Name: ____________________

Student ID number: ____________________

Midterm 1

Statistics 153 Introduction to Time Series

March 7th, 2019

General comments:

1. Flip this page only after the midterm has started.
2. Before handing in, write your name one **every** sheet of paper!
3. Anyone caught cheating on this midterm will receive a failing grade and will also be reported to the University Office of Student Conduct. In order to guarantee that you are not suspected of cheating, please keep your eyes on your own materials and do not converse with others during the midterm.

Name: ____________________

1. Consider the following model for time series data $X_t = X_{t-1} + Z_t + \delta$, where δ is some non-zero constant and Z_t is white noise with variance σ^2.

   (a) Give the definition of weak and strong stationarity.

   (4 Points)

   A sequnce of random variables (X_t) is strongly stationary if for any choice of times $t_1, ..., t_k$ and lag h, $(X_{t_1}, ..., X_{t_k})$ is equal in distribution of $(X_{t_1+h}, ..., X_{t_k+h})$.

   The sequence if weakly stationary if X_t all have common mean, and for all choices t, s, and h, $Cov(X_t, X_{t+h}) = Cov(X_s, X_{s+h})$.

   (b) Show that there exist no stationary solution for X_t in the above model.

   (2 Points)

   We observe that $EX_t = EX_{t-1} + \delta$. Since $\delta \neq 0$, $EX_t \neq EX_{t-1}$, so the expectations are not the same for all t.

Name: ____________________

(c) From now on suppose that $X_0 = 0$. Compute the mean and the variance of X_t for all $t > 0$.

(3 Points)

Observe that X_t can be written as

$$X_t = t\delta + \sum_{k=1}^{t} Z_k \tag{1}$$

Taking expectations, we find that $EX_t = t\delta$, since the Z_k's are zero mean.

Next, notice that X_{t-1} and Z_t are uncorrelated, since the Z_t's are assumed to be uncorrelated. Therefore, $Var(X_t) = Var(X_{t-1}) + Var(Z_t) = Var(X_{t-1}) + \sigma^2$. Hence, $Var(X_t) = t\sigma^2$, noting that $Var(X_0) = 0$.

(d) Is X_t homoscedastic? Explain.

(1 Points)

X_t is not homoscedastic because the variance grows with t.

Name: ____________________

(e) Propose an invertible function $f(\cdot)$ such that the transformed data $f(X_t)$ has approximately constant variance. Explain.

Hint: You may assume that all your observations are positive.

(3 Points)

Let $\mu_t = EX_t$, so $Var(X_t) = C\mu_t$ where $C = \sigma^2/\delta$.

We use a variance stabilizing tranformation. We consider a function f, and do a Taylor expansion about μ_t:

$$f(X_t) \approx f(\mu_t) + f'(\mu_t)(X_t - \mu_t) \tag{2}$$

So

$$\begin{aligned} Var(f(X_t)) &\approx [f'(\mu_t)]^2 Var(X_t) && (3) \\ &= [f'(\mu_t)]^2 C\mu_t && (4) \end{aligned}$$

We want the variance to be constant, so we should choose $f(x)$ to satisfy $[f'(x)]^2 = 1/x$, and conclude that $f(x) = \sqrt{x}$.

In summary, we proopose a variance stabilizing transform of $f(X_t) = \sqrt{X}_t$.

Name: ______________________

(f) Propose an invertible transformation of X_t such that it is stationary. Explain.

(3 Points)

Observe that $\nabla X_t = X_t - X_{t-1} = Z_t + \delta \sim WN(\delta, \sigma^2)$. Hence ∇X_t is stationary.

Name: ____________________

2. Consider the stationary, zero-mean AR(1) model $X_t = 0.5X_{t-1} + Z_t$ and the MA(1) model $W_t = 0.5Z_{t-1} + Z_t$, where Z_t is some white noise with variance σ^2.

