Yuyi Zhang

May 31, 2020

Spring 2020 CPS Quarter Term B

Instructor: Danielle Mauro

Project 2:   
Emergency Facilities Readiness Project

ALY 6050\_Introduction to Enterprise Analytics

# **Introduction**

In project 2, University administers are making emergency evacuation planning for disaster victims to 5 different local hospitals. In this report, we are going to go through project 2 with two different distributed simulations. In each distribution, we are going to conduct an exploratory data analysis to see what the effective elements for time transport time would be.

# **Analysis**

## **Triangular Probability Distribution**

Perform a simulation analysis consisting of 5,000 simulations to determine:

### Average number of victims that can be expected at each hospital.

Because the total number of victims are triangular distributed, so the mean value of total victims equals to (min + max + mode)/3.

A screenshot of a cell phone

Description automatically generated

*Figure 1*. Average Number of Victims at each Hospitals

### For each hospital, the average total time (in hours) needed to transport all victims.

Since the total number of victims are known, the total time in hours of transport for each hospital would be calculated as below.

A picture containing knife

Description automatically generated

*Figure 2*. Total Time of Transport for each Hospitals

### For part (a) above, create a chart the displays the Law of Large Numbers in action for the Beth Israel Medical. (Law of large numbers: As the number of trials becomes larger, the observed averages approach to the theoretical average.)

In this part, 5000 simulations of victims are generated by rtriangle() function. 20% of them are being transported to Beth Israel Medical. Then plot the 5000 simulations as below,

A screenshot of a cell phone

Description automatically generated

*Figure 3*. Law of Large Numbers chart

The average number was calculated in part (a) as mu.bim = 26.667. In this chart, the end number is 26.773, which is very close to theoretical number.

### For the Beth Israel Medical hospital, perform an exploratory data analysis of the total transport time by:

#### Calculating a 95% confidence interval for the total transport time.

According to properties of the sampling distribution, the equation of mean and standard deviation is:

According to the Central Limit Theorem, the interval estimator would be calculated below when the confidence interval is 95%:

Substituting and into the interval estimator formula, we obtain: (6.609, 6.778).

#### Determining a probability distribution that best fits the total transport time (in hours).

A screenshot of a cell phone

Description automatically generated

*Figure 4*. Histogram of Total Transport Time

Based on the histogram above, we assume that the total transport time is normally distributed.

#### Supporting your assertion in part (ii) by creating a frequency distribution and performing a Chi-squared Goodness of fit test.

To test whether the variables fit on Normal distribution with the level of significance: α = 0.05. State null and alternative hypotheses as below:

Null hypothesis H0: the total transport time follows the Normal distribution

Alternative hypothesis H1: the total transport time does not follow the Normal distribution

A screenshot of a cell phone

Description automatically generated

*Figure 5*. Chi-squared test result

Compare p-value and significance value. In this case, the p-value (1) is larger than the significance level (α = 0.05). So there isn’t sufficient evidence to reject H0. In another word, the total transport time fits in the Normal distribution with mean value of 6.693 and standard deviation of 3.051.

### Let 𝒕 denote the average transport time (in minutes) per victim for the entire process of transporting all victims. Perform an exploratory data analysis of 𝒕.

The average transport time would be calculated as below,

𝒕 = 0.2\*15 + 0.15\*10 + 0.3\*7 + 0.25\*15 + 0.1\*20 = 12.35

Therefore, 𝒕 is a constant value of 12.35.

## **Normal Probability Distribution**

Perform a simulation analysis consisting of 5,000 simulations to determine:

### Average number of victims that can be expected at each hospital.

Because the total number of victims are normally distributed with mean and standard deviation known, so the average number of victims are calculated as below:

A screenshot of a cell phone

Description automatically generated

*Figure 6*. Average Number of Victims at each Hospitals

### For each hospital, the average total time (in hours) needed to transport all victims.

Since the total number of victims are known, the total time in hours of transport for each hospital would be calculated as below.

A picture containing knife

Description automatically generated

*Figure 7*. Total Time of Transport for each Hospitals

### For part (a) above, create a chart the displays the Law of Large Numbers in action for the Beth Israel Medical. (Law of large numbers: As the number of trials becomes larger, the observed averages approach to the theoretical average.)

In this part, 5000 simulations of victims are generated by rnorm() function. 20% of them are being transported to Beth Israel Medical. Then plot the 5000 simulations as below,

A screenshot of a cell phone

Description automatically generated

*Figure 8*. Law of Large Numbers chart

The average number was calculated in part (a) as mu.bim = 30. In this chart, the end number is 30.228, which is very close to theoretical number.

