STAT 443 Lab 3

Caden Hewlett

2024-01-26

Question 1

We'll first simulate the data for $\{Z_t\}_{t \in \mathbb{Z}}$, where

$$X_t = Z_t - 1.3 Z_{t-1} + 0.4 Z_{t-2}$$

```
coefs = c(-1.3, 0.4)

# simulate the MA(2) process
simulated_data <- arima.sim(n = 1000,
                            model = list(ma = coefs),
                            sd = sqrt(0.01))
```

Part 1

Create a time series plot of the data.

```
p1data = fortify.zoo(simulated_data)
p1 <- ggplot(p1data, aes(x = Index, y = simulated_data)) +
  geom_line(color = "#E63946", linewidth = 0.65) +
  labs(
    title = "Simulated Data From an MA(2) Process",
    subtitle = "Assumed Z ~ WN(0, 1)",
    x = "Value",
    y = "Index"
  ) + theme_bw() +
  theme(panel.grid.minor = element_line(
    color = "grey90",
    linetype = "dashed",
    linewidth = 0.5
  ))
print(p1)
```

Simulated Data From an MA(2) Process
Assumed Z ~ WN(0, 1)
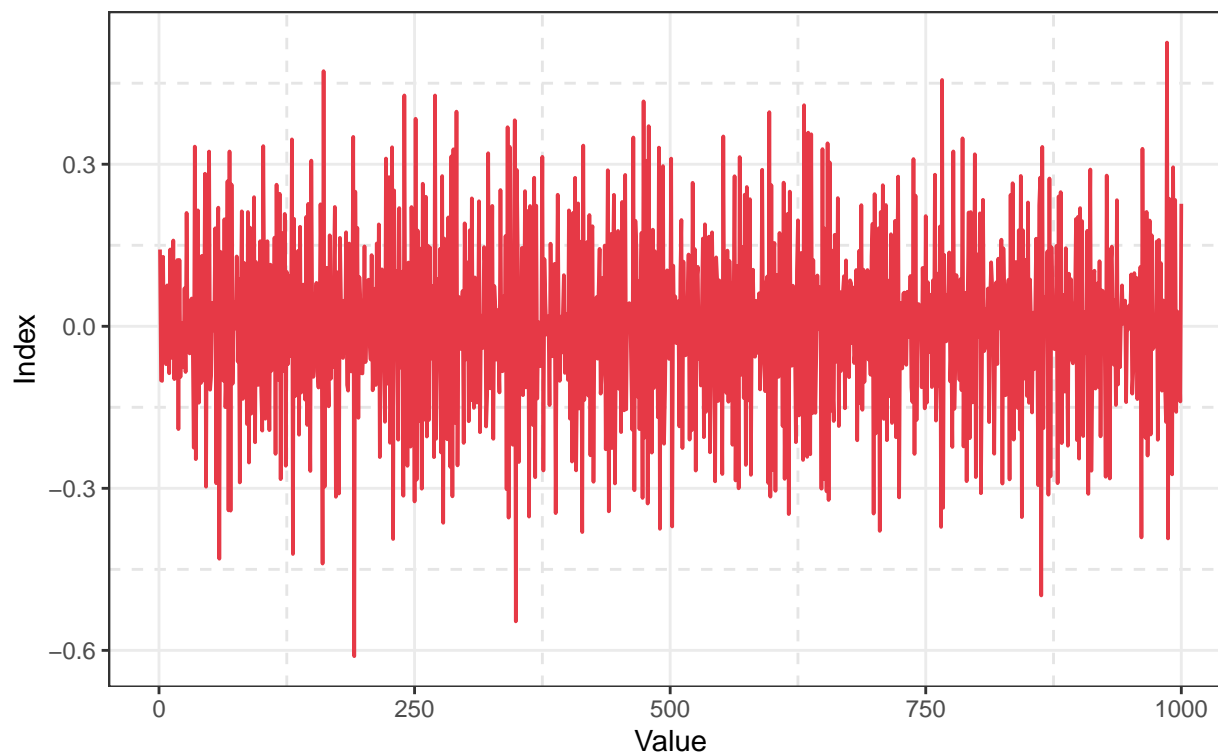
What would we expect the acf of $\{X_t\}$ to look like?

I would expect the ACF to (obviously) have a value of 1 for $h = 0$. Further, I would expect the autocorrelation for lag values $h \in \{1\}$ to be negative. The reason is this: if an innovation occurs at time t, we can consider our model forward-shifted by two to think about what is going to happen two time steps from now.
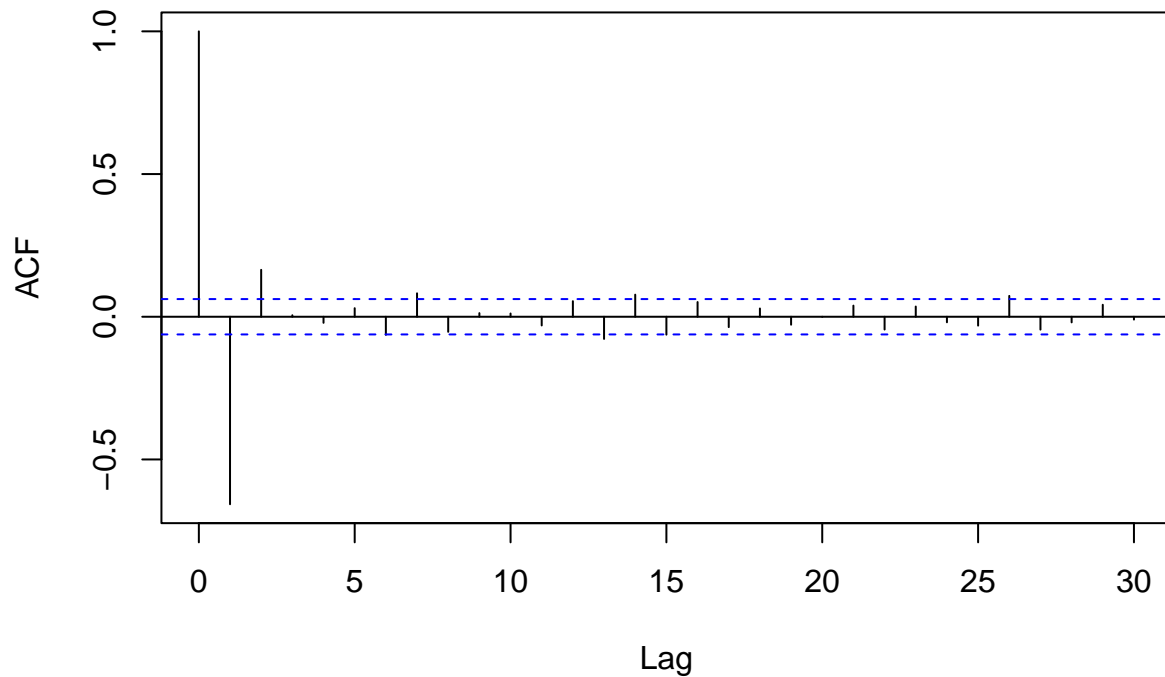
$$X_{t+2} = Z_{t+2} - 1.3Z_{t+1} + 0.4Z_t$$

.

