

Variability and Waiting Formulas

λ average arrival rate (customers per hour)

$1/\lambda$ average time between arrivals (interarrival time); $a = 1/\lambda$

μ average service rate (customers per hour)

$1/\mu$ average service time per customer; $p = 1/\mu$

Utilization = $\frac{\lambda}{\mu} = \frac{p}{a}$ when there is a single server.

Idle time = 1 – utilization

Coefficient of variation of the arrival process:

$$CV_a = \frac{\text{Standard deviation of interarrival time}}{\text{Average interarrival time}}$$

Coefficient of variation of the service process:

$$CV_p = \frac{\text{Standard deviation of service time}}{\text{Average service time}}$$

The coefficient of variation of an exponential distribution is equal to 1.

For a single server, when the utilization is less than 100%, the approximation formula for the steady-state Expected Waiting Time in Queue is

$$W_q = \text{Service time} \times \left(\frac{\text{Utilization}}{1 - \text{Utilization}} \right) \times \left(\frac{CV_a^2 + CV_p^2}{2} \right)$$

M/M/1 Formulas (require that $\lambda < \mu$):

$$\text{Utilization} = \frac{\lambda}{\mu}$$

In steady state:

$$\text{Expected waiting time in queue: } W_q = \frac{\lambda}{\mu(\mu - \lambda)}$$

$$\text{Expected waiting time in the system: } W = W_q + \frac{1}{\mu} = \frac{1}{\mu - \lambda}$$

$$\text{Expected number of customers in queue: } L_q = \frac{\lambda^2}{\mu(\mu - \lambda)}$$

$$\text{Expected number of customers in the system: } L = \frac{\lambda}{\mu - \lambda}$$

Little's Law: $L = \lambda W$

Determining important characteristics of queueing models if we know W_q

Expected number of customers in queue: $L_q = \lambda W_q$

Expected waiting time in the system $W = W_q + \frac{1}{\mu}$

Expected number of customers in the system: $L = \lambda W$

Multiple Servers

Suppose there are m identical servers

Single queue feeding first available server on customer first-come first-serve basis

λ is the arrival rate to the m -server system

Utilization = $\frac{\lambda}{m\mu}$

Approximation formula:

$$W_q = \left(\frac{\text{Service time}}{m} \right) \times \left(\frac{\text{Utilization}^{\sqrt{2(m+1)} - 1}}{1 - \text{Utilization}} \right) \times \left(\frac{CV_a^2 + CV_p^2}{2} \right)$$