Proceedings of the 2020 IISE Annual Conference
L. Cromarty, R. Shirwaiker, P. Wang, eds.

Estimating True Demand at a Local Hunger Relief Organization

Abstract ID: 790752

Kehinde Odubela, Steven Jiang and Lauren Davis
Department of Industrial & Systems Engineering
North Carolina A & T State University
Greensboro, NC 27411, USA

Abstract

Hunger relief organizations are mostly non-profits that collect food from various sources and redirect them to the people in need. This is to combat the prevalent food insecurity affecting children, the unemployed, students, seniors and so on. Previous research has focused on the demand/donation side of food rescue operations, but the distribution or supply side - especially in reducing the uncertainty associated with food demand - has received significantly lower attention. In this study, we obtained data from a local hunger relief organization, specifically a food pantry to develop estimates of the demand they expect to receive in the future. To do this, we fit the growth of the food pantry client population to a logistic growth model to obtain a good fit. We then obtained data for frequency of visits to develop estimates of the number of visits expected in the future, using time series models. This will be combined with the allocation policy for food distribution to develop estimates of true demand. This study has merit for hunger relief organizations. It will aid decision making relative to food distribution, while also providing data for planning purposes.

Keywords
Demand, Logistic Growth, Hunger Relief, Food Pantry

1. Introduction

Food banking is one part of the broad safety net available to people requiring food assistance. Food banks receive food and support from the government and other donors such as manufacturers, farmers, grocery stores etc. through a network of branches. The food is then distributed to agencies such as food pantries, group homes, and soup kitchens who then distribute the food to people in need. An individual or household is considered food insecure when they have limited access to safe and nutritious food required for an active, healthy life [1, 2]. In their study, [3] found that households that were food insecure at some time during the year were food insecure on an average of 7 out of the 12 months in the year. Food insecurity is also prevalent among college students. One study found that approximately half of 2- and 4-year college students surveyed were food insecure [4].

The studies above point to the changing population of those requiring food assistance all year round. Even so, food banks must keep up with changing demand and supply patterns. They work to ensure the limited food they receive is equitably distributed based on the need in their service area. However, data-driven estimates of the true demand by the people in need are not available. Previous research has focused on developing forecasts of the amount of donation food banks expect to receive using time series [5] and other machine learning models [6]. However, research is scant that has studied the distribution or supply side of food bank operations, especially in reducing the uncertainty associated with food demand.

In this research, we use data from a local food pantry in Greensboro, North Carolina, to develop estimates of the amount of food they expect to distribute in the future. This food pantry typically opens twice a week – on Tuesdays and Thursdays - for client visits. Only for special purposes does it open three times a week. Clients are made to register on their first visit during which they answer questions about their household size, demographics, type of utensils they have for cooking, and so on. The food pantry operates a client-choice model whereby those seeking food assistance choose for themselves what products they want. This model has the potential to reduce food waste as clients do not have to take items they already have or foods they cannot eat for health or other personal reasons. It also makes ordering easier and ultimately upholds the dignity of the clients. More so, clients are provided a shopping list indicating the type of foods available from which they can choose. However, there is a cap on the amount of each food category that the client can choose from. The food pantry has an allocation policy based on the household size of the clients requiring food assistance. It also has a limit on how frequently the same client can visit, usually no more than once a week or twice a month.

This study has benefit for hunger relief organizations as it provides a methodology for estimating the true demand they expect to receive in the near future. This will aid decision making relative to food distribution while also providing data for planning purposes. Ultimately, the organization can better serve its clients and improve on its effort to close the hunger gap.

2. Method

2.1 Data Collection

We obtained visitation data from a local food pantry to determine how many people or households who are registered that use the food pantry's services. The data was aggregated on a weekly basis for the period for which the food pantry was opened from February through December 2019. The food pantry did not open in August 2019. Open days ranged from one to three times per week. Data for 36 weeks was obtained in total. Sample data for February through April is shown in Table 1 below.

