

最新精品软件普及丛书



Visual FoxPro

中文版易学易用专辑

5.0

法篇

- 实用培训教程系列
- 最新精品软件
- 有六大特点
- 帮助你步入人生的成功坦途
- 帮助你实现业务的梦幻构想

● 宜晨 主编 ● 品辰工作室 路可 蒋啸立 编著



电子工业出版社

PUBLISHING HOUSE OF ELECTRONIC INDUSTRY
http://www.phei.com.cn

内 容 简 介

本书介绍用 GPSS 语言进行离散事件系统仿真的编程方法。这种仿真技术在生产管理、仓库管理、设备维修、商业营销系统、公共服务事业、港口设计等领域有广泛应用。

本书可作为大专院校管理类专业、计算机应用专业的参考书,也可供有关工程技术人员参考。

管理系统仿真 (GPSS 语言)

贾兰香 陈宝谦 编
荆洪刚 杨兴宴

*

天津大学出版社出版

(天津大学内)

邮编:300072

河北省永清第一胶印印刷厂印刷

新华书店天津发行所发行

*

开本:787×1092 毫米¹/₁₆ 印张:11³/₄ 字数:294 千

1997 年 7 月第一版 1997 年 7 月第一次印刷

印数:1—1000

ISBN 7-5618-0966-2

F·116 定价:13.00 元

前 言

仿真又称作模拟,这个术语译自英文 Simulation。仿真是一项新技术,它的含义简要说明如下:人们要研究一个现实的系统模型或设想中的系统模型,由于模型比较复杂,还常常带有不确定性因素(概率因素),因此难于用解析方法求得定量的解答,于是人们便设计一个与原模型在逻辑上完全相当的物理模型或数学模型,这样设计出的模型称作仿真模型。对仿真模型进行实验或抽样研究,得出系统的某些特性和指标,便把它们作为现实系统或设想系统的特性和指标,这就是仿真。由于所研究系统的多样性,所以仿真技术和方法也呈现出多样性。

本书介绍的管理系统仿真,是指利用计算机对离散事件系统的仿真。这类系统的特点是:首先,系统是动态的,即系统的状态随着时间的进程而变化,并且只是在离散的时间点上改变;同时系统中含有随机因素,多工序生产流水线是这类系统的一个简单例子。离散事件系统仿真在生产管理、公共服务事业、交通运输、港口管理等部门有广泛应用。

仿真语言 GPSS(General Purpose Simulation System)是用于离散事件系统仿真的一种专门计算机语言。它有几种版本,本书全面介绍了微机版本 GPSS/PC 的编程方法。

本书的初稿,曾作为讲义,在南开大学国际企业管理系讲授过多次。此次出版,作了适当修改。

本书可作为大学管理专业、计算机专业、运筹学等相关专业的教材或教学参考书,也可供有关的工程技术人员参考。

编者

1996.9

目 录

第一章 离散系统仿真方法概述	(1)
§ 1 引言	(1)
§ 2 几个例子	(1)
§ 3 手工仿真方法	(4)
§ 4 系统仿真在计算机上的实现	(7)
§ 5 随机数的产生	(7)
§ 6 随机事件和随机变量的抽样.....	(10)
§ 7 关于仿真结果的可靠性和精度估计.....	(12)
第二章 仿真语言 GPSS 编程方法(一)	(13)
§ 1 多工序生产模型的仿真.....	(13)
§ 2 输出报表及其计算方法.....	(20)
§ 3 多个加工设备与多个输入.....	(22)
§ 4 随机变量的抽样、函数	(29)
§ 5 GPSS 的一些基本概念	(33)
§ 6 优先权.....	(36)
§ 7 仿真程序在计算机中的执行过程.....	(41)
§ 8 语句的重新定义.....	(43)
第二章习题	(47)
第三章 仿真语言 GPSS 编程方法(二)	(49)
§ 1 保留值与变量.....	(49)
§ 2 检测语句.....	(53)
§ 3 表.....	(58)
§ 4 选择语句.....	(64)
§ 5 矩阵.....	(67)
§ 6 分裂、装配与匹配	(70)
§ 7 循环、缓冲、强占.....	(75)
§ 8 用户链.....	(84)
§ 9 种子、实体个数限制	(87)
第三章习题	(89)
第四章 应用实例选编	(91)
§ 1 生产管理问题.....	(91)
§ 2 库存管理问题.....	(96)
§ 3 交通运输问题	(108)
§ 4 服务窗口问题	(127)

§ 5 工程管理问题	(138)
附录一 上机操作	(146)
附录二 GPSS/PC 语句规则和功能	(150)
§ 1 语句的一般形式和标准字符	(150)
§ 2 块语句格式	(152)
§ 3 控制语句	(165)
§ 4 定义语句	(167)
附录三 概率统计基础知识	(174)
§ 1 概率论基础知识	(174)
§ 2 数理统计基础知识	(177)
参考文献	(179)

第一章 离散系统仿真方法概述

§ 1 引言

仿真也称模拟,它是 Simulation 的译名。仿真是系统分析的一种手段或方法。对于不同类型的问题,仿真方法和内容可能有很大的不同,从这种意义上讲仿真的含义是比较广泛的,因而也是不很确定的。本书介绍的仿真是指对离散系统的仿真。关于离散系统的概念和特点,我们在介绍了几个实例之后再来说明。

在工业管理、交通运输、公共服务事业等许多部门中提出的系统分析问题,其系统模型通常都比较复杂,系统中包含着多个环节和成份,与系统相联系的数量有些是随机变量,系统本身还是动态的,即系统随着时间的变化而变化。所以尽管这类模型本质上也是数学模型,但是现有的数学理论难于对这类复杂系统进行分析。在这种情况下,仿真是唯一可行的一种定量分析方法。

