Introduction to Queueing Theory

Prof. N. Selvaraju
Department of Mathematics
Indian Institute of Technology Guwahati, India

Lecture - 22
Erlang and Phase-Type Distributions

Hi and hello, everyone; what we have seen so far are Queueing models or Queueing systems which can be modelled by a stochastic process with Poisson process arrivals and exponentially distributed service times, and simple variations there of that is what we have seen. Variations mean we talked about this bulk arrival or bulk service queues that we have. So, which we have handled through this Poisson exponential model basically together if you can call, but in any system that if you look at it, especially in practice like you would see that the exponential distribution assumptions whether it is for the arrival process or it is for the service process this will be limiting our the applicability of this particular models especially the service time assumptions on exponential distribution because this Erlang himself has found out it is not just that it is a new thing that it has to be discovered. So, especially the service one like this may be a limitation of this model if we are to apply this. So, what do we do? Now, we will try to go beyond systems having these Poisson exponential assumptions; then, if so, then how do we handle such queues where the interarrival times and or service time distributions are not exponential. There could be many ways in which they can be handled; what we can look at, for example, one way is to look to express this distribution in terms of exponentials and whether these distributions can be represented through an exponential distribution; if so, the advantage could be then whatever we have done so far, the approach which we have carried out with respect to the Markovian queues like they will all still hold good. And such a method suppose if you can use this idea like if you can express the given distribution in terms of exponential, then one can use what is called as generally method of stages. Of course, we may not use the terminology, but that is what we are going to do is basically that, and when you do that, you can still be within the Markovian framework. So, our approach will remain the same. If that is not possible, then you can try to handle the actual distribution itself; if that is amenable in some sense, that amenability will come to that later, which we will also do later. So, that is what is our study in the last portion of this course, which is the $M/G/1$ or $G/M/1$ models and how one can handle that. But if that G that the general distribution that we have is too complex, and it is not amenable to even that kind of treatment, then one has to use some kind of approximation methods based on the mean and variance, and you may not deal with the distribution itself, but you may be limiting to mean and variance, and this is we are not going to cover in this course. So, this is beyond our scope. So, the first approach is this one this is what we thought. We are trying to express some sort of general distribution which is not exponential, but we are trying to express it in terms of exponential. So, within that framework, the Erlang distributed interarrival or service times form the starting point of dealing with more general distributions, and what kind of distribution this approach can handle is what we will see now.

Now, what is this Erlang distribution or in notation, we will be using this E_k; what is that mean is as follows.

- Recall that $T \sim Gamma(\alpha, \theta)$ with PDF

$$f_T(t) = \frac{\theta^\alpha}{\Gamma(\alpha)} t^{\alpha-1} e^{-\theta t}, \quad t > 0; \quad \alpha, \theta > 0.$$

▶ Here, $E(T) = \alpha/\theta$ and $Var(T) = \alpha/\theta^2$.

- Now, consider a special class of these distributions where α is restricted to positive integer. Specifically, put $\alpha = k$ and $\theta = k\mu$, where k is any arbitrary positive integer and μ is any arbitrary positive constant. This, gives the Erlang family of probability distributions with PDF and CDF (with parameters k and μ)

$$f(t) = \frac{(\mu k)^k}{(k-1)!} t^{k-1} e^{-k\mu t}, \quad t > 0,$$
$$F(t) = 1 - \sum_{n=0}^{k-1} e^{-k\mu t} \frac{(k\mu t)^n}{n!}, \quad t \geq 0 \qquad \text{[and } F(t) = 0 \text{ for } t \leq 0]$$

