# Problem Set 8

## Statistics 509 – Winter 2022

## Due to Canvas on Wednesday, March 30 by Midnight

**Instructions.** You may work in teams, but you must turn in your own work/code/results. Also for the problems requiring use of the R-package, you need to include a copy of your R-code. This provides us a way to give partial credit in case the answers are not totally correct.

**1.** Suppose have $Y$ is random variable taking values/probabilities according to the following table.

| Values | Probabilities |
|---|---|
| -2 | $\frac{1}{6}$ |
| 0 | $\frac{1}{6}$ |
| 1 | $\frac{1}{6}$ |
| 2 | $\frac{1}{2}$ |

and $X$ is random variable defined as $X = |Y|$.

**(a)** Derive the best predictor of $Y$ based on $X$, and give the MSPE for this predictor.

**(b)** Derive the best linear predictor of $Y$ based on $X$, and give the MSPE for this predictor.

**(c)** Explain why the MSPE in **(b)** is larger than that given in **(a)**.

**2.** Problem 7 on page 357 in Ruppert/Matteson, with the revision for part **(c)** below.

**(c)** Suppose $\rho(1) = .3,\ \rho(2) = .2$. Find $\phi_1, \phi_2, \rho(3)$.

*Hint* For non-singular matrix $A$, can generate the inverse of $A$ via command $solve(A)$.

**3.** For this problem there is a data set consisting of Russell 2000 weekly data over 4 years from 2015 to 2019 with the name RUT_03_2015-03_2019 – it is found in the Data folder on Canvas.

**(a)** Generate plot of the sample auto-correlation of the Russell 2000 weekly returns, and discuss what is observed from these plots. Also, carry out a Box-Ljung test on whether the auto-correlations are all zero out to lag 10, and summarize your results

**(b)** Derive the optimal AR model utilizing the ar function in R. Provide a comprehensive set of diagnostics and give a clear summary of findings/conclusions.

**(c)** Utilizing output from **(b)**, do a prediction out to 12 weeks ahead along with 95% confidence intervals – provide numerical values along with a plot of predictions/confidence intervals.

**(d)** Repeat **(b)** for the optimal ARMA model. Provide a comprehensive set of diagnostics and give a clear summary of findings/conclusions.

**(e)** Utilizing output from **(d)**, do a prediction out to 12 weeks ahead along with 95% confidence intervals – provide numerical values along with a plot of predictions/confidence intervals.