

文章编号:1007-6735(2012)03-0252-05

银行客户到访的非泊松性

李 超, 郭进利

(上海理工大学 管理学院, 上海 200093)

摘要: 通过分析某银行网点 19 天所有客户到访的真实记录,挖掘银行客户到访的间隔时间分布、单日客户到访问隔时间分布、细分客户到访问隔时间分布等统计特征,发现银行客户的到访蕴含着胖尾的统计规律,并不是先前排队论假设的泊松过程。统计结果显示,人类访问银行的行为具有明显偏离泊松分布的胖尾特性,幂指数在 2~3 的范围内。这一结果为针对幂律事件间隔分布的排队理论的建立奠定了实证基础,为下一步银行排队问题的分析作了探索。

关键词: 人类动力学; 幂律分布; 排队论; 间隔时间分布; 阵发性

中图分类号: N 941 **文献标志码:** A

Non-Poisson Properties of Bank Customers Visiting

LI Chao, GUO Jin-li

(Business School, University of Shanghai for Science and Technology, Shanghai 200093, China)

Abstract: Through analyzing the real data of bank customers visiting in a bank outlet in 19 days, it was discovered that there were many forms of non-Poisson characters appearing in the data of bank customers visiting, such as inter-arrival time distribution of all customers in 19 days, total customers visiting in one day, customers visiting for personal banking business in 19 days and customers visiting for corporate banking business in 19 days. These characters were different from those in the hypothesis of queuing theory: the coming of customers could be well approximated by Poisson processes. These distributions denoted the pattern of bank customers visiting follows non-Poisson statistics or heavy tailed distribution. It is found that most of these distribution exponents are between 2 and 3. This result establishes the empirical foundation of new queuing theory with power-law inter-arrival time distribution, and also explores the analysis of bank queuing in further studies.

Key words: human dynamics; power-law distribution; queuing theory; inter-arrival time distribution; bursts

收稿日期: 2012-02-20
基金项目: 国家自然科学基金资助项目(70871082);上海市重点学科建设资助项目(S30504);上海市研究生创新基金资助项目(JWCXSL1022)
作者简介: 李 超(1985-),男,硕士研究生.研究方向:人类行为动力学. E-mail: haluckha@sina.com
郭进利(联系人),男,教授.研究方向:人类行为动力学、工业工程、供应链管理、质量管理. E-mail: phd5816@163.com

人每天要做很多事情,包括收发电子邮件、前往银行处理金融事务等.除了明显的每日或季节性周期以外,找到人类动力学的规律似乎是困难的.在过去的 100 年中,对人类行为的研究一直是社会学、心理学、经济学等学科共同关注的问题^[1],但人类自身的复杂性和多样性,对一切科学的尝试来说都是巨大的挑战.事实上,过去学科的绝大多数命题和结论都是定性描述^[2].早前的研究中,由于数据量比较小,且人类行为具有很多随机性,因此均假设人类行为的发生是随机和稳态的,即可以看作一个泊松过程.此模型中,对于某给定的个体,连续两次行为之间的间隔时间分布是一个指数分布^[3].Barabási 等从记录人类活动历史的数据库中挖掘出人类行为的统计规律,发现人类行为所呈现出的规律并不与泊松过程一致:很长时间内可能了无一事,而这些长长的空白与空白之间则被阵发的密集活动所填充^[4],呈现胖尾分布,且幂指数大多分布在1~3之间.文献[5]给出评论时间间隔分布的实证研究,发现人们对于某个话题的兴趣逐渐消失,评论的时间间隔服从幂律分布,提出了一个兴趣逐渐消失的人类行为动力学模型.文献[6]对钱学森的邮件通讯记录进行了统计,统计结果表明其邮件回复时间间隔分布和写信时间间隔分布均具有幂律的胖尾规律,但已经超出了普适类的范围.文献[7]从群体和个人两个角度,分析借阅过程中图书的借阅时间和用户连续两次借阅或还回图书的时间间隔等统计特征,探寻群体行为与个体行为的关系.

