STOCKHOLM UNIVERSITY
Department of Statistics
Econometrics II, Time Series Analysis, ST223G
Autumn semester 2019

# Written Re-examination in Econometrics II, Time Series Analysis

| | |
|---|---|
| Date | 2020-02-13 |
| Hour: | 16.00-21.00 |
| Examiner: | Jörgen Säve-Söderbergh |
| Allowed tools: | 1) Textbook: Wooldridge, J.M. *Introductory Econometrics: A Modern Approach*, Cengage, Boston.<br>2) Textbook: Montgomery, D.C., Jennings, C.L., and Kulachi, M., *Introduction to Time Series Analysis and Forecasting*, John Wiley & Sons, New Jersey.<br>3) Pocket calculator<br>4) Notes written in the text book are allowed. |

- Note that no formula sheet is provided.
- Passing rate: 50% of overall total, which is 100 points. For detailed grading criteria, see the course description.
- The maximum number of points for each problem is stated after each question. If not indicated otherwise, to obtain the maximum number of points on each problem, detailed and clear solutions are required. Answers may be given in English or Swedish.

**Good luck!**

1. The following observations have been recorded.

3 9 1 7 2 8 4 5 7 4

We think of these observations as having been generated by an underlying random process. Your task is to estimate the sample autocorrelation function (sample ACF) $r_k$ for the first two lags.

(a) Calculate $r_1$ och $r_2$. (8p)

(b) Although this is not done in our textbook *Montgomery et al*, you should also compute the sample partial autocorrelation function (sample PACF). This can be done by using a recursive formula. We use the notation $r_{ii}$ for the sample PACF. From the recursive formula we have

$$r_{11} = r_1 \quad \text{and} \quad r_{22} = \frac{r_2 - r_{11}r_1}{1 - r_{11}r_1}$$

Use this to calculate $r_{11}$ and $r_{22}$. (2p)

(c) The sample ACF $r_1$ is an estimator of the theoretical ACF $\rho_1$ defined on page 30 (in the first edition), also $r_2$ is an estimator of $\rho_2$.

Perform a hypothesis test of the null hypothesis $H_0 : \rho_k = 0$ for $k = 1, 2$. (10p)

2. The production of beer in the U.S.A. between the first quarter 1975 and the last quarter 1982 has been observed. A researcher has fitted an unknown model to this data (the model is unknown to you). After some discussion you realize that the model is some member of the ARIMA family. However, the researcher wants some help with the residual analysis.

   The researcher has collected the following table of the sample autocorrelation function of the residuals $r_e(k)$ from the estimation of the ARIMA model.

| k | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| $r_e(k)$ | -0.02292 | -0.27576 | 0.08407 | -0.14912 | -0.08831 | 0.07549 |
| k | 7 | 8 | 9 | 10 | 11 | 12 |
| $r_e(k)$ | 0.03141 | 0.10318 | -0.07282 | -0.15111 | -0.15678 | -0.03024 |

   (a) What can we tell the researcher about the individual sample autocorrelations? How can we do a hypothesis test? State the null hypothesis and suggest a test statistic. (6p)

   (b) Use the Ljung-Box goodness-of-fit statistic (found on page 57, equation (2.40), in the first edition) to test a set of sample autocorrelations. Consider the cases where $K = 6$ and $K = 12$. What null hypothesis are you testing? State the null hypothesis clearly for both cases of $K$. (12p)

3. Ambjörn is an archaeologist who has participated in excavations close to Sancerre in France. The expedition has noted the famous wines of Sancerre during the last summers. However the financiers, Horizon Europe, are worried. An official at Horizon Europe has been given the task to evaluate Ambjörn's expeditions consumption of wine. He finds the following:

| Year | 2015 | 2016 | 2017 | 2018 | 2019 |
|---|---|---|---|---|---|
| Number of bottles of Sancerre | 178 | 192 | 211 | 228 | 274 |

