**Name: Muhammad Shoaib Akhter Qadri**

**Enrollment No: 02-131212-009**

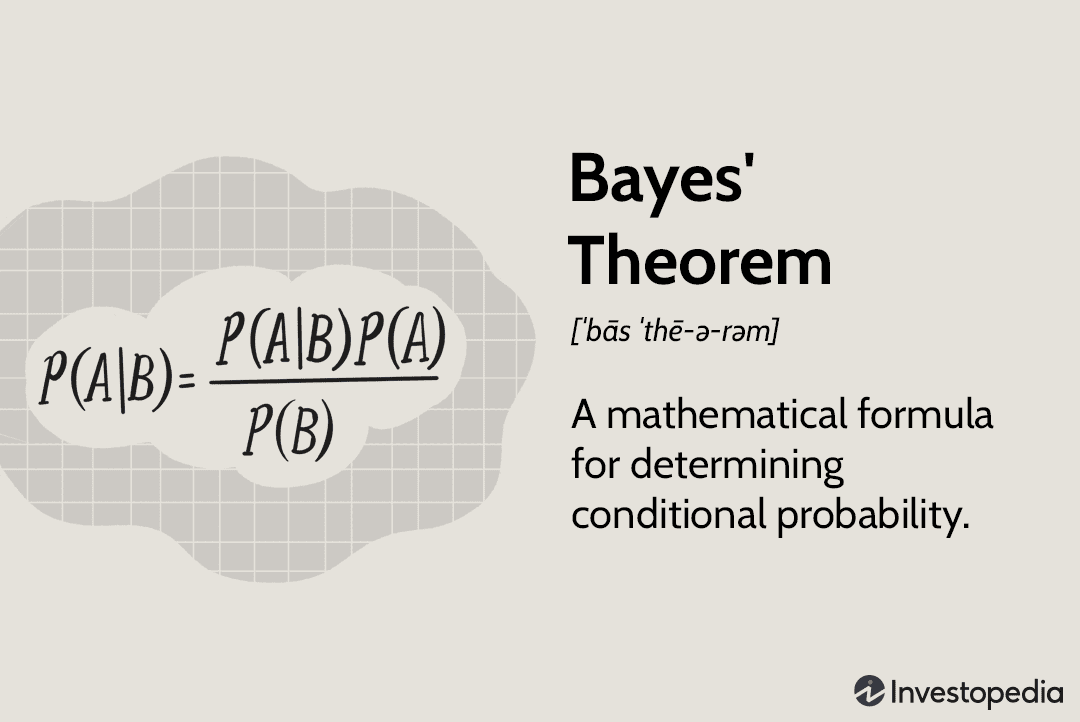
**Roll No: 79290**

**Assignment No : 03**

Q1. Describe at least 2 applications of Bayesian Probability in Computer Science / Software Engineering.

### **Applications of Bayes' theorem:**

In more practical terms, Bayes' theorem allows scientists to combine a priori beliefs about the probability of an event (or an environmental condition, or another metric) with empirical (that is, observation-based) evidence, resulting in a new and more robust posterior probability distribution.



**Pollutant Removal Infrastructure Performance:**

In this hypothetical example, we are trying to improve our understanding of how effective stormwater management infrastructure systems are at removing sediment from stormwater runoff. While sediment often carries nutrients, metals, and other contaminants, sediment itself is also a pollutant in many environmental systems.

#### Predicting Water Quality Conditions:

#### Water quality is often measured by the concentration of one or more in situ pollutants (such as nutrients, bacteria, and organic compounds), and the suitability of a particular water body for its intended use (such as drinking water, recreation, or agricultural use) depends on whether or not the measured pollutant concentrations exceed water quality standard numeric limits. Because these pollutants often cannot be measured directly, scientists typically measure indicators that serve as potential surrogates for the pollutant of concern.

Q2. Describe the mechanism which is used to apply Bayesian Probability in your stated applications in Q1.

#### Understanding Pollutant Removal Infrastructure Performance:

Figure 1 presents an example of how Bayes' theorem can be applied to solve environmental problems. In this hypothetical example, we are trying to improve our understanding of how effective stormwater management infrastructure systems are at removing sediment from stormwater runoff. While sediment often carries nutrients, metals, and other contaminants, sediment itself is also a pollutant in many environmental systems. In this problem, we represent the fraction of sediment removed by a stormwater management system as θ. Figure 1 presents the evolution of this understanding in a Bayesian framework, beginning with the development of a prior probability distribution. The prior probability distribution for θ is based on pollutant removal rate values in a published database documenting hundreds of studies , and is expressed in Fig. 1 first as a histogram of historic values (Fig. 1*a*), and then as a dashed line approximating the pollutant removal rate prior probability distribution (Fig. 1*b*). Hypothetical sediment removal rates from a new study site are then introduced through a likelihood function (solid line in Fig. 1*c*), and finally the posterior probability distribution is calculated using Bayes' theorem (and represented by a dotted line in Fig. 1*d*).

Chart

Description automatically generated with medium confidence

#### Predicting Water Quality Conditions:

Water quality is often measured by the concentration of one or more in situ pollutants (such as nutrients, bacteria, and organic compounds), and the suitability of a particular water body for its intended use (such as drinking water, recreation, or agricultural use) depends on whether or not the measured pollutant concentrations exceed water quality standard numeric limits. Because these pollutants often cannot be measured directly, scientists typically measure indicators that serve as potential surrogates for the pollutant of concern. The strength of the relationship between an indicator concentration and the concentration of the pollutant it supposedly represents varies widely depending on the type of pollutant. For example, in recreational and shellfish-harvesting waters throughout the United States, water quality is based on the concentration of nonpathogenic fecal indicator bacteria (FIB) such as fecal coliforms and *Escherichia coli*. These bacteria are used as a conservative indicator of fecal contamination and of the potential presence of harmful waterborne pathogens, which, while more directly linked to human and environmental health, are also much more difficult and costly to measure. Regardless of the specific pollutant and associated indicator, it is clear that not only the pollutant-indicator relationship, but also the spatial and temporal frequency of sampling and other factors might collectively contribute to uncertainty and variability in environmental condition forecasts. Here, we present a Bayesian approach to assessing water quality conditions using fecal coliform concentration measurements (reported in organisms per 100 ml) in a shellfish harvesting area as an example.

Like many other pollutants, FIB concentrations are commonly assumed to follow a lognormal LN (μ, σ) probability distribution with log-concentration mean (*μ*) and log-concentration standard deviation (σ). While this common probability model acknowledges natural spatial and temporal variability in FIB dispersion patterns, it (like other simple probability models) often fails to explicitly acknowledge other, more subtle sources of variability, including intrinsic sources arising from FIB concentration measurements and how FIB concentrations are calculated, all of which can lead not only to uncertainty in FIB concentration predictions, but to uncertainty in probability distribution parameters (that is, μ and σ) as well. In a Bayesian framework, we can explicitly acknowledge these uncertainties by first placing a prior probability distribution on the population parameters μ and σ (which may account for a priori beliefs about their potential values), then developing a likelihood function for μ and σ based on empirical evidence (in this case, using water quality samples), and, finally, deriving a joint posterior probability distribution for both. Results of this procedure are presented in [Fig. 2](https://www.accessscience.com/content/article/aYB100249#fYB100249FG0020), which includes a smoothed contour plot of the joint posterior probability density for the fecal coliform log-concentration mean (μ) and standard deviation (σ) for a sample site in eastern North Carolina.

