运筹学

Operational Research

OR

第三章 随机服务理论

北院喷泉

本章篇目

制作与教学

河北工业大学管理学院 孔造杰

kongzj@sina.com

Exit

§ 3-1 随机服务系统

- 系统的输入与输出是随机变量
- A.k.Erlang 于1909~1920年发表了一系列根据话务量计算电话机键配置的方法，为随机服务理论奠定了基础
- 又称为排队论(*Queuing Theory*)或拥塞理论(*Congestion Theory*)

与服务系统性能相关的特性

- 服务系统存在来自两个矛盾方面的要求
  - 顾客希望服务质量好，如排队等待时间短，
  - 系统运营方希望设备利用率高，损失率低
- 给用户一个经济上能够承受的满意的质量
- 哪些系统特性会影响系统的性能？
  - 服务机构的组织方式与服务方式
  - 顾客的输入过程和服务时间分布
  - 系统采用的服务规则

3.1.1 服务机构的组织方式与服务方式

- 单台制和多台制
- 并联服务和串联服务
- 串并联服务、网络服务
- 单人服务和成批服务

与服务系统性能相关的特性

3.1.2 输入过程和服务时间

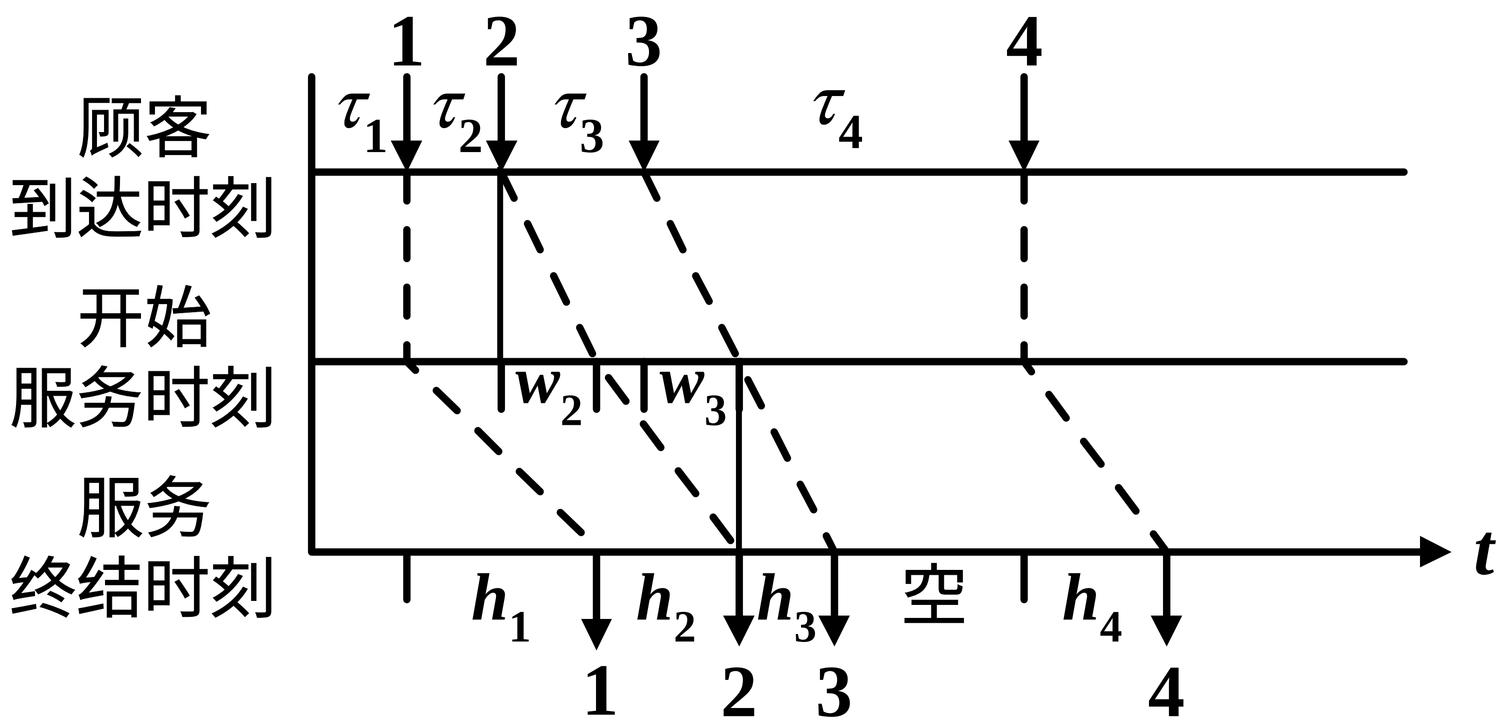
- 顾客单个到达或成批到达
- 顾客到达时间间隔的分布和服务时间的分布
- 顾客源是有限的还是无限的

