Exam 1 is Wed. Sept. 18. **You are allowed 4 sheets of notes and a calculator.** The exam covers HW1-3 and Q1-3. Numbers refer to types of problems on exam.

A *time series* $Y_1, ..., Y_n$ consists of observations Y_t collected sequentially in time.

A *stochastic process* $\{Y_t, t \in \tau\}$ is a collection of random variables where we will usually use $\tau = \{0, \pm 1, \pm 2, ..., \} = \mathbb{Z}$, the set of integers.

The *mean function* $\mu_t = E(Y_t)$ for $t \in \mathbb{Z}$.

The *autocovariance function* $\gamma_{t,s} = Cov(Y_t, Y_s) = E[(Y_t - \mu_t)(Y_s - \mu_s)] = E(Y_t Y_s) - \mu_t \mu_s$ for $t, s \in \mathbb{Z}$.

The *autocorrelation function* $\rho_{t,s} = Corr(Y_t, Y_s) = \dfrac{Cov(Y_t, Y_s)}{\sqrt{Var(T_t)Var(Y_s)}} = \dfrac{\gamma_{t,s}}{\gamma_{t,t}\gamma_{s,s}}$ for $t, s \in \mathbb{Z}$.

Know that $\gamma_{t,s} = \gamma_{s,t}$, $\rho_{t,s} = \rho_{s,t}$, Cauchy Schwarz inequality $\gamma_{t,s} \leq \sqrt{\gamma_{t,t}\gamma_{s,s}}$ and $|\rho_{t,s}| \leq 1$.

Know that $Cov(\sum_{i=1}^m c_i Y_{t_i}, \sum_{j=1}^n d_j Y_{t_j}) = \sum_{i=1}^m \sum_{j=1}^n c_i d_j Cov(Y_{t_i}, Y_{t_j})$.

Know that $Var(\sum_{i=1}^n c_i Y_{t_i}) = \sum_{i=1}^n c_i^2 Var(Y_{t_i}) + 2 \sum_{i=2}^n \sum_{j=1}^{i-1} c_i c_j Cov(Y_{t_i}, Y_{t_j})$.

1) Be able to compute $Cov(k + aX + bY + cZ, h + dX + eY + fZ) = Cov(aX + bY + cZ, dX + eY + fZ) = adCov(X, X) + aeCov(X, Y) + afCov(X, Z) + bdCov(Y, X) + beCov(Y, Y) + bfCov(Y, Z) + cdCov(Z, X) + ceCov(Z, Y) + cfCov(Z, Z)$. Simplify using $Cov(W, W) = V(W)$ and $Cov(U, W) = Cov(W, U)$. See HW1 F. Here the additive constants k and h could depend on time, eg $k = 4 + 5t$.

2) Be able to compute $E(Y_t) = \mu_t, \quad Var(Y_t) = \gamma_{t,t}, \quad Cov(Y_t, Y_s) = \gamma_{t,s}$ and $Corr(Y_t, Y_s) = \rho_{t,s}$ for simple stochastic processes.

A process $\{Y_t\}$ is *strictly stationary* if the joint distribution of $Y_{t_1}, ..., Y_{t_n}$ is the same as the joint distributions of $Y_{t_1-k}, ..., Y_{t_n-k}$ for all choices of time points $t_1, ..., t_n$ and for all choices of time lag k.

A process $\{Y_t\}$ is **stationary** if a) $E(Y_t) = \mu_t \equiv \mu$ is constant over time, and b) $\gamma_{t,t-k} = \gamma_{0,k}$ for all times t and lags k. Hence the covariance function $\gamma_{t,s}$ depends only on the absolute difference $|t - s|$.

Know: For a stationary process $\{Y_t\}$, write the *autocovariance function* as $\gamma_k = Cov(Y_t, Y_{t-k})$ and the *autocorrelation function* as $\rho_k = corr(Y_t, Y_{t-k})$. Note that the mean function $E(Y_t) = \mu$ and the variance function $V(Y_t) = Var(Y_t) = \gamma_0$ are constant and do not depend on t. The autocovariance function γ_k and the autocorrelation function ρ_k depend on the lag k but not on the time t.