仿真的思想是比较古老的,但由于仿真方法本身有很大的运算工作量,所以只有在电子计算机出现之后系统仿真才能在实际上付诸应用。

本章介绍仿真方法的原理。在这里除了涉及到概率论与数理统计的初步知识之外,不需要其他方面的数学知识。

用仿真方法作系统分析,要经过这样几个工作步骤:首先根据现实的系统和分析的要求建立系统模型;第二步是编制计算机程序并进行计算;最后对计算结果进行分析评价。如果发现计算结果不甚合理,这大都是因为模型不完善引起的。在这种情况下应修改模型并重新计算。系统建模没有成文的规则可以遵循,主要是靠经验和技巧。但从原则上说,应避免两种错误倾向:一种倾向是对现实问题作过分的简化,致使模型本身不能很好地反映实际问题,这样计算结果的可靠性就很差了;另一种倾向是在模型中保留了过多的繁琐的细节,而从逻辑上说其中的许多细节是不必要的,这样由于模型的复杂导致了编制程序和计算上的困难,而计算结果的可靠性并未因此而提高。

系统仿真的计算机程序,虽然可以用 FORTRAN、PASCAL 和 BASIC 等通用语言来编写,但这是比较艰难的工作。正因为如此,自 60 年代以后出现了几种用于仿真问题的计算机语言。其中比较流行的 GPSS 是适用于离散系统的一种仿真语言。用这类专用语言编写仿真程序就容易得多了。第二、三章我们将详细介绍 GPSS 语言的程序编制方法。

§ 2 几个例子

这一节通过几个例子来说明系统仿真研究哪一类的问题。

(一)商店模型

一个理发店有四个理发师(服务员),店中为等待的顾客设置了六个座位。当一个顾客到来

时,如果有的服务员正空闲着,这个顾客便立即接受服务;如果服务员都在忙着但等待座位有空位,这个顾客便坐下来等待;如果座位无空位,这个顾客便离去。服务员完成一个顾客的服务时,如果有等待的顾客,服务员便指定一个并对他进行服务;如果没有等待的顾客,服务员便空闲下来,直到新的顾客到来,再开始进行服务。第 $k-1$ 个顾客与第 k 个顾客到来间隔时间记作 X_k ,假定这是一列独立同分布的随机变量,它们的分布函数 $F(x)$ 是已知的。服务员对第 k 个顾客服务花费的时间记作 Y_k ,假定这也是一列独立同分布的随机变量,它们的分布函数 $G(x)$ 也是已知的。

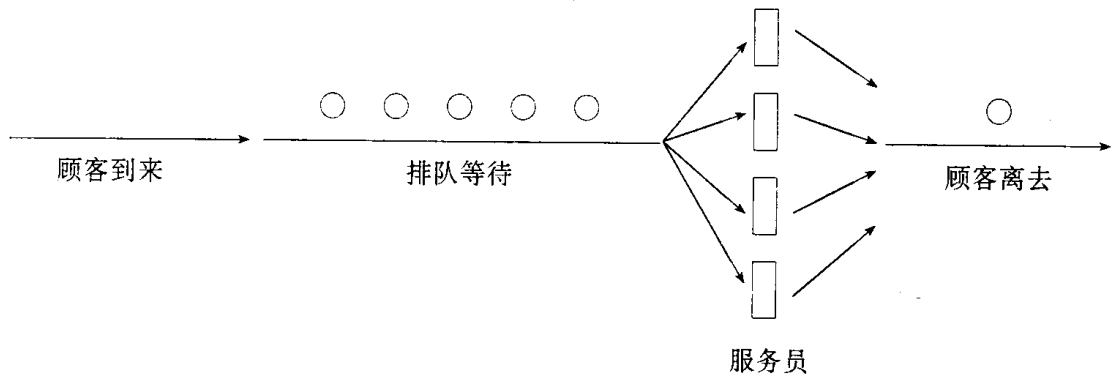


图 1-1

顾客、服务员连同以上所述的有关规则一起构成了一个系统,如图 1-1 所示。这种系统叫排队系统,也叫随机服务系统。假定一天营业八个小时,全天接待的顾客数记为 Z 。由于系统中含有随机变量,因而 Z 也是随机变量。 Z 的数值代表着商店的收入,是人们关心的一个量。现在的问题是希望求出 Z 的数学期望 $E(Z)$ 。但是用分析方法不能求出 $E(Z)$ 。如果这个模型确实符合现实生活中某个理发店的情况,那么可以使用统计抽样方法来估计 $E(Z)$ 。随机变量 Z 是由整个排队系统决定的,对 Z 抽样就是观察理发店一天的结果。现在对理发店观察了 n 天,得到每天接待顾客是 $Z_1, Z_2, \dots, Z_n$ 。于是用样本均值 $\bar{Z} = \frac{1}{n} \sum_{i=1}^n Z_i$ 作为数学期望 $E(Z)$ 的估计。

如果我们考虑的模型只是一种方案,现实中并没有这个系统。比如说,原来的理发店只有三个服务员,为了增加盈利拟增加一个服务员,这样就需研究四个服务员的情形。这时,通常意义下的统计抽样方法便无法使用了。就必须使用系统仿真方法,它实质上是通过计算机来实现统计抽样的。

(二)多工序生产模型

工业生产流程很多都是多工序生产。下面我们考虑一条生产线上的 A、B 两道工序。A 工序接受上一道工序的半成品,上一工序每间隔时间 X 送出一件半成品(工件)。A 工序有一台加工机器,加工一个工件所需时间为 Y ,加工后的工件送至 B 工序再加工, B 工序有两台加工机器,每台机器加工一个工件所需时间为 Z ,加工过程如图 1-2 所示。

X, Y, Z 一般并不是常数,而应看作一个随机变量,经过统计分析,我们假定它们的分布函数是已知的。这三个随机变量的分布,反映了相应工序的生产特点。

一个班日(八小时)加工的工件数记作 L ,在工序 A、B 前等待的工件个数分别记作 Q_1, Q_2 。这 Q_1, Q_2 个工件通常要暂时放在半成品库中。 L 表示产量, Q_1, Q_2 反映占用生产流动资金的数量。我们希望了解 L, Q_1 和 Q_2 的数值。这些数值都是随机变量,利用仿真方法可以求出它们的