3.1.3 服务规则

- 损失制（消失系统）
- 等待制：先到先服务(FIFO)，后到先服务，随机服务，优先权服务
- 混合制
- 逐个到达，成批服务；成批到达，逐个服务

§3-2 随机服务过程

- 单台服务系统、等待制、先到先服务
- 顾客在系统中的总时长：逗留时间=等待时长+服务时长
- 等待时长与顾客到达率和服务时长有关



§3-2 随机服务过程

当服务台连续不断服务时，有如下关系： $w_{i+1}+\tau_{i+1}= w_i+h_i$

- w_i+h_i 表示了累计的未完成的服务时长，一般地有

$$w_{i+1}=\begin{cases} w_i+h_i-\tau_{i+1} & \text{if } w_i+h_i-\tau_{i+1}\geq 0 \\ 0 & \text{if } w_i+h_i-\tau_{i+1}<0 \end{cases}$$

- τ_i, w_i, h_i 可通过写实来获得，另一种写实法

  $\alpha(t)$ 代表时段(0, t)中累计到达顾客数

  $\beta(t)$ 代表时段(0, t)中累计接受服务的顾客数

  $\gamma(t)$ 代表时段(0, t)中累计服务完毕的顾客数

- 则在任意考察时刻 t，有

  正在等待的顾客数：$L(t)= \alpha(t)-\beta(t)$

  正在接受服务的顾客数：$L_s(t)= \beta(t)-\gamma(t)$

  系统中逗留的顾客数：$N(t)= \alpha(t)-\gamma(t)$

- 上述关系是普遍成立的，与服务台设置和服务规则无关
- 下面分析等待顾客数、等待时间和顾客到达率的关系
  - 到达率定义为单位时间内平均到达的顾客数，即

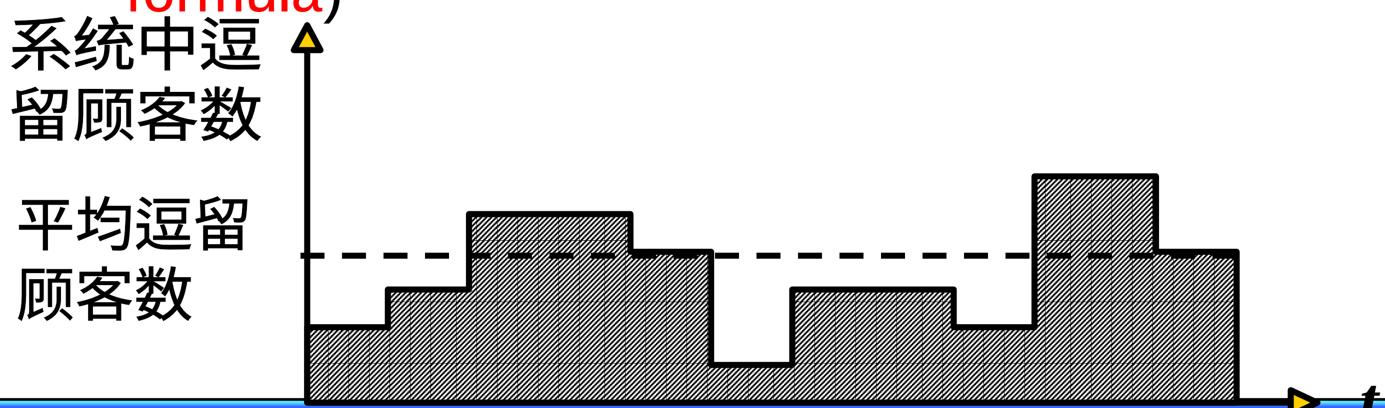
$$\lambda_t = \alpha(t)\ /\ t$$

  - 令 $\delta(t)$ 表示在时段(0, t)内到达系统内顾客的总逗留时长
  - 则每一个顾客的平均逗留时间为

$$W_{st} = \delta(t)\ / \alpha(t)$$

  - 系统中平均逗留顾客数可表达为

$$L_{st} = \delta(t)\ /\ t = (\alpha(t)\ /\ t\,)(\delta(t)\ / \alpha(t)\ ) = \lambda_t W_{s\,t}$$ (Little formula)



