}}$ is normally distributed with mean zero and variance one. We can hence represent the change in our stock price over the time T as the stochastic process

$$\sigma\sqrt{T}N(0,1)$$

We can then represent the terminal value of our stock price over the time T as the stochastic process

$$S_0 + \sigma\sqrt{T}N(0,1)$$

A more realiztic stock price process

The Markov property

The value of X_t given X_s is that determined by $X_t - X_s$ and so the behaviour of X_t is totally unaffected by the values of X_r for r less than s. This is known as the Markov property.

Normal distribution notation

$\phi(0,1)$ Normal distribution with mean zero and variance of one

$\phi(0,2)$ Normal distribution with mean zero and variance of two

$\phi(0,2) = \sqrt{2}\phi(0,1)$

$\phi(0,t) = \sqrt{t}\phi(0,1)$

If a process has a variance of σ^2 per unit of time then we get

$\phi(0,\sigma^2 t) = \sigma^2 \sqrt{t}\phi(0,1)$

Risk and Pricing Solutions

Brownian Motion

We say that a stochastic process W_t is a Brownian motion if it has the following properties.

- $W_0 = 0$
- For every s less than t, $W_t - W_s$ is normally distributed with variance $(t-s)$.

From these properties it follows that

$$W_t - W_s = \phi(0, t-s).$$

In the limit as $(t-s)$ tends to zero we get

$$dW_t = (W_t - W_s) = \phi(0, t-s) = \sqrt{(t-s)}\phi(0,1).$$

We can then represent a process with non unit variance as

$$\sigma dW_t = \sigma(W_t - W_s) = \sigma\phi(0, t-s) = \sigma\sqrt{(t-s)}\phi(0,1)$$

Stochastic differential equation

We extend this notion to define a family X of random variables X_t that satisfy the stochastic differential equation.

$$dX_t = \mu(t, X_t)dt + \sigma(t, X_t)dW_t$$

Dropping the parameters of the drift and volatility we get a general stochastic differential equation for a variable with a constant growth or drift rate plus some noise.

$dX_t = \mu dt + \sigma dW_t$

Proofs

Show that $E[X+Y] = E[X] + E[Y]$

If X is a random variable with sample space $\{x_1, x_2, \ldots\ldots, x_m\}$and Y is an independent random variable with sample space $\{y_1, y_2, \ldots\ldots, y_n\}$ then the sample space of X+Y is

$\{x_1, y_1\}, \{x_1, y_2\}, \ldots\ldots, \{x_1, y_n\}$

$\{x_2, y_1\}, \{x_2, y_2\}, \ldots\ldots, \{x_2, y_n\}$

.

.

.

$\{x_m, y_1\}, \{x_m, y_2\}, \ldots\ldots, \{x_m, y_n\}$

The expectation of the sum of the two variables in then given by

$\sum_{i=1}^{m}\sum_{j=1}^{n}(X(x_i) + Y(y_j)).p(x_i, y_j)$

Multiplying out we get

$$\sum_{i=1}^{m}\sum_{j=1}^{n} X(x_i).p(x_i, y_j) + \sum_{i=1}^{m}\sum_{j=1}^{n} Y(y_j)\,p(x_i, y_j)$$

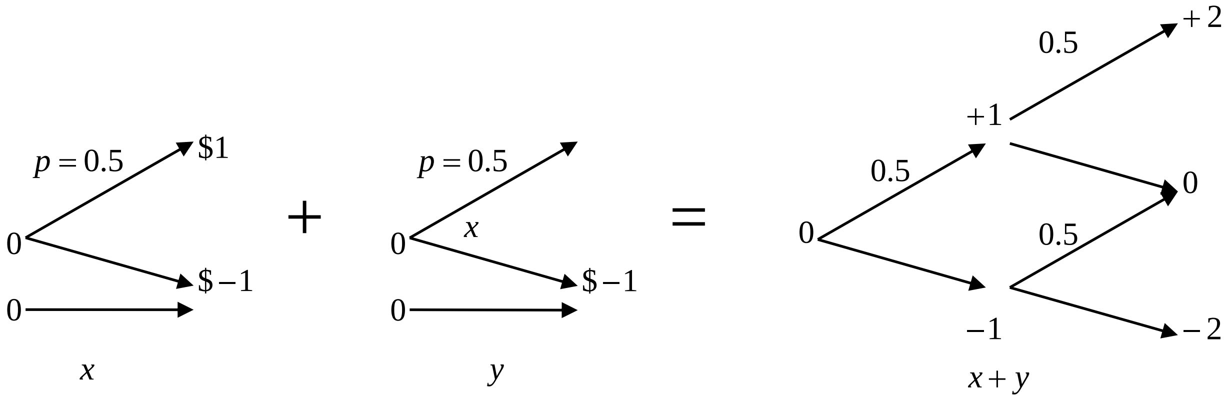
Noting that $\sum_{j=1}^{n} p(x_i, y_j) = p(x_i)$ and $\sum_{i=1}^{m} p(x_i, y_j) = p(y_j)$