访问银行办理业务,是人的行为之一.近来,银行客户排队问题越来越受到各方的关注,国家监管部门和各家银行纷纷出台各种措施来缓解排队等候现象.然而,国内市场研究公司 CTR 于 2007 年进行的调查表明,排队问题在大多数银行仍普遍存在,国有四大行平均等候 41 min,而在 5 家股份制银行平均等候 14 min.为分析银行排队问题,文献[8]用蒙特卡洛模拟对银行柜台多服务器单队列(M/M/C)的服务状况进行了动态模拟,提出在不同的服务强度水平下,当到达速率不同时服务对策,为改进商业银行客户服务水平提供参考.文献[9]利用银行叫号机实测队长系列排队数据,运用排队论所给出的改善银行排队状况的窗口设置设想与例证,可以作为银行排队管理改善的依据和一种简易可行的方法.目前大部分研究都用到了一个重要假设:银行客户到访是一个泊松过程.现有的研究还没有对银行客户到访问隔时间分布进行分析,本文根据某银行

网点客户到达时间的真实记录,深入挖掘银行客户到访的时间间隔分布、细分客户到访的时间间隔分布、单日所有客户到访的时间间隔分布的统计特征,发现银行客户到访的时间间隔分布并非指数分布,而是满足阵发性的幂律分布.

本文的发现有助于理解社会中人群到访的行为特征,为建立针对幂律事件间隔分布的排队理论奠定实证基础,具有重要的理论意义,同时为解决银行排队拥挤问题作出理论探索,具有一定的商业化潜力.由于社会中广泛存在类似于银行客户到访的顾客到达行为,结合原有针对社会系统的实证研究,这些统计结果暗示从复杂的人类行为中可以揭示普遍的规律,发现支配人类活动时间选择的机制,并获得控制和预测人类行为的能力.这必将具有重大的实用化潜力.

1 数据来源与统计方法

通过人工统计某月国内某银行网点大厅(营业时间为 8:30~16:30,中午保留少量窗口,周末照常营业)叫号机吐出的每张小票的数据,以每个客户的真实到达时间为研究对象.该数据库中客户到访共 10 403 人次,客户所办业务可分成 7 类,分别用 A(个人业务)、B(理财业务)、G(公司业务)、J(公司解现)、W(外币业务)、V(VIP 客户)、T(领取吞卡)作为代表.记录每个客户所得小票的打印时间,精确到秒,以此作为到达时间,跨度为某月 1 日到 19 日.各个客户所办业务的详细资料如表 1 所示,比较大的数据量足够反映真实的统计规律.为简便起见,对幂律分布的判断分析用一元线性回归和最小二乘法得到拟合直线的斜率即幂指数.

表 1 19 天中客户到达时间记录数据概况
Tab.1 Records of customer arrival time in 19 days

记录编码	编码含义	访问量/人次	平均每天访问量/人次
A	个人业务	7 239	381
B	理财业务	76	4
G	公司业务	2 329	122
J	公司解现	166	9
W	外币业务	320	17
V	VIP 贵宾	85	4
T	领取吞卡	188	10
共计	7 种	10 403	547

图1显示的是在19天的时间范围内每天各个时刻所有客户的访问量.由于一部分客户会在银行刚开门和马上要关门时进入大厅,因此统计8:00~17:00每小时的客户访问量.所有客户访问银行具有明显的规律性,每天的访问量曲线形状大致相同呈“M”形状,波峰和波谷出现在大致相同的时刻.其中,休息日的访问量要明显少于工作日.另外,银行访问还表现出明显的波动性.统计发现客户访问量的几个峰值出现在10:00~11:00、13:00~14:00两个时间段.显然,两个时间段是个人及公司前往银行办理事务的高峰期,也是人类一天中活动的活跃期.访问量的低谷出现在12:00~13:00的中午休息时间和16:00~17:00银行即将停止营业这一个时间段.这样的统计规律与客户工作生活的时间规律很好吻合.