§ 3-2　随机服务过程

- 系统处于稳态时的利特尔公式：$L_s = \lambda W_s$
- 利特尔公式也是普遍成立的，已知其中任两个量，可以求出另一个量
- 利特尔公式的分解：

$$L_s = \lambda W_s = \lambda (W_q + \bar{h}) = L_q + L_n$$

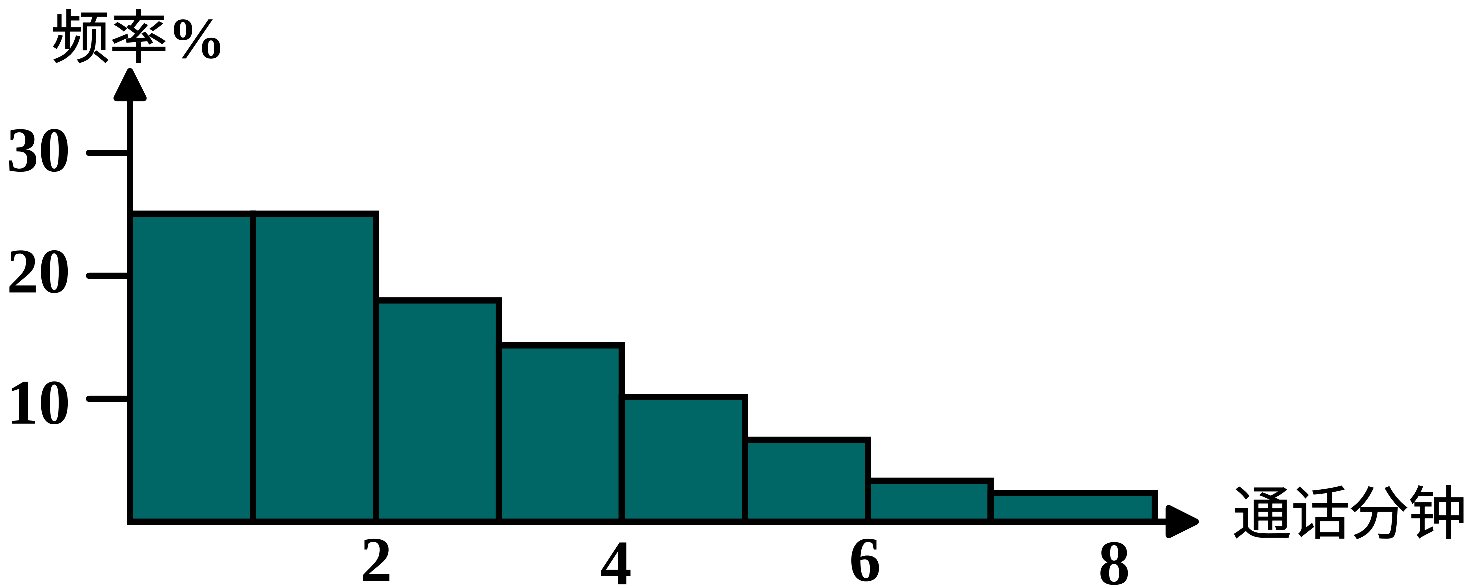
$$L_q = \lambda W_q \qquad L_n = \lambda \bar{h}$$

  - W_q 是顾客的平均排队等待时间
  - L_q 是排队等待的平均队长
  - $\bar{h}$ 是顾客的平均服务时长
  - L_n 是同时接受服务的平均顾客数(即平均服务台占用数)

§3-3 服务时间与到达间隔时间

3.3.1 概述

- 顾客的服务时间由于多种原因具有不确定性，最好的描述方法就是概率分布；同样顾客到达的间隔时间也具有一定的概率分布
- 服务时间和到达间隔时间服从什么分布？可以先通过统计得到经验分布，然后再做理论假设和检验
- 经验分布一般采用直方图来表示，如下图



若统计区间分得越细，样本越多，则经验分布的轮廓越接近曲线

- 一般服务时间和到达间隔时间都是非负的连续实变量，令 h 代表服务时间，τ 代表到达间隔时间，t 为给定的时间，则它们的概率分布函数分别表示为

  $F(t)=P\{h\le t\}$　　$F(t)=P\{\tau\le t\}$

- 它们的概率密度函数为　$f(t)=F'(t)$，具有性质：

  $f(t)\ge 0$，$\int f(t)\,dt=1$

- 服务时间落在区间(a, c)的概率为

$$P\{a<h\le c\}=\int_a^c f(t)dt=F(c)-F(a)$$

- 服务时间落在区间$(t, t+\Delta t)$的概率为　$P\{t<h\le t+\Delta t\}=f(t)\Delta t$
- 平均服务时长和平均到达间隔时长

$$\bar{h}=E[h]=\int_0^\infty tf(t)dt \qquad \bar{\tau}=E[\tau]=\int_0^\infty tf(t)dt$$

