Bruce Campell NCSU ST 534 HW 2

Probems 1.15, 1.20, 1.27, 2.3

Shumway, Robert H.; Stoffer, David S. Time Series Analysis and Its Applications: With R Examples (Springer Texts in Statistics)

26 September, 2017

1.15

Let w_t, for $t = 0, \pm 1, \pm 2, \ldots$ be a normal white noise process, and consider the series $x_t = w_t w_{t-1}$. Determine the mean and autocovariance function of x_t, and state whether it is stationary.

$$\mu_t = E[x_t] = E[w_t w_{t-1}]$$

and since $w_t \perp w_{t-1}$

$$E[w_t w_{t-1}] = E[w_t]\, E[w_{t-1}]$$

$\forall t\ E[w_t] = 0$ so we have that the mean function of x_t is $\mu_t = 0$

The autocovariance of x_t is defined by

$$\gamma(s, t) = E[(x_t - \mu_t)(x_s - \mu_s)]$$

since $\mu_t = \mu_s\ \forall s, t$ we have that

$$\gamma(s, t) = E[x_t x_s] = E[w_t\, w_{t-1}\, w_s\, w_{s-1}]$$

There are several cases we need to consider when evaluating the expectation

$$\begin{cases} s = t \\ s = t - 1 \\ s = t + 1 \\ |s - t| > 1 \end{cases}$$

When $s = t$ $\gamma(s, t) = \gamma(t, t) = E[w_t^2 w_{t-1}^2] = E[w_t^2]\, E[w_{t-1}^2]$ since $w_t \perp w_{t-1}$ Now since w_t is a mean zero white noise process we have

1

$$(\frac{w_t}{\sigma_w})^2 \sim \chi_1^2 \implies E[w_t^2] = \sigma_w^2 E[(\frac{w_t}{\sigma_w})^2] = \sigma_w^2 \times 1$$

Where we've used that $E[\chi_1^2] = 1$ in the last step. We have a similar result for $E[w_{t-1}^2]$ so $\gamma(t,t) = (\sigma_w^2)^2$

When $s = t-1$ $\gamma(s,t) = \gamma(t-1,t) = E[w_{t-1}w_{t-2}\, w_t w_{t-1}] = E[w_{t-1}^2]\, E[w_{t-2}\, w_t] = (\sigma_w^2) \times 0 = 0$ since $w_{t-2} \perp w_{t-1} \perp w_t$ A similar calculation for the case $s = t+1$ yields that $\gamma(t+1,t) = 0$

Lastly, when $|s - t| > 1$ we have no overlapping terms in the expectation $E[w_t\, w_{t-1}\, w_s\, w_{s-1}]$ so in that case we also have $\gamma(s,t) = 0$

We see that

$$\gamma(s,t) = \begin{cases} (\sigma_w^2)^2 & |s = t \\ 0 & |s \neq t \end{cases}$$

Rewriting this

$$\gamma(h) = \begin{cases} (\sigma_w^2)^2 & |h = 0 \\ 0 & |h \neq 0 \end{cases}$$