Table 1: Sample Data

Month	Week	0 Visit	1 Visit	2 Visits	Total	New Clients
Feb	1		24	0	24	24
Feb	2	22	32	0	54	30
Feb	3	51	14	0	65	11
Feb	4	57	17	0	74	9
Mar	5	74	0	0	74	0
Mar	6	68	11	0	79	5
Mar	7	71	11	0	82	3
Mar	8	76	10	0	86	4
Apr	9	81	9	0	90	4
Apr	10	84	14	0	98	8
Apr	11	89	14	0	103	5
Apr	12	94	18	0	112	9

The data was analyzed to obtain the number of no-shows, one and two visits per client recorded per week. No-show (0 visit) refers to a client that registered (and visited at least once) at some point in the past but did not visit subsequently. The number of new clients was simply calculated as the difference between the total number of visitors in week, *t* and the immediate past month, *t*-1.

2.2 Population Growth Models

We considered two simple, deterministic models of population growth to represent the growth in the client population that visited the food pantry – exponential and logistic growth models. The models are discretized with weekly time step based on the data collected from the food pantry. For both models, the population size (N_t) in week *t* is expressed in terms of the total population (N_{t-1}) in the immediate past week (*t*-1) added to the number of new clients (growth), which is a fraction of N_{t-1}. The exponential growth model assumes the client population continues to grow at a constant annual rate (R_{max}) [7]. This is represented in equation (1) below:

$$N_t = N_{t-1} + N_{t-1}R_{max} \tag{1}$$

The exponential growth model does not prove true forever because it depends on infinite amount of resources. However, populations can witness exponential growth before approaching what is known as the carrying capacity (*K*). That is, when the population size gets large enough, the growth rate begins to slow down because resources are more quickly used up. The growth rate ultimately levels off and the growth curve makes an S-shaped curve – the logistic growth curve. *K* is an input in the logistic growth model and is the maximum population size a specific environment can support, that is, the point at which the growth curve plateaus. The parameter, z is the exponent which governs the shape of the curve. The standard logistic curve is obtained when z = 1 [7]. See equation (2) below for the logistic growth model.

$$N_t = N_{t-1} + N_{t-1}R_{max} \left[1 - \left(\frac{N_{t-1}}{K}\right)^z\right] \tag{2}$$

Ideal exponential growth curve is J-shaped and logistic growth curve is S-shaped as shown in figure 1 below.

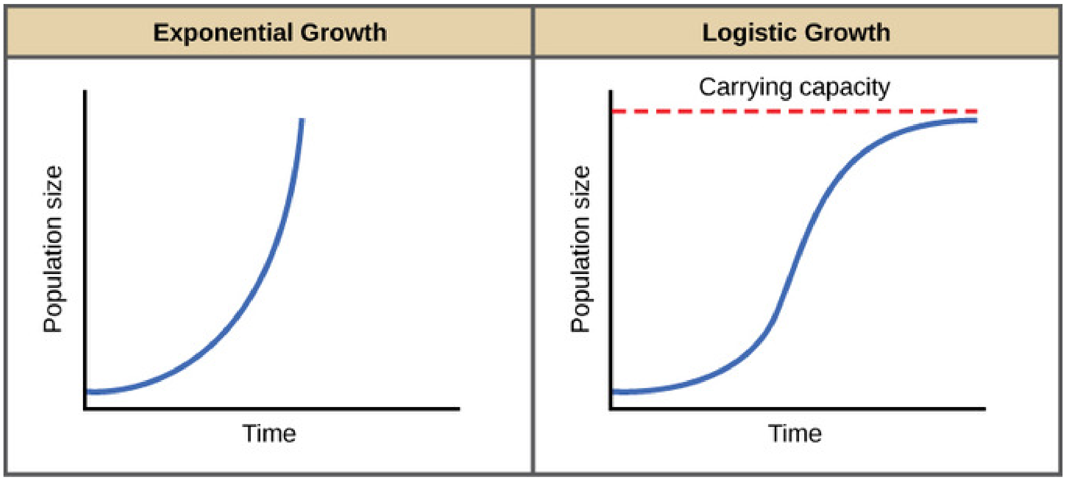

Figure 1: Exponential and Logistic Growth Curves [8]

2.3 Forecast Models
Of the total population registered to use the food pantry, we know that only a fraction visits the food pantry once a week, and a much smaller fraction visits twice a week. Unlike the population which continues to grow over time, the number of clients who visit, at least once, can increase or decrease from month to month. Therefore, we applied time series models to examine if the number of visits have a time-dependent structure, and thereafter, develop estimates of the number of weekly visits.

