Math 625.714
Stochastic Control and Forecasting
Final Project Submission

Richard N. Shea

August 21, 2018

Introduction

In this project I've taken retail sales data from the well-known Rossman Store items used in Kaggle [1] competitions and apply a Dynamic Harmonic Regression model as discussed in Young [2] and McElroy [3]. Thus, there are two main task at hand. Filter non-stationary data and work our way up to the Harmonic modelling.

Also, after we have our forecast, we then need some inventory control, hopefully of the HJB type [4], which will be discussed in its own entity.

Part I. Non-stationary Data

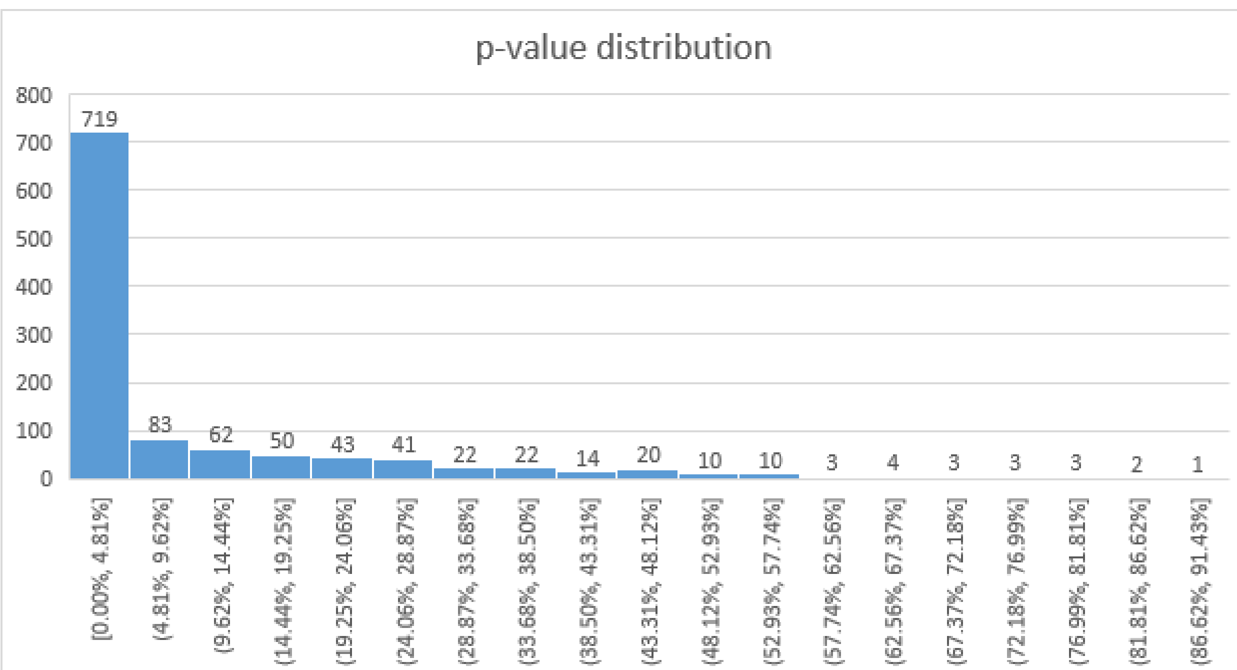
To filter the data, I used a Priestley-Subba Rao (PSR) Test [5] in R to run a custom search for nonstationary data. See Abraham & Thavaneswaran [6] for Kalman filtering & PSR on missing observations (even when a store is closed, say, every Saturday).

The PSR test centers around the time-varying Fourier spectrum f(t,w) where t is time and w is frequency. For a stationary time series, the time-varying spectrum is, not surprisingly, a constant function of time. The PSR test investigates how "non-constant" f(t,w) is a function of time. It does this by looking at the logarithm of an estimator of f(t,w), i.e. obtaining Y(t, w) = log{F(t,w)}, and the variance of Y(t, w) is approximately constant. Here the logarithm acts as a variance stabilizer permitting us to focus on changes in the mean structure of Y. These actions permit us to write Y(t, w) as a linear model with constant variance and test the constancy of f using a standard one-way analysis of variance (ANOVA). [7]

For the PSR steps (1115 stores with %5 error);

1. Take the time series and subtract by the sample mean (this is re-centering).
2. Then segment into blocks (default is 9)
3. spectral density function then estimates (average of orthogonal sinusoidal tapers).
4. Then we ask how homogeneous are these (SDF) estimates?

