[STAT 4540] Exam-2

Michael Ghattas

4/3/2022

Problem 1

```r
library(itsmr)

df <- load("/Users/Home/Documents/Michael_Ghattas/School/CU_Boulder/2022/Spring 2022/STAT - 4540/Exams/2/Exam2data.RData");
```

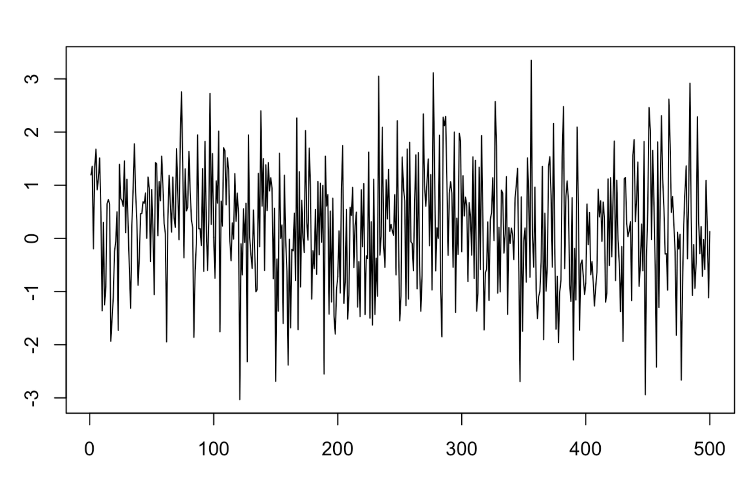
```r
dat <- data.frame(x = x)
x = dat[, 1]

mean(x)
```

```
## [1] 0.2005136
```

