# 考慮服務壓力係數下 M/M/s/k 等候模式之建構與推導

**Consider the Construction and the Study that the Service Pressure Coefficient Descends the M/M/S/K Queuing Model**

藍俊雄[1]　郭美貝[2]

## 摘要

對很多企業來說，因為服務的消費與生產是同時發生的，所以服務需求的變動難以掌握。顧客通常是隨機地到達，並立即提出服務的需求，若服務能力在顧客到達已飽和，顧客便需排隊等候，由於到達率與服務時間需求是變動的，因而導致等候線的形成。若要在市場上保有競爭力，加快服務的速度將是重要的課題。對於第一線的服務人員來說，由於要直接與顧客面對面，尤其是當等待服務的人很多時，服務人員的壓力可能因此而改變。故本文欲透過以生死過程(Birth-and-death process)建構的 M/M/s/k 等候模式，推導出六項等候特徵方程式，並從壓力係數(Pressure coefficient)的變化進一步探討服務人員工作效率的變動。此外，一般研究在討論壓力情況下皆以無限等候容量來探討，顯然實務上是不合理的情況，因此本研究嘗試討論壓力情況下有上限的服務率與有負面的服務率加以討論，並將有限容量一併列入探討。

**關鍵字：** 等候理論、生死過程、壓力係數、有限容量

## Abstract

For lots of enterprises, the ministrant consumption and production exist simutaneously, so the fluctuation of service demand is hard to control. A customer usually randomly arrives, stands together and puts forward a ministrant need namely. If the service ability is already full, the customer then starts to wait. Because the arriving rate and service time is fluctuant, the waiting line happens. To speed up the ministrant speed which is an important topic, if you want to preserve the competition ability in the service market. A customer wants a direct service especially in whenever a lot of customers in wait , therefore the service pressure is then emerging. Thus, this article would like to develope the whole new characteristic equations under the considerations of service pressure for M/M/s/k model through the Birth-and-death process, and further inquired the fluctuation of the work efficiency from the various pressure coefficients. In addition, a common discussions are infinite waiting capacity and no pressure, obviously those situations are nonpratical. Therefore, this research tries to discuss the pressure condition and finite waiting capacity to make it near pratical.

[1]南華大學管理科學研究所教授
[2]南華大學管理科學研究所碩士班研究生

***Keywords*****:** Queuing Theory, Birth-and-death Process, pressure coefficient, finite capacity

## 1. 前言

由於社會結構與經濟環境急遽改變，人們面臨的壓力愈來愈大，鮮少有人願意把時間浪費在等候上，因此探討等待相關議題的迫切性與日俱增 Katz, et al. (1991)。舉凡在收銀機前等候結帳、加油站前等候加油、電影院前買票等候入場、等候大眾運輸工具、車輛經過高速公路收費站等，均是顯見的有形的等候線。無形的等候線如：工廠中等待修復的故障機器、等待計算機中央處理器的訊息、等待生產的訂單等，雖看不到有形的排隊隊伍，但實質上亦是一種典型的等候線。

目前大多的企業都非常強調提升自己的服務品質，減少顧客的等候時間。因此大多的企業便選擇了增加自己的服務設施；例如量販店增加結帳櫃檯、收銀員及收銀機，工廠增加生產線上的設備。這種方式固然可以縮短顧客的等候時間，不過同時也增加了企業的支出。所以要如何在改善服務品質的同時也要注意因等候所增加的成本，從中取得均衡便是成為等候線分析中所要討論的重點問題之ㄧ。

