STAT 4440 Final Exam

Deadline: May 1, 11:59 pm

This is an open book/notes exam. You need to show all your work for credit. You may not discuss, copy or share this exam with anybody else, not even your best friends, your family, or other people that may have nothing to do with your exam.

1. (10 points) Stationarity.

   (a) Let $\{X_t\}$ be a stationary series with autocovariance function γ_k for $k > 0$. Given the constant $c_1, c_2, \ldots, c_s$, decide whether $Y_t = c_1 X_t + c_2 X_{t-1} + \cdots + c_s X_{t-s+1}$ is stationary or not.

   (b) Let $\{X_t\}$ be a stationary ARMA(p, q) process with ψ-weights $\psi_1^{(X)}, \psi_2^{(X)}, \psi_3^{(X)}, \ldots$. Find the ψ-weights for process $Y_t = c_1 X_t + c_2 X_{t-1} + \cdots + c_s X_{t-s+1}$.

2. (10 points) Model specification. (Note that ARMA(p, q) and seasonal ARMA$(p, q) \times (P, Q)_s$ models must be both stationary and invertible.)

   (a) Specify p, d, q for the ARIMA model and find the values of the parameters (ϕ's and θ's in the ARMA(p, q) model):

   $$Y_t = 2.8Y_{t-1} - 2.6Y_{t-2} + 0.8Y_{t-3} + e_t - 0.76e_{t-1} - 0.2e_{t-2}$$

   (b) Specify p, d, q, P, D, Q and the period s for the multiplicative seasonal ARIMA model and find the values of the parameters (ϕ's, θ's, Φ's, and Θ's in the seasonal ARMA$(p, q) \times (P, Q)_s$ model):

   $$Y_t = -0.4Y_{t-1} + Y_{t-4} + 0.4Y_{t-5} + e_t - 0.6e_{t-1} - 0.5e_{t-4} + 0.3e_{t-5}$$

3. (30 points) Consider the model in the midterm exam:

   $$Y_t - \mu = 0.9(Y_{t-1} - \mu) - 0.5(Y_{t-2} - \mu) + e_t - 0.6e_{t-1}$$

   (a) Find the partial autocorrelation function, and evaluate ϕ_{11}, ϕ_{22}, and ϕ_{33}. Show all working.

   (b) Suppose that we have 5 observations generated by this process: $y_1 = 21, y_2 = 19, y_3 = 18, y_4 = 20, y_5 = 21$. Calculate method of moments estimates of $\mu, \gamma_0, \rho_1, \rho_2, \rho_3$ manually.

   (c) Suppose that we also observed the corresponding residuals $e_1 = 1, e_2 = -1, e_3 = 0, e_4 = 2, e_5 = 1$. Compute the forecasts for the next 5 observations $\hat{Y}_5(1), \hat{Y}_5(2), \hat{Y}_5(3), \hat{Y}_5(4), \hat{Y}_5(5)$.

(d) Assume $\sigma_e^2 = 1$. Calculate the variance of each of the forecasts made in (3c).

(e) Calculate 80% prediction limits for each of your forecasts in (3c).

(f) Suppose we obtained two new observations $y_6 = 19, y_7 = 18$. Compute the update of the forecasts $\hat{Y}_7(1), \hat{Y}_7(2), \hat{Y}_7(3)$.

4. Time Series data analysis.

The data "Nebraska_LaborForce_UnemploymentRate.csv" provides the labor force population and unemployment rate of Nebraska from January 1976 to March 2019. It is collected from the U.S. Bureau of Labor Statistics website (https://www.bls.gov/bls/unemployment.htm). You can load the data into R using the following command:

```
dat = read.csv("Nebraska_LaborForce_UnemploymentRate.csv")
labor = ts(dat$LaborForce,start=1976,frequency=12)
unemployment = ts(dat$UnemploymentRate,start=1976,frequency=12)
```

(a) (25 points) Fit an appropriate model to `labor`. Forecast the next 3 years with 95% prediction limits.

(b) (25 points) Fit an appropriate model to `unemployment`. Forecast the next 3 years with 95% prediction limits.

Notes:

(a) If you fit a model and later decide that the model is unsuitable, do not remove the model from your answer. Include it, discuss why it was unsuitable, and then move on to another model.

(b) You should consider the following stages: data cleaning, transformation, deterministic trend model, stochastic process model, seasonal model, model specification, parameter estimation, model diagnostics, forecasting.

(c) Data analysis is not only made by code. You need to provide explicit explanation for each step, and there should be a conclusion. Lack of explanation will result in loss of points.