

An Introduction to Forecasting with Markov Chains

As a marketer, you need to know how your customers and markets will evolve. You want to be able to answer questions such as which brands, stores, or vendors will customers choose in the future. With just a little data, these types of issues can often be illuminated using a technique known as a Markov chain. Marketers have found Markov chains to be powerful tools for anticipating, forecasting, and experimenting with market changes.

Suppose that you categorize your customers and prospective customers into three groups:

- Segment A - prospects that you would like to turn into new customers; they represent the overall target population for your product or service
- Segment B - new customers acquired within a given time period
- Segment C - acquired customers that have remained customers for a given time period

You can think about how people move from one segment to another in terms of origin and destination. Markov chains work on the probability that a given person starts in one segment at a point in time and moves to another segment in the next period. This information can be represented conveniently in a table like this:

| To: | From: | |
|------|-------|------|
| | A | B |
| A | 95% | 5% |
| B | 20% | 0% |
| C | 10% | 100% |
| Sums | 100% | 100% |

Note that the percentages in each column add up to 100%, indicating that each person must stay in their original segment or migrate to one of the other two. (The percentages are also known as Transition Probabilities, and the table is called a Transition Matrix.)

This table tells us how the percentage of people in each group will evolve over time. For example, from one period to the next, 95% of Segment A prospects will remain in Segment A; 20% of Segment B customers will defect and fall back into Segment A; and 10% of Segment C customers will also fall back into Segment A. From this we can calculate the size of Segment A in the next time period as:

$$\begin{aligned}
 \text{Segment A} &= & 95\% & & * (\text{Segment A}) \\
 &+ & 20\% & & * (\text{Segment B}) \\
 &+ & 10\% & & * (\text{Segment C})
 \end{aligned}$$

We can do similar calculations for each group:

$$\begin{aligned}
 \text{Segment B} &= & 5\% & & * (\text{Segment A}) \\
 &+ & 0\% & & * (\text{Segment B}) \\
 &+ & 0\% & & * (\text{Segment C})
 \end{aligned}$$

The above tells us that in each time period, 5% of the prospects in Segment A will become newly acquired customers.

$$\begin{aligned}
 \text{Segment C} &= & 0\% & & * (\text{Segment A}) \\
 &+ & 90\% & & * (\text{Segment B}) \\
 &+ & 80\% & & * (\text{Segment C})
 \end{aligned}$$

This tells us that 90% of Segment B customers acquired in the previous period will remain Segment B customers (i.e., 10% attrition rate) and that 80% of Segment C customers from the previous period will continue to be Segment C (i.e., 20% attrition rate).

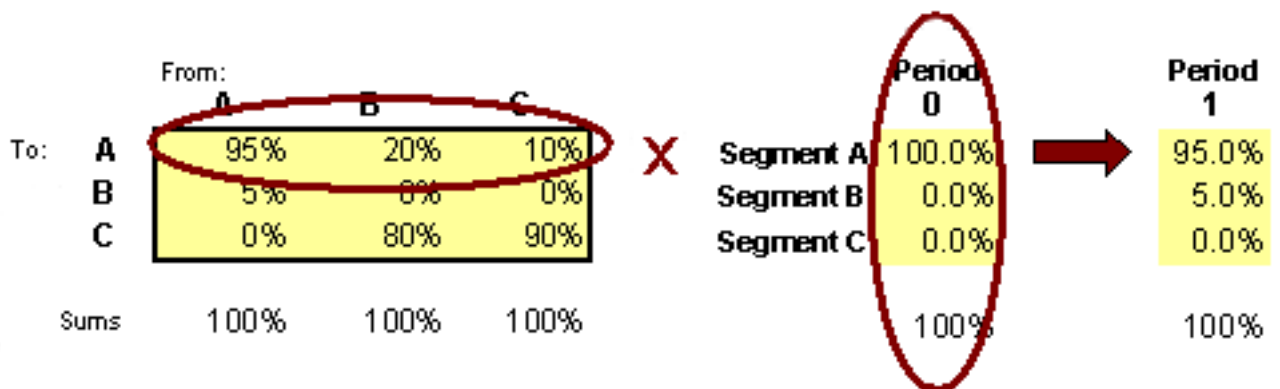
Forecasting Changes in Segment Size

If we know (or can reasonably guess) the percent of consumers that began in each segment, we can calculate what the percent will be for each segment in future time periods.

Let's look at how this calculation works in a spreadsheet. The column below contains initial segment data, referred to as "Period 0."

| | |
|------------------|-----------------|
| Segment A | Period 0 |
| Segment B | 100% |
| Segment C | 0% |
| | 0% |
| | 100% |

To find the percent in each segment for the next period (Period 1), we multiply the percentages in the Period 0 column by each of the rows of the Transition Matrix shown above. The math works like this:



Segment A in Period 1 = 95% + 20% + 10%
 * (Segment A in Period 0)
 * (Segment B in Period 0)
 * (Segment C in Period 0)

By applying the appropriate formulas, you can calculate Segment B and Segment C shares in all three groups for Period 1. These calculations produce a new column that can be used for computing market share in the next period, "Period 2".