(a) For each of Z_t, W_t, and X_t give the ACVF and ACF function.

i. For Z_t:

(1 Points)

By definition of white noise, the ACVF fucntion $Cov(Z_t, Z_{t+h})$ is zero if $h > 0$ and σ^2 otherwise. Its ACF is then 0 if $h > 0$ and 1 otherwise.

ii. For W_t:

(2 Points)

We compute

$$Var(W_t) = 0.25Var(Z_{t-1}) + Var(Z_t) = 1.25\sigma^2 \tag{5}$$

Next,

$$\begin{aligned} Cov(W_t, W_{t+1}) &= Cov(0.5Z_{t-1} + Z_t, 0.5Z_t + Z_{t+1}) && (6) \\ &= 0.5Var(Z_t) = 0.5\sigma^2 && (7) \end{aligned}$$

Finally observe that $Cov(W_t, W_{t+h}) = 0$ if $h > 1$.
In summary, the ACVF is $1.25\sigma^2$ if $h = 0$, $0.5\sigma^2$ if $h = 1$, and 0 otherwise. The ACF is then 1 if $h = 0$, 0.4 if $h = 1$, and 0 otherwise.

Name: ____________________

iii. For X_t:

(2 Points)

We recall the MA representation of the AR(1) process, as

$$X_t = \sum_{j=0}^{\infty} (0.5)^j Z_{t-j} \tag{8}$$

Hence we compute the ACVF as

$$ACVF(h) = \sigma^2 \sum_{j=0}^{\infty} 0.5^j 0.5^{j+h} = \sigma^2 0.5^h \sum_{j=0}^{\infty} 0.5^{2j} = \sigma^2 \frac{0.5^h}{1 - 0.5^2} \tag{9}$$

for $h \geq 0$. The ACF is then

$$ACF(h) = 0.5^h, \quad h \geq 0 \tag{10}$$

Name: ____________________

(b) For each of Z_t, W_t, and X_t give the approximate mean and variance of its sample ACF at lag 2 for $n = 100$ observations.

Hint: Recall Bartlett's formula $W_{ij} = \sum_{m=1}^{\infty} (\rho(m+i) + \rho(m-i) - 2\rho(i)\rho(m)) (\rho(m+j) + \rho(m-j) - 2\rho(j)\rho(m))$

i. For Z_t:

(2 Points)

Bartlett's formula says that the sample ACF at lag 2, r_2, is approximately normal with mean $\rho(2)$ and variance $W_{22}/100$,

$$W_{22} = \sum_{m=1}^{\infty} (\rho(m+2) + \rho(m-2) - 2\rho(2)\rho(m))^2 \tag{11}$$

So for Z_t, W_{22} is given by 1 since $\rho(m) = 0$ for $m > 0$.
In sum, $E\hat{r}_2 \approx 0$ and $Var(\hat{r}_2) \approx 1/100$.

Name: ____________________

ii. For W_t:

(4 Points)

For W_t, we have $\rho(m) = 0$ for $m \geq 1$, so we compute

$$W_{22} = \rho(0)^2 + \rho(-1)^2 + \rho(1)^2 = 2\rho(1)^2 = 1 + 2(0.4)^2$$

So $Var(\hat{r}_2) \approx 1.32/100$, and $E\hat{r}_2 \approx 0$.

Name: ____________________

iii. For X_t:

(4 Points)

Recall that $\rho(h) = 0.5^{|h|}$. Hence, here

$$W_{22} = \sum_{m=1}^{\infty} (0.5^{m+2} + 0.5^{|m-2|} - 2(0.5)^{2+m})^2 \tag{12}$$

$$= (0.5^3 + 0.5^1 - 2(0.5)^3)^2 + \sum_{m=2}^{\infty} (0.5^{m+2} + 0.5^{m-2} - 2(0.5)^{2+m})^2 \tag{13}$$

$$= (0.375)^2 + \sum_{m=2}^{\infty} (0.5)^{2m} (0.5^2 + 0.5^{-2} - 2(0.5)^2)^2 \tag{14}$$

$$= (0.375)^2 + (3.75)^2 \sum_{m=2}^{\infty} (0.5)^{2m} \tag{15}$$

$$= (0.375)^2 + (3.75)^2 \frac{1}{12} = \frac{21}{16} \tag{16}$$

In summary, $E\hat{r}_2 \approx (0.5)^2$ and $Var(\hat{r}_2) \approx W_{22}/100 = 21/1600$.

