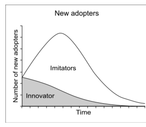
New Product Forecast - Bass Model

The Bass product forecast method has been successfully used to predict the market shares of various newly introduced products, as well as mature ones. Bass model assumes that the adoption rate of a product is contributed by two factors :

Innovators (denoted by p): Consumers that adopt the product independent of social influences.

Imitators (denoted by q): Consumers who adopt the product because of influence from early adopters.

In the early life cycle of a product p > q, which means adoption rate of product in innovators is much higher than the adoption rate in imitators. At some point in the life cycle of a good product, the influence of the early adopters becomes sufficiently strong to drive many other imitators to adopt the product (i.e, p <= q).



Assumptions of Bass Model

1. Diffusion process is binary (consumer either adopts,or waits to adopt).
2. Maximum potential number of buyers (m) is constant, eventually, all m will buy the product.
3. No repeat purchase, or replacement purchase.
4. The impact of the word-of-mouth is independent of adoption time.
5. Innovation is considered independent of substitutes.
6. The marketing strategies supporting the innovation are not explicitly included.

Mathematical Formulation of Bass Model

p - Coefficient of innovation.
q - Coefficient of imitation
m - Potential Market (ultimate number of purchasers of the product).
t - Time interval.
$s(t)$ - Sale of the product at time interval t.
$S(t)$ - Cumulative sales of the product till time interval t $(0 < x <= t)$.

The portion of product adopters at any time t is given by a probability density function as

$$f(t) = \frac{s(t)}{M}$$

The cumulative portion of product adopters upto time t is given by a cumulative density function as

$$F(t) = \frac{S(t)}{M}$$

The derivative of a cdf at a point yields the value of pdf at that point, so

$$\frac{\partial F}{\partial t} = f(t)$$

Because the total number of adopters is 100% (or 1) of the potential market, the number of adopters at time t who have not yet adopted is given by .

$$1 - F(t)$$

The portion of market that adopts the product at time t given that they have not yet adopted is represented as

$$\frac{f(t)}{1 - F(t)}$$

According to Bass Model, the portion of customers that adopts the product at ant time t, given that they have not yet adopted is a linear function of innovators and imitators.

$$\frac{f(t)}{1 - F(t)} = p + q * F(t)$$

Estimation m, p, q from historical sales:

$$\frac{f(t)}{1 - F(t)} = p + q\ F$$

$$\frac{s(t)/m}{1 - S(t)/m} = p + q\ S(t)/m$$

$$s(t) = |p + q\ S(t)/m|(m - S(t)|$$

$$s(t) = \beta_0 + \beta_1 * S(t) + \beta_2 * S(t)^2$$

$$\beta_0 = pm$$
$$\beta_1 = q - p$$
$$\beta_2 = -q/m$$

Equation (BASS) may be estimated by a regression of sales against cumulative sales. Once the coefficients in the regression (β0,β1,β2) are obtained, the equations above may be inverted to determine the values of (m,p,q).

$$\beta_1 = q - p = -m\beta_2 - \frac{\beta_0}{m}$$

$$\beta_2 m^2 + \beta_1 m + \beta_0 = 0$$

$$m1 = \frac{-\beta_1 + \sqrt{\beta_1^2 - 4\beta_0\beta_2}}{2\beta_1}$$

$$m1 = \frac{-\beta_1 - \sqrt{\beta_1^2 - 4\beta_0\beta_2}}{2\beta_1}$$

$$m = max(m1, m2)$$
$$p = \frac{\beta_0}{m}$$
$$q = -m\beta_2$$

Solving model for F(t) assuming F(0) = 0:

$$\frac{dF/dt}{1 - F} = p + q\ F.$$

$$F = \frac{p[e^{(p+q)t} - 1)}{pe^{(p+q)t} + q}$$

Please refer the following link for full derivation: https://srdas.github.io/MLBook/productForecastingBassModel.html