   (a) The official starts the investigation by producing a forecast of the number of wine bottles that the expedition is expected to consume this summer. On the boofshelf he has his old time series book: Montgomery *et al*, *Introduction to Time Series Analysis and*

*Forecasting.* On page 191 in the first edition he finds the model equations as follows for exponential smoothing with a trend

$$
\begin{aligned}
l_t &= \alpha y_t + (1-\alpha)(l_{t-1} + b_{t-1}) \\
b_t &= \gamma(l_t - l_{t-1}) + (1-\gamma)b_{t-1}
\end{aligned}
$$

Further reading he decides on letting $\alpha = 0.1$ and $\gamma = 0.1$.
To start the iteration he needs starting values, $l_0$ and $b_0$. He fits a regression model to the entire data material and and uses the estimated intercept as $l_0$ and the estimated regression coefficient as $b_0$. Which forecast for 2020 does the official produce? (10p)

(b) The official realises that the estimated regression model can also produce forecasts, besides the starting values in (a). Calculate a forecast using the regression model. Which forecast does the official produce? (5p)

(c) Now, the official has two forecasts. Which forecast would you recommend that the official uses in his report to Horizon Europe? Discuss the merits of each forecast in a non-technical way. (5p)

4. The following model was found to fit a time series well.

$$(1-B)y_t = \delta + (1-\theta B)\varepsilon_t$$

where $\varepsilon$ are independent with expected value 0 and with variance $\sigma^2$.

(a) What model is this? Explain each one of the parameters. (2p)

(b) Rewrite the model in difference equation-form (the form we use ordinarily). (4p)

(c) Is the model stationary? Is there any easy fix for this? (2p)

(d) Give a general comment on the autocorrelation function (ACF) $\rho_k, k = 1, 2, \ldots$ and the partial autocorrelation function (PACF) $\phi_{kk}, k = 1, 2, \ldots$ for this model. Explain using words or graphs. (4p)

(e) Calculate the entire infinite AR representation or give at least a few terms. Is there any connection with exponential smoothing? Explain. (8p)

5. Assume the model

$$y_t = a_t - 0.1a_{t-1} + 0.21a_{t-2}$$

where $a_t$ are independent and normally distributed with expected value 0 and with known variance $\sigma_a^2$.

(a) What kind of model is this? (2p)

(b) Compute $E(y_t)$. (4p)

(c) Compute $\text{Var}(y_t)$. (4p)

(d) Compute the first and second order autocorrelation $\rho_1$ and $\rho_2$. (10p)

# Appendix

## Rules for the covariance

The *covariance* between $X$ and $Y$ is defined as

$$\text{Cov}(X,Y) = E\{(X - E(X))(Y - E(Y))\}$$

The following formulas serves you well in this course.

(a) Relationship between the variance and the covariance

$$\text{Cov}(X,X) = \text{Var}(X).$$

(b) The order in which $X$ and $Y$ are mentioned in the covariance does not matter. We always have a symmetry:

$$\text{Cov}(X,Y) = \text{Cov}(Y,X).$$

(c) Multiplicative constants can be factored:

$$\text{Cov}(aX,Y) = a\text{Cov}(X,Y),$$

even

$$\text{Cov}(aX,bY) = ab\text{Cov}(X,Y).$$

(d) Additive constants has no effect on the covariance:

$$\text{Cov}(X+a,Y+b) = \text{Cov}(X,Y).$$

(e) The covariance between two sums

$$\text{Cov}(X+Y,Z+W) = \text{Cov}(X,Z)+\text{Cov}(X,W)+\text{Cov}(Y,Z)+\text{Cov}(Y,W).$$

A special case of (e) is

$$\text{Cov}(X,Z+W) = \text{Cov}(X,Z) + \text{Cov}(X,W).$$

(f) The covariance between two (different) linear combinations of $X$ and $Y$; $aX + bY$ and $cX + dY$

$$\text{Cov}(aX+bY,cX+dY) = ac\text{Var}(X)+bd\text{Var}(Y)+(ad+bc)\,\text{Cov}(X,Y).$$

(g) Important special case of (f): $V(aX+bY) = a^2V(X) + b^2V(Y) + 2ab\text{Cov}(X,Y)$.