Since the mean function is zero and the autocovariance is only a function of the lag between s, t we have that x_t is second order stationary.
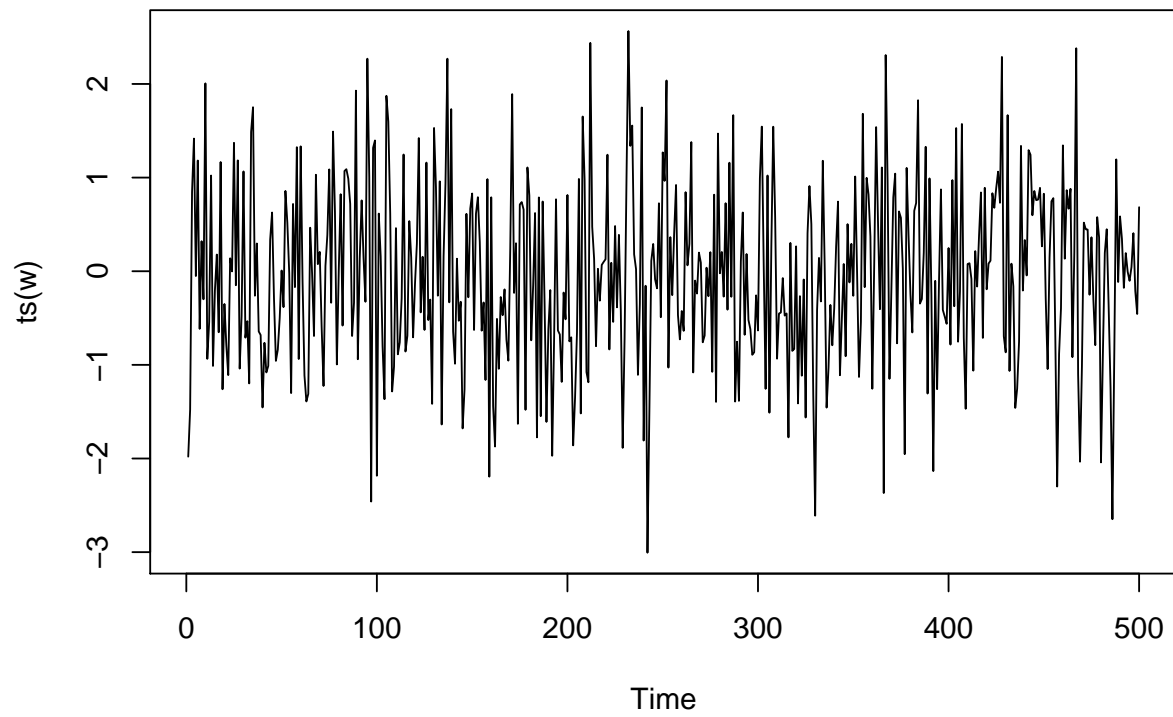
1.20

Simulate a series of $n = 500$ Gaussian white noise observations as in Example 1.8 and compute the sample ACF, $\hat{\rho}(h)$, to lag 20. Compare the sample ACF you obtain to the actual ACF, $\rho(h)$.
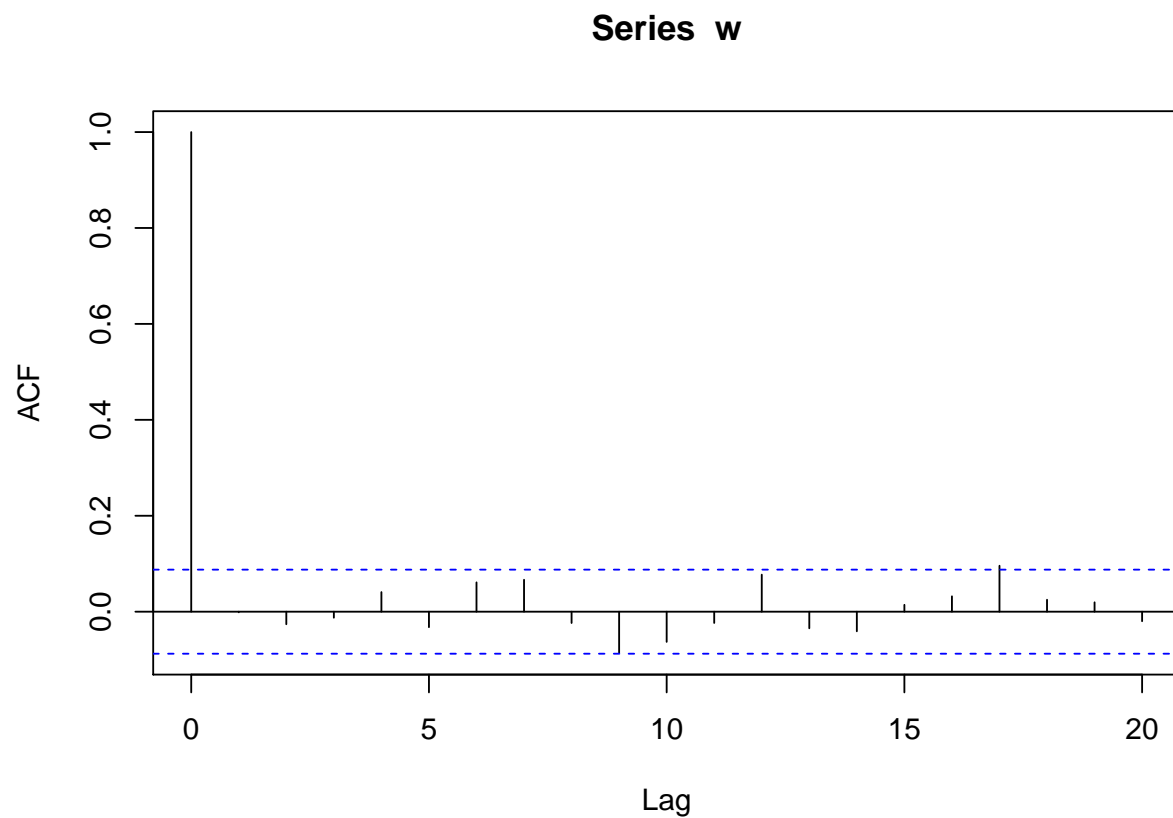
In the plot the dotted line is at $\pm\frac{z_{(0.05)}}{\sqrt{n}}$, this is the level $\alpha = 0.05$ Wald test for the hypothesis $H_0 : acf(i) = 0$.

```
w = rnorm(500, 0, 1)
plot(ts(w), main = "simulated series")
```

