Assignment – Module 4

1. Determine the autocorrelation at lag 3 for the data given below. Check if this correlation is significant at 95% confidence level.

   14,000 17,700 17,500 15,500 20,500 18,100 15,800 14,900 16,300
   14,900 17,600 17,000 17,300 18,300 19,100 17,900 19,400 22,900
   16,200 14,300

2. For the data given below, estimate the spectral densities for p=1,2 and 3 with notations followed in the lectures, for a maximum lag of 2

Year	1	2	3	4	5	6	7	8	9	10	11
Peak flow (m3/sec)	2160	3210	3070	4000	3830	978	6090	1150	6510	3070	3360

3. Statistical properties (in Mm^3) of streamflow at a site in the three seasons of a year are given below

	Season I	Season II	Season III
Mean	35	15	8
Std. Devaition	40	10	6
Lag one Correlation	0.43	0.67	0.5

   The lag one correlation is the correlation of flows with those of the previous season. Using a Non-stationary, First Order Markov Model, generate streamflow data for 3 years at the site. State the assumptions you make in using such a model.

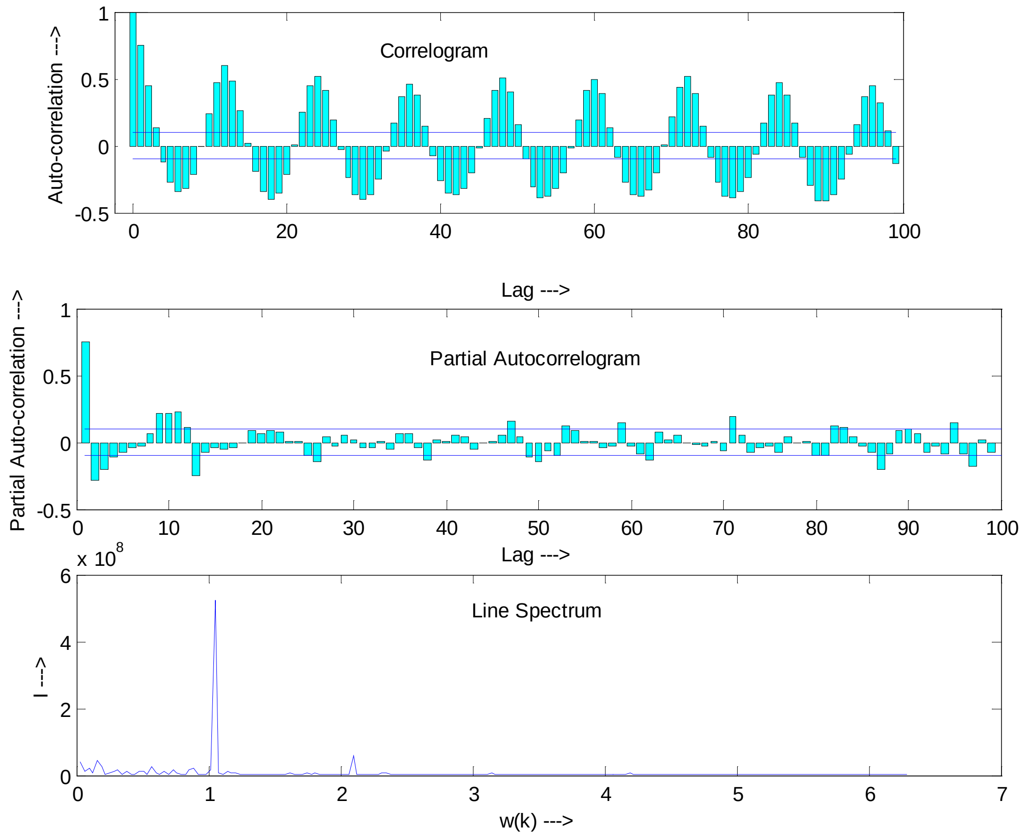
4. Given the auto-correlations, r_1 = -0.671 and r_2 = 0.463, obtain the initial estimates of parameters of an ARIMA (2,1,0) model using Yule Walker equations.

5. Express the following two models in the form ARIMA (p,d,q):

   a) $(1-0.2B)(1-B) X_t = (1 - 0.5B) e_t$

   b) $(1-B) X_t = (1 - 0.2B) e_t$

6. Monthly streamflow data for 33 years is analysed. The correlogram (autocorrelation function), partial autocorrelations and the line spectrum, with usual notations, are shown in the figure above. (a) Identify (determine) the periodicities in the data, (b) Which model of the ARMA family is best suited for this data ? Why?