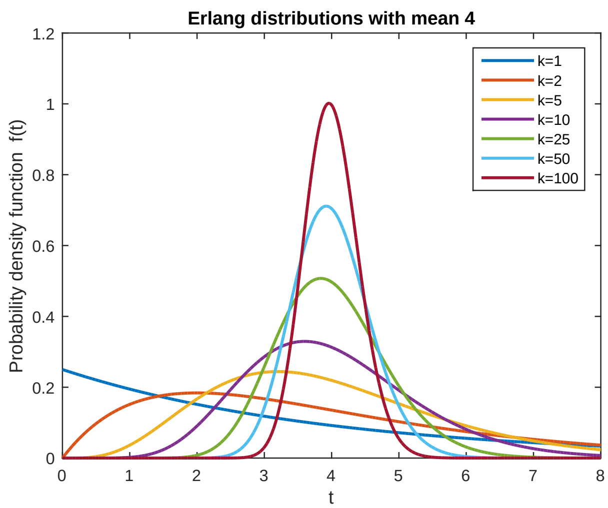
with the mean and the variance as $E(T) = \frac{1}{\mu}$, and $Var(T) = \frac{1}{k\mu^2}$.

So, one can easily get the coefficient of variation in terms of only k; you do not need to keep both the parameters in that is the advantage and also the mean we are keeping it as $1/\mu$. So, this the parameter in this thing which is this μthat is occurring here would be the rate of arrival or rate of service or rate of arrival that would still we want to retain it in that form, and that is why we are making this adjustment to this θ parameter also k. In general, sometimes, even if you do not make $\theta = k\mu$ kind of adjustment, we can still call it as an Erlang; as long as the α parameter is k which is a positive integer, but we will always keep it in $\theta = k\mu$ form unless we see otherwise that is what you keep this in mind.

- For a particular value of k, the distribution is referred to as an Erlang type-k or E_k distribution.

When you say now E k distribution with mean $1/\mu$, then you have specified both the parameters of the Erlang, which is k, which is coming in E_k and μ which is what is this one. So, this θ parameter would automatically take the form $k\mu$.

So, that is the form that we will be dealing with Erlang whenever we are deviating from this particular form of Erlang, and we may not call this $k\mu$. Suppose if you keep this as μ, that can still be called Erlang, but that is then a special case of gamma as well; this is also a special case of gamma. But, like whenever we happen that way like we will mention that otherwise, in general, we would take θ to be equal to $k\mu$ form is the common form of Erlang, which goes without saying. But one need not get fixated with this k mu factor alone. So, we need to be a little liberal on that side. So, this is what we define to be Erlang distribution.
Now, if you look at what has, we said that it might give a better fit, or whatever, why it can be considered to be more general than exponential. So, here we have plotted the probability density function of an Erlang distribution with simply mean as 4. So, 1 by mu, so mu is taken to be 0.25; that is all we have, and for different cases, we have plotted the curve.



- When $k = 1$, the Erlang reduces to an exponential distribution with mean $\frac{1}{\mu}$.
- As k increases, the Erlang becomes more symmetrical and more closely centered around its mean.
- For $k \to \infty$, the Erlang becomes deterministic with value $\frac{1}{\mu}$.
- In practice, the Erlang family provides more flexibility in fitting a distribution to real data than the exponential family.

So, this is what you are observing here by looking at certain plots of this distribution function. Now, as you can see as it can fit with different μ's and different k's; it can fit quite a bit of general distributions that can be modelled through this or at least approximately, if not exactly.

- The Erlang distribution is useful in queueing analysis because of its relationship to the exponential distribution.

  We know from $f_T(t) = \frac{\theta^\alpha}{\Gamma(\alpha)} t^{\alpha-1} e^{-\theta t}, \quad t > 0; \quad \alpha, \theta > 0.$ itself or when $\alpha = k$, which means when this α is an integer parameter, then you can relate this distribution of gamma distribution with an integer-valued α to its exponential to some exponential distribution, and that relationship is what is we exploited and this Erlang distribution whenever it is used in the queueing analysis. What is the fact?

- Fact: The sum of k IID exponential random variables with mean $\dfrac{1}{k\mu}$ is an E_k distribution.