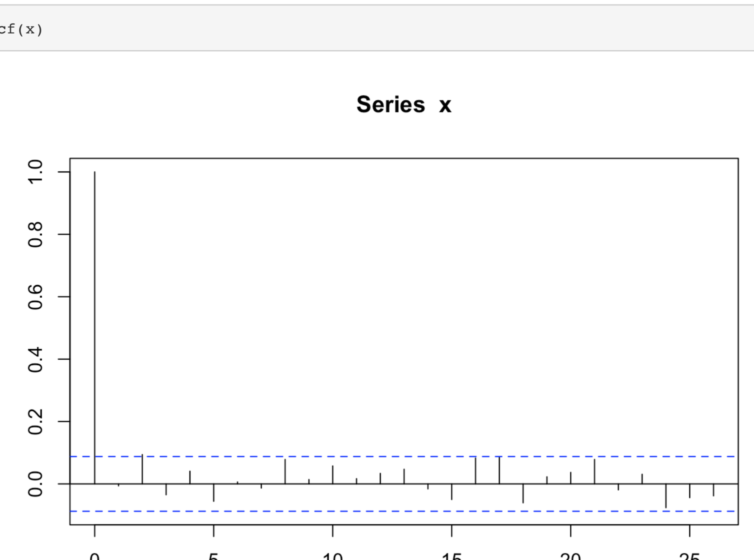
```r
sd(x)
```

```
## [1] 1.13315
```

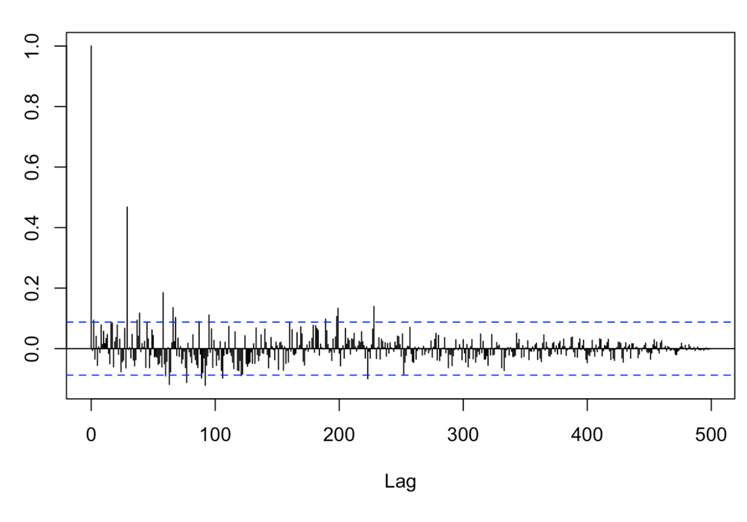
```r
plot(x)
```



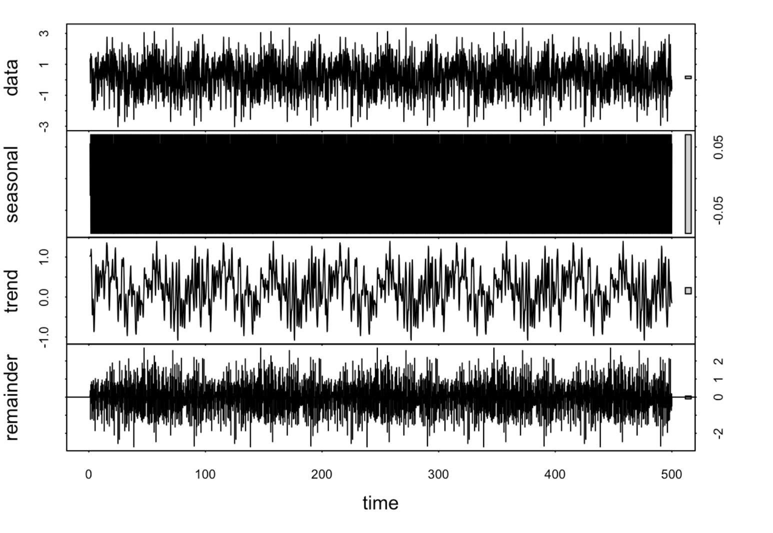
```r
acf(x)
```



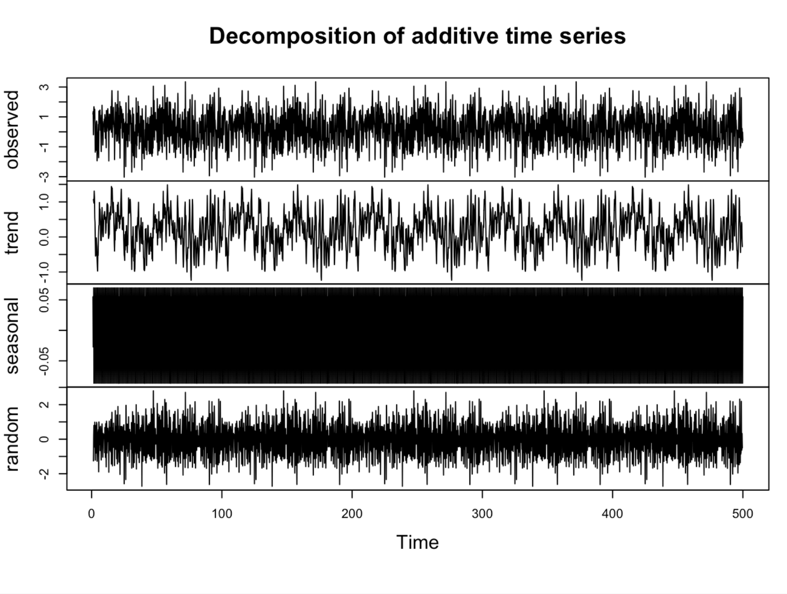
```r
acf(x, lag.max = 500)
```



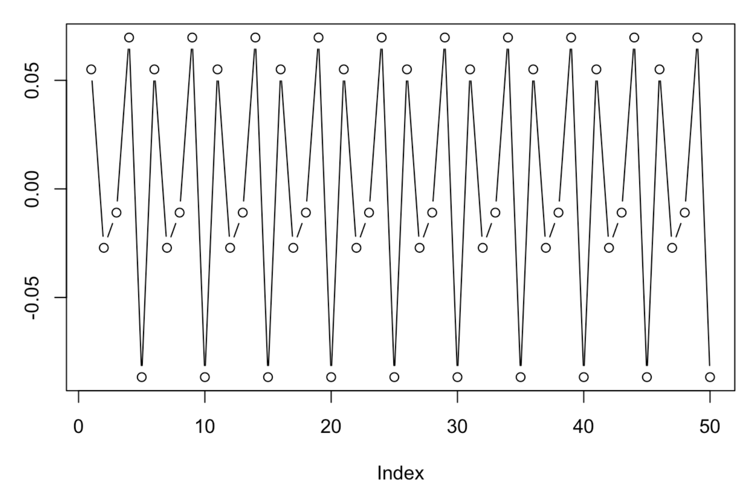
```r