Department of Statistics

Stockholms universitet

# Correction sheet

**Date:** 13/02/2020

**Room:** Uggleviksalen

**Exam:** Econometrics 2

**Course:** Econometrics

**Anonymous code:** 0021-MJK

☒ I authorise the anonymous posting of my exam, in whole or in part, on the department homepage as a sample student answer.

**NOTE! ALSO WRITE ON THE BACK OF THE ANSWER SHEET**

**Mark answered questions**

| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | Total number of pages |
|---|---|---|---|---|---|---|---|---|---|
| x | x | x | x | x | | | | | 10 |
| Teacher's notes 20 | 12 | 20 | 20 | 20 | | | | | |

| Points | Grade | Teacher's sign. |
|---|---|---|
| 92 | A | JSS |

**SU, DEPARTMENT OF STATISTICS**

Room: Uggleviks salen Anonymous code: 0021-MJK Sheet number: 1

Uppgift 1. 3 9 1 7 2 8 4 5 7 4 sample ACF $r_k$ $k=1$ $k=2$

a) Beräkna $r_1$ och $r_2$.

Formel

$r_k = \hat{\rho}_k = \frac{c_k}{c_0}$ $k = 0,1,2,\dots k.$

$c_k = \hat{\gamma}_k = \frac{1}{T}\sum_{t=1}^{T-k}(y_t-\bar{y})(y_{t+k}-\bar{y}) = \frac{1}{T}\left(\sum y_t y_{t+k} - \bar{y}\sum y_t - \bar{y}\sum y_{t+k} + (T-k)\bar{y}^2\right)$

OBSERVERA ATT FÖR SAMTLIGA BERÄKNINGAR AV $C_k$ SÅ MINSKAR T MED 1 NÄR K ÖKAR MED 1. MEN VÄRDET PÅ $\bar{Y}$ KONSTANT I SAMTLIGA BERÄKNINAR ALLTSÅ $\bar{Y} = \frac{\sum Y_t}{T}$.

$Y_1 = 3$
$Y_2 = 9$
$Y_3 = 1$
$Y_4 = 7$
$Y_5 = 2$
$Y_6 = 8$
$Y_7 = 4$
$Y_8 = 5$
$Y_9 = 7$
$Y_{10} = 4$

$T = 10$
$\bar{Y} = 5$

$k=0 \Rightarrow c_0 = \frac{1}{10}\sum_{t=1}^{10}(y_t-\bar{y})(y_t-\bar{y}) = \frac{1}{10}\left(\sum y_t^2 - T\cdot\bar{y}^2\right) = \frac{1}{10}\left(314-(10\cdot 5^2)\right) = 6,4$

$k=1 \Rightarrow c_1 = \frac{1}{10}\sum_{t=1}^{9}(y_t-\bar{y})(y_{t+1}-\bar{y}) = \frac{1}{10}\left(\sum y_t y_{t+1} - \bar{y}\sum y_t - \bar{y}\sum y_{t+1} + (T-1)\bar{y}^2\right) =$

$= \frac{1}{10}\left(188-(5\cdot 46)-(5\cdot 47)+(9\cdot 5^2)\right) = -5,2$

$k=2 \Rightarrow c_2 = \frac{1}{10}\sum_{t=1}^{8}(y_t-\bar{y})(y_{t+2}-\bar{y}) = \frac{1}{10}\left(\sum y_t y_{t+2} - \bar{y}\sum y_t - \bar{y}\sum y_{t+2} + (T-2)\bar{y}^2\right) =$

