Introduction to Enterprise Analytics

ALY6050, WINTER 2019

MODULE 2 PROJECT ASSIGNMENT

Emergency Facilities Readiness Project

SUBMITTED BY: SHIVANI ADSAR
NUID: 001399374
SUBMITTED TO: PROF. RICHARD HE
DATE: 01/13/2019

Introduction

The assignment gives us an opportunity to learn and perform practical implementations on some of the simulation techniques. These simulations are further used for decision making purposes and draw meaningful insights from the data which would help in understanding the efficiency of emergency facilities.

Analysis

Part I

a) In order to analyse the average number of victims, we have generated 5000 random numbers using the RAND() function. Using the concept of Triangular Probability Distribution, since we have limited information about the sample and population.

We have used the given values, ie. 20, 300 and 80 as minimum, maximum and peak values respectively. Moreover, we have generated 5000 triangular random values, using the following formulas:

A- Min + SQRT((max-min)*(peak-min))*random value
B- Max-SQRT((max-min)*(max-peak)*(1-random value))
C- (Peak-min)/(max-min)
We have used the condition, IF(random value<C,A,B) for generating the X2 values. Using these X2 values, we have plotted the histogram for triangular probability distribution.

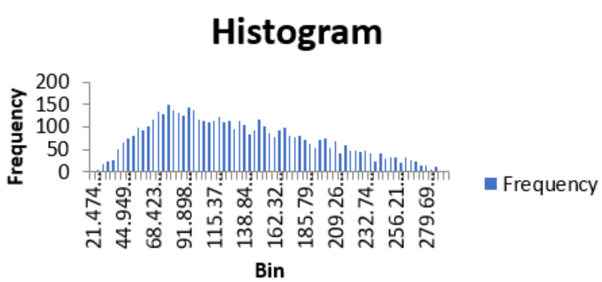


Fig.1 : Histogram for triangular distribution

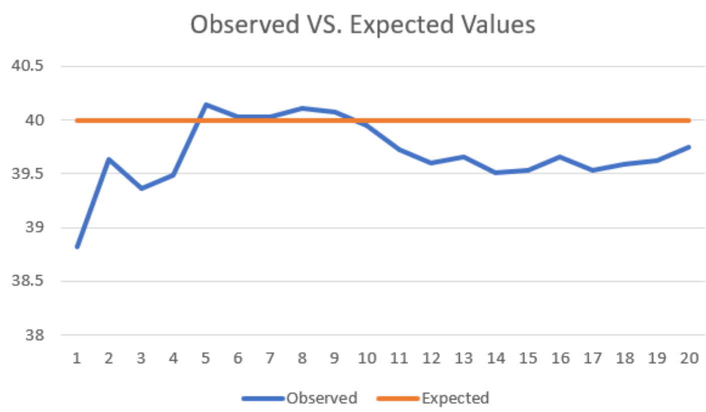
Hospital	Allocation of Disaster Victims
Beth Israel Medical	30%
Tufts Medical	15%
Massachusetts General	20%
Boston Medical	25%
Brigham and Women's	10%

Data Provided

b) Using the percentage of allocation of disaster victims for each hospital, we have used the formula, (frequency * % of allocation of disaster victims). The values generated have been used to calculate the total number of victims. In order to calculate the average time that would take to transfer the victims to the hospitals, we have used the formula,
CONVERT(LN(1-Random value)* Average value* Average in minutes,"min","hr")

This formula is used to convert the time from minutes to hours using the average time in minutes as mentioned. Hence, we have calculated the average of total time required to transport the victims to the hospital.

(c) Using the law of large numbers which states that the average of observed values will approach the theoretical values with the increase in the number of experiments. We have performed the action on Beth Israel Medical, hospital by calculating the Observed and Expected values. The expected values are calculated using the formula, ((min+max+peak)/3)*0.3.



Observation: As can be observed, the expected values are almost closer to the observed values.

