Exam 3

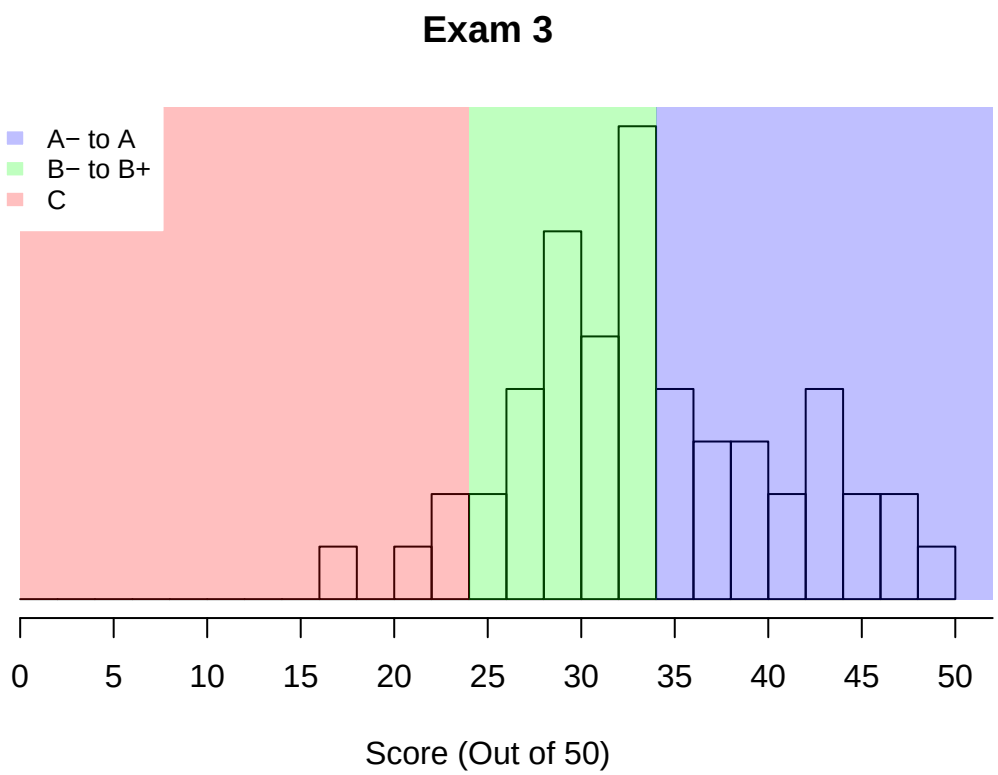
5/3/19

There are 5 questions, each of which has several parts, as well as a bonus question. Neither the questions nor the parts are necessarily in order from easiest to most difficult. Make sure you have taken a look at and attempted all of the questions in the allotted time. Stop working and immediately turn in your exam when time has been called.

Name: _____

Question	Maximum Possible Points
1	10
2	9
3	6
4	7
5	18
Bonus	1
Total	50

The figure below shows a histogram of scores. The mean was 34.1, the median was 33, and the standard deviation was 7.15. The rank correlations between the first and second, second and third, and first and third exam scores scores were 0.58, 0.49, and 0.58.



1. The Univariate State-Space Model and ARIMA

The basic linear state-space model is given by:

$$
\begin{aligned}
y_t &= a x_t + \boldsymbol{z}_t' \boldsymbol{\gamma} + v_t && \text{Observation Equation} \\
x_t &= \phi x_{t-1} + \boldsymbol{z}_t' \boldsymbol{\upsilon} + w_t && \text{State Equation,}
\end{aligned}
$$

where $v_t \overset{i.i.d.}{\sim} \mathcal{N}\left(0, \sigma_v^2\right)$, $w_t \overset{i.i.d.}{\sim} \mathcal{N}\left(0, \sigma_w^2\right)$ and $x_1 \sim \mathcal{N}\left(\mu, \sigma_x^2\right)$ and $t = 1, \ldots, n$.

(a) This state-space model includes an $\mathbf{AR}(1)$ model as a special case. How is the $\mathbf{AR}(1)$ model obtained? Indicate which parameters need to be fixed, and the values they need to be fixed to. Don't forget μ and σ_x^2!

$$\boldsymbol{\gamma} = \mathbf{0},\ \boldsymbol{\upsilon} = \mathbf{0},\ a = 1,\ \sigma_v^2 = 0$$

Full credit was given for $\mu = 0$ or $\mu = \mu$ abd $\sigma_x^2 = \frac{\sigma_w^2}{1-\phi^2}$ or $\sigma_x^2 = \sigma_x^2$, since I did not specify if the $\mathbf{AR}(1)$ model should be mean-zero but we generally worked with mean-zero models in class and we often interpret $\sigma_x^2 = \mathbb{V}[x_t]$. I also accepted $\mu = x_1$ if $\sigma_x^2 = 0$.