数学期望的近似值。

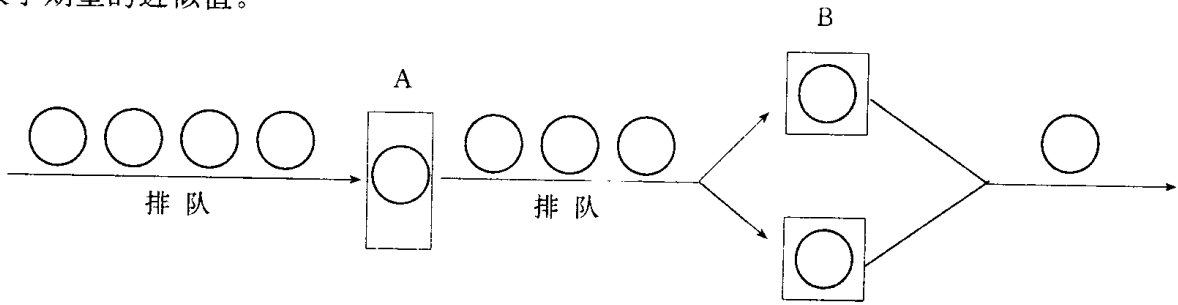


图 1-2

不同的生产工艺和不同的管理方法,表现为系统中的有关参数的改变。仿真方法对不同的方案进行研究,可以为管理决策和工艺设计提供依据。

(三)车间生产的组织问题

一个服装车间,有 50 台缝纫机,45 个缝纫工人每人使用一台。缝纫机正常工作一段时间后总会出现或大或小的故障,须维修后才能使用。车间有两个机器维修工。缝纫机连续无故障的持续工作时间即前后两次发生故障的间隔时间记为 X ,维修一台机器花费的时间记为 Y ,假定 X, Y 是随机变量,分布是已知的。模型如图 1-3 所示。

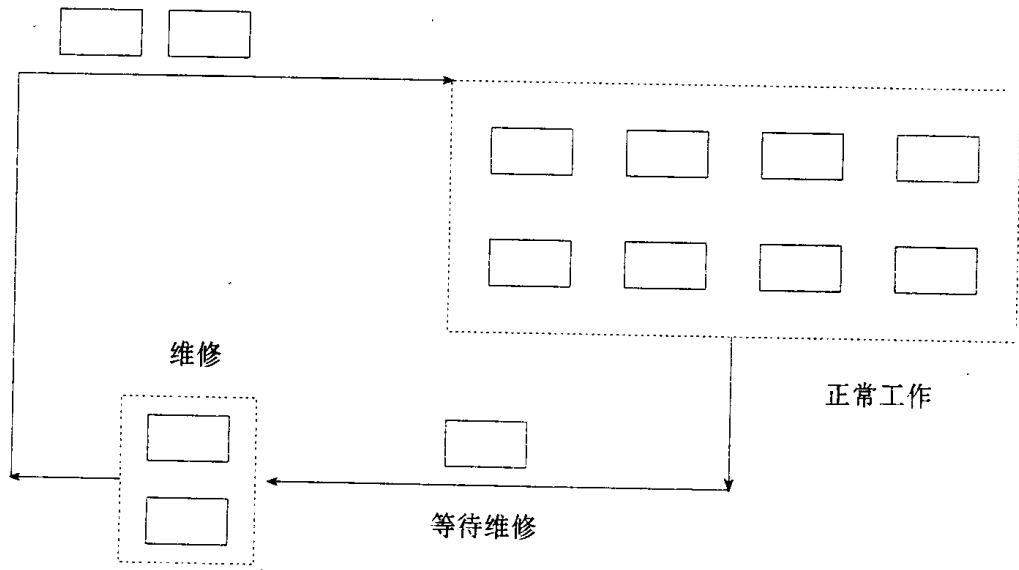


图 1-3

用仿真方法可以对这个系统的运行行为进行分析,进而可以进行经济分析。同样地,分析不同方案,便可为管理决策提供依据。

(四)离散事件系统的特点

上面三个例子有一些共同的特点,这些特点有以下四个方面:

1. 系统是动态的

系统在一定的时间范围中运行,随着时间的变化,系统的状态在不断变化。例如商店模型,在时间 t_1 等待的顾客有三个,但到时间 t_2 等待的顾客可能是五个。

2. 系统中含有随机因素

系统中有些参数是随机变量。在很多场合必须认为系统中的有关数量是随机变量而不是

常量,这样模型才能较好地符合实际情况,但同时这也增加了系统分析工作的困难,导致常规数学方法无能为力,而必须采用仿真方法。

3. 系统中包含排队环节

在商店模型中,顾客常常要排队等待服务;在多工序生产模型中,工件也常常要排队等待加工;在第三个例子中,发生故障的机器要排队等待维修,维修好的机器排队等待使用。因为系统中包含排队环节,所以这种系统叫作排队系统。

4. 在一定意义下系统具有离散性

对于商店模型,我们用服务员的忙或闲,以及排队等待的顾客个数来描述系统的状态。系统随时间的变化而变化反映为系统状态的变化。第 n 个顾客到来的时刻记为 x_n ,开始为第 n 个顾客服务的时刻记为 y_n ,为第 n 个顾客服务完的时刻记为 z_n 。 x_n, y_n, z_n 是时间轴上的一些离散点,只有当通过这些点时系统的状态才可能发生变化。比如在 z_n 前一瞬间四个服务员都在忙,并且队中有三个等待的顾客,刚完成服务的服务员立即接受队中的一个顾客并开始对其进行服务,因此在时刻 z_n 如果没有新的顾客到来,则在 z_n 稍后的瞬间,四个服务员都在忙,而等待的顾客只有两个了。这表明经过时刻 z_n 系统的状态发生了变化,等待的顾客由三个变成了两个,“一个顾客到来”,“对一个顾客开始服务”和“完成一个顾客的服务”这些都称作事件,只有当某个事件发生时系统的状态才可能发生变化,而事件只在 x_n, y_n, z_n 这些离散点上发生,所以这类系统也叫作离散事件系统,简称离散系统。