So the innovation is going to "ripple" forward, but be multiplied by a factor of -1.3 and become negatively-valued. So I would *roughly* expect a negative correlation for lags of 1 and 2, descending into noise for lags greater than 2.

Part 2

Here is the acf for the simulated data:

```
acf(simulated_data)
```

Series simulated_data

The behaviour of this ACF is sort of what I expected, but a little bit different. Interestingly, after running it a few times I began to notice that the autocorrelation for lag $h = 1$ was "more negative" than I would have expected. Unsurprisingly, the process quickly devolves into white noise for $h > 2$, however it retains the "up and down" alternating pattern due to the negative coefficient in the model. Another interesting observation was that the observed ACF value for lag $h = 2$ was consistently moderately/weakly positive over repeated trials. I have a few ideas on why this is happening (to be discussed later.)

Part 3

Create the ACF Correlogram using `ARMAacf`.

```
#  theoretical ACF for an MA(2) process
acf_values <- ARMAacf(ma = c(-1.3, 0.4), lag.max = 10)

# plotted theoretical ACF
plot(acf_values, x = 0:(length(acf_values)-1), type = "h",
     main = "Theoretical ACF of Simulated MA(2) Process",
     xlab = "Lag", ylab = "ACF", ylim = c(-0.8, 1))
abline(h = 0, lty = 'dashed')
abline(h = -0.2, lty = 'dashed', col = 'red')
abline(h = 0.2, lty = 'dashed', col = 'red')
```

Theoretical ACF of Simulated MA(2) Process
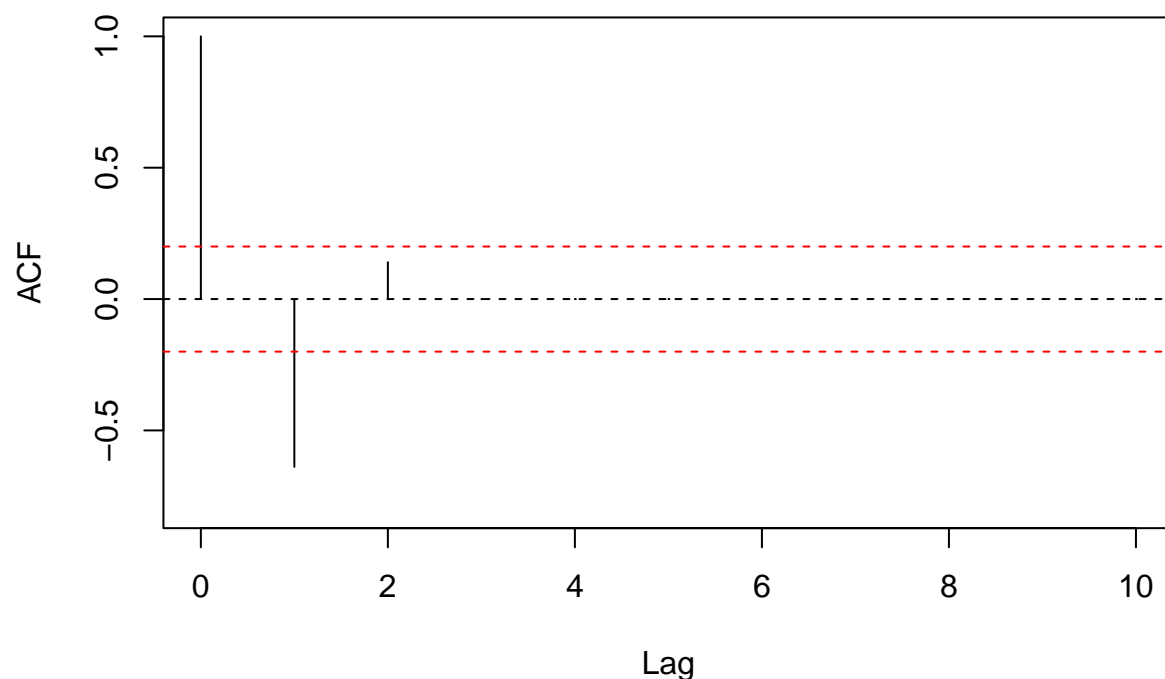
Part 4

Explain the behaviour of the sample acf here.

To explain the AVF values, perhaps it is best to use an example.

Let's say there was an innovation with value 1 at time t, such that

$$X_t = Z_t - 1.3Z_{t-1} + 0.4Z_{t-2} = 1 + -1.3(0) + 0.4(0) = 1$$

And, for the sake of simplicity, Z_{t-1} and Z_{t-2} are both zero.

The effects of this innovation are going to "ripple forward." Let's say that time step $t+1$ has no innovation, and is a more regular value, e.g. $Z_{t+1} = 0.5$. In this case,

$$X_{t+1} = Z_{t+1} - 1.3Z_t + 0.4Z_{t-1} = 0.5 + -1.3(1) + 0.4(0) = -0.8$$

So, we can consider the pair $\{Z_t, Z_{t+1}\} = \{1, -0.8\}$. They're going to be negatively correlated on average (i.e. under this data production schema), as we can see in the ACF plot above for $h = 1$.

Let's continue with this example. Let's say that there is again no innovation at time step $t+2$, so the value of $Z_{t+2} = 0.3$. In this case,

$$X_{t+2} = Z_{t+2} - 1.3Z_{t+1} + 0.4Z_t = 0.3 + -1.3(-0.8) + 0.4(1) = 1.73$$

Now, there will be a weakly positive correlation between $\{Z_t, Z_{t+2}\} = \{1, 1.73\}$ for a lag of 2, which we also see in the ACF plot.

Then for further iterations, the process devolves into noise, which makes sense considering it is an $MA(2)$ process. This is reflected in the sample acf plot as zero autocorrelation.
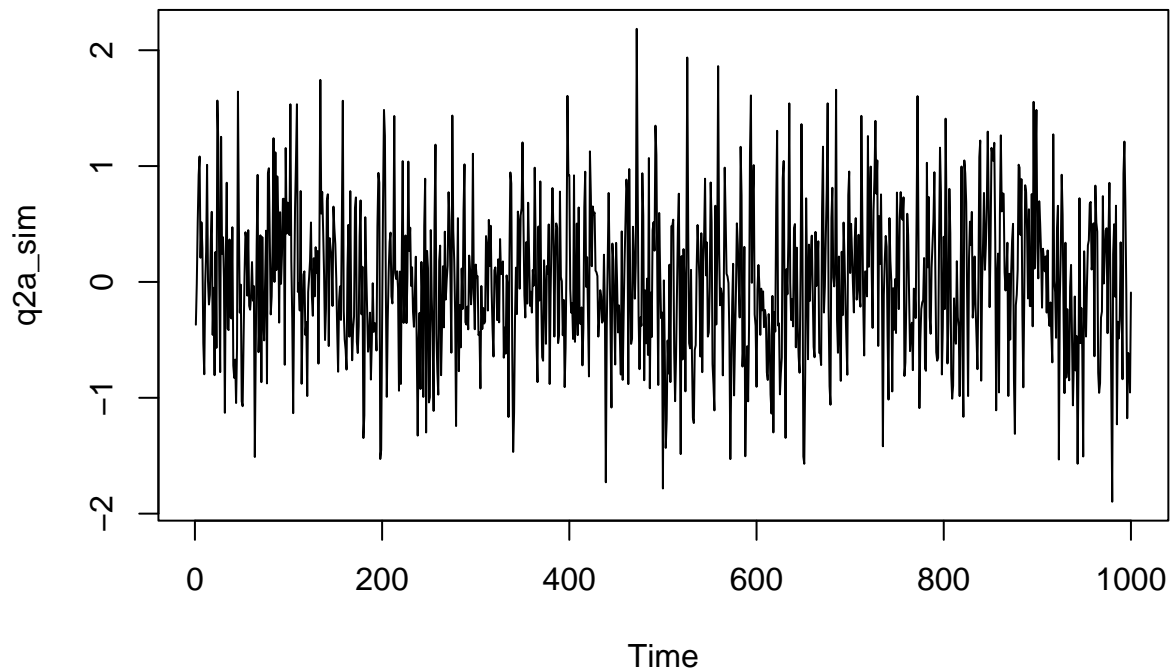
Question 2