3) Be able to show that a simple process $\{Y_t\}$ is stationary: show that $\mu_t = E(Y_t) = \mu$ is constant for all t and that $\gamma_{t,t-k} = Cov(Y_t, Y_{t-k})$ depends only on the lag k, not on time t. Note that $\gamma_{t,t-k} \equiv \sigma^2$ is allowed, as a constant function of k. See HW1 D,G.

4) For a simple stationary process, be able to compute $E(Y_t) = \mu$, $V(Y_t) = \gamma_0$, $\gamma_k = Cov(Y_t, Y_{t-k})$ and $\rho_k = corr(Y_t, Y_{t-k})$. See HW1 Gb.

5) Be able to show that a simple process $\{Y_t\}$ is not stationary: try to show that $E(Y_t)$ depends on t. If this fails, try to show that $Var(Y_t)$ depends on t. If this fails, show that $\gamma_{t,t-k} = Cov(Y_t, Y_{t-k})$ depends on t. Stop as soon as one of these functions depends on t. See HW1 E.

Know that for a stationary process $\{Y_t\}$, $\gamma_k = \gamma_{-k}$, $\rho_k = \rho_{-k}$, $\gamma_0 = Var(Y_t)$, $\rho_0 = 1$, $|\gamma_k| \leq \gamma_0$ and $|\rho_k| \leq 1$.

For a *white noise process* $\{e_t\}$, the e_t are iid (independent and identically distributed). If $E(e_t) = \mu$ and $V(e_t) = \sigma^2$ exist, then the white noise process is stationary (and strictly stationary) with $\gamma_0 = \sigma^2$ and $\gamma_k = 0$ for $k \neq 0$. **After chapter 3,** assume that for a white noise, $E(e_t) = 0$ and $V(e_t) = \sigma^2$, unless told otherwise.

6) Be able to plot a simple time series, possibly defined from computer output. See HW1 C.

The plot of a stationary time series should oscillate about the mean and the variability of sequences of similar length should be similar.

A deterministic trend has $\mu_t = g(t)$ where a linear trend is $\mu_t = a + bt$ and a quadratic trend is $\mu_t = a + bt + ct^2$. With a stochastic trend, there appears to be a nonconstant mean pattern, often linear or quadratic, when in fact there is no deterministic trend.

Know: $\overline{Y} = \frac{1}{n} \sum_{i=1}^n Y_i$ estimates $\mu = E(Y_t)$ and the sample autocorrelation function r_k estimates ρ_k for a stationary time series. Usually "sample" will be omitted.

Know: the plot of k vs r_k for $k = 1, ..., J$, where usually $10 \leq J \leq n/4$ with horizontal lines at $\pm 2/\sqrt{n}$ added to the plot, is called the autocorrelation function, ACF.

Let the "past" $\Psi_{t-1} = \{Y_{t-1}, Y_{t-2}, ...\} \cup \{e_{t-1}, e_{t-2}, ...\}$. Let the time series model be $Y_t = m_t + e_t = E(Y_t|\Psi_{t-1}) + e_t$. Then the tth fitted value $\hat{Y}_t = \hat{m}_t = \hat{E}(Y_t|\Psi_{t-1})$, and the tth residual $\hat{e}_t = Y_t - \hat{Y}_t$. For a good model, $\hat{Y}_{t+k}$ should predict Y_{t+k} better than $\overline{Y}$ for small lags k, and the residuals $\{\hat{e}_t\}$ should behave a lot like the white noise $\{e_t\}$ that the residuals from a good model estimate.