### For the Beth Israel Medical hospital, perform an exploratory data analysis of the total transport time by:

#### Calculating a 95% confidence interval for the total transport time.

Substituting and into the interval estimator formula,

we obtain: (7.444, 7.605).

#### Determining a probability distribution that best fits the total transport time (in hours).

A screenshot of a cell phone

Description automatically generated

*Figure 9*. Histogram of Total Transport Time

Based on the histogram above, we assume that the total transport time is normally distributed.

#### Supporting your assertion in part (ii) by creating a frequency distribution and performing a Chi-squared Goodness of fit test.

To test whether the variables fit on Normal distribution with the level of significance: α = 0.05. State null and alternative hypotheses as below:

Null hypothesis H0: the total transport time follows the Normal distribution

Alternative hypothesis H1: the total transport time does not follow the Normal distribution

A picture containing bird

Description automatically generated

*Figure 10*. Chi-squared test result

Compare p-value and significance value. In this case, the p-value (1) is larger than the significance level (α = 0.05). So there isn’t sufficient evidence to reject H0. In another word, the total transport time fits in the Normal distribution with mean value of 7.525 and standard deviation of 2.905.

### Let 𝒕 denote the average transport time (in minutes) per victim for the entire process of transporting all victims. Perform an exploratory data analysis of 𝒕.

The total transport time of each hospital would be simulated by rnorm() function with mean and sd provided by instructions. So 𝒕 is calculated as below:

20% \* rnorm.bim + 15% \* rnorm.tm + 30% \* rnorm.mg + 25% \* rnorm.bm + 10% \* rnorm.bw

Then perform an exploratory data analysis of the total transport time by:

#### Calculating a 95% confidence interval for the total transport time.

Substituting and into the interval estimator formula,

we obtain: (12.297, 12.458).

#### Determining a probability distribution that best fits the average transport time (in hours).

A screenshot of a cell phone

Description automatically generated

*Figure 11*. Histogram of Average Transport Time

Based on the histogram above, we assume that the average transport time is normally distributed.

#### Supporting your assertion in part (ii) by creating a frequency distribution and performing a Chi-squared Goodness of fit test.

To test whether the variables fit on Normal distribution with the level of significance: α = 0.05. State null and alternative hypotheses as below:

Null hypothesis H0: the average transport time follows the Normal distribution

Alternative hypothesis H1: the average transport time does not follow the Normal distribution

A picture containing bird

Description automatically generated

*Figure 12*. Chi-squared test result

Compare p-value and significance value. In this case, the p-value (2.2e-16) is smaller than the significance level (α = 0.05). So there is sufficient evidence to reject H0. In another word, the average transport time doesn’t fit in the Normal distribution with mean value of 12.378 and standard deviation of 1.673.

# **Conclusions**

## **Differences between (1) and (2)**

|  |  |  |
| --- | --- | --- |
|  | Triangular Distribution A picture containing table, boat, man, people  Description automatically generated  =133.333  =59.757 | Normal Distribution A screenshot of a cell phone  Description automatically generated  =150  =30 |
| a | 26.667 | 30 |
| b | 6.667 | 7.5 |
| c | 26.773 | 30.228 |
| d.i | (6.609, 6.778) | (7.444, 7.605) |
| d.iii | A screenshot of a cell phone  Description automatically generated =6.693  =3.051 | A screenshot of a cell phone  Description automatically generated =7.525  =2.905 |
| 𝒕 | A close up of a mans face  Description automatically generated 𝒕 = 12.35 | A screenshot of a cell phone  Description automatically generated (12,297, 12,458)  =12.378  =1.673 |

According to the comparison listed above, the triangular distribution provides a lower mean value with larger standard deviation. Because the transport time in (1) is fixed, so the 𝒕 is a constant value. In (2), the 𝒕 is normally distributed.

## **Planning opinions**

In the end, 𝒕 value in (1) 12.35, is very close to the mean value (12.387) of 𝒕 in (2) with a very small standard deviation of 1.673. So, no matter how do we distribute the total victims in the beginning, the average transport time would approaching to a similar value, 12.36. In this simulation, we could know which elements that would affect the 𝒕 value, and which are not.

## **Providing additional information**

In planning, reducing the average total transport time would be our goal. As we mentioned earlier, no matter what distribution we used to simulate victims, the transport time would approach the same value in the end. Having minor standard deviation in transport time for each hospital wouldn’t change the 𝒕 value too. Therefore, decreasing each victim’s transport time would be a good place to start, for example, provide more ambulances would approach this goal.