# Formative Assessmber #7

BORROMEO, S, MAYO, Z, MERCADO, C, SINOCRUZ, A, & TAGAYTAY, G

2025-04-10

*Github Link:* https://github.com/sophiaborromeo/Probability-FA7/blob/main/fa7.md

## Campus-Related Issues Using Exponential Distribution

### 1. Identify a Practical Campus Problem

For this assessment, the group identified the time between students arriving to queue for an elevator in the Education Building inside the FEU Manila campus—a practical campus problem where events occur randomly over time.

### 2. Collect Data

The data collection started at 1:40 PM and ended at 2:10 PM.

The table below shows the time intervals between the arrival of each person. The time is standardized.

Table 1: Time Between Student Arrivals at the Elevator Queue (in Seconds)

| Person | Arrival_Time_Seconds |
|---|---|
| 1 | 62.00 |
| 2 | 45.00 |
| 3 | 22.78 |
| 4 | 2.00 |
| 5 | 104.20 |
| 6 | 43.00 |
| 7 | 104.20 |
| 8 | 115.34 |
| 9 | 91.63 |
| 10 | 29.25 |
| 11 | 36.01 |
| 12 | 7.36 |
| 13 | 13.82 |
| 14 | 61.91 |
| 15 | 11.75 |
| 16 | 2.81 |
| 17 | 15.46 |
| 18 | 75.10 |
| 19 | 85.47 |

| Person | Arrival_Time_Seconds |
|---|---|
| 20 | 10.99 |
| 21 | 13.60 |
| 22 | 131.94 |
| 23 | 41.26 |
| 24 | 28.84 |
| 25 | 46.69 |
| 26 | 21.80 |
| 27 | 53.86 |
| 28 | 202.49 |
| 29 | 120.00 |
| 30 | 40.91 |

## 3. Verify if Exponential Distribution is Applicable

**Instruction 1: Check if events occur randomly and independently over time.**

Because the group recorded interarrival times, this implies independence—since each record is a time interval not influenced by the last one. Since students come from random locations with different purposes in queueing, this also indicates random occurrence. It's safe to assume the events are independent and random over time. Therefore, we can apply the exponential distribution.

**Instruction 2: Identify the average rate of event occurrence per unit of time.**

The formula for lambda is:

$$\text{Lambda} = \frac{n}{\sum(\text{arrival times})}$$

Where $n$ is the number of events, and $\sum$(waiting times) is the sum of all the arrival times. Through manual calculation, we get:

$$\text{Lambda} = \frac{30}{1,641.47}$$

$$\text{Lambda} = 0.0182763$$

For number 4, we will be calculating lambda using a statistical tool.

## 4. Compute Key Parameters

### Mean (Expected Value)

For the exponential distribution, the mean (expected value $\mu$) is the reciprocal of the rate $\lambda$:

$$\lambda = \frac{\text{Number of Observations}}{\text{Total Arrival Time}}$$

$$\mu = \frac{1}{\lambda}$$

Mean arrival time: **54.716 seconds**.

### Probability Density Function (PDF)

$$f(x) = \lambda e^{-\lambda x}, \quad \text{for } x \geq 0$$

Table 2: PDF Values for Each Data Point

| Time_Seconds | PDF |
|---|---|
| 62.00 | 0.00589 |
| 45.00 | 0.00803 |
| 22.78 | 0.01205 |
| 2.00 | 0.01762 |
| 104.20 | 0.00272 |
| 43.00 | 0.00833 |
| 104.20 | 0.00272 |
| 115.34 | 0.00222 |
| 91.63 | 0.00342 |
| 29.25 | 0.01071 |
| 36.01 | 0.00946 |
| 7.36 | 0.01598 |
| 13.82 | 0.01420 |
| 61.91 | 0.00590 |
| 11.75 | 0.01474 |
| 2.81 | 0.01736 |
| 15.46 | 0.01378 |
| 75.10 | 0.00463 |
| 85.47 | 0.00383 |
| 10.99 | 0.01495 |
| 13.60 | 0.01425 |
| 131.94 | 0.00164 |
| 41.26 | 0.00860 |
| 28.84 | 0.01079 |
| 46.69 | 0.00779 |
| 21.80 | 0.01227 |
| 53.86 | 0.00683 |
| 202.49 | 0.00045 |
| 120.00 | 0.00204 |
| 40.91 | 0.00865 |