  So, what you have here, suppose this is what is your scenario. So, each one of them is an exponential with $k\mu$. So, each one of them is an exponential with $k\mu$. So, what you have here as if some arrival happens here and something moves here, or the distribution the time if it is interarrival time or service time. This duration suppose if you think it is an interarrival time or service time. Suppose you think this service consists of some k stages independent stages, and each of the stages is exponentially distributed with parameter $k\mu$, then this total time can be represented as the sum of this k number of IID exponential with parameter $k\mu$. So, that is what is the fact. So, this is what we will give you an E_k distribution. So, what this gives us is that you can look at these stages or steps or, in general, this is what is called phases. So, this is suppose if you think of this as service, then as soon as it enters into the service, this is in the $1st$ phase, $2nd$ phase, and so on into the kth phase. So, remember there is a single server suppose whatever is the server, only one server or two. The server that is handling here would also handle the $2nd$ phase, would also handle the $3rd$ phase, would also handle the kth phase before he takes up another customer; that is how this thing will be looked at. So, the same service facility is what is this serving different phases; each of these phases can be considered to be IID exponential with parameter $k\mu$.

► This makes it possible to take advantage of the Markovian property of the exponential distribution (though Erlang is not Markovian).

Because we know that Erlang distribution by itself is not Markovian, but Erlang distribution if you view it as if its a sum of some number of IID exponential, then the properties of this exponential distribution that gave rise to this Markovian property nature or Markovian models can still be exploited to handle this kind of distribution. So, as you see that this in that sense say, for example, if there is some laboratory test and you are going getting into that laboratory there is one technician, and then that technician has to perform multiple tests on you, and each of this test takes an exponential amount of time, say the total service time unless he finishes with you he cannot go to the testing the other person. Suppose this is a scenario. So, here basically, naturally, you have phases that are exponential, but in many practical applications, this may not be the case, but even then, the uses of phases help from a mathematical perspective because we can then benefit from the properties of the exponential distribution or the Markovian models.

- The Erlang distribution is suitable in many applications, for example, when the service consists of many phases (or steps) performed by the same server with each step following an IID exponential distribution.

Just like we have seen in this laboratory test kind of situation, or even if it is not there, it will be useful to think as if the service is consisting of phases. The only thing is the overall distribution should be approximately at least close to some Erlang distribution with for some k and some μ. But, the drawback that you will be having with that is that as the number of phases increases, the state space of the model; if you have to bring this within the Markovian framework, then the state space of the model would also increase, and the model becomes complex to handle it. So, there is always a trade-off how much effort you want to put in if you want to bring it under the Markovian ambit; if not, maybe it may be useful to use some other way directly taking that as a distribution. So, that is what is the plus and minus of viewing this Erlang as a sum of exponential just to try to take advantage of the properties or the address that the exponential distribution might give you.

- The idea of phases (of exponentials) to construct distributions can be generalized to include a much wider class of distributions than just the Erlang. These are known as **phase-type** distributions.
- The distributions that uses the concept of phases are: Erlang, hyperexponential, hypoexponential (or generalized Erlang), Coxian and a general phase-type distribution.

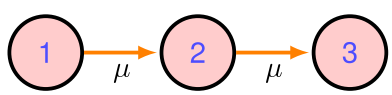
So, this is what is called phase-type distribution and what is the advantage if you want to use the phase-type distribution. As we already said that you can still be within the Markovian framework. So, you can use all the theories that you have developed so far to analyze such models, which are given in terms of phase-type distribution, whether it is interarrival times or service times.

- And also, the theoretical property that this phase-type distribution has is this a class of phase-type distributions is dense in the field of all positive valued distributions. What does that mean is it can be used to approximate in a quite good approximation it can give you for any positive valued distributions.

