一、问题引入与分析

2001 B题 公交车调度

该条公交线路上行方向共 14 站,下行方向共 13 站,第 3-4 页给出的是典型的一个工作日两个运行方向各站上下车的乘客数量统计。公交公司配给该线路同一型号的大客车每辆标准载客 100 人,据统计客车在该线路上运行的平均速度为 20 公里/小时。运营调度要求,乘客候车时间一般不要超过 10 分钟,早高峰时一般不要超过 5 分钟,车辆满载率不应超过 120%,一般也不要低于 50%。

试根据这些资料和要求,为该线路设计一个便于操作的全天(工作日)的公交车调度方案,包括两个起点站的发车时刻表;一共需要多少辆车;这个方案以怎样的程度照顾到了乘客和公交公司双方的利益;等等。



试根据这些材料和要求,为该线路设计一个便于操作的全天(工作日)的公交车调度方案,包括: 1.两个起点站的发车时刻表:

2. 一共需要多少辆车;

这个方案以怎样的程度照顾到了乘客和公交公司双方的利益等等.



2.问题分析:

对于第一个问题---公交车的调度方案,可以有多种方法,比如多目标规划。

但要建立具体、明确、完整的数学模型,以及模型的求解,如何采集运营数据是关键的。

在实际中,公交车具有随机性,特别是顾客的到来是随机的,是无法精确预测的,这就决定了公交车的数量并不是完全确定,而是随着不同时段顾客的多少而变化。因此顾客和公交车之间存在随机因素。

很明显,如果公交车数量过多,减少了顾客排队等待的时间,使顾客的满意度很高,却会造成公交车公司的浪费,长时间的空跑最终会使公交车公司不堪重负;

另一方面,如果公交车数量较少,很多顾客得不到服务,或者很多顾客花了很长的时间去排队等候,造成顾客的不满意度提高,这样公交车公司会失去顾客。



2.排队系统描述

排队系统又称为随机服务系统,是研究服务过程和拥挤现象的随机模型。

排队系统的共同特征:

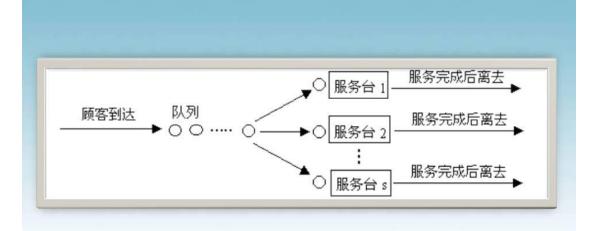
- ✓ 请求服务的人或者物,即顾客;
- ✓ 为顾客服务的人或者物,即服务员或服务台;
- ✓ 顾客到达系统的时刻是随机的,为每一位顾客提供服务的时间是随机的,因而整个排队系统的状态也是随机的。

排队系统的几种形式:



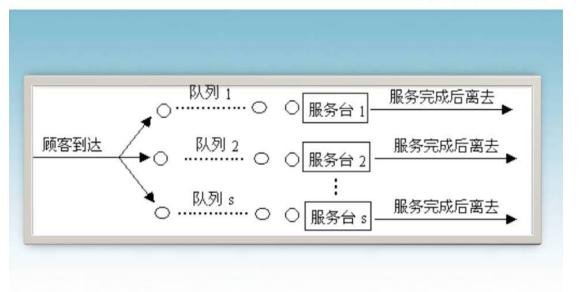
单服务台排队系统





单队列——S个服务台的并联排队系统





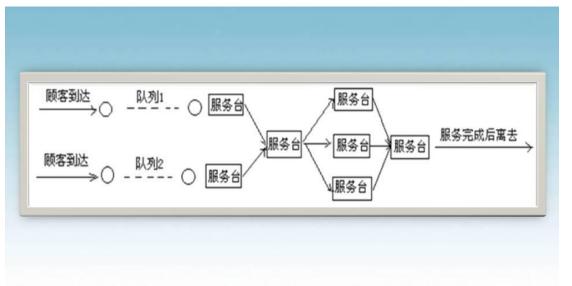
S个队列——S个服务台的并联排队系统





单队列——多个服务台的串联排队系统

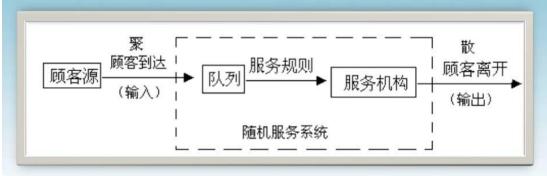




多队列——多个服务台的混联排队系统



尽管各种排队系统具体形式不同,但都可以这样描述:



随机服务系统

顾客源里的每个顾客按一定方式到达服务系统,首先加入 队列排队等待接受服务,然后服务台按一定规则从队列中 选择顾客进行服务,获得服务的顾客立即离开。



3.排队系统的基本组成部分

排队系统是由输入过程、排对规则、服务机构组成。

- (1)输入过程,也称为顾客流,指要求服务的顾客是按怎样的规律到达排队系统。一般可以从3个方面来描述一个输入过程。
- (i) <mark>顾客总体数,</mark>又称顾客源、输入源。这是指顾客的来源.顾客源可以是有限的,也可以是无限的。
- (ii) 顾客到达方式, 这是描述顾客是怎样来到系统的,是单个 到达,还是成批到达。
- (iii) 顾客流的概率分布,或称相继顾客到达的时间间隔的分布。 这是求解排队系统有关运行指标问题时,首先需要确定的 指标。顾客流的概率分布一般有定长分布、二项分布、泊 松流(最简单流)、爱尔朗分布等若干种。

- (2)排对规则 指服务台从队列中选取顾客进行服务的顺序。
- 一般可以分为损失制、等待制和混合制等3大类。
- (i) 损失制 指如果顾客到达排队系统时,所有服务台都被先到的顾客占用,那么他们就自动离开系统不再回来。
- (ii)等待制 指当顾客来到系统时,所有服务台都不空,顾客加入排队行列等待服务。等待制中,服务台在选择顾客进行服务时常有如下四种规则:
 - a.先到先服务 按顾客到达的先后顺序对顾客进行服务。
 - b.先到后服务
 - c. 随机服务 即当服务台空闲时,不按照排队序列而随意指 定某个顾客接受服务。
 - d.优先权服务



- (iii)混合制 这是等待制与损失制相结合的一种服务规则,一般是指允许排队,但又不允许队列无限长下去. 具体说来,大致有三种:
 - ▶a.队长有限.当排队等待服务的顾客人数超过规定数量时,后来的顾客就自动离去,另求服务,即系统的等待空间是有限的.
 - ▶b.等待时间有限.即顾客在系统中的等待时间不超过某一给定的长度T,当等待时间超过T时,顾客将自动离去,并不再回来.
 - ▶c.逗留时间(等待时间与服务时间之和)有限.



(3)服务机构

- (i)服务台数量及构成形式. 从数量上说,服务台有单服务台和多服务台之分. 从构成形式上看,服务台有: ①单队---单服务台式; ②单队---多服务台并联式; ③单队---多服务台串联式; ④单队---多服务台并串联混合式; ⑤多队---多服务台并联式; 以及多队多服务台并串联混合式等等.
- (ii)服务方式. 这是指在某一时刻接受服务的顾客数,它有单个服务和成批服务两种.
- (iii) 服务时间的分布.在多数情况下,对每一个顾客的服务时间是一随机变量.



4.排队系统的主要数量指标

排队论主要研究与排队有关的数量指标的概率规律性;系统的优化问题;统计推断,根据资料合理建立模型。目的是正确设计和有效运行各个服务系统,使之发挥最佳效益。所以必须确定判断系统运行优劣的基本数量指标.