### Cumulative Distribution Function (CDF)

$$F(x) = 1 - e^{-\lambda x}, \quad \text{for } x \geq 0$$

Table 3: CDF Values for Each Data Point

| Time_Seconds | CDF |
|---|---|
| 62.00 | 0.67798 |
| 45.00 | 0.56064 |
| 22.78 | 0.34054 |
| 2.00 | 0.03589 |
| 104.20 | 0.85109 |
| 43.00 | 0.54428 |

| Time_Seconds | CDF |
|---|---|
| 104.20 | 0.85109 |
| 115.34 | 0.87852 |
| 91.63 | 0.81263 |
| 29.25 | 0.41409 |
| 36.01 | 0.48218 |
| 7.36 | 0.12586 |
| 13.82 | 0.22320 |
| 61.91 | 0.67745 |
| 11.75 | 0.19325 |
| 2.81 | 0.05006 |
| 15.46 | 0.24614 |
| 75.10 | 0.74654 |
| 85.47 | 0.79030 |
| 10.99 | 0.18197 |
| 13.60 | 0.22008 |
| 131.94 | 0.91031 |
| 41.26 | 0.52956 |
| 28.84 | 0.40968 |
| 46.69 | 0.57400 |
| 21.80 | 0.32862 |
| 53.86 | 0.62632 |
| 202.49 | 0.97530 |
| 120.00 | 0.88844 |
| 40.91 | 0.52654 |

## 5. Interpret the Results

**Determine the likelihood of an event happening within a specific timeframe.**

We can use the exponential Cumulative Distribution Function (CDF) to find the probabilities,

**a. Probability of a student arriving within 30 seconds:**

$$F(x) = 1 - e^{-\lambda x}$$

```
lambda <- 30 / sum(time_secs)
time_30 <- 30
prob_within_30 <- 1 - exp(-lambda * time_30)
```

Probability a student arrives within 30 seconds: **42.21%**

**b. Probability of student arriving more than 60 seconds:**

$$P(X > x) = e^{-\lambda x}$$

```
time_60 <- 60
prob_more_than_60 <- exp(-lambda * time_60)
```

Probability of student arriving more than 60 seconds: **33.4%**

**Implications**

**a. Expected Arrival Time:** Based on the data, the mean waiting time between student arrivals is approximately:

```
mean_wait <- 1 / lambda
```

On average, a student arrives every **54.72 seconds** .

**b. Real-World Implications on Campus**

*Factor Affecting Queuing*

- It was observed that the intervals between arrivals of the FEU community, especially students arriving at the elevator were shorter at around 1:30 PM to 1:45 PM since it was the usual starting time of classes (Peak Periods). This means that the queue was longer. As time passed by, specifically after 1:45 PM (Off-Peak Periods), the intervals between student arrivals started to increase. This is because, as more professors and students proceeded to their respective classes, fewer students or professors required the use of the elevator, causing a larger interval between arrivals. As a result, the queue got thinner, and waiting times between students grew, as can be inferred from the data obtained.

*Elevator Usage Efficiency:*

- There is only a 42% chance that a student or professor arrives within 30 seconds and as the group observed there is a chance that elevators often leave with only one or two students resulting in a wastage of energy and time.

*Queue Management:*

- With a mean student arrival time of *54.7 seconds*, a line can form during peak hours. This implies that during rush hours, such as at the beginning of classes, the intervals between student arrivals can be shorter, resulting in a more congested queue and possibly faster boarding times. Conversely, at off-peak times, when students are arriving in smaller numbers, waiting times between arrivals can become longer. The average waiting time of 54.7 seconds serves as a valuable benchmark for estimating how long a group of students will have to wait to enter the elevator. By knowing these trends, the school can better predict peak periods and schedule the elevator system appropriately, resulting in smoother operations and reduced wait times during peak periods.