Fig.2 : Observed Vs. Expected Values

(d) Using the random values generated earlier, we have calculated the descriptive statistics. The left tail and right tail values are calculated using the bin values generated. The relative frequency has been calculated using the formula, Observed Frequency/ Sum of frequencies. Since, it seems that the data has an exponential probability distribution, we have calculated using the formula, EXPON.DIST(Right Tail, 1/ Mean,1), and the expected frequency has been calculated, using the formula, (Exponential Probability *Sum of observed frequencies). We have used the formula, for calculating the chi-square, (Observed-Expected values)^2/Expected value.

i) The confidence interval was calculated to be 1.66595 for 95 % .

ii) The total time in minutes, is calculated using the formula, (Total Victims/ Total Time) and the exponential probability distribution was plotted as:

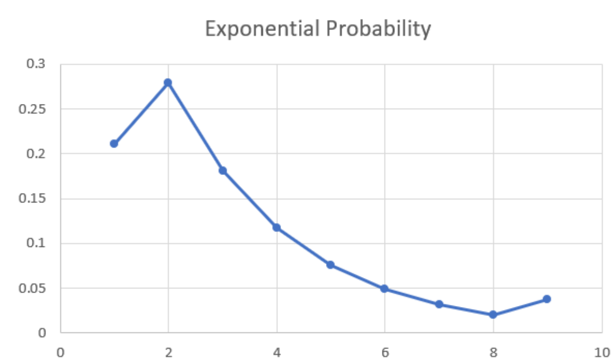


Fig.3 : Exponential Probability Distribution

iii) We have used the Chi-Square Goodness Fit test for performing the hypothesis,

Ho: Our data is exponential ; Ha: Our data is not exponential

Chi Square Goodness of Fit Test	
Tstats	21.07801666
Level of si	0.05
df	6
pvalue	0.001776329

On performing further analysis, it can be seen that, the p-value is lesser than the significance level. Hence, we will reject the null hypothesis, and our data is not exponential.

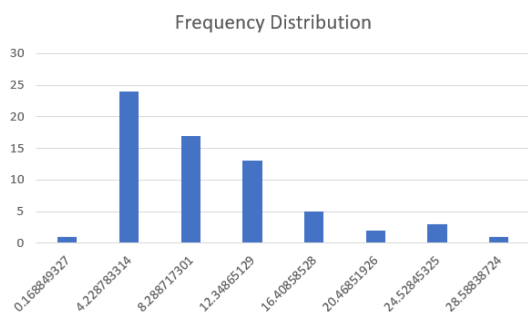


Fig.4 : Frequency Distribution

e) On performing the exploratory analysis for the average transport time required per victim, it could be seen that,

Descriptive Statistics Of Time	
Mean	9.306925
Standard Error	0.982693
Median	6.959259
Mode	#N/A
Standard Deviation	8.2803185
Sample Variance	68.563675
Kurtosis	1.3214324
Skewness	1.4318855
Range	32.479472
Minimum	0.1688493
Maximum	32.648321
Sum	660.79168
Count	71
Largest(1)	32.648321
Smallest(1)	0.1688493
Confidence Level(95.0%	1.9599194

Fig.5 : Descriptive Statistics of Time

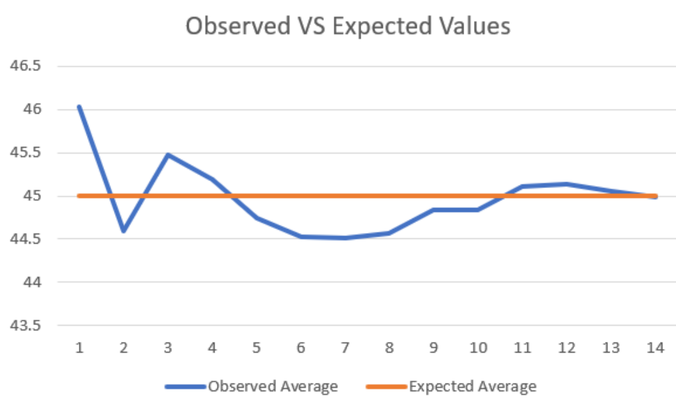
Part 2

(a) In order to analyse the average number of victims, we have generated 5000 random numbers using the RAND() function. We have plotted the X, by using the formula, NORM.INV(Mean, SD, 1), using mean as 150 and Standard Deviation as 50.