X <- ts(data = dat, start = c(1,1), end = c(500,1), frequency = 5)
out <- decompose(x = X, type = "additive")
decomposed <- stl( X, s.window = "periodic" )

plot(decomposed)
```



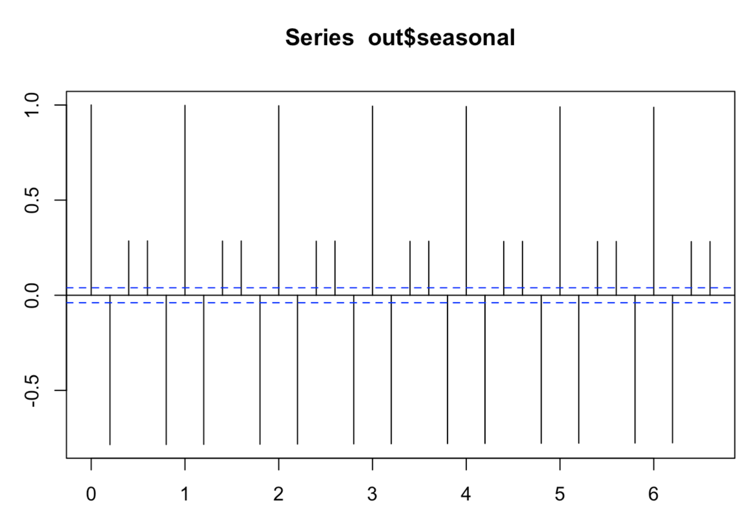
```r
plot(out)
```



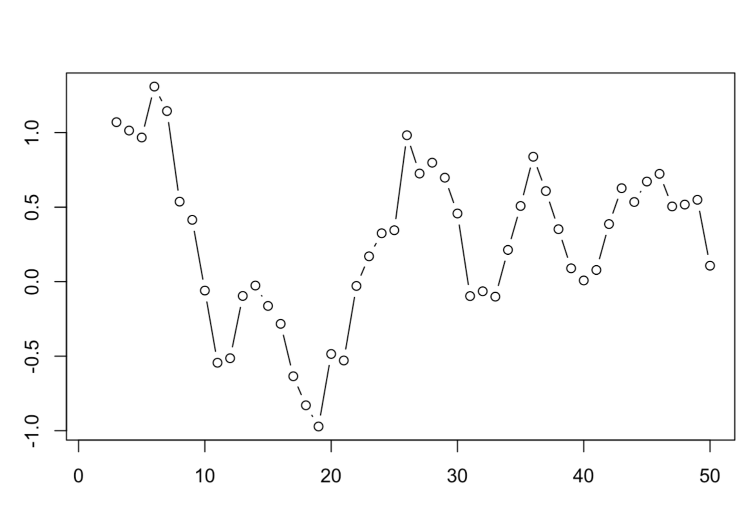
```r
plot(out$seasonal[1:50], type = "b")
```



