Statistical Approach - Case Study & Assignment 1

2022-09-15

## Uniform Distribution

**Case Study: City of Medville Uses Statistical Approach to Estimate Costs for Its Pilot Project To encourage sports and fitness among students from kindergarten to 12th grade, the education department of the city of Medville, Pennsylvania, conceived a 12-month pilot project to provide special free training, nutrition, and sports gear to the students of a select 10 schools. The goal of this project was to cover 70% of the student population under the new program. The initial challenge was to figure out the funds required to run this project and also the plan to carry out the project work. For scope management, the project management committee divided the student population in different age groups and estimated the cost for students in each age group.**

**The table below depicts the various student age groups and the cost estimates.**

|  |  |
| --- | --- |
| **Table 1: Table 1 Estimated Project Cost for Various Student Age Groups** | |
| **Student Age Group** | **Estimated Cost Per Student** |
| **Less than 10 years old** | **$2,000** |
| **10 to 15 years old** | **$5,000** |
| **More than 15 years old** | **$3,000** |

**The project assumed that the total population of students (2,000 students) was normally distributed with a mean age of 12 and a standard deviation of 3. The following statistical calculations for normal distribution were used to make decisions.**

### Determine Target Age Group for Initial Project Pilot

For normal distribution,

* 1 − σ covers roughly 68% of the population, which implies 68% of the total 2,000 students fall in the age group 9 to 15 (12 +/– 3).
* 2 − σ covers roughly 95% of the population, which implies 95% of the total 2,000 students fall in the age group 6 to 18 (12 +/– 6).

Because the goal of the pilot project was to cover 70% of the student population, students in age group 6 to 18 were selected for the initial pilot.

### Estimate Project Costs for the Target Age Group

The target age group contained the student population from all three population bands listed in Table 1 above. Thus, cost estimates pertaining to those population bands or age groups had to be considered for calculating costs for the target age group (6 to 18 years old). The project could be figured it out using the R function norm or Excel NORMDIST function as follows:

We will be using norm function in R to obtain some of the probability values: \* pnorm(X, µ, σ)

X here is a random variable representing either the lower or upper end of the age-group range), µ = mean age in the age group, σ = standard deviation.

**The percentage of target students belonging to age group 6 to 10 years old will be**

k= pnorm(10,12,3)-pnorm(6,12,3)

k

## [1] 0.2297424

**Cost Allocation for 6- to 10 year old students**

k1=(2000 \* k \* 2000)  
k1

## [1] 918969.6

Percentage of target students belonging to age group 10 to 15 years

m=pnorm(15,12,3)-pnorm(10,12,3)  
m

## [1] 0.5888522

Cost Allocation for 10 to 15 year old students

m1=(5000 \* m \* 2000)  
m1

## [1] 5888522

Percentage of target students belonging to age group 10 to 15 years

o=1-k-m  
o

## [1] 0.1814054

Cost Allocation for 10 to 15 year old students

o1=(3000 \* o \* 2000)  
o1

## [1] 1088432

Now, the total estimated cost for all target students for the initial pilot project will be the percentage of target students belonging to the age group 6 to 18 years

k1+o1+m1

## [1] 7895924

## Poisson Distribution

A cannabis manufacturing firm with limited financial resources wants to establish its packing unit and determine if they need to acquire an additional machine that is averaging a downtime of 45 min per shift (each shit runs for 8 hours) for the past year. As per their previous year’s production rate they cannot have machines idle for 2 hours or more a day to meet their target production. As a project manager, what would you suggest to the firm?

Probability for having 2 or more

ppois(2, lambda=1.75, lower =FALSE)

## [1] 0.2560303

1- ppois(2, lambda=1.75)

## [1] 0.2560303

### Assignment 1:

**Note-Answer Questions 1-3 based on solution of the Case Study – City of Medville**

1. **What approach was used by the city of Meadville to estimate the overall project cost?**
2. Here to estimate the cost for the overall project the statistical in which theoretical approach is used.
3. **Define the scope of this project.**

In this project, the scope is to calculate the cost of the student in the different age groups.

