## STAT 520, Fall 2015: Exam I

This exam was posted online Tuesday October 13, 2015 at noon. Complete the exam, scan your solutions, and email to Yawei Liang (yliang@email.sc.edu) by noon Wednesday October 14, as usual. You are to work completely independly on this exam; however you may use notes, your textbook, Google, etc.

1. Let $\{Y_t\}$ be a stationary process with constant mean $E(Y_t) = \mu$ and autocorrelation function $\rho_k$. Define $\overline{Y} = \frac{1}{n}\sum_{t=1}^{n} Y_t$ to be the sample mean of $Y_1, Y_2, \ldots, Y_n$. Recall that

$$\operatorname{var}(\overline{Y}) = \frac{\gamma_0}{n}\left[1 + 2\sum_{k=1}^{n-1}\left(1 - \frac{k}{n}\right)\rho_k\right],$$

   where $\gamma_0 = \operatorname{var}(Y_t)$. Find $\operatorname{var}(\overline{Y})$ when $\{Y_t\}$ is ARMA(1,1) with parameters $\phi$ and $\theta$. Find one pair of values $(\phi, \theta)$ that makes $\operatorname{var}(\bar{Y})$ smaller than that for white noise when $n = 3$. By what fraction is the variance reduced?

2. Let $\{e_t\}$ be mean-zero white noise, as usual, and consider the series

$$Y_t = 1.25Y_{t-1} - 0.125Y_{t-2} - 0.125Y_{t-3} + e_t + 0.25e_{t-1}.$$

   (a) Write the series as $\phi(B)(1-B)^d Y_t = \theta(B)e_t$. That is, find $\phi(B)$, $\theta(B)$, and $d$ so that the differenced series is stationary and invertible. You may use an online root finder for help if you want. Hint: as shown in class, root finders will rewrite the equation $1 + c_1B + c_2B^2 + c_3B^3 = 0$ as $a(B - r_1)(B - r_2)(B - r_3) = 0$. You need to divide both sides of this latter equation by $ar_1r_2r_3$ and multiply by $(-1)^3 = -1$ to get it in the form $(1 - B/r_1)(1 - B/r_2)(1 - B/r_3) = 0$.

   (b) Obtain the population autocorrelation function (ACF) for the appropriately differenced series $W_t = \nabla^d Y_t$. You can use `ARMAacf` in R and report a plot.

   (c) Simulate $n = 200$ from the series $\{Y_t\}$ assuming $e_t \overset{iid}{\sim} N(0, \sigma_e^2)$ where $\sigma_e = 1$ and plot it along with the sample ACF. Do the same for the differenced series $\{\nabla Y_t\}$. R code to simulate from an ARIMA(1,1,1) is in Chapter 5; you could modfiy this.

3. Let $\{Y_t\}$ be an AR(1) process with $|\phi| < 1$. Find the autocorrelation function for $W_t = \nabla Y_t = Y_t - Y_{t-1}$ in terms of $\phi$ and $\sigma_e^2$. What does this simplify to at lag $k = 0$, i.e. what is $\operatorname{var}(W_t)$?

4. An underground temperature probe was placed about a half kilometer away from geothermal borehole in Iceland. Degrees Celsius $Y_t$ were recorded daily for 200 days in 2014. This will read the data into R:

```
library(TSA)
y=ts(read.table("http://people.stat.sc.edu/hansont/stat520/thermal.txt"),start=1)
```

   (a) Plot the temperature $Y_t$ versus time $t$ and accompanying sample ACF. Does the data appear to be stationary? Are there any pronounced trends? Describe.

   (b) Fit a straight-line regression model to the data for detrending purposes and plot the residuals $r_t = Y_t - \hat{\beta}_0 - \hat{\beta}_1 t$ versus time $t$. Does a line appear to fit okay?

   (c) Fit a quadratic regression model to the data for detrending purposes and plot the residuals $r_t = Y_t - \hat{\beta}_0 - \hat{\beta}_1 t - \hat{\beta}_2 t^2$ versus time $t$, along with the sample ACF of the residuals. Do the residuals appear to be stationary?

   (d) Now plot the first difference of the residuals $\nabla r_t = r_t - r_{t-1}$ from part (c) versus time $t$ along with the sample ACF. Comment on whether stationarity is finally achieved.

   (e) Based on the ACF from part (d), what might a plausible model for $w_t = \nabla r_t$ be? Why?