Name: ____________________

Student ID number: ____________________

# PRACTICE FINAL EXAM - FALL 2019

STAT 153 - Introduction to Time Series

December 16, 2019

## General comments:

1. Solutions in Red
2. Explanations in Blue

Name: ____________________

1. (a) Define (weak) stationarity.

   A process $\{X_t\}$ is weakly stationary if:

   - $E(X_t)$ is constant and does not depend on time $t$
   - The autocovariance $cov(X_t, X_s) = \gamma_X(h)$ only depends on $s, t$ through their difference $h = |s - t|$

   (b) Do all white noise processes follow weak stationarity? If yes, explain why. If no, provide a counter example.

   Yes. White noise is defined as having mean zero, constant variance, and $cov(X_s, X_t) = 0$ $\forall s \neq t$. It is weakly stationary as mean zero means the expectation is constant, and the ACVF does not depend on anything but $h$: $\gamma_X(h) = \begin{cases} \sigma^2 & h = 0 \\ 0 & h \neq 0 \end{cases}$.

   (c) Are all stationary processes considered white noise? If yes, explain why. If no, provide a counter example.

   No. If $Z_t$ is white noise, define $Y_t = Z_t + 4$. $Y_t$ has the same ACVF as Z, but is no longer mean zero, so it is not white noise.

   As for other counter examples, any ARMA(p,q) is stationary if $\phi(B)$ has no roots equal to 1 in magnitude. ARMA(p,q) is not white noise unless $p = q = 0$.

Name: ____________________

2. Consider the time series $X_t = a + bt + ct^2 + cos(2\pi ft) + Z_t$, where $a, b, c \in R$, $f = 1/3$, and $\{Z_t\}$ is white noise.

   (a) $\{X_t\}$ is not stationary. Give an example of positive integers $j, k$ that make $\{Y_t\}$ stationary, where $Y_t = \nabla_j \nabla_k X_t$.

   SOLUTION: we choose one of the lags $j, k$ to be 3 (or any multiple of 3) and the other can be any positive integer. For example: j=1, k=3; or j=k=3; or j=33, k=17; etc.

   EXPLANATION: First,

$$\begin{aligned}
\nabla_k X_t &= X_t - X_{t-k} \\
&= a + bt + ct^2 + cos(2\pi ft) + Z_t - (a + b(t-k) + c(t-k)^2 + cos(2\pi f(t-k)) + Z_t) \\
&= cos(2\pi ft) + Z_t + bk + 2ckt - ck^2 - cos(2\pi f(t-k)) - Z_{t-k} \\
Y_t &= \nabla_j \nabla_k X_t \\
&= \nabla_j [cos(2\pi ft) + Z_t + bk + 2ckt - ck^2 - cos(2\pi f(t-k)) - Z_{t-k}] \\
&= [cos(2\pi ft) + Z_t + bk + 2ckt - ck^2 - cos(2\pi f(t-k)) - Z_{t-k}] - \\
&\quad [cos(2\pi f(t-j)) + Z_{t-j} + bk + 2ck(t-j) - ck^2 - cos(2\pi f(t-j-k)) - Z_{t-j-k}] \\
&= 2cjk + cos(2\pi ft) - cos(2\pi f(t-k)) - cos(2\pi f(t-j)) + cos(2\pi f(t-j-k)) \\
&\quad + Z_t - Z_{t-k} - Z_{t-j} + Z_{t-j-k}.
\end{aligned}$$