Since in queueing systems, we deal with only nonnegative random variables or positive valued random variables, because it is basically whether you are looking at either service time or interarrival times, this cannot be negative valued. So, it is positive valued random variables then this any such distribution that you can have with the appropriate choice of this k and μ Erlang or in general even if it is not falling within this Erlang or hyper-exponential, hypo-exponential or Coxian distributions then one can use a general phase-type distributions right to approximate that particular distribution using a phase-type distribution. And the phase-type distribution brings in the concept of phases and phases of exponentials; then, the general Markovian theory would apply. So, particularly if you see if I have a phase-type distribution, then I can as well approximate any general distribution that you may have to handle with respect to this a phase-type using a phase-type distribution. And that is why the popularity of these phase-type distributions are there in many of the queueing analyses that one does, which takes you closer to the reality than, say, an $M/M/1$ or a simple exponential model or even Erlang model or anything of that sort. And since the last late 1970s or the early 80s in the last four decades, you could see like a large number of research work has been done on these phase-type distributions, their properties, and how they can be applied successfully to queueing theory to develop queueing models which can model quite good very well the the the real-life distributions that you may encounter in this case. So, that is what is the advantage of phase-type. So, up to this much phase-type distributions is what is the generalization that one can make, and you can still be using the Markovian approach, which we can call as Markovian approach Markovian model approach, and that is the beauty of such a wider class of distribution that you can think.

Now, let us see how one can view this Erlang as a phase-type distribution in general. So, this is we have already said that Erlang is a sum of IID exponential, but we can also look at an alternative way, thinking that exponentials arise naturally as holding times of a process in a Markov chain. So, if I take a continuous-time Markov chain and if I look at the sojourn time or holding time of the process in a particular state, that is exponential. So, I can think of this how long it is being held in a particular phase as if it is being held in a particular state of a Markov chain. So, that is the view that we can take to generate an Erlang or to contribute to the alternative view of this Erlang distribution.

- Consider the following three-state continuous-time Markov chain (with an absorbing state at 3).



The associated Q matrix is $Q = \begin{bmatrix} -\mu & \mu & 0 \\ 0 & -\mu & \mu \\ 0 & 0 & 0 \end{bmatrix}$.

- Suppose that the process starts in state 1 (i.e. $\mathbf{p}(0) = (1, 0, 0)$) and let T be the time to absorption. Then, T follows an E_2 distribution with mean $\frac{2}{\mu}$ because
  ▶ The sojourn time in state 1 and state 2 are exponentials with mean $\frac{1}{\mu}$ each.
  ▶ T is the sum of these two IID exponentials with mean $\frac{1}{\mu}$ and hence E_2.

  So, this is an alternative view of this one, but here, if we have to stick to our definition of exponential, then we ideally have to write, I mean, take the rates as 2μ, 2μ here. So, the matrix will have, instead of μ, 2μ everywhere, and then this E_2 distribution will have then mean as $1/\mu$. So, that is what it is, but just to illustrate that you should not get fixed with that idea alone that you know we are given in general here even if it is μ, μ you can still call it Erlang no problem, but with the mean is different.

You have to be careful what is its mean and its variance or covariance; I mean coefficient of variation accordingly; this is a deviation; this is not the common one. The normal one will have 2μ here. So, all the rates will be 2μ's. So, this is because we want to illustrate in terms of CTMC. So, let us not have 2μ; instead of that, we could have mu itself to illustrate this purpose. So, this is what is an Erlang distribution.

So, here what we have done is we have used that the phase is exponential; we are using the fact that in a Markov chain, the residence time or the holding time or sojourn time of the process in a particular state is exponential. So, we are using that to construct a Markov chain with an absorbing state and the remaining as non absorbing states; they will not be recurrent in any case because the chain has to ultimately get should get absorbed into this absorbing state. So, we construct a Markov chain of that form and the corresponding Q matrix with an appropriate initial distribution; suppose if the initial distribution if you vary, you will get different distribution; you can make using that. But, this Erlang will always have $\mathbf{p}(0) = (1, 0, 0)$ as the initial distribution and $\begin{bmatrix} -\mu & \mu & 0 \\ 0 & -\mu & \mu \\ 0 & 0 & 0 \end{bmatrix}$ as the Q matrix of the Markov chain. So, the Q and $\mathbf{p}$ can be used appropriately to generate distribution. So, for this particular Q for this particular $\mathbf{p}$, you will get the Erlang distribution. Now the basic idea of generating a phase-type distribution is this. What is that?