For simplicity, I didn't take account of weak stationary p-values. Observe the p-distribution below



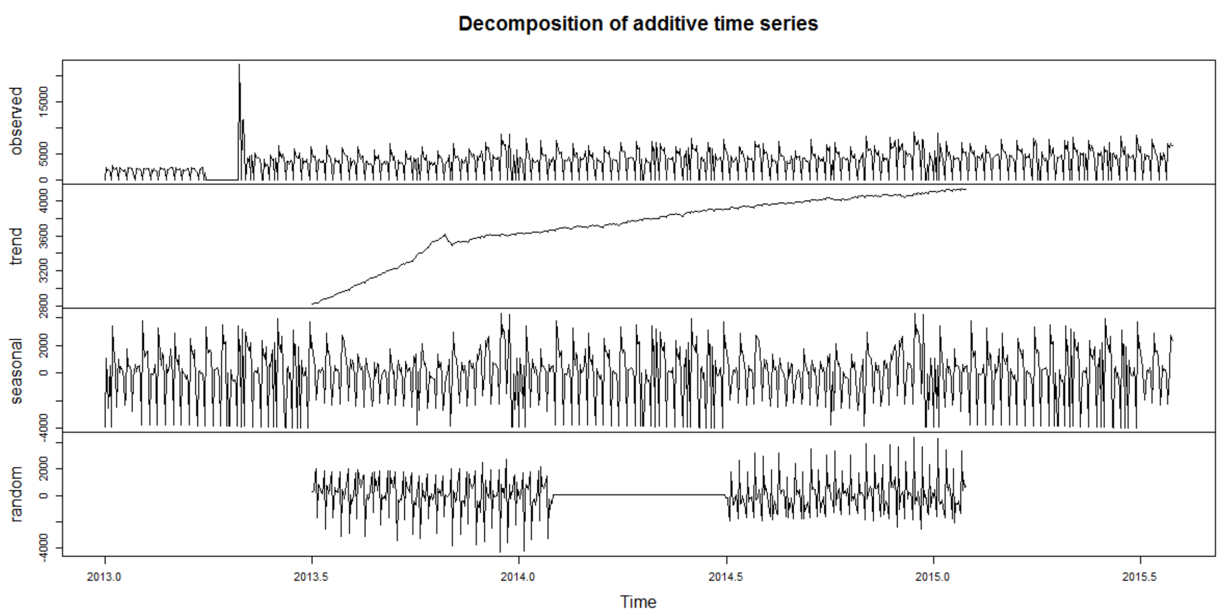
(*1 - significance probability*) that the data is nonstationary. Therefore, we will filter for p = 0 to avoid possible weak stationary stores.

Example:

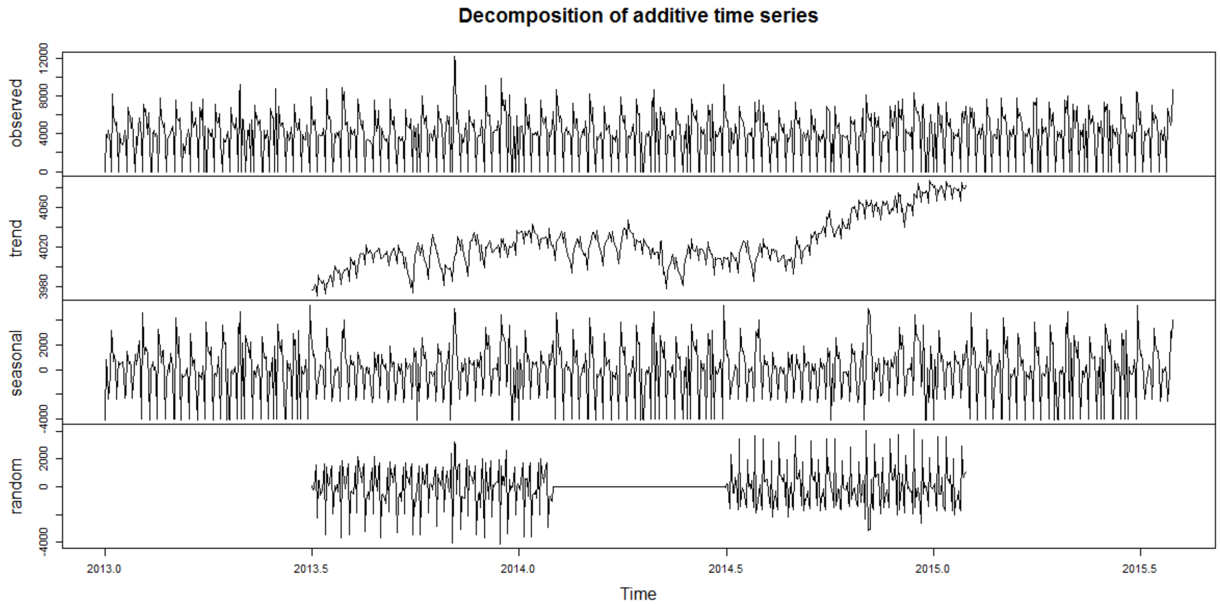
```
Priestley-Subba Rao stationarity Test for store.xts
----------------------------------------------------
Samples used               : 942
Samples available          : 936
Sampling interval          : 1
SDF estimator              : Multitaper
  Number of (sine) tapers  : 5
  Centered                 : TRUE
  Recentered               : FALSE
Number of blocks           : 9
Block size                 : 104
Number of blocks           : 9
p-value for T              : 0
p-value for I+R            : 1.943193e-06
p-value for T+I+R          : 0
```

Below is an example of two stores with completed opposite P values. Observe the Trend component.

Store 105 (p = 0)



Store 1047 (p = 0.91423)



Observe that is over the whole-time domain. We could probably find nonstationary segments of Store 1047, but I wanted only a handful of stores from the Rossman data and had to make a reasonable decision.

There are 14 stores with zero p-values.

Store	p-value
105	0
126	0
136	0
183	0
192	0
349	0
364	0
589	0
663	0
700	0
708	0
837	0
897	0
909	0

Pick top and bottom store to send to CAPTAIN

Store	Sales
909	6512493
349	5711451
192	5608471
126	5069540
663	3623173
589	3537175
364	3510451
136	3414045
105	3348791
700	3297857
708	3145529
183	2945836
897	2717511
837	2608765

Part II. CAPTAIN

Here is an attempt to model daily data, an example that has not been found in Young's DHR work.

From CAPTAIN_TVPMOD manual [8]:

The DHR model is similar to a Fourier analysis, but with coefficients that evolve smoothly in time. The model is,

$$y_t = S_t + e_t = \sum_{j=0}^{\left[\frac{s}{2}\right]} \{a_{jt}\cos(\omega_j t) + b_{jt}\sin(\omega_j t)\} + e_t \tag{3.8}$$

with,

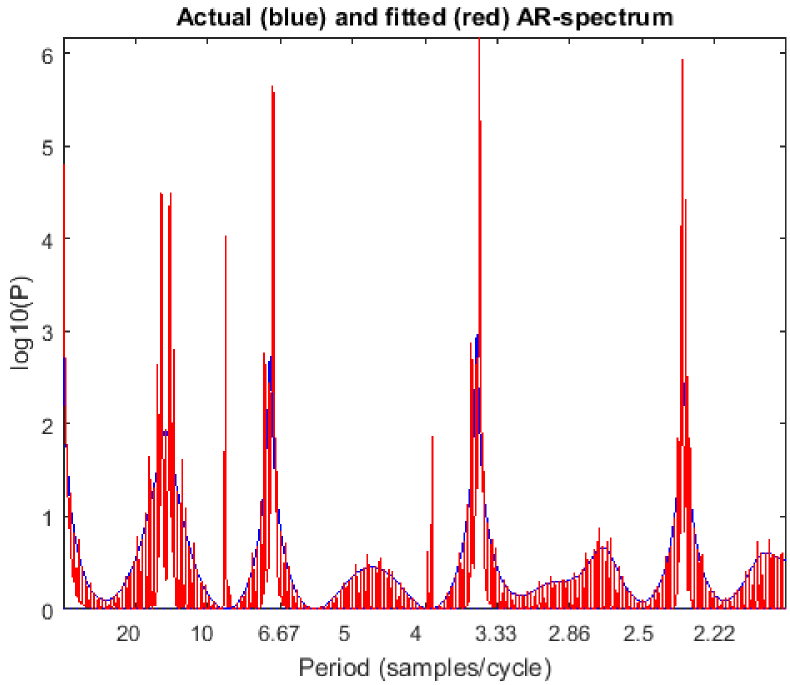
$$\omega_j = \frac{2\pi j}{s} \qquad j = 1,\ 2,\ \ldots,\ \left[\frac{s}{2}\right] \tag{3.9}$$

