# Brownian motion

Math 478

May 8, 2019

## 1 General characterization

A stochastic process $W_t, t \geq 0$ is a Brownian motion if

- $W_0 = 0$.
- $W_t$ has independent stationary increments.
- $W_t - W_s \sim N(0, t-s)$ for $t \geq s$.

We also say a stochastic process $X_t, t \geq 0$ is a Brownian motion with variance parameter $\sigma^2$ (this definition seems to be found only in Ross and not elsewhere) if

- $X_0 = 0$.
- $X_t$ has independent stationary increments.
- $X_t - X_s \sim N(0, \sigma^2(t-s))$ for $t \geq s$.

Observe how this definition is very similar to the one of Poisson process. The difference is in the distribution of the increments : it is Normal with mean 0 and variance equaling the time duration $t-s$ here versus being Poisson with mean $\lambda(t-s)$ for the Poisson process.

An important consequence of the definition is that the state space of Brownian motion is $\mathbb{R}$ : it is a continuous state space process, the first we've ssen so far in this course. Some other facts connected to this include Brownian motion having continuous but nowhere differentiable path. That is, with probability 1, the path $W_s, 0 \leq s \leq t$ is continuous but nowhere differentiable ! The proper context to give a proof for this fact is in a stochastic analysis course.

Also observe that the dynamical description is not given for Brownian motion, unlike the Poisson process case. Recall that the dynamical description for the Poisson process is closely connected to the Poisson approximation for the Binomial. Intuitively for the Brownian motion case if such a description is given it should be related to the Normal approximation for the Binomial. On the other hand, this approximation does not lend it self well to the the dynamical approach presentation. We'll see this more clearly in the section below.

## 1.1 Brownian motion as the limit of symmetric random walk

Recall that the symmetric random walk $S_k$ is given as

- $S_0 = 0$
- $S_k$ has iid increments
- $P(S_{k+1} - S_k = 1) = P(S_{k+1} - S_k = -1) = \frac{1}{2}$.

We present $S_k$ this way to draw the obvious connection to Brownian motion. Now $S_k$ is only defined for integral time points $k$. We can use $S_k$ to define the value of a process $X_t$ at a general time point $t$ as followed : for a given $dt$

$$X_t = \Delta x \sum_{k=1}^{\lfloor \frac{t}{dt} \rfloor} \Delta S_k,$$

where $\Delta S_k := S_{k+1} - S_k$ and $\Delta x$ is some scaling constant depending on $dt$.

The idea is to divide $[0, t]$ into subintervals with length $dt$ and run the usual symmetric random walk along the grid points until $t$ with step size $\Delta x$. Note that without the scaling constant $\Delta x$ the value of $X_t$ can get very large if $dt$ is small. In fact, we want to choose $\Delta x$ so that $E(X_t)$ and $Var(X_t) = t$ is independent of $dt$. By construction $E(X_t) = 0$ and $Var(X_t) = (\Delta x)^2 \frac{t}{4dt}$ so the choice of $\Delta X = 2\sqrt{dt}$ accomplishes this goal.

We can obviously repeat this idea for the increments $X_t - X_s$ instead of just $X_t$:

$$X_t - X_s = \Delta x \sum_{k=1}^{\lfloor \frac{t-s}{dt} \rfloor} \Delta S_k.$$

Again with the choice of $\Delta X = 2\sqrt{dt}$ we have $E(X_t - X_s) = 0$ and $Var(X_t - X_s) = t - s$. The increments can be viewed as a way of defining $X_t$ once $X_s$ is given.

Now for any fixed $t$, by the Central Limit Theorem the distribution of $X_t$ converges to $N(0,t)$ as $dt \to 0$. Similarly, the distribution of $X_t - X_s$ converges to $N(0, t-s)$ and by the independence of random walk increments we also get the independence of the increments of $X_t$. But notice how this is a finite dimensional description. That is for any finite collection of $t_1, t_2, \cdots, t_n$ we know the joint distribution of $X_{t_1}, X_{t_2}, \cdots, X_{t_n}$ as $dt \to 0$. On the other hand, it is not clear what we can immediately say about the distribution of the path $X_u, s \leq u \leq t$ as $dt \to 0$ (e.g. what is the probability that a path continuous ? ) This remains true no matter how close we push $t$ to $s$, because we still have uncountably many time points between $s$ and $t$. That is we do not directly have a dynamical description of $X_t - X_s$ from the limit of a symmetric random walk.

