A Brief Introduction to Queuing Theory

With an Application to Biological Processes

Prepared by:
Arthur Shih
Student, Chemical Engineering
University of Michigan

Prepared for:
Dan Burns Jr.
Professor, Mathematics
University of Michigan

Date Prepared:
December 14, 2009

Queuing Theory Background

Queuing theory is the mathematical approach of studying and analyzing waiting lines, also known as queues. Its primary usage was initially intended for studying queues in transportation, telephone traffic, commerce, services, etc. This section introduces the basic concepts of queuing theory and its useful function as models.

Three Components

There are three basic components of a waiting line: (1) arrivals, (2) servers, and (3) the queue. In general terms, arrivals require some form of service, servers serve arrivals and the queue is where arrivals wait for an open server. Figure 1 shows the simplest queuing model: a single queue.

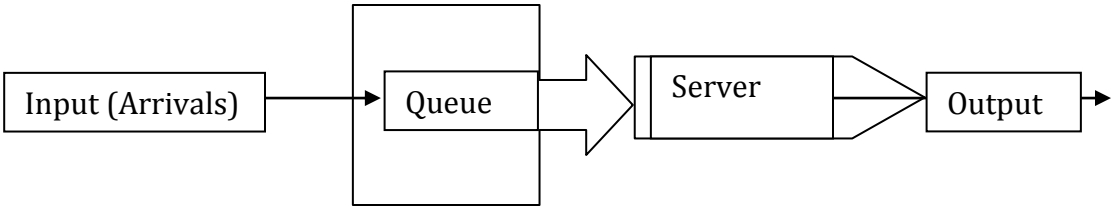

Figure 1: the simplest queue model: a single queue.

Factors

There are factors that affect each of the three components of a waiting line. It is necessary in engineering to identify and understand these factors in order to gain a broader understanding of any given queuing system. These factors, in turn, are able to help us determine optimal distribution models for each of the three components.
To illustrate these factors, a fast food joint will be used as an example.

Arrival Factors

An arrival distribution model tells us when arrivals are expected at certain times of the day. Arrivals are affected by lots of factors such as weather, politics, cleanliness of the restaurant, etc. For example, if customers in a fast food joint have not finished their meal, we do not expect them to come back and buy more while still eating; also, if there

is horrible weather, the arrival distribution model itself may change dramatically for that day.

Server Factors

A service distribution model tells us the expected service times at certain times of the day. This distribution depends on the type of service given, amount of service each arrival requests, the quality of the service, etc. In a fast food joint, the service distribution is dependent primarily on everything that takes place behind ordering counter. These include: number of ordering counters, organization of its service system, the productivity of employees working, etc.

Queue Factors

A queue distribution model tells us how long certain numbers of customers have to wait at certain times of the day. This too has factors that affect its distribution, but the biggest factors of the queue distribution depends primarily on the arrival distribution model and service distribution model. Other factors include arrival behavior and service priority. In a fast food joint, customers arriving tend to join the shortest line. Also, if customers wait too long, they may decide to leave.

Cost Factors

This component of the queuing system is dependent on arrival, server, and queue factors. Depending on how well the three components work for the system, costs can be driven up by poor efficiency where one or more components fail to be as efficient as the others. The system is the most efficient when costs are reduced to its minimum. This occurs when all services flow smoothly through the system at the same rate.

Single Queue Notation

The notation for a basic queue system is in the form A/B/m/n where:

A: arrival statistic

B: service time statistic

m: number of servers

n: number of customers (if the number of customers is assumed to be infinity, then this component is omitted)

Some common modeling choices for A and B are:

M: Markov/exponential

D: Deterministic timing

G: General (arbitrary)

Geom: Geometric wait time

M/M/1 Queues

The most common and basic queuing system is the M/M/1 queuing system. This "Markovian" queue (as shown in Figure 2) models with a Poisson arrival process and exponential service times.