and,

$$\begin{pmatrix} a_{jt} \\ a'_{j1t} \end{pmatrix} = \begin{pmatrix} \alpha_j & \beta_j \\ 0 & \gamma_j \end{pmatrix} \begin{pmatrix} a_{jt-1} \\ a'_{jt-1} \end{pmatrix} + \begin{pmatrix} \eta_{jt} \\ \eta'_{jt} \end{pmatrix} \qquad \mathrm{NVR}(\eta_{jt}) = \mathrm{NVR}(\eta'_{jt}) \tag{3.10}$$

The focus will be on (3.9). That is, for s = 365, we choose our [s/2] = [(s-1)/2] = [(365 – 1)/2] = 182 since 365 is odd.

The dhropt() for NVR values works nicely and was able to fit the AR spectrum as seen below



However, when plugging these NVR values into the dhr() function, it fails. It doesn't appear to handle large values of s. This is most likely due to memory issues.

The conclusion is that the dhr.m algorithm can't handle (s/2) > 60. I'm still waiting to hear back from Young on this (I sent him my error), but I believe for memory issues, it was designed to handle weekly or monthly data. That is max[s/2] = 52. The manual doesn't state if it handles a multi-seasonal vector input. This would allow for a faster partition breakdown period search. We now attempt a DHR in R.

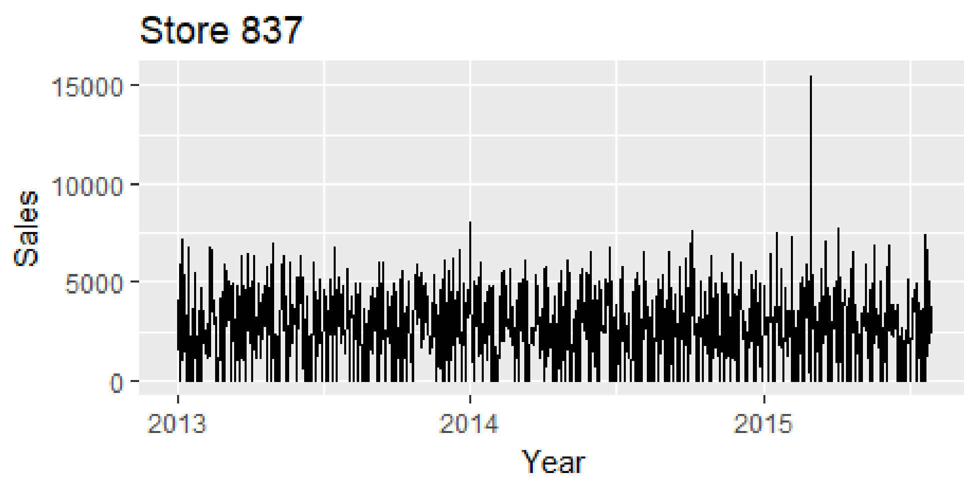
Part III. R

In R we must covert the data to a univariate time series objects because just about every package won't read the data as a frame. For this, we mainly use Hyndman's R forecasting packages [9]

We start up with some naïve forecasting and build up a noise model and then attempt to transition into a multi-seasonal model to handle the daily periods.