Name: ____________________

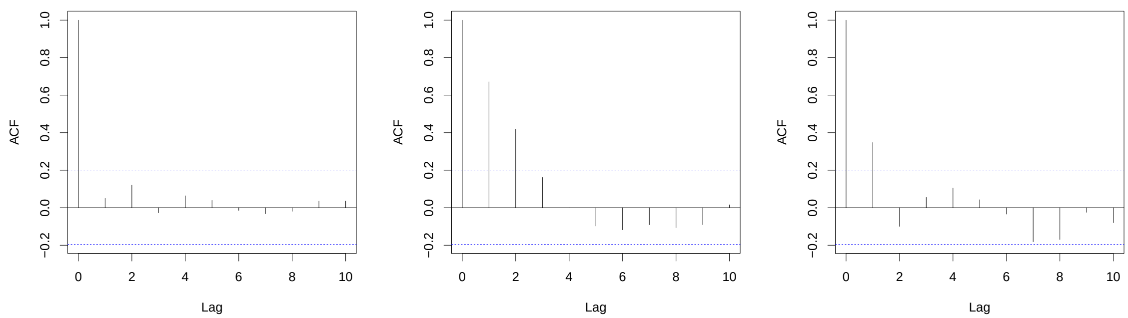

Figure 1: Sample ACFs of different time series data.

(c) Figure 1 shows sample ACFs for each of the three models for $n = 100$ observations. Which figure corresponds to which process? Explain.

(3 Points)

The first one looks the most like white noise, since the ACF values are all below the blue line.

The second looks most like the AR process X_t, since we computed the ACF(h) to decrease geometrically as $(0.5)^h$.

The last looks most like the MA process W_t since we computed the ACF to be 0.4 at $h = 1$ and 0 for $h > 1$.

Name: ____________________

3. For zero mean time series data $\{X_t\}$ consider the model $(1 - 0.2B)(X_t - 0.5X_{t-1}) = (Z_t - 0.6Z_{t-1} + 0.05Z_{t-2})$, where $\{Z_t\}$ is white noise with variance $\sigma^2 = 4$.

   (a) Identify $\{X_t\}$ as an ARMA(p,q) model and give its MA and AR polynomials.

   (4 Points)

Writing X_t in polynomial and canceling common factors gives:

$$(1 - .2B)(1 - .5B)X_t = (1 - .1B)(1 - .5B)Z_t$$
$$(1 - .2B)X_t = (1 - .1B)Z_t$$

We have an ARMA(1,1) model. The AR polynomial is $1 - .2B$, and the MA polynomial is $1 - .1B$.

Name: ____________________

(b) Is the model invertible and causal?

(2 Points)

The AR polynomial $1 - .2B$ has a root of 5, and the MA polynomial $1 - .1B$ has a root of 10. Both have magnitude greater than 1, and so this model is both causal and invertible.

(c) Find its unique stationary solution.