I gave full credit if you got everything right, took off 1 point if you got almost everything right, 2 points if there were some serious issues but you tried, and 3 points if you did nothing or very little.

(b) Set $\phi = 1$, $\boldsymbol{\gamma} = \mathbf{0}$ and $\boldsymbol{\upsilon} = \mathbf{0}$. Using the lag-operator notation $Bx_t = x_{t-1}$ solve the state equation for x_t.

$$x_t = \left(\frac{1}{1-B}\right) w_t$$

(c) Plug your expression from (b) into the observation equation, and multiply through to eliminate any lag operators that appear as denominators. Collect all of the terms that involve y_t on the left-hand side, and put the rest on the right-hand side. Define z_t to be equal to the left-hand side.

$$
\begin{aligned}
y_t &= a\left(\frac{1}{1-B}\right) w_t + v_t \\
\underbrace{(1-B)\, y_t}_{z_t} &= a w_t + (1-B)\, v_t
\end{aligned}
$$

Some students got confused about the request that you define z_t to be equal to the left-hand side, because there is already a z_t in the model - this was an oversight on my part. I meant for you to define a new time series process to make (d) easier, but had forgotten we already had a z_t, especially because we set $\boldsymbol{\gamma} = \mathbf{0}$ and $\boldsymbol{\upsilon} = \mathbf{0}$ in (b). I can see why this was confusing, so I did not deduct any points if I saw that people were trying to use the $\boldsymbol{z}_t$ from the original state equation with $\boldsymbol{\gamma} \neq \mathbf{0}$.

(d) Derive the autocovariance function $\gamma_z(h)$ of z_t.

Because z_t is a moving average process, we have:

$$\gamma_z(h) = \begin{cases} a^2\sigma_w^2 + 2\sigma_v^2 & h = 0 \\ -\sigma_v^2 & |h| = 1 \\ 0 & \text{otherwise} \end{cases}$$

Although I did not explicitly ask for it, I was hoping you would notice that the model from (b) corresponds to an **ARIMA**$(0, 1, 1)$ model for y_t!

2. State-Space Model Smoothing

Let's consider the a tiny example of the basic linear state-space model from the previous part, with $a = 1$, $|\phi| < 1$, $\mu = 0$, $\sigma_x^2 = \sigma_w^2/\left(1-\phi^2\right)$ and $n = 2$. Remember that we know that the joint distribution of the time series and the latent states is the nicely structured normal distribution,

$$\begin{pmatrix} \boldsymbol{y} \\ \boldsymbol{x} \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} a\mathbb{E}[\boldsymbol{x}] \\ \mathbb{E}[\boldsymbol{x}] \end{pmatrix}, \begin{pmatrix} a^2\mathbb{V}[\boldsymbol{x}] + \sigma_v^2\boldsymbol{I}_n & a\mathbb{V}[\boldsymbol{x}] \\ a\mathbb{V}[\boldsymbol{x}] & \mathbb{V}[\boldsymbol{x}] \end{pmatrix}\right).$$

(a) Write the equation for the smoothed values of the states from the notes. Simplify as much as you can without multiplying or inverting any matrices.

$$\begin{aligned} \mathbb{E}[\boldsymbol{x}|\boldsymbol{y}] =& \mathbb{E}[\boldsymbol{x}] + a\mathbb{V}[\boldsymbol{x}]\left(a^2\mathbb{V}[\boldsymbol{x}] + \sigma_v^2\boldsymbol{I}_n\right)^{-1}(\boldsymbol{y} - a\mathbb{E}[\boldsymbol{x}]) \\ =& \begin{pmatrix} 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 1 & \phi \\ \phi & 1 \end{pmatrix}\left(\begin{pmatrix} 1 & \phi \\ \phi & 1 \end{pmatrix} + \frac{\sigma_v^2\left(1-\phi^2\right)}{\sigma_w^2}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\right)^{-1}\begin{pmatrix} y_1 - 0 \\ y_2 - 0 \end{pmatrix} \end{aligned}$$

(b) Based on your intuition and your answer to (b), describe what happens to the smoothed values as $\sigma_v^2 \to 0$, but all of the other parameters are held constant in one sentence.