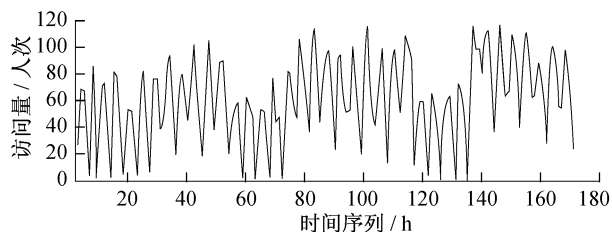


图1 19天内各个时刻的客户访问量

Fig.1 Customer visits per hour in 19 days

为消除银行晚上不营业这一明显事实对客户到访问隔时间的影响,对数据作如下处理:第二天第一个客户的到访问隔时间将减去16 h(银行从16:30到第二天8:30不营业).由于一部分客户会在8:30之前进入大厅,若作差后的间隔时间小于0,则认为该客户的到达间隔时间为1 s,这样多天的客户到访可以看成是一个连续的过程.经过对比,后文出现的概率分布图的趋势在数据处理前后都无明显改变.图2显示的是数据处理后,19天内所有客户到访问隔时间图,图中 τ 表示间隔时间.注意那些最长的像长钉一样扎在水平轴上的竖线,它们对应于那些很长的到访问隔时间(可以长达近2 h).这些最长的竖线是以周为周期出现的,并且都是出现在休息日的中午时间段,表明休息日中午几乎没有客户前往银行办理业务.而对比工作日中午银行仍有客户连续到访,但到访问量相比每天的高峰时段要明显减少.进一步观察发现,在这些长钉一般竖线之间,都有几段很小的区域几乎贴着横轴,且以天为周期出现,这是每天早晨银行开门时客户集中点击叫号机打印小票的结果,并且只维持较短的时间.由此可以判断,银行客户到访会在很长时间都没有发生,也会有在短时间内阵发.

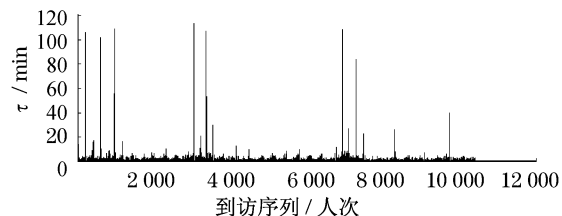


图2 19天内所有客户到访问隔时间

Fig.2 Inter-arrival time of total customers in 19 days

2 客户到访行为的统计特征

银行业务是现代社会的易于获取的服务,每个人都会隔一段时间访问银行.不同的人,访问银行的相隔时间会有不同的差异,有的人一天会访问好多次,有的人很长一段时间内都不会亲自访问银行.根据上文分析,一个网点接受该地区客户访问单位时间内的客户数量会有很大的涨落,类似情况也发生在客户到访的间隔时间上,有的时段一个客户也没有,而有的时段几秒钟就有一个客户到访.这些现象再次说明,人类的行为并不能被认为是一种泊松过程,而是大部分不同客户之间都有着某种共同的特性,经过研究,发现这种特性与泊松分布有很大的差异.本文将一段时间内所有客户或同一类型客户看成一个整体,考察客户到访银行的间隔时间统计特征,从而获得该网点客户访问银行的总体行为特征.

2.1 所有客户到访问隔时间分布

统计结果显示,银行客户到访的间隔时间分布满足 $P(\tau) \propto \tau^{-\alpha}$ (α 为幂指数),幂指数约为3,如图3所示.客户访问量的高峰一般出现在10:00~11:00、13:00~14:00点之间,访问量的低峰出现在12:00~13:00、16:00~17:00之间,这主要和该网点员工的工作时间与休息时间有关.大部分的客户到访问隔时间较短,像周末中午时间段出现很长的到访问隔时间是很少的,正是时间间隔的不均匀性造成了幂律分布的形成.

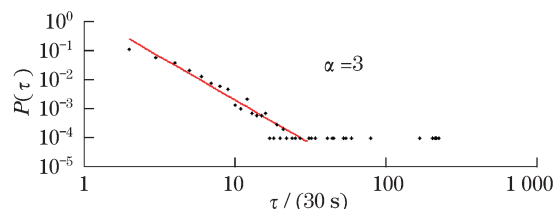


图3 双对数坐标下所有银行客户到访的间隔时间分布

Fig.3 Inter-arrival distribution of bank customers visiting in 19 days under double logarithmic coordinates

2.2 单日客户到访的间隔时间分布

图4是某个休息日和某个工作日客户到访的间隔时间分布.统计结果显示,客户到访的间隔时间分布也符合幂律分布,而幂指数分别约为2.45和2.5.该休息日的客户访问量为417人次,其中个人业务392人次,公司业务3人次,其它22人次;该工作日的客户访问量为620人次,其中个人业务364人次,公司业务203人次,其它53人次.而休息日客户到访平均间隔时间为69 s,工作日客户到访平均间隔时间为

47 s.休息日最大到访问隔时间为6 060 s,工作日最大到访问隔时间为597 s,两天最小到访问隔时间皆为1 s.客户访问量、客户结构、客户到访平均间隔时间等都有较大差异,而幂指数十分相近.接着计算数据库中每一天的客户到访问隔时间分布,发现都近似服从幂律分布,而幂指数范围为2.3~2.5之间.休息日与工作日的幂指数并没有较大差异,这表明银行客户到访有着共同的特性,而且对于休息日、工作日并无差别,形成这种特性的机制还尚不清楚.

