STAT 720 sp 2019 hw 1

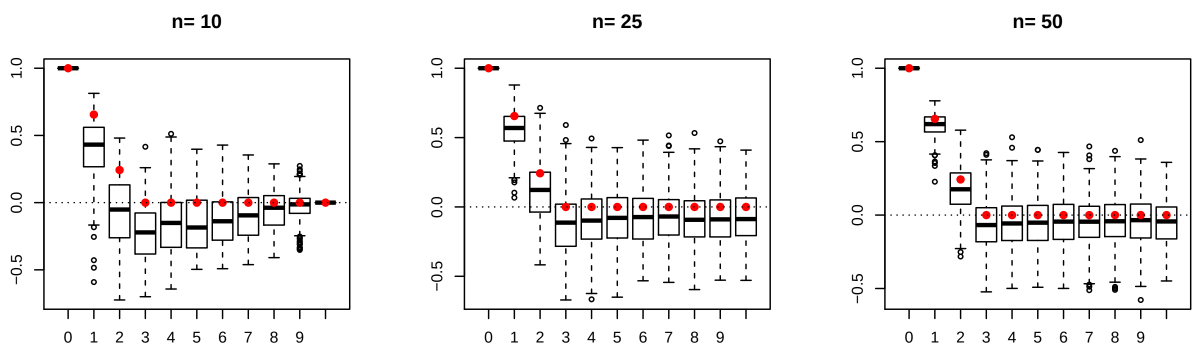
due: Wednesday, Feb 6th, 2019

1. Do problems 1.1, and 1.15 from B&D Intro.

2. Let $Y_t = Z_t + \theta_1 Z_{t-1} + \theta_2 Z_{t-2}$, for $t \in \mathbb{Z}$, where $\{Z_t, t \in \mathbb{Z}\}$ is $\mathrm{WN}(0, \sigma^2)$.

   (a) Find the autocorrelation function $\rho(\cdot)$ of $\{Y_t, t \in \mathbb{Z}\}$.

   (b) Set $\theta_1 = 0.9$ and $\theta_2 = 0.5$ and let $\{Z_t, t \in \mathbb{Z}\}$ be independent $\mathrm{Normal}(0, 1)$ random variables. Then, for the sample sizes $n = 10$, $n = 25$, and $n = 50$, generate 500 realizations of $Y_1, \dots, Y_n$. On each realization, compute the sample autocorrelation function at lags $h = 0, 1, \dots, 10$. Then make boxplots of the estimated values of the autocorrelations from the 500 simulated time series with the true values of the autocorrelations overlaid. The resulting plots should look like this:



   Turn in your plots and your R code. In addition, comment on the performance of the sample autocorrelation function as an estimator of the true autocorrelation function as the sample size increases.

3. Let $Y_t = m_t + \varepsilon_t$ for $t \in \mathbb{Z}$, where $m_t = c_0 + c_1 t + c_2 t^2$ and $\{\varepsilon_t, t \in \mathbb{Z}\}$ is a stationary time series with mean zero.

   (a) Show that the series $\nabla^2 Y_t, t \in \mathbb{Z}$ is stationary.

   (b) Find choices of $a_2 = a_{-2}$, $a_1 = a_{-1}$, and a_0 such that the filter

$$\hat{m}_t = \sum_{j=-2}^{2} a_j Y_{t-j} \tag{1}$$

   is an unbiased estimator of m_t for all t.

   (c) Following part (b), set $a_0 = 18/12$ and a_1 and a_2 accordingly. Then apply the smoothing in (1) to the R data set `uspop`, using $Y_{-1} = Y_0 = Y_1$ and $Y_{n+1} = Y_{n+2} = Y_n$ to take care of estimation at the start and end of the series. Plot the time series with the estimated $\hat{m}_1, \dots, \hat{m}_n$ overlaid and make a plot of the residuals $Y_t - \hat{m}_t$ versus t for all $t = 1, \dots, n$. Supply the R code and the plots.

(d) Since the trend in the `uspop` data appears to be quadratic, and since quadratic trends can be eliminated by differencing the series with the operator ∇^2, apply the differencing ∇^2 to the `uspop` data and plot the resulting series. Comment on whether you think the trend has been eliminated by the differencing.

4. Show that for any random variables $X_1, \dots, X_m$ and real numbers $a_1, \dots, a_m$, we have

$$\mathrm{Var}(\sum_{i=1}^{m} a_i X_i) = \sum_{i=1}^{m}\sum_{j=1}^{m} a_i a_j \,\mathrm{Cov}(X_i, X_j).$$

5. Let $\{X_t, t \in \mathbb{Z}\}$ be a stationary time series with autocovariance function $\gamma(\cdot)$ and show that

$$\mathrm{Var}(\sqrt{n}\bar{X}_n) = \sum_{h=-(n-1)}^{n-1} \left(1 - \frac{|h|}{n}\right)\gamma(h),$$

where $\bar{X}_n = (X_1 + \cdots + X_n)/n$.

6. Consider the $\mathrm{MA}(q)$ process defined by

$$X_t = \theta_0 Z_t + \theta_1 Z_{t-1} + \cdots + \theta_q Z_{t-q}, \quad \text{for} \quad t \in \mathbb{Z},$$

where $\{Z_t, t \in \mathbb{Z}\}$ is $\mathrm{WN}(0, \sigma^2)$.

(a) Show that

$$\mathrm{Cov}(X_t, X_{t+h}) = \begin{cases} \sigma^2 \sum_{j=0}^{q-|h|} \theta_j \theta_{j+|h|} & \text{for } |h| \le q \\ 0 & \text{for } |h| > q \end{cases}$$

for all $t \in \mathbb{Z}$.

(b) Show that

$$\lim_{n\to\infty} \mathrm{Var}(\sqrt{n}\bar{X}_n) = \sigma^2 \left(\sum_{j=0}^{q} \theta_j\right)^2.$$

(c) Set $(\theta_0, \theta_1, \theta_2, \theta_3) = (1, 0.8, 0.5, -0.2)$ and let $\{Z_t, t \in \mathbb{Z}\}$ be independent $\mathrm{Normal}(0, 1)$ random variables.

i. Give the asymptotic variance of the quantity $\sqrt{n}\bar{X}_n$ as $n \to \infty$ based on this time series.

ii. Use R to generate a realization of length $n = 100$ of the time series; produce a plot of your time series using `plot(X,type="l")`. *Hint: You will need to generate $n + q = 103$ values from the* $\mathrm{Normal}(0, 1)$ *distribution and then use these to construct the realizations* $X_1, \dots, X_n$.

iii. Under the same settings, generate 1000 realizations of length $n = 100$ of the times series and compute the sample means of each of the 1000 realizations. Then compute the variance of $\sqrt{n}\bar{X}_n$ over the 1000 realizations. Compare this value to the theoretical asymptotic variance of $\sqrt{n}\bar{X}_n$.