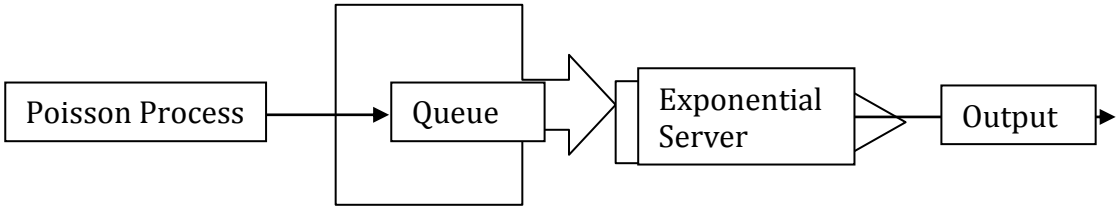

Figure 2: the M/M/1 queuing system

The M/M/1 queuing model contains Poisson process arrivals and exponentially distributed service times. These specifications of the M/M/1 process make it a reasonable model for a wide variety of situations because arrival processes of customers are often accurately described by a Poisson process.

Poisson Process Overview

Poisson Process Definitions

The Poisson Process is a purely random arrival process and can be thought of as a representation of a number of discrete arrivals. This number of distinct arrivals is described by the so called counter process N(t) which tells us the number of independent arrivals between the time increment (0,t). By altering the characteristics of the time increment, the Poisson process can be defined in three different, but equivalent ways:

1. Poisson process as a pure birth process;
   The probability of an arrival arriving within the infinitely small time interval *dt* is λdt. Given that all arrivals are independent, we can conclude that the probability of arriving within *dt* is independent of arrivals outside *dt*.

2. The number of arrivals N(t) in a finite interval of time t obeys the Poisson[λt] distribution, which allows us to calculate the probability of a certain number of arrivals within a time interval (0,t). Poisson[λt] is given by:

$$P[N(t) = n] = (\lambda t)^n/n!*e^{-\lambda t}$$

3. The arrival times between each arrival (interarrival time) are independent and obeys the Exp[λ] distribution, which allows us to calculate the probability of needing to wait at least t units of time for the next arrival. Exp[λ] is given by:

$$P[\text{interarrival time} > t] = e^{-\lambda t}$$

Two Useful Properties of Poisson Process

Superposition

The superposition of two Poisson processes with arrival rate of $\lambda_1, \lambda_2, \ldots, \lambda_n$ is a Poisson process with an arrival rate of $\lambda = \lambda_1 + \lambda_2 + \ldots + \lambda_n$.

Proof

It is given that the probability of independent Poisson arrivals are $\lambda_{1,\ldots,} \lambda_n$ where $n > 0$, and N(t) within the interval *dt* is $\lambda_1 dt + \ldots + \lambda_n dt$, respectively.

The superimposed probability of n Poisson processes is then $(\lambda_1 + \ldots + \lambda_n)$ *dt* from the distribution property. From this, we are able to conclude that the superposition of a Poisson process is simply the sum of the arrival rates, λ.

Figure 3 illustrates a Poisson superposition of 3 queues with independent probabilities of p_1, p_2, and p_3 where $p_1 + p_2 + p_3 = 1$

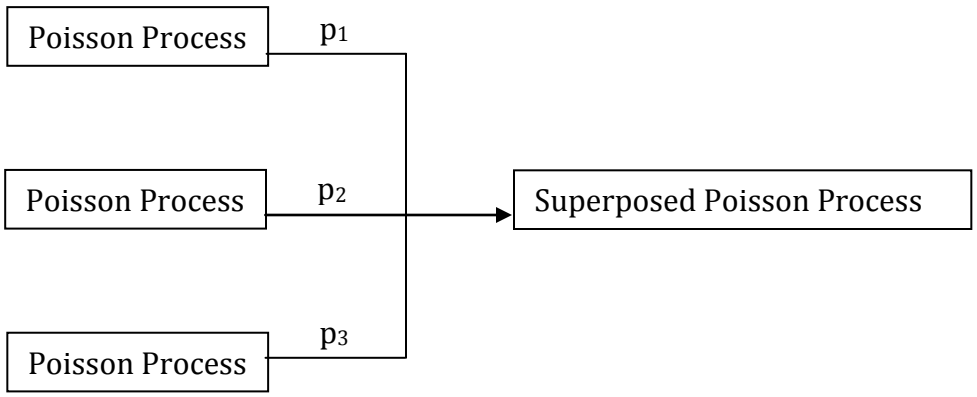

Figure 3: Poisson process superposition

Random Split

A Poisson process with arrival rate can be randomly split into multiple independent subprocesses with probabilities $p_1, p_2, \ldots, p_n$. Where $p_1 + p_2 + \ldots + p_n = 1$. Each of the resulting independent Poisson subprocesses will now have arrival rates of,