2.3.1 ARIMA Models
The autoregressive integrated moving average (ARIMA) assumes that subsequent values of a variable linearly depend on its past values and the values of past, random shocks [9]. To make a forecast, a multivariate model is considered when a relationship exists between the predicted values (dependent variable) and one or more independent variables. The aim is to forecast the dependent variable with more accuracy compared to using only the time series values. This is also known as the ARIMAX model, where 'X' represents the endogenous or independent variable. Conversely, a univariate model is used in the absence of an independent variable [10].

2.3.2 Holt Winters' Exponential Smoothing
The Holt Winter's method is a generalization of the exponential smoothing methods. It is useful for modeling univariate time series with trend or seasonal components. Past observations are assigned an exponentially decreasing weight, and these weights are summed to generate the forecast [11].

3. Results and Analysis
3.1 Growth Model Fitting
The growth model was built based on the weekly visitation data collected from the food pantry. This involved finding model parameters which provided the best fit to the data. Initial population size, N_0 = 24 is the population in the first week used as the base week and designated as t = 0. It also serves as the intercept of the growth curve. Several values were considered for the parameters (except the initial population) until a model with the best fit was obtained. These are R_{max} (maximum growth rate), K (carrying capacity) and z (exponent). The logistic growth model in equation (2) was applied to the data. The parameter values are shown in table 2 below.

Table 2: Model Parameters

Description	Parameter	Value
Max rate of increase	R_{max}	0.12
Initial population	N_0	24
Carrying capacity	K	750
Exponent	z	2

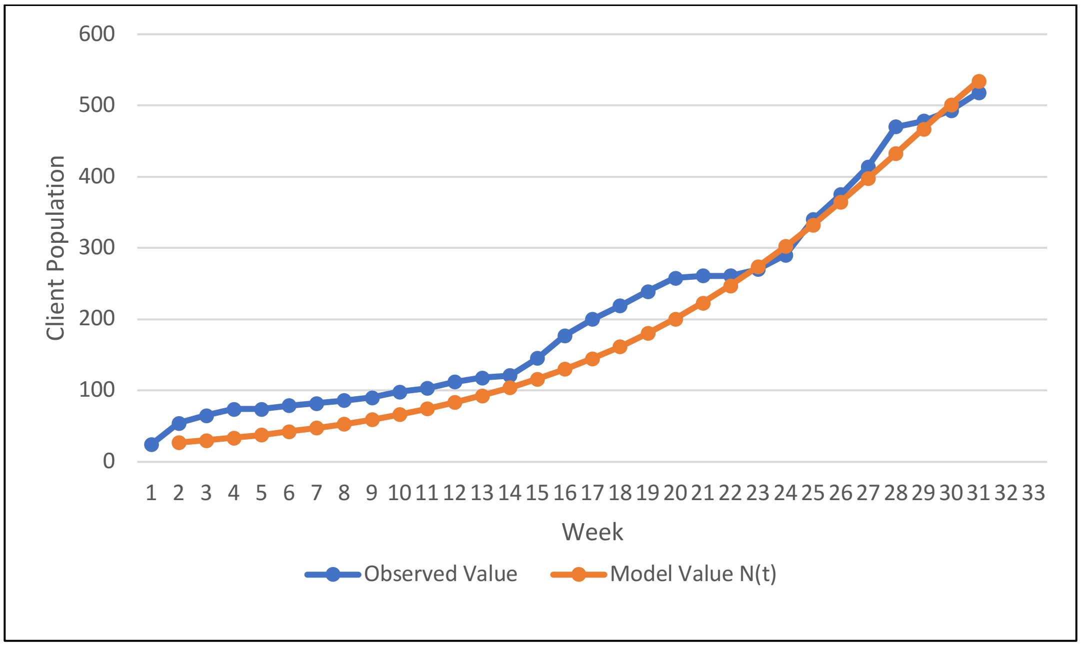

Figure 2: Observed and model values of client population size

As can be observed from figure 2, the logistic growth model captures the population growth fairly well, even though the model lags the observed values for the most part. The growth curve demonstrates apparent slowing of the growth rate as the population approaches current carrying capacity.