Fortunately, Excel can perform these types of calculations for us. If you [download the sample file Segment.xls](#) and click on the tab "Segment", you will see that we have generated a table of segment share calculations for 60 periods. (Note that we've broken up the table here into four sections to facilitate viewing.)

| | | | | | | | | | | | | | | | |
|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|
| 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 |
| 100.0% | 95.0% | 91.3% | 88.0% | 85.3% | 82.9% | 80.9% | 79.2% | 77.7% | 76.5% | 75.4% | 74.5% | 73.7% | 73.0% | 72.4% | 71.9% |
| 0.0% | 5.0% | 4.6% | 4.6% | 4.4% | 4.3% | 4.1% | 4.0% | 4.0% | 3.9% | 3.8% | 3.8% | 3.7% | 3.7% | 3.6% | 3.6% |
| 0.0% | 0.0% | 4.0% | 7.4% | 10.3% | 12.8% | 14.9% | 16.8% | 18.3% | 19.7% | 20.8% | 21.8% | 22.8% | 23.3% | 23.9% | 24.5% |
| 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 | 27 | 28 | 29 | 30 | |
| 71.5% | 71.1% | 70.8% | 70.5% | 70.3% | 70.1% | 70.0% | 69.8% | 69.7% | 69.6% | 69.5% | 69.4% | 69.4% | 69.3% | 69.3% | |
| 3.6% | 3.6% | 3.6% | 3.5% | 3.5% | 3.5% | 3.5% | 3.5% | 3.5% | 3.5% | 3.5% | 3.5% | 3.5% | 3.5% | 3.5% | |
| 24.9% | 25.3% | 25.6% | 25.8% | 26.2% | 26.4% | 26.5% | 26.7% | 26.8% | 26.9% | 27.0% | 27.1% | 27.2% | 27.2% | 27.3% | |
| 31 | 32 | 33 | 34 | 35 | 36 | 37 | 38 | 39 | 40 | 41 | 42 | 43 | 44 | 45 | |
| 69.2% | 69.2% | 69.1% | 69.1% | 69.1% | 69.1% | 69.1% | 69.0% | 69.0% | 69.0% | 69.0% | 69.0% | 69.0% | 69.0% | 69.0% | |
| 3.5% | 3.5% | 3.5% | 3.5% | 3.5% | 3.5% | 3.5% | 3.5% | 3.5% | 3.5% | 3.5% | 3.5% | 3.5% | 3.5% | 3.4% | |
| 27.3% | 27.4% | 27.4% | 27.4% | 27.4% | 27.5% | 27.5% | 27.5% | 27.5% | 27.5% | 27.5% | 27.5% | 27.5% | 27.6% | 27.6% | |
| 46 | 47 | 48 | 49 | 50 | 51 | 52 | 53 | 54 | 55 | 56 | 57 | 58 | 59 | 60 | |
| 69.0% | 69.0% | 69.0% | 69.0% | 69.0% | 69.0% | 69.0% | 69.0% | 69.0% | 69.0% | 69.0% | 69.0% | 69.0% | 69.0% | 69.0% | |
| 3.4% | 3.4% | 3.4% | 3.4% | 3.4% | 3.4% | 3.4% | 3.4% | 3.4% | 3.4% | 3.4% | 3.4% | 3.4% | 3.4% | 3.4% | |
| 27.6% | 27.6% | 27.6% | 27.6% | 27.6% | 27.6% | 27.6% | 27.6% | 27.6% | 27.6% | 27.6% | 27.6% | 27.6% | 27.6% | 27.6% | |

