AN EMPIRICAL DISTRIBUTION
IN CONTINUOUS CASE

Abstract: Probability is useful among other things since we use it to make sense of the world around us. We use it to make inferences about the things that we do not see directly. And this is done by using probability distributions. We have already seen some of the distributions. Several distributions have been developed by some transformations on the existing distributions. First part of the paper is the estimating one of the distributions developed by smoothing techniques of the histograms and in the second part is the cumulative distribution function of estimated distribution.

Key Words: Random variable, continuous probability distribution, arrival rate, density function, histograms, Cumulative Probability distribution.

Introduction Random variable of interest is “Staying only n number of arrivals in the system in particular time interval”. The arrivals stay in the system while going in queue to service and questioning for some data in the system (before leaving the system). Instead of asking “how many arrivals take place in a particular time interval (Poisson)”, we ask for “how likely the system have n number of arrivals in particular time interval”. Since X is continuous, the PDF should be a function. We have to make inferences about this unknown function. This means the probability distribution that takes into account of measurements those we have surveyed for a considerable period of time. So the output of the inference problem is to come up with the distributions of X. We plot the histograms for different number of arrivals from which we find the density curves.

Staying only that ‘n’ no. of arrivals in the system in a particular interval:

Zero arrivals happen in the system in the following ways:

If no person enters the system ⇒ 0 arrivals stay in the system

If n persons enters the system in particular
Interval and n persons leave the system ⇒ 0 arrivals stay in the system
who receives the service in that particular
interval. And the system is idle for some
time.

If n-1 persons enters the system in particular
interval and n-1 persons leave the system \( \Rightarrow 0 \) arrivals stay in the system
after receiving the service in that particular
interval.

\[ \vdots \]

\[ \vdots \]

and so on.

‘0’ arrivals stay in the system not only at starting time, system may idle for some time when
entered arrivals have taken the service and left the system.

The following figure 1 is the histogram of the data when the system is with 0 no. of arrivals in
different time periods.

FIGURE-1
The following figure 2 is the density function obtained from the histogram of data when the system is with 0 no. of arrivals in different time periods.

**FIGURE-2**

![Figure 2: Density function graph](image)

Series-1 represents the probability density function graph of new random variable when \( n = 0 \).

\[
X \text{- represents time, } Y \text{- represents } f(x).
\]

The Poisson density function with some extra normalizing constant satisfying this graph. The arrival rate \( \lambda \) is taken as normalizing constant. ‘0’ arrivals stay in the system not only at starting time, we may have the system staying with zero arrivals after some time also. When entered arrivals have taken the service and there are no arrivals for some time then the system is idle for some time.

\[
f(x) = \frac{e^{-\lambda x} (\lambda x)^0}{0!}
\]

\[
f(x) = e^{-\lambda x} \lambda
\]

For chosen random variable, this \( f(x) = e^{-\lambda x} \lambda \) is to be probability density functions if the function satisfies the following two conditions.

\[
f(x) \geq 0 \forall x
\]

\[
\int_{-\infty}^{\infty} e^{-\lambda x} \lambda dx = 1.
\]
i.e. $e^{-\lambda x} \geq 0 \forall x$ since exponential gives positive values and $\lambda$ is intensity factor (arrival rate) which is positive

Now $\int_{-\infty}^{\infty} e^{-\lambda x} \lambda \, dx$

$= \lambda \left[ \frac{e^{-\lambda x}}{-\lambda} \right]_{0}^{\infty}$

$= -\left[ e^{-\infty} - e^{0} \right] = 1$

$\therefore f(x) = e^{-\lambda x}$

$\therefore f(x) = e^{-\lambda x}$ is the probability density function for the assumed random variable for $n = 0$.

For $n = 1$:

One arrival happens in the system in the following ways:

or

The system stays with only one arrival in the following ways:

If only one person enters the system $\Rightarrow$ system stays with 1 arrival.

particular interval

If $n$ persons enter the system in

a particular interval and $n-1$ persons $\Rightarrow$ system stays with 1 arrival.

leave the system after service.