Predicting the sales using p,q,m & t values:

$$s(t) = mf(t) = \frac{dF}{dt} = \frac{e^{(p+q)t}p(p+q)^2}{[pe^{(p+q)t} + q]^2}$$

Case Study - Quarterly sales of iphone

IPhone Quarterly Sales

QUARTER	SALES	YEAR	CUM_SALES	CUM_SALES_SQUARED	TIME_INDEX
2007_Q3	0.27	2007	0.27	0.0729	1
2007_Q4	1.12	2007	1.39	1.9321	2
2008_Q1	2.32	2008	3.71	13.7641	3
2008_Q2	1.70	2008	5.41	29.2681	4
2008_Q3	0.72	2008	6.13	37.5769	5
2008_Q4	6.89	2008	13.02	169.5204	6
2009_Q1	4.36	2009	17.38	302.0644	7
2009_Q2	3.79	2009	21.17	448.1689	8
2009_Q3	5.21	2009	26.38	695.9044	9
2009_Q4	7.37	2009	33.75	1139.0625	10

Finding m,p,q values

```
## Estimating beta's of the quadratic equation
model = lm(iphone_sales$SALES ~ iphone_sales$CUM_SALES + iphone_sales$CUM_SALES_SQUARED)
b = model$coefficients

m1 = (-b[2]+sqrt(b[2]^2-4*b[1]*b[3]))/(2*b[3])
m2 = (-b[2]-sqrt(b[2]^2-4*b[1]*b[3]))/(2*b[3])
m = max(m1,m2);
p = b[1]/m
q = -m*b[3]
print(paste0('Coefficient of innovation - p: ',round(p,7)))
```

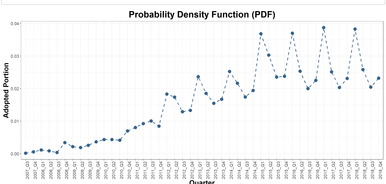
```
## [1] "Coefficient of innovation - p: 0.0022111"
```

```
print(paste0('Coefficient of imitation - q: ',round(q,7)))
```

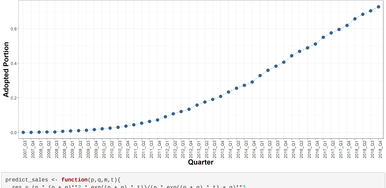
```
## [1] "Coefficient of imitation - q: 0.1105407"
```

```
print(paste0('Potential Market Size ',round(m),' Million Units.'))
```

```
## [1] "Potential Market Size 2020 Million Units."
```



Predicting Sales using p,q,m & t values




```
predict_sales <- function(p,q,m,t){
  res = (p * (p + q)**2 * exp((p + q) * t))/(p * exp((p + q) * t) + q)**2
  return(res * m)
}

preds = c()
for(qtr in seq(0:60)) preds = c(preds,predict_sales(p,q,m,qtr))
iphone_sales$PRED_SALES = preds[1:nrow(iphone_sales)]


### Sales Trend across quarters
ggplot(iphone_sales,aes(x = QUARTER, y = SALES,group = 1)) +
  geom_point(color = 'steelblue',size = 4) +
  geom_line(color = 'steelblue',size = 1, linetype = 'dashed') + theme_bw() +
  geom_point(aes(y = PRED_SALES),color = 'red',size = 4) +
  geom_line(aes(y = PRED_SALES), color = 'red',size = 1,linetype="dashed") + theme_bw() +
  theme(axis.text.x = element_text(size=12,angle = 90),
        axis.title.x = element_text(size=20,face="bold"),
        axis.text.y = element_text(size=12),
        axis.title.y = element_text(size=20,face="bold"),
        plot.title = element_text(size=25,face="bold",hjust = 0.5)) +
  labs(title = 'Actual Sales Vs Bass Predicted Sales',x = 'Quarter',y = 'Million Units')
```