   We need $0 = cos(2\pi ft) - cos(2\pi f(t-k)) - cos(2\pi f(t-j)) + cos(2\pi f(t-j-k))$, so choose $k = 1/f = 3$ (or any multiple of 3) as $cos(x) = cos(x + 2\pi\ell)$ for any integer $\ell$:

$$\begin{aligned}
\nabla_k X_t Y_t &= 2cjk + cos(2\pi ft) - cos(2\pi f(t-1/f)) - cos(2\pi f(t-j)) + cos(2\pi f(t-j-1/f)) \\
&\quad + Z_t - Z_{t-k} - Z_{t-j} + Z_{t-j-k} \\
&= 2cjk + cos(2\pi ft) - cos(2\pi ft - 2\pi) - cos(2\pi f(t-j)) + cos(2\pi f(t-j) - 2\pi)) \\
&\quad + Z_t - Z_{t-1/f} - Z_{t-j} + Z_{t-j-1/f} \\
&= 2cjk + cos(2\pi ft) - cos(2\pi ft) - cos(2\pi f(t-j)) + cos(2\pi f(t-j)) \\
&\quad + Z_t - Z_{t-1/f} - Z_{t-j} + Z_{t-j-1/f} \\
&= 2cjk + Z_t - Z_{t-1/f} - Z_{t-j} + Z_{t-j-1/f}
\end{aligned}$$

   Thus $Y_t$ is a constant plus white noise terms, and is thus stationary. So, we choose one of the lags $j, k$ to be 3 (or any multiple of 3) and the other can be any positive integer. For example: j=1, k=3; or j=k=3, etc. (You need not show all this work. If $j = k = 3$ is obvious to you, then state that. I give a full derivation here because it's the formal solution key ;)

   (b) What is the best predictor of $X_n$ given $X_1, ..., X_{n-1}$?

   The distribution of $X_t$ does not involve other X's, so $E(X_n|X_1, ..., X_{n-1}) = E(X_n) = E(a + bn + cn^2 + cos(2\pi fn) + Z_n) = a + bn + cn^2 + cos(2\pi fn)$.

Name: ____________________

3. Consider the equation $X_t - X_{t-4} + 0.5X_{t-8} = Z_t$ where $\{Z_t\}$ is white noise.

   (a) Let $Y_t = X_{4t}$ (such that $Y_0 = X_0$, $Y_1 = X_4$, $Y_2 = X_8$, .... ). Show that $\{Y_t\}$ is an AR(2) process.

   SOLUTION: replace $t$ with $4t$ in the original equation:

   $$X_{4t} - X_{4t-4} + 0.5X_{4t-8} = Z_4t$$

   now the time indices for $Y_t$ are 1/4 that of $X_t$'s (i.e $Y_t = X_{4t}$):

   $$Y_t - Y_{t-1} + 0.5Y_{t-2} = Z_4t.$$

   $\{Z_{4t}\}$ is clearly a white noise process, thus the difference equation implies $\{Y_t\}$ is an AR(2) process.

   ———

   SOLUTION via the HINT: We are also told that $Y_t = X_{4t}$, such that $Y_1 = X_4$, $Y_2 = X_8$, etc. So for $t = 1, 2, ...$:

   $$\begin{aligned} Y_1 &= X_4 = X_0 - 0.5X_{-4} + Z_4 \\ Y_2 &= X_8 = X_4 - 0.5X_0 + Z_8 \\ Y_3 &= X_{12} = X_8 - 0.5X_4 + Z_{12} \end{aligned}$$

   combining those three lines:

   $$Y_3 = X_{12} = Y_2 - 0.5Y_1 + Z_{12}.$$

   You can see this holds for all $t$. Hence we get the same difference equation as in the first explanation above.

   (b) What are the roots of the AR polynomial of $Y_t$? Is $Y_t$ causal? Does it have a unique stationary solution? (If you were not able to complete part a, assume $\phi_1 = 1$ and $\phi_2 = -0.5$.)

The AR polynomial is $\phi(B) = 1 - B + 0.5B^2$, so the roots are

$$\begin{aligned} \frac{-(-1) \pm \sqrt{(-1)^2 - 4(0.5)(1)}}{2(0.5)} &= \frac{1 \pm \sqrt{1-2}}{1} \\ &= 1 \pm \sqrt{-1} \\ &= 1 \pm i \end{aligned}$$

The magnitude (absolute value) of the roots are $\sqrt{1^2 + (\pm 1)^2} = \sqrt{1+1} = \sqrt{2}$. As $\sqrt{2} > 1$, this AR(2) has a unique, causal, stationary solution.

