



排队模型

凯里学院
余英

模型要点

- 1、掌握排队模型的基本概念
- 2、了解常见的分布函数及生灭过程
- 3、掌握典型排队系统模型的结构及应用

排队模型的基本概念

一、引言

1、什么是排队模型（排队论）？

排队论是研究拥挤现象的一门学科。

它是在研究各种排队系统概率规律性的基础上，解决有关排队系统的最优化设计（静态）和最优控制（动态）问题。

现实生活中的排队系统

序号	到达的顾客	要求服务内容	服务机构
1	不能运转的机器	修理	修理技工
2	修理技工	领取修配零件	发放修配零件的管理员
3	病人	诊断或做手术	医生(或包括手术台)
4	电话呼唤	通话	交换台
5	文件搞	打字	打字员
6	提货单	提取存货	仓库管理员
7	驶入港口的货船	装(卸)货	装(卸)货码头(泊位)
8	上游河水进入水库	放水, 调整水位	水闸管理员

2、排队论的起源与应用领域

- 1)、20世纪初Bell电话公司为减少用户呼叫，研究电话线路合理配置问题；
- 2)、1909年丹麦工程师A.K.Erlang受热力学统计平衡概念启发发表论文《概率论与电话交换》，解决上述问题；
- 3)、应用于：通讯系统、交通运输、机器维修、库存控制、计算几设计等领域。

二、排队系统的特征及其组成

- 1、排队系统的特征即拥挤现象的共性
 - 1)、有请求服务的人或物
 - 2)、有为顾客服务的人或物
 - 3)、具有随机性
 - 4)、服务的数量超过服务机构的容量

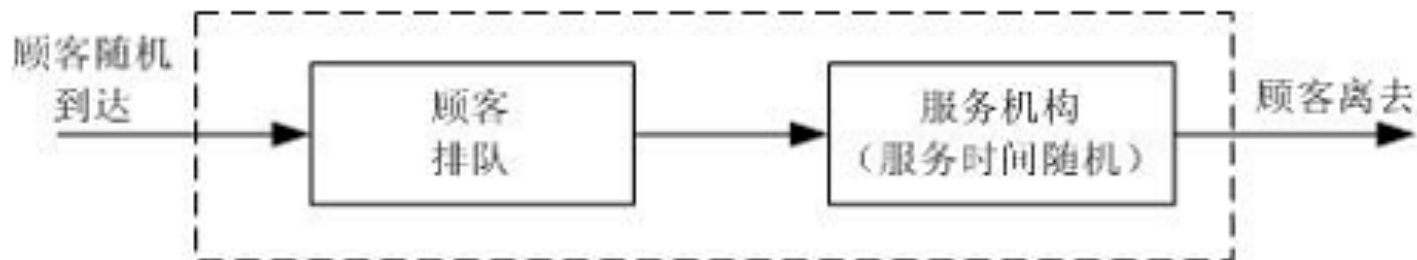


图1 排队模型

2、排队系统的三大基本组成部分

1)、输入过程(顾客到达的方式)

- a、顾客的总体（顾客源）的组成可能是有限的，也可能是无限的；
- b、顾客相继到达的时间间隔可以是确定的，也可以是随机的，对于随机的情形，要知道单位时间内的顾客到达数或相继到达的间隔时间的概率分布；
- c、输入过程可以是平稳的（描述相继到达的间隔时间分布和所含参数（如期望值、方差等）都是与时间无关的），否则成为非平稳的，我们研究平稳的。

2、排队系统的三大基本组成部分

2)、排队规则

- a、顾客到达时，如所有服务台都被占用，在这种情形下，顾客可以随即离去，也可以排队等待，前者成为损失制，后者成为等待制，我们研究后者；其次还有混合制，它是介于等待制和损失制之间的；
- b、从占有的空间来看，有的系统要规定容量（即允许进入排队系统的顾客数）的最大限，有的没有这种限制

2、排队系统的三大基本组成部分

3)、服务过程

- a、可以是没有服务员，单个的，多个的，对于多个的，它们之间可以是平行排列（并列）的，也可以是前后排列（串列）的，也可以是混合的；
- b、服务时间可以是确定的，也可以是随机的，对于后者要知道它的概率分布；
- c、服务时间可以是平稳的，也可以是非平稳的，我们研究前者；
- d、对于等待制，服务规则又可以分为先到先服务（**FCFS**），后到先服务（**LCFS**），随机服务和有优先权的服务。