$$p_1 \lambda, p_2 \lambda, \ldots, p_n \lambda$$

Proof

Subprocesses represent a random selection of arrivals from the original Poisson process. Since the subprocesses are randomly chosen from points from a Poisson process, we can safely state that each subprocess is a Poisson process themselves arrival rates of λp_1, $\lambda p_2, \ldots, \lambda p_n$,

Let $N_n(I_n)$ = number of arrivals from subprocess n in interval I_n.

Also, denote $I = I_1 \cap I_2 \cap \ldots \cap I_n$.

We can write:

$$\begin{cases} N_1(I_2) = N_1(I) + N_1(I_1 \cap I_2^c \cap \ldots \cap I_n^c) \\ N_2(I_2) = N_2(I) + N_2(I_2 \cap I_1^c \cap \ldots \cap I_n^c) \\ \ldots \\ \ldots \\ \ldots \\ N_n(I_n) = N_n(I) + N_2(I_n \cap I_1^c \cap \ldots \cap I_{n-1}^c) \end{cases}$$

As shown above, the intervals $I_1 \cap I_2^c \cap \ldots \cap I_n^c$ to $I_n \cap I_1^c \cap \ldots \cap I_{n-1}^c$ are non-overlapping and therefore, independent.

This proves that each subprocess are Poisson processes with intensities $p_i \lambda$.

Figure 4 illustrates a queuing system with a random split into two queues.

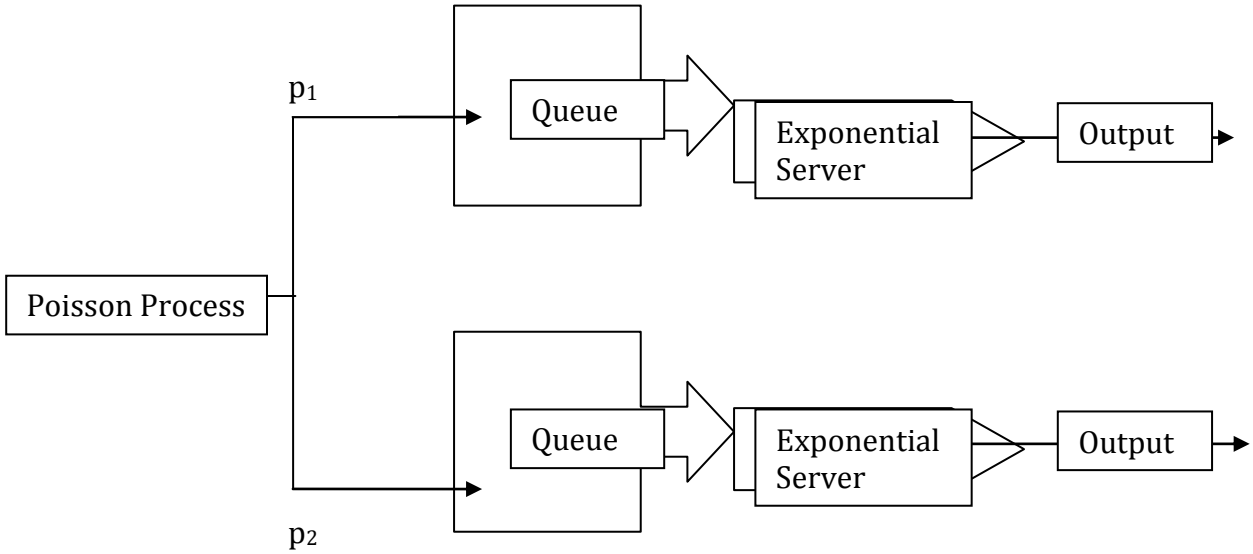

Figure 4: Random split of a Poisson process

An Application to a Biological Process

Case Study: Relationship between Insulin Level and Number of Insulin Receptors

Research by Cagin Kandemir-Cavas, Dokuz Eylul, University Department of Statistics
and Levent Cavas, Department of Chemistry, Biochemistry Division