- The phase-type distribution in question is defined as the time to absorption in a specified CTMC.

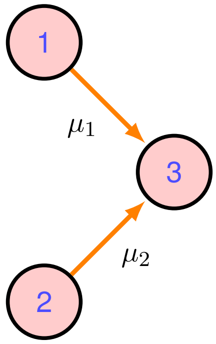
  So, you are defining the duration that follows a phase-type distribution as a time to absorption in a particular continuous-time Markov chain. So, you define appropriate CTMC.

- By choosing different CTMCs and by choosing different initial probability distributions, we can construct different phase-type distributions.

- These distributions are not exponential in general, but they can be analyzed using the CTMC theory, taking advantage of the underlying exponential distributions.

  So, this is the idea that a phase-type distribution is defined as the one corresponding to the time to absorption in a specified Markov chain, and by picking these Markov chains and by picking even with the initial distributions, we can construct different phase-type distributions that are the basic ideas of phase-type distributions.

- Let us now look at generating a hyperexponential distribution. How do we generate a hyper-exponential distribution? We know hyper-exponential distribution is really a mixture of two exponential distributions, two independent exponentials with parameters say mu_1 and mu_2, for example. Consider the following CTMC and its Q matrix.



$$Q = \left[\begin{array}{cc:c} -\mu_1 & 0 & \mu_1 \\ 0 & -\mu_2 & \mu_2 \\ \hdashline 0 & 0 & 0 \end{array}\right]$$

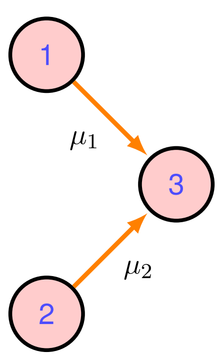
- Let the chain starts in state 1 with probability q and in state 2 with probability $1-q$, i.e., the initial state vector is $\mathbf{p}(0) = (q, 1-q, 0)$.

  We will always assume that the chain will not start in state 3, then you are not going to get anything; the time is 0. If it has to be represented as a time to absorption, it has to start in a non-absorbing state. So, that will be with some probability. So, that is why many a time this third quantity may not be given this $(q, 1-q)$ alone this will and this the non-absorbing states probabilities starting probabilities if you sum they will give you a proper distribution. So, this $\mathbf{p}(0)$ would be $(q, 1-q, 0)$ in this particular case.

- Then T, the time to absorption, follows a hyperexponential (H_2) distribution (being a mixture of two independent exponentials).
  ▶ Here with probability q, the time to absorption is exponential with mean $\frac{1}{\mu_1}$ and with probability $1-q$, the time to absorption is exponential with mean $\frac{1}{\mu_2}$.

  And this hyper-exponential distribution will be denoted by this notation H_k hyper-exponential the number of in this case you can think this as the phases. So, you are representing a hyper-exponential in terms of exponential. So, this is what we have here. So, this is another example or another using this Markov chain viewpoint that how one can generate using exponential, but so far, what we have seen these are two simple examples. Now, what we will do? We will give a formal approach as to how one can determine the distribution of time to absorption in a general CTMC. So, to illustrate that, we take this particular hyper-exponential example itself, and we will explain that process.

Now, suppose if I call



as the hyper-exponential Markov chain just for the sake of illustration.

- Then the forward Kolmogorov equations for the hyperexponential Markov chain are given by $\mathbf{p}'(t) = \mathbf{p}(t)\mathbf{Q}$ which when written explicitly are:

$$\begin{aligned} p_1'(t) &= -\mu_1 p_1(t) \\ p_2'(t) &= -\mu_2 p_2(t) \\ p_3'(t) &= \mu_1 p_1(t) + \mu_2 p_2(t) \end{aligned}$$

- Solving the first differential equation, we get $p_1(t) = qe^{-\mu_1 t}$, since $p_1(0) = q$.