In general,

If $(n-1)$ persons enter the system in

particular interval and $(n-2)$ persons $\Rightarrow$ system stays with 1 arrival.
interval system after receiving their service

......

......

and so on.

The following figure 3 is the density function of the data obtained from the smoothing the histogram when the system is with 1 no. of arrivals in different time periods. The Poisson density function with some extra normalizing constant satisfying this curve. The arrival rate $\lambda$ is taken as the normalizing constant.

**FIGURE-3**

Series-2 represents the probability density function graph of new random variable when $n = 1$.

$X$- represents time, $Y$- represents $f(x)$.

$$f(x) = \frac{e^{-\lambda x} (\lambda x)^\lambda}{1!}$$

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

For chosen random variable, this is to be probability density function if
1. \( f(x) \geq 0 \forall x \)

i.e. \( e^{-\lambda x} \lambda x \lambda \geq \forall x \)

Since exponential gives all positive values for \( t \geq 0 \) and \( \lambda \geq 0 \)

2. Now \( = \int_{-\infty}^{\infty} e^{-\lambda x} \lambda x \lambda dx \)

\[ = \int_{0}^{\infty} e^{-\lambda t} \lambda t \lambda dt \]

\( v = e^{-\lambda t} \quad u = \lambda t \);

\[ \int_{0}^{\infty} e^{-\lambda t} \lambda dt \]

\[ = \lambda \left[ \frac{\lambda t e^{-\lambda t}}{-\lambda} - \int_{0}^{\infty} e^{-\lambda t} \lambda dt \right]_{0}^{\infty} \]

\[ = \left[ -te^{-\lambda t} + \frac{e^{-\lambda t}}{-\lambda} \right]_{0}^{\infty} \]

\[ = -\lambda \left[ e^{-a} + \frac{e^{0}}{-\lambda} \right] \]

\[ = -\lambda \left[ 0 - \frac{1}{\lambda} \right] \]

\[ = 1 \]

\[ \int_{0}^{\infty} \lambda e^{-\lambda x} dx = 1 \]
\( f(x) \geq 0 \forall x \)

\( \therefore f(x) \) is the density function for chosen random variable.

**For n = 2:**

**Two arrivals happen in the system in the following ways:**

or

The system stays with only two arrivals in the following ways:

If two persons enter the system in a particular interval and no service has been taken. \( \Rightarrow \) System has two arrivals at that particular interval.

If n persons enter the system in Particular interval and n-2 persons leave in that interval.

the system after their service

. . .

. . .

And so on.

The following figure 4 is the density function of the data obtained from the smoothing the histogram when the system is with 2 no. of arrivals in different time periods. The Poisson density function with some extra normalizing constant satisfying this curve. The arrival rate \( \lambda \) is taken as the normalizing constant.
Series-3 represents the probability density function graph of new random variable when \( n = 2 \).

\[ f(x) = \frac{e^{-\lambda x} (\lambda x)^2 \lambda}{2!} \]

For chosen random variable, this \( f(x) = \frac{e^{-\lambda x} (\lambda x)^2 \lambda}{2!} \) is to be probability density function if

\[ f(x) \geq \forall x \]

\[ \frac{e^{-\lambda x} (\lambda x)^2 \lambda}{2!} \geq \forall x \]

Since the exponential gives positive values, \( t \geq 0 \) and \( \lambda \geq 0 \)

Now \( \int_{-\infty}^{\infty} f(x)dx \)

\[ = \int_{0}^{\infty} e^{-\lambda t} (\lambda t) \, dt \]
\[\lambda^3 \int_0^\infty \left( \frac{e^{-\lambda} t^2}{-\lambda} - \int_0^\infty \frac{e^{-\lambda} t}{-\lambda} \, dt \right) \, dt\]

\[\lambda \left[ \frac{e^{-\lambda} t^2}{-\lambda} + \frac{e^{-\lambda} t}{-\lambda^2} - \int_0^\infty \frac{e^{-\lambda} t}{-\lambda} \, dt \right] \]

\[\lambda \left[ \frac{e^{-\lambda} t}{-\lambda} + \frac{e^{-\lambda} t^2}{\lambda^2} - \frac{e^{-\lambda} t}{-\lambda} \right] \]

\[\lambda \left[ \frac{e^{-\lambda} t}{-\lambda} \right]_0^\infty = e^{-\alpha} - e^0 = 1\]

\[\int_0^\infty \frac{e^{-\lambda x} (\lambda x) \lambda}{2!} \, dx = 1\]