A *response plot* of $\hat{Y}_t$ vs Y_t should scatter about the identity line with unit slope and zero intercept with no other pattern if the model is adequate. The vertical deviation of Y_t from the identity line is the residual $Y_t - \hat{Y}_t = \hat{e}_t$. A *residual plot* of $\hat{Y}_t$ vs $\hat{e}_t$ or of t vs $\hat{e}_t$ should scatter about the $\hat{e}_t = 0$ line, with no other pattern if the model is adequate. The evenly spaced t tends to make the residual plot of t vs $\hat{e}_t$ to look "spikey."

In this class, a plot of a vs b means a is on the horizontal axis and b is on the vertical axis.

7) Be able to sketch a "good" response plot and residual plot.

Know: a *moving average* $MA(q)$ times series is $Y_t = \mu - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \cdots - \theta_q e_{t-q} + e_t$ where $\theta_q \neq 0$. Often $\mu = 0$.

Know: an *autoregressive* $AR(p)$ times series is $Y_t = \phi_0 + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \cdots + \phi_p Y_{t-p} + e_t$ where $\phi_p \neq 0$. Often $\phi_0 = 0$.

Know: an *autoregressive moving average* $ARMA(p, q)$ times series is
$Y_t = \tau + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \cdots + \phi_p Y_{t-p} - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \cdots - \theta_q e_{t-q} + e_t$ where $\theta_q \neq 0$ and $\phi_p \neq 0$. Often $\tau = 0$.

Want the $ARMA(p, q)$ model to be stationary and invertible. Let $Z_t = Y_t - \mu$ where $\mu = E(Y_t)$ if $\{Y_t\}$ is stationary and μ is some origin otherwise. Then stationarity implies that $Z_t = \sum_{j=1}^\infty \psi_j e_{t-j} + e_t$, which is an $MA(\infty)$ representation, where the $\psi_j \to 0$ rapidly as $j \to \infty$. Invertibility implies that $Z_t = \sum_{j=1}^\infty \pi_j Z_{t-j} + e_t$, which is an $AR(\infty)$ representation, where the $\pi_j \to 0$ rapidly as $j \to \infty$. Thus if the $ARMA(p, q)$ model is stationary and invertible, then Y_t depends almost entirely on nearby lags of Y_t and e_t, not on the distant past.

Let the complex number $W = W_1 + W_2 \quad i$ have modulus $|W| = W_1^2 + W_2^2$.

The backshift operator or lag operator B satisfies $BW_t = W_{t-1}$ and $B^j W_t = W_{t-j}$. Then the $MA(q)$ model is $Z_t = e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \cdots - \theta_q e_{t-q} = (1 - \theta_1 B - \cdots - \theta_q B^q) e_t = \theta(B) e_t$. The $AR(p)$ model is $e_t = Z_t - \phi_1 Z_{t-1} - \phi_2 Z_{t-2} - \cdots - \phi_p Z_{t-p} = (1 - \phi_1 B - \cdots - \phi_p B^p) Z_t$, or $\phi(B) Z_t = e_t$. The $ARMA(p,q)$ model is $Z_t - \phi_1 Z_{t-1} - \phi_2 Z_{t-2} - \cdots - \phi_p Z_{t-p} = e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \cdots - \theta_q e_{t-q}$, or $\phi(B) Z_t = \theta(B) e_t$. Consider $\theta(B)$ and $\phi(B)$ as polynomials in B.

An $ARMA(p,q)$ model is invertible if all of the roots of the polynomial $\theta(B) = 0$ have modulus > 1, and stationary if all of the roots of the polynomial $\phi(B) = 0$ have modulus > 1. Hence the roots of both polynomials lie outside the unit circle. An $AR(p)$ model is always invertible and an $MA(q)$ model is always stationary. For the $AR(1)$ model, need $|\phi_1| < 1$. For the $MA(1)$ model, need $|\theta_1| < 1$. For the $ARMA(1,1)$ model, need $|\phi_1| < 1$ and $|\theta_1| < 1$.

