HABIB UNIVERSITY	**Activity 07** **Dr. Shafayat Abrar** **CE/MATH 362/322-L1**	**Course: Statistics and Inferencing** **Due Date: Sept. 28, 2023** **Due Time: 11:00 PM**

Reading Material:

The sum of two **exponential random variables** often appears in practical scenarios involving waiting times, system failures, and service processes. Here are a few examples:

- **Queueing Systems (e.g., M/M/1 Queue)**

  Scenario: In a system where two tasks need to be completed sequentially, both with exponentially distributed service times (e.g., waiting for service at a bank and then processing a transaction).

  Example: Suppose the time to be served by a teller follows an exponential distribution $X_1 \sim \text{Exp}(\lambda_1)$, and the time to process your transaction follows another exponential distribution $X_2 \sim \text{Exp}(\lambda_2)$. The total time spent at the bank is the sum $X_1 + X_2$, which is the sum of two independent exponential random variables.

- **Reliability of Systems with Two Components in Series**

  Scenario: Two components in a series system must both work for the system to function and each component has an exponentially distributed time to failure.

  Example: If the lifetime of the first component is $X_1 \sim \text{Exp}(\lambda_1)$ and the second component is $X_2 \sim \text{Exp}(\lambda_2)$, the system will fail when the first or second component fails. The total time until the system fails is the sum $X_1 + X_2$.

- **Telecommunications: Time to Transfer Packets**

  Scenario: In packet-switched networks, the transmission times of packets are often modeled as exponentially distributed due to the random nature of packet arrivals and transfers.

  Example: If two routers each have exponentially distributed transmission times $X_1 \sim \text{Exp}(\lambda_1)$ and $X_2 \sim \text{Exp}(\lambda_2)$, the total time to transfer a packet through both routers is $X_1 + X_2$, the sum of two exponential random variables.

- **Poisson Process: Time Between Events**

  Scenario: In a Poisson process, the time between events (such as phone calls arriving at a call center) is exponentially distributed. If you're interested in the total time between two events, this is the sum of two exponential variables.

  Example: If the time until the first event follows $X_1 \sim \text{Exp}(\lambda)$ and the time until the second event follows $X_2 \sim \text{Exp}(\lambda)$, the total time between the first and second events is $X_1 + X_2$, which has a Gamma distribution with shape parameter 2 and rate λ.

- **Maintenance Systems: Time Until a Certain Number of Failures**

  Scenario: In maintenance systems, the time to complete two repairs, where each repair time is exponentially distributed, is of interest.

  Example: If the time to complete the first repair is $X_1 \sim \text{Exp}(\lambda_1)$ and the second is $X_2 \sim \text{Exp}(\lambda_2)$, the total time to complete both repairs is the sum $X_1 + X_2$.

In general, the sum of two independent exponential random variables leads to a **Gamma** distribution if both have the same rate, or a more complex distribution (known as **hypoexponential** distribution.) if the rates differ. The proof of Gamma distribution is shown below. The proof of hypoexponential distribution and its usage in obtaining the mean-square-error estimator is the subject of this activity.

If $X \sim \exp(\lambda_1), Y \sim \exp(\lambda_2)$ and $\lambda_1 = \lambda_2 = \lambda$, the sum $Z = X + Y$ has pdf given by the convolution

$$
\begin{aligned}
f_Z(z) &= \int_{-\infty}^{\infty} f_Y(z-x) f_X(x) dx \\
&= \lambda^2 \int_0^z e^{-\lambda(z-x)} e^{-\lambda x} dx \\
&= \lambda^2 e^{-\lambda z} \int_0^z 1 dx \\
&= \lambda^2 z e^{-\lambda z}, z \geq 0
\end{aligned}
$$

which is the **two-parameter Gamma**$(2, \lambda)$ distribution.

Problem 01: Suppose we observe $Y = X + V$, where X and V are independent random variables with exponential distributions with parameters λ_1 and λ_2 ($\lambda_1 \neq \lambda_2$). That is, the pdfs of X and V are $f_X(x) = \lambda_1 e^{-\lambda_1 x}$ for $x \geq 0$ and $f_V(v) = \lambda_2 e^{-\lambda_2 v}$ for $v \geq 0$, respectively.

(a) Using the fact that the pdf of the sum of two independent random variables is the convolution of the individual pdfs, show that

$$f_Y(y) = \frac{\lambda_1 \lambda_2}{\lambda_2 - \lambda_1} e^{-\lambda_2 y} \left(e^{(\lambda_2 - \lambda_1) y} - 1 \right), \quad y \geq 0$$

which is the **two-parameter hypoexponential** distribution.

(b) Establish that

$$f_{X,Y}(x, y) = \lambda_1 \lambda_2 e^{(\lambda_2 - \lambda_1)x - \lambda_2 y},$$

for $x \geq 0$ and $y \geq 0$.

(c) Show that the optimal mean-squares estimate of X given $Y = y$

$$\widehat{X}_{\text{MSE}} = \mathrm{E}[X \mid Y] = \frac{1}{\lambda_1 - \lambda_2} - \frac{y\, e^{-\lambda_1 y}}{e^{-\lambda_2 y} - e^{-\lambda_1 y}}.$$