Please note that these are practice questions The actual final exam questions will be different.

[1] Let $(Z_t \colon t \in \mathbb{Z})$ be a sequence of independent zero mean normal random variables with variance σ^2 and let a, b and c be constants. Which of the following processes are weakly stationary? For each weakly stationary process specify the mean and the ACVF.

(a) $X_t = a + bZ_t + cZ_{t-1}$;

(b) $X_t = Z_t \cos(ct) + Z_{t-1} \sin(ct)$;

(c) $X_t = a + bZ_0$;

(d) $X_t = Z_t Z_{t-1}$.

[2] Let $(X_t \colon t \in \mathbb{Z})$ be a stochastic process given by the equations

$$X_t = b_0 + b_1 t + Z_t, \qquad t \in \mathbb{Z},$$

where b_0, b_1 are fixed constants and $(Z_t \colon t \in \mathbb{Z}) \sim \mathrm{IID}(0, \sigma^2)$.

(a) Show that $(X_t \colon t \in \mathbb{Z})$ is a nonstationary process.
(b) Show that the differenced sequence $(\nabla X_t \colon t \in \mathbb{Z})$ is weakly stationary. Compute its ACVF $\gamma_{\nabla X}$.
(c) Repeat part (b) if $(Z_t \colon t \in \mathbb{Z})$ is replaced by a general weakly stationary process $(Y_t \colon t \in \mathbb{Z})$ with mean μ and ACVF γ_Y.

[3] Download the file `jj.dat` from the course website. It contains 84 quarterly earnings per share for the U.S. company Johnson & Johnson collected from 1960 until 1980. Interpret these values as observations of the random variables

$$V_t = \text{earnings per share in quarter } t, \qquad t = 1, \ldots, 84,$$

and set $X_t = \ln V_t$, $t = 1, \ldots, 84$.

(a) Plot the data.
(b) Fit the regression model

$$X_t = bt + a_1 Q_1(t) + a_2 Q_2(t) + a_3 Q_3(t) + a_4 Q_4(t) + Z_t \tag{1}$$

to the data, where $Q_i(t) = 1$ if time t corresponds to quarter $i = 1, 2, 3, 4$, and zero otherwise.
(c) What is the interpretation of the parameters $b, a_1, \ldots, a_4$? What happens if you include an intercept term into the model specified by (1)?

(d) Plot the residuals

$$\hat{Z}_t = X_t - \hat{b}t + \hat{a}_1 Q_1(t) + \hat{a}_2 Q_2(t) + \hat{a}_3 Q_3(t) + \hat{a}_4 Q_4(t).$$

Comment on the quality of the fitted model.

[4] Consider the two-sided moving average filter

$$W_t = \frac{1}{2q+1} \sum_{j=-q}^{q} X_{t+j}, \qquad t \in \mathbb{Z},$$

where $q \in \mathbb{N}_0$, and apply it to a stochastic process $(X_t \colon t \in \mathbb{Z})$ given by letting $X_t = m_t + Y_t$.

(a) If $m_t = b_0 + b_1 t$, show that an application of W_t passes m_t undistorted.

(b) If, moreover, $(Y_t \colon t \in \mathbb{Z})$ is an independent sequence with mean zero and variance σ^2, compute $E[W_t]$ and $\mathrm{Var}(W_t)$. Compare these quantities with $E[X_t]$ and $\mathrm{Var}(X_t)$.

[5] Simulate 200 values $x_1, \ldots, x_{200}$ of the stochastic process

$$X_t = \sin\left(\frac{t}{20}\right) + Z_t, \qquad t \in \mathbb{Z},$$

where $(Z_t \colon t \in \mathbb{Z})$ is a sequence of independent identically distributed random variables with mean zero and variance .25. In all tests below use a significance level $\alpha = .05$.

(a) Plot the simulated data.

(b) Check the data for whiteness using the sample ACF (Method 1).

(c) Check the data for whiteness using a version of the Portmanteau test (Method 2). In R, you may apply the command `Box.test` with `type="Ljung"` and `lag=20`.

(d) Check the data for normality using the qq plot (Method 4). The critical value for the test statistic R^2 can be obtained from $P(R^2 < .987) = .05$.

[6] Download the files `prob6a.dat` and `prob6b.dat` from the course website. Each file contains 200 simulated values obtained from two different stochastic processes.

(a) Plot both data sets.

(b) Utilize qq plots to decide which of the two data sets is obtained from a normal distribution (visual inspection of the plots suffices).

(c) Estimate the mean and variance of both processes. Is there correlation in the data?

[7] Download the file `prob7.dat` from the course website. It contains 100 simulated values of some stochastic process.

(a) Use the rank test (Method 3) to decide whether there is a linear trend in the data.
(b) Remove the trend and test the resulting residuals for whiteness with a method of your choice.

[8] Let $(X_t \colon t \in \mathbb{Z})$ be a stochastic process given by the equations

$$X_t = 3 + Z_t + .5Z_{t-1}, \qquad t \in \mathbb{Z},$$

where $(Z_t \colon t \in \mathbb{Z})$ are independent, identically distributed normal random variables with zero mean and unit variance.