(五) 广义的顾客和服务员概念

在前面提到多工序生产模型,我们可以把工件叫作顾客,把加工机器叫作服务员,把对工件的加工叫作对顾客服务。在这个系统中顾客经过两次排队接受两种服务,这是由两个排队子系统构成的一个串联排队系统。对于第三个例子,我们可以把缝纫机叫顾客,把缝纫工和维修工叫服务员,把使用缝纫机正常工作和对缝纫机进行维修叫服务。顾客在这个系统接受两种服务,这也是由两个排队子系统构成的一个复杂系统。顾客在这个系统所走的路线成环状,系统中的顾客总数是有限的而且是不变的(50个)。这样的系统叫有限源排队系统。

今后我们在广泛的意义下使用“顾客”、“服务员”、“服务”这些术语。顾客可以是人也可以是工件或机器,服务员同样也可以有各种含义。实际上多工序生产模型也可以看作一个商店模型,其中A工序表示对顾客理发,B工序表示顾客买东西。以后我们在介绍不同类型的排队系统和仿真方法时一般都使用顾客和服务员这样的术语,只是在介绍某些应用实例时才使用该问题本身的术语。

§ 3 手工仿真方法

这一节通过两个简单例子来说明什么是系统仿真。

问题一:用八种零件各一件装配成一件成品,每件零件中都有正品和副品,从一种零件中随机地抽取一件为正品的概率是 $5/6$,为副品的概率是 $1/6$ 。在装配一件成品时所用零件若全为正品则成品是一等品,若有一个零件是副品则成品是二等品,若有两个以上零件是副品则成品是三等品,假定装配成品时不对零件进行检验。问成品的一等品率、二等品率和三等品率是多少?

约定以下随机事件:

$A_i = \left\{ \text{从第 } i \text{ 种零件中抽取一件为正品} \right\}, (i=1, 2, \dots, 8)$

$B_j = \left\{ \text{从成品中抽取一件为 } j \text{ 等品} \right\}, (j=1, 2, 3)$

这个问题是求概率 $P(B_1), P(B_2), P(B_3)$ 。

为从概念上进行比较,对这个问题我们介绍三种解法。

1. 解析解法

解析解法即理论解法。一个问题如果能够使用解析解法,必须优先考虑使用它。这种方法得出的解是精确解。

现在这个例子很简单,能够使用理论解法。根据假设条件, $A_1, A_2, \dots, A_8$ 是相互独立的事件,并且 $P(A_i) = 5/6$ 。又

$B_1 = \left\{ \text{事件 } A_1, A_2, \dots, A_8 \text{ 同时发生} \right\}$

$B_2 = \left\{ \text{事件 } A_1, A_2, \dots, A_8 \text{ 中恰有一个不发生} \right\}$

所以得

$$P(B_1) = \left(\frac{5}{6}\right)^8$$

$$P(B_2) = 8\left(\frac{1}{6}\right)\left(\frac{5}{6}\right)^7$$

$$P(B_3) = 1 - P(B_1) - P(B_2)$$

2. 统计抽样方法

当问题比较复杂,理论解法不存在时,用统计方法有时是可能的。假定我们的例子完全符合现实的生产情况。统计方法就是直接对成品进行抽样。比如我们抽查 100 件成品,而且能检验出每件是一等品、二等品或三等品。如果在这 100 件成品中有 r_1 件一等品, r_2 件二等品, r_3 件三等品,那么就认为 $P(B_j) = r_j/100$ ($j=1, 2, 3$)。

如果例子中的条件只是一种假定条件,那么上述统计方法便不能实现。比如,现实的生产情况是 $P(A_1) = \frac{4}{5}, P(A_2) = P(A_3) = \dots = P(A_8) = \frac{5}{6}$ 。企业想通过改进工艺来使 $P(A_1) = \frac{5}{6}$, 这样成品的质量水平会有所提高。改进工艺需付出一定费用。成品水平的提高又会增加收益。现在需要论证这项措施对企业是否有利,从而决定是否采取这项措施。对这种情况,上述意义下的统计抽样方法就不能实现了。

3. 仿真方法

仿真方法的要点是:(1)针对原有问题,构成一个新问题,要使新问题与原问题在数学上完全等价,只是实际背景不同,同时新问题能够实现统计抽样。(2)对新问题进行统计抽样,得出所求的解。最后即以得出的新问题的解作为原问题的解。

对于我们的例题,我们构造出以下的新问题。同时投掷八颗骰子,记:

$B_1 = \left\{ \text{八颗骰子都不出现点 } 1 \right\}$

$B_2 = \left\{ \text{八颗 中恰有一颗是 } 1 \right\}$

$B_3 = \left\{ \text{八颗 中有二个以上是 } 1 \right\}$

求 $P(B_j)$ ($j=1, 2, 3$)。

由于{掷一个骰子出现点1}与{从零件中抽出一件为副品}这两个随机事件的概率相同,都是1/6。即这两个随机事件在概率的意义下等价。容易看出新问题同原问题在数学上是完全等价的,尽管它们各自的背景不同。原问题是产品装配,新问题是掷骰子游戏。

对新问题作统计抽样是很容易的事。每次同时掷八颗骰子共掷 n 次,累计每次的情况,如果在 n 次试验中 B_1, B_2, B_3 分别出现了 s_1, s_2, s_3 次,那么便认为 $P(B_j) = s_j/n$ 。

从数理统计的观点说, s_j/n 只是概率 $P(B_j)$ 的一个估计并不能保证它们严格相等。但是当样本容量 n 足够大时, s_j/n 很好地接近于 $P(B_j)$,所以就可认为 $P(B_j) = s_j/n$ 。同时也认为这就是原问题的解。我们在这里所作的是用掷骰子试验对产品装配问题进行仿真。可见所谓仿真,就是在一个与原模型等价的假想的模型上进行统计抽样(试验)用以模仿在真实模型上的统计抽样。所以仿真方法实质上还是统计抽样方法。

