Forecasting Model for Sporadic Distributor based Markets

Ahmed El-azzamy
Stanley Park
Agenda

- Company Settings
- Motivation / Background
- Literature Review
- Methodology
- Results
- Conclusion
Company Settings

Company Information

Industry

Fast Moving Consumer Goods (FMCG) Company

Geographical Operation

Global Region

Overall Process of Goods

Goods Flow → Information Flow

(FMCG) Factory

Distributor

Retailers (Region A)

Retailers (Region B)

Retailers (Region C)

Order & Fulfillment Mgmt.

Demand Management Team

Order

Forecast

Shipment

Sell-out

Order
Motivation / Background

- Wide discrepancy between forecast and actual shipment impacts business (ex. missing sales, high inventory costs)

1. Sample of one SKU
Key Question

• What is the best model to forecast long & short term shipments for sporadic distributor based markets?

• How can we link distributor data to improve the E2E (End to End) supply network?
**Literature Review – Croston’s Method**

**Traditional Forecasting method**

- **Moving Average**
  - The method forecasts based on average of most recent observations (ex. 3MA observes 3 weeks of demand to forecast)
  - Weight is same to each observation

- **Exponential Smoothing**
  - The forecast uses “Alpha” factor, which indicates the value of new information
  - Put more weight to recent observation
    - If Alpha increases, more weight on new information

Since there are too many zero values in demand, it could lead to poor forecast

**Croston’s Method**

**Intro**
- 1972, Croston JD developed the forecasting method
- Developed to forecast demands that have multiple 0 values

**Approach**
- Forecast two data independently and aggregates in order to determine average demand per period

**Non-zero demand time series** → **Demand size** → **Average Demand per time period**

**Zero demand time series** → **Inter-arrival time**
Literature Review – Multi-Tier Regression Analysis

- Demand-Driven Forecasting: A Structured Approach to Forecasting, Charles W. Chase, Jr.
- Methods of linking downstream data into forecasting model
- CPG Industry Case
- Demand & Supply Model Through Using Linear Regression
- Demand Forecast → Supply Forecast

(1) Demand (D)
\[ D = \beta_d0 \ \text{Constant} + \beta_d1 \ \text{Trend} + \beta_d2 \ \text{Seasonality} + \beta_d3 \ \text{Price} + \beta_d4 \ \text{Advertising} + \beta_d5 \ \text{Sales} + \beta_d6 \ % \ \text{ACV Feature} + \beta_d7 \ \text{FSI} + \beta_d8 \ \text{Store Distribution} + \beta_d9 \ \text{Competitive Price} + \cdots \beta_dn \]

(2) Supply (S)
\[ S = \beta_s0 \ \text{Constant} + \beta_s1D(lag1 - n) + \beta_s2\ \text{Trend} + \beta_s3 \ \text{Seasonality} + \beta_s4 \ \text{Gross Dealer Price} + \beta_s5 \ \text{Factory Rebates} + \beta_s6 \ \text{Cash Discount} + \beta_s7 \ \text{Coop Advertising} + \beta_s8 \ \text{Trade Promotions} + \cdots \beta_sn \]
## Method – Downstream Data Acquisition

<table>
<thead>
<tr>
<th>Category</th>
<th>Data Point</th>
<th>Note</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Manufacturer / Distributor</strong></td>
<td>Sell-out</td>
<td>Shipments from distributor to retailer</td>
</tr>
<tr>
<td></td>
<td>Inventory level</td>
<td>Inventory level of distributor</td>
</tr>
<tr>
<td></td>
<td>Customer Fill Rate</td>
<td>Satisfied orders / Requested orders</td>
</tr>
<tr>
<td></td>
<td>Sales Target (or Building Block)</td>
<td>Sales target set by manufacturer</td>
</tr>
<tr>
<td><strong>Retailer</strong></td>
<td>Price</td>
<td>Not available</td>
</tr>
<tr>
<td></td>
<td>POS data</td>
<td>Not available</td>
</tr>
<tr>
<td></td>
<td>Competition Prices</td>
<td>Not available</td>
</tr>
<tr>
<td></td>
<td>Premium Displays</td>
<td>Not available</td>
</tr>
<tr>
<td></td>
<td>Advertising</td>
<td>Not available</td>
</tr>
<tr>
<td><strong>General</strong></td>
<td>Events</td>
<td>National events</td>
</tr>
</tbody>
</table>
## Method – Multi-Tier Regression Model