而等候理論是在 1910 年左右由丹麥的數學家兼電腦工程師 A.K.Erlang 首先提出來的，他當時是研究電話線路忙碌與閒置的問題。他想要知道線路忙碌（或佔線）的期望值，因此利用機率論上的一些假設做為分析的條件，於 1917 年發表一篇「自動電話交換機機率理論問題之解答」，發展成一套分析和研究排隊等候的工具，用來解決電話線路擁擠問題，奠定了等候理論的基礎。由於當時的社會經濟結構較為單純，對此理論的應用亦較少見。直到第二次世界大戰後，等候理論才被廣泛及大量加以應用於實際問題中。例如生活中常碰到的問題，機場班機等候起飛、港口之容量、門診排隊等候時間、高速公路收費站、速食店服務櫃臺、銀行收付櫃臺等等，都可以透過等候理論的分析做為決策的依據。

Hiller and Lieberman (2002)提及，等候理論是在研究各種不同狀況的等候，它使用等候模式(Queuing Model)來表示實務上各種類型的等候系統（包括某些種類的等候線），各模式的公式指出相對應的等候系統將如何進行，包括在各種不同環境下將發生的平均等待時間。

廖慶榮(1994)指出，等候理論(Queuing Theory 或 Waiting Line Theory)是研究等候系統(Queuing System)之問題，只要有人或物等待接受服務，等候系統就自然形成，而其共通現象就是等。

林照雄(1977)於「作業研究」一書中提到，等候線可用其可能之長度來作為表示，一般可分為無限長度等候線(Infinite Queue)及有限長度等候線(Finite Queue)二種。前者在等候線上之數量沒有限制，因此要求服務數量雖多，排列長度雖長，假如服務機構之能力，可以應付顧客需求，仍視為合理等候線。但當服務機構無法應付顧客需求時，此

時必須設法增加處理站來改善，以維持顧客合理之等候系統。在有限等候線系統，等候線顧客有一定限制，若超出此項限制，後來之顧客必須要離開等候線。

大部分基本的等候模式都假設系統的輸入（到達的顧客）和輸出（離開的顧客），是遵從生死過程(birth-and-death process)，機率理論中這個重要的過程在多種不同的領域都有應用案例。但是在等候理論中，名詞「生」是指一位新顧客到達等候系統，而死是指一位已完成服務的顧客離開。而 M/M/s/k 模式即為其一。

而對於第一線的服務人員來說，由於要直接與顧客面對面，尤其是當等待服務的人很多時，臨櫃人員的壓力亦會隨著等待的人數增加而產生變化，然而有些人在壓力情況下反而會使服務率下降,此類問題鮮有被論及。此外，一般研究在討論壓力情況下皆以無限等候容量來探討，但實際上等候的空間並非無限大，很顯然在實務上欠缺合理性。基於上述情況，本研究主要目的概述如下：

(1) 運用等候理論建構壓力考量下之服務率函數。

(2) 推導壓力情況下 M/M/s/k 之六項特性方程式。

(3) 討論壓力情況下有上限的服務率與有負面的服務率，並將有限容量一併列入探討。

## 2. 假設及符號說明

### 2.1 研究假設

本研究之基本研究假設如下：

(1) 等候系統擁有 $s$ 個服務人員，系統最多容納 $k$ 位顧客，且到達時間間隔與服務時間均服從指數分配。

(2) 任何到達的顧客發現等候線是滿的，就拒絕進入到系統並且永遠離開。

(3) 因系統容量有限，系統中有 $n$ 位顧客時的平均到達率為

$$\lambda_n = \begin{cases} \lambda & , n = 0, 1, \ldots, k-1 \\ 0 & , n \geq k \end{cases}$$

(4) 因系統中有 $s$ 個服務者，所以最多可有 $s$ 位顧客同時接受服務，但當顧客務 $n$ 小於 $s$ 時，僅能有 $n$ 位顧客接受服務，其餘 $(s-n)$ 個服務者則為閒置狀態。當有 $n$ 位顧客在系統時的平均服務率為

$$\mu_n = \begin{cases} n\mu & , n = 0, 1, \ldots, s \\ \left[\dfrac{n(s+1)}{s(n+1)}\right]^{\alpha} s\mu & , n = s+1, \ldots, k \end{cases}$$