## 1.2 Finite dimensional distribution of Brownian motion

The independent increment property of Brownian motion leads to the fact that $[W(t_1), W(t_2), \cdots, W(t_n)]$ has a multivariate Normal distribution (recall Ross problem 3.16 for the case when $n = 2$). The multivariate Normal distribution is completely determined by its mean and covariance matrix. Indeed $X_1, X_2, \cdots, X_n$ has multivariate normal distribution with mean $\boldsymbol{\mu}$ and (invertible) covariance matrix $\Sigma$ if their joint density has the form :

$$f_{\mathbf{X}}(\mathbf{x}) = \frac{1}{\sqrt{(2\pi)^n det(\Sigma)}} e^{-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^T \Sigma^{-1} (\mathbf{x}-\boldsymbol{\mu})}.$$

Here $\mathbf{x} := [x_1, x_2, \cdots, x_n]^T, \boldsymbol{\mu} = [\mu_1, \mu_2, \cdots, \mu_n]^T$.

It is clear that for $[W(t_1), W(t_2), \cdots, W(t_n)], \boldsymbol{\mu} = \mathbf{0}$. How about $\Sigma$? Here we just need to decide the covariance between $W(t_i), W(t_j)$ for any $t_i, t_j$. We have

$$\begin{aligned} \Sigma_{ii} &= Var(W(t_i)) = t_i \\ \Sigma_{ij} &= Cov(W(t_i), W(t_j)) = \min(t_i, t_j). \end{aligned}$$

The last equality comes from the fact that if $s < t$

$$Cov(W_s, W_t) = Cov(W_s, W_s + (W_t - W_s)) = Cov(W_s, W_s) = s,$$

by independence of $W_s$ and $W_t - W_s$.

**Example 1.1.** In a bicycle race between two competitors, let $Y(t)$ denote the amount of time (in seconds) by which the racer that started in the inside position is ahead when $100t$ percent of the race has been completed, $0 \leq t \leq 1$. (So $t$ here does not quite have the meaning of time as we usually think it. Also if $Y(t)$ is negative this means the inside racer is behind). $Y(t, 0 \leq t \leq 1)$ can be effectively modeled as a Brownian motion with variance parameter $\sigma^2$.

a) If the inside racer is leading by $\sigma$ seconds at the midpoint of the race, what is the probability that she is the winner?

b) If the inside racer wins the race by a margin of $\sigma$ seconds, what is the probability that she was ahead at the midpoint?

Ans:

a) Mid point of the race means $t = 1/2$. The probability is

$$\begin{aligned} P(Y(1) > 0 \Big| Y(1/2) = \sigma) &= P(Y(1) - Y(1/2) > -\sigma \Big| Y(1/2) = \sigma) \\ &= P(Y(1) - Y(1/2) > -\sigma) = P(N(0, \frac{\sigma^2}{2}) > -\sigma) \\ &= P(Z > -\frac{\sqrt{2}}{\sigma}), \end{aligned}$$

where $Z$ denotes the standard Normal. Observe how the independent increment is utilized to solve this question.

b) Notice in this question, because we're conditioning on $Y(1)$, independent increment cannot be used directly:

$$P(Y(1/2) > 0 \Big| Y(1) = \sigma) = ?$$

We need to use the conditional density of $Y(1/2)|Y(1)$ to compute this probability. To ease notation, let $X = Y(1/2), Y = Y(1)$. We can quote the result of Ross problem 3.16 part b for the case of bivariate Normal distribution as followed : the conditional density of $X$ given $Y = \sigma$ is normal with mean $\mu_x + (\rho\sigma_x/\sigma_y)(\sigma - \mu_y)$ and variance $\sigma_x^2(1 - \rho^2)$. Here $\mu_x = \mu_y = 0, \sigma_x = \frac{\sigma}{\sqrt{2}}, \sigma_y = \sigma, \rho = \frac{\sqrt{2}}{2}$. Hence

$$X|Y = \sigma \sim N(\frac{\sigma}{2}, \frac{\sigma^2}{4}).$$

We can then compute the probability using the density of this Normal distribution.

