Final Exam

1.

1.1.

Expectation:

$$E\left(X_{T+1} \mid \{X_t\}_{t=1}^{T}\right) = E\left(0.2 \cdot X_T + \eta_{T+1}\right) = 0.2 \cdot E\left(X_T\right) + E\left(\eta_{T+1}\right)$$

Some observations:

- X_T is in our time series, so it is not a random variable but rather a value. The expectation of a value is itself.
- η_{T+1} is taken from a normal distribution with a mean of 0, so we expect it to be 0 within some variance.

$$E\left(X_{T+1} \mid \ldots\right) = 0.2 \cdot E\left(X_T\right) + E\left(\eta_{T+1}\right) = 0.2 \cdot X_T$$

1.2.

Constructing a 95% confidence interval around the point at X_{T+1}:

$$\mathrm{CI}_1 \Leftrightarrow E\left(X_{T+1} \mid \ldots\right) \pm 95\% \cdot \sqrt{\frac{V\left(X_{T+1}\right)}{T+1}}$$

$$V\left(X_{T+1} \mid \ldots\right) = V\left(0.2 \cdot X_T + \eta_{T+1}\right) = 0.4 \cdot V\left(X_T\right) + V\left(\eta_{T+1}\right) + 2 \cdot coV\left(X_T, \eta_{T+1}\right)$$

Some observations:

- X_T is in our time series, so it is not a random variable but rather a value. The variance of a value is 0.
- η_{T+1} is a normal distribution, its variance is a known quantity. It is σ^2. In this case 0.25^2.
- Given the two above, $coV\left(X_T, \eta_{T+1}\right) = 0$.

$$0.4 \cdot V\left(X_T\right) + V\left(\eta_{T+1}\right) + coV\left(X_T, \eta_{T+1}\right) = 0.4 \cdot 0 + 0.25^2 + 2 \cdot 0 = 0.0625$$

$$E\left(X_{T+1} \mid \ldots\right) \pm 95\% \cdot \sqrt{\frac{V\left(X_{T+1}\right)}{T+1}} = 0.2 \cdot X_T \pm 95\% \cdot \sqrt{\frac{0.0625}{T+1}} = 0.2 \cdot X_T \pm 95\% \cdot \frac{0.25}{\sqrt{T+1}}$$

1.3.

We would need to know the variance of that value 0.2, which i'll call γ_0. As it stands we do not need to know that because the assumption we have made is $V\left(\gamma_0\right) \approx 0$ within an acceptable tolerance.

1.4.

Expectation:

$$E\left(Y_{T+2} \mid \{X_t\}_{t=1}^T, \{Y_t\}_{t=1}^T\right) = E\left(0.5 \cdot Y_{T+1} + 0.7 \cdot X_{T+1} + \epsilon_{T+2} - 0.2 \cdot \epsilon_{T+1}\right)$$
$$\cdots = 0.5 \cdot E\left(Y_{T+1}\right) + 0.7 \cdot E\left(X_{T+1}\right) + E\left(\epsilon_{T+2}\right) - 0.2 \cdot E\left(\epsilon_{T+1}\right)$$

Some observations:

- X_{T+1} is known.
- ϵ_{T+1} and ϵ_{T+2} are taken from a normal distribution with a mean of 0, so we expect them to be 0 within some variance.
- Y_{T+1} is a recursive call.

$$\cdots = 0.5 \cdot E\left(0.5 \cdot Y_T + 0.7 \cdot X_T + \epsilon_{T+1} - 0.2 \cdot \epsilon_T\right) + 0.14 \cdot X_T$$
$$\cdots = 0.25 \cdot E\left(Y_T\right) + 0.7 \cdot E\left(X_T\right) + E\left(\epsilon_{T+1}\right) - 0.2 \cdot E\left(\epsilon_T\right) + 0.14 \cdot X_T$$

Some more observations:

- X_T and Y_T are in the time series, and as such are values not random variables. Their expectations are themselves.
- ϵ_T is taken from a normal distribution with a mean of 0, so we expect it to be 0 within some variance.

$$E\left(Y_{T+2} \mid \dots\right) = 0.25 \cdot Y_T + 0.84 \cdot X_T$$

1.5.