三、排队模型的分类（符号表示）

我们采用Kendall记号

顾客相继到达时间间隔分布/服务时间分布/服务台数目/排队系统允许的最大顾客容量（系统容量）/顾客总体数量（顾客源数量）/排队规则

说明：如果Kendall记号中略去后3项，表示 $x/y/z/\infty/\infty/\text{FCFS}$
相继到达时间间隔和服务时间分布的符号如下：

M——负指数分布

D——确定型

E_k——**k**阶爱尔朗分布

GI——一般相互独立的时间间隔分布

G——一般服务时间分布

四、排队模型的数量指标

1、平均队长(L_s)：指在系统中的顾客数（包括正被服务的顾客和排队等待的顾客）的期望值。

2、平均排队长(L_q)：指系统中排队等候服务的顾客数的期望值。

$$L_s = L_q + \text{正被服务的顾客数}$$

3、平均逗留时间(W_s)：指一个顾客在系统中的停留时间期望值。

4、平均等待时间(W_q)：指一个顾客在系统中排队等待的时间的期望值。

$$W_s = W_q + \text{服务时间}$$

5、忙期：指从顾客到达空闲服务机构起到服务机构再次空闲止这段时间长度，即服务机构连续繁忙的时间长度。

6、系统的状态概率 $[P_n(t)]$ ：指系统中的顾客数为 n 的概率。

7、稳定状态： $\lim_{t \rightarrow \infty} P_n(t) \rightarrow P_n$

四、排队模型的数量指标

- 8、 λ_n —— 系统有 n 个顾客时的平均到达率
- 9、 μ_n —— 系统有 n 个顾客时的平均服务率
- 10、 λ —— 对任何 n 都是常数的平均到达率
- 11、 μ —— 对任何 n 都是常数的平均服务率
- 12、 ρ —— 服务强度，或称使用因子，平均到达率与服务台与平均服务率的乘积的比值
- 13、系统的状态——系统中的顾客数，如果系统中有 n 个顾客，就说系统的状态是 n ，系统的状态是随着时间在变化的
- 14、 $p_n(t)$:时刻 t 系统状态为 n 的概率，稳态时系统状态为 n 的概率用 p_n 表示。

五、常见的分布函数及生灭过程

■ 1、poisson流

定义：设 $N(t)$ 为时间 $[0, t]$ 内到达系统的顾客数，如果满足下面三个条件：

a、平稳性：在 $[t, t+\Delta t]$ 内有一个顾客到达的概率为 $\lambda \Delta t + o(\Delta t)$ ；

b、独立性（无后效性）：任意两个不相交区间内顾客到达情况相互独立；

c、普遍性：在 $[t, t+\Delta t]$ 内多于一个顾客到达的概率为 $o(\Delta t)$ ；

则称 $\{N(t), t \geq 0\}$ 为poisson流。

■ 2、poisson分布

设 $N(t)$ 为时间 $[0, t]$ 内到达系统的顾客数，则 $\{N(t), t \geq 0\}$ 为poisson流的充要条件是：

$$p\{N(t) = n\} = \frac{(\lambda t)^n}{n!} e^{-\lambda t} (n = 1, 2, \dots)$$

五、常见的分布函数及生灭过程

■ 3、负指数分布

定理：设 $N(t)$ 为时间 $[0, t]$ 内到达系统的顾客数，则 $\{N(t), t \geq 0\}$ 为参数为 λ 的poisson流的充要条件是：相继到达时间间隔服从相互独立的参数为 λ 的负指数分布。

■ 4、k阶爱尔朗分布

设 v_1, v_2, \dots, v_k 是 k 个相互独立的随机变量，服从相同参数 $k\mu$ 的负指数分布，那么

$$T = v_1 + v_2 + \dots + v_k$$

服从 k 阶爱尔朗分布。

五、常见的分布函数及生灭过程

■ 5、生灭过程

定义：设 $\{N(t), t \geq 0\}$ 为一随机过程，若 $N(t)$ 的概率分布具有以下性质：

a、假设 $N(t) = n$, 则从时刻 t 起到下一个顾客到达时刻止的时间服从参数为 λ_n 的负指数分布, $n=0,1,2,\dots$

b、假设假设 $N(t) = n$, 则从时刻 t 起到下一个顾客离去时刻止的时间服从参数为 μ_n 的负指数分布, $n=0,1,2,\dots$

c、同一时刻时只有一个顾客到达或离去。
则称 $\{N(t), t \geq 0\}$ 为一个生灭过程。

五、常见的分布函数及生灭过程

生灭过程中 C_n 与 p_0 的推导及应用

根据系统平稳状态时“流入=流出”原理，得到如下任一状态下的平衡方程：

$$0 \quad \mu_1 p_1 = \lambda_0 p_0$$

$$1 \quad \lambda_0 p_0 + \mu_2 p_2 = (\lambda_1 + \mu_1) p_1$$

$$2 \quad \lambda_1 p_1 + \mu_3 p_3 = (\lambda_2 + \mu_2) p_2$$

$$\dots \quad \dots$$

$$n-1 \quad \lambda_{n-2} p_{n-2} + \mu_n p_n = (\lambda_{n-1} + \mu_{n-1}) p_{n-1}$$