- As $\sigma_v^2 \to 0$, the smoothed values will shrink towards the observed values and become less smooth.
- Decreasing σ_v^2 makes the conditional means of the states more closely track the observed data, and it makes the conditional variances of the states smaller.

(c) Based on your intuition and your answer to (b), describe what happens to the smoothed values as $\sigma_w^2 \to 0$, but all of the other parameters are held constant in one sentence.

- As $\sigma_w^2 \to 0$, the smoothed values will shrink more towards $\mathbf{0}$, the mean of the state model, and become smoother.
- Decreasing σ_w^2 makes the conditional means of the states vary more smoothly and less like the observed data, and makes the conditional variances of the states smaller.

(d) Based on your intuition and your answer to (b), describe what happens to the smoothed values as $\phi \to 0$, but all of the other parameters are held constant in one sentence.

- If $\phi > 0$ and $\phi \to 0$, the smoothed values will shrink less towards each other and and become less smooth, whereas if $\phi < 0$ and $\phi \to 0$, the smoothed values will shrink more towards each other and become smoother.
- The smoothed values will become proportional to observed values of $\boldsymbol{y}$ as $\phi \to 0$

3. Stochastic Volatility Model Properties

The stochastic volatility model is given by

$$
\begin{aligned}
y_t &= \exp\{h_t/2\}\, v_t && \text{Observation Equation} \\
(h_t - \mu_h) &= \phi\,(h_{t-1} - \mu_h) + w_t && \text{State Equation,}
\end{aligned}
$$

where $|\phi| < 1$, $v_t \overset{i.i.d.}{\sim} \mathcal{N}(0,1)$, $w_t \overset{i.i.d.}{\sim} \mathcal{N}\left(0, \sigma_w^2\right)$ and $h_1 \sim \mathcal{N}\left(\mu_h, \sigma_w^2/\left(1-\phi^2\right)\right)$.

(a) Derive $\mu_y = \mathbb{E}[y_t]$.

$$
\begin{aligned}
\mu_y &= \mathbb{E}[y_t] \\
&= \mathbb{E}[\exp\{h_t/2\}\, v_t] \\
&= \mathbb{E}[\exp\{h_t/2\}]\,\mathbb{E}[v_t] = 0
\end{aligned}
$$

(b) Derive $\gamma_y(h) = \mathrm{Cov}[y_{t-h}, y_t]$ for $h > 0$.

If $h > 0$,

$$
\begin{aligned}
\gamma_y(h) &= \mathrm{Cov}[y_{t-h}, y_t] \\
&= \mathbb{E}[y_t y_{t-h}] \\
&= \mathbb{E}[\exp\{(h_t + h_{t-h})/2\}\, v_t v_{t-h}] \\
&= \mathbb{E}[\exp\{(h_t + h_{t-h})/2\}\, v_{t-h}]\,\mathbb{E}[v_t] \\
&= 0.
\end{aligned}
$$

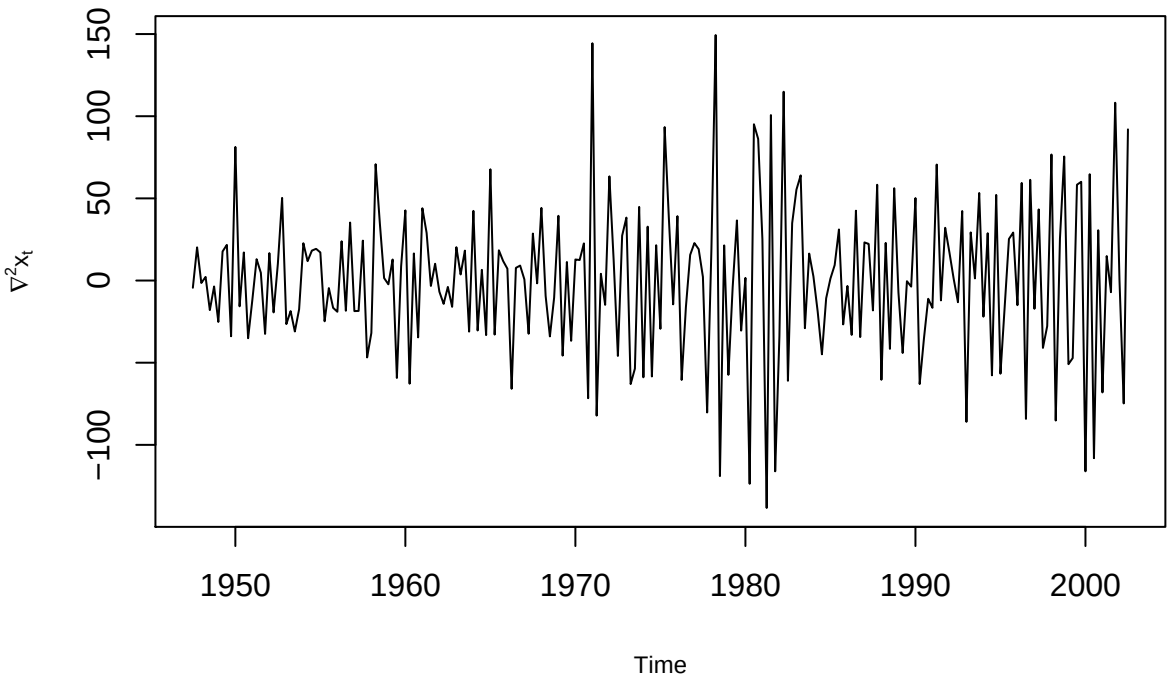
(c) Are successive values of y_t correlated under the stochastic volatility model? Answer by writing "yes" or "no."

