第六讲: 排队模型和排队系统仿真

数学模型和算法的应用与 MATLAB 实现

周吕文

中国科学院力学研究所

2017 年 7 月 8 日

微信公众号: 超级数学建模

Notes

历史

排队论又称随机服务系统，是研究系统随机聚散现象和随机服务系统工作过程的数学理论和方法，是运筹学的一个分支。

排队论的基本思想是 1909 年丹麦数学家 A.K. 埃尔朗在解决自动电话设计问题时开始形成的，当时称为话务理论。

现实生活中如排队买票、病人排队就诊、轮船进港、高速路上汽车排队通过收费站、机器等待修理等都属于排队论问题。

Notes

定义

通过对服务对象到来及服务时间的统计研究，

得出这些数量指标（等待时间、排队长度、忙期长短等）的统计规律，

然后根据这些规律来改进服务系统的结构或重新组织被服务对象，

使得服务系统既能满足服务对象的需要，又能使机构的费用最经济或某些指标最优。

Notes

应用

CUMCM 2009B 的眼科病床的合理安排问题

MCM 2005B 收费站最佳配置问题

ICM 2017D 机场安检问题

Notes

排队论基本构成与指标

排队论的基本构成

输入过程：描述顾客按照怎样的规律到达排队系统。顾客总体（有限/无限）、到达的类型（单个/成批）、到达时间间隔。

排队规则：指顾客按怎样的规定次序接受服务。常见的有等待制、损失制、混合制、闭合制。

服务机构：服务台的数量；服务时间服从的分布。

排队系统的数量指标

队长：系统中的平均顾客数（包括正在接受服务的顾客）。

等待队长：系统中处于等待的顾客的数量。

等待时间：等待时间包括顾客的平均逗留时间。

忙期：连续保持服务的时长。

Notes

数学表示

排队论中的符号表示

$$A/B/C/n$$

A 输入过程，B 服务时间，C 服务台数，n 系统容量。

排队论表示实例　$M/M/S/\infty$

输入过程是 Poisson 流

服务时间服从负指数分布

系统有 S 个服务台平行服务

系统容量为无穷大的等待制排队系统

Notes

等待制模型 $M/M/S/\infty$

顾客到达规律服从参数为 λ 的 Poisson 分布

$$P\{X(t)=k\}=\frac{(\lambda t)^k e^{-\lambda t}}{k\,!}$$

$[0,t]$ 时间内到达的顾客平均数为 λt。

顾客接受服务的时间服从参数为 μ 的负指数分布

$$f(t)=\mu e^{-\mu t},\ (t>0)$$

每个顾客接受服务的平均时间为 $1/\mu$。

Notes

等待制模型 $M/M/S/\infty$：$S=1$

系统的服务强度 & 无顾客的概率 & 有 n 个顾客的概率

$$\rho=\frac{\lambda}{\mu},\quad p_0=1-\rho=1-\frac{\lambda}{\mu},\quad p_n=(1-\rho)\rho^n$$

平均队长

$$L_s=\sum_{n=0}^{\infty} n\cdot p_n=(1-\rho)\sum_{n=0}^{\infty} n\rho^n=\frac{\rho}{1-\rho}=\frac{\lambda}{\mu-\lambda}$$

平均等待队长

$$L_q=\sum_{n=1}^{\infty}(n-1)\cdot p_n=(1-\rho)\sum_{n=1}^{\infty}(n-1)\rho^n=\frac{\rho^2}{1-\rho}=\frac{\lambda^2}{\mu(\mu-\lambda)}$$

Notes

等待制模型 $M/M/S/\infty$: $S=1$

平均逗留时间

$$W_s = \frac{1}{\mu - \lambda}$$

平均等待时间

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

Little 公式

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

Notes

等待制模型 $M/M/S/\infty$: $S=1$

案例

某机关接待室只有 1 名对外接待人员。
来访人员按照 Poisson 流到达，到达速率为 $\mu = 8$ 人/小时。
接待人员的服务速率间服 $\lambda = 9$ 人/小时的负指数分布。

