Solutions Exercise 3 - Forecasting

Inventory Management

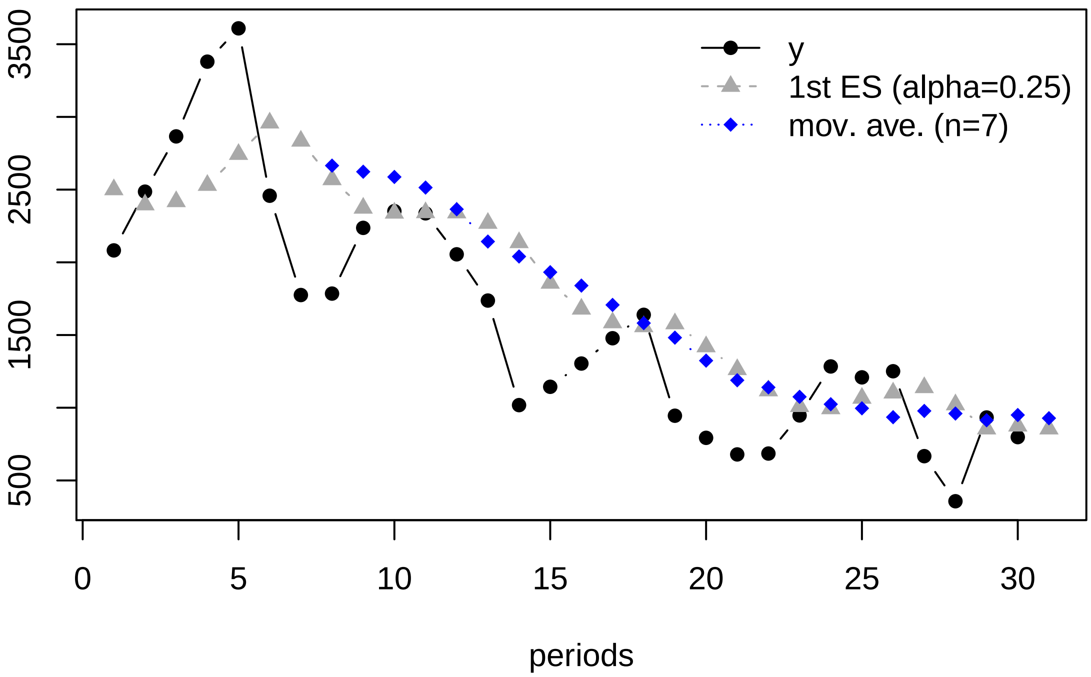
Thomas Kirschstein

Forecasting for dense time series (I)

1. Apply 1st-order exponential smoothing (with $\alpha = 0.25$ and initial forecast $2,500$) and moving averages (with $n = 7$).

period	y	yhat.moving.ave	yhat.1es
1	2082	NA	2500
2	2486	NA	2396
3	2866	NA	2418
4	3380	NA	2530
5	3609	NA	2743
6	2458	NA	2959
7	1775	NA	2834
8	1785	2665	2569
9	2237	2623	2373
10	2352	2587	2339
11	2337	2514	2342
12	2055	2365	2341
13	1737	2143	2269
14	1018	2040	2136
15	1144	1932	1857
16	1304	1840	1679
17	1478	1707	1585
18	1639	1582	1558
19	945	1482	1578
20	793	1324	1420
21	679	1189	1263
22	685	1140	1117
23	947	1075	1009
24	1284	1024	994
25	1209	996	1066
26	1251	935	1102
27	667	978	1139
28	357	960	1021
29	933	914	855
30	798	950	875
31	NA	928	855

