Using Survival Modeling for Turn-Time Predictions in Foodservice Settings

Mark P. Legg
Lisa Slevitch
Oklahoma State University

ABSTRACT

Within the competitive foodservice industry, the ability to accurately predict the length of the meal process known as turn-time is critical to the success of the firms in the industry. This is traditionally done through multiple least squares (linear regression) technique. However, linear regression lack the characteristics needed to accurately predict time durations, while survival models were designed for that purpose. This study utilized simulated data of a dine-in restaurant to test and compare the ability of linear regression to five survival models (proportional hazard models) to accurately predict the duration of turn-time. The results from the simulated trials show that while some of the survival models held marginal improvements, linear regression performed adequately for predicting duration of turn-time as compared to the survival models. For practitioners interested in the practical ease of the models, linear regression is recommended while practitioners interested in incremental improvements may opt for survival models.

KEY WORDS: Survival modeling, proportional hazards models, turn-time, revenue management, simulations
INTRODUCTION

As the competition within the foodservice industry has increased, the need to accurately predict factors that are key determinants of restaurants’ success becomes more important. One of such critical factors affecting consumers’ perceptions of a restaurant service quality is the amount of time the meal process will take (Richard, Sundaram, & Allaway, 1994). This duration of the meal process is known as turn-time (Thompson & Kwortnik, 2008).

The importance of turn-time is two-fold: first, it is a key determinant in patrons’ decision making process; second, it plays a vital role in practitioners’ ability of efficiently managing the dining floor and making profits (Kimes & Thompson, 2004). The lack of the latter can potentially lead to higher marginal costs due to greater periods in the lag between usages of tables and higher probabilities of patrons being turned away during busy meal periods (Sill, 1991). Thus, the foodservice facilities that are able to accurately predict turn-time can develop more consistent models for their revenue management systems.

Given the importance of the topic, it is no surprise that the question how to effectively predict turn-time in commercial foodservice has been actively explored by hospitality researchers and practitioners (Kimes & Thompson, 2004). Nevertheless, certain areas of the turn-time theory are flawed. Often, predictions of turn-time lack reliability. As patrons make their way through their meal process, a variety of factors that hold influence over turn-time may occur and interfere (Thompson & Kwortnik, 2008). These influential on turn-time variables are known as time-varying variables (Klien & Moeschberger, 2003). Examples of these factors vary drastically, from patrons abruptly ordering dessert or quickly asking for a check. That is why predicting turn-time with time-varying variables has typically caused lack of consistency in traditional regression studies (Golub, 2007).

The significance of the current study stems from the attempt to address this gap by employing survival analysis that has not been commonly used in hospitality research. While traditional regression techniques lack robustness in measuring time-processes with time-varying variables, survival analysis was developed for measuring time-processes and can take into account time-varying variables (Golub, 2007). A number of studies in a variety of fields have shown that survival analysis predicted time better than traditional regression methods (Gokovali, Bahar & Kozak, 2007).

To reiterate, the objectives of the current study can be summarized as following:

- To examine applications of survival analysis techniques for turn-time in foodservice settings.
- To compare survival modeling results of turn-time with the most widely used traditional regression method, multiple least-squares (MLS).

LITERATURE REVIEW

Turn-time

Plentora of studies have found reliable predictions of the quality of food, delivery of service, and costs of meals through traditional regression techniques based on well-known scales (Becker, 1991; Cardello, 1995; Ganoos, 1993). However, the duration of turn-time itself has been difficult to predict consistently (Kimes, Wirtz, & Noone, 2002). Managers knowing the length of the meal process based on characteristics of the party are able to book reservations with
fewer lapses in table usage time as well as know how to fine-tune the waiting staff schedule (Kimes et al., 2002). Additionally, the ability to accurately predict turn-time has been suggested to improve service quality (Thompson & Kwortnik, 2008).