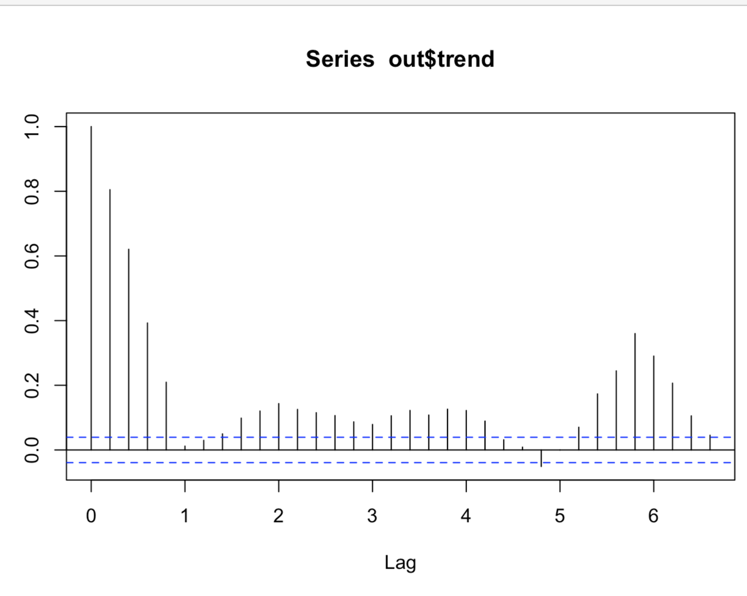
```r
acf(out$seasonal, na.action = na.pass)
```



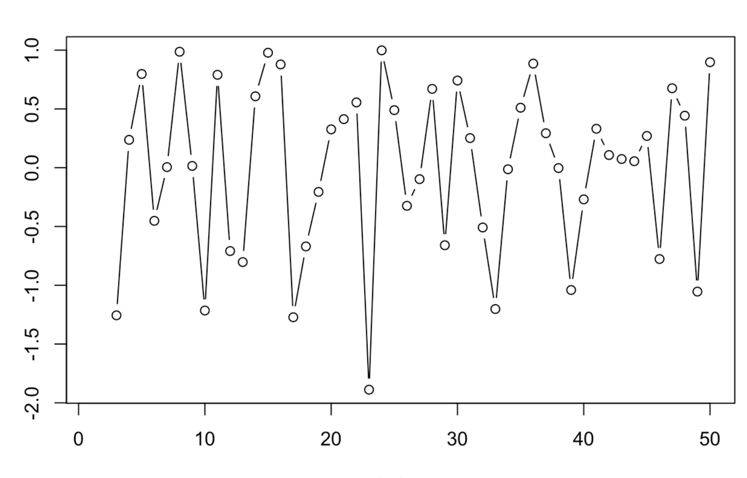
```r
plot(out$trend[1:50], type = "b")
```