<table>
<thead>
<tr>
<th>Regression Analysis</th>
<th>Dependent Variables (=y)</th>
<th>Independent Variables (=x)</th>
<th>Metrics</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>• Shipments</td>
<td>• Sell-out</td>
<td>• RMSE¹ / A-MAPE² / MD³</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Inventory level</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Customer Fill Rate</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>• National holiday or event</td>
<td></td>
</tr>
</tbody>
</table>

- **Apply various lag time between shipment and sell-out**
  - Since there is lead time of approx. 1 month in shipping, the distributor might ship 1 month prior
  - The team applied lag time of 1 ~ 3 month

- **Analyze the relationship between distributor’s inventory and shipment**
  - Identify how inventory level affects the shipment

1. Root Mean Square Error  
2. Mean Absolute Percentage Error  
3. Adjusted Mean Absolute Percentage Error  
4. Mean Deviation
Method – Metrics

1. RMSE\(^1\) = \left(\frac{1}{n}\sum_{t=1}^{n}(Y_t - F_t)\right)^{1/2}

2. MAPE\(^2\) = \frac{1}{n}\sum_{t=1}^{n}\frac{|Y_t - F_t|}{Y_t}

   \[\text{A-MAPE}\(^3\) = \frac{1}{n}\sum_{t=1}^{n}\frac{|Y_t - F_t|}{\frac{1}{n}\sum_{t=1}^{n}Y_t}\]

3. MD\(^4\) = \sum_{t=1}^{n}\frac{Y_t - F_t}{n}

Approach #1:
Multi-regression Analysis

Approach #2:
Croston Method

Internal Method
Objective:
Understand the behavior of explanatory variables overtime

Results:

1. There is a lag between Sellout and shipment
2. Distributor Inventory increased at the end of the time horizon due to sellout decrease.
Objective:
Understand correlation between explanatory variables

Observations in the underlying dataset:
1. Events doesn’t correlate with number of shipments
2. Number of shipments is highly correlated with sellout (.84) and Inventory (.41)
3. Building block and inventory are highly correlated which may cause multicollinearity
Results- Regression Analysis on Category Level

Objective:
Identify the lag between Shipment and sellout

Results:
1. Shipments are more correlated with sellout of t-1
2. There are some outliers, we flagged them as an event to see their effect on the model
Results - Category level Model

1. We predicted sellout using the following equation
2. Model can predict up to building block availability date
3. We used the predicted sellout as an independent variable to predict shipments
4. Used a simulated inventory value as an explanatory variable
5. Flagged event dates, and added it as another independent variable
6. Predicted Sellout with an 83% $R^2$, and shipments with 78% $R^2$

\[
\text{Sellout} = 8291.22 - 45.18 \times \log(BB + 1) - 6818.99 \times \text{(Category 2)} - 1788.95 \times \text{(Category 3)} - 7033.47 \times \text{(Category 4)} - 5170.5 \times \text{(Category 5)}
\]

\[
\text{Shipment volume} = 8.399e+03 + 2.552e-01 \times \text{(Sellout t - 1)} - 6.535e-02 \times \text{(Inventory t - 1)} - 7.957e+02 \times \text{Events} - 7.228e+03 \times \text{(Category 2)} - 1.691e+03 \times \text{(Category 3)} - 6.978e+03 \times \text{(Category 4)} - 5.023e+03 \times \text{(Category 5)}
\]
Results - Measuring Accuracy on Category Level

Category Results Summary:

1. Model REMSE has decreased for 3 out of 5 categories
2. Proposed model is less biased in some categories than original forecast
3. Since the model is taking building block as main variable, it’s all dependent on its accuracy.
Results - Comparing with Croston

- Ran Croston Method on category level
- Compared results over three month
- Improvement in 3 out of 5 categories