# 2 Processes related to Brownian motion

A Brownian motion is like a building block which one can use to build other processes. (To be precise, it is the white noise "process" $dW_t$, the differential of the Brownian motion which is the building block. We won't cover the white noise process here, because once again it belongs to a stochastic calculus treatment). If $W_t$ is a Brownian motion, we can have Brownian motion with drift

$$X_t = \mu t + \sigma W_t,$$

Geometric Brownian motion:

$$X_t = e^{\mu t + \sigma W_t},$$

the Ornstein-Uhlenbeck process:

$$X_t = e^{-\alpha t/2} W_{e^{\alpha t}}$$

and Brownian bridge:

$$X_t = W_t | W_T = x, t \leq T,$$

to name a few. The Brownian motion with drift is easy to understand. Geometric Brownian motion can be viewed as the exponential of Brownian motion with drift, but it is deeper than that. Geometric Brownian motion is the model for exponential growth under influence of white noise:

$$\begin{aligned} dX_t &= (\mu + \frac{1}{2}\sigma^2) X_t dt + \sigma X_t dW_t \\ X_0 &= 1. \end{aligned}$$

It is used in financial model for asset price such as the Black-Scholes model. Indeed the Black-Scholes formula is just about giving the explicit formula for the computation

$$E(e^{-rT} \max(X_T - K, 0)),$$

where $K$ is a positive constant and $X_t = e^{(r - \frac{1}{2}\sigma^2)t + \sigma W_t}$ is a Geometric Brownian motion. Brownian bridge is Brownian motion conditional on its terminal value $W_T$. It can also be viewed as a "clamped" Brownian motion (since we also have $W_0 = 0$) and hence the term bridge. Finally Ornstein Uhlenbeck process is at heart a time-changed Brownian motion. Ornstein-Uhlenbeck process is also the continuous version of the AR(1) model (autoregressive of order 1) :

$$\begin{aligned} dX_t &= 2\alpha X_t dt + \sqrt{\alpha} dW_t \\ X_0 &= 0. \end{aligned}$$

We will discuss more about Brownian bridge and Ornstein-Uhlenbeck process below.

# 3 Hitting times and the maximum of Brownian motion

As mentioned earlier in the Poisson process chapter, one of the most interesting questions one can ask about a stochastic process is about its time dimension. Brownian motion is (very) different from Poisson process because it has nowhere differentiable paths (while Poison paths are piecewise constant and increasing). Thus given a level $a > 0$, ($a < 0$ has a similar result due to the symmetry of Brownian motion) asking for the distribution of $T_a := \inf\{t \geq 0 : W_t = a\}$ is an interesting question (also with practical applications). Closely related to this question is the question about the distribution of the running max of a Brownian motion: $X_t = \max_{0\leq u\leq t} W_t$.

First a remark before we compute the distribution of $T_a$. It's also an interesting (or amazing) feature of Brownian motion that $T_a := \inf\{t \geq 0 : W_t > a\}$. The first time the Brownian motion hits $a$ (from below) is also the first time it crosses over $a$ (thus the scenario where it hits $a$ and "bounces" down has probability 0).

The main idea for the computation of $P(T_a \leq t)$ is the reflection principle: for every path of $W$ that crosses $a$ before time $t$ and reaches above $a$ at time $t$, there is a reflected path ( obtained by reflecting the part of the previous path on the interval $[T_a, t]$ over the line $y = a$ ) that crosses $a$ at the same time but reaches below $a$ at time $t$. In terms of probability, this reflection principle is

$$P(T_a \leq t, W_t > a) = P(T_a \leq t, W_t < a) = P(T_a \leq t, W_t \leq a).$$

We have

$$\begin{aligned}
P(T_a \leq t) &= P(T_a \leq t|W_t > a)P(W_t > a) + P(T_a \leq t|W_t \leq a)P(W_t \leq a) \\
&= P(T_a \leq t|W_t > a)P(W_t > a) + P(T_a \leq t, W_t \leq a) \\
&= P(T_a \leq t|W_t > a)P(W_t > a) + P(T_a \leq t, W_t \geq a) \\
&= P(T_a \leq t|W_t > a)P(W_t > a) + P(T_a \leq t|W_t \geq a)P(W_t > a) \\
&= 2P(W_t > a),
\end{aligned}$$

where the last equality comes from the fact that $P(T_a \leq t|W_t \geq a) = 1$ which in turn follows from the continuity of $W_t$ in $t$. Lastly we can compute $P(W_t > a)$ from the density of $W_t$ and differentiate that with respect to $t$ to obtain the density of $T_a$.

Now let $X_t = \max_{0\leq s\leq t} W_s$. Then

$$P(X_t \geq x) = P(W_s \geq x, \text{ some } s, 0 \leq s \leq t) = P(T_x \leq t),$$

which we obtained above.

