CS-417
COMPUTER SYSTEMS MODELING

Spring Semester 2020

Batch: 2016-17

(LECTURE # 23)

FAKHRA AFTAB
LECTURER
DEPARTMENT OF COMPUTER & INFORMATION SYSTEMS ENGINEERING
NED UNIVERSITY OF ENGINEERING & TECHNOLOGY

Recap of Lecture # 22

M/M/1/K Queue Analysis

Example Problems

Chapter # 6 (Cont'd)

FUNDAMENTALS OF QUEUING MODELS

QUEUING MODELS INVOLVING NONEXPONENTIAL DISTRIBUTIONS

- The assumption of exponential inter-arrival times implies that arrivals occur randomly (a Poisson input process).

- Furthermore, the actual service-time distribution frequently deviates greatly from the exponential form, particularly when the service requirements of the customers are quite similar.

- Therefore, it is important to have available other queuing models that use alternative distributions.

- Unfortunately, the mathematical analysis of queuing models with non-exponential distributions is much more difficult.

QUEUING MODELS INVOLVING NONEXPONENTIAL DISTRIBUTIONS

➢The *M/G*/1 Model

Assumptions

1. The queuing system has a *single server* and

2. A *Poisson input process* (exponential inter-arrival times) with a *fixed* mean arrival rate λ.

3. The customers have *independent* service times with the *same* probability distribution.

- In fact, it is only necessary to know (or estimate) the mean $1/\mu$ and variance σ^2 of this distribution.

QUEUING MODELS INVOLVING NONEXPONENTIAL DISTRIBUTIONS

- The readily available steady-state results for this general model are the following:

$$P_0 = 1 - \rho,$$

$$L_q = \frac{\lambda^2\sigma^2 + \rho^2}{2(1 - \rho)},$$

$$L = \rho + L_q,$$

$$W_q = \frac{L_q}{\lambda},$$

$$W = W_q + \frac{1}{\mu}.$$

MODELS INVOLVING NONEXPONENTIAL DISTRIBUTIONS

- The formula for L_q is one of the most important results in queuing theory because of
  - its ease of use and
  - the prevalence of $M/G/1$ queuing systems in practice.

- This equation for L_q (or its counterpart for W_q)
  - commonly referred to as **Pollaczek-Khintchine formula**,
  - named after two pioneers in the development of queuing theory
  - who derived the formula independently in the early 1930s.

- The model does not provide a closed-form expression for P_n because of analytical intractability.

MODELS INVOLVING NONEXPONENTIAL DISTRIBUTIONS

- For any fixed expected service time $1/\mu$, notice that L_q, L, W_q, and W all increase as σ^2 is increased.

- This result is important because it indicates that
  - the consistency of the server has a major bearing on the performance of the service facility —
  - not just the server's average speed.

- When the service-time distribution is exponential, $\sigma^2 = 1/\mu^2$, and the preceding results will reduce to the corresponding results for the $M/M/1$ model.

UNIFORM DISTRIBUTION

- X is a uniform random variable if the PDF of X is

$$f_X(x) = \begin{cases} \dfrac{1}{(b-a)} & a \leq x < b \\ 0 & otherwise \end{cases}$$

where the two parameters are **b > a**.

Theorem

- If X is a uniform random variable with parameters **a** and **b > a**.
- The CDF of X is

$$F_X(x) = \begin{cases} 0 & x \leq a \\ \dfrac{(x-a)}{(b-a)} & a < x \leq b \\ 1 & x > b \end{cases}$$

- The expected value of X is E[X] = **(b+a)/2**
- The variance of X is Var[X] = **$(b - a)^2/12$**

Example Problem 1

M/G/1

Consider the following single-server queue: the inter-arrival time is exponentially distributed with a mean of 10 minutes and the service time has the uniform distribution with a maximum of 9 minutes and a minimum of 7 minutes, find out:

(i) mean wait in the queue,

(ii) mean number in the queue,

(iii) the mean wait in the system,

(iv) mean number in the system and

(v) proportion of time the server is idle.

Answers

i) 1.602

ii) 16.02 mins

iii) 24.02 mins

iv) 2.402

v) 0.2

MODELS INVOLVING NONEXPONENTIAL DISTRIBUTIONS

The *M/D/s* Model

- The $M/D/s$ model often provides a reasonable representation for this kind of situation,
  - because it assumes that all service times actually equal some fixed *constant* (the *degenerate* service- time distribution) and
  - that we have a *Poisson* input process with a fixed mean arrival rate λ.

- When there is just a single server, the $M/D/1$ model is just the special case of the $M/G/1$ model where $\sigma^2 = 0$, so that the *Pollaczek-Khintchine formula* reduces to

$$L_q = \frac{\rho^2}{2(1-\rho)},$$

- where L, W_q, and W are obtained from L_q as just shown.

MODELS INVOLVING NONEXPONENTIAL DISTRIBUTIONS

- Notice that these L_q and W_q are exactly *half* as large as those for the exponential service-time case of the $M/M/1$ model, where $\sigma^2 = 1/\mu^2$, so decreasing σ^2 can *greatly* improve the measures of performance of a queuing system.