And \(f(x) \geq 0\) for all \(x\)

\[\therefore \text{For chosen random variable } f(x) \text{ is the density function.}\]

For \(n = 3\):

There are three arrivals happen in the system in the following ways:

or

The system stays with only three arrival in the following ways:

Only three persons enter the system \(\Rightarrow\) system has three arrivals in that particular interval.

If \(n\) persons enter the system in

Particular interval and \(n-3\) \(\Rightarrow\) system has 3 arrivals in that interval.
Persons leave the system in that particular interval after their service.

If (n-1) persons enter the system in particular interval and (n-4) persons leave the system in that interval after receiving their service in that interval.

......

......

and so on.

The following figure 5 is the density function of the data obtained from the smoothing the histogram when the system is with 3 no. of arrivals in different time periods. The Poisson density function with some extra normalizing constant satisfying this curve. The arrival rate $\lambda$ is taken as the normalizing constant.

**FIGURE-5**

![Series-4](image)

Series-4 represents the probability density function graph of new random variable when $n = 3$.

X- represents time, Y- represents f(x).
\[
f(x) = \frac{e^{-\lambda x} (\lambda x)^3}{3!}
\]

For chosen random variable, this \(f(x) = \frac{e^{-\lambda x} (\lambda x)^3}{3!}\) is to be probability density function if

\[f(x) \geq 0 \forall x\]

i.e. \(f(x) = \frac{e^{-\lambda x} (\lambda x)^3}{3!} \geq 0\)

Since exponential gives all positive values, \(t \geq 0\) and \(\lambda \geq 0\)

Now \(\int_{-\infty}^{\infty} \frac{e^{-\lambda x} (\lambda x)^3}{3!} dx\)

\(= \int_{0}^{\infty} \frac{e^{-\lambda x} (\lambda x)^3}{3!} dx\)

\(= 1\)

\(\therefore f(x)\) is density function for chosen random variable.

For \(n = 4\):

Four arrivals happen in the system in the following ways:

\[\text{or}\]

The system stays with only four arrivals in the following ways:

If only four persons enter the system \(\Rightarrow\) system stays with four arrivals.

If \(n\) persons enter the system in particular interval and \(n-4\) \(\Rightarrow\) system stays with 4 arrivals.
persons leave the system in that particular interval after their service.

If (n-1) persons enter the system in particular interval and (n-5) persons system after receiving the service in that interval

......

......

and so on.

The following figure 6 is the density function of the data obtained from the smoothing the histogram when the system is with 4 no. of arrivals in different time periods. The Poisson density function with some extra normalizing constant satisfying this curve. The arrival rate λ is taken as the normalizing constant.

FIGURE-6

Series-5 represents the probability density function graph of new random variable when n = 4.
X- represents time, Y- represents f(x).

\[ f(x) = \frac{e^{-\lambda x} (\lambda x)^4 \lambda}{4!} \]

For chosen random variable, this is \( f(x) = \frac{e^{-\lambda x} (\lambda x)^4 \lambda}{4!} \) to be probability density function if

\[ (1) f(x) \geq 0 \forall x \]

i.e. \( f(x) = \frac{e^{-\lambda x} (\lambda x)^4 \lambda}{4!} \geq 0 \)

Since exponential gives all the values, \( t \geq 0 \) and \( \lambda \geq 0 \)

Now

\[ \int_{-\infty}^{\infty} e^{-\lambda x} (\lambda x)^4 \lambda \, dx \]

\[ = \int_{0}^{\infty} e^{-\lambda x} (\lambda x)^4 \lambda \, dx \]

\[ = 1 \]

\[ \therefore f(x) \text{ is density function for chosen random variable.} \]

n arrivals happen in the system in the following ways:

or

The system stays with n persons in the following ways.
Series 2 represents the probability density function graph when $n = 1$.