(b) This has generated the frequencies and bins, which are further used for calculating the average values. Using the percentage of allocation of disaster victims for each hospital, we have used the formula, (frequency * % of allocation of disaster victims). The values generated have been used to calculate the total number of victims. In order to calculate the average time that would take to transfer the victims to the hospitals, we have used the formula,

CONVERT(LN(1-Random value)* Average value* Average in minutes,"min","hr")This formula is used to convert the time form minutes to hours using the average time in minutes as mentioned.

c) Using the law of large numbers which states that the average of observed values will approach the theoretical values with the increase in the number of experiments. We have performed the action on Beth Israel Medical, hospital by calculating the Observed and Expected values. The expected values are calculated using the formula, ((mean value)*0.3.



Observation: As can be observed, the expected values are almost closer to the observed values.

(d)

i) After calculating the descriptive statistics, it can be seen that the confidence interval is 0.413649 for 95%.

ii) Since, it seems that the data has an normal probability distribution, we have calculated using the formula, EXPON.DIST(Right Tail, 1/ Mean,1), and the expected frequency has been calculated, using the formula, (Exponential Probability *Sum of observed frequencies). We have used the formula, for calculating the chi-square, (Observed-Expected values)^2/Expected value.

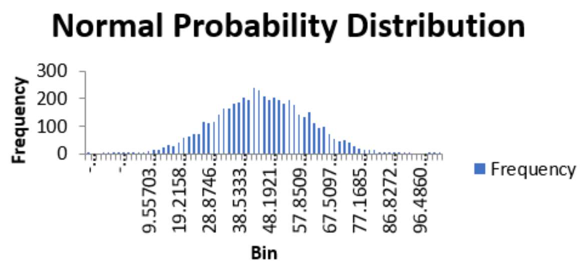


Fig.: Normal Probability Distribution

iii) We have used the Chi-Square Goodness Fit test for performing the hypothesis,

Ho: Our data is normally distributed ; Ha: Our data is not normally distributed

Chi-Square Goodness Fit Test	
test-statistics	16964.96803
level of significa	0.05
df	7.00
p-value	0.00

On performing further analysis, it can be seen that, the p-value is lesser than the significance level. Hence, we will reject the null hypothesis, and our data is not normally distributed.

e) On calculating the descriptive statistics for the time taken for the victims, the following was observed:

Descriptive Statistics	
Mean	0.014193404
Standard Error	0.001632457
Median	0.007462687
Mode	0.002877729
Standard Deviation	0.01375533
Sample Variance	0.000189209
Kurtosis	-0.020260071
Skewness	1.132056131
Range	0.045001605
Minimum	0
Maximum	0.045001605
Sum	1.007731656
Count	71
Largest(1)	0.045001605
Smallest(1)	0
Confidence Level(95	0.003255833

We have anaysed the descriptive statistics on the basis of mean, Standard deviation, count and confidence intervals.

Conclusion

1. It can be observed that the simulations had exponential and normal probability distributions. The descriptive analysis was calculated and compared between the two simulations. The mean, standard deviation and confidence intervals were used as important parameters to determine the efficiency of victims being commuted to the hospitals on emergency. In exponential probability distribution, the events take place continuously and independently, having constant time whereas normal distribution has a symmetric distribution with the normal curve equals 1.
2. The information from simulations can be used for planning crisis management where the victims can be relieved at the earliest during emergency situations to nearby towns. We would perform, gamma, normal and exponential simulations which would give us the descriptive statistics on planning in emergent conditions. Moreover, This would help in analyzing the chi-square goodness test for finding out the appropriate method with maximum effectiveness.
3. The simulations can be changed, by varying the random numbers for every average calculations done in the data. This would help in giving the effectiveness for random set of numbers and would generate un-biased results.

Reference

1. Illowsky, B. (n.d.). Introduction to Statistics. Retrieved from https://courses.lumenlearning.com/introstats1/chapter/the-exponential-distribution/.
2. A. K. (2019, December 22). Gamma Function - Intuition, Derivation, and Examples. Retrieved from https://towardsdatascience.com/gamma-function-intuition-derivation-and-examples-5e5f72517dee.