$$n \quad \lambda_{n-1} p_{n-1} + \mu_{n+1} p_{n+1} = (\lambda_n + \mu_n) p_n$$

$$\dots \quad \dots$$

五、常见的分布函数及生灭过程

- 由上述方程可求得

$$0 \quad p_1 = p_0 \lambda_0 / \mu_1$$

$$1 \quad p_2 = \lambda_1 p_1 / \mu_2 + (\mu_1 p_1 - p_0 \lambda_0) / \mu_2 \\ = p_0 \lambda_0 \lambda_1 / (\mu_2 \mu_1)$$

$$2 \quad p_3 = \lambda_2 p_2 / \mu_3 + (\mu_2 p_2 - p_1 \lambda_1) / \mu_3 \\ = p_0 \lambda_2 \lambda_1 \lambda_0 / (\mu_3 \mu_2 \mu_1)$$

...

$$n-1 \quad p_n = \lambda_{n-1} p_{n-1} / \mu_n + (\mu_{n-1} p_{n-1} - p_{n-2} \lambda_{n-2}) / \mu_n \\ = p_0 \lambda_{n-2} \lambda_{n-1} \dots \lambda_0 / (\mu_n \mu_{n-1} \dots \mu_1)$$

$$n \quad p_{n+1} = \lambda_n p_n / \mu_{n+1} + (\mu_n p_n - p_{n-1} \lambda_{n-1}) / \mu_{n+1} \\ = p_0 \lambda_n \lambda_{n-1} \dots \lambda_0 / (\mu_{n+1} \mu_n \dots \mu_1)$$

五、常见的分布函数及生灭过程

- 记 $c_n = \frac{\lambda_{n-1}\lambda_{n-2}\dots\lambda_0}{\mu_n\mu_{n-1}\dots\mu_0} (n = 1, 2, \dots)$

- 则平稳状态的分布为 $p_n = c_n p_0$ 。由此可得 $p_0 = \frac{1}{1 + \sum_{n=1}^{\infty} c_n}$
- 生灭过程排队系统的各项指标，即

$$l = \sum_{n=0}^{\infty} n p_n, l_q = \sum_{n=c}^{\infty} (n - c) p_n, w = \frac{l}{\lambda_e}, w_q = \frac{l_q}{\lambda_e}$$

其中 λ_e 是整体平均到达率

五、常见的分布函数及生灭过程

6、经验分布

例1 某服务机构单服务台，先到先服务，对41顾客记录到达时刻 τ 和服务时间 s （单位：分钟）如下表，表中第1号顾客到达时刻为0。全部服务时间为127（分钟）。

(1)	(2)	(3)	(4)	(5)	(1)	(2)	(3)	(4)	(5)	(1)	(2)	(3)	(4)	(5)
i	τ_i	s_i	t_i	w_i	i	τ_i	s_i	t_i	w_i	i	τ_i	s_i	t_i	w_i
1	0	5	2	0	5	12	2	7	10	9	36	1	2	0
2	2	7	4	3	6	19	4	3	5	10	38	2	7	0
3	6	1	5	6	7	22	3	4	6	11	45	5	2	0
4	11	9	1	2	8	26	3	10	5	12	47	4	2	3

(1)	(2)	(3)	(4)	(5)	(1)	(2)	(3)	(4)	(5)	(1)	(2)	(3)	(4)	(5)
i	τ_i	s_i	t_i	w_i	i	τ_i	s_i	t_i	w_i	i	τ_i	s_i	t_i	w_i
13	49	1	3	5	23	86	6	2	2	33	117	4	4	7
14	52	2	9	3	24	88	5	4	6	34	121	2	6	7
15	61	1	1	0	25	92	1	3	7	35	127	1	2	3
16	62	2	3	0	26	95	3	6	5	36	129	6	1	2
17	65	1	5	0	27	101	2	4	2	37	130	3	3	7
18	70	3	2	0	28	105	2	1	0	38	133	5	2	7
19	72	4	8	1	29	106	1	3	1	39	135	2	4	10
20	80	3	1	0	30	109	2	5	0	40	139	4	3	8
21	81	2	2	2	31	114	1	2	0	41	142	1		9
22	83	3	3	2	32	116	8	1	0					