The lack in ability to precisely predict turn-time of patrons is attributed to various characteristics that hinder predictions of time in traditional regression models (Kimes et al., 2002; Gokovali et al., 2007). First, while time can only be positive, traditional regression models may result in negative prediction values of time (Gokovali et al., 2007).

Second, the duration of the meal process is dependent on variables that cannot be easily predicted in advance as they occur and change during the process of the meal. These time-varying variables are problematic because they can violate the assumptions of conditional independence in least-squares models (Gokovali et al., 2007). As changes occur during the duration of the event, least-squares models results will be bias as the time-varying variables influence the duration of the event, but the predicted values remains static (Gokovali et al., 2007; Stock & Watson, 2007).

Lastly, time does not always follow a normal distribution (Peister, 2007). Modeling non-normal distributions through least-squares will violate the normality assumption and cause bias predictions in the results (Stock & Watson, 2007).

Even though traditional regression models lack robustness when measuring time-varying variables, studies have utilized an assortment of traditional techniques to find the factors that influence turn-time (Bell & Pliner, 2003; Kimes & Mutkoski, 1989; Ma & Jatoi, 2007). It was found that the size of the party dining at a table not only influenced turn-time but also held a direct influence on the per-patron calorie intake (Bell & Pliner, 2003). As patrons increased their calorie intake they typically stayed longer, this could be due to the additional time it takes processing non-entrée (Wansink, Payne & North, 2007). Lastly, types of payment, meals and time of arrival can also influence the duration of the meal-process (Kimes & Mutkoski, 1989; Kimes & Thompson, 2005).

However, as each of the turn-time studies contributed to the ability to understand turn-time, the results of these studies come with a degree of limitations due to the limited robustness in the models used in the studies (Gokovali et al., 2007). To adequately remedy these limitations, survival analysis is suggested and discussed in the next section.

Survival Analysis

Survival analysis measures the probability of how long it will take a given outcome to occur for a group of similar individuals known as failure times (Cox & Oates, 1984). Furthermore, survival analysis also measures the probability that the given event in question will last until the next point in time (Gokovali et al., 2007). It should be noted though that the incident itself does not need to hold negative connotations (Bland, 2004).

The most useful benefits to survival models are keen to their growth in popularity. First, survival models are able to take time-varying variables into account within the modeling process (Golub, 2007). This is mostly done through proportional hazard (PH) models (Barros et al., 2010). Second, survival models are not restricted by the assumption that the distributions of the variables in the data need to be normal (Sloot & Verschuren, 1990). Third, survival models only produce positive predictions of time (Gokovali et al., 2007). It follows, that time has the potential to not follow a normal distribution, needs to be positive in predictions and is influenced by time-varying variables. Therefore, this study expects that survival models will outperform MLS model with predicting of turn-time.
Hypothesis: Predictions of the duration of turn-time through survival models will be more precise than predictions of the duration of turn-time through traditional regression models.

METHODS

In order to test the predictive abilities of MLS and the survival models, simulated data was developed via probability distributions for creating the values of the variables influencing turn-time. Turn-time was set as the total time spent in the restaurant. This includes time spent waiting, talking and eating from each part of the meal process.

A total of 1000 random dine-in party observations were created for the simulated dine-in restaurant that was assumed to be open seven days a week from 8AM to 8PM with the end part of the week excluding Sunday being busier. In addition, breakfast, lunch and dinner were set as the busiest times of the day while late-morning, mid-afternoon and late-dinner were not as busy. The meal process for the restaurant is broken out into waiting-time, beverage, appetizer, entrée, dessert and check-time. Lastly, parties were randomly set to opt out of some of the meal process.

The hypothesis testing was conducted through 1000 simulation trials of comparing turn-time prediction models. The turn-time prediction models included MLS model as the traditional model and five parametric proportional hazard models as the survival models. The parametric distributions utilized in the PH models were the more widely used distributions (Weibull, Exponential, Gamma, Log-normal, and Log-logistic) for predicting time (Klien & Moeschberger, 2003). For a descriptive list of the distributions properties please see Table 1.