来访人员的平均等待时间，等候的平均人数

$$S = 1, \quad \lambda = 8, \quad \mu = 9$$

$$W_q = \frac{\lambda}{\mu(\mu - \lambda)} = \frac{8}{9(9-8)} = \frac{8}{9} = 0.89$$

$$L_q = \lambda W_q = \frac{64}{9} = 7.1$$

Notes

等待制模型 $M/M/S/\infty$: $S>1$

服务能力和强度

$$S\mu, \quad \rho = \frac{\lambda}{s\mu}$$

服务台都空闲的概率

$$p_0 = \left[\sum_{k=0}^{S-1} \frac{(S\rho)^k}{k!} + \frac{(S\rho)^S \rho}{S!(1-\rho)} \right]^{-1}$$

平均队长

$$L_s = S\rho + \frac{(S\rho)^S \rho}{S!(1-\rho)^2} \cdot p_0$$

Notes

等待制模型 $M/M/S/\infty$: $S>1$

平均逗留时间

$$W_s = \frac{L_s}{\lambda}$$

平均等待时间

$$W_q = W_s - \frac{1}{\mu}$$

平均等待队长

$$L_q = \lambda W_q$$

Notes

等待制模型 $M/M/S/\infty$: $S = 1$

案例

来访人员按照 Poisson 流到达，到达速率为 $\mu = 20$ 人/小时。
接待人员的服务速率间服 $\lambda = 9$ 人/小时的负指数分布。
为使来访问者等待不超过半小时，最少应配置几名接待员？

```
lambda = 20; mu = 9; s = 3;
rho = lambda/(s*mu);
p0 = 1./( sum((s*rho).^k./factorial(k)) + ...
          (s*rho)^s/(factorial(s)*(1-rho)) );
Ls = s*rho + (s*rho)^s*rho/(factorial(s)*(1-rho)^2)*p0;
Ws = Ls/lambda;
Wq = Ws - 1/mu
```


Notes

其它模型

损失制模型 M/M/S/S

顾客到达服从泊松分布，服务台服务时间服从负指数分布，
当 S 个服务台被占用后，顾客自动离开，不再等待。

混合制模型 M/M/S/K

顾客到达服从泊松分布，服务台服务时间服从负指数分布，
系统容量为 K，当 K 个位置被占用时，顾客自动离开。

闭合制模型 M/M/S/K/K

顾客到达服从泊松分布，服务台服务时间服从负指数分布，
系统容量和潜在的顾客数都为 K。

Notes

单服务台

开始服务, 到达, 离开时刻和服务, 等待时长的关系

$$\text{服务时刻}(i) = \max\left\{\text{到达时刻}(i), \text{离开时刻}(i-1)\right\}$$

$$\text{离开时刻}(i) = \text{服务时刻}(i) + \text{服务时长}(i)$$

$$\text{等待时长}(i) = \text{离开时刻}(i) - \text{到达时刻}(i)$$

Notes

多服务台

开始服务, 到达, 离开时刻和服务, 等待时长的关系

$$\text{服务时刻}(i) = \max\left\{\text{到达时刻}(i), \min\{\text{服务台空闲时刻}\}\right\}$$

$$\text{所使用服务台}(i) = k, \quad \text{其中 } k \text{ 使 服务台空闲时刻}(k) = \min$$

$$\text{离开时刻}(i) = \text{服务时刻}(i) + \text{服务时长}(i)$$

$$\text{服务台空闲时刻}(k) = \text{离开时刻}(i)$$

$$\text{等待时长}(i) = \text{离开时刻}(i) - \text{到达时刻}(i)$$

Notes

自动取款机问题

问题

银行计划安置取款机, A 机价格和平均服务率都是 B 机的 2 倍. 应购置 1 台 A 机还是 2 台 B 机?

