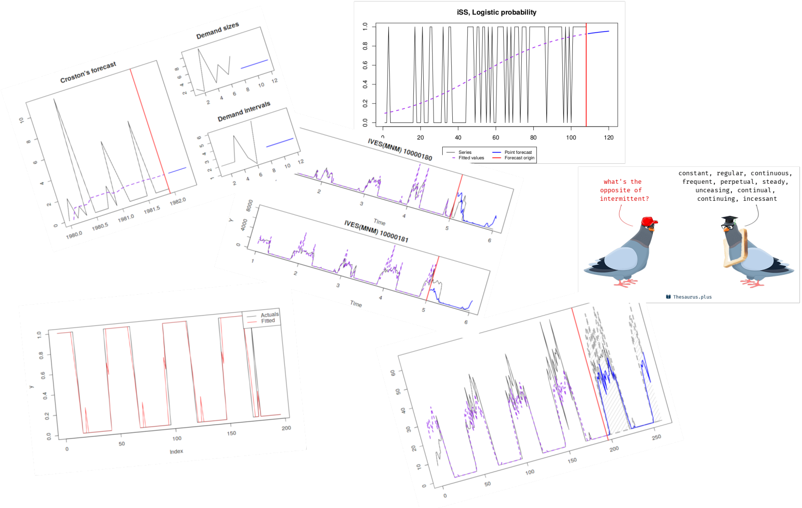
Intermittent state space model saga



The original poster of "Intermittent state space model" saga

Previous episodes (ISF2017)...

The original poster of Star Wars, Episode IV

Previous episodes (ISF2018)...

The original poster of Star Wars, Episode V

And now... a spin-off!

All images are from "Family Guy"

What about those sweet melons? Using mixture models for demand forecasting in retail

Ivan Svetunkov

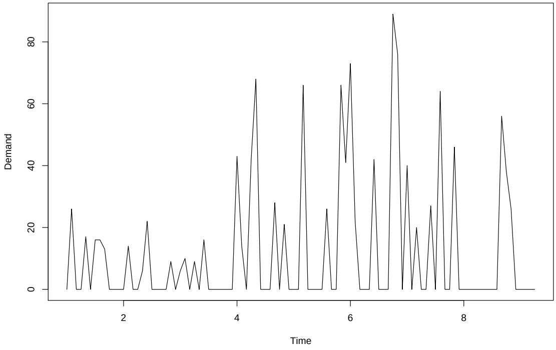
ISF2019

18 June 2019

Marketing Analytics and Forecasting
Lancaster University Management School

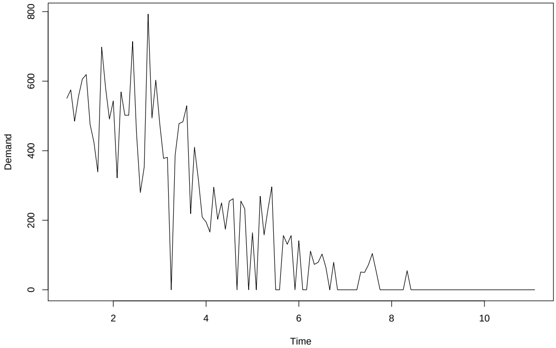
Introduction

A typical intermittent demand in wholesale looks like:



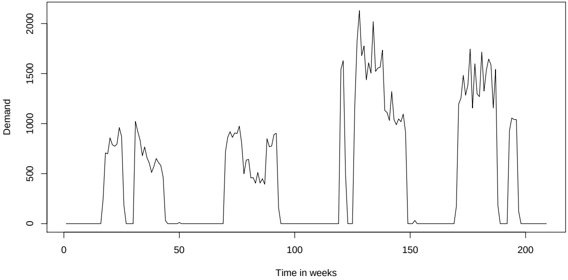
Introduction

In retail we might get trends, i.e. demand obsolescence:



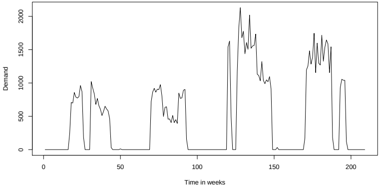
Introduction

But there are products like this...