Kandemir-Cavas and Cavas used the principles of Queuing theory to look into the chemical kinetics between insulin and insulin receptors. I present in this section the part of their research that uses most the introduction to queuing theory laid out in the first half of this paper.

Introduction to Insulin and Diabetes

Insulin is an important hormone that controls blood glucose concentrations in the body, and like most other hormones in the body, sufficient amounts are necessary for proper bodily function. Insulin helps convert glucose in the blood into glycogen, which is stored in the liver or muscles. Without insulin, the body will have difficulty utilizing glucose for energy. Given that carbohydrates, such as glucose, are the body's primary source of energy, insulin deficiencies pose a big problem. If the body is unable to store its primary energy source for future use, then the buy is said to have diabetes.

There are two types of diabetes: type 1 and type 2. Patients with type 1 diabetes produce no insulin due to the destruction of insulin-producing beta cells in the pancreas by the immune system. Patients with type 2 diabetes have beta cells that produce insulin normally, but the insulin receptors do not recognize the insulin. In turn, the produced insulin does not affect cells in the body. Both types of diabetes make the body unable to convert glucose to glycogen for storage.

About 7.8 % of the general population in the U.S. has diabetes, as stated by the National Diabetes Statistics 2007. Also, about 95% of all diabetes patients have type-2 diabetes., which can also be denoted as P(Type2|has diabetes). Using these two probabilities, we are able to put together a probability joint function, as shown in Table 1.

X / Y	Type 1 diabetes	Type 2 diabetes	Neither	Row Sum
Has diabetes	0.0741	0.0039	0	0.078
No diabetes	0	0	0.922	0.922
Column Sum	0.0741	0.0039	0.922	1

Table 1: Typical probabilities for the general U.S. population regarding diabetes. In this table, X=Type of diabetes, Y=Has diabetes or not.

Use of Queuing Theory in Biological Diabetes

In the past few years, statistical processes such as queuing theory have been used in clinical trials to reveal measures of an organ's function. In the case for insulin and diabetes, the Cava applied the M/M/c queue model to investigate the relationships between insulation concentration and the insulation receptor count. Cava's goal was to "use queuing theory to find the optimum number of insulin receptors and bring up the concept of metabolic energy balance and optimal energy use" (Cavas, 33). In other words, Cavas believes that in order to solve the problem of diabetes, one must look into and understand the source first.

The application of queuing theory to this kind of application requires us to redefine the meaning of some variables, as shown in Table 2. A schematic for the insulin queue system is shown in Figure 5.

Queuing System	Insulin System (units)
Cost	Energy value (ATP)
Number of servers	Number of insulin receptors (c)
Customers	Insulin level (μU/mL)
Arrival rate	Arrival rate of insulin (λ)
Service rate	Insulin-insulin receptor complex/time (μ)

Table 2: Components of a queuing system expressed for the Insulin System.