The models were compared through various measures that examined the turn-time predictive properties of the models in the simulations. The measurements include how well the models survival curves fit the actual survival curve including the best match and average difference. Additional measurements included how well the models turn-time predictions matched the actual turn-time. This was done through comparing the models to determine which had the best match for each observation, which model held the best log-likelihood scores, which model had the furthest difference in time prediction and how many exact turn-time predictions (less than 1% error) each model had. Lastly, other measurements included how many negative matches each model had for each simulation run and which model held the closest in predicted overall mean to the actual overall mean.

RESULTS

Simulation Results

With the random values from the probability distributions being set to follow the literature on the typical values for the influential turn-time variables, it was no surprise that the basic descriptive results from the simulations followed a similar path (Kimes, 2004; Kimes et al.,
2002; Kimes & Thompson, 2005). The means of the variables from the simulations are summed up in Table 2. As expected Thursday, Friday, and Saturdays were the busier days, while breakfast, lunch and dinner were the busier times (Kimes, 2004). The time talking was more than the times eating as patrons spend more time chatting when eating in larger parties (Bell & Pliner, 2003). The average turn-time of 65.6 minutes is well within what other studies reported for turn-time (Kimes et al., 2002; Thompson & Kwortnik, 2008).

<table>
<thead>
<tr>
<th>Day of the Week</th>
<th>Means</th>
<th>Party Size</th>
<th>Means</th>
<th>Arrival Time</th>
<th>Means</th>
</tr>
</thead>
<tbody>
<tr>
<td>Monday</td>
<td>38.12</td>
<td>1 Patron</td>
<td>37.71</td>
<td>8:00:00 AM</td>
<td>69.11</td>
</tr>
<tr>
<td>Tuesday</td>
<td>110.61</td>
<td>2 Patrons</td>
<td>354.19</td>
<td>9:00:00 AM</td>
<td>104.10</td>
</tr>
<tr>
<td>Wednesday</td>
<td>170.86</td>
<td>3 Patrons</td>
<td>362.53</td>
<td>10:00:00 AM</td>
<td>57.65</td>
</tr>
<tr>
<td>Thursday</td>
<td>211.16</td>
<td>4 Patrons</td>
<td>164.89</td>
<td>11:00:00 AM</td>
<td>38.60</td>
</tr>
<tr>
<td>Friday</td>
<td>224.05</td>
<td>5 Patrons</td>
<td>56.68</td>
<td>12:00:00 PM</td>
<td>143.07</td>
</tr>
<tr>
<td>Saturday</td>
<td>192.82</td>
<td>6 Patrons</td>
<td>17.54</td>
<td>1:00:00 PM</td>
<td>71.56</td>
</tr>
<tr>
<td>Sunday</td>
<td>52.38</td>
<td>7 Patrons</td>
<td>5.40</td>
<td>2:00:00 PM</td>
<td>38.80</td>
</tr>
</tbody>
</table>

Waiting Times 8 or more Patrons 1.08 3:00:00 PM 30.70
Check-in 4.48 Time 4:00:00 PM 98.79
Beverage 1.92 Eating 18.68 5:00:00 PM 161.03
Appetizer 8.80 Talking 25.48 6:00:00 PM 19.33
Entrée 12.52 Turn-Time 65.80 7:00:00 PM 59.20
Dessert 6.85 Ordered 8:00:00 PM 7.48
Check-out 2.21 Appetizer 43.64
Entrée 947.91
Dessert 60.78

Table 2. Means of Variables

Note: Means are based on 1000 simulation trials

A closer inspection of turn-time was conducted through graphing of the average of the survival curves of turn-time with the lower and upper 95% confidence intervals in Figure 1. While the survival curves indicated that turn-time was fairly normal as the mean of turn-time was within the range of the median, there appeared to be departures from normality. These departures from normality included a shift early in the survival curve as well as the tail ends of the survival curve being fatter than the tail ends of a normal distribution.

