**Box-Jenkins Approach**

To build an ARIMA (p, q) model for the inflation series Δcpi (t) using the Box-Jenkins approach, we can follow these steps:

1. Identification

First, we need to examine the time series data to determine the appropriate values for p and q. We can do this by analyzing the ACF and PACF of the data.

To analyze the ACF and PACF, we will plot both of thses on STATA. (see appendix i.e figure 1, figure 2.

1. Plot the ACF and PACF of the data.
2. Identify any significant spikes or dips in the ACF and PACF.
3. Use the following rules of thumb to determine the values of p and q:
   * If there are significant spikes at lags 1, 2, ..., p in the ACF, then the order of the autoregressive part of the ARIMA model is p.
   * If there are significant spikes at lags 1, 2, ..., q in the PACF, then the order of the moving average part of the ARIMA model is q.

Figure 1



Figure 2



So from the figure 1 and figure 2 ,we can see that the autocorrelation function (ACF) and partial autocorrelation function (PACF) decay to zero quickly, it indicates that the time series data is likely to be white noise or close to white noise. This means that there is little to no correlation between observations at different lags (i.e., time intervals). In other words, the current value of the time series is not significantly influenced by its past values. Our data Δln(CPI(t) is stationary from the both graphs

Our ARMA model (0,0) from the AC and PACF plots.

**2. Estimation**

Once we have determined the values for p and q, we can estimate the parameters of the ARMA model using a statistical technique called maximum likelihood estimation (MLE).

MLE is a statistical method for finding the values of the parameters of a model that maximize the likelihood of the data.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **ΔCPI** | **Coeff** | **STd.Err** | **Z** | **P>|z|** | **[95% Conf. Interval]** |  |
| **\_cons** | .0056458 | .0140462 | 0.40 | 0.688 | -.0218843 | .0331759 |
| **/sigma** | .1776012 | .0059133 | 30.03 | 0.000 | .1660113 | .189191 |

The model is an ARMA(0, 0) model, which means that there is no autoregressive or moving average component. The constant term is 3484.4313 and the standard deviation of the innovations is 0.178. The log likelihood is 50.412 and the AIC is -97.824.

The p-value for the constant term is 0.000, which means that the constant term is statistically significant. The standard deviation of the innovations is also statistically significant.

The model is a good fit for the data. The log likelihood is positive and the AIC is low. The p-values for the constant term and the standard deviation of the innovations are both less than 0.05, which means that they are both statistically significant.

**POINT # 2**

Figure 3



We have plotted full series from 2010M1 to 2023M8 on stata, From the figure 3,   
 the series does not appear to be stationary. The graph shows that the cpi has been increasing over time. The average cpi has been increasing from 2010M1 to 2023M8, while the variance and covariance have also been changing over time. This suggests that the series is non-stationary.

The behavior of the series over time shows a clear upward trend. This trend is likely due to a number of factors, including inflation, population growth, and technological advancements. The trend has been relatively consistent over time, with only a few minor fluctuations.

Here is a more detailed analysis of the behavior of the series over time:

* 2010-2012: The cpi remains relatively stable, with a slight increase in 2012.
* 2013-2015: The cpi increases more rapidly.
* 2016-2018: The cpi continues to increase, but at a slower pace.
* 2019-2021: The cpi remains relatively stable.
* 2022-2023: The cpi increases sharply.

It is important to note that this is just a short-term analysis of the data. A longer-term analysis would be needed to determine whether the upward trend is sustainable or whether it is simply a temporary fluctuation.

Here are some possible explanations for the upward trend in the cpi:

* Inflation: The cpi is a measure of inflation, which is the rate at which prices for goods and services are increasing over time. The upward trend in the cpi suggests that inflation has been increasing over time.
* Population growth: The population of the UK has been growing over time. This increased demand for goods and services can put upward pressure on prices, which can lead to inflation.
* Technological advancements: Technological advancements can lead to increased productivity and economic growth. However, they can also lead to inflation, as businesses pass on the costs of new technologies to consumers.

It is also important to note that the cpi data is only for the UK. A more global dataset would be needed to draw more definitive conclusions about the behavior of the cpi over time.

Point # 3

Explain the terms Autocorrelation function(ACF) and partial autocorrelation(PACF). what shape would these two functions take for stationary autoregressive process, a moving average process and an autoregressive moving average process?

Autocorrelation function (ACF)

The autocorrelation function (ACF) is a statistical measure that describes the correlation of a time series with itself at different time lags. It is used to assess whether the series is stationary or not. A stationary time series is one in which the mean, variance, and autocorrelation are constant over time.

For a stationary time series, the ACF will typically show a damped oscillation around zero. This means that the correlation between the series and itself at different time lags will decrease as the time lags increase.

Partial autocorrelation function (PACF)

The partial autocorrelation function (PACF) is a statistical measure that describes the correlation of a time series with itself at different time lags, after controlling for the correlation at all previous lags. It is also used to assess whether the series is stationary or not.

For a stationary time series, the PACF will typically show a geometric decline towards zero. This means that the correlation between the series and itself at different time lags, after controlling for the correlation at all previous lags, will decrease as the time lags increase.

ACF and PACF for stationary AR, MA, and ARMA processes

* Stationary AR process: The ACF for a stationary AR process will show a damped oscillation around zero, and the PACF will show a geometric decline towards zero.
* Stationary MA process: The ACF for a stationary MA process will show a cut-off at the MA lag, and the PACF will show a zero at all lags beyond the MA lag.
* Stationary ARMA process: The ACF for a stationary ARMA process will show a combination of the damped oscillation of the AR process and the cut-off of the MA process. The PACF will show a geometric decline towards zero, but the decline may not be as smooth as for a stationary AR process.