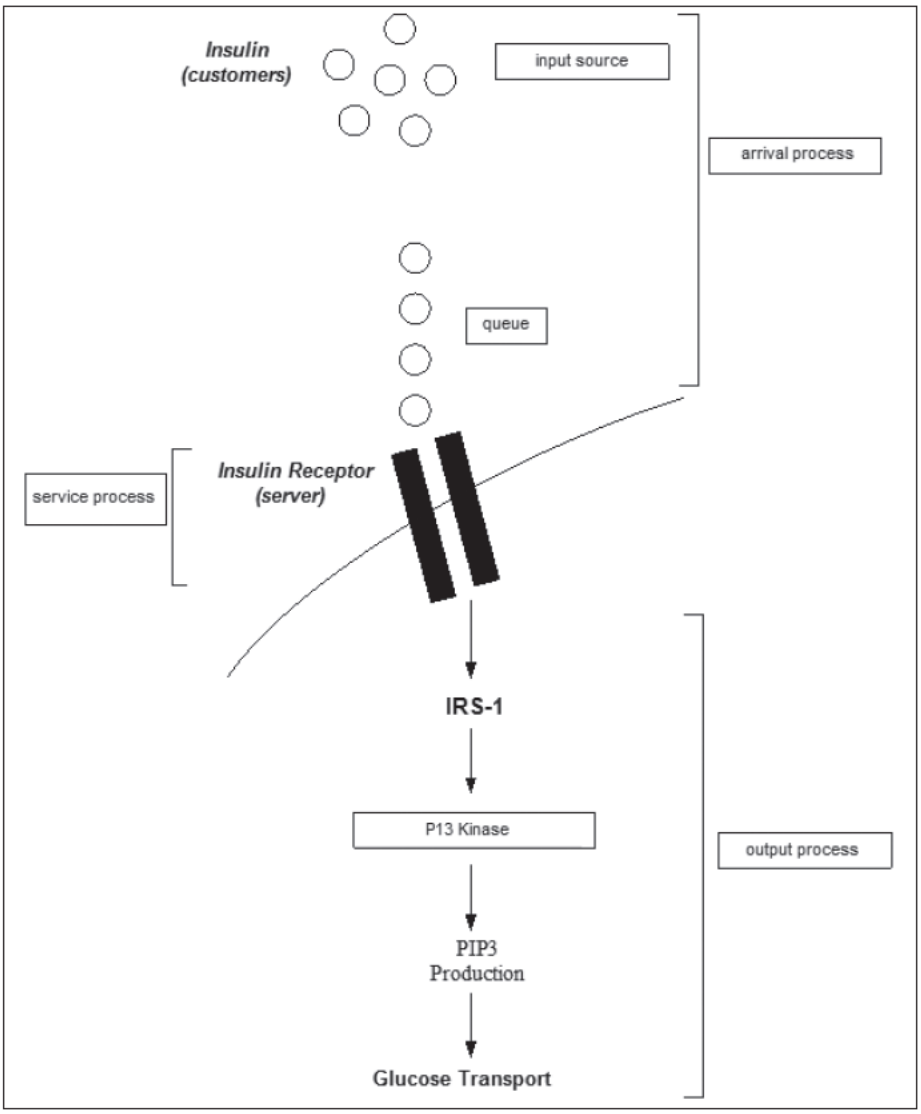

Figure 5: Schematic for the Insulin queuing system.

In the case for the insulin queuing system, the system would be the most efficient if the energy value were minimized, as discussed earlier in the paper.

The M/M/c queue model has a parameter of $1/\lambda$ for arrivals and a parameter of $1/\mu$ for the exponential distribution. The system has c servers and services by first-come, first-served. Also, we assume that the waiting line is infinitely long because the fourth component was omitted in the naming of this queuing model. It is assumed that the arrival rate does not affect the speed of the servers. Using these statements, a mathematical approach using queuing theory was taken up by Cavas.

Mathematical Approach

Service Rate

Given that the number of parallel receptors in the insulin queuing system is c, the service rate, μ, of the system can be expressed as:

$$\begin{cases} n\mu \ if \ n \leq c \\ c\mu \ if \ n > c \end{cases}$$

This is so because the service rate is limited by the number of parallel receptors, c.

Energy Model

An energy model can be calculated by taking the expectation of each of the processes consume energy in this process. In this system, two processes use energy: energy used in the production of insulin and energy used when insulin attaches onto an insulin receptor. Based off this, we are able to say that:

$$E(C_T(x)) = E(C_O(x)) + E(C_w(x))$$

Where $E(C_T(x))$ = expected total energy usage for each insulin receptor
$E(C_O(x))$ = expected amount of energy used due to bonding of insulin to insulin receptor
$E(C_w(x))$ = expected amount of energy used to overproduce extra insulin per unit time.

In this study, it was assumed that the amount of energy used when an insulin molecule comes into contact with an insulin receptor is 1. It was also assumed 2 ATP is used for each over-produced insulin.