顾客平均每分钟到达 1 位, A 型机的平均服务时间为 0.9, B 型机为 1.8 分钟, 顾客到达间隔和服务时间都服从指数分布.

Notes

单服务台

mm1.m

```
n = 100000;                       % 模拟顾客总数
mu = 1; muA = 0.9;                % 到达率和服务率
tarr = cumsum(exprnd(mu,1,n));    % 到达时刻
tsrv = exprnd(muA,1,n);           % 服务时长
tsta = zeros(1,n);                % 初始化服务时刻
tlea = zeros(1,n);                % 初始化离开时刻
twat = zeros(1,n);                % 初始化等待时长
tsta(1) = tarr(1);                % 首位顾客服务时刻=到达时刻
tlea(1) = tsta(1) + tsrv(1);      % 首位顾客离开时刻
twat(1) = tlea(1) - tarr(1);      % 首位顾客等待时长=0
for i = 2:n
    % 服务时刻 = max{到达时刻, 上一个顾客离开时刻}
    tsta(i) = max(tarr(i),lea(i-1));
    tlea(i) = tsta(i) + tsrv(i);% 离开时刻=服务时刻+服务时长
    twat(i) = tlea(i) - tarr(i);% 等待时长=离开时刻-到达时刻
end
```

Notes

两服务台

mm2.m

```
n = 100000;                        % 模拟顾客总数
mu = 1; muB = 1.8;                 % 到达率和服务率
tarr = cumsum(exprnd(mu,1,n));     % 到达时刻
tsrv = exprnd(muB,1,n);            % 服务时长
tsta = zeros(1,n); tlea = zeros(1,n);% 初始化服务/离开时刻
twat = zeros(1,n);                 % 初始化等待时长
last = [0 0];                      % 初始化服务台结束服务时刻
for i = 1:n
    [minval, k] = min(last);  % 找出最快结束服务的服务台时刻
    tsta(i) = max(tarr(i),minval);% 服务时刻
    tlea(i) = tsta(i) + tsrv(i);  % 离开时刻
    last(k) = tlea(i);            % 服务台结束服务时刻
    twat(i) = tlea(i) - tarr(i);  % 等待时长
end
```

Notes

2013HIMCM-B: 银行服务问题

银行经理正试图通过提供更好的服务来提高顾客满意度. 管理层期待顾客平均等待时间小于 2 分钟, 平均队列长度 (等待队列的长度) 是 2 人或者更少. 银行估算, 每天大约为 150 名顾客提供服务. 现有的到达和服务时间如下表所示:

表: 到达时间

到达间隔 (min.)	0	1	2	3	4	5
概率	0.10	0.15	0.10	0.35	0.25	0.05

表: 服务时间

服务时间 (min.)	1	2	3	4
概率	0.25	0.20	0.40	0.15

建立一个该系统的数学模型。

根据经理的指引, 确定当前顾客是否对服务满意. 如果不满意, 通过模型对服务进行微小的改变, 以达到经理的目标.

除了比赛格式的论文以外, 给经理写一封简短的, 非技术性的, 大约 1-2 页的信, 给出你的最终建议.

Notes

2013HIMCM-B: 银行服务问题

由时间间隔 $t = [0 \;\; 1 \;\; 2]$ 和概率 $p = [0.2 \; 0.3 \; 0.5]$ 得到各到顾客达时间间隔

$$p' = \text{cumsum}([0.5 \; 0.3 \; 0.2]) = [0.5 \; 0.8 \; 1.0]$$

$$R = \text{rand}(1,5) = [0.1 \; 0.9 \; 0.2 \; 0.4 \; 0.8];$$

$$R(R < 0.5) = 2, \quad R(R < 0.8) = 1, \quad R(p < 1.0) = 0$$

由到达时间间隔得到各顾客到达时刻

$$\text{间隔} = [0 \; 1 \; 3 \; 2] \quad \Longrightarrow \text{时刻} = \text{cumsum}(\text{间隔}) = [0 \; 1 \; 4 \; 6]$$

