Statistics Final Take Home Exam Executive Summaries

Group 7: Thomas Welles, Hongsen Wang, Xinzhou Xia, Chunxi Wang

12/12/2018

Break Down of Work

Thomas Welles: All Executive Summaries Completed
Problem 1- Analysis of Number Waiting in Queue
Problem 2- Analysis of Number of Customers Lost
Problem 3- All Analysis Completed

Hongsen Wang: Problem 1- Analysis of Time in System
Problem 2- Analysis of Waiting time in Queue

Xinzhou Xia: Problem 1- Analysis of Time Waiting in Queue
Problem 2- Analysis of Number of Customers in Waiting Room

Chunxi Wang: Problem 1- Analysis of Utilization
Problem 2- Analysis of Utilization

Table of Contents of Executive Summaries

Executive Summary of Problem 1

Problem number one entails simulating a response and analysis for an operations manager in charge of airport security/operation. Key categories of airport security performance and efficiency were measured and statistically analyzed in order to obtain a final regression model able to predict airport efficiency given some airport parameters. This summary will highlight a single performance measure, Number Waiting in Queue, as an example of the analysis performed.

Data was collected using the ARENA simulation program. A total of 5 runs were compared with the following parameters

Run	Run No.	Arrival Dist Par	Chk Dist Par	Security Par
Original	1	3	2.5	0.95
Test 2 Gr 7	2	2.7	1.5	0.94
Run 1 Gr 7	3	2	.75	.96
Run 2 Gr 7	4	1.5	2	0.96
Run 3 Gr 7	5	3	.75	.95

For each run a sample of 17 replications were recorded for 1. Time in System, 2. Waiting Time in Queue, 3. Number Waiting in Queue, 4. Resource Utilization. The data set may be found in the supplied data files with the report.

The goal of the analysis was to develop for each response a simple regression model that allowed for proper prediction of the response given a few airport parameters, such as arrival mean and check distribution parameter.

The analysis of The number waiting in Q is presented here as an example. The number waiting in Q was first checked for normality, then an analysis of variance was completed with a Tukey Comparison test to determine which means differed, this was verified with a Kruskal-Wallis non-parametric test, and finally the regression model was developed.

Two regression models were successfully developed to accurately predict the number of individuals waiting in the queue. As an operations manager, arrival time is easily predicted from a flight schedule and is invaluable in both regression models. In order to properly predict the number waiting in queue, the operations manager also needs the utilization and the triangular mode, or if he/she can accurately predict the security check parameter all that is needed is the security parameter and the arrival mean. In order to manage the number of people in queue it is advised to monitor passenger arrival and gauge the attitude and alert levels of the TSA agents. This provides the most predictive model of the queue. The regression models for the number waiting in Q are as follows:

1. No. in Q = 291 - 126.5 Arr. Mean + 207.3 Utilization. This regression represents 68% percent of the data.
2. No. in Q = -9543 - 187.79 Arr. Mean + 10696 Sec. Prob. This regression represents 96% of the data.

Therefore it is recommended to obtain an estimate of arrival mean and of the security parameter (the alert level of the airport) in order to predict the efficiency of the airport operation. With knowledge of these simple parameters the operations manager will be able to properly staff the airport to maintain a utilization, number in Q, and wait times within desired levels. A combination of arrival mean and security parameter were found to be highly correlated to all measures of performance, resulting in simple predictions.

Executive Summary of Problem 2

Problem number two is meant to simulate consultation with a small barbershop business owner. The business owner currently has a one man barbershop with a 5 person, small, waiting room. The barber is currently losing customers and therefore revenue. Consultation is provided to direct the owner of the barbershop how to expand, ie add another barber, or increase waiting room size.

The solution first began by simulating a total of 24 runs with 5 iterations a piece in ARENA with each run utilizing a unique set of 4 factors. The factors are as follows:

Factor	Factor Description	Values
A	Number of Chairs in Waiting Room	5,14,20
B	Number of Barbers Working	1,2
C	Interval Arrival Mean (min)	15.143, 20.143
D	Service Time (min)	10.143, 13.143

For each iteration the responses of Customers Wait Time, Number of Customers Waiting, Customers Lost, and Barber Utilization were recorded for analysis.

Each parameter underwent a regression analysis and analysis of variance. For simplicity the executive summary will only discuss the number of the customers lost. The customers lost is the primary concern of the business owner, since it is a direct measure of lost revenue of the company.

A general linear regression for the number of customers lost was created using a stepwise selection procedure for each of the generation factors A-D shown above. The regression equation is shown below.