simulated series
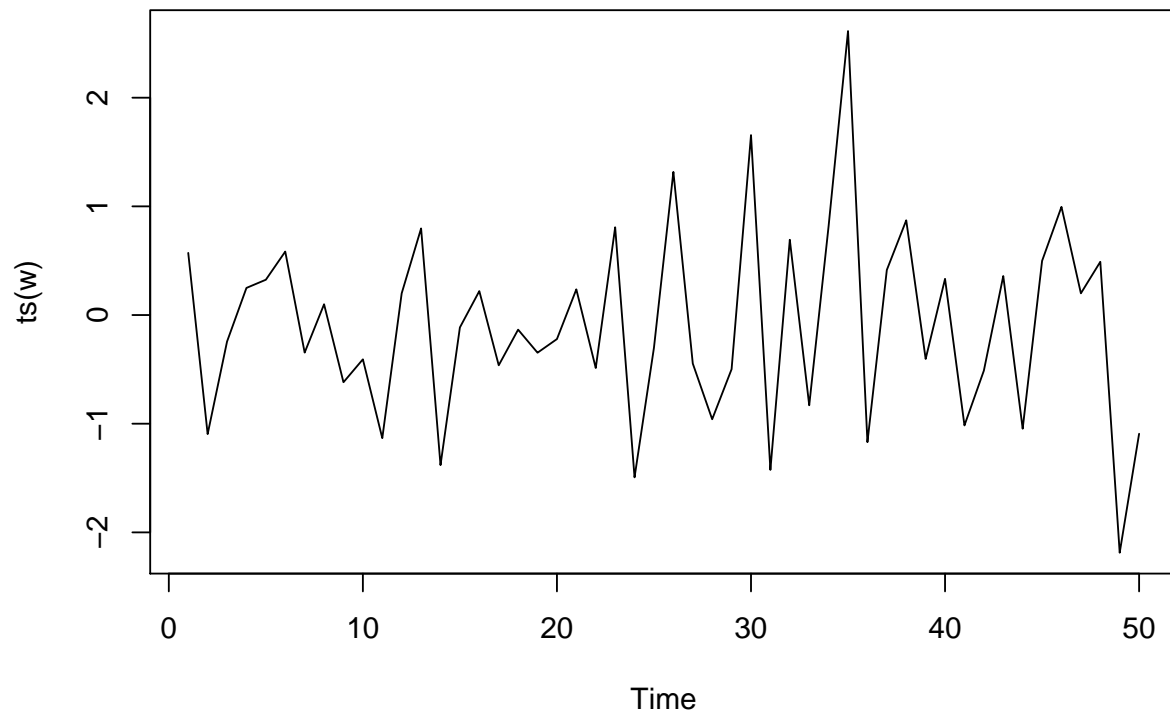


```r
acf(w, lag.max = 20)
```