其中 $\alpha$ 為壓力係數。吾人考慮此一函數關係是因為其滿足三個基本條件：

(1) 當顧客人數 $n$ 無限增加時，系統服務率 $\mu_n$ 之極限值為一有限數值 $(s+1)^{\alpha} s^{1-\alpha} \mu$ 。亦即

$$\lim_{n\to\infty}\left[\frac{n(s+1)}{s(n+1)}\right]^{a} s\mu = (s+1)^{a} s^{1-a}\mu$$

(2) 當顧客人數 $n$ 等於服務人員數 $s$ 時，系統服務率 $\mu_n$ 之值等於數值 $s\mu$。

(3) 當壓力係數 $\alpha$ 等於 0 時，有 $n$ 位顧客在系統時的平均服務率為

$$\mu_n = \begin{cases} n\mu \quad , \quad n = 0, 1, \ldots, s \\ s\mu \quad , \quad n = s+1, \ldots, k \end{cases}$$

即為不考慮壓力係數時的平均服務率。

### 2.2 符號說明

茲將本研究所建構之數學模式符號說明如下：

$s$ ：櫃臺最大數目

$k$ ：最大容忍系統人數

$\mu_n$ ：有 $n$ 個人到達此系統時之服務率

$n$ ：系統中的人數

$\alpha$ ：壓力係數

$P_o$: 系統內沒有人的機率

$P_n$: 系統有 $n$ （n≧1）位顧客機率

$\lambda_n$: 系統中有 $n$ 個人時之顧客到達率

$L_q$: 平均系統中在等候的人數

$L$: 平均系統中的人數（包含被服務者）

$W_q$: 每位顧客在等候線上的期望等候時間

$W$: 每位顧客在系統中的期望等候時間

## 3. $M/M/s/k$ 等候模式建構

本文討論的等候空間是有限制的，亦即在系統中的顧客數目是不允許超過某個特定的數目（在此以 $k$ 表示），以致於等候線的容量為 $(k-s)$。任何到達的顧客發現等候線是滿的，就拒絕進入到系統並且永遠離開。從生死過程的觀點來看，系統中有 $n$ 位顧客時的平均到達率為

$$\lambda_n = \begin{cases} \lambda & , n = 0, 1, \ldots, k-1 \\ 0 & , n \geq k \end{cases}$$

而當有 $n$ 位顧客在系統時的平均服務率為

$$\mu_n = \begin{cases} n\mu & , \quad n = 0,1,\ldots, s \\ \left[\dfrac{n(s+1)}{s(n+1)}\right]^a s\mu & , \quad n = s+1,\ldots, k \end{cases}$$

其中 $\alpha$ 為壓力係數。為了簡化符號，令

$$c_n = \frac{\lambda_{n-1}\cdots\lambda_0}{\mu_n\cdots\mu_1}, \quad n = 1, 2, \ldots, k$$

吾人可推得

$$c_n = \frac{\lambda_{n-1}\cdots\lambda_0}{\mu_n\cdots\mu_1}$$

$$= \frac{\lambda^n}{\left\{\left[\dfrac{n(s+1)}{s(n+1)}\right]^a s\mu\right\}\left\{\left[\dfrac{(n-1)(s+1)}{sn}\right]^a s\mu\right\}\cdots\left\{\left[\dfrac{(s+1)(s+1)}{s(s+2)}\right]^a s\mu\right\}\cdot s\mu\cdot(s-1)\mu\cdots 2\mu}$$

$$= \frac{\lambda^n}{s^{n-s}s!\mu^n\left[\dfrac{(s+1)}{s}\right]^{a(n-s)}\left\{\dfrac{n}{n+1}\cdot\dfrac{n-1}{n}\cdots\dfrac{s+1}{s+2}\right\}^a}$$

$$= \frac{\rho^n s^{a(n-s)}(n+1)^a}{s^{(n-s)}s!(s+1)^{a(n-s+1)}}$$

其中 $\rho = (\lambda/\mu)$。若對 $n = 0$ 定義 $c_n = 1$，可得系統在狀態 $n$ 時的穩定狀態機率為

$$P_n = c_n P_0$$

$$= \begin{cases} \dfrac{\rho^n}{n!}P_0 & , n = 1, \cdots, s \\ \dfrac{\rho^n(n+1)^a}{s^{(1-a)(n-s)}s!(s+1)^{a(n-s+1)}}P_0 & , n = s+1, \cdots, k \end{cases}$$