$= \frac{1}{10}\left(220-(5\cdot 39)-(5\cdot 38)+(8\cdot 5^2)\right) = 3,5$

$r_0 = \frac{c_0}{c_0} = \frac{6,4}{6,4} = 1$

$r_1 = \frac{c_1}{c_0} = -\frac{5,2}{6,4} = -0,8125$

$r_2 = \frac{c_2}{c_0} = \frac{3,5}{6,4} = 0,546875$

Svar: $r_1 = -0,8125$ $r_2 = 0,546875$

**SU, DEPARTMENT OF STATISTICS**

**Room:** Ugglevikssalen **Anonymous code:** 0021-MJK **Sheet number:** 2

Uppgift 1.

b) Beräkna sample PACF med formeln: $r_{11} = r_1$ $r_{22} = \frac{r_2 - r_{11} r_1}{1 - r_{11} r_1}$

$r_{11} = r_1 = -0,8125$

$r_{22} = \frac{0,546875 - (-0,8125)^2}{1 - (-0,8125)^2} = -\frac{1}{3}$

Svar: $r_{11} = -0,8125$ $r_{22} = -\frac{1}{3}$ R

c) Testa $\rho_1$ och $\rho_2$ med nollhypotesen $H_0: \rho_k = 0$

$r \sim N(0, \frac{1}{n})$

$H_0: \rho_k = 0$ teststatistika: $z = \frac{r_k}{\sqrt{\frac{1}{n}}} = r_k \cdot \sqrt{n}$

$H_A: \rho_k \neq 0$

signifikansnivå: $\alpha = 0,05$

$z_{krit} = z_{\alpha/2} = z_{0,05/2} = z_{0,025} = 1,96$ från tabell Montgomery.

Beslutsregel förkasta $H_0$ om $|z_{OBS}| > z_{krit}$.

$H_0: \rho_1 = 0$
$H_A: \rho_2 \neq 0$

$z_{OBS} = -0,8125 \cdot \sqrt{10} \approx -2,569$ R

$|-2,569| > 1,96 \Rightarrow$ alltså $H_0$ förkastas.

$H_0: \rho_2 = 0$
$H_A: \rho_2 \neq 0$

$z_{OBS} = 0,546875 \cdot \sqrt{10} \approx 1,729$ R

$|1,729| < 1,96 \Rightarrow$ alltså kan ej förkastas.

Svar: $H_0: \rho_1 = 0$ förkastas men $H_0: \rho_2 = 0$ kan ej förkastas. R

**SU, DEPARTMENT OF STATISTICS**

Room: Ugglevikssalen Anonymous code: 0021-HJK Sheet number: 3

Uppgift 2. OBSERVATIONER 1975 – 1982 varje kvartal. Q1, Q2, Q3, Q4

$n = 8 \cdot 4 = 32$

| k | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| $r_e(k)$ | -0,02292 | -0,27576 | 0,08407 | -0,14912 | -0,08831 | 0,07549 |
| k | 7 | 8 | 9 | 10 | 11 | 12 |
| $r_e(k)$ | 0,03141 | 0,10318 | -0,07282 | +0,15111 | -0,15678 | -0,03024 |

Någon typ av ARIMA modell

b) Ljungbox

Finns det autokorrelation i residualerna vi kan testa med Ljung-Box test.

$H_0: \rho_{e1} = \rho_{e2} = \rho_{e3} = \rho_{e4} = \rho_{e5} = \rho_{e6} = \rho_{e7} = \rho_{e8} = \rho_{e9} = \rho_{e10} = \rho_{e11} = \rho_{e12} = 0$

$H_A$: Minst 1 av $\rho_{e(k)}$ är skild från 0.