Finally, the distribution of $T_{-b}$ for some $b > 0$ is the same as the distribution of $T_b$ by the symmetry of Brownian motion. We have

$$P(T_{-b} < T_a) = \frac{a}{a+b}.$$

This follows from the (symmetric) gambler's ruin problem and the fact that Brownian motion can be approximated as the limit of a symmetric random walk.

# 4 Gaussian processes

Brownian motion enjoys many nice properties. In this and the next section, we discuss some generalizations of Brownian motion. That is, we consider only a class of processes that retain one of these properties and investigate particular examples of such processes.

As discussed above, Brownian motion has multivariate Normal distribution as its finite dimensional distribution. We define any process with such finite dimensional distribution as a Guassian process. There are several things to keep in mind about multivariate Normal distribution in regards to Guassian process:

a. A multivariate Normal distribution is completely specified by its mean and covariance matrix. Therefore, a Gaussian process $X_t$ is completely specified by the mean function $\mu(t) = E(X_t), t \geq 0$ and covariance function $\sigma(s,t) = Cov(X_s, X_t), 0 \leq s \leq t$.

b. If $\mathbf{X} \in \mathbb{R}^n$ has a multivariate Normal distribution with mean $\boldsymbol{\mu}$ and covariance matrix $\Sigma$ then $\mathbf{Y} = A\mathbf{X} \in \mathbb{R}^m$ also has a multivariate distribution with mean $A\boldsymbol{\mu}$ and covariance matrix $A\Sigma A^T$.

c. If $\mathbf{X}$ has a multivariate Normal distribution and $\Sigma = I_n$ is the covariance matrix of $\mathbf{X}$ then the elements of $\mathbf{X}$ are also independent. That is zero-covariance implies independence under multivariate Normal distribution assumption (which is of course not true in general).

d. If $\mathbf{X} = [\mathbf{X}_1, \mathbf{X}_2]^T$ has a multivariate Normal distribution then $\mathbf{X}_1 \big| \mathbf{X}_2$ also has a multivariate Normal distribution with mean with mean $\boldsymbol{\mu_1} + \Sigma_{12}\Sigma_{22}^{-1}(\mathbf{X}_2 - \boldsymbol{\mu}_2)$ and covariance matrix $\Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}$.

Thus Brownian motion is a Gaussian process with mean function $\mu(t) = 0$ and covariance function $\sigma(s,t) = \min(s,t)$. This can be viewed as an alternative definition of Brownian motion in the following way. Note that this implies for $r < s < t$ :

$$\begin{aligned} Cov(X_t - X_s, X_s - X_r) &= Cov(X_t, X_s) + Cov(X_s, X_r) \\ &- Cov(X_s, X_s) - Cov(X_t, X_r) \\ &= s + r - s - r = 0. \end{aligned}$$

Thus we can deduce the fact that Brownian motion has independent increments from its Gaussian process characterization.

There are of course many other Gaussian processes besides Brownian motion (as many as one can envision a reasonable structure of the mean vector $\boldsymbol{\mu}$ and the covariance matrix $\Sigma$ ! ). Our first example is the Brownian bridge, which has applications in other areas, e.g. Monte Carlo simulation and financial modelling.

$X_t, 0 \le t \le T$ is a Brownian bridge if $X_t = W_t \big| W_T = x, 0 \le t \le T$, for some value of $x$. Thus a Brownian bridge is a Brownian motion conditional on its terminal value being specified. From property d of the multivariate Normal distribution above, a Brownian bridge is a Gaussian process. We can use the formula provided in property d to compute its finite dimensional mean vectors and covariance matrices.

$$E(X_t) = E(W_t | W_T = x) = 0 + tT^{-1}(x - 0) = \frac{t}{T}x,$$

where $\boldsymbol{\mu_1} = 0, \Sigma_{12} = t, \Sigma_{22} = T, \mathbf{X}_2 = x, \boldsymbol{\mu}_2 = 0$. On the other hand, for $s < t < T$ and $\mathbf{X} = (W_s, W_t, W_T)^T$

$$\Sigma = \begin{bmatrix} s & s & s \\ s & t & t \\ s & t & T \end{bmatrix}.$$

Thus

$$\Sigma_{11} = \begin{bmatrix} s & s \\ s & t \end{bmatrix}, \Sigma_{12} = \begin{bmatrix} s \\ t \end{bmatrix}, \Sigma_{21} = \begin{bmatrix} s & t \end{bmatrix}, \Sigma_{22} = T.$$