Is this intermittent? Is this seasonal? Maybe both? Trends?

Introduction



It can be considered as intermittent, as demands occur at random.

The probability of occurrence p_t is high in some weeks.

But it is low in the others.

If we know, when the next demand will happen, we can set $p_t = 1$

What might drive the demand in this case?

Splitting the demand into two parts... $y_t = o_t z_t$

Demand sizes:

- Seasonality;
- Promotional activities;
- Prices.

Demand occurrence:

- Seasonality;
- Promotional activities;
- Prices.

Mixture distribution model

Mixture distribution model

This is not a new idea, it has been known long in statistics and econometrics.

Hua et al. (2007) use a mixture of logistic regression (demand occurrence) and a bootstrap (demand sizes).

Snyder et al. (2012) used a mixture of Hurdle Shifted Poisson and Geometric distributions.

Jiang et al. (2019) used Poisson-based mixture distribution models together with logistic regression.

Mixture distribution model

The mixture distribution model can be summarised as:

$$\begin{array}{l} y_t = o_t z_t \\ z_t \text{ is a statistical model} \\ o_t \sim \text{Bernoulli}\left(p_t\right) \\ p_t \text{ is another statistical model} \end{array}, \qquad (1)$$

where z_t is the demand size, o_t is the demand occurrence $\{0, 1\}$, p_t is the probability of occurrence.

Mixture distribution model

Examples for z_t:

- Normal linear regression;
- Log normal linear regression;
- Normal regression after Box-Cox transform;
- Negative binomial regression;
- Poisson regression;
- Anything else for positive data.

Normal distribution does not make sense from statistical point of view.

ETS and ARIMA are out of the scope of this presentation.

Mixture distribution model

Examples for p_t:

- Logistic regression;
- Probit regression.

Mix the demand sizes and the demand occurrence in order to obtain the full model

Mixture distribution model

An example of a mixture of the logistic and the log normal:

$$\begin{aligned} & y_t = o_t z_t \\ & z_t \sim \log \mathcal{N}(\mu_{z,t}, \sigma_z^2) \\ & \mu_{z,t} = \mathbf{B}'\mathbf{X}_t \\ & o_t \sim \text{Bernoulli}\left(p_t\right) \\ & p_t = \frac{1}{1+\exp(-\mu_{p,t})} \\ & \mu_{p,t} = \mathbf{A}'\mathbf{X}_t \end{aligned}, \quad (2)$$

where $\mathbf{B}$ and $\mathbf{A}$ are the vectors of parameters and $\mathbf{X}_t$ is the vector of explanatory variables.

Mixture distribution model

This can be estimated using the likelihood approach.

But the cases of $o_t = 0$ need to be considered as “missing data”, when estimating z_t.

Svetunkov & Boylan (under review) show that the likelihood in this case can be calculated as:

$$\begin{aligned} \ell(\boldsymbol{\theta}, \sigma_\epsilon^2 | \mathbf{Y}) = & \sum_{o_t=1} \log f_z(z_t) - \frac{T_0}{2} \log(2\pi e \sigma_z^2) \\ & + \sum_{o_t=1} \log(p_t) + \sum_{o_t=0} \log(1-p_t) \end{aligned} , \quad (3)$$

where T_0 is the number of zeroes.

Mixture distribution model

Maximising the likelihood (3), we can estimate the parameters of the model.

Information criteria (i.e. AIC) can be calculated based on (3).

One of the options – estimate different distribution models and select the best for your data.

Compare mixture distribution with the normal linear regression?

Mixture distribution model

The forecast of the demand sizes: $\hat{z}_t = \mu_{z,t}$

The forecast of the demand occurrence: $\hat{p}_t = \frac{\mu_{p,t}}{1+\mu_{p,t}}$

The final point forecast (conditional mean) is: $\hat{y}_t = \hat{p}_t \hat{z}_t$

Bonus: parametric prediction intervals.