须指出我们这个例子中有一个巧合。零件的正品率是5/6,这恰好可以用掷骰子来仿真。如果正品率是0.37,用掷骰子来仿真就不行了,对此以后再来考虑。

这个例子只是对随机事件仿真,下面介绍一个系统仿真的例子。

问题二:考虑一个很简单的排队问题,其模型如下:

- (1)系统有一个服务员,服务一个顾客所需时间是常数10(时间单位:分钟。下同)。
- (2)顾客到来的间隔时间 X 是随机变量,它以1/5的概率取值5,6,8,10,12。第一个顾客到来时刻与 X 同分布。
- (3)排队长度不限,先到先服务。

现要求对系统的运行过程进行仿真,截止到对第三个顾客服务完的时刻,并指出当时系统的状态,如图1-4所示。

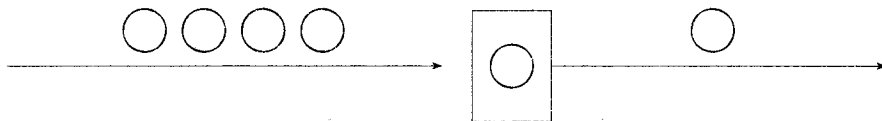


图1-4

首先考虑随机变量 X 的抽样方法。对此考虑袋子游戏:做五个形状相同的小球,各球分别写上数5,6,8,10,12。将球放入一个口袋,从袋中随机地摸出一球得到的数记为 Y ,显然随机变量 Y 与 X 同分布。利用这个袋子游戏可以对上述排队系统进行仿真。

- (1)对 X 抽样两次,比如说得8,12。这表示第一位顾客在时刻8到来,第二个在时刻20到来。
- (2)在时钟时间8~18,服务于第一个顾客。
- (3)在时间18~20,服务员空闲。
- (4)对 X 抽样,比如说得6。即第三个顾客在时刻26到来。在时间20~30服务于第二个顾客。
- (5)对 X 抽样,比如说得5,即第四个顾客在时刻31到来。在时间30~40服务于第三个顾客。第三个顾客在队中的等待时间是4。
- (6)对 X 抽样两次,比如说得8,10。即第五,六个顾客到来时刻是39,49。

在时刻40服务完了三个顾客。此时系统中还有两个顾客在排队等待,它们分别是在时刻31,39到来的。

通过以上例子看出,对于系统仿真,从作法上来说就是作两件事:第一,对于系统中的随机变量或随机事件找出抽样方法;第二,按照系统本身的规则,让系统动态地运行起来。

§ 4 系统仿真在计算机上的实现

在 § 3 我们已经初步介绍了用手工方法进行系统仿真的具体作法。利用计算机作仿真只是用计算机来执行手工方法的那些步骤。在计算机中实现随机变量的抽样,我们已在上节作了介绍。在此基础上现在来介绍计算机仿真的计算过程。考虑的模型是电视机生产线的最后一个工序——调试工序。

假定有一个调试工,调试一台电视机所需的时间或长或短有较大的差别。应将这个时间看作一个随机变量,记它为 X ,通过对历史资料的分析可以确定出 X 的分布函数 $F(x)$ 。装配完的电视机由上一道工序送到调试工序。到来间隔 Y 是随机变量,其分布函数 $G(y)$ 也是已知的。假定时间以分为单位,工人一个工班为八小时即 480 分。对系统在一个班的运行情况进行仿真。在这里,我们使用顾客和服务员这样广义的术语,通过仿真计算出以下几个数字指标:

(1)本工序一班的产量;

(2)顾客平均等待时间;

(3)服务员的服务强度。服务员有时在工作,有时可能空闲。服务强度 = $\frac{\text{累计工作时间}}{480}$;

(4)平均队长和最大队长。

下面通过一个框图(见图 1-5)说明仿真计算的具体步骤。图中 t_i 表示第 i 个顾客到来时刻, s_i 表示开始对第 i 个顾客服务的时刻, h_i 表示服务完第 i 个顾客的时刻。

通过仿真计算平均队长和最大队长其步骤比较复杂,所以框图中没有加入这部分内容。在仿真语言 GPSS 中这两项指标由系统自动计算,不需为此编写程序。

如果用通用语言(例如 FORTRAN)编写仿真程序,按照框图的流程编写即可。对于如此简单的模型,框图的流程已经比较复杂了,而实际应用问题其模型要复杂得多。可见用通用语言编写仿真程序是一项艰难的工作。

§ 5 随机数的产生

前已指出仿真方法实质上是统计抽样方法,在计算机上实现仿真,需考虑在计算机上如何实现随机事件和随机变量的抽样。在 § 5, § 6 我们将介绍有关方法。

(一)伪随机数

回忆随机变量分布拟合检验问题。设有随机变量 X ,对 X 抽样得到样本值 $x_1, x_2, \dots, x_n$ 。另有一个分布函数 $F(x)$,利用样本值 $x_1, \dots, x_n$ 与函数 $F(x)$,进行分布拟合检验,可以断定 X 的分布函数是 $F(x)$ 或不是 $F(x)$ 。

现在制订一种算法(通常都是递推算法)按这种算法产生一系列数 $x_1, x_2, \dots, x_n$,把这组数“看作”某个随机变量 X 的样本。另外已知一个分布函数 $F(x)$ 。按照分布函数拟合检验的手续进行检验,如果能得到结论: X 的分布函数是 $F(x)$,那么就可以把 $x_1, x_2, \dots, x_n$ 看作是分布函数为 $F(x)$ 的随机变量的抽样。我们就是使用这样的算法在计算机中得到随机变量的样本。