Constructing a 95% confidence interval around the point at X_{T+1}:

$$\text{CI}_2 \Leftrightarrow E\left(Y_{T+2} \mid \dots\right) \pm 95\% \cdot \sqrt{\frac{V\left(Y_{T+2}\right)}{T+2}}$$

$$V\left(Y_{T+2} \mid \dots\right) = V\left(0.5 \cdot Y_{T+1} + 0.7 \cdot X_{T+1} + \epsilon_{T+2} - 0.2 \cdot \epsilon_{T+1}\right)$$

Some observations:

- All covariances involving one of the error terms are 0, as they are taken from an *iid* distribution.
- The variances of the error terms are equal; both equal to σ^2 of the normal distribution they are taken from. That value is 0.0625.

$$\cdots = V\left(0.5 \cdot Y_{T+1} + 0.7 \cdot X_{T+1}\right) + V\left(\epsilon_{T+2}\right) - 0.4 \cdot V\left(\epsilon_{T+1}\right)$$
$$\cdots = V\left(0.5 \cdot Y_{T+1} + 0.7 \cdot X_{T+1}\right) + 0.0375$$
$$\cdots = V\left(Y_{T+1}\right) + 1.4 \cdot V\left(X_{T+1}\right) + 2 \cdot coV\left(Y_{T+1}, X_{T+1}\right) + 0.0375$$

Another observation:

- $V\left(Y_{T+1}\right)$ is a recursive call.

$$\cdots = V\left(0.5 \cdot Y_T + 0.7 \cdot X_T + \epsilon_{T+1} - 0.2 \cdot \epsilon_T\right) + 1.4 \cdot V\left(X_{T+1}\right) + 2 \cdot coV\left(Y_{T+1}, X_{T+1}\right) + 0.0375$$
$$\cdots = V\left(Y_T\right) + 1.4V\left(\cdot X_T\right) + V\left(\epsilon_{T+1}\right) - 0.4 \cdot V\left(\epsilon_T\right) + 1.4 \cdot V\left(X_{T+1}\right) + 2 \cdot coV\left(Y_{T+1}, X_{T+1}\right) + 0.0375$$
$$\cdots = V\left(Y_T\right) + 1.4V\left(\cdot X_T\right) + 1.4 \cdot V\left(X_{T+1}\right) + 2 \cdot coV\left(Y_{T+1}, X_{T+1}\right) + 0.075$$

Another observation:

- Both the variances of Y_T and X_T are 0, as they are values.

$$\dots = 1.4 \cdot V\left(X_{T+1}\right) + 2 \cdot coV\left(Y_{T+1}, X_{T+1}\right) + 0.075$$

Dealing with the covariance of Y_{T+1} and X_{T+1}.

The expression for Y_t includes X_{t-1}, but not X_t. The error terms of Y_t are *iid* normal, and thus do not include X_t. Finally, time cannot run backwards, so neither Y_{t-1} nor X_{t-1} are functions of X_t. It is safe to say that $coV\left(Y_{T+1}, X_{T+1}\right) = 0$.

Finally, the value of $V\left(X_{T+1}\right)$ is known.

$$V\left(Y_{T+2} \mid \dots\right) = 1.4 \cdot 0.0625 + 0.075 = 0.1625$$

$$\dots = E\left(Y_{T+2} \mid \dots\right) \pm 95\% \cdot \sqrt{\frac{0.1625}{T+2}}$$

$$\dots = 0.25 \cdot Y_T + 0.84 \cdot X_T \pm 95\% \cdot \sqrt{\frac{0.1625}{T+2}}$$

2.

```
knitr::opts_chunk$set(eval=TRUE, cache=TRUE)
library(tidyverse)
library(stargazer)
library(forecast)
```

Data:

```
DATA <- inner_join(read.csv("data/PCESC96.csv"),
                   read.csv("data/UNRATE.csv"), by = "DATE") |>
  mutate(MONTH = DATE |>
    str_split("-") |>
    map(as.numeric) |>
    map(\(date) date[1] * 12 + date[2] - 1) |>
    unlist()) |>
  mutate(YEAR = MONTH / 12) |>
    mutate(DIF_PCESC96 = PCESC96 - PCESC96 |> lag(n=1)) |>
    mutate(DIF_UNRATE  = UNRATE - UNRATE |> lag(n=1)) |>
  na.omit()
```

Best Model:

I estimated 36 models in total, because it only takes my computer a second-ish to complete those arima models. This choice was one of computational convenience.