到达间隔分布表

到达间隔 (分钟)	次数
1	6
2	10
3	8
4	6
5	3
6	2
7	2
8	1
9	1
10以上	1
合计	40

服务时间分布表

服务时间 (分钟)	次数
1	10
2	10
3	7
4	5
5	4
6	2
7	1
8	1
9以上	1
合计	41

平均间隔时间：
 $=142/40=3.55(\text{分钟/人})$

平均服务时间：
 $127/41=3.12(\text{分钟/人})$

平均到达率：
 $41/142=0.28(\text{人/分钟})$

平均服务率：
 $41/127=0.32(\text{人/分钟})$

六、典型排队系统模型的结构及应用

■ M/M/C等待制排队模型研究要点：

- a、系统意义
- b、状态转移速度图与状态转移速度矩阵
- c、状态概率方程
- d、系统的基本数量指标

Passion分布

设 $N(t)$ 表示在时间 $[0, t)$ 内到达顾客数;

令 $P_n(t_1, t_2)$ 表示在时间区间 $[t_1, t_2)$ ($t_2 > t_1$) 内有 n (≥ 0) 个顾客到达的概率, 即

$$P_n(t_1, t_2) = P\{N(t_2) - N(t_1) = n\} \quad (t_2 > t_1, n \geq 0)$$

Passion分布的三条件:

(1) 无后效性: 不相重叠的时间区间内顾客到达数相互独立

$$(2) \quad P_1(t, t + \Delta t) = \lambda \Delta t + o(\Delta t)$$

$$(3) \quad \sum_{n=2}^{\infty} P_n(t, t + \Delta t) = o(\Delta t)$$

情况	$[0, t)$		$[t, t+\Delta t)$		$[0, t+\Delta t)$	
	个数	概率	个数	概率	个数	概率
(A)	n	$P_n(t)$	0	$1-\lambda\Delta t+o(\Delta t)$	n	$P_n(t) (1-\lambda\Delta t+o(\Delta t))$
(B)	$n-1$	$P_{n-1}(t)$	1	$\lambda\Delta t$	n	$P_{n-1}(t)\lambda\Delta t$
(C)	$n-2$	$P_{n-2}(t)$	2	} $o(\Delta t)$	n	} $o(\Delta t)$
	$n-3$	$P_{n-3}(t)$	3		n	
	\dots	\dots	\dots		\dots	
	0	$P_0(t)$	n		n	

在上述条件下，研究顾客到达数 n 的概率分布

$$P_n(t+\Delta t) = P_n(t) (1-\lambda\Delta t+o(\Delta t))$$

$$+ P_{n-1}(t)\lambda\Delta t$$

$$+ o(\Delta t)$$

$$P_n(t+\Delta t) = P_n(t)(1-\lambda\Delta t) + P_{n-1}(t)\lambda\Delta t + o(\Delta t)$$

$$[P_n(t+\Delta t) - P_n(t)]/\Delta t = -\lambda P_n(t) + \lambda P_{n-1}(t) + [o(\Delta t)]/\Delta t$$

$$\text{令 } \Delta t \rightarrow 0$$

$$\begin{cases} dP_n(t)/dt = -\lambda P_n(t) + \lambda P_{n-1}(t) \\ P_n(0) = 0 \end{cases} \quad (n \geq 1)$$

$$\begin{cases} dP_0(t)/dt = -\lambda P_0(t) \\ P_0(0) = 1 \end{cases} \quad (n=0)$$

$$\therefore P_0(t) = e^{-\lambda t}$$

$$P_n(t) = [(\lambda t)^n e^{-\lambda t}] / n! \quad t > 0, n = 0, 1, 2, \dots$$

$$E[N(t)] = \lambda t; \quad \text{Var}[N(t)] = \lambda t.$$

负指数分布

$$f_T(t) = \begin{cases} \lambda e^{-\lambda t}, & t \geq 0 \\ 0, & t < 0 \end{cases}$$

$$\begin{aligned} E(T) &= \int_0^{\infty} t \lambda e^{-\lambda t} dt = - \int_0^{\infty} t e^{-\lambda t} d(-\lambda t) = - \int_0^{\infty} t d e^{-\lambda t} \\ &= -[t e^{-\lambda t}]_0^{\infty} - \int_0^{\infty} e^{-\lambda t} dt = \int_0^{\infty} e^{-\lambda t} dt = - \frac{1}{\lambda} \int_0^{\infty} e^{-\lambda t} d(-\lambda t) \\ &= - \frac{1}{\lambda} e^{-\lambda t} \Big|_0^{\infty} = \frac{1}{\lambda} \end{aligned}$$

$$\begin{aligned} V(T) &= E(T^2) - [E(T)]^2 = \int_0^{\infty} t^2 \lambda e^{-\lambda t} dt - \frac{1}{\lambda^2} \\ &= -[t^2 e^{-\lambda t}]_0^{\infty} - \int_0^{\infty} 2t e^{-\lambda t} dt - \frac{1}{\lambda^2} = 2 \int_0^{\infty} t e^{-\lambda t} dt - \frac{1}{\lambda^2} \\ &= \frac{2}{\lambda^2} - \frac{1}{\lambda^2} = \frac{1}{\lambda^2} \end{aligned}$$