Figure 1. Turn-time Survival Curve
Hypothesis Testing

To test the hypothesis, parameter estimates of each of the six models were first conducted for testing the predictive properties in the simulations. The parameters estimates for each of the six models are displayed side-by-side in Table 3. The models mostly held similarity on the variables parameter estimates with differences between their exact strength and subtle differences in a few of the variables. The similarity in parameter estimates was to be expected as three of the survival models were derivations of one another with the exponential distribution being a special case of the Weibull and gamma distributions (Hassett & Stewart, 1999).

The simulations recorded in Table 4 revealed some interesting findings. While the log-likelihood scores and predicted precision indicate that exponential model had the best overall scores with the MLS being in the middle of the pack. In addition, MLS was in the middle of the set with how well its overall average of turn-time matched the actual average of turn-time. Even though MLS did not perform as poorly as the review of literature on survival models would indicate, it was the only model to have negative predictions of turn-time.
Lastly, to see how well each model predicted turn-time for each time interval of the survival curve, the models log-likelihood scores were summed from the simulations for every time-interval on the survival curve. The model with the best summed log-likelihood score at each time interval is highlighted in Table 5. It is interesting to note that all the models with the exception of the Weibull model outperformed the other models at some point over the time period.

To sum the hypothesis results, the exponential and Log-normal proportional hazard models fit the simulated data better than MLS model; while Gamma and Log-logistic PH models were roughly equivalent and the Weibull proportional hazard model fit worse. In addition, MLS model outperformed the PH models in some areas, but was mostly average in its ability to predict turn-time.

Furthermore, the hypothesis stated that MLS would be outperformed by all the proportional hazard models. Given that this was not the case, the hypothesis is rejected. However, it is important to mention that some PH models did outperform MLS model in a majority of the tests.

**DISCUSSION**

While previous studies indicate that MLS would have difficulty predicting at or near the levels of the survival models (Gokovali et al., 2007), it did not lag behind the other tested models in many of the testing categories. Even though MLS did produce some negative predictions, it has particular advantages that may outweigh its negatives. First, multiple least-squares is available in most software packages. Furthermore, with MLS ease of usage and its ability to
predict duration of turn-time that is on par with some of the survival models, it clearly is a beneficial model for practitioners to use.

It should be noted, however, that for practitioners whom are constantly looking for improved and more accurate predictions, survival analysis can be a viable option. The hypothesis testing of this study revealed that a few of the proportional hazard models predicted duration of turn-time better than multiple least-squares. Furthermore, each small degree of improvement in turn-time predictions can considerably increase the efficiency a foodservice establishment, particularly the service and reservation systems.

This study shows that in varying degrees each of the PH models positively stood out in comparison with the traditional MLS method. Such advantage practitioners could utilize when looking to improve the precision of turn-time predictions. Moreover, different survival models may be more effective in various contexts. For example, establishments with dominant proportion of patrons staying longer may utilize the gamma PH model to improve predictions of turn-time. On the other hand, establishments that have a large number of patrons who tend to leave earlier than the expected may look to log-normal PH model to classify the attributes of these patrons to improve waiting times. Operators with a stable meal process may consider utilizing exponential PH model for improving the predictions of turn-time.

Given that a majority of the models stood out in the predictions of turn-time at some point, practitioners looking for optimal predictions may opt for hybrid turn-time prediction models. As each model brings a degree of positives, mixing of these models may also limit the downside of each of the models. In addition, future research may look to stratified survival models by attempting to find the optimal combination of variables utilized for splitting the survival models.

This study comes with some limitations. First, the simulated data used in this study may not fully represent all possible outcomes and their actual effects played out at a dine-in restaurant. Second, previous studies have noted that the type of restaurant greatly affects how long patrons stay (Kivela, 1997). This study assumed a traditional sit down dine-in restaurant for the simulations in determining turn-time. Other types of restaurants’ turn-time values may follow paths that potentially vary when compared to the values in this study.

REFERENCES