开始服务, 到达, 离开时刻和服务, 等待时间的关系:

$$\text{开始服务的时刻}(i) = \max\left\{\text{到达时刻}(i), \text{离开时刻 (i-1)}\right\}$$

$$\text{离开时刻}(i) = \text{开始服务的时刻}(i) + \text{服务时间}(i)$$

$$\text{等待时间}(i) = \text{离开时刻}(i) - \text{到达时刻}(i) - \text{服务时间}(i)$$

Notes

2013HIMCM-B: 银行服务问题

计算 Tarrival, Tservice

```
n = 150;
ta = [5 4 3 2 1 0];  pa = [0.05 0.25 0.35 0.10 0.15 0.10];
ts = [  4 3 2 1  ];  ps = [     0.15 0.40 0.20 0.25     ];
pacum = cumsum(pa);
pscum = cumsum(ps);
Tarrival = rand(1,n);
for i = 1:length(pa)
    Tarrival(Tarrival<pacum(i)) = ta(i);
end
Tarrival = cumsum(Tarrival);

Tservice = rand(1,n);
for i = 1:length(ps)
    Tservice(Tservice<pscum(i)) = ts(i);
end
```

Notes

2013HIMCM-B: 银行服务问题

计算 Tarrival, Tservice

```
Tstart = zeros(1,n); Tleave = zeros(1,n);
Twait  = zeros(1,n); line   = zeros(1,n);

Tstart(1) = Tarrival(1);
Tleave(1) = Tstart(1) + Tservice(1);
Twait(1) = Tleave(1) - Tarrival(1) - Tservice(1);
line(1) = 0;

for i = 2:n
   Tstart(i) = max(Tleave(i-1), Tarrival(i));
   Tleave(i) = Tstart(i) + Tservice(i);
   Twait(i) = Tleave(i) - Tarrival(i) - Tservice(i);

   k = i-1;
   while ( k>0 )&&( Tarrival(i)<Tleave(k) )
       line(i) = line(i) + 1;
       k = k - 1;
   end
end
```

Notes

2013HIMCM-B: 银行服务问题

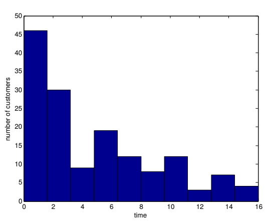


图: 等待时间分布

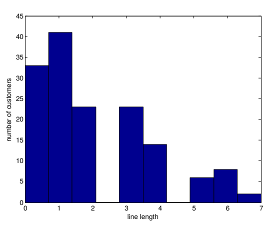
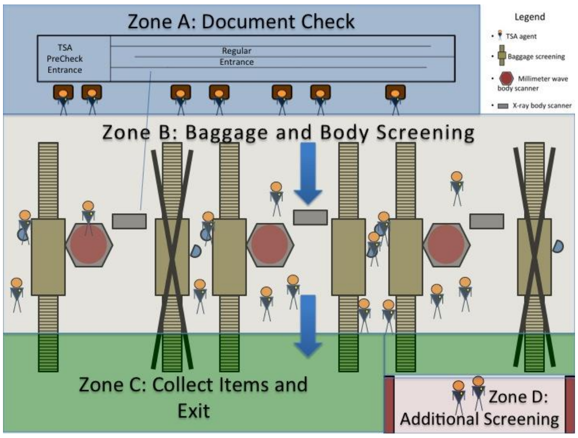