其中

$$P_0 = \left[ \sum_{n=0}^{s} \frac{\rho^n}{n!} + \sum_{n=s+1}^{k} \frac{\rho^n (n+1)^a}{s^{(1-a)(n-s)} s! (s+1)^{a(n-s+1)}} \right]^{-1}$$

從以上求得之穩定狀態機率 $P_n$，等候線的期望長度(不包括正在被服務的顧客) $L_q$ 為

$$L_q = \sum_{n=s+1}^{k} (n-s) P_n$$

$$= \sum_{n=s+1}^{k} (n-s) \frac{\rho^n (n+1)^a}{s^{(1-a)(n-s)} s! (s+1)^{a(n-s+1)}} P_0$$

$$= \frac{\rho^s P_0}{s!(s+1)^a} \left\{ \sum_{n=s+1}^{k} (n-s) \frac{\rho^{(n-s)} [(n-s)+(s+1)]^a}{s^{(1-a)(n-s)} (s+1)^{a(n-s)}} \right\}$$

$$= \frac{\rho^s P_0}{s!(s+1)^a} \sum_{i=1}^{k-s} i \frac{\rho^i (i+s+1)^a}{s^{(1-a)i} (s+i)^{ai}}$$

$$= \frac{\rho^s P_0}{s!(s+1)^a} \sum_{i=1}^{k-s} i (i+s+1)^a \left[ \frac{\rho}{s^{(1-a)} (s+1)^a} \right]^i \tag{3.1}$$

而在等候系統內的期望顧客數為

$$L = \sum_{n=0}^{k} n P_n$$

$$= \sum_{n=0}^{s} n P_n + \sum_{n=s+1}^{k} n P_n$$

$$= \sum_{n=0}^{s} n P_n + \sum_{n=s+1}^{k} (n-s) P_n + \sum_{n=s+1}^{k} s P_n$$

$$= \sum_{n=0}^{s} n P_n + L_q + s \left( 1 - \sum_{n=0}^{s} P_n \right)$$

$$= L_q + \sum_{n=0}^{s} n \frac{\rho^n}{n!} P_0 + s \left( 1 - \sum_{n=0}^{s} \frac{\rho^n}{n!} P_0 \right)$$

$$= L_q + P_0 \sum_{n=1}^{s} \frac{\rho^n}{(n-1)!} + s - s P_0 \sum_{n=0}^{s} \frac{\rho^n}{n!}$$

$$=L_q+P_0\rho\sum_{n=1}^{s}\frac{\rho^n}{(n-1)!}+s-sP_0\sum_{n=0}^{s}\frac{\rho^n}{n!}$$

$$=L_q+P_0\rho\sum_{n=0}^{s-1}\frac{\rho^n}{n!}+s-sP_0\sum_{n=0}^{s}\frac{\rho^n}{n!}$$

$$=L_q+P_0\rho\left(\sum_{n=0}^{s}\frac{\rho^n}{n!}-\frac{\rho^s}{s!}\right)+s-sP_0\sum_{n=0}^{s}\frac{\rho^n}{n!}$$

$$=L_q+s+P_0(\rho-s)\sum_{n=0}^{s}\frac{\rho^n}{n!}-P_0\frac{\rho^{s+1}}{s!} \qquad (3.2)$$

在一個穩定狀態的等候過程，由於有下列關係式(Little's formula)

