# Application of nonlinear least squares: Estimating parameters of the Bass model

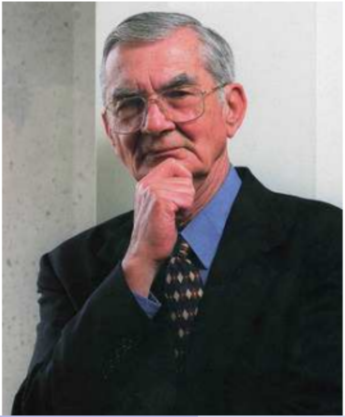
Beáta Stehlíková

# Bass model

# Bass model

## Frank Bass (1926-2006)

Mathematical models in marketing



# Bass model

F. Bass, A New Product Growth for Model Consumer Durables, Management Science, Vol. 15 (January 1969)

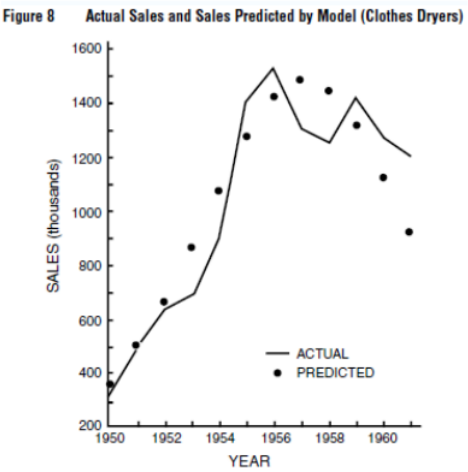
One of 10 papers in the selection *Top 10 Most Influential Papers published in the 50-year history of Management Science* (2004)

> *Perhaps the first thing to notice... is the title. It contains a typo. The correct title should be* ***A New Product Growth Model for Consumer Durables****. I suppose that I was so excited about having the paper accepted for publication that I failed to carefully proofread the galley proofs. (Frank Bass, 2004)*

Mathematical model for **a new product on the market**

# Bass model

Example from the original Bass' paper:



# Bass model: idea

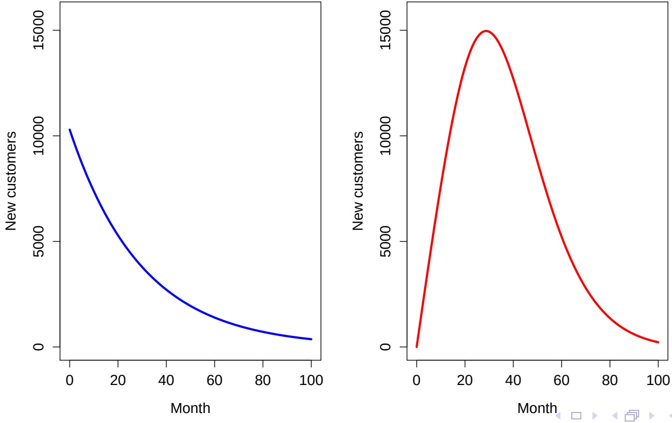
## Innovators and imitators

The basic idea - there are *two types of customers*:

- *innovators* - the buy the product using the information about the product, advertisement, etc.
- *imitators* - their decision is based on experience of other people, their ratings, etc.

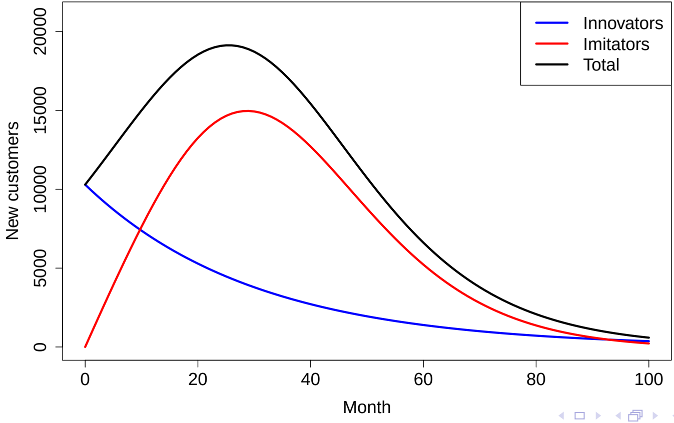
# Bass model: idea

Innovators and immitators - number of new customers:



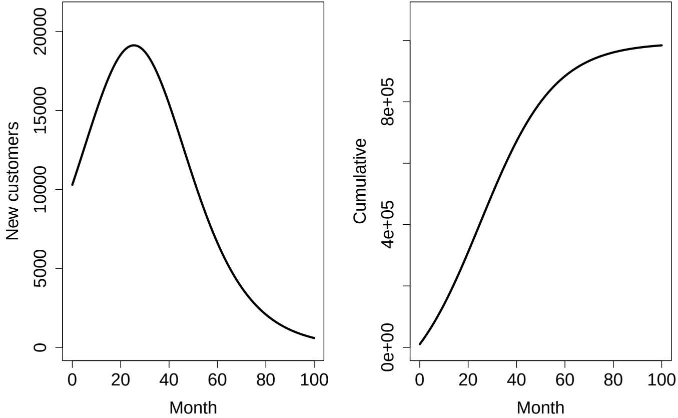
# Bass model: idea

New customers - total:



# Bass model: idea

New customers and cumulative number of customers:



# Bass model: mathematical formulation

- Continuous time $t$
- $F(t) =$ ratio of people (out of the total market volume) that bought the product by the time $t$
- $f(t) =$ ratio of people (out of the total market volume) that bought the product at the time $t$, we have? $f(t) = F'(t)$
- **Assumption:** Probability that a person buys the product at time $t$, id he did not buy it before, is $p + qF(t)$
- **Parameters:** $p$ given the effect of innovators, $q$ the effect of imitators
- We obtain the equation:

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

pričom $F(0) = 0$

# Bass model: solution

We have an ordinary differential equation for $F(t)$:

$$\frac{F'(t)}{1 - F(t)} = p + qF(t), F(0) = 0,$$

which can be solved by separation of variables:

$$\frac{dF}{(1 - F)(p + qF)} = dt \Rightarrow F(t) = \frac{1 - e^{-(p+q)t}}{1 + \frac{q}{p} e^{-(p+q)t}}$$

# Bass model: solution

Corresponding function $f(t) = F'(t)$:

$$f(t) = \frac{(p+q)^2 e^{-(p+q)t}}{p\left[1 + \frac{q}{p} e^{-(p+q)t}\right]^2}$$

and its maximum - loosely speaking: at which time we sell the highest number of products - for $q > p > 0$:

$$t_{peak} = \frac{\ln(q/p)}{p+q}$$

# Bass model: solution

Exercises:

- Derive the expressions for $f(t)$ a $F(t)$
- Derive the expression for $t_{peak}$ from the previous slide, i.e., for the case of d $q > p > 0$.
- What happens for $q < p$? What is an intuitive explanation of this?

# Bass model: parameters

As a first approximation - left: $p$, right: $q$

| | | |
|---|---|---|
| Baseline case: | | |
| US, consumer, durable, launch in '76 ... | 0.016 | 0.409 |
| For other cases, multiply by the following factors ... | | |
| Cellular telephone | 0.226 | 0.635 |
| Non durable product | 0.689 | 0.931 |
| Industrial | 1.058 | 1.149 |
| Non commercial innovation | 0.365 | 2.406 |
| Western Europe | 0.464 | 0.949 |
| Asia | 0.595 | 0.743 |
| Other regions | 0.796 | 0.699 |
| For each year after 1976, multiply by ... | 1.021 | 1.028 |

Christophe Van den Bulte: *Want to know how diffusion speed varies across countries and products? Try using a Bass model.* PDMA Visions 26(4) 2002, pp. 12-15

# Bass model: parameters

We define a function:

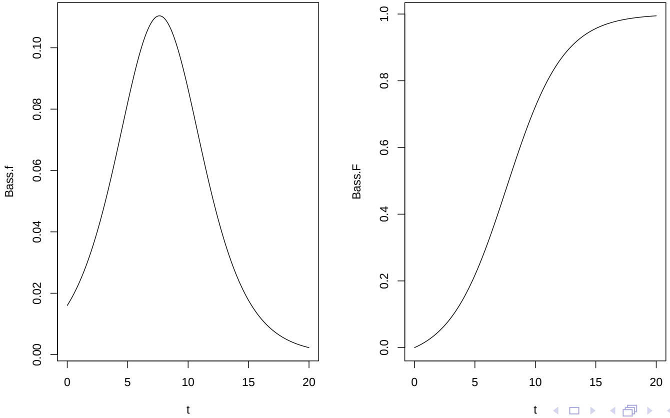
```
graphs <- function(p,q,t.max) {
  t <- seq(from=0,to=t.max,by=0.01)
  Bass.f <- ((p+q)^2/p)*exp(-(p+q)*t)/
                       (1+(q/p)*exp(-(p+q)*t))^2
  Bass.F <- (1-exp(-(p+q)*t))/(1+(q/p)*exp(-(p+q)*t))
  par(mfrow=c(1,2))
  plot(t, Bass.f, type="l"); plot(t, Bass.F, type="l")
}
```