图: 队长分布

Notes

ICM2017-D: 优化机场安检口旅客通行



Notes

ICM2017-D: 优化机场安检口旅客通行

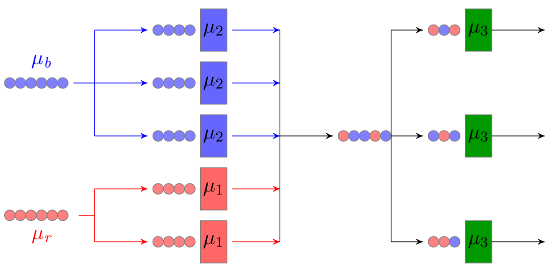
问题

建立一个或多个模型，研究旅客通过安检口的流量，确定瓶颈，明确判断当前流程问题区域位置。

设计两个或更多对现有系统德潜在改进，提高旅客通信，减少等待时间。模拟这些变化展示改进如何影响流程。

Notes

排队系统：$\mu_r = 10$, $\mu_b = 13$, $\mu_1 = 12$, $\mu_2 = 9$, $\mu_3 = 16$



Notes

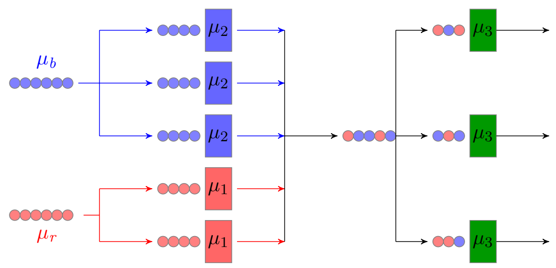
多服务并联

mms.m

```
function [tlea,twat,qlen] = mms(tarr, type, mus)
narr = length(tarr); nsvr = length(mus);
last = zeros(nsvr,1);
[tsta, tlea, twat, qlen] = deal(zeros(narr,1));
rndm = zeros(nsvr,narr);
for k = 1:nsvr; rndm(k,:) = exprnd(mus(k)*type); end
for i = 1:narr
    [minemp, ksvr(i)] = min(last);
    tsta(i) = max(tarr(i), minemp);
    tsvr(i) = rndm(ksvr(i),i);
    tlea(i) = tsta(i) + tsvr(i);
    last(ksvr(i)) = tlea(i);
    twat(i) = tlea(i) - tarr(i);
    j = i - 1;
    while j>0 && tarr(i)<tlea(j)
        if ksvr(j)==ksvr(i); qlen(i) = qlen(i) + 1; end
        j = j - 1;
    end
end
```

Notes

串并混合系统: $\mu_r = 10,\ \mu_b = 13,\ \mu_1 = 12,\ \mu_2 = 9,\ \mu_3 = 16$



Notes

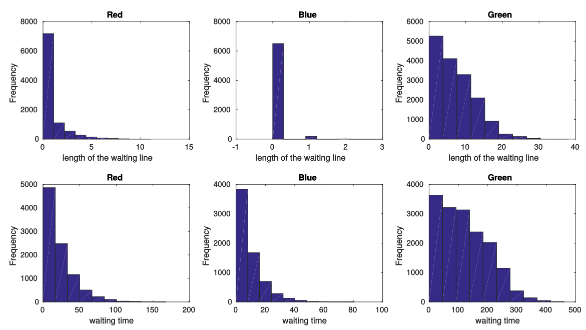
串并联混合系统

mms.m

```
n1 = 2; n2 = 3; n3 = 3;
mu1 = 12; mu2 = 9; mu3 = 16;
muR =  9.19; muB = 12.95;
nR = ceil(24*3600/muR); nB = ceil(24*3600/muB);
tArrR = cumsum(exprnd(muR,nR,1));
tArrB = cumsum(exprnd(muB,nB,1));
tArr = [tArrR; tArrB];
type = [0.8*ones(nR,1); 1.2*ones(nB,1)];
[tLeaR,tWatR,qLenR] = mms(tArrR, ones(nR,1), mu1*ones(n1,1));
[tLeaB,tWatB,qLenB] = mms(tArrB, ones(nB,1), mu2*ones(n2,1));
[tArrG, order] = sort([tLeaR; tLeaB]);
[tLeaG,tWatG,qLenG] = mms(tArrG,type(order), mu3*ones(n3,1));
tLeaG(order) = tLeaG;
tWatG(order) = tWatG;
qLenG(order) = qLenG;
hist(qLenR); ylabel('Frequency');
```

Notes

结果



Notes

Thank You!!!

Notes