(a) Show that $(X_t \colon t \in \mathbb{Z})$ is strictly stationary.
(b) Construct a 95% confidence interval for $E[X_t]$.
(c) Simulate 200 values of $(X_t \colon t \in \mathbb{Z})$ with the R command

```
arima.sim(list(order=c(0,0,1), ma=.5), n=200)
```

and compute the sample mean $\bar{x}$. (Don't forget to incorporate that the mean is equal to 3.)
(d) Fit a model to the mean-corrected data $y_t = x_t - \bar{x}$, $t = 1, \ldots, 200$, with the R command `arima(yt, order=c(0,0,1))`.
(e) Use the estimated coefficients to calculate the 95% confidence interval for the simulated data.

[9] Reconsider the stochastic process $(X_t \colon t \in \mathbb{Z})$ from Problem 8.

(a) Compute the ACF $\rho(h)$ of $(X_t \colon t \in \mathbb{Z})$.
(b) Compute a 95% confidence interval for $\rho(3)$ using the theoretical values.

[10] Identify the following models as ARMA(p, q) models, and determine whether they are causal and/or invertible:

(a) $X_t = .80X_{t-1} - .15X_{t-2} + Z_t - .30Z_{t-1}$.
(b) $X_t = X_{t-1} - .50X_{t-2} + Z_t - Z_{t-1}$.

[11] Let $(X_t)_{t \in \mathbb{N}_0}$ be a stochastic process given by the equations

$$X_0 = Z_0, \qquad X_t = \phi X_{t-1} + Z_t, \quad t \in \mathbb{N},$$

where $(Z_t \colon t \in \mathbb{N}_0) \sim \text{WN}(0, \sigma^2)$.

(a) Find the mean and the variance of X_t, $t \in \mathbb{N}_0$. Is the process weakly stationary?
(b) Show that

$$\text{Corr}(X_{t+h}, X_t) = \phi^h \sqrt{\frac{\text{Var}(X_t)}{\text{Var}(X_{t+h})}}, \qquad h \in \mathbb{N}_0.$$

[12] Download the file `recruit.dat` from the course website. It contains the 453 data points of monthly recruitment discussed in Example 3.3.5.

(a) Use the R command `predict` to forecast the recruitment for the next 24 months;

(b) Use only the last 100 observations to refit the AR(2) model and to forecast the next 24 months of recruitment;

(c) Compare the results.

[13] Download the file `soi.dat` from the course website. Analyze the data repeating the steps performed in Example 3.3.5 and the previous problem.

[14] Generate $n = 500$ observations from the ARMA model

$$X_t = .9X_{t-1} + Z_t - .9Z_{t-1}, \qquad t \in \mathbb{Z},$$

where $(Z_t \colon t \in \mathbb{Z})$ denotes a sequence of independent standard normal random variables.

(a) Plot the simulated data and compute the sample ACF and PACF.

(b) Fit an ARMA(1,1) model to the data. What happened and how do you explain the results?

[15] Consider Wölfer's sunspot numbers which can be found as `sunspots.dat` on the course webpage.

(a) Assuming an AR(2) model, find the Yule-Walker estimates of ϕ_1, ϕ_2 and σ^2.

(b) Provide 95% confidence intervals for ϕ_1 and ϕ_2.

[16] Consider the cardiovascular mortality series `mort.dat` discussed in Problem 18.

(a) Fit an AR(2) model to the data using Yule-Walker and compare the parameter estimates with the ones obtained in Problem 18.

(b) Compare the corresponding standard errors.

[17] The model fitted to a data set $x_1, \ldots, x_{100}$ is $X_t + .4X_{t-1} = Z_t$, where the Z_t are uncorrelated with zero mean and unit variance. The sample ACF and PACF of the residuals are given in the table below.

(a) Are these values compatible with whiteness of the residuals?

(b) If not, suggest a better model for the X_t and give estimates of the parameters.

Lag	1	2	3	4	5	6	7	8	9	10	11	12
ACF	.799	.412	.025	−.228	−.316	−.287	−.198	−.111	−.056	−.009	.048	.133
PACF	.799	−.625	−.044	.038	−.020	−.077	−.007	−.061	−.042	.089	.052	.125

[18] Let $(X_t\colon t \in \mathbb{Z})$ be a first-order autoregressive process with autoregressive parameter $\phi = .8$.

(a) Compute and plot the spectral density $f_X(\omega)$ of $(X_t\colon t \in \mathbb{Z})$.
(b) Generate an AR(1) process of length $n = 200$ in R and compute the periodogram.
(c) Repeat the previous steps with $\phi = -.8$.

[19] Download the file `sunspots.dat` from the course website. It contains the sunspot numbers recorded for the years 1770–1869.

(a) Perform a periodogram analysis of the sunspot data identifying the predominant periods.
(b) Obtain confidence intervals for the identified periods and interpret your findings.

[20] Use the spectral densities derived in Examples 4.2.1–4.2.3 to compute the ACVF of

(a) a white noise process $(Z_t\colon t \in \mathbb{Z})$;
(b) the moving average process $(X_t\colon t \in \mathbb{Z})$ given by $X_t = \frac{1}{2}(Z_t + Z_{t-1})$;
(c) an AR(2) process.