Teststatistika $Q_{LB} = n(n+2)\sum_{k=1}^{K}\left(\frac{\hat{\rho}_{e(k)}^2}{n-k}\right) \approx \chi^2_{(K)}$ $\chi^2_{krit} = \chi^2_{0,05}$

Beslutsregel: Förkasta $H_0$ om $\chi^2_{OBS} > \chi^2_{0,05}(K)$

$Q_{LB} = 32(32-2)\sum_{k=1}^{12}\left(\frac{\hat{\rho}_{(k)}^2}{32-k}\right) = 960\sum_{k=1}^{12}\left(\frac{\hat{\rho}_{(k)}^2}{n-k}\right) =$

$= 960\left[\frac{(-0,02292)^2}{31} + \frac{(-0,27576)^2}{30} + \ldots + \frac{-0,03024}{21}\right] = 6,5782$

R

$\chi^2_{krit} = \chi^2_{0,05}(12) = 18,55$
← från tabell

$\chi^2_{OBS} < \chi^2_{krit}$ alltså kan $H_0$ inte förkastas.

Svar: Vi kan genom Ljung Box hypotestest inte påvisa någon autokorrelation i residualerna. R

Jag förstår inte riktigt vad det syftas på att jag ska göra jag hade gjort Ljung Box testet i både a) och b)

$H_0: \rho_{e1} = \rho_{e2} = \rho_{e3} = \rho_{e4} = \rho_{e5} = \rho_{e6}$ $Q_{LB} = 960\sum_{k=1}^{6}\left(\frac{\hat{\rho}_{(k)}^2}{(32-k)}\right) = 3,93373$ R

$H_A$: Minst en $\rho_{(k)}$ är skild från 0. $\chi^2_{krit} = \chi^2_{0,05}(6) = 12,59$

Vi kan ej förkasta $H_0$.

Man kan testa varje autokorrelation individuellt

SU, DEPARTMENT OF STATISTICS

Room: Uggleviksssalen Anonymous code: 0021-MJK Sheet number: 4

Uppgift 2.

Svar: Vi kan ej förkasta $H_0$ för vare sig $k=6$ eller $k=12$.

Vi kan alltså inte påvisa autokorrelation i residualerna.

**SU, DEPARTMENT OF STATISTICS**

Room: Ugglevikssalen Anonymous code: 0021-MDK Sheet number: 5

Uppgift 3.

| Year | 2015 | 2016 | 2017 | 2018 | 2019 |
|---|---|---|---|---|---|
| Number of Bottles | 178 | 192 | 211 | 228 | 274 |

a) Forecast 2020. Holt's method $\alpha = 0{,}1$ $\gamma = 0.1$

$l_t = \alpha y_t + (1-\alpha)(l_{t-1} + b_{t-1})$

$b_t = \gamma(l_t - l_{t-1}) + (1-\gamma) b_{t-1}$

skatta Startvärdena $l_0$ och $b_0$.

Skatta $l_0$ och $b_0$ med OLS. där $l_0 = a$ (interceptet) och $b_0 = b$ (riktningskoefficienten) År = x Antal flaskor = y

$a = \bar{y} - b\bar{x} = 216{,}6 - (22{,}8 \cdot 3) = 148{,}2$

$b = \frac{\Sigma x_i y_i - n\bar{x}\bar{y}}{\Sigma x_i^2 - n\bar{x}^2} = \frac{3477 - (5 \cdot 3 \cdot 216{,}6)}{55 - (5 \cdot 3^2)} = 22{,}8$

$2015 = 1 = x_1$ $178 = y_1$
$2016 = 2 = x_2$ $192 = y_2$
$2017 = 3 = x_3$ $211 = y_3$
$2018 = 4 = x_4$ $228 = y_4$
$2019 = 5 = x_5$ $274 = y_5$
$N = 5$

$l_0 = a = 148.2$ R $b_0 = b = 22{,}8$ D

$l_1 = \alpha y_1 + (1-\alpha)(l_0 + b_0) = 0{,}1 \cdot 178 + 0{,}9(148{,}2 + 22{,}8) = 171{,}7$
$b_1 = \gamma(l_1 - l_0) + (1-\gamma) b_0 = 0{,}1(171{,}7 - 148{,}2) + 0{,}9 \cdot 22{,}8 = 22{,}87$