The competition

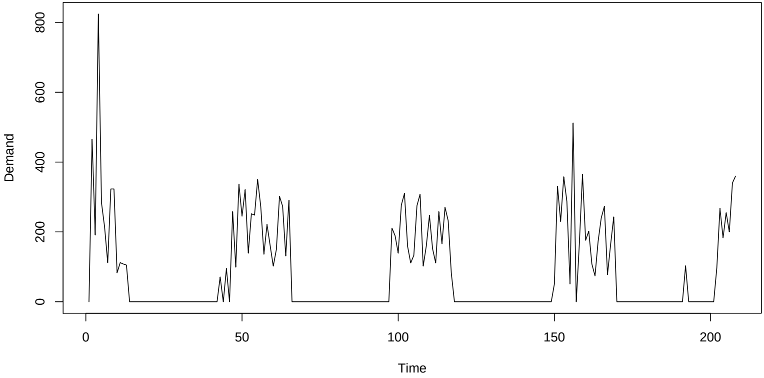
Simulation setting

The data simulated from mixture of Logistic and LogNormal

- 1000 weekly series;
- 208 observations: 156 in-sample, 52 holdout;
- Random number of zeroes between 10 and 42 for each period;
- Fixed origin;
- Deterministic seasonality;
- Promotions;
- Rounded up values.

Simulation setting

Series like this:



The models

Models included:

`alm()` function from `greybox` package v0.5.2 for `R`.

- Benchmark: normal linear regression (`distribution="dnorm"`);
- Normal + logistic (`distribution="dnorm", occurrence="plogis"`);
- Log normal + logistic (`distribution="dlnorm", occurrence="plogis"`);

The models

Silly benchmarks:

- Negative binomial + logistic (`distribution="dnbinom", occurrence="plogis"`);
- Poisson + logistic (`distribution="dpois", occurrence="plogis"`);
- Model selected using AIC.
- ETSX(A,N,N) (`es()` from `smooth` package for R) – just for fun...
- iETS_I (`es()` from `smooth`) – because Nikos asked...

Dummies for weeks as explanatory variables + promotions.

Produce mean forecasts and the upper bounds.

The evaluation

Forecasts evaluation:

- Relative RMSE for point forecasts:

$$\text{RelRMSE} = \frac{\text{RMSE}_a}{\text{RMSE}_b}, \tag{4}$$

where

$$\text{RMSE} = \sqrt{\frac{1}{h}\sum_{j=1}^{h}(y_{t+j} - \hat{y}_{t+j})^2} \tag{5}$$

The evaluation

Forecasts evaluation:

- Relative Mean Interval Score (MIS, inspired by Gneiting and Raftery, 2007);

$$\text{RelMIS} = \frac{\text{MIS}_a}{\text{MIS}_b}, \quad (6)$$

where $\text{MIS} = \frac{1}{h} \sum_{j=1}^{h} \text{IS}$ and

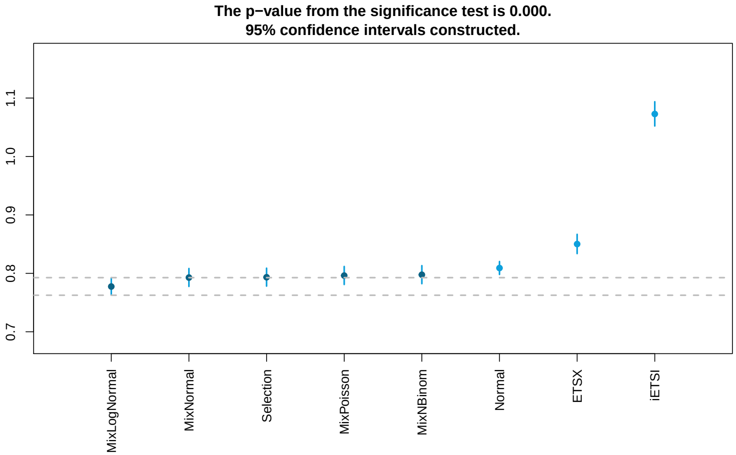
$$\text{IS} = \begin{array}{l} (u_{t+j} - l_{t+j}) + \frac{2}{\alpha}(l_{t+j} - y_{t+j}) 1\{y_{t+j} < l_{t+j}\} + \\ \frac{2}{\alpha}(y_{t+j} - u_{t+j}) 1\{y_{t+j} > u_{t+j}\} \end{array}, \quad (7)$$

