Exam in STK4060/STK9060 – Time Series, spring 2006

Your answer to the problems below will be the main basis for your evaluation, and the point of departure for the oral examination. You are supposed to work individually on the problems, but discussions are of course permitted.
The deadline is **Monday 29. May at 1700**. Your answer should preferably be sent as an attachment (in .pdf format) to an email to eivind.damsleth@hydro.com. Alternatively you can put it in my mailbox on the left hand side in the room next to the Expedition at the 7. floor in N.H.Abel's building. Handwritten answers are OK, as long as they can be easily read. Attach only those parts of your computer printouts that are needed to support your answers, and make sure that it's clear what part of the printouts you refer to.

Good luck! Eivind Damsleth

- 1 -

The file kpi.txt contains 243 monthly observations of the consumer goods price index (KPI) from the Central Bureau of Statistics (CBS), from Jan. 1986 to March 2006 (both included). The content is also shown in Appendix 1.

a) Identify a suitable SARIMA model for this series, and estimate its parameters. Explain the choices you make (Box-Cox transform, differencing, proposed AR and MA order for the ordinary and seasonal parameters).

b) Use the model you have found to forecast the KPI for the period Jan. 2005-March 2006 (obs. 229-243) based on the data up to and including December 2004 (obs. no. 228). Also find the associated ca. 95 % confidence intervals for each forecast. Compare with the observed values for the same period. Conclusion?
This comparison will probably give a somewhat optimistic picture of the model's forecasting ability. Why?

c) The KPI-values are presented with only one decimal. You can assume that CBS uses standard rules for rounding. Assuming that the rounding errors are uniformly distributed in the interval <-0.05,0.05]. Then, what is the lower limit of the standard deviation for the 1-step forecast error (= residual standard deviation if there is no Box-Cox transformation in your model)?

d) One possible definition of the annual inflation in Norway (in %) is: $Y_t = \frac{X_t - X_{t-12}}{X_{t-12}} \cdot 100\%$,

where X_t is the KPI value for month t and Y_t is the increase (decrease) in % over the last 12 months. Note that as long as the annual inflation is close to 0, a 1. order Taylor expansion of $\ln(\frac{X_t}{X_{t-12}}) = \ln(1+Y_t/100)$ shows that $Y_t \approx \ln(\frac{X_t}{X_{t-12}}) \cdot 100\% = (1-\mathrm{B}^{12})\ln(X_t) * 100\%$ .We can assume the approximation to be an equality for our purpose. This may be useful in the following.

Norges Bank has set as a target for the inflation: The annual inflation should be approximately 2.5%. Lately it has been significantly lower. Find a suitable SARIMA model for $\{Y_t\}$, either by deduction from the model you found in part b) or by identifying/estimating a separate model for $\{Y_t\}$. Use the data up to and including March 2006 to forecast the annual inflation at the end of December 2006 and at the end of December 2007, with associated ca. 95% confidence intervals. What is (approximately) the probability that Norges Bank reaches its inflation target, i.e. that the inflation becomes > 2.5% at these two point in time?

In an industrial process a number of different dry substances are mixed in a blending drum, according to a given weight recipe. After blending the drum is discharged, and the result is weighed. The discharge process is not perfect, so that some residue will be left in the drum. This residue will then constitute a part of the next batch. The amount of residue will vary from one batch to the next.

The weighing of the input materials has some measurement error. The measurement precision (the standard deviation of the measurement error) for the weighing equipment is known. There is also some measurement error in the output, again with a known standard deviation.

Use the following notation:

I_t = Total actual (correct) input weight to batch no. t (unobserved)
μ_{I} = Required total input weight according to the recipe, so that $\mathrm{E}I_t = \mu_{\mathrm{I}}$ (known)
Z_{It} = Total input weight error, so that $Z_{It} = I_t - \mu_{\mathrm{I}}$ (unobserved)
σ_I = Standard deviation for Z_{It}. (Known, calculated as the square root of the sum of the variances for the individual weighing processes)

R_t = Residue in the blender after discharging batch no. t (unobserved)
R_t is assumed to follow a statuonary AR(1) process: $(R_t - \mu_R) = \varphi(R_{t-1} - \mu_R) + Z_{Rt}$,
where Z_{Rt} WN$(0, \sigma_{\mathrm{R}}^2)$ and the mean μ_R is known.