Name: ____________________

4. Show how the Yule Walker equations can estimate the parameters of an AR(1) model $(\phi, \sigma^2)$ with the sample ACVF values $\hat{\gamma}(0) = 3, \hat{\gamma}(1) = 2$.

   As $q = 0$, we have two equations

$$\gamma(0) - \phi\gamma(-1) = \sigma^2$$
$$\gamma(1) - \phi\gamma(0) = 0$$

   and by plugging in the sample ACVF values (adding all the hats), we can estimate the parameters. We start by finding $\hat{\phi}$ with the second equation:

$$\begin{aligned} 0 &= \hat{\gamma}(1) - \hat{\phi}\hat{\gamma}(0) \\ \hat{\phi}\hat{\gamma}(0) &= \hat{\gamma}(1) \\ \hat{\phi} &= \frac{\hat{\gamma}(1)}{\hat{\gamma}(0)} \\ \hat{\phi} &= \frac{2}{3} \end{aligned}$$

   and then finish by finding $\hat{\sigma}^2$ with the first equation:

$$\begin{aligned} \hat{\sigma}^2 &= \hat{\gamma}(0) - \hat{\phi}\hat{\gamma}(-1) \\ \hat{\sigma}^2 &= \hat{\gamma}(0) - \hat{\phi}\hat{\gamma}(1) \\ \hat{\sigma}^2 &= 3 - \frac{2}{3}(2) \\ \hat{\sigma}^2 &= \frac{9}{3} - \frac{4}{3} \\ \hat{\sigma}^2 &= \frac{5}{3} \end{aligned}$$

   Thus $\hat{\phi} = \frac{2}{3}$ and $\hat{\sigma}^2 = \frac{5}{3}$.

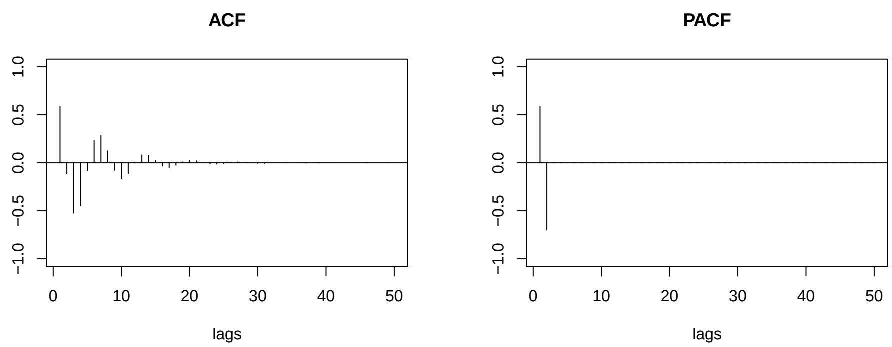
Name: ____________________

5. Below are the theoretical ACF plots and PACF plots for three different models. Note that there are no bars for lag 0 in either the ACF or PACF plots. Also note that these are simple MSARMA models