Using a normal regression with intercept (average of the series) as a benchmark.

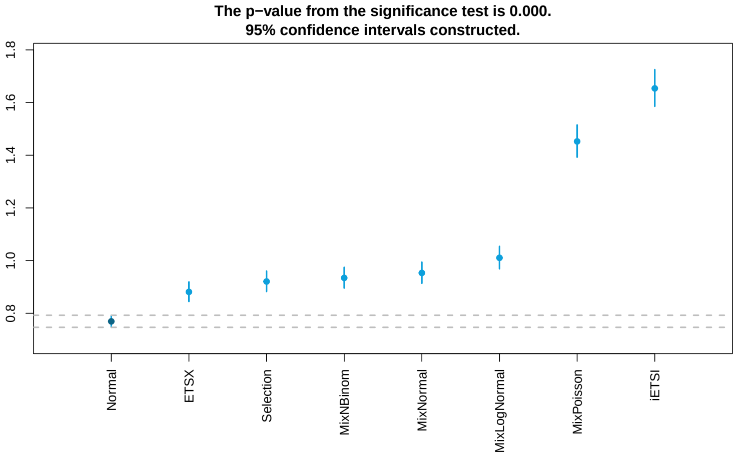
The simulation results

Methods	Mean values RelRMSE	Mean values RelMIS	Median Values RelRMSE	Median Values RelMIS
Normal	0.809	**0.769**	0.802	**0.781**
MixNormal	0.793	0.953	0.783	0.940
MixLogNormal	**0.777**	1.011	**0.768**	1.003
MixPoisson	0.796	1.453	0.780	1.467
MixNBinom	0.798	0.934	0.786	0.938
Selection	0.793	0.921	0.784	0.931
ETSX	0.850	0.881	0.858	0.904
$iETS_I$	1.073	1.654	1.021	1.278

The simulation results. MCB on RelRMSE



The simulation results. MCB on RelMIS



The real data experiment

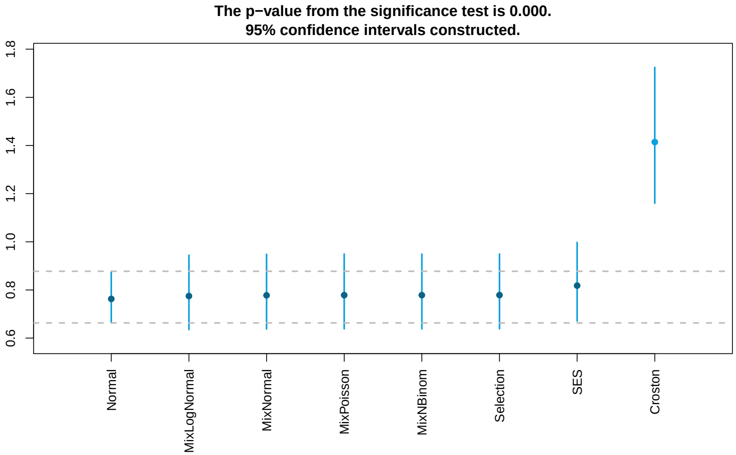
An experiment on a real data

- Tomato sales data;
- Weekly;
- 261 observations: 209 in-sample, 52 for the holdout;
- Fixed origin;
- 24 products that have non-zeroes in the holdout;