$$\sum_{i=1}^{m} x_i p(x_i) + \sum_{j=1}^{n} x_j p(y_j)$$

Therefore we can note that

$E[X+Y] = E[X] + E[Y]$

The following figure shows the a specific example approach



We have random variable x with sample space $\{x_1, x_2\} = \{1,0\}$ and a second identically distributed random variable $y$with sample space$\{y_1, y_2\} = \{1,0\}$. The sample space of the joint distribution $x + y$ is given by the set of pairs

$$\{x_1, y_1\}, \{x_1, y_2\}$$

$$\{x_2, y_1\}, \{x_2, y_2\}$$

The expectation of the sum of the variables is then given by

$$\sum_{i=1}^{2}\sum_{j=1}^{2}(x_i + y_j)p(x_i, y_j)$$

Multiplying out

$$\sum_{i=1}^{2}\sum_{j=1}^{2} x_i\, p(x_i, y_j) + \sum_{i=1}^{2}\sum_{j=1}^{2} y_j\, p(x_i, y_j)$$

Noting that $\sum_{j=1}^{n} p(x_i\,, y_j) = p(x_i)$ and $\sum_{i=1}^{m} p(x_i\,, y_j) = p(y_j)$

$$\sum_{i=1}^{m} x_i p(x_i) + \sum_{j=1}^{n} x_j p(y_j)$$

Therefore we can note that

$$\boldsymbol{E}[\boldsymbol{X} + \boldsymbol{Y}] = \boldsymbol{E}[\boldsymbol{X}] + \boldsymbol{E}[\boldsymbol{Y}]$$

Risk and Pricing Solutions

Show that the expectation of the sum of n iid random variables is $n.E[X_n]$

We can calculate the expectation of the sum of n identically distributed random variables denoted by $X_1, X_2, \ldots, X_n$ as $E[X_1] + E[X_2] + \ldots + E[X_n]$ which is equal to

$$n.E[X_n]$$

Show that the $E[aX + b] = aE[X] + b$

$E[aX + b] = \sum_{\omega\in\Omega}(aX(\omega) + b)P(\omega)$ From definition 1

$= \sum_{\omega\in\Omega}(aX(\omega))P(\omega) + \sum_{\omega\in\Omega} bP(\omega)$ By multiplying out the brackets

$= a\sum_{\omega\in\Omega}(X(\omega))P(\omega) + \sum_{\omega\in\Omega} bP(\omega)$ From the properties of summation

$= aE[X] + b\sum_{\omega\in\Omega} P(\omega)$ From definition 1

$= aE[X] + b.1$ From axioms of probability

$$= aE[X] + b$$

Show that $Var[X] = E[(X-\mu)^2] = E[X^2] - (E[X])^2$

Let $\mu = E[X]$

$E[(X-\mu)^2] = \sum_{\omega\epsilon\Omega}(X(\omega) - \mu)^2 \; P(\omega)$ From definition

$=\sum_{\omega\epsilon\Omega}\left(\left(X(\omega)\right)^2 - 2\mu X(\omega) + \mu^2\right)P(\omega)$ Multiplying out

$=\sum_{\omega\epsilon\Omega}\left(X(\omega)\right)^2 P(\omega) + \sum_{\omega\epsilon\Omega} -2\mu X(\omega)P(\omega) + \sum_{\omega\epsilon\Omega}\mu^2 P(\omega)$

$= E[X^2] + \sum_{\omega\epsilon\Omega} -2\mu X(\omega)P(\omega) + \sum_{\omega\epsilon\Omega}\mu^2 P(\omega)$ From definition 3