$l_2 = \alpha y_2 + (1-\alpha)(l_1 + b_1) = 0{,}1 \cdot 192 + 0{,}9(171{,}7 + 22{,}87) = 194{,}313$
$b_2 = \gamma(l_2 - l_1) + (1-\gamma) b_1 = 0{,}1(194{,}313 - 171{,}7) + 0{,}9 \cdot 22{,}87 = 22{,}8443$

$l_3 = \alpha y_3 + (1-\alpha)(l_2 + b_2) = 0{,}1 \cdot 211 + 0{,}9(194{,}313 + 22{,}8443) = 216{,}54157$
$b_3 = \gamma(l_3 - l_2) + (1-\gamma) b_2 = 0{,}1(216{,}54157 - 194{,}313) + 0{,}9 \cdot 22{,}8443 = 22{,}782727$

$l_4 = \alpha y_4 + (1-\alpha)(l_3 + b_3) = 0{,}1 \cdot 228 + 0{,}9(216{,}54157 + 22{,}782727) = 238{,}1918673$
$b_4 = \gamma(l_4 - l_3) + (1-\gamma) b_3 = 0{,}1(238{,}1918673 - 216{,}54157) + 0{,}9 \cdot 22{,}782727 = 22{,}66748403$

$l_5 = \alpha y_5 + (1-\alpha)(l_4 + b_4) = 0{,}1 \cdot 274 + 0{,}9(238{,}1918673 + 22{,}66748403) = 262{,}1752162$
$b_5 = \gamma(l_5 - l_4) + (1-\gamma) b_4 = 0{,}1(262{,}1752162 - 238{,}1918673) + 0{,}9 \cdot 22{,}66748403 = 22{,}80087052$

Forecast för $\hat{y}_{2020} = \hat{y}_{T+\tau}(T) = \hat{y}_{5+1}(5) = l_5 + 1 \cdot b_5 = 284{,}9760867 \approx 295.$ R

← diskret värde avrunda uppåt!

Svar: Det predikerade värdet för antal flaskor är 295 st.

SU, DEPARTMENT OF STATISTICS

Room: Ugglevikssalen Anonymous code: 0021-MJK Sheet number: 6.

b) estimera $Y_{2020}$ med hjälp av OLS. Uppgift 3.

Vi använder modellen $Y_t = a + b \cdot t$.

Vi vet från uppgift a) att:

$a = 148,2$ $b = 22,8$

$\hat{Y}_{2020} = 148,2 + 22,8 \cdot 6 = 285$

Svar: Det predikterade värdet för antal flaskor är 285 st. R

c) Vilken modell ska användas?

Den modell med lägst värde på MSE är den bäst lämpade eftersom att den är mest pricksäker. Det är inte så konstigt i tanke på att $MSE = \hat{\sigma}_\varepsilon^2$ vi vill ha en så låg spridning som möjligt på residualerna då det innebär en lägre felmarginal för skattningarna.

Bra svar!

SU, DEPARTMENT OF STATISTICS

Room: Ugglevikssalen Anonymous code: 0021-MJK Sheet number: 7

Uppgift 4 $(1-B)y_t = \delta + (1-\theta B)\varepsilon_t$ $\varepsilon_t \overset{iid}{\sim} N(0, \sigma^2)$

a) identifiera modellen. $B^d y_t = y_{t-d}$

$(1-B)y_t = \delta + (1-\theta B)\varepsilon_t \iff$

$y_t - y_{t-1} = \delta + \varepsilon_t - \theta\varepsilon_{t-1}$

$(-\theta\varepsilon_{t-1})$ säger ~~att det~~ är en MA(1).

$(y_{t-1})$ säger att $\phi_1 = 1$ villkoret innebär att modellen inte kan uttryckas $y_t = \delta + y_{t-1} + \varepsilon_t - \theta\varepsilon_{t-1}$ för då är den inte stationär.