Series 3 represents the probability density function graph when $n = 2$.

Series 4 represents the probability density function graph when $n = 3$.

Series 5 represents the probability density function graph when $n = 4$.

$X$ - Axis represents time; $Y$ - represents $f(x)$.

If $n$ persons enters in to the system $\Rightarrow$ $n$ persons are there in the
in a particular interval and no service system in the particular interval
has started at that particular interval

If $n+1$ persons enter the system in a $n$ persons are there in the
particular interval and 1 person has $\Rightarrow$ system in that particular interval.
taken the service

If \( n+2 \) persons enter the system in a particular interval and 2 persons have taken the service \( \Rightarrow n \) persons are there in that particular interval.

\\ldots

\\ldots

and so on.

Therefore in general for the chosen random variable “Staying only \( n \) number of persons in the system in particular interval” the probability density function is

\[
f(x) = \begin{cases} 
\frac{e^{-\lambda x} (\lambda x)^n}{n!} & \text{when } x \geq 0 \\
0 & \text{otherwise}
\end{cases}
\]

So we developed one probability distribution by using histograms of area of probability density.

To extend this distribution theory, the cumulative distributive function and interesting properties of all statistical distributions mean, variance and standard deviation are studied which are widely used in several fields insurance, management, business and finance etc.

**Cumulative distributive function (CDF) \( F_x(x) \)**

It is the probability that the random variable takes values less than or equal to certain number.

\[
F_x(x) = P(x \leq x) = \int_{-\infty}^{x} f_x(t)dt
\]

And derivative of CDF is equal to the density

\[
\frac{d}{dx} (F_x(x)) = f_x(x)
\]
The density function of assumed Random variable is

\[ f(x) = \frac{e^{-\lambda x} (\lambda x)^n}{n!} \]

\[ P(X \leq x) = f_X(x) = 1 - \int_x^\infty \frac{\lambda e^{-\lambda x}}{n!} (\lambda x)^n \, dx \]

For \( n = 0 \) the density function is \( f(x) = \lambda e^{-\lambda x} \)

\[ F_X(x) = 1 - \int_x^\infty \lambda e^{-\lambda x} \, dx \]
\[ = 1 - \left[ \frac{\lambda e^{-\lambda x}}{-\lambda} \right]_x^\infty \]
\[ = 1 + \left[ e^{-\lambda x} - e^{-\lambda x} \right] \]
\[ f_X(x) = 1 - e^{-\lambda x} \]

This gives up to certain time internal \([0, x]\) the probability of having ‘0’ arrivals in the system

For \( n = 1 \)

\[ F_X(x) = p(X \leq x) = 1 - \int_x^\infty \frac{\lambda e^{-\lambda x} \lambda x}{1!} \, dx \]
\[ = 1 - \lambda^2 \int_x^\infty xe^{-\lambda x} \, dx \]
\[ = 1 - \lambda^2 \left[ \frac{-xe^{-\lambda x}}{\lambda} - \frac{e^{-\lambda x}}{-\lambda} \right]_x^\infty \]
\[ = 1 - \lambda^2 \left[ \frac{-xe^{-\lambda x}}{\lambda} - \frac{e^{-\lambda x}}{\lambda^2} \right]_x^\infty \]
This gives up to certain time interval \([0, x]\) the probability of having one arrival in the system.