No.

4. An Application of the Stochastic Volatility Model

```
library(astsa)
data(gnp)
time <- time(gnp)
gnp <- c(gnp) - mean(gnp)
n <- length(gnp)
```

We're going to keep working with the `gnp` data from the `astsa` package, and continue to focus on the second differences $\nabla^2 x_t$, which are plotted again below.



(a) Suppose we fit a stochastic volatility model to the second differences, using the default priors specified by the `stochvol` package.

```
library(stochvol)
sims <- svsample(diff(gnp, d = 2),
                 draws = 100000)
```

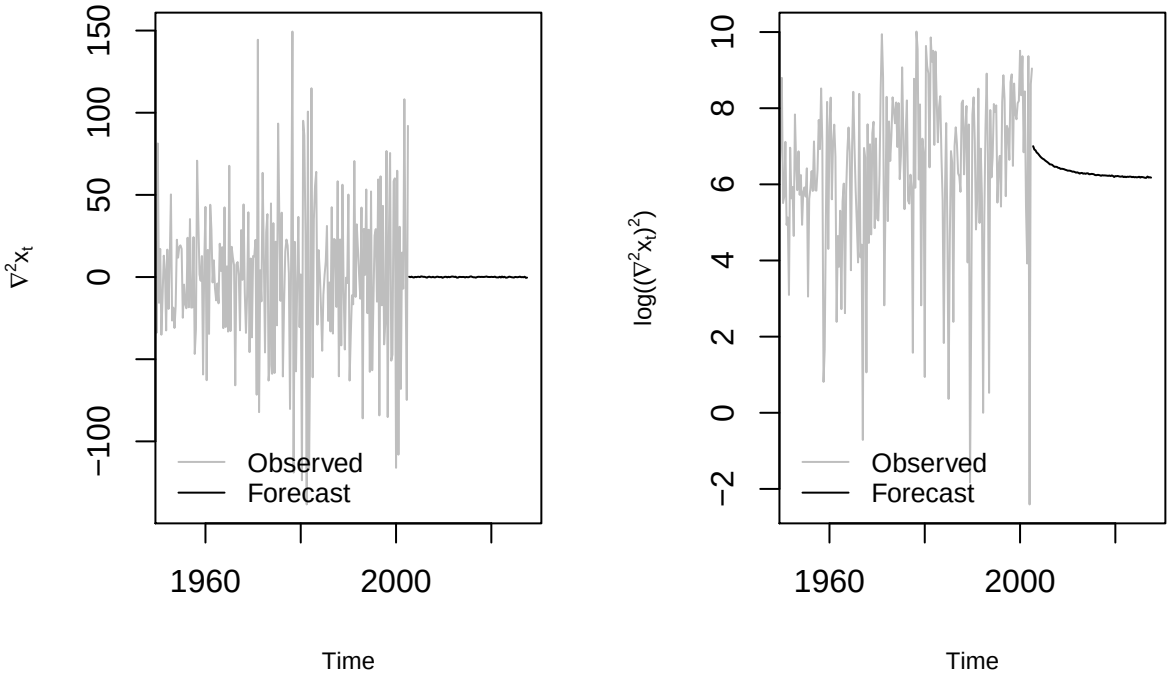
This `svsample` command returns an object `sims` that contains:

- $100,000$ simulated values of ϕ: $\phi^{(1)}, \ldots, \phi^{(1)}$
- $100,000$ simulated values of μ_h: $\mu_h^{(1)}, \ldots, \mu_h^{(1)}$
- $100,000$ simulated values of σ_w^2: $\left(\sigma_w^2\right)^{(1)}, \ldots, \left(\sigma_w^2\right)^{(1)}$

Using the output of the `svsample` command, how would you approximate $\mathbb{E}\left[\phi | \nabla^2 x_1, \ldots, \nabla^2 x_n\right]$? Explain using an equation, supplemented by at most one sentence. No need to write out the `R` code.