第三节 单服务台负指数分布排队系统的分析

一、M/M/1 模型

1、假设

- (1) 顾客到达的间隔时间满足参数为 λ 的负指数分布
- (2) 服务时间满足参数为 μ 的负指数分布 ($\lambda < \mu$)
- (3) 服务机构是单服务台
- (4) 顾客源是无限的，顾客相互独立
- (5) 单队排列，且对队长没有限制

2、 P_n 的计算

情况	在时刻 t 顾客数	在区间 $(t, t+\Delta t)$		在时刻 $t+\Delta t$ 顾客数
		到达	离去	
(A)	n	×	×	n
(B)	$n+1$	×	0	n
(C)	$n-1$	0	×	n
(D)	n	0	0	n

0表示发生（1个），×表示没有发生

$$\begin{aligned}P_n(t+\Delta t) = & P_n(t)(1-\lambda\Delta t)(1-\mu\Delta t) \\& + P_{n+1}(t)(1-\lambda\Delta t)\mu\Delta t \\& + P_{n-1}(t)\lambda\Delta t(1-\mu\Delta t) \\& + P_n(t)\lambda\Delta t\mu\Delta t\end{aligned}$$

整理得：

$$P_n(t+\Delta t)=P_n(t)(1-\lambda\Delta t-\mu\Delta t)+P_{n+1}(t)\mu\Delta t+P_{n-1}(t)\lambda\Delta t+o(t)$$

$$[P_n(t+\Delta t)-P_n(t)]/\Delta t=\lambda P_{n-1}(t)+\mu P_{n+1}(t)-(\lambda+\mu)P_n(t) \quad (1)$$

$$\Delta t \rightarrow 0 \Rightarrow dP_n(t)/dt=\lambda P_{n-1}(t)+\mu P_{n+1}(t)-(\lambda+\mu)P_n(t)$$

考虑 $P_0(t)$ 的情况：

$$P_0(t+\Delta t)=P_0(t)(1-\lambda\Delta t)+P_1(t)(1-\lambda\Delta t)\mu\Delta t$$

$$\Delta t \rightarrow 0 \Rightarrow dP_0(t)/dt=-\lambda P_0(t)+\mu P_1(t) \quad (2)$$

由 $dP_n(t)/dt=0$ 得到

$$\begin{cases} -\lambda P_0+\mu P_1=0 & (3) \\ \lambda P_{n-1}+\mu P_{n+1}-(\lambda+\mu)P_n=0 & (4) \end{cases}$$

由式(3)得 $P_1 = \frac{\lambda}{\mu} P_0$

通过求解可得 $P_n = \left(\frac{\lambda}{\mu}\right)^n P_0, \quad n = 0, 1, 2, \dots$

$$\sum_{n=1}^{\infty} P_n = P_0(1 + \rho + \rho^2 + \dots) = P_0 \frac{1}{1 - \rho} = 1 \quad \rho = \frac{\lambda}{\mu} < 1$$

$$P_0 = 1 - \rho$$

$$P_n = \rho^n (1 - \rho), \quad n \geq 1$$

参数意义:

λ —— 单位时间内到达的平均顾客数

μ —— 单位时间内服务的平均顾客数

ρ —— 服务强度

3、M/M/1参数计算

(1) 系统中平均顾客数 (L_s)

$$\begin{aligned} L_s &= \sum_{n=0}^{\infty} nP_n = \rho(1-\rho) + 2\rho^2(1-\rho) + 3\rho^3(1-\rho) + \dots \\ &= (1-\rho)(\rho + 2\rho^2 + 3\rho^3 + \dots) \end{aligned}$$

$$\left. \begin{aligned} \text{记 } S &= \rho + 2\rho^2 + 3\rho^3 + \dots \\ \rho S &= \rho^2 + 2\rho^3 + 3\rho^4 + \dots \end{aligned} \right\} \Rightarrow$$

$$(1-\rho)S = \rho + \rho^2 + \rho^3 + \dots = \frac{\rho}{1-\rho}$$

$$L_s = \frac{\rho}{1-\rho} = \frac{\lambda}{\mu - \lambda}$$

(2) 队列中等待的平均顾客数 (L_q)

$$L_q = \sum_{n=1}^{\infty} (n-1)P_n = \sum_{n=1}^{\infty} nP_n - \sum_{n=1}^{\infty} P_n = L_s - \rho = \frac{\lambda^2}{\mu(\mu - \lambda)}$$