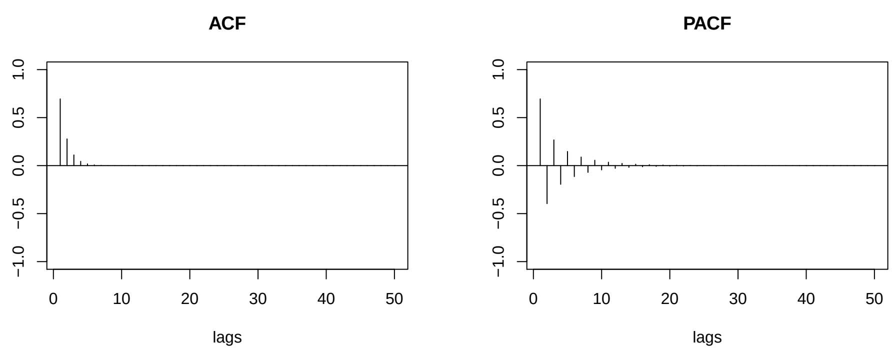
$$ARMA(p,q)x(P,Q)_S$$

such that $p + q + P + Q \leq \mathbf{2}$, so possibilities include AR(1), AR(2), ARMA(1,1), ARMA(1,0)x(0,1)$_{12}$, ARMA(0,0)x(1,1)$_{20}$, etc. Above each pair of plots in the blanks provided, write in the appropriate values of p, q, P, Q, and S, where S is the period length for the seasonal ARMA component. Let $S = 0$ if no seasonal component is present.

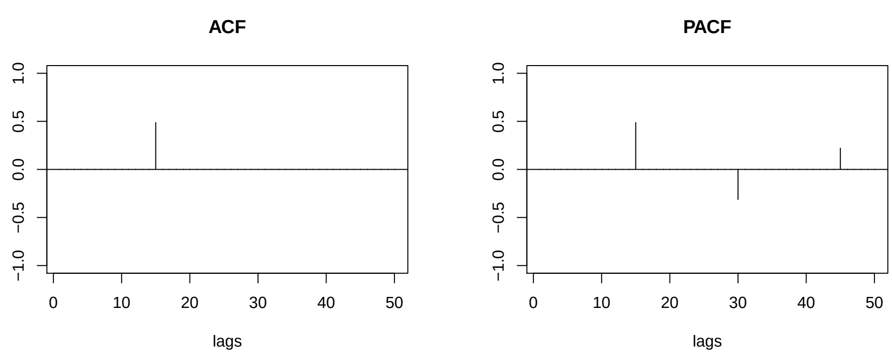
(a) ARMA( _____ , _____ )x( _____ , _____ )_____ AR(2) = ARMA(2,0)x(0,0)$_0$




(b) ARMA( _____ , _____ )x( _____ , _____ )_____ ARMA(1,1)x(0,0)$_0$

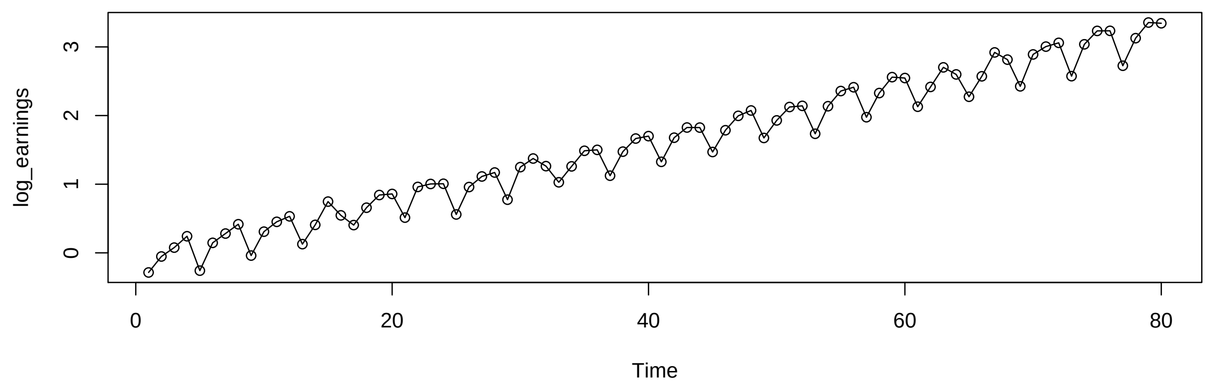



(c) ARMA( _____ , _____ )x( _____ , _____ )_____ MA(1)$_{15}$ = ARMA(0,0)x(0,1)$_{15}$



Name: ____________________

6. These are the quarterly earnings for Mom and Pop's Diner. We have already taken the log of the original data to stabilize the variance.