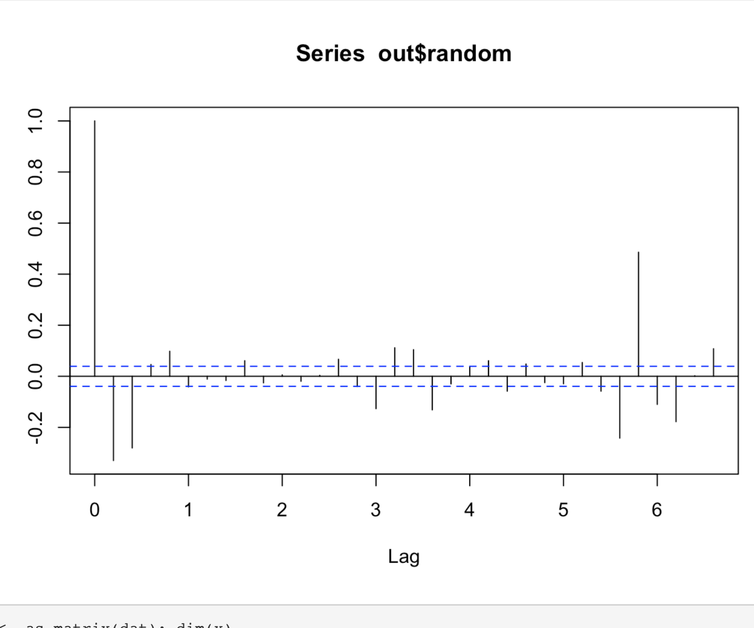
```r
acf(out$trend, na.action = na.pass)
```



```r
plot(out$random[1:50], type = "b")
```



```r
acf(out$random, na.action = na.pass)
```



```r
x<- as.matrix(dat); dim(x)
```

```
## [1] 500   1
```

```r
cos.c <- rep(cos(3 * pi * 1:500 / 500), 1)
c1 <- as.matrix(cos.c); dim(c1)
```

```
## [1] 500   1
```

```r
sin.c <- rep(sin(3 * pi * 1:500 / 500), 1)
s1 <- as.matrix(sin.c); dim(s1)
```

```
## [1] 500   1
```

```r
cos.c2 <- rep(cos(5 * pi * 1:500 / 500), 1)
c2 <- as.matrix(cos.c2); dim(c2)
```

```
## [1] 500   1
```

```r
sin.c2 <- rep(sin(5 * pi * 1:500 / 500), 1)
s2 <- as.matrix(sin.c2); dim(s2)
```

```
## [1] 500   1
```

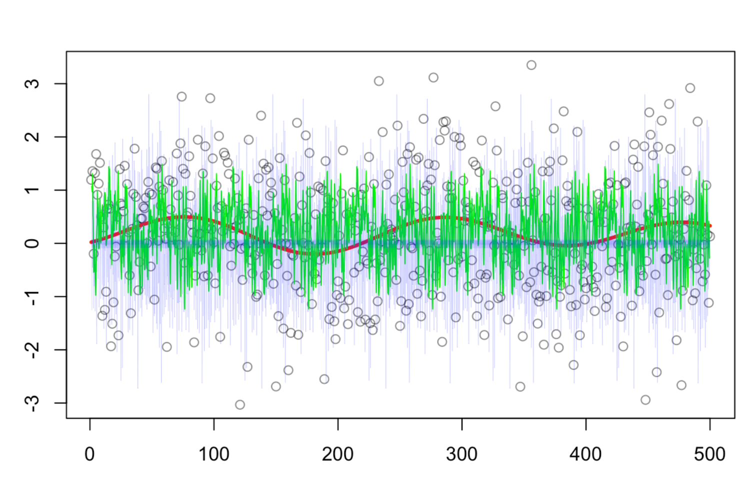
```r
data <- data.frame(cbind(x, c1, s1, c2, s2))

fit <- lm(x ~ c1 + s1 + c2 + s2, data = data)

summary(fit)
```

```
##
## Call:
## lm(formula = x ~ c1 + s1 + c2 + s2, data = data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -3.2632 -0.7736 -0.0321  0.8161  3.2930
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.17386    0.05336   3.258 0.001199 **
## c1           0.08219    0.07069   1.163 0.245543
## s1           0.02752    0.07423   0.371 0.711028
## c2          -0.23885    0.07069  -3.379 0.000786 ***
## s2           0.16104    0.07199   2.237 0.025727 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.118 on 495 degrees of freedom
## Multiple R-squared:  0.03475,    Adjusted R-squared:  0.02695
## F-statistic: 4.456 on 4 and 495 DF,  p-value: 0.001514
```