O_t = Actual output weight from batch no. t (unobserved)

Y_t = Measured output weight from batch no. t (Observed). The weighting equipment is unbiased, so that $\mathrm{E}(Y_t|O_t) = O_t$
Z_{Yt} = Measurement error on output, i.e. $Z_{Yt} = Y_t - O_t$ (unobserved)
σ_Y = Standard deviation for Z_{Yt} : $\sigma_Y = \mathrm{Stdv}(Y_t|O_t)$. (known)

The dynamics of this industrial process can then be written:

$$\begin{aligned} I_{t+1} &= \mu_I + Z_{I\,t+1} \\ O_{t+1} &= I_{t+1} + R_t + R_{t+1} \\ R_{t+1} - \mu_R &= \varphi(R_t - \mu_R) + Z_{R\,t+1} \\ Y_{t+1} &= O_{t+1} + Z_{Y\,t+1} \end{aligned} \qquad (1)$$

This industrial process can be formulated as a *state-space* model

$$\begin{aligned} \mathbf{X_{t+1}} &= \mathbf{F}\mathbf{X_t} + \mathbf{V_{t+1}} \\ Y_t &= \mathbf{G}\mathbf{X_{t+1}} + W_t \end{aligned} \qquad (2)$$

with state vector $\mathbf{X_t}=(1, I_t, O_t, R_t)'$.

a) Show that: $\mathbf{F} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ \mu_I & 0 & 0 & 0 \\ \mu_I - (1-\varphi)\mu_R & 0 & 0 & (1-\varphi) \\ (1-\varphi)\mu_R & 0 & 0 & \varphi \end{bmatrix}$, $\mathbf{G} = (0,0,1,0)$, $V_t = \begin{bmatrix} 0 \\ Z_{It} \\ Z_{It} - Z_{Rt} \\ Z_{Rt} \end{bmatrix}$ and $W_t = Z_{Yt}$,

and that (2) fulfils the requirements for a *state-space* representation.

b) What are the conditions on the parameters to make this *state-space* representation stable?

c) Try to give an informal ”physical” interpretation of the parameter φ. What do values close to +1 mean? Close to –1? And near 0?

d) The (unobserved) series of actual output from the blending process, $\{O_t\}$, can be modeled as an ARMA process. Find this process, and give the relation between the parameters in this ARMA process and the parameters in the *state-space* representation. What happens when $\varphi \to 1$? And $\varphi \to 0$?

e) Similarly, the series with the observed, measured output, $\{Y_t\}$, can also be modeled as an ARMA process. Which? And what is the relation between the parameters in this ARMA process and those in the *state-space* representation?

f) The file batch.txt (the content is also given in App. 2) contains 200 sequential observations from $\{Y_t\}$. For these data we have: μ_I = 1000, σ_I = 20, μ_R = 20 and σ_Y = 20.

Identify an ARMA model for these data, and estimate its parameters. Are the results consistent with your findings in part e)? Use the results from part e) to find estimates for the parameters φ and σ_R in the *state-space* representation.

Appendix 1: Kpi.txt – Consumer goods price index, monthly, Jan. 1986 – March 2006

	1986-89	1990-93	1994-97	1998-2001	2002-2005	2006
Jan.	63.8	81.9	90.7	98.9	109.0	115.6
Feb.	64.0	82.3	91.0	99.3	109.3	116.6
March	64.5	83.2	91.5	99.8	109.7	116.9
April	64.9	83.2	91.6	99.9	109.7	
May	65.0	83.4	91.7	99.7	110.0	
June	66.1	83.7	91.9	100.0	110.1	
July	66.8	83.7	92.1	100.1	109.9	
Aug.	67.0	83.7	92.1	99.8	109.6	
Sep.	68.0	84.4	92.5	100.5	110.2	
Oct.	68.3	85.1	92.6	100.6	110.6	
Nov.	68.5	85.1	92.6	100.7	111.0	
Dec.	68.8	84.9	92.6	100.8	113.1	
Jan.	69.8	85.3	93.1	101.2	113.3	
Feb.	70.5	85.5	93.4	101.4	114.6	
March	71.2	86.1	94.0	102.1	113.8	
April	71.4	86.4	94.0	102.4	112.9	
May	71.6	86.6	94.1	102.2	112.3	
June	71.9	86.6	94.4	102.3	112.0	
July	72.2	86.7	94.3	102.0	111.6	
Aug.	72.3	86.6	94.1	101.7	111.9	
Sep.	73.3	87.3	94.7	102.6	112.5	
Oct.	73.4	87.3	94.6	103.1	112.4	
Nov.	73.6	87.3	94.6	103.5	112.6	
Dec.	73.9	87.3	94.6	103.6	112.6	
Jan.	74.8	87.3	94.1	104.1	112.4	
Feb.	75.2	87.5	94.2	104.6	112.6	
March	76.3	88.3	94.6	104.7	113.1	
April	76.5	88.5	95.0	105.1	113.3	
May	76.7	88.6	95.1	105.1	113.4	
June	77.1	88.8	95.2	105.7	113.4	
July	77.1	88.9	95.6	105.4	113.3	
Aug.	77.0	88.7	95.5	105.3	113.0	
Sep.	78.0	89.1	95.9	106.2	113.7	
Oct.	78.1	89.2	96.3	106.3	114.0	
Nov.	78.1	89.3	96.3	106.8	114.0	
Dec.	78.0	89.2	96.2	106.7	113.8	
Jan.	78.6	89.5	97.0	107.6	113.6	
Feb.	78.9	89.8	97.3	108.4	113.7	
March	79.6	90.6	97.5	108.6	114.2	
April	80.0	90.8	97.4	109.1	114.8	
May	80.3	90.8	97.7	109.6	115.2	
June	80.7	90.9	97.9	109.7	115.3	
July	80.8	90.8	97.7	108.9	114.9	
Aug.	80.6	90.6	97.7	108.1	115.1	
Sep.	81.2	91.0	98.1	108.7	116.0	
Oct.	81.4	91.0	98.3	108.6	116.0	
Nov.	81.4	91.0	98.4	108.7	116.0	
Dec.	81.3	90.9	98.5	108.9	115.9	