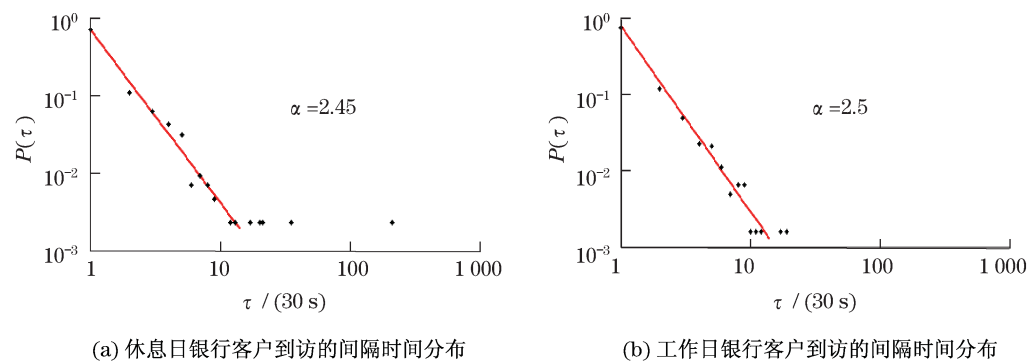


图4 某个休息日和某个工作日客户到访的间隔时间分布

Fig.4 Inter-arrival distributions of bank customers visiting on an off day and on a working day

3 细分客户到访行为的统计特征

银行个人业务和公司业务具有很大的差别,那么个人业务客户和公司业务客户前往银行办理业务的行为是否还会保持着某种特有的性质,接下来本文分析细分客户到访行为的分布情况,见图5.由于数据量的关系无法对其它业务进行合理的分析,待今后能得到更多

相关数据时再次讨论.细分客户到访的间隔时间分布是一种特定客户的分析(用户都是某银行网点的客户),每一类型的客户访问银行都受到很多因素的影响,例如工作时间、休息时间等.所以,可能和其他类型客户访问银行的情况在幂指数上存在一定的偏差,如退休职工办理个人业务时对于工作日或者休息日的选择并没有差异,而公司业务一般是在工作日前往银行办理,这种分析反映了社会某一个层面的人类活动特性.

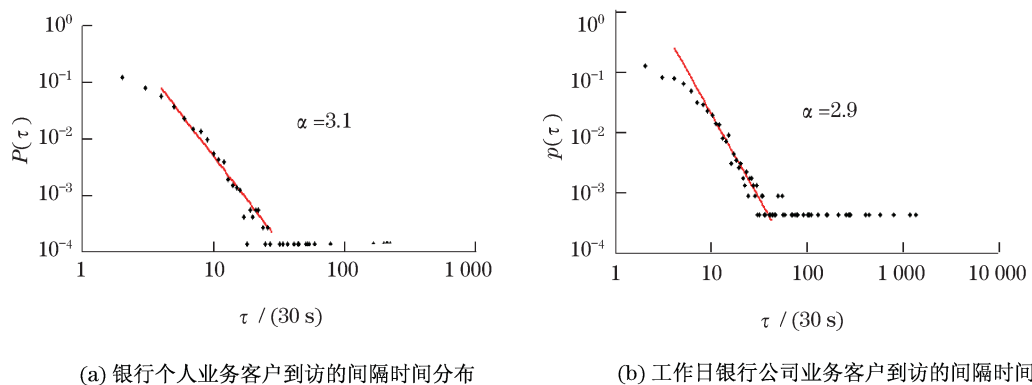


图5 19天内银行细分客户到访的间隔时间分布

Fig.5 Inter-arrival distributions of customers visiting for personal banking business and for corporate banking business in 19 days