Series w

The true acf is zero for all $i \neq 0$ Most of non-zero the entries in the empirical acf from the simulated data are within the bounds of $\pm \frac{z_{(0.05)}}{\sqrt{n}}$.
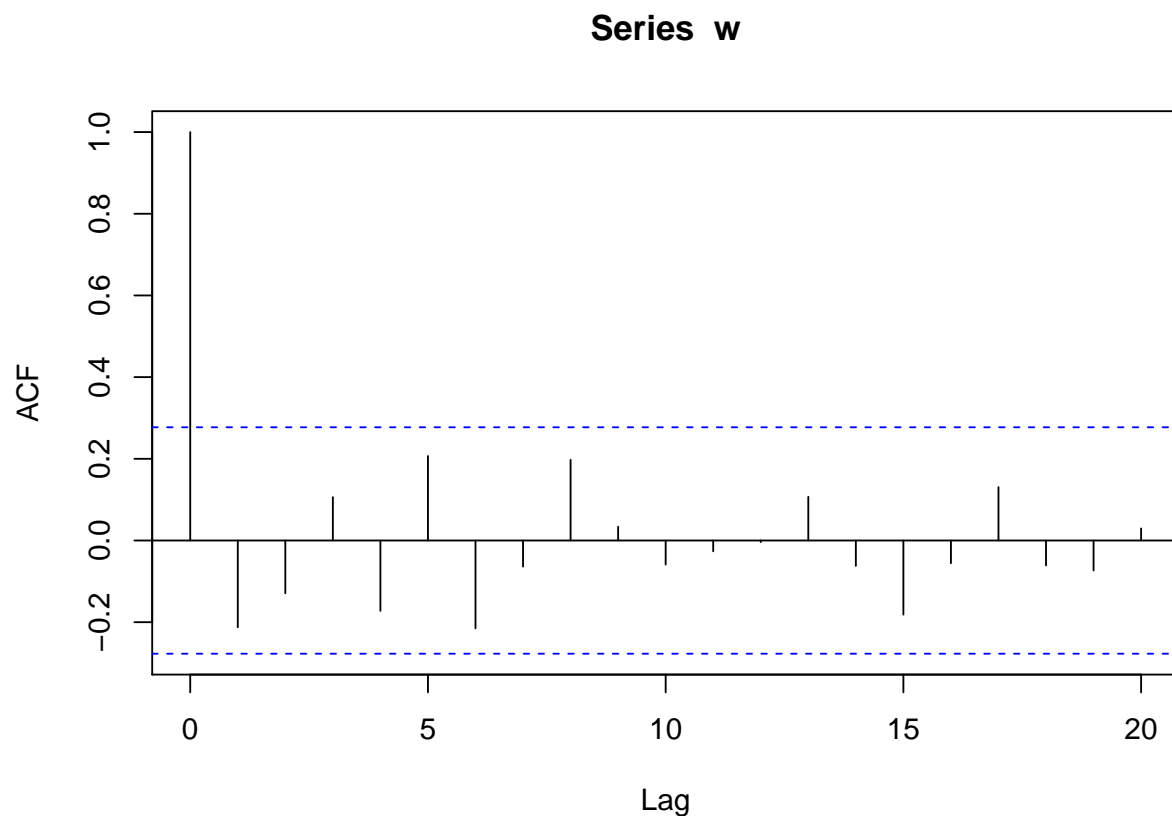
Repeat part (a) using only n = 50. How does changing n affect the results?

```r
w = rnorm(50, 0, 1)
plot(ts(w), main = "simulated series n=50")
```

simulated series n=50



```r
acf(w, lag.max = 20)
```

Series w

Again, most of non-zero the entries are within the bounds of $\pm\frac{z_{(0.05)}}{\sqrt{n}}$. Note that because n is smaller the scale of $\pm\frac{z_{(0.05)}}{\sqrt{n}}$ is larger and generally the acf values for $i \neq 0$ are increased. If we do the same experiment for 5000 samples we'll see this effect - in the other direction - more dramatically.

```r
w = rnorm(5000, 0, 1)
plot(ts(w), main = "simulated series n=5000")
```