(3) 顾客逗留时间 (W_s)

$$W_s = E[W] = \frac{1}{\mu - \lambda}$$

(4) 队列中顾客等待时间 (W_q)

$$W_q = W_s - \frac{1}{\mu} = \frac{\lambda}{\mu(\mu - \lambda)}$$

Ls: 平均顾客数

Ws: 顾客逗留时间

Lq: 队列中等待的平均顾客数

Wq: 顾客等待时间

它们的相互关系如下:

$$\begin{aligned} L_s &= \lambda W_s, & L_q &= \lambda W_q \\ W_s &= W_q + \frac{1}{\mu}, & L_s &= L_q + \frac{\lambda}{\mu} \end{aligned}$$

其中 $L_s = \lambda W_s$, $L_q = \lambda W_q$

称为**little**公式, 它是排队论中的一个重要公式。

例3 100个工作小时内每小时来就诊的病人数 n 出现次数如下

到达的病 人数 n	出现次数 t_n
0	10
1	28
2	29
3	16
4	10
5	6
6	1
合计	100

100个完成手术的病例所用时间 v (小时)出现的次数如下

为病人完成手术 时间 v (小时)	出现次数 t_v
0.0-0.2	38
0.2-0.4	25
0.4-0.6	17
0.6-0.8	9
0.8-1.0	6
1.0-1.2	5
1.2以上	0
合计	100

解：

$$(1) \text{病人平均到达率} = \frac{\sum n f_n}{100} = 2.1 (\text{人/小时})$$

$$\text{每次手术平均时间} = \frac{\sum v f_v}{100} = 0.4 (\text{小时/人})$$

$$\text{每小时完成手术人数 (平均服务率)} = \frac{1}{0.4} = 2.5 (\text{人/小时})$$

$$(2) \text{取 } \lambda = 2.1, \quad \mu = 2.5 \qquad \text{则 } \rho = \frac{\lambda}{\mu} = \frac{2.1}{2.5} = 0.84$$

$$(3) L_s = \frac{2.1}{2.5 - 2.1} = 5.25 (\text{人})$$

$$L_q = 0.84 \times 5.25 = 4.41 (\text{人})$$

$$W_s = \frac{1}{2.5 - 2.1} = 2.5 (\text{小时})$$

$$W_q = \frac{0.84}{2.5 - 2.1} = 2.1 (\text{小时})$$

二、M/M/1/N/∞ 模型

假定系统最大容量为N，单服务台情形排队等待的顾客最多为N-1，下面只考虑稳态情形：

$$\begin{cases} \mu P_1 = \lambda P_0 \\ \mu P_{n+1} + \lambda P_{n-1} = (\lambda + \mu) P_n, & n \leq N-1 \\ \mu P_N = \lambda P_{N-1} \end{cases}$$

解得：

$$P_0 = \frac{1 - \rho}{1 - \rho^{N+1}} \quad \rho \neq 1$$

$$P_n = \frac{1 - \rho}{1 - \rho^{N+1}} \rho^n \quad n \leq N$$

根据上式我们可以推导出系统的各项指标:

$$(1) \text{ 队长 } L_s = \sum_{n=0}^N n P_n = \frac{\rho}{1-\rho} - \frac{(N+1) \rho^{N+1}}{1-\rho^{N+1}}, \quad \rho \neq 1$$

$$(2) \text{ 队列长 } L_q = \sum_{n=1}^N (n-1) P_n = L_s - (1-P_0)$$

有效到达率 $\lambda_e = \lambda(1-P_N)$

可以验证: $1-P_0 = \lambda_e / \mu$

带入公式

$$(3) \text{ 顾客逗留时间 } W_t = \frac{L_s}{\mu (1-P_0)} = \frac{L_q}{\lambda (1-P_N)} + \frac{1}{\mu}$$