1. **Do you think the city made a wise decision to use this approach for cost estimation? Why do you think so?**

As we know from the case study, they must use the statistical approach to make their best decision which indeed makes sense. Because in the above we must use the normal distribution which is used to estimate the cost of the age groups.

By using the statistical approach in this project, it provides an accurate outcome, can reduce the bias, can predict a better cost prediction, etc. these are the reasons which help to project manager to take a wise decision.

1. **What is the empirical rule in standard normal distribution?**

The empirical rule in the standard normal distribution is also called as three sigma rule. This means that all the data lies between the three standard deviations from the mean.

Here Standard deviation is denoted as **σ,** and the mean is denoted as **µ.**

* This rule predicts the data about 68% which data lies between the one standard deviation (**µ + σ).**
* It predicts the data about 95% of which data lies between the two-standard deviation (**µ +2 σ)**.
* It predicts the data about 99.7% of which data lies between the three standard deviations (**µ +3 σ)**.

1. **The mean duration of the activities of a project is 10 days with a standard deviation of 2 days. Using the empirical rule estimate the percentage of project activities with a duration between 7 and 10 days.**

Mean – 10

Sd – 2

m=pnorm(10,10,2)-pnorm(7,10,2)  
m

## [1] 0.4331928

So around 43.33% percentage of project activities with a duration between 7 and 10 day

1. **Calculate the different empirical rules for a data with mean project completion of 15 days and a standard deviation of 3 days, assuming the observed data follows a normal distribution.**

Mean = 15

Sd = 3

99.7 % - Mean +- 3Sd

15 +- 3(3)

15 +- 9

It means the data lies between the 6 days and 24 days where 6 is the lower bound and 24 is the upper bound.

95 % - Mean +- 2Sd

15 +- 2(3)

15 +- 6

It means the data lies between the 9 days and 21 days where 9 is the lower bound and 21 is the upper bound.

68 % - Mean +- 3Sd

15 +- (3)

15 +- 3

It means the data lies between the 12 days and 18 days where 12 is the lower bound and 18 is the upper bound.

1. **Explain the characteristics of Poisson distribution**

* Poisson distribution is of discrete distribution nature.
* In Poisson distribution, each event is independent of other events and throughout the experiment, the mean of the events remains constant and discrete events occur over intervals of time.
* Range of each even at intervals is from 0 to infinity.

1. **What is the primary purpose of using the Lean Six Sigma approach in project management?**
2. The main objective of using the lean six sigma approach in project management are as follows:

* Decrease the error(defects) in the project.
* Improving the quality of the project such as having optimized results with lesser defects.
* Improving the productivity of the project
* Improving consistency in the project
* Keep in mind the client’s needs and satisfaction and result

1. **What are the advantages of using analytics in project management?**
2. The merits of using analytics in project management are:

* **Pattern findings**

With the help of analytics, the project managers collect the data and highlight the work that was completed, and it also helps to find the patterns from the analysis.

* **Easy Availability**

In today’s world, there is a technology widely available with tools like analytical tools which help the project manager to use their knowledge and understanding of the data in the visual analysis to predict future outcomes and its behavior.

* **Data-Driven Decision**

Due to having open-sourced data analytical tools, the predictive data is being used by the project manager to provide a good decision that has a good impact on the project.

* **Improving the success rate of the project**

When the project is in the performing stage based on the analysis, they can able to generate an understanding to check whether the project is working better and based on this analysis what strategy and decision to make to maintain the success rate of the project or even if the project not performing well it helps to improve its efficiency rate.

1. **What does PDSA stand for?**
2. PDSA stands for Plan, Do, Study, Act.

Plan: Initiation of the plan

Do: Plan execution and its documentation

Study: Based on the execution plan analysis is to be done.

Act: Based on the result of analysis next steps can be executed.