simulated series n=5000


```r
acf(w, lag.max = 20)
```

Series w
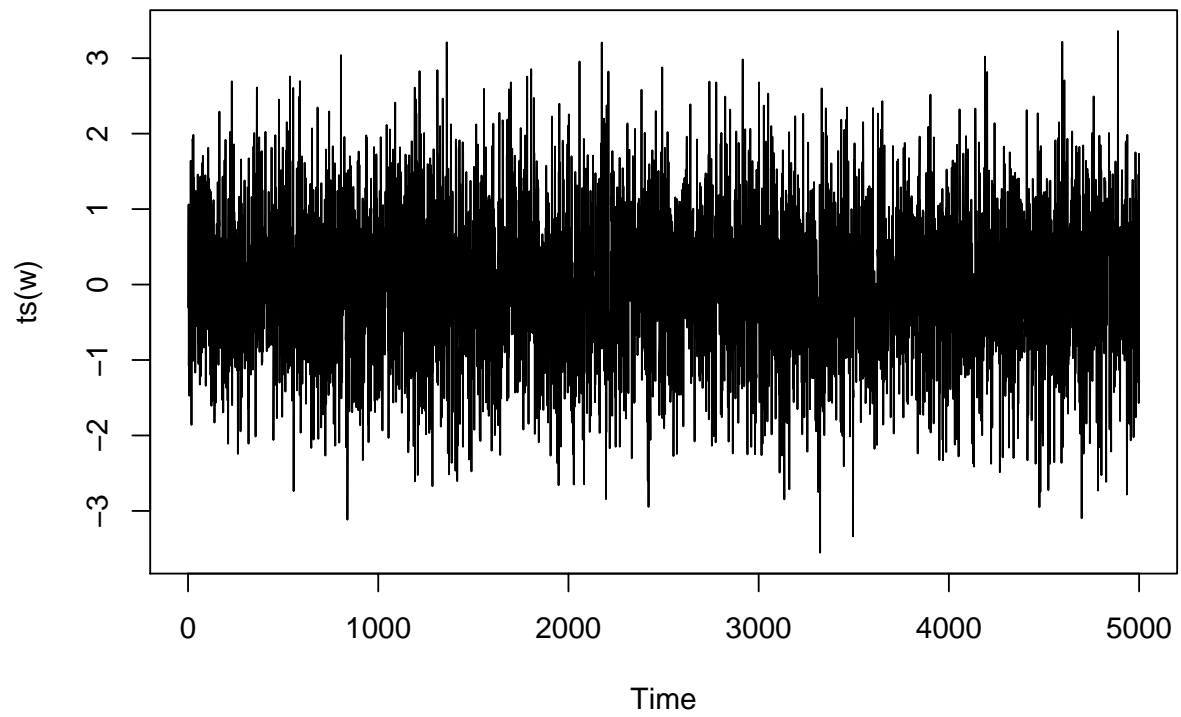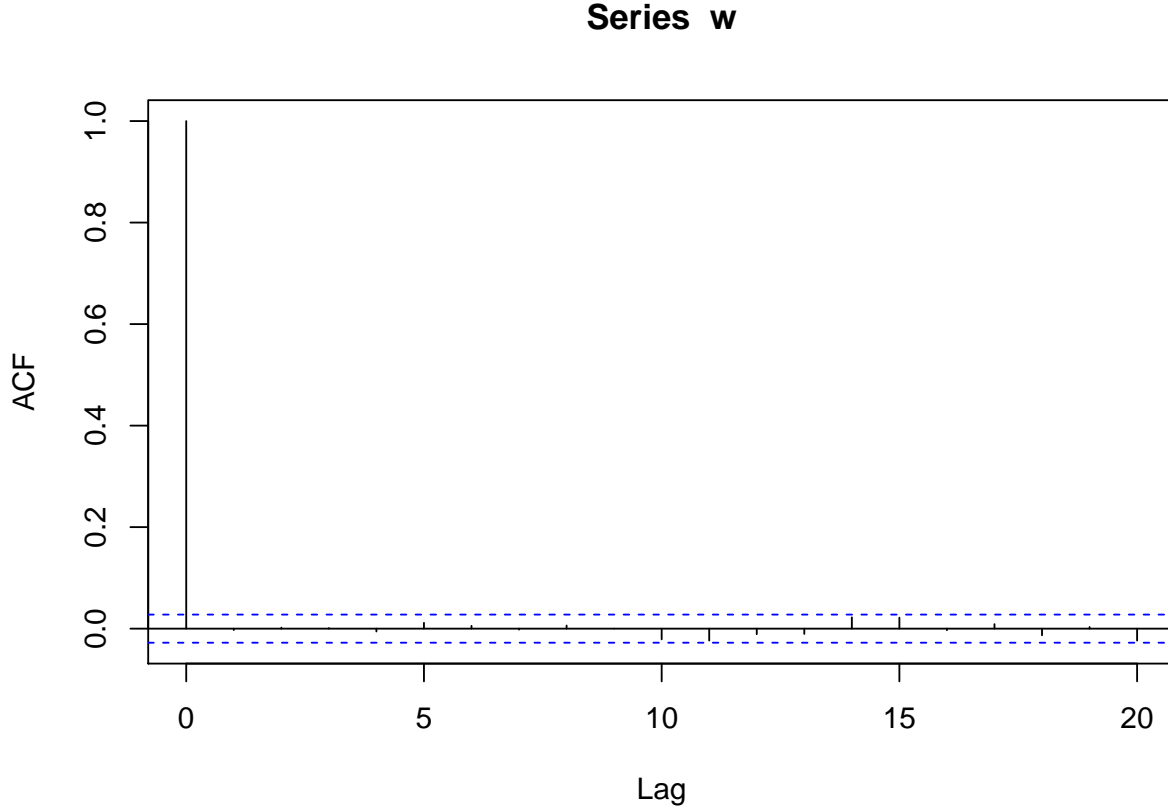
Since we have more samples, there's more evidence to provide in assessing the correlation between x_t and x_s, and the entries are closer to the true values.