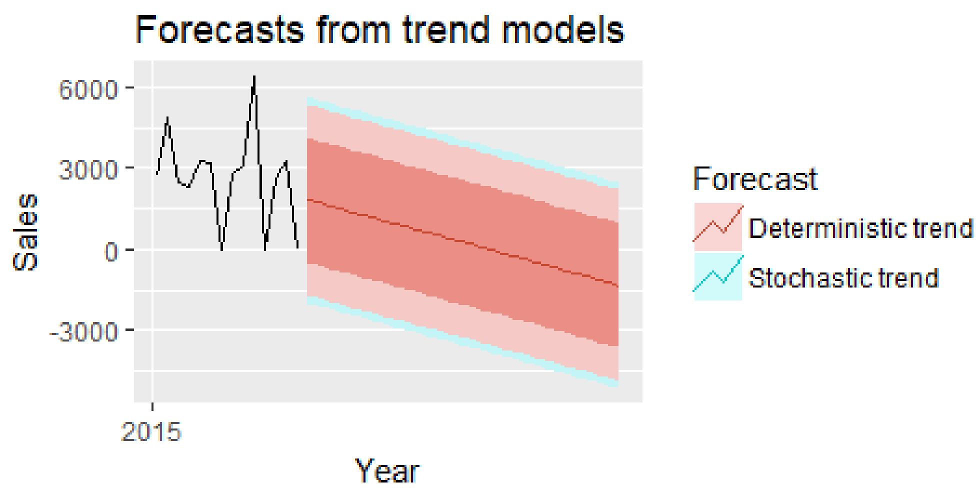
We model trend deterministically and stochastically.

Example for Store 837:



Deterministically vs. Stochastically:

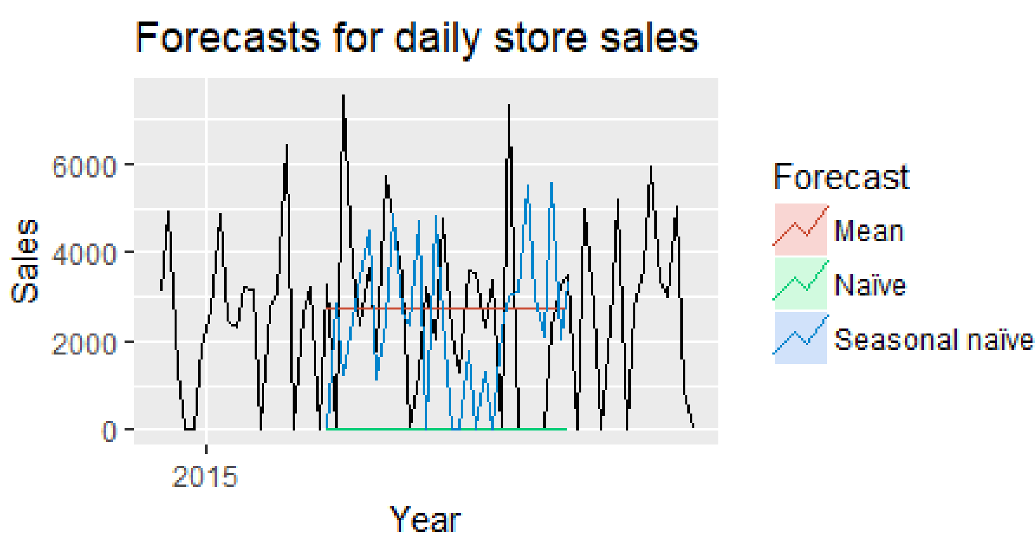
We create a test window from 2013 to the beginning of 2015. We then look at a 30-day window projected through the rest of the data set. Observer that the Stochastic trend (should) have wider prediction intervals because their errors (assumed) to be also non-stationary.



This forecast is of course a poor and conservative one that clearly won't meet customer demand and produce stock outs.

Next, the attempt is to use a mean, random walk (zero drift) and seasonal onto this Store. We can label all three naïve forecasting to mimic the interpolation of a time slice.