For an $AR(2)$ model, want $\phi_2 \pm \phi_1 < 1$ and $|\phi_2| < 1$. For an $MA(2)$ model, want $\theta_2 \pm \theta_1 < 1$ and $|\theta_2| < 1$. For an ARMA(2,2) model want the conditions for the $AR(2)$ and $MA(2)$ model to both hold. For an $ARMA(1,2)$ model, want $|\phi_1| < 1$ and the conditions for the $MA(2)$ model to hold. For an $ARMA(2,1)$ model, want $|\theta_1| < 1$ and the conditions for the $AR(2)$ model to hold.

Let τ_i stand for θ_i or ϕ_i. Let k stand for q or p, and let $\psi(B) = 1 - \tau_1 B - \tau_2 B^2 - \cdots - \tau_k B^k$ stand for $\phi(B)$ or $\theta(B)$. A necessary but not sufficient condition for the roots of $\psi(B) = 0$ to all be greater than 1 in modulus is $\tau_1 + \cdots + \tau_k < 1$ and $|\tau_k| < 1$.

Consider the sequence of AR models

$AR(1): Y_t = \phi_{01} + \phi_{11} Y_{t-1} + e_t,$

$AR(2): Y_t = \phi_{02} + \phi_{12} Y_{t-1} + \phi_{22} Y_{t-2} + e_t,$

$\vdots$

$AR(k): Y_t = \phi_{0k} + \phi_{1k} Y_{t-1} + \cdots + \phi_{kk} Y_{t-k} + e_t.$

Know: the population partial autocorrelation function (PACF) is a plot of k vs ϕ_{kk} and the sample PACF is a plot of k vs $\hat{\phi}_{kk}$ for $k = 1, 2, ..., J$ where $10 \leq J \leq n/4$ with horizontal lines at $\pm 2/\sqrt{n}$ added to the plot.

8) **Know for final**: Given the ACF and PACF for several time series, be able to identify $AR(p)$ and $MA(q)$ models. If only one of the several time series is an $ARMA(p,q)$ model, be able to pick it out. See HW2 E) and Q2.

Tips: Assume the model is stationary and invertible. Then the $AR(p)$ model is an $MA(\infty)$ model, the $MA(q)$ model is an $AR(\infty)$ model, and the $ARMA(p,q)$ model is an $MA(\infty)$ model and an $AR(\infty)$ model.

a) For an $AR(p)$ time series, the population ACF decays exponentially or like a damped sinusoidal to 0 rapidly. The population PACF has a spike at lag p and usually at lags 1 to $p-1$. The ϕ_{kk} are 0 for $k > p$, so the PACF cuts off to 0 after lag p.

b) For an $MA(q)$ time series, the population ACF has a spike at lag q and usually at lags 1 to $q-1$. The ρ_k are 0 for $k > q$, so the ACF cuts off to 0 after lag q. The population PACF decays exponentially or like a damped sinusoidal to 0 rapidly.

c) For an $ARMA(p,q)$ time series, the ACF and PACF decay exponentially or like a damped sinusoidal to 0 rapidly. For the ACF there may be spikes at lags up to

$\min(0, q-p)$ and the quick decay starts for $k > \min(0, q-p)$. For the PACF there may be spikes at lags up to $\min(0, p-q)$ and the quick decay starts for $k > \min(0, p-q)$.

For a white noise, $se(r_k) = se(\hat{\phi}_{kk}) = 1/\sqrt{n}$ for $k \geq 1$, so the horizontal lines at $\pm 2/\sqrt{n}$ are at $\pm 2se$ and act like pointwise 95% confidence intervals (CIs) if the time series is a white noise. If the time series is a white noise, about 95% of the values will be between the horizontal lines, but about 5% will be outside.

Warning: R starts the ACF at lag 0 where $\hat{\rho}_0 = \rho_0 = 1$. Ignore this 0 lag when identifying a model for the time series.

Output in symbols and for a data set are shown below.