$$(4) \text{ 顾客等待时间 } W_q = W_s - \frac{1}{\mu}$$

例4 单人理发馆有六个椅子接待客人。当6个椅子都坐满时，后来的顾客不进店就离开。顾客平均到达率为3人/小时，理发需时平均15分钟。则：

$N=7$ 为系统中最大的顾客数， $\lambda=3$ 人/小时， $\mu=4$ 人/小时

(1) 求某顾客一到达就能理发的概率。

相当于没有顾客，所求 概率为

$$P_0 = \frac{1 - 3/4}{1 - (3/4)^8} = 0.2778$$

(2) 求需要等待的顾客数的期望值。

$$L_s = \frac{3/4}{1 - 3/4} - \frac{8(3/4)^8}{1 - (3/4)^8} = 2.11$$

$$L_q = L_s - (1 - P_0) = 2.11 - (1 - 0.2778) = 1.39$$

(3) 求有效到达率。

$$\lambda_e = \mu(1 - P_0) = 4(1 - 0.2778) = 2.89 \text{ (人/小时)}$$

(4) 求一顾客在理发馆内逗留的时间。

$$W_s = L_s / \lambda_e = 2.11 / 2.89 = 0.73 \text{ 小时} = 43.8 \text{ 分钟}$$

(5) 在可能到达的顾客中有百分之几不等待就离开。

$$P_7 = \left(\frac{\lambda}{\mu}\right)^7 \left(\frac{1 - \lambda / \mu}{1 - (\lambda / \mu)^8}\right) = \left(\frac{3}{4}\right)^7 \left(\frac{1 - \frac{3}{4}}{1 - \left(\frac{3}{4}\right)^8}\right) \approx 3.7\%$$

第四节 多服务台指数分布排队系统的分析

一、M/M/c

规定各服务台工作相互独立且平均分配服务率相同, 即

$$\mu_1 = \mu_2 = \dots = \mu_c = \mu$$

整个服务机构的平均服务率为
$$\begin{cases} c\mu, & (n \geq c) \\ n\mu, & (n < c) \end{cases}$$

令 $\rho = \frac{\lambda}{c\mu}$, 只有当 $\frac{\lambda}{c\mu} < 1$ 时才不会排成无限的队列. 称它为

这个系统的服务强度, 或服务机构的平均利用率.

$$\begin{cases} \mu P_1 = \lambda P_0 \\ (n+1)\mu P_{n+1} + \lambda P_{n-1} = (\lambda + n\mu)P_n, & (1 \leq n < c) \\ c\mu P_{n+1} + \lambda P_{n-1} = (\lambda + c\mu)P_n, & (n \geq c) \end{cases}$$

这里 $\sum_{i=0}^{\infty} P_i = 1$, 且 $\rho \leq 1$.

用递推法解上述差分方程，可求得状态概率。

$$P_0 = \left[\sum_{k=0}^{c-1} \frac{1}{k!} \left(\frac{\lambda}{\mu} \right)^k + \frac{1}{c!} \cdot \frac{1}{1-\rho} \cdot \left(\frac{\lambda}{\mu} \right)^c \right]^{-1}$$

$$P_n = \begin{cases} \frac{1}{n!} \left(\frac{\lambda}{\mu} \right)^n P_0 & (n < c) \\ \frac{1}{c! c^{n-c}} \left(\frac{\lambda}{\mu} \right)^n P_0 & (n \geq c) \end{cases}$$

根据上式我们可以推导出系统的各项指标：

$$L_s = L_q + \frac{\lambda}{\mu}, \quad L_q = \sum_{n=c+1}^{\infty} (n-c) P_n = \frac{(c\rho)^c \rho}{c(1-\rho)^2} P_0$$

$$\text{因为 } \sum_{n=c+1}^{\infty} (n-c) P_n = \sum_{n'=1}^{\infty} n' P_{n'+c} = \sum_{n'=1}^{\infty} \frac{n'}{c! c^{n'}} (c\rho)^{n'+c} P_0 = \text{右边}$$

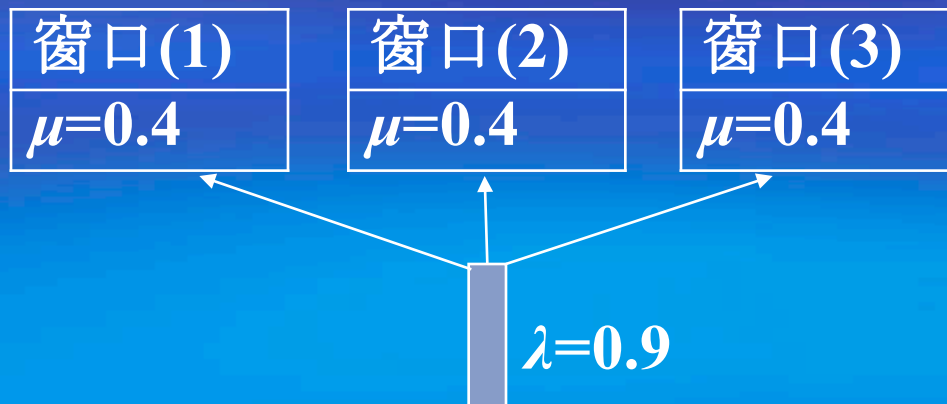
$$W_q = \frac{L_q}{\lambda}, \quad W_s = \frac{L_s}{\lambda}$$

例6 某售票所有三个窗口，顾客到达服从Passion过程，平均到达率每分钟 $\lambda=0.9$ (人)，服务(售票)时间服从负指数分布，平均服务率每分钟 $\mu=0.4$ (人)。