The algorithm used for finding the best model given the parameters is as follows:

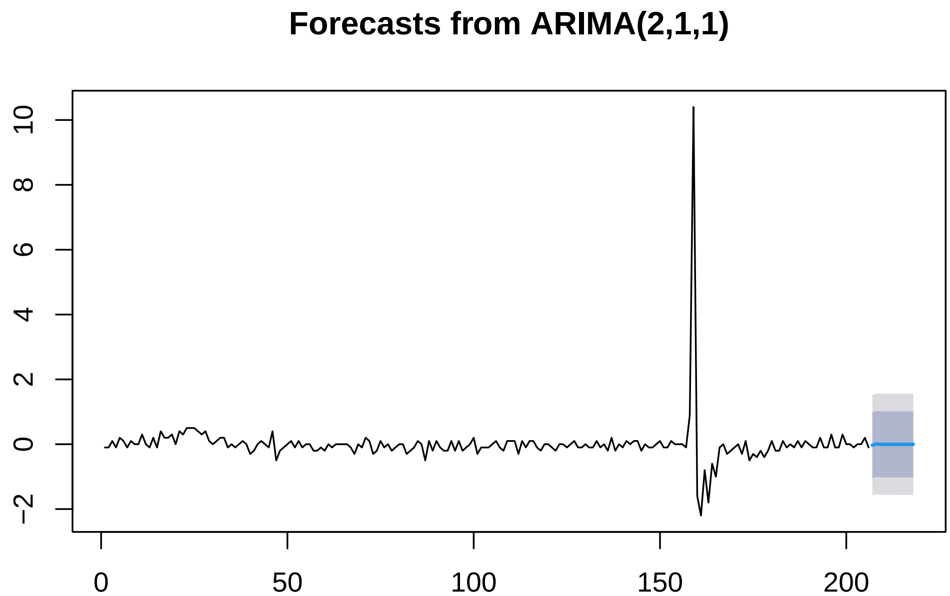
```
H <- 12

BEST_XMODEL <- NULL
XFORECAST <- NULL
```

```
for (i in 1:6) {
  for (j in 1:6) {
    XMODEL <- DATA$DIF_UNRATE |>
      Arima(order = c(i, 1, j))

    if (BEST_XMODEL |> is.null() || BEST_XMODEL$aic > XMODEL$aic) {
      BEST_XMODEL <- XMODEL

      XFORECAST <- BEST_XMODEL |>
      forecast(h = H)
    }

    remove(XMODEL)
  }
}

BEST_MODEL <- NULL
FORECAST <- NULL

for (i in 1:6) {
  for (j in 1:6) {
    MODEL <- DATA$DIF_PCESC96 |>
      Arima(order = c(i, 1, j), xreg = cbind(
        lag(DATA$DIF_UNRATE, n = 1)
      ))

    if (BEST_MODEL |> is.null() || BEST_MODEL$aic > MODEL$aic) {
      BEST_MODEL <- MODEL

      FORECAST <- MODEL |>
      forecast(h = H, xreg = XFORECAST$mean |> as.vector())
    }

    remove(MODEL)
  }
  remove(j)
}
remove(i)
```