(1)排队系统主要数量指标

等待时间、忙期、队长(queue length).



- (i)等待时间 从顾客到达时刻起到他开始接受服务止这段时间 称为等待时间.等待时间是个随机变量.从顾客到达时刻起到他接 受服务完成止这段时间称为逗留时间,也是随机变量.
- (ii) 忙期 忙期是指从顾客到达空闲着的服务机构起,到服务机构再次成为空闲止的这段时间,即服务机构连续忙的时间.这是个随机变量,是服务员最为关心的指标,因为它关系到服务员的服务强度.与忙期相对的是闲期,即服务机构连续保持空闲的时间.在排队系统中,忙期和闲期总是交替出现的.

(iii)队长 队长是指系统中的顾客数(排队等待的顾客数与正在接受服务的顾客数之和);排队长是指系统中正在排队等待服务的顾客数.队长和排队长一般都是随机变量.



(2)数量指标的常用记号

(i)主要数量指标

 L_s ——平均队长,即稳态系统任一时刻的 所有顾客数的期望值;

 L_q ——平均等待队长,即稳态系统任一时刻等待服务的顾客数的期望值;

W——平均逗留时间,即(在任意时刻)进入 稳态系统的顾客逗留时间的期望值;

 W_q ——平均等待时间,即(在任意时刻)进入 稳态系统的顾客等待时间的期望值.



(ii)其它常用数量指标

- s ——系统中并联服务台的数目;
- λ ——平均到达率;
- 1/2——平均到达时间间隔;
 - μ ——平均服务率;
- 1/μ ——平均服务时间;
 - N ——稳态系统任一时刻的状态(即系统中 所有顾客数);
 - U——任一顾客在稳态系统中的逗留时间;
 - Q——任一顾客在稳态系统中的等待时间;



 $P_n = P(N = n)$: 稳态系统任意时刻状态为n的概率;特别当n=0时(系统中顾客数为0), P_0 即稳态系统所有服务台全部空闲的概率;

ho ——服务强度,即每个服务台单位时间内的平均服务时间,一般有 $ho = \lambda/\mu s$,这是衡量排队系统繁忙程度的重要尺度.



5. 排队系统的符号描述

描述符号: X/Y/Z/A/B/C

X—顾客相继到达的间隔时间的分布; 常用下 列符号:

M——表示到达的过程为泊松过程或负指数分布;

D——表示定长输入;

 E_k ——表示 K 阶爱尔朗分布;

G——表示一般相互独立的随机分布.

Y一服务时间的分布; 所用符号与X相同.



描述符号: X/Y/Z/A/B/C

Z—服务台个数; "1"表示单个服务台, "s" (s>1) 表示多个服务台.

A—系统容量限制(默认为 ∞);如系统有K个等待位子,则0<K< ∞ ,当K=0时,说明系统不允许等待,即为损失制.K= ∞ 时为等待制系统,此时一般 ∞ 省略不写.K为有限整数时,表示为混合制系统.

B一顾客源数目(默认为∞);分有限与无限两种,∞表示顾客源无限,一般∞也可省略不写.

C-服务规则, 常用下列符号:

FCFS: 表示先到先服务的排队规则; LCFS: 表示后到先服务的排队规则; PR: 表示优先权服务的排队规则。



例如:某排队问题为 $M/M/S/\infty/\infty$ /FCFS,

则表示顾客到达间隔时间为负指数分布(泊松流);

服务时间为负指数分布;有s(s>1)个服务台;

系统等待空间容量无限(等待制);

顾客源无限,采用先到先服务规则.

通常排队问题仅用上述表达形式中的前3个符号.

例如,某排队问题为M/M/S.

如不特别说明则均理解为系统等待空间容量无限; 顾客源无限,先到先服务,单个服务的等待制系统.