Now, we have the following process, for $\{Z_t\}_{t\in\mathbb{Z}}$ for $Z_t \sim \text{WN}(0, \sigma = \sqrt{0.4})$.

Part 1

We generate the data and plot the sample acf for $X_t = Z_t + 0.25Z_{t-1}$ as follows:
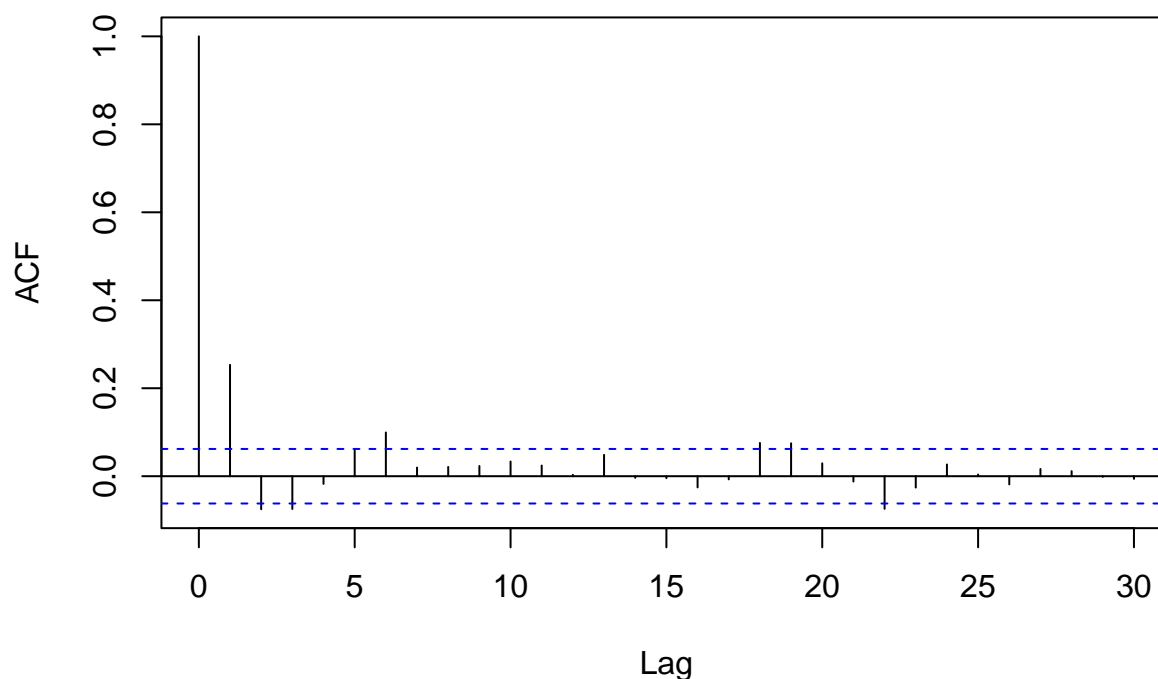
```r
### MA(1) proc with a coefficient of 0.4
q2a_sim <- arima.sim(n = 1000, model = list(ma = c(0.25)), sd = sqrt(0.4))
plot(q2a_sim, main = "Plot of Simulated MA(1) Process with Coefficient of 0.25")
```


Plot of Simulated MA(1) Process with Coefficient of 0.25

```r
acf(q2a_sim)
```

Series q2a_sim

Comments: There appears to be a moderately high (> 0.2 error bound, appx. 0.35) autocorrelation for lag $h = 1$, which declines somewhat quickly into pure noise. This makes sense, considering the Z_{t-1} coefficient is not "scaling up" the data, so the effect of an innovation will decline over time somewhat rapidly.
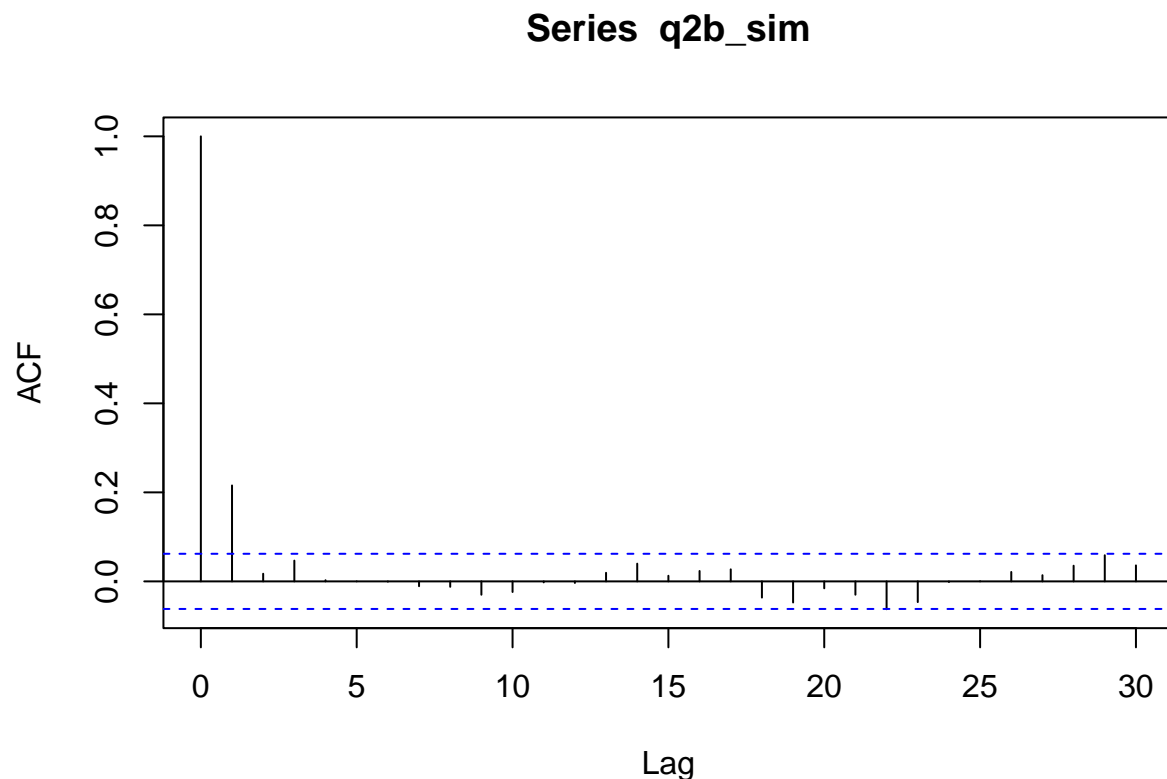
Part 2