1.27

A concept used in geostatistics, see Journel and Huijbregts (1978) or Cressie (1993), is that of the variogram, defined for a spatial process x_s, $s = (s1, s2)$, for $s1, s2 = 0, \pm 1, \pm 2, ...$, as $V_x(h) = \frac{1}{2} E[(x_s + h - x_s)^2]$, where $h = (h1, h2)$, for $h1, h2 = 0, \pm 1, \pm 2, ...$ Show that, for a stationary process, the variogram and autocovariance functions can be related through $V_x(h) = \gamma(0) - \gamma(h)$, where $\gamma(h)$ is the usual *lagh* covariance function and $0 = (0, 0)$. Note the easy extension to any spatial dimension.

$$2 \times V_x(h) = Var(x_{s_h} - x_s) = E[(x_{s_h} - x_s)^2] - E[(x_{s_h} - x_s)]^2$$

Now since we have second order stationarity $\mu_s = E[x_s] = \mu \;\; \forall s$ and $E[(x_{s+h} - x_s)] = 0 \;\forall s, h$

Now multiplying out the terms in the first part of the expression for the variaogram

$$2 \times V_x(h) = E[(x_{s_h} - x_s)^2] = E[x_{s+h}^2] - 2E[x_s x_{s+h}] + E[x_s^2]$$

8

Now $E[x_{s+h}^2] = E[x_s^2] = \gamma(0)$ and $E[x_s x_{s+h}] = \gamma(h)$ so $E[(x_{s_h} - x_s)^2] = 2\gamma(0) - 2\gamma(h)$ and

$$V_x(h) = \frac{1}{2}E[(x_s + h - x_s)^2] = \gamma(0) - \gamma(h)$$

When x_s is stationary.

2.3

Repeat the following exercise six times and then discuss the results. Generate a random walk with drift, (1.4), of length $n = 100$ with $\delta = .01$ and $sigma_w = 1$. Call the data x_t for $t = 1, ..., 100$. Fit the regression $x_t = \beta t + w_t$ using least squares. Plot the data, the mean function (i.e., $\mathfrak{t}_t = .01t$) and the fitted line, $\hat{x}_t = \hat{\beta}t$, on the same graph. Discuss your results.