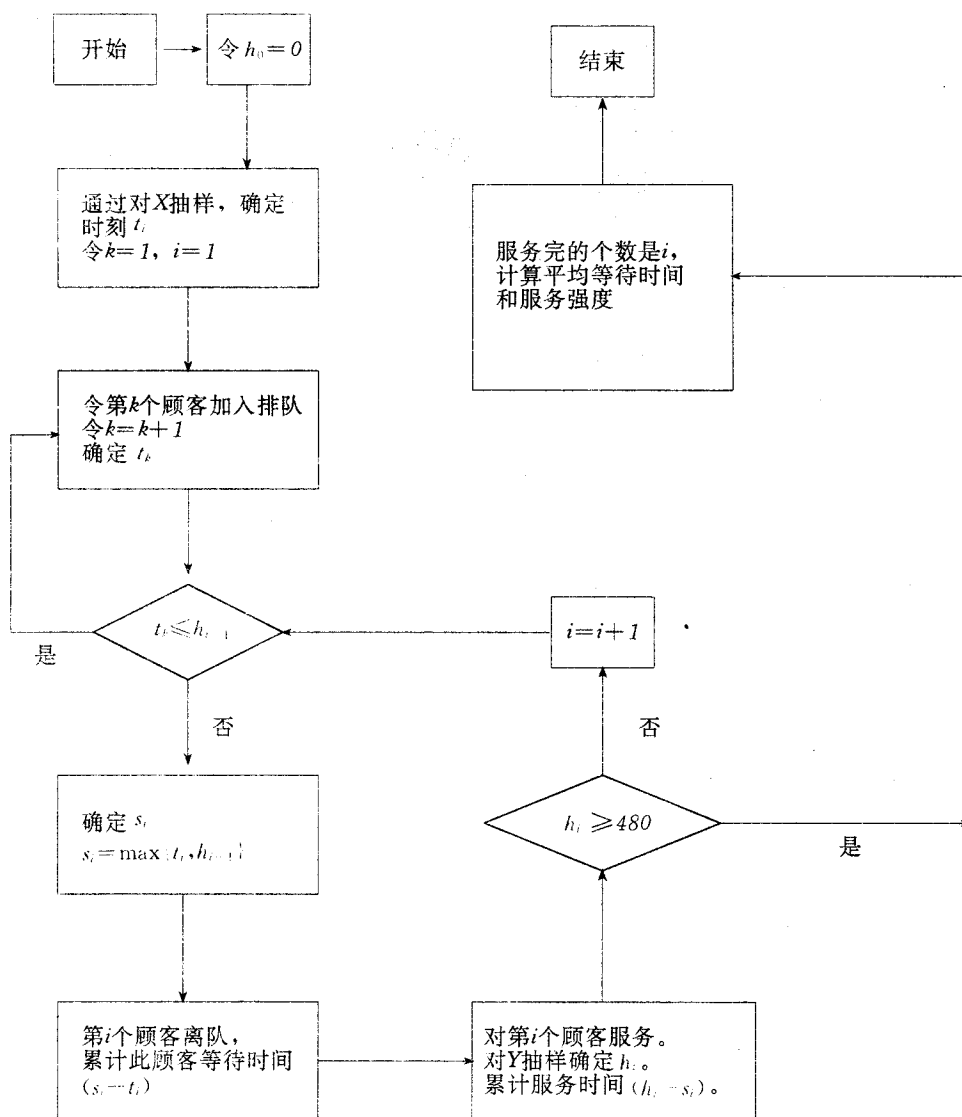


图 1-5

随机变量真正的样本值叫随机数,用人为算法产生的样本值叫伪随机数。对于伪随机数当它的算法给定之后,伪随机数 $x_1, x_2, \dots, x_n$ 中的每一个的数值就完全被确定了,不管人们是否已经计算出了它们。在这一点上来说,伪随机数与随机数有本质的不同。

但是从数理统计的观点来说,伪随机数既然已经通过了分布函数拟合检验,就可以把它们当作具有给定分布函数 $F(x)$ 的随机变量的样本。

今后我们只涉及伪随机数,故把伪字略去,一律叫随机数。

由附录三 § 1 第(五)段我们知道,有许多随机变量可以表示为 $(0,1)$ 均匀分布随机变量的函数。因此只需着重研究产生 $(0,1)$ 均匀分布随机数的算法。然后通过随机变量的变换产生其它分布的随机数。

(二)(0,1)均匀分布随机数的产生

产生随机数的方法已经有了很多,现在还是一个研究专题,仍有一些新方法在不断出现。下面介绍三种产生(0,1)均匀分布随机数的算法。

1. 乘同余法

对于字长为 32 位二进制的计算机,记 $M=2^{15}=32\,768$,取定一个质数,例如 3,5,7 等。不妨取 5,确定一个正整数 k 使满足条件 $5^k \leq 2^{15} < 5^{k+1}$,得 $k=6$ 。令 $t=5^k=15\,625$ 。产生随机数 $x_1, x_2, x_3, \dots$ 的算法如下:

$$r_0 = 1$$

$$\begin{cases} r_n = t \cdot r_{n-1} \pmod{32\,768} \\ x_n = r_n / 32\,768 \quad (n=1, 2, 3, \dots) \end{cases}$$

其中等式 $r_n = t \cdot r_{n-1} \pmod{32\,768}$ 是同余式,式中 t, r_{n-1} 是已知正整数,按此式计算 r_n ,其作法是:用 32 768 去除乘积 $t \cdot r_{n-1}$,设整数商为 a ,余 d , $0 \leq d < 32\,768$,即

$$t \cdot r_{n-1} = 32\,768a + d$$

按同余式的规则 r_n 等于余数 d 。

每选定一个 t 就得了一个算法,但是这个算法究竟是否可用,还须经过分布拟合检验。

2. 种子生成法

假定我们限制在八位正整数之内运算。

选定一个奇数 S (不超过四位)作为乘数,选取一个奇数作为 r_1 (不超过四位), r_1 叫种子。这种算法也是一种迭代算法(递推算法)。对于 $n=1, 2, 3, \dots$ 。算法步骤如下:

(1)计算乘积 $r_n \cdot S$,这个乘积的值不超过八位整数,不足八位时前边增写适当个数的 0 凑足八位。

(2)取乘积 $r_n \cdot S$ 这个八位数“中间”的四位,在四位数前边添加上小数点,作为随机数 x_n 。

(3)取乘积 $r_n \cdot S$ 的末四位作为下次使用的乘子 r_{n+1} 。

上述算法每执行一遍(1)(2)(3)产生一个随机数。例如,取乘数 $S=5\,167$,种子 $r_1=3\,729$,计算前 5 个随机数如表 1-1 所示。

表 1-1

n	r_n	$r_n S$	x_n
1	3729	19 267 743	0.2677
2	7743	40 008 081	0.0080
3	8081	41 754 527	0.7545
4	4527	23 391 009	0.3910
5	1009	05 213 503	0.2135

种子和乘数 S 的选取虽说具有很大的任意性,但是正如我们在乘同余法中已经强调的那样,每个算法经过检验之后方可使用。当然别人已经考察过的算法,我们可以直接使用不必再作检验。

3. 乘同余法的另一种形式

假定我们使用的计算机,一个整数用 32 位二进制表示。32 位中用一位表示正负号,31 位表示整数的绝对值。 $2^{31}=2\,147\,483\,648$ 。

取一个质数例如取 5, 确定 k 使满足条件 $5^k \leq 2^{31} < 5^{k+1}$, 得 $k=13, 5^{13}=1\ 220\ 703\ 125$ 。取 $S=5^{13}, r_1=1$ 。由以下步骤递推地生成 r_{n+1} 与随机数 x_n 。

(1) 计算乘积 $S \cdot r_n$, 并将它表示为

$$S \cdot r_n = a \cdot 2^{32} + d$$

其中 a 是非负整数, d 是整数, 满足条件 $-2^{31} \leq d < 2^{31}$ 。

(2) 取 $r_{n+1} = |d|$ (d 的绝对值)。

$$x_n = r_{n+1} / 2^{31}$$

乘积 $S \cdot r_n$ 的值一般会超过 2^{31} , 所以运算会出现“上溢”, 按照计算机内部运算规则, 乘积 $S \cdot r_n$ 的计算值恰好是 d 。利用这个结果, 得到一个很简单的产生随机数的算法:

取 $r_1=1$, 对 $n=1, 2, 3, \dots$

$r_{n+1} = \text{abs}(S \cdot r_n)$ (在计算机运算的意义下)

$$x_n = r_{n+1} / 2^{31}$$

§ 6 随机事件和随机变量的抽样

(一) 完备事件组的抽样

一组随机事件 $A_1, A_2, \dots, A_n$, 如果它们两两互不相容, 即 $P(A_i A_j) = 0$, 并且 $\sum_{i=1}^n P(A_i) = 1$, 那么这一组随机事件叫完备事件组。在一次试验中这组事件中必有一个发生而且只有一个发生。如何对这组事件进行抽样呢?

为简单, 设 A_1, A_2, A_3 是一个完备事件组, 并且 $P(A_1)=0.2, P(A_2)=0.1, P(A_3)=0.7$ 。又设随机变量 X 服从 $(0, 1)$ 上的均匀分布。考虑随机事件 $B_1 = \{0 \leq X \leq 0.2\}, B_2 = \{0.2 < X \leq 0.3\}, B_3 = \{0.3 < X \leq 1.0\}$ 。显然 B_1, B_2, B_3 是一个完备事件组, 并且 $P(B_1)=P(A_1), P(B_2)=P(A_2), P(B_3)=P(A_3)$ 。这样可以用 B_1, B_2, B_3 来仿真 A_1, A_2, A_3 。具体作法是产生 $(0, 1)$ 均匀分布的随机数 x_i 。

若 $0 \leq x_i \leq 0.2$, 便认为事件 A_1 发生; 若 $0.2 < x_i \leq 0.3$ 便认为 A_2 发生; 若 $0.3 < x_i \leq 1.0$ 便认为 A_3 发生。

(二) 任意分布随机变量的抽样

设随机变量 X 服从 $(0, 1)$ 上的均匀分布, 随机变量 Y 的分布函数 $F(x)$ 是已知的, 现在考虑 Y 的随机数产生方法。由附录三 § 1 知, 若 $F(x)$ 是严格单增的, 用 $F^{-1}(x)$ 表示它的反函数, 那么 $F^{-1}(X)$ 的分布函数为 $F(x)$, 由此得 Y 的随机数产生方法: 产生 $(0, 1)$ 均匀分布随机数 x_i , 令 $y_i = F^{-1}(x_i)$, 则 y_i 即随机变量 Y 的随机数。使用这种变换方法, 需要求出反函数 $F^{-1}(x)$ 的表达式。

例 1 随机变量 Y 服从参数为 λ 的指数分布, 试给出抽样算法。

Y 的分布函数为

$$F(x) = \begin{cases} 1 - e^{-\lambda x}, & x \geq 0 \\ 0, & x < 0 \end{cases}$$

在 $(0, +\infty)$ 上 $F(x)$ 的反函数是 $F^{-1}(x) = -\frac{1}{\lambda} \ln(1-x), 0 < x < 1$, 所以, 若 x_i 是 $(0, 1)$ 均匀分

布的随机数,那么 $y_i = -\frac{1}{\lambda} \ln(1-x_i)$ 便是指数分布随机变量 Y 的随机数。

例 2 设 Y 服从 (a, b) 上的均匀分布,试给出 Y 的抽样算法。

Y 的分布函数是

$$F(x) = \begin{cases} 1, & x \geq b \\ \frac{x-a}{b-a}, & a < x < b \\ 0, & x \leq a \end{cases}$$

在区间 (a, b) 上, $F(x)$ 的反函数 $F^{-1}(x)$ 是

$$F^{-1}(x) = (b-a)x + a$$

因此,若 x_i 是 $(0, 1)$ 均匀分布的随机数,那么 $y_i = (b-a)x_i + a$ 便是 (a, b) 上均匀分布的随机数。

(三)任意分布随机变量的抽样(续)