Now we can plot:

```
graphs(0.016, 0.409, 20)   # baseline from the table
```

# Bass model: parameters

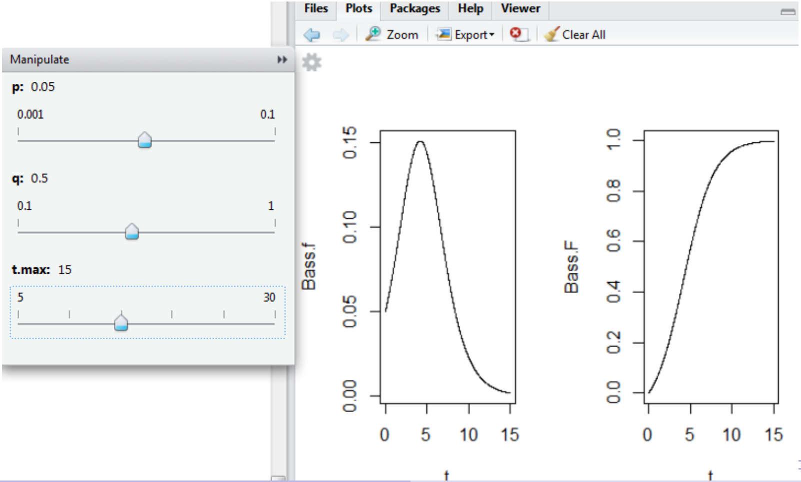
## Output



# Interactive graphs

# Bass model: interactive choice of parameters

We show how to create something like:



# Library manipulate for R-studio

Load (if needed, install) library `manipulate`:

```
library(manipulate)
```

Example of how to use the function `manipulate`:

```
manipulate(graphs(p,q,t.max),
           p=slider(min=0.001, max=0.1, step=0.001),
           q=slider(min=0.1, max=1, step=0.01),
           t.max=slider(min=5, max=30, step=5))
```

Among the parameters of `slider` we can specify also the initial value - using `initial=...` - try this.

For more info: `?slider`

# Estimating parameters of the Bass model in R

# Example - data

Source: *P. S. P. Cowpertwait, A. V. Metcalfe, Introductory Time Series with R. Springer 2009. Kapitola 3.3.4, str. 52-54*

We are going to model **the sales of VCRs in the USA in 1980-1989**:

```
T <- 1:10  # time, year = 1979 + T
Sales <- c(840,1470,2110,4000,7590,10950,10530,9470,
           7790,5890)
```

Model: $Sales(t) = M \times f(t)$, where $M$ are the total sales.

**Exercise 1:** Plot the time evolution of the sales.

# Estimation of parameters

Function `nls` - estimates parameters using the method of nonlinear least squares.

We need initial values of parameters, we take:

- `p` a `q` from the table in the slides or using `manipulate`
- `M` equal to the total sales so far (they start decreasing and the total cummulative sales will not be much higher) or somewhat higher

# Estimation of parameters

We use `nls`:

```
Bass.nls<-nls(Sales ~ M*(((P+Q)^2/P)*exp(-(P+Q)*T))/
                (1+(Q/P)*exp(-(P+Q)*T))^2,
            # add P and Q below
            start=c(list(M=sum(Sales),P=...,Q=...)))
summary(Bass.nls)
```

# Estimation of parameters

Output:

```
##
## Formula: Sales ~ M * (((P + Q)^2/P) * exp(-(P + Q) * T))
##     exp(-(P + Q) * T))^2
##
## Parameters:
##    Estimate Std. Error t value Pr(>|t|)
## M 6.798e+04  3.128e+03   21.74 1.10e-07 ***
## P 6.594e-03  1.430e-03    4.61  0.00245 **
## Q 6.381e-01  4.140e-02   15.41 1.17e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1
##
## Residual standard error: 727.2 on 7 degrees of freedom
```

# Estimation of parameters

**Exercise 2:** Compare the actual sales with the fitted and make a prediction for the following years.

**Remark:** How to access the estimated parameters:

```
coef(Bass.nls)
```

```
##             M            P            Q
## 6.798093e+04 6.593972e-03 6.380909e-01
```

```
as.vector(coef(Bass.nls))
```

```
## [1] 6.798093e+04 6.593972e-03 6.380909e-01
```