- 平均服务时长的倒数为服务率，平均间隔时长的倒数为到达率

$$\mu=1/\bar{h} \qquad \lambda=1/\bar{\tau}$$

3.3.2 常用的概率分布

1、定长分布

- 流水线的加工时间

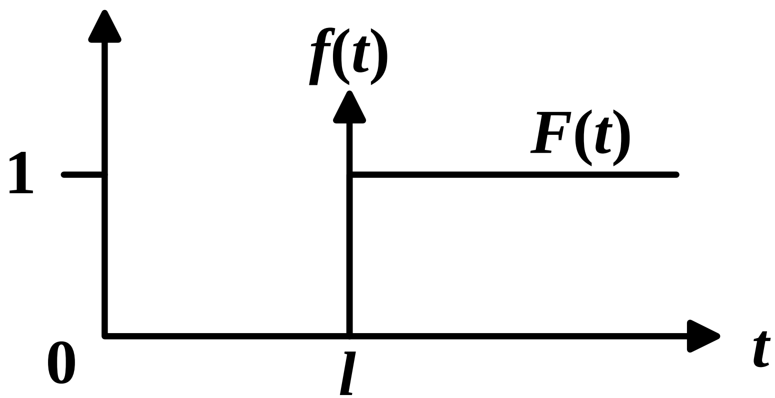
$$F(t)=\begin{cases}1 & 当 t\ge l\\ 0 & 当 t<l\end{cases}$$

2、负指数分布

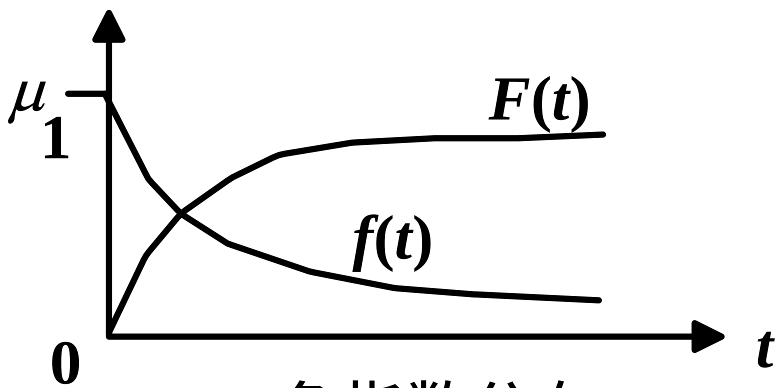
一类最常用的分布，如上述通话时长，可靠度等

$$F(t)=1-e^{-\mu t}$$

$$f(t)=\mu e^{-\mu t} \qquad \bar{h}=\int_0^{\infty} t\mu e^{-\mu t}dt=1/\mu$$



定长分布



负指数分布

3.3.2 常用的概率分布

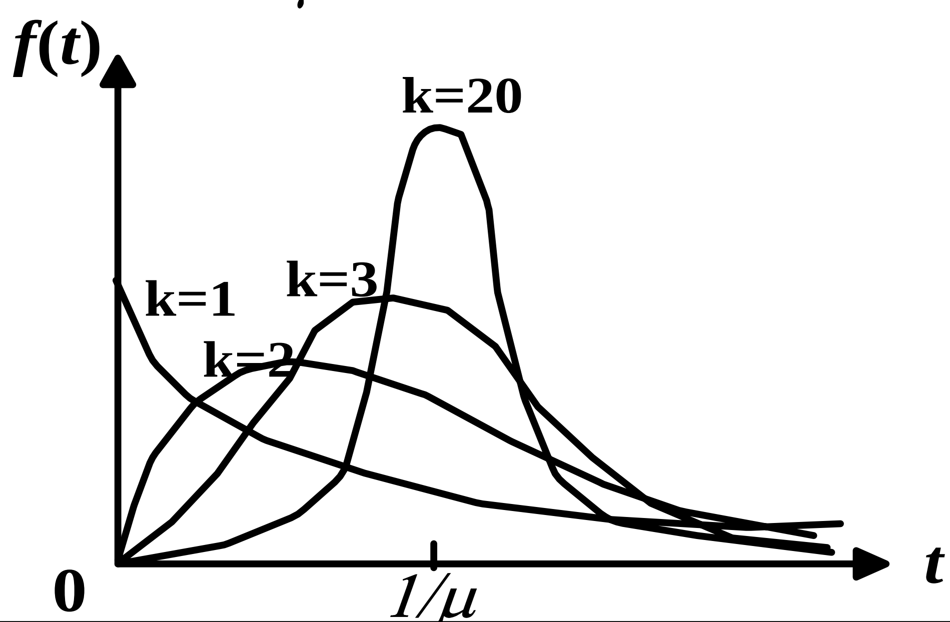
3、爱尔朗分布(Erlang分布)