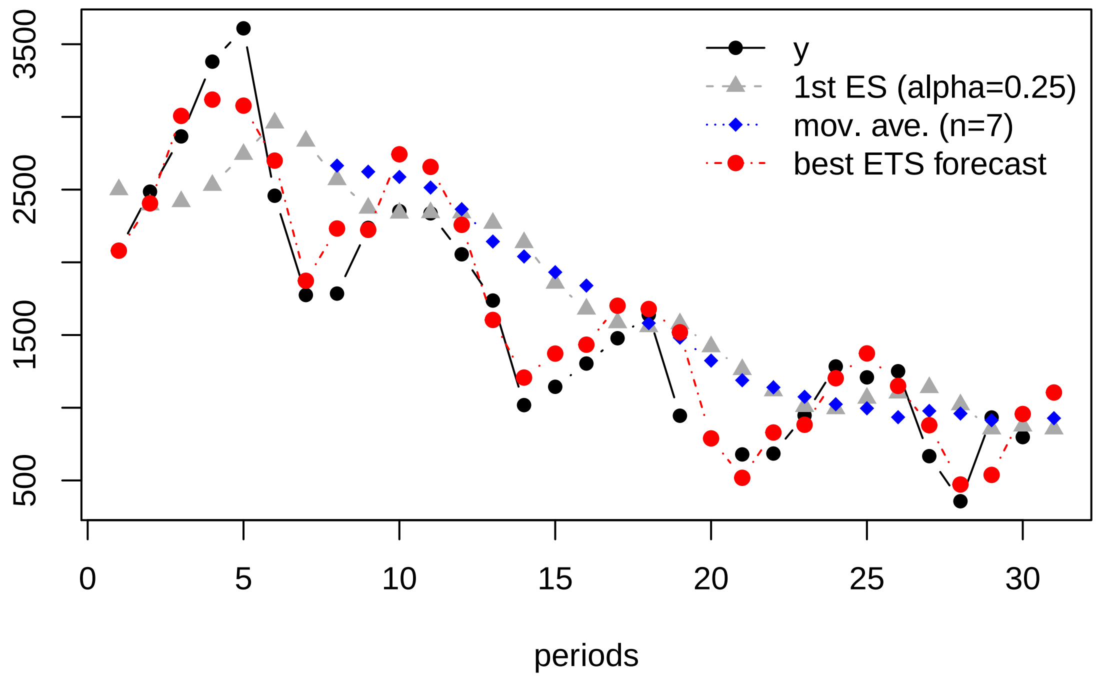
2. Plot the time series. Which structural properties do you observe? Are the aforementioned forecasting models appropriate to forecast the time series?



The time series shows a declining trend and seasonality pattern with a 7 period rhythm. Thus, neither 1st order exponential smoothing nor the moving average method are appropriate forecasting techniques.

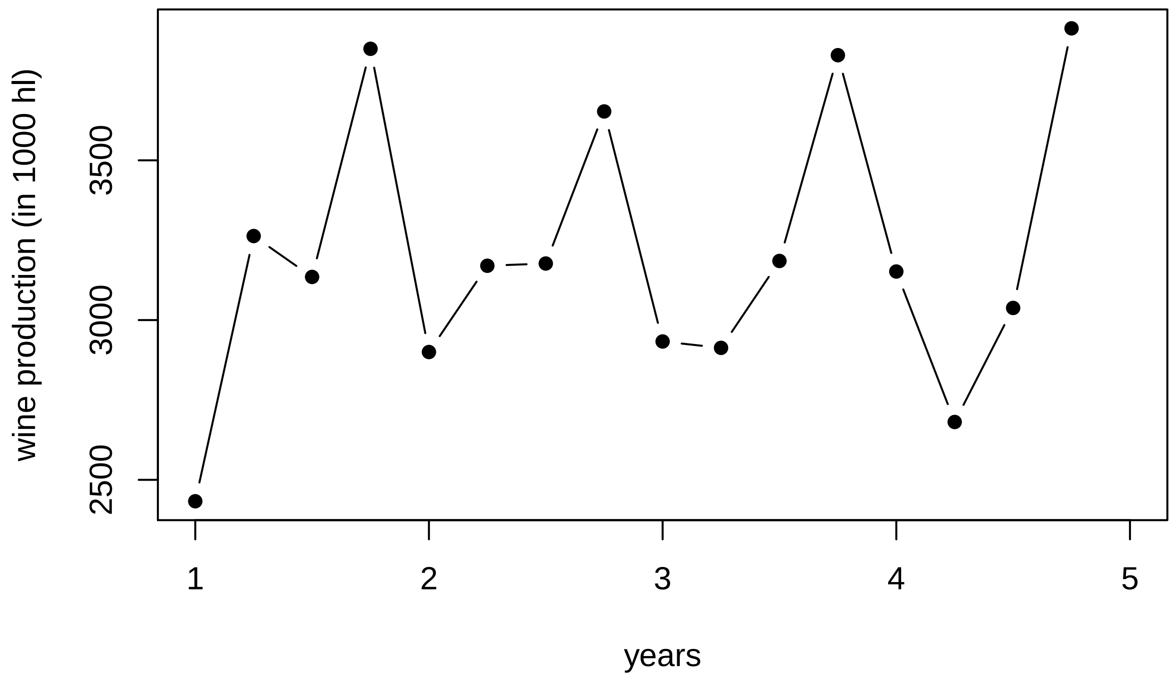
3. Do you recognize the time series?

