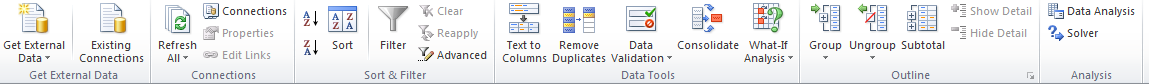
Retention Exercise Walkthrough

In this guide, we will develop two models that can be used to forecast customer retention. We will first develop a model based on a geometric process, which assumes a constant churn probability. We will then generalize this model to allow for the churn probability to change over time. For each of the models, you will find a worksheet on which you will conduct your work (labeled “working”) and a worksheet that is complete and has been annotated (“COMPLETE”).

In both cases, we will use the Solver add-in. This should appear on the Data ribbon in Excel.



If you do not see this, you can add Solver as follows. Go to File 🡪 Options 🡪 Add-Ins 🡪 Manage: Excel Add-Ins. Click the box next to Solver Add-In to check it.

Model 1: Geometric Model

To begin, we need to calculate the probability that an individual will drop service after t months. In cell D3, enter the formula:

=$K$1\*(1-$K$1)^(A3-1)

This corresponds to a geometric process in which the monthly probability that service is cancelled is given by the probability in cell K1. This customer does not drop service (with a probability of 1 – K1) for A3 – 1 months, and then drops service in month A3 with a probability of K1. Copy this formula until cell D9.

In cell D10, we need to calculate the probability of a customer maintaining service beyond the seven months. This can be written as:

=(1-$K$1)^(A9)

Now that we have calculated the probability associated with an individual dropping service after a given number of months, we move on to calculate the log likelihood of the observed data. In cell E3, enter the following:

=C3\*LN(D3)

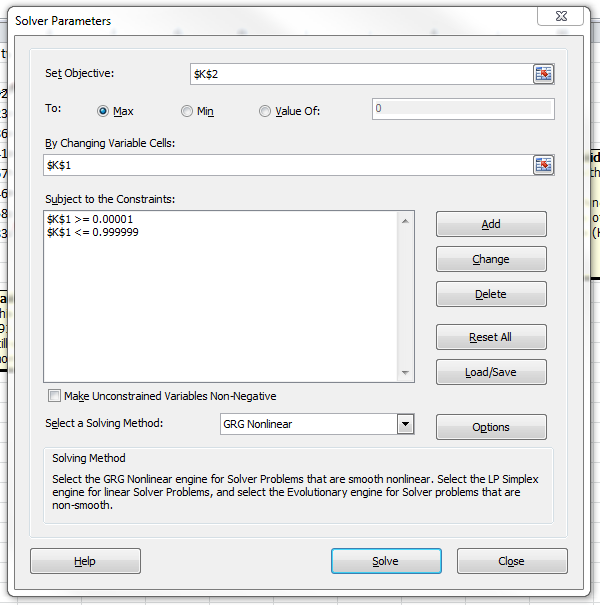
For each of the C3 customers who dropped service after A3 months, the log of the likelihood associated with this observation is D3. Copy this formula through cell E9. In cell E10, the formula must consider the number of customers who have not yet dropped service, which is given in B9. In E10, the formula is:

=B9\*LN(D10)

In cell K2, construct the log likelihood of the data by summing the calculations in E3 through E10:

=sum(E3:E10)

To estimate the maximum likelihood estimate for p, click on the Solver add-in on the Data ribbon. Set the objective cell to K2 and the objective to “Max” by changing variable cells K1. Since K1 is a probability, it is bounded by 0 and 1, so we need to add constraints. In specifying the constraints, Excel does not allow for constraints to be < or >, but instead <= or >=. To work around this, we will specify the constraints as K1>=.00001 and K1<=.99999. You can add the constraints by clicking on “Add”. The Solver menu screen should appear as follows:



Click on Solve to estimate the geometric model, the results of which are on the “COMPLETE” worksheet.

Model 2: Time Trend

To build the model with a linear time trend, our approach will be to make the probability of dropping service a function of time. We begin in column D by calculating the probability that a customer still has service at the end of a time period. In D3, enter the following:

=D2\*(1/(1+EXP($L$1+($L$2\*A3))))

This assumes that the probability of having service after the first month is the product of (1) the probability that you had service through the previous month (given by D2) and (2) the probability that you keep service through the current month. If we use a binary logit model to specify the probability that you drop service in a given month, given by:

*p*=exp(L1+(L2\*A3))/(1+ exp(L1+(L2\*A3)))

then the probability that you keep service is given by 1 – p, or

1 – *p* = (1/(1+EXP($L$1+($L$2\*A3))))

Copy this formula through cell D14.

Once we have calculated the probability that a customer maintains service through month t, we can then calculate the probability that a customer maintains service until month *t* and drops service in month *t*. In cell E3, enter:

=D2-D3

and copy this formula down through cell E9. To account for those customers who maintain service beyond seven months, in E10, enter:

=D9

To calculate the log likelihood, as for the geometric model, we combine the number of customers who drop service in a given month with the log of the probability with which this happens. In F3, enter:

=C3\*LN(E3)

and copy this formula through F9. In F10, to account for those who have not yet dropped service, enter:

=B9\*LN(E10)

Based on the current values of L1 and L2, we can forecast the number of remaining customers in column L. In G1, enter:

=1000\*D2

Copy this formula through cell G14.

In L3, calculate the log likelihood of data:

=SUM(F3:F10)

Using Solver, set the objective as maximizing cell L3 by changing cells L1 and L2. Be sure that the box labeled “Make unconstrained variables non-negative” is not checked. Run Solver.

On the Comparison worksheet, you will find a comparison of the geometric model and the time trend model through the first 12 months of data.