  Similarly, solving the second differential equation, we get $p_2(t) = (1-q)e^{-\mu_2 t}$.

  Substituting in the third, we get

$$p_3'(t) = q\mu_1 e^{-\mu_1 t} + (1-q)\mu_2 e^{-\mu_2 t}.$$

- Note that, by definition, $p_3(t)$ is the CDF of T, i.e., $p_3(t) = P\{T \leq t\}$ and this means that $p_3'(t)$ is the PDF of T. The PDF given above is a H_2 distribution, as expected.

Now, **alternatively**, like here in this particular case,

$$\begin{aligned} p_1'(t) &= -\mu_1 p_1(t) \\ p_2'(t) &= -\mu_2 p_2(t) \\ p_3'(t) &= \mu_1 p_1(t) + \mu_2 p_2(t) \end{aligned}$$

is what is the case that we have. So, we have $\mathbf{p}'(t) = \mathbf{p}(t)\mathbf{Q}$ forward Kolmogorov equation from where you solve and to get $p_3^{'}(t)$ which is corresponding to the absorbing state to get the time to absorption of this. So, which means that basically, what you are doing, you are looking at the non-absorbing states here, which are 1 and 2, and you have

$$\begin{aligned} p_1'(t) &= -\mu_1 p_1(t) \\ p_2'(t) &= -\mu_2 p_2(t) \end{aligned}$$

system of equations that you are solving corresponding to the probabilities q and $1-q$. And once we have these expressions for $p_1^{'}(t)$ and $p_2^{'}(t)$, I can simply substitute into this the absorbing states the Kolmogorov equation to get the density of the time to absorption of this Markov chain; that is what you see here. Now, if you adapt this, if you write this generically, this is what you will have.

- Let $\tilde{\mathbf{p}}(t) = (p_1(t), p_2(t))$ (without the absorbing state) and let $\tilde{Q}$ be the corresponding matrix from Q. That is,

$$\tilde{Q} = \begin{bmatrix} -\mu_1 & 0 \\ 0 & -\mu_2 \end{bmatrix}, \quad \text{and} \quad \tilde{\mathbf{p}}'(t) = \tilde{\mathbf{p}}(t)\tilde{Q}.$$

- The solution to this matrix system of differential equations is

$$\tilde{\mathbf{p}}(t) = \tilde{\mathbf{p}}(0)e^{\tilde{Q}t}, \quad \text{where} \quad \tilde{\mathbf{p}}(0) = (q, 1-q),$$

where

$$
\begin{aligned}
e^{\tilde{Q}t} &= I + \tilde{Q}t + \frac{(\tilde{Q}t)^2}{2!} + \dots \\
&= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} -\mu_1 t & 0 \\ 0 & -\mu_2 t \end{pmatrix} + \begin{pmatrix} \mu_1^2 t^2/2 & 0 \\ 0 & \mu_2^2 t^2/2 \end{pmatrix} + \dots \\
&= \begin{pmatrix} e^{-\mu_1 t} & 0 \\ 0 & e^{-\mu_2 t} \end{pmatrix}.
\end{aligned}
$$

- We then obtain finally

$$\tilde{\mathbf{p}}(t) = \tilde{\mathbf{p}}(0)e^{\tilde{Q}t} = (qe^{-\mu_1 t}, (1-q)e^{-\mu_2 t}),$$

  which is the same as the one obtained earlier.

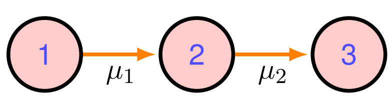
- The PDF of T is obtained as before using the last differential equation $p_3'(t) = \mu_1 p_1(t) + \mu_2 p_2(t)$.

- In general, a phase-type distribution is described using the parameters α and $\tilde{Q}$, where $\alpha = \tilde{\mathbf{p}}(0)$ and $\tilde{Q}$ as above (corresponding to the non-absorbing states).
  ♦ A phase-type distribution is denoted by the notation $PH(\alpha, \tilde{Q})$.