$\Delta y_t = y_t - y_{t-1}$ innebär att man differentierat för att uppnå stationäritet. Alltså det är en I(1).

Modellen är en ARIMA(p, d, q)-modell.

Denna modellen är en ARIMA(0,1,1).

Svar: Det är en ARIMA(0,1,1)-modell.

b) Formulera modellen i differentierad ekvationsform på det vanliga sättet.

$y_t - y_{t-1} = \delta + \varepsilon_t - \theta\varepsilon_{t-1} \iff$

$\Delta y_t = \delta - \theta\varepsilon_{t-1} + \varepsilon_t$.

Svar: $\Delta y_t = \delta - \theta\varepsilon_{t-1} + \varepsilon_t$.

**SU, DEPARTMENT OF STATISTICS**

Room: Ugglevikssalen Anonymous code: 0021-HJK Sheet number: 8

Uppgift 4.

c) Är modellen stationär? Finns det någon enkel lösning?

Modellen uttryckt i:

$y_t = \delta + y_{t-1} - \theta \varepsilon_{t-1} + \varepsilon_t$ är inte stationär eftersom $\phi_1 = 1$.

För att den ska vara stationär krävs att $|\phi_1| < 1$.

När $|\phi_1| > 1$ så kan man invertera modellen så att $|\frac{1}{\phi_1}| < 1$.

Detta är inte en okomplicerad procedur och det är heller inte aktuellt i detta fall då $|\phi_1| = 1$.

När $\phi_1 = 1$ så kan vi differentiera modellen för att uppnå stationäritet. Alltså $y_t - y_{t-1} = \Delta y_t$.

Man kan tänka att $y_t$ ärver hela $y_{t-1}$ värde, plus att $y_t$:s värde påverkas av ytterligare faktorer. För att bli av med den ackumulerande delen ($y_{t-1}$) så är det enklaste att bara subtrahera $y_t$ med $y_{t-1}$.

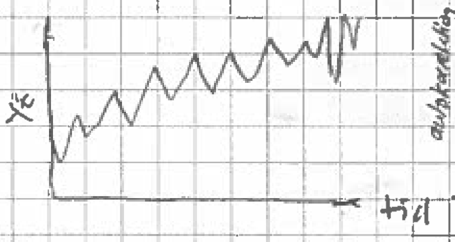
**SU, DEPARTMENT OF STATISTICS**

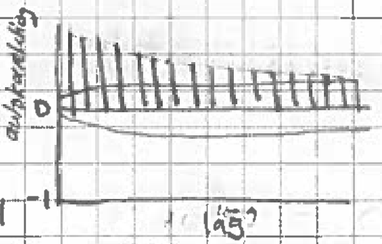
Room: Uggleviksssalen Anonymous code: 0021-HJK Sheet number: 4

Uppgift 4.

a) Kommentera ACF $\rho_k$ $k=1,2,\ldots$ och PACF $\phi_{kk}$ $k=1,2\ldots$ för denna modell. Beskriv med ord eller Bilder.

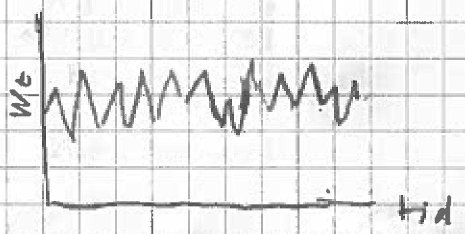
① $Y_t = Y_{t-1} + \delta - \varepsilon_{t-1} + \varepsilon_t$. Denna tidsserie kan tänkas se ut ungefär såhär på en plott samt ACF.