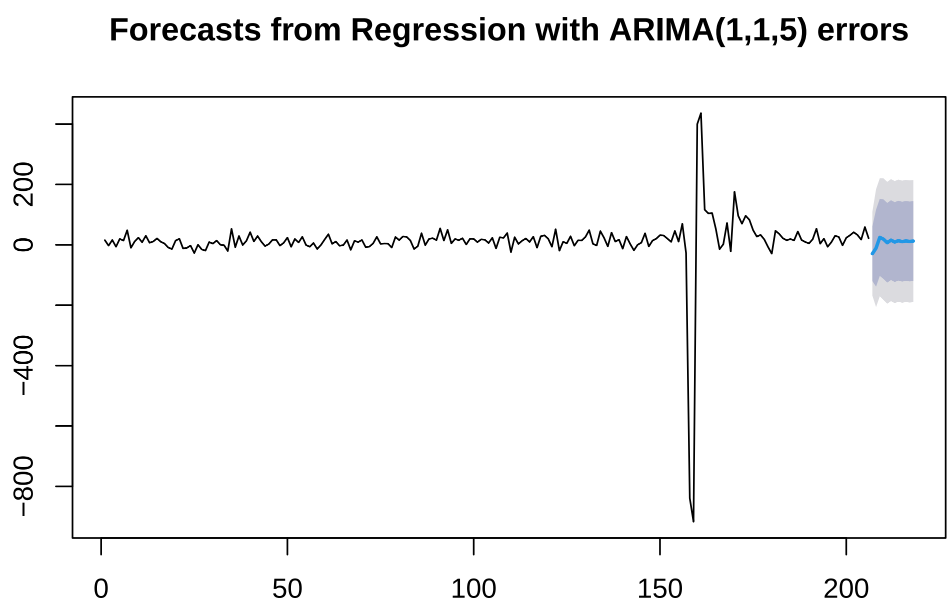
The Forecast:

Here is the best model:

```
XFORECAST |> plot()
```




```
FORECAST |> plot()
```

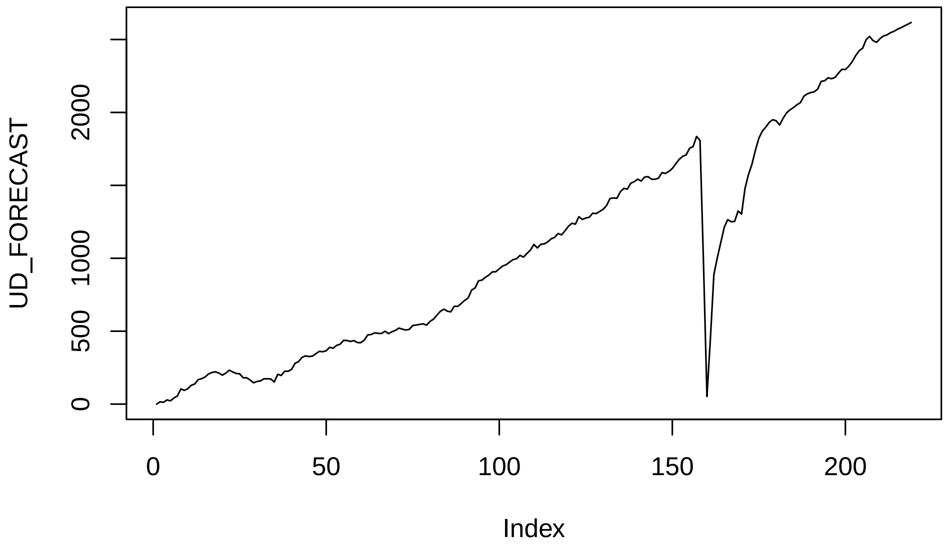


I've written a little function to un-difference the predicted values.

(It is named `discrete_integral` as differencing is also known as the "discrete derivative", sometimes as well the derivative is called the "continuous difference")

```
discrete_integral <- \(c) \(d) {
  xs <- c
  for (x in d$x) xs <- c(xs,x + xs |> tail(n=1))
  for (x in d$mean) xs <- c(xs,x + xs |> tail(n=1))
  xs
}

UD_FORECAST <- FORECAST |> discrete_integral(DATA$PCESC96 |> head(n=1))()
CI_HIGH <- (FORECAST$upper |> tail.matrix(n=1))[,2]
CI_LOW <- (FORECAST$lower |> tail.matrix(n=1))[,2]
```

```r
UD_FORECAST |> plot(type="l")
```



The mean prediction at 12 months is 1.0408235×10^4.

The AIC of this model is 2330.286294.

The confidence interval around this value is: $1.0408235 \times 10^4 + \{214.7010514, -189.8112843\}$.