I would approximate $\mathbb{E}\left[\phi | \nabla^2 x_1, \ldots, \nabla^2 x_n\right]$ using $\frac{1}{100,000} \sum_{i=1}^{100,000} \phi^{(i)}$.

(b) Approximate forecasts of the second differences $\mathbb{E}\left[\nabla^2 x_{n+k} | \nabla^2 x_1, \ldots, \nabla^2 x_n\right]$ and approximate forecasts of the log magnitudes of the second differences $\mathbb{E}\left[\log\left(\left(\nabla^2 x_{n+k}\right)^2\right) | \nabla^2 x_1, \ldots, \nabla^2 x_n\right]$ based on this stochastic volatility model are given in the plots below. Based the plotted forecasts, does assuming a stochastic volatility model for $\nabla^2 x_t$ help us forecast $\nabla^2 x_t$ or the log magnitudes $\log\left(\left(\nabla^2 x_t\right)^2\right)$?



Assuming a stochastic volatility model helps us forecast the log magnitudes $\log\left(\left(\nabla^2 x_t\right)^2\right)$ but not the values $\nabla^2 x_t$.

(c) In one sentence, explain why your answer in (b) makes sense given what we know about the stochastic volatility model the previous question.

It makes sense that the stochastic volatility model does not help us forecast the values $\nabla^2 x_t$ because successive values $\nabla^2 x_t$ are uncorrelated under the stochastic volatility model.

5. Stationarity for Multivariate Time Series

Consider the following model for two simultaneously observed time series, $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$:

$$
\begin{aligned}
x_{t1} &= w_t + w_{t-1} \\
x_{t2} &= v_t + v_{t-1},
\end{aligned}
$$

where

$$
\begin{pmatrix} w_t \\ v_t \end{pmatrix} \overset{i.i.d}{\sim} \mathcal{N}\left(\mathbf{0}, \begin{pmatrix} \sigma_v^2 & \sigma_v \sigma_w \rho^t \\ \sigma_v \sigma_w \rho^t & \sigma_w^2 \end{pmatrix}\right),
$$

and $|\rho| < 1$.

(a) What is the covariance function of x_{t1}?

$$
\gamma_{x_1}(h) = \begin{cases} 2\sigma_v^2 & h = 0 \\ \sigma_v^2 & |h| = 1 \\ 0 & \text{otherwise} \end{cases}
$$

(b) Is x_{t1} stationary? Write "yes" or "no."

Yes.

(c) Is x_{t2} stationary? Write "yes" or "no," but also explain your answer in at most one sentence.

Yes, it is exactly the same type of process as x_{t1}.

(d) What is the cross-covariance function of x_{t1} and x_{t2}?

$$\gamma_{x_1 x_2}(h) = \begin{cases} \sigma_v \sigma_w \left(\rho^t + \rho^{t-1}\right) & h = 0 \\ \sigma_v \sigma_w \rho^t & h = 1 \\ \sigma_v \sigma_w \rho^{t-1} & h = -1 \\ 0 & \text{otherwise} \end{cases}$$

(e) Are these processes jointly stationary? Write "yes" or "no."

No.

(f) Suppose we change our model for v_t and w_t. If

$$\begin{pmatrix} w_t \\ v_t \end{pmatrix} \overset{i.i.d}{\sim} \mathcal{N}\left(\mathbf{0}, \begin{pmatrix} \sigma_v^2 & \sigma_v\sigma_w\rho \\ \sigma_v\sigma_w\rho & \sigma_w^2 \end{pmatrix}\right),$$

where $|\rho| < 1$. What is the cross-covariance function of x_{t1} and x_{t2}?

$$\gamma_{x_1 x_2}(h) = \begin{cases} 2\sigma_v\sigma_w\rho & h = 0 \\ \sigma_v\sigma_w\rho & |h| = 1 \\ 0 & \text{otherwise} \end{cases}$$

(g) If we assume the distribution for the errors given in (g), are these processes jointly stationary? Write "yes" or "no."

Yes.

Bonus

In one sentence, explain whether you generally prefer the ARMA modeling framework or the state-space modeling framework and why. You will receive full credit for your answer as long as it is not something that is false.