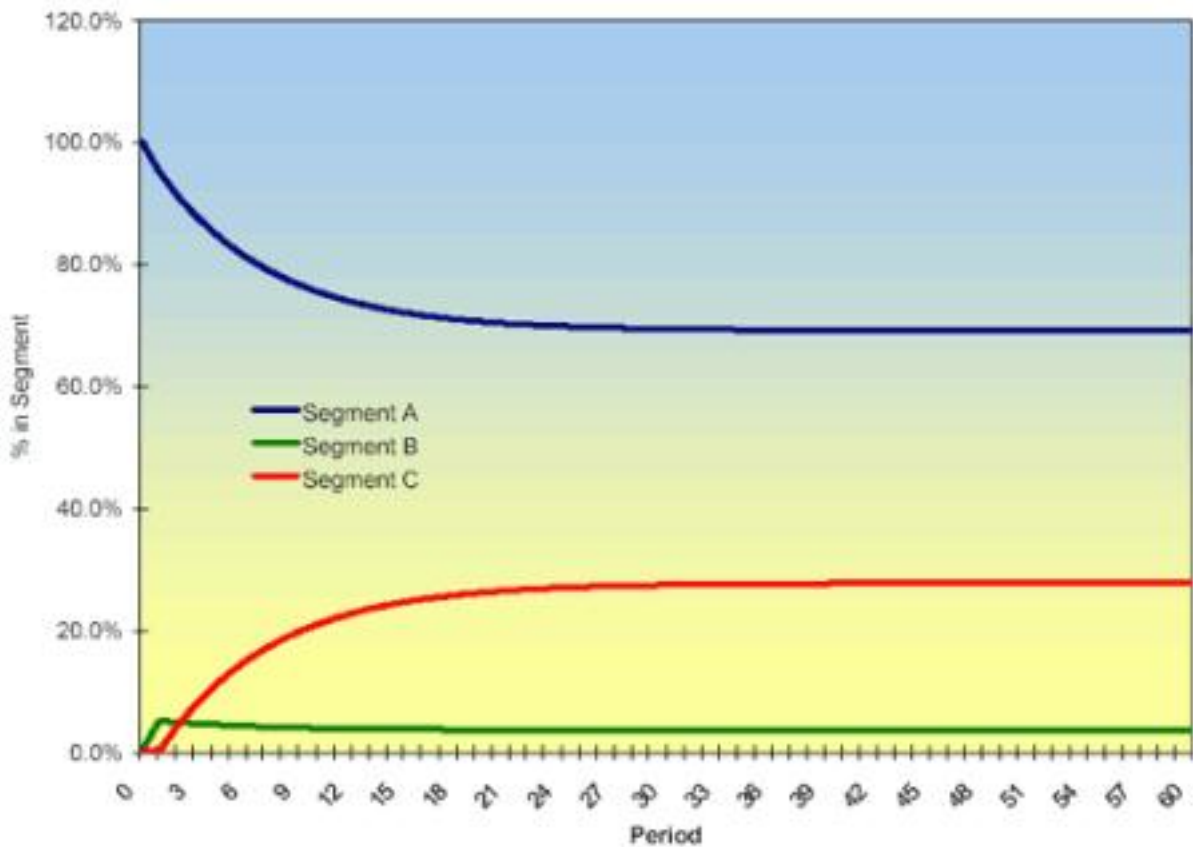
Looking at the table, you can see that the patterns begin to become stable after approximately 20 periods and completely stable by Period 45 with the following values:

Period 45

- 69%
- 3.4%
- 27.6%

This stable state is called "equilibrium."

We can graph the shift between segments over time on the "Graph" tab. You will see the graph shown below.



Transition Matrix vs. Starting Shares

To get a sense of the role played by the Transition Matrix in determining the eventual equilibrium shares, let's look at what happens if we assume that at Period 0 we start with 100% of the market in the Segment B.

- Segment A**
- Segment B**
- Segment C**

Period 0

- 0%
- 100%
- 0%

Change the share values to the above numbers for Period 0 in the "Segment" tab.

Notice that up to the 30th period or so, the segment shares are quite different from those that we saw earlier, but that by Period 53 they settle down to the same values:

Period 53

69%
3.4%
27.6%

In fact, no matter what values you insert for Period 0 (try it yourself; remember that the sum must equal 100%), segment shares will eventually attain the same equilibrium values.

This powerful property of Markov chains suggests that it is more important to focus on the probabilities in the Transition Matrix than it is to fuss over the initial distribution of market share. By experimenting with different transition probabilities, you can understand their effect on long-term market share.

For example, what if we were able to increase our retention rate of Segment C customers from 90% to 92% (via actions such as improved customer service or loyalty programs)? The Transition Mmatrix would now look like this:

| | | |
|------------|--------------|----------|
| | From: | |
| | A | B |
| To: | A | 95% |
| | B | 5% |
| | C | 0% |
| Sums | 100% | 100% |

And the new equilibrium looks like this:

Period 53

64.5%
3.2%
32.3%

Compare this with the equilibrium segment shares in the original matrix:

Period 53

69%
3.4%
27.6%

An improvement in Segment C retention rate from 90% to 92% resulted in an increase in Segment C customers from 27.6% to 32.2% — a gain in share of 4.6% (with a small decline in Segment B). Try experimenting with changes in the original Transition Matrix to see their impact on equilibrium segment share.

Other Applications

By modifying the Markov chain template, you can model things like brand switching, store choice, and customer relationship management. With the addition of data on customer profitability by segment, Markov-based forecasting models can be adapted to determine which segments should be targeted for increased marketing investment. (See the example at [Will This CRM Program Pay Out Soon?](#))

There is an IMPORTANT caveat to bear in mind regarding Markov chains. We are assuming that Transition Probabilities don't change over time. Therefore, when discussing results with others, be sure to warn: "This is what happens if things don't change." If you want to model what happens if something causes the Transition Probabilities to change at a certain point in time, run one Markov chain with the original assumptions, then take the last segment shares before the change and use those as the Period 0 shares when you adjust the Matrix.

For more information about Markov chains and their application, we refer you to the book and software package *Decision Making with Insight* by Dr. Sam L. Savage.