Results

TORA software was used to calculate values displayed in Figure 6. From some literature search, Cavas gave μ and λ values of 12.3 μU/mL min and 6.6 μU/mL min, respectively. Using these values along with the assumptions given and calculated earlier, the TORA software was able to calculate the following parameters shown in Table 3.

Denotation	Description	Units
Lq	Expected amount of insulin level in queue	μU/mL
Ls	Expected amount of insulin level in system	μU/mL
Wq	Waiting time of insulin before service	Minutes
Ws	Total waiting time during service with receptor	Minutes
C	Total energy spent for insulin queuing system	# of ATP

Table 3: Parameters calculated by the TORA software.

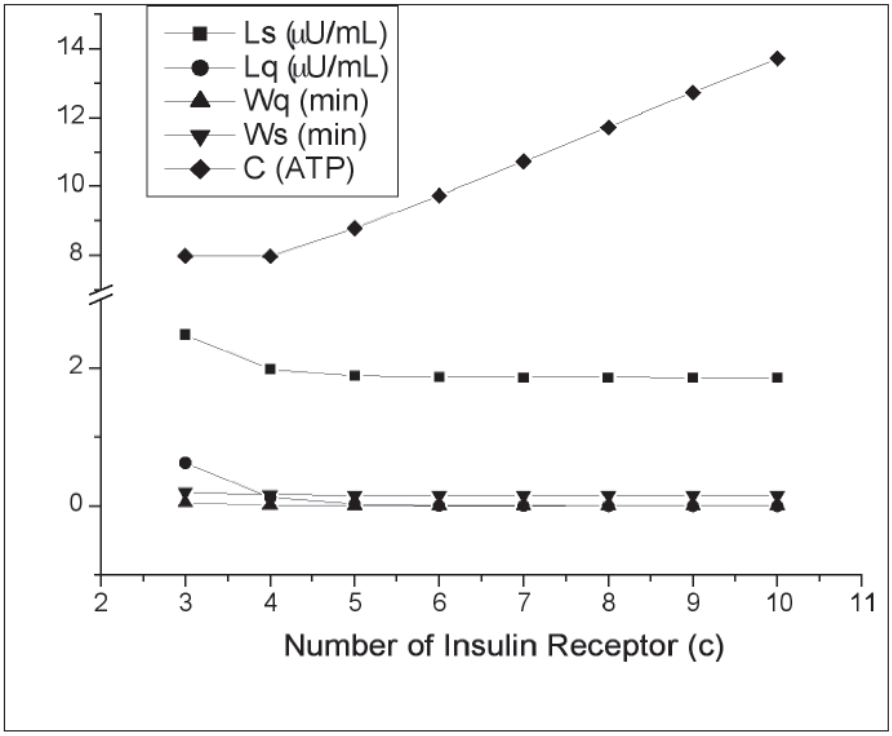

Figure 6: Relationship of several parameters vs. the number of insulin receptors.

Data Analysis

In the data generated by the TORA software, we notice that the amount of energy used (C) increases when the number of insulin receptors is greater than four. We also notice that the expected amount of insulin level in the system and expected number of insulin level in the queue decreases between three and four insulin receptors. Therefore, an insulin receptor number of 4 achieves both minimum wait time in the queue and minimum costs. Just to give this number some context, there are about 103 receptors on a typical red blood cell.

Resources

Baker, S.L. (2006). Queuing Theory 1. Retrieved (2009, December)

Calkins, K.G. (2005, May 10). *Queuing tTheory and the Poisson Distribution*. Retrieved from http://www.andrews.edu/~calkins/math/webtexts/prod10.htm

Kandermir-Cavas, C. (2006). An Application of queueing theory to the relationship between insulin level and number of insulin receptors. *Turkish Journal of Biochemistry*, 32(1), 32-38.

Ng, C.H. (2008). *Queueing Modelling Fundamentals with Applications in Communication Networks*. Chichester, England: Wiley.

Ross, S. (2009). *A First Course in Probability*. Upper Saddle River, NJ: Pearson Prentice Hall.