(a) How could you model the trend and seasonality of this time series, without using differencing? Write the suggested equation.

The trend appears linear, and there is some seasonality effect every 4 periods (annual). Thus a potential model could be

$$signal(t) = a + bt + \beta_1 Q_1 + \beta_2 Q_2 + \beta_3 Q_3$$

where $Q_j$ are indicator functions for each quarter. You could instead mimic these with a sinusoid with period 4:

$$signal(t) = a + bt + A\cos(2\pi t/4) + B\sin(2\pi t/4)$$

(b) How could you remove the trend and seasonality of this time series using differencing? Briefly explain why this removes the trend and seasonality.

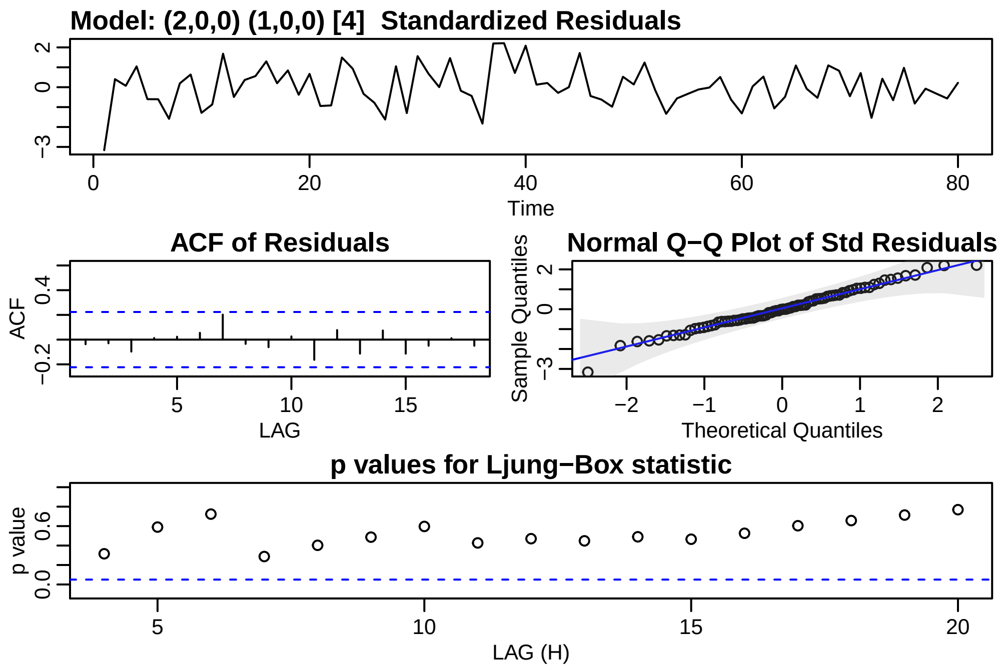
We can take a lag 4 difference:

$$\nabla_4 X_t = X_t - X_{t-4}$$

as the lag 4 difference will remove seasonality with period 4, and any difference will remove the linear trend.

Name: ____________________

(c) So, instead, the data were de-trended by someone else, and they modeled the noise with auto.arima(): ARMA$(2,0)x(1,0)_4$. You have been taught by Professor Fisher that this is a good tool, but perhaps not the best final model, so you are a tad suspicious. Below are the sarima() diagnostic plots. Comment briefly on each plot. Then, conclude whether or not this model seems to fit the noise well.




Residual plot - shows no pattern and fairly constant variance

ACF plot - no significant lags/spikes shown

QQ plot - no points fall outside of the gray bands, so nothing disagrees with the normality of ARMA's residuals (looking good in this plot is not required for white noise, but indicates we may have Gaussian noise, which is nice to work with!)

Ljung-Box - all visible p-values are large, indicating reasonable in-sample fit

All plots indicate reasonable fit. Therefore, this model does seem to fit the noise well.