<table>
<thead>
<tr>
<th>Category</th>
<th>RMSE</th>
<th>MD</th>
<th>A-MAPE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Original Forecast</td>
<td>Model Forecast</td>
<td>Croston</td>
</tr>
<tr>
<td>Cat 1</td>
<td>6.1</td>
<td>5.3</td>
<td>5.4</td>
</tr>
<tr>
<td>Cat 2</td>
<td>0.7</td>
<td>0.9</td>
<td>0.8</td>
</tr>
<tr>
<td>Cat 3</td>
<td>3.9</td>
<td>4.1</td>
<td>4</td>
</tr>
<tr>
<td>Cat 4</td>
<td>1.9</td>
<td>1.6</td>
<td>2.1</td>
</tr>
<tr>
<td>Cat 5</td>
<td>2.7</td>
<td>2.5</td>
<td>2.9</td>
</tr>
</tbody>
</table>
Results - Split on SKU level

➢ Used demand model to predict expected sellout
➢ Predicted shipment using supply model
➢ Split Category value on SKU based on how each SKU represents in total category
➢ On the shaped demand; on average 67 SKU improved more than 20% in terms of RMSE

<table>
<thead>
<tr>
<th></th>
<th>Oct</th>
<th>Nov</th>
<th>Average</th>
</tr>
</thead>
<tbody>
<tr>
<td>20%~</td>
<td>51</td>
<td>44</td>
<td>48</td>
</tr>
<tr>
<td>10~20%</td>
<td>3</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>0~10%</td>
<td>4</td>
<td>10</td>
<td>7</td>
</tr>
<tr>
<td>-20%~0%</td>
<td>37</td>
<td>37</td>
<td>37</td>
</tr>
<tr>
<td>Total</td>
<td>95</td>
<td>96</td>
<td>96</td>
</tr>
</tbody>
</table>
Conclusion

➢ With the underlying dataset, Multiple regression analysis shows

➢ Collaboration with distributor is an imperative factor of the success of this method of linking downstream

➢ Croston method helps setting a smooth inventory level.

Recommendation

➢ For Future research; Machine learning can be utilized to cluster items, and used different forecasting methods accordingly
Thank You
### Final Results in $\$

<table>
<thead>
<tr>
<th>Category</th>
<th>$/unit</th>
<th>RMSE Improvement (Unit)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Category 1</td>
<td>$20.00</td>
<td>800</td>
</tr>
<tr>
<td>Category 2</td>
<td>$48.36</td>
<td>-200</td>
</tr>
<tr>
<td>Category 3</td>
<td>$42.70</td>
<td>-200</td>
</tr>
<tr>
<td>Category 4</td>
<td>$56.71</td>
<td>300</td>
</tr>
<tr>
<td>Category 5</td>
<td>$53.20</td>
<td>200</td>
</tr>
</tbody>
</table>

**RMSE difference in $**: $25,442

**Forecast error difference of**: > $80 million

* Assume 1 distributor in 1~2 countries (i.e. 40 countries for potential savings)
* Assume 20 categories
Scatter plot matrix for shipment and other three numeric predictors in 1,000 unit

<table>
<thead>
<tr>
<th></th>
<th>Sellout</th>
<th>Shipment</th>
<th>Inv</th>
<th>BB</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Corr:</strong></td>
<td>0.817</td>
<td>0.676</td>
<td>0.648</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.568</td>
<td>0.587</td>
<td>0.752</td>
<td></td>
</tr>
</tbody>
</table>

© 2018 MIT Center for Transportation & Logistics | Page 21
Results- Simulation based on target inventory

Objective:
To investigate the replenishment the products that fall below the target inventory level

Result:
Brand 2 had improved by 33%
Other brands accuracy dropped significantly

<table>
<thead>
<tr>
<th></th>
<th>RMSE Simulation</th>
<th>RMSE Internal</th>
<th>Improvement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Brand1</td>
<td>7,094</td>
<td>3,311</td>
<td>-53%</td>
</tr>
<tr>
<td>Brand2</td>
<td>4,644</td>
<td>6,176</td>
<td>33%</td>
</tr>
<tr>
<td>Brand3</td>
<td>8,523</td>
<td>1,701</td>
<td>-80%</td>
</tr>
</tbody>
</table>
Model

