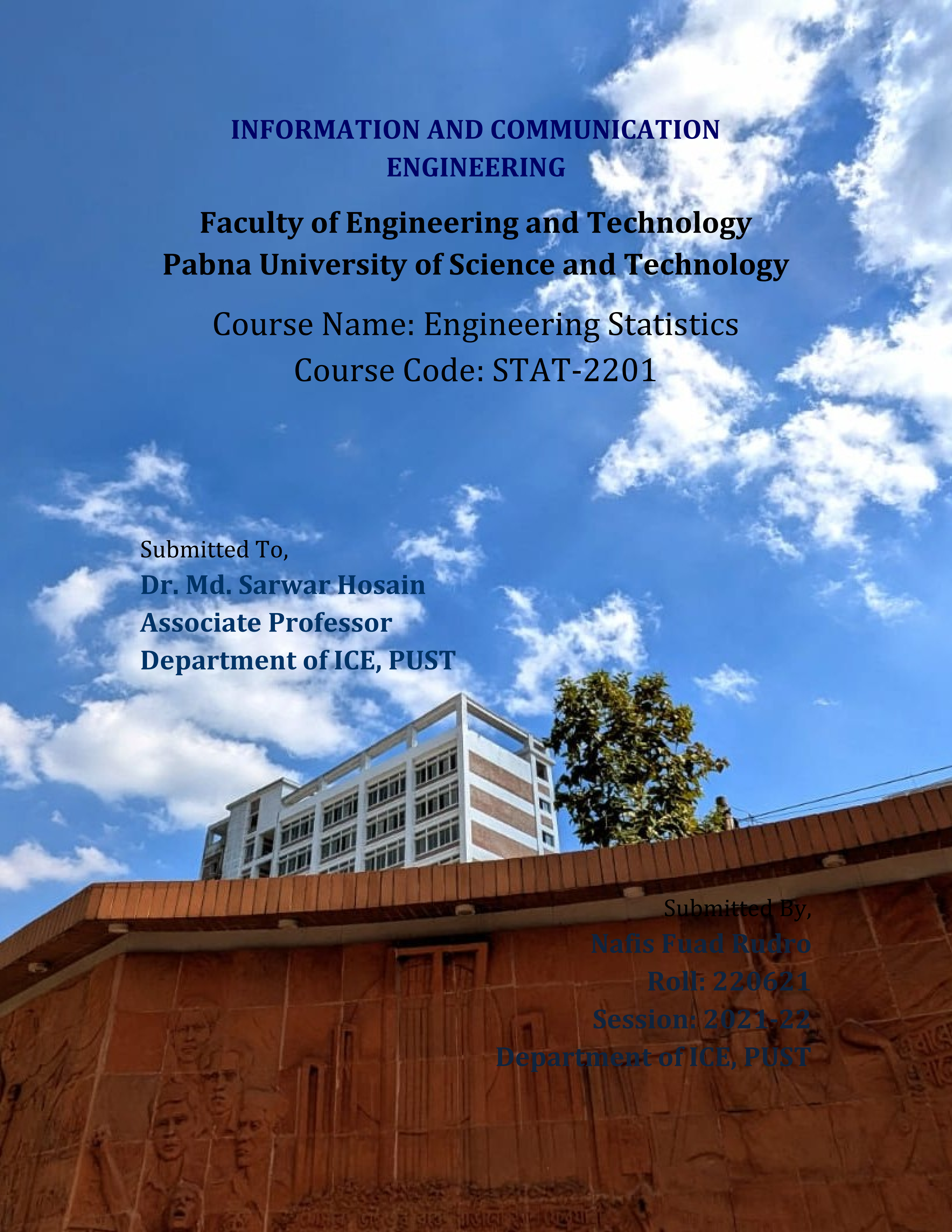

# INFORMATION AND COMMUNICATION ENGINEERING

## Faculty of Engineering and Technology
## Pabna University of Science and Technology

Course Name: Engineering Statistics
Course Code: STAT-2201

Submitted To,
**Dr. Md. Sarwar Hosain**
**Associate Professor**
**Department of ICE, PUST**


Submitted By,
**Nafis Fuad Rudro**
**Roll: 220621**
**Session: 2021-22**
**Department of ICE, PUST**

# Maximum Likelihood Estimation (MLE)

# Illustration from Poisson and Normal Distributions

## 1. Maximum Likelihood Estimation (MLE) from Poisson Distribution

### Theory of MLE from Poisson Distribution

The **Poisson distribution** models the probability of a given number of events happening in a fixed interval of time or space, given a fixed rate of occurrence. It is defined by the following probability mass function (PMF):

$$P(X = x) = \frac{\lambda^x e^{-\lambda}}{x!}, \quad x = 0, 1, 2, \ldots$$

Where: $\lambda$ is the rate parameter (mean of the distribution), $x$ is the number of occurrences, $e$ is Euler's number.

The **Maximum Likelihood Estimation (MLE)** is a method used for estimating the parameters of a statistical model. The idea behind MLE is to choose the parameter values that maximize the likelihood of the observed data.