Now, we generate the data and plot the sample acf for $X_t = Z_t + 4Z_{t-1}$ as follows:

```
### MA(1) proc with a coefficient of 4
q2b_sim <- arima.sim(n = 1000, model = list(ma = c(4)), sd = sqrt(0.4))
plot(q2b_sim, main = "Plot of Simulated MA(1) Process with Coefficient of 4")
```

Plot of Simulated MA(1) Process with Coefficient of 4


```r
acf(q2b_sim)
```
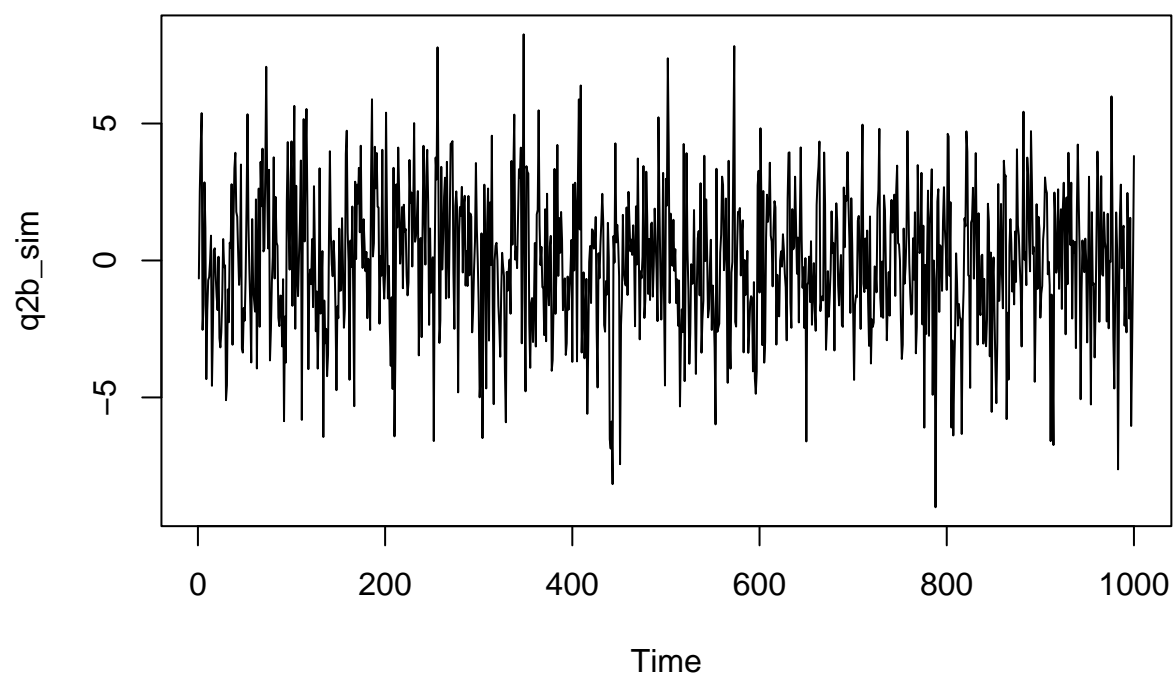
Series q2b_sim

Comments Again, we see a decently positive correlation. Importantly, however, the magnitude of the data itself is much greater in the plot of the simulations.

Part 3

Repeat multiple times. What do you notice?

Despite the magnitude of the time series being vastly different for the two simulation groups, the correlogram (acf plots) reflect similar autocorrelation coefficients for the initial lag $h = 1$ component, and both decline into noise. It's interesting that despite the difference in magnitude the two simulations have extremely similar acfs.

Question 3

Now, we have the following process:

$$X_t = \alpha X_{t-1} + Z_t, \text{ where } \alpha = -0.5$$

Where, $\forall t \in \mathbb{Z}, Z_t \sim N(0, (0.01)^2)$. We can identify this as an $AR(1)$ process.

Part 1

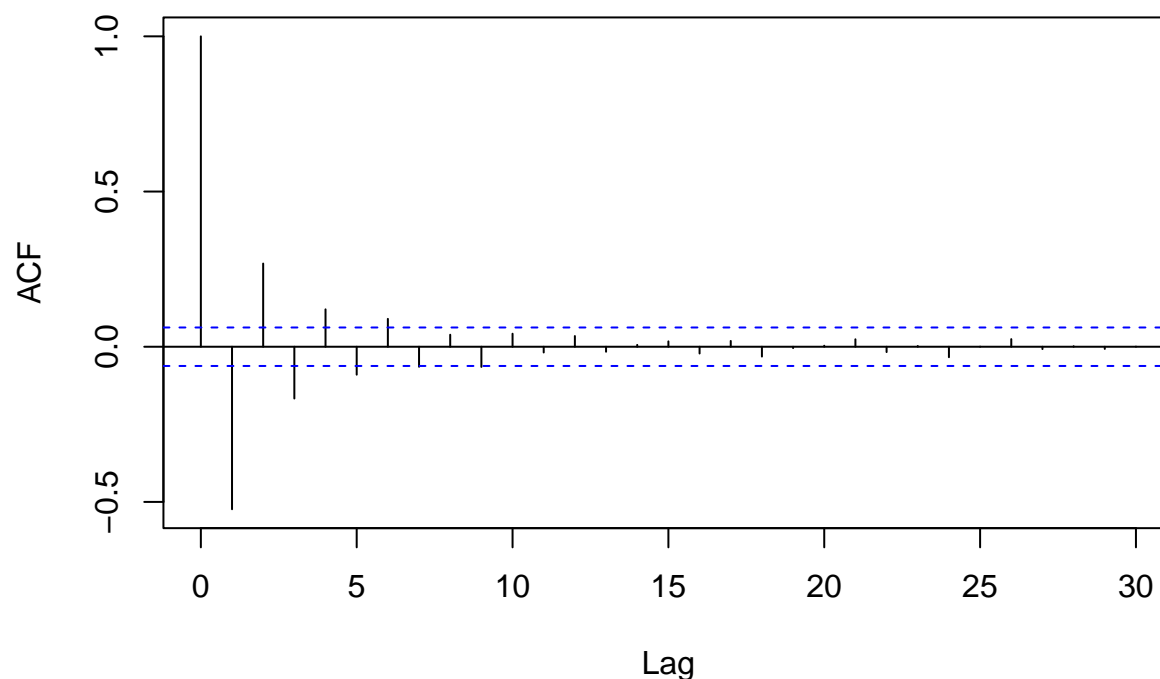Create a time series plot of your simulated data. Plot the sample acf, and comment on its behaviour.

```
### AR(1) proc
alpha = -0.5
q3a_sim <- arima.sim(n = 1000, model = list(ar = c(alpha)), sd = 0.01)

plot(q3a_sim, main = "Plot of Simulated AR(1) Process with Coefficient of -0.5")
```