用上一段介绍的方法来产生分布函数 $F(x)$ 的随机数,必须求出反函数 $F^{-1}(x)$ 的表达式。但是有很多情况 $F^{-1}(x)$ 的表达式不能求出,例如对于正态分布就是如此。另外,在仿真语言 GPSS 中没有初等函数的运算功能,不能计算对数值 $\ln(1-x_i)$ 。所以对于指数分布的随机数便不能像例 1 那样来做。

现在介绍反函数 $F^{-1}(x)$ 的一种近似计算方法——线性插值法。仍假定 $F(x)$ 在整个数轴上或者在某一个区间上是严格单增的。在 $F(x)$ 的单增区间上选定一组分点 $t_1 < t_2 < \dots < t_n$, 计算函数值 $S_i = F(t_i)$, t_1, t_n 要满足条件 $S_1 = 0$ 或 $S_1 \approx 0$, $S_n = 1$ 或 $S_n \approx 1$ 。 (t_i, S_i) 作为函数 $F(x)$ 图形上的一个点,记作 P_i ,过这些点作一条折线(见图 1-6),这条折线代表的函数记为 $G(x)$,函数 $G(x)$ 可以作为分布函数 $F(x)$ 的近似。 $G(x)$ 是一个分段线性函数,其表达式是

$$G(x) = \frac{S_{i+1} - S_i}{t_{i+1} - t_i} (x - t_i) + S_i, (t_i \leq x \leq t_{i+1}), i = 1, 2, \dots, n-1.$$

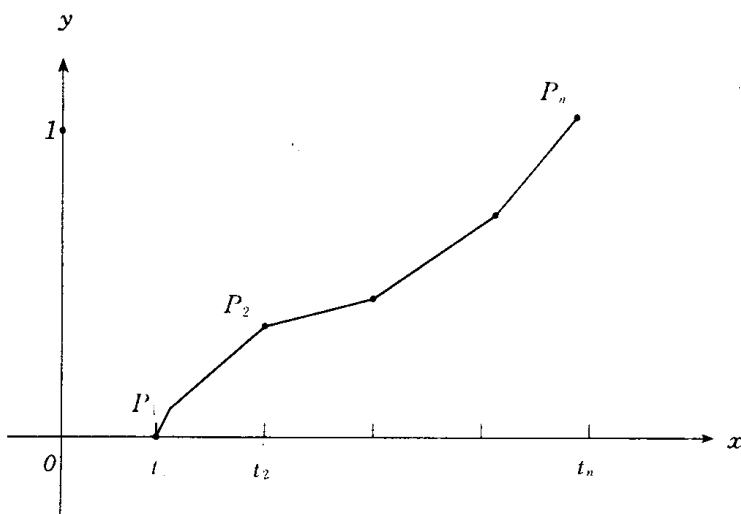


图 1-6

在 GPSS 语言中,容易实现对于分布函数为 $G(x)$ 的随机变量的抽样。

§ 7. 关于仿真结果的可靠性和精度估计

由一个实际问题建立起一个模型,在这个过程中必然经过一些简化并添加一些假定。应该注意的是经过简化建立的模型必须在主要方面能较好地反映实际情况。所作的假定也应该是合理的。仿真结果的可靠性主要取决于模型的合理性。建模工作没有现成的方法可循,主要是靠专业知识、经验和技巧以及对模型本身的深入理解。所以建立一个较好的模型需要专业人员和系统分析人员一起合作才能完成。仿真结果经分析认为不合理,这大都是模型不合理造成的,须修改后再试算。这种反复往往需要进行多次才能得出比较完善的模型和合理的仿真结果。

假定仿真模型是正确的,现在来估计仿真结果的精度。我们曾经指出仿真方法实质上是统计抽样方法。通过仿真要计算的某些量大都是随机变量。例如 § 4 模型中我们要求电视机调试工序一班的产量,这个量实际上是随机变量,将它记为 X 。确切地说我们要计算的是 X 的数学期望 $E(X)$ 。进行一次仿真计算得到的是 X 的一个样本。为要估计 $E(X)$,需进行多次抽样。这就是说需要在改变随机数的情况下进行多次仿真。假定我们这样得出了 X 的 n 个样本 $x_1, x_2, \dots, x_n$ 。于是取 $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$ 作为 $E(x)$ 的估计。现在分析这个估计的精度。

样本 $x_1, x_2, \dots, x_n$ 有两重含义。从实际计算上说,它们都是已知的具体的数值;另外,在抽样之前不能断定某次抽样必定得到某个数值,因此从理论概念上说又需把样本 x_i 看作一个随机变量,并且它与母体 X 是同分布的。

记 X 的方差 $D(X)$ 为 σ^2 , 根据中心极限定理,当 $n \rightarrow \infty$ 时有

$$P\left\{\frac{\bar{x} - E(X)}{\sigma/\sqrt{n}} \leq x\right\} \rightarrow \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{u^2}{2}} du$$

所以当 n 适当大时近似地有

$$P\left\{\frac{|\bar{x} - E(X)|}{\sigma/\sqrt{n}} \leq 2\right\} \approx \frac{1}{\sqrt{2\pi}} \int_{-2}^2 e^{-\frac{u^2}{2}} du = 0.95$$

这就是说以 0.95 的概率使下式成立

$$|\bar{x} - E(X)| \leq \frac{2\sigma}{\sqrt{n}}$$

记 $\epsilon = 2\sigma/\sqrt{n}$, 当 n 很大时 ϵ 很小, ϵ 表示精度。上式表明以 0.95 的概率使下式成立

$$|\bar{x} - E(X)| < \epsilon$$

此式给出了用 $\bar{x}$ 代替 $E(X)$ 时的精度估计。这个精度估计式只是一种理论分析,在实际应用时人们常常不考虑这个问题。