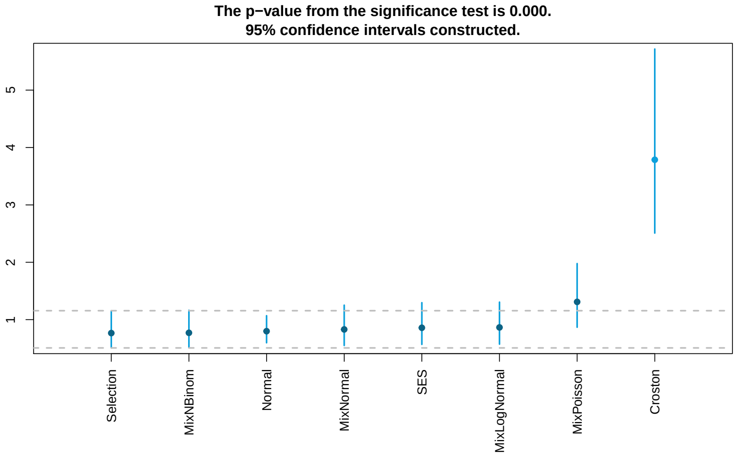
The real data results

Methods	Mean values RelRMSE	Mean values RelMIS	Median Values RelRMSE	Median Values RelMIS
Normal	**0.763**	**0.799**	**0.757**	0.776
MixNormal	0.778	0.830	0.770	0.645
MixLogNormal	0.775	0.865	0.763	0.670
MixPoisson	0.778	1.309	0.773	1.328
MixNBinom	0.778	0.770	0.761	0.643
Selection	0.779	0.766	0.760	**0.639**
ETSX	0.818	0.858	0.810	0.798
iETS	1.414	3.787	1.445	3.399

The real data results. MCB on RelRMSE



The real data results. MCB on RelMIS



Conclusions

Conclusions

- Regression approach is reasonable for retail data;
- Normal linear regression seems to work okay;
- Using mixture distribution models is a promising direction;
- We need more data!
- We really need more data!
- And explanatory variables!
- Add AR, I, MA components;
- Variables / model selection.

Thank you for your attention!

Ivan Svetunkov

i.svetunkov@lancaster.ac.uk

https://forecasting.svetunkov.ru

twitter: @iSvetunkov

Marketing Analytics and Forecasting

Lancaster University Management School

The simulation results expanded

Methods	Mean values		Median Values	
	Coverage	RelRange	Coverage	RelRange
Normal	**0.930**	0.728	**0.942**	0.731
MixNormal	0.891	0.469	0.904	0.483
MixLogNormal	0.873	0.461	0.885	0.476
MixPoisson	0.789	**0.337**	0.808	**0.347**
MixNBinom	0.888	0.510	0.904	0.524
Selection	0.891	0.506	0.904	0.568
ETSX	0.924	0.715	0.923	0.721
$iETS_I$	0.812	0.648	0.885	0.728

CMAF

Gneiting, T., Raftery, A. E., 2007. Strictly proper scoring rules, prediction, and estimation. Journal of the American Statistical Association 102 (477), 359–378.

Hua, Z. S., Zhang, B., Yang, J., Tan, D. S., 2007. A new approach of forecasting intermittent demand for spare parts inventories in the process industries. Journal of the Operational Research Society 58 (1), 52–61.

Jiang, A., Tam, K. L., Guo, X., Zhang, Y., 2019. A New Approach to Forecasting Intermittent Demand Based on the Mixed Zero-Truncated Poisson Model. Journal of Forecasting. URL `https://onlinelibrary.wiley.com/doi/abs/10.1002/for.2614`

Snyder, R. D., Ord, J. K., Beaumont, A., apr 2012. Forecasting the intermittent demand for slow-moving inventories: A modelling approach. International Journal of Forecasting 28 (2), 485–496. URL `http://dx.doi.org/10.1016/j.ijforecast.2011.03.`

CMAF

009http://linkinghub.elsevier.com/retrieve/pii/S0169207011000781