统计结果显示,个人业务客户到访的间隔时间分布服从幂律分布,幂指数约为3.1.公司业务客户到访的间隔时间分布服从幂律分布,幂指数约为2.9.这种差异可以从个人业务和公司业务的特点上解释.个人业务在总的业务量上占了约70%,银行开门后密集访问的客户群和银行当日歇业前最后一个客户都是个人业务客户,正是个人业务客户到访时间间隔的不均匀性使间隔时间分布出现了幂律.图6给出的个人业务客户到访问隔时间图,可以发现个人业务客户的间隔时间图与所有客户的间隔时间图的形状十分相似.那些长钉一般的竖线是由于个人业务客户在周末中午到访问隔时间较大造成的,这些竖线与图2中长钉般竖线的长度及出现的位置基本一致.而公司员工一般集中在工作日前往银行办理业务,很少会在休息日处理,除非有紧急事务,因此每逢周末都会出现相比个人业务客户更长的间隔时间(见图7).正是由于个人业务客户和公司业务客户到访时间间隔的极不均匀性形成了幂律.

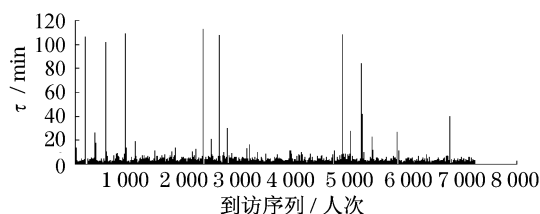


图6 19天内个人业务客户到访问隔时间

Fig.6 Inter-arrival time of customers visiting for personal banking business in 19 days

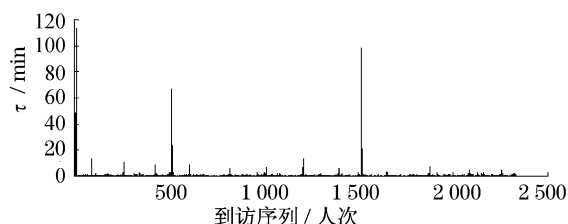


图7 19天内公司业务客户到访问隔时间

Fig.7 Inter-arrival time of customers visiting for corporate banking business in 19 days

4 小 结

本文主要讨论银行客户的到访行为模式和个人业务客户、公司业务客户到访的间隔时间分布的统计特性.由分析结果可以得出银行客户到访的间隔时间分布、单日银行客户到访问隔时间分布、个人业务客户到访问隔时间分布、公司业务客户到访问隔时间分布都是满足 $P(\tau) \propto \tau^{-\alpha}$ (α 在2到3之间)的幂律分布.因此,有理由认为银行客户到访的时间间隔分布具有幂律特性.统计结果再次说明人类行为可能在很多个层面上都偏离泊松过程.这个发现为排队理论的创新奠定了实证基础,具有一定的意义,同时也为解决银行排队拥挤问题奠定了理论基础.就排队论视角而言,银行排队问题还牵涉客户等待时间、服务时间、队长等,后续工作将对这些问题进行研究.进而,将实证某高校财务事务排队系统和研究统计的时间单位对排队指标分布形状的影响.

参考文献:

- [1] 曹杰.行为科学[M].北京:科学技术文献出版社,1987.
- [2] 李楠楠,周涛,张宁.人类动力学基本概念与实证分析[J].复杂系统与复杂性科学,2008,5(2):15-24.
- [3] Haight F A. Handbook of the Poisson distribution[M]. New York: Wiley, 1967.
- [4] Barabási A L. The origin of bursts and heavy tails in human dynamics[J]. Nature, 2005, 435(7039): 207-211.
- [5] 郭进利.博客评论的人类行为动力学实证研究和建模[J].计算机应用研究,2011,28(4):1422-1424.
- [6] 李楠楠,张宁,周涛.人类通信模式中基于时间统计的实证研究[J].复杂系统与复杂性科学,2008,5(3):43-47.
- [7] 樊超,郭进利.基于图书借阅的人类行为标度律分析[J].图书情报工作,2010,54(1):35-39.
- [8] 颜薇娜.基于蒙特卡洛模拟的商业银行排队问题研究[J].技术经济与管理,2009(1):20-22.
- [9] 杨米沙.银行排队系统数据分析及窗口设置优化研究[J].武汉理工大学学报(信息与管理工程版),2008,30(4):20-22.