1- Predicting distributor sellout

Coefficients:

| Model     | Estimate | Std. Error | t value | Pr(>|t|) |
|-----------|----------|------------|---------|----------|
| (Intercept) | 8291.22  | 533.96     | 15.528  | < 2e-16  *** |
| log(BB + 1) | -45.18   | 58.13      | -0.777  | 0.438    |
| CategoryCategory2 | -6818.99 | 314.82     | -21.660 | < 2e-16  *** |
| CategoryCategory3 | -1788.95 | 301.25     | -5.938  | 1.59e-08 *** |
| CategoryCategory4 | -7033.47 | 416.75     | -16.877 | < 2e-16  *** |
| CategoryCategory5 | -5170.50 | 303.98     | -17.009 | < 2e-16  *** |

---

Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1258 on 169 degrees of freedom
Multiple R-squared:  0.8308,  Adjusted R-squared:  0.8258
F-statistic:  166 on 5 and 169 DF,  p-value:  < 2.2e-16

2- Predicting company shipment

Coefficients:

| Model     | Estimate | Std. Error | t value | Pr(>|t|) |
|-----------|----------|------------|---------|----------|
| (Intercept) | 8.399e+03 | 1.141e+03  | 7.361   | 7.88e-12  *** |
| CategoryCategory2 | -7.228e+03 | 1.047e+03  | -6.906  | 9.95e-11  *** |
| CategoryCategory3 | -1.691e+03 | 5.081e+02  | -3.327  | 0.00108  **  |
| CategoryCategory4 | -6.978e+03 | 1.029e+03  | -6.778  | 2.00e-10  *** |
| CategoryCategory5 | -5.023e+03 | 7.546e+02  | -6.657  | 3.85e-10  *** |
| Sellout    | 2.552e+01 | 1.101e+01  | 2.319   | 0.02161  *  |
| Inv        | -6.535e+02 | 2.255e+02  | -2.898  | 0.00427  ** |
| Events     | -7.957e+02 | 3.605e+02  | -2.207  | 0.02867  *  |

---

Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1792 on 167 degrees of freedom
Multiple R-squared:  0.7563,  Adjusted R-squared:  0.7461
F-statistic:  74.03 on 7 and 167 DF,  p-value:  < 2.2e-16
Forecasting Techniques

➢ Simulation

➢ Croston and it’s variation

➢ Multi-Tiered Causal Analysis

\[
(1) \text{Demand } (D) = \beta_0 \text{ Constant} + \beta_1 \text{ Trend} + \beta_2 \text{ Seasonality} + \beta_3 \text{ Price} + \beta_4 \text{ Advertising} + \beta_5 \text{ Sales} \\
+ \beta_6 \% \text{ ACV Feature} + \beta_7 \text{ FSI} + \beta_8 \text{ Store Distribution} + \beta_9 \text{ Competitive Price} + \cdots \beta_n
\]

\[
(2) \text{Supply } (S) = \beta_0 \text{ Constant} + \beta_1 D(\text{lag1}-n) + \beta_2 \text{ Trend} + \beta_3 \text{ Seasonality} + \beta_4 \text{ Gross Dealer Price} \\
+ \beta_5 \text{ Factory Rebates} + \beta_6 \text{ Cash Discount} + \beta_7 \text{ Coop Advertising}
\]
Model with building Block

Coefficients:

|                  | Estimate | Std. Error | t value | Pr(>|t|) |
|------------------|----------|------------|---------|----------|
| (Intercept)      | 8.434e+03| 1.183e+03  | 7.131   | 2.93e-11 | ***     |
| CategoryCategory2| -7.256e+03| 1.076e+03  | -6.745  | 2.42e-10 | ***     |
| CategoryCategory3| -1.673e+03| 5.310e+02  | -3.151  | 0.00193  | **      |
| CategoryCategory4| -7.012e+03| 1.073e+03  | -6.534  | 7.52e-10 | ***     |
| CategoryCategory5| -5.056e+03| 8.065e+02  | -6.269  | 3.02e-09 | ***     |
| Sellout          | 2.532e-01| 1.118e-01  | 2.266   | 0.02477  | *       |
| Inv              | -6.373e-02| 2.647e-02  | -2.408  | 0.01714  | *       |
| BB               | -1.148e-02| 9.736e-02  | -0.118  | 0.90632  |         |
| Events           | -7.879e+02| 3.676e+02  | -2.143  | 0.03355  | *       |

---

Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1798 on 166 degrees of freedom
Multiple R-squared:  0.7563,  Adjusted R-squared:  0.7446
F-statistic:  64.4 on 8 and 166 DF,  p-value: < 2.2e-16