Plot of Simulated AR(1) Process with Coefficient of –0.5
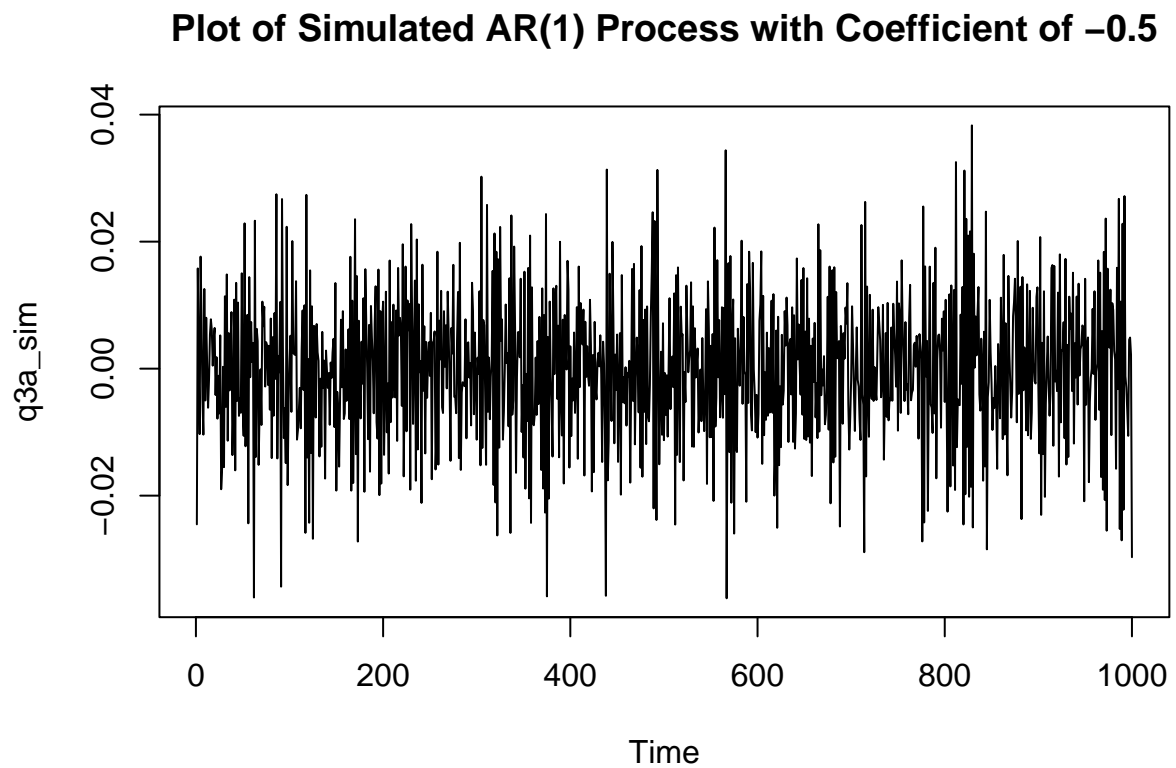

```
acf(q3a_sim)
```

Series q3a_sim

Comments: The ACF for the $AR(1)$ process is quite interesting. The observed acf as a function of lag h alternates in sign in each iteration (i.e. $\{+, -, +, -, \ldots\}$) and, with some error due to noise, tends to gradually decline in magnitude as $h \uparrow$.