€ 陕西斜板大学

泊松流

在排队论中,经常用Poisson流描述顾客到达规律。

设 N(t) 表示在时间区间 [0,t) 内到达的顾客数, 若满足:

- (1) 独立性: 在任意两个不相交区间内顾客到达数相互独立:
- (2) 平稳性: 在 $[t,t+\Delta t)$ 内有一个顾客到达的概率 为 $\lambda \Delta t + o(\Delta t)$;
- (3) 普通性: 在 $[t,t+\Delta t)$ 内多于一个顾客到达的概率为 $o(\Delta t)$;

则 N(t) 称为Poisson流。



定理1 N(t) 为Poisson流的充分必要条件是:

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

定理2 到达过程为参数为λ的Poisson流的充分必要条件是: 顾客相继到达的时间间隔服从相互独立的参数为λ的负指数分布。

定理1说明,若顾客到达为Poisson流,则到达顾客数的分布恰为 Poisson分布。但是,分析顾客到达情况很不方便,实际中比较 容易得到和进行分析的是顾客相继到达的时间间隔。

定理2说明,顾客相继到达的时间间隔服从相互独立的参数为λ的负指数分布,与到达过程为参数为λ的Poisson流是等价的。



三、M/M/S排队论模型

M/M/s排队模型是指s个服务员的排队系统,顾客到来间隔时间服从参数为 λ 的负指数分布,服务时间服从参数为 μ 的负指数分布,且有隐含假定:

- •顾客到来间隔时间是独立同分布的:
- •服务时间也是独立同分布的;
- •并且独立于输入过程;
- •排队规则是等待制;



按排队论的基本构成特征,来求解该排队模型 的主要数量指标:

(1)基本构成

(i) 顾客到达规律

X(t) 表示在 $(t,t+\Delta t)$ 时间到达的顾客数,称为排队系统的输入过程.

它服从参数为 λt 的泊松分布,即:

$$P(X(t) = k) = \frac{(\lambda t)^k}{k!} e^{-\lambda t}, k = 0, 1, 2 \cdots$$

其平均值为 λt ,即单位时间内到达的顾客数为 λ ,称之为平均到达率。



 $\{s_n \mid s_n = \tau_{n+1} - \tau_n\}$ 表示顾客到达间隔时间序列,其中 τ_n 表示第n个顾客的到来时刻.

定理2表明: X(t) 服从参数 λt 为的泊松分布的充要条件是到达间隔时间序列 $\{s_n\}$ 独立同分布且服从负指数分布.

(ii) 服务时间

记Z为服务时间, Z服从参数为 μ 的负指数分布:

$$P(Z \le t) = \begin{cases} 1 - e^{-\mu t} & t \ge 0 \\ 0 & t \le 0 \end{cases}$$

则 $EZ=1/\mu$ 即为每个顾客平均服务时间为 $1/\mu$,从而单位时间内被服务的顾客的平均数为 μ ,称为<mark>平均服务率</mark> .

(iii) 排队规则

按顾客的到达的先后顺序服务,即先到先服务. 满足以上三个条件的模型在排队论中记为 M/M/s模型, 其中 s 为服务员的个数.

(2)数量特征(只讨论s=1情形)

(i) 平均队长
$$L_s = \frac{\rho}{1-\rho} = \frac{\lambda}{\mu-\lambda}$$

稳态下系统内等待服务的顾客数,其数学期望称为**平均等待队长,即**

$$L_q = \frac{\rho^2}{1-\rho} \quad (其中 \rho = \frac{\lambda}{\mu} (\rho < 1) 称为服务强度.)$$



(ii)平均逗留时间和平均等待时间

平均逗留时间为
$$W_s = \frac{1}{\mu - \lambda} = \frac{L_s}{\lambda}$$
 平均等待时间为 $W_q = \frac{1}{\mu - \lambda} - \frac{1}{\mu} = \frac{\lambda}{\mu(\mu - \lambda)}$

公式
$$L_s = \lambda W_s$$
 $L_q = \lambda W_q$
$$L_s = L_q + \frac{\lambda}{\mu} \quad W_s = W_q + \frac{1}{\mu}$$

称为Little公式.