# Modelling sales of movies

# Sales of movies

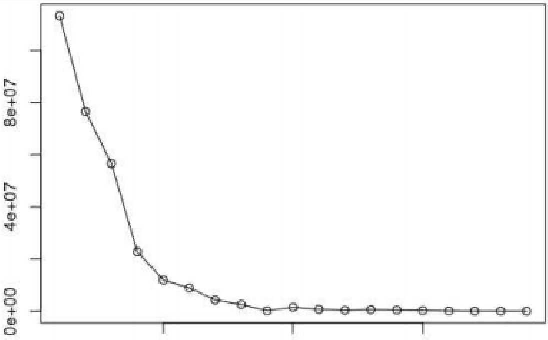
Again, it makes sense to distinguish:

- some people go to see a movie because they know about it and want to see it (*innovators* in Bass model)
- others decide to see it because of the recommendation of those who have already seen it (*imitators* in Bass model)

# Two kinds of movies (I.)

Advertisement, people are waiting for the movie - a lot of people come to see it in the first days, e.g. *The Hobbit: An Unexpected Journey (2012)*

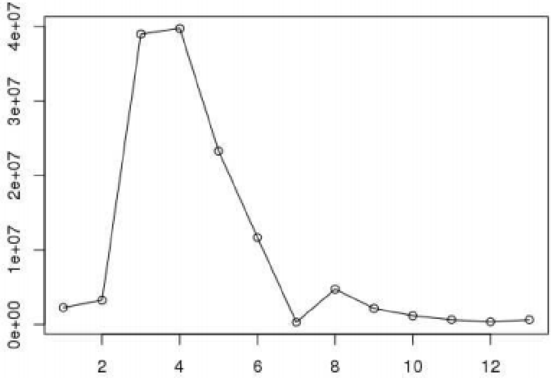
Sales by weeks:




# Two kinds of movies (II.)

A movie becomes popular later, e.g. *The Blair Witch Project (1999)*

Sales by weeks:




# Example

Data can be found for example at `http://www.the-numbers.com`

We will use weekly data (1.-7. day, 8.-14. day, etc.)

For example:

**Daily Chart Record**

| Date | Rank | Gross | % Change | Theaters | Per Theater | Total Gross | Days |
|---|---|---|---|---|---|---|---|
| 2003/12/17 | 1 | $34,450,834 | | 3,703 | $9,303 | $34,450,834 | 1 |
| 2003/12/18 | 1 | $17,019,987 | -51% | 3,703 | $4,596 | $51,470,821 | 2 |
| 2003/12/19 | 1 | $21,811,549 | +28% | 3,703 | $5,890 | $73,282,370 | 3 |
| 2003/12/20 | 1 | $27,492,053 | +26% | 3,703 | $7,424 | $100,774,423 | 4 |
| 2003/12/21 | 1 | $23,326,111 | -15% | 3,703 | $6,299 | $124,100,534 | 5 |
| 2003/12/22 | 1 | $13,563,208 | -42% | 3,703 | $3,663 | $137,663,742 | 6 |
| 2003/12/23 | 1 | $12,476,242 | -8% | 3,703 | $3,369 | $150,139,984 | 7 |
| 2003/12/24 | 1 | $7,544,400 | -40% | 3,703 | $2,037 | $157,684,384 | 8 |
| 2003/12/25 | 1 | $13,986,220 | +85% | 3,703 | $3,777 | $171,670,604 | 9 |
| 2003/12/26 | 1 | $19,152,196 | +37% | 3,703 | $5,172 | $190,822,800 | 10 |
| 2003/12/27 | 1 | $17,249,267 | -10% | 3,703 | $4,658 | $208,072,067 | 11 |
| 2003/12/28 | 1 | $14,196,641 | -18% | 3,703 | $3,834 | $222,268,708 | 12 |
| 2003/12/29 | 1 | $10,490,000 | -26% | 3,703 | $2,833 | $232,759,000 | 13 |
| 2003/12/30 | 1 | $9,615,000 | -8% | 3,703 | $2,597 | $241,937,000 | 14 |
| 2003/12/31 | 1 | $7,500,000 | -22% | 3,703 | $2,025 | $249,400,000 | 15 |

150139984 (days 1–7)

92233724 (days 8–14)

# Example

```
# The LOTR: Return of the King
Sales <- c(150139984,92233724,52192378,20100138,
           15302761,9109110,7300394,5612861,5995863,
           3809753,3140000,4062251,2923806,2055943)
```

**Exercise:** Estimate parameters of the Bass model and compare the actual and the fitted sales.