Part 2

Let's increase the magnitude of α. We'll inspect $\alpha \in \{-0.75, -0.9, -0.999^-\}$ in the plots to follow. We will only be plotting the ACFs.
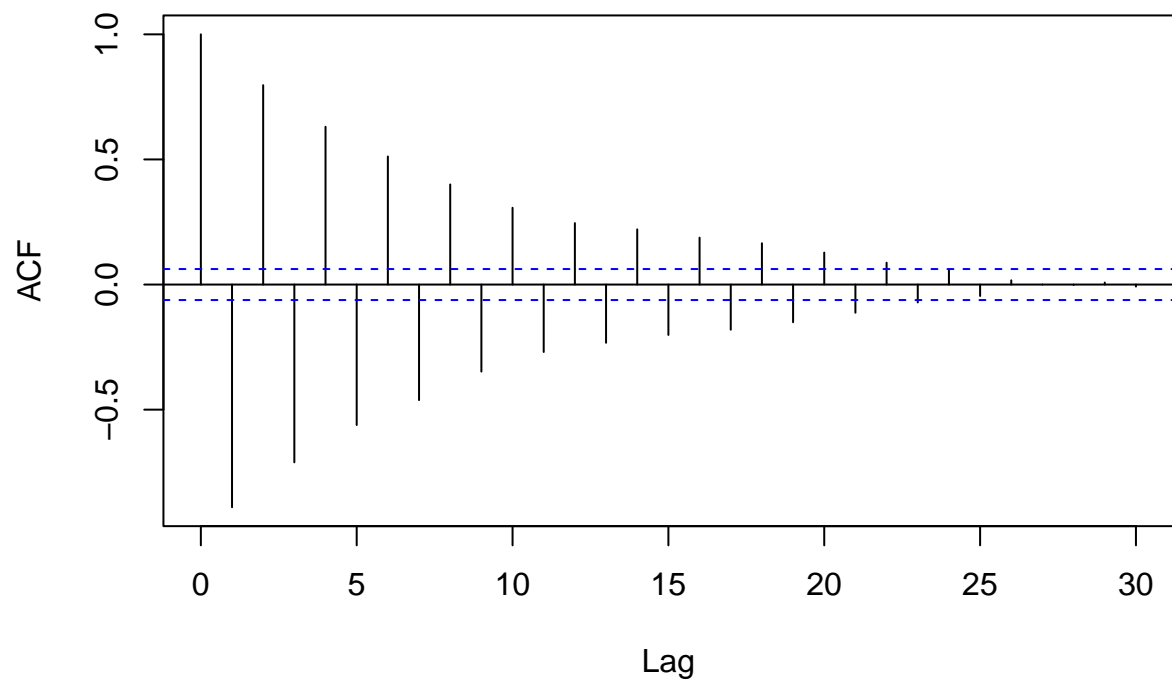
```r
### AR(1) proc
alpha_75 <- arima.sim(n = 1000, model = list(ar = c(-0.75)), sd = 0.01)
### AR(1) proc
alpha_90 <- arima.sim(n = 1000, model = list(ar = c(-0.90)), sd = 0.01)
### AR(1) proc
alpha_100 <- arima.sim(n = 1000, model = list(ar = c(-0.999999)), sd = 0.01)
# comparsion
acf(alpha_75)
```
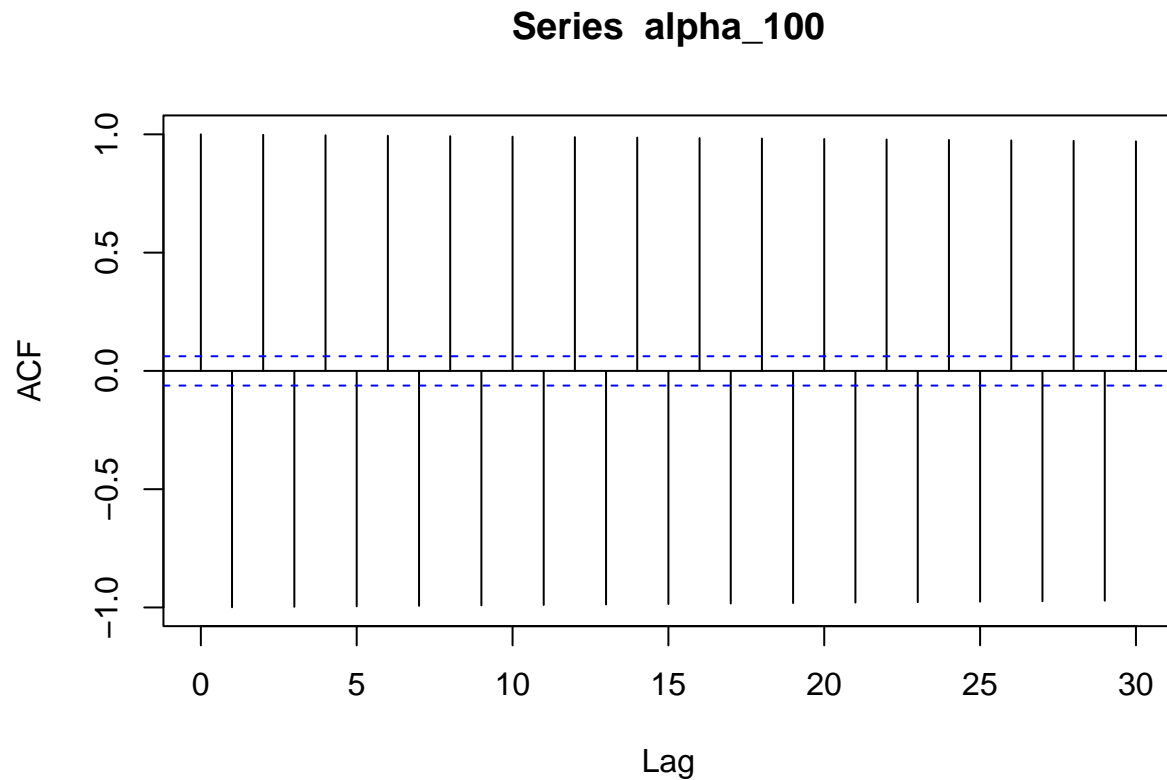
Series alpha_75
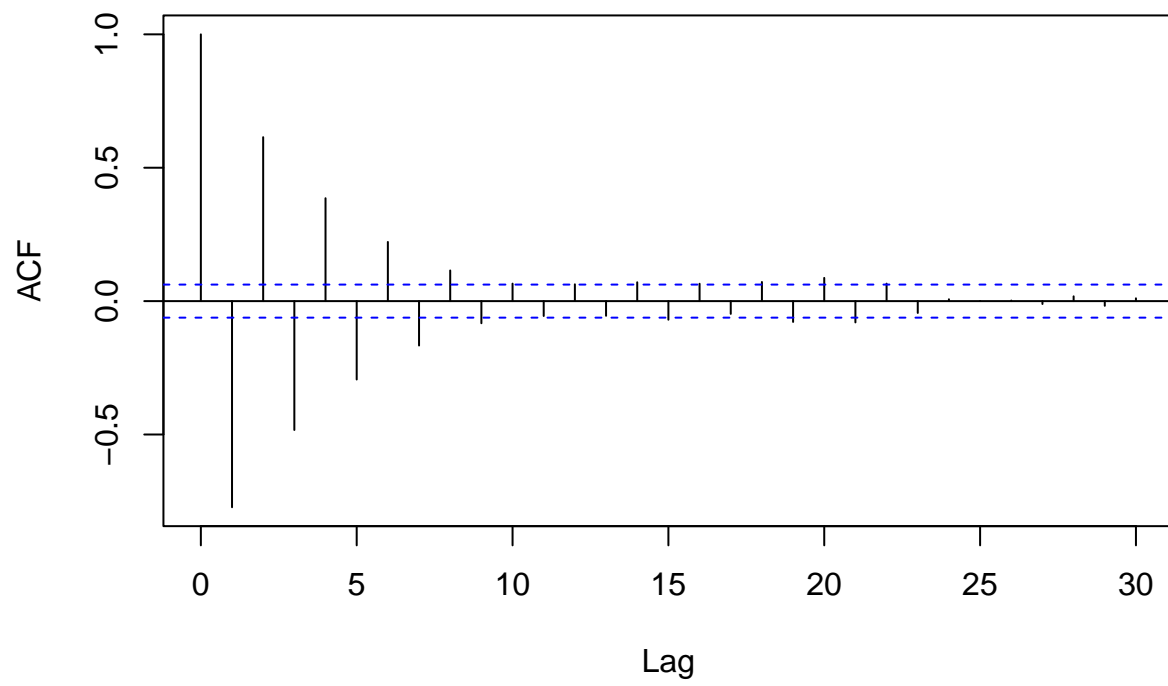

```
acf(alpha_90)
```

Series alpha_90


```
acf(alpha_100)
```

Series alpha_100


From these ACFs, we still see the alternating process as observed before, however, as $|\alpha| \to 1$, the magnitude of the $\rho(h)$ values declines at a slower and slower rate. On attempting $\alpha = -1$, we would lose stationarity. So, interestingly $(\alpha \to -1) \implies \forall h, |\rho(h)| = 1$. In other words, this is why we couldn't have a process with alpha at exactly -1, because it would be non-stationary.