*Facility Optimization:*

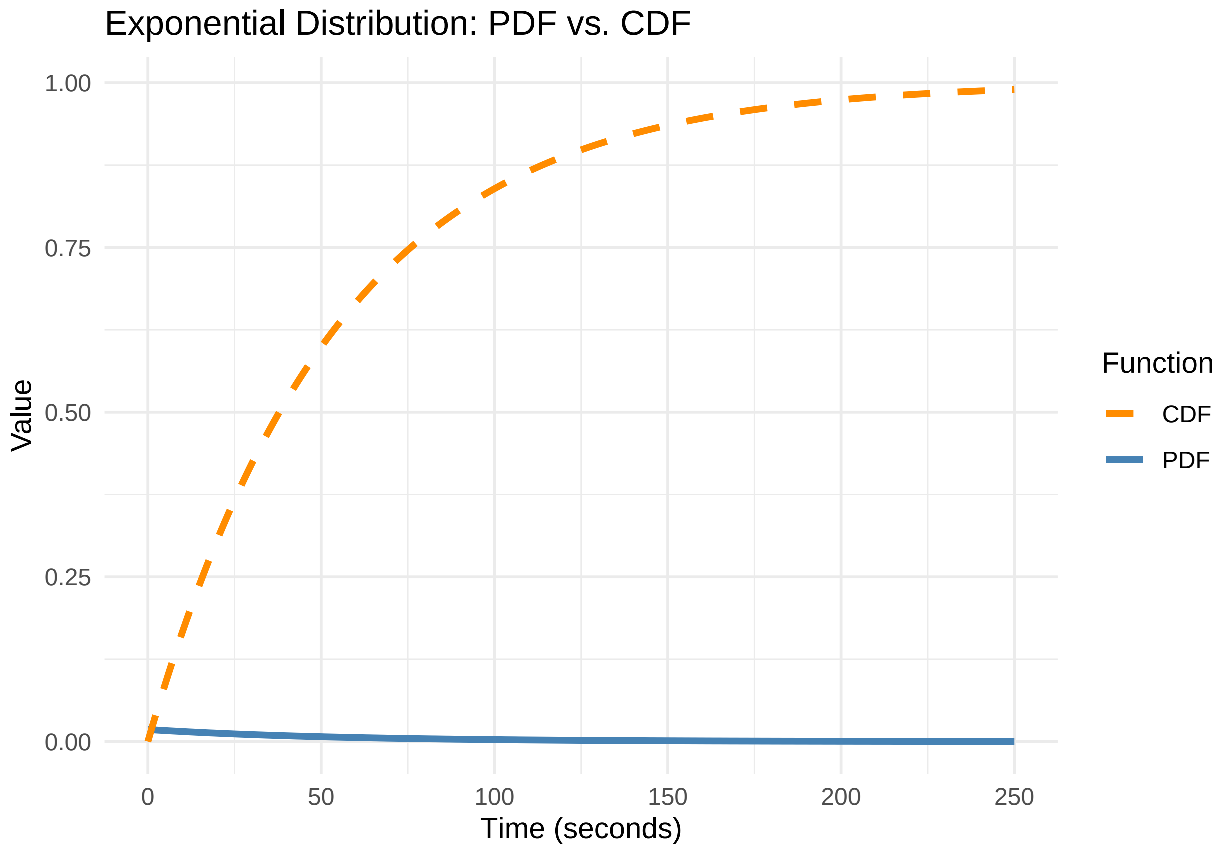
- Elevator scheduling or idle-time automation could be improved using real-time arrival predictions.
- Digital displays showing estimated wait times could enhance the student experience.

*Energy Efficiency:*

- Since there is a 33% chance of waiting more than 60 seconds, the elevator could delay activation unless more people arrive, optimizing both energy and time.

## Visualization:

```r
library(ggplot2)

# Define lambda
lambda <- 30 / sum(time_secs)

# Define a sequence of time values (x)
x_vals <- seq(0, 250, by = 1)

# Compute PDF and CDF for each x
pdf_vals <- lambda * exp(-lambda * x_vals)
cdf_vals <- 1 - exp(-lambda * x_vals)

# Combine into a data frame
exp_data <- data.frame(
  Time = x_vals,
  PDF = pdf_vals,
  CDF = cdf_vals
)

# Plot
ggplot(exp_data, aes(x = Time)) +
  geom_line(aes(y = PDF, color = "PDF"), linewidth = 1.2) +
  geom_line(aes(y = CDF, color = "CDF"), linewidth = 1.2, linetype = "dashed") +
  labs(title = "Exponential Distribution: PDF vs. CDF",
       x = "Time (seconds)",
       y = "Value",
       color = "Function") +
  theme_minimal() +
  scale_color_manual(values = c("PDF" = "steelblue", "CDF" = "darkorange"))
```




**Interpretation**

The graph shows the relationship between the Cumulative Distribution Function (CDF) and the Probability Density Function (PDF) of an exponential distribution. The CDF, in the orange dashed line, rises slowly and approaches 1 asymptotically, showing that as time goes by, the likelihood of the event increases. Conversely, the PDF, shown as the blue solid line, starts with a larger value and then decays exponentially, consistent with the fact that the chance of occurrence of the event decreases as time increases. This is an example or characteristic of exponential distributions, in which events have higher chances of occurrence earlier than later.