Name: ____________________

(d) Below are some model fit diagnostics of the ARMA$(2,0)x(1,0)_4$ compared to the client's preference for no noise model (i.e. ARMA(0,0)). Defend your preference for the more complicated model to your client using the information below, and be sure to distinguish in-sample versus out-of-sample. Note that RMSE is the square root of the mean squared error.

| Noise Model | AIC | AICc | BIC | Cross-validated RMSE |
|---|---|---|---|---|
| ARMA$(2,0)x(1,0)_4$ | -1.94 | -1.93 | -1.79 | 0.017 |
| ARMA$(0,0)$ | -1.83 | -1.83 | -1.77 | 0.048 |

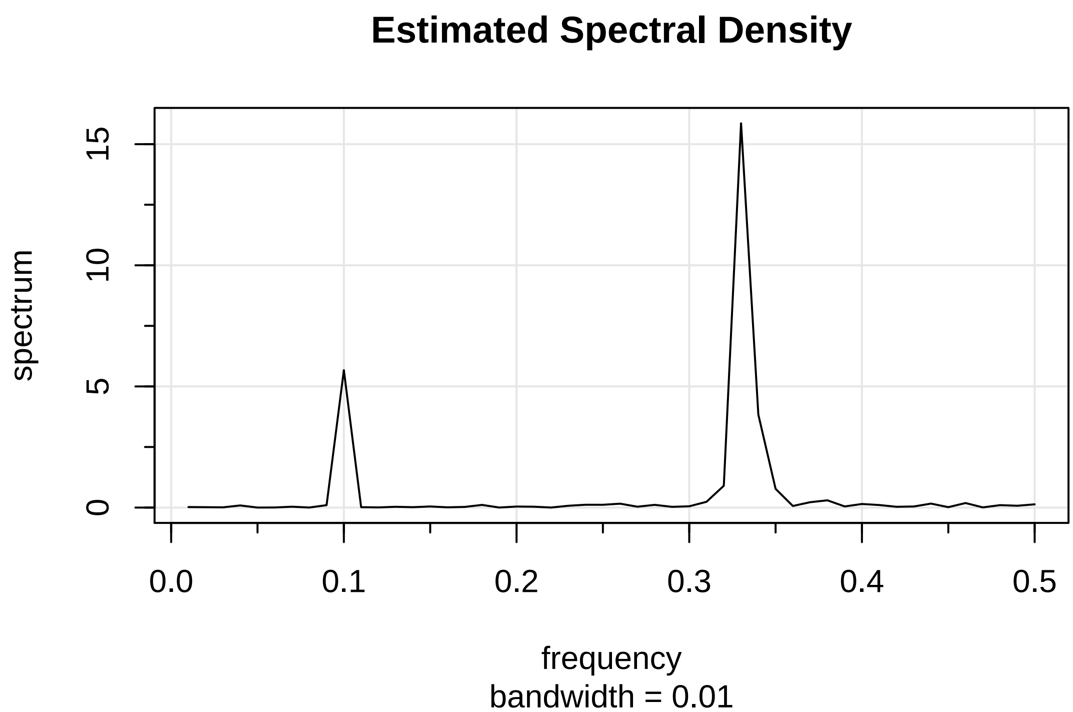
The information criteria penalize against having too large a model, in-sample, and lower values are better. We see that the more complicated model has lower IC values despite the penalty, meaning the data suggest a more complicated model.

Cross validation is the standard process for estimating out-of-sample performance. Here, cross-validation shows a large drop in forecast error when using the more complicated model. Therefore, if the client wants to have the best forecasts of future (log) earnings, the ARMA$(2,0)x(1,0)_4$ process should be used even though it is more complex.

Name: ____________________

7. Finding frequencies and periods.

   (a) The plot below is the spectral density of a process estimated from data. Does this process seem like it could be modeled with sinusoids? If so, what are the PERIODS the sinusoids should have?