Part 3

Now, we inspect how the process changes as $\alpha \uparrow 0$. We will inspect $\alpha \in \{-0.1, -0.01, -1 \times 10^{-9}\}$
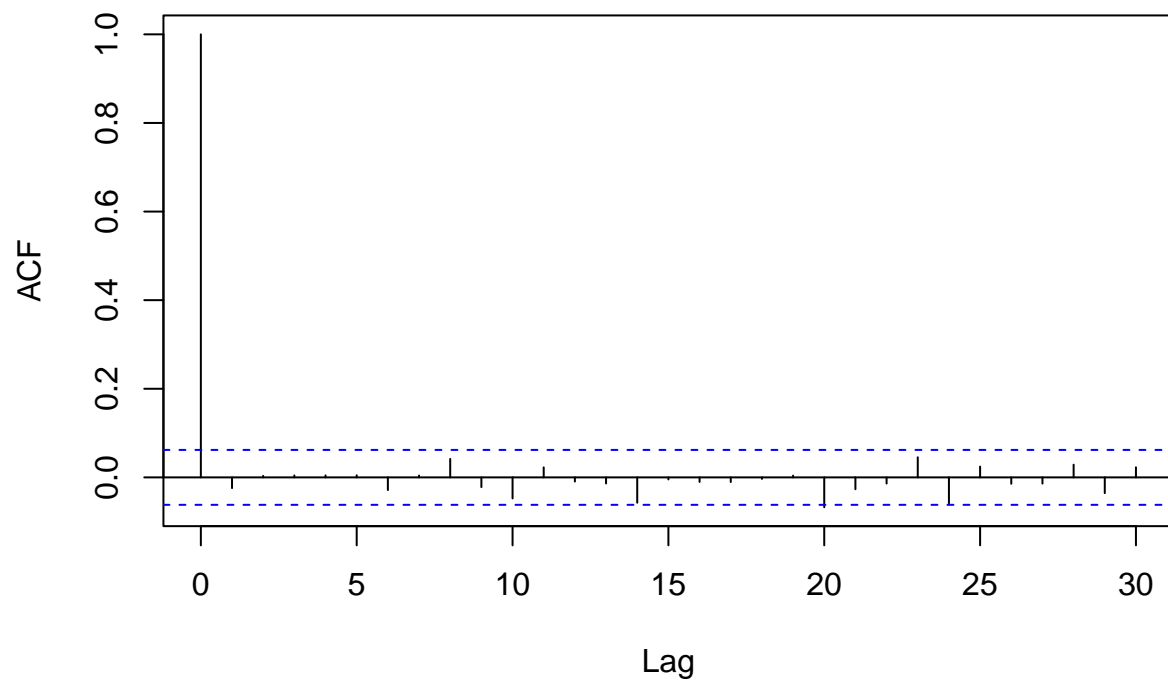
```
### AR(1) proc
alpha_1 <- arima.sim(n = 1000, model = list(ar = c(-0.1)), sd = 0.01)
### AR(1) proc
alpha_01 <- arima.sim(n = 1000, model = list(ar = c(-0.01)), sd = 0.01)
### AR(1) proc
alpha_smol <- arima.sim(n = 1000, model = list(ar = c(-1e-9)), sd = 0.01)
# comparsion
acf(alpha_1)
```
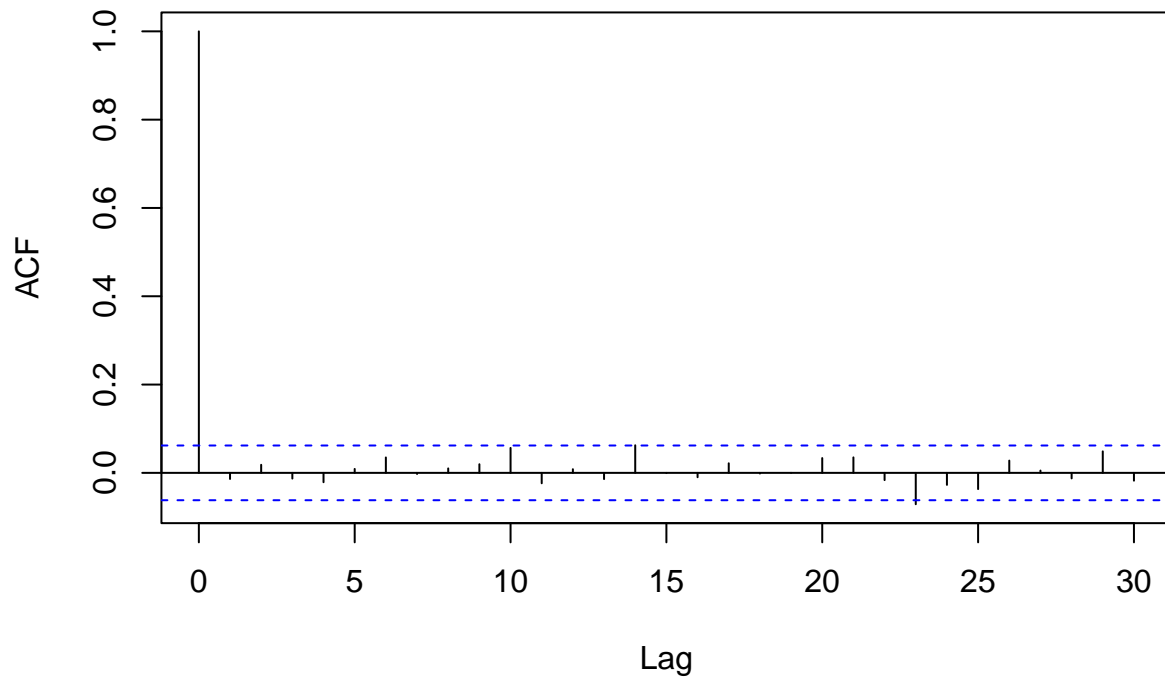
Series alpha_1
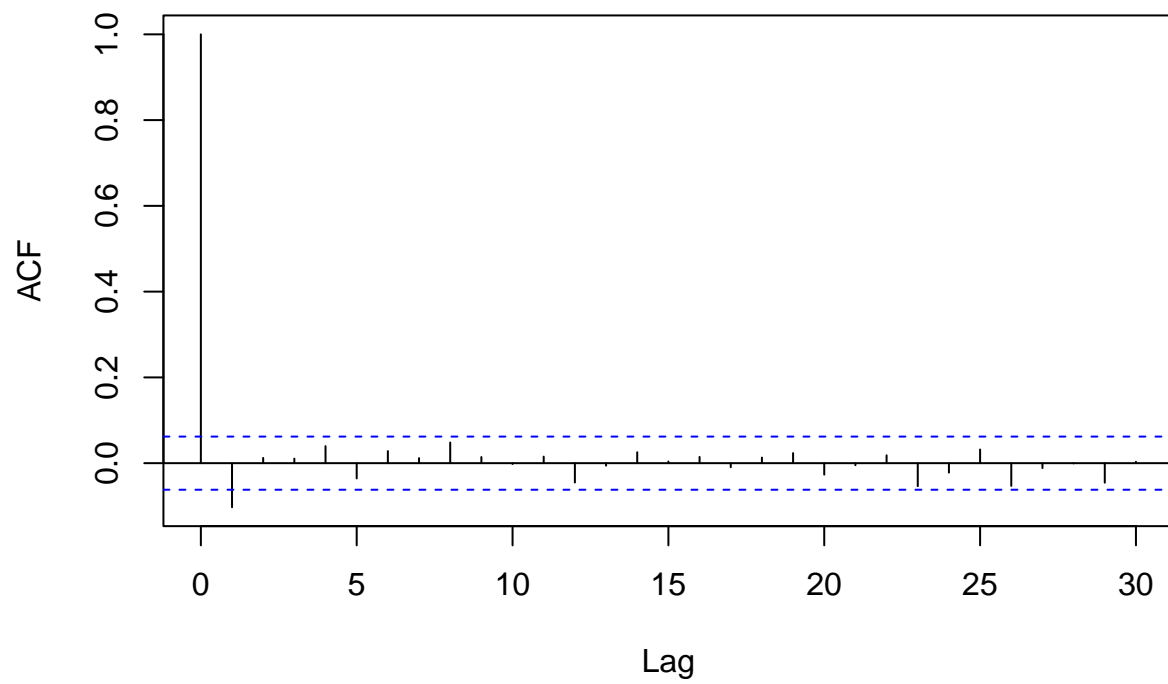

```r
acf(alpha_01)
```

Series alpha_01


```
acf(alpha_smol)
```

Series alpha_smol

Comments It seems that as α goes to zero, the acf declines into noise. This is evident of the serial dependence of the time series decilining more and more rapidly, since each time step will be multiplied by a smaller coefficient, thereby having little dependence.

For example, for $\alpha \approx 0$, we'd have:

$$X_t = \alpha X_{t-1} + Z_t = (0)X_{t-1} + Z_t \approx Z_t$$

So as α goes to zero, the $AR(1)$ process goes to simply Z_t.