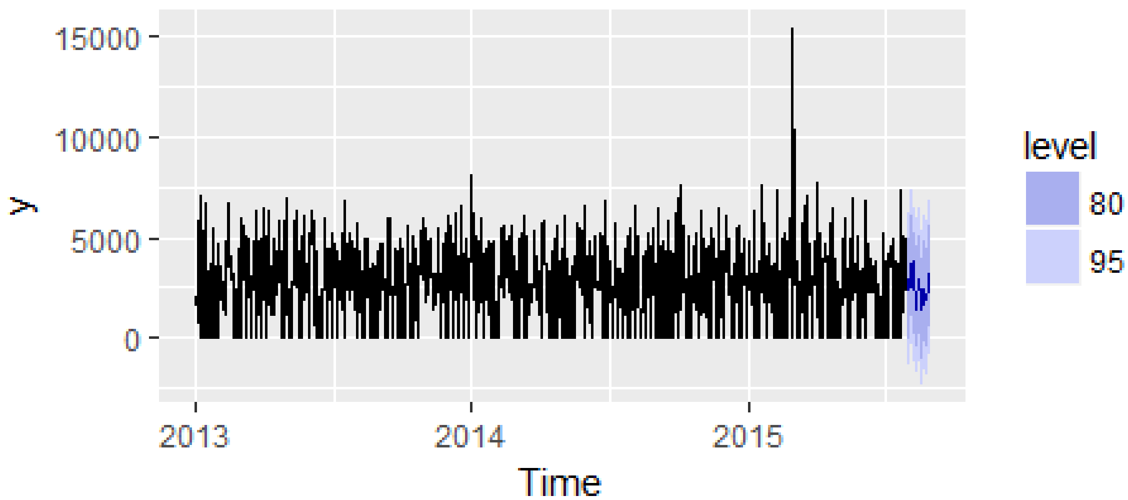
Observe below with metrics.



Store 837			
	MAE	MPE	MAPE
mean	507.22	15.53	15.53
naïve	2167.88	100	100
Seasonal naïve	3269	100	100

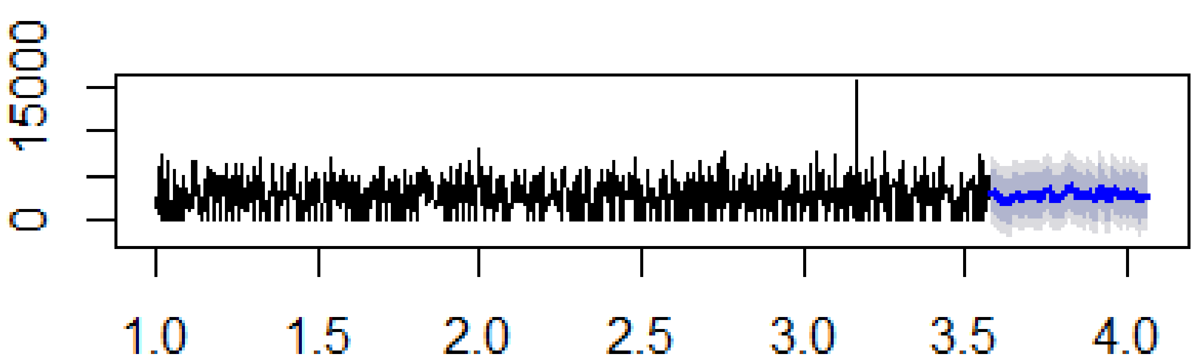
Clearly the mean forecasting function wins in the naïve sense.

We now try a more Fourier (random) fit in R:



The MAE drops to 1471.403, but nowhere near the mean forecast performance.

Another attempt is to try to fit the data using a multi-seasonal time series object.



The MAE stays around 1502.94, but the computational time drops to seconds. Overall, a Moving average model would be the best univariate model to use in R for the Rossman data.

Part III. Control

The idea for control on the Rossman data is to introduce a common deterministic control method and add noise in some way using R.

In inventory management, economic order quantity (EOQ) is the order quantity that minimizes the total holding costs and ordering costs. The model was developed by Ford W. Harris in 1913, but R. H. Wilson, a consultant who applied it extensively, and K. Andler are given credit for their in-depth analysis [10].

We want to determine the optimal number of units to order so that we minimize the total cost associated with the purchase, delivery and storage of the product.

Economic Order Quantity

$$Q^* = \sqrt{\frac{2DK}{h}}$$

Where D is annual demand quantity, K is the fixed cost per order, and h is the holding cost per unit.