$L=\lambda W$ 與 $L_q=\lambda W_q$

因此，每一位顧客在系統內及在等候線上的時間

$$W=\frac{L}{\lambda}$$

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

可藉由帶入公式(3.1)與(3.2)之 $L$ 與 $L_q$ 值求得。

## 4. 結論

Hiller 針對壓力情況所建構的系統平均服務速率的模式中，考慮了壓力係數。然而，當壓力愈大時，$\mu_n$ 愈大，亦即服務率會越大。此種情況並不合理顧客數增加，服務率 $\mu_n$ 不可能會無限增加。因此，本文提出一個極限值存在的合理模型，並透過生死過程證明 $M/M/s/k$ 等候模式之六項特徵方程式。經由 $M/M/s/k$ 壓力下的等候理論推論研究得知，

(1) 系統內沒有人的機率

$$P_0=\left[\sum_{n=0}^{s}\frac{\rho^n}{n!}+\sum_{n=s+1}^{k}\frac{\rho^n(n+1)^a}{s^{(1-a)(n-s)}s!(s+1)^{a(n-s+1)}}\right]^{-1}$$

(2) 系統有 $n$ （n≧1）位顧客機率為

$$P_n = \begin{cases} \dfrac{\rho^n}{n!} P_0 & , n = 1, \cdots, s \\ \dfrac{\rho^n (n+1)^a}{s^{(1-a)(n-s)} s! (s+1)^{a(n-s+1)}} P_0 & , n = s+1, \cdots, k \end{cases}$$

(3) 平均系統中在等候線上的人數為

$$L_q = \frac{\rho^s P_0}{s!(s+1)^a} \sum_{i=1}^{k-s} i(i+s+1)^a \left[ \frac{\rho}{s^{(1-a)}(s+1)^a} \right]^i$$

(4) 平均系統中的人數（包含被服務者）為

$$L = L_q + s + P_0(\rho - s) \sum_{n=0}^{s} \frac{\rho^n}{n!} - P_0 \frac{\rho^{s+1}}{s!}$$

(5) 每位顧客在系統中的期望等候時間為：

$$W = \frac{L}{\lambda} = \frac{1}{\lambda} \left[ L_q + s + P_0(\rho - s) \sum_{n=0}^{s} \frac{\rho^n}{n!} - P_0 \frac{\rho^{s+1}}{s!} \right]$$

(6) 每位顧客在等候線上的期望等候時間為：

$$W_q = \frac{L_q}{\lambda} = \frac{1}{\lambda} \left\{ \frac{\rho^s P_0}{s!(s+1)^a} \sum_{i=1}^{k-s} i(i+s+1)^a \left[ \frac{\rho}{s^{(1-a)}(s+1)^a} \right]^i \right\}$$

## 參考文獻


1. 林照雄(1997)，「作業研究」，台北：三民書局。
2. 葉桂珍(1992)，「數量方法」，台北：三民書局。
3. 楊創發(2000)，「門診等候時間縮短之研究—運用行動通訊技術」，國立中正大學資訊管理學系碩士論文。
4. 楊朝欽(1996)，「醫院掛號作業效率之決策研究-等候理論之應用」，大葉大學事業經營研究所碩士論文。
5. 廖慶榮(2005)，「作業研究」，台北：華泰書局。
6. David R. Anderson, Dennis J. Sweeney, Thomas A. Williams (2003),“An Introduction to Management Science Quantitative Approaches to Decision Making”, *Thomson South-Western.*
7. Fitzsimmons, James and Mona Fitzsimmons (1998), “Service Management: Operations, Strategy, and Information Technology” , *Boston: Irwin /McGraw Hill. Chapt. 11: Managing Queues.*

8. Frederi ck S. Hiller, Gerald J. Lieberman (2001), "Introduction to Operations Research", *McGraw Hill Higher Education*.
9. Katz, K. L., Larson, B.M., and Larson, R.C. (1991), "Prescription of the Waiting-in line Blues: Enlighten and engage ", *Sloan Management Review*, 32(4), pp.44-53.