(4 Points)

We need to "solve for" X_t. Since the model is causal, we can invert the AR polynomial:

$$\begin{aligned}
\frac{1}{1-.2B}(1-.2B)X_t &= \frac{1}{1-.2B}(1-.1B)Z_t \\
\implies X_t &= (\sum_{j\geq 0}(\frac{B}{5})^j)(Z_t - \frac{1}{10}Z_t) \\
&= \sum_{j\geq 0}(\frac{1}{5})^j Z_{t-j} - \frac{1}{10}\sum_{k\geq 0}(\frac{1}{5})^k Z_{t-1-k} \\
&= \sum_{j\geq 0}(\frac{1}{5})^j Z_{t-j} - \frac{1}{2}\sum_{k\geq 0}(\frac{1}{5})^{k+1} Z_{t-(k+1)} \\
&= \sum_{j\geq 0}(\frac{1}{5})^j Z_{t-j} - \frac{1}{2}\sum_{k\geq 1}(\frac{1}{5})^k Z_{t-k} \\
&= Z_t + \frac{1}{2}\sum_{j\geq 1}(\frac{1}{5})^j Z_{t-j}
\end{aligned}$$

Name: ____________________

(d) Compute its ACVF.

(4 Points)

The Z_t term isn't being multiplied by a $\frac{1}{2}$, so in its current form we can't treat the ACVF calculation like that of a causal AR(1) process. However, a small trick will simplify the calculation.

Rewriting, $X_t = Z_t + \frac{1}{2}\sum_{j\geq 1}(\frac{1}{5})^j Z_{t-j} = \frac{1}{2}Z_t + \frac{1}{2}(\sum_{j\geq 0}(\frac{1}{5})^j Z_{t-j}) = \frac{1}{2}(Z_t + Y_t)$. Note that Y_t is the MA(∞) form of a causal AR(1) model with $\phi = \frac{1}{5}$, whose ACVF we already know. Thus, using bilinearity to expand the covariance:

$$\begin{aligned} Cov(X_{t+h}, X_t) &= \frac{1}{4}Cov(Z_{t+h} + Y_{t+h}, Z_t + Y_t) \\ &= \frac{1}{4}[Cov(Z_{t+h}, Z_t) + Cov(Z_{t+h}, Y_t) + Cov(Z_t, Y_{t+h}) + Cov(Y_{t+h}, Y_t)] \end{aligned}$$

Now, we can use this to treat two cases:

h=0. All covariance terms are nonzero:

$$Cov(X_t, X_t) = \frac{1}{4}[4 + 4 + 4 + \frac{4}{1 - \frac{1}{5}^2}] = 4 + \frac{1}{24} = \frac{97}{24}$$

$\boldsymbol{h > 0}$. Now we can ignore the first two terms.

$$\begin{aligned} Cov(X_{t+h}, X_t) &= \frac{1}{4}[Cov(Z_t, Y_{t+h}) + Cov(Y_{t+h}, Y_t)] \\ &= \frac{1}{4}[4(\frac{1}{5})^h + \frac{4(\frac{1}{5})^h}{1 - \frac{1}{5}^2}] = (1 + \frac{25}{24})(\frac{1}{5})^h = (\frac{49}{24})(\frac{1}{5})^h \end{aligned}$$

Name: ____________________

(e) Assume someone wants to use this model to predict weekly car sales. On average the company sells 100 cars per week. Two weeks ago they sold 95 cars and last week they sold 101 cars. Based on this, what is the best linear predictor of car sales next week?
Hint: You do not have to compute the actual value, it is enough to write down a linear system of equations that needs to be solved.

(3 Points)

By causality, we know X_t has a MA(∞) representation and therefore is mean zero. However, we're given that the average of car sales Y_t is 100 per week. Assuming this is (approximately) the population mean, we instead model $Y_t - 100 = X_t$ using the ARMA process.