3.2 Measure of Goodness of Fit

For the population growth model, the goodness of fit was measured using the two-sample Kolmogorov-Smirnov (K-S) test to determine whether the standardized residuals are normally distributed [7]. The two-sample K-S test is a non-parametric test that quantifies the distance between the empirical distribution functions of two samples. The null distribution of the statistic is estimated using the null hypothesis that the samples are drawn from the same distribution and that the two distributions are identical.

For the logistic growth model developed in this study, $D = 0.233$ was obtained as the two-sample K-S statistic with p-value = 0.342. Therefore, at significance level $\alpha = 0.05$, we fail to reject the null hypothesis that the two distributions are identical. Therefore, the logistic growth model developed properly captures the population growth of the clients using the food pantry data considered in this study. Consequently, we can also represent the number of new visitors to the food pantry on a weekly basis.

3.3 Forecasting Number of Visitors

Data for the weekly number of visits to the food pantry was used to develop estimates of how many visits they should expect in the future. Before applying the time series models discussed in section 2.3 above, February data – the first month - was removed because this month had significantly larger number of visitors (relative to the total population) compared to other months. Therefore, data for 31 weeks starting from March was used to build the time series models. The data was divided into training and testing sets (80:20 split). Three time series models – Holt Winters' Exponential

Smoothing, ARIMA and ARIMAX were considered, and the models were built with Python and R programs. We employed walk-forward validation in the modeling process such that the training set expands each time step and the test set is fixed one-step ahead. This ensures that the model is updated each time step new data is added considering the small dataset presently available for our study.

3.3.1 ARIMA Models

We explored the possibility of predicting the number of visits with the registered population size at time, *t*. This was because the number of visits and population size were highly correlated (correlation coefficient, $r = 0.76$). Therefore, we fit the ARIMAX model to the visitation data with population as the endogenous variable. A seasonal decomposition of the data indicated the presence of trend and seasonal components. This is shown in the figure below.

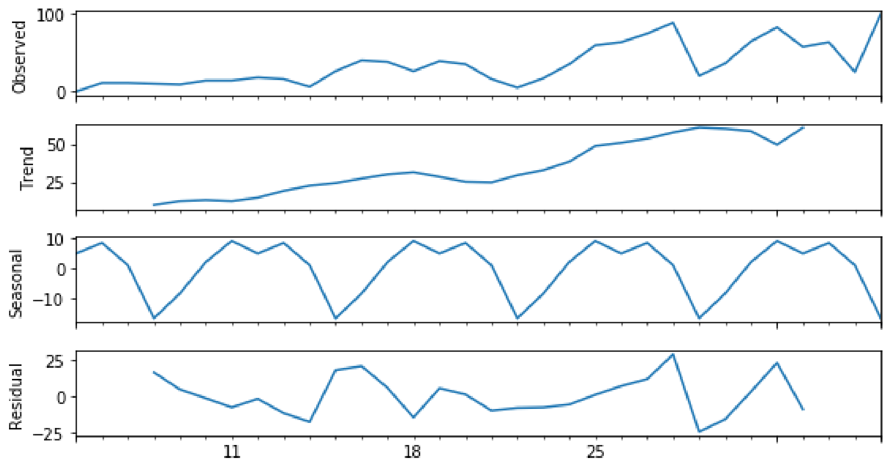

Figure 3: Decomposition of the time series of number of visits to the food pantry

We grid-searched a suite of ARIMA hyperparameters to identify the combination of p, d, q (trend, seasonality and white noise) parameters with the best performance based on the root mean square error (RMSE). ARIMA (0,1,1) was identified as the best performing combination and was applied to the training set. Mean Absolute Percentage Error (MAPE) was used to measure the accuracy of the models.

Based on the test set, the ARIMAX model returned a MAPE of 56.6%, which is very high. The univariate ARIMA model was also applied only to the visitation data. The test set revealed a MAPE of 48.4%. The ARIMAX model has a limitation in that it requires an observation for the endogenous variable for every future value of the time series it forecasts. However, the pure ARIMA model can forecast farther into the future, even though the skill of the model degrades with time when used to forecast for a much longer horizon.