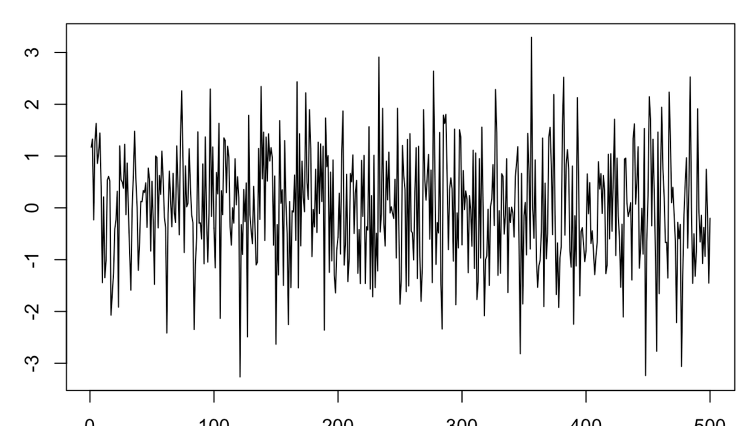
```r
coef(fit)
```

```
## (Intercept)          c1          s1          c2          s2
##  0.17385832  0.08219102  0.02751739 -0.23884794  0.16104127
```

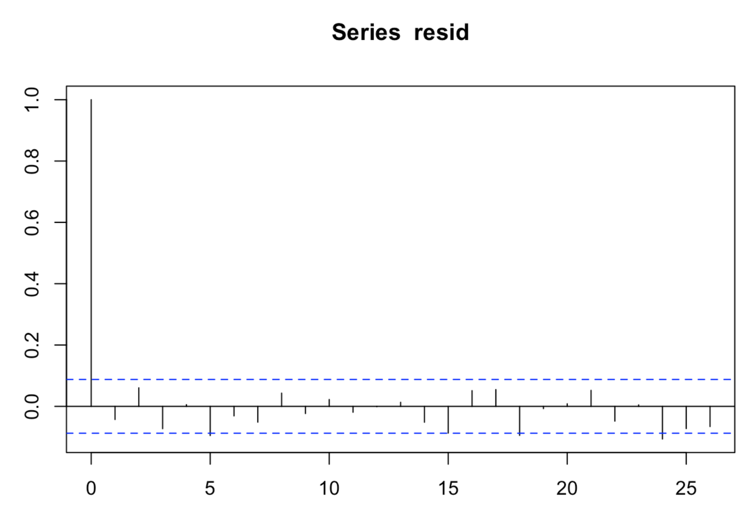
```r
plot(x, type = "p", col = gray(0, .5))
lines(fit$fitted, col = "red", lwd = 3)
lines(out$trend, col = "green", lwd = 1)
lines(out$seasonal, col = "blue", lwd = 0.1)
lines(out$random, col = "blue", lwd = 0.1)
```



```r
resid <- fit$resid
plot(resid, type = "l")
```



```r
acf(resid)
```



I do not believe the data is white noise, as we can see that there is correlation in the points once we set the max lag to 500. Additionally the data seems to indicate that there is a none stationary trend and seasonality imbeded within the random part of the data. Once the trend/seasonal elements have been removed, further examination of the residuals is needed.

problem 2:

a) Assuming a stationary μ, γ^{ACVF} we choose the criteria: $\min \mathbb{E}\,Loss$, i.e. $L(x,y) = (x-g)^2$. If $X = X_{n+h}$ and $Y = x_1, \ldots, x_n$, where the minimizer g of $\mathbb{E}(X-g(y))^2$ is $g(Y) = \mathbb{E}(X|Y)$, and $\mathbb{E}[\vec{X}|\vec{Y}] = \vec{\mu}_X + \Sigma_{XY}\Sigma_{YY}^{-1}(\vec{Y} - \vec{\mu}_X)$. We also concluded that the best linear predictor of X_{n+h} is $(P_n X_{n+h}) = a_0 + \sum_{i=1}^{n} a_j X_{n+1-j}$. with $P_n X_{n+h} = \vec{a}^T \vec{X}_n = \gamma_n(h)^T \Gamma_n^{-1} \vec{X}_n$. We looked at the Durbin-Levinson model, the Innovations algorithm, and the Infinite Past approach, MLE, and YW. We also analyzed how $\tilde{P}_n X_{n+h} = \lim_{m \to \infty} P_{m,n} X_{n+h}$, where $P_{n,m} X_{n+h}$ is our BLE.