例1.某医院急诊室同时只能诊治一个病人,诊治时间 服从指数分布,每个病人平均需要15分钟。病人按泊 松分布到达,平均每小时到达3人。试对此排队系统 进行分析.

解 对此排队系统分析如下:

① 先确定参数值: 这是单服务系统有,

平均到达率 $\lambda = 3 \text{ / h}$

平均服务率 $\mu = 60/15$ 人/h=4人/h



故服务强度为: $\rho = \frac{\lambda}{\mu} = 0.75$

② 计算稳态概率:

$$P_0 = 1 - \rho = 1 - 0.75 = 0.25$$

这就是急诊室空闲的概率,也是病人不必等待立即就能就诊的概率.

而病人需要等待的概率则为

$$P = 1 - P_0 = 0.75$$

这也是急诊室繁忙的概率.



③ 计算系统主要工作指标

急诊室内外的病人平均数: $L_s = \frac{\lambda}{\mu - \lambda} = \frac{3}{4 - 3}$ 人 = 3人

急诊室外排队等待的病人平均数:

$$L_a = L_s \rho = 3 \times 0.75 \text{ } = 2.25 \text{ } \text{ }$$

病人在急诊室内外平均逗留时间:

$$W_s = \frac{1}{\mu - \lambda} = \frac{1}{4 - 3}h = 1h = 60$$
 min

病人平均等候时间:

$$W_q = \rho W_s = 0.75h = 45 \text{min}$$



④ 为使病人平均逗留时间不超过半小时,那么平均 服务时间应减少多少?

由于
$$W_s = \frac{1}{\mu - \lambda} \le \frac{1}{2}$$

代入
$$\lambda=3$$
 解得: $\mu \geq 5$

平均服务时间为:
$$\frac{1}{\mu} \le \frac{1}{5}h = 12 \min$$

15-12=3min

即平均服务时间至少应减少3min.



⑤ 若医院希望候诊的病人90% 以上都能有座位,则候诊室至少应安置多少座位?

设应该安置*x* 个座位, 加上急诊室的一个座位, 共有*x*+1个。要使90% 以上的候诊病人有座位, 相当于使"来诊的病人数不多于 *x*+1个"的概率 不少于90%, 即

$$P(N \le x+1) = 1 - P(N > x+1) \ge 0.9$$

$$\therefore P(N > x+1) \le 0.1$$



$$P(N \le x+1) = 1 - P(N > x+1) \ge 0.9$$

$$\therefore P(N > x + 1) \le 0.1$$

$$P(N > x + 1) = 1 - \sum_{n=0}^{x+1} P_n = 1 - \sum_{n=0}^{x+1} (1 - \rho) \rho^n$$

$$= 1 - (1 - \rho) \sum_{n=0}^{x+1} \rho^n = 1 - (1 - \rho) \cdot \frac{1 - \rho^{x+2}}{1 - \rho} = \rho^{x+2}$$

$$\therefore \rho^{x+2} \leq 0.1$$

上式两边取对数

$$(x+2)\log\rho \leq \log 0.1$$

因为 ρ <1 故

$$x + 2 \ge \frac{\log 0.1}{\log \rho} = \frac{-1}{\log 0.75} = 8$$

所以 $x \ge 6$

即候诊室至少应安置6个座位.