Cust Lost = 3.108 + 1.592 Arr Mean_15.143 - 1.592 Arr Mean_20.143 - 1.875 Service Mean_10.143
+ 1.875 Service Mean_13.143 + 5.617 Wait Room Sz_5 - 2.633 Wait Room Sz_14
- 2.983 Wait Room Sz_20 + 2.908 Num. Barbers_1 - 2.908 Num. Barbers_2

The regression indicates a high level of significance for every regressor. However, when the interaction between regressors for the number of customers lost is included it becomes clear that after the waiting room size has increased to 14, the interaction between all other factors is greatly reduced. The number of seats in the waiting room therefore is seen to be the first place of expansion for the barbershop.

Based upon the regression models, correlation, and interaction of factors leading to number of customers lost it is recommended that the waiting room size be increased from 5 to 14 seats. Lost customers lead to lost revenue. If each customer spends an average of $15 and with a current set up of 1 barber and 5 waiting seats, 16 customers are being lost every 5 days. Therefore, per year, $17500 in potential revenue is lost. Decreasing service time is difficult because you do not want to sacrifice the barber experience or quality of services. Hiring another barber would decrease the number the customers lost, but the gained revenue may not cover his/her expenses and salary. Additional space would also be needed to bring in another barber, which could increase cost. However, if instead, that additional space was utilized for a larger waiting room, which may only cost a few thousand a year, the net benefit would be greatly increased. A larger waiting room has the greatest effect for the smallest investment and does not sacrifice customer experience. Increasing the waiting area beyond 14 seats has been shown to have negligible benefit. Therefore the waiting room should be increased from 5 to 15 seats, increasing yearly revenue by 10 to 15 thousand dollars. Complete analysis and data set can be found in the attached minitab file labeled Prob_2_Barbershop_customerslost.

The four best cases for the lowest number of customers lost were then selected to determine equivalence of mean and variance

The cases that were selected were as follows

Case No.	Arrival Mean (min)	Service Mean (min)	Waiting Room Size	Number of Barbers
5	15.143	10.143	14	1
8	20.143	13.143	14	1
14	15.143	13.143	5	2
17	15.143	10.143	14	2

All variances were found to be the same for the best four cases. The mean number of customers lost was also found to be the same by a Tukey Comparison for all of the runs. However, it is important to note that the Confidence Interval (CI) for the number of customers lost for case/run 14 does not cover zero. Therefore, on average case 14 will on average have at least one customer leave. All other cases CI cover zero and therefore no lost customers is possible. Therefore this strengthens the recommendation in the summary of the number of customers lost to focus on increasing waiting room size instead of first hiring another barber.

