Lecture 9

Conditional Law of Probability

Conditional probability is define as the probability of some events occuring in relationship with one or more other events.

- It is sometimes useful to interpret P[A] as our knowledge of the occurrence of event A before an experiment takes place.
- Sometimes, we refer to P[A] as the a priori probability of A.

Given the two events E and F with $P\{F\} > 0$, the conditional probability of E given F, $P\{E|F\}$, is defined as

$$P\{E|F\} = \frac{P\{EF\}}{P\{F\}}, \quad P\{F\} > 0$$

If E is a subset of (i.e., contained in) F, then $P\{EF\} = P\{E\}$.

The two events, E and F, are *independent* if, and only if,

$$P\{E|F\} = P\{E\}$$

In this case, the conditional probability law reduces to

$$P\{EF\} = P\{E\}P\{F\}$$

➢ ***Example 1:***

You are playing a game in which another person is rolling a die. You cannot see the die, but you are given information about the outcomes. Your job is to predict the outcome of each roll. Determine the probability that the outcome is a 6, given that you are told that the roll has turned up an even number.

Let $E = \{6\}$, and define $F = \{2, 4, \text{or } 6\}$. Thus,

$$P\{E|F\} = \frac{P\{EF\}}{P\{F\}} = \frac{P\{E\}}{P\{F\}} = \left(\frac{1/6}{1/2}\right) = \frac{1}{3}$$

Note that $P\{EF\} = P\{E\}$ because E is a subset of F.

Memoryless Property

Exponential distribution exhibits a unique property known as Memoryless property.

• Let T be a random variable denoting service time of a server.

• Suppose that a job currently with the server has already consume service time t.

• We are interested in the probability that the job will stay with the server for additional time s; i.e. we wish to calculate the conditional probability P[T > t + s | T > t]

i.e. the probability that the service time of a server is greater than the time job has already consumed plus the additional time it will stay with the server given that the service time of server is greater than the time job has already consumed. So by putting the values in the formula,

$$P[T > t + s \mid T > t] = \frac{P[T > t + s]}{P[T > t]} = \frac{e^{-\lambda(t+s)}}{e^{-\lambda t}} = e^{-\lambda s}$$

Here from the result we can see that it is independant of t which means that the probability of spending additional service time is

not depend on the time that has already been spent with the server. In other words we can also say that the past has no bearing on future & hence the name is called Memoryless property.
The Memoryless property for exponentially distributed lifetime would mean that the probability of a component serviving for some additional time would be independant of how long it has been operating.

Recall Formulae

• If λ is the rate at which Poisson events occur, then the distribution of time between successive arrivals, t,
$f(t) = \lambda e^{-\lambda t}$, $t >= 0$. This is the Probability Density Function (PDF).
• $1/\lambda$ is the average time interval between successive events.
• The Cumulative Distribution Function (CDF) of Exponential Distribution is given by:

$$P\{t \leq T\} = \int_0^T \lambda e^{-\lambda t}\, dt$$
$$= 1 - e^{-\lambda T}$$

• Additionally, $P(t > T) = e^{-\lambda T}$
• $P(a < t \leq b) = e^{-\lambda a} - e^{-\lambda b}$
• Memoryless Property: $P(X > x + k \mid X > x) = P(X > k)$

Percentile

• A percentile is a measure used in statistics indicating the value below which a given percentage of observations in a group of observations falls.
• For example, the 20th percentile is the value (or score) below

which 20% of the observations may be found. Equivalently, 80% of the observations are found above the 20th percentile. Lets suppose you are performing in a certain test & you might know that you scored 67 out of 90 in a test. But that has no real meaning unless you know what percentile you fall into.If you know that your score is in the 90^{th} percentile taht means that you score better than 90% of the people who took the test.

- The 25th percentile is also called the first quartile.
- The 50th percentile is generally the median.
- The 75th percentile is also called the third quartile.

rth Percentile Value

The rth percentile value $\prod(r)$, for any random variable X, is defined by:

$$P[X \leq \prod(r)] = r/100$$

Thus the 90^{th} percentile value of an exponential random variable is defined by:

$P[X \leq \prod(90)] = 0.9$ or $1 - e^{-\lambda\Pi(90)} = 0.9$; put the value of r = 90

Hence,

$$e^{-\lambda\Pi(90)} = 0.1$$

Taking natural logarithm on both sides,

$$-\lambda\prod(\mathbf{90}) = \ln(0.1)$$

$$\prod(\mathbf{90}) = \frac{\ln(0.1)^{-1}}{\lambda}$$

$$\prod(\mathbf{90}) = E[X]\ln(10)$$

$$\underline{\prod(\mathbf{90}) = \mathbf{2.3\,E[X]}}$$

here $1/\lambda$ is replaced by the E[X] which is also called the expected value, average value or mean value. So the r^{th} percentile value for r = 90 is 2.3 times the average or expected value.

Example Problem 1

Personnel of the “Farout Engineering Company” use an online terminal to make routine engineering calculations. If the time each engineer spends in the session at a terminal has an exponential distribution with an average value of 36 minutes, find:

a) The probability that engineer will spend 30 minutes or less at the terminal.

b) The probability that an engineer will use it for more than an hour.

c) If the engineer has already been at the terminal for 30 minutes, what is the probability that he will spend more than another hour at the terminal?

d) 90% of the session ends in “R” minutes. What is R ?