The time series shows the daily number of persons newly infected with SARS-CoV2 from April 14th till May, 13th 2020 in Germany.



Forecasting for dense time series (II)

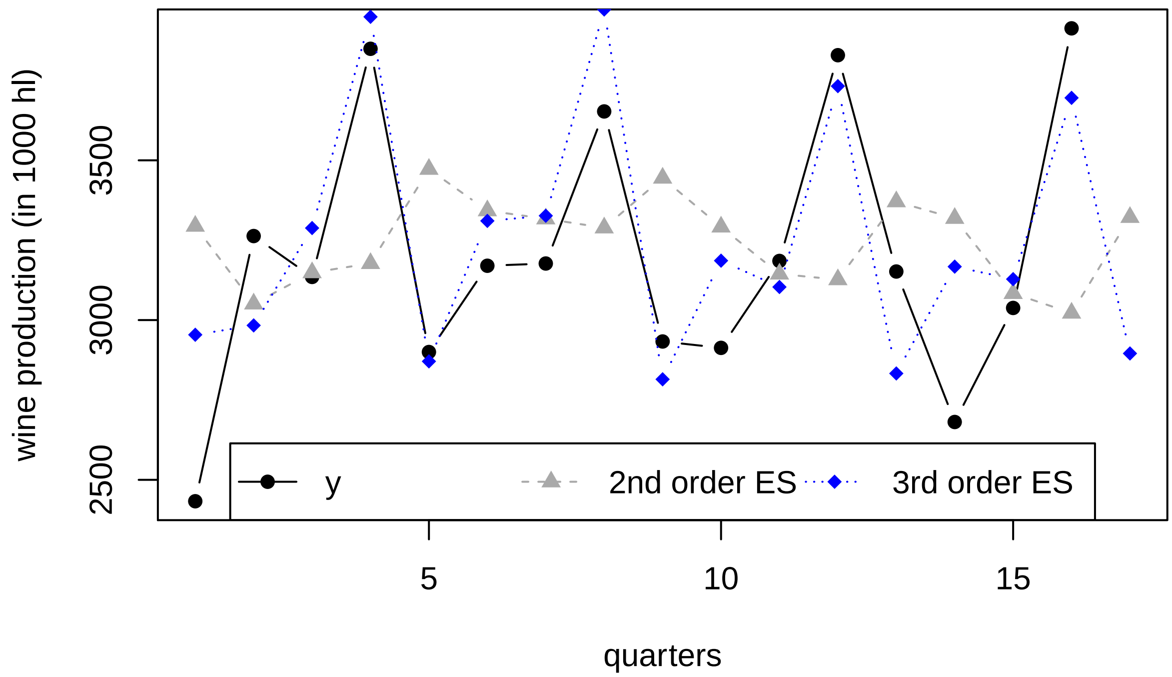
1. Display the time series graphically. Which typical pattern becomes obvious in the time series?



Stationary time series with seasonal fluctuations.

2. Calculate forecasts for the time series with 2nd-order and 3rd-order exponential smoothing (with $\alpha = \beta = \gamma = 0.3$). Display the forecasts in the same diagram. *Hint*: Initialize the forecasts with the average over all observations for a_0, the average slope ($b_0 = \frac{y_{16} - y_1}{16}$, and $c_{-3:0} = (-340, -200, -50, +600)$)

period	y	yhat.2es	yhat.3es
1	2433	3294	2954
2	3263	3051	2983
3	3135	3149	3288
4	3849	3177	3949
5	2900	3472	2871
6	3170	3342	3310
7	3177	3317	3326
8	3653	3289	3972
9	2933	3445	2815
10	2913	3292	3186
11	3185	3144	3103
12	3829	3127	3732
13	3152	3371	2833
14	2681	3319	3167
15	3038	3084	3128
16	3913	3022	3695
17	NA	3322	2896



3. Compare the forecasts' accuracies by an appropriate measure. Which forecasting method should be chosen for the time series?