For \(n = 2\)

\[
F_x(x) = p(X \leq x) = 1 - \int_x^\infty \frac{\lambda e^{-\lambda x}(\lambda x)^2}{2!} dx
\]

\[
= 1 - \lambda^3 \int_x^\infty \frac{x^2 e^{-\lambda x}}{2!} dx
\]

\[
= 1 - \frac{(\lambda x)^2 e^{-\lambda x}}{2!} - \lambda xe^{-\lambda x} - e^{-\lambda x}
\]

\[
F_x(x) = \begin{cases} 
1 - \sum_{k=0}^{2} \frac{e^{-\lambda x} (\lambda x)^k}{k!} & \text{when } x \geq 0 \\
0 & \text{otherwise}
\end{cases}
\]

This gives up to certain time interval \([0, x]\) the probability of having two arrivals in the system.

For \(n = 3\)

\[
F_x(x) = p(X \leq x) = 1 - \int_x^\infty \frac{\lambda e^{-\lambda x}(\lambda x)^3}{3!} dx
\]

\[
= 1 - \lambda^4 \int_x^\infty \frac{x^3 e^{-\lambda x}}{3!} dx
\]
\[ F(x) = 1 - \frac{\lambda x^3 e^{-\lambda x}}{3!} - \frac{\lambda x^2 e^{-\lambda x}}{2!} - \lambda x e^{-\lambda x} - e^{-\lambda x} \]

\[ F_x(x) = \begin{cases} 
1 - \sum_{k=0}^{\infty} \frac{e^{-\lambda x} (\lambda x)^k}{k!} & \text{when } x \geq 0 \\
0 & \text{otherwise} 
\end{cases} \]

This gives up to certain time interval \([0, x]\) the probability of having three arrivals in the system.

For \( n = 4 \)

\[ F_x(x) = p(X \leq x) = 1 - \int_{-\infty}^{x} \frac{\lambda e^{-\lambda x} (\lambda x)^4}{4!} dx \]

\[ = 1 - \frac{\lambda}{x} \int_{-\infty}^{x} x^4 e^{-\lambda x} dx \]

\[ = 1 - \frac{\lambda}{4!} \sum_{k=0}^{\infty} \frac{e^{-\lambda x} (\lambda x)^k}{k!} \]

\[ F_x(x) = \begin{cases} 
1 - \sum_{k=0}^{\infty} \frac{e^{-\lambda x} (\lambda x)^k}{k!} & \text{when } x \geq 0 \\
0 & \text{otherwise} 
\end{cases} \]

This gives up to certain time interval \([0, x]\) the probability of having four arrivals in the system.

In general CDF function defined as

\[ f(x) = \begin{cases} 
1 - \sum_{k=0}^{\infty} \frac{e^{-\lambda x} (\lambda x)^k}{k!} & \text{when } x \geq 0 \\
0 & \text{otherwise.} 
\end{cases} \]

Graph of cumulative distribution function

Figure - 8
Series 1 is the cumulative distribution function graph when \( n = 1 \).

Series 2 is the cumulative distribution function graph when \( n = 2 \).

Series 3 is the cumulative distribution function graph when \( n = 3 \).

Series 4 is the cumulative distribution function graph when \( n = 4 \).

**CONCLUSION**

In this paper we proposed Probability density function and CDF for the assumed random variable “staying \( n \) number of arrivals in the system”, Instead of asking how many arrivals takes place.

For the assumed a continuous random variable, the density function proposed is

\[
f(x) = \begin{cases} 
    \frac{e^{-\lambda x} (\lambda x)^n}{n!} & \text{when } x \geq 0, \\
    0 & \text{otherwise}
\end{cases}
\]

Corresponding Cumulative Distributive Function

\[
F_x(x) = \begin{cases} 
    1 - \sum_{k=0}^{n} \frac{e^{-\lambda x} (\lambda x)^k}{k!} & \text{when } x \geq 0, \\
    0 & \text{otherwise}
\end{cases}
\]
FUTURE WORK

We can find mean, variance and standard deviation of the above probability distribution.

REFERENCES


[23] Ajeet Kumar Pandey, Srinivas Panchangam, Jessy George Smith, Safety analysis of Automatic Door operation for Metro Train: A case study. 9th International conference on Heterogeneous Networking for quality, reliability Robustness, January 11-12-2013, Greater Noida, India.