在平衡条件下系统中顾客数为n的概率

$$P_n = (1 - \rho)\rho^n$$
, $n = 0, 1, 2, 3, \dots$

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

平均等待队长
$$L_q = \rho L_s = \rho \frac{\lambda}{\mu - \lambda} = \frac{\rho^2}{1 - \rho}$$

(其中
$$\rho = \frac{\lambda}{\mu} (\rho < 1)$$
 称为**服务强度**)

平均逗留时间为 $W_s = \frac{1}{\mu - \lambda}$

平均等待时间为
$$W_q = \rho W_s = \rho \frac{1}{\mu - \lambda}$$

顾客在系统中的逗留时间T,服从参数为 $\mu - \lambda$ 的负指数分布,即有

$$P\{T>t\}=e^{-(\mu-\lambda)t}$$
 , $t\geq 0$



例2 承接例1,假设医院增强急诊室的服务能力,使 其同时能诊治两个病人,且平均服务率相同,试分析 该系统工作情况,并且,例1、例2的结果进行比较。

解 这相当于增加了一个服务台,故有:

$$s = 2 \qquad \lambda = 3 \text{ J/h} \qquad \mu = 4 \text{ J/h}$$

$$\rho = \frac{\lambda}{\mu} = 0.75, \qquad \rho^* = \frac{\lambda}{S\mu} = \frac{3}{2 \times 4} = 0.375$$

$$P_0 = \left[1 + 0.75 + \frac{(0.75)^2}{2!(1 - 0.375)}\right]^{-1} = \frac{1}{2.2} = \frac{5}{11} = 0.45$$



病人必须等候的概率,即系统状态N≥2的概率:

$$P(N \ge 2) = \frac{(0.75)^2}{2! (1 - 0.375)} \times \frac{5}{11} \approx 0.20$$



两个系统的比较

指标	s=1 系统	s=2 系统
P(Q>0)	0.75	0.20
等待队长 L_q	2.25 人	0.12 人
队长L。	3 人	0.87 人
逗留时间 W_s	60 min	17.4 min
等待时间 W_q	45 min	2.4 min
系统空闲概率 P ₀	0.25	0.45



M/M/1排队模型练习

练习1. 某车间的工具仓库只有一个管理员,平均有4人/h来领工具,到达过程为Poisson流;领工具的时间服从负指数分布,平均为6min。试求

- (1) 仓库内没有人领工具的概率
- (2) 仓库内领工具的工人的平均数
- (3) 排队等待领工具的工人的平均数
- (4) 工人在系统中的平均花费时间
- (5) 工人平均排队时间