Thus, we can use the defining equations of the BLP for mean-zero RV's:

$$E((X_t - \beta_1 X_{t-1} - \beta_2 X_{t-2})X_{t-1}) = 0 \quad ; \quad E((X_t - \beta_1 X_{t-1} - \beta_2 X_{t-2})X_{t-2}) = 0$$

Putting this in matrix form, we have:

$$\begin{bmatrix} \gamma(0) & \gamma(1) \\ \gamma(1) & \gamma(0) \end{bmatrix} \beta = \begin{bmatrix} \gamma(1) \\ \gamma(2) \end{bmatrix}$$

Plugging in values:

$$\begin{bmatrix} 97/24 & 49/24(1/5) \\ 49/24(1/5) & 97/24 \end{bmatrix} \beta = \begin{bmatrix} 49/24(1/5) \\ 49/24(1/25) \end{bmatrix}$$

We can solve for β to obtain the weights of our BLP. So, the BLP of X_t is $\beta_1(1)+\beta_2(-5)$, and as $X_t = Y_t - 100$ our prediction for sales is $\hat{Y}_t = \beta_1(1) + \beta_2(-5) + 100$.

Name: ____________________

4. A scientist considers the model $X_t = m_t + s_t + W_t$ for some time series data, where $m_t = a\,t + b$ is a linear trend function with parameters a, b and s_t is a seasonal component with period 2, that is, $s_t = s_{t+2}$ for all t. W_t is some zero mean stationary process.

   (a) First, the scientist wants to estimate the trend function m_t using a filter of the form $1 + \alpha B + \beta B^2 + \gamma B^3$, where B denotes the backshift operator and α, β, γ are parameters. How should she chose the parameters α, β, γ such that the filtered time series is an unbiased estimator of the trend m_t, that is, $E((1 + \alpha B + \beta B^2 + \gamma B^3)X_t) = m_t$?
   Hint: First, argue that without loss of generality you can assume that $s_1 + s_2 = 0$.

   (5 Points)

First, we tackle the hint. If $s_t + s_{t+1} = \delta \neq 0$, let $b' = b + \frac{\delta}{2}$ and $s'_t = s_t - \frac{\delta}{2}$. Then,

$$X_t = at + b + s_t + W_t = at + b' + s'_t + W_t$$

and we may assume WLOG that $s_t + s_{t+1} = 0$.

Upon taking the expectation and setting $s_{t-2} = s_t$, we have:

$$E((1 + B + \beta B^2 + \gamma B^3)X_t) = (1 + \alpha + \beta + \gamma)(at + b) - (\alpha + 2\beta + 3\gamma) + [(1 + \beta)s_t + (\alpha + \gamma)s_{t-1}]$$

Since we want $s_t + \alpha s_{t-1} + \beta s_t + \gamma s_{t-1} = 0$, by the hint one way to do that is to have $1 + \beta = \alpha + \gamma$.

Therefore, we have three equations we want to satisfy:

$$1 + \alpha + \beta + \gamma = 1 \tag{1}$$
$$\alpha + 2\beta + 3\gamma = 0 \tag{2}$$
$$1 + \beta = \alpha + \gamma \tag{3}$$

Substituting (3) into (1), we see:

$$(1 + \beta) = \frac{1}{2} \implies \beta = -\frac{1}{2}$$

Then, plugging $\beta = -\frac{1}{2}$ into (2) and adding it to (3) gives $\gamma = \frac{1}{4}$. Solving for α then gives $\alpha = \frac{1}{4}$.

Therefore, one possible filter is $\alpha = \gamma = \frac{1}{4},\ \beta = -\frac{1}{2}$.

Name: ____________________

(b) Is X_t a stationary process? Explain.

(1 Points)

In general, no.

$$E(X_t) = at + b + s_t$$
$$E(X_{t-2}) = a(t-2) + b + s_{t-2} = at + b + s_t - 2a$$

Therefore, stationarity requires that $a = 0$. A similar analysis shows that $s_t = s_{t-1}$ and so s_t must be a constant. The model is stationary only when it has no linear trend or seasonality at all - ie. it's just white noise plus a constant.

(c) Propose a transformation using differencing to make the process stationary. Explain.