Solution

first task is to find out appropriate probability distribution to be used.

Since in this example you have been provided with this information that the terminal is having exponential distribution with an average value of 36 minutes.

It means that you have been provided with the expected value or the average value (i.e. 36 mins) which is also been represented by $1/\lambda$.

$E[X] = 36$ min Therefore, $\lambda = 1/36$

1) $P(t \leq 30) = 1 - e^{-30/36} = 0.5654$

2) $P(t > 60) = e^{-60/36} = 0.1888$

3) It does not matter. If the engineer is using the terminal already, according to memoryless property, it has no impact that he will use it for another 1 hour.

Hence $P(t > 60) = 0.1888$
4) $R = \prod(90) = 2.3\ E[X] = 2.3 * 36 = 82.8$ minutes

Example Problem 2
Requests arrive at a database server randomly, every 12 msecs on the average, following negative-exponential distribution. Determine the probability that the inter-arrival time of requests:
- does not exceed 4 msecs
- does exceed 4 msecs
- falls between 3 msecs and 6 msecs (inclusive)

Answers:
- 0.2834
- 0.7165
- 0.17227

Example Problem 3
Consider a web server with Poisson arrival stream at an average rate of 60/hour. Calculate the probability that the inter-arrival time is:
- longer than 4 min,
- between 2 and 6 min,
- shorter than 8 min.

Answers:
- 0.0183
- 0.1328
- 0.9996

Example Problem 4
Listen at 26:00

Example

Cars arrive at a gas station randomly every 2 minutes, on the average. Determine the probability that the interarrival time of cars does not exceed 1 minute.

The desired probability is of the form $P\{x \leq A\}$, where $A = 1$ minute in the present example. The determination of this probability is the same as computing the CDF of x—namely,

$$P\{x \leq A\} = \int_0^A \lambda e^{-\lambda x}\,dx$$
$$= -e^{-\lambda x}\Big|_0^A$$
$$= 1 - e^{-\lambda A}$$

The arrival rate for the example is computed as

$$\lambda = \tfrac{1}{2} \text{ arrival per minute}$$

Thus,

$$P\{x \leq 1\} = 1 - e^{-\left(\frac{1}{2}\right)(1)} = .3934$$

<u>Tasks</u>

Q1) Derive the generalized formula for rth Percentile value for the random variable X of Exponential Distribution. Also determine the value of ∏x(95).

Q2) A modem transmits a data packet over phone line every 50 milliseconds, on the average, following negative-exponential distribution. Calculate the probability (up to 4 decimal places) that the inter-arrival time of transmissions

- does not exceed 30 msec.
- exceeds 50 msec.
- is between 30 msec and 50 msec.

<u>Bayes' Theorem</u>

- Bayes' theorem can be used to calculate the probability that a certain event will occur or that a certain proposition is true, given that we already know a related piece of information.
- P(B) is called the prior probability of B. ; where B represents

an event.

• $P(B|A)$, being called the conditional probability, is also known as the posterior probability of B.

• Let us briefly examine how Bayes' theorem is derived: We can deduce a further equation from the following product rule:-

$$P(A \wedge B) = P(A|B)P(B)$$

• Note that due to the commutativity of ∧, we can also write

$$P(A \wedge B) = P(B|A)P(A)$$

• Hence, we can deduce:-

$$P(B|A)P(A) = P(A|B)P(B)$$

• This can then be rearranged to give Bayes' theorem:-

$$P(B|A) = \frac{P(A|B) \cdot P(B)}{P(A)}$$

➢Example: Medical Diagnosis

• Let us examine a simple example to illustrate the use of Bayes' theorem for the purposes of medical diagnosis.

• When one has a cold, one usually has a high temperature (let us say, 80% of the time).

- A: "I have a high temperature" and
- B: "I have a cold"

• Therefore, we can write this statement of posterior probability as

$$P(A|B) = 0.8$$

▪ Note that in this case, we are using A and B to represent pieces of data that could each either be a hypothesis or a piece of evidence.

▪ Now, let us suppose that we also know that at any one time around 1 in every 10,000 people has a cold, and that 1 in every 1000 people has a high temperature.

▪ We can write these prior probabilities as

- P(A) = 0.001
- P(B) = 0.0001
-

• Now suppose that you have a high temperature.
• What is the likelihood that you have a cold?
• This can be calculated very simply by using Bayes' theorem:

$$
\begin{aligned}
P(B|A) &= \frac{P(A|B)\cdot P(B)}{P(A)} \\
&= \frac{0.8\cdot 0.0001}{0.001} \\
&= 0.008
\end{aligned}
$$

• Hence, we have shown that just because you have a high temperature does not necessarily make it very likely that you have a cold—in fact; the chances of cold are just 8 in 1000.

• Task: Recall the Axioms of Probability
So Axioms are the facts to be accepted without any prove.
and theorems are consequences that follow logically from the definitions & Axioms. & All the theorems has a prove that refers to definitions, Axioms & other theorems.
Although there are dozens of definitions & theorems of probability. There are only 3 axioms of probability theory.

Recall those 3 axioms & identify them.