Statistics 519, Winter Quarter 2020

Problem Set 3

Problem 7 (4 points for part a and 5 points for part b). Let $\{Y_t\}$ be a stationary process with ACVF $\gamma_Y(\cdot)$, and let k be any positive integer.

a. Show that $\{\nabla^k Y_t\}$ is a stationary process (Appendix D at the end of Chapter 5 of Cryer & Chan defines and discusses the unit-lag difference operator ∇; see also overheads III–77 and III–78). Note: you need *not* determine an explicit expression for the ACVF for $\{\nabla^k Y_t\}$.

b. Show that, if $X_t = m_t + Y_t$, where $m_t = \sum_{j=0}^{k} c_j t^j$ is a kth order polynomial (which implies that $c_k \neq 0$), then $\{\nabla^k X_t\}$ is a stationary process, whereas $\{\nabla^{k-1} X_t\}$ is not.

Problem 8 (8 points). Perform an analysis of the atmospheric carbon dioxide (CO_2) time series from Mauna Loa, Hawaii, along the same lines as the analysis of the accidental deaths (AD) series starting with lecture overhead III–82 (the 3rd set of `R` code on the course Web site has the code used to analyze the AD series). You can read the CO_2 data directly into `R` using

```
co2 <- read.table("http://faculty.washington.edu/dbp/s519/Data/co2-1958-2019.txt")
```

and then access the data values and the years associated with them via, respectively, `co2$V2` and `co2$V1`. Alteratively, you can access the data either via a link on item 9 in the list on the Data page of the course Web site or by going directly to

`http://faculty.washington.edu/dbp/s519/Data/co2-1958-2019.txt`

Please feel free to alter choices that were made in the analysis of the AD series if you deem them to be inappropriate for your analysis of the CO_2 series. Create and turn in plots that correspond to overheads III–85, III–87, III–89 (but just show the seasonal pattern for a single year, as is done in III-107), III–92, III–94, III–96, III–97 and III–98, along with brief descriptions of the steps you took in your analysis (please also turn in the code you used to do your analysis). Finally, state briefly your conclusions about how well the simple modeling approach worked for the CO_2 series.

Problem 9 (8 points). Here we consider a time series $\{x_t\}$ measuring ambient noise in the ocean from one second to the next. You can read this series directly into `R` using

```
on.ts <- scan("http://faculty.washington.edu/dbp/s519/Data/ocean-noise.txt")
```

Alteratively, you can access the data either via a link on item 15 in the list on the Data page of the course Web site or by going directly to

`http://faculty.washington.edu/dbp/s519/Data/ocean-noise.txt`

Create plots of (a) the time series $\{x_t\}$, (b) its unit-lag scatter plot (i.e., x_{t+1} versus x_t) and (c) its sample ACF out to lag 20, along with lines showing 95% confidence intervals for the ACF under the null hypothesis that $\{x_t\}$ is a realization of an IID noise process. Examine the null hypothesis by subjecting $\{x_t\}$ to the portmanteau, turning point, difference-sign, rank and runs tests. State the results of each test and your overall conclusion about the viability of the IID noise hypothesis. How do these formal tests compare with the informal test given in your plot of the sample ACF? (The 4th set of `R` code on the course Web site has code used to compute the five tests for the examples considered in the course overheads.)

Solutions are due Friday, January 31, at the beginning of the class.