We calculate root mean squared forecasting error (RMSFE) and the root median squared logarithmized accuracy ratio (RMSLAR).

period	y	yhat.2es	yhat.3es	squ.FE.2ES	squ.FE.3ES	squ.LAR.2ES	squ.LAR.3ES
1	2433	3294	2954	741321	271441	0.0917954	0.0376496
2	3263	3051	2983	44944	78400	0.0045128	0.0080492
3	3135	3149	3288	196	23409	0.0000199	0.0022706
4	3849	3177	3949	451584	10000	0.0368164	0.0006579
5	2900	3472	2871	327184	841	0.0324072	0.0001010
6	3170	3342	3310	29584	19600	0.0027918	0.0018677
7	3177	3317	3326	19600	22201	0.0018596	0.0021007
8	3653	3289	3972	132496	101761	0.0110177	0.0070092
9	2933	3445	2815	262144	13924	0.0258882	0.0016862
10	2913	3292	3186	143641	74529	0.0149602	0.0080251
11	3185	3144	3103	1681	6724	0.0001679	0.0006803
12	3829	3127	3732	492804	9409	0.0410182	0.0006584
13	3152	3371	2833	47961	101761	0.0045121	0.0113851
14	2681	3319	3167	407044	236196	0.0455710	0.0277539
15	3038	3084	3128	2116	8100	0.0002258	0.0008523
16	3913	3022	3695	793881	47524	0.0667631	0.0032860

Based on these values the RMSFE for 2nd order ES and 3rd order ES is 494 and 253, respectively. Likewise, the corresponding RMSLAR values are 0.114 and 0.047. Thus, both accuracy measures imply that (as expected) 3rd order ES delivers more accurate forecasts.

Forecasting sporadic time series (intermittent demands)

1. Calculate forecasts for the year 2008 at the of Dec. 2007.

First, we estimate the expected non-zero demand by calculating the mean of the non-zero demand values such that $\mu = 3.5$. Second, the distribution of interarrival times based on the time series is derived:

```
##
##   0   1   2   4   5   7
## 0.2 0.4 0.1 0.1 0.1 0.1
```

Thus, at the end of 2007, for the first 7 months the following expected demands can be forecasted (ignoring the last two zero-demand periods in 2007):

month	forecast
Jan	0.70
Feb	1.40
Mar	0.35
Apr	0.00
May	0.35
Jun	0.35
Jul	0.00
Aug	0.35

If we take the last two zero-demand periods in 2007, we know that the inter-arrival time is at least 2, such that the forecasts change to:

month	forecast
Jan	0.875
Feb	0.000
Mar	0.875
Apr	0.875
May	0.000
Jun	0.875

2. Assume that in the first 3 months of 2008 no lubricants are sold. What are the forecasts for the next 3 months in 2008?

Here, we have to calculate the expected demand under the condition that inter-arrival time is at least 3 (again ignoring the last two zero-demand periods in 2007). The conditional inter-arrival time distribution is then

```
##
##         4         5         7
## 0.3333333 0.3333333 0.3333333
```

Thus, for April to August, the following expected demands can be forecasted:

month	forecast
Apr	0.00
May	1.17
Jun	1.17
Jul	0.00
Aug	1.17

Taking into account the zero-demand periods at the end of 2007, inter-arrival is at least 5 leading to the following estimates

```
##
##   5   7
## 0.5 0.5
```