Exercise. *Apply this procedure to an appropriate E_2 distribution case and obtain $p_3'(t) = 4\mu^2 te^{-2\mu t}$.*

Remember now, in this particular case; we are now putting it as if this a proper the usual E 2 distribution where the mean is 1 by mu and the number of phases is k which means each of these phases will have the rate as k mu and hence here if it is it will be 2 mu that is what is the case. So, you can try this again using this process, whatever we have described just now here. Now, let us use this to describe or get to know another distribution which is called as hypo-exponential distribution. What is that? Which you have already seen to be the sum of independent exponential again, you are expressing this distribution in terms of exponential.

- Consider the following continuous-time Markov chain, where $\mu_1 \neq \mu_2$, and the system starts in state 1.



  The associated Q matrix is

$$Q = \left[\begin{array}{cc:c} -\mu_1 & \mu_1 & 0 \\ 0 & -\mu_2 & \mu_2 \\ \hdashline 0 & 0 & 0 \end{array}\right] \qquad \text{[with } \mathbf{p}(0) = (1,0,0)\text{]}$$

- The time to absorption is the convolution of two nonidentical exponential random variables.
  ♦ Note that this is not an Erlang, since an Erlang requires identical exponentials.

- One can derive $\mathbf{p}(t)$ in a similar way and get

$$\begin{aligned} p_1(t) &= e^{-\mu_1 t} \\ p_2(t) &= \frac{\mu_1}{\mu_2 - \mu_1} e^{-\mu_1 t} - \frac{\mu_1}{\mu_2 - \mu_1} e^{-\mu_2 t} \\ \text{and therefore} \quad p_3'(t) &= \mu_2 p_2(t) = \frac{\mu_1 \mu_2}{\mu_2 - \mu_1} e^{-\mu_1 t} - \frac{\mu_1 \mu_2}{\mu_2 - \mu_1} e^{-\mu_2 t}, \end{aligned}$$

  where $p_3'(t)$ is the probability density function (of the time to absorption) which is a two-term hypoexponential distribution.

- We now have a method for obtaining more general distributions and still use Markovian methods (and this can be of great help in many situations!).

  So, this is the continuous-time version, but again you could have you have a what you called as discrete phase-type distributions to handle the discrete distribution that may arise. But, since mainly we are concerned about continuous distribution, we did not mention it, but that is what is the case that you may have here. So, this is now even if it is, for example, if I change the initial distribution to be something else right instead of $(1, 0, 0)$, I can make say $(1/2, 1/2, 0)$ suppose then you will get another distribution, or the behaviour like you could have either go here, or it can go here with certain rates if you think all this form that you will get different CTMCs and it can give you different distributions which can quite get approximated in terms of phases. So, that is what is the advantage of using these phase-type distributions that one can still be within the Markovian framework when you are trying to analyze this distribution. What we will do is that we are not going to deal with phase-type distribution because the theory becomes much more complex though it mirrors the scalar quantities; now, you have to handle all matrices. So, what is called as quasi birth-death processes, $M/G/1$ type or $G/M/1$ type model these are all you can have matrix versions of that, but they will be very difficult to handle which we are not going to do it in this introductory course.

- We will now take up some Erlangian queueing systems to illustrate the ideas.

  Now, you can easily generalize this Erlangian analysis that we are doing how we are doing this Erlangian queues to say hypo-exponential, hyper-exponential or Coxian, or phase-type distributions in a similar way, but the complexity might increase, but in a similar way, this can be handled that is what you need to understand. So, to illustrate that what we will pick it up is that we will take up only the Erlangian case, and then we will talk about how one can analyze this Erlangian case this queues with Erlang distribution which will form the general exponential. And we already know Erlang itself can give a lot of flexibility in modelling when you are trying to use the general distributions for this, which we will take it up later.
  Thank you bye.