(3 Points)

Since our trend is a sum of linear and seasonal components, we can difference appropriately to handle each. In fact, since differencing by any lag destroys a linear trend, we can just difference by lag 2, the period, to get rid of both trends at the same time

$$\nabla_2 X_t = at + b + s_t + W_t - [a(t-2) + b + s_{t-2} + W_{t-2}] = 2a + W_t - W_{t-2}$$

A linear sum of stationary process is stationary.

Name: ____________________

(d) For the stationary process W_t the scientist considers two different models:

- an MA(1) model,
- an AR(1) model.

For both of these choices identify the transformed data from (4c) as some ARMA model.
Hint: It is enough to state the orders of the respective ARMA models with explanation.

(6 Points)

MA(1) model. $W_t = (1 - \theta B)Z_t$. Let $Y_t = \nabla_2 X_t$. Then,

$$Y_t - 2a = (1 - B^2)W_t = (1 - B^2)(1 - \theta B)Z_t = (1 - \theta B - B^2 + \theta B^3)Z_t$$

So $Y_t = \nabla_2 X_t$ is a MA(3) model.

AR(1) model. $Z_t = (1 - \phi B)W_t$. Then,

$$\begin{aligned} Y_t - 2a &= (1 - B^2)W_t \\ \implies (1 - \phi B)(Y_t - 2a) &= (1 - B^2)(1 - \phi B)W_t \\ &= (1 - \phi B)(Y_t - 2a) = (1 - B^2)Z_t \end{aligned}$$

Therefore, Y_t is a ARMA(1,2) process.

Name: ____________________

5. For each statement, indicate whether it is true or false and give a short explanation.
   You only get points when both, True/False and the explanation, are correct.

   (a) For the sample autocorrelations of $n = 1,000$ i.i.d. white noise random variables at lags $h = 1, \ldots, 100$, you expect on average 5 of them to be larger than 1.96 in absolute value.

   [ ] True [x] False

   Explanation: You expect on average 5 of them to be larger than $1.96/\sqrt{n} = 0.0196$ in absolute value.

   (b) The sample autocorrelations of an AR(1) process with i.i.d. white noise are (for large sample size) approximately i.i.d..

   [ ] True [x] False

   Explanation: By Bartlett's formula they are going to be correlated and hence not i.i.d..

   (c) Applying a linear (time invariant) filter to a stationary process results again in a stationary process.

   [x] True [ ] False

   Explanation: This follows easily by bilinearity of the covariance.

   (d) When you want to fit a seasonal parametric function of the form $s_t = a_0 + \sum_{f=1}^{k} (a_f \cos(2\pi ft/d) + b_f \sin(2\pi ft/d))$ with parameters $a_0, a_1, \ldots, a_k, b_1, \ldots, b_k$ it can be helpful to chose $k > d/2$.

   [ ] True [x] False

   Explanation: For $k \geq d/2$ function s_t has at least d parameters and thus, every d-periodic seasonal function can be written in that form. There is no point in choosing $k > d/2$.

   (e) A time series $\{X_t\}$ where X_t follows a Gaussian distribution for each t is a Gaussian process.

   [ ] True [x] False

   Explanation: One also needs that $(X_{t_1}, \ldots, X_{t_k})$ is multivariate Gaussian for every collection $t_1, \ldots, t_k$.

   (f) Whether a time series is invertible or not is fully determined by its finite dimensional distributions.

   [ ] True [x] False

   Explanation: Invertibility is a property of the relation of the white noise $\{Z_t\}$ and the time series $\{X_t\}$, not of $\{X_t\}$ alone.

   (g) Whether a time series is strongly stationary or not is fully determined by its mean and covariance function.

   [ ] True [x] False

   Explanation: In general this depends on the full finite dimensional distributions.

   (h) Whether a Gaussian process is strongly stationary or not is fully determined by its mean and covariance function.

Name: ____________________

[x] True [ ] False

Explanation: The Gaussian distribution is already fully determined by mean and covariance structure and hence the assertion is true.

(8 Points)