$$f(t)=\frac{\mu k(\mu kt)^{k-1}}{(k-1)!}e^{-\mu kt}$$

- 一种代表性更广的分布

k 为整数，称为 k 阶爱尔朗分布；当 $k=1$ 时，退化为负指数分布；$k\to\infty$ 时趋向定长分布

爱尔朗分布实际上是 k 个独立同分布的负指数分布随机变量的和的分布，即 k 个服务台的串联，每个服务台的平均服务时长为 $1/k\mu$



3.3.3 负指数分布的特点

- 负指数分布之所以常用，是因为它有很好的特性，使数学分析变得方便
- 无记忆性。指的是不管一次服务已经过去了多长时间，该次服务所剩的服务时间仍服从原负指数分布

证：令 h 代表服务时间，t_0 代表服务已过去的时间，则服务剩余时间为 $h-t_0$，它的分布函数为

$$P\{h-t_0 \le t \mid h > t_0\} = P\{h \le t+t_0 \mid h > t_0\}$$

$$= \frac{P\{t_0 < h \le t+t_0\}}{P\{h > t_0\}} = \frac{P\{h \le t+t_0\} - P\{h \le t_0\}}{P\{h > t_0\}}$$

$$= \frac{1-e^{-\mu(t+t_0)} - (1-e^{-\mu t_0})}{1-(1-e^{-\mu t_0})} = 1-e^{-\mu t} \qquad \text{Q.E.D}$$

3.3.3 负指数分布的特点

- 一个服从负指数分布的服务，在下一瞬间结束的概率

  应用无记忆性

$$P\{h \le t_0 + \Delta t | h > t_0\} = 1 - e^{-\mu \Delta t}$$

$$= 1 - [1 - \mu \Delta t + (\mu \Delta t)^2 / 2! - (\mu \Delta t)^3 / 3! + \cdots]$$

$$= \mu \Delta t + o(\Delta t)$$

- 在 Δt 内服务终结的概率只与 μ 和 Δt 成正比，与 t_0 无关；因此 μ 又称为终结率，或离去率
- 同理，在 Δt 内服务不终结的概率为 $1-\mu \Delta t + o(\Delta t)$
- n 个独立同分布(负指数)的服务台同时被占用，在 Δt 内只有一个服务台终结的概率为

$$C_n^1(\mu\Delta t + o(\Delta t))(1 - \mu \Delta t + o(\Delta t))^{n-1} = n\mu\Delta t + o(\Delta t)$$

- 在 Δt 内有 $k>1$ 个服务台终结的概率为 $o(\Delta t)$，称为普通性

$$C_n^k(\mu \Delta t + o(\Delta t))^k (1 - \mu\Delta t + o(\Delta t))^{n-k} = o(\Delta t)$$

Have a rest

Would you like coffee or tea?

Help yourself

作业：教材

上一节 下一节 输入过程

§ 3-4 输入过程

- 即顾客到达的分布，可用相继到达顾客的间隔时间描述，也可以用单位时间内到达的顾客数描述
  - 间隔时间服从定长分布
  - 单位时间内到达的顾客数服从波松分布(法国数学家 *Poisson*, 1837)
  - 间隔时间服从爱尔朗分布
  - 一般独立同分布

3.4.1 波松输入过程及其特点

- (0, *t*)时间内到达 *k* 个顾客的个数服从波松分布，若λ 为到达率，则波松分布函数为：

$$P_k(t) = \frac{(\lambda t)^k}{k!} e^{-\lambda t}$$

电话呼叫的到达，商店的顾客到达，十字路口的汽车流，港口到达的船只，机场到达的飞机等

3.4.1 波松输入过程及其特点

(1) 平稳性：顾客到达数只与时间区间长度有关

(2) 无后效性：不相交的时间区间内所到达的顾客数是独立的

(3) 普通性：在 Δt 时间内到达一个顾客的概率为 $\lambda\Delta t + o(\Delta t)$，到达两个或两个以上顾客的概率为 $o(\Delta t)$；即两个顾客不可能同时到达

- 波松过程具有可迭加性
  - 即独立的波松分布变量的和仍为波松分布

设两个波松分布为：$P_i(t) = \dfrac{(\lambda_1 t)^i}{i!} e^{-\lambda_1 t}, P_j(t) = \dfrac{(\lambda_2 t)^j}{j!} e^{-\lambda_2 t}$