100 Points
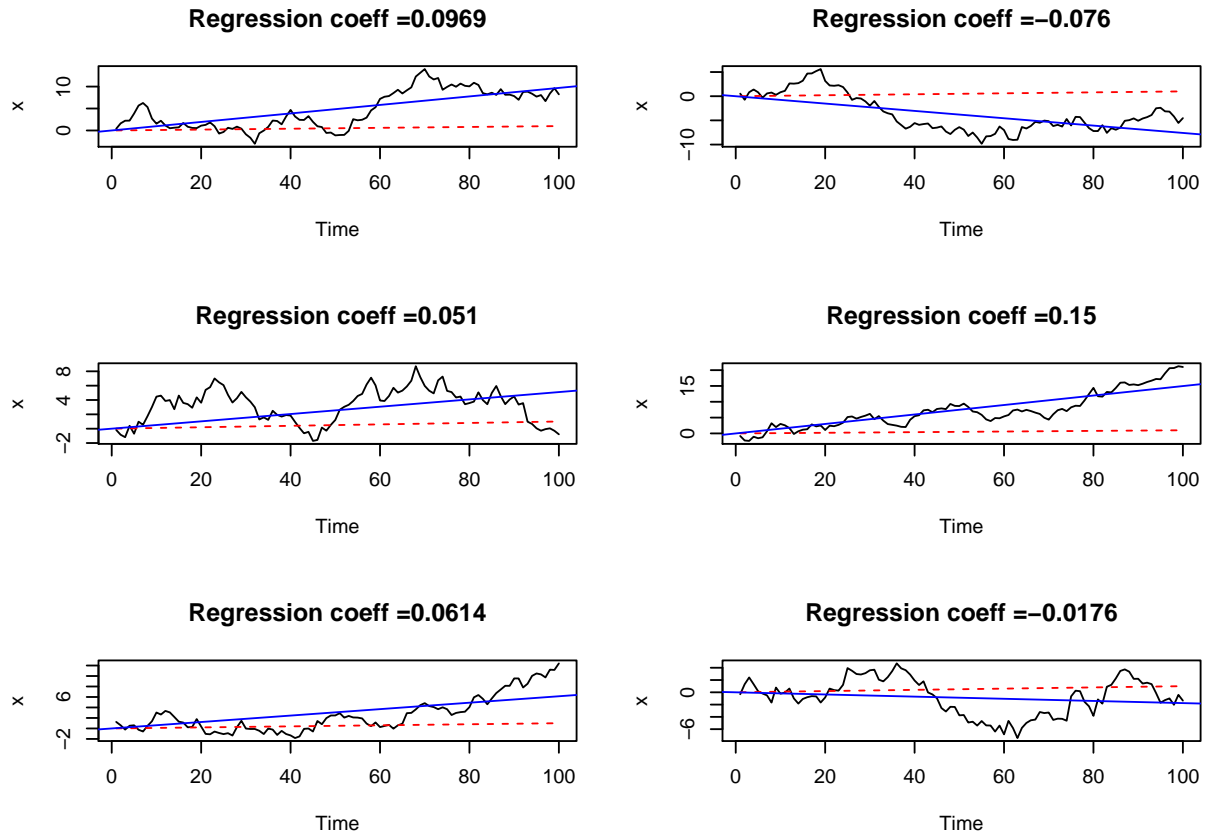
The regression line does a good job of identifying the trend in the series. We see that the regression line and the drift do not agree. The regression line does a much better job of identifying the trend in this case. If we reduce the variance of the white noise process in the random walk with drift model we would expect to see greater alignment between the drift

term and the regression. Also, if we increase the number of points in the series we would expect to achieve greater alignment in between the drift t and the regression.

