Some General Results from QT

$W_i =$ Waiting time for $i$th arrival  
$S_i =$ Service time for $i$th arrival  
$A_i =$ Inter-arrival time between $(i-1)$th and $i$th arrival  
$\lambda =$ Average customer arrival rate  
$\mu =$ Average customer service rate  

Recursive relationship:

$$W_i = W_{i-1} + S_{i-1} - A_i$$

(1)

Little’s Law

Relates steady state mean system size to steady state average customer waiting times. Given by:

$$L = \lambda W$$

(2)

where $L$ is the steady state size of the system given by:

$$L = \lim_{T \to \infty} \frac{\int_0^T N(t)dt}{T}$$

(3)

and $W$ is the steady state waiting time in the system given by:

$$W = \lim_{m \to \infty} \frac{\int_0^T N(t)dt}{m}$$

(4)

where $N(t)$ is the number of customers in the system at time point $t$, which lies between 0 and $T$ and $m$ is the total number of customer arrivals between 0 and $T$. From (3) and (4), (2) can be derived given that the steady state value of $\lambda$ is:

$$\lambda = \lim_{T \to \infty, m \to \infty} \frac{m}{T}$$

(5)

Some general results from Little's Law in a G/G/c QS:

$$\rho = \frac{\lambda}{c \mu}$$

(6)

The system is in steady state when $\rho < 1$. 

Ergodicity

A stochastic process \( N(t) \) is Ergodic if with probability 1 it can be said that all its measures can be determined or well approximated from a single realization \( N_0(t) \) of the process. Every process of interest to us is ergodic. That is to say, that the time average of \( N(t) \) for a single run of the simulation will be equal to the ensemble average of \( N(t) \), the ensemble average being the average taken across an ensemble of simulation runs at steady state. It is in effect the steady state average.

So we can define the probability of \( N(t) = n \), i.e. there being \( n \) number of customers in the system at the time point \( t \) as follows:

\[
Pr\{N(t) = n\} = p_n(t)
\]

hence, we can say that for a G/G/c QS:

\[
L = E[N] = \sum_{n=0}^{\infty} np_n \quad \text{and} \quad L_q = E[N_q] = \sum_{n=c+1}^{\infty} (n-c)p_n
\]

where there \( L_q \) is the expected number of customers in queue and the number of servers is \( c \).

From Little’s Law we can say that \( L_q = \lambda W_q \) where \( W_q \) is the number of customers in queue. So we can say:

\[
L - L_q = \lambda(W - W_q) = \lambda E[S] = \frac{\lambda}{\mu} = r
\]

Here \( r \) is the load on the system (expected number of customers in service) at steady state and for \( c \) servers we can say that the load on each server is \( r/c \) which we denote by \( \rho \) - same as (6).

Also we can reduce (8) in the following way:

\[
L - L_q = E[N] - E[N_q] = \sum_{n=1}^{\infty} np_n - \sum_{n=c+1}^{\infty} (n-c)p_n = c \sum_{n=1}^{\infty} p_n = c(1 - p_0)
\]

where \( p_0 \) is the probability of there being 0 customers in the system. Combining (9) and (10) we get another general result:

\[
\rho = 1 - p_0
\]

In the next class we will discuss the nature of the customer arrival process and the inter-arrival rate.