3.3.2 Holt Winters' Exponential Smoothing

To generate the forecast of the number of visits, a seasonal period of 5 weeks was used as these generated lower percentage error compared to 4 weeks. The test set returned a MAPE of 37.11%.

4. Discussion and Conclusion

In this study, we obtained visitation data for a local hunger relief organization. The data was aggregated weekly to develop estimates of the total number of clients using the food pantry's services and the new clients that were added every week for the period considered. We then modeled the population growth by fitting the data to a logistic growth model. We obtained a model with a good fit and therefore, can develop estimates of the population growth in the future.

Of the total registered population, we recognize that a fraction visits the food pantry one or more times in a week. Hence, we used time series models to capture the changing number of visits received every week. Even though we expected the introduction of the population variable to improve the accuracy of the forecast, the univariate time series model performed better than the multivariate model, despite a high correlation between the number of visits and the population size. More so, both ARIMA models performed lower than the Holt-Winter's method. We note the small size of the data available for our study. Therefore, the accuracy of these models can be improved as more data becomes available. Machine learning models will also be considered, even though this study revealed trend and seasonal components inherent in the number of visits made to the food pantry.

Fitting the population growth to a distribution and developing estimates of the number of visits to the food pantry are the first steps in our quest to estimate the true demand at this local hunger relief organization. As a next step, we will obtain data for the allocation policy of the food pantry, based on household size, since the quantity of food given to the client depends on the number of people in the client's household. We can then obtain the true demand by combining the distribution of the client population, the frequency of visits by the registered clients and the allocation policy used for food distribution by the food pantry.

Data-driven estimates of the true demand received by a hunger relief organization helps to reduce the uncertainty surrounding meeting the needs of those requiring food assistance. This research provides a method for estimating the true demand for charitable food that can be applied to other agencies involved in the fight against hunger. We will also apply this technique to the broader agency-food bank relationship since larger food banks also treat individual agencies as clients that have specific times when they visit for shopping. Agencies are also sometimes categorized as small, medium and large, similar to the way a food pantry allocates food to its clients based on household size. This research contributes to the supply side of food bank operations which has received less attention among researchers over the years.

References

1. Coleman-Jensen, A., Gregory, C., Singh, A. (2013). "Household food security in the United States in 2012," USDA-ERS Economic Research Report (173).
2. Campbell, C.C., (1991). "Food insecurity: a nutritional outcome or a predictor variable?" Journal of Nutrition.121(3),408–415.
3. C Alisha Coleman-Jensen, Matthew P. Rabbitt, Christian A. Gregory, and Anita Singh. (2017). Household Food Security in the United States in 2016, ERR-237, U.S. Department of Agriculture, Economic Research Service.
4. Broton, Katharine & Goldrick-Rab, Sara. (2017). Going Without: An Exploration of Food and Housing Insecurity among Undergraduates. Educational Researcher. 47. 0013189X1774130.
5. Lauren B. Davis, Steven X. Jiang, Shona D. Morgan, Isaac A. Nuamah, Jessica R. Terry. (2016). Analysis and prediction of food donation behavior for a domestic hunger relief organization. International Journal of Production Economics 182, 26-37.
6. Farrokhvar, L., Ansari, A., & Kamali, B. (2018). Predictive models for charitable giving using machine learning techniques. PloS one, 13(10), e0203928. doi:10.1371/journal.pone.0203928.
7. Huber, H. and Laake, J. (2002). Trends and status of harbor seals in Washington state: 1978-99. In *AFSC Quarterly Report*, 1-13. Retrieved from http://www.afsc.noaa.gov/Quarterly/ond2002/ond02feature.pdf
8. Environmental Limits to Population Growth. (2016). OpenStax, Biology. OpenStax CNX. May 18, 2016. Retrieved from http://cnx.org/contents/185cbf87-c72e-48f5-b51e-f14f21b5eabd@10.12.
9. Box, G. E. P. & Jenkins, G. M. (1976). Time Series Analysis Forecasting and Control, 2nd edition. San Francisco: Holden-Day.
10. Pereira, A. (2004). Performance of time-series methods in forecasting the demand for red blood cell transfusion. Transfusion. 44. 739-46.
11. Winters, P.R. (1960). Forecasting sales by exponentially weighted moving averages. Management, 6, 324–342.