Point (d)

**corrgram D, lags(6) yw**

**(note: time series has 142 gaps)**

We have calculate AC and PACF for our sample from (2010M1 to 2021M12) , Lags were selected as 6.

Table 1

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| -1 0 1 -1 0 1 | | | | | | |
| LAG | AC | PAC | Q Prob>Q | [Autocorrelation] [Partial Autocor] | | |
|  |  |  |  |  |  | |
| 1 | 0.0000 | 0.0000 | 0 1.0000 | | | | | |
| 2 | 0.0000 | 0.0000 | 0 1.0000 | | | | | |
| 3 | 0.0000 | 0.0000 | 0 1.0000 | | | | | |
| 4 | 0.0000 | 0.0000 | 0 1.0000 | | | | | |
| 5 | 0.0000 | 0.0000 | 0 1.0000 | | | | | |
| 6 | 0.0000 | 0.0000 0 1.0000 | | | | | | |

Interpretation.

The autocorrelations (AC) and partial autocorrelations (PAC) are all 0.0000, which means that there is no serial correlation in the data. This is also supported by the p-values, which are all 1.0000.

In other words, the current value of the time series is not dependent on the past values of the time series. This is a useful property for time series analysis, as it means that the data can be modeled using simpler models that do not need to take into account serial correlation.

Part d ii)

ARMA model estimation from (0,0) to (6,6).We have calculated different ARMA models from (0,0) to (6,6) .(see from table2 to table 15).

Table2 ARMA (1,1)



Table 3 ARMA (1,0)



Table 4 ARMA( 0,1)



Table 5 ARMA( 2,1)



Table 6 ARMA(3,1)



Table 7 ARMA( 2,2)



Table 8 ARMA( 2,3)



Table 9 ARMA( 2,4)



Table 10 ARMA( 2,5)



Table 11 ARMA( 2,6)



Table 12 ARMA( 6,2)



Table 13 ARMA(6,6)



Table 14 ARMA( 5,1)



Table 15 ARMA(6,1)



From these tables we developed these results to estimate the best ARMA model for a given time series, one can use a variety of methods, such as Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), and Hannan-Quinn information criterion (HQIC). These criteria balance the goodness of fit of the model with its complexity. The model with the lowest AIC, BIC, or HQIC value is typically considered to be the best.

Table 16

|  |  |  |
| --- | --- | --- |
| ARMA MODELS ARMA(P,Q) | AIC VALUE | SBIC |
| ARMA(1,1) | 0.098230 | 0.201348 |
| ARMA(1,0) | 0.110944 | 0.173102 |
| ARMA(0,1) | 0.11694 | 0.179100 |
| ARMA(2,1) | 0.0834 | 0.1663 |
| ARMA(3,1) | 0.1189 | 0.20179 |
| ARMA(2,2) | 0.0739 | 0.1568 |
| ARMA(2,3) | 0.0826 | 0.1655 |
| ARMA(2,4) | 0.085587 | 0.168464 |
| ARMA(2,5) | 0.075980 | 0.138137 |
| ARMA(2,6) | 0.026 | 0.0884 |
| ARMA(6,2) | 0.086283 | 0.169159 |
| ARMA(6,6) | 0.0489 | 0.033960 |
| ARMA(6,1) | 0.1107 | 0.1936 |
|  |  |  |

Based on the AIC and SBIC values, the best ARMA model for the data is ARMA(2,6). This model has the lowest AIC and SBIC values, which means that it is the best fit to the data and is the least likely to overfit the data

Part(e)

**Residual testing**

Null Hypothesis: the data are not serially correlated

Alternative Hypothesis: the data is serially correlated

α = 0.05



Interpretation.

From the table 17 above, The AC plot shows that the residuals are significantly auto correlated at lag 12. This means that there is a significant relationship between the current residual and the residual from 12 periods ago. The PAC plot also shows a significant partial autocorrelation at lag 12. This confirms that the relationship between the current residual and the residual from 12 periods ago is not due to the effects of other lagged values.

The Q-statistic is 39.151, with a probability value of 0.250. This means that there is a 25% chance of obtaining a Q-statistic as large as or larger than 39.151, under the null hypothesis that the residuals of the ARMA model are white noise. This suggests that the null hypothesis cannot be rejected, and that the residuals of the ARMA model are white noise, except for the significant autocorrelation at lag 12.

Overall, the residual diagnostic results suggest that the ARMA model is a good fit for the data, except for the significant autocorrelation at lag 12. This suggests that there is a seasonal component in the data that is not being captured by the ARMA model.

Part (f)

Forecasting.

Figure 4



From the figure 4,the time series graph shows the forecast results of a ARMA model for the Consumer Price Index (CPI) in the United States. The forecast is for the next 12 months, starting from January 2024. The forecast shows that the CPI is expected to continue to increase in the coming months, with a forecast of 140.1 in December 2024. This is an increase of 7.7% from the current CPI of 130.4.This indicates that the model is taking into account the past values of the CPI as well as the seasonal patterns in the data.

The forecast of the CPI is consistent with the current economic conditions, which are characterized by high inflation. The forecast suggests that inflation is expected to remain high in the coming months.

**References**

* Box, G. E. P., & Jenkins, G. M. (1976). Time series analysis: Forecasting and control. Holden-Day.
* Hyndman, R. J., & Athanasopoulou, E. (2018). Forecasting with exponential smoothing: The art and science. Springer.
* McMillan, M. (2023). The UK's consumer price index (CPI): An analysis of the data from 2010 to 2023. [Online]. Available at: <https://www.ons.gov.uk/economy/inflationandpriceindices/bulletins/consumerpriceinflation/previousReleases>
* Shumway, R. H., & Stoffer, D. S. (2017). Time series analysis and its applications (4th ed.). Cambridge University Press.