解:本题属于M/M/1系统

$$\lambda = 4$$
 , $\mu = \frac{60}{6} = 10$, $\rho = \frac{\lambda}{\mu} = 0.4$

(1) 仓库内没有人领工具的概率

$$P_0 = 1 - \rho = 1 - 0.4 = 0.6$$

(2) 仓库内领工具的工人的平均数

$$L_s = \frac{\rho}{1-\rho} = \frac{0.4}{1-0.4} = \frac{2}{3}(\text{L})$$

(3) 排队等待领工具的工人的平均数

$$L_q = \frac{\rho^2}{1-\rho} = \frac{4}{15}(\text{L})$$

(4) 工人在系统中的平均花费时间

$$W_s = \frac{L_s}{\lambda} = \frac{2}{3} \cdot \frac{1}{4} = 10 \text{(min)}$$

(5) 工人平均排队时间

$$W_q = \frac{L_q}{\lambda} = \frac{4}{15} \cdot \frac{1}{4} = 4(\min)$$



- (6) 仓库内恰有3人来领工具的概率
- (7) 仓库内至少有1个人的概率
- (8) 顾客在仓库内逗留时间超过10min的概率

解: (6) 仓库内恰有3人来领工具的概率为

$$P_3 = (1 - \rho)\rho^3 = (1 - \frac{2}{5})(\frac{2}{5})^3 = 0.038$$

(7) 仓库内至少有1个人的概率为

$$P\{N \ge 1\} = 1 - P_0 = \rho = \frac{2}{5} = 0.4$$

- (6) 仓库内恰有3人来领工具的概率
- (7) 仓库内至少有1个人的概率
- (8) 顾客在仓库内逗留时间超过10min的概率

解: (8) 顾客在仓库内逗留时间超过10min的概率

顾客在系统中的逗留时间T,服从参数为 $\mu-\lambda$ 的负指数分布 ,即有

$$P\{T>t\}=e^{-(\mu-\lambda)t}$$
 $t\geq 0$

$$\therefore P\{T > 10\} = e^{-\frac{1}{6}(10-4)} = e^{-1} = 0.3679$$



练习2. 某单人理发店顾客到达为Poisson流,平均到达间隔为20分钟,理发时间服从负指数分布,平均为15分钟。求:

- (1) 顾客来理发不必等待的概率
- (2) 理发店内顾客平均数
- (3) 顾客在理发店内平均逗留时间
- (4) 若顾客在店内平均逗留时间超过1.25小时,则店主将考虑增加设备及理发师,那么平均到达率提高多少时店主才会考虑增加呢?

解: 本题属于M/M/1系统

$$\lambda = 3$$
 , $\mu = \frac{60}{15} = 4$, $\rho = \frac{\lambda}{\mu} = \frac{3}{4}$

(1) 顾客来理发不必等待的概率

$$p_0 = 1 - \rho = 1 - 0.75 = 0.25$$

(2) 理发店内顾客平均数

$$L_s = \frac{\rho}{1-\rho} = \frac{\lambda}{\mu - \lambda} = \frac{3}{4-3} = 3(\lambda)$$



(3) 顾客在理发店内平均逗留时间

$$W_s = \frac{1}{\mu - \lambda} = \frac{1}{4 - 3} = 1(小时)$$

(4) 由
$$W_s = \frac{1}{\mu - \lambda} > 1.25$$
 可得

$$\lambda > 3.2$$

$$\therefore$$
 3.2-3=0.2(人/小时)

平均到达率提高0.2人/小时,店主才会考虑增加.



忙期和闲期

在平衡状态下,忙期B和闲期I一般均为随机变量,求它们的分布比较麻烦,因此我们一般考虑平均忙期 \overline{B} 和平均闲期 \overline{I}

由于忙期和闲期出现的概率分布为 ρ 和 $1-\rho$,所以在一段时间内可以认为忙期和闲期的总长度之比为 ρ : $1-\rho$



又因为忙期和闲期是交替出现的, 所以在充分长的时间 里, 它们出现的平均次数应是相同的。

于是,忙期的平均长度 \overline{B} 和闲期的平均长度 \overline{I} 之比也 应是 $\rho: 1-\rho$

$$\frac{\overline{B}}{\overline{I}} = \frac{\rho}{1 - \rho}$$

$$\frac{\overline{B}}{\overline{I}} = \frac{\rho}{1 - \rho}$$

又因为到达为Poisson流时,根据负指数分布的无记忆性和 到达与服务相互独立的假设,可以证明从系统空闲时刻起 到下一个顾客到达时刻止(即闲期)的时间间隔也服从参 数为λ的负指数分布,且与到达时间间隔相互独立。

因此,平均闲期应为 $\frac{1}{\lambda}$

$$\overline{B} = \frac{\rho}{1-\rho} \cdot \overline{I} = \frac{\rho}{1-\rho} \cdot \frac{1}{\lambda} = \frac{1}{\mu-\lambda}$$

美西种核大学

$$\overline{B} = \frac{\rho}{1 - \rho} \cdot \overline{I} = \frac{\rho}{1 - \rho} \cdot \frac{1}{\lambda} = \frac{1}{\mu - \lambda}$$

$$W_{s} = \frac{L}{\lambda} = \frac{1}{\mu - \lambda}$$

比较上面两式可知,平均逗留时间 = 平均忙期

这一结果是显然的,顾客在系统中逗留的时间越长,服务员连续忙的时间也就越长。因此,一个顾客在系统内的平均逗留时间应等于服务员平均连续忙的时间。