7. Based on the following streamflow statistics (in appropriate units), generate 50 values of streamflow data using the first order, seasonal, Markov model. Compute and plot the correlogram and spectral density function for the generated data. ρ_j is the correlation with next month, j+1

Month, j	$\overline{x}_j$	s_j	ρ_j
Oct	5.02	2.31	0.61
Nov	6.50	3.38	0.58
Dec	7.33	3.23	0.50
Jan	6.42	2.95	0.31
Feb	5.35	2.62	0.38
Mar	5.02	1.66	0.37
Apr	6.42	1.80	0.44
May	10.70	2.89	0.34
Jun	12.76	3.32	0.17
Jul	9.05	3.26	0.65
Aug	4.44	1.47	0.93
Sep	3.29	1.22	0.51

8. Monthly streamflows (in Mm^3) at a location for the year 2010-11 (upto March 2011) are given below

Year :2010

Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
19	15	39	102	90	29	90	46	30	66	80	89

Year : 2011

Jan	Feb	Mar
82	17	26

Using a double moving average of order 4, obtain the forecast of streamflow for April 2011.

9. The table below gives a time series composed of 60 values. Plot this series to identify its trend. Compute the first differences and plot the resulting series. (a)Plot the autocorrelations upto 15 lags for both the original and the differenced series. Also determine the first 3 partial auto correlations for the original data. (b) Plot the power spectrum of the original and the differenced data

Period	Observation	Period	Observation	Period	Observation
1	9.56	21	60.50	41	85.28
2	12.48	22	63.29	42	84.44
3	13.64	23	66.55	43	86.59
4	18.80	24	68.65	44	88.05
5	25.04	25	72.66	45	90.83
6	30.33	26	71.25	46	93.05
7	34.08	27	65.48	47	94.65
8	40.10	28	62.68	48	96.66
9	42.40	29	56.60	49	96.30
10	41.36	30	49.90	50	96.09
11	39.25	31	49.82	51	99.27
12	38.20	32	51.87	52	104.77
13	41.47	33	57.74	53	105.51
14	46.14	34	58.24	54	105.19
15	52.62	35	58.31	55	109.16
16	59.01	36	59.91	56	110.78
17	60.20	37	62.61	57	115.77
18	58.53	38	69.07	58	122.75
19	56.98	39	77.36	59	126.85
20	57.82	40	80.39	60	132.57

10. The inflow data (in million cubic feet) to a reservoir is given below

Year	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Jan	Feb	Mar	Apr	May
1940-41	29,243	87,671	77,982	14,465	12,527	23,769	5,976	590	2,391	1,245	1,339	3,039
41-42	29,019	76,623	56,931	18,032	13,877	2,633	3,616	580	731	258	792	1,644
42-43	14,916	97,800	56,286	24,710	13,262	3,714	3,957	2,253	897	588	785	6,690
43-44	17,800	107,200	24,321	29,113	38,651	12,887	3,717	1,781	1,264	1,148	610	1,851
44-45	4,665	78,799	34,444	9,453	12,560	8,985	4,078	1,816	959	562	836	950
1945-46	5,860	71,985	31,280	17,474	12,322	4,454	2,153	1,446	1,162	925	1,006	996
46-47	20,720	77,108	113,153	27,862	14,683	16,006	6,914	1,259	1,068	766	1,321	1,112
47-48	1,745	55,929	58,237	39,271	21,552	2,236	3,063	1,548	952	275	1,759	3,725
48-49	16,421	55,903	96,601	23,747	15,347	7,138	4,026	1,622	841	727	731	1,870
49-50	13,942	44,977	51,101	24,715	15,400	4,822	2,727	1,831	1,346	482	78	789
1950-51	7,855	91,462	44,170	49,455	14,721	6,409	2,176	1,608	920	432	1,157	2,087
51-52	15,052	55,179	44,183	16,420	20,326	5,104	2,321	722	952	609	623	1,050
52-53	9,430	36,096	55,821	10,194	24,359	2,914	4,109	1,701	865	720	696	260
53-54	9,658	104,243	100,664	15,400	35,493	3,509	1,446	677	198	168	603	4,314
54-55	20,101	80,172	74,683	22,160	22,692	2,133	3,549	1,718	610	248	879	8,858
1955-56	23,448	26,032	25,025	27,574	29,877	8,225	4,114	2,030	1,163	584	848	2,748
56-57	28,459	87,026	57,194	14,943	27,650	20,217	2,463	341	529	535	469	5,178
57-58	16,196	83,511	41,085	9,148	11,072	14,093	3,159	1,469	1,059	634	1,493	5,670
58-59	14,394	131,338	51,913	31,080	18,479	7,267	3,441	1,334	896	304	688	1,552
59-60	26,616	196,988	54,466	52,557	13,280	6,893	3,481	1,337	582	818	1,988	3,325

Develop candidate ARIMA models and select the best model for the data series based on following methods.

- Maximum likelihood (ML)
- Minimum mean square error (MMSE)

Validate the models using following tests

- Significance of residual mean
- Significance of periodicities
- Cumulative periodogram test or Bartlett’s test
- White noise test
  - Whittle’s test
  - Portmanteau test