We thus have

$$\begin{aligned} \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21} &= \begin{bmatrix} s & s \\ s & t \end{bmatrix} - \frac{1}{T}\begin{bmatrix} s \\ t \end{bmatrix}\begin{bmatrix} s & t \end{bmatrix} \\ &= \begin{bmatrix} s - \frac{s^2}{T} & s - \frac{st}{T} \\ s - \frac{st}{T} & t - \frac{t^2}{T} \end{bmatrix} = \begin{bmatrix} \frac{s}{T}(T-s) & \frac{s}{T}(T-t) \\ \frac{s}{T}(T-t) & \frac{t}{T}(T-t) \end{bmatrix}. \end{aligned}$$

We thus also have that Brownian bridge $X_t$, conditional on $W_T = x$ is a Gaussian process with mean function $\mu(t) = \frac{t}{T}x$ and covariance function $\sigma(s,t) = \frac{s}{T}(T-t)$. Notice that this implies $X(T) = x$ since $Var(X(T)) = \sigma(T,T) = 0$ and $E(X(T)) = \mu(T) = x$.

The mean function structure of the Brownian bridge suggests the following:

**Proposition 4.1.** *$X(t) = W_t - \frac{t}{T}W_T, 0 \le t \le T$ is a Brownian bridge with terminal value $X(T) = 0$. It is also independent of $W_T$.*

Remark: Note that in this formulation, $W_T$ is treated as a random variable. Also the conditioning on $W_T$ is implied since without knowing $W_T$ we cannot know the value of $X(t)$. Also note the interesting fact on how two terms of a Brownian bridge are always positively correlated.

First, it is clear that $X(t)$ is a Gaussian process, since it is a linear combination of the $W_t$ process. It is also clear that $X(T) = 0$. We actually just need to verify the mean and covariance structure of $X(t)$ :

$$\begin{aligned}
\mu(t) &= E(X_t) = E(W_t - \frac{t}{T}W_T) = 0 \\
\sigma(s,t) &= Cov(X_s, X_t) = Cov(W_s - \frac{s}{T}W_T, W_t - \frac{t}{T}W_T) \\
&= Cov(W_s, W_t) - \frac{s}{T}Cov(W_T, W_t) - \frac{t}{T}Cov(W_s, W_T) + \frac{st}{T^2}Cov(W_T, W_T) \\
&= s - \frac{st}{T} - \frac{st}{T} + \frac{st}{T} \\
&= \frac{s}{T}(T-t).
\end{aligned}$$

Finall, the independence with $W_T$ again can be verified with the covariance structure:

$$Cov(X_t, W_T) = Cov(W_t - \frac{t}{T}W_T, W_T) = t - \frac{t}{T}T = 0.$$

To recover the general terminal value condition of Brownian bridge we just need to add $x$ to the original formula:

**Proposition 4.2.** *$X(t) = W_t - \frac{t}{T}(W_T - x), 0 \le t \le T$ is a Brownian bridge with terminal value $X(T) = x$. It is also independent of $W_T$.*

Lastly, we can also specify a general initial value:

**Proposition 4.3.** *$X(t) = (W_t + a) + \frac{t}{T}[(b-a) - W_T], 0 \le t \le T$ is a Brownian bridge with initial value $X(0) = a$ and terminal value $X(T) = b$. It is also independent of $W_T$.*

We leave both straightforward verications to the reader.

# 5 Stationary and weakly stationary processes

A process is stationary if its finite dimensional distribution does not change under a time translation. Specifically, $X_t$ is stationary if for any $t_i, i = 1, \cdots, n$ and $s > 0$

$$(X_{t_1}, X_{t_2}, \cdots, X_{t_n}) \stackrel{d}{=} (X_{t_1+s}, X_{t_2+s}, \cdots, X_{t_n+s}).$$

We have seen the stationary concept before, in the increments of Poisson process and Brownian motion. Thus for some fixed $L, X(t) := N(t+L) - N(t), t \geq 0$ and $X(T) := W(t+L) - W(t)$ are two examples of stationary processes. If we consider discrete times, then any ergodic Markov chain whose initial distribution is its stationary distribution is also stationarsy. Another trivial example is an iid sequence of RV is a (discrete time) stationary process.