$= E[X^2] + -2\mu\sum_{\omega\epsilon\Omega} X(\omega)P(\omega) + \mu^2\sum_{\omega\epsilon\Omega} P(\omega)$ Properties of summations

$$= E[X^2] - 2\mu\mu + \mu^2 \sum_{\omega\epsilon\Omega} P(\omega)$$

$= E[X^2] - 2\mu\mu + \mu^2$ Axioms of probability

$= E[X^2] - \mu^2$

$= E[X^2] - (E[X])^2$

Risk and Pricing Solutions

Show that $Var[aX] = a^2Var[X]$

$$Var[aX] = E[(aX - E[aX])^2]$$

$= E[(aX - aE[X])^2]$ From definition 2

$= E[(aX - a\mu)^2]$ Letting $\mu = E[X]$

$= \sum_{\omega\in\Omega}(aX(\omega) - a\mu)^2 \; P(\omega)$ From definition

$= \sum_{\omega\in\Omega} a^2(X(\omega) - \mu)^2 \; P(\omega)$

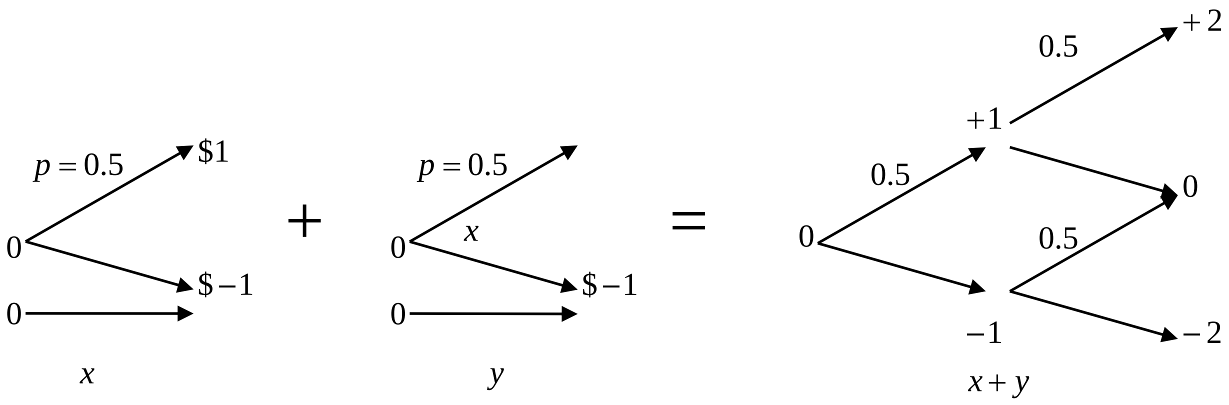
$$= a^2 \sum_{\omega\in\Omega}(X(\omega) - \mu)^2 \; P(\omega)$$

$= a^2Var[X]$ From definition 4

Risk and Pricing Solutions

Show that $\boldsymbol{Var[x+y] = Var[x] + Var[y]}$

The following diagram shows the general approach.



We have random variable x with sample space $\{x_1, x_2\} = \{1,0\}$and another identically distributed random variable y with sample space$\{y_1, y_2\} = \{1,0\}$. The sample space of the joint distribution $x + y$is given by the set of pairs

$$\{x_1, y_1\}, \{x_1, y_2\}$$

$$\{x_2, y_1\}, \{x_2, y_2\}$$

Therefore

$$Var[x+y] = E[(x+y)^2] - \{E[x+y]\}^2$$

$$Var[x+y] = E[(x^2 + 2xy + y^2)] - \{E[x] + E[Y]\}^2$$

$$Var[x+y] = E[x^2] + E[y^2] + E[2xy] - \{E[x] + E[Y]\}^2$$

$$Var[x+y] = E[x^2] + E[y^2] + E[2xy] - E[x]^2 - E[y]^2 - 2E[x][y]$$

$$Var[x+y] = E[x^2] + E[y^2] + 2E[x][y] - E[x]^2 - E[y]^2 - 2E[x][y]$$

$$Var[x+y] = E[x^2] - E[x]^2 + E[y^2] - E[y]^2$$

$$Var[x+y] = Var[x] + E[y]$$