令 $n = i + j$, 在 $(0,t)$ 内来到 n 个顾客的概率为

$$P\{x_1 + x_2 = n\} = \sum_{j=0}^{\infty} P\{x_1 = j\} \bullet P\{x_2 = n - j\}$$

$$= \sum_{j=0}^{\infty} \frac{(\lambda_1 t)^j}{j!} e^{-\lambda_1 t} \bullet \frac{(\lambda_2 t)^{n-j}}{(n-j)!} e^{-\lambda_2 t} = \frac{(\lambda_1 + \lambda_2)^n t^n}{n!} e^{-(\lambda_1+\lambda_2)t}$$

3.4.1 波松输入过程及其特点

波松过程的到达间隔时间为负指数分布

- 令 τ 代表间隔时间，则概率 $P\{\tau > t\}$代表时间区间(0, t)内没有顾客来的概率；由波松分布可知

$$P_0(t) = P\{\tau > t\} = e^{-\lambda t}$$

- 故间隔时间 τ 的分布为 $P\{\tau \le t\} = 1 - e^{-\lambda t}$

3.4.2 马尔科夫链

- 马尔科夫链(*Markov Chain*)又简称马氏链，是一种离散事件随机过程。用数学式表达为

$$P\{X_{n+1}=x_{n+1} | X_1=x_1, X_2=x_2, \ldots, X_n=x_n\} = P\{X_{n+1}=x_{n+1} | X_n=x_n\}$$

- X_{n+1}的状态只与 X_n的状态有关，与 X_n 前的状态无关，具有无记忆性，或无后效性，又称马氏性
- 状态转移是一步一步发生的，一步状态转移概率

$$P_{ij}(t) = P\{X_{n+1}=j | X_n=i\}$$

例 3.4.2

一售货员出售两种商品A和B，每日工作 8 小时。购买每种商品的顾客到达过程为波松分布，到达率分别为 λ_A=8人/日，λ_B=16人/日，试求：(1) 1小时内来到顾客总数为 3 人的概率；(2) 三个顾客全是购买B类商品的概率。

解：(1)总到达率为 λ_A+ λ_B=24人/日，1 小时=1/8 日，故

$$P_3(1/8)=\frac{(24\times1/8)^3}{3!}e^{-24\times1/8}=0.224$$

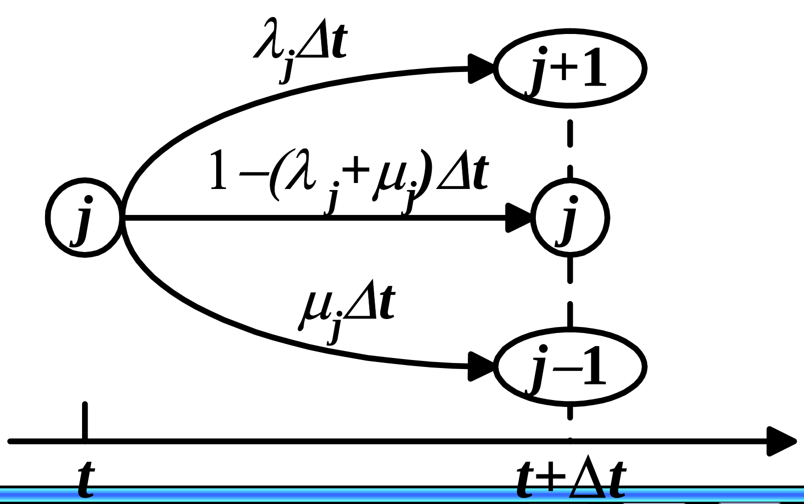
(2) 3 个顾客全是购买 B 类商品的概率为

$$P_{B3}(1/8)P_{A0}(1/8)=\frac{(16\times1/8)^3}{3!}e^{-16\times1/8}\times e^{-8\times1/8}$$