month	forecast
Jun	1.75
Jul	0.00
Aug	1.75

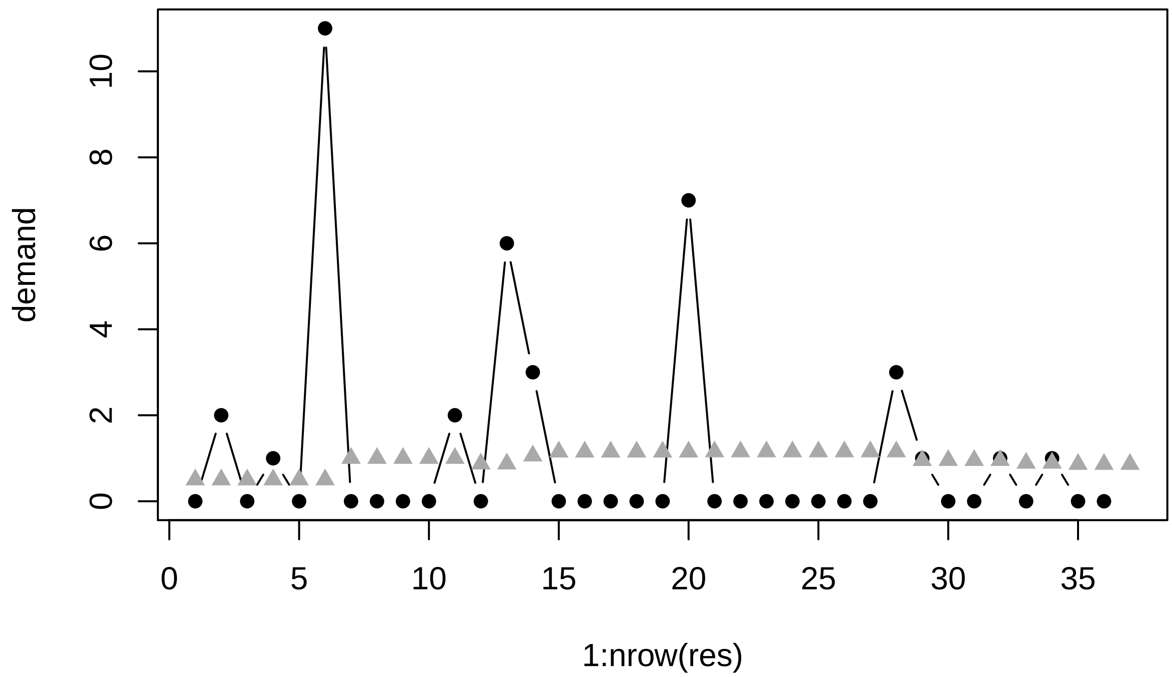
3. Another approach to forecast sporadic tiem series is proposed by Croston (see here for an brief introduction). Calculate forecasts by based on Croston's method for the first 3 month of 2008 either by hand or by using R.

Croston's forecasting method is basically a 1st order exponential smoothing technique applied to the non-zero demand time series and the inter-arrival time series. Thus, first both time series must be created and afterwards a 1st order ES is applied. Here we set the smoothing parameter to $\alpha = 0.1$ and initialize the estimate with the first observations of each time series. Note that Croston calculates the interarrival time as the difference in the indices of two subsequent periods with non-zero demand. Thus, the inter-arrival times are higher by 1 compared to the inter-arrival times used above.

	int.arr	nz.dem	fc.int.arr	fc.nz.dem	fc.dem
0	NA	NA	NA	NA	NA
1	2	1	2.000	1.000	0.500
2	2	11	2.000	1.000	0.500
3	5	2	2.000	2.000	1.000
4	2	6	2.300	2.000	0.870
5	1	3	2.270	2.400	1.057
6	6	7	2.143	2.460	1.148
7	8	3	2.529	2.914	1.152
8	1	1	3.076	2.923	0.950
9	3	1	2.868	2.730	0.952
10	2	1	2.881	2.557	0.888
11	NA	NA	2.793	2.402	0.860

Extended to the original time series yields the following forecasts can be deduced:

y	forecast
0	0.500
2	0.500
0	0.500
1	0.500
0	0.500
11	0.500
0	1.000
0	1.000
0	1.000
0	1.000
2	1.000
0	0.870
6	0.870
3	1.057
0	1.148
0	1.148
0	1.148
0	1.148
0	1.148
7	1.148
0	1.152
0	1.152
0	1.152
0	1.152
0	1.152
0	1.152
0	1.152
3	1.152
1	0.950
0	0.952
0	0.952
1	0.952
0	0.888
1	0.888
0	0.860
0	0.860
NA	0.860



Note that the forecasts based on Croston's method do not change until a new non-zero demand occurs.