Label	coef	SE	LCI	UCI	Z	p-value
ar1	$\hat{\phi}_1$	$se(\hat{\phi}_1)$	95% CI	for ϕ_1	$Z_{ar,1} = \hat{\phi}_1/se(\hat{\phi}_1)$	for Ho: $\phi_1 = 0$
$\vdots$						
arp	$\hat{\phi}_p$	$se(\hat{\phi}_p)$	95% CI	for ϕ_p	$Z_{ar,p} = \hat{\phi}_p/se(\hat{\phi}_p)$	for Ho: $\phi_p = 0$
ma1	$\hat{\theta}_1$	$se(\hat{\theta}_1)$	95% CI	for θ_1	$Z_{ma,1} = \hat{\theta}_1/se(\hat{\theta}_1)$	for Ho: $\theta_1 = 0$
$\vdots$						
maq	$\hat{\theta}_q$	$se(\hat{\theta}_q)$	95% CI	for θ_q	$Z_{ma,q} = \hat{\theta}_q/se(\hat{\theta}_q)$	for Ho: $\theta_q = 0$
intercept	$\hat{\mu}_Y$	$se(\hat{\mu}_Y)$	95% CI	for μ_Y	$Z_o = \hat{\mu}_Y/se(\hat{\mu}_Y)$	for Ho: $\mu_Y = 0$

```
                coef        se         LCI        UCI          Z         pval
ar1        0.8912619 0.1929058  0.50545023  1.2770735  4.6201918 3.833855e-06
ar2       -0.4617117 0.1192590 -0.70022956 -0.2231938 -3.8715053 1.081653e-04
ma1       -0.1006720 0.2020894 -0.50485084  0.3035068 -0.4981559 6.183742e-01
ma2        0.1817279 0.1152152 -0.04870248  0.4121583  1.5772912 1.147285e-01
intercept  0.1786932 0.1242312 -0.06976925  0.4271556  1.4383920 1.503229e-01
```

9) **Know for final:** Let τ_k be μ_Y, ϕ_k or θ_k. The large sample 95% CI for τ is given by (LCI, UCI). Know how to get the 95% CI from output. The coef for ark is $\hat{\phi}_k$ and the coef for mak is $\hat{\theta}_k$, where for R the model is $Y_t = \tau + \phi_1 Y_{t-1} + \cdots + \phi_p Y_{t-p} + \theta_1 e_{t-1} + \cdots + \theta_q e_{t-q} + e_t$. For an AR(p) model, $\tau = \phi_0$ and for an MA(q) model, $\tau = \mu$. **Warning:** the $\hat{\theta}_k$ in R is $-\hat{\theta}_k$ in the book. Want the 95% CI for θ_q and for ϕ_p to be inside of $(-1, 1)$. The above output is for an $ARMA(2,2)$ model, the 95% CI for ϕ_2 is $(-0.700, -0.223)$ and the 95% CI for θ_2 is $(-0.049, 0.412)$. See HW3 Ag) and Q3.

10) **Know for final:** Let τ_k be μ_Y, ϕ_k or θ_k. See HW3 Ah) and Q3. The 4 step test of hypotheses for $Ho: \tau_k = 0$ is i) State the hypotheses Ho: $\tau_k = 0$ Ha: $\tau_k \neq 0$.
ii) Find the test statistic $Z = \hat{\tau}_k / se(\hat{\tau}_k)$ from output.
iii) Find pval = the estimated p–value from output.
iv) State whether you reject Ho or fail to reject Ho and give a conclusion. Reject Ho if pval $\leq \alpha$ and fail to reject Ho if pval $> \alpha$. **Use $\alpha = 0.05$ if α is not given.** For $\tau_k = \phi_k$ (θ_k), conclude Y_{t-k} (e_{t-k}) is needed in the model if Ho is rejected. Conclude Y_{t-k} (e_{t-k}) is not needed in the model given the other terms are in the model if you fail to reject Ho. If $\tau_k = \mu_Y$, the constant is needed in the model if Ho is rejected. The constant is not needed in the model if you fail to reject Ho. Note: $|Z| \geq 2$ will have pval < 0.05.