Executive Summary of Problem 3

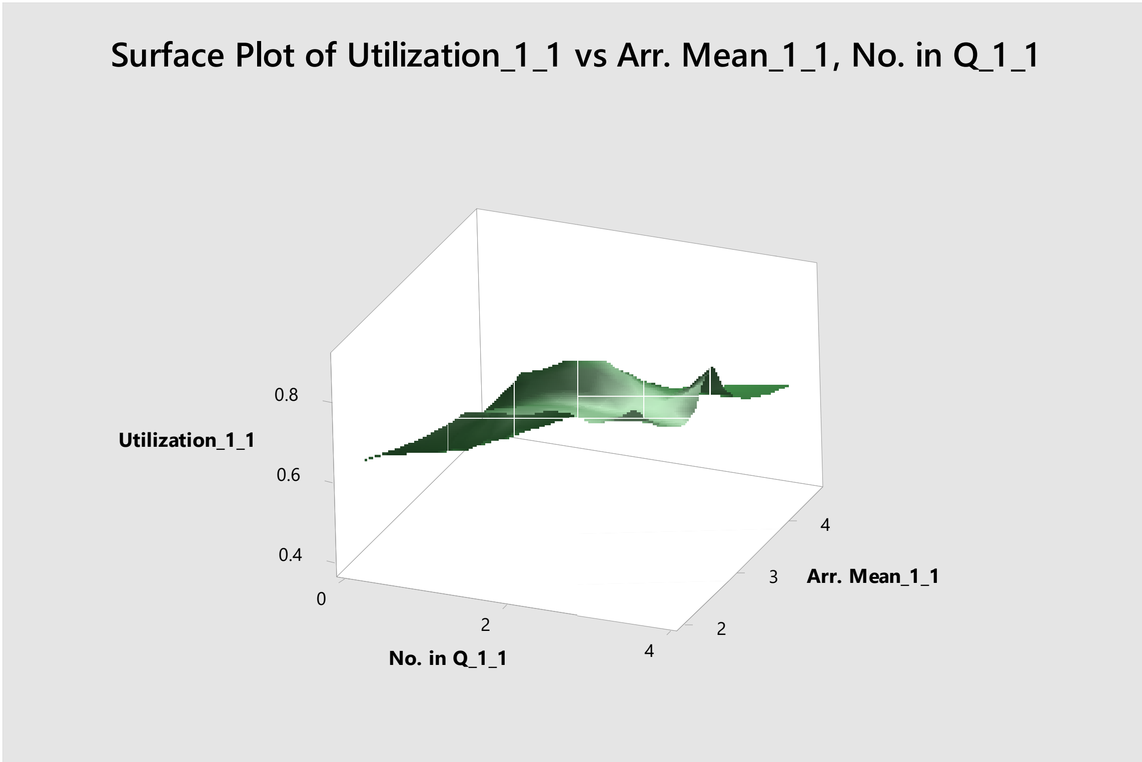
Problem number three of the take home exam was focused on analysis of the data created in take home exam number two. The utilization of resources, the activity of the TSA agents, was analyzed to provide actionable feedback to the operations manager of an airport. The airport must maintain smooth operation and not allow the waiting lines to become too large to overwhelm the staff. Therefore it is crucial for the operations manager to be able to predict the staff utilization. The operations manager must know when to open addition service lanes, bring on additional staff, and when to reduce staff to prevent idle time. This problem is finding the balance between cost of operation and efficiency. If personnel are not active consistently the airport is losing money, as individuals are paid to do nothing, and if the personnel are always overwhelmed with work, the number of people in secure areas may become too large and safety will be placed at risk.

Data was collected using the ARENA simulation program for Airport Security. The data was pooled from Take Home Test 2 into a single 300 data line long excel file. The large data file was used to create a correlation matrix among all generating factors and responses. In this case the generating factors are arrival mean, Check Distribtution parameter, security parameter; and the responses are waiting time, waiting in Q, number in Q, and utilization. This is identical to the first problem.

A smaller, more manageable data was then created by taking the first 12 entries for each data group in the original set. The utilization response was analyzed only, as it is indicative of all other response variables. The utilization response in the small data set was first analyzed with a one-way anova with tukey comparison. The initial ten groups were able to then be separated into three classifications based on level of utilization; high, medium, and low utilization. A discriminant analysis was performed on the three groups. The three groups were easily and corrected sorted back into three groups with significant distance between groups. Details of the discriminant analysis may be seen in the complete report. A forward selection regression model was then performed for all responses and factors generating utilization. Using a significance level of alpha=0.05 for the forward selection of terms a final regression was found as follows:

Utilization_1_1 = 0.8023 - 0.14387 Arr. Mean_1_1 + 0.13216 Triang. Mode_1_1
+ 0.05834 No. in Q_1_1

The arrival mean, triangular mode (or the check distribution), and the number currently in Q were found to be the most significant regressors. The triangular mode was then dropped out of the regression equation to develop a response surface model. Triangular mode was dropped as it is the most difficult of the three parameters to obtain and in minitab a response surface for utilization may only be created for two regressors. The response surface and equation follows:



Utilization_1_1 = 0.7734 - 0.0820 Arr. Mean_1_1 + 0.10593 No. in Q_1_1

The response surface allows for fast visual prediction of the utilization and also a visual depiction of how changes in arrival mean and number in Q will affect the utilization of the staff.

Utilization level is clearly distinguished between high and low by the response surface plot provided immediately preceding this summary. Utilization can be predicted from the arrival mean and the number in the queue. These values are known, easily obtained or easily predicted. The arrival mean can be predicted from the known daily aircraft schedule and is able to be updated automatically as aircraft arrive and depart. Therefore this value is fairly well known ahead of creating a workers schedule and may be updated continuously. The number in queue can be counted and updated. The number in queue is also highly correlated with the alert level of the airport. Therefore prior knowledge of the number in queue is available. Both regressors can be predicted before creating a workers schedule and can be updated continuously with live data. These values can then accurately predict the utilization of airport workers. If the predicted utilization approaches 1, more workers are needed for that particular shift to keep the airport running smoothly. If predicted utilization is low then fewer workers may be assigned to work. This keeps the airport running efficiently at all times.

Complete analysis and data set can be found in the attached minitab file labeled Prob_3_util.