ANSWER: We would expect the sinusoids to have periods of 10 and 3.

EXPLANATION: the spectral density shows significant frequencies at 0.1 and about 1/3. Period is the reciprocal of frequency, so the significant periods are approximately 10 and 3.

Name: ____________________

(b) Below are the Fourier coefficients for a time series of length 13. This time series was generated from a single sinusoid plus white noise. What approximately was the frequency of the sinusoid?

$0.29 + 0.00i$

$-0.50 + 0.52i$

$0.36 - 0.15i$

$-1.31 + 0.01i$

$0.00 - 0.01i$

$3.85 + 5.17i$

$0.22 - 0.35i$

$0.22 + 0.35i$

$3.85 - 5.17i$

$0.00 + 0.01i$

$-1.31 - 0.01i$

$0.36 + 0.15i$

$-0.50 - 0.52i$

ANSWER: frequency $j/n = 5/13$.

EXPLANATION: The values above are $b_0, ..., b_12$. Only $b_0, ..., b_7$ provide unique information. We could use these values to build a periodogram. Recall the definition of the periodogram:

$$I(j/n) = \frac{|b_j|^2}{n}$$

and recall that the magnitude of a complex number $a + bi$ is the $\sqrt{a^2 + b^2}$. Of the complex numbers above, the largest one clearly is the sixth one down, $3.85 + 5.17i$, which is $b_5$ (because the list starts with $b_0$!). The other $b_j$ in the first half are relatively small, so the periodogram would only have one large spike at $j = 5$. Thus the important frequency here is (approximately) $j/n = 5/13$.

Name: ____________________

8. Calculate the spectral density of AR(1) and simplify such that no imaginary component remains.

By theorem the formula for ARMA's spectral density is

$$f_X(\lambda) = \sigma_Z^2 \frac{|\theta(e^{-2\pi i j\lambda})|^2}{|\phi(e^{-2\pi i j\lambda})|^2}.$$

For AR(1), $\theta(B) = 1$ and $\phi(B) = 1 - \phi B$. Thus $\theta(e^{-2\pi i j\lambda}) = 1$ and $\phi(e^{-2\pi i j\lambda}) = 1 - \phi e^{-2\pi i(1)\lambda}$.

So we remove the imaginary components by taking the magnitude, then simplify:

$$\begin{aligned}
|1 - \phi e^{-2\pi i\lambda}|^2 &= |1 - \phi(\cos(2\pi\lambda) - i\sin(2\pi\lambda))|^2 \\
&= |1 - \phi\cos(2\pi\lambda) + \phi i\sin(2\pi\lambda)|^2 \\
&= Re(1 - \phi\cos(2\pi\lambda) + \phi i\sin(2\pi\lambda))^2 + Im(1 - \phi\cos(2\pi\lambda) + \phi i\sin(2\pi\lambda))^2 \\
&= (1 - \phi\cos(2\pi\lambda))^2 + (\phi\sin(2\pi\lambda))^2 \\
&= 1 - 2\phi\cos(2\pi\lambda) + \phi^2\cos^2(2\pi\lambda) + \phi^2\sin^2(2\pi\lambda) \\
&= 1 - 2\phi\cos(2\pi\lambda) + \phi^2(\cos^2(2\pi\lambda) + \sin^2(2\pi\lambda)) \\
&= 1 - 2\phi\cos(2\pi\lambda) + \phi^2
\end{aligned}$$

Note that there are other ways to simplify this quantity, and due to the various sine/cosine identities, your simplification of this expression may be different than mine.

Regardless, now plug it in and the solution is

$$f_X(\lambda) = \sigma_Z^2 \frac{1}{1 - 2\phi\cos(2\pi\lambda) + \phi^2}.$$

Name: ____________________

FORMULA SHEET

On the actual final, the formula sheet will be attached to the back. For this practice exam, the formula sheet will be posted on bCourses in the "Final Exam Preparation" folder by end of the workday on Friday 12/13/19.