On the other hand, note in particular that Brownian motion and Poisson process are NOT stationary. Similarly, a symmetric random walk is not stationary. Stationarity is an important concept in time series modeling. On the other hand, this definition is too strict and many process of interest does not satisfy the property. We thus define a weaker concept, not surprisingly, referred to as weak stationarity : a process is weakly stationary if its mean function $\mu(t)$ is constant and its covariance function $\sigma(s,t)$ is only a function of $t - s$. It is easy to check that strong stationarity implies weak stationarity but not vice versa. On the other hand, a Gaussian process is uniquely defined by its mean and covariance function. Thus if a process is weakly stationary AND it is Gaussian then it is also strongly stationary. An example of this is the Ornstein-Uhlenbeck process.

Let $W_t$ be a Brownian motion. Then $X_t = e^{-\alpha t/2} W_{e^{\alpha t}}$ is an Ornstein-Uhlenbeck (OU) process. We have

$$\begin{aligned} E(X_t) &= 0; \\ Cov(X_s, X_t) &= e^{-\alpha(t+s)/2} Cov(W_{e^{\alpha s}}, W_{e^{\alpha t}}) \\ &= e^{-\alpha(t+s)/2} e^{-\alpha/2[(t+s)-|t-s|]} = e^{-\alpha|t-s|/2}, \end{aligned}$$

where we used the fact that $\min(a,b) = a + b - |a-b|$. This shows that the Ornstein-Uhlenbeck is weakly stationary. Since $W_t$ has multivariate Normal distribution, $W_{e^{\alpha t}}$ also has multivariate Normal distribution. Thus the Ornstein-Uhlenbeck process is a Gaussian process and it is strongly stationary.

As mentioned above, the OU process is the continuous version of AR(1) model. In discrete time, the AR(1) model is as followed:

$$X_{n+1} = \lambda X_n + Z_{n+1}, |\lambda| < 1,$$

where $Z_i, i \geq 0$ are uncorrelated with mean 0. We also specify that the variance of $Z_i, i \geq 1$ to be $\sigma^2$ and variance of $Z_0$ to be $\frac{\sigma^2}{1-\lambda^2}$, for reasons that we will see soon. We have

$$X_n = \sum_{i=0}^{n} \lambda^{n-i} Z_i,$$

and thus $E(X_n) = 0$. For $m < n$

$$\begin{aligned}
Cov(X_m, X_n) &= Cov(\sum_{i=0}^{m} \lambda^{m-i} Z_i, \sum_{i=1}^{n} \lambda^{n-i} Z_i) \\
&= Cov(\sum_{i=0}^{m} \lambda^{m-i} Z_i, \sum_{i=0}^{m} \lambda^{n-i} Z_i) \\
&= \sigma^2 \lambda^{m+n} (\sum_{i=1}^{m} \lambda^{-2i} + \frac{1}{1-\lambda^2}) \\
&= \sigma^2 \lambda^{m+n} (\sum_{i=-m}^{-1} \lambda^{2i} + \frac{1}{1-\lambda^2}) \\
&= \sigma^2 \lambda^{m+n} (\frac{\lambda^{-2m} - 1}{1-\lambda^2} + \frac{1}{1-\lambda^2}) \\
&= \sigma^2 \frac{\lambda^{n-m}}{1-\lambda^2}.
\end{aligned}$$

Thus the AR(1) model is weakly stationary. It is clear that it is not strongly stationary.

Remark: If we do not want to start $Z_0$ with a particular variance (which can be viewed as a initial distribution to jump start the process into stationarity right away), we can view $X_n$ as

$$X_n = \sum_{i=-\infty}^{n} \lambda^{n-i} Z_i.$$

Following the same procedure as above, we can show that $E(X_n) = 0$,

$$Cov(X_m, X_n) = \sigma^2 \frac{\lambda^{n-m}}{1-\lambda^2}, m < n.$$

Our last example is of stationarity is a moving average process. Let $Z_1, Z_2, \cdots$ be uncorrelated with mean 0 and variance $\sigma^2$. Then

$$X_n = \frac{1}{k+1} \sum_{i=n-k}^{n} Z_i, n \geq k$$

is called a moving average process. We have $E(X_n) = 0$ and for $m < n$

$$\begin{aligned} Cov(X_m, X_n) &= Cov(\frac{1}{k+1} \sum_{i=m-k}^{m} Z_i, \frac{1}{k+1} \sum_{i=n-k}^{n} Z_i) \\ &= 0 \text{ if } m \leq n-k \\ &= \frac{(m-n+k+1)\sigma^2}{(k+1)^2} \text{ if } m > n-k. \end{aligned}$$

Thus $Cov(X_m, X_n)$ is only a function of $n - m$ and the moving average process is weakly stationary.