Appendix 2: Batch.txt – 200 (synthetic) data from an industrial mixing process

1	941.1	51	980.0	101	1013.6	151	996.4

2	1006.9	52	970.9	102	1002.6	152	1054.2
3	1002.0	53	1030.0	103	975.3	153	975.1
4	1032.8	54	1021.8	104	1053.1	154	1017.0
5	965.0	55	952.8	105	1031.6	155	989.0
6	988.4	56	1040.9	106	956.2	156	1044.3
7	954.3	57	944.7	107	998.2	157	1002.1
8	1046.2	58	1040.1	108	1052.5	158	988.1
9	983.8	59	1024.4	109	970.7	159	1031.8
10	1007.6	60	957.7	110	1061.2	160	931.8
11	1005.9	61	1019.2	111	1021.0	161	1026.9
12	969.0	62	1036.3	112	1010.9	162	978.7
13	993.1	63	980.1	113	1013.2	163	1028.2
14	1004.2	64	962.0	114	967.4	164	933.1
15	1037.6	65	1025.4	115	1019.5	165	1058.3
16	1025.5	66	946.8	116	928.9	166	1058.7
17	912.5	67	1008.8	117	1060.1	167	890.4
18	1045.2	68	972.3	118	926.1	168	1043.6
19	961.7	69	1008.5	119	978.6	169	960.3
20	989.9	70	1055.5	120	1029.1	170	1045.8
21	1011.6	71	997.1	121	1032.0	171	960.3
22	934.2	72	1070.5	122	1020.2	172	1000.6
23	1033.5	73	921.8	123	1009.6	173	1008.8
24	970.2	74	1064.9	124	960.9	174	962.8
25	973.3	75	1000.2	125	1041.7	175	1042.8
26	1027.3	76	1005.3	126	932.8	176	1022.5
27	950.7	77	974.9	127	1072.8	177	1009.7
28	1002.4	78	984.3	128	968.1	178	980.3
29	934.8	79	956.8	129	1010.2	179	1036.3
30	1017.5	80	1066.1	130	989.6	180	990.3
31	1001.7	81	971.5	131	1045.7	181	946.3
32	967.1	82	1015.8	132	929.3	182	1035.7
33	1058.7	83	1003.8	133	1061.2	183	990.9
34	969.5	84	991.2	134	924.0	184	986.4
35	1033.6	85	974.5	135	993.6	185	1057.3
36	959.9	86	1049.0	136	1019.9	186	974.4
37	1007.1	87	975.9	137	1017.1	187	1019.8
38	928.8	88	1034.7	138	965.5	188	1002.5
39	1006.0	89	982.1	139	982.8	189	1087.1
40	932.9	90	986.2	140	1066.1	190	987.3
41	1021.8	91	1021.6	141	970.4	191	1036.5
42	992.8	92	950.9	142	1074.7	192	1003.0
43	1014.8	93	1039.1	143	1031.3	193	1054.3
44	956.4	94	965.6	144	959.5	194	990.4
45	1033.8	95	992.7	145	1037.3	195	997.6
46	980.2	96	1004.0	146	1014.4	196	1026.9
47	1000.4	97	995.5	147	1045.3	197	966.8
48	957.7	98	1002.0	148	981.6	198	1027.0
49	1093.3	99	1004.4	149	1048.7	199	959.6
50	975.8	100	1021.3	150	992.2	200	1057.1