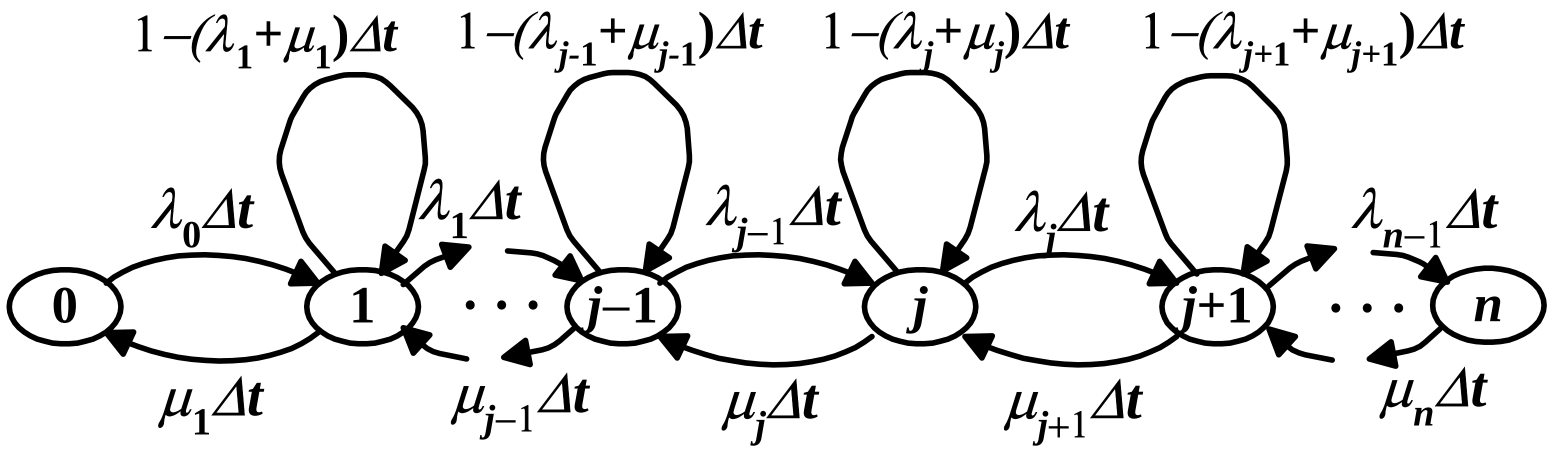
$$=0.0664$$

§3-5 生灭过程

- 一种描述自然界生灭现象的数学方法，如细菌的繁殖和灭亡，人口的增减，生物种群的灭种现象等
- 采用马氏链
  - 令 $N(t)$代表系统在时刻 t 的状态，下一瞬间 $t+\Delta t$ 系统的状态只能转移到相邻状态，或维持不变，如图所示
  - 三种转移是不相容的，三者必居其一
  - 只有具有无记忆性和普通性的过程(分布)才适用马氏链
  - 令 $P_j(t)=P\{N(t)=j\}$代表系统在时刻 t 处于状态 j 的概率



生灭过程的马氏链



- 根据马氏链，应用全概率公式，有状态转移概率方程

$$p_j(t+\Delta t)=p_{j-1}(t)\lambda_{j-1}\Delta t+p_{j+1}(t)\mu_{j+1}\Delta t$$
$$+p_j(t)(1-\lambda_j\Delta t-\mu_j\Delta t)+o(\Delta t)\qquad j=1,2,\cdots,n-1$$

- 另有两个边界方程

$$p_0(t+\Delta t)=p_1(t)\mu_1\Delta t+p_0(t)(1-\lambda_0\Delta t)+o(\Delta t)$$
$$p_n(t+\Delta t)=p_{n-1}(t)\lambda_{n-1}\Delta t+p_n(t)(1-\mu_n\Delta t)+o(\Delta t)$$

生灭方程的推导过程

- 将上述三个差分方程化为微分方程

$$\lim_{\Delta t\to 0}\frac{p_j(t+\Delta t)-p_j(t)}{\Delta t}=\lambda_{j-1}p_{j-1}(t)+\mu_{j+1}p_{j+1}(t)-(\lambda_j+\mu_j)p_j(t)\qquad j=1,2,\cdots,n-1$$

即 $$p'_j(t)=\lambda_{j-1}p_{j-1}(t)+\mu_{j+1}p_{j+1}(t)-(\lambda_j+\mu_j)p_j(t)$$

同理有 $$p'_0(t)=\mu_1 p_1(t)-\lambda_0 p_0(t)$$

$$p'_n(t)=\lambda_{n-1}p_{n-1}(t)-\mu_n p_n(t)$$

- 上述三个方程是动态方程，当系统处于稳态时，系统处于统计平衡状态，即状态概率不随时间变化，从而状态概率导数为 0；令上三个方程左侧为 0，得稳态方程组

$$\mu_1 p_1-\lambda_0 p_0=0\qquad j=0\qquad (1)$$

$$\mu_{j+1}p_{j+1}+\lambda_{j-1}p_{j-1}-(\lambda_j+\mu_j)p_j=0\qquad 1\le j<n\qquad (2)$$