problem 3:

a) $$f(\lambda) = \frac{1}{2\pi}\sum_{h=-2}^{2} e^{-ih}\gamma(h) = \frac{1}{2\pi}\left(\gamma(2)e^{i2\lambda} + \gamma(1)e^{i\lambda} + \gamma(0) + \gamma(1)e^{-i\lambda} + \gamma(2)e^{-i2\lambda}\right)$$

$$= \frac{1}{2\pi}\left[\gamma(0) + \gamma(1)(e^{i\lambda} + e^{-i\lambda}) + \gamma(2)(e^{i2\lambda} + e^{-i2\lambda}\right)$$

$$= \boxed{\frac{1}{2\pi}\left[\gamma(0) + \gamma(1)\cos(\lambda) + \gamma(2)\cos(2\lambda)\right]}$$

b)

(i)

- $\gamma_Z(t+h,t) = Cov(Z_{t+h}, Z_t) = Cov(X_{t+h} + Y_{t+h}, X_t + Y_t)$
  $= Cov(X_{t+h}, X_t) + Cov(X_{\cancel{t+h}}, Y_t) + Cov(Y_{\cancel{t+h}}, X_t) + Cov(Y_{t+h}, Y_t)$
  $= Cov(X_{t+h}, X_t) + Cov(Y_{t+h}, Y_t) = \gamma_X(t+h,t) + \gamma_Y(t+h,t)$
- $\gamma_Z(h) = \boxed{\gamma_X(h) + \gamma_Y(h)}$

(ii)

- $\gamma_Z(h) = \gamma_X(h) + \gamma_Y(h)$
- $\gamma_Z(h) = \int_{-\pi}^{\pi} e^{ih\lambda}\, dF_Z(\lambda)$

$$\gamma_Z(h) = \int_{-\pi}^{\pi} e^{ih\lambda}\, dF_X(\lambda) \quad \int_{-\pi}^{\pi} e^{ih\lambda}\, dF_Y(\lambda)$$
$$= \int_{-\pi}^{\pi} e^{ih\lambda}\left(dF_X(\lambda) + dF_Y(\lambda)\right)$$

- $dF_Z(\lambda) = dF_X(\lambda) + dF_Y(\lambda)$
- $F_Z(\lambda) = \int_{-\pi}^{\lambda} dF_Z(\nu) = \int_{-\pi}^{\lambda}\left(dF_Z(\nu) + dF_Y(\nu)\right)$
  $= \boxed{F_X(\lambda) + F_Y(\lambda)}$

(iii) Yes! As now we have to consider Y_{t-1} instead of Y_t.

Problem 4:

- AR(1) $\Rightarrow$ $X_t = 0.4996\, X_{t-1} + Z_t$
- MA(4) $\Rightarrow$ $X_t = 0.4996 Z_{t-1} + 0.2524 Z_{t-2} + 0.1292 Z_{t-3} + 0.0663_{t-4}$

While it is hard to deduce the better model from the variance, MLE, or AIC which is the better model, the model output of the MA(4)/ARMA(0,4) process has coefficients that seem to sum to a higher level of significance than the AR(1)/ARMA(1,0) process, thus indicating $\boxed{\text{ARMA(0,4)}}$ is a better fit as a model. Additionally we can note that the PACF plot would have a value $= 0$ at lags > 1 for AR(1), though we can see that at lags $\{4, 5, 6\}$ the PACF > 0.

By the CU Honor Code I agree:

Michael Ghattas
Apr/3rd/2022