现设顾客到达后排成一队，依次向空闲的窗口购票，

如下图。是一个 $M/M/c$ 型的系统，其中 $c=3$, $\frac{\lambda}{\mu} = 2.25$,

$$\rho = \frac{\lambda}{c\mu} = \frac{2.25}{3} (< 1) \text{ 符合要求的公式。}$$



代入公式得

(1) 整个售票所空闲的概率

$$P = \frac{1}{\frac{(2.25)^0}{0!} + \frac{(2.25)^1}{1!} + \frac{(2.25)^2}{2!} + \frac{(2.25)^3}{3!} \cdot \frac{1}{1 - 2.25/3}} = 0.0748$$

(2) 平均队长

$$L_q = \frac{(2.25)^3 \cdot 3/4}{3!(1/4)^2} \times 0.0748 = 1.70$$

(3) 平均等待时间和逗留时间

$$W_q = 1.70 / 0.90 = 1.89 \text{ 分钟}$$

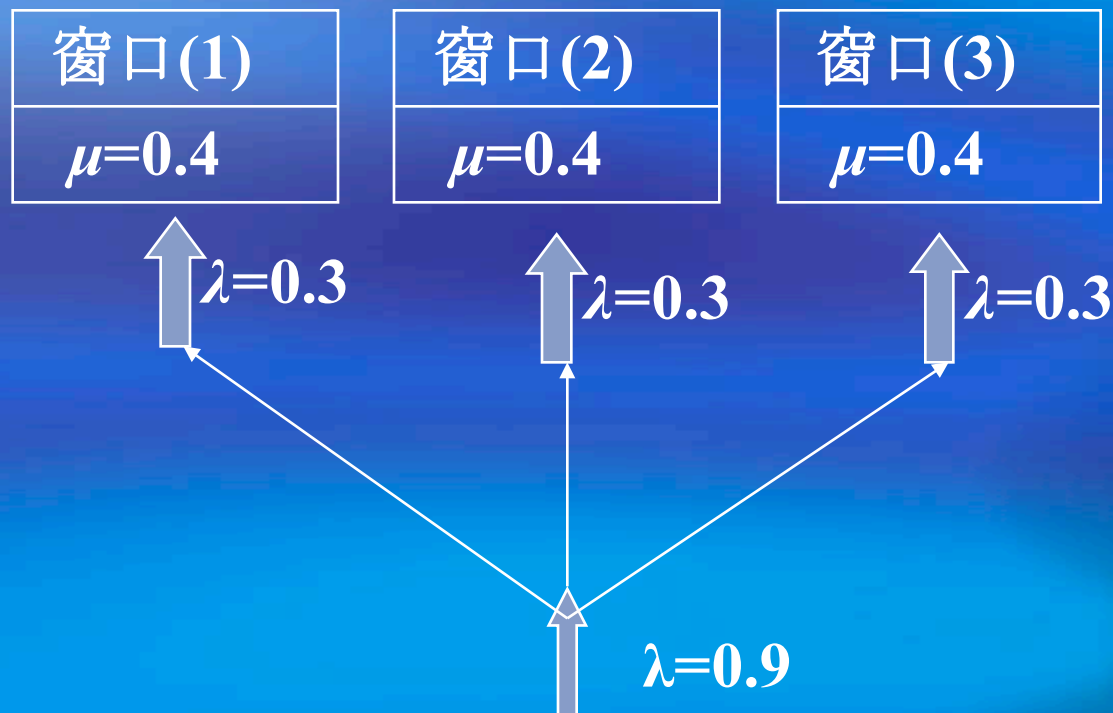
$$W_s = 1.89 + 1/4 = 4.39 \text{ 分钟}$$

(4) 顾客到达后必须等待（即系统中顾客数已有3人）的概率

$$P(n \geq 3) = \frac{(2.25)^3}{3!(1/4)} \times 0.0748 = 0.57$$

M/M/c型系统和c个M/M/1系统的比较

上例中，排队方式不变，但顾客到达后在每个窗口前各排一队，且进入队列后坚持不换，这就形成3个队列，如下图二每个队列平均到达率为 $\lambda = \lambda = \lambda = 0.9/3 = 0.3$ （每分钟），这样原来的系统就变成3个M/M/1型的子系统。



现按M/M/1型解决这个问题，并与上表比较：

指标 \ 模型	(1)M/M/3型	(2)M/M/1型
服务台空闲的概率	0.0748	0.25(每个子系统)
顾客必须等待的概率	$P(n \geq 3) = 0.57$	0.75
平均队列	1.70	2.25(每个子系统)
平均队长	3.95	9.00(整个系统)
平均逗留时间	4.39(分钟)	10(分钟)
平均等待时间	1.89(分钟)	7.5(分钟)

从表中各指标的对比可以看出单队比三队有显著的优越性