# Markov analytics of brand dynamics; technical notes
**Joel Rubinson, February 2, 2023**

## Markov assumptions

The key Markov idea is that you have a process where the next action has a probability of occurrence that is dependent on the last action, not the full journey of how the consumer got there. That implies that buyers of a given brand have a vector of probabilities of what their next purchase will be. This process then compounds and, since all transition probabilities are non-zero, there is also a non-zero probability that a consumer can go between any pair of brands in k steps. This will lead the system to eventually reach a steady state that we can calculate, that meet the detailed balance condition that the absolute number of consumers switching from brand i to brand j are equal for all i, j. This is under the assumption that the transition probabilities do not change. This is a typical assumption of stationarity of underlying probabilities that we make here as well.

## Building a mathematical framework

We start with an illustrative matrix **M** (shown at right). This is a prototypical brand switching matrix that can be viewed as a Markov transition matrix. A switching matrix is basically a cross-tab of what all consumers bought on their last two purchases (all brands down the side, and all brands across the top). Note that the diagonal represents the repeat rate of each respective brand and the off-diagonal terms are the switching probabilities of going from one brand to any other brand. Second, in practice, there are non-zero probabilities that a consumer can go from any brand to any other brand in k steps. Then, the Perron-Frobenius theorem applies, which says that we will be able to find one and only one equilibrium vector of steady state market shares and that this is irrespective of the market shares at any point in time.

| | A | B | C | D | E | F |
|---|---|---|---|---|---|---|
| A | 0.800 | 0.020 | 0.060 | 0.060 | 0.030 | 0.030 |
| B | 0.123 | 0.650 | 0.093 | 0.100 | 0.025 | 0.005 |
| C | 0.061 | 0.094 | 0.550 | 0.148 | 0.074 | 0.073 |
| D | 0.061 | 0.094 | 0.151 | 0.500 | 0.097 | 0.097 |
| E | 0.031 | 0.047 | 0.076 | 0.098 | 0.601 | 0.151 |
| F | 0.062 | 0.047 | 0.075 | 0.105 | 0.158 | 0.552 |

Examples of how to read this:
--repeat rates for Brand A, B, C are 80%, 65%, 55%, respectively
--3% of those who bought brand A switched to brand F. 6.2% of those who bought Brand F switched to Brand A. Note that this might be the same number of buyers if brand A is twice as large as brand F

**Finding equilibrium market shares.** Think of switching matrix, **M**, as something that can transform a vector of market shares from time t to time t+1 via the equation. $\mathbf{M}*v(t) = v(t+1)$.

All switching matrices are square and have eigenvector/eigenvalue structures…but in a Markov matrix, there is one pair that is particularly powerful.

**What are eigenvalues and eigenvectors?** An eigenvector is a special kind of vector that solves the equation, $\mathbf{M}*v_1 = \lambda_1 v_1$ (1),where $\lambda_1$ is the eigenvalue associated with that particular eigenvector $v_1$. Out of the infinite number of possible market share configurations, if there are G brands, at most, there will be only G eigenvalue/eigenvector pairs. One stands out; there is always one eigenvector whose eigenvalue is 1.

Plugging λ = 1 into equation (1), we get  M*v  = 1*v   (2)
In words, you wind up with the same shares you started with when you find this magic vector where v(t) = v(t+1) = v(t+2), etc., which is the definition of steady state.  It can also be proven that the steady state shares are independent of current market shares.  That is a powerful statement and can be used to spot brands that are likely to trend up or down from their current share.
The table below shows the comparison of Numerator data vs. "eigen-predicted" market shares for brands of frozen pizza (actuals from Numerator receipt scanning data).

Table 1: predicted vs. actual shares

| brand | actual share | eigen share |
|---|---|---|
| 1 | 14.8% | 14.8% |
| 2 | 19.0% | 19.3% |
| 3 | 10.1% | 10.2% |
| 4 | 5.6% | 5.7% |
| 5 | 4.0% | 4.0% |
| 6 | 3.6% | 3.5% |
| 7 | 2.5% | 2.5% |
| 8 | 1.7% | 1.7% |
| 9 | 3.4% | 3.4% |
| 10 | 2.1% | 2.1% |
| 11 | 2.6% | 2.6% |
| 12 | 30.6% | 30.4% |

**Simulating brand penetration without knowing brand market shares.** Most readers are familiar with the principle that brand penetration and market share are strongly correlated but how can penetration be estimated without knowing who the big vs. small brands are?  Actually, it IS possible to predict penetration for each brand with high accuracy just by knowing the Markov switching matrix.   From the Markov switching matrix, one can construct the Fundamental Matrix. This is a matrix of waiting times which is based on:

- Creating a new switching matrix by removing the row and column of the switching matrix that contains the brand of interest (conventionally called the Q sub-matrix)
- The Fundamental matrix is then $(I-Q)^{-1}$, where I is the identity matrix.
- This then gives the waiting times from each brand to the brand that as removed.  I took a weighted average of those waiting times

For our application, the waiting time of interest is how many cycles does it take for competitive brands to "send their customers" to your brand? Then if we know how long a cycle is (average category purchase cycle), we can calculate the half-life of waiting times. It is then trivial to use an exponential decay formula to estimate the build in monthly penetration (consumers sent to your brand). Note that the Fundamental matrix of waiting times is related to what is used to construct the Next Generation matrix for estimating $R_0$ in contagion modeling that tells us how fast a disease will spread. Systematically eliminating one brand at a time, we get a Fundamental Matrix for each brand which leads to predictions of 10 month penetration that are compared to actuals from Numerator receipt scanning data.

Table 2: Predicted vs. actual penetration

| Frozen pizza brand | Predicted 10 month penetration | Actual 10 Month penetration |
|---|---|---|
| 1 | 41.5% | 37.9% |
| 2 | 45.1% | 39.7% |
| 3 | 28.4% | 29.9% |
| 4 | 15.2% | 14.5% |
| 5 | 12.9% | 13.5% |
| 6 | 13.5% | 14.5% |
| 7 | 7.8% | 8.5% |
| 8 | 5.5% | 6.9% |
| 9 | 10.7% | 11.4% |
| 10 | 7.6% | 7.7% |
| 11 | 8.5% | 10.6% |

Source for actual data: Numerator receipt scanning

Why do repeat and brand to brand transition rates lead to accurate penetration estimates? Think of a stochastic process with balls in a box in an arcade game bouncing around due to air flowing from the bottom of the box where there is a hole at the top.  Most balls will bounce to all corners of the box but eventually, a ball will randomly bounce out of the box. If the hole is large, that will happen sooner than if it is small. Think of the balls as consumers and think of the sides of the box as the collection of competitors where consumers can repeat purchase or switch between those brands, staying in the box. Eventually, a consumer will bounce out of the box and the time it takes depends on the size of the hole which is determined by the sub system of competitors' repeat rates and "intra-box" transition probabilities (e.g., higher repeat rates = smaller hole). The Fundamental matrix is letting you calculate the average time it takes for the balls to bounce out and go into your box.

This stochastic process that can be raised to the kth power so even very small transition probabilities lead to average wait times that can be calculated and then converted into exponential decay functions based on the wait times using a half-life calculation.