For a set of independent and identically distributed (i.i.d.) observations $X_1, X_2, \ldots, X_n$, the likelihood function $L(\lambda)$ is the joint probability of observing the data. The likelihood function for the Poisson distribution is given by the product of the individual probabilities for each observation:

$$L(\lambda) = \prod_{i=1}^{n} \frac{\lambda^{x_i} e^{-\lambda}}{x_i!}$$

The log-likelihood function is the natural logarithm of the likelihood function:

$$\ell(\lambda) = \ln(L(\lambda)) = \sum_{i=1}^{n} (x_i \ln(\lambda) - \lambda - \ln(x_i!))$$

To find the MLE for $\lambda$, we differentiate the log-likelihood function with respect to $\lambda$ and set the derivative equal to zero.

## Proof of MLE for Poisson Distribution

Let's differentiate the log-likelihood function:

$$\ell(\lambda) = \sum_{i=1}^{n} (x_i \ln(\lambda) - \lambda - \ln(x_i!))$$

Now, differentiate with respect to $\lambda$:

$$\frac{d}{d\lambda}\ell(\lambda) = \sum_{i=1}^{n} \left(\frac{x_i}{\lambda} - 1\right)$$

Setting the derivative equal to zero to find the critical point:

$$\sum_{i=1}^{n} \left(\frac{x_i}{\lambda} - 1\right) = 0$$

This simplifies to:

$$\sum_{i=1}^{n} x_i = n\lambda$$

Thus:

$$\hat{\lambda} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

The MLE for $\lambda$ is the sample mean of the observed data.

## Math Example for Poisson Distribution

Suppose we have the following data representing the number of events observed in 5 independent time intervals: $X_1 = 2, X_2 = 3, X_3 = 4, X_4 = 3, X_5 = 2$.

First, calculate the sample mean:

$$\hat{\lambda} = \frac{1}{5}(2+3+4+3+2) = \frac{14}{5} = 2.8$$

Thus, the MLE estimate for $\lambda$ is 2.8.

# 2. Maximum Likelihood Estimation (MLE) from Normal Distribution

## Theory of MLE from Normal Distribution

The **Normal distribution** is one of the most commonly used distributions in statistics. Its probability density function (PDF) is given by:

$$f(x|\mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

Where:

$\mu$ is the mean (location

parameter), $\sigma^2$ is the variance

(scale parameter), $x$ is the

observation.

The likelihood function for a set of independent observations $X_1, X_2, \dots, X_n$ is:

$$L(\mu, \sigma^2) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x_i-\mu)^2}{2\sigma^2}}$$

The log-likelihood function is:

$$\ell(\mu, \sigma^2) = \sum_{i=1}^{n} \left( -\frac{1}{2} \ln(2\pi\sigma^2) - \frac{(x_i - \mu)^2}{2\sigma^2} \right)$$

To estimate $\mu$ and $\sigma^2$, we differentiate the log-likelihood function with respect to $\mu$ and $\sigma^2$, and set the derivatives equal to zero.

## Proof of MLE for Normal Distribution

Let's first differentiate the log-likelihood function with respect to $\mu$:

$$\frac{d}{d\mu}\ell(\mu,\sigma^2)=\sum_{i=1}^{n}\frac{x_i-\mu}{\sigma^2}$$

Setting the derivative equal to zero:

$$\sum_{i=1}^{n}(x_i-\mu)=0$$

This simplifies to:

$$\hat{\mu}=\frac{1}{n}\sum_{i=1}^{n}x_i$$

Thus, the MLE for $\mu$ is the sample mean.

Next, differentiate the log-likelihood function with respect to $\sigma^2$:

$$\frac{d}{d\sigma^2}\ell(\mu,\sigma^2)=\sum_{i=1}^{n}\left(-\frac{1}{2\sigma^2}+\frac{(x_i-\mu)^2}{2\sigma^4}\right)$$

Setting the derivative equal to zero:

$$\sum_{i=1}^{n}\left(-\frac{1}{2\sigma^2}+\frac{(x_i-\mu)^2}{2\sigma^4}\right)=0$$

This simplifies to:

$$\hat{\sigma^2}=\frac{1}{n}\sum_{i=1}^{n}(x_i-\hat{\mu})^2$$

Thus, the MLE for $\sigma^2$ is the sample variance.

## Math Example for Normal Distribution

Suppose we have the following data representing the heights of 6 individuals (in cm): $X_1 = 162, X_2 = 170, X_3 = 168, X_4 = 165, X_5 = 171, X_6 = 169$.

First, calculate the sample mean $\hat{\mu}$:

$$\hat{\mu} = \frac{1}{6}(162 + 170 + 168 + 165 + 171 + 169) = \frac{1005}{6} = 167.5$$

Next, calculate the sample variance $\hat{\sigma}^2$:

$$\hat{\sigma^2} = \frac{1}{6}\sum_{i=1}^{6}(x_i - 167.5)^2 = \frac{1}{6}\left((162 - 167.5)^2 + (170 - 167.5)^2 + \cdots + (169 - 167.5)^2\right)$$

$$\hat{\sigma^2} = \frac{1}{6}(30.25 + 6.25 + 0.25 + 6.25 + 12.25 + 2.25) = \frac{57.5}{6} = 9.5833$$

Thus, the MLE for $\mu$ is 167.5 cm, and the MLE for $\sigma^2$ is 9.5833 cm$^{\dagger}$.