Part 4

What happens to the sample acf if the parameter α changes sign? Experiment with this, changing the sign of α in your simulations. Explain what is happening.

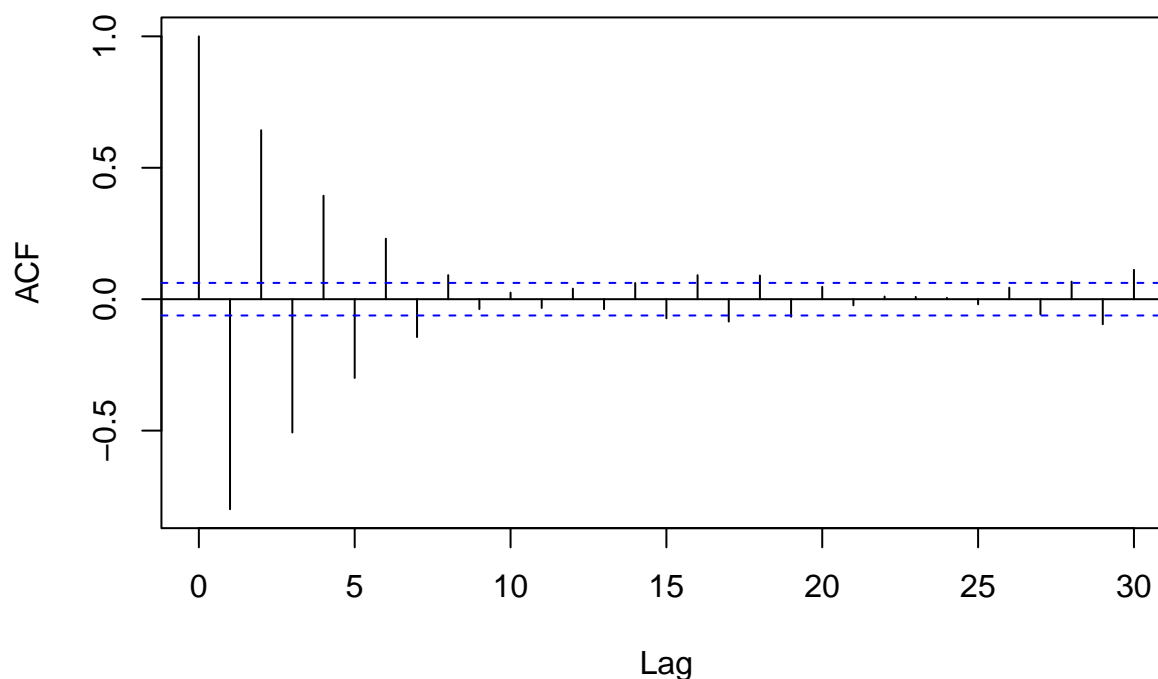For this, we'll compare $\alpha = \pm 0.8$.

```
### AR(1) proc
alpha_plus <- arima.sim(n = 1000, model = list(ar = c(0.8)), sd = 0.01)
### AR(1) proc
alpha_minus <- arima.sim(n = 1000, model = list(ar = c(-0.8)), sd = 0.01)
acf(alpha_plus)
```

Series alpha_plus
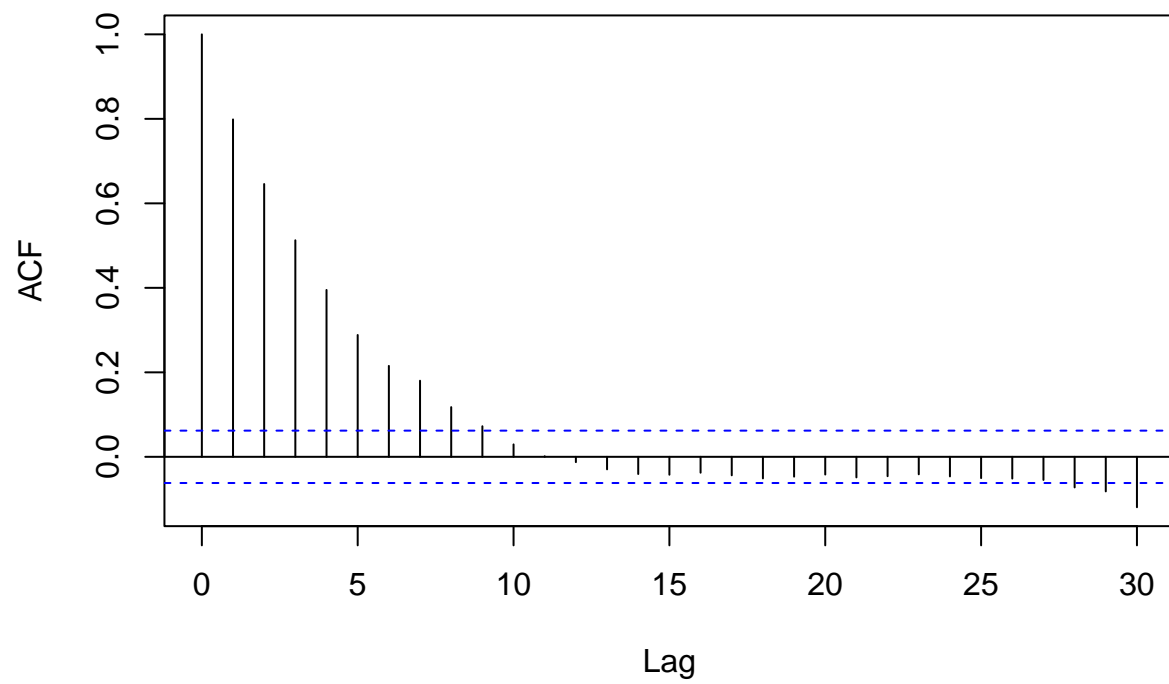

```r
acf(alpha_minus)
```

Series alpha_minus

Comments As we change the sign of α to be positive, we no longer see the alternating pattern in the sign of the acf, merely a consistent decline in serial dependence as a function of lag.

To explain this change, let's use some dummy variables. For simplicity, $Z_t = 0$, and our $X_{t-1} = 1$.

From this setup, for $\alpha = 0.8$, we'd have:

$$X_t = \alpha X_{t-1} + Z_t = (0.8)(1) + 0 = 0.8$$

Continuing this process (assuming zero noise), we'd have X_t values of $\{0.8, 0.64, 0.512, ..\}$ so then auto-correlation (for time t) would be found by comparing pairs $\{0.8, 0.64\}$ for $h = 1$, $\{0.8, 0.512\}$ for $h = 2$ and so forth. As you can see, the correlation will be decreasing over time, but never changing signs.

Conversely, if α was negative we'd have X_t values of $\{0.8, -0.64, 0.512, -0.4096, \dots\}$ so then auto-correlation (for time t) would be found by comparing pairs $\{0.8, -0.64\}$ for $h = 1$, $\{0.8, 0.512\}$ for $h = 2$, $\{0.8, -0.4096\}$ for $h = 3$ and so forth. As you can see, the correlation will both be decreasing over time, and changing signs for odd powers, which is reflected in the corresponding acf plot.