Below is an example where we've reduced the variance of the white noise fro 1 to 0.5 and have increased the length of the series from 100 to 10000. We see the long term agreement is much better for some of these examples.
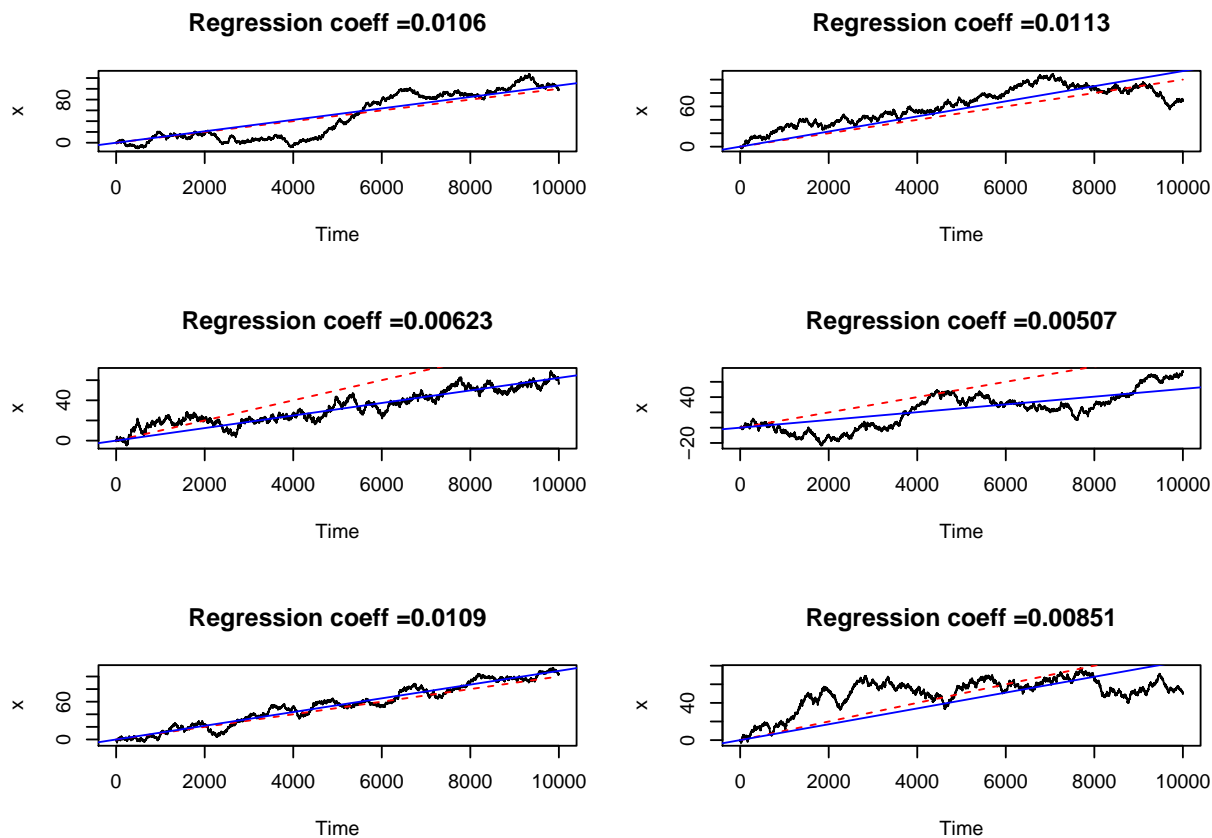
Here we run a simulation where we generate 5000 random walk with drift time series and fit the regression line. We assemble the fitted slopes and display the empirical distribution with an overly of a normal distribution with mean and variance provided by the a mean and standard deviated estimated from the sample of fitted values.
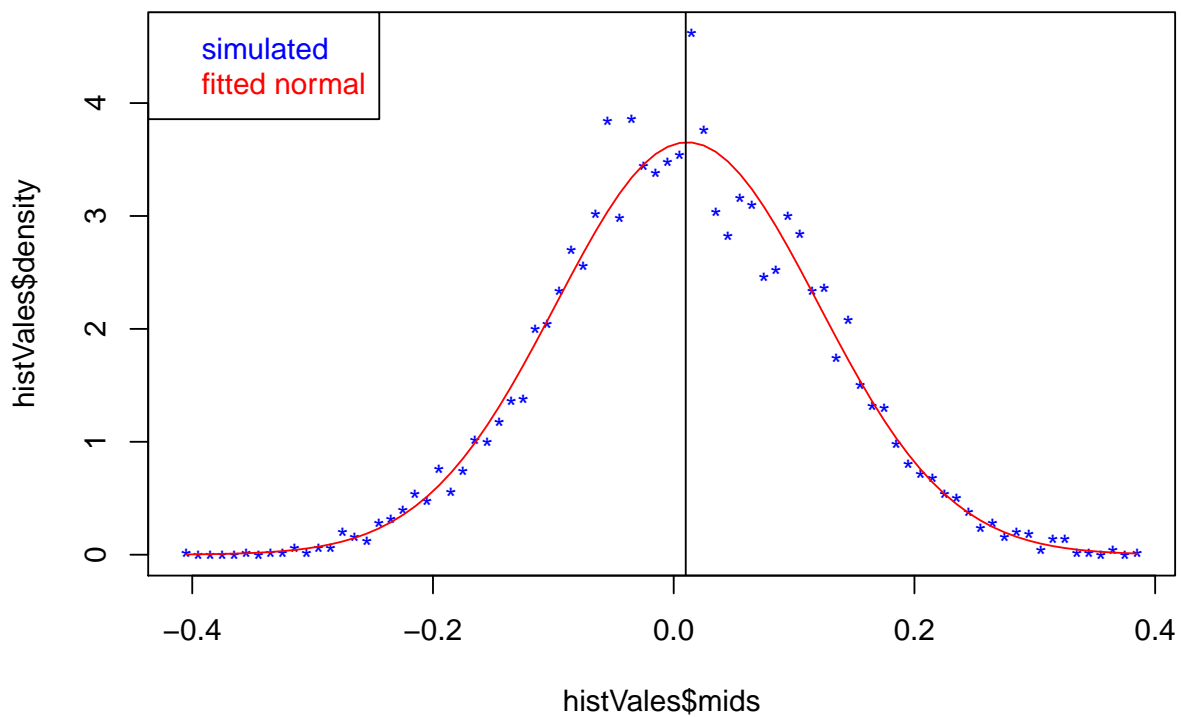
If we lower the standard deviation of the white noise we see the same distribution but with a much smaller variance about the drift.