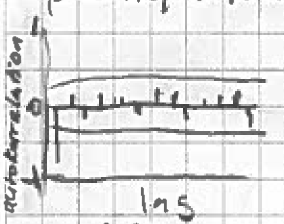




Man ser på plotten att det är en positiv lutning och att medelvärdet ändrar sig med tiden – alltså ej stationär och autokorrelationen är hög och svagt avtagande.

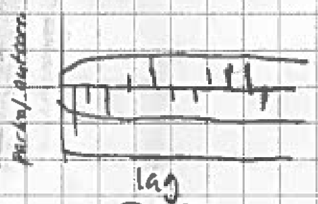
② $\Delta Y_t = Y_t - Y_{t-1} = \delta - \theta\varepsilon_{t-1} + \varepsilon_t$. Samma tidsserie differentieras och plotten ser ut ungefär såhär.




Det är ingen signifikant skillnad på medelvärdet oavsett tidsperiod. Modellen är stationär.
Låt oss kalla denna tidsserie $w_t$ $w_t = \Delta Y_t = \delta - \theta\varepsilon_{t-1} + \varepsilon_t$.
$w_t$ ser ut i formen som en MA(1) och därför kommer ACF samt PACF bete sig som att det vore en MA(1)-modell.
Nu vet vi inte om $\theta_1$ är negativ eller positiv, men för ett positivt värde på $\theta_1$ så kommer ACF samt PACF se ut såhär typ.




ACF




PACF

I både ACF samt PACF så kommer första spiken vara negativ och de laggar med ett värde $k \geq 1$ kommer inte att vara signifikanta.

R

**SU, DEPARTMENT OF STATISTICS**

Room: Ugglaviksssalen Anonymous code: 0021-MJK Sheet number: 10

Uppgift 5. $y_t = a_t - 0{,}1a_{t-1} + 0{,}21a_{t-2}$ $a_t \overset{iid}{\sim} N(0, \sigma_a^2)$

a) identifiera modellen.

Svar: Det är en second-order moving average process, MA(2)-modell
Allmänt beskriven enligt: $y_t = \mu + \varepsilon_t - \theta_1\varepsilon_{t-1} - \theta_2\varepsilon_{t-2}$ R

b) Beräkna $E(y_t)$

$E(y_t) = E(a_t - 0{,}1a_{t-1} + 0{,}21a_{t-2}) = E(a_t) - 0{,}1E(a_{t-1}) + 0{,}2E(a_{t-2}) =$

$= 0 - 0{,}1 \cdot 0 + 0{,}2 \cdot 0 = 0$

Svar: $E(y_t) = 0$ R

c) Beräkna $V(y_t)$

$V(y_t) = V(a_t - 0{,}1a_{t-1} + 0{,}21a_{t-2}) = V(a_t) + 0{,}1^2V(a_{t-1}) + 0{,}21^2V(a_{t-2}) =$

$= \sigma_a^2 + 0{,}1^2\sigma_a^2 + 0{,}21^2\sigma_a^2 = \sigma_a^2(1 + 0{,}1^2 + 0{,}21^2) = 1{,}0541\sigma_a^2$

Svar: $V(y_t) = 1{,}0541\sigma_a^2$ R

d) Beräkna $\rho_1$ och $\rho_2$

$\theta_1 = 0{,}1 \quad \theta_2 = -0{,}21$

$\rho_1 = \frac{-\theta_1 + \theta_1\theta_2}{1 + \theta_1^2 + \theta_2^2} = \frac{-0{,}1 + (0{,}1 \cdot (-0{,}21))}{1 + 0{,}1^2 + (-0{,}21)^2} = \frac{-0{,}079}{1{,}0541} \approx 0{,}0749454511$ R

$\rho_2 = \frac{-\theta_2}{1 + \theta_1^2 + \theta_2^2} = \frac{-(-0{,}21)}{1{,}0541} = \frac{0{,}21}{1{,}0541} \approx 0{,}1992220852$ R

Svar: Avrundat till 4 decimaler $\rho_1 = 0{,}0750$ och $\rho_2 = 0{,}1992$