We use a package called 'SCperf' [11].

We can enter in the demand, holding cost, and penalty shortage.

Example:

d = 4000 #Let our mean demand be 4,000 from our forecast

c = 300000 # Total holding cost

hc = 0.3

p = 100 # zero shortage penalty.

EOQ(d,c,hc,p)

Q T TVC

89443 22 26833

Thus, our order quantity is 89443 for the year. With 22 days between ordering and a total variable cost of $26883.

We then can use the stochastic newsboy() function to add a normal SD term to the function.

When the demand is a random variable with normal distribution, the optimal stocking quantity that minimize the expected cost is: $Q = m + z * sd$, where z is known as the safety factor and $Q - m = z * sd$ is known as the safety stock.

Inputs:

m Mean demand during the period s

sd Standard deviation of demand during the period

p The selling price, where $p > c$

c The unit cost

s The salvage value (default:0), where $s < c$

We get the following outputs:

Q Optimal order-up-to quantity

SS Safety stock

ExpC Expected cost

ExpP Expected profit

CV Coefficient of variation of the demand

FR Fill rate, the fraction of demand served from stock
z Safety factor

However, for the Rossman data, we don't have enough model fit to assume normality.

With made up date we have the following example:

m = 100 # mean demand

sd = 30 #

p = 4 # selling price

c = 1 # unit cost.

```
> Newsboy(m,sd,p,c)
```

```
     Q      SS    ExpC    ExpP      CV      CR      FR       z
120.23   20.23   38.13  261.87    0.30    0.75    0.96    0.67
```

Observe that SS is what we are really after on the fulfillment side.

This R package was designed for instructional use, and real-world inventory packages are custom made by software consultants such as JDA [12].

Part IV: Conclusion

Daily non-stationary data needs to handle multi-period fitting. This is assuming we don't want to apply differencing. With CAPTAIN it may be worth investigating aggregating weekly data, but then we need to handle the philosophical problem of a "true" PSR test for non-stationary data.

I think it's worth adding in more components (since Rossman can supply this on daily basis also) to the model. With daily sales, there is just too much noise to justify needing more feature (detection). See [13].

The Control part was its own entity to grab a concept from Oksendal [14] or Young for learning purposes. See [15].

Overall this project has been a great experience in hands on learning of stochastic forecasting.

References

[1] – Kaggle Data Science competitions: www.kaggle.com

[2] - Young_et_al-1999-Journal_of_Forecasting

[3] - Modeling of Holiday Effects and Seasonality in Daily Time Series – Tucker S. McElroy

[4] - Deterministic and stochastic optimal inventory control with logistic stock-dependent demand rate Control - A. Tsoularis - Int. J. Mathematics in Operational Research, Vol. 6, No. 1, 2014

[5] - PSR https://www.rdocumentation.org/packages/aTSA/versions/3.1.2/topics/stationary.test

[6] - A NONLINEAR TIME SERIES MODEL AND ESTIMATION OF MISSING OBSERVATIONS - Abraham & Thavaneswaran- Ann. Inst. Statist. Math.

[7] - University of Bristol – PSR Test: https://people.maths.bris.ac.uk/~magpn/Research/LSTS/TOS.html

[8] - Guide to TVPMOD - Handbook Part 2, 3 July 2017, C J Taylor, D J Pedregal and P C Young

[9] - Forecasting: Principles & Practice Time Variable Parameter Models - Rob J Hyndman, www.datacamp.com

[10] - Economic Order Quantity: https://en.wikipedia.org/wiki/Economic_order_quantity

[11] - 'SCperf' Inventory Control Package (R): https://cran.r-project.org/web/packages/SCperf/SCperf.pdf

[12] - JDA: Inventory Optimization: https://jda.com/solutions/manufacturing-distribution-solutions/manufacturing-planning/inventory-optimization

[13] - Forecasting Rossmann Store Leading 6-month Sales; Sen Lin, Eric Yu, Xiuzhen Guo

[14] - Stochastic Differential Equations; Bernt Øksendal – Springer. ISBN-10: 3540047581

[15] - Porteus E. L. (2002) Foundations of Stochastic Inventory Theory, Stanford University Press, Stanford, CA.