- For the multiple-server version of this model ($M/D/s$), a complicated method is available for deriving the steady-state probability distribution of the number of customers in the system and its mean [assuming $\rho = \lambda/(s\mu) < 1$].

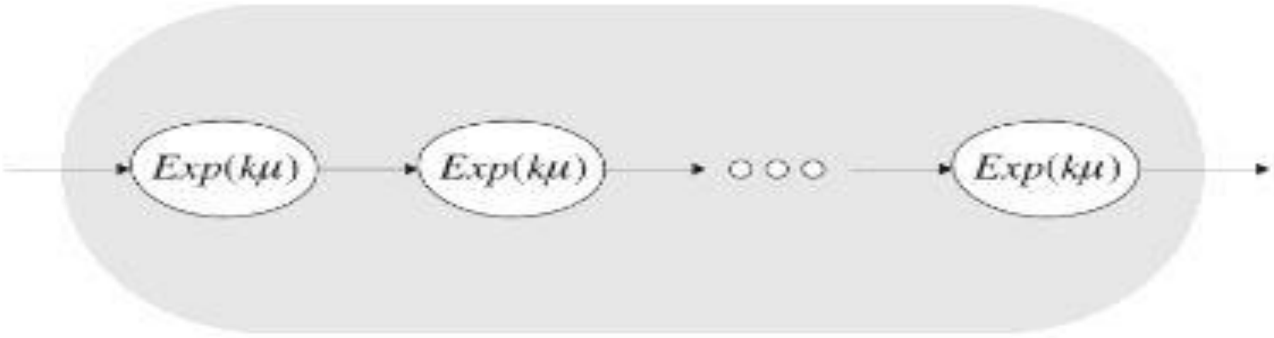
MODELS INVOLVING NONEXPONENTIAL DISTRIBUTIONS

- Notice that these L_q and W_q are exactly *half* as large as those for the exponential service-time case of the $M/M/1$ model, where $\sigma^2 = 1/\mu^2$, so decreasing σ^2 can *greatly* improve the measures of performance of a queuing system.

- For the multiple-server version of this model ($M/D/s$), a complicated method is available for deriving the steady-state probability distribution of the number of customers in the system and its mean [assuming $\rho = \lambda/(s\mu) < 1$].

Erlang-n Distribution

- Think of putting exponential distribution in series.
- If a random variable X is the sum of n-identical exponential random variables of service times n/μ, then X is said to have an Erlang-n distribution. Service time of a single server is given by 1/μ.
- The customer has to visit each stage in the facility to complete the service.



- A generalized Erlang Distribution is the sum of exponential random variables with different rates (also called a hypo-exponential distribution)

Erlang-n Distribution

- An Erlang random variable X with scale parameter α and n stages has probability density function:

$$f(x) = \frac{x^{n-1}e^{-x/\alpha}}{\alpha^n (n-1)!} \qquad x > 0.$$

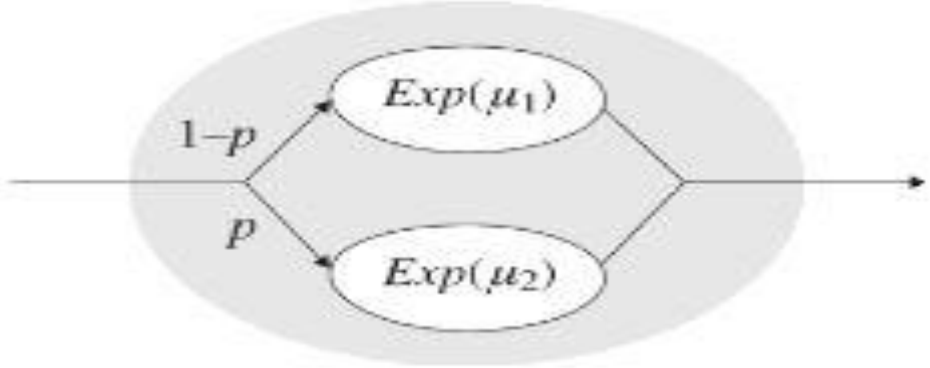
- The cumulative distribution function on the support of X is:

$$F(x) = P(X \le x) = 1 - \sum_{i=0}^{n-1} \frac{e^{-x/\alpha}x^n}{\alpha^n n!} \qquad x > 0.$$

- The population mean and variance are given by E[X] = nα V [X] = nα² respectively. α is given by 1/μ.

Hyper-Exponential Distribution

- Think of putting exponential distribution in parallel.



- A random variable X is hyper-exponentially distributed if X is with probability p_i , $i = 1, \ldots, k$ an exponential random variable X_i with mean $1/\mu_i$. For this random variable we use the notation $H_k(p_1, \ldots, p_k; \mu_1, \ldots, \mu_k)$, or simply H_k. The density is given by:

$$f(t) = \sum_{i=1}^{k} p_i \mu_i e^{-\mu_i t}, \qquad t > 0,$$

Hyper-Exponential Distribution

- The cumulative distribution function on the support of X is:

$$F(x) = P(X \leq x) = 1 - \sum_{i=1}^{n} p_i e^{-x/\alpha_i} \qquad x > 0.$$

α is given by 1/μ.