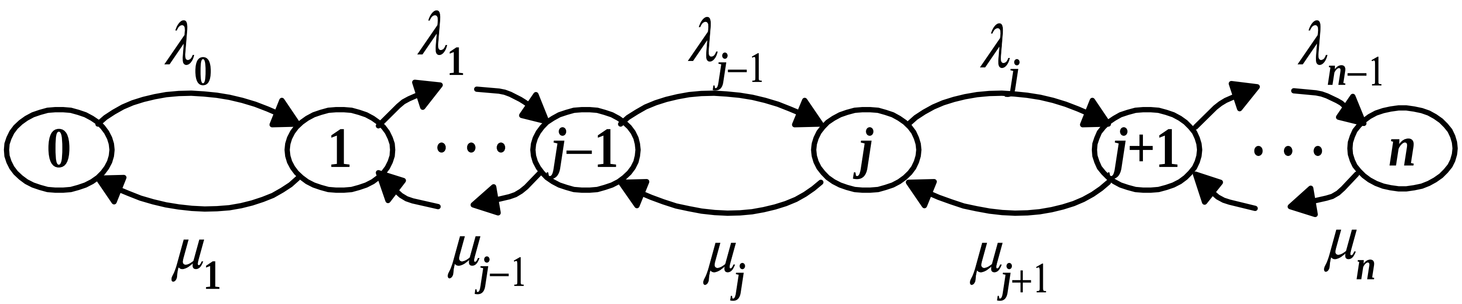
$$\mu_n p_n-\lambda_{n-1}p_{n-1}=0\qquad j=n\qquad (3)$$

生灭过程稳态解

$$\mu_1 p_1 - \lambda_0 p_0 = 0 \qquad j = 0 \qquad (1)$$

$$\mu_{j+1} p_{j+1} + \lambda_{j-1} p_{j-1} - (\lambda_j + \mu_j) p_j = 0 \qquad 1 \le j < n \qquad (2)$$

$$\mu_n p_n - \lambda_{n-1} p_{n-1} = 0 \qquad j = n \qquad (3)$$



- 方程(1), (2), (3) 与稳态状态态转移图一一对应；递归解如下：

$$p_1 = \frac{\lambda_0}{\mu_1} p_0 \qquad \text{令 } j = 1, \text{将 } p_1 \text{ 代入 (2)式得：} p_2 = \frac{\lambda_0 \lambda_1}{\mu_1 \mu_2} p_0$$

依次递推，得 $p_j = \dfrac{\lambda_{j-1}}{\mu_j} p_{j-1} = \dfrac{\lambda_0 \lambda_1 \cdots \lambda_{j-1}}{\mu_1 \mu_2 \cdots \mu_j} p_0$ (归纳法)

由 $\sum_0^n p_j = 1$, 得 $p_0 = \left[1 + \sum_{j=1}^{n} \dfrac{\lambda_0 \lambda_1 \cdots \lambda_{j-1}}{\mu_1 \mu_2 \cdots \mu_j} \right]^{-1}$

满足生灭过程的条件

- 系统的输入过程和服务过程具有平稳、无记忆性和普通性
- 服务台是独立的、相同的、并联的
- 波松输入过程和负指数服务时长就具有这些性质
  - 可以用马氏链来描述系统的状态转移
  - 这种系统称为生灭服务系统，一般用 M/M/*n* 表示，又称为标准服务系统；
  - 标准服务系统的形式很多，但都是基于生灭方程，关键是找出 λ_j，μ_j 的不同表达式，将它们代入生灭方程
- 标准服务系统的表示法：(X/Y/Z: A/B/C)，X 表示输入过程，Y 表示服务过程，Z 表示并联服务台的个数，A 表示顾客源，B 表示系统容量，C 表示服务规则
  - (M/M/*n*: ∞/ *m*/FIFO) 表示波松输入，负指数服务时长，*n* 个并联服务台，顾客源无穷，系统容量为 *m*, $m \geq n$，先到先服务
  - D 定长分布；E_k k阶爱尔兰分布；G 一般独立分布

§3-6 纯增过程

- 令生灭过程中所有消亡率 μ_j=0， 即只有顾客到达，没有顾客离去
- 令所有 λ_j=λ，且系统服务台无限，即 $n\to\infty$
- 容易得到下列微分方程组

$$\frac{dP_0(t)}{dt}=-\lambda P_0(t) \qquad j=0$$

$$\frac{dP_j(t)}{dt}=-\lambda P_j(t)+\lambda P_{j-1}(t) \qquad j\geq 1$$

由此解出 $P_0(t)=e^{-\lambda t}$

解一次型微分方程，并递推得

$$P_j(t)=\frac{(\lambda t)^j}{j!}e^{-\lambda t} \qquad j=1,2,\cdots$$

该问题没有稳态解，正是波松过程

